Intermediate Social Statistics Hilary 2009 Lecture 4: Binary Discrete Dependent Variable

Raymond Duch

Nuffield College

February 16, 2009

The Likelihood Function

- In this example is Labour incumbent vote and takes on a value of 1 or 0.
- The independent variable, is income category that ranges in value from 10 (10,000) to 100 (100,000 or greater).

$$lnL = \sum_{y_{i=1}}^{N=10} y_i lnF(\alpha + \beta_1 Income) + \sum_{y_{i=0}}^{N=10} (1 - y_i) ln[1 - F(\alpha + \beta_1 Income)]$$

The Likelihood Function

- This can simply be calculated by hand
- Remember that the CDF for the logit is

$$F(\alpha + \beta_1 Income) = \frac{1}{1 + e^{-(\alpha+\beta_1 Income)}}$$

which is equivalent to

$$F(\alpha + \beta_1 Income) = \frac{e^{(\alpha+\beta_1 Income)}}{1 + e^{(\alpha+\beta_1 Income)}}$$

The Likelihood Function

Lets try α=-.02 and β_1=1.2

Table: α=-.02 and β_1=1.2

Labour Vote	Income Category	Likelihood
1	10	-0.313
1	20	-0.371
1	30	-0.437
0	40	-0.913
0	50	-0.798
1	60	-0.693
0	70	-0.598
0	80	-0.513
0	90	-0.437
0	100	-0.371
Log Likelihood		-5.44

The Likelihood Function by Hand

Lets try α=-.05 and β_1=1.0

Table: α=-.05 and β_1=1.0

Labour Vote	Income Category	Likelihood
1	10	-0.474
1	20	-0.693
1	30	-0.974
0	40	-0.313
0	50	-0.201
1	60	-2.126
0	70	-0.078
0	80	-0.0485
0	90	-0.0297
0	100	-0.0181
Log Likelihood		-4.958

The Likelihood Function by Hand

Lets try α=-.086 and β_1=3.879

Table: α=-.086 and β_1=3.879

Labour Vote	Income Category	Likelihood
1	10	-0.047
1	20	-0.109
1	30	-0.241
0	40	-0.936
0	50	-0.505
1	60	-1.526
0	70	-0.111
0	80	-0.048
0	90	-0.021
0	100	-0.009
Log Likelihood		-3.55

The Likelihood Function by Hand

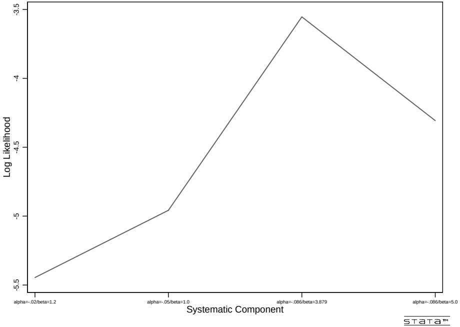
Lets try α=-.086 and β_1=5

Table: α=-.086 and β_1=5

Labour Vote	Income Category	Likelihood
1	10	-0.016
1	20	-0.037
1	30	-0.085
0	40	-1.75
0	50	-1.103
1	60	-0.776
0	70	-0.308
0	80	-0.142
0	90	-0.063
0	100	-0.027
Log Likelihood		-4.308

The Likelihood Function by Hand

Figure: Plot of likelihoods suggesting a maximum at α=-.086 and β_1=3.879



Marginal Effects

- want the marginal effects of a change in some x on the probability of y = 1
- note that the parameters of the model, β, are not the marginal effect of x on y
- in general:

$$\frac{\partial E[y|\mathbf{x}]}{\partial \mathbf{x}} = \{\frac{dF(\mathbf{x}\beta)}{d(\mathbf{x}\beta)}\}\beta = f(\mathbf{x}\beta)\beta$$

Marginal Effects

- note: values of the marginal effects (partials of y wrt x are) will change with x
- Linear case: marginal effect of x on y does not change anywhere over the course of that line. It is always β.
- Non-linear case:

$$\frac{d\Lambda(\mathbf{x}\beta)}{d(\mathbf{x}\beta)} = \frac{e^{\mathbf{x}\beta}}{(1+e^{\mathbf{x}\beta})^2} = \Lambda(\mathbf{x}\beta)[1-\Lambda(\mathbf{x}\beta)]$$

or:

$$\frac{\partial E[y|\mathbf{x}]}{\partial \mathbf{x}} = \Lambda(\mathbf{x}\beta)[1-\Lambda(\mathbf{x}\beta)]\beta$$

Marginal Effects

- Note that the marginal effects in the previous equation are conditional on the value of the x variable of interest
- But also conditional on the value of all the other independent variables in model (plus their coefficients)
- The following example calculates the marginal effects for logit regression equation using the Stata logistic regression results for the UK 2004 model introduced at the beginning of the lecture:
  - Labour vote = retnat + class + union + southwest + urban + lrself + own + ε

Marginal Effects

Figure: Stata Estimation of Logit Labour Vote Model

```
Tuesday January 29 13:11:38 2008   Page 1

                                          ___  ____  ____  ____  ____tm
                                         /__    /   ____/   /   ____/
                                        ___/   /   /___/   /   /___/
                                          Statistics/Data Analysis

1 . logit  incumvote retnat class union southwest urban lrself own

  Iteration 0:   log likelihood =  -507.77976
  Iteration 1:   log likelihood =  -445.90699
  Iteration 2:   log likelihood =  -443.71375
  Iteration 3:   log likelihood =  -443.69615
  Iteration 4:   log likelihood =  -443.69615

  Logistic regression                               Number of obs   =        785
                                                       LR chi2( 7)      =     128.17
                                                       Prob > chi2      =     0.0000
  Log likelihood =  -443.69615                         Pseudo R2        =     0.1262
```

incumvote	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
retnat	-.7038763	.1034497	-6.80	0.000	-.906634	-.5011186
class	-.3854525	.0861783	-4.47	0.000	-.5543589	-.216546
union	.5060247	.1915331	2.64	0.008	.1306267	.8814226
southwest	-.933915	.3541623	-2.64	0.008	-1.62806	-.2397697
urban	.3036703	.1060421	2.86	0.004	.0958315	.5115091
lrself	-.1339321	.0369293	-3.63	0.000	-.2063122	-.0615519
own	-.5231139	.1897608	-2.76	0.006	-.8950382	-.1511896
_cons	1.97143	.4262088	4.63	0.000	1.136076	2.806784

Mean Values of Variables in Model

Figure: Variables in the Logit Model

Tuesday January 29 12:43:44 2008 Page 1

Statistics/Data Analysis

Variable	Obs	Mean	Std. Dev.	Min	Max
retnat	1454	2.102476	.8110265	1	3

1 . tabulate retnat

retnat	Freq.	Percent	Cum.
better	411	28.27	28.27
same	483	33.22	61.49
worse	560	38.51	100.00
Total	1,454	100.00	

2 . tabulate class

class	Freq.	Percent	Cum.
working	600	43.20	43.20
lower middle	259	18.65	61.84
middle	461	33.19	95.03
upper middle	62	4.46	99.50
upper	7	0.50	100.00
Total	1,389	100.00	

3 . tabulate union

union	Freq.	Percent	Cum.
non-union	1,157	77.65	77.65
union	333	22.35	100.00
Total	1,490	100.00	

4 . tabulate southwest

southwest	Freq.	Percent	Cum.
0	1,370	91.33	91.33
1	130	8.67	100.00
Total	1,500	100.00	

5 . tabulate urban

urban	Freq.	Percent	Cum.
rural	493	33.11	33.11
small-medium town	576	38.68	71.79
large town	420	28.21	100.00
Total	1,489	100.00	

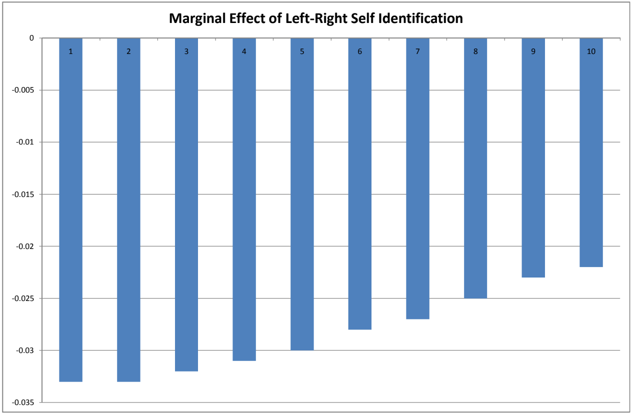
Calculating Marginal Effects for Left-Right Variable

- set all other independent variables to their means – a in previous slide
- calculate marginal effect for each value of left-right variable (with all other variables at their mean):

$$\frac{e^{\hat{\beta}_1(LR=1)+\Sigma_{j=1}^{J}\hat{\phi}_j\bar{Z_{ji}}}}{1+e^{\hat{\beta}_1(LR=1)+\Sigma_{j=1}^{J}\hat{\phi}_j\bar{Z_{ji}}}} \times [1 - \frac{e^{\hat{\beta}_1(LR=1)+\Sigma_{j=1}^{J}\hat{\phi}_j\bar{Z_{ji}}}}{1+e^{\hat{\beta}_1(LR=1)+\Sigma_{j=1}^{J}\hat{\phi}_j\bar{Z_{ji}}}}] \times \hat{\beta}_1 \qquad (1)$$

Generating Marginal Effects

Figure: Stata Estimation of Logit Labour Vote Model



Generating Point Estimates

- Interpretations of the parameter effects β are sensitive to the values of X
- A more straightforward and informative strategy is to generate some interesting predictions
- Here are the Stata logistic regression results for the UK 2004 model introduced at the beginning of the lecture
- Labour vote = retnat + class + union + southwest + urban + lrself + own + ε

Generating Point Estimates

Figure: Stata Estimation of Logit Labour Vote Model

```
  Tuesday January 29 13:11:38 2008   Page 1

                                                      Statistics/Data Analysis

1 . logit  incumvote retnat class union southwest urban lrself own

  Iteration 0:   log likelihood =  -507.77976
  Iteration 1:   log likelihood =  -445.90699
  Iteration 2:   log likelihood =  -443.71375
  Iteration 3:   log likelihood =  -443.69615
  Iteration 4:   log likelihood =  -443.69615

  Logistic regression                               Number of obs   =        785
                                                      LR chi2( 7)     =     128.17
                                                      Prob > chi2     =     0.0000
  Log likelihood =  -443.69615                        Pseudo R2       =     0.1262
```

incumvote	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
retnat	-.7038763	.1034497	-6.80	0.000	-.906634	-.5011186
class	-.3854525	.0861783	-4.47	0.000	-.5543589	-.216546
union	.5060247	.1915331	2.64	0.008	.1306267	.8814226
southwest	-.933915	.3541623	-2.64	0.008	-1.62806	-.2397697
urban	.3036703	.1060421	2.86	0.004	.0958315	.5115091
lrself	-.1339321	.0369293	-3.63	0.000	-.2063122	-.0615519
own	-.5231139	.1897608	-2.76	0.006	-.8950382	-.1511896
_cons	1.97143	.4262088	4.63	0.000	1.136076	2.806784

Generating Point Estimates

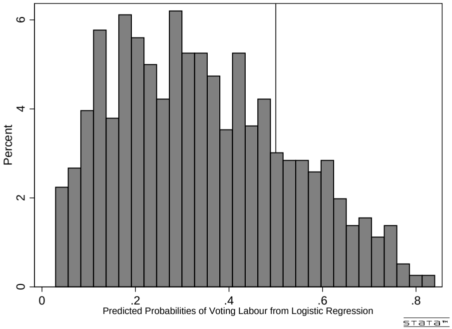
Using these estimates we can generate a predicted Labour vote probability for any respondent in our data set.

$$Prob(LabourVote_i = 1|\mathbf{x_i}) = \frac{e^{\mathbf{x}\beta}}{1 + e^{\mathbf{x}\beta}}$$

- Here is a histogram of the predicted vote for all respondents in the UK 2004 survey based on each of their individual characteristics
- How does this differ from the histogram of predicted votes generated by OLS regression?

Generating Point Estimates

Figure: Predicted Probabilities from Logistic Regression of Labour Vote



Generating Meaningful Predictions

Figure: Variables in the Logit Model

Tuesday January 29 12:43:44 2008 Page 1

Statistics/Data Analysis

retnat	1454	2.102476	.8110265	1	3

1 . tabulate retnat

retnat	Freq.	Percent	Cum.
better	411	28.27	28.27
same	483	33.22	61.49
worse	560	38.51	100.00
Total	1,454	100.00	

2 . tabulate class

class	Freq.	Percent	Cum.
working	600	43.20	43.20
lower middle	259	18.65	61.84
middle	461	33.19	95.03
upper middle	62	4.46	99.50
upper	7	0.50	100.00
Total	1,389	100.00	

3 . tabulate union

union	Freq.	Percent	Cum.
non-union	1,157	77.65	77.65
union	333	22.35	100.00
Total	1,490	100.00	

4 . tabulate southwest

southwest	Freq.	Percent	Cum.
0	1,370	91.33	91.33
1	130	8.67	100.00
Total	1,500	100.00	

5 . tabulate urban

urban	Freq.	Percent	Cum.
rural	493	33.11	33.11
small-medium town	576	38.68	71.79
large town	420	28.21	100.00
Total	1,489	100.00	

Generating Meaningful Predictions

$$Prob(y_i = 1|\mathbf{x_i}) = \frac{e^{\mathbf{x_i}\beta}}{1 + e^{\mathbf{x_i}\beta}} = \Lambda(\mathbf{x_i}\beta)$$

- where we define a vector of x values for $\mathbf{x_i}$ for a particular interesting case
- we use the β vector to generate the predicted probability of voting Labour for this particular "individual".

Generating Meaningful Predictions

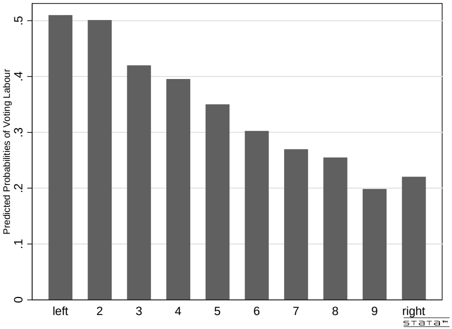
Table: Left-Wing Urban Working Class Male with Favorable View of Economy

Variable	Value	Coefficient	Value
Retnat	1	-.703	-.703
Class	1	-.385	-.385
Union	0	.506	0
Southwest	0	-.933	0
Urban	1	.303	.303
Lrself	1	-.133	-.133
Own	1	-.523	-.523
Constant	1	1.97	1.97
Systematic Component			0.529

$$Prob(y_i = 1|\mathbf{x_i}) = \frac{e^{0.529}}{1 + e^{0.529}}$$
$$= .63$$

Generating Meaningful Predictions

Figure: Predicted Probabilities of Voting Labour by Left-Right Self Identification



Illustrating Meaningful Predictions: The Economic Vote

- Generating estimated changes in probabilities associated with meaningful changes in the independent variables.
- Here is an illustration of estimating the "economic vote":
  - the change in the probabilities of voting for the incumbent government (Labour) when economic evaluations shift one unit on a three-unit economic evaluation scale.

Generating Meaningful Predictions

The economic vote for any individual (EV_i) in the sample data set is the following:

$$EV_i = \frac{e^{\hat{\beta}_1(Retnat=1)+\Sigma_{j=1}^{J}\hat{\phi}_j Z_{ji}}}{1+e^{\hat{\beta}_1(Retnat=1)+\Sigma_{j=1}^{J}\hat{\phi}_j Z_{ji}}} - \frac{e^{\hat{\beta}_1(Retnat=2)+\Sigma_{j=1}^{J}\hat{\phi}_j Z_{ji}}}{1+e^{\hat{\beta}_1(Retnat=2)+\Sigma_{j=1}^{J}\hat{\phi}_j Z_{ji}}} \qquad (2)$$

Illustrating Meaningful Predictions: The Economic Vote

Table: Left-Wing Urban Working Class Male with Indifferent View of Economy

Variable	Value	Coefficient	Value
Retnat	2	-.703	-1.406
Class	1	-.385	-.385
Union	0	.506	0
Southwest	0	-.933	0
Urban	1	.303	.303
Lrself	1	-.133	-.133
Own	1	-.523	-.523
Constant	1	1.97	1.97
Systematic Component			-0.174

Illustrating Meaningful Predictions: The Economic Vote

$$
\begin{aligned}
Prob(y_i = 1|\mathbf{x_i}) &= \frac{e^{-0.174}}{1 + e^{-0.174}} \\
&= .46
\end{aligned}
$$

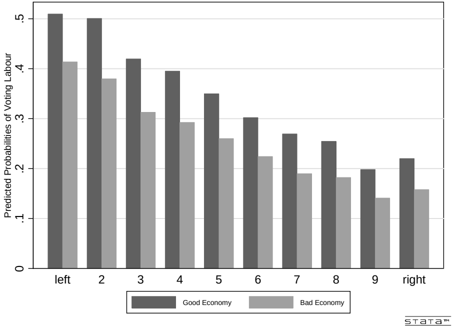
$$
\begin{aligned}
EVi &= Prob(y_i = 1|\mathbf{x_i}) - Prob(y_i = 1|\hat{\mathbf{x}}_\mathbf{i}) \\
&= \frac{e^{0.529}}{1 + e^{0.529}} - \frac{e^{-0.174}}{1 + e^{-0.174}} \\
&= 0.63 - 0.46 \\
&= 0.17
\end{aligned}
$$

Illustrating Meaningful Predictions: The Economic Vote

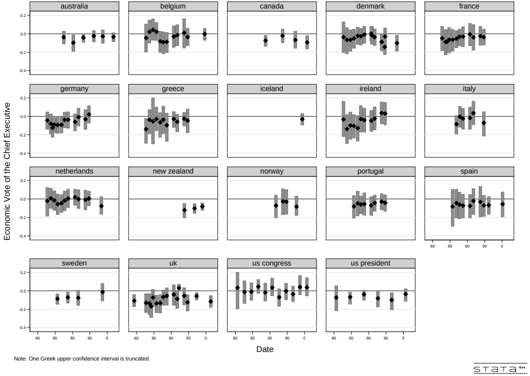
- Each of these estimated individual "economic votes" can be summarized for the whole sample
- simply take the average of these estimated changes in vote probabilities over the whole sample
- $EV = \Sigma_{n=1}^{N} EV_i / N$

Illustrating Meaningful Predictions: The Economic Vote

Figure: Predicted Probabilities of voting Labour for Left-Right and Good-Bad Economic Perceptions



Further Illustrations: The Economic Vote



Further Illustrations: The Economic Vote