Problem Set 9

Nicholas Wu

Spring 2021

Problem 1

(a) We know

$$P(T_i = 0|z_i) = Pr(z_i'\gamma + e_{0i} \leq 0|z_i)$$
$$= Pr(e_{0i} \leq -z_i'\gamma|z_i)$$
$$= \Phi(-z_i'\gamma)$$

Also

$$f(y_i, T_i = 1|x_i, z_i) = P(T_i = 1|y_i^*, x_i, z_i)f(y_i^*|x_i, z_i)$$
$$f(y_i, T_i = 1|x_i, z_i) = P(e_{0i} > -z_i'\gamma|e_{1i} = y_i^* - \beta x_i, x_i, z_i)\frac{1}{\sigma}\phi\left(\frac{y_i^* - \beta x_i}{\sigma}\right)$$
$$f(y_i, T_i = 1|x_i, z_i) = \left(1 - \Phi\left(\frac{-z_i'\gamma - \frac{\rho}{\sigma^2}(y_i^* - \beta x_i)}{\sqrt{1 - (\rho/\sigma)^2}}\right)\right)\frac{1}{\sigma}\phi\left(\frac{y_i^* - \beta x_i}{\sigma}\right)$$
$$= \Phi\left(\frac{z_i'\gamma + \frac{\rho}{\sigma^2}(y_i^* - \beta x_i)}{\sqrt{1 - (\rho/\sigma)^2}}\right)\frac{1}{\sigma}\phi\left(\frac{y_i^* - \beta x_i}{\sigma}\right)$$

So the likelihood function is

$$L = \left(\prod_{i|T_i=0}\Phi(-z_i'\gamma)\right)\left(\prod_{i|T_i=1}\Phi\left(\frac{z_i'\gamma + \frac{\rho}{\sigma^2}(y_i^* - \beta x_i)}{\sqrt{1 - (\rho/\sigma)^2}}\right)\frac{1}{\sigma}\phi\left(\frac{y_i^* - \beta x_i}{\sigma}\right)\right)$$

(b) $\hat{\gamma}$ is the probit estimator, so

$$\hat{\gamma} = \arg\max_{\gamma}\frac{1}{N}\sum_i\left(T_i\log(1 - \Phi(-z_i'\gamma)) + (1 - T_i)\log\Phi(-z_i'\gamma)\right)$$

The FOC is

$$\frac{1}{n}\sum_i\left(\frac{-\phi(-z_i'\gamma)(-z_i)T_i}{1 - \Phi(-z_i'\gamma)} + \frac{\phi(-z_i'\gamma)(-z_i)(1 - T_i)}{\Phi(-z_i'\gamma)}\right) = 0$$
$$\frac{1}{n}\sum_i z_i\left(\frac{\phi(-z_i'\gamma)T_i}{1 - \Phi(-z_i'\gamma)} - \frac{\phi(-z_i'\gamma)(1 - T_i)}{\Phi(-z_i'\gamma)}\right) = 0$$

For $\hat{\beta}$ and the coefficient of $\lambda(z_i'\gamma)$ (call it $\hat{c}$), we get

$$\hat{\beta}, \hat{c} = \arg\min_{\beta, c} \frac{1}{N}\sum_i (y_i - \beta' x_i + c\lambda(z_i'\hat{\gamma}))^2$$

these have the FOC:

$$\frac{1}{N}\sum_i 2 \begin{pmatrix} x_i \\ \lambda(z_i'\hat{\gamma}) \end{pmatrix} (y_i - \beta' x_i + c\lambda(z_i'\hat{\gamma})) = 0$$

$$\frac{1}{N}\sum_i \begin{pmatrix} x_i \\ \lambda(z_i'\hat{\gamma}) \end{pmatrix} (y_i - \beta' x_i + c\lambda(z_i'\hat{\gamma})) = 0$$

So altogether:

$$\frac{1}{N}\sum_i \begin{pmatrix} z_i \left(\frac{\phi(-z_i'\gamma)T_i}{1-\Phi(-z_i'\gamma)} - \frac{\phi(-z_i'\gamma)(1-T_i)}{\Phi(-z_i'\gamma)} \right) \\ \begin{pmatrix} x_i \\ \lambda(z_i'\hat{\gamma}) \end{pmatrix} (y_i - \beta' x_i + c\lambda(z_i'\hat{\gamma})) \end{pmatrix} = 0$$

and these are our equations.

Problem 2

Let $e|x$ be distributed with cdf F.

$$Q(\beta) = E[yI[x'\beta \geq 0] + (1-y)I[x'\beta < 0]]$$

$$= E[E[y|x]I[x'\beta \geq 0] + (1 - E[y|x])(1 - I[x'\beta \geq 0])]$$

$$= E[E[y|x]I[x'\beta \geq 0] + 1 - E[y|x] - (1 - E[y|x])I[x'\beta \geq 0])]$$

$$= E[(2E[y|x] - 1)I[x'\beta \geq 0] + 1 - E[y|x]]$$

$$= E[(E[y|x] - E[1-y|x])I[x'\beta \geq 0]] + E[1 - E[y|x]]$$

$$= E[(P(y=1|x) - P(y=0|x))I[x'\beta \geq 0]] + E[1 - E[y|x]]$$

$$= E[(P(e \geq -x'\beta_0|x) - P(e < -x'\beta_0|x))I[x'\beta \geq 0]] + E[1 - E[y|x]]$$

$$= E[(1 - F(-x'\beta_0) - F(-x'\beta_0))I[x'\beta \geq 0]] + E[1 - E[y|x]]$$

$$= E[(1 - 2F(-x'\beta_0))I[x'\beta \geq 0]] + E[1 - E[y|x]]$$

Further, we know that $(1 - 2F(x'\beta_0))I[x'\beta \geq 0] \leq \max(1 - 2F(x'\beta_0), 0)$, so

$$Q(\beta) \leq E[\max(1 - 2F(-x'\beta_0), 0)] + E[1 - E[y|x]]$$

Now, we note that since the conditional median of e given x is 0, $F(0) = 1/2$, so $(1-2F(x'\beta_0))$ is nonnegative iff $x'\beta_0 \geq 0$. So for $\beta = \beta_0$, $(1 - 2F(x'\beta_0))I[x'\beta_0 \geq 0] = \max(1 - 2F(x'\beta_0), 0)$. Therefore,

$$Q(\beta_0) = E[(1 - 2F(-x'\beta_0))I[x'\beta \geq 0]] + E[1 - E[y|x]] = E[\max(1 - 2F(-x'\beta_0), 0)] + E[1 - E[y|x]]$$

Therefore, since $Q(\beta)$ is at most this bound, and $Q(\beta_0)$ equals the bound, we have Q is maximized at β_0.