ziologit postestimation — Postestimation tools for ziologit

Postestimation commands

The following postestimation commands are available after `ziologit`:

Command	Description
`contrast`	contrasts and ANOVA-style joint tests of estimates
`estat ic`	Akaike's, consistent Akaike's, corrected Akaike's, and Schwarz's Bayesian information criteria (AIC, CAIC, AICc, 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
`etable`	table of estimation results
* `forecast`	dynamic forecasts and simulations
* `hausman`	Hausman's specification test
`lincom`	point estimates, standard errors, testing, and inference for linear combinations of coefficients
* `lrtest`	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`	probabilities, linear predictions and their SEs, etc.
`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

* `forecast`, `hausman`, and `lrtest` are not appropriate with `svy` estimation results.

predict

Description for predict

`predict` creates a new variable containing predictions such as probabilities, linear predictions, and standard errors.

Menu for predict

Statistics > Postestimation

Syntax for predict

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

predict [*type*] *stub** [*if*] [*in*], scores

statistic	Description
Main	
`pmargin`	marginal probabilities of levels, $\Pr(y_j = h)$; the default
`pjoint1`	joint probabilities of levels and susceptibility, $\Pr(y_j = h, s_j = 1)$
`pcond1`	probabilities of levels conditional on susceptibility, $\Pr(y_j = h \vert s_j = 1)$
`ps`	probability of susceptibility, $\Pr(s_j = 1)$
`pns`	probability of nonsusceptibility, $\Pr(s_j = 0)$
`xb`	linear prediction
`xbinfl`	linear prediction for inflation equation
`stdp`	standard error of the linear prediction
`stdpinfl`	standard error of the linear prediction for inflation equation

If you do not specify `outcome()`, `pmargin`, `pjoint1`, and `pcond1` (with one new variable specified) assume `outcome(#1)`.

You specify one or k new variables with `pmargin`, `pjoint1`, and `pcond1`, where k is the number of outcomes.

You specify one new variable with `ps`, `pns`, `xb`, `xbinfl`, `stdp`, and `stdpinfl`.

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

Options for predict

Main

`pmargin`, the default, calculates the predicted marginal probabilities of outcome levels, $\Pr(y_j = h)$.

`pjoint1` calculates the predicted joint probabilities of outcome levels and susceptibility, $\Pr(y_j = h, s_j = 1)$.

`pcond1` calculates the predicted probabilities of outcome levels conditional on susceptibility, $\Pr(y_j = h \vert s_j = 1)$.

With `pmargin`, `pjoint1`, and `pcond1`, you can compute predicted probabilities for one or for all outcome levels. When you specify one new variable, `predict` computes probabilities for the first outcome level. You can specify the `outcome(#`*i*`)` option to obtain probabilities for the ith level. When you specify multiple new variables or a stub, `predict` computes probabilities for all outcome levels. The behavior of `predict` with one new variable is equivalent to specifying `outcome(#1)`.

`ps` and `pns` calculate the predicted marginal probability of susceptibility $[\Pr(s_j = 1)]$ and of nonsusceptibility $[\Pr(s_j = 0)]$, respectively.

In econometrics literature, probabilities of susceptibility and nonsusceptibility are known as probabilities of participation and nonparticipation. Similarly to `predict` after `zioprobit`, you can use options `ppar` and `pnpar` to compute these probabilities. Options `ppar` and `pnpar` produce identical results to the respective options `ps` and `pns` but label new variables as `Pr(participation)` and `Pr(nonparticipation)` instead of `Pr(susceptible)` and `Pr(nonsusceptible)`.

`xb` calculates the linear prediction for the ordered logit equation, which is $\mathbf{x}_j\boldsymbol{\beta}$ if `offset()` was not specified with `ziologit` and is $\mathbf{x}_j\boldsymbol{\beta} + \text{offset}_j^{\beta}$ if `offset()` was specified.

`xbinfl` calculates the linear prediction for the inflation equation, which is $\mathbf{z}_j\boldsymbol{\gamma}$ if `offset()` was not specified in `inflate()` and is $\mathbf{z}_j\boldsymbol{\gamma} + \text{offset}_j^{\gamma}$ if `offset()` was specified in `inflate()`.

`stdp` calculates the standard error of the linear prediction for the ordered logit equation.

`stdpinfl` calculates the standard error of the linear prediction for the inflation equation.

`outcome(`*outcome*`)` specifies the outcome for which predicted probabilities are to be calculated. `outcome()` should contain either one value of the dependent variable or one of `#1`, `#2`, ..., with `#1` meaning the first category of the dependent variable, `#2` meaning the second category, etc. `outcome()` is allowed only with `pmargin`, `pjoint1`, and `pcond1`.

`nooffset` is relevant only if you specified `offset(`*varname*`)` with `ziologit` or within the `inflate()` option. It modifies the calculations made by `predict` so that they ignore the offset variable; that is, the linear prediction for the main regression equation is treated as $\mathbf{x}_j\boldsymbol{\beta}$ rather than as $\mathbf{x}_j\boldsymbol{\beta} + \text{offset}_j^{\beta}$ and the linear prediction for the inflation equation is treated as $\mathbf{z}_j\boldsymbol{\gamma}$ rather than as $\mathbf{z}_j\boldsymbol{\gamma} + \text{offset}_j^{\gamma}$.

`scores` calculates equation-level score variables.

The first new variable will contain $\partial \ln L/\partial(\mathbf{x}_j\boldsymbol{\beta})$. In the absence of independent variables in the main equation, this variable is not stored.

The second new variable will contain $\partial \ln L/\partial(\mathbf{z}_j\boldsymbol{\gamma})$.

When the dependent variable takes k different values, the third new variable through new variable $k+1$ will contain $\partial \ln L/\partial(\kappa_h)$ for $h = 0, 1, \ldots, k-2$.

margins

Description for margins

margins estimates margins of response for probabilities and linear predictions.

Menu for margins

Statistics > Postestimation

Syntax for margins

margins [*marginlist*] [, *options*]

margins [*marginlist*] , predict(*statistic* ...) [predict(*statistic* ...) ...] [*options*]

statistic	Description
default	marginal probabilities for each outcome
pmargin	marginal probabilities of levels, $\Pr(y_j = h)$; the default
pjoint1	joint probabilities of levels and susceptibility, $\Pr(y_j = h, s_j = 1)$
pcond1	probabilities of levels conditional on susceptibility, $\Pr(y_j = h \mid s_j = 1)$
ps	probability of susceptibility, $\Pr(s_j = 1)$
pns	probability of nonsusceptibility, $\Pr(s_j = 0)$
xb	linear prediction
xbinfl	linear prediction for inflation equation
stdp	not allowed with margins
stdpinfl	not allowed with margins

pmargin, pjoint1, and pcond1 default to the first outcome.

Statistics not allowed with margins are functions of stochastic quantities other than e(b).

For the full syntax, see [R] **margins**.

Remarks and examples

stata.com

The ZIOL model allows all the predictions and marginal effects available with the standard `ologit` model (see [R] **ologit postestimation**), along with additional predictions and marginal effects related to the inflation equation for susceptibility. The probabilities of susceptibility and nonsusceptibility can be calculated using options `ps` and `pns`, respectively. If you prefer an alternative terminology of probabilities of participation and nonparticipation, you can instead use options `ppar` and `pnpar`, which will produce identical numerical results but label variables as `Pr(participation)` and `Pr(nonparticipation)` instead of `Pr(susceptible)` and `Pr(nonsusceptible)`.

▷ Example 1: Average marginal effect of gender on probability of nonsusceptibility

In example 1 of [R] **ziologit**, we fit a model for levels of cigarette consumption.

```
. use https://www.stata-press.com/data/r18/tobacco
(Fictional tobacco consumption data)
. ziologit tobacco education income age i.female,
> inflate(education income age i.female i.parent i.religion)
  (output omitted )
```

This model parallels the zero-inflated ordered probit (ZIOP) model that was fit in example 1 of [R] **zioprobit**.

To continue the comparison between the ZIOL and ZIOP models, we re-create example 1 from [R] **zioprobit postestimation** by using `margins` to estimate the average marginal effect of gender on the probability of nonsusceptibility (being an excess zero) for individuals with a college degree (17 years of education) and a smoking parent.

```
. margins, predict(pns) dydx(female) at(education = 17 parent = 1)
Average marginal effects                              Number of obs = 15,000
Model VCE: OIM
Expression: Pr(nonsusceptible), predict(pns)
dy/dx wrt:  1.female
At: education = 17
    parent    =  1
```

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf. interval]	
female						
Female	.085421	.010096	8.46	0.000	.0656333	.1052087

```
Note: dy/dx for factor levels is the discrete change from the base level.
```

Despite the differences between the ZIOL and ZIOP models, the conclusion is the same: women with a college degree and a smoking parent are expected to have an approximately 8.5% higher chance of being genuine nonsmokers (excess zeros) than comparable men.

◁

▷ Example 2: Predicted probabilities of conditional zeros

Next, we consider the effect of income on the probability of zero tobacco consumption, conditional on susceptibility. These would-be smokers are known as conditional zeros. In example 1 of [R] **ziologit**, we saw that increasing income raises a smoker's odds of increased tobacco consumption dramatically, so we expect to see a larger fraction of conditional zeros at the lower end of the income scale.

We examine conditional probabilities of zero consumption for incomes ranging from $10,000 to $60,000, and we use the `noatlegend` option to suppress the default legend because we know the values 1 to 6 correspond to income in tens of thousands of dollars.

```
. margins, predict(pcond1 outcome(0)) at(income = (1/6)) noatlegend

Predictive margins                                     Number of obs = 15,000
Model VCE: OIM

Expression: Pr(tobacco=0|susceptible=1), predict(pcond1 outcome(0))

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .5923634   .0027586   214.73   0.000     .5869566    .5977702
          2  |   .5393818   .0025948   207.87   0.000      .534296    .5444676
          3  |   .4854668   .0024651   196.94   0.000     .4806354    .4902982
          4  |   .4306299   .0023953   179.78   0.000     .4259352    .4353245
          5  |   .3741538   .0024547   152.42   0.000     .3693427    .3789649
          6  |   .3152985   .0026294   119.91   0.000     .3101449     .320452
------------------------------------------------------------------------------
```

The influence of income is dramatic: susceptible individuals (potential smokers) who earn $10,000 a year are almost twice as likely to refrain from smoking as potential smokers who earn $60,000 per year (59% versus 32%).

◁

Methods and formulas

See *Methods and formulas* in [R] **ziologit** for the model definition and notation. Specifically, see (1) for the formula for the probability of susceptibility, $\Pr(s_j = 1|\mathbf{z}_j)$; see (2) for the formula for the probabilities of outcome levels conditional on susceptibility, $\Pr(y_j = h|s_j = 1, \mathbf{x}_j)$; and see (4) for the formula for the marginal probabilities of outcome levels, $\Pr(y_j = h|\mathbf{z}_j, \mathbf{x}_j)$.

The joint probability of susceptibility and outcome $y_j = h$ can be expressed as

$$\Pr(y_j = h,\, s_j = 1|\mathbf{z}_j, \mathbf{x}_j) = \Pr\left(s_j = 1|\mathbf{z}_j\right) \Pr(y_j = h|s_j = 1, \mathbf{x}_j)$$

for $h = 0, 1, \ldots, H$.

Reference

Kelley, M. E., and S. J. Anderson. 2008. Zero inflation in ordinal data: Incorporating susceptibility to response through the use of a mixture model. *Statistics in Medicine* 27: 3674–3688. https://doi.org/10.1002/sim.3267.

Also see

[R] **ziologit** — Zero-inflated ordered logit regression

[U] 20 Estimation and postestimation commands