Practice Problem for Limited Dependent Variables

Francis J. DiTraglia

May 9, 2021

The following is based on a question from the 2020 "PhD Application Exam."

1. Let $(y_1, x_1), \cdots, (y_N, x_N)$ be a collection of iid observations where $y_i \in \{0, 1\}$ and x_i is continuously distributed. Suppose that $p(x_i) \equiv \mathbb{P}(y_i = 1|x_i) = F(\alpha + \beta x_i)$ where $F(z) = e^z/(1 + e^z)$ and (α, β) are unknown parameters.

   (a) Derive an expression for the partial effect of x_i on $p(x_i)$ in this model.

   Solution: We have

   $$\frac{d}{dx}p(x) = \frac{\partial}{\partial x}F(\alpha + \beta x) = F'(\alpha + \beta x)\beta$$

   so all that remains is to calculate F'. By the quotient rule,

   $$F'(z) = \frac{d}{dz}\left(\frac{e^z}{1 + e^z}\right) = \frac{e^z(1 + e^z) - e^z e^z}{(1 + e^z)^2} = \frac{e^z}{(1 + e^z)^2}$$

   Therefore,

   $$\frac{d}{dx}p(x) = \left\{\frac{\exp(\alpha + \beta x)}{[1 + \exp(\alpha + \beta x)]^2}\right\}\beta$$

   (b) Write out the log-likelihood function $\ell_N(\alpha, \beta)$ for this model, simplifying your result as far as possible.

   Solution: The likelihood of a single observation is given by

   $$L_i(\alpha, \beta) = f(y_i|x_i, \alpha, \beta) = F(\alpha + \beta x_i)^{y_i}\left[1 - F(\alpha + \beta x_i)\right]^{1 - y_i}$$

   and the corresponding log-likelihood is

   $$\ell_i(\alpha, \beta) = \log L_i(\alpha, \beta) = y_i \log\left[F(\alpha + \beta x_i)\right] + (1 - y_i)\log\left[1 - F(\alpha + \beta x_i)\right].$$

   Substituting the definition of F and simplifying, we obtain

   $$\begin{aligned}\ell_i(\alpha, \beta) &= y_i \log\left[\frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)}\right] + (1 - y_i)\log\left[1 - \frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)}\right] \\ &= y_i(\alpha + \beta x_i) - y_i \log\left[1 + \exp(\alpha + \beta x_i)\right] + (1 - y_i)\log(1) \\ &\quad - (1 - y_i)\log\left[1 + \exp(\alpha + \beta x_i)\right] \\ &= y_i(\alpha + \beta x_i) - \log\left[1 + \exp(\alpha + \beta x_i)\right]\end{aligned}$$

Because our observations are iid, the log-likelihood function equals the sum of the likelihoods of each observation. Hence,

$$\ell_N(\alpha, \beta) = \sum_{i=1}^{N} \{y_i(\alpha + \beta x_i) - \log[1 + \exp(\alpha + \beta x_i)]\}$$

(c) Using your answer to the preceding part, derive the first-order conditions for the maximum likelihood estimators of α and β. Simplify your results as far as possible.

Solution: Differentiating,

$$\begin{aligned}\frac{\partial \ell_N}{\partial \alpha} &= \sum_{i=1}^{N} \frac{\partial}{\partial \alpha} \ell_i(\alpha, \beta) = \sum_{i=1}^{N} \frac{\partial}{\partial \alpha} \{y_i(\alpha + \beta x_i) - \log[1 + \exp(\alpha + \beta x_i)]\} \\ &= \sum_{i=1}^{N} \left[y_i - \frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)} \right] = \sum_{i=1}^{N} [y_i - F(\alpha + \beta x_i)]\end{aligned}$$

and similarly

$$\begin{aligned}\frac{\partial \ell_N}{\partial \beta} &= \sum_{i=1}^{N} \frac{\partial}{\partial \beta} \ell_i(\alpha, \beta) = \sum_{i=1}^{N} \frac{\partial}{\partial \beta} \{y_i(\alpha + \beta x_i) - \log[1 + \exp(\alpha + \beta x_i)]\} \\ &= \sum_{i=1}^{N} \left[y_i x_i - \frac{\exp(\alpha + \beta x_i) x_i}{1 + \exp(\alpha + \beta x_i)} \right] = \sum_{i=1}^{N} [y_i - F(\alpha + \beta x_i)]\, x_i\end{aligned}$$

Therefore, the first-order conditions are

$$\sum_{i=1}^{N} \left[y_i - F(\widehat{\alpha} + \widehat{\beta} x_i) \right] \begin{bmatrix} 1 \\ x_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$