

EL PROCESO DE DIFUSIÓN LOGNORMAL PARA MODELIZAR PATRONES DE CRECIMIENTO

Nuria Rico Castro¹, Desirée Romero Molina²

Departamento de Estadística e Investigación Operativa, Universidad de Granada.
Facultad de Ciencias, Avda. Fuentenueva sn. 18071, Granada, España.

ABSTRACT

This chapter describes two models: the homogeneous lognormal diffusion process and the non homogeneous lognormal diffusion process with particular interest in the case the exogenous factor is polynomial type. These models are presented as powerful tools in the study of observed paths that grow exponentially over time: An overview of them is provided, starting from its definition, reviewing the main characteristics and dealing with some relevant aspects on the estimation of parameters. In this regard, we discuss the use of a family of parametric functions, which contain as particular cases, the mean function, median and quantiles, along with their conditional versions. These functions allow us model and forecast. We illustrate the estimation procedure with an example, using the model to the amount of emissions of methane to the atmosphere for a time period.

KEY WORDS: Temporal evolution, exponential growth, parametric functions, estimation prediction

RESUMEN

En este capítulo se describen dos modelos: el proceso de difusión lognormal homogéneo y el proceso de difusión lognormal no homogéneo prestando especial interés al caso en el que el factor exógeno es de tipo polinómico. Estos modelos se presentan como potentes herramientas en el estudio de trayectorias observadas que presenten un crecimiento exponencial a lo largo del tiempo: Se ofrece una visión general sobre ellos, partiendo de su definición, revisando sus características principales y tratando algunos aspectos relevantes sobre la estimación de parámetros. En este aspecto, se expone el uso de una familia de funciones paramétricas, las cuales contienen, como casos particulares, a la función media, moda y de cuantiles, junto con sus versiones condicionadas. Estas funciones permiten la modelización así como la realización de predicciones. Se ilustra el procedimiento de estimación con un ejemplo, aplicando los modelos a la serie de emisiones de metano a la atmósfera durante un periodo de tiempo.

PALABRAS CLAVE: Evolución temporal, Crecimiento exponencial, Funciones paramétricas, Estimación, Predicción.

1. INTRODUCCIÓN

Dado un conjunto de observaciones de una variable cuantitativa que evoluciona en el tiempo, existen diversos procedimientos enfocados a la búsqueda de un patrón matemático que represente las observaciones, esto es, el ajuste de una curva de tipo determinista a dichos datos. En este tipo de modelo no se considera ninguna componente aleatoria. La elección del modelo determinista en cada situación dependerá de la evolución observada en los datos. Concretamente, si el conjunto de datos mostrara una tendencia exponencial, se podría ajustar una curva determinista exponencial, obteniendo los valores para los parámetros del modelo matemático que mejor representen la trayectoria observada. Sin embargo, dadas las perturbaciones naturales que afectan a las variables cuantitativas, en general los datos mostrarán desviaciones sobre la curva determinista. Esto es lógico, ya que las curvas deterministas son rígidas, son modelos matemáticos totalmente determinados, como su nombre indica, mientras que los datos fácilmente se ven afectados por cuestiones externas o simplemente por ruido. De este hecho surge la idea de utilizar el modelo determinista añadiendo un ruido blanco, es decir, sumando un proceso gaussiano. La solución de la ecuación resultante al añadir el ruido blanco es un proceso de difusión. Así, la parte estocástica que se introduce en el modelo trata de modelar las desviaciones que se observan en los datos con respecto al modelo determinista.

Si se realiza esta operación con el modelo asociado a la curva exponencial, conocido como modelo malthusiano o de Malthus, se origina el proceso lognormal o logarítmico-normal homogéneo. El modelo determinista malthusiano^[17] postula un crecimiento exponencial de la población: Malthus estudia este modelo alarmado por la predicción de un exponencial -y por tanto rápido- crecimiento del número de habitantes junto con un lineal -y por tanto lento- crecimiento de la cantidad de alimentos disponibles. A lo largo de los siglos, se ha comprobado que la teoría malthusiana no es aplicable a la población humana, debido a un gran número de razones. Sin embargo, este modelo teórico se considera de gran utilidad para la modelización del crecimiento de otro tipo de poblaciones^[3], como por ejemplo la reproducción de seres vivos bajo condiciones ideales de falta de depredadores y exceso de espacio y alimento.

¹ nrco@ugr.es

² deromero@ugr.es

Introduciendo ambiente aleatorio en el modelo malthusiano se deriva, pues, el proceso de difusión lognormal^[2].^[4] Este proceso de difusión toma el nombre de *lognormal* puesto que la distribución de la densidad de transición del mismo es de tipo lognormal, aunque también se podría llamar proceso *exponencial* dado que la tendencia del proceso y la curva determinista a la que va asociado es la curva exponencial. Sin embargo, este modelo no solamente ha sido aplicado al estudio del crecimiento de poblaciones biológicas, sino que puede encontrarse en la literatura su aplicación en diversas áreas de conocimiento tales como economía^[1], ingeniería nuclear^[16], ingeniería mecánica^[18] y astrofísica^[23], entre otras.

En algunas ocasiones, además, se da la circunstancia de que hay disponible información sobre ciertas variables externas que influyen en el comportamiento de la variable bajo estudio. El ruido observado en los datos contiene la influencia de dichas variables. En estos casos, sería útil poder incluir en el modelo dicha influencia y que esto permita controlar el ruido o las desviaciones. Si se conocen las variables externas, o algunas posibles, se pueden incluir en el modelo utilizando, en la media infinitesimal del proceso, una función $h(t)$ que incorpore esta información. Este modelo es el llamado proceso de difusión lognormal no homogéneo o con factores exógenos. Los factores exógenos que se incluyen en el modelo suelen ser funciones dependientes del tiempo^{[5], [6]}. El uso de estas funciones permite^[22] un mejor ajuste así como cierto control sobre la evolución del proceso, en la medida en que estas funciones sean conocidas y controlables.

Una de las ventajas que presenta el proceso de difusión lognormal no homogéneo es que, utilizando una función $h(t)$ adecuada, puede servir para modelar no sólo datos con crecimiento exponencial. Así, puede aplicarse para variables que presentan un crecimiento de tipo sigmoidal, como en el proceso tipo Gompertz^{[11], [20]}, donde se toma como función del tiempo $h(t)$ una doble exponencial. También se ha aplicado a datos que muestran una tendencia inicial sigmoidal que cambia, a partir de cierto momento, a una exponencial. En este caso el proceso se denomina proceso mixto Gompertz-lognormal^[19] y $h(t)$ es una función exponencial combinada con una doble exponencial. Estos casos son expresiones concretas de procesos de difusión lognormales no homogéneos.

En general, se aborda el estudio del proceso de difusión lognormal tomando la función que afecta a la media infinitesimal $h(t)$ como combinación lineal de funciones, lo que permite obtener expresiones más sencillas y que no son aplicables a cualquier tipo de función. Sobre este modelo no homogéneo hay numerosos estudios que cubren diversas cuestiones sobre inferencia a partir de muestreo discreto, incluyendo estudios sobre tiempos de primer paso^{[12], [13]} y estimaciones puntuales tanto máximo verosímiles como insesgadas de mínima varianza^{[10], [14], [15]} para las funciones de interés.

Un caso aún más general es el que considera que las variables externas que influyen son desconocidas, o bien no se dispone de información precisa sobre su comportamiento. En este caso se puede utilizar, para la modelización, el proceso no homogéneo introduciendo los factores exógenos como una función de tipo polinómica^{[7], [8], [9]}. En este capítulo nos centramos en este último caso, de especial utilidad cuando no existe información externa disponible.

El capítulo está organizado en dos secciones. En la primera sección se describe, desde el punto de vista formal, el proceso de difusión lognormal homogéneo y se muestran los principales resultados sobre inferencia para tres funciones de interés: la función media, la función moda y la función de cuantiles, considerando también las versiones condicionadas de estas tres funciones. A continuación, se aplican estos resultados a la evolución de las emisiones de metano a la atmósfera durante un periodo de tiempo.

En la segunda sección, se resumen los principales resultados sobre el proceso de difusión lognormal no homogéneo, revisando por una parte la expresión de las funciones y la estimación máximo verosímil en el caso general y por otra parte abordando la inclusión de factores exógenos de tipo polinómico. Este análisis es conveniente para abordar el caso en que se observan desviaciones a la tendencia del proceso lognormal homogéneo y sin embargo no hay disponible ninguna información adicional sobre la variable observada. Por último, se aplican estos resultados para la obtención de un modelo no homogéneo para las emisiones de metano a la atmósfera, a partir de los datos obtenidos^[21] para los años comprendidos entre 1860 y 1994.

2. EL PROCESO DE DIFUSIÓN LOGNORMAL HOMOGÉNEO

En esta sección se describe el proceso de difusión lognormal homogéneo, partiendo de su definición y de la función de densidad de transición. A continuación, se dan las expresiones de las funciones de interés media, moda y de cuantiles, así como de sus versiones condicionadas considerando el caso particular de que la distribución inicial sea degenerada y dando una expresión general para todas ellas en forma de función paramétrica. Seguidamente, se revisa la estimación máximo verosímil de los parámetros y de las funciones paramétricas descritas. Por último, se aplican los resultados para modelar la emisión de metano a la atmósfera durante un periodo de tiempo.

1.1 Definición del proceso de difusión lognormal homogéneo

Desde un punto de vista formal, la definición del proceso de difusión lognormal, en su versión homogénea, puede hacerse a partir de su media y varianza infinitesimales, $A_1(x)$ y $A_2(x)$ respectivamente. Así, se define el proceso de difusión lognormal homogéneo como un proceso de Markov $\{X(t): t_0 \leq t \leq T\}$ en tiempo continuo, con espacio de estados continuo, trayectorias continuas casi seguro y momentos infinitesimales $A_1(x) = mx$ y $A_2(x) = \sigma^2 x^2$, donde $m \in \mathbb{R}$ y $\sigma > 0$.

Para el proceso estocástico así definido, la función de densidad de transición es

$$f(x, t | y, s) = \frac{1}{x\sqrt{2\pi\sigma^2(t-s)}} \exp\left(-\frac{\left[\ln\left(\frac{x}{y}\right) - (t-s)\left(m - \frac{\sigma^2}{2}\right)\right]^2}{2\sigma^2(t-s)}\right)$$

esto es, la ley de probabilidad de la variable $X(t)$ condicionada a que $X(s) = y$, con $s < t$ es de tipo lognormal unidimensional (\mathcal{L}_1), lo cual se puede expresar:

$$X(t) | X(s) = y \sim \mathcal{L}_1\left(\ln(y) + \left(m - \frac{1}{2}\sigma^2\right)(t-s); \sigma^2(t-s)\right).$$

Esta distribución, junto con la distribución inicial, permite el cálculo de las distribuciones finito-dimensionales. Por ejemplo, tomando la distribución inicial degenerada $P[X(t_0) = x_0] = 1$, entonces las distribuciones finito-dimensionales serán de tipo lognormal unidimensional.

En la siguiente sección se dan las expresiones de las funciones media, moda y de cuantiles del proceso, así como las de sus versiones condicionadas. Estas funciones son las más utilizadas para realizar la modelización de trayectorias observadas, así como la predicción de los valores futuros.

1.2 Funciones paramétricas de interés

Para la modelización de trayectorias a partir de este modelo así como para realizar predicciones, existen diversas funciones que pueden ser utilizadas. Entre ellas, se destacan las funciones media, moda y de cuantiles. La función media, $E[X(t)]$, permite la predicción puntual sobre el valor esperado. Su expresión en el proceso de difusión lognormal homogéneo, y para $P[X(t_0) = x_0] = 1$, es la siguiente:

$$E[X(t)] = x_0 \exp(m(t - t_0))$$

La función moda $Mo[X(t)]$ por su parte, permitirá la predicción del valor más probable. La expresión de la función moda para el proceso lognormal homogéneo es

$$Mo[X(t)] = x_0 \exp\left(m(t - t_0) - \frac{3}{2}\sigma^2(t - t_0)\right)$$

La función de cuantiles de orden α , $C_\alpha[X(t)]$, habitualmente se utiliza para obtener un intervalo de predicción, esto es, la estimación de un intervalo en el cual la variable toma valores con determinada probabilidad. La expresión de la función de cuantiles para el proceso de difusión lognormal homogéneo es

$$C_\alpha[X(t)] = x_0 \exp\left(m(t - t_0) - \frac{1}{2}\sigma^2(t - t_0) + z_\alpha \sigma \sqrt{t - t_0}\right)$$

donde z_α es el cuantil de orden α de la distribución normal estándar.

Estas expresiones pueden reescribirse, tomando $a = m - \frac{\sigma^2}{2}$. En la tabla siguiente se tienen las expresiones para las funciones en general y para el caso en que la distribución inicial es degenerada, con esta notación.

En algunas situaciones, en las cuales se dispone únicamente de trayectorias observadas del proceso, la inferencia puede hacerse a través de las versiones condicionadas de las funciones anteriores, esto es, las funciones media condicionada, moda condicionada y de cuantiles condicionada. En la tabla siguiente se resumen las expresiones de estas funciones.

Esta notación sobre las funciones de interés media, moda, de cuantiles, así como sus versiones condicionadas, permiten que las funciones puedan todas escribirse en la forma general

$$\theta(C, A(t, s), B(t, s), l) = C \exp(A(t, s)a + B(t, s)\sigma^l)$$

con $C > 0$, $A(t, s) \in \mathbb{R}$, $B(t, s) \in \mathbb{R}y$ $l \in \mathbb{N}$ tal y como se resume en la tabla 3.

Función	Expresión general	Caso de distribución inicial degenerada $P[X(t_0) = x_0] = 1$
Media: $E[X(t)]$	$M(t) = E[X(t_0)] \exp\left((t - t_0)a + \frac{1}{2}(t - t_0)\sigma^2\right)$	$x_0 \exp\left((t - t_0)a + \frac{1}{2}(t - t_0)\sigma^2\right)$
Moda: $Mo[X(t)]$	$Mo(t) = E[X(t_0)] \exp\left((t - t_0)a - (t - t_0)\sigma^2\right)$	$x_0 \exp\left((t - t_0)a - (t - t_0)\sigma^2\right)$
Cuantil de orden α : $Cuantil_\alpha[X(t)]$	$C_\alpha(t) = E[X(t_0)] \exp\left((t - t_0)a + z_\alpha \sqrt{(t - t_0)}\sigma\right)$	$x_0 \exp\left((t - t_0)a + z_\alpha \sqrt{(t - t_0)}\sigma\right)$

Tabla 1. Expresiones de las funciones media, moda y de cuantiles del proceso de difusión lognormal homogéneo.

La inferencia sobre las funciones de interés puede llevarse a cabo considerando esta clase de funciones θ , que contiene las funciones de interés como casos particulares. Así, obteniendo los estimadores de máxima verosimilitud de los parámetros a y σ^2 , se tendrá $\hat{\theta}$ el estimador de máxima verosimilitud de θ de forma sencilla, aplicando el principio de invarianza funcional.

Función	Expresión
Media condicionada: $E[X(t) X(s) = x_s]$	$x_s \exp\left((t - s)a + \frac{1}{2}(t - s)\sigma^2\right)$
Moda condicionada: $Mo[X(t) X(s) = x_s]$	$x_s \exp\left((t - s)a - (t - s)\sigma^2\right)$
Cuantil de orden α condicionada: $Cuantil_\alpha[X(t) X(s) = x_s]$	$x_s \exp\left((t - s)a + z_\alpha \sqrt{(t - s)}\sigma\right)$

Tabla 2. Expresiones de las funciones media, moda y de cuantiles condicionadas del proceso de difusión lognormal homogéneo.

Función	C	$A(t, s)$	$B(t, s)$	l
Media	x_0	$(t - t_0)$	$\frac{1}{2}(t - t_0)$	2
Media condicionada	x_s	$(t - s)$	$\frac{1}{2}(t - s)$	2
Moda	x_0	$(t - t_0)$	$-(t - t_0)$	2
Moda condicionada	x_s	$(t - s)$	$-(t - s)$	2
Cuantil de orden α	x_0	$(t - t_0)$	$z_\alpha \sqrt{(t - t_0)}$	1
Cuantil de orden α condicionada	x_s	$(t - s)$	$z_\alpha \sqrt{(t - s)}$	1

Tabla 3. Valores de los argumentos de la función θ para las funciones media, moda y de cuantiles del proceso de difusión lognormal homogéneo, así como sus versiones condicionadas.

1.3 Estimación máximo verosímil

Consideremos los instantes de tiempo t_1, t_2, \dots, t_n en los cuales se observa el proceso, esto es, las variables $X(t_1), X(t_2), \dots, X(t_n)$ obteniéndose los valores x_1, x_2, \dots, x_n . Se define la siguiente transformación de los valores

observados $v_i = \frac{\ln(x_i) - \ln(x_{i-1})}{\sqrt{t_i - t_{i-1}}}$ para $i = 2, \dots, n$, de forma que a partir de ellos se obtienen los estimadores de máxima verosimilitud de los parámetros desconocidos a y σ^2 como

$$\hat{a} = \frac{1}{t_n - t_1} \ln\left(\frac{x_n}{x_1}\right) \quad \text{y} \quad \hat{\sigma}^2 = \frac{1}{n-1} (v - \hat{a}u)'(v - \hat{a}u)$$

donde $v = (v_1, \dots, v_n)'$ y $u = (\sqrt{t_2 - t_1}, \dots, \sqrt{t_n - t_{n-1}})'$.

Estos estimadores son independientes, suficientes y completos, y sus distribuciones son

$$\hat{a} \sim N\left(a, \frac{\sigma^2}{t_n - t_{n-1}}\right) \quad \text{y} \quad \frac{(n-1)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-2}^2.$$

A partir de estos estimadores, se obtienen los estimadores de máxima verosimilitud de la familia de funciones θ , como

$$\hat{\theta}(C, A(t, s), B(t, s), l) = C \exp(A(t, s)\hat{a} + B(t, s)\hat{\sigma}^l).$$

1.4 Aplicación a los datos sobre emisiones de metano a la atmósfera

Stern y Kaufman presentan en 1998 un estudio sobre las emisiones globales de metano producidas por el hombre en el periodo de tiempo comprendido entre 1860 y 1994. En este estudio, los autores proporcionan la estimación para este periodo de tiempo sobre el total de las emisiones producidas por el hombre, así como para siete componentes. El propósito de este estudio tenía en cuenta las estimaciones del Intergovernmental Panel on Climate Change.

La figura 1 representa la serie cronológica propuesta, donde se puede observar una tendencia de tipo exponencial.

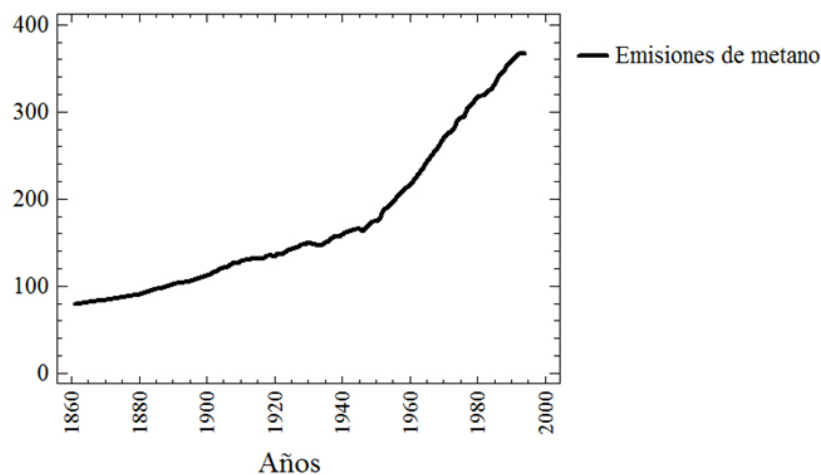


Figura 1: Evolución de las emisiones humanas de metano a la atmósfera (Tg) en los años 1860-1994.

Esta gráfica muestra la evolución de emisiones de metano a la atmósfera, medidas en teragramos ($1Tg = 10^{12}gr$) a lo largo de los años, mostrando los datos con un crecimiento de tipo exponencial, por lo que el proceso de difusión lognormal homogéneo puede ser adecuado para la modelización.

Para hacer el ajuste de las funciones de interés, consideramos el primer dato como fijo, esto es $P[X(t_1) = x_1] = 1$, con x_1 el primer valor observado en la serie y t_1 el primer año de observación. Así, se realiza la estimación máximo verosímil de los parámetros, resultando los valores $\hat{a} = 0.0115238235$ y $\hat{\sigma}^2 = 0.00008354483352128336$.

Con estos valores, se obtiene la estimación máximo verosímil de la familia de funciones θ y en particular se tiene la estimación máximo verosímil de las funciones media, moda y de cuantiles. En la figura 2 se pueden observar

estas funciones, junto con la trayectoria observada. Para representar la función de cuantiles se han seleccionado los valores de $\alpha = 0,05$ y $\alpha = 0,95$.

Como puede observarse en la figura 2, las funciones media y moda estimadas en este caso son muy similares, de forma que gráficamente son indistinguibles. Ambas representan en general el tipo de crecimiento, es decir, la tendencia creciente y la curvatura presente en los datos. Sin embargo, hay intervalos de tiempo en los cuales el ajuste no es adecuado. Esto sugiere el uso de información externa a la observación de la emisión de metano. Incorporar información en este modelo puede revertir en una estimación de un modelo mejor que represente de forma más ajustada y precisa la tendencia.

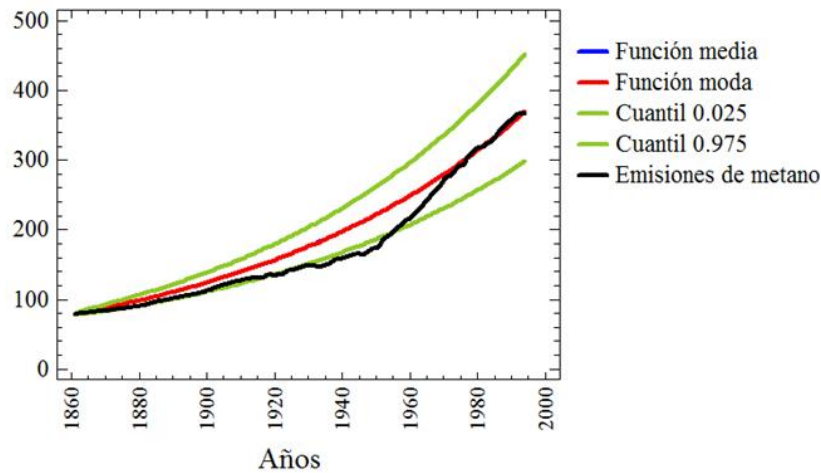


Figura 2. Funciones paramétricas media, moda y de cuantiles estimadas.

2. EL PROCESO DE DIFUSIÓN LOGNORMAL NO HOMOGÉNEO

La modelización con el proceso de difusión lognormal homogéneo, cuyos momentos infinitesimales no dependen del tiempo, se ve limitada cuando se observan trayectorias que presentan, en algunos intervalos, desviaciones con respecto a la tendencia. En estos casos, incluir ciertos factores exógenos que influyen en la variable de interés, considerando un proceso no homogéneo, en general hará que mejore la modelización y a su vez el fenómeno que intenta describirse podrá explicarse de una forma más acertada.

En esta sección se expone, en primer lugar, la definición del proceso de difusión lognormal no homogéneo en el caso general en que los factores exógenos se consideran una combinación lineal de funciones del tiempo. La descripción del proceso en este caso se hace en términos de los momentos infinitesimales del proceso, dando también la expresión de la función de densidad de transición y, a partir de ella, estableciendo la expresión de las funciones paramétricas de interés: media, moda y de cuantiles, así como sus versiones condicionadas. Estas funciones pueden expresarse como casos particulares de una familia más general, como en el caso del proceso homogéneo.

A continuación, se expone el caso particular en que los factores exógenos son desconocidos y se aproximan mediante una función de tipo polinómico. Para este caso, se reescriben las expresiones de las funciones paramétricas de interés y se aborda la inferencia máximo verosímil para esta clase de funciones. Una vez obtenidos los estimadores máximo verosímiles, se expone la formulación recursiva de los mismos, de modo que se puede abordar el problema de la elección del grado del polinomio realizando sucesivos ajustes y comparándolos entre sí para obtener el valor óptimo para el grado del polinomio.

Por último, se aplican estos resultados al caso de las emisiones de metano en un periodo de tiempo, obteniendo el polinomio que mejor ajuste proporciona a los datos.

2.1 Definición del proceso de difusión lognormal no homogéneo

El proceso de difusión lognormal con factores exógenos se define como un proceso de difusión no homogéneo en el tiempo $\{X(t): t_0 \leq t \leq T\}$ con momentos infinitesimales $A_1(x, t) = h(t)x$ y $A_2(x, t) = \sigma^2 x^2$, donde $\sigma > 0$ y la función que aparece en la media infinitesimal, $h(t)$, es una función continua en $[t_0, T]$, a diferencia del proceso de

difusión homogéneo donde la media infinitesimal es $A_1(x) = mx$.

La función $h(t)$ representa los factores exógenos que se incluyen en el modelo, los cuales son variables relacionadas con el proceso de manera que afectan a su comportamiento. Para poder incluirlas en el modelo, debe conocerse la función que regula el comportamiento de estos factores o, al menos, el valor de la integral de esta función, durante el periodo observado y así como durante el periodo de predicción.

Dado que es frecuente considerar más de una fuente externa de información, la función $h(t)$ suele considerarse como una combinación lineal de funciones continuas F_j , que son llamados factores, $h(t) = \sum_{j=1}^q \beta_j F_j(t)$ donde $\beta_j \in \mathbb{R}$ para $j = 1, \dots, q$.

En este proceso de difusión, para $s < t$, la función de densidad de transición viene dada por

$$f(x, t | y, s) = \frac{1}{x\sqrt{2\pi\sigma^2(t-s)}} \exp\left(\frac{-\left[\ln\left(\frac{x}{y}\right) - \left(\beta_0 - \left(\frac{\sigma^2}{2}\right)\right)(t-s) - \sum_{j=1}^q \beta_j \int_s^t F_j(\tau) d\tau\right]^2}{2\sigma^2(t-s)}\right)$$

que puede expresarse

$$f(x, t | y, s) = \frac{1}{x\sqrt{2\pi\sigma^2(t-s)}} \exp\left(\frac{-\left[\ln\left(\frac{x}{y}\right) - \bar{u}(t, s)'a\right]^2}{2\sigma^2(t-s)}\right)$$

tomando $a = \left(\beta_0 - \frac{\sigma^2}{2}, \beta_1, \dots, \beta_q\right)'$ y $\bar{u}(t, s) = \left(t-s, \int_s^t F_1(\tau) d\tau, \dots, \int_s^t F_q(\tau) d\tau\right)'$.

A partir de esta densidad de transición y de la distribución inicial del proceso, las distribuciones finito-dimensionales pueden obtenerse. Así, considerando el caso habitual $P[X(t_0) = x_0] = 1$ y denotando $\bar{u}(t) = \bar{u}(t, t_0)$, la distribución de $X(t)$ es lognormal:

$$X(t) \sim \Lambda_1(\ln(x_0) + \bar{u}(t)'a; \sigma^2(t-t_0))$$

y así, la función media del proceso puede expresarse como

$$M(t) = E[X(t)] = x_0 \exp\left(\bar{u}(t)'a + \frac{1}{2}(t-t_0)\sigma^2\right)$$

la función moda por su parte puede ser expresada de la siguiente forma:

$$Mo(t) = \text{Moda de } X(t) = x_0 \exp(\bar{u}(t)'a - (t-t_0)\sigma^2)$$

y la función de cuantiles se expresará

$$C_\alpha(t) = \text{Cuantil de orden } \alpha \text{ de } X(t) = x_0 \exp(\bar{u}(t)'a + z_\alpha \sigma \sqrt{t-t_0})$$

donde z_α es el cuantil de orden α de una distribución normal estándar.

Función	C	$A(t, s)$	$B(t, s)$	l
Media	x_0	$\bar{u}(t)$	$\frac{1}{2}(t-t_0)$	2
Media condicionada ($t > s$)	x_s	$\bar{u}(t, s)$	$\frac{1}{2}(t-s)$	2
Moda	x_0	$\bar{u}(t)$	$-(t-t_0)$	2
Moda condicionada ($t > s$)	x_s	$\bar{u}(t, s)$	$-(t-s)$	2
Cuantil de orden α	x_0	$\bar{u}(t)$	$z_\alpha \sqrt{t-t_0}$	1
Cuantil de orden α condicionada ($t > s$)	x_s	$\bar{u}(t, s)$	$z_\alpha \sqrt{t-s}$	1

Tabla 4. Valores de los argumentos de la función θ para las funciones media, moda y de cuantiles del proceso de difusión lognormal no homogéneo, así como sus versiones condicionadas.

Para realizar la estimación, estas funciones pueden expresarse en términos de la clase de funciones θ que se expresa

$$\theta(C, A(t, s), B(t, s), l) = C \exp(A(t, s)'a + B(t, s)\sigma_k^l)$$

de forma que para cada función de interés los valores de C , $A(t, s)$, $B(t, s)$ y l se resumen en la tabla 4.

No obstante, no siempre se dispone de información sobre los factores exógenos. En este tipo de situaciones se puede utilizar una aproximación a estas funciones del tiempo mediante el uso de factores exógenos polinómicos.

2.2 Caso de factor exógeno aproximado por una función de tipo polinómico

La función $h(t)$ se expresa como la suma $\sum_{j=0}^k \beta_j^{(k)} P_j^{(k)}$ donde $P_j^{(k)}$ es un polinomio de grado k y $\beta_j^{(k)}$ son parámetros fijos reales, $j = 1, \dots, k$.

La ventaja de usar esta función $h(t)$ reside principalmente en que no es necesario conocer el comportamiento del factor exógeno. Además, otra de las ventajas es que se puede realizar un procedimiento recursivo de ajuste, de forma que se puede determinar el grado del polinomio $h(t)$ realizando sucesivos ajustes de polinomios cada vez de grado mayor, hasta encontrar el que mejor ajuste proporcione.

En efecto, uno de los interrogantes que se plantea en el uso de este tipo de función es la forma de determinar el grado del polinomio. Esta cuestión se puede abordar realizando un ajuste recursivo, esto es, en cada paso se realiza el ajuste de un polinomio de un grado superior hasta que se obtiene el grado para el cual es mejor el ajuste de los datos. Este esquema de recursividad permite además utilizar los cálculos de cada paso realizado para el ajuste en el paso siguiente.

La metodología que se sigue en este caso parte de la suposición de que no existe información adicional a la observación de los valores de la variable endógena, la cual es observada de forma discreta, proporcionando los valores x_1, x_2, \dots, x_n en los instantes de tiempo t_1, t_2, \dots, t_n . Si consideramos $P[X(t_1) = x_1] = 1$, entonces

$$\ln\left(\frac{E[X(t)]}{x_1}\right) = \int_{t_1}^t h(s) ds = H(t)$$

y se pueden considerar los valores $f_i = \ln\left(\frac{x_i}{x_1}\right)$, $i = 1, \dots, n$ como una aproximación a $H(t_i)$. Así, con estos valores se ajusta un polinomio de grado $k + 1$, $P(t) = \sum_{i=1}^{k+1} a_i Q_i(t)$ y se aproxima el proceso de difusión lognormal $\{X(t): t_0, t, T\}$ con momentos infinitesimales $A_1(x, t) = h(t)x$ y $A_2(x, t) = \sigma^2 x^2$ por el proceso de difusión lognormal con factores exógenos de tipo polinómico, cuyos momentos infinitesimales vendrán dados por

$$A_1^{(k)}(x, t) = \left[\sum_{j=0}^k \beta_j^{(k)} P_j^{(k)}(t)\right]x \quad \text{y} \quad A_2^{(k)}(x, t) = \sigma_k^2 x^2$$

donde $P_j^{(k)}(t) = a_{j+1} Q'_{j+1}(t)$, $j = 1, \dots, k$ y $P_0^{(k)}(t) = 1$.

2.3 Funciones paramétricas de interés

En el caso del proceso de difusión lognormal con factor exógeno de tipo polinómico de grado k , las funciones, para modelización y predicción; media, moda y de cuantiles, y sus versiones condicionadas, se expresan tal y como se resume en la tabla 5.

Donde

- $a_k = (a_0^{(k)}, \beta_1^{(k)}, \dots, \beta_k^{(k)})'$ con $a_0^{(k)} = \beta_0^{(k)} - \frac{1}{2} \sigma_k^l$.
- $\bar{u}_k(t, s) = \left(t - s, \int_s^t P_1^{(k)}(\tau) d\tau, \dots, \int_s^t P_k^{(k)}(\tau) d\tau\right)'$ y $\bar{u}_k(t) = \bar{u}_k(t, t_0)$.
- z_α es el cuantil de orden α de una distribución normal estándar.

Estas funciones pueden expresarse, al igual que ocurre en el caso homogéneo, como casos particulares de una clase de funciones paramétricas:

$$\theta_k(C, A_k(t, s), B(t, s), l) = C \exp(A_k(t, s)' a_k + B(t, s) \sigma_k^l)$$

donde $C > 0$, $A(t, s) \in \mathbb{R}^{k+1}$, $B(t, s) \in \mathbb{R}$ y $l \in \mathbb{N}$.

Función	Expresión
Media: $E[X^{(k)}(t)]$	$M_k(t) = E[X^{(k)}(t_0)] \exp\left(\bar{u}_k(t)' a_k + \frac{1}{2}(t - t_0) \sigma_k^2\right)$
Media condicionada: $E[X^{(k)}(t) X^{(k)}(s) = x_s]$	$M_k(t) = E[X^{(k)}(s)] \exp\left(\bar{u}_k(t, s)' a_k + \frac{1}{2}(t - s) \sigma_k^2\right)$
Moda: $Mo[X^{(k)}(t)]$	$Mo_k(t) = E[X^{(k)}(t_0)] \exp(\bar{u}_k(t)' a_k - (t - t_0) \sigma_k^2)$
Moda condicionada: $Mo[X^{(k)}(t) X^{(k)}(s) = x_s]$	$Mo_k(t s) = E[X^{(k)}(s)] \exp(\bar{u}_k(t, s)' a_k - (t - s) \sigma_k^2)$
Cuantil de orden α : $C_\alpha[X^{(k)}(t)]$	$C_\alpha^k(t) = E[X^{(k)}(t_0)] \exp\left(\bar{u}_k(t)' a_k + z_\alpha \sqrt{(t - t_0) \sigma_k^2}\right)$
Cuantil de orden α condicionada: $C_\alpha[X^{(k)}(t) X^{(k)}(s)]$	$C_\alpha^k(t s) = E[X^{(k)}(s)] \exp\left(\bar{u}_k(t, s)' a_k + z_\alpha \sqrt{(t - s) \sigma_k^2}\right)$

Tabla 5. Expresión de las funciones media, moda y de cuantiles, así como sus versiones condicionadas, en el proceso de difusión lognormal con un polinomio de grado k como factor exógeno.

Los valores de C , $A_k(t, s)$, $B(t, s)$ y l que dan lugar a las funciones media, moda y de cuantiles se resumen en la tabla 6 considerando el caso de $P[X(t_0) = x_0] = 1$ y valores en $X(s) = x_s$ cuando $t > s$.

Función	C	$A_k(t, s)$	$B(t, s)$	l
Media	x_0	$\bar{u}_k(t)$	$\frac{1}{2}(t - t_0)$	2
Media condicionada ($t > s$)	x_s	$\bar{u}_k(t, s)$	$\frac{1}{2}(t - s)$	2
Moda	x_0	$\bar{u}_k(t)$	$-(t - t_0)$	2
Moda condicionada ($t > s$)	x_s	$\bar{u}_k(t, s)$	$-(t - s)$	2
Cuantil de orden α	x_0	$\bar{u}_k(t)$	$z_\alpha \sqrt{t - t_0}$	1
Cuantil de orden α condicionada ($t > s$)	x_s	$\bar{u}_k(t, s)$	$z_\alpha \sqrt{t - s}$	1

Tabla 6. Valores de los argumentos de la función θ para las funciones media, moda y de cuantiles, así como sus versiones condicionadas, en el proceso de difusión lognormal con un polinomio de grado k como factor exógeno.

2.4 Estimación máximo verosímil

Consideramos x_1, x_2, \dots, x_n una muestra discreta observada en los tiempos t_1, t_2, \dots, t_n , con $n > k + 2$ y suponemos $P[X^{(k)}(t_1) = x_1] = 1$. Después de transformar los datos de la siguiente forma $v_1 = x_1$ y $v_i = \frac{1}{\sqrt{t_i - t_{i-1}}} \ln\left(\frac{x_i}{x_{i-1}}\right)$ para $i = 1, \dots, n$, se tiene que los estimadores de máxima verosimilitud para los parámetros a_k y σ_k^2 son:

$$\hat{a}_k = V_k v \quad \text{y} \quad \hat{\sigma}_k^2 = \frac{1}{n-1} v' H_k v$$

donde $v = (v_1, \dots, v_n)'$, $V_k = (U_k U_k')^{-1} U_k$ y $H_k = I_{n-1} - U_k' V_k$, siendo U_k una matriz de rango $k + 1$ dada por $U_k = (u_2^{(k)}, \dots, u_n^{(k)})$ donde $u_i^{(k)} = \frac{1}{\sqrt{t_i - t_{i-1}}} \bar{u}^{(k)}(t_i, t_{i-1})$.

A partir de estos estimadores se obtiene $\hat{\theta}_k$, el estimador de máxima verosimilitud para las funciones θ_k de forma inmediata:

$$\hat{\theta}_k(C, A(t, s), B(t, s), l) = C \exp(A'(t, s) \hat{a}_k + B(t, s) \hat{\sigma}_k^l).$$

Este estimador contiene, como casos particulares, los estimadores de máxima verosimilitud de las funciones media, moda y cuantiles, así como de sus versiones condicionadas, seleccionando adecuadamente los valores que vienen expresados en la tabla 6.

2.5 Recursividad de las expresiones de los estimadores

Una de las cuestiones que se formulan en caso de que el factor exógeno sea un polinomio, es la elección de k , el grado de dicho polinomio. La respuesta dependerá del comportamiento de los datos. Una posibilidad es realizar un procedimiento por pasos, donde se introduce sucesivamente en $h(t)$ un polinomio de grado superior en cada paso. Para llevar a cabo esta estrategia, es de especial interés poder utilizar expresiones recursivas que simplifiquen los cálculos a la hora de realizar la estimación de las funciones de interés con un polinomio de grado $k + 1$, $P_{k+1}^{(k+1)}$, utilizando los cálculos realizados para la estimación con el polinomio de grado k , $P_k^{(k)}$. Así, como puede verse en [CITA], al introducir $P_{k+1}^{(k+1)}$, la nueva información que se añade sobre los factores exógenos queda reflejada en la matriz $U_{k+1} = (U_k' | d_{k+1}')$ donde la componente $i - 1$ -ésima del vector d_{k+1} viene dada por $\frac{1}{\sqrt{t_i - t_{i-1}}} \int_{t_{i-1}}^{t_i} P_{k+1}^{(k+1)(\tau)} d\tau$ con $i = 2, \dots, n$.

Teniendo en cuenta las expresiones para invertir una matriz por partes, y denotando $e_{k+1} = d_{k+1}' H_k d_{k+1}$, se obtienen las siguientes expresiones recursivas para los estimadores de máxima verosimilitud:

$$\hat{a}_{k+1} = \frac{1}{e_{k+1}} \left[\begin{array}{c} (e_{k+1} I_{k+1} + V_k d_{k+1} d_{k+1}' U_k') \hat{a}_k + (-V_k d_{k+1}) d_{k+1}' v \\ -d_{k+1}' U_k' \\ 1 \end{array} \right]$$

$$\hat{\sigma}_{k+1}^2 = \hat{\sigma}_k^2 - \frac{v' H_k d_{k+1} d_{k+1}' H_k v}{e_{k+1}}$$

De esta forma se simplifica el procedimiento recursivo de cálculo, realizándose la estimación de los sucesivos parámetros involucrados en la definición del proceso de forma que pueden utilizarse los cálculos de un paso para los cálculos necesarios en el paso siguiente.

2.6 Aplicación a los datos sobre emisiones de metano a la atmósfera

Para aplicar estos resultados a los datos sobre emisiones de metano a la atmósfera, notemos que nuestro objetivo pasa por determinar el grado del polinomio que actúa como factor exógeno, por lo que elegiremos un criterio que tenga en cuenta tanto el ajuste de la función a los datos como el error cometido en predicción. Para ello, como es usual, se realizará el ajuste con los datos de la serie salvo el último dato conocido, el cual será utilizado para comprobar el error cometido con cada modelo en la predicción de este valor. Así, la serie de datos de los años 1860 a 1993 es la que servirá para la estimación de los parámetros. Una vez obtenidas las estimaciones, se mide el error cuadrático medio para cada uno de los modelos, así como el error absoluto de predicción. Con esta información, se seleccionará el mejor modelo para los datos en estudio.

En primer lugar, consideremos, como en el ajuste del proceso de difusión lognormal homogéneo,

$P[X(t_1) = x_1] = 1$ y consideremos los valores $f_i = \ln\left(\frac{x_i}{x_1}\right)$ como los valores que se desean ajustar. Tomamos $f_1 = H(t_1) = 0$ y $Q_i(t) = (t - t_1)^j$, $j = 1, 2, \dots$ como base generadora de los polinomios.

Como primera aproximación al grado del polinomio, puede realizarse un ajuste de regresión lineal simple a los datos f_i obtenidos, de manera que pueden observarse el número de veces en los que los puntos en que la trayectoria observada se cortan con el modelo lineal. En este caso, pueden ser válidos 4 o 5 puntos de corte, lo que indicaría que el grado del polinomio que actúa como factor exógeno podría ser 3 o 4.

Con estas premisas, se realiza el ajuste del proceso de difusión lognormal con factores exógenos de tipo polinómico, donde recursivamente se aumenta k , el grado del polinomio en una unidad, desde el valor $k = 1$ hasta $k = 5$. Para cada modelo, se obtienen las estimaciones de máxima verosimilitud de los parámetros y así se obtienen las funciones media, moda y de cuantiles.

En la tabla 7 se resumen los valores de la estimación máximo verosímil de los parámetros. En la tabla 8, se observa el error cuadrático medio que se comete con la modelización según cada modelo y el error absoluto cometido para la predicción del último valor de la serie, cuyo valor es conocido. Los valores menores se han resaltado en negrilla.

A la vista de los resultados que se resumen en la tabla 8, puede observarse que los modelos donde $k = 3$ y $k = 4$ son los que presentan un mejor ajuste. El error cuadrático medio es menor para las funciones media y moda cuando $k = 4$, pero para las versiones condicionadas de las funciones media y moda, el ajuste será mejor para $k = 3$.

El error de predicción, para las funciones media y moda condicionadas es menor con $k = 3$. Para las versiones

sin condicionar, el valor más bajo de error de predicción con la función media se da con $k = 4$ y para la función moda se da para $k = 1$.

Modelo	Estimación de $a = (\beta_0 - \sigma^2/2, \beta_1, \dots, \beta_k)'$	Estimación de σ^2
$k = 1$	$\hat{a} = (0.0071421, 0.8236728)'$	$\hat{\sigma}^2 = 7.7145276 \times 10^{-5}$
$k = 2$	$\hat{a} = (0.0066739, -1.4806552, -0.135798)'$	$\hat{\sigma}^2 = 7.7101451 \times 10^{-5}$
$k = 3$	$\hat{a} = (0.0108849, 1.6291457, 1.9672737, 2.8234522)'$	$\hat{\sigma}^2 = 7.4570513 \times 10^{-5}$
$k = 4$	$\hat{a} = (0.0020905, 0.722, 0.7454921, 0.7653233, 0.7760967)'$	$\hat{\sigma}^2 = 6.5996458 \times 10^{-5}$
$k = 5$	$\hat{a} = (0.003296, 0.9105836, 1.1430773, 5.7305022, 0.1740628, 0.4579898)'$	$\hat{\sigma}^2 = 6.5864993 \times 10^{-5}$

Tabla 7: Estimación máximo verosímil de los modelos con factor exógeno de tipo polinómico, con polinomios de grado 1 a 5.

Por lo tanto, con la idea de utilizar el modelo que mejor represente la serie, para las emisiones de metano a la atmósfera el modelo seleccionado dependerá de si se desea utilizar las funciones media y moda, en cuyo caso debe utilizarse el modelo con $k = 4$, o bien sus versiones condicionadas, en cuyo caso se seleccionará un modelo con $k = 3$.

En la figura 3 pueden observarse las estimaciones de las funciones media y moda, así como los cuantiles de orden 0.025 y 0.975 con el proceso de difusión lognormal con factor exógeno polinómico de grado 4. La desviación que se observa de los valores reales con respecto al modelo es muy pequeña y en general puede apreciarse un ajuste bastante preciso de la curva a los datos. Este modelo es el que presenta el menor error cuadrático medio para el ajuste de la serie $M(t | t_0)$ y moda $Mo(t | t_0)$ y sus versiones condicionadas: $M(t | s)$ y $Mo(t | s)$ en los diferentes modelos con factor exógeno polinómico al usar las funciones media, moda y de cuantiles sin condicionar.

Modelo	Error cuadrático medio de ajuste				Error absoluto de predicción			
	$M(t t_0)$	$M(t s)$	$Mo(t t_0)$	$Mo(t s)$	$M(t t_0)$	$M(t s)$	$Mo(t t_0)$	$Mo(t s)$
$X^1(t)$	81.692	7.252	87.494	7.133	4.033	2.114	1.738	2.071
$X^2(t)$	83.747	7.124	86.907	7.005	3.852	1.935	1.912	1.892
$X^3(t)$	54.649	6.774	53.342	6.663	2.142	0.269	3.409	0.255
$X^4(t)$	22.184	7.242	14.02	7.145	1.587	3.205	6.455	3.241
$X^5(t)$	22.797	7.343	14.498	7.247	2.088	3.7	6.939	3.736

Tabla 8: Error cuadrático medio de ajuste y error absoluto de predicción para las funciones de interés media

La figura 4 muestra los datos observados junto con la funciones media, moda y de cuantiles condicionadas, para el modelo con $k = 3$. Al utilizar las funciones condicionadas, cada valor se estima a partir del anterior, por lo que las estimaciones serán muy precisas y el error cometido se reduce. Las funciones media y moda condicionadas son muy similares entre sí, de forma que aparecen indistinguibles en el gráfico.

En resumen, en la modelización de los datos de emisiones humanas de metano a la atmósfera en el periodo de tiempo 1869-1994, el proceso de difusión lognormal homogéneo resulta insuficiente en cuanto a que no representa adecuadamente los valores observados en un intervalo amplio de tiempo. El proceso de difusión lognormal no homogéneo precisa, por su parte, del conocimiento del comportamiento de los factores exógenos que se quieran incorporar al modelo. En caso de no conocerse tales factores exógenos, como es el caso en la serie de emisiones de metano, puede utilizarse una aproximación polinómica a dicho comportamiento, siendo posible realizar en este caso una estimación recursiva, lo que facilita la selección del grado del polinomio que actúa como factor exógeno.

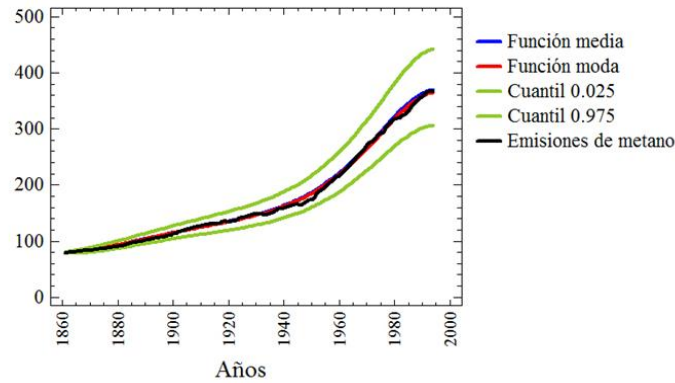


Figura 3. Funciones media, moda y de cuantiles ($\alpha = 0.025$ y 0.975). Modelo con $k = 4$.

En el caso de las emisiones de metano a la atmósfera, el modelo seleccionado como mejor para la estimación, es el que considera el polinomio de grado 4 como factor exógeno, en el caso de utilizar las funciones sin condicionar para modelizar o hacer predicciones.

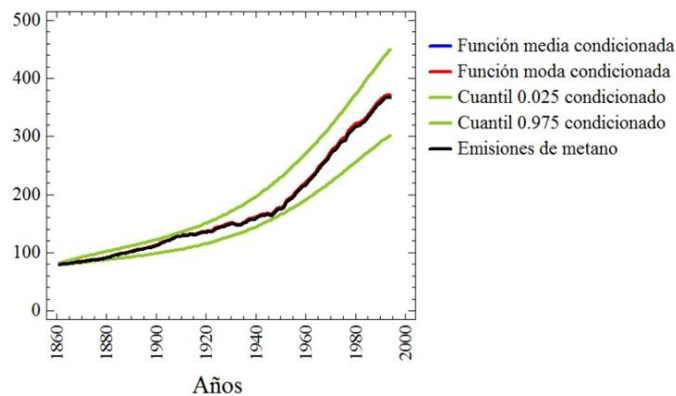


Figura 4. Funciones media, moda y de cuantiles condicionadas ($\alpha = 0.025$ y 0.975). Modelo con $k = 3$.

Si se seleccionan las funciones condicionadas para realizar el ajuste y las consiguientes predicciones, el modelo que presenta mejores resultados es el que considera el polinomio de grado 3 como factor exógeno.

En ambos casos, el ajuste de los datos mediante las funciones media y moda (o versiones condicionadas) resulta muy preciso y las predicciones realizadas no se alejan de la observación real, manifestándose como buenos modelos para explicar los datos observados.

Agradecimientos: Las autoras de este capítulo desean hacer constar su agradecimiento a los profesores Dña. Patricia Román y D. Francisco Torres, por su trabajo en la línea de investigación desarrollada conjuntamente, parte de la cual queda resumida en este texto.

Este trabajo ha sido subvencionado en parte por el Ministerio de Economía y Competitividad, MTM 2011/28962.

REFERENCIAS

- [1] BLACK, F. & SCHOLES, M. (1973): The pricing of options and corporate liabilities. **Journal of Political Economy**, 81, 637-654.
- [2] CAPOCELLI, R. M. & RICCIARDI, L.M. (1974): Growth with regulation in random environment. **Kybernetick**, 15, 147-157.

- [3] CAPOCELLI, R. M. & RICCIARDI, L.M. (1974): A diffusion model for population growth in random environment. **Theoretical Population Biology**, 5, 28-41.
- [4] CAPOCELLI, R. M. & RICCIARDI, L.M. (1975): A note on growth processes in random environment. **Biol. Cybernetics**, 18, 105-109.
- [5] GUTIÉRREZ, R. ANGULO, J.M., GONZÁLEZ, A. & PÉREZ, R. (1991): Inference in lognormal multidimensional diffusion process with exogenous factors: application to modelling un economics. **Applied Stochastic Models and Data Analysis**, 7, 295-396.
- [6] GUTIÉRREZ, R., GONZÁLEZ, A. & TORRES, F. (1997): Estimation in multivariate lognormal diffusion process with exogenous factors. **Applied Statistics**, 46, 140-146.
- [7] GUTIÉRREZ, R., RICO, N., ROMÁN, P., ROMERO, D., SERRANO, J.J. & TORRES, F. (2006): Approximating the nonhomogeneous lognormal diffusion process via polynomial exogenous factors. **Cybernetics and Systems: An International Journal**, 37, 293-309.
- [8] GUTIÉRREZ, R., RICO, N., ROMÁN, P., ROMERO, D. & TORRES, F. (2003): Lognormal diffusion process with polynomial exogenous factors. **Proc. of the ISI 54th Session**, 324-325.
- [9] GUTIÉRREZ, R., RICO, N., ROMÁN, P., ROMERO, D. & TORRES, F. (2004): Use of lognormal diffusion process with exogenous factors without external information. **Proc. of the 17th European Meeting on Cybernetics and Systems Research**, 163-168.
- [10] GUTIÉRREZ, R., ROMÁN, P., ROMERO, D. & TORRES, F. (2003): Forecasting for the univariate lognormal diffusion process with exogenous factors. **Cybernetics and Systems**, 34, 709-724.
- [11] GUTIÉRREZ, R., ROMÁN, P., ROMERO, D. & TORRES, F. (2006): A Gompertz-type Growth Stochastic Model. **Proc. of the 18th European Meeting on Cybernetics and Systems Research**, 204-208.
- [12] GUTIÉRREZ, R., ROMÁN, P. & TORRES, F. (1993): Inference on univariate lognormal diffusion processes via first-passage times. **Proc. of the 6th International Symposium on ASMDA**, 330-339.
- [13] GUTIÉRREZ, R., ROMÁN, P., & TORRES, F. (1999): Inference and first-passage time for the lognormal diffusion process with exogenous factors: application to modelling in economics. **Applied Stochastic Models in Bussines and Industry**, 15, 325-332.
- [14] GUTIÉRREZ, R., ROMÁN, P., & TORRES, F. (2001): Inference on some parametric functions in the univariate lognormal diffusion process with exogenous factors. **Test**, 10, 357-373.
- [15] GUTIÉRREZ, R., ROMÁN, P., & TORRES, F. (2001): UMVU estimation of the trend and covariance functions for the univariate lognormal diffusion process with exogenous factors. **Proc. of the 10th International Symposium on ASMDA**, 1, 516-521.
- [16] HUANG, W.N. (1973): Non-linear creep buckling with random temperature variations. **Nuclear Engineering and Desing**, 25, 432-437.
- [17] MALTHUS, T.R. (1846): **Ensayo sobre el principio de la población**. Madrid, España.
- [18] NODA, S. & HOSHIYA, M. (1998): Kriging of lognormal stochastic field. **Journal of Engineering Mechanics**, 124, 1775-1784.
- [19] RICO, N., ROMÁN, P., ROMERO, D. & TORRES, F. (2009): Gompertz-lognormal diffusion process for modelling the accumulated nutrients dissolved in the growth of *Capiscum Annum*. **Proc. of the 20th Annual Conference of The International Environmetrics Society**.
- [20] ROMÁN-ROMÁN, P., ROMERO, D., RUBIO, M.A. & TORRES-RUIZ, F. (2012): Estimating the parameters of a Gompertz-type diffusion process by means of Simulated Annealing, **Applied Mathematics and Computation**, 218, 5121-5131.

- [21] STERN, D.I. & KAUFMAN, R.K. (1998): Annual estimates of global anthropogenic methane emissions: 1960-1994. J. Carbon dioxide information analysis center, Oak Ridge National Laboratory, U.S Department of Energy, **Trends Online: a compendium of data on global change**. Oak Ridge, Tenn. USA.
- [22] TINTNER, G., & SENGUPTA, J.K. (1972): **Stochastic Economics**. Academic Press, N. York.
- [23] VÁZQUEZ, S. & GARCÍA, N. (2001): The probability distribution function of column density in molecular clouds. **Astrophysical Journal**, 557, 727-735.