

Taller #10
Econometría 06216
Modelos Logit y Probit

Profesor: Julio César Alonso

Monitores: Paúl Semaán

Francisco Quevedo

Notas:

- o Recuerde que sólo tres preguntas, seleccionadas al azar, serán calificadas.
- o Este taller es para ser entregado en los primeros 10 minutos de la clase del próximo **10 de Mayo**.

INSTRUCCIONES:

- Este taller debe ser escrito en computador y entregado en papel.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.

Trece aspirantes a una beca estatal para subsidiar programas de doctorado obtuvieron diferentes resultados en sus pruebas cuantitativas y verbales. Seis de los trece aspirantes fueron admitidos al programa. Con base en los datos del archivo T10-01-06.xls. Conteste las siguientes preguntas. De acuerdo con la información anterior:

1. Estime el modelo MLP para estimar la probabilidad de admisión al programa con base en los puntajes cuantitativos y cualitativos de la prueba.
2. ¿Considera usted que el modelo estimado en el punto 1, es un modelo satisfactorio? Sustente suficientemente su respuesta.
3. Estime el modelo Logit y reporte sus resultados en una tabla, no olvide mostrar en ésta cuál es la razón de máxima verosimilitud y el logaritmo de la función de máxima verosimilitud. Comente sobre la significancia conjunta e individual de los coeficientes.
4. Estime el modelo Probit y reporte sus resultados en una tabla, no olvide mostrar en ésta cuál es la razón de máxima verosimilitud y el logaritmo de la función de máxima verosimilitud. Comente sobre la significancia conjunta e individual de los coeficientes.
5. Encuentre el efecto marginal de los resultados en las pruebas cuantitativas y verbales para ambos modelos para cada uno de los trece individuos, a su vez determine el efecto marginal promedio.
6. Suponga que una hora de estudio más tiende a aumentar en 2 puntos el puntaje verbal y en un punto el puntaje cuantitativo. Además suponga que se cuenta con una hora más para ser dedicada a estudiar. Ahora responda (muestre los cálculos o razonamiento que sustenta su respuesta):
 - a. ¿Cuál es el individuo que tendría un mayor incremento en la probabilidad de ser admitido dado que dedica el tiempo disponible a estudiar en el área cuantitativa?
 - b. ¿Cuál es el individuo que tendría un mayor incremento en la probabilidad de ser admitido dado que dedica el tiempo disponible a estudiar en el área verbal?
 - c. Suponiendo que los individuos son racionales, ¿cuál sería el efecto promedio sobre la probabilidad de ser aceptados si los exámenes se aplazarán por un día? (suponiendo que esto brinda 12 horas más de estudio a cada aspirante, y todos los aspirantes hacen uso de estas horas para estudiar).

Taller #10

Respuestas Sugeridas
Econometría 06216
Modelos Logit y Probit

Profesor: Julio César Alonso

Monitores: Paúl Semaán

Francisco Quevedo

Notas:

- o Recuerde que sólo tres preguntas, seleccionadas al azar, serán calificadas.
- o Este taller es para ser entregado en los primeros 10 minutos de la clase del próximo 10 de Mayo.

INSTRUCCIONES:

- Este taller debe ser escrito en computador y entregado en papel.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.

Trece aspirantes a una beca estatal para subsidiar programas de doctorado obtuvieron diferentes resultados en sus pruebas cuantitativas y verbales. Seis de los trece aspirantes fueron admitidos al programa. Con base en los datos del archivo T10-01-06.xls. Conteste las siguientes preguntas.

De acuerdo con la información anterior:

1. Estime el modelo MLP para estimar la probabilidad de admisión al programa con base en los puntajes cuantitativos y cualitativos de la prueba.

Al estimar el modelo por MCO, se obtienen los resultados reportados en la tabla 1. NO obstante, conocemos que por construcción el modelo estimado posee problemas de heteroscedasticidad. Para resolver el problema se emplea el método de MCP, empleando las ponderaciones: $w_i = \sqrt{\hat{y}_i(1 - \hat{y}_i)}$. Donde $\hat{y}_i$ corresponde al valor estimado por el método MCO. Ahora, es importante resaltar que en este caso obtenemos para 3 casos (estudiante 2, 5 y 13) probabilidades estimadas mayores que uno o negativas. Una solución es asignar el valor de uno para aquellas probabilidades estimadas mayores que uno y el valor de cero si la probabilidad estimada es negativa. Pero esto implicará que $w_i = 0$ y por tanto la división estará interminada.

Así los casos de los estudiantes estudiante 2, 5 y 13 son eliminados obteniendo el resultado que se reporta en la tabla 1 bajo la columna correspondiente a la estimación por MCP.

Tabla 1. Estimación modelo MLP

VARIABLE DEPENDIENTE: Admitido,
Estadísticos t entre paréntesis

	MCO
constante	-2.86733 (-3.42) ***
Q_i	0.00313 (2.98) ***
V_i	0.00234 (3.44) ***
R^2	64.93
R^2 Ajustado	57.92
F Global	9.26 ***
# de Obs.	13

(*) nivel de significancia: 10%

(**) nivel de significancia: 5%

(***) nivel de significancia: 1%

MCO: Mínimo Cuadrado Ordinario

2. ¿Considera usted que el modelo estimado en el punto 1, es un modelo satisfactorio? Sustente suficientemente su respuesta.

A pesar de que estadísticamente el modelo aparenta ser tener un buen ajuste, de todas formas como lo discutimos en clase el modelo MLP no es satisfactorio debido a los múltiples problemas que presenta, siendo {estos principalmente la no normalidad en la distribución del término de error, heteroscedasticidad, etc. Por lo tanto se estima el modelo por medio de los MCP, para eliminar el problema de heteroscedasticidad, tal como lo discutimos en clase.

3. Estime el modelo Logit y reporte sus resultados en una tabla, no olvide mostrar en ésta cuál es la razón de máxima verosimilitud y el logaritmo de la función de máxima verosimilitud. Comente sobre la significancia conjunta e individual de los coeficientes.

La estimación de éste modelo se reporta en la tabla 2.

Este modelo se puede expresar de la forma:

$$P(\text{Admitido} = 1 | x_i) = \Phi(\beta^T x_i^T)$$

Donde $\Phi(z) = \int_{-\infty}^z \phi(v) dv = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{v^2}{2}} dv$, es decir, la función de densidad de la normal

estándar, x_i corresponden al vector fila compuesto por algunas características de los estudiantes admitidos, y por consiguiente, el método de máxima verosimilitud para la estimación de los coeficientes viene definido por:

$$Max_{\beta} P(y|X) = Max_{\beta} \left[\prod_{i=1}^n \Phi(\hat{\beta}^T x_i^T) \right] \text{ equivalente a } Max_{\beta} l(y|X) = Max_{\beta} \left[\sum_{i=1}^n \ln[\Phi(\hat{\beta}^T x_i^T)] \right]$$

Los resultados se muestran en la Tabla 1, al final del taller.

Los coeficientes estimados no son significativos conjunta e individualmente. Por su parte, la prueba de Wald, al tener un valor tan bajo, no nos permite rechazar la hipótesis nula de que todos los coeficientes asociados a variables explicativas son conjuntamente iguales a cero.

Tabla 2. Estimación de modelos Logit y Probit

	VARIABLE DEPENDIENTE: Admitido, Estadísticos t entre paréntesis			
	Probit EMV		Logit EMV	
	Coefficientes	Efecto Marginal	Coefficientes	Efecto Marginal
constante	-45.13056 (-0.64)		-80.9004 (-0.72)	
Q_i	0.0461 (0.59)	0.0036	0.0834 (0.68)	0.0040
V_i	0.0266 (0.82)	0.0021	0.0469 (0.82)	0.0023
LRI	0.78834		0.78318	
Wald	0.88		0.96	
ln(L)	-1.89908		-1.94539	
# de Obs.	13		13	

(*) nivel de significancia: 10%

(**) nivel de significancia: 5%

(***) nivel de significancia: 1%

Wald: corresponde al estadístico de Wald que comprueba la significancia conjunta de todas las pendientes.

EMV: Estimadores de Máxima Verosimilitud

4. Estime el modelo Probit y reporte sus resultados en una tabla, no olvide mostrar en ésta cuál es la razón de máxima verosimilitud y el logaritmo de la función de

máxima verosimilitud. Comente sobre la significancia conjunta e individual de los coeficientes.

Recuerden que la diferencia fundamental entre los modelos Logit y Probit es el supuesto sobre la distribución de los errores (normal en Probit, Logística en Logit). Así, tenemos:

$$P(\text{Admitidos} = 1 | x_i) = \Lambda(\beta^T x_i^T)$$

Donde $\Lambda(z) = \frac{e^z}{1+e^z}$, es decir, la función de densidad logística, x_i corresponden al vector

fila compuesto por las mismas características de los estudiantes admitidos i que fueron consideradas en el ejercicio anterior, y por consiguiente, el método de máxima verosimilitud para la estimación de los coeficientes viene definido por:

$$Max_{\beta} P(y|X) = Max_{\beta} \left[\prod_{i=1}^n \Lambda(\hat{\beta}^T x_i^T) \right] \text{ Equivalente a } Max_{\beta} l(y|X) = Max_{\beta} \left[\sum_{i=1}^n \ln[\Lambda(\hat{\beta}^T x_i^T)] \right]$$

Los resultados se muestran en la Tabla 2.

Los coeficientes estimados no son significativos conjunta e individualmente. Por su parte, la prueba de Wald, al tener un valor tan bajo, no nos permite rechazar la hipótesis nula de que todos los coeficientes asociados a variables explicativas son conjuntamente iguales a cero.

5. Encuentre el efecto marginal de los resultados en las pruebas cuantitativas y verbales para ambos modelos para cada uno de los trece individuos, a su vez determine el efecto marginal promedio.

Tabla 3 Efecto marginales Modelo Probit

Modelo Probit				
Coeficientes Probit		0.0461	0.0266	-45.131
(i)	($\beta^T x(i)$)	Q(i)	V(i)	
1	4.5505357	5.87E-07	3.38E-07	
2	-8.1458203	7.18E-17	4.14E-17	
3	-3.4055233	5.58E-05	3.21E-05	
4	4.3686507	1.32E-06	7.6E-07	
5	-9.2505873	4.82E-21	2.77E-21	
6	0.0060147	0.018408	0.010598	
7	5.0679747	4.87E-08	2.81E-08	
8	2.9283297	0.000253	0.000146	
9	-3.6017333	2.81E-05	1.62E-05	
10	-1.5736503	0.005337	0.003073	
11	-1.4620883	0.006322	0.00364	
12	-0.3994283	0.016997	0.009786	
13	9.7240577	5.4E-23	3.11E-23	
Media del Efecto Marginal		0.003646	0.002099	

Tabla 4. Efectos marginales Modelo Logit

Modelo Logit					
Coefficientes Logit		0.0834	0.0469	-80.9004	
(i)	($\beta \cdot x(i)$)	$\Lambda(\beta \cdot x(i))$	Q(i)	V(i)	
1	8.26192	0.9997419	2.15E-05	1.21E-05	
2	-14.45778	5.261E-07	4.39E-08	2.47E-08	
3	-5.858314	0.0028479	0.000237	0.000133	
4	7.843923	0.999608	3.27E-05	1.84E-05	
5	-16.54348	6.535E-08	5.45E-09	3.06E-09	
6	0.027393	0.5068479	0.020842	0.01172	
7	9.252786	0.9999042	7.99E-06	4.49E-06	
8	5.185359	0.9944332	0.000462	0.00026	
9	-6.59249	0.0013687	0.000114	6.41E-05	
10	-2.469864	0.077998	0.005997	0.003372	
11	-2.525063	0.0741198	0.005722	0.003218	
12	-0.649439	0.343116	0.018794	0.010569	
13	17.43211	1	2.24E-09	1.26E-09	
Media del Efecto Marginal		0.004018	0.002259		

6. Suponga que una hora de estudio más tiende a aumentar en 2 puntos el puntaje verbal y en un punto el puntaje cuantitativo. Además suponga que se cuenta con una hora más para ser dedicada a estudiar. Ahora responda (muestre los cálculos o razonamiento que sustenta su respuesta):

- a. ¿Cuál es el individuo que tendría un mayor incremento en la probabilidad de ser admitido dado que dedica el tiempo disponible a estudiar en el área cuantitativa?

Nota: la respuesta correcta para esta pregunta en todos sus numerales es que no existe efecto marginal, dado que los coeficientes no son significativos.

Por otro lado, es incorrecto suponer que el efecto de la primera unidad es igual al efecto de la segunda unidad extra y así sucesivamente. Por eso era necesario evaluar la derivada en cada unidad.

Suponiendo que los coeficientes si fueran significativos la respuesta sería el individuo 6 con un efecto aproximadamente igual a 0.02.

- b. ¿Cuál es el individuo que tendría un mayor incremento en la probabilidad de ser admitido dado que dedica el tiempo disponible a estudiar en el área verbal?

Nota: la respuesta correcta para esta pregunta en todos sus numerales es que no existe efecto marginal, dado que los coeficientes no son significativos.

Por otro lado, es incorrecto suponer que el efecto de la primera unidad es igual al efecto de la segunda unidad extra y así sucesivamente. Por eso era necesario evaluar la derivada en cada unidad.

- c. Suponiendo que los individuos son racionales, ¿cuál sería el efecto promedio sobre la probabilidad de ser aceptados si los exámenes se aplazarán por un día? (suponiendo que esto brinda 12 horas más de estudio a cada aspirante, y todos los aspirantes hacen uso de estas horas para estudiar).

Nota: la respuesta correcta para esta pregunta en todos sus numerales es que no existe efecto marginal, dado que los coeficientes no son significativos.

Por otro lado, es incorrecto suponer que el efecto de la primera unidad es igual al efecto de la segunda unidad extra y así sucesivamente. Por eso era necesario evaluar la derivada en cada unidad.