

Un Estimador Máximo-Entrópico de Modelos Logit Jerárquicos

Pedro Donoso S.

Laboratorio de Modelamiento del Transporte y Uso del Suelo (LABTUS), Departamento de Ingeniería Civil, Universidad de Chile.
Fono: (56-2) 978 0531; E-mail: pdonoso@labtus.cl

Louis de Grange C.

Escuela de Ingeniería Civil Industrial, Universidad Diego Portales, Santiago de Chile.
Fono: (56-2) 676 8118; E-mail: louis.degrange@udp.cl

RESUMEN

En este trabajo presentamos un nuevo enfoque para estimar modelos de demanda agregados de tipo Logit Jerárquicos. Para ello, formulamos modelos de programación matemática no lineal con componentes entrópicas cuyas condiciones de optimalidad proporcionan modelos tipo Logit Jerárquicos, equivalentes a los modelos basados en la teoría de la utilidad aleatoria, y cuyos multiplicadores de Lagrange corresponden a los parámetros de los modelos Logit Jerárquicos; a estos parámetros les denominamos máximo entrópicos.

A partir de simulaciones, concluimos que los estimadores máximo entrópicos presentan mejores propiedades estadísticas que los clásicos estimadores máximo verosímiles, específicamente en lo que a sesgo se refiere, especialmente en muestras de menor tamaño que son normalmente utilizadas para calibrar modelos Logit Jerárquicos en la práctica.

Palabras Claves: Modelos Logit, Logit Jerárquico, Multiplicadores de Lagrange, Máxima Entropía, Máxima Verosimilitud, Sesgo.

1. INTRODUCCIÓN

Los modelos de elección discreta Logit Multinomial (LMN) y Jerárquico (LJ) son usados frecuentemente para representar la demanda en sistemas de Transporte y de Uso de Suelo; por ejemplo, la elección de destinos y rutas de viaje, modos de transporte y opciones de localización de las personas. Tradicionalmente, estos modelos se han deducido del paradigma de usuario racional que consume o utiliza aquellos servicios que maximizan su utilidad, y de la asignación de una determinada ley de probabilidad a dicha utilidad, y se han estimado valiéndose de criterios estadísticos clásicos, siendo el más común el de maximización de la verosimilitud.

También ambos modelos se han deducido como solución a determinados problemas de maximización de la Entropía con restricciones, siendo sus parámetros los multiplicadores de Lagrange de estas restricciones. Estos parámetros, que en este trabajo se denominan estimadores máximo-entrópicos (ME), son estimadores de los parámetros poblacionales alternativos a los de máxima verosimilitud (MV).

El enfoque de máxima entropía se ha utilizado predominantemente para formular modelos agregados de demanda por transporte, tanto en usuarios como en opciones de elección. En este ámbito, destacan los modelos de distribución espacial de viajes, como el modelo gravitacional doblemente acotado propuesto por Wilson (1970) y otros posteriores que introducen cambios y extensiones del mismo (Fotheringham 1986; Fang y Tsao, 1995; De Grange et al, 2009). Otras aplicaciones relevantes de este enfoque corresponden a los denominados modelos combinados, los cuales integran simultáneamente diferentes decisiones (generación de viajes, elección de destino, modo y ruta), y cuyos orígenes datan de principios de la década de los 80 (Anas, 1983; Boyce et al, 1983; Boyce et al, 1988; Safwat y Magnanti, 1988; Brice, 1989; Oppenheim, 1995; Abrahamsson y Lundqvist, 1999; Boyce y Bar-Gera, 2003; De Cea et al, 2005; García y Marín, 2005; De Cea et al, 2008). En cada uno de los casos reportados en la literatura especializada que siguen este enfoque, el modelo de demanda es la solución de un problema de optimización equivalente, con un conjunto de restricciones lineales de no negatividad y de conservación de flujo. La solución de este problema es un modelo LMN o LJ dependiendo de la forma de la función objetivo y de las restricciones. Este problema es de maximización de la entropía, pero la mayor parte de los parámetros del modelo se consideran exógenos al problema.

Los parámetros exógenos, se determinan aplicando diversas técnicas estadísticas de calibración de modelos econométricos, siendo la más utilizada la de maximización de la verosimilitud. Cuando el modelo es LMN esta estrategia es coherente debido a que los estimadores MV y ME son idénticos, según se deduce del

trabajo de Anas (1983); sin embargo, esta estrategia podría ser inapropiada cuando se trata un modelo LJ.

En este trabajo se analiza la equivalencia o diferencias entre los estimadores ME y MV en el caso del modelo LJ, aspecto que no ha sido reportado previamente en la literatura especializada. Junto a una formulación y breve análisis teórico de ambos estimadores, estos se calculan en un gran número de bases de datos, mediante simulaciones de Monte Carlo con parámetros poblacionales conocidos. Se analiza el sesgo, eficiencia y el error cuadrático medio (ECM) de ambos estimadores.

2. FORMULACIÓN Y ESTIMACIÓN DE MODELOS LOGIT JERÁRQUICOS

2.1 Formulación del Modelo LJ

En el modelo LJ (Williams, 1977), la utilidad que provee la alternativa a del grupo g para el individuo q del tipo i es:

$$U_{agqi} = V_{agi} + V_{gi} + \varepsilon_{agqi} + \varepsilon_{gqi} \quad \forall a, g, q=1, \dots, N_i^0, i \quad (1)$$

donde N_i^0 representa el número de individuos de tipo i . Si $N_i^0 = 1 \forall i$ el modelo es desagregado, pudiéndose omitir por simplicidad el subíndice q ; si no, el modelo es agregado. Los términos V_{agi} y V_{gi} son las componentes determinísticas de la utilidad que el individuo de tipo i percibe de la alternativa a y del grupo g , respectivamente. Se asume que ambos términos son funciones lineales de atributos (x y w) y de parámetros (β y γ), genéricos o específicos por alternativa o grupo y tipo de individuo:

$$V_{agi} = \sum_k \beta_k x_{agik} \quad \forall a, g, i \quad (2)$$

$$V_{gi} = \sum_m \gamma_m w_{gim} \quad \forall g, i \quad (3)$$

Las variables ε_{agqi} son v.a.i.i.d. que se distribuyen Gumbel $(0, \mu_g)$, donde $\mu_g > 0$. Por su parte, las variables ε_{gqi} son tales que $\varepsilon'_{gqi} = \varepsilon^*_{gqi} + \varepsilon_{gqi} \forall g, q, i$, donde, ε^*_{gqi} son v.a.i.i.d. que se distribuyen Gumbel $(0, \lambda)$, donde $\lambda > 0$ y ε^*_{gqi} son v.a.i.i.d. que se distribuyen Gumbel $(0, \mu_g)$. El modelo LJ de la probabilidad de que el individuo del tipo i elija la alternativa a del grupo g ($p_{ag/i}$) está dado por:

$$p_{ag/i} = p_{a/gi} p_{g/i} \quad \forall a, g, i \quad (4)$$

donde $p_{a/gi}$ es la probabilidad de que el individuo del tipo i elija la alternativa a , dado que eligió el grupo g , y $p_{g/i}$ es la probabilidad de que el individuo del tipo i elija el grupo g . Si se denota

$$V_{gi}^* = \frac{1}{\mu_g} \ln \left(\sum_a \exp(\mu_g V_{agi}) \right) \quad \forall g, i, \quad \text{ambas}$$

probabilidades se definen como:

$$p_{a/gi} \equiv \frac{\exp(\mu_g V_{agi})}{\sum_{a'} \exp(\mu_g V_{a'gi})} \quad \forall a, g, i \quad (5)$$

$$p_{g/i} \equiv \frac{\exp(V_{gi}^* + V_{gi})}{\sum_{g'} \exp(V_{g'i}^* + V_{g'i})} \quad \forall g, i \quad (6)$$

En la especificación de (6) se ha supuesto que $\lambda = 1$ para así poder identificar los parámetros de este modelo. Con este supuesto, debe verificarse que $\mu_g \geq 1 \forall g$. Si $\mu_g = 1 \forall g$, entonces el modelo LJ se reduce al modelo LMN.

2.2 Estimación del modelo LJ por Maximización de la Verosimilitud

Los estimadores máximo-verosímiles (MV) de los parámetros del modelo Logit Jerárquico definido por (4)-(5)-(6) se obtienen como solución del siguiente problema de optimización:

$$\begin{aligned} \text{Máx}_{\mu, \beta, \gamma} \ln L = & \sum_{i,g} N_{gi}^0 \ln p_{g/i}(\mu, \beta, \gamma) + \\ & \sum_{i,g,a} N_{agi}^0 \ln p_{a/gi}(\mu, \beta) \end{aligned} \quad (7)$$

En la expresión (7) N_{agi}^0 es el número observado de individuos de tipo i que eligieron la alternativa a del grupo g , cumpliéndose que $N_i^0 = \sum_{g,a} N_{agi}^0 \forall i$ y $N_{gi}^0 = \sum_a N_{agi}^0 \forall g, i$

Las condiciones de primer orden de (7) son:

$$\frac{\partial \ln L}{\partial \beta_k} = \sum_{i,g} \frac{N_{gi}^0}{p_{g/i}} \frac{\partial p_{g/i}}{\partial \beta_k} + \sum_{i,g,a} \frac{N_{agi}^0}{p_{a/gi}} \frac{\partial p_{a/gi}}{\partial \beta_k} = 0 \quad \forall k \quad (8)$$

$$\frac{\partial \ln L}{\partial \mu_{g'}} = \sum_{i,g} \frac{N_{gi}^0}{p_{g/i}} \frac{\partial p_{g/i}}{\partial \mu_{g'}} + \sum_{i,a} \frac{N_{agi}^0}{p_{a/gi}} \frac{\partial p_{a/gi}}{\partial \mu_{g'}} = 0 \quad \forall g' \quad (9)$$

$$\frac{\partial \ln L}{\partial \gamma_m} = \sum_{i,g} \frac{N_{gi}^0}{p_{g/i}} \frac{\partial p_{g/i}}{\partial \gamma_m} = 0 \quad \forall m \quad (10)$$

Es simple demostrar que:

$$\frac{\partial p_{g/i}}{\partial \beta_k} = p_{g/i} \left(\sum_a p_{a/gi} x_{agik} - \sum_{g'} p_{g'/i} \sum_a p_{a/g'i} x_{ag'ik} \right) \quad \forall g, i, k \quad (11)$$

$$\frac{\partial p_{a/gi}}{\partial \beta_k} = p_{a/gi} \mu_g \left(x_{agik} - \sum_{a'} p_{a'/gi} x_{a'gik} \right) \quad \forall a, g, i, k \quad (12)$$

Luego, reemplazando (11) y (12) en (8) se deduce la siguiente ecuación asociada a β_k :

$$\sum_{i,g,a} \left(N_{gi}^0 p_{g/i} p_{a/gi} + (\mu_g - 1) (N_{gi}^0 p_{a/gi} - N_{agi}^0) \right) x_{agik} = \sum_{i,g,a} N_{agi}^0 x_{agik} \quad \forall k \quad (13)$$

Si $\mu_g = 1 \forall g$, el modelo Logit Jerárquico se transforma en un modelo Logit Multinomial (LMN), y la expresión (13) se reduce a la conocida expresión que establece que la suma de los valores del atributo x_k para las alternativas elegidas por los individuos es igual a la suma predicha por las probabilidades de elección estimadas. De esta condición se deduce que en un modelo LMN, los estimadores máximo-verosímiles permiten reproducir los valores promedio de las variables explicativas (por ejemplo, tiempos de viaje, costos, etc.), y si se especifican constantes específicas por alternativa, las particiones de mercado.

Sin embargo, de la expresión (13), apreciamos que esta condición no se verifica en el caso del modelo Logit Jerárquico, pues ello requiere el cumplimiento de la siguiente condición adicional:

$$\sum_{i,g,a} (\mu_g - 1) (N_{gi}^0 p_{a/gi} - N_{agi}^0) x_{agik} = 0 \quad \forall k \quad (14)$$

Por otra parte, si definimos una constante específica β_{0ag} para cada alternativa a al interior de cada grupo g , la ecuación (15) de este parámetro es:

$$\sum_i N_i^0 p_{g'/i} p_{a'/g'i} + (\mu_{g'} - 1) \sum_i (N_{g'i}^0 p_{a'/g'i} - N_{a'g'i}^0) = \sum_i N_{a'g'i}^0 \quad \forall a', g' \quad (15)$$

Por lo tanto, de (15) se concluye que en el Logit Jerárquico no se reproducen las particiones observadas de cada alternativa dentro de cada grupo, porque ello requeriría que $\mu_g = 1 \quad \forall g$ (modelo Logit Multinomial) o, alternativamente, que $\sum_i N_{g'i}^0 p_{a'/g'i} = \sum_i N_{a'g'i}^0 \quad \forall a', g'$, es decir, que se reprodujesen las particiones observadas de cada alternativa, condicional en la elección de cada grupo. Además, si se suma sobre a' en ambos lados de (15), se deduce:

$$\sum_i N_i^0 p_{g'/i} = \sum_i N_{g'i}^0 \quad \forall g' \quad (16)$$

Por lo tanto, se reproducen las particiones observadas de cada nido cuando se especifican constantes específicas por alternativa. Siguiendo una deducción similar se obtiene la misma conclusión si se especifican constantes específicas por nido.

2.2 Estimación del Modelo LJ por Maximización de la Entropía

Considérese el siguiente problema de optimización, denominado de Maximización de la Entropía con Probabilidades Jerárquicas (MEPJ):

$$\begin{aligned} \text{Máx} \quad & - \sum_{a,g,i} N_i^0 p_{ag/i} \ln p_{ag/i} = \\ & - \sum_{a,g,i} N_i^0 p_{g/i} p_{a/gi} \ln(p_{g/i} p_{a/gi}) \end{aligned} \quad (17)$$

sujeto a:

$$- \sum_{i,a} N_i^0 p_{g/i} p_{a/gi} \ln p_{a/gi} = \quad (18)$$

$$- \sum_{i,a} N_{agi}^0 \ln \left(\frac{N_{agi}^0}{N_{gi}^0} \right) \quad \forall g \quad \left(\frac{1}{\mu_g} \right) \quad (19)$$

$$\sum_{i,g,a} N_i^0 p_{g/i} p_{a/gi} x_{agik} = \sum_{i,g,a} N_{agi}^0 x_{agik} \quad \forall k \quad (\beta_k) \quad (19)$$

$$\sum_{i,g} N_i^0 p_{g/i} w_{gim} = \sum_{i,g} N_{gi}^0 w_{gim} \quad \forall m \quad (\gamma_m) \quad (20)$$

$$\sum_a p_{a/gi} = 1 \quad \forall g, i \quad (\alpha_{gi}) \quad (21)$$

$$\sum_g p_{g/i} = 1 \quad \forall i \quad (\rho_i) \quad (22)$$

En esta formulación $\left(\frac{1}{\mu_g} \right), (\beta_k), (\gamma_m), (\alpha_{gi})$ y (ρ_i) representan a

los multiplicadores de Lagrange de las restricciones (18) a (22) respectivamente. El problema MEPJ tiene una estructura similar al problema de maximización de la entropía que genera un modelo Logit Multinomial (Anas, 1983), salvo por (18), que restringe la Entropía dentro de cada grupo considerado.

Es útil notar que la Entropía total puede descomponerse de la siguiente forma:

$$\begin{aligned} & - \sum_{a,g,i} N_i^0 p_{ag/i} \ln p_{ag/i} = \\ & - \sum_{i,g} N_i^0 p_{g/i} \ln p_{g/i} - \sum_{a,g,i} N_i^0 p_{g/i} p_{a/gi} \ln p_{a/gi} \end{aligned} \quad (23)$$

Si se verifica la condición (18) se obtiene:

$$\begin{aligned} & - \sum_{a,g,i} N_i^0 p_{ag/i} \ln p_{ag/i} = \\ & - \sum_{i,g} N_i^0 p_{g/i} \ln p_{g/i} - \sum_{a,i,g} N_{agi}^0 \ln \left(\frac{N_{agi}^0}{N_{gi}^0} \right) \end{aligned} \quad (24)$$

Luego, la función objetivo (17) puede ser reemplazada por $-\sum_{i,g} N_i^0 p_{g/i} \ln p_{g/i}$, dado que el segundo término del lado derecho de (24) es una constante. La función Lagrangeana del problema MEPJ (con la nueva función objetivo) es:

$$\begin{aligned} L = & - \sum_{i,g} N_i^0 p_{g/i} \ln p_{g/i} + \\ & \sum_g \frac{1}{\mu_g} \left(- \sum_{i,a} N_i^0 p_{g/i} p_{a/gi} \ln p_{a/gi} + \sum_{i,a} N_{agi}^0 \ln \left(\frac{N_{agi}^0}{N_{gi}^0} \right) \right) + \\ & + \sum_k \beta_k \left(\sum_{i,g,a} N_i^0 p_{g/i} p_{a/gi} x_{agik} - \sum_{i,g,a} N_{agi}^0 x_{agik} \right) + \\ & + \sum_m \gamma_m \left(\sum_{i,g} N_i^0 p_{g/i} w_{gim} - \sum_{i,g} N_{gi}^0 w_{gim} \right) + \\ & + \sum_{i,g} \alpha_{gi} \left(\sum_a p_{a/gi} - 1 \right) + \sum_i \rho_i \left(\sum_g p_{g/i} - 1 \right) \end{aligned} \quad (25)$$

A partir de las condiciones de optimalidad de (25) se deduce que la solución $p_{a/gi}$ y $p_{g/i}$ cumple con las ecuaciones (5), (6), (2) y (3) y por lo tanto, corresponde a un modelo LJ.

Los estimadores ME de este modelo son los multiplicadores de Lagrange $\left(\frac{1}{\mu_g} \right), (\beta_k), (\gamma_m)$, los cuales se obtienen al reemplazar (5) y (6) en las restricciones (18), (19) y (20), y resolver numéricamente estas ecuaciones, por ejemplo mediante el método de Newton en nuestro caso. Como los estimadores ME satisfacen la condición (19), a diferencia de los estimadores MV, permiten reproducir los valores promedio de las variables explicativas y si se especifican constantes específicas por alternativa, las particiones de mercado. Esto permite concluir que los estimadores MV son diferentes a los estimadores ME.

3. ANÁLISIS DE LOS ESTIMADORES MV Y ME MEDIANTE SIMULACIÓN

Para comparar ambos enfoques de estimación de parámetros (Máxima Verosimilitud y Máxima Entropía), realizamos una serie de simulaciones de Monte Carlo con un modelo jerárquico de distribución y partición modal conjunta, considerando diferentes tamaños muestrales.

El árbol simulado se muestra en la Figura 1. La simulación consideró un total de 30 zonas origen y destino, y un total de 4 modos de transporte (Auto, Bus, Taxi y Metro), los cuales están disponibles para cualquier destino.

Por simplicidad, también supusimos que los parámetros μ son igual para todos los destinos y no se consideraron atributos especiales por cada nido.

Además, para identificar los parámetros del modelo, se asignó un valor igual a 1 al parámetro λ de la Figura 1.

Las funciones de utilidad de los cuatro modos de transporte incluyeron las variables tiempo de viaje y costo de viaje, cuyos parámetros (exceptuando las constantes modales) son genéricos, es decir, iguales para los cuatro modos. En síntesis, las utilidades de cada modo fueron las siguientes:

$$V_{Auto} = \beta_{Auto}^0 + \beta^{Tiempo} T_{Auto} + \beta^{Costo} C_{Auto} \tag{26}$$

$$V_{Bus} = \beta_{Bus}^0 + \beta^{Tiempo} T_{Bus} + \beta^{Costo} C_{Bus} \tag{27}$$

$$V_{Taxi} = \beta_{Taxi}^0 + \beta^{Tiempo} T_{Taxi} + \beta^{Costo} C_{Taxi} \tag{28}$$

$$V_{Metro} = \beta_{Metro}^0 + \beta^{Tiempo} T_{Metro} + \beta^{Costo} C_{Metro} \tag{29}$$

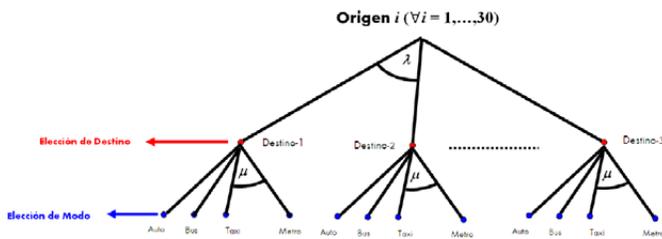


FIGURA 1: Estructura del Árbol Jerárquico usado en la Simulación

Los valores de los parámetros considerados para las simulaciones se reportan en la Tabla 1, donde se aprecia que el valor subjetivo del tiempo (VST) es $(-0,25/-0,006) = 41,47$.

TABLA 1: Valores de los Parámetros usados en las Simulaciones de Monte Carlo

Parámetro	Valor
β_{Auto}^0	0,9
β_{Auto}^0	0
β_{Taxi}^0	0,5
β_{Metro}^0	0,4
β^{Tiempo}	-0,25
β^{Costo}	-0,006
λ	1
$1/\mu$	0,5

Fuente: Elaboración propia. Los valores son similares a los reportados en De Cea et al (2008)

Respecto a las variables explicativas (tiempos y costos), los valores fueron extraídos de la EOD-2001 del Gran Santiago. En la Tabla 2 presentamos la media y la desviación estándar de estas variables para cada alternativa de transporte considerada:

TABLA 2: Valores de los Parámetros usados en las Simulaciones de Monte Carlo

Modo	Variable	Media (*)	Desv. Estándar (*)
Auto	Tiempo de Viaje	16	11
	Costo	2.031	138
Taxi	Tiempo de Viaje	17	11
	Costo	2.279	148
Bus	Tiempo de Viaje	54	12
	Costo	409	25
Metro	Tiempo de Viaje	45	7
	Costo	833	73

Fuente: EOD 2001 del Gran Santiago
 (*): Tiempo de viaje en minutos y costo en pesos del 2009.

Consideramos tres tamaños muestrales: 1.000, 5.000 y 10.000 observaciones. Para cada tamaño muestral, realizamos 1.000 simulaciones de Monte-Carlo, lo que nos permitió por lo tanto obtener 1.000 estimadores Máximo Verosímiles (MV) y Máximo Entrópicos (ME) de los parámetros para cada uno de los tres tamaños muestrales.

Los problemas de maximización de la verosimilitud y de la entropía fueron resueltos numéricamente aplicando el método de Newton. El criterio de convergencia utilizado en todas las simulaciones fue de un 0,1%, es decir, que la diferencia porcentual entre las estimaciones de cada parámetro obtenidas de dos iteraciones sucesivas del método, no superase el valor de 0,001.

A partir de estos 1.000 estimadores de MV y ME, construimos histogramas que permiten analizar simultáneamente su sesgo, varianza y error cuadrático medio (ECM). Los resultados de la estimación del parámetro μ se muestran en la Figura 2, y los resultados de la estimación del VST se presentan en la Figura 3. En ambas figuras, la línea punteada indica el valor del parámetro poblacional (usado en la simulación), mientras que la línea negra muestra el histograma de estimaciones MV y la línea azul el histograma de estimaciones ME.

A partir de la Figura 2 observamos claramente que la estimación por MV de μ es sesgada pero consistente, mientras que la estimación por ME presenta, para las muestras analizadas, un sesgo significativamente inferior, aunque es menos eficiente (mayor varianza) que MV. Analizando la Figura 3, observamos que tanto el estimador MV como el ME del Valor Subjetivo del Tiempo (VST) son insesgados, aunque claramente el estimador ME es más eficiente. Estos resultados se confirman en la Tabla 3, donde se exponen el sesgo, varianza y error cuadrático medio de los estimadores MV y ME, para los tres tamaños muestrales analizados (1.000, 5.000 y 10.000).

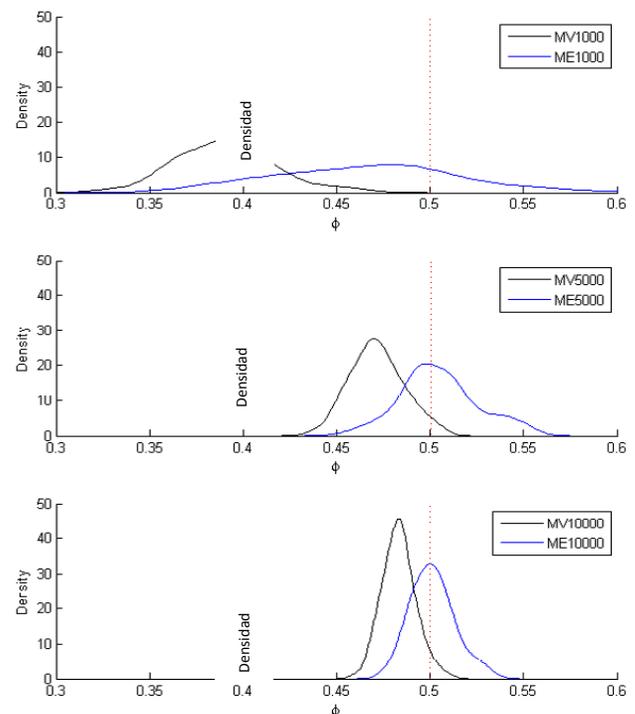


FIGURA 2: Resultados de la Estimación del Parámetro $\phi = 1/\mu$ mediante MV y ME

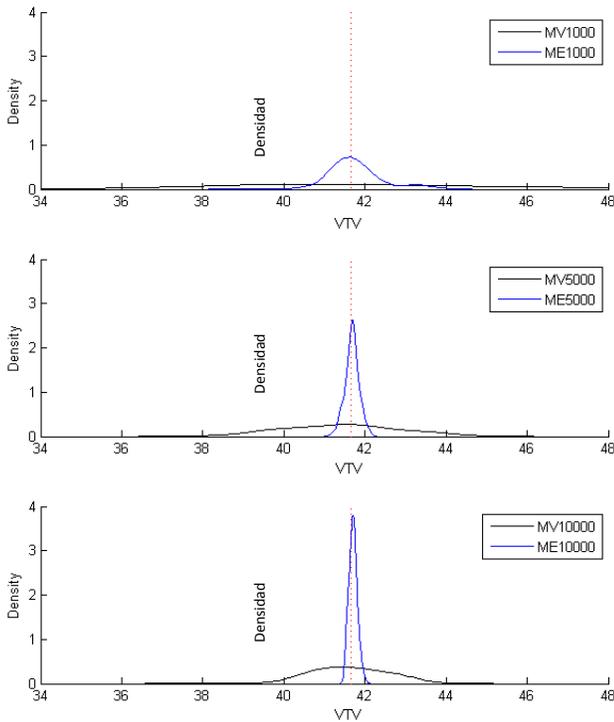


FIGURA 3: Resultados de la Estimación del VST = $\beta^{Tiempo} / \beta^{Costo}$ mediante MV y ME

TABLA 3: Síntesis de Resultados de los Estimadores MV y ME

Método	Parámetro	Muestra	Sesgo	Varianza	ECM (*)
MV	$1/\mu$	1.000	0.01199	0.01280	0.02479
	$1/\mu$	5.000	0.00081	0.00102	0.00183
	$1/\mu$	10.000	0.00028	0.00036	0.00064
	VST	1.000	0.33730	13.35662	13.69392
	VST	5.000	0.00201	2.69824	2.70025
	VST	10.000	0.00189	1.14187	1.14376
ME	$1/\mu$	1.000	0.00118	0.00351	0.00468
	$1/\mu$	5.000	0.00002	0.00045	0.00047
	$1/\mu$	10.000	0.00000	0.00013	0.00013
	VST	1.000	0.09107	1.54540	1.63647
	VST	5.000	0.01681	0.15821	0.17502
	VST	10.000	0.00570	0.05632	0.06202

(*): el ECM se define como la varianza más el cuadrado del sesgo.

A partir de los resultados de la Tabla 3 y de las Figuras 4 y 5, es posible concluir que el ECM del estimador MV es sistemáticamente mayor al ECM del estimador ME, particularmente en muestras de menor tamaño.

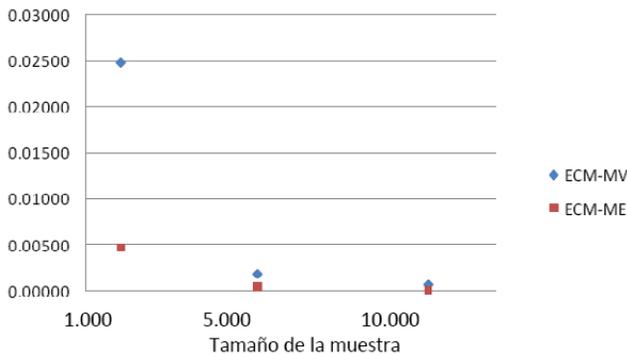


FIGURA 4: Error Cuadrático Medio (ECM) de Estimaciones MV y ME del Parámetro $\phi = 1/\mu$

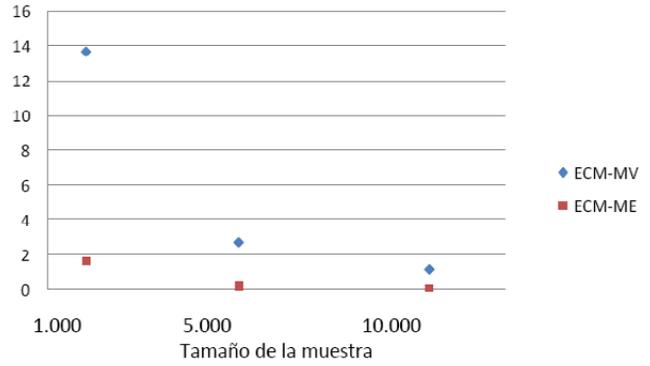


FIGURA 5: Error Cuadrático Medio (ECM) de Estimaciones MV y ME del VST

4. SÍNTESIS Y CONCLUSIONES

El método tradicional de estimación de los parámetros de los modelos Logit Multinomial y Logit Jerárquico es el de maximización de la Verosimilitud. En el contexto de modelos agregados de demanda se han formulado problemas de maximización de la Entropía cuyas soluciones son estos modelos Logit, con parámetros que pueden interpretarse como estimadores de los parámetros poblacionales, motivo por el cual en este trabajo se denominan estimadores máximo-entrópicos. Del trabajo de Anas de 1983 se deduce que estos estimadores coinciden con los estimadores máximo-verosímiles cuando se trata del modelo Logit Multinomial.

En este trabajo se analiza esta equivalencia en el caso modelos de demanda agregados del tipo Logit Jerárquicos.

En primer término, se formuló un problema general de maximización de la entropía de estructura jerárquica y se dedujo que su solución es un modelo Logit Jerárquico. A continuación, se demostró que los estimadores máximo-entrópicos son diferentes a los máximo-verosímiles, particularmente porque estos últimos no reproducen los valores promedio de las variables explicativas ni las proporciones de mercado observadas (muestra de calibración), mientras que los estimadores máximo-entrópicos sí lo hacen por construcción.

A continuación, se analizaron empíricamente ambos estimadores. Se efectuaron simulaciones de Monte Carlo de este modelo jerárquico con parámetros poblacionales dados para distintos tamaños de la base de datos de elecciones y se calcularon ambos estimadores. Los resultados obtenidos muestran que el estimador máximo-entrópico es superior al estimador máximo-verosímil en cuanto a sesgo y eficiencia, particularmente en el caso de muestras de menor tamaño.

Si bien es necesario analizar otras estructuras jerárquicas a las consideradas en esta investigación aplicada, es posible concluir que el enfoque de máxima entropía es una alternativa efectiva para estimar los parámetros de modelos agregados de demanda del tipo Logit Jerárquico, y especialmente en muestras pequeñas como normalmente ocurre en el estado de la práctica.

REFERENCIAS

- Abrahamsson, T. and L. Lundqvist (1999): Formulation and Estimation of Combined Network Equilibrium Models with Applications to Stockholm. **Transportation Science**, 33, 80-100.
- Anas, A. (1983). Discrete Choice Theory, Information Theory and the Multinomial Logit and Gravity Models. **Transportation Research**, 17B, 13-23.
- Boyce, D., Le Blanc L., Chon K., Lee Y. and Lin, K. (1983). Implementation and computational issues for combined models of location, destination, mode and route choice. **Environment and Planning**, 15A, 1219-1230.
- Boyce, D. E., Leblanc, L.J. and Chon, K.S. (1988). Network equilibrium models of urban location and travel choices: a retrospective survey, **Journal of Regional Science**, 28, 159-183.
- Boyce, D. and H. Bar-Gera (2003). Validation of Multiclass Urban Travel Forecasting Models Combining Origin-Destination, Mode, and Route Choices, **Journal of Regional Science**, 43, 517-540.
- Brice, S. (1989). Derivation of nested transport models within a mathematical programming framework. **Transportation Research**, 23B, 19 – 28.
- De Cea, J. Fernandez, J.E., Soto, A. and Dekock, V. (2005). Solving network equilibrium on multimodal urban transportation networks with multiple user classes. **Transport Reviews**, 25, 293-317.
- De Cea, J., J.E. Fernandez and L. De Grange (2008) Combined models with hierarchical demand choices: a multi-objective entropy optimization approach. **Transport Reviews**, 28, 415-438.
- De Grange, L., Ibeas, A. and Gonzalez, F. (2009). A Hierarchical Gravity Model with Spatial Correlation: Mathematical Formulation and Parameter Estimation. **Networks and Spatial Economics**. (DOI: 10.1007/s11067-008-9097-0).
- Fang, S. C. and S. J. Tsao (1995). Linearly-constrained entropy maximization problem with quadratic cost and its applications to transportation planning problems. **Transportation Science**, 29, 353-365.
- Fotheringham, A.S. (1986). Modeling hierarchical destination choice. **Environment and Planning**, 18A, 401-418.
- García, R. and Marín, A. (2005). Network equilibrium with combined modes: models and solution algorithms. **Transportation Research**, 39B, 223-254.
- Oppenheim, N. (1995). Urban Travel Demand Modeling. John Wiley & Sons, New York.
- Safwat, K. and Magnanti, T. (1988). A combined trip generation, trip distribution, modal split and traffic assignment model. **Transportation Science**, 22, 14-30.
- Williams, H.C.W.L. (1977) On the formation of travel demand models and economic evaluation measures of user benefit. **Environment and Planning**, 9A, 285-344.
- Wilson, A. G. (1970). **Entropy in Urban and Regional Modeling**. Pion, London.