Title

stata.com

asclogit postestimation — Postestimation tools for asclogit

Description
Options for predict
Options for estat mfx
Methods and formulas
Syntax for predict
Syntax for estat
Remarks and examples
Also see
Menu for predict
Menu for estat
Stored results

Description

The following postestimation commands are of special interest after `asclogit`:

Commands	Description
`estat alternatives`	alternative summary statistics
`estat mfx`	marginal effects

The following standard postestimation commands are also available:

Commands	Description
`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)
`estimates`	cataloging estimation results
`hausman`	Hausman's specification test
`lincom`	point estimates, standard errors, testing, and inference for linear combinations of coefficients
`lrtest`	likelihood-ratio test
`nlcom`	point estimates, standard errors, testing, and inference for nonlinear combinations of coefficients
`predict`	predicted probabilities, estimated linear predictor and its standard error
`predictnl`	point estimates, standard errors, testing, and inference for generalized predictions
`test`	Wald tests of simple and composite linear hypotheses
`testnl`	Wald tests of nonlinear hypotheses

Special-interest postestimation commands

`estat alternatives` displays summary statistics about the alternatives in the estimation sample.

`estat mfx` computes probability marginal effects.

Syntax for predict

```
predict [type] newvar [if] [in] [, statistic options]

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

statistic	Description
Main	
pr	probability that each alternative is chosen; the default
xb	linear prediction
stdp	standard error of the linear prediction

options	Description
Main	
*k(# | observed)	condition on # alternatives per case or on observed number of alternatives
altwise	use alternativewise deletion instead of casewise deletion when computing probabilities
nooffset	ignore the offset() variable specified in asclogit

*k(# | observed) may be used only with pr.

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 computes the probability of choosing each alternative conditioned on each case choosing k() alternatives. This is the default statistic with default k(1); one alternative per case is chosen.

xb computes the linear prediction.

stdp computes the standard error of the linear prediction.

k(# | observed) conditions the probability on # alternatives per case or on the observed number of alternatives. The default is k(1). This option may be used only with the pr option.

altwise specifies that alternativewise deletion be used when marking out observations due to missing values in your variables. The default is to use casewise deletion. The xb and stdp options always use alternativewise deletion.

nooffset is relevant only if you specified offset(*varname*) for asclogit. It modifies the calculations made by predict so that they ignore the offset variable; the linear prediction is treated as $\mathbf{x}\boldsymbol{\beta}$ rather than as $\mathbf{x}\boldsymbol{\beta} + \text{offset}$.

scores calculates the scores for each coefficient in e(b). This option requires a new variable list of length equal to the number of columns in e(b). Otherwise, use the *stub** option to have predict generate enumerated variables with prefix *stub*.

Syntax for estat

Alternative summary statistics

```
estat alternatives
```

Marginal effects

```
estat mfx [if] [in] [, options]
```

options	Description
Main	
varlist(*varlist*)	display marginal effects for *varlist*
at(mean [*atlist*] | median [*atlist*])	calculate marginal effects at these values
k(#)	condition on the number of alternatives chosen to be #
Options	
level(#)	set confidence interval level; default is level(95)
nodiscrete	treat indicator variables as continuous
noesample	do not restrict calculation of means and medians to the estimation sample
nowght	ignore weights when calculating means and medians

Menu for estat

Statistics > Postestimation > Reports and statistics

Options for estat mfx

Main

varlist(*varlist*) specifies the variables for which to display marginal effects. The default is all variables.

at(mean [*atlist*] | median [*atlist*]) specifies the values at which the marginal effects are to be calculated. *atlist* is

[[*alternative*:*variable* = #] [*variable* = #] [*alternative*:*offset* = #] [...]]

The default is to calculate the marginal effects at the means of the independent variables by using the estimation sample, at(mean). If offset() is used during estimation, the means of the offsets (by alternative) are computed by default.

After specifying the summary statistic, you can specify a series of specific values for variables. You can specify values for alternative-specific variables by alternative, or you can specify one value for all alternatives. You can specify only one value for case-specific variables. You specify values for the offset() variable (if present) the same way as for alternative-specific variables. For example, in the choice dataset (car choice), income is a case-specific variable, whereas dealer is an alternative-specific variable. The following would be a legal syntax for estat mfx:

```
. estat mfx, at(mean American:dealer=18 income=40)
```

When nodiscrete is not specified, at(mean [atlist]) or at(median [atlist]) has no effect on computing marginal effects for indicator variables, which are calculated as the discrete change in the simulated probability as the indicator variable changes from 0 to 1.

The mean and median computations respect any if or in qualifiers, so you can restrict the data over which the statistic is computed. You can even restrict the values to a specific case, for example,

```
. estat mfx if case==21
```

k(#) computes the probabilities conditioned on # alternatives chosen. The default is one alternative chosen.

Options

level(#) sets the confidence level; default is level(95).

nodiscrete specifies that indicator variables be treated as continuous variables. An indicator variable is one that takes on the value 0 or 1 in the estimation sample. By default, the discrete change in the simulated probability is computed as the indicator variable changes from 0 to 1.

noesample specifies that the whole dataset be considered instead of only those marked in the e(sample) defined by the asclogit command.

nowght specifies that weights be ignored when calculating the medians.

Remarks and examples

stata.com

Remarks are presented under the following headings:

Predicted probabilities
Obtaining estimation statistics

Predicted probabilities

After fitting a McFadden's choice model with alternative-specific conditional logistic regression, you can use predict to obtain the estimated probability of alternative choices given case profiles.

▷ Example 1

In example 1 of [R] **asclogit**, we fit a model of consumer choice of automobile. The alternatives are nationality of the automobile manufacturer: American, Japanese, or European. There is one alternative-specific variable in the model, dealer, which contains the number of dealerships of each nationality in the consumer's city. The case-specific variables are sex, the consumer's sex, and income, the consumer's income in thousands of dollars.

```
. use http://www.stata-press.com/data/r13/choice
. asclogit choice dealer, case(id) alternatives(car) casevars(sex income)
  (output omitted)
. predict p
(option pr assumed; Pr(car))
. predict p2, k(2)
(option pr assumed; Pr(car))
. format p p2 %6.4f
```

```
. list car choice dealer sex income p p2 in 1/9, sepby(id)
```

	car	choice	dealer	sex	income	p	p2
1.	American	0	18	male	46.7	0.6025	0.8589
2.	Japan	0	8	male	46.7	0.2112	0.5974
3.	Europe	1	5	male	46.7	0.1863	0.5437
4.	American	1	17	male	26.1	0.7651	0.9293
5.	Japan	0	6	male	26.1	0.1282	0.5778
6.	Europe	0	2	male	26.1	0.1067	0.4929
7.	American	1	12	male	32.7	0.6519	0.8831
8.	Japan	0	6	male	32.7	0.1902	0.5995
9.	Europe	0	2	male	32.7	0.1579	0.5174

◁

Obtaining estimation statistics

Here we will demonstrate the specialized `estat` subcommands after `asclogit`. Use `estat alternatives` to obtain a table of alternative statistics. The table will contain the alternative values, labels (if any), the number of cases in which each alternative is present, the frequency that the alternative is selected, and the percent selected.

Use `estat mfx` to obtain marginal effects after `asclogit`.

▷ Example 2

We will continue with the automobile choice example, where we first list the alternative statistics and then compute the marginal effects at the mean income in our sample, assuming that there are five automobile dealers for each nationality. We will evaluate the probabilities for females because `sex` is coded 0 for females, and we will be obtaining the discrete change from 0 to 1.

```
. estat alternatives
   Alternatives summary for car
```

index	Alternative value	label	Cases present	Frequency selected	Percent selected
1	1	American	295	192	65.08
2	2	Japan	295	64	21.69
3	3	Europe	295	39	13.22

```
. estat mfx, at(dealer=0 sex=0) varlist(sex income)
Pr(choice = American|1 selected) = .41964329
```

variable	dp/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
casevars							
sex*	.026238	.068311	0.38	0.701	-.107649	.160124	0
income	-.007891	.002674	-2.95	0.003	-.013132	-.00265	42.097

(*) dp/dx is for discrete change of indicator variable from 0 to 1

```
Pr(choice = Japan|1 selected) = .42696187
```

variable	dp/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
casevars							
sex*	-.161164	.079238	-2.03	0.042	-.316468	-.005859	0
income	.005861	.002997	1.96	0.051	-.000014	.011735	42.097

```
(*) dp/dx is for discrete change of indicator variable from 0 to 1
Pr(choice = Europe|1 selected) = .15339484
```

variable	dp/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
casevars							
sex*	.134926	.076556	1.76	0.078	-.015122	.284973	0
income	.00203	.001785	1.14	0.255	-.001469	.00553	42.097

```
(*) dp/dx is for discrete change of indicator variable from 0 to 1
```

The marginal effect of `income` indicates that there is a lower chance for a consumer to buy American automobiles with an increase in income. There is an indication that men have a higher preference for European automobiles than women but a lower preference for Japanese automobiles. We did not include the marginal effects for `dealer` because we view these as nuisance parameters, so we adjusted the probabilities by fixing dealer to a constant, 0.

◁

Stored results

`estat mfx` stores the following in `r()`:

Scalars

`r(pr_`*alt*`)` scalars containing the computed probability of each alternative evaluated at the value that is labeled X in the table output. Here *alt* are the labels in the macro `e(alteqs)`.

Matrices

`r(`*alt*`)` matrices containing the computed marginal effects and associated statistics. There is one matrix for each alternative, where *alt* are the labels in the macro `e(alteqs)`. Column 1 of each matrix contains the marginal effects; column 2, their standard errors; column 3, their z statistics; and columns 4 and 5, the confidence intervals. Column 6 contains the values of the independent variables used to compute the probabilities `r(pr_`*alt*`)`.

Methods and formulas

The deterministic component of the random-utility model can be expressed as

$$\begin{aligned}\boldsymbol{\eta} &= \mathbf{X}\boldsymbol{\beta} + (\mathbf{z}\mathbf{A})' \\ &= \mathbf{X}\boldsymbol{\beta} + (\mathbf{z} \otimes \mathbf{I}_J)\,\mathrm{vec}(\mathbf{A}') \\ &= (\mathbf{X},\ \mathbf{z} \otimes \mathbf{I}_J) \begin{pmatrix} \boldsymbol{\beta} \\ \mathrm{vec}(\mathbf{A}') \end{pmatrix} \\ &= \mathbf{X}^*\boldsymbol{\beta}^*\end{aligned}$$

where $\mathbf{X}$ is the $J \times p$ matrix containing the alternative-specific covariates, $\mathbf{z}$ is a $1 \times q$ vector of case-specific variables, $\boldsymbol{\beta}$ is a $p \times 1$ vector of alternative-specific regression coefficients, and $\mathbf{A} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_J)$ is a $q \times J$ matrix of case-specific regression coefficients (with one of the $\boldsymbol{\alpha}_j$ fixed to the constant). Here $\mathbf{I}_J$ is the $J \times J$ identity matrix, vec() is the vector function that creates a vector from a matrix by placing each column of the matrix on top of the other (see [M-5] **vec( )**), and $\otimes$ is the Kronecker product (see [M-2] **op_kronecker**).

We have rewritten the linear equation so that it is a form that we all recognize, namely, $\boldsymbol{\eta} = \mathbf{X}^*\boldsymbol{\beta}^*$, where

$$\mathbf{X}^* = (\mathbf{X},\ \mathbf{z} \otimes \mathbf{I}_J)$$

$$\boldsymbol{\beta}^* = \begin{pmatrix} \boldsymbol{\beta} \\ \mathrm{vec}(\mathbf{A}') \end{pmatrix}$$

To compute the marginal effects, we use the derivative of the log likelihood $\partial\ell(\mathbf{y}|\boldsymbol{\eta})/\partial\boldsymbol{\eta}$, where $\ell(\mathbf{y}|\boldsymbol{\eta}) = \log \Pr(\mathbf{y}|\boldsymbol{\eta})$ is the log of the probability of the choice indicator vector $\mathbf{y}$ given the linear predictor vector $\boldsymbol{\eta}$. Namely,

$$\begin{aligned} \frac{\partial \Pr(\mathbf{y}|\boldsymbol{\eta})}{\partial \mathrm{vec}(\mathbf{X}^*)'} &= \Pr(\mathbf{y}|\boldsymbol{\eta}) \frac{\partial \ell(\mathbf{y}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} \frac{\partial \boldsymbol{\eta}}{\partial \mathrm{vec}(\mathbf{X}^*)'} \\ &= \Pr(\mathbf{y}|\boldsymbol{\eta}) \frac{\partial \ell(\mathbf{y}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}'} \left(\boldsymbol{\beta}^{*\prime} \otimes \mathbf{I}_J\right) \end{aligned}$$

The standard errors of the marginal effects are computed using the delta method.

Also see

[R] **asclogit** — Alternative-specific conditional logit (McFadden's choice) model

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