

Bibliotecas: el coste de las suscripciones y aportaciones bayesianas

Agustín ALONSO RODRÍGUEZ
Real Centro Universitario
«Escorial-María Cristina»
San Lorenzo del Escorial

Resumen: Tomando el artículo de Bergstrom sobre los costes de las suscripciones a las revistas de economía por parte de las bibliotecas en Estados Unidos, como punto de partida, se presenta el modelo de Stock-Watson para estos datos, y se completa el modelo propuesto con el punto de vista bayesiano. El resultado final muestra las semejanzas entre la estimación mínimo cuadrática ordinaria y la bayesiana.

Abstract: Taking the article of Bergstrom on libraries subscriptions to economic journals in the US as an starting point, the paper presents the model of Stock-Watson for these data. The proposed model is then taken into consideration from a bayesian perspective. The final result shows the similarities of estimates in both methods: frequentist and bayesian.

Palabras clave: Bibliotecas, suscripciones, revistas de economía, Teorema de Bayes, estimación mínimo cuadrática, estimación bayesiana, distribución de probabilidad a priori, distribución de probabilidad a posteriori, MCMC, programa R, programa WinBUGS.

Key words: Libraries, journals subscriptions, economic journals, Bayes' Theorem, statistical estimation methods, prior probability distributions, posterior probability distributions, MCMC, R statistical program, WinBUGS program.

Sumario:

- I. Introducción.**
- II. Los costes de las suscripciones para las bibliotecas.**
- III. Razón de la situación.**
- IV. Un modelo para la demanda de suscripciones por parte de las bibliotecas.**
- V. La aportación bayesiana.**
- VI. Implementación bayesiana.**

VII. Detalles estadísticos.

VIII. Aplicación del programa WinBUGS al modelo de las suscripciones de las bibliotecas.

IX. Conclusión.

X. Bibliografía.

I. INTRODUCCIÓN

Theodore C. Bergstrom, en un artículo famoso y controvertido sobre las suscripciones a las revistas de economía por parte de las bibliotecas, animaba a un boicot de las revistas publicadas por editoriales comerciales, dado lo elevado de los costes.

El autor, comparando los precios de las suscripciones a las revistas publicadas por asociaciones profesionales y universidades con las publicadas por editoriales comerciales, consideraba injustificados los altos precios cargados por estas últimas, a los que tacha de abusivos.

II. LOS COSTES DE LAS SUSCRIPCIONES PARA LAS BIBLIOTECAS

El artículo publicado en 2001 presenta los resultados de comparar 180 revistas de carácter científico en el ámbito de la economía y en los Estados Unidos. Toma como referencias el precio por página, el precio por cita, el *ranking* de la revista por las citas y otras nueve variables más.

El artículo es de obligada lectura. Aquí señalaremos algunos de los resultados más notables.

En primer lugar, el autor divide las publicaciones en dos grandes grupos: publicaciones de asociaciones profesionales y de editoriales universitarias, publicaciones sin carácter de lucro, y revistas de editoriales comerciales.

Las seis revistas más citadas por el *Social Science Citation Index* son de carácter no lucrativo, con un coste medio de suscripción anual de 180 dólares. Si se consideran las veinte revistas más citadas, sólo cinco pertenecen al sector comercial, siendo el coste medio de estas cinco revistas de 1.660 dólares anuales.

Con las tarifas de suscripción del año 2001, Bergstrom elabora las dos tablas siguientes.

Tabla 1. Precios suscripciones y citas – Revistas sin ánimo de lucro

Título revista	Precio bibliotecas	Precio por página	Precio por cita	Precio por cita reciente	Ranking
AEA Journals*	\$140	\$0.03	\$0.01	\$0.12	1
Econometrica	\$214	\$0.14	\$0.03	\$0.93	2
J Political Ec	\$175	\$0.10	\$0.03	\$0.69	3
Quarterly J Ec	\$198	\$0.13	\$0.05	\$0.70	4
J Finance	\$207	\$0.07	\$0.05	\$0.63	5
J Consumer Res	\$99	\$0.23	\$0.04	\$0.90	6
Ec Journal	\$321	\$0.16	\$0.13	\$1.29	8
Rev Ec Studies	\$180	\$0.22	\$0.08	\$2.34	11
Rev Ec Statistics	\$200	\$0.29	\$0.09	\$1.15	12
Amer J Ag Ec	\$134	\$0.11	\$0.07	\$1.01	14

Notas:

1. AEA Journals: *The American Economic Review*, *Journal of Economic Perspectives* y *Journal of Economic Literature*. Los costes por página y cita se han calculado utilizando el total de páginas y citas de las tres revistas.
2. El precio por página se obtiene al dividir el precio de la suscripción por el número de páginas de la publicación en el año 2000.
3. El precio por cita es resultado de dividir el precio de la suscripción por el número de citas en 1998, recogido por el *Social Science Citation Index*.
4. El precio por cita reciente se obtiene al dividir el precio de la suscripción por el número de veces que los volúmenes de las revistas publicados en 1996 y 1997 son citados en 1998.
5. El *ranking* de la revista es resultado del número de veces que la revista fue citada en 1998.

Tabla 2. Precios suscripciones y citas – Revistas de editoriales comerciales

Titulo Revista	Precio bibliotecas	Precio por página	Precio por cita	Precio por cita reciente	Ranking
J Financial Ec	\$1429	\$0.73	\$0.53	\$7.85	7
J Ec Theory	\$1800	\$0.90	\$0.72	\$10.40	9
J Econometrics	\$2020	\$0.87	\$0.81	\$8.74	10
J Monetary Econ	\$1078	\$0.80	\$0.58	\$9.71	13
J Public Ec	\$1546	\$0.72	\$1.08	\$10.66	19
World Development	\$1548	\$1.35	\$1.10	\$7.04	20
European Ec Rev	\$1189	\$0.65	\$0.96	\$6.83	21
J Env Ec & Mgmt	\$650	\$1.02	\$0.56	\$3.90	22
J Health Ec	\$865	\$0.98	\$0.90	\$5.41	28
Ec Letters	\$1592	\$1.04	\$1.03	\$17.12	29

Notas, como en la Tabla 1.

Para completar la información, el autor hace referencia a los descuentos ofrecidos a los suscriptores individuales. Así, las diez más importantes revistas no-lucrativas ofrecen un precio medio de suscripción de 60 dólares al año. Mientras que las diez revistas más importantes de carácter comercial ofrecen suscripciones que van desde 85 a 1.187 dólares al año, con un promedio de 360 dólares.

Otro detalle que resalta Bergstrom es el referente al número de suscriptores. Las de carácter no-lucrativo tienen un gran número de suscriptores individuales. Por ejemplo, *Econometrica* tenía unos 2.900, suscriptores individuales y 2.400 suscriptores institucionales. La *Review of Economic Studies*, unos 850 individuales y 2.000 institucionales. En lo que respecta a las editoriales comerciales, las cifras de suscriptores se guardan como secretas.

En relación con las revistas de menor prestigio, las diferencias entre las revistas no-lucrativas y las comerciales tienen parecidas características.

Si pasamos a los costes de publicación, Bergstrom presenta como resumen, los datos de la tabla 3, referidos a una muestra de 297 publicaciones.

Tabla 3.- Estadísticas por el tipo de propietario

Tipo de editorial	Número de revistas	Coste total	Total páginas	Total citas	Precio por pagina	Precio por cita
No-lucrativas	91	\$11,644	66,304	75,330	\$0.15	\$0.15
Blackwell	46	\$11,807	23,574	6,335	\$0.50	\$1.86
Comerciales	160	\$100,381	113,646	40,402	\$0.88	\$2.48
Totales	297	\$123,832	203,524	122,067		

La presencia de Blackwell (actualmente absorbida por Wiley) se debe al hecho de que esta editorial publica revistas de su propiedad, pero también de asociaciones no-lucrativas.

A manera de resumen, las editoriales comerciales cargan alrededor de seis veces más por página, y 16 veces más por cita que las editoriales no-lucrativas.

Otra característica que analiza Bergstrom es la efectividad de los costes de las suscripciones. Mientras las editoriales no-lucrativas proporcionan la mayor parte de la información utilizada por los economistas, las editoriales comerciales absorben la mayor parte del presupuesto de las bibliotecas. En cifras, si una biblioteca se subscribiese a todas las revistas de economía no-lucrativas ello le supondría menos del 10% del presupuesto, permitiendo el acceso a más del 60% de los artículos citados en el ámbito de la economía. Sin embargo, la suscripción a las revistas publicadas por las editoriales comerciales, Blackwell excluida, consumiría el 80% del presupuesto, permitiendo el acceso a sólo la tercera parte de los artículos citados. Cf. la Tabla 4.

Tabla 4. Reparto de costes, páginas y citas

Tipo de editorial	Total costes	Total páginas	Total citas
No lucrativas	9%	33%	62%
Blackwell	10%	12%	5%
Comerciales	81%	56%	33%

Esta situación no es exclusiva de la disciplina económica, sino que es común con otras ramas de la ciencia.

III. RAZÓN DE LA SITUACIÓN

Hasta aquí los hechos en el año 2001. A continuación, el autor se pregunta cómo ha podido llegarse a esta situación, y dado que ahora aparece una varia-

ble que posteriormente entrará en el modelo econométrico, como es la edad de las revistas, se hace un breve resumen sobre el tema.

Ocho de las diez revistas más citadas de carácter no-lucrativo fueron fundadas antes de 1933, y nueve más, antes de 1945. Y por lo que respecta a ocho de las diez más citadas de las editoriales comerciales fueron fundadas entre 1969 y 1974, y el total de las diez, entre 1966 y 1982.

Esta evolución es resultado de la gran expansión de los estudios y de los profesionales de la economía, y de la incapacidad a aumentar sus posibilidades las editoriales no-lucrativas ante el incremento de la investigación publicable.

Al principio, las editoriales comerciales fijaron precios más competitivos que los actuales, que, como se ha indicado, absorbían en el 2001 el 80% del presupuesto de las bibliotecas. Para explicar tal situación Bergstrom acude a la *Teoría de juegos*, ejemplarizándola en la *Parábola de la reunión anual de los anarquistas*, que, a pesar de su anarquía, es decir de su falta de coordinación, se ven forzados a una situación que permite al dueño del hotel donde se celebra la reunión anual, alzar los precios sin que por ello se anule la reunión anual de los anarquistas.

Aleccionados por la parábola, las editoriales comerciales de las revistas económicas con éxito han descubierto que pueden alzar los precios de las suscripciones por encima de sus costes medios. Tal subida puede que reduzca el número de suscriptores, pero aumenta los beneficios, ya que el efecto final es mayor que la reducción de suscripciones, siendo ello posible ante la falta de coordinación de las bibliotecas e instituciones no-lucrativas.

Esta falta de coordinación en la Comunidad Académica es la que motiva al autor a proponer una serie de actuaciones conjuntas, que intenten reducir el 80% del presupuesto bibliotecario destinado a la adquisición de revistas que sólo aportan el 30% de las citas que se utilizan en el ámbito de la economía. El lector puede encontrar estas medidas en el artículo tantas veces citado del *Journal of Economic Perspectives*, accesible también en la página web del autor: <http://www.econ.ucsb.edu/~tedb/>

IV UN MODELO PARA LA DEMANDA DE SUSCRIPCIONES POR PARTE DE LAS BIBLIOTECAS

Las ideas y hallazgos de Bergstrom no han caído en el vacío, y buena prueba de ello es el hecho de su popularización, al ser objeto de estudio y ejemplo en textos de econometría, como Stock-Watson: *Introduction to Econometrics* (2007), segunda edición, y Kleiber-Zeilis: *Applied Econometrics with R* (2008).

Stock y Watson utilizan los datos del artículo de Bergstrom para establecer un modelo, inicialmente no lineal, para explicar la demanda de suscripciones a las revistas técnicas en economía.

Su razonamiento es directo y sencillo. Los economistas profesionales siguen de cerca la investigación en materia económica que aparece en las revistas de la especialidad, bien por suscripción individual o acudiendo a las bibliotecas. Desde el punto de vista de estas últimas, los autores se preguntan por la elasticidad de demanda de las bibliotecas respecto a las revistas en economía en los Estados Unidos.

Para contestar a esta pregunta, estudian la relación entre el número de suscripciones y los precios de las suscripciones. Al fin y al cabo, el producto final de las revistas técnicas son las ideas que contienen, cuantificando su precio no directamente pero sí indirectamente a partir del número de veces que tales ideas se citan por los especialistas en otras publicaciones. Dado el interés por estimar las elasticidades de la demanda, la especificación apropiada del modelo se hace en logaritmos. Es decir, el modelo inicialmente adoptado es:

$$\log(\text{subs}) = \beta_1 + \beta_2 \log(\text{preciocita}) + u$$

Con los datos de Bergstrom, las variables en consideración son:

- a) el número de suscripciones bibliotecarias: subs,
- b) el precio de la cita: precio de la suscripción dividido por el número de citas.

Para evitar el posible sesgo causado por la ausencia de variables, Stock-Watson incluyen tres variables de control, a saber, la edad de la revista y el número de caracteres anuales de la publicación, así como la variable de interacción resultante de estas últimas.

Tenemos así el modelo final propuesto:

$$\log(y) = \beta_1 + \beta_2 \log(x_2) + \beta_3 \log(x_3) + \beta_4 \log(x_2 \times x_3) + \beta_5 \log(x_4) + u \quad (1)$$

tras hacer

y: número de suscripciones bibliotecarias

x_2 : precio de las citas

x_3 : edad de la revista = 2000 - año de fundación de la revista

x_4 : número de caracteres anuales de la revista = caracteres por página multiplicado por número de páginas de la revista, dividiendo el total por 10^6 .

La formulación en (1) es la especificación lineal final a la que llegan Stock y Watson, que inicialmente consideraron para estos datos una formulación no-lineal.

Con ayuda del programa R, el modelo estimado, con algunos de sus estadísticos, es el siguiente:

```

Coefficients:
(Intercept)      3.43352    0.31486   10.905 < 2e-16 ***
x2              -0.89891    0.16151   -5.566 9.66e-08 ***
x3               0.37351    0.08897    4.198 4.28e-05 ***
I(x2 * x3)       0.14096    0.04483    3.144 0.00196 **
x4               0.22947    0.10513    2.183 0.03038 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6876 on 175 degrees of freedom
Multiple R-squared:  0.6339,    Adjusted R-squared:  0.6255
F-statistic: 75.75 on 4 and 175 DF,  p-value: < 2.2e-16

```

A la vista de estos resultados, Stock y Watson establecen las siguientes conclusiones:

- la demanda es menos elástica para las revistas más antiguas que para las de fundación reciente;
- la demanda es mayor para las revistas con mayor número de caracteres, permaneciendo constantes el precio y la edad;
- la evidencia remite a una especificación lineal, quedando descartada la no-lineal inicialmente asumida, por presentar coeficientes estadísticamente no-significativos (Cf. STOCK-WATSON, *Introduction to Econometrics*, segunda edición, p. 290.)

Para posterior comparación, el intervalo del 95% de confianza para los coeficientes del modelo viene dado por

	2.5 %	97.5 %
(Intercept)	2.81211505	4.0549264
x2	-1.21766428	-0.5801552
x3	0.19791431	0.5491154
I(x2 * x3)	0.05247246	0.2294457
x4	0.02198764	0.4369445

Tenemos así que la demanda de suscripciones depende de la edad de la revista, lo que se puede interpretar afirmando que las bibliotecas muestran mayor preferencia por las revistas consolidadas, y que la demanda es muy

inelástica con relación al precio, especialmente para las revistas más antiguas (para una revista de 80 años de existencia la elasticidad es -0.28 , con un error estándar de $0,06$, y para la demanda de una revista con cinco años de existencia, la elasticidad pasa a ser -0.67 , con un error estándar de $0,08$). El disponer de la investigación más reciente es, por tanto, para las bibliotecas, una necesidad no un lujo. Stock y Watson terminan su análisis relacionando esta inelasticidad, o adición de las bibliotecas, con la adición a los cigarrillos, que varía entre $-0,3$ y $-0,5$, si bien, como ellos señalan, la adición de las bibliotecas no daña a la salud. Cf. ID., *ibíd.*

V. LA APORTACIÓN BAYESIANA

El enfoque bayesiano también arranca con un modelo para los datos observados, $y = (y_1, y_2, \dots, y_n)$, dado un vector de parámetros desconocido θ . El modelo en formato funcional se da como una distribución de probabilidad $f(y; \theta)$. Al poner el énfasis en el vector de parámetros tenemos la función de verosimilitud, escrita frecuentemente como $L(\theta; y)$, para resaltar el cambio de roles de θ e y .

Puede ocurrir que L no sea una función de distribución de probabilidad de θ , dados los valores de y , es decir, que $\int L(\theta; y) d\theta$ puede no ser igual a la unidad, e incluso puede que no sea finita. No obstante, dados los valores concretos de y , es posible encontrar un $\hat{\theta}$ que maximice la función de verosimilitud. Es decir, $\hat{\theta} = \arg \max_{\theta} L(\theta; y)$. Tal valor $\hat{\theta}$ es la estimación máximoverosimil de θ . Estas ideas se remontan a Ronald A. Fisher, el inventor del método, ya en 1922.

En el enfoque bayesiano, en lugar de considerar el vector de parámetros θ desconocido, pero de valor constante, se considera como una magnitud aleatoria, es decir, como una variable aleatoria.

La forma de hacer operativa esta idea consiste en atribuir a θ una distribución de probabilidad, que resume en sí la información que se conozca al margen de la aportada por los datos muestrales: y . Esta distribución de probabilidad recibe el nombre técnico de distribución de probabilidad *a priori*.

Como la función de verosimilitud tiene parámetros θ , también la distribución de probabilidad *a priori* tiene sus parámetros, η , conocidos como hiperparámetros, en vista a diferenciarlos de los parámetros θ de la función de verosimilitud.

Para empezar, supongamos conocidos los hiperparámetros. Ello nos permite escribir la función de probabilidad *a priori* como $\pi(\theta) = \pi(\theta | \eta)$.

En estas circunstancias, la inferencia respecto a θ queda recogida en la distribución de probabilidad *a posteriori*

$$\begin{aligned}
 p(\theta | y) &= \frac{p(y, \theta)}{p(y)} = \frac{p(y; \theta)}{\int p(y, \theta) d\theta} = \\
 &= \frac{f(y | \theta) \pi(\theta)}{\int f(y | \theta) \pi(\theta) d\theta} = \frac{p(y | \theta) \pi(\theta)}{p(y)} \quad (2)
 \end{aligned}$$

fórmula que se conoce como el Teorema de Bayes (1763).

Conviene destacar que tanto la contribución de los datos, mediante la función de verosimilitud f , como la opinión *a priori*, presente en π son los componentes de la función de probabilidad *a posteriori* en (2). De hecho, la distribución de probabilidad *a posteriori* es simplemente el producto de la función de verosimilitud por la distribución de probabilidad *a priori*, normalizando el producto para que la integral sea igual a la unidad, y así ser una distribución de probabilidad válida.

Llegados aquí, quizás sea apropiado presentar la versión discreta de la formulación en (2). Sea A el evento de interés, y $B_j : j = 1, 2, \dots, J$ una colección de eventos disjuntos o mutuamente excluyentes. Dadas las probabilidades $P(B_j)$ y la probabilidad condicionada $P(A | B_j)$, podemos escribir

$$P(B_j | A) = \frac{P(A, B_j)}{P(A)} = \frac{P(A, B_j)}{\sum_{j=1}^{j=J} P(A, B_j)} = \frac{P(A | B_j) P(B_j)}{\sum_{j=1}^{j=J} P(A | B_j) P(B_j)} \quad (3)$$

siendo $P(A, B_j)$ la probabilidad conjunta de que A y B_j ocurran, también escrita como $P(A \cap B_j)$. (Alonso Rodríguez, 2008). Ésta es la formulación del Teorema de Bayes en los textos introductorios de estadística.

Al comparar (2) con (3) se puede observar la correspondencia de las expresiones al hacer que B_j juegue el papel de θ y A el papel del vector de datos y .

VI. IMPLEMENTACIÓN BAYESIANA

Una forma concreta de evaluar la distribución de probabilidad *a posteriori* de los parámetros es acudiendo al programa WinBUGS. Las cuatro letras de BUGS hacen referencia a *Bayesian inference Using Gibbs Sampling* (Spiegelhalter, Thomas, Best y Lunn, 2003). No es la única adaptación del método Gibbs para la obtención de muestras. JAGS: *Jet Another Gibbs Sampler* (Plummer, 2003) es otra implementación del método.

WinBUGS utiliza el procedimiento Gibbs (Geman y Geman, 1984; Gelfand y Smith, 1990, Casella y George, 1992) y el algoritmo Metropolis

(Metropolis, Rosenbluth, Rosenbluth, Teller y Teller, 1953) para generar cadenas de Markov, al muestrear distribuciones condicionadas de probabilidad.

Una nueva y más flexible revisión de WinBUGS es OpenBUGS (Spiegelhalter, Thomas, Best y Lunn, 2004). Para R, la versión adaptada es R2WinBUGS (Sturz, Ligges y Gelman). Basado en OpenBUGS está el programa BRugs. Con el tiempo se espera fundir ambos programas, BRugs y R2WinBUGS en un solo paquete para R.

WinBUGS es un programa gratuito desarrollado por especialistas en probabilidad y estadística en la Unidad de Investigación del Consejo de Bioestadística en la Universidad de Cambridge, Reino Unido.

El programa WinBUGS utiliza una sintaxis muy parecida a la utilizada por el programa R. Es más, como se ha indicado antes, existen dos adaptaciones a R del programa WinBUGS: BRugs y R2WinBUGS.

WinBUGS utiliza una forma concreta del algoritmo Monte Carlo conocida como *Markov Chain Monte Carlo*, abreviadamente MCMC. En esencia, WinBUGS genera muestras aleatorias secuenciales de las distribuciones de cada uno de los parámetros del modelo, condicionadas a los parámetros de otro modelo y a los datos muestrales en consideración.

La utilización del programa requiere le especificación de tres elementos:

- un modelo, en la sintaxis de WinBUGS,
- los datos muestrales, y
- la especificación de los valores iniciales para los parámetros.

Tanto WinBUGS como JAGS generan resultados en formato MCMC, que permiten ser analizados, con gran lujo de detalles, por otros dos paquetes: CODA y BOA, también adaptados a R. Los nombres de estos paquetes hacen referencia al análisis de los resultados en formato MCMC: BOA, *Bayesian Output Analysis* (Smith, 2004) y CODA: *Output analysis and diagnostics for MCMC* (Plummer, Best, Cowles y Vines, 2004).

La utilización de BRugs y R2WinBUGS suponen el tener que formular el modelo a utilizar en el formato establecido por WinBUGS, y el establecimiento de los otros dos pasos antes mencionados. Para facilitar aún más la aplicación de la visión bayesiana, existen otros dos programas de carácter general, que liberan al analista de tener que especificar los anteriores pasos. Son *MCMCpack* (Martin, Quinn y Park, 2000) y *Zelig: Everyone's Statistical Software* (Imai, King y Lau, 2000). Ambos paquetes requieren tan sólo el formular los modelos a simular en el lenguaje básico de la regresión establecido por R.

VII. DETALLES ESTADÍSTICOS

Si volvemos a la formulación en (2), cabe escribirla también como:

$$p(\theta | y) \propto f(y | \theta)\pi(\theta) \quad (4)$$

que podemos enunciar en palabras diciendo que la distribución de probabilidad *a posteriori* de los parámetros es proporcional a la función de verosimilitud multiplicada por la probabilidad *a priori* de los parámetros.

Para aplicar el programa R2WinBUGS es necesario formular el modelo en el formato establecido por WinBUGS. En esta sección presentamos los fundamentos para poder hacerlo.

El modelo lineal general clásico en formato matricial se escribe como:

$$y = X\beta + u \quad (5)$$

siendo:

y , de dimensiones $n \times 1$, el vector de observaciones
 X , de dimensiones $n \times k$, la matriz de regresores o variables explicativas, con la primera columna igual a la unidad;
 β , de dimensiones $k \times 1$, el vector de parámetros; y
 u , de dimensiones $n \times 1$, el vector término de error del modelo.

En el enfoque bayesiano, tanto β como u , son variables aleatorias, que, por tanto, requieren una distribución de probabilidad *a priori*.

En el modelo lineal general (5), los supuestos clásicos permiten escribir, $E(u | X) = 0$ lo que nos lleva a que $E(y | X) = X\beta$.

La distribución *a posteriori* de β , tras obtener y y X , es $p(\beta | y, X)$, y en virtud del teorema de Bayes pasa a ser:

$$p(\beta | y, X) \propto p(y, X | \beta)p(\beta)$$

expresión que requiere conocer la distribución de probabilidad conjunta de y y X , dados β : $p(y, X | \beta)$, y la probabilidad de β : $p(\beta)$.

Como $y = X\beta + u$, si logramos establecer la probabilidad conjunta X y u , dados β , entonces podremos obtener la de y y X . Ahora bien, $p(u | X)$ requiere el cumplimiento de $E(u | X) = 0$, y en el contexto bayesiano tenemos que pasar del supuesto de la media condicionada anterior a una distribución de probabilidad completa. La forma más directa de establecerla requiere exigir que u y X sean independientes, y no simplemente incorrelacionados, es lo que implica $E(u | X) = 0$.

En estas condiciones, podemos escribir:

$$p(u, X) = p(u)p(X)$$

Si intercambiamos las variables u y X por y y X , utilizando (5), llegamos a

$$p(y, X | \beta) = p(y - X\beta)p(X | \beta)$$

Se evidencia así que la función de probabilidad para los errores del modelo, $p(u)$, depende de la naturaleza de la variable y . Si es continua o aproximadamente continua y toma valores en la línea recta, una distribución inicialmente apropiada es la distribución normal. Si y es discreta o binaria o no-negativa otras posibilidades existen. Cabe incluso la transformación de los valores de y para acercarlos a los valores de la distribución normal.

Comencemos, pues, suponiendo que las u tienen distribución normal, con sus valores independientes e idénticamente distribuidos con media cero y varianza constante. En el enfoque bayesiano, en lugar de la varianza se utiliza el recíproco de la varianza, con el nombre *precisión*, y símbolo *tau*:

$\tau = \frac{1}{\sigma^2}$. *Tau* es un hiperparámetro.

Tenemos así la distribución de probabilidad *a priori*

$$p(u | \tau) \propto \tau^{n/2} \exp\{-(\tau/2)u'u\}, \text{ siendo } u'u = \sum_{i=1}^n u_i^2.$$

De aquí se sigue:

$$p(y, X | \beta, \tau) \propto \tau^{n/2} \exp\{-(\tau/2)(y - X\beta)'(y - X\beta)\} p(X | \beta) \quad (6)$$

Si la distribución de X , $p(X | \beta)$ no depende de β ó τ , entonces X es estrictamente exógena, y en este supuesto $p(X)$ puede ser absorbido como factor de proporcionalidad en la anterior expresión, y puede ser eliminado, llegando a

$$p(y, X | \beta, \tau) = p(y | X, \beta, \tau) \propto \tau^{n/2} \exp\{-(\tau/2)(y - X\beta)'(y - X\beta)\} \quad (7)$$

tras eliminar $p(X)$ que no depende de β ni de τ .

Bajo estos supuestos de normalidad, independencia y homoscedasticidad del modelo lineal general clásico, tenemos, las dos formas siguientes de representar el modelo lineal general:

$$y = X\beta + u; \quad u | X, \beta, \tau \sim N(0, \tau I_n) \quad (8)$$

o también

$$Y | X, \beta, \tau \sim N(X\beta, \tau I_n) \quad (9)$$

que expresan el hecho de que dados X, β y τ , la variable aleatoria Y tiene una distribución multinormal, con media $X\beta$ y matriz de varianzas-covarianzas $(1/\tau)I_n$.

En esta segunda manera de presentar el modelo lineal, (9), el término de error u es irrelevante, siendo innecesario nombrarlo. Sin embargo, hay que decir que en la econometría tradicional, la formulación preferida es la primera (8).

Distribuciones *a priori* de β y τ .

Una vez establecida la distribución del término de error u , pasemos a considerar las distribuciones de los parámetros.

Para dar mayor relevancia a los datos y menos a las opiniones del investigador, se opta por distribuciones *a priori* para β y τ que se denominan *vagas* o no-informativas.

De entre las posibles, hay una particularmente interesante por un doble motivo. Primero porque va a permitir la inferencia *a posteriori* respecto a β y τ , que al menos numéricamente, coincide con los resultados de la econometría tradicional. Y segundo, esta distribución *a priori* permite derivaciones analíticas exactas para la obtención de las distribuciones marginales *a posteriori* de β y τ . Nos referimos a la *distribución uniforme*, formulada como:

$$p(\beta, \tau) \propto \frac{1}{\tau}; -\infty < \beta < +\infty; \tau > 0 \quad (10)$$

Por el teorema de Bayes, al multiplicar la función de verosimilitud (7) por la distribución *a priori* (10) obtenemos la distribución de probabilidad *a posteriori* de los $(k + 1)$ parámetros β y τ , considerados como variables aleatorias. Es decir:

$$p(\beta, \tau | y, X) \propto \tau^{n/2-1} \exp\{-(\tau/2)(y - X\beta)'(y - X\beta)\} \quad (11)$$

o equivalentemente

$$p(\beta, \tau | y, X) \propto \tau^{n/2-1} \exp\{-(\tau/2)(\beta - \hat{\beta})'X'X(\beta - \hat{\beta})\} \exp\{-(\tau/2)\hat{u}'\hat{u}\} \quad (12)$$

siendo $\hat{\beta} = (X'X)^{-1}X'y$, y $\hat{u} = y - X\hat{\beta}$, los estimadores de mínimos cuadrados ordinarios.

El vector de residuos es de importancia crucial a la hora de comprobar si los supuestos del modelo lineal general son consistentes con las observaciones. Y se puede demostrar que la distribución marginal de τ , al integrar respecto a β , en (12) es la *distribución gamma*. Es decir:

$$p(\tau | y, X) \propto \tau^{\nu/2-1} \exp\{-\tau \nu s^2 / 2\} \quad (13)$$

siendo $\nu = n - k$ los grados de libertad, y $s^2 = \hat{u}'\hat{u} / \nu$ la varianza residual.

VIII. APLICACIÓN DEL PROGRAMA WINBUGS AL MODELO DE LAS SUSCRIPCIONES DE LAS BIBLIOTECAS

Como se ha indicado antes, para aplicar el programa WinBUGS se requiere el establecimiento de un modelo, los datos muestrales y el establecimiento de los valores iniciales de los parámetros.

El modelo, dadas las explicaciones anteriores, y el modelo establecido como modelo final por Stock-Watson, permite la siguiente formulación, en la sintaxis establecida por WinBUGS, cf. LANCASTER, T., (2004), p. 375.

```
model{
  for (i in 1:N){
    y[i] ~ dnorm(mu[i], tau)
    mu[i] <- b1 + b2*x2[i] + b3*x3[i] + b4*x4[i] + b5*x5[i]
  }
  b1 ~ dnorm(0, 1.0E-6)
  b2 ~ dnorm(0, 1.0E-6)
  b3 ~ dnorm(0, 1.0E-6)
  b4 ~ dnorm(0, 1.0E-6)
  b5 ~ dnorm(0, 1.0E-6)
  tau ~ dgamma(1.0E-3, 1.0E-3)
}
```

N. B. La variable de interacción $x_2 \times x_3$ pasa a ser aquí la variable x_4 , y la anterior variable x_4 , pasa a ser ahora la variable x_5 .

Sobre este modelo como base, vamos a generar tres Cadenas de Markov, de 50.000 muestras o valores muestrales, y valores iniciales diferentes para cada una, sobre las que el programa WinBUGS establecerá la distribución de probabilidad *a posteriori* para los parámetros.

Los valores iniciales, en el formato de WinBUGS, quedan almacenados con el nombre *inits*, y son los siguientes:

```
inits=
list(list(b1=0,b2=0,b3=0,b4=0,b5=0,tau=1),list(b1=3.0,b2=0.8,
b3=0.3,b4=0.1,b5=0.2,tau=2.1),list(b1=1.0,b2=1.0,b3=0,b4=0,
b5=0,tau=2))
```

N. B. El segundo conjunto de valores iniciales se inspira en los valores estimados por mínimos cuadrados ordinarios, en el modelo de Stock-Watson.

Invocamos por primera vez el programa WinBUGS mediante la instrucción:

```
simulación1 =
bugs(data,inits,parameters,model.file="biblio.txt",n.chains=3,n.
iter=50000,          digits=5,bugs.directory="c:/Archivos      de
programa/WinBUGS14/",debug=TRUE)
```

Los resultados quedan guardados como `simulación1`. Para verlos, basta con invocarlos:

```
simulación1
```

y los resultados de la simulación, es decir, la distribución de probabilidad *a posteriori* de los parámetros arroja los siguientes resultados

```
Inference for Bugs model at "biblio.txt", fit using WinBUGS,
3 chains, each with 50000 iterations (first 25000 discarded),
n.thin = 75
n.sims = 1002 iterations saved
```

	mean	sd	2.5%	25%	50%	75%	97.5%	Rhat	n.eff
b1	3.4	0.3	2.8	3.2	3.4	3.7	4.0	1	1000
b2	-0.9	0.2	-1.2	-1.0	-0.9	-0.8	-0.6	1	1000
b3	0.4	0.1	0.2	0.3	0.4	0.4	0.6	1	660
b4	0.1	0.0	0.1	0.1	0.1	0.2	0.2	1	1000
b5	0.2	0.1	0.0	0.2	0.2	0.3	0.4	1	1000
tau	2.1	0.2	1.7	2.0	2.1	2.3	2.6	1	850
deviance	377.0	3.5	372.2	374.5	376.4	378.8	385.6	1	1000

For each parameter, n.eff is a crude measure of effective sample size, and Rhat is the potential scale reduction factor (at convergence, Rhat=1).

DIC info (using the rule, $pD = \bar{D} - \hat{D}$)

$pD = 6.0$ and $DIC = 383.1$

DIC is an estimate of expected predictive error (lower deviance is better).

De esta tabla, que por razones de espacio sólo brinda un decimal para el valor esperado o medio de la distribución de cada uno de los parámetros, la primera columna (luego se verán más decimales), hay que resaltar el valor del estadístico *Rhat* de Gelman y Rubin (1992); su valor, igual a la unidad, indica que las cadenas se entrecruzan y convergen correctamente a la distribución de probabilidad *a posteriori* de cada parámetro.

Para estudiar mejor los resultados, acudimos al programa *CODA*. Ahora, debido a la instrucción `codaPkg=TRUE`, los resultados de la simulación se guardarán para luego poderlos mostrar con más detalle.

La nueva invocación de WinBUGS es la siguiente:

```
simulación2 =
bugs(data, inits, parameters, model, file="biblio.txt", digits=5, coda
Pkg=T, n.chains=3, n.iter=50000, bugs.directory="c:/Archivos de
programa/WinBUGS14/", debug=TRUE)
```

Al invocar `simulación2`, el programa muestra los tres ficheros en los que han quedado almacenados los resultados de cada cadena

```
simulación2 =
[1] "g:/Anuario09B/coda1.txt" "g:/Anuario09B/coda2.txt"
[3] "g:/Anuario09B/coda3.txt"
```

Para ver estos resultados, los leemos y guardamos como simulaciones, mediante la instrucción:

```
simulaciones = read.bugs(simulación2)
```

y con

```
summary(simulaciones)
```

obtenemos:

1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
b1	3.4298	0.32381	0.010229	0.009951
b2	-0.9042	0.16014	0.005059	0.004538
b3	0.3747	0.09162	0.002894	0.002694
b4	0.1424	0.04412	0.001394	0.001284
b5	0.2271	0.10729	0.003389	0.003524
deviance	377.0137	3.52112	0.111236	0.113146
tau	2.1188	0.21511	0.006796	0.006253

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
b1	2.79008	3.2035	3.4290	3.6658	4.0270
b2	-1.20795	-1.0190	-0.9096	-0.7947	-0.5951
b3	0.20495	0.3107	0.3751	0.4373	0.5509
b4	0.06051	0.1118	0.1428	0.1733	0.2287
b5	0.02153	0.1518	0.2310	0.3015	0.4405
deviance	372.20250	374.5000	376.4000	378.8000	385.5875
tau	1.72402	1.9760	2.1105	2.2618	2.5550

A la vista de estos resultados, y al compararlos con los obtenidos al estimar el modelo de Stock-Watson por mínimos cuadrados ordinarios, ver columna **Mean** y columna **Estimate** de ambos resúmenes (**summary**), vemos la coincidencia casi absoluta de los valores, con la salvedad de que en el enfoque bayesiano hemos obtenido distribuciones de probabilidad para cada parámetro, en lugar del único valor dado por la estimación mínimo cuadrática.

Ya que anteriormente calculamos el intervalo de confianza para los coeficientes estimados del modelo de Stock-Watson, podemos comparar sus valores con los percentiles dados en el **summary(simulaciones)**, y también las coincidencias son notables.

Si nos detenemos en el valor de τ , cuya media es 2.1188, recordando la relación $\hat{\tau} = \frac{1}{\hat{\sigma}} = 2.1188$, podemos escribir $\hat{\sigma} = \frac{1}{\hat{\tau}} = 0.47197$, es decir $\hat{\sigma} = 0.6870$. En la estimación de este parámetro por mínimos cuadrados, en el output de R, aparece $\hat{\sigma} = 0.6876$, **Residual standard error**.

Con el programa CODA varias son las posibilidades de análisis que ahora están disponibles, y con gráficos en alta resolución. Por ejemplo:

qqmath(simulaciones), que muestra un gráfico *qqnorm*, para comprobar visualmente la normalidad de los resultados;

cumuplot(simulaciones), para ver las trayectorias seguidas por las simulaciones;

geweke.diag(simulaciones), para aplicar el contraste de Geweke;

heidel.diag(simulaciones), para aplicar el contraste de Heidelberg-Welch;

gelman.diag(simulaciones), para aplicar este contraste;

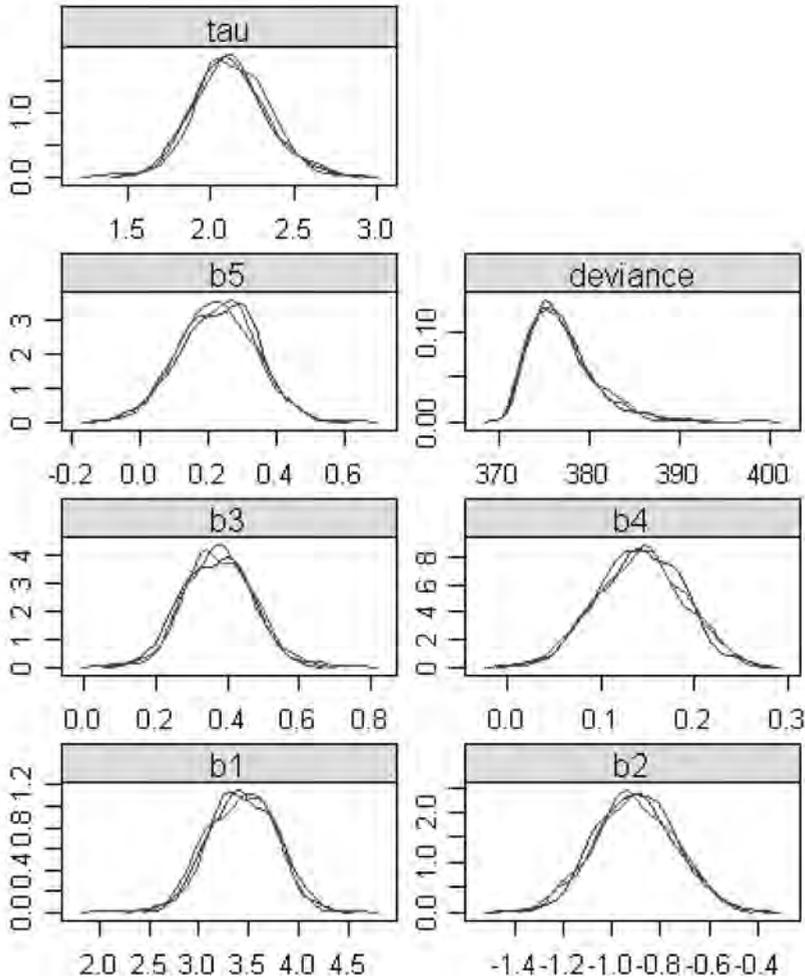
plot(simulaciones), para obtener una visión gráfica global de los resultados; etc.

Aquí nos limitaremos a ver gráficamente las distribuciones de los parámetros; la instrucción

densityplot(simulaciones)

genera las representaciones de la figura 1.

FIGURA 1



La figura 1 muestra la buena mezcla de las 3 cadenas, y su buena adecuación a la distribución normal, en lo que se refiere a los parámetros b_1, b_2, b_3, b_4, b_5 y a la distribución gamma en lo referente a τ .

IX. CONCLUSIÓN

Además de hacernos eco del artículo de T. Bergstrom, y del modelo desarrollado por Stock y Watson, en este trabajo hemos querido mostrar cómo el

enfoque bayesiano permite llegar a los resultados del enfoque “frecuentista”, tradicional en econometría.

Aunque el establecimiento de probabilidades *a priori* denominadas *no-informativas*, permite el acercamiento entre ambos enfoques, con todo no es lo más deseable desde el punto de vista bayesiano, que valora grandemente la opinión de los especialistas en cuanto conocedores de sus ámbitos de especialización.

Dados los avances en el ámbito de la informática y la estadística, sería de desear una mayor aceptación del enfoque bayesiano en econometría, ya que permite abrir un campo de posibilidades de grandes consecuencias, al establecer la forma de *revisar* las opiniones de los expertos, traducidas en las probabilidades *a priori*, un hecho de gran importancia en economía.

X. BIBLIOGRAFÍA

- ALBERT, J., *Bayesian Computation with R*, Springer, New York 2007.
- ALONSO RODRÍGUEZ, A., «El resurgir de Thomas Bayes», en *Anuario Jurídico y Económico Escurialense* 41 (2008) 327-360.
- BERGSTROM, T. C., «Free Labour for Costly Journals?», en *Journal of Economic Perspectives*, 15, 4 (2001) 183-198. También en: <http://www.econ.ucsb.edu/~tedb/>
- CARLIN, B. P., y LOUIS, T. A., *Bayesian Methods for Data Analysis*, 3.^a ed., CRC Press, Boca Ratón, FL. 2009.
- CASELLA, G., y GEORGE, G., «Explaining the Gibbs Sampler», en *American Statistician*, 46 (1992) 167-174.
- GELFAND, A. E., y SMITH, A. F. M., «Sampling-based Approaches to Calculating Marginal Densities», *Journal of the American Statistical Association* 85 (1990) 398-409.
- GELMAN, A., y RUBIN, D. B., «Inference from iterative simulations using multiple sequences (with discussion)», en *Statistical Science* 7 (1992) 457-511.
- GELMAN, A., y HILL, J., *Data Analysis using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, New York 2007.
- GEMAN, S., y GEMAN, D., «Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images», en *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6 (1984) 721-741.
- GILKS, W. R.; RICHARDSON, S., y SPIEGELHALTER, D. J., *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC, Boca Ratón, Florida 1996.
- GREENBERG, E., *Introduction to Bayesian Econometrics*, Cambridge University Press, New York 2008.
- IMAI, K.; KING, G., y LAU, O., «Zelig: Everyone’s Statistical Software, R version 3.3-1, 2008», en *CRAN: The Comprehensive R Archive Network*, URL:<http://cran.r-project.org/web/packages/Zelig/>, Viena.
- KLEIBER, C., y ZEILEIS, C., *Applied Econometrics with R*, Springer, New York 2008.
- KOOP, G., *Bayesian Econometrics*, John Wiley, Chichester, UK 2003.
- LANCASTER, T., *An Introduction to Modern Bayesian Econometrics*, Blackwell Publishing, Oxford 2004.

- LYNCH, S. M., *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*, Springer, New York 2007.
- MARTIN, A.D.; QUINN, K. M., y PARK, J. H., *MCMCpack Package, R version 0.9-5, 2008*, en *CRAN: The Comprehensive R Archive Network*,
URL:<http://cran.r-project.org/web/packages/MCMCpack/>, Viena.
- METROPOLIS, N.; ROSENBLUTH, A.; ROSENBLUTH, M.; TELLER, H., y TELLER E., «Equation of State Calculations by Fast Computing Machines», *Journal of Chemical Physics* 21 (1953) 1087-1092.
- PLUMMER, A., *JAGS: Jet Another Gibbs Sampler, version 1.0.3, 2008*,
URL: <http://calvin.iarc.fr/~martyn/software/jags/>
- PLUMMER, M.; BEST, N.; COWLES, K., y VINES, K., 2004, «CODA: Output Analysis and Diagnostics for MCMC, R package version 0.13-3, 2008», en *CRAN: The Comprehensive R Archive Network*,
URL:<http://cran.r-project.org/web/packages/coda/>, Viena.
- R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Viena, Austria, 2008. URL <http://www.r-project.org>.
- SMITH, B. J., «BOA: Bayesian Output Analysis program for MCMC, 2004. R package, version 1.1.7-2, 2008», en *CRAN: The Comprehensive R Archive Network*,
URL:<http://cran.r-project.org/web/packages/boa/>, Viena.
- SPIEGELHALTER, D. J.; THOMAS, A.; BEST, N.G., y LUNN, D., *WinBUGS Version 1.4 Users Manual*, MRC Biostatistics Unit, Cambridge 2003.
– *WinBUGS Version 2.0 Users Manual*, MRC Biostatistics Unit, Cambridge 2004,
URL: <http://mathstat.helsinki.fi/openbugs/>
- STOCK, J. H., y WATSON, M. W., *Introduction to Econometrics*, 2.^a ed., Pearson-Addison Wesley, Boston 2007.
- STURZ, S.; LIGGES, U.; GELMAN A.; GORJANC, G., y KERMAN, J., «R2WinBUGS, R version 2.1-8, 2008», en *CRAN: The Comprehensive R Archive Network*,
URL:<http://cran.r-project.org/web/packages/R2WinBUGS/>, Viena.
- SU, Y-S., y YAJIMA, M., «R2JAGS, version 0.01-13, 2008», en *CRAN: The Comprehensive R Archive Network*,
URL:<http://cran.r-project.org/web/packages/R2JAGS/>, Viena.