

**glm postestimation** — Postestimation tools for glm

[Postestimation commands](#)
[predict](#)
[margins](#)
[Remarks and examples](#)
[Methods and formulas](#)
[References](#)
[Also see](#)

## Postestimation commands

The following postestimation commands are available after `glm`:

Command	Description
<code>contrast</code>	contrasts and ANOVA-style joint tests of estimates
* <code>estat ic</code>	Akaike's and Schwarz's Bayesian information criteria (AIC and BIC)
<code>estat summarize</code>	summary statistics for the estimation sample
<code>estat vce</code>	variance–covariance matrix of the estimators (VCE)
<code>estat (svy)</code>	postestimation statistics for survey data
<code>estimates</code>	cataloging estimation results
† <code>forecast</code>	dynamic forecasts and simulations
† <code>hausman</code>	Hausman's specification test
<code>lincom</code>	point estimates, standard errors, testing, and inference for linear combinations of coefficients
<code>linktest</code>	link test for model specification
*† <code>lrtest</code>	likelihood-ratio test
<code>margins</code>	marginal means, predictive margins, marginal effects, and average marginal effects
<code>marginsplot</code>	graph the results from margins (profile plots, interaction plots, etc.)
<code>nlcom</code>	point estimates, standard errors, testing, and inference for nonlinear combinations of coefficients
<code>predict</code>	predictions, residuals, influence statistics, and other diagnostic measures
<code>predictnl</code>	point estimates, standard errors, testing, and inference for generalized predictions
<code>pwcompare</code>	pairwise comparisons of estimates
<code>suest</code>	seemingly unrelated estimation
<code>test</code>	Wald tests of simple and composite linear hypotheses
<code>testnl</code>	Wald tests of nonlinear hypotheses

\* `estat ic` and `lrtest` are not appropriate after `glm`, `irls`.

† `forecast`, `hausman`, and `lrtest` are not appropriate with `svy` estimation results. `forecast` is also not appropriate with `mi` estimation results.

# predict

## Description for predict

`predict` creates a new variable containing predictions such as expected values, linear predictions, standard errors, residuals, Cook's distance, diagonals of the "hat" matrix, weighted averages, differences between the observed and fitted outcomes, and equation-level scores.

## Menu for predict

Statistics > Postestimation

## Syntax for predict

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

<i>statistic</i>	Description
Main	
<u>mu</u>	expected value of $y$ ; the default
<u>xb</u>	linear prediction $\eta = \mathbf{x}\hat{\beta}$
<u>eta</u>	synonym of <u>xb</u>
<u>stdp</u>	standard error of the linear prediction
<u>anscombe</u>	<a href="#">Anscombe (1953)</a> residuals
<u>cooksd</u>	Cook's distance
<u>deviance</u>	deviance residuals
<u>hat</u>	diagonals of the "hat" matrix
<u>likelihood</u>	a weighted average of standardized deviance and standardized Pearson residuals
<u>pearson</u>	Pearson residuals
<u>response</u>	differences between the observed and fitted outcomes
<u>score</u>	first derivative of the log likelihood with respect to $\mathbf{x}_j\beta$
<u>working</u>	working residuals

<i>options</i>	Description
Options	
<u>nooffset</u>	modify calculations to ignore offset variable
<u>adjusted</u>	adjust deviance residual to speed up convergence
<u>standardized</u>	multiply residual by the factor $(1 - h)^{-1/2}$
<u>studentized</u>	multiply residual by one over the square root of the estimated scale parameter
<u>modified</u>	modify denominator of residual to be a reasonable estimate of the variance of <i>depvar</i>

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

`mu`, `xb`, `stdp`, and `score` are the only statistics allowed with `svy` estimation results.

## Options for predict

## Main

`mu`, the default, specifies that `predict` calculate the expected value of  $y$ , equal to  $g^{-1}(\mathbf{x}\hat{\beta})$  [ $ng^{-1}(\mathbf{x}\hat{\beta})$  for the binomial family].

`xb` calculates the linear prediction  $\eta = \mathbf{x}\hat{\beta}$ .

`eta` is a synonym for `xb`.

`stdp` calculates the standard error of the linear prediction.

`anscombe` calculates the [Anscombe \(1953\)](#) residuals to produce residuals that closely follow a normal distribution.

`cooks` calculates Cook's distance, which measures the aggregate change in the estimated coefficients when each observation is left out of the estimation.

`deviance` calculates the deviance residuals. Deviance residuals are recommended by [McCullagh and Nelder \(1989\)](#) and by others as having the best properties for examining the goodness of fit of a GLM. They are approximately normally distributed if the model is correct. They may be plotted against the fitted values or against a covariate to inspect the model's fit. Also see the `pearson` option below.

`hat` calculates the diagonals of the "hat" matrix, analogous to linear regression.

`likelihood` calculates a weighted average of standardized deviance and standardized Pearson residuals.

`pearson` calculates the Pearson residuals. Pearson residuals often have markedly skewed distributions for nonnormal family distributions. Also see the `deviance` option above.

`response` calculates the differences between the observed and fitted outcomes.

`score` calculates the equation-level score,  $\partial \ln L / \partial (\mathbf{x}_j \beta)$ .

`working` calculates the working residuals, which are response residuals weighted according to the derivative of the link function.

## Options

`nooffset` is relevant only if you specified `offset(varname)` for `glm`. 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  $\mathbf{x}_j \mathbf{b} + \text{offset}_j$ .

`adjusted` adjusts the deviance residual to speed up the convergence to the limiting normal distribution. The adjustment deals with adding to the deviance residual a higher-order term that depends on the variance function family. This option is allowed only when `deviance` is specified.

`standardized` requests that the residual be multiplied by the factor  $(1 - h)^{-1/2}$ , where  $h$  is the diagonal of the hat matrix. This operation is done to account for the correlation between `deprvar` and its predicted value.

`studentized` requests that the residual be multiplied by one over the square root of the estimated scale parameter.

`modified` requests that the denominator of the residual be modified to be a reasonable estimate of the variance of `deprvar`. The base residual is multiplied by the factor  $(k/w)^{-1/2}$ , where  $k$  is either one or the user-specified dispersion parameter and  $w$  is the specified weight (or one if left unspecified).

# margins

## Description for margins

`margins` estimates margins of response for expected values and linear predictions.

## Menu for margins

Statistics > Postestimation

## Syntax for margins

```
margins [marginlist] [, options]
```

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

<i>statistic</i>	Description
<u>mu</u>	expected value of $y$ ; the default
<u>xb</u>	linear prediction $\eta = \mathbf{x}\hat{\beta}$
<u>eta</u>	synonym for <code>xb</code>
<u>stdp</u>	not allowed with <code>margins</code>
<u>anscombe</u>	not allowed with <code>margins</code>
<u>cooksd</u>	not allowed with <code>margins</code>
<u>deviance</u>	not allowed with <code>margins</code>
<u>hat</u>	not allowed with <code>margins</code>
<u>likelihood</u>	not allowed with <code>margins</code>
<u>pearson</u>	not allowed with <code>margins</code>
<u>response</u>	not allowed with <code>margins</code>
<u>score</u>	not allowed with <code>margins</code>
<u>working</u>	not allowed with <code>margins</code>

Statistics not allowed with `margins` are functions of stochastic quantities other than  $\mathbf{e}(\mathbf{b})$ .

For the full syntax, see [R] [margins](#).

## Remarks and examples

[stata.com](http://www.stata.com)

Remarks are presented under the following headings:

*Predictions*

*Other postestimation commands*

## Predictions

### ▷ Example 1

After `glm` estimation, `predict` may be used to obtain various predictions based on the model. In [example 2](#) of [\[R\] glm](#), we mentioned that the complementary log–log link seemed to fit the data better than the logit link. Now, we go back and obtain the fitted values and deviance residuals:

```
. use https://www.stata-press.com/data/r17/ldose
. glm r ldose, family(binomial n) link(logit)
  (output omitted)
. predict mu_logit
(option mu assumed; predicted mean r)
. predict dr_logit, deviance
. quietly glm r ldose, f(binomial n) l(cloglog)
. predict mu_cl
(option mu assumed; predicted mean r)
. predict dr_cl, d
. format mu_logit dr_logit mu_cl dr_cl %9.5f
. list r mu_logit dr_logit mu_cl dr_cl, sep(4)
```

	r	mu_logit	dr_logit	mu_cl	dr_cl
1.	6	3.45746	1.28368	5.58945	0.18057
2.	13	9.84167	1.05969	11.28067	0.55773
3.	18	22.45139	-1.19611	20.95422	-0.80330
4.	28	33.89761	-1.59412	30.36942	-0.63439
5.	52	50.09584	0.60614	47.77644	1.28883
6.	53	53.29092	-0.12716	54.14273	-0.52366
7.	61	59.22216	1.25107	61.11331	-0.11878
8.	60	58.74297	1.59398	59.94723	0.32495

In six of the eight cases,  $|\text{dr\_logit}| > |\text{dr\_cl}|$ . The above represents only one of the many available options for `predict`. See [Hardin and Hilbe \(2018\)](#) for a more in-depth examination.

◀

## Other postestimation commands

### □ Technical note

After `glm` estimation, you may perform any of the postestimation commands that you would perform after any other kind of estimation in Stata; see [\[U\] 20 Estimation and postestimation commands](#). Below, we test the joint significance of all the interaction terms.

```

. use https://www.stata-press.com/data/r17/beetle, clear
. glm r beetle#c.ldose, family(binomial n) link(cloglog)
  (output omitted)
. testparm i.beetle beetle#c.ldose
( 1)  [r]2.beetle = 0
( 2)  [r]3.beetle = 0
( 3)  [r]2.beetle#c.ldose = 0
( 4)  [r]3.beetle#c.ldose = 0

      chi2( 4) = 249.69
      Prob > chi2 = 0.0000

```

If you wanted to print the variance–covariance matrix of the estimators, you would type `estat vce`.

If you use the `linktest` postestimation command, you must also specify the `family()` and `link()` options; see [R] [linktest](#). □

## Methods and formulas

We follow the terminology used in *Methods and formulas* of [R] [glm](#).

The deviance residual calculated by `predict` following `glm` is  $r_j^D = \text{sign}(y_j - \hat{\mu}_j) \sqrt{d_j^2}$ .

The Pearson residual calculated by `predict` following `glm` is

$$r_j^P = \frac{y_j - \hat{\mu}_j}{\sqrt{V(\hat{\mu}_j)}}$$

where  $V(\hat{\mu}_j)$  is the family-specific variance function.

$$V(\hat{\mu}_j) = \begin{cases} \hat{\mu}_j(1 - \hat{\mu}_j/m_j) & \text{if binomial or Bernoulli } (m_j = 1) \\ \hat{\mu}_j^2 & \text{if gamma} \\ 1 & \text{if Gaussian} \\ \hat{\mu}_j^3 & \text{if inverse Gaussian} \\ \hat{\mu}_j + k\hat{\mu}_j^2 & \text{if negative binomial} \\ \hat{\mu}_j & \text{if Poisson} \end{cases}$$

The response residuals are given by  $r_j^R = y_j - \hat{\mu}_j$ . The working residuals are

$$r_j^W = (y_j - \hat{\mu}_j) \left( \frac{\partial \eta}{\partial \mu} \right)_j$$

and the score residuals are

$$r_j^S = \frac{y_j - \hat{\mu}_j}{V(\hat{\mu}_j)} \left( \frac{\partial \eta}{\partial \mu} \right)_j^{-1}$$

Define  $\widehat{W} = V(\widehat{\mu})$  and  $X$  to be the covariate matrix.  $h_j$ , then, is the  $j$ th diagonal of the hat matrix given by

$$\widehat{H} = \widehat{W}^{1/2} X (X^T \widehat{W} X)^{-1} X^T \widehat{W}^{1/2}$$

As a result, the likelihood residuals are given by

$$r_j^L = \text{sign}(y_j - \hat{\mu}_j) \{h_j(r_j^{P'})^2 + (1 - h_j)(r_j^{D'})^2\}^{1/2}$$

where  $r_j^{P'}$  and  $r_j^{D'}$  are the standardized Pearson and standardized deviance residuals, respectively. By *standardized*, we mean that the residual is divided by  $\{1 - h_j\}^{1/2}$ .

Cook's distance is an overall measure of the change in the regression coefficients caused by omitting the  $i$ th observation from the analysis. Computationally, Cook's distance is obtained as

$$C_j = \frac{(r_j^{P'})^2 h_j}{k(1 - h_j)}$$

where  $k$  is the number of regressors, including the constant.

Anscombe residuals are given by

$$r_j^A = \frac{A(y_j) - A(\hat{\mu}_j)}{A'(\hat{\mu}_j)\{V(\hat{\mu}_j)\}^{1/2}}$$

where

$$A(\cdot) = \int \frac{d\mu}{V^{1/3}(\mu)}$$

Deviance residuals may be adjusted (`predict, adjusted`) to make the following correction:

$$r_j^{Da} = r_j^D + \frac{1}{6}\rho_3(\theta)$$

where  $\rho_3(\theta)$  is a family-specific correction. See [Hardin and Hilbe \(2018\)](#) for the exact forms of  $\rho_3(\theta)$  for each family.

## References

- Anscombe, F. J. 1953. Contribution of discussion paper by H. Hotelling "New light on the correlation coefficient and its transforms". *Journal of the Royal Statistical Society, Series B* 15: 229–230. <https://doi.org/10.1111/j.2517-6161.1953.tb00136.x>.
- Hardin, J. W., and J. M. Hilbe. 2018. *Generalized Linear Models and Extensions*. 4th ed. College Station, TX: Stata Press.
- McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*. 2nd ed. London: Chapman & Hall/CRC.
- Newson, R. B. 2013. [Attributable and unattributable risks and fractions and other scenario comparisons](#). *Stata Journal* 13: 672–698.

## Also see

- [R] [glm](#) — Generalized linear models
- [R] [regress postestimation](#) — Postestimation tools for regress
- [U] [20 Estimation and postestimation commands](#)