



Modelos matemáticos para la dinámica
de parásitos, bacterias y virus, bajo
una estrategia de variación antigénica

Sebastián Castro

Maestría en Ingeniería Matemática, Facultad
de Ingeniería, Universidad de la República

Abril, 2017

Director Académico:

Enrique Cabaña, Instituto de Matemática y Estadística Rafael Laguardia (IMERL), Facultad de Ingeniería, Universidad de la República.

Directores de Tesis:

- Gustavo Guerberoff, Instituto de Matemática y Estadística Rafael Laguardia (IMERL), Facultad de Ingeniería, Universidad de la República.
- Fernando Alvarez, Departamento de Biología Celular y Molecular, Sección Biomatemáticas, Facultad de Ciencias, Universidad de la República.

Tribunal:

- Rodney Colina, Laboratorio de Virología Molecular, Regional Norte - Universidad de la República, Uruguay.
- Gustavo Sibona, Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Argentina.
- José Vieitez, Departamento de Matemática y Estadística, Regional Norte - Universidad de la República, Uruguay.

Resumen

Este trabajo de Maestría en Ingeniería Matemática trata sobre la utilización de modelos matemáticos para la descripción y análisis de la dinámica de una población de bacterias, parásitos o virus, que sigue un proceso de variación antigénica, en su interacción con el sistema inmune del hospedero que lo aloja. El Capítulo 1 comienza con la descripción del proceso de variación antigénica y se muestran datos de un ejemplo real. Luego se presentan las teorías matemáticas que sirven de base para algunos de los principales modelos que se han utilizado para describir este fenómeno y que se presentarán más adelante: (sistemas de) ecuaciones diferenciales, lineales y no lineales; y procesos de ramificación, de un tipo y multitypo. En el Capítulo 2 se presenta uno de los ejemplos que más influencia ha tenido: un modelo macroscópico para la dinámica del Virus de Inmunodeficiencia Humana / Síndrome de Inmunodeficiencia Adquirida (VIH/SIDA), basado en un sistema de ecuaciones diferenciales no lineales (del tipo presa-predador) y un proceso estocástico que registra la cantidad de estas ecuaciones en el tiempo. Se desarrollan en forma progresiva las distintas versiones del modelo, de manera analítica y mediante simulaciones computacionales en algún caso, con el objetivo de reflejar las principales características del fenómeno que se quiere modelar. El Capítulo 3 contiene la presentación y análisis de un modelo más reciente basado en un proceso de ramificación de dos tipos y pensado en principio para un proceso de variación antigénica general. Nuevamente se presentan en forma progresiva distintas versiones del modelo, incorporando mayor información biológica en cada etapa, y se analizan de forma matemática y mediante simulaciones. Por último, en el Capítulo 4 se realizan algunas observaciones generales sobre los modelos presentados y algunos temas no abordados. Además, se destacan los modelos presentados como ejemplos de interacción entre la Biología y la Matemática, en general y en este tema en particular.

Índice general

1. Introducción	5
1.1. Variación antigénica y ejemplos	5
1.2. Modelos matemáticos para la dinámica de poblaciones biológicas	7
1.2.1. Modelos macroscópicos basados en ecuaciones diferen-	
ciales	9
1.2.2. Modelos microscópicos basados en procesos de ramifi-	
cación	17
2. Modelos macroscópicos de variación antigénica	25
2.1. Introducción	25
2.1.1. Virus, sistema inmune y evolución	26
2.1.2. VIH/SIDA y variación antigénica	27
2.2. Un modelo para la dinámica de VIH/SIDA	29
2.2.1. Un modelo simple de variación antigénica	29
2.2.2. Un modelo con inmunidad específica y reactividad cru-	
zada	33
2.2.3. El modelo de umbral de diversidad antigénica	36
2.2.4. Algunas extensiones y discusiones sobre el modelo de	
umbral de diversidad antigénica	39
3. Modelos microscópicos de variación antigénica	42
3.1. Introducción	42
3.2. Un modelo microscópico simple de variación antigénica	43
3.2.1. Probabilidades de extinción	44
3.3. Generalización del modelo	47
3.3.1. Muerte de células resistentes	48
3.3.2. Crecimiento de la probabilidad de respuesta inmune	49
3.4. Autocontrol y ondas de parasitemia	52

3.4.1. Parámetros aleatorios	53
3.4.2. Simulaciones	55
3.5. Discusión	59
4. Comentarios finales	60

Capítulo 1

Introducción

Supuesto diálogo entre un/a matemático/a y un/a biólogo/a.

Biólogo/a: *al principio yo estudiaba matemática. Pero dejé y me dediqué a la biología.*

Matemático/a: *ah, no sabía. ¿Y por qué fue?*

Biólogo/a: *bueno, lo que pasaba era que si bien yo podía darme cuenta de cuáles afirmaciones eran verdaderas y cuáles eran falsas, lo que no podía hacer era decidir cuáles eran las importantes.*

Matemático/a: *es muy interesante lo que me decís, porque originalmente yo me había dedicado a la biología pero me cambié al campo de la matemática.*

Biólogo/a: *¿ah, sí? ¿Y por qué?*

Matemático/a: *porque si bien yo podía decidir cuáles de las afirmaciones eran importantes y separarlas de las triviales, mi problema era que no podía diferenciar las que eran ciertas de las que no.^a*

^aAdaptado de [25] sobre un supuesto diálogo entre Albert Einstein (como físico) y Henri Poincaré (como matemático), y siguiendo la analogía de [15].

1.1. Variación antigénica y ejemplos

Diversos virus, bacterias y parásitos han desarrollado mecanismos para evadir la respuesta del sistema inmune y de esta manera sobrevivir durante períodos de tiempo prolongados en el hospedero que los aloja. Entre estos mecanismos uno de los más relevantes es el de *variación antigénica*, el cual consiste en la alteración de proteínas de superficie del agente infeccioso para evitar ser reconocidos por el sistema inmune.

Cuando un organismo es expuesto a un antígeno particular, se estimula una respuesta inmune y se generan anticuerpos específicos para combatirlo. A partir de entonces el sistema inmune tiene la capacidad de “recordar” ese antígeno particular y si el mismo patógeno intenta reinfectar al mismo

hospedero, los anticuerpos específicos actuarán rápidamente para eliminarlo. Sin embargo, si el patógeno es capaz de alterar sus antígenos de superficie entonces podrá evadir la respuesta del sistema inmune re infectando al hospedero mientras se generan nuevos anticuerpos para el “nuevo” antígeno identificado. Mediante la repetición de esta estrategia durante el curso de una infección, algunos agentes infecciosos son capaces de permanecer en el hospedero durante períodos prolongados hasta eventualmente causarle la muerte.

El proceso de variación antigénica ha sido principalmente identificado y estudiado en algunas bacterias, parásitos como *Trypanosoma brucei* causante de la enfermedad del sueño en humanos y *Plasmodium falciparum* responsable de la malaria, y en ciertos virus donde *Influenza* y *VIH* son dos de los ejemplos más importantes.

En la Figura 1.1 se observa el proceso de una infección del parásito *Trypanosoma gambiense* (una subespecie del parásito *Trypanosoma brucei*) en un hombre de 26 años de edad. Allí¹ se observa la evolución del número de tripanosomas que circulan en el torrente sanguíneo (*parasitemia*) a intervalos diarios, luego de la infección. La estrategia de variación antigénica del parásito es responsable de la periodicidad observada, aunque no siempre debe esperarse tanta regularidad en estas *ondas de parasitemia* en cuanto a la altura, ancho, número y tiempo de aparición, sino que existe gran variedad de perfiles de parasitemia [5]. A medida que el número de tripanosomas se multiplica (aproximadamente cada 8 horas mediante la división en dos células hijas), se generan anticuerpos contra ellos. Luego de 5 a 7 días, estos anticuerpos eliminan la mayor parte de los tripanosomas y la población de parásitos disminuye hasta que aparece una nueva onda de parasitemia generada por variantes del tripanosoma que no son afectados por los anticuerpos generados para la población anterior. De esta manera, el sistema inmune empieza nuevamente su tarea de reconocimiento y combate al “nuevo” agente infeccioso. Este proceso puede continuar en forma indefinida, hasta la eventual muerte del hospedero. Durante el tiempo que dura la infección el sistema inmune mantiene esta dinámica de combate contra la población de parásitos, sufriendo un desgaste en la producción de anticuerpos que a la larga puede resultar inútil.

¹Descripción y figura tomadas de http://tryps.rockefeller.edu/trypsru2_avaiation_intro.html.

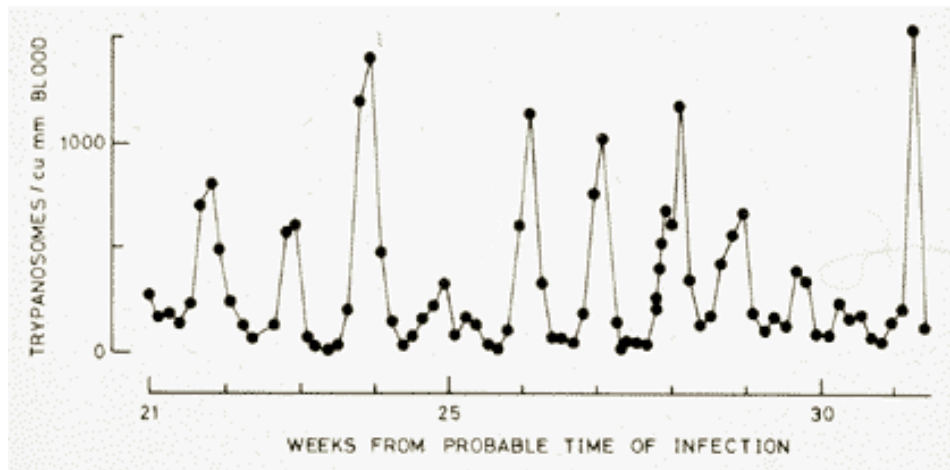


Figura 1.1: Ciclos de variación en la concentración de *Trypanosoma brucei gambiense* en sangre, medidos por Ross y Thompson (1910) en un individuo infectado [28].

La dinámica entre los agentes infecciosos y el sistema inmune involucra varios componentes diferentes. En este escenario, los principios que gobiernan esta dinámica y el resultado de la infección no pueden ser entendidos únicamente por razonamientos verbales o gráficos. Los *modelos matemáticos* representan una herramienta importante para capturar un conjunto básico de supuestos y desarrollarlos hasta llegar a conclusiones lógicas precisas. A su vez, a menudo permiten generar nuevas hipótesis, sugerir experimentos y medir parámetros críticos [35].

1.2. Modelos matemáticos para la dinámica de poblaciones biológicas

Un modelo matemático es una descripción formal, habitualmente a través de una o varias funciones y/o ecuaciones, de un fenómeno del mundo real. El propósito general de un modelo de este tipo es el de lograr un mejor entendimiento o descripción, y en algunos casos poder hacer predicciones acerca del comportamiento futuro del fenómeno. El planteo del modelo requiere la

utilización del conocimiento acerca del problema de estudio para la identificación de variables relevantes, la formulación de supuestos que simplifiquen lo suficiente el fenómeno para hacerlo manejable pero no demasiado alejado de la realidad, y la habilidad necesaria para la identificación de ecuaciones que relacionen las variables y respeten los supuestos asumidos. Estas ecuaciones a menudo son el reflejo directo de *leyes naturales* conocidas o supuestas (*hipótesis*), así como de la evidencia de *datos experimentales*.

Luego de formulado un modelo, la siguiente etapa consiste por lo general en *desarrollarlo* con el objetivo de deducir conclusiones sobre las variables de interés interpretándolas en el contexto del problema de estudio, y eventualmente haciendo predicciones. En una última etapa, estas conclusiones y predicciones serán contrapuestas con el conocimiento que se tiene del fenómeno así como de la evidencia empírica que se recabe para la evaluación del modelo. Si las conclusiones o predicciones no se condicen adecuadamente con la realidad, será necesario modificar el modelo o formular uno nuevo. Es importante tener siempre presente que un modelo nunca es una representación exacta de la realidad, sino más bien una *idealización* de la misma. Un buen modelo simplifica la realidad lo suficiente como para permitir la manipulación matemática pero no demasiado como para aún obtener conclusiones útiles.

En particular, un modelo para el crecimiento de una población se basa en uno o más *supuestos* acerca de cómo cambia la misma con el tiempo, en función de varios aspectos como el tamaño de la población actual, la disponibilidad de recursos del ambiente, la presencia o no de otras poblaciones que interactúan, la presencia de enfermedades y la respuesta a las mismas, etc. Si el tamaño de la población inicial es pequeño entonces un modelo *aleatorio* puede ser más apropiado para tener en cuenta la posibilidad de que la población se extinga debido al azar. Para tamaños poblacionales suficientemente grandes algunas dinámicas poblacionales pueden ser mejor descritas por modelos *determinísticos* [6].

Los modelos matemáticos determinísticos a menudo toman la forma de una *ecuación diferencial*, es decir, una ecuación que contiene una función desconocida y algunas de sus derivadas. En especial cuando lo que se desea modelar es la evolución o *dinámica temporal* de una o más variables de interés en función de sus valores actuales. En este contexto se puede representar por $x(t)$ a una función que representa el tamaño o densidad poblacional de

interés a tiempo t , se asume que dicha función es lo suficientemente “suave” (existen las derivadas del orden necesario) y se denota por $\dot{x}(t)$ a su derivada respecto de t , que indica cómo crece la población con el tiempo. Varios de estos modelos determinísticos tienen su análogo estocástico. Una de las principales diferencias entre ambos tipos de modelos es que los modelos determinísticos predicen un resultado con certeza mientras que los modelos estocásticos dan la probabilidad de cada posible resultado. En estos últimos y manteniendo la notación anterior, $x(t)$ representa una variable aleatoria para cada tiempo t , y más en general, $\{x(t) : t \geq 0\}$ un *proceso estocástico*. Otra distinción relevante proviene de considerar si el planteo de los modelos se hace a nivel global, utilizando ecuaciones directamente sobre la dinámica de toda la población, o si se hace a nivel individual describiendo el comportamiento de cada integrante de la población y posteriormente observando la evolución resultante del conjunto de la misma. Estos dos enfoques dan lugar, respectivamente, a los *modelos macroscópicos* y *microscópicos* de variación antigénica que serán considerados en este trabajo.

1.2.1. Modelos macroscópicos basados en ecuaciones diferenciales

Un primer modelo sencillo para la dinámica de una población se basa en el supuesto de que el crecimiento o decrecimiento depende únicamente del tamaño poblacional y es proporcional al mismo. Este puede ser un supuesto razonable para una población de bacterias o animales en condiciones ideales (recursos ilimitados, ausencia de predadores, inmunidad a enfermedades, etc.). El *modelo de crecimiento exponencial* asume estos supuestos a través de la siguiente ecuación diferencial con valor inicial:

$$\dot{x}(t) = kx(t), x(0) = x_0 > 0 \quad (1.1)$$

El parámetro k indica el sentido del crecimiento: si $k > 0$ la población aumenta, si $k < 0$ la población disminuye. Eventualmente este parámetro puede representar la tasa de crecimiento efectivo, definida como la diferencia entre la tasa de nacimiento y la de mortalidad ($k = r - d$). Reescribiendo la ecuación como $\dot{x}(t)/x(t) = k$, integrando y despejando se obtiene la solución $x(t) = x_0 e^{kt}$. De acuerdo con este modelo, a largo plazo la población crece arbitrariamente (“*explota*”) si $k > 0$ y se extingue si $k < 0$. Esto es, si $k > 0$, $\lim_{t \rightarrow \infty} x(t) = \infty$, y si $k < 0$, $\lim_{t \rightarrow \infty} x(t) = 0$. Este modelo no

acepta soluciones intermedias y resulta muy rígido en general. Sin embargo, puede resultar útil en situaciones donde los recursos son extremadamente abundantes, la escala de tiempo de observación es suficientemente pequeña para asumir $\dot{x}(t)/x(t)$ constante y el tamaño de la población inicial (x_0) es pequeño. Es un modelo razonable para estimar la tasa de crecimiento de un parásito cuando es introducido en la sangre de un individuo (como el parásito de la *malaria*), del número de nuevos casos de infección al *comienzo de una epidemia*, y de la tasa de extinción de especies en peligro, entre otras [6]. Además, como se verá más adelante, el modelo de crecimiento exponencial puede ser un componente dentro de un modelo más complejo.

En particular, un modelo más realista debe incorporar la observación de que muchas poblaciones tienen un crecimiento exponencial al inicio pero la población se estabiliza cuando se aproxima a la capacidad máxima que el ambiente puede soportar, conocida como *capacidad de carga* y denotada por el parámetro K . O que la población disminuye hacia esta cantidad si en algún momento la supera. En términos matemáticos, debe cumplirse que $\dot{x}(t) \approx kx(t)$ si $x(t)$ es pequeña (crecimiento exponencial al inicio), y $\dot{x}(t) < 0$ si $x(t) > K$. Un modelo sencillo que incorpora ambos supuestos es el *modelo logístico*, que se basa en la siguiente ecuación diferencial:

$$\dot{x}(t) = kx(t) \left(1 - \frac{x(t)}{K} \right), x(0) = x_0 > 0 \quad (k, K > 0) \quad (1.2)$$

Observar que si $x(t)$ es pequeña en comparación con K ($x(t) \ll K$), entonces $x(t)/K \approx 0$ y $\dot{x}(t) \approx kx(t)$ (crecimiento exponencial para poblaciones pequeñas). En cambio, si $x(t) > K$ entonces $1 - \frac{x(t)}{K} < 0$, $\dot{x}(t) < 0$ y la población disminuye. Además si $x(t) \approx K$ entonces $\dot{x}(t) \approx 0$. Estas observaciones muestran que $x(t) = K$ es un *equilibrio estable* del modelo, en el sentido de que si la población alcanza esta cantidad entonces pequeños cambios en la misma la harán regresar al valor de partida. Otra solución o *equilibrio* es $x(t) = 0$, sin embargo esta es *inestable*.

También es posible obtener una *solución explícita* para este modelo, al igual que para el exponencial. Para ello se reescribe la ecuación (1.2) como $\frac{\dot{x}(t)}{x(t)(1-x(t)/K)} = k$, se separa en fracciones simples $\frac{1}{x(t)(1-x(t)/K)} = \frac{K}{x(t)(K-x(t))} =$

$\frac{1}{x(t)} + \frac{1}{K-x(t)}$, se integra a ambos lados y se despeja obteniendo

$$x(t) = \frac{K}{1 + \left(\frac{K-x_0}{x_0}\right) e^{-kt}} = \frac{Kx_0}{x_0 + (K-x_0)e^{-kt}} = \frac{Kx_0e^{kt}}{(K-x_0) + x_0e^{kt}} \quad (1.3)$$

Se observa que esta función predice que a largo plazo la población alcanza su capacidad de carga: $\lim_{t \rightarrow \infty} x(t) = K$.

El modelo logístico representa una generalización del modelo exponencial, ya que este último puede derivarse del primero tomando capacidad de carga infinita ($K = \infty$). Del mismo modo, pueden obtenerse otros modelos de crecimiento o dinámica poblacional que generalicen al logístico en algún aspecto significativo que este no refleje.

Por ejemplo, para algunas poblaciones existe un nivel mínimo “m” donde la población se extinguirá si cae por debajo del mismo. Esta condición puede *modelarse* incorporando en la ecuación del modelo logístico el factor $(1 - m/x(t))$ y planteando la ecuación diferencial:

$$\dot{x}(t) = kx(t) \left(1 - \frac{x(t)}{K}\right) \left(1 - \frac{m}{x(t)}\right), x(0) = x_0 > 0 \quad (1.4)$$

Nuevamente, separando en fracciones simples, integrando y despejando se obtiene la *solución explícita* del modelo

$$x(t) = \frac{m(K-x_0) + K(x_0-m)e^{kt}}{K-x_0 + (x_0-m)e^{kt}} \quad (1.5)$$

El modelo predice que si $x_0 > m$, a largo plazo se alcanza la capacidad de carga: $\lim_{t \rightarrow \infty} x(t) = K$. Observar que de acuerdo con (1.4), si $m < x(t) < K$ entonces $\dot{x}(t) > 0$ y la población crece, y si $0 < x(t) < m$ entonces decrece. Esto muestra que $x(t) = m$ es un equilibrio inestable y que $x(t) = 0$ y $x(t) = K$ son equilibrios estables del modelo.

Otra posibilidad consiste en flexibilizar el modelo logístico incorporando una tasa de crecimiento inicial variable en el tiempo mediante una función $k(t)$. Esto es

$$\dot{x}(t) = k(t)x(t) \left(1 - \frac{x(t)}{K}\right), x(0) = x_0 > 0 \quad (1.6)$$

Operando en forma similar a los casos anteriores, no es difícil mostrar que la solución a la ecuación (1.6) está dada por

$$x(t) = \frac{Kx_0}{x_0 + (K - x_0)e^{-\int_0^t k(s)ds}} \quad (1.7)$$

Un modelo de dinámica poblacional en apariencia distinto está dado por la *función de Gompertz* que verifica la ecuación diferencial:

$$\dot{x}(t) = k \log\left(\frac{K}{x(t)}\right) x(t), x(0) = x_0 > 0 \quad (1.8)$$

Observar que de acuerdo a este modelo (considerando $k > 0$) si para algún t , $x(t) > K$, entonces la población disminuye ($\dot{x}(t) < 0$), mientras que si $x(t) < K$ la población aumenta. La solución $x(t) = K$ resulta en consecuencia un equilibrio estable. En forma similar a los casos anteriores puede obtenerse la solución explícita a este modelo:

$$x(t) = K \left(\frac{x_0}{K}\right)^{e^{-kt}} \quad (1.9)$$

que además verifica $\lim_{t \rightarrow \infty} x(t) = K$. Algunos ejemplos de las funciones consideradas pueden verse en la Figura 1.2.

Modelos lineales para la dinámica de dos o más poblaciones

Hasta aquí los distintos modelos presentados analizan la dinámica de una población en función de su tamaño y de otros componentes del ambiente, pero no toman en cuenta la interacción con otras poblaciones. Por ejemplo, si se desea modelar la dinámica de dos poblaciones biológicas que interactúan, un primer modelo sencillo podría estar dado por el siguiente *sistema de ecuaciones diferenciales lineales*:

$$\begin{aligned} \dot{x}_1(t) &= ax_1(t) + bx_2(t) \\ \dot{x}_2(t) &= cx_1(t) + dx_2(t) \end{aligned} \quad (1.10)$$

donde $x_1(t)$ y $x_2(t)$ denotan los tamaños poblacionales de ambas poblaciones de interés a tiempo t , y los parámetros a , b , c y d poseen una interpretación biológica directa. Si las especies *compiten* es razonable suponer $b, c < 0$ ya que el incremento en el tamaño de la población de una especie disminuye el crecimiento de la otra. Un segundo modelo plantea una relación de *cooperación*

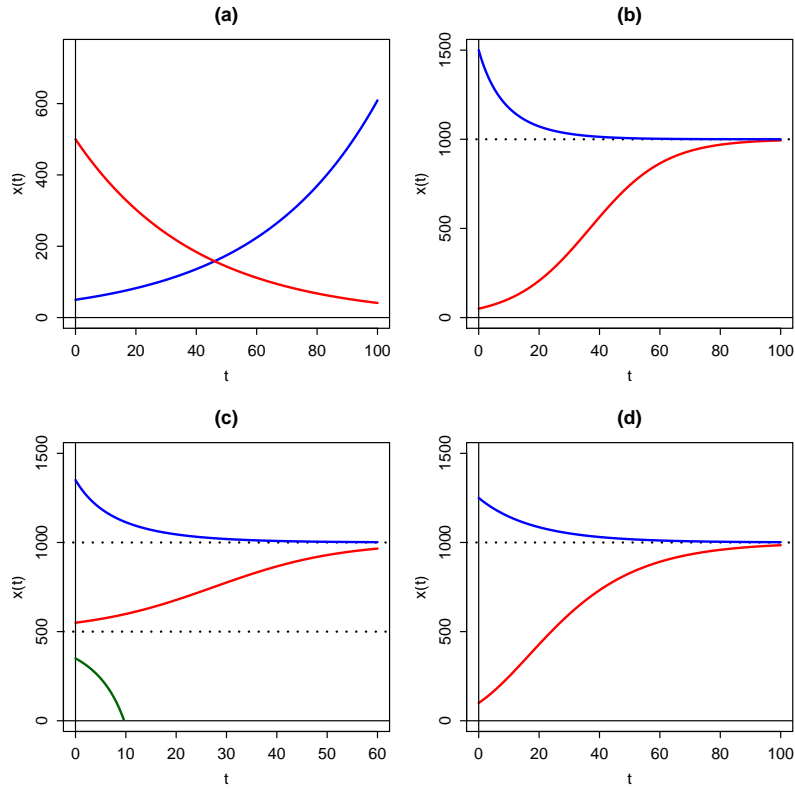


Figura 1.2: (a) modelo exponencial ($k > 0$ en azul y $k < 0$ en rojo); (b) modelo logístico ($x_0 > K$ en azul y $x_0 < K$ en rojo); (c) modelo logístico con mínimo ($x_0 > K$ en azul, $m < x_0 < K$ en rojo y $0 < x_0 < m$ en verde); (d) modelo de Gompertz ($x_0 > K$ en azul y $x_0 < K$ en rojo).

y en ese caso resulta adecuado tener $b, c > 0$. Por último, existe también una posibilidad de modelar una relación de tipo *predador-presa* bajo este modelo lineal. Si por ejemplo, la población 1 es la presa y 2 la predadora, entonces es razonable tener $b < 0$ y $c > 0$.

Un aspecto interesante de estos modelos es que puede obtenerse una *solución explícita* del sistema que permita estudiar sus propiedades, teniendo además una generalización del modelo de crecimiento exponencial de una población presentado anteriormente. Para observar esto en un contexto más general de n poblaciones que interactúan, se plantea un sistema lineal de

ecuaciones diferenciales:

$$\begin{aligned}
 \dot{x}_1(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \dots + a_{1n}x_n(t) \\
 \dot{x}_2(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \dots + a_{2n}x_n(t) \\
 &\vdots \\
 \dot{x}_n(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \dots + a_{nn}x_n(t)
 \end{aligned} \tag{1.11}$$

Definiendo $x(t) = (x_1(t), \dots, x_n(t))^{tr}$, $\dot{x}(t) = (\dot{x}_1(t), \dots, \dot{x}_n(t))^{tr}$ y $A = ((a_{ij}))_{i,j=1,\dots,n}$, el sistema (1.11) puede escribirse en forma compacta en notación matricial como

$$\dot{x}(t) = Ax(t) \tag{1.12}$$

Agregando la condición inicial $x(0) = (x_1(0), \dots, x_n(0))^{tr} = x_0$, esta ecuación es una generalización de la ecuación (1.1) para el modelo de crecimiento exponencial de una población. De hecho, la solución de la misma está dada por $x(t) = e^{At}x_0$, donde la *exponencial matricial* se encuentra definida como

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots = \sum_{j=0}^{\infty} \frac{A^j}{j!} \tag{1.13}$$

En la práctica, para el cálculo de e^{At} se utiliza la *descomposición diagonal* en vectores y valores propios o la *forma canónica de Jordan*. Dependiendo de las condiciones iniciales y de la magnitud y signo de los coeficientes del modelo (dados por x_0 y A respectivamente), el modelo permite representar situaciones de extinción o de coexistencia de las poblaciones. El análisis detallado de algunos ejemplos sencillos en el caso de dos poblaciones puede verse en [13] como aplicaciones de la descomposición de matrices.

Modelos no lineales para la dinámica de dos o más poblaciones

El modelo macroscópico más conocido para la dinámica de dos (o más) poblaciones que interactúan en la forma de *presa-predador* es el de *Lotka-Volterra*. La versión clásica (ver [6, 8]), para la dinámica de una población de presas $x(t)$ y una de predadores $y(t)$ a tiempo t , se basa en las siguientes hipótesis:

- en ausencia de depredadores, la población de presas crece exponencialmente a una tasa $a > 0$: $\dot{x}(t) = ax(t)$;
- en ausencia de presas, la población de depredadores disminuye exponencialmente a una tasa $b > 0$: $\dot{y}(t) = -by(t)$;
- en presencia de presas y depredadores, adicionalmente a las tasas anteriores de crecimiento y decrecimiento, hay una disminución en la población de presas y un aumento en la de predadores, ambas con una tasa proporcional a la frecuencia de encuentros entre los individuos de las dos especies (parámetros $p, q > 0$ respectivamente). Se supone además que la frecuencia de encuentros es proporcional al producto de ambas poblaciones $x(t)y(t)$.

Estos supuestos dan lugar al siguiente sistema *no lineal* de ecuaciones diferenciales

$$\begin{aligned}\dot{x}(t) &= ax(t) - px(t)y(t) = x(t)(a - py(t)) \\ \dot{y}(t) &= -by(t) + qx(t)y(t) = y(t)(-b + qx(t))\end{aligned}\tag{1.14}$$

A diferencia del sistema lineal de ecuaciones diferenciales, no existe una solución explícita para este modelo. Sin embargo, el mismo puede resolverse en forma aproximada numéricamente al tiempo que es posible estudiar algunas de sus propiedades más importantes como puntos de equilibrio, estabilidad y órbitas asociadas (ver [6,8] y [33]). Algunas generalizaciones pueden obtenerse sustituyendo el crecimiento/decrecimiento exponencial de una población en ausencia de otra por un término logístico, por ejemplo, o alguno de los presentados anteriormente. Esto, sin embargo, puede modificar algunos aspectos de la dinámica del modelo, por ejemplo al pasar de una órbita periódica a una en forma de espiral [6,8].

En la Figura 1.3 se muestra un ejemplo, tomado de [33], de la dinámica de interacción de conejos (presa) y lobos (predador). El modelo (A) se corresponde con el sistema presa-predador (1.14) con los parámetros $a = 0.08$, $p = 0.001$, $b = 0.02$ y $q = 0.00005$; mientras que en el modelo (B) se reemplaza el crecimiento exponencial de conejos en ausencia de lobos, por uno del tipo logístico con capacidad de soporte $K = 5000$. Es decir, la primer ecuación en (1.14) se reemplaza por $\dot{x}(t) = ax(t)(1 - \frac{x(t)}{K}) - px(t)y(t)$. En ambos casos las

condiciones iniciales son $x(0) = 1000, y(0) = 40$. Las soluciones numéricas de ambos sistemas se obtuvieron utilizando el paquete `odeSolve` del software `R` [27, 31, 32]. El modelo de Lotka-Volterra es la base de los modelos macroscópicos de variación antigénica donde los *virus* representan a las presas y el *sistema inmune* al predador y será retomado en el Capítulo 2.

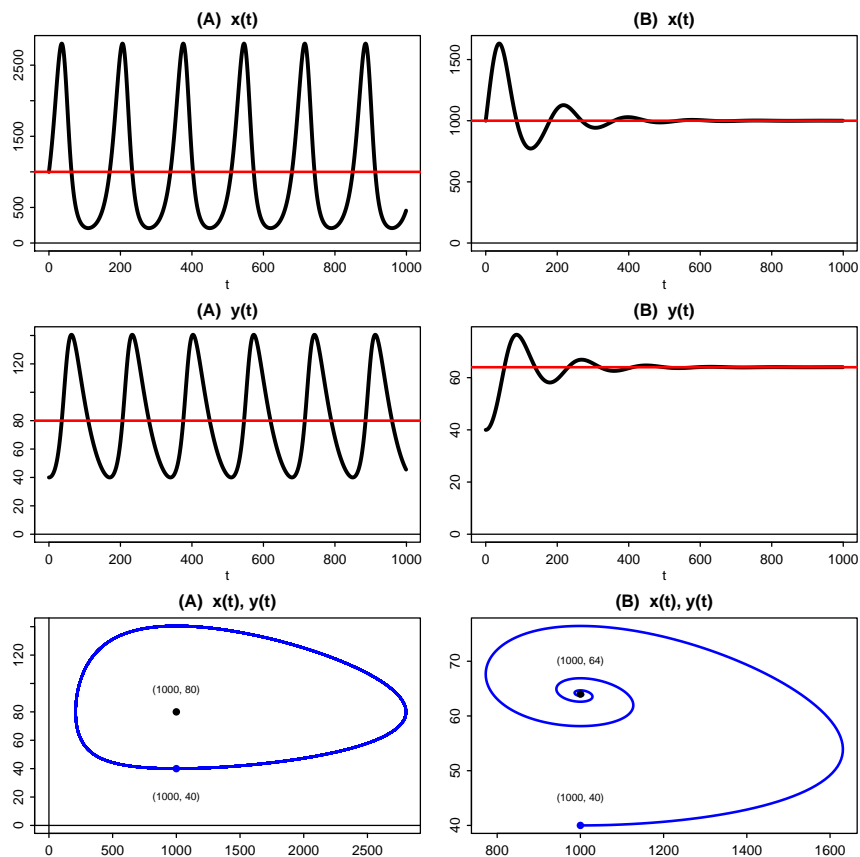


Figura 1.3: Dos ejemplos de modelos Lotka-Volterra: (A) crecimiento exponencial de conejos en ausencia de lobos, y (B): crecimiento logístico de conejos en ausencia de lobos. Las líneas rojas horizontales representan los niveles de equilibrio $\dot{x}(t) = \dot{y}(t) = 0$.

Dado un modelo determinístico para la dinámica de una población con solución $x(t)$, en algunos casos es posible identificar un estocástico que tiene como valor esperado de su solución a $x(t)$ (ver ejemplos en [2] y [14]). Esta

analogía resulta interesante pero no será desarrollada en profundidad en este capítulo, ya que el objetivo es el de presentar las herramientas básicas para los modelos de variación antigénica que se desarrollan en los Capítulos 2 y 3.

1.2.2. Modelos microscópicos basados en procesos de ramificación

En esta subsección se presentan modelos para la dinámica poblacional desde una perspectiva diferente. En lugar de plantear ecuaciones para el tamaño total de la(s) población(es) de interés, como en la sección anterior, se modela directamente la dinámica de cada individuo de la población utilizando descripciones o supuestos acerca de su vida y forma de reproducción, y luego se observa el resultado agregado de toda la población.

En este contexto, los *procesos de ramificación* son modelos microscópicos y estocásticos para la dinámica de poblaciones. En particular, los procesos de ramificación en tiempo discreto, también conocidos como *procesos de Galton-Watson*, son ejemplos de *cadena de Markov* en tiempo discreto; donde la variable tiempo y el espacio de estados son conjuntos discretos (habitualmente los enteros no negativos). Siguiendo la descripción realizada en [2], en el proceso de ramificación de Galton-Watson X_0 denota el tamaño total de la población de interés en la generación 0 (o a tiempo 0), y X_n , $n = 1, 2, \dots$, el correspondiente tamaño poblacional en la n -ésima generación. A su vez el proceso $\{X_n : n = 0, 1, \dots\}$ tiene espacio de estados (toma valores en el conjunto) $\{0, 1, 2, \dots\}$ y utiliza los siguientes supuestos:

- (i) cada individuo j en la generación n tiene $Y_n^{(j)}$ descendientes en la siguiente generación, donde $Y_n^{(j)}$ es una variable aleatoria (*va*) que toma valores en $\{0, 1, \dots\}$ con probabilidades respectivas (para todo j) $p_k = P(Y_n^{(j)} = k)$, $k = 0, 1, \dots$
- (ii) cada individuo tiene cierta cantidad (aleatoria) de descendientes *independientemente* del resto de los individuos;
- (iii) la misma distribución de descendencia se mantiene en todas las generaciones y para todos los individuos, en otras palabras, $\{Y_n^{(j)} : j = 1, 2, \dots, n; n = 0, 1, \dots\}$ son variables aleatorias *iid* (independientes e idénticamente distribuidas), con distribución común Y .

Observar que de acuerdo con esta descripción, si en alguna generación el tamaño poblacional es cero ($X_n = 0$), la población se extingue ($X_{n+j} = 0$ para $j = 1, 2, \dots$). En la terminología habitual de procesos estocásticos, “0” es un *estado absorbente* del proceso. Otra observación importante es que X_n puede escribirse en términos de las *va iid* $Y_n^{(1)}, Y_n^{(2)}, \dots$ mediante la relación

$$X_n = \sum_{j=1}^{X_{n-1}} Y_n^{(j)} = Y_n^{(1)} + \dots + Y_n^{(X_{n-1})}, \quad (1.15)$$

que muestra que X_n es una suma de *va iid* con valores en los enteros no negativos, donde la cantidad de sumandos (el tamaño poblacional en la generación anterior, X_{n-1}), es a su vez una variable aleatoria. En este contexto y con el objetivo de hallar algunos aspectos de la distribución de X_n , resulta de utilidad el uso de las *funciones generatrices de probabilidades (fgp)* [9].

Para simplificar la presentación, comencemos suponiendo un solo individuo en la población inicial ($X_0 = 1$) y denotemos por $f(t)$ a la *fgp* de Y . Esto es, $f(t) = E(t^Y) = \sum_{k=0}^{\infty} p_k t^k$, $t \in [0, 1]$. Además, sea $h_n(t)$ la *fgp* de X_n , donde en particular $h_0(t) = t$, dado que $X_0 = 1$ y $h_1(t) = f(t)$ debido a que $X_1 \sim Y$. Veamos que la *fgp* de X_n es la composición n -ésima de f , es decir,

$$h_n(t) = f^n(t) = f(f(\dots(f(t))\dots)) = f^{n-1}(f(t)) \quad (1.16)$$

La deducción puede realizarse mediante inducción. Asumiendo que $h_{n-1}(t) = f^{n-1}(t)$ y escribiendo $h_{n-1}(t) = E(t^{X_{n-1}}) = \sum_{k=0}^{\infty} P(X_{n-1} = k)t^k$, se obtiene la *fgp* de $X_n = \sum_{j=1}^{X_{n-1}} Y_n^{(j)} = Y_n^{(1)} + \dots + Y_n^{(X_{n-1})}$:

$$\begin{aligned} h_n(t) &= E(t^{X_n}) = \sum_{j=0}^{\infty} P(X_n = j)t^j = \\ &= \sum_{j=0}^{\infty} t^j \sum_{k=0}^{\infty} P(X_n = j | X_{n-1} = k) P(X_{n-1} = k) = \\ &= \sum_{j=0}^{\infty} t^j \sum_{k=0}^{\infty} P(Y_n^{(1)} + \dots + Y_n^{(k)} = j | X_{n-1} = k) P(X_{n-1} = k) = \\ &= \sum_{k=0}^{\infty} P(X_{n-1} = k) \sum_{j=0}^{\infty} P(Y_n^{(1)} + \dots + Y_n^{(k)} = j)t^j = \sum_{k=0}^{\infty} P(X_{n-1} = k) E\left(t^{Y_n^{(1)} + \dots + Y_n^{(k)}}\right) = \end{aligned}$$

$$\begin{aligned} \sum_{k=0}^{\infty} P(X_{n-1} = k) E(t^{Y_n^{(1)}}) \dots E(t^{Y_n^{(k)}}) &= \sum_{k=0}^{\infty} P(X_{n-1} = k) (E(t^Y))^k = \\ \sum_{k=0}^{\infty} P(X_{n-1} = k) (f(t))^k &= h_{n-1}(f(t)) = f^{n-1}(f(t)) \end{aligned} \quad (1.17)$$

Observar que esta derivación está basada en que $X_0 = 1$. Si en general $X_0 = N$ (población inicial arbitraria), el proceso podría considerarse como la suma de N procesos de ramificación independientes, $X_n = \sum_{i=1}^N X_{i,n}$, cada uno con *fgp* $f^n(t)$. Dado que N está fijo, es sencillo observar que en este caso la *fgp* de X_n resulta:

$$h_n(t) = (f^n(t))^N, \text{ si } X_0 = N. \quad (1.18)$$

Salvo en algunos casos particulares, como el de la *distribución geométrica* de progenie, en general no ha resultado posible obtener explícitamente la *fgp* de X_n a partir de la *fgp* de la distribución de descendencia $f(t)$ [2, 16]. Sin embargo, las ecuaciones (1.16) y (1.18) resultan de utilidad para estudiar algunas propiedades importantes del proceso, en especial la *probabilidad de extinción última* del proceso: $\lim_{n \rightarrow \infty} P(X_n = 0)$. Para ese propósito, se asumen las siguientes hipótesis sobre la distribución de descendencia:

$$0 < p_0 \text{ y } 0 < p_0 + p_1 < 1 \quad (1.19)$$

lo cual implica que existe probabilidad positiva de no descendencia, y de más de un descendiente. Si $p_0 = 0$, en cada generación habría al menos un nacimiento con lo cual no hay posibilidad de extinción. En el caso $0 < p_0$ y $p_0 + p_1 = 1$ la *fgp* de progenie es lineal, $f(t) = p_0 + p_1 t$, y puede probarse por inducción que $f^n(t) = 1 - (p_1)^n + p_1^n t$ (ver [2]), con lo cual $P(X_n = 0) = 1 - (p_1)^n \rightarrow 1$ si $n \rightarrow \infty$. Las hipótesis en (1.19) excluyen estas dos posibilidades.

Entonces, la *fgp* $f(t) = \sum_{k=0}^{\infty} p_k t^k$ verifica las siguientes propiedades:

1. $m = E(Y) = f'(1) = \lim_{t \rightarrow 1^-} f'(t) = \sum_{k=1}^{\infty} k p_k$.
2. $f(0) = p_0 > 0$, $f(1) = 1$.
3. $f(t)$ es continua para todo $t \in [0, 1]$.
4. $f(t)$ es infinitamente diferenciable en $(0, 1)$.

5. $f'(t) = \sum_{k=1}^{\infty} kp_k t^{k-1} > 0$ para $t \in (0, 1]$.

6. $f''(t) = \sum_{k=2}^{\infty} k(k-1)p_k t^{k-2} > 0$ para $t \in (0, 1]$.

Observar que m es el valor esperado de nacimientos de cualquier individuo de la población. Estas propiedades implican que la función f es continua, estrictamente positiva y estrictamente convexa en $[0, 1]$. Por lo tanto, su gráfico $y = f(t), t \in [0, 1]$ puede cortar la recta identidad $y = t$ en uno o dos puntos (ver Figura 1.4). Estos son llamados *puntos fijos* de f , ya que verifican la ecuación $f(t) = t$. Las propiedades anteriores son utilizadas para hallar la probabilidad última de extinción del proceso de ramificación.

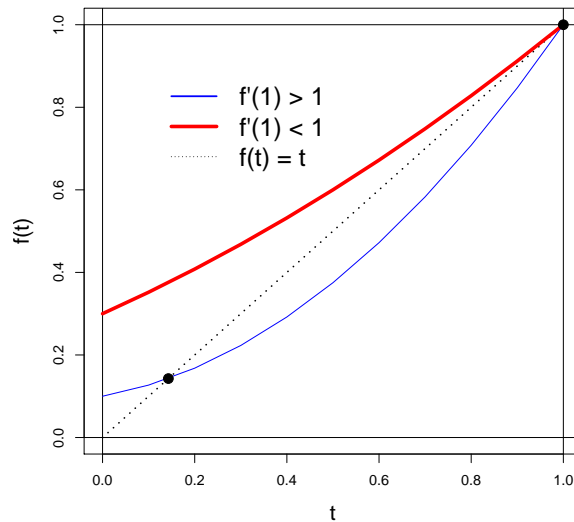


Figura 1.4: dos ejemplos de funciones generatrices de probabilidades (*fgp*): $f(t) = 0.1 + 0.2t + 0.7t^2$ (línea fina azul) y $f(t) = 0.3 + 0.5t + 0.2t^2$ (línea gruesa roja).

Se observa que como f' es continua y estrictamente creciente en $[0, 1]$, si $m = f'(1) \leq 1$ entonces $f'(t) < 1$ para $t \in [0, 1)$. Recíprocamente, si $f'(t) < 1$ en $[0, 1)$ entonces $m = f'(1) \leq 1$. Por otro lado, si $m \leq 1$ entonces $f(t)$ tiene un único punto fijo, en $t = 1$, en el intervalo $[0, 1]$. Esto se deduce de la propiedad anterior, ya que como $f'(t) < 1$ en $[0, 1)$, integrando entre

t y 1 en esta desigualdad se obtiene $f(1) - f(t) = 1 - f(t) < 1 - t$, lo cual implica que $t < f(t)$ para $t \in [0, 1)$. Por tanto, en el caso $m = f'(1) \leq 1$ el único punto fijo de $f(t)$ en $[0, 1]$ es $t = 1$. En la Figura 1.4 puede verse un ejemplo de esta situación.

Si fuera posible obtener una fórmula explícita para la fgp de X_n , $h_n(t) = \sum_{j=0}^{\infty} P(X_n = j)t^j$, entonces se tendría la probabilidad de extinción del proceso para cada tiempo n haciendo la evaluación $h_n(0) = P(X_n = 0)$. Si $X_0 = 1$, $h_n(0) = f^n(0)$ y si $X_0 = N$, $h_n(0) = (f^n(0))^N$. Luego sería directo obtener la probabilidad última de extinción tomando el límite: $\lim_{n \rightarrow \infty} P(X_n = 0)$. Como se mencionó anteriormente, esto rara vez es posible. Sin embargo, el siguiente resultado muestra que esta probabilidad puede obtenerse hallando los puntos fijos de la fgp de la distribución de progenie y que cualitativamente la extinción depende del número esperado de descendientes por individuo.

Teorema 1.2.1. *Se tiene un proceso de ramificación $\{X_n : n = 0, 1, \dots\}$, con $X_0 = 1$, cuya fgp de la distribución de progenie $f(t)$ verifica las propiedades (1) a (6) anteriores. Si $m = f'(1) \leq 1$ entonces $\lim_{n \rightarrow \infty} P(X_n = 0) = 1$, y si $m > 1$ entonces existe un único $0 < q < 1$ que verifica $f(q) = q$ tal que $\lim_{n \rightarrow \infty} P(X_n = 0) = q$.*

De acuerdo con este resultado, si $m \leq 1$ la probabilidad de extinción última es 1. Si $m > 1$, existe entonces una probabilidad positiva igual a $1 - q$ de que el proceso no se extinga nunca.

Demostración. Para cada $n = 0, 1, \dots$ se considera la sucesión $p_0(n) = P(X_n = 0)$. Dado que si $n < m$, el suceso $\{X_n = 0\}$ está incluido en $\{X_m = 0\}$, la sucesión $p_0(n)$ es *monótona creciente* y por definición está *acotada superiormente* por 1. Por lo tanto, tiene límite: $q := \lim_{n \rightarrow \infty} P(X_n = 0) \in [0, 1]$. Como se vio en (1.17), $h_n(0) = f(h_{n-1}(0))$, con lo cual $p_0(n) = f(p_0(n-1))$ y por la continuidad de f tomando límite a ambos lados se obtiene $q = f(q)$. Es decir, el límite q es un punto fijo de f . Como f es estrictamente convexa y $f(0) = p_0 > 0$, a lo sumo puede tener dos puntos fijos en $[0, 1]$, uno de los cuales es 1.

Si $m \leq 1$, el único punto fijo de f en $[0, 1]$ es 1 como se observó anteriormente, con lo cual $q = 1$. Si $m > 1$, la pendiente de la recta tangente al gráfico de $f(t)$ en $t = 1$ es mayor a 1 y en este caso $0 < q < 1$ (ver Figura 1.4). Resta verificar que, en estas condiciones, $\lim_{n \rightarrow \infty} p_0(n) = q < 1$. Supongamos que $p_0(n) \rightarrow 1$ cuando $n \rightarrow \infty$. En ese caso, para n suficientemente

grande, $p_0(n) > q$. Pero como en el intervalo $(q, 1)$, $t > f(t)$, se tiene que $p_0(n) > f(p_0(n)) = p_0(n + 1)$, lo cual contradice que $p_0(n)$ sea una sucesión creciente. Por lo tanto, $\lim_{n \rightarrow \infty} p_0(n) = q < 1$. \square

Habitualmente se dice que el proceso es *supercrítico* cuando $m > 1$, *crítico* si $m = 1$ y *subcrítico* cuando $m < 1$. El teorema anterior puede extenderse al caso $X_0 = N > 1$, en cuyo caso la fgp de X_n es $(f^n(t))^N$ y la probabilidad de extinción en la generación n respectiva es $(f^n(0))^N = (p_0(n))^N$. De esta manera, si $m \leq 1$ entonces $\lim_{n \rightarrow \infty} P(X_n = 0) = \lim_{n \rightarrow \infty} (p_0(n))^N = 1$; y si $m > 1$, $\lim_{n \rightarrow \infty} P(X_n = 0) = \lim_{n \rightarrow \infty} (p_0(n))^N = q^N < 1$. Observar que en este último caso si la población inicial es suficientemente grande, la probabilidad de extinción última puede volverse tan cercana a 0 como se quiera.

En la práctica, la convergencia $p_0(n) \rightarrow q$ es en general rápida y, de este modo, hay una probabilidad grande de que el proceso se extinga muy pronto o continúe para siempre [9]. Desde otra perspectiva, el tamaño esperado de la población en la n -ésima generación está dado por $EX_n = h'_n(1)$. Utilizando la expresión recursiva $h_n(t) = h_{n-1}(f(t))$ hallada en (1.17) y la *regla de la cadena* se obtiene

$$EX_n = h'_{n-1}(f(1))f'(1) = h'_{n-1}(1)m = h'_{n-2}(1)m^2 = \dots = m^n \quad (1.20)$$

No es sorprendente que el proceso tienda a extinguirse cuando $m < 1$ aunque a priori podría no ser evidente que una situación estable a largo plazo fuese imposible, aún cuando $m = 1$. Si $m > 1$ se espera un crecimiento geométrico de la población de acuerdo con (1.20), aunque no importa qué tan grande sea m siempre habrá una probabilidad positiva de extinción. De otra forma, puede mostrarse que salvo el estado “0” que es un estado absorbente y por lo tanto recurrente, los demás estados $k = 1, 2, \dots$ son todos transitorios [2]. Por lo tanto, cuando el número esperado de descendientes es mayor a uno ($m > 1$), se cumple que o bien la población se extingue con probabilidad positiva, $\lim_{n \rightarrow \infty} P(X_n = 0) = q$, $0 < q < 1$, o bien “explota” con probabilidad (positiva) complementaria: $P(\lim_{n \rightarrow \infty} X_n = \infty) = 1 - q$.

En el proceso de ramificación de Galton-Watson de “un solo tipo” considerado hasta ahora, cada individuo produce descendencia y es reemplazado por su progenie. Esta descendencia es del mismo tipo que sus ancestros, en el sentido de tener la misma distribución de probabilidad de progenie en cada generación. En un *proceso de ramificación multitypo*, cada individuo podría

generar diferentes tipos de individuos en la población y existe una distribución de descendencia correspondiente a cada tipo de individuos. El elemento clave de toda la teoría sigue siendo que cada individuo genera descendencia en forma *independiente* del resto. Este tipo de modelos puede utilizarse para la dinámica de dos o más poblaciones que interactúan de alguna manera, en forma análoga a la generalización que se planteó en la subsección anterior.

Un proceso de ramificación de Galton-Watson multitypo $\{X(n) : n = 0, 1, \dots\}$, donde $X(n) = (X_1(n), X_2(n), \dots, X_k(n))^{tr}$, es un *vector aleatorio* con k diferentes tipos de individuos. El elemento aleatorio $X_i(n)$ denota el número de individuos de tipo i en la generación n , y tiene asociado k *va* $\{Y_{i1}, Y_{i2}, \dots, Y_{ik}\}$ donde cada elemento Y_{ij} indica el número (aleatorio) de descendientes de tipo j de un ancestro de tipo i . La *fgp* asociada con la distribución de descendencia de $X_i(0) = 1$ es una función $f_i : [0, 1]^k \rightarrow [0, 1]$, definida como

$$f_i(t_1, t_2, \dots, t_k) = E(t_1^{Y_{i1}} t_2^{Y_{i2}} \dots t_k^{Y_{ik}}) = \sum_{s_1=0}^{\infty} \sum_{s_2=0}^{\infty} \dots \sum_{s_k=0}^{\infty} P(Y_{i1} = s_1, Y_{i2} = s_2, \dots, Y_{ik} = s_k) t_1^{s_1} t_2^{s_2} \dots t_k^{s_k} \quad (1.21)$$

para cada $i = 1, 2, \dots, k$. Considerando que $X_i(0) = \delta_i$ significa que inicialmente hay un sólo individuo de tipo i en la población (y ninguno de otro tipo), la *fgp* para $X_i(0) = \delta_i$ es $f_i^0(t_1, t_2, \dots, t_k) = t_i$. Denotando como $f_i^n(t_1, t_2, \dots, t_k)$ a la *fgp* de $X_i(n)$ respectiva, $f_i^1(t_1, t_2, \dots, t_k) = f_i(t_1, t_2, \dots, t_k)$ por ejemplo, se define la *fgp* vectorial $f : [0, 1]^k \rightarrow [0, 1]^k$,

$$f(t_1, \dots, t_k) = (f_1(t_1, \dots, t_k), f_2(t_1, \dots, t_k), \dots, f_k(t_1, \dots, t_k)) \quad (1.22)$$

La función f tiene un punto fijo en $(1, \dots, 1)$, dado que para cada i , $f_i(1, \dots, 1) = 1$. En forma análoga al proceso de Galton-Watson de un solo tipo, la probabilidad de extinción última depende de si f tiene otro punto fijo en $[0, 1]^k$, lo cual a su vez depende del número esperado de descendientes de cada individuo. Sea m_{ij} el número esperado de descendientes de tipo j provenientes de un individuo de tipo i , esto es,

$$m_{ij} = E(X_j(1) | X(0) = \delta_i), \text{ para } i, j = 1, 2, \dots, k, \quad (1.23)$$

los cuales, a su vez, se pueden escribir en términos de las *fgp* respectivas,

$$m_{ij} = \frac{\partial f_i(t_1, \dots, t_k)}{\partial t_j} \Big|_{t_1=1, \dots, t_k=1} \quad (1.24)$$

Sea $M = ((m_{ij}))_{i,j=1,2,\dots,k}$ la matriz de valores esperados. De acuerdo con [2], si M es *regular* ($M^p > 0$ para algún entero positivo p) entonces tiene un autovalor o valor propio λ de módulo máximo de cuya magnitud depende la probabilidad de extinción última del proceso de ramificación multitipo.

Teorema 1.2.2. *Asumiendo que la fgp $f(t_1, \dots, t_k)$ es una función no lineal de las variables t_1, \dots, t_k y que la matriz de valores esperados M es regular con valor propio dominante λ , entonces:*

- si $\lambda \leq 1$, $\lim_{n \rightarrow \infty} P(X(n) = 0 | X(0) = \delta_i) = 1$;
- si $\lambda > 1$, existe un único vector $q = (q_1, \dots, q_k)$, $0 < q_i < 1, i = 1, \dots, k$ tal que $\lim_{n \rightarrow \infty} P(X(n) = 0 | X(0) = \delta_i) = q_i$, siendo q un punto fijo de la función f ($f(q) = q$).

Una demostración en el caso de dos poblaciones (tipos) puede verse en [16]. La segunda parte del teorema puede extenderse fácilmente al caso de una población inicial arbitraria, $X(0) = (N_1, \dots, N_k)^{tr}$, $N_i \geq 0, i = 1, \dots, k$, utilizando la hipótesis de independencia. En este caso,

$$\lim_{n \rightarrow \infty} P(X(n) = 0 | X(0) = (N_1, \dots, N_k)^{tr}) = q_1^{N_1} \dots q_k^{N_k} \quad (1.25)$$

Nuevamente en forma análoga al proceso de un solo tipo, se observa que $(0, \dots, 0)$ es un estado absorbente y puede probarse que todos los demás estados son transitorios [2]. En consecuencia, bajo este tipo de procesos a largo plazo sólo se puede esperar una de dos cosas: “extinción” o “explosión” de la población. En el Capítulo 3 se presenta un modelo de variación antigénica basado en un proceso de ramificación de dos tipos [12].

Capítulo 2

Modelos macroscópicos de variación antigénica

Un pastor de ovejas se encuentra cuidando un rebaño numeroso y un hombre se acerca y le habla.

Hombre: *Si adivino el número correcto de ovejas que tiene, ¿me da una?*

Pastor: *Por favor, intente.*

Hombre: *Mmm... (piensa y dice) ... 183.*

Pastor: *Así es, dice el pastor sorprendido.*

El hombre toma una oveja, saluda y comienza a caminar. Entonces, el pastor le dice...

Pastor: *Espere. Si adivino cuál es su profesión, ¿puedo recuperar mi oveja?*

Hombre: *Claro, me parece justo.*

Pastor: *Usted debe ser un biólogo matemático.*

Hombre: *Correcto. ¿Cómo lo supo?*

Pastor: *Porque se llevó a mi perro.^a*

^aTomado de [23].

2.1. Introducción

Varios modelos del tipo macroscópicos se han propuesto para la dinámica de virus, parásitos o bacterias en presencia de variación antigénica. Las distintas variantes dependen del tipo de población que se intenta modelar así como de los supuestos que se asumen sobre el comportamiento de la misma y del sistema inmune. En este capítulo se describe uno de los que ha tenido mayor repercusión, conocido como el *modelo de umbral de diversidad* para la dinámica del VIH y su progresión hacia la enfermedad del SIDA,

siguiendo [19–21, 23, 24]. Primero se describe en términos generales el accionar de los virus y del sistema inmune, luego se analiza el caso particular del VIH/SIDA y por último se plantea y analiza el modelo en etapas progresivas.

2.1.1. Virus, sistema inmune y evolución

Los *virus* son agentes infecciosos microscópicos incapaces de reproducirse por sí mismos pero con posibilidades de manipular células del organismo que los hospeda para que lo haga. Las células podrían eventualmente dedicar todos sus recursos a producir nuevas partículas de virus hasta morir. Pero éstos pueden tener estrategias más sutiles: “decirle” al hospedero que los reproduzca pero a una tasa lenta sin poner en peligro la supervivencia de la célula. En principio, cualquier organismo que tenga la habilidad de reproducir información genética puede ser explotado por los virus.

Si bien todos los virus necesitan infectar células para lograr su reproducción, una vez que ingresan al hospedero no se replican sin oposición. La *evolución* ha equipado a todos los vertebrados, entre ellos el ser humano, con un complejo e ingenioso *sistema inmune* para combatir los virus y otras enfermedades infecciosas. Este sistema es el conjunto de estructuras y procesos biológicos en el interior de un organismo que le permiten mantener el equilibrio interno frente a agresiones internas y externas (*homeostasis*).

La columna vertebral del sistema inmune está constituida por unas pequeñas células llamadas linfocitos, que pueden clasificarse en células *B* y *T*. Un microlitro (μl) de sangre humana contiene alrededor de 2500 linfocitos y en un humano adulto existen alrededor de 10^{12} de estas células. Si un patógeno extraño, por ejemplo un virus, ingresa en el cuerpo humano se encontrará con células *B*, algunas de las cuales serán capaces de enlazar las distintas partículas virales en un mecanismo de “llave-cerradura”.

Estas células específicas, se vuelven activas y comienzan a dividirse aumentando en número y generando variantes incluso con mejores respuestas inmunes. Luego de cierto tiempo, la respuesta de células *B* puede alcanzar una especificidad óptima para las partículas virales invasoras. Los otros elementos claves del sistema inmune son las células *T*, tanto las células auxiliares *CD4* como las células “asesinas” *CD8*, que tienen la responsabilidad de reconocer y eliminar células infectadas.

El sistema inmune de los vertebrados, con sus células *B* y *T*, ha sido moldeado a lo largo de millones de años de evolución. La exposición continua a

agentes infecciosos ha promovido la *selección* de hospederos que sobreviven a las infecciones existentes. Pero la evolución también actúa para los microorganismos, seleccionando aquellas mutaciones que pueden superar la respuesta del sistema inmune y establecer infecciones con éxito. En su mayoría, además, estos tienen tiempos de generación más cortos y tasas de *mutación* más altas que sus hospederos, y se reproducen más rápidamente, lo cual constituye una ventaja desde el punto de vista evolutivo.

2.1.2. VIH/SIDA y variación antigénica

A comienzos de la década de 1980 se registró un incremento de pacientes que morían por enfermedades infecciosas que normalmente deberían ser controladas por el sistema inmune. La nueva enfermedad, caracterizada por un severo debilitamiento del sistema inmune y transmitida por contacto con secreciones genitales o transfusiones de sangre, fue denominada *síndrome de inmunodeficiencia adquirida* (SIDA). Diversos esfuerzos llevaron al descubrimiento del agente causante del SIDA: una clase completamente nueva de *retrovirus*, al que llamaron *virus de inmunodeficiencia humana* (VIH).

El VIH tiene como objetivo principal la eliminación de las células auxiliares $CD4$, un componente esencial de la respuesta inmune. Mediante la infección y agotamiento de esta población de células, el virus ataca y debilita el sistema inmune que debería combatirlo y controlarlo. Un adulto saludable tiene alrededor de 1000 células $CD4$ por μl de sangre, y una cantidad menor a 200 es utilizada como definición de SIDA (aún sin la presencia de enfermedades infecciosas). Se ha observado que existe una alta correlación entre la *carga de virus* (abundancia en sangre) durante el primer año de infección y el tiempo de supervivencia del paciente infectado.

Más allá de variaciones individuales, el patrón general de progresión de la enfermedad en la infección por VIH (en ausencia de tratamiento) se puede dividir en tres fases, como se resume en la Figura 2.1.¹

1^a Primeras semanas de infección: alta carga de virus y algunos síntomas similares a la gripe. Las células $CD4$ disminuyen al comienzo y luego retornan a niveles casi normales.

¹Fuente: <http://www.youngdayschool.edu.uy/webliceo/Imagenes/480px-Hiv-timecourse-es.png>

- 2ª Duración variable (de pocos meses a varios años, 8 a 10 en promedio): fase esencialmente asintomática donde el virus se replica y las células CD_4 disminuyen continuamente.
- 3ª Fase final: desarrollo del SIDA. Alta carga de virus y número de células CD_4 menor a 200 por μl . Aparecen enfermedades infecciosas que el sistema inmune debilitado no puede combatir y el paciente puede morir a causa de estas.

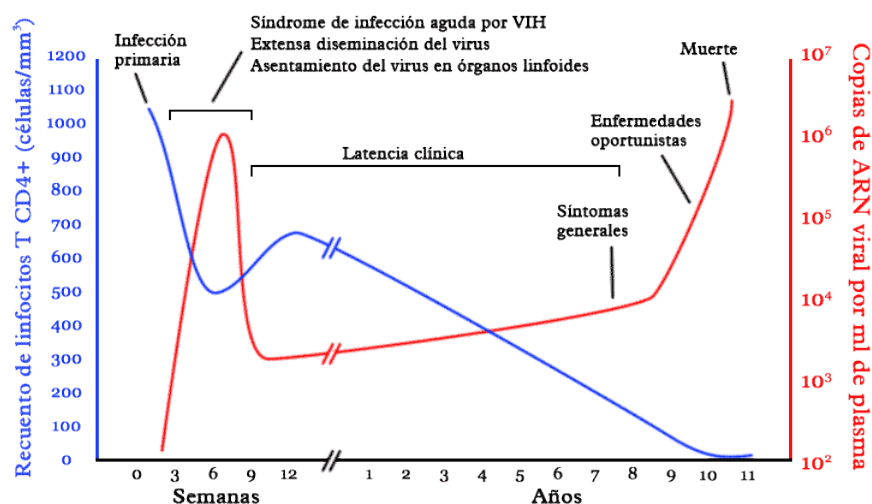


Figura 2.1: las tres fases generales de progresión del VIH/SIDA en ausencia de tratamiento.

Si bien existen alrededor de 20 drogas que inhiben la reproducción del VIH, estas tienen escaso efecto individual por el desarrollo de resistencia del virus, aunque con efectos positivos al aplicarse simultáneamente. Aunque no es posible erradicar el virus una vez que este ha infectado al paciente, en algunos casos es posible aumentar la calidad y esperanza de vida del paciente en forma considerable. Los tratamientos exitosos llevan a una enorme disminución en la carga del virus, sin embargo, estos son costosos y la mayor parte de las personas infectadas viven en países pobres.

Desde la aparición de la enfermedad y con el objetivo de la generación de tratamientos eficaces, uno de los principales intereses ha sido comprender los mecanismos de progresión del VIH/SIDA. En particular, entender por

qué existe un período de tiempo asintomático tan largo y variable que en última instancia da lugar al desarrollo de la enfermedad.

Una *teoría* presentada por primera vez en 1990 en [19], con algunas modificaciones en sucesivas publicaciones como [20, 21, 23, 24], plantea que el mecanismo clave de la progresión de la enfermedad es la *evolución* del virus en pacientes individuales. En el transcurso de una infección por VIH, el virus genera continuamente nuevas variantes a través de *mutaciones* que por algún período de tiempo no son reconocidas por el sistema inmune y por lo tanto logran escapar de su respuesta. Como se ha mencionado en el Capítulo 1, este proceso es conocido como *variación antigénica*.

En el proceso de una infección, el número de diferentes variantes antigénicas (*diversidad antigénica*) tiende a aumentar con el tiempo y permite al virus evolucionar y evadir la respuesta inmune durante períodos prolongados. Una vez que la diversidad antigénica supera cierto *umbral*, el sistema inmune es incapaz de continuar combatiendo el virus y se desarrolla el SIDA.

2.2. Un modelo para la dinámica de VIH/SIDA

Esta teoría fue formalmente planteada por sus autores a través de un modelo matemático, el *modelo de umbral de diversidad*, el cual será presentado a continuación en forma progresiva siguiendo principalmente las versiones más recientes en [23, 24].

2.2.1. Un modelo simple de variación antigénica

El siguiente modelo simple de variación antigénica describe la dinámica de un patógeno viral, o similar, al cual se le opone una respuesta inmune específica a cada variante del mismo. Sea v_i el tamaño de la población de la variante i y x_i la magnitud de la respuesta inmune específica (ambas funciones a tiempo t , $v_i = v_i(t)$, $x_i = x_i(t)$). La dinámica de interacción entre estas dos magnitudes se modela con el siguiente sistema de ecuaciones diferenciales del tipo “presa-predador” (el virus representa a la presa y el sistema inmune al predador). Para $i = 1, \dots, n$ variantes del virus:

$$\begin{aligned} \dot{v}_i &= r v_i - p x_i v_i; & r, p &> 0 \\ \dot{x}_i &= c v_i - b x_i; & c, b &> 0 \end{aligned} \tag{2.1}$$

donde $n = n(t)$ es un proceso estocástico que modela el supuesto de que las mutaciones generan constantemente nuevas variantes que pueden escapar de las respuestas inmunes existentes y $\dot{v}_i = \frac{d}{dt}v_i(t)$ y $\dot{x}_i = \frac{d}{dt}x_i(t)$ representan la dinámica de crecimiento de las variables de interés.

En ausencia de respuesta inmune, el crecimiento del virus es exponencial a tasa “ r ”. La respuesta inmune es estimulada a la tasa “ cv_i ”, proporcional a la abundancia del virus, elimina el virus a tasa “ px_iv_i ” y decae exponencialmente, en ausencia de estimulación del virus, a tasa “ b ”. En el modelo hay n variantes del virus, combatidas únicamente por n respuestas inmunes específicas. Las variantes evolucionan de manera *independiente*.

Como se mencionó en el Capítulo 1, sistemas de ecuaciones diferenciales como (2.1) no tienen solución analítica. Sin embargo, los mismos pueden resolverse numéricamente en forma aproximada. En las Figuras 2.2 y 2.3 se muestra un ejemplo de la dinámica del modelo (2.1), obtenida utilizando el paquete `odeSolve` de R [27, 31, 32].

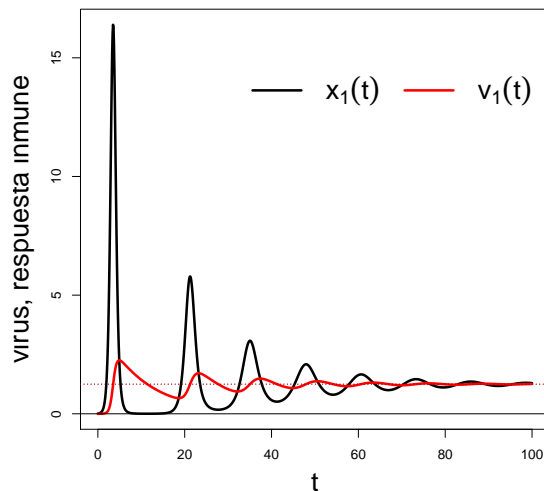


Figura 2.2: ejemplo de dinámica del modelo (2.1) con parámetros $r = 2.5, p = 2, c = b = 0.1$ y condiciones iniciales $v_1(0) = 0.1$ y $x_1(0) = 0$.

La dinámica comienza con una sola variante del virus, v_1 , la cual inicialmente crece exponencialmente a tasa r . El virus estimula la respuesta

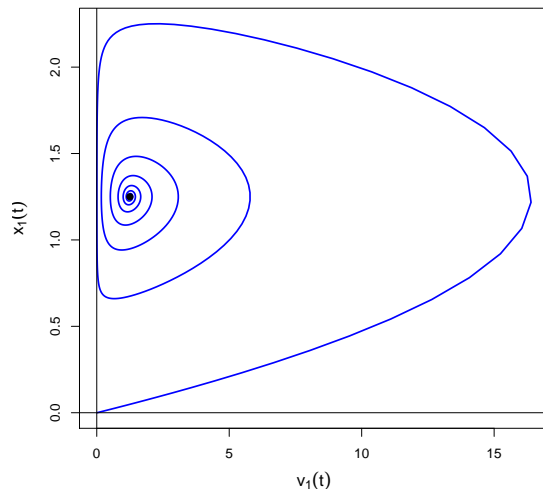


Figura 2.3: Trayectoria $(v_1(t), x_1(t))$ solución de (2.1) con parámetros $r = 2.5, p = 2, c = b = 0.1$ y condiciones iniciales $v_1(0) = 0.1$ y $x_1(0) = 0$.

inmune específica, x_1 , que reduce el crecimiento del virus y eventualmente pone fin a la expansión viral. El sistema se va estabilizando en oscilaciones amortiguadas hacia el equilibrio

$$v_1^* = \frac{br}{cp}; x_1^* = \frac{r}{p}, \quad (2.2)$$

que surge de resolver el sistema no lineal $\dot{v}_1 = \dot{x}_1 = 0$. En efecto, de la primera ecuación se obtiene $v_i(r - px_i) = 0$, con lo cual $x_i = r/p$, y de la segunda se despeja $v_i = (b/c)x_i$ y al reemplazar por el valor de x_i anterior se obtiene la solución (2.2).

En el modelo se asume que las mutaciones generan continuamente nuevas variantes virales que escapan de las respuestas inmunes específicas ya existentes. En este ejemplo, la generación de nuevas variantes se hace de acuerdo con un *proceso de Poisson* (homogéneo) de intensidad λ , con lo cual la probabilidad de que aparezca una nueva variante en el intervalo infinitesimal $[t, t + dt]$, es aproximadamente λdt [2, 16]. Es más realista, sin embargo, considerar que la probabilidad de aparición de nuevas variantes es proporcional a la carga de virus [23] y así se ha hecho en algunas versiones del modelo [19, 20]. Ca-

da vez que aparece una nueva variante, a tiempo aleatorio con distribución exponencial de intensidad λ , se resuelve el mismo sistema “presa-predador” que da la dinámica correspondiente. En la Figura 2.4 se observa el resultado de una simulación completa del modelo de variación antigénica (2.1).

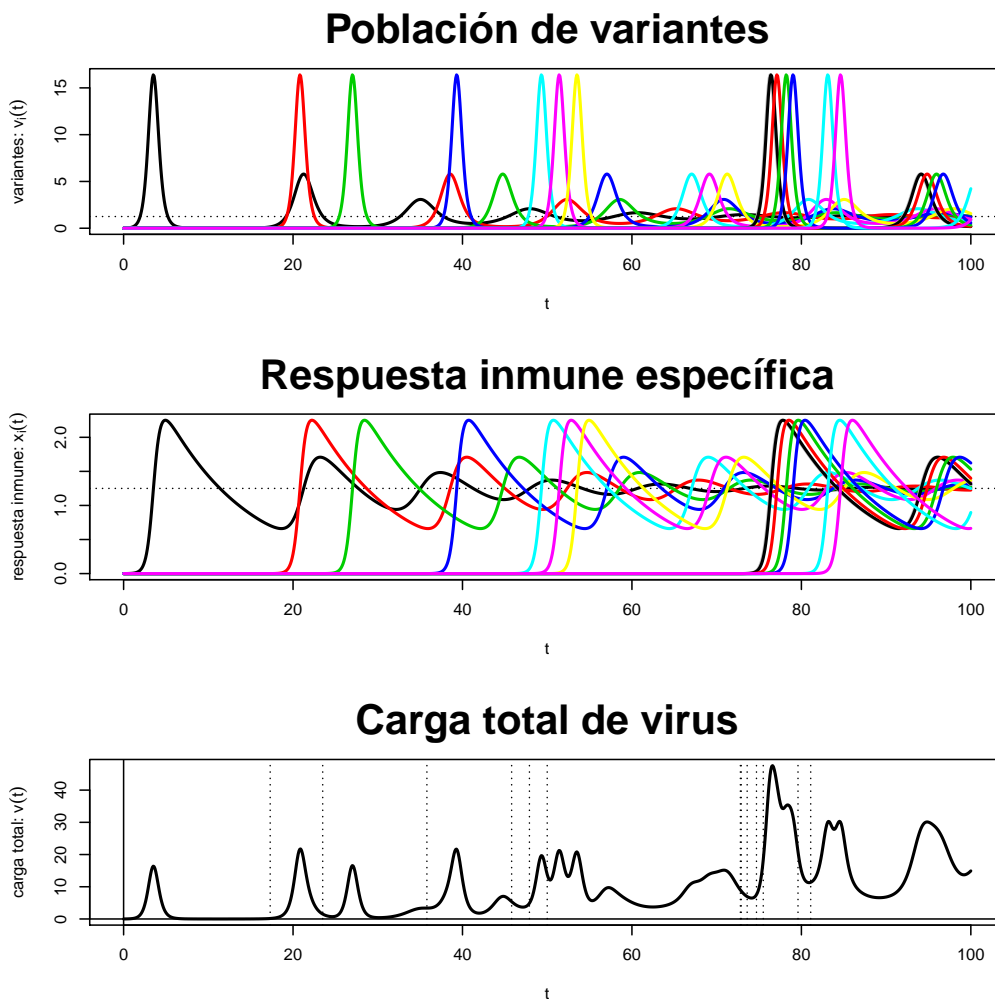


Figura 2.4: simulación del modelo de variación antigénica (2.1) con parámetros $r = 2.5, p = 2, c = 0.1, b = 0.1, \lambda = 0.1$ y condiciones iniciales $v_1(0) = 0.1$ y $x_1(0) = 0$. Los tiempos de aparición de nuevas variantes se representan con líneas punteadas verticales en el tercer gráfico.

En la simulación del modelo se observa que aparece una nueva variante, v_2 , que escapa de la respuesta inmune, x_1 y crece sin control. Pero induce una respuesta inmune, x_2 , que la combate y estabiliza luego de cierto tiempo. Mientras tanto, otra nueva variante, v_3 , emerge y estimula una respuesta x_3 , y así sucesivamente. El resultado es una sucesión de diferentes variantes antigénicas que crecen durante un tiempo antes de ser controladas por el sistema inmune. Si hay n variantes del virus y todas ellas están en el equilibrio dado por (2.2), la carga total del virus ($v = \sum_{i=1}^n v_i$) resulta una función creciente de la diversidad antigénica (n):

$$v = \frac{brn}{cp} \quad (2.3)$$

El sistema (2.1) define uno de los modelos más simples de variación antigénica, donde cada variante del virus, v_i , es controlada por una respuesta inmune específica, x_i . Esto implica que la dinámica de cualquier variante es independiente de todas las otras variantes, lo cual no es realista en la mayoría de los casos.

2.2.2. Un modelo con inmunidad específica y reactividad cruzada

Una extensión directa del modelo anterior consiste en incluir *respuesta inmune de reactividad cruzada* que pueda reconocer algunas (o todas) las variantes del virus. Denotando por “ z ” la magnitud de esta respuesta inmune, se plantea el siguiente modelo para $i = 1, \dots, n$ variantes ($r, p, q, c, b, k > 0$):

$$\begin{aligned} \dot{v}_i &= v_i(r - px_i - qz) \\ \dot{x}_i &= cv_i - bx_i \\ \dot{z} &= kv - bz \end{aligned} \quad (2.4)$$

La respuesta inmune de reactividad cruzada, “ z ”, es estimulada por todas las variantes del virus, cada variante i contribuye con la tasa “ kv_i ”, y decae exponencialmente a tasa “ b ”. Además, elimina la variante i del virus a tasa “ $qv_i z$ ”. Las dinámicas de las variantes individuales ya no son independientes entre sí.

Como las variantes que van apareciendo no escapan completamente a todas las respuestas inmunes existentes, su tasa de crecimiento es menor a la del modelo anterior. Para n variantes, la carga total de equilibrio del virus es

$$v^* = \frac{brn}{cp + kqn}, \quad (2.5)$$

que se obtiene nuevamente de resolver las ecuaciones $\dot{v}_i = \dot{x}_i = \dot{z} = 0$.

Al despejar en las tres ecuaciones de (2.4) se obtiene, respectivamente, $z = (r - px_i)/q$, $x_i = (c/b)v_i$ y $v = (b/k)z$. Si se sustituyen las dos primeras ecuaciones en la tercera resulta $v = \frac{b}{k} (r - p(\frac{c}{b}v_i)) / q$ y al sumar en $i = 1, \dots, n$, se obtiene $nv = \frac{b}{k} (nr - \frac{pc}{b}v) / q$. Luego, despejando v de esta expresión se llega a (2.5).

Nuevamente la carga total (v) es una función creciente de la diversidad antigénica (n), aunque se satura para valores grandes de n ($v^* \rightarrow br/kq$ si $n \rightarrow \infty$). La carga máxima posible de equilibrio es $v_{\text{máx}} = br/kq$, que representa la carga de equilibrio en presencia únicamente de respuesta inmune de reactividad cruzada (si $p = 0$ entonces $v^* = v_{\text{máx}}$). Por lo tanto, el incremento en la diversidad antigénica elimina el efecto de la respuesta inmune específica.

En este modelo de respuesta inmune específica y de reactividad cruzada, las variantes individuales del virus se caracterizan por ser diferentes en relación a la respuesta inmune específica a la variante, pero todas ellas son reconocidas por la respuesta inmune de reactividad cruzada. En la realidad, la situación es menos clara y algunos anticuerpos son capaces de reconocer varias (aunque no todas) las variantes virales. El modelo presentado solo considera las dos posibilidades extremas de respuesta inmune completamente específica o completamente de reactividad cruzada. Sin embargo, un espectro amplio de posibilidades se puede obtener asignando distintos valores paramétricos para balancear la importancia relativa de las respuestas inmunes específicas o de reactividad cruzada.

En la Figura 2.5 se muestra una simulación del modelo (2.4).² Nuevamente la simulación comienza con una sola variante viral, pero en este caso la misma induce tanto una respuesta inmune específica como de reactividad

²Figura tomada de [24].

cruzada. Las nuevas variantes que aparecen sucesivamente (de acuerdo con un proceso de Poisson como antes) escapan de las respuestas inmunes específicas existentes pero no de la de reactividad cruzada. La carga de equilibrio tiene una tendencia creciente con la diversidad antigénica pero se satura para niveles altos de esta.

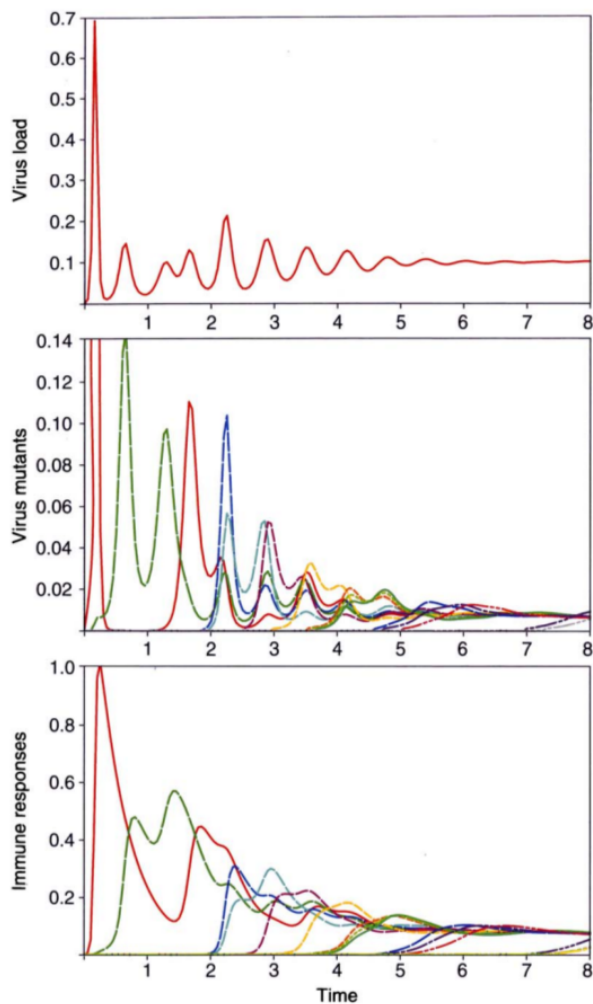


Figura 2.5: simulación del modelo de variación antigénica (2.4) con parámetros $r = 2.5$, $p = 2$, $c = k = 1$, $b = 0.1$, $\lambda = 0.1$, $q = 2.4$ y condiciones iniciales $v_1(0) = 0.1$ y $x_1(0) = 0$.

2.2.3. El modelo de umbral de diversidad antigénica

Los dos modelos anteriores describen la dinámica de cualquier virus o agente infeccioso que establece una infección persistente en su hospedero y puede mutar para escapar de la respuesta inmune. Es decir, que siguen una estrategia de variación antigénica. El siguiente modelo agrega la característica de que el virus es capaz de *debilitar la respuesta inmune*, lo cual lo hace más específico para VIH por su continua eliminación de células $CD4$ que ayudan a las células B y $CD8$ a montar la respuesta inmune contra el virus.

Para $i = 1, \dots, n$ variantes del virus y r, p, q, c, b, u, k parámetros positivos, se considera el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned}\dot{v}_i &= v_i(r - px_i - qz) \\ \dot{x}_i &= cv_i - bx_i - uvx_i \\ \dot{z} &= kv - bz - uvz\end{aligned}\tag{2.6}$$

El parámetro “ u ” caracteriza la habilidad del virus para debilitar la respuesta inmune y este efecto se resume en los términos “ uvx_i ” y “ uvz ”.

Asumiendo que los niveles de respuesta inmune (x_i y z) se estabilizan en tiempos más cortos comparados con aquellos en los que la población total de virus cambia, se consideran los niveles de equilibrio que se obtienen de resolver las ecuaciones $\dot{x}_i = \dot{v}_i = 0$ en (2.6). Esto es, $cv_i - (b + uv)x_i = 0$ y $kv - (b + uv)z = 0$, de donde se obtienen las respuestas inmunes de equilibrio (a corto plazo):

$$x_i^* = \frac{cv_i}{b + uv}, \quad i = 1, \dots, n; \quad z^* = \frac{kv}{b + uv}\tag{2.7}$$

Una vez que la respuesta inmune ha alcanzado estos niveles, la evolución de la población total de virus se obtiene reemplazando los valores de equilibrio anteriores en la primer ecuación de (2.6) y sumando en todas las variantes:

$$\begin{aligned}\dot{v} &= \sum_{i=1}^n \dot{v}_i = \sum_{i=1}^n v_i \left(r - p \frac{cv_i}{b + uv} - q \frac{kv}{b + uv} \right) = vr - \frac{pc}{b + uv} \sum_{i=1}^n v_i^2 - \frac{kqv^2}{b + uv} = \\ &= \frac{v}{b + uv} \left(r(b + uv) - pcv \sum_{i=1}^n \frac{v_i^2}{v^2} - kqv \right) = \frac{v}{b + uv} (rb + uvr - pcvD - kqv),\end{aligned}$$

donde la variable

$$D = \sum_{i=1}^n (v_i/v)^2 \quad (2.8)$$

denota el *índice de diversidad de Simpson*, un indicador muy utilizado en Ecología para medir la biodiversidad de un hábitat. Este índice verifica $1/n \leq D \leq 1$, y representa una medida inversa de diversidad. Se puede interpretar como la probabilidad de que dos partículas del virus elegidas al azar (con reposición) pertenezcan a la misma variante.

La primer acotación puede probarse utilizando la *desigualdad de Cauchy-Schwarz*,

$$\begin{aligned} v^2 &= \left(\sum_{i=1}^n v_i \right)^2 = \langle (v_1, \dots, v_n), (1, \dots, 1) \rangle^2 \leq \\ &\| (v_1, \dots, v_n) \|^2 \| (1, \dots, 1) \|^2 = n \sum_{i=1}^n v_i^2 \end{aligned}$$

Mientras que la segunda es consecuencia directa de la no negatividad de cada término v_i ,

$$\sum_{i=1}^n v_i^2 \leq v^2 = \left(\sum_{i=1}^n v_i \right)^2 = \sum_{i=1}^n v_i^2 + \sum_{i,j} v_i v_j$$

Claramente, si hay una sola variante del virus, $D = 1$, y si hay n variantes con la misma frecuencia, digamos $v_i = k$ para todo $i = 1, \dots, n$, entonces $D = \sum_{i=1}^n \frac{k^2}{(nk)^2} = 1/n$ (que además es el único caso donde se alcanza este valor, de acuerdo con la (des)igualdad de Cauchy-Schwarz).

En conclusión, la población total de virus evoluciona de acuerdo a

$$\dot{v} = \frac{v}{b + uv} [rb - v(cpD + kq - ru)] \quad (2.9)$$

que converge, para valores fijos de D , al nivel

$$v^* = \frac{rb}{kq + cpD - ru} \quad (2.10)$$

El término “ kq ” representa la eficacia de la respuesta inmune de reactividad cruzada. Mientras que “ cpD ” caracteriza la respuesta inmune específica de cada variante, cuya eficacia depende de la diversidad antigénica de la población de virus caracterizada por D . La ecuación (2.10) muestra que al aumentar la diversidad (disminuye D), tiende a aumentar la población total del virus.

El modelo tiene tres regiones paramétricas distintas que se corresponden con tres cursos de infección cualitativamente diferentes:

- (i) si $ru > kq + cp \Rightarrow$ **enfermedad inmediata**. Una sola variante puede superar el efecto combinado de la respuesta inmune específica y de reactividad cruzada. No es necesaria la variación antigénica para evadir el combate del sistema inmune.
- (ii) si $kq > ru \Rightarrow$ **infección crónica sin enfermedad**. El efecto solamente de la respuesta inmune de reactividad cruzada es suficiente para controlar el virus. El proceso de variación antigénica incrementa la carga del virus pero el sistema inmune es capaz de controlarlo.
- (iii) si $kq + cp > ru > kq \Rightarrow$ **infección crónica y enfermedad luego de largo período de incubación**. El efecto combinado de la respuesta inmune específica y la de reactividad cruzada puede controlar cualquier variante, pero la respuesta inmune de reactividad cruzada solamente no es suficiente para controlar el virus.

El caso (iii) es el más interesante. Al comienzo la diversidad antigénica es baja y la carga de virus se estabiliza en algún valor de equilibrio. La diversidad antigénica aumenta con el tiempo, reduciendo el efecto de la respuesta inmune específica, y eventualmente es tan grande que la carga total de virus aumenta sin control. Al principio $D = 1$ y tiende a disminuir durante la infección a medida que aumenta el número de variantes. El sistema inmune pierde el control cuando D cae por debajo de un **valor crítico** dado por:

$$D < \frac{ru - kq}{cp} \quad (\iff kq + Dcp - ru < 0) \quad (2.11)$$

Esta desigualdad define el “umbral de diversidad antigénica”. En el caso particular de que las poblaciones de virus tienen la misma frecuencia, vimos

que $D = \frac{1}{n}$ y al reemplazar en (2.11) se obtiene $n > \frac{cp}{ru-kq}$, una aproximación del número crítico de variantes que el sistema inmune puede tolerar en función de las características del virus (ru) y del sistema inmune (cp y kq).

Una vez que se excede el umbral la población de virus escapa del control del sistema inmune y tiende a crecer arbitrariamente. Este proceso podría interpretarse como el desarrollo de la enfermedad de inmunodeficiencia, caracterizada por alta carga del virus y agotamiento casi total de las células $CD4$. La razón principal de la pérdida de control del sistema inmune es la interacción asimétrica entre la diversidad inmunológica y la viral. Cada variante del virus puede debilitar toda respuesta inmune, pero la respuesta inmune específica puede solamente atacar la variante respectiva.

El resultado de una simulación para los casos (ii) y (iii) del modelo (2.4), puede verse en la Figura 2.6.³ En el caso (iii), el incremento en la diversidad antigénica permite a la población de virus escapar de la respuesta inmune luego de un largo período de incubación.

2.2.4. Algunas extensiones y discusiones sobre el modelo de umbral de diversidad antigénica

Los propios autores han considerado algunas extensiones del modelo de “umbral de diversidad antigénica” incluyendo diferentes tasas de replicación o propiedades inmunológicas para diferentes variantes del virus, así como la consideración de parámetros que cambian durante el transcurso de la infección. Estas modificaciones llevan a un mayor realismo así como a una mayor complejidad en el análisis del modelo. Si bien surgen nuevos elementos a considerar, como la correlación entre la carga de virus y la diversidad antigénica en pacientes individuales según la importancia relativa de la respuesta inmune específica o de reactividad cruzada, algunos elementos claves de la teoría, como la existencia de algún tipo de umbral de diversidad antigénica, se mantienen (por más detalles ver [23], Capítulo 13).

Por otro lado, en [30] otros autores realizaron sus propias simulaciones de una versión del modelo criticando la fuerte dependencia de los resultados

³Figura tomada de [24].

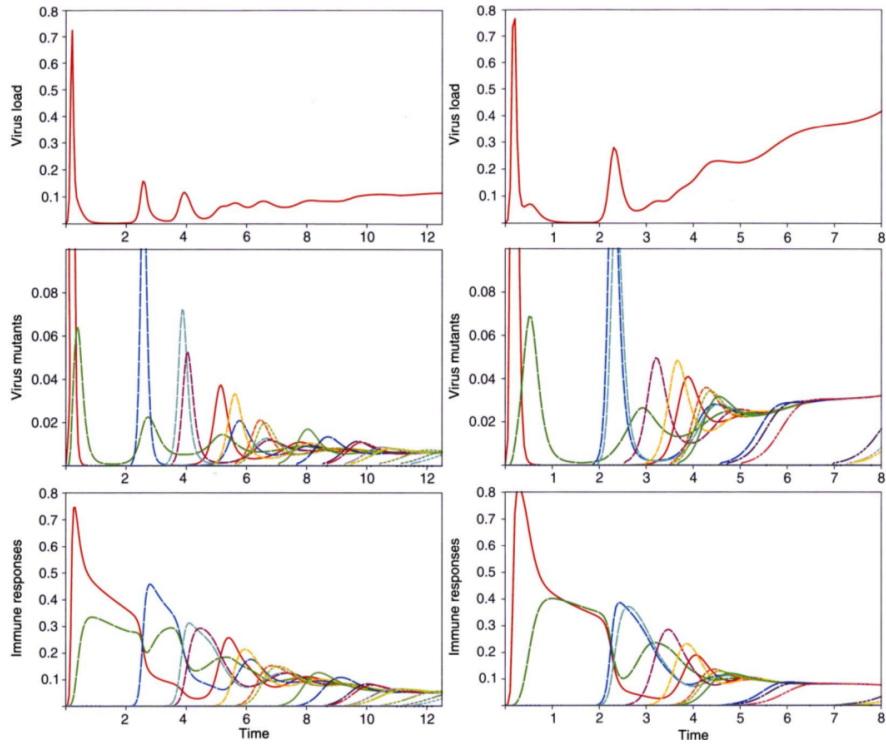


Figura 2.6: simulación del modelo de umbral de diversidad antigénica (2.6) con parámetros: en (ii) $r = 2.3, p = 2, c = k = u = 1, b = 0.1, \lambda = 0.1, q = 2.4$ y en (iii) $r = 2.5, p = 2, c = k = u = 1, b = 0.1, \lambda = 0.1, q = 2.4$. Condiciones iniciales en ambos casos $v_1(0) = 0.1$ y $x_1(0) = 0$.

en algunos parámetros y condiciones iniciales, así como cierta contradicción con lo observado en casos reales en relación con algunos aspectos del tiempo de incubación de la enfermedad. En un artículo posterior, los autores del modelo respondieron a estas críticas señalando que la respuesta a algunas de ellas ya estaba incluida en varios de sus artículos y mostrando que la supuesta contradicción con la realidad se debía a la consideración de una versión simple del modelo, pero que no ocurría con el presentado en este capítulo [17].

De acuerdo con Nowak en [24], desde la presentación del modelo cada hecho confirmado de la biología del VIH está de acuerdo con un modelo evolutivo de la progresión de la enfermedad. Concluye entonces, que en términos generales es probable que este modelo proporcione el mecanismo correcto de la progresión de la enfermedad del VIH/SIDA.

Capítulo 3

Modelos microscópicos de variación antigénica

Una persona le plantea a un/a biólogo/a y a un/a matemático/a el problema de hacer un té con los siguientes elementos: una garrafa, una pava (caldera) con agua, fósforos, una taza, un saquito de té y una cucharita. Cada uno, sin saber la respuesta del otro/a, responde:

Biólogo/a: *Primero, pongo la pava con agua arriba de la garrafa. Enciendo el fósforo y con él la garrafa. Espero a que hierva el agua. Pongo el saquito de té dentro de la taza. Vierto el agua dentro de la taza y revuelvo con la cucharita para que el saquito de té tiña el agua.*

Matemático/a: *Primero, pongo la pava con agua arriba de la garrafa. Enciendo el fósforo y con él la garrafa. Espero a que hierva el agua. Pongo el saquito de té dentro de la taza. Vierto el agua dentro de la taza y revuelvo con la cucharita para que el saquito de té tiña el agua.*

Bien, responde el examinador/a. Ahora les planteo otro problema: supongamos que les doy el agua hervida y les pido que hagan un té. ¿Qué harían?

Biólogo/a: *Bueno, en ese caso pongo el saquito de té dentro de la taza. Vierto en ella el agua ya hervida y revuelvo con la cucharita para que el saquito de té tiña el agua.*

Matemático/a: *Espero a que el agua se enfríe y vuelvo al caso anterior, que ya tengo resuelto. ^a*

^aTomado de [25].

3.1. Introducción

En este capítulo se presenta un modelo microscópico estocástico para describir la dinámica intra-hospedero de parásitos (o similar), que siguen una estrategia de variación antigénica. El modelo fue presentado recientemente en [12] y la descripción a continuación sigue esta referencia, en algunos casos con un desarrollo más detallado de los resultados principales y sus deducciones.

nes.

Un ejemplo de aplicación es el de *Tripanosoma Brucei*, responsable de la enfermedad del sueño en humanos, pero puede pensarse para otro tipo de parásitos, bacterias y virus. El enfoque está basado en un *proceso de ramificación multitypo en tiempo discreto* con dos tipos de células parasitarias definidas de acuerdo a la capacidad del sistema inmune de reconocerlas: células parasitarias *sensibles* (tipo 1) reconocidas por el sistema inmune pero que podrían dejar de serlo, y células parasitarias *resistentes* (tipo 2) no reconocidas aún por el sistema inmune pero que eventualmente podrían volverse sensibles. La versión más sencilla del modelo contiene solamente tres parámetros y permite la determinación de la probabilidad de extinción del proceso en función de estos, aunque no es capaz de reproducir las habituales *ondas de parasitemia* asociadas al proceso de variación antigénica que se observa en casos reales. Sin embargo, mediante la incorporación de un mecanismo dependiente de la abundancia del parásito, el modelo puede reproducir estos fenómenos. A su vez, el modelo permite algunas generalizaciones para incorporar información biológica, a costas de una eventual mayor complejidad. En este capítulo se describen en forma progresiva las principales características de estos modelos y al final se presentan algunos ejemplos mediante simulaciones.

3.2. Un modelo microscópico simple de variación antigénica

Como es mencionado anteriormente, en el modelo se asume la existencia de dos tipos de células parasitarias las cuales son definidas de acuerdo a la capacidad del sistema inmune de reconocerlas (células tipo 1 y 2). Además, se consideran tres parámetros $\delta, \mu, p \in [0, 1]$, definidos de la siguiente manera. La población de células prolifera por división binaria donde cada descendiente de una célula sensible (tipo 1) muere, independientemente, con probabilidad δ y sobrevive con probabilidad $1 - \delta$. A su vez, cada célula sobreviviente puede convertirse en resistente (tipo 2), en forma independiente, con probabilidad μ . Por otro lado, cada célula resistente crea un clan o linaje al dividirse en dos células resistentes en cada generación, las cuales permanecen sin ser reconocidas, en forma independiente entre generaciones, con probabilidad p .

En otras palabras, la probabilidad de que el sistema inmune adquiriera la habilidad para reconocer un clan particular de células resistentes es $1 - p$. En resumen, el parámetro δ mide la eficiencia de la respuesta inmune contra las células sensibles; el parámetro μ representa la probabilidad de aparición de una nueva variante (célula resistente), y el parámetro p está relacionado al tiempo que demora el sistema inmune en reconocer una nueva variante.

Es importante remarcar que los diferentes clanes de células resistentes, así como las células sensibles, evolucionan todas de manera independiente entre sí. Sin embargo, las células resistentes en un clan dado no se comportan de manera independiente ya que el sistema inmune en un momento particular reconoce o no a todas las células del clan. Es por esto que el proceso no es estrictamente un proceso de ramificación multitipo como el presentado en el Capítulo 1. Sin embargo, como veremos en la próxima subsección, puede asimilarse a uno que sí lo sea con el propósito de obtener las probabilidades de extinción del proceso en función de los parámetros.

3.2.1. Probabilidades de extinción

Un proceso de ramificación multitipo, con las mismas probabilidades de extinción que el modelo de variación antigénica anterior, se obtiene colapsando a una generación cada clan de células resistentes. En este proceso las distribuciones de descendencia son diferentes según el tipo de células consideradas.

Cada célula sensible (tipo 1) da lugar a:

- dos células tipo 1 con probabilidad $(1 - \mu)^2(1 - \delta)^2$ (ambas células sobreviven y ninguna muta),
- dos células tipo 2 con probabilidad: $\mu^2(1 - \delta)^2$,
- una célula tipo 1 y una célula tipo 2 con probabilidad: $2\mu(1 - \mu)(1 - \delta)^2$,
- una célula tipo 1 con probabilidad: $2(1 - \mu)\delta(1 - \delta)$,
- una célula tipo 2 con probabilidad: $2\mu\delta(1 - \delta)$ y
- ninguna célula descendiente con probabilidad: δ^2 .

Por otro lado, cada célula resistente (tipo 2) da lugar a 2^N células sensibles (tipo 1) con probabilidad $p^{N-1}(1-p)$, $N = 1, 2, \dots$. Esto es, la distribución de descendencia de células tipo 2 es *geométrica* de parámetro p , donde los descendientes son células de tipo 1.

Para obtener las probabilidades de extinción del proceso comenzando con una célula tipo 1 o tipo 2, o eventualmente cualquier cantidad de células, se consideran las respectivas funciones generatrices de probabilidades en una generación y sus puntos fijos. Para cada $n = 0, 1, 2, \dots$ se denota por $Z_1(n)$ y $Z_2(n)$ el número de células tipo 1 y 2 respectivamente, presentes en la generación n en el modelo colapsado. En el caso particular $n = 1$ se define $Z_1 = Z_1(1)$ y $Z_2 = Z_2(1)$ para simplificar la notación. Entonces, las funciones generatrices de probabilidades (*fgp*) de Z_1 y Z_2 son:

$$f_1(s, t) = E(s^{Z_1} t^{Z_2} | Z_1(0) = 1, Z_2(0) = 0) = (\delta + (1 - \delta)[\mu t + (1 - \mu)s])^2,$$

$$f_2(s, t) = E(s^{Z_1} t^{Z_2} | Z_1(0) = 0, Z_2(0) = 1) = (1 - p) \sum_{k=1}^{\infty} s^{2k} p^{k-1}, \quad (3.1)$$

para $s, t \in [0, 1]$. En (3.1) la segunda ecuación es directa de acuerdo a la definición de *fgp* y las probabilidades de descendencia de una célula de tipo 2, y la primera puede obtenerse de igual manera realizando algunas operaciones:

$$f_1(s, t) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} s^n t^m P(Z_1 = n, Z_2 = m | Z_1(0) = 1, Z_2(0) = 0) =$$

$$\delta^2 + s2(1-\mu)\delta(1-\delta) + t2\mu\delta(1-\delta) + st2\mu(1-\mu)(1-\delta)^2 + s^2(1-\mu)^2(1-\delta)^2 + t^2\mu^2(1-\delta)^2 =$$

$$\delta^2 + (1-\delta)^2[2\mu(1-\mu)st + (1-\mu)^2s^2 + \mu^2t^2] + 2\delta(1-\delta)[(1-\mu)s + \mu t] =$$

$$\delta^2 + (1-\delta)^2[\mu t + (1-\mu)s]^2 + 2\delta(1-\delta)[\mu t + (1-\mu)s] = (\delta + (1-\delta)[\mu t + (1-\mu)s])^2.$$

Denotando por q_1 y q_2 las probabilidades de extinción del proceso comenzando con una célula tipo 1 o una de tipo 2, respectivamente, se tiene el siguiente resultado como caso particular de lo presentado en el Capítulo 1.

Teorema 3.2.1. *La probabilidad de extinción del proceso, $q = (q_1, q_2)$, es la solución más cercana al origen en $[0, 1]^2$, de la ecuación*

$$(f_1(s, t), f_2(s, t)) = (s, t).$$

Teniendo en cuenta que $f_1(1, 1) = f_2(1, 1) = (1, 1)$, dependiendo de los valores de los parámetros δ, μ y p puede ocurrir que $(1, 1)$ sea el único punto fijo (representando la extinción segura del proceso) o exista otra solución en $(0, 1)^2$ y la probabilidad de no extinción sea positiva. Con el objetivo de encontrar los valores paramétricos que distinguen ambas situaciones, buscamos $(s_0, t_0) \in [0, 1]^2$ que satisfagan

$$\begin{aligned} (\delta + (1 - \delta)[\mu t_0 + (1 - \mu)s_0])^2 &= s_0, \\ (1 - p) \sum_{k=1}^{\infty} s_0^{2^k} p^{k-1} &= t_0. \end{aligned} \quad (3.2)$$

De estas ecuaciones, obtenemos que la probabilidad de no extinción es positiva sí y sólo si existe $s_0 \in (0, 1)$ tal que

$$\left(\delta + (1 - \delta) \left[\mu(1 - p) \sum_{k=1}^{\infty} s_0^{2^k} p^{k-1} + (1 - \mu)s_0 \right] \right)^2 = s_0.$$

Para cada $s \in [0, 1]$ se considera la función

$$g(s) = \left(\delta + (1 - \delta) \left[\mu(1 - p) \sum_{k=1}^{\infty} s^{2^k} p^{k-1} + (1 - \mu)s \right] \right)^2.$$

Se observa que para $\delta, \mu, p \in (0, 1)$, $g(\cdot)$ es una función estrictamente creciente ($g'(s) > 0$) y convexa ($g''(s) > 0$), con $g(0) = \delta^2$ y $g(1) = 1$. Por lo tanto, existe $s_0 \in (0, 1)$ que satisface $g(s_0) = s_0$, sí y solo si $g'(1) > 1$, donde

$$g'(1) = 2(1 - \delta) \left[2\mu(1 - p) \sum_{k=0}^{\infty} (2p)^k + (1 - \mu) \right].$$

Dado que $\sum_{k=0}^{\infty} (2p)^k$ converge a $(1 - 2p)^{-1}$ si $p < 1/2$ y diverge a ∞ si $p \geq 1/2$, se distinguen dos casos de interés.

1. $p \geq 1/2$: $g'(1) = \infty$ y existe $s_0 \in (0, 1)$ tal que $g(s_0) = s_0$. La probabilidad de no extinción es positiva y se dice que el proceso se encuentra en la región *supercrítica*,

2. $p < 1/2$: $g'(1) = \frac{2(1-\delta)}{1-2p}(1 + \mu - 2p)$. En este caso, la existencia o no de una probabilidad positiva de no extinción depende de los valores relativos de los parámetros δ, μ, p .

Una forma de representar la superficie crítica en el segundo caso consiste en introducir para cada par (δ, p) el *valor crítico* de μ , $\mu_c(\delta, p) \in [0, 1]$, como aquel que verifica

$$\frac{2(1-\delta)}{1-2p}(1 + \mu_c(\delta, p) - 2p) = 1. \quad (3.3)$$

De donde surge que la probabilidad de extinción es 1 o estrictamente menor a 1, dependiendo de si $\mu \leq \mu_c(\delta, p)$ o $\mu > \mu_c(\delta, p)$ respectivamente. Este análisis se resume en el siguiente resultado donde se utiliza el hecho que $\mu_c(\delta, p)$ verifica (3.3) sí y sólo sí $\mu_c(\delta, p) = \frac{1-2p}{2(1-\delta)}(2\delta - 1)$.

Teorema 3.2.2. *Sea*

$$\mu_c(\delta, p) = \begin{cases} \min \left\{ 1, \max \left\{ 0, \frac{1-2p}{2(1-\delta)}(2\delta - 1) \right\} \right\} & \text{si } p < 1/2, \\ 0 & \text{si } p \geq 1/2. \end{cases}$$

Entonces, para cada (δ, p, μ) se cumple:

- (i) *el proceso es subcrítico (probabilidad de extinción igual a 1) si $\mu \leq \mu_c(\delta, p)$,*
- (ii) *el proceso es supercrítico (probabilidad de no extinción estrictamente positiva) si $\mu > \mu_c(\delta, p)$.*

3.3. Generalización del modelo

El modelo anterior puede ser generalizado para incluir situaciones más realistas, permitiendo en particular la muerte natural de células resistentes (a través de un conjunto de parámetros Δ) y mecanismos más eficientes de respuesta inmune para reconocer clanes resistentes (mediante un conjunto de parámetros \mathbf{p}).

Sean \mathcal{N} el número de células que comprenden un clan resistente dado justo en el momento que es reconocido por el sistema inmune (y por tanto se vuelve sensible), $\Theta(\Delta, \mathbf{p})$ la esperanza de esta variable aleatoria, siendo

el caso de interés $\Theta(\Delta, \mathbf{p}) > 1$, y $P_{\Delta, \mathbf{p}}$ su distribución de probabilidad. La descendencia de las células tipo 1 no se modifica y nuevamente para obtener las probabilidades de extinción o supervivencia, cada clan resistente se colapsa a una generación. Las *fgp* de (3.1) ahora son:

$$f_1(s, t) = (\delta + (1 - \delta)[\mu t + (1 - \mu)s])^2,$$

$$f_2(s, t) = \sum_{k=0}^{\infty} s^k P_{\Delta, \mathbf{p}}(\mathcal{N} = k).$$

Los mismos cálculos de la sección anterior llevan a que el proceso se extingue con probabilidad 1 sí y sólo si $2(1 - \delta)[\mu\Theta(\Delta, \mathbf{p}) + (1 - \mu)] \leq 1$, y se obtiene entonces el resultado general siguiente.

Teorema 3.3.1. *Sea $\mu_c(\delta, \Delta, \mathbf{p}) = 0$ si $\Theta(\Delta, \mathbf{p}) = \infty$ y*

$$\mu_c(\delta, \Delta, \mathbf{p}) = \min \left\{ 1, \max \left\{ 0, \frac{(2\delta - 1)}{2(1 - \delta)(\Theta(\Delta, \mathbf{p}) - 1)} \right\} \right\},$$

si $\Theta(\Delta, \mathbf{p}) < \infty$. Entonces, para cada $(\delta, \mu, \Delta, \mathbf{p})$ se cumple:

- (i) *el proceso es subcrítico (probabilidad de extinción igual a 1) si $\mu \leq \mu_c(\delta, \Delta, \mathbf{p})$,*
- (ii) *el proceso es supercrítico (probabilidad de no extinción estrictamente positiva) si $\mu > \mu_c(\Delta, \delta, \mathbf{p})$.*

En las próximas subsecciones se presentan dos casos particulares de interés, representando diferentes posibilidades de generalización del modelo: (i) *muerte de células resistentes* y (ii) *crecimiento de la probabilidad de respuesta inmune*. En ambos casos la distribución de descendencia de las células de tipo 1 permanece incambiada.

3.3.1. Muerte de células resistentes

Se continua asumiendo que en cada generación, las células tipo 2 de un clan dado permanecen siendo resistentes con probabilidad p o se vuelven sensibles con probabilidad $1 - p$. Adicionalmente, asumimos ahora que cada célula resistente dentro del clan prolifera por división binaria y su descendencia muere, independientemente, con probabilidad Δ (la versión anterior del

modelo se corresponde con $\Delta = 0$). Denotando por \mathcal{M} a la variable aleatoria que registra el número de generaciones hasta que un clan resistente dado se vuelve sensible, la distribución de esta variable aleatoria es *geométrica* de parámetro $1 - p$. Se observa que, en esta versión del modelo, previo a la generación \mathcal{M} el clan de descendientes de una célula de tipo 2 dada se comporta como un proceso de ramificación (de un sólo tipo) con distribución de descendencia *binomial* de parámetros 2 y $1 - \Delta$. Por lo tanto, el número medio de descendientes de cada una de estas células es $2(1 - \Delta)$. Luego, utilizando el resultado sobre valor esperado del proceso de ramificación del Capítulo 1 se obtiene:

$$\Theta(\Delta, p) = \sum_{m=1}^{\infty} P(\mathcal{M} = m) E(\mathcal{N} | \mathcal{M} = m) = \sum_{m=1}^{\infty} (1-p)p^{m-1} [2(1-\Delta)]^m. \quad (3.4)$$

Por lo tanto, teniendo en cuenta el resultado de suma de una *serie geométrica*, en esta versión del modelo se tiene

$$\Theta(\Delta, p) = \begin{cases} \frac{2(1-\Delta)(1-p)}{1-2p(1-\Delta)} & \text{si } 2p(1-\Delta) < 1, \\ \infty & \text{si } 2p(1-\Delta) \geq 1. \end{cases} \quad (3.5)$$

Por último, la superficie crítica en el espacio de parámetros (δ, μ, Δ, p) , que separa la región subcrítica de la supercrítica, se obtiene reemplazando (3.5) en el resultado general del Teorema 3.3.1.

$$\mu_c(\delta, \Delta, p) = \begin{cases} \min \left\{ 1, \max \left\{ 0, \frac{(2\delta-1)[1-2p(1-\Delta)]}{2(1-\delta)(1-2\Delta)} \right\} \right\} & \text{si } 2p(1-\Delta) < 1, \\ 0 & \text{si } 2p(1-\Delta) \geq 1. \end{cases}$$

3.3.2. Crecimiento de la probabilidad de respuesta inmune

En este caso, la generalización pasa por introducir un mecanismo por el cual la respuesta inmune aumenta su efectividad con el tamaño de los clanes resistentes. Esta extensión no afecta la intensidad de la respuesta inmune (regulada por el parámetro δ) sino el tiempo de espera hasta que la respuesta inmune aparece (a través del parámetro p).

En este sentido, el número de generaciones \mathcal{M} hasta que el sistema inmune reconoce un nuevo clan resistente es una variable aleatoria con la siguiente distribución de probabilidades: $P(\mathcal{M} = 1) = 1 - p_1$ y $P(\mathcal{M} = m) = p_1 p_2 \dots (1 - p_m)$, para $m = 1, 2, \dots$, donde p_i es la probabilidad de que el sistema inmune no reconozca al clan resistente en la generación i dado que no fue reconocido previamente. En la versión previa, más simple del modelo, $p_i = p$ para todo $i = 1, 2, \dots$ (caso *homogéneo*). En este contexto, la efectividad creciente de la respuesta inmune durante la infección puede ser explícitamente modelada eligiendo una sucesión no decreciente $p_{i+1} \leq p_i$, para todo $i = 1, 2, \dots$. Sin embargo, la elección de estos parámetros puede hacer más difícil la obtención de la superficie crítica en el espacio de parámetros. En los próximos tres ejemplos se exploran diversas situaciones donde esto es posible.

Ejemplo 1: tamaño uniformemente acotado de clanes resistentes.

Para simplificar, se supone que no hay muerte natural de células tipo 2 ($\Delta = 0$). Para $N_0 > 1$ y $p \in (0, 1/2)$ fijos se considera $p_i = p$ si $i \leq N_0$ y $p_i = 0$ si $i > N_0$. Es decir, a partir de cierto número de generaciones (N_0) el sistema inmune es capaz de reconocer cada clan de células resistentes con probabilidad 1. En este caso entonces,

$$\begin{aligned} \Theta(p, N_0) &= E(2^{\mathcal{M}}) = \sum_{m=1}^{\infty} 2^m P(\mathcal{M} = m) = \sum_{m=1}^{\infty} 2^m p_1 p_2 \dots p_{m-1} (1 - p_m) = \\ &= \sum_{m=1}^{N_0} 2^m p^{m-1} (1 - p) + 2^{N_0+1} p^{N_0} (1 - 0) = 2(1 - p) \sum_{m=1}^{N_0} (2p)^{m-1} + 2^{N_0+1} p^{N_0} = \\ &= 2(1 - p) \frac{1 - (2p)^{N_0}}{1 - 2p} + 2^{N_0+1} p^{N_0} = \frac{2(1 - p) - (2p)^{N_0+1}}{1 - 2p}, \end{aligned}$$

dando lugar a la superficie crítica

$$\mu_c(\delta, p, N_0) = \min \left\{ 1, \max \left\{ 0, \frac{(2\delta - 1)(1 - 2p)}{2(1 - \delta)[1 - (2p)^{N_0+1}]} \right\} \right\}.$$

Se observa que $\mu_c(\delta, p, N_0) \geq \mu_c(\delta, p)$ (valor crítico de μ en el caso homogéneo $p_i = p$ para todo i), con lo cual el efecto de esta modificación es un incremento de la región subcrítica en relación al modelo simple de variación

antigénica. Es decir, el proceso se extingue con probabilidad 1 en situaciones donde antes no lo hacía.

Ejemplo 2: crecimiento exponencial de la probabilidad de respuesta inmune.

Nuevamente se supone que no hay muerte natural de células tipo 2. Dados $p, \alpha \in (0, 1)$ fijos, se define $p_1 = p$ y $p_{i+1} = \alpha p_i$ para $i = 1, 2, \dots$. Esta situación se corresponde con un decaimiento exponencial de la probabilidad de que el sistema inmune no reconozca un clan resistente dado, en función del número de generaciones. Se tiene ahora que

$$\begin{aligned} P(\mathcal{M} = m) &= p_1 p_2 \dots p_{m-1} (1 - p_m) = p \alpha p \alpha^2 p \dots \alpha^{m-2} p (1 - \alpha^{m-1} p) = \\ &= p^{m-1} \alpha^{\sum_{i=1}^{m-2} i} (1 - \alpha^{m-1} p) = p^{m-1} \alpha^{\frac{(m-2)(m-1)}{2}} (1 - p \alpha^{m-1}), \end{aligned}$$

para $m = 1, 2, \dots$ y con este resultado se llega a

$$\begin{aligned} \Theta(p, \alpha) &= E(2^{\mathcal{M}}) = \sum_{m=1}^{\infty} 2^m p^{m-1} \alpha^{\frac{(m-2)(m-1)}{2}} (1 - p \alpha^{m-1}) = \\ &= 2 \sum_{m=1}^{\infty} (2p)^{m-1} \alpha^{\frac{(m-2)(m-1)}{2}} - \sum_{m=1}^{\infty} (2p)^m \alpha^{\frac{(m-2)(m-1)}{2} + m-1} = \\ &= 2 \sum_{m=1}^{\infty} (2p)^{m-1} \alpha^{\frac{(m-2)(m-1)}{2}} - \left(-1 + \sum_{m=0}^{\infty} (2p)^m \alpha^{\frac{(m-1)m}{2}} \right) = \\ &= 1 + \sum_{m=0}^{\infty} (2p)^m \alpha^{\frac{m^2 - m}{2}} = 1 + \sum_{m=0}^{\infty} \left(\frac{2p}{\sqrt{\alpha}} \right)^m (\sqrt{\alpha})^{m^2}. \end{aligned} \quad (3.6)$$

De esta expresión se puede obtener la superficie crítica $\mu_c(\delta, p, \alpha)$ reemplazando en el resultado general. Un hecho a resaltar es que en este caso el proceso puede estar en el régimen subcrítico incluso para valores de p arbitrariamente cercanos a 1.

Ejemplo 3: muerte de células resistentes y crecimiento exponencial de probabilidad de respuesta inmune.

Consiste en agregar al ejemplo anterior el efecto de la muerte de células resistentes a través del parámetro Δ como en la subsección 3.3.1. Combinando los resultados (3.4) y (3.6) se obtiene el número esperado de células que componen un clan resistente dado cuando este se vuelve sensible:

$$\Theta(p, \alpha) = 1 + (1 - \Delta) \sum_{m=0}^{\infty} \left(\frac{2p(1 - \Delta)}{\sqrt{\alpha}} \right)^m (\sqrt{\alpha})^{m^2}.$$

Como se menciona en [12], los ejemplos presentados de las generalizaciones del modelo simple de variación antigénica producen modificaciones en la estructura de las regiones crítica, subcrítica y supercrítica en el espacio de parámetros. Sin embargo, en todos los casos el proceso de ramificación muestra inestabilidad en el límite asintótico, en el sentido de que el proceso crece a tasa exponencial o se extingue (posiblemente a una tasa subexponencial). En otras palabras, todos los estados son transitorios a excepción de la extinción de la población de parásitos, y en consecuencia cualquier tipo de régimen de oscilación poblacional como los observados en presencia de variación antigénica son extremadamente inverosímiles. Los autores concluyen entonces que es necesario realizar algunas modificaciones a las distintas versiones de los modelos presentados anteriormente con el objetivo de obtener condiciones que favorezcan la aparición de *ondas de parasitemia*.

3.4. Autocontrol y ondas de parasitemia

En esta sección se presentan las extensiones del modelo que permiten la aparición de regímenes oscilatorios con cierta frecuencia para un número razonable de generaciones (en el orden de cientos o miles). Para simplificar, estas extensiones se realizan primero sobre la versión simple del modelo de variación antigénica y luego se agrega la versión con muerte natural de células resistentes.

3.4.1. Parámetros aleatorios

Diversos parásitos pueden controlar su densidad poblacional mediante mecanismos que favorecen infecciones prolongadas incrementando con ello las posibilidades de transmisión. Estos mecanismos de *autoregulación* poblacional son incorporados en el modelo considerando los parámetros δ y Δ como variables aleatorias, donde sus distribuciones dependen de los tamaños de las poblaciones de células sensibles y resistentes. Adicionalmente, el parámetro μ que gobierna la tasa de mutación también se puede modelar como una variable aleatoria dependiente de la densidad poblacional de manera tal que, cuando la población es pequeña la tasa de mutación se incrementa evitando la extinción, y por el contrario disminuye cuando la población de células es muy elevada, de modo de evitar la muerte del hospedero.

Como los parámetros δ, Δ y μ pertenecen al intervalo $[0, 1]$ (representan probabilidades), un modelo de probabilidad posible en estos casos es el de la distribución *Beta*. Esto es habitual, en particular, cuando se utiliza el enfoque de la inferencia bayesiana [10]. Adicionalmente, los autores muestran que al utilizar esta distribución sobre un proceso de Galton-Watson se obtiene una cota superior del tipo logística para el número medio de células en la población [12]. Entienden entonces que esta situación representa una población viral autocontrolada en ausencia de respuesta inmune y por lo tanto da respaldo a la elección de la distribución Beta para modelar los cambios de Δ y δ en función de las generaciones.

Una variable aleatoria X con distribución Beta de parámetros $a, b > 0$, denotada por $X \sim \text{Beta}(a, b)$, tiene función de densidad

$$f_X(x; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} 1_{\{0 \leq x \leq 1\}},$$

siendo $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$, $z > 0$, la función *Gamma*. Un aspecto importante a destacar de esta distribución es la posibilidad de representar situaciones bastante diversas según la elección de parámetros a, b . En términos cualitativos, se observa que si $a = b$ se obtiene una densidad simétrica alrededor del valor $1/2$, si a es mucho más grande que b ($a \gg b$) la distribución se concentra en 1 y si $b \gg a$ entonces la probabilidad está concentrada alrededor de 0. En este sentido se muestran algunos ejemplos en la Figura 3.1.

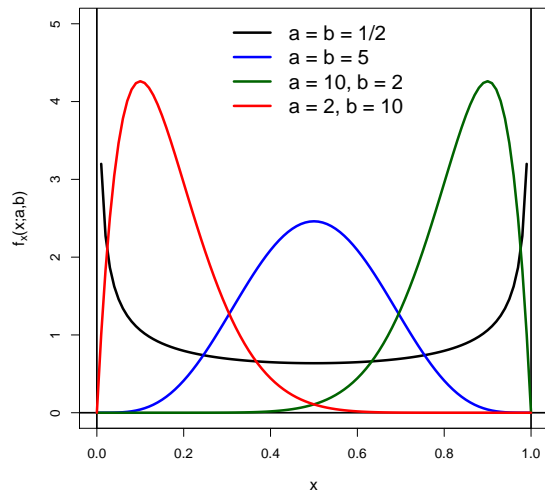


Figura 3.1: ejemplos de distintas funciones de densidad con distribución Beta, según posibles valores de sus parámetros.

Variación de δ y Δ

Se consideran los valores fijos $r_\Delta > 0$, $0 < \Delta_{\text{mín}} < 1/2 < \Delta_{\text{máx}} < 1$ y se denota por $Q(n)$ al número de células no reconocidas por el sistema inmune (células tipo 2) en la generación n . Sea $X_\Delta(n) \sim \text{Beta}(Q(n)/r_\Delta, 1)$, esto es, una variable aleatoria con distribución Beta de parámetros $a = Q(n)/r_\Delta$ y $b = 1$. Entonces para cada $n = 1, 2, 3, \dots$ se define

$$\Delta(n) = \Delta_{\text{mín}} + (\Delta_{\text{máx}} - \Delta_{\text{mín}})X_\Delta(n),$$

donde $\Delta_{\text{mín}}$ representa la probabilidad subyacente de muerte natural de células resistentes y el término $(\Delta_{\text{máx}} - \Delta_{\text{mín}})X_\Delta(n)$ incorpora un elemento de autocontrol en la población. En efecto, si para un n dado la población de células tipo 2 ($Q(n)$) es grande entonces $X_\Delta(n)$ tomará un valor cercano a 1 y $\Delta(n)$ será cercano a $\Delta_{\text{máx}}$. Esto tenderá a reducir la población en las próximas generaciones. Por otro lado, si $Q(n) \approx 0$ entonces $\Delta(n) \approx \Delta_{\text{mín}}$ y la población de células tipo 2 tenderá a crecer.

En forma similar se define la actualización del parámetro δ . Se fijan los parámetros $r_\delta > 0$, $1/2 < \delta_{\text{mín}} < \delta_{\text{máx}} < 1$ y se denota por $R(n)$ al número

de células reconocidas por el sistema inmune (células tipo 1) en la generación n . Se toma $X_\delta(n) \sim \text{Beta}(R(n)/r_\delta, 1)$ y se define

$$\delta(n) = \delta_{\text{mín}} + (\delta_{\text{máx}} - \delta_{\text{mín}})X_\delta(n).$$

Se interpreta el término $\delta_{\text{mín}}$ como la probabilidad subyacente de muerte de células sensibles por la respuesta inmune, mientras que $(\delta_{\text{máx}} - \delta_{\text{mín}})X_\delta(n)$ toma en cuenta el mecanismo de autocontrol de la población de células parasitarias sensibles en la generación n .

Se observa una diferencia importante entre la variación de $\Delta(n)$ y la de $\delta(n)$. Como $\delta(n) > 1/2$ para todo n , se produce constantemente una reducción en la población de células tipo 1 por el accionar del sistema inmune. En cambio $\Delta(n)$ varía tomando valores menores y mayores a $1/2$, lo cual enlentece pero no necesariamente detiene el crecimiento poblacional.

Variación de μ

En forma similar a los casos anteriores, se considera μ como una variable aleatoria que se actualiza en cada generación n mediante

$$\mu(n) = \mu_{\text{mín}} + (\mu_{\text{máx}} - \mu_{\text{mín}})X_\mu(n),$$

donde $0 \leq \mu_{\text{mín}} \leq \mu_{\text{máx}} \leq 1$ y para $r_\mu > 0$ fijo, $X_\mu \sim \text{Beta}(r_\mu/R(n), 1)$. El efecto de tener $\mu(n)$ variable es, en este caso, el de prevenir que se produzcan un número grande de eventos de mutación durante el pico de una infección y, en contrapartida, inducir una alta tasa de mutación cuando la población es pequeña disminuyendo la probabilidad de extinción.

3.4.2. Simulaciones

En esta sección se muestran algunas simulaciones de distintas versiones del modelo presentado en el capítulo. En la implementación de las simulaciones se sigue el algoritmo utilizado en [11] y se utiliza el entorno estadístico y lenguaje de programación R [27]. Al igual que en [12], el objetivo consiste en visualizar las dinámicas del proceso y en particular analizar las condiciones que promueven el surgimiento de ondas de parasitemia. Allí se destaca, además, que en la elección de valores de los parámetros se utilizan los resultados analíticos obtenidos sobre probabilidades de extinción. En particular,

se señala que el comportamiento oscilatorio puede ocurrir con mayor probabilidad en la superficie crítica.

Algoritmo para las simulaciones

Sean:

- $v(n)$: número de variantes no reconocidas a tiempo n ,
- $N_1(n), N_2(n), \dots, N_{v(n)}(n)$: número de células de cada una de las variantes no reconocidas,
- $R(n)$: número de células reconocidas (tipo 1) a tiempo n ,
- $Q(n)$: número de células no reconocidas (tipo 2) a tiempo n ,
- $Z(n)$: número total de células a tiempo n .

Entonces, para cada $n = 1, 2, \dots$

- $R(n+1) = \sum_{i=1}^{R(n)} \left[Y_1^{(i)}(1 - W_1^{(i)}) + Y_2^{(i)}(1 - W_2^{(i)}) \right] + \sum_{j=1}^{v(n)} 2N_j(n)(1 - X^{(j)})$,
- $Q(n+1) = \sum_{i=1}^{R(n)} \left[Y_1^{(i)}W_1^{(i)} + Y_2^{(i)}W_2^{(i)} \right] + \sum_{j=1}^{v(n)} 2N_j(n)X^{(j)}$,
- $v(n+1) = \sum_{i=1}^{R(n)} \left[Y_1^{(i)}W_1^{(i)} + Y_2^{(i)}W_2^{(i)} \right] + \sum_{j=1}^{v(n)} X^{(j)}$,
- $Z(n+1) = R(n+1) + Q(n+1)$.

Donde las variables aleatorias $Y_1^{(j)}, Y_2^{(j)}$ tienen distribución *Bernoulli* de parámetro $1 - \delta(n)$; $Y_1^{(j)}, Y_2^{(j)}$ *Bernoulli* de parámetro $\mu(n)$ y $X^{(j)}$ *Bernoulli* de parámetro $p(n)$, todas ellas independientes entre sí.

Modelo simple de variación antigénica (parámetros fijos)

A modo de ejemplo, se realizan dos simulaciones de la versión simple del modelo con los parámetros $\delta(n) = \delta = 0.6$, $\mu(n) = \mu = 0.004$ y $p(n) = p = 0.65$, y las condiciones iniciales $R(0) = 5 \times 10^3$ y $Q(0) = v(0) = N_1(0) = 1$ (las mismas que en [12]). Para estos valores de δ y p el valor crítico de μ es 0 con lo cual esta configuración corresponde a la región supercrítica (probabilidad estrictamente positiva de no extinción). En la Figura 3.2 se observan dos realizaciones del modelo (utilizando los mismos parámetros pero distintas *semillas* en la simulación), donde en un caso la población de parásitos se extingue y en la otra aumenta indiscriminadamente (“explota”, tomando como criterio que esto ocurre cuando la misma supera el umbral de 10^7 , de acuerdo con [12]).

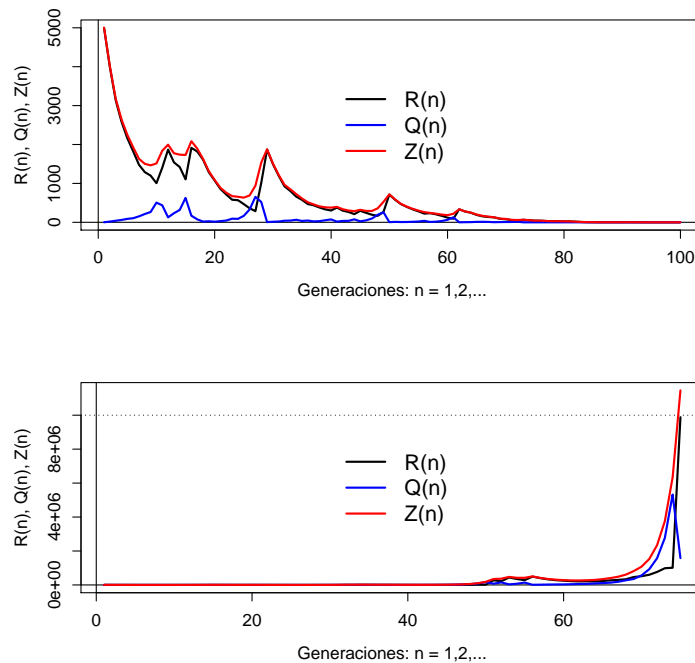


Figura 3.2: dos ejemplos, uno de *extinción* y otro de *explosión*, de distintas realizaciones del modelo simple de variación antigénica (mismos parámetros fijos), en la región supercrítica.

Autocontrol y ondas de parasitemia (parámetros aleatorios)

En la Figura 3.3 se muestra el resultado de una simulación del modelo con autocontrol utilizando los parámetros $\delta_{\text{mín}} = 0.6$, $\delta_{\text{máx}} = 0.95$, $r_{\delta} = 10^4$, $r_{\mu} = 10^2$, $\mu_{\text{mín}} = 0.004$, $\mu_{\text{máx}} = 0.05$ y $p = 0.65$. Se observa en este ejemplo, a diferencia de lo que ocurre con el modelo de parámetros fijos, que cuando la población de bacterias alcanza valores extremos (altos o bajos) surgen los mecanismos de autocontrol que evitan (o al menos demoran) la extinción o explosión poblacional. En este sentido, el modelo es capaz de reproducir el comportamiento de algunas poblaciones de parásitos que siguen una estrategia de variación antigénica en su interacción con el sistema inmune.

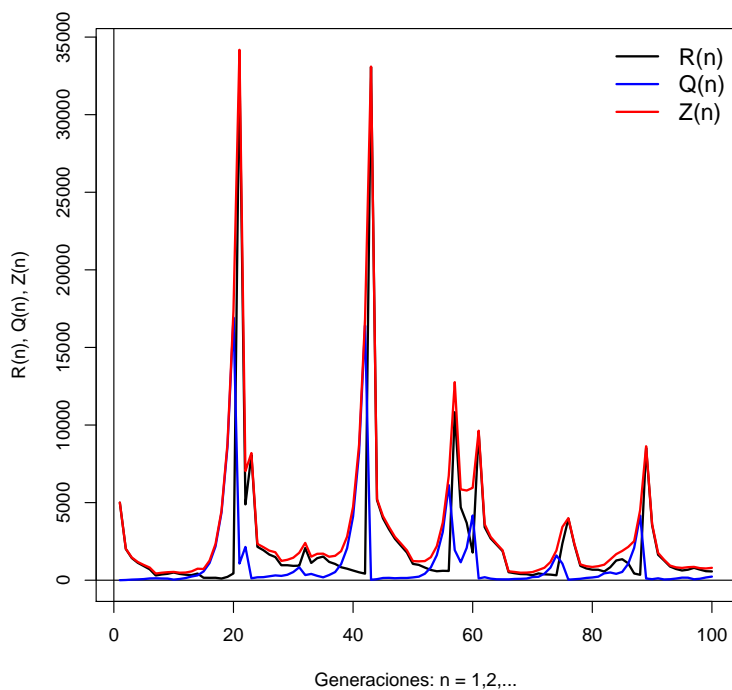


Figura 3.3: ejemplo de realización del modelo de variación antigénica con autocontrol (parámetros aleatorios).

3.5. Discusión

Los autores del modelo presentado en este capítulo destacan que el mismo considera diversos aspectos entre los cuales destacan:

- eficiencia de la respuesta inmune contra las células sensibles;
- tasa de aparición de nuevas variantes resistentes;
- muerte natural de células resistentes y
- tiempo aleatorio de demora del sistema inmune para reconocer una nueva variante.

Además, para la versión más simple del modelo donde los parámetros son fijos, se caracteriza analíticamente las regiones críticas, subcríticas y supercríticas en los correspondientes espacios paramétricos. Esto les permite concluir que, bajo su modelo, los fenómenos ondulatorios observados en los procesos de variación antigénica (denominados *ondas de parasitemia*) no pueden ser explicados únicamente por la interacción entre la producción de nuevas variantes de parásitos y la respuesta inmune que intenta eliminarlas. Esto debido a que el sistema es muy inestable y se extingue o crece en forma exponencial. Para que este fenómeno no suceda es necesario introducir mecanismos de autoregulación poblacional, mediante la consideración de parámetros aleatorios que dependen del estado del proceso (tamaños poblacionales). Por ejemplo, destacan que el modelo puede ser extendido para incluir situaciones de mayor complejidad y realismo, permitiendo que tanto la efectividad de la respuesta inmune como la demora en su aparición dependan del tamaño de la población de células resistentes.

Capítulo 4

Comentarios finales

En biología la historia del objeto de investigación importa mucho y no puede ser despreciada. Cada problema atacado es particular y casi único, y en consecuencia la existencia de leyes generales formalizadas en lenguaje matemático son prácticamente inexistentes... sus formas de atacar problemas parecen, si no contradictorias, sí distintas: las matemáticas enfatizan el rigor y la precisión, sus resultados definitivos no cambian y se mantienen sólidos por miles de años... Por el contrario, en biología nuestro conocimiento de los sistemas que estudiamos es muy reciente y las teorías biológicas evolucionan de manera rápida.

Este planteamiento inmediatamente indica la necesidad del trabajo multi- e interdisciplinario, que integre especialistas de diferentes disciplinas biológicas y matemáticas que sean capaces de subordinar sus intereses disciplinarios naturales a la resolución de un problema común de naturaleza biológica... Debo subrayar que el potencial científico de las interacciones entre matemáticas y biología puede ser desarrollado únicamente si se le fomenta y promueve entre nuestros estudiantes tanto de matemáticas como de biología.^a

^aTomado de [34].

De acuerdo con [22], muchas preguntas biológicamente importantes no se ocupan principalmente de los mecanismos moleculares del reconocimiento del sistema inmune sino con la dinámica poblacional de la respuesta del mismo. Tales preguntas no pueden ser por lo general respondidas solamente mediante métodos experimentales y requieren la ayuda de modelos matemáticos. Estos difieren de las teorías verbales en el hecho de dar una conexión explícita y precisa entre los supuestos y las conclusiones. Entonces, el hecho de formular un modelo obliga a explicitar los supuestos asumidos y a realizarse preguntas que de otro modo quizá podrían pasarse por alto.

Luego de realizar una introducción histórica y biológica breve del fenómeno de variación antigénica en el primer capítulo, en los capítulos siguientes se

presentaron algunos de los esfuerzos principales que se han desarrollado para la modelización de la dinámica de parásitos, bacterias y virus, que siguen una estrategia de variación antigénica con el objetivo de evadir la respuesta inmune del hospedero. Lo expuesto en este trabajo de tesis no pretende ser una presentación exhaustiva sobre el tema sino más bien una muestra de algunos de sus principales metodologías. Además de plantear y desarrollar cada uno de los modelos siguiendo lo expuesto por sus autores, se ha intentado justificar los resultados que se consideraron más relevantes y que en varios casos no estaban desarrollados en los respectivos artículos originales. Al mismo tiempo, se presentaron simulaciones computacionales propias que permiten ayudar en la interpretación y alcance de los modelos.

Por otro lado, no se han considerado algunos aspectos importantes pensando en la aplicación de estos modelos a casos reales particulares. Por ejemplo, dado uno modelo (como alguno de los presentados en capítulos previos) y datos sobre la dinámica de algún parásito o similar que sigue una estrategia de variación antigénica, ¿cómo *estimar* los parámetros del modelo de manera que el *ajuste* sea el mejor posible en algún sentido? O más en general, ¿cómo evaluar y comparar en estos casos el ajuste de diversos modelos a un mismo conjunto de datos reales? Estos temas, vinculados con lo que habitualmente se hace desde la *inferencia estadística*, no han sido considerados en este trabajo. Ello se debe, principalmente, a que intentar hacerlo extendería considerablemente el alcance del trabajo ya que los propios autores de los modelos presentados no han expuesto la forma de abordar estos problemas.

Durante los distintos capítulos fueron apareciendo varios temas y subtemas de la Matemática, como *ecuaciones diferenciales* y *procesos de ramificación* (o más en general procesos estocásticos), que a veces pueden parecer algo alejados entre sí, quizá porque así se presenten en la formación de estudiantes y porque a menudo comunidades o grupos distintos se dediquen a cada uno de ellos, en ocasiones con poca interacción. Posiblemente la necesidad o exigencia de la investigación especializada en Matemática también lleve a eso, trazando algunas fronteras que vistas desde otras disciplinas, como la Biología, usuarias de la Matemática como herramienta para la modelización y comprensión de sus objetos de estudio, carecen de sentido.

Uno de los objetivos en este trabajo fue el de partir de un “problema biológico” (el de variación antigénica) e ir recorriendo los principales aborda-

jes matemáticos del mismo y, con ello, describir las herramientas utilizadas a medida que fueran apareciendo, tratando de ignorar estas fronteras. Estos son solo ejemplos de los principios planteados al comienzo del capítulo sobre posibles vínculos entre la Biología y la Matemática, dos disciplinas que se encuentran en creciente interacción.

Claro que existe un conjunto mucho más amplio de conexiones, como puede verse, por ejemplo, en [2, 6, 14, 24, 34]. Como menciona el sugerente título en [15] y retomando el epígrafe del primer capítulo, (para la Matemática) la *Biología es la nueva Física*.

Bibliografía

- [1] Alizon, Samuel; Magnus, Carsten. *Modelling the Course of an HIV Infection: Insights from Ecology and Evolution*, Viruses, 4, 1984-2013, 2012.
- [2] Allen, Linda. *An Introduction to Stochastic Processes with Applications to Biology*, CRC Press, 2nd edition, 2010.
- [3] Anderson, Roy M. *Complex dynamic behaviours in the interaction between parasite populations and the host's immune system*, International Journal of Parasitology, 28, 551-566, 1998.
- [4] Antia, Rustom; Nowak, Martin; Anderson, Roy. *Antigenic variation and the within-host dynamics of parasites*, Proceedings of the National Academy of Sciences, Volume 93, pages 985-989, 1996.
- [5] Barry, J.; Turner, C. *The Dynamics of Antigenic Variation and Growth of African Trypanosomes*, Parasitology Today, Volume 7, number 8, 1991.
- [6] Brauer, Fred; Castillo-Chavez, Carlos. *Mathematical Models in Population Biology and Epidemiology*, Springer, 2nd edition, 2012.
- [7] Dobrow, Robert. *Introduction to Stochastic Processes with R*, John Wiley & Sons, 2016.
- [8] Edwards, C. Henry; Penney, David. *Ecuaciones Diferenciales*, Prentice Hall, 2a edición, 2001.
- [9] Feller, William. *Introducción a la teoría de probabilidades y sus aplicaciones*, Volumen I, 1973.
- [10] Gelman, Andrew; Carlin, John; Stern, Hal; Dundson, David; Vehtari, Aki; Rubin, Donald. *Bayesian Data Analysis*, 3rd edition, 2013.

- [11] Guerberoff, Gustavo. *Un modelo microscópico para la dinámica de la variación antigénica*, I Jornadas de Estadística Aplicada, homenaje a Ariel Roche, La Paloma, 25 de octubre de 2013.
- [12] Guerberoff, Gustavo; Alvarez-Valin, Fernando. *A stochastic microscopic model for the dynamics of antigenic variation*, Journal of Theoretical Biology, 380, 489-498, 2015.
- [13] Grossman, Stanley; Flores, José. *Algebra Lineal*, McGrawHill, 7a edición, 2012.
- [14] Haccou, Patsy; Jagers, Peter; Vatutin, Vladimir. *Branching Processes: Variation, Growth, and Extinction of Populations*, Cambridge University Press, 2005.
- [15] Hunter, Philip. *Biology is the new physics*, EMBO reports, Science and Society, 11(5), 350-352, 2010.
- [16] Karlin, Samuel; Taylor, Howard. *A first course in Stochastic Processes*, Academic Press, 2nd edition, 1975.
- [17] May, Robert; Stekel, Dov; Nowak, Martin. *Antigenic Diversity Thresholds and Hazard Functions*, Mathematical Biosciences, 139, 59-68, 1997.
- [18] Mugnier, Monica; Cross, George; Papavasiliou, Nina. *The in vivo dynamics of antigenic variation in Trypanosoma brucei*, Science, 347, 1470, 2015.
- [19] Nowak, Martin; May, Robert; Anderson, Roy. *The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease*, AIDS, 4:1095-1103, 1990.
- [20] Nowak, Martin; May, Robert. *Mathematical Biology of HIV Infections: Antigenic Variation and Diversity Threshold*, Mathematical Biosciences, 106:1-21, 1991.
- [21] Nowak, Martin; May, Robert. *AIDS pathogenesis: mathematical models of HIV and SIV infections*, AIDS, Volume 7 (suppl. 1), 1993.
- [22] Nowak, Martin; Bangham, Charles. *Population Dynamics of Immune Responses to Persistent Viruses*, Science, Volume 272, 1996.

- [23] Nowak, Martin; May, Robert. *Virus dynamics. Mathematical principles of immunology and virology*, Oxford University Press, 2000.
- [24] Nowak, Martin. *Evolutionary Dynamics. Exploring the equations of life*, Belknap Press, 2006.
- [25] Paenza, Adrián. *Matemática...¿Estás ahí?*, Siglo XXI editores, 2005.
- [26] Perelson, Alan; Ribeiro, Ruy. *Modeling the within-host dynamics of HIV infection*, BMC Biology, 11:96, 2013.
- [27] R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>, 2015.
- [28] Ross, Ronald; Thomson, David. *A Case of Sleeping Sickness Studied by Precise Enumerative Methods: Regular Periodical Increase of the Parasites Disclosed*, Proceedings of the Royal Society of London, serie B: 82, 411-415, 1910.
- [29] Savill; N., Seed; J. *Mathematical and statistical analysis of the Trypanosoma brucei slender to stumpy transition*, Parasitology, 128, 53-67, 2004.
- [30] Stilianakis, Nikolaos; Schenzle, Dieter; Dietz, Klaus. *On the Antigenic Diversity Threshold Model for AIDS*, Mathematical Biosciences, 121, 235-247, 1994.
- [31] Soetaert, Karline; Petzoldt, Thomas; Setzer, Woodrow. *Solving Differential Equations in R: Package deSolve*, Journal of Statistical Software, Volume 33, Issue 9, 2010.
- [32] Soetaert, Karline; Cash, Jeff; Mazzia, Francesca. *Solving Differential Equations in R*, Springer, 2012.
- [33] Stewart, James. *Cálculo: trascendentes tempranas*, Sexta edición, 2008.
- [34] Velasco Hernandez, Jorge. *El gen, la forma, el virus y la idea: una perspectiva personal de la biología matemática*, Miscelanea Matematica, 32, 5-38, 2000.
- [35] Wodarz, Dominik; Nowak, Martin. *Mathematical models of HIV pathogenesis and treatment*, BioEssays, 24:1178-1187, 2002.