

Elección de un modelo explicativo del TIEPI a través de las técnica de decisión multicriterio Promethee

Ester Gutiérrez Moya¹

¹ Lda. Ciencias y Técnicas Estadísticas. Lda. Admón y Dirección de Empresas.
Dpto. Organización Industrial y Gestión de Empresas.
Escuela Superior de Ingenieros, egm@esi.us.es

RESUMEN

En este trabajo se lleva a cabo la resolución del problema de elegir un modelo que explique el TIEPI en función ciertas variables exógenas, a través del método multicriterio Promethee. Una vez analizados cinco modelos, con formas funcionales distintas, representativos del TIEPI, se plantea la disyuntiva de la elección de uno de ellos, la cual se abordará por medio de un conjunto de estadísticos y con la ayuda de la técnica de decisión Promethee, el cual es un procedimiento que permite sintetizar la información disponible y obtener una ordenación sobre el conjunto estudiado.

Palabras clave: Criterios de Información, TIEPI, Promethee.

1. Introducción.

Este trabajo realiza una análisis de determinados estadísticos que ayudan a suministrar información acerca de distintos modelos, lineales en los parámetros, que explican la continuidad del servicio eléctrico en España.

Se han planteado cinco modelos que, a nuestro juicio, explican la variabilidad de la continuidad del servicio eléctrico en España, a través de los datos de los principales cuatro grupos eléctricos que operan en el mercado español de la electricidad.

La elección de un modelo de regresión o serie cronológica es una tarea compleja, ya que es preciso distinguir algunos aspectos tales como la existencia de la multicolinealidad, heterocedasticidad, autocorrelación, no normalidad. Una vez superadas estas pruebas nos planteamos cinco modelos y analizamos el grado de la validez de los mismos.

El análisis se abordará a través de una serie de criterios de información. En la literatura estadística se han definido numerosos índices o criterios que ayudan a la elección de un modelo, esta consideración confiere al problema planteado el alcanzar una dimensión multicriterio, por lo que parece apropiado utilizar un método multicriterio, en este caso se utiliza el método Promethee.

2. Criterios de Información.

Una pregunta quizás tan importante como la modelización de un problema es conocer qué variables son importantes. En los modelos de regresión la inclusión de un número elevado de variables en el modelo siempre contribuye a mejorar el grado de explicación alcanzado por éste, por lo cual se hace necesario utilizar cambios que valoren la inclusión de estas nuevas variables tanto en la mejora que se produce en su coeficiente de determinación como en la reducción que sufren los grados de libertad del modelo estimado. Con esta finalidad se plantean distintos estadísticos que permiten seleccionar las variables que deben ser explicativas en un modelo econométrico. Para este análisis se han utilizado los siguientes:

Coeficiente de Determinación Ajustado, \bar{R}^2 .

$$\bar{R}^2 = 1 - \frac{SCE/(N-k)}{SCT/(N-1)} = 1 - \left(R^2 \frac{N-1}{N-k} \right) \quad (1)$$

Suma de Cuadrados de los Residuos, SCR.

$$SCR = \sum_{i=1}^N e_i^2 \quad (2)$$

Criterio de Información de Akaike, IA.

$$IA = \frac{2k}{N} + \log\left(\frac{SCR}{N}\right) \quad (3)$$

Criterio de Predicción de Amemiya, AM.

$$AM = \frac{SCR}{N-k}(N+k) \quad (4)$$

Criterio de Schwarz, SCH.

$$SCH = \frac{k}{N} \log N + \log\left(\frac{SCR}{N}\right) \quad (5)$$

donde N es el número de observaciones y k el número de variables exógenas del modelo.

3. Análisis de los criterios de información.

Para seleccionar alguno/s de los modelos como modelo explicativo de las variaciones del TIEPI, establecemos algunos criterios para su elección, tal como aparece reflejado en la tabla 1.

	Lineal	Log lineal	Log lin	Lin log	Recíproco
R ² corregido	0.783196	0.797462	0.759986	0.768853	0.425184
SCR	1.012162	0.117157	0.138835	0.996114	2.889990
Akaike	0.452377	-1.453954	-1.284183	0.561395	1.376541
Amemiya	1.4793137	0.2236633	0.2650486	1.2814217	3.7157014
Schwarz	0.597238	-1.212520	-1.042749	0.754542	1.473115

Tabla 1: Criterios de Información.

Como podemos observar en la tabla, la mayoría de los modelos analizados alcanzan en sus índices de información valores similares, salvo el modelo inverso o recíproco. El modelo que alcanza mejores valores en todos los índices de información considerados es el modelo log lineal, modelo de regresión 2. Este modelo log lineal depende de cuatro variables explicativas, la potencia instalada, la longitud de las líneas de baja tensión, el número de empleados y número de clientes (todas ellas logaritmizadas). Si atendemos al *principio de parsimonia* enunciado anteriormente, quizás se debería considerar la posibilidad de elegir entre otro u otros modelos, beneficiándose de esta forma por ejemplo de su simplicidad funcional y de menor número de variables explicativas necesarias.

Ya que puede darse el caso de que no se pueda acceder a todos los datos que sean requeridos por las variables del modelo seleccionado, por eso es aconsejable tener en cuenta otros modelos que pese a no poseer tan buenos índices de información sean más fáciles de analizar.

Así, si sabiendo que el modelo que mejor explica el TIEPI, eligiéramos otro modelo de entre los cuatro restantes para salvar alguna de las circunstancias comentadas anteriormente, no nos resultaría fácil su elección ya que por ejemplo el modelo de regresión lineal es aconsejable desde el punto de vista del coeficiente de determinación corregido pero no desde el criterio suma de cuadrado de los residuos, ni tampoco del resto de los otros criterios. Esto pone de manifiesto, el hecho de que existen varios modelos que resultan favorables en algunos índices pero no en todos ellos, de forma que estamos ante la opción de elegir un modelo de entre varios, los cuales se encuentran en conflicto cuando una serie de índices se pretenden conseguir, esto es, un problema de decisión multicriterio.

4. Aplicación del método Promethee en la elección de un modelo de la continuidad del suministro eléctrico.

A continuación, se procederá a dar solución a este problema de carácter multicriterio planteado, con la ayuda del soporte informático denominado Decision Lab.

El problema multicriterio planteado es dar solución a la elección de un/os modelo/os de entre cinco posibles que pueden encontrarse en conflicto cuando cinco criterios de información se pretenden conseguir. Para lo cual se tendrá que definir en primer lugar la denominada *matriz de decisión*, la cual es prácticamente igual a la tabla 1, siendo necesario expresar si el objetivo de cada criterio es maximizarlo o minimizarlo, el tipo de criterio generalizado y el peso asignado.

En primer lugar, desde el punto de vista de su maximización o minimización, habría que decir que cuanto mayor sea el valor alcanzado por el coeficiente de determinación corregido mejor

será el ajuste del modelo hacia los datos reales, luego su objetivo será maximizarlo. Sin embargo, los restantes criterios como son función todos de SCR (Suma de los cuadrados de los residuos) , su objetivo será minimizarlos, ya que expresa la variación residual o no explicada de los valores del TIEPI con respecto a la línea o recta de regresión.

Respecto al tipo de criterio asignado a cada criterio, en todos resulta ser el mismo, el criterio tipo 1, el cual considera que cualquier diferencia entre los valores es significativa. En cuanto a la importancia dada a cada indicador ésta no ha sido la misma para todos así, al coeficiente de determinación corregido se le ha asignado el peso 1, mientras que al resto de los índices la mitad por representar la misma información. Los valores alcanzados por cada uno de los modelos en cada uno de los criterios se presenta en la siguiente matriz de decisión:

Se puede observar, en la tabla 2, que el modelo mejor situado es el modelo log lineal, por ser el que alcanza mejor evaluación en todos los índices. En segundo lugar, se podría elegir al modelo log-lin , teniendo en cuenta que el número de variables consideradas por el modelo es el mismo y en tercer lugar el modelo lineal, el cual reduce a dos el número de variables explicativas, hace depender la continuidad del suministro eléctrico exclusivamente de la longitud de las líneas de baja y alta tensión.

Lineal	3	Lin-log	4
Log-lineal	1	Recíproco	5
Log-lin	2		

Tabla 2: Ordenación de los modelos según Promethee

Como era de esperar a la vista de los resultados el modelo recíproco es el que menor información proporciona acerca del TIEPI. A la vista de los resultados, los tres modelos que mayor información aportan a la explicación del índice de continuidad del suministro eléctrico, TIEPI, son el modelo log-lineal, modelo log-lin y modelo lineal. Considerando que los dos primeros modelos dependen del mismo número de variables explicativas, cuatro, coincidiendo tres de ellas en ambos, éstas son, la potencia instalada en las centrales eléctricas, el número de empleados y el número de clientes. Respecto a la longitud de las líneas de transporte y distribución de la energía eléctrica el modelo log lineal incluye la longitud de las líneas de baja tensión y el modelo lin-log las líneas de alta tensión. Esta circunstancia de coincidencia de tres de las cuatro variables consideradas para los dos modelos mejor evaluados es un argumento consistente para la utilización de dichos modelos.

Por otra parte, el tercer modelo aconsejado desde el punto de vista de su grado de información acerca del TIEPI es el modelo lineal, este es el modelo más simple de entre los mencionados hasta ahora, ya que en primer lugar no necesita ninguna transformación de sus variables ya que todas son lineales, y en segundo lugar, suministra suficiente información (el grado de ajuste del modelo es del 78 por ciento) necesitando tan sólo dos factores, éstos son la longitud de las líneas de baja y alta tensión, circunstancia que minimiza el riesgo de accesibilidad a los datos.

Este análisis no quedaría completo, sin estudiar como afectaría al resultado una variación en los pesos asignados a los distintos indicadores, de esta forma para comprobar la robustez del análisis efectuado, se realiza un análisis de la sensibilidad de los pesos asignados a cada modelo como aparece reflejado en la tabla nº.

Este análisis proporciona información acerca de los intervalos en los que pueden variar los pesos (expresado en valor absoluto y relativo) sin que ello altere la ordenación de los modelos establecida, tal como se refleja en la tabla 3. Así, por ejemplo, podríamos incrementar en dos puntos porcentuales cualquiera de los indicadores sin que el decremento en estos dos puntos en cualquiera del resto de los modelos afecte a la ordenación ya establecida de los mismos. Por lo tanto, este análisis pone de manifiesto la estabilidad de la ordenación de los modelos ante cambios sensibles en los pesos de los criterios de información.

Criterio	Pesos	Intervalo	Pesos (%)	Intervalo (%)
\bar{R}^2	1	[0, 1'5]	33'33%	[0, 42'86]
SCR	0'5	[0, 1'5]	16'67%	[0, 37'50]
Akaike	0'5	[0, +∞]	16'67%	[0, 100]
Amemiya	0'5	[0, 1'5]	16'67%	[0, 37'50]
Schwarz	0'5	[0, +∞]	16'67%	[0, 100]

Tabla 3: Análisis de la Sensibilidad.

Una vez resuelto el problema de decisión multicriterio y analizada la estabilidad del resultado, se propondrían tres modelos para dar respuesta a las variaciones de la continuidad del suministro eléctrico, los cuales se expresan en la tabla 4 a través de la siguiente estimación mínimo cuadrática:

Modelo	Ecuación de la estimación
Log-lineal	$\ln \hat{y}_i = 2.41 + 0.86 \ln X_{1i} + 1.28 \ln X_{2i} - 0.46 \ln X_{5i} - 1.32 \ln X_{6i}$
Log-lin	$\ln \hat{y}_i = 0.64 + 2.54 \cdot 10^{-4} X_{1i} + 3.26 \cdot 10^{-4} X_{3i} - 1.51 \cdot 10^{-4} X_{5i} - 1.14 \cdot 10^{-6} X_{6i}$
Lineal	$\hat{y}_i = 1.87 + 1.73 \cdot 10^{-5} X_{2i} - 1.17 \cdot 10^{-4} X_{3i}$

Tabla 4: Modelos de Económicos Seleccionados.

donde X_1 : potencia instalada expresada en megavatios.

X_2 : longitud de las líneas de baja y media tensión, expresada en kilómetros.

X_3 : longitud de las líneas de alta tensión, expresada en kilómetros

X_4 : energía facturada, expresada en gigavatios/hora.

X_5 : número de personas empleadas.

X_6 : número de clientes.

En la estimación de los modelos, se aprecia una especificación correcta de las variables, a priori, pues los signos de los coeficientes de regresión se corresponden con lo esperado en el ámbito teórico: de esta forma en los dos primeros modelos (log-lineal y log-lin) se verifica que nuevas incorporaciones de la potencia instalada en centrales eléctricas incrementan la ocurrencia de interrupciones en la continuidad del suministro eléctrico. Por otra parte, se verifica que la longitud de las líneas de transporte y distribución de la electricidad propician

un incremento del tiempo de interrupción equivalente de la potencia instalada. Además, en los modelos log lineal y log lin, un incremento en el número de empleados (personal técnico y administrativo) y el número de clientes provoca un aumento en el tiempo en el que la red está disponible.

Los parámetros estimados de los tres modelos seleccionados reflejan una aproximación al concepto de *elasticidad*, el cual para cada uno de los modelos tendrá una interpretación distinta, ya que los tres modelos elegidos son no lineales en las variables (modelo log-lineal y modelo lin-log) o lineales en las variables (modelo lineal) pero todos los modelos planteados y seleccionados sí son lineales en los parámetros. Así, cualquiera de los coeficientes de regresión anteriormente expresados respecto a cada una de las variables representa la variación del tiempo de interrupción de la potencia instalada ante cambios en cada una de las variables explicativas, es decir, de que forma el TIEPI es sensible a las variaciones de las mismas.

A partir de los resultados obtenidos para cada uno de los tres modelos, pueden realizarse las siguientes observaciones, según orden de preferencia del modelo:

Modelo Log lineal: el cual relaciona el TIEPI con la potencia instalada, longitud de las líneas de baja tensión, el número de empleados y clientes (todas ellas logaritmizadas). Desde el punto de vista de sus coeficientes de regresión habría que comentar, por ejemplo, el coeficiente elasticidad- potencia instalada es de 0.86, lo cual implica que para un aumento del un punto porcentual por MW de capacidad, la posible interrupción del suministro en promedio aumenta aproximadamente en 0.86 por ciento. En cuanto a los coeficientes de la longitud de las líneas de baja, el signo de sus coeficientes elasticidades nos indica, que un incremento en un punto porcentual por Km de línea de alta tensión aumenta el TIEPI por término medio un 1.28 por ciento. La razón de este comportamiento distinto de ambas variables, que parecen a priori tener efectos similares es que las líneas de alta tensión son fundamentales en el transporte de energía eléctrica ya que es la vía a través de la cual la energía generada en origen se suministra a través de las redes, constituye la primera fase de desplazamiento de la energía, denominada *transporte*. De esta forma, un *corte* de suministro eléctrico en la anteriormente fase mencionada supondría importantes efectos a nivel técnico, cuyas consecuencias soportarían un también importante número de usuarios (domésticos e industriales). Respecto al coeficiente de elasticidad-número de empleados, un incremento en un punto porcentual en el número de empleados tiene como consecuencia un descenso en 0.46 por ciento del TIEPI, esto es consecuencia de que cuanto mayor sea el número de empleados ya sean de carácter técnico o administrativo antes quedará reestablecida la incidencia. Por último, el dato del coeficiente de elasticidad del número de clientes parece favorecer el TIEPI, la explicación a este comportamiento no quedaría muy clara y para ello sería preciso desagregar su número en clientes industriales y domésticos, ya que parte de los clientes industriales poseen plantas de generación propias.

Modelo Log-lin: este modelo o ecuación de regresión explicativa del TIEPI depende como en el modelo log lineal de cuatro factores o variables, siendo tres de ellas coincidentes (potencia instalada, número de empleados y número de clientes), a excepción del cuarto factor, la longitud de las líneas de baja tensión.

Modelo Lineal: Este modelo, a nuestro juicio, es el más favorable, ya que depende sólo de dos

factores los cuales explican suficientemente y según los criterios de información el TIEPI. En cuanto a los coeficientes de la longitud de las líneas de alta y baja tensión, el signo de sus coeficientes elasticidades no es el mismo, ya que un incremento en un

5. Conclusiones.

Del trabajo realizado se concluye que si se tuviera que elegir algún modelo que nos ayudará a entender la continuidad o discontinuidad del suministro eléctrico en España, sería adecuados, desde el punto de vista de la significación estadística, tres modelos: el modelo log-lineal, modelo lin-log y modelo lineal. Teniendo cada uno de los modelos características propias que favorecen o dificultan su elección, tales como el grado de asociación lineal, el número de variables exógenas consideradas, así como la información suministrada de los criterios de información. Todas estas consideraciones han sido tratadas a través de la formulación de un problema multicriterio mediante la técnica Promethee, que ha ayudado a la elección del modelo log-lineal.

Agradecimientos

Se desea agradecer a la profesora María Teresa Arévalo Quijada las sugerencias y comentarios brindados para la realización de este trabajo.

Referencias

- [1] Eviews. User Guide. Version 2.0. Quantitative Micro Software (1995).
- [2] Informe Anual de Endesa, S.A. 1998, 1999, 2000, 2001.
- [3] Informe Anual de Hidroeléctrica del Cantábrico, S.A. 1998, 1999, 2000, 2001.
- [4] Informe Anual de Iberdrola, S.A. 1998, 1999, 2000, 2001.
- [5] Informe Anual de Unión Fenosa, S.A. 1998, 1999, 2000, 2001.
- [6] Mc. Quarrrie (2002), "Regresión and Times Series Selection". World. Scientific
- [7] Memoria de Red Eléctrica, 1998, 1999, 2000, 2001.

*V Congreso de Ingeniería de Organización
Valladolid-Burgos, 4-5 Septiembre 2003*