Serie Apuntes de Clase ΩBΓ

N° 10
Diciembre de 2014

MODELOS DE ELECCIÓN NOMINAL. APLICACIONES EN STATA 14.

Rafael Bustamante Romaní

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú, DECANA DE AMÉRICA

FACULTAD DE CIENCIAS ECONÓMICAS

La **Serie Apuntes de Clase Omega Beta Gamma** tiene por objetivo difundir los materiales de enseñanza generados por los docentes que tienen a su cargo el desarrollo de las asignaturas que forman parte de los Planes de Estudios de las Escuelas Académico-Profesionales de la Facultad de Ciencias Económicas de la Universidad Nacional Mayor de San Marcos. Estos documentos buscan proporcionar a los estudiantes la explicación de algunos temas específicos que son abordados en su formación universitaria.

Encargados de la serie:

Bustamante Romaní, Rafael.
rbustamanter@unmsm.edu.pe

Cisneros García, Juan Manuel.
jcisnerosg@unmsm.edu.pe

Facultad de Ciencias Económicas.
Universidad Nacional Mayor de San Marcos.
Calle Germán Amézaga N° 375.
Ciudad Universitaria, Lima 1. Perú.

La **Serie Apuntes de Clase ΩΒΓ** es promovida y desarrollada por un colectivo de docentes del Departamento de Economía de la Universidad Nacional Mayor de San Marcos.

Modelos de Elección Ordinal.

Aplicaciones en Stata 14.

Rafael Bustamante Romaní[◊]

Resumen

Los modelos Multinomiales introducidos por McFadden en 1974 son herramientas de análisis que, con base en el supuesto de que la variable categórica de interés sigue una distribución Multinomial. Estos modelos utilizan el método de Máxima Verosimilitud para estimar las probabilidades asociadas a cada elección, dadas las características particulares de los individuos o los atributos de las elecciones, resumidas en los regresores. Estas notas son una introducción a los modelos de elección Ordinal, con aplicaciones a Stata 14. Asimismo se presentan todas las herramientas en proceso de estimación y análisis de indicadores de bondad de ajuste, parsimonia entre otros.

Palabras Claves: Modelos de elección discreta, especificaciones, Logit Multinomial, Probit Multinomial

Clasificación JEL: C2, C25

[◊] Estudios concluidos de Doctorado en Economía con mención en los Recursos Naturales (c), Universidad Nacional Autónoma de México. MBA Gerencial (c), CENTRUM Pontificia Universidad Católica del Perú. Maestría en Economía con mención en Finanzas, Universidad Nacional Mayor de San Marcos. B. Sc. Economía, UNMSM. Profesor Auxiliar del Departamento de Economía de la UNMSM. Investigador asociado al Instituto de Investigaciones FCE - UNMSM. Contacto: rbustamanter@unmsm.edu.pe

I. INTRODUCCION

Cuando la variable dependiente es discreta, pero sus valores indican un orden, no es correcto realizar la estimación de la misma a través de los modelos presentados en el apartado anterior, ya que la inclusión de la información que aporta el orden de las alternativas en la especificación del modelo permite obtener unos mejores resultados (Bravo & Vásquez, 2008).

Las variables ordinales son a menudo codificadas como enteros consecutivas de 1 al número de categorías, no sería correcto el uso de un modelo de regresión clásico, ya que codificadas las posibles alternativas como 1, 2, ...(j+1), ..., J, se estaría considerando la diferencia entre (j+1) y (j+2) como la existente entre 1 y 2, lo cual no tiene por qué ser así ya que los números utilizados en la codificación solo representan un orden dentro de una clasificación. Así, con modelos de salida ordinal es mejor usar modelos que eviten el supuesto de que las distancias entre las categorías sean iguales, ahora nos enfocaremos en un logit y probit que consideren esta ordenación, modelos introducidos por McKelvey y Zavoina (1975) en términos de una variable latente.

Cuando las salidas son ordinales o nominales la dificultad de explicar más de dos respuestas se incrementa. Una variable puede ser ordenada de cierta manera cuando consideramos un tema, y ordenada de otra manera cuando consideramos un tema diferente. Millar y Volker (1985) descubrieron como diferentes supuestos sobre el ordenamiento de ocupaciones, proyectan diferentes resultados. Una variable podría reflejar ordenamiento sobre más de una dimensión tal como escalas de actitudes, que reflejen ambas la intensidad y dirección de opinión. Más aún es muy común que encuestas incluyan la categoría "no sabe, no opina", lo cual probablemente no corresponda a la categoría intermedia en una escala, aun cuando en el análisis uno este tentado a colocarla como tal, sobre todo cuando la propuesta de ordenamiento es ambigua, el modelo de salidas nominales podría ser considerado.

Estos modelos pueden ser desarrollados de diferentes maneras, pero cada una de ellas nos conduce al mismo resultado. Además de que el modelo de regresión binaria (MRB) puede ser visto como un caso especial de los modelos de regresión Ordinales (MRO), en el cual la variable endógena solo tiene dos categorías.

II. MODELO DE VARIABLE LATENTE

El modelo de regresión ordinal es comúnmente presentado como un modelo de variable latente. Definida y ∗ como una variable latente cuyo rango va desde -∞ a ∞.

$$y^*_i = X_i\beta + \varepsilon_i \qquad 1.$$

Donde la variable endógena toma los siguientes valores:

$$y_i = \text{m}, \ \text{r}_{m-1} \le y_i < r_m \ \forall\text{m=1,2,3,....J} \qquad 2.$$

O también de manera extendida:

$$y_i = \begin{cases} 1, & -\infty = \text{r}_0 \le y_1^* < \text{r}_1 \\ 2, & \text{r}_1 \le y_1^* < \text{r}_2 \\ 3, & \text{r}_2 \le y_1^* < \text{r}_3 \\ . & \\ . & \\ . & \\ J, & \text{r}_{J-1} \le y_J^* < \text{r}_J = +\infty \end{cases} \qquad 3.$$

Asimismo, _se establecen puntos de corte (r_i) en los cuales se sitúa el índice de performance.

Si $y_1^* < \text{r}_1$, el individuo se ubica en la categoría 1; si $\text{r}_1 \le y_1^* < \text{r}_2$ el individuo se sitúa en la categoría 2, si $\text{r}_2 \le y_1^* < \text{r}_3$ el individuo se sitúa en la categoría 3, si $\text{r}_3 \le y_1^* < \text{r}_4$ el individuo se sitúa en la categoría 4 y así hasta llegar a la última categoría.

De esta forma se requerirán tantos puntos de corte como categorias haya, menos uno. Téngase en cuenta que las distancias entre los valores de corte no pueden asumirse como uniformes, razón por la cual cualquier tipo de regresión lineal no debe ser usado.

Donde los puntos de corte rj son estimados. Como ejemplo, podríamos tener la siguiente pregunta en una encuesta: ¿Una mujer trabajadora establece un fuerte y seguro vínculo con su hijo, así como una mujer que no trabaja?

Las posibles respuestas podrían ser:

1=Desacuerdo Total

2=Desacuerdo

3=Acuerdo

4= total acuerdo

La variable latente continua puede imaginarse como el grado de aceptación a favor de que las mujeres trabajadoras son buenas madres.

$$y_i = \begin{cases} 1, & -\infty = r_0 \leq y_1^* < r_1 \\ 2, & r_1 \leq y_1^* < r_2 \\ 3, & r_2 \leq y_1^* < r_3 \\ 4, & r_3 \leq y_J^* < r_4 = +\infty \end{cases} \qquad 4.$$

La probabilidad de una variable observada dado el valor de x, corresponde a la región en la que la distribución de y ∗ cae entre rm−1y rm

$$\mathrm{Prob}(y_m = m / x) = \mathrm{Prob}(r_{m-1} \leq y_1^* < r_m) \qquad 5.$$

Sustituyendo $x\beta + \varepsilon$ por y ∗ y usando algo de algebra obtenemos la formula estándar que predice la probabilidad en el MRO, cuya interpretación es la probabilidad de que el individuo se encuentre en la categoría m dadas las variables explicativas x.

$$\Pr ob(y_m = m \ / x) = F(r_m - x\beta) - F(r_{m-1} - x\beta) \quad 6.$$

Supuesto:
Donde F es la función de probabilidad acumulada para ε. En el probit ordinal, F es una normal con $Var(\varepsilon_i)=1$, en el logit ordinal, F es una logistica con $Var(\varepsilon_i)=\frac{\pi^2}{3}$. Notar que cuando y=1 el termino $F(-\infty - x\beta)=0$ y cuando y=J el primer termino de $F(\infty - x\beta)=1$.

Donde, comúnmente, $F(.)$ puede ser normal estándar o logística, lo que da lugar a los modelos probit o logit ordenado, respectivamente. Para que todas las probabilidades sean positivas, debe ser cierto que $0 < r_1 < r_2 < r_3 ..... < r_J$. Estos puntos de corte son estimados por el modelo junto con los β y hacen posible determinar las probabilidades estimadas de estar en cada categoría[1]

De hecho, si los a estimados son significativamente diferentes de cero, ello implica que las categorías son definitivamente ordenadas (Orihuela, 2011).

Como en el caso binomial, los coeficientes no tienen un significado individual sino dentro del argumento de la función de densidad. No obstante, su signo indicará la dirección de la relación con la probabilidad de estar en la categoría más alta, y la inversa de la misma en el caso de la categoría más baja2. Las categorías intermedias tendrán efectos impacto que no se puede definir a priori.

De hecho, y tal como se observa en las expresiones que siguen para el efecto impacto de una variable continua, solo se puede adelantar el signo, sin ambigüedad, para los dos casos extremos (Orihuela, 2011).

[1] Cabe mencionar que las probabilidades de estar en cada una de las cuatro categorías, para un mismo conjunto de valores de las variables explicativas, deben sumar 1.

2 Es decir, un coeficiente positivo indica que la variable explicativa correspondiente tiene una relación positiva con la categoría más alta, y negativa con la más baja.

$$\frac{\partial \Pr ob(y_i = 1/x)}{\partial x_k} = [-f(r_1 - x_i\beta)]\beta_k$$

$$\frac{\partial \Pr ob(y_m = 2/x)}{\partial x_k} = [-f(r_2 - x_i\beta) + F(r_1 - x_i\beta)]\beta_k$$

$$\frac{\partial \Pr ob(y_m = 3/x)}{\partial x_k} = [-f(r_3 - x_i\beta) + f(r_2 - x_i\beta)]\beta_k$$

.

.

.

$$\frac{\partial \Pr ob(y_i = J-1/x)}{\partial x_k} = [-f(r_J - x_i\beta) + f(r_{J-1} - x_i\beta)]\beta_k$$

$$\frac{\partial \Pr ob(y_i = J/x)}{\partial x_k} = [f(r_{J-1} - x_i\beta)]\beta_k \qquad 7.$$

Téngase en cuenta que los efectos impacto de las probabilidades de estar en cada una de las J categorías, ante cambios de una misma variable explicativa, deben sumar cero, ya que consisten en un juego de suma cero en lo que se refiere al impacto final sobre dichas probabilidades (Bravo & Vásquez, 2008).

Cabe mencionar, por último, que las particularidades anteriores también aplican a la interpretación de los RP. Así, por ejemplo, un coeficiente positivo implica que cuanto mayor sea el regresor asociado, mayor será el RP de la categoría más alta frente a las de menor valoración.

Comparando estas ecuaciones con las de un MRB se observa que el MRO es idéntico a la regresión binaria, veamos:

use fuerza[3]

[3] Se puede acceder a este archivo en el siguiente link: https://www.dropbox.com/s/3y270m59f0xpliw/fuerza.dta.DTA?dl=0&preview=fuerza.dta.DTA

```
. d

Contains data from C:\Users\finance.bussines\Documents\Dropbox\ECONOMETRIA II\
  obs:           753
 vars:            22                         2 Mar 1999 11:30
 size:        36,897
--------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
--------------------------------------------------------------------------------
inlf            byte   %9.0g                  =1 if in lab frce, 1975
hours           int    %9.0g                  hours worked, 1975
kidslt6         byte   %9.0g                  # kids < 6 years
kidsge6         byte   %9.0g                  # kids 6-18
age             byte   %9.0g                  woman's age in yrs
educ            byte   %9.0g                  years of schooling
wage            float  %9.0g                  est. wage from earn, hrs
repwage         float  %9.0g                  rep. wage at interview in 1976
hushrs          int    %9.0g                  hours worked by husband, 1975
husage          byte   %9.0g                  husband's age
huseduc         byte   %9.0g                  husband's years of schooling
huswage         float  %9.0g                  husband's hourly wage, 1975
faminc          float  %9.0g                  family income, 1975
mtr             float  %9.0g                  fed. marg. tax rte facing woman
motheduc        byte   %9.0g                  mother's years of schooling
fatheduc        byte   %9.0g                  father's years of schooling
unem            float  %9.0g                  unem. rate in county of resid.
city            byte   %9.0g                  =1 if live in SMSA
exper           byte   %9.0g                  actual labor mkt exper
nwifeinc        float  %9.0g                  (faminc - wage*hours)/1000
lwage           float  %9.0g                  log(wage)
expersq         int    %9.0g                  exper^2
--------------------------------------------------------------------------------
Sorted by:

.
```

```
. tab inlf

   =1 if in |
  lab frce, |
       1975 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        325       43.16       43.16
          1 |        428       56.84      100.00
------------+-----------------------------------
      Total |        753      100.00
```

ologit inlf kidslt6 kidsge6 age hushrs faminc, nolog

```
. logit inlf kidslt6 kidsge6 age hushrs faminc, nolog

Logistic regression                               Number of obs   =        753
                                                  LR chi2(5)      =      83.38
                                                  Prob > chi2     =     0.0000
Log likelihood = -473.18104                       Pseudo R2       =     0.0810

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     kidslt6 |  -1.378262   .1902921    -7.24   0.000    -1.751227   -1.005296
     kidsge6 |  -.1189967   .0658062    -1.81   0.071    -.2479744     .009981
         age |  -.0718954   .0123235    -5.83   0.000     -.096049   -.0477419
      hushrs |  -.0003241   .0001333    -2.43   0.015    -.0005854   -.0000629
      faminc |   .0000211   6.89e-06     3.06   0.002     7.60e-06    .0000346
       _cons |   4.086941   .6892715     5.93   0.000     2.735994    5.437888
------------------------------------------------------------------------------

. ologit inlf kidslt6 kidsge6 age hushrs faminc, nolog

Ordered logistic regression                       Number of obs   =        753
                                                  LR chi2(5)      =      83.38
                                                  Prob > chi2     =     0.0000
Log likelihood = -473.18104                       Pseudo R2       =     0.0810

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     kidslt6 |  -1.378262   .1902921    -7.24   0.000    -1.751227   -1.005296
     kidsge6 |  -.1189967   .0658062    -1.81   0.071    -.2479744     .009981
         age |  -.0718954   .0123235    -5.83   0.000     -.096049   -.0477419
      hushrs |  -.0003241   .0001333    -2.43   0.015    -.0005854   -.0000629
      faminc |   .0000211   6.89e-06     3.06   0.002     7.60e-06    .0000346
-------------+----------------------------------------------------------------
       /cut1 |  -4.086941   .6892715                     -5.437888   -2.735994
------------------------------------------------------------------------------

.
```

Los coeficientes y sus desviaciones estándar son los mismos pero el intercepto para el logit, es reportado, mientras que para el ologit ese intercepto es reemplazado por el punto de corte del mismo nivel pero de signo opuesto (Bernardí Cabrer Borrás & Amparo Sancho Pérez, Guada, 2001).

En Stata, la identificación del modelo de regresión ordinal asume que el intercepto es cero y así los valores de los puntos de corte son estimados.

El modelo de regresión ordinal puede también ser desarrollado como un modelo de probabilidad no lineal sin recurrir a la idea de variable latente. Para mostrar esto, primero definimos el odds de que la variable explicada es menor o igual a "m" vs que sea mayor que "m" dado las variables exogenas "x": Por ejemplo, podríamos calcular el ratio de odds de desagrado o fuerte desagrado, versus el agrado o fuerte agrado. Así el logaritmo del odds es igual a (Beltran Barco, Arlette; Castro Carlin, Juan;, 2010):

$$\Omega_{\leq m/>m} = \frac{\Pr ob(y \leq m / x)}{\Pr ob(y > m / x)} \qquad 8.$$

Para una simple variable independiente y tres categorías en la explicada, donde el intercepto fue fijado en "0", tendríamos:

$$Ln(\frac{\Pr ob(y \leq 1 / x)}{\Pr ob(y > 1 / x)}) = r_1 - \beta x_1 \qquad 9.$$

$$Ln(\frac{\Pr ob(y \leq 2 / x)}{\Pr ob(y > 2 / x)}) = r_2 - \beta x_2 \qquad 10.$$

Parece confuso que el modelo substraiga xb en lugar de añadirlo, esto es consecuencia del cálculo del logit de y ≤ m vs y > m.

Aquí un ejemplo basado en la encuesta realizada entre 1977 y 1989 de General Social Survey, donde el tema y pregunta tratado fue: "¿Una madre trabajadora puede establecer una relación afectiva de calidad con su hijo (a) en función de variables socioeconómicas como la edad, prestigio ocupacional, la raza, el generó del hijo? (Orihuela, 2011)"

use relaciones, clear[4]

[4] Se puede acceder a este archivo en el siguiente link: https://www.dropbox.com/s/wipcdjbata0p1gb/relaciones.dta.dta?dl=0&preview=relaciones.dta.dta

```
. d

Contains data from C:\Users\finance.bussines\Documents\Dropbox\ECONOMETRIA II\sesion 4\warm.dta
  obs:         2,293                         77 & 89 General Social Survey
 vars:             7                         16 Nov 2009 14:25
 size:        16,051                         (_dta has notes)
------------------------------------------------------------------------------------------------
              storage  display    value
variable name   type   format     label      variable label
------------------------------------------------------------------------------------------------
warm            byte   %10.0g     lab1       las madres tienen fuertes relaciones con los hijos
yr89            byte   %10.0g     yrlbl      encuesta anual: 1=1989 0=1977
male            byte   %10.0g     lab2       genero: 1=hombre 0=mujer
white           byte   %10.0g     lab3       raza 1=blanco 0=no blanco
age             byte   %10.0g                edad en años
ed              byte   %10.0g                años de educación
prst            byte   %10.0g                prestigio ocupacional
------------------------------------------------------------------------------------------------
Sorted by:  warm

.
```

tab warm

```
. table warm
```

las madres tienen fuertes relaciones con los hijos	Freq.
FD	297
D	723
A	856
FA	417

tab warm

```
. tab warm

 las madres |
     tienen |
   fuertes  |
relaciones  |
    con los |
      hijos |      Freq.     Percent        Cum.
------------+-----------------------------------
         FD |        297       12.95       12.95
          D |        723       31.53       44.48
          A |        856       37.33       81.81
         FA |        417       18.19      100.00
------------+-----------------------------------
      Total |      2,293      100.00
```

Usando los datos estimamos el siguiente modelo:

$$\Pr ob(\mathrm{warm}=\mathrm{m}/\ x_i)=\mathrm{F}(\mathrm{r}_m-x_i\beta)-\mathrm{F}(\mathrm{r}_{m-1}-x_i\beta) \quad 11.$$

Donde:

$$x\beta=\beta_{yr89}yr89+\beta_{male}male+\beta_{white}white+\beta_{ed}ed+\beta_{prst}prst\beta_{prst}prst \quad 12.$$

Aquí las estimaciones sean con ologit, oprobit, pueden ser comparadas con el outreg:

```
ologit warm yr89 male white age ed prst,nolog
 outreg using ordenado,replace
oprobit warm yr89 male white age ed prst,nolog
 outreg using ordenado,append
```

```
. ologit warm yr89 male white age ed prst,nolog

Ordered logistic regression                       Number of obs   =       2293
                                                  LR chi2(6)      =     301.72
                                                  Prob > chi2     =     0.0000
Log likelihood = -2844.9123                       Pseudo R2       =     0.0504

------------------------------------------------------------------------------
        warm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        yr89 |   .5239025   .0798989     6.56   0.000     .3673036    .6805014
        male |  -.7332997   .0784827    -9.34   0.000     -.887123   -.5794765
       white |  -.3911595   .1183808    -3.30   0.001    -.6231816   -.1591373
         age |  -.0216655   .0024683    -8.78   0.000    -.0265032   -.0168278
          ed |   .0671728    .015975     4.20   0.000     .0358624    .0984831
        prst |   .0060727   .0032929     1.84   0.065    -.0003813    .0125267
-------------+----------------------------------------------------------------
       /cut1 |  -2.465362   .2389128                     -2.933622   -1.997102
       /cut2 |   -.630904   .2333156                     -1.088194   -.1736138
       /cut3 |   1.261854    .234018                      .8031871    1.720521
------------------------------------------------------------------------------

. outreg using ordenado,replace
(note: file ordenado.doc not found)
```

warm	yr89	0.524
		(6.56)**
	male	-0.733
		(9.34)**
	white	-0.391
		(3.30)**
	age	-0.022
		(8.78)**
	ed	0.067
		(4.20)**
	prst	0.006
		(1.84)
cut1	_cons	-2.465
		(10.32)**
cut2	_cons	-0.631
		(2.70)**
cut3	_cons	1.262
		(5.39)**
N		2,293

* p<0.05; ** p<0.01

```
. oprobit warm yr89 male white age ed prst,nolog

Ordered probit regression                         Number of obs   =       2293
                                                  LR chi2(6)      =     294.32
                                                  Prob > chi2     =     0.0000
Log likelihood =  -2848.611                       Pseudo R2       =     0.0491

------------------------------------------------------------------------------
        warm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        yr89 |   .3188147   .0468521     6.80   0.000     .2269863    .4106431
        male |  -.4170287   .0455461    -9.16   0.000    -.5062974     -.32776
       white |  -.2265002   .0694776    -3.26   0.001    -.3626738   -.0903267
         age |  -.0122213   .0014427    -8.47   0.000    -.0150489   -.0093937
          ed |   .0387234   .0093241     4.15   0.000     .0204485    .0569983
        prst |    .003283    .001925     1.71   0.088    -.0004899    .0070559
-------------+----------------------------------------------------------------
       /cut1 |  -1.428578   .1387749                     -1.700572   -1.156585
       /cut2 |  -.3605589   .1369224                     -.6289219   -.0921959
       /cut3 |   .7681637   .1370569                      .4995371     1.03679
------------------------------------------------------------------------------

. outreg using ordenado,append
```

warm	yr89	0.524
		(6.56)**
	male	-0.733
		(9.34)**
	white	-0.391
		(3.30)**
	age	-0.022
		(8.78)**
	ed	0.067
		(4.20)**
	prst	0.006
		(1.84)
cut1	_cons	-2.465
		(10.32)**
cut2	_cons	-0.631
		(2.70)**
cut3	_cons	1.262
		(5.39)**
N		2,293
warm	yr89	0.319
		(6.80)**
	male	-0.417
		(9.16)**
	white	-0.227
		(3.26)**
	age	-0.012
		(8.47)**
	ed	0.039
		(4.15)**
	prst	0.003
		(1.71)
cut1	_cons	-1.429
		(10.29)**
cut2	_cons	-0.361
		(2.63)**
cut3	_cons	0.768
		(5.60)**
N		2,293

* p<0.05; ** p<0.01


```
.
end of do-file

.
```

Como en el análisis de los modelos de regresión binaria, la diferencia estriba en que los coeficientes tienen una razón de 1.7, es decir, solo hay diferencia en escala, sin embargo los z-test, son los mismos y no se ven afectados por la escala.

III. PRUEBAS DE HIPÓTESIS

Para las pruebas de hipótesis podremos usar el test de wald, el de máxima verosimilitud, o usar el fitstat para elegir el mejor modelo.

ologit warm yr89 male white age ed prst,nolog

test male

```
. test male

 ( 1)  [warm]male = 0

           chi2(  1) =   87.30
         Prob > chi2 =    0.0000
```

ologit warm yr89 male white age ed prst,nolog

test age white male

```
. test age white male

 ( 1)  [warm]age = 0
 ( 2)  [warm]white = 0
 ( 3)  [warm]male = 0

           chi2(  3) =  166.62
         Prob > chi2 =    0.0000
```

ologit warm yr89 male white age ed prst,nolog
lrtest, saving(0)
ologit warm yr89 white age ed prst,nolog
lrtest

```
.
. lrtest
You ran lrtest using the old syntax.  Click here to learn about the new syntax.


Likelihood-ratio test                                 LR chi2(1)  =     88.73
(Assumption: . nested in LRTEST_0)                    Prob > chi2 =    0.0000

.
```

ologit warm yr89 male white age ed prst,nolog
lrtest, saving(0)
ologit warm yr89 ed prst,nolog
lrtest

```
. lrtest
You ran lrtest using the old syntax.  Click here to learn about the new syntax.


Likelihood-ratio test                                 LR chi2(3)  =    171.58
(Assumption: . nested in LRTEST_0)                    Prob > chi2 =    0.0000
```

ologit warm yr89 male white age ed prst,nolog
fitstat

```
. fitstat

Measures of Fit for ologit of warm

Log-Lik Intercept Only:      -2995.770   Log-Lik Full Model:          -2844.912
D(2284):                      5689.825   LR(6):                         301.716
                                         Prob > LR:                       0.000
McFadden's R2:                   0.050   McFadden's Adj R2:               0.047
ML (Cox-Snell) R2:               0.123   Cragg-Uhler(Nagelkerke) R2:      0.133
McKelvey & Zavoina's R2:         0.127
Variance of y*:                  3.768   Variance of error:               3.290
Count R2:                        0.432   Adj Count R2:                    0.093
AIC:                             2.489   AIC*n:                        5707.825
BIC:                        -11982.891   BIC':                         -255.291
BIC used by Stata:            5759.463   AIC used by Stata:            5707.825

.
```

IV. SUPUESTO DE PARALELISMO

Antes de discutir la interpretación, es importante entender un supuesto que está implícito en el modelo de regresión ordenado (MRO), conocido como paralelismo de la regresión, y para el modelo ologit, el supuesto de odds proporcional.

$$\begin{aligned}
&\Pr ob(y_i = 1/x_i) = \mathrm{F}(\mathrm{r}_1 - x_i\beta)\\
&\Pr ob(y_i = 2/x_i) = \mathrm{F}(\mathrm{r}_2 - x_i\beta) - \mathrm{F}(\mathrm{r}_1 - x_i\beta)\\
&.\\
&.\\
&.\\
&\Pr ob(y_i = m/x_i) = \mathrm{F}(\mathrm{r}_m - x_i\beta) - \mathrm{F}(\mathrm{r}_{m-1} - x_i\beta)
\end{aligned} \qquad 13.$$

Las ecuaciones presentadas pueden ser usadas para calcular la probabilidad acumulada, lo cual tienen la siguiente forma:

$$\Pr ob(y_i < m/x_i) = \mathrm{F}(\mathrm{r}_m - x_i\beta)\ \forall\mathrm{=m=1,2,3,...J\text{-}1} \qquad 14.$$

En esta ecuación se muestra que el MRO es equivalente para J-1 regresiones binarias con el supuesto de que las pendientes o coeficientes son idénticos a lo largo de cada regresión.

Por ejemplo, si tenemos cuatro categorías en nuestra endógena y una variable independiente las ecuaciones serian:

$$\begin{aligned}
&\Pr ob(y_i < 1/x_i) = \mathrm{F}(\mathrm{r}_1 - x_i\beta)\\
&\Pr ob(y_i < 2/x_i) = \mathrm{F}(\mathrm{r}_2 - x_i\beta)\\
&\Pr ob(y_i < 3/x_i) = \mathrm{F}(\mathrm{r}_3 - x_i\beta)\\
&\Pr ob(y_i < 4/x_i) = \mathrm{F}(\mathrm{r}_4 - x_i\beta)
\end{aligned} \qquad 15.$$

El intercepto no se encuentra en las ecuaciones dado que se ha asumido que β0 = 0, cada curva de probabilidad diferirá únicamente en su inclinación hacia la derecha o izquierda, es decir, son paralelas como consecuencia de que el parámetro β es el mismo en cada ecuación. De esta manera el supuesto de paralelismo implica que, β1 = β2 =... = βJ−1. El grado de paralelismo se asume con parámetros muy cercanos entre sí.

El comando "omodel", de Wolfe y Gould (1998) calcula una aproximación del test LR, en el que se compara el logaritmo de la verosimilitud del ologit (o oprobit) para la obtención de un set de J-1 modelos binarios estimados con ologit (o oprobit), haciendo un ajuste para la correlación entre las salidas binarias definidas por y ≤ m.

La hipótesis nula será: Existencia del paralelismo en la regresión.

findit omodel
Para realizar la instalación del comando
omodel logit warm yr89 male white age ed prst

```
. omodel logit warm yr89 male white age ed prst

Iteration 0:   log likelihood = -2995.7704
Iteration 1:   log likelihood = -2846.4532
Iteration 2:   log likelihood = -2844.9142
Iteration 3:   log likelihood = -2844.9123

Ordered logit estimates                           Number of obs   =       2293
                                                  LR chi2(6)      =     301.72
                                                  Prob > chi2     =     0.0000
Log likelihood = -2844.9123                       Pseudo R2       =     0.0504

------------------------------------------------------------------------------
        warm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        yr89 |   .5239025   .0798988     6.56   0.000     .3673037    .6805013
        male |  -.7332997   .0784827    -9.34   0.000    -.8871229   -.5794766
       white |  -.3911595   .1183808    -3.30   0.001    -.6231815   -.1591374
         age |  -.0216655   .0024683    -8.78   0.000    -.0265032   -.0168278
          ed |   .0671728    .015975     4.20   0.000     .0358624    .0984831
        prst |   .0060727   .0032929     1.84   0.065    -.0003813    .0125267
-------------+----------------------------------------------------------------
       _cut1 |  -2.465362   .2389126          (Ancillary parameters)
       _cut2 |   -.630904   .2333155
       _cut3 |   1.261854   .2340179
------------------------------------------------------------------------------

Approximate likelihood-ratio test of proportionality of odds
across response categories:
         chi2(12) =      48.91
       Prob > chi2 =     0.0000

.
```

Uno no puede determinar si el coeficiente de algunas variables es idénticos a lo largo de las ecuaciones binarias, mientras que los coeficientes de otras variables difieren. Al final un test de wald elaborado por Brant (1990) es útil pues el test asume el paralelismo de la regresión para cada variable individual.

ologit warm yr89 male white age ed prst,nolog
brant,detail

```
. brant,detail

Estimated coefficients from j-1 binary regressions

                 y>1         y>2         y>3
 yr89      .9647422   .56540626   .31907316
 male    -.30536425  -.69054232  -1.0837888
white    -.55265759  -.31427081  -.39299842
  age     -.0164704  -.02533448  -.01859051
   ed     .10479624   .05285265   .05755466
 prst    -.00141118   .00953216   .00553043
_cons     1.8584045   .73032873  -1.0245168

Brant Test of Parallel Regression Assumption

    Variable |      chi2   p>chi2    df
-------------+--------------------------
         All |     49.18    0.000    12
-------------+--------------------------
        yr89 |     13.01    0.001     2
        male |     22.24    0.000     2
       white |      1.27    0.531     2
         age |      7.38    0.025     2
          ed |      4.31    0.116     2
        prst |      4.33    0.115     2
----------------------------------------

A significant test statistic provides evidence that the parallel
regression assumption has been violated.

.
```

La chi-cuadrado de 49.18 de Brant es muy cercano al valor de 48.91 del test LR, sin embargo, Brant muestra que las mayores violaciones son por yr89 y male, las cuales producen el problema.

V. ANÁLISIS DE PROBABILIDADES Y CAMBIOS MARGINALES

El MRO es no lineal, entonces, no hay una sola aproximación que pueda describir totalmente la relación entre una variable y las probabilidades, por lo tanto, se debería considerar cada uno de estos métodos antes de decidir que aproximación es más efectiva en nuestra aplicación. En el MRO, $y^{*} = x\beta + \varepsilon$,el cambio marginal en y ∗ con respecto a xk es: Siendo y una variable latente (cuya medida es desconocida), el cambio marginal no puede ser interpretado sin la estandarización, mediante la desviación estándar de y*.

ologit warm yr89 male white age ed prst,nolog
listcoef, std help
Voy obtener una tabla de efectos marginales en términos de ratios de odds.

```
. listcoef,std help

ologit (N=2293): Unstandardized and Standardized Estimates

 Observed SD: .9282156
   Latent SD: 1.9410634

-------------------------------------------------------------------------------
        warm |       b         z     P>|z|    bStdX    bStdY   bStdXY      SDofX
-------------+-----------------------------------------------------------------
        yr89 |   0.52390    6.557   0.000   0.2566   0.2699   0.1322     0.4897
        male |  -0.73330   -9.343   0.000  -0.3658  -0.3778  -0.1885     0.4989
       white |  -0.39116   -3.304   0.001  -0.1287  -0.2015  -0.0663     0.3290
         age |  -0.02167   -8.778   0.000  -0.3635  -0.0112  -0.1873    16.7790
          ed |   0.06717    4.205   0.000   0.2123   0.0346   0.1094     3.1608
        prst |   0.00607    1.844   0.065   0.0880   0.0031   0.0453    14.4923
-------------------------------------------------------------------------------
       b = raw coefficient
       z = z-score for test of b=0
   P>|z| = p-value for z-test
   bStdX = x-standardized coefficient
   bStdY = y-standardized coefficient
  bStdXY = fully standardized coefficient
   SDofX = standard deviation of X

.
```

Podemos observar que en 1989 el apoyo hacia las madres que trabajan fue de 0.27 desviaciones estándar mayores que en 1977, manteniendo las demás variables constantes.

Por cada desviación estándar en que se incremente la educación, se incrementa el apoyo para las madres que trabajan en 0.11 desviaciones estándar, manteniendo las demás variables constantes.

Predicción de Probabilidades

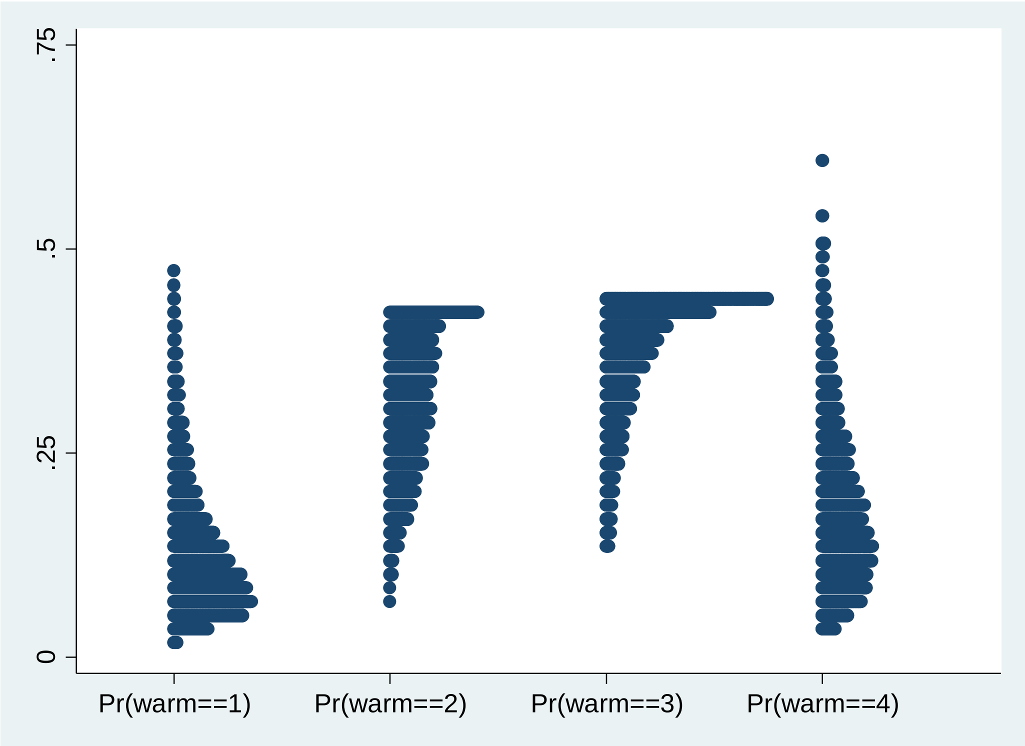
Predecimos las probabilidades como:

$$\hat{P}r(y = m/x) = F(\hat{r}_m - x\hat{\beta}) - F(\hat{r}_{m-1} - x\hat{\beta})$$

Con probabilidades acumuladas:

$$\hat{P}r(y \leq m/x) = F(r_m - x\hat{\beta})$$

predict fd_dlogit d_logit a_logit falogit
dotplot fdlogit dlogit alogit falogit, ylabel(0(.25).75)



Las probabilidades predichas para las categorías extremas tienden a ser menos que 0.25, la mayor cantidad de las predicciones para las categorías intermedias caen entre 0.25 y 0.5, solo unas cuantas tienden a ser mayores que 0.5.

VI. PREDICCIÓN DE PROBABILIDADES CON PRVALUE

La predicción de probabilidades para individuos con un conjunto de características puede ser calculada mediante "prvalue", por ejemplo, nosotros podríamos desear, examinar las probabilidades predichas para individuos con las siguientes características:

Hombres de la clase trabajadora en 1977 quienes están cerca de retirarse, es decir tiene una edad de 50 años.Mujeres jóvenes con elevada educación y prestigiosos trabajos. Individuo promedio en 1977. Individuo promedio en 1989.

ologit warm yr89 male white age ed prst

prvalue, x(yr89=0 male=1 prst=20 age=64 ed=16) rest(mean)

Estimo la probabilidad, para las cuatro categorías, de aquellas personas que son varones y son hombres, además tiene un nivel de prestigo de 20 y edad 64 años.

prvalue, x(yr89=1 male=0 prst=80 age=30 ed=24) rest(mean)
prvalue, x(yr89=0) rest(mean)
prvalue, x(yr89=1) rest(mean)

```
. prvalue, x( yr89=0 male=1 prst=20 age=64 ed=16) rest(mean)

ologit: Predictions for warm

Confidence intervals by delta method

                              95% Conf. Interval
     Pr(y=FD|x):      0.2317   [ 0.1776,    0.2857]
     Pr(y=D|x):       0.4221   [ 0.3942,    0.4500]
     Pr(y=A|x):       0.2723   [ 0.2249,    0.3198]
     Pr(y=FA|x):      0.0739   [ 0.0523,    0.0954]

        yr89      male     white        age         ed       prst
x=         0         1  .8765809         64         16         20

.
. prvalue, x( yr89=1 male=0 prst=80 age=30 ed=24) rest(mean)

ologit: Predictions for warm

Confidence intervals by delta method

                              95% Conf. Interval
     Pr(y=FD|x):      0.0164   [ 0.0106,    0.0222]
     Pr(y=D|x):       0.0781   [ 0.0554,    0.1008]
     Pr(y=A|x):       0.3147   [ 0.2636,    0.3658]
     Pr(y=FA|x):      0.5908   [ 0.5143,    0.6673]

        yr89      male     white        age         ed       prst
x=         1         0  .8765809         30         24         80

.
. prvalue, x( yr89=0) rest(mean)

ologit: Predictions for warm

Confidence intervals by delta method

                              95% Conf. Interval
     Pr(y=FD|x):      0.1336   [ 0.1176,    0.1496]
     Pr(y=D|x):       0.3577   [ 0.3348,    0.3806]
     Pr(y=A|x):       0.3737   [ 0.3517,    0.3957]
     Pr(y=FA|x):      0.1349   [ 0.1195,    0.1504]

         yr89        male       white         age          ed          pr
x=          0   .46489315    .8765809   44.935456   12.218055   39.5852

.
. prvalue, x( yr89=1) rest(mean)

ologit: Predictions for warm

Confidence intervals by delta method

                              95% Conf. Interval
     Pr(y=FD|x):      0.0837   [ 0.0711,    0.0963]
     Pr(y=D|x):       0.2802   [ 0.2571,    0.3032]
     Pr(y=A|x):       0.4277   [ 0.4046,    0.4507]
     Pr(y=FA|x):      0.2085   [ 0.1855,    0.2315]

         yr89        male       white         age          ed          pr
x=          1   .46489315    .8765809   44.935456   12.218055   39.5852

.
```

VII. PREDICCIÓN DE PROBABILIDADES CON PRTAB

En algunos casos nos puede ser de necesario calcular las probabilidades predichas para todas las combinaciones de un conjunto de variables independientes categóricas, por ejemplo, si estamos interesados en ver la importancia del género y de los años cuando las preguntas fueron realizadas:

prtab yr89 male

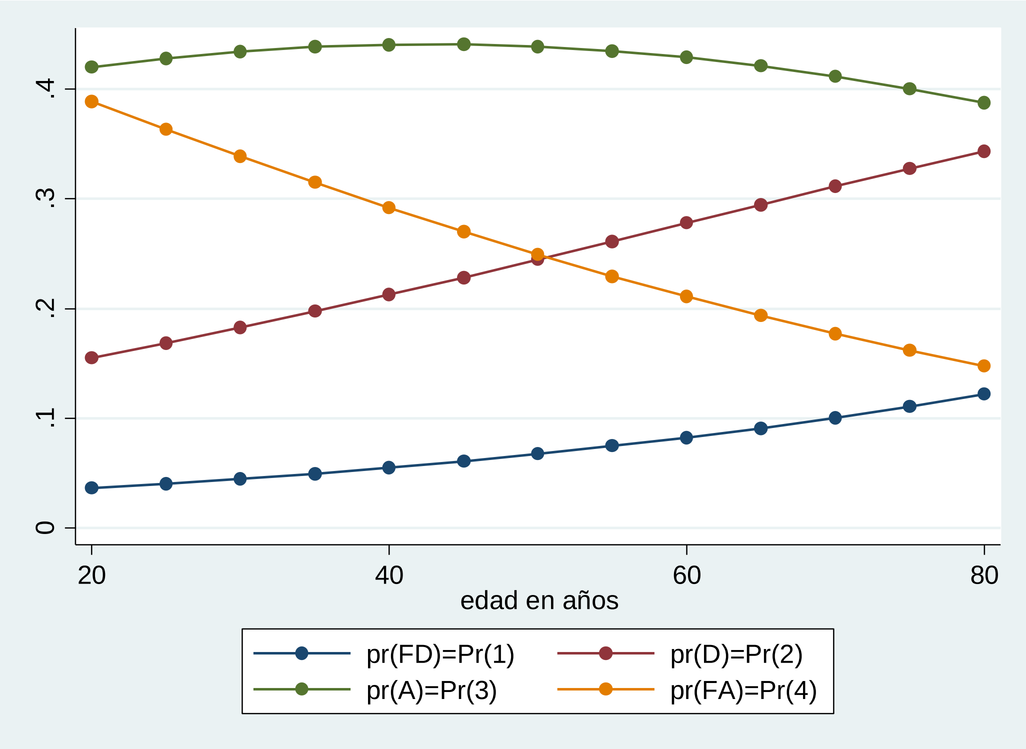
```
. prtab yr89 male

ologit: Predicted probabilities for warm

Predicted probability of outcome 1 (FD)

----------------------------------------
encuesta  |        genero:
anual:    |        1=hombre
1=1989    |        0=mujer
0=1977    |     mujer   hombre
----------+-----------------------------
     1977 |    0.0989   0.1859
     1989 |    0.0610   0.1191
----------------------------------------

Predicted probability of outcome 2 (D)

----------------------------------------
encuesta  |        genero:
anual:    |        1=hombre
1=1989    |        0=mujer
0=1977    |     mujer   hombre
----------+-----------------------------
     1977 |    0.3083   0.4026
     1989 |    0.2282   0.3394
----------------------------------------

Predicted probability of outcome 3 (A)

----------------------------------------
encuesta  |        genero:
anual:    |        1=hombre
1=1989    |        0=mujer
0=1977    |     mujer   hombre
----------+-----------------------------
     1977 |    0.4129   0.3162
     1989 |    0.4406   0.3904
----------------------------------------

Predicted probability of outcome 4 (FA)

----------------------------------------
encuesta  |        genero:
anual:    |        1=hombre
1=1989    |        0=mujer
0=1977    |     mujer   hombre
----------+-----------------------------
     1977 |    0.1799   0.0953
     1989 |    0.2703   0.1510
----------------------------------------

          yr89        male       white          a
x=   .39860445   .46489315    .8765809    44.935

.
```

Las salidas las podemos reorganizar en la siguiente tabla, donde se observa claramente como los hombres probablemente tienden a estar más en desacuerdo, comparados con las mujeres, al hecho de que las madres trabajadoras tiendan a tener una calidad relación con sus hijos como una madre que no trabaja. También se observa que entre 1977 y 1989 hubo un cambio en la opinión, tanto para hombres como para mujeres, hacia una actitud más positiva respecto a la pregunta (Orihuela, 2011):

VIII. PREDICCIÓN DE PROBABILIDADES CON PRGEN (PARA RALIZAR GRAFICOS DE PROBABILIDADES)

Grafiquemos las probabilidades lo cual nos será de mucha utilidad en los este tipo de modelos, por ejemplo si consideramos una mujer en 1989 y mostramos como las predicciones de sus probabilidades son afectadas por la edad:

```
prgen age, from(20) to (80) gen(Probw89) x(male=0 yr89=1) ncases(13)
desc w89*
graph tw sc Probw89p1 Probw89p2 Probw89p3 Probw89p4 Probw89x, connect(l l l
l)
```

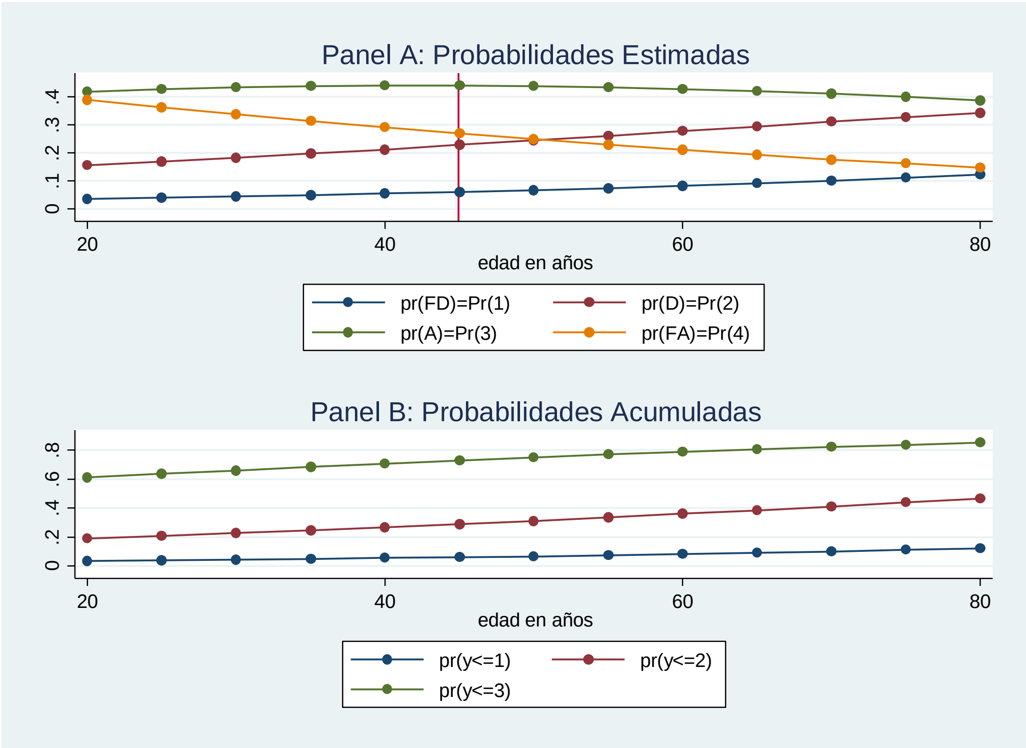


En este ejemplo "w98x" tendrá los valores de "age", para el rango de 20- 80, la p# variable contiene la predicción de la probabilidad para la opción # de la endógena. Cuando el modelo es ordinal, "prgen" también calcula las probabilidades acumuladas, las que son indicadas por w89s#, la cual es la suma de probabilidades para las características 1 y 2 (Orihuela, 2011).

En el gráfico, la edad de 44.93 marca el promedio en el Panel A. Observamos que cuando la edad se incrementa, la probabilidad de "SA", decrece rápidamente mientras que la probabilidad de "D" y "SD" se incrementan, la gráfica del Panel B muestra la probabilidad acumulada (Orihuela, 2011).

```
graph tw sc w89p1 w89p2 w89p3 w89p4 w89x, connect(l l l l) ///
xline(44.93) title(Panel A: Probabilidades Estimadas)
graph save temp1,replace
graph tw sc w89s1 w89s2 w89s3 w89x, connect(l l l l) ///

title(Panel B: Probabilidades Acumuladas)
graph save temp2,replace
graph combine temp1.gph temp2.gph ,col(1)
```

Cambios en las Probab



IX. PROBABILIDADES PREDICHAS CON PRCHANGE

Cuando existen múltiples variables en el modelo no es practico dibujarlas, para ello es útil usar prchange como resumen de los efectos de cada variable sobre la endógena. El cambio marginal en la probabilidad es calculado como:

$$\frac{\partial \Pr ob(y_i = m / x_i)}{\partial x_k} = \frac{\partial \mathrm{F}(\mathrm{r}_m - x_i\beta)}{\partial x_k} - \frac{\partial \mathrm{F}(\mathrm{r}_{m-1} - x_i\beta)}{\partial x_k} \qquad 16.$$

La cual es la pendiente de la curva que relaciona x_k a Pr(y = m/x), manteniendo las otras variables constantes. En nuestro ejemplo, nosotros consideraremos el efecto marginal de la edad, para mujeres en 1989, manteniendo en su media a las demás variables. Esto corresponde a la pendiente de las curvas en el Panel A del grafico anterior evaluado sobre la línea vertical. Con prchange, el calculo seria el siguiente (Orihuela, 2011):

prchange age, x(male=0 yr89 =1) rest(mean)

```
. prchange age,   x( male=0  yr89 =1) rest(mean)

ologit: Changes in Probabilities for warm

age
            Avg|Chg|         FD          D           A          FA
Min->Max  .16441458   .10941909   .21941006  -.05462247  -.27420671
   -+1/2  .00222661   .00124099   .00321223   -.0001803  -.00427291
  -+sd/2   .0373125    .0208976   .05372739  -.00300205  -.07162295
MargEfct  .00222662   .00124098   .00321226  -.00018032  -.00427292

               FD          D           A          FA
Pr(y|x)  .06099996  .22815652  .44057754  .27026597

          yr89     male    white      age       ed     prst
   x=        1        0  .876581  44.9355  12.2181  39.5853
sd_x=  .489718  .498875  .328989   16.779  3.16083  14.4923

.
```

Lo primero que debemos notar es la fila denotada por Pr(y/x), la cual es la probabilidad predicha para los valores fijados en x() y en rest(). En la fila de efectos marginales se listan las pendientes de las curvas de probabilidades en el punto de intersección con la línea vertical de la figura anterior (Orihuela, 2011).

Por ejemplo, la pendiente de "FD" es de 0.0124098, mientras que la pendiente de "A" es negativa y muy pequeña, pero no corresponde exactamente a la cantidad de cambio en probabilidad para el cambio en una unidad en la variable independiente. Sin embargo cuando la curva de probabilidad es aproximadamente lineal, el efecto marginal puede ser usado para resumir el efecto de una unidad de cambio en la variable exógena sobre la probabilidad de ocurrencia de un evento (Bravo & Vásquez, 2008).

El cambio marginal también puede ser analizado con mfx, este comando no calcula los efectos del conjunto de variables independientes y solo estima el efecto marginal para una categoría por vez, la cual es especificada en la opción predict(outcome(#)). Veamos ésto con una estimación del ologit y considerando las mismas variables (Beltran Barco, Arlette; Castro Carlin, Juan;, 2010).

mfx compute, at(male=1 yr89=1) predict(outcome(1))

```
. mfx compute, at(male=0 yr89=1) predict(outcome(1))

warning: no value assigned in at() for variables white age ed prst;
   means used for white age ed prst

Marginal effects after ologit
      y  = Pr(warm==1) (predict, outcome(1))
         =  .06099996
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
   yr89*|  -.0378526      .00601   -6.30   0.000  -.049633 -.026072         1
   male*|   .0581355      .00731    7.95   0.000   .043803  .072468         0
  white*|   .0197511       .0055    3.59   0.000   .008972   .03053   .876581
     age|    .001241      .00016    7.69   0.000   .000925  .001557   44.9355
      ed|  -.0038476      .00097   -3.96   0.000  -.005754 -.001941   12.2181
    prst|  -.0003478      .00019   -1.83   0.068  -.000721  .000025   39.5853
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

El impacto marginal de la edad es de 0.001241 como lo muestra la figura lo cual es comparable con el resultado obtenido con "prchange". La ventaja de usar "mfx" es que podemos obtener las desviaciones estándar inherentes a cada cambio marginal (Greene, 1997).

Bibliografía

Colin Cameron , A., & Trivedi, P. (2005). *Microeconometrics: Methods and Applications.* (C. U. Press, Ed.) New York.

Orihuela, A. (2011). *Stata Avanzado Aplicado a la Investigación Económica.* Grupo Iddea, Lima.

Beltran Barco, A. (2001). *Econometria de Corte Transversal.* Notas de Clase.

Beltran Barco, Arlette; Castro Carlin, Juan;. (2010). *Modelos de datos de panel y variables dependientes limitadas: teoría y práctica.* (U. d. Pacífico, Ed.)

Bernardí Cabrer Borrás, & Amparo Sancho Pérez, Guada. (2001). *Microeconometría y Decisión.* Ediciones Pirámide, .

Bravo, D., & Vásquez, J. (2008). *Microeconometria Aplicada.* Notas de Clase, Centro Micro Datos., Santiago. Obtenido de http://www.academia.edu/9494003/MICROECONOMETR%C3%8DA_CON_STATA

Greene, W. (1997). *Análisis Econometrico* (Tercera ed.). Prentice Hall.