

## PRÁCTICA 6: MODELOS PROBIT Y LOGIT (II)

- Contrastes de hipótesis
- Porcentaje de aciertos en la predicción
- Comparación de modelos

- Se ha estimado el siguiente modelo logit:

```
=====
Dependent Variable: IMASD
Method: ML - Binary Logit (Quadratic hill climbing)
Sample(adjusted): 1 1462
Included observations: 1462 after adjusting endpoints
Convergence achieved after 5 iterations
Covariance matrix computed using second derivatives
=====
      Variable      Coefficient Std. Error z-Statistic Prob.
=====
              C      -8.132017    0.736705   -11.03836    0.0000
              BIG       0.677001    0.224421    3.016656    0.0026
            CONCEN       0.009303    0.002929    3.176155    0.0015
             LSALE       0.565199    0.059229    9.542605    0.0000
            PROPEX       1.406879    0.349063    4.030439    0.0001
=====
Mean dependent var    0.504104    S.D. dependent var 0.500154
S.E. of regression    0.407299    Akaike info criter 1.014575
Sum squared resid    241.7055    Schwarz criterion  1.032659
Log likelihood       -736.6546    Hannan-Quinn criter 1.021321
Restr. log likelihood -1013.332    Avg. log likelihood -0.503868
LR statistic (4 df)   553.3546    McFadden R-squared 0.273037
Probability(LR stat) 0.000000
=====
Obs with Dep=0          725    Total obs          1462
Obs with Dep=1          737
=====
```

Piense en:

- o ¿Qué utilidad tienen los estadístico Akaike, Schwarz, Hannan-Quinn?

- Se ha estimado el modelo logit sin las variables BIG y LSALE, obteniéndose:

```

=====
Redundant Variables: BIG LSALE
=====
F-statistic          254.0976      Probability          0.000000
Log likelihood ratio 405.3496      Probability          0.000000
=====

Test Equation:
Dependent Variable: IMASD
Method: ML - Binary Logit (Quadratic hill climbing)
Sample: 1 1462
Included observations: 1462
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives
=====
      Variable      Coefficient Std. Error z-Statistic Prob.
=====
              C      -0.850478    0.112407   -7.566034    0.0000
            CONCEN      0.015556    0.002532    6.142798    0.0000
            PROPEX      3.206805    0.369130    8.687469    0.0000
=====
Mean dependent var    0.504104    S.D. dependent var 0.500154
S.E. of regression    0.472703    Akaike info criter 1.289096
Sum squared resid     326.0113    Schwarz criterion  1.299946
Log likelihood        -939.3295    Hannan-Quinn criter 1.293144
Restr. log likelihood -1013.332    Avg. log likelihood -0.642496
LR statistic (2 df)   148.0050    McFadden R-squared 0.073029
Probability(LR stat)  0.000000
=====
Obs with Dep=0         725      Total obs          1462
Obs with Dep=1         737
=====

```

Piense en lo siguiente:

- o ¿Qué se está contrastando en el cuadro anterior?
- o ¿Cómo se ha obtenido el logaritmo de la razón de verosimilitud?
- o ¿Qué distribución sigue el estadístico logaritmo de la razón de verosimilitud bajo la hipótesis nula?
- o ¿Cuál es la conclusión del contraste?

- Se ha obtenido el siguiente cuadro que compara valores predichos 0 ó 1 con valores observados

```

=====
Dependent Variable: IMASD
Method: ML - Binary Logit (Quadratic hill climbing)
Sample(adjusted): 1 1462
Included observations: 1462 after adjusting endpoints
Prediction Evaluation (success cutoff C = 0.5)
=====

```

|                | Estimated Equ |        |       | Constant Probability |        |       |
|----------------|---------------|--------|-------|----------------------|--------|-------|
|                | Dep=0         | Dep=1  | Total | Dep=0                | Dep=1  | Total |
| P(Dep=1) <= C  | 598           | 220    | 818   | 0                    | 0      | 0     |
| P(Dep=1) > C   | 127           | 517    | 644   | 725                  | 737    | 1462  |
| Total          | 725           | 737    | 1462  | 725                  | 737    | 1462  |
| Correct        | 598           | 517    | 1115  | 0                    | 737    | 737   |
| % Correct      | 82.48         | 70.15  | 76.27 | 0.00                 | 100.00 | 50.41 |
| % Incorrect    | 17.52         | 29.85  | 23.73 | 100.00               | 0.00   | 49.59 |
| Total Gain*    | 82.48         | -29.85 | 25.85 |                      |        |       |
| Percent Gain** | 82.48         | NA     | 52.14 |                      |        |       |

```

=====

```

Responda:

- o ¿Cómo se asignan las predicciones de ceros?
- o ¿Cómo se asignan las predicciones de unos?
- o ¿Qué importancia tiene C en el cuadro anterior? (ver dos primeras filas)
- o ¿Cuál es el porcentaje de aciertos y cómo se ha obtenido?
- o ¿Cuál habría sido el porcentaje de aciertos en un "modelo ingenuo"? (recuerde dicho "modelo ingenuo")
- o ¿Cuál es la ganancia relativa en términos de aciertos con respecto al "modelo ingenuo"?

- Se ha vuelto a estimar la especificación anterior utilizando un probit

```

=====
Dependent Variable: IMASD
Method: ML - Binary Probit (Quadratic hill climbing)
Sample(adjusted): 1 1462
Included observations: 1462 after adjusting endpoints
Convergence achieved after 5 iterations
Covariance matrix computed using second derivatives
=====

```

| Variable | Coefficient | Std. Error | z-Statistic | Prob.  |
|----------|-------------|------------|-------------|--------|
| C        | -4.791662   | 0.419082   | -11.43372   | 0.0000 |
| BIG      | 0.409806    | 0.131290   | 3.121390    | 0.0018 |
| CONCEN   | 0.005298    | 0.001702   | 3.112597    | 0.0019 |
| LSALE    | 0.332141    | 0.033929   | 9.789316    | 0.0000 |
| PROPEX   | 0.858762    | 0.196385   | 4.372840    | 0.0000 |

```

=====
Mean dependent var      0.504104      S.D. dependent var 0.500154
S.E. of regression      0.407438      Akaike info criter 1.014963
Sum squared resid       241.8702      Schwarz criterion  1.033046
Log likelihood          -736.9379      Hannan-Quinn criter 1.021708
Restr. log likelihood   -1013.332      Avg. log likelihood -0.504061
LR statistic (4 df)     552.7881      McFadden R-squared 0.272758
Probability(LR stat)    0.000000
=====
Obs with Dep=0          725      Total obs          1462
Obs with Dep=1          737
=====

```

Observe:

- o Similitudes entre los estadísticos de comparación de los modelos probit y logit
- Llamemos IMASDLP, IMASDL, IMASDP a la predicción de la variable dependiente en un modelo lineal de probabilidad, un modelo logit y un modelo probit respectivamente. Se han representado los primeros 60 valores junto con la variable IMASD original.

| IMASD    | IMASDLP  | IMASDL   | IMASDP   |
|----------|----------|----------|----------|
| 1.000000 | 0.871987 | 0.888857 | 0.886711 |
| 0.000000 | 0.351760 | 0.334812 | 0.338408 |
| 1.000000 | 0.923612 | 0.914158 | 0.916820 |
| 0.000000 | 0.098685 | 0.108998 | 0.106102 |
| 0.000000 | 0.336893 | 0.322624 | 0.322745 |
| 0.000000 | 0.352248 | 0.341041 | 0.340800 |
| 0.000000 | 0.279795 | 0.252800 | 0.258002 |
| 1.000000 | 0.261347 | 0.232562 | 0.237473 |
| 1.000000 | 0.502026 | 0.540336 | 0.530794 |
| 1.000000 | 0.972801 | 0.931770 | 0.938316 |
| 1.000000 | 0.809577 | 0.869975 | 0.862670 |
| 0.000000 | 0.359724 | 0.346531 | 0.348407 |
| 0.000000 | 0.328399 | 0.309616 | 0.311816 |
| 0.000000 | 0.096031 | 0.108353 | 0.104845 |
| 1.000000 | 0.430915 | 0.449104 | 0.440776 |
| 1.000000 | 0.162299 | 0.149909 | 0.149955 |
| 0.000000 | 0.181496 | 0.165445 | 0.165650 |
| 0.000000 | 0.157527 | 0.147804 | 0.146848 |
| 0.000000 | 0.107835 | 0.115218 | 0.112184 |
| 0.000000 | 0.142757 | 0.134852 | 0.136536 |
| 0.000000 | 0.729004 | 0.806948 | 0.795722 |
| 1.000000 | 0.400656 | 0.398425 | 0.398663 |
| 0.000000 | 0.364140 | 0.354412 | 0.354597 |
| 1.000000 | 0.785528 | 0.829426 | 0.822181 |
| 0.000000 | 0.299326 | 0.276470 | 0.279091 |
| 0.000000 | 0.365629 | 0.355815 | 0.356235 |
| 0.000000 | 0.505955 | 0.550303 | 0.537522 |
| 0.000000 | 0.211555 | 0.188616 | 0.190827 |
| 0.000000 | 0.444004 | 0.457788 | 0.454413 |
| 0.000000 | 0.189715 | 0.171870 | 0.172440 |
| 0.000000 | 0.253488 | 0.228079 | 0.231310 |
| 0.000000 | 0.151178 | 0.142021 | 0.141556 |
| 0.000000 | 0.458833 | 0.477396 | 0.475803 |
| 0.000000 | 0.354826 | 0.343139 | 0.343492 |
| 0.000000 | 0.387176 | 0.383140 | 0.382656 |
| 1.000000 | 0.480186 | 0.511659 | 0.502867 |
| 0.000000 | 0.129833 | 0.127048 | 0.126011 |
| 0.000000 | 0.471917 | 0.499727 | 0.491913 |
| 1.000000 | 1.012200 | 0.944696 | 0.952610 |
| 0.000000 | 0.384528 | 0.382573 | 0.380376 |
| 0.000000 | 0.415640 | 0.420662 | 0.418458 |
| 0.000000 | 0.199778 | 0.177745 | 0.180097 |
| 0.000000 | 0.108243 | 0.114517 | 0.112051 |
| 1.000000 | 0.294235 | 0.270015 | 0.273208 |
| 0.000000 | 0.501149 | 0.536455 | 0.532779 |
| 1.000000 | 0.472768 | 0.499883 | 0.494096 |
| 0.000000 | 0.431366 | 0.441701 | 0.438519 |
| 0.000000 | 0.246635 | 0.219634 | 0.223369 |
| 0.000000 | 0.384649 | 0.382564 | 0.380467 |
| 1.000000 | 0.730364 | 0.806828 | 0.798288 |
| 0.000000 | 0.293740 | 0.267318 | 0.271862 |
| 0.000000 | 0.270051 | 0.239800 | 0.249507 |
| 0.000000 | 0.335881 | 0.320836 | 0.321354 |
| 0.000000 | 0.273381 | 0.246487 | 0.250492 |
| 0.000000 | 0.213069 | 0.190332 | 0.192346 |
| 1.000000 | 0.555788 | 0.617684 | 0.603538 |
| 1.000000 | 1.066147 | 0.958570 | 0.967653 |
| 1.000000 | 1.060104 | 0.957485 | 0.966064 |
| 1.000000 | 0.915003 | 0.909928 | 0.911704 |
| 1.000000 | 0.845020 | 0.873040 | 0.868808 |

- o Comente estos resultados (identifique aciertos, errores, predicciones sin sentido y grado de discrepancia entre los tres modelos)