Comparing Logit and Probit Coefficients between Models and Across Groups

Richard Williams, Notre Dame Sociology, rwilliam@ND.Edu
https://www3.nd.edu/~rwilliam/

August 2012 Annual Meetings of the American Sociological Association

I. Comparing coefficients across models

$$V(y^*) = V(\alpha + x\beta) + V(\varepsilon_{y^*}) = V(\alpha + x\beta) + \pi^2/3 = V(\alpha + x\beta) + 3.29$$

```
. quietly logit ybinary x1
. listcoef, std

logit (N=500): Unstandardized and Standardized Estimates

 Observed SD: .50035659
   Latent SD: 2.3395663

  Odds of: 1 vs 0

-------------------------------------------------------------------------------
     ybinary |      b         z     P>|z|    bStdX    bStdY   bStdXY      SDofX
-------------+-----------------------------------------------------------------
          x1 |  0.73887   10.127   0.000   1.4777   0.3158   0.6316     2.0000
-------------------------------------------------------------------------------

. quietly logit  ybinary x2
. listcoef, std

logit (N=500): Unstandardized and Standardized Estimates

 Observed SD: .50035659
   Latent SD: 2.3321875

  Odds of: 1 vs 0

-------------------------------------------------------------------------------
     ybinary |      b         z     P>|z|    bStdX    bStdY   bStdXY      SDofX
-------------+-----------------------------------------------------------------
          x2 |  0.48868   10.134   0.000   1.4660   0.2095   0.6286     3.0000
-------------------------------------------------------------------------------

. quietly logit  ybinary x1 x2
. listcoef, std

logit (N=500): Unstandardized and Standardized Estimates

 Observed SD: .50035659
   Latent SD: 5.3368197

  Odds of: 1 vs 0

-------------------------------------------------------------------------------
     ybinary |      b         z     P>|z|    bStdX    bStdY   bStdXY      SDofX
-------------+-----------------------------------------------------------------
          x1 |  1.78923    9.815   0.000   3.5785   0.3353   0.6705     2.0000
          x2 |  1.17314    9.714   0.000   3.5194   0.2198   0.6595     3.0000
-------------------------------------------------------------------------------
```

```
. corr, means

(obs=500)

    Variable |         Mean    Std. Dev.          Min          Max
-------------+----------------------------------------------------
           y |     5.51e-07     3.000001    -8.508021     7.981196
     ybinary |         .488     .5003566            0            1
          x1 |    -2.19e-08            2     -6.32646     6.401608
          x2 |     3.57e-08            3    -10.56658     9.646875


             |        y  ybinary       x1       x2
-------------+------------------------------------
           y |   1.0000
     ybinary |   0.7923   1.0000
          x1 |   0.6667   0.5248   1.0000
          x2 |   0.6667   0.5225   0.0000   1.0000
```

```
. webuse nhanes2f, clear
. khb logit diabetes black || weight

Decomposition using the KHB-Method

Model-Type:  logit                                 Number of obs     =   10335
Variables of Interest: black                       Pseudo R2         =    0.02
Z-variable(s): weight
------------------------------------------------------------------------------
    diabetes |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
black        |
     Reduced |   .6038012   .1236714     4.88   0.000     .3614098    .8461926
        Full |   .5387425   .1241889     4.34   0.000     .2953368    .7821483
        Diff |   .0650587   .0132239     4.92   0.000     .0391403    .0909771
------------------------------------------------------------------------------
```

```
. khb logit jobenjoy race || gpa ses sex educjob educimportant luckimportant sbprevent

Decomposition using the KHB-Method

Model-Type:  logit                                 Number of obs     =    6731
Variables of Interest: race                        Pseudo R2         =    0.08
Z-variable(s): gpa ses sex educjob educimportant luckimportant sbprevent
```

jobenjoy	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
race						
Reduced	-.5727334	.10607	-5.40	0.000	-.7806269	-.3648399
Full	-.4833004	.1095584	-4.41	0.000	-.6980309	-.26857
Diff	-.089433	.0349898	-2.56	0.011	-.1580117	-.0208542

II. Comparing coefficients across groups

Case 1: True coefficients are equal, residual variances differ

	Group 0	Group 1
True coefficients	$y_i^* = x_{i1} + x_{i2} + x_{i3} + \varepsilon_i$	$y_i^* = x_{i1} + x_{i2} + x_{i3} + 2\varepsilon_i$
Standardized Coefficients	$y_i^* = x_{i1} + x_{i2} + x_{i3} + \varepsilon_i$	$y_i^* = .5x_{i1} + .5x_{i2} + .5x_{i3} + \varepsilon_i$

Case 2: True coefficients differ, residual variances differ

	Group 0	Group 1
True coefficients	$y_i^* = x_{i1} + x_{i2} + x_{i3} + \varepsilon_i$	$y_i^* = 2x_{i1} + 2x_{i2} + 2x_{i3} + 2\varepsilon_i$
Standardized Coefficients	$y_i^* = x_{i1} + x_{i2} + x_{i3} + \varepsilon_i$	$y_i^* = x_{i1} + x_{i2} + x_{i3} + \varepsilon_i$

Case 3: True coefficients differ, residual variances differ even more

	Group 0	Group 1
True coefficients	$y_i^* = x_{i1} + x_{i2} + x_{i3} + \varepsilon_i$	$y_i^* = 2x_{i1} + 2x_{i2} + 2x_{i3} + 3\varepsilon_i$
Standardized Coefficients	$y_i^* = x_{i1} + x_{i2} + x_{i3} + \varepsilon_i$	$y_i^* = \frac{2}{3}x_{i1} + \frac{2}{3}x_{i2} + \frac{2}{3}x_{i3} + \varepsilon_i$

Allison's example: Apparent differences in effects across groups may be an artifact of differences in residual variability

Table 1: Results of Logit Regressions Predicting Promotion to Associate Professor for Male and Female Biochemists (Adapted from Allison 1999, p. 188)

	Men		*Women*		*Ratio of*	*Chi-Square*
Variable	*Coefficient*	*SE*	*Coefficient*	*SE*	*Coefficients*	*for Difference*
Intercept	-7.6802***	.6814	-5.8420***	.8659	.76	2.78
Duration	1.9089***	.2141	1.4078***	.2573	.74	2.24
Duration squared	-0.1432***	.0186	-0.0956***	.0219	.67	2.74
Undergraduate selectivity	0.2158***	.0614	0.0551	.0717	.25	2.90
Number of articles	*0.0737****	*.0116*	*0.0340***	*.0126*	*.46*	*5.37**
Job prestige	-0.4312***	.1088	-0.3708*	.1560	.86	0.10
Log likelihood	-526.54		-306.19			
Error variance	3.29		3.29			

**p < .05, **p < .01, *** p < .001*

Allison's solution: Add delta to adjust for differences in residual variability

Table 2: Logit Regressions Predicting Promotion to Associate Professor for Male and Female Biochemists, Disturbance Variances Unconstrained (Adapted from Allison 1999, p. 195)

	All Coefficients Equal		*Articles Coefficient Unconstrained*	
Variable	*Coefficient*	*SE*	*Coefficient*	*SE*
Intercept	-7.4913***	.6845	-7.3655***	.6818
Female	-0.93918**	.3624	-0.37819	.4833
Duration	1.9097***	.2147	1.8384***	.2143
Duration squared	-0.13970***	.0173	-0.13429***	.01749
Undergraduate selectivity	0.18195**	.0615	0.16997***	.04959
Number of articles	0.06354***	.0117	0.07199***	.01079
Job prestige	-0.4460***	.1098	-0.42046***	.09007
δ	*-0.26084**	*.1116*	*-0.16262*	*.1505*
Articles x Female			*-0.03064*	*.0173*
Log likelihood	-836.28		-835.13	

**p < .05, **p < .01, *** p < .001*

Alternative (and broader) solution: Heterogeneous Choice Models

With heterogeneous choice (aka Location-Scale) models, the dependent variable can be ordinal or binary. For a binary dependent variable, the model (Keele & Park, 2006) can be written as

$$\Pr(y_i = 1) = g\left(\frac{x_i\beta}{\exp(z_i\gamma)}\right) = g\left(\frac{x_i\beta}{\exp(\ln(\sigma_i))}\right) = g\left(\frac{x_i\beta}{\sigma_i}\right)$$

In the above formula,

- g stands for the link function (in this case logit; probit is also commonly used, and other options are possible, such as the complementary log-log, log-log and cauchit).
- x is a vector of values for the ith observation. The x's are the explanatory variables and are said to be the determinants of the choice, or outcome.
- z is a vector of values for the ith observation. The z's define groups with different error variances in the underlying latent variable. The z's and x's need not include any of the same variables, although they can.
- β and γ are vectors of coefficients. They show how the x's affect the choice and the z's affect the variance (or more specifically, the log of σ).
- The numerator in the above formula is referred to as the choice equation, while the denominator is the variance equation. These are also referred to as the location and scale equations. Also, the choice equation includes a constant term but the variance equation does not.
- The conventional logit and probit models, which do not have variance equations, are special cases of the above, where σ_i = 1 for all cases.
- Allison's model is a special case of a heterogeneous choice model, where the dependent variable is a dichotomy and both the variance and choice equations include the same dichotomous grouping variable.

In Stata, heterogeneous choice models can be estimated via the user-written routine `oglm`.

```
. * oglm replication of Allison's Table 2, Model 2 with interaction added:
. use "http://www.indiana.edu/~jslsoc/stata/spex_data/tenure01.dta", clear
(Gender differences in receipt of tenure (Scott Long 06Jul2006))
. keep if pdasample
(148 observations deleted)
. oglm tenure female year yearsq select articles prestige f_articles, het(female)

Heteroskedastic Ordered Logistic Regression       Number of obs   =       2797
                                                  LR chi2(8)      =     415.39
                                                  Prob > chi2     =     0.0000
Log likelihood = -835.13347                       Pseudo R2       =     0.1992

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
tenure       |
      female |  -.3780597   .4500207    -0.84   0.401    -1.260084    .5039646
        year |   1.838257   .2029491     9.06   0.000     1.440484     2.23603
      yearsq |  -.1342828    .017024    -7.89   0.000    -.1676492   -.1009165
      select |   .1699659   .0516643     3.29   0.001     .0687057    .2712261
    articles |   .0719821   .0114106     6.31   0.000     .0496178    .0943464
    prestige |  -.4204742   .0961206    -4.37   0.000    -.6088671   -.2320813
  f_articles |  -.0304836   .0187427    -1.63   0.104    -.0672185    .0062514
-------------+----------------------------------------------------------------
lnsigma      |
      female |   .1774193   .1627087     1.09   0.276     -.141484    .4963226
-------------+----------------------------------------------------------------
       /cut1 |   7.365285   .6547121    11.25   0.000     6.082073    8.648497
------------------------------------------------------------------------------
. display "Allison's delta = " (1 - exp(.1774193)) / exp(.1774193)
-.16257142
```

```
. * Hauser & Andrew's original LRPC program
. * Code has been made more efficient and readable,
. * but results are the same.  Note that it
. * actually estimates and reports
. * lambda - 1 rather than lamba.
. program define lrpc02
  1.         tempvar theta
  2.         version 8
  3.         args lnf  intercepts lambdaminus1 betas
  4.         gen double `theta' = `intercepts' + `betas' + (`lambdaminus1' * `betas')
  5.         quietly replace `lnf' = ln(exp(`theta')/(1+exp(`theta'))) if $ML_y1==1
  6.         quietly replace `lnf' = ln(1/(1+exp(`theta'))) if $ML_y1==0
  7. end
. * Hauser & Andrews original LRPC parameterization used with Allison's data
. * Results are identical to Allison's Table 2, Model 1
. ml model lf lrpc02 ///
>         (intercepts: tenure = male female, nocons) ///
>         (lambdaminus1: female, nocons) ///
>         (betas: year yearsq select articles prestige, nocons), max nolog
. ml display

                                                  Number of obs   =       2797
                                                  Wald chi2(2)    =     180.60
Log likelihood = -836.28235                       Prob > chi2     =     0.0000

------------------------------------------------------------------------------
      tenure |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
intercepts   |
        male |  -7.490506   .6596634   -11.36   0.000    -8.783422   -6.197589
      female |  -6.230958   .6205863   -10.04   0.000    -7.447285   -5.014631
-------------+----------------------------------------------------------------
lambdaminus1 |
      female |  -.2608325   .1080502    -2.41   0.016    -.4726069   -.0490581
-------------+----------------------------------------------------------------
betas        |
        year |   1.909544   .1996937     9.56   0.000     1.518151    2.300936
      yearsq |  -.1396868   .0169425    -8.24   0.000    -.1728935   -.1064801
      select |   .1819201   .0526572     3.45   0.001     .0787139    .2851264
    articles |   .0635345    .010219     6.22   0.000     .0435055    .0835635
    prestige |  -.4462074    .096904    -4.60   0.000    -.6361357   -.2562791
```