

# Intermediate Social Statistics Hilary 2010 Lecture 4: Binary Discrete Dependent Variable

Raymond Duch

Nuffield College

February 8, 2010

# 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).

$$\ln L = \sum_{y_i=1}^{N=10} y_i \ln F(\alpha + \beta_1 \text{Income}) + \sum_{y_i=0}^{N=10} (1 - y_i) \ln [1 - F(\alpha + \beta_1 \text{Income})]$$

# The Likelihood Function

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

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

which is equivalent to

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

# The Likelihood Function

Lets try  $\alpha = -.02$  and  $\beta_1 = 1.2$

Table:  $\alpha = -.02$  and  $\beta_1 = 1.2$

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  $\alpha = -.05$  and  $\beta_1 = 1.0$

Table:  $\alpha = -.05$  and  $\beta_1 = 1.0$

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  $\alpha = -.086$  and  $\beta_1 = 3.879$

Table:  $\alpha = -.086$  and  $\beta_1 = 3.879$

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

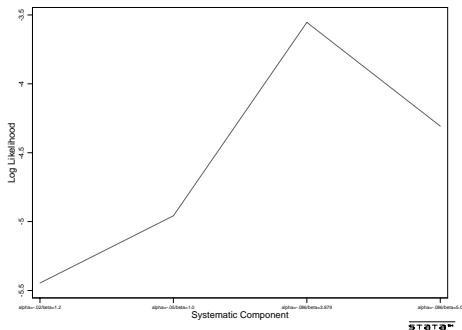
Lets try  $\alpha = -.086$  and  $\beta_1 = 5$

Table:  $\alpha = -.086$  and  $\beta_1 = 5$

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  $\alpha = -.086$  and  $\beta_1 = 3.879$



# 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,  $\beta$ , are not the marginal effect of  $x$  on  $y$
- in general:

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

## 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  $\beta$ .
- 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:

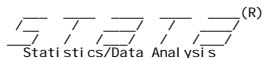
$$\frac{\partial E[y|\mathbf{x}]}{\partial \mathbf{x}} = \Lambda(\mathbf{x}\beta)[1 - \Lambda(\mathbf{x}\beta)]\beta$$

## 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 =  $\text{retnat} + \text{class} + \text{union} + \text{southwest} + \text{urban} + \text{lrsel} + \text{own} + \varepsilon$

# Marginal Effects

Monday February 8 14:57:51 2010 Page 1



Logistic regression

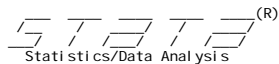
Number of obs = 785  
 LR chi2(7) = 128.17  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.1262

Log Likelihood = **-443.69615**

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
lrsel f	-.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

Monday February 8 15:19:15 2010 Page 1



- ```
1 . end of do-file
2 . tabulate retnat
```

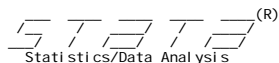
| retnat | Freq.        | Percent       | Cum.          |
|--------|--------------|---------------|---------------|
| better | <b>411</b>   | <b>28.27</b>  | <b>28.27</b>  |
| same   | <b>483</b>   | <b>33.22</b>  | <b>61.49</b>  |
| worse  | <b>560</b>   | <b>38.51</b>  | <b>100.00</b> |
| Total  | <b>1,454</b> | <b>100.00</b> |               |

- ```
3 . tabulate class
```

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

# Mean Values of Variables in Model

Monday February 8 15:24:43 2010 Page 1



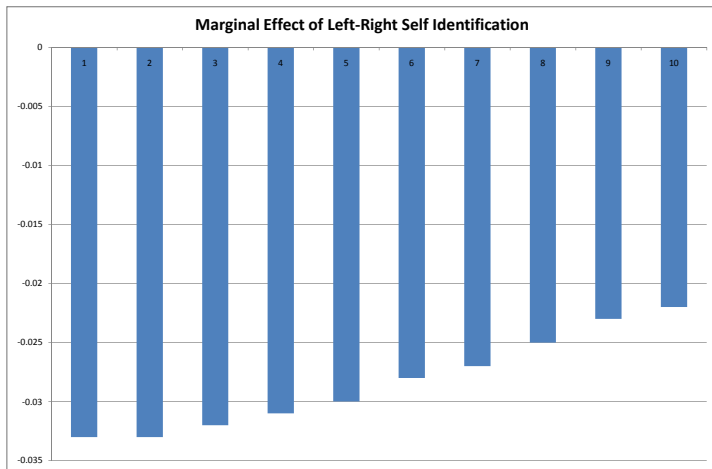
Total		<b>1,489</b>	<b>100.00</b>	
1. tabulate union				
union		Freq.	Percent	Cum.
0		<b>1,157</b>	<b>77.65</b>	<b>77.65</b>
1		<b>333</b>	<b>22.35</b>	<b>100.00</b>
Total		<b>1,490</b>	<b>100.00</b>	
2. tabulate southwest				
southwest		Freq.	Percent	Cum.
0		<b>1,370</b>	<b>91.33</b>	<b>91.33</b>
1		<b>130</b>	<b>8.67</b>	<b>100.00</b>
Total		<b>1,500</b>	<b>100.00</b>	
3. tabulate urban				
urban		Freq.	Percent	Cum.
small		<b>493</b>	<b>33.11</b>	<b>33.11</b>
medium		<b>576</b>	<b>38.68</b>	<b>71.79</b>
large		<b>420</b>	<b>28.21</b>	<b>100.00</b>
Total		<b>1,489</b>	<b>100.00</b>	

## 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):

$$\frac{e^{\hat{\beta}_1(LR=1)+\sum_{j=1}^J \hat{\phi}_j \bar{Z}_{ji}}}{1 + e^{\hat{\beta}_1(LR=1)+\sum_{j=1}^J \hat{\phi}_j \bar{Z}_{ji}}} \times \left[ 1 - \frac{e^{\hat{\beta}_1(LR=1)+\sum_{j=1}^J \hat{\phi}_j \bar{Z}_{ji}}}{1 + e^{\hat{\beta}_1(LR=1)+\sum_{j=1}^J \hat{\phi}_j \bar{Z}_{ji}}} \right] \times \hat{\beta}_1 \quad (1)$$

# Generating Marginal Effects



# Generating Point Estimates

- Interpretations of the parameter effects  $\beta$  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 + lrsel  
+ own +  $\varepsilon$

# Generating Point Estimates

Figure: Stata Estimation of Logit Labour Vote Model

Monday February 8 14: 57: 51 2010 Page 1

 Statistics/Data Analysis<sup>(R)</sup>

Logistic regression

Number of obs = 785  
 LR chi2(7) = 128.17  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.1262

Log likelihood = -443.69615

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
lrsel f	-.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

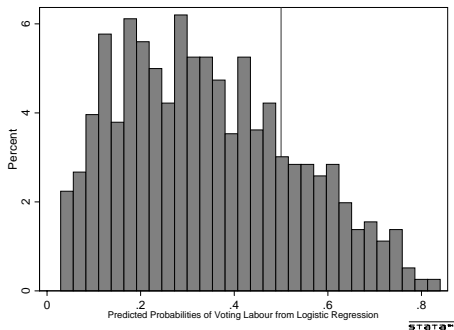
Using these estimates we can generate a predicted Labour vote probability for any respondent in our data set.

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

- 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?

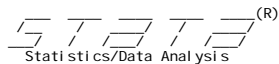
# Generating Point Estimates

Figure: Predicted Probabilities from Logistic Regression of Labour Vote



# Generating Meaningful Predictions

Monday February 8 15:19:15 2010 Page 1



- ```
1 . end of do-file
2 . tabulate retnat
```

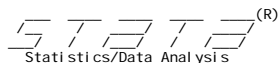
| retnat | Freq.        | Percent       | Cum.          |
|--------|--------------|---------------|---------------|
| better | <b>411</b>   | <b>28.27</b>  | <b>28.27</b>  |
| same   | <b>483</b>   | <b>33.22</b>  | <b>61.49</b>  |
| worse  | <b>560</b>   | <b>38.51</b>  | <b>100.00</b> |
| Total  | <b>1,454</b> | <b>100.00</b> |               |

- ```
3 . tabulate class
```

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

# Generating Meaningful Predictions

Monday February 8 15:24:43 2010 Page 1



Total		<b>1,489</b>	<b>100.00</b>	
1. tabulate union				
union		Freq.	Percent	Cum.
0		<b>1,157</b>	<b>77.65</b>	<b>77.65</b>
1		<b>333</b>	<b>22.35</b>	<b>100.00</b>
Total		<b>1,490</b>	<b>100.00</b>	
2. tabulate southwest				
southwest		Freq.	Percent	Cum.
0		<b>1,370</b>	<b>91.33</b>	<b>91.33</b>
1		<b>130</b>	<b>8.67</b>	<b>100.00</b>
Total		<b>1,500</b>	<b>100.00</b>	
3. tabulate urban				
urban		Freq.	Percent	Cum.
small		<b>493</b>	<b>33.11</b>	<b>33.11</b>
medium		<b>576</b>	<b>38.68</b>	<b>71.79</b>
large		<b>420</b>	<b>28.21</b>	<b>100.00</b>
Total		<b>1,489</b>	<b>100.00</b>	

# 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  $\beta$  vector to generate the predicted probability of voting Labour for this particular "individual".

# Generating Meaningful Predictions

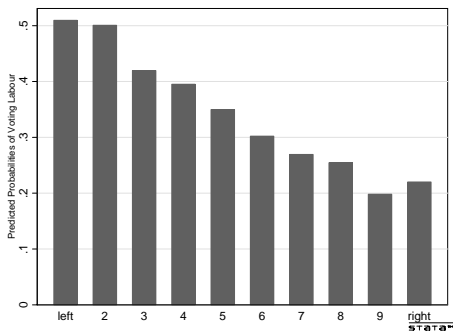
Table: Left-Wing Urban Working Class Male with Favorable View of Economy

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

$$\begin{aligned}
 \text{Prob}(y_i = 1 | \mathbf{x}_i) &= \frac{e^{0.529}}{1 + e^{0.529}} \\
 &= .63
 \end{aligned}$$

# 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:

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

# 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

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

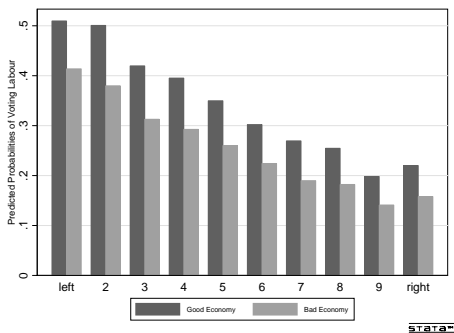
$$\begin{aligned} EV_i &= \text{Prob}(y_i = 1 | \mathbf{x}_i) - \text{Prob}(y_i = 1 | \hat{\mathbf{x}}_i) \\ &= \frac{e^{0.529}}{1 + e^{0.529}} - \frac{e^{-0.174}}{1 + e^{-0.174}} \\ &= 0.63 - 0.46 \\ &= 0.17 \end{aligned}$$

# Illustrating Meaningful Predictions: The Economic Vote

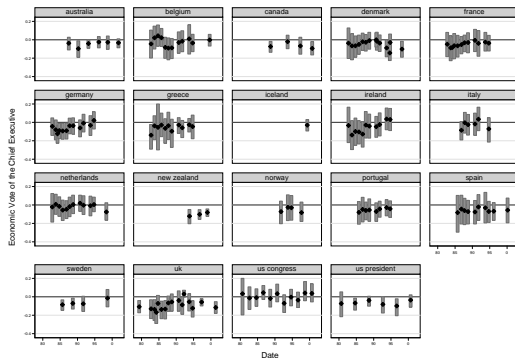
- Each of these estimated individual "economic votes" can be summarized for the whole sample
- simply take the average of these estimated changes in vote probabilities over the whole sample
- $EV = \sum_{n=1}^N EV_i / N$

# Illustrating Meaningful Predictions: The Economic Vote

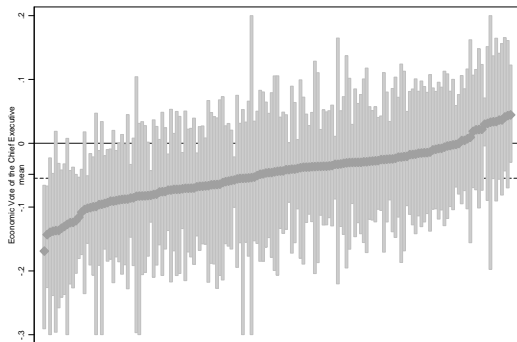
**Figure:** Predicted Probabilities of voting Labour for Left-Right and Good-Bad Economic Perceptions



# Further Illustrations: The Economic Vote



# Further Illustrations: The Economic Vote



Confidence bounds greater than .2 and less than -.3 truncated for display

STATA™