Serie Apuntes de Clase ΩΒΓ

N° 09
Noviembre 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.

ΩBΓ

CURIOSITAS, DUBITARE, INVESTIGARE

Modelos de Elección Nominal.

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 Multinomial, con aplicaciones a Stata 13. 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.INTRODUCION

Una elección o salida es nominal[1] cuando se asumen categorías no ordenadas. Por ejemplo, el estado marital puede ser agrupado nominalmente como divorciado, no casado, casado o viudo. Las ocupaciones puedes ser organizadas como profesional, empleador (trabajador de oficina o contratante), obrero (o trabajador técnico), artesano y sirviente. En algunos casos se suele tratar las salidas nominales como ordenadas o no ordenadas, por ejemplo, si tu respuesta fuese Totalmente de acuerdo, de acuerdo, en desacuerdo, totalmente en desacuerdo y no sabe no opina, la categoría "no sabe no opina", invalida el modelo ordinal. Se podría decidir usar un modelo de regresión nominal cuando el supuesto del paralelismo de la regresión es rechazada. En general, si uno es consciente del ordenamiento de la variable dependiente, la perdida potencial de eficiencia en el uso de modelos de salida nominal es mayor que la ganancia por evitar el sesgo.

Por ejemplo, queremos conocer cuáles son los aspectos que explican el auto reporte de salud de los individuos. En la Encuesta de Protección Social de Chile, por ejemplo, los entrevistados ordenan su salud en 6 categorías: 1) Excelente; 2) Muy buena; 3) Buena; 4) Regular; 5) Mala; y 6) Muy mala. Se trata, obviamente, de una variable ordinal puesto que al ir desde 1 a 6 claramente la salud de la persona es peor, y viceversa. Otra variable ordinal que podría ser objeto de nuestro interés, es el número de personas que trabajan en la empresa. En la encuesta CASEN[2], por ejemplo, la respuesta a esta pregunta consta de las siguientes categorías: A) 1 persona; B) de 2 a 5 personas; C) de 6 a 9 personas; D) de 10 a 49 personas; E) de 50 a 199 personas; F) 200 y más personas.

[1] Encuestas realizadas en Chile
[2] Encuestas realizadas en Chile

El modelo logit multinomial que es uno de los más usados en los modelos de regresión multinomial. La dificultad de usarlo es que este incluye muchos de parámetros, y fácilmente podría estar sobreestimado. Además de que existe dificultad nace por el cálculo no lineal del modelo lo cual conduce a problemas de interpretación.

Nuestro punto de partida es el proceso a través del cual una persona escoge entre diferentes alternativas de acuerdo con aquella que le dé la utilidad más alta. Esta utilidad no es directamente observable pero suponemos que se puede representar como una función lineal de un conjunto de determinantes (Orihuela, 2011).

II. MODELO LOGIT MULTINOMIAL

El modelo puede ser imaginado como una estimación simultánea y binaria de logits, para todas las comparaciones posibles de categorías dependientes.

Por ejemplo, dejemos que "ocupación" sea una salida nominal con la categoría "S" para trabajos manuales, "E" para trabajos de oficina o empleadores, y "P" para trabajos profesionales. Asumimos que hay una simple variable independiente que mide los años de educación "Ed" (Orihuela, 2011).

Tomando el primer ejemplo, supongamos que se desea analizar los determinantes del tipo de ocupación, de forma tal que la variable dependiente se define como:

$$Y_i = \text{Ocupación del Jefe de Hogar}$$
$$= \begin{cases} 1 \text{ Sirviente} \\ 2 \text{ Obrero} \\ 3 \text{ Artesano} \\ 4 \text{ Empleador} \\ 5 \text{ Profesional} \end{cases} \qquad (1)$$

De esta manera, se tiene en total j categorías no ordenadas ya que, a priori, no se puede establecer cuáles de ellas pueden ser consideradas mejores que otras[3]. Lo que sí suponemos es que cada agente de la muestra ha elegido la opción o categoría que le reporta mayor utilidad y que, de acuerdo con lo especificado en (1.), esta puede ser representada como:

$$Y^*_{ij} = U_{ij}^* = X_{ij}\beta_j + \varepsilon_{ij} \qquad (2)$$

La probabilidad de que el i-ésimo agente elija la k-ésima categoría corresponde a la Probabilidad de que esta sea la que mayor utilidad le brinda. Formalmente (Beltran Barco, Arlette; Castro Carlin, Juan;, 2010):

$$\begin{aligned} &\Pr(Y_i = k) = \Pr(Y^*_{ik} > Y^*_{ij}) \quad \forall \mathrm{j} \neq \mathrm{k} \\ &= \Pr(\mathrm{X}_{ik}\,\beta_k + \varepsilon_{ij} > \mathrm{X}_{ij}\,\beta_j + \varepsilon_{ik}) \ \forall \mathrm{j} \neq \mathrm{k} \\ &= \Pr(\varepsilon_{ij} - \varepsilon_{ik} > \mathrm{X}_{ik}\,\beta_k - \mathrm{X}_{ij}\,\beta_j) \ \forall \mathrm{j} \neq \mathrm{k} \end{aligned} \qquad (3)$$

Para facilitar el análisis, y dado que las categorías no pueden ser relacionadas de

[3] Muchas veces es difícil determinar si las categorías de elección son efectivamente ordenadas o no, o quizás tienen la condición de secuencialidad que veremos más adelante. En ese caso, será mejor elegir el modelo menos restrictivo, es decir, realizar la estimación como si se tratara de categorías no ordenadas.

acuerdo con algún ordenamiento específico, resulta conveniente elegir una categoría base o referencial[4].

A partir de ella (a la que llamaremos "categoría m") se puede especificar la probabilidad de escoger alguna de las otras categorías, utilizando un conjunto de-modelos binomiales donde las opciones por elegir son solo la categoría en cuestión y la base.

La probabilidad de elegir la k-ésima alternativa en este contexto binomial corresponde a la probabilidad que se obtiene del modelo multinomial, pero reescalada tomando en cuenta solo las dos categorías. Formalmente (Beltran Barco, Arlette; Castro Carlin, Juan;, 2010):

$$\frac{\Pr(Y_i = k)}{\Pr(Y_i = k) + \Pr(Y_i = m)} = F(\mathrm{X}_{ik}\,\beta_k) \tag{4}$$

En la expresión anterior, F (.) corresponde a la función de densidad de la diferencia de los errores de las ecuaciones explicativas de la utilidad que reportan las alternativas k y la m. Esto último se deriva de la expresión (3.) si tenemos en cuenta que los coeficientes de la categoría base han sido normalizados en cero. Es decir, corresponde a la probabilidad de que la categoría k reporte más utilidad que la categoría base:

$$\Pr(Y^*_{ik} > Y^*_{im}) = \Pr(\varepsilon_{ik} - \varepsilon_{im} > \mathrm{X}_{im}\,\beta_m - \mathrm{X}_{ik}\,\beta_k)\ \forall \mathrm{j} \neq m \tag{5}$$

Pero dado β_m que esta normalizado en cero tenemos:

$$\Pr(Y^*_{ik} > Y^*_{ij}) = \Pr(\varepsilon_{ik} - \varepsilon_{im} > -\mathrm{X}_{ik}\,\beta_k)\ \forall \mathrm{j} \neq m \tag{6}$$

[4] La elección de la categoría base no resulta ser un procedimiento trivial, dado que la interpretación de resultados se hará tomándola como referencia. Por ello, generalmente se escoge como base una categoría neutral (en el ejemplo, no tener ocupación) o aquella que es el centro de interés del investigador

Lo anterior implica que:

$$\frac{\Pr(Y_i = k)}{\Pr(Y_i = m)} = \frac{F(\mathrm{X}_{ik}\,\beta_k)}{1 - F(\mathrm{X}_{ik}\,\beta_k)} = G(\mathrm{X}_{ik}\,\beta_k) \tag{7}$$

De donde se puede derivar la probabilidad de escoger la categoría m si es que evaluamos la sumatoria sobre todas las categorías menos la base:

$$\sum_{k=1}^{J-1} \frac{\Pr(Y_i = k)}{\Pr(Y_i = m)} = \sum_{k=1}^{J-1} \frac{F(\mathrm{X}_{ik}\,\beta_k)}{1 - F(\mathrm{X}_{ik}\,\beta_k)} = \sum_{k=1}^{J-1} G(\mathrm{X}_{ik}\,\beta_k) \tag{8}$$

$$\sum_{k=1}^{J-1} \frac{\Pr(Y_i = k)}{\Pr(Y_i = m)} = \frac{1 - \Pr(Y_i = m)}{\Pr(Y_i = m)} = \frac{1}{\Pr(Y_i = m)} - 1 = \sum_{j=k}^{J-1} G(\mathrm{X}_{ik}\,\beta_k) \tag{9}$$

$$\Pr(Y_i = m) = \frac{1}{1 + \sum_{k=1}^{J-1} G(\mathrm{X}_{ik}\,\beta_k)} \tag{10}$$

A partir de la expresión anterior es posible hallar la probabilidad de escoger una alternativa k cualquiera:

$$\Pr(Y_i = k) = G(\mathrm{X}_{ik}\,\beta_k)\Pr(Y_i = m) = \frac{G(\mathrm{X}_{ik}\,\beta_k)}{1 + \sum_{k=1}^{J-1} G(\mathrm{X}_{kj}\,\beta_k)} \tag{11}$$

Como ya se mencionó, el modelo no ordenado más utilizado por su simplicidad operativa es el logit multinomial. Para esto, suponemos

$$F(\mathrm{X}_{ik}\,\beta_k) = \frac{\exp(\mathrm{X}_{ik}\,\beta_k)}{1 + \exp(\mathrm{X}_{ik}\,\beta_k)} \tag{12}$$

Asumo que la función de Distribución Acumulativa es una Logit. Tomando en cuenta este resultado, tenemos:

$$\Pr(Y_i = k) = \frac{\exp(X_{ik}\,\beta_k)}{1+\sum_{k=1}^{J-1}\exp(X_{kj}\,\beta_k)} \tag{13}$$

Nótese que la especificación anterior es lo suficientemente general como para admitir un conjunto distinto de variables explicativas y parámetros para cada categoría. Frente a esto, es común suponer que existe un único conjunto de regresores (o características) y un vector de coeficientes distinto para cada categoría5. Esto es suficiente para explicar cómo un mismo agente con características X_j deriva un nivel de utilidad distinto de cada categoría, y que estos niveles no tienen por qué ser los mismos que los de otro individuo en la muestra con características X_j. Cada agente, por tanto, puede maximizar su utilidad eligiendo una categoría J particular, la que solo tiene que ser igual a aquella que elija otro individuo con las mismas características.

Si introducimos esta simplificación, es posible reexpresar la probabilidad de elegir la k-ésima alternativa como:

$$\Pr(Y_i = k) = \frac{\exp(X_i\,\beta_k)}{1+\sum_{j=1}^{J-1}\exp(X_i\,\beta_j)} \tag{14}$$

A partir de la expresión anterior, se puede construir el ratio de probabilidad (RP) para dos categorías cualesquiera (Ratio de Odds):

$$RP(k,k+1) = \frac{\Pr(Y_i = k)}{\Pr(Y_i = k+1)} = \exp(X_i(\beta_k - \beta_{k+1})) \tag{15}$$

[5] Para un mismo agente, las características reciben pesos distintos para construir el nivel de utilidad asociado a cada categoría.

$$RP(k,\mathrm{m}) = \frac{\Pr(Y_i = k)}{\Pr(Y_i = m)} = \exp(\mathrm{X}_i\,\beta_k) \tag{16}$$

$$\ln RP(k,\mathrm{m}) = \mathrm{X}_{ki}\,\beta_k \tag{17}$$

La estimación de este modelo implica obtener un total de J-1 ecuaciones, una para cada categoría, excepto la categoría base. A cada ecuación corresponde un vector de coeficientes (β_k) y, de acuerdo con la expresión anterior, cada coeficiente recoge el efecto de un cambio marginal en el regresor correspondiente sobre el logaritmo del ratio de probabilidades de la k-ésima categoría respecto a la categoría base. Por lo mismo, si el i-ésimo regresor se incrementa en una unidad, el RP de la k-ésima categoría respecto a la categoría base se incrementa en ($\exp^{\beta_{ik}} - 1$) *100 por ciento. (Orihuela, 2011)

III. PROPIEDAD DE ALTERNATIVAS IRRELEVANTES

Se refiere a que cada una de las alternativas posibles esta netamente diferenciada y valorada de manera independiente por el que toma la decisión. Supuesto altamente restrictivo que es difícil de cumplirse en la evaluación empírica de los modelos.

Una de las principales desventajas de esta clase de modelos es que se ve afectado por lo que se conoce como la propiedad de independencia de alternativas irrelevantes (IIA, por sus siglas en inglés). Si divido una categoría ya existente en dos muy parecidas, debería esperarse que ambas se repartieran la probabilidad de ser escogida que antes tenía la que ya estaba presente, mientras que el resto de alternativas mantuvieran la misma probabilidad de ser elegidas. No obstante, y de acuerdo con la propiedad de IIA, el modelo logit multinomial reasigna las probabilidades de ocurrencia entre el total de categorías existentes, incluyendo la nueva. **Por lo mismo, no es apropiado cuando se sabe que se tienen categorías que son sustitutas cercanas** (Orihuela, 2011).

De acuerdo con la propiedad de IIA, la aplicación del modelo multinomial no ordenado logístico supone que el ratio de probabilidades entre dos alternativas no depende de las demás categorías. Para verificar si la inclusión de determinada categoría afecta la consistencia de nuestros estimados (y, con esto, los ratios de probabilidad), es posible utilizar una prueba de la clase de Hausman (Greene, 1997).

Basados en este resultado, es evidente que la presencia de una categoría base nos permite una interpretación directa del signo y magnitud de determinado coeficiente como el efecto que tiene el regresar en cuestión sobre la probabilidad de elegir la k-ésima alternativa respecto a la categoría base. Para esto, basta recordar que los coeficientes de la categoría base han sido normalizados en cero. Mientras las probabilidades predichas serán obtenidas con la categoría base (m), cambiar la base de la categoría podría confundir a algunos, dado que los resultados de los parámetros tienden a ser algo diferentes. Solo habría un cambio en la parametrización mas no en la estimación de las probabilidades predichas, dado que estas serán las mismas, sea cual sea la categoría base (Beltran Barco, Arlette; Castro Carlin, Juan;, 2010).

IV. APLICACIONES EN STATA

En 1982 General Social Survey, pregunto a 337 personas sobre su nivel de ocupación, categorizando de cinco maneras las respuestas: Sirviente“S”, Obrero “O”, Artesano“A”, Empleado “E” y Profesional“P”. Tres variables independientes son consideradas, “raza” que indica raza del encuestado, “ed” que indica años de educación del encuestado y “exper” que mide los años de experiencia laboral (Orihuela, 2011)[6].

```
. describe

Contains data from ocupacion.dta
  obs:           337                          1982 General Social Survey
 vars:             4                          18 Nov 2009 00:05
 size:         1,348                          (_dta has notes)
-----------------------------------------------------------------------------
              storage  display    value
variable name   type   format     label      variable label
-----------------------------------------------------------------------------
ocupacion       byte   %11.0g     ocupa      ocupación
raza            byte   %10.0g     raz        raza, 1=blanco
ed              byte   %10.0g                años de educación
exper           byte   %10.0g                años de experiencia
-----------------------------------------------------------------------------
Sorted by:  ocupacion

. sum

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   ocupacion |       337    3.397626    1.367913          1          5
        raza |       337    .9169139    .2764227          0          1
          ed |       337    13.09496    2.946427          3         20
       exper |       337    20.50148    13.95936          2         66
```

[6] Puede obtener la base de datos a través del siguiente link: https://www.dropbox.com/s/f7qs0nvacvjg0hr/ocupacion_stata13.dta?dl=0

```
end of do-file

. tab ocupacion, missing

  ocupación |      Freq.     Percent        Cum.
------------+-----------------------------------
  sirviente |         31        9.20        9.20
     obrero |         69       20.47       29.67
   artesano |         84       24.93       54.60
  empleador |         41       12.17       66.77
profesional |        112       33.23      100.00
------------+-----------------------------------
      Total |        337      100.00
```

mlogit ocupacion raza ed exper,b(5), nolog

```
. mlogit ocupacion raza ed exper, b(5) nolog

Multinomial logistic regression                   Number of obs   =        337
                                                  LR chi2(12)     =     166.09
                                                  Prob > chi2     =     0.0000
Log likelihood = -426.80048                       Pseudo R2       =     0.1629

------------------------------------------------------------------------------
   ocupacion |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sirviente    |
        raza |  -1.774306   .7550543    -2.35   0.019    -3.254186   -.2944273
          ed |  -.7788519   .1146293    -6.79   0.000    -1.003521   -.5541826
       exper |  -.0356509    .018037    -1.98   0.048    -.0710028    -.000299
       _cons |   11.51833   1.849356     6.23   0.000     7.893659      15.143
-------------+----------------------------------------------------------------
obrero       |
        raza |  -.5378027   .7996033    -0.67   0.501    -2.104996    1.029391
          ed |  -.8782767   .1005446    -8.74   0.000     -1.07534   -.6812128
       exper |  -.0309296   .0144086    -2.15   0.032     -.05917    -.0026893
       _cons |   12.25956   1.668144     7.35   0.000     8.990061    15.52907
-------------+----------------------------------------------------------------
artesano     |
        raza |  -1.301963    .647416    -2.01   0.044    -2.570875   -.0330509
          ed |  -.6850365   .0892996    -7.67   0.000    -.8600605   -.5100126
       exper |  -.0079671   .0127055    -0.63   0.531    -.0328693    .0169351
       _cons |   10.42698   1.517943     6.87   0.000     7.451864    13.40209
-------------+----------------------------------------------------------------
empleador    |
        raza |  -.2029212   .8693072    -0.23   0.815    -1.906732     1.50089
          ed |  -.4256943   .0922192    -4.62   0.000    -.6064407   -.2449479
       exper |   -.001055   .0143582    -0.07   0.941    -.0291967    .0270866
       _cons |   5.279722   1.684006     3.14   0.002     1.979132    8.580313
-------------+----------------------------------------------------------------
profesional  |  (base outcome)
------------------------------------------------------------------------------

.
end of do-file

.
```

Por defecto mlogit deja como categoría base a la variable con mayor cantidad de observaciones. De otra forma, uno puede seleccionar la categoría base con el comando **"basecategory()".** Si nuestro interés es saber cómo la raza afecta la ubicación de los trabajadores entre artesanos y sirvientes, lo cual no fue estimado en la salida anterior, pero puede ser calculado estimando el modelo mlogit con una categoría diferente, sin embargo es más fácil usar "listcoef", el cual presenta las estimaciones para todas las combinaciones de categorías (Orihuela, 2011).

mlogit (N=337): Factor Change in the Odds of ocupacion

Variable: raza (sd=.27642268)

Odds comparing Alternative 1 to Alternative 2	b	z	P>$\lvert z\rvert$	e^b	e^bStdX
sirvient-obrero	-1.23650	-1.707	0.088	0.2904	0.7105
sirvient-artesano	-0.47234	-0.782	0.434	0.6235	0.8776
sirvient-empleado	-1.57139	-1.741	0.082	0.2078	0.6477
sirvient-profesio	-1.77431	-2.350	0.019	0.1696	0.6123
obrero -sirvient	1.23650	1.707	0.088	3.4436	1.4075
obrero -artesano	0.76416	1.208	0.227	2.1472	1.2352
obrero -empleado	-0.33488	-0.359	0.720	0.7154	0.9116
obrero -profesio	-0.53780	-0.673	0.501	0.5840	0.8619
artesano-sirvient	0.47234	0.782	0.434	1.6037	1.1395
artesano-obrero	-0.76416	-1.208	0.227	0.4657	0.8096
artesano-empleado	-1.09904	-1.343	0.179	0.3332	0.7380
artesano-profesio	-1.30196	-2.011	0.044	0.2720	0.6978
empleado-sirvient	1.57139	1.741	0.082	4.8133	1.5440
empleado-obrero	0.33488	0.359	0.720	1.3978	1.0970
empleado-artesano	1.09904	1.343	0.179	3.0013	1.3550
empleado-profesio	-0.20292	-0.233	0.815	0.8163	0.9455
profesio-sirvient	1.77431	2.350	0.019	5.8962	1.6331
profesio-obrero	0.53780	0.673	0.501	1.7122	1.1603
profesio-artesano	1.30196	2.011	0.044	3.6765	1.4332
profesio-empleado	0.20292	0.233	0.815	1.2250	1.0577

Variable: ed (sd=2.9464271)

Odds comparing Alternative 1 to Alternative 2	b	z	P>$\lvert z\rvert$	e^b	e^bStdX
sirvient-obrero	0.09942	0.972	0.331	1.1045	1.3404
sirvient-artesano	-0.09382	-0.962	0.336	0.9105	0.7585
sirvient-empleado	-0.35316	-3.011	0.003	0.7025	0.3533
sirvient-profesio	-0.77885	-6.795	0.000	0.4589	0.1008
obrero -sirvient	-0.09942	-0.972	0.331	0.9054	0.7461
obrero -artesano	-0.19324	-2.494	0.013	0.8243	0.5659
obrero -empleado	-0.45258	-4.425	0.000	0.6360	0.2636
obrero -profesio	-0.87828	-8.735	0.000	0.4155	0.0752
artesano-sirvient	0.09382	0.962	0.336	1.0984	1.3184
artesano-obrero	0.19324	2.494	0.013	1.2132	1.7671
artesano-empleado	-0.25934	-2.773	0.006	0.7716	0.4657
artesano-profesio	-0.68504	-7.671	0.000	0.5041	0.1329
empleado-sirvient	0.35316	3.011	0.003	1.4236	2.8308
empleado-obrero	0.45258	4.425	0.000	1.5724	3.7943
empleado-artesano	0.25934	2.773	0.006	1.2961	2.1471
empleado-profesio	-0.42569	-4.616	0.000	0.6533	0.2853
profesio-sirvient	0.77885	6.795	0.000	2.1790	9.9228
profesio-obrero	0.87828	8.735	0.000	2.4067	13.3002
profesio-artesano	0.68504	7.671	0.000	1.9838	7.5264
profesio-empleado	0.42569	4.616	0.000	1.5307	3.5053

Variable: exper (sd=13.959364)

Odds comparing Alternative 1 to Alternative 2	b	z	P>$\lvert z\rvert$	e^b	e^bStdX
sirvient-obrero	-0.00472	-0.271	0.786	0.9953	0.9362
sirvient-artesano	-0.02768	-1.660	0.097	0.9727	0.6795
sirvient-empleado	-0.03460	-1.837	0.066	0.9660	0.6170
sirvient-profesio	-0.03565	-1.977	0.048	0.9650	0.6079
obrero -sirvient	0.00472	0.271	0.786	1.0047	1.0681
obrero -artesano	-0.02296	-1.829	0.067	0.9773	0.7258
obrero -empleado	-0.02987	-1.954	0.051	0.9706	0.6590
obrero -profesio	-0.03093	-2.147	0.032	0.9695	0.6494
artesano-sirvient	0.02768	1.660	0.097	1.0281	1.4717
artesano-obrero	0.02296	1.829	0.067	1.0232	1.3779
artesano-empleado	-0.00691	-0.495	0.621	0.9931	0.9080
artesano-profesio	-0.00797	-0.627	0.531	0.9921	0.8947
empleado-sirvient	0.03460	1.837	0.066	1.0352	1.6208
empleado-obrero	0.02987	1.954	0.051	1.0303	1.5174
empleado-artesano	0.00691	0.495	0.621	1.0069	1.1013
empleado-profesio	-0.00106	-0.073	0.941	0.9989	0.9854
profesio-sirvient	0.03565	1.977	0.048	1.0363	1.6449
profesio-obrero	0.03093	2.147	0.032	1.0314	1.5400
profesio-artesano	0.00797	0.627	0.531	1.0080	1.1176
profesio-empleado	0.00106	0.073	0.941	1.0011	1.0148

Por ejemplo, cuando la variable explicativa experiencia se incremente en una unidad, el coeficiente sirviente-obrero puede ser interpretado como la probabilidad de ser sirviente es 0.559 veces, respecto a la probabilidad de ser overo, manteniendo las demás variables explicativas constantes.

```
. listcoef raza, pvalue(0.05) help

mlogit (N=337): Factor Change in the Odds of ocupacion when P>|z| < 0.05

Variable: raza (sd=.27642268)

Odds comparing    |
Alternative 1     |
to Alternative 2  |        b         z     P>|z|      e^b    e^bStdX
------------------+------------------------------------------------------
sirvient-profesio |  -1.77431    -2.350    0.019    0.1696    0.6123
artesano-profesio |  -1.30196    -2.011    0.044    0.2720    0.6978
profesio-sirvient |   1.77431     2.350    0.019    5.8962    1.6331
profesio-artesano |   1.30196     2.011    0.044    3.6765    1.4332
-------------------------------------------------------------------------
       b = raw coefficient
       z = z-score for test of b=0
   P>|z| = p-value for z-test
     e^b = exp(b) = factor change in odds for unit increase in X
 e^bStdX = exp(b*SD of X) = change in odds for SD increase in X
```

4.1 PRUEBAS DE HIPÓTESIS

En el Modelo No Lineal Multinomial (MNLM) uno puede testear los coeficientes de manera individual y reportar los z-statistic, con un test de Wald o con un test LR. Hay buenas razones para testear los coeficientes de manera grupal. Al testear el hecho de que una variable no tenga efectos requiere que el test para J-1 coeficientes sea simultáneamente iguales a cero. Luego el testeo de que las variables independientes como un grupo sean diferentes entre estimaciones, requiere un test de K coeficientes (Bravo & Vásquez, 2008).

4.1.1 TESTEO DE EFECTOS DE VARIABLES INDEPENDIENTES

Con J categorías dependientes, hay J-1 coeficientes no redundantes asociados con cada variable independiente. Por ejemplo para nuestro logit de ocupación hay cuatro coeficientes asociados con educación “ed”, $\beta_{2,M/P},\beta_{2,B/P},\beta_{2,C/P},\beta_{2,W/P}$ La hipótesis de que X_k no tiene efectos sobre la variable dependiente puede ser escrita como:

$$
\begin{aligned}
&H_o : \beta_{2,\mathrm{S}/P} = \beta_{2,\mathrm{O}/P} = \beta_{2,\mathrm{A}/P} = \beta_{2,\mathrm{E}/P} = 0 \\
&H_o : \beta_{\mathrm{exper,S}/P} = \beta_{\mathrm{exper,O}/P} = \beta_{\mathrm{exper,A}/P} = \beta_{\mathrm{exper,E}/P} = 0 \\
&H_o : \beta_{\mathrm{raza,S}/P} = \beta_{\mathrm{raza,O}/P} = \beta_{\mathrm{raza,A}/P} = \beta_{\mathrm{raza,E}/P} = 0
\end{aligned}
\qquad (18)
$$

Donde b es la categoría base y como $\beta_{k,b/b} = 0$, la hipótesis impone restricciones de J-1 parámetros, esta hipótesis puede ser falseada con los test de LR y Wald.

Test de Ratio de Verosimilitud

El LR-test involucra:

- Estimar el modelo completo incluyendo todas las variables, resultantes en el estadístico del ratio de verosimilitud LRsr.
- Estimar el modelo restricto excluyendo las variables xk, y obtener el ratio de verosimilitud LRr.
- Calcular la diferencia LR = LRsr ◎ LRr, el cual es distribuido como una chi-cuadrado con J-1 grados de libertad.

El cálculo puede ser hecho con el comando "lrtest"

```
mlogit ocupacion raza ed exper, b(5) nolog
lrtest, saving(0)
mlogit ocupacion ed exper, b(5) nolog
lrtest
```

```
Likelihood-ratio test                          LR chi2(4)  =      8.10
(Assumption: . nested in LRTEST_0)             Prob > chi2 =    0.0881

.
end of do-file

.

.  mlogtest,lr

**** Likelihood-ratio tests for independent variables (N=337)

 Ho: All coefficients associated with given variable(s) are 0.
```

	chi2	df	P>chi2
raza	8.095	4	0.088
ed	156.937	4	0.000
exper	8.561	4	0.073

mlogit ocupacion raza ed exper, nolog

mlogtest,lr

El efecto de la raza en la ocupación es significativa al 0.1 de significancia, pero no lo es al 0.05. El efecto de la educación es significativa al 0.01. O más formal, la hipótesis de que todos los coeficientes asociados con la educación son simultáneamente iguales a cero puede ser rechazada al 0.01.

4.1.2 TEST DE WALD

Aunque el LR test es generalmente considerado superior, si el modelo es complejo, la muestra es muy grande, es muy costoso usar este test. Alternativamente, K test de Wald puede ser calculado usando "test", sin ninguna estimación adicional, por ejemplo:

mlogit ocupacion raza ed exper, b(5) nolog

test raza

test ed

test exper

```
. test raza

 ( 1)  [sirviente]raza = 0
 ( 2)  [obrero]raza = 0
 ( 3)  [artesano]raza = 0
 ( 4)  [empleador]raza = 0
 ( 5)  [profesional]o.raza = 0
       Constraint 5 dropped

           chi2(  4) =    8.15
         Prob > chi2 =    0.0863

.
. test ed

 ( 1)  [sirviente]ed = 0
 ( 2)  [obrero]ed = 0
 ( 3)  [artesano]ed = 0
 ( 4)  [empleador]ed = 0
 ( 5)  [profesional]o.ed = 0
       Constraint 5 dropped

           chi2(  4) =   84.97
         Prob > chi2 =    0.0000

.
. test exper

 ( 1)  [sirviente]exper = 0
 ( 2)  [obrero]exper = 0
 ( 3)  [artesano]exper = 0
 ( 4)  [empleador]exper = 0
 ( 5)  [profesional]o.exper = 0
       Constraint 5 dropped

           chi2(  4) =    7.99
         Prob > chi2 =    0.0918

.
```

Una forma de resumir lo anterior es:

mlogtest, wald

La lógica del test de Wald o LR puede ser extendida para testear el efecto de que dos o más variables independientes sean simultáneamente cero.

mlogit ocupacion raza ed exper, b(5) nolog

lrtest, saving(0)

mlogit ocupacion raza, b(5) nolog

lrtest

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


Likelihood-ratio test                                 LR chi2(8)  =    160.77
(Assumption: . nested in LRTEST_0)                    Prob > chi2 =    0.0000

.
```

O también:

mlogit ocupacion raza ed exper, b(5) nolog

mlogtest, lr set(ed exper)

```
. test [obrero]

 ( 1)  [obrero]raza = 0
 ( 2)  [obrero]ed = 0
 ( 3)  [obrero]exper = 0

           chi2(  3) =   76.39
         Prob > chi2 =    0.0000
```

Si ninguna de las variables independientes afecta significativamente al odds de la categoría m vs la categoría n, nosotros decimos que m y n son indistinguibles con respecto a las variables en el modelo 1. Que las categorías m y n sean indistinguibles corresponde a probar la siguiente hipótesis:

$$H_o : \beta_{1,m/P} = \beta_{2,m/P} = ... = \beta_{k-1,m/P} = \beta_{k,m/P} = 0 \qquad (19)$$

La cual será testeada con Wald o LR. Ambos test proveen resultados muy similares. Si las dos categorías son indistinguibles con respecto a las variables en el modelo, entonces podríamos obtener estimadores más eficientes, asociándolas. Para testear esto usamos “mlogtest”.

4.1.3 TEST DE WALD PARA CATEGORÍAS COMBINADAS

El comando usa el test de Wald, para la hipótesis nula de que dos categorías pueden ser no significativas, para todas las combinaciones de categorías que existan. Por ejemplo:

```
mlogit ocupacion raza  ed exper, b(5) nolog
test raza
test ed
test exper
```

```
. mlogit ocupacion raza  ed exper, b(5) nolog

Multinomial logistic regression                   Number of obs   =        337
                                                  LR chi2(12)     =     166.09
                                                  Prob > chi2     =     0.0000
Log likelihood = -426.80048                       Pseudo R2       =     0.1629

------------------------------------------------------------------------------
   ocupacion |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sirviente    |
        raza |  -1.774306   .7550543    -2.35   0.019    -3.254186   -.2944273
          ed |  -.7788519   .1146293    -6.79   0.000    -1.003521   -.5541826
       exper |  -.0356509    .018037    -1.98   0.048    -.0710028    -.000299
       _cons |   11.51833   1.849356     6.23   0.000     7.893659      15.143
-------------+----------------------------------------------------------------
obrero       |
        raza |  -.5378027   .7996033    -0.67   0.501    -2.104996    1.029391
          ed |  -.8782767   .1005446    -8.74   0.000     -1.07534   -.6812128
       exper |  -.0309296   .0144086    -2.15   0.032     -.05917   -.0026893
       _cons |   12.25956   1.668144     7.35   0.000     8.990061    15.52907
-------------+----------------------------------------------------------------
artesano     |
        raza |  -1.301963    .647416    -2.01   0.044    -2.570875   -.0330509
          ed |  -.6850365   .0892996    -7.67   0.000    -.8600605   -.5100126
       exper |  -.0079671   .0127055    -0.63   0.531    -.0328693    .0169351
       _cons |   10.42698   1.517943     6.87   0.000     7.451864    13.40209
-------------+----------------------------------------------------------------
empleador    |
        raza |  -.2029212   .8693072    -0.23   0.815    -1.906732     1.50089
          ed |  -.4256943   .0922192    -4.62   0.000    -.6064407   -.2449479
       exper |   -.001055   .0143582    -0.07   0.941    -.0291967    .0270866
       _cons |   5.279722   1.684006     3.14   0.002     1.979132    8.580313
-------------+----------------------------------------------------------------
profesional  |  (base outcome)
------------------------------------------------------------------------------

.
. test raza

 ( 1)  [sirviente]raza = 0
 ( 2)  [obrero]raza = 0
 ( 3)  [artesano]raza = 0
 ( 4)  [empleador]raza = 0
 ( 5)  [profesional]o.raza = 0
       Constraint 5 dropped

           chi2(  4) =    8.15
         Prob > chi2 =    0.0863

.
. test ed

 ( 1)  [sirviente]ed = 0
 ( 2)  [obrero]ed = 0
 ( 3)  [artesano]ed = 0
 ( 4)  [empleador]ed = 0
 ( 5)  [profesional]o.ed = 0
       Constraint 5 dropped

           chi2(  4) =   84.97
         Prob > chi2 =    0.0000

.
. test exper

 ( 1)  [sirviente]exper = 0
 ( 2)  [obrero]exper = 0
 ( 3)  [artesano]exper = 0
 ( 4)  [empleador]exper = 0
 ( 5)  [profesional]o.exper = 0
       Constraint 5 dropped

           chi2(  4) =    7.99
         Prob > chi2 =    0.0918

.
```

Una forma de resumir lo anterior es:

mlogtest, wald

```
. mlogtest, wald

**** Wald tests for independent variables (N=337)

 Ho: All coefficients associated with given variable(s) are 0.

             |      chi2   df   P>chi2
-------------+-------------------------
        raza |     8.149    4    0.086
          ed |    84.968    4    0.000
       exper |     7.995    4    0.092
---------------------------------------
```

La lógica del test de Wald o LR puede ser extensa para testear el efecto de que dos o más variables independientes sean simultáneamente cero.

mlogit ocupacion raza ed exper, b(5) nolog

lrtest, saving(0)

mlogit ocupacion raza, b(5) nolog

lrtest

O también:

mlogit ocupacion raza ed exper, b(5) nolog

mlogtest, lr set(ed exper)

Si ninguna de las variables independientes afecta significativamente al odds de la categoría m vs la categoría n, nosotros decimos que m y n son indistinguibles con respecto a las variables en el modelo 1. Que las categorías m y n sean indistinguibles corresponde a probar la siguiente hipótesis:

La cual será testeada con Wald o LR. Ambos test proveen resultados muy similares. Si las dos categorías son indistinguibles con respecto a las variables en el modelo, entonces podríamos obtener estimadores más eficientes, asociándolas. Para testear esto usamos "mlogtest".

4.1.4 TEST LR PARA CATEGORÍAS COMBINADAS

El test LR que combina m y n se calcula estimando el modelo completo sin restricciones y obteniendo el estadístico LR^2_{sr} y luego uno restricto en la cual la categoría m es usada como la base y todos los otros coeficientes excepto la constante en la ecuación de categoría n son ceros, obteniéndose el estadístico LR^2_r. El estadístico final es la diferencia $LR^2_{sr,r} = LR^2_{sr} - LR^2_r$, el cual se distribuye como una chi-cuadrada con K grados de libertad.

```
mlogit ocupacion raza  ed exper, b(5) nolog
mlogtest, lrcomb
```

Este comando puede usar restricciones, para ver esto, nosotros usamos el test que compara la categoría sirviente con obrero. Primero calculamos el modelo completo y guardamos los resultados:

```
mlogit ocupacion  raza ed exper, b(5) nolog
mlogtest, combine
```

Podemos rechazar la hipótesis de que la categoría para sirviente y profesional son indistinguibles, pero no podemos rechazar que sirviente y obrero sean indistinguibles.

Para testear que la categoría de sirviente sea indistinguible de la categoría base profesional digitamos

mlogtest, combine

```
. mlogtest, combine

**** Wald tests for combining alternatives (N=337)

 Ho: All coefficients except intercepts associated with a given pair
     of alternatives are 0 (i.e., alternatives can be combined).

Alternatives tested|      chi2    df   P>chi2
-------------------+-------------------------
sirvient-   obrero |     3.994     3    0.262
sirvient-artesano  |     3.203     3    0.361
sirvient-empleado  |    11.951     3    0.008
sirvient-profesio  |    48.190     3    0.000
  obrero-artesano  |     8.441     3    0.038
  obrero-empleado  |    20.055     3    0.000
  obrero-profesio  |    76.393     3    0.000
artesano-empleado  |     8.892     3    0.031
artesano-profesio  |    60.583     3    0.000
empleado-profesio  |    22.203     3    0.000
---------------------------------------------
```

Aquí nos sale todas las combinaciones posibles de pruebas de significancia estadística.

mlogit ocupacion raza ed exper, nolog

lrtest, saving(lrf)

Luego construimos la restricción

constraint define 999 [sirviente]

mlogit ocupacion raza ed exper, base(2) constraint(999) nolog

```
. mlogit ocupacion raza ed exper, base(2) constraint(999) nolog

Multinomial logistic regression                   Number of obs   =        337
                                                  Wald chi2(9)    =      88.31
Log likelihood = -428.84791                       Prob > chi2     =     0.0000

 ( 1)  [sirviente]o.raza = 0
 ( 2)  [sirviente]o.ed = 0
 ( 3)  [sirviente]o.exper = 0

------------------------------------------------------------------------------
   ocupacion |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sirviente    |
        raza |          0  (omitted)
          ed |          0  (omitted)
       exper |          0  (omitted)
       _cons |  -.8001193   .2162194    -3.70   0.000    -1.223901   -.3763371
-------------+----------------------------------------------------------------
obrero       |  (base outcome)
-------------+----------------------------------------------------------------
artesano     |
        raza |  -.2381783   .4978563    -0.48   0.632    -1.213959    .7376021
          ed |   .1599345   .0693853     2.31   0.021     .0239418    .2959273
       exper |   .0242824   .0113959     2.13   0.033     .0019469    .0466179
       _cons |  -1.969087   1.054935    -1.87   0.062    -4.036721     .098547
-------------+----------------------------------------------------------------
empleador    |
        raza |   .8829927    .843371     1.05   0.295    -.7699841    2.535969
          ed |   .4195709   .0958978     4.38   0.000     .2316147     .607527
       exper |   .0312007   .0143598     2.17   0.030     .0030561    .0593454
       _cons |  -7.140306   1.623401    -4.40   0.000    -10.32211   -3.958498
-------------+----------------------------------------------------------------
profesional  |
        raza |   1.097459   .6877939     1.60   0.111    -.2505923     2.44551
          ed |   .8445092    .093709     9.01   0.000     .6608429    1.028176
       exper |    .032303   .0133779     2.41   0.016     .0060827    .0585233
       _cons |  -12.42143   1.569897    -7.91   0.000    -15.49837   -9.344489
------------------------------------------------------------------------------

.
```

Donde observamos que restricción es impuesta y así calculamos el test de verosimilitud.

lrtest, using(lrf)

```
. lrtest, using(lrf)
You ran lrtest using the old syntax.  Click here to learn about the new syntax.


Likelihood-ratio test                                 LR chi2(3)  =      4.09
(Assumption: . nested in LRTEST_lrf)                  Prob > chi2 =    0.2514

.
```

V. TEST DE INDEPENDENCIA DE LAS ALTERNATIVAS IRRELEVANTES (IIA)

En el modelo de regresión logística multinomial, las razones de ocurrencia (ratios de odds) de cada par de valores de la variable dependiente no deben ser afectadas por el resto de alternativas posibles (Añadir o eliminar alternativas, no debe modificar el cociente). Se trata del supuesto de independencia de alternativas irrelevantes, que requiere que cada una de las posibles alternativas sea valorada de manera independiente por el que toma la decisión. Si no se cumple este supuesto el modelo multinomial no tiene lugar pues da lugar a coeficientes incorrectos. (Orihuela, 2011)

Tanto el MNLM y el condicional tienen como supuesto la independencia de alternativas irrelevantes, mostramos este supuesto en términos del modelo logit multinomial[7] (Orihuela, 2011).

[7] Es habitual en investigación empírica la utilización del modelo probit multinomial cuando se vulnera el supuesto de alternativas irrelevantes (En Stata el modelo probit multinomial se calcula con el comando mprobit pero cuando se incumple la condición mencionada se puede usar el comando asmprobit. El modelo probit asumen que los errores son normales por lo tanto los errores de las distintas alternativas están correlacionados y por lo tanto se supone que el modelo probit no se ve afectado por la vulneración del supuesto. Sin embargo autores como Long y Fresse (2006) argumenta que el modelo multinomial probit adolece de los mismos problemas que presenta el modelo multinomial logit.

$$\frac{\Pr(y=m/x)}{\Pr(y=n/x)}=\exp(x(\beta_{m/b}-\beta_{n/b})) \qquad (20)$$

Donde el Odds no depende de otras categorías que sean viables. En este sentido, estas categorías alternativas son irrelevantes, lo que significa que al añadir o borrar una categoría esta no afectara la cantidad de Odds en las categorías principales. Este punto es explicado a menudo con un ejemplo de transporte en autobuses rojos/azules: Supongamos que se tiene que elegir entre un autobús rojo y un carro para ir a trabajar y que el Odds de tomar el autobús comparado con el carro es de 1/1. La IIA implica que el oods deberá mantenerse 1/1 entre estas dos alternativas, aún si una nuevo compañía de autobuses azules llega al pueblo, autobuses de características idénticas a la compañía de autobuses roja. Así, las probabilidades de manejar un carro pueden ser tan pequeñas aún al añadir diferentes colores de autobuses. Más razonable, seria esperar que el Odds de comparar un autobús rojo y un carro debería reducirse a 1/2 ya que la mitad de personas que subían al autobús rojo, ahora se esperaría que suban al azul.

Hay dos test que tratan el supuesto IIA. Hausman y McFadden (1984) propusieron un test tipo Hausman. Y McFadden, Tye y Train (1976) propusieron una aproximación al test de ratio de verosimilitud, que fue implantado por Small y Hsiao (1985). Ambos, asumían que el MNLM es estimado con la categoría base "b", y existían por tanto J-1 test a ser calculados excluyendo cada uno las principales categorías para formar un modelo restricto. Para cambiar la categoría base, el test puede ser calculado excluyendo b. El resultado del test difieren dependiendo de cuál es la categoría base fue usada para estimar el modelo (Orihuela, 2011).

5.1 TEST DE HAUSMAN-MCFADDEN (1984).

El test de Hausman implica los siguientes pasos:

Estimar el modelo completo con todas las J categorías incluidas y obtener el estimador $\hat{\beta}_{sr}$

Estimar un modelo restricto eliminando una a una las diferentes categorías y obtener el estimador $\hat{\beta}_{r}$.

Dejar que $\hat{\beta}^{*}_{sr}$ sea una sub muestra de $\hat{\beta}_{sr}$ luego de eliminar los coeficientes no estimados en el modelo restricto.

Si se estima el modelo con el conjunto de elección completo (f) y restringido (s). Bajo la hipótesis nula de que se cumple IIA, los coeficientes en ambos casos deberían ser similares. Se contrasta la significatividad de dicha diferencia.

$$H = (\hat{\beta}_r - \hat{\beta}^{*}_{sr})' \left[\hat{V}_r - \hat{V}^{*}_{sr} \right]^{-1} (\hat{\beta}_s - \hat{\beta}^{*}_{sr}) \sim \chi^2$$

Donde H es distribuido asintóticamente como una chi-cuadrado con grados de libertad iguales a las filas de $\hat{\beta}_r$ si IIA es verdadero.

Donde $\hat{V}_r$ y $\hat{V}^{*}_{sr}$ son la matriz de varianzas y covarianzas estimada en cada caso.

Si rechazamos la hipótesis nula de validez de IIA, debemos emplear un modelo alternativo al MNL.

Los valores signicativos de H indican que el supuesto de IIA ha sido violado.

```
mlogit ocupacion raza ed exper, b(5) nolog
mlogtest, hausman base
```

```
. mlogtest, hausman base

**** Hausman tests of IIA assumption (N=337)

 Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

 Omitted |      chi2   df   P>chi2   evidence
---------+------------------------------------
sirvient |     7.324   12    0.835   for Ho
  obrero |     0.320   12    1.000   for Ho
artesano |   -14.436   12     ---    ---
empleado |    -5.541   11     ---    ---
profesio |    -0.119   12     ---    ---
----------------------------------------------
 Note: If chi2<0, the estimated model does not
 meet asymptotic assumptions of the test.

.
```

Cinco test fueron reportados, los primeros cuatro corresponden a la exclusión de una de las cuatro categorías no base. La quinta es calculada reestimando el modelo usando la más categoría más grande como categoría base. Ninguna rechaza la H0, de que la IIA sea verdadera. Los resultados difieren considerablemente al cambiar la categoría base. Tres de los test estadísticos son negativos, lo cual es común encontrar, un resultado así, presenta evidencias de que la IIA no ha sido violada. Un mayor sentimiento de la variabilidad de los resultados puede ser visto corriendo el mlogit con una categoría diferente y volviendo a calcular el test.

Bondad de ajuste

Para medir el ajuste, podemos usar el **fitstat** como medida de análisis.

```
fitstat

easures of Fit for mlogit of ocupacion

og-Lik Intercept Only:      -509.844   Log-Lik Full Model:          -426.800
(321):                       853.601   LR(12):                       166.087
                                       Prob > LR:                      0.000
cFadden's R2:                  0.163   McFadden's Adj R2:              0.131
_ (Cox-Snell) R2:              0.389   Cragg-Uhler(Nagelkerke) R2:     0.409
ount R2:                       0.501   Adj Count R2:                   0.253
IC:                            2.628   AIC*n:                        885.601
IC:                        -1014.646   BIC':                         -96.246
IC used by Stata:            946.722   AIC used by Stata:            885.601
```

VI. ANÁLISIS DE PROBABILIDADES Y CAMBIOS MARGINALES

Mientras el MNLM es una simple extensión matemática del modelo binario, la interpretación se dificulta por la gran cantidad de posibles comparaciones que se pueden hacer. Aún en, nuestro ejemplo con cinco categorías, nosotros tendríamos muchas comparaciones por hacer. Pero existen comandos que nos proporcionan de herramientas muy poderosas para llevar a cabo dicha tarea.

6.1 PREDICCIÓN DE LAS PROBABILIDADES CON "PREDICT"

Las probabilidades son obtenidas de la siguiente manera:

mlogit ocupación raza ed exper, b(5) nolog

predict probM probC probB probW probP

describe prob*

summarize prob*

```
. predict probM probC probB probW probP
(option pr assumed; predicted probabilities)

. describe prob*

              storage  display     value
variable name   type   format      label      variable label
--------------------------------------------------------------------------------
probM           float  %9.0g                  Pr(ocupacion==sirviente)
probC           float  %9.0g                  Pr(ocupacion==obrero)
probB           float  %9.0g                  Pr(ocupacion==artesano)
probW           float  %9.0g                  Pr(ocupacion==empleador)
probP           float  %9.0g                  Pr(ocupacion==profesional)

.


. summarize prob*

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       probM |        337    .0919881     .059396   .0010737   .3281906
       probC |        337    .2047478    .1450568   .0012066   .6974148
       probB |        337    .2492582    .1161309   .0079713    .551609
       probW |        337    .1216617    .0452844   .0083857   .2300058
       probP |        337    .3323442    .2870992   .0001935   .9597512
```

6.2PREDICCIÓN DE LAS PROBABILIDADES CON "PRVALUE"

Predecir las probabilidades para un individuo con características específicas puede realizarse con el comando "prvalue". Por ejemplo, podríamos querer calcular las probabilidades de cada categoría ocupacional comparando a los de raza negra con los de raza blanca, con educación y experiencia promedia.

mlogit ocupacion raza ed exper, b(5) nolog

quietly prvalue, x(raza 0) rest(mean) save

prvalue, x(raza 1) rest(mean) dif

```
.  prvalue, x( raza 0) rest(mean) save

mlogit: Predictions for ocupacion

Confidence intervals by delta method

                                95% Conf. Interval
  Pr(y=sirvient|x):    0.2168   [ 0.0426,    0.3910]
  Pr(y=obrero|x):      0.1363   [ 0.0004,    0.2722]
  Pr(y=artesano|x):    0.4387   [ 0.2343,    0.6431]
  Pr(y=empleado|x):    0.0877   [-0.0303,    0.2058]
  Pr(y=profesio|x):    0.1204   [-0.0019,    0.2426]

         raza          ed       exper
x=          0   13.094955   20.501484

.



. quietly prvalue, x( raza 0) rest(mean) save

.
. prvalue, x(raza 1) rest(mean) dif

mlogit: Change in Predictions for ocupacion

Confidence intervals by delta method

                     Current     Saved    Change   95% CI for Change
  Pr(y=sirvient|x):   0.0860    0.2168   -0.1309  [-0.3056,   0.0439]
  Pr(y=obrero|x):     0.1862    0.1363    0.0498  [-0.0897,   0.1893]
  Pr(y=artesano|x):   0.2790    0.4387   -0.1597  [-0.3686,   0.0491]
  Pr(y=empleado|x):   0.1674    0.0877    0.0797  [-0.0477,   0.2071]
  Pr(y=profesio|x):   0.2814    0.1204    0.1611  [ 0.0277,   0.2944]

                 raza          ed       exper
Current=            1   13.094955   20.501484
  Saved=            0   13.094955   20.501484
   Diff=            1           0           0

.
```

6.3 PREDICCIÓN DE PROBABILIDADES CON "PRTAB"

Si se desea predecir las probabilidades para todas las combinaciones posibles de un conjunto de variables categóricas independientes, se usa el comando "prtab". Por ejemplo si quisiéramos conocer como la respuesta de los blancos y negros difieren en sus probabilidades de tener un trabajo de sirvientes, conforme aumentan los años de educación. Entonces digitamos el siguiente comando:

prtab ed raza, novarlbl outcome(1)

```
. prtab ed raza, novarlbl outcome(1)

mlogit: Predicted probabilities of outcome

-----------------------------------------
          |            raza
       ed |   no blanco       blanco
----------+------------------------------
        3 |      0.2847       0.1216
        6 |      0.2987       0.1384
        7 |      0.2988       0.1417
        8 |      0.2963       0.1431
        9 |      0.2906       0.1417
       10 |      0.2814       0.1366
       11 |      0.2675       0.1265
       12 |      0.2476       0.1104
       13 |      0.2199       0.0883
       14 |      0.1832       0.0632
       15 |      0.1393       0.0401
       16 |      0.0944       0.0228
       17 |      0.0569       0.0120
       18 |      0.0310       0.0060
       19 |      0.0158       0.0029
       20 |      0.0077       0.0014
-----------------------------------------

          raza          ed       exper
x=   .91691395   13.094955   20.501484
```

La tabla muestra una sustancial diferencia entre blancos y negros en la probabilidad de tener trabajos de sirvientes, y como estas probabilidades son afectadas según aumentan los años de educación. Sin embargo, dado el número de categorías para "ed", el graficar las probabilidades predichas con "prgen" será la manera más útil de examinar estos resultados.

6.4 PREDICCIÓN DE PROBABILIDADES CON "PRGEN"

Las probabilidades predichas pueden ser graficadas usando los mismos métodos considerados para los modelos de **regresión ordinal**. Luego de estimar el modelo, nosotros usaremos "prgen" para calcular las probabilidades predichas para blancos con experiencia laboral promedia e incrementos en los años de educación de seis a veinte años.

```
variable name    type     format      label       variable label
--------------------------------------------------------------------
whtx             float    %9.0g                   Años de Educación
whtp1            float    %9.0g                   pr(sirvient)=Pr(1)
whtp2            float    %9.0g                   pr(obrero)=Pr(2)
whtp3            float    %9.0g                   pr(artesano)=Pr(3)
whtp4            float    %9.0g                   pr(empleado)=Pr(4)
whtp5            float    %9.0g                   pr(profesio)=Pr(5)
whts1            float    %9.0g                   pr(y<=1)
whts2            float    %9.0g                   pr(y<=2)
whts3            float    %9.0g                   pr(y<=3)
whts4            float    %9.0g                   pr(y<=4)
whts5            float    %9.0g                   pr(y<=5)

.
end of do-file

. do "C:\Users\Usuario\AppData\Local\Temp\STD04000000.tmp"

. prgen ed, x(raza=0) from(6) to (20) gen(nwht) ncases(15)

mlogit: Predicted values as ed varies from 6 to 20.

         raza          ed       exper
x=          0   13.094955   20.501484

.
end of do-file

.
```

Las variables nwhtp1 whtp1 contienen las probabilidades predichas de tener trabajos de sirvientes para los de raza negros y blancos. La gráfica de éstas, pueden suministrar información más clara que los resultados que proporciona el comando "prtab".

prgen ed, x(raza=1) from(6) to (20) gen(wht) ncases(15)

describe wht*

```
. prgen ed, x(raza=1) from(6) to (20) gen(wht) ncases(15)

mlogit: Predicted values as ed varies from 6 to 20.

          raza         ed      exper
x=           1  13.094955  20.501484

.
. describe wht*

              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------
whtx            float   %9.0g                 Años de educaciçon
whtp1           float   %9.0g                 pr(sirvient)=Pr(1)
whtp2           float   %9.0g                 pr(obrero)=Pr(2)
whtp3           float   %9.0g                 pr(artesano)=Pr(3)
whtp4           float   %9.0g                 pr(empleado)=Pr(4)
whtp5           float   %9.0g                 pr(profesio)=Pr(5)
whts1           float   %9.0g                 pr(y<=1)
whts2           float   %9.0g                 pr(y<=2)
whts3           float   %9.0g                 pr(y<=3)
whts4           float   %9.0g                 pr(y<=4)
whts5           float   %9.0g                 pr(y<=5)

.
```

prgen ed, x(raza=0) from(6) to (20) gen(nhwht) ncases(15)

describe nhwht*

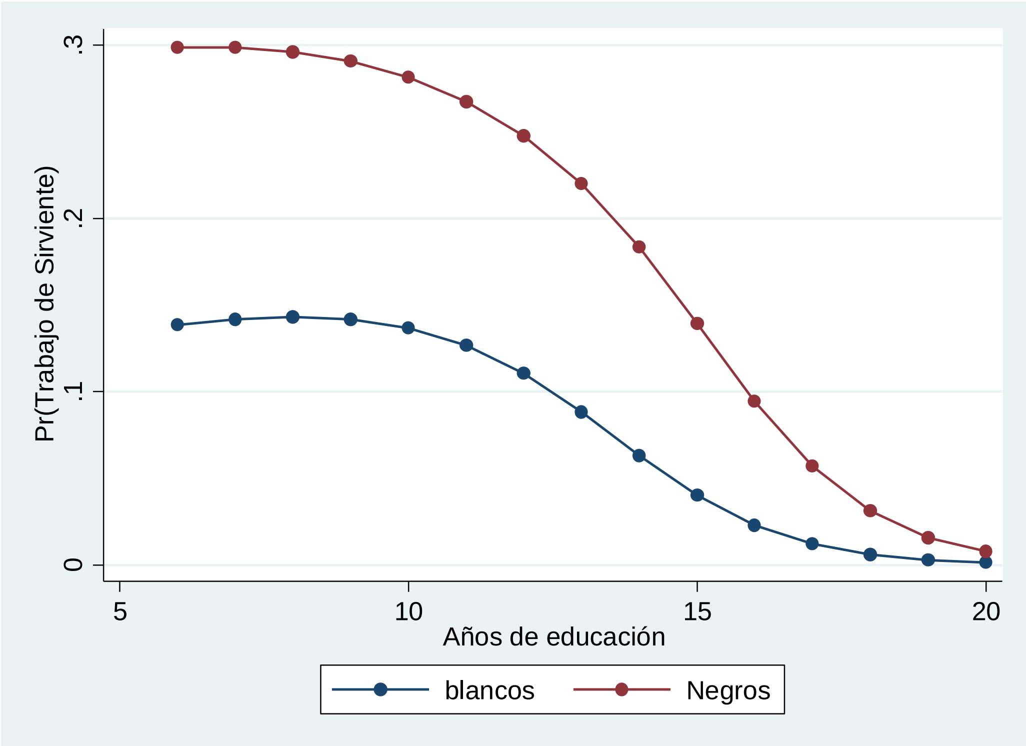
```
mlogit: Predicted values as ed varies from 6 to 20.

         raza         ed      exper
x=          0  13.094955  20.501484

.
.
.
. describe nhwht*

              storage   display    value
variable name   type    format     label      variable label
---------------------------------------------------------------------------
nhwhtx          float   %9.0g                 Años de educaçon
nhwhtp1         float   %9.0g                 pr(sirvient)=Pr(1)
nhwhtp2         float   %9.0g                 pr(obrero)=Pr(2)
nhwhtp3         float   %9.0g                 pr(artesano)=Pr(3)
nhwhtp4         float   %9.0g                 pr(empleado)=Pr(4)
nhwhtp5         float   %9.0g                 pr(profesio)=Pr(5)
nhwhts1         float   %9.0g                 pr(y<=1)
nhwhts2         float   %9.0g                 pr(y<=2)
nhwhts3         float   %9.0g                 pr(y<=3)
nhwhts4         float   %9.0g                 pr(y<=4)
nhwhts5         float   %9.0g                 pr(y<=5)
```

```
label var whtp1 "blancos"
label var nhwhtp1 "Negros"
set textsize 125
tw sc whtp1 nwhtp1 nwhtx, connect(ss ss) ///
xtitle(Años de educación) ///
ytitle(Pr(Trabajo de Sirviente))
```



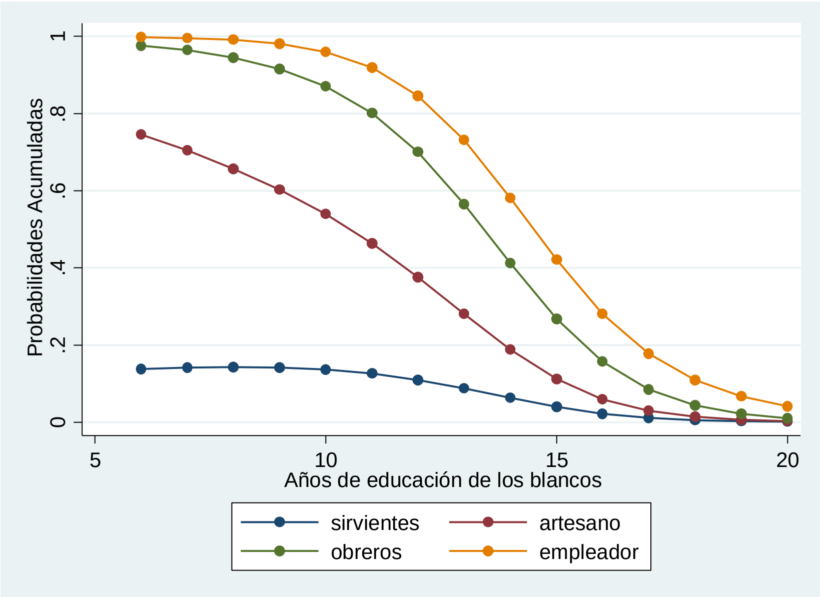
Aunque las categorías nominales no están ordenadas, el grafico que suma las probabilidades puede ser una útil manera de revelar las probabilidades predichas para todas las categorías. Para esto construimos una gráfica que muestre como la educación afecta la probabilidad de cada ocupación para blancos. A continuación realizamos los siguientes comandos en el dofile[8].

```
label var whts1 "sirvientes"
label var whts2 "artesano"
label var whts3 "obreros"
label var whts4 "empleador"
label var whts5 "profesional"
set textsize 125
tw sc whts1 whts2 whts3 whts4 whtx, c(ss ss ss ss) ///
```

[8] Cuando hay toda una rutina de comandos, es preferible escribirlos en un Dofile, el cual es un documento que sirve para realizar programas en Stata.

xtitle(Años de educación de los blancos) ///

ytitle(Probabilidades Acumuladas)



El gráfico muestra las cuatro probabilidades acumuladas de las cuatro categorías. Conforme aumenta los años de educación la línea más baja etiquetada con sirvientes grafica las probabilidades de tener un trabajo de sirviente según varía los años de educación. Es la misma información que se presentó en el gráfico anterior para las personas de raza blanca. La siguiente línea, etiquetada como artesano grafica la suma de probabilidades de tener trabajo un trabajo de **sirviente o artesano**. De esta manera el área entre la línea roja y azul, es la probabilidad de tener trabajo de artesano.

La desventaja de estos comandos, es que puede tomar un largo tiempo su estimación, luego de la estimación multinomial, si el número de observaciones y las variables independientes, son muchas (Orihuela, 2011).

6.6 CAMBIO EN LAS PROBABILIDADES PREDICHAS

Cambios marginales y discretos pueden ser usados de la misma manera que en modelos de salidas ordinales. Como antes, ambas pueden ser calculadas con "prchange".

Cambio Marginal: Asimismo podemos definir el cambio marginal como:

$$\frac{\partial \Pr(y=m/x)}{\partial x_k} = \Pr(y=m/x)\left[\beta_{k,m/J} - \sum_{j=1}^{J}\beta_{k,m/J} * \Pr(y=j/x)\right]$$

Dado que esta ecuación combina todos los $\beta_{k,j/J}$, el valor de los cambios marginales dependen de los valores de todas las variables del modelo. Más aun cuando el valor de x_k cambia, el signo del impacto marginal puede cambiar. Por ejemplo, en algún punto, el efecto marginal de la educación sobre tener una ocupación de sirvientes podría ser positivo, mientras que en otro punto dicho efecto podría ser negativo (Colin Cameron & Trivedi, 2005).

Cambio Discreto Podemos definir el cambio discreto como:

$$\frac{\Delta \Pr(y=m/x)}{\Delta x_k} = \Pr(y=m/x, x_k = x_E) - \Pr(y=m/x, x_k = x_s)$$

Donde la magnitud del cambio depende de los niveles de todas las variables y del tamaño del cambio que es realizado.

Los J cambios discretos de los coeficientes para una variable (uno por cada categoría) pueden ser resumidos calculando un promedio del valor absoluto de los cambios a través de todas las categorías.

$$\bar{\Delta} = \frac{1}{J}\sum_{j=1}^{J}\left|\frac{\Delta \Pr(y = m / x)}{\Delta x_k}\right|$$

Donde el valor absoluto es tomado porque la suma de los cambios sin tomar el valor absoluto es necesariamente cero.

7. LOS CAMBIOS MARGINALES SON LISTADOS EN LAS FILAS DEL EFECTO MARGINAL

Para variables que no son binarias, los cambios discretos son reportados sobre el rango completo de las variables (reportado como Min ➔ Max), para cambios en una unidad centrada alrededor de los valores base (reportado como -+1/2) y para cambios en una desviación estándar centrada alrededor de los valores base (reportado como -+sd/2). Si la opción “uncentered” es usada, los cambios comienza en la opción especificada por x() y rest() y se incrementan en una unidad o una desviación estándar desde allí. Para variables binarias, el cambio discreto de 0 a 1 es la cuantía apropiada y es la única cuantía presentada.

Vemos en el resultado para raza, que para aquellos que tienen educación y experiencia promedio, la probabilidad predicha de tener un trabajo profesional es de 0.16 veces mayor para blancos que para negros. El cambio promedio es listado en la primera columna. Por ejemplo, para el hecho de ser blanco, el cambio absoluto promedio en la probabilidad de varias categorías laborales de ser blanco como oposición a ser negro es de 0.12 (Colin Cameron & Trivedi, 2005).

```
mlogit: Changes in Probabilities for ocupacion

raza
        Avg|Chg|     sirvient       obrero     artesano     empleado     profesio
0->1   .11623582  -.13085523   .04981799  -.15973434   .07971004     .1610615

ed
            Avg|Chg|     sirvient       obrero     artesano     empleado     profesio
Min->Max   .39242268  -.13017954  -.70077323  -.15010394   .02425591   .95680079
   -+1/2   .05855425  -.02559762  -.06831616  -.05247185   .01250795   .13387768
  -+sd/2    .1640657  -.07129153  -.19310513  -.14576758   .03064777   .37951647
MargEfct   .05894859  -.02579097  -.06870635  -.05287415   .01282041   .13455107

exper
            Avg|Chg|     sirvient       obrero     artesano     empleado     profesio
Min->Max   .12193559  -.11536534  -.18947365   .03115708   .09478889   .17889298
   -+1/2   .00233425  -.00226997  -.00356567   .00105992    .0016944   .00308132
  -+sd/2   .03253578  -.03167491  -.04966453   .01479983   .02360725   .04293236
MargEfct   .00233427  -.00226997  -.00356571   .00105992   .00169442   .00308134

          sirvient      obrero    artesano    empleado    profesio
Pr(y|x)  .09426806   .18419114   .29411051   .16112968   .26630062

          raza        ed     exper
   x=  .916914    13.095   20.5015
sd_x=  .276423   2.94643   13.9594
```

El cambio marginal también puede ser calculado con "mfx", que al igual que **"prchange"**, calcula el cambio manteniendo todo el conjunto de variables independientes en su media. Hay que notar que no solo nos permite calcular el efecto de un conjunto de variables en el modelo, sino que también estima los efectos marginales para una categoría a la vez (Bravo & Vásquez, 2008):

mfx compute, predict(outcome(1))

```
. mfx compute, predict(outcome(1))

Marginal effects after mlogit
      y  = Pr(ocupacion==sirviente) (predict, outcome(1))
         =  .09426806
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.	]	X
raza*	-.1308552	.08914	-1.47	0.142	-.305562	.043852	.916914
ed	-.025791	.00688	-3.75	0.000	-.039269	-.012313	13.095
exper	-.00227	.00126	-1.80	0.071	-.004737	.000197	20.5015

(*) dy/dx is for discrete change of dummy variable from 0 to 1

.

Estos resultados son para la categoría "sirvientes", obteniéndose similares resultados que "prchange" en relación a las variables continuas y discretas. La ventaja, una vez más, es que podemos obtener los valores de las desviaciones estándar, la desventaja es que nos puede tomar un largo tiempo su estimación, luego de la estimación multinomial, si el número de observaciones y las variables independientes, son numerosas.

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