TA session #5
ECON 342

J. Marcelo Ochoa

March 3, 2009

Problem 1. Consider the Logit model with $\mathbb{P}(y_i = 1|x_i) = \Lambda(\beta_0 + \beta_1 x_i)$, where the scalar random variable x_i is binary, taking values of 0 and 1.

(a) Show that this model can be written as a linear probability model $\mathbb{P}(y_i = 1|x_i) = \gamma_0 + \gamma_1 x_i$ and derive γ_0 and γ_1 as functions of β_0 and β_1

(b) Show that the MLE of γ_0 and γ_1 is equal to the least squares estimate of y_i on x_i with intercept.

Solution. Note that:

$$\mathbb{P}(y_i = 1|x_i = 0) = \frac{\exp[\beta_0]}{1 + \exp[\beta_0]}$$

and,

$$\mathbb{P}(y_i = 1|x_i = 1) = \frac{\exp[\beta_0 + \beta_1]}{1 + \exp[\beta_0 + \beta_1]}$$

Then,

$$\mathbb{P}(y_i = 1|x_i) = \gamma_0 + \gamma_1 x_i$$

with,

$$\gamma_0 = \frac{\exp[\beta_0]}{1 + \exp[\beta_0]}; \qquad \gamma_1 = \frac{\exp[\beta_0 + \beta_1]}{1 + \exp[\beta_0 + \beta_1]} - \frac{\exp[\beta_0]}{1 + \exp[\beta_0]}$$

The least squares estimate of y_i on x_i and a constant solves,

$$\min_{b_0, b_1} \sum_{i=1}^{N} \left(y_i - b_0 - b_1 x_i\right)^2$$

Let n_{kl} be number of observations that record a pair $(y_i = k, x_i = l)$, e.g., there are n_{11} observations with $(y_i = 1, x_i = 1)$. Then, the sum of squared residuals can be written as,

$$n_{11}(1 - b_0 - b_1)^2 + n_{10}(1 - b_0)^2 + n_{01}(b_0 + b_1)^2 + n_{00}(b_0)^2$$

The first-order conditions for this problem are,

$$-2n_{11}(1-\hat{b}_0-\hat{b}_1)-2n_{10}(1-\hat{b}_0)+2n_{01}(\hat{b}_0+\hat{b}_1)+2n_{00}\hat{b}_0=0 \tag{1}$$
$$-2n_{11}(1-\hat{b}_0-\hat{b}_1)+2n_{01}(\hat{b}_0+\hat{b}_1)=0 \tag{2}$$

From (2) we have:

$$\hat{b}_0+\hat{b}_1=\frac{n_{11}}{n_{11}+n_{01}}$$

which replaced into (1) we have:

$$\underbrace{(n_{01}+n_{11})(\hat{b}_0+\hat{b}_1)}_{n_{11}}+(n_{10}+n_{00})\hat{b}_0=n_{11}+n_{10}\Rightarrow\hat{b}_0=\frac{n_{10}}{n_{10}+n_{00}}$$

Then, the LS estimates equal,

$$\begin{aligned}\hat{b}_0&=\frac{n_{10}}{n_{10}+n_{00}}\\ \hat{b}_1&=\frac{n_{11}}{n_{11}+n_{01}}-\frac{n_{10}}{n_{10}+n_{00}}\end{aligned}$$

The ML estimate of (γ_0,γ_1) solves,

$$\max\sum_{i=1}^{N}\left[y_i\ln\mathbb{P}\left(y_i=1|x_i\right)+(1-y_i)\ln\mathbb{P}\left(y_i=0|x_i\right)\right]$$

Using the notation above we can re-write the log-likelihood as,

$$n_{11}\ln(\gamma_0+\gamma_1)+n_{10}\ln\gamma_0+n_{01}\ln(1-\gamma_0-\gamma_1)+n_{00}\ln(1-\gamma_0)$$

The first-order conditions yield,

$$\begin{aligned}&\frac{n_{11}}{\hat{\gamma}_0+\hat{\gamma}_1}+\frac{n_{10}}{\hat{\gamma}_0}-\frac{n_{01}}{1-\hat{\gamma}_0-\hat{\gamma}_1}-\frac{n_{00}}{1-\hat{\gamma}_0}=0\\ &\frac{n_{11}}{\hat{\gamma}_0+\hat{\gamma}_1}-\frac{n_{01}}{1-\hat{\gamma}_0-\hat{\gamma}_1}=0\end{aligned}$$

then,

$$\begin{aligned}\hat{\gamma}_0&=\frac{n_{10}}{n_{10}+n_{00}}\\ &\frac{n_{11}}{n_{11}+n_{01}}-\frac{n_{10}}{n_{10}+n_{00}}\end{aligned}$$

Thus, we can conclude that:

$$\hat{\gamma}_0=\hat{b}_0\ \text{ and }\hat{\gamma}_1=\hat{b}_1$$

as desired. $\Diamond$

Problem 2. Let t_i^* denote the duration of some event, such as the time to graduate from Grad School, measured in continuous time. Consider the following model for t_i^*,

$$\begin{aligned} t_i^* &= \exp[\mathbf{x}_i\beta + u_i] \\ t_i &= \min(t_i^*, c) \end{aligned}$$

where $c > 0$ is a known censoring constant, and $u_i \sim N(0, \sigma^2)$.

(a) Find $\mathbb{P}(t_i = c|\mathbf{x}_i)$.

(b) Write the full density of $\ln(t_i)$ given $\mathbf{x}_i$ and define the MLE estimator.

(c) Obtain the log-likelihood function if the censoring time is potentially different for each person, so that

$$t_i = \min(t_i^*, c_i)$$

where c_i is observed for each i. Assume that u_i is independent of $(\mathbf{x}_i, c_i)$.

Solution. By definition:

$$\mathbb{P}(t_i = c|\mathbf{x}_i) = \mathbb{P}(t_i^* \geq c|\mathbf{x}_i)$$

Which is equivalent to say,

$$\mathbb{P}(\ln t_i^* \geq \ln c|\mathbf{x}_i)$$

Since

$$\ln t_i^* = \mathbf{x}_i\beta + u_i$$

then,

$$\begin{aligned} \mathbb{P}(\mathbf{x}_i\beta + u_i \geq \ln c|\mathbf{x}_i) =& \mathbb{P}(u_i \geq \ln c - \mathbf{x}_i\beta|\mathbf{x}_i) \\ =& \mathbb{P}\left(\frac{u_i}{\sigma} \geq \frac{\ln c - \mathbf{x}_i\beta}{\sigma}|\mathbf{x}_i\right) \\ =& 1 - \Phi\left(\frac{\ln c - \mathbf{x}_i\beta}{\sigma}\right) \qquad \left(\text{since } \tfrac{u_i}{\sigma} \sim N(0,1)\right) \end{aligned}$$

When $\ln t_i^* < c$ then the density of $\ln t_i < \ln c$ is the same as that of $\ln t_i^*$, then the pdf of this $\ln t_i$ is:

$$f(\ln t_i|\mathbf{x}_i) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(t_i - \mathbf{x}_i\beta)^2}{2\sigma^2}\right], & \text{if } ; \ln t_i < \ln c \\ 1 - \Phi\left(\frac{\ln c - \mathbf{x}_i\beta}{\sigma}\right), & \text{if } ; \ln t_i = \ln c \end{cases}$$

Then, the likelihood function can be written as:

$$L(\beta) = \sum_{\ln t_i = \ln c} \ln\left[1 - \Phi\left(\frac{\ln c - \mathbf{x}_i\beta}{\sigma}\right)\right] + \sum_{\ln t_i < \ln c} \ln \frac{1}{\sigma}\phi\left(\frac{\ln t_i - \mathbf{x}_i\beta}{\sigma}\right)$$

If the censoring parameter is different for each individual, we observe it, and we assume that it is independent, then the likelihood function is:

$$L(\beta) = \sum_{\ln t_i = \ln c_i} \ln \left[1 - \Phi\left(\frac{\ln c_i - \mathbf{x}_i \beta}{\sigma} \right) \right] + \sum_{\ln t_i < \ln c_i} \ln \frac{1}{\sigma} \phi \left(\frac{\ln t_i - \mathbf{x}_i \beta}{\sigma} \right)$$

◊