Interactions and Subset Models

Dave Armstrong

09/07/2021

First, let's load the data

```
use 0705_Result.dta, clear
```

I think you might be better off using the Inverse Hyperbolic Sine transformation for the `aid_gdp` variable. So, let's look at the first model.

```
gen aid_trans = asinh(aid_gdp)
logit q179_china c.aid_trans##i.ethnic_chinese
Logistic regression                                     Number of obs =  9,043
                                                        LR chi2(3)    = 257.79
                                                        Prob > chi2   = 0.0000
Log likelihood = -6113.7169                             Pseudo R2     = 0.0206

------------------------------------------------------------------------------------------
               q179_china | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
--------------------------+---------------------------------------------------------------
                aid_trans |  -.0866912     .00639   -13.57   0.000    -.0992154    -.074167
         1.ethnic_chinese |   .6997498   .1187404     5.89   0.000     .4670228    .9324768
                          |
ethnic_chinese#c.aid_trans |
                        1 |   .1476459    .119794     1.23   0.218    -.0871461    .3824379
                          |
                    _cons |   .0250886   .0263287     0.95   0.341    -.0265147    .0766919
------------------------------------------------------------------------------------------
```

In this case, the interaction term is insignificant, though what we're really interested in is the second difference (see Carlisle Rainey's article in PSRM). You could get this with a combination of `margins` and `mlincom` from the SPost package.

```
margins, at(aid_trans  =(0 12.8) ethnic_chinese=(0 1)) post

Adjusted predictions                                     Number of obs = 9,043
Model VCE: OIM

Expression: Pr(q179_china), predict()
1._at: aid_trans      =    0
       ethnic_chinese =    0
2._at: aid_trans      =    0
       ethnic_chinese =    1
3._at: aid_trans      = 12.8
       ethnic_chinese =    0
4._at: aid_trans      = 12.8
       ethnic_chinese =    1
```

```
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .5062718   .0065811    76.93   0.000      .493373    .5191706
          2  |   .6736716   .0254539    26.47   0.000     .6237829    .7235603
          3  |   .2526443   .0132997    19.00   0.000     .2265774    .2787112
          4  |   .8183278   .2263146     3.62   0.000     .3747593    1.261896
------------------------------------------------------------------------------


mlincom (4-3) - (2-1)

             |   lincom    pvalue        ll        ul
-------------+----------------------------------------
           1 |    0.398     0.082    -0.051     0.848
```

The output of `mlincom` shows

$$\begin{aligned}\Delta\Delta = &(Pr(y=1|\text{Aid=High,Chinese}) - Pr(y=1|\text{Aid=High, Not Chinese}))\\ &-(Pr(y=1|\text{Aid=Low,Chinese}) - Pr(y=1|\text{Aid=Low, Not Chinese}))\end{aligned}$$

If the second difference is significant, then there is a significant interaction regardless of the significance of the product regressor in the model.

An alternative would be to consider the difference in first derivatives:

$$\delta\delta = \frac{\partial Pr(y=1|\text{Chinese})}{\partial \text{Aid}} - \frac{\partial Pr(y=1|\text{Not Chinese})}{\partial \text{Aid}}$$

This result is substantively similar:

```
. margins, dydx(aid_trans) at(ethnic_chinese = (0 1)) post

Average marginal effects                                 Number of obs = 9,043
Model VCE: OIM

Expression: Pr(q179_china), predict()
dy/dx wrt:  aid_trans
1._at: ethnic_chinese = 0
2._at: ethnic_chinese = 1

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
aid_trans    |
         _at |
          1  |  -.0210482   .0014892   -14.13   0.000     -.023967   -.0181293
          2  |   .0126185   .0231892     0.54   0.586    -.0328315    .0580685
------------------------------------------------------------------------------

. mlincom 2-1
```

```
             |   lincom    pvalue        ll        ul
-------------+----------------------------------------
           1 |    0.034     0.147    -0.012     0.079
```

It is difficult to derive the same test from the subset models. What Stata is doing is generating four predicted probabilities and the full variance-covariance matrix of those predicted probabilities.

$$b = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix}$$

and

$$V = \begin{bmatrix} v(p_1) & & & \\ cov(p_2, p_1) & v(p_2) & & \\ cov(p_3, p_1) & cov(p_3, p_2) & v(p_3) & \\ cov(p_4, p_1) & cov(p_4, p_2) & cov(p_4, p_3) & v(p_4) \end{bmatrix}$$

Then, the `mlincom` function is calculating $(p_4 - p_3) - (p_2 - p_1) = p_1 - p_2 - p_3 + p_4$. We could generate this difference and its standard error by making:

$$A = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$

and then calculating:

$$\Delta\Delta = Ab$$

and its standard error:

$$SE(\Delta\Delta) = \sqrt{A'VA}$$

The problem that you run into using the subset models is that you have to assume that the lower-left and upper right 2×2 sub-matrices of V are $\mathbf{0}$ because you can generate the terms to create a first difference and its standard error. Here's how it would work out.

```
quietly logit q179_china c.aid_trans##i.ethnic_chinese
quietly margins, at(aid_trans = (0 12.8) ethnic_chinese=(0 1)) post
matrix b_full = e(b)
matrix v_full = e(V)

quietly logit q179_china aid_trans if ethnic_chinese == 0
quietly margins, at(aid_trans = (0 12.8)) post
matrix b0 = e(b)
matrix v0 = e(V)

quietly logit q179_china aid_trans if ethnic_chinese == 1
quietly margins, at(aid_trans = (0 12.8)) post
matrix b1 = e(b)
matrix v1 = e(V)
```

```
matrix z = J(2,2,0)
```

Then, we could move over into Mata

```
mata
v_full = st_matrix("v_full")
b_full = st_matrix("b_full")'

v_sub = st_matrix("v0"), st_matrix("z")\st_matrix("z"), st_matrix("v1")
b_sub = st_matrix("b0")'\st_matrix("b1")'

A = (1\-1\-1\1)


DD_full = A'*b_full
se_DD_full = sqrt(A'*v_full*A)
z_DD_full = DD_full/se_DD_full

DD_sub = A'*b_sub
se_DD_sub = sqrt(A'*v_sub*A)
z_DD_sub = DD_sub/se_DD_sub
```

The results, then are:

```
: DD_full
  .3982837076

: z_DD_full
  1.736854039

: DD_sub
  .3982859389

: z_DD_sub
  1.736874446
```

The two results here are nearly identical. So, even though the coefficients appear to be different in the two subset models:

```
. logit q179_china aid_trans if ethnic_chinese == 0
...
   aid_trans |  -.0866912     .00639   -13.57   0.000    -.0992154    -.074167

. logit q179_china aid_trans if ethnic_chinese == 1
...
   aid_trans |   .0609559   .1196239     0.51   0.610    -.1735026    .2954144
```

The second difference from the two subset models suggests that the effect of `aid_trans` is not significantly different for Chinese and non-Chinese respondents.