ESTUDIOS DE ECONOMÍA APLICADA VOL. 27-2 2009 PÁGS. 1-22

Estimación Bayesiana en modelos de producción con frontera determinista*

FRANCISCO JAVIER ORTEGA IRIZO y JESÚS BASULTO SANTOS
Facultad de Ciencias Económicas y Empresariales
UNIVERSIDAD DE SEVILLA
e-mail: fjortega@us.es; basulto@us.es

RESUMEN

Como consecuencia de no ser válidos, en los modelos de producción con frontera determinista, las condiciones usuales de regularidad (que justifican la consistencia y normalidad asintótica de los estimadores máximo verosímiles), desconocemos las propiedades generales de estos estimadores. Una alternativa son los métodos de inferencia bayesiana que, gracias al algoritmo de Gibbs, son relativamente fáciles de aplicar. En nuestro trabajo proponemos una distribución a priori no informativa para este modelo y, por simulación, analizamos el comportamiento de los estimadores e intervalos bayesianos.

Palabras clave: modelos de producción con frontera, modelo Half-Normal, estimación bayesiana, algoritmo de Gibbs.

Bayesian Estimation in Deterministic Frontier Production Models

ABSTRACT

As a consequence of non valid regularity conditions, we unknown the general properties (consistency and asymptotic normality) of the maximun likelihood estimators in the deterministic frontier production models. An alternative to these estimators are the bayesian inference methods. Thanks to Gibbs' algorithm, these methods are relativity easy to apply. In our work, we propose a prior noninformative distribution for the deterministic frontier models, and, by simulation, we study the Bayesian estimators and intervals behaviour.

Keywords: Frontier Production Models, Half-Normal Model, Bayesian Estimation, Gibbs Sampling.

Clasificación JEL: C11, C2.

Artículo recibido en diciembre de 2008 y aceptado en abril de 2009.
Artículo disponible en versión electrónica en la página www.revista-eea.net, ref. ə-27205.

*** Agradecimientos:** Los autores queremos expresar nuestro agradecimiento a los evaluadores, cuyas sugerencias han contribuido a mejorar la versión inicial del trabajo.

1. INTRODUCCIÓN

La estimación de los modelos de producción con frontera es uno de los temas más interesantes del área de economía aplicada, pudiéndose situar su origen en el artículo de Aigner and Chu (1968). Posteriormente, el tema ha suscitado numerosos trabajos de investigación. En este tipo de modelos, se expresa un determinado output y en función de una serie de inputs $x_1,...,x_k$ a través de un modelo de regresión. Por ejemplo, para un único input x, el modelo quedaría expresado como $y=\beta_1+\beta_2 x+\varepsilon,$ donde β_1 y β_2 son parámetros desconocidos y ε es una perturbación aleatoria. La particularidad de estos modelos es que la parte determinista (en nuestro ejemplo $\beta_1+\beta_2 x$) representa la frontera de producción o valor máximo alcanzable de output para unos inputs dados, mientras que la perturbación aleatoria (diferencia entre la producción real y la máxima posible) representaría el grado de ineficiencia en el proceso productivo. De esta forma, la perturbación aleatoria ha de ser necesariamente negativa y así la hipótesis habitual de normalidad de las perturbaciones no tiene cabida en los modelos de producción con frontera. Un ejemplo de aplicación de estos modelos puede verse en García et al. (2004) donde se analiza la eficiencia de la producción de la flota pesquera del golfo de Cádiz.

Alternativamente, podemos utilizar la formulación dual de este problema, en la que la variable dependiente es el coste de producción y la perturbación aleatoria sería positiva, representando la parte determinista del modelo el coste mínimo de producción para unos inputs dados.

Esta formulación, denominada habitualmente modelo de producción con frontera determinista, presenta una importante dificultad de estimación, debido al hecho de que la modelización de la perturbación unilateral rompe las hipótesis habituales de regularidad asumidas para la obtención de las propiedades asintóticas del estimador máximo-verosímil. Así, podremos obtener los estimadores máximo-verosímiles resolviendo un problema de optimización con restricciones, pero la determinación de las propiedades estadísticas del estimador es un problema complicado. Otro inconveniente que se ha destacado acerca de este modelo es su gran sensibilidad a la presencia de outliers y/o valores extremos (Simar, 2007). Para evitar esta cuestión, se han propuesto métodos de estimación que permiten dejar un cierto porcentaje de observaciones "por encima" de la frontera estimada (Cazals et al., 2002 y Daouia and Simar, 2004).

Alternativamente, se han planteado en la literatura los denominados modelos de producción con frontera estocástica. En ellos, se introducen dos perturbaciones, siendo el modelo para un único input de la forma $y=\beta_1+\beta_2 x+v,$ donde $v=\varepsilon+u.$ En este caso, $\varepsilon<0$ es una medida de la ineficiencia, mientras que $u\in\mathbb{R}$ refleja factores aleatorios, como por ejemplo errores de medida. Con esta formulación, se permite que algunas observaciones puedan quedar por encima de la frontera y ello sería debido a factores aleatorios. En estos modelos, se verifican las condiciones de regularidad habituales, por lo que podemos aplicar las propiedades asintóticas del

estimador máximo-verosímil para llevar a cabo nuestras inferencias. El problema principal de esta formulación es que, una vez estimado el modelo, no podemos identificar qué parte de cada residuo $\hat{v}$ se debe a ineficiencia y qué parte se debe a errores aleatorios, es decir, no podemos obtener medidas individuales de eficiencia; en este caso, sólo es posible estimar niveles medios para el grupo analizado. Una posible salida a este problema se ofrece en Jondrow et al. (1982). En Førsund et al. (1980) se analiza con mayor detalle las ventajas e inconvenientes de cada una de las formulaciones.

Los modelos con frontera estocástica presentan otro inconveniente. Al ser el término de perturbación ν la suma de una variable negativa ε más otra variable u que toma valores en $\mathbb{R}$ (y que siempre se supone que sigue una distribución normal de media nula) resulta que la distribución de ν es asimétrica a la izquierda (en el caso de los modelos de coste de producción la asimetría es a la derecha). Para obtener los estimadores máximo verosímiles de los parámetros, se toma como punto de partida los estimadores mínimo cuadráticos (MCO). Puede ocurrir que los residuos MCO presenten asimetría "errónea" (es decir asimetría positiva en el caso de modelos de producción o asimetría negativa en el caso de modelos de coste) y en este caso aparecen dificultades para obtener los estimadores máximo verosímiles.

Este hecho se ha considerado como una prueba inicial de diagnóstico del modelo, es decir, la presencia de asimetría "errónea" significaría que el modelo está mal especificado y no es adecuado a los datos observados (Green, 1993). Por ejemplo, el programa LIMDEP 7 muestra un mensaje de error puesto que no es posible continuar con el proceso de estimación a través del método de máxima verosimilitud. Sin embargo, recientes estudios simulados han puesto de manifiesto que en el modelo con frontera estocástica la asimetría "errónea" puede aparecer con probabilidades no despreciables, que pueden llegar a ser hasta del 30% y que en ciertos casos se necesita un tamaño muestral superior a 1000 para que esta probabilidad descienda al 5% (Simar, 2007).

El objetivo principal de este trabajo es aportar una solución a las dificultades de estimación que se presentan en el modelo con frontera determinista a través de las técnicas de inferencia Bayesiana. Como sabemos, este enfoque de la inferencia se basa en la combinación de la información aportada por los datos y el modelo probabilístico supuesto para los mismos (a través de la función de verosimilitud) con la información subjetiva acerca del parámetro aportada por el investigador (a través de la distribución a priori). No obstante, si no queremos asumir ninguna información inicial acerca de los parámetros, podemos derivar una distribución a priori "no informativa" o "referencia a priori" a través de algún proceso matemático formal aplicado a la función de verosimilitud del modelo. Esta opción conduce a las técnicas bayesianas *objetivas*, en el preciso sentido de que los resultados obtenidos sólo dependen del modelo asumido y de los datos observados. Una excelente y moderna revisión de las técnicas bayesianas, y en particular de las técnicas bayesianas objetivas puede verse en Bernardo (2009).

La primera propuesta de obtención de una distribución a priori a partir de la verosimilitud es la conocida Regla de Jeffreys (Jeffreys, 1946, 1961), aplicable solo a modelos regulares. Posteriormente, han sido muchas las propuestas alternativas; una revisión bastante exhaustiva puede encontrarse en Kass and Wasserman (1996). Cabe destacar las referencias a priori de Berger&Bernardo (Bernardo, 1979, 2005, Berger and Bernardo, 1992) y las denominadas *probability matching priors* (Datta and Gosh, 1995, Datta and Sweeting, 2005).

Nosotros vamos a trabajar dentro del marco de las técnicas bayesianas objetivas, pero vamos a usar una distribución a priori no informativa que se obtiene aplicando la generalización de la Regla de Jeffreys propuesta en Ortega y Basulto (2003), que es aplicable tanto a modelos regulares como no regulares.

Veremos que, aunque debido a la naturaleza complicada del problema, no podemos obtener fórmulas explícitas para nuestras inferencias (estimadores, intervalos, etc), sí que podremos aplicar el algoritmo de Gibbs para obtener una muestra simulada de la distribución a posteriori, lo que nos permitirá estimar cualquier característica interesante de nuestro modelo.

A partir de aquí, en la sección 2 presentamos el modelo de producción así como su función de verosimilitud bajo las hipótesis asumidas para el término de perturbación; en la sección 3 obtenemos la distribución a priori no informativa y, a partir de ella, el núcleo de la densidad a posteriori conjunta; en la sección 4 presentamos la formulación del algoritmo de Gibbs para nuestro problema, que se basa en la simulación de muestras de las distribuciones condicionadas unidimensionales; en la sección 5 aplicamos la metodología descrita a dos ejemplos con datos simulados para ayudar a comprender la naturaleza del problema y efectuamos un estudio tipo Monte Carlo para analizar las propiedades del método de estimación propuesto, mientras que en la sección 6 aplicamos nuestra propuesta al ejemplo estudiado en Coelli et al (1998), pág. 192. Por último, en la sección 7 recogemos las conclusiones principales obtenidas a partir de nuestro trabajo.

2. PLANTEAMIENTO DEL MODELO

Consideremos el modelo de producción especificado como:

$$y_i = \mathbf{x_i}'\boldsymbol{\beta} + \varepsilon_i, \quad i = 1,...,n$$

donde y_i es la producción, $\boldsymbol{\beta} \in \mathbb{R}^k$ es un vector de parámetros, $\mathbf{x_i}$ es el correspondiente vector de variables exógenas, n es el tamaño de la muestra y $\varepsilon_i < 0$ es una perturbación aleatoria que mide la ineficiencia de la observación *i*-ésima. Asumiremos la hipótesis de que $-\varepsilon_i$ sigue una distribución half-normal $(0, \sigma^2)$ (Johnson et al. 1994, Pewsey 2002, 2004), y por tanto la función de densidad viene dada por:

$$f(-\varepsilon_i) = \frac{2}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(-\varepsilon_i)^2\right\}, \ -\varepsilon_i > 0,$$

con lo que la densidad inducida para y_i sería:

$$f(y_i \mid \mathbf{x_i}, \boldsymbol{\beta}, \sigma) = \frac{2}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y_i - \mathbf{x_i'}\boldsymbol{\beta})^2\right\}, \ y_i \le \mathbf{x_i'}\boldsymbol{\beta}.$$

Por tanto, la función de verosimilitud para la muestra de tamaño n del modelo quedaría como:

$$L(\boldsymbol{\beta}, \sigma \mid \mathbf{y}, \mathbf{X}) \propto (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_i (y_i - \mathbf{x_i'}\boldsymbol{\beta})^2\right\}, \ y_i \le \mathbf{x_i'}\boldsymbol{\beta} \quad \forall i = 1,...,n$$

donde $\mathbf{y} = (y_1,...,y_n)' \in \mathbb{R}^n$ y $X' = [x_1 \mid ... \mid x_n] \in \mathrm{M}_{k\times n}$, que expresaremos como

$$L(\boldsymbol{\beta}, \sigma \mid \mathbf{y}, \mathbf{X}) \propto (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_i (y_i - \mathbf{x_i'}\boldsymbol{\beta})^2\right\}, \ \boldsymbol{\beta} \in \mathrm{B},$$

donde $\mathrm{B} = \left\{\boldsymbol{\beta} \in \mathbb{R}^k / \, y_i \le \mathbf{x_i'}\boldsymbol{\beta} \ \ \forall i = 1,...,n\right\}$, es decir, B es el espacio paramétrico o conjunto de valores de $\boldsymbol{\beta}$ compatibles con los datos observados según el modelo establecido.

3. DISTRIBUCIONES A PRIORI Y A POSTERIORI

Como hemos visto en la sección anterior, el conjunto de parámetros desconocidos del modelo es $(\boldsymbol{\beta}, \sigma)$, donde $\boldsymbol{\beta} \in \mathbb{R}^k$ y $\sigma > 0$. El proceso de inferencia bayesiana se basa en la determinación de la distribución a posteriori conjunta del vector paramétrico $(\boldsymbol{\beta}, \sigma)$ dados los datos, que notaremos por $\pi(\boldsymbol{\beta}, \sigma \mid \mathbf{y}, \mathbf{X})$. Como es conocido, para obtener la distribución a posteriori es necesario previamente elegir una distribución a priori conjunta para el vector de parámetros $(\boldsymbol{\beta}, \sigma)$, que notaremos por $\pi(\boldsymbol{\beta}, \sigma)$. Para obtener esta distribución a priori conjunta, en primer lugar obtendremos las condicionadas unidimensionales.

Considerando $\boldsymbol{\beta}$ fijo, el modelo es regular por lo que la distribución a priori para σ (condicionada a $\boldsymbol{\beta}$) la obtendremos a partir de la Regla de Jeffreys:

$$\pi(\sigma \mid \boldsymbol{\beta}) \propto \sqrt{-E\left[\frac{\partial^2 \ell}{\partial \sigma^2}\right]} \propto \sigma^{-1},$$

donde $\ell = \ell(\boldsymbol{\beta}, \sigma) = \text{Ln}(L(\boldsymbol{\beta}, \sigma \mid \mathbf{y}, \mathbf{X}))$.

Ahora, considerando fijos σ y $\boldsymbol{\beta}_{(j)} = (\beta_1, ..., \beta_{j-1}, \beta_{j+1}, ..., \beta_k)$, el modelo es no regular (obsérvese que el soporte depende de β_j) y por tanto, siguiendo la propuesta de Ortega y Basulto (2003), pág. 377, la distribución a priori se obtendrá en base a la derivada primera del logaritmo de la verosimilitud (en vez de utilizar la derivada segunda como se hace en la Regla de Jeffreys). Concretamente:

$$\pi(\beta_j \mid \boldsymbol{\beta}_{(j)}, \boldsymbol{\sigma}) \propto \left| E\left[\frac{\partial \ell}{\partial \beta_j} \right] \right| = \left| \sigma^{-2} \sum_i (E[y_i \in] - x_i' \beta)\, x_{ij} \right|,$$

donde x_{ij} es la componente *j*-ésima del vector $\mathbf{x_i}$ y $|z|$ representa el valor absoluto de *z* (observemos que el uso del valor absoluto garantiza la obtención de una distribución positiva). Puesto que y_i sigue una distribución half-normal de parámetros $\mathbf{x_i'\beta}$ y σ^2, se verifica $E[y_i] = x_i'\beta - \sqrt{2/\pi}\,\sigma$ (Greene, 1998), por lo que

$$\pi(\beta_j \mid \boldsymbol{\beta}_{(j)}, \boldsymbol{\sigma}) \propto \left| \sigma^{-1} \sqrt{2/\pi} \sum_i x_{ij} \right| \propto 1.$$

Ahora, a partir de las distribuciones condicionadas unidimensionales, planteamos la obtención de una distribución a priori conjunta compatible, que no siempre tiene por qué existir (Arnold y otros (1999), págs. 8 y 133, Ortega y Basulto (2003), pág. 381). Para ello, usamos el siguiente resultado.

Consideremos un modelo biparamétrico, cuyos parámteros unidimensionales llamaremos genéricamente θ_1 y θ_2. Si se verifica que los núcleos de las distribuciones a priori condicionadas son de la forma:

$$\text{Núcleo}(\pi(\theta_1 \mid \theta_2)) = f(\theta_1)$$
$$\text{Núcleo}(\pi(\theta_2 \mid \theta_1)) = g(\theta_2)$$

entonces, es fácil ver que existe una distribución a priori conjunta consistente con ambas condicionadas cuya expresión viene dada por

$$\pi(\theta_1, \theta_2) \propto f(\theta_1) \times g(\theta_2).$$

En nuestro modelo, se verifica esta condición en todas las distribuciones condicionadas. Así, aplicando el resultado para el caso bidimensional sucesivamente obtendríamos $\pi(\boldsymbol{\beta} \mid \sigma) \propto 1$ y como $\pi(\sigma \mid \boldsymbol{\beta}) \propto \sigma^{-1}$ podemos concluir que la distribución a priori conjunta compatible con las condicionadas viene dada por

$$\pi(\boldsymbol{\beta}, \sigma) \propto 1 \times \sigma^{-1} \propto \sigma^{-1}.$$

El núcleo de la distribución a posteriori conjunta, lo obtendremos multiplicando la verosimilitud por la distribución a priori, con lo que obtendremos:

$$\pi(\boldsymbol{\beta},\sigma \mid \mathbf{y},\mathbf{X}) \propto (\sigma^2)^{-(n+1)/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_i (y_i - \mathbf{x_i'}\boldsymbol{\beta})^2\right\},\ \boldsymbol{\beta} \in \mathrm{B},$$

que, alternativamente, podemos expresar como:

$$\pi(\boldsymbol{\beta},\sigma \mid \mathbf{y},\mathbf{X}) \propto (\sigma^2)^{-(n+1)/2} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right\},\ \boldsymbol{\beta} \in \mathrm{B}.$$

La manipulación algebraica de esta distribución es complicada, debido a la restricción $\boldsymbol{\beta} \in B$, que hace que no sea fácil conocer los límites de integración y que lleva también a que las distribuciones marginales de cada componente β_j estén definidas a trozos. Para evitar esta dificultad y dar respuesta al problema de la inferencia, utilizaremos el algoritmo de Gibbs (Gelfand and Smith,1990), ya que, como veremos en el epígrafe siguiente, no resulta excesivamente complicado simular muestras de las distribuciones unidimensionales de cada parámetro condicionadas al resto de parámetros.

4. FORMULACIÓN DEL ALGORITMO DE GIBBS

Aplicaremos el algoritmo de Gibbs, obteniendo muestras simuladas de las distribuciones $\sigma \mid \boldsymbol{\beta},\mathbf{y},\mathbf{X}$ y $\beta_j \mid \boldsymbol{\beta}_{(\mathbf{j})},\sigma,\mathbf{y},\mathbf{X}$, $j = 1, \ldots, k$.

Observando que la distribución a posteriori conjunta puede expresarse como

$$\pi(\boldsymbol{\beta},\sigma \mid \mathbf{y},\mathbf{X}) \propto (\sigma^{-2})^{\frac{n+1}{2}} \exp\left\{-\sigma^{-2}\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right\},\ \boldsymbol{\beta} \in \mathrm{B},$$

obtenemos inmediatamente que $\sigma^{-2} \mid \boldsymbol{\beta},\mathbf{y},\mathbf{X}$ sigue una distribución Gamma, concretamente:

$$\sigma^{-2} \mid \boldsymbol{\beta},\mathbf{y},\mathbf{X} \sim Ga\left(\frac{n+3}{2}, \frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right)$$

Para obtener las distribuciones $\beta_j \mid \boldsymbol{\beta}_{(\mathbf{j})},\sigma,\mathbf{y},\mathbf{X}$, consideremos $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y}$ y $\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ (el estimador y los residuos mínimo cuadráticos del modelo de regresión sin restricciones). Teniendo en cuenta que:

$$(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) = \hat{\mathbf{u}}'\hat{\mathbf{u}} + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\mathbf{X'X}\,(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}),$$

la distribución a posteriori conjunta puede expresarse como:

$$\pi(\boldsymbol{\beta},\sigma \mid \mathbf{y},\mathbf{X}) \propto (\sigma^2)^{-(n+1)/2} \exp\left\{-\frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{2\sigma^2}\right\} \exp\left\{-\frac{1}{2\sigma^2}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right\},\ \boldsymbol{\beta}\in \mathrm{B}$$

de donde obtenemos que

$$\boldsymbol{\beta} \mid \sigma,\mathbf{X},\mathbf{y} \sim N_k(\hat{\boldsymbol{\beta}},\sigma^2(\mathbf{X}'\mathbf{X})^{-1}),\ \boldsymbol{\beta}\in B,$$

es decir, una distribución Normal k-dimensional truncada en el subconjunto $B \subseteq \mathbb{R}^k$.

En principio, podrían generarse valores de esta distribución a través del procedimiento:

1. Generar un valor $\boldsymbol{\beta}_0 \sim N_k(\hat{\boldsymbol{\beta}},\sigma^2(\mathbf{X}'\mathbf{X})^{-1})$.
2. Si $\boldsymbol{\beta}_0 \in B$, aceptar dicho valor; si no, rechazarlo y volver a 1.

Sin embargo, en la práctica, es frecuente que la probabilidad de rechazo de este procedimiento sea tan elevada que el método no resulte viable. Una solución para este problema es analizar las distribuciones condicionadas unidimensionales.

Para obtener muestras simuladas de $\beta_j \mid \boldsymbol{\beta}_{(j)},\sigma,\mathbf{y},\mathbf{X}$, reordenando los parámetros, podemos considerar la siguiente partición:

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_j \\ \boldsymbol{\beta}_{(j)} \end{pmatrix} \ y\ \ \Sigma = \begin{pmatrix} \sigma_j^2 & \boldsymbol{\sigma}'_{j(j)} \\ \boldsymbol{\sigma}_{j(j)} & \Sigma_{j(j)} \end{pmatrix},$$

donde hemos llamado $\Sigma = \sigma^2(X'X)^{-1}$. Si en el modelo no hubiese restricciones, es conocido que se tendría

$$\beta_j \mid \boldsymbol{\beta}_{(j)},\sigma,\mathbf{y},\mathbf{X} \sim N(\hat{\beta}_j + \Sigma_{j(j)}^{-1}(\boldsymbol{\beta}_{(j)} - \hat{\boldsymbol{\beta}}_{(j)}), \sigma_j^2 - \boldsymbol{\sigma}'_{j(j)}\Sigma_{j(j)}^{-1}\boldsymbol{\sigma}_{j(j)}).$$

Ahora, si suponemos que todas las variables explicativas toman valores positivos, las restricciones $y_i \le \mathbf{x}_i'\boldsymbol{\beta},\ i=1,...,n$ son equivalentes a

$$\beta_j \ge \frac{y_i - \mathbf{x}'_{i(j)}\boldsymbol{\beta}_{(j)}}{x_{ij}}, i=1,...,n \Leftrightarrow \beta_j \ge \max_i \left\{\frac{y_i - \mathbf{x}'_{i(j)}\boldsymbol{\beta}_{(j)}}{x_{ij}}\right\} \Leftrightarrow \beta_j \ge b(\boldsymbol{\beta}_{(j)},X,\mathbf{y}),$$

donde $b(\boldsymbol{\beta}_{(j)},X,\mathbf{y}) = \max_i \left\{(y_i - \mathbf{x}'_{i(j)}\boldsymbol{\beta}_{(j)})/x_{ij}\right\}$, y como consecuencia, obtenemos:

$$\beta_j \mid \boldsymbol{\beta}_{(j)},\sigma,\mathbf{y},\mathbf{X} \sim N(\hat{\beta}_j + \Sigma_{j(j)}^{-1}(\boldsymbol{\beta}_{(j)} - \hat{\boldsymbol{\beta}}_{(j)}), \sigma_j^2 - \boldsymbol{\sigma}'_{j(j)}\Sigma_{j(j)}^{-1}\boldsymbol{\sigma}_{j(j)}),\ \beta_j \ge b(\boldsymbol{\beta}_{(j)},X,\mathbf{y}),$$

es decir, las distribuciones condicionadas unidimensionales serán normales truncadas; la generación de valores simulados de esta distribución puede llevarse a cabo fácilmente siguiendo el algoritmo propuesto en Devroye (1986), p. 380.

En los modelos de producción, es habitual medir las variables en términos de logaritmos, por lo que aparecen con frecuencia valores negativos para las covariables; si no imponemos la hipótesis de que las variables explicativas sean positivas, la distribución condicionada sería la misma, cambiando el recorrido del parámetro, que sería de la forma $b_1(\boldsymbol{\beta}_{(j)}, X, \mathbf{y}) \geq \beta_j \geq b_2(\boldsymbol{\beta}_{(j)}, X, \mathbf{y})$, donde

$$b_1(\boldsymbol{\beta}_{(j)}, X, \mathbf{y}) = \min_{i \,/\, x_{ij}<0} \left\{ (y_i - \mathbf{x}'_{i(j)}\boldsymbol{\beta}_{(j)}) \big/ x_{ij} \right\} \text{ y}$$

$$b_2(\boldsymbol{\beta}_{(j)}, X, \mathbf{y}) = \max_{i \,/\, x_{ij}>0} \left\{ (y_i - \mathbf{x}'_{i(j)}\boldsymbol{\beta}_{(j)}) \big/ x_{ij} \right\}$$

(obsérvese que si algún x_{ij} es cero, entonces la observación *i*-ésima no interviene en el conjunto de restricciones que debe verificar el parámetro β_j). Este cambio sólo supone una ligera dificultad adicional en la programación del algoritmo de Gibbs. No obstante, siempre podremos efectuar cambios de escala en las variables explicativas, de forma que tomen valores superiores a 1, y así sus logaritmos siempre serán positivos.

El algoritmo de Gibbs para nuestro modelo podría formularse por tanto como se especifica a continuación. Dados los valores $\boldsymbol{\beta}^{\mathbf{m}}$ y $(\sigma^2)^m$, obtenidos en la etapa *m*-ésima, los valores generados en la etapa siguiente vendrían dados por:

1. Generar un valor v^{m+1} de la distribución $\sigma^{-2} \,|\, \boldsymbol{\beta}^{\mathbf{m}}, \mathbf{y}, \mathbf{X}$. Tomar $(\sigma^2)^{m+1} = 1/v^{m+1}$.
2. Generar un valor β_j^{m+1} de la distribución
$$\beta j \,|\, \beta_1^{m+1}, ..., \beta_{j-1}^{m+1}, \beta_{j+1}^{m}, ..., \beta_k^{m}, (\sigma^2)^{m+1}, \mathbf{y}, \mathbf{X}, \;\; j = 1, ..., k.$$
3. $\boldsymbol{\beta}^{\mathbf{m+1}} = (\beta_1^{m+1}, ..., \beta_k^{m+1})'$.

Como valor inicial $\boldsymbol{\beta}^{\mathbf{0}}$ un buen candidato sería el estimador máximo-verosímil del modelo, aunque para obtenerlo es necesario resolver numéricamente un problema de optimización restringido, que en algunas ocasiones resulta complicado aún contando con software específico para tal tarea. Otra opción que hemos ensayado con éxito es modificar una de las componentes del estimador de mínimos cuadrados sin restricciones para que el punto así obtenido pertenezca a la frontera de la región factible. Para ello, basta sustituir, por ejemplo, la primera componente de $\hat{\boldsymbol{\beta}}$ por $\hat{\beta}_1^* = \max_i \left\{ (y_i - \mathbf{x}'_{i(1)}\hat{\boldsymbol{\beta}}_{(1)}) \big/ x_{i1} \right\}$.

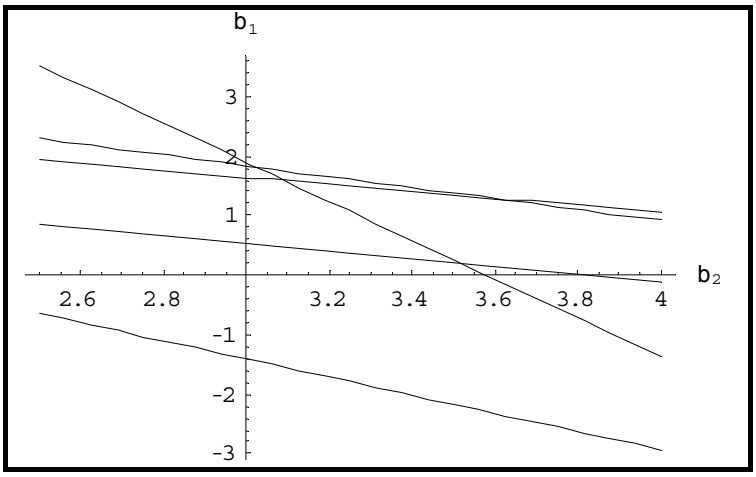
5. SIMULACIONES

En primer lugar, hemos llevado a cabo algunos ejemplos simulados simples que nos ayudan a comprender la formulación del modelo y nos ofrecen una primera evaluación de los resultados obtenidos a través del método de estimación que hemos propuesto. A pesar de que el modelo estudiado presenta complicaciones, la estimación bayesiana con ayuda del algoritmo de Gibbs, ofrece una solución relativamente fácil de llevar a la práctica y, como veremos, en general proporciona resultados satisfactorios.

Ejemplo 1: Para entender mejor la naturaleza del problema y las complicaciones que suponen las restricciones, comenzaremos con un modelo sencillo, en el que la regresión sólo tendrá una variable explicativa (más un parámetro de ordenada en el origen), y obtendremos una muestra simulada de tamaño $n=5$. El modelo de regresión será por tanto de la forma $y_i=\beta_1+\beta_2 x_i+\varepsilon_i,\quad i=1,...,n,$ donde asumimos las hipótesis reseñadas en el epígrafe 2. Hemos simulado una muestra de tamaño 5 de un modelo Uniforme en [0,4] para la variable X. A partir de estos valores, hemos simulado la muestra correspondiente de la variable Y para $\beta_1=2,\ \beta_2=3$ y $\sigma=2.$ La muestra resultante se recoge en el apéndice.

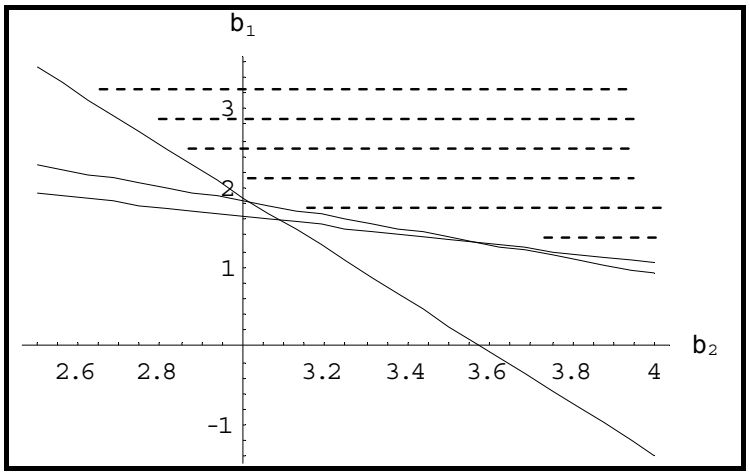
Las restricciones que impone el modelo son $y_i\le\beta_1+\beta_2 x_i,\ i=1,...,n$ o, equivalentemente $\beta_1\ge y_i-\beta_2 x_i,\ i=1,...,n.$ La representación gráfica de las rectas $\beta_1=y_i-\beta_2 x_i,\ i=1,...,n,$ (donde el eje OY corresponde a β_1 y el OX a β_2) se muestra en la figura 1. Los puntos que verifican la desigualdad correspondiente a cada una de las observaciones son los que se sitúan en el semiplano superior de cada una de las rectas, por lo que la región factible sería la intersección de estos cinco semiplanos. Como podemos apreciar en la figura, en este ejemplo de las cinco restricciones dos resultan redundantes, concretamente las correspondientes a las observaciones primera y quinta.

Figura 1.
Restricciones sobre los parámetros (ejemplo 1).



Una vez eliminadas estas dos rectas, mostramos en la figura 2 la representación gráfica de la región factible para el problema.

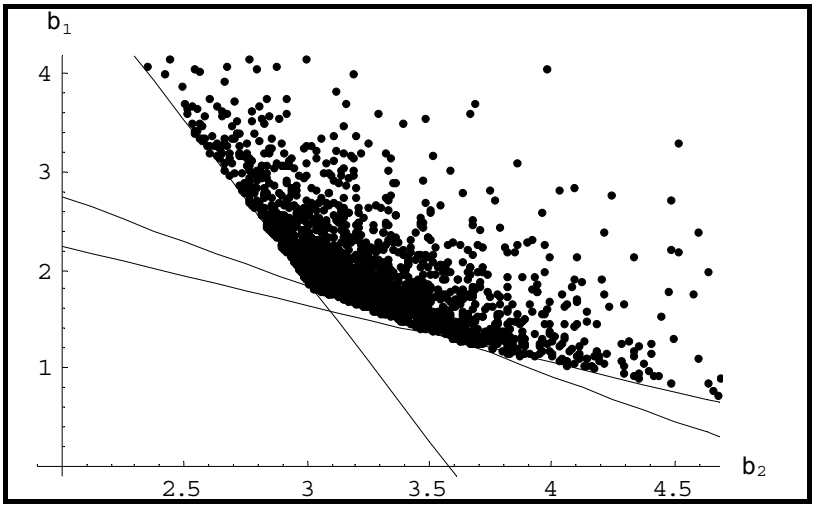
Figura 2.
Región Factible resultante (ejemplo 1).



Para esta muestra, el estimador de mínimos cuadrados no restringido resulta ser $\hat{\beta}_1 = 0.679517$ y $\hat{\beta}_2 = 3.158500,$ que no pertenece a la región factible y que, como podemos apreciar, se aleja bastante del verdadero valor del parámetro en el caso de la ordenada en el origen. Como vimos en el epígrafe 4, tenemos que $\boldsymbol{\beta} \mid \sigma, \mathbf{X}, \mathbf{y} \sim N_2(\hat{\boldsymbol{\beta}}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}),\ \boldsymbol{\beta} \in B,$ donde B representa el conjunto de puntos pertenecientes a la región factible. Así, la distribución a posteriori marginal de $\boldsymbol{\beta}$ concentrará una mayor masa de probabilidad en la zona de la frontera de la región factible más cercana al punto definido por las coordenadas de $\hat{\boldsymbol{\beta}}$. En la figura 3 ofrecemos el diagrama de dispersión correspondiente a una muestra simulada de la

distribución de $\boldsymbol{\beta} \mid \sigma, \mathbf{X}, \mathbf{y}$ de tamaño 2000, donde podemos apreciar cómo los puntos se acumulan efectivamente en la zona de mayor probabilidad. A partir de dicha muestra, las estimaciones bayesianas de los parámetros resultan ser $\hat{\beta}_1^B = 2.03414$ y $\hat{\beta}_2^B = 3.29046$.

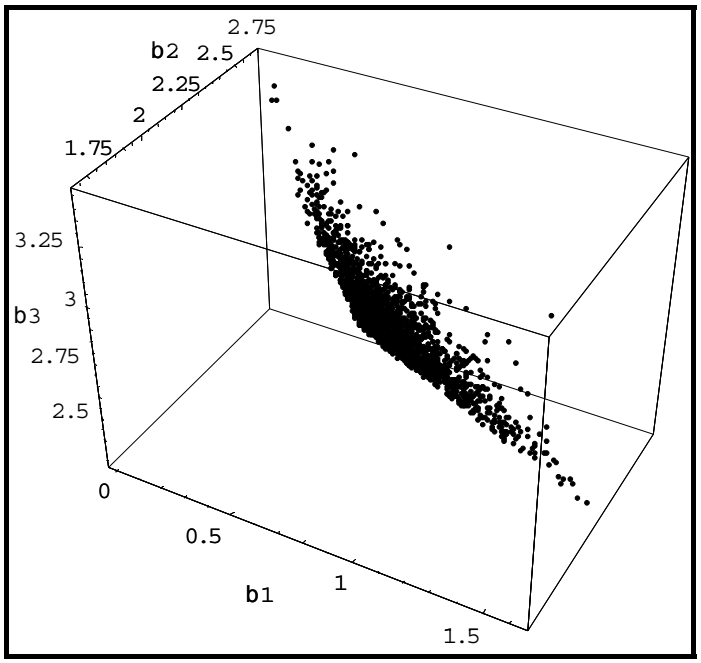
Figura 3.
Muestra simulada de la distribución a posteriori (ejemplo 1).



Ejemplo 2: En este segundo ejemplo, consideraremos un modelo con dos variables explicativas y una ordenada en el origen, para el que obtendremos una muestra simulada de tamaño $n = 50$. El modelo de regresión será por tanto de la forma $y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i, \quad i = 1,...,n,$ donde asumimos las hipótesis reseñadas en el epígrafe 2. Hemos simulado muestras de tamaño 50 de modelos Uniforme en [0,3] y [0,2] para las variables X_2 y X_3 respectivamente . A partir de estos valores, hemos simulado la muestra correspondiente de la variable Y para $\beta_1 = 1,$ $\beta_2 = 2,$ $\beta_3 = 3$ y $\sigma = 3.$ La muestra resultante se recoge en el apéndice.

Para estos datos, el estimador de mínimos cuadrados sin restricciones resulta ser el vector $\hat{\boldsymbol{\beta}} = (-1.8687, 2.3293, 2.9546)'$, que no pertenece a la región factible y que se aleja bastante del verdadero valor del parámetro en el caso de la ordenada en el origen. Aplicando el algoritmo de Gibbs, hemos obtenido una muestra simulada de la distribución de $\boldsymbol{\beta} \mid \sigma, \mathbf{X}, \mathbf{y}$ de tamaño 2000, cuyo diagrama de dispersión se ofrece en la figura 4, en la que podemos apreciar de forma aproximada cuál sería la zona de mayor probabilidad de la región factible. A partir de dicha muestra, las estimaciones bayesianas de los parámetros resultan ser $\hat{\boldsymbol{\beta}}^{\mathbf{B}} = (0.8336, 2.2042, 2.8547)'$.

Figura 4.
Muestra simulada de la distribución a posteriori (ejemplo 2).



A continuación, llevamos a cabo un estudio tipo Monte Carlo, en el que para distintos valores de *n*, hemos generado 1000 muestras de tamaño *n* de nuestro modelo a las que hemos aplicado el Algoritmo de Gibbs (con muestras de tamaño *m* = 2000), con el objetivo de estudiar el sesgo y el Error Cuadrático Medio (ECM) de los estimadores, la longitud de los intervalos simétricos de probabilidad 0.95 (Long.) , así como la proporción de aciertos de dichos intervalos (Ac.), es decir, la probabilidad de cubrimiento en sentido frecuencialista de los intervalos simétricos bayesianos. Hemos utilizado un modelo con una variable explicativa más una ordenada en el origen, fijando los valores $\beta_1 = \beta_2 = \sigma = 1$. Este procedimiento se ha hecho dos veces; en la primera, los valores de la variable X han sido simulados como realizaciones independientes de un modelo $U(0,1)$; en la segunda, se ha utilizado un modelo $U(0,5)$, con el objeto de analizar la posible influencia de la dispersión en los datos de las variables explicativas. Los resultados obtenidos se ofrecen en las tablas 1 y 2 respectivamente.

Tabla 1.
Propiedades de los estimadores y los intervalos probabilísticos
($\beta_1 = \beta_2 = \sigma = 1,\ X_i \sim U(0,1)$).

	β_1				β_2			
	Sesgo	**ECM**	**Long.**	**Ac.**	**Sesgo**	**ECM**	**Long.**	**Ac.**
$n = 5$	–0.109178	0.266588	1.496830	0.827	–0.037798	0.776655	2.750790	0.843
$n = 10$	–0.044969	0.112170	1.113830	0.906	0.022238	0.318605	2.022170	0.916
$n = 15$	–0.027264	0.082346	0.969248	0.928	–0.003651	0.219527	1.715920	0.933
$n = 20$	–0.004177	0.028785	0.596147	0.948	–0.007213	0.090985	1.093980	0.951
$n = 50$	0.000192	0.005300	0.250071	0.950	–0.002060	0.016951	0.470160	0.948
$N = 100$	–0.002023	0.002097	0.150775	0.949	0.000871	0.006734	0.283299	0.950

En cuanto al sesgo, podemos apreciar que los valores son bastante cercanos a 0, salvo en el caso del parámetro β_1 con tamaño muestral $n = 5,$ en el que el sesgo se sitúa, en valor absoluto, en torno al 11% con respecto al verdadero valor del parámetro. Cabe destacar también la persistencia del signo negativo en el caso de β_1 para casi todos los tamaños muestrales (aunque con valores absolutos muy cercanos a cero a partir de $n = 20),$ hecho que no se manifiesta para el parámetro β_2.

Con respecto al ECM y la longitud de los intervalos, se obtienen valores bastante razonables, observándose en general el comportamiento esperado de descenso en ambas cantidades a medida que aumenta el tamaño muestral.

Por otra parte, podemos apreciar cómo la probabilidad de cubrimiento en sentido frecuencialista de los intervalos bayesianos es cercana a la probabilidad nominal de los mismos (0.95) sólo para tamaños muestrales a partir de 20. Para muestras pequeñas ($n = 5$ y $n = 10),$ vemos que la probabilidad de cubrimiento es sistemática y apreciablemente inferior a la probabilidad nominal. El estudio simulado parece mostrar que en este modelo y con la distribución a priori propuesta, los intervalos bayesianos se comportan también como intervalos clásicos asintóticamente, no ocurriendo lo mismo para pequeñas muestras.

Tabla 2.
Propiedades de los estimadores y los intervalos probabilísticos
($\beta_1 = \beta_2 = \sigma = 1,\ X_i \sim U(0,5)$).

	β_1				β_2			
	Sesgo	**ECM**	**Long.**	**Ac.**	**Sesgo**	**ECM**	**Long.**	**Ac.**
n = 5	–0.113131	0.207657	1.311950	0.824	–0.003060	0.028518	0.530942	0.830
n = 10	–0.047413	0.105605	1.071350	0.920	0.006306	0.012793	0.390117	0.904
n = 15	–0.015787	0.117725	1.308470	0.946	0.002587	0.011762	0.433535	0.940
n = 20	–0.009287	0.015468	0.447562	0.944	–0.000373	0.002523	0.194748	0.945
n = 50	–0.011658	0.006770	0.273231	0.940	0.002456	0.000540	0.084527	0.944
n = 100	–0.000481	0.001393	0.133449	0.952	0.000430	0.000176	0.050188	0.960

El efecto de la variabilidad de los datos de X se manifiesta de manera clara en el ECM y la longitud del intervalo para el parámetro β_2; como podemos apreciar, ambas magnitudes disminuyen significativamente cuando aumentamos la dispersión en la variable explicativa. También se produce una disminución de estas cantidades para β_1, si bien en este caso los cambios son mucho menos importantes.

El comportamiento del sesgo del parámetro β_1, así como el de la probabilidad de cubrimiento de los intervalos para ambos parámetros, llevan a pensar en la posibilidad de introducir correcciones al método de estimación propuesto para el caso de pequeñas muestras, tema que será objeto de futuras investigaciones.

6. UN EJEMPLO ILUSTRATIVO

En esta sección, vamos a aplicar el método de estimación descrito a un modelo de producción, cuyos datos se analizan en Whiteman and Pearson (1993) y en Coelli et al (1998), pág. 192. Los datos hacen referencia a empresas de telecomunicación en 21 países en 1990. Se considera un solo output *y* (basado en los ingresos) y dos inputs: el factor capital x_1 (medido a través de los kilómetros de línea) y el factor trabajo x_2 (número de empleados). Los datos considerados se ofrecen en el apéndice.

En esta situación, estimamos un modelo de producción tipo Cobb-Douglas, es decir, $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, donde las variables y, x_1, x_2 están medidas en logaritmos y la perturbación $-\varepsilon$ sigue una distribución half-Normal (0, σ^2), $\sigma > 0$.

Las estimaciones de los parámetros, junto con su desviación típica e intervalos bayesianos simétricos de probabilidad 0.95 se ofrecen en la tabla 3. Como punto de partida para el algoritmo de Gibbs se ha utilizado el estimador de mínimos cuadra-

dos no restringido, al que hemos modificado la primera componente para obligarlo a pertenecer a la frontera de la región factible. Hemos generado una muestra de tamaño 10000 de la distribución a posteriori conjunta, a partir de la cual hemos obtenido las estimaciones ofrecidas en la tabla 3.

Tabla 3.
Resultados de la estimación del modelo en telecomunicaciones.

Parámetro	Estimación	Desv. Típica	Intervalo
β_0	–5.841250	0.663918	(–6.78780, –4.31347)
β_1	0.932661	0.209273	(0.461015, 1.283140)
β_2	–0.082434	0.194559	(–0.429136, 0.345061)
σ^2	0.377713	0.134278	(0.197077, 0.709509)

Una vez estimados los parámetros, uno de los objetivos importantes de este tipo de modelos es la estimación de una medida de eficiencia-ineficiencia en cada una de las observaciones (en este caso, en cada uno de los países considerados). En nuestro modelo de producción de tipo Cobb-Douglas, es habitual medir la eficiencia de la observación *i*-ésima a través de la expresión $\exp\{\hat{\varepsilon}_i\} = \exp\left\{y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}\right\}$, que puede tomar valores en el intervalo $(0,1]$, correspondiendo el valor 1 a la máxima eficiencia. Los resultados correspondientes se ofrecen en la tabla 4. Observemos que España se sitúa en penúltimo lugar, tan sólo por debajo de Turquía, con un coeficiente de eficiencia muy por debajo de países como Inglaterra, Suecia, Dinamarca, Noruega u Holanda.

Tabla 4.
Eficiencia estimada para cada uno de los países.

País	Eficiencia est.	País	Eficiencia est.
Suiza	0.943287	Francia	0.653460
Reino Unido	0.908947	Irlanda	0.648132
Suecia	0.873255	Bélgica	0.608050
Dinamarca	0.857525	Italia	0.567722
Holanda	0.789518	Alemania	0.529563
Noruega	0.781926	Portugal	0.514585
Australia	0.740416	Japón	0.500081
Finlandia	0.697198	Austria	0.479902
Islandia	0.686110	España	0.372007
Canadá	0.681828	Turquía	0.156290
Nueva Zelanda	0.660958		

Para finalizar, destacamos que en Coelli et al (1998) se estiman los parámetros planteando un modelo de producción con frontera estocástica, siendo todos los resultados obtenidos bastante similares a los que hemos presentado en base al modelo de frontera determinista (téngase en cuenta que, puesto que hemos considerado unidades distintas para las covariables, las estimaciones de las ordenadas en el origen no son directamente comparables). De hecho, en la pág, 198 se indica que "These results indicate that the vast majority of residual variation is due to the inefficiency effect, u_i, and that the random error, v_i, is approximately zero."

7. CONCLUSIONES

La conclusión principal que podemos obtener a partir de este trabajo es que el uso de las técnicas Bayesianas permite resolver con cierta comodidad un problema bastante complicado, para el que aún no existe una respuesta satisfactoria desde el punto de vista de la estimación máximo-verosímil. En particular, hemos comprobado que, aunque el tratamiento de la distribución a posteriori conjunta es complicado, las condicionadas unidimensionales siguen distribuciones conocidas y que pueden ser simuladas con facilidad (distribuciones normales truncadas para el caso de los parámetros de localización y gamma para el caso del inverso de la varianza de las perturbaciones); este hecho nos permite formular de forma satisfactoria el algoritmo de Gibbs para este problema.

A través de los ejemplos de las secciones 5 y 6, en una primera aproximación, podemos ver cómo los resultados en general son aceptables. El estudio de simulación efectuado, incide en esta misma apreciación inicial, con la salvedad de que en un futuro debemos plantear la posibilidad de introducir correcciones para tamaños muestrales pequeños.

Otra generalización interesante a este artículo sería considerar que las perturbaciones (cambiadas de signo), sigan un modelo más flexible que incluya a la distribución half-normal como caso particular. Una posibilidad es considerar una distribución normal truncada a partir de un nuevo parámetro desconocido $\gamma \geq 0$. Otra opción tal vez más interesante es considerar la familia triparamétrica skew-normal introducida en Azzalini (1985), que incluye a familia normal como caso particular. La distribución half-normal puede obtenerse también como una distribución límite especial de la familia skew-normal (Pewsey, 2002).

REFERENCIAS BIBLIOGRÁFICAS

AIGNER, D.J. and CHU, S.F. (1968): "On Estimating the Industry Production Function", *American Economic Review*, 58, pp. 826-839.

ARNOLD, B.C.; CASTILLO, E. and SARABIA, J.M. (1999): *Conditional Specification of Statistical Models*, New York: Springer Series in Statistics, Springer-Verlag.

AZZALINI, A. (1985): "A Class of Distributions Which Includes the Normal Ones", *Scandinavian Journal of Statistics*, 12, pp. 171-178.

BERGER, J.O. and BERNARDO, J.M. (1992): "On the Development of reference priors", *Bayesian Statistics 4*, Oxford: University Press, pp. 35-60.

BERNARDO, J.M. (1979): "Reference Posterior Distributions for Bayesian Inference" (with discussion), *Journal of the Royal Statistical Society Series B,* 41, pp. 113-147.

BERNARDO, J.M. (2005): "Reference Analysis", *Handbook of Statistics* 25, Amsterdam: Elsevier, pp. 17-90.

BERNARDO, J.M. (2009): "Modern Bayesian Inference: Foundations and Objective Methods", *Philosophy of Statistics*, Amsterdam: Elsevier (to appear). Disponible en: http://www.uv.es/~bernardo/2009PhilScience.pdf [20/04/2009].

CAZALS, C.; FLORENS J.P. and SIMAR, L. (2002): "Nonparametric Frontier Estimation: a Robust Approach", *Journal of Econometrics*, 106, pp. 1-25.

COELLI, T.; PRASADA RAO, D.S. and BATTESE, G.E. (1998): *An Introduction to Efficiency and Productivity Analysis*, Boston: Kluwer Academic Publishers.

DAOUIA, A. and SIMAR, L. (2004): "Nonparametric Efficiency Analysis: a Multivariate Conditional Quantile Approach", *Journal of Econometrics,* 140, pp. 375-400.

DATTA, G.S. and GOSH, J.K. (1995): "On priors Providing Frequentist Validity for Bayesian Inference", *Biometrika*, 82, pp.37-45.

DATTA, G.S. and SWEETING, T.J. (2005): "Probability Matching Priors", *Handbook of Statistics* 25, Amsterdam: Elsevier, pp. 91-114.

DEVROYE, L. (1986): *Non-Uniform Random Variate Generation*, New-York: Springer-Verlag.

FØRSUND, F.R.; NOVELL, C.A.K. and SCHMIDT, P. (1980): "A Survey of Frontier Production Functions and of their Relationship to Efficiency Measurement", *Journal of Econometrics,* 13, pp. 5-25.

GARCÍA, J.J.; CASTILLA, D. and JIMÉNEZ, R. (2004): "Determination of Technical Efficiency of Fisheries by Stochastic Frontier Models: a Case on the Gulf of Cádiz (Spain)", *Journal of Marine Science, 61*, pp. 416-421.

GELFAND, A.E. and SMITH, A.F. (1990): "Sampling-Based Approaches to Calculating Marginal Densities", *Journal of the American Statistical Association,* 85, pp. 398-409.

GREENE, W.H. (1993): "The Econometric Approach to Efficiency Analysis", *Measurement of Productive Efficiency: Techniques and Applications,* Chapter 2, pp. 68-109, New York: Oxford University Press.

GREENE, W.H. (1998): *Análisis Econométrico (3nd ed.)*, Madrid: Prentice-Hall.

JEFFREYS, H. (1946): "An Invariant Form for the Priori Probability in Estimation Problems", *Proceedings of the Royal Society of London, Series A,* 186, pp. 453-461.

JEFFREYS, H. (1961): *Theory of Probability (3nd ed.)*, London: Oxford University Press.

JOHNSON, N.; KOTZ, S. and BALAKRISHNAN, N. (1994): *Continuous Univariate Distributions, Vol.1 (2nd ed.)*, New York: Wiley.

JONDROW, J.; NOVELL, C.A.K.; MATEROV, I.S. and SCHMIDT, P. (1982): "On Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model", *Journal of Econometrics*, 19, pp. 233-238.

KASS, R.E. and WASSERMAN, L., (1996): "The Selection of prior Distributions by Formal Rules", *Journal of the American Statistical Association*, 91, pp. 1343-1370.

ORTEGA, F.J. y BASULTO, J. (2003): "Distribuciones a priori unidimensionales en Modelos No Regulares", *Estadística Española*, 45 (154), pp. 363-383.

PEWSEY, A. (2002): "Large-Sample Inference for the General Half-Normal Distribution", *Communications in Statistics-Theory and Methods,* 31, pp. 1045-1054.

PEWSEY, A. (2004): "Improved likelihood based inference for the general half-normal distribution", *Communications in Statistics-Theory and Methods,* 33, pp. 197-204.

SIMAR, L. (2007): "How to Improve the Performances of DEA/FDH Estimators in the Presence of Noise?", *Journal of Productivity Analysis,* 28, pp. 183-201.

WHITEMAN, J. and PEARSON, K. (1993): "Benchmarking Telecommunications Using Data Envelopment Análisis", *Australian Economic Papers*, 12, pp. 97-105.

APÉNDICE: DATOS UTILIZADOS EN LOS EJEMPLOS

Muestra de datos simulados (Ejemplo 1)

Y	2.442977	11.695199	4.580773	3.400216	3.165631
X	0.636940	3.267755	0.915114	0.585950	1.523870

Muestra de datos simulados (Ejemplo 2)

Y	{4.62587, 5.9591, 3.64545, 4.53587, 7.06785, 5.65241, 1.79605, 6.35308,1.43932, 3.22241, 5.39405, 9.54582, 3.28809, 10.0456, 0.657027, –0.733859, –0.346423, –3.22865, 4.15841, 7.12289, 7.23358, 4.07836, 9.02886, 4.32932,1.42179, 2.12037, 6.38586, 4.89608, 6.49352, 4.5048,4.45847, 2.18042,0.363753, 8.76576, 2.47284, 1.52509, 8.58874, 6.01232, 3.79583, 9.16072,6.79729, 6.35966, 5.2574, -3.75637, 11.8284, 4.74722, 6.67094, 2.21107,7.35288, –1.47465}
X_2	{1.80846, 0.192743, 1.28582, 1.843, 1.46173, 1.88556, 1.79588, 1.01287,1.9491, 0.313742, 2.36173, 2.77974, 0.505303, 2.38196, 0.378946, 0.796419, 0.70683, 0.0737809, 1.33566, 1.68444, 2.28006, 1.26077, 1.80297, 0.112218 ,0.471599, 1.06803, 0.517146, 1.26922, 2.00987, 2.18247, 1.72126, 0.256345, 0.0607656, 1.86873, 2.35953, 0.476604, 2.55546, 2.48677, 1.98058, 2.68019, 1.84863, 2.41299, 0.644927, 0.995741, 2.56858, 1.15222, 1.84196, 0.883523, 2.09698, 0.0841912}
X_3	{0.883209, 1.74287, 0.0580738, 0.601147, 1.7357, 1.57197, 0.0175634, 1.35533, 0.16268, 1.25424, 0.313922, 1.69748, 0.84229, 1.46745, 1.0815, 0.0888206, 0.412339, 0.803619, 1.36912, 1.32067, 1.18437, 0.214603, 1.97113, 1.26455, 0.301158, 0.471734, 1.91306, 0.663401, 0.565458, 0.899761, 1.89549, 1.30807, 0.402779, 1.64552, 1.58157, 1.61059, 1.56049, 0.178079, 0.500071, 1.52177, 1.14815, 1.37446, 1.13096, 0.201098, 1.96378, 1.15986, 1.15983, 0.93655, 1.66262, 0.688123}

Datos de producción e inputs en telecomunicaciones (1990)

País	Índice Prod.	Líneas (10^3 km)	Empleados (10^3 personas)
Alemania	1.73	2998.1	212
Australia	0.74	776.7	85
Austria	0.24	322.3	18
Bélgica	0.36	399.0	26
Canadá	1.26	1529.6	105
Dinamarca	0.39	291.1	18
España	0.59	1260.3	75
Finlandia	0.29	267.0	20
Francia	2.06	2808.5	156
Irlanda	0.11	98.3	13
Islandia	0.02	12.6	2
Italia	1.48	2235.0	118
Japón	2.73	5323.6	277
Holanda	0.77	694.0	32
Nueva Zelanda	0.16	147.3	17
Noruega	0.27	213.2	15
Portugal	0.19	237.9	23
Reino Unido	2.53	2540.4	227
Suecia	0.71	584.9	42
Suiza	0.56	394.3	22
Turquía	0.15	689.3	36