Suggested Answers
Problem Set 7
Economics 60303

Bill Evans
Spring 2013

1. a) The decision rule is that y_i=1 if $c_i \geq c^*$ where $c^* \sim N[\mu, \sigma^2]$. Note that

$\Pr[y_i = 1] = \Pr[c_i \geq c^*] = \Pr[(c^* - \mu)/\sigma = z \leq (c_i - \mu)/\sigma] = \Phi[(c - \mu)/\sigma]$ which can be reduced to read $\Pr[y_i = 1] = \Phi[-(\mu/\sigma) + (c_i/\sigma)]$. Let $\beta_1 = 1/\sigma$ and $\beta_0 = -\mu/\sigma$ then we can rewrote the decision rule as $\Pr[y_i = 1] = \Phi[\beta_0 + \beta_1 c_i]$. Notice that $\hat{\beta}_1$ should be >0 and hence $\partial \Pr[y_i = 1]/\partial c_i > 0$ -- as the cholesterol level increases the probability of getting a prescription increases. The log likelihood function for observation i is then

$$L_i = \Phi[\beta_0 + \beta_1 CLS_i]^{y_i} (1 - \Phi[\beta_0 + \beta_1 CLS_i])^{1-y_i}$$

b) Let β_0 vary across people so it now equals $\beta_0^i = -\mu_i/\sigma = -x_i\beta/\sigma$

2. a.

$$\Pr(y_i > a) = \int_a^\infty \alpha \left(\frac{y_m^\alpha}{y^{\alpha+1}}\right) dy = \int_a^\infty \alpha y_m^\alpha y^{-\alpha-1} dy = \left(\frac{\alpha}{\alpha}\right) y_m^\alpha y^{-\alpha} \Big|_a^\infty = y_m^\alpha y^{-\infty}$$

$$E[y_i] = \int_a^\infty \alpha y \left(\frac{y_m^\alpha}{y^{\alpha+1}}\right) dy = \int_a^\infty \alpha y_m^\alpha y^{-\alpha} dy = \left(\frac{\alpha}{-\alpha+1}\right) y_m^\alpha y^{-\alpha-1} \Big|_a^\infty =$$

b.

$$\left(\frac{\alpha}{1-\alpha}\right) y_m^\alpha y^{-\alpha+1} \Big|_a^\infty = \left(\frac{\alpha}{1-\alpha}\right)\left(\left[y_m^\alpha \infty^{-\alpha+1}\right] - \left[y_m^\alpha y_m^{\ -\alpha+1}\right]\right) = \left(\frac{y_m \alpha}{1-\alpha}\right)$$

$$\ell = \sum_i \ln\left[\alpha\left(\frac{y_m^\alpha}{y^{\alpha+1}}\right)\right] = \sum_i [\ln(\alpha) + \alpha[\ln(y_m^\alpha) - (\alpha+1)\ln(y_i)]$$

c.

$$= n\ln(\alpha) + n\alpha\ln(y_m^\alpha) - (\alpha+1)\sum_i \ln(y_i)$$

b. $\dfrac{\partial \ell}{\partial \alpha} = \dfrac{n}{\alpha} + n\ln(y_m^\alpha) - \sum_i \ln(y_i) = 0$ which implied that $\hat{\alpha} = \dfrac{n}{\ln(y_m^\alpha) - \bar{y}}$

3. $\ell = N_{11}\ln(P_1) + N_{01}\ln(1-P_1) + N_{10}\ln(P_0) + N_{00}\ln(1-P_0)$

The first order conditions are

$$\frac{\partial \ell}{\partial P_1} = \frac{N_{11}}{P_1} - \frac{N_{01}}{1-P_1} = 0 \text{ which implies that } \hat{P}_1 = \frac{N_{11}}{N_{11}+N_{01}}$$

$$\frac{\partial \ell}{\partial P_0} = \frac{N_{11}}{P_0} - \frac{N_{00}}{1-P_0} = 0 \text{ which implied that } \hat{P}_0 = \frac{N_{10}}{N_{10}+N_{00}}$$

Note that $\Phi(\beta_0 + \beta_1) - \Phi(\beta_0) = \hat{P}_1 - \hat{P}_0$. Looking ahead to the next problem, note that we can write $\hat{P}_1 = \frac{N_{11}}{N_{11}+N_{01}}$ as $\bar{y}_1 = \frac{\sum_i x_i y_i}{\sum_i x_i}$ and $\hat{P}_0 = \frac{N_{10}}{N_{10}+N_{00}}$ as $\bar{y}_0 = \frac{\sum_i (1-x_i) y_i}{\sum_i (1-x_i)}$

4. Continue with problem 3. The 2 x 2 table below provides the number of observations in each pair of (X,Y) combinations:

$$\hat{P}_1 = \frac{N_{11}}{N_{11}+N_{01}} = \frac{30}{30+15} = 0.667 \text{ and } \hat{P}_0 = \frac{N_{10}}{N_{10}+N_{00}} = \frac{20}{20+35} = 0.364$$

$\Phi(\beta_o) = 0.364$ which implies that $\Phi^{-1}(0.364) = \hat{\beta}_0 = -0.348$

$\Phi(\beta_o + \beta_1) = 0.667$ which implies that $\Phi^{-1}(0.667) = (\hat{\beta}_o + \hat{\beta}_1) = 0.432$ and
$\hat{\beta}_1 = 0.432 - \hat{\beta}_0 = 0.432 - -0.348 = 0.78$

5. In a previous problem set we showed that given a linear regression $y_i = \beta_0 + x_i\beta_1 + \varepsilon_i$ were x_i is dichotomous, then the estimate for β_1 is

$$\hat{\beta}_1 = \bar{y}_1 - \bar{y}_o = \hat{P}_1 = \frac{N_{11}}{N_{11}+N_{01}} - \hat{P}_0 = \frac{N_{10}}{N_{10}+N_{00}}$$

6. Note that from problem 3, regardless of the distribution of the error in the discrete choice model $\hat{P}_1 = \frac{N_{11}}{N_{11}+N_{01}}$ and $\hat{P}_0 = \frac{N_{10}}{N_{10}+N_{00}}$. Therefore, $\hat{P}_0 = \frac{e^{\beta_0}}{1+e^{\beta_0}}$ and $\hat{\beta}_0 = \ln(\hat{p}_0 / (1-\hat{p}_0))$ and $\hat{\beta}_1 = \ln(\hat{p}_1 / (1-\hat{p}_1)) - \hat{\beta}_0$.

7.

OLS coefficient or marginal effect on worka variable (standard error)		
OLS (Linear probability)	Marginal effect on worka from probit	Marginal effect on worka from logit
-0.0838 (0.0073)	-0.0838 (0.0076)	-0.0838 (0.0076)

8. $\ell = N_{11}\ln(P_1) + N_{01}\ln(1-P_1) + N_{10}\ln(P_0) + N_{00}\ln(1-P_0)$

$$\ell = N_{11}\ln[\Phi(\beta_0+\beta_1)] + N_{01}\ln[1-\Phi(\beta_0+\beta_1)] + N_{10}\ln[\Phi(\beta_0)] + N_{00}\ln[1-\Phi(\beta_0)]$$

In this case, N_{01}=0. There is no mother with a twin on the 1st birth without a second child. Therefore, the log likelihood reduces to

$$\ell = N_{11}\ln[\Phi(\beta_0+\beta_1)] + N_{10}\ln[\Phi(\beta_0)] + N_{00}\ln[1-\Phi(\beta_0)]$$

Taking the first order conditions

$$\frac{\partial\ell}{\partial\beta_0} = \frac{N_{11}\phi(\beta_0+\beta_1)}{\Phi(\beta_0+\beta_1)} + \frac{N_{10}\phi(\beta_0)}{\Phi(\beta_0)} - \frac{N_{00}\phi(\beta_0)}{1-\Phi(\beta_0)} = 0$$

$$\frac{\partial\ell}{\partial\beta_1} = \frac{N_{11}\phi(\beta_0+\beta_1)}{\Phi(\beta_0+\beta_1)} = 0$$

Notice that $\frac{\partial\ell}{\partial\beta_1}$ will only equal zero if $\hat{\beta}_1 \to \infty$ so functionally, the parameter is undefined -- the numerator will approach zero and the denominator will approach 1. In general, if the (2x2) box of (y x x) has a missing off-diagonal element, then $\hat{\beta}_1 \to \infty$ or $\hat{\beta}_1 \to -\infty$ depending on what element of the box is zero. In this case, the estimate is undefined.