

## ¿Cómo ilustrar con la ayuda del ordenador la construcción de diferentes tipos de modelos para describir datos reales de crecimiento poblacional?

Marta Ginovart

email: [marta.ginovart@upc.edu](mailto:marta.ginovart@upc.edu)

Departamento de Matemática Aplicada, Universitat Politècnica de Catalunya,  
Barcelona

### RESUMEN

El objetivo de este trabajo es diseñar e implementar en el aula un conjunto de tareas para, con la ayuda del ordenador, investigar y describir un conjunto de datos reales correspondiente a un crecimiento poblacional. Se ilustra paso a paso la construcción de modelos empíricos y de modelos pseudo-mecanicistas en este contexto. Se presentan diferentes estrategias o métodos para modelizar este conjunto de datos, y se constata cómo estas distintas aproximaciones generan una variedad de respuestas plausibles a la situación que se estudia. Se fomenta la aplicación de conocimientos previos de los estudiantes para la construcción e interpretación de estos modelos, así como se propicia la comprensión y profundización de ideas matemáticas implicadas en sus formulaciones y reparametrizaciones.

*Crecimiento poblacional, Datos reales, Modelo empírico, Modelo pseudo-mecanicista, Funciones sigmoideas*

## Introducción

La enseñanza y el aprendizaje de las matemáticas en cualquier contexto deberían promover el desarrollo del pensamiento y la posibilidad de un uso adecuado de la tecnología disponible hoy en día. Procesos como el favorecer o impulsar la curiosidad, el pensamiento crítico, el razonamiento, así como formas de verificación, refutación, y deducción deberían encontrarse en las actividades de aula que proponemos a nuestros estudiantes. Asimismo, hay que tener en cuenta que para algunos de estos procesos y su entrenamiento la ayuda de un ordenador puede ser una herramienta muy valiosa y útil. De hecho, en el siglo XXI, la computación es más que un soporte o ayuda para la actividad científica, la computación está cambiando la forma fundamental en que se practica la ciencia y también en la manera en cómo se está aprendiendo y se está enseñando [1,4,10]. La computación nos permite obtener y analizar grandes volúmenes de datos, considerar y resolver problemas inaccesibles hasta ahora, construir modelos sofisticados y testarlos, visualizar fenómenos o realizar experimentos difíciles o imposibles en los laboratorios, entre otras muchas opciones.

Está ampliamente aceptado que el pensamiento matemático surge y se desarrolla en una compleja interacción de lenguajes y representaciones. Existe un término o concepto relativamente nuevo que es el "pensamiento computacional" [8,11]. A pesar de que todavía está en discusión su definición, parece centrarse en ideas o tópicos de la informática con relación a procesos de resolución de problemas, tales como: el reconocimiento de patrones o tendencias, la generación de pautas o esquemas, la abstracción (composición, descomposición, generalización y especialización), la modelización, el diseño de algoritmos (secuencia, iteración y selección), el análisis de datos y la visualización. Con todo ello, el "pensamiento computacional" puede ser considerado como una habilidad fundamental para todas las personas, no sólo para los informáticos, y en particular, algo muy atractivo para los matemáticos y para cualquiera persona que esté involucrada o inmersa en la enseñanza y el aprendizaje de las matemáticas. Por ejemplo, sería deseable diseñar y desarrollar estrategias acertadas y eficientes para abordar la descripción de datos reales y generar modelos que permitieran explicar o representar estos datos, a la vez que conocer cómo utilizar los recursos computacionales que están disponibles hoy en día en los ordenadores con los que trabajamos o que encontramos en las aulas para que nos ayuden en estas actividades. El ordenador debería ser visto como un "compañero" indispensable con el cual se debería poder entrenar habilidades diversas y facilitar la aplicación de conceptos o conocimientos previos para la consecución de soluciones a problemas planteados. El potencial de los programas disponibles en la mayoría de los ordenadores que tenemos, o bien de acceso libre desde Internet, no debería permanecer inexplorado e inexplorado cuando, por ejemplo, actividades relacionadas con la matemática aplicada, modelos cuantitativos y métodos numéricos se llevan a cabo en el aula.

En el contexto del Proyecto de Matemáticas del Milenio, una iniciativa de divulgación y educación matemática, se puede encontrar la página Web *nRich* (<http://nrich.maths.org/>) que contiene una lista con algunas de las características que hacen que una tarea sea una "tarea rica", teniendo en consideración que, fundamentalmente, es cómo la tarea es utilizada en la clase que la hace realmente una "tarea rica". Resumiendo parte del contenido de esta web podríamos señalar que "tareas ricas" son, por ejemplo, las que:

- ✓ son accesibles a una amplia gama de estudiantes
- ✓ atraen a los estudiantes con algún punto de partida interesante o descubrimientos iniciales intrigantes
- ✓ ofrecen oportunidades para conseguir un éxito inicial
- ✓ ofrecen diferentes niveles de desafío
- ✓ permiten utilizar diferentes métodos o estrategias y obtener una variedad de respuestas plausibles
- ✓ tienen el potencial de reforzar y ampliar habilidades de los estudiantes
- ✓ posibilitan el ahondar en la comprensión de ideas matemáticas
- ✓ fomentan la creatividad y la aplicación imaginativa de conocimientos
- ✓ tienen la capacidad de revelar patrones o conducir a generalizaciones

- ✓ alientan a los estudiantes a desarrollar confianza e independencia.

En este contexto o marco de actuación, el análisis de datos reales y la construcción de diferentes tipos de modelos para describir o representar estos datos son buenas opciones para poder diseñar “tareas ricas” y conseguir que los estudiantes puedan ganar la confianza suficiente para aplicar sus conocimientos en una variedad de contextos, al mismo tiempo que encajan indudablemente con la necesidad de utilizar el ordenador y sus recursos computacionales. Por ejemplo, en el campo de la “biología matemática” o “matemática biológica”, la capacidad de diseñar actividades de clase que involucrasen modelos cuantitativos con conceptos y herramientas matemáticas para investigar sistemas dinámicos sería muy deseable y valorada. Además, este tipo de actividades pueden justificar y dar espacio para introducir métodos computacionales complementarios a los métodos más clásicos y tradicionales [3,7]. La posibilidad de sugerir contextos o situaciones reales para explorar ideas matemáticas y desarrollar modelos matemáticos aplicables a esas realidades es de gran interés para la actividad diaria en el aula, con independencia del nivel educativo en el cual nos encontremos (secundaria, bachillerato o universidad). En un contexto académico donde la biología está presente como un campo de aplicación, los modelos se pueden utilizar para poder analizar y comprender los fenómenos, o bien para diseñar y construir instrumentos que hagan posible una “experimentación” virtual con el fin de mejorar de forma iterativa nuestra representación de la realidad. Es por eso que cada nivel educativo debería encontrar aquellos modelos apropiados con los que se pudieran trabajar atendiendo a los conocimientos previos que se requieran, tanto de matemáticas como de biología.

Hay varias categorías de clasificaciones para los modelos en función de criterios diversos. Por ejemplo, si en el modelo se considera el tiempo, el modelo es un modelo dinámico. Los modelos de crecimiento poblacional, evoluciones temporales de la cantidad de individuos que configuran una población en un determinado entorno, se reconocen como modelos dinámicos primarios, y tienen un gran interés en el ámbito de la biología y la predicción. Los modelos mecanicistas o heurísticos, en el grupo de modelos primarios, son modelos cuyo desarrollo viene de la comprensión de los procesos bioquímicos o biológicos subyacentes que rigen el fenómeno poblacional y donde, a los parámetros implicados en estos modelos, se les supone un significado biológico, mientras que los modelos empíricos son funciones matemáticas que simplemente se ajustan a las observaciones del fenómeno a lo largo del tiempo [9]. Las funciones lineales o los polinomios, cuando se usan para ajustar datos observados por medio de un método de regresión, son ejemplos clásicos de modelos empíricos.

Teniendo en cuenta que la mayoría de los modelos con más tradición en el contexto académico para representar las evoluciones temporales de las poblaciones son modelos continuos en la clase de los modelos empíricos, tiene un valor añadido trabajar con estos modelos pero demandando o reivindicando que los parámetros involucrados en el modelo tengan un significado biológico, en sintonía con el contexto del conjunto de datos reales que se quiera representar. Este tipo de modelo no es completamente un modelo mecanicista o heurístico, por lo que se llama modelo pseudo-mecanicista o pseudo-heurístico [9]. Tratar con este tipo de modelo pseudo-mecanicista es una buena ocasión para vincular herramientas y conceptos matemáticos con ideas biológicas, y también una magnífica oportunidad de hacer uso de los ordenadores y aprovechar la ayuda que nos pueden proporcionar en este contexto.

El objetivo general de este trabajo es diseñar e implementar en el aula un conjunto de “tareas ricas” para llevar a cabo, con la ayuda del ordenador, la investigación y representación de un conjunto de datos reales (una evolución temporal de una población microbiana creciendo en un entorno cerrado), ilustrando tanto la construcción de modelos empíricos como de modelos pseudo-mecanicista en este contexto de aplicación. Uno de los principales propósitos en el diseño de este conjunto de tareas es configurar un entorno adecuado para mostrar diferentes estrategias o métodos para describir este conjunto de datos reales, y constatar cómo estos distintos enfoques o aproximaciones pueden generar una variedad de respuestas plausibles a la realidad que nos ocupa. Las tareas diseñadas deberían fomentar aplicaciones imaginativas de los conocimientos previos que los estudiantes tienen con el fin de hacer frente a la construcción e interpretación de estos modelos, partiendo de unas primeras aproximaciones iniciales y no complicadas para lograr modelos relativamente simples, y avanzar después en el uso de modelos más sofisticados y complejos, profundizando en la comprensión de las ideas matemáticas implicadas, mostrando como la ayuda que diversos programas de ordenador ofrecen resulta ser indispensable y muy valiosa en todo este proceso. Con todo esto los

objetivos específicos son:

- a) Conocer que conocimientos previos pueden utilizar los estudiantes para construir modelos que describan crecimientos poblacionales, y que programas accesible en los ordenadores que utilizamos, y que ya conocen, se pueden usar para llevar a cabo el análisis de un conjunto de datos reales relativos al crecimiento de una población.
- b) Diseñar (e implementar en el aula) un conjunto de tareas agrupadas en distintos módulos para poder desarrollar en distintos entornos de computación, como son una hoja de cálculo, un programa matemático y un programa estadístico, procedimientos o acciones que posibiliten la investigación de un conjunto de datos reales, con el fin de obtener diferentes tipos de modelos apropiados para la modelización del crecimiento microbiano, tanto modelos empíricos como modelos pseudo-mecanicistas.
- c) Hacer uso de algunos programas de acceso libre en Internet ya concebidos específicamente para tratar con modelos pseudo-mecanicistas en este contexto de crecimiento de poblaciones formadas por microorganismos.

## Material y métodos

Los participantes en este estudio fueron un grupo de 50 estudiantes de tercer año de una titulación en el ámbito de la Ingeniería de Biosistemas de la Universitat Politècnica de Catalunya (UPC), en la Escuela Superior de Agricultura de Barcelona.

El currículum docente que estos estudiantes ya han cursado previamente en llegar a este tercer curso está conformado, entre otras, con las siguientes asignaturas obligatorias: Matemáticas I y II, Física I y II, Química I y II, Biología General, Microbiología y Estadística. Esta preparación previa garantiza un buen conocimiento de algunos sistemas biológicos y de un conjunto de herramientas e ideas matemáticas básicas para abordar la modelización. En el contexto de Matemáticas I (primer semestre) este grupo de estudiantes estudió y se entrenó con contenidos del ámbito del álgebra, la geometría y el cálculo diferencial, y en el contexto de Matemáticas II (segundo semestre) con el cálculo integral, la resolución analítica y numérica de ecuaciones diferenciales ordinarias básicas, y en el reconocimiento de algunos modelos computacionales [7]. En el contexto de Estadística los estudiantes tuvieron múltiples posibilidades de practicar con la descripción numérica y gráfica de conjuntos de datos y de utilizar diversos métodos básicos de inferencia estadística.

Las actividades diseñadas, y que se implementaron en el aula, se llevaron a cabo en el contexto de la asignatura obligatoria "Programación y resolución de problemas a la ingeniería" que se cursa en el sexto semestre del grado de Ingeniería de Sistemas Biológicos de la UPC.

La Tabla 1 muestra el conjunto de datos que fueron analizados y utilizados para la construcción de modelos empíricos y modelos pseudo-mecanicista. La colección de tareas secuenciales, agrupadas en distintos módulos, para guiar la descripción de estos datos se proporcionó y gestionó a través del campus virtual Atenea, la plataforma de soporte a la docencia de la UPC, durante tres sesiones consecutivas de dos horas de duración cada una en aula informática. Cada estudiante tenía un ordenador con acceso a una hoja de cálculo (Excel), un programa matemático utilizado en las dos asignaturas de matemáticas previas, Matemáticas I y II (Maple) y un programa estadístico utilizado previamente en la asignatura de Estadística (R), junto con una conexión a Internet por si era necesario buscar información específica. Las actividades fueron diseñadas para ser llevadas a cabo de forma individual. Sin embargo, en sesiones de laboratorio y con grupos reducidos de alumnos, comentarios, sugerencias, e interacciones por parte del profesor se llevaron a cabo siempre que se consideró necesario o conveniente. A la vez no se prohibió expresamente que los estudiantes pudieran hablar e interactuar entre ellos durante las sesiones de trabajo.

Las respuestas de los estudiantes con respecto a los análisis y las diferentes modelizaciones del crecimiento microbiano fueron elaboradas individualmente y recolectadas a través de documentos tipo texto dando respuestas a cuestiones formuladas, hojas de cálculo comentadas, y salidas de programas matemáticos y estadísticos insertados en un texto explicativo. También entrevistas o conversaciones personales durante el desarrollo de las

sesiones en el aula informática permitieron recoger información sobre el desarrollo de la actividad.

**Tabla 1.** Evolución temporal de una población microbiana que crece en un medio líquido (1 mL) con una cantidad inicial de azúcar y sin ninguna otra adición de nutrientes durante el período de tiempo controlado. Los datos se obtuvieron mediante un método indirecto basado en la densidad óptica. El experimento lo llevó a cabo un estudiante en el laboratorio de microbiología de la Escola Superior d'Agricultura de Barcelona (UPC) y formaba parte de su Trabajo Final de Grado (*TFG de Ingeniería Alimentaria: "Efectos de las concentraciones de glucosa y etanol iniciales en el crecimiento de Saccharomyces cerevisiae"*. J. Cañadillas (2014) Tutoras: R. Carbo y M. Ginovart).

Observación	Tiempo (horas)	Numero de microorganismos
1	0	145349
2	3	146217
3	6	139333
4	9	143620
5	12	168557
6	15	287768
7	18	972270
8	21	2996236
9	24	4444266
10	27	5953756
11	30	7245644
12	33	7614686
13	36	8187928
14	39	10214427
15	42	11842517
16	45	13650985
17	48	12837014

## **Conjunto de tareas para investigar el crecimiento de la población y construir diferentes tipos de modelos empíricos y modelos pseudo-mecanicistas**

A continuación se detallan las distintas partes o módulos de la actividad diseñada que aglutinan las diferentes tareas propuestas a los estudiantes para investigar, describir y modelizar el crecimiento microbiano reportado en la Tabla 1. El formato de presentación de estas tareas pretende reflejar fielmente la documentación que se fue entregando a los estudiantes y que ellos utilizaron en el aula durante el desarrollo de la actividad.

### **Cuestiones previas**

Estas cuestiones preliminares se deberían responder sin leer nada más del guion de esta actividad, con el único propósito de que reflexiones sobre lo que has estudiado y aprendido hasta ahora, y sobre lo que serías capaz de aplicar en este tipo de problema. Debemos investigar y describir el conjunto de datos de la Tabla 1, ¿Qué estrategias o recursos de los que ya conoces podrías utilizar para analizar y describir estos datos? ¿Qué tipo de función o modelo podrías ajustar y que programas podrías utilizar para ello?

## Parte A: Representación gráfica

Lleva a cabo un estudio exploratorio de los datos y decide la mejor forma de representarlos. ¿Es mejor trabajar con los datos originales o bien se puede imaginar alguna transformación lineal o no lineal para estos datos? ¿Las magnitudes de los valores observados en el inicio del crecimiento son muy distintas de las que se obtienen al finalizar el experimento? ¿Qué función matemática permite modificar o modular este tipo de comportamiento?

## Parte B: ¿Crecimiento lineal? ¿Crecimiento exponencial?

¿Se observa un crecimiento lineal para esta población? Justifica la respuesta. ¿Se observa un crecimiento exponencial de microorganismos? Justifica la respuesta.

## Parte C: Modelo empírico y funciones polinómicas

Utiliza funciones polinómicas para describir esta evolución temporal. ¿Qué tipos de polinomios permiten tener un buen ajuste para estos datos? Para empezar prueba con un polinomio de grado dos. ¿Qué se observa? ¿Crees que es posible mejorar este ajuste? Prueba con polinomios de grado 3, grado 4 y grado 5. ¿Qué se puede decir acerca de la calidad de estos ajustes? ¿Qué información puedes extraer del valor de los parámetros implicados en este tipo de modelo empírico que son los coeficientes de los polinomios? ¿Qué ventajas y desventajas tiene hacer uso de funciones polinómicas para ajustar o describir este tipo de crecimiento poblacional?

## Parte D: Fases o etapas en la evolución temporal del crecimiento poblacional

Utilizaremos a partir de ahora el conjunto de datos que se obtiene aplicando la transformación logarítmica en base 10. Pasamos de tener  $N_t$ , que es el número de microorganismos correspondientes a la observación  $t$ , a trabajar con  $\text{LOG}_{10}(N_t)$ .

**Comentario:** Para esta parte de la actividad los estudiantes deberían utilizar los conocimientos adquiridos en microbiología y conectarlos con conceptos matemáticos. En poblaciones microbianas en cultivo cerrado se puede identificar una sucesión de fases o etapas caracterizadas por las variaciones de la tasa (o velocidad) de crecimiento: i) fase de latencia o “lag phase” con una tasa de crecimiento nula, ii) fase de aceleración con un aumento progresivo de esta tasa de crecimiento, iii) fase exponencial o “log phase” con una tasa de crecimiento positiva, constante y máxima, iv) fase de retraso con la disminución progresiva de la tasa de crecimiento, v) fase estacionaria con una tasa de crecimiento (aproximadamente) nula, y vi) la fase de declive o muerte con una tasa de crecimiento negativa.

Representa gráficamente este conjunto de datos transformados por el logaritmo en base 10. ¿Qué se observa? Construye la gráfica correspondiente a la variación del crecimiento para esta población, es decir en cada instante temporal  $t$  representa  $\text{LOG}_{10}(N_t) - \text{LOG}_{10}(N_{t+1})$ . ¿Puedes identificar fases o etapas distintas en esta evolución temporal atendiendo a la variación del crecimiento? ¿De forma aproximada, que etapas crees que puedes caracterizar fácilmente con los datos disponibles combinando la información que proporciona la curva del crecimiento con la gráfica de las variaciones en el crecimiento?

## Parte E: Modelo pseudo-mecanicista y funciones lineales a trozos

Utiliza funciones lineales o rectas para describir de forma aproximada las tres fases principales de una curva de crecimiento poblacional como la que tenemos: fase de latencia, fase exponencial y fase estacionaria. ¿Crees que estas tres rectas ajustadas a cada una de estas tres fases pueden dar una respuesta plausible a la descripción de estos datos? Contrasta y compara gráficamente los datos observados y los datos simulados con el modelo lineal a trozos que has construido. ¿Qué puedes comentar? ¿Puedes asignar significado biológico a los parámetros implicados en esta descripción matemática de la evolución de la población?

¿Cuáles son los parámetros con significado biológico que has obtenido para este modelo pseudo-mecanicista? ¿Este tipo de aproximación que has desarrollado de esta forma genera siempre una única respuesta posible y plausible? ¿Por qué?

Un trabajo de Buchanan y otros autores [2] en el ámbito de la microbiología predictiva propuso hace unos años una función simple para modelizar estas curvas de crecimiento microbiano en cultivo cerrado mediante el uso de funciones lineales o rectes para describir las tres fases del crecimiento: latencia, exponencial y estacionaria. Los parámetros implicados en ese modelo de Buchanan son  $\text{LOG}_{10}N_0$ ,  $\mu_{\text{max}}$  (o velocidad máxima de crecimiento específico), lag (o duración de la fase de latencia) y  $\text{LOG}_{10}N_{\text{max}}$ . Veremos más adelante, en otros apartados de esta actividad, como se puede trabajar con este tipo de modelo pseudo-mecanicista con el programa estadístico R y que información complementaria y extra se puede obtener con él respecto al ajuste que has realizado en la hoja de cálculo.

## Parte F: Construcción del modelo logístico discreto en una hoja de cálculo

Teniendo en cuenta que el número de microorganismos ahora está expresado en unidades logarítmicas en base 10 (aunque también podríamos considerar los datos originales), definiremos la variable  $M_t$ , para simplificar la notación, como el  $\text{LOG}_{10}(N_t)$ , siendo  $N_t$  el tamaño de la población correspondiente a la observación  $t$ . Vamos, a partir de los valores  $M_t$ , a trabajar en la construcción paso a paso de un modelo logístico discreto [5].

Esta nueva manera de investigar y describir estos datos se basa en una perspectiva diferente a la utilizada hasta este momento y nos conducirá a modelos que hacen uso de funciones sigmoideas, funciones matemáticas que tiene una forma de "S" y para las que su función derivada tiene aproximadamente forma de campana de Gauss. ¿Reconoces en los datos que has representado anteriormente estas características?

Para la construcción del modelo logístico discreto, usaremos el conjunto de valores  $M_t = \text{Log}_{10}(N_t)$  y regresiones lineales sobre estos valores previamente operados. Un buen inicio para la construcción de este tipo de modelo discreto es fijar la atención en la siguiente ecuación:

$$\text{valor futuro} = \text{valor actual} + \text{cambio} ,$$

la cual, utilizando la variable que tenemos definida, se escribe como

$$M_{t+1} = M_t + \Delta M_t .$$

El propósito de este proceso iterativo es, paso a paso y desde un valor inicial  $M_0$ , encontrar una aproximación para este  $\Delta M_t$  que pueda reproducir los datos de manera razonable. Por lo tanto, la cuestión es conocer o intuir como podemos proponer algún ajuste apropiado para este cambio o incremento. Considerando que

$$\Delta M_t = M_{t+1} - M_t ,$$

revisa las gráficas de las observaciones y de las variaciones de crecimiento que ya habías construido anteriormente en la hoja de cálculo. ¿Qué observas? Se puede intuir que hay un crecimiento exponencial al inicio de la evolución (crecimiento rápido), pero que a medida que nos vamos acercando a un máximo  $K$  (capacidad máxima que puede soportar el sistema) este crecimiento se frena o ralentiza, y por tanto el crecimiento disminuye. Con esta idea se puede ensayar una expresión como:

$$\Delta M_t = M_{t+1} - M_t = r M_t (K - M_t) .$$

Para explorar cómo se pueden encontrar los valores de  $r$  y  $K$ , representa en un diagrama de dispersión estos datos, situando en el eje de ordenadas  $M_{t+1} - M_t$  y en el eje de abscisas  $M_t (K - M_t)$ . Utiliza una regresión lineal con ordenada en el origen igual a cero sobre esos puntos para encontrar el valor de  $r$ . Para ello puedes ensayar diferentes valores numéricos para la constante  $K$ , y ver que valores se van obteniendo para la constante  $r$  (o pendiente de la recta ajustada en ese diagrama de dispersión). Una vez consideres que has probado suficientes valores distintos y que has constatado que algunos valores dan mejores resultados que otros, escoge los valores que consideres que mejor pueden funcionar para construir a partir de estos datos la expresión:

$$M_{t+1} = M_t + r M_t (K - M_t) .$$

Con este modelo discreto construido, que se conoce como modelo logístico discreto, vamos a construir una serie de valores simulados que compararemos con los valores observados. Los valores simulados empezaran con  $M_0$ , el primer valor observado, y a partir de aquí vamos obteniendo valores sucesivos:

$$M_1 = M_0 + r M_0 (K - M_0)$$

$$M_2 = M_1 + r M_1 (K - M_1)$$

$$M_3 = M_2 + r M_2 (K - M_2)$$

y así, repetidamente, hasta construir la serie simulada completa.

Representa gráficamente los valores observados y los valores simulados. ¿Qué puedes decir? ¿Cómo puedes valorar la bondad de ajuste de los diferentes modelos que se pueden ir obteniendo con valores distintos para  $r$  y  $K$ ? ¿Qué crees que deberías hacer para encontrar el mejor de los modelos posibles de este tipo?

## Parte G: Ajuste del modelo logístico continuo en una hoja de cálculo

Existe una familia de modelos matemáticos que hacen uso de funciones sigmoideas, siendo el modelo logístico (continuo) uno de ellos. El modelo llamado de crecimiento logístico, fue introducido por Pierre François Verhulst en 1838 y supone que la razón o tasa de crecimiento es proporcional conjuntamente tanto a la población misma como a la cantidad faltante para llegar a la máxima población sustentable por el medio en el cual la población subsiste. Con esto, definimos dicho modelo, siendo  $N=N(t)$  el número de individuos de una población, como la solución de la siguiente ecuación diferencial ordinaria:

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

donde  $r$  se define como la tasa de crecimiento intrínseco, y  $K$  es la capacidad sustentable o sostenible (capacidad de carga) que es el máximo valor que puede alcanzar  $N$ . Para  $N \ll K$ ,  $\frac{N}{K}$  es próximo a cero, y por tanto  $\left(1 - \frac{N}{K}\right)$  próximo a 1, con lo que se tiene que  $\frac{dN}{dt} \cong rN$  y por tanto el crecimiento es del tipo que da el modelo Malthus o modelo exponencial, mientras que, cuando  $N$  tiende a  $K$ , es decir la población  $N$  se encuentra muy próxima a su capacidad de carga,  $\frac{N}{K} \cong 1$  y por tanto  $\left(1 - \frac{N}{K}\right)$  próximo a 0, por lo que entonces  $\frac{dN}{dt} \cong 0$ .

Reescribe la expresión utilizada en la tarea anterior  $M_{t+1} = M_t + r M_t (K - M_t)$  en la forma

$$M_{t+1} = aM_t (1 - M_t/b).$$

¿Qué valor tienen  $a$  y  $b$  en función de  $r$  y  $K$ ? Con esta reescritura de la expresión utilizada anteriormente, se constata que la expresión encontrada en el modelo discreto se corresponde con el patrón o estructura que tiene la expresión en el modelo continuo, por lo que ambas se conocen con el nombre de modelo logístico (discreto o continuo).

Supongamos ahora que  $x$  representa el tamaño de la población y  $t$  representa el tiempo, este modelo logístico queda formalizado por la solución de la ecuación diferencial ordinaria que ya conoces:

$$x'(t) = rx(t) \left(1 - \frac{x(t)}{K}\right)$$

$$\frac{dx}{dt} = rx \left(1 - \frac{x}{K}\right)$$

donde la constante  $r$  define la tasa de crecimiento y  $K$  es la capacidad de carga

Resuelve esta ecuación diferencial ordinaria con la ayuda del programa matemático Maple (la resolución analítica, a mano, ya se ejercitó en la asignatura de Matemáticas II) y comprueba que hay maneras distintas para escribir su solución, como por ejemplo

$$x(t) = \frac{K}{1 + Ce^{-rt}}$$

o bien como

$$x(t) = \frac{Kx_0e^{rt}}{K + x_0(e^{rt} - 1)}$$

Representa gráficamente alguna de estas soluciones fijando valores para la población inicial  $x_0$  y los parámetros  $r$  y  $K$ . ¿Qué efecto tienen estos parámetros sobre la forma de la curva representada?

Volvemos a los datos de la Tabla 1 originales sin transformar por el logaritmo (aunque también podríamos considerar los datos originales) y vamos a ajustar una función logística con la ayuda del *Solver*, un complemento de Excel útil para problemas de optimización. Encontrareis el *Solver* en la opción de “Datos” del Excel, y en caso de que no esté activado se puede activar como un complemento más de la hoja de cálculo. Para hacer uso del *Solver* trabajaremos con la columna del tiempo, los datos originales (el número de microorganismos), y construiremos una columna con los datos generados por la función o modelo logístico que queremos obtener, dando para empezar unos valores “preliminares” a los parámetros implicados en el modelo. Ajustar esta función o modelo a los datos experimentales sería conseguir los mejores valores posibles para los parámetros involucrados, es decir los que minimizaran el error o discrepancia entre valores observados y valores predichos por el modelo. Para ello añadiremos una columna más con el valor absoluto de la diferencia entre valor observado y valor predicho por la función logística y calcularemos la suma de todas estas discrepancias. En un rango de celdas de la hoja de cálculo situaremos el valor de los parámetros que recoge la función, el cual se podrá ir variando, hasta conseguir una combinación de valores para los parámetros que minimice la suma de las discrepancias, lo cual, es precisamente, lo que nos proporciona la opción *Solver* de la hoja de cálculo en este contexto, la solución aproximada e iterada de un problema de minimización. Fija valores preliminares para estos parámetros  $r$  y  $K$  que permitan obtener una primera aproximación al conjunto de datos, lo cual puede ser valorado de forma gráfica con una representación conjunta de valores observados y valores predichos. Prueba diversas opciones de parámetros antes de proceder al uso del *Solver* ¿Qué obtienes? Con el uso del *Solver*, ¿qué valores se consiguen para los parámetros implicados? ¿Cómo valoras este ajuste?

## Parte H: Función de Gompertz y su reparametrización para obtener un modelo pseudo-mecanicista

Existen otras funciones sigmoideas que se presentan por medio de sus expresiones matemáticas y sus representaciones gráficas, o como soluciones de otras ecuaciones diferenciales. Por ejemplo, la solución de la siguiente ecuación diferencial ordinaria

$$y'(t) = \frac{dy}{dt} = r \cdot y(t) \cdot \left( \ln \left( \frac{K}{y(t)} \right) \right)$$

es la función de Gompertz, la cual incluye los mismos parámetros biológicos que la ecuación logística (o de Verhulst): la tasa intrínseca de crecimiento  $r$  y la capacidad de carga  $K$ , e intenta describir el mismo tipo de evolución temporal. Se observa que cuando  $y(t)$  está próximo a  $K$ , el cociente  $K/y(t)$  es próximo a 1, con lo que el  $\ln$  de valores cercanos a 1 es próximo a 0.

Encuentra la solución de esta ecuación diferencial con la ayuda del programa matemático Maple, y comprueba que se obtiene una función de la forma:

$$y(t) = a \cdot \exp(-\exp(b - c \cdot t)) = ae^{-e^{b-c \cdot t}}$$

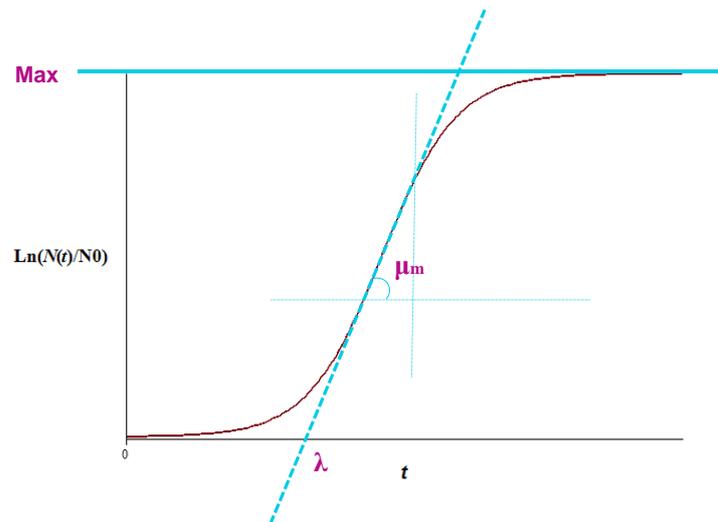
donde  $a$ ,  $b$  y  $c$  son constantes o parámetros. Esta función de Gompertz que has obtenido es una función de la familia de funciones sigmoideas, al igual que lo era la función logística que obtuviste anteriormente.

Representa gráficamente esta función Gompertz para valores diferentes de  $a$ ,  $b$  y  $c$  para constatar que tiene la forma de “S”. Su forma, parecida a la función logística, justifica que sea otro de los modelos que se pueda utilizar para describir y representar el crecimiento microbiano que nos ocupa.

Como ya sabemos que los microorganismos pueden crecer de forma exponencial durante un periodo de tiempo, es útil la transformación logaritmo sobre el tamaño de la población ( $N=N(t)$ ). Ahora, vamos a trabajar con el logaritmo natural (o logaritmo neperiano) del tamaño relativo de la población, es decir, vamos a considerar una nueva variable, una transformación de  $N(t)$ :

$$y(t) = \text{Ln} \left( \frac{N(t)}{N_0} \right)$$

Esto significa que tenemos una curva genérica del tipo que se encuentra representada en la Figura 1.



**Figura 1.** Curva genérica de crecimiento con la identificación de parámetros con significado biológico: la velocidad específica máxima de crecimiento ( $\mu_m$ ), el tiempo de latencia ( $\lambda$  o "lag"), y el valor final o máximo posible que se puede alcanzar.

Las tres fases principales de la curva de crecimiento pueden ser descritas por tres parámetros:

- 1) La velocidad máxima de crecimiento específico ( $\mu_m$ ) que se define como la tangente en el punto de inflexión de la curva.
- 2) El tiempo de latencia ( $\lambda$ ) o "lag" que se define como el valor de la intersección de esta tangente con el eje de abscisas.
- 3) La asíntota determinada por el valor Max, el máximo valor que puede alcanzar la población.

Un estudio de la primera derivada y segunda derivada de la función nos permite identificar el rol que juegan los parámetros que intervienen en sus definiciones. La identificación del significado de los parámetros de estas funciones de Gompertz, seguida de una conveniente reparametrización, nos permiten disponer de parámetros con un significado biológico conectado en el contexto de crecimiento microbiano que tenemos, y por tanto obtener un modelo pseudo-mecanicista para este contexto [12].

El propósito de esta tarea es reescribir, con la ayuda del programa Maple, la función o expresión de Gompertz substituyendo los parámetros  $a$ ,  $b$ , y  $c$  dados en la expresión anterior que no tienen significado biológico, por las constantes  $\mu_m$ ,  $\lambda$ , y Max que sí que tienen un significado biológico. Considerando que sabrías resolver a mano (con lápiz y papel) esta tarea, pero que ahora el propósito es agilizar los cálculos implicados en el proceso evitando errores de transcripción o simplificación, podemos utilizar la ayuda que nos proporcionan programas matemáticos como el Maple para la resolución de este tipo de tarea. Esquematiza e implementa en tu hoja de trabajo los pasos a realizar en Maple para reparametrizar la función de Gompertz y obtener el modelo de crecimiento de Gompertz con parámetros biológicos.

## Parte I: Modelos pseudo-mecanicistas de crecimiento microbiano y R

Para finalizar esta actividad se exploraran las opciones que proporcionan el programa estadístico R y sus paquetes que se han ido diseñando y publicando, accesibles desde Internet.

En los últimos años se han ido desarrollado paquetes o programas específicos para tratar con la modelización de curvas biológicas, y en particular para utilizar en crecimientos microbianos (<http://cran.r-project.org/web/packages/nlsMicrobio/nlsMicrobio.pdf>). Dentro del proyecto R existe el paquete *nlsmicrobio* desarrollado por un grupo de científicos que trabajan en el campo de la microbiología predictiva que permite el ajuste de curvas experimentales en este contexto de aplicación. Entre los diferentes modelos que se pueden encontrar para modelizar el crecimiento de una población microbiana, se encuentra el modelo Buchanan [2] y el modelo de Gompertz modificado introducido por Gibson y otros autores [6] y reparametrizado por Zwietering y otros autores [12]. Estos dos modelos trabajan con la evolución del logaritmo decimal del número de microorganismos (LOG<sub>10</sub>N) en función del tiempo.

Revisa y utiliza el programa que encontraras en el campus virtual Atenea que te facilitará el uso de R y de este paquete *nlsmicrobio* para analizar y ajustar los datos de la Tabla 1 con los dos modelos que ya conoces, el modelo de Buchanan y el de Gompertz reparametrizado. ¿Qué información de salida proporciona este programa R cuando se ejecuta con los datos de la Tabla 1? ¿Qué información complementaria aporta este tipo de programa respecto a lo que ya has conseguido previamente en las tareas anteriores?

Si comparas los resultados del modelo lineal a tres fases que obtuviste con la hoja de cálculo Excel en la primera parte de la actividad con los resultados que proporciona el programa R para el modelo Buchanan, es razonable pensar que los valores obtenidos en las dos resoluciones deberían ser similares. ¿Es realmente así? Si utilizas la información de salida de R que nos da la estimación puntual de los parámetros y la estimación por intervalos de confianza al 95%, ¿qué puedes comentar? ¿Hay algún valor obtenido para el modelo lineal a tres fases con la hoja de cálculo que no se corresponda de forma directa con la información que proporciona el programa R? ¿Cuál es? Razona o justifica el porqué de esta discordancia.

En relación a la salida que proporciona el programa R con relación al modelo Gompertz modificado, es decir la función de Gompertz reparametrizada con parámetros con significado biológico, ¿qué puedes comentar? Una comparación directa de las dos expresiones, la obtenida por ti en la tarea anterior con el programa Maple y la que nos da el paquete *nlsmicrobio* de R, ¿nos permite afirmar de forma inmediata que se tratan de la misma función? Demuestra que la expresión que hemos obtenido en la tarea anterior se corresponde con la que nos proporciona la salida del programa R si se considera que  $y(t) = Ln\left(\frac{N(t)}{N_0}\right)$  y que hay que aplicar el cambio de base entre logaritmos.

Desde tu perspectiva, ¿qué ventajas tiene esta modelización con R del conjunto de datos de la Tabla 1 con respecto a las aproximaciones desarrolladas en las anteriores tareas? ¿Tiene alguna desventaja?, en caso afirmativo, ¿cuál o cuáles?

Si tuvieras que analizar otro conjunto de datos de crecimiento microbiano, a las preguntas que se formularon inicialmente, antes de empezar a desplegar esta actividad:

*“¿Qué estrategias o recursos de los que ya conoces podrías utilizar para analizar y describir estos datos?” “¿Qué tipo de función o modelo podrías ajustar y que programas podrías utilizar para ello?”*,

ahora, después de haber desplegado las distintas partes de esta actividad, ¿cuál sería tu prioridad u orden en el conjunto de tareas a realizar para dar una respuesta a la modelización de estos nuevos datos?

## Resultados y discusión

Esta actividad subdividida en un conjunto de tareas agrupadas en nueve partes distinguibles se llevó a cabo en el aula informática durante tres sesiones de dos horas de duración cada una, por lo que la extensión y diversidad de los resultados conseguidos fue notable.

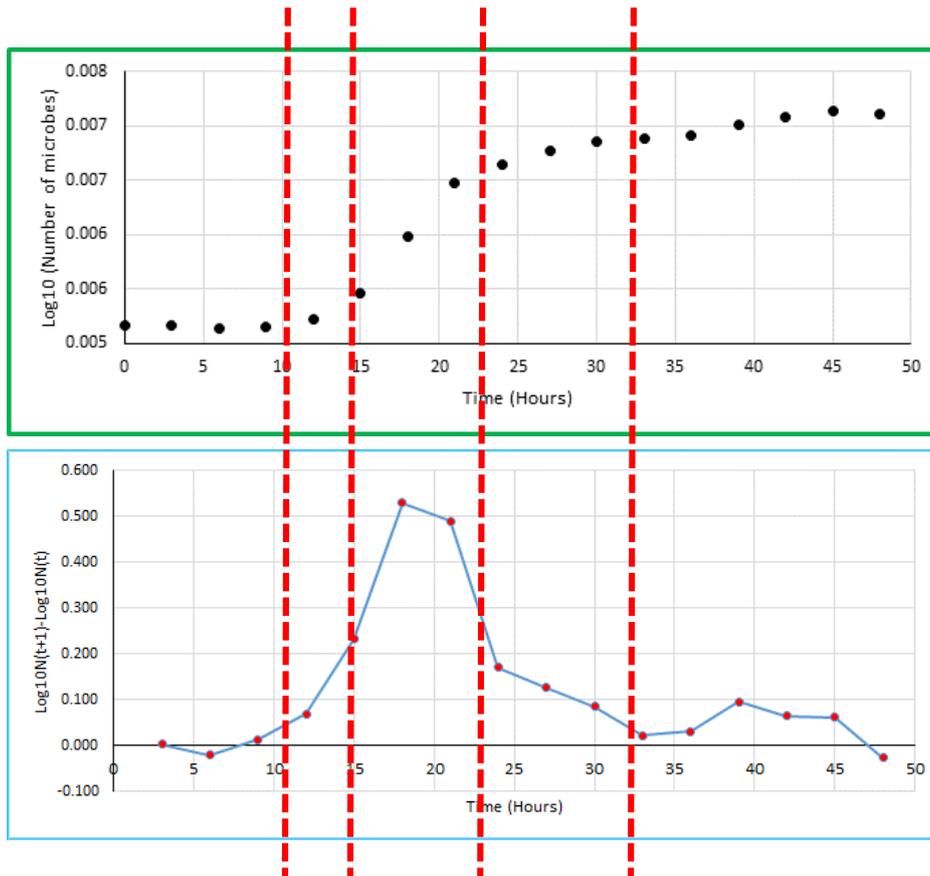
Representaciones gráficas de la evolución temporal del número de individuos de la población y de las transformaciones realizadas con los datos, definiciones y manipulaciones de funciones matemáticas para construir y formular distintos modelos, conjuntamente con el cálculo de los parámetros correspondientes implicados en estos modelos, así la asignación de significado biológico para la población microbiana en el caso de los modelos de pseudo-mecanicistas, configuraron el conjunto de resultados que los estudiantes obtuvieron, analizaron y discutieron, en conexión con los conocimientos de biología que tenían. Únicamente se han escogido algunos de los resultados generados a partir de que este conjunto de tareas para presentar y discutir a continuación en este apartado.

Los estudiantes llevaron a cabo un estudio exploratorio de los datos y decidieron cuál era la mejor forma de representarlos considerando la naturaleza de las observaciones experimentales que se manejaban. Los datos originales que se tenían al comienzo de la evolución temporal eran valores de orden  $10^5$  y al finalizar la evolución temporal se alcanzaron valores de  $10^7$ , con lo que la magnitud del rango que se debía representar era considerable. Los incrementos (como valores absolutos) de un tiempo de muestro a otro tiempo de muestreo eran también valores de magnitudes muy distintas, dependiendo del momento temporal de la evolución. La representación gráfica de los datos originales en diferentes escalas y a través de algunas transformaciones no lineales fueron opciones inspeccionadas por los estudiantes. Teniendo en cuenta que en el contexto de la microbiología es frecuente que las poblaciones microbianas se expresen habitualmente en unidades logarítmicas (en base 10), las transformaciones con logaritmos en distintas bases fueron las estrategias más ensayadas por los estudiantes en la hoja de cálculo.

Ajustes con polinomios de distintos grados para describir el conjunto de datos originales que procura la misma hoja de cálculo fueron examinados y comentados por los estudiantes, discutiendo alguna ventaja sobre este método y enfatizando las desventajas que tenía este enfoque o aproximación, donde el concepto y propósito de lo que debe ser un modelo no quedaba bien simbolizado. Se constató lo que significaba trabajar con modelos de tipo empírico y cuáles eran sus limitaciones, y se valoró muy positivamente lo que podría llegar a ofrecer un modelo pseudo-mecanicista en ese contexto. Los coeficientes de los polinomios ajustados no mostraban ni podían incorporar ninguna de las ideas que se encuentran tras el fenómeno que pretendemos describir, una población que llega a un entorno o espacio y que utiliza los recursos energéticos que encuentra, posibilitando la reproducción de los individuos, aumentando el tamaño de la población y gastando los nutrientes del entorno que no van a ser repuestos, hasta llegar a una situación no favorable en la que ya no es posible que la población siga creciendo. A la incapacidad de estos modelos de poder explicar el fenómeno biológico, se añadió el hecho que algunos de los polinomios ajustados alcanzaban valores negativos, perdiendo entonces todo sentido la expresión matemática para representar la evolución de la población.

La utilización de la transformación logaritmo decimal sobre el número de microorganismos observados a lo largo del tiempo junto con los incrementos observados en cada uno de los muestreos permitió la identificación de las diferentes fases de crecimiento que se suceden durante la evolución temporal del tamaño de la población como muestra la Figura 2.

Un primer acercamiento a la formulación de un modelo sencillo que pueda recoger las características o tendencias principales de este tipo de crecimiento pasó por trabajar únicamente con las tres fases más relevantes y principales de la curva, y utilizar funciones lineales a trozos para la descripción aproximada de este crecimiento. Los estudiantes estudiaron con la hoja de cálculo la utilización de líneas rectas para describir cada una de estas tres fases caracterizadas por variaciones significativas de la tasa de crecimiento: primero la fase de latencia con una tasa de crecimiento nula, después la fase exponencial con una tasa de crecimiento constante, y finalmente la fase estacionaria sin crecimiento (Figura 3).

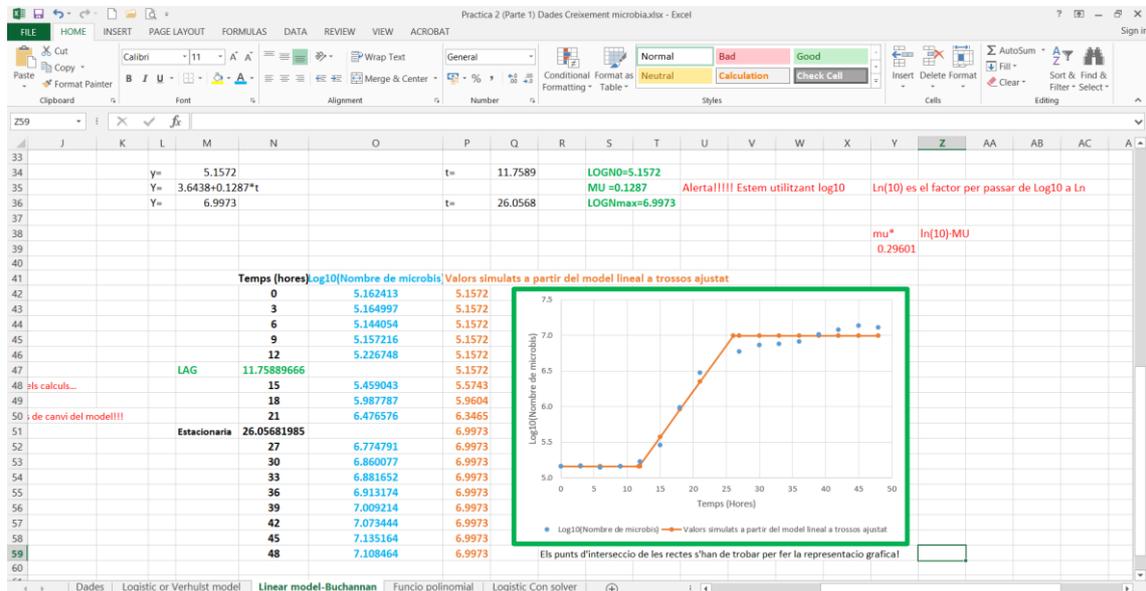


**Figura 2:** La combinación de la información que proporcionan los datos transformados previamente por la función logaritmo en base 10 y la de los incrementos correspondientes para cada uno de los instantes de tiempo muestreados permitieron detectar las distintas fases del crecimiento poblacional que se van sucediendo. Las rectas verticales indican de forma aproximada estos cambios de fase consecutivos: i) fase de latencia o “lag phase” con una tasa de crecimiento nula, ii) fase de aceleración con un aumento de esta tasa de crecimiento, iii) fase exponencial o “log phase” con una tasa de crecimiento positiva, constante y máxima, iv) fase de retardo o desaceleración con la disminución de la tasa de crecimiento, v) fase estacionaria con una tasa de crecimiento (aproximadamente) nula.

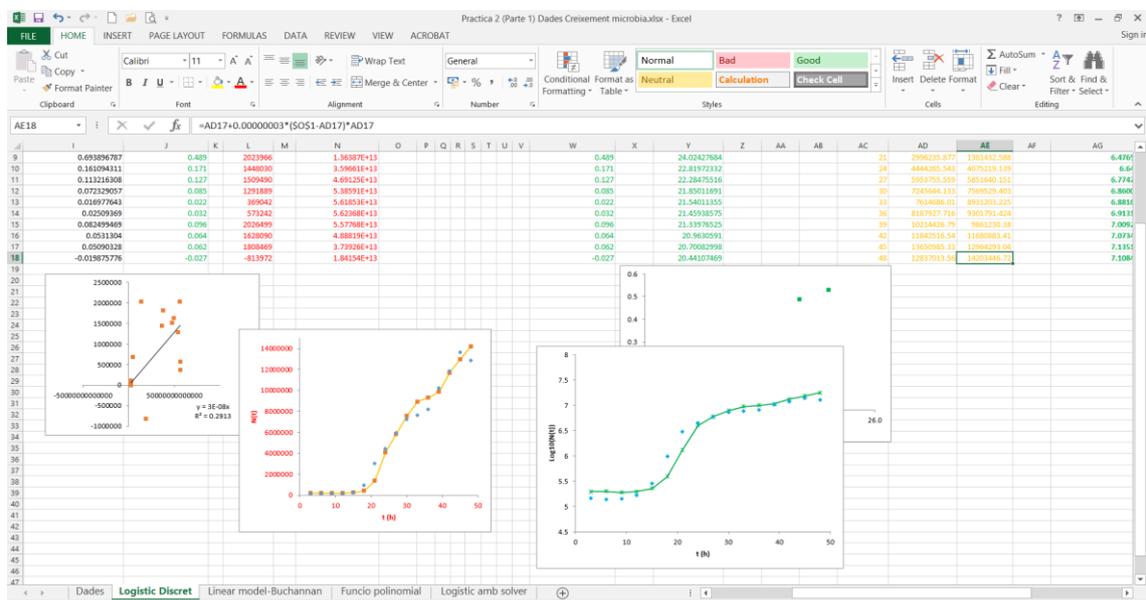
La lectura del documento complementario a la que los estudiantes pudieron acceder de Buchanan y otros autores [2] que lleva por título “*When is simple good enough: a comparison of the Gompertz, Baranyi, and three-phase linear models for fitting bacterial growth curves*”, ilustró y puso de manifiesto que el modelo lineal a trozos que ellos habían construido sintonizaba con una de las posibles respuestas que otros investigadores presentaban y aceptaban en el entorno de la microbiología predictiva. El modelo lineal de tres fases fue el primer modelo pseudo-mecanicista que se obtenía, pues en su formulación se reconocían los parámetros con significado biológico como son el logaritmo decimal de la población inicial y de la población final, la duración de la fase de latencia y la velocidad específica máxima de crecimiento. Los estudiantes pudieron reconocer el trabajo realizado por ellos en una hoja de cálculo como una opción inmediata y simple de modelización, y ubicarlo en un contexto más amplio, pues este artículo que se referenció presentaba además otros modelos de interés en la biotecnología y en la microbiología predictiva aplicada a la industria agroalimentaria.

Tanto la construcción paso a paso de la función logística discreta por medio de conjunto de cálculos y aproximaciones lineales sobre los datos manejados que permitió introducir la idea de una simulación discreta (Figura 4), como el uso y ajuste de la función logística continua con la ayuda del Solver de la hoja de cálculo Excel (Figura 5), posibilitó la introducción de las ideas y conceptos que subyacen detrás del modelo logístico, ya que se incidió en la naturaleza de las condiciones que lo generan o definen. Se constató que la función logística discreta permitía representar de forma razonable tanto los datos originales como los datos transformados por la

función logaritmo (Figura 4). En relación al uso de la función logística continua, se observó que con los datos transformados por el logaritmo decimal no fue posible encontrar un buen ajuste del modelo logístico continuo, mientras que cuando se trabajó con los datos originales el ajuste que se consiguió fue sensiblemente mejor (Figura 5).



**Figura 3:** Captura de pantalla de una hoja de cálculo con la implementación de las tareas que condujeron a la obtención de un modelo pseudo-mecanicista, el modelo lineal a trozos o modelo de Buchanan (Buchanan et al. 1997).



**Figura 4:** Captura de pantalla de una hoja de cálculo con la implementación de las tareas que condujeron a la simulación de datos con un modelo logístico discreto.

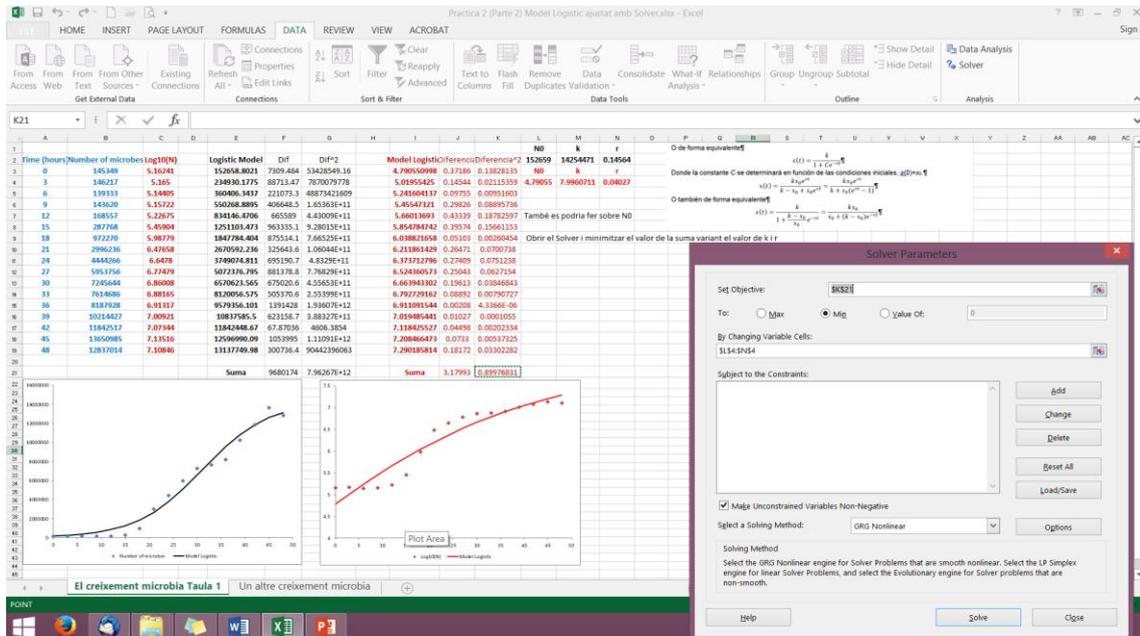


Figura 5: Captura de pantalla de una hoja de cálculo con la implementación de las tareas que condujeron al ajuste del modelo logístico continuo.

Las respuestas de los estudiantes a la parte de la actividad que hacía referencia al estudio de otra función sigmoidea distinta a la función logística, la función de Gompertz y su reparametrización para obtener un modelo pseudo-mecanicista, fue sorprendentemente interesante. Permitió que los estudiantes utilizaran conocimientos repetidos de forma incesante durante sus etapas anteriores (primera derivada, segunda derivada, punto de inflexión, pendiente de la recta tangente en un punto, asíntota,...) pero ahora en un contexto que daba sentido a todo aquello que habían aprendido, tanto en matemáticas como en microbiología. El propósito de esta parte era reescribir, con la ayuda del programa Maple, la función o expresión de Gompertz substituyendo los parámetros matemáticos que la definen,  $a$ ,  $b$ , y  $c$ , pero que no tienen significado biológico, por los parámetros biológicos,  $\mu_m$ ,  $\lambda$ , y  $Max$ , que sí que tienen un significado en el contexto del crecimiento poblacional (Figura 1). Una secuencia de pasos realizada con el programa Maple que permitía reescribir la función logística y obtener este modelo pseudo-mecanicista con los tres parámetros  $\mu_m$ ,  $\lambda$ , y  $Max$  se presenta a continuación:

$$>y(t) := a \cdot e^{-e^{(b-c \cdot t)}}$$

$$y := t \rightarrow a e^{-e^{b-ct}}$$

$$>\frac{d}{dt} y(t)$$

$$a c e^{-ct+b} e^{-e^{-ct+b}}$$

$$>\frac{d}{dt} \frac{d}{dt} y(t)$$

$$-a c^2 e^{-ct+b} e^{-e^{-ct+b}} + a c^2 (e^{-ct+b})^2 e^{-e^{-ct+b}}$$

$$>solve\left(\frac{d}{dt} \frac{d}{dt} y(t) = 0, t\right)$$

$$\frac{b}{c}$$

$$>eval\left(\frac{d}{dt} y(t), t = \frac{b}{c}\right)$$

$$a c e^{-1}$$

$$> mumax = a c e^{-1}$$

$$mumax = a c e^{-1}$$

$$> \text{solve}(mumax = a c e^{-1}, c)$$

$$\frac{mumax}{e^{-1} a}$$

$$> c := \frac{mumax}{e^{-1} a}$$

$$c := \frac{mumax}{e^{-1} a}$$

$$> 0 = y\left(\frac{b}{c}\right) + mumax \cdot \left(t - \frac{b}{c}\right)$$

$$0 = e^{-1} a + mumax \left(t - \frac{b e^{-1} a}{mumax}\right)$$

$$> \text{solve}\left(0 = y\left(\frac{b}{c}\right) + mumax \cdot \left(t - \frac{b}{c}\right), t\right)$$

$$\frac{e^{-1} a (b - 1)}{mumax}$$

$$> lag = \frac{e^{-1} a (b - 1)}{mumax}$$

$$lag = \frac{e^{-1} a (b - 1)}{mumax}$$

$$> \text{solve}\left(lag = \frac{e^{-1} a (b - 1)}{mumax}, b\right)$$

$$\frac{e^{-1} a + lag mumax}{e^{-1} a}$$

$$> b := \frac{e^{-1} a + lag mumax}{e^{-1} a}$$

$$b := \frac{e^{-1} a + lag mumax}{e^{-1} a}$$

$$> \text{simplify}\left(\frac{e^{-1} a + lag mumax}{e^{-1} a}\right)$$

$$\frac{lag mumax e + a}{a}$$

$$> b := \frac{lag mumax e + a}{a}$$

$$b := \frac{lag mumax e + a}{a}$$

$$> a := Max;$$

$$a := Max$$

$$> y(t)$$

$$Max e^{-e} \frac{lag mumax e + Max}{Max} - \frac{mumax t}{e^{-1} Max}$$

$$> \text{simplify}$$

$$> y(t) := Max e^{-e} \frac{-mumax (t - lag) e + Max}{Max}$$

Ahora considerando que  $y(t) = \text{Ln}\left(\frac{N(t)}{N_0}\right)$ ,  $Max = \text{Ln}\left(\frac{N_{Max}}{N_0}\right)$ ,  $mumax = \mu_{Max}$ ,  $lag = \lambda$ , tenemos:

$$\text{Ln}\left(\frac{N(t)}{N_0}\right) = \text{Ln}\left(\frac{N_{Max}}{N_0}\right) e^{-e^{\left(\frac{-\mu_{Max}(t-\lambda)e_{+1}}{\text{Ln}\left(\frac{N_{Max}}{N_0}\right)}\right)}}$$

$$\text{Ln}(N(t)) = \text{Ln}(N_0) + (\text{Ln}(N_{Max}) - \text{Ln}(N_0)) e^{-e^{\left(\frac{\mu_{Max}(\lambda-t)e}{\text{Ln}(N_{Max}) - \text{Ln}(N_0)} + 1\right)}}$$

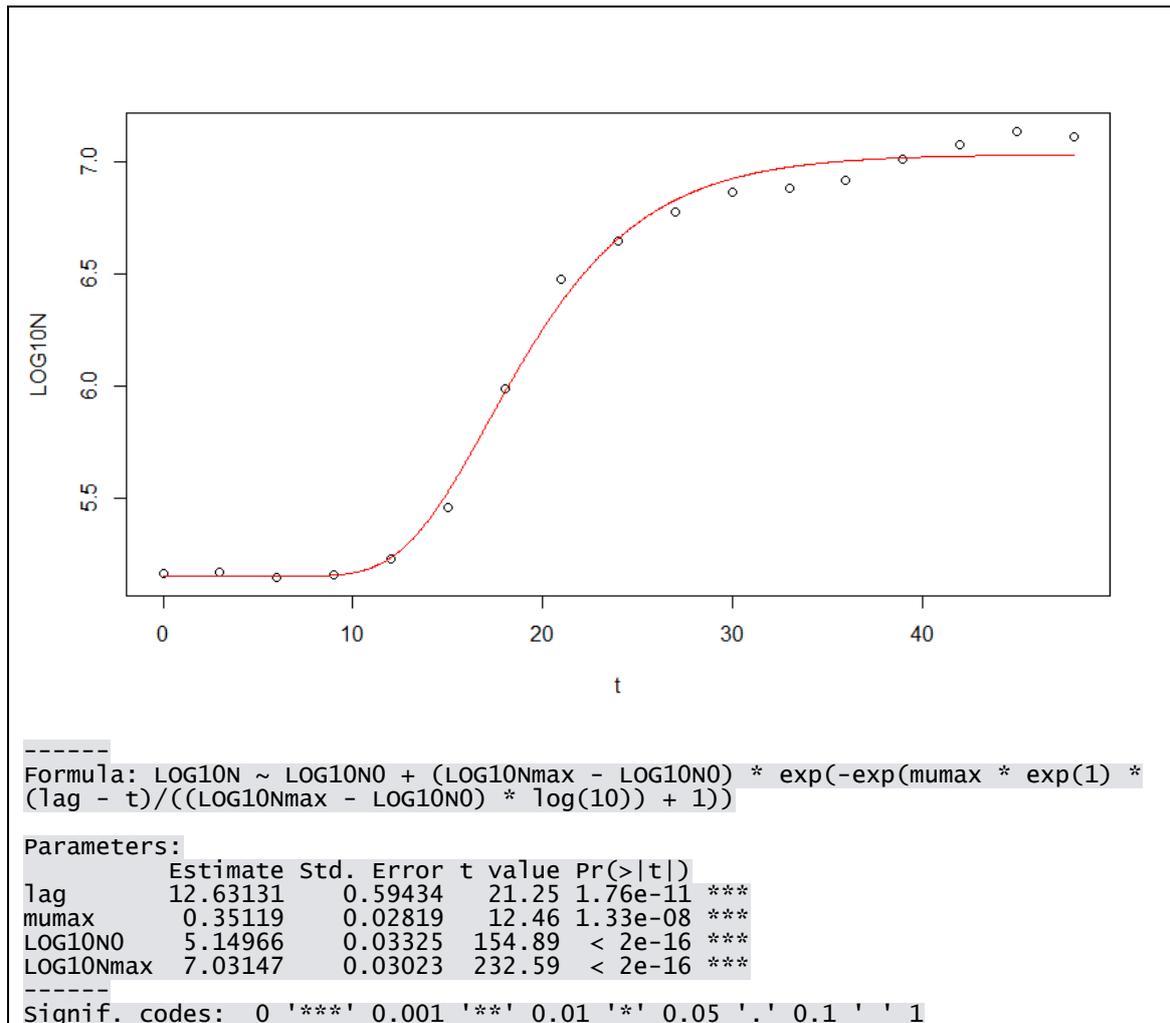
expresión que es la formulación de un modelo pseudo-mecanicista para el crecimiento poblacional a partir de la función de Gompertz.

En relación a la utilización del paquete *nlsmicrobio* del programa R, y sin poder entrar a considerar todos los detalles y resultados de esta parte de la actividad, indicar únicamente que una parte de la salida que proporcionó R, y que los estudiantes discutieron, se presenta en la Figura 6, con la representación gráfica del ajuste de los datos por la función Gompertz reparametrizada y la estimación puntual de los valores de los parámetros implicados en su formulación. Además se obtuvieron otros resultados de tipo estadístico que completaron el estudio de este ajuste. Realizando previamente el cambio de base en los logaritmos implicados, los estudiantes identificaron la función de Gompertz reparametrizada que aparece en la salida de R y que se corresponde con la función que obtuvieron previamente con la ayuda del Maple. En esa tarea se trabajó con el logaritmo natural (o neperiano) y el ajuste realizado por R se efectúa con los datos originales transformados previamente con el logaritmo decimal ( $\text{Ln}(x) = \text{Log}_{10}(x) \cdot \text{Ln}(10)$ ), lo que permitió una vez más la aplicación de conocimientos previos. Además de esta salida relativa al modelo Gompertz, R también proporcionó otros resultados relativos al ajuste con el modelo lineal a trozos (o modelo Buchanan), el otro modelo ajustado en la hoja de cálculo por los estudiantes en el inicio de la actividad. Las salidas de tipo estadístico (intervalos de confianza, análisis de residuales,...) también fueron comentadas a la luz de los conocimientos previos que tenían los estudiantes procedentes de la asignatura de estadística cursada con anterioridad.

En relación a la valoración de esta actividad por parte de los estudiantes, los comentarios proporcionados fueron todos positivos en el sentido que habían percibido un escenario de aprendizaje de relevancia en el mundo real, una modelización que se utiliza en el ámbito de la microbiología predictiva. Se destacó positivamente el hecho de haber podido examinar desde varias perspectivas la cuestión planteada, a la vez que poder utilizar una variedad de recursos accesibles en sus ordenadores. Todos estos recursos conducían a soluciones distintas, pero todas ellas con sentido y aceptables en el contexto del trabajo a realizar, lo cual fue muy apreciado (e incluso algo sorprendente para algunos). En el ámbito académico, y por necesidad de evaluar las respuestas como correctas o incorrectas en el día a día, quizás hay un exceso de cuestiones planteadas que conducen a una única solución correcta, y los estudiantes están más entrenados a conseguir “la” respuesta que a conseguir “las” posibles respuestas.

Destacaron también que tuvieron la oportunidad de reflexionar sobre aspectos o ideas previamente presentadas en otras asignaturas pero que no habían estado, hasta ese momento, trabajadas de forma conjunta, estableciendo las conexiones y las atractivas relaciones que habían verificado en el contexto de esta modelización. Sorprendió constatar que, aunque se conocían los recursos informáticos utilizados (la hoja de cálculo, el análisis de datos de la hoja de cálculo, el programa matemático Maple y el programa estadístico R), se tuvieron algunas dificultades iniciales con su uso, por lo que la actividad reforzó y amplió sus habilidades digitales. La habilidad en el uso de los ordenadores, indudablemente, ayuda a los estudiantes en su educación en general, y en particular, en contextos matemáticos aplicados como éste, el de la modelización, ayuda que no puede ser de ninguna manera obviada si se quiere trabajar en un contexto real (y no en un contexto artificial preparado de antemano con valores o resultados simples). La opción del *Solver* en la hoja de cálculo Excel y el paquete *nlsmicrobio* del proyecto R fueron las dos novedades o recursos computacionales desconocidos por estos estudiantes que apreciaron notablemente.

El desarrollo de esta actividad supuso la posibilidad de ir alentando a los estudiantes a desarrollar su confianza, a la vez que su independencia y trabajó autónomo. El uso del ordenador y de sus recursos hizo posible repetir las veces que fuera necesario un cálculo o probar con múltiples elecciones de valores y poner a prueba lo que sucedía con lo experimentado y obtenido, sin importar si en la primera elección o ejecución se acertaba o no, reforzando la confianza en el trabajo individual realizado. El poder reiterar de forma rápida o, deshacer y volver hacer, en caso de que no se hiciera bien a la primera, de forma automatizada, permitió que fueran tomando confianza en todo el proceso de modelización con datos reales, una característica muy apreciada en una actividad de entrenamiento y aprendizaje como ésta.



**Figura 6:** Captura de pantalla de una parte de la salida de resultados que proporciona el programa R, con la utilización del paquete *nlsmicrobio*, cuando se aplica a los datos de la Tabla 1 transformados previamente a partir del logaritmo decimal.

El conjunto de las distintas tareas planteadas que configuraron las nueve partes de la actividad ofreció muchas oportunidades para refrescar y aplicar conocimientos previos, y se convirtió en un “viaje” matemático-biológico a consumir por los estudiantes. La secuencia de pasos de dificultad creciente, desde unos descubrimientos iniciales de patrones o tendencias de los datos (modelos simples), hasta la consecución de descripciones más completas y sofisticadas del crecimiento de la población (modelos empíricos y modelos pseudo-mecanicistas), permitió que fuera una actividad accesible a una amplia gama de estudiantes, ofreciendo oportunidades de éxito inicial y diferentes niveles de desafío. Se ahondó en la comprensión de ideas matemáticas ligadas a fenómenos de poblaciones microbianas, lo que fue realmente atrayente

para estos estudiantes de Ingeniería de Biosistemas. Se constató que los resultados conseguidos por ellos en las distintas tareas propuestas tenían plenamente sentido en el ámbito de la modelización y la microbiología predictiva (de interés en biotecnología y en la industria agroalimentaria).

La actividad desarrollada, formada por un conjunto de tareas que pueden ser clasificadas como “tareas ricas” después de su implementación en el aula, ha permitido llevar a cabo con éxito la modelización de un conjunto de datos reales de crecimiento poblacional de interés en el estudio de biosistemas. Esta secuencia de tareas, planteadas y agrupadas en distintas partes para configurar una actividad completa en el ámbito de aplicación escogido, podría ser readaptada también con éxito al estudio de otros fenómenos o procesos.

## Referencias bibliográficas

- [1] Bennison, A., Goos, M. (2010): “Learning to Teach Mathematics with Technology: A Survey of Professional Development Needs, Experiences and Impacts”. *Mathematics Education Research Journal* 22, 31-56.
- [2] Buchanan, R.L., Whiting, R.C., Damert, W.C. (1997): “When is simple good enough: a comparison of the Gompertz, Baranyi, and three-phase linear models for fitting bacterial growth curves”. *Food Microbiology* 14, 313-326.
- [3] Caspersen, M.E., Nowack, P. (2014): “Model-based thinking and practice: a top-down approach to computational thinking”. In: *Koli Calling '14 Proceedings of the 14th Koli Calling International Conference on Computing Education Research*. ACM New York, pp.147-15.
- [4] Clark-Wilson, A., Robutti, O., Sinclair, N. (Eds.) (2014): “The mathematics teacher in the digital era. An international perspective on technology focused professional development”. Springer, Dordrecht.
- [5] de Vries, G., Hillen, T., Lewis, M., Müller, J., Schönfisch, B. (2006): “A course in mathematical biology: quantitative modeling with mathematical & computational methods”. Society for Industrial and Applied Mathematics, Philadelphia.
- [6] Gibson, A.M., Bratchell, N., Roberts, T.A. (1988): “Predicting microbial growth: growth responses of salmonellae in a laboratory medium as affected by pH, sodium chloride and storage temperature”. *International Journal of Food Microbiology*, 6, 155-178.
- [7] Ginovart, M. (2014): “Discovering the power of individual-based modelling in teaching and learning: the study of a predator-prey system”. *Journal of Science Education and Technology* 23, 496-513.
- [8] Papert, S. (1996): “An exploration in the space of mathematics educations”. *International Journal of Computers for Mathematical Learning* 1, 95-123.
- [9] Perez-Rodriguez, F. (2014): “Development and application of predictive microbiology models in food”. In: Granato D. and Ares G. (Editors) *Mathematical and statistical methods in food science and technology*. IFT Press, Wiley Blackwell. Pp. 321-361.
- [10] Shiflet, A.B., Shiflet, G.W. (2014): “Introduction to computational science: Modelling and simulation for the science”. Princeton University Press. Princeton.
- [11] Wing, J.M. (2006): “Computational thinking”. *Communications of the ACM* 49, 33-35.
- [12] Zwietering, M.H., Jongenburger, I., Rombouts, F.M., Van't Riet, K. (1990): “Modeling of the bacterial growth curve”. *Applied and Environmental Microbiology*, 56, 1875-1881.