Un modelo Arima de la difusión de un diario de la prensa española: el caso de *El Faro de Vigo*[1]

Prof. Dr. Valentín A. Martínez Fernández*, Prof. Dr. Carlos Bouza Herrera**,

Prof. Javier Orosa González* y Prof. Oscar Juanatey Boga*

*Área de Comercialización e Investigación de Mercados. Universidad de A Coruña

**Área de Estadística. Universidad de La Habana

Abstract

This article studies the diffusion of the one most important newspapers in Spain, it´s the most ancient, it´s called: *El Faro de Vigo.*

First of all we do a descriptive analysis before modelling the diffusion of *El Faro de Vigo* using Box-Jenkins methodology. We identify the ARIMA model which has the best explanation of the time serie studied. Data belongs to Oficina de Justificación de la Difusión (OJD) and econometric software used is EViews 5.1.

The model $ARMA\ (2,2)*(1,2)^{12}$ is best ARIMA model for the diffusion of *El Faro de Vigo* in 1991-98..

Keywords: journalist marketing, diffusion, ARIMA models, Augmented Dickey-Fuller test

JEL code: C220 y M310

Resumen

El presente documento de investigación centra su estudio en la difusión de uno de los principales diarios de España, hablamos del decano de los periódicos españoles: *El Faro de Vigo.*

[1] Los autores quieren expresar su agradecimiento a D. Ángel Durández y a D. Eusebio Serrano, presidente y director de la Oficina de Justificación de la Difusión (OJD) por su inestimable colaboración en la obtención de los datos necesarios para la realización de la presente investigació

Después de realizar un análisis descriptivo, se procede a la modelización econométrica de la serie objeto de estudio. Para ello se utiliza, la metodología de Box-Jenkins y en concreto se identifica y estima el modelo ARIMA que mejor explica el comportamiento de la serie. Los datos han sido extraídos de la Oficina de Justificación de la Difusión (OJD) y el paquete econométrico utilizado para su procesamiento es el EViews 5.1.

De todos los posibles modelos ARIMA, el modelo que mejor aproxima los datos disponibles es el $ARMA\ (2,2)*(1,2)^{12}$.

Palabras clave: marketing periodístico, difusión, modelos ARIMA, test de Dickey-Fuller Aumentado

Códigos JEL: C220 y M310

1.Introducción

Fundado por Ángel de Lema y Marina, uno de los precursores de la empresa periodística española, *El Faro de Vigo* publicó su primer número el 3 de noviembre de 1853, lo que le convierte en el diario decano de la prensa editada en España. En 1986 fue adquirido por Prensa Ibérica, grupo de comunicación que contaba a finales de 2004 con catorce cabeceras y de la cuales una de ellas, *La Opinión de A Coruña*, se edita también en la comunidad autónoma española de Galicia.

El Faro de Vigo, con una difusión media en 2003 de 42.242 ejemplares, ocupa el segundo lugar en el ranking de difusión de los diarios comercializados en Galicia y el décimo quinto del mercado español (OJD, 2005). Ello, junto con los 287.000 lectores diarios que le reconoce el Estudio General de Medios (EGM) en la oleada de marzo de 2004 lo consolidan con el periódico líder en el sur de Galicia a través de seis ediciones que segmentan dicho mercado geográfico y que corresponden a las áreas de Vigo, Pontevedra, Arousa, Ourense, Morrazo y Deza-Tabeirós-Montes. Así, la distribución geográfica de su difusión, también en 2003, era del 91,91 por ciento en la provincia de Pontevedra, el 3,08 en la de Ourense, el 3,01 en la de A Coruña y el 0,67 en la de Lugo. Mientras que el 0,81 por ciento restante corresponde a otras provincias españolas y el 0,52 al extranjero, fundamentalmente el norte de Portugal (OJD, 2004). Igualmente es significativo el ratio de rotación de cada ejemplar al ser 6,8 lectores, toda vez que en España este se sitúa en 3,1 y en Galicia en el 3,5 (AEDE, 2005)

La difusión de *El Faro de Vigo* se obtiene fundamentalmente por medio de la comercialización de sus ejemplares en los puntos de venta y que representa el 77,87 por ciento de la misma, mientras que la suscripción de carácter individual nada más supone el 13,36 por ciento, lo cual lo distancia de los diarios de marcado acento local publicados en Galicia y que cuentan, como nota común, con una suscripción superior al

25 por ciento y entre los que el diario *El Progreso*, editado en Lugo, se muestra como paradigmático al contar con un 32,76 por ciento de su difusión destinada al citado tipo de suscripción (AEDE, 2005)

Por otra parte, en 1998 *El Faro de Vigo* fue galardonado con el Premio Stendhal para el periodismo y la comunicación en Europa.

La antigüedad en el mercado periodístico español, su ámbito de difusión que trasciende lo meramente local al contar con un carácter interprovincial, pero, a la vez, propio de un área determinada como es la que corresponde al sur de Galicia, junto con el hecho de que, al igual de su cabecera hermana *La Opinión de A Coruña*, no pertenece a una empresa familiar sino a uno de los principales grupos periodísticos españoles tanto por el número de cabeceras cómo por el volumen de negocio, al tiempo de presentar una difusión que no refleja a priori un alto grado de fidelización al tener un bajo porcentaje de suscripción, pero, sin embargo, con un importante arraigo e influencia social reflejada en el número de sus lectores y en el correspondiente índice de rotación muestran a Faro de Vigo como un diario que reúne las características básicas para llevar a cabo un análisis de la difusión desde la perspectiva metodológica que se aborda en este trabajo de investigación y en orden a los objetivos planteados en el mismo.

2.Marco teórico

Box y Jenkins desarrollaron un cuerpo metodológico destinado a identificar, estimar y diagnosticar modelos dinámicos de series temporales en los que la variable tiempo juega un papel fundamental. Este se originó en el estudio de series de tiempo que caracterizaban el proceso de la contaminación de la bahía de San Francisco. Una parte importante de esta metodología está pensada para liberar al investigador de la tarea de especificación de los modelos. Son los propios datos temporales de la variable objeto de estudio los que indican las características de la estructura probabilística subyacente. En parte, estos procedimientos se contraponen a la forma tradicional de identificar y especificar un modelo apoyándose en las teorías subyacentes al fenómeno analizado. Convenientemente utilizados, estos conceptos y procedimientos constituyen una herramienta útil para ampliar y complementar los conocimientos econométricos básicos.

La consideración exclusiva de los valores pasados de una determinada variable para explicar su evolución presente y futura supone ventajas e inconvenientes.

La principal ventaja reside en el hecho de no necesitar nada más que una serie de tiempo y no de varias como es el caso de la econometría tradicional. El inconveniente más importante es la pérdida de capacidad de análisis, puesto que se renuncia a la utilización de otras variables con las que puede existir algún tipo de relación causal.

Un marco adecuado para el estudio de datos de este tipo son los modelos autorregresivos, tal y como señalan Chatfield (2003) y Rynkiewicz (2004). Dentro de los modelos univariantes aparecen los llamados modelos ARIMA. Esta palabra es un acrónimo: Autorregresive, Integrated and Moving Average (modelos Autorregresivos, Integrados y de Medias Móviles). La complejidad del modelo ARIMA es caracterizada en la notación usual de ARIMA(p, d, q) que es una combinación de los modelos de las medias móviles MA(q), la identificación del orden de la diferenciación I(d) y el proceso autorregresivo AR(p). El modelo matemático es escrito en su forma general como:

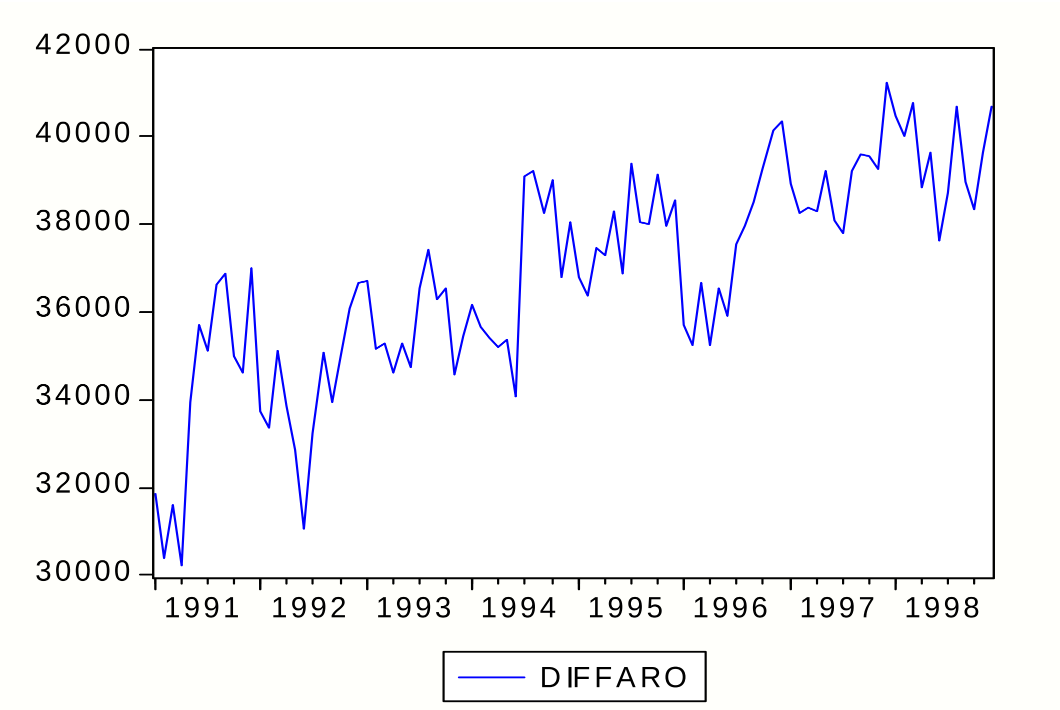
$$Y_T = \varphi_1 Y_{T-1} + \varphi_2 Y_{T-2} + \Lambda + \varphi_{Ps+p+Ds+d} Y_{T-Ps-p-sD-d} + \\ \delta + U_T + \theta_1 U_{T-1} + \Lambda + \phi_{Qs+q} U_{T-sQ-q} \quad (1)$$

Es de interés del presente trabajo estudiar las series temporales originadas por las ventas de *El Faro de Vigo*. La fuente de la que se han extraído los datos, es la Oficina de Justificación de la Difusión[2] (OJD) y el paquete econométrico utilizado para su procesamiento es el EViews 5.1.

3.Análisis descriptivo

Como base se toman una serie de datos temporales. El primer paso en su estudio es el análisis del gráfico que estos determinan a lo largo del tiempo. Esto constituye un análisis descriptivo preliminar. El gráfico 1, permite observar la evolución de la difusión del diario *El Faro de Vigo* en el periodo 1991:01-1998:12. Un simple análisis sugiere la existencia de una tendencia creciente de la serie

Gráfico 1. Difusión de *El Faro de Vigo* en el periodo 1991-98



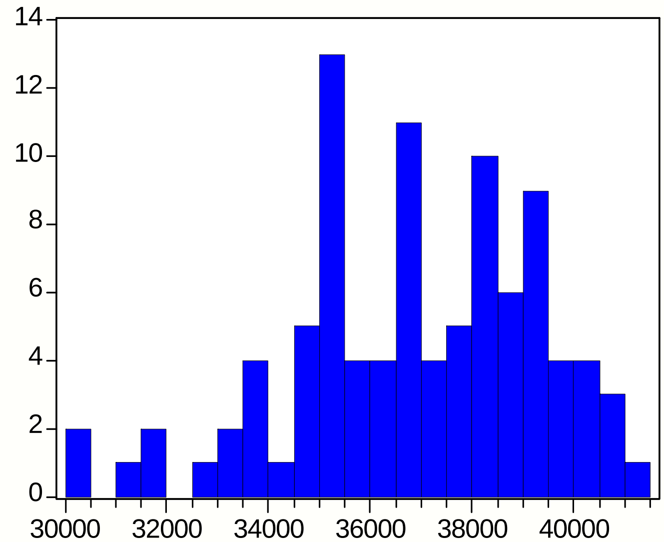
Fuente: Elaboración propia

---

[2] La Oficina de Justificación de la Difusión (OJD) es la entidad constituida por empresas periodísticas, agencias de publicidad y anunciantes con el objeto principal de obtener y facilitar información útil y puntual de la difusión y distribución de las publicaciones periódicas, para uso de los anunciantes, agencias de publicidad, editores y demás personas o entes interesados. El Reglamento de la Oficina de Justificación de la Difusión permite la realización de controles parciales voluntarios.

La media de la difusión de *El Faro de Vigo* para el periodo objeto de estudio es de 36.829,39 ejemplares. Esta constituye el llamado ' nivel' de la serie. Se nota que alrededor de 1996 la difusión de este diario se posiciona por encima del nivel. La serie alcanza un valor máximo en diciembre de 1997 con 41.227 ejemplares y un valor mínimo en abril de 1991 con 30.232 ejemplares. La desviación típica de la serie asciende a 2.418,637. Las principales características numéricas y algunos estadísticos notables de la serie objeto de estudio aparecen recogidos en el gráfico 2:

Gráfico 2. Principales estadísticos de la difusión de *El Faro de Vigo*



Series: DIFFARO	
Sample 1991M01 1998M12	
Observations 96	
Mean	36829.39
Median	36833.50
Maximum	41227.00
Minimum	30232.00
Std. Dev.	2418.637
Skewness	-0.519324
Kurtosis	2.941778
Jarque-Bera	4.328719
Probability	0.114823

Fuente: Elaboración propia

Estos indican que la asimetría y la kurtosis son cercanos al esperado de una variable con distribución normal que deberían tomar los valores cero y tres respectivamente. El test de Jarque-Bera refleja que los datos se ajustan adecuadamente a los provenientes de una distribución normal.

El siguiente paso en el análisis de una serie temporal consiste en analizar si la secuencia de valores exhibe algún patrón a lo largo del tiempo o si éste puede ser considerado aleatorio. Si pueden encontrarse patrones de regularidad en el comportamiento de la serie es posible determinar un modelo matemático que permita hacer predicciones sobre las realizaciones futuras de la serie. Al estimar válida la existencia de un patrón de regularidad es factible utilizar algún modelo probabílistico para describirla. El modelo matemático que es empleado para describir el comportamiento de la secuencia de variables aleatorias es llamado *proceso estocástico*. La naturaleza de tales procesos se caracteriza por el hecho de que establecer el futuro de la serie no puede ser pronosticado con exactitud usando la información del pasado. Tal es el caso del problema aquí abordado. Modelar adecuadamente una serie temporal mediante un proceso estocástico permite hacer predicciones y valorar el error de tales predicciones a través del uso de un instrumental científico.

4.Modelo Arima de la difusión de *El Faro de Vigo*

Un modelo importante para modelar las series temporales son los llamados procesos estocásticos estacionarios. Muchas series de tiempo exhiben un comportamiento no estacionario en problemas financieros y económicos y los precios de artículos. Sin embargo, este tipo de serie puede ser convertida en estacionaria cuando se toman las diferencias de primer o segundo orden. Una serie es estacionaria si sus propiedades no cambian a lo largo del tiempo (equilibrio estadístico). Un modelo popular para el análisis de tales series es el llamado ARIMA que describe series de tiempo no estacionarias que pueden convertirse en estacionarais usando estas diferencias.

Dadas las características de las series de venta, en el presente documento de investigación se muestra la construcción de un modelo ARIMA para explicar el comportamiento de la difusión del diario *El Faro de Vigo*. Para ello se dispone de una serie de tiempo mensual para el periodo 1991:01-2000-12. Este periodo se ha dividido en dos subperiodos: a) periodo de estimación del modelo que va desde 1991 hasta 1998 y b) periodo de predicción 1999-2000.

La complejidad de un modelo ARIMA es caracterizada por el número de parámetros que envuelve. Si usamos demasiados parámetros esto conduce a modelos que son malos predictores y se dice que se ha 'sobre-ajustado' el modelo. En la selección de un adecuado modelo para representar la serie de tiempo debemos seguir el llamado principio de la parsimonia que establece utilizar como primera aproximación un modelo bien simple.

La construcción de un modelo ARIMA comprende el cumplimiento de cuatro fases sucesivas: identificación, estimación, validación y predicción.

a) Fase de Identificación

El objetivo de la presente fase es la identificación de las distintas partes del modelo. Esto es, la componente autorregresiva (*p*), la que indica el retardo de la variable dependiente, y la correspondiente a las medias móviles (*q*), el retardo del término de error. En algunos casos se incluye en la especificación del modelo ARIMA el parámetro *d* que determina el grado de las diferencias usado para lograr la estacionalidad.

Antes de proceder a la identificación de los valores de *p* y *q* parece oportuno recordar que los modelos ARIMA forman parte de la denominada metodología Box-Jenkins y por lo tanto estas herramientas han sido desarrolladas para procesos estocásticos estacionarios (Box y Jenkins, 1970).

Así, lo primero que ha de hacerse es convertir la serie de observaciones que constituyen la base de datos en una serie estacionaria. Este proceso se conoce como 'estabilización' de la serie. Para ello deben efectuarse algunas transformaciones que permitan establecer la existencia de regularidades y su naturaleza. Esto quiere decir que se debe previamente realizar un conjunto de transformaciones en busca de la citada estacionariedad.

Si se juzga que toda serie de tiempo puede ser considerada como la realización particular de un proceso estocástico subyacente, una serie de tiempo estacionaria es aquella que se corresponde a la realización de un proceso estocástico estacionario. Un proceso estocástico es estacionario[3], si su media y varianza son constantes en el tiempo. Además, se estima que las autocorrelaciones son independientes del tiempo, o sea que el valor de la covarianza entre dos periodos depende exclusivamente de la distancia entre los dos periodos y no del tiempo en el cual se ha calculado (Gujarati, 1997). Cuando una serie de tiempo no cumple lo anteriormente expuesto se considera: serie de tiempo no estacionaria.

Varios métodos son recurrentes para realizar estas transformaciones. Estos son denominados 'filtros' y su uso en los procesos estocásticos se antojan de gran utilidad a juicio de Peña, (1987) y Elliot (1994).

Transformación para estabilizar la varianza

Al objeto de obtener, eliminar o reducir la inestabilidad de la varianza es necesario realizar una transformación en la serie objeto de este estudio. En este orden de cosas, puede demostrarse que cuando existe una relación entre las medias, μ_i , de un conjunto de variables aleatorias, x_i , y sus desviaciones típicas, σ_i , del tipo:

$$\sigma_i = k\mu_i^{\alpha} \tag{2}$$

al transformar las variables x_i en nuevas variables y_i mediante:

[3] La literatura econométrica denomina a estos procesos estocásticos como débilmente estacionarios.

$$y_i = \frac{x_i^{1-\alpha} - 1}{1-\alpha} \tag{3}$$

estas nuevas variables y_i tienen varianza constante. La transformación (3) se conoce como transformación (filtro) de Box-Cox y depende de un parámetro $\lambda = 1-\alpha$, e incluye básicamente las transformaciones mediante potencias de la variable. (Box y Cox, 1964). En efecto, la transformación es aparte de una constante, $y_i = x_i^{\lambda}$. Al escribirla de esta forma, esta transformación incluye también el logaritmo, ya que cuando α es muy próximo a uno, o λ próximo a cero, puede comprobarse tomando límites en la expresión (3) que se obtiene el logaritmo de la variable.

Por lo tanto, se lleva a cabo la transformación que toma *lambda* igual a cero que se materializa aplicando logaritmos con lo que se ha pasado de la serie DIFFARO, difusión del diario *El Faro de Vigo*, a la serie LDIFFARO.

Transformaciones para estabilizar la media

El punto de partida es que la estabilización de la media en una serie no estacionaria puede requerir la aplicación de diferencias regulares y diferencias estacionales. Son muy usados el ajuste de polinomios y el de alisado exponencial, pero el más corrientemente utilizado es el de las diferencias sucesivas.

La llamada diferencia de primer orden está dada por calcular el ‘retardo de orden 1’:

$$\nabla x_{t+1} = x_{t+1} - x_t \tag{4}$$

la de orden 2 es lograda por el ‘retardo de orden 2’y es obtenida por diferenciar la nueva serie restando los nuevos valores consecutivos:

$$\nabla^2 x_{t+2} = x_{t+2} - x_{t+1} \tag{5}$$

A continuación se procede a verificar si la serie LDIFFARO, es estacionaria o no estacionaria y en caso de ser no estacionaria el orden de integración de la serie (d).

Cuando una serie de tiempo es no estacionaria y su primera diferencia es estacionaria se dice que la original es una serie integrada de orden (grado) d=1, y se denota por I(1). La serie original se dice que tiene una raíz unitaria y se denomina *caminata aleatoria.* En general, una serie de tiempo que necesita diferenciarse *d* veces, se dice que es una serie integrada de orden *d* y se representa I(d). (Gujarati, 1997)

Para la realización del estudio del orden de integración de una serie pueden utilizarse distintos métodos: análisis del gráfico de la serie, estudio de la función de autocorrelación simple (fas) y las pruebas de raíces unitarias (Maddala, 1996). La fas mide la correlación entre los valores de la serie distanciados en un lapso de tiempo h. Esta consiste en el conjunto de coeficientes de autocorrelación con h=1,...,n/2, (n es el numero de observaciones). Si hay estacionalidad los valores separados entre si por intervalos iguales deben estar correlacionados significativamente en alguna forma.

De los diferentes tests existentes para verificar la presencia de una raíz unitaria en una serie temporal, el test de Dickey-Fuller Aumentado es el más utilizado. (Dickey y Fuller, 1979). Por tato, de aceptarse que la serie tiene raíz unitaria la estacionalidad puede ser considerada verdadera.

Una prueba de naturaleza intuitiva como es la visión del gráfico de la serie LDIFFARO, apunta la hipótesis de que nos encontramos ante una serie no estacionaria en media. Este sugiere que existe una tendencia creciente. Por lo que respecta a la función de autocorrelación simple (fas), se observan autocorrelaciones positivas, con decrecimiento lento y gradual. Esto es un signo evidente de la presencia de una serie no estacionaria. No obstante, para mejorar el estudio se procede a realizar el test de Dickey-Fuller Aumentado

Test de Dickey-Fuller Aumentado

Dickey y Fuller (1979) establecieron un test para hacer contraste de raíz unitaria. Este es muy sensible a la presencia de 'outliers' lo que hace que su uso deba ser realizado con cuidado. Este se basa en el uso de una regresión simple. Una mejoría es el test de Dickey-Fuller Aumentado que consiste en el cálculo de una regresión para la primera diferencia de la serie. En el modelo se consideran como variables independientes la propia serie en el periodo anterior y las diferencias en los términos retardados. También se suele incluir una constante y una tendencia temporal.

$$\Delta y_t = \alpha_1 y_{t-1} + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + ... + \beta_k \Delta y_{t-k} + \gamma + \Psi_t \qquad (6)$$

En el test de Dickey-Fuller Aumentado (Dickey y Fuller, 1981), se verifica que el valor del coeficiente del término retardado en la regresión sea significativamente distinto de cero [H1: La serie no tiene raíz unitaria]. Si se rechaza esta hipótesis la serie contiene una raíz unitaria [H0] y se estaría ante una serie estacionaria. Este test proporciona el valor t del coeficiente (estadístico *tau*) y su nivel de significación. El estadístico *tau*, no se distribuye como una *t de Student* por lo que para hacer el test se utilizan los valores críticos particulares de su distribución. (McKinnon, 1991).

Se procede a la realización del test de Dickey-Fuller Aumentado para la cual se utiliza un modelo con ordenada en el origen y se toma un máximo de retardos igual a once (0<h<12). El estadístico *tau* (–2.980618) obtenido presenta un resultado inferior en valor absoluto al 1% de los valores críticos de Mckinnon (-3.501445). Por lo tanto, no puede rechazarse la hipótesis nula de que la serie tiene una raíz unitaria y es necesario diferenciar la serie LDIFFARO para convertirla en estacionaria.

Como se ha demostrado de distintas formas, la serie LDIFFARO no es estacionaria y por lo tanto no puede emplearse una metodología diseñada exclusivamente para procesos estocásticos estacionarios. Se necesita transformar una serie no estacionaria en media en otra que sí lo es.

Por lo que respecta al componente regular, la transformación de una serie de tiempo no estacionaria en estacionaria, se realiza mediante la utilización del operador de primera diferencia d=1 es decir, se toman primeras diferencias lo que supone una pérdida de grados de libertad. En resumen, se pasa de la serie LDIFFARO a la serie DLDIFFARO.

Llegado a este punto, el estudio de la función de autocorrelación simple apunta a la presencia de una componente estacional de orden 12. Sin duda esta componente tiene su origen en la naturaleza mensual de la serie objeto de estudio.

El último paso consiste en el desmantelamiento de la componente estacional de la serie. Para ello se procede a tomar diferencias de orden 12 y se construye una nueva serie a la que se denomina DDLDIFFARO. La construcción de esta nueva serie trae consigo una nueva pérdida de grados de libertad.

Nuevamente, ha de llevarse a cabo la realización del test de Dickey-Fuller Aumentado y para ello se utiliza un modelo con ordenada en el origen y se toma un máximo de retardos igual a once. En este caso, el estadístico *tau* (–8.690150) obtenido presenta un resultado superior en valor absoluto al 1% de los valores críticos de McKinnon (-3.513344). Por lo tanto, puede rechazarse la hipótesis nula de que la serie DDLDIFPRO tiene una raíz unitaria; es decir, se está ante una serie de tiempo estacionaria.

En resumen, se ha transformado una serie de tiempo no estacionaria (DIFFARO) en otra estacionaria (DDLDIFFARO) con el objetivo de poder aplicar la metodología Box-Jenkins específicamente diseñada para procesos estocásticos estacionarios.

Identificación de la parte autorregresiva (*p*) y la de medias móviles (*q*)

Una vez estabilizada la serie original y tener una serie estacionaria, se está en condiciones de proceder a la identificación de las distintas partes del modelo: parte autorregresiva [AR] y parte de medias móviles [MA]. Para ello se emplean dos instrumentos de naturaleza estadística como son: la función de autocorrelación simple y la función de autocorrelación parcial. Una vez calculadas la función de autocorrelación simple estimada y la función de autocorrelación parcial estimada, se procede a su comparación con la función de autocorrelación simple teórica y la función de autocorrelación parcial teórica de distintos procesos ARMA con el objeto de identificar el proceso estocástico que ha generado la serie objeto de estudio. Esto es llevado a cabo al compararles con un catálogo de patrones que tipifican a los distintos modelos propuestos, ver Chatfield (2003).

La interpretación de la función de autocorrelación simple (*fas*) y de la función de autocorrelación parcial (*fap*) muestrales es compleja por tres razones:

1.Cuando existe autocorrelación las estimaciones de las autocorrelaciones están a su vez correladas, lo que introduce una pauta de variación aleatoria en la *fas* y en la *fap* muestrales que se superpone a la verdadera pauta existente.

2.Los límites de confianza que han sido utilizados para juzgar si las autocorrelaciones son distintas de cero, $2/\sqrt{T}$, son asintóticos, y son poco precisos para las primeras autocorrelaciones.

3.Para procesos mixtos ARMA, la estructura de la *fas* y la *fap* teórica es muy complicada, y extremadamente difícil estimar el orden del proceso, incluso si se conocen los valores teóricos.

No obstante, en esta etapa no es necesario decidir cual es el orden del modelo, basta con seleccionar un conjunto de modelos ARMA que parezca adecuado para representar los rasgos principales de la serie objeto de estudio. Posteriormente, se procederá a la estimación de los modelos seleccionados para elegir aquel que resulte más adecuado.

El resultado de la comparación entre fas y fap muestrales y teóricas presenta un modelo con dos componentes autorregresivas, una de orden uno y otra de orden veinte y cuatro y una de medias móviles de orden uno.

Al objeto de profundizar en el análisis y una vez efectuada la anterior operación se procede a realizar pruebas de estimación con diferentes modelos ARMA. El resultado de estas pruebas no permite encontrar un modelo que supere al mencionado $ARMA\ (2,2)*(1,2)^{12}$.

b) Fase de Estimación

Una vez identificados los componentes del proceso estocástico es cuando se inicia la fase de estimación cuya finalidad principal es la obtención de los estimadores de los parámetros.

En primer lugar, el valor del estadístico *F* (13.36850) con una probabilidad prácticamente nula, permite rechazar la hipótesis de nulidad conjunta de los parámetros. (Véase tabla 2)

Tabla 1. Resultados de la estimación del modelo

Variable	Coefficientet	Std. Error	t-Statitstic	Probability
AR(2)	-0.302030	0.090307	-3.344467	0.0014
AR(12)	-0.302156	0.088747	-3.404699	0.0011
MA(2)	-0.980000	0.000206	-4749.118	0.0000
SMA(2)	0.750527	0.101707	7.379333	0.0000

Fuente: Elaboración propia

Los estadísticos *t* (-3.344467) (-3.404699) (-4749.118) (7.379333) presentan valores claramente superiores al valor estándar dos. Esto implica una elevada precisión de los estimadores. Sus probabilidades asociadas son muy próximas a cero por lo que puede descartarse la hipótesis de nulidad individual de cualquiera de los parámetros.

Por lo que respecta a la bondad del ajuste, el valor del R cuadrado ajustado (0.346439) indica la existencia de un buen ajuste si se tiene en cuenta que se está en el ámbito de un modelo ARIMA. (Véase tabla 2)

Por último, indicar los valores de los estimadores (-0.302030), (-0.302156), (-0.980000) y (0.750527).

c) Fase de Validación

El objetivo de esta fase es comprobar si el modelo se adecua a los datos. En caso de que sea así, estaremos en condiciones de pasar a la fase de predicción; de lo contrario será necesario una reformulación del modelo.

Para la validación del modelo es necesario estudiar el cumplimiento de los siguientes requisitos.

Tabla 2. Resultados de la validación del modelo

R-squared	0.374449	Mean dependent var	-0.001393
Adjusted R-squared	0.346439	S.D. dependent var	0.031679
S.E. of regression	0.025610	Akaike info criterion	-4.436966
Sum squared resid	0.043944	Schwarz criterion	-4.309491
Log likelihood	161.5123	F-statistic	13.36850
Durbin-Watson stat	2.343631	Prob(F-statistic)	0.000001

Fuente: Elaboración propia

En primer lugar, los residuos del modelo deben comportarse como un *ruido blanco*. En este orden de cosas, se aprecia que en el Correlograma la Función de Autocorrelación y la Función de Autocorrelación Parcial se mueven dentro de las bandas de fluctuación.

En segundo lugar, se verifica que los coeficientes del modelo son estadísticamente significativos. Para ello han sido empleados los estadísticos *F* y *t*. Las probabilidades de ambos estadísticos, muy próximas a cero, llevan a descartar las hipótesis de nulidad conjunta y nulidad individual de los parámetros, respectivamente.

En tercer lugar, para analizar la bondad del ajuste es utilizado el R cuadrado ajustado. El valor de este estadístico (0.346439) indica la existencia de un buen ajuste si se tiene en cuenta el hecho de que se está en el ámbito de la Econometría de Series de Tiempo.

Una vez verificado el cumplimiento de los anteriores requisitos cabe concluir que la elaboración del modelo $ARMA\ (2,2)*(1,2)^{12}$, se ajusta a los datos. Esto quiere decir, que se está en disposición de realizar predicciones.

d) Fase de Predicción

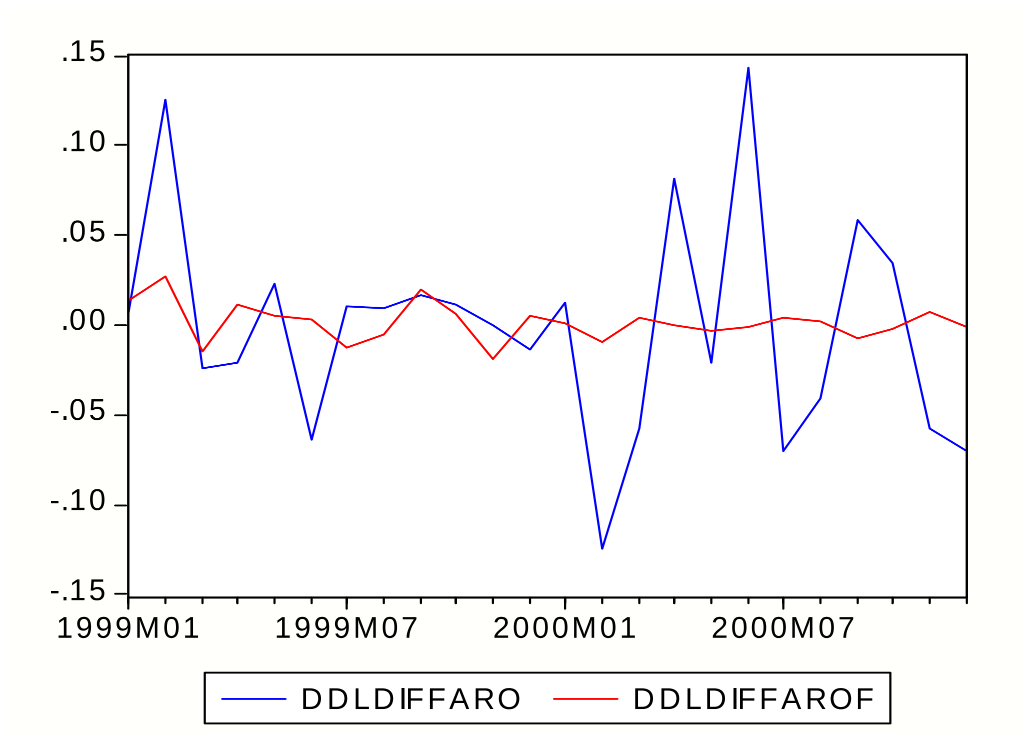
La realización de estimaciones de valores futuros de la variable en función de los ajustes hechos del modelo matemático se denomina predicción. La predicción en la serie temporal permitiría establecer una proyección. En este caso, lo que corresponde a la venta de ejemplares del diario. De efectuarse un estudio consecuente podrá establecerse el efecto de las políticas de difusión en las ventas a partir de conocer cual es el retardo. Si se cuenta con un retardo h, se sabrá que las implicaciones de una nueva política, por ejemplo una promoción de ventas, aparecerá sólo en el periodo h posterior.

A modo de síntesis, cabe decir que de la consulta de los datos certificados por la Oficina de la Justificación de la Difusión, en sus respectivas actas de control, se ha extraído la serie mensual de la difusión de *El Faro de Vigo* (DIFFARO). Esta serie ha sido dividida en dos periodos: el de estimación, que va desde el primer mes de 1991 hasta el último de 1998 y el de predicción, que comprende los años 1999 y 2000.

En el gráfico 3, aparece representada la difusión real de *El Faro de Vigo* (línea azul) y la difusión estimada con el modelo planteado en este trabajo (línea roja). Esta representación pone de manifiesto la buena capacidad predictiva del modelo ARIMA obtenido.

Es obvio que la evaluación de la capacidad predictiva del modelo no arroja resultados tan satisfactorios como si se hubiese procedido a analizar un modelo econométrico causal, pero no es menos cierto que de todos los posible modelos ARIMA que son susceptibles de ser construidos conforme a la serie objeto de estudio, éste es el que predice con una mayor adecuación a la realidad.

Gráfico 3. Comparación entre la difusión real y la estimada con el modelo ARIMA



Fuente: Elaboración propia

5.Conclusiones y futuras líneas de investigación

A continuación se presenta de forma ordenada y sistemática las conclusiones obtenidas gracias a la realización del presente trabajo.

1. La difusión de *El Faro de Vigo* ha experimentado, aunque con oscilaciones, un importante crecimiento en el periodo 1991-98. Pasando de los 31.843 ejemplares al inicio de periodo objeto de estudio a los 41.618 de diciembre de 1998. Se ha determinado que hay una tendencia positiva en esta serie.
2. La media de la difusión de *El Faro de Vigo* para el periodo es de 36.829,39 ejemplares. La serie alcanza un valor máximo en diciembre de 1997 con 41.227 ejemplares y un valor mínimo en abril de 1991 con 30.232 ejemplares. La desviación típica de la serie asciende a 2.418,637.
3. De todos los posibles modelos ARIMA, el modelo definido como $ARMA\ (2,2)*(1,2)^{12}$ representa la construcción matemática que mejor se ajusta a los datos disponibles. Este establece la regularidad (movimiento suavizado) del comportamiento de la serie tomando en cuenta los resultados del periodo analizado.

Con la presente investigación se ha verificado que la difusión de *El Faro de Vigo* sigue un patrón de comportamiento determinístico y, en cierta medida, puede ser explicada a partir de la difusión en los periodos anteriores. Esto quiere decir, que existe un fuerte componente de fidelidad en las pautas de comportamiento de los lectores de este diario.

El presente documento se integra dentro de una línea de trabajo de nuestro equipo de investigación. Nuestro objetivo es por un lado probar la metodología empleada en el resto de diarios españoles y por otro, proponer diferentes metodologías econométricas como la cointegración para intentar mejores modelizaciones del comportamiento de la difusión de este diario.

Bibliografía

Abraham, B y Ledolter, J. (1983): *Statistical Methods for Forecasting*, John Wiley and Sons, Nueva York.

AEDE (2005), *Libro Blanco de la Prensa Diaria 2005*, AEDE, Madrid

Álvarez, N. (1997): *Introducción a la evolución de la metodología de la Econometría*. Madrid: Universidad Nacional de Educación a Distancia.

Amemiya, T. (1985): *Advanced Econometrics*, Basil Blackwell, Oxford.

Ansley, C. F. (1983): Comentario a "Forecasting Structural Time Series with Structural and Box-Jenkins Models", Journal of Business and Economic Statistic, vol. 1, 307-309.

Asociación de Editores de Diarios Españoles, (2005), *Libro Blanco de la Prensa Diaria 2005*, AEDE, Madrid

Berndt, E. R. (1991): *The Practice of Econometrics: Classic and Contemporany*. Addison-Wesley.

Box, G. E. P., y Cox, D. R. (1964): "An Analysis of Transformations", *Journal of the Royal Statistical Society*, Series B, vol.26, pp. 211-243.

Box, G. E. P., y Jenkins. G. M.(1970): *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day.

Cancelo, J. R. (1990): "Actualización del modelo de predicción de la circulación fiduciaria: la diferencia estacional y la memoria del modelo reconsideradas", mimeo, Oficina de Operaciones del Banco de España.

Cancelo, J. R. y Espasa, A. (1991): "Common Trend Analysis for Consumer Prices Within and Across South-European Economies", trabajo presentado en la reunión anual de la *Association of Southern European Economic Theorists*, Atenas, noviembre 1.991.

Canova, F. (1992): “An Alternative Approach to Modeling and Forecasting Seasonal Time Series”, *Journal of Business and Economic Statistic*, vol. 10, 97-108.

Chatfield, c. (2003): The Analysis of Time Series: An Introduction. Ed. Chapman and Hall, Londres.

Chow, G. C. (1.983): *Econometric Methods.* New York: McGraw-Hill.

Dagum, E. B. (1978): “Modeling, Forecasting and Seasonally Adjusting Economic Time Series with the X11ARIMA Method” The Statistician, vol. 27, 203-216.

Dagum, E. B. (1980): The X11ARIMA Seasonal Adjustment Method, Statistic Canada, Ottawa.

Dagum, E. B. (1988): The X11ARIMA/88 Seasonal Adjustment Method, Statistics Canada, Ottawa.

Dickey, D. A.; Fuller, W. A. (1979). Distribution of the Estimators for Autorregresive Time Series with a Unit Root. *Journal of the American Statistical Association*, Vol. 74, 427-431.

Dickey, D. A.; Fuller, W. A. (1981). Likelihood Ratio Statistics for Autorregresive Time Series with a Unit Root. *Econometrica*, Vol. 49: 4, 1057-1072.

Elliot R.J. (1994): Exact adaptative filters for Markov cahins observed in Gaussian noise. Automatica, 30,9, 1399-1408.

Espasa, A. (1979): Modelos ARIMA univariantes con análisis de intervención para las series de agregados monetarios (saldos medios mensuales) M3 y M2”, Cuadernos Económicos de ICE, núms. 11-12, 109-146.

Espasa, A. y Cancelo, J. R. (1989): “Modelos univariantes en el análisis económico”, *Revista Española de Economía*, vol. 6, núm. 1-2, 85-107.

Espasa, A. y Peña, D. (1990): "Los modelos ARIMA, el estado de equilibrio en variables económicas y su estimación", *Investigaciones Económicas*, vol. XIV, 191-211.

Espasa, A. y Peña, D. (1991): "ARIMA Models, the Steady State of Economic Variables and their Estimation", Working Paper 91-02, Departamento de Economía, Universidad Carlos III de Madrid.

Granger, C. W. J. (ed, 1990): *Modelling Economic Series*, Clarendon Press, Oxford.

Goldberger, A. S. (1991): *A Course in Econometrics.* Cambridge: Harvard University Press.

Gujarati, D. (1997): *Econometría*. Tercera edición. Santafé de Bogotá, Colombia: Mc Graw Hill.

Harvey, A. C. y Todd, P. (1983): "Forecasting Economic Time Series with Structural and Box-Jenkins Models: a case study", *Journal of Business and Economic Statistic*, vol. 1, 299-315.

Hendry, D. F. y Richard, J. F. (1983): "The Econometric Analysis of Economic Time Series", *International Statistical Review*, vol. 51, 111-163.

Johnston, J. (1973): *Métodos de Econometría*. Londres: Macmillan.

Maddala, G. S. (1996). *Introducción a la Econometría* (2nd ed.) México D.F.: Ed. Prentice-Hall Iberoamericana.

Maravall, A. (1986): "Revisions in ARIMA Signal Extraction", Journal of the American Statistical Association, vol. 81, 736-740.

Maravall, A. (1988): "The Use of ARIMA Models in Unobserved Components Estimation" en W. Barnett et al. (eds., 1.988), 171-196.

McKinnon, J. (1991). Critical Values for Cointegration Tests, in Engle, R. F. y Granger, C. W. J. (eds.); *Long Run Economic Relatioships: Readings in Cointegration*. New York: Ed. Oxford University Press.

Mills, T. C. (1990): *Time Series Techniques for Economists.* Cambridge: Cambridge University Press.

Mills, T. C. (1993): *The Econometric Modelling of Financial Time Series*. Cambridge: Cambridge University Press.

Novales, A. (1993): *Econometría*. 2ª edición. Madrid: Editorial McGraw-Hill.

Oficina de Justificación de la Difusión (2004), *Acta de Control 2003*, OJD, Madrid

Oficina de Justificación de la Difusión (2005), *Acta de Control 2004*, OJD, Madrid

Otero, J. M. (1993): *Econometría. Series Temporales y Predicción.* Madrid: Editorial AC.

Peña, D. (1987): "Sobre la interpretación de modelos ARIMA univariantes", *Trabajos de Estadística*, vol. 4, núm. 2, 19-45.

Pulido, A. (1993): *Modelos Econométricos*. 4ª edición. Madrid: Ediciones Pirámide.

Rynkiewicz, J.: (2004): Estimation of linear autoregressive models with Markov-switching , The EM algorithm revisited. Investigación Operacional, 25, 2. 188-173.

Uriel, E. (1985): *Análisis de series temporales: modelos ARIMA*, Paraninfo, Madrid.