ESTUDIOS DE ECONOMÍA APLICADA VOL. 22 – 3,2004 ART. 22309

Desagregación espacial para pequeñas áreas. Un modelo bayesiano normal-gamma.

ROJO GARCÍA, J.L. Y SANZ GÓMEZ, J.A.
Universidad de Valladolid

RESUMEN

La desagregación de magnitudes económicas de grandes áreas en sus valores para áreas menores (por ejemplo, NUTS III en la UE) en las que la insuficiencia estadística invalida los procedimientos contables, es un problema relevante en la Estadística aplicada. En este trabajo se propone para su resolución un modelo bayesiano, con verosimilitud normal.

A diferencia de propuestas anteriores de los autores, en que se utilizaban distribuciones de Laplace (primera ley del error) para la verosimilitud, este trabajo emplea distribuciones de la familia normal-gamma, lo que permite obtener soluciones explícitas sin recurrir al muestreo de Gibbs. Asimismo, se utilizan distribuciones a priori informativas, en la línea de las propuestas de las g-priors de Zellner, en lugar de distribuciones no informativas.

Los resultados se ilustran con una aplicación a la Contabilidad Regional de España, concretándola en un reparto provincial para Castilla y León. La calidad del procedimiento se evalúa mediante procedimientos de simulación, en ausencia de valores observables para las magnitudes.

Palabras clave: Análisis bayesiano, Desagregación espacial, Simulación Monte-Carlo

Spatial Dissagregation for Small Areas. A Bayesian Normal-Gamma Model

ABSTRACT

Spatial economic dissagregation for small areas (districts, counties or provinces (NUTS III administrative units of the EU)) is a relevant item in Applied Statistics. This paper proposes a Bayesian approach, using a normal likelihood.

In a previous work, authors used a Laplace (first law of error distribution) likelihood in producing the small area estimates, but the actually used normal-gamma family allows to obtain explicit solutions without the application of the Gibbs sampling method. Moreover, we make use of informative priors, in the line of the Zellner's g-priors, instead of non-informative distributions.

We conduct a limited study relative to the Regional Accounts of Spain, showing a provincial disaggregation of the annual added-value for 'Castilla y León'. Additionally the performance of the procedure is evaluated by means of Monte-Carlo simulation.

Keywords: Bayesian Analysis, Small areas, Monte-Carlo simulation
JEL Classification: R12, C11, C13, C15

Artículo recibido en abril de 2004 y aceptado en octubre de 2004
La referencia electrónica de este artículo en la página www.revista-eea.net, es e-22309

ISSN 1697-5731 (online) – 1133-3197 (print)

1. INTRODUCCIÓN

Este artículo construye un método que permite desagregar magnitudes correspondientes a zonas geográficas agregadas entre áreas más pequeñas, con la ayuda de indicadores, cuando no se dispone de posibilidades de desagregación por procedimientos contables. Habitualmente, la denominación "*área pequeña*" (*small area*, en su expresión inglesa, más habitual) se emplea en los procedimientos ligados a la realización de encuestas por muestreo, cuando el número de observaciones muestrales correspondientes a un cierto dominio geográfico o sociológico es escaso o incluso es igual a cero. Resulta imposible en esos casos estimar magnitudes para tales dominios o áreas (y, aunque sea posible, los errores son demasiado grandes), por lo que se han desarrollado procedimientos complementarios.

Las aplicaciones del procedimiento obtenido en este trabajo son múltiples, y, como más adelante veremos, no se orientarán a resolver los problemas de insuficiencia muestral en encuestas, aunque nuestra propuesta bien pudiera utilizarse con ese objetivo. En concreto, los autores lo empleamos para provincializar datos económicos regionales, de acuerdo con los valores de ciertos indicadores provinciales representativos, distribuyendo la Contabilidad Regional del INE entre las 9 provincias de Castilla y León. Con ello, conseguimos ampliar al año 2002 los actuales datos provinciales de macromagnitudes proporcionados por dicha Contabilidad.

De hecho, es esta la motivación inicial del trabajo propuesto, esto es, salvar el retardo temporal (de un año) en las estimaciones provinciales frente a las regionales en un contexto, dentro del ámbito de la Economía Aplicada, que exige valoraciones más actualizadas de las macromagnitudes provinciales para el análisis económico de las Comunidades autónomas españolas.

Asimismo, los autores lo empleamos para realizar desagregaciones subsectoriales de una magnitud sectorial agregada, posibilidad que, aunque no se presenta en este trabajo, permite completar la Contabilidad Regional de 2002 para Castilla y León (actualmente desagregada a 6 ramas), de manera que presente la misma desagregación sectorial que la proporcionada por el INE para 2001 (a 30 ramas de actividad). Esta Comunidad autónoma es un paradigma de lo que la expresión *área pequeña* sugiere, en el sentido de que una (relativamente) pequeña economía, como es la castellano-leonesa, resulta desagregada a 30 ramas de actividad con nuestro método.

Como más adelante señalaremos, el objetivo no es el de replicar la elaboración de la Contabilidad Regional o Provincial, ya que los indicadores utilizados por el I.N.E. son sólo parcialmente conocidos, sino avanzar en una estimación alternativa de mayor utilidad en el análisis de coyuntura, por su menor desfase temporal. Por otro lado, no parece metodológicamente adecuado valorar un método de estimación con los resultados de otro, por lo que la validación que realizaremos se apoya en un ejercicio de simulación.

En otro trabajo (Rojo y Sanz, 2002) los autores proponen utilizar un procedimiento relacionado con el actual para resolver procedimientos de congruencia que combinen métodos de estimación bottom-up con otros top-down, esto es, para ajustar predicciones correspondientes a varias pequeñas áreas (bottom-up) con predicciones globales obtenidas para el área agregada (top-down). En estos casos los indicadores son aproximaciones (no congruentes con la predicción global) de las magnitudes para las primeras.

El lector interesado puede encontrar referencias a otros métodos de estimaciones para pequeñas áreas, especialmente en Ghosh y Rao (1994) o Rao

(1999, 2000 y 2003) [1]. En términos generales, estos métodos pueden descomponerse en directos e indirectos. Los métodos directos son métodos de reponderación de resultados muestrales, basados en estimaciones de mínima chi-cuadrado, y se usan en procedimientos demográficos.

Por su parte, los indirectos se dividen en dos: los que se basan en modelos implícitos que suponen la aceptación para las áreas pequeñas de las regularidades observadas en las áreas mayores (los así llamados estimadores sintéticos o estimadores compuestos), y los que se basan en modelos explícitos (estimadores empíricos de Bayes y estimadores jerárquicos de Bayes, *HB*). Nuestro trabajo, aun no derivándose de ninguno de los anteriores, ni por la modelización propuesta ni por las técnicas de estimación implicadas, se insertaría adecuadamente en los métodos basados en modelos explícitos. Guarda una fuerte relación con los modelos *HB,* pero el papel jugado por el parámetro de precisión t no se ajusta a la estructura "vertical" de estos modelos, en el sentido de que es tanto un parámetro de la distribución a priori de los parámetros del modelo como uno de los parámetros del mismo modelo de indicadores.

En el siguiente apartado, describimos el problema a resolver y fijamos las notaciones básicas que utilizaremos. Establecemos asimismo las distribuciones a priori para los parámetros implicados, básicamente pertenecientes a la familia de distribuciones normal-gamma. En el tercero, formulamos la función de verosimilitud del modelo de indicadores que subyace al problema. El cuarto apartado construye, a partir de las distribuciones a priori y verosimilitud ya obtenidas, la distribución a posteriori. La complejidad en la formulación de esta distribución deriva de la necesidad de hacer operativa la restricción que implica que el modelo debe cuadrar las cifras desagregadas con el agregado total.

El quinto apartado proporciona las distribuciones marginales a posteriori. Su importancia deriva del hecho de que, bajo una función de pérdida cuadrática, el estimador óptimo es la esperanza de dichas distribuciones. La elección de la familia normal-gamma simplifica la obtención de dichas marginales, que resultan ser distribuciones gamma o multivariantes de Student, cuyos momentos de primer y segundo orden se establecen en el apartado séptimo.

En el apartado sexto, intermedio entre los dos descritos en el párrafo anterior, se propone una estimación (asignación) de los parámetros que intervienen en las distribuciones a priori. Como se discutirá más detenidamente en el mismo, se rechaza la idoneidad de las distribuciones neutrales a priori, optando por distribuciones informativas bajo el supuesto del mantenimiento distribucional de la información histórica.

Los apartados octavo y noveno se dedican a la implementación y validación práctica del método obtenido. En el octavo se ilustra su funcionamiento con una desagregación parcial de la Contabilidad Regional de España, en concreto la desagregación provincial del valor añadido en la Construcción en Castilla y León para el año 2002. Esta ilustración pretende únicamente mostrar la factibilidad del procedimiento, y no su precisión, ya que los valores provinciales no son conocidos y, como dijimos en la introducción, no nos planteamos calibrar nuestra estimación con la que realiza el INE, al ser desconocidos los indicadores concretos que se utilizan.

En el apartado noveno se propone la validación del procedimiento con un ejercicio de simulación, en el que se evalúa la precisión del resultado para 168 escenarios, que combinan diversos grados de calidad de los indicadores empleados con diferentes ponderaciones de las áreas desagregadas y con distintas dinámicas de crecimiento de la magnitud.

[1] La última de estas referencias es un magnífico resumen del actual estado de la cuestión.

Finalmente, el trabajo se cierra con el establecimiento de las conclusiones y algunas propuestas de trabajo en el futuro próximo.

2. PLANTEAMIENTO DEL PROBLEMA. DISTRIBUCIÓN A PRIORI

Se dispone de N áreas desagregadas para las que se desea estimar los valores de una cierta magnitud, que escribiremos como $v_{11},\ldots v_{1N}$, y que son realizaciones respectivas de N variables aleatorias, que por comodidad designaremos con el mismo símbolo.

Se conocen, además, los valores correspondientes a esas áreas desagregadas para un año base, $v_{01},\ldots v_{0N}$, valores éstos no aleatorios (o suponemos que trabajamos con la distribución condicionada por los mismos), los correspondientes a ambos años (base y corriente) para el área agregada, que denominaremos V_0 y V_1, asimismo no aleatorios[2], y los valores (no aleatorios) de K indicadores para las N áreas desagregadas, tanto para el año base ($j=0$) como para el corriente ($j=1$), que denotaremos como z_{ji}^k , $i=1,\ldots,N,\quad k=1,\ldots,K,\quad j=0,1$

Fijaremos las siguientes notaciones: Llamamos $a=(V_1-V_0)/V_0$ a la tasa de variación de la magnitud para el área agregada, Z la matriz $N\times K$ que agrupa las tasas de variación de los indicadores,

$$Z=(z_{ik}),\text{ donde } z_{ik}=\frac{z_{1i}^k-z_{oi}^k}{z_{0i}^k},\ i=1,\ldots,N,\ k=1,\ldots,K \qquad [1]$$

y d_i, $i=1,\ldots,N$ van a denotar las tasas de variación de la magnitud para las áreas desagregadas

$$d_i=\frac{v_{1i}-v_{0i}}{v_{0i}},\ i=1,\ldots,N\ , \qquad [2]$$

que agruparemos en el vector $d=(d_1,\ldots,d_N)'$. Utilizaremos el mismo nombre para describir la variable aleatoria y su realización, lo que no se presta a confusión.

La idea básica es que, a priori, los crecimientos de las áreas desagregadas no deben ser muy diferentes del crecimiento del área agregada, ya que éste resulta de la agregación de aquéllos. Más aún, en el reparto del PIB o del VAB de un área agregada en áreas menores (uno de los problemas para los que los autores diseñan este método) la demanda de aquélla es (parcialmente) satisfecha por la oferta de éstas por lo que, a priori, podría aceptarse un valor esperado similar al del área agregada.

En definitiva, para d supondremos una distribución a priori, dada la precisión, $t>0$

$$p(d\,|\,t)\propto t^{N/2}\ \exp\left[-\frac{t}{2}\left[(d-ai)'P(d-ai)\right]\right]\quad,\quad d\in\mathbb{R}^N \qquad [3]$$

siendo $i=(1,\ldots,1)'$ de dimensión N . Se trata de una normal N -variante, con matriz de esperanzas $E[d\,|\,t]=ai$ y matriz de varianzas-covarianzas $Cov(d\,|\,t)=t^{-1}P^{-1}$.

Nótese que la suma de valores de la magnitud para las áreas desagregadas debe coincidir con el valor para el área agregada. En definitiva, la distribución a priori de

[2] La aleatoriedad de $v_{11},\ldots v_{1N}$ no contradice la no aleatoriedad de V_1. Se supone, básicamente, que se trata de un modelo de reparto, y no de agregación, en el sentido de que se genera inicialmente el valor agregado y posteriormente se desagrega con el modelo de reparto. Puede, por tanto, suponerse que el valor agregado es no aleatorio o que trabajamos con la distribución condicionada por dicho valor.

d que utilizaremos es la correspondiente a [3], pero restringida por $v_{11}+\cdots+v_{1N}=V_1$. Más adelante tendremos en cuenta este hecho.

Supondremos que P es una matriz conocida. Posteriormente estudiaremos cómo obtener valores razonables para la misma a partir de datos históricos. Obsérvese que, en relación con el trabajo de Rojo y Sanz (2002), este planteamiento supone una mejora ya que la asignación de correlaciones lineales a priori entre las variables permite utilizar información sobre las relaciones estructurales entre las áreas desagregadas. Obviamente, las relaciones estructurales a valorar deben ser a priori porque, tanto en el trabajo citado como en el actual, la distribución a posteriori ya estima las correlaciones (a posteriori, por supuesto) entre los elementos de d.

Por su parte, t es un parámetro de precisión cuya distribución marginal a priori supondremos gamma[3] $g(b,a)$,

$$p(t) \propto t^{a-1} \; e^{-tb} \;,\; a>0 \;,\; b>0 \;,\; t>0 \qquad [4]$$

donde los parámetros a y b de esta distribución se fijarán más adelante. Recuérdese que, para una distribución gamma, $E[t]=a/b$ y que $Var(t)=a/b^2$. Asimismo, cálculos directos (véase, por ejemplo, Broemeling (1985) o Zellner (1971)) proporcionan los momentos de t^{-1},

$$E\left[t^{-1}\right] = b(a-1)^{-1} \;,\; a>1 \qquad [5]$$

y

$$Var(t^{-1}) = b^2(a-1)^{-2}(a-2)^{-1} \;,\; a>2 \qquad [6]$$

Volviendo sobre las distribuciones, la priori conjunta de d y t resulta, en definitiva, una normal-gamma,

$$p(d,t) \propto t^{(N/2)+a-1} \; \exp\left\{-\frac{t}{2}\left[2b+(d-ai)'P(d-ai)\right]\right\} \;,\; t>0 \;,\; d\in \mathbb{R}^N \qquad [7]$$

Puede integrarse en t para obtener la distribución a priori marginal de d. Por integración directa de [7],

$$p(d) \propto \left[2b+(d-ai)'P(d-ai)\right]^{-(N/2)-a} \;,\; d\in \mathbb{R}^N \qquad [8]$$

Se trata de una distribución de Student multivariante[4] (*MS-t* a partir de ahora), con $2a$ grados de libertad, matriz de posición ai y matriz de precisión (escala) $(a/b)P$, de donde[5] la matriz de varianzas y covarianzas resulta $\frac{b}{a-1}P^{-1}$. Recuérdese, no obstante, que a las distribuciones a priori [7] y [8] hay que añadirlas la restricción inducida para d, a partir de $v_{11}+\cdots+v_{1N}=V_1$.

---

[3] La notación que utilizaremos para la función de densidad de una $g(a,p), a>0, p>0$ es

$$f(x)=\frac{a^p}{\Gamma(p)}x^{p-1}e^{-ax} \text{ para } x>0$$

[4] No existe una definición única para las distribuciones multivariantes de Student. Aquí tomamos la propuesta por Zellner, 1971, pág. 383.

[5] Dada una distribución MS-t, con matriz de escala A y n grados de libertad, su matriz de varianzas y covarianzas resulta $\frac{n}{n-2}A^{-1}$ para $n>2$. Véase, por ejemplo, Zellner, 1971, pág. 383.

3. FUNCIÓN DE VEROSIMILITUD

Se buscan aquí los elementos diferenciadores de cada una de las áreas desagregadas, a través de un modelo de indicadores para las tasas de variación. Estos modelos son frecuentes, aunque no los únicos, en las estimaciones para pequeñas áreas. Véase, por ejemplo, Stasny y otros (1991), para estimaciones de la cosecha por condados en Kansas, donde los indicadores son datos de las explotaciones, o Battese y otros (1988) para estimaciones similares para condados de Iowa donde los indicadores son ahora número de pixeles coloreados en fotos de satélite, o aplicaciones de la estimación EBLUP (Empirical Best Linear Unbiased Prediction, véase Rao (2003), cap. 7) En concreto, y con las notaciones de [1] y [2], el modelo que utilizaremos es

$$d_i = \sum_{k=1}^{K} z_{ik} d_k + e_i \;\; , \; i = 1,\ldots,N$$

o, en notación matricial, $d = Zd + e$, siendo $d = (d_1,\ldots,d_K)'$ un vector de parámetros para los que se supone una distribución a priori condicionada por t, normal, e independiente a priori de $(d \mid t)$, con matriz de esperanzas $E[d \mid t] = \bar{d}$, y matriz de varianzas y covarianzas $Cov(d \mid t) = t^{-1} P_d^{-1}$, esto es,

$$p(d \mid t) \propto t^{K/2} \;\; \exp\left\{-\frac{t}{2}\left[(d-\bar{d})' P_d (d-\bar{d})\right]\right\} \;\; , \;\; d \in \square^K \qquad [9]$$

Obsérvese que, haciendo el mismo ejercicio que nos permitió derivar la expresión [8], puede comprobarse que la distribución a priori marginal de d es una MS-t,

$$p(d) \propto \left[2b + (d-\bar{d})' P_d (d-\bar{d})\right]^{-(K/2)-a} \qquad , \qquad d \in \square^K \qquad [10]$$

con $2a$ grados de libertad, matriz de posición $\bar{d}$, matriz de escala $(a/b)P_d$, y por tanto matriz de varianzas y covarianzas $Cov_p(d) = \frac{b}{a-1} P_d^{-1}$.

En cuanto a las perturbaciones del modelo, supondremos que la matriz de perturbaciones, $e = (e_1,\ldots,e_N)'$ sigue una distribución condicionada por t, $(e \mid t) \to N_N(0, t^{-1} I_N)$, esto es, la correspondiente a un modelo lineal clásico, como es habitual.

La distribución marginal de e se obtiene entonces como hemos obtenido las de d y de d, resultando

$$p(e) \propto \left[2b + e'e\right]^{-(N/2)-a} \quad , \quad e \in \square^N \qquad [11]$$

también una MS-t, con $2a$ grados de libertad, matriz de posición nula, matriz de escala $(a/b)I_N$ y, por tanto, matriz de varianzas y covarianzas $\frac{b}{a-1} I_N$.

Estos resultados nos serán de utilidad más adelante. Pero si nos centramos por el momento en la verosimilitud del modelo, resulta entonces, dados los datos D (que en este caso agrupan los valores de los indicadores, cuyas tasas se recogen en la matriz Z),

$$L(d,d,t \mid D) \propto t^{N/2} \;\; \exp\left\{-\frac{t}{2}\left[(d-Zd)'(d-Zd)\right]\right\} \; , \;\; d \in \square^N \;\; , \;\; t \in \square^+ \;\; , \;\; d \in \square^K \qquad [12]$$

4. DISTRIBUCIÓN A POSTERIORI RESTRINGIDA

Multiplicando la distribución a priori[6] (producto a su vez de las expresiones [3], [4] y [9]) por la verosimilitud expresada en [12], resulta la distribución a posteriori

$$p(d,d,t \mid D) \propto t^{N+(K/2)+a-1} \exp\left\{-\frac{t}{2}\left[2b+(d-ai)'P(d-ai)+(d-\bar{d})'P_d(d-\bar{d})+(d-Zd)'(d-Zd)\right]\right\} \qquad [13]$$

para $d \in \mathbb{R}^N$, $t \in \mathbb{R}^+$, $d \in \mathbb{R}^K$, sujeta a la restricción $v_{11}+\cdots+v_{1N}=V_1$.

Escrita la restricción en términos de tasas, resulta $a = w_1 d_1 + \cdots + w_N d_N$, siendo $w_i = v_{0i}/V_0$, $i=1,\ldots,N$. Puede escribirse en la forma de una restricción para d_N, $d_N = a^* - w_1^* d_1 - \cdots - w_{N-1}^* d_{N-1}$

o en forma matricial,

$$d_N = a^* - v^{*\prime}\ \bar{d} \qquad [14]$$

con $v^* = (w_1^*,\ldots,w_{N-1}^*)'$, $\bar{d} = (d_1,\ldots,d_{N-1})'$, siendo

$$w_i^* = w_i / w_N \quad , \quad i=1,\ldots,N-1 \quad , \quad a^* = a/w_N \quad . \qquad [15]$$

Como la distribución de d es continua, la distribución restringida degenera en un hiperplano. Lo que haremos[7] es sustituir la restricción en la expresión [13] de la distribución a posteriori, y expresarla como una distribución no restringida de $N-1$ tasas, $\bar{d} = (d_1,\ldots d_{N-1})'$, obteniendo finalmente la distribución de d_N a través de la restricción. La elección de la última tasa es arbitraria pero no condiciona el resultado.

---

[6] $p(d,d,t) = p(t)p(d,d \mid t) = p(t)p(d \mid t)p(d \mid t)$ por la independencia a priori de $(d \mid t)$ y $(d \mid t)$

[7]Dada una distribución continua, sujeta a una restricción lineal, equivale a sustituir la restricción en el modelo probabilístico (que ahora dependerá de una variable menos) e imponer, obviamente, la restricción cuando se obtiene separadamente la variable sustituida. Ilustremos este hecho con un sencillo ejemplo. Supongamos que $f(x,y,z)$ es la densidad de una variable tridimensional (X,Y,Z), e impongamos la restricción $aX+bY+cZ=d$. Obtengamos la densidad de (X,Y), sujeta a la restricción. Haciendo la transformación $U=X$, $V=Y$, $T=aX+bY+cZ$, la transformación en $\mathbb{R}^3$ inducida resulta $u=x$, $v=y$, $t=ax+by+cz$, esto es, $x=u$, $y=v$, $z=(t-au-bv)/c$, con módulo del jacobiano igual a $|1/c|$. La densidad conjunta de (U,V,T) resulta $f_{(U,V,T)}(u,v,t)=f\left(u,v,\frac{t-au-bv}{c}\right)\left|\frac{1}{c}\right|$. Entonces,

$$f_{(U,V|T=d)}(u,v)=\frac{f_{(U,V,T)}(u,v,d)}{f_T(d)} \propto f\left(u,v,\frac{d-au-bv}{c}\right).$$

Como $U=X$, $V=Y$, escribiendo esta densidad en (X,Y) resulta

$$f_{(X,Y|T=d)}(x,y) \propto f\left(x,y,\frac{d-ax-by}{c}\right).$$

En definitiva, basta con sustituir en la conjunta la restricción $z=(d-ax-by)/c$.

Realicemos dicha sustitución. Llamando $W = \begin{pmatrix} I_{N-1} \\ -v^{*}{}' \end{pmatrix}$, donde v^* es la matriz definida en la expresión [15] e $\bar{i} = (1,\dots,1)'$ la matriz de unos de dimensión $N-1$, se comprueba[8] que

$$d - ai = W(\bar{d} - a\bar{i}) \qquad [16]$$

de donde $(d-ai)'P(d-ai) = (\bar{d}-a\bar{i})'W'PW(\bar{d}-a\bar{i})$ y

$$(d-Zd)'(d-Zd) = \left[(d-ai)+(ai-Zd)\right]'\left[(d-ai)+(ai-Zd)\right] =$$
$$= (d-ai)'(d-ai) + (ai-Zd)'(ai-Zd) + 2(d-ai)'(ai-Zd) =$$
$$= (\bar{d}-a\bar{i})'W'W(\bar{d}-a\bar{i}) + 2(\bar{d}-a\bar{i})'W'(ai-Zd) + (ai-Zd)'(ai-Zd)$$

Agrupando términos y tras unos sencillos cálculos resulta

$$(d-ai)'P(d-ai) + (d-Zd)'(d-Zd) =$$
$$= \left[\bar{d} - \left(a\bar{i} - T^{-1}W'(ai-Zd)\right)\right]' T\left[\bar{d} - \left(a\bar{i} - T^{-1}W'(ai-Zd)\right)\right] + (ai-Zd)' M_r (ai-Zd) \qquad [17]$$

siendo $T = W'SW$, con $S = P + I_N$ y $M_r = I_N - WT^{-1}W'$. Llamando $\bar{m}_r = a\bar{i} - T^{-1}W'(ai-Zd)$, y sustituyendo la expresión anterior [17] en la distribución a posteriori recogida en [13], la posteriori conjunta restringida se escribe como

$$p(\bar{d},d,t \mid D) \propto t^{N+(K/2)+a-1} \cdot$$
$$\cdot \exp\left\{-\frac{t}{2}\left[2b + (d-\bar{d})'P_d(d-\bar{d}) + (ai-Zd)'M_r(ai-Zd) + (\bar{d}-\bar{m}_r)'T(\bar{d}-\bar{m}_r)\right]\right\} \qquad [18]$$

donde $\bar{d} \in \mathbb{R}^{N-1}$, $t \in \mathbb{R}^+$, $d \in \mathbb{R}^K$, añadiendo la restricción que recoge [14] y que permite obtener d_N.

Debe tenerse en cuenta que $\bar{m}_r$ depende de d.

5. DISTRIBUCIONES MARGINALES A POSTERIORI RESTRINGIDAS

En este apartado obtenemos las distribuciones marginales a posteriori de $\bar{d}$, t y d. Recuérdese que, suponiendo una función de pérdida cuadrática, los estimadores óptimos (de Bayes) para la priori anterior son las esperanzas de la distribución a posteriori. La elección de la familia normal-gamma facilita la obtención de dichas distribuciones que, en todo caso, van a ser MS-t (las de $\bar{d}$ y d) o gamma (la de t). Este hecho nos permite derivar de forma inmediata su esperanza, que es conocida explícitamente.

---

[8] Para probar este resultado basta comprobar que $d_N - a = -v^{*}{}'(\bar{d}-a\bar{i})$. En efecto,

$$-v^{*}{}'(\bar{d}-a\bar{i}) = -w_1^*(d_1-a) - \cdots - w_{N-1}^*(d_{N-1}-a) = -\frac{w_1d_1 + \cdots + w_{N-1}d_{N-1}}{w_N} + a(w_1^* + \cdots + w_{N-1}^*) =$$
$$= d_N - a^* + a(w_N^{-1} - 1) = d_N - a$$

5.1 Distribución marginal a posteriori restringida para t

Teniendo en cuenta que $\int_{\mathbf{R}^{N-1}} e^{-\frac{t}{2}(\bar{d}-\bar{m}_r)'T(\bar{d}-\bar{m}_r)}d\bar{d} \propto \sqrt{\left|t^{-1}T^{-1}\right|} \propto t^{-\frac{N-1}{2}}$, ya que T no depende de los parámetros y es de dimensión $N-1$, la distribución a posteriori restringida para d y t resulta, integrando en $\bar{d}$ la distribución a posteriori restringida [18],

$$p(d,t|D) \propto t^{\frac{N+1}{2}+\frac{K}{2}+a-1} e^{-\frac{t}{2}\left[2b+\left(d-\bar{d}\right)'P_d\left(d-\bar{d}\right)+(ai-Zd)'M_r(ai-Zd)\right]}, d \in \mathbf{R}^K, \quad t>0 \quad [19]$$

Llamando $Q_d = P_d + Z'M_rZ$ y $d_0 = Q_d^{-1}\left[P_d\bar{d} + aZ'M_ri\right]$, el exponente de la exponencial de la anterior expresión [19] resulta

$$2b+\left(d-\bar{d}\right)'P_d\left(d-\bar{d}\right)+(ai-Zd)'M_r(ai-Zd) = 2b+(d-d_0)'Q_d(d-d_0)+a^2i'M_ri+\bar{d}'P_d\bar{d}-d_0'Q_dd_0$$

Teniendo en cuenta que $\int_{\mathbf{R}^k} e^{-\frac{t}{2}(d-d_0)'Q_d(d-d_0)}dd \propto \sqrt{\left|t^{-1}Q_d^{-1}\right|} \propto t^{-K/2}$, si integramos en d la expresión [19], obtenemos la marginal a posteriori restringida de t ,

$$p(t|D) \propto t^{\frac{N+1}{2}+a-1} e^{-\frac{t}{2}\left[2b+a^2i'M_ri+\bar{d}'P_d\bar{d}-d_0'Q_dd_0\right]}, \quad t \in \mathbf{R}^+$$

Esta distribución es una $g(b',a')$, donde

$$a' = \frac{N+1}{2} + a \qquad [20]$$

y

$$b' = b + \frac{a^2i'M_ri + \bar{d}'P_d\bar{d} - d_0'Q_dd_0}{2} \qquad [21]$$

5.2 Distribución marginal a posteriori restringida para d

Si integramos en t la posteriori conjunta restringida para (d,t) de [19], obtenemos

$$p(d|D) \propto \left[2b+(d-d_0)'Q_d(d-d_0)+a^2i'M_ri+\bar{d}'P_d\bar{d}-d_0'Q_dd_0\right]^{-\frac{N+1}{2}-a}, \quad d \in \mathbf{R}^K$$

Se trata de una distribución MS-t de dimensión K, con $n_d = N+1+2a-K$ grados de libertad, matriz de posición

$$d_0 = Q_d^{-1}\left[P_d\bar{d} + aZ'M_ri\right] \qquad [22]$$

(donde recuérdese, $Q_d = P_d + Z'M_rZ$), y matriz de la forma cuadrática

$\frac{N+1+2a-K}{2b+a^2i'M_ri+\bar{d}'P_d\bar{d}-d_0'Q_dd_0}Q_d$, ya que $Q_dd_0 = P_d\bar{d} + aZ'M_ri$.

Su matriz de varianzas y covarianzas vale

$$Cov(d) = \frac{2b+a^2i'M_ri+\bar{d}'P_d\bar{d}-d_0'Q_dd_0}{N+1+2a-K-2}Q_d^{-1} \qquad [23]$$

Obviamente hay que imponer la condición $n_d > 1$ para que la matriz de posición sea a su vez matriz de esperanzas, y de que $n_d > 2$ para que exista la anterior matriz de varianzas y covarianzas.

5.3 Distribución marginal a posteriori restringida para $\bar{d}$

Esta distribución marginal puede extraerse desde la distribución a posteriori conjunta restringida [18]. No obstante, resulta más operativo y analíticamente más tratable obtener la distribución a posteriori de d sin restringir, e imponer posteriormente la restricción.

Obtengamos en primer lugar la posteriori sin restringir para (d,t). Desarrollando los dos miembros de la posteriori sin restringir [13] que dependen de d, $\left(d-\bar{d}\right)'P_d\left(d-\bar{d}\right)=d'P_d d-2d'P_d\bar{d}+\bar{d}'P_d\bar{d}$ y $(d-Zd)'(d-Zd)=d'd-2d'Z'd+d'Z'Zd$, cuya suma, tras sencillas aunque tediosas operaciones, resulta

$$\left(d-\bar{d}\right)'P_d\left(d-\bar{d}\right)+(d-Zd)'(d-Zd)= \qquad [24]$$

$$=\left(d-P_d^*\left[P_d\bar{d}+Z'd\right]\right)'\left(P_d^*\right)^{-1}\left(d-P_d^*\left[P_d\bar{d}+Z'd\right]\right)-\left[P_d\bar{d}+Z'd\right]'P_d^*\left[P_d\bar{d}+Z'd\right]+\bar{d}'P_d\bar{d}+d'd$$

donde hemos llamado $P_d^*=\left[P_d+Z'Z\right]^{-1}$.

La integral de $-t/2$ veces la forma cuadrática en d de la expresión anterior [24], es proporcional a $\sqrt{\left|t^{-1}P_d^*\right|}\ \propto\ t^{-K/2}$, de donde la marginal de (d,t) sin restringir resulta

$$p(d,t\mid D)\propto t^{N+a-1}\exp\left\{-\frac{t}{2}\left[2b+(d-ai)'P(d-ai)-\left(P_d\bar{d}+Z'd\right)'P_d^*\left(P_d\bar{d}+Z'd\right)+\bar{d}'P_d\bar{d}+d'd\right]\right\},\ d\in\mathbb{R}^N \qquad [25]$$

La integral en la recta positiva de la densidad anterior [25], proporciona la distribución a posteriori de d sin restringir,

$$p(d\mid D)\propto \left[2b+(d-ai)'P(d-ai)-\left(P_d\bar{d}+Z'd\right)'P_d^*\left(P_d\bar{d}+Z'd\right)+\bar{d}'P_d\bar{d}+d'd\right]^{-N-a}, \qquad [26]$$

que es una MS-t.

Como indicábamos anteriormente, restrinjamos ahora esta distribución. La restricción equivale a la expresión [16], $d-ai=W\left(\bar{d}-a\bar{i}\right)$, por lo que escribiremos la densidad de d no restringida en función de la diferencia $d-ai$ para realizar la correspondiente sustitución. Así,

$$d'd=\left[(d-ai)+ai\right]'\left[(d-ai)+ai\right]=(d-ai)'(d-ai)+2a(d-ai)'i+a^2i'i$$

y

$$\left(P_d\bar{d}+Z'd\right)'P_d^*\left(P_d\bar{d}+Z'd\right)=\left[P_d\bar{d}+Z'(d-ai)+Z'ai\right]'P_d^*\left[P_d\bar{d}+Z'(d-ai)+Z'ai\right]=$$

$$(d-ai)'ZP_d^*Z'(d-ai)+2(d-ai)'ZP_d^*\left[P_d\bar{d}+Z'ai\right]+\left[P_d\bar{d}+Z'ai\right]'P_d^*\left[P_d\bar{d}+Z'ai\right]$$

de donde la expresión que figura entre corchetes en el denominador de la densidad a posteriori sin restringir de d, [26], resulta

$$2b+\bar{d}'P_d\bar{d}+(d-ai)'\left[I_N+P-ZP_d^*Z'\right](d-ai)+2(d-ai)'\left(ai-ZP_d^*\bar{d}^*\right)+a^2N-\left(\bar{d}^*\right)'P_d^*\bar{d}^*, \qquad [27]$$

donde $\bar{d}^{*} = P_d \bar{d} + aZ'i$. Recordando que $S = I_N + P$, $T = W'SW$ y que la restricción equivalía a $d - ai = W\left(\bar{d} - a\bar{i}\right)$, podemos escribir la expresión anterior [27], tras unos cálculos inmediatos, como

$$2b + a^2N - (\bar{d}^{*})'P_d^{*}\bar{d}^{*} + \bar{d}'P_d\bar{d} - \left[ai - ZP_d^{*}\bar{d}^{*}\right]'W\left[T - W'ZP_d^{*}Z'W\right]^{-1}W'\left[ai - ZP_d^{*}\bar{d}^{*}\right] +$$
$$+\left(\bar{d} - m_{\bar{d}}\right)'\left[T - W'ZP_d^{*}Z'W\right]\left(\bar{d} - m_{\bar{d}}\right)$$

donde

$$m_{\bar{d}} = a\bar{i} - \left[T - W'ZP_d^{*}Z'W\right]^{-1}W'\left[ai - ZP_d^{*}\bar{d}^{*}\right]. \qquad [28]$$

En conclusión, $\bar{d}$ tiene una distribución MS-t, de dimensión $N-1$, grados de libertad $n_{\bar{d}} = N + 2a - 1$, posición $m_{\bar{d}}$ y matriz de la forma cuadrática

$$\frac{N + 2a - 1}{2b + a^2N - (\bar{d}^{*})'P_d^{*}\bar{d}^{*} + \bar{d}'P_d\bar{d} - \left(ai - ZP_d^{*}\bar{d}^{*}\right)'W\left[T - W'ZP_d^{*}Z'W\right]^{-1}W'\left(ai - ZP_d^{*}\bar{d}^{*}\right)}\left[T - W'ZP_d^{*}Z'W\right]$$

de donde la matriz de varianzas y covarianzas resulta

$$\frac{2b + a^2N - \bar{d}^{*}{}'P_d^{*}\bar{d}^{*} + \bar{d}'P_d\bar{d} - \left(ai - ZP_d^{*}\bar{d}^{*}\right)'W\left[T - W'ZP_d^{*}Z'W\right]^{-1}W'\left(ai - ZP_d^{*}\bar{d}^{*}\right)}{N + 2a - 3}\cdot\left[T - W'ZP_d^{*}Z'W\right]^{-1}. \qquad [29]$$

6. ESTABLECIMIENTO DE LOS PARÁMETROS DE LAS DISTRIBUCIONES A PRIORI

Las distribuciones a priori implicadas son la marginal de t, [4], y las condicionadas por t para d y d, respectivamente [3] y [9]. Como puede verse en dichas expresiones, hemos descartado el establecimiento de distribuciones a priori no informativas, neutrales en el sentido de Jeffreys (ver Box y Tiao (1973). Esta propuesta puede verse también en Zellner (1971) o Gamerman (1997), entre otros).

Esta posibilidad ha sido descartada por los autores ya que, en nuestra opinión, las propuestas que implican distribuciones a priori no informativas o aproximadamente no informativas están más bien diseñadas para problemas ligados a muestras amplias, en las que el peso de la verosimilitud es mucho más importante que el de la distribución a priori con lo que, básicamente, es indiferente desde el punto de vista práctico qué distribución a priori usar (siempre que no se restrinjan con la misma los posibles valores de la variable en estudio).

Sin embargo, nuestra experiencia en modelización económica desde una perspectiva bayesiana, especialmente en problemas macroeconómicos, apunta a que la información estadística disponible no es tan amplia como para que la verosimilitud reduzca a un papel simbólico el peso jugado por la distribución a priori. Ello apunta a la conveniencia de utilizar informaciones a priori informativas, siempre que sea posible, utilizando para ello información histórica o de otro tipo.

Este criterio nos conduce a la utilización de distribuciones conjugadas y, en concreto, hemos optado por la familia normal-gamma. Claro está que, dentro de esta familia, pueden considerarse distribuciones "aproximadamente"' no informativas. Esta opción puede verse, por ejemplo, en Young (1996), Spiegelhalter et al. (1999), Broemeling (1985), Rojo y Sanz (1999, 2000 y 2002). En nuestro caso, supondría utilizar como distribuciones a priori ciertas distribuciones normales muy planas para $d \mid t$ y para $d \mid t$ (podría optarse por suponer independencia a priori y tomar varianzas muy grandes, con lo que las curvas de nivel serían “circunferencias" muy abiertas) y para la distribución marginal de t considerar una distribución gamma con densidad prácticamente plana entre cero e infinito (con nuestra parametrización, un valor grande de a y uno pequeño para b, lo que produciría una varianza grande, y asimismo una media grande). Esta opción es, en nuestra opinión, rechazable por los mismos motivos apuntados en la crítica a la utilización de las distribuciones neutrales en problemas como el nuestro.

En definitiva, nos plantearemos utilizar distribuciones a priori informativas, lo que implica la asignación de valores razonables a los parámetros (nuisance) implicados. La idea no es diferente, básicamente, de la que sugiere la obtención de las denominadas *g-priors* (Zellner (1986)) basada en la suposición de que el modelo que se establece para el periodo de estimación es correcto, básicamente, para periodos históricos.

Los parámetros a los que vamos a asignar valores son los dos parámetros incluidos en la distribución a priori de t, [4], esto es, a y b, la matriz de parámetros P que interviene en la distribución a priori [3], de d condicionada por t y la matriz P_d que interviene en la distribución [9] de los parámetros, d, del modelo de indicadores, así como la matriz de posición a priori de la distribución de dichos parámetros, $\bar{d}$.

Como vimos en [5] y [6], los dos primeros momentos de t^{-1} son[9] $E(t^{-1}) = \dfrac{b}{a-1}$, $(a>1)$ y $Var(t^{-1}) = \dfrac{b^2}{(a-1)^2(a-2)}$, $(a>2)$. Por tanto, podemos obtener

$$a = 2 + \frac{E^2(t^{-1})}{Var(t^{-1})} \qquad \text{y} \qquad b = (a-1)\cdot E(t^{-1}) \tag{30}$$

Además, la matriz de varianzas y covarianzas de la distribución a priori, [8], de d es

$$Cov_p(d) = \frac{b}{a-1}P^{-1} \tag{31}$$

de donde

$$P = \frac{b}{a-1}\left(Cov_p(d)\right)^{-1} \tag{32}$$

y la matriz de varianzas y covarianzas de la distribución a priori, [9], de d es

$$Cov_p(d) = \frac{b}{a-1}P_d^{-1} \tag{33}$$

de donde

$$P_d = \frac{b}{a-1}\left(Cov_p(d)\right)^{-1} \qquad [32]$$

Los parámetros a priori se asignan suponiendo que el modelo probabilístico implicado se mantiene para los años anteriores, en los que se suponen conocidos los valores que toman las magnitudes, tanto para las áreas desagregadas como, por

[9] La utilización de los momentos de t^{-1} es preferible a la de los momentos de t porque es más intuitivo pensar en términos de la varianza que de la precisión. En este caso, la elicitación de los momentos de la distribución a priori de t^{-1} no parece complicada con nuestra propuesta, por lo que no parece necesario recurrir a procedimientos de elicitación de cuantiles (véase Kadane et al. (1980)).

supuesto, para el área agregada. Así, suponemos conocidas las tasas de variación para las N áreas desagregadas y para H periodos históricos, que agrupamos en una matriz D_H,

$$D_H = \begin{pmatrix} d_{11} & \cdots & d_{1N} \\ \vdots & \ddots & \vdots \\ d_{H1} & \cdots & d_{HN} \end{pmatrix} \quad [35]$$

y, llamando A_H a la matriz columna de tasas agregadas, $A_H = (a_1,\ldots,a_H)'$ podemos "centrar" dichas tasas históricas mediante

$$D_{Hc} = D_H - A_H i'_N = \begin{pmatrix} d_{11}-a_1 & \cdots & d_{1N}-a_1 \\ \vdots & \ddots & \vdots \\ d_{H1}-a_H & \cdots & d_{HN}-a_H \end{pmatrix}$$

Se puede entonces asignar como matriz de varianzas-covarianzas a priori de d la que resulta de las varianzas y covarianzas obtenidas para el periodo histórico,

$$\widetilde{Cov}_p(d) = \frac{1}{H} D'_{Hc} \left[I_H - \frac{1}{H} i_H i'_H \right] D_{Hc} \quad [36]$$

Por otro lado, si suponemos que el modelo de indicadores que subyace a la función de verosimilitud es adecuado también para el periodo histórico, y suponemos conocidas las tasas de variación de los indicadores para dicho periodo, llamando Z_h, $h = 1,\ldots,H$, a la matriz de tasas de los indicadores para h, (Z_h tiene la misma estructura y tamaño que la matriz Z de [1]) y d_h, $h = 1,\ldots,H$ a las tasas de variación del área desagregada (d_h es la transpuesta de la fila h-ésima de la matriz [35], D_H) podemos calcular

$$\hat{d}_h = (Z_h' Z_h)^{-1} Z'_h d_h, \; h = 1,\ldots,H \quad [37]$$

A partir de estas *valoraciones* (que no estimaciones) de d para los periodos históricos podemos asignar a $\bar{d}$ el valor

$$\hat{\bar{d}} = \frac{1}{H}\sum_{h=1}^{H} \hat{d}_h$$

y, si llamamos $\Delta = \left(\hat{d}_1,\ldots,\hat{d}_H\right)$ a la matriz $K \times H$ con las H valoraciones históricas de d, podemos asignar

$$\widetilde{Cov}_p(d) = \frac{1}{H}\Delta'\Delta - \frac{1}{H^2}\Delta i_H i'_H \Delta' = \frac{1}{H}\Delta\left[I_H - \frac{1}{H} i_H i'_H \right]\Delta' \quad [38]$$

Además, de las H regresiones anteriores, y utilizando las valoraciones [37], obtenemos H asignaciones de t^{-1},

$$\tilde{t}_h^{-1} = \frac{\left(d_h - Z_h \hat{d}_h\right)'\left(d_h - Z_h \hat{d}_h\right)}{N-k}, \qquad h = 1,\ldots,H$$

Construyendo sus dos primeros momentos

$$\overline{t^{-1}} = \frac{1}{H}\sum_{h=1}^{H} \tilde{t}_h^{-1} \quad \text{y} \quad s_{t^{-1}}^2 = \frac{1}{H}\sum_{h=1}^{H} \tilde{t}_h^{-1\,2} - \left(\overline{t^{-1}}\right)^2$$

asignamos, teniendo en cuenta [30]

$$\tilde{a} = 2 + \frac{\left(\overline{t^{-1}}\right)^2}{s_{t^{-1}}^2} \qquad \text{y} \qquad \tilde{b} = (\tilde{a} - 1)\overline{t^{-1}}$$

Podemos ahora asignar, utilizando [31] y [36]

$$\tilde{P} = \frac{\tilde{b}}{\tilde{a}-1}\left(\tilde{C}ov_p(d)\right)^{-1}$$

y, utilizando [33] y [38],

$$\tilde{P}_d = \frac{\tilde{b}}{\tilde{a}-1}\left(\tilde{C}ov_p(d)\right)^{-1}$$

7. ESTIMADORES DE BAYES

En esta sección explicitamos las estimaciones de las variables implicadas supuesta la función de pérdida cuadrática. Nótese que, puesto que para dicha función de pérdida el estimador de Bayes es la esperanza de la distribución a posteriori, y estas distribuciones han sido obtenidas en el apartado 6, nos bastará con explicitar su valor.

7.1 Estimación de los coeficientes, d, del modelo de indicadores

De la expresión [22] tenemos que, si los grados de libertad de la distribución a posteriori son mayores que la unidad, esto es, si $n_d = N+1+2a-K > 1$, la esperanza existe y coincide con el parámetro de posición, esto es,

$$\tilde{d} = d_0 = Q_d^{-1}\left(P_d\bar{d} + aZ'M_r i\right)$$

7.2 Estimación de la precisión, t

Como su distribución a posteriori es una $g(b', a')$, donde a' y b' vienen definidos en [20] y [21], respectivamente, su esperanza resulta

$$\tilde{t} = \frac{2b + a^2 i'M_r i + \bar{d}'P_d\bar{d} - d_0'Q_d d_0}{N+1+2a}$$

7.3 Estimadores de los valores de la magnitud y de sus tasas de variación

Obtenemos previamente los estimadores de d. Recuérdese que d_N se obtiene a través de la restricción [14]. Por tanto, la estimación de d_N resulta

$$E(d_N) = a^* - v^{*\prime} E\left(\bar{d}\right)$$

donde $E\left(\bar{d}\right)$ es la esperanza de la distribución a posteriori restringida de $\bar{d}$, [28].

Entonces

$$\tilde{d} = \begin{pmatrix} E\left(\bar{d}\right) \\ E(d_N) \end{pmatrix} = \begin{pmatrix} m_{\bar{d}} \\ a^* - v^{*\prime} m_{\bar{d}} \end{pmatrix}$$

donde $m_{\bar{d}}$ es el parámetro de posición [28] y será la esperanza de $\bar{d}$ siempre que $n_{\bar{d}} = N+2a-1 > 1$, cosa que, evidentemente, ocurre.

Por otro lado, para $i = 1,\ldots,N-1$

$$Var(d_N) = v^{*\prime}\, Cov(\bar{d})\, v^* \qquad \text{y} \qquad Cov(d_i, d_N) = -v^{*\prime}\left(Cov(\bar{d})\right)_i$$

siendo $\left(Cov(\bar{d})\right)_i$ la columna i-ésima de $Cov(\bar{d})$. De forma agrupada puede escribirse

$$Cov(d) = \begin{pmatrix} I_{N-1} \\ -v^{*\prime} \end{pmatrix} Cov\left(\bar{d}\right)\left(I_{N-1}, \quad -v^{*\prime}\right) \qquad [39]$$

esto es, $Cov(d) = WCov\left(\bar{d}\right)W'$, donde $Cov\left(\bar{d}\right)$ es la expresión [29].

Para recuperar los valores añadidos, puede invertirse la expresión [2]. En expresión matricial,

$$v_1 = v_0 + v_{od} \cdot d$$

siendo

$$v_1 = (v_{11},\ldots,v_{1N})', \quad v_0 = (v_{01},\ldots,v_{0N})', \quad v_{0d} = \begin{pmatrix} v_{01} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & v_{0N} \end{pmatrix}$$

Podemos escribir la expresión análoga para las $N-1$ primeras áreas desagregadas, $\bar{v}_1 = \bar{v}_0 + \bar{v}_{0d} \cdot \bar{d}$.

Si calculamos los momentos de primer y segundo orden,

$$E(v_1) = v_0 + v_{0d}E(d) \qquad \text{y} \qquad E(\bar{v}_1) = \bar{v}_0 + \bar{v}_{0d}E(\bar{d})$$

y

$$Cov(v_1) = v_{0d}Cov(d)v_{0d} \qquad \text{y} \qquad Cov(\bar{v}_1) = \bar{v}_{0d}Cov(\bar{d})\bar{v}_{0d}$$

Desde el punto de vista distribucional, si hacemos el cambio de variable

$$\bar{v}_1 - E(\bar{v}_1) = \bar{v}_{0d}\left(\bar{d} - E(\bar{d})\right)$$

con transformación inversa $\bar{d} - E(\bar{d}) = \bar{v}_{0d}^{-1}\left(\bar{v}_1 - E(\bar{v}_1)\right)$ se obtiene sin dificultad que $\bar{v}_1$ tiene una distribución marginal a posteriori restringida MS-t, con grados de libertad $n_{\bar{v}_1} = n_{\bar{d}} = N + 2a - 1$, matriz de posición $m_{\bar{v}_1} = \bar{v}_0 + \bar{v}_{0d}m_{\bar{d}}$ y matriz de varianzas y covarianzas $Cov(\bar{v}_1) = \bar{v}_{0d}\,Cov(\bar{d})\bar{v}_{0d}$.

En definitiva,

$$\tilde{v}_1 = v_0 + v_{0d}\tilde{d}\ .$$

8. UNA ILUSTRACIÓN NUMÉRICA DEL PROCEDIMIENTO

El Instituto Nacional de Estadística de España proporciona en junio de cada año las estimaciones de crecimiento a 6 ramas[10] para las regiones (NUTS-II) correspondientes al año anterior. Ciertamente se trata de unas primeras estimaciones, que se modifican de forma importante cada año, y que sólo tras cuatro años de provisionalidad se estiman de forma definitiva, pero también es cierto que las políticas económicas regionales se apoyan en estos datos, por lo que a pesar de su provisionalidad adquieren una consistencia "casi-oficial".

El primer avance en las estimaciones de este organismo para las provincias (NUTS-III en la terminología europea) se producen con un retardo de un año adicional, esto es, año y medio después de la finalización del periodo correspondiente. La evaluación estadística coyuntural de los ritmos de crecimiento de las provincias exige, por tanto, la estimación anticipada de sus crecimientos anuales de una forma congruente con los ritmos aceptados para el conjunto regional. Esta estimación se puede realizar con la ayuda de indicadores provinciales, que ya están disponibles cuando se presentan las estimaciones regionales.

No obstante, para regiones de pequeña dimensión económica, o constituidas por numerosas provincias, las pautas de los indicadores provinciales son más bien erráticas, debido a su escasa significación estadística. El procedimiento

[10] Ramas Agraria, Energética, Industrial, Construcción, Servicios de mercado y Servicios de no mercado. Se añade a las mismas los SIFMI (Servicios de intermediación financiera) que no están sectorializados.

desarrollado en los anteriores apartados garantiza unas estimaciones más regulares, y estima la participación de cada indicador en el seno del propio procedimiento.

En la siguiente ilustración se muestra al lector la utilización del procedimiento para desagregar el valor añadido bruto a precios básicos constantes (vabpb) de Castilla y León correspondiente al año 2002 entre las 9 provincias que la componen. En el momento de redactar este artículo, el INE proporciona dicha magnitud para los años 1995 a 2002 en el conjunto regional, y hasta 2001 para las provincias[11]. La estimación se realiza para el sector de la Construcción (sector 2.3 de la clasificación A6) pero puede realizarse para las 6 ramas de actividad de esta clasificación, como hemos señalado anteriormente.

Los indicadores utilizados son el consumo (aparente) de cemento (Tm.), los números de viviendas visadas, iniciadas y terminadas, y la población ocupada en el sector[12]. No se ha estudiado hasta el momento completar el procedimiento con una adecuada metodología de selección de indicadores, trabajo que los autores prevén abordar en un futuro próximo.

Resulta obvio que no se trata de replicar la estimación que el INE realizará en julio de este año[13]. En primer lugar, porque los indicadores utilizados no son, seguramente, los mismos, y en segundo porque las estimaciones del INE son en términos corrientes. Además, tiene poco sentido comparar entre sí distintas estimaciones como muestra de la calidad del método. De hecho, en este apartado se trata únicamente de ilustrar su funcionamiento. Es en el siguiente donde se realiza un ejercicio de simulación, en el que se evalúa la calidad del método propuesto de una forma más sistemática.

La tabla 1 muestra los resultados básicos de nuestra estimación. Como puede verse, se estiman las tasas de crecimiento provinciales sectoriales, tasas que son congruentes con las variaciones de los indicadores (no se presentan éstas, aunque los autores han realizado las pertinentes comprobaciones). Las tasas permiten obtener los valores estimados para 2002, que se presentan en negrita. Obviamente, cumplen la restricción de que la suma de los valores provinciales coincide con el valor regional.

[11] Los datos originales para las provincias son en precios básicos corrientes. El procedimiento de deflación está implementado por los autores. Asimismo, se utiliza una serie regional y provincial desde 1986, no sólo desde 1995, tomada de HISPADAT (2003) en el ámbito regional, y estimada por los autores en el provincial.

[12] Obtenidos de Oficemen el Consumo aparente de cemento, de la Consejería de Fomento de la Junta de Castilla y León los indicadores de viviendas y del INE (E.P.A.) los datos de ocupación en el sector.

[13] Y que estará a disposición de los lectores cuando este trabajo vea la luz. Por otra parte, dicha estimación reestimará los valores regionales de 2002 y anteriores, así como los provinciales de 2001 y anteriores, por lo que cambiará completamente las pautas numéricas de lo realizado en este apartado.

Tabla 1. Valor Añadido Bruto a precios básicos en Construcción (precios constantes). Miles euros de 1995			
	2001 (v_0)	2002(v_1)	Tasa (*d*)
Ávila	188270 (v_{01})	**195074 (v_{11})**	**3.61% (d_1)**
Burgos	304661 (v_{02})	**320966 (v_{12})**	**5.35% (d_2)**
León	488908 (v_{03})	**508583 (v_{13})**	**4.02% (d_3)**
Palencia	161429 (v_{04})	**171019 (v_{14})**	**5.94% (d_4)**
Salamanca	407939 (v_{05})	**427708 (v_{15})**	**4.85% (d_5)**
Segovia	168264 (v_{06})	**177338 (v_{16})**	**5.39% (d_6)**
Soria	73828 (v_{07})	**76449 (v_{17})**	**3.55% (d_7)**
Valladolid	537280 (v_{08})	**557657 (v_{18})**	**3.79% (d_8)**
Zamora	228060 (v_{09})	**235421 (v_{19})**	**3.23% (d_9)**
Castilla y León	2558640 (V_0)	2670217 (V_1)	4.36% (a)

Contabilidad Regional de España y estimaciones propias
En negrita, los valores estimados con nuestro procedimiento

Por otro lado, el procedimiento proporciona la matriz de varianzas y covarianzas a posteriori entre los valores para las áreas desagregadas, a partir de las cuales se puede obtener la correspondiente matriz de correlaciones. Nótese que al ser no aleatorios los valores para el periodo base, las correlaciones lineales entre dichos valores son las mismas que las que existen entre las tasas de crecimiento provincial para las distintas provincias. La siguiente tabla muestra la matriz de correlaciones de las tasas de crecimiento provincial.

Tabla 2. Matriz de correlaciones interprovincial									
	Ávila	Burgos	León	Palencia	Salamanca	Segovia	Soria	Valladolid	Zamora
Ávila	1	0.257	-0.091	-0.124	-0.108	-0.387	0.281	-0.290	-0.019
Burgos	0.257	1	-0.217	-0.279	0.099	-0.116	0.134	-0.397	-0.324
León	-0.091	-0.217	1	-0.325	-0.516	0.283	-0.180	0.228	-0.254
Palencia	-0.124	-0.279	-0.325	1	-0.041	0.044	-0.188	-0.116	0.213
Salamanca	-0.108	0.099	-0.516	-0.041	1	-0.067	-0.154	-0.524	-0.002
Segovia	-0.387	-0.116	0.283	0.044	-0.067	1	-0.179	-0.064	-0.420
Soria	0.281	0.134	-0.180	-0.188	-0.154	-0.179	1	-0.043	-0.022
Valladolid	-0.290	-0.397	0.228	-0.116	-0.524	-0.064	-0.043	1	-0.153
Zamora	-0.019	-0.324	-0.254	0.213	-0.002	-0.420	-0.022	-0.153	1

Puede observarse que no existen correlaciones importantes entre la actividad constructora de las provincias, como corresponde a un mercado que trabaja para la demanda local.

Los coeficientes (d_i, $i = 1,\ldots,5$) de los indicadores utilizados (Consumo de cemento, Viviendas visadas, iniciadas y terminadas, y ocupados en el sector) se presentan en la siguiente tabla, así como las correlaciones a posteriori entre dichos coeficientes estimados. Como ya hemos indicado, es nuestra intención estudiar, en un trabajo posterior, las cuestiones relacionadas con la selección de variables para este tipo de modelos, que permitiría estudiar las causas de ciertos signos negativos que aparecen asociados a las viviendas visadas y terminadas[14]. Obsérvese que las correlaciones entre coeficientes estimados que muestra la tabla son, más bien, pequeñas.

Tabla 3. Coeficientes y matriz de correlaciones a posteriori de los indicadores

a_i		Cons. Cemento	V. Visadas	V. Iniciadas	V. Terminadas	Ocupados
0.056	Consumo de Cemento	1	0.432	-0.378	0.140	0.203
-0.091	Viviendas Visadas	0.432	1	-0.436	0.329	0.310
0.048	Viviendas Iniciadas	-0.378	-0.436	1	-0.092	-0.442
-0.038	Viviendas Terminadas	0.140	0.329	-0.092	1	0.032
0.138	Ocupados	0.203	0.310	-0.442	0.032	1

9. UNA EVALUACIÓN DEL PROCEDIMIENTO

El apartado anterior presenta un ejemplo que ilustra que el procedimiento es factible y proporciona una adecuada evaluación de los estimadores de los parámetros y de las dispersiones de éstos. La utilización prevista del método para la desagregación de las Contabilidades nacionales o regionales (provincias, en nuestro caso) en pequeñas áreas dificulta la evaluación numérica de la precisión del procedimiento, en el sentido de que los valores numéricos para las áreas desagregadas no son conocidos, sino únicamente estimaciones teóricas. Por ejemplo, magnitudes como el consumo privado, la formación bruta de capital o el valor añadido por la Construcción, son cantidades estimadas para cada una de las áreas desagregadas, y no calculadas, de la misma forma que ocurre para las áreas agregadas.

La evaluación de la calidad del procedimiento, además de en su solidez formal, se sustentará, entonces, en procedimientos de simulación. En este apartado realizaremos una breve evaluación del método propuesto en cuanto a

1. La homogeneidad u heterogeneidad en el tamaño de las áreas desagregadas.
2. La homogeneidad u heterogeneidad en su comportamiento, en concreto, en las tasas de variación anual de la magnitud
3. La mayor o menor calidad de los indicadores.

Como se verá en el desarrollo de este apartado, son posibles evaluaciones del método para otros escenarios de los parámetros o hiperparámetros a priori. Pero hemos tratado de elegir los determinantes más importantes, en nuestra opinión, del método empleado. Se verá también que el método únicamente se evalúa para dos áreas desagregadas, y que el modelo de indicadores utiliza un único indicador. La primera limitación no parece afectar a la calidad del método, sino únicamente a la complejidad de los cálculos. En cuanto a la segunda, la multiplicación del número de indicadores dificultaría enormemente la valoración de los resultados, si se realizara junto con los tres items seleccionados, por lo que se ha preferido no incluirlo en este momento en la evaluación.

[14] Nótese también que las correlaciones lineales entre las Viviendas visadas y el resto de los indicadores son negativas, pero el problema no se resuelve suprimiendo este indicador, sino que se traslada a los otros dos indicadores de viviendas (visadas y terminadas). Probablemente hay un problema de periodificación de la actividad, o una necesidad de inclusión de indicadores que valoren directamente las Obras Públicas, y no sólo indirectamente.

Supongamos, pues, que disponemos de dos áreas desagregadas, que numeraremos con los índices 1 y 2. Esto es, $N = 2$ en nuestro caso. Llamaremos, con la notación habitual en este artículo, d_1 y d_2 a las tasas de variación de la magnitud para las áreas desagregadas, $d_i = \frac{v_{1i} - v_{0i}}{v_{0i}} = \frac{v_{1i}}{v_{0i}} - 1, \quad i = 1,2$.

La primera de las disyuntivas establecidas más arriba (**homogeneidad u heterogeneidad en cuanto al tamaño de las áreas desagregadas**) la valoraremos mediante la asignación para el periodo base de un peso, p ($0 < p < 1$) a la primera área desagregada, esto es, asignando un valor a $p = \frac{v_{01}}{v_{01} + v_{02}}$.

Aunque los posibles valores de p forman un continuo, optaremos en la simulación por ciertas especificaciones discretas, en concreto, $\boxed{p = 0.01, 0.05, 0.10, 0.20, 0.30, 0.40 \text{ y } 0.50}$

Obsérvese que las primeras especificaciones corresponden a áreas muy diferentes en tamaño y las últimas a áreas más similares, así como que no se opta, por motivos obvios, por pesos superiores a 0.5, ya que sería equivalente cambiar los papeles de las dos áreas.

El segundo de los criterios establecidos más arriba (**homogeneidad u heterogeneidad en la dinámica de las áreas desagregadas, esto es, en las tasas** d_1 y d_2) la evaluaremos estableciendo $d_1 = k \cdot d_2$, donde tomaremos $\boxed{k = -2, -1, -0.5, -0.1, 0.1, 0.5, 1 \text{ y } 2}$ de manera que los cuatro valores extremos de k indican mayor dinámica (positiva o negativa) en la primera de las áreas, y los cuatro centrales, mayor dinámica en la segunda. Ello nos permite, aunque las relaciones no son lineales, fijar el valor de d_2 y establecer la relación entre d_1 y d_2 a través de k. En este trabajo tomaremos $d_2 = 0.01$, esto es, un crecimiento del 1%.

Nótese que, con lo propuesto hasta ahora, dispondríamos en la simulación de las tasas para las áreas desagregadas, $d_2 = 0.01$ y $d_1 = k \cdot d_2$ para cada una de las especificaciones de k. Asimismo, dispondríamos de la tasa de variación de la magnitud del área agregada mediante la especificación de p, $a = p \cdot d_1 + (1-p) \cdot d_2$. Nos faltaría especificar los valores de los indicadores para ambas áreas desagregadas.

Como indicamos anteriormente, el modelo de indicadores que constituye la verosimilitud en la propuesta bayesiana constará de un único indicador, esto es,

$$d_1 = z_1 d + e_1, \quad d_2 = z_2 d + e_2$$

donde e_1 y e_2 siguen distribuciones condicionadas por t independientes y normales, $e_i \big| t \to N(0, t^{-1/2})$, $i = 1,2$, z_1 y z_2 son las tasas de variación del indicador en las áreas desagregadas (véase [1]) y d es el (único) parámetro de la expresión funcional del modelo.

La simplificación de tomar un único indicador permite trasladar los problemas de escala a la tasa de variación del indicador, con lo que en la simulación se puede tomar $d = 1$. Obviamente, ello impide valorar la sensibilidad del procedimiento al comportamiento de múltiples indicadores, valoración que podría ser objeto de otro ejercicio de simulación. La inclusión de este nivel de análisis en la propuesta actual multiplicaría el número de escenarios, dificultando la visibilidad de los resultados.

Puesto que tomamos $d = 1$, podemos escribir $z_1 = d_1 - e_1$ y $z_2 = d_2 - e_2$, lo que nos permite obtener (las tasas de variación de) los indicadores en cada simulación sin más que simular la perturbación del modelo de indicadores. Esta simulación puede realizarse asignando un valor a t y simulando la distribución condicionada por t (véase [40]), simulación que presenta la ventaja de la independencia de e_1 y e_2. No obstante, parece más prudente simular directamente e_1 y e_2 mediante la distribución marginal conjunta de ambos, que obtuvimos en [11]. Si escribimos la expresión para $N = 2$, resulta

$$p(e) \propto \frac{1}{\left[2b + e_1^2 + e_2^2\right]^{a+1}} \qquad [41]$$

Las distribuciones de ambas perturbaciones no son independientes. De Zellner (1971), pág. 386 obtenemos

$$\sqrt{a / b} \cdot e_2 \to t_{2a} \qquad \text{y} \qquad \sqrt{(2a+1)\big/\left(2b+e_2^2\right)} \cdot e_1 \to t_{2a+1} \qquad [42]$$

La primera expresión de [42] proporciona la distribución marginal de e_2, y la segunda la distribución condicionada $(e_1|e_2)$, por lo que la simulación puede realizarse en ese orden. Obsérvese asimismo que aunque la distribución no depende de t, depende de a y b. Tomaremos para a un valor arbitrario, pero que no impida la existencia de momentos de primer y segundo orden (a =3) y asignaremos $b = s_i^2(a-1)$ teniendo en cuenta [5], donde s_i son valores que permiten valorar el tercero de los anteriores criterios, **la mayor o menor calidad de los indicadores**. En concreto, si llamamos s a la desviación estándar de e_i, $i = 1,2$ para su distribución condicionada por t, tomaremos $\boxed{s = 0.005,\ 0.015 \text{ y } 0.025}$ Aunque estas cifras puedan parecer pequeñas, obsérvese que las tasas de variación están expresadas en tantos por uno con lo que, por ejemplo, $s_2 = 0.025$ implica que la perturbación modifica la tasa de variación desagregada en ± 5% con probabilidad 0.95, siendo dicha tasa de variación del 1 por ciento. La calidad del indicador sería, en este caso, débil.

En definitiva, los escenarios de simulación resultan ser 168. Para cada uno de ellos, se realizan 500 simulaciones, evaluando la calidad del procedimiento mediante los estadísticos[15] MAPE (Mean Absolute Percentage Error) y RMSPE (Root of Mean Square Percentage Error), calculados como

$$MAPE(d_j) = \frac{1}{500}\sum_{i=1}^{500}\left|\frac{d_j^{(i)} - d_j}{d_j}\right|, \qquad RMSPE(d_j) = \frac{1}{500}\sum_{i=1}^{500}\left(\frac{d_j^{(i)} - d_j}{d_j}\right)^2$$

siendo d_j, $j = 1,2$ los valores de d_1, d_2 para cada uno de los 168 escenarios de simulación y $d_j^{(i)}$, $j = 1,2$, $i = 1,\ldots,500$ los valores estimados para ambas tasas en cada una de las 500 iteraciones en cada escenario.

Los dos estadísticos caminan en la misma dirección y muestran, básicamente, los mismos hechos destacables, por lo que nuestros comentarios y tablas se referirán únicamente al primero de ellos. Así, la tabla siguiente presenta los resultados en cada uno de los 168 escenarios. Cada celda contiene tres valores, que corresponden a los distintos valores de s utilizados. Como puede verse, tanto $MAPE(d_1)$ como $MAPE(d_2)$ aumentan según aumenta[16] s esto es, según se deteriora la calidad de los indicadores. El procedimiento, por tanto, muestra una sensibilidad razonable a la calidad del modelo de verosimilitud.

Una segunda conclusión es que las áreas desagregadas más pequeñas sufren de una peor estimación de sus tasas de crecimiento. Visto en la tabla, puede observarse que según se desciende en la tabla, mejora $MAPE(d_1)$y empeora

[15] Véanse Bryan (1999) o Campbell (2002) para una discusión sobre estos estadísticos y otros complementarios y/o alternativos.

[16] Según descendemos dentro de una misma celda.

$MAPE(d_2)$. El descenso en la tabla equivale a un aumento del peso de la primera área desagregada en el total, pasando v_{01} de ser el 1% de $V_0 = v_{01} + v_{02}$ hasta el 50%. En todo caso, no puede sorprender que normalmente $MAPE(d_2)$ sea inferior a $MAPE(d_1)$, puesto que el peso del primer área es siempre inferior o igual al de la segunda (escenarios con pesos superiores a 0.5 supondrían cambiar los papeles de las áreas primera y segunda, por lo que no es necesaria su simulación)

Finalmente, la dinámica relativa entre áreas afecta a cada una de ellas individualmente, pero no a la otra. Como hemos indicado anteriormente, d_2 se toma siempre igual a 0.01 (crecimientos del 1%), mientras que d_1 va variando desde el -2% hasta el -0.1% (en el lado negativo) y luego desde el 0.1% hasta el 2%. En la tabla, estas variaciones de d_1 se producen según nos movemos hacia la derecha en cada una de ellas y, puede observarse que $MAPE(d_2)$ no presenta cambios sistemáticos, mientras que sí lo hace $MAPE(d_1)$, en el sentido de que aumenta cuando nos acercamos a la parte central de la tabla (de derecha hacia el centro o desde la izquierda asimismo hacia el centro). Ello indicaría una mejora en la estimación de d_1 según el área primera es más dinámica (valores de k próximos a 2 o a -2) mientras que la estimación empeoraría cuando k toma valores próximos a cero.

Puede, finalmente, observarse que hay una relación aritmética entre la última fila de $MAPE(d_1)$y la correspondiente de $MAPE(d_2)$, pero es una simple coincidencia derivada del hecho de que hay sólo 2 áreas desagregadas. En concreto, se demuestra sin dificultad que para $p = 0.5$ (las dos áreas son entonces iguales en el periodo base) $|k| \cdot MAPE(d_1) = MAPE(d_2)$ que es el resultado que se observa en la tabla 4, en la cual, la primera columna recoge p, el tamaño relativo, con $v_{01} = p \cdot V_0$.

Tabla 4. Calidad del procedimiento según k, p y s

p	Dinámica relativa, k (d_1= k·d_2)															
	MAPE (d_1)								MAPE (d_2)							
	-2	**-1**	**-0.5**	**-0.1**	**0.1**	**0.5**	**1**	**2**	**-2**	**-1**	**-0.5**	**-0.1**	**0.1**	**0.5**	**1**	**2**
0.01	0.740	1.007	1.542	5.543	4.555	0.558	0.184	0.255	0.015	0.010	0.008	0.006	0.005	0.003	0.002	0.005
	0.767	1.066	1.718	7.351	6.870	1.143	0.552	0.376	0.015	0.011	0.009	0.007	0.007	0.006	0.006	0.008
	0.827	1.278	2.159	10.770	9.721	1.937	1.022	0.518	0.017	0.013	0.011	0.011	0.010	0.010	0.010	0.010
0.05	0.716	0.958	1.452	5.237	4.433	0.577	0.184	0.243	0.075	0.050	0.038	0.028	0.023	0.015	0.010	0.026
	0.728	1.010	1.609	7.323	6.812	1.185	0.554	0.351	0.077	0.053	0.042	0.039	0.036	0.031	0.029	0.037
	0.758	1.224	2.249	10.248	10.044	1.715	0.918	0.510	0.080	0.064	0.059	0.054	0.053	0.045	0.048	0.054
0.1	0.669	0.898	1.365	4.983	4.134	0.543	0.180	0.237	0.149	0.100	0.076	0.055	0.046	0.030	0.020	0.053
	0.686	1.006	1.539	7.200	6.581	1.131	0.546	0.317	0.153	0.112	0.086	0.080	0.073	0.063	0.061	0.071
	0.768	1.213	2.260	9.578	9.900	2.003	0.940	0.472	0.171	0.135	0.126	0.106	0.110	0.111	0.104	0.105
0.2	0.604	0.796	1.172	4.446	3.822	0.510	0.186	0.209	0.302	0.199	0.147	0.111	0.096	0.064	0.046	0.105
	0.622	0.868	1.515	6.286	6.125	1.104	0.515	0.294	0.311	0.217	0.189	0.157	0.153	0.138	0.129	0.147
	0.732	1.174	1.793	9.256	9.406	1.855	0.861	0.485	0.366	0.294	0.224	0.231	0.235	0.232	0.215	0.242
0.3	0.531	0.694	1.061	3.860	3.453	0.465	0.172	0.185	0.455	0.298	0.227	0.165	0.148	0.100	0.074	0.158
	0.546	0.806	1.380	6.334	5.927	1.048	0.521	0.301	0.468	0.345	0.296	0.271	0.254	0.225	0.223	0.258
	0.610	1.090	1.765	9.356	9.299	1.679	0.855	0.443	0.523	0.467	0.378	0.401	0.399	0.360	0.366	0.380
0.4	0.446	0.602	0.900	3.340	2.918	0.394	0.151	0.154	0.595	0.401	0.300	0.223	0.195	0.131	0.100	0.205
	0.468	0.677	1.181	5.214	5.607	0.968	0.495	0.266	0.624	0.451	0.394	0.348	0.374	0.323	0.330	0.355
	0.577	0.942	1.775	8.230	7.993	1.564	0.734	0.401	0.769	0.628	0.592	0.549	0.533	0.521	0.489	0.535
0.5	0.372	0.492	0.770	2.934	2.377	0.343	0.127	0.136	0.743	0.492	0.385	0.293	0.238	0.171	0.127	0.272
	0.408	0.586	1.001	4.708	4.447	0.786	0.409	0.238	0.816	0.586	0.500	0.471	0.445	0.393	0.409	0.477
	0.459	0.782	1.424	7.090	6.766	1.317	0.670	0.356	0.918	0.782	0.712	0.709	0.677	0.659	0.670	0.711

En cada triple celda, los valores se corresponden (de arriba hacia abajo), respectivamente, con s=0.005, s=0.015 y s=0.025

10. CONCLUSIONES

En este trabajo, se propone un método para desagregar magnitudes en espacios geográficos o sectoriales menores, con la ayuda de indicadores. El método consiste en buscar el estimador de Bayes para una función de pérdida cuadrática, dada una distribución a priori para los parámetros de interés perteneciente a la familia normal-gamma.

Como puede verse en el apartado segundo, la consecuencia de una tal elección para la distribución a priori conjunta es que las distribuciones a priori marginales son, bien distribuciones gamma, o bien multivariantes de Student, lo que facilita tanto su utilización como la imposición de la restricción que impone que la suma de las magnitudes de las áreas desagregadas debe ser igual al total agregado.

En el apartado tercero proponemos la función de verosimilitud, a través de un modelo de indicadores con una distribución condicionada normal, esto es, una vez más en el ámbito de las distribuciones normal-gamma.

Los apartados cuarto y quinto se dedican al cálculo de las distribuciones marginales a posteriori. La imposición de la restricción de agregación, aun haciendo más tediosos los cálculos, no complica los resultados, proporcionando unas distribuciones multivariantes de Student o gamma, según los parámetros.

Una atención especial se ha dedicado en el apartado sexto a la asignación de valores razonables a ciertos parámetros que intervienen en las distribuciones a priori y que no son objeto de interés (*nuisance parameters*). Los autores concluyen que la utilización de distribuciones no informativas, neutrales en el sentido de Jeffreys, crea una sobredispersión que no es corregida por la verosimilitud, por lo que resulta preferible la utilización de distribuciones a priori informativas, con una adecuada elección de los valores a asignar a dichos parámetros en función de la información histórica, y aceptando que el modelo actual ha sido asimismo correcto en el pasado histórico.

El apartado séptimo culmina el establecimiento de las estimaciones, sintetizando los resultados anteriores. La utilización de la familia normal-gamma permite que los estimadores tengan expresiones explícitas, como esperanzas de la distribuciones marginales a posteriori restringidas, sin que sea necesario recurrir a procedimientos ergódicos de estimación de medias, como el muestreo de Gibbs. Ello simplifica la implementación del procedimiento, aumentando la velocidad de cálculo.

Los dos últimos apartados se orientan a la implementación práctica del método obtenido. En el primero de ellos, apartado octavo, se realiza una ilustración numérica, que está en la base de la motivación inicial de la búsqueda del método. Se trata de estimar los valores añadidos a precios básicos para las 9 provincias de Castilla y León en la rama de la Construcción para el año 2002, conocidos dichos valores para 2001 y el agregado regional en 2002. Se presentan los resultados obtenidos, que podrían extenderse de la misma forma al resto de las 6 ramas de actividad en que se describen las economías provinciales por parte de la Contabilidad Regional de España. Las tablas muestran los resultados obtenidos, con una implementación en SPLUS realizada por los autores.

Como en dicho apartado se indica, el método debe completarse con una implementación de la selección de indicadores, de manera que sea posible discutir acerca de la adecuación de los signos de los coeficientes y de la superación de efectos colineales entre los indicadores. Este trabajo queda pendiente de futuras ampliaciones del método.

También queda pendiente la obtención de estimaciones óptimas plurianuales. Nótese que con el método que proponemos, la desagregación se realiza para un único año, siendo de utilidad dudosa su extensión a varios años, al menos si es necesario predecir los valores futuros de los indicadores, ya que la falta de calidad estadística de muchos de los indicadores para pequeñas áreas debilita asimismo la calidad de la predicción de los mismos. Con todo, si se dispusiera de indicadores para varios años, el procedimiento podría enlazarse secuencialmente para desagregar los valores agregados de varios años, aunque el procedimiento no sería óptimo, por carecer de simultaneidad.

Asimismo debiera analizarse la sensibilidad de las estimaciones a las asignaciones a priori efectuadas. Esta labor queda pendiente para futuras ampliaciones, especialmente en relación con la determinación del horizonte temporal del periodo histórico considerado.

En el apartado noveno, se evalúa la calidad numérica del procedimiento. Para ello, se simulan distintos escenarios que difieren en la calidad de los indicadores, en el grado de similitud en el tamaño de las áreas desagregadas y en la mayor o menor dinámica de las mismas, con un total de 168 escenarios diferentes. Para cada uno de ellos, se realizan 500 simulaciones de los indicadores y se estima la magnitud para las áreas desagregadas, evaluando el grado de aproximación mediante los estadísticos MAPE y RMSPE, aunque sólo se presentan los valores del primero de ellos.

Básicamente, los resultados obtenidos sugieren en primer lugar que la calidad de la estimación es muy sensible a la calidad de los indicadores (como sería de esperar). Una segunda conclusión es que las áreas desagregadas más pequeñas (proporcionalmente) sufren de una peor estimación de sus tasas de crecimiento, esto es, que la mejor calidad de la estimación se presenta cuando las áreas desagregadas son similares en tamaño. Finalmente, una tercera apunta a la mejora en la calidad de la estimación cuanto más dinámica es un área desagregada, sin afectar a la calidad de la estimación del resto de las áreas. Qué duda cabe que esta validación y evaluación del procedimiento debe complementarse en escenarios más complejos, pero proporciona una valoración básicamente favorable al método empleado.

REFERENCIAS BIBLIOGRÁFICAS

BATTESE, G.E., Harter, R.M., y Fuller, W.A. (1988) ``An Error Component Model for Prediction of County Crop Areas Using Survey and Satelite Date'' Journal of the American Statistical Association, 83, págs. 28-36

BOX, G.E.P. y TIAO, G.C. (1973) Bayesian Inference in Statistical Analysis, Addison Wesley, Reading, Mass.

BROEMELING, L.D. (1985) Bayesian Analysis of Linear Models, Marcel Dekker, Inc., New York.

BRYAN, Th. (1999) Evaluating Small-Area Population Estimation Results with Loss Functions and Optimisation Criteria, Bulletin of the I.S.I., 52nd. Session, Proceedings, Tome LVIII, Finland

CAMPBELL, P. R. (2002) Evaluating Forecast Error in State Population Projections Using Census 2000 Counts, US Bureau of the Census, Population Division, Working Paper Series, No. 57. Washington.

GAMERMAN, D. (1997) Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference, Chapman & Hall.

GOSH, M. y RAO, J.N.K. (1994) ``Small area estimation: an appraisal''. Statistical Science, 9, pp.55-93.

HISPADAT (2003), Base de datos HISPADAT, Proyecto HISPALINK, diciembre 2003 (http://www.hispalink.org)

KADANE, J.B., Dickey, J.M., Winkler, R.L., Smith, V.S. y Peters, S.C. (1980) ``Interactive Elicitation of Opinion for a Normal Linear Model'' JASA, 75, págs. 845-854.

RAO, J.N.K. (1999) ``Some recents advances in model-based small area estimation''. Survey Methodology, 25, 2, pp. 175-186.

RAO, J.N.K. (2000) ``Statistical methodology for indirect estimations in small areas''. Seminario Internacional de Estadística de Euskadi, 39. Eustat.

RAO, J.N.K. (2003) Small Area estimation, Wiley Interscience, Wiley Series in Survey Methodology

ROJO, J.L. y SANZ, J.A. (1999): ``Una propuesta bayesiana para la distribución de Contabilidades regionales por procedimientos indirectos'', en el libro Cambios Regionales en la U.E. y Nuevos Retos Territoriales, AECR, ISBN: 84-607-3322-X, Madrid, 2001, páginas 1-19.

ROJO, J.L. y SANZ, J.A. (2000) ``Estimaciones de magnitudes en pequeñas áreas'' Trabajo subvencionado por la Dirección General de Estadística de la Junta de Castilla y León (no publicado).

ROJO, J.L. y SANZ, J.A. (2002) ``Estimaciones para pequeñas áreas: un enfoque bayesiano a un problema de distribución''. Estudios de Economía aplicada. 20-I, pp. 217-240.

SPIEGELHALTER, D., THOMAS, A. y BEST, N. (1999) WinBUGS, V. 1.2: User Manual MRC Biostatistics Unit, Cambridge, U.K.

STANSI, E., GOEL, P.K., y RUMSEY, D.J. (1991) ``County Estimates of Wheat Production'', Survey Methodology, 17, págs. 211-225

YOUNG, M.R. (1996) ``Robust Seasonal Adjustment by Bayesian Modelling'', Journal of Forecasting, 15, pp. 355-367.

ZELLNER, A. (1971) An Introduction to Bayesian Inference in Econometrics, J. Wiley & Sons, Inc. N. York.