**Trabajo Práctico N° 4:**

**Modelos para Variables Categóricas Ordenadas.**

**Ejercicio 1: Predicción de Calificaciones de Clientes.**

*Considerar el ejercicio del Problem Set anterior con el mismo título que éste. Repetir el análisis utilizando un modelo ordenado.*

**(a)** *Abrir y describir la base.*

Variable | Obs Mean Std. dev. Min Max

-------------+---------------------------------------------------------

nps | 42,019 8.369975 2.263878 1 10

marital\_st~e | 0

gender\_code | 0

edad | 42,020 52.16497 12.56996 19 101

branch\_desc | 0

-------------+---------------------------------------------------------

segmento | 0

operaciones | 42,020 1.728439 1.476585 1 31

mes | 42,020 6.736292 3.241668 1 12

nps\_anterior | 0

hora | 42,020 11.7812 1.743031 7 18

-------------+---------------------------------------------------------

dia | 42,020 14.91792 8.634796 1 31

dia\_semana | 0

espera | 42,020 10.89938 10.70589 0 60

cliente | 42,020 21372.36 12335.51 1 42760

**(b)** *Generar una variable que clasifique a los clientes en función de si son promotores, detractores o neutrales.*

clasificaci |

on | Freq. Percent Cum.

------------+-----------------------------------

Detractor | 6,265 14.91 14.91

Neutral | 9,579 22.80 37.71

Promotor | 26,175 62.29 100.00

------------+-----------------------------------

Total | 42,019 100.00

**(c)** *Analizar cómo cambia la variable de espera en función de la clasificación de los clientes.*



**(d)** *Tomar una muestra del 10% de los datos. Estimar un logit multinomial ordenado para predecir cómo cambian las clasificaciones en función de la espera, condicionando en explicativas que se considere relevantes.*

Logit multinomial ordenado (betas):

Ordered logistic regression Number of obs = 4,202

LR chi2(7) = 394.03

Prob > chi2 = 0.0000

Log likelihood = -3659.6981 Pseudo R2 = 0.0511

-------------------------------------------------------------------------------

clasificacion | Coefficient Std. err. z P>|z| [95% conf. interval]

--------------+----------------------------------------------------------------

\_Igender\_co\_2 | -.0667762 .0687925 -0.97 0.332 -.201607 .0680546

edad | .0163998 .0026606 6.16 0.000 .0111852 .0216144

\_Isegmento\_2 | .7313334 .8363563 0.87 0.382 -.9078948 2.370562

\_Isegmento\_3 | .2147268 .1144579 1.88 0.061 -.0096065 .4390601

\_Isegmento\_4 | -.478007 .1271562 -3.76 0.000 -.7272286 -.2287855

\_Isegmento\_5 | -.3697909 .1299945 -2.84 0.004 -.6245754 -.1150063

espera | -.0359647 .0030773 -11.69 0.000 -.041996 -.0299333

--------------+----------------------------------------------------------------

/cut1 | -1.429497 .178061 -1.77849 -1.080504

/cut2 | -.1286401 .1758049 -.4732113 .2159311

-------------------------------------------------------------------------------

Logit multinomial ordenado (odds ratios):

Ordered logistic regression Number of obs = 4,202

LR chi2(7) = 394.03

Prob > chi2 = 0.0000

Log likelihood = -3659.6981 Pseudo R2 = 0.0511

-------------------------------------------------------------------------------

clasificacion | Odds ratio Std. err. z P>|z| [95% conf. interval]

--------------+----------------------------------------------------------------

\_Igender\_co\_2 | .9354045 .0643488 -0.97 0.332 .8174161 1.070424

edad | 1.016535 .0027045 6.16 0.000 1.011248 1.02185

\_Isegmento\_2 | 2.077849 1.737822 0.87 0.382 .4033725 10.7034

\_Isegmento\_3 | 1.239523 .1418732 1.88 0.061 .9904395 1.551249

\_Isegmento\_4 | .6200178 .0788391 -3.76 0.000 .4832464 .7954992

\_Isegmento\_5 | .6908788 .0898104 -2.84 0.004 .5354887 .8913605

espera | .9646744 .0029686 -11.69 0.000 .9588736 .9705103

--------------+----------------------------------------------------------------

/cut1 | -1.429497 .178061 -1.77849 -1.080504

/cut2 | -.1286401 .1758049 -.4732113 .2159311

-------------------------------------------------------------------------------

Note: Estimates are transformed only in the first equation to odds ratios.

**(e)** *Calcular los efectos marginales.*

Efectos marginales en Logit multinomial ordenado (clasificación 1):

Marginal effects after ologit

y = Pr(clasificacion==1) (predict, pr outcome(1))

= .13947074

------------------------------------------------------------------------------

variable | dy/dx Std. err. z P>|z| [ 95% C.I. ] X

---------+--------------------------------------------------------------------

\_Igend~2\*| .0079445 .00812 0.98 0.328 -.007962 .023851 .682532

edad | -.0019683 .00032 -6.13 0.000 -.002597 -.001339 51.9412

\_Isegm~2\*| -.0671776 .05629 -1.19 0.233 -.177512 .043157 .001666

\_Isegm~3\*| -.0261068 .0141 -1.85 0.064 -.053748 .001534 .580438

\_Isegm~4\*| .063994 .01889 3.39 0.001 .02698 .101008 .179914

\_Isegm~5\*| .0486811 .01869 2.60 0.009 .012049 .085313 .149929

espera | .0043164 .00038 11.41 0.000 .003575 .005058 11.1349

------------------------------------------------------------------------------

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

Efectos marginales en Logit multinomial ordenado (clasificación 2):

Marginal effects after ologit

y = Pr(clasificacion==2) (predict, pr outcome(2))

= .23365369

------------------------------------------------------------------------------

variable | dy/dx Std. err. z P>|z| [ 95% C.I. ] X

---------+--------------------------------------------------------------------

\_Igend~2\*| .0076247 .00788 0.97 0.333 -.007815 .023064 .682532

edad | -.0018677 .00031 -6.00 0.000 -.002478 -.001258 51.9412

\_Isegm~2\*| -.0833501 .08871 -0.94 0.347 -.25722 .09052 .001666

\_Isegm~3\*| -.0242958 .01287 -1.89 0.059 -.049518 .000926 .580438

\_Isegm~4\*| .0511031 .01252 4.08 0.000 .026558 .075648 .179914

\_Isegm~5\*| .0401242 .01326 3.03 0.002 .014131 .066118 .149929

espera | .0040958 .00038 10.66 0.000 .003342 .004849 11.1349

------------------------------------------------------------------------------

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

Efectos marginales en Logit multinomial ordenado (clasificación 3):

Marginal effects after ologit

y = Pr(clasificacion==3) (predict, pr outcome(3))

= .62687557

------------------------------------------------------------------------------

variable | dy/dx Std. err. z P>|z| [ 95% C.I. ] X

---------+--------------------------------------------------------------------

\_Igend~2\*| -.0155692 .01599 -0.97 0.330 -.046903 .015764 .682532

edad | .003836 .00062 6.17 0.000 .002618 .005054 51.9412

\_Isegm~2\*| .1505277 .14494 1.04 0.299 -.133553 .434608 .001666

\_Isegm~3\*| .0504026 .02693 1.87 0.061 -.002378 .103183 .580438

\_Isegm~4\*| -.1150971 .0312 -3.69 0.000 -.176246 -.053948 .179914

\_Isegm~5\*| -.0888053 .03183 -2.79 0.005 -.151197 -.026414 .149929

espera | -.0084122 .00072 -11.67 0.000 -.009825 -.006999 11.1349

------------------------------------------------------------------------------

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

**(f)** *Repetir el análisis con un Probit multinomial ordenado y comparar.*

Probit multinomial ordenado:

Ordered probit regression Number of obs = 4,202

LR chi2(7) = 450.86

Prob > chi2 = 0.0000

Log likelihood = -3631.2859 Pseudo R2 = 0.0585

-------------------------------------------------------------------------------

clasificacion | Coefficient Std. err. z P>|z| [95% conf. interval]

--------------+----------------------------------------------------------------

\_Igender\_co\_2 | .0079382 .040891 0.19 0.846 -.0722067 .0880831

edad | .0088488 .0015378 5.75 0.000 .0058348 .0118628

\_Isegmento\_2 | .3043314 .5146773 0.59 0.554 -.7044176 1.31308

\_Isegmento\_3 | .1132939 .0664574 1.70 0.088 -.0169602 .2435479

\_Isegmento\_4 | -.3667121 .0740819 -4.95 0.000 -.5119099 -.2215144

\_Isegmento\_5 | -.3928859 .0764502 -5.14 0.000 -.5427255 -.2430463

espera | -.0202896 .0018409 -11.02 0.000 -.0238977 -.0166814

--------------+----------------------------------------------------------------

/cut1 | -.9159207 .1025571 -1.116929 -.7149126

/cut2 | -.1473535 .1017522 -.3467842 .0520771

-------------------------------------------------------------------------------

Efectos marginales en Progit multinomial ordenado (clasificación 1):

Marginal effects after oprobit

y = Pr(clasificacion==1) (predict, pr outcome(1))

= .13680723

------------------------------------------------------------------------------

variable | dy/dx Std. err. z P>|z| [ 95% C.I. ] X

---------+--------------------------------------------------------------------

\_Igend~2\*| -.001742 .00899 -0.19 0.846 -.019358 .015873 .680866

edad | -.0019388 .00034 -5.73 0.000 -.002602 -.001276 52.1171

\_Isegm~2\*| -.055947 .07726 -0.72 0.469 -.207383 .095489 .001428

\_Isegm~3\*| -.0250324 .01481 -1.69 0.091 -.054056 .003991 .567111

\_Isegm~4\*| .0905003 .02035 4.45 0.000 .050614 .130386 .186578

\_Isegm~5\*| .099153 .02183 4.54 0.000 .05637 .141936 .147787

espera | .0044455 .00041 10.79 0.000 .003638 .005253 11.1349

------------------------------------------------------------------------------

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

Efectos marginales en Probit multinomial ordenado (clasificación 2):

Marginal effects after oprobit

y = Pr(clasificacion==2) (predict, pr outcome(2))

= .23532568

------------------------------------------------------------------------------

variable | dy/dx Std. err. z P>|z| [ 95% C.I. ] X

---------+--------------------------------------------------------------------

\_Igend~2\*| -.0012622 .00649 -0.19 0.846 -.013992 .011467 .680866

edad | -.0014084 .00025 -5.61 0.000 -.001901 -.000916 52.1171

\_Isegm~2\*| -.0520378 .09116 -0.57 0.568 -.230701 .126626 .001428

\_Isegm~3\*| -.0179079 .01044 -1.71 0.086 -.038376 .00256 .567111

\_Isegm~4\*| .0518055 .0091 5.69 0.000 .033965 .069646 .186578

\_Isegm~5\*| .053895 .00872 6.18 0.000 .036803 .070987 .147787

espera | .0032294 .00032 10.13 0.000 .002605 .003854 11.1349

------------------------------------------------------------------------------

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

Efectos marginales en Probit multinomial ordenado (clasificación 3):

Marginal effects after oprobit

y = Pr(clasificacion==3) (predict, pr outcome(3))

= .6278671

------------------------------------------------------------------------------

variable | dy/dx Std. err. z P>|z| [ 95% C.I. ] X

---------+--------------------------------------------------------------------

\_Igend~2\*| .0030042 .01548 0.19 0.846 -.02734 .033349 .680866

edad | .0033472 .00058 5.75 0.000 .002207 .004487 52.1171

\_Isegm~2\*| .1079848 .16839 0.64 0.521 -.22206 .43803 .001428

\_Isegm~3\*| .0429403 .02522 1.70 0.089 -.006492 .092373 .567111

\_Isegm~4\*| -.1423059 .02913 -4.88 0.000 -.199406 -.085205 .186578

\_Isegm~5\*| -.1530479 .03018 -5.07 0.000 -.212194 -.093902 .147787

espera | -.007675 .0007 -11.01 0.000 -.009042 -.006308 11.1349

------------------------------------------------------------------------------

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

**(g)** *Realizar un test de la significatividad de las variables.*

Stata.

**Ejercicio 2:** **Modelo Secuencial.**

**(a)** *Considerar la base de datos “nlsw88.dta”. En la misma, hay datos de un grupo de mujeres de entre 30 y 40 años para estudiar los patrones de la fuerza laboral. Estimar un logit secuencial con la decisión de educación utilizando el comando seqlogit y mostrar que se pueden obtener los mismos resultados estimando varios modelos logit por separado.*

Logit secuencial:

Number of obs = 2,244

LR chi2(9) = 108.50

Log likelihood = -2882.1386 Prob > chi2 = 0.0000

------------------------------------------------------------------------------

educ\_cat | Coefficient Std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

\_2\_3\_4v1 |

race |

Black | -.9151569 .1282466 -7.14 0.000 -1.166516 -.6637983

Other | -.4910998 .5511525 -0.89 0.373 -1.571339 .5891394

|

south |

South | -.4175069 .1259601 -3.31 0.001 -.6643841 -.1706298

\_cons | 2.250353 .0952967 23.61 0.000 2.063574 2.437131

-------------+----------------------------------------------------------------

\_3\_4v2 |

race |

Black | -.173837 .1131414 -1.54 0.124 -.3955902 .0479161

Other | 1.745005 .6241267 2.80 0.005 .5217389 2.968271

|

south |

South | -.1495226 .0968386 -1.54 0.123 -.3393228 .0402777

\_cons | .1079773 .0617595 1.75 0.080 -.0130691 .2290237

-------------+----------------------------------------------------------------

\_4v3 |

race |

Black | -.3065161 .1648533 -1.86 0.063 -.6296227 .0165905

Other | -.3798123 .4723054 -0.80 0.421 -1.305514 .5458893

|

south |

South | .4052292 .138966 2.92 0.004 .1328609 .6775975

\_cons | .0396236 .0855118 0.46 0.643 -.1279765 .2072237

------------------------------------------------------------------------------

Logit (High School):

Logistic regression Number of obs = 2,244

LR chi2(3) = 78.50

Prob > chi2 = 0.0000

Log likelihood = -904.78566 Pseudo R2 = 0.0416

------------------------------------------------------------------------------

hs | Coefficient Std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

race |

Black | -.9151569 .1282466 -7.14 0.000 -1.166516 -.6637983

Other | -.4910998 .5511525 -0.89 0.373 -1.571339 .5891394

|

south |

South | -.4175069 .1259601 -3.31 0.001 -.6643841 -.1706298

\_cons | 2.250353 .0952967 23.61 0.000 2.063574 2.437131

------------------------------------------------------------------------------

Logit (Junior College):

Logistic regression Number of obs = 1,910

LR chi2(3) = 18.95

Prob > chi2 = 0.0003

Log likelihood = -1314.2871 Pseudo R2 = 0.0072

------------------------------------------------------------------------------

sc | Coefficient Std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

race |

Black | -.173837 .1131414 -1.54 0.124 -.3955902 .0479161

Other | 1.745005 .6241267 2.80 0.005 .5217389 2.968271

|

south |

South | -.1495226 .0968386 -1.54 0.123 -.3393228 .0402777

\_cons | .1079773 .0617595 1.75 0.080 -.0130691 .2290237

------------------------------------------------------------------------------

Logit (College):

Logistic regression Number of obs = 967

LR chi2(3) = 11.05

Prob > chi2 = 0.0114

Log likelihood = -663.06592 Pseudo R2 = 0.0083

------------------------------------------------------------------------------

c | Coefficient Std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

race |

Black | -.3065161 .1648533 -1.86 0.063 -.6296227 .0165905

Other | -.3798123 .4723054 -0.80 0.421 -1.305514 .5458893

|

south |

South | .4052292 .138966 2.92 0.004 .1328609 .6775974

\_cons | .0396236 .0855118 0.46 0.643 -.1279765 .2072236

------------------------------------------------------------------------------

**(b)** *Considerar la base de datos “gss.dta”. La misma posee datos de la encuesta GSS (General Social Survey). Esta encuesta realiza investigaciones científicas básicas sobre la estructura y el desarrollo de la sociedad estadounidense con un programa de recopilación de datos diseñado tanto para monitorear el cambio social dentro de Estados Unidos como para comparar a Estados Unidos con otras naciones. Iniciado en 1972, el GSS contiene un núcleo estándar de preguntas demográficas, de comportamiento y de actitud, además de temas de especial interés. Muchas de las preguntas centrales se han mantenido sin cambios desde 1972 para facilitar los estudios de tendencias temporales, así como la replicación de hallazgos anteriores. En este ejercicio, se utilizan datos de educación similares a los del inciso anterior. Estimar un logit secuencial, interpretar los resultados y mostrar el efecto de la educación del padre en las decisiones de educación en cada transición.*

Logit secuencial:

Number of obs = 9,842

LR chi2(18) = 2461.15

Log likelihood = -9530.0004 Prob > chi2 = 0.0000

--------------------------------------------------------------------------------------

degree | Coefficient Std. err. z P>|z| [95% conf. interval]

---------------------+----------------------------------------------------------------

\_1\_2\_3v0 |

south | -.7967635 .0736484 -10.82 0.000 -.9411116 -.6524153

coh | .7483053 .3414704 2.19 0.028 .0790356 1.417575

|

c.coh#c.coh | -.0482221 .0400122 -1.21 0.228 -.1266445 .0302003

|

paeduc | .1124402 .0778119 1.45 0.148 -.0400684 .2649488

|

c.paeduc#c.coh | .0469452 .0369009 1.27 0.203 -.0253792 .1192696

|

c.paeduc#c.coh#c.coh | -.0050879 .0041484 -1.23 0.220 -.0132187 .0030428

|

\_cons | -1.782385 .6862366 -2.60 0.009 -3.127385 -.4373864

---------------------+----------------------------------------------------------------

\_2\_3v1 |

south | .0469273 .0521384 0.90 0.368 -.055262 .1491166

coh | .3228634 .4189998 0.77 0.441 -.498361 1.144088

|

c.coh#c.coh | -.0371565 .0445171 -0.83 0.404 -.1244084 .0500954

|

paeduc | .1222627 .0808644 1.51 0.131 -.0362286 .280754

|

c.paeduc#c.coh | .0188174 .0344105 0.55 0.584 -.0486259 .0862607

|

c.paeduc#c.coh#c.coh | -.000731 .0035726 -0.20 0.838 -.0077331 .0062712

|

\_cons | -3.497795 .956858 -3.66 0.000 -5.373202 -1.622388

---------------------+----------------------------------------------------------------

\_3v2 |

south | .0710731 .0976914 0.73 0.467 -.1203984 .2625446

coh | .9594559 .8457289 1.13 0.257 -.6981422 2.617054

|

c.coh#c.coh | -.1700969 .0872356 -1.95 0.051 -.3410755 .0008818

|

paeduc | .3357249 .1775429 1.89 0.059 -.0122528 .6837027

|

c.paeduc#c.coh | -.1217749 .0719208 -1.69 0.090 -.262737 .0191873

|

c.paeduc#c.coh#c.coh | .0155494 .0071984 2.16 0.031 .0014408 .0296579

|

\_cons | -.6964155 2.011413 -0.35 0.729 -4.638713 3.245882

--------------------------------------------------------------------------------------