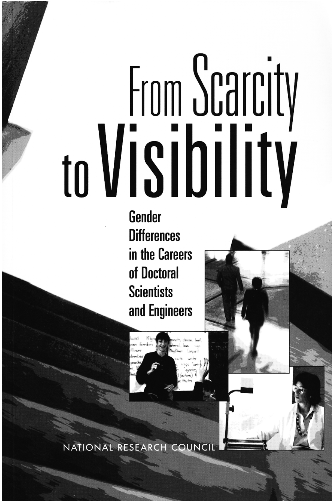
Comparing group effects in nonlinear models: Statistical problems and substantive insights

Scott Long, Indiana University

April 18, 2007

The problem

Women in science and statistical "fine points"



The statistical and substantive problem

Traditional methods don't work

A common solution for comparing groups is to:

1. Estimate a model for men
2. Estimate a model for women
3. Compare the coefficients

This approach with binary outcomes has two limitations:

1. The NRC committee thought that the coefficients did not have a clear link to the substance of the problems.
2. Allison sent me a working paper that said: "Differences in the estimated coefficients tell us nothing about the differences in the underlying impact of [publications] on [tenure for] the two groups."

Objectives

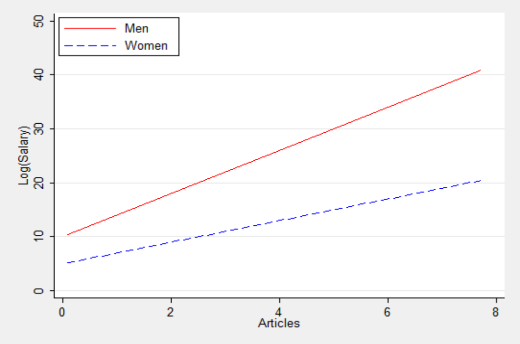
Comparing groups in logit and probit

1. Comparing β's across groups in the LRM.
2. Review of the logit and probit models for binary outcomes.
3. Focus on interpretation using predicted probabilities.
4. Discuss methods for comparing groups with binary logit and probit.
5. Consider problems with the usual test comparing β's across groups.
6. Show how groups can be compared using predicted probabilities.

Group comparisons in the LRM - I

A Chow test comparing structural coefficients

$$
\begin{array}{ll}
\text{Men:} & y = \alpha^m + \beta^m_{articles} articles + \beta^m_{prestige} prestige + \varepsilon \\
\text{Women:} & y = \alpha^w + \beta^w_{articles} articles + \beta^w_{prestige} prestige + \varepsilon
\end{array}
$$



Group comparisons in the LRM - II

A Chow test comparing structural coefficients

1. Do men and women have the same return for education?

$$H_0\colon\ \beta^m_{articles} = \beta^w_{articles}$$

2. We compute a Chow test:

$$z = \frac{\widehat{\beta}^m_{articles} - \widehat{\beta}^w_{articles}}{\sqrt{Var\left(\widehat{\beta}^m_{articles}\right) + Var\left(\widehat{\beta}^w_{articles}\right)}}$$

3. Or, we might test:

$$H_0\colon\ \alpha^m = \alpha^w;\ \beta^m_{articles} = \beta^w_{articles}; \beta^m_{prestige} = \beta^w_{prestige}$$

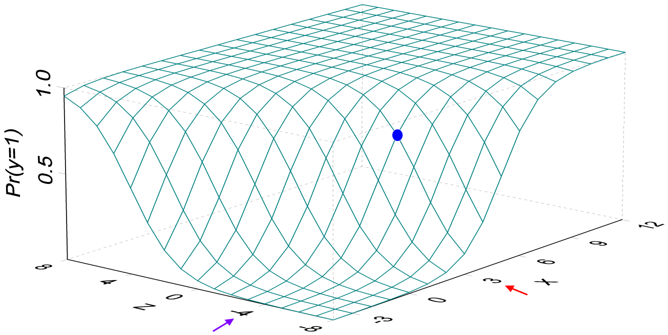
This does not imply that $R^2_m = R^2_w$.

Overview

Group comparisons in logit and probit

1. While testing H_0: $\beta^m_{articles} = \beta^w_{articles}$ is appropriate in LRM, it is not in the BRM.
2. Here, the test confounds:
   2.1 Group differences in the effect of x.
   2.2 Group differences in unobserved heterogeneity.
3. Allison's test assumes that the effects of other variables are equal across groups.
4. Alternatively, I propose tests based on comparing predicted probabilities.

$$\Pr\left(y=1 \mid \mathbf{x}\right)=F\left(\beta_0+\beta_x x+\beta_z z\right)$$



Logit

$$\begin{aligned}\Pr\left(y=1\mid \mathbf{x}\right) &= \Lambda\left(\beta_0+\beta_x x+\beta_z z\right)\\ &= \frac{\exp\left(\beta_0+\beta_x x+\beta_z z\right)}{1+\exp\left(\beta_0+\beta_x x+\beta_z z\right)}\end{aligned}$$

Probit

$$\begin{aligned}\Pr\left(y=1\mid \mathbf{x}\right) &= \Phi\left(\beta_0+\beta_x x+\beta_z z\right)\\ &= \int_{-\infty}^{\beta_0+\beta_x x+\beta_z z}\frac{1}{\sqrt{2\pi}}\left(\frac{-t^2}{2}\right)dt\end{aligned}$$

Example: gender differences in tenure

Descriptive statistics for data as career years

Variable	Mean	StdDev	Minimum	Maximum	Label
tenure	0.12	0.33	0.00	1.00	Is tenured?
female	0.38	0.49	0.00	1.00	Scientist is female?
year	4.33	3.09	1.00	22.00	Years in rank.
yearsq	28.29	44.18	1.00	484.00	Years in rank squared.
select	4.97	1.43	1.00	7.00	Selectivity of bachelor's.
articles	7.21	6.74	0.00	73.00	Total number of articles.
prestige	2.63	0.77	0.65	4.80	Prestige of department.
presthi	0.05	0.21	0.00	1.00	Prestige is 4 or higher?

N = 2945

Model 1: logit estimates

Using only a dummy variable for group membership

$$\Pr(\text{tenure} = 1 \mid \mathbf{x}) = \Lambda \begin{pmatrix} \beta_0 + \beta_1 \text{female} + \beta_2 \text{year} + \beta_3 \text{yearsq} \\ +\beta_4 \text{select} + \beta_5 \text{articles} + \beta_6 \text{presthi} \end{pmatrix}$$

```
logit (N=2945): Factor Change in Odds

  Odds of: Tenure vs NoTenure
```

	b	z	P>|z|	e^b	e^bStdX	SDofX
female	-0.37354	-2.941	0.003	0.6883	0.8341	0.4856
year	0.93246	11.003	0.000	2.5407	17.8420	3.0903
yearsq	-0.05380	-8.936	0.000	0.9476	0.0928	44.1814
select	0.12314	2.876	0.004	1.1310	1.1931	1.4339
articles	0.05091	6.562	0.000	1.0522	1.4097	6.7449
presthi	-0.94447	-2.555	0.011	0.3889	0.8219	0.2077
constant	-5.77055	-16.379	0.000			

Odds ratios for interpretation

Model 1: a dummy variable for group membership

Odds:

$$\text{Odds}\left(x, z\right) = \frac{\Pr\left(y = 1 \mid x, z\right)}{\Pr\left(y = 0 \mid x, z\right)}$$

Odds ratios:

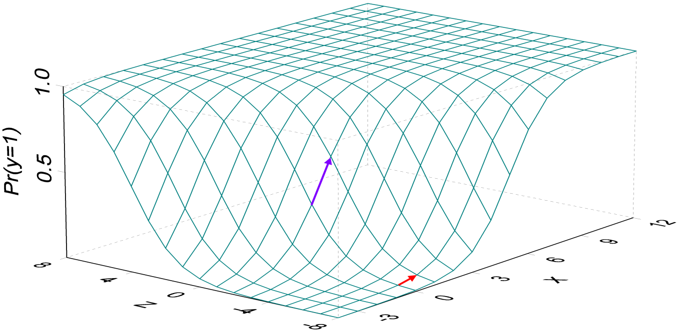
$$\frac{\text{Odds}\left(x + 1, z\right)}{\text{Odds}\left(x, z\right)} = \exp\left(\beta_x\right)$$

1. For articles: $\exp\left(\beta_{\text{articles}}\right) = 1.05$.
   - For each additional article, the odds of tenure increase by a factor of 1.05, holding all other variables constant.
2. For female: $\exp\left(\beta_{\text{female}}\right) = 0.69$.
   - Being a female scientist decreases the odds of tenure by a factor of .69, holding all other variables constant.

Odds ratios compared to changes in probabilities

Binary logit

Both arrows are the same factor change in the odds.



Predicted probabilities at a given x

Model 1: a dummy variable for group membership

1. The predicted probability at specific values of the independent variables:

$$\Pr(\text{tenure} = 1 \mid \mathbf{x}) = \Lambda \left(\begin{array}{c} \beta_0 + \beta_1 \text{female} + \beta_2 \text{year} + \beta_3 \text{yearsq} \\ + \beta_4 \text{select} + \beta_5 \text{articles} + \beta_6 \text{presthi} \end{array} \right)$$

2. For example, the probability of tenure for non-publishing women in year 7, with selectivity 4 and low prestige:

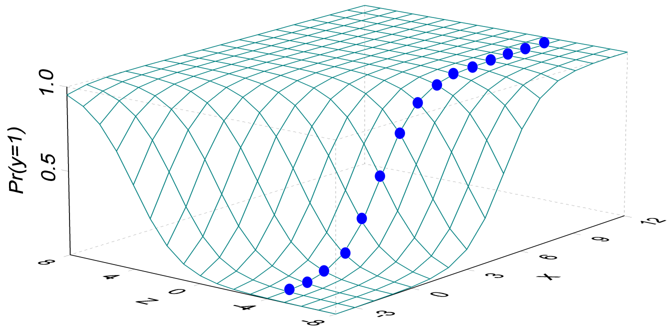
$$0.16 = \Lambda \left(\begin{array}{c} \beta_0 + \beta_1 (1) + \beta_2 (7) + \beta_3 (49) \\ + \beta_4 (4) + \beta_5 (0) + \beta_6 (.05) \end{array} \right)$$

3. Extending this idea, plots of probabilities can be constructed...

Plotting predicted probabilities as one variable changes

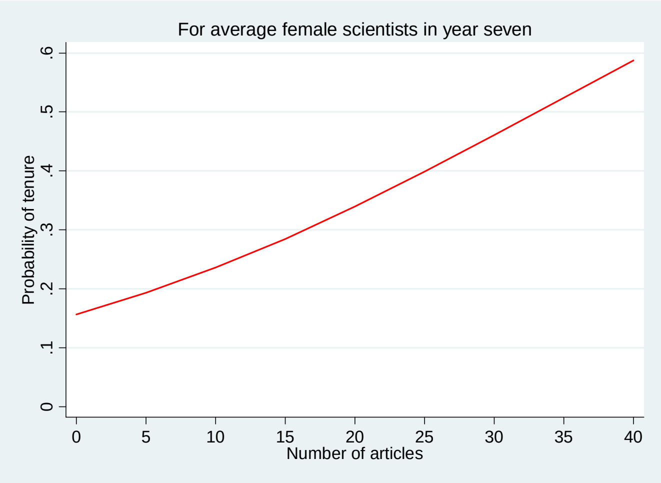
For a simple model with two predictors

$$\Pr\left(y = 1 \mid \mathbf{x}\right) = F\left(\beta_0 + \beta_x x + \beta_z z\right)$$



Plotting predicted probabilities as one variable changes

Model 1: a dummy variable for group membership



Confidence intervals for predicted probabilities

Binary logit and probit

1. **Confidence intervals** for predictions:

$$\left[\Pr\left(y = 1 \mid \mathbf{x}\right)_{\text{LowerBound}},\ \Pr\left(y = 1 \mid \mathbf{x}\right)_{\text{UpperBound}} \right]$$

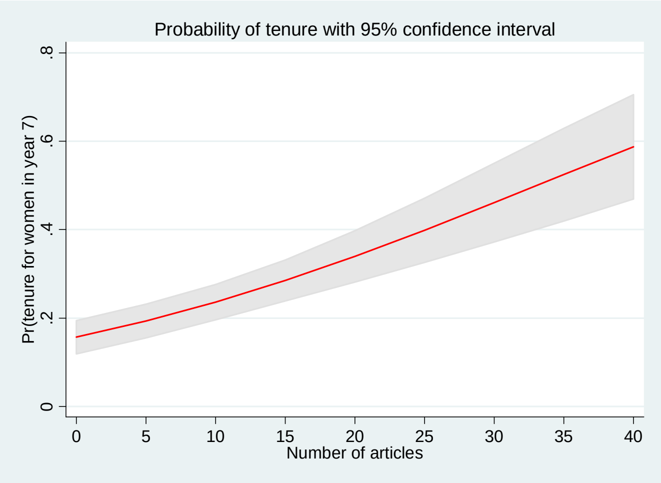
2. **Delta method** is easy is fast:

$$Var\left[\widehat{\Pr}\left(y = 1 \mid \mathbf{x}\right)\right] = \left[\frac{\partial F\left(\mathbf{x}\widehat{\boldsymbol{\beta}}\right)}{\partial \widehat{\boldsymbol{\beta}}}\right]^{T} Var(\widehat{\boldsymbol{\beta}}) \left[\frac{\partial F\left(\mathbf{x}\widehat{\boldsymbol{\beta}}\right)}{\partial \widehat{\boldsymbol{\beta}}}\right]$$

3. **Bootstrap method** requires at least 1,000 replications to get reliable results.

Confidence intervals for predicted probabilities

Model 1: a dummy variable for group membership



Approaches for making group comparisons:

1. Include a dummy variable for group.

$$\beta_{\text{female}} \text{ in the prior model.}$$

2. Allow the effects of the x's to differ by group.

$$\text{Let } \beta^{\text{m}}_{\text{articles}} \text{ and } \beta^{\text{w}}_{\text{articles}} \text{ differ.}$$

   2.1 Test the equality of coefficients.

$$\beta^{\text{m}}_{\text{articles}} = \beta^{\text{w}}_{\text{articles}}$$

   2.2 Compare predictions by across groups.

$$\Pr\left(y = 1 \mid \mathbf{x}\right)_m = \Pr\left(y = 1 \mid \mathbf{x}\right)_w$$

1. In the BRM:

$$\begin{aligned} \text{Men:} \quad & \Pr(y=1) = \Lambda\left(\alpha^{m} + \beta^{m}_{articles} articles + \beta^{m}_{prestige} prestige\right) \\ \text{Women:} \quad & \Pr(y=1) = \Lambda\left(\alpha^{w} + \beta^{w}_{articles} articles + \beta^{w}_{prestige} prestige\right) \end{aligned}$$

2. Can we use a Chow-type test?

$$H_0\colon\ \beta^{m}_{articles} = \beta^{w}_{articles}$$

3. Due to an identification problem Allison (1999) argues that this test tell us nothing about the underlying impact of x for the two groups.

Regression on a latent y* - I

Binary logit and probit

1. **Structural model** with a latent y^*:

$$y^* = \alpha + \beta x + \varepsilon$$

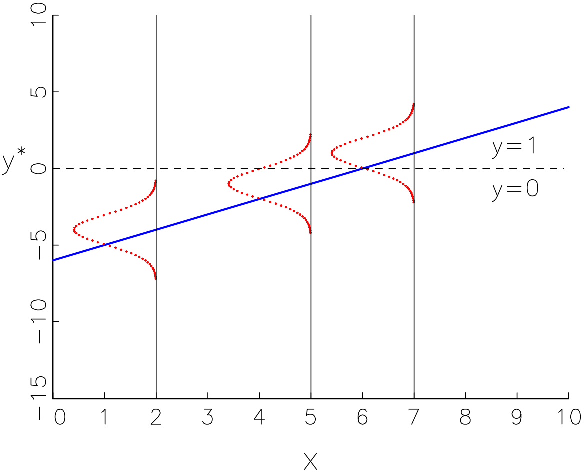
2. **Error** ε is normal(0,1) for probit; ε is logistic(0,$\pi^2/3$) for logit.
3. **Observed y and latent y*** are linked by:

$$y = \begin{cases} 1 & \text{if } y^* > 0 \\ 0 & \text{if } y^* \leq 0 \end{cases}$$

4. **Graphically**,

Regression on a latent y* - II

Binary logit and probit



Regression on a latent y* - III

Binary logit and probit

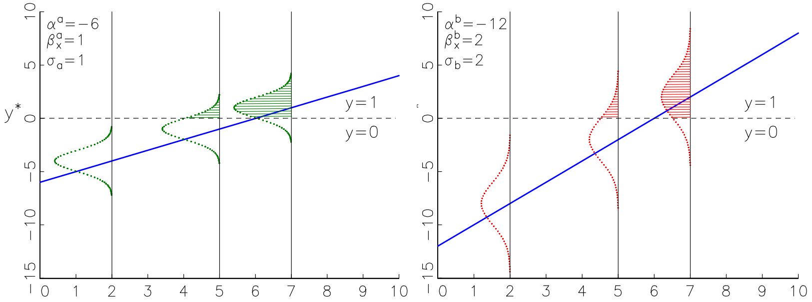
5. **Pr(y=1)** depends on the error distribution *and* the coefficients:

$$\begin{aligned} \Pr\left(y = 1 \mid x\right) &= \Pr\left(y^* > 0 \mid x\right) \\ &= \Pr\left(\varepsilon < [\alpha + \beta x] \mid x\right) \end{aligned}$$

6. There is an **identification problem** that can be illustrated graphically.

Regression on a latent y* - IV

Binary logit and probit



7. In terms of $\Pr(y = 1)$, these are empirically indistinguishable:

   Green A change in x of 1 when $\beta_x^a = 1$ and $\sigma_a = 1$.

   Red A change in x of 1 when $\beta_x^b = 2$ and $\sigma_b = 2$.

Identification and group comparisons - I

Identification of betas, error variance and probabilities

1. Let y^* be the latent variable associated with receipt of tenure:

$$\begin{array}{ll} \text{Men:} & y^* = \alpha^m + \beta^m_{\text{articles}}\text{articles} + \varepsilon_m \\ \text{Women:} & y^* = \alpha^w + \beta^w_{\text{articles}}\text{articles} + \varepsilon_w \end{array}$$

2. Assume the coefficients for articles are equal:

$$\beta^m_{\text{articles}} = \beta^w_{\text{articles}}$$

3. But, assume women have more unobserved heterogeneity:

$$\sigma_w > \sigma_m$$

4. Now estimate the model...

5. Software makes implicit assumptions:

$$\text{Logit: } \quad Var(\varepsilon) = \frac{\pi^2}{3}$$

$$\text{Probit: } \quad Var(\varepsilon) = 1$$

6. What is the effect of these assumptions?

7. With probit, ε is rescaled so that:

$$Var\left(\frac{\varepsilon}{\sigma}\right) = Var\left(\widetilde{\varepsilon}\right) = 1$$

8. For men, the estimated model for probit is:

$$\begin{aligned}\frac{y^*}{\sigma_m} &= \frac{\alpha^m}{\sigma_m} + \frac{\beta^m_{\text{articles}}}{\sigma_m}\text{articles} + \frac{\varepsilon_m}{\sigma_m} \\ &= \widetilde{\alpha}^m + \widetilde{\beta}^m_{\text{articles}}\text{articles} + \widetilde{\varepsilon}_m, \text{ where } \widetilde{\sigma}_m = 1\end{aligned}$$

9. For women, the estimated model for probit is:

$$\begin{aligned}\frac{y^*}{\sigma_w} &= \frac{\alpha^w}{\sigma_w} + \frac{\beta^w_{\text{articles}}}{\sigma_w}\text{articles} + \frac{\varepsilon_w}{\sigma_w} \\ &= \widetilde{\alpha}^w + \widetilde{\beta}^w_{\text{articles}}\text{articles} + \widetilde{\varepsilon}_w, \text{ where } \widetilde{\sigma}_w = 1\end{aligned}$$

Identification and group comparisons - IV

Identification of betas, error variance and probabilities

10. We want to test:

$$H_0\colon\ \beta^m_{\text{articles}} = \beta^w_{\text{articles}}$$

11. But, we test:

$$H_0\colon\ \widetilde{\beta}^m_{\text{articles}} = \widetilde{\beta}^w_{\text{articles}}$$

12. Unless $\sigma^2_m = \sigma^2_w$,

$$\widetilde{\beta}^m_{\text{articles}} = \widetilde{\beta}^w_{\text{articles}}$$

does not imply

$$\beta^m_{\text{articles}} = \beta^w_{\text{articles}}$$

Aside: rescaling errors in logit

1. The model is:

$$y^* = \alpha + \beta_{\text{articles}}\text{articles} + \varepsilon$$

2. We rescale the errors so that:

$$Var\left(\widetilde{\varepsilon}\right) = \frac{\pi^2}{3} \text{ rather than 1 for probit}$$

3. This leads to the equation that is estimated:

$$\frac{\pi}{\sqrt{3}}\frac{y^*}{\sigma} = \frac{\pi}{\sqrt{3}}\frac{\alpha}{\sigma} + \frac{\pi}{\sqrt{3}}\frac{\beta_{\text{articles}}}{\sigma}\text{articles} + \frac{\pi}{\sqrt{3}}\frac{\varepsilon}{\sigma}$$

Alternatives for testing group differences - I

Binary logit and probit

Two distinct approaches address the identification problem.

1. Allison's test of H_0: $\beta_x^m = \beta_x^w$, disentangles the β's and $Var(\varepsilon)$.
   1.1 The test requires the strong assumption:

   $$\beta_z^m = \beta_z^w \quad \text{or equivalently} \quad \frac{\beta_z^m}{\beta_z^w} = 1$$

   1.2 Then, the ratio of estimated $\widetilde{\beta}$ gives us information on the unobserved variances:

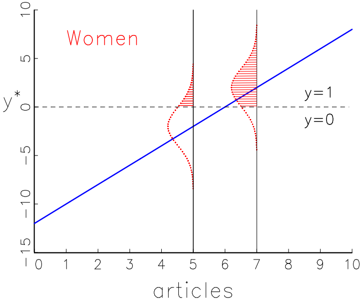
   $$\frac{\widetilde{\beta}_z^m}{\widetilde{\beta}_z^w} = \frac{\beta_z^m / \sigma_m}{\beta_z^w / \sigma_w} = \frac{\sigma_w}{\sigma_m}$$

   1.3 This provides leverage to test:

   $$H_0 : \beta_x^m = \beta_x^w$$

2. Alternatively, since the probabilities are invariant to $Var(\varepsilon)$, I propose testing

$$H_0\colon\ \Pr\left(y=1 \mid \mathbf{x}\right)_m = \Pr\left(y=1 \mid \mathbf{x}\right)_w$$



1. Let $w = 1$ for women, else 0 and $wx = w \times x$;
   let $m = 1$ for men, else 0 and $mx = m \times x$.

$$\Pr(y = 1) = F\left(\alpha^{w} w + \beta_{x}^{w} wx + \alpha^{m} m + \beta_{x}^{m} mx\right)$$

2. Then:

$$\begin{aligned} \Pr(y = 1 \mid \mathbf{x})_{w} &= F\left(\alpha^{w} + \beta_{x}^{w} x\right) \quad \text{if } w = 1, m = 0 \\ \Pr(y = 1 \mid \mathbf{x})_{m} &= F\left(\alpha^{m} + \beta_{x}^{m} x\right) \quad \text{if } w = 0, m = 1 \end{aligned}$$

3. The gender difference in the probability of tenure is:

$$\Delta_{m-w}(\mathbf{x}) = \Pr(y = 1 \mid \mathbf{x})_{m} - \Pr(y = 1 \mid \mathbf{x})_{w}$$

Model 2: Logit estimates and Chow-type test

Articles only by group

Start with a simple model with only publications predicting tenure:

```
logit (N=2945): Factor Change in Odds

  Odds of: Tenure vs NoTenure

MEN                  b         z     P>|z|     e^b    e^bStdX      SDofX
    constant  -2.69315   -23.000   0.000    0.0677    0.2704     0.4856
    articles   0.09810     9.928   0.000    1.1031    1.7854     5.9089

WOMEN                b         z     P>|z|     e^b    e^bStdX      SDofX
    constant  -2.47327   -18.299   0.000    0.0843    0.3009     0.4856
    articles   0.04215     4.259   0.000    1.0430    1.2855     5.9592
```

Comparing groups with predicted probabilities and CIs - I

Model 2: articles only by group

To compare groups at different levels of articles:

1. Compute differences:

$$\begin{aligned}\Delta_{m-w}\left(\text{articles}\right) &= \Pr\left(y=1 \mid \text{articles}\right)_m \\ &\quad - \Pr\left(y=1 \mid \text{articles}\right)_w\end{aligned}$$

2. Compute confidence intervals by delta or bootstrap:

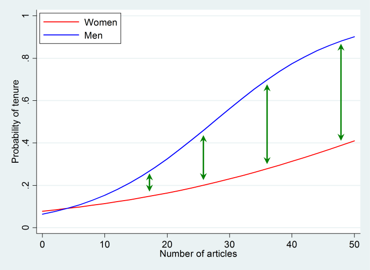
$$\left[\Delta_{m-w}\left(\text{articles}\right)_{\text{LowerBound}}, \Delta_{m-w}\left(\text{articles}\right)_{\text{UpperBound}}\right]$$

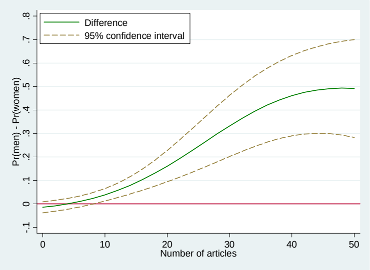
3. With one RHS variable, we can plot all comparisons.

Comparing groups with predicted probabilities and CIs - II

Model 2: articles only by group

4. Moving from predictions for each group to differences in predictions:





Logit and probit with additional independent variables - I

Effects of additional variables for predicted probabilities

Adding variables introduces substantial complications for interpretation:

1. With two independent variables:

$$\Pr(y = 1 \mid x, z) = F(\alpha + \beta_x x + \beta_z z)$$

2. Setting $z = Z^*$ changes the intercept in an equation with only x:

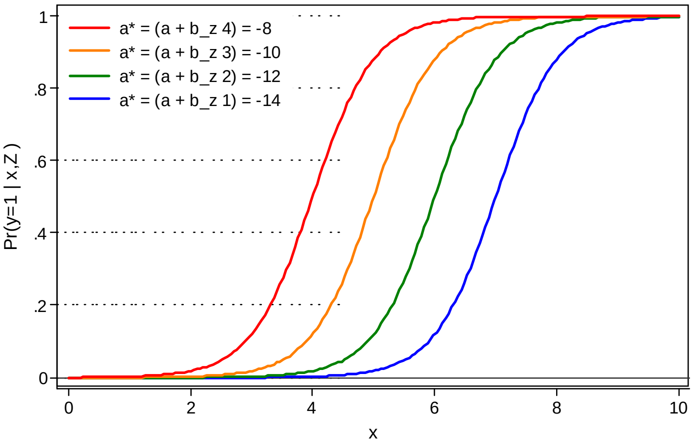
$$\begin{aligned} \Pr(y = 1 \mid x, Z^*) &= F(\alpha + \beta_x x + \beta_z Z^*) \\ &= F([\alpha + \beta_z Z^*] + \beta_x x) \\ &= F(\alpha^* + \beta_x x) \end{aligned}$$

3. Predictions depend on the levels of each variable in the model.

Logit and probit with additional independent variables - II

Effects of additional variables for predicted probabilities

Graphically, predictions at a single x depend on z:



Comparing groups with additional independent variables

Control variables change the intercept

1. For a given $z = Z$:

$$\begin{aligned} \text{Men:} \quad & \Pr\left(y = 1 \mid x, Z^*\right)_m = F\left(\alpha^{*m} + \beta_x^m x\right) \\ \text{Women:} \quad & \Pr\left(y = 1 \mid x, Z^*\right)_w = F\left(\alpha^{*w} + \beta_x^w x\right) \end{aligned}$$

2. Differences in probabilities for a given x depends on the level of other variables:

$$\Delta_{m-w}\left(x, Z^*\right) = \Pr\left(y = 1 \mid x, Z^*\right)_m - \Pr\left(y = 1 \mid x, Z^*\right)_w$$

Model 3: logit estimates

Articles and prestige by group

Using a dummy variable indicating high prestige jobs:

```
logit (N=2945): Factor Change in Odds

  Odds of: Tenure vs NoTenure

MEN                   b         z     P>|z|     e^b    e^bStdX      SDofX
    constant   -2.68452   -22.871    0.000   0.0683    0.2715     0.4856
    articles    0.10011    10.014    0.000   1.1053    1.8068     5.9089
     presthi   -0.72952    -1.713    0.087   0.4821    0.8934     0.1545

WOMEN                b         z     P>|z|     e^b    e^bStdX      SDofX
    constant   -2.54377   -17.900    0.000   0.0786    0.2907     0.4856
    articles    0.05724     4.995    0.000   1.0589    1.4065     5.9592
     presthi   -1.63483    -2.433    0.015   0.1950    0.7923     0.1424
```

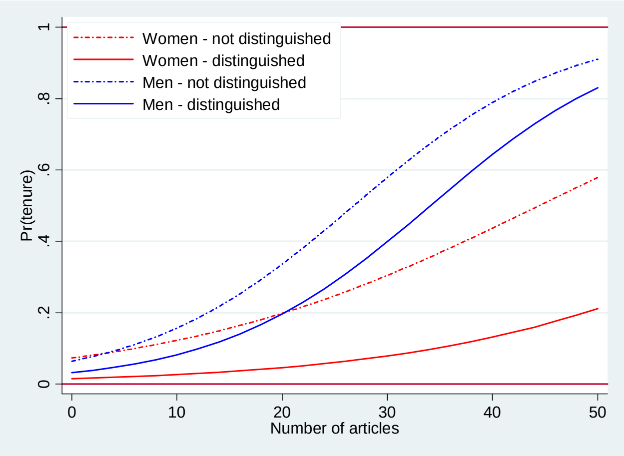
Chow tests: $X^2_{\text{articles}}(1) = 1.1,\ p = .30;\ X^2_{\text{presthi}}(1) = .03,\ p = .86.$

Allison tests: $X^2_{\text{articles}}(1) = 1.7,\ p = .19;\ X^2_{\text{presthi}}(1) = 1.6,\ p = .21.$

Predicted probabilities by group and prestige level - I

Model 3: articles and prestige by group

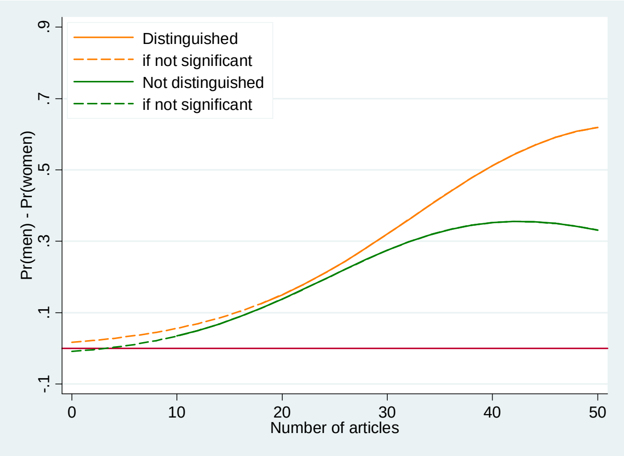
This shows all predictions from this model.



Predicted probabilities by group and prestige level - II

Model 3: articles and prestige by group

Alternatively, we can plot $\Delta_{m-w}(\text{articles}, \text{presthi})$:



Model 4: logit estimates

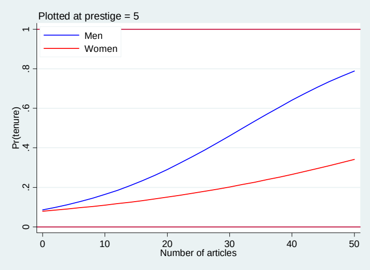
All variables by group

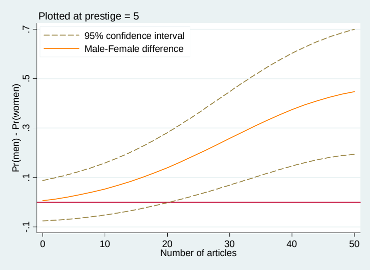
```
logit (N=2945): Factor Change in Odds

  Odds of: Tenure vs NoTenure

MEN                     b         z     P>|z|      e^b    e^bStdX      SDofX
    constant     -5.82375   -11.551    0.000    0.0030     0.0591     0.4856
        year      1.07188     9.084    0.000    2.9209    27.7967     3.1019
      yearsq     -0.06540    -7.513    0.000    0.9367     0.1043    34.5630
      select      0.21072     3.687    0.000    1.2346     1.7448     2.6416
    articles      0.07355     6.836    0.000    1.0763     1.5444     5.9089
    prestige     -0.37700    -3.645    0.000    0.6859     0.5865     1.4151

WOMEN                   b         z     P>|z|      e^b    e^bStdX      SDofX
    constant     -4.20721    -6.675    0.000    0.0149     0.1296     0.4856
        year      0.76851     6.123    0.000    2.1565     9.9209     2.9858
      yearsq     -0.04176    -4.930    0.000    0.9591     0.2430    33.8754
      select      0.03444     0.504    0.614    1.0350     1.0936     2.5975
    articles      0.03570     2.982    0.003    1.0363     1.2371     5.9592
    prestige     -0.34818    -2.288    0.022    0.7060     0.6212     1.3673
```

Chow tests: $X^2_{articles}(1){=}1.1$, p=.30. $X^2_{prestige}(1){=}.03$, p=.86.

Allison tests: $X^2_{articles}(1){=}1.7$, p=.19. $X^2_{prestige}(1){=}1.6$, p=.21

Moving from by group predicted probabilities to group differences:

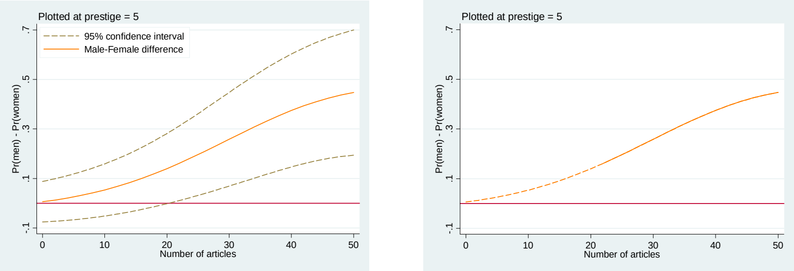




Plotting probabilities and group differences II

Model 4: all variables by group

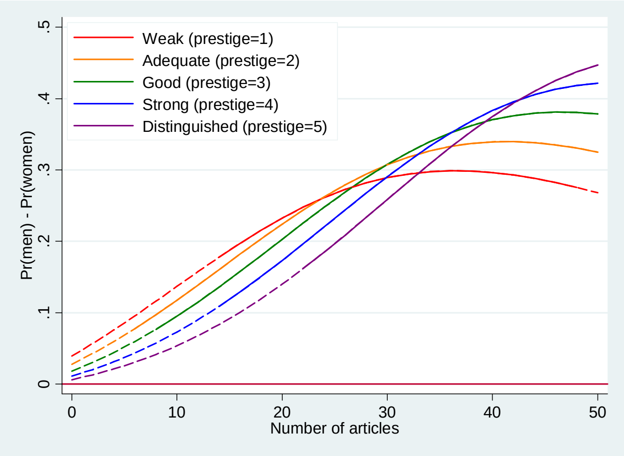
A dashed line indicates the differences is not significant:



Discrete change as prestige and articles varies - I

Model 4: all variables by group

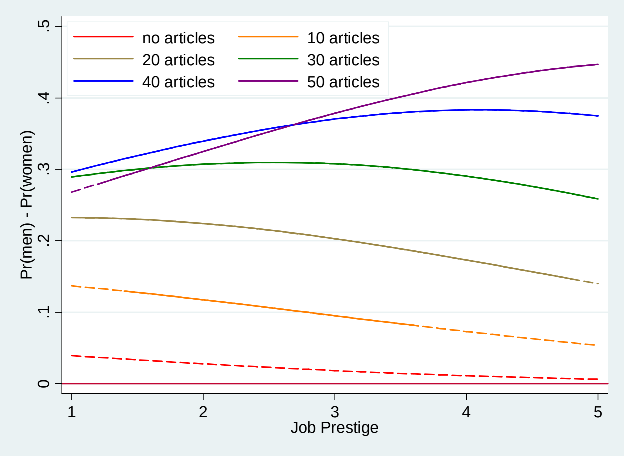
Holding all other variables constant:



Discrete change as prestige and articles varies - II

Model 4: all variables by group

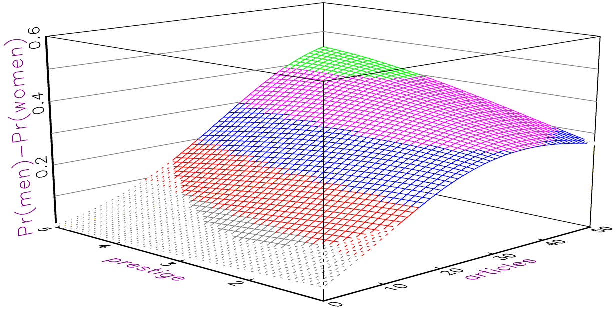
Or, reversing job prestige and articles in the graph:



Discrete change as prestige and articles varies - III

Model 4: all variables by group

Equivalently, in three dimensions:



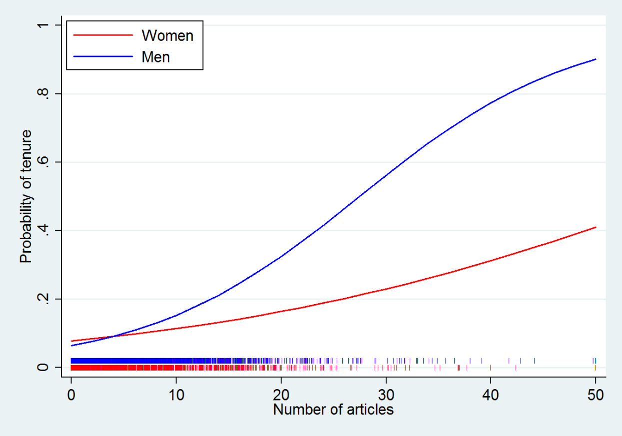
Summary

Basic issues in group comparison in logit and probit

1. In the LRM we can use standard test to compare coefficients across groups.
2. In the BRM, these tests are problematic due to identification issues.
3. We can make stronger assumptions to test the coefficients.
4. Tests comparing predicted probabilities are unaffected by the identification problem.
5. But you must deal with the interpretation of nonlinear models since a single test is not possible.

Is there sufficient data to draw conclusions?

The distribution of articles



Installing the SPost programs in Stata

In Stata while connected to the internet:

```
findit spost9
```

and follow instructions to install the packages:

```
spost9_ado :  Stata 9 SPost ado files.
spost_nd :  Long - ND 2007 sample files.
```

References

1. Allison, Paul D. 1999. "Comparing Logit and Probit Coefficients Across Groups." Sociological Methods and Research 28:186-208.
2. Chow, G.C. 1960. "Tests of equality between sets of coefficients in two linear regressions." Econometrica 28:591-605.
3. Long, J.S. 2005. Comparing Groups using Predicted Outcomes. Working paper.
4. Long, J.S. and Freese, J. 2005. Regression Models for Categorical and Limited Dependent Variables with Stata. Second Edition. College Station, TX: Stata Press.
5. Long, J. Scott, Paul D. Allison, and Robert McGinnis. 1993. "Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity." American Sociological Review 58:703-722.
6. Xu, Jun and Long, J. 2005, Confidence intervals for predicted outcomes in regression models for categorical outcomes. The Stata Journal 5: 537-559.

Confidence intervals for predictions - I

CIs computed by delta method or bootstrap

Delta

1. Take a Taylor series expansion of $G(\widehat{\boldsymbol{\beta}}) = \Pr(y = 1 \mid \mathbf{x}) = F(\mathbf{x}\boldsymbol{\beta})$:
$$G(\widehat{\boldsymbol{\beta}}) \approx G(\boldsymbol{\beta}) + (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T G'(\boldsymbol{\beta})$$
2. Where
$$G'(\boldsymbol{\beta}) = \begin{bmatrix} \frac{\partial F(\mathbf{x}\boldsymbol{\beta})}{\partial \beta_0} & \frac{\partial F(\mathbf{x}\boldsymbol{\beta})}{\partial \beta_1} & \cdots & \frac{\partial F(\mathbf{x}\boldsymbol{\beta})}{\partial \beta_K} \end{bmatrix}^T$$
3. Under standard assumptions, $G(\widehat{\boldsymbol{\beta}})$ is distributed normally around $G(\boldsymbol{\beta})$ with a variance
$$Var\left[G(\widehat{\boldsymbol{\beta}})\right] = G'(\widehat{\boldsymbol{\beta}})^T Var(\widehat{\boldsymbol{\beta}}) G'(\widehat{\boldsymbol{\beta}})$$

Confidence intervals for predictions - II

CIs computed by delta method or bootstrap

4. Or:

$$Var\left[\widehat{\Pr}\left(y=1 \mid \mathbf{x}\right)\right] = \left[\frac{\partial F\left(\mathbf{x}\widehat{\boldsymbol{\beta}}\right)}{\partial\widehat{\boldsymbol{\beta}}}\right]^T Var(\widehat{\boldsymbol{\beta}}) \left[\frac{\partial F\left(\mathbf{x}\widehat{\boldsymbol{\beta}}\right)}{\partial\widehat{\boldsymbol{\beta}}}\right]$$

5. For discrete change:

$$\begin{aligned} Var\left[F(\widehat{\boldsymbol{\beta}}|\mathbf{x}_a) - F\left(\widehat{\boldsymbol{\beta}}|\mathbf{x}_b\right)\right] \\ = \left\{\left[\frac{\partial F(\boldsymbol{\beta}|\mathbf{x}_a)}{\partial\boldsymbol{\beta}^T} Var(\widehat{\boldsymbol{\beta}}) \frac{\partial F(\boldsymbol{\beta}|\mathbf{x}_a)}{\partial\boldsymbol{\beta}}\right] - \left[\frac{\partial F(\boldsymbol{\beta}|\mathbf{x}_a)}{\partial\boldsymbol{\beta}^T} Var(\widehat{\boldsymbol{\beta}}) \frac{\partial F(\boldsymbol{\beta}|\mathbf{x}_b)}{\partial\boldsymbol{\beta}}\right]\right\} \\ - \left\{\left[\frac{\partial F(\boldsymbol{\beta}|\mathbf{x}_b)}{\partial\boldsymbol{\beta}^T} Var(\widehat{\boldsymbol{\beta}}) \frac{\partial F(\boldsymbol{\beta}|\mathbf{x}_a)}{\partial\boldsymbol{\beta}}\right] - \left[\frac{\partial F(\boldsymbol{\beta}|\mathbf{x}_b)}{\partial\boldsymbol{\beta}^T} Var(\widehat{\boldsymbol{\beta}}) \frac{\partial F(\boldsymbol{\beta}|\mathbf{x}_b)}{\partial\boldsymbol{\beta}}\right]\right\} \end{aligned}$$

Bootstrap

To get stable results, you need to use at least 1,000 replications.