

# Probit

Econ 350: The University of Chicago

This Draft: December 13, 2013

**Exercise 1.1** (*Identification of the Probit Model*) Model a bivariate, static, discrete choice through a Probit model. The convention is that in this model the unobserved variable is normally distributed. (i) Show that you can normalize the threshold that defines the agent's decision without loss of generality; (ii) why is the normalization without loss of generality?; (iii) what other normalization can you make without loss of generality in your Probit model?; (iv) what does this normalization implies with respect to the distribution of the unobserved variable? (v) are you able to identify all the parameters of the model?; (vi) how does identification and the normalizations relate to each other? Hint: think of the scalar and spatial identification issues that the structure of a Probit model generates.

Answer:

This model is static so the answer neglects the time subscript. Individual  $i = 1, \dots, N$  has two choices,  $\{0, 1\}$ , that imply two potential outcomes,  $y_i^0, y_i^1$ . These outcomes depend linearly on observed and unobserved variables (from the perspective of the econometrician):

$$\begin{aligned} y_i^0 &= \mathbf{X}_i^0 \beta_0 + \epsilon_0 \\ y_i^1 &= \mathbf{X}_i^1 \beta_1 + \epsilon_1 \end{aligned} \tag{1}$$

where  $X_i^j$  is a  $1 \times k_j$  vector of independent variables for  $j = 0, 1 \forall i = 1, \dots, N$ ,  $\beta_j$  is a  $k_j \times 1$  vector of coefficients and  $\epsilon_j$ ,  $j = 0, 1$  is a random perturbation, which the econometrician ignores. In the Probit model, the assumptions is that these perturbations have a joint normal distribution with mean zero. Individual  $i$  chooses 1 iff  $y_i^1 - y_i^0 \geq a$ , for  $a \in \mathbb{R}$ . The econometrician only observes whether the individual chooses 0 or 1. It is possible, then, to model the choice rule as follows:

$$d_i = \begin{cases} 1 & \text{if } y_i^* \geq a \\ 0 & \text{if } y_i^* < a \end{cases} \tag{2}$$

where  $y_i^* \equiv (\mathbf{X}_i^1 \beta_1 - \mathbf{X}_i^0 \beta_0) + (\epsilon_1 - \epsilon_0) \equiv \mathbf{X}_i \beta + \epsilon_i$ .  $\mathbf{X}_i, \beta$  are of dimensions  $1 \times k_1 + k_0$  and  $k_1 + k_0 \times 1$ , respectively. By construction,  $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ , where  $\sigma^2$  is a function of the variances and the covariance of  $\epsilon_0, \epsilon_1$ .

A simple way to asses identification is the following. The principal ingredients of the model's identification are the choice probabilities. This actually have the same components so we focus on the following

$$\begin{aligned} \Pr(d_i = 1 | X_i) &= \Pr(y_i^* \geq a | X_i) \\ &= \Pr(\mathbf{X}_i \beta + \epsilon_i \geq a | X_i) \\ &= \Pr(\epsilon_i \geq a - \mathbf{X}_i \beta) \\ &= \Pr\left(\frac{\epsilon_i}{\sigma} \geq \frac{a - \mathbf{X}_i \beta}{\sigma}\right). \end{aligned} \tag{3}$$

As is standard, the first element in  $X_i$  is a constant term across individuals. Thus, the first element in  $\beta$  cannot be identified from  $a$ , i.e. spatial identification of  $a$  from the first element of  $\beta$  is impossible. The identification impossibility implies that the assumption  $a = 0$  does not imply loss of generality. Likewise, scalar identification of each element of  $\beta$  from  $\sigma$  is impossible. The usual assumption here is  $\sigma^2 = 1$ , which implies that  $\epsilon_i$  has a normal standard distribution.  $\sigma^2 = 1$  is as arbitrary as any other assumption so standard normality of  $\epsilon_i$  is as arbitrary as any other distribution. The identification of the parameters of the Probit, thus, is up to scale. A scale set to 1 typically. The identification of the rest of the parameters relies on a set of  $k_1 + k_0$  first order conditions. In particular, provided that the matrix  $X = [X_1, \dots, X_N]$ , the first order conditions are a set of  $k_1 + k_0$  that solve for the  $k_1 + k_0$  parameters in  $\beta$ .

**Exercise 1.3** (*Estimation of the Probit Model*) From Exercise 1.1 you have clear the setup of the Probit model. Make sure you know what the correct parametric assumptions are and what can you identify in the model. Simulate all the data necessary to estimate a Probit model. The instructions are loose in purpose because we want to evaluate your ability to create the data and estimate the model from scratch. Hint: simulate a single independent variable. This is sufficient for the purposes of this exercise.

Answer: The likelihood of individual  $i$  is:  $L_i(\beta|X_i) = \Phi(\mathbf{X}_i\beta)^{d_i} \cdot \Phi(-\mathbf{X}_i\beta)^{1-d_i}$ . The sample log likelihood is:  $l(\beta|X) = \sum_{i=1}^N d_i \cdot \log \Phi(\mathbf{X}_i\beta) + (1 - d_i) \log \Phi(-\mathbf{X}_i\beta)$ . For the rest of the exercise see the code “probit.py”.