### Postestimation commands

The following postestimation commands are available after `nbreg` and `gnbreg`:

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

* `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 numbers of events, incidence rates, probabilities, linear predictions, standard errors, and predicted values.

Menu for predict

Statistics > Postestimation

Syntax for predict

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

```
predict [ `type` ] { stub* | newvar `reg` newvar `disp` } [ `if` ] [ `in` ], scores
```

<table>
<thead>
<tr>
<th>statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main</td>
<td>number of events; the default</td>
</tr>
<tr>
<td>n</td>
<td>incidence rate (equivalent to predict ..., n nooffset)</td>
</tr>
<tr>
<td>ir</td>
<td>probability Pr(y_j = n)</td>
</tr>
<tr>
<td>pr(n)</td>
<td>probability Pr(a ≤ y_j ≤ b)</td>
</tr>
<tr>
<td>xb</td>
<td>linear prediction</td>
</tr>
<tr>
<td>stdp</td>
<td>standard error of the linear prediction</td>
</tr>
</tbody>
</table>

In addition, relevant only after gnbreg are the following:

<table>
<thead>
<tr>
<th>statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main</td>
<td>predicted values of α_j</td>
</tr>
<tr>
<td>alpha</td>
<td>predicted values of lnα_j</td>
</tr>
<tr>
<td>lnalpha</td>
<td>standard error of predicted lnα_j</td>
</tr>
</tbody>
</table>

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

n, the default, calculates the predicted number of events, which is exp(x_jβ) if neither offset(varname_o) nor exposure(varname_e) was specified when the model was fit; exp(x_jβ + offset_j) if offset() was specified; or exp(x_jβ) × exposure_j if exposure() was specified.

ir calculates the incidence rate exp(x_jβ), which is the predicted number of events when exposure is 1. This is equivalent to specifying both the n and the nooffset options.
\texttt{pr(n)} calculates the probability $Pr(y_j = n)$, where $n$ is a nonnegative integer that may be specified as a number or a variable.

\texttt{pr(a,b)} calculates the probability $Pr(a \leq y_j \leq b)$, where $a$ and $b$ are nonnegative integers that may be specified as numbers or variables;

- $b$ missing ($b \geq \cdot$) means $+\infty$;
- \texttt{pr(20,\cdot)} calculates $Pr(y_j \geq 20)$;
- \texttt{pr(20,b)} calculates $Pr(y_j \geq 20)$ in observations for which $b \geq \cdot$ and calculates $Pr(20 \leq y_j \leq b)$ elsewhere.

\texttt{pr(\cdot,b)} produces a syntax error. A missing value in an observation of the variable $a$ causes a missing value in that observation for \texttt{pr(a,b)}.

\texttt{xb} calculates the linear prediction, which is $x_j\beta$ if neither \texttt{offset()} nor \texttt{exposure()} was specified; $x_j\beta + \text{offset}_j$ if \texttt{offset()} was specified; or $x_j\beta + \ln(\text{exposure}_j)$ if \texttt{exposure()} was specified; see \texttt{nooffset} below.

\texttt{stdp} calculates the standard error of the linear prediction.

\texttt{alpha, lnalpha, and stdplna} are relevant after \texttt{gnbreg} estimation only; they produce the predicted values of $\alpha_j$, $\ln\alpha_j$, and the standard error of the predicted $\ln\alpha_j$, respectively.

\texttt{nooffset} is relevant only if you specified \texttt{offset()} or \texttt{exposure()} when you fit the model. It modifies the calculations made by \texttt{predict} so that they ignore the offset or exposure variable; the linear prediction is treated as $x_j\beta$ rather than as $x_j\beta + \text{offset}_j$ or $x_j\beta + \ln(\text{exposure}_j)$. Specifying \texttt{predict \dots, nooffset} is equivalent to specifying \texttt{predict \dots, ir}.

\texttt{scores} calculates equation-level score variables.

- The first new variable will contain $\partial \ln L / \partial (x_j\beta)$.
- The second new variable will contain $\partial \ln L / \partial (\ln\alpha_j)$ for \texttt{dispersion(mean)} and \texttt{gnbreg}.
- The second new variable will contain $\partial \ln L / \partial (\ln\delta)$ for \texttt{dispersion(constant)}. 
margins

Description for margins

margins estimates margins of response for numbers of events, incidence rates, probabilities, linear predictions, and predicted values.

Menu for margins

Statistics > Postestimation

Syntax for margins

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

statistic        Description

n               number of events; the default
ir              incidence rate (equivalent to predict ..., n nooffset)
pr(n)           probability $Pr(y_j = n)$
pr(a,b)         probability $Pr(a \leq y_j \leq b)$
xb              linear prediction
stdp            not allowed with margins

In addition, relevant only after gnbreg are the following:

statistic        Description

alpha           predicted values of $\alpha_j$
lnalpha         predicted values of $\ln\alpha_j$
stdplna         not allowed with margins

Statistics not allowed with margins are functions of stochastic quantities other than $e(b)$. For the full syntax, see [R] margins.
Remarks and examples

After `nbreg` and `gnbreg`, `predict` returns the expected number of deaths per cohort and the probability of observing the number of deaths recorded or fewer.

```
. use https://www.stata-press.com/data/r16/rod93
. nbreg deaths i.cohort, nolog
```

```
  Negative binomial regression
  Number of obs = 21
  LR chi2(2)   = 0.14
  Dispersion = mean
  Prob > chi2 = 0.9307
  Log likelihood = -108.48841
  Pseudo R2    = 0.0007

  deaths | Coef.   Std. Err.     z    P>|z|   [95% Conf. Interval]
---------+----------------------------------------------------------------
  cohort  |                                                                
     1960-1967 |  .0591305   .2978419   0.20   0.843  -.5246289    .642898
     1968-1976 | -.0538792   .2981621  -0.18   0.857  -.6382662   .5305077
     _cons   |  4.435906   .2107213   21.05   0.000     4.0229    4.848912
---------+----------------------------------------------------------------
      /lnalpha | -1.207379   .3108622  -1.8166 0.0707  -.8021297  -.4071934
       alpha  |  .29898   .0929416   .3195 0.0014   .1625683   .5498555

LR test of alpha=0: chibar2(01) = 434.62 Prob => chibar2 = 0.000
```

```
. predict count
(option n assumed; predicted number of events)
. predict p, pr(0, deaths)
. summarize deaths count p
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>deaths</td>
<td>21</td>
<td>84.66667</td>
<td>48.84192</td>
<td>10</td>
<td>197</td>
</tr>
<tr>
<td>count</td>
<td>21</td>
<td>84.66667</td>
<td>4.00773</td>
<td>80</td>
<td>89.57143</td>
</tr>
<tr>
<td>p</td>
<td>21</td>
<td>.4991542</td>
<td>.2743702</td>
<td>.0070255</td>
<td>.9801285</td>
</tr>
</tbody>
</table>

The expected number of deaths ranges from 80 to 90. The probability \( Pr(y_i \leq \text{deaths}) \) ranges from 0.007 to 0.98.

Methods and formulas

In the following, we use the same notation as in \[R\] `nbreg`.

Methods and formulas are presented under the following headings:

*Mean-dispersion model*

*Constant-dispersion model*

**Mean-dispersion model**

The equation-level scores are given by

\[
\text{score}(x\beta)_j = p_j(y_j - \mu_j)
\]

\[
\text{score}(\tau)_j = -m \left\{ \frac{\alpha_j(\mu_j - y_j)}{1 + \alpha_j \mu_j} - \ln(1 + \alpha_j \mu_j) + \psi(y_j + m) - \psi(m) \right\}
\]

where \( \tau_j = \ln\alpha_j \) and \( \psi(z) \) is the digamma function.
Constant-dispersion model

The equation-level scores are given by

\[
\begin{align*}
score(\mathbf{x}\beta)_j &= m_j \{\psi(y_j + m_j) - \psi(m_j) + \ln(p)\} \\
score(\tau)_j &= y_j - (y_j + m_j)(1 - p) - score(\mathbf{x}\beta)_j
\end{align*}
\]

where \(\tau_j = \ln\delta_j\).

Reference


Also see

[R] nbreg — Negative binomial regression
[U] 20 Estimation and postestimation commands