Title

slogit — Stereotype logistic regression

Syntax
Menu
Description
Options
Remarks and examples
Stored results
Methods and formulas
References
Also see

Syntax

`slogit` *depvar* [*indepvars*] [*if*] [*in*] [*weight*] [, *options*]

options	Description
Model	
`dimension(#)`	dimension of the model; default is `dimension(1)`
`baseoutcome(# | lbl)`	set the base outcome to # or *lbl*; default is the last outcome
`constraints(numlist)`	apply specified linear constraints
`collinear`	keep collinear variables
`nocorner`	do not generate the corner constraints
SE/Robust	
`vce(vcetype)`	*vcetype* may be `oim`, `robust`, `cluster` *clustvar*, `opg`, `bootstrap`, or `jackknife`
Reporting	
`level(#)`	set confidence level; default is `level(95)`
`nocnsreport`	do not display constraints
display_options	control column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
maximize_options	control the maximization process; seldom used
`initialize(initype)`	method of initializing scale parameters; *initype* can be `constant`, `random`, or `svd`; see *Options* for details
`nonormalize`	do not normalize the numeric variables
`coeflegend`	display legend instead of statistics

indepvars may contain factor variables; see [U] **11.4.3 Factor variables**.

`bootstrap`, `by`, `fp`, `jackknife`, `rolling`, `statsby`, and `svy` are allowed; see [U] **11.1.10 Prefix commands**.

Weights are not allowed with the `bootstrap` prefix; see [R] **bootstrap**.

`vce()` and weights are not allowed with the `svy` prefix; see [SVY] **svy**.

`fweight`s, `iweight`s, and `pweight`s are allowed; see [U] **11.1.6 weight**.

`coeflegend` does not appear in the dialog box.

See [U] **20 Estimation and postestimation commands** for more capabilities of estimation commands.

Menu

Statistics > Categorical outcomes > Stereotype logistic regression

Description

`slogit` fits maximum-likelihood stereotype logistic regression models as developed by Anderson (1984). Like multinomial logistic and ordered logistic models, stereotype logistic models are for use with categorical dependent variables. In a multinomial logistic model, the categories cannot be ranked, whereas in an ordered logistic model the categories follow a natural ranking scheme. You can view stereotype logistic models as a compromise between those two models. You can use them when you are unsure of the relevance of the ordering, as is often the case when subjects are asked to assess or judge something. You can also use them in place of multinomial logistic models when you suspect that some of the alternatives are similar. Unlike ordered logistic models, stereotype logistic models do not impose the proportional-odds assumption.

Options

Model

`dimension(#)` specifies the dimension of the model, which is the number of equations required to describe the relationship between the dependent variable and the independent variables. The maximum dimension is $\min(m-1, p)$, where m is the number of categories of the dependent variable and p is the number of independent variables in the model. The stereotype model with maximum dimension is a reparameterization of the multinomial logistic model.

`baseoutcome(#` | *lbl*`)` specifies the outcome level whose scale parameters and intercept are constrained to be zero. The base outcome may be specified as a number of a label. By default, `slogit` assumes that the outcome levels are ordered and uses the largest level of the dependent variable as the base outcome.

`constraints(`*numlist*`)`, `collinear`; see [R] **estimation options**.

By default, the linear equality constraints suggested by Anderson (1984), termed the corner constraints, are generated for you. You can add constraints to these as needed, or you can turn off the corner constraints by specifying `nocorner`. These constraints are in addition to the constraints placed on the ϕ parameters corresponding to `baseoutcome(#)`.

`nocorner` specifies that `slogit` not generate the corner constraints. If you specify `nocorner`, you must specify at least `dimension()` $\times$ `dimension()` constraints for the model to be identified.

SE/Robust

`vce(`*vcetype*`)` specifies the type of standard error reported, which includes types that are derived from asymptotic theory (`oim`, `opg`), that are robust to some kinds of misspecification (`robust`), that allow for intragroup correlation (`cluster` *clustvar*), and that use bootstrap or jackknife methods (`bootstrap`, `jackknife`); see [R] ***vce_option***.

If specifying `vce(bootstrap)` or `vce(jackknife)`, you must also specify `baseoutcome()`.

Reporting

`level(#)`; see [R] **estimation options**.

`nocnsreport`; see [R] **estimation options**.

display_options: noomitted, vsquish, noemptycells, baselevels, allbaselevels, nofvlabel, fvwrap(#), fvwrapon(*style*), cformat(%*fmt*), pformat(%*fmt*), sformat(%*fmt*), and nolstretch; see [R] **estimation options**.

Maximization

maximize_options: difficult, technique(*algorithm_spec*), iterate(#), [no]log, trace, gradient, showstep, hessian, showtolerance, tolerance(#), ltolerance(#), nrtolerance(#), nonrtolerance, and from(*init_specs*); see [R] **maximize**. These options are seldom used.

Setting the optimization type to technique(bhhh) resets the default *vcetype* to vce(opg).

initialize(constant | random | svd) specifies how initial estimates are computed. The default, initialize(constant), is to set the scale parameters to the constant $\min(1/2, 1/d)$, where d is the dimension specified in dimension().

initialize(random) requests that uniformly distributed random numbers between 0 and 1 be used as initial values for the scale parameters. If you specify this option, you should also use set seed to ensure that you can replicate your results; see [R] **set seed**.

initialize(svd) requests that a singular value decomposition (SVD) be performed on the matrix of regression estimates from mlogit to reduce its rank to the dimension specified in dimension(). slogit uses the reduced-rank components of the SVD as initial estimates for the scale and regression coefficients. For details, see *Methods and formulas*.

nonormalize specifies that the numeric variables not be normalized. Normalization of the numeric variables improves numerical stability but consumes more memory in generating temporary double-precision variables. Variables that are of type byte are not normalized, and if initial estimates are specified using the from() option, normalization of variables is not performed. See *Methods and formulas* for more information.

The following option is available with slogit but is not shown in the dialog box:

coeflegend; see [R] **estimation options**.

Remarks and examples

stata.com

Remarks are presented under the following headings:

Introduction
One-dimensional model
Higher-dimension models

Introduction

Stereotype logistic models are often used when subjects are requested to assess or judge something. For example, consider a survey in which consumers may be asked to rate the quality of a product on a scale from 1 to 5, with 1 indicating poor quality and 5 indicating excellent quality. If the categories are monotonically related to an underlying latent variable, the ordered logistic model is appropriate. However, suppose that consumers assess quality not just along one dimension, but rather weigh two or three latent factors. Stereotype logistic regression allows you to specify multiple equations to capture the effects of those latent variables, which you then parameterize in terms of observable characteristics. Unlike with multinomial logit, the number of equations you specify could be less than $m - 1$, where m is the number of categories of the dependent variable.

Stereotype logistic models are also used when categories may be indistinguishable. Suppose that a consumer must choose among A, B, C, or D. Multinomial logistic modeling assumes that the four choices are distinct in the sense that a consumer choosing one of the goods can distinguish its characteristics from the others. If goods A and B are in fact similar, consumers may be randomly picking between the two. One alternative is to combine the two categories and fit a three-category multinomial logistic model. A more flexible alternative is to use a stereotype logistic model.

In the multinomial logistic model, you estimate $m-1$ parameter vectors $\widetilde{\boldsymbol\beta}_k$, $k=1,\dots,m-1$, where m is the number of categories of the dependent variable. The stereotype logistic model is a restriction on the multinomial model in the sense that there are d parameter vectors, where d is between one and $\min(m-1,p)$, and p is the number of regressors. The relationship between the stereotype model's coefficients $\boldsymbol\beta_j$, $j=1,\dots,d$, and the multinomial model's coefficients is $\widetilde{\boldsymbol\beta}_k = -\sum_{j=1}^{d}\phi_{jk}\boldsymbol\beta_j$. The ϕs are scale parameters to be estimated along with the $\boldsymbol\beta_j$s.

Given a row vector of covariates $\mathbf{x}$, let $\eta_k = \theta_k - \sum_{j=1}^{d}\phi_{jk}\mathbf{x}\boldsymbol\beta_j$. The probability of observing outcome k is

$$\Pr(Y_i = k) = \begin{cases} \dfrac{\exp(\eta_k)}{1+\sum_{l=1}^{m-1}\exp(\eta_l)} & k<m \\ \dfrac{1}{1+\sum_{l=1}^{m-1}\exp(\eta_l)} & k=m \end{cases}$$

This model includes a set of θ parameters so that each equation has an unrestricted constant term. If $d=m-1$, the stereotype model is just a reparameterization of the multinomial logistic model. To identify the ϕs and the $\boldsymbol\beta$s, you must place at least d^2 restrictions on the parameters. By default, `slogit` uses the "corner constraints" $\phi_{jj}=1$ and $\phi_{jk}=0$ for $j\neq k$, $k\le d$, and $j\le d$.

For a discussion of the stereotype logistic model, see Lunt (2005).

One-dimensional model

▷ Example 1

We have 2 years of repair rating data on the make, price, mileage rating, and gear ratio of 104 foreign and 44 domestic automobiles (with 13 missing values on repair rating). We wish to fit a stereotype logistic model to discriminate between the levels of repair rating using mileage, price, gear ratio, and origin of the manufacturer. Here is an overview of our data:

```
. use http://www.stata-press.com/data/r13/auto2yr
(Automobile Models)

. tabulate repair

     Repair |
     rating |      Freq.     Percent        Cum.
------------+-----------------------------------
       Poor |          5        3.70        3.70
       Fair |         19       14.07       17.78
    Average |         57       42.22       60.00
       Good |         38       28.15       88.15
  Excellent |         16       11.85      100.00
------------+-----------------------------------
      Total |        135      100.00
```

The variable `repair` can take five values, 1, ..., 5, which represent the subjective rating of the car model's repair record as *Poor*, *Fair*, *Average*, *Good*, and *Excellent*.

We wish to fit the one-dimensional stereotype logistic model

$$\eta_k = \theta_k - \phi_k\left(\beta_1 \texttt{foreign} + \beta_2 \texttt{mpg} + \beta_3 \texttt{price} + \beta_4 \texttt{gratio}\right)$$

for $k < 5$ and $\eta_5 = 0$. To fit this model, we type

```
. slogit repair foreign mpg price gratio
Iteration 0:   log likelihood = -173.78178  (not concave)
Iteration 1:   log likelihood = -164.77316
Iteration 2:   log likelihood =  -161.7069
Iteration 3:   log likelihood = -159.76138
Iteration 4:   log likelihood = -159.34327
Iteration 5:   log likelihood = -159.25914
Iteration 6:   log likelihood = -159.25691
Iteration 7:   log likelihood = -159.25691
Stereotype logistic regression                  Number of obs     =        135
                                                Wald chi2(4)      =       9.33
Log likelihood = -159.25691                     Prob > chi2       =     0.0535
 ( 1)  [phi1_1]_cons = 1
```

repair	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
foreign	5.947382	2.094126	2.84	0.005	1.84297	10.05179
mpg	.1911968	.08554	2.24	0.025	.0235414	.3588521
price	-.0000576	.0001357	-0.42	0.671	-.0003236	.0002083
gratio	-4.307571	1.884713	-2.29	0.022	-8.00154	-.6136017
/phi1_1	1	(constrained)				
/phi1_2	1.262268	.3530565	3.58	0.000	.5702904	1.954247
/phi1_3	1.17593	.3169397	3.71	0.000	.5547394	1.79712
/phi1_4	.8657195	.2411228	3.59	0.000	.3931275	1.338311
/phi1_5	0	(base outcome)				
/theta1	-6.864749	4.21252	-1.63	0.103	-15.12114	1.391639
/theta2	-7.613977	4.861803	-1.57	0.117	-17.14294	1.914981
/theta3	-5.80655	4.987508	-1.16	0.244	-15.58189	3.968786
/theta4	-3.85724	3.824132	-1.01	0.313	-11.3524	3.637922
/theta5	0	(base outcome)				

```
(repair=Excellent is the base outcome)
```

The coefficient associated with the first scale parameter, ϕ_{11}, is 1, and its standard error and other statistics are missing. This is the corner constraint applied to the one-dimensional model; in the header, this constraint is listed as `[phi1_1]_cons = 1`. Also, the ϕ and θ parameters that are associated with the base outcome are identified. Keep in mind, though, that there are no coefficient estimates for `[phi1_5]_cons` or `[theta5]_cons` in the `ereturn` matrix `e(b)`. The Wald statistic is for a test of the joint significance of the regression coefficients on `foreign`, `mpg`, `price`, and `gratio`.

The one-dimensional stereotype model restricts the multinomial logistic regression coefficients $\widetilde{\boldsymbol{\beta}}_k$, $k = 1, \ldots, m-1$ to be parallel; that is, $\widetilde{\boldsymbol{\beta}}_k = -\phi_k \boldsymbol{\beta}$. As Lunt (2001) discusses, in the one-dimensional stereotype model, one linear combination $\mathbf{x}_i\boldsymbol{\beta}$ best discriminates the outcomes of the dependent variable, and the scale parameters ϕ_k measure the distance between the outcome levels and the linear predictor. If $\phi_1 \geq \phi_2 \geq \cdots \phi_{m-1} \geq \phi_m \equiv 0$, the model suggests that the subjective assessment of the dependent variable is indeed ordered. Here the maximum likelihood estimates of the ϕs are not monotonic, as would be assumed in an ordered logit model.

We test that $\phi_1 = \phi_2$ by typing

```
. test [phi1_2]_cons = [phi1_1]_cons
 ( 1)  - [phi1_1]_cons + [phi1_2]_cons = 0
           chi2(  1) =    0.55
         Prob > chi2 =    0.4576
```

Because the two parameters are not statistically different, we decide to add a constraint to force $\phi_1 = \phi_2$:

```
. constraint define 1 [phi1_2]_cons = [phi1_1]_cons
. slogit repair foreign mpg price gratio, constraint(1) nolog
Stereotype logistic regression                   Number of obs   =        135
                                                 Wald chi2(4)    =      21.28
Log likelihood = -159.65769                      Prob > chi2     =     0.0003
 ( 1)  [phi1_1]_cons = 1
 ( 2)  - [phi1_1]_cons + [phi1_2]_cons = 0
```

repair	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
foreign	7.166515	1.690177	4.24	0.000	3.853829	10.4792
mpg	.2340043	.0807042	2.90	0.004	.0758271	.3921816
price	-.000041	.0001618	-0.25	0.800	-.0003581	.000276
gratio	-5.218107	1.798717	-2.90	0.004	-8.743528	-1.692686
/phi1_1	1	(constrained)				
/phi1_2	1	(constrained)				
/phi1_3	.9751096	.1286563	7.58	0.000	.7229478	1.227271
/phi1_4	.7209343	.1220353	5.91	0.000	.4817494	.9601191
/phi1_5	0	(base outcome)				
/theta1	-8.293452	4.645182	-1.79	0.074	-17.39784	.8109368
/theta2	-6.958451	4.629292	-1.50	0.133	-16.0317	2.114795
/theta3	-5.620232	4.953981	-1.13	0.257	-15.32986	4.089392
/theta4	-3.745624	3.809189	-0.98	0.325	-11.2115	3.720249
/theta5	0	(base outcome)				

(repair=Excellent is the base outcome)

The ϕ estimates are now monotonically decreasing and the standard errors of the ϕs are small relative to the size of the estimates, so we conclude that, with the exception of outcomes *Poor* and *Fair*, the groups are distinguishable for the one-dimensional model and that the quality assessment can be ordered.

◁

Higher-dimension models

The stereotype logistic model is not limited to ordered categorical dependent variables; you can use it on nominal data to reduce the dimension of the regressions. Recall that a multinomial model fit to a categorical dependent variable with m levels will have $m - 1$ sets of regression coefficients. However, a model with fewer dimensions may fit the data equally well, suggesting that some of the categories are indistinguishable.

▷ Example 2

As discussed in [R] **mlogit**, we have data on the type of health insurance available to 616 psychologically depressed subjects in the United States (Tarlov et al. 1989; Wells et al. 1989). Patients may have either an indemnity (fee-for-service) plan or a prepaid plan, such as an HMO, or may be uninsured. Demographic variables include age, gender, race, and site.

First, we fit the saturated, two-dimensional model that is equivalent to a multinomial logistic model. We choose the base outcome to be 1 (indemnity insurance) because that is the default for `mlogit`.

```
. use http://www.stata-press.com/data/r13/sysdsn1
(Health insurance data)

. slogit insure age male nonwhite i.site, dim(2) base(1)

Iteration 0:   log likelihood = -534.36165
Iteration 1:   log likelihood = -534.36165

Stereotype logistic regression                    Number of obs   =        615
                                                  Wald chi2(10)   =      38.17
Log likelihood = -534.36165                       Prob > chi2     =     0.0000

 ( 1)  [phi1_2]_cons = 1
 ( 2)  [phi1_3]_cons = 0
 ( 3)  [phi2_2]_cons = 0
 ( 4)  [phi2_3]_cons = 1

------------------------------------------------------------------------------
      insure |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
dim1         |
         age |    .011745   .0061946     1.90   0.058    -.0003962    .0238862
        male |  -.5616934   .2027465    -2.77   0.006    -.9590693   -.1643175
    nonwhite |  -.9747768   .2363213    -4.12   0.000    -1.437958   -.5115955
             |
        site |
          2  |  -.1130359   .2101903    -0.54   0.591    -.5250013    .2989296
          3  |   .5879879   .2279351     2.58   0.010     .1412433    1.034733
-------------+----------------------------------------------------------------
dim2         |
         age |   .0077961   .0114418     0.68   0.496    -.0146294    .0302217
        male |  -.4518496   .3674867    -1.23   0.219     -1.17211     .268411
    nonwhite |  -.2170589   .4256361    -0.51   0.610     -1.05129    .6171725
             |
        site |
          2  |   1.211563   .4705127     2.57   0.010     .2893747    2.133751
          3  |   .2078123   .3662926     0.57   0.570     -.510108    .9257327
-------------+----------------------------------------------------------------
     /phi1_1 |          0  (base outcome)
     /phi1_2 |          1  (constrained)
     /phi1_3 |          0  (omitted)
-------------+----------------------------------------------------------------
     /phi2_1 |          0  (base outcome)
     /phi2_2 |          0  (omitted)
     /phi2_3 |          1  (constrained)
-------------+----------------------------------------------------------------
     /theta1 |          0  (base outcome)
     /theta2 |   .2697127   .3284422     0.82   0.412    -.3740222    .9134476
     /theta3 |  -1.286943   .5923219    -2.17   0.030    -2.447872   -.1260134
------------------------------------------------------------------------------
(insure=Indemnity is the base outcome)
```

For comparison, we also fit the model by using `mlogit`:

```
. mlogit insure age male nonwhite i.site, nolog

Multinomial logistic regression                   Number of obs     =        615
                                                  LR chi2(10)       =      42.99
                                                  Prob > chi2       =     0.0000
Log likelihood = -534.36165                       Pseudo R2         =     0.0387
```

insure	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Indemnity	(base outcome)					
Prepaid						
age	-.011745	.0061946	-1.90	0.058	-.0238862	.0003962
male	.5616934	.2027465	2.77	0.006	.1643175	.9590693
nonwhite	.9747768	.2363213	4.12	0.000	.5115955	1.437958
site						
2	.1130359	.2101903	0.54	0.591	-.2989296	.5250013
3	-.5879879	.2279351	-2.58	0.010	-1.034733	-.1412433
_cons	.2697127	.3284422	0.82	0.412	-.3740222	.9134476
Uninsure						
age	-.0077961	.0114418	-0.68	0.496	-.0302217	.0146294
male	.4518496	.3674867	1.23	0.219	-.268411	1.17211
nonwhite	.2170589	.4256361	0.51	0.610	-.6171725	1.05129
site						
2	-1.211563	.4705127	-2.57	0.010	-2.133751	-.2893747
3	-.2078123	.3662926	-0.57	0.570	-.9257327	.510108
_cons	-1.286943	.5923219	-2.17	0.030	-2.447872	-.1260134

Apart from having opposite signs, the coefficients from the stereotype logistic model are identical to those from the multinomial logit model. Recall the definition of η_k given in the *Remarks and examples*, particularly the minus sign in front of the summation. One other difference in the output is that the constant estimates labeled `/theta` in the `slogit` output are the constants labeled `_cons` in the `mlogit` output.

Next we examine the one-dimensional model.

```
. slogit insure age male nonwhite i.site, dim(1) base(1) nolog
Stereotype logistic regression                  Number of obs     =        615
                                                Wald chi2(5)      =      28.20
Log likelihood = -539.75205                     Prob > chi2       =     0.0000
 ( 1)  [phi1_2]_cons = 1
```

insure	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
age	.0108366	.0061918	1.75	0.080	-.0012992	.0229723
male	-.5032537	.2078171	-2.42	0.015	-.9105678	-.0959396
nonwhite	-.9480351	.2340604	-4.05	0.000	-1.406785	-.489285
site						
2	-.2444316	.2246366	-1.09	0.277	-.6847113	.1958481
3	.556665	.2243799	2.48	0.013	.1168886	.9964415
/phi1_1	0	(base outcome)				
/phi1_2	1	(constrained)				
/phi1_3	.0383539	.4079705	0.09	0.925	-.7612535	.8379613
/theta1	0	(base outcome)				
/theta2	.187542	.3303847	0.57	0.570	-.4600001	.835084
/theta3	-1.860134	.2158898	-8.62	0.000	-2.28327	-1.436997

(insure=Indemnity is the base outcome)

We have reduced a two-dimensional multinomial model to one dimension, reducing the number of estimated parameters by four and decreasing the model likelihood by ≈ 5.4.

`slogit` does not report a model likelihood-ratio test. The test of $d = 1$ (a one-dimensional model) versus $d = 0$ (the null model) does not have an asymptotic χ^2 distribution because the unconstrained ϕ parameters (`/phi1_3` in this example) cannot be identified if $\boldsymbol{\beta} = \mathbf{0}$. More generally, this problem precludes testing any hierarchical model of dimension d versus $d - 1$. Of course, the likelihood-ratio test of a full-dimension model versus $d = 0$ is valid because the full model is just multinomial logistic, and all the ϕ parameters are fixed at 0 or 1.

◁

❑ Technical note

The stereotype model is a special case of the reduced-rank vector generalized linear model discussed by Yee and Hastie (2003). If we define $\eta_{ik} = \theta_k - \sum_{j=1}^{d} \phi_{jk}\mathbf{x}_i\boldsymbol{\beta}_j$, for $k = 1, \ldots, m-1$, we can write the expression in matrix notation as

$$\boldsymbol{\eta}_i = \boldsymbol{\theta} + \boldsymbol{\Phi}\left(\mathbf{x}_i\mathbf{B}\right)'$$

where $\boldsymbol{\Phi}$ is a $(m-1) \times d$ matrix containing the ϕ_{jk} parameters and $\mathbf{B}$ is a $p \times d$ matrix with columns containing the $\boldsymbol{\beta}_j$ parameters, $j = 1, \ldots, d$. The factorization $\boldsymbol{\Phi}\mathbf{B}'$ is not unique because $\boldsymbol{\Phi}\mathbf{B}' = \boldsymbol{\Phi}\mathbf{M}\mathbf{M}^{-1}\mathbf{B}'$ for any nonsingular $d \times d$ matrix $\mathbf{M}$. To avoid this identifiability problem, we choose $\mathbf{M} = \boldsymbol{\Phi}_1^{-1}$, where

$$\boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{\Phi}_1 \\ \boldsymbol{\Phi}_2 \end{pmatrix}$$

and $\boldsymbol{\Phi}_1$ is $d \times d$ of rank d so that

$$\boldsymbol{\Phi}\mathbf{M} = \begin{pmatrix} \mathbf{I}_d \\ \boldsymbol{\Phi}_2\boldsymbol{\Phi}_1^{-1} \end{pmatrix}$$

and $\mathbf{I}_d$ is a $d \times d$ identity matrix. Thus the corner constraints used by `slogit` are $\phi_{jj} \equiv 1$ and $\phi_{jk} \equiv 0$ for $j \neq k$ and $k, j \leq d$.

❑

Stored results

slogit stores the following in e():

Scalars	
e(N)	number of observations
e(k)	number of parameters
e(k_indvars)	number of independent variables
e(k_out)	number of outcomes
e(k_eq)	number of equations in e(b)
e(k_eq_model)	number of equations in overall model test
e(df_m)	Wald test degrees of freedom
e(df_0)	null model degrees of freedom
e(k_dim)	model dimension
e(i_base)	base outcome index
e(ll)	log likelihood
e(ll_0)	null model log likelihood
e(N_clust)	number of clusters
e(chi2)	χ^2
e(p)	significance
e(ic)	number of iterations
e(rank)	rank of e(V)
e(rc)	return code
e(converged)	1 if converged, 0 otherwise
Macros	
e(cmd)	slogit
e(cmdline)	command as typed
e(depvar)	name of dependent variable
e(indvars)	independent variables
e(wtype)	weight type
e(wexp)	weight expression
e(title)	title in estimation output
e(clustvar)	name of cluster variable
e(out#)	outcome labels, $\# = 1, \ldots,$ e(k_out)
e(chi2type)	Wald; type of model χ^2 test
e(labels)	outcome labels or numeric levels
e(vce)	*vcetype* specified in vce()
e(vcetype)	title used to label Std. Err.
e(opt)	type of optimization
e(which)	max or min; whether optimizer is to perform maximization or minimization
e(ml_method)	type of ml method
e(user)	name of likelihood-evaluator program
e(technique)	maximization technique
e(properties)	b V
e(predict)	program used to implement predict
e(marginsnotok)	predictions disallowed by margins
e(footnote)	program used to implement the footnote display
e(asbalanced)	factor variables fvset as asbalanced
e(asobserved)	factor variables fvset as asobserved
Matrices	
e(b)	coefficient vector
e(outcomes)	outcome values
e(Cns)	constraints matrix
e(ilog)	iteration log (up to 20 iterations)
e(gradient)	gradient vector
e(V)	variance–covariance matrix of the estimators
e(V_modelbased)	model-based variance
Functions	
e(sample)	marks estimation sample

Methods and formulas

`slogit` obtains the maximum likelihood estimates for the stereotype logistic model by using `ml`; see [R] **ml**. Each set of regression estimates, one set of $\boldsymbol{\beta}_j$s for each dimension, constitutes one `ml model` equation. The $d \times (m-1)$ ϕs and the $(m-1)$ θs are `ml` ancillary parameters.

Without loss of generality, let the base outcome level be the mth level of the dependent variable. Define the row vector $\boldsymbol{\phi}_k = (\phi_{1k}, \ldots, \phi_{dk})$ for $k = 1, \ldots, m-1$, and define the $p \times d$ matrix $\mathbf{B} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_d)$. For observation i, the log odds of outcome level k relative to level m, $k = 1, \ldots, m-1$ is the index

$$\begin{aligned} \ln\left\{\frac{\Pr(Y_i = k)}{\Pr(Y_i = m)}\right\} &= \eta_{ik} = \theta_k - \boldsymbol{\phi}_k\left(\mathbf{x}_i\mathbf{B}\right)' \\ &= \theta_k - \boldsymbol{\phi}_k\boldsymbol{\nu}_i' \end{aligned}$$

The row vector $\boldsymbol{\nu}_i$ can be interpreted as a latent variable reducing the p-dimensional vector of covariates to a more interpretable $d < p$ dimension.

The probability of the ith observation having outcome level k is then

$$\Pr(Y_i = k) = p_{ik} = \begin{cases} \dfrac{e^{\eta_{ik}}}{1 + \sum_{j=1}^{m-1} e^{\eta_{ij}}}, \text{ if} & k < m \\ \dfrac{1}{1 + \sum_{j=1}^{m-1} e^{\eta_{ij}}}, \text{ if} & k = m \end{cases}$$

from which the log-likelihood function is computed as

$$L = \sum_{i=1}^{n} w_i \sum_{k=1}^{m} I_k(y_i)\ln(p_{ik}) \tag{1}$$

Here w_i is the weight for observation i and

$$I_k(y_i) = \begin{cases} 1, \text{ if observation } y_i \text{ has outcome } k \\ 0, \text{ otherwise} \end{cases}$$

Numeric variables are normalized for numerical stability during optimization where a new double-precision variable $\widetilde{x}_j$ is created from variable x_j, $j = 1, \ldots, p$, such that $\widetilde{x}_j = (x_j - \bar{x}_j)/s_j$. This feature is turned off if you specify `nonormalize`, or if you use the `from()` option for initial estimates. Normalization is not performed on byte variables, including the indicator variables generated by [R] **xi**. The linear equality constraints for regression parameters, if specified, must be scaled also. Assume that a constraint is applied to the regression parameter associated with variable j and dimension i, β_{ji}, and the corresponding element of the constraint matrix (see [P] **makecns**) is divided by s_j.

After convergence, the parameter estimates for variable j and dimension i—$\widetilde{\beta}_{ji}$, say—are transformed back to their original scale, $\beta_{ji} = \widetilde{\beta}_{ji}/s_j$. For the intercepts, you compute

$$\theta_k = \widetilde{\theta}_k + \sum_{i=1}^{d} \phi_{ik} \sum_{j=1}^{p} \frac{\widetilde{\beta}_{ji}\bar{x}_j}{s_j}$$

Initial values are computed using estimates obtained using `mlogit` to fit a multinomial logistic model. Let the $p \times (m-1)$ matrix $\widetilde{\mathbf{B}}$ contain the multinomial logistic regression parameters less the $m-1$ intercepts. Each ϕ is initialized with constant values $\min(1/2, 1/d)$, the `initialize(constant)` option (the default), or, with uniform random numbers, the `initialize(random)` option. Constraints are then applied to the starting values so that the structure of the $(m-1) \times d$ matrix $\boldsymbol{\Phi}$ is

$$\boldsymbol{\Phi} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{m-1} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_d \\ \widetilde{\boldsymbol{\Phi}} \end{pmatrix}$$

where $\mathbf{I}_d$ is a $d \times d$ identity matrix. Assume that only the corner constraints are used, but any constraints you place on the scale parameters are also applied to the initial scale estimates, so the structure of $\boldsymbol{\Phi}$ will change accordingly. The ϕ parameters are invariant to the scale of the covariates, so initial estimates in $[\,0,1\,]$ are reasonable. The constraints guarantee that the rank of $\boldsymbol{\Phi}$ is at least d, so the initial estimates for the stereotype regression parameters are obtained from $\mathbf{B} = \widetilde{\mathbf{B}}\boldsymbol{\Phi}(\boldsymbol{\Phi}'\boldsymbol{\Phi})^{-1}$.

One other approach for initial estimates is provided: `initialize(svd)`. It starts with the `mlogit` estimates and computes $\widetilde{\mathbf{B}}' = \mathbf{UDV}'$, where $\mathbf{U}_{m-1\times p}$ and $\mathbf{V}_{p\times p}$ are orthonormal matrices and $\mathbf{D}_{p\times p}$ is a diagonal matrix containing the singular values of $\widetilde{\mathbf{B}}$. The estimates for $\boldsymbol{\Phi}$ and $\mathbf{B}$ are the first d columns of $\mathbf{U}$ and $\mathbf{VD}$, respectively (Yee and Hastie 2003).

The score for regression coefficients is

$$\mathbf{u}_i(\boldsymbol{\beta}_j) = \frac{\partial L_{ik}}{\partial \boldsymbol{\beta}_j} = \mathbf{x}_i \left(\sum_{l=1}^{m-1} \phi_{jl} p_{il} - \phi_{jk} \right)$$

the score for the scale parameters is

$$u_i(\phi_{jl}) = \frac{\partial L_{ik}}{\partial \phi_{jl}} = \begin{cases} \mathbf{x}_i\boldsymbol{\beta}_j(p_{ik}-1), & \text{if } l = k \\ \mathbf{x}_i\boldsymbol{\beta}_j p_{il}, & \text{if } l \neq k \end{cases}$$

for $l = 1, \ldots, m-1$; and the score for the intercepts is

$$u_i(\theta_l) = \frac{\partial L_{ik}}{\partial \theta_l} = \begin{cases} 1 - p_{ik}, & \text{if } l = k \\ -\,p_{il}, & \text{if } l \neq k \end{cases}$$

This command supports the Huber/White/sandwich estimator of the variance and its clustered version using `vce(robust)` and `vce(cluster` *clustvar*`)`, respectively. See [P] **_robust**, particularly *Maximum likelihood estimators* and *Methods and formulas*.

`slogit` also supports estimation with survey data. For details on VCEs with survey data, see [SVY] **variance estimation**.

References

Anderson, J. A. 1984. Regression and ordered categorical variables (with discussion). *Journal of the Royal Statistical Society, Series B* 46: 1–30.

Lunt, M. 2001. sg163: Stereotype ordinal regression. *Stata Technical Bulletin* 61: 12–18. Reprinted in *Stata Technical Bulletin Reprints*, vol. 10, pp. 298–307. College Station, TX: Stata Press.

——. 2005. Prediction of ordinal outcomes when the association between predictors and outcome differs between outcome levels. *Statistics in Medicine* 24: 1357–1369.

Tarlov, A. R., J. E. Ware, Jr., S. Greenfield, E. C. Nelson, E. Perrin, and M. Zubkoff. 1989. The medical outcomes study. An application of methods for monitoring the results of medical care. *Journal of the American Medical Association* 262: 925–930.

Wells, K. B., R. D. Hays, M. A. Burnam, W. H. Rogers, S. Greenfield, and J. E. Ware, Jr. 1989. Detection of depressive disorder for patients receiving prepaid or fee-for-service care. Results from the Medical Outcomes Survey. *Journal of the American Medical Association* 262: 3298–3302.

Yee, T. W., and T. J. Hastie. 2003. Reduced-rank vector generalized linear models. *Statistical Modelling* 3: 15–41.

Also see

[R] **slogit postestimation** — Postestimation tools for slogit

[R] **logistic** — Logistic regression, reporting odds ratios

[R] **mlogit** — Multinomial (polytomous) logistic regression

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

[R] **oprobit** — Ordered probit regression

[R] **roc** — Receiver operating characteristic (ROC) analysis

[SVY] **svy estimation** — Estimation commands for survey data

[U] **20 Estimation and postestimation commands**