



UNIVERSIDAD AUTÓNOMA DE GUERRERO

Unidad Académica de Matemáticas

Maestría en Matemáticas Aplicadas

**Comparación de dos modelos
dinámicos del virus de hepatitis B
mediante inferencia bayesiana**

T E S I S

PARA OBTENER EL GRADO DE:

Maestro en Matemáticas Aplicadas

PRESENTA:

Laima Imbert Rodríguez

DIRECTOR DE TESIS:

Dr. Ramón Reyes Carreto.

Msc. Cruz Vargas De León

Dedicatoria

A la memoria de mi abuelo Lázaro Martí, que con gusto se hubiera leído también mi segunda tesis. A la memoria de mi abuela paterna Zenaida, que se sentiría orgullosa de mi, por haber seguido sus pasos. A mis padres que gracias a ellos estoy cumpliendo otro de mis sueños. A mi hermana, que siempre trato de ser mejor por ella. A mi abuela Nilda que por ella me levanto todos los días con más ganas de vivir. A toda mi familia. A mis compañeros de maestría que me apoyaron durante todo este tiempo. A mis profesores por su valiosa enseñanza. A todos aquellos que creyeron en mí dedico esta tesis.

Agradecimientos

A mis padres, mis grandes amores, por haber siempre confiado en mí y por todo lo que han hecho en pos de mi formación profesional.

A Laura mi pequeña gigante, por ser tan especial.

A mi abuela Nilda, mi más valioso tesoro que hizo de mí la persona que soy hoy.

A toda mi familia por quererme tanto.

A mis compañeros de maestría Dannier, Edgar, Eduard, Noemí y Omar por todo su apoyo incondicional.

A los amigos que me acompañaron en este magnífico viaje Ely, Yadira, Enisdey, Alexey, Alian, Jose Miguel y José Luis.

A mis tutores, Ramón Reyes Carreto y Cruz Varga De León, por su dedicación y por permitirme trabajar a su lado.

A mis profesores de la Maestría de Matemática Aplicada por sus consejos y valiosa enseñanza.

A todas las personas lindas de México que hicieron posible mi superación profesional.

A la Universidad Autónoma de Guerrero por darme la oportunidad de continuar mis estudios.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por haberme brindado el apoyo económico necesario para realizar mi posgrado.

A todo aquel que de una forma u otro aportó su granito de arena para que este sueño se haya hecho realidad.

A todos ustedes mil gracias.

Resumen

La hepatitis B es una infección hepática potencialmente mortal causada por el virus de la hepatitis B (VHB). La organización mundial de la salud señala que hay 240 millones de personas que padecen infección crónica causada por el VHB. En epidemiología, existe un importante uso de modelos de ecuaciones diferenciales para describir de manera concisa fenómenos complejos de ciertas enfermedades usando parámetros con interpretación biológica. En las últimas décadas ha crecido el interés por modelar la dinámica viral a través de ecuaciones diferenciales, y la integración de datos clínicos, como la carga viral, a estos modelos.

Este trabajo tiene como objetivo *comparar los modelos dinámicos de infección viral básico y estandarizado, mediante el enfoque bayesiano, utilizando mediciones de carga viral en un grupo de pacientes con VHB, con y sin efecto de terapia a través de la aplicación de lamivudina*. En esta metodología estadística se usa la función de pérdida cuadrática para determinar el estimador Bayes y los intervalos creíbles para los parámetros de interés. Se determinan los números reproductivos básicos de infección de los modelos con y sin efecto de terapia. Por último, se determina el mejor modelo utilizando el criterio de información de la devianza y el factor Bayes.

Abstract

Hepatitis B is a life-threatening liver infection caused by the hepatitis B virus (HBV). The world health organization points out that there are 240 million people suffering from chronic infection caused by HBV. In epidemiology, there is an important use of differential equation models to describe the form of the complex phenomena of certain diseases with the parameters with the biological interpretation. In recent decades, interest in viral dynamics has increased through differential equations, and the integration of clinical data, such as viral load, with these models.

The objective of this work is to compare the models of basic and standardized viral infection, using the Bayesian approach, the use of viral load in a group of patients with HBV, and the effect of therapy through the application of lamivudine. In this statistical methodology, the quadratic loss function is used to determine the Bayes estimator and the credible intervals for the parameters of interest. The basic reproductive numbers of the infection of both models are determined, with and without therapy effect. Finally, the best model is determined using the information criterion of the deviance and the Bayes factor.

Índice general

1. Introducción	1
1.1. La hepatitis B	1
1.2. Replicación viral	3
1.3. Carga viral	3
1.4. Problemática	4
1.5. Objetivos	5
2. Marco Teórico	7
2.1. Modelo simple de decaimiento viral	7
2.2. Modelo básico de infecciones virales	8
2.3. Modelo de infecciones virales estandarizado	10
2.4. Métodos de solución numérica	11
2.4.1. Runge-Kutta de segundo orden	11
2.4.2. Runge-Kutta de cuarto orden	13
2.4.3. Algoritmo de Runge-Kutta para sistemas de ecuaciones diferenciales	14
2.5. Análisis bayesiano	15
2.5.1. Inferencia bayesiana	15
2.5.2. Cadenas de Markov Monte Carlo	18
2.5.3. Estimadores bayesianos	19
2.5.4. Intervalos creíbles	20
2.5.5. Diagnósticos de convergencia	20
2.5.6. Inferencia bayesiana para ecuaciones diferenciales	22
2.5.7. Selección de modelos	23

3. Metodología	26
3.1. Modelo básico de infecciones virales con efecto de terapia	27
3.2. Modelo básico de infecciones virales post-terapia	28
3.3. Modelo estandarizado de la infección del VHB con efecto de terapia	28
3.4. Modelo estandarizado de infecciones virales post-terapia	29
3.5. Datos del tratamiento de lamivudina	29
3.6. Condiciones iniciales	30
3.7. Valores iniciales de los parámetros para los modelos de la dinámica de infección del VHB	31
3.8. Solución numérica	32
3.9. Estimación de parámetros	33
3.9.1. Distribuciones a priori	33
3.9.2. Distribuciones a posteriori	34
3.9.3. Estimaciones de cantidades relacionadas a la dinámica viral.	35
4. Resultados y discusión	36
4.1. Estimación de parámetros del modelo básico de infecciones virales con terapia	36
4.1.1. Distribución a priori de los parámetros del modelo básico y estan- darizado, con terapia.	37
4.1.2. Estimaciones bayesianas de los parámetros del modelo básico con terapia	37
4.1.3. Diagnósticos de convergencia	39
4.2. Estimación de parámetros del modelo básico de infecciones virales post- terapia	41
4.2.1. Distribución a priori de los parámetros de los modelo básico y es- tandarizado, post-terapia.	41
4.2.2. Estimaciones bayesianas de los parámetros del modelo básico post- terapia	43
4.2.3. Diagnósticos de convergencia	44
4.3. Estimación de parámetros del modelo estandarizado con terapia	45

4.3.1. Estimaciones bayesianas de los parámetros del modelo estandarizado con terapia	45
4.3.2. Diagnósticos de convergencia	46
4.4. Estimación de parámetros del modelo estandarizado post-terapia	47
4.4.1. Estimaciones bayesianas de los parámetros del modelo estandarizado post-terapia	48
4.4.2. Diagnósticos de convergencia	49
4.5. Estimaciones del número reproductivo básico de la infección	50
4.6. Selección de modelo	51
4.7. Discusión	53
4.7.1. Comparación de resultados	53
5. Conclusiones	56
Anexos	61

Índice de figuras

1.1. Ciclo de replicación viral de la hepatitis B.	4
2.1. Diagrama básico de infección viral.	8
3.1. Esquema que describe la metodología de trabajo.	27
4.1. Distribución a priori de los parámetros β y κ para los modelos (básico y estandarizado) con tratamiento.	38
4.2. Distribución a priori de los parámetros m , n y σ para los modelos (básico y estandarizado) con tratamiento.	39
4.3. Distribución a priori de los parámetros β , κ y σ para los modelos (básico y estandarizado) post-terapia.	42
4.4. Distribución del número reproductivo básico de la infección para los modelos estandarizado y básico, con terapia.	51
4.5. Distribución del número reproductivo básico de la infección para los modelos estandarizado y básico, post-terapia.	52
4.6. Ajuste entre los datos de carga viral y la solución numérica utilizando las estimaciones bayesianas de los modelos estandarizado y básico, con terapia.	55
4.7. Ajuste entre los datos de carga viral y la solución numérica utilizando las estimaciones bayesianas de los modelos estandarizado y básico, post-terapia.	55
1. Trazas y densidades del modelo básico con tratamiento.	61
2. Trazas y densidades del modelo básico con tratamiento.	62
3. Distribuciones a posteriori de los parámetros del modelo básico con tratamiento.	62

4.	Gráfico de diagnóstico de convergencia de Gelman-Rubin de las estimaciones para los parámetros del modelo básico con tratamiento.	63
5.	Gráfico de diagnóstico de convergencia de Geweke para el modelo básico con tratamiento.	64
6.	Trazas y densidades del modelo básico post-terapia.	65
7.	Distribuciones a posteriori de los parámetros del modelo básico post-terapia.	66
8.	Gráfico de diagnóstico de convergencia de Gelman-Rubin de las estimaciones para los parámetros del modelo básico post-terapia.	67
9.	Gráfico de diagnóstico de convergencia de Geweke para el modelo básico post-terapia.	68
10.	Trazas y densidades del modelo estandarizado con tratamiento.	69
11.	Trazas y densidades del modelo estandarizado con tratamiento.	70
12.	Distribuciones a posteriori de los parámetros del modelo estandarizado con tratamiento.	70
13.	Gráfico de diagnóstico de convergencia de Gelman-Rubin de las estimaciones para los parámetros del modelo estandarizado con tratamiento. . .	71
14.	Gráfico de diagnóstico de convergencia de Geweke para el modelo estandarizado con tratamiento.	72
15.	Trazas y densidades del modelo estandarizado post-terapia.	73
16.	Distribuciones a posteriori de los parámetros del modelo estandarizado post-terapia.	74
17.	Gráfico de diagnóstico de convergencia de Gelman-Rubin de las estimaciones para los parámetros del modelo básico post-terapia.	75
18.	Gráfico de diagnóstico de convergencia de Geweke para el modelo estandarizado post-terapia.	76

Índice de tablas

2.1. Interpretación del FB según Jeffreys.	25
3.1. Respuesta a la terapia y el nivel del virus que vuelve después de que se detuvo la terapia.	29
4.1. Distribuciones a priori de los parámetros del modelo básico y estandarizado, ambos con tratamiento.	37
4.2. Estimaciones bayesianas del modelo Básico con efecto de tratamiento. . .	40
4.3. Diagnóstico de Gelman-Rubin de las estimaciones para los parámetros del modelo básico con terapia.	40
4.4. Diagnóstico de Geweke de las estimaciones para los parámetros del modelo básico con terapia.	41
4.5. Distribuciones de los parámetros del modelo básico y estandarizado, ambos en la post-terapia.	42
4.6. Estimaciones bayesianas del modelo básico post-terapia.	43
4.7. Diagnóstico de Gelman-Rubin de las estimaciones para los parámetros del modelo básico post-terapia.	44
4.8. Diagnóstico de Geweke de las estimaciones para los parámetros del modelo básico post-terapia.	45
4.9. Estimaciones bayesianas para los parámetros del modelo estandarizado con tratamiento.	46
4.10. Diagnóstico de convergencia de las estimaciones para los parámetros del modelo estandarizado con tratamiento.	47

4.11. Diagnóstico de Geweke de las estimaciones para los parámetros del modelo estandarizado con terapia.	47
4.12. Estimaciones bayesianas para los parámetros del modelo estandarizado post-terapia.	48
4.13. Diagnóstico de Gelman-Rubin de las estimaciones para los parámetros del modelo básico post-terapia.	49
4.14. Diagnóstico de Geweke de las estimaciones para los parámetros del modelo estandarizado post-terapia.	49
4.15. Estimaciones del número reproductivo básico de la infección, para los modelos básico y estandarizado, con terapia.	50
4.16. Estimaciones del número reproductivo básico de la infección, para los modelos básico y estandarizado, post-terapia.	51
4.17. Criterios de comparación de los modelos estandarizado y básico, con terapia.	52
4.18. Criterios de comparación de los modelos estandarizado y básico, post-terapia.	53
4.19. Estimaciones de los parámetros de los modelos estandarizado y básico, con terapia.	54
4.20. Estimaciones de los parámetros de los modelos estandarizado y básico, post-terapia.	54

Introducción

En la biología matemática, existe un amplio uso de modelos de ecuaciones diferenciales para describir fenómenos complejos de enfermedades infecciosas usando parámetros con interpretación biológica. Este trabajo se enfocó en el estudio en una de las enfermedades que más afecta a la población mundial, es el caso de la hepatitis B.

1.1. La hepatitis B

La hepatitis B es una infección hepática potencialmente mortal causada por el virus de la hepatitis B (VHB). La enfermedad se transmite por contacto con sangre o fluidos corporales de una persona contagiada. Este virus infecta a las células del hígado, que puede dar lugar tanto a un cuadro agudo como a una enfermedad crónica. La organización mundial de la salud (OMS) señala que hay 240 millones de personas que padecen infección crónica por el VHB. Más de 686,000 personas mueren cada año como consecuencia de la hepatitis B, incluido por cirrosis y cáncer hepático. En 2015 se registraron en México 24 mil casos de hepatitis, de éstos cerca de 1,750 casos de hepatitis B (WHO, 2018).

¿Qué es el hígado?

El hígado es un importante órgano que está presente tanto en el ser humano, como en los animales vertebrados. Está situado en la parte superior derecha del abdomen, debajo del diafragma, segrega la bilis esencial para la digestión de las grasas y cuenta con otras

muchas funciones, entre ellas la síntesis de proteínas plasmáticas, almacenamiento de vitaminas, glucógeno y función desintoxicante.

Para diagnosticar un problema del hígado se realiza una biopsia cuando es difícil determinar un problema del hígado mediante pruebas sanguíneas o técnicas de imagenología, como el ultrasonido y los rayos X. Es más común que la biopsia del hígado se realice para calcular el nivel de daño hepático, un proceso llamado estadificación. La estadificación ayuda a guiar el tratamiento. La biopsia del hígado es un procedimiento que consiste en extirpar una pequeña muestra del hígado para examinarlo bajo un microscopio y analizar si hay daño o enfermedad. Los tres tipos principales de biopsia del hígado son la percutánea, la transvenosa y la laparoscópica (Ripoll & Bañares, 2012).

Biopsia hepática percutánea

La técnica más comúnmente usada para recolectar una muestra de hígado es la biopsia hepática percutánea. Para este método, se introduce una aguja hueca a través del abdomen y dentro del hígado para extirpar un pequeño pedazo de tejido. La biopsia percutánea tiene como finalidad la toma de una muestra de una lesión localizada en una víscera o estructura ósea determinada para analizarla y obtener un diagnóstico de posible enfermedad.

Biopsia hepática transvenosa

La biopsia hepática transvenosa se utiliza cuando la sangre de una persona coagula con lentitud o cuando existe exceso de líquido en el abdomen, una alteración llamada ascitis. Se hace una pequeña incisión en el cuello y se introduce en la vena yugular una vaina, que es un tubo hueco de diseño especial.

Biopsia hepática laparoscópica

Los médicos usan una biopsia hepática laparoscópica para obtener una muestra de tejido de una o varias zonas específicas del hígado o cuando existe el riesgo de propagación de cáncer o infección. La cirugía laparoscópica es una técnica que evita el hacer una incisión grande, al realizar una o pocas incisiones pequeñas.

La biopsia percutánea es el mejor medio para la evaluación del daño hepático, pero su aplicación presenta diversos inconvenientes: es cara, no es fiable, es dolorosa, y no está exenta de riesgo a enfermedades oportunistas.

1.2. Replicación viral

La infección comienza con el virus adhiriéndose a la superficie celular. La envoltura del virus se fusiona con la membrana celular, liberando el núcleo al citoplasma de la célula. Posteriormente, las proteínas del núcleo se separan de la doble cadena de ácido desoxirribonucleico (ADN) y el genoma viral se desplaza hacia el interior del núcleo. Luego la enzima polimerasa viral completa los fragmentos que hacen falta, hasta generar la cadena doble de ADN circular cerrado de modo covalente (cccADN). Luego, comienza la transcripción del ADN viral mediante una enzima de ácido ribonucleico (ARN) polimerasa de la célula, para formar el ARN que servirá de plantilla para el ADN genómico viral y para los ARN mensajeros (ARNm), que darán origen a las proteínas virales. Estos ARN salen del núcleo y son traducidos como proteínas virales estructurales y como una cadena ARN pre-genómica, la cual es encapsulada por las proteínas del núcleo. Posteriormente, dentro del núcleo, la cadena de ARN viral es transcrita a una cadena de ADN por la misma enzima polimerasa viral que completó la cadena doble de ADN inicialmente. Las proteínas que constituyen el antígeno de superficie de la hepatitis B (HBsAg) son sintetizadas en el retículo endoplasmático rugoso de donde saldrán posteriormente las cápsides con ellas en su superficie. Finalmente, el virus es secretado fuera de la célula. En la Figura 1.1 se presenta un esquema de la replicación del virus de la hepatitis B (Collier & Oxford, 2008). Para el tratamiento VHB se emplea la lamivudina, este es un fármaco que se utiliza para el tratamiento de las infecciones víricas, es un análogo del nucleósido citidina que actúa inhibiendo la enzima transcriptasa inversa del virus de la inmunodeficiencia humana.

1.3. Carga viral

La carga viral es determinante en el estudio del paciente VHB crónico y en la evaluación de la eficacia del tratamiento. Un examen de biología molecular permite cuantificar el VHB que porta un paciente. La carga viral de VHB permite predecir y monitorear la respuesta al tratamiento antiviral

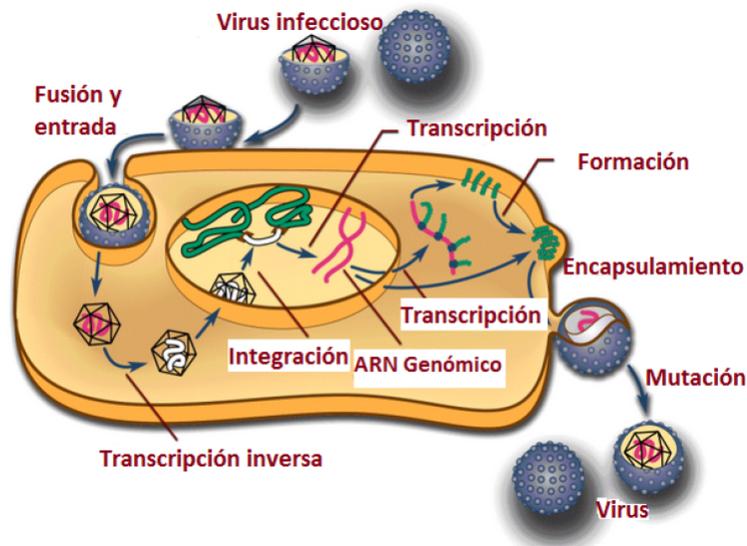


Figura 1.1: Ciclo de replicación viral de la hepatitis B.

1.4. Problemática

La eficacia del tratamiento antiviral de la infección de hepatitis B (VHB) puede analizarse mediante el uso de modelos matemáticos. Diversos modelos matemáticos se han desarrollado para dichos estudios. El modelo básico de infección viral (BVIM, por sus siglas en inglés) es utilizado ampliamente en estudios de la dinámica de infección del virus de hepatitis B. Este modelo supone que el proceso de infección sigue la ley de acción de masas¹ que se describe mediante un sistema de ecuaciones diferenciales ordinarias no lineales.

En el presente trabajo se abordan los modelos matemáticos denominados modelo BVIM y Modelo basado en incidencia estándar (ABVIM, por sus siglas en inglés) que, desde el punto de vista médico, mejora al modelo BVIM puesto que a diferencia de éste no supone que el tamaño de hígado pequeño puede ser más resistente a la infección del virus que un hígado de tamaño grande.

La biopsia percutánea ha resultado ser el mejor medio para evaluar el daño hepático de un individuo; sin embargo, es importante tomar en cuenta que este tipo de estudio resulta

¹La ley de acción de masas establece que para una reacción química reversible en equilibrio a una temperatura constante, una relación determinada de concentraciones de reactivos y productos tiene un valor constante.

poco conveniente porque es caro, doloroso, no es fiable y no está exento de riesgos a enfermedades oportunistas, es por esto que lo que se hace es monitorear la carga viral en sangre periférica para cuantificar el daño hepático.

Con este trabajo se determinó el porcentaje de efectividad del tratamiento de lamivudina² en enfermos de hepatitis B utilizando solamente datos clínicos y de laboratorio, mediante la medición de la carga viral total de individuos; es decir, sin necesidad de realizar ningún tipo de biopsia de hígado.

1.5. Objetivos

Objetivo general

Comparar los modelos dinámicos de infección viral básico y estandarizado, mediante el enfoque bayesiano, utilizando mediciones de carga viral en un grupo de pacientes con VHB, con y sin efecto de terapia a través de la aplicación de lamivudina.

Objetivos específicos

- Estimar los parámetros en los modelos BVIM y ABVIM, bajo tratamiento antiviral y el efecto post-terapia ³ .
- Obtener los intervalos creíbles para los parámetros de interés de ambos modelos.
- Determinar los números reproductivos básicos de infección de ambos modelos.
- Identificar el mejor modelo mediante la comparación de los criterios estadísticos.

Para lograr los anteriores objetivos, el presente trabajo se estructura de la siguiente manera: en el Capítulo 2 se aborda el marco teórico de los modelos BVIM y ABVIM; uno de los métodos de solución más común de ecuaciones diferenciales ordinarias no lineales,

²La lamivudina es un fármaco que se utiliza para el tratamiento de las infecciones víricas. Es un análogo del nucleósido citidina que actúa inhibiendo la enzima transcriptasa inversa del virus de la inmunodeficiencia humana, causante del SIDA. De esta manera impide que el virus se replique (Pascuet, Creus, Martín & del Cacho, 2011).

³Luego de retirado el tratamiento de lamivudina

el método Runge-Kutta; la inferencia bayesiana para la estimación de parámetros que utiliza los resultados de la solución de las ecuaciones diferenciales en la función de verosimilitud y distribuciones a priori adecuadas, así como algunas pruebas estadísticas de diagnóstico para asegurar la validez de los resultados obtenidos y de algunos criterios de selección de modelos. En el Capítulo 3, se describen los dos modelos bajo estudio; se presentan los datos del tratamiento de lamivudina, los valores de los parámetros resultados de la literatura, y las condiciones iniciales para resolver el problema de falta de solución analítica del sistema de ecuaciones diferenciales; las distribuciones a priori para los parámetros de interés, y finalmente, las distribuciones a posteriori obtenidas. En el Capítulo 4 se presentan los resultados de las estimaciones bayesianas obtenidas de los dos modelos empleados en el estudio, algunas pruebas de validez de los resultados, y al final se presenta una comparación y discusión entre ambos resultados. Finalmente, en el Capítulo 5 se presentan las conclusiones a las que se llegan en el presente trabajo.

Marco Teórico

En este Capítulo se enuncian los preliminares que fueron necesarios para encaminar este trabajo, se dan a conocer los modelos determinísticos que se utilizaron, el enfoque bayesiano, su inferencia, así como toda la teoría necesaria que fundamenta la investigación.

2.1. Modelo simple de decaimiento viral

El modelo más simple para describir el decaimiento de la carga viral en pacientes (Perelson & Nelson, 1999) es el siguiente:

$$\frac{dv(t)}{dt} = p - \mu v(t) \quad (2.1)$$

donde $v(t)$ es logaritmo base 10 de la concentración de partículas virales; p y μ son parámetros positivos, que determinan la producción y decaimiento viral, respectivamente. El modelo tiene como condición inicial $v(t_0) = v_0$ y al resolver mediante el método de separación de variables se logra la solución exacta:

$$v(t) = (p/\mu) + (v_0 - p/\mu) * \exp(-\mu * (t - t_0)). \quad (2.2)$$

A través de la solución anterior se obtuvo una ecuación para determinar la “vida media” ($t_{1/2}$) de la carga viral:

$$t_{1/2} = t_0 + \frac{1}{\mu} * \ln \frac{(v_0 - \frac{p}{\mu})}{(\frac{v_0}{2} - \frac{p}{\mu})}.$$

2.2. Modelo básico de infecciones virales

Una forma más detallada de analizar el ciclo de replicación viral es mediante el modelo BVIM el cual es ampliamente usado para estudiar la dinámica de la infección del VHB. Este modelo asume que el proceso de infección sigue la ley de acción de masas y está fundamentado en un sistema de tres ecuaciones diferenciales ordinarias no lineales.

En la Figura 2.1 las células susceptibles a la infección, conocida como células blanco (X), están infectadas por el virus (V) con una tasa constante β . Se supone que las células susceptibles proliferan a una tasa λ y mueren a una tasa δ por célula. La infección produce células infectadas (Y), que producen nuevos virus a una tasa κ , y mueren a una tasa α por célula. Los virus libres se eliminan a una tasa μ por virión ¹.

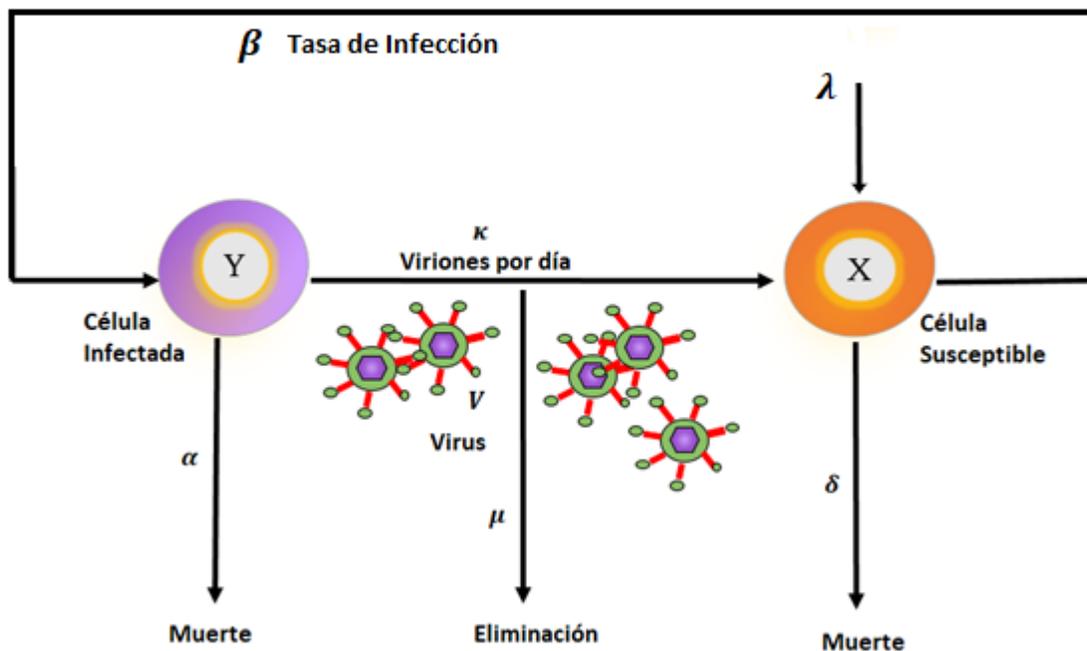


Figura 2.1: Diagrama básico de infección viral.

¹ Se denomina virión a la partícula vírica morfológicamente completa e infecciosa (Collier & Oxford, 2008).

El modelo matemático básico (BVIM) de ecuaciones diferenciales ordinarias para células no infectadas x , células infectadas y , y virus libre v , es el siguiente:

$$\begin{aligned}\frac{dx}{dt} &= \lambda - \delta x - \beta vx, \\ \frac{dy}{dt} &= \beta vx - \alpha y, \\ \frac{dv}{dt} &= \kappa y - \mu v,\end{aligned}\tag{2.3}$$

donde x , y y v son números de células no infectadas (susceptible), células infectadas y virus libres, respectivamente. Se supone que las células no infectadas se reproducen a una tasa constante λ , mueren a una tasa de δ , y se infectan a razón de β . Se tiene que β es una tasa constante que describe el proceso de infección. Las células infectadas se producen a una tasa de β y mueren a una tasa α . Las partículas virales se producen a partir de células infectadas a razón de κ y se eliminan a una tasa μ . Este modelo se utiliza para el estudio de la dinámica de los agentes infecciosos tales como hepatitis B. En el caso del Modelo (2.3), el número reproductivo básico de la infección:

$$R_0 = \frac{\lambda\beta\kappa}{\alpha\delta\mu}$$

es proporcional al tamaño del órgano $\frac{\lambda}{\delta}$ (Nowak, Bonhoeffer & McDade, 1996).

Teniendo en cuenta el R_0 los posibles escenarios de la infección viral son los siguientes:

- Si $R_0 \leq 1$, el modelo anterior tienen un punto de equilibrio:

$$E_f = \left(\frac{\lambda}{\delta}, 0, 0 \right)$$

- Si $R_0 > 1$, el modelo anterior tiene dos puntos de equilibrio:

$$E_f = \left(\frac{\lambda}{\delta}, 0, 0 \right)$$

$$E^* = \left(\frac{\alpha\mu}{\beta\kappa}, \frac{\lambda}{\alpha} \left(1 - \frac{1}{R_0} \right), \frac{\delta}{\beta} (R_0 - 1) \right)$$

Se conoce que el virus en el plasma tiene una vida media de 24 horas, que es el tiempo en el que la concentración viral se reduce a la mitad; la producción total es de 10^{11} y la carga viral de 2×10^{11} . En el caso de las células infectadas estas tienen una vida media de 10 a 100 días.

El número reproductivo básico de la infección del BVIM es proporcional al número total de células del hígado antes de la infección. Esto sugiere que el BVIM puede no ser un modelo razonable para la descripción de la infección del virus VHB ya que implica que un individuo con el hígado más pequeño tal vez sea más resistente a las infecciones de virus que un individuo con un órgano más grande. En consecuencia, el significado práctico de R_0 es biológicamente cuestionable; el BVIM puede no ser un modelo apropiado para la interpretación de infección por el virus VHB.

2.3. Modelo de infecciones virales estandarizado

El modelo matemático estandarizado ABVIM está dado por (Min, Su & Kuang, 2008):

$$\begin{aligned}\frac{dx}{dt} &= \lambda - \delta x - \beta v \frac{\mathbf{x}}{\mathbf{x}+\mathbf{y}}, \\ \frac{dy}{dt} &= \beta v \frac{\mathbf{x}}{\mathbf{x}+\mathbf{y}} - \alpha y, \\ \frac{dv}{dt} &= \kappa y - \mu v,\end{aligned}\tag{2.4}$$

Las variables x , y , v y los parámetros λ , δ , β , α , κ y μ son definidos de la misma forma que los del modelo básico de infección viral (BVIM). Sin embargo, aquí el número reproductivo básico de la infección R_0^* no depende del número total de células del hígado de la persona. En el caso del Modelo (2.4), el número reproductivo básico de la infección:

$$R_0^* = \frac{\beta \kappa}{\alpha \mu}.$$

Se observa que el R_0^* no depende del número total de células del hígado de la persona. Se tiene en cuenta que para el R_0^* los posibles escenarios de la infección viral son los siguientes :

- Si $R_0^* \leq 1$ entonces el modelo anterior tienen un punto de equilibrio:

$$E_f = \left(\frac{\lambda}{\delta}, 0, 0 \right)$$

- Si $R_0^* > 1$, el modelo anterior tiene dos puntos de equilibrio:

$$E_f = \left(\frac{\lambda}{\delta}, 0, 0 \right)$$

$$E^* = \left(\frac{\lambda}{\delta + \alpha(R_0^* - 1)}, \frac{\lambda(R_0^* - 1)}{\delta + \alpha(R_0^* - 1)}, \frac{\lambda \kappa (R_0^* - 1)}{\mu[\delta + \alpha(R_0^* - 1)]} \right)$$

Dado que los sistemas de ecuaciones diferenciales (2.3) y (2.4) no tienen solución exacta se usan métodos de solución numérica.

2.4. Métodos de solución numérica

Los métodos de Runge-Kutta emplean indirectamente el algoritmo de Taylor. Estos métodos evalúan $f(x, y)$ en más de un punto en la proximidad de (x_n, y_n) en lugar de evaluar derivadas de $f(x, y)$, las cuales se necesitarían para el uso directo del algoritmo por series de Taylor.

La derivación de estos métodos se acompaña de la suposición de un algoritmo particular con ciertos coeficientes indeterminados. Los valores de estos términos constantes se encuentran igualando la fórmula de Runge-Kutta de orden p al algoritmo de Taylor de orden p (Rodríguez, 1999).

2.4.1. Runge-Kutta de segundo orden

Los métodos de Runge-Kutta logran una exactitud del procedimiento de una serie de Taylor, sin requerir el cálculo de derivadas superiores. Probablemente uno de los procedimientos más difundidos, y a la vez más exactos, para obtener la solución numérica del problema de valor inicial. Se parte de:

$$y_{n+1} = y_n + h[k_1f(x_n, y_n) + k_2f(x_n + \alpha h, y_n + \beta hf(x_n, y_n))] \quad (2.5)$$

la cual es equivalente al algoritmo cuadrático de Taylor:

$$y_{n+1} = y_n + h \left[f(x_n, y_n) + \frac{h}{2} f'(x_n, y_n) \right] \quad (2.6)$$

Igualando las Ecuaciones (2.5) y (2.6) para resolver:

$$k_1f(x_n, y_n) + k_2f(x_n + \alpha h, y_n + \beta hf(x_n, y_n)) = f(x_n, y_n) + \frac{h}{2} f'(x_n, y_n) \quad (2.7)$$

y expandiendo el segundo término del lado izquierdo en la ecuación anterior en una expansión lineal de Taylor, obtenemos:

$$f(x_n + \alpha h, y_n + \beta h f(x_n, y_n)) = f(x_n, y_n) + \alpha h f_x(x_n, y_n) + \beta h f(x_n, y_n) f_y(x_n, y_n) \quad (2.8)$$

Y luego sustituimos la Ecuación (2.8) en la (2.7) para obtener:

$$\begin{aligned} k_1 f(x_n, y_n) + k_2 [f(x_n, y_n) + \alpha h f_x(x_n, y_n) + \beta h f(x_n, y_n) f_y(x_n, y_n)] \\ = f(x_n, y_n) + \frac{h}{2} [f(x_n, y_n) f_y(x_n, y_n)] \end{aligned}$$

donde el nuevo lado derecho se obtiene de reemplazar f' por $f_x + f f_y$ y así obtenemos tres ecuaciones:

$$\begin{aligned} k_1 + k_2 &= 1 \\ k_2 \alpha &= \frac{1}{2} \\ k_2 \beta &= \frac{1}{2} \end{aligned}$$

Dado que se tiene 4 parámetros por determinar y solo 3 ecuaciones, se tiene una familia de soluciones en términos de uno de ellos $(k_1, k_2, \alpha, \beta)$. Si hacemos $k_2 \neq 0$ como parámetro independiente, se obtiene la siguiente solución:

$$\begin{aligned} k_1 &= 1 - k_2 \\ \alpha &= \frac{1}{2k_2} \\ \beta &= \frac{1}{2k_2} \end{aligned}$$

y sustituyendo en la Ecuación (2.5), obtenemos:

$$y_{n+1} = y_n + \frac{h}{2} [f(x_n, y_n) + f(x_n + h, y_n + h f(x_n, y_n))]$$

Para simplificar un poco, podemos hacer $x_n + h = x_{n+1}$ y $f(x_n, y_n) = f_n$ y así simplificar la notación:

$$y_{n+1} = y_n + \frac{h}{2} [f_n + f(x_{n+1}, y_n + h f_n)]$$

que no es más que el método mejorado de Euler.

2.4.2. Runge-Kutta de cuarto orden

La fórmula utilizada por este método se emplea para calcular la nueva aproximación y_{n+1} (no olvidar que aquí la variable independiente es x y la variable dependiente es y):

$$y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

donde

$$\begin{aligned}k_1 &= hf(x_n, y_n) \\k_2 &= hf\left(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}\right) \\k_3 &= hf\left(x_n + \frac{h}{2}, y_n + \frac{k_2}{2}\right) \\k_4 &= hf(x_n + h, y_n + k_3)\end{aligned}$$

Este método se basa en una expansión de Taylor truncada hasta términos de cuarto orden. En algunas ocasiones es llamado el método de Kutta-Simpson porque se reduce a la Simpson si $y'(x)$ es independiente de y . Esto es, si $y'(x) = f(x)$, tenemos:

$$\begin{aligned}k_1 &= hf(x_n) \\k_2 &= hf(x_n + h/2) \\k_3 &= hf(x_n + h/2) \\k_4 &= hf(x_n + h)\end{aligned}$$

y

$$y_{n+1} = y_n + \frac{h}{6}[f(x_n) + 4f(x_n + h/2) + f(x_n, y_n)]$$

y si utilizamos un intervalo de $h^* = h/2$, podemos escribir:

$$y_{n+1} = y_n + \frac{h^*}{3}[f(x_n) + 4f(x_n + h^*) + f(x_n + 2h^*)]$$

Que es la Regla de Simpson $\frac{1}{3}$ para integración.

2.4.3. Algoritmo de Runge-Kutta para sistemas de ecuaciones diferenciales

Dado que ahora tenemos p ecuaciones diferenciales de orden 1, existirán p variables que denotarán la solución al sistema. Para lograr lo anterior, manejamos dos subíndices: el primero indicará la variable y el segundo indicará el número de la aproximación.

Considere el siguiente sistema de ecuaciones diferenciales (Rodríguez, 1999):

$$\begin{aligned} x_1(t) &= f_1(t, x_1, x_2, \dots, x_p) = f_1(t, x) \\ &\vdots \\ x_p(t) &= f_p(t, x) \end{aligned}$$

Las condiciones iniciales las podemos expresar de la siguiente forma:

$$\begin{aligned} x_1^{(0)} &= \alpha_1 \\ &\vdots \\ x_p^{(0)} &= \alpha_p \end{aligned}$$

con la fórmula recursiva

$$x^{(j+1)} = x^{(j)} + \frac{1}{6}[k_1 + 2k_2 + 2k_3 + k_4]$$

y las k toman la siguiente forma:

$$\begin{aligned} k_1 &= hf(t_j, x^{(j)}) \rightarrow k_1^{(i)} = hf_i(t_j, x^{(j)}) \\ k_2 &= hf\left(t_j + \frac{h}{2}, x^{(j)} + \frac{1}{2}k_1\right) \rightarrow k_2^{(i)} = hf_i\left(t_j + \frac{h}{2}, x^{(j)} + \frac{1}{2}k_1\right) \\ k_3 &= hf\left(t_j + \frac{h}{2}, x^{(j)} + \frac{1}{2}k_2\right) \rightarrow k_3^{(i)} = hf_i\left(t_j + \frac{h}{2}, x^{(j)} + \frac{1}{2}k_2\right) \\ k_4 &= hf\left(t_j + h, x^{(j)} + k_3\right) \rightarrow k_4^{(i)} = hf_i\left(t_j + h, x^{(j)} + k_3\right) \end{aligned}$$

donde $i = 1, \dots, p$. Conocer la solución exacta o numérica de sistemas de ecuaciones diferenciales es necesario para la estimación de parámetros tanto en el contexto clásico, como para el bayesiano.

2.5. Análisis bayesiano

La utilización de modelos bayesianos constituye un enfoque alternativo a problemas resueltos mediante métodos convencionales de inferencia. La estadística bayesiana combina la inferencia estadística y la teoría de decisión con incertidumbre, la misma introduce al modelo una distribución a priori $f(\theta)$, que expresa el conocimiento o ignorancia acerca del parámetro θ .

2.5.1. Inferencia bayesiana

La inferencia bayesiana, puede utilizarse para estimar parámetros, seleccionar modelos y hacer predicción usando las soluciones de los modelos de ecuaciones diferenciales e incluso cuando dichos modelos son no lineales y no tienen solución analítica.

Supongamos que observamos un proceso $y = (y_1, \dots, y_n)$ en los tiempos discretos $t_1, \dots, t_n \in [0, T[$ tal que

$$y_i = f(X_{\theta}(t_i)) + \varepsilon_i, \quad \varepsilon_i \sim_{i.i.d.} \mathcal{N}(0, \sigma^2),$$

donde X_{θ} es la solución del siguiente sistema de ecuaciones diferenciales ordinarias:

$$\frac{dX_{\theta}}{dt} = F(X_{\theta}, t, \theta), \quad X_{\theta}(t_0) = X_0 \quad (2.9)$$

$\theta \in A \subset \mathbb{R}^d$ es un vector de parámetros desconocidos, y $F : \mathbb{R}^d \times [0, T[\times A \mapsto \mathbb{R}^d$ es una función conocida. Tomamos $\sigma^2 \in S \subset \mathbb{R}^+$

Suponemos que la función F en el lado derecho de problema del valor inicial (2.9) sigue las condiciones de regularidad del teorema de Picard (Süli & Mayers, 2003). También se asume para el espacio paramétrico que A y S son conjuntos compactos.

En estadística bayesiana, la estimación, la predicción y la selección del modelo se basa en la función de verosimilitud:

$$P_{Y|\Phi}(y|\phi) = \sigma^{-n} (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - X_{\theta}(t_i))^2 \right\}$$

donde $\Phi = (\Theta, \Sigma)$ es una variable aleatoria con una realización particular $\phi = (\theta, \sigma)$. Esta expresión involucra el cálculo aproximado de X_θ , que es solución de (2.9).

En la práctica, el sistema (2.9) se resuelve utilizando un solucionador numérico y la inferencia se realiza, no en el modelo exacto anterior, pero si en un modelo aproximado,

$$y_i = f(X_\theta^h(t_i)) + \varepsilon_i, \quad \varepsilon_i \sim_{i.i.d.} \mathcal{N}(0, \sigma^2), \quad (2.10)$$

donde y_i son las observaciones en el tiempo t_i y $X_\theta^h(t_i)$ denota la solución aproximada de (2.9) proporcionada por la solución numérica (siendo h un parámetro de precisión de la solución, típicamente su tamaño de paso), luego se supone que:

$$y_i \sim \mathcal{N}(X_\theta^h(t_i), \sigma^2)$$

La verosimilitud derivada del modelo (2.10) se expresa de la siguiente manera:

$$P_{Y|\Phi}^h(y|\phi) = \sigma^{-n} (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - X_\theta^h(t_i))^2 \right\}$$

Existe una necesidad real de comprender y controlar el error cometido al trabajar con $P_{Y|\Phi}^h(y|\phi)$ en lugar de $P_{Y|\Phi}(y|\phi)$ (Capistrán, Christen & Donnet, 2016).

Para hacer inferencia sobre ϕ , generalmente se cuenta con información acerca de su valor, antes de ver los datos. Esto se representa en la distribución a priori.

Definición 2.5.1 La distribución a priori $P_\Phi(\phi)$ mide el grado de conocimiento inicial que se tiene de los parámetros en estudio.

La distribución a priori es una distribución conjunta en el caso de que ϕ tiene dimensión mayor a 1. Esto implica la búsqueda de esta distribución. Comúnmente se considera la distribución a priori como el producto de las distribuciones marginales, se supone los parámetros ϕ_i y ϕ_j , son independientes con $i \neq j$. Cuando el conocimiento inicial del parámetro ϕ puede ser muy vago, da lugar a distribuciones a priori no informativas.

Cuando se cuenta con base de datos pequeñas, la mayor información recae en las distribuciones a priori que se asigna a los parámetros. Es por esta razón la importancia de

investigar cual es la distribución indicada para cada parámetro. A menudo, un análisis bayesiano es equivalente a un análisis clásico con una distribución a priori no informativa, porque estas no van a ofrecer mucha información sobre los parámetros.

Definición 2.5.2 Una distribución de ϕ es no informativa si no contiene información sobre ϕ , es decir, no establece si unos valores de ϕ son más favorables que otros.

Las distribuciones informativas, son resultado de estudios empíricos previos, conocimiento del investigador por intervalos de estimación de momentos y supuesto de simetría.

Una distribución a priori pertenece a cierta clase o familia de distribuciones paramétricas D ; entonces diremos que esa distribución es conjugada respecto a la verosimilitud si la distribución a posteriori también pertenece a la clase D . La idea de la distribución conjugada es que la distribución a posteriori tiene la misma forma algebraica que la distribución a priori. Luego, está claro que si se elige una distribución a priori $P_{\Phi}(\phi)$ con la misma estructura de $P_{\Phi}^h(\phi|y)$, pensando en esto como función de ϕ entonces, la distribución a posteriori tendrá la misma forma.

Definición 2.5.3 Teniendo la creencia a priori de que la función de distribución de probabilidad es $P_{\Phi}(\phi)$ y de que una observación Y con verosimilitud $P_{Y|\Phi}^h(y|\phi)$ la probabilidad a posteriori es definida como $P_{\Phi|Y}^h(\phi|y) \propto P_{\Phi}(\phi)P_{Y|\Phi}^h(y|\phi)$.

La distribución posterior del modelo aproximado, es:

$$P_{\Phi}^h(\phi|y) = \frac{P_{Y|\Phi}^h(y|\phi)P_{\Phi}(\phi)}{P_Y^h(y)} \quad (2.11)$$

donde $P_Y^h(y) = \int P_{Y|\Phi}^h(y|\phi)P_{\Phi}(\phi)d\phi$ y ϕ es vector de parámetros.

Para generar muestras de las distribuciones a posteriori y estimar cantidades de interés se utilizan los métodos de cadenas de Markov Monte Carlo (MCMC). En los métodos MCMC se simulan valores de una densidad propuesta, que no tiene que ser necesariamente parecida a la densidad a posteriori. En muchos casos la distribución a posteriori se

hace muy compleja de determinar, ya que intervienen integrales múltiples de gran complejidad es por esta razón que se acude a algoritmos computacionales como el método de cadenas de Markov Monte Carlo.

2.5.2. Cadenas de Markov Monte Carlo

Las cadenas de MCMC son métodos de simulación para generar muestras de las distribuciones a posteriori y estimar cantidades de interés a posteriori. Son algoritmos que permiten obtener una muestra de una distribución de probabilidad f sin necesidad de simular directamente la distribución [Robert & Casella, 2010].

Una cadena de Markov es un proceso estocástico $\{\phi^{(1)}, \phi^{(2)}, \dots, \phi^{(T)}\}$ tal que

$$f(\phi^{(t+1)}|\phi^{(t)}, \dots, \phi^{(1)}) = f(\phi^{(t+1)}|\phi^{(t)})$$

es decir, la distribución de ϕ en el estado $t + 1$ dados todos los estados previos de ϕ depende solo del estado anterior, $\phi^{(t)}$. Además $f(\phi^{(t+1)}|\phi^{(t)})$ es independiente del tiempo t .

Con el fin de generar una muestra de $f(\phi)$ se desea construir una cadena de Markov con dos propiedades, la primera es que $f(\phi^{(t+1)}|\phi^{(t)})$ debe ser fácil de generar y la segunda es que el distribución estacionaria debe ser la distribución a posteriori de interés.

Suponiendo que hemos construido una cadena de Markov con estos requisitos, entonces el algoritmo para obtener la cadena MCMC consiste en:

1. Seleccionar un valor inicial.
2. Generar T resultados hasta generar la distribución estacionaria.
3. Monitorear la convergencia del algoritmo usando diagnósticos de convergencia. Si no se obtiene la convergencia, el diagnóstico lo detecta y se generan más observaciones.
4. Desechar las primeras B observaciones.

5. Considerar $\{\phi^{(B+1)}, \phi^{(B+2)}, \dots, \phi^{(T)}\}$
6. Graficar la distribución a posteriori(usualmente se pone atención en las distribuciones marginales univariadas)
7. Resumir la distribución a posteriori (media, mediana, desviación estándar, cuantiles)

2.5.3. Estimadores bayesianos

Después de obtener la distribución a posteriori, la estimación de ϕ se determina considerando la función de pérdida $L(\hat{\phi}, a)$.

Definición 2.5.4 Sean $f(\phi|y)$ la distribución a posteriori y $L(\phi, a)$ una función de pérdida que en términos generales, da el costo de decidir que el parámetro tiene el valor a , cuando de hecho es igual a ϕ . El estimador a puede ser elegido para minimizar la pérdida esperada posterior,

$$E[L(a|y)] = \int L(\phi, a)f(\phi|y)d\phi$$

La función de pérdida es difícil de especificar, ya que la pérdida no siempre es fácilmente medible. Algunas de las funciones de pérdida más utilizadas son:

- a) La función de pérdida cuadrática:

$$L(\phi, a) = (\phi - a)^2$$

produce como estimador bayesiano la media de la distribución a posteriori.

- b) La función de pérdida lineal absoluta:

$$L(\phi, a) = |\phi - a|$$

produce como estimador bayesiano la mediana de la distribución a posteriori.

c) La función de pérdida todo/nada:

$$L(\phi, a) = \begin{cases} 0 & \text{si } \phi = a \\ 1 & \text{si } \phi \neq a \end{cases}$$

produce como el estimador bayesiano la moda de la distribución a posteriori (Gelman et al., 2014).

2.5.4. Intervalos creíbles

En el enfoque bayesiano, la estimación por intervalos creíbles se define por una evaluación simple de las distribuciones a posteriori marginales de los parámetros.

Definición 2.5.5 Un intervalo de credibilidad al $100(1 - \alpha)\%$ para θ es un subconjunto C de Θ tal que

$$1 - \alpha \leq P(C|y) = \begin{cases} \int_C f(\theta|y) d\theta & \text{caso continuo} \\ \sum_{\theta \in C} f(\theta|y) & \text{caso discreto} \end{cases}$$

Así, la distribución a posteriori $f(\theta|y)$ permite calcular probabilidades en Θ , además admite hablar de que la probabilidad de θ esté en C . Al determinar un conjunto creíble para θ , una propiedad deseable es que la longitud del intervalo sea pequeña (Cepeda, Aguilar & Cervantes, 2008).

2.5.5. Diagnósticos de convergencia

Cuando se utilizan los métodos de MCMC se espera que las cadenas alcancen la convergencia a una distribución estacionaria que es la distribución posterior de interés. Sin embargo, esto no siempre ocurre, es por esto que se necesitan pruebas estadísticas para determinar la convergencia de las cadenas.

Diagnóstico de Gelman-Rubin

Al generar múltiples cadenas de Markov, cada una con condiciones iniciales diferentes, un diagnóstico de convergencia es el de Gelman y Rubin. Este propone una prueba de convergencia basada en 2 o más cadenas paralelas, cada una partiendo de diferentes valores iniciales que están sobre-dispersos con respecto a la verdadera distribución posterior. Su método se basa en una comparación entre las varianzas dentro y entre cadenas para cada variable (esencialmente un análisis clásico de varianza).

Los diagnósticos de Gelman y Rubin reportados por la librería *coda* del software R (Plummer, Best & Vines, 2006), son el 50 % y el 97,5 % cuantiles de la distribución de muestreo para este factor de reducción de escala potencial. Se simulan $m \geq 2$ cadenas de longitud $2n$ con valores iniciales sobredispersos. Estos cuantiles se estiman a partir de la segunda mitad de cada cadena, luego de realizar el quemado. Si el cuantil superior es aproximadamente 1, se puede diagnosticar la convergencia. En este caso, las estadísticas resumidas y las estimaciones de densidad pueden calcularse combinando el último 50 % de iteraciones de todas las cadenas. El estadístico de Gelman-Rubin es la reducción de escala potencial se estima de la siguiente forma:

$$\text{Estadístico de Gelman-Rubin } \sqrt{\hat{R}} = \sqrt{\frac{\widehat{Var}(\theta)}{W}}$$

$$\text{Varianza en las cadenas } W = \frac{1}{m(n-1)} \sum_{j=1}^m \sum_{i=1}^n (\theta_j^i - \bar{\theta}_j)^2$$

$$\text{Varianza estimada } \widehat{V}(\theta) = \left(1 - \frac{1}{n}\right) W + \frac{1}{n} B$$

$$\text{Varianza entre cadenas } B = \frac{n}{(m-1)} \sum_{j=1}^m (\bar{\theta}_j - \bar{\theta})^2$$

que se puede interpretar como la medición del factor por el cual la desviación estándar de θ podría reducirse al extender la cadena. El factor \hat{R} disminuye a 1 como la longitud de la cadena tiende al infinito, por lo que \hat{R} debería estar cerca de 1 si las cadenas han convergido aproximadamente a la distribución objetivo. Gelman sugiere que \hat{R} debe ser menor que 1.1 o 1.2 (Rizzo, 2007).

Cuando declaramos la convergencia aproximada, en realidad estamos concluyendo que

cada secuencia individual parece estacionaria y que las secuencias observadas se han mezclado bien entre sí.

Diagnóstico de Geweke

El diagnóstico de Geweke verifica la convergencia de la media de cada parámetro por separado de los valores de una sola cadena. En esta prueba se aplica una prueba de Z para comprobar si las medias estimadas a partir de dos submuestras diferentes de la muestra total, son iguales. Usualmente se compara el 10 % inicial y el 50 % final de la muestra total. Para la varianza asintótica de la media muestral se usa una estimación de la densidad de la teoría espectral. Para un conjunto de generado de valores $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(T)}$ de un parámetro θ de interés, se calcula la media de la muestra $\bar{\theta}$, seguida por el cálculo de la distribución espectral $S_{\theta}(\omega)$ para series de tiempo. Entonces el error estándar de la media está dado por $\sqrt{S_{\theta}(0)/T}$ y por lo tanto el estadístico Z se define como

$$Z = \frac{\bar{\theta}^A - \bar{\theta}^B}{\sqrt{\frac{S_{\theta}^A(0)}{T_A} + \frac{S_{\theta}^B(0)}{T_B}}}$$

Asintóticamente Z tiene distribución normal estándar, donde $\bar{\theta}^A, \bar{\theta}^B$ son las medias muestrales de las dos submuestras mencionadas, T_A y T_B son los tamaños, y $S_{\theta}^A(0)$ y $S_{\theta}^B(0)$ son las varianzas.

Los parámetros con $|Z| \geq 2$ tienen medias diferentes de las primeras y últimas iteraciones, y por lo tanto indican no convergencia de la cadena MCMC (Ntzoufras, 2009).

2.5.6. Inferencia bayesiana para ecuaciones diferenciales

La librería *deBIInfer* brinda un marco bayesiano para la inferencia paramétrica en ecuaciones diferenciales. Este enfoque ofrece una metodología rigurosa para la estimación de parámetros, así como para modelar el vínculo entre las condiciones iniciales y parámetros del modelo. Para realizar la inferencia utilizando *deBIInfer*, el usuario debe especificar las

estructuras de datos que representan el modelo de ecuaciones diferenciales, la verosimilitud y las distribuciones a priori para aquellos parámetros que se desean estimar. La librería *deBInfer* toma esa información y muestrea las distribuciones a posteriores usando cadenas de MCMC, resolviendo numéricamente el modelo de ecuaciones diferenciales dentro del procedimiento MCMC. El procedimiento MCMC para *deBInfer* ofrece realizaciones de Metropolis-Hastings independientes y aleatorias, y se implementa completamente en R (Boersch, Ryan & Johnson, 2017).

2.5.7. Selección de modelos

En la actualidad es posible resolver problemas de gran complejidad utilizando varios modelos estadísticos ya sean clásicos o bayesianos, lo que lleva a decidir cuál de todos es el modelo más adecuado. Para esto se tienen medidas y criterios que ayudan a determinar que modelo es mejor en el sentido de ajuste o de predicción.

Criterios para la selección de modelos

Desde un punto de vista frecuentista la devianza es la diferencia en log-verosimilitudes entre el modelo saturado y el modelo ajustado.

Definición 2.5.6 La devianza para la selección de modelos bayesianos se define como:

$$D(\phi_M, M) = -2 \log[f(y|\phi_M, M)] + 2 \log g(y) \quad (2.12)$$

donde y son los datos, ϕ_M el vector de parámetros bajo el modelo M , $f(y|\phi)$ es la función de verosimilitud y $\log g(y)$ denota un término de notación totalmente especificado por los datos.

Dempster propuso comparar la media posteriori de la devianza. Siguiendo esta idea Spiegelhalter et al. (2002) propuso el criterio de información de la devianza (DIC por sus siglas en inglés) basado en la distribución a posteriori de $D(\phi)$

Definición 2.5.7 El DIC se define como:

$$DIC(M) = D(\bar{\phi}_M, M) + 2d_M$$

donde $\bar{\phi}$ es la media posterior del parámetro en el modelo M .

La ventaja del DIC sobre otros criterios, en el caso de selección de modelos bayesianos, es que el DIC se calcula fácilmente a partir de las muestras generadas por una cadena MCMC.

Factor de Bayes

El factor Bayes se utiliza para comparar modelos bayesianos. Supongamos un problema en el que se tiene que elegir entre dos posibles modelos M_1 y M_2 , una vez que se ha observado una muestra y .

Definición 2.5.8 El factor Bayes (FB) se define como:

$$FB = \frac{f(y|\theta_{M_1})}{f(y|\theta_{M_2})}$$

donde $f(y|\theta_M)$ es la verosimilitud marginal bajo el modelo M , $M \in (M_1, M_2)$ que está dada por

$$f(y|M) = \int f(y|\theta_M)f(\theta_M|M)d\theta_M \quad (2.13)$$

$f(y|\theta_M)$ es la verosimilitud bajo el modelo M que tiene parámetro θ_M y $f(\theta_M|M)$ es la a priori de θ_M dado el modelo M . Existen varios métodos para estimar (2.13), en este trabajo se realizó vía MCMC.

El FB es similar a la prueba de la razón de verosimilitud, pero en lugar de maximizar la verosimilitud, el FB realiza un promedio ponderado en la distribución de los parámetros (Kass & Raftery, 1995). En el caso del FB, Jeffreys estableció una escala de interpretación como se muestra en la Tabla 2.1:

BF	Fuerza de la evidencia a favor de M_1
1 – 3	Negativa (apoya M_2)
3 – 20	Positiva
20 – 150	Fuerte
> 150	Muy Fuerte

Tabla 2.1: Interpretación del FB según Jeffreys.

Estimación de la verosimilitud marginal

Un método para aproximar la verosimilitud marginal es utilizar los valores simulados, vía MCMC, de la distribución a posteriori mediante la siguiente ecuación:

$$\int \frac{1}{f(y|\boldsymbol{\theta}_m, m)} f(\boldsymbol{\theta}_m|y, m) d\boldsymbol{\theta}_m = \int \frac{1}{f(y|\boldsymbol{\theta}_m, m)} \frac{f(y|\boldsymbol{\theta}_m, m) f(\boldsymbol{\theta}_m|m)}{f(y|m)} d\boldsymbol{\theta}_m = \frac{1}{f(y|m)}$$

(Newton & Raftery, 1994) propusieron una estimación de la verosimilitud marginal. Este estimador, se obtiene en base a la media armónica de las verosimilitudes calculadas en cada paso de un algoritmo MCMC y está definido como:

$$\hat{f}_1(y|m) = \left(\frac{1}{T} \sum_{t=1}^T \{f(y|\boldsymbol{\theta}_m^{(t)}, m)\} \right)^{-1} \quad (2.14)$$

Gelfand y Dey (1994) generalizaron esta idea con:

$$\hat{f}_2(y|m) = \left(\frac{1}{T} \sum_{t=1}^T \frac{g(\boldsymbol{\theta}_m^{(t)})}{f(y|\boldsymbol{\theta}_m^{(t)}, m) f(\boldsymbol{\theta}_m^{(t)}|m)} \right)^{-1} \quad (2.15)$$

donde $g(\boldsymbol{\theta})$ es una distribución con colas más delgadas que el producto de la distribución a priori y la verosimilitud. Una elección sugerida para $g(\boldsymbol{\theta})$ es una distribución multivariada normal con media y varianza igual a la media y varianza a posteriori.

Metodología

En este trabajo se toman las ideas y los datos propuestos por Min, Su y Kuang (2009) para estimar los parámetros del modelo básico de infecciones virales con efecto de terapia y el modelo estandarizado de infecciones virales con efecto de terapia. Para las estimaciones de los parámetros en ambos modelos se utiliza el enfoque bayesiano, con el objetivo de realizar comparaciones de los resultados obtenidos por ambos modelos.

Se estudia un grupo de pacientes con hepatitis B crónica, que recibieron 100 mg de lamivudina una vez al día. El estudio se realizó durante 72 semanas (48 semanas de tratamiento y 24 libre de tratamiento). Para facilitar el entendimiento de la metodología usada en este trabajo se realizó el esquema que se muestra en la Figura 3.1 donde se describen los pasos y se detallan los métodos utilizados.

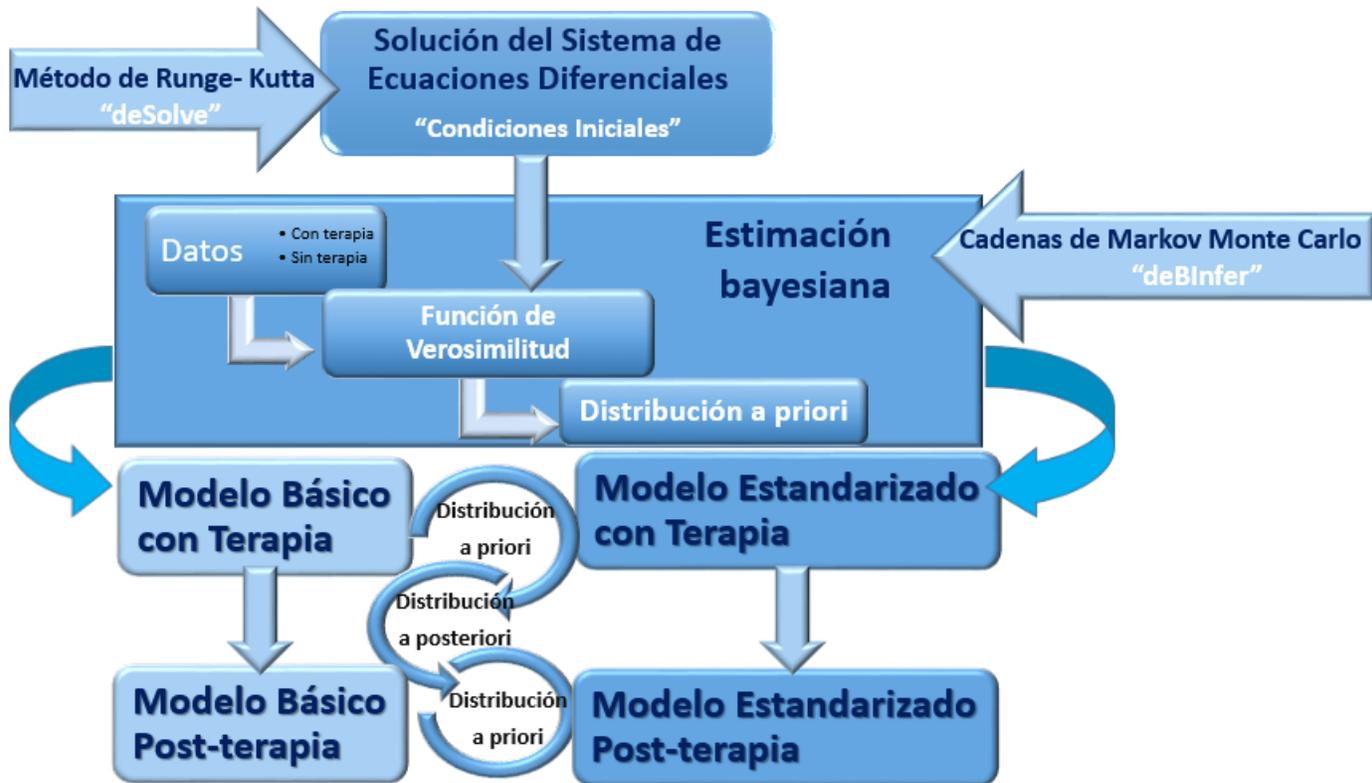


Figura 3.1: Esquema que describe la metodología de trabajo.

3.1. Modelo básico de infecciones virales con efecto de terapia

Para el tratamiento con el fármaco lamivudina, el BVIM de la terapia anti-VHB del paciente se describe por el siguiente modelo:

$$\begin{aligned}
 \frac{dx}{dt} &= \lambda - \delta x - (1 - m)\beta vx \\
 \frac{dy}{dt} &= (1 - m)\beta vx - \alpha y \\
 \frac{dv}{dt} &= (1 - n)\kappa y - \mu v
 \end{aligned}
 \tag{3.1}$$

donde m y n representan la eficacia de la terapia, dichos parámetros asumen el valor de cero cuando la terapia es ineficiente y la unidad cuando la terapia es efectiva. Estos parámetros proporcionan una medida de que tan efectivo es el fármaco que se aplica al

paciente, en el caso de m se tiene que este inhibe la tasa de contagio, mientras que n inhibe la tasa de replicación.

3.2. Modelo básico de infecciones virales post-terapia

Al retirar la terapia, m y n toman el valor de cero, entonces el BVIM post-terapia se describe por el siguiente modelo:

$$\begin{aligned}\frac{dx}{dt} &= \lambda - \delta x - \beta vx \\ \frac{dy}{dt} &= \beta vx - \alpha y \\ \frac{dv}{dt} &= \kappa y - \mu v\end{aligned}\tag{3.2}$$

3.3. Modelo estandarizado de la infección del VHB con efecto de terapia

El modelo estandarizado (3.2) tiene más sentido biológico debido a que el número reproductivo básico no considera el tamaño del órgano. Al aplicar el tratamiento con el fármaco lamivudina, el ABVIM de la terapia anti-VHB del paciente se describe por el siguiente modelo:

$$\begin{aligned}\frac{dx}{dt} &= \lambda - \delta x - (1 - m)\beta v \frac{x}{x + y} \\ \frac{dy}{dt} &= (1 - m)\beta v \frac{x}{x + y} - \alpha y \\ \frac{dv}{dt} &= (1 - n)\kappa y - \mu v\end{aligned}\tag{3.3}$$

donde se tiene que m y n representan la eficacia de la terapia, estos parámetros toman el valor de uno cuando el fármaco es efectivo y cero cuando el fármaco es ineficiente. Estos parámetros, al igual que en el Modelo (3.1) representan una medida de que tan efectivo es el fármaco que se aplica al paciente.

3.4. Modelo estandarizado de infecciones virales post-terapia

Al retirar la terapia m y n toman el valor de cero, entonces el ABVIM post-terapia se describe de la manera siguiente:

$$\begin{aligned}\frac{dx}{dt} &= \lambda - \delta x - \beta v \frac{x}{x+y} \\ \frac{dy}{dt} &= \beta v \frac{x}{x+y} - \alpha y \\ \frac{dv}{dt} &= \kappa y - \mu v\end{aligned}\tag{3.4}$$

3.5. Datos del tratamiento de lamivudina

Una vez que se le aplicó el tratamiento de lamivudina, se midió el logaritmo de la carga viral a los pacientes. Los datos obtenidos fueron los siguientes:

Durante la terapia													
Semanas	0	1	2	4	6	8	12	18	24	30	36	42	48
Carga Viral	9.8	7.8	6.6	5.6	5.1	4.8	4.4	4.3	4.2	4.0	4.15	4.2	4.5
Post-terapia													
Semanas	52	60	72										
Carga Viral	7.0	8.0	8.2										

Tabla 3.1: Respuesta a la terapia y el nivel del virus que vuelve después de que se detuvo la terapia.

En el estudio, los pacientes con hepatitis B crónica recibieron 100 mg de lamivudina una vez al día. El estudio comprendió 48 semanas de tratamiento y 24 semanas de seguimiento libre del tratamiento. Si bien el inicio de la terapia y los niveles virales disminuyen rápidamente, el virus regresa tan pronto como se retira el medicamento.

3.6. Condiciones iniciales

Con el objetivo de estimar los parámetros de los modelos BVIM y ABVIM mediante el enfoque bayesiano, primero se resuelven los sistemas de ecuaciones diferenciales mediante métodos numéricos, para ello se utilizan los datos de la Tabla 3.1.

a) Durante la terapia y las siguientes condiciones iniciales, según Min, Su y Kuang:

- Un hígado humano contiene 2×10^{11} hepatocitos. Un paciente tiene aproximadamente un total de 3000 ml plasma. Como consecuencia, se supone que el tamaño del órgano está dado por:

$$\frac{\lambda}{\delta} = \frac{2 \times 10^{11}}{3000}$$

Por lo tanto, la cantidad de células sanas inicial es:

$$x_0 = \left(\frac{2 \times 10^{11}}{3000} \right)$$

- La cantidad de células infectadas inicialmente es:

$$y_0 = 0$$

- La carga viral inicial según los datos de la Tabla 3.1 es: $v_0 = 9.8$

b) Una vez retirada la terapia las condiciones iniciales para los modelos (3.2) y (3.4), son las siguientes:

- La cantidad de células sanas inicial después de retirada la terapia es:

$$x_0 = 549.4172$$

- La cantidad de células infectadas inicial después de retirada la terapia es:

$$y_0 = 45125200$$

- La carga viral inicial después de retirada la terapia es:

$$v_0 = 4.5$$

3.7. Valores iniciales de los parámetros para los modelos de la dinámica de infección del VHB

Sobre la base de los datos clínicos de la información obtenida en la literatura sobre el tema, se seleccionan los valores iniciales de los parámetros de la siguiente manera:

- Dado que la vida media de un hepatocito es aproximadamente la mitad de un año, se supone que

$$\delta = -\frac{\ln(0.5)}{183} \approx 0.00379$$

entonces, la tasa de muerte de las células sanas es $\delta \approx 0.00379 \times 10^{-3}$.

- Se sabe que la vida media de un virus es aproximadamente un día. La tasa de declinación de las partículas virales es $\mu = 0.67$.
- El número de células no infectadas es $\lambda = 252667$, que sería la producción de células sanas.
- La tasa de infección es $\beta = 0.005$, la cual describe el proceso de propagación de la enfermedad cuando las células sanas entran en contacto con los virus libres.
- La tasa de muerte de las células infectadas es $\alpha = 3.38\delta$, que es proporcional a la muerte de las células sanas, por lo que se toma un $\alpha = 0.0128$.
- La producción de partículas virales $\kappa = 6.24$, lo que significa que por cada célula infectada se van a generar 6.24 partículas virales.
- La eficacia de la terapia que bloquea la producción de células infectadas es $m = 0$.
- La eficacia de la terapia que bloquea la producción de partículas virales es $n = 0.99982$.
- Supongamos que, antes del tratamiento, el paciente se encuentra en un estado persistente de la infección, es decir:

$$E^* = (\bar{x}, \bar{y}, \bar{v}) = \left(\frac{\lambda}{\delta + \alpha(R_0^* - 1)}, \frac{\lambda(R_0^* - 1)}{\delta + \alpha(R_0^* - 1)}, \frac{\lambda\kappa(R_0^* - 1)}{\mu[\delta + \alpha(R_0^* - 1)]} \right)$$

donde \bar{x} , \bar{y} y \bar{v} son las medias de las células sanas, células infectadas y virus, respectivamente.

Entonces:

$$\frac{\bar{y}}{\bar{x} + \bar{y}} = \frac{R_0^* - 1}{R_0^*}$$

en una infección crónica por el VHB entre 5 y 40 por ciento de todos los hepatocitos pueden infectarse. En consecuencia, podemos elegir $R_0^* = 1.33$. Antes de la terapia con medicamentos, se supone que $m = n = 0$, y el paciente está en el estado estable E^* , por lo tanto:

$$\kappa = \frac{\mu v(0)[\delta + (R_0^* - 1)]}{\lambda(R_0^* - 1)}$$

$$\beta = \frac{\alpha \mu R_0^*}{\kappa}$$

donde $\kappa = 7.948337$, luego podemos determinar que $\beta = 0,002$.

De acuerdo a los modelos de infecciones virales con efecto de tratamiento (3.1) y (3.2) se plantea el siguiente vector de parámetros:

$$\theta = (\lambda, \delta, \beta, \alpha, \kappa, \mu, m, n)$$

En el estudio de la enfermedad de hepatitis B los parámetros que se desean estimar en los modelos (3.1) y (3.3) son m y n porque proporcionan información de que tan eficaz es el efecto de la lamivudina para el paciente, β y κ porque determinan qué tan alta es la tasa de transmisión y la producción viral de un determinado paciente, sin tener que someterlo a la biopsia del hígado. Por consiguiente, en esta investigación se consideran $(\lambda, \delta, \alpha, \mu)$ parámetros fijos y se desean estimar β, κ, m y n ; mientras que para los modelos (3.2) y (3.4) se estimarán β y κ debido a que en estos modelos no existe efecto de terapia.

3.8. Solución numérica

Dado que los modelos (3.1), (3.2), (3.3) y (3.4) no tiene solución exacta se obtiene una aproximación numérica mediante el método de Runge-Kutta. Para obtener la solución

aproximada de los modelos BVIM y ABVIM con y sin tratamiento, se utiliza la librería deSolve del software R, usando las condiciones iniciales y valores iniciales de los parámetros establecidos en los apartados anteriores (Soetaert, Petzoldt, & Setzer, 2010).

3.9. Estimación de parámetros

Los datos del logaritmo de las cargas virales durante las 48 semanas que se aplicó el tratamiento se muestran en la Tabla 3.1. En este trabajo se supone el siguiente modelo estadístico:

$$y_i = f(X_{\theta}(t_i)) + \varepsilon_i, \quad \text{con } i = 1, 2, \dots, 13. \quad (3.5)$$

donde $\theta = (\beta, \kappa, m, n)$, X_{θ} es la solución numérica de los modelos BVIM y ABVIM, i es el tiempo medido en semanas y los y_i son el logaritmo de la carga viral.

La transformación logarítmica de la carga viral se usa para estabilizar la varianza del error de medición y el algoritmo de estimación.

Se trabaja bajo el supuesto que los datos se distribuyen lognormal, con media X_{θ}^h y varianza constante desconocida σ^2

$$P_{Y|\Phi}(y|\phi) = (y_i\sigma)^{-13}(2\pi)^{-13/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{13} (\log(y_i) - X_{\theta}^h(t_i))^2 \right\} \quad (3.6)$$

3.9.1. Distribuciones a priori

Según Vegvari, Cauët, Lawrence y Anderson (2016) la tasa de infección β tiene distribución gamma y la eficacia del tratamiento n tiene distribución beta. En el presente trabajo se asume que la distribución conjunta del vector de parámetro ϕ es el producto de distribuciones marginales de cada parámetro, es decir, que se asumen que son independientes. Las distribuciones a priori utilizadas son las siguientes:

- $\beta \sim \Gamma(a_1, b_1)$.

$$f(\beta, a_1, b_1) = \frac{1}{b_1^{a_1} \Gamma(a_1)} \beta^{a_1-1} e^{-\frac{\beta}{b_1}}$$

- $m \sim B(a_2, b_2)$.

$$f(m, a_2, b_2) = \frac{\Gamma(a_2 + b_2)}{\Gamma(a_2) \Gamma(b_2)} m^{a_2-1} (1-m)^{b_2-1}$$

- $n \sim B(a_3, b_3)$.

$$f(n, a_3, b_3) = \frac{\Gamma(a_3 + b_3)}{\Gamma(a_3) \Gamma(b_3)} n^{a_3-1} (1-n)^{b_3-1}$$

- $\kappa \sim \Gamma(a_4, b_4)$.

$$f(\kappa, a_4, b_4) = \frac{1}{b_4^{a_4} \Gamma(a_4)} \kappa^{a_4-1} e^{-\frac{\kappa}{b_4}}$$

- $\sigma \sim \Gamma(a_5, b_5)$.

$$f(\sigma, a_5, b_5) = \frac{1}{b_5^{a_5} \Gamma(a_5)} \sigma^{a_5-1} e^{-\frac{\sigma}{b_5}}$$

3.9.2. Distribuciones a posteriori

La distribución a posteriori del modelo aproximado, sería de la forma:

$$P_{\Phi}^h(\phi|y) = \frac{P_{Y|\Phi}^h(y|\phi) P_{\Phi}(\phi)}{P_Y^h(y)}$$

donde

$$P_Y^h(y) = \int P_{Y|\Phi}^h(y|\phi) P_{\Phi}(\phi) d\phi$$

$\phi = (\beta, \kappa, m, n, \sigma)$ es vector de parámetros y

$$P_{\Phi}(\phi) \propto \beta^{a_1-1} e^{-\frac{\beta}{b_1}} \times \kappa^{a_4-1} e^{-\frac{\kappa}{b_4}} \times m^{a_2-1} (1-m)^{b_2-1} \times n^{a_3-1} (1-n)^{b_3-1} \times \sigma^{a_5-1} e^{-\frac{\sigma}{b_5}}$$

La distribución a posteriori es

$$P_{\Phi}(\phi|y) \propto (y_i \sigma)^{-13} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{13} (\log(y_i) - X_{\theta}^h(t_i))^2 \right\} \times \beta^{a_1-1} e^{-\frac{\beta}{b_1}} \times \kappa^{a_4-1} e^{-\frac{\kappa}{b_4}} \times m^{a_2-1} (1-m)^{b_2-1} \times n^{a_3-1} (1-n)^{b_3-1} \times \sigma^{a_5-1} e^{-\frac{\sigma}{b_5}}$$

La distribución a posteriori del vector de parámetros θ se aproxima utilizando cadenas de Markov. Se generaron 3 cadenas para cada modelo, con 40000 iteraciones de las cuales

se descartaron 20000. Se utiliza la librería *deBInfer* para determinar las estimaciones y los intervalos creíbles. Para verificar la convergencia se utiliza el diagnóstico de Gelman y Rubin que está implementado en el software **R** utilizando la librería *coda* (Plummer, Best, Cowles & Vines, 2006).

3.9.3. Estimaciones de cantidades relacionadas a la dinámica viral.

Para realizar las estimaciones del número reproductivo básico de la infección se tomaron las a posteriori de cada uno de los parámetros estimados y se dejaron los demás fijos, se obtiene así una estimación bayesiana para el R_0 y para R_0^* y sus respectivos intervalos de credibilidad.

Para estimar el número de partículas virales producidas por una célula infectada de determinó el cociente entre la producción de partículas virales y la tasa de muerte de las células infectadas. Se tomó la a posteriori de κ y el valor fijo de α para lograr así una estimación bayesiana del tamaño de la explosión.

Resultados y discusión

En este capítulo se muestran los resultados de las estimaciones bayesianas obtenidas en el modelo estandarizado y el modelo básico, con terapia y post-terapia de lamivudina. También se realiza una comparación estadística entre el modelo estandarizado y básico.

4.1. Estimación de parámetros del modelo básico de infecciones virales con terapia

Se desea realizar las estimaciones de los parámetros de interés en el modelo básico con efecto de terapia, que es el Modelo (3.1). Para realizar las estimaciones bayesianas se utiliza la librería *deBInfer*.

4.1.1. Distribución a priori de los parámetros del modelo básico y estandarizado, con terapia.

Las distribuciones a priori con los hiperparámetros que se muestran en la Tabla 4.1. Para las distribuciones de los parámetros se tuvo en cuenta las distribuciones que se propusieron en la literatura consultada y de acuerdo a la naturaleza de los parámetros. Los valores de los hiperparámetros de cada una de las distribuciones se les asignaron según la media y la varianza de cada una de las distribuciones trabajadas, de forma tal que la distribución a priori le diera un sentido biológico al parámetro que se desea estimar.

Parámetros	Distribución	Hiperparámetros
β	Gamma	$f(\beta, 1, 15)$
κ	Gamma	$f(\kappa, 17, 2)$
m	Beta	$f(m, 1, 1000)$
n	Beta	$f(n, 100, 1)$
σ	Gamma	$f(\sigma, 1, 20)$

Tabla 4.1: Distribuciones a priori de los parámetros del modelo básico y estandarizado, ambos con tratamiento.

Una mejor forma de ver las distribuciones a priori propuestas consiste en observar el comportamiento de su función de distribución, como se muestra en las Figuras 4.1 y 4.2

4.1.2. Estimaciones bayesianas de los parámetros del modelo básico con terapia

Luego de obtenida la solución numérica del modelo (3.1) y utilizando la función *ode* de la librería *deSolve*, se determina la función de verosimilitud que se muestra en la ecuación (3.6). Se emplean las distribuciones a priori de la Tabla 4.1 para determinar las estimaciones bayesianas de los parámetros del modelo básico con terapia y sus intervalos creíbles.

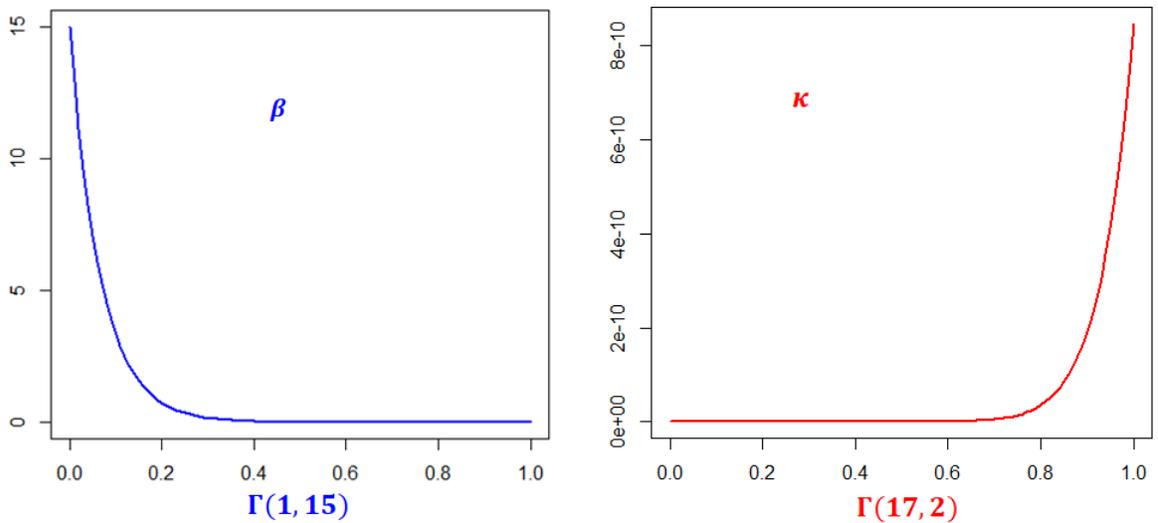


Figura 4.1: Distribución a priori de los parámetros β y κ para los modelos (básico y estandarizado) con tratamiento.

Se utilizan 3 cadenas cada una de 40000 iteraciones. Las estimaciones se presentan en la Tabla 4.2.

La tasa de infección β tuvo un valor estimado de 0.0665, la cual describe el proceso de propagación de la enfermedad cuando las células sanas entran en contacto con los virus libres. La producción de partículas virales κ resultó con un valor estimado de 7.9742, lo que significa que por cada célula infectada se van a generar aproximadamente unas 8 partículas virales. La eficacia de la terapia que bloquea la producción de células infectadas m se estimó con un valor de 0.0005, lo que indica que este efecto de terapia está siendo eficaz solo un 0.05%. Por otro lado, la eficacia de la terapia que bloquea la producción de partículas virales n dio un valor estimado de 0.9997, lo que muestra que esta terapia está siendo efectiva 99.97%, bloqueando así la producción viral.

De los resultados del modelo básico con efecto de tratamiento se concluye que las estimaciones bayesianas obtenidas fueron buenas porque tienen semejanza a los valores que se consultaron anteriormente en la literatura, por lo tanto, podrían ser de utilidad para futuros trabajos en la medicina.

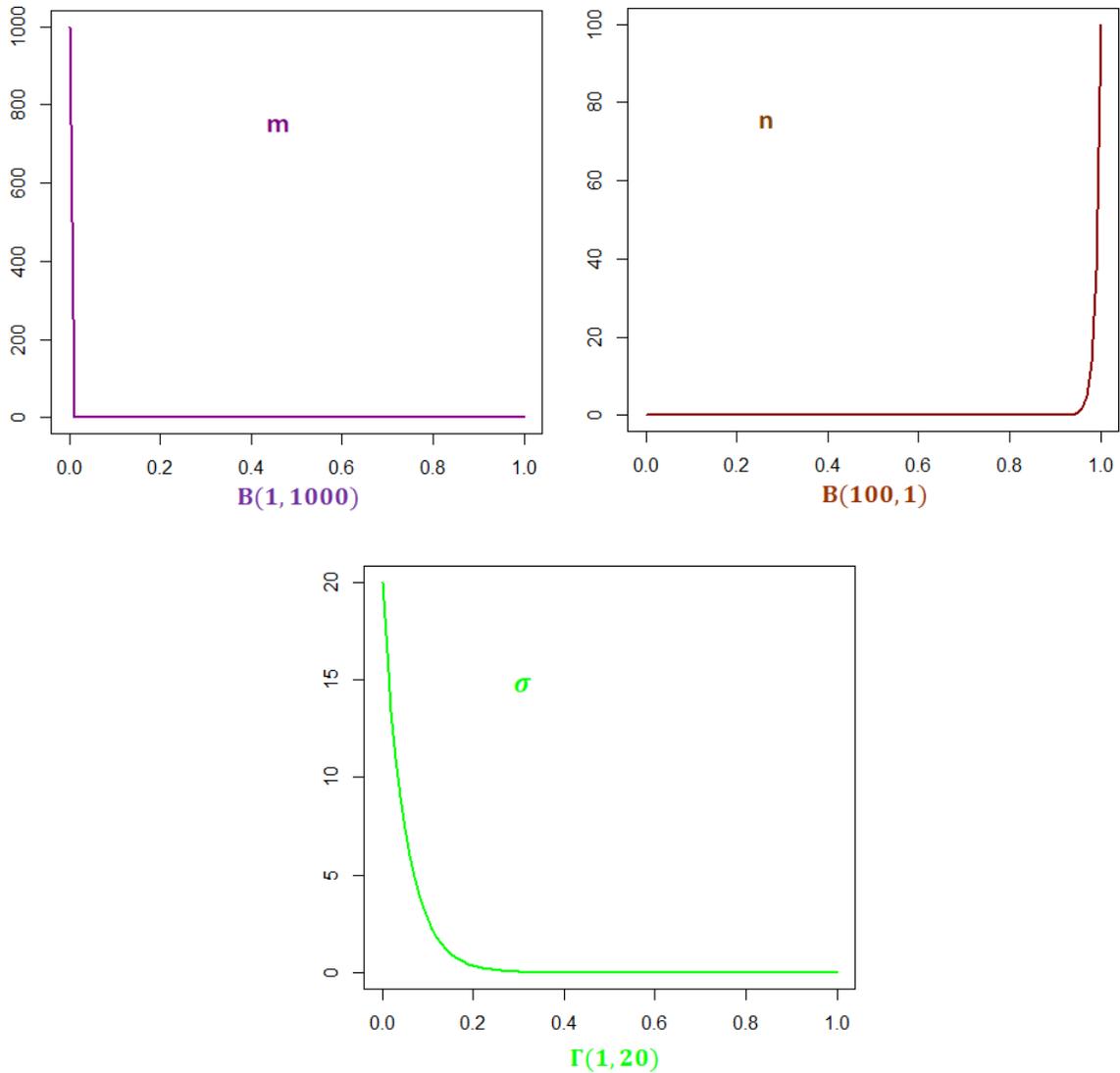


Figura 4.2: Distribución a priori de los parámetros m , n y σ para los modelos (básico y estandarizado) con tratamiento.

4.1.3. Diagnósticos de convergencia

Para validar los resultados de las estimaciones anteriores en las Figuras 1 y 2 de los anexos se muestran las trazas de las tres cadenas resultantes. Se observa gráficamente que las mismas tienen un buen mezclado, lo indica que existe convergencia en las cadenas. En la Figura 3 de los anexos se observan las distribuciones a posteriori de cada una de las estimaciones y las tres cadenas resultantes se superponen muy bien una encima

Parámetros	media	sd	Intervalos creíbles		
			2.5 %	50 %	97.5 %
β	0.0665	0.0666	1.644e-03	0.0457	0.2456
κ	7.9742	1.9807	4.601e+00	7.8000	12.3028
m	0.0005	0.0007	4.724e-07	0.00023	0.0025
n	0.9997	0.0022	9.995e-01	0.9997	0.9998
$sd.v$	0.0503	0.0507	1.268e-03	0.0346	0.1893

Tabla 4.2: Estimaciones bayesianas del modelo Básico con efecto de tratamiento.

de la otra.

Se aplicó el diagnóstico de convergencia de Gelman y Rubin, para cada uno de los parámetros del Modelo 3.1. Se obtuvo la estimación puntual y el límite superior del intervalo creíble (I.C) como se muestra en la Tabla 4.3 y en la Figura 5 de los anexos. En esta figura se puede apreciar que a partir de la segunda mitad el comportamiento de cada una de las gráficas es cercano a 1, de igual manera los valores que se observan en la tabla todos se encuentran por debajo de 1.2, lo que garantiza la convergencia en las cadenas de los parámetros para modelo básico con efecto de terapia. El diagnóstico de

Parámetros	Estimación Puntual	I.C Superior
β	1.00	1.00
κ	1.01	1.03
m	1.00	1.00
n	1.02	1.05
σ	1.00	1.00

Tabla 4.3: Diagnóstico de Gelman-Rubin de las estimaciones para los parámetros del modelo básico con terapia.

convergencia de Geweke dio como resultado $|Z| < 2$ para las cadenas de cada uno de los parámetros, como se muestra en la Tabla 4.4 y en la Figura 5, lo que indica medias iguales para las primeras y últimas iteraciones, y por lo tanto existe convergencia de la cadena MCMC para los parámetros del modelo básico con terapia.

Parámetros	β	κ	m	n	σ
$ Z $	-0.3533	0.6344	-1.4211	-0.4049	1.3878

Tabla 4.4: Diagnóstico de Geweke de las estimaciones para los parámetros del modelo básico con terapia.

4.2. Estimación de parámetros del modelo básico de infecciones virales post-terapia

Una vez obtenida las estimaciones para el modelo básico con efecto de tratamiento, se realiza el mismo procedimiento para obtener las estimaciones de los parámetros de interés para el modelo post-terapia, que es el Modelo (3.2). De igual manera se utiliza la librería *deBInfer*.

4.2.1. Distribución a priori de los parámetros de los modelo básico y estandarizado, post-terapia.

Se asignan las distribuciones a priori con los hiperparámetros que se observan en la Tabla 4.5. Los valores de los hiperparámetros de cada una de las distribuciones se asignaron realizando una prueba de bondad de ajuste según el conocimiento previo que se obtuvo de las estimaciones del modelo básico con efecto de terapia, de forma tal que la distribución a priori le diera un sentido biológico al parámetro que se desea estimar, ya que la muestra con la que se cuenta es más pequeña que para el modelo con terepia, por tanto el mayor peso recae en las distribuciones a priori.

Una mejor forma de ver las distribuciones a priori propuesta es mediante la gráficas de su distribución, como se muestra en la Figura 4.3.

Parámetros	Distribución	Hiperparámetros
β	Gamma	$f(\beta, 0.59, 14.66)$
κ	Gamma	$f(\kappa, 16.01, 1.97)$
σ	Gamma	$f(\kappa, 1.02, 20.66)$

Tabla 4.5: Distribuciones de los parámetros del modelo básico y estandarizado, ambos en la post-terapia.

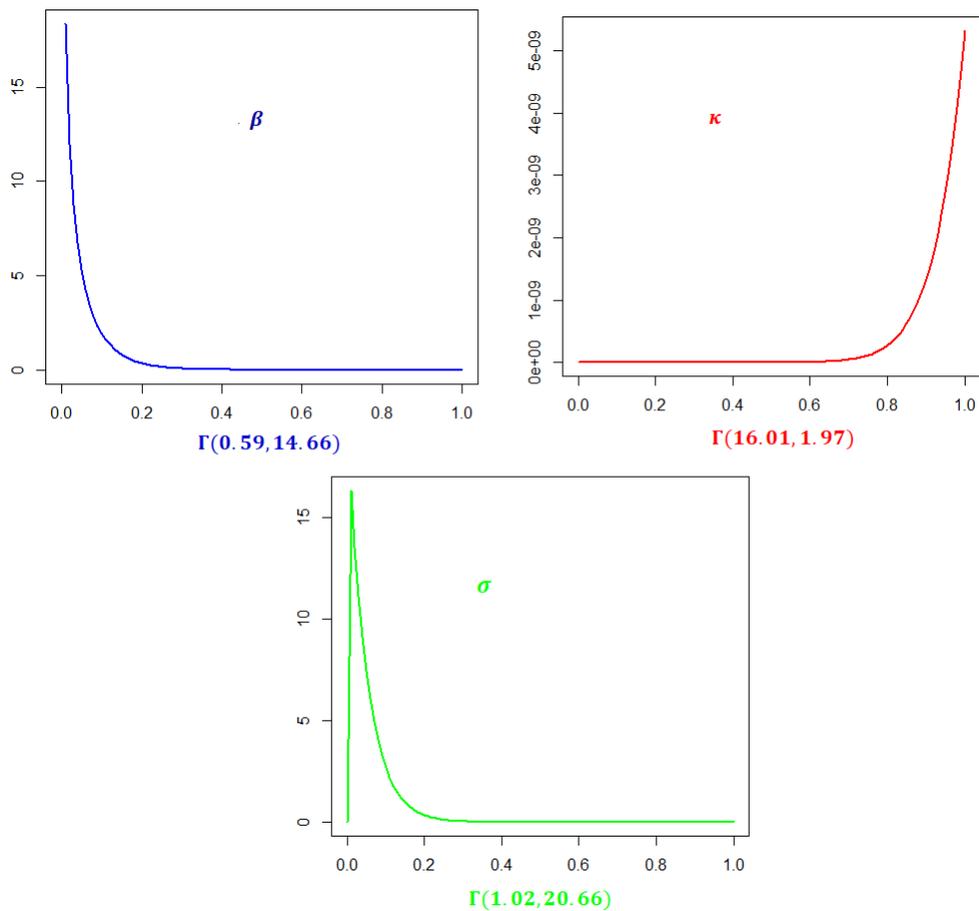


Figura 4.3: Distribución a priori de los parámetros β , κ y σ para los modelos (básico y estandarizado) post-terapia.

4.2.2. Estimaciones bayesianas de los parámetros del modelo básico post-terapia

Una vez determinada la solución numérica del Modelo (3.2) utilizando la función *ode* de la librería *deSolve*, se determina la función de verosimilitud que se muestra en la ecuación (3.6). Se utilizan las distribuciones a priori que se observan en la Tabla 4.5 para determinar las estimaciones bayesianas de los parámetros del modelo básico post-terapia y sus respectivos intervalos creíbles. Para realizar las estimaciones se emplean 3 cadenas de 10000 iteraciones. Los resultados de las estimaciones se presentan en la Tabla 4.6.

En los resultados del modelo básico post-terapia se obtuvieron buenas estimaciones bayesianas teniendo en cuenta los valores que se consultaron y sus correspondientes intervalos creíbles para cada uno de los parámetros.

La producción de partículas virales κ resultó con una estimación de 7.9102, es decir que por cada célula infectada se van a producir aproximadamente 8 virus. Para la tasa de infección β se obtuvo un valor estimado de 0.0322, la cual indica el proceso de propagación de la enfermedad en el momento que las células sanas entran en contacto con los viriones.

Parámetros	media	sd	Intervalos creíbles		
			2.5 %	50 %	97.5 %
β	0.0322	0.0497	0.0001	0.0123	0.1804
κ	7.9102	1.8967	4.5797	7.7649	12.0670
σ	0.0467	0.0501	0.0013	0.0317	0.1845

Tabla 4.6: Estimaciones bayesianas del modelo básico post-terapia.

4.2.3. Diagnósticos de convergencia

La Figura 7 de los anexos muestra las distribuciones a posteriori de cada una de las estimaciones y las tres cadenas resultantes se superponen bien una encima de la otra. En la Figura 6 de los anexos se presentan las trazas de las tres cadenas resultantes, gráficamente se muestra que las mismas tienen un buen mezclado, lo que llevaría a pensar que existe convergencia en las cadenas.

Para garantizar la convergencia de las cadenas, se aplicó el diagnóstico de convergencia de Gelman y Rubin, tal como se muestra en la Tabla 4.7 y en la Figura 8 de los anexos. En la figura se muestra que a partir de la segunda mitad el comportamiento de cada una de las gráficas es bien cercano a 1, de igual forma en la tabla todos los valores se encuentran por debajo de 1.2, lo que garantiza la convergencia en las cadenas de los parámetros para modelo básico post-terapia.

Parámetros	Estimación Puntual.	I.C Superior
β	1.00	1.00
κ	1.01	1.03
σ	1.00	1.00

Tabla 4.7: Diagnóstico de Gelman-Rubin de las estimaciones para los parámetros del modelo básico post-terapia.

También se aplicó el diagnóstico de convergencia de Geweke que arrojó como resultado $|Z| < 2$ para cada uno de los parámetros, como se muestra en la Tabla 4.8 y en la Figura 9 de los anexos, indicando así medias iguales para las primeras y últimas iteraciones, es decir, se concluye que existe convergencia de la cadena MCMC para los parámetros del modelo básico post-terapia.

Parámetros	β	κ	σ
$ Z $	0.6565	0.6096	0.7842

Tabla 4.8: Diagnóstico de Geweke de las estimaciones para los parámetros del modelo básico post-terapia.

4.3. Estimación de parámetros del modelo estandarizado con terapia

Se obtiene la solución numérica del modelo de ecuaciones diferenciales, con las condiciones y los valores iniciales. Luego de obtenida la solución numérica del modelo (3.3) utilizando la función *ode* de la librería *deSolve*, se determina la función de verosimilitud que se muestra en la ecuación (3.6).

4.3.1. Estimaciones bayesianas de los parámetros del modelo estandarizado con terapia

Con las distribuciones a priori que se observan en la Tabla 4.1 se determinan las estimaciones para los parámetros del modelo estandarizado con terapia y sus respectivos intervalos creíbles. Las estimaciones se presentan en la Tabla 4.9.

En el modelo estandarizado con efecto terapia se puede concluir que se obtuvieron satisfactorias estimaciones bayesianas según los valores que se consultaron previamente en la literatura, además se determinaron sus correspondientes intervalos creíbles para cada uno de los parámetros, brindando así un conjunto de valores con sentido biológico que será gran utilidad para futuras investigaciones en la rama de la biología.

Para la tasa de infección β se obtuvo un valor estimado de 0.0405, la cual describe el proceso de propagación de la enfermedad cuando los hepatocitos sanos entran en contacto

con los viriones. La eficacia de la terapia que bloquea la producción de células infectadas m se estimó con un valor de 0.0010, lo que indica que este efecto de terapia está siendo eficaz solo un 0.101 %. Sin embargo, la eficacia de la terapia que bloquea la producción de partículas virales n dio un valor estimado de 0.9960, lo que muestra que esta terapia está siendo efectiva 99.60 %, bloqueando así los virus. La producción de partículas virales κ dio con un valor estimado es de 8.0082, lo que nos dice que por cada célula infectada se generan unos 8 viriones.

Parámetros	media	sd	Intervalos creíbles		
			2.5 %	50 %	97.5 %
β	0.0405	0.291	7.353e-03	0.0150	0.0735
κ	8.0082	1.940	5.039e+00	7.7550	12.0030
m	0.0010	0.015	7.765e-07	0.0002	0.0025
n	0.9960	0.019	9.930e-01	0.9968	0.9984
σ	0.0498	0.056	1.846e-03	0.0339	0.1869

Tabla 4.9: Estimaciones bayesianas para los parámetros del modelo estandarizado con tratamiento.

4.3.2. Diagnósticos de convergencia

Con el interés de validar los resultados anteriores en las Figuras 10 y 11 de los anexos se presentan las trazas de las tres cadenas resultantes. Se observa gráficamente que las mismas tienen un buen mezclado, lo que llevaría a pensar que existe convergencia en las cadenas. En la Figura 12 de los anexos se observan las distribuciones a posteriori de cada una de las estimaciones y las tres cadenas resultantes se superponen una encima de la otra.

En la Figura 13 se observan las gráficas del diagnóstico de convergencia de Gelman y Rubin que a partir de la segunda mitad el comportamiento de cada una de ellas es bien cercano a 1, de igual manera en la Tabla 4.10 todos los valores se comportan por debajo de 1.2, lo que garantiza la convergencia de los parámetros del modelo estandarizado con

efecto de tratamiento.

Luego aplicando el diagnóstico de convergencia de Geweke se obtiene $|Z| < 2$ para las

Parámetros	Estimación Puntual.	I.C Superior
β	1.02	1.07
κ	1.01	1.03
m	1.00	1.00
n	1.10	1.26
σ	1.00	1.00

Tabla 4.10: Diagnóstico de convergencia de las estimaciones para los parámetros del modelo estandarizado con tratamiento.

cadena de cada uno de los parámetros, como se muestra en la Tabla 4.11 y la Figura 14, es decir, existe convergencia de la cadena MCMC para los parámetros del modelo estandarizado con terapia, ya que las medias para las primeras y últimas iteraciones son iguales.

Parámetros	β	κ	m	n	σ
$ Z $	-0.6245	-0.5387	-0.2357	0.6142	1.4357

Tabla 4.11: Diagnóstico de Geweke de las estimaciones para los parámetros del modelo estandarizado con terapia.

4.4. Estimación de parámetros del modelo estandarizado post-terapia

Una vez retirada la terapia se determina la solución numérica del Modelo (3.4) utilizando la función *ode* de la librería *deSolve* y la función de verosimilitud que se muestra en la Ecuación (3.6). Se utilizan las distribuciones a priori que se observan en la Tabla 4.5 pa-

ra determinar las estimaciones bayesianas de los parámetros del modelo estandarizado post-terapia y sus respectivos intervalos de credibilidad.

4.4.1. Estimaciones bayesianas de los parámetros del modelo estandarizado post-terapia

Para realizar las estimaciones se utilizan 3 cadenas de 10000 iteraciones. Las estimaciones se presentan en la Tabla 4.12. En los resultados del modelo estandarizado post-terapia se obtuvieron buenas estimaciones bayesianas teniendo en cuenta los valores que se consultaron en la literatura y sus correspondientes intervalos creíbles para cada uno de los parámetros resultante.

La tasa de infección β obtuvo un valor estimado de 0.0471, la cual describe el proceso de propagación de la enfermedad cuando las células sanas entran en contacto con los virus libres. La producción de partículas virales κ resultó con un valor estimado de 8.0758, lo que significa que por cada célula infectada se van a generar aproximadamente unas 8 partículas virales.

Parámetros	media	sd	Intervalos creíbles		
			2.5 %	50 %	97.5 %
β	0.0471	0.0641	4.534e-05	0.0232	0.2394
κ	8.0758	1.9022	4.711e+00	7.9333	12.2693
σ	0.0466	0.0534	4.751e-04	0.0296	0.1923

Tabla 4.12: Estimaciones bayesianas para los parámetros del modelo estandarizado post-terapia.

4.4.2. Diagnósticos de convergencia

Se muestran en la Figura 15 de los anexos las trazas de las tres cadenas resultantes. Gráficamente se muestra que las mismas tienen un buen mezclado, lo que llevaría a pensar que existe convergencia entre las cadenas. En la Figura 16 de los anexos se observan las distribuciones a posteriori de cada una de las estimaciones y las tres cadenas resultantes se superponen muy bien una encima de la otra.

Para garantizar la convergencia de las cadenas, se aplicó el diagnóstico de convergencia de Gelman y Rubin, tal como se muestra en la Tabla 4.13 y en la Figura 17 de los anexos. Se observa en la figura que a partir de la segunda mitad el comportamiento de cada una de las gráficas es bien cercano a 1, de igual manera en la tabla todos los valores se encuentran por debajo de 1.2, lo que garantiza la convergencia de los parámetros del modelo estandarizado post-terapia. El diagnóstico de convergencia de Geweke dio como

Parámetros	Estimación Puntual.	I.C Superior
β	1	1.00
κ	1	1.01
σ	1	1.00

Tabla 4.13: Diagnóstico de Gelman-Rubin de las estimaciones para los parámetros del modelo básico post-terapia.

resultado $|Z| < 2$ para las cadenas de cada uno de los parámetros, como se muestra en la Tabla 4.14, lo que indica medias iguales para las primeras y últimas iteraciones, y por lo tanto existe convergencia de la cadena MCMC para los parámetros del modelo estandarizado post-terapia.

Parámetros	β	κ	σ
$ Z $	0.3381	-0.3976	-1.3845

Tabla 4.14: Diagnóstico de Geweke de las estimaciones para los parámetros del modelo estandarizado post-terapia.

4.5. Estimaciones del número reproductivo básico de la infección

De los resultados presentados en la Tabla 4.15 se obtiene un R_0^* para el modelo estandarizado de $R_0^* = 0.02020532$ que significa que no existe presencia de la infección, es decir, que una vez aplicada la terapia ésta bloquea la enfermedad de la VHB. Para el modelo básico se obtiene un $R_0 = 1034193$ que significa que si existe presencia de dicha enfermedad y es una contradicción teniendo en cuenta que se aplicó el tratamiento, por tanto existe una sobre estimación del R_0 dependiendo del tamaño del hígado.

		R_0^* del Modelo Estandarizado	$R_0^* = (1 - m)(1 - n) \frac{\beta\kappa}{\alpha\mu}$		
		media	2.5 %	50 %	97.5 %
R_0^*	0.0202	0.0100	0.01634	0.0537	
		R_0 del Modelo Básico	$R_0 = (1 - m)(1 - n) \frac{\lambda\beta\kappa}{\alpha\delta\mu}$		
		media	2.5 %	50 %	97.5 %
R_0	1034193	25036.45	701057.37	3857697.12	

Tabla 4.15: Estimaciones del número reproductivo básico de la infección, para los modelos básico y estandarizado, con terapia.

Al retirar la terapia de lamivudina la carga viral comienza a aumentar rápidamente, por lo tanto el número reproductivo básico de la infección debe ser mayor que 1. De la Tabla 4.16 se obtiene un $R_0^* = 23.43823$ para el modelo estandarizado post-terapia que significa la presencia de la infección y es consistente con lo que está ocurriendo en la Figura 4.7. Para el modelo básico post-terapia se obtiene un $R_0 = 733433839$ que significa la presencia de la enfermedad, pero se puede notar que es un valor extremadamente grande, por lo tanto de igual manera, existe una sobre estimación del R_0 dependiendo del tamaño del hígado. Sin embargo, se puede notar en la Figura 4.5 que en ambos casos su distribución mantiene igual comportamiento.

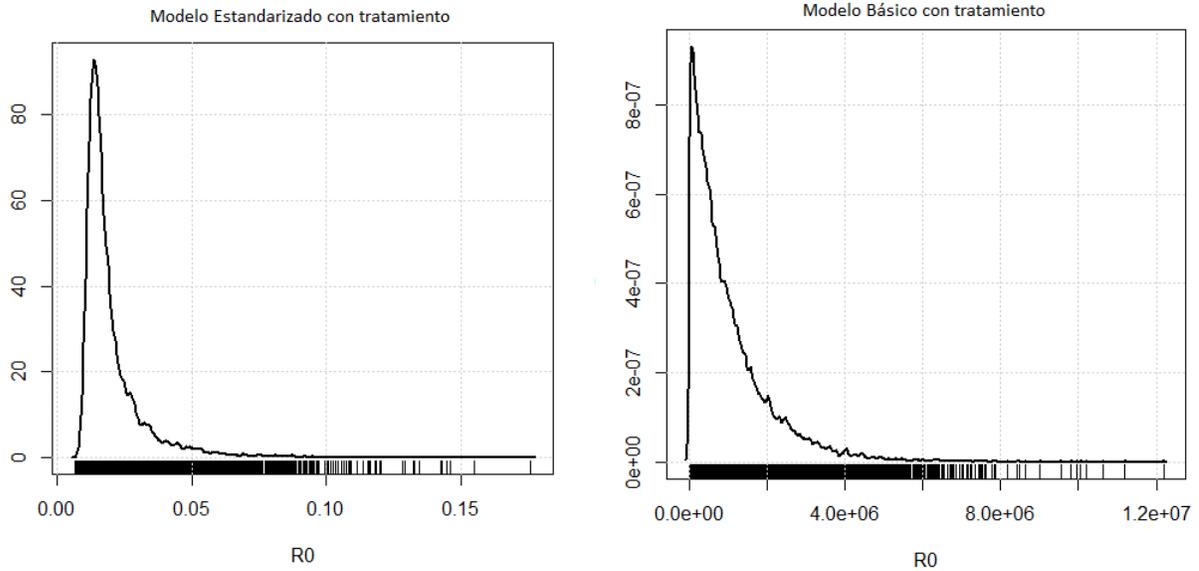


Figura 4.4: Distribución del número reproductivo básico de la infección para los modelos estandarizado y básico, con terapia.

		R_0^* del Modelo Estandarizado		$R_0^* = \frac{\beta\kappa}{\alpha\mu}$	
	media	2.5 %	50 %	97.5 %	
R_0^*	23.43823	1.658906	17.074157	76.849241	
		R_0 del Modelo Básico		$R_0 = \frac{\lambda\beta\kappa}{\alpha\delta\mu}$	
	media	2.5 %	50 %	97.5 %	
R_0	733433839	4012751	282302447	3912493363	

Tabla 4.16: Estimaciones del número reproductivo básico de la infección, para los modelos básico y estandarizado, post-terapia.

4.6. Selección de modelo

¿Cómo seleccionar el mejor modelo de los dos? Para esto se utiliza los criterios estadísticos bayesianos de comparación de modelos que se muestran en la Tabla 4.17 para los modelos con terapia y en la Tabla 4.18 para los modelos post-terapia.

Después de obtener los resultados de los métodos estadísticos para los criterios de selec-

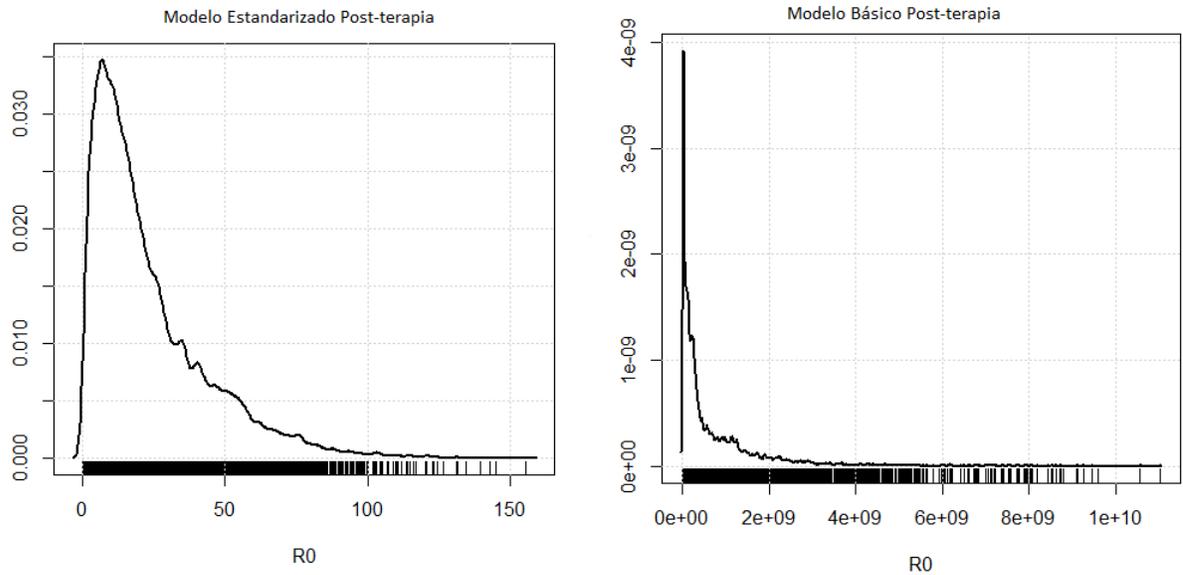


Figura 4.5: Distribución del número reproductivo básico de la infección para los modelos estandarizado y básico, post-terapia.

ción de modelos, se puede concluir con certeza que el mejor modelo es el estandarizado, con terapia y post-terapia. Se observa que el DIC , es menor para el modelo estandarizado que para el modelo básico, como se muestra en las Tablas 4.17 y 4.18. El resultado del factor Bayes, confirma que el mejor modelo es el estandarizado, puesto que indica una fuerte evidencia a favor de dicho modelo.

Criterios de Comparación	Mod. Estandarizado	Mod. Básico
DIC	6.505935	6.66652
Mod. Básico vs Mod.Estandarizado		
FB	12.7316	

Tabla 4.17: Criterios de comparación de los modelos estandarizado y básico, con terapia.

Criterios de Comparación	Mod. Estandarizado	Mod.Básico
<i>DIC</i>	2.505935	2.520765
Mod. Básico vs Mod.Estandarizado		
<i>FB</i>	16.39591	

Tabla 4.18: Criterios de comparación de los modelos estandarizado y básico, post-terapia.

4.7. Discusión

Una vez obtenida las estimaciones de los parámetros para cada uno de los modelos y decidir cuál fue el mejor atendiendo los criterios estadísticos, se compraron los resultados obtenidos con la literatura consultada.

4.7.1. Comparación de resultados

Se muestra a continuación en la Tabla 4.19 las estimaciones de los parámetros en los modelos básico y estandarizado con efecto de tratamiento, se puede observar que las estimaciones para cada uno de los parámetros son parecidas. De igual manera que ocurre para los modelos post-terapia que se muestra en la Tabla 4.20.

El número de partículas virales (tamaño de la explosión) producidas por una célula infectada en el modelo estandarizado con terapia es $\frac{\kappa}{\alpha} = 625.645$.

Para determinar el tamaño de la explosión en el modelo estandarizado post-terapia, el número de partículas virales producidas por una célula infectada aumenta $\frac{\kappa}{\alpha} = 630.927$.

En la Figura 4.6 se puede observar un buen ajuste entre los datos y la solución numérica obtenida utilizando las estimaciones bayesianas de los modelos estandarizado y básico con terapia, ambos modelos arrojan resultados similares, ya que una vez aplicado el tratamiento la carga viral comienza a disminuir. Sin embargo, el modelo estandarizado puede interpretar mejor los datos clínicos en términos biológicos, ya que no implica la afirmación

Estimaciones			
Parámetros	Valores Iniciales	Mod. Estandarizado	Mod. Básico
β	0.002	0.04051	0.0665614
κ	7.948337	8.00825	7.9742659
m	3.040533e-09	0.00101	0.0005006
n	0.9941047	0.99606	0.9997058
σ	0.05	0.04984	0.0503244

Tabla 4.19: Estimaciones de los parámetros de los modelos estandarizado y básico, con terapia.

Estimaciones			
Parámetros	Valores Iniciales	Mod. Estandarizado	Mod. Básico
β	0.002	0.04717	0.03224
κ	7.948337	8.07586	7.91029
σ	0.05	0.04667	0.04672

Tabla 4.20: Estimaciones de los parámetros de los modelos estandarizado y básico, post-terapia.

absurda de que una persona con un hígado más pequeño puede ser más resistente a las infecciones virales que una persona con un hígado grande. Luego de retirada la terapia la carga viral aumenta rápidamente. En la Figura 4.7 se muestra un buen ajuste entre los datos y la solución numérica utilizando las estimaciones bayesianas de los modelos estandarizado y básico, post-terapia, ambos modelos arrojan resultados similares.

Para el número reproductivo básico de la infección del modelo estandarizado, con y post-terapia, los resultados obtenidos en esta investigación tienen igual valores a los establecidos por Min, Su y Kuang (2009) para el modelo estandarizado con terapia y post-terapia. Los resultados estadísticos obtenidos en este trabajo, luego de la selección de modelos son congruentes con los biológicos, ya que con terapia y post-terapia, el mejor modelo es el estandarizado. Esta afirmación coincide con las conclusiones obtenidas por Min, Su

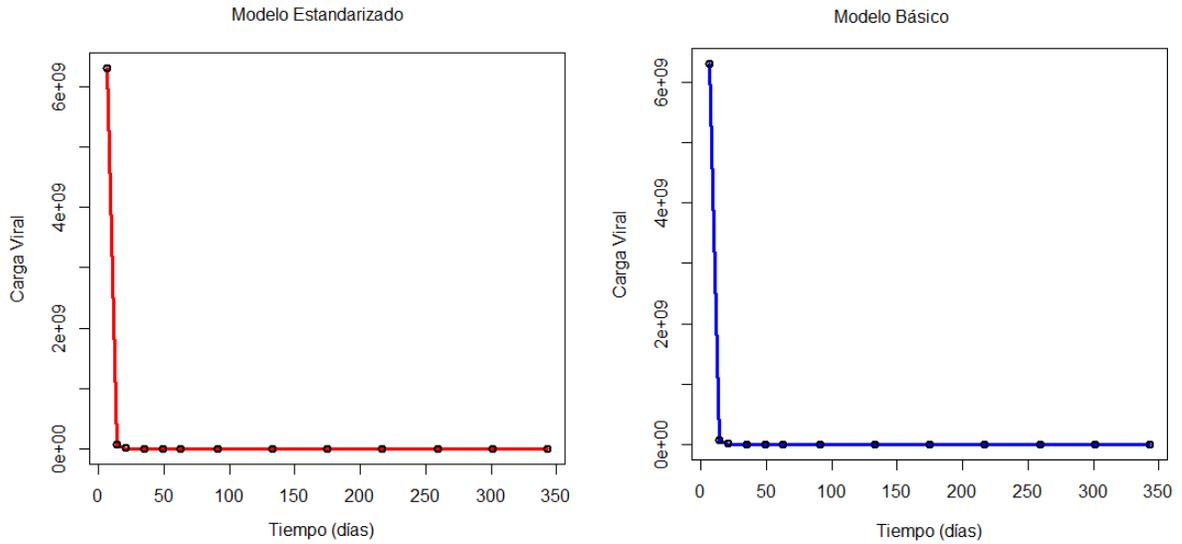


Figura 4.6: Ajuste entre los datos de carga viral y la solución numérica utilizando las estimaciones bayesianas de los modelos estandarizado y básico, con terapia.

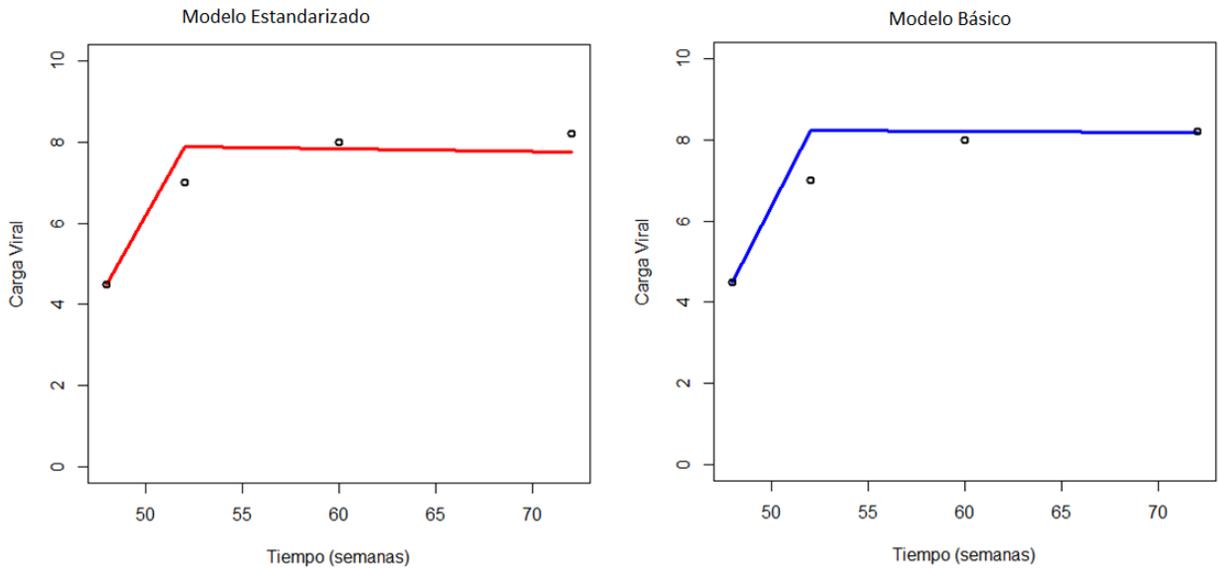


Figura 4.7: Ajuste entre los datos de carga viral y la solución numérica utilizando las estimaciones bayesianas de los modelos estandarizado y básico, post-terapia.

y Kuang (2009), en el sentido de que el modelo estandarizado es mejor que el básico porque no depende del tamaño del órgano.

Conclusiones

- La estimación bayesiana de los parámetros en los modelos dinámicos BVIM y AB-VIM, utilizando mediciones de carga viral total en pacientes con VHB positiva, con y sin terapia con lamivudina, presenta un buen ajuste a los datos reales.
- Acorde con la literatura y con los resultados de este trabajo, los parámetros de eficacia de tratamiento para las tasas de infección y la producción viral, así como la tasa de infección y la producción de partículas virales, incluidos en modelos matemáticos que describen el proceso de infección del VHB, pueden representarse mediante distribuciones de probabilidad beta y gamma, respectivamente.
- Los intervalos creíbles obtenidos del R_0 permiten al investigador concluir si hay cura o persiste la infección, ya que no incluye al uno.
- Los resultados de los números reproductivos de la infección para el modelo básico con terapia y post-terapia, muestran una sobreestimación de la presencia de la infección de VHB, debido a que dependen del tamaño del órgano.

- Utilizando criterios estadísticos para seleccionar modelos, se concluye que el mejor es el modelo estandarizado con efecto de terapia y post-terapia.
- Entre los futuros trabajos de investigación sobre VHB se sugiere utilizar alternativas diferentes a la solución del sistema de ecuaciones diferenciales y otros modelos dinámicos utilizando esta misma metodología bayesiana.

Referencias

1. Boersch-Supan, P. H., Ryan, S. J., & Johnson, L. R. (2017). deBInfer: Bayesian inference for dynamical models of biological systems in R. *Methods in Ecology and Evolution*, 8(4), 511-518.
2. Capistrán, M. A., Christen, J. A., & Donnet, S. (2016). Bayesian analysis of ODEs: Solver optimal accuracy and Bayes factors. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1), 829-849.
3. Collier, L., & Oxford, J. (2008). *Virología humana* (No. 616-022). McGraw-Hill,.
4. Gelfand, A. E., & Dey, D. K. (1994). Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 501-514.
5. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., Rubin, D. B. (2014). *Bayesian data analysis*. CRC press
6. Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the american statistical association*, 90(430), 773-795.
7. Kass, R. E. (1988). Asymptotics in Bayesian computation (with discussion). *Bayesian Statistics* 3, 261-278.
8. Min, L., Su, Y., and Kuang, Y. (2008). Mathematical analysis of a basic virus infection model with application to HBV infection. *The Rocky Mountain Journal of Mathematics*, 1573-1585.

9. Newton, M. A., & Raftery, A. E. (1994). Approximate Bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society. Series B (Methodological)*, 3-48.
10. Nowak, M. A., Bonhoeffer, S., Hill, A. M., Boehme, R., Thomas, H. C., & McDade, H. (1996). Viral dynamics in hepatitis B virus infection. *Proceedings of the National Academy of Sciences*, 93(9), 4398-4402.
11. Nowak, M., Nowak, M. A., & May, R. M. (2000). *Virus dynamics: mathematical principles of immunology and virology*. Oxford university press.
12. Ntzoufras, I. (2011). *Bayesian modeling using WinBUGS (Vol. 698)*. John Wiley & Sons.
13. Pascuet, E. R., Creus, M. T., Martín, M. T., & del Cacho, E. D. C. (2011). Características de los fármacos antirretrovirales. *Enfermedades infecciosas y microbiología clínica*, 29(5), 362-391.
14. Perelson, A. S., and Nelson, P. W. (1999). Mathematical analysis of HIV-1 dynamics in vivo. *SIAM review*, 41(1), 3-44.
15. Perelson, A. S. (2002). Modelling viral and immune system dynamics. *Nature Reviews Immunology*, 2(1), 28.
16. Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R news*, 6(1), 7-11.
17. Rizzo, M.(2007) *Statistical computing with R. Computer Science and Data Analysis Serie*.
18. Ripoll, C., & Bañares, R. (2012). Biopsia hepática. *Medicine-Programa de Formación Médica Continuada acreditado*, 11(12), 677-681.
19. Robert, C. P., Casella, G., & Casella, G. (2010). *Introducing Monte Carlo methods with r (Vol. 18)*. New York: Springer.
20. Rodríguez, M. A. (1999). *Ecuaciones diferenciales ordinarias*.

21. Siegfried, K. I., & Sansó, B. (2006). Two Bayesian methods for estimating parameters of the von Bertalanffy growth equation. *Environmental Biology of Fishes*, 77(3-4), 301-308.
22. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(3), 485-493.
23. Stone, M. (1979). Comments on model selection criteria of Akaike and Schwarz. *Journal of the Royal Statistical Society. Series B (Methodological)*, 276-278.
24. Süli & Mayers, D. F.(2003). *An Introduction to Numerical Analysis*. Cambridge University Press. Cambridge. UK.
25. Ushakova, A., Pettersen, F. O., Mæland, A., Lindqvist, B. H., and Kvale, D. (2013). Bayesian estimation of HIV-1 dynamics in vivo. *Mathematical medicine and biology: a journal of the IMA*, 32(1), 39-56.
26. Soetaert, K. E. R., Petzoldt, T., & Setzer, R. W. (2010). Solving differential equations in R: package deSolve. *Journal of Statistical Software*, 33.
27. Vegvari, C., Cauët, E., Hadjichrysanthou, C., Lawrence, E., Weverling, G. J., de Wolf, F., & Anderson, R. M. (2016). Using clinical trial simulators to analyse the sources of variance in clinical trials of novel therapies for acute viral infections. *PloS one*, 11(6), e0156622.
28. WHO. <http://www.who.int/mediacentre/factsheets/>

Anexos

En este anexo se muestran las trazas y las distribuciones a posteriori de los parámetros estimados en el modelo básico con tratamiento, para realizar las estimaciones se utilizaron los 13 datos de carga viral de los pacientes durante el tratamiento de lamivudina. Además, se muestran los gráfico de los diagnósticos de convergencia de Gelman-Rubin y Geweke.

Trazas y distribuciones a posteriori

En cada uno de los gráficos se puede observar un buen mezclado de las cadenas y las distribuciones a priori generan distribuciones a posteriori cerradas

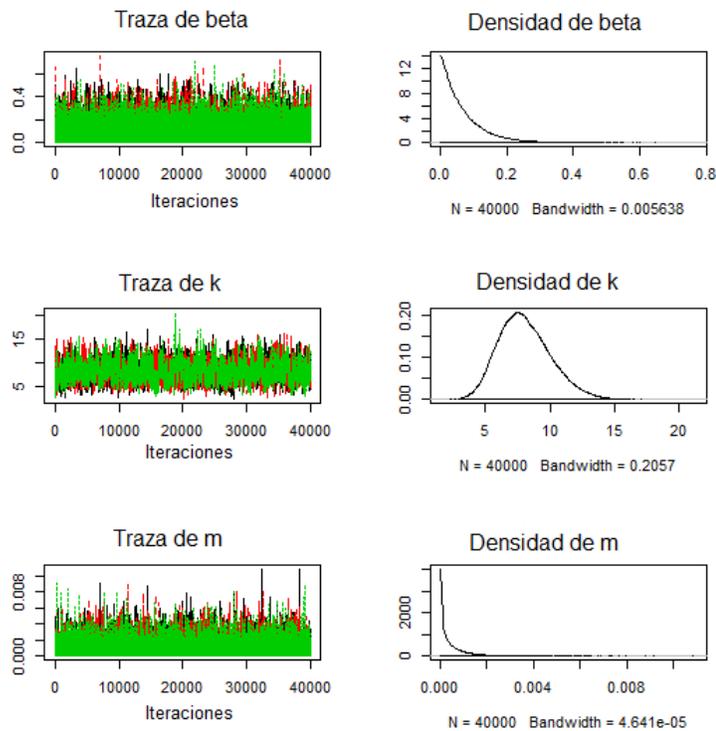


Figura 1: Trazas y densidades del modelo básico con tratamiento.

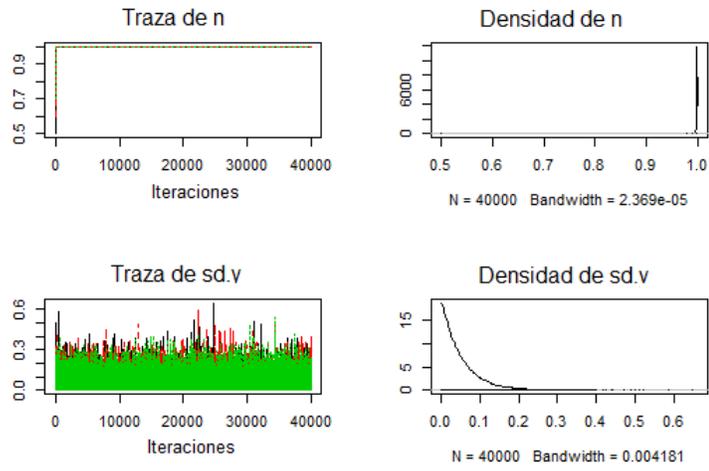


Figura 2: Trazas y densidades del modelo básico con tratamiento.

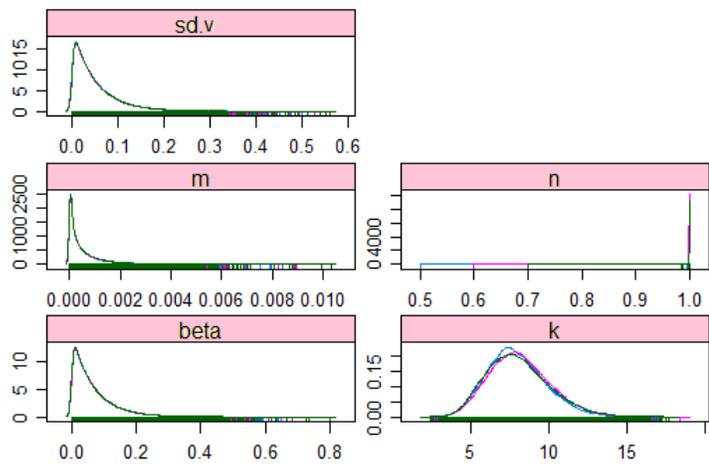


Figura 3: Distribuciones a posteriori de los parámetros del modelo básico con tratamiento.

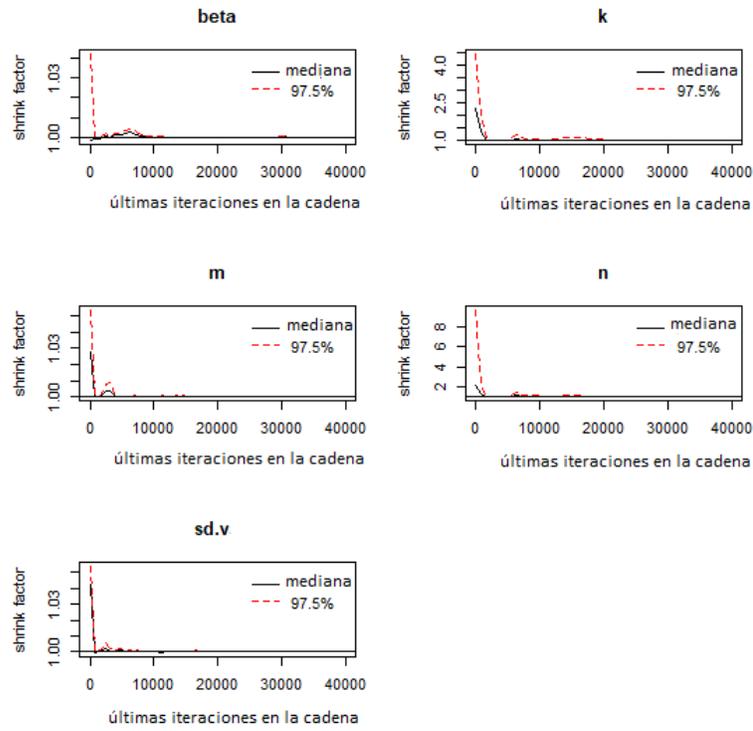


Figura 4: Gráfico de diagnóstico de convergencia de Gelman-Rubin de las estimaciones para los parámetros del modelo básico con tratamiento.

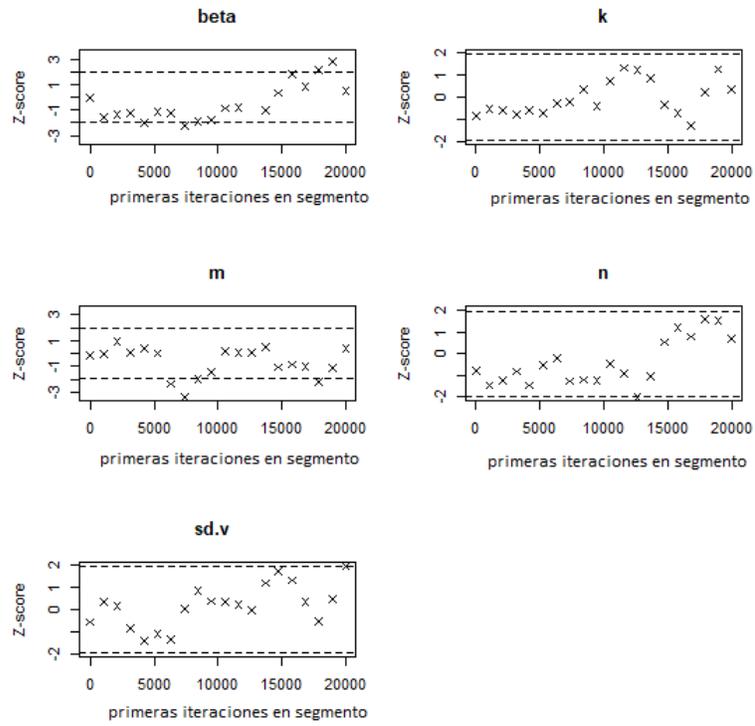


Figura 5: Gráfico de diagnóstico de convergencia de Geweke para el modelo básico con tratamiento.

En este anexo se muestran las trazas y las distribuciones a posteriori de los parámetros estimados en el modelo básico post-tratamiento, para realizar las estimaciones se utilizaron los 4 datos de carga viral de los pacientes después de suspendido el tratamiento de lamivudina. Además, se muestran los gráfico de los diagnósticos de convergencia de Gelman-Rubin y Geweke.

Trazas y distribuciones a posteriori

En cada uno de los gráficos se puede observar un buen mezclado de las cadenas y las distribuciones a priori generan distribuciones a posteriori cerradas

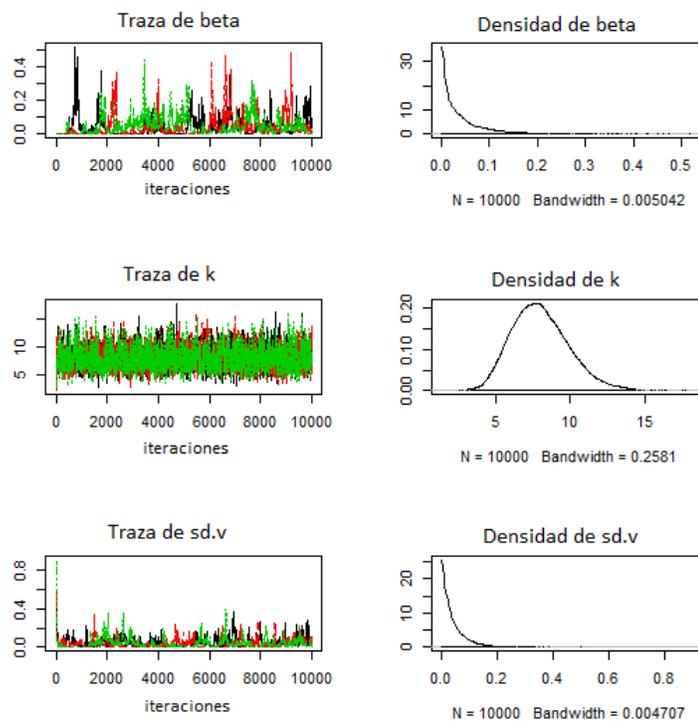


Figura 6: Trazas y densidades del modelo básico post-terapia.

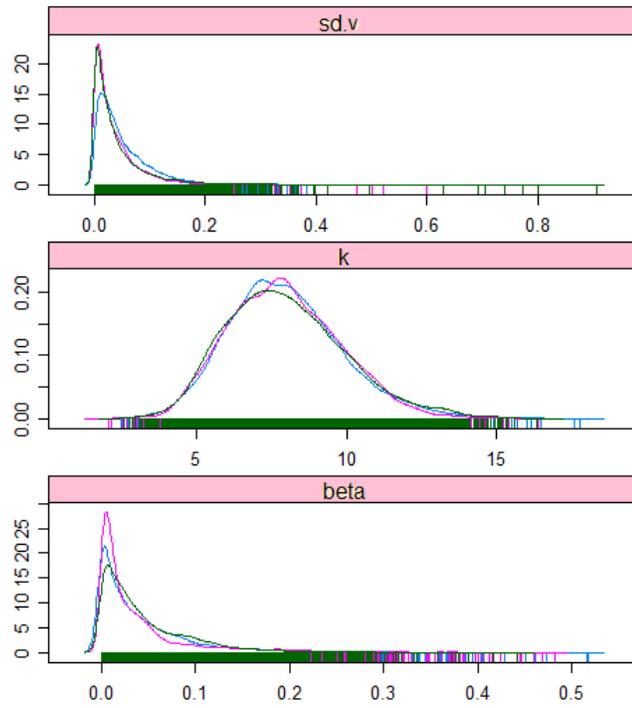


Figura 7: Distribuciones a posteriori de los parámetros del modelo básico post-terapia.

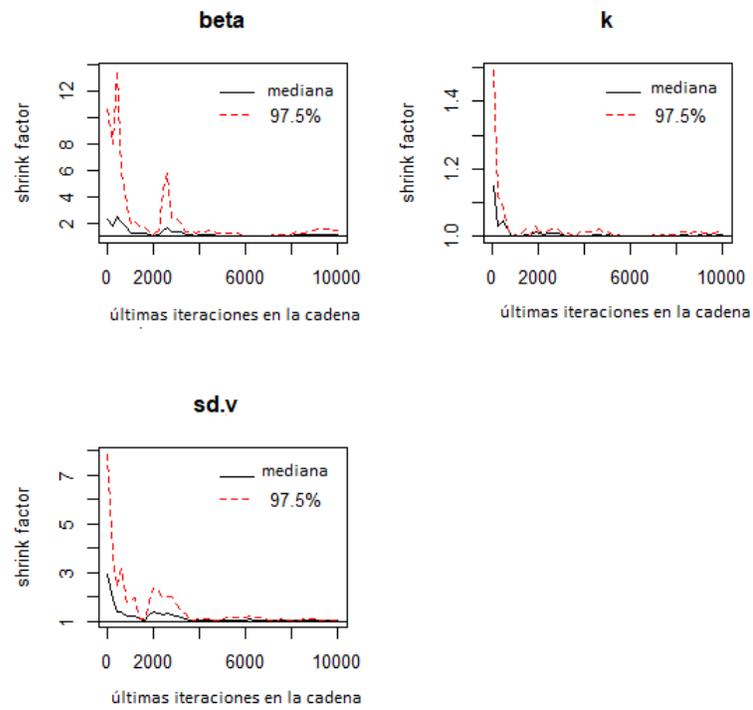


Figura 8: Gráfico de diagnóstico de convergencia de Gelman-Rubin de las estimaciones para los parámetros del modelo básico post-terapia.

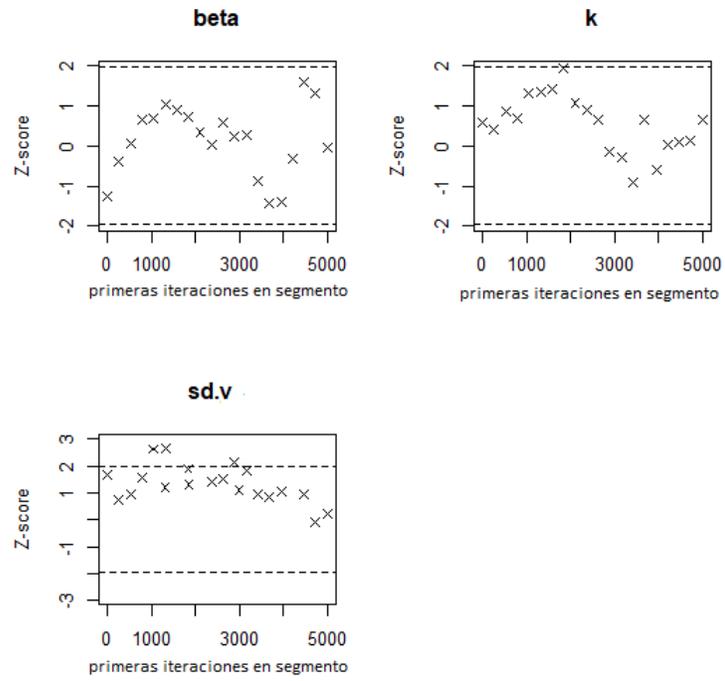


Figura 9: Gráfico de diagnóstico de convergencia de Geweke para el modelo básico post-terapia.

En este anexo se muestran las trazas y las distribuciones a posteriori de los parámetros estimados en el modelo estandarizado con tratamiento, para realizar las estimaciones se utilizaron los 13 datos de carga viral de los pacientes durante el tratamiento de lamivudina. Se muestran los gráfico de los diagnósticos de convergencia de Gelman-Rubin y Geweke.

Trazas y distribuciones a posteriori

En cada uno de los gráficos se puede observar un buen mezclado de las cadenas y las distribuciones a priori generan distribuciones a posteriori cerradas

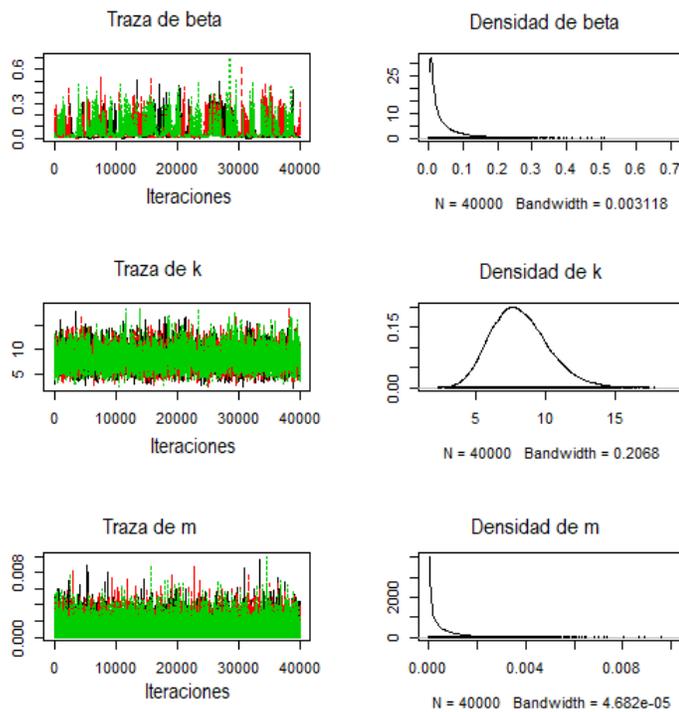


Figura 10: Trazas y densidades del modelo estandarizado con tratamiento.

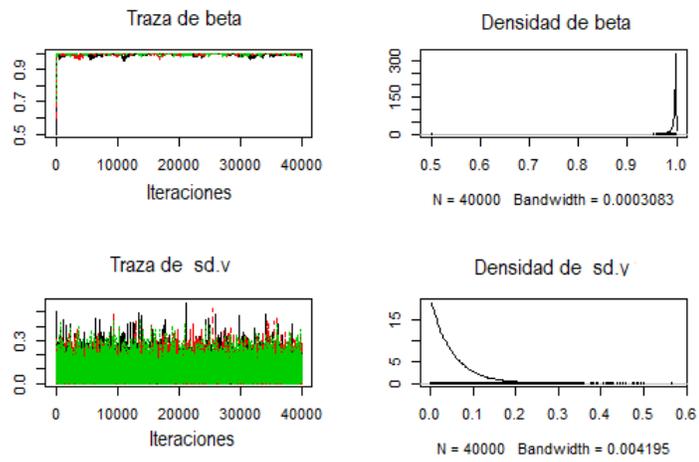


Figura 11: Trazas y densidades del modelo estandarizado con tratamiento.

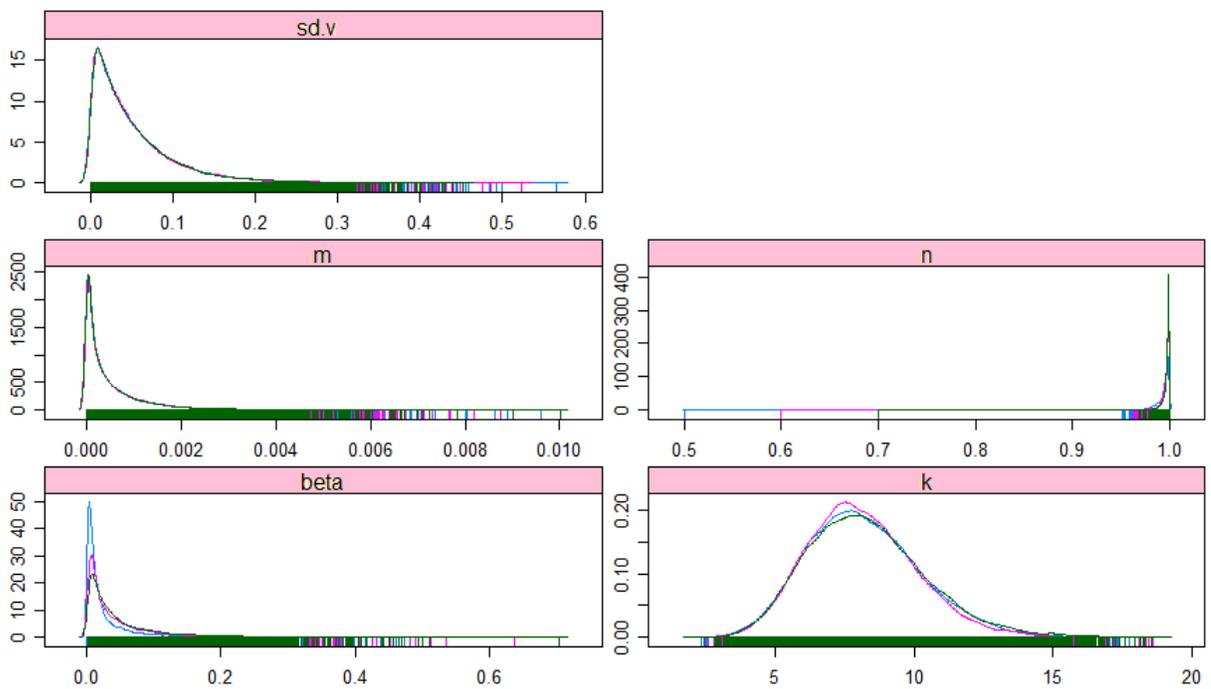


Figura 12: Distribuciones a posteriori de los parámetros del modelo estandarizado con tratamiento.

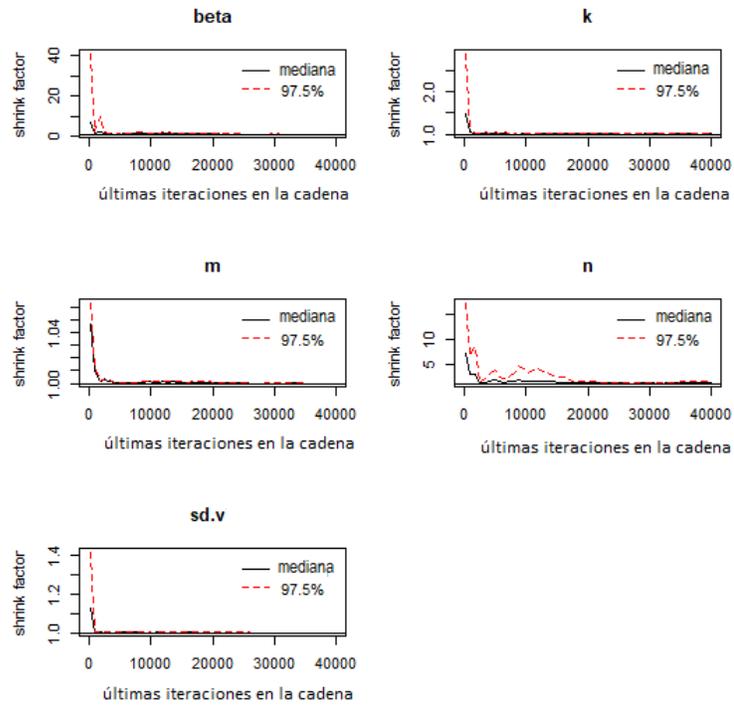


Figura 13: Gráfico de diagnóstico de convergencia de Gelman-Rubin de las estimaciones para los parámetros del modelo estandarizado con tratamiento.

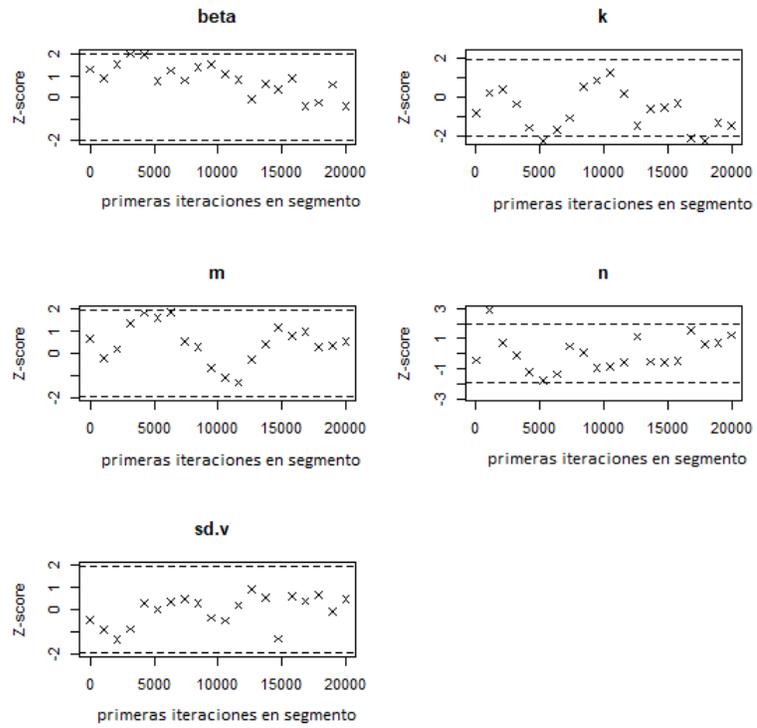


Figura 14: Gráfico de diagnóstico de convergencia de Geweke para el modelo estandarizado con tratamiento.

En este anexo se muestran las trazas y las distribuciones a posteriori de los parámetros estimados en el modelo estandarizado post-tratamiento, para realizar las estimaciones se utilizaron los 4 datos de carga viral de los pacientes después de suspendido el tratamiento de lamivudina. Se muestran los gráfico de los diagnósticos de convergencia de Gelman-Rubin y Geweke.

Trazas y distribuciones a posteriori

En cada uno de los gráficos se puede observar un buen mezclado de las cadenas y las distribuciones a priori generan distribuciones a posteriori cerradas

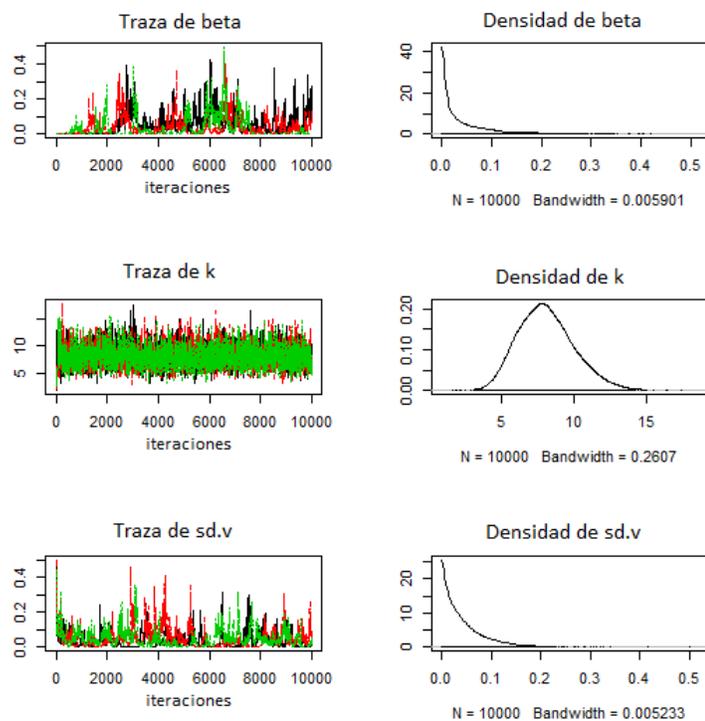


Figura 15: Trazas y densidades del modelo estandarizado post-terapia.

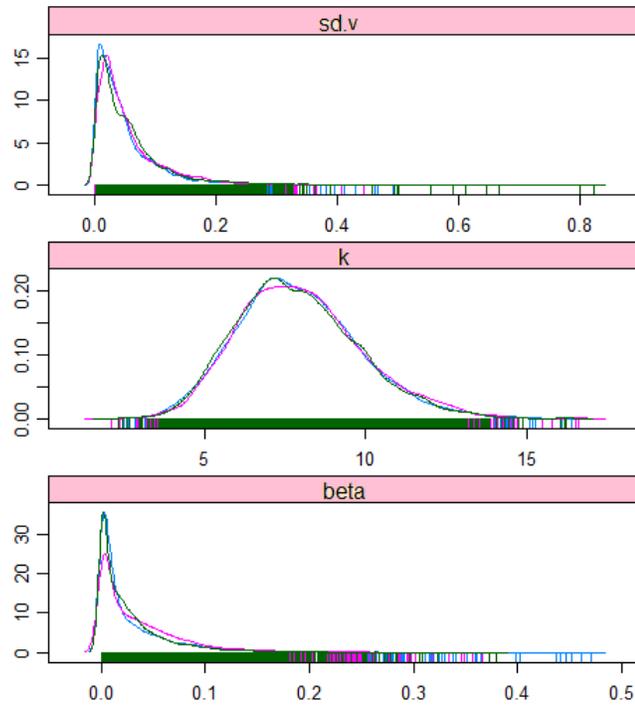


Figura 16: Distribuciones a posteriori de los parámetros del modelo estandarizado post-terapia.

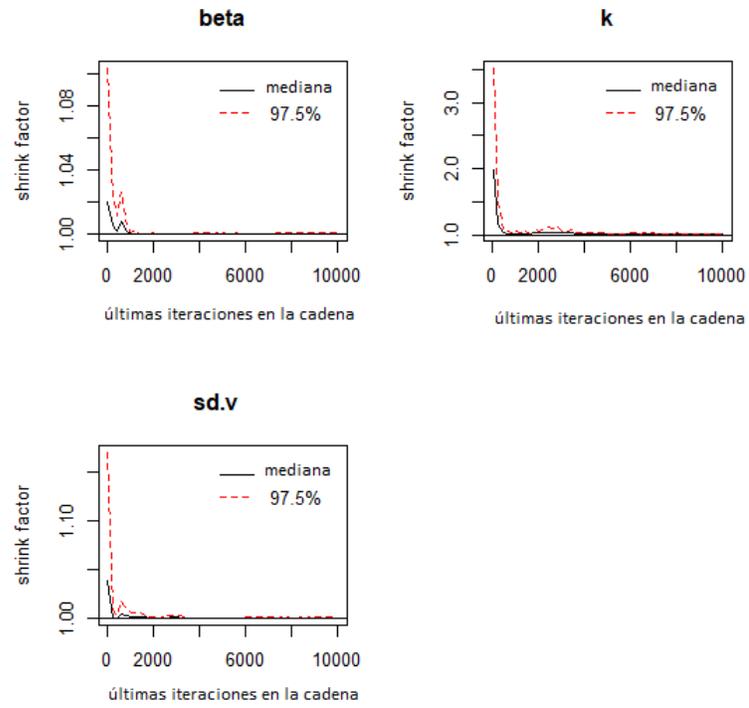


Figura 17: Gráfico de diagnóstico de convergencia de Gelman-Rubin de las estimaciones para los parámetros del modelo básico post-terapia.

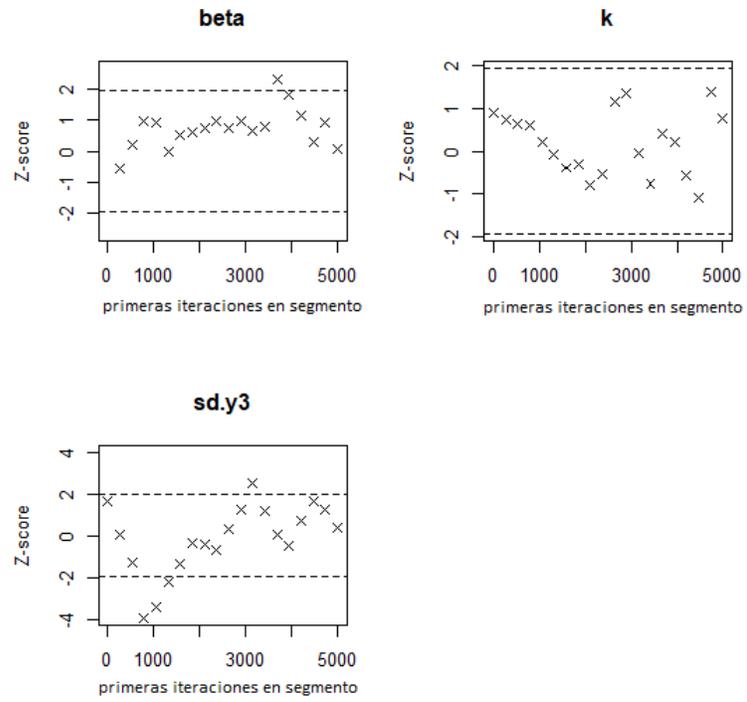


Figura 18: Gráfico de diagnóstico de convergencia de Geweke para el modelo estandarizado post-terapia.