



Universidad Autónoma de Guerrero

Facultad de Matemáticas

Maestría en Matemáticas Aplicadas

**Estimación Bayesiana de parámetros en
un modelo para la dinámica de infección
del VISH - KS661 *in vitro***

T E S I S

PARA OBTENER EL GRADO DE:

Maestro en Matemáticas Aplicadas

PRESENTA:

Lic. Blanca Noemi Cortez Galeana

DIRECTORES DE TESIS:

Dr. Ramón Reyes Carreto

M.C. Cruz Vargas de León

diciembre 13 de 2018.

Dedicatoria

Dedico esta tesis a mis padres:

M. Guadalupe Galeana Rumbo
Julian Cortez Rios

Sara Rumbo Márquez
Rafael Galeana Rios

Porque sin ustedes simplemente, no estaría aquí, ¡Los amo!

Agradecimientos

A mi madre y padre, por todos los sacrificios que hicieron para darme la oportunidad de estudiar y por todas sus valiosas enseñanzas.

A mis hermanos, porque siempre estuvieron conmigo en los momentos más difíciles.

A Jesús Emigdio Flores Terrero por todos sus consejos y su apoyo incondicional.

A la familia Flores por cuidar de mí y todo el apoyo brindado.

A mis compañeros de maestría: Laima, Dannier, Edgar, Eduard Y Omar.

A mis asesores; el Dr. Ramón Reyes Carreto y el M.C. Cruz Vargas de León por su tiempo, dedicación y permitirme trabajar a su lado.

A mis revisores de tesis, la Dra. María Guzmán Martínez, el Dr. Flaviano Godínez Jaimes y el Dr. Raúl Peralta Rodríguez, por el tiempo y consejos que me brindaron para un mejor trabajo.

Al Dr. Marco Antonio Taneco Hernández por su apoyo y disponibilidad durante mis estudios.

Al Programa de Incorporación de Mujeres Indígenas en Posgrados para el Fortalecimiento Regional del Estado de Guerrero, por el apoyo económico y asesoramiento.

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

Índice general

Resumen	VIII
Abstract	IX
1 Introducción	1
1.1 Planteamiento del problema	2
Planteamiento del problema	2
1.2 Objetivos	3
Objetivos	3
2 Marco teórico	4
2.1 Biología de los retrovirus	4
2.1.1 Estructura de los virus	5
2.1.2 Ciclo de replicación del retrovirus	6
2.2 Modelos matemáticos en infecciones virales	7
2.2.1 Modelo básico de infección <i>in vivo</i>	8
2.2.2 Modelo de infección <i>in vitro</i>	10
2.3 Índice de la Multiplicidad de Infección (MOI)	12
2.4 Estimación Bayesiana	14
2.4.1 Modelo estadístico	14
2.4.2 Estimación de parámetros	15
2.4.3 Diagnósticos de convergencia	16
2.4.3.1 Diagnóstico de Gelman–Rubin	17
2.4.3.2 Diagnóstico de Geweke	18
3 Metodología	19

3.1	Experimento	19
3.2	Datos	21
3.3	Modelo <i>ad-hoc</i> de infección del VISH - KS661 <i>in vitro</i>	22
3.4	Inferencia Bayesiana	23
4	Resultados y Discusión	26
4.1	Resultados	26
4.1.1	Estimación mediante el método de mínimos cuadrados	26
4.1.2	Estimación Bayesiana	29
4.2	Discusión	34
5	Conclusiones	39
	Anexos	40
A	Gráficos y cuadros correspondientes al MOI 2×10^{-3}	41
B	Gráficos y cuadros correspondientes al MOI 2×10^{-6}	44
	Referencias	49

Índice de figuras

2.1	Estructura de un virus	5
2.2	Ciclo de replicación del retrovirus	7
2.3	Diagrama del modelo básico de infecciones virales <i>in vivo</i>	9
2.4	Representación esquemática para el modelo <i>in vitro</i>	11
3.1	Pocillos.	19
3.2	Representación esquemática de las nueve mediciones	20
4.1	Datos reales y los resultados de la solución numérica usando mínimos cuadrados para el MOI 2×10^{-3}	28
4.2	Comparación de los datos reales con los resultados de la solución numérica usando mínimos cuadrados para el MOI 2×10^{-6}	29
4.3	Solución numérica del modelo <i>ad-hoc</i> usando resultados de la estimación Bayesiana para los datos del MOI 2×10^{-3}	31
4.4	Datos reales y solución numérica del modelo <i>ad-hoc</i> usando resultados de la estimación Bayesiana	33
4.5	Intervalo creíble para R_0^{ad-hoc} y valores de R_0^{ad-hoc} obtenidos por ambos métodos para el MOI 2×10^{-3}	35
4.6	Comparación de los datos reales, solución obtenida por Iwami et al. (2012) y las soluciones numéricas del modelo <i>ad-hoc</i> usando mínimos cuadrados e inferencia Bayesiana para el MOI 2×10^{-3}	35
4.7	Comparación de los datos reales, solución obtenida por Iwami et al. (2012) y las soluciones numéricas del modelo <i>ad-hoc</i> usando mínimos cuadrados e inferencia Bayesiana para el MOI 2×10^{-6}	37
A.1	Diagnóstico de convergencia de Gelman–Rubin	42
A.2	Trazas y distribuciones <i>a posteriori</i> para β_{50} , a , y k	42

A.3	Trazas y distribuciones <i>a posteriori</i> para k_{50} y σ^2	43
B.1	Diagnóstico de convergencia de Gelman–Rubin	45
B.2	Trazas y distribuciones <i>a posteriori</i> para β_{50} , a , y k	45
B.3	Trazas y distribuciones <i>a posteriori</i> para k_{50} y σ^2	46

Índice de cuadros

2.1	Supuestos de la infección por virus <i>in vitro</i> e <i>in vivo</i>	10
3.1	Datos del experimento <i>in vitro</i>	21
4.1	Condiciones iniciales para las variables del modelo <i>ad-hoc</i>	26
4.2	Estimaciones obtenidas mediante el método de mínimos cuadrados	27
4.3	Estimaciones Bayesianas usando los datos del MOI 2×10^{-3}	30
4.4	Estimaciones Bayesianas usando los datos del MOI 2×10^{-6}	32
4.5	Estimaciones obtenidas mediante el método de mínimos cuadrados e inferencia Bayesiana (MOI 2×10^{-3})	34
4.6	Estimaciones obtenidas mediante el método de mínimos cuadrados e inferencia Bayesiana (MOI 2×10^{-6})	36
A.1	Test de Gelman–Rubin para el MOI 2×10^{-3}	41
B.1	Test de Gelman–Rubin para el MOI 2×10^{-6}	44

Resumen

El Virus de Inmunodeficiencia Humana-1 (VIH-1) es uno de los causantes del SIDA. El VIH-1 causa la mayoría de las infecciones por el VIH alrededor del mundo. El desarrollo y uso de modelos matemáticos para comprender cuantitativamente la cinética del VIH-1 ha resultado ser una herramienta útil para determinar la patogenicidad y la transmisibilidad de este virus, predecir el curso de la enfermedad y evaluar los efectos de la terapia antiviral. Ante la limitación de estudios directos para investigar la patogenicidad del VIH-1 *in vivo*, se han utilizado modelos para explicar la dinámica del Virus de Inmunodeficiencia Simio/Humano (VISH) *in vitro*.

El estudio del VISH-KS661 (cepa del VISH altamente patógena) proporciona información importante para la comprensión de la patogénesis del VIH-1. Los objetivos de este trabajo son: 1) estimar bajo la perspectiva bayesiana los parámetros del modelo que explican la dinámica viral del VISH-KS661 *in vitro*; 2) estimar el número reproductivo básico de la infección, a partir de las muestras *a posteriori* de los parámetros de interés; 3) estimar las cantidades relevantes de la cinética viral como: promedio de vida y vida media de las células infectadas, y el tamaño de la explosión viral total e infecciosa. Para cumplir con los objetivos antes señalados, se usan los datos de un experimento del VISH-KS661 *in vitro* utilizados por [Iwami et al. \(2012\)](#), así como el modelo de la dinámica viral, que describe las células T CD4⁺, células infectadas, virus no infeccioso y virus infeccioso. Para la estimación de los parámetros, se utiliza la media de la distribución *a posteriori* obtenida por simulación con el método MCMC, así como los intervalos creíbles de los parámetros.

Abstract

The Human Immunodeficiency Virus-1 (HIV-1) is one of the causes of AIDS. HIV-1 causes the majority of HIV infections around the world. The development and use of mathematical models to quantitatively understand the kinetics of HIV-1 has proved to be a useful tool for determining the pathogenesis and transmissibility of this virus, predicting the course of the disease and evaluating the effects of antiviral therapy. In view of the limitations of direct studies to investigate the pathogenicity of HIV-1 *in vivo* the models have been used to explain the dynamics of the Simian/Human Immunodeficiency Virus (SHIV) *in vitro*. The study of SHIV-KS661 (highly pathogenic SHIV strains) provides important information for the understanding of the pathogenesis of HIV-1. The goals of this work are: 1) to estimate the parameters using Bayesian approach of an viral dynamics model to experimental data of SHIV-KS661 *in vitro*; 2) estimate the basic reproductive number of the infection, using the *a posteriori* distribution of the parameters of interest; 3) estimate relevant amounts of viral kinetics as: average life of the infected cells, burst size of the total and infectious viral particles. In order to fulfill the aforementioned objectives, we used the data from cell cultures on the infection of HSC-F cells with SHIV-KS661 *in vitro* obtained in [Iwami et al. \(2012\)](#), and as well a viral dynamics model, which describes the CD4⁺ T cells, infected cells, non-infectious virus and infectious virus. For the estimation of the parameters, we computed the *a posteriori* distribution using MCMC methods. Finally estimate the mean and credible intervals of the parameters.

Introducción

Diversos modelos matemáticos se han formulado con base a la ley de acción de masas, que rige la propagación de un virus dentro de un tejido u órgano. Estos modelos se han utilizado para determinar el impacto del virus en el sistema inmunológico y para probar la capacidad de respuesta del sistema inmune al tratamiento.

El desarrollo de una comprensión cuantitativa de la cinética viral es útil para determinar la patogénesis y la transmisibilidad del virus, predecir el curso de la enfermedad y evaluar los efectos de la terapia antiviral. Sin embargo, la disponibilidad de datos en estudios clínicos, animales y de cultivos celulares ha sido bastante limitada. Muchos estudios de la cinética de infección viral se han basado únicamente en las mediciones del recuento total o infeccioso del virus. Se introduce un modelo matemático que considera la carga viral tanto infecciosa como total, así como la fracción de células infectadas y no infectadas dentro de un cultivo celular, y se aplican para analizar los datos en el tiempo de una infección por el Virus de Inmunodeficiencia Simio/Humano (VISH) *in vitro* ([Iwami et al., 2012](#)).

Una de las cepas del VISH altamente patógenas, es VISH-KS661. Para describir la cinética *in vitro* de cultivos celulares con infección por el VISH-KS661 en el sistema experimental, se generalizó el modelo básico que describe las células T CD4⁺, células infectadas y virus libres, ahora considerando si el virus es infeccioso o no infeccioso.

En este trabajo se utiliza un modelo estadístico para analizar la dinámica del VISH-KS661 desde la perspectiva de los modelos Bayesianos, el reto consiste en estimar parámetros de un sistema

de Ecuaciones Diferenciales Ordinarias (EDO), incluso cuando dicho sistema sea no lineal y no tenga solución analítica.

Para la estimación de los parámetros del sistema de EDO, se asignarán distribuciones *a priori* adecuadas para cada uno de ellos y así obtener la distribución *a posteriori* de los parámetros mediante la aproximación de Cadenas de Markov Monte Carlo (MCMC, por sus siglas en inglés), dicho método nos ayudará a muestrear las distribuciones *a posteriori* de los parámetros.

1.1. Planteamiento del problema

El desarrollo y uso de modelos matemáticos para comprender cuantitativamente la cinética del Virus de Inmunodeficiencia Humana-1 (VIH-1) ha resultado de bastante utilidad para determinar la patogenia y la transmisibilidad del virus, predecir el curso de la enfermedad y evaluar los efectos de la terapia antiviral ([Putter et al., 2002](#)).

Los análisis de datos experimentales que utilizan modelos matemáticos, mediante la estimación de los parámetros de la infección, han proporcionado información cuantitativa sobre la cinética de las infecciones virales, particularmente para el VIH-1. Los datos virológicos disponibles, incluso para experimentos *in vitro*, a menudo se han visto limitados, ya que muchos análisis de modelado se han basado solo en datos de carga viral total o datos de carga viral infecciosa.

Se han desarrollado y utilizado los VISH para investigar la patogenicidad del VIH-1 *in vivo* ante la limitación de estudios directos sobre el VIH-1.

Varios estudios que modelan la cinética de la infección viral se han basado únicamente en medidas del recuento total o infeccioso de virus ([Lewin et al., 2001](#); [Nowak et al., 1996](#); [Tsiang y Gibbs, 1999](#)).

El modelo propuesto por [Iwami et al. \(2012\)](#) que estudia la cinética viral *in vitro* de la infección del VISH-KS661, consta de 4 ecuaciones diferenciales ordinarias que describen la dinámica de las células blanco, células infectadas, carga viral no infecciosa e infecciosa.

El problema consiste en estimar los parámetros que representan las tasas de transmisión infecciosa, de muerte de las células infectadas, carga viral total e infecciosa del modelo matemático *ad-hoc* que considera la carga viral total en lugar de la carga viral no infecciosa, agregando las tasas de cosechas y los resultados del experimento de infección por VISH-KS661 *in vitro*, para proporcionar información cuantitativa sobre la cinética de la infección por VISH que podría

utilizarse para mejorar la comprensión de la patogénesis de VIH-1.

1.2. Objetivos

Objetivo General:

Estimar, mediante el enfoque Bayesiano, los parámetros del modelo *ad-hoc* de infección del VISH-KS661 *in vitro* propuesto por [Iwami et al. \(2012\)](#), utilizando las concentraciones de células blanco, células infecciosas, carga viral total y carga viral infecciosa.

Objetivos Particulares:

- Estimar los parámetros relacionados con las tasas de infección, muerte, carga viral total e infecciosa y sus intervalos creíbles.
- Estimar el valor del número reproductivo básico y su intervalo creíble.
- Estimar las medidas relevantes de la cinética viral del VISH-KS661.

Con el propósito de lograr los objetivos antes mencionados, este trabajo de investigación está estructurado en cinco capítulos, mismos que se describen brevemente:

En el Capítulo 1 se menciona la introducción, el planteamiento del problema y los objetivos de este trabajo de tesis. En el Capítulo 2 se describe el marco teórico de la biología del retrovirus, su importancia, los componentes del virus y el ciclo de replicación del retrovirus, los modelos matemáticos en infecciones virales, el modelo básico de infección por virus *in vivo*, el modelo de infección por virus *in vitro*, el índice de multiplicidad de infección (MOI, por sus siglas en inglés); el modelo estadístico y la teoría de inferencia Bayesiana, funciones de pérdida, y criterios de convergencia. En el Capítulo 3 se presenta el experimento, los datos obtenidos, los modelos matemáticos y estadístico utilizados, las distribuciones *a priori* y las distribuciones *a posteriori*. Los resultados y la discusión se presentan en el Capítulo 4. Finalmente, en el Capítulo 5 se presentan las conclusiones de este trabajo.

Marco teórico

2.1. Biología de los retrovirus

La palabra latina *virus* significa “fluido venenoso” y eso fue precisamente lo que los primeros virólogos creían que eran estos patógenos. En la segunda mitad del siglo XIX se lograron enormes avances en el estudio de los microbios ([Collier, 2008](#)).

- Charles Chamberland diseñó y creó un filtro que retendría hasta las bacterias más pequeñas;
- Iwanowski (en Rusia) y Beijerinck (en Holanda) demostraron que se podía transmitir una enfermedad de plantas mediante extractos de líquidos que habían pasado a través del filtro de Chamberland, y que por tanto, no podía contener bacterias.

Entonces fue obvio que había otros agentes vivos más pequeños que cualquier bacteria conocida, pero capaces de multiplicarse, los cuales podían causar una amplia gama de enfermedades en vegetales y animales.

Pasteur utilizó animales para desarrollar la vacuna contra la rabia y trabajar la inmunización contra la viruela. Sin embargo, además de consideraciones humanitarias, para gran parte del trabajo experimental utilizar animales resulta insatisfactorio en las áreas de costos y reproducibilidad, de modo que los biólogos siempre han tratado de sustituirlo con pruebas *in vitro* (tubo

de ensayo). Esto se logró en 1928 cuando por primera vez se cultivó un virus en suspensiones de tejido renal molido.

2.1.1. Estructura de los virus

Los virus se componen de dos o tres partes (ver Figura 2.1):

- a) Su material genético, el cual porta la información hereditaria, que puede ser, el ADN o el ARN:
 - ADN (**ácido desoxirribonucleico**) el cual es una macromolécula que colabora con el ARN para sintetizar proteínas vitales para el desarrollo de todas las características y funciones grabadas en el ADN.
 - ARN (**ácido ribonucleico**) es una macromolécula que ayuda al ADN en las funciones de trasmisión de genes y de sintetizar proteínas, a los virus con este génoma también se les conoce como retrovirus.
- b) Una cubierta protéica, y
- c) En algunos una bicapa lipídica.

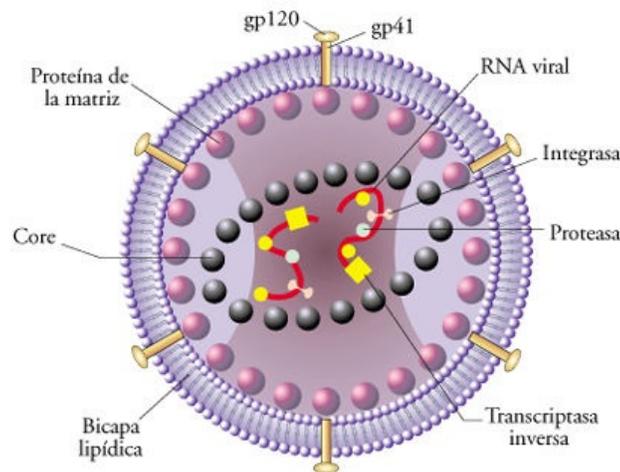


Figura 2.1: Estructura de un virus, tomada de [Andrea \(2012\)](#).

En 1910 Peyton Rous descubrió un retrovirus cuando colaboraba en el *Rockefeller Institute for Medical Research, en Nueva York*. Este agente era el virus del sarcoma aviar, después hasta

1930 se descubrieron otros retrovirus, los cuales afectaban a ratones y otros mamíferos ([Collier, 2008](#)).

El VISH-KS661 al igual que el VIH es un retrovirus, es decir, su material genético es el ARN, además de una enzima; transcriptasa inversa, que sintetiza ADN a partir de ARN. Además las infecciones con cepas del VISH-KS661 altamente patógenas en experimentos con animales, han mostrado manifestaciones clínicas estables en la mayoría de los animales infectados, similares a las infecciones humanas por VIH.

2.1.2. Ciclo de replicación del retrovirus

El retrovirus infecta a las células que tengan en su superficie la molécula $CD4^+$ (una proteína que pertenece a algunas células del sistema inmunológico y que el VIH utiliza como receptor). La gp120 viral (molécula compuesta por una proteína unida a uno o varios glúcidos) reconoce y se une específicamente al $CD4^+$, y de este modo el virus se une a la membrana celular.

Tras esta unión, el nucleóide viral se introduce en la célula y, mediante el proceso de transcripción inversa, el ARN viral se transforma en ADN de doble hebra. Este ADN viral es transportado al núcleo de la célula, donde se inserta o se integra al ADN de los cromosomas de la célula. Cuando se producen los estímulos necesarios, se desencadena el proceso de formación de nuevos viriones (partícula vírica morfológicamente completa e infecciosa). El ADN viral integrado en los cromosomas de la célula huésped se sirve de los mecanismos de replicación de ésta para su transcripción a ARN mensajero (ARNm) y a nuevas hebras de ARN genómico viral. Se produce entonces la traducción del ARNm viral a proteínas virales, y el ensamblaje de viriones nuevos dentro de la célula. Las partículas de los retrovirus creadas se liberan de la célula tomando en su salida parte de la membrana de la célula para utilizarla como cubierta. La replicación del retrovirus puede producir la muerte de los linfocitos T $CD4^+$ (uno de los distintos tipos de glóbulos blancos) ([Lomba, 2014](#)), una representación gráfica de dicho proceso se presenta en la Figura 2.2.

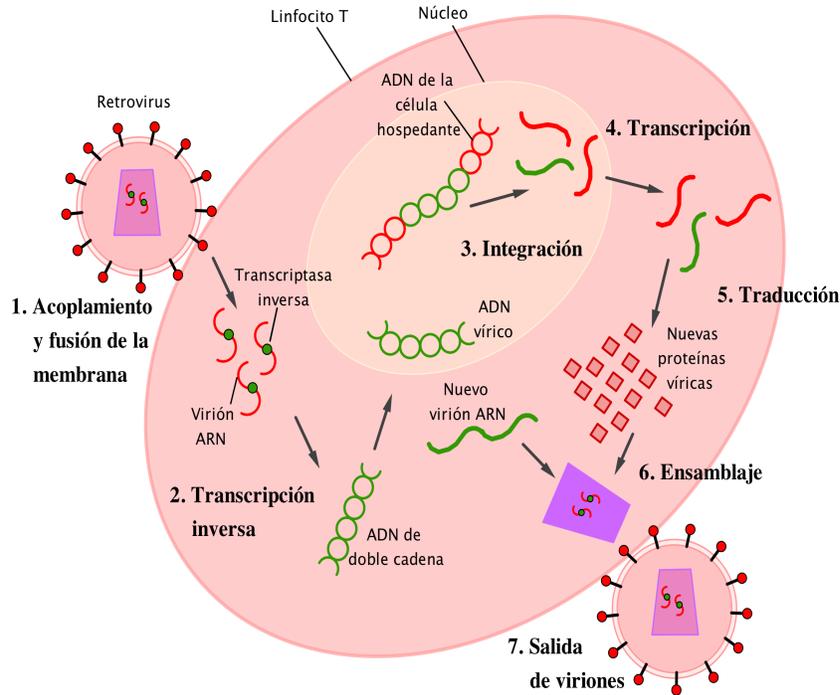


Figura 2.2: Ciclo de replicación del retrovirus.

2.2. Modelos matemáticos en infecciones virales

Diversos modelos matemáticos se han formulado con base a la ley de acción de masas que rige la propagación de un virus dentro de un tejido u órgano. Estos modelos se han utilizado para describir la dinámica del virus, entender la interacción del virus con el sistema inmune y cuantificar la eficacia de un tratamiento (Perelson, 2002).

De acuerdo con Montesinos y Hernández (2007) los modelos matemáticos son una de las herramientas utilizadas hoy en día para el estudio de problemas en medicina, biología, fisiología, bioquímica, epidemiología, entre otras áreas del conocimiento. Sus objetivos primordiales son describir, explicar y predecir procesos. Su aplicación con frecuencia se limita por la falta de conocimientos acerca de los principios de la modelación matemática.

La importancia de los modelos matemáticos para enfermedades infecciosas es evidente debido a que:

- a) La construcción de modelos revela algunas veces relaciones que no son obvias a primera vista;
- b) Una vez construido el modelo matemático es posible extraer de él propiedades y características de las relaciones entre los elementos que de otra forma permanecerían ocultas;
- c) En la mayor parte de los problemas de enfermedades infecciosas del mundo real no es factible experimentar con la realidad, ya que puede ser muy costoso, peligroso, inmoral o incluso imposible. Por lo tanto, es natural intentar superar esta dificultad con la construcción de un modelo que describa de manera adecuada las características básicas de la epidemia y entonces usar el modelo para predecir las consecuencias de introducir cambios específicos;
- d) La función principal de un modelo para una enfermedad infecciosa consiste en proveer un medio que posibilite entender la dispersión de una enfermedad infecciosa a través de una población bajo diferentes escenarios.

2.2.1. Modelo básico de infección *in vivo*

El modelo básico de infecciones virales ha sido fundamental para entender la dinámica de las infecciones producidas por los retrovirus, los cuales atacan ciertos glóbulos blancos importantes para la función del sistema inmune, conocidos como células T CD4⁺. El modelo básico para células T y dinámica del virus es un sistema de tres ecuaciones diferenciales, que representan las concentraciones para: células blanco, células infectadas y virus libres (Allen, 2007; Choisy et al., 2006).

El modelo básico de infección se explica de la siguiente forma: las células blanco, x , son infectadas por el virus, v , a una tasa de transmisión infecciosa, β . Las células blanco proliferan a una tasa constante λ y mueren a una tasa δ . Las células infectadas, y , producen nuevos viriones a una tasa p y mueren a una tasa a . Finalmente, los viriones declinan a una tasa c , ver Figura 2.3.

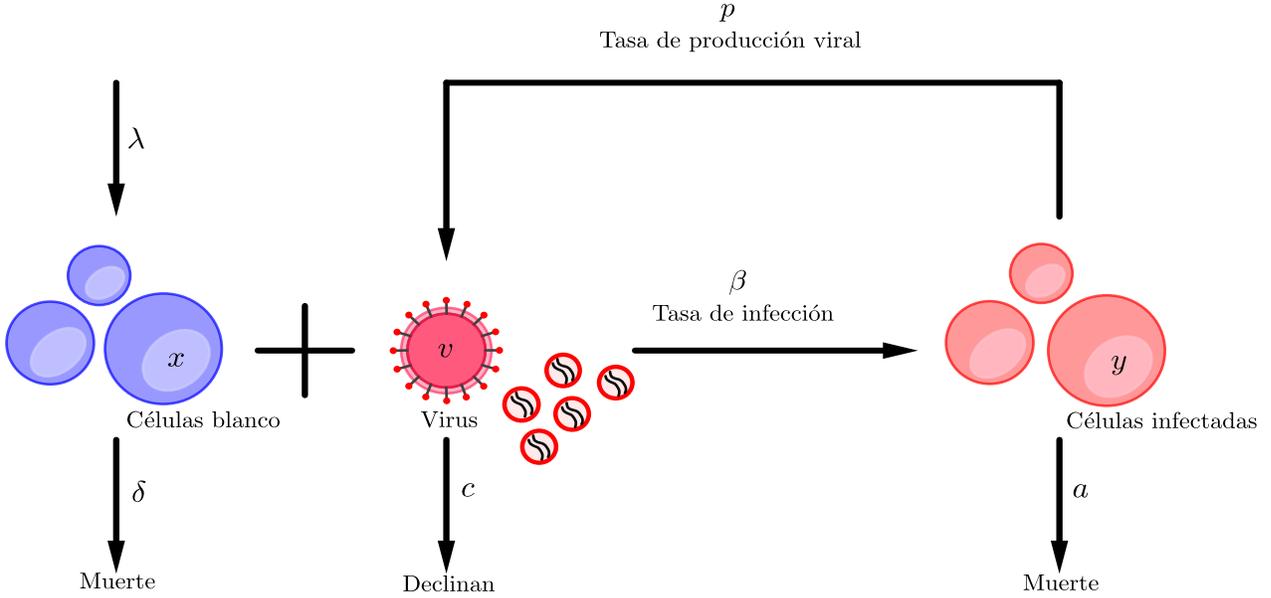


Figura 2.3: Diagrama del modelo básico de infecciones virales *in vivo*.

Las ecuaciones que describen el modelo básico de infecciones virales *in vivo* (ver [Perelson, 2002](#)) mostrado en la Figura 2.3 son:

$$\begin{aligned}
 \frac{dx}{dt} &= \lambda - \delta x - \beta xv \\
 \frac{dy}{dt} &= \beta xv - ay \\
 \frac{dv}{dt} &= py - cv
 \end{aligned}
 \tag{2.1}$$

En el sistema de EDO (2.1), el número reproductivo básico de la infección, R_0 , está determinado por la siguiente expresión:

$$R_0 = \frac{\lambda \beta p}{a \delta c}$$

R_0 proporciona información acerca de la velocidad con que una enfermedad puede propagarse en una población determinada, además de que es el umbral para determinar si hay o no infección y se interpreta como el número de infecciones secundarias que se obtienen al introducir una célula infectada en un medio susceptible.

Los posibles escenarios de la infección viral son:

- Si $R_0 \leq 1$, el sistema tiene un punto de equilibrio en:

$$E_1 = \left(\frac{\lambda}{\delta}, 0, 0 \right)$$

- Si $R_0 > 1$, el sistema tiene dos punto de equilibrio en: E_1 y E_2 , donde

$$E_2 = \left(\frac{\lambda}{\delta R_0}, (R_0 - 1) \frac{c\delta}{\beta p}, (R_0 - 1) \frac{\delta}{\beta} \right)$$

Si $R_0 \leq 1$, la infección declinará; mientras que si $R_0 > 1$ la infección persistirá (Putter et al., 2002).

2.2.2. Modelo de infección *in vitro*

A diferencia del modelo por infección *in vivo*, en el modelo *in vitro*, la dinámica de las células susceptibles está ausente (ver el sistema ecuaciones (2.2)). Cuando se trabaja con modelos de infección *in vitro* los supuestos de los modelos cambian debido a la diferencia entre los datos *in vitro* e *in vivo*. Algunas suposiciones se muestran en el Cuadro 2.1.

Cuadro 2.1: Supuestos de la infección por virus *in vitro* e *in vivo*

<i>in vivo</i>	<i>in vitro</i>
<ul style="list-style-type: none"> • Se encuentra dentro de un organismo vivo. • No se controlan las condiciones, aumenta la complejidad. • El tiempo de observación es el tiempo real de una infección en un órgano o tejido, el cual varía de individuo a individuo. • Después de obtener resultados <i>in vitro</i>, se puede llevar a cabo un mejor estudio <i>in vivo</i> para estudiar el efecto de terapias antivirales. 	<ul style="list-style-type: none"> • Se encuentra en un ambiente controlado fuera del organismo vivo. • Al ser un ambiente controlado, se simplifica la complejidad. • El tiempo de observación es limitado. • Se utiliza para estudiar el efecto de terapias antivirales y dar indicios del comportamiento <i>in vivo</i>.

Se sabe que el material genético del VISH-KS661 es ARN, el cual no es lo suficientemente estable, por lo cual al replicarse este suelen ocurrir errores, que producen que una fracción de las partículas sea infecciosa (p), las cuales pierden infectividad a una tasa r_I , convirtiéndose en partículas no infecciosas ($1 - p$). Lo que implica que la fracción de partículas $1 - p$ no tienen la capacidad de infectar nuevas células blanco, lo anterior queda descrito por el diagrama de la Figura 2.4.

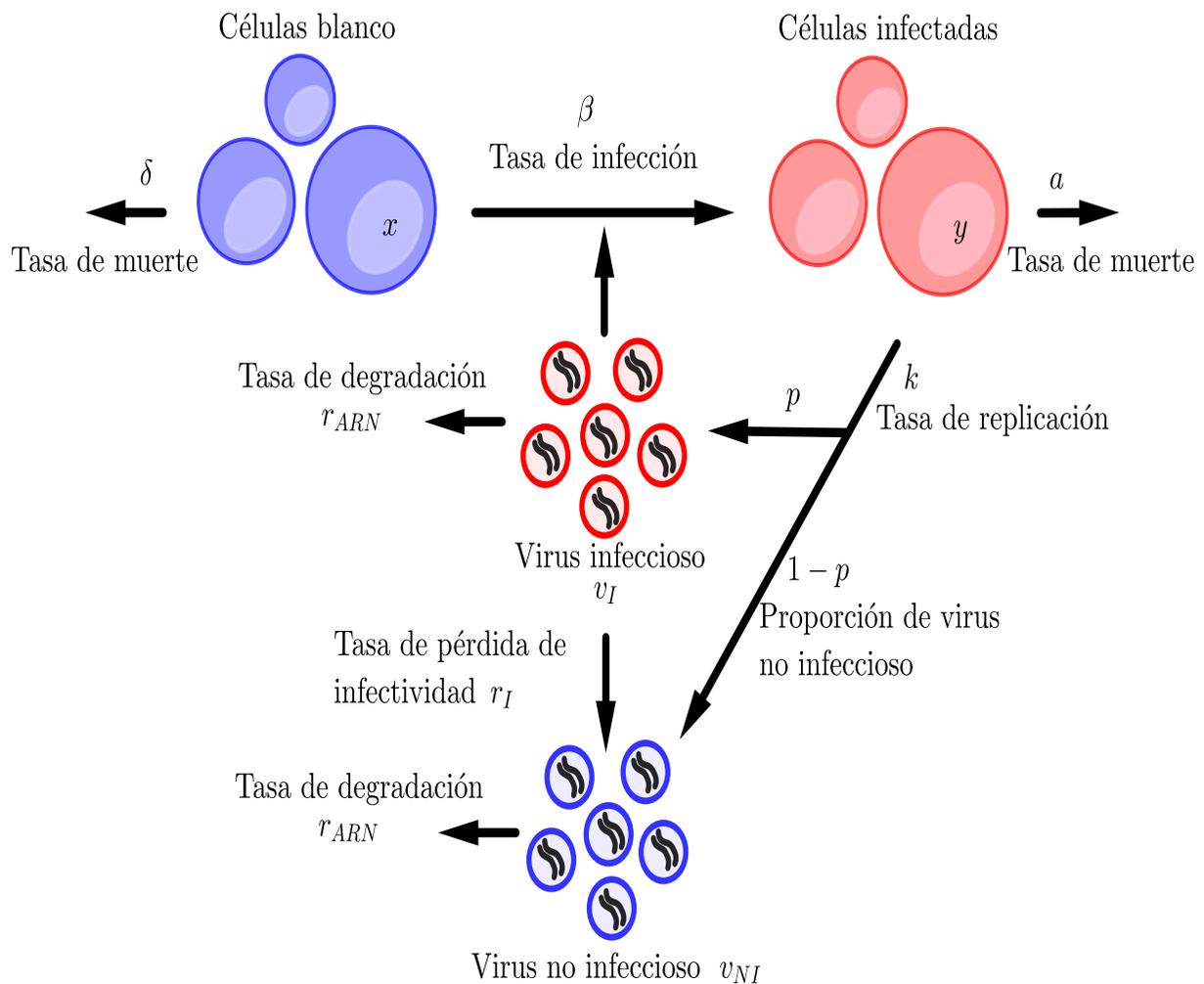


Figura 2.4: Representación esquemática para el modelo *in vitro* de la infección del VISH-KS661.

Para describir la cinética *in vitro* de la infección del VISH-KS661, [Iwami et al. \(2012\)](#) ampliaron el modelo básico (2.1) y propusieron el siguiente modelo extendido:

$$\begin{aligned}
 \frac{dx}{dt} &= -\beta xv_I - \delta x \\
 \frac{dy}{dt} &= \beta xv_I - ay \\
 \frac{dv_I}{dt} &= pk y - r_I v_I - r_{ARN} v_I \\
 \frac{dv_{NI}}{dt} &= (1-p)ky + r_I v_I - r_{ARN} v_{NI}
 \end{aligned} \tag{2.2}$$

donde x y y son el número de células blanco e infectadas, v_I y v_{NI} son el número de copias de ARN de virus infeccioso y no infeccioso, respectivamente. Los parámetros δ , a , r_{ARN} y β representan la tasa de mortalidad de las células blanco, la tasa de mortalidad de las células infectadas, la tasa de degradación del ARN viral y la tasa de transmisión infecciosa de las células blanco por virus, respectivamente. El parámetro k es la tasa de replicación de partículas virales liberadas por las células infectadas.

El número reproductivo básico de la infección propuesto por [Iwami et al. \(2012\)](#) para el modelo (2.2) es:

$$R_0^* = \frac{\beta k x_0}{a(r_I + r_{ARN})}$$

donde, x_0 es el número inicial de células blanco.

2.3. Índice de la Multiplicidad de Infección (MOI)

El índice de la multiplicidad de infección (MOI, por sus siglas en inglés), es un parámetro clave en la evolución de los virus, ya que determina procesos tales como el intercambio genético entre genomas o la intensidad de selección sobre los genes virales. Este índice es la relación entre el número de viriones y el número de células blanco presentes en un espacio definido; es de gran utilidad cuando en un experimento se requiere que cada célula blanco en el cultivo esté infectada; un MOI pequeño se utiliza cuando se requieren múltiples ciclos de infección.

El MOI es el cociente del número total de viriones entre el número total de células blanco del cultivo:

$$MOI = \frac{\text{Número de viriones}}{\text{Número de células blanco}}$$

Es decir, el número de viriones que se agregan por célula blanco durante la infección. Por ejemplo, si se agregan cien mil viriones a un millón de células blanco el MOI es de 0.1 y si se agregan un millón de viriones el MOI es de 1, esto no significa que cada célula en el cultivo recibe un virión. El número de viriones que recibe cada célula en diferentes MOIs, está dado por distribución Poisson ([Medina et al., 2016](#)).

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Donde $P(k)$ es la fracción de células infectadas por k partículas víricas, y λ es el MOI. La ecuación puede ser simplificada para calcular la probabilidad de células no infectadas ($k = 0$), de células con una única infección ($k = 1$) y de células con infección múltiple ($k > 1$):

$$P(k = 0) = e^{-\lambda}$$

$$P(k = 1) = \lambda e^{-\lambda}$$

$$P(k > 1) = 1 - [e^{-\lambda}(\lambda + 1)]$$

Ejemplo: Si tenemos 6.46×10^6 células/ml y las infectamos con un MOI de 2×10^{-3} , ¿cuántas células reciben 0, 1 y más de un virión? La fracción de células no infectadas, aquellas que reciben 0 partículas, es

$$P(0) = e^{-2 \times 10^{-3}} = 0.99,$$

es decir, en un cultivo de 6.46×10^6 células/ml se tienen 6,447,093 células no infectadas.

Con el mismo MOI de 2×10^{-3} , el número de células que reciben 1 partícula vírica se calcula mediante

$$P(1) = 2 \times 10^{-3}(e^{-2 \times 10^{-3}}) = 2 \times 10^{-3}(0.99) = 1.99 \times 10^{-3},$$

es decir, en un cultivo de 6.46×10^6 células/ml se tiene que 12,894.19 células reciben 1 partícula viral.

El número de células que reciben más de una partícula se calcula mediante

$$P(> 1) = 1 - [e^{-2 \times 10^{-3}}(2 \times 10^{-3} + 1)] = 1.99 \times 10^{-6},$$

esto es, en un cultivo de 6.46×10^6 células/ml, aproximadamente 13 células reciben más de una partícula viral.

2.4. Estimación Bayesiana

Uno de los teoremas más importantes de la teoría Bayesiana es el Teorema de Bayes, pues constituye la base de la estadística Bayesiana. El enfoque Bayesiano de los problemas inversos es de gran interés para la cuantificación de la incertidumbre en presencia de datos. [Capistrán et al. \(2016\)](#) desarrollaron un análisis Bayesiano de los parámetros de las ecuaciones diferenciales ordinales, donde consideran la solución aproximada mediante algún método numérico y el error derivado del método utilizado. Por consiguiente, se utiliza la distribución *a posteriori* aproximada en lugar de la *a posteriori* exacta.

2.4.1. Modelo estadístico

Se asume que los datos observados $\mathbf{y} = (y_1, \dots, y_n)$ en los tiempos discretos $t_1, \dots, t_n \in [0, T]^n$,

$$y_i = f(X_\theta(t_i)) + \varepsilon_i, \quad \varepsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2), \quad (2.3)$$

donde X_θ es la solución del sistema de ecuaciones diferenciales 2.4

$$\frac{dX_\theta}{dt} = F(X_\theta, t, \theta), \quad X_\theta(t_0) = X_0, \quad (2.4)$$

donde $\theta \in A \subset \mathbb{R}^d$ es un vector de parámetros desconocidos y $F : \mathbb{R}^p \times [0, T] \times A \rightarrow \mathbb{R}^p$ es una función conocida. Tomando $\sigma^2 \in S \subset \mathbb{R}^+$ se supone que la función F en el lado derecho del problema de valor inicial (2.4) sigue las condiciones de regularidad del teorema de Picard ([Süli y Mayers, 2003](#)). También se asume por el espacio paramétrico, que A y S son conjuntos compactos.

En estadística Bayesiana, para obtener la distribución *a posteriori* se utiliza la distribución *a priori* y la función de verosimilitud. Cuando los errores de la Ecuación (2.3) tiene distribución normal, la función de verosimilitud de los datos es:

$$P_{\mathbf{Y}|\Phi}(\mathbf{y}|\theta, \sigma) = \sigma^{-n} (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(X_\theta(t_i)))^2 \right\}, \quad (2.5)$$

donde $\Phi = (\Theta, \Sigma)$ es un vector aleatorio con realización particular $\phi = (\theta, \sigma)$. Esta expresión involucra la obtención de X_θ mediante una solución numérica de (2.4).

Como consecuencia, la inferencia se realiza, no con la solución exacta sino a través del siguiente modelo aproximado,

$$y_i = f(X_\theta^h(t_i)) + \varepsilon_i, \quad \varepsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2), \quad (2.6)$$

donde X_θ^h es la solución aproximada de (2.4) proporcionada por la solución numérica, h es el parámetro de precisión de la solución numérica. La nueva verosimilitud derivada del modelo (2.6) se expresa de la siguiente manera:

$$P_{\mathbf{Y}|\Phi}^h(\mathbf{y}|\theta, \sigma) = \sigma^{-n} (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(X_\theta^h(t_i)))^2 \right\}. \quad (2.7)$$

La distribución *a posteriori* es:

$$P_{\Phi|\mathbf{Y}}^h(\theta, \sigma|\mathbf{y}) = \frac{P_{\mathbf{Y}|\Phi}^h(\mathbf{y}|\theta, \sigma) P_\Phi(\theta, \sigma)}{P_{\mathbf{Y}}^h(\mathbf{y})},$$

donde $P_{\mathbf{Y}}^h(\mathbf{y}) = \int P_{\mathbf{Y}|\Phi}^h(\mathbf{y}|\theta, \sigma) P_\Phi(\theta, \sigma) d\theta d\sigma$ es la constante de normalización, también llamada verosimilitud marginal de los datos \mathbf{y} .

2.4.2. Estimación de parámetros

Una vez obtenida la distribución *a posteriori* del vector de parámetros, dada la solución por aproximación MCMC. Se necesita un estimador puntual que resuma toda la información de la distribución *a posteriori*. Este se puede obtener al considerar una función de pérdida $\mathcal{P}(\mathcal{T}, \theta)$, la cual es una función no negativa que indica la pérdida que se tiene al considerar un estimador \mathcal{T} en lugar del verdadero estimador θ .

Las funciones de pérdida más utilizada son:

- Pérdida cuadrática: $\mathcal{P}(\mathcal{T}, \theta) = (\mathcal{T} - \theta)^2$.
- Pérdida absoluta: $\mathcal{P}(\mathcal{T}, \theta) = |(\mathcal{T} - \theta)|$.
- Pérdida todo o nada:
$$\begin{cases} \mathcal{P}(\mathcal{T}, \theta) = 0, & \text{si } (\mathcal{T} - \theta) \leq \epsilon, \\ \mathcal{P}(\mathcal{T}, \theta) = 1, & \text{si } (\mathcal{T} - \theta) > \epsilon. \end{cases}$$

Dichas funciones de pérdida conducen a distintos estimadores Bayesianos, los cuales tienen la finalidad de minimizar la pérdida esperada $E_\theta[\mathcal{P}(\mathcal{T}, \theta)]$, esto es,

$$\begin{aligned} \int_{\Theta} E_\theta[\mathcal{P}(\mathcal{T}, \theta)]\pi(\theta)d\theta &= \int_{\Theta} \left[\int_Y \mathcal{P}(\mathcal{T}, \theta)\pi(\mathbf{y}|\theta) \prod_{i=1}^n dy_i \right] \pi(\theta)d\theta \\ &= \int_Y \left[\int_{\Theta} \mathcal{P}(\mathcal{T}, \theta)\pi(\theta|\mathbf{y})d\theta \right] \pi(\mathbf{y}) \prod_{i=1}^n dy_i. \end{aligned}$$

Minimizar la doble integral es equivalente a minimizar la pérdida esperada *a posteriori* $E_\theta[\mathcal{P}(\mathcal{T}, \theta)|\mathbf{y}]$ (Mood et al., 1974). Entonces, se define el estimador Bayes de θ como el estimador que minimiza la pérdida esperada *a posteriori*.

Teorema 2.1. *Si la función de pérdida es cuadrática, entonces el estimador Bayes es la esperanza a posteriori:*

$$\theta^*(x) = E(\theta|\mathbf{y}).$$

Demostración. Utilizando la función de pérdida cuadrática, observamos que el estimador Bayesiano es la media *a posteriori* (ver Gómez y Delicado, 2016):

$$\begin{aligned} E_\theta[\mathcal{P}(\mathcal{T}, \theta)|\mathbf{y}] &= E_\theta[(\mathcal{T} - \theta)^2|\mathbf{y}] \\ &= E_\theta[(\mathcal{T}^2 - 2\mathcal{T}\theta + \theta)^2|\mathbf{y}] \\ &= E_\theta[\mathcal{T}^2|\mathbf{y}] - 2E_\theta[\mathcal{T}\theta|\mathbf{y}] + E_\theta[\theta^2|\mathbf{y}] \\ &= \mathcal{T}^2 - 2\mathcal{T}E_\theta[\theta|\mathbf{y}] + E_\theta[\theta^2|\mathbf{y}]. \end{aligned}$$

Derivando respecto a \mathcal{T} e igualando a cero:

$$2\mathcal{T} - 2E_\theta[\theta|\mathbf{y}] + 0 = 0 \implies \mathcal{T} = E_\theta[\theta|\mathbf{y}],$$

es un mínimo, dado que la segunda derivada respecto a \mathcal{T} es $2 > 0$. ■

2.4.3. Diagnósticos de convergencia

Para verificar la convergencia de las cadenas se utilizan algunos diagnósticos, ya sean gráficos o pruebas estadísticas que den indicios de que las cadenas de Markov convergen. Entre las pruebas estadísticas, están el diagnóstico de Gelman–Rubin y el de Geweke.

2.4.3.1. Diagnóstico de Gelman–Rubin

Gelman y Rubin proponen una prueba de convergencia basada en 2 o más cadenas paralelas, cada una partiendo de diferentes valores iniciales con respecto a la verdadera distribución a *posteriori*. Su método se basa en una comparación entre las varianzas dentro y entre cadenas para cada variable (Gelman et al., 2013).

Para obtenerlo se generan $m \geq 2$ cadenas, cada una con $2n$ iteraciones y con puntos iniciales dispersos. Se eliminan las n primeras iteraciones y se trabaja con las iteraciones restantes.

Sea ψ un estadístico que estima algún parámetro de la distribución objetivo. De las m cadenas generadas con las restantes n iteraciones se calculan las estimaciones de la varianza entre y dentro de las cadenas para estimar la varianza de ψ .

- Varianza entre las cadenas (varianza de los promedios dentro de las cadenas):

$$B = \frac{n}{(m-1)} \sum_{j=1}^m (\bar{\psi}_{\cdot j} - \bar{\psi}_{\cdot\cdot})^2,$$

donde

$$\bar{\psi}_{\cdot j} = \frac{1}{n} \sum_{i=1}^n \psi_{ij}, \quad \bar{\psi}_{\cdot\cdot} = \frac{1}{m} \sum_{j=1}^m \bar{\psi}_{\cdot j}.$$

- Varianza dentro de la cadenas (promedio de las varianzas dentro de las cadenas):

$$W = \frac{1}{m} \sum_{j=1}^m s_j^2, \quad s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (\psi_{ij} - \bar{\psi}_{\cdot j})^2.$$

Las estimaciones de la varianza entre y dentro de las cadenas se combinan para estimar un límite superior para la varianza de ψ :

$$\hat{V}(\psi) = \frac{n-1}{n} W + \frac{1}{n} B.$$

Finalmente, se estima el Estadístico de Gelman-Rubin:

$$\sqrt{\hat{R}} = \sqrt{\frac{\hat{V}(\psi)}{W}},$$

que puede interpretarse como la medición del factor por el cual la desviación estándar de ψ podría reducirse al extender la cadena. El factor $\sqrt{\hat{R}}$ tiende a 1 cuando la cadena tiende a infinito, entonces $\sqrt{\hat{R}}$ debería estar cerca de 1 si las cadenas convergen aproximadamente a la distribución objetivo. El valor de \hat{R} debería ser menor que 1.1 o 1.2 (Rizzo, 2007).

2.4.3.2. Diagnóstico de Geweke

Geweke (1992) sugiere un criterio para evaluar la convergencia de las cadenas en base a la comparación de medias, en intervalos diferentes, después de una fase de calentamiento del algoritmo. Si la convergencia fue alcanzada, los comportamientos en esos intervalos deben ser semejantes.

Considere $\psi = t(\theta)$ una función real de un parámetro de interés θ , para el cual fue obtenida una muestra a partir del algoritmo MCMC. Sea m el número de iteraciones correspondiente a la fase de calentamiento y $m + n$ el total de iteraciones. Se construyen las medias $\bar{\psi}_a$ y $\bar{\psi}_b$ en base a los grupos de iteraciones $n_a < n$ y $n_b < n$.

Luego,

$$z_G = \frac{\bar{\psi}_a - \bar{\psi}_b}{\sqrt{Var(\bar{\psi}_a) + Var(\bar{\psi}_b)}} \xrightarrow{D} N(0, 1).$$

De esta forma, valores grandes de z_G indican falta de convergencia. Sin embargo, valores pequeños no significan que existe convergencia. Luego, la decisión debe ser tomada conjuntamente con otros criterios y con la inspección visual. Para implementar el criterio de Geweke, basta con efectuar sólo una cadena larga del algoritmo MCMC.

Metodología

3.1. Experimento

Para lograr los objetivos de este estudio se tomaron los datos obtenidos por [Iwami et al. \(2012\)](#), que provienen de experimentos, cada uno realizado en dos pocillos con una concentración inicial de células de 6.46×10^6 células/ml en cada pocillo (ver Figura 3.1). Los cultivos de células se inocularon en dos diferentes MOIs, 2.0×10^{-3} y 2.0×10^{-6} , de VISH-KS661 y se encubaron durante cuatro horas.

Después de la inoculación, las células se lavaron tres veces para eliminar el medio de infección. El sobrenadante (fracción líquida soluble de una muestra después de la centrifugación o precipitación de sólidos insolubles) del cultivo se recogió durante nueve días junto con una pequeña fracción de células (5.5%) para el recuento de células blanco y células infectadas. Las células restantes se lavaron suavemente tres veces y se pusieron en un medio fresco, libre de virus (infección). Cabe destacar que se realizaron nueve mediciones para cada MOI con la misma concentración de células inicial, las cuales quedan representadas en la Figura 3.2.

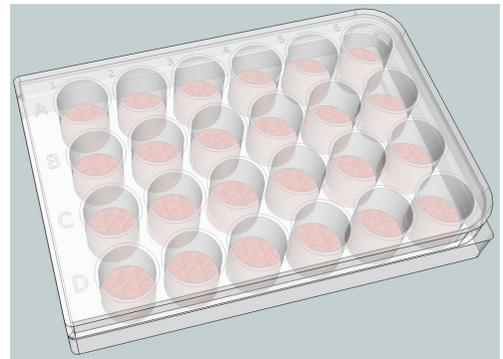


Figura 3.1: Pocillos.

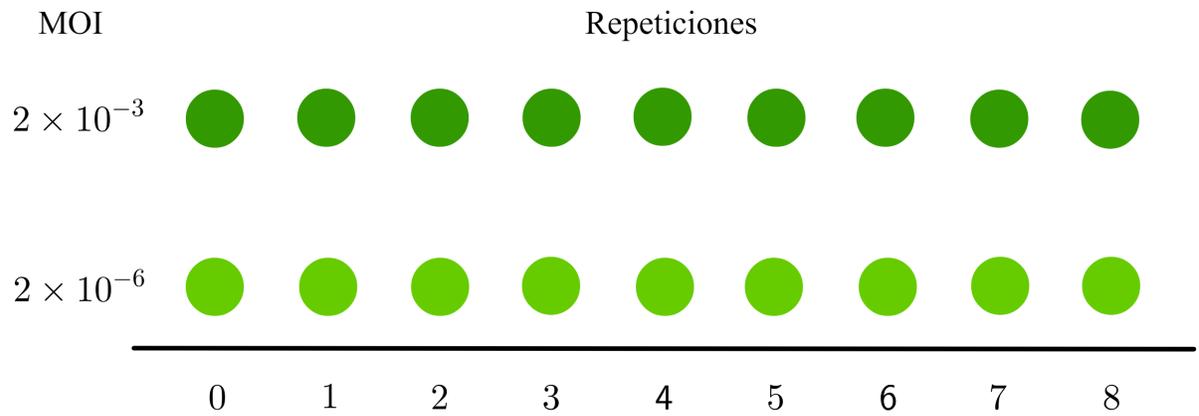


Figura 3.2: Representación esquemática de las nueve mediciones para cada MOI.

En [Iwami et al. \(2012\)](#) se pueden encontrar las nueve mediciones para las concentraciones de células blanco (células Nef-negativas HSC-F), células infectadas (células Nef-positivas HSC-F), total de la carga viral VISH-KS661 y carga viral infecciosa de VISH-KS661, para cada MOI, éstas mediciones se muestran en el Cuadro 3.1.

3.2. Datos

Cuadro 3.1: Datos del experimento *in vitro*.

MOI	Día de medición								
	0	1	2	3	4	5	6	7	8
Concentración de células Nef-negativas HSC-F (células/ml)									
2×10^{-3}	5470829	6044623	2690861	1012828	223584	42130	58470	10386	10270
2×10^{-6}	3357117	2583058	4557411	35074989	1334060	1896048	1022157	307908	153360
Concentración de células Nef-positivas HSC-F (células/ml)									
2×10^{-3}	d.l.	d.l.	439139	1167172	736416	177870	41530	19614	9730
2×10^{-6}	d.l.	d.l.	d.l.	d.l.	65940	383952	347843	232092	86640
Total de la carga viral VISH-KS661 (copias de ARN/ml)									
2×10^{-3}	9180000	331000000	2840000000	4050000000	3140000000	1120000000	1540000000	202000000	56500000
2×10^{-6}	10170	800536	4600000	54200000	322000000	1300000000	1210000000	603000000	2750000000
Carga viral infecciosa de VISH-KS661 (TCID ₅₀)									
2×10^{-3}	40	4064	40960	81920	163840	20480	2560	160	d.l.
2×10^{-6}	40	40	80	640	5120	1280	1280	640	1280

d.l.: designa muestras en las que la concentración estaba por debajo del límite de detección.

3.3. Modelo *ad-hoc* de infección del VISH - KS661 *in vitro*

El modelo extendido (2.2) describe las tasas de cambio para las células blanco (x), células infectadas (y), virus infecciosos (v_I) y virus no infecciosos (v_{NI}); sin embargo, en la práctica resulta difícil disponer del conteo de virus no infeccioso por lo que es común utilizar el total de la carga viral, razón por la cual se realiza un cambio de variable $v_{ARN} = v_I + v_{NI}$ para trabajar con el siguiente modelo *ad-hoc* de infección del VISH - KS661 *in vitro* (Iwami et al., 2012), el cual incluye la carga viral total (v_{ARN})

$$\begin{aligned}
 \frac{dx}{dt} &= -\beta_{50}xv_{50} - \delta x - hx \\
 \frac{dy}{dt} &= \beta_{50}xv_{50} - ay - hy \\
 \frac{dv_{ARN}}{dt} &= ky - r_{ARN}v_{ARN} - r_c v_{ARN} \\
 \frac{dv_{50}}{dt} &= k_{50}y - r_I v_{50} - r_{ARN}v_{50} - r_c v_{50}
 \end{aligned} \tag{3.1}$$

donde, v_I del modelo (1.2) se cambia por v_{50} ; h es la tasa de cosecha de las células blanco e infectadas, r_c es la tasa de cosecha del total de partículas virales y partículas virales infecciosas, k y k_{50} son las tasas de producción de virus total y virus infeccioso, respectivamente, la tasa de transmisión infecciosa, β se renombra como β_{50} . Los parámetros restantes significan lo mismo que en el modelo extendido (2.2). El modelo *ad-hoc* (3.1) es un sistema de ecuaciones diferenciales ordinarias que no tiene solución exacta.

El número reproductivo básico de la infección propuesto por Iwami et al. (2012) para el modelo *ad-hoc*, está dado por

$$R_0^{ad-hoc} = \frac{\beta_{50}k_{50}x_0}{(a+h)(r_I+r_{ARN}+r_c)}.$$

Estos autores realizan estudios empíricos sobre los parámetros de dicho modelo. De los 9 parámetros a estimar ($\beta_{50}, \delta, h, a, k, k_{50}, r_{ARN}, r_c, r_I$), estos autores muestran que fijando los parámetros δ, h, r_{ARN}, r_c y r_I , y estimando β_{50}, a, k y k_{50} mediante mínimos cuadrados se obtienen resultados coherentes y con interpretación biológica.

Para obtener h (cosecha celular) y r_c (cosecha de virus) se aproxima el muestreo de células y virus como un decaimiento exponencial continuo. δ, r_I y r_{ARN} se determinaron mediante mediciones en experimentos separados. Se fijaron las tasas de degradación y pérdida de infectividad de los virus (r_{ARN} y r_I), así como la tasa de mortalidad (δ).

Los parámetros β_{50} (tasa de transmisión infecciosa), a (tasa de muerte de las células infectadas), k (total de la carga viral) y k_{50} (carga viral infecciosa) se estimaron mediante regresión no lineal de mínimos cuadrados con su respectivo intervalo de confianza del 95 %, ajustando el modelo de ecuaciones diferenciales al conjunto de datos *in vitro*.

Este trabajo de tesis se enfoca en estudiar los parámetros β_{50}, a, k y k_{50} y el resto se consideraron fijos con los valores tomados de [Iwami et al. \(2012\)](#). Además se aplica logaritmo a las series de datos reales. Para el $MOI = 2 \times 10^{-3}$ no se tomaron en cuenta los días 1 y 8; y para el $MOI = 2 \times 10^{-6}$ los días 1, 2 y 3, ya que en esos días la concentración de las muestras estaban por debajo del límite de detección para células infecciosas HSC-F o la carga viral infecciosa de VISH-KS661.

3.4. Inferencia Bayesiana

La inferencia Bayesiana puede utilizarse para estimar parámetros, seleccionar modelo y hacer predicción.

Dado el modelo estadístico

$$\mathbf{y}(t_i) = \mathbf{X}_{\boldsymbol{\theta}}^h(t_i) + \boldsymbol{\varepsilon}(t_i), \quad i = 1, \dots, 9, \quad (3.2)$$

donde $\mathbf{y}(t_i)$ representa los datos del experimento *in vitro* en el tiempo i por infección del VISH-KS661, $\boldsymbol{\theta} = (\beta_{50}, a, k, k_{50})$ es el vector de parámetros de interés, $\mathbf{X}_{\boldsymbol{\theta}}^h(t_i)$ la solución numérica del sistema de ecuaciones (3.1) (es la función que satisface el sistema de ecuaciones diferenciales) y $\boldsymbol{\varepsilon}(t_i)$ es el error aleatorio en la i -ésima observación, los cuales se distribuyen normal, con vector de media cero y matriz de varianzas y covarianzas $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$, de manera que

$$\mathbf{y}(t_i) \sim \mathcal{N}_4[\mathbf{X}_{\boldsymbol{\theta}}^h(t_i), \boldsymbol{\Sigma}]. \quad (3.3)$$

Se sabe que la inferencia Bayesiana sobre los parámetros se basa en la función de verosimilitud $l[\cdot|\cdot]$, que es la función que proporciona toda la información de los datos (Capistrán et al., 2016). La verosimilitud de la distribución (3.3), queda planteada de la siguiente forma:

$$l[\mathbf{y}(t_i)|\boldsymbol{\theta}, \boldsymbol{\Sigma}] = \prod_{i=1}^9 \frac{1}{(2\pi)^{q/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2} [\mathbf{y}(t_i) - \mathbf{X}_{\boldsymbol{\theta}}^h(t_i)]^T \boldsymbol{\Sigma}^{-1} [\mathbf{y}(t_i) - \mathbf{X}_{\boldsymbol{\theta}}^h(t_i)] \right\}.$$

Por teorema de Bayes la distribución *a posteriori* conjunta está dada por la siguiente expresión

$$P[\boldsymbol{\theta}, \boldsymbol{\Sigma}|\mathbf{y}(t_i)] = \frac{l[\mathbf{y}(t_i)|\boldsymbol{\theta}, \boldsymbol{\Sigma}]P(\boldsymbol{\theta}, \boldsymbol{\Sigma})}{\int l[\mathbf{y}(t_i)|\boldsymbol{\theta}, \boldsymbol{\Sigma}]P(\boldsymbol{\theta}, \boldsymbol{\Sigma})\partial\boldsymbol{\theta}\partial\boldsymbol{\Sigma}} \quad (3.4)$$

$$\propto l[\mathbf{y}(t_i)|\boldsymbol{\theta}, \boldsymbol{\Sigma}]P(\boldsymbol{\theta})P(\boldsymbol{\Sigma}), \quad (3.5)$$

donde se supone independencia entre $\boldsymbol{\theta}$ y $\boldsymbol{\Sigma}$.

Así, la distribución *a posteriori* marginal de $\boldsymbol{\theta}$ es

$$P[\boldsymbol{\theta}|\mathbf{y}(t_i)] = \int P[\boldsymbol{\theta}, \boldsymbol{\Sigma}|\mathbf{y}(t_i)]\partial\boldsymbol{\Sigma}. \quad (3.6)$$

No existe información sobre las distribuciones *a priori* de los parámetros β_{50} , a , k , k_{50} pero sí se conoce que son valores mayores que cero, esto se debe a que son tasas de muerte, de producción y de infección. Considerando independencia entre los parámetros que conforman el vector $\boldsymbol{\theta}$, la distribución *a priori* conjunta está dada por:

$$P(\boldsymbol{\theta}) = P(\beta_{50})P(a)P(k)P(k_{50}), \quad (3.7)$$

donde $\boldsymbol{\theta} = (\beta_{50}, a, k, k_{50})$.

Se utilizan las siguientes distribuciones *a priori*, donde el uso de la distribución *log-normal* se sugiere en Agbaje et al. (2003), la *gamma* en Capistrán et al. (2016) y para la constante de la matriz de varianza y covarianza Putter et al. (2002) recomienda usar una *gamma*.

- Tasa de transmisión infecciosa: β_{50} , se distribuye *log-normal*(μ_0, τ_0^2), es decir,

$$P(\beta_{50}) = \frac{1}{\sqrt{2\pi\tau_0}\beta_{50}} e^{-\frac{(\log \beta_{50} - \mu_0)^2}{2\tau_0^2}},$$

donde $0 \leq \beta_{50} < \infty$, $-\infty < \mu_0 < \infty$ y $\tau_0 > 0$.

- Tasa de muerte de las células infectadas: a , se distribuye $gamma(\alpha_0, \beta_0)$, es decir,

$$P(a) = \frac{1}{\Gamma(\alpha_0)\beta_0^{\alpha_0}} a^{\alpha_0-1} e^{-a/\beta_0},$$

donde $\Gamma(\cdot)$ es la función Gamma¹, $0 \leq a < \infty$ y $\alpha_0, \beta_0 > 0$.

- Tasa de producción total del virus: k , se distribuye $log - normal(\mu_1, \tau_1^2)$, es decir,

$$P(k) = \frac{1}{\sqrt{2\pi\tau_1}k} e^{-\frac{(\log k - \mu_1)^2}{2\tau_1^2}},$$

donde $0 \leq k < \infty$, $-\infty < \mu_1 < \infty$ y $\tau_1 > 0$.

- Tasa de producción de virus infeccioso: k_{50} , se distribuye $gamma(\alpha_1, \beta_1)$, es decir,

$$P(k_{50}) = \frac{1}{\Gamma(\alpha_1)\beta_1^{\alpha_1}} k_{50}^{\alpha_1-1} e^{-k_{50}/\beta_1},$$

donde $0 \leq k_{50} < \infty$ y $\alpha_1, \beta_1 > 0$.

- Constante de la matriz de varianzas y covarianzas: σ^2 , se distribuye $gamma(\alpha_2, \beta_2)$, es decir,

$$P(\sigma^2) = \frac{1}{\Gamma(\alpha_2)\beta_2^{\alpha_2}} \sigma^{2\alpha_2-1} e^{-\sigma^2/\beta_2},$$

donde $0 \leq \sigma^2 < \infty$ y $\alpha_2, \beta_2 > 0$.

La distribución *a posteriori* conjunta de los parámetros se obtiene mediante la aproximación de MCMC. A partir de ésta se obtiene una aproximación a la distribución *a posteriori* marginal para cada uno de los 4 parámetros. Posteriormente, se obtiene el estimador Bayes para los parámetros de interés utilizando la función de pérdida cuadrática, esto es, la media de la distribución marginal *a posteriori*. En este trabajo se utiliza la función `de_mcmc` del paquete `deBInfer` y la función `ode` del paquete `deSolve` para obtener los estimadores Bayes del vector de parámetros.

¹Función Gamma: $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$

Resultados y Discusión

Los resultados de este capítulo se basan en el modelo *ad-hoc* (transformación logarítmica) y en el logaritmo de los datos del Cuadro 3.1.

4.1. Resultados

4.1.1. Estimación mediante el método de mínimos cuadrados

En este apartado se presentan los resultados de la estimación de los parámetros usando mínimos cuadrados, el cual es el método más común para la estimación de parámetros de un sistema de ecuaciones diferenciales ordinarias no lineales.

Al modelo *ad-hoc* (3.1) se le imponen las condiciones iniciales de la Cuadro 4.1

Cuadro 4.1: Condiciones iniciales para las variables del modelo *ad-hoc*.

Variable con logaritmo	MOI 2×10^{-3}	MOI 2×10^{-6}
$x(0)$	$\log(5470829)$	$\log(3357117)$
$y(0)$	$\log(10^{-10})$	$\log(10^{-10})$
$v_{ARN}(0)$	$\log(9180000)$	$\log(10170)$
$v_{50}(0)$	$\log(40)$	$\log(40)$

donde, x, y, v_{ARN} y v_{50} son las células blanco, células infectadas, carga viral total del VISH-KS661 y carga viral infecciosa del VISH-KS661 en $t = 0$, respectivamente.

Los valores iniciales para los parámetros de interés se toman de [Iwami et al. \(2012\)](#).

Se utilizó la función `ode` del paquete `deSolve` ([Soetaert et al., 2010](#)) del software `R` 3.4.0 ([R Core Team, 2017](#)) para resolver numéricamente, mediante la función `lsoda`, el modelo *ad-hoc* y obtener \mathbf{X}_{θ}^h . Para obtener la estimación de mínimos cuadrados se utiliza la función `nmkb` de la función `dfoptim` usando `R`.

En el Cuadro 4.2 se muestran las estimaciones puntuales obtenidas mediante el método de mínimos cuadrados para los cuatro parámetros de interés y la suma de los cuadrados de los residuos, tomando en cuenta las cuatro series para cada MOI. Bajo estos resultados, en las Figuras 4.1 y 4.2 se muestran los gráficos relativos al ajuste entre los datos reales y la solución numérica considerando los MOIs 2×10^{-3} y 2×10^{-6} .

Los números reproductivos básicos para los dos MOIs, útiles para diagnosticar la infección, a través de este modelo *ad-hoc* fueron 6.93 y 6.22, respectivamente; es decir, el número de células recién infectadas generadas intrínsecamente por una sola célula infecciosa al inicio de la infección es de aproximadamente 7 y 6 células para los MOIs 2×10^{-3} y 2×10^{-6} , respectivamente.

Cuadro 4.2: Estimaciones obtenidas mediante el método de mínimos cuadrados.

Parámetro		Estimadores de mínimos cuadrados	
		MOI 2×10^{-3}	MOI 2×10^{-6}
Tasa de transmisión infecciosa	β_{50}	1.87×10^{-5}	1.05×10^{-4}
Tasa de muerte de células infectadas	a	1.37	0.49
Tasa de producción total del virus	k	3.94×10^4	3.01×10^4
Tasa de producción de virus infeccioso	k_{50}	0.8	0.08

La Figura 4.1 compara los datos experimentales (en escala logarítmica) para la concentración de células blanco, células infectadas, total de la carga viral y carga viral infecciosa. La solución numérica usando las estimaciones por mínimos cuadrados es representada por la línea roja y los datos reales por los puntos azules. Se observa que el comportamiento de las células infectadas y de la carga viral total expresado por los datos reales y la solución numérica es mejor aproximado que el comportamiento de las células blanco y carga viral infecciosa. Además, se observa que

en el día 4 la carga viral total e infecciosa alcanzan su máxima producción viral y luego decaen rápidamente, esto se debe a que llega un momento en que ya no hay suficientes células blanco para que los virus se sigan replicando.

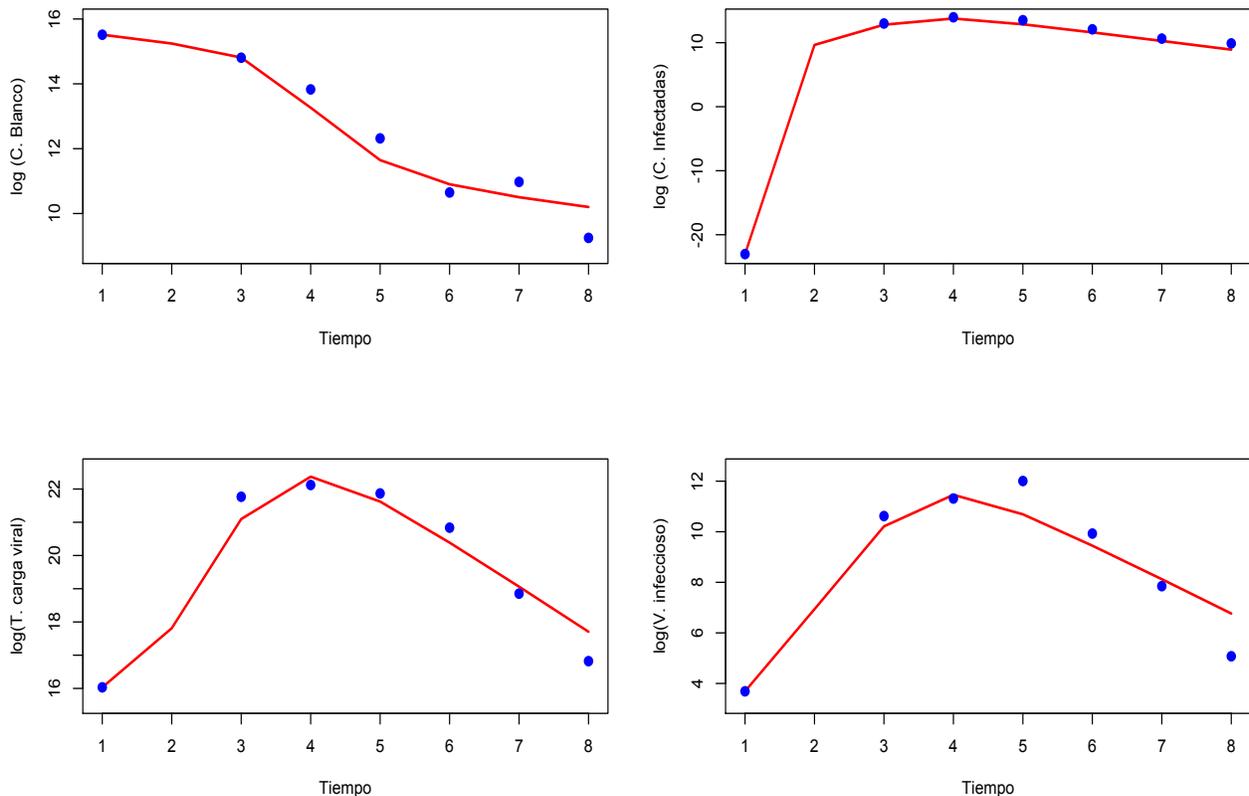


Figura 4.1: Datos reales y los resultados de la solución numérica usando mínimos cuadrados para el MOI 2×10^{-3} .

La Figura 4.2 compara los datos experimentales (escala logarítmica) para la concentración de células blanco, células infectadas, total de la carga viral y carga viral infecciosa. La solución numérica usando las estimaciones por mínimos cuadrados se representa por la línea roja y los datos reales por los puntos azules. Se observa, al igual que para el MOI 2×10^{-3} , el comportamiento de las células infectadas y del total de la carga viral expresado por los datos reales y la solución numérica es mejor aproximado que el comportamiento de las células blanco y carga viral infecciosa.

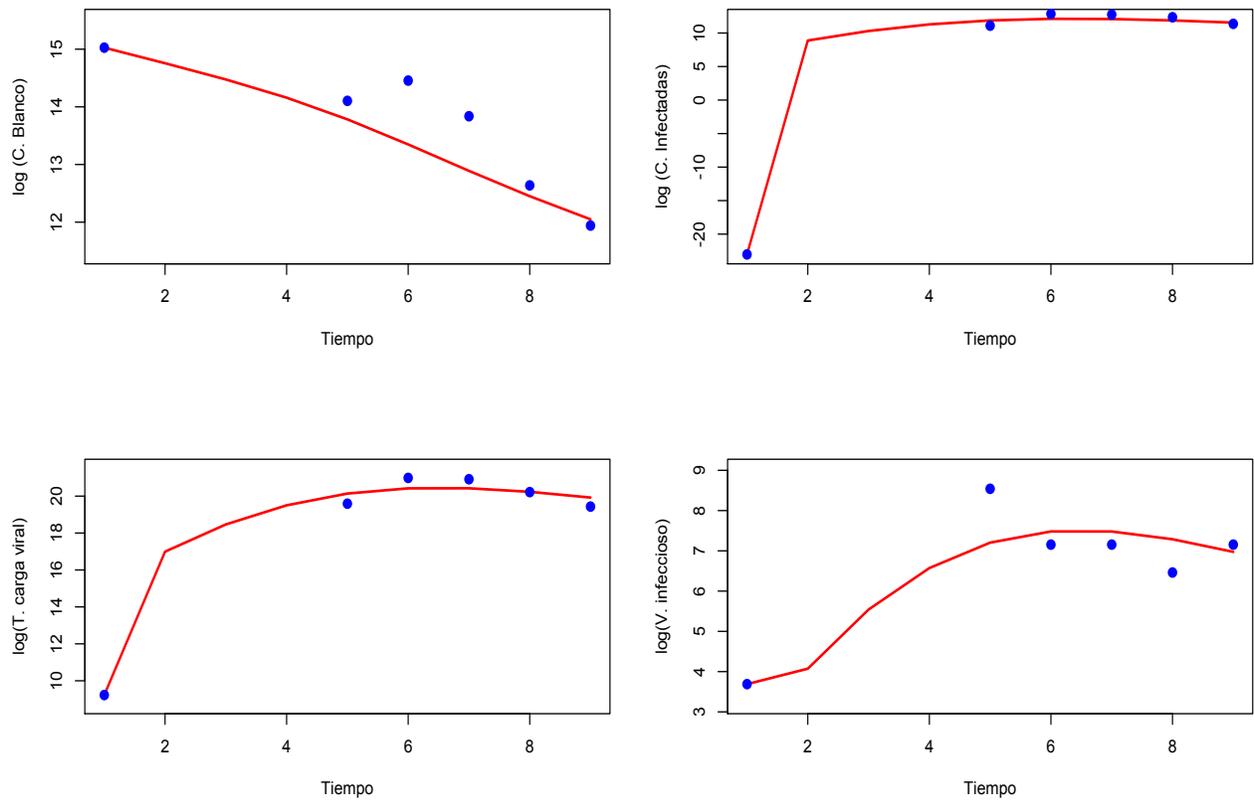


Figura 4.2: Comparación de los datos reales con los resultados de las solución numérica usando mínimos cuadrados para el MOI 2×10^{-6} .

4.1.2. Estimación Bayesiana

En esta sección se muestran los resultados, mediante el enfoque Bayesiano de la estimación puntual de los parámetros de interés, intervalos creíbles y el test de Gelman-Rubin tomando en cuenta los datos correspondientes a los MOI 2×10^{-3} y 2×10^{-6} .

Los resultados que se presentan enseguida se obtuvieron usando la paquetería `deBInfer` (Boersch-Supan et al., 2017) del software R, considerando la solución numérica \mathbf{X}_θ^h así como las condiciones y valores iniciales utilizados en la estimación por mínimos cuadrados en el Cuadro 4.2.

a) Para obtener la estimación Bayesiana del MOI 2×10^{-3} se consideran las distribuciones *a priori* siguientes

$$\beta_{50} \sim \log - normal(-11.25983, (0.1)^2),$$

$$a \sim gamma(1, 1),$$

$$k \sim \log - normal(10.5759, (0.01)^2),$$

$$k_{50} \sim gamma(1/2, 1/2),$$

$$\sigma^2 \sim gamma(1, 20).$$

Los resultados se presentan en el Cuadro 4.3.

Cuadro 4.3: Estimaciones Bayesianas usando los datos del MOI 2×10^{-3} .

Parámetro		Valor	Intervalo creíble del 95 %
Tasa de transmisión infecciosa	β_{50}	1.296×10^{-5}	$1.060 \times 10^{-5} - 1.571 \times 10^{-5}$
Tasa de muerte de células infectadas	a	1.018	$2.652 \times 10^{-2} - 3.844$
Tasa de producción total del virus	k	3.914×10^4	$3.834 \times 10^4 - 3.993 \times 10^4$
Tasa de producción de virus infeccioso	k_{50}	1.032	$2.170 \times 10^{-3} - 5.261$
Varianza	σ^2	0.04	$1.49 \times 10^{-3} - 0.18$

El cuadro A.1 del **Anexo A** muestra el resultado del test de Gelman–Rubin, utilizando tres MCMC, tiene un valor menor a 1.1 para cada uno de los parámetros, lo que implica que hay convergencia entre las Cadenas de Markov, y esto puede verse para cada uno de los parámetros de manera gráfica en las Figura A.1 y Figura A.2 del mismo anexo.

En la columna izquierda de la Figura A.2 (ver **Anexo A**) se muestran las trazas de tres Cadenas de Markov para los parámetros de interés, en las que se observan que las tres Cadenas tienen una buena mezcla y convergen, es decir, cada cadena recorre todo el espacio de parámetros. La columna derecha de la misma figura muestra las distribuciones *a posteriori* marginales para los cuatro parámetros de interés. Al comparar las trazas de las tres Cadenas de Markov para cada uno de los parámetros de interés con su respectiva distribución *a posteriori* existe congruencia entre la media de las trazas y distribución *a posteriori* marginal.

Tomando en cuenta los resultados de la estimación Bayesiana de los parámetros de interés, los valores y condiciones iniciales dados en el apartado de la estimación por mínimos cuadrados, se presentan las gráficas que muestran la solución numérica del modelo *ad-hoc* y los datos reales.

La Figura 4.3 muestra las gráficas de los comportamientos entre la solución numérica del modelo *ad-hoc* usando los resultados de la estimación Bayesiana, representada por la línea roja y los datos reales por los puntos azules (escala logarítmica), para la concentración de células blanco, células infectadas, carga viral total y carga viral infecciosa. Se observa que el comportamiento de las células blanco e infectadas expresado por los datos reales y la solución numérica es mejor aproximado que el comportamiento de la carga viral total e infecciosa. Además, en las carga viral total y la carga viral infecciosa, mostrada en la parte inferior de la misma figura, se observa que la solución numérica alcanza los puntos máximos de los datos reales, a diferencia de la solución mediante mínimos cuadrados; aunque en los días 7 y 8, los puntos de los datos reales están por debajo de la solución numérica usando las estimaciones de Bayes. En los días 4 y 5, la solución para las carga viral total e infecciosa alcanzan su máxima producción viral y luego decaen, respectivamente. Es decir, es el mismo comportamiento que se observa en la solución por mínimos cuadrados.

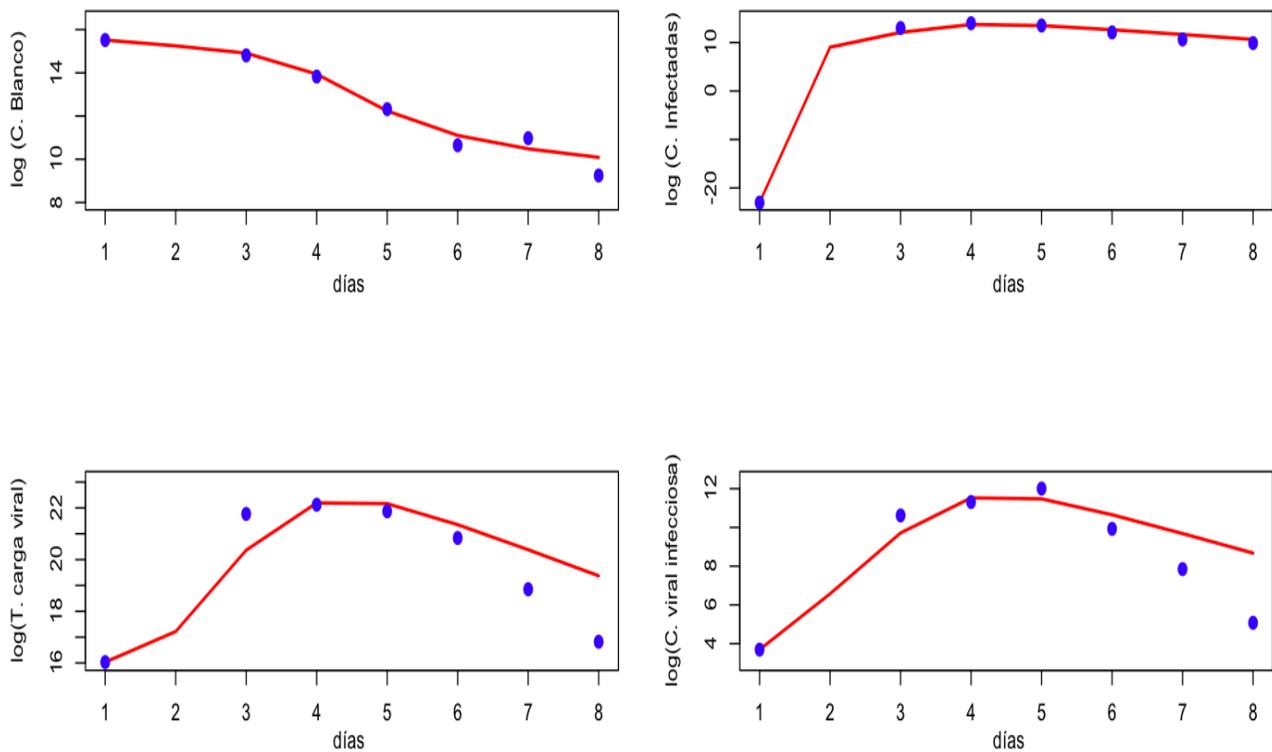


Figura 4.3: Solución numérica del modelo *ad-hoc* usando resultados de la estimación Bayesiana para los datos del MOI 2×10^{-3} .

Para caracterizar la cinética viral, es recomendable calcular el número reproductivo básico para el sistema que representa el modelo *ad-hoc* (R_0^{ad-hoc}). Utilizando los resultados de las estimaciones Bayesianas, el valor de dicho número reproductivo es $R_0^{ad-hoc} = 8.22$, es decir, el número de células recién infectadas generadas intrínsecamente por una simple célula infecciosa al inicio de la infección es de aproximadamente 8 células (Baccam et al., 2006; Beauchemin et al., 2008; Nowak et al., 1997; Perelson et al., 1996; Ribeiro et al., 2010; Stafford et al., 2000).

b) Para obtener la estimación Bayesiana usando los datos del MOI 2×10^{-6} se consideran las siguientes distribuciones *a priori*.

$$\beta_{50} \sim \log - normal(-9.155069, (0.1)^2),$$

$$a \sim gamma(0.49, 1),$$

$$k \sim \log - normal(11.31, (0.01)^2),$$

$$k_{50} \sim gamma(0.32, 4),$$

$$\sigma^2 \sim gamma(1, 20).$$

Los resultados se presentan en el Cuadro 4.3.

Cuadro 4.4: Estimaciones Bayesianas usando los datos del MOI 2×10^{-6} .

Parámetro		Valor	Intervalo creíble del 95 %
Tasa transmisión infecciosa	β_{50}	1.06×10^{-4}	$8.68 \times 10^{-5} - 1.28 \times 10^{-4}$
Tasa de muerte de células infectadas	a	0.5	$2.20 \times 10^{-4} - 2.56$
Tasa de producción total del virus	k	3.01×10^4	$2.95 \times 10^4 - 3.06 \times 10^4$
Tasa de producción de virus infeccioso	k_{50}	7.6×10^{-2}	$9.96 \times 10^{-7} - 0.49$
Varianza	σ^2	0.04	$6.12 \times 10^{-4} - 9.71$

El Cuadro B.1 del **Anexo B** muestra el resultado del test de Gelman–Rubin, utilizando tres MCMC, el cual arroja un valor menor a 1.1 para cada uno de los parámetros, lo que implica que hay convergencia entre las Cadenas de Markov Monte Carlo, y esto puede verse para cada uno de los parámetros de manera gráfica en las Figuras B.1 y B.2 (ver **Anexo B**).

Tomando en cuenta los resultados de la estimación Bayesiana de los parámetros de interés, los valores y condiciones iniciales dados en el apartado de la estimación por mínimos cuadrados, se presentan las gráficas que muestran el comportamiento entre la solución numérica del modelo *ad-hoc* y los datos reales.

La Figura 4.4 muestra las gráficas de los comportamientos entre la solución numérica del modelo *ad-hoc* usando los resultados de la estimación Bayesiana y los datos reales (escala logarítmica) para la concentración de células blanco, células infectadas, carga viral total y carga viral infecciosa. La línea roja corresponde a ésta solución numérica y los puntos azules a los datos reales. Se observa que la solución numérica se aproxima mejor al comportamiento de las células infectadas, carga viral total y carga viral infecciosa. Además, en el gráfico de las células blanco mostrada en la parte superior izquierda de la misma figura, en los días 5, 6 y 7, los puntos de los datos reales están por encima de la solución numérica usando las estimaciones de Bayes, mismo comportamiento que se observa en la solución por mínimos cuadrados.

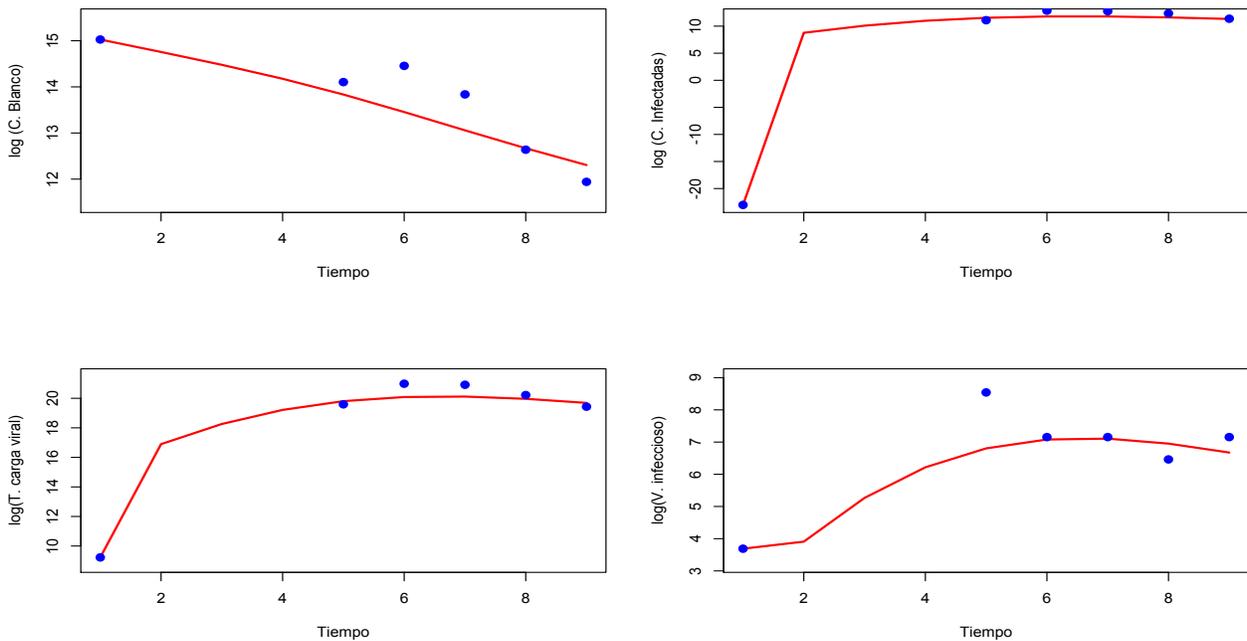


Figura 4.4: Datos reales y solución numérica del modelo *ad-hoc* usando resultados de la estimación Bayesiana para el $\text{MOI } 2 \times 10^{-6}$.

Utilizando los resultados de las estimaciones Bayesianas, el valor de dicho número reproductivo es $R_0^{ad-hoc} = 5.86$, es decir, el número de células recién infectadas generadas intrínsecamente por

una simple célula infecciosa al inicio de la infección es de aproximadamente 6 células.

4.2. Discusión

En los Cuadros 4.5 y 4.6 se realizan comparaciones entre los resultados de las estimaciones de mínimos cuadrados y Bayesianas para los MOIs 2×10^{-3} y 2×10^{-6} , respectivamente.

En el Cuadro 4.5, puede observarse que en la estimación Bayesiana todos los valores de los parámetros tienen el mismo exponente que los resultados de las estimaciones por el método de mínimos cuadrados. Así mismo, se observa que los valores de los estimadores por mínimos cuadrados están dentro de los intervalos creíbles (excepto la tasa de transmisión infecciosa, β_{50}).

Cuadro 4.5: Estimaciones obtenidas mediante el método de mínimos cuadrados e inferencia Bayesiana (MOI 2×10^{-3}).

Parámetro	V. L.*	Estimador		Intervalo creíble del 95 %
		M. C.**	Bayes	
β_{50}	4.95×10^{-5}	1.87×10^{-5}	1.29×10^{-5}	$1.06 \times 10^{-5} - 1.57 \times 10^{-5}$
a	1.18	1.37	1.01	$2.65 \times 10^{-2} - 3.84$
k	2.61×10^4	3.94×10^4	3.91×10^4	$3.83 \times 10^4 - 3.99 \times 10^4$
k_{50}	0.22	0.8	1.03	$2.17 \times 10^{-3} - 5.26$
R_0^{ad-hoc}	7.01	6.92	8.22	0.18 - 14.00

* Valor de la literatura (obtenido por [Iwami et al. \(2012\)](#))

** Mínimos cuadrados

Considerando la condición inicial de células blanco (HSC-F), los mismos valores iniciales de los parámetros en el modelo *ad-hoc*, se comparan los resultados de los números reproductivo básico de la infección, R_0^{ad-hoc} , por el método de mínimos cuadrados e inferencia Bayesiana para los valores de los cuatro parámetros estimados por ambos métodos. En la última fila del Cuadro 4.5 se presentan las estimaciones del número reproductivo básico. Ambos resultados son valores mayores que uno (ver Figura 4.5), lo que se interpreta, dada una población de células infectadas, la infección es muy probable que se extienda a toda la población de células blanco.

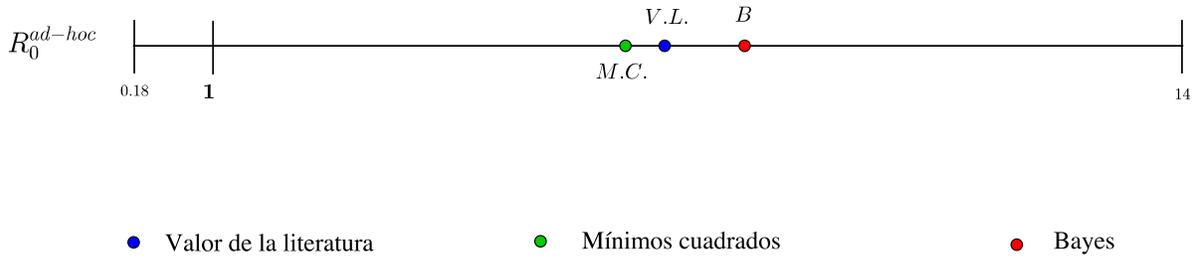


Figura 4.5: Intervalo creíble para R_0^{ad-hoc} y valores de R_0^{ad-hoc} obtenidos por ambos métodos para el MOI 2×10^{-3} .

Se retomaron los resultados de las estimaciones de los parámetros de interés por los métodos antes mencionados con los valores y condiciones iniciales descritos en la Subsección 4.1.1, se presenta la gráfica (Figura 4.6) que muestra una comparación de las soluciones numéricas del modelo *ad-hoc* por ambos métodos junto con los datos reales al considerar el MOI 2×10^{-3} .

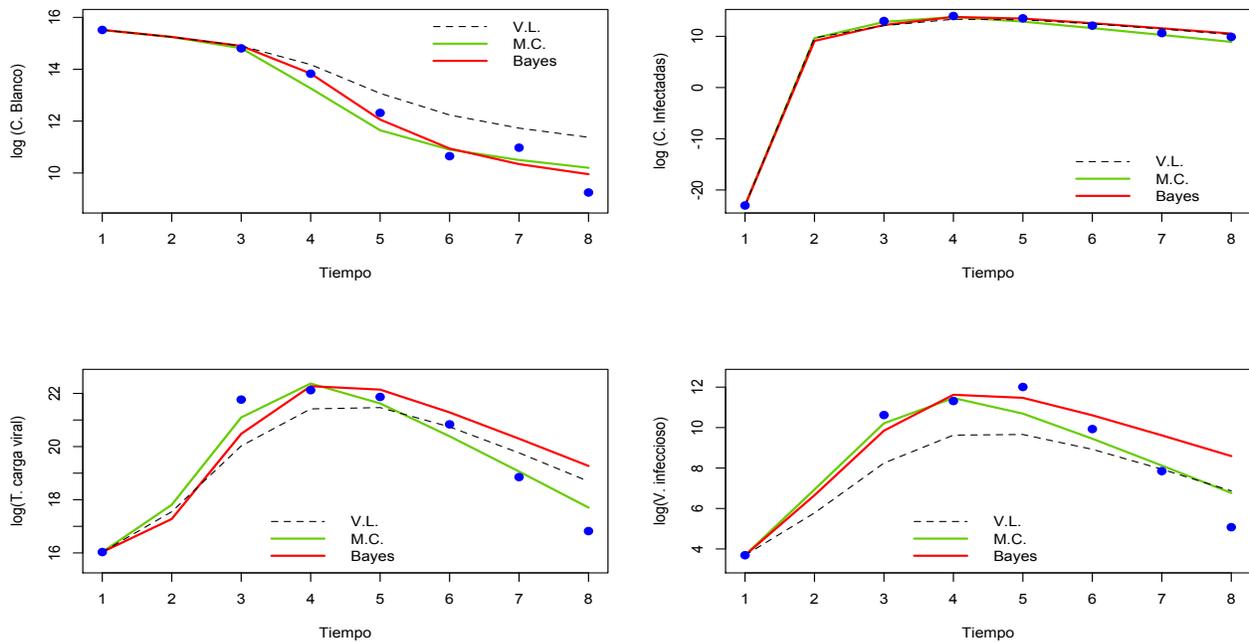


Figura 4.6: Comparación de los datos reales, solución obtenida por [Iwami et al. \(2012\)](#) y las soluciones numéricas del modelo *ad-hoc* usando mínimos cuadrados e inferencia Bayesiana para el MOI 2×10^{-3} .

La Figura 4.6 muestra la comparación de las soluciones numéricas del modelo *ad-hoc*, por los métodos ya mencionados. El objetivo es visualizar que método da como resultado un mejor ajuste de la solución numérica a los datos reales. En el gráfico de la células blanco el mejor ajuste se obtiene mediante inferencia Bayesiana, en el gráfico de células infectadas ambos métodos tienen un buen ajuste, para el total de la carga viral y virus infeccioso, la solución Bayesiana alcanza mejor los puntos máximos, pero en los días 7 y 8 la solución Bayesiana se aleja de los datos reales, pasa algo similar con la solución de mínimos cuadrados en el día 8.

En el Cuadro 4.6 puede observarse que en la estimación Bayesiana todos los valores de los parámetros tienen el mismo exponente que los resultados de las estimaciones por el método de mínimos cuadrados, al igual que en el MOI 2×10^{-3} . Así mismo, se observa que los valores de los estimadores por mínimos cuadrados están dentro de los intervalos creíbles.

Cuadro 4.6: Estimaciones obtenidas mediante el método de mínimos cuadrados e inferencia Bayesiana (MOI 2×10^{-6}).

Parámetro	V. L.*	Estimador		Intervalo creíble del 95 %
		M. C.**	Bayes	
β_{50}	4.95×10^{-5}	1.05×10^{-4}	1.06×10^{-4}	$8.68 \times 10^{-5} - 1.28 \times 10^{-4}$
a	1.18	0.49	0.5	$2.20 \times 10^{-4} - 2.56$
k	2.61×10^4	3.01×10^4	3.01×10^4	$2.95 \times 10^4 - 3.06 \times 10^4$
k_{50}	0.22	0.08	0.07	$9.96 \times 10^{-7} - 0.49$
R_0^{ad-hoc}	7.01	6.22	5.86	$6.12 \times 10^{-4} - 9.71$

* Valor de la literatura (obtenido por [Iwami et al. \(2012\)](#))

** Mínimos cuadrados

Al considerar la condición inicial de células blanco (HSC-F), los mismos valores iniciales de los parámetros en el modelo *ad-hoc*, se comparan los resultados de los números reproductivo básico de la infección, R_0^{ad-hoc} , por el método de mínimos cuadrados e inferencia Bayesiana para los valores de los cuatro parámetros estimados por ambos métodos. En la última fila del Cuadro 4.6 se presentan las estimaciones del número reproductivo básico. Ambos resultados son valores mayores que uno, lo cual quiere decir que, dada una población de células infectadas, es muy probable que la infección se extienda a toda la población de células blanco.

Se retoman los resultados de las estimaciones de los parámetros de interés por los métodos antes mencionados con los valores y condiciones iniciales descritos en la Subsección 4.1.1, se presentan las gráficas que muestran la comparación de las soluciones numéricas del modelo *ad-hoc* por ambos métodos y los datos reales del MOI 2×10^{-6} .

La Figura 4.7 muestra la comparación de las soluciones numéricas del modelo *ad-hoc*, por los métodos ya mencionados. El objetivo es comparar gráficamente los resultados de las soluciones numéricas, de los métodos de estimación, con los datos reales. Para las células blanco, puede apreciarse un mejor ajuste mediante estimación Bayesiana; para las células infectadas y total de la carga viral, la solución de ambos métodos de estimación describen bien y de manera similar el comportamiento de los datos reales; y para la carga viral infecciosa, la estimación Bayesiana describe mejor el comportamiento de los datos reales.

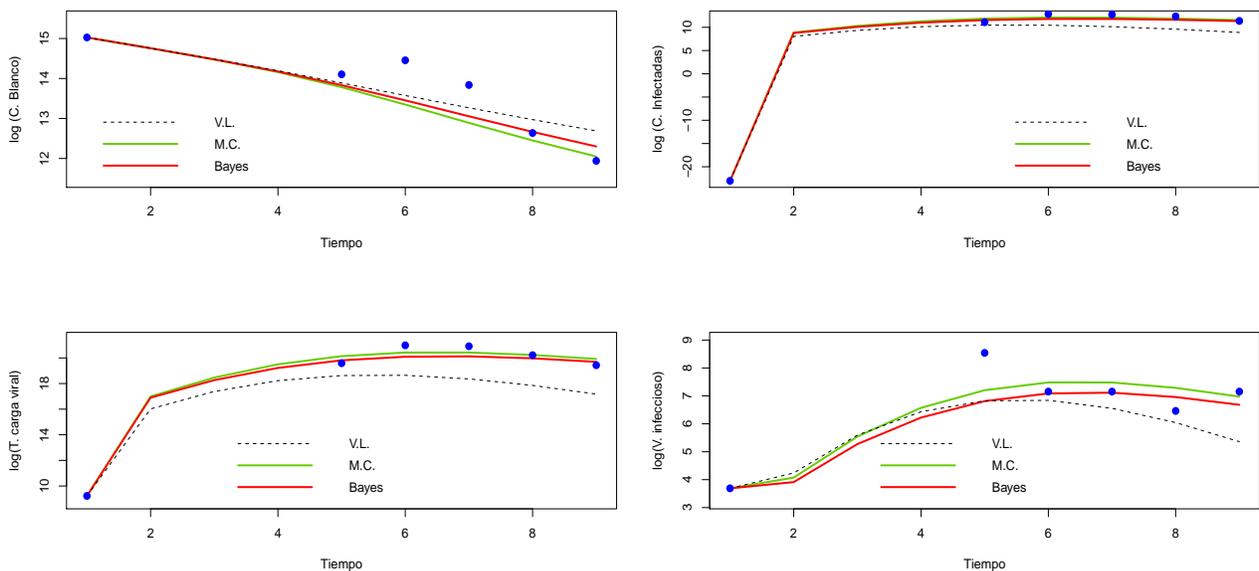


Figura 4.7: Comparación de los datos reales, solución obtenida por [Iwami et al. \(2012\)](#) y las soluciones numéricas del modelo *ad-hoc* usando mínimos cuadrados e inferencia Bayesiana para el MOI 2×10^{-6} .

De acuerdo con [Iwami et al. \(2012\)](#) las medidas relevantes de la cinética viral del VISH- KS661 *in vitro* están relacionadas con el parámetro a que representa la tasa de muerte de las células infectadas. Los resultados de este trabajo usando estimación Bayesiana, para el MOI 2×10^{-3}

a es igual a 1.018 y β_{50} es igual a 1.296×10^{-5} ; mientras que para el MOI 2×10^{-6} a es igual a 0.5 y β_{50} es igual a 1.06×10^{-4} .

El promedio de vida de las células infectadas está dado por el inverso de la tasa de muerte (a^{-1}). Para este trabajo el promedio de vida de las células infectadas para el MOI 2×10^{-3} es aproximadamente un día; mientras que para el MOI 2×10^{-6} es aproximadamente 2 días.

La vida media de la células infectadas está dada por $\log(\frac{2}{a})$. En este trabajo es aproximadamente 16 hrs (0.68 día) para el MOI 2×10^{-3} ; mientras que para el MOI 2×10^{-6} es aproximadamente 33 hrs (1.3 día).

El número total de partículas virales que produce una célula infectada (burst size) está dado por la tasa de producción total de virus (k) entre la tasa de muerte de la células infectadas a . Para este trabajo resulta que para el MOI 2×10^{-3} una célula infectada produce 2.58×10^4 virus (IC: 11738.48 – 932747.64); mientras que para el MOI 2×10^{-6} una célula infectada produce 6.02×10^4 virus (IC: 13,380.48 – 7,805,161.65). En este trabajo se observa que a mayor MOI una célula infectada produce menos virus.

El número de partículas virales infecciosas que produce una célula infectada (burst size) está dado por la tasa de producción de virus infeccioso k_{50} , entre la tasa de muerte de la células infectadas a . Para este trabajo resulta que para el MOI 2×10^{-3} una célula infectada produce 1.01 virus infecciosos (IC: 0.00096 – 25.05); mientras que para el MOI 2×10^{-6} una célula infectada produce 0.14 virus infecciosos (IC: 0.001 – 14.769). Para este trabajo, a mayor MOI, una célula infectada produce más virus infecciosos.

En este trabajo, considerando la estimación obtenida de la tasa de transmisión infecciosa ($\beta_{50} = 1.296 \times 10^{-5}$) para el MOI 2×10^{-3} , significa que se requiere el encuentro de aproximadamente 10,000 células blanco con una partícula viral en un ml para producir 129 células infectadas; mientras que para el MOI 2×10^{-6} la estimación obtenida para la tasa de transmisión infecciosa es de $\beta_{50} = 1.06 \times 10^{-4}$, esto es, se requiere el encuentro de aproximadamente 10,000 células blanco con una partícula viral en un ml para producir 1,060 células infectadas. Entonces, entre más pequeña es la tasa de transmisión infecciosa, se producen más células infectadas.

Estos resultados pueden ser de utilidad para futuras investigaciones sobre la cinética viral del VISH-KS661 *in vitro*.

Conclusiones

La estimación Bayesiana permite incorporar el conocimiento empírico o la experiencia que se tiene respecto a los parámetros desconocidos, en términos de distribuciones *a priori*; otra ventaja se refiere al tamaño de muestra. En este estudio, el tamaño de muestra es pequeño, y los parámetros, en razón de que se interpretan como tasas (infección, muerte y producción viral) todos se consideran positivos.

Utilizando el modelo matemático *ad-hoc* para la infección por VISH-KS661 en células HSC-F, el conjunto de datos experimentales y la metodología Bayesiana, resulta posible estimar los parámetros que caracterizan las interacciones célula – virus *in vitro*, así como las medidas relevantes de la cinética viral de dicho virus. Con base en estos resultados, se obtiene una buena descripción cuantitativa de la cinética viral del VISH-KS661, mismos que son consistentes con resultados de estudios previos del virus. De igual manera, comparando estos resultados con los obtenidos por mínimos cuadrados, para el MOI 2×10^{-6} la solución numérica utilizando las estimaciones Bayesiana ajusta mejor al comportamiento de los datos reales del experimento.

El número reproductivo básico de la infección R_0^{ad-hoc} , obtenido con los resultados de los estimadores Bayes para los MOIs 2×10^{-3} y 2×10^{-6} , resultaron similares a estudios previos obtenidos por otros métodos. Dado que el R_0^{ad-hoc} obtenido en ambos MOIs y el referente a la literatura

(Iwami et al., 2012) son mayor que uno para los datos analizados (*in vitro*), se concluye que la infección va a persistir hasta que ya no queden más células blanco que infectar y las células infectadas mueran.

Respecto a las medidas relevantes de la cinética viral del VISH-KS661 se concluye que, el número total de partículas virales que produce una célula infectada (burst size) para los MOIs 2×10^{-3} y 2×10^{-6} es de 2.58×10^4 y 6.02×10^4 virus, respectivamente; el número de partículas virales infecciosas que produce una célula infectada para los MOIs 2×10^{-3} y 2×10^{-6} es de 1.01 y 0.14 virus, respectivamente . Estos resultados son congruentes con los obtenidos por Iwami et al. (2012).

Gráficos y cuadros correspondientes al MOI 2×10^{-3}

En este anexo se presentan los cuadros y los gráficos correspondientes al test de Gelman-Rubin, así como la trazas y densidades marginales *a posteriori* de cada parámetro de interés para los MOI 2×10^{-3} .

Cuadro A.1: Test de Gelman–Rubin para el MOI 2×10^{-3} .

	\widehat{R}	Límite superior del I.C.
β_{50}	1.00	1.00
a	1.00	1.00
k	1.06	1.18
k_{50}	1.00	1.00
σ^2	1.00	1.00

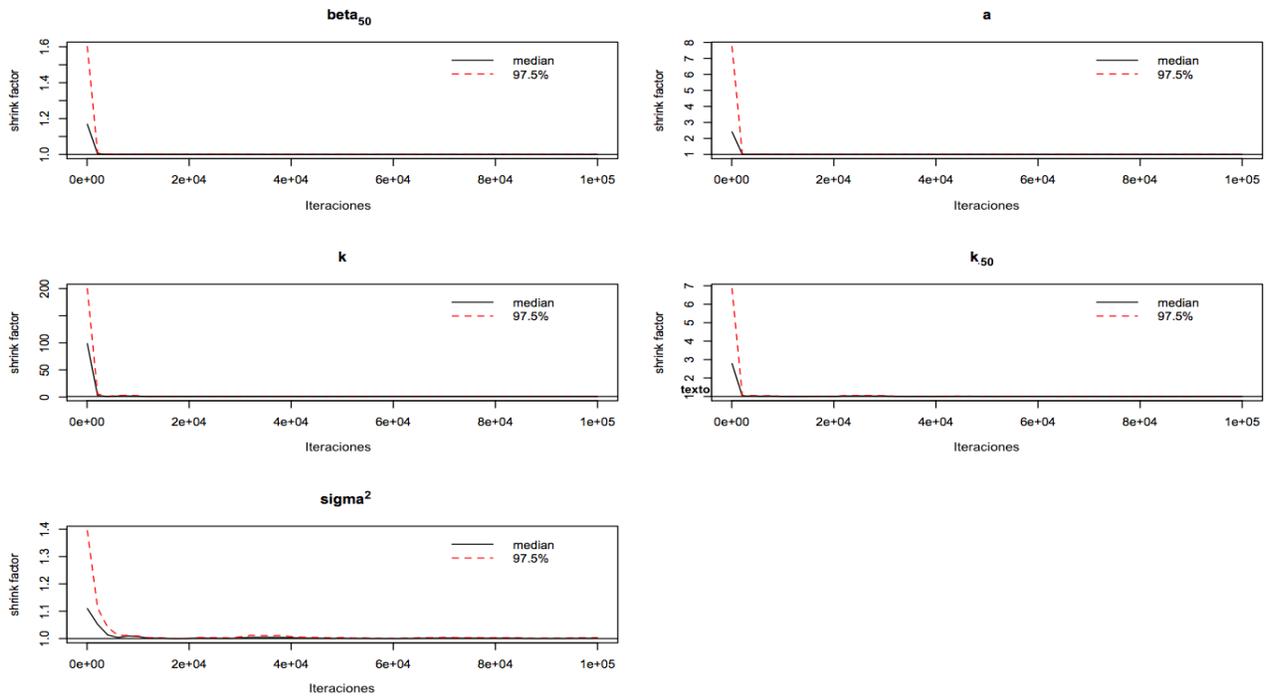


Figura A.1: Diagnóstico de convergencia de Gelman–Rubin para β_{50} , a , k , k_{50} y σ^2 .

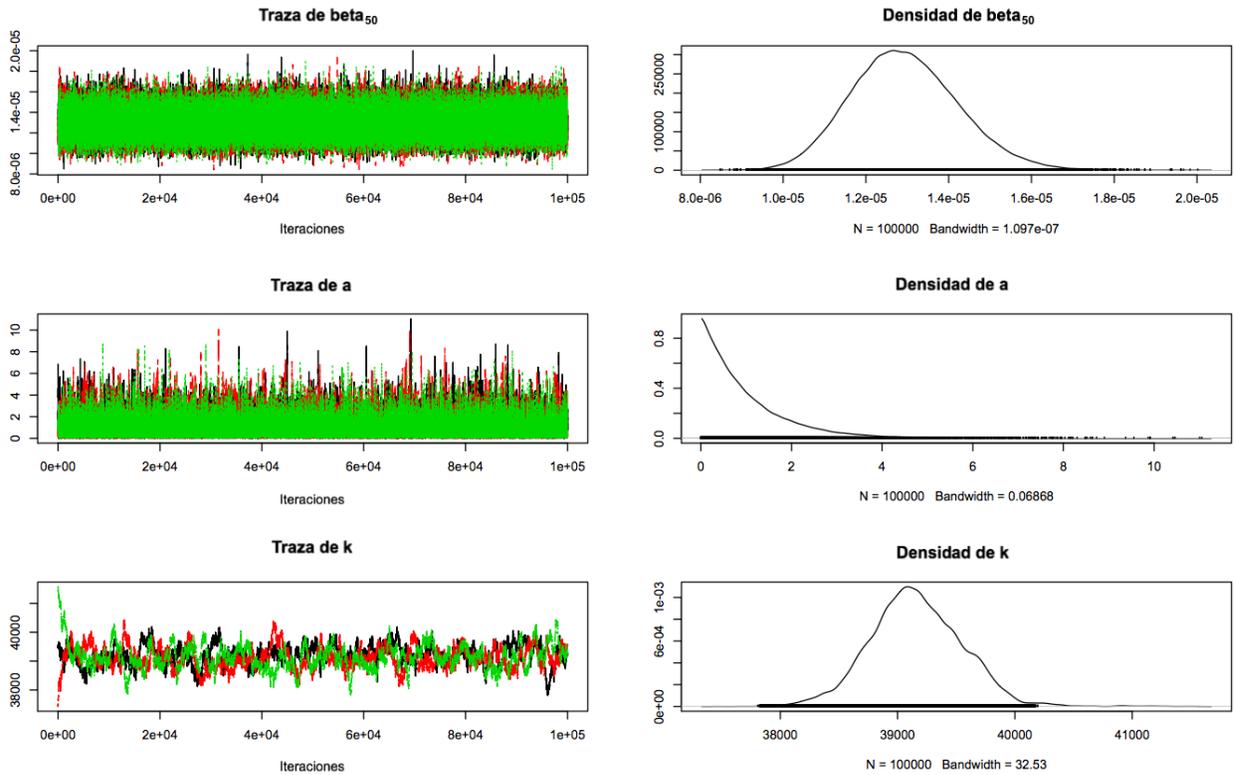


Figura A.2: Trazas y distribuciones *a posteriori* para β_{50} , a , y k .

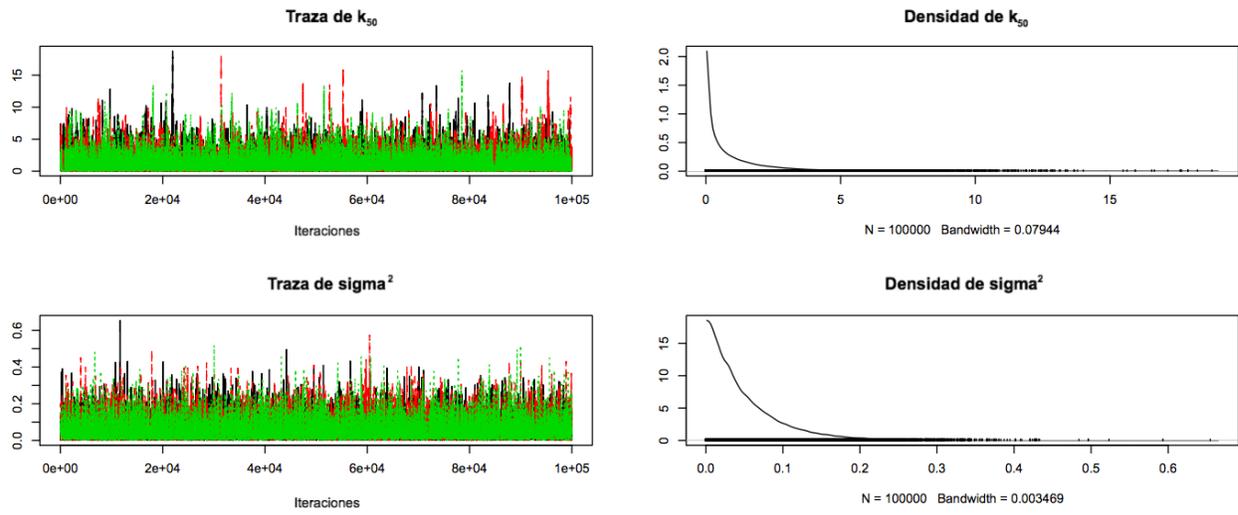


Figura A.3: Trazas y distribuciones *a posteriori* para k_{50} y σ^2 .

Gráficos y cuadros correspondientes al MOI 2×10^{-6}

En este anexo se presentan los cuadros y los gráficos correspondientes al test de Gelman-Rubin, así como la trazas y densidades marginales *a posteriori* de cada parámetro de interés para los MOI 2×10^{-6} .

Cuadro B.1: Test de Gelman–Rubin para el MOI 2×10^{-6} .

	\widehat{R}	Límite superior del I.C.
β_{50}	1.00	1.00
a	1.00	1.01
k	1.01	1.04
k_{50}	1.01	1.00
σ^2	1.00	1.00

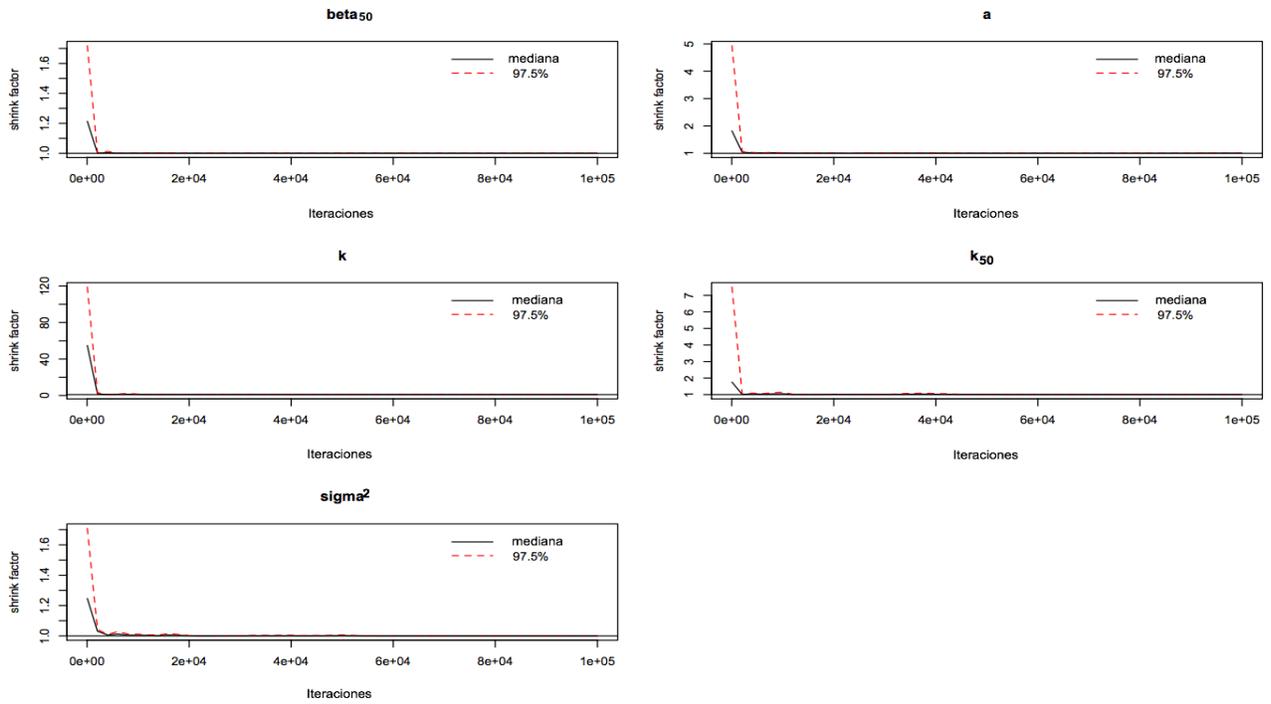


Figura B.1: Diagnóstico de convergencia de Gelman–Rubin para β_{50} , a , k , k_{50} y σ^2 .

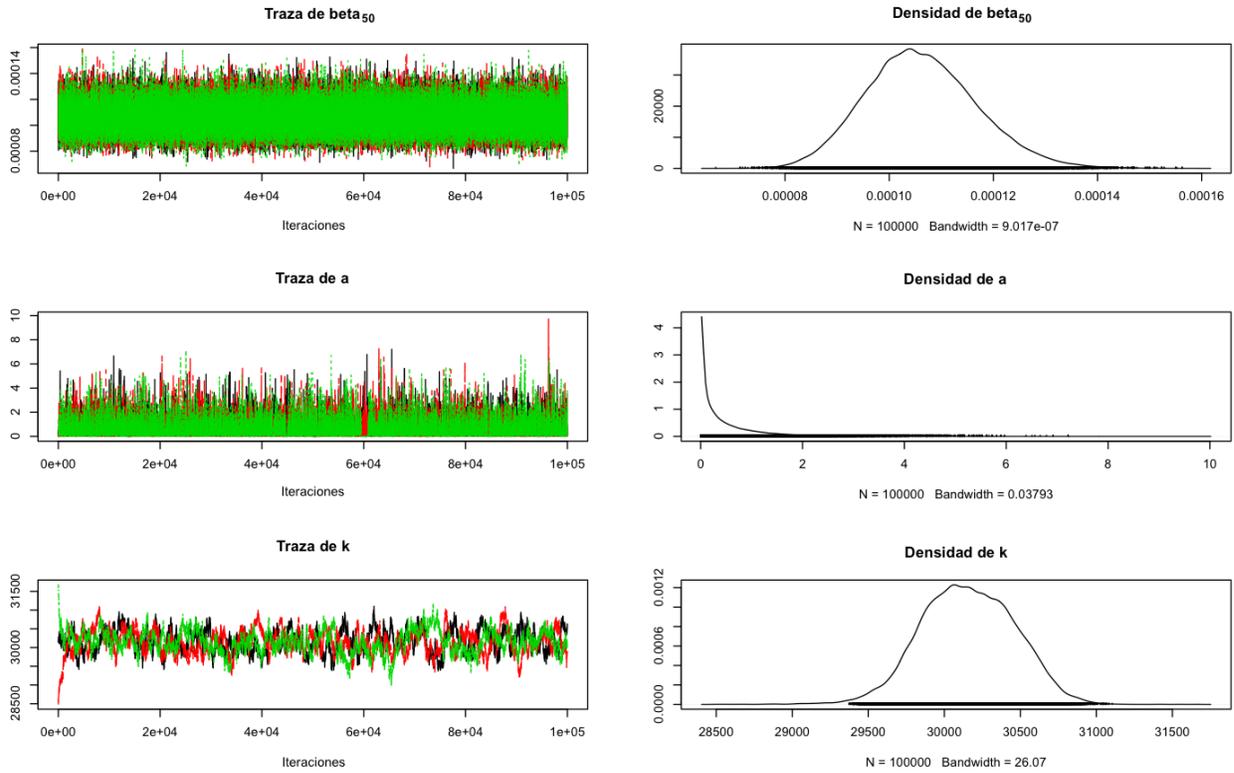


Figura B.2: Trazas y distribuciones *a posteriori* para β_{50} , a , y k

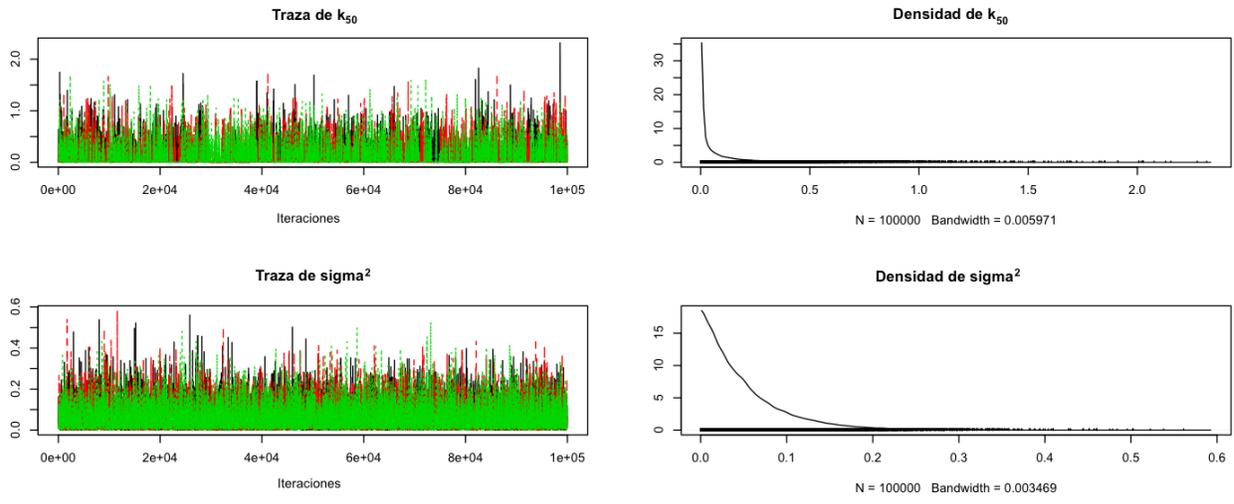


Figura B.3: Trazas y distribuciones *a posteriori* para k_{50} y σ^2 .

Referencias

- Agbaje, O., Luzio, S., Albarrak, A., J Lunn, D., R Owens, D., y Hovorka, R. (2003). Bayesian hierarchical approach to estimate insulin sensitivity by minimal model. *Clinical science (London, England : 1979)*, 105:551–60.
- Allen, L. (2007). *An Introduction to Mathematical Biology*. Pearson/Prentice Hall.
- Andrea (2012). VIH e inmunidad. <https://andreabiologia.wordpress.com/>.
- Baccam, P., Beauchemin, C., Macken, C. A., Hayden, F. G., y Perelson, A. S. (2006). Kinetics of influenza a virus infection in humans. *Journal of Virology*, 80(15):7590–7599.
- Beauchemin, C. A., McSharry, J. J., Drusano, G. L., Nguyen, J. T., Went, G. T., Ribeiro, R. M., y Perelson, A. S. (2008). Modeling amantadine treatment of influenza a virus in vitro. *Journal of Theoretical Biology*, 254(2):439 – 451.
- Boersch-Supan, P. H., Ryan, S. J., y Johnson, L. R. (2017). deBinfer: Bayesian inference for dynamical models of biological systems. *Methods in Ecology and Evolution*, 8:511–518.
- Capistrán, M. A., Christen, J. A., y Donnet, S. (2016). Bayesian analysis of odes: Solver optimal accuracy and bayes factors. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):829–849.
- Choisy, M., Guégan, J.-F., y Rohani, P. (2006). *Mathematical Modeling of Infectious Diseases Dynamics*, pages 379–404. John Wiley & Sons, Inc.
- Collier, L. (2008). *Virología humana*. McGraw-Hill Interamericana de España S.L.
- Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., y Rubin, D. (2013). *Bayesian Data Analysis, Third Edition*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press.

- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In Bernardo, J. M., Berger, J., Dawid, A. P., y Smith, J. F. M., editors, *Bayesian Statistics 4*, pages 169–193. Oxford University Press, Oxford.
- Gómez, G. y Delicado, P. (2016). Curso de inferencia y decisión. <http://www-eio.upc.edu/~delicado/docencia/IyDapuntes.pdf>.
- Iwami, S., Holder, B. P., Beauchemin, C. A., Morita, S., Tada, T., Sato, K., Igarashi, T., y Miura, T. (2012). Quantification system for the viral dynamics of a highly pathogenic simian/human immunodeficiency virus based on an *in vitro* experiment and a mathematical model. *Retrovirology*, 9(1):18.
- Lewin, S. R., Ribeiro, R. M., Walters, T., Lau, G. K., Bowden, S., Locarnini, S., y Perelson, A. S. (2001). Analysis of hepatitis b viral load decline under potent therapy: Complex decay profiles observed. *Hepatology*, 34(5):1012–1020.
- Lomba, S. (2014). Desenmascarando a las ITS. <http://itsenfermeria.blogspot.com/2014/10/como-se-produce-realmente-la-infeccion.html>.
- Medina, L. Q., Guerrero, J., Palomar, L., y Giovanna, R. (2016). Bioestadística en virología: índice multiplicidad de infección (moi).
- Montesinos, O. A. y Hernández, C. M. (2007). Modelos matemáticos para enfermedades infecciosas. *Salud Pública de México*, 49:218 – 226.
- Mood, A., Graybill, F., y Boes, D. (1974). *Introduction to the Theory of Statistics*. International Student edition. McGraw-Hill.
- Nowak, M. A., Bonhoeffer, S., Hill, A. M., Boehme, R., Thomas, H. C., y McDade, H. (1996). Viral dynamics in hepatitis b virus infection. *Proceedings of the National Academy of Sciences of the United States of America*, 93(9):4398–4402.
- Nowak, M. A., Lloyd, A. L., Vasquez, G. M., Wiltout, T. A., Wahl, L. M., Bischofberger, N., Williams, J., Kinter, A., Fauci, A. S., Hirsch, V. M., y Lifson, J. D. (1997). Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection. *Journal of Virology*, 71(10):7518–7525.
- Perelson, A. (2002). Modelling viral and immune system dynamics. 2:28–36.

- Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M., y Ho, D. D. (1996). Hiv-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271(5255):1582–1586.
- Putter, H., Heisterkamp, S. H., Lange, J. M. A., y Wolf, F. d. (2002). A bayesian approach to parameter estimation in hiv dynamical models. *Statistics in Medicine*, 21(15):2199–2214.
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ribeiro, R. M., Qin, L., Chavez, L. L., Li, D., Self, S. G., y Perelson, A. S. (2010). Estimation of the initial viral growth rate and basic reproductive number during acute hiv-1 infection. *Journal of Virology*, 84(12):6096–6102.
- Rizzo, M. (2007). *Statistical Computing with R*. Chapman & Hall/CRC The R Series. CRC Press.
- Soetaert, K., Petzoldt, T., y Setzer, R. W. (2010). Solving differential equations in r: Package desolve. *Journal of Statistical Software*, 33(9):1–25.
- Stafford, M. A., Corey, L., Cao, Y., Daar, E. S., Ho, D. D., y Perelson, A. S. (2000). Modeling plasma virus concentration during primary hiv infection. *Journal of Theoretical Biology*, 203(3):285 – 301.
- Süli, E. y Mayers, D. (2003). *An Introduction to Numerical Analysis*. Cambridge University Press.
- Tsiang, M. y Gibbs, C. S. (1999). *Analysis of Hepatitis B Virus Dynamics and Its Impact on Antiviral Development*, chapter 28, pages 361–377. Humana Press.
- Weidemann, F., Dehnert, M., Koch, J., Wichmann, O., y Höhle, M. (2014). Bayesian parameter inference for dynamic infectious disease modelling: rotavirus in germany. *Statistics in Medicine*, 33(9):1580–1599.