Title

slogit postestimation — Postestimation tools for slogit

Description

The following postestimation commands are available after `slogit`:

Command	Description
`contrast`	contrasts and ANOVA-style joint tests of estimates
`estat ic`	Akaike's and Schwarz's Bayesian information criteria (AIC and BIC)
`estat summarize`	summary statistics for the estimation sample
`estat vce`	variance–covariance matrix of the estimators (VCE)
`estat` (svy)	postestimation statistics for survey data
`estimates`	cataloging estimation results
`lincom`	point estimates, standard errors, testing, and inference for linear combinations of coefficients
`lrtest`[1]	likelihood-ratio test
`margins`	marginal means, predictive margins, marginal effects, and average marginal effects
`marginsplot`	graph the results from margins (profile plots, interaction plots, etc.)
`nlcom`	point estimates, standard errors, testing, and inference for nonlinear combinations of coefficients
`predict`	predicted probabilities, estimated index and its approximate standard error
`predictnl`	point estimates, standard errors, testing, and inference for generalized predictions
`pwcompare`	pairwise comparisons of estimates
`suest`	seemingly unrelated estimation
`test`	Wald tests of simple and composite linear hypotheses
`testnl`	Wald tests of nonlinear hypotheses

[1] `lrtest` is not appropriate with `svy` estimation results.

Syntax for predict

predict [*type*] {*stub** | *newvar* | *newvarlist*} [*if*] [*in*] [, *statistic* outcome(*outcome*)]

predict [*type*] {*stub** | *newvarlist*} [*if*] [*in*], scores

statistic	Description
Main	
pr	probability of one or all of the dependent variable outcomes; the default
xb	index for the kth outcome
stdp	standard error of the index for the kth outcome

If you do not specify outcome(), pr (with one new variable specified), xb, and stdp assume outcome(#1).

You specify one or k new variables with pr, where k is the number of outcomes.

You specify one new variable with xb and stdp.

These statistics are available both in and out of sample; type predict ... if e(sample) ... if wanted only for the estimation sample.

Menu for predict

Statistics > Postestimation > Predictions, residuals, etc.

Options for predict

Main

pr, the default, calculates the probability of each of the categories of the dependent variable or the probability of the level specified in outcome(*outcome*). If you specify the outcome(*outcome*) option, you need to specify only one new variable; otherwise, you must specify a new variable for each category of the dependent variable.

xb calculates the index, $\theta_k - \sum_{j=1}^{d} \phi_{\mathrm{jk}} \mathbf{x}_i \boldsymbol{\beta}_j$, for outcome level $k \neq$ e(i_base) and dimension $d =$ e(k_dim). It returns a vector of zeros if $k =$ e(i_base). A synonym for xb is index. If outcome() is not specified, outcome(#1) is assumed.

stdp calculates the standard error of the index. A synonym for stdp is seindex. If outcome() is not specified, outcome(#1) is assumed.

outcome(*outcome*) specifies the outcome for which the statistic is to be calculated. equation() is a synonym for outcome(): it does not matter which you use. outcome() or equation() can be specified using

#1, #2, ..., where #1 means the first category of the dependent variable, #2 means the second category, etc.;

the values of the dependent variable; or

the value labels of the dependent variable if they exist.

scores calculates the equation-level score variables. For models with d dimensions and m levels, $d + (d + 1)(m - 1)$ new variables are created. Assume $j = 1, \ldots, d$ and $k = 1, \ldots, m$ in the following.

The first d new variables will contain $\partial \ln L / \partial(\mathbf{x}\boldsymbol{\beta}_j)$.

The next $d(m-1)$ new variables will contain $\partial \ln L/\partial \phi_{jk}$.

The last $m-1$ new variables will contain $\partial \ln L/\partial \theta_k$.

Remarks and examples

stata.com

Once you have fit a stereotype logistic model, you can obtain the predicted probabilities by using the `predict` command for both the estimation sample and other samples; see [U] **20 Estimation and postestimation commands** and [R] **predict**.

`predict` without arguments (or with the `pr` option) calculates the predicted probability of each outcome of the dependent variable. You must therefore give a new variable name for each of the outcomes. To compute the estimated probability of one outcome, you use the `outcome(`*outcome*`)` option where *outcome* is the level encoding the outcome. If the dependent variable's levels are labeled, the outcomes can also be identified by the label values (see [D] **label**).

The `xb` option in conjunction with `outcome(`*outcome*`)` specifies that the index be computed for the outcome encoded by level *outcome*. Its approximate standard error is computed if the `stdp` option is specified. Only one of the `pr`, `xb`, or `stdp` options can be specified with a call to `predict`.

▷ Example 1

In example 2 of [R] **slogit**, we fit the one-dimensional stereotype model, where the *depvar* is `insure` with levels $k=1$ for outcome *Indemnity*, $k=2$ for *Prepaid*, and $k=3$ for *Uninsure*. The base outcome for the model is *Indemnity*, so for $k \neq 1$ the vector of indices for the kth level is

$$\eta_k = \theta_k - \phi_k \left(\beta_1 \texttt{age} + \beta_2 \texttt{male} + \beta_3 \texttt{nonwhite} + \beta_4 \texttt{2.site} + \beta_5 \texttt{3.site}\right)$$

We estimate the group probabilities by calling `predict` after `slogit`.

```
. use http://www.stata-press.com/data/r13/sysdsn1
(Health insurance data)

. slogit insure age male nonwhite i.site, dim(1) base(1) nolog
  (output omitted )

. predict pIndemnity pPrepaid pUninsure, p

. list pIndemnity pPrepaid pUninsure insure in 1/10
```

	pIndem~y	pPrepaid	pUnins~e	insure
1.	.5419344	.3754875	.0825782	Indemnity
2.	.4359638	.496328	.0677081	Prepaid
3.	.5111583	.4105107	.0783309	Indemnity
4.	.3941132	.5442234	.0616633	Prepaid
5.	.4655651	.4625064	.0719285	.
6.	.4401779	.4915102	.0683118	Prepaid
7.	.4632122	.4651931	.0715948	Prepaid
8.	.3772302	.5635696	.0592002	.
9.	.4867758	.4383018	.0749225	Uninsure
10.	.5823668	.3295802	.0880531	Prepaid

Observations 5 and 8 are not used to fit the model because `insure` is missing at these points, but `predict` estimates the probabilities for these observations since none of the independent variables is missing. You can use `if e(sample)` in the call to `predict` to use only those observations that are used to fit the model.

◁

Methods and formulas

predict

Let level b be the base outcome that is used to fit the stereotype logistic regression model of dimension d. The index for observation i and level $k \neq b$ is $\eta_{ik} = \theta_k - \sum_{j=1}^{d} \phi_{jk} \mathbf{x}_i \boldsymbol{\beta}_j$. This is the log odds of outcome encoded as level k relative to that of b so that we define $\eta_{ib} \equiv 0$. The outcome probabilities for this model are defined as $\Pr(Y_i = k) = e^{\eta_{ik}} / \sum_{j=1}^{m} e^{\eta_{ij}}$. Unlike in `mlogit`, `ologit`, and `oprobit`, the index is no longer a linear function of the parameters. The standard error of index η_{ik} is thus computed using the delta method (see also [R] **predictnl**).

The equation-level score for regression coefficients is

$$\frac{\partial \ln L_{ik}}{\partial \mathbf{x}_i \boldsymbol{\beta}_j} = \left(\sum_{l=1}^{m-1} \phi_{jl} p_{il} - \phi_{jk} \right)$$

the equation-level score for the scale parameters is

$$\frac{\partial \ln L_{ik}}{\partial \phi_{jl}} = \begin{cases} \mathbf{x}_i \boldsymbol{\beta}_j (p_{ik} - 1), & \text{if } l = k \\ \mathbf{x}_i \boldsymbol{\beta}_j p_{il}, & \text{if } l \neq k \end{cases}$$

for $l = 1, \ldots, m-1$; and the equation-level score for the intercepts is

$$\frac{\partial \ln L_{ik}}{\partial \theta_l} = \begin{cases} 1 - p_{ik}, & \text{if } l = k \\ -p_{il}, & \text{if } l \neq k \end{cases}$$

Also see

[R] **slogit** — Stereotype logistic regression

[U] **20 Estimation and postestimation commands**