UCD GEARY INSTITUTE DISCUSSION PAPER SERIES

Very simple marginal effects in some discrete choice models[*]

Kevin J. Denny

School of Economics & Geary Institute
University College Dublin

21st July 2009

Abstract

I show a simple back-of-the-envelope method for calculating marginal effects in binary choice and count data models. The approach suggested here focuses attention on marginal effects at different points in the distribution of the dependent variable rather than representative points in the joint distribution of the explanatory variables. For binary models, if the mean of the dependent variable is between 0.4 and 0.6 then dividing the logit coefficient by 4 or multiplying the probit coefficient by 0.4 should be moderately accurate.

1. Introduction

Limited dependent variable models are widely used in the analysis of survey data. Unlike linear regression model, the magnitudes of the parameters are not easily interpreted. For this reason it is common to present estimates of marginal effects: the effect of a small change in the covariates on the probability of a particular outcome. Statistical packages are increasingly including one or more methods for doing this. This note presents a very simple way of doing this for a number of leading limited dependent variable models, probit, logit, Poisson and negative binomial regression, which can be calculated on the back of an envelope.

2. Binary choice models

Say one is estimating a binary choice model given by:

$$E(Y/X) = F(bX) \tag{1}$$

Y is a binary variable, X is a design matrix including a constant. The model is probit or logit depending on the choice of F. The marginal effect for x_1 then is:

$$\frac{\partial E(Y/X)}{\partial x_1} = f(bX)b_1 \tag{2}$$

Where f(.) is the density function corresponding to the distribution function F(.). Hence the marginal effect is the product of the relevant coefficient and a scale factor which will be common to all variables but which will vary from one observation to the next. Standard solutions are to evaluate the scale factor at (i) the mean of the X's, the marginal effect at the mean (MEM-X) or at (ii) each observation and take the average, the average marginal effect (AME). The former is what Stata's *mfx* command produces while the latter is what the *margeff* procedure due to Bartus (2005) does. AME takes longer to calculate but is arguably more intuitively appealing.

Anderson & Newell (2003) present an alternative solution to this problem. They note that if all the X variables are normalized to have a mean of 0, the scale factor depends only on the estimated constant. By simply plugging the constant into the appropriate density function, an estimated scale factor is generated[1]. Whether this method is of much practical use is unclear: since the constant can be any real number a large table of possible values must be consulted. Moreover in practice many researchers will not wish to normalize their X variables. For example if one wishes to estimate models on different sub-populations one would need to normalize the data for each model. Alternatively if the estimation sample changed because one added a variable with a different pattern of missing values from the existing variables then, again, re-normalization would be required.

A simpler solution which may be more convenient is to evaluate at the mean of the dependent variable ("MEM-Y").

$$\bar{Y} \equiv F(\overline{bX}\ ) \tag{3}$$

That is $\overline{bX}$ is the scalar value of the single index that would generate the sample mean. Inverting (3) and substituting into (2) implies

$$\frac{\partial E(Y\,/\,X)}{\partial x_1} \approx f(F^{-1}(\bar{Y}))b_1 \tag{4}$$

Hence to calculate marginal effects one multiplies the coefficient by a scalar which is a function only of the mean of the dependent variable. The function is symmetric around 0.5. The values of the scale factor for various values of the mean are given in Table 2 below. So if

---

[1] It is also possible to generate estimated asymptotic variances for the marginal effects with this approach, given an estimate of the variance/covariance matrix of the coefficients but this is probably not its main attraction.

the mean of the dependent variable is between 0.4 and 0.6 then dividing the logit coefficient by 4 or multiplying the probit coefficient by 0.4 should be moderately accurate.

The approach suggested does not generate variances/t ratios for the marginal effects. In practise the t ratios for marginal effects and for the underlying coefficients seldom seem to differ by much. Where they do differ by much, such that one was significantly different from zero and the other was not, interpretation would be somewhat problematic.

3. Applications:

To illustrate the method suggested here I use the "union" dataset provided with Stata and model union membership as function of age, grade, not_smsa, south and southXt [2]. The mean of the dependent variable is .222 so I use the nearest value in Table 2, 0.3 for probit and 0.174 for logit. Table 1 shows the different marginal effects for one variable (south). The first row has the regular probit and logit coefficients. The last row, MEM-Y, gives the estimates with the method suggested in equation (4). For both logit and probit the difference with conventional marginal effects is less than 1% i.e. less than one percentage point, good enough for most purposes. Proportionally the gap is slightly smaller for probit.

In many empirical models of individual behaviour one includes both age and its square and one may be interested in the marginal effect of one year. This will depend on both coefficients, say b_A and b_{AA}, respectively. One could use the following approximation and evaluate at whatever values of age that one is interested in.

$$\frac{\partial E(Y/X)}{\partial Age} \approx f(F^{-1}(\bar{Y}))(b_A + 2b_{AA}Age) \qquad (5)$$

---

[2] Within Stata type "webuse union" followed by "probit union age grade not_smsa south southXt" for the probit model.

4. Extension to count data models

It is straightforward to extend this approach to count data models such as Poisson and Negative binomial since the conditional mean function is of the form:

$$E(Y/X) = e^{X\beta} \tag{6}$$

In this case the marginal effect is

$$\frac{\partial(E(Y/X))}{\partial x_1} = \beta_1 e^{X\beta} \tag{7}$$

Approximating this at the mean of Y yields

$$\frac{\partial(E(Y/X))}{\partial x_1} \approx \beta_1 \bar{Y} \tag{8}$$

Alternatively, one could evaluate at any other point of the distribution of Y that one is interested in. Note that if one is interested in evaluating the elasticity then one can evaluate at either the mean of the relevant X variable or at some other value that is of interest:

$$\frac{\partial(E(Y/x))}{\partial x_1} \frac{\bar{x_1}}{(E(Y/x))} \approx \beta_1 \bar{x_1} \tag{9}$$

5. Concluding remarks

This note provides a simple, back-of-the-envelope, method to estimate marginal effects in several popular limited dependent variable models. The approximation suggested here will not coincide with either the conventional "marginal effects at the mean" or the "average marginal effects" because of the non-linear functional form. They are different parameters but will probably be quite close in practice. The approach suggested here focuses attention on marginal effects at different points in the distribution of the dependent variable rather than representative points in the joint distribution of the explanatory variables.

Table 1: Marginal effects with different methods: an example with P=.222

	Probit	Logit
Coefficient	-.4036	-.7144
MEM-X	-.1145	-.1162
AME	-.1143	-.1162
MEM-Y	-.1208	-.1243

Table 2: Scale factors as a function of the mean of the dependent variable

P	1-P	logit	probit
0.010	0.990	0.010	0.027
0.025	0.975	0.024	0.058
0.050	0.950	0.048	0.103
0.075	0.925	0.069	0.142
0.100	0.900	0.090	0.175
0.125	0.875	0.109	0.206
0.150	0.850	0.128	0.233
0.175	0.825	0.144	0.258
0.200	0.800	0.160	0.280
0.225	0.775	0.174	0.300
0.250	0.750	0.188	0.318
0.275	0.725	0.199	0.334
0.300	0.700	0.210	0.348
0.325	0.675	0.219	0.360
0.350	0.650	0.228	0.370
0.375	0.625	0.234	0.379
0.400	0.600	0.240	0.386
0.425	0.575	0.244	0.392
0.450	0.550	0.248	0.396
0.475	0.525	0.249	0.398
0.500	0.500	0.250	0.399

References

Anderson, S. & R.G. Newell (2003) Simplified marginal effects in discrete choice models *Economics Letters*, 81, 323-326

Bartus, T. (2005) Estimation of marginal effects using margeff, *Stata Journal*, 5(3), 309-329