Monitoria Econometria II

REC 2312 - 1º Semestre 2015

Monitor(a):Victória Mazás Martinez

4ª Lista de exercícios - modelos não lineares

*Lembrando que a resolução é apenas uma síntese simplificada das respostas esperadas.

1. Suponha que você tenha que estimar um modelo no qual sua variável dependente seja uma *dummy* que assume o valor 1 se o indivíduo de 18 anos completou o Ensino Médio e zero caso contrário. Discorra sobre como seria a melhor forma de estimar o modelo, dado que queremos saber o efeito sobre a probabiliade de conclusão do E.M. aos 18 anos. Descreva detalhadamente as etapas necessárias para a estimação e as suposições sobre o modelo.

   Como não se deseja obter probabilidade fora do intervalo [0,1], devemos utilizar o modelo *Probit* ou *Logit*. A diferença entre eles encontra-se apenas na suposição sobre a distribuição do termo de erro; caso acredite-se que ele segue uma distribuição Normal, utiliza-se o modelo Probit, caso acredite-se que ele segue uma distribuição logistica, utiliza-se o *Logit* . Em ambos os casos, assume-se que : $P(y = 1|x) = G(X\beta)$, onde $G(z) = \Phi(z)$, é uma distribuição acumulada de uma normal padrão no caso do *Probit* e $G(z) = exp(z)/[1 + exp(z)]$ no caso do *Logit*. A estimação dos parâmetros é feita pela maximização da função de verossimilhança:

$$L = \prod_{i=1}^{n} f(y|x_i;\beta) = \prod_{i=1}^{n} [G(x_i\beta)]^{yi}[1 - G(x_i\beta)]^{1-y_i}$$

   Que é o mesmo que maximizar a função de log-verossimilhança:

$$L = \sum_{i=1}^{n} \ln f(y|x_i;\beta) = \sum_{i=1}^{n} y_i[G(x_i\beta)] + (1-y_i)\ln[1-G(x_i\beta)]$$

Efetuando-se a maximização ...

2. Discorra sobre a diferença na interpretação dos coeficientes nos modelos de probabilidade Linear e não Linear (Probit ou Logit)

   No modelo de probabilidade linear o coeficiente estimado pode ser interpretado diretamente como a mudança marginal na probabilidade de um evento ocorrer dada uma mudança em x. Contudo, o mesmo não ocorre com os modelos *Logit* e *Probit*. Nestes modelos apenas o sinal do coeficiente pode ser considerado, ou seja, ele dará um indicativo da direção do efeito (observe a forma do efeito marginal para variáveis contínuas na questão abaixo, como g(.) é sempre positivo, β determinará o sinal do efeito marginal. A magnitude do efeito marginal dependerá do valor do vetor x).

3. Apresente a forma dos efeito marginais nos modelos não lineares (obs: eles diferem nos casos contínuo, discreto e para dummys).

   Efeitos Marginais :

   (a) Para o caso de variáveis contínuas:

   $$Pr(y_i = 1|x) = G(\beta_0 + \beta X)$$

   Derivando em x, temos:

   $$\frac{\partial G(.)}{\partial x_i} = g(.)\beta_i$$

   onde g(.) é uma função de densidade de probabiliade (que dependendo da hipótese sobre a distribuição do termo de erro pode ser logistica ou normal).

   (b) para o caso de variável dummy:

   $$G(\hat{\beta}_0 + \hat{\beta}_1) - G(\hat{\beta}_0)$$

(c) Se x_i é discreto :

$$G(\hat{\beta}_0 + \hat{\beta}_1(c+1)) - G(\hat{\beta}_0 + \hat{\beta}_1(c))$$

4. Dado que o efeito marginal em regressões não lineares depende do valor do vetor $\mathbf{x}$, apresente duas formas de calcular uma medida sumarizada do impacto marginal de uma variável explicativa.

   (a) Efeito Marginal na Média:

   $$\frac{\partial Pr}{\partial x_j} = g(\hat{\beta}_0 + \hat{\beta}_1\bar{x}_1 + \hat{\beta}_2\bar{x}_2)\beta_j$$

   Onde o $\bar{x}_i$ representa o valor atribuído ao indivíduorepresentativo que neste caso é o indivíduo médio.

   (b) Média do Efeito Marginal :

   $$E[\partial Pr/\partial x_j]$$

   Por analogia :

   $$1/N\sum_{i=1}^{N} g_i\beta_j$$

5. O efeito parcial de x_j sobre $Pr(\mathbf{x})$ depende de $\mathbf{x}$. Contudo o efeito relativo não depende de $\mathbf{x}$. Mostre que para variáveis contínuasx_j e x_h, a taxa dos efeitos parciais é constante.

$$\frac{\partial Pr/\partial x_j}{\partial Pr/\partial x_h} = g(\mathbf{x}\boldsymbol{\beta})\beta_j/g(\mathbf{x}\boldsymbol{\beta})\beta_h = \frac{\beta_j}{\beta_h}$$

6. Defina *grad* como uma variável *dummy* informando se um estudante- atleta de uma grande universidade se formará em cinco anos. Sejam *nmem* e *sat* a nota média do ensino médio e a nota no exame SAT, respectivamente. Defina *estudo* como o número de horas gastas por semana em uma sala de estudos organizada. Suponha que usando os dados de 420 estudantes-atletas, observa-se o seguinte modelo logit:

$\hat{P}(grad = 1|nmen, sat, estudo) = \Lambda(-1,17 + 0,24nmem + 0,00058sat + 0,073estudo)$,

onde $\Lambda(z) = exp(z)/[1 + exp(z)]$ é a função logit. Mantendo fixos $nmem$ e sat em 3,0 e 1.200, respectivamente, compute a diferença estimada na probabilidade de formatura de alguém que passou dez horas por semana em uma sala de estudos e de alguém que passou cinco horas por semana.

Como horas é uma variável discreta, pelo exercício acima, sabemos que seu efeito será calculado da seguinte maneira

$$G(\hat{\beta}_0 + \hat{\beta}_1(c + k)) - G(\hat{\beta}_0 + \hat{\beta}_1(c)),$$

portanto substituindo os valores dados e calculando, obteremos

$$\Lambda(7.24) - \Lambda(6.875) = 0.000315433$$

7. Apresente a justificativa para o uso de variável latente para estimar um modelo e a estrutura do modelo sob este enfoque.

O enfoque da variável latente trata o uso de variáveis dicotômicas basicamente como um problema de medida, ou seja, existe uma variável latente contínua (Y*) que não observamos diretamente de forma que não podemos mensurá-la. O que nós observamos é um indicador da variável latente.

O modelo sobre este enfoque é descrito da seguinte maneira:

$$Y^* = \beta_0 + \beta_1 X_i + \epsilon_i$$
$$Y_i = 0, se Y_i^* \leq 0,$$
$$Y_i = 1, se Y_i^* > 0,$$

$$P(Y_i = 1) = P(Y_i^* > 0) = P(X_i\beta + \epsilon_i > 0) = P(\epsilon_i \leq X_i\beta)$$