



TESIS DOCTORAL

UNIVERSIDAD
DE GRANADA

Ajuste y estudio de fenómenos dinámicos mediante procesos de difusión de tipo logístico.

Aplicación a fenómenos de crecimiento.

Antonio Jesús Barrera García

Departamento de Estadística e Investigación Operativa

Universidad de Granada

2017

Editor: Universidad de Granada. Tesis Doctorales
Autor: Antonio Jesús Barrera García
ISBN: 978-84-9163-682-3
URI: <http://hdl.handle.net/10481/48830>



UNIVERSIDAD
DE GRANADA

Antonio Jesús Barrera García

Ajuste y estudio de fenómenos dinámicos mediante procesos de difusión de tipo logístico.

Aplicación a fenómenos de crecimiento.

Memoria presentada por Antonio Jesús Barrera García, y realizada en el Departamento de Estadística e Investigación Operativa de la Universidad de Granada, bajo la dirección de D^a. Patricia Román Román y D. Francisco de Asís Torres Ruiz, Catedráticos de dicho Departamento, con objeto de aspirar al grado de Doctor.

Antonio Jesús Barrera García

Dra. D^a. Patricia Román Román

Dr. D. Francisco de Asís Torres Ruiz

En sincero agradecimiento a
los directores de este trabajo.

A mi familia.

Índice general

Introducción	15
Objetivos y estructura de la tesis	19
Trabajos derivados de esta tesis	21
1. Procesos de difusión de tipo logístico	23
1.1. Introducción	23
1.2. La ecuación logística	24
1.3. Procesos de difusión a partir de la ecuación logística	26
1.4. Proceso de difusión con media logística	30
1.5. Proceso de difusión gaussiano con media logística	34
1.5.1. Modelo gaussiano	35
Características del proceso	36
1.5.2. Metodología y ajuste	37
Aproximación de la media infinitesimal	38
Aproximación de la varianza infinitesimal	39
1.5.3. Aplicaciones	42
Crecimiento de cultivos microorgánicos	42
2. Proceso de difusión logístico generalizado	45
2.1. Introducción	45
2.2. Una generalización de la ecuación logística	46
2.3. Difusión logística generalizada	47
2.3.1. Características principales	48
Función de densidad de transición	48
Distribuciones finito-dimensionales y momentos	50
2.4. Aplicabilidad del modelo	52
2.4.1. Parámetros desconocidos	53
El modelo	56
2.4.2. Influencia de datos redundantes	56
2.4.3. Resumen del procedimiento	58
2.5. Aplicaciones	60
2.5.1. Modelo con datos sin truncar	60
2.5.2. Modelo con datos truncados	63

2.5.3. Equivalencia paramétrica de modelos logísticos	64
2.6. Extensión a otros modelos	66
3. Proceso de difusión hiperbolástico de tipo I	69
3.1. Introducción	69
3.2. El modelo	70
3.2.1. Reformulación de la curva H_1	70
3.2.2. El proceso de difusión H_1	72
3.3. Estimación del modelo	74
3.3.1. Aplicación del algoritmo <i>Firefly</i>	76
Parámetros iniciales y acotación del espacio paramétrico	79
3.4. Aplicaciones	81
3.4.1. Aplicación a datos simulados	81
3.4.2. Aplicación a datos reales	88
Conclusiones	95
Líneas futuras	96
A. Cálculo de las ecuaciones de verosimilitud	99
B. Código en R para el algoritmo Firefly	101
Bibliografía	103

Índice de tablas

2.1. Estimaciones para dos modelos basados en la curva logística.	65
3.1. Resultados del algoritmo <i>Firefly</i> sobre datos simulados.	82
3.2. Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para α vs. γ	87
3.3. Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para α vs. δ	88
3.4. Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para γ vs. δ	88
3.5. Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para α vs. n	89
3.6. Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para γ vs. n	89
3.7. Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para δ vs. n	89
3.8. Parámetros estimados y errores relativos absolutos de la aplicación del algoritmo <i>Firefly</i> a datos reales.	92
3.9. Aplicación real: estimaciones de los parámetros tras aplicar el algoritmo <i>Firefly</i>	92

Índice de figuras

1.1. Curvas logísticas para diversos valores de α y β	25
1.2. Trayectorias simuladas, media observa y varianza observada. . .	44
1.3. (a) Media observada y función media ajustada. (b) Función $\dot{v}_{\hat{\theta}}(t) - 2a_*(t)v_{\hat{\theta}}(t)$	44
1.4. (a) Varianzas observada y ajustada. (b) Trayectorias simuladas del nuevo proceso.	44
2.1. Comparación entre datos no truncados y truncados mediante la técnica descrita (figuras arriba, de izquierda a derecha). Abajo, de izquierda a derecha, imágenes de los datos originales (truncados y no) por $\rho_{\hat{\alpha}}$. Nótese que los provenientes de los datos truncados son susceptibles de un ajuste mucho más sencillo, en este caso, de tipo lineal.	58
2.2. De arriba a abajo y de izquierda a derecha: Trayectorias simuladas (rojo); datos modificados mediante <i>logit</i> (rojo) y ajuste (verde); trayectorias del nuevo proceso; comparación de medias (original rojo, nuevo azul).	62
2.3. De arriba a abajo y de izquierda a derecha: datos originales, media observada vía ρ vs. ajuste, trayectorias del nuevo proceso, comparación de medias.	64
2.4. Ajuste de dos modelos de tipo logístico sobre datos simulados. Franja central, en gris, datos simulados. Siguiendo la flecha, se llega a la gráfica superior: logit y funciones ajustadas (azul polinomio, violeta hiperbolástico I). Con ellas se generan dos difusiones respectivas, azul y violeta en el centro. Abajo, comparativa de medias observada y simuladas.	65
3.1. Trayectorias simuladas del proceso de difusión de tipo H_1	81
3.2. Primer ejemplo de simulación: evolución (desde gris claro a negro) de 80 generaciones de 40 luciérnagas para $\eta = 0.5$, $\lambda = 0.8$, $\mu = 0.8$ y $\sigma = 0.015$ (de izquierda a derecha, y de arriba a abajo). Valores originales, en rojo.	84

3.3. Primer ejemplo de simulación: evolución de las luciérnagas a lo largo de las últimas 60 generaciones, para cada parámetro (en el mismo orden que la figura anterior). En verde, evolución de la mejor luciérnaga; en rojo, de la peor.	85
3.4. Primer ejemplo de simulación: para cada par de parámetros, proyección de la trayectoria bidimensional representando la evolución de la última luciérnaga (la mejor) durante las generaciones (empezando en el punto de color rojo claro y acabando en rojo oscuro).	86
3.5. Primer ejemplo de simulación: (a): valor de la función objetivo para la mejor luciérnaga. (b): comparación entre las medias observada y ajustada (en rojo).	86
3.6. Aplicación real: fluorescencia vs. ciclo.	91
3.7. Aplicación real: evolución de las estimaciones de los parámetros mediante el algoritmo <i>Firefly</i>	93
3.8. Aplicación real: (a) Media de la fluorescencia original (gris) vs. fluorescencia simulada (rojo) vs. fluorescencia teórica ajustada (verde). (b) Log-fluorescencia de los datos originales (gris) y logaritmo de las trayectorias simuladas del proceso (rojo). . . .	93

Introducción

La modelización matemática es una herramienta fundamental a la hora de analizar y describir los fenómenos ocurridos en la naturaleza o la sociedad. El estudio de fenómenos dinámicos comprende diversas áreas de las matemáticas, así como exige un conocimiento profundo de la naturaleza del fenómeno en estudio. Mediante las herramientas adecuadas es posible entender tanto el funcionamiento de un proceso dinámico, como intentar predecir con unos márgenes mínimos posibles evoluciones en el futuro. La aplicación de esta metodología abarca desde la biomedicina hasta las finanzas, pasando, obviamente, por la física y otras ciencias aplicadas al mundo real.

En el contexto particular de los fenómenos de crecimiento, cabe destacar el interés por el crecimiento de poblaciones, humanas o no, que dio origen a algunas funciones entendidas como curvas de crecimiento, por expresar estas la evolución del número de elementos de una población durante un intervalo de tiempo. Surgieron así, entre los siglos XVIII y XIX, ejemplos de modelos de población como el de Malthus, de carácter exponencial (véase Malthus [17]) o el de Verhulst (presentado en Verhulst [33]). El modelo malthussiano considera un crecimiento poblacional sin límite, lo que lo hace adecuado para entornos carentes de depredadores y con exceso de alimentos (esto es, los elementos de la población no eran eliminados ni competían por comida con otros). Sin embargo, es a todas luces evidente que tal modelo no parece adecuado para estudiar el crecimiento de la población humana. Surge así el modelo de Verhulst, que extiende y mejora al de Malthus en el sentido de incluir a una cota superior en la cual el crecimiento se equilibra, alcanzando así la población la llamada capacidad máxima sostenible.

El modelo de Verhulst se conoce como modelo logístico (o curva logística) y desde su descubrimiento se ha convertido en una de las curvas sigmoidales más

sencillas, dotada de un punto de inflexión en el cual se alcanza la mitad de la capacidad máxima sostenible. Sobre la curva existe un estudio en profundidad de Feller [9], considerado pionero pues introduce la curva logística en un contexto biológico más amplio (recordemos que esta curva, como tantas otras, surge del intento de estudiar el crecimiento de la población).

La simplicidad y efectividad de la curva logística hace que se hayan desarrollado numerosas variantes que bien podrían emplearse con los mismos propósitos, pero con mayor éxito (numerosos ejemplos pueden consultarse en el trabajo de Tsoularis y Wallace [30]). Un caso particular de esto lo constituye la curva hiperbolística de tipo I, introducida por Tabatabai, Williams y Bursac [29] y con un interesante desempeño en el ajuste de datos de tipo sigmoideal. Las curvas hiperbolísticas (existen de tres tipos) se han aplicado con éxito en modelos matemáticos relacionados con crecimiento tumoral o de células madre embrionarias, entre otros (pueden consultarse, respectivamente, los trabajos de Eby, Tabatabai y Bursac [6] y de Tabatabai et al. [27]).

La característica común de todas las curvas de tipo logístico radica en la existencia de una función exponencial cuyo argumento es una u otra función del tiempo, dependiente, en general, de uno o más parámetros.

Por otro lado, el uso de modelos deterministas como medio para la descripción de fenómenos dinámicos, y en particular de fenómenos de crecimiento, se enfrenta, sin embargo, a un problema evidente. En efecto, la complejidad inherente al fenómeno implica la consideración de numerosos factores que, si bien pueden tener un origen físico, no siempre son conocidos o, caso de serlo, cuantificables. Esto es, la dinámica del fenómeno a estudio normalmente consta de elementos que están fuera del alcance de un modelo matemático pero que, claramente, ejercen una influencia notable en la evolución del conjunto. Por lo tanto, la única forma de tomar en consideración tales efectos es mediante el uso de modelos estocásticos entre los que cabe destacar los procesos de difusión, usados con éxito en numerosas áreas de investigación.

A este respecto, los trabajos de Feller [7, 8] supusieron en su momento un gran avance en el campo de la modelización matemática pues aunaban los desarrollos de Kolmogorov en teoría de la probabilidad (sus ecuaciones adelantada y atrasada, por ejemplo) con las teorías de Volterra sobre dinámica de poblaciones. En concreto, el objetivo desde aquel momento era conse-

guir procesos estocásticos a partir de las ecuaciones propias de algún modelo determinista. Para ello, generalmente se añadía un ruido a la ecuación diferencial ordinaria y se estudiaba el comportamiento del proceso resultante, en la mayoría de casos de expresión inaccesible.

Como ejemplo, pueden consultarse los trabajos de Capocelli y Ricciardi [4, 5], donde consideran el modelo malthusiano con una tasa de crecimiento modificada mediante la adición de un ruido blanco (un proceso gaussiano delta-correlado). La solución es el proceso de difusión lognormal, en el cual la media sigue un comportamiento exponencial.

Sin embargo, es el aumento de la variabilidad lo que dificulta la obtención de procesos de difusión con una tendencia similar a la de los modelos deterministas que los originan. Así, en el caso de la extensión aleatoria del modelo logístico, los trabajos ya citados de Feller o Capocelli y Ricciardi indican que las difusiones resultantes presentan un comportamiento adecuado con variabilidades bajas, pero son muy sensibles a altas variaciones de manera que no se logran modelos estocásticos con funciones media exactamente iguales a los modelos deterministas de partida. Además, para algunos procesos no es posible disponer de herramientas básicas para su estudio, como es el caso de las densidades de transición.

Un paso más en el desarrollo de difusiones con un comportamiento medio prefijado, lo encontramos en los trabajos de Gutiérrez-Jáimez et al. [15], construyendo un proceso de difusión cuya función media es la curva Gompertz, Román-Román, Romero y Torres-Ruiz [22] con un proceso de media la curva de von Bertalanffy o, en el caso de una difusión con media logística, el trabajo de Román-Román y Torres-Ruiz [23]. La clave para la construcción de estos procesos de difusión está en considerar una media infinitesimal no solo dependiente del estado sino también del tiempo, y expresar esta influencia temporal en función de la curva original que se pretende sea la media del proceso final. El planteamiento de la ecuación diferencial estocástica es sencillo, y su solución resulta ser el proceso de difusión lognormal con factores exógenos (son estos factores exógenos los representados mediante la función temporal insertada en la media infinitesimal). La buena noticia es que este proceso ha sido estudiado en profundidad, por ejemplo por Gutiérrez et al. [14], obteniéndose de manera explícita tanto las expresiones de su solución, de la función de densidad de

transición de probabilidad, así como de otras funciones asociadas. Esto hace posible otros desarrollos, habida cuenta de la disponibilidad de tales funciones. Ello conlleva, de forma natural, la consideración de estudios que requieran expresiones explícitas como las que proporciona este enfoque metodológico.

En un contexto eminentemente práctico, es lógico preguntarse por la posibilidad de obtener una difusión a partir de los datos generados por la observación de un fenómeno dinámico con un comportamiento de tipo sigmoïdal. A este respecto, parece adecuado lograr, a partir de observaciones de un fenómeno, los momentos infinitesimales que caracterizan al proceso.

Llegados aquí, y volviendo al plano determinista, surge la cuestión relativa a si merece la pena el desarrollo de un marco teórico de carácter estocástico enfocado únicamente en un tipo de curva (la curva logística, en este caso). Tal y como se ha comentado al inicio de esta introducción, existen numerosas variantes de la curva logística, por lo que la posibilidad de unificar este tipo de curvas en una misma expresión, sugiere la propuesta de una generalización desde un punto de vista funcional, es decir, considerar un operador que tome funciones con determinadas características y devuelva una curva de tipo logístico. De esta manera se logra una visión general de numerosas curvas de comportamiento sigmoïdal, que puede extenderse a modelos aleatorios, consiguiendo así las expresiones genéricas de distintas características del proceso de difusión asociado.

Con todo ello, el resultado es un único modelo aleatorio mediante el cual obtener procesos de difusión que sigan una determinada tendencia de tipo logístico.

Otra cuestión pendiente, una vez desarrollado el proceso generalizado, estriba en considerar procedimientos de inferencia para trabajar así en aplicaciones prácticas. A tal fin, el método de máxima verosimilitud, ampliamente usado, resulta adecuado. Sin embargo, el sistema de ecuaciones resultantes (ecuaciones de verosimilitud) puede llegar a ser de enorme complejidad, como es el caso de las derivadas del estudio de un proceso de difusión con función media la curva hiperbolística de tipo I. Para resolverlas y completar así el procedimiento de definición teórica de un proceso y su adaptación para un uso real en aplicaciones, es posible emplear algoritmos metaheurísticos que permitan minimizar una función determinada en un espacio paramétrico defi-

nido previamente. En efecto, Román-Román et al. [20] emplean el algoritmo *Simulated Annealing* como herramienta para proceder con la estimación de los parámetros de un proceso de difusión con media una curva Gompertz. O también, Román-Román y Torres-Ruiz [21] obtienen los parámetros de un proceso con media de tipo Richards mediante el algoritmo *Variable Neighborhood Search*.

Objetivos y estructura de la tesis

Con todo esto, podemos ahora establecer claramente los objetivos fundamentales de este trabajo, que son:

1. Construcción de procesos de difusión cuya función media sea una curva de tipo logístico y determinación del mismo a partir de los datos muestrales suministrados en aplicaciones prácticas.
2. Construcción de un proceso de difusión generalizado que permita estudiar diversos patrones de comportamiento sigmoidales derivados de la curva logística. En ese sentido, determinar procedimientos generales para su estudio y particularización a curvas concretas de gran aplicación hoy día, como es el caso de las curvas hiperbolásticas, así como plantear futuras extensiones a curvas de tipo plurisigmoidal u otros tipos de curvas mediante la aplicación de la metodología aquí introducida.
3. Estudiar los problemas de inferencia para los procesos considerados en los apartados anteriores, abordando la introducción de métodos metaheurísticos cuando la dificultad de resolución de las ecuaciones de verosimilitud así lo aconseje.
4. Mostrar la validez de los métodos y procedimientos introducidos mediante ejemplos de simulación y aplicación a casos reales.

Y a fin de alcanzar los objetivos propuestos, la tesis se estructura de la siguiente manera:

En el capítulo 1 se realiza, de forma somera, una introducción histórica a la ecuación logística y los motivos esgrimidos para su desarrollo, poniendo el

foco en su solución, la conocida curva logística. Posteriormente se consideran variaciones clásicas de la ecuación logística (véase a este respecto el trabajo de recopilación de Tsoularis y Wallace [30]), dando lugar a curvas específicas para distintos campos de estudio. Una vez presentada la curva logística, se analiza su introducción en un contexto aleatorio mediante la presentación de los avances más importantes en el desarrollo de procesos estocásticos de difusión a partir de la curva logística, desde los trabajos pioneros de Feller [9], pasando por Capocelli y Ricciardi [5] y Giovanis y Skiadas [13], que trabajan empleando la estimación por máxima verosimilitud desde el muestreo continuo de las trayectorias. Además se consideran los problemas asociados a tales desarrollos, recurriendo a la literatura existente al respecto. Posteriormente, se aborda la solución al problema relativo a conseguir procesos de difusión que extiendan de manera estocástica la curva logística (véase el trabajo de Román-Román y Torres-Ruiz [23]). A tal efecto, se introduce la teoría conducente a un proceso lognormal con factores exógenos cuya función media es la curva logística. Finalmente, se analiza en profundidad el caso particular de un proceso gaussiano con media logística en un contexto aplicado. Para ello se discutirá la posibilidad de estimación de sus funciones media y varianza, aprovechando la condición de normalidad que permite disponer de distintas expresiones de manera explícita.

Tomando como punto de partida las numerosas modificaciones que se han realizado de la curva logística, y por ende, de su proceso de difusión asociado, en el capítulo 2 se establece el marco teórico para la definición de un proceso de difusión logístico generalizado. En concreto, se aborda la introducción de una generalización, desde un punto de vista funcional de la curva logística, al contrario de que otras generalizaciones que se sitúan en un contexto paramétrico. Se consigue, a este respecto, un operador que da lugar a distintas curvas de tipo logístico según actúe sobre una u otra función no acotada superiormente, citando el caso de la curva hiperbolística de tipo I (puede consultarse Tabatabai, Williams y Bursac [29]), que será punto de partida del último capítulo de este trabajo. Tal enfoque permitirá estudiar las características de cualquier curva de tipo logístico en base a la función que la defina dentro del contexto de la curva logística generalizada. Con este fin, se obtienen las características más importantes, empleando para

ello resultados teóricos como el de Ricciardi [19]. A todo esto ha de sumarse la posible extensión a un modelo aleatorio cuya función media sea la recién introducida curva logística generalizada. A tal efecto se aplicará la metodología del capítulo anterior, la cual, sin embargo, se encontrará con un problema en el proceso de estimación de los parámetros de la curva logística generalizada. La construcción misma de la difusión nos permitirá solventar esta problemática de manera efectiva.

Finalmente, el capítulo 3 toma como punto de partida una función particular que, siendo argumento del operador definido en el capítulo 2, dará lugar a la curva hiperbolástica de tipo I, mencionada en el párrafo anterior. Esta curva es presentada y empleada, con la metodología apropiada, para construir un proceso de difusión hiperbolástico de tipo I. A continuación se aborda la cuestión relativa a la inferencia en tal proceso. La complejidad de las ecuaciones de verosimilitud llevará a considerar un procedimiento de optimización mediante algoritmos metaheurísticos. Es aquí donde se introduce el algoritmo *Firefly* (sobre el cual puede consultarse el trabajo de Yang [34]), describiendo su funcionamiento y aplicación al caso particular de la difusión hiperbolástica de tipo I recién definida. Con esta herramienta se conseguirá resolver el problema de inferencia, lo que se muestra al final del capítulo con dos aplicaciones, la primera empleando datos simulados, y la segunda usando datos reales procedentes del campo de la biología molecular.

Trabajos derivados de esta tesis

A consecuencia de la persecución de tales objetivos, además de esta memoria, se han obtenido resultados expuestos en congresos nacionales e internacionales, así como en revistas indexadas en el *Journal Citation Reports (JCR)*, todos realizados por el autor de esta memoria y sus directores. En concreto:

- Artículos de revistas citadas en el índice JCR
 - *Fitting dynamic growth models of biological phenomena from sample observations through Gaussian diffusion processes*, publicado en la revista *Biosystems* (situada en cuartil Q2), versando sobre la construcción y ajuste de procesos gaussianos a partir de observaciones.

- *A hyperbolic type-I diffusion process: parameter estimation by means of the Firefly algorithm*, enviado para su revisión a la revista *Biosystems* y donde se construye un proceso de difusión con media la curva hiperbólica de tipo I, realizando un estudio de inferencia mediante el uso del algoritmo metaheurístico *Firefly*.
- Contribuciones en congresos de ámbito nacional
 - SEIO¹ 2012: se realiza una primera introducción a la construcción y ajuste de procesos gaussianos, considerando una acotación de su varianza infinitesimal.
 - SEIO 2013: se presenta el proceso de difusión hiperbólico de tipo I.
 - RSME² 2015: se aborda el uso de algoritmos metaheurísticos el procedimiento de inferencia relacionado con el proceso hiperbólico.
 - Contribuciones en congresos de ámbito internacional
 - BIOCOMP³ 2012: se extiende el trabajo realizado en la construcción y ajuste del proceso gaussiano con media logística.
 - ERCIM⁴ 2013: se aborda una generalización funcional de la curva logística, al fin de construir un proceso de difusión cuya tendencia englobe los distintos comportamientos de carácter sigmoidal.
 - COMPSTAT⁵ 2014: empleando la misma metodología, se introduce una generalización de la curva Weibull.
 - ASMDA⁶ 2017: se desarrolla el concepto de proceso de difusión logístico generalizado, así como sus potenciales aplicaciones.

¹Congreso de la Sociedad Española de Estadística e Investigación Operativa

²Congreso de la Real Sociedad Matemática Española

³Mathematical Modeling and Computational Topics in Biosciences

⁴ERCIM Working Group on Computational and Methodological Statistics

⁵International Conference on Computational Statistics

⁶Applied Stochastic Models and Data Analysis

Capítulo 1

Procesos de difusión de tipo logístico

1.1. Introducción	23
1.2. La ecuación logística	24
1.3. Procesos de difusión a partir de la ecuación logística . .	26
1.4. Proceso de difusión con media logística	30
1.5. Proceso de difusión gaussiano con media logística . . .	34

1.1. Introducción

En el capítulo inicial de este trabajo se presentan los elementos básicos que conformarán los desarrollos posteriores, incluyendo las dos cuerpos teóricos principales, a saber, curvas de crecimiento y procesos de difusión. Se consideran su evolución histórica y la fusión de ambos al objeto de enfrentar los problemas relacionados con la dinámica de poblaciones. A tal efecto, en la segunda sección se introduce la curva logística desde su perspectiva natural determinista, repasando sus orígenes y propiedades más importantes, además de enumerando aquellas otras curvas derivadas de ella que tengan o hayan tenido un papel relevante en el estudio de los fenómenos de crecimiento.

Por otro lado, en la sección inmediatamente siguiente, entra el concepto de proceso de difusión, que se convertirá en elemento central de esta memoria. Su inclusión en el campo de la dinámica poblacional, por parte de Feller

establece las bases del tratamiento estocástico de fenómenos de crecimiento, principalmente de carácter biológico, aún hoy vigente. La unificación entre la teoría de Volterra y los trabajos de Kolmogorov da lugar a numerosos avances e intentos para introducir modelos estocásticos que describan fenómenos exhibiendo distintos comportamientos, en particular, aquellos de tipo logístico. Sin embargo, los primeros modelos se encontraban con dificultades teóricas que limitaban su aplicación. Por ello, en las dos últimas secciones se comentan los avances a este respecto, haciendo especial hincapié en los procesos gaussianos, que, por su naturaleza, permiten un tratamiento teórico más exhaustivo.

1.2. La ecuación logística

El modelo logístico, inicialmente propuesto en Verhulst [33], viene dado por la ecuación diferencial

$$\frac{dN(t)}{dt} = rN(t) \left(1 - \frac{N(t)}{K} \right), \quad (1.1)$$

donde $N(t)$ es el tamaño de la población en el instante t , r la ratio de crecimiento intrínseco y K la capacidad máxima sostenible, parámetro relativo al concepto de nivel de saturación al que puede llegar una población debido a factores como la competencia o la escasez de recursos. Fue esta idea la que permitió a Verhulst desarrollar su modelo a partir del exponencial, propuesto por Malthus [17] y con el inconveniente de predecir un aumento de la población sin restricciones.

La solución de (1.1), para una población inicial $N_0 = N(t_0)$, es

$$N(t) = K \left(\left(\frac{K}{N_0} - 1 \right) e^{rt_0} e^{-rt} + 1 \right)^{-1}.$$

En general, la ecuación logística (1.1), para $X(t_0) = x_0$, suele escribirse como la ecuación de Bernoulli

$$\frac{dx(t)}{dt} = \alpha x(t) - \beta x(t)^2, \quad (1.2)$$

donde $\alpha = r$ y $\alpha/\beta = K$ según la notación original. La solución se expresa

entonces como

$$x(t) = \frac{\alpha}{\beta} \left(1 + \left(\frac{\alpha}{\beta x_0} - 1 \right) e^{-\alpha(t-t_0)} \right)^{-1}. \quad (1.3)$$

Esta curva presenta, como principales características, un comportamiento creciente para $\alpha > \beta x_0$ con $\lim_{t \rightarrow \infty} x(t) = \alpha/\beta$, así como una evolución decreciente de la ratio de crecimiento relativo $\dot{x}(t)/x(t)$ según aumenta la población, anulándose cuando $x(t) = \alpha/\beta$.

Además, la curva posee un punto de inflexión en el instante

$$t_* = t_0 + \frac{1}{\alpha} \log \left(\frac{\alpha}{\beta x_0} - 1 \right),$$

en el cual la población alcanza exactamente la mitad de la capacidad máxima sostenible, esto es, $x(t_*) = \alpha/(2\beta)$.

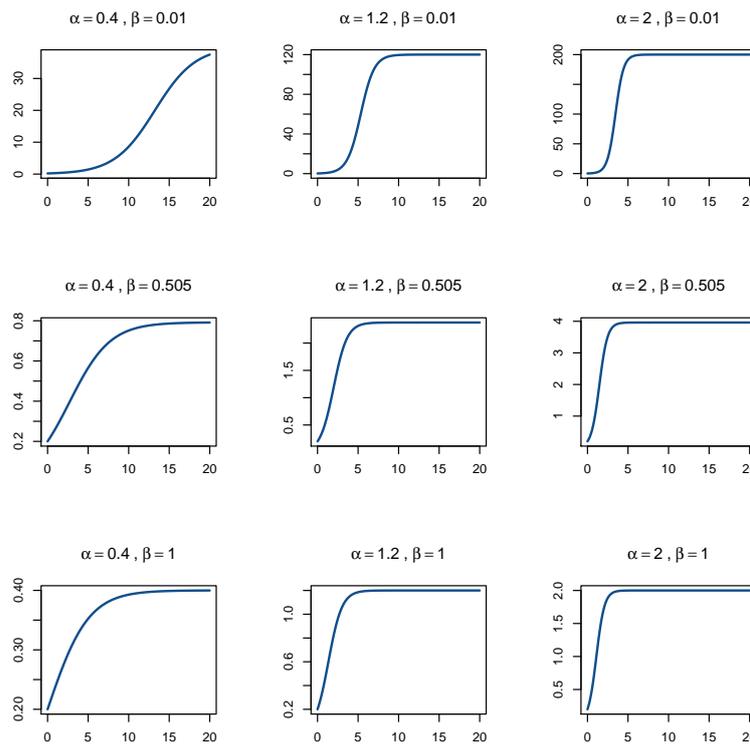


Figura 1.1: Curvas logísticas para diversos valores de α y β .

Estas propiedades hacen del logístico un modelo apropiado y susceptible

de ser aplicado en numerosos campos de investigación relacionados con el crecimiento o la evolución de poblaciones. Así, constantemente se han sucedido numerosas variantes en función del problema a tratar, resultando en un amplio espectro de curvas con características específicas pero compartiendo un mismo origen de tipo logístico. Un detallado resumen de muchos de estos modelos y sus características principales, así como del mismo modelo logístico clásico, puede encontrarse en Tsoularis y Wallace [30]. En particular, cabe destacar el de von Bertalanffy, surgido al objeto de introducir en el logístico factores metabólicos relacionados con el crecimiento animal, y dado por la siguiente ecuación

$$\frac{dN(t)}{dt} = rN(t)^{\frac{2}{3}} \left(1 - \left(\frac{N(t)}{K} \right)^{\frac{1}{3}} \right).$$

Posteriormente, y tratando de describir ciertas pautas del crecimiento vegetal, esta expresión fue extendida a su vez hasta el conocido modelo de Richards,

$$\frac{dN(t)}{dt} = rN(t) \left(1 - \left(\frac{N(t)}{K} \right)^{\beta} \right),$$

que deriva en la ecuación logística cuando $\beta = 1$.

El creciente número de modelos derivados de la curva logística ha llevado a considerar versiones generalizadas, desde un punto de vista paramétrico, como, por ejemplo, el visto en Turner et al. [32],

$$\frac{dN(t)}{dt} = rN(t)^{1+\beta(1-\gamma)} \left(1 - \left(\frac{N(t)}{K} \right)^{\beta} \right)^{\gamma},$$

para $\gamma, \beta > 0$ con $\gamma < 1 + 1/\beta$.

1.3. Procesos de difusión a partir de la ecuación logística

El desarrollo de modelos deterministas para la descripción de fenómenos dinámicos de crecimiento se enfrenta al dilema, desde sus inicios, de la necesidad que pueda existir de considerar aquellos factores generadores de incertidumbre, que escapan al control del modelo clásico, bien sea por una

falta de conocimiento del fenómeno en sí, dificultad o imposibilidad teórica para incluirlos en las ecuaciones, limitación o errores en la toma de medidas, o, simplemente, la misma naturaleza aleatoria de los mismos. Los trabajos de Volterra sobre dinámica de poblaciones, a principios del siglo XX, hacen uso de ecuaciones diferenciales ordinarias como las vistas hasta ahora en la definición de curvas de crecimiento. Estas modelan fenómenos de nacimiento, muerte e interacción de individuos, entendiendo que el mayor tamaño de las poblaciones es suficiente razón como para ignorar las fluctuaciones aleatorias o no controlables por el modelo.

Por otro lado, en la misma época, Kolmogorov, en su vasto programa de formalización de la teoría de la probabilidad, plantea de manera rigurosa la dinámica estocástica, empleando para ello cadenas de Markov en tiempo continuo y obteniendo así las conocidas ecuaciones adelantada y atrasada que llevan su nombre. Estos trabajos proporcionan las herramientas formales para el tratamiento de los fenómenos dinámicos basándose en probabilidades y ampliando así la capacidad descriptiva de los modelos.

La unificación de estas dos visiones llega por parte de Feller [7], que aplica la teoría de Kolmogorov a los trabajos de Volterra, sentando así las bases para el estudio probabilístico de los fenómenos dinámicos de crecimiento, aún en la actualidad muy utilizado.

Recordemos que un proceso de difusión se define como un proceso estocástico $\{X_t : t_0 \leq t \leq T\}$, de Markov, en tiempo continuo y con espacio de estados continuo, trayectorias continuas casi seguro y con las siguientes propiedades para todo $\epsilon > 0$ y todo x , siendo $F(x, t|y, s) = P(X_t \leq x | X_s = y)$ su función de distribución de transición para $t > s$,

$$\begin{aligned} \lim_{h \downarrow 0} \frac{1}{h} \int_{|y-x| > \epsilon} F(dy, t+h|x, t) &= 0 \\ \lim_{h \downarrow 0} \frac{1}{h} \int_{|y-x| \leq \epsilon} (y-x)F(dy, t+h|x, t) &= A_1(x, t) \\ \lim_{h \downarrow 0} \frac{1}{h} \int_{|y-x| \leq \epsilon} (y-x)^2 F(dy, t+h|x, t) &= A_2(x, t) \end{aligned}$$

para funciones A_1 y A_2 , los momentos truncados, por unidad de tiempo, de los incrementos condicionados (siempre existen).

En el caso particular de tiempo continuo, las condiciones anteriores pueden simplificarse con la existencia de $\delta > 0$ tal que para todo x se tiene

$$\lim_{h \downarrow 0} \frac{1}{h} \int |y - x|^{2+\delta} F(dy, t + h | x, t) = 0$$

$$\lim_{h \downarrow 0} \frac{1}{h} \int (y - x) F(dy, t + h | x, t) = A_1(x, t)$$

$$\lim_{h \downarrow 0} \frac{1}{h} \int (y - x)^2 F(dy, t + h | x, t) = A_2(x, t)$$

donde ahora A_1 y A_2 son los momentos, por unidad de tiempo, de los incrementos condicionados, que coinciden con los anteriores y se conocen como momentos infinitesimales. En efecto, de

$$A_n(x, t) = \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E}((X_{t+h} - X_t)^n | X_t = x)$$

para $n = 1, 2$, se tiene, para un incremento pequeño h ,

$$\begin{aligned} \mathbb{E}(X_{t+h} - X_t | X_t = x) &\approx A_1(x, t) h, \\ \mathbb{E}((X_{t+h} - X_t)^2 | X_t = x) &\approx A_2(x, t) h, \end{aligned}$$

y entonces

$$\begin{aligned} \mathbb{V}(X_{t+h} - X_t | X_t = x) &\approx A_2(x, t) h - A_1(x, t)^2 h^2 \\ \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{V}(X_{t+h} - X_t | X_t = x) &= A_2(x, t), \end{aligned}$$

lo que justifica los nombres de media y varianza infinitesimal para A_1 y A_2 respectivamente, que coinciden con los truncados para $n = 1, 2$.

Estas funciones caracterizan al proceso de difusión, que puede expresarse en forma de ecuación diferencial estocástica

$$dX_t = A_1(X_t, t) dt + A_2(X_t, t)^{\frac{1}{2}} dW_t,$$

donde W_t es el proceso estándar de Wiener. La solución de esta ecuación es única bajo condiciones de tipo Lipschitz, siendo un proceso de Markov y, más concretamente, una difusión bajo condiciones de continuidad de las funciones (véase Arnold [2]).

Por otro lado, bajo ciertas condiciones de regularidad, la función de densidad de transición, $f := f(x, t|y, s)$, para $s \leq t$, es solución de las ecuaciones de Fokker-Planck y Kolmogorov, respectivamente,

$$\begin{aligned} \frac{\partial}{\partial t}f + \frac{\partial}{\partial x}(A_1(x, t)f) - \frac{1}{2} \frac{\partial^2}{\partial x^2}(A_2(x, t)f) &= 0 \\ \frac{\partial}{\partial s}f + A_1(y, s) \frac{\partial}{\partial y}f + \frac{A_2(y, s)}{2} \frac{\partial^2}{\partial y^2}f &= 0. \end{aligned}$$

Volviendo ahora al citado trabajo de Feller, el autor introduce un ruido intrínseco en la ecuación logística (1.2), estudiando así la ecuación diferencial estocástica

$$dN_t = (\alpha N_t - \beta N_t^2) dt + \gamma \sqrt{N_t} dW_t, \quad (1.4)$$

siendo $\gamma > 0$ y W el proceso estándar de Wiener. Sin embargo, las ecuaciones de Kolmogorov asociadas al proceso de difusión solución de (1.4) no han sido resueltas.

Numerosos modelos han sido introducidos desde entonces siguiendo una metodología similar. Una versión general de la ecuación diferencial estocástica asociada al crecimiento logístico aparece en el trabajo de Schurz [25]. Un resumen de algunos de ellos puede verse en Tuckwell y Koziol [31]. Entre ellos, cabe destacar el modelo de Capocelli y Ricciardi [5],

$$dN_t = (a - b \log N_t) N_t dt + c N_t dW_t,$$

obtenido tras incluir una función de regulación en el modelo Malthusiano. Aquí consiguen obtener la función de densidad de transición; sin embargo, el modelo solo exhibe un comportamiento aproximadamente logístico para variabilidades muy bajas, o exactamente logístico para variabilidad nula.

Como puede observarse, lo que se hace hasta el momento es partir de la ecuación (1.2) e introducir una componente de volatilidad asociada a un ruido (proceso de Wiener). Estos modelos dan soluciones que, para volatilidades

bajas, se acercan a la curva logística, pero no así cuando la variabilidad es alta. Además, en la mayoría de los casos no pueden resolverse las ecuaciones de Kolmogorov, lo que impide contar con la función de densidad de transición, necesaria para procedimientos de inferencia o tiempos de primer paso. Y aún teniéndola, debe tenerse en cuenta la dificultad que entraña la obtención de la función media, por lo que nunca se puede garantizar con certeza absoluta la cercanía al comportamiento logístico.

Dadas las circunstancias, los procedimientos de estimación empleados para estos modelos suelen basarse en técnicas de muestreo continuo, como se observa en Giovanis y Skiadas [13], donde consideran la estimación por máxima verosimilitud desde el muestreo continuo de las trayectorias.

Ante todo ello, se hace necesario establecer el marco teórico para definir procesos con un comportamiento medio de tipo logístico y que además permitan desarrollar procedimientos de estimación adecuados. Los trabajos de Román-Román y Torres-Ruiz [23] y Barrera-García, Román-Román y Torres-Ruiz [3] avanzan hacia estos objetivos consiguiendo procesos con media logística y caracterizados por funciones que, bajo condiciones de normalidad, puedan ser estimadas de forma completa desde los datos.

1.4. Proceso de difusión con media logística

El problema referente a la resolución de las ecuaciones de Kolmogorov representa una limitación importante pues condiciona la obtención explícita de la función de densidad de transición de probabilidad, necesaria en procedimientos de inferencia, como es el caso de la máxima verosimilitud a partir del muestre discreto de las trayectorias. Incluso aún cuando se obtiene tal densidad, suele ocurrir que no se consiga la expresión del valor esperado de la difusión solución. Estos inconvenientes apenas han permitido realizar estudios con un enfoque teórico o, en el mejor de los casos, abordando una aplicación práctica basándose en derivaciones de la ecuación diferencial estocástica mediante muestreo continuo o a través de la discretización de la ecuación diferencial. Es por ello que se hacía necesario un enfoque distinto pero de tal manera que se mantuviera la idea de obtener un proceso cuyo comportamiento se asemejara al de tipo logístico, resultando, sin embargo,

en unas ecuaciones de Kolmogorov resolubles. A tal efecto, y en el caso de la curva logística, encontramos el trabajo de Román-Román y Torres-Ruiz [23], donde, siguiendo la metodología de Gutiérrez-Jáimez et al. [15] para la curva Gompertz y de Román-Román, Romero y Torres-Ruiz [22] para la curva de Von Bertalanffy, parten de la ecuación de Langevin de un modelo malthusiano

$$dX_t = h(t)X_t dt + \sigma X_t dW_t \quad (1.5)$$

con fertilidad $h(t)$ dependiente del tiempo. Tal función $h(t)$ se define como el cociente de la derivada de una curva $c(t)$ y la curva misma, por lo que se tiene $h(t) = \dot{c}(t)/c(t)$. De la ecuación (1.5), imponiendo la condición de continuidad sobre $h(t)$, (véase, por ejemplo Arnold [2]) se obtiene como solución un proceso lognormal con factores exógenos. Este tipo de procesos ha sido ampliamente estudiado (véase, por ejemplo, el trabajo de Gutiérrez et al. [14]) debido a las numerosas posibilidades que ofrece para la obtención de características explícitas, como es el caso de la función media $\mathbb{E}(x)$, que cumple

$$\mathbb{E}(x) = \mathbb{E}(X_0) \exp \int_{t_0}^t h(s) ds$$

para la condición inicial $X_0 = X_{t_0}$. Nótese ahora que, por la definición de $h(t)$, esta función media es exactamente la curva $c(t)$. Por otro lado, para el proceso lognormal con factores exógenos se puede calcular de manera explícita su función de densidad de transición, solventando así el problema de otros enfoques previos.

Más en detalle, y antes de la obtención de la difusión según las ideas expuestas anteriormente, Román-Román y Torres-Ruiz [23] proponen una reformulación de la curva logística (1.3), basándose en la necesidad de conseguir una cota dependiente del valor inicial x_0 . Esto se debe a que en numerosas situaciones de aplicación real pueden existir individuos cuyas características a estudio exhiban un comportamiento de tipo logístico similar, pero con diferentes estados iniciales y, cabe esperar, capacidades máximas (valores límites). Por lo tanto, con esta idea, definiendo una versión general de la curva logística $f(t) = a/(1 + b e^{-t})^{-1}$ para $a, b, c > 0$ e imponiendo la condición inicial

$f(t_0) = x_0$, se llega a la expresión, para $t \geq t_0$,

$$f(t) = x_0 \frac{1 + b e^{-ct_0}}{1 + b e^{-ct}}.$$

Esta función es similar a (1.3) para $b = (\alpha/(\beta x_0) - 1)e^{\alpha t_0}$ y $c = \alpha$, verificando que su cota es ahora dependiente del valor inicial, y conservando la propiedad inherente a la curva logística referente a la inflexión, alcanzándose esta justo en la mitad del crecimiento total.

Una vez reformulada la curva en el contexto de una cota dependiente del valor inicial, se propone la metodología explicada al inicio de la sección para obtener un proceso de difusión X_t tal que su función media condicionada al valor inicial cumpla $\mathbb{E}(X_t | X_{t_0} = x_0) = f(t)$ para todo t en el espacio $[t_0, T]$. Para ello, se define una función $h : [t_0, T] \rightarrow \mathbb{R}$ como

$$h(t) = \frac{\dot{f}(t)}{f(t)} = \frac{bc}{e^{ct} + b}.$$

Partiendo ahora de una generalización del modelo de crecimiento de Malthus, se considera una fertilidad dependiente del tiempo, $h(t) + \sigma W_t$, donde $\sigma > 0$ y W_t es el proceso estándar de Wiener, para lograr así la ecuación de Langevin, expresada en la forma de la ecuación diferencial estocástica (1.5) con distribución inicial $X_{t_0} = x_0$ (degenerada, en el sentido $P(X_{t_0} = x_0) = 1$) independiente de W_t para todo $t \geq t_0$. La solución de (1.5), vía integración en el sentido de Itô, es un proceso de difusión lognormal con factores exógenos (proceso no homogéneo) con valores en \mathbb{R}^+ y caracterizado por el *drift* $h(t)x$ y el coeficiente de difusión $\sigma^2 x^2$. De Gutiérrez et al. [14] se obtiene la distribución condicionada de tal proceso,

$$(X_t | X_s = y) \sim \Lambda \left(\log y + \log \frac{1 + b e^{-cs}}{1 + b e^{-ct}} - \frac{\sigma^2}{2}(t-s); \sigma^2(t-s) \right),$$

donde $(X_t | X_s = y)$ es la variable X_t condicionada a que en el instante s tenga valor y , esto es, $X_s = y$, para $s \leq t$.

La intención final de todo este desarrollo es aprovechar la disponibilidad de la función de densidad de transición, así como de otros momentos de interés del proceso, para abordar problemas de inferencia. A tal efecto, considerando

el método de máxima de verosimilitud, la función a estudiar resulta en la forma

$$\mathbb{L}(\mu_1, \sigma_1^2, b, c, \sigma^2) = \prod_{i=1}^d \left(f_1(x_{i1}) \prod_{j=2}^n f(x_{ij}, t_{ij} | x_{ij-1}, t_{ij-1}) \right),$$

para una muestra de d trayectorias en tiempos t_{ij} con $i = 1, \dots, d$, $j = 1, \dots, n$ y $t_{i1} = t_1$, observando valores $x(t_{ij}) = x_{ij}$. En este caso, f_1 es la densidad de la variable $X_{t_1} \sim \Lambda(\mu_1, \sigma_1^2)$. Si se considerase una distribución inicial degenerada, esto es $P(X_{t_1} = x_1) = 1$, se tendría

$$\mathbb{L}(\mu_1, \sigma_1^2, b, c, \sigma^2) = \prod_{i=1}^d \prod_{j=2}^n f(x_{ij}, t_{ij} | x_{ij-1}, t_{ij-1}).$$

La función de verosimilitud cuenta entonces con los parámetros añadidos μ_1 y σ_1^2 , si bien estos dependen únicamente de los valores iniciales y por tanto no influyen en la estimación de los demás parámetros b, c y σ^2 .

De esta función se obtiene el sistema de ecuaciones de verosimilitud, en este caso de gran complejidad para su resolución debido a la naturaleza de la densidad, la cantidad de parámetros y el volumen de datos que pueda generar la observación del fenómeno. Por ello, es necesario el uso de métodos numéricos para resolver las ecuaciones y conseguir así las estimaciones máximo verosímiles de los parámetros. Con este fin, Román-Román y Torres-Ruiz [23] proponen trabajar con la expresión del punto de inflexión de la curva logística y su relación con los parámetros, aprovechando el hecho de que, en ese punto, la población alcanza la mitad de su capacidad máxima. Definiendo $\gamma = e^{-c}$ y $\eta = 1/b$, si la inflexión se alcanza en t_* , obtienen las relaciones

$$\begin{aligned} \eta &= \gamma^{t_*}, \\ \gamma &= \left(\frac{k}{x_0} - 1 \right)^{\frac{1}{t_0 - t_*}}, \end{aligned}$$

siendo la cota superior k y el valor inicial x_0 . De esta manera, estimando un valor para t_* consiguen el de γ y de estos el de η . Para tal estimación del punto de inflexión, se proponen, básicamente, dos métodos; bien tomando la media de los instantes en los que se alcanza la mitad del valor total para cada

trayectoria, bien considerando una aproximación de la derivada de la función a partir de la media y calculando el instante en el que se alcanza el máximo.

A fin de contrastar este modelo, los autores realizan un estudio comparativo con el propuesto por Giovanis y Skiadas [13], exhibiendo un comportamiento de tipo logístico para variabilidad baja, pero con el inconveniente de carecer de expresión para la densidad de transición ni la función media, abordando cuestiones de estimación mediante técnicas de muestreo continuo.

Mostrando mayor precisión en las estimaciones, el modelo propuesto por los autores es aplicado a datos simulando un cultivo microorgánico, aprovechando así el hecho conocido de que estos presentan patrones logísticos en su crecimiento.

1.5. Proceso de difusión gaussiano con media logística

Una vez obtenido un proceso de difusión de manera que su función media sea la curva logística y además la función de densidad de transición de probabilidad sea calculable de manera explícita, cabe preguntarse por las posibilidades prácticas de aplicar este tipo de modelos a situaciones reales. Con ello se podrían realizar estudios propios del campo en el que se aplica sobre la base de un modelo que describa adecuadamente el comportamiento a estudio. Además, la modelización matemática permitiría el desarrollo de procedimientos de simulación que sean de utilidad en las primeras etapas de nuevas investigaciones. Para alcanzar estos objetivos, los modelos obtenidos mediante procedimientos teóricos como el mencionado anteriormente, deben ajustarse a los datos observados, esto es, la tendencia de la difusión debe asemejarse en lo posible a la tendencia observada en los datos. Esto, sin embargo, requiere de la identificación de la tendencia del fenómeno dinámico a estudio, en particular, del fenómeno de crecimiento susceptible de ser descrito por alguna curva de crecimiento conocida, como por ejemplo la curva logística. Tal aproximación heurística no es, desafortunadamente, posible en todos los casos e, incluso cuando lo es, cabe la posibilidad de encontrar varios modelos aplicables, es decir, diferentes procesos estocásticos que se ajusten correctamente a los datos

observados.

Además, el desarrollo de tales procesos se basa en la introducción de un ruido en la ecuación diferencial ordinaria del correspondiente modelo determinista (por ejemplo, una curva de crecimiento). El resultado es una ecuación diferencial estocástica cuya solución es un proceso con varianza infinitesimal proporcional al cuadrado de la variable de estado, esto es, algo del tipo $\sigma^2 x^2$, que no necesariamente se ajuste a los datos disponibles.

Por lo tanto, aún cuando un modelo puede ser útil en la descripción de un fenómeno dinámico basándose en su tendencia, el hecho de no controlar el ajuste de la variabilidad puede incurrir en problemas a la hora de trabajar con otras características de interés como pueden ser los tiempos de primer paso, donde la variabilidad juega un papel crucial.

Con todo esto en mente, se juzga necesario construir modelos estocásticos que tomen en consideración toda la información disponible a fin de determinar, de la manera más aproximada posible, la forma de los momentos infinitesimales (*drift* y coeficiente de difusión) que caracterizan al proceso de difusión.

En esta sección se construirá un proceso de difusión gaussiano a partir de las observaciones, en diversos instantes, del fenómeno a estudio. La condición de normalidad permitirá trabajar con numerosas características explícitas del proceso, consiguiendo así ajustar el mismo adecuadamente a la tendencia y la variabilidad observadas. Todo ello permitiría abordar problemas de ajuste, predicción o tiempos de primer paso, entre otros.

1.5.1. Modelo gaussiano

El modelo dinámico que consideraremos viene definido por la siguiente ecuación diferencial estocástica,

$$dX_t = a(t)X_t dt + b(t)dW_t, \quad (1.6)$$

donde a y b son funciones continuas definidas en un intervalo real $[t_0, T]$, con $b(t) > 0$, $t_0, T \in \mathbb{R}^+$, W_t el proceso estándar de Wiener, y la condición inicial $X_{t_0} = x_0 \in \mathbb{R}^+$ (en el sentido de $P(X_{t_0} = x_0) = 1$) o $X_{t_0} \sim N(\mu_0, \sigma_0^2)$.

La solución de (1.6) es el proceso de difusión gaussiano

$$X_t = \left(X_{t_0} + \int_{t_0}^t b(s) e^{-\int_{t_0}^s a(u) du} dW_s \right) \exp \int_{t_0}^t a(s) ds, \quad (1.7)$$

caracterizado por los momentos infinitesimales

$$\begin{aligned} A_1(x, t) &= a(t)x, \\ A_2(t) &= b^2(t). \end{aligned} \quad (1.8)$$

Nótese que el caso relativo a una distribución inicial degenerada correspondería a la situación real en la cual todas las trayectorias del proceso comenzaran en el mismo valor x_0 .

Características del proceso

Al objeto de ajustar el modelo a los datos, debemos abordar dos funciones clave, a saber, la función media y la función varianza. Es más, de cara a obtener la expresión de la densidad de transición de probabilidad, serán necesarias las distribuciones finito-dimensionales, y para estas, la función de covarianza habrá de ser calculada.

Así pues, definiendo la función media del proceso (1.7) como $m(t) := \mathbb{E}(X_t)$, esta verifica la siguiente ecuación diferencial ordinaria,

$$\dot{m}(t) - a(t)m(t) = 0, \quad (1.9)$$

con la condición inicial $m_0 = \mathbb{E}(X_{t_0})$. Por lo tanto, se tendrá

$$m(t) = m_0 e^{\int_{t_0}^t a(s) ds}.$$

La función varianza, $v(t) := \mathbb{V}(X_t)$, verifica, por otro lado,

$$\dot{v}(t) - 2a(t)v(t) = b^2(t), \quad (1.10)$$

con la condición inicial $v_0 = \mathbb{V}(X_{t_0})$, luego

$$v(t) = e^{2\int_{t_0}^t a(s) ds} \left(v_0 + \int_{t_0}^t b^2(s) e^{-2\int_{t_0}^s a(u) du} ds \right). \quad (1.11)$$

Por otro lado, para cualesquiera instantes $s, t \in [t_0, T]$, la función covarianza viene dada por la expresión

$$c(s, t) = e^{\int_{t_0}^s a(u) du + \int_{t_0}^t a(u) du} \left(v_0 + \int_{t_0}^{s \wedge t} b^2(u) e^{-2\int_{t_0}^u a(r) dr} du \right),$$

donde $s \wedge t = \min(s, t)$.

Ahora, puesto que el proceso de difusión es gaussiano, sabemos que el vector $(X_{t_1}, \dots, X_{t_n})'$, para n instantes $t_1, \dots, t_n \in [t_0, T]$, sigue una distribución normal multivariante n -dimensional $N_n(\mathbf{m}, \mathbf{C})$ donde $\mathbf{m} = (m(t_1), \dots, m(t_n))'$ es el vector de medias y $\mathbf{C} = (c_{ij})$ es la matriz $n \times n$ de covarianzas con elementos $c_{ij} = c(t_i, t_j)$ para $i, j \in \{1, \dots, n\}$.

Finalmente, de las distribuciones marginales uni y bidimensional, se obtiene, para instantes $\tau < t$, la función de densidad de transición

$$f(x, t | y, \tau) = \left(2\pi e^{2\int_{t_0}^t a(\theta) d\theta} \int_{\tau}^t e^{-2\int_{t_0}^u a(\theta) d\theta} b^2(u) du \right)^{-\frac{1}{2}} \times \exp \left(-\frac{(x - ye^{\int_{\tau}^t a(\theta) d\theta})^2}{2e^{2\int_{t_0}^t a(\theta) d\theta} \int_{\tau}^t e^{-2\int_{t_0}^u a(\theta) d\theta} b^2(u) du} \right).$$

1.5.2. Metodología y ajuste

Supongamos un conjunto de números reales $\{x_{ij}\}$, para $i = 1, \dots, d$ y $j = 1, \dots, n$, resultantes de la observación, en los instantes t_1, \dots, t_n , de un fenómeno dinámico, esto es, la evolución en el tiempo de una determinada característica de interés (por ejemplo, el peso de un animal, el volumen de un tumor, el número de células con una propiedad concreta, etc.) para una muestra de d individuos.

El objetivo es construir un proceso de difusión gaussiano que explique la dinámica del fenómeno haciendo uso de la información aportada por los datos

observados; todo ello sin seleccionar previamente ningún modelo predeterminado ni hacer restricciones sobre el mismo.

Debido a que la difusión viene completamente caracterizada por sus momentos infinitesimales, así como por la distribución inicial, enfocaremos la construcción del proceso en la obtención de las funciones a y b en (1.8). Por otro lado, los datos iniciales de la muestra determinarán la distribución inicial.

Aproximación de la media infinitesimal

De la expresión (1.9) se deduce que la función a , siendo m diferenciable y tal que $m(t) \neq 0$ para todo t , puede expresarse como

$$a(t) = \frac{\dot{m}(t)}{m(t)}. \quad (1.12)$$

Por lo tanto, para calcular a debemos encontrar una función diferenciable que se aproxime a la función media del proceso. A tal efecto caben distintas posibilidades, a saber,

1. Seleccionar, caso de existir, una curva de crecimiento conocida o, de manera general, una función ρ -paramétrica

$$m_\rho : [t_0, T] \rightarrow \mathbb{R}$$

que describa adecuadamente el patrón de crecimiento del fenómeno a estudio. Algunos ejemplos serían la curva exponencial, para el caso de crecimientos rápidos sin acotación, o las curvas logística, Gompertz o de Von Bertalanffy, cubriendo patrones de crecimiento sigmoideal, acotado y con un punto de inflexión. El parámetro ρ (que puede ser uni o multidimensional) sería estimado, por ejemplo, mediante regresión por mínimos cuadrados sobre el conjunto de datos disponible, siendo el valor obtenido $\hat{\rho}$.

2. Desde el conjunto formado por la media de los datos para cada $j = 1, \dots, n$, esto es, los valores $m_j = d^{-1} \sum_{i=1}^d x_{ij}$, puede aproximarse una función m_* mediante interpolación.

Siguiendo entonces (1.12), definimos la aproximación de $a(t)$ como, para el primer caso,

$$a_* \equiv \frac{\dot{m}_\rho}{m_\rho},$$

o, para el segundo caso,

$$a_* \equiv \frac{\dot{m}_*}{m_*}.$$

Aproximación de la varianza infinitesimal

Una vez obtenida la aproximación de la función a , partimos de la ecuación (1.10) con a_* y buscamos una aproximación de la función b , cuyo cuadrado será la varianza infinitesimal del proceso. A tal fin parece lógico considerar la variabilidad aportada por los datos, esto es, la varianza muestral dada por el conjunto de valores

$$v_j = d^{-1} \sum_{i=1}^d (x_{ij} - m_j)^2, \quad (1.13)$$

para $j = 1, \dots, n$.

Sin embargo, al contrario que con la media muestral, la varianza de las observaciones puede presentar un alto grado de fluctuación por su propia naturaleza, condicionando así la elección para el procedimiento de aproximación de la función v . De hecho, intentar obtener una función v_* , aproximación de v , mediante técnicas de interpolación desde los valores (1.13) puede llevar a resultados erróneos si la variabilidad de la muestra es alta. Esto se debe a que el procedimiento de interpolación considera una función diferenciable, lo que contrasta con la naturaleza de los datos.

Por todo ello, en el caso de la varianza infinitesimal, es recomendable seguir el método consistente en buscar una función θ -paramétrica, con θ parámetro uni o multidimensional,

$$v_\theta : [t_0, T] \rightarrow \mathbb{R}^+$$

diferenciable en $[t_0, T]$ y cuya forma funcional puede ser sugerida por la misma disposición geométrica de los datos en un diagrama de dispersión estándar. Como en el caso de la función media, los parámetros de la función candidata pueden ser estimados según distintos procedimientos al objeto de

encontrar finalmente la función $v_{\hat{\theta}}$, aproximación de v . De esta manera, la aproximación de b vendría dada por

$$b_*(t) = \left(\dot{v}_{\hat{\theta}}(t) - 2 a_*(t) v_{\hat{\theta}}(t) \right)^{\frac{1}{2}}. \quad (1.14)$$

Nótese ahora que la función $v_{\hat{\theta}}$ seleccionada debe cumplir, para todo t en el intervalo de observación $[t_0, T]$,

$$\dot{v}_{\hat{\theta}}(t) > 2 a_*(t) v_{\hat{\theta}}(t). \quad (1.15)$$

Obviamente, esta condición es restrictiva en lo que concierne a la aplicación de la metodología propuesta al tipo de datos disponibles. De hecho, puede condicionar la elección de la función sugerida v_{θ} (por ejemplo, en el caso de procesos con tendencia positiva, esta función debería ser creciente) o incluso el rango de variación de los parámetros involucrados.

Para abordar esta cuestión, supongamos existe un intervalo contenido en $[t_0, T]$ en el cual no se cumple la condición (1.15) para la función v_{θ} . En ese caso, una propuesta inicial sería considerar b_* como el valor absoluto

$$b_{*1}(t) := \left| \dot{v}_{\hat{\theta}}(t) - 2 a_*(t) v_{\hat{\theta}}(t) \right|^{\frac{1}{2}}.$$

Sin embargo, de (1.11) se tiene que

$$v(t) = m^2(t) \left(\frac{v_0}{m_0^2} + \int_{t_0}^t b^2(u) m^{-2}(u) du \right),$$

por lo que, para las funciones finales b_* , b_{*1} , m_{\bullet} (con $\bullet = \hat{\rho}, *$, según se elija el método para aproximar la media), se tiene, para todo $t \in [t_0, T]$,

$$\int_{t_0}^t b_{*1}^2(u) m_{\bullet}^{-2}(u) du \geq \int_{t_0}^t b_*^2(u) m_{\bullet}^{-2}(u) du.$$

Es decir, el uso alternativo de la función b_{*1} puede conducir a un incremento importante de la variabilidad del proceso, haciéndole perder representatividad.

Para abordar esta cuestión, nos planteamos medir ese incremento y considerar entonces una constante de normalización que compense el exceso

de variabilidad derivado del uso de b_{*1} . Nótese ahora que, con la desigualdad anterior, ya conseguimos el primer objetivo. En efecto, el incremento de variabilidad viene dado por

$$\left(\int_{t_0}^T b_{*1}^2(u) m_{\bullet}^{-2}(u) du \right) \left(\int_{t_0}^T b_{*1}^2(u) m_{\bullet}^{-2}(u) du \right)^{-1},$$

por lo que definimos la constante (para cada elección de $v_{\hat{\theta}}$, m_{\bullet} y a_{*}),

$$\Delta_{v,m,a} := \left(\int_{t_0}^T b_{*1}^2(u) m_{\bullet}^{-2}(u) du \right)^{\frac{1}{2}} \left(\int_{t_0}^T b_{*1}^2(u) m_{\bullet}^{-2}(u) du \right)^{-\frac{1}{2}} \in \mathbb{R}. \quad (1.16)$$

De esta manera, a partir de la expresión (1.14), obtenemos finalmente

$$b_{*} \equiv \Delta_{v,m,a} b_{*1}. \quad (1.17)$$

Finalmente, si consideramos los casos de funciones paramétricas m_{ρ} o v_{θ} para la media y varianza respectivamente, debemos abordar el procedimiento de estimación por máxima verosimilitud al objeto de obtener las estimaciones paramétricas $\hat{\rho}$ o $\hat{\theta}$. Gracias al carácter gaussiano del proceso, teníamos acceso a la función de densidad de transición, por lo que la función de verosimilitud, definida como

$$\mathbb{L}(\rho, \theta) = \prod_{i=1}^d \prod_{j=2}^n f(x_{ij}, t_{ij} | x_{ij-1}, t_{ij-1}),$$

para valores $\{x_{ij}\}$ observados en instantes t_{ij} con $j = 1, \dots, n$ e $i = 1, \dots, d$ para d individuos (trayectorias del proceso), resulta

$$\begin{aligned} \mathbb{L}(\rho, \theta) &= \prod_{i=1}^d \prod_{j=2}^n \left(2\pi \left(v_{\theta}(t_{ij}) - \frac{m_{\rho}(t_{ij})}{m_{\rho}(t_{ij-1})} v_{\theta}(t_{ij-1}) \right) \right)^{-\frac{1}{2}} \\ &\quad \times \exp \left(- \frac{\left(x_{ij} - x_{ij-1} \frac{m_{\rho}(t_{ij})}{m_{\rho}(t_{ij-1})} \right)^2}{2 \left(v_{\theta}(t_{ij}) - \frac{m_{\rho}(t_{ij})}{m_{\rho}(t_{ij-1})} v_{\theta}(t_{ij-1}) \right)} \right). \end{aligned}$$

Este procedimiento, sin embargo, no es recomendable en tanto que, a las

restricciones teóricas sobre las funciones sugeridas para la media y varianza, se uniría la complejidad creciente de las ecuaciones de verosimilitud debido al número de parámetros involucrados y la misma forma funcional de las funciones sugeridas.

En el capítulo 3 abordaremos esta cuestión en un caso particular, proponiendo una metodología para su resolución mediante algoritmos metaheurísticos.

1.5.3. Aplicaciones

En esta sección pondremos en práctica los desarrollos teóricos anteriores aplicando la metodología propuesta a un conjunto de datos simulando el crecimiento de un cultivo de microorganismos.

Crecimiento de cultivos microorgánicos

Los datos considerados aquí proceden de la simulación realizada por Román-Román y Torres-Ruiz [23] en su estudio del proceso logístico resultante de (1.5) y comentado ya en este trabajo. Tales datos consisten en medidas de densidad de población en un cultivo, contando el número de individuos x por mililitro de solución acuosa al día durante 50 días, y considerando una ratio de crecimiento intrínseco de 0.25 por día, así como una densidad de equilibrio de 1000 individuos por mililitro. Tomando en total 50 cultivos (esto es, 50 trayectorias del proceso), partiendo todos en un mismo instante inicial (en días) $t_0 = 0$ con densidad inicial de 5 individuos por mililitro (es decir, estableciendo una distribución inicial degenerada en $x_0 = 5$ para $t_0 = 0$), se simulan las trayectorias para $\sigma = 0.01$.

Las gráficas de 1.2 muestran las trayectorias simuladas, así como la media y varianza observada.

Claramente, la media observada sugiere, por su forma sigmoïdal, el uso de la función logística como candidata a aproximar la función media del proceso. Por lo tanto, proponemos la función

$$m_{\rho}(t) = k(1 + b e^{-ct})^{-1},$$

donde $\rho = (k, b, c)$ es el parámetro de dimensión 3 para $k, a, b \in \mathbb{R}$. Mediante

estimación por mínimos cuadrados sobre los datos simulados, obtenemos las estimaciones $\hat{\rho} = (\hat{k}, \hat{b}, \hat{c})$ para $\hat{k} = 981.5475$, $\hat{b} = 195.3095$ y $\hat{c} = 0.2506$. La figura 1.3a muestra la media observada junto con la función propuesta $m_{\hat{\rho}}$ del proceso gaussiano.

Definimos ahora la media infinitesimal del proceso

$$a_{\star}(t) = \frac{\dot{m}_{\hat{\rho}}(t)}{m_{\hat{\rho}}(t)} = \hat{b}\hat{c}(e^{\hat{c}t} + \hat{b})^{-1}.$$

Por otro lado, para obtener b consideramos una aproximación de v de tipo exponencial,

$$v_{\theta}(t) = \exp\left(\sum_{i=1}^3 \theta_i t^i\right) - \exp\left(\sum_{i=1}^3 \theta_i t_0^i\right),$$

para el parámetro $\theta = (\theta_1, \theta_2, \theta_3)$. En este caso, aplicando un procedimiento de estimación por mínimos cuadrados sobre los datos simulados, obtenemos las estimaciones para $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ en los valores $\hat{\theta}_1 = 0.4477$, $\hat{\theta}_2 = -0.0076$ y $\hat{\theta}_3 = 3.9924 \times 10^{-5}$.

Ahora, con la función $v_{\hat{\theta}}$ comprobamos que, en efecto y como era de esperar, no se verifica la condición (1.15) en todo el intervalo temporal. En la figura 1.3b se muestra la función $\dot{v}_{\hat{\theta}}(t) - 2a_{\star}(t)v_{\hat{\theta}}(t)$ en $[0, 50]$, donde ya se aprecian los subintervalos que incumplen la condición. Siguiendo por tanto la metodología, calculamos $\Delta_{v,m,a}$ mediante (1.16) para $m_{\hat{\rho}}$, $v_{\hat{\theta}}$ y a_{\star} , obteniendo $\Delta = 0.3159$. Definimos con ello la función

$$b_{\star}(t) = 0.3159 \left| \dot{v}_{\hat{\theta}}(t) - 2a_{\star}(t)v_{\hat{\theta}}(t) \right|^{\frac{1}{2}}.$$

Completamos así el proceso gaussiano con media logística como solución de la ecuación diferencial estocástica

$$dX_t = a_{\star}(t)X_t dt + b_{\star}(t)dW_t,$$

con condición inicial $X_{t_0} = x_0 = 5$ en el intervalo $[0, 50]$. En la figura 1.4a se muestran la varianza observada y la obtenida tras el procedimiento.

Finalmente, para comprobar lo adecuado que es el modelo gaussiano obtenido, simulamos trayectorias de este nuevo proceso, presentadas en la figura 1.4b. Puede apreciarse el comportamiento similar al de los datos originales

(simulaciones mostradas en la figura 1.2). El resultado es pues un modelo aproximado del patrón de crecimiento mostrado en el fenómeno original.

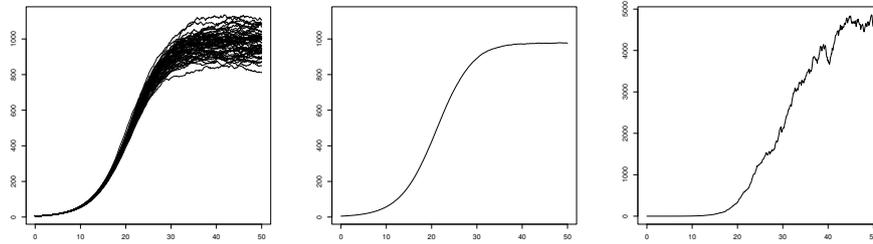


Figura 1.2: Trayectorias simuladas, media observada y varianza observada.

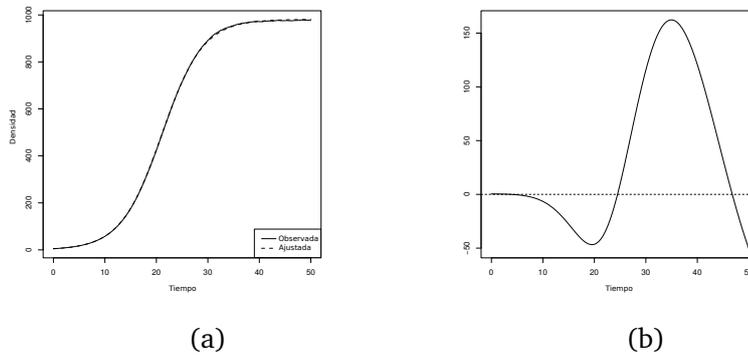


Figura 1.3: (a) Media observada y función media ajustada. (b) Función $\hat{v}_\theta(t) - 2a_*(t)v_\theta(t)$.

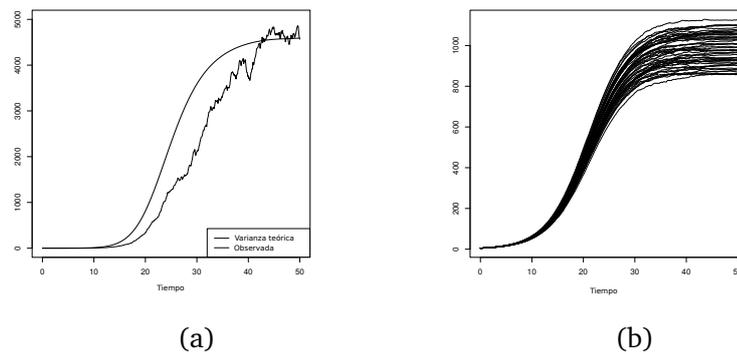


Figura 1.4: (a) Varianzas observada y ajustada. (b) Trayectorias simuladas del nuevo proceso.

Capítulo 2

Proceso de difusión logístico generalizado

2.1. Introducción	45
2.2. Una generalización de la ecuación logística	46
2.3. Difusión logística generalizada	47
2.4. Aplicabilidad del modelo	52
2.5. Aplicaciones	60
2.6. Extensión a otros modelos	66

2.1. Introducción

Buena parte de la importancia de la curva logística estriba en el gran número de modificaciones que de la misma han ido surgiendo, lo que ha derivado en una amplia variedad de modelos que aportan descripciones precisas de fenómenos concretos con un alto grado de sofisticación. Esto lleva a plantearse, de manera natural, las posibles generalizaciones de la curva logística, al objeto de estudiar todo su potencial y ver algunas de sus modificaciones como casos particulares. En efecto, las generalizaciones de la curva se han producido en un sentido paramétrico, esto es, introduciendo nuevos parámetros en la curva como forma de extenderla al ámbito más general posible. Véase a este respecto el trabajo de Tsoularis y Wallace [30] presentando las curvas de tipo logístico más usadas.

En este capítulo consideramos, sin embargo, un enfoque funcional de la generalización. Es decir, procedemos a extender el modelo logístico mediante la introducción de una función (paramétrica), en lugar de solo unos parámetros. Algunos casos concretos de esta función darán lugar a curvas particulares derivadas de la logística.

2.2. Una generalización de la ecuación logística

En lugar de una generalización paramétrica de la curva logística, consideremos el modelo genérico desde un punto de vista funcional. Para ello, sea $g : [t_0, T] \rightarrow \mathbb{R}$ una función continua, integrable en el intervalo de observación $[t_0, T] \subset \mathbb{R}$ para $t_0 \geq 0$ y tal que su primitiva verifique $\int_{t_0}^t g(u) du \rightarrow \infty$ cuando $t \rightarrow \infty$. Expresamos ahora el modelo de crecimiento clásico de Malthus mediante un parámetro referente a la ratio de crecimiento $\alpha > 0$, un inhibidor $\beta > 0$ (en general se tiene $\beta < \alpha$) y la función recién definida g . De esta manera se tiene la ecuación diferencial ordinaria

$$\frac{dx(t)}{dt} = (\alpha x(t) - \beta x^2(t)) g(t) \quad (2.1)$$

con condición inicial $x(t_0) = x_0 \in \mathbb{R}$. La solución a esta ecuación es la curva

$$x(t) = \frac{\alpha}{\beta} \left(1 + \left(\frac{\alpha}{\beta x_0} - 1 \right) \exp \left(-\alpha \int_{t_0}^t g(u) du \right) \right)^{-1},$$

creciente solo cuando $\alpha > \beta x_0$, con límite $\lim_{t \rightarrow \infty} x(t) = \alpha/\beta$ y exhibiendo un punto de inflexión, si $\alpha > \beta x_0$, en el instante t_* tal que

$$\int_{t_0}^{t_*} g(u) du = \frac{1}{\alpha} \log \left(\frac{\alpha}{\beta x_0} - 1 \right).$$

Nótese que, para distintas funciones g , se pueden obtener curvas conocidas. En efecto,

- Si $g(t) = \delta \in \mathbb{R}$, entonces la curva $x(t)$ es la curva logística clásica.

- Si $g(t) = \sum_{i=0}^n a_i x^i$ es un polinomio de grado n con coeficientes $a_i \in \mathbb{R}$, la curva resultante será multisigmoidal.
- Si $g(t) = \delta + \theta (1 + t^2)^{-\frac{1}{2}}$, para $\delta, \theta \in \mathbb{R}$, la curva $x(t)$ se convierte en la curva hiperbolística de tipo I, introducida por Tabatabai, Williams y Bursac [29] y que trataremos en el capítulo 3.

Poniendo el foco en la expresión de la curva, y no ya en la ecuación diferencial ordinaria que la genera, podemos reescribirla, sin pérdida de generalidad, en la forma

$$x(t) = a \left(1 + b e^{-g(t)}\right)^{-1},$$

donde a y b son parámetros reales y $g(t) := \int^t g(u) du$ es una primitiva de g . Ahora, $\lim_{t \rightarrow \infty} x(t) = a$ y el punto de inflexión se alcanza en el instante t_* tal que

$$g(t_*) = \log b.$$

Podemos introducir ahora el objeto matemático con el que trabajaremos. Para ello, sea \mathcal{C}_∞^1 el espacio de funciones reales f con primera derivada y tales que $\lim_{t \rightarrow \infty} f(t) = \infty$. Definimos ahora el *operador logístico* $L_{ab} : \mathcal{C}_\infty^1 \rightarrow \mathcal{C}_\infty^\infty$ como

$$L_{ab} f(t) := a \left(1 + b e^{-f(t)}\right)^{-1},$$

para cada par de puntos $a, b \in \mathbb{R}$ y donde $\mathcal{C}_\infty^\infty$ es el espacio de funciones reales infinitamente derivables. A efectos de notación, entenderemos

$$L_{ab}^{-1} f(t) := \frac{1}{L_{ab} f(t)}, \quad \dot{L}_{ab} f(t) := \frac{d}{dt} L_{ab} f(t).$$

Este enfoque basado en operadores funcionales nos permite unificar notación e identificar claramente cada curva de tipo logístico, caracterizada por los parámetros a y b , así como por la función g que compone con la exponencial en la expresión general de la curva.

2.3. Difusión logística generalizada

Siguiendo la metodología presentada en la sección 1.4, el objetivo es obtener un proceso de difusión cuya función media sea, en este caso, la curva

logística generalizada. Tal proceso dependerá, por tanto, de la función g .

En efecto, definimos la función $h : [t_0, T] \rightarrow \mathbb{R}$ como

$$h(t) = \mathcal{L}_{ab} g(t) \mathcal{L}_{ab}^{-1} g(t) = b g(t) (e^{g(t)} + b)^{-1}, \quad (2.2)$$

continua en un intervalo real $[t_0, T]$, para $b \in \mathbb{R}$. Con esto, consideramos la ecuación diferencial estocástica (1.5), esto es,

$$dX_t = b g(t) (e^{g(t)} + b)^{-1} X_t dt + \sigma X_t dW_t$$

para el parámetro¹ $\sigma > 0$, W_t el proceso de Wiener y condición inicial $X_{t_0} = X_0$, siendo X_0 una variable aleatoria positiva e independiente de W_t para todo $t \geq t_0$. La solución, recordemos, habida cuenta la continuidad de las funciones $h(t)x$ y σx , era un proceso de difusión $\{X_t : t_0 \leq t \leq T\}$ en tiempo continuo y con espacio de estados continuo, en concreto, el proceso lognormal con factores exógenos

$$X_t = \mathcal{L}_{ab} g(t) \exp\left(-\frac{\sigma^2}{2}(t - t_0) + \sigma(W_t - W_{t_0})\right),$$

caracterizado por los momentos infinitesimales $A_1(x, t) = h(t)x$ y $A_2(x, t) = \sigma^2 x^2$.

2.3.1. Características principales

Función de densidad de transición

Para obtener la función de densidad de transición de probabilidad del proceso de difusión recién definido, hacemos uso del siguiente resultado debido a Ricciardi [19], en el que establece las condiciones necesarias para

¹Es necesario advertir que la metodología propuesta en este capítulo pone el foco únicamente en el comportamiento de la tendencia de los datos, sin considerar el apartado relativo a la variabilidad de los mismos (que se representa en la ecuación diferencial estocástica mediante el parámetro σ).

exista una transformación² del tipo

$$\begin{aligned}x' &= \psi(x, t), & t' &= \phi(t), \\y' &= \psi(y, s), & s' &= \phi(s),\end{aligned}$$

de manera que

$$f(x, t|y, s) = \frac{\partial \psi(x, t)}{\partial x} f'(x', t'|y', s'), \quad (2.3)$$

donde f' es la conocida función de densidad de transición del proceso estándar de Wiener,

$$f'(x', t'|y', s') = (2\pi(t' - s'))^{-\frac{1}{2}} \exp\left(-\frac{(x' - y')^2}{2(t' - s')}\right).$$

El resultado comentado es el siguiente:

Teorema 1 (Ricciardi, 1976). *Un proceso de difusión con densidad de transición $f(x, t|y, s)$ y momentos infinitesimales $A_1(x, t)$ y $A_2(x, t)$ es transformable al proceso estándar de Wiener si y solo si existen funciones arbitrarias $C_1(t)$ y $C_2(t)$ tales que*

$$A_1(x, t) = \frac{1}{4} \frac{\partial A_2(x, t)}{\partial x} + \frac{A_2(x, t)^{\frac{1}{2}}}{2} \left(C_1(t) + \int_z^x \frac{C_2(t)A_2(y, t) + \frac{\partial A_2(y, t)}{\partial t}}{A_2(y, t)^{\frac{3}{2}}} dy \right).$$

En ese caso, la transformación es

$$\begin{aligned}x' &= \psi(x, t) = k_1^{\frac{1}{2}} e^{-\frac{1}{2} \int_{t_0}^t C_2(s) ds} \int_z^x A_2(y, t)^{-\frac{1}{2}} dy - \frac{k_1^{\frac{1}{2}}}{2} \int_{t_2}^t C_1(s) e^{-\frac{1}{2} \int_{t_0}^s C_2(u) du} ds + k_2, \\t' &= \phi(t) = k_1 \int_{t_1}^t e^{-\int_{t_0}^s C_2(u) du} ds + k_3,\end{aligned}$$

donde z es un valor en el intervalo de definición del proceso, $t_i \in \mathbb{R}^+$ y k_i son

²En el contexto de este resultado, se usará la notación original empleando el apóstrofo únicamente como distinción entre variables, sin relación alguna con derivadas o traspuestas.

constantes arbitrarias bajo la restricción $k_1 > 0$.

Aplicando este teorema se llega a la ecuación

$$h(t)x = \left(\frac{\sigma^2}{2} + \frac{\sigma}{2} C_1(t) \right) x + \frac{x \log \frac{x}{z}}{2} C_2(t),$$

que se verifica para las funciones

$$\begin{aligned} C_1(t) &= \frac{2h(t)}{\sigma} - \sigma, \\ C_2(t) &= 0. \end{aligned}$$

Por lo tanto, la transformación pedida es

$$\begin{aligned} \psi(x, t) &= \frac{\sqrt{k_1}}{\sigma} \left(\log \frac{x}{z} - \int_{t_2}^t h(u) du + \frac{\sigma^2}{2} (t - t_2) \right) + k_2, \\ \phi(t) &= k_1(t - t_1) + k_3. \end{aligned}$$

Finalmente, con tales transformaciones y la expresión (2.3), usando la definición (2.2), se tiene la función de densidad de transición del proceso de difusión logístico generalizado,

$$f(x, t | y, s) = \frac{1}{x} (2\pi\sigma^2(t-s))^{-\frac{1}{2}} \exp \left(- \frac{\left(\log \frac{x}{y} - \log \frac{L_{ab} g(t)}{L_{ab} g(s)} + \frac{\sigma^2}{2} (t-s) \right)^2}{2\sigma^2(t-s)} \right).$$

Por lo tanto, el proceso condicionado $(X_t | X_s = y)$ para $s < t$, útil en procedimientos de inferencia, sigue una distribución lognormal, esto es,

$$(X_t | X_s = y) \rightsquigarrow \Lambda_1 \left(\log y + \log \frac{L_{ab} g(t)}{L_{ab} g(s)} - \frac{\sigma^2}{2} (t-s); \sigma^2(t-s) \right).$$

Distribuciones finito-dimensionales y momentos

Para el estudio de las distribuciones finito-dimensionales, téngase en cuenta que la elección de la distribución inicial está condicionada por el hecho de obtener un proceso lognormal. Esto se asegura, como puede verse en Arnold [2], mediante una distribución degenerada o bien lognormal. Suponiendo

entonces $P(X_{t_0} = x_0) = 1$, esto es, una distribución inicial degenerada en el valor x_0 , la distribución unidimensional viene dada por la densidad³

$$\begin{aligned} f(x, t) &= f(x, t | x_0, t_0) \\ &= \frac{1}{x (2\pi\sigma^2(t - t_0))^{\frac{1}{2}}} \exp\left(-\frac{\left(\log \frac{x}{x_0} - \log \frac{L_{ab} g(t)}{L_{ab} g(t_0)} + \frac{\sigma^2}{2}(t - t_0)\right)^2}{2\sigma^2(t - t_0)}\right). \end{aligned}$$

Por la propiedad de Markov, asegurada en este tipo de procesos, la distribución bidimensional conjunta del par (X_t, X_s) , con $s < t$, vendrá dada por la densidad

$$\begin{aligned} f(x, t; y, s) &= f(y, s) f(x, t | y, s) \\ &= \frac{1}{2\pi xy} |\Sigma_1|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (\log \mathbf{x} - \boldsymbol{\mu})' \Sigma_1^{-1} (\log \mathbf{x} - \boldsymbol{\mu})\right), \end{aligned}$$

donde se han definido los vectores

$$\begin{aligned} \log \mathbf{x} &= (\log x, \log y)', \\ \boldsymbol{\mu} &= \left(\log x_0 + \log \frac{L_{ab} g(t)}{L_{ab} g(t_0)} - \frac{\sigma^2}{2}(t - t_0), \log x_0 + \log \frac{L_{ab} g(s)}{L_{ab} g(t_0)} - \frac{\sigma^2}{2}(s - t_0) \right)', \end{aligned}$$

y la matriz

$$\Sigma_1 = \sigma^2 \begin{pmatrix} t - t_0 & s - t_0 \\ s - t_0 & s - t_0 \end{pmatrix}.$$

En general, para cualesquiera s, t , la densidad conjunta seguirá una distribución lognormal bidimensional $\Lambda_2(\boldsymbol{\mu}, \Sigma)$ con

$$\Sigma = \sigma^2 \begin{pmatrix} t - t_0 & (t \wedge s) - t_0 \\ (t \wedge s) - t_0 & s - t_0 \end{pmatrix}.$$

Por otro lado, aprovechando estas distribuciones y la función de densidad de transición calculada anteriormente, podemos considerar la variable $Z_t =$

³En el caso de una distribución inicial lognormal, los cálculos son posibles, si bien más complejos que cuando se considera una distribución inicial degenerada. Por ello, es recomendable abordarlo mediante ecuaciones diferenciales estocásticas.

$\log X_t$, que es normal, pues $(X_t | X_s)$ es lognormal. Por lo tanto,

$$(Z_t | X_s = y) \rightsquigarrow N_1 \left(\log y + \log \frac{L_{ab} g(t)}{L_{ab} g(s)} - \frac{\sigma^2}{2}(t-s); \sigma^2(t-s) \right).$$

Así, el momento condicionado no centrado de orden k viene dado por

$$\begin{aligned} \mathbb{E}(X_t^k | X_s) &= \mathbb{E}(e^{k \log X_t} | X_s) \\ &= X_s^k \exp \left(k \log \frac{L_{ab} g(t)}{L_{ab} g(s)} + \frac{k\sigma^2(t-s)}{2} (k-1) \right), \end{aligned}$$

y con él podemos calcular el momento no centrado de orden k ,

$$\begin{aligned} \mathbb{E}(X_t^k) &= \mathbb{E}(\mathbb{E}(X_t^k | X_{t_0})) \\ &= \mathbb{E}(X_{t_0}^k) \exp \left(k \log \frac{L_{ab} g(t)}{L_{ab} g(t_0)} + \frac{k\sigma^2(t-t_0)}{2} (k-1) \right). \end{aligned}$$

En particular, la función media vendrá dada, para el caso concreto de distribución inicial degenerada en x_0 , esto es, $P(X_{t_0} = x_0) = 1$, por

$$\mathbb{E}(X_t) = \mathbb{E}(x_0) \exp \int_{t_0}^t h(u) du = L_{ab} g(t),$$

es decir, la curva logística generalizada en torno a la cual construimos el proceso de difusión.

Finalmente, usando $\mathbb{E}(X_t^2)$, calculamos la función varianza para $t \geq t_0$,

$$\mathbb{V}(X_t) = \left(\frac{L_{ab} g(t)}{L_{ab} g(t_0)} \right)^2 \mathbb{V}(X_{t_0}) e^{\sigma^2(t-t_0)} L_{ab} g(t)^2 (e^{\sigma^2(t-t_0)} - 1).$$

2.4. Aplicabilidad del modelo

Desde un punto de vista práctico, los desarrollos anteriores apenas tienen utilidad si no se conoce una forma funcional explícita de g o de su primitiva. Es por ello que el siguiente paso consistirá en estudiar la manera de conseguir tal función a partir de los datos observados.

Consideremos una muestra discreta $\{x_{ij}\}$ para $i = 1, \dots, d$ y $j = 1, \dots, n$ de

n observaciones de d trayectorias distintas realizadas en instantes de tiempo t_j , exhibiendo un comportamiento susceptible de ser aproximado por una curva de tipo logístico, esto es, una evolución en su tendencia de tipo sigmooidal. Supuestas observadas las trayectorias en los mismos instantes t_j , puede calcularse la media de los valores de cada trayectoria para cada instante de tiempo, de manera que se obtenga el conjunto $\{m_j\}$ de la media observada, donde, claro es, $m_j = d^{-1} \sum_{i=1}^d x_{ij}$.

Buscamos entonces una función real g , continua en el intervalo $[t_1, t_n]$, y parámetros $a, b \in \mathbb{R}$, tales que, para todo $j = 1, \dots, n$,

$$L_{ab} g(t_j) = m_j. \quad (2.4)$$

Despejando g en (2.4), se obtiene

$$g(t_j) = \log b - \log \left(\frac{a}{m_j} - 1 \right).$$

Por lo tanto, el conjunto $\{g(t_j)\}$ puede entenderse como una realización de la función g . Definamos ahora la función continua a -paramétrica $\rho_a : (0, a) \rightarrow \mathbb{R}$ como

$$\rho_a(x) = -\log \left(\frac{a}{x} - 1 \right).$$

Nótese que ρ_a no es más que la función *logit* reescalada en el intervalo $(0, a)$.

Finalmente, introducimos el conjunto de puntos de \mathbb{R}^2 por donde pasaría g ,

$$D := \left\{ (t_j, \log b + \rho_a(m_j)) \right\}_{j=1}^n \subset \mathcal{G}g \quad (2.5)$$

donde $\mathcal{G}g$ es el grafo de g . Se sigue entonces que una función candidata a g debe ajustarse necesariamente al conjunto D .

2.4.1. Parámetros desconocidos

El conocimiento completo de D está sujeto al conocimiento de los parámetros a y b de la curva. Con la información de la que se dispone, es posible obtener una estimación \hat{a} del parámetro a sabiendo que este representa la

capacidad máxima sostenible de la población cuyo crecimiento describe la curva de tipo logístico. Es decir, a es el límite de la curva cuando $t \rightarrow \infty$. Por ello, y con los datos provenientes de las observaciones, basta definir

$$\hat{a} = \sup_{j=1, \dots, n} m_j.$$

Nótese que, aún así, pueden existir $i_1 \in \{1, \dots, d\}$ y $j_1 \in \{1, \dots, n\}$ tales que $x_{i_1 j_1} \geq \hat{a}$, es decir, que alguna trayectoria corte a la asíntota horizontal $x = \hat{a}$. Sin embargo, este hecho puede ignorarse debido a que el objetivo es encontrar una curva de tipo logístico que se ajuste a los valores medios de los datos observados. Será suficiente, pues, trabajar con los valores m_j .

Esta definición de \hat{a} es puramente teórica, en el sentido de trabajar con un supremo. Obviamente, durante la puesta en práctica del procedimiento, se consideraría el máximo de los m_j más una constante $\varepsilon > 0$ a fin de evitar valores nulos para el logaritmo durante el cómputo de $\rho_{\hat{a}}$.

La presencia del parámetro b , por el contrario, supone un impedimento a la hora de seguir el desarrollo de la metodología. En efecto, no existen elementos que puedan aportar información suficiente como para estimar el valor de b , habida cuenta de que no conocemos siquiera la curva (que se construirá al final, una vez obtenida la función g).

Para tratar este inconveniente, sin embargo, no será necesario ningún procedimiento de estimación. Basta con prescindir de b en la definición del conjunto D , sin perjuicio por ello.

En efecto, es evidente el hecho de que el procedimiento de ajuste de una función a los datos (independientemente del método usado) es invariante por traslaciones, esto es, si una función $z(t)$ se ajusta a un conjunto S , entonces $z(t)+k$, con k constante, se ajustará al conjunto trasladado $S+k = \{s+k; s \in S\}$.

Sea por tanto $k = \log b$. Si una función real paramétrica diferenciable $\phi_{\star}^{\tau}(t)$, definida en el intervalo de observación y con τ parámetro uni o multi-dimensional, se ajusta al conjunto de puntos $C := \{(t_j, \rho_{\hat{a}}(m_j))\}_{j=1}^n$, es decir, existe $\hat{\tau}$ tal que, para todo j ,

$$\phi_{\star}^{\hat{\tau}}(t_j) = \rho_{\hat{a}}(m_j) + \varepsilon_j,$$

haciendo mínimo el vector de residuos $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ del procedimiento de

estimación, entonces la función

$$\phi^{\hat{\tau}} \equiv k + \phi_{\star}^{\hat{\tau}}$$

se ajustará a D (reemplazando a por \hat{a} en (2.5)). Es decir, para todo j ,

$$\phi^{\hat{\tau}}(t_j) = k + \phi_{\star}^{\hat{\tau}}(t_j) = \log b + \rho_{\hat{a}}(m_j) + \varepsilon_j = \log b - \log\left(\frac{\hat{a}}{m_j} - 1\right) + \varepsilon_j.$$

Obsérvese, mediante la ecuación anterior, que la idea de prescindir de b no significa que no se tenga en cuenta durante el proceso de ajuste. En realidad, se está realizando una estimación implícita de b por medio de la tautología $b = e^{\log b}$, aplicada en (2.4), es decir, ajustando una función ϕ a los puntos C estamos consiguiendo una estimación encubierta de b .

Así, $\forall t \geq t_0$ y una hipotética función $\phi^{\hat{\tau}}$ ajustada a D (para $a = \hat{a}$),

$$L_{\hat{a}1} \phi_{\star}^{\hat{\tau}}(t) = a \left(1 + \exp(-\phi_{\star}^{\hat{\tau}}(t))\right)^{-1} = a \left(1 + b \exp(-\phi^{\hat{\tau}}(t))\right)^{-1} = L_{\hat{a}b} \phi^{\hat{\tau}}(t).$$

Es decir, la curva logística generalizada es la misma para $\phi^{\hat{\tau}}$ con parámetros a y b , que para $\phi_{\star}^{\hat{\tau}}$ con parámetros a y 1. Resulta ahora claro que, en la estimación de τ se incluye la estimación implícita de b .

Por otro lado, es interesante destacar que, una vez obtenida la función $\phi_{\star}^{\hat{\tau}}$, podemos calcular el punto de inflexión de la curva logística generalizada $L_{\hat{a}1} \phi_{\star}^{\hat{\tau}}$ fácilmente. En efecto, si t_{\star} es el punto de inflexión, entonces, por la naturaleza logística, se habrá de alcanzar el valor $\hat{a}/2$. Es decir,

$$L_{\hat{a}b} \phi^{\hat{\tau}}(t_{\star}) = L_{\hat{a}1} \phi_{\star}^{\hat{\tau}}(t_{\star}) = \frac{\hat{a}}{2}.$$

Y de la última igualdad podemos deducir entonces que

$$\phi_{\star}^{\hat{\tau}}(t_{\star}) = 0,$$

es decir, el punto de inflexión anula a la función $\phi_{\star}^{\hat{\tau}}$. Además, como la curva coincide con aquella para la aproximación (no obtenida pues dependería de b) $\phi^{\hat{\tau}}$, podemos afirmar que ambas presentan la misma inflexión.

El modelo

Con la intención de definir el modelo de difusión para la curva logística generalizada, seguimos ahora la metodología de Román-Román y Torres-Ruiz [23] y definimos una función de fertilidad h_* para el modelo malthusiano, a la que se añade un ruido σW_t , al objeto de conseguir la ecuación diferencial estocástica cuya solución tenga por función media la curva logística generalizada.

Teniendo en cuenta que $\dot{\phi}^\tau \equiv \dot{\phi}_*^\tau$, podemos definir entonces una función $h_* : [t_0, t_n] \rightarrow \mathbb{R}$ como

$$h_*(t) = L_{\hat{a}b} \phi_*^{\hat{t}}(t)^{-1} \frac{d}{dt} L_{\hat{a}b} \phi_*^{\hat{t}}(t) = \frac{d}{dt} \log L_{\hat{a}1} \phi_*^{\hat{t}}(t) = \dot{\phi}_*^{\hat{t}}(t) (1 + \exp \phi_*^{\hat{t}}(t))^{-1}.$$

La ecuación diferencial estocástica (1.5), con la función h_* , tiene por solución un proceso de difusión lognormal con factores exógenos, con función media

$$\mathbb{E}(x_t) = x_0 L_{\hat{a}1} \phi_*^{\hat{t}}(t_0)^{-1} L_{\hat{a}1} \phi_*^{\hat{t}}(t) = L_{\hat{a}1} \phi_*^{\hat{t}}(t),$$

atendiendo al hecho de partir de una distribución inicial degenerada en x_0 , luego se tiene $m_1 = x_0$ y, por lo tanto, $\phi_*^{\hat{t}}(t_0) = \rho_{\hat{a}}(x_0)$.

2.4.2. Influencia de datos redundantes

Como se ha visto hasta ahora, nuestro enfoque práctico consiste en capturar el comportamiento sigmoïdal de los datos mediante el ajuste de una función $\phi_*^{\hat{t}}$, con la que posteriormente construir un proceso de tipo logístico. Sin embargo, debido a su propia naturaleza, los datos están sujetos a perturbaciones aleatorias que provocan una variabilidad creciente en su comportamiento. Cabe preguntarse entonces qué influencia tiene esta variabilidad en la metodología descrita y, de ser importante, cómo tratarla para asegurar buenos resultados.

En efecto, el carácter sigmoïdal de las observaciones consta de tres fases claramente diferenciadas, a saber, un crecimiento inicial lento, una fase de crecimiento exponencial en la que se cruza el umbral de la mitad de la capacidad máxima sostenible, y una estabilización final que se refleja en el comportamiento asintótico de la curva.

Es justo ahí donde los datos dejan de aportar información para nuestros propósitos. Puesto que existe una asíntota horizontal en a , sabemos que los valores medios, en ausencia de ruido, evolucionarán acercándose cada vez más a a . Sin embargo, la presencia de fluctuaciones aleatorias hace que estos valores oscilen, alterando la estructura interna del conjunto de observaciones y afectando al proceso de ajuste de la función candidata ϕ_{\star}^{τ} .

Por ello, se hace necesario trabajar con un conjunto truncado de observaciones, manteniendo únicamente aquellos puntos que aportan información sobre el comportamiento sigmoïdal de los datos. Esto lo conseguimos restringiendo el intervalo de observación para así centrarlo en torno al tiempo de inflexión t_{\star} . En efecto, una propuesta es considerar los índices de aquellos valores m_j que estén por encima de $\frac{a-x_0}{2}$, tomar el mínimo k y obtener así el primer instante t_k en el cual la población supera la mitad de su total. Es decir, sea

$$k = \min_{u=1,\dots,n} (u : 2x_u \geq a - x_0), \quad (2.6)$$

entonces tomaremos el intervalo de observación

$$\begin{cases} [t_0, t_{2k-1}] & \text{si } 2k \leq n, \\ [t_{2k-n}, t_n] & \text{si } 2k > n. \end{cases}$$

En efecto, contraemos el intervalo original $[t_0, t_n]$ por la derecha o la izquierda, según exista un *exceso* de datos, esto es, valores que no aportan información sobre el comportamiento sigmoïdal (valores cercanos a un comportamiento asintótico) pero que, en cambio, podrían alterar los procesos de ajuste propios de la metodología propuesta.

Una prueba elemental de este procedimiento puede observarse en la figura [2.1](#).

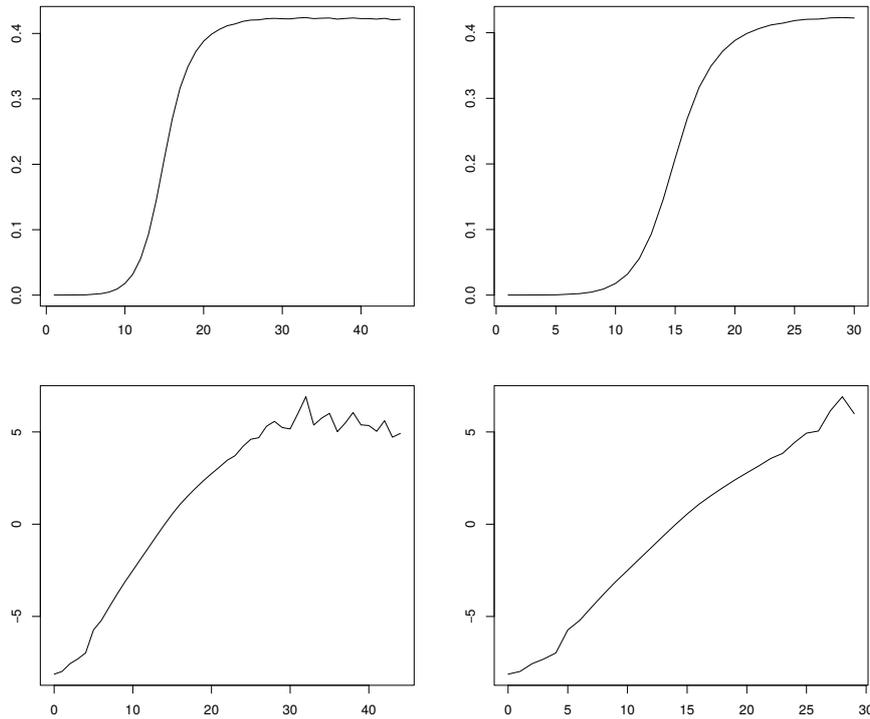


Figura 2.1: Comparación entre datos no truncados y truncados mediante la técnica descrita (figuras arriba, de izquierda a derecha). Abajo, de izquierda a derecha, imágenes de los datos originales (truncados y no) por $\rho_{\hat{a}}$. Nótese que los provenientes de los datos truncados son susceptibles de un ajuste mucho más sencillo, en este caso, de tipo lineal.

A pesar de que este procedimiento puede ser aplicado tanto a los valores iniciales como finales del conjunto de observaciones, el truncamiento en la parte inicial de los datos no tiene por qué ser tan necesaria como la final. En efecto, los valores cercanos a t_0 pueden presentar variaciones debido al comportamiento asintótico de ρ_a , pero la fluctuación aleatoria tiene menor influencia en estos puntos que en aquellos del tramo final del intervalo de observación, habida cuenta de la variabilidad creciente con la que se define el modelo estocástico.

2.4.3. Resumen del procedimiento

A continuación, y para clarificar lo expuesto en las secciones anteriores, se resumen los pasos del procedimiento propuesto:

1. A partir de datos observados $\{x_{ij}\}_{i=1,\dots,d,j=1,\dots,n}$ se calculan los puntos medios (conjunto de la media observada)

$$m_j = \frac{1}{d} \sum_{i=1}^d x_{ij}.$$

2. Se calcula $\hat{a} = \max_{j=1,\dots,n} m_j + \epsilon$ donde $\epsilon > 0$ impide problemas computacionales en el cálculo de logaritmos.
3. Mediante la observación de la gráfica de los m_j se concluye si es necesario o no truncar los datos hasta restringirse únicamente a aquellos que describen un comportamiento sigmoideal, desechando los valores que redundan en el carácter asintótico de la muestra.

- a) Si se deciden truncar datos, habrá de hacerse según el cálculo de $k = \min_{u=1,\dots,n} (u : 2x_u \geq a - x_0)$ y eligiendo

$$\begin{cases} [t_0, t_{2k-1}] & \text{si } 2k \leq n, \\ [t_{2k-n}, t_n] & \text{si } 2k > n. \end{cases}$$

- b) Si no se decide truncar, debe proseguirse con todo el intervalo $[t_0, t_n]$.

4. A partir de los m_j se calcula el conjunto $\{\rho_{\hat{a}}(m_j)\}_j$ donde

$$\rho_{\hat{a}}(m_j) = -\log\left(\frac{\hat{a}}{m_j} - 1\right).$$

5. A partir de la visualización del conjunto $\{\rho_{\hat{a}}(m_j)\}_j$ se sugiere una función paramétrica ϕ_{\star}^{τ} que debe ajustarse a los datos mediante cualquier procedimiento válido, sea por ejemplo un ajuste de regresión no lineal por el método de mínimos cuadrados. Con esto se consigue $\hat{\tau}$ y se completa la función $\phi_{\star}^{\hat{\tau}}$.
6. Ahora, la curva de tipo logístico $L_{\hat{a}1} \phi_{\star}^{\hat{\tau}}(t)$ se ajusta a los datos correspondientes a las medias de las observaciones, $\{m_j\}_j$.

7. Con ella, definimos la función $h_*(t) = \dot{L}_{\hat{a}_1} \phi_*^{\hat{\tau}}(t) L_{\hat{a}_1}^{-1} \phi_*^{\hat{\tau}}(t)$ en todo el intervalo de observación (independientemente de si se han truncado o no los datos).
8. Dado un valor σ apropiado según la variabilidad observada del fenómeno, se establece la ecuación diferencial estocástica

$$dX_t = h_*(t)X_t dt + \sigma X_t dW_t$$

con condición inicial $X_{t_0} = x_0$, donde x_0 es el valor inicial de los datos, $x_0 = m_1$.

9. La solución a tal ecuación es el proceso de difusión logístico generalizado

$$X_t = L_{\hat{a}_1} \phi_*^{\hat{\tau}}(t) e^{-\frac{\sigma^2}{2}(t-t_0) + \sigma(W_t - W_{t_0})},$$

de tipo lognormal con factores exógenos.

10. Entre otras características, se tiene, para todo t ,

$$\mathbb{E}(X_t) = L_{\hat{a}_1} \phi_*^{\hat{\tau}}(t).$$

2.5. Aplicaciones

2.5.1. Modelo con datos sin truncar

Al objeto de aplicar la metodología descrita, simulamos trayectorias de un proceso de difusión con media logística del tipo (1.5) con las siguientes características:

- 200 trayectorias;
- Definido en el intervalo $[0, 50]$ con paso 0.1;
- Presentando distribución inicial degenerada en $x_0 = 0.2$;
- $b = 24$;
- $c = 0.15$;

- $\sigma = 0.03$.

La media de los datos simulados (en este caso, observaciones simuladas) presenta un comportamiento sigmoïdal de tipo logístico. Apliquemos pues la metodología descrita en este capítulo al fin de conseguir un proceso de difusión cuya media se ajuste a la de los datos observados.

Para ello, en primer lugar, definimos la cota superior de esta media, \hat{a} , como

$$\hat{a} = \max_{j=1,\dots,n} m_j + 0.00001 = 4.8442,$$

donde el valor 0.00001 se impone para evitar logaritmos en 0 durante el cálculo de $\rho_{\hat{a}}$, que se realiza acto seguido, según la definición. El conjunto de valores obtenido $\{\rho_{\hat{a}}(m_j)\}_j$ puede observarse representado en azul en la segunda figura superior de 2.2. Como era de esperar, estos valores se distribuirán según un patrón lineal, habida cuenta de la naturaleza de ρ_a como inversa de la función logística, siendo los m_j valores obtenidos según tal función logística.

Aún con todo, la variabilidad creciente hace que los valores fluctúen en la parte final del intervalo de observación, por lo que a la hora de proponer una función que se ajuste a los mismos, no nos limitaremos a una recta, si no a un polinomio de grado mayor o igual que 1. Esto se traduce en sugerir la función

$$\phi_{\star}^{\tau}(t) = \tau_1 + \tau_2 t^{\tau_3} - \tau_2 t_0^{\tau_3} = \tau_1 + \tau_2 t^{\tau_3},$$

para el vector paramétrico $\tau = (\tau_1, \tau_2, \tau_3)$. Nótese ahora que podemos reducir en un parámetro este vector, puesto que hemos de imponer la condición inicial $\phi_{\star}^{\tau}(t_0) = \rho_a(x_0) = \rho_a(m_1)$, que resulta en

$$\hat{\tau}_1 = -\log\left(\frac{\hat{a}}{m_1} - 1\right) = -3.1451.$$

Podemos ahora ajustar la función ϕ_{\star}^{τ} para τ_2 y τ_3 a los datos $\{t_j, \rho_{\hat{a}}(m_j)\}_j$ mediante regresión no lineal por mínimos cuadrados, por ejemplo. Los resultados son $\hat{\tau}_2 = 0.0963$ y $\hat{\tau}_3 = 1.1368$. La función $\phi_{\star}^{\hat{\tau}}$ se ha representado en color rojo en la gráfica superior derecha de la figura 2.2, donde puede observarse el ajuste adecuado a los valores $\rho_{\hat{a}}(m_j)$.

Una vez obtenida la función $\phi_{\star}^{\hat{\tau}}$ y comprobada su adecuación, pasamos a

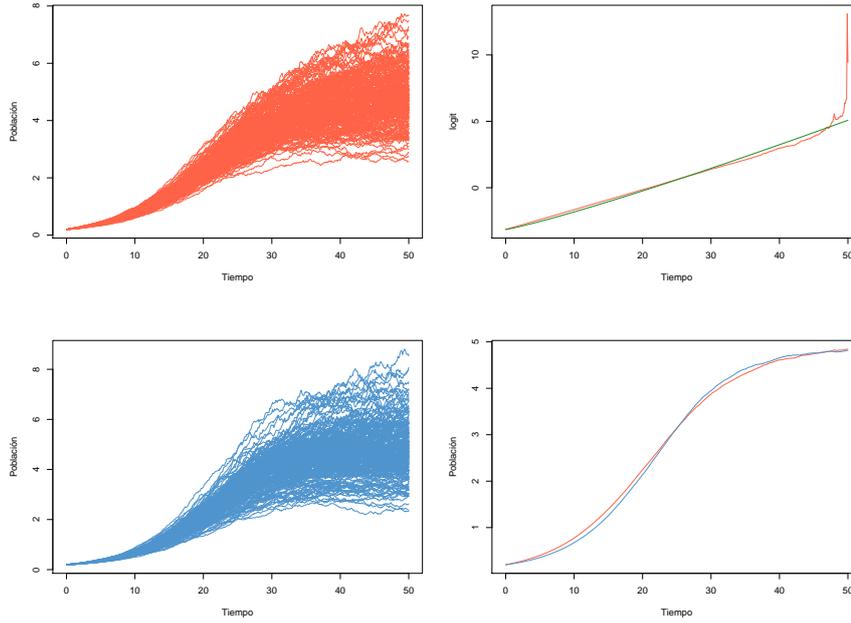


Figura 2.2: De arriba a abajo y de izquierda a derecha: Trayectorias simuladas (rojo); datos modificados mediante *logit* (rojo) y ajuste (verde); trayectorias del nuevo proceso; comparación de medias (original rojo, nuevo azul).

construir el proceso de difusión. Para ello, definimos la función

$$h_*(t) = \dot{\phi}_*^{\hat{\tau}}(t) (1 + \exp \phi_*^{\hat{\tau}}(t))^{-1} = \hat{\tau}_2 \hat{\tau}_3 t^{\hat{\tau}_3 - 1} (1 + e^{\hat{\tau}_1 + \hat{\tau}_2 t^{\hat{\tau}_3}})^{-1}.$$

Con ella, planteamos la ecuación diferencial estocástica

$$dX_t = h_*(t)X_t dt + \sigma X_t dW_t$$

para el mismo σ , sujeta a la condición inicial $X_{t_0} = x_0 = 0.2$. La solución, como se vio en secciones anteriores, es un proceso de difusión lognormal con factores exógenos cuya función media es la curva $L_{\hat{\alpha}_1} \phi_*^{\hat{\tau}}(t)$. Simulando el mismo número de trayectorias de este proceso, en el mismo intervalo de tiempo, obtenemos la gráfica inferior izquierda de la figura 2.2, muy similar a los datos originales (observaciones simuladas), encima en color rojo. De hecho, tomando la media de los valores simulados del nuevo proceso, podemos compararla gráficamente con la propia de los datos originales, resultando en la gráfica inferior derecha de la misma figura.

2.5.2. Modelo con datos truncados

Por otro lado, al objeto de estudiar el desempeño de la metodología propuesta referente al truncado de datos irrelevantes en cuanto a la tendencia de tipo sigmoideal del modelo, simulamos ahora trayectorias bajo las siguientes condiciones:

- 200 trayectorias;
- En el intervalo $[0, 50]$ con paso 0.1;
- Distribución inicial degenerada en $x_0 = 0.2$;
- $b = 24$;
- $c = 1.5$;
- $\sigma = 0.03$.

Es decir, mantenemos las condiciones del ejemplo anterior, pero cambiando el parámetro c por 1.5. Esto hace que la inflexión tenga lugar mucho antes, por lo que el proceso presenta su comportamiento asintótico ya en las primeras etapas de su evolución. Por ello, solo nos interesan los datos referentes a esos instantes iniciales, entendiendo que la fase de evolución asintótica del proceso no nos aporta información relevante, pudiendo, en cambio, influir notablemente en el procedimiento de ajuste a los datos $\rho_{\hat{a}}(m_j)$.

En efecto, el intervalo de observación consta de 501 instantes t_0, \dots, t_{500} . Así pues, siguiendo (2.6) y el método diseñado, obtenemos $k = 21$. Como $2k = 42 < 501$, truncamos el intervalo de observación a $[t_0, t_{41}]$. Aplicando ahora la metodología al conjunto de observaciones medias $\{m_j\}_{j=0, \dots, 41}$ para la misma función paramétrica ϕ_{\star}^{τ} del ejemplo anterior, obtenemos $\hat{\tau}_1 = -3.1067$, $\hat{\tau}_2 = 1.5140$ y $\hat{\tau}_3 = 0.9759$.

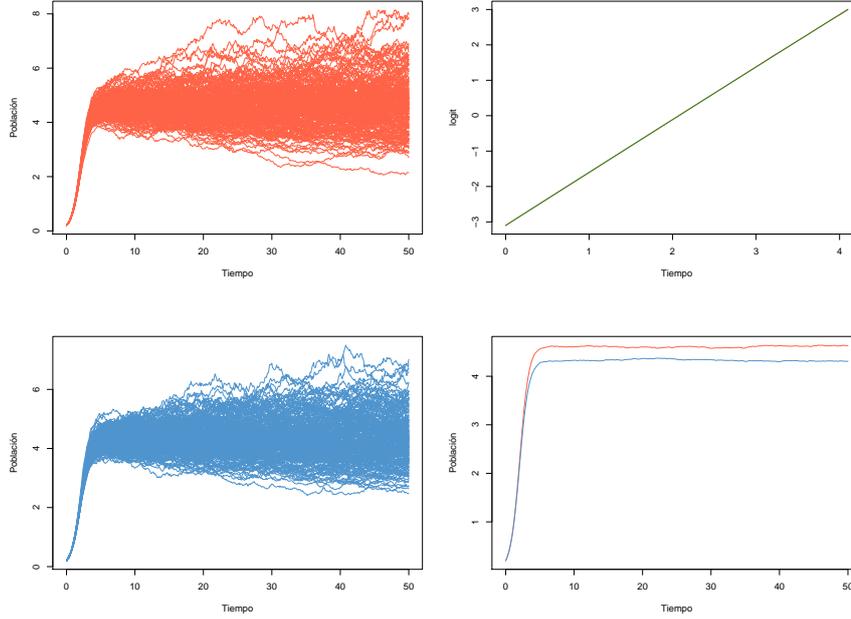


Figura 2.3: De arriba a abajo y de izquierda a derecha: datos originales, media observada vía ρ vs. ajuste, trayectorias del nuevo proceso, comparación de medias.

2.5.3. Equivalencia paramétrica de modelos logísticos

Una aplicación práctica de la metodología descrita consiste en establecer una equivalencia entre los valores de los parámetros de un modelo de tipo logístico y los de otro. Aprovechando la curva logística generalizada definida aquí, podemos considerar dos modelos concretos y aplicarles el procedimiento a cada uno de ellos sobre datos simulados. Estos dos modelos serán:

- Modelo M_1 : $\phi_{*1}^{\tau_1}(t) = \tau_1^1 + \tau_2^1 t^{\tau_3^1} - \tau_2^1 t_0^{\tau_3^1}$;
- Modelo M_2 : $\phi_{*2}^{\tau_2}(t) = \tau_1^2 + \tau_2^2 t + \tau_3^2 \operatorname{arcsenh} t - \tau_2^2 t_0 - \tau_3^2 \operatorname{arcsenh} t_0$.

El modelo M_1 es el usado en los ejemplos anteriores, mientras que el M_2 genera la curva hiperbolástica de tipo I (véanse a este respecto el capítulo 3 y Tabatabai, Williams y Bursac [29]).

De la condición $\phi_{*1}^{\hat{\tau}}(t_0) = \rho_{\hat{a}}(x_0)$ y con $\hat{a} = 5.0001$, se tiene que

$$\hat{\tau}_1^1 = \hat{\tau}_1^2 = -\log\left(\frac{\hat{a}}{m_1} - 1\right) = -3.1780.$$

Tras esto, ajustamos la función de cada uno de los dos modelos a los datos $\rho_{\hat{a}}(m_j)$ para $j = 1, \dots, 501$, obteniendo las estimaciones de la siguiente tabla:

	τ_1	τ_2	τ_3
M_1	-3.1780	0.0695	1.2380
M_2	-3.1780	0.1895	-0.2134

Tabla 2.1: Estimaciones para dos modelos basados en la curva logística.

En la figura 2.4 se observan los datos simulados (en el centro en gris), así como las funciones de cada modelo ajustadas a los datos transformados por ρ (gráfica superior). Una vez hecho esto, se pueden simular dos difusiones logísticas generalizadas, una para cada modelo, presentando sus trayectorias en las gráficas central y derecha de la franja central de la figura (azul y violeta). Finalmente, en la gráfica inferior, se comparan las medias de estos dos procesos con la observada (datos simulados).

En resumen, con esta aplicación es posible definir los rangos paramétricos respectivos de distintos modelos de base logística para que se comporten de igual manera.

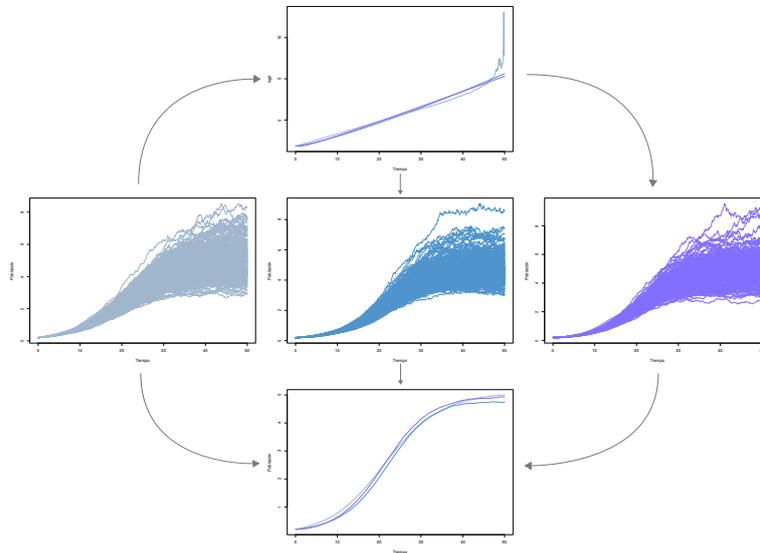


Figura 2.4: Ajuste de dos modelos de tipo logístico sobre datos simulados. Franja central, en gris, datos simulados. Siguiendo la flecha, se llega a la gráfica superior: logit y funciones ajustadas (azul polinomio, violeta hiperbolástico I). Con ellas se generan dos difusiones respectivas, azul y violeta en el centro. Abajo, comparativa de medias observada y simuladas.

2.6. Extensión a otros modelos

La metodología propuesta surge de la necesidad de obtener un proceso de difusión con media de tipo logístico, que describa el comportamiento de las observaciones de un fenómeno dinámico de crecimiento. Para superar las limitaciones de la curva logística clásica, consideramos una generalización funcional de la misma, de manera que, bajo determinadas condiciones, puedan obtenerse curvas con características específicas pero basando siempre su comportamiento en el modelo sigmoideal de la logística.

La metodología propuesta puede, sin embargo, extenderse a otro tipo de curvas susceptibles de ser generalizadas mediante el añadido de una función genérica en las ecuaciones diferenciales que las definen. Tal es el caso de la curva Weibull, cuya generalización obtenemos como solución de

$$\frac{dx(t)}{dt} = (\alpha - x(t))g(t)$$

para $\alpha > 0$ y una función continua g tal que su primitiva $g(t) \rightarrow \infty$ cuando $t \rightarrow \infty$. Es decir,

$$W_{ab}(t) = a(1 - be^{-g(t)})$$

para parámetros $a, b \in \mathbb{R}$. Definiendo la fertilidad malthusiana

$$h(t) = \dot{W}_{ab}(t) W_{ab}^{-1}(t),$$

la solución de la ecuación diferencial estocástica (1.5) es un proceso de difusión lognormal con factores exógenos cuya función media es exactamente la curva de Weibull generalizada W_{ab} .

La aplicación práctica de este modelo se enfrenta a los mismos inconvenientes que el estudiado para la curva logística. En efecto, despejando g se llega a

$$g(t) = \log b - \log \left(1 - \frac{W_{ab}(t)}{a} \right),$$

por lo que, para obtener una realización de la función y así ajustar una candidata, habría que prescindir del término $\log b$. Con esto, tomando los valores medios observados m_j como puntos de la curva, se debería ajustar una función

al conjunto de puntos

$$\left\{ \left(t_j, -\log \left(1 - \frac{m_j}{a} \right) \right) \right\}_j.$$

El procedimiento restante es análogo al estudiado para la curva logística generalizada.

Capítulo 3

Proceso de difusión hiperbolástico de tipo I

3.1. Introducción	69
3.2. El modelo	70
3.3. Estimación del modelo	74
3.4. Aplicaciones	81

3.1. Introducción

La construcción de modelos matemáticos como los descritos en este trabajo puede enfrentarse a distintos problemas asociados, entre los que destacan los relacionados con la naturaleza del fenómeno cuyo comportamiento pretenden describir. Uno de ellos es la restricción que provocan los puntos de inflexión fijos. Esta propiedad, presente en numerosas curvas de tipo sigmoideal, resta flexibilidad a las mismas, disminuyendo su capacidad de adaptación a los datos.

Un ejemplo de curvas que evitan este problema es el de las curvas hiperbolásticas de tipo I, II y III (H_1, H_2 y H_3 respectivamente), habiendo sido la primera de ellas introducida ya en este trabajo a modo de ejemplo. Desarrolladas por Tabatabai, Williams y Bursac [29], estas curvas presentan, como se ha comentado, inflexiones móviles que las dotan de mayor capacidad de ajuste a los datos.

En este sentido, pueden consultarse los trabajos de Eby, Tabatabai y Bursac [6] sobre crecimiento del carcinoma sólido de Ehrlich bajo determinados tratamientos, y en el cual obtiene descripciones más precisas con la curva H_3 que con otras clásicas como Gompertz o Weibull. De la misma forma, se emplean curvas hiperbolásticas para describir el comportamiento de células madre embrionarias (en Tabatabai et al. [27]) o para generar modelos más sofisticados si cabe como el oscilobolástico (véase Tabatabai, Eby y Bursac [26]) o el modelo T, en Tabatabai et al. [28].

Siguiendo la metodología de este trabajo, introduciremos a continuación, como caso particular para una curva logística generalizada, el proceso de difusión hiperbolástico de tipo I (basado pues en la logística).

Hecho esto, se desarrollarán procedimientos de inferencia para el proceso, lo que llevará a problemas relacionados con la imposibilidad de resolver, analítica y numéricamente, las ecuaciones de verosimilitud. Esto nos llevará a considerar el uso de algoritmos metaheurísticos para resolver el problema, siendo el algoritmo *Firefly* el empleado en este caso.

Introducido por Yang [34], este algoritmo se encuentra bajo un importante desarrollo, apareciendo en aplicaciones de ingeniería u optimización (véanse los trabajos de Alb et al. [1] o Kavousi-Fard, Samet y Marzbani [16]), además de ser objeto de modificaciones como las de Gandomi et al. [12] o Zhang et al. [36].

En este capítulo introducimos pues la difusión hiperbolástica de tipo I, y tratamos su procedimiento de inferencia mediante el algoritmo *Firefly*, realizando finalmente un estudio de simulación, y aplicando la metodología a un caso real procedente de la biología molecular.

3.2. El modelo

3.2.1. Reformulación de la curva H_1

La curva H_1 es solución de la ecuación diferencial de Bernoulli

$$\frac{dx(t)}{dt} = M^{-1}x(t)(M - x(t))\left(\rho + \frac{\theta}{\sqrt{1+t^2}}\right), \quad (3.1)$$

para $t \geq t_0$, $\theta \in \mathbb{R}$, y condición inicial $X(t_0) = x_0 > 0$. Aquí, $x(t)$ representa el tamaño de la población en el instante t y M su capacidad máxima sostenible (*carrying capacity*), mientras que los parámetros θ y ρ determinan, en conjunto, la ratio de crecimiento. Nótese que cuando $\theta = 0$ la expresión resultante es la conocida ecuación diferencial logística. Por otro lado, puede observarse que la ecuación (3.1) es de la forma (2.1) para $\alpha = 1$, $\beta = M^{-1}$ y

$$g(t) = \rho + \frac{\theta}{\sqrt{1+t^2}}.$$

Es decir, podemos considerar aquí los desarrollos del capítulo 2. En efecto, la solución de (3.1) es

$$x(t) = M \left(1 + a \exp(-\rho t - \theta \operatorname{arcsenh}(t)) \right)^{-1}, \quad (3.2)$$

donde

$$a = x_0^{-1}(M - x_0) \exp(\rho t_0 + \theta \operatorname{arcsenh}(t_0)),$$

esto es, $L_{M\alpha} g(t)$, un caso particular de curva logística generalizada para

$$g(t) = \rho t + \theta \operatorname{arcsenh} t,$$

donde $g(t) = \int_0^t g(u) du$.

Nótese que $\lim_{t \rightarrow \infty} x(t) = M$ si $\rho > 0$, y $\lim_{t \rightarrow \infty} x(t) = 0$ si $\rho < 0$. Luego, si $\rho > 0$, la curva crece hasta la asíntota M , mientras que $\rho < 0$ conduce a una evolución decreciente.

Hemos considerado que la asíntota M es independiente del valor inicial, algo que puede suponer una importante restricción para algunas aplicaciones. En la práctica, se dan situaciones en las cuales el fenómeno de crecimiento a estudio puede mostrar un crecimiento de tipo sigmoideal (susceptible de ser descrito por una curva H_1) para diferentes trayectorias iniciadas en distintos puntos (por ejemplo, el peso de los individuos en una población). Por esta razón, y al objeto de modelar situaciones en las cuales el valor límite depende de los valores iniciales, consideraremos una reformulación de la curva (3.2).

En efecto, dados $\eta = 1/a$, $\lambda = e^{-\rho}$, y $\mu = e^{-\theta}$, obtenemos

$$x(t) = x_0 \frac{\eta + \xi(t_0)}{\eta + \xi(t)}, \quad (3.3)$$

para $t \geq t_0$, $\eta, \mu > 0$, y donde $\xi(t) = \lambda^t \mu^{\operatorname{arcsenh}(t)}$. En lo que sigue, trataremos con curvas crecientes. Este hecho conduce a establecer las condiciones $0 < \lambda < 1$ y $\mu < c_\lambda(t)$, $\forall t \geq t_0$, donde $c_\lambda(t) = \lambda^{-\sqrt{1+t^2}}$ es una función creciente para todo t . Así, vamos a considerar la condición $\mu < c_\lambda(t_0)$.

Otra importante característica de la curva H_1 es la movilidad de sus inflexiones, lo cual la dota de una flexibilidad superior a la de otras curvas comunes en este campo de estudio. Ello permite un modelado más preciso de los fenómenos que exhiban un comportamiento sigmoideal.

En particular, de $\ddot{x}(t) = 0$ se sigue que las inflexiones de la curva verifican

$$\frac{2\eta}{\eta + \xi(t)} - 1 = \frac{t \log \mu}{(1 + t^2)^{3/2}} \left(\log \lambda + \frac{\log \mu}{\sqrt{1 + t^2}} \right)^{-2}.$$

3.2.2. El proceso de difusión H_1

En esta sección, vamos a introducir el proceso de difusión asociado a la curva (3.3). Para ello, siguiendo la metodología propuesta inicialmente en la sección 1.4 y calculando $h(t)$ según el enfoque de la curva logística generalizada (sección 2.3) en el caso (3.2.1), obtenemos

$$\begin{aligned} h(t) &= ag(t) \left(e^{g(t)} + a \right)^{-1} \\ &= a \left(\rho + \theta (1 + t^2)^{-\frac{1}{2}} \right) \left(e^{\rho t + \theta \operatorname{arcsenh} t} + a \right)^{-1} \\ &= -\dot{\xi}(t) (\eta + \xi(t))^{-1}, \end{aligned}$$

donde la última expresión se establece según la reformulación propuesta (3.3). Con esta función, planteamos, según (1.5), la ecuación diferencial estocástica

$$dX_t = -\dot{\xi}(t) (\eta + \xi(t))^{-1} X_t dt + \sigma X_t dW_t$$

para $\sigma > 0$.

Teniendo en cuenta que h es continua y acotada, esta ecuación verifica las

condiciones para la existencia y unicidad de solución, que será un proceso de difusión continuo, tomando valores en \mathbb{R}^+ y caracterizado por los momentos infinitesimales (*drift* y coeficiente de difusión) $A_1(x, t) = -\dot{\xi}(t)(\eta + \xi(t))^{-1} x$ y $A_2(x) = \sigma^2 x^2$.

Además, es posible encontrar la forma explícita de la solución, que resulta ser

$$X_t = x_0 \frac{\eta + \xi(t_0)}{\eta + \xi(t)} \exp\left(-\frac{\sigma^2}{2}(t - t_0) + \sigma(W_t - W_{t_0})\right), \quad (3.4)$$

considerando la condición inicial $X_{t_0} = x_0$, independiente de W_t para $t \geq t_0$.

Al ser un caso particular para la curva logística generalizada, podemos afirmar que este proceso es un proceso de difusión lognormal con factores exógenos, ampliamente estudiado y ya citado en este trabajo.

Siguiendo ahora la sección 2.3.1, podemos obtener fácilmente las distribuciones finito-dimensionales del proceso, así como su función de densidad de transición de probabilidad.

En efecto, tanto si la distribución inicial es lognormal $\Lambda_1(\mu_0, \sigma_0^2)$ como si es degenerada en un punto x_0 , todas las distribuciones finito-dimensionales serán normales.

Considerando el caso inicial lognormal, tendremos que para todo $n \in \mathbb{N}$ y $t_1 < \dots < t_n$, el vector $(X_{t_1}, \dots, X_{t_n})'$ sigue una distribución lognormal $\Lambda_n(\zeta, \Sigma)$, donde las componentes del vector ζ y la matriz Σ son, respectivamente,

$$\zeta_i = \mu_0 + \log \frac{\eta + \xi(t_0)}{\eta + \xi(t_i)} - \frac{\sigma^2}{2}(t_i - t_0)$$

y

$$\sigma_{ij} = \sigma_0^2 + \sigma^2((t_i \wedge t_j) - t_0),$$

para $i, j = 1, \dots, n$.

Por otro lado, la función de densidad de transición se obtiene de la distribución bidimensional para $s < t$, siendo

$$f(x, t|y, s) = \frac{1}{x \sqrt{2\pi\sigma^2(t-s)}} \exp\left(-\frac{\left(\log \frac{x}{y} - \log \frac{\eta + \xi(s)}{\eta + \xi(t)} + \frac{\sigma^2}{2}(t-s)\right)^2}{2\sigma^2(t-s)}\right)$$

esto es, la distribución condicionada es lognormal,

$$(X_t | X_s = y) \rightsquigarrow \Lambda_1 \left(\log y + \log \frac{\eta + \xi(s)}{\eta + \xi(t)} - \frac{\sigma^2}{2}(t-s); \sigma^2(t-s) \right).$$

De aquí era posible obtener los momentos (condicionados o no) no centrados de orden $k \geq 0$.

En este caso particular, y a efectos de simplificar notación, podemos definir la función

$$G_k(t|z, \tau) = z^k \left(\frac{\eta + \xi(\tau)}{\eta + \xi(t)} \right)^k \exp \left(\frac{k(k-1)\sigma^2}{2}(t-\tau) \right),$$

con lo cual,

$$\begin{aligned} \mathbb{E}(X_t^k) &= G_k(t | \mathbb{E}(X_{t_0}), t_0), \\ \mathbb{E}(X_t^k | x_\tau = y) &= G_k(t | y, \tau). \end{aligned}$$

En particular, la función media del proceso, así como la media condicionada a un valor inicial x_0 vienen dadas por

$$\begin{aligned} \mathbb{E}(X_t) &= \mathbb{E}(X_{t_0}) \frac{\eta + \xi(t_0)}{\eta + \xi(t)}, \\ \mathbb{E}(X_t | X_{t_0} = x_0) &= x_0 \frac{\eta + \xi(t_0)}{\eta + \xi(t)}, \end{aligned}$$

respectivamente. Nótese que ambas son curvas H_1 como se describieron previamente en (3.3). Esto justifica el uso del modelo propuesto, sobre todo con fines predictivos.

3.3. Estimación del modelo

Consideremos una muestra discreta del proceso, basada en d trayectorias, para instantes t_{ij} , ($i = 1, \dots, d$, $j = 1, \dots, n_i$) con $t_{i1} = t_1$, $i = 1, \dots, d$. Esto es, observamos las variables $x_{t_{ij}}$, cuyos valores $\mathbf{x} = \{x_{ij}\}_{i=1, \dots, d; j=1, \dots, n_i}$, hacen a la muestra apropiada para llevar a cabo un estudio inferencial.

Considerando el caso más general, en el cual la distribución inicial es

lognormal, $X_{t_1} \rightsquigarrow \Lambda(\mu_1, \sigma_1^2)$, el logaritmo de la función de verosimilitud L_x es

$$\log L_x(\mu_1, \sigma_1^2, \eta, \lambda, \mu, \sigma^2) = \sum_{i=1}^d \log f_1(x_{i1}) + \sum_{i=1}^d \sum_{j=2}^{n_i} \log f(x_{ij}, t_{ij} | x_{i,j-1}, t_{i,j-1})$$

donde f_1 es la función de densidad de la variable X_{t_1} . Ahora, notando $N = \sum_{i=1}^d n_i$, tenemos

$$\begin{aligned} \log L_x(\mu_1, \sigma_1^2, \eta, \lambda, \mu, \sigma^2) &= -\frac{N}{2} \log(2\pi) - \frac{d}{2} \log \sigma_1^2 - \frac{N-d}{2} \log \sigma^2 - \sum_{i=1}^d \log x_{i1} \\ &\quad - \frac{1}{2\sigma_1^2} \sum_{i=1}^d (\log x_{i1} - \mu_1)^2 - \sum_{i=1}^d \sum_{j=2}^{n_i} \log x_{ij} - \frac{1}{2} \sum_{i=1}^d \sum_{j=2}^{n_i} \log \Delta_{ij} \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{1}{\Delta_{ij}} \left(R_{ij} - T_{ij}^{\eta, \lambda, \mu} + \frac{\sigma^2}{2} \Delta_{ij} \right)^2 \quad (3.5) \end{aligned}$$

donde

$$T_{ij}^{\eta, \lambda, \mu} = \log \left(\frac{\eta + \xi(t_{i,j-1})}{\eta + \xi(t_{ij})} \right),$$

$$R_{ij} = \log \frac{x_{ij}}{x_{i,j-1}},$$

$$\Delta_{ij} = t_{ij} - t_{i,j-1}.$$

De (3.5), las estimaciones máximo-verosímiles de μ_1 y σ_1^2 son

$$\hat{\mu}_1 = \frac{1}{d} \sum_{i=1}^d \log x_{i1} \quad \text{y} \quad \hat{\sigma}_1^2 = \frac{1}{d} \sum_{i=1}^d (\log x_{i1} - \hat{\mu}_1)^2.$$

Sin embargo, estimar el resto de parámetros conlleva algunas dificultades. Concretamente, el sistema de ecuaciones resultante (ver apéndice A) es extremadamente complejo y no presenta una solución explícita. Por lo tanto, deben emplearse procedimientos numéricos para encontrar una solución aproximada. Además, resultaría imposible llevar a cabo un estudio general del sistema de ecuaciones al objeto de comprobar las condiciones de convergencia para el método numérico que se propusiera, ya que este sistema depende de la mues-

tra de datos elegida, lo que puede llevar a comportamientos impredecibles. Finalmente, a todo esto habría que añadir la necesidad de elegir una solución inicial.

Por todas estas razones, proponemos el uso del algoritmo metaheurístico de búsqueda local *Firefly*, con la finalidad de maximizar la función de verosimilitud. El estudio y desarrollo de los algoritmos de optimización metaheurísticos es un campo de intensa actividad investigadora, por lo que existen numerosos ámbitos de estudio en los que se aplican, incluyendo la estimación en procesos de difusión. Por ejemplo, Román-Román et al. [20] usan el algoritmo *Simulated Annealing* para estimar los parámetros de un proceso de difusión de tipo Gompertz (su tendencia sigue esa curva), mientras que Román-Román y Torres-Ruiz [21] sugieren el algoritmo *Variable Neighborhood Search* en combinación con *Simulated Annealing* para el estudio de una difusión con tendencia de tipo Richards.

Una vez hemos obtenido las estimaciones $\hat{\mu}_1$ y $\hat{\sigma}_1^2$, el objetivo principal consiste en maximizar la función $\log L_x(\hat{\mu}_1, \hat{\sigma}_1^2, \eta, \lambda, \mu, \sigma^2)$. Así, de (3.5), consideraremos la función objetivo

$$f_o(\lambda, \mu, \eta, \sigma) = -\frac{d(N-1)}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{1}{\Delta_{ij}} \left(R_{ij} - T_{ij}^{\eta, \lambda, \mu} + \frac{\sigma^2}{2} \Delta_{ij} \right)^2.$$

3.3.1. Aplicación del algoritmo *Firefly*

El algoritmo *Firefly* es un algoritmo estocástico metaheurístico, inspirado en la naturaleza y basado en poblaciones, que ha sido desarrollado por Yang [34]. Ha sido aplicado con éxito en numerosos campos de investigación como la optimización o la ingeniería (véanse los trabajos de Kavousi-Fard, Samet y Marzbani [16] y Niknam, Azizipanah-Abarghooee y Roosta [18]) debido a su eficiencia y la capacidad para tratar con problemas NP-complejos, llegando a convertirse en una de las herramientas más importantes en inteligencia de enjambres (*swarm intelligence*), un campo de investigación propio de la inteligencia artificial, enfocado en el comportamiento colectivo de sistemas descentralizados y auto-organizados. De hecho, la investigación sobre el algoritmo *Firefly* y sus aplicaciones se encuentra en un estado de importante desarrollo. Como ejemplo, pueden consultarse los recientes trabajos de Fister

et al. [10] y Yang y He [35], así como las sofisticadas modificaciones realizadas por Fister et al. [11] y Gandomi et al. [12].

El algoritmo trabaja sobre subgrupos de la población, consiguiendo así tratar, de manera eficiente, problemas de optimización no lineal y multimodal. De hecho, el algoritmo *Firefly* puede verse como una generalización de algoritmos más conocidos como los *Particle Swarm Optimization*, *Differential Evolution* o *Simulated Annealing*.

La inspiración para desarrollar este algoritmo procede de los destellos de luz que emiten las luciérnagas, producidos por un fenómeno de bioluminiscencia y usados por los insectos para atraerse entre ellos. Por lo tanto, en términos elementales, podemos considerar n luciérnagas distribuidas aleatoriamente en el espacio de búsqueda, cada una de ellas con una intensidad de luz asociada que resulta ser dependiente de su posición. Así, en cada generación, las luciérnagas son atraídas (el atractivo se considera directamente proporcional a la intensidad de luz) por las más brillantes de su entorno en función de la distancia entre ellas. Es decir, si una luciérnaga brilla con una alta intensidad, las más cercanas serán atraídas y volarán hacia ella (su posición).

Con el movimiento, la atracción entre luciérnagas se ve alterada debido a las nuevas distancias e intensidades lumínicas de las mismas. De esta manera, al finalizar cada generación, se ordenan las luciérnagas de la menos brillante a la más brillante. Todo el procedimiento se repite durante un número fijo de generaciones, en cada una de las cuales se reduce la aleatoriedad de los movimientos de las luciérnagas (un vuelo menos errático al haber mayor seguridad acerca del brillo de aquella que causa la atracción).

Analizando esto con más detalle, es obvio que la intensidad lumínica, y por ende el atractivo, varían de acuerdo a la distancia entre luciérnagas, así como el medio en el que estas se mueven y las propiedades que presenta para absorber la luz (por ejemplo, durante un día de niebla, sería más difícil percibir la luz emitida por una luciérnaga vecina). El algoritmo *Firefly* usa estas características al objeto de buscar los máximos locales y globales de una función dada, definiendo el espacio de búsqueda como el dominio de la función objetivo y la intensidad lumínica proporcional al valor de la función en ese punto (ocupado por la luciérnaga). Por lo tanto, las medidas de atracción entre luciérnagas debe decrecer cuando la distancia y/o el coeficiente de absorción

(del medio) aumentan. Normalmente, esto se formula como

$$\beta(r) = \beta_0 \exp(-\gamma r^2),$$

donde β_0 es el atractivo en el instante $t = 0$ y γ es el coeficiente de absorción del medio.

Por lo tanto, si una luciérnaga i detecta la mayor intensidad lumínica de la luciérnaga j a una distancia r_{ij} , entonces se dará un movimiento de i hacia j según la expresión¹

$$x_k^i \rightarrow x_k^i + \beta(r_{ij})(x_k^j - x_k^i) + \alpha \epsilon_k, \quad (3.6)$$

donde x_k^i y x_k^j son las k -ésimas componentes de las posiciones de la i -ésima y j -ésima luciérnagas, respectivamente, ϵ_k es un valor aleatorio de la forma $u - \frac{1}{2}$ para u proveniente de una distribución uniforme (también puede sustituirse por una distribución normal o de otro tipo), y α es el parámetro de aleatorización, el cual controla la influencia estocástica en el movimiento, viéndose reducido de manera recursiva en cada generación por un factor $\delta \in (0, 1)$, de la forma $\alpha \rightarrow \delta \alpha$.

Debe tenerse en cuenta que el algoritmo *Firefly* está esencialmente caracterizado por tres parámetros, a saber: el parámetro de aleatorización α , el atractivo inicial β_0 y el coeficiente de absorción lumínica γ . Esto puede verse como un comportamiento asintótico; en efecto, por ejemplo, si $\gamma \rightarrow 0$, esto resulta en un atractivo constante $\beta = \beta_0$, que lleva a un caso especial del algoritmo *Particle Swarm Optimization*. Y cuando $\gamma \rightarrow \infty$, el algoritmo *Firefly* se convierte en un camino aleatorio, una versión paralela al algoritmo *Simulated Annealing*.

¹Se emplea aquí notación recursiva.

Algoritmo 1 Pseudocódigo del algoritmo *Firefly*

```

Definir función objetivo  $f(\mathbf{x})$  donde  $\mathbf{x} = (x_1, \dots, x_d)'$ 
Asignar valores para  $\gamma$ ,  $\beta_0$ ,  $\alpha$ ,  $\delta$  y MáxGeneración (número máximo de
generaciones)
Generar la población inicial de luciérnagas  $\mathbf{x}_i$  para  $i = 1, 2, \dots, n$ 
Determinar la intensidad lumínica  $I_i$  en  $\mathbf{x}_i$  según  $f(\mathbf{x}_i)$ 
while  $t < \text{MáxGeneración}$  do
  for  $i = 1 : n$  todas las  $n$  luciérnagas do
    for  $j = 1 : i$  todas las  $n$  luciérnagas do
      if  $I_i < I_j$  then
        Mover luciérnaga  $i$  hacia  $j$  de acuerdo a (3.6)
      end if
      Variar el atractivo con la distancia  $r$  a través de  $\exp(-\gamma r^2)$ 
      Actualizar la intensidad lumínica evaluando las nuevas soluciones
    end for
  end for
  Ordenar las luciérnagas (desde la de menor hasta la de mayor valor de
  la función objetivo) y encontrar la mejor (i.e., la última)
  Actualizar  $\alpha$  aplicando el factor de reducción  $\delta$ 
end while
Procesar los resultados y visualizar

```

Parámetros iniciales y acotación del espacio paramétrico

En lo que respecta a los parámetros iniciales del algoritmo, y siguiendo los comentarios de Yang [34], se pueden realizar las siguientes consideraciones para cada actor del algoritmo:

- α : El parámetro de aleatorización controla el carácter estocástico del movimiento de las luciérnagas. Normalmente toma valores entre 0 y 1, siendo 0.2 un valor recomendado.
- δ : Este parámetro reduce α en cada generación. Suele tomarse un valor entre 0.95 y 0.99.
- β_0 : Dado que el atractivo a distancia $r = 0$ es habitualmente considerado de valor 1, $\beta_0 = 1$ se tomará en la mayoría de casos.
- γ : El coeficiente de absorción del medio es un parámetro crítico que influye de manera notable en la velocidad de convergencia del algoritmo.

Simula las condiciones del entorno en el cual las luciérnagas pueden detectar la luz. En la mayoría de los casos, los valores apropiados van de 1 a 10.

- n : El número de luciérnagas puede variar en un amplio rango dependiendo, por ejemplo, del número esperado de máximos locales. Por supuesto, un valor alto de n asegura un mayor alcance en el espacio paramétrico.
- Generaciones: Normalmente un valor entre 50 y 100 es suficiente, aunque pueden obtenerse buenos resultados incluso con valores por debajo de 50.

En cualquier caso, el alto grado de eficiencia del algoritmo *Firefly* nos permite aplicarlo con distintos valores iniciales sin incurrir en costes computacionales excesivos.

Por otro lado, para determinar la población inicial de luciérnagas, deben seleccionarse, de manera aleatoria, n puntos en el espacio paramétrico. Para facilitar este proceso, nos proponemos acotar este espacio. Y a tal fin, emplearemos la reformulación de la curva establecida en la sección 3.2.1, así como la información proporcionada por las trayectorias muestrales. De esta manera, para cada parámetro tenemos lo siguiente:

- λ : Está entre 0 y 1, al objeto de garantizar el crecimiento estricto de la curvas consideradas.
- μ : Verifica $0 < \mu < \lambda^{-\sqrt{1+t_0^2}}$ para λ previamente establecido.
- η : Este parámetro puede ser acotado teniendo en cuenta la asíntota de la curva, que verifica $k = x_0 \left(1 + \frac{\xi(t_0)}{\eta}\right)$. De ahí, y si denotamos como k_i al valor máximo de la i -ésima trayectoria, se verifica la expresión

$$\xi(t_0) \left[\max_{i=1, \dots, d} \left(\frac{k_i}{x_{i0}} - 1 \right) \right]^{-1} < \hat{\eta} < \xi(t_0) \left[\min_{i=1, \dots, d} \left(\frac{k_i}{x_{i0}} - 1 \right) \right]^{-1}, \quad (3.7)$$

donde x_{i0} es el valor inicial de la i -ésima trayectoria. En la práctica, cuando $t_0 \neq 0$, el algoritmo debe elegir valores $\tilde{\lambda}$ y $\tilde{\mu}$ para construir finalmente un intervalo para η . De otro modo, en el caso $t_0 = 0$, el intervalo no dependerá de λ ni μ .

- σ : Con valores altos, las trayectorias presentan una mayor variabilidad alrededor de la media del proceso, pudiendo resultar en un modelo de tipo H_1 inutilizable. Algunas simulaciones llevadas a cabo para distintos valores de σ sugieren que valores entre 0 y 0.5 con aceptables, lo que puede derivar en trayectorias con un crecimiento de tipo H_1 .

3.4. Aplicaciones

3.4.1. Aplicación a datos simulados

En esta sección realizaremos algunos estudios mediante simulaciones. Esto ayudará a validar los procedimientos descritos anteriormente, en tanto que se aplican para estimar los parámetros del proceso simulado.

El patrón general de simulaciones está basado en la generación de 30 trayectorias del proceso de difusión H_1 en el intervalo $[0, 50]$ de (3.4). Todas las trayectorias tienen la misma longitud, y $t_i = (i - 1) \cdot 0.1$, $i = 1, \dots, 501$ son los instantes en los cuales se han realizado las observaciones. En todos los ejemplos, este patrón se repetirá 50 veces.

Un primer ejemplo de simulación se desarrolla tomando una distribución inicial degenerada en el punto $x_0 = 0.1$ y parámetros $\eta = 0.5$, $\lambda = 0.8$, $\mu = 0.8$ y $\sigma = 0.015$. La figura 3.1 muestra las trayectorias simuladas.

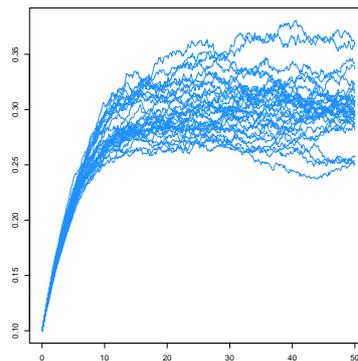


Figura 3.1: Trayectorias simuladas del proceso de difusión de tipo H_1 .

Hemos considerado $n = 40$ luciérnagas, esto es, 40 4-uplas del espacio paramétrico, seleccionando cada una como sigue: primero, para λ , elegimos

un valor aleatorio (distribuido de manera uniforme) entre 0 y 1, y después, para μ , un valor aleatorio en el intervalo $(0, \frac{1}{\lambda})$ dado que $t_0 = 0$ en este caso. La 4-upla se completa seleccionando un valor aleatorio para η en el intervalo $(0.3562706, 0.5612756)$ (en virtud de (3.7)) y otro para σ en el intervalo $(0, 0.5)$.

Por otro lado, con respecto a los parámetros del algoritmo *Firefly*, consideramos los siguientes valores:

- 80 generaciones.
- $\alpha = 0.2$.
- $\delta = 0.97$.
- $\beta_0 = 1$.
- $\gamma = 1$.

	λ	μ	η	σ	f_o
Valor real	0.8	0.8	0.5	0.015	55554.17
Valor estimado	0.7873909	0.8184537	0.5001565	0.0149564	55554.58
Error relativo abs.	0.0157614	0.0230670	0.0003129	0.0029034	0.0000073

Tabla 3.1: Resultados del algoritmo *Firefly* sobre datos simulados.

La tabla 3.1 contiene los valores reales y estimados de los parámetros, así como el valor de la función objetivo f_o en el punto resultante. Como medida del error de estimación, hemos incluido el error relativo absoluto, es decir, la diferencia en valor absoluto entre el valor real y el estimado, dividida por el valor real. De aquí ya pueden observarse buenos resultados en términos de valores estimados y errores relativos absolutos.

El comportamiento del algoritmo a lo largo de las generaciones se muestra con más detalles en las figuras 3.2 y 3.3. Para cada parámetro del proceso, esto es, cada coordenada de las luciérnagas, cada línea en la figura 3.2 representa una generación (de colores grises claro a negro oscuro) mostrando los valores estimados del parámetro para cada luciérnaga. La figura 3.3 muestra, para cada parámetro del proceso, la evolución de sus valores estimados para

cada luciérnaga durante las últimas 60 generaciones. Las líneas verde y roja representan la evolución de la estimación según la mejor y peor luciérnaga, respectivamente.

Debido a que el espacio paramétrico es tetradimensional, resulta imposible visualizar la evolución de las trayectorias de las luciérnagas. Por esta razón, en la figura 3.4 hemos incluido las gráficas de la dinámica de la mejor luciérnaga en cada generación, para todos los pares de parámetros posibles.

Por otro lado, la figura 3.5a muestra la evolución de la función objetivo f_o para la mejor luciérnaga en cada generación. Obviamente esta gráfica es creciente, debido a que la mejor luciérnaga de cada generación mejora el valor de la función objetivo (en el sentido de máximos o mínimos locales) respecto a la generación anterior. Finalmente, la figura 3.5b muestra la tendencia de los datos y la función media del proceso de difusión H_1 .

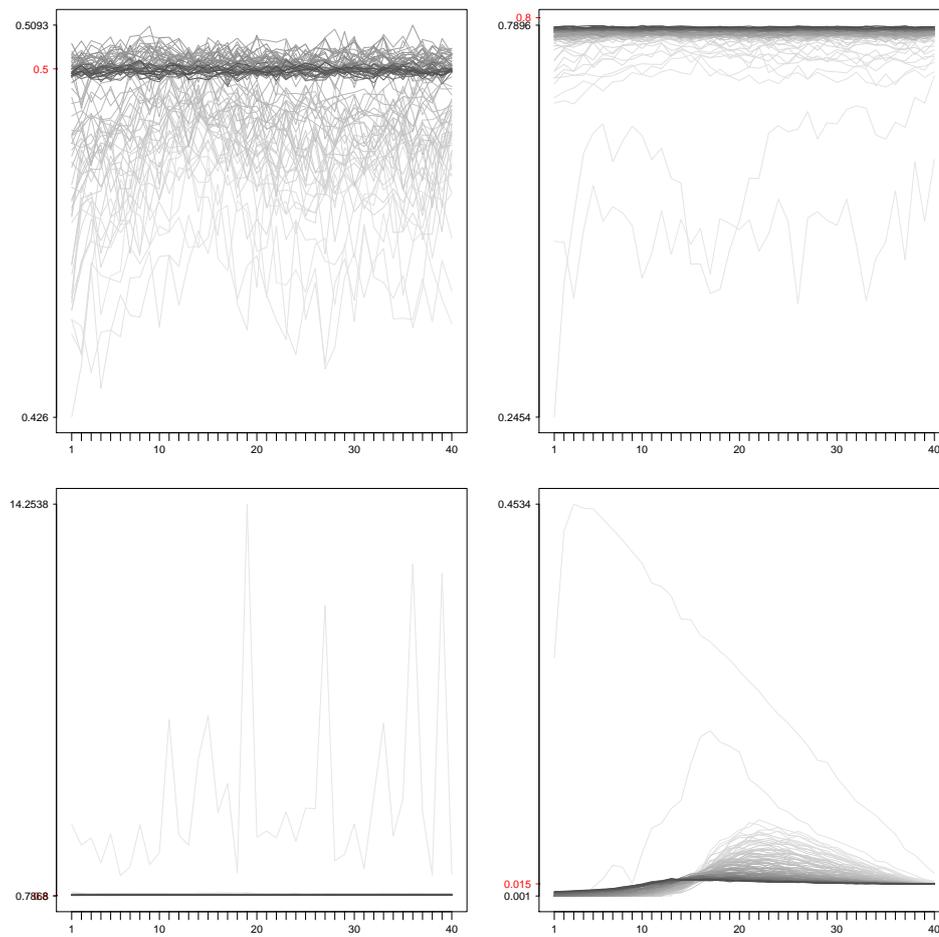


Figura 3.2: Primer ejemplo de simulación: evolución (desde gris claro a negro) de 80 generaciones de 40 luciérnagas para $\eta = 0.5$, $\lambda = 0.8$, $\mu = 0.8$ y $\sigma = 0.015$ (de izquierda a derecha, y de arriba a abajo). Valores originales, en rojo.

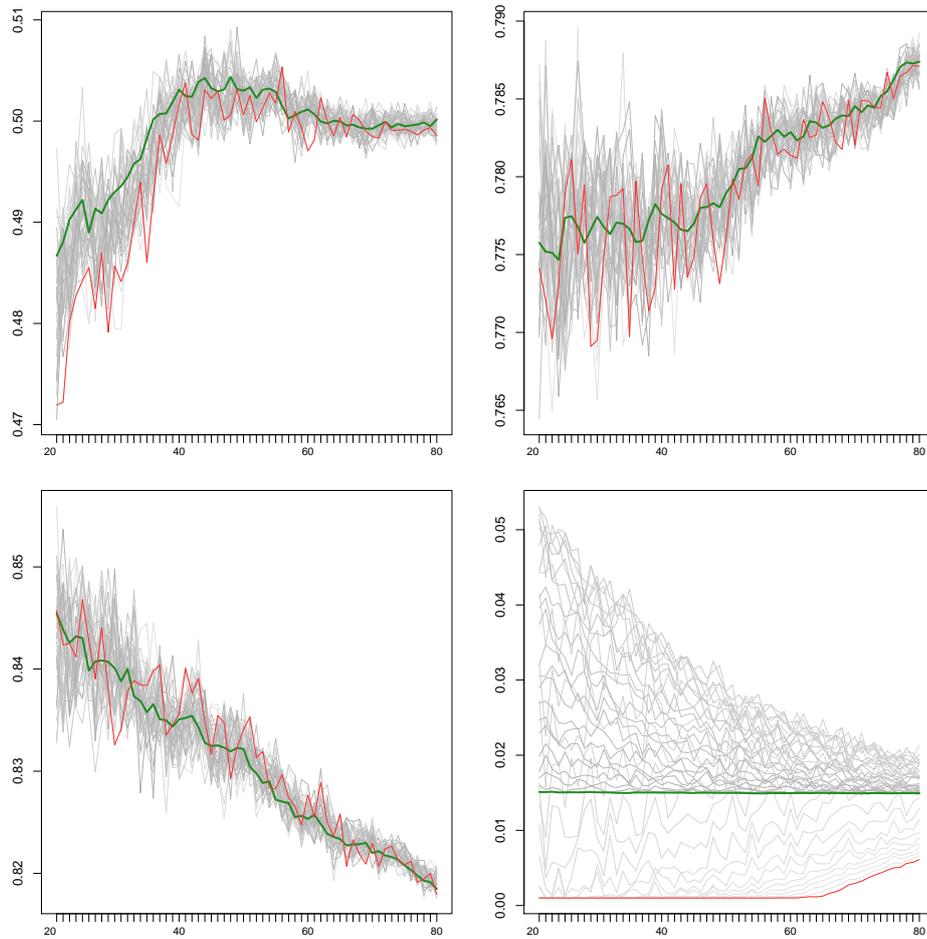


Figura 3.3: Primer ejemplo de simulación: evolución de las luciérnagas a lo largo de las últimas 60 generaciones, para cada parámetro (en el mismo orden que la figura anterior). En verde, evolución de la mejor luciérnaga; en rojo, de la peor.

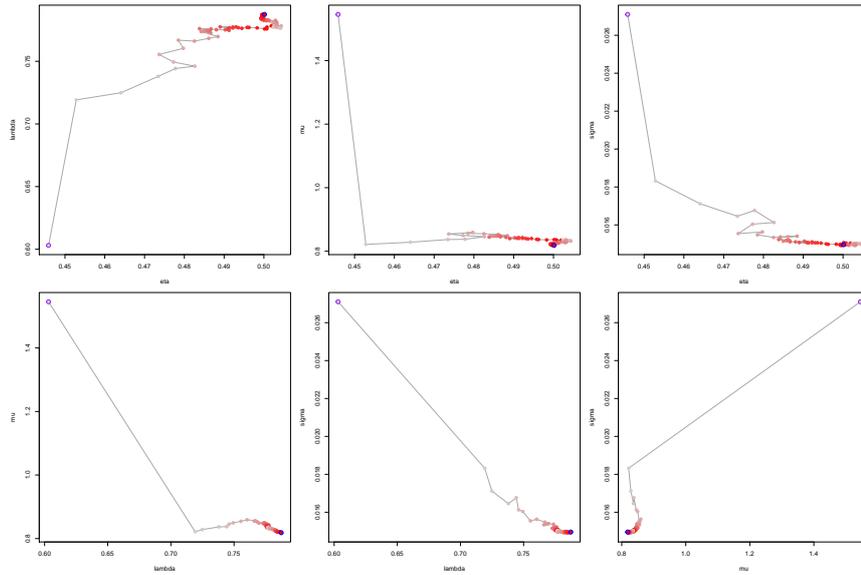


Figura 3.4: Primer ejemplo de simulación: para cada par de parámetros, proyección de la trayectoria bidimensional representando la evolución de la última luciérnaga (la mejor) durante las generaciones (empezando en el punto de color rojo claro y acabando en rojo oscuro).

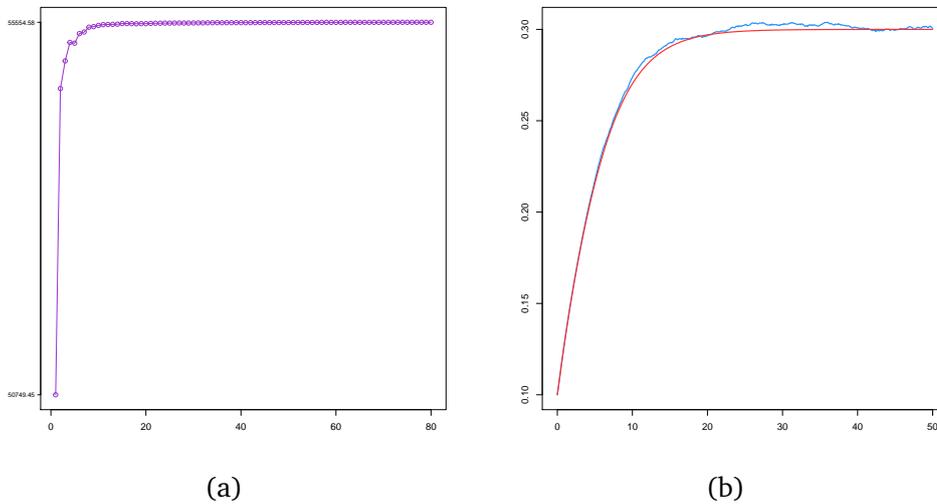


Figura 3.5: Primer ejemplo de simulación: (a): valor de la función objetivo para la mejor luciérnaga. (b): comparación entre las medias observada y ajustada (en rojo).

Por otra parte, y para observar el comportamiento del algoritmo con diferentes valores de los parámetros, simulamos un nuevo conjunto de datos tomando $\eta = 0.0003$, $\lambda = 0.6$, $\mu = 0.8$, $\sigma = 0.025$ y $x_0 = 0.000125$. Sobre

estos datos, aplicamos el algoritmo *Firefly*, variando sus parámetros α, γ, δ , así como el número de luciérnagas n , y consideramos 60 generaciones.

Posteriormente, analizamos el comportamiento de los errores absolutos relativos a cada par de parámetros considerados en la aplicación del algoritmo *Firefly*. Estos errores se han calculado tomando la media sobre todos los demás parámetros.

La tabla 3.2 muestra tales errores variando α y γ . Puede verse que, en general, el error aumenta cuando lo hace γ , independientemente del valor de α , siendo entre 1 y 5 los valores más apropiados para γ . En cambio, para cada γ , los mayores errores están asociados a valores extremos de α , siendo 0.2 y 0.4 aquellos para los cuales los errores observados son menores.

α	γ				
	1	5	10	20	35
0.1	0.0199	0.0311	0.0295	0.0502	0.0454
0.2	0.0089	0.0095	0.0123	0.0220	0.0395
0.4	0.0073	0.0048	0.0111	0.0270	0.0394
0.6	0.0125	0.0115	0.0142	0.0298	0.0495
0.8	0.0112	0.0160	0.0274	0.0479	0.0642
0.9	0.0126	0.0209	0.0306	0.0459	0.0548

Tabla 3.2: Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para α vs. γ .

La tabla 3.3 compara los errores relativos absolutos en función de α y δ . Estos dos parámetros están relacionados debido a que marcan el nivel de aleatoriedad con el cual las luciérnagas se mueven en el espacio. A priori, parece lógico pensar que cuando se establece un valor bajo de α (baja aleatoriedad), este debe disminuir lentamente en las últimas etapas del algoritmo, lo cual está asociado a valores altos de δ . Puede darse la situación inversa cuando se toman valores iniciales altos de α , en cuyo caso un δ pequeño resultaría apropiado. Estas ideas intuitivas encuentran confirmación después de observar esta tabla. De manera más específica, la combinación óptima viene dada por $\alpha = 0.2$ (tal y como sugiere Yang [34]) y $\delta = 0.99$.

En la tabla 3.4 se consideran los parámetros δ y γ . Se puede observar como, en general, para γ fijo, los errores crecen cuando lo hace δ . Más allá del valor de δ , los errores más bajos están asociados a valores de γ entre 1 y 5.

Las tablas 3.5 y 3.7 muestran la combinación entre el número de luciérnagas

α	δ			
	0.9	0.95	0.97	0.99
0.1	0.0331	0.0090	0.0052	0.0079
0.2	0.0167	0.0075	0.0047	0.0043
0.4	0.0115	0.0114	0.0114	0.0254
0.6	0.0109	0.0151	0.0166	0.0442
0.8	0.0135	0.0337	0.0426	0.0748
0.9	0.0099	0.0351	0.0461	0.0767

Tabla 3.3: Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para α vs. δ .

γ	δ			
	0.9	0.95	0.97	0.99
1	0.0276	0.0229	0.0203	0.0228
5	0.0186	0.0161	0.0252	0.0435
10	0.0243	0.0238	0.0343	0.0489
20	0.0331	0.0364	0.0443	0.0614
35	0.0516	0.0574	0.0494	0.0635

Tabla 3.4: Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para γ vs. δ .

n y otros parámetros. En todos los casos, se observa que el error siempre decrece cuando aumenta n , independientemente del resto de parámetros (lo cual parece a todas luces lógico).

Por otro lado, sea cual sea el valor de n , los menores errores están asociados a valores centrales de α , en especial $\alpha = 0.2$ y $\alpha = 0.4$, como puede observarse en la tabla 3.5.

Para un n fijo, la tabla 3.6 muestra que los mejores valores para γ son entre 1 y 10, con $\gamma = 1$ el más recomendable. Finalmente, la tabla 3.7 no muestra grandes diferencias en el error cuando varía δ .

3.4.2. Aplicación a datos reales

Con la finalidad de probar la metodología en una aplicación práctica, consideramos en esta sección un caso real en el contexto de la biología molecular, en particular de la técnica qPCR (*Quantitative Polymerase Chain Reaction*). Este método se emplea para amplificar piezas de ADN o ARN (ácidos nucleicos) aprovechando para ello las propiedades de la enzima polimerasa.

α	n				
	5	10	20	40	60
0.1	0.1484	0.0387	0.0070	0.0025	0.0013
0.2	0.0787	0.0167	0.0031	0.0014	0.0013
0.4	0.0687	0.0150	0.0074	0.0021	0.0018
0.6	0.0788	0.0248	0.0098	0.0054	0.0057
0.8	0.0936	0.0331	0.0187	0.0116	0.0094
0.9	0.0795	0.0380	0.0202	0.0126	0.0128

Tabla 3.5: Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para α vs. n .

γ	n				
	5	10	20	40	60
1	0.0820	0.0365	0.0195	0.0075	0.0059
5	0.0831	0.0363	0.0234	0.0112	0.0103
10	0.0955	0.0399	0.0191	0.0173	0.0126
20	0.1584	0.0545	0.0214	0.0148	0.0144
35	0.1672	0.0747	0.0334	0.0185	0.0113

Tabla 3.6: Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para γ vs. n .

δ	n				
	5	10	20	40	60
0.9	0.0911	0.0362	0.0170	0.0071	0.0038
0.95	0.0843	0.0386	0.0187	0.0080	0.0071
0.97	0.0784	0.0381	0.0259	0.0162	0.0148
0.99	0.1097	0.0568	0.0318	0.0217	0.0201

Tabla 3.7: Segundo ejemplo simulado: error relativo absoluto de \hat{f}_o para δ vs. n .

Existen diferentes técnicas qPCR que permiten la amplificación y cuantificación simultánea del producto obtenido. En particular, la PCR cinética (*kinetic PCR*) cuantifica el producto desde el ciclo en el cual se alcanza un determinado umbral de amplicones de ADN (*DNA amplicon*, fragmentos de ADN origen o producto de un proceso de amplificación o replicación) durante la llamada fase exponencial. Al objeto de calcular este ciclo, la técnica (así como otras) emplea la monitorización de la fluorescencia, y considera el primer ciclo en el que esta sobrepasa un cierto nivel. Rutledge y Cote [24] establecen los criterios para un desarrollo matemático de la técnica kPCR, poniendo el foco en el umbral del logaritmo de la fluorescencia, *log-fluorescencia*. La cuantificación absoluta se logra empleando una curva estándar construida mediante la amplificación de cantidades conocidas de ADN objetivo (*target DNA*) y representando gráficamente los valores obtenidos para el ciclo C_t frente a la concentración de ADN.

Como puede verse, la determinación del ciclo C_t es crítica. Debido a que estamos tratando con un fenómeno dinámico, podemos considerar el ajuste de una difusión de tipo H_1 al objeto de modelar, en cada instante de tiempo (en este caso, ciclos), el nivel de fluorescencia.

Los datos (disponibles como material suplementario en la referencia mencionada anteriormente) corresponden a los niveles de fluorescencia emitidos por diferentes amplificaciones replicadas en una concentración de ADN dada (de hecho, los autores consideran 6 concentraciones, de las cuales trataremos la primera de estas) durante 45 ciclos (figura 3.6). Debemos puntualizar que, dadas las características del proceso, los datos considerados aquí serán los de la fluorescencia, no sus logaritmos.

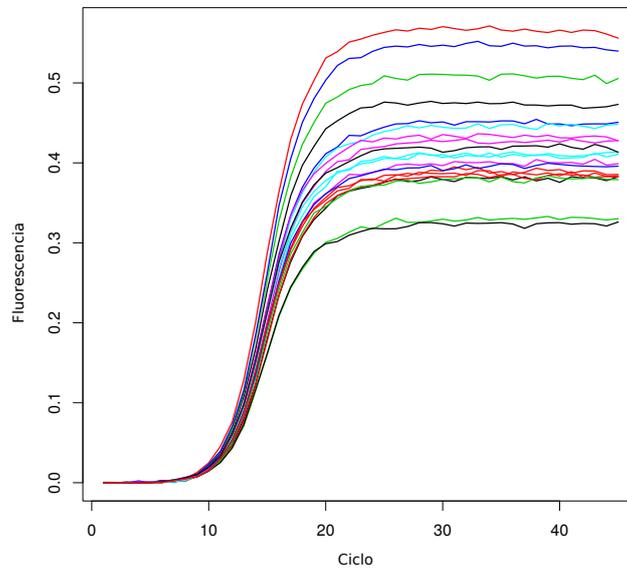


Figura 3.6: Aplicación real: fluorescencia vs. ciclo.

Con la finalidad de aplicar el algoritmo *Firefly* para estimar los parámetros del proceso, consideramos, como parámetros iniciales, una combinación de valores sugerida por el estudio de simulación realizado previamente. En concreto, $\alpha = 0.2, 0.4$; $\gamma = 1, 5$ y $\delta = 0.9, 0.95, 0.97, 0.99$. Además, tomaremos $n = 20$ luciérnagas durante 110 generaciones.

Los resultados se presentan en la tabla 3.8. Para cada combinación de valores iniciales del algoritmo, calculamos los errores relativos absolutos entre los datos observados de fluorescencia y las correspondientes provenientes del proceso, una vez estimados sus parámetros. De esta tabla podemos concluir que la combinación óptima de parámetros iniciales para el algoritmo es $\alpha = 0.2$, $\gamma = 1$ y $\delta = 0.99$.

α	γ	δ	$\hat{\eta}$	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\sigma}$	Error
0.2	1	0.9	0.00031	0.46869	1.89661	0.02499	0.22549
		0.95	0.00029	0.46847	1.88574	0.02500	0.22092
		0.97	0.00030	0.46274	1.94778	0.02500	0.23296
		0.99	0.00030	0.44580	2.27941	0.02500	0.17487
	5	0.9	0.00030	0.44996	2.05263	0.02500	0.26564
		0.95	0.00029	0.46062	1.93521	0.02500	0.26507
		0.97	0.00030	0.47582	1.80397	0.02500	0.22722
		0.99	0.00032	0.43198	2.45740	0.02499	0.20623
0.4	1	0.9	0.00032	0.46943	1.88993	0.02500	0.23309
		0.95	0.00029	0.44280	2.26997	0.02499	0.22014
		0.97	0.00034	0.44229	2.33465	0.02499	0.22091
		0.99	0.00032	0.45137	2.21121	0.02500	0.19202
	5	0.9	0.00028	0.48461	1.66431	0.02489	0.29521
		0.95	0.00029	0.42797	2.53649	0.02500	0.18941
		0.97	0.00028	0.46580	1.86436	0.02500	0.29821
		0.99	0.00031	0.45165	2.17721	0.02500	0.19100

Tabla 3.8: Parámetros estimados y errores relativos absolutos de la aplicación del algoritmo *Firefly* a datos reales.

De estos resultados, en la tabla 3.9 se presentan las estimaciones de los parámetros del proceso.

	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\eta}$	$\hat{\sigma}$
Valor estimado	0.483548	1.6539	0.0002902	0.025

Tabla 3.9: Aplicación real: estimaciones de los parámetros tras aplicar el algoritmo *Firefly*.

En la figura 3.7 se muestra la estabilización de los 4 parámetros estimados para la última replicación.

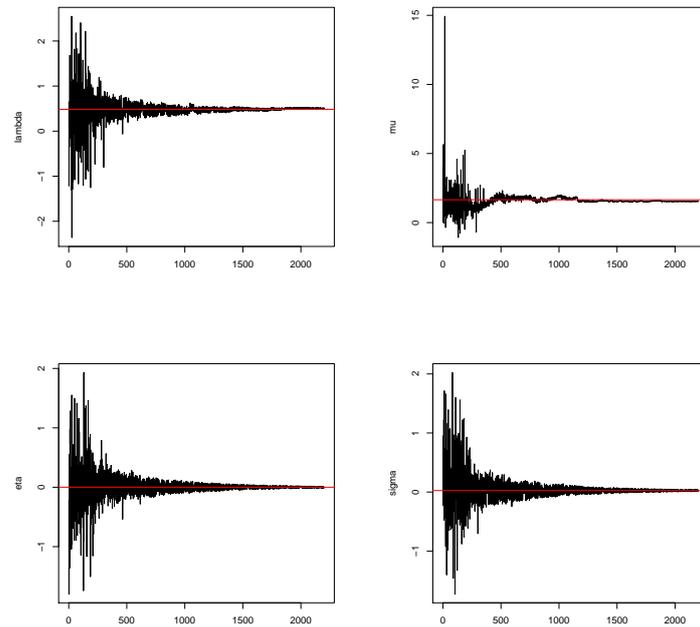


Figura 3.7: Aplicación real: evolución de las estimaciones de los parámetros mediante el algoritmo *Firefly*.

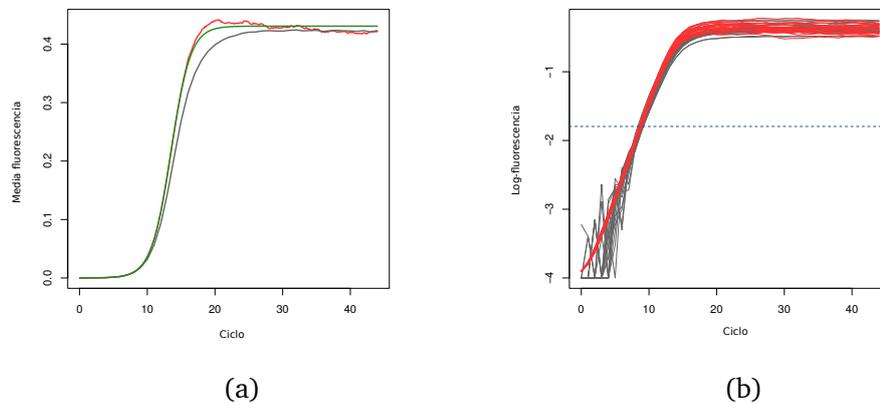


Figura 3.8: Aplicación real: (a) Media de la fluorescencia original (gris) vs. fluorescencia simulada (rojo) vs. fluorescencia teórica ajustada (verde). (b) Log-fluorescencia de los datos originales (gris) y logaritmo de las trayectorias simuladas del proceso (rojo).

Una vez se han obtenido las estimaciones de los parámetros, simulamos un proceso de difusión de tipo H_1 , según la teoría desarrollada en el capítulo

2, bajo las condiciones del experimento. La figura 3.8a muestra la media de los valores observados (tendencia de los datos) junto con la función media teórica (curva H_1) y la media de las trayectorias simuladas.

Con este método podemos simular la fluorescencia del procedimiento de amplificación ante diferentes reacciones, lo cual nos permite emplear datos simulados en lugar de reales, con el impacto que ello puede tener en la reducción de costes, tanto de tiempo como económicos. Aún cuando el proceso simulado difiere del original en los primeros ciclos, la cuestión clave aquí es si podemos encontrar los ciclos en los cuales se alcanza el umbral de fluorescencia predeterminado (véase la figura 3.8b, mostrando los valores de log-fluorescencia para los datos originales y el logaritmo de las trayectorias simuladas del proceso estimado). Esta cuestión puede abordarse empleando técnicas asociadas con el estudio de variables temporales en el contexto de los procesos de difusión, como puede ser el caso de los tiempos de primer-paso.

Conclusiones

Este trabajo surge con los objetivos fundamentales de establecer el marco teórico para la definición de un proceso logístico generalizado, esto es, un proceso de difusión cuya media sea una curva de tipo logístico, expresable en base a una función derivable y no acotada superiormente. Además, deben considerarse las dificultades derivadas de su aplicación a datos (simulados o reales) mediante procedimientos de inferencia.

Con todo ello, la memoria agrupa los trabajos realizados para conseguir, en primer lugar, un proceso de difusión cuya función media sea la curva logística clásica, así como la estimación de los momentos infinitesimales que lo caracterizan para el caso de un proceso gaussiano.

Una vez conseguido el paso de un modelo determinista a uno estocástico, con ciertas garantías para abordar estudios de inferencia, se plantea la cuestión central del trabajo, y que parte del hecho de tener numerosas variantes de la curva logística y la necesidad de agrupar en un solo modelo estocástico todos los posibles comportamientos. Para ello, se introduce previamente el concepto de curva logística generalizada, desde un enfoque funcional, mediante un operador definido en un espacio de funciones derivables y no acotadas superiormente. Con esto se consiguen las expresiones finales de la difusión en base a las funciones que caracterizan a cada curva de tipo logístico, así como sus parámetros propios. A partir de ahí se emplea la metodología de la primera parte del trabajo para construir el proceso de difusión lognormal con factores exógenos, que en este caso dependerá de una función que le permitirá adoptar distintas tendencias sigmoidales. El principal obstáculo en este caso llega a la hora de valorar la aplicabilidad del modelo cuando se dispone de datos provenientes de observaciones que presenten un comportamiento de tipo sigmoidal. La solución propuesta hace uso de una estimación paramétri-

ca implícita que permite conseguir otra curva de tipo logístico distinta pero que, con los parámetros adecuados, se comporta como la original, que era inaccesible por constar de un parámetro imposible de estimar a partir de los datos.

Finalmente, se considera el caso particular de la curva hiperbolística de tipo I para construir la difusión asociada y abordar el estudio de inferencia. En este caso se enfrenta al problema derivado de la complejidad de las ecuaciones de verosimilitud, por lo que se introduce un novedoso algoritmo metaheurístico, el algoritmo *Firefly*, para resolver el sistema y conseguir así las estimaciones máximo-verosímiles de los parámetros del proceso.

Líneas futuras

Como cuestiones a tratar en futuras investigaciones, pueden destacarse, principalmente, el estudio de estimación para la variabilidad de procesos más complejos (aquí se aborda la cuestión para procesos gaussianos), así como el uso de otras varianzas infinitesimales que no dificulten la obtención de características del proceso (recuérdese que la metodología aquí estudiada trata el proceso lognormal con factores exógenos).

Por otro lado, del último ejemplo del trabajo se puede esperar el estudio de tiempos de primer paso, y quizás obtener sus expresiones para el proceso de difusión logístico generalizado, lo que llevaría a obtener resultados dependientes de funciones y, por lo tanto, con un amplio espectro de aplicación y estudio.

Además, el marco teórico de la curva logística generalizada, puede extenderse a otros modelos siguiendo la misma metodología. Cabe destacar la propuesta básica realizada para el modelo Weibull y presentada por el autor de esta memoria y sus directores en el congreso COMPSTAT 2014. De igual manera, sería posible extender a otros modelos y obtener una colección de difusiones generalizadas *canónicas* que engloben la mayoría de comportamientos dinámicos habituales en la literatura.

Por otro lado, debe tenerse en cuenta la finalidad práctica de estas metodologías, es decir, su utilidad en otros campos de investigación. En efecto, la modelización debe entenderse como una traslación a un contexto matemático.

Como añadido final, y solo quizás desde una perspectiva algo más filosófica, obtener modelos matemáticos generalizados permitiría la introducción de nuevos enfoques para describir un mismo fenómeno dinámico. Así, se podría considerar un conjunto paramétrico sin origen físico y entonces compararlo, habida cuenta el sentido físico del fenómeno, con sus parámetros originales, abordando así cuestiones propias del fenómeno.

Apéndice A

Cálculo de las ecuaciones de verosimilitud

Del logaritmo de la función de verosimilitud (3.5), las estimaciones de η, λ, μ y σ^2 se obtienen como soluciones del sistema de ecuaciones

$$\begin{aligned} \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{R_{ij} - T_{ij}^{\eta, \lambda, \mu}}{\Delta_{ij} S_{ij}^{\eta, \lambda, \mu}} W_{ij}^{\eta, \lambda, \mu} + \frac{\sigma^2}{2} \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{W_{ij}^{\eta, \lambda, \mu}}{S_{ij}^{\eta, \lambda, \mu}} &= 0 \\ \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{R_{ij} - T_{ij}^{\eta, \lambda, \mu}}{\Delta_{ij} S_{ij}^{\eta, \lambda, \mu}} V_{ij}^{\eta, \lambda, \mu} + \frac{\sigma^2}{2} \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{V_{ij}^{\eta, \lambda, \mu}}{S_{ij}^{\eta, \lambda, \mu}} &= 0 \\ \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{R_{ij} - T_{ij}^{\eta, \lambda, \mu}}{\Delta_{ij} S_{ij}^{\eta, \lambda, \mu}} Z_{ij}^{\eta, \lambda, \mu} + \frac{\sigma^2}{2} \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{Z_{ij}^{\eta, \lambda, \mu}}{S_{ij}^{\eta, \lambda, \mu}} &= 0 \\ \sigma^4 \sum_{i=1}^d \Delta_{i1}^{in_i} + 4\sigma^2(N-d) - 4 \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{(R_{ij})^2}{\Delta_{ij}} - 4 \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{(T_{ij}^{\eta, \lambda, \mu})^2}{\Delta_{ij}} + 8 \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{R_{ij} T_{ij}^{\eta, \lambda, \mu}}{\Delta_{ij}} &= 0 \end{aligned}$$

donde

$$\begin{aligned} S_{ij}^{\eta, \lambda, \mu} &= (\eta + \xi(t_{i,j-1}))(\eta + \xi(t_{ij})) \\ V_{ij}^{\eta, \lambda, \mu} &= \lambda^{-1} [t_{i,j-1} \xi(t_{i,j-1})(\eta + \xi(t_{ij})) - t_{ij} \xi(t_{ij})(\eta + \xi(t_{i,j-1}))] \\ W_{ij}^{\eta, \lambda, \mu} &= \xi(t_{ij}) - \xi(t_{i,j-1}) \\ Z_{ij}^{\eta, \lambda, \mu} &= \mu^{-1} [\operatorname{arcsenh}(t_{i,j-1}) \xi(t_{i,j-1})(\eta + \xi(t_{ij})) - \operatorname{arcsenh}(t_{ij}) \xi(t_{ij})(\eta + \xi(t_{i,j-1}))] \end{aligned}$$

Resolviendo la última ecuación para σ , tenemos

$$\frac{\sigma^2}{2} = U^{-\frac{1}{2}} \Gamma_{\eta, \lambda, \mu}^{\frac{1}{2}}$$

donde

$$\Gamma_{\eta, \lambda, \mu} = \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{1}{\Delta_{ij}} (R_{ij} - T_{ij}^{\eta, \lambda, \mu})^2 \quad \text{y} \quad U = 4(N - d) + \sum_{i=1}^d \Delta_{i1}^{in_i}.$$

Esta expresión depende únicamente de η , λ y μ . Sustituyendo en las otras ecuaciones, se consigue el sistema

$$\Lambda_{\mathcal{J}}^{\eta, \lambda, \mu} + U^{-\frac{1}{2}} \Gamma_{\eta, \lambda, \mu}^{\frac{1}{2}} \Psi_{\mathcal{J}}^{\eta, \lambda, \mu} = 0, \quad \mathcal{J} = W, V, Z,$$

donde

$$\Lambda_{\mathcal{J}}^{\eta, \lambda, \mu} = \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{R_{ij} - T_{ij}^{\eta, \lambda, \mu}}{\Delta_{ij} S_{ij}^{\eta, \lambda, \mu}} \mathcal{J}_{ij}^{\eta, \lambda, \mu} \quad \text{y} \quad \Psi_{\mathcal{J}}^{\eta, \lambda, \mu} = \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{\mathcal{J}_{ij}^{\eta, \lambda, \mu}}{S_{ij}^{\eta, \lambda, \mu}}, \quad \mathcal{J} = W, V, Z.$$

La solución del sistema de ecuaciones anterior nos da las estimaciones máximo-verosímiles $\hat{\eta}$, $\hat{\lambda}$ y $\hat{\mu}$, desde las cuales, $\hat{\sigma}^2 = 2U^{-\frac{1}{2}} \Gamma_{\hat{\eta} \hat{\lambda} \hat{\mu}}^{\frac{1}{2}}$.

Apéndice B

Código en R para el algoritmo Firefly

Se aplica a minimizar la función objetivo `fun.obj` asociada al proceso de difusión hiperbólico de tipo I, simulando `Ntray` trayectorias durante `gens` generaciones (datos en una matriz `X`) desde `x0` para 4 parámetros `eta`, `lambda`, `mu`, `sigma` y parámetros del algoritmo `alg.alfa`, `alg.beta0`, `alg.beta`, `alg.delta`, `alg.gamma`.

```
m <- NULL
for (i in 1:Ntray){
  m <- c(m,max(X[,i])/x0 - 1)
}

eta.sol <- c(1/max(m),1/min(m))
lambda.sol <- c(0,1)
sigma.sol <- c(0.001,0.5)

eta.rango <- diff(eta.sol)
lambda.rango <- diff(lambda.sol)
sigma.rango <- diff(sigma.sol)

eta.n <- runif(n) * eta.rango + eta.sol[1]
lambda.n <- runif(n) * lambda.rango + lambda.sol[1]
sigma.n <- runif(n) * sigma.rango + sigma.sol[1]
mu.n <- runif(n) * 1/lambda.n^(sqrt(1+t0^2))

L.n <- rep(0,n)

for (i in 1:gens){

  F.n <- Vectorize(fun.obj)(eta.n,lambda.n,mu.n,sigma.n)

  L.n <- sort(F.n)
  index <- order(F.n)
```

```

eta.n <- eta.n[index]
lambda.n <- lambda.n[index]
mu.n <- mu.n[index]
sigma.n <- sigma.n[index]

eta.o <- eta.n
lambda.o <- lambda.n
mu.o <- mu.n
sigma.o <- sigma.n

L.o <- L.n

for (i in 1:n){
  for (j in 1:n){
    r <- sqrt((eta.n[i] - eta.o[j])^2 + (lambda.n[i] -
      lambda.o[j])^2 + (mu.n[i] - mu.o[j])^2 + (sigma
      .n[i] - sigma.o[j])^2)
    if (L.n[i] < L.o[j]){
      alg.beta0 <- 1
      alg.beta <- alg.beta0 * exp(-alg.gamma*r
      ^2)
      eta.n[i] <- eta.n[i] * (1 - alg.beta) +
      eta.o[j] * alg.beta + alg.alfa * (runif
      (1) - 0.5)
      lambda.n[i] <- lambda.n[i] * (1 - alg.beta
      ) + lambda.o[j] * alg.beta + alg.alfa *
      (runif(1) - 0.5)
      mu.n[i] <- mu.n[i] * (1 - alg.beta) + mu.o
      [j] * alg.beta + alg.alfa * (runif(1) -
      0.5)
      sigma.n[i] <- sigma.n[i] * (1 - alg.beta) +
      sigma.o[j] * alg.beta + alg.alfa * (
      runif(1) - 0.5)
    }
  }
}

for (i in 1:n){
  if (eta.n[i] <= eta.sol[1]){ eta.n[i] <- eta.sol[1] }
  if (eta.n[i] >= eta.sol[2]){ eta.n[i] <- eta.sol[2] }
  if (lambda.n[i] <= lambda.sol[1]){ lambda.n[i] <- lambda.
  sol[1] }
  if (lambda.n[i] >= lambda.sol[2]){ lambda.n[i] <- lambda.
  sol[2] }
  if (mu.n[i] <= mu.sol[1]){ mu.n[i] <- mu.sol[1] }
  if (mu.n[i] >= mu.sol[2]){ mu.n[i] <- mu.sol[2] }
  if (sigma.n[i] <= sigma.sol[1]){ sigma.n[i] <- sigma.sol[1]
  }
  if (sigma.n[i] >= sigma.sol[2]){ sigma.n[i] <- sigma.sol[2]
  }
}

alg.alfa <- alg.alfa*alg.delta
}

```

Bibliografía

- [1] M. Alb, P. Alotto, C. Magele, W. Renhart, K. Preis y B. Trapp. “Firefly Algorithm for Finding Optimal Shapes of Electromagnetic Devices”. En: *IEEE Transactions on Magnetics* 52.3 (mar. de 2016), págs. 1-4. DOI: [10.1109/TMAG.2015.2483058](https://doi.org/10.1109/TMAG.2015.2483058) (vid. pág. 70).
- [2] L. Arnold. *Stochastic differential equations: theory and applications*. New York: Wiley-Interscience, 1974 (vid. págs. 29, 31, 50).
- [3] A. Barrera-García, P. Román-Román y F. Torres-Ruiz. “Fitting dynamic growth models of biological phenomena from sample observations through Gaussian diffusion processes”. En: *Biosystems* 112.3 (2013). Selected papers presented at BIOCOMP 2012, págs. 284-291. DOI: [10.1016/j.biosystems.2012.12.007](https://doi.org/10.1016/j.biosystems.2012.12.007) (vid. pág. 30).
- [4] R. Capocelli y L. Ricciardi. “A diffusion model for population growth in random environment”. En: *Theoretical Population Biology* 5.1 (1974), págs. 28-41 (vid. pág. 17).
- [5] R. Capocelli y L. Ricciardi. “Growth with regulation in random environment”. En: *Biological Cybernetics* 15.3 (1974), págs. 147-157 (vid. págs. 17, 20, 29).
- [6] W. M. Eby, M. A. Tabatabai y Z. Bursac. “Hyperbolic modeling of tumor growth with a combined treatment of iodoacetate and dimethylsulphoxide”. En: *BMC Cancer* 10.1 (2010), pág. 509 (vid. págs. 16, 70).
- [7] W. Feller. “Die Grundlagen der Volterraschen Theorie Des Kampfes Ums Dasein in Wahrscheinlichkeitstheoretischer Behandlung”. En: *Acta Biotheoretica* 5.1 (1939), págs. 11-40 (vid. págs. 16, 27).
- [8] W. Feller. “Diffusion processes in one dimension”. En: *Trans. Amer. Math. Soc* 77.1 (1953), págs. 1-31 (vid. pág. 16).
- [9] W. Feller. “On the logistic law of growth and its empirical verifications in biology”. En: *Acta Biotheoretica* 5.2 (1940), págs. 51-66. DOI: [10.1007/BF01602862](https://doi.org/10.1007/BF01602862) (vid. págs. 16, 20).

- [10] I. Fister, I. Fister Jr, X.-S. Yang y J. Brest. “A comprehensive review of firefly algorithms”. En: *Swarm and Evolutionary Computation* (2013) (vid. pág. 76).
- [11] I. Fister, X.-S. Yang, J. Brest y I. Fister Jr. “On the Randomized Firefly Algorithm”. En: *Cuckoo Search and Firefly Algorithm*. Springer, 2014, págs. 27-48 (vid. pág. 77).
- [12] A. Gandomi, X.-S. Yang, S. Talatahari y A. Alavi. “Firefly algorithm with chaos”. En: *Communications in Nonlinear Science and Numerical Simulation* 18.1 (2013), págs. 89-98 (vid. págs. 70, 77).
- [13] A. Giovanis y C. Skiadas. “A stochastic logistic innovation diffusion model studying the electricity consumption in Greece and the United States”. En: *Technological Forecasting and Social Change* 61.3 (1999), págs. 235-246 (vid. págs. 20, 30, 34).
- [14] R. Gutiérrez, N. Rico, P. Román y F. Torres. “Approximate and generalized confidence bands for some parametric functions of the lognormal diffusion process with exogenous factors”. En: *Scientiae Mathematicae Japonicae* 64 (2006), págs. 843-859 (vid. págs. 17, 31, 32).
- [15] R. Gutiérrez-Jáimez, P. Román, D. Romero, J. J. Serrano y F. Torres. “A new Gompertz-type diffusion process with application to random growth”. En: *Mathematical Biosciences* 208.1 (2007), págs. 147-165. DOI: [10.1016/j.mbs.2006.09.020](https://doi.org/10.1016/j.mbs.2006.09.020) (vid. págs. 17, 31).
- [16] A. Kavousi-Fard, H. Samet y F. Marzbani. “A new hybrid Modified Firefly Algorithm and Support Vector Regression model for accurate Short Term Load Forecasting”. En: *Expert Systems with Applications* (2014) (vid. págs. 70, 76).
- [17] T. Malthus. *An Essay on the Principle of Population*. London: J. Johnson, 1798 (vid. págs. 15, 24).
- [18] T. Niknam, R. Azizipanah-Abarghooee y A. Roosta. “Reserve constrained dynamic economic dispatch: a new fast self-adaptive modified firefly algorithm”. En: *Systems Journal, IEEE* 6.4 (2012), págs. 635-646 (vid. pág. 76).
- [19] L. M. Ricciardi. “On the transformation of diffusion processes into the Wiener process”. En: *Journal of Mathematical Analysis and Applications* 54.1 (1976), págs. 185-199. DOI: [http://dx.doi.org/10.1016/0022-247X\(76\)90244-4](http://dx.doi.org/10.1016/0022-247X(76)90244-4) (vid. págs. 21, 48).
- [20] P. Román-Román, D. Romero, M. Rubio y F. Torres-Ruiz. “Estimating the parameters of a Gompertz-type diffusion process by means of Simulated Annealing”. En: *Applied Mathematics and Computation* 218.9 (2012), págs. 5121-5131 (vid. págs. 19, 76).

- [21] P. Román-Román y F. Torres-Ruiz. “A stochastic model related to the Richards-type growth curve. Estimation by means of simulated annealing and variable neighborhood search”. En: *Applied Mathematics and Computation* 266 (2015), págs. 579-598. DOI: <http://dx.doi.org/10.1016/j.amc.2015.05.096> (vid. págs. 19, 76).
- [22] P. Román-Román, D. Romero y F. Torres-Ruiz. “A diffusion process to model generalized von Bertalanffy growth patterns: Fitting to real data”. En: *Journal of Theoretical Biology* 263.1 (2010), págs. 59-69. DOI: [10.1016/j.jtbi.2009.12.009](http://dx.doi.org/10.1016/j.jtbi.2009.12.009) (vid. págs. 17, 31).
- [23] P. Román-Román y F. Torres-Ruiz. “Modeling logistic growth by a new diffusion process: application to biological systems”. En: *Biosystems* 110 (2012), págs. 9-21 (vid. págs. 17, 20, 30, 31, 33, 42, 56).
- [24] R. Rutledge y C. Cote. “Mathematics of quantitative kinetic PCR and the application of standard curves”. En: *Nucleic Acids Research* 31.16 (2003), e93-e93 (vid. pág. 90).
- [25] H. Schurz. “Modeling, analysis and discretizations of stochastic logistic equations”. En: *International Journal of Numerical Analysis and Modeling* 4.2 (2007), págs. 178-197 (vid. pág. 29).
- [26] M. Tabatabai, W. Eby y Z. Bursac. “Oscillabolastic model, a new model for oscillatory dynamics, applied to the analysis of Hes1 gene expression and Ehrlich ascites tumor growth”. En: *Journal of Biomedical Informatics* 45.3 (jun. de 2012), págs. 401-407. DOI: [10.1016/j.jbi.2011.11.016](http://dx.doi.org/10.1016/j.jbi.2011.11.016) (vid. pág. 70).
- [27] M. A. Tabatabai, Z. Bursac, W. M. Eby y K. P. Singh. “Mathematical modeling of stem cell proliferation”. En: *Medical & Biological Engineering & Computing* 49.3 (mar. de 2011), págs. 253-262. DOI: [10.1007/s11517-010-0686-y](http://dx.doi.org/10.1007/s11517-010-0686-y) (vid. págs. 16, 70).
- [28] M. A. Tabatabai, W. M. Eby, K. P. Singh y S. Bae. “T model of growth and its application in systems of tumor-immune dynamics”. En: *Mathematical Biosciences and Engineering* 10.3 (jun. de 2013), págs. 925-938. DOI: [10.3934/mbe.2013.10.925](http://dx.doi.org/10.3934/mbe.2013.10.925) (vid. pág. 70).
- [29] M. Tabatabai, D. Williams y Z. Bursac. “Hyperbolastic growth models: theory and application”. En: *Theoretical Biology and Medical Modelling* 2.1 (2005), pág. 14. DOI: [10.1186/1742-4682-2-14](http://dx.doi.org/10.1186/1742-4682-2-14) (vid. págs. 16, 20, 47, 64, 69).
- [30] A. Tsoularis y J. Wallace. “Analysis of logistic growth models”. En: *Mathematical Biosciences* 179 (2002), págs. 21-55 (vid. págs. 16, 20, 26, 45).

- [31] H. C. Tuckwell y J. A. Koziol. “Logistic population growth under random dispersal”. En: *Bulletin of Mathematical Biology* 49.4 (1987), págs. 495-506. DOI: [http://dx.doi.org/10.1016/S0092-8240\(87\)80010-1](http://dx.doi.org/10.1016/S0092-8240(87)80010-1) (vid. pág. 29).
- [32] M. E. Turner, E. L. Bradley, K. A. Kirk y K. M. Pruitt. “A theory of growth”. En: *Mathematical Biosciences* 29.3 (1976), págs. 367-373. DOI: [http://dx.doi.org/10.1016/0025-5564\(76\)90112-7](http://dx.doi.org/10.1016/0025-5564(76)90112-7) (vid. pág. 26).
- [33] P. F. Verhulst. “Notice sur la loi que la population suit dans son accroissement”. En: *Curr. Math. Phys* 10 (1838), págs. 113+ (vid. págs. 15, 24).
- [34] X.-S. Yang. *Nature-Inspired Metaheuristic Algorithms*. Luniver Press, 2008 (vid. págs. 21, 70, 76, 79, 87).
- [35] X.-S. Yang y X. He. “Firefly algorithm: recent advances and applications”. En: *International Journal of Swarm Intelligence* 1.1 (2013), págs. 36-50 (vid. págs. 76, 77).
- [36] L. Zhang, L. Liu, X.-S. Yang y Y. Dai. “A Novel Hybrid Firefly Algorithm for Global Optimization”. En: *PLoS ONE* 11.9 (sep. de 2016), págs. 1-17. DOI: [10.1371/journal.pone.0163230](https://doi.org/10.1371/journal.pone.0163230) (vid. pág. 70).