_title — Obtain predictions, residuals, etc., after estimation programming command

## Syntax

After `regress`

```
._predict [type] newvar [if] [in] [, xb stdp stdf stdr hat cooksd residuals rstandard rstudent nolab]
```

After single-equation (SE) estimators

```
._predict [type] newvar [if] [in] [, xb stdp nooffset nolabel]
```

After multiple-equation (ME) estimators

```
._predict [type] newvar [if] [in] [, xb stdp stddp nooffset nolabel

        equation(eqno[, eqno])]
```

## Description

`_predict` is for use by programmers as a subroutine for implementing the `predict` command for use after estimation; see [R] `predict`.

## Options

`xb` calculates the linear prediction from the fitted model. That is, all models can be thought of as estimating a set of parameters $b_1, b_2, \ldots, b_k$, and the linear prediction is $\hat{y}_j = b_1 x_{1j} + b_2 x_{2j} + \cdots + b_k x_{kj}$, often written in matrix notation as $\hat{y}_j = x_j b$. For linear regression, the values $\hat{y}_j$ are called the predicted values, or for out-of-sample predictions, the forecast. For logit and probit, for example, $\hat{y}_j$ is called the logit or probit index.

It is important to understand that the $x_{1j}, x_{2j}, \ldots, x_{kj}$ used in the calculation are obtained from the data currently in memory and do not have to correspond to the data on the independent variables used in fitting the model (obtaining the $b_1, b_2, \ldots, b_k$).

`stdp` calculates the standard error of the prediction after any estimation command. Here the prediction is understood to mean the same thing as the “index”, namely, $x_j b$. The statistic produced by `stdp` can be thought of as the standard error of the predicted expected value, or mean index, for the observation’s covariate pattern. This is also commonly 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] `predict`.

1
predict — Obtain predictions, residuals, etc., after estimation programming command

stdr calculates the standard error of the residuals.
hat (or leverage) calculates the diagonal elements of the projection hat matrix.
cooksd calculates the Cook’s $D$ influence statistic (Cook 1977).
residuals calculates the residuals.
rstudent calculates the standardized residuals.

nooffset may be combined with most statistics and specifies that the calculation be made, ignoring any offset or exposure variable specified when the model was fit. This option is available, even if not documented, for predict after a specific command. If neither the offset(varname) option nor the exposure(varname) option was specified when the model was fit, specifying nooffset does nothing.
nolabel prevents _predict from labeling the newly created variable.

stddp is allowed only after you have previously fit a multiple-equation model. The standard error of the difference in linear predictions ($x_1^j b - x_2^j b$) between equations 1 and 2 is calculated. Use the equation() option to get the standard error of the difference between other equations.

equation(eqno[ , eqno]) is relevant only when you have previously fit a multiple-equation model. It specifies the equation to which you are referring.

If you do not specify equation(), the results are the same as if you specified equation(#1).

Other statistics refer to between-equation concepts; stddp is an example. You might then specify equation(#1,#2) or equation(income,hours). When two equations must be specified, equation() is required.

Methods and formulas

See Methods and formulas in [R] predict and [R] regress.

Reference


Also see

[R] predict — Obtain predictions, residuals, etc., after estimation
[U] 20 Estimation and postestimation commands