

INTERPRETACION PROBABILISTICA DE LA ESCALA SEMANTICA EN EXPERIMENTOS DE PREFERENCIAS DECLARADAS

Juan de Dios Ortúzar y Rodrigo A. Garrido
Departamento de Ingeniería de Transporte
Pontificia Universidad Católica de Chile
Casilla 306 Código 105 Santiago 22, Chile
Fono: (56-2) 552 2375 Anexos 4270, 4930
Fax: (56-2) 552 4054

RESUMEN

Los experimentos de elección generalizada, utilizando una escala semántica (típicamente de cinco puntos) son fáciles de diseñar y efectuar, y entregan mayor información sobre las preferencias de los entrevistados que los de elección tradicional. Sin embargo, sus resultados son muy dependientes de la interpretación probabilística que hace el analista de la escala semántica utilizada por los individuos para expresar sus preferencias. Este problema se analiza en detalle utilizando una variedad de técnicas estadísticas: regresión lineal, probit ordinal, logit binario y un novedoso enfoque que hemos denominado regresión lineal optimizadora de escala. Los métodos se comparan en base a la bondad de ajuste de los modelos y calidad estadística de sus parámetros, haciendo especial referencia a la razón entre los de tiempos y el de costos, que permite estimar valores subjetivos del tiempo.

En base a lo anterior se definió un diseño factorial fraccional cuyos efectos principales quedan totalmente cubiertos con nueve opciones (ver Kocur et al, 1982). El diseño consideró a cada factor como la diferencia entre los atributos de auto (o bus) y semimetro, salvo en el caso de la variable Intervalo para usuarios de auto (en que el Intervalo es el de semimetro); otra peculiaridad de la opción auto fue que la variable Costo de viaje no era la tarifa como en los otros dos casos, sino que correspondía a la suma del costo de bencina y el costo de estacionamiento (en \$/hr), asumiendo que se debía permanecer ocho horas en el lugar de destino como se señalaba en el cuestionario (ver Ortúzar y Garrido, 1994).

La Figura 1 muestra el instrumento de medición utilizado para el experimento de escalamiento. Este comienza con una breve descripción del viaje hipotético a realizar (hay cuestionarios distintos para usuarios de auto y de bus), y luego se pide al individuo que evalúe las nueve opciones presentadas encerrando en un círculo el número de la sentencia que refleje más fielmente su preferencia en cada caso, de acuerdo a la siguiente escala de cinco puntos: 1: *Siempre en auto*; 2: *Probablemente en auto*; 3: *Me daría igual*; 4: *Probablemente en semimetro*; 5: *Siempre en semimetro* (obviamente "auto" fue reemplazado por "bus" en el formulario para usuarios de bus).

Los niveles de las diferencias entre los atributos (es decir, variables de semimetro menos las variables del modo utilizado) para los formularios de auto y bus, se muestran en la Tabla 1. Se presentan las diferencias entre los atributos, en vez de sus valores absolutos, debido a que el diseño experimental construye las opciones en base a combinaciones de éstas.

Tabla N°1: Diferencias entre atributos consideradas para el diseño

Atributo de Auto o Bus menos Atributo de Semimetro	Nivel de Diferencias entre Atributos					
	Formulario Bus			Formulario Auto		
	Bajo	Medio	Alto	Bajo	Medio	Alto
Costo de viaje (\$)	-10	60	80	20	80	560
Tiempo de viaje (min)	15	25	-	5	15	-
Distancia de caminata (cuadras)	-7	-3	0	0,5	3,5	7,5
Intervalo (min)	-3	2	-	3	8	-

Es evidente que las diferencias mostradas en la Tabla 1 permiten una infinidad de valores posibles para cada atributo (por ejemplo, muchas tarifas de bus y semimetro pueden diferir en \$ -10 y ser consideradas como "tarifa baja"); no obstante, se utilizaron valores similares a los imperantes en el mercado a fin de plantear opciones lo más realistas posibles. El proceso de simulación que apoya al diseño (ver Fowkes y Wardman, 1988) y la encuesta piloto realizada previo a la experiencia definitiva, fueron de gran ayuda para esta definición.

3. PLANTEAMIENTO DEL PROBLEMA

El análisis de los datos de escalamiento consiste en encontrar una relación cuantitativa entre los atributos de cada opción y la respuesta expresada en la escala semántica. Para encontrar esta relación es necesario asociar, a cada sentencia m ($m = 1, \dots, M$) de la escala, un valor numérico que la represente: sea éste R_m . Con esta transformación, es posible postular un modelo lineal del tipo:

$$\theta_0 + \theta_1 X_1 + \theta_2 X_2 + \dots + \theta_K X_K = r_j \quad (1)$$

en que θ_0 es una constante, X_k es la diferencia entre los k -ésimos atributos de los modos alternativos en la opción considerada; θ_k es el coeficiente asociado a X_k y r_j representa alguna transformación de la respuesta del individuo j ; es decir, define una correspondencia única entre la escala semántica y la escala numérica R_m . Por lo tanto, al aplicar el instrumento se obtienen los valores escogidos de la variable dependiente R_m , y conociendo los valores de los atributos X_k es posible realizar un *análisis de regresión múltiple* para estimar los valores de θ_k .

Como existen múltiples escalas numéricas asociables a la escala semántica de respuestas, claramente los resultados del análisis (coeficientes estimados, razones entre éstos y bondad de ajuste del modelo) dependen de la elección de R_m ; esto pone de manifiesto la importancia de elegir correctamente la escala. Varios autores (ver por ejemplo, Bates y Roberts, 1983) han utilizado la escala $R_1 = 2,197$; $R_2 = 0,847$; $R_3 = 0,000$; $R_4 = -0,847$; $R_5 = -2,197$, que corresponde a la transformada Berkson-Theil de las siguientes probabilidades de elección asociadas a cada sentencia de la escala semántica: 0,1; 0,3; 0,5; 0,7; 0,9; otros, como Kocur et al (1982), han utilizado $R_1 = 1$; $R_2 = 2$; $R_3 = 3$; $R_4 = 4$; $R_5 = 5$. Así, realmente no existe consenso acerca de cuál es una escala "apropiada" para un estudio determinado. Debido a esto, parece importante investigar si su elección afecta en forma significativa los resultados del análisis. De hecho, Garrido (1993) demuestra que dependiendo del conjunto de atributos del experimento y de las características socioeconómicas de los individuos, es posible producir fuertes diferencias en la razón entre coeficientes (como el valor subjetivo del tiempo, VST) usando dos escalas distintas.

A modo de ejemplo, en las Tablas 3 y 4 se presentan los VST (razón entre las variables tiempo y costo, ver Gaudry et al, 1989) obtenidos a partir de las tres escalas de probabilidad definidas en la Tabla 2, a las que se aplicó la transformada de Berkson-Theil. Como se puede apreciar que la escala escogida tiene gran influencia en los resultados derivables a partir de los valores de los coeficientes estimados; los VST no sólo difieren entre sí, sino que además provienen de modelos con distinta bondad de ajuste (ver los coeficientes de determinación). Así por ejemplo, en la Tabla 3 se observa una diferencia de 137,6% entre los VST de viaje para las mismas escalas. A su vez, la Tabla 4 muestra una diferencia de 24,5% en los VST de caminata entre las escalas 1 y 2; cabe destacar que esta última no es sólo importante en términos porcentuales, sino que también en valor absoluto (16,2 \$/min), lo que podría tener dramáticas consecuencias en la evaluación de un proyecto de transporte.

Tabla N°2: Escalas utilizadas en el análisis de sensibilidad

	Escala 1	Escala 2	Escala 3
R_1	0,010	0,300	0,200
R_2	0,400	0,450	0,400
R_3	0,500	0,500	0,500
R_4	0,600	0,850	0,880
R_5	0,990	0,950	0,970

Tabla N°3: VST (\$/min) para alumnos usuarios de bus

Tipo de tiempo	Escala 1	Escala 2	Escala 3
Viaje	1,73	3,98	4,11
Espera ⁽¹⁾	18,67	23,89	23,24
Caminata ⁽²⁾	21,63	24,91	24,74
R^2	0,44	0,46	0,45

(1) Asumiendo el tiempo promedio de espera como la mitad del intervalo

(2) Asumiendo una velocidad media de caminata de 4 km/hr

Tabla N°4: VST (\$/min) para funcionarios usuarios de auto

Tipo de tiempo	Escala 1	Escala 2	Escala 3
Viaje	29,87	34,93	35,02
Espera ⁽¹⁾	42,05	66,08	65,06
Caminata ⁽²⁾	66,25	82,45	80,99
R^2	0,31	0,36	0,36

(1) Asumiendo el tiempo promedio de espera como la mitad del intervalo

(2) Asumiendo una velocidad media de caminata de 4 km/hr

4. EL MODELO PROBIT ORDINAL

Una forma de evitar el problema descrito en la sección anterior es utilizar un paradigma de modelación que no requiera conocer *a priori* una escala numérica para la estimación del modelo, sino que permita plantearla como un conjunto adicional de parámetros a ser estimados. McKelvey y Zavoina (1975) desarrollaron un modelo con esta característica, denominado "probit ordinal", cuya fundamentación teórica es la siguiente. Sea Y una variable que represente el grado de preferencia de una alternativa sobre otra, en una opción presentada en un experimento de escalamiento; es decir, si A y B son el par de alternativas a comparar, Y representaría cuán preferible sería A sobre B (o vice-versa). Supóngase, además, que esta variable satisface el siguiente modelo lineal:

$$Y_j = X_j\theta + \epsilon_j \quad (2)$$

con X_j un vector k-dimensional de variables independientes para el individuo j ($j = 1, \dots, n$); es decir, son los atributos del individuo y de las alternativas a comparar; θ es el vector de coeficientes del modelo, y ϵ_j es un término de error aleatorio que se supone con distribución Normal multivariada, valor esperado cero y matriz de varianza-covarianza $\sigma^2 I$, en que I es a la matriz identidad. Sin embargo, debido a que las técnicas estadísticas existentes son incapaces de medir el verdadero valor de Y_j , solamente es posible observar una versión ordinal de esta variable, denominada Z_j , que no satisface el modelo lineal de la expresión (2). Se asumirá que Z_j es una variable que cuenta con M categorías de respuesta R_1, R_2, \dots, R_M , que se relacionan con la variable inobservable Y_j de la siguiente forma. Supongamos que $\mu_0, \mu_1, \dots, \mu_M$, son $M+1$ números reales, con:

$$\begin{aligned} \mu_0 &= -\infty \\ \mu_M &= +\infty \end{aligned}$$

Además, estos números cumplen la siguiente condición:

$$\mu_0 \leq \mu_1 \leq \dots \leq \mu_{M-1} \leq \mu_M$$

tal que:

$$\begin{aligned} Z_j \in R_m &\Leftrightarrow \mu_{m-1} < Y_j \leq \mu_m \\ &\text{para } 1 \leq j \leq n \end{aligned} \quad (3)$$

Ahora bien, ya que Z_j es ordinal, puede ser representada por una serie de variables mudas:

$$\begin{aligned} Z_{jm} &= \begin{cases} 1 & \text{si } Z_j \in R_m \\ 0 & \text{en otro caso} \end{cases} \\ &\text{para } 1 \leq j \leq n, \quad 1 \leq m \leq M \end{aligned}$$

De las expresiones (2) y (3) se puede escribir directamente la función de probabilidad de la variable dependiente observada Z_j :

para $1 \leq m \leq M$ y $1 \leq j \leq n$

$$\begin{aligned} \mu_{m-1} < Y_j < \mu_m \Rightarrow \mu_{m-1} < \sum_{k=0}^K \theta_k X_{kj} + e_j < \mu_m \\ \Rightarrow \frac{\mu_{m-1} - \sum_{k=0}^K \theta_k X_{kj}}{\sigma} < \frac{e_j}{\sigma} < \frac{\mu_m - \sum_{k=0}^K \theta_k X_{kj}}{\sigma} \end{aligned} \quad (4)$$

donde X_0 vale 1 (es decir, θ_0 es una constante específica). Ahora, dado que e_j tiene una distribución Normal multivariada se puede escribir:

$$Pr(Z_{jm} = 1) = Pr(Z_j \in R_m) = \Phi\left(\frac{\mu_m - \sum_{k=0}^K \theta_k X_{kj}}{\sigma}\right) - \Phi\left(\frac{\mu_{m-1} - \sum_{k=0}^K \theta_k X_{kj}}{\sigma}\right) \quad (5)$$

en que $\Phi(\cdot)$ representa la función Normal acumulada estándar. Finalmente, se debe tener presente que cualquier transformación lineal de la variable inobservable Y_j , si es aplicada a la serie μ_m , conduce exactamente al mismo modelo de la expresión (5). Así, para identificar los coeficientes del modelo se asumirá, sin pérdida de generalidad, que $\mu_1 = 0$ y $\sigma = 1$. Con este supuesto se obtiene el modelo definitivo:

$$Pr(Z_{jm} = 1) = \Phi\left(\mu_m - \sum_{k=0}^K \theta_k X_{kj}\right) - \Phi\left(\mu_{m-1} - \sum_{k=0}^K \theta_k X_{kj}\right) \quad (6)$$

Así, el problema de calibración consiste en estimar $M+K-1$ coeficientes $\mu_2, \dots, \mu_{M-1}, \theta_0, \dots, \theta_K$. Más antecedentes acerca del proceso de calibración y de los estadígrafos asociados al modelo pueden consultarse en McKelvey y Zavoina (1975) o Terza (1985).

5. PRINCIPALES RESULTADOS DE LA MODELACION

A fin de comparar los resultados de modelos con distintas escalas de respuesta (variable dependiente), a continuación se presentan los coeficientes de los siguientes cuatro modelos:

- REG1: Modelo de regresión lineal cuya escala de respuesta es la transformada Berkson-Theil de la serie $R_1 = 0,1$; $R_2 = 0,3$; $R_3 = 0,5$; $R_4 = 0,7$; $R_5 = 0,9$. Aún cuando se sabe que una escala fijada a priori por el modelador podría tener gran influencia en los resultados de la modelación, se decidió analizar esta escala porque ha sido utilizada en varios estudios (ver por ejemplo Fowkes y Tweddle, 1988; Bates y Roberts, 1983; Ortúzar y Garrido, 1994).

- b) REG2: Modelo de regresión lineal cuya escala de respuesta se determina durante el proceso de maximización del ajuste del modelo, considerando cada valor de la escala (R_1, R_2, \dots, R_5) como una variable adicional. Esto otorga, en general, $M-1$ grados de libertad al problema de mínimos cuadrados en un modelo con una escala de M puntos.
- c) BIN: Modelo logit binario que utiliza la siguiente transformación de la escala semántica (ver Figura 1): Cualquier respuesta marcada a la izquierda de "Me daría igual", corresponde directamente a una elección del modo habitual (auto o bus); a su vez, cualquier respuesta marcada a la derecha de "Me daría igual", corresponde a una elección de semimetro. Las respuestas marcadas "Me daría igual" se eliminan.
- d) PRO: Modelo probit ordinal.

Dado que estos modelos difieren en algunas de sus características, tales como forma funcional, tratamiento de la información o distribución estocástica de los errores, no es posible comparar directamente sus coeficientes. Por esta razón es conveniente comparar estos resultados en base a los VST calculados de acuerdo al procedimiento estándar (Gaudry et al, 1989).

La Tabla 5 presenta los valores óptimos para la escala encontrada mediante REG2 para dos grupos (estudiantes usuarios de bus y funcionarios usuarios de auto); también se reproducen los resultados para REG1 a fin de facilitar su comparación.

Tabla N°5: Valores para la escala óptima (REG2)

Escala	REG1	Estudiantes usuarios de bus	Funcionarios usuarios de auto
R_1	0,1	0,284	0,228
R_2	0,3	0,286	0,278
R_3	0,5	0,500	0,500
R_4	0,7	0,714	0,722
R_5	0,9	0,900	0,842

Se aprecia que los valores de la escala de REG2 no son simétricos con respecto al punto "Me daría igual" (como lo son los valores de la escala de REG1). Esto podría parecer, a primera vista, contrario a lo esperado. Sin embargo, si se examina cuidadosamente el significado de la escala se llega a la conclusión que los puntos de ésta no tienen por qué ser simétricos y, más aún, no tienen por qué estar igualmente distanciados unos de otros. En efecto, los valores de la escala óptima reproducen aquel valor de cada punto para los cuales hay un mejor ajuste del modelo; éste asume que las preferencias son declaradas en una escala discreta (de cinco puntos en este caso) sin embargo no es capaz de discriminar preferencias "intermedias" (o continuas) en las cuales, por ejemplo, dos individuos marcaron el punto 1 ("Siempre en auto") aún cuando

uno de ellos haya estado muy seguro de su decisión y el otro haya tenido una preferencia mayor que "Probablemente en auto" (punto 2) pero poco menor que "Siempre en auto"; claramente el punto 1 no representa fielmente las preferencias de ambos. Por otra parte la utilidad asociada, por ejemplo, a "Probablemente en semimetro" puede diferir del punto de indiferencia entre ambos modos en una cantidad muy distinta de la que lo está el punto "Probablemente en auto", dependiendo de los valores de los parámetros de la utilidad para los individuos y de la diferencia entre los niveles de servicio de cada modo.

En las Tablas 6 y 7 a continuación, se presentan los VST obtenidos a partir de los mejores modelos estimados para dos grupos de la muestra (alumnos que usan bus y funcionarios que utilizan auto). En cada tabla se presentan también los estadígrafos de bondad de ajuste general de cada modelo. Para el modelo de regresión y para el probit ordinal se calculó R^2 , que corresponde al coeficiente de determinación ajustado en el primer caso (Wonnacott y Wonnacott, 1979) y a un valor directamente comparable en el segundo (McKelvey y Zavoina, 1975). Finalmente, para el modelo logit binario se calculó el índice ρ^2 (Tardiff, 1976) que obviamente no es comparable con el anterior.

Tabla N°6: VST (\$/min) para alumnos usuarios de bus

Tipo de tiempo	REG1	REG2	PRO	BIN
Viaje	4,01 (2,80)	2,81 (3,06)	1,68 ⁽³⁾ (1,28)	3,79 ⁽³⁾ (1,84)
Espera ⁽¹⁾	20,66 (3,24)	22,17 (3,25)	16,90 (2,87)	17,82 (2,10)
Caminata ⁽²⁾	23,65 (5,94)	23,55 (5,89)	20,74 (6,30)	19,26 (4,96)
R^2, ρ^2	0,48	0,49	0,51	0,42

(1) Asumiendo el tiempo promedio de espera como la mitad del intervalo

(2) Asumiendo una velocidad media de caminata de 4 km/hr

(3) Coeficiente no es significativamente distinto de cero al 95% de confianza

En la Tabla 6 se observa que todos los VST de REG1 y REG2 son significativamente distintos de cero al 95% de confianza; además, REG1 presenta estadísticos de significación menores que REG2 excepto para el VST de caminata. Por otro lado el VST de viaje no es significativamente distinto de cero al 95% de confianza en los modelos probit ordinal y logit binario. Se observa también que los VST de REG1 y REG2 son mayores que los correspondientes VST de los otros dos modelos (considerando los valores significativamente distintos de cero).

Con respecto a la bondad de ajuste general de los modelos, se aprecia que el modelo probit ordinal explica un 2% más de la variabilidad de los datos que REG2 y un 3% más que REG1. A su vez, el modelo logit binario es notablemente más explicativo que su modelo de referencia

(sólo constantes). Por otra parte, se observa que los cuatro modelos adjudican la misma importancia relativa a los VST; es decir, el tiempo más valorado resulta ser el de caminata, seguido por el de espera y el de viaje (es interesante comentar que el orden de los primeros se revierte en los modelos para alumnos usuarios de auto).

En el caso de funcionarios usuarios de auto (Tabla 7), se observa que los VST de los cuatro modelos presentan buenos estadígrafos de significación al 95 % de confianza; también se aprecia que éstos son más altos en REG1 y REG2, y al igual que en los casos anteriores, REG1 presenta estadísticos de significación menores que REG2. Además, se aprecia que la importancia relativa de los VST es la misma en los cuatro modelos: caminata seguido por espera y en tercer lugar tiempo de viaje (este resultado se repite en el caso de funcionarios usuarios de bus). Al igual que en el caso anterior, REG1 y REG2 entregaron VST más altos y mejores estadígrafos de significación que los otros dos modelos.

Tabla N°7: VST (\$/min) para funcionarios usuarios de auto

Tipo de tiempo	REG1	REG2	PRO	BIN
Viaje	33,38 (4,54)	35,98 (5,15)	26,60 (3,81)	22,28 (2,75)
Espera ⁽¹⁾	63,38 (2,35)	68,63 (2,41)	56,53 (2,03)	60,57 (1,96)
Caminata ⁽²⁾	76,68 (6,79)	80,21 (6,79)	75,56 (6,42)	65,18 (6,18)
R^2, ρ^2	0,36	0,37	0,38	0,33

(1) Asumiendo el tiempo promedio de espera como la mitad del intervalo de semímetro;

(2) Asumiendo una velocidad media de caminata de 4 km/hr.

Con respecto a la bondad de ajuste general de los modelos, se observa que el modelo probit ordinal explica la variabilidad de los datos sólo en un 1 % más que REG2 y un 2 % más que REG1. El modelo logit binario, por su parte, es significativamente más explicativo que el modelo de sólo constantes.

Una observación detallada de la Tabla 5 sugiere la posibilidad de examinar si el número original de puntos de la escala semántica (cinco) es apropiado. Por ejemplo, si sólo se ocupa un valor para representar los puntos 1 y 2 de la escala de REG2 (que resultaron muy cercanos), es interesante ver qué consecuencias puede traer esta aparente pérdida de información; notar, por otro lado, que en la nueva escala se tendría un parámetro menos que calibrar. La Tabla 8 muestra los valores óptimos de la escala obtenida cuando R_1 y R_2 son reemplazados por un solo punto R'_1 . En esta escala, como en las anteriores, el valor de la probabilidad asociada a R_3 fue fijado en 0,5 ya que corresponde al punto de indiferencia.

Tabla N°8: Valores de la escala óptima de cuatro puntos

Escala	Estudiantes usuarios de bus	Funcionarios usuarios de auto
R'_1	0,277	0,121
R_2	0,500	0,500
R_4	0,716	0,776
R_5	0,899	0,922

Como se ve, los valores de los puntos de la nueva escala están más espaciados que los de la Tabla 5, lo que no sugiere una nueva fusión de puntos.

La Tabla 9 presenta los VST derivados a partir de los mejores modelos REG2 y PRO en los dos grupos utilizando una escala de cuatro puntos. Se puede ver que el estadígrafo R^2 en la muestra de estudiantes es considerablemente más alto que el correspondiente valor para la escala de cinco puntos. Los VST cambiaron levemente sin seguir alguna regla fácil de descubrir. En el caso de la muestra de funcionarios se puede ver que aunque REG2 mantuvo su bondad de ajuste, PRO logró una considerable mejora. Además, en este caso todos los VST incrementaron sus valores y en general resultaron más significativos.

Tabla N°9: VST (\$/min) utilizando la escala de cuatro puntos

Muestra	Estudiantes usuarios de bus		Funcionarios usuarios de auto	
	REG2	PRO	REG2	PRO
Valor del Tiempo				
Viaje	3,63 (2,53)	2,16 (1,57)	35,42 (4,93)	31,03 (4,14)
Espera	24,57 (3,69)	18,07 (2,81)	72,81 (2,62)	66,37 (2,20)
Caminata	24,88 (6,10)	21,72 (6,08)	83,20 (6,68)	86,05 (6,53)
R^2, ρ^2	0,51	0,54	0,37	0,43

6. CONCLUSIONES

El principal objetivo de esta investigación era la búsqueda de la escala de medición más adecuada que se debe asociar a la escala semántica de respuesta de un experimento de escalamiento. Para esto se analizaron cuatro enfoques diferentes para el mismo problema: la regresión lineal (en dos versiones), el modelo probit ordinal y el modelo logit binario. Los modelos obtenidos en cada segmentación de la muestra considerada fueron sometidos a una comparación en términos de la significación estadística de sus parámetros, similitud de resultados y bondad de ajuste general. En base a la comparación de los modelos obtenidos en cada caso (Tablas 6 y 7) se puede inferir lo siguiente:

- Los modelos adjudican el mismo orden de importancia relativa a los distintos VST en cada submuestra considerada.
- La mayoría de los VST derivados del modelo logit binario resultaron menos significativos que los valores correspondientes obtenidos mediante los otros modelos. La mayoría de los VST obtenidos a partir del modelo REG1 (escala de medición imputada a priori) resultaron menos significativos que los valores correspondientes en el modelo REG2 (escala de medición producto de la maximización de R^2); obviamente REG2 siempre presenta una mejor bondad de ajuste general que REG1.
- El modelo REG2 dio como resultado parámetros de VST significativos al 95 % de confianza en todos los casos (salvo el VST de espera para funcionarios usuarios de bus). Esto indica que tiene mayor robustez que los demás modelos. El modelo probit ordinal presentó en todos los casos una bondad de ajuste general levemente superior a los modelos de regresión. En efecto, la proporción de variación explicada por este modelo fue en promedio un 1,8% superior en el caso de REG2 y un 3,3% superior al modelo REG1.
- A partir de los resultados de REG2 fue posible realizar una redefinición de la escala de medición, logrando una escala de cuatro puntos que permitió un mejor ajuste de los modelos calibrados (tanto de regresión lineal como probit)

Tomando en consideración lo señalado en estos puntos, se puede concluir que las diferencias entre los modelos permiten descartar REG1 en favor de REG2, bajo el supuesto que el proceso de optimización involucrado no consume gran cantidad de recursos. Por otra parte el modelo logit binario requiere una transformación de la escala (ver sección 5) que tiene las siguientes desventajas:

- Pierde gran cantidad de información, ya que agrupa respuestas diversas que señalan distintos grados de preferencia.
- Elimina a las observaciones que se encuentran en situación de indiferencia, lo cual no sólo disminuye el tamaño muestral sino que además deja fuera del análisis a un subconjunto importante de respuestas (las situaciones de indiferencia contribuyen enormemente al análisis de situaciones de iso-utilidad entre las alternativas planteadas).

Los dos aspectos anteriores juegan un rol importante en la significación estadística de los parámetros del modelo, lo cual se refleja en parte en los resultados presentados en las Tablas 6 y 7. Las desventajas mencionadas, tanto de tipo teórico como práctico, llevan a concluir que no es conveniente utilizar este tipo de modelo para el análisis de experimentos de escalamiento.

Con respecto a los modelos de regresión lineal y probit ordinal, se debe tener presente que ambos tienen una fuerte sustentación teórica, que se ve reflejada en la evidencia empírica aquí presentada. Sin embargo, aun cuando el modelo probit ordinal es más flexible que el anterior (ver sección 3), las diferencias entre REG2 y PRO (ver sección 5) no justifican el mayor grado de complejidad y la necesidad de software especializado que el segundo requiere. Además el uso de REG2 permitió detectar la fusión de la escala descrita anteriormente.

Por todo lo anterior, se recomienda utilizar el enfoque de regresión lineal múltiple optimizador de escala, dada su simplicidad y abundante disponibilidad de software de calibración. No obstante, se debe tener presente que este enfoque requiere de un proceso de optimización para la elección de la escala numérica, como se explicó en la sección 5.

AGRADECIMIENTOS

Los autores desean expresar su agradecimiento al Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) y a la Dirección de Investigación de la Pontificia Universidad Católica de Chile (DIUC), por el apoyo recibido durante esta investigación.

REFERENCIAS

Bates, J.J. y Roberts, M. (1983) Recent experience with models fitted to stated preference data. *Proceedings 11th PTRC Summer Annual Meeting*, University of Sussex, Inglaterra, 4-7 Julio 1983.

Fowkes, A.S. y Tweddle, G. (1988) A computer guided stated preference experiment for freight mode choice. *Proceedings 16th PTRC Summer Annual Meeting*, University of Bath, Inglaterra, 12-16 Septiembre 1988.

Fowkes, A.S. y Wardman, M. (1988) The design of stated preference travel choice experiments, with special reference to interpersonal taste variations. *Journal of Transport Economics and Policy* XXII, 27-44.

Garrido, R.A. (1993) Incidencia de la escala semántica en la determinación de razones entre coeficientes en modelos de regresión lineal con preferencias declaradas. *Apuntes de Ingeniería* 50, 49-57.

Gaudry, M.J.L., Jara-Díaz, S.R. y Ortúzar, J. de D. (1989) Value of time sensitivity to model specification. *Transportation Research* 23B, 151-158.

Kocur, G., Adler, T., Hyman, W. y Aunet, B. (1982) Guide to forecasting travel demand with direct utility assessment. Report No. UMTA-NH-11-1-82, US Department of Transportation, Washington, DC.

McKelvey, R.D. y Zavoina, W. (1975) A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology* 4, 103-120.

Ortúzar, J. de D. y Garrido, R.A. (1994) A practical assessment of stated preference methods. *Transportation* 21 (en imprenta).

Tardiff, T.J. (1976) A note on goodness-of-fit statistics for probit and logit models. *Transportation* 5, 377-388.

Terza, J.V. (1985) Ordinal probit: a generalization. *Communications in Statistics - Theory and Method* 14, 1-11.

Wonnacott, R.J. y Wonnacott, T.H. (1979) *Econometrics*. John Wiley & Sons, Nueva York.

Suponga que Ud. debe realizar un viaje a las 8:30 A.M. desde su hogar hasta un destino situado a 10 km. de distancia, donde debe permanecer durante aproximadamente 8 hrs. Existe un servicio de buses, con paradero a una cuadra de su hogar y con un paradero de bajada aproximadamente a una cuadra de su destino. Además se dispone de un servicio de semi metro con una estación ubicada a una cuadra de su destino.

Por favor evalúe las siguiente nueve situaciones referidas a dicho viaje.

FACTORES DEL BUS				FACTORES DEL SEMI - METRO				RESPUESTA			
Tarifa (\$)	Frecuencia (minutos entre pasadas)	Tiempo de viaje (min)	Tarifa (\$)	Frecuencia (minutos entre pasadas)	Distancia entre el hogar y la estación más cercana	Tiempo de viaje (min)	Encierre en un círculo la opción más cercana a su preferencia:				
							Siempre en bus	Probablemente en bus	Me daría igual	Probablemente en semi-metro	Siempre en semi-metro
1. 90	5	40	100	8	1 cuadra	25	1	2	3	4	5
2. 90	5	50	100	8	4 cuadras	25	1	2	3	4	5
3. 90	5	40	100	3	8 cuadras	25	1	2	3	4	5
4. 180	5	50	120	3	4 cuadras	25	1	2	3	4	5
5. 180	5	50	120	8	4 cuadras	25	1	2	3	4	5
6. 180	5	40	120	8	8 cuadras	25	1	2	3	4	5
7. 180	5	40	100	8	1 cuadra	25	1	2	3	4	5
8. 180	5	40	100	3	4 cuadras	25	1	2	3	4	5
9. 180	5	50	100	8	8 cuadras	25	1	2	3	4	5

Figura N° 1: Instrumento de Medición para PD-Escalamiento