

“Análisis de Incertidumbre y Sensibilidad: Un Modelo de Transmisión de VIH/SIDA en
Puerto Rico”

Por

Yetsabel Auccaille Quispe

Tesis sometida en cumplimiento parcial de los requisitos para el grado de

Maestría en Ciencias

En

Matemáticas (Estadística)

Universidad de Puerto Rico

Recinto Universitario de Mayagüez

Departamento de Ciencias Matemáticas

Julio 2011

Aprobado por:

Dámaris Santana Morant, Ph.D.
Presidente, Comité Graduado

Fecha

Edgardo Lorenzo González, Ph.D.
Miembro, Comité Graduado

Fecha

Karen Ríos Soto, Ph.D.
Miembro, Comité Graduado

Fecha

Julio C. Quintana Díaz, Ph.D.
Miembro, Comité Graduado

Fecha

Gloribell Ortiz, M.S.M.
Representante, Estudios Graduados

Fecha

Omar Colón Reyes, Ph.D.
Director Interino
Departamento de Ciencias Matemáticas

Fecha

Abstract

We studied the behavior of HIV/AIDS in Puerto Rico using epidemiological models. We adjusted the model proposed by Blower et al. (1991) to the data that is available for Puerto Rico to assess the factors that contribute to the prevalence of the virus in Puerto Rico. The model is designed to assess the epidemiological effects of the heterosexual transmission of HIV and the transmission of HIV among intravenous drug users. The model takes into account certain risky behaviors like: the rate of sharing needles per year (for men and women that use intravenous drugs), the rate of change of sex partners per year (for men and women that use intravenous drug and are sexually active), the rate of change of sex partners per year (for men and women that does not use intravenous drugs), and epidemiological parameters like: the heterosexual transmission efficiency per partnership, the HIV transmission efficiency per needle injection, the average adult incubation time, and the average adult survival time. We estimate the cumulative number of AIDS cases after ten years in Puerto Rico and identified the parameters that contributed most to the uncertainty in the estimation using uncertainty analysis and sensitivity analysis. These methods require estimates of the model parameters for Puerto Rico and their performance depends on the efficient exploration of the parameter space. Seaholm et al. (1988) and Blower and Dowlatabadi (1994) proposed the use of Latin Hypercube Sampling for this purpose. We proposed a modification of the Latin Hypercube Sampling Method using randomized Halton sequences. This method wereshown to be more efficient in the sensitivity analysis, in our model for Puerto Rico, in the identification of the parameters that contributed most to the uncertainty of the estimation as well as in

the convergence of the mean of some of the distributions of interest. The sensitivity analysis shows that the parameters that affect the cumulative number of AIDS in the next ten years in Puerto Rico are: the heterosexual transmission efficiency per partnership of the HIV virus, the average adult incubation time of the VIH virus and the rate of sex partners per year for women that use intravenous drugs and are sexually active.

Resumen

Este trabajo de investigación se realizó con el propósito de estudiar el comportamiento del VIH/SIDA en Puerto Rico usando modelos epidemiológicos. Ajustamos el modelo matemático determinístico propuesto por Blower et al. (1991) a los datos disponibles para Puerto Rico para evaluar los factores que afectan la prevalencia del virus en Puerto Rico. Este modelo está diseñado para evaluar el efecto epidemiológico de la transmisión heterosexual y la transmisión entre usuarios de droga intravenosa (UDI), tomando en cuenta parámetros de conductas de riesgo tales como: la tasa de intercambio de inyectables por año (en mujeres y hombres UDI), y la tasa de cambio de parejas sexuales por año (en mujeres que son UDI y sexualmente activas y en hombres UDI y sexualmente activos), y tasa de cambio de parejas sexuales por año (en mujeres No-UDI y en hombres No-UDI). También se consideran parámetros epidemiológicos como: la eficiencia de transmisión del VIH por inyectable dado que el inyectable está infectado, eficiencia de transmisión heterosexual por contacto sexual dado que la pareja está infectada, el tiempo promedio de incubación del virus y el tiempo promedio de sobrevivencia en adultos. Usando el modelo estimamos el número de casos acumulados de SIDA en adultos en Puerto Rico para los próximos diez años a partir del año 2010 e identificamos los parámetros que contribuyen más a la incertidumbre de estos estimados usando análisis de incertidumbre y análisis de sensibilidad. Estos métodos requieren estimados de los parámetros para Puerto Rico y su desempeño depende de la eficiencia con la que se explore el espacio muestral de estos parámetros. El método de muestreo por Hipercubo Latino Seaholm et al. (1988), Blower y Dowlatabadi (1994) se ha utilizado con este fin. Proponemos una modificación a este método usando secuencias de Halton aleatorizadas

que resultó ser más eficiente en el análisis de sensibilidad en nuestro modelo para Puerto Rico en cuanto a la identificación de parámetros importantes que contribuyen a la incertidumbre de los estimados y en ocasiones con respecto a la convergencia del promedio de las algunas de distribuciones de interés. El análisis de sensibilidad reveló que los parámetros de entrada importantes que afectan el número de casos acumulados de SIDA en Puerto Rico, para los próximos diez años son la eficiencia de transmisión del VIH, el tiempo promedio de incubación del virus del VIH y la tasa de cambio de parejas sexuales para mujeres que usan droga intravenosas y son sexualmente activas.

DEDICATORIA

*Desde el fondo de mi corazón,
dedico este esfuerzo de vida a **DIOS**,
por la fortaleza que me dio en momentos
que más necesitaba.*

*Con todo el amor infinito dedico a
mis padres **Juan y Marbila** por su
invalorable esfuerzo, motivación y
confianza, que me brindaron para la
conclusión de este trabajo.*

*A mi hermano Juan Hemersón,
por su apoyo y confianza.*

*Quien prefiere los valores del espíritu, elige los más divinos, quien los del cuerpo (elige) los
humanos.*

AGRADECIMIENTOS

Agradezco

A la Dra. Dámaris Santana Morant, por su apoyo y comprensión en este trabajo. Por la experiencia y el ejemplo brindado para ser buenos estadísticos y ser mejores personas.

Por incentivar y transmitirme sus propuestas.

Al Dr. Julio C. Quintana Díaz, por su sugerencia de continuar el estudio del VIH/SIDA para la población de Puerto Rico desde el punto de vista epidemiológico.

A la Dra. Karen Ríos Soto, por su contribución, guía y sugerencias particularmente respecto a los modelos epidemiológicos.

Índice General

Abstract.	ii
Resumen	iv
DEDICATORIA	vi
AGRADECIMIENTOS	vii
Índice general	viii
Índice de figuras	xi
Índice de tablas.	xvii
1 PRELIMINARES	1
1.1 Introducción.	1
1.2 Justificación.	4
1.3 Antecedentes.	6
1.4 Objetivos.	10
1.4.1 Objetivo General.	10
1.4.2 Objetivos Específicos.	11
2 REVISIÓN LITERARIA	13
2.1 Modelos Epidemiológicos.	13
2.2 Análisis de Incertidumbre.	17
2.3 Análisis de Sensibilidad.	19

2.4 Exploración del espacio muestral de los parámetros que entran al modelo.....	24
2.4.1 Muestreo Aleatorio.....	25
2.4.2 Muestreo por Hipercubo Latino.....	25
2.4.3 Secuencia de Halton.....	26
3 Modelo matemático de transmisión del VIH/SIDA en Puerto Rico	29
3.1 Ecuaciones del modelo para Puerto Rico.....	36
3.2 Estimación puntual y de la función de densidad de los parámetros	48
4 Análisis del modelo VIH/SIDA para Puerto Rico	58
4.1 Análisis para el valor puntual de los parámetros.....	58
4.2 Muestreo por hipercubo latino modificado.....	63
4.3 Análisis de incertidumbre.....	65
4.4 Análisis de sensibilidad.....	70

5 Conclusiones y trabajo futuro	80
6 Limitaciones	83
Apéndice.	84
Bibliografía	102

Índice de Figuras

2.1 Secuencia de Halton y la permutación aleatoria de la secuencia de Halton	28
3.1 Tamaño inicial de los estratos del modelo para Puerto Rico.	32
3.2 Diagrama de flujo del modelo para Puerto Rico	37
4.1 Nivel de seroprevalencia (%) en hombres y mujeres UDI.	59
4.2 Nivel de seroprevalencia (%) en hombres y mujeres UDI.	60
4.3 Seroprevalencia en los subgrupos UDI (lado izquierdo) y UDISA (lado derecho).	61
4.4 Seroprevalencia en los subgrupos en hombres y mujeres (No-UDI)SA	62
4.5 Gráfica de los intervalos visitados por MHL y MHL modificado, respectivamente	64

4.6 Histograma de frecuencia del número de casos acumulados de SIDA en adultos, para los próximos 10 años. Izquierda superior: Aleatorio n=1000; derecha superior: Aleatorio n=100; izquierda inferior: MHL n=100 y derecha inferior: MHL modificado n=100.	67
4.7 Histograma de frecuencia del número de casos acumulados de SIDA en adultos, para los próximos 10 años. Izquierda superior: Aleatorio n=1000; derecha superior: Aleatorio n=200; izquierda inferior: MHL n=200 y derecha inferior: MHL modificado n=200.	69
4.8 Convergencia del promedio de casos acumulados de SIDA en adultos, para los próximos 10 años. Izquierda superior: Aleatorio n=1000; derecha superior: Aleatorio n=200; izquierda inferior: MHL n=200 y derecha inferior: MHL modificado n=200.	75
4.9 Convergencia del promedio de la eficiencia de transmisión de VIH por contacto heterosexual, para los próximos 10 años. Izquierda superior: Aleatorio n=1000; derecha superior: Aleatorio n=200; izquierda inferior: MHL n=200 y derecha inferior: MHL modificado n=200.	76

4.10	Convergencia del promedio de la tasa de cambio de parejas sexuales para mujeres UDISA, para los próximos 10 años. Izquierda superior: Aleatorio n=1000; derecha superior: Aleatorio n=200; izquierda inferior: MHL n=200 y derecha inferior: MHL modificado n=200.	77
4.11	Convergencia del promedio del tiempo de incubación del virus del VIH, para los próximos 10 años. Izquierda superior: Aleatorio n=1000; derecha superior: Aleatorio n=200; izquierda inferior: MHL n=200 y derecha inferior: MHL modificado n=200.	78
A.1	Exploración del espacio muestral generado por el muestreo aleatorio n = 1000, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1	84
A.2	Exploración del espacio muestral generado por el muestreo aleatorio n = 1000, para los parámetros v vs s y C_{m2} vs K_{h3}	85
A.3	Histograma de los parámetros de entrada del modelo i_m , β_{mh} , C_{m2} , C_{m3} , i_h , K_{h2} , K_{h3} , β_1 , v y s generado por el muestreo aleatorio n = 1000.	86

A.4 Exploración del espacio muestral generado por el muestreo aleatorio n = 100, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2} vs K_{h3}	87
A.5 Histograma de los parámetros de entrada del modelo i_m , β_{mh} , C_{m2} , C_{m3} , i_h , K_{h2} , K_{h3} , β_1 , v y s generado por el muestreo aleatorio n = 100.	88
A.6 Exploración del espacio muestral generado por el muestreo aleatorio n = 200, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2} vs K_{h3}	89
A.7 Histograma de los parámetros de entrada del modelo i_m , β_{mh} , C_{m2} , C_{m3} , i_h , K_{h2} , K_{h3} , β_1 , v y s generado por el muestreo aleatorio n = 200.	90
A.8 Exploración del espacio muestral generado por el muestreo hipercubo latino n = 100, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2} vs K_{h3}	91
A.9 Histograma de los parámetros de entrada del modelo i_m , β_{mh} , C_{m2} , C_{m3} , i_h , K_{h2} , K_{h3} , β_1 , v y s generado por el muestreo por hipercubo latino n = 100.	92

A.10 Exploración del espacio muestral generado por el muestreo hipercubo latino n = 200, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2} vs K_{h3}	93
A.11 Histograma de los parámetros de entrada del modelo i_m , β_{mh} , C_{m2} , C_{m3} , i_h , K_{h2} , K_{h3} , β_1 , v y s generado por el muestreo por hipercubo latino n = 200.	94
A.12 Exploración del espacio muestral generado por el muestreo hipercubo latino modificado n = 100, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2} vs K_{h3}	95
A.13 Histograma de los parámetros de entrada del modelo i_m , β_{mh} , C_{m2} , C_{m3} , i_h , K_{h2} , K_{h3} , β_1 , v y s generado por el muestreo por hipercubo latino modificado n = 100.	96
A.14 Exploración del espacio muestral generado por el muestreo hipercubo latino modificado n = 200, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2} vs K_{h3}	97

A.15 Histograma de los parámetros de entrada del modelo $i_m, \beta_{mh}, C_{m2}, C_{m3}, i_h, K_{h2}, K_{h3}, \beta_1, v$ y s generado por el muestreo por hipercubo latino modificado $n = 200$ 98

Índice de tablas

3.1 Definición de los parámetros para el modelo	37
3.2 Estimados puntuales de los parámetros	55
3.3 Función de distribución de los parámetros	56
4.1 Estadística descriptiva del número de casos acumulados de SIDA en adultos para $n = 1000$ y $n = 100$	67
4.2 Estadística descriptiva del número de casos acumulados de SIDA en adultos para $n = 1000$ y $n = 200$	69
4.3 Prueba de correlación parcial para $n = 1000$ y $n = 100$	71
4.4 Prueba de correlación parcial para $n = 1000$ y $n = 200$	72
A.1 Vía de utilización de droga	99

A.2 Diagnóstico de salud física por sexo de las personas que usan droga	
con inyectable	100
A.3 Promedio de años de uso de droga intravenosa	101

Capítulo 1

PRELIMINARES

1.1. Introducción

Una preocupación mundial en salud pública de las últimas décadas es el Virus de Inmunodeficiencia Humano (VIH) y el Síndrome de Inmunodeficiencia Adquirida (SIDA), que por sus características y distribución es considerado una pandemia. El SIDA es una enfermedad causada por el VIH. El virus ataca el sistema inmunológico, principal mecanismo de defensa del organismo contra las agresiones externas de virus, bacterias, hongos microscópicos y parásitos. El virus infecta principalmente las células CD4, también conocidas como células T, éstas son glóbulos blancos que dirigen las operaciones de defensa para luchar contra las enfermedades (Álvarez y Bravo, 2006).

El VIH/SIDA, no discrimina, es decir, puede afectar a cualquier persona, de

cualquier cultura, edad, género, situación económica, religión, opción política o grupo social. Existen poblaciones que tienden a ser más susceptibles, como por ejemplo, las personas involucradas en el trabajo sexual, usuarios de droga intravenosa, homosexuales y poblaciones de reclusos (personas en cárceles) entre otros. Una de las consecuencias de esta enfermedad es, que no sólo afecta a la persona que lo contrae, sino que deteriora progresivamente su entorno familiar, económico y social; además genera el rechazo de la sociedad, la discriminación, la homofobia y por ende el aislamiento y el daño psicológico (Álvarez y Bravo, 2006).

A finales del 2008, el Programa Conjunto de las Naciones Unidas sobre el VIH/SIDA (ONUSIDA) estimó que a nivel mundial habían 33.4 millones de personas que vivían con VIH/SIDA, de los cuales 31.3 millones eran adultos, 15.7 millones eran mujeres, y 2.1 millones eran niños, donde la mayoría de estas infecciones ocurrieron en el oeste del África. El número de adultos infectados con VIH en el 2008 fue de 2.7 millones y de niños infectados fue de 0.43 millones. Durante el año 2008 el SIDA causó 2 millones de muertes mundialmente. A fines del 2008 las mujeres representaron el 50 % de todos los adultos que vivían con VIH en todo el mundo. Para el 2009 alrededor del 67 % de personas con VIH vivían en África Subsahariana (ONUSIDA,2009).

Por otro lado, en los Estados Unidos a finales del 2007 el número estimado de personas que vivían con VIH/SIDA en 34 estados y 5 áreas dependientes era de 571,378, de los cuales 549,196 eran adultos y adolescentes, y 2,736 eran niños menores de 13 años (CDC).

El VIH/SIDA en Puerto Rico

Desde que inició el sistema de reporte de los casos de VIH (no-SIDA), en abril del 2003 y hasta el 30 de septiembre de 2005, se han reportado al Sistema de Vigilancia 4,558 casos de VIH (no-SIDA) en Puerto Rico. Aproximadamente el 65.2% del total de casos reportados correspondieron a diagnósticos en hombres. El porcentaje mayor de diagnósticos de VIH ocurrió en el grupo de 30 a 39 años de edad (34.4%), seguido por el grupo de 20 a 29 años (39.1%) y el grupo de 40 a 49 años (20.8%). La mayoría de los casos de VIH (no-SIDA) reportados al Sistema de Vigilancia fueron atribuidos al uso de drogas inyectables (41.8%), seguido por el contacto heterosexual sin protección (31.4%) y a hombres que tienen sexo con hombres (10%) (Marrero y Miranda, 2005).

Para fines del 2005, el Programa de Vigilancia de SIDA estimó en casi 23 mil los

casos de hombres con SIDA y en 7 mil los casos de mujeres, de los cuales el 76.5 % representa los hombres. Desde 1994 se ha observado una tendencia descendente en el número de casos nuevos de SIDA. Frente a los 2,687 en 1993, en el 2004 se detectaron 781 casos, y en el primer semestre de 2005 la cifra se redujo a 266 nuevos casos, esta reducción está asociada a la utilización de los tratamientos antirretrovirales. La mitad de los casos de SIDA en Puerto Rico se debe en gran parte al uso de drogas intravenosas como modo principal de contagio. Sin embargo, el modo de transmisión entre hombres y mujeres es muy distinto. Mientras el uso de droga intravenosa es el modo de transmisión más frecuente entre el 63 % de los hombres, entre las mujeres esa proporción se reduce al 37 %. En el caso de las mujeres el modo principal de transmisión es el contacto heterosexual (61 %), posiblemente por la prostitución y el sexo sin protección, en cambio, en los hombres el contacto heterosexual es considerado el modo de transmisión sólo en el 14 % (Lobato y Soto, 2006).

1.2. Justificación

Hernández Maldonado M. (1990), Varas Díaz N. (2002) y Resto Olivo J. (2003) investigaron la enfermedad del SIDA en Puerto Rico, desde el punto de vista

educativo, social y filosófico, respectivamente. Sherry Deren et al. (2004) compararon la incidencia del VIH entre usuarios de droga puertorriqueños mediante inyectable en Bayamón, e East Harlem, New York mediante el análisis de seroconversión usando variables predictoras sociodemográficas, de conductas de riesgo relacionada al consumo de drogas y de conducta sexual. Torres Velásquez, B. (2009) utilizó modelos de regresión lineal múltiple e identificó el uso de drogas por vía intravenosa y el contacto heterosexual como factores importantes para la predicción del número de muertes por SIDA en Puerto Rico.

Este trabajo de investigación se realizó con el propósito de estudiar el comportamiento del VIH/SIDA en Puerto Rico usando modelos epidemiológicos. Se plantea un modelo matemático determinístico basado en un sistema de ecuaciones diferenciales ordinarias para evaluar los factores que afectan la prevalencia del virus en Puerto Rico. Este modelo está diseñado para evaluar el efecto epidemiológico de la transmisión heterosexual y la transmisión entre usuarios de droga intravenosa (UDI), tomando en cuenta parámetros de conductas de riesgo tales como: la tasa de intercambio de inyectables por año (en mujeres UDI y en hombres UDI), la tasa de cambio de parejas sexuales por año (en mujeres que son UDI sexualmente activas y en hombres UDI sexualmente activos), y la tasa de cambio de parejas sexuales por año (en mujeres No-UDI y en hombres No-UDI).

Además se consideran parámetros epidemiológicos como: la eficiencia de transmisión del VIH por inyectable dado que el inyectable está infectado, eficiencia de transmisión heterosexual por contacto sexual dado que la pareja está infectada, el tiempo promedio de incubación del virus y el tiempo promedio de sobrevivencia en adultos. Usando el modelo estimamos el número de casos acumulados de SIDA en adultos en Puerto Rico, para los próximos 10 años e identificamos los parámetros que contribuyen más a la incertidumbre de estos estimados usando análisis de incertidumbre y análisis de sensibilidad.

Estos métodos requieren estimados de los parámetros para Puerto Rico y su desempeño depende de la eficiencia con la que se explore el espacio muestral de estos parámetros. Varios métodos han sido utilizados con este fin (Seaholm et al. 1988, Blower and Dowlatabadi 1994). Compararemos la eficiencia de éstos y uno nuevo que proponemos en nuestro modelo de Puerto Rico.

1.3. Antecedentes

Hernández Maldonado M. (1990), Varas Díaz N. (2002) y Resto Olivo J. (2003) investigó la enfermedad del SIDA en Puerto Rico desde el punto de vista educativo,

social y filosófico, respectivamente. Hernández Maldonado M. (1990) evalúa la efectividad de la guía de enseñanza del VIH/SIDA "Learning to be Healthy" en niños de segundo y tercer grado en el municipio de Barceloneta, donde el 87% de los niños reconoció que el SIDA no se transmite si una persona que lo tiene tose o estornuda. Varas Díaz, N. (2002) estudió la epidemia de estigma, que consiste en la interpretaciones negativas que se hacen de las personas seropositivas y de la enfermedad como tal. Resto Olivo J. (2003) investigó dos estrategias de recaudación de fondos para una organización sin fines de lucro que ofrece servicios en el área de VIH/SIDA, mediante solicitudes de donativos por correo a individuos y solicitudes de donativos a fundaciones filantrópicas.

Desde la perspectiva estadística Torres Velásquez, B. (2009) diseñó un modelo de regresión múltiple para predecir el número de muertes por SIDA en el tiempo, que utiliza como variables predictivas los tipos de transmisión del SIDA en Puerto Rico: usuarios de droga intravenosa, contacto heterosexual, contacto homosexual, desórdenes de coagulación, receptores de sangre, componentes de tejido, exposición perinatal, casos de VIH en menores de 13 años y casos pediátricos. Mediante el método de regresión escalonada hacia adelante se obtuvo que la variable más importante para predecir el número de muertes de SIDA es el uso de droga vía intravenosa. En cuanto a la población de mujeres, las variables predictivas más

importantes que explican el número de muertes por SIDA fueron el contacto heterosexual y la exposición perinatal. En el caso de los hombres, las variables que mejor explican el número de muertes por SIDA fueron el uso de droga intravenosa y el contacto heterosexual.

Utilizando modelos matemáticos de ecuaciones diferenciales ordinarias, Blower et al. (1991) y Blower and Dowlatabadi (1994) estudiaron la dinámica del VIH/SIDA en la ciudad de Nueva York. El modelo consistió de 34 ecuaciones diferenciales ordinarias, de las cuales 15 ecuaciones corresponden a la población de mujeres, 15 ecuaciones a la población de hombres y 4 ecuaciones a la población de niños. Las 15 ecuaciones correspondientes a la población de mujeres modelan la dinámica de transmisión de tres grupos: mujeres susceptibles, mujeres con VIH y mujeres con SIDA con 5 ecuaciones para cada grupo. La estructura de las 5 ecuaciones para la población susceptible de mujeres presenta los siguientes grupos de riesgo: usuarias de droga que comparten jeringuillas infectadas con extraños, usuarias de droga que comparten jeringuillas infectadas con amigos, usuarias de droga y sexualmente activas que comparten jeringuillas infectadas con extraños, usuarias de droga y sexualmente activas que comparten jeringuillas infectadas con amigos, mujeres que no usan droga mediante jeringuilla (contacto heterosexual). Las ecuaciones diferenciales ordinarias para la población de hombres tienen la

misma estructura que la población de mujeres. Las ecuaciones diferenciales para la población de niños representan los niños que nacen infectados de madres usuarias de droga por jeringuilla y de madres que no son usuarias de droga mediante jeringuilla, y la población de niños que nacen con SIDA de madres que son usuarias de droga intravenosa y de madres que no son usuarias de droga intravenosa.

Usando el modelo estimaron el número de casos acumulados de SIDA en adultos y el número de casos acumulados de SIDA en niños para 30 años. La variabilidad en los resultados se exploró usando análisis de sensibilidad y análisis de incertidumbre. Para esto utilizaron el método de Muestreo por Hipercubo Latino (Stein, M. 1987) para la exploración del espacio muestral de los parámetros que entraron al modelo. Después de definir una función de distribución para cada parámetro aplicaron el muestreo por Hipercubo Latino para generar una matriz de valores iniciales. Para cada conjunto de valores iniciales de los parámetros se obtuvo una solución para el modelo. El método supone que los parámetros de entrada son estadísticamente independientes, ya que el muestreo por Hipercubo Latino usa distribuciones marginales.

El análisis de incertidumbre se hizo usando estadística descriptiva (media, varianza, desviación estándar, histogramas) de las variables respuesta (número

de casos de SIDA en adultos y número de casos pediátricos). Para el análisis de sensibilidad se calculó el coeficiente de correlación parcial (CCP) entre los parámetros de entrada y las variables respuesta . En el análisis de sensibilidad e incertidumbre de los resultados estimados por Blower et al. (1991) se identificó que las variables que afectan la imprecisión en la estimación del número de casos acumulados de SIDA en adultos fue la eficiencia de transmisión heterosexual tanto en hombres y mujeres, el tiempo promedio de incubación del virus, la eficiencia de transmisión del VIH por compartir jeringuillas infectadas, la tasa de cambio de parejas sexuales por año (mujeres No-UDI) y la tasa de cambio de parejas sexuales por año (hombres No-UDI).

1.4. Objetivos

1.4.1. Objetivo General

- Construir un modelo similar al de Blower et al. (1991) para la población VIH/SIDA en Puerto Rico y evaluar los factores que afectan la prevalencia del virus en Puerto Rico.

1.4.2. Objetivos Específicos

- Adquirir datos actualizados sobre la población de VIH/SIDA en Puerto Rico a través del Departamento de Salud de Puerto Rico y otras fuentes.
- Estimar puntualmente los parámetros iniciales de entrada al modelo, utilizando la base de datos recopilada.
- Definir funciones de distribución para los distintos parámetros que entran al modelo.
- Modificar el método de Muestreo por Hipercubo Latino usando la secuencia de Halton para mejorar su eficiencia.
- Generar la matriz de valores iniciales de los parámetros usando los métodos de muestreo Aleatorio, Hipercubo Latino e Hipercubo Latino Modificado.
- Comparar la eficacia de los métodos de muestreo Aleatorio, Hipercubo Latino e Hipercubo Latino Modificado en la exploración del espacio muestral de los parámetros que entran al modelo.
- Estudiar la distribución del número de casos acumulados de SIDA en adultos de Puerto Rico obtenidos por el modelo e identificar los parámetros que

contribuyen más a la incertidumbre en la estimación mediante el Análisis de Incertidumbre y Análisis de Sensibilidad usando los métodos de muestreo Aleatorio, Hipercubo Latino e Hipercubo Latino modificado para la exploración del espacio muestral de los parámetros.

Capítulo 2

REVISIÓN LITERARIA

2.1. Modelos Epidemiológicos

La construcción de modelos matemáticos es una de las herramientas utilizadas para estudiar problemas en medicina, biología, fisiología, bioquímica, epidemiología, entre otras áreas del conocimiento. Sus objetivos primordiales son describir, explicar y predecir fenómenos y procesos en dichas áreas. Los modelos matemáticos para enfermedades infecciosas se utilizan como herramienta para mejorar la comprensión de un sistema, para prevenir futuras situaciones de enfermedades, determinar la prevalencia e incidencia y contribuir a tomar decisiones objetivas para controlar o erradicar las enfermedades. En los modelos epidemiológicos se parte del supuesto de que los individuos se encuentran en uno de varios estados epidemiológicos posibles. En función de dichos estados, la población puede dividirse en grupos como por ejemplo: individuos susceptibles (S), infectados (I) o removi-

dos (R), entre otros que dependerán del problema de interés. La dinámica entre estos grupos se modela usando los modelos epidemiológicos. A continuación describimos tres de los modelos más utilizados.

Modelos principales de transmisión de enfermedades infecciosas

La dinámica de la transmisión de las enfermedades infecciosas se representa a través de los principales modelos epidemiológicos: Susceptible-infecciosa (SI), susceptible-infeccioso-susceptible (SIS) y susceptible-infeccioso -removido (SIR). El modelo SI (por ejemplo un brote sencillo sin nacimientos ni muertes), consiste en un sistema de dos ecuaciones diferenciales:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta I(t)S(t)}{N}, \\ \frac{dI}{dt} &= \frac{\beta I(t)S(t)}{N}\end{aligned}\tag{2.1}$$

donde $N = S + I$.

La ecuación presenta dos variables dependientes: el número de personas susceptibles (S) y el número de personas infectadas (I), donde $\frac{dS}{dt}$ es la razón de cambio de individuos susceptibles y $\frac{dI}{dt}$ la razón de cambio de los infectados en la población. Por ende las soluciones de estas ecuaciones no-lineales nos darán el número de susceptibles ($S(t)$) e infectados ($I(t)$) en tiempo t , respectivamente. Los contactos

entre los susceptibles e infectados se pueden modelar por acción de masa SI o como se representa en el sistema 2.1 por incidencia estándar $\frac{SI}{N}$. El parámetro β es la tasa de transmisión de la enfermedad el cual es el producto del número de contactos de un individuo por unidad de tiempo y la probabilidad de que el contacto sea efectivo, o sea que transmita la enfermedad. La razón de cambio de la población de individuos infectados ($\frac{dI}{dt}$) es igual al número de individuos recientemente infectados (Montesinos y Hernández, 2007).

El modelo SIS puede formularse como un sistema de dos ecuaciones diferenciales, como se ilustra a continuación:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta I(t)S(t)}{N} + \mu I(t), \\ \frac{dI}{dt} &= \frac{\beta I(t)S(t)}{N} - \mu I(t)\end{aligned}\tag{2.2}$$

donde $N = S + I$.

La ecuación del modelo SIS difiere del modelo SI porque se agrega el término $\mu I(t)$ en la ecuación, donde μ representa la tasa de recuperación de un individuo (Montesinos y Hernández, 2007).

Por último el modelo SIR , en su forma más simple (brote sencillo) puede formularse como un conjunto de ecuaciones diferenciales, tal y como se muestra

a continuación:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta I(t)S(t)}{N}, \\ \frac{dI}{dt} &= \frac{\beta I(t)S(t)}{N} - \mu I(t), \\ \frac{dR}{dt} &= \mu I(t)\end{aligned}\tag{2.3}$$

donde $N = S + I + R$.

A diferencia del sistema 2.2 aquí se añade un elemento adicional para cuantificar el número de recuperados por unidad de tiempo. Por ende la razón de cambio del número de individuos removidos (recuperados) $\frac{dR}{dt}$ es proporcional al número de individuos que se recuperan $\mu I(t)$.

La mejor manera de modelar algunas enfermedades infantiles consiste en emplear un modelo SIR puesto que la infección lleva a una inmunidad vitalicia. Para la mayor parte de las enfermedades de transmisión sexual (*ETS*) resulta más útil el modelo *SIS*, toda vez que tan sólo un número reducido de *ETS* confiere inmunidad tras la infección. El VIH es una excepción clara y todavía puede describirse de forma adecuada, al menos en el mundo occidental, mediante el modelo *SI* (De Pereda, 2010) y (Montesinos y Hernández, 2007).

2.2. Análisis de Incertidumbre

El análisis de incertidumbre de un modelo tiene por objeto proporcionar mediciones cuantitativas de la incertidumbre que tienen los valores finales (en adelante variables respuesta) del modelo como consecuencia de la incertidumbre en los valores iniciales de los parámetros de entrada. Cada valor de los parámetros de entrada genera un valor de la variable respuesta, por tanto la incertidumbre en la variable respuesta depende de la incertidumbre de los valores de los parámetros de entrada. En el análisis de incertidumbre se examinan las variables respuesta del modelo mientras se varían los valores de los parámetros de entrada. Formalmente, sea k el número de parámetros de entrada al modelo. Se busca generar n conjuntos de parámetros iniciales denotados por $X_i = [x_{i1}, x_{i2}, \dots, x_{ik}]$, $i = 1, 2, \dots, n$. Cada conjunto ($i = 1, 2, \dots, n$) se genera usando la distribución conjunta $f(x_{i1}, \dots, x_{ik})$. En la mayoría de los análisis es difícil especificar esta distribución. Blower et al. (1991) supone la independencia de los parámetros e identifica una función de distribución para cada parámetro de entrada, f_j , $j = 1, 2, \dots, k$, por lo que la distribución conjunta es $\prod_{j=1}^k f_j(x_{ij})$. Cada distribución refleja el grado de creencia de la ubicación apropiada de los valores de los parámetros de entrada y su forma debe tener una justificación biológica.

Luego se genera un conjunto de valores para los parámetros de entrada a partir de la distribución conjunta, es decir se obtiene $X_i = [x_{i1}, x_{i2}, \dots, x_{ik}]$, $i = 1, 2, \dots, n$. Cada valor inicial X_i proporciona valores para las variables respuesta del modelo $Y_i = [y_1(X_i), y_2(X_i), \dots, y_m(X_i)]$, donde m es el número de ecuaciones en el modelo. Se definió la matriz $X = [x_1, x_2, \dots, x_k]$ como la matriz $n \times k$ que contiene los valores iniciales de cada parámetro de entrada, donde $x_j = [x_{1j}, x_{2j}, \dots, x_{nj}]'$, $j = 1, 2, \dots, k$ representa los n valores del j -ésimo parámetro de entrada del modelo y $Y = [y_1, y_2, \dots, y_m]$ como la matriz $n \times m$ que contiene las variables respuesta del modelo donde $y_l = [y_l(X_1), y_l(X_2), \dots, y_l(X_n)]'$, con $l = 1, 2, \dots, m$ representa los n valores de la l -ésima variable respuesta.

Luego de realizar las n simulaciones se estudia la incertidumbre en los valores finales de las ecuaciones usando estadísticos descriptivos tales como media, mediana, varianza, máximo, mínimo, percentiles, e histogramas de los y_l , donde $l = 1, 2, \dots, m$ (Helton et al., 2006).

2.3. Análisis de Sensibilidad

El análisis de sensibilidad es una extensión del análisis de incertidumbre donde se busca identificar los parámetros de entrada que contribuyen más (debido a la incertidumbre de su estimación) a la imprecisión de la predicción de las variables respuesta en el modelo. Para identificar estos parámetros importantes se utilizan métodos como Diagramas de Dispersión, Correlación y Correlación Parcial que se describen a continuación (Helton et al., 2006).

Diagrama de Dispersión

El diagrama de dispersión es el método más simple para el análisis de sensibilidad ya que sólo implica graficar cada vector x_j con cada vector y_l , que sean de interés. El diagrama de dispersión nos permite estudiar la relación entre las variables bajo estudio. Revela la relación lineal o no lineal, los picos y la interacción de variables que facilitan el entendimiento del comportamiento del modelo (Helton et al., 2006).

Correlación y correlación parcial

El coeficiente de correlación (CC) y el coeficiente de correlación parcial (CCP) establecen una medida del grado de relación lineal entre la variable respuesta y la variable predictiva (en nuestro caso parámetros de entrada). Sea el modelo de regresión lineal múltiple

$$y_l = \mathbf{X}\boldsymbol{\beta} + \varepsilon,$$

donde

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{pmatrix},$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix},$$

y

$$\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

aquí,

y_l : Es un vector columna de dimensión $n \times 1$.

\mathbf{X} : Es una matriz $n \times (k + 1)$.

$\boldsymbol{\beta}$: Es el vector de coeficientes de regresión de dimensión $(k + 1)$.

$\boldsymbol{\varepsilon}$: Es un vector aleatorio de dimensión $n \times 1$, con $\varepsilon_i \sim N(0, \sigma^2)$.

El coeficiente de correlación r_{x_j, y_l} mide el grado de relación lineal entre la variable respuesta y_l cuando ésta depende de un solo parámetro de entrada. Supongamos que ese parámetro de entrada es x_j , el modelo de regresión para este caso es:

$$y_l = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

y el coeficiente de correlación viene dado por:

$$r_{x_j, y_l} = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(y_l(X_i) - \bar{y}_l)}{\sqrt{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2} \sqrt{\sum_{i=1}^n (y_l(X_i) - \bar{y}_l)^2}},$$

donde

$$\mathbf{X} = \begin{pmatrix} 1 & x_{1j} \\ 1 & x_{2j} \\ \vdots & \vdots \\ 1 & x_{nj} \end{pmatrix}$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_j \end{pmatrix}$$

$$\bar{y}_l = \frac{\sum_{i=1}^n y_l(X_i)}{n},$$

y

$$\bar{x}_j = \frac{\sum_{i=1}^n x_{ij}}{n}.$$

Cuando hay más de un parámetro de entrada se considera el coeficiente de correlación parcial porque establece la medida del grado de relación lineal entre la variable respuesta y_l y cada una de los parámetros de entrada x_j , manteniendo constante los demás parámetros de entrada. El coeficiente de correlación parcial (CCP) entre x_j y y_l se obtiene usando los siguientes modelos de regresión, donde \widehat{x}_j es el valor estimado del j -ésimo parámetro de entrada y \widehat{y}_l es el valor estimado

de la l -ésima variable respuesta.

$$\widehat{x}_j = \mathbf{X}_{(-j)}\mathbf{C}_{(-j)}$$

$$\widehat{y}_l = \mathbf{X}_{(-j)}\mathbf{B}_{(-j)}$$

$$\mathbf{C} = [c_0 \ c_1 \ \dots \ c_k]'$$

$$\mathbf{B} = [b_0 \ b_1 \ \dots \ b_k]'$$

$\mathbf{X}_{(-j)}$: Es una matriz $n \times (k + 1)$ que no contiene el j -ésimo parámetro de entrada.

$\mathbf{C}_{(-j)}$: Es el vector que no contiene el valor del coeficiente de regresión del j -ésimo parámetro de entrada.

$\mathbf{B}_{(-j)}$: Es el vector que no contiene el valor del coeficiente de regresión del j -ésimo parámetro de entrada.

Entonces definimos nuevas variables $(x_{ij} - \widehat{x}_j)$ y $(y_l - \widehat{y}_l)$. El coeficiente de correlación parcial (CCP) δ_{x_j, y_l} entre x_j y y_l es el coeficiente de correlación entre $(x_{ij} - \widehat{x}_j)$ y $(y_l - \widehat{y}_l)$.

Así el CCP mide el grado de asociación lineal entre las nuevas variables definidas con el efecto lineal de las otras variables eliminadas.

El CCP caracteriza la estrecha relación entre dos variables después de haber realizado la corrección para el efecto lineal de las otras variables en el análisis. En particular, el CCP proporciona la medida de aquella variable importante (parámetro de entrada) que tiende a excluir los efectos de los otros parámetros de

entrada (Helton & Davis, 2000)

2.4. Exploración del espacio muestral de los parámetros que entran al modelo

Para estudiar la variación del número acumulado de SIDA en adultos en Puerto Rico obtenidos por el modelo a través del análisis de incertidumbre y análisis de sensibilidad se simularon valores iniciales de los parámetros que entran en el modelo usando la distribución conjunta $f(x_{i1}, \dots, x_{ik})$ que por la hipótesis de independencia es $\prod_{j=1}^k f(x_{ij})$. Para simular estos valores iniciales se utilizó muestreo aleatorio, muestreo por hipercubo latino y una modificación del muestreo por hipercubo latino. A continuación se describen los dos primeros métodos, sin embargo, el muestreo por hipercubo latino modificado se describe en la Sección 4.2. Aquí se describe la secuencia de Halton que usamos como base para modificar el muestreo por hipercubo latino.

2.4.1. Muestreo Aleatorio

La muestra aleatoria de los parámetros de entrada se genera muestreando independientemente cada parámetro x_{ij} ($j = 1, \dots, k$), usando $f_j(x_{ij})$ ($j = 1, \dots, k$), respectivamente. Casi todos los programas estadísticos tienen funciones predefinidas para generar muestras aleatorias de las distribuciones más conocidas. En nuestro caso usamos Matlab.

2.4.2. Muestreo por Hipercubo Latino

En este método el rango de cada parámetro de entrada se divide en n intervalos con igual probabilidad $\frac{1}{n}$ de acuerdo a su función de distribución. Para cada parámetro y en cada intervalo se selecciona al azar un valor inicial para obtener $n \times k$ valores iniciales. Los n valores iniciales del primer parámetro son asociados en forma aleatoria y sin reemplazo con los n valores iniciales del segundo parámetro. Los n valores iniciales del tercer parámetro se asignan de la misma forma, continuando el proceso con los demás parámetros hasta formar n vectores fila de orden k , es decir la matriz X de $n \times k$ valores iniciales.

Su ventaja principal es que el número de simulaciones n no crece exponen-

cialmente con el número de parámetros. Muchos autores, entre ellos Blower et al. (1991) lo utilizan porque aparentemente cubre bien el espacio muestral de los parámetros con menos simulaciones, si comparamos con el método aleatorio que requiere un gran número de simulaciones, particularmente cuando hay muchos parámetros. Nos parece que una desventaja del muestreo por hipercubo latino está en la formación de los vectores. El hecho de que la asociación se haga sin reemplazo elimina combinaciones de valores iniciales que deberían ser examinadas. Por ejemplo, si consideremos el caso de dos parámetros, ambos, con distribución uniforme $(0, 1)$ y $n = 5$. El método de hipercubo latino podría escoger el par de valores iniciales (w_1, v_1) , donde $w_1 > 0.8$. Como el muestreo es sin reemplazo un valor de la magnitud de $w_1 > 0.8$ ya no sería considerado como valor inicial con algún otro valor del parámetro v_1 (Helton & Davis, 2000) y (Mckay et al., 1979).

2.4.3. Secuencia de Halton

Las secuencias de Halton son generalizaciones en cualquier dimensión $s \geq 1$ de las secuencias de Van der Corput. La idea consiste en considerar la función radical inversa simultáneamente en bases diferentes. Estas bases son un conjunto de números $b_1 \dots b_s$ primos entre sí. El n -ésimo término de la secuencia de Van

Der Corput en base b , denotado por $\phi_b(n)$, está definido como sigue: primero, se escribe n en base b :

$$n = (a_k \dots a_1 a_0)_b = a_0 + a_1 b + \dots + a_k b^k, \quad (2.4)$$

y luego se calcula

$$\phi_b(n) = (0.a_0 a_1 \dots a_k)_b = \frac{a_0}{b} + \frac{a_1}{b^2} + \dots + \frac{a_k}{b^{k+1}}. \quad (2.5)$$

La secuencia de Halton (Halton 1960) en las bases $b_1 \dots b_s$ es $(\phi_{b_1}(n), \dots, \phi_{b_s}(n))_{n=1}^{\infty}$, Halton demostró que sus secuencias son de baja discrepancia, es decir tienen la propiedad de cubrir un espacio de dimensión s de una manera más uniforme que una secuencia de puntos aleatorios. Esta propiedad es la que nos lleva a crear un nuevo método para la exploración del espacio de los parámetros al que llamamos Hipercubo Latino modificado que describiremos en la sección 4.2. Las secuencias de Halton utilizan un número primo por cada dimensión del espacio a muestrear. En la práctica se utiliza la lista creciente de números primos (2 para la primera dimensión, 3 para la segunda, 5 para la tercera, etc.). Una aplicación de la secuencia de Halton es la integración numérica (Okten, 2010).

Un problema conocido de la secuencia de Halton es la correlación entre las secuencias a medida que se aumentan las bases. Para remediar el problema se han propuesto permutaciones particulares de los dígitos de las secuencias (Tuffin,

1996) y (Okten et al., 2009) para los cuales se rompe la correlación manteniendo la propiedad de tener secuencias de baja discrepancia. Okten et al. (2009) propone permutar los dígitos de las secuencias aleatoriamente y demuestra con simulaciones que estas permutaciones rompen la correlación, pero todavía es un problema abierto el demostrar que son secuencias de baja discrepancia. Con una permutación aleatoria P de los dígitos $0, \dots, b_i - 1$, la ecuación 2.5 se convierte en

$$\phi_{b_i}(n) = \frac{P_i(a_0)}{b_i} + \frac{P_i(a_1)}{b_i^2} + \dots + \frac{P_i(a_k)}{b_i^{k+1}}. \quad (2.6)$$

La Figura 2.1 muestra al lado izquierdo la gráfica de la secuencia de Halton original de las bases 11 vs 13 y al lado derecho la gráfica de las secuencias aleatorizadas de las bases 11 vs 13. Se observa que las secuencias de Halton aleatorizadas rompen la correlación.

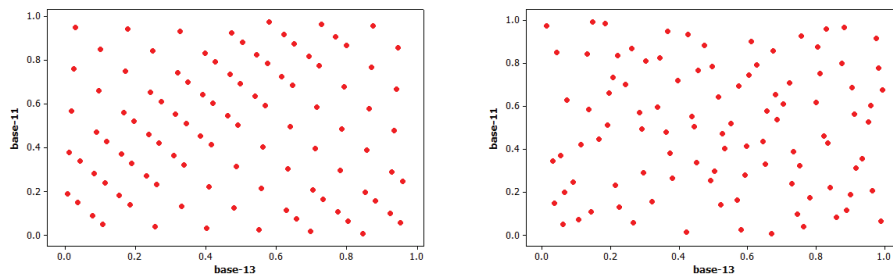


Figura 2.1: Secuencia de Halton y la permutación aleatoria de la secuencia de Halton

Capítulo 3

Modelo matemático de transmisión del VIH/SIDA en Puerto Rico

El modelo que usamos para la transmisión del VIH/SIDA en Puerto Rico es el modelo matemático determinístico propuesto por Blower et al. (1991), con algunas diferencias, debido a la disponibilidad de datos para Puerto Rico. El modelo formulado por Blower et al. (1991) estratifica la comunidad usuaria de droga mediante jeringuilla de la ciudad de Nueva York (población susceptible) en dos grupos, mujeres y hombres. Cada uno de éstos los divide en 5 subgrupos, de acuerdo a su conducta de riesgo. Los subgrupos para hombres y mujeres son los siguientes: usuarios de droga que comparten jeringuilla con extraños, usuarios de droga que comparten jeringuilla con amigos, usuarios de droga que comparten jeringuilla con extraños y son sexualmente activas, usuarios de droga que comparten jeringuilla con amigos y son sexualmente activas y no-usuarios de droga mediante jeringuilla

llas. En consecuencia Blower et al. (1991) tiene 10 subgrupos, que comprenden 20 parámetros de entrada. Para Puerto Rico, la población susceptible también es la población usuaria de droga mediante jeringuilla, que se estima en 45,294 usuarios de droga (consumo de Heroína). De éstos 9,470 son mujeres y 35,824 son hombres con razón de masculinidad de 4 hombres por cada mujer, según la Administración de Servicios de Salud Mental y Contra la Adicción (ASSMCA), según estudio realizado en el año 2009 sobre "Transtornos de sustancias y uso de servicios en Puerto Rico". Cada grupo (mujeres y hombres) se divide en 3 subgrupos de acuerdo a la conducta de riesgo: usuarios de droga por inyectable (UDI)(únicamente un factor de riesgo), usuarios de droga por inyectable sexualmente activos (UDISA)(dos factores de riesgo) y los que no usan droga mediante inyectable (No-UDI) pero que son sexualmente activos (No-UDI)SA. Tenemos entonces 6 subgrupos, que incorporan 11 parámetros de entrada. El modelo de Blower et al. (1991) incluye supuestos biológicos que restringen algunos parámetros, por ejemplo, la eficiencia de transmisión heterosexual (durante el contacto heterosexual) que se asume igual para hombres y mujeres. En nuestro modelo también consideramos esta restricción, por lo tanto tenemos 10 parámetros en total. Se describe cada uno de éstos parámetros en la Sección 3.1.

Se desconoce el tamaño inicial de los 4 estratos de los usuarios de droga

mediante jeringuilla por lo que utilizamos el método presentado por Blower y Dowlatabadi (1994), donde el tamaño inicial de los subgrupos de usuarios de droga por jeringuilla se genera aleatoriamente con funciones de probabilidad independientes. El total de mujeres UDI es de 9,470, se generó la variable aleatoria P_1 (uniforme entre 0 y 1) y se tomó $P_1 \times 9,470$ como la población de mujeres UDI y $(1 - P_1) \times 9,470$ como la población de mujeres UDISA. Hacemos lo mismo con los hombres, pero con P_2 y el total de 35,824 (ver Figura 3.1). El tamaño inicial de los dos estratos que conciernen a hombres y mujeres no-usuarios de droga mediante jeringuilla (que son las parejas sexuales de hombres y mujeres usuarios de droga por jeringuilla) se determina escogiendo M , como el tamaño de la población de mujeres, aleatoriamente de una distribución uniforme sobre el rango de 17,912 y 35,824. Para calcular este rango se utiliza la razón de sexo en la población total de UDI para Puerto Rico que es 4:1 (hombres: mujeres). El tamaño de la población de hombres (H) se determina con la ecuación: $H = 45,294 - M$. El análisis de incertidumbre y el análisis de sensibilidad requieren simulaciones por lo que para cada simulación, el tamaño de la población de UDI en el tiempo inicial (tiempo cero) es 45,294, el tamaño de los 4 subgrupos de riesgo de UDI es variado para cada corrida. Del mismo modo el tamaño de la población (No-UDI)SA en el tiempo inicial es 45,294 (Blower et al., 1991) y el tamaño de los subgrupos de H y M es

variado para cada corrida.

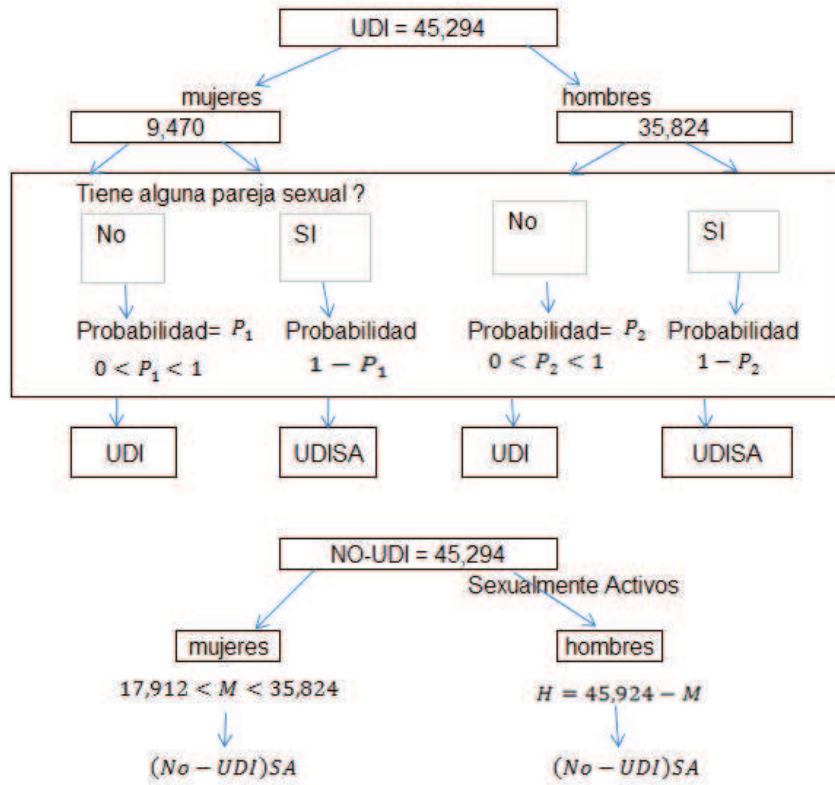


Figura 3.1: Tamaño inicial de los estratos del modelo para Puerto Rico

El modelo está diseñado para reflejar la dinámica de transmisión entre mujeres y hombres UDI, mujeres y hombres UDISA y mujeres y hombres (No-UDI)SA. El modelo consiste de 18 ecuaciones diferenciales ordinarias no-lineales, con 10 parámetros. Estas ecuaciones definen la transmisión del virus dentro y entre los

6 subgrupos de riesgo: 4 subgrupos de usuarios de droga por jeringuilla y dos subgrupos no-usuarios de droga mediante jeringuilla.

Usamos el modelo para explicar el patrón de seroprevalencia en la población UDI, para explorar el efecto del riesgo de transmisión heterosexual sobre el riesgo de infección del VIH en usuarios de droga por jeringuilla, para predecir el número de casos acumulados de SIDA en adultos en Puerto Rico para los próximos 10 años, desde el año 2010, para evaluar la variabilidad de la predicción, y para identificar las variables importantes que contribuyen a la imprecisión de la predicción.

Matriz de Contacto Sexual

Blower et al. (1991) definen 4 formas diferentes en que los individuos de un subgrupo particular eligen pareja sexual entre los individuos de cualquier otro subgrupo del sexo opuesto. La relación entre los grupos se puede definir usando matrices de contacto sexual. La matriz de contacto sexual está basada en una suposición particular, de cómo los diferentes subgrupos seleccionan parejas sexuales. Uno de estos modelos es la "selección perfecta negativa de pareja sexual" donde los individuos eligen pareja del sexo opuesto únicamente en grupos diferentes al

que ellos pertenecen (por ejemplo, un usuario de droga sólo elige parejas que no usan drogas). El otro es la "selección perfecta positiva de pareja sexual" donde los individuos buscan parejas sexuales exclusivamente con su misma característica (por ejemplo, un usuario de droga sólo elige parejas que también usan drogas). Así una matriz de contacto sexual de selección perfecta positiva de pareja sexual es equivalente a la matriz identidad, los coeficientes de la diagonal son unos y los demás coeficientes son ceros. Blower et al. (1991) utilizó esta última como la forma de selección de parejas. Otro modelo, que es el que usamos nosotros para Puerto Rico es "la selección proporcional de pareja sexual" donde los individuos eligen parejas del sexo opuesto en proporción a la disponibilidad de individuos del sexo opuesto que están sexualmente activas (Ríos y Castillo, 2011). En nuestro caso la matriz de contacto sexual para mujeres se define como:

$$M_m = \begin{bmatrix} \text{prob mujer UDISA elija} & \text{prob mujer UDISA elija} \\ \text{hombre UDISA} & \text{hombre (No-UDI)SA} \\ \text{prob mujer (No-UDI)SA elija} & \text{prob mujer (No-UDI)UDISA elija} \\ \text{hombre UDISA} & \text{hombre (No-UDI)SA} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{N_{h2}(t)}{N_{h2}(t)+N_{h3}(t)} & \frac{N_{h3}(t)}{N_{h2}(t)+N_{h3}(t)} \\ \frac{N_{h2}(t)}{N_{h2}(t)+N_{h3}(t)} & \frac{N_{h3}(t)}{N_{h2}(t)+N_{h3}(t)} \end{bmatrix}$$

donde:

$N_{h2}(t)$ es el total de hombres UDISA en el tiempo t

$N_{h3}(t)$ es el total de hombres (No-UDI)SA en el tiempo t

Similarmente, la matriz de contacto sexual para hombres se define como:

$$M_h = \begin{bmatrix} \text{prob hombre UDISA elija} & \text{prob hombre UDISA elija} \\ \text{mujer UDISA} & \text{mujer (No-UDI)SA} \\ \text{prob hombre (No-UDI)SA elija} & \text{prob hombre (No-UDI)UDISA elija} \\ \text{mujer UDISA} & \text{mujer (No-UDI)SA} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{N_{m2}(t)}{N_{m2}(t)+N_{m3}(t)} & \frac{N_{m3}(t)}{N_{m2}(t)+N_{m3}(t)} \\ \frac{N_{m2}(t)}{N_{m2}(t)+N_{m3}(t)} & \frac{N_{m3}(t)}{N_{m2}(t)+N_{m3}(t)} \end{bmatrix}$$

donde:

$N_{m2}(t)$ es el total de mujeres UDISA en el tiempo t

$N_{m3}(t)$ es el total de mujeres (No-UDI)SA en el tiempo t

Las matrices de contacto sexual deben cumplir la siguientes restricciones:

- i) Los elementos de las matrices M_m y M_h deben estar entre 0 y 1.
- ii) La suma de las filas de las matrices de contacto sexual M_m y M_h debe ser igual a 1.

3.1. Ecuaciones del modelo para Puerto Rico

Como se explicó en la sección anterior el modelo propuesto está basado en el modelo de Blower et al. (1991). A continuación se presenta la dinámica de transmisión del modelo en forma gráfica. La población se estratifica de acuerdo a su estado epidemiológico: susceptibles (S), infectados (I) y SIDA (A). Además, se estratifica por género: mujeres (m) y hombres (h) y a su vez por su conducta de riesgo: UDI, UDISA y (NO-UDI)SA. En la Figura 3.2 se muestra un diagrama de flujo que representa la dinámica de transmisión para la población de mujeres en el modelo que propuesto (la dinámica es análoga para la población de hombres).

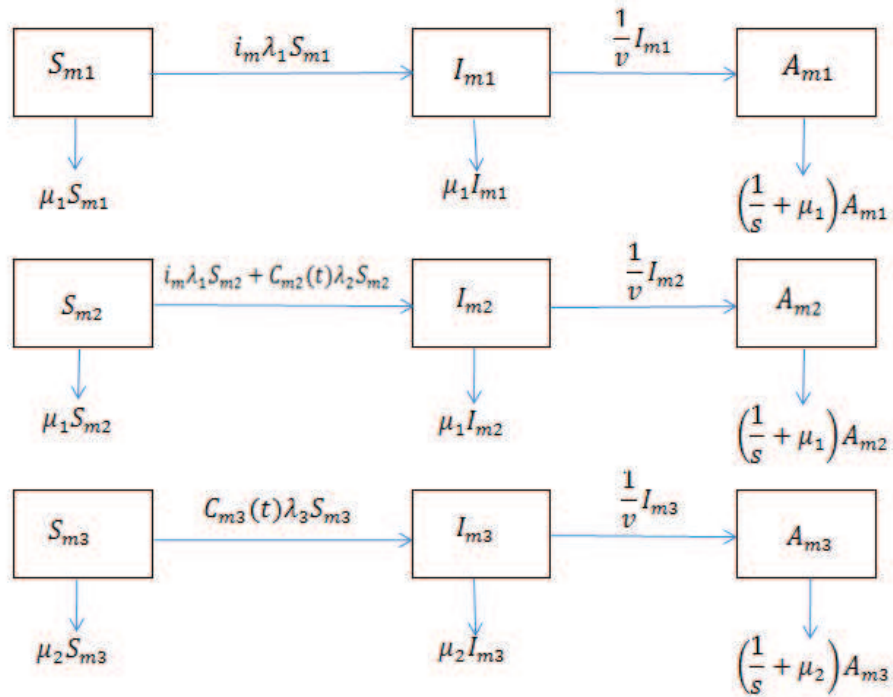


Figura 3.2: Diagrama de flujo del modelo para Puerto Rico

En el cuadro 3.1 presentamos un resumen de los parámetros de nuestro modelo.

Cuadro 3.1: Definición de los parámetros para el modelo

β_1	Eficiencia de transmisión del VIH por inyectable (inyectable infectado)
β_{mh}^*	Eficiencia de transmisión heterosexual por asociación (contacto heterosexual) de mujer a hombre dado que la mujer está infectada
β_{hm}^*	Eficiencia de transmisión heterosexual por asociación (contacto heterosexual) de hombre a mujer dado que el hombre está infectado
$C_{m2}(t)$	Tasa de cambio de parejas sexuales por año (mujer UDI) al tiempo t
$C_{m3}(t)$	Tasa de cambio de parejas sexuales por año (mujer No-UDI) al tiempo t
$K_{h2}(t)$	Tasa de cambio de parejas sexuales por año (hombres UDI) al tiempo t
$K_{h3}(t)$	Tasa de cambio de parejas sexuales por año (hombres No-UDI) al tiempo t
i_m	Tasa de intercambio de inyectables por año (mujeres UDI)
i_h	Tasa de intercambio de inyectables por año (hombres UDI)
v	Tiempo promedio de incubación del virus (años)
s	Tiempo promedio de sobrevivencia en adultos (años)
λ_1	Probabilidad de tener VIH por compartir inyectables por mujer UDI
λ_2	Probabilidad de tener VIH por contacto heterosexual por mujer UDISA
λ_3	Probabilidad de tener VIH por contacto heterosexual por mujer (No-UDI)SA
μ_1	Tasa de mortalidad natural en hombres y mujeres UDI
μ_2	Tasa de mortalidad natural en hombres y mujeres (No-UDI)SA

* $\beta_{mh} = \beta_{hm}$ Blower et al.(1991)

Las mujeres susceptibles UDI (m_1), UDISA (m_2), (N0-UDI)SA (m_3) pasan a la categoría de infectadas (I) con el virus del VIH de acuerdo a los términos $i_m\lambda_1$, $(i_m\lambda_1 + C_{m2}(t)\lambda_2)$ y $C_{m3}(t)\lambda_3$ (que explicamos más adelante), respectivamente, o mueren por razones ajenas al VIH (causas naturales) de acuerdo a los parámetros μ_1 o μ_2 . De modo similar las infectadas(os) pasan a la categoría de enfermas(os) con SIDA por la tasa $\frac{1}{s}$ o mueren por razones ajenas al SIDA, por la tasa μ_2 , mientras que $\frac{1}{s}$ representa las muertes por SIDA.

Las siguientes ecuaciones describen el modelo propuesto. La razón de cambio del tamaño de la población susceptible de mujeres (S_{m1}) y hombres (S_{h1}) que son UDI es:

$$\frac{dS_{m1}}{dt} = -i_m\lambda_1 S_{m1} - \mu_1 S_{m1}, \quad (3.1)$$

$$\frac{dS_{h1}}{dt} = -i_h\alpha_1 S_{h1} - \mu_1 S_{h1}. \quad (3.2)$$

El primer término de la ecuación representa infección con el VIH por jeringuilla, que es el producto de la tasa de intercambio de inyectables por mujer (i_m) y la probabilidad de adquirir VIH por intercambio de inyectables con otros UDI por mujer (λ_1). El segundo término representa muerte natural de la población de susceptibles (con parámetro μ_1). La probabilidad de adquirir VIH por intercambio de inyectables con otros UDI por mujer (λ_1) es igual a la eficiencia de transmisión

del VIH mediante inyectable dado que el inyectable está infectado (β_1) multiplicado por la seroprevalencia en el grupo de hombres y mujeres (sexualmente activos y no sexualmente activos) en la población UDI al tiempo t . La interpretación de la Ecuación 3.2 para la población susceptible de hombres es similar a la de mujeres con parámetros i_h y α_1 análogos a i_m y λ_1 , respectivamente.

La razón de de cambio del tamaño de la población susceptible de mujeres (S_{m2}) y hombres (S_{h2}) que tienen dos factores de riesgo (UDI y son sexualmente activos) es:

$$\frac{dS_{m2}}{dt} = -i_m \lambda_1 S_{m2} - C_{m2}(t) \lambda_2 S_{m2} - \mu_1 S_{m2}, \quad (3.3)$$

$$\frac{dS_{h2}}{dt} = -i_h \alpha_1 S_{h2} - K_{h2}(t) \alpha_2 S_{h2} - \mu_1 S_{h2}. \quad (3.4)$$

Como en la Ecuación 3.1, el primer término representa las infectadas con VIH por jeringuilla. En este caso, el segundo término representa mujeres infectadas con VIH por transmisión heterosexual y se define como el producto de la tasa de cambio de parejas sexuales por mujer UDI [$C_{m2}(t)$] (que definiremos luego) y la probabilidad de adquirir VIH por transmisión heterosexual para mujer (λ_2). El tercer término representa la muerte natural de mujeres donde la tasa de mortalidad natural (No -VIH) es (μ_1). La probabilidad de adquirir VIH por transmisión heterosexual por mujer (λ_2) es igual a la eficiencia de transmisión heterosexual

de hombre a mujer por contacto sexual (dado que el hombre estaba infectado) (β_{hm}) multiplicado por la seroprevalencia ponderada $\sum_j M_m(m2, j, t)P_h(j, t)$; donde $P_h(j, t)$ es la seroprevalencia en hombres sexualmente activos del subgrupo j (j es igual a hombres UDISA u hombres (No-UDI)SA y $M_m(m2, j, t)$ es la probabilidad de que una mujer UDI sexualmente activa tenga contacto sexual con un hombre del subgrupo j en el tiempo t . En efecto se precisaron los parámetros $C_{m2}(t)$ y λ_2 , para los hombres $K_{h2}(t)$ y α_2 como:

$$C_{m2}(t) = C_{m2}(t-1) \times \frac{N_{h2}(t)K_{h2}(t-1)M_h(h2, m2, t-1) + N_{h3}(t)K_{h3}(t-1)M_h(h3, m2, t-1)}{N_{h2}(t-1)K_{h2}(t-1)M_h(h2, m2, t-1) + N_{h3}(t-1)K_{h3}(t-1)M_h(h3, m2, t-1)}, \quad (3.5)$$

$$K_{h2}(t) = K_{h2}(t-1) \times \frac{N_{m2}(t)C_{m2}(t-1)M_m(m2, h2, t-1) + N_{m3}(t)C_{m3}(t-1)M_m(m3, h2, t-1)}{N_{m2}(t-1)C_{m2}(t-1)M_m(m2, h2, t-1) + N_{m3}(t-1)C_{m3}(t-1)M_m(m3, h2, t-1)}, \quad (3.6)$$

$$\lambda_2 = \beta_{hm} * (M_m(m2, h2, t)P_h(h2, t) + M_m(m2, h3, t)P_h(h3, t)), \quad (3.7)$$

y

$$\alpha_2 = \beta_{mh} * (M_h(h2, m2, t)P_m(m2, t) + M_h(h2, m3, t)P_m(m3, t)). \quad (3.8)$$

donde

$$\begin{aligned} P_h(h2, t) &= \frac{I_{h2} + A_{h2}}{S_{h2} + I_{h2} + A_{h2}} \\ P_h(h3, t) &= \frac{I_{h3} + A_{h3}}{S_{h3} + I_{h3} + A_{h3}} \\ P_m(m2, t) &= \frac{I_{m2} + A_{m2}}{S_{m2} + I_{m2} + A_{m2}} \\ P_m(m3, t) &= \frac{I_{m3} + A_{m3}}{S_{m3} + I_{m3} + A_{m3}} \end{aligned} \quad (3.9)$$

y A e I representan los subgrupos de enfermos con SIDA e infectados con VIH respectivamente, que se describirán más adelante. La interpretación de la Ecuación 3.4 de la población susceptible de hombres que tienen dos factores de riesgo es similar a la de mujeres pero usando sus parámetros respectivos.

La razón de cambio del tamaño de la población susceptible de mujeres (S_{m3}) y hombres (S_{h3}) que tienen un factor de riesgo (transmisión heterosexual, (No-UDI)SA) es:

$$\frac{dS_{m3}}{dt} = -C_{m3}(t)\lambda_3 S_{m3} - \mu_2 S_{m3}, \quad (3.10)$$

$$\frac{dS_{h3}}{dt} = -K_{h3}(t)\alpha_3 S_{h3} - \mu_2 S_{h3}. \quad (3.11)$$

El primer término representa el grupo de las mujeres susceptibles que se infectan con VIH por transmisión heterosexual, el cual es el producto de la tasa de cambio de parejas sexuales por mujer No-UDI [$C_{m3}(t)$] y la probabilidad de adquirir VIH por transmisión heterosexual por mujer (λ_3) y el número de mujeres susceptibles S_{m3} . El segundo término representa el número de mujeres que emigran de esta población por causas ajenas al VIH y que es el producto de la tasa de mortalidad natural (No -VIH) (μ_2) y el número de mujeres susceptibles S_{m3} . La probabilidad de adquirir VIH por transmisión heterosexual por mujer No-UDI (λ_3) es igual a la eficiencia de transmisión heterosexual de hombre a mujer por contacto sexual (dado que el hombre estaba infectado) (β_{hm}) multiplicado por la seroprevalencia ponderada $\sum_j M_m(m3, j, t)P_h(j, t)$; donde $P_h(j, t)$ es la seroprevalencia en hombres sexualmente activos del subgrupo j (j es igual a hombres UDISA u hombres (No-UDI)SA) y $M_m(m3, j; t)$ es la probabilidad que una mujer (No-UDI)SA tenga contacto sexual con un hombre del subgrupo j en el tiempo t . Ahora se explican los parámetros $C_{m3}(t)$, $K_{h3}(t)$, λ_3 y α_3 dados por:

$$\begin{aligned}
C_{m3}(t) &= C_{m3}(t-1) \\
&\times \frac{N_{h2}(t)K_{h2}(t-1)M_h(h2, m3, t-1) + N_{h3}(t)K_{h3}(t-1)M_h(h3, m3, t-1)}{N_{h2}(t-1)K_{h2}(t-1)M_h(h2, m3, t-1) + N_{h3}(t-1)K_{h3}(t-1)M_h(h3, m3, t-1)},
\end{aligned} \tag{3.12}$$

$$\begin{aligned}
K_{h3}(t) &= K_{h3}(t-1) \\
&\times \frac{N_{m2}(t)C_{m2}(t-1)M_m(m2, h3, t-1) + N_{m3}(t)C_{m3}(t-1)M_m(m3, h3, t-1)}{N_{m2}(t-1)C_{m2}(t-1)M_m(m2, h3, t-1) + N_{m3}(t-1)C_{m3}(t-1)M_m(m3, h3, t-1)},
\end{aligned} \tag{3.13}$$

$$\lambda_3 = \beta_{hm} * (M_m(m3, h2, t)P_h(h2, t) + M_m(m3, h3, t)P_h(h3, t)), \tag{3.14}$$

y

$$\alpha_3 = \beta_{mh} * (M_h(h3, m2, t)P_m(m2, t) + M_h(h3, m3, t)P_m(m3, t)). \tag{3.15}$$

La razón de cambio del tamaño de población de los seis subgrupos de mujeres y

hombres infectados con VIH se presenta en las siguientes ecuaciones:

$$\frac{dI_{m1}}{dt} = i_m \lambda_1 S_{m1} - \frac{1}{v} I_{m1} - \mu_1 I_{m1}, \quad (3.16)$$

$$\frac{dI_{h1}}{dt} = i_h \alpha_1 S_{h1} - \frac{1}{v} I_{h1} - \mu_1 I_{h1}, \quad (3.17)$$

$$\frac{dI_{m2}}{dt} = (i_m \lambda_1 + C_{m2}(t) \lambda_2) S_{m2} - \frac{1}{v} I_{m2} - \mu_1 I_{m2}, \quad (3.18)$$

$$\frac{dI_{h2}}{dt} = (i_h \alpha_1 + K_{h2}(t) \alpha_2) S_{h2} - \frac{1}{v} I_{h2} - \mu_1 I_{h2}, \quad (3.19)$$

$$\frac{dI_{m3}}{dt} = C_{m3}(t) \lambda_3 S_{m3} - \frac{1}{v} I_{m3} - \mu_2 I_{m3}, \quad (3.20)$$

$$\frac{dI_{h3}}{dt} = K_{h3}(t) \alpha_3 S_{h3} - \frac{1}{v} I_{h3} - \mu_2 I_{h3}. \quad (3.21)$$

La razón de cambio del número de mujeres UDI infectadas ($\frac{dI_{m1}}{dt}$) es igual al número de mujeres recientemente infectadas ($i_m \lambda_1 S_{m1}$) menos el número de mujeres que mueren por causas naturales ($\mu_1 I_{m1}$) o porque desarrollaron el SIDA [$\frac{1}{v} I_{m1}$]. La razón de cambio del número de mujeres UDISA infectadas ($\frac{dI_{m2}}{dt}$) es igual al número de mujeres recientemente infectadas [$i_m \lambda_1 + C_{m2}(t) \lambda_2$] menos el número de mujeres que mueren por causas naturales ($\mu_1 I_{m2}$) o porque desarrollaron el SIDA [$\frac{1}{v} I_{m2}$]. La razón de cambio del número de mujeres (No-UDI)SA infectadas ($\frac{dI_{m3}}{dt}$) es igual al número de mujeres recientemente infectadas [$C_{m3}(t) \lambda_3 S_{m3}$] menos el número de mujeres que mueren por causas naturales ($\mu_2 I_{m3}$) o porque desarrollaron el SIDA [$\frac{1}{v} I_{m3}$]. El parámetro v es el tiempo promedio de incubación del virus (VIH). La interpretación de las ecuaciones para los hombres es similar a las de mujeres.

La razón de cambio del tamaño de la población de los tres subgrupos de mujeres y hombres con SIDA se presentan en las siguientes ecuaciones:

$$\frac{dA_{m1}}{dt} = \frac{1}{v}I_{m1} - \left(\frac{1}{s} + \mu_1\right)A_{m1}, \quad (3.22)$$

$$\frac{dA_{h1}}{dt} = \frac{1}{v}I_{h1} - \left(\frac{1}{s} + \mu_1\right)A_{h1}, \quad (3.23)$$

$$\frac{dA_{m2}}{dt} = \frac{1}{v}I_{m2} - \left(\frac{1}{s} + \mu_1\right)A_{m2}, \quad (3.24)$$

$$\frac{dA_{h2}}{dt} = \frac{1}{v}I_{h2} - \left(\frac{1}{s} + \mu_1\right)A_{h2}, \quad (3.25)$$

$$\frac{dA_{m3}}{dt} = \frac{1}{v}I_{m3} - \left(\frac{1}{s} + \mu_2\right)A_{m3}, \quad (3.26)$$

$$\frac{dA_{h3}}{dt} = \frac{1}{v}I_{h3} - \left(\frac{1}{s} + \mu_2\right)A_{h3}. \quad (3.27)$$

La razón de cambio del número de mujeres UDI con SIDA ($\frac{dA_{m1}}{dt}$) es igual al número de mujeres que progresan a la etapa del SIDA [$\frac{1}{v}I_{m1}$] menos el número de mujeres que mueren por causas naturales ($\mu_1 I_{m1}$) o porque mueren de SIDA [$\frac{1}{s}A_{m1}$]. La razón de cambio del número de mujeres UDISA con SIDA ($\frac{dA_{m2}}{dt}$) es igual al número de mujeres que progresan a la etapa del SIDA [$\frac{1}{v}I_{m2}$] menos el número de mujeres que mueren por causas naturales ($\mu_1 I_{m2}$) o porque mueren de SIDA [$\frac{1}{s}A_{m2}$]. La razón de cambio del número de mujeres (No-UDI)SA con SIDA ($\frac{dA_{m3}}{dt}$) es igual al número de mujeres que progresan a la etapa del SIDA [$\frac{1}{v}I_{m3}$] menos el número de mujeres que mueren naturalmente ($\mu_2 I_{m3}$) o porque mueren de SIDA [$\frac{1}{s}A_{m3}$]. El parámetro s es el tiempo promedio de sobrevivencia desde el diagnóstico del SIDA hasta su

muerte. La interpretación de las ecuaciones para los hombres es similar a la de las mujeres.

Blower et al. (1991), especifica la tasa de cambio de parejas sexuales y el número de hombres N_h y mujeres N_m sexualmente activos y establece que la suma total de parejas sexuales en cada grupo debe satisfacer la siguiente restricción, que también se consideraron en el modelo propuesto:

$$\sum_j N_{mj}C_{mj}(t) = \sum_j N_{hj}K_{hj}(t), \quad (3.28)$$

donde;

$$N_{mj}(t) = S_{mj}(t) + I_{mj}(t) + A_{mj}(t), \quad (3.29)$$

$$N_{hj}(t) = S_{hj}(t) + I_{hj}(t) + A_{hj}(t), \quad (3.30)$$

donde $j=UDISA$ o $(No-UDI)SA$. En forma general podemos expresar la tasa de cambio de parejas sexuales en proporción a la disponibilidad del sexo opuesto de la siguiente manera:

$$C_{mi}(t) = C_{mi}(t-1) \frac{\sum_j N_{hj}(t)K_{hj}(t)M_h(j, i, t-1)}{\sum_j N_{hj}(t-1)K_{hj}(t-1)M_h(j, i, t-1)} \quad (3.31)$$

$$K_{hi}(t) = K_{hi}(t-1) \frac{\sum_j N_{mj}(t)C_{mj}(t)M_m(j, i, t-1)}{\sum_j N_{mj}(t-1)C_{mj}(t-1)M_m(j, i, t-1)} \quad (3.32)$$

donde $i=UDISA$ o $(No-UDI)SA$.

3.2. Estimación puntual y de la función de densidad de los parámetros

Para el análisis de nuestro modelo necesitamos estimadores puntuales de cada uno de los parámetros. El análisis de incertidumbre y el análisis de sensibilidad requieren conocer la distribución conjunta de los parámetros de interés. Si se asume que los parámetros son independientes (Blower et al., 1994) sólo se requiere conocer la distribución marginal de cada parámetro, pues la distribución conjunta es el producto de éstas. El Cuadro 3.2 muestra los estimados puntuales que tenemos disponibles para cada uno de los parámetros de nuestro modelo. Incluimos también, los estimados usados por Blower et al. (1994) con fines informativos. El Cuadro 3.3 muestra las distribuciones que se utilizaron. Más adelante mostramos la información detallada que se obtuvo para cada uno de los parámetros.

La indagación de los datos se realizó por la búsqueda de estudios, artículos, tesis, monografías, etc. por medio del internet, llamadas telefónicas y visita al Departamento de Vigilancia SIDA de Puerto Rico, al Recinto de Ciencias Médicas y a organizaciones privadas dedicadas al VIH/SIDA en Puerto Rico. Se descubrió que para los parámetros de interés en el modelo propuesto hay muy poca información

actualizada para la población de Puerto Rico. Algunos de estos estimados corresponden a estudios sobre Puerto Rico, otros sobre Estados Unidos que incluyen Puerto Rico y otros sólo sobre Estados Unidos. Es importante hacer notar que esto impone limitaciones en la inferencia para Puerto Rico pues los valores de los parámetros no necesariamente reflejan la población UDI en Puerto Rico. Ahora la información detallada para cada parámetro.

β_1

La tasa de transmisión del VIH mediante inyectable (β_1) es el cociente entre el número total de hombres y mujeres UDI con VIH y la población UDI. La Administración de Servicios de Salud Mental y Contra la Adicción (ASSMCA) estima la población UDI para Puerto Rico de las personas que son atendidas en los programas de tratamiento por abuso de sustancias y salud mental durante el año fiscal 2009 – 2010 (Apéndice A.4). En el año fiscal 2009 – 2010 (1-julio-2009 hasta 30-junio-2010) el número de usuarios de droga por inyectable fue de 9,310 de los cuales el número de infectados con VIH fue 510 usuarios, por lo que el estimado puntual que usamos para $\beta_1 = 510/9,310$. Blower et al.(1991) asignó la función de distribución triangular para β_1 , con pico en cero y rango de valores entre 0 y 1 por un estudio realizado para la comunidad de UDI de Nueva York. Por no tener información adicional se utilizó la misma distribución que usa Blower et al.(1991)

para este estudio.

β_{mh} y β_{hm}

Morbidity and Mortality Weekly Report (MMWR) realizó un estudio sobre “Incidence and Diagnoses of HIV Infection-Puerto Rico (2006).” En éste se estima la eficiencia de transmisión heterosexual por asociación (de mujer a hombre dado que la mujer está infectada) β_{mh} en 0.25. Baggaley, et al.(2010) estudió el riesgo de transmisión del VIH por contacto heterosexual y hombre con hombre en Estados Unidos, donde estimó que β_{mh} varía entre 0.225 – 0.574. Blower et al. (1991) estimó este valor entre 0 y 0.5 y le asignó la función de distribución uniforme por un estudio desarrollado para la comunidad de UDI de Nueva York. Consecuentemente se utilizó 0.25 para el valor puntual y la función de distribución de probabilidad uniforme para β_{mh} , con valor mínimo en 0 y valor máximo de 0.5.

i_m y i_h

La tasa de intercambio de inyectables por año para mujeres UDI (i_m) y para hombres UDI (i_h) es el número de jeringuillas que intercambia un individuo durante un año. Andía, J. et al. (2008) examinó la influencia del intercambio de jeringuilla entre UDI puertorriqueños en Nueva York y Puerto Rico, y calculan el valor puntual estimado para i_m en 1,999.2 inyectables por año y i_h en 2,137 inyectables por año. Iniciativa Comunitaria, institución dedicada al servicio social por medio

del proyecto Punto Fijo se encarga del intercambio de jeringuillas en Puerto Rico. El Sr. Francisco Torres coordinador del proyecto informó que el número de jeringuillas que usa un individuo anualmente varía entre 12 y 3,650. Esta información está basada en su experiencia y no en datos recopilados por su agencia. Blower et al.(1991) estimó i_m entre 13 y 5,265 e i_h entre 13 y 3,120 y le asignó una distribución sesgada a la izquierda a ambos parámetros. Definimos, entonces, los estimados para nuestro modelo como: $i_m = 1,999.2$ e $i_h = 2,137$ y le asignamos la distribución Gamma sesgada a la izquierda entre 12 y 3,650. Usamos la distribución Gamma por su facilidad de uso.

C_{m2} y C_{m3}

La tasa de cambio de parejas sexuales por año para mujeres UDISA ($C_{m2}(t)$) es 12 según estimado por Latka, M. et al. (2006) para mujeres de Baltimore, Miami, New York y San Francisco. La tasa de cambio de parejas sexuales en puertorriqueños (hombres y mujeres) UDISA está en el rango de 1 a 120 (Deren, S. et al. 2006). Blower et al. (1991) asignó a la tasa de cambio de parejas sexuales por año de mujeres UDISA ($C_{m2}(t)$) la distribución sesgada a la izquierda y con rango entre 1 y 100. En consecuencia el valor puntual para $C_{m2}(t) = 12$ y se asignó la distribución Gamma sesgada a la izquierda entre 1 y 120. La tasa de cambio de parejas sexuales para mujeres (No- UDI)SA ($C_{m3}(t)$) se estima en 8 parejas por año según Stevens,

S. (2006) para mujeres de Arizona. La tasa de cambio de parejas sexuales varía entre 1 y 15 por año (Stein, 1987). Blower et al. (1991) asignó a la tasa de cambio de parejas sexuales por año de mujeres (No-UDI)SA ($C_{m3}(t)$), la distribución sesgada a la izquierda y con rango entre 1 y 20. En nuestro caso definimos la función de densidad de probabilidad para la tasa de cambio de parejas sexuales por año de mujeres (No-UDI)SA usando la distribución Gamma sesgada a la izquierda entre 1 y 15 y valor puntual de 8 parejas por año.

K_{h2} y K_{h3}

La tasa de cambio de parejas sexuales por año para hombres UDISA ($K_{h2}(t)$) se estima en 27 (Purcell, D. et al. 2006, estudio conducido en Baltimore, Miami, New York y San Francisco de los Estados Unidos). La tasa de cambio de parejas sexuales en puertorriqueños UDISA está en el rango de 1 a 120 (Deren, S. et al. 2006). Blower et al.(1991) asignó a ($K_{h2}(t)$) la distribución sesgada a la izquierda y rango entre 1 y 100. En consecuencia definimos la función de densidad de probabilidad Gamma sesgada a la izquierda para la tasa de cambio de parejas sexuales por año para hombres UDISA de Puerto Rico entre 1 y 120 y usamos 27 parejas sexuales por año como valor puntual. La tasa de cambio de parejas sexuales para hombres (No-UDI)SA ($K_{h3}(t)$) se estima en 7 parejas por año (Johansson, B. 2010, una aplicación de Monte Carlo realizado en Estados Unidos). Este parámetro varía entre

1 y 20 por año según (Bobashev, G. 2010, aplicado en la población de Carolina del Norte, Estados Unidos). Blower et al.(1991) le asignó la distribución sesgada a la izquierda y con rango entre 1 y 38. Por tanto, para este modelo se utilizó como valor puntual 7 parejas sexuales por año y se le asignó una distribución Gamma sesgada a la izquierda entre 1 y 20.

v

El tiempo promedio de incubación del VIH (*v*) es el tiempo desde el momento de la infección del VIH hasta la aparición de los síntomas del SIDA. Chevret et al.(1992) estima el periodo de incubación del VIH en hombres homosexuales infectados de Francia, en 10 años y rango entre 1 y 15 años. Blower et al.(1991) en un estudio realizado para la población UDI de Nueva York estimó a *v* entre 1.36 y 10 años y le asignó una distribución Weibull. Basado en esta literatura para nuestro parámetro se utilizó 10 años como valor puntual y le asignamos una distribución Weibull entre 1 y 15 años.

S

El tiempo promedio de sobrevivencia (*s*) es el tiempo de vida de un individuo después de ser diagnosticado con SIDA hasta su muerte. El tiempo promedio de sobrevivencia estimado en "HIV Surveillance Report 2008" es de 3 años o más, esto según un estudio hecho para Estados Unidos y Puerto Rico por el Centro de

Control de Enfermedades (CDC). Lee et al. (2001) investigó sobre la sobrevivencia después del diagnóstico del SIDA en adolescentes y adultos de los Estados Unidos en los años de 1984 hasta 1997, y encontró que el periodo de sobrevivencia varía entre 1 y 5 años. Blower et al. (1991) en un estudio realizado para la población UDI de Nueva York estimó s entre 1 y 5 años y le asignó una distribución sesgada a la izquierda. En consecuencia, para este estudio se utilizó la distribución Gamma sesgada a la izquierda entre 1 y 5 años y el valor puntual de 3 años.

μ_1 y μ_2

Blower et al. (1991) calculó la tasa de mortalidad natural de los UDI y UDISA (μ_1) asumiendo el tiempo promedio que duran inyectándose droga los UDI. Para la población de usuarios de droga por inyectable en Puerto Rico este tiempo es 19 años, según ASSMCA ver (Apéndice A.4). Por otro lado Blower et al. (1991) calculó la tasa de mortalidad para el subgrupo de los (No-UDI)SA (μ_2) asumiendo el tiempo promedio de actividad sexual de una persona (No-UDI)SA, y estimó este periodo en 50 años. Para nuestro modelo se utilizó el tiempo promedio de 19 años para calcular la tasa de mortalidad natural ajena al VIH de los UDI y UDISA para la población de Puerto Rico (μ_1) y el periodo en 50 años para estimar la tasa de mortalidad de los (No-UDI)SA (μ_2), y son fijas en el modelo según Blower et al. (1991).

En el Cuadro 3.2 se muestran los valores puntuales estimados para la población de Puerto Rico y valores puntuales que utilizó Blower et al.(1991). Note que todos los valores puntuales para Puerto Rico son mayores que los de Blower, lo que tendrá un efecto en la seroprevalencia de la enfermedad en Puerto Rico según presentaremos en la Sección 4.1.

Cuadro 3.2: Estimados puntuales de los parámetros

Parámetros	Blower et al.(1991)	Modelo para Puerto Rico	Referencia
β_1	0.01	0.05+	ASSMCA
β_{mh}	0.08	0.25++	CDC.Incidence and diagnoses of HIV infection Puerto Rico, 2006
β_{hm}	0.24	0.25++	CDC.Incidence and diagnoses of HIV infection Puerto Rico, 2006
$C_{m2}(t)$	2.00	12.00+++	Mary H. Latka et al. (2006)
$C_{m3}(t)$	0.57	8.00+++	Sally J. Steves and Rosi Andrade (2006)
$K_{h2}(t)$	1.00	27.00+++	David W. Purcell et al. (2006)
$K_{h3}(t)$	1.00	7.00+++	Sara Soorapanth et al.(2010)
i_m	300.00	1,999.20+	Jonny R. Andía et al. (2008)
i_h	230.00	2,137.90+	Jonny R. Andía et al. (2008)
v	8.00	10.00+++	Sylvie Chevret et al. (1992)
s	1.00	3.00++	Lisa M. Lee et al. (2001)
μ_1	35.00	19.00+	ASSMCA (2009-2010)
μ_2	50.00	50.00+++	Blower et al. (1991)

+Datos correspondientes a estudios sobre Puerto Rico.

++Datos correspondientes a estudios sobre Estados Unidos que incluyen Puerto Rico

+++Datos correspondientes a estudios sólo en Estados Unidos

Cuadro 3.3: Función de distribución de los parámetros

Parámetro	Mínimo	Máximo	Forma funcional
β_I	0	1.0	Triangular(pico en 0)
β_{mh}	0	0.5	Uniforme
β_{hm}	0	0.5	Uniforme
$C_1(t)$	1	120.0	Gamma sesgada a la izquierda
$C_2(t)$	1	15.0	Gamma sesgada a la izquierda
$K_1(t)$	1	120.0	Gamma sesgada a la izquierda
$K_2(t)$	1	20.0	Gamma sesgada a la izquierda
I_m	12	3650.0	Gamma sesgada a la izquierda
I_h	12	3650.0	Gamma sesgada a la izquierda
v	1	15.0	Weibull
s	1	5.0	Gamma sesgada a la izquierda

Capítulo 4

Análisis del modelo VIH/SIDA para Puerto Rico

4.1. Análisis para el valor puntual de los parámetros

El comportamiento del modelo se explora para el estimado puntual de los parámetros, que se obtuvo en la literatura (ver Cuadro 3.2). El tamaño inicial de los subgrupos se genera según explicamos en la introducción del Capítulo 3. Además al tiempo cero $I_{h2} = 1$, lo que significa que hay un hombre UDISA infectado. Se calculó la seroprevalencia con los valores puntuales estimados de los parámetros y el tamaño de los subgrupos generados aleatoriamente $S_{m1} = 8650$, $S_{m2} = 820$, $S_{m3} = 20187$, $S_{h1} = 22654$, $S_{h2} = 13170$, $S_{h3} = 25107$, $I_{m1} = I_{m2} = I_{m3} = I_{h1} = I_{h3} = 0$, $A_{m1} = A_{m2} = A_{m3} = A_{h1} = A_{h2} = A_{h3} = 0$. Se estudió la seroprevalencia en mujeres y hombres UDI, UDISA y (No-UDI)SA. La seroprevalencia es el número o proporción de personas UDI que están infectados con VIH o enfermos con SIDA, como resultado de pruebas serológicas de detección de la infección. El resultado para este escenario particular se muestra en la Figura 4.1; donde se observa que

el nivel de seroprevalencia en la población de UDI de Puerto Rico crece durante los tres primeros años, tiempo en el cual todos los individuos susceptibles tanto UDI Y UDISA se infectan con VIH o se enferman de SIDA. Aproximadamente el nivel de seroprevalencia es del 60% en los dos primeros años tanto en el grupo de hombres y mujeres. El número estimado de casos acumulados de SIDA para Puerto Rico después de 10 años para estos estimados puntuales es 19,148.

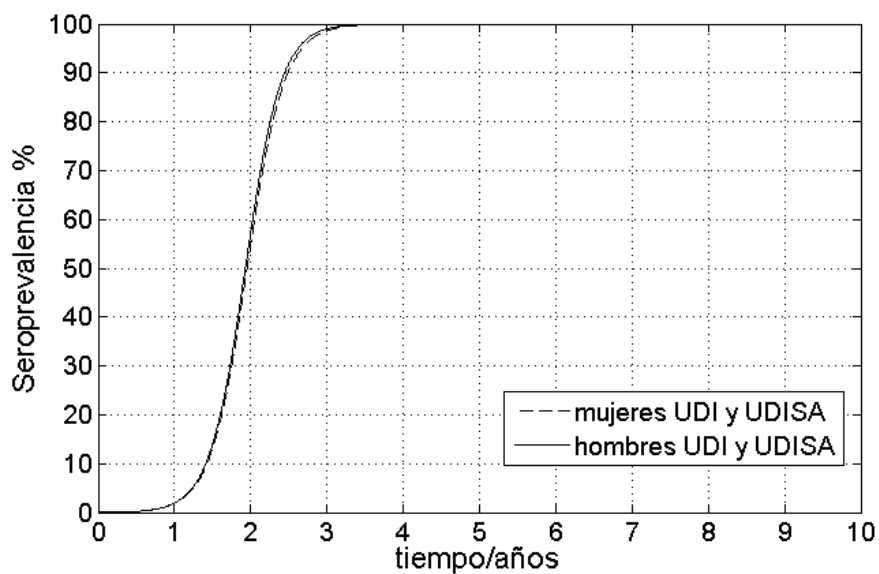


Figura 4.1: Nivel de seroprevalencia (%) en hombres y mujeres UDI.

La Figura 4.2 muestra el nivel de seroprevalencia que estimó Blower et al. (1991), donde el tamaño de los subgrupos de susceptibles fueron los siguientes: $X_{m1} = 12,500$, $X_{m2} = 8,250$, $X_{m3} = 12,500$, $X_{m4} = 16,750$, $X_{m5} = 100,000$, $X_{h1} = 63,750$,

$X_{I2} = 37,500$, $X_{I3} = 11,250$, $X_{I4} = 37,500$, $X_{I5} = 50,000$ y donde usó los valores puntuales del Cuadro 3.2 para los parámetros de entrada.

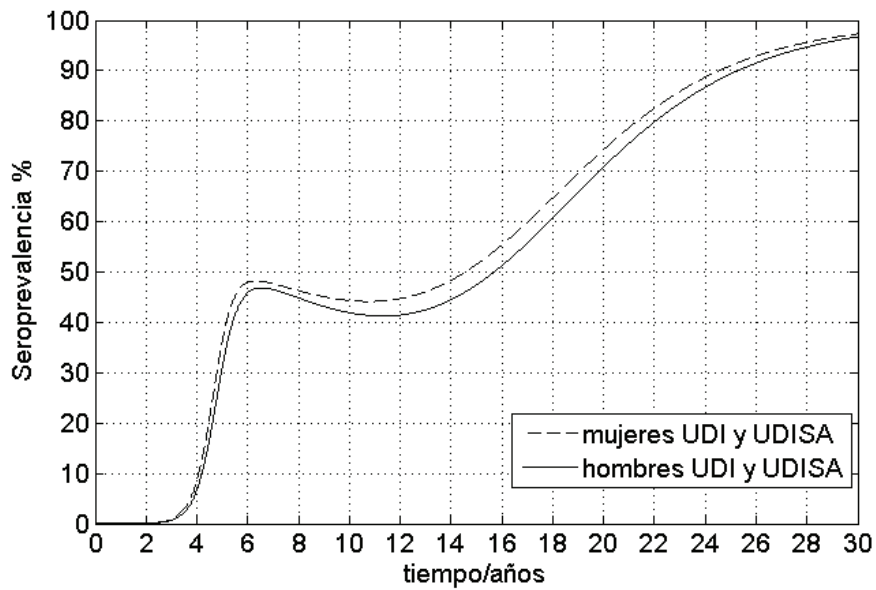


Figura 4.2: Nivel de seroprevalencia (%) en hombres y mujeres UDI.

En la Figura 4.1, se observa que la seroprevalencia después de tres años es el 100 %, es decir que todos los hombres y mujeres de la población UDI de Puerto Rico se infectan con VIH/SIDA en tres años, esto en comparación a la seroprevalencia de hombres y mujeres UDI de la población de Nueva York estimada por Blower et al. (1991) donde tarda 30 años el alcanzar una seroprevalencia de cerca de 100 %. Esto podría deberse a que los valores de los parámetros de entrada para Puerto

Rico son todos mayores que los que usó Blower et al.(1991).

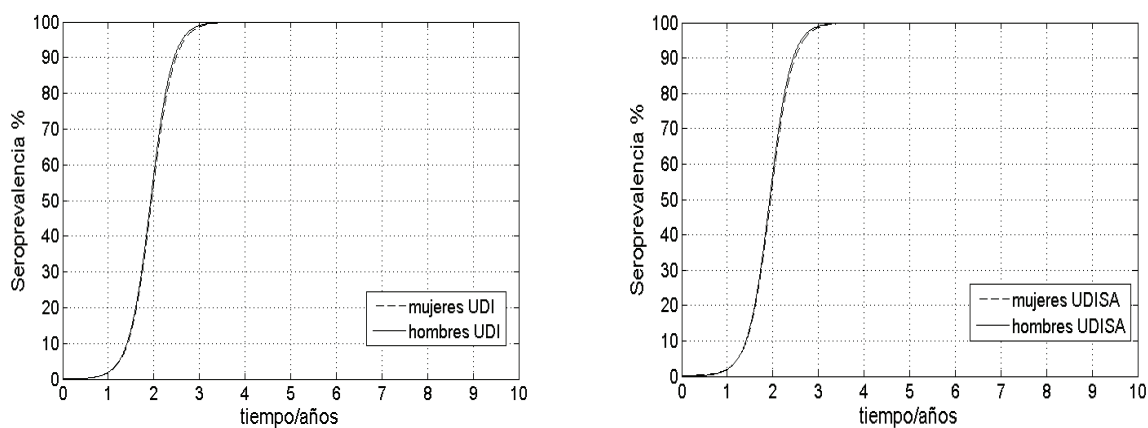


Figura 4.3: Seroprevalencia en los subgrupos UDI (lado izquierdo) y UDISA (lado derecho).

En la Figura 4.3 se observa que el nivel de seroprevalencia tanto en el subgrupo de UDI y UDISA es similar al nivel de seroprevalencia del grupo total de UDI (es decir UDI Y UDISA). Notamos que la diferencia del nivel de seroprevalencia en mujeres (UDI Y UDISA) y el total de mujeres UDI es muy pequeña, son valores multiplicados por 10^{-15} , lo mismo pasa para los subgrupos de hombres (UDI y UDISA) y el total de hombres UDI.

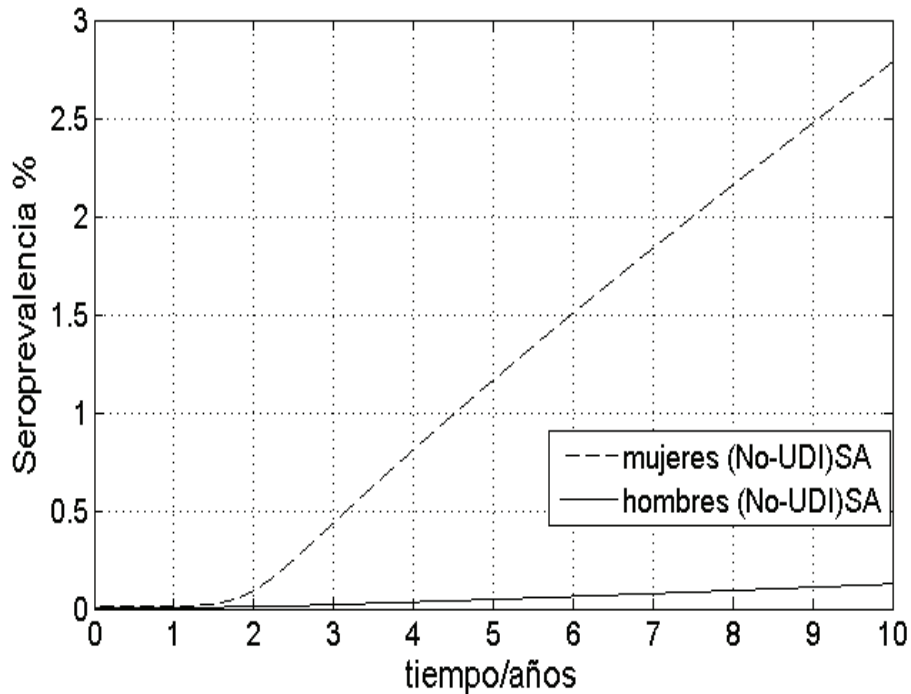


Figura 4.4: Seroprevalencia en los subgrupos en hombres y mujeres (No-UDI)SA

En la Figura 4.4 se muestra el nivel de seroprevalencia del subgrupo de mujeres (No-UDI)SA es mayor que el subgrupo de hombres, significa que el número de mujeres infectadas con VIH o se enferman con SIDA es mayor al número de hombres infectados con VIH o enfermos de SIDA. Sin embargo, el nivel de seroprevalencia en hombres (No-UDI)SA y mujeres (No-UDI)SA es baja, es decir, pocos (No-UDI)SA se infectan.

4.2. Muestreo por hipercubo latino modificado

Como se menciona antes, para el análisis de incertidumbre y el análisis de sensibilidad debemos generar vectores aleatorios $X_i = [x_{i1}, x_{i2}, \dots, x_{ik}]$, $i = 1, 2, \dots, n$ de la distribución conjunta $f(x_{i1}, \dots, x_{ik})$. Asumiendo la independencia de los parámetros la distribución conjunta es $\prod_{j=1}^k f_j(x_{ij})$. Si usamos solo muestreo aleatorio debemos muestrear independientemente de cada $f_j(x_{ij})$ $j = 1, \dots, k$. Si usamos muestreo por hipercubo latino modificado, cada $f_j(x_{ij})$ se divide en n intervalos con igual probabilidad $\frac{1}{n}$. De cada intervalo se selecciona un valor para obtener $n \times k$ valores iniciales. Las columnas de esta matriz se permutan para obtener la matriz de parámetros iniciales deseada. Lo importante es que la meta es explorar eficientemente el espacio de los parámetros. El método aleatorio requiere n grande, particularmente cuando se tienen muchos parámetros. El método por hipercubo latino aparentemente requiere un n menor, pero como discutimos antes, pensamos que potencialmente hay combinaciones de parámetros que no exploran. Sin embargo las secuencias de Halton tienen la propiedad de cubrir un espacio de dimensión s de una manera más uniforme que una secuencia de puntos aleatorios, por lo que se propuso usar la secuencia de Halton para determinar los intervalos donde se hará el muestreo de hipercubo latino modificado en vez de usar las

permutaciones sin reemplazo.

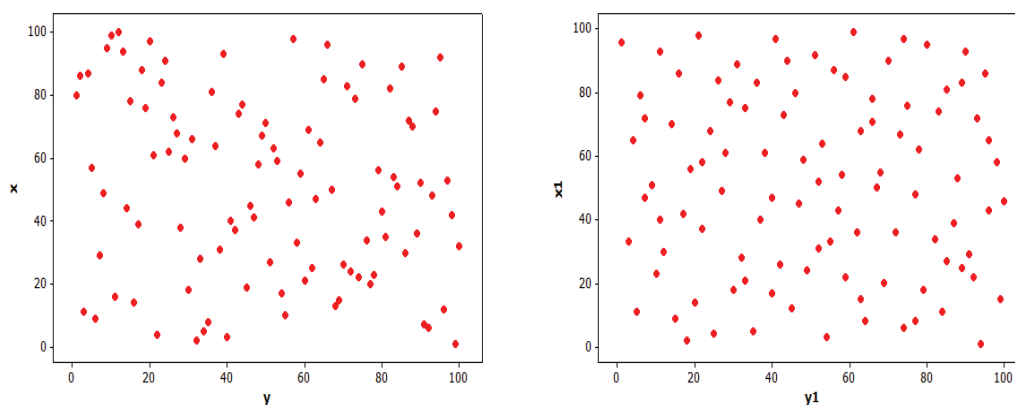


Figura 4.5: Gráfica de los intervalos visitados por MHL y MHL modificado, respectivamente

En la Figura 4.5 (izquierda) se muestra los intervalos que visitaría MHL si el modelo tuviera dos parámetros y estuviéramos haciendo 100 simulaciones. A la derecha se muestran los intervalos que visitaría MHL modificado. Note que los correspondientes a MHL modificado se ven mejor distribuidos, por lo que exploran mejor el espacio muestral. Como mencionamos en la Sección 2.4.3 las secuencias de Halton utilizan un número primo por cada dimensión del espacio a muestrear. En nuestro modelo para Puerto Rico tenemos 10 parámetros de los que debemos muestrear por lo que usamos las bases primas $2, 3, \dots, 29$. A partir de la base 7 se observa el problema de correlación, que también mencionamos en

la Sección 2.4.3, por lo que usamos el método Halton aleatorizado con la bases primas $2, 3, \dots, 29$ para determinar los intervalos donde se hará el muestreo de hipercubo latino modificado.

4.3. Análisis de incertidumbre

Se analizó la incertidumbre del número de casos acumulado de SIDA en Puerto Rico, usando el muestreo aleatorio, el muestreo por hipercubo latino y el muestreo por hipercubo latino modificado, y calculando las estadísticas descriptivas del número de casos acumulados de SIDA en Puerto Rico en los próximos 10 años, a partir del año 2010. Por la ley de los números grandes, consideramos los resultados generados por el muestreo aleatorio para 1000 simulaciones como los resultados que se pueden considerar verdaderos. De aquí que hay un 95% de confianza de que el promedio del número de casos acumulados de SIDA en Puerto Rico en los próximos 10 años esté entre 44,723 y 46,128 casos. En la esquina superior izquierda de la Figura 4.6 se muestra el histograma del número de casos acumulados de SIDA en Puerto Rico para los próximos 10 años basado en el método aleatorio $n = 1000$. En las Figuras 4.6 y 4.7 se muestran, también, los histogramas que se obtienen usando los métodos de muestreo aleatorio, hipercubo latino e hipercubo latino

modificado ($n = 100, n = 200$). En los Cuadros 4.1 y 4.2 se presentan las estadísticas descriptivas para cada método. Notamos que las estadísticas descriptivas son similares, pero el método LHS y LHS modificado no capturan los valores inferiores de la distribución de casos acumulados de SIDA. La diferencia entre los métodos se hace visible en el análisis de sensibilidad.

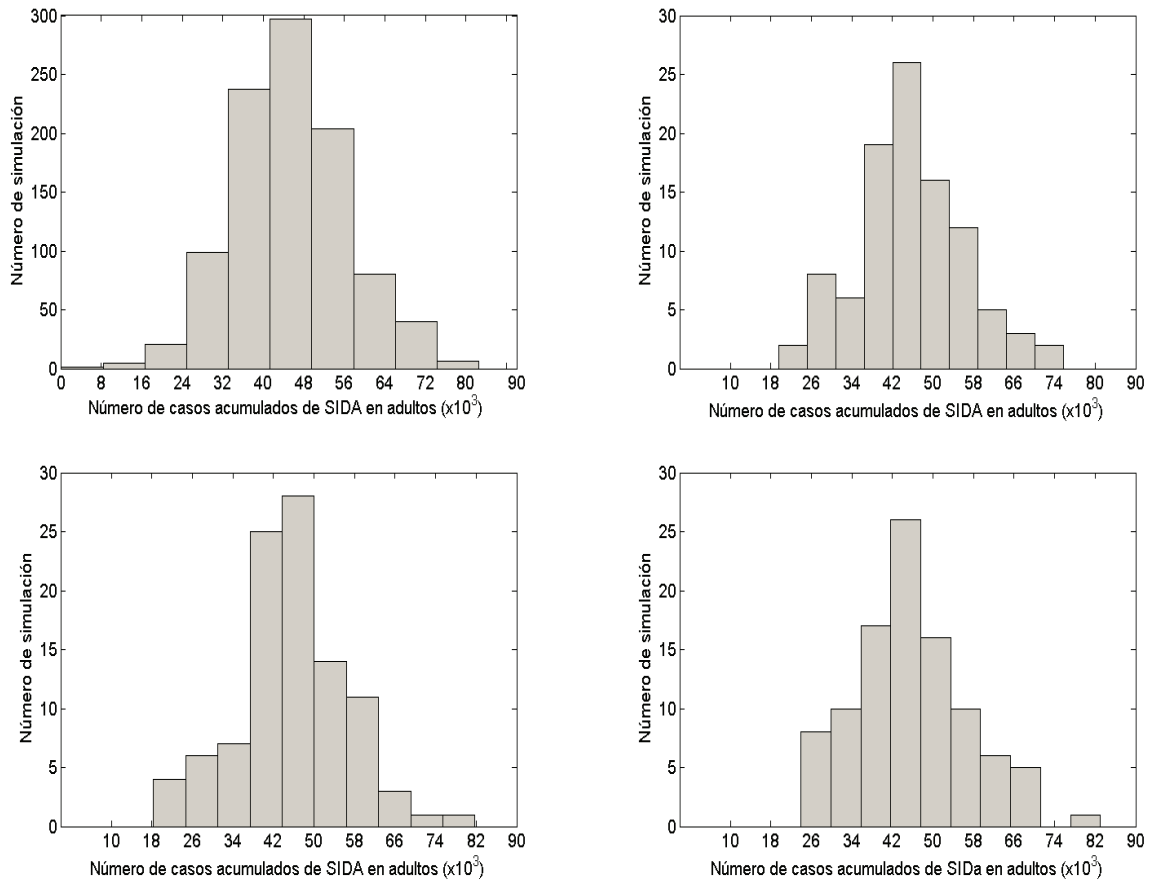


Figura 4.6: Histograma de frecuencia del número de casos acumulados de SIDA en adultos, para los próximos 10 años. Izquierda superior: Aleatorio $n=1000$; derecha superior: Aleatorio $n=100$; izquierda inferior: LHS $n=100$ y derecha inferior: LHS modificado $n=100$.

Cuadro 4.1: Estadística descriptiva del número de casos acumulados de SIDA en adultos para $n = 1000$ y $n = 100$

Estadístico	$n = 1000$ Aleatorio	$n = 100$		
		Aleatorio	LHS	MHL modifi- cado
Mínimo	149	19,573	18,321	23,912
Máximo	82,519	75,698	81,782	83,070
Media	45,426	45,258	45,589	45,904
Mediana	45,276	44,642	44,888	45,213
Desviación estándar	11,272	10,559	11,251	11,320
Intervalo de confianza al 95 %	(44,723-46,128)	(43,178-47,338)	(43,383-47,794)	(43,674-48,134)

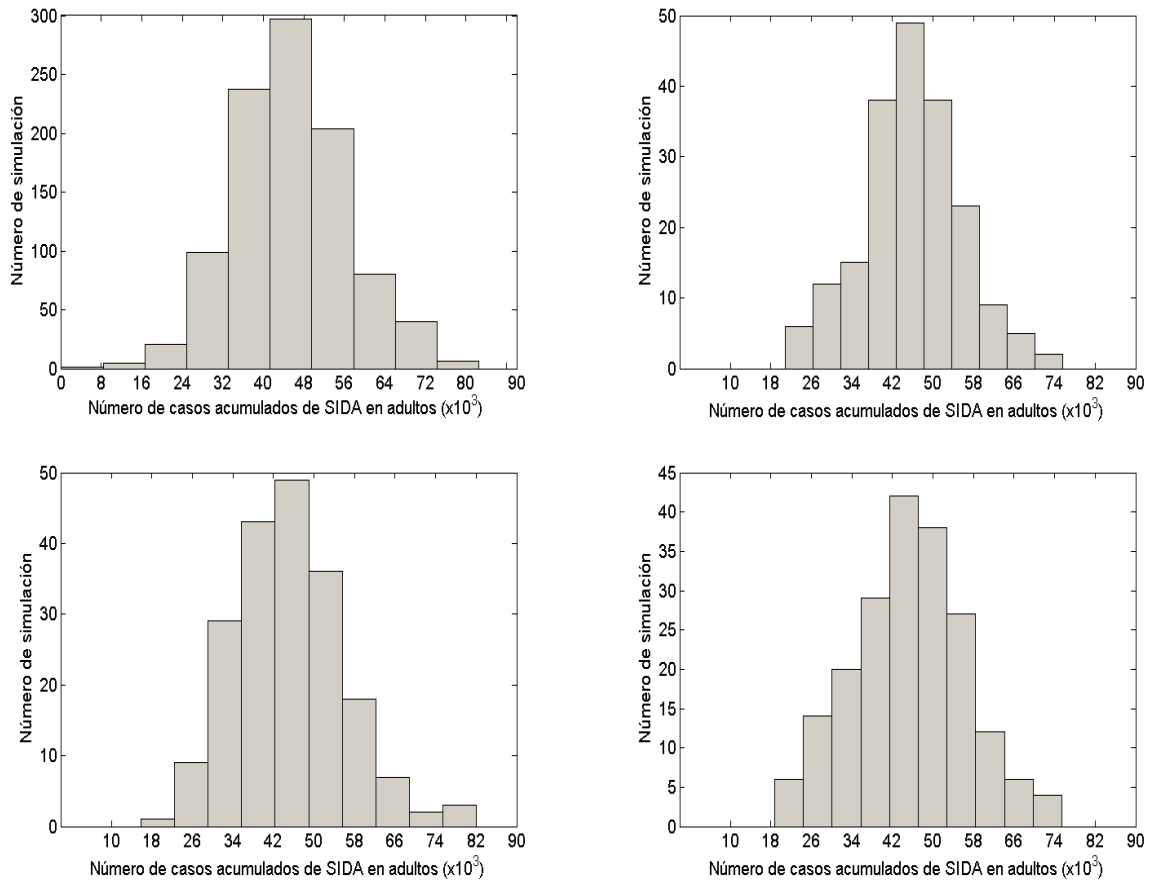


Figura 4.7: Histograma de frecuencia del número de casos acumulados de SIDA en adultos, para los próximos 10 años. Izquierda superior: Aleatorio $n=1000$; derecha superior: Aleatorio $n=200$; izquierda inferior: MHL $n=200$ y derecha inferior: MHL modificado $n=200$.

Cuadro 4.2: Estadística descriptiva del número de casos acumulados de SIDA en adultos para $n = 1000$ y $n = 200$

Estadístico	$n = 1000$ Aleatorio	$n = 200$		
		Aleatorio	MHL	MHL modificado
Mínimo	149	20,899	15,836	18,736
Máximo	82,519	75,642	82,091	75,425
Media	45,426	45,902	44,987	45,282
Mediana	45,276	45,374	44,655	44,786
Desviación estándar	11,272	10,100	10,825	11,208
Intervalo de confianza al 95 %	(44,723-46,128)	(44,492-47,313)	(43,475-46,499)	(43,721-46,843)

4.4. Análisis de sensibilidad

El coeficiente de correlación parcial (CCP) es calculado para el muestreo aleatorio, muestreo por Hipercubo Latino y MHL modificado entre el número acumulado de casos de SIDA para los próximos 10 años y cada uno de los parámetros de entrada. En el Cuadro 4.3 se muestran los resultados para el método aleatorio

($n = 1000$ y $n = 100$) y para los métodos MHL y MHL modificado ($n = 100$). Los asteriscos marcan los parámetros que se consideran importantes a distintos niveles de significancia. Según el método aleatorio $n = 1000$ los parámetros que más contribuyen a la incertidumbre de los estimados del número de casos acumulados de SIDA son: eficiencia de transmisión heterosexual por asociación, tiempo promedio de incubación del virus del VIH y la tasa de cambio de parejas sexuales por año de mujeres (No-UDI)SA. El muestreo por Hipercubo Latino modificado coincide al identificar los mismos parámetros importantes mientras que el MHL y el método aleatorio $n = 100$ identifican otros adicionales. El signo de los CCP identifica la relación cualitativa entre la variable de entrada y la variable respuesta. El valor positivo de los CCP para la mayoría de las variables implica que cuando el valor de la variable entrada crece, el número de casos acumulados de SIDA en adultos también crece. El número de casos acumulados de SIDA en adultos decrece cuando el período de incubación se alarga, porque las personas siguen siendo contagiosas durante un período más largo y por tanto la tasa de progresión de la enfermedad disminuye. Si aumentamos el número de simulaciones a 200 (ver Cuadro 4.4) notamos que los resultados del MHL modificado concuerdan con el aleatorio $n = 1000$, mientras que el MHL y el aleatorio $n = 200$ aún no concuerdan.

Cuadro 4.3: Prueba de correlación parcial para $n = 1000$ y $n = 100$

Parámetros	$n = 1000$	$n = 100$		
		Aleatorio	MHL	MHL modificado
$\beta_{mh} = \beta_{hm}$	0.7177***	0.7479***	0.7175***	0.7519***
v	-0.8280***	-0.8172***	-0.8915***	-0.8703***
$C_{m3}(t)$	0.3826***	0.2400*	0.4644***	0.4532***
$K_{h3}(t)$	0.1853	0.2997**	0.1183	0.1298
$K_{h2}(t)$	-0.0359	-0.0941	0.2533*	0.0296
$C_{m2}(t)$	0.0444	0.2121*	0.0661	-0.0236
i_m	0.0438	0.1867	0.0060	0.0862
β_1	0.1920	0.1584	0.1496	0.1491
i_h	0.1477	0.1748	0.0329	0.0448
s	-0.0376	0.1699	-0.0245	-0.0124

* $\alpha = 0.05$; ** $\alpha = 0.01$ y *** $\alpha = 0.001$

Cuadro 4.4: Prueba de correlación parcial para $n = 1000$ y $n = 200$

Parámetros	$n = 1000$	$n = 200$		
		Aleatorio	MHL	MHL modificado
$\beta_{mh} = \beta_{hm}$	0.7177***	0.7214***	0.7101***	0.7489***
v	-0.8280***	-0.8312***	-0.8001***	-0.8671***
$C_{m3}(t)$	0.3826***	0.3369***	0.1888	0.4377***
$K_{h3}(t)$	0.1853	0.2457*	0.2643**	0.1663
$K_{h2}(t)$	-0.0359	-0.0243	-0.0159	-0.0432
$C_{m2}(t)$	0.0444	0.0075	-0.0151	-0.0507
i_m	0.0438	0.1333	-0.0015	0.0791
β_1	0.1920	0.0788	0.1803	0.1700
i_h	0.1477	-0.0019	0.1292	0.0787
s	-0.0376	-0.0291	-0.0516	-0.1022

* $\alpha = 0.05$; ** $\alpha = 0.01$ y *** $\alpha = 0.001$

En el Apéndice A se presentan los histogramas de los parámetros de entrada y algunas gráficas de la exploración del espacio muestral de los parámetros para $n = 1000$, $n = 100$ y $n = 200$ cuando se usa el muestreo aleatorio. En los apéndices B y C se muestran las gráficas correspondientes a MHL y MHL modificado, respectivamente.

Convergencia a la media de la distribución para los parámetros importantes del modelo

Se analiza la convergencia a la media de la distribución del número de casos acumulados de SIDA en adultos para la población de Puerto Rico y de los parámetros importantes que afectan el número de casos acumulados de SIDA generado por los tres métodos de muestreo Aleatorio: $n = 200$, MHL: $n = 200$, MHL modificado: $n = 200$ y comparamos con el promedio de casos acumulados de SIDA generado por el muestreo aleatorio de tamaño $n = 1000$. Para esto se grafican las medias móviles (*running means*).

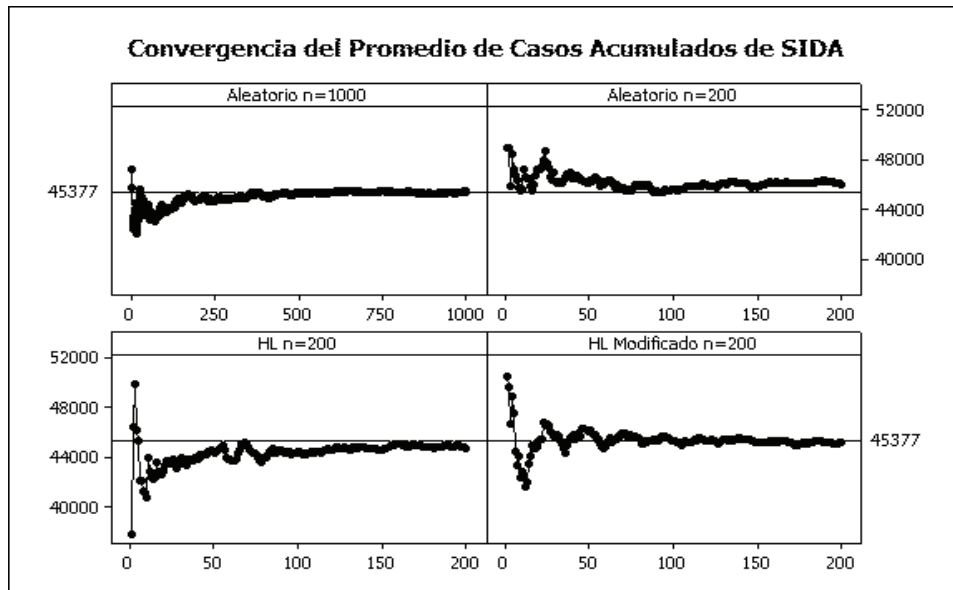


Figura 4.8: Convergencia del promedio de casos acumulados de SIDA en adultos, para los próximos 10 años. Izquierda superior: Aleatorio n=1000; derecha superior: Aleatorio n=200; izquierda inferior: MHL n=200 y derecha inferior: MHL modificado n=200.

En la Figura 4.8 se observa que el promedio de casos acumulados de SIDA estimado por el muestreo por hipercubo latino modificado (n=200) se aproxima al promedio de casos acumulados de SIDA estimado por el muestreo aleatorio (n=1000) mientras que el muestreo aleatorio (n=200) y el muestreo por hipercubo latino (n=200) se quedan por encima y por debajo de este promedio, respectivamente.

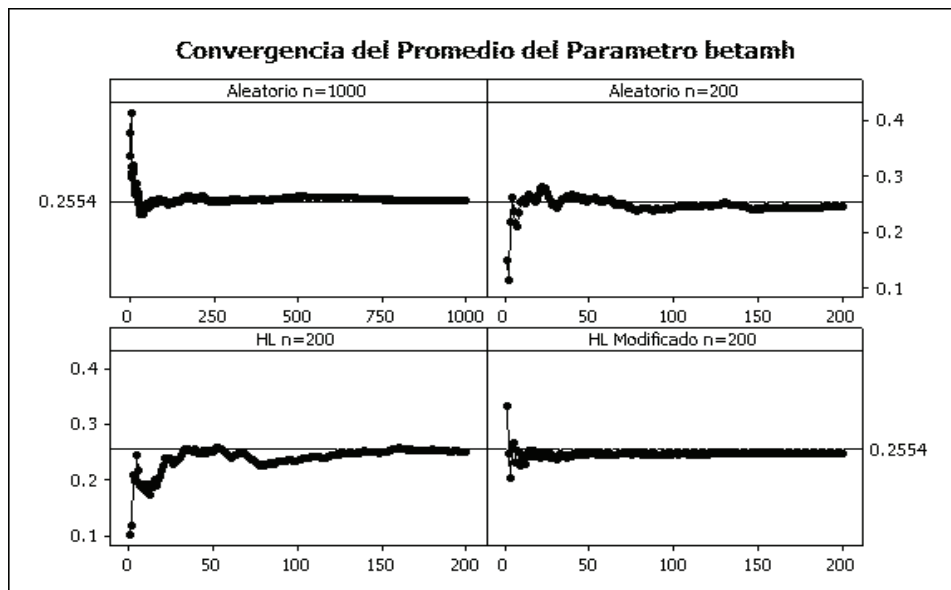


Figura 4.9: Convergencia del promedio de la eficiencia de transmisión de VIH por contacto heterosexual, para los próximos 10 años. Izquierda superior: Aleatorio $n=1000$; derecha superior: Aleatorio $n=200$; izquierda inferior: MHL $n=200$ y derecha inferior: MHL modificado $n=200$.

En la Figura 4.9 se observa que el promedio de la distribución de la eficiencia de transmisión de VIH por contacto heterosexual generada por el muestreo aleatorio, hipercubo latino e hipercubo latino modificado ($n=200$) tiende a estar por debajo del promedio de la distribución de la eficiencia de transmisión de VIH por contacto heterosexual generada por el muestreo aleatorio ($n=1000$). Sin embargo, en el muestreo por hipercubo latino modificado el promedio se estabiliza mucho más rápido que en el muestreo aleatorio y el muestreo por hipercubo latino.

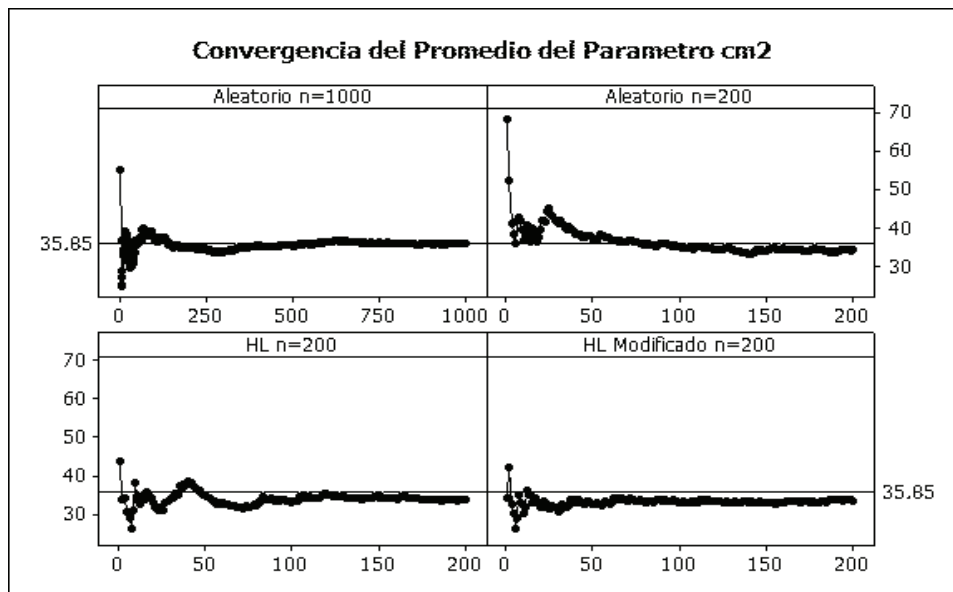


Figura 4.10: Convergencia del promedio de la tasa de cambio de parejas sexuales para mujeres UDISA, para los próximos 10 años. Izquierda superior: Aleatorio $n=1000$; derecha superior: Aleatorio $n=200$; izquierda inferior: MHL $n=200$ y derecha inferior: MHL modificado $n=200$.

El comportamiento en la Figura 4.10 es similar al de la Figura 4.9, esta vez para la tasa de cambio de parejas sexuales para mujeres UDISA. En la Figura 4.10 se observa que el promedio de la distribución de la tasa de cambio de parejas sexuales para mujeres UDISA generada por el muestreo aleatorio, hipercubo latino e hipercubo latino modificado ($n=200$) tiende a estar por debajo del promedio de la distribución de la tasa de cambio de parejas sexuales para mujeres UDISA generada por el muestreo aleatorio ($n=1000$). Sin embargo, en el muestreo por

hipercubo latino modificado el promedio se estabiliza mucho más rápido que en el muestreo aleatorio y el muestreo por hipercubo latino.

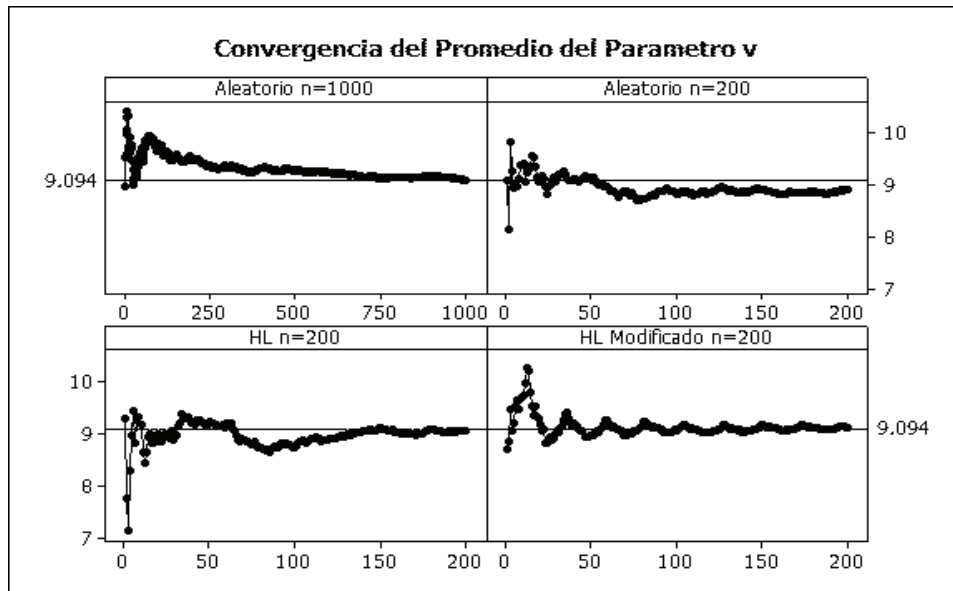


Figura 4.11: Convergencia del promedio del tiempo de incubación del virus del VIH, para los próximos 10 años. Izquierda superior: Aleatorio $n=1000$; derecha superior: Aleatorio $n=200$; izquierda inferior: MHL $n=200$ y derecha inferior: MHL modificado $n=200$.

En la Figura 4.11 se observa que el promedio de la distribución del tiempo promedio de incubación del virus generada por el muestreo por hipercubo latino y por el muestreo por hipercubo latino modificado ($n = 200$) se aproxima al promedio de la distribución del tiempo promedio de incubación del virus generada por el muestreo aleatorio ($n = 1000$) aunque hay un patrón de no convergencia

que disminuye según aumenta el número de simulaciones para el muestreo por hipercubo latino modificado. Por otro lado, se observa que el muestreo aleatorio ($n=200$) subestima el promedio de la distribución.

Capítulo 5

Conclusiones y trabajo futuro

- Para los valores puntuales de los parámetros, los hombres y mujeres UDI se infectan al mismo ritmo y después de 3 años se infectan casi todos. En cambio los hombres y mujeres (No-UDI)SA se infectan menos, siendo las mujeres las que se infectan más. El número estimado de casos acumulados de SIDA para Puerto Rico después de 10 años y a partir del año 2010 es de 19,148.
- Para los valores puntuales de los parámetros, la seroprevalencia en mujeres y hombres UDI y UDISA es similar en ambos subgrupos.
- El análisis de sensibilidad revela que los parámetros de entrada importantes que afectan el número de casos acumulados de SIDA, para los próximos 10 años son la eficiencia de transmisión del VIH (β_{mh}), el tiempo promedio de incubación del virus del VIH (v) y la tasa de cambio de parejas sexuales para mujeres UDISA (C_{m3}). Estos resultados demuestran que los parámetros de conducta sexual y biológicos son importantes para predecir la imprecisión

del número de casos acumulados de SIDA.

- La exploración del espacio muestral con MHL modificado parece ser más uniforme según las gráficas.
- Para el análisis de sensibilidad vemos que el MHL modificado con ($n = 100$ y $n = 200$) concluye lo mismo que el muestreo aleatorio con $n = 1000$, lo que representa un ahorro en el tiempo de simulación.
- Para el análisis de incertidumbre vemos que el MHL y el MHL modificado falla en estimar la parte inferior del histograma del número de casos acumulados de SIDA.
- Explorar el espacio de los parámetros usando el menor número de simulaciones es de suma importancia, particularmente cuando se tienen muchos parámetros. En este trabajo el número de parámetros es relativamente pequeño por lo que podemos comparar los resultados de MHL y MHL modificado contra los del método aleatorio. La eficiencia en el análisis de sensibilidad del MHL modificado sobre el MHL podría deberse a que nuestro modelo tiene pocos parámetros.
- La convergencia al promedio del número de casos acumulados de SIDA en adultos, de la eficiencia de transmisión del VIH por contacto heterosexual, del

tiempo promedio de incubación del virus del VIH y de la tasa de cambio de parejas sexuales para mujeres UDISA parece ser más rápida para el muestreo HL modificado, que para el muestreo aleatorio e HL cuando $n = 200$.

- Como un trabajo futuro se propone explorar la eficiencia del MHL modificado en problemas con más parámetros de entrada ya que el número de parámetros de entrada para este estudio es relativamente pequeño.

Capítulo 6

Limitaciones

El estudio de la dinámica del SIDA en Puerto Rico usando modelos epidemiológicos requiere el tener conjuntos de datos actualizados para la población de Puerto Rico de modo que la inferencia estadística refleje con cierto nivel de confianza la realidad del país. En nuestro estudio usamos conjuntos de datos que obtuvimos de diferentes instituciones dedicadas al SIDA en Puerto Rico como el Departamento de Salud de Puerto Rico e Iniciativa Comunitaria, pero también usamos datos de estudios hechos en Estados Unidos y Puerto Rico, o sólo en Estados Unidos debido a la falta de información para Puerto Rico. El conjunto de datos para Puerto Rico sólo refleja la situación de los pacientes que visitan las instituciones dedicadas al SIDA por lo que la información que tenemos es extremadamente limitada no sólo en términos de municipios, pero también de años. La falta de información para este tipo de estudios no es particular a Puerto Rico, ocurre a nivel mundial, sin embargo entendemos que se deben hacer esfuerzos en Puerto Rico para mejorar la recopilación de información.

Apéndice

A.1 Muestreo aleatorio

Exploración del espacio muestral de los parámetros $n = 1000$

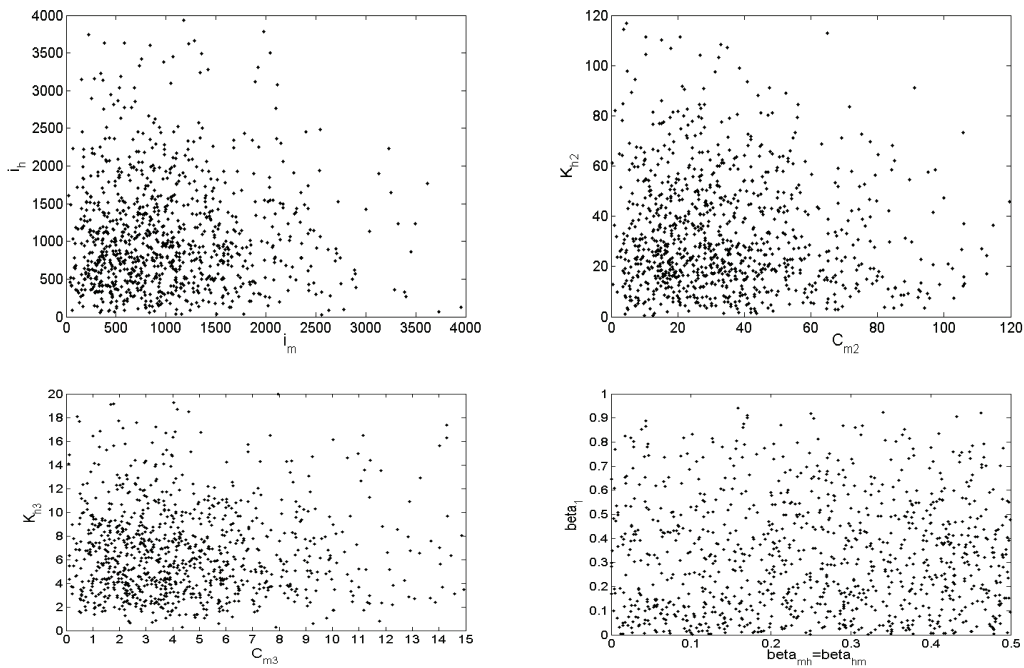


Figura A.1: Exploración del espacio muestral generado por el muestreo aleatorio

$n = 1000$, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_{hm} .

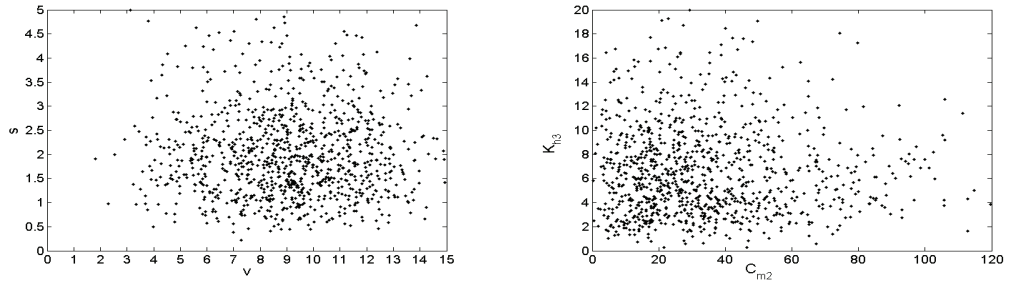


Figura A.2: Exploración del espacio muestral generado por el muestreo aleatorio

$n = 1000$, para los parámetros v vs s y C_{m2} vs K_{h3} .

Histograma de los parámetros de entrada $n = 1000$

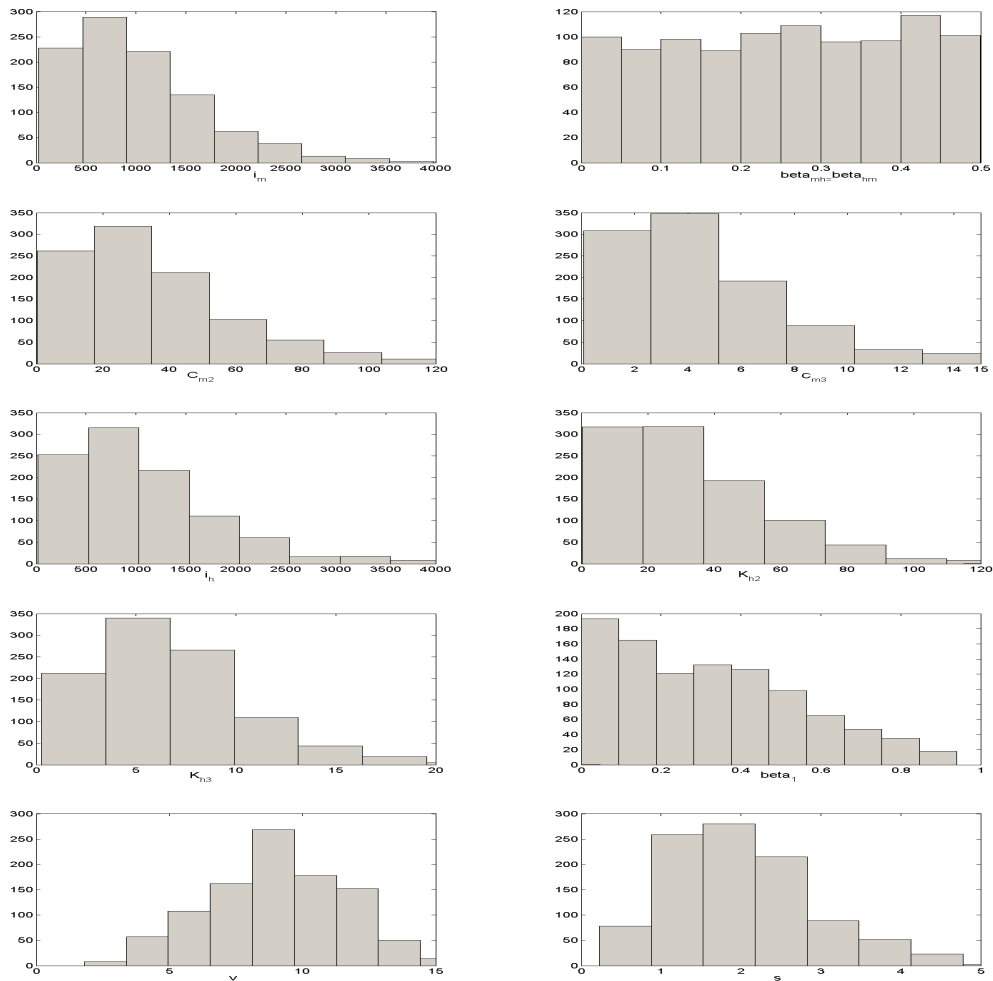


Figura A.3: Histograma de los parámetros de entrada del modelo $i_m, \beta_{mh}, C_{m2},$

$C_{m3}, i_h, K_{h2}, K_{h3}, \beta_1, v$ y s generado por el muestreo aleatorio $n = 1000$.

Exploración del espacio muestral de los parámetros $n = 100$

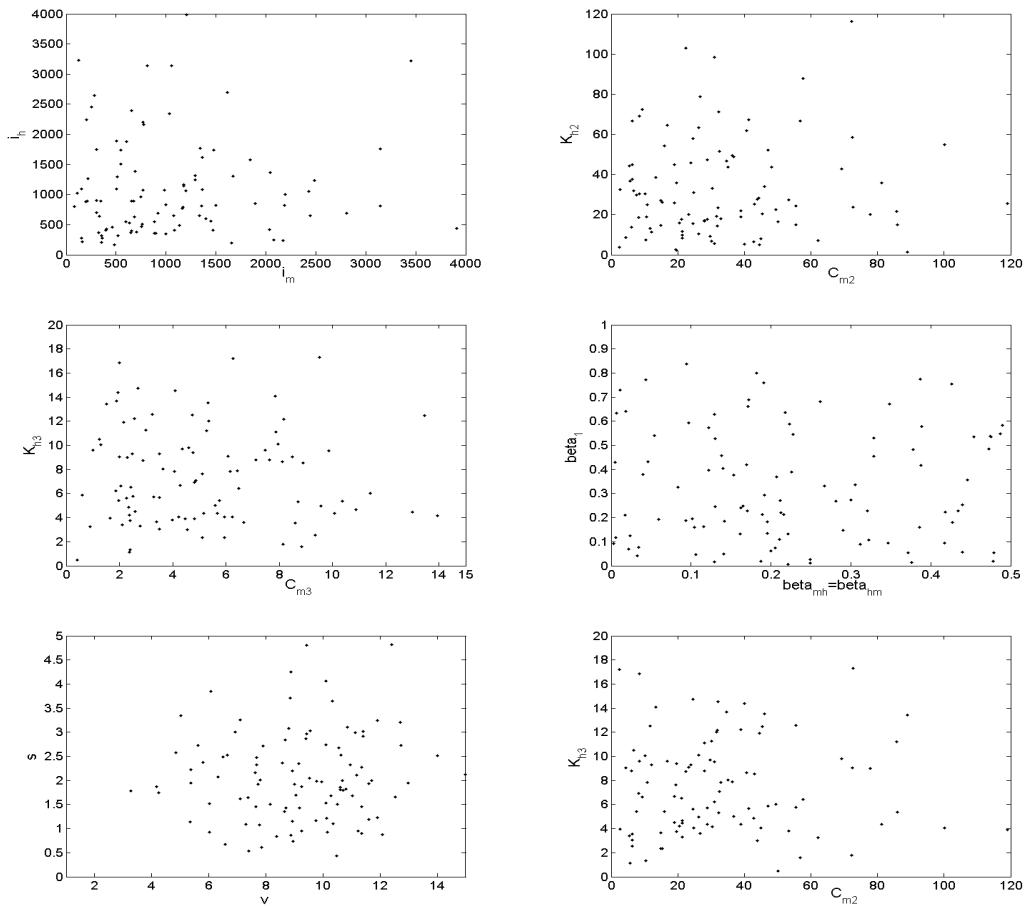


Figura A.4: Exploración del espacio muestral generado por el muestreo aleatorio

$n = 100$, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2}

vs K_{h3} .

Histograma de los parámetros de entrada $n = 100$

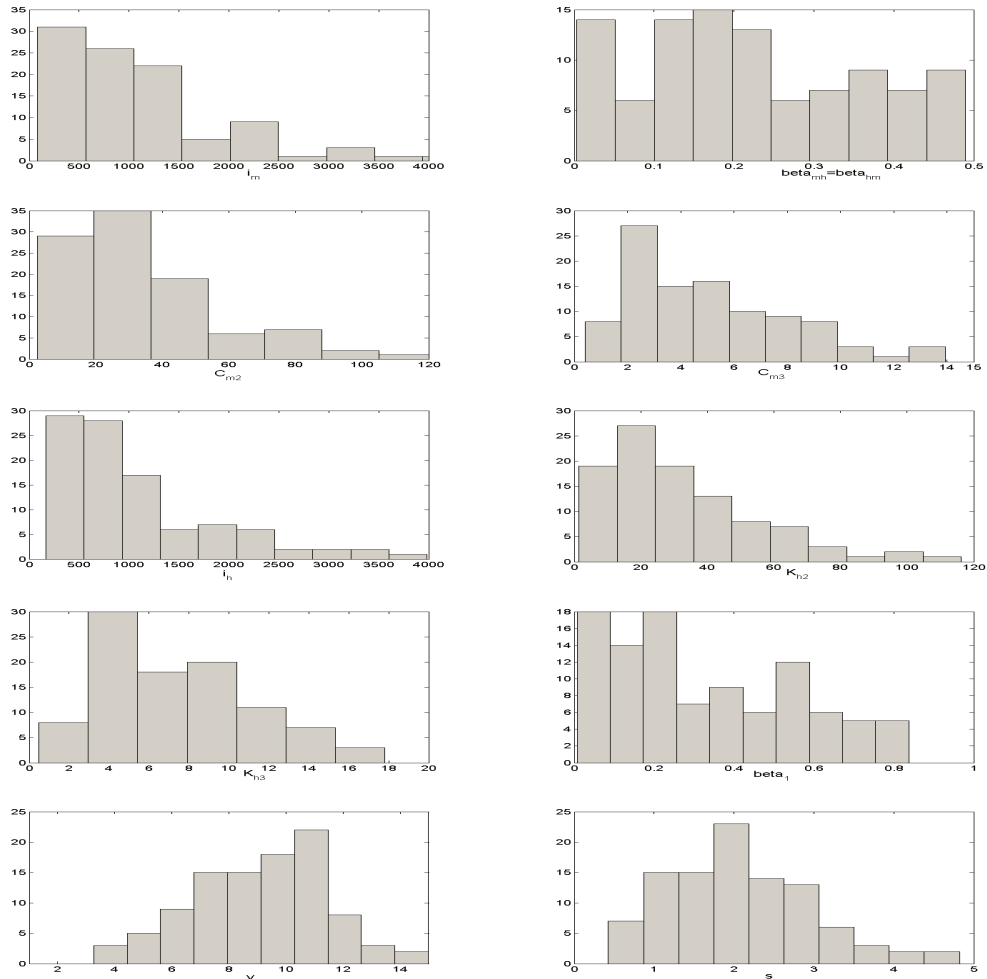


Figura A.5: Histograma de los parámetros de entrada del modelo $i_m, \beta_{mh}, C_{m2},$

$C_{m3}, i_h, K_{h2}, K_{h3}, \beta_1, v$ y s generado por el muestreo aleatorio $n = 100$.

Exploración del espacio muestral de los parámetros $n = 200$

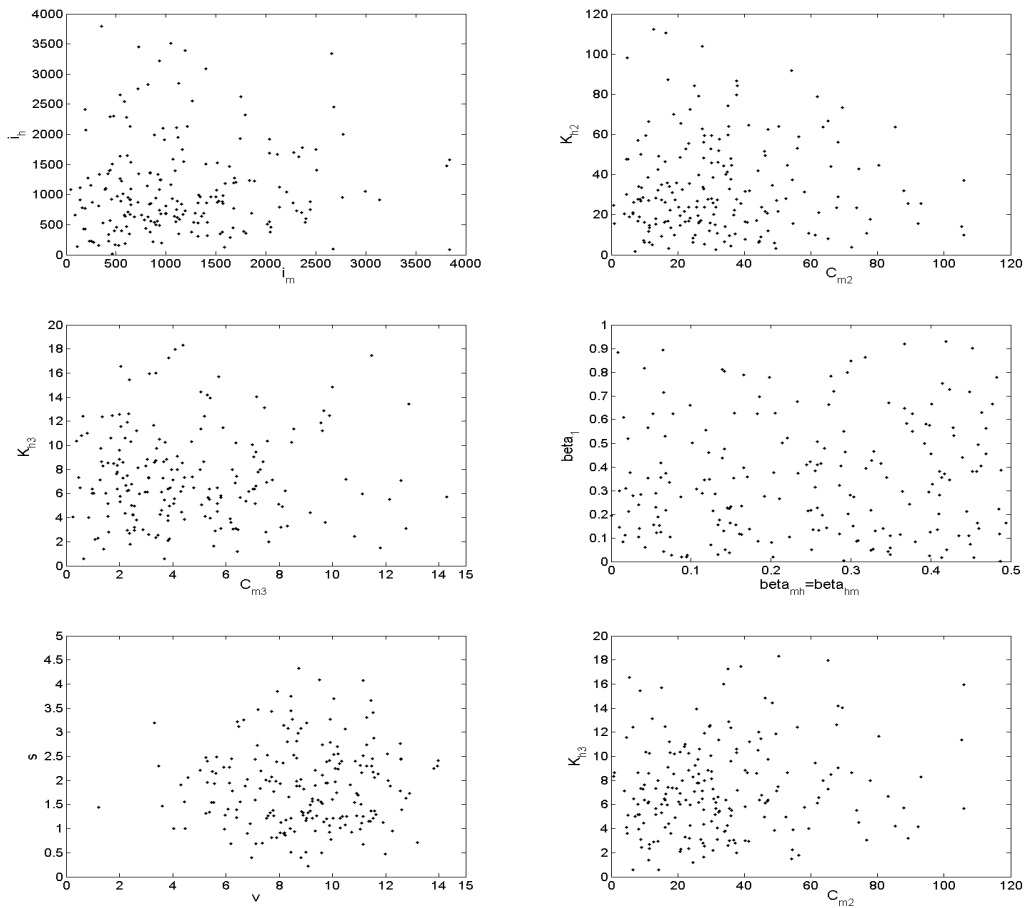


Figura A.6: Exploración del espacio muestral generado por el muestreo aleatorio $n = 200$, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2} vs K_{h3} .

Histograma de los parámetros de entrada $n = 200$

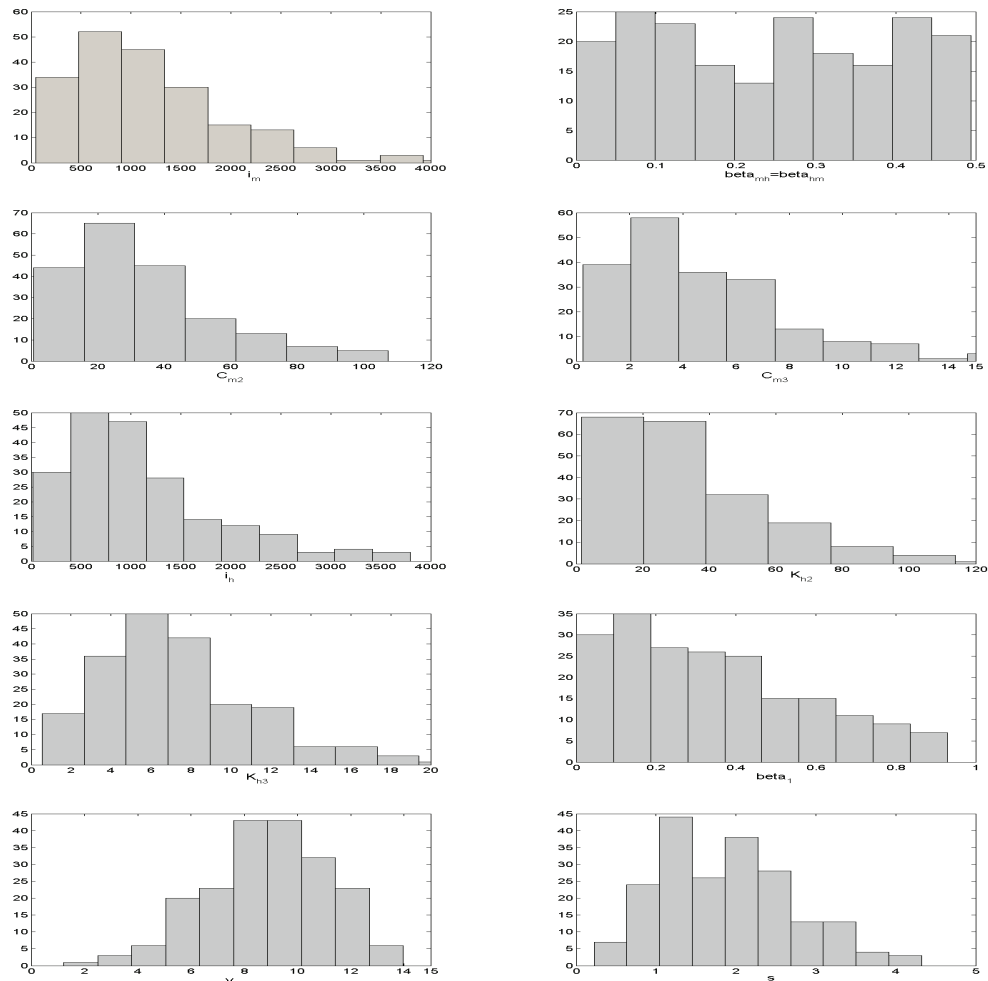


Figura A.7: Histograma de los parámetros de entrada del modelo $i_m, \beta_{mh}, C_{m2},$

$C_{m3}, i_h, K_{h2}, K_{h3}, \beta_1, v$ y s generado por el muestreo aleatorio $n = 200$.

A.2 Muestreo por hipercubo latino

Exploración del espacio muestral de los parámetros $n = 100$

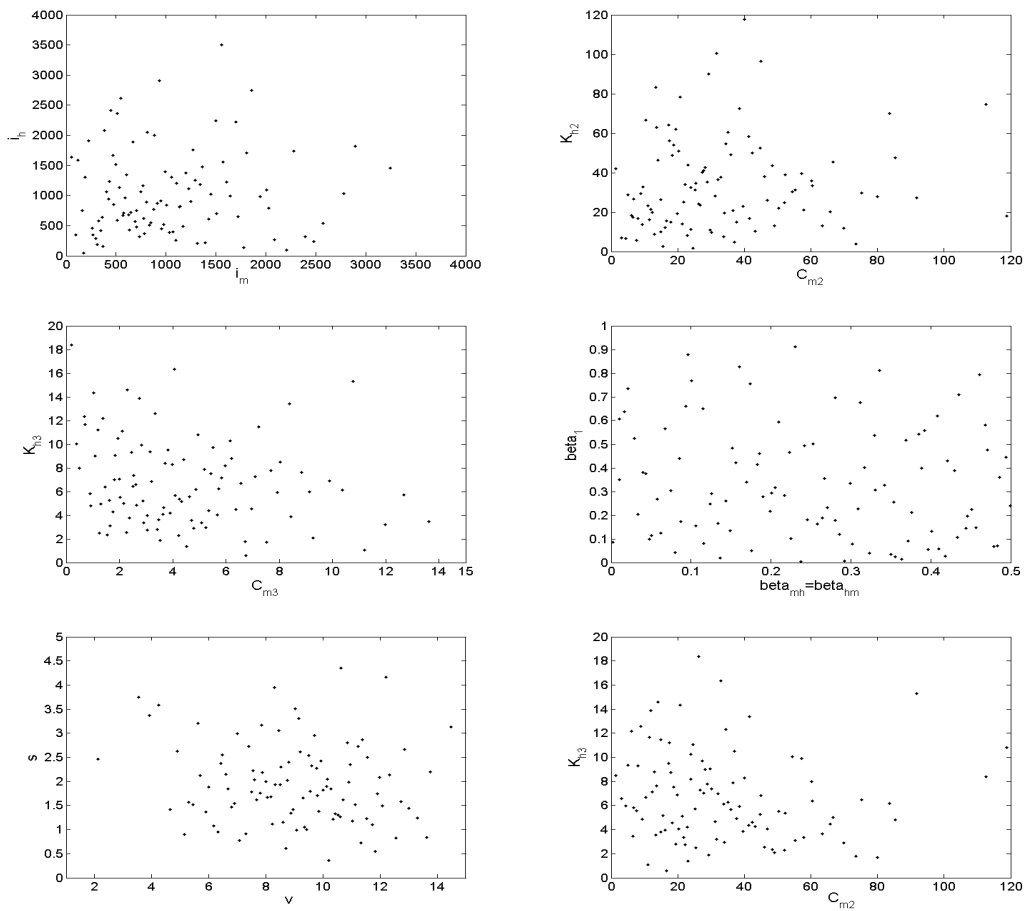


Figura A.8: Exploración del espacio muestral generado por el muestreo

hipercubo latino $n = 100$, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh}

vs β_l , v vs s y C_{m2} vs K_{h3} .

Histograma de los parámetros de entrada $n = 100$

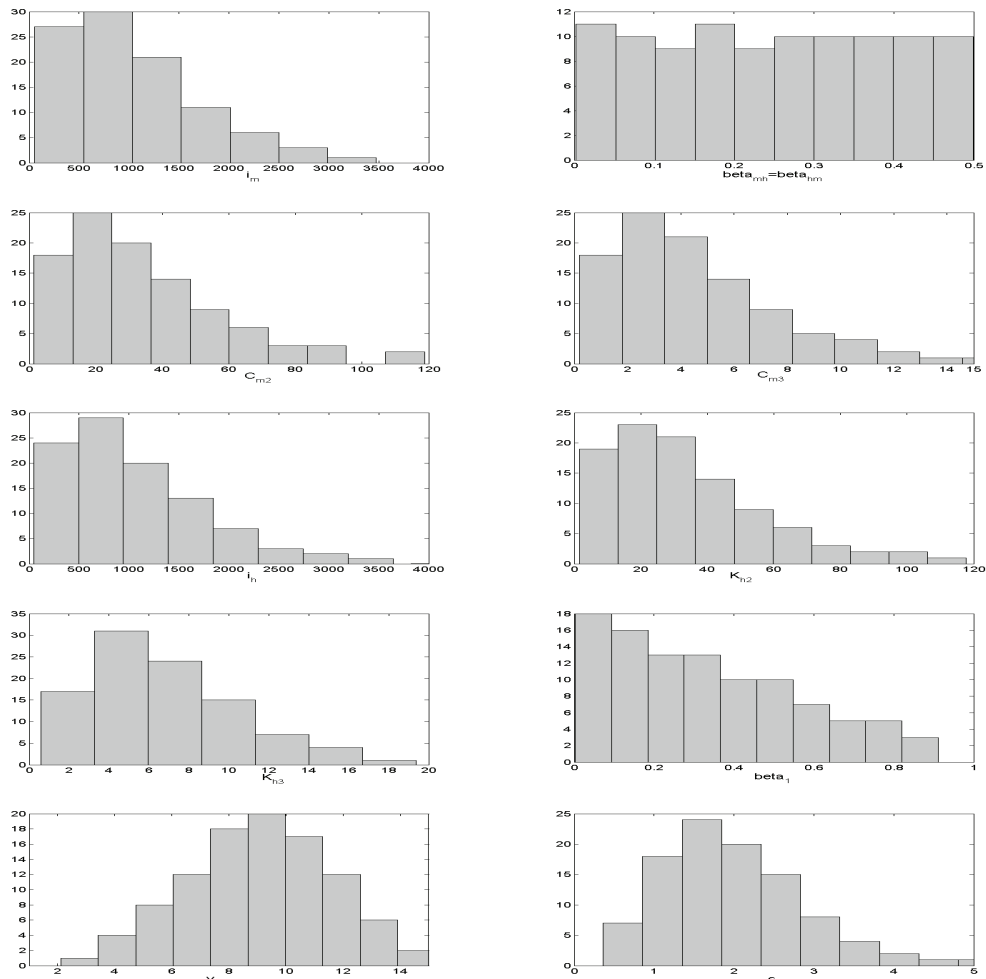


Figura A.9: Histograma de los parámetros de entrada del modelo $i_m, \beta_{mh}, C_{m2}, C_{m3}, i_h, K_{h2}, K_{h3}, \beta_1, v$ y s generado por el muestreo por hipercubo latino $n = 100$.

Exploración del espacio muestral de los parámetros $n = 200$

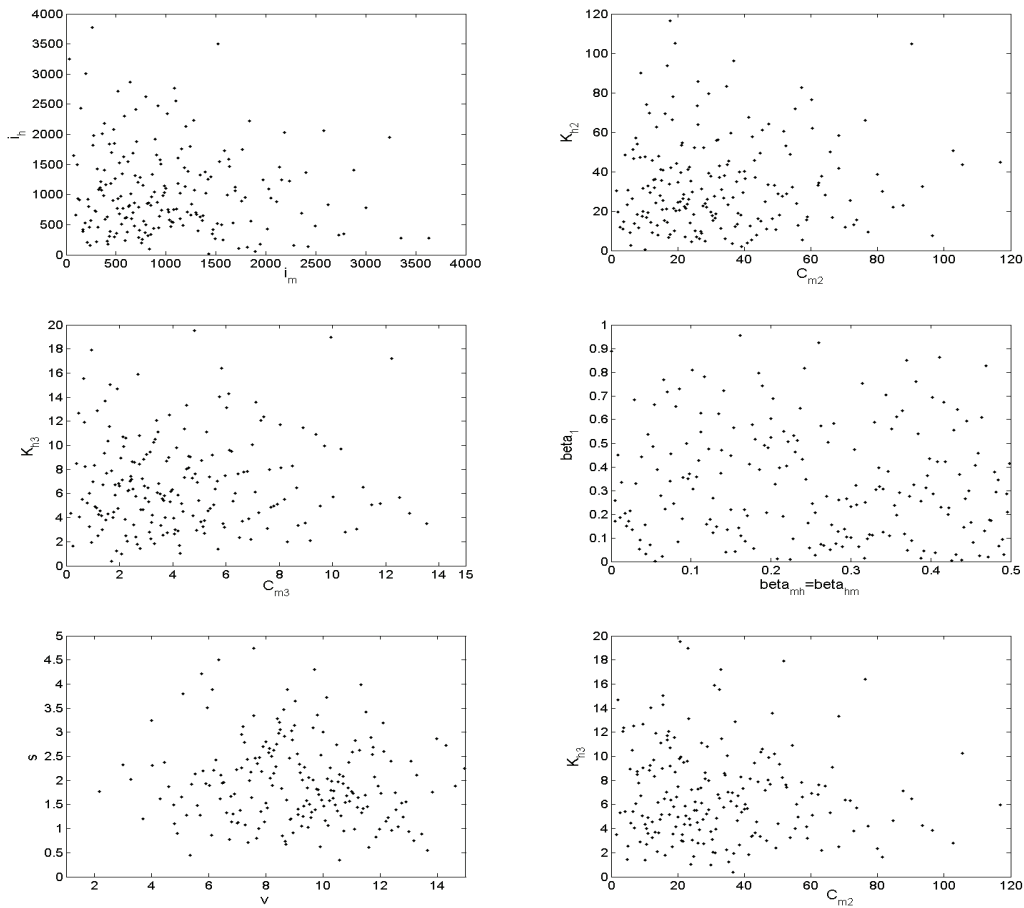


Figura A.10: Exploración del espacio muestral generado por el muestreo

hipercubo latino $n = 200$, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh}

vs β_1 , v vs s y C_{m2} vs K_{h3} .

Histograma de los parámetros de entrada $n = 200$

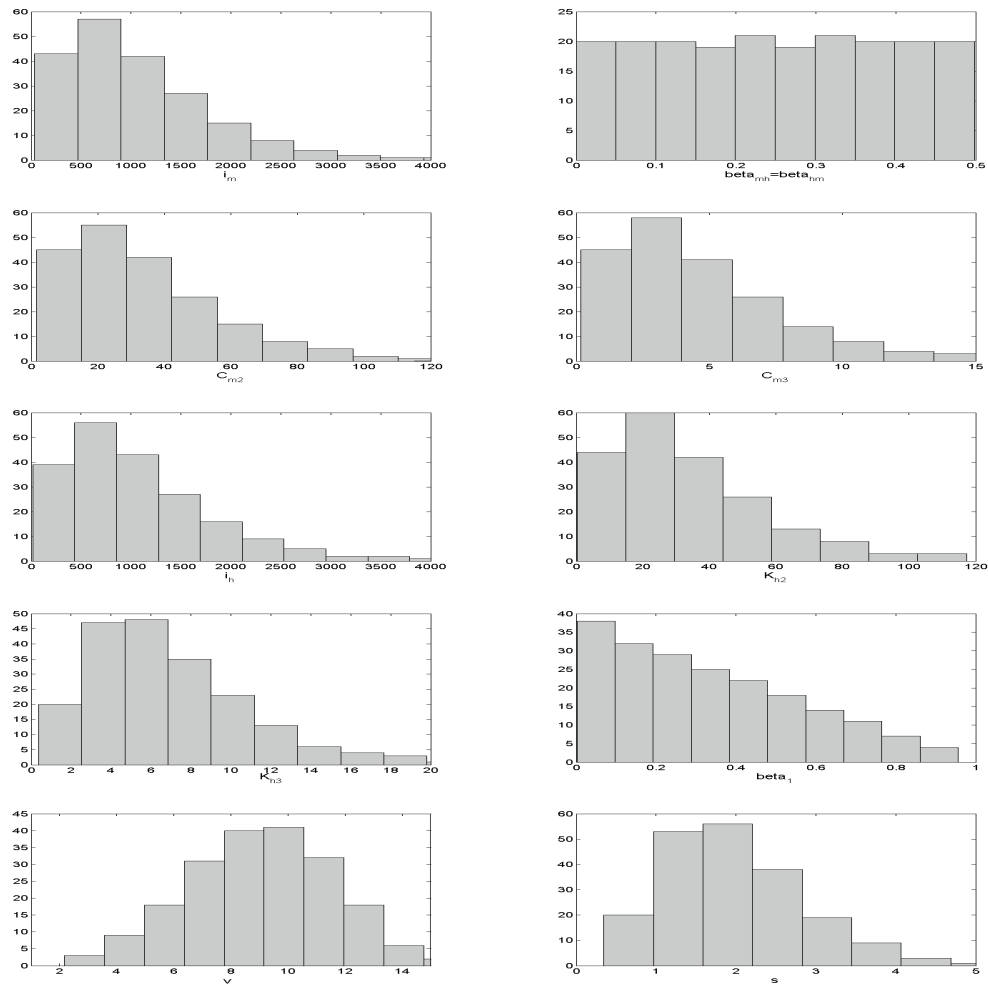


Figura A.11: Histograma de los parámetros de entrada del modelo $i_m, \beta_{mh}, C_{m2}, C_{m3}, i_h, K_{h2}, K_{h3}, \beta_1, v$ y s generado por el muestreo por hipercubo latino $n = 200$.

A.3 Muestreo por hipercubo latino modificado

Exploración del espacio muestral de los parámetros $n = 100$

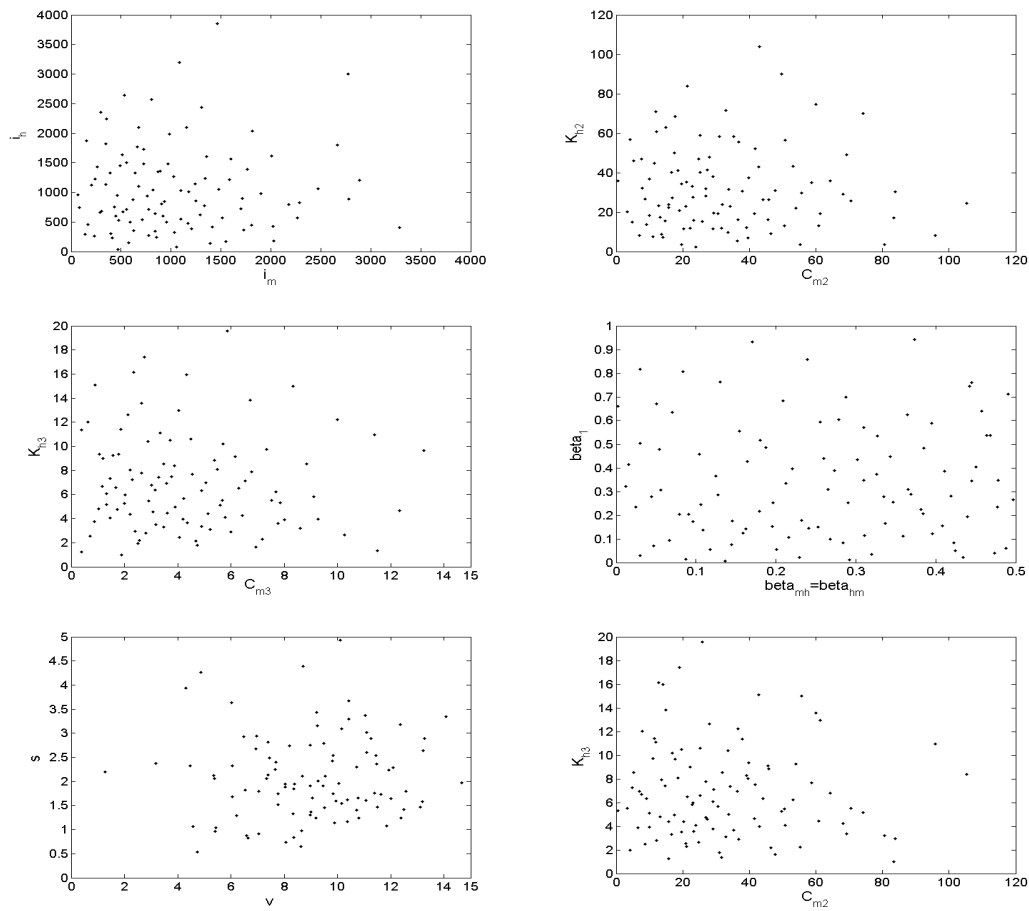


Figura A.12: Exploración del espacio muestral generado por el muestreo hipercubo latino modificado $n = 100$, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3} vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2} vs K_{h3} .

Histograma de los parámetros de entrada $n = 100$

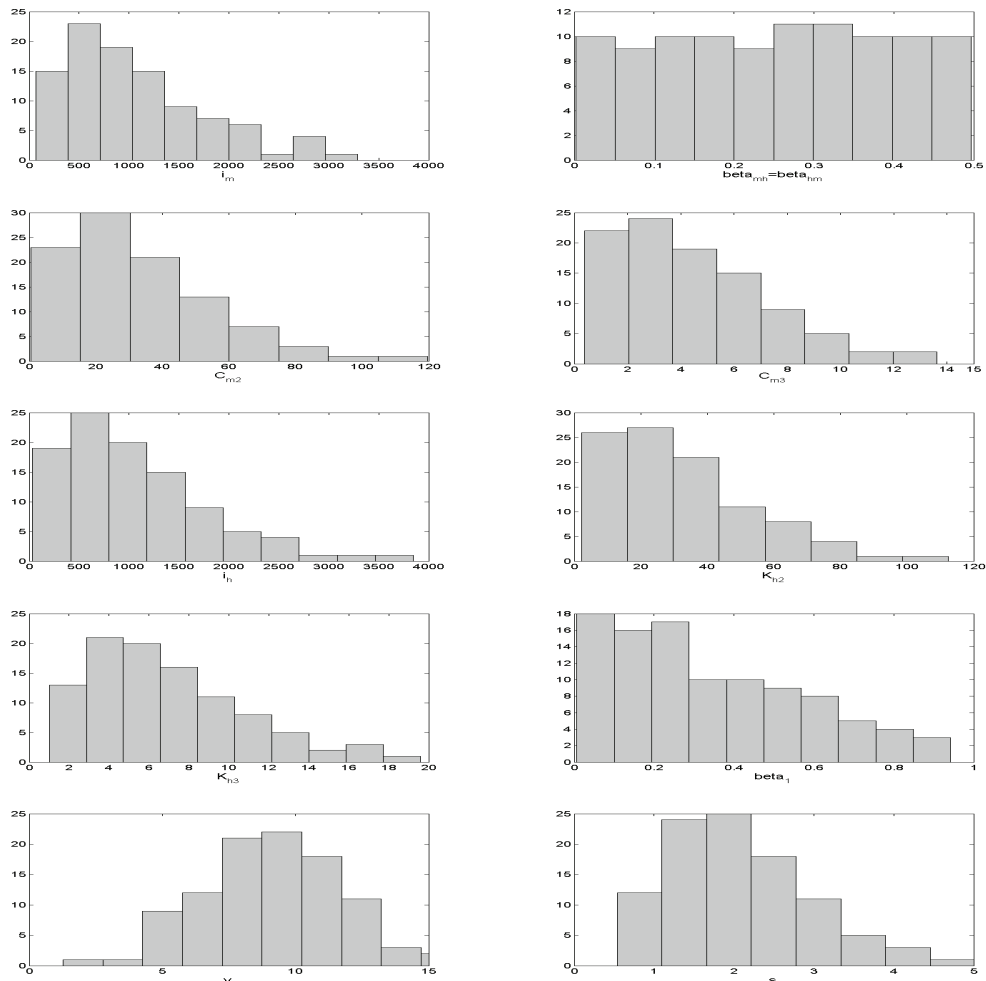


Figura A.13: Histograma de los parámetros de entrada del modelo $i_m, \beta_{mh}, C_{m2},$

$C_{m3}, i_h, K_{h2}, K_{h3}, \beta_1, v$ y s generado por el muestreo por hipercubo latino

modificado $n = 100$.

Exploración del espacio muestral de los parámetros $n = 200$

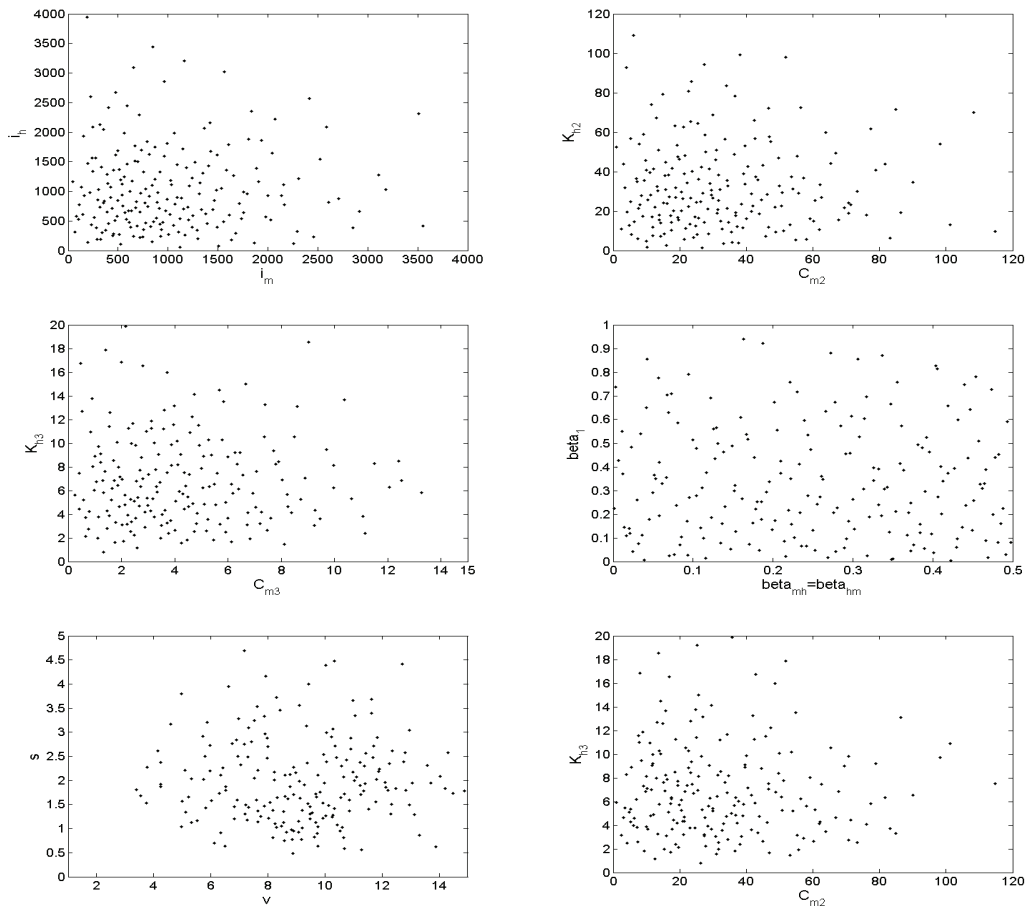


Figura A.14: Exploración del espacio muestral generado por el muestreo

hipercubo latino modificado $n = 200$, para los parámetros i_m vs i_h , C_{m2} vs K_{h2} , C_{m3}

vs K_{h3} , β_{mh} vs β_1 , v vs s y C_{m2} vs K_{h3} .

Histograma de los parámetros de entrada $n = 200$

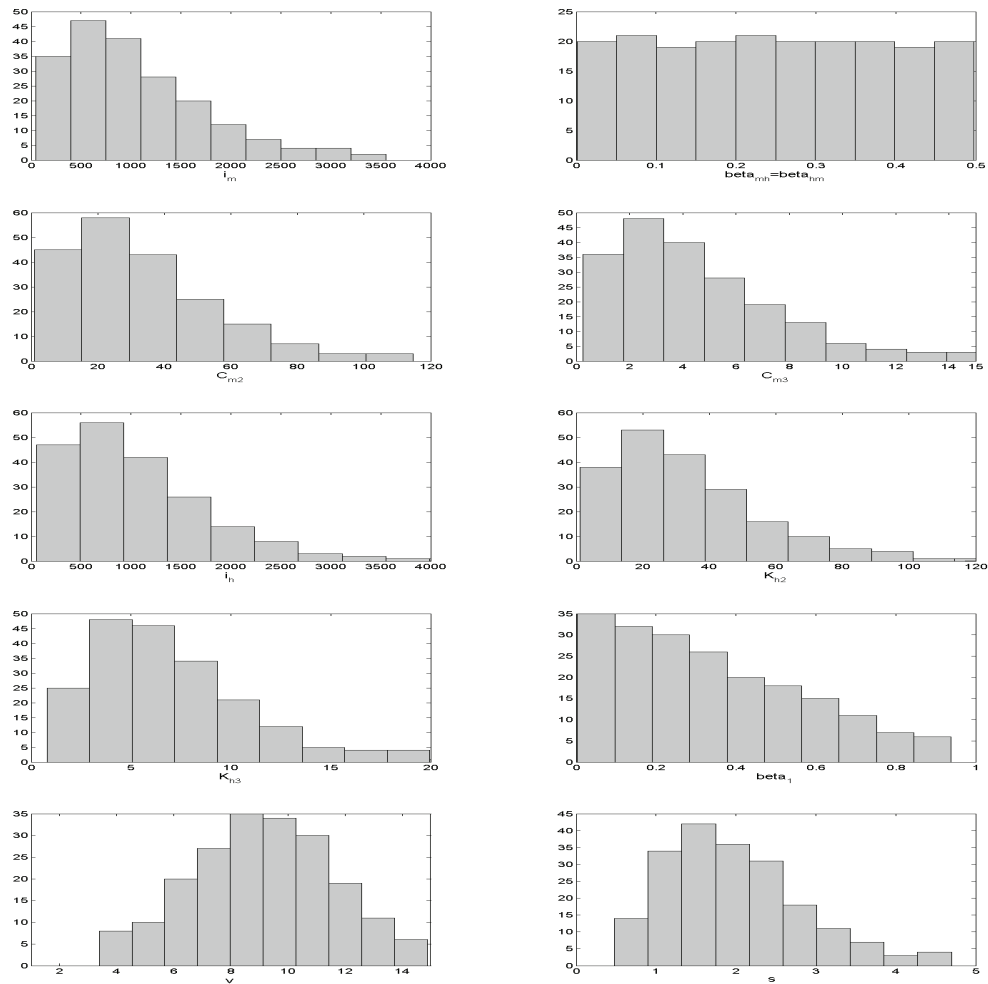


Figura A.15: Histograma de los parámetros de entrada del modelo $i_m, \beta_{mh}, C_{m2},$

$C_{m3}, i_h, K_{h2}, K_{h3}, \beta_1, v$ y s generado por el muestreo por hipercubo latino

modificado $n = 200$.

A.4 Datos de la Administración de Servicios de Salud Mental y Contra la Adicción (ASSMCA) año para el periodo de 1-julio-2009 hasta 30-junio-2010

A.1 Vía de utilización de droga

Modalidad de uso	Sexo			Total
	Femenino	Masculino	Transgénero	
Fumada	959	6,634	0	7,593
Intravenosa/inyectado	968	8,342	0	9,310
Nasal/inhalar	509	3,293	0	3,802
No aplica	9,905	10,637	1	20,543
No informó	19	39	0	58
Oral	718	7,602	0	8,320
Otro	0	2	0	2
Total	13,078	36,549	1	49,628

A.2 Diagnóstico de salud física por sexo de las personas que usan droga con
inyectable

Diagnóstico	Sexo		
	Femenino	Masculino	Total
Asma	195	730	925
Cáncer	9	28	37
Ceguera	0	4	4
Diabetes/hipoglicemia	26	215	241
Enfermedad cardiovascular	18	61	79
ETS	9	22	31
Epilepsia	6	19	25
Hepatitis A	9	95	104
Hepatitis B	19	200	219
Hepatitis C	172	1,875	2,047
Hipertensión	33	151	184
Infección con VIH	64	446	510
No aplica	338	4,073	4,411
Otro diagnóstico	62	371	410
Otro impedimento físico	3	15	18
Paciente de SIDA	3	28	31
Síndromes metabólicos	1	0	1

A.3 Promedio de años de uso de droga intravenosa

<hr/>	
Tiempo de	
uso	
<hr/>	
Media en años	19.0506
<hr/>	

Bibliografía

- [1] Administración de servicios de salud mental y contra la adicción (ASSMCA). Transtornos de sustancias y uso de servicios en Puerto Rico: Encuesta de Hogares-2008.
- [2] Álvarez Jurgenson, J. L. y Bravo García. E. (2006). SIDA: Aspectos de salud pública, de <http://www.salud.gob.mx/inidades/cdi/documentos/manualSIDA2006.pdf>.
- [3] Baggaley, R., White, R. G. & Boily, M. C. (2010). "HIV transmission risk through anal intercourse: systematic review, meta-analysis and implications for HIV prevention." *International Journal of Epidemiology* **39**, 1048 – 1063.
- [4] Boily, M. C., Poulin, R. & Masse, B. (2000). "Some methodological issues in the study of sexual networks." **12**, 558 – 571
- [5] Blythe, S., Busenberg, S. & Castillo Chavez, C. (1995). "Affinity in Paired Event Probability." *Mathematical Biosciences* **128**, 265 – 284.

- [6] Blower, S. M., Hartel, D., Dowlatabadi, H., Anderson, R. M. & May, R.M.(1991). "Drugs, sex and HIV: a mathematical model for New York City." *Phil. Trans. R. Soc. Lond. (Series B)* **331**, 171 – 187.
- [7] Blower, S. M. & Dowlatabadi, H. (1994). "Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: an HIV model, as an example." *International Statistical Review*, **62**, pp. 229 – 243. Printed in Great Britain.
- [8] Bobashev, G., Zule, W.A & Morris, R. J. (2010). "Projecting sexual and injecting HIV risks into future outcomes with agente-based modeling." RTI international.
- [9] Centers for disease control and prevention,
<http://www.cdc.gov/hiv/topics/surveillance/basic.htm>.
- [10] Chevret, S., Costagliola, D., Lefrere, J & Valleron, A. (1992). "A new approach to estimating AIDS incubation times: results in homosexual infected men." *Journal of Epidemiology and Community Health*. **46**, 582 – 586.
- [11] Deren, S., Kang, S. Y., Cólón, H. M., Andía J. F. & Robles, R.R. (2004). "HIV incidence among high-risk puertorican drug users: A comparison of East

Harlem, New York, and Bayamón, Puerto Rico." *J Acquir Immune Defic Syndr.* **36**, 10671074.

- [12] De Pereda Sebastian, D. (2010). Modelización matemática de la difusión de una epidemia de peste porcina entre granjas. Tesis doctoral.
- [13] Helton, J. C. & Davis, F. J.(2000). "Sampling-Based Methods for Uncertainty and Sensitivity Analysis". *Albuquerque, New Mexico 87185-0779*: Sandia National Laboratories.
- [14] Helton, J.C. Johnson, Sallaberry, C.J. & Storlied, C.B.(2006)."Survey of Sampling-Based Methods for uncertainty and Sensitivity Analysis". *Albuquerque, New Mexico 87185-0779*: Sandia National Laboratories.
- [15] Hernández Madonado, M. (1990). Efectividad de la guía para la enseñanza de la prevención del VIH/SIDA en niños del segundo y tercer grado en el municipio de Barceloneta.
- [16] Latka, M., Metsch, L. R., Mizuno, Y., Tobin, K., Mackenzie, S., Arnsten, J. H. & Gourevitch, M. N. (2006). "Unprotected sex among HIV-positive injection drug-using women and their serodiscordant male partners: role of personal and partnership influences." *J. Acquir Immune defic syndr.* **42**, 222 – 228.

- [17] Lee Pont, F. & Blower, S. M. (1991). "The supply and demand dynamics of sexual behavior: implications for heterosexual HIV epidemics." *Journal of acquired immune deficiency syndromes* **5**, 987 – 999.
- [18] Lee, L., Karon, J., selik, R., Neal, J. & Fleming, P. (2001). "Survival after AIDS diagnosis in adolescents and adults during the treatment Era, United Stated, 1984-1997." *Jama*. **285**, 1308 – 1315.
- [19] Lobato M. y Soto D.(2006). El SIDA en Puerto Rico de <http://www.tendenciaspr.com>.
- [20] Marrero E., Cruz, M. y Miranda, S.(2005). Perfil epidemiológico para la prevención del VIH en Puerto Rico.
- [21] Mckay, M. D., Conover, W. J. & Beckman, R. J. (1979). "A Comparison of three methods for selecting values of input variables in the analysis of output from a computer code". *Technometrics* **21**, 239 – 245.
- [22] Montesinos, O. y Hernández, C. (2007). "Modelos matemáticos para enfermedades infecciosas." *Salud Pública Mex* 2007; **49**, 218 – 226.
- [23] Okten, G., (2010). "Randomly permuted and random-started Halton Sequences."Florida State University, <http://www.okten@math.fsu.edu>.

- [24] Okten, G., Shah, M. & Goncharov, Y. (2009). "Random and deterministic digit permutations of the Halton Sequences." *Elsevier Science*, 1 – 20.
- [25] ONUSIDA: Informe sobre la epidemia mundial del SIDA (2009), <http://www.avert.org/estadisticas-sida.htm>.
- [26] Purcell, D. W., Mizuno, Y., Metsch, R. G., Tobin, K., Knight, K. & Latka, M. H. (2006). "Unprotected sexual behavior among heterosexual HIV-positive injection drug using men: Associations by partner type and partner serostatus." *Journal of Urban health: Bulletin of the New York academy of medicine*.
- [27] Resto Olivo, J. (2003). Evaluación de dos recaudación de fondos para una organización sin fines de lucro que ofrece servicios en el área de VIH/SIDA. Tesis doctoral.
- [28] Ríos Soto, K., Song, B. & Castillo Chavez, C. (2011). "Epidemic spread of influenza viruses: The impact of transient populations on disease dynamics." *Mathematical biosciences and engineering* **8**, 201 – 224.
- [29] Seaholm, S., Ackerman, E. & Shu-Chen Wu (1988). "Latin Hypercube Sampling and the sensitivity analysis of the Monte Carlo epidemic model." *International Journal of Biomedical Computing* **23**, 97 – 112.

- [30] Stein, M. (1987). "Large sample properties of simulations using latin hypercube sampling." *Technometrics* **29**, 143 – 151.
- [31] Stevens, S. & Andrade, R.(2006). "Teoría feminista: un marco teórico para la educación efectiva en la prevención del VIH en mujeres drogadictas." *Southwest Institute for research on Women*. University of Arizona. 151 – 161.
- [32] Soorapanth, S. & Chick, S. E. (2010). "Cost-Utility analysis of behavioral interventions for HIV-infected persons to reduce HIV transmission in the USA." 2433 – 2443.
- [33] Torres Velsquez B. C. (2009). La transmisión de la epidemia del VIH/SIDA en Puerto Rico: Modelo de Regresión. Tesis de maestría.
- [34] Tuffin, B.(1996). "A new permutation choice in Halton sequences, in: Monte Carlo and Quasi-Monte Carlo." *Springer Verlag, New York,1997*, **127**, 427 – 435.
- [35] Varas Díaz, N.(2002). Peligrosidad encarnada: Estigama y VIH/SIDA en Puerto Rico. Tesis doctoral
- [36] Vargas Jerez, J.C., Espitia Cuchango, H. E. y Galeano Ureña, C. H. (2010). "Ajuste de los modelos *S.I.R* y *S.E.I.R* empleando optimización para el modelamiento de la transmisión de la gripe A(H1N1)." *INGENIUM* **7** – 15.

[37] Wang, X. & Hickernell, F. J. (1998). "Randomized Halton Sequences."