Alastair Hall | **ECON61001: Econometric Methods** | Semester 1 2020-21

Solutions to Problem Set for Tutorial 9

1.(a) Since y_i only takes two values $\{0,1\}$, it follows that given x_i, u_i can only take two values $1 - x_i'\beta_0$ and $-x_i'\beta_0$ and

$$\begin{aligned} P(u_i = 1 - x_i'\beta_0 \,|\, x_i) &= P(y_i = 1 \,|\, x_i) \;=\; x_i'\beta_0, \\ P(u_i = -x_i'\beta_0 \,|\, x_i) &= P(y_i = 0 \,|\, x_i) \;=\; 1 - x_i'\beta_0. \end{aligned}$$

Therefore, we have

$$E[u_i \,|\, x_i] \;=\; (1 - x_i'\beta_0)x_i'\beta_0 + (-x_i'\beta_0)(1 - x_i'\beta_0) \;=\; 0.$$

1.(b) Using part (a), we have $Var[u_i \,|\, x_i] = E[u_i^2 \,|\, x_i]$ where

$$E[u_i^2 \,|\, x_i] \;=\; (1 - x_i'\beta_0)^2(x_i'\beta_0) + (-x_i'\beta_0)^2(1 - x_i'\beta_0) \;=\; (x_i'\beta_0)(1 - x_i'\beta_0).$$

1.(c) As shown in the answer to part (a), u_i is a discrete random variable conditional on x_i, and so does not have a normal distribution.

2.(a) Since this is binary response model, there are only two outcomes for y_i, namely $\{0,1\}$. We are given that $P(y_i = 1|x_i) = \Lambda(x_i'\beta_0)$ and so $P(y_i = 0|x_i) = 1 - \Lambda(x_i'\beta_0)$. Recall that the conditional LF is:

$$CLF_N(\beta) \;=\; \prod_{i=1}^{N} p(y_i|x_i;\beta)$$

where $p(y_i|x_i)$ is the conditional probability function for y_i given x_i evaluated at the sample values. We have:

$$\begin{aligned} p(y_i|x_i;\beta) &= \Lambda(x_i'\beta) \qquad \text{if } y_i = 1 \\ &= 1 - \Lambda(x_i'\beta) \qquad \text{if } y_i = 0 \end{aligned}$$

Note that we can write:

$$p(y_i|x_i;\beta) \;=\; \{\Lambda(x_i'\beta)\}^{y_i} \, \{1 - \Lambda(x_i'\beta)\}^{1-y_i},$$

and so the likelihood function is:

$$CLF(\beta) \;=\; \prod_{i=1}^{N} \left(\{\Lambda(x_i'\beta)\}^{y_i} \, \{1 - \Lambda(x_i'\beta)\}^{1-y_i} \right).$$

(b) The (conditional) log likelihood function is $CLLF_N(\beta) = ln[CLF_N(\beta)]$ which, using part (a), is

$$CLLF_N(\beta) \;=\; \sum_{i=1}^{N} \{\, y_i ln[\Lambda(x_i'\beta)] \,+\, (1-y_i)ln[1 \,-\, \Lambda(x_i'\beta)\,\}\,.$$

(c) Since $\partial e^z/\partial z = e^z$, it follows that:

$$\frac{\partial\Lambda(z)}{\partial z} \;=\; \Lambda(z) \,-\, \{\Lambda(z)\}^2 \;=\; \Lambda(z)\{1 \,-\, \Lambda(z)\}$$

If we now set $z = x_i'\beta$ and use the chain rule then

$$\frac{\partial\Lambda(z)}{\partial x_{i,\ell}} \;=\; \left(\frac{\partial\Lambda(z)}{\partial z}\right)\frac{\partial z}{\partial x_{i,\ell}}.$$

Since $\partial z/\partial x_{i,\ell} = \beta_\ell$, we obtain:

$$\frac{\partial P(y_i = 1|x_i;\beta)}{\partial x_{i,\ell}} \;=\; \Lambda(x_i'\beta)\{1 \,-\, \Lambda(x_i'\beta)\}\beta_\ell.$$

3.(a) Using $y \sim N(X\beta_0, \sigma_0^2 I_T)$ and Definition 2.3 in the Lecture Notes (Section 2.1), if follows that the pdf of y is given by:

$$f_y(y;\,\theta) \;=\; (2\pi\sigma^2)^{-T/2} exp\{-(y-X\beta)'(y-X\beta)/2\sigma^2\}.$$

Taking logs, we obtain:

$$LLF_T(\theta) = -\frac{T}{2}ln[2\pi] - \frac{T}{2}ln[\sigma^2] - \frac{(y-X\beta)'(y-X\beta)}{2\sigma^2}.$$

(b) Recall that the score equations are:

$$\frac{\partial LLF_T(\theta)}{\partial\theta}\Big|_{\theta=\hat{\theta}_T} = 0$$

where $\hat{\theta}_T$ denotes the MLE of θ_0. Differentiating the LLF (using the results in Lemma 2.2 in Section 2.2. of the Lecture Notes) we obtain:

$$\frac{\partial LLF_T(\theta)}{\partial\theta} \;=\; \begin{bmatrix} \frac{\partial LLF_T(\theta)}{\partial\beta} \\ \frac{\partial LLF_T(\theta)}{\partial(\sigma^2)} \end{bmatrix} \;=\; \begin{bmatrix} \frac{1}{\sigma^2}X'(y-X\beta) \\ -\frac{T}{2\sigma^2} \,+\, \frac{1}{2\sigma^4}(y-X\beta)'(y-X\beta) \end{bmatrix}.$$

Therefore the score equations imply MLE's satisfy:

$$\begin{bmatrix} X'(y-X\hat{\beta}_T) \\ -T\hat{\sigma}_T^2 \,+\, (y-X\hat{\beta}_T)'(y-X\hat{\beta}_T) \end{bmatrix} \;=\; 0.$$

(c) From $X'(y-X\hat{\beta}_T) = 0$ it follows that $\hat{\beta}_T = (X'X)^{-1}X'y$. From $-T\hat{\sigma}_T^2 \,+\, (y-X\hat{\beta}_T)'(y-X\hat{\beta}_T) = 0$, it follows that $\hat{\sigma}_T^2 = (y-X\hat{\beta}_T)'(y-X\hat{\beta}_T)/T$.

(d) It can be recognized that the OLS and MLE estimators of β_0 are the same, but that the OLS and MLE estimators of σ_0^2 are different. Notice that as the OLS estimator of σ_0^2 is unbiased, it must follow that the MLE of σ_0^2 is actually a biased estimator (although it is consistent).