ologit postestimation — Postestimation tools for ologit[+]

[+]Postestimation after `ologit` includes features that are part of StataNow.

Postestimation commands

The following postestimation command is of special interest after `ologit`:

Command	Description
[+][*] `estat parallel`	test the parallel lines assumption (proportional odds assumption)

[+]This command is part of StataNow.

[*]`estat parallel` is not available after models fit with constraints, offsets, time-series operators, `iweight`s, `pweight`s, `vce()`'s *vcetype*s other than `oim` or `opg`, or any of the following prefixes: `bayes`, `bayesboot`, `bootstrap`, `fmm`, `jackknife`, `mi estimate`, `rolling`, `statsby`, and `svy`.

The following postestimation commands are available after `ologit`:

Command	Description
`contrast`	contrasts and ANOVA-style joint tests of parameters
`estat ic`	Akaike's, consistent Akaike's, corrected Akaike's, and Schwarz's Bayesian information criteria (AIC, CAIC, AICc, and BIC, respectively)
`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
`etable`	table of estimation results
[*] `forecast`	dynamic forecasts and simulations
[*] `hausman`	Hausman's specification test
`lincom`	point estimates, standard errors, testing, and inference for linear combinations of parameters
`linktest`	link test for model specification
[*] `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 parameters
`predict`	probabilities, linear predictions and their SEs, etc.
`predictnl`	point estimates, standard errors, testing, and inference for generalized predictions
`pwcompare`	pairwise comparisons of parameters
`suest`	seemingly unrelated estimation
`test`	Wald tests of simple and composite linear hypotheses
`testnl`	Wald tests of nonlinear hypotheses

[*]`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 probabilities, linear predictions, and standard errors.

Menu for predict

Statistics > Postestimation

Syntax for predict

predict [*type*] { *stub** | *newvar* | *newvarlist* } [*if*] [*in*] [, *statistic*
outcome(*outcome*) nooffset]

predict [*type*] *stub** [*if*] [*in*] , scores

statistic	Description
Main	
pr	predicted probabilities; the default
xb	linear prediction
stdp	standard error of the linear prediction

You specify one or *k* new variables with pr, where *k* is the number of outcomes. If you specify one new variable and you do not specify outcome(), then outcome(#1) is assumed.

You specify one new variable with xb and stdp.

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

Main

pr, the default, computes the predicted probabilities for all outcomes or for a specific outcome. To compute probabilities for all outcomes, you specify *k* new variables, where *k* is the number of categories of the dependent variable. Alternatively, you can specify *stub**; in which case, pr will store predicted probabilities in variables *stub*1, *stub*2, ..., *stubk*. To compute the probability for a specific outcome, you specify one new variable and, optionally, the outcome value in option outcome(); if you omit outcome(), the first outcome value, outcome(#1), is assumed.

Say that you fit a model by typing *estimation_cmd* y x1 x2, and y takes on four values. Then, you could type predict p1 p2 p3 p4 to obtain all four predicted probabilities; alternatively, you could type predict p* to generate the four predicted probabilities. To compute specific probabilities one at a time, you can type predict p1, outcome(#1) (or simply predict p1), predict p2, outcome(#2), and so on. See option outcome() for other ways to refer to outcome values.

xb calculates the linear prediction. You specify one new variable, for example, predict linear, xb. The linear prediction is defined, ignoring the contribution of the estimated cutpoints.

stdp calculates the standard error of the linear prediction. You specify one new variable, for example, predict se, stdp.

outcome(*outcome*) specifies for which outcome the predicted probabilities are to be calculated. outcome() should contain either one value of the dependent variable or one of #1, #2, ..., with #1 meaning the first category of the dependent variable, #2 meaning the second category, etc. outcome() is available only with the default pr option.

nooffset is relevant only if you specified offset(*varname*) for ologit. 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 $\mathbf{x}_j\mathbf{b} + \text{offset}_j$. nooffset is not allowed with scores.

scores calculates equation-level score variables. The number of score variables created will equal the number of outcomes in the model. If the number of outcomes in the model was k, then

the first new variable will contain $\partial \ln L/\partial(\mathbf{x}_j\mathbf{b})$;

the second new variable will contain $\partial \ln L/\partial\kappa_1$;

the third new variable will contain $\partial \ln L/\partial\kappa_2$;

...

and the kth new variable will contain $\partial \ln L/\partial\kappa_{k-1}$, where κ_i refers to the ith cutpoint.

margins

Description for margins

margins estimates margins of response for probabilities and linear predictions.

Menu for margins

Statistics > Postestimation

Syntax for margins

margins [*marginlist*] [, *options*]

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

statistic	Description
default	probabilities for each outcome
pr	probability for a specified outcome
xb	linear prediction
stdp	not allowed with margins

pr defaults to the first outcome.

Statistics not allowed with margins are functions of stochastic quantities other than e(b).

For the full syntax, see [R] **margins**.

estat

Description for estat parallel

`estat parallel` is part of StataNow. It performs several tests of the parallel lines assumption, also known as the proportional odds assumption, after fitting an ordered logit model. The supported tests are the Brant, likelihood-ratio, score, Wald, and Wolfe–Gould tests.

Menu for estat

Statistics > Postestimation

Syntax for estat parallel

`estat parallel` [, *options*]

options	Description
`all`	display all five tests; the default
`brant`	display the Brant test
`lr`	display the likelihood-ratio test
`score`	display the score test
`wald`	display the Wald test
`wgould`	display the Wolfe–Gould test

`collect` is allowed; see **[U] 11.1.10 Prefix commands**.

Options for estat parallel

Main

`all` displays the results of all five tests of the proportional odds assumption: Brant, likelihood-ratio, score, Wald, and Wolfe–Gould. `all` is the default.

`brant` displays the results of the Brant test for proportional odds. This is a Wald test constructed by fitting multiple logistic regression models to a dichotomized outcome variable.

`lr` displays the results of the likelihood-ratio test for proportional odds. This test compares the likelihood of the proportional odds model (as the reduced model) with that of the generalized ordered logit model (as the full model). The generalized ordered logit model permits different regression slopes for different levels of the outcome variable.

`score` displays the results of the score test for proportional odds. This is the score test of the equality of coefficients of predictors across outcome categories in a generalized ordered logit model. It does not require fitting the generalized ordered logit model.

`wald` displays the results of the Wald test for proportional odds. The Wald test is constructed by fitting the generalized ordered logit model and then testing whether the regression slopes are constant over different levels of the outcome variable.

`wgould` displays the results of the Wolfe–Gould test for proportional odds. This is a likelihood-ratio test constructed by fitting multiple logistic regression models to a dichotomized outcome variable.

Remarks and examples

See [U] 20 Estimation and postestimation commands for instructions on obtaining the variance–covariance matrix of the estimators, predicted values, and hypothesis tests. Also see [R] lrtest for performing likelihood-ratio tests.

Remarks are presented under the following headings:

predict: Predicting the probabilities of ordinal outcomes

▷ Example 1

In example 2 of [R] **ologit**, we fit the model `ologit rep77 foreign length mpg`. The `predict` command can be used to obtain the predicted probabilities.

We type `predict` followed by the names of the new variables to hold the predicted probabilities, ordering the names from low to high. In our data, the lowest outcome is "poor", and the highest is "excellent". We have five categories, so we must type five names following `predict`; the choice of names is up to us:

```
. use https://www.stata-press.com/data/r19/fullauto
(Automobile models)
. ologit rep77 foreign length mpg
  (output omitted)
. predict poor fair avg good exc
(option pr assumed; predicted probabilities)
. list exc good make model rep78 if rep77>=., sep(4) divider
```

	exc	good	make	model	rep78
3.	.0033341	.0393056	AMC	Spirit	.
10.	.0098392	.1070041	Buick	Opel	.
32.	.0023406	.0279497	Ford	Fiesta	Good
44.	.015697	.1594413	Merc.	Monarch	Average
53.	.065272	.4165188	Peugeot	604	.
56.	.005187	.059727	Plym.	Horizon	Average
57.	.0261461	.2371826	Plym.	Sapporo	.
63.	.0294961	.2585825	Pont.	Phoenix	.

The eight cars listed were introduced after 1977, so they do not have 1977 repair records in our data. We predicted what their 1977 repair records might have been using the fitted model. We see that, based on its characteristics, the Peugeot 604 had about a $41.65 + 6.53 \approx 48.2\%$ chance of a good or an excellent repair record. The Ford Fiesta, which had only a 3% chance of a good or an excellent repair record, in fact, had a good record when it was introduced in the following year.

◁

❑ Technical note

For ordered logit, `predict, xb` produces $S_j = x_{1j}\beta_1 + x_{2j}\beta_2 + \cdots + x_{kj}\beta_k$. The ordered-logit predictions are then the probability that $S_j + u_j$ lies between a pair of cutpoints, κ_{i-1} and κ_i. Some handy formulas are

$$\Pr(S_j + u_j < \kappa) = 1/(1 + e^{S_j - \kappa})$$
$$\Pr(S_j + u_j > \kappa) = 1 - 1/(1 + e^{S_j - \kappa})$$
$$\Pr(\kappa_1 < S_j + u_j < \kappa_2) = 1/(1 + e^{S_j - \kappa_2}) - 1/(1 + e^{S_j - \kappa_1})$$

Rather than using `predict` directly, we could calculate the predicted probabilities by hand. If we wished to obtain the predicted probability that the repair record is excellent and the probability that it is good, we look back at `ologit`'s output to obtain the cutpoints. We find that "good" corresponds to the interval `/cut3` $< S_j + u <$ `/cut4` and "excellent" to the interval $S_j + u >$ `/cut4`:

```
. predict score, xb
. generate probgood = 1/(1+exp(score-_b[/cut4])) - 1/(1+exp(score-_b[/cut3]))
. generate probexc = 1 - 1/(1+exp(score-_b[/cut4]))
```

The results of our calculation will be the same as those produced in the previous example. We refer to the estimated cutpoints just as we would any coefficient, so `_b[/cut3]` refers to the value of the `/cut3` coefficient; see **[U] 13.5 Accessing coefficients and standard errors**.

❑

estat parallel: Testing the proportional odds assumption (StataNow)

Introduction

Ordered logit models rely on the parallel lines assumption, which states that the cumulative probability curves for each outcome category plotted against a predictor should be parallel. Mathematically, this means that the effect of the predictor on each outcome category can be expressed as a single coefficient; only the intercepts (cutpoints) vary by outcome category. With multiple predictors, the lines are actually planes, and there is a vector of coefficients—one coefficient for each predictor.

In the context of an ordinal logit model, the parallel lines assumption is more commonly known as the proportional odds assumption, which equivalently states that the odds of an ordinal outcome being in the next-highest category are the same for all outcome categories. In what follows, we will refer to the assumption as the proportional odds assumption.

Several tests have been proposed in the literature to assess the proportional odds assumption. They can be divided into two main groups. One group uses a generalized ordered logit model (GOLM) that allows the effects of each predictor (coefficients) to vary across the outcome categories (Agresti 2010, 121). The other group uses multiple logistic regression models fit to a dichotomized outcome variable. `estat parallel` supports five tests of the proportional odds assumption after fitting an ordered logit model by using `ologit`: Brant, score, likelihood-ratio, Wald, and Wolfe–Gould. The score, likelihood-ratio, and Wald tests fall into the first group, and the Brant and Wolfe–Gould tests fall into the second group.

The score test is the most commonly used test for proportional odds, in part because it "evaluates the rate of change of the log likelihood only at the null hypothesis", which avoids fitting a more complicated GOLM (Agresti 2010, 70). The Wald and likelihood-ratio tests require fitting the GOLM. The Wald test of proportional odds is the Wald test that the coefficients for each predictor are constant over all categories. The likelihood-ratio test compares the likelihood of the GOLM with that of the proportional-odds logit model fit by `ologit`.

To test the proportional odds assumption without fitting the GOLM, Brant (1990) proposed fitting separate binary logistic regression models to each dichotomized outcome. In these separate logistic regression models, the outcome is 1 if the response is greater than or equal to the higher category; otherwise, it is 0. By combining the results of multiple binary logistic regressions, Brant constructs a Wald-type test of the proportional odds assumption. Wolfe and Gould (1998) devised a test that also involves fitting multiple binary logistic regression models, but in their case, the results are combined to construct a likelihood-ratio test.

Under the null hypothesis of proportional odds, each of the five test statistics follows a χ^2 distribution with degrees of freedom equal to the difference between the number of parameters in the GOLM and the proportional odds model.

There are cases when some tests may not be available. When the GOLM fails to converge, the Wald and likelihood-ratio tests cannot be performed. The Brant and Wolfe–Gould tests cannot be performed when one of the separate binary models fails to converge, such as when there are data separation issues (Tang, He, and Tu 2023). Liu et al. (2023) compare the performance of the five tests and find them to have similar properties when the sample size is large, but the Brant and Wolfe–Gould tests appear to control the type I error better when the sample size is small relative to the number of predictors. A simulation study by Buis and Williams (2013) found the Wolfe–Gould test to do the best job controlling type I error, particularly when the sample size is small or the number of outcome categories is large.

By default, `estat parallel` reports all five tests, but you can specify the options to display specific tests. The command reports missing values for the tests that could not be performed.

If the proportional odds assumption is violated, Harrell (2022) states that the results of a proportional odds model are still meaningful and can be used, for example, to determine which level of a categorical predictor variable is associated with the most favorable response. Alternatives to the proportional odds model include the GOLM and the stereotype logistic regression model developed by Anderson (1984). The GOLM can be fit using the community-contributed `gologit2` command (Williams 2006), whereas the stereotype model is fit using the `slogit` command.

Testing for proportional odds

In example 2 of [R] **ologit**, we fit a proportional odds model to the 1977 repair records of 66 cars using a variation of the automobile dataset described in **[U] 1.2.2 Example datasets**. Here we test the proportional odds assumption from that model.

▷ Example 2: Perform all proportional-odds tests

We continue with example 1 above and use `estat parallel` to conduct tests of the proportional odds assumption.

```
. estat parallel

Tests of proportional odds assumption

Number of obs              = 66
Number of predictors       =  3
Number of outcome levels =  5
Degrees of freedom         =  9

                  Test        chi2   P>chi2
 
                 Brant     9.89864    0.359
      Likelihood-ratio     26.2457    0.002
                 Score     22.5998    0.007
                  Wald     9.08136    0.430
           Wolfe-Gould     19.1206    0.024
```

By default, `estat parallel` performs five tests of the proportional odds assumption. The null hypothesis of each test is that the proportional odds assumption is valid, so failure to reject the null signifies that we do not have evidence that this assumption has been violated.

The five different tests do not all agree. The Brant and Wald tests do not indicate a violation of the proportional odds assumption, but the other tests do indicate a departure from it. To investigate what is going on, we examine the distribution of `rep77` by using the `tabulate` command.

```
. tabulate rep77, nolabel

     Repair
record 1977        Freq.     Percent        Cum.

          1            3        4.55        4.55
          2           11       16.67       21.21
          3           27       40.91       62.12
          4           20       30.30       92.42
          5            5        7.58      100.00

      Total           66      100.00
```

There are only 3 observations for the lowest category with the value of 1 for `rep77` and only 5 observations for the highest category with the value of 5. The tests performed by `estat parallel` use the results of the GOLM and its dichotomized approximation, both of which are sensitive to categories with few observations. For the purpose of demonstration, below we collapse the highest and lowest categories of `rep77` into adjacent categories and refit the `ologit` model.

We define a new variable, r7, that collapses categories 1 and 2 into a single category and likewise collapses categories 4 and 5. We then fit an ordered logit model using foreign, length, and mpg to predict the new outcome r7.

```
. generate r7 = rep77
(8 missing values generated)
. replace r7 = 2 if r7 == 1
(3 real changes made)
. replace r7 = 4 if r7 == 5
(5 real changes made)
. ologit r7 foreign length mpg
Iteration 0:  Log likelihood = -70.110919
Iteration 1:  Log likelihood = -59.651579
Iteration 2:  Log likelihood = -59.473823
Iteration 3:  Log likelihood = -59.473274
Iteration 4:  Log likelihood = -59.473274
Ordered logistic regression                            Number of obs =     66
                                                       LR chi2(3)    =  21.28
                                                       Prob > chi2   = 0.0001
Log likelihood = -59.473274                            Pseudo R2     = 0.1517
```

r7	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
foreign	2.804298	.8376428	3.35	0.001	1.162548	4.446047
length	.0900057	.0265638	3.39	0.001	.0379417	.1420698
mpg	.2766396	.0913218	3.03	0.002	.0976522	.455627
/cut1	22.10749	6.759157			8.859785	35.35519
/cut2	24.37357	6.880472			10.88809	37.85904

We now run estat parallel to test the proportional odds assumption of the collapsed model.

```
. estat parallel
Tests of proportional odds assumption
Number of obs            = 66
Number of predictors     =  3
Number of outcome levels =  3
Degrees of freedom       =  3
```

Test	chi2	P>chi2
Brant	2.01015	0.570
Likelihood-ratio	2.53614	0.469
Score	2.40529	0.493
Wald	2.50717	0.474
Wolfe–Gould	2.50687	0.474

After we collapse the outcome into three categories, none of the five tests indicates a problem with the proportional odds assumption.

◁

Stored results

`estat parallel` stores the following in `r()`:

Scalars

`r(df)`	degrees of freedom for the χ^2 tests of the parallel lines assumption
`r(brant)`	test statistic for the Brant test
`r(p_brant)`	p-value for the Brant test
`r(lr)`	test statistic for the likelihood-ratio test
`r(p_lr)`	p-value for the likelihood-ratio test
`r(score)`	test statistic for the score test
`r(p_score)`	p-value for the score test
`r(wald)`	test statistic for the Wald test
`r(p_wald)`	p-value for the Wald test
`r(wgould)`	test statistic for the Wolfe–Gould test
`r(p_wgould)`	p-value for the Wolfe–Gould test

Matrices

`r(results)`	χ^2 statistics and p-values

Methods and formulas

Please read *Methods and formulas* in [R] **ologit** before reading this section.

`estat parallel` performs five tests of the proportional odds assumption: Brant, likelihood-ratio, score, Wald, and Wolfe–Gould. All five tests have the same null hypothesis: that the proportional odds assumption holds and the slope of each predictor is the same across outcome categories. Under the null hypothesis, all five test statistics follow a χ^2 distribution with $(k-2)\times p$ degrees of freedom, where k is the number of outcome categories and p is the number of predictors.

The score, Wald, and likelihood-ratio tests are constructed using the GOLM, which allows predictors to have different slopes for each outcome category (Agresti 2010, 121). In this model, category $i=1$ is defined as the minimum value of the outcome variable, $i=2$ as the next ordered value, and so on, for the k categories.

According to the GOLM, the probability of observing outcome y_j for any j in $1,\dots,n$ is

$$\begin{aligned}\Pr(y_j=i|\mathbf{x}_j) &= \Pr(\kappa_{i-1} < \mathbf{x}_j\boldsymbol{\beta}_i + u_{ij} \le \kappa_i)\\ &= \frac{1}{1+\exp(-\kappa_i+\mathbf{x}_j\boldsymbol{\beta}_i)} - \frac{1}{1+\exp(-\kappa_{i-1}+\mathbf{x}_j\boldsymbol{\beta}_i)}\end{aligned}$$

where u_{ij} is assumed to be logistically distributed random error. The constant term for each outcome is modeled as a cutpoint, denoted $\kappa_1,\kappa_2,\dots,\kappa_{k-1}$. κ_0 is defined as $-\infty$ and κ_k as $+\infty$.

The proportional odds model is a special case of the GOLM; in the proportional odds model, the vector of coefficients $\boldsymbol{\beta}$ is not indexed by outcome i. Written mathematically, the null hypothesis of the score, Wald, and likelihood-ratio tests is $H_0\colon \boldsymbol{\beta}_1=\boldsymbol{\beta}_2=\cdots=\boldsymbol{\beta}_{k-1}=\boldsymbol{\beta}$. Conceptually, under the null hypothesis, the GOLM reduces to the proportional odds model.

Both the Wald and likelihood-ratio tests require fitting the GOLM to obtain maximum likelihood estimates $\widehat{\boldsymbol{\beta}}_i$. The likelihood-ratio test compares the log likelihood of the GOLM with that of the proportional odds model by using a χ^2 test. See [R] **lrtest** for details about the likelihood-ratio test. The Wald test is conducted by fitting the GOLM and testing $\boldsymbol{\beta}_1=\boldsymbol{\beta}_2=\cdots=\boldsymbol{\beta}_{k-1}$. See [R] **test** for details about the Wald test. Liu et al. (2023) provide results from a simulation study in a scenario with multiple predictors and limited sample size, where the likelihood-ratio test demonstrates inflated type I error under H_a and the Wald test has low power under H_0.

The score test is calculated under the null hypothesis, which does not require the estimation of $\widehat{\boldsymbol{\beta}}_i$. However, Peterson and Harrell (1990) note that the score test can be anticonservative and may perform poorly when some outcome categories have few observations, and Wolfe and Gould (1998) caution that the score test has slower asymptotic convergence than the likelihood-ratio test.

If the GOLM fails to converge (which can occur when the number of categories or predictors is large relative to the number of observations), the likelihood-ratio and Wald tests cannot be performed. In this case, `estat parallel` issues a warning message and reports the Wald and likelihood-ratio statistics as missing. The score test can be computed even when maximum likelihood estimates for $\boldsymbol{\beta}_i$ are not available.

The Brant and Wolfe–Gould tests are based on a series of separate binary logit models for each dichotomized outcome y^*_{ij}, where $y^*_{ij} = 1$ if $y_j \leq i$ (for categories $i = 1, \ldots, k-1$) and $y^*_{ij} = 0$ if $y_j > i$. The development of these tests was spurred by challenges fitting the GOLM. The probability of a given dichotomized outcome is

$$\Pr(y^*_{ij} = 1|\mathbf{x}_j) = \Pr(\mathbf{x}_j\boldsymbol{\beta}^*_i + u^*_{ij} \leq \kappa^*_i)$$

where u^*_{ij} is assumed to be logistically distributed random error and $\boldsymbol{\beta}^*_i$ and κ^*_i take the place of $\boldsymbol{\beta}_i$ and κ_i from the GOLM.

The separate (but inherently correlated) binary logit models are used to obtain maximum likelihood estimates $\widehat{\boldsymbol{\beta}^*_i}$ for $i = 1, \ldots, k-1$. Wolfe and Gould (1998) note that $E(y^*_{ij}|\mathbf{x}_j) = \Pr(y_j \leq i|\mathbf{x}_j)$, and Brant (1990) states that $E(\widehat{\boldsymbol{\beta}^*_i}) \approx \boldsymbol{\beta}_i$. The Brant and Wolfe–Gould tests use the $\widehat{\boldsymbol{\beta}^*_i}$ estimates from these binary models to construct Wald and likelihood-ratio tests, respectively. The null hypothesis for both of these tests is $H^*_0\colon \boldsymbol{\beta}^*_1 = \boldsymbol{\beta}^*_2 = \cdots = \boldsymbol{\beta}^*_{k-1} = \boldsymbol{\beta}^*$, which is an asymptotic approximation of the null hypothesis from the earlier Wald and likelihood-ratio tests.

Conceptually, when we create $k-1$ dichotomized outcomes, $\mathbf{y}^*_1, \ldots, \mathbf{y}^*_{k-1}$, we are storing the dichotomized outcome in a wide format; each $\mathbf{y}^*_i$ is the response variable of a separate logit model. Computationally, we transform the separate outcomes into a long format, where each observation becomes $k-1$ rows: one for each outcome except the last. The result is a column vector of $(k-1) \times n$ dichotomized outcomes that we denote as $\mathbf{y}^*$. We index the resulting dataset by h, where $h = 1, \ldots, (k-1) \times n$. We also create $k-1$ new indicator variables for each observation h in the long dataset, $z_{1h}, \ldots, z_{(k-1)h}$, where each z_{ih} equals 1 if y^*_h corresponds to y^*_{ij} for corresponding row j in the wide dataset and otherwise, $z_{ih} = 0$. Written mathematically, this is expressed as $z_{ih} = \mathbb{1}[\{h \bmod (k-1)\} + (k-1) \times 0^{h \bmod (k-1)} = i]$, where indicator function $\mathbb{1}(A) = 1$ if A is true and $\mathbb{1}(A) = 0$ otherwise.

This enables us to fit a single logit model to the long dataset, yielding parameter estimates that are mathematically equivalent to the estimates from separate logit models. The probability of dichotomized outcome y^*_h is

$$\Pr(y^*_h = 1|\mathbf{x}_h) = \Pr\{(\mathbf{z}_h \otimes \mathbf{x}_h)\boldsymbol{\beta}^*_{\text{long}} - \mathbf{z}_h\boldsymbol{\kappa}^* + u^*_h \leq 0\}$$

where $\mathbf{z}_h = (z_{1h}, \ldots, z_{(k-1)h})$, $\boldsymbol{\beta}^*_{\text{long}} = (\boldsymbol{\beta}^{*\prime}_1, \ldots, \boldsymbol{\beta}^{*\prime}_{k-1})'$, $\boldsymbol{\kappa}^* = (\kappa^*_1, \ldots, \kappa^*_{k-1})'$, $\otimes$ is the Kronecker product (see [M-2] **op_kronecker** for details), and u^*_h is defined similarly to y^*_h.

After estimating $\widehat{\boldsymbol{\beta}^*_{\text{long}}}$ by fitting a single logit model to the long dataset, we use the method of Brant (1990, 1,173) to create a Wald-type test statistic. To calculate the Wolfe–Gould test statistic, we must additionally fit a reduced logit model under H_0^*. The probability of a given dichotomized outcome in the reduced model is

$$\Pr(y_h^* = 1|\mathbf{x}_h) = \Pr(\mathbf{x}_h\boldsymbol{\beta}^* - \mathbf{z}_h\boldsymbol{\kappa}^* + u_h^* \le 0)$$

We then use estimates $\widehat{\boldsymbol{\beta}^*_{\text{long}}}$ and $\widehat{\boldsymbol{\beta}^*}$ to calculate the likelihood under the GOLM and the ordered logit model, respectively; the Wolfe–Gould test statistic is twice the difference between the log likelihoods.

Liu et al. (2023) performed a simulation study with limited sample sizes where they observed that the Brant and Wolfe–Gould tests controlled the type I error better than the score, Wald, and likelihood-ratio tests. If, however, one of the binary logit models fails to converge (which can occur because of perfect prediction, for instance), the Brant and Wolfe–Gould tests cannot be performed. In this case, `estat parallel` issues a warning message and reports their test statistics as missing.

Acknowledgments

The development of `estat parallel` was inspired by community-contributed commands `oparallel` by Maarten L. Buis (2019), `brant` in the `spost13_ado` package by J. Scott Long and Jeremy Freese (2014), and `omodel` by Rory Wolfe and William Gould (1998). Also see `gologit2` by Richard Williams (2006) and `gologit` by Vincent Kang Fu (1998) for fitting GOLMs that relax the proportional odds assumption.

References

Agresti, A. 2010. *Analysis of Ordinal Categorical Data*. 2nd ed. Hoboken, NJ: Wiley. https://doi.org/10.1002/9780470594001.

Anderson, J. A. 1984. Regression and ordered categorical variables (with discussion). *Journal of the Royal Statistical Society*, B ser., 46: 1–30. https://doi.org/10.1111/j.2517-6161.1984.tb01270.x.

Brant, R. 1990. Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics* 46: 1171–1178. https://doi.org/10.2307/2532457.

Buis, M. L. 2019. oparallel: Stata module providing post-estimation command for testing the parallel regression assumption. Statistical Software Components S457720, Department of Economics, Boston College. https://ideas.repec.org/c/boc/bocode/s457720.html.

Buis, M. L., and R. Williams. 2013. Using simulation to inspect the performance of a test in particular tests of the parallel regressions assumption in ordered logit models. Presented at the 2013 German Stata Users Group meeting, Potsdam, Germany, June 7. https://www.stata.com/meeting/germany13/abstracts/materials/de13_buis.pdf.

Fagerland, M. W., and D. W. Hosmer, Jr. 2017. How to test for goodness of fit in ordinal logistic regression models. *Stata Journal* 17: 668–686.

Fu, V. K. 1998. sg88: Estimating generalized ordered logit models. *Stata Technical Bulletin* 44: 27–30. Reprinted in *Stata Technical Bulletin Reprints*, vol. 8, pp. 160–164. College Station, TX: Stata Press.

Harrell, F. E., Jr. 2022. Assessing the proportional odds assumption and its impact. *Statistical Thinking Blog*. https://www.fharrell.com/post/impactpo/.

Liu, A., H. He, X. M. Tu, and W. Tang. 2023. On testing proportional odds assumptions for proportional odds models. *General Psychiatry* 36: e101048. https://doi.org/10.1136/gpsych-2023-101048.

Long, J. S., and J. Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata*. 3rd ed. College Station, TX: Stata Press.

Peterson, B., and F. E. Harrell, Jr. 1990. Partial proportional odds models for ordinal response variables. *Journal of the Royal Statistical Society*, C ser., 39: 205–217. https://doi.org/10.2307/2347760.

Tang, W., H. He, and X. M. Tu. 2023. *Applied Categorical and Count Data Analysis*. 2nd ed. Boca Raton, FL: CRC Press.

Williams, R. 2006. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata Journal* 6: 58–82.

Wolfe, R., and W. W. Gould. 1998. sg76: An approximate likelihood-ratio test for ordinal response models. *Stata Technical Bulletin* 42: 24–27. Reprinted in *Stata Technical Bulletin Reprints*, vol. 7, pp. 199–204. College Station, TX: Stata Press.

Also see

[R] **ologit** — Ordered logistic regression

[U] 20 Estimation and postestimation commands