Modelos ordinales en gretl

Microeconomía Cuantitativa

R. Mora

Departmento of Economía
Universidad Carlos III de Madrid

Esquema

Introduction

El Probit ordenado y la estimación MV

- Considera m resultados observados: $y = 0, 1, \ldots, m$.
- Considera un modelo de variable latente sin constante:

$$y^* = x'\beta + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0,1)$$

- Define $m-1$ puntos de corte: $\alpha_1 < \ldots < \alpha_{m-1}$
- No observamos y^*, pero observamos elecciones según la siguiente regla

$$\begin{aligned}
y &= 0 \text{ if } y^* \leq \alpha_1 \\
y &= 1 \text{ if } \alpha_1 < y^* \leq \alpha_2 \\
&\vdots \\
y &= m \text{ if } \alpha_{m-1} < y^*
\end{aligned}$$

Estimación por MV

- Puesto que:

$$\Pr(y=0|x) = \Pr\left(x'\beta + \varepsilon \leq \alpha_1\right) = 1 - \Phi\left(x'\beta - \alpha_1\right)$$
$$\Pr(y=1|x) = \Phi\left(x'\beta - \alpha_1\right) - \Phi\left(x'\beta - \alpha_2\right)$$
$$\vdots$$
$$\Pr(y=m|x) = \Phi\left(x'\beta - \alpha_{m-1}\right)$$

- Entonces, para una muestra de N observaciones, la verosimilitud es:

$$\mathrm{L}(b) = \prod_{i=1}^{N} \left\{ \left(1 - \Phi\left(x_i'b - \alpha_1\right)\right)^{1(y_i=0)} \times \right.$$
$$\left(\Phi\left(x_i'b - \alpha_1\right) - \Phi\left(x_i'b - \alpha_2\right)\right)^{1(y_i=1)} \times \ldots \times$$

Comandos básicos en `gretl` para la estimación del probit Ordinal

- `probit`: estima por MV el modelo probit ordinal si la variable dependiente no es binaria pero es discreta
- `omit/add`: contrasta tests de significatividad
- `$yhat`: proporciona estiamaciones de la variable latente
- `$lnl`: proporciona la log-verosimilitud del último modelo estimado
- `logit`: estima por MV el modelo logit ordinal si la variable dependiente no es binaria pero es discreta

- en esta Sesión vamos a parender a utilizar el comando `probit` para estimar modelos probit ordinales

Probit ordinal en gretl

probit *depvar* *indvars* --robust --verbose --p-values

- *depvar* debe tomar valores enteros no negativos, o estar explícitamente marcada como discreta.
  - Caso de que tenga valores no enteros, es recodificada.
- por defecto, los errores estándar se obtienen con la inversa del opuesto del hesiano
- opciones:
  1. `--robust`: matriz de covarianzas robusta
  2. `--verbose`: muestra todos los pasos del alogoritmo de maximización

Estimaciones

- Se obtienen estimaciones MV para los coeficientes β y para los valores o puntos de corte α_m.
  - éstos se representan en `gretl` como `cut1`, `cut2`, etc.
- `$uhat` da los residuos generalizados; `$yhat` da $x'\widehat{\beta}$.
  - It is thus possible to compute an unbiased estimate of the latent variable U simply by adding the two together.
- el output muestra el test χ^2_q para la nula de que todas las pendientes son iguales a cero

Ejemplo: Datos simulados

Actividad

- $y^* = 0.07 \times educ - 1.0 \times kids + \varepsilon$, donde $\varepsilon \sim \mathcal{N}(0,1)$

$$y = 0 \text{ if } y^* \leq 0.5 \text{ (inactiva)}$$
$$y = 1 \text{ if } 0.5 < y^* \leq 2.5 \text{ (tiempo parcial)}$$
$$y = 2 \text{ if } 2.5 < y^* \text{ (tiempo completo)}$$

- la educación da más utilidad cuanto más trabajas
- tener un niño da más utilidad cuanto menos trabajas

probit *work educ kids*

```
Model 1: Ordered Probit, using observations 1-5000
Dependent variable: work

             coefficient   std. error      z       p-value
  --------------------------------------------------------
  educ         0.0795682    0.00402396    19.77    5.03e-87  ***
  kids        -0.955750     0.0366670    -26.07    8.94e-150 ***

  cut1         0.656064     0.0677545      9.683   3.56e-22  ***
  cut2         2.63981      0.0764705     34.52    3.93e-261 ***

Mean dependent var   0.657000   S.D. dependent var   0.599852
Log-likelihood      -3903.096   Akaike criterion     7814.193
Schwarz criterion    7840.262   Hannan-Quinn         7823.330

Number of cases 'correctly predicted' = 3107 (62.1%)
Likelihood ratio test: Chi-square(2) = 1054.95 [0.0000]

Test for normality of residual -
  Null hypothesis: error is normally distributed
  Test statistic: Chi-square(2) = 0.108987
  with p-value = 0.946965
```

El MLP puede dar malos resultados

```
Model 2: OLS, using observations 1-5000
Dependent variable: work
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio    p-value
  --------------------------------------------------------
  const       0.251884     0.0292683      8.606   1.00e-17  ***
  educ        0.0354561    0.00169676    20.90    4.97e-93  ***
  kids       -0.431542     0.0152135    -28.37    3.02e-164 ***

Mean dependent var    0.657000   S.D. dependent var   0.599852
Sum squared resid     1457.253   S.E. of regression   0.540024
R-squared             0.189855   Adjusted R-squared   0.189531
F(2, 4997)            651.6896   P-value(F)           3.2e-252
Log-likelihood       -4012.480   Akaike criterion     8030.960
Schwarz criterion     8050.512   Hannan-Quinn         8037.813
```

omit no funciona para la significatividad conjunta

- no podemos usar omit para contrastar la significatividad conjunta de todos las pendientes con el test de Wald

omit educ kids –wald

```
  -

No independent variables left after omissions

Error executing script: halting
 > omit educ kids --wald
```

- el resultado del test LR aparece en el output, y además podemos calcularlo "a mano"

Test LR para la significatividad de las pendientes

```
# estimating unrestricted model and storing loglikelihood
probit work educ kids --quiet
scalar lur= $lnl

# estimating restricted model and storing loglikelihood
probit work const --quiet
scalar lr= $lnl

# computing the LR statistic and p-value
scalar LR=2*(lur-lr)
scalar pval = pvalue(X, 1, LR)
```

- el resultado coincide con el output en `probit`:

```
? printf " LR = %.8g\n p-value = %.8g\n", LR,pval
 LR = 1054.9521
 p-value = 2.0415848e-231
```

Efectos marginales

Efectos parciales sobre las probabilidades

- Como interpretar los resultados?
  - interpretar la ecuación de la variable latente
  - $\frac{\partial \mathrm{E}(y|x)}{\partial x_j} = \frac{\partial \Pr(y=0|x)}{\partial x_j} \times 0 + \frac{\partial \Pr(y=1|x)}{\partial x_j} \times 1 + \ldots + \frac{\partial \Pr(y=m|x)}{\partial x_j} \times m$
- Alternativamente, podríamos computar los efectos sobre las probabilidades de los diferentes categorías ordenadas
  - si x_j es discreta computamos como en el caso binario de un cambio discreto en x_j.
  - en general, los efectos parciales para las probabilidades intermedias son pequeños y no significativos.

Efectos Marginales: Cambio discreto

Cambio discreto de x_0 a x_1

- estima el modelo y guarda $\widehat{\beta}$ y $\widehat{\alpha}$
- predice la función índice en los dos puntos $x_0'\widehat{\beta}$ y $x_1'\widehat{\beta}$
- genera los efectos marginales individuales

$$\Delta\widehat{\Pr}(y=0|x) = \Phi\left(x_0'\widehat{\beta}-\widehat{\alpha}_1\right) - \Phi\left(x_1'\widehat{\beta}-\widehat{\alpha}_1\right) = -\Delta\Phi\left(x'\widehat{\beta}-\widehat{\alpha}_1\right)$$

$$\Delta\widehat{\Pr}(y=1|x) = \Delta\Phi\left(x'\widehat{\beta}-\widehat{\alpha}_1\right) - \Delta\Phi\left(x'\widehat{\beta}-\widehat{\alpha}_2\right)$$

$$\vdots$$

$$\Delta\widehat{\Pr}(y=m|x) = \Delta\Phi\left(x'\widehat{\beta}-\widehat{\alpha}_{m-1}\right)$$

- calcula las medias

Ejemplo: tener un niño *adicional*

```
# marginal effects of having an additional kid
probit work educ kids --quiet
genr beta=$coeff[1:2]
genr alpha=$coeff[3:4]
series kids0=kids
matrix x0={educ,kids0}
series kids1=kids0+1
matrix x1={educ,kids1}
series x1b = x1*beta
series x0b = x0*beta

series Mg_kid0 = (cdf(N,x0b-alpha[1])-cdf(N,x1b-alpha[1]))
series Mg_kid1 = (cdf(N,x1b-alpha[1])-cdf(N,x0b-alpha[1])) \
                -(cdf(N,x1b-alpha[2])-cdf(N,x0b-alpha[2]))
series Mg_kid2 = (cdf(N,x1b-alpha[2])-cdf(N,x0b-alpha[2]))
```

summary *Mg_kid** −−simple

```
summary Mg_kid* --simple
Summary statistics, using the observations 1 - 5000

                     Mean         Minimum         Maximum        Std. Dev.
Mg_kid0           0.31409         0.13797         0.36398         0.057924
Mg_kid1          -0.25621        -0.31999        -0.13648         0.058980
Mg_kid2         -0.057872        -0.13917      -0.0014979         0.050060
```

- tener un niño adicional aumenta la probabilidad de no trabajar en más de 30 puntos porcentuales
- este incremento se produce a costa de reducciones de 25.6 puntos porcentuales en la probabilidad de trabajar a tiempo parcial y de 5.8 puntos porcentuales en la probabilidad de trabajar a tiempo completo

Efectos marginales: Cambios infinitesimales

Aproximación con cálculo

- estima el modele y guarda $\widehat{\beta}$ y $\widehat{\alpha}$
- predice la función $x'\widehat{\beta}$ y calcula su media $\overline{x'\widehat{\beta}}$
- genera la aproximación:

$$\frac{\partial \widehat{\Pr}(y=0|x)}{\partial x_j} = -\phi\left(\overline{x'\widehat{\beta}} - \widehat{\alpha}_1\right)\widehat{\beta}_j$$

$$\frac{\partial \widehat{\Pr}(y=1|x)}{\partial x_j} = \left(\phi\left(\overline{x'\widehat{\beta}} - \widehat{\alpha}_1\right) - \phi\left(\overline{x'\widehat{\beta}} - \widehat{\alpha}_2\right)\right)\widehat{\beta}_j$$

$$\frac{\partial \widehat{\Pr}(y=2|x)}{\partial x_j} = \phi\left(\overline{x'\widehat{\beta}} - \widehat{\alpha}_2\right)\widehat{\beta}_j$$

Ejemplo de la aproximación

```
# one extra year of education: calculus approximation
genr beta=$coeff[1:2]
genr alpha=$coeff[3:4]
matrix X={educ,kids}
series Xb=X*beta
scalar meanXb=mean(Xb)
scalar Mg_educ0=-pdf(N,meanXb-alpha[1])*$coeff(educ)
scalar Mg_educ1= (pdf(N,meanXb-alpha[1]) \
         -pdf(N,meanXb-alpha[2]))*$coeff(educ)
scalar Mg_educ2= pdf(N,meanXb-alpha[2])*$coeff(educ)
```

```
? printf " Mg_educ0 = %.8g\n Mg_educ1 = %.8g\n Mg_educ2 = %.8g\n", \
          Mg_educ0, Mg_educ1, Mg_educ2

 Mg_educ0 = -0.030702828
 Mg_educ1 = 0.023540482
 Mg_educ2 = 0.0071623465
```

Efectos sobre la variable ordinal

Efectos sobre y

- una vez hemos obtenido los efectos marginales sobre las probabilidades, el efecto esperado sobre la variable ordinal y is fácil de obtener
- para regresor continuo:

$$\frac{\partial \widehat{\mathrm{E}}(y|x)}{\partial x_j} = \frac{\partial \widehat{\Pr}(y=1|x)}{\partial x_j} \times 1 + \frac{\partial \widehat{\Pr}(y=2|x)}{\partial x_j} \times 2$$

- para regresor discreto:

$$\Delta \widehat{\mathrm{E}}(y|x) = \Delta \widehat{\Pr}(y=1|x) + 2 \times \Delta \widehat{\Pr}(y=2|x)$$

```
# marginal effects on work
scalar EMg_kid=mean(Mg_kid1)+2*mean(Mg_kid2)
scalar EMg_educ=Mg_educ1+2*Mg_educ2
```

```
? printf " EMg_kid = %.8g\n EMg_educ = %.8g\n", EMg_kid, EMg_educ
 EMg_kid = -0.371958
 EMg_educ = 0.037865175
```

Resumen

- gretl permite la estimación por MV del probit ordinal y del logit ordinal
- los efectos marginales son fáciles de computar