Title

stata.com

xtintreg postestimation — Postestimation tools for xtintreg

Description Syntax for predict Menu for predict Options for predict
Remarks and examples Also see

Description

The following postestimation commands are available after `xtintreg`:

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)
`estimates`	cataloging estimation results
`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`	predictions, residuals, influence statistics, and other diagnostic measures
`predictnl`	point estimates, standard errors, testing, and inference for generalized predictions
`pwcompare`	pairwise comparisons of estimates
`test`	Wald tests of simple and composite linear hypotheses
`testnl`	Wald tests of nonlinear hypotheses

Syntax for predict

`predict` [*type*] *newvar* [*if*] [*in*] [, *statistic* `nooffset`]

statistic	Description
Main	
`xb`	linear prediction assuming a zero random effect, the default
`stdp`	standard error of the linear prediction
`stdf`	standard error of the linear forecast
`pr0(`*a*`,`*b*`)`	$\Pr(a < y < b)$ assuming a zero random effect
`e0(`*a*`,`*b*`)`	$E(y \mid a < y < b)$ assuming a zero random effect
`ystar0(`*a*`,`*b*`)`	$E(y^*)$, $y^* = \max\{a, \min(y_j, b)\}$ assuming a zero random effect

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

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 linear prediction. It 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 linear forecast. This 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**.

`pr0(`*a*`,`*b*`)` calculates estimates of $\Pr(a < y < b|\mathbf{x} = \mathbf{x}_{it}, \nu_i = 0)$, which is the probability that y would be observed in the interval (a, b), given the current values of the predictors, $\mathbf{x}_{it}$, and given a zero random effect. In the discussion that follows, these two conditions are implied.

a and *b* may be specified as numbers or variable names; *lb* and *ub* are variable names;
`pr0(20,30)` calculates $\Pr(20 < y < 30)$;
`pr0(`*lb*`,`*ub*`)` calculates $\Pr(lb < y < ub)$; and
`pr0(20,`*ub*`)` calculates $\Pr(20 < y < ub)$.

a missing ($a \geq .$) means $-\infty$; `pr0(.,30)` calculates $\Pr(-\infty < y < 30)$;
`pr0(`*lb*`,30)` calculates $\Pr(-\infty < y < 30)$ in observations for which $lb \geq .$
(and calculates $\Pr(lb < y < 30)$ elsewhere).

b missing ($b \geq .$) means $+\infty$; `pr0(20,.)` calculates $\Pr(+\infty > y > 20)$;
`pr0(20,`*ub*`)` calculates $\Pr(+\infty > y > 20)$ in observations for which $ub \geq .$
(and calculates $\Pr(20 < y < ub)$ elsewhere).

`e0(`*a*`,`*b*`)` calculates estimates of $E(y \mid a < y < b, \mathbf{x} = \mathbf{x}_{it}, \nu_i = 0)$, which is the expected value of y conditional on y being in the interval (a, b), meaning that y is truncated. *a* and *b* are specified as they are for `pr0()`.

`ystar0(`*a*`,`*b*`)` calculates estimates of $E(y^*|\mathbf{x} = \mathbf{x}_{it}, \nu_i = 0)$, where $y^* = a$ if $y \leq a$, $y^* = b$ if $y \geq b$, and $y^* = y$ otherwise, meaning that y^* is the censored version of y. *a* and *b* are specified as they are for `pr0()`.

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

Remarks and examples

stata.com

Example 1

In example 1 of [XT] **xtintreg**, we fit a random-effects model of wages. Say that we want to know how union membership status affects the probability that a worker's wage will be "low", where low means a log wage that is less than the 20th percentile of all observations in our dataset. First, we use `centile` to find the 20th percentile of `ln_wage`:

```
. use http://www.stata-press.com/data/r13/nlswork5
(National Longitudinal Survey.  Young Women 14-26 years of age in 1968)
. xtintreg ln_wage1 ln_wage2 i.union age grade south##c.year, intreg
  (output omitted)
. centile ln_wage, centile(20)
```

Variable	Obs	Percentile	Centile	— Binom. Interp. — [95% Conf.	Interval]
ln_wage	28534	20	1.301507	1.297063	1.308635

Now we use `margins` to obtain the effect of union status on the probability that the log of wages is in the bottom 20% of women. Given the results from `centile` that corresponds to the log of wages being below 1.30. We evaluate the effect for two groups: 1) women age 30 living in the south in 1988 who graduated high school, but had no more schooling, and 2) the same group of women, with the exception that they are college graduates (`grade=16`).

```
. margins, dydx(union) predict(pr0(.,1.30))
> at(age=30 south=1 year=88 grade=12 union=0)
> at(age=30 south=1 year=88 grade=16 union=0)
Conditional marginal effects                    Number of obs   =      19224
Model VCE    : OIM
Expression   : Pr(ln_wage1<1.30), predict(pr0(.,1.30))
dy/dx w.r.t. : 1.union
1._at        : union           =           0
               age             =          30
               grade           =          12
               south           =           1
               year            =          88
2._at        : union           =           0
               age             =          30
               grade           =          16
               south           =           1
               year            =          88
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.union						
_at						
1	-.0787117	.0060655	-12.98	0.000	-.0905999	-.0668235
2	-.0378758	.0035595	-10.64	0.000	-.0448523	-.0308993

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

For the first group of women, according to our fitted model, being in a union lowers the probability of being classified as a low-wage worker by almost 7.9 percentage points. Being a college graduate attenuates this effect to just under 3.8 percentage points.

◁

Also see

[XT] **xtintreg** — Random-effects interval-data regression models

[U] 20 Estimation and postestimation commands