

# PERFIL DE AUDIENCIA DE UN PROGRAMA DE TELEVISIÓN: ANÁLISIS A TRAVÉS DE MODELOS DE REGRESIÓN TOBIT

---

GILBERT SWINNEN, JUAN CARLOS GÁZQUEZ ABAD, DAVID JIMÉNEZ CASTILLO, ELVIRA SÁEZ GONZÁLEZ, GEMA MARÍA MARÍN CARRILLO

[gilbert.swinnen@uhasselt.be](mailto:gilbert.swinnen@uhasselt.be), [jcgazque@ual.es](mailto:jcgazque@ual.es), [djcasti@ual.es](mailto:djcasti@ual.es), [esaez@ual.es](mailto:esaez@ual.es), [gmarin@ual.es](mailto:gmarin@ual.es)

Universidad de Hasselt, Universidad de Almería

## RESUMEN

*La descripción del perfil de los telespectadores de un programa de televisión puede proporcionar información muy útil, tanto para los responsables de las cadenas como para los anunciantes. En la literatura de marketing no es habitual la utilización de modelos estadísticos muy sofisticados para determinar dicho perfil. Este trabajo en curso trata de analizar la utilidad que tienen los modelos de regresión Tobit censurados como herramienta que permite determinar el perfil de la audiencia de un programa de televisión. Para ello, se utilizarán los datos de audiencia de un debate político de un panel formado por 1706 individuos (750 familias) combinado con información socio-demográfica sobre dichos individuos recogida a través de cuestionario.*

## PALABRAS CLAVE

*Televisión, Comportamiento de visionado, Perfil de audiencia, Modelos de regresión Tobit*

## 1. Introducción

La televisión es un elemento primordial para la mayor parte de las personas, ya que se trata de un recurso primario de información, entretenimiento, notoriedad política y conocimiento cultural (Stewart, 1999). La información sobre la preferencia que muestra la audiencia por los distintos programas existentes es muy importante tanto para las cadenas como para los anunciantes. Esto ha motivado que un elevado número de trabajos se hayan centrado en analizar el perfil dinámico de audiencia utilizando modelos de elección de programas. Existe consenso en la literatura al señalar que la preferencia de los individuos por ver un determinado programa es explicada parcialmente por la lealtad hacia cada canal y hacia cada tipo de programa (Cohen, 2002; Schachar, 2000; Zubayr, 1999). Este trabajo en curso no pretende analizar el comportamiento dinámico de elección de programa de los espectadores, sino que se centra en un determinado programa de televisión (un debate político) y trata de hallar la relación que existe entre la preferencia hacia el programa y las características del televidente. El modelo propuesto puede ayudar tanto a los ejecutivos a la hora de maximizar sus *ratings* (adaptando las características del programa a dicho perfil de audiencia), como a los anunciantes al predecir y conocer la composición de dicha audiencia.

## 2. El modelo de regresión Tobit

Uno de los principales problemas asociado con la utilización de paneles de audiencia es la existencia de un gran número de “no espectadores”, lo que se refleja en la existencia de una gran cantidad de ceros en la variable dependiente. En este caso, los métodos de regresión convencionales no son adecuados (Greene, 1997), ya que el método de estimación debe considerar que el modelo se encuentra restringido a este valor cero. Es por ello que utilizamos un modelo de regresión Tobit (censurado) que nos va a permitir utilizar tanto datos de espectadores como de no espectadores (tiempo de visionado = 0) en el mismo modelo.

Sea,  $Y_i$  el número observado de minutos vistos por cada miembro  $i$  del panel. Entonces, el modelo limitado con variable dependiente (censurada) viene dado por:

$$Y_i = \begin{cases} X_i B + u_i & \text{si } X_i B + u_i > 0 \\ 0 & \text{en otro caso,} \end{cases} \quad [1]$$

Donde  $X_i$  es el vector de variables explicativas,  $B$  es el vector de coeficientes que han de estimarse, mientras que el término de error es independiente e idénticamente distribuido con una normal  $(0, \sigma^2)$ .

El modelo supone la existencia de una variable latente  $Y_i^* = X_i B + u_i$ , existente, únicamente, cuando es positiva, y con media  $E(Y_i^*) = X_i B$ . Si sustituimos por tanto:

$$Y_i = \begin{cases} Y_i^* & \text{si } Y_i^* > 0 \\ 0 & \text{en otro caso} \end{cases} \quad [2]$$

La función de probabilidad tiene dos partes. La primera de ellas representa la probabilidad de ocurrencia para observaciones con valor cero (no espectadores), mientras que la segunda representa la probabilidad de ocurrencia para aquellas observaciones con un valor 1 o superior (aquellos individuos del panel que han estado viendo el programa al menos 1 minuto).

La probabilidad de ocurrencia de las observaciones que toman valor cero, es decir, aquellos individuos que no han estado viendo el programa, viene dada por:

$$\Pr(Y_i = 0) = \Pr(u_i < X_i B) = (1 - F_i). \quad [3]$$

Para los individuos que sí han estado viendo el programa, esta probabilidad viene dada por la siguiente expresión:

$$\Pr(Y_i > 0) * f(Y_i | Y_i > 0) = [\sigma^2 (2\pi)]^{-1/2} e^{-(1/2 \sigma^2)(Y_i - X_i B)} \quad [4]$$

$$L = \prod_0 (1 - F_i) \prod_1 [2\pi\sigma^2]^{-1/2} e^{-(1/2 \sigma^2)(Y_i - X_i B)} \quad [5]$$

La función de probabilidad viene definida por:

mientras que la función de verosimilitud toma la siguiente forma:

$$\ln L = \sum_0 \ln(1 - F_i) + \sum_1 \ln[(2\pi\sigma^2)^{-1/2}] - \sum_1 (2\sigma)^{-2} (Y_i - X_i B)^2. \quad [6]$$

El procedimiento de estimación por máxima verosimilitud proporciona estimaciones consistentes y óptimas de los parámetros del modelo *Tobit*. Para la estimación de dichos parámetros utilizaremos el software LIMDEP.

### 3. Metodología de la investigación

#### 3.1. Tipo de datos

Este trabajo está basado en un panel de audiencia en Bélgica, proporcionado por NV Audimetrie, que recoge el comportamiento de visionado por individuo, así como las características demográficas de cada uno de ellos. Este panel utiliza individuos de más de 3 años. El audímetro permite conocer qué programa está viendo cada individuo, así como el tiempo de visionado. La principal ventaja es que la información no está basada en la “capacidad de recuerdo” que el individuo pueda tener de los programas que ha estado viendo, así como del tiempo dedicado a cada uno de ellos.

Además, se dispone de datos socio-demográficos de los individuos del panel (e.g., edad, sexo, profesión, tamaño familiar, nivel de educación, orientación política, etc.).

El tamaño total del panel de audiencia en Bélgica es de 1500 familias, repartidas equitativamente entre la región Flamenca del país (de habla neerlandesa) y la región Valona (de habla francesa). Sin embargo, y dado que este trabajo va a estar basado en un programa de debate político emitido en la zona flamenca, únicamente vamos a utilizar los datos del panel correspondientes a dicha parte del país. En total, se trata de 750 familias flamencas que suponen un total de 1706 individuos<sup>1</sup>.

Los datos proceden de la audiencia total de 4 shows de debate político emitidos en 4 semanas consecutivas durante el mes de noviembre de 1996. El show era emitido los domingos de 11 horas a 12:30 horas de la mañana, y se trata de un debate en el que se analizaban las noticias que se habían producido durante esa semana. En el debate participaban personajes del ámbito social y político implicado en dichas noticias. Los responsables de la cadena de TV tenían la impresión de que el programa tenía una imagen algo “anticuada” y que era principalmente visto por aquellas personas que tenían el tiempo y la formación suficiente como para comprender lo que allí se decía. Sin embargo, su intención era ampliar la audiencia tanto en términos de edad, como de capacidad intelectual y clase social de los individuos.

---

<sup>1</sup> En la recogida de datos se contó, igualmente, con la colaboración del presentador del programa y el departamento de investigación de la VRT (Radio y Televisión de Flandes)

### 3.2. Variables

La base de datos contiene información del tiempo total (minutos) que cada individuo ha visto la televisión. El tiempo de visionado fue contabilizado para cuatro programas consecutivos y varía entre 0 y 360 minutos. El tiempo total de visionado será utilizado como la variable dependiente en el modelo desarrollado. Con el objetivo de explicar el comportamiento de visionado, vamos a emplear 9 variables explicativas, algunas de ellas métricas y otras no métricas que hemos transformado en variables *dummy* (tabla 1).

TABLA 1  
Variables explicativas: significado y medición

Variable	Medición
Tamaño familiar	De 1 a 8 y $\geq 9$
Edad	Número de años
Sexo	1= hombre, 2= mujer
Interés por la política	Puntuación entre 0 y 25 (basado en encuesta)
Interés por la cultura	Puntuación entre 0 y 25 (basado en encuesta)
Interés por el deporte	Puntuación entre 0 y 25 (basado en encuesta)
Nivel de educación	5 variables <i>dummy</i> para 6 categorías: educación primaria, secundaria baja, secundaria alta, técnica baja, técnica alta, licenciatura + superior
Profesión	8 variables <i>dummy</i> para 9 categorías: otros, estudiante, empleado "white collar", ejecutivos y profesionales, autónomo, empleado "blue collar", parado, ama de casa, jubilado
Hábitat	3 variables <i>dummy</i> para 4 categorías: rural, rural centro, zona regional urbanizada, ciudad + aglomeración
Clase social	3 variables <i>dummy</i> para 4 categorías: baja, baja media, media alta, alta

### 4. Resultados previos

En la fase inicial del trabajo en la que nos encontramos, hemos realizado una primera estimación de los parámetros del modelo. Uno de los problemas con los que nos encontramos viene derivado de la existencia de un elevado número de variables *dummy*, utilizadas para transformar las variables no métricas. Esto puede ocasionar importantes problemas de multicolinealidad en el modelo, por lo que para evitar dichos problemas, hemos introducido cada una de estas variables de forma secuencial (*stepwise way*) y hemos utilizado sólo aquellas que tienen un impacto significativo. De este modo, sólo tres de las variables *dummy* calculadas han sido utilizadas. La tabla 2 muestra los resultados iniciales de la estimación del modelo

TABLA 2  
Estimación del modelo Tobit

Variable	B (coeficiente)	Error estándar	t
Constante	-124,4589	31,1579	-3,994*
Tamaño familiar	-17,3583	4,5039	-3,854*
Sexo	-39,9617	9,4116	-4,246*
Edad	1,9936	0,3989	4,998*
Interés por la cultura	0,0240	0,3922	0,061
Interés por los deportes	-0,1638	0,3959	-0,414
Interés por la política	1,0066	0,4254	2,336**
Clase media-alta	36,6347	10,9529	3,345*
Jubilado	44,9427	14,9785	3,000*
Licenciado (o superior)	43,9461	11,9679	3,672*

\*p<.01 \*\*p<.05

La interpretación de los resultados de la estimación nos proporciona las primeras conclusiones:

- La preferencia por ver el programa se ve afectada de forma positiva por la edad del individuo, de tal modo que por cada año más que tenga un individuo, se espera que vea 2 minutos más el programa.
- La preferencia por ver el programa se ve afectada de forma positiva por el interés político (un incremento de un punto en la escala de interés por la política supone un incremento esperado de 1 minuto del tiempo que el individuo dedica a ver el programa).

- Se espera que los hombres vean el programa durante 40 minutos más que las mujeres.
- Las personas de clase social media alta ven el programa 37 minutos más que los individuos del resto de clases sociales.
- Los individuos jubilados pasan un mayor tiempo (45 minutos aproximadamente) viendo el programa que los individuos del resto de profesiones.
- Finalmente, el interés por la cultura y los deportes no tiene un impacto significativo en el tiempo que los individuos pasan viendo el programa.

Algunos de los resultados pueden ser explicados por el hecho de que el programa sea emitido el domingo por la mañana, franja horaria en la que muchas personas (sobre todo aquellas familias más numerosas) suelen estar realizando otras actividades. En general, los resultados permiten conocer aquellos aspectos sobre los que los responsables de la programación han de poner más atención. Así, y con el objetivo de atraer una mayor audiencia, los resultados sugieren desarrollar un programa más atractivo para las mujeres (p.e., incluyendo una presentadora), la gente joven, y haciéndolo más comprensible para el público en general.

Desde el punto de vista de los anunciantes, los resultados que proporciona nuestro modelo pueden ser utilizados para conocer la conveniencia o no de insertar publicidad en determinados programas en función del tipo de audiencia más interesada en los mismos (si bien este no es el objetivo de este trabajo, ya que en este caso particular no se permite la inclusión de publicidad comercial en las cadenas de emisión nacional).

## 5. Conclusiones

Este trabajo en curso pretende analizar la utilidad de los modelos de regresión *Tobit* (censurados) en el análisis de las preferencias de visionado de los telespectadores. El modelo tiene la propiedad fundamental de que permite incluir aquellas observaciones que para la variable dependiente tengan un valor cero. Además, el modelo permite analizar el impacto significativo que tienen determinadas características de los individuos del panel, además de interpretar estos resultados de forma muy intuitiva. La utilidad de los resultados puede ser muy elevada, tanto para los responsables de las cadenas (al poder determinar los cambios y ajustes que deben realizar en el formato de sus programas), como para los anunciantes (al poder optimizar su presupuesto publicitario). Nuestro objetivo es continuar con el análisis del panel de audiencia, estimando los efectos marginales de cada variable, así como desarrollando modelos *Tobit* de coeficientes paramétricos y semiparamétricos que permitan conocer, por una parte, las diferencias existentes entre cada individuo y, por otra, agrupar la audiencia en segmentos a partir de las características socio-demográficas de las que disponemos.

## Bibliografía

- CLANCEY, M. (1994). "The Television Audience Examined," Winner of ARF's 1994 Richard L. Lysaker Price For Outstanding Research on Media, pgs. 77-87.
- COHEN, J. (2002). "Television Viewing Preferences: Programs, Schedules, and the Structure of Viewing Choices Made by Israeli Adults", *Journal of Broadcasting & Electronic Media*, Vol. 46, n° 2, pgs. 204-221.
- DANAHER, P. Y MAWHINNEY, D. (2001). "Optimizing Television Program Schedules Using Choice Modeling. *Journal of Marketing Research*, Vol. XXXVIII, n° August, pgs. 298-312.
- GREENE, W. (1997). *Econometric Analysis*, Prentice Hall.
- HOOGHE, M. (2002). "Watching Television and Civic Engagement. Disentangling the Effects of Time, Programs, and Stations," *Harvard International Journal of Press/Politics*, Vol. 7, n° 2, pgs. 84-104.
- KIM, CH. (2002). "Identifying Viewer Segments for Television Programs", *Journal of Advertising Research*, Vol. 42 n° February, pgs. 51-66.
- RAO, V. (1975). "Taxonomy of Television Programs Based on Viewing Behavior", *Journal of Marketing Research*, Vol. 12, n° August, pgs. 355-358.
- RUST, R. Y MARK, A. (1984). "An Audience Flow Model of Television Viewing Choice", *Marketing Science*, Vol. 3, n° 2, pgs. 113-124.
- SHACHAR, R. Y EMERSON, J. (2000). "Cast Demographics, Unobserved Segments, and Heterogeneous Switching Costs in a Television Viewing Choice Model", *Journal of Marketing Research*, Vol. XXXVII, n° May, pgs. 173-186.
- STEWART, J. (1999). "Interactive television at home: Television meets the Internet," en: *Cathy Toscan Jens Jensen* (eds), Aalborg University Press.