

20) Logit and Probit I

Vitor Kamada

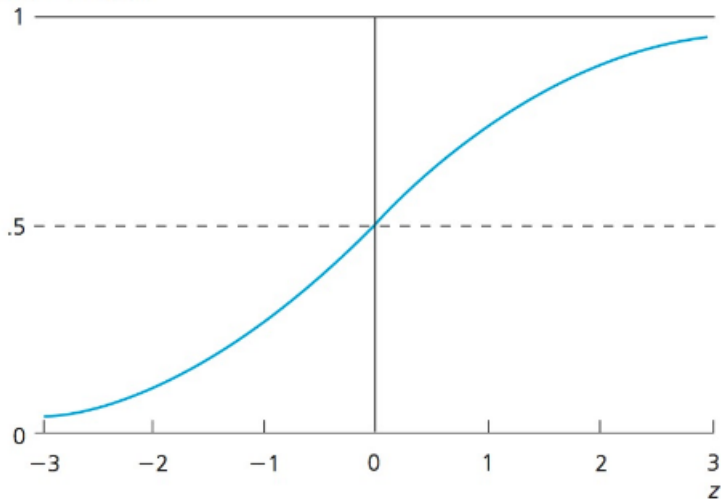
October 2018

Tables, Graphics, and Figures from **Introductory Econometrics**

Wooldridge, J. M. (2015): Chapter 17 - Limited
Dependent Variable Models

Graph of the Logistic Function

$$G(z) = \exp(z) / [1 + \exp(z)]$$



Specifying Logit and Probit Models

$$P(y = 1|x) = G(\beta_0 + x\beta)$$

$$0 < G(z) < 1$$

$$G(z) = \frac{e^z}{1+e^z} = \Lambda(z)$$

$$G(z) = \Phi(z) = \int_{-\infty}^z \phi(v) dv$$

$$\phi(v) = (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{z^2}{2}\right)$$

$$y^* = \beta_0 + \mathbf{x}\beta + e$$

$$y = 1[y^* > 0]$$

$$e \perp \mathbf{x}$$

$$e \sim N(0, 1) \text{ or } \textit{Logistic}(0, 1)$$

$$\therefore 1 - G(-z) = G(z)$$

Response Probability for y

$$\begin{aligned}P(y = 1|x) &= P(y^* > 0|x) = P[e > -(\beta_0 + x\beta)|x] \\&= 1 - G[-(\beta_0 + x\beta)] = G(\beta_0 + x\beta)\end{aligned}$$

$$\frac{\partial p(x)}{\partial x_j} = g(\beta_0 + x\beta)\beta_j$$

$$g(z) = \frac{dG}{dz}(z) > 0$$

$G(\cdot)$ is a strictly increasing cdf

$$\phi(0) = \frac{1}{\sqrt{2\pi}} \approx .4, \quad \lambda(0) = \frac{\exp(0)}{[1+\exp(0)]^2} \approx .25$$

Partial Effect of Continuous Variables on the Response Probability

$$\frac{\partial p(x)}{\partial x_j} = g(\beta_0 + x\beta)\beta_j$$

$$\frac{\partial p(x)}{\partial x_j} / \frac{\partial p(x)}{\partial x_h} = \frac{\beta_j}{\beta_h}$$

Binary: $\frac{\Delta p(x)}{\Delta x_1} = G(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)$

$$- G(\beta_0 + \beta_2 x_2 + \dots + \beta_k x_k)$$

$$f(y|x_i; \beta) = [G(x_i\beta)]^y [1 - G(x_i\beta)]^{1-y}$$

$$\ell_i(\beta) = y_i \log[G(x_i\beta)] + (1 - y_i) \log[1 - G(x_i\beta)]$$

$$\mathcal{L}(\beta) = \sum_{i=1}^n \ell_i(\beta)$$

$$\text{Pseudo } R^2 = 1 - \frac{\mathcal{L}_{ur}}{\mathcal{L}_0}$$

\mathcal{L}_0 the model with only an intercept

$$|\mathcal{L}_{ur}| \leq |\mathcal{L}_0|$$

$$\tilde{y}_i = 1 \text{ if } G(\hat{\beta}_0 + x_i\hat{\beta}) \geq .5$$

$$\tilde{y}_i = 0 \text{ if } G(\hat{\beta}_0 + x_i\hat{\beta}) < .5$$

	(y_i, \tilde{y}_i)	
Correct Prediction	0,0	1,1
Incorrect Prediction	0,1	1,0

The % correctly predicted is the % of times that $y_i = \tilde{y}_i$

Table 17.1 Estimates of Labor Force Participation

Independent Variables	LPM (OLS)	Logit (MLE)	Probit (MLE)
<i>nwifeinc</i>	-.0034 (.0015)	-.021 (.008)	-.012 (.005)
<i>educ</i>	.038 (.007)	.221 (.043)	.131 (.025)
<i>exper</i>	.039 (.006)	.206 (.032)	.123 (.019)
<i>exper</i> ²	-.00060 (.00019)	-.0032 (.0010)	-.0019 (.0006)
<i>age</i>	-.016 (.002)	-.088 (.015)	-.053 (.008)
<i>kidslt6</i>	-.262 (.032)	-1.443 (.204)	-.868 (.119)
<i>kidsge6</i>	.013 (.014)	.060 (.075)	.036 (.043)
<i>constant</i>	.586 (.152)	.425 (.860)	.270 (.509)
Percentage correctly predicted	73.4	73.6	73.4
Log-likelihood value	—	-401.77	-401.30
Pseudo <i>R</i> -squared	.264	.220	.221

Partial Effect at the Average (PEA)

$$\Delta \hat{P}(y = 1|x) \approx [g(\hat{\beta}_0 + x\hat{\beta})\hat{\beta}_j]\Delta x_j$$

$$g(\hat{\beta}_0 + \bar{x}\hat{\beta}) = g(\hat{\beta}_0 + \hat{\beta}_1\bar{x}_1 + \hat{\beta}_2\bar{x}_2 + \dots + \hat{\beta}_k\bar{x}_k)$$

Partial Effect of x_j for the “average” person in the sample.

Average Partial Effect (APE) or Average Marginal Effect (AME)

$$n^{-1} \sum_{i=1}^n [g(\hat{\beta}_0 + \mathbf{x}_i \hat{\beta}) \hat{\beta}_j]$$

AME is the average of the nonlinear function rather than the nonlinear function of the average

$$g[E(\mathbf{x}\beta)] \neq E[g(\mathbf{x}\beta)]$$

Table 17.2 Average Partial Effects for the Labor Force Participation Models

Independent Variables	LPM	Logit	Probit
<i>nwifeinc</i>	-.0034 (.0015)	-.0038 (.0015)	-.0036 (.0014)
<i>educ</i>	.038 (.007)	.039 (.007)	.039 (.007)
<i>exper</i>	.027 (.002)	.025 (.002)	.026 (.002)
<i>age</i>	-.016 (.002)	-.016 (.002)	-.016 (.002)
<i>kids116</i>	-.262 (.032)	-.258 (.032)	-.261 (.032)
<i>kidsge6</i>	.013 (.014)	.011 (.013)	.011 (.013)

Figure 17.2 Estimated Response Probabilities with Respect to Education for the Linear Probability and Probit Models

