Title

stata.com

> **intreg postestimation —** Postestimation tools for intreg

Description

The following postestimation commands are available after `intreg`:

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	predictions, residuals, influence statistics, and other diagnostic measures
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*] *newvar* [*if*] [*in*] [, *statistic* <u>nooff</u>set]

> predict [*type*] { *stub** | *newvar*$_{\text{reg}}$ *newvar*$_{\text{lnsigma}}$ } [*if*] [*in*] , <u>sc</u>ores

statistic	Description
Main	
xb	linear prediction; the default
stdp	standard error of the prediction
stdf	standard error of the forecast
<u>pr</u>(*a*,*b*)	$\Pr(a < y_j < b)$
e(*a*,*b*)	$E(y_j \mid a < y_j < b)$
<u>ystar</u>(*a*,*b*)	$E(y_j^*)$, $y_j^* = \max\{a, \min(y_j, b)\}$

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

stdf is not allowed with svy postestimation results.

where *a* and *b* may be numbers or variables; *a* missing ($a \geq .$) means $-\infty$, and *b* missing ($b \geq .$) means $+\infty$; see [U] **12.2.1 Missing values**.

Menu for predict

Statistics > Postestimation > Predictions, residuals, etc.

Options for predict

> Main

xb, the default, calculates the linear prediction.

stdp calculates the standard error of the prediction, which can be thought of as the standard error of the predicted expected value or mean for the observation's covariate pattern. The standard error of the prediction is also referred to as the standard error of the fitted value.

stdf calculates the standard error of the forecast, which is the standard error of the point prediction for 1 observation. It is commonly referred to as the standard error of the future or forecast value. By construction, the standard errors produced by stdf are always larger than those produced by stdp; see *Methods and formulas* in [R] **regress postestimation**.

pr(*a*,*b*) calculates $\Pr(a < \mathbf{x}_j \mathbf{b} + u_j < b)$, the probability that $y_j \mid \mathbf{x}_j$ would be observed in the interval (a, b).

a and *b* may be specified as numbers or variable names; *lb* and *ub* are variable names;
pr(20,30) calculates $\Pr(20 < \mathbf{x}_j \mathbf{b} + u_j < 30)$;
pr(*lb*,*ub*) calculates $\Pr(lb < \mathbf{x}_j \mathbf{b} + u_j < ub)$; and
pr(20,*ub*) calculates $\Pr(20 < \mathbf{x}_j \mathbf{b} + u_j < ub)$.

a missing ($a \geq .$) means $-\infty$; pr(.,30) calculates $\Pr(-\infty < \mathbf{x}_j \mathbf{b} + u_j < 30)$;
pr(*lb*,30) calculates $\Pr(-\infty < \mathbf{x}_j \mathbf{b} + u_j < 30)$ in observations for which $lb \geq .$
and calculates $\Pr(lb < \mathbf{x}_j \mathbf{b} + u_j < 30)$ elsewhere.

b missing ($b \geq .$) means $+\infty$; pr(20,.) calculates $\Pr(+\infty > \mathbf{x}_j\mathbf{b} + u_j > 20)$;
pr(20,ub) calculates $\Pr(+\infty > \mathbf{x}_j\mathbf{b} + u_j > 20)$ in observations for which $ub \geq .$
and calculates $\Pr(20 < \mathbf{x}_j\mathbf{b} + u_j < ub)$ elsewhere.

e(a,b) calculates $E(\mathbf{x}_j\mathbf{b} + u_j \mid a < \mathbf{x}_j\mathbf{b} + u_j < b)$, the expected value of $y_j|\mathbf{x}_j$ conditional on
$y_j|\mathbf{x}_j$ being in the interval (a, b), meaning that $y_j|\mathbf{x}_j$ is truncated.
a and b are specified as they are for pr().

ystar(a,b) calculates $E(y_j^*)$, where $y_j^* = a$ if $\mathbf{x}_j\mathbf{b} + u_j \leq a$, $y_j^* = b$ if $\mathbf{x}_j\mathbf{b} + u_j \geq b$, and
$y_j^* = \mathbf{x}_j\mathbf{b} + u_j$ otherwise, meaning that y_j^* is censored. a and b are specified as they are for pr().

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

scores calculates equation-level score variables.

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

The second new variable will contain $\partial \ln L / \partial \ln\sigma$.

Remarks and examples

▷ Example 1

We continue with example 1 of [R] **intreg**.

```
. use http://www.stata-press.com/data/r13/intregxmpl
. intreg wage1 wage2 age c.age#c.age nev_mar rural school tenure
(output omitted )
```

By default, the predict command produces the linear prediction, which in this case is the expected
wage for each individual.

```
. predict w1
(option xb assumed; fitted values)
```

We can use the e(a,b) option to compute the expected wage, conditional on it being larger than
$5,000:

```
. predict w2, e(5,.)
```

The probability of earning more than $5,000 might vary with age. We can use margins to compute
the marginal means for those probabilities for different ages.

```
. margins, predict(pr(5,.)) at(age=(20(5)50))
```

Predictive margins   Number of obs   =   488
Model VCE   : OIM

Expression   : Pr(y>5), predict(pr(5,.))
1._at   : age   =   20
2._at   : age   =   25
3._at   : age   =   30
4._at   : age   =   35
5._at   : age   =   40
6._at   : age   =   45
7._at   : age   =   50

	Margin	Delta-method Std. Err.	z	P>|z|	[95% Conf. Interval]
_at						
1	.8912598	.0151773	58.72	0.000	.8615127	.9210068
2	.9104568	.0103467	87.99	0.000	.8901775	.930736
3	.9160005	.0120025	76.32	0.000	.892476	.9395251
4	.9096667	.0136693	66.55	0.000	.8828753	.9364581
5	.8894289	.0206992	42.97	0.000	.8488593	.9299985
6	.8491103	.0447429	18.98	0.000	.7614159	.9368048
7	.7781644	.0970557	8.02	0.000	.5879387	.9683902

We can visualize these results by using `marginsplot`:

```
. qui margins, predict(pr(5,.)) at(age=(20(5)50))
. marginsplot
  Variables that uniquely identify margins: age
```
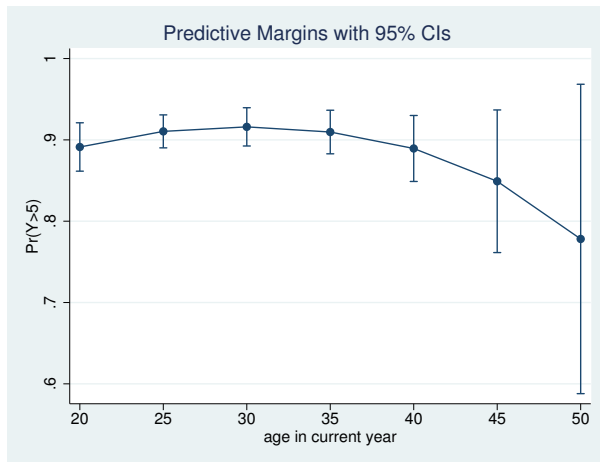

The probability increases until age 30, and it decreases after that age.

◁

Also see

[R] **intreg** — Interval regression

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