



La perspectiva estudiantil sobre el desempeño del profesor: un modelo de regresión logística ordinal

María Noel Rodríguez-Ayán¹
Unidad Académica de Educación Química
Facultad de Química, Universidad de la República.

RESUMEN

La regresión logística es una técnica estadística de creciente uso en investigaciones en Ciencias de la Salud y en Educación. El objetivo de este trabajo es presentar un modelo de regresión logística ordinal empleando variables educativas, de modo de ofrecer al lector algunos lineamientos metodológicos a partir de datos empíricos, con un enfoque aplicado. Desde una perspectiva constructivista se modela el criterio “percepción estudiantil sobre el desempeño del profesor como facilitador del aprendizaje” por medio de regresión logística ordinal, empleando el programa SPSS. Los datos proceden de una encuesta aplicada a 244 estudiantes de primer año de la Facultad de Química, Uruguay. Las calificaciones estudiantiles, así como su opinión general sobre las actitudes y estrategias del profesor, resultaron ser buenos predictores del criterio. Se discuten detalles sobre la elaboración y selección de modelos, transformación de variables, interpretación de parámetros, análisis de efectos de interacción y confusión, evaluación del ajuste del modelo y empleo de estadísticos de diagnóstico mediante patrones de covariables. Se comparan los resultados de la regresión logística binaria y la ordinal empleando SPSS.

Palabras clave: Regresión logística ordinal, variables ordinales, Educación Superior, evaluación docente.

ABSTRACT

Logistic regression is a statistical analytical technique of increasing use in health care and educational research. The purpose of this paper is to present a logistic regression model for ordinal response variables within an educational environment, in order to provide readers with guidelines from real data and a practical point of view. Based on a constructivist framework, the criterion “students’ perception of the teacher as a facilitator of learning” is modeled by ordinal logistic regression, employing the statistical package SPSS. Data were collected in a survey carried out among 244 students of the first year of the School of Chemistry, Uruguay. Students’ grades as well as their general opinion towards teachers’ strategies and attitudes proved to be effective predictors of the criterion. Details of model-building procedure, transformation of predictors, parameters interpretation, examination of interaction and confounding effects, goodness-of-fit assessment and regression diagnosis based on covariate patterns are discussed. Results of both ordinal and binary logistic regression with SPSS are also compared.

Keywords: Ordinal logistic regression, ordinal response variables, Higher Education, teachers’ assessment.

¹ e-mail: mayan@fq.edu.uy

Facultad de Química, Universidad de la República. CC 1157 CP 11800 Montevideo-Uruguay



1.- Introducción

El objetivo de este trabajo es presentar un modelo de regresión logística ordinal, proporcionado directrices sistematizadas para que los usuarios de esta técnica puedan adoptar decisiones respecto a su elección. El trabajo se enmarca en el estudio de la perspectiva estudiantil sobre el desempeño del profesor en la Universidad.

El uso de la regresión logística para modelar variables educativas se ha incrementado en la última década. En la predicción del rendimiento académico los antecedentes muestran la mejor adecuación de esta técnica frente a la regresión lineal (García, Alvarado y Jiménez, 2000). Sin embargo, su uso en contextos educativos ha estado más orientado a la predicción de la deserción/retención universitaria que a otras cuestiones, tal vez porque en estos casos las variables dependientes son típicamente dicotómicas o categóricas (Peng, So, Stage y St. John, 2002). El aumento de la popularidad de la regresión logística puede ser atribuido en parte al fácil acceso a programas de estadística que ofrecen análisis de este tipo (Peng, Lee e Ingersoll, 2002). Estos autores realizaron un relevamiento de los artículos publicados en el *Journal of Educational Research* durante 1990-2000, encontrando 8 publicaciones en las que se había empleado esta técnica. El tema central, común a todos los trabajos, fue la predicción del rendimiento escolar, abarcando kindergarten, educación primaria y secundaria. En todos los casos la variable dependiente fue dicotómica.

Peng, So, Stage y St. John (2002) relevaron el uso de la regresión logística en tres publicaciones de Educación Superior durante 1988-1999: *Research in Higher Education*, *The Review of Higher Education* y *The Journal of Higher Education*. Encontraron 52 artículos en los que se emplearon los siguientes análisis: logit (47), tobit (1), logit y tobit (1) y regresión logística polinómica (3). Entre los temas abordados hubo 29 artículos sobre matrícula y retención en la Universidad. Solamente 3 artículos abordaron cuestiones sobre los profesores universitarios. A partir de dicho relevamiento los autores detectaron algunas fallas metodológicas, así como información incompleta, en los resultados de algunos autores: ausencia de resultados sobre estudios de interacción entre predictoras, reportes de efectos de interacción procedentes de análisis de efectos de *confusión*, ausencia de resultados de estadísticos de diagnóstico, ausencia de errores típicos de las estimaciones impidiendo al lector analizar su estabilidad, confusión entre odds ratio y probabilidad. Pedhazur (1997) menciona esta confusión y la interpretación de los parámetros de la ecuación de regresión logística de manera análoga a los parámetros de una regresión lineal como los principales errores en que se incurre al emplear la regresión logística.

En este trabajo se analiza la relación entre el criterio “percepción que tienen los estudiantes de la Facultad de Química (FQ) de la figura del docente como sujeto facilitador de su aprendizaje” y un conjunto de predictoras propias de los estudiantes (descriptivas y de opinión/actitud hacia la gestión de los profesores). Aunque la evaluación docente debe realizarse a partir de información procedente de diversas fuentes - enfoque multiestratégico - (Cohen y Manion, 1994; Fernández, Mateo y Muñiz, 1996; Hernández, Fernández y Baptista, 1998; Mateo y Fernández, 1993; Muñiz, García y Virgos, 1991; Muñoz, Ríos de Deus y Abalde, 2002), los cuestionarios de opinión estudiantil constituyen la fuente de información que cuenta con mayor apoyo empírico y la que tiene mayor garantía de fiabilidad y validez de la información recabada (Fernández y otros, 1996).



García-Valcárcel (2001) señala que el docente universitario debe saber qué sucede en el aula, cómo aprenden los estudiantes, cómo organizar el espacio y el tiempo, qué estrategias didácticas son más convenientes. Ejemplos de acciones docentes que propician la motivación para el aprendizaje y por tanto un mayor rendimiento académico y un mayor grado de satisfacción en los alumnos son: presentar objetivos de la asignatura, cuidar la claridad y el ritmo de exposición, mostrar buena disposición para evacuar consultas, mostrarse accesible a los alumnos, flexibilizar los contenidos programáticos para adaptarse a los intereses, necesidades y experiencias de los alumnos, atender a las capacidades cognitivas que se desean desarrollar, etc. (Alonso Tapia, 1997; García-Valcárcel, 1991).

Dado su carácter exploratorio, los resultados presentados pueden servir de referencia para estudios futuros. Las conclusiones constituyen insumos para optimizar la estrategia de indagación por encuesta, como parte de la evaluación docente Institucional. Se explica el procedimiento de construcción de los modelos de regresión logística, la interpretación de las estimaciones, la evaluación del ajuste y estadísticos de diagnóstico, considerando patrones de covariables. Se comparan resultados de regresiones logísticas binaria y ordinal empleando SPSS.

2.- Método

2.1.- Participantes

Participaron 244 estudiantes de una asignatura de primer año de la FQ, común a todas las carreras: Química Farmacéutica (QF), Bioquímica Clínica (BC), Química (Q), Ingeniería Química (IQ) e Ingeniería de Alimentos (IA).

2.2.- Instrumento de medida

Encuesta piloto conteniendo delimitación de objetivos, instrucciones sobre cómo contestar, variables descriptivas y 11 ítems de opinión medidos en una escala Likert de 5 puntos (5 Totalmente de acuerdo, 1 Totalmente en desacuerdo).

Cuadro 1			
Definición y operacionalización de variables			
Nombre	Definición conceptual	Definición operativa	Escala
Variable dependiente			
FACIL	Percepción estudiantil sobre el docente como facilitador de su aprendizaje	"El docente resulta un sujeto facilitador de su aprendizaje"	1 TD 2 D 3 I 4 A 5 TA
Variables independientes			
CARR	Carrera(s)	"Carrera(s)"	1QF 2BC 3IQ 4IA 5Q 6QFyBC 7IQeIA



SITCUR	Situación curricular del alumno	Diferencia entre el año de ingreso a la FQ y el año de la encuesta	0 ingreso 1 1 año atraso 2 2 años atraso
ESTUD	Hábitos de estudio del alumno	"Habitualmente repasa los temas"	1 TD 2 D 3 I 4 A 5 TA
AUTO	Autoevaluación del alumno respecto a su capacidad de resolver situaciones nuevas	"Se siente más capacitado que antes para resolver situaciones nuevas"	1 TD 2 D 3 I 4 A 5 TA
NOTAS	Calificaciones en el parcial	"Puntaje en el parcial"	continua 0-24 puntos
DIF	Dificultad del curso según el alumno	Constructo con 3 indicadores	continua 3-15 puntos
DOC	Conformidad con las actitudes y estrategias didácticas docentes	Constructo con 5 indicadores	continua 5-25 puntos

Las variables latentes DIF y DOC se obtuvieron por análisis de componentes principales con rotación varimax de 8 ítems medidos en escala Likert de 5 puntos. El puntaje de cada factor se obtuvo por suma de los puntajes en los ítems, siendo el α de Cronbach 0,62 y 0,81 respectivamente.

2.3.- Procedimiento

La participación fue voluntaria y anónima. Se suministró el cuestionario en clase, ofreciendo garantías de confidencialidad.

2.4.- Análisis

Se realizaron regresiones ordinales de probabilidad acumulada y regresiones binarias. Se asume que existe una variable continua, subyacente a las medidas ordinales del criterio, que constituye la percepción estudiantil acerca del rol del docente. Por lo tanto se prefiere la regresión ordinal frente a la de categorías adyacentes (Agresti, 1989) o a la multinomial, técnica que no tiene en cuenta la información ordinal contenida en los datos (Agresti, 1989; Bender & Grouven, 1997; Manor, Mathews & Power, 2000). Los análisis incluyen definición de variables dependiente e independientes, estadística descriptiva, transformación de variables, análisis univariado y multivariado, estudio de interacciones, diagnóstico y validación del modelo. El paquete estadístico empleado fue el SPSS versión 11.0.

3.- Resultados

3.1.- Estadística descriptiva y transformación de variables

La Tabla 1 muestra la media, la desviación típica y el estadístico de Kolmogorov-Smirnov (KS) con la corrección de la significación de Lilliefors para las variables continuas.



Tabla 1 Estadísticos de variables continuas			
	NOTAS	DOC	DIF
Media	8,5	16,1	7,9
Desv. típ	3,39	3,5	2,22
KS (sig)	0,095 (0,000)	0,139 (0,000)	0,145 (0,000)

La Tabla 2 muestra las frecuencias de las variables categóricas. Como algunas categorías tienen muy pocos casos, estas variables fueron transformadas. SITCUR se dicotomizó en NSITCUR: ingreso/atraso curricular; CARR se dicotomizó en CARRDIC: carreras biológicas (QF, BC y Q) / ingenieriles (IQ e IA); FACIL se transformó en NFACIL con 3 categorías: TD y D / I / TA y A; ESTUD y AUTO se transformaron en NESTUD y NAUTO con 4 categorías: TD y D / I / A / TA.

Tabla 2 Frecuencias de variables categóricas							
SITCUR		CARR		FACIL	ESTUD	AUTO	
ingresos	181	QF	75	TD	13	18	4
		BC	38	D	43	43	16
1 año de atraso	54	IQ	75	I	85	51	78
		IA	41	A	95	99	116
más de 1 año de atraso	9	Q	9	TA	8	33	30
		QF y BC	5				
		IQ e IA	1				

3.2. Análisis univariado

Se tomó una significación univariada $< 0,25$ como criterio de inclusión de una predictora en el análisis multivariado, para contemplar la posibilidad de que alguna variable con asociación univariada débil, en presencia de otras variables, pudiera tornarse una predictora importante. Los resultados se resumen en la Tabla 3.

Tabla 3 Asociación de NFACIL con cada predictora					
Variable	Estadístico				Sig.
	tau-b	tau-c	Gamma	d	
CARRDIC	0,100	0,114	0,175	0,114	0,096
NSITCUR	-0,162	-0,161	-0,309	-0,210	0,005
NAUTO	0,283	0,286	0,427	0,283	0,000
NESTUD	0	0	0	0	0,997
		Kruskal-Wallis			
NOTAS - DOC - DIF	16,934	- 101,067	- 26, 818		0,000



3.3.- Regresión logística uni y multivariada

Se realizaron regresiones logísticas ordinales univariadas y multivariada (modelo 1) con todas las variables excepto NESTUD, a efectos de comparar los coeficientes de regresión (Tabla 4).

	Coeficiente		Error típico		Wald	
	uni	Mod 1	uni	Mod 1	uni	Mod 1
NOTAS	0,150	0,108	0,037	0,045	16,346	5,804
DOC	0,529	0,508	0,058	0,062	82,566	67,314
DIF	-0,287	-0,026	0,058	0,071	24,504	0,133
NAUTO	0,798	0,286	0,161	0,185	24,683	2,390
NSITCUR	-0,684	-0,408	0,273	0,325	6,302	1,576
CARRDIC	0,394	0,0608	0,239	0,283	2,710	0,046

Los estadísticos de Wald univariados fueron todos significativos, lo cual es esperable dadas las asociaciones significativas de la Tabla 3. En la regresión multivariada los valores de Wald de NOTAS, DOC y NAUTO son superiores a 2, significativos, en tanto que DIF, NSITCUR y CARRDIC no contribuyen al modelo. Esto sugiere la eliminación de estas tres variables, con miras a obtener un modelo más parsimonioso.

NOTAS y DOC prácticamente mantienen los mismos coeficientes de regresión (considerando los errores típicos). Ello significa que las restantes variables, al ser agregadas, no afectan estos efectos, ninguna de las restantes es confundidora de los efectos de NOTAS o DOC. La variable NAUTO en cambio presenta un coeficiente más bajo en el modelo 1 que en el univariado. Esto significa que su relación con NFACIL depende de la inclusión de otras variables en la ecuación (efecto de confusión).

Se realizaron dos regresiones multivariadas adicionales: eliminando las variables DIF, NSITCUR y CARRDIC (modelo 2) y eliminando además NAUTO (modelo 3). En la Tabla 5 se comparan los coeficientes de regresión y su IC 95% para todos los modelos.

Modelo	NOTAS	DOC	NAUTO
Uni	0,150 (0,077-0,223)	0,529 (0,415-0,643)	0,798 (0,483-1,113)
1	0,108 (0,02-0,195)	0,508 (0,386-0,629)	0,286 (-0,077-0,649)
2	0,122 (0,004-0,206)	0,513 (0,395-0,630)	0,293 (-0,066-0,653)
3	0,135 (0,053-0,217)	0,527 (0,411-0,643)	-



NOTAS y DOC tienen prácticamente los mismos coeficientes en todos los modelos. NAUTO tiene un estadístico de Wald ligeramente superior a 2 en los modelos 1 y 2. Su exclusión de la ecuación (modelo 3) no afecta los parámetros de NOTAS y DOC, por lo que se eligió el modelo 3 como más parsimonioso. No interesó por tanto indagar cuál de las variables (NOTAS, DOC o ambas) es confundidora del efecto de NAUTO. La prueba de líneas paralelas para el modelo 3 da un valor de χ^2 (2 gl) de 1,689, con una significación de 0,43, por lo tanto se acepta la hipótesis de paralelismo.

La significación de los términos eliminados también se analizó comparando los estadísticos de razón de verosimilitud de los modelos multivariados. Puesto que la técnica PLUM no permite el agregado de variables en bloques para la comparación de $-2LL$ se realizaron dos regresiones binarias, tomando como variable dependiente NFACIL dicotomizada de acuerdo con los siguientes criterios: NFACDIC1 (1 Desacuerdos / 0 Indecisos y Acuerdos) y NFACDIC2 (1 Desacuerdos e Indecisos / 0 Acuerdos). La codificación se eligió de manera tal de estimar en las regresiones binarias la ods de “categoría $\leq k$ / categoría $> k$ ”, igual que PLUM. Las regresiones binarias se realizaron en tres bloques consecutivos, agregando las siguientes variables en cada uno: NOTAS y DOC (bloque 1), NAUTO (bloque 2) y DIF, NSITCUR y CARRDIC (bloque 3). En la Tabla 6 se muestra $-2LL$ y la significación del cambio en $-2LL$ de cada modelo respecto al inmediatamente anterior:

Variable dependiente	- 2LL para modelos anidados			
	- 2LL inicial	-2LL Bloque 1 (sig)	-2LL Bloque 2 (sig)	-2LL Bloque 3 (sig)
NFACDIC1	262,877	166,631 (0,000)	165,583 (0,306)	164,631 (0,813)
NFACDIC2	332,314	231,991 (0,000)	229,635 (0,125)	222,642 (0,072)

Las diferencias entre los bloques 3 y 2 (3 gl) y entre bloques 1 y 2 (1 gl) no son significativas, resultado consistente con que el modelo más parsimonioso es el que contiene NOTAS y DOC (modelo 3, bloque 1). Las dos ecuaciones de regresión binaria son y los dos logits acumulados de PLUM son:

$$\text{Logit}(NFACDIC1 = 1) = 7,735 - 0,547 * DOC - 0,083 * NOTAS \quad (1)$$

$$\text{Logit}(NFACDIC2 = 1) = 10,863 - 0,544 * DOC - 0,168 * NOTAS \quad (2)$$

$$\text{Logit}(NFACIL = 1) = 7,835 - 0,537 * DOC - 0,135 * NOTAS \quad (3)$$

$$\text{Logit}(NFACIL \leq 2) = 10,286 - 0,537 * DOC - 0,135 * NOTAS \quad (4)$$

La diferencia en los signos de los coeficientes se debe a que en la regresión binaria se modela $\text{Log}(P(Y = 1)) = cte + \sum \beta_i X_i$, en tanto que en la ordinal $\text{Log}(P(Y < k)) = cte - \sum \beta_i X_i$. El coeficiente de NOTAS en (1) es la mitad de su valor en las restantes, pero considerando el rango de error (Tabla 5), los intervalos se solapan razonablemente. Para DOC los coeficientes y sus rangos son prácticamente iguales en los 4 casos. Todo esto es coherente con que el test de pruebas paralelas realizado en la regresión



ordinal no haya sido significativo, por lo cual se asume que los coeficientes de las predictoras son los mismos con independencia de la ubicación de los puntos de corte.

3.4.- Selección de la escala adecuada para las variables continuas

Habiendo predictoras continuas hay que verificar si se cumple el supuesto de linealidad con el logit. Se categorizaron las predictoras en cuartiles y se diseñaron tres vectores ortogonales, representando las componentes lineal, cuadrática y cúbica de la tendencia del logit. Se realizaron dos regresiones ordinales, una con NOTAS y las tres componentes para DOC, y otra a la inversa. Entre los vectores ortogonales solamente las componentes lineales resultaron significativas, por lo que se concluye que hay linealidad del logit con NOTAS y con DOC.

3.5.- Estudio de componentes de interacción y multicolinealidad

Para investigar interacción se realizó la regresión ordinal agregando el término NOTAS*DOC al modelo 3. Su estadístico de Wald fue ser 0,088, no significativo. La colinealidad se buscó estudiando la correlación de Pearson entre NOTAS y DOC. El coeficiente, aunque estadísticamente significativo, fue de 0,170, revelando una asociación muy débil entre las dos variables. La conclusión es que no hay interacción ni colinealidad.

3.6.- Evaluación del modelo

Además de los ya estudiados estadísticos de Wald y $\Delta(-2LL)$ se analizaron estadísticos pseudo R^2 , el ajuste del modelo a los datos experimentales y la eficacia predictiva del modelo.

3.6.1.- Pseudo R^2

Los valores de pseudo R^2 de Cox y Snell, Nagelkerke y McFadden son razonables: 0,445, 0,504 y 0,275, si bien deben tomarse con precaución dado que ninguno de ellos explica la varianza de manera análoga al coeficiente R^2 de la regresión lineal.

3.6.2.- Ajuste del modelo a los datos

Se estudió el ajuste para cada logit, de manera análoga a lo sugerido por Hosmer y Lemeshow (1989) para la regresión logística politómica. Interesa determinar para cada logit el estadístico de Hosmer y Lemeshow (HS) - recomendado cuando hay predictoras continuas - y los estadísticos de Pearson (P) y Desviación (D). Para los dos últimos es necesario considerar los patrones de covariables y no los sujetos individualmente, dado que el número de configuraciones de covariables es 126 y al número de casos 244 (Hosmer y Lemeshow, 1989; Peng, So y otros, 2002). Dado que en el menú de SPSS HS es una opción en la regresión binaria, en tanto que P y D son opciones de la regresión ordinal, para obtener las tres estimaciones se hizo uso de ambos procedimientos. Se realizaron dos regresiones ordinales y dos regresiones binarias, empleando NFACDIC1 y NFACDIC2 (Tabla 7).



P y D tienen significación 1, pero ello no es informativo pues las predictoras son continuas y hay un número elevado de casillas con frecuencia cero. Esto invalida la hipótesis de que si el ajuste del modelo a los datos es bueno, ambos estadísticos se distribuyen según chi-cuadrado. Se hizo el cociente entre cada estadístico y el número de grados de libertad: tanto P como D resultan inferiores a los grados de libertad para ambos logits, lo cual indica que el ajuste es bueno. HS tampoco es significativo, indicando que los modelos se ajustan bien a los datos.

Tabla 7				
Estadísticos de bondad de ajuste				
Estadístico	Regresión binaria		Regresión ordinal	
	NFACDIC1	NFACDIC2	NFACDIC1	NFACDIC2
P (123 gl)		-	68,117	80,175
D (123 gl)		-	79,245	87,637
HS (8 gl)	5,343	8,072		-

3.6.3.- Eficacia predictiva del modelo

La Tabla 8 muestra la clasificación y la Tabla 9 algunas medidas de eficacia predictiva.

Tabla 8				
Clasificación				
		Categoría Pronosticada (CP)		
		D	I	A
D	recuento	30	23	3
	res. tip.	6,4	,4	-4,4
NFACIL I	recuento	13	40	32
	res. tip	-,5	1,4	-1,0
A	recuento	0	29	74
	res. tip	-4,3	-1,6	4,1

Tabla 9		
Medidas de eficacia predictiva		
Proporción de:	Operación	%
A reales correctamente pronosticados	74/103	71,8
D reales correctamente pronosticados	30/56	53,6
A reales pronosticados como D	0	0
D reales pronosticados como A	3/56	5,4
pronósticos de A acertados	74/109	67,9
pronósticos de D acertados	30/43	69,8
I reales pronosticados como A o D	45/85	52,9

Entre los D reales (Tabla 8) la tendencia es que su pronóstico también sea D (res.>1,96), o a que no sea A (<-1,96). Análogamente, entre los A el pronóstico tiende a ser A o a no ser D. A y D tienen pronósticos bastante acertados y es muy poco probable que una



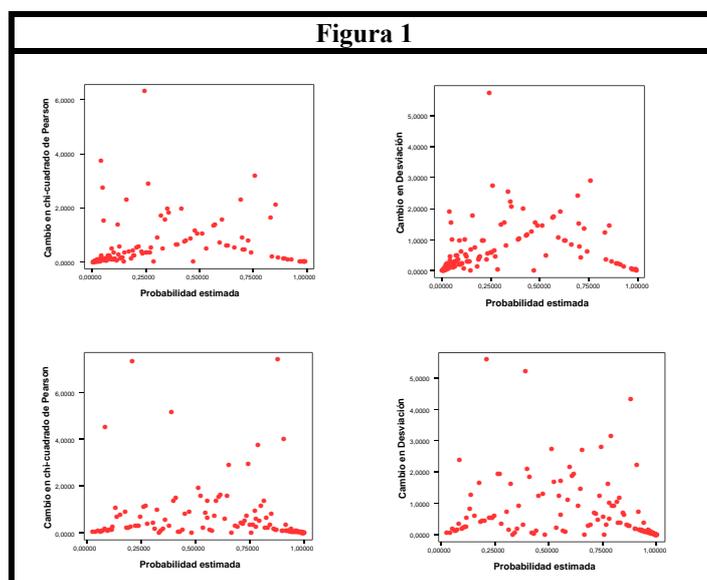
categoría extrema de la escala sea pronosticada como la opuesta. Los I no tienen residuos significativos.

El modelo tiene una buena capacidad predictiva para A: 71,8%, no así de D (53,6%) (Tabla 9). En cuanto a errores, no hay ningún A pronosticado como D y solamente un 5,4 % de D tiene pronósticos de A. Casi un 70% de los pronósticos de A y D son pronósticos correctos. Estos porcentajes deben interpretarse con precaución, puesto que las probabilidades pronosticadas basadas en una ecuación derivada de los mismos datos tienden a arrojar valores inflados, (Pedhazur, 1997). El modelo tiene una capacidad de atribución de opiniones a sujetos indecisos del 52,9%.

La Tabla 8 tiene un valor relativo, pues los residuos se calculan a partir de la hipótesis de independencia entre las variables sin tomar en cuenta que son variables ordinales. Por lo tanto se estimaron también tau-b, tau-c, d (entre 0,53 y 0,563) y Gamma (0,781) para la asociación de NFACIL con CP, lo que muestra que hay clara asociación positiva entre categorías reales y estimadas.

3.7.- Estadísticos de diagnóstico

Se trabajó separadamente con las ecuaciones 3 y 4 para identificar los grupos de casos que afectan el ajuste y/o los parámetros estimados. Los estadísticos de diagnóstico no se pueden estimar a partir de datos individuales, por las mismas razones expuestas al estimar P y D. Estas estimaciones, así como los ΔP y ΔD producidos por la eliminación de cada patrón, fueron realizadas de acuerdo con las recomendaciones de Hosmer y Lemeshow (1989) y los estadísticos así obtenidos coinciden exactamente con los resultados de la Tabla 7. Peng y So (2002) señalan como un error de SPSS que en el menú Regresión Logística se consideren casos y no patrones al estimar estadísticos de bondad de ajuste. Con PLUM SPSS estima P y D correctamente a partir de patrones. Por tanto, en el modelado de variables dicotómicas, una alternativa para evaluar el ajuste es emplear PLUM para estimar P y D. De hecho, la regresión binaria no es más que un caso particular de regresión ordinal, con un número de categorías igual a 2. En la Figura 1 se muestran las gráficas de ΔP y ΔD para ambos logits.





El límite para considerar un ajuste pobre es ΔP o $\Delta D > 4$ (Hosmer y Lemeshow, 1989). Para el primer logit hay sólo un patrón cuyos ΔP y ΔD están muy apartados del resto. Para el segundo logit hay 4 patrones con $\Delta P > 4$, de los cuales 3 también tienen $\Delta D > 4$. En la Tabla 10 se resumen las características de los 5 patrones. ΔP , ΔD y ΔB fueron estimados según las ecuaciones 5.11, 5.12 y 5.10 de Hosmer y Lemeshow (1989). La suma de VD, probabilidad pronosticada, ΔP , ΔD y ΔB se informan *para la regresión correspondiente* pues VD es diferente en cada caso y las 4 estimaciones dependen de la ecuación de regresión.

Patrón n	N° de casos	NOTAS	DOC	Suma de VD	Prob. estimada	ΔP	ΔD	ΔB	Lev.
62	2	8	15	2 ^(a)	0,24 ^(a)	6,3 ^(a)	5,7 ^(a)	0,093 ^(a)	0,015
3	1	1	16	0 ^(b)	0,88 ^(b)	7,4 ^(b)	4,3 ^(b)	0,16 ^(b)	0,021
76	6	9	18	5 ^(b)	0,39 ^(b)	5,2 ^(b)	5,2 ^(b)	0,25 ^(b)	0,046
99	4	11	19	3 ^(b)	0,21 ^(b)	7,3 ^(b)	5,6 ^(b)	0,35 ^(b)	0,046
117	2	14	20	1 ^(b)	0,085 ^(b)	4,5 ^(b)	2,3 ^(b)	0,13 ^(b)	0,027

(a) estimaciones según ecuación 3 (b) estimaciones según ecuación 4

Ningún patrón tiene ΔB elevado, aunque sí ΔP y ΔD . Los patrones 3 y 117 están en las regiones de probabilidad pronosticada en las que es esperable ΔP elevados (Hosmer y Lemeshow, 1989). Los patrones 62, 76 y 99 tienen probabilidades estimadas de 0,24, 0,39 y 0,21 y probabilidades reales (cocientes suma de VD/número de sujetos) de 1, 0,83 y 0,75. Estas diferencias pueden contribuir a generar ΔP elevados: son configuraciones de baja probabilidad pronosticada y valor de variable dependiente contrario a lo que el modelo predice. Conceptualmente, las configuraciones de las predictoras son plausibles y la eliminación de cualquiera de estos patrones no modifica sensiblemente los parámetros de la ecuación de regresión, por lo que no hay razones para excluirlos de los datos.

4.- Discusión de resultados y conclusiones

Las ecuaciones 3 y 4 estiman el logit de “como máximo la opinión k/como mínimo la opinión k”. Al ser negativos los coeficientes de las predictoras, valores altos de éstas se asocian con una mayor probabilidad de categorías de respuesta “elevadas” en la variable dependiente. Sujetos con NOTAS altas, controlando por DOC, tienden a visualizar al docente como un mediador de su proceso de aprendizaje. Análogamente, sujetos con alto grado de conformidad con las estrategias y actitudes docentes, controlando por NOTAS, tienen mayor probabilidad de considerar al docente como facilitador del aprendizaje. Como las predictoras son independientes, es válido afirmar que los alumnos con puntajes altos en NOTAS y en DOC son los que tienen más probabilidad de considerar que el docente facilita su aprendizaje.

Para una mejor interpretación de estos resultados se recurre a las OR. Como no hay efectos de interacción, se puede calcular la OR de cada predictor exponenciando cada coeficiente de regresión. Dadas las escalas de medida de las variables continuas (Cuadro 1) consideramos 5 y 3 puntos respectivamente como una variación de interés en NOTAS y DOC. Tomemos la ecuación 3, que modela la ods de “NFACIL=1/NFACIL>1”. En este caso se está estudiando la ods de *no considerar al docente como un guía para el aprendizaje*, con respecto a afirmar que sí lo es o a estar indeciso. Desde una perspectiva constructivista, según



la cual se concibe al docente como un facilitador del proceso de aprendizaje del alumno, este cociente de probabilidades representa de alguna manera un “índice de fracaso” del profesor en su gestión docente. Esta interpretación obviamente adolece de ser reduccionista, por cuanto la evaluación de la gestión docente no sólo debe nutrirse de las opiniones estudiantiles. Pero se puede considerar como un índice de fracaso, el cual debe ser contrastado con otras medidas e índices antes de pretender arribar a cualquier conclusión sobre la gestión docente. Bajo esos supuestos, la ecuación estima el índice de fracaso docente en función de las calificaciones (NOTAS) y de la opinión general de los alumnos sobre el profesor (DOC).

En una población estudiantil con una misma opinión general hacia el profesor, cada 5 puntos de aumento en NOTAS la diferencia de logits es $-0,675$, lo que equivale a una OR de $0,51$. El índice de fracaso docente se hace la mitad al aumentar 5 puntos las calificaciones. Si controlamos por NOTAS, la OR para DOC es $0,20$, considerando 3 puntos de variación en DOC. El índice de fracaso docente se reduce a la quinta parte al aumentar 3 puntos la opinión sobre el docente, cuando las calificaciones se mantienen en un mismo nivel. Las OR para la ecuación 4 son exactamente las mismas, ya que se cumple el supuesto de líneas paralelas.

5.- Referencias

- Agresti, A. (1989). Tutorial on Modeling Ordered Categorical Response Data. *Psychological Bulletin* 105(2), 290-301.
- Alonso Tapia, J. (1997). *Motivar para el aprendizaje*. Barcelona: Edebé.
- Bender, R. y Grouven, U. (1997). Ordinal logistic regression in medical research. *Journal of the Royal College of Physicians of London*, 31(5), 546-551.
- Cohen, L. y Manion, L. (1994). *Research Methods in Education*. London: Routledge.
- Fernández, J; Mateo, M.A. y Muñiz, J. (1996). Valoración por parte del profesorado de la evaluación docente realizada por los alumnos. *Psicothema*, 8(1), 167-172.
- García, M.V.; Alvarado, JM y Jiménez, A. (2000). La predicción del rendimiento académico: regresión lineal versus regresión logística *Psicothema*, 12(2), 248-252
- García-Valcárcel, A. (1991). El comportamiento de los profesores universitarios en el aula. *Studia Pedagogica*, 23, 135-153.
- García-Valcárcel, A. (2001). *La función docente del profesor universitario, su formación y desarrollo profesional*. En *Didáctica universitaria*. García-Valcárcel (coord.). Madrid: Editorial La Muralla.
- Hernández, R.; Fernández, C. y Baptista, P. (1998). *Metodología de la investigación*. México: McGraw-Hill.
- Hosmer, DW y Lemeshow, S (1989). *Applied Logistic Regression*. USA: John Wiley & Sons.
- Jovell, A.J. (1995). *Análisis de regresión logística*. Madrid: CIS.



- Manor, O.; Matthews, S. y Power, C. (2000). Dichotomous or categorical response? Analysing self-rated health and lifetime social class. *International Journal of Epidemiology* 29, 149-157.
- Mateo, M.A. y Fernández, J. (1993). Dimensiones de la calidad de la enseñanza universitaria. *Psicothema*, 5(2), 265-275.
- Muñiz, J.; García, A. y Virgos, J.M. (1991). Escala de la Universidad de Oviedo para la evaluación del profesorado. *Psicothema*, 3(2), 269-281.
- Muñoz, J.M.; Ríos de Deus, M.P. & Abalde, E. (2002). Evaluación docente vs. evaluación de la calidad. *Revista Electrónica de Investigación y Evaluación Educativa*, 8(2), 103-134. En: http://www.uv.es/RELIEVE/v8n2/RELIEVEv8n2_4.htm. Consultado el 10 de agosto 2004.
- Pardo, A. (2002). *Análisis de datos categóricos*. Universidad Autónoma de Madrid.
- Pedhazur, E. (1997). *Multiple Regression in Behavioral Research*. Fort Worth: Hartcourt Brace College Publishers.
- Peng, C.J; Lee, K.L e Ingersoll, G.M. (2002). An introduction to logistic regression. Analysis and reporting. *The Journal of Educational Research*, 96(1), 3-14
- Peng, C.J y So, TSH (2002). Statistical Software Applications and Review. Modeling strategies in Logistic Regression with SAS, SPSS, Systat, BMDP, Minitab and Stata. *Journal of Modern Applied Statistical Methods*, 1(1), 147-156.
- Peng, C.J.; So, T.S.:H; Stage, F.K. & St. John, E.P. (2002). The use and interpretation of logistic regression in Higher Education Journals: 1988-1999. *Research in Higher Education*, 43(3), 259-293.

Agradecimiento

A la Profesora María Teresa Coello, por su revisión de una versión anterior de este trabajo, en el curso de Regresión Logística del Doctorado en Metodología.