



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

MODELO EPIDEMIOLÓGICO PROBABILISTA CON EXTENSIÓN A
GRUPOS METAPOBLACIONALES

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICO

P R E S E N T A :

CARLOS IGNACIO HERRERA NOLASCO

TUTOR

DR. MARCO ARIELI HERRERA VALDEZ



CIUDAD UNIVERSITARIA, Cd. Mx. ABRIL, 2020



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*Dedicado a Margarita Nolasco Tapia y Maria Nolasco Tapia.
Por todo, gracias.*

Agradecimientos

A la Facultad de Ciencias de la UNAM por la formación que recibí durante mi estancia en la carrera en matemáticas. A mis sinodales, por tomarse el tiempo de leer mi trabajo junto con todas sus observaciones. A mi asesor, Marco, por toda su ayuda y guía tanto fuera como dentro de las aulas, prometo mejorar todavía más.

Índice general

1. Resumen general	9
2. Modelos epidemiológicos tipo SIR	11
2.1. Modelación epidemiológica	11
2.2. Modelos SIR, SI y Metapoblacionales	11
2.2.1. El modelo SIR	12
2.2.2. El modelo SI	16
2.2.3. Extensión metapoblacional del SIR	17
2.2.4. El problema del SIR y sus derivados	20
3. Modelos discretos en poblaciones pequeñas	23
3.1. Visión probabilista de procesos epidemiológicos	23
3.1.1. Cadenas de transmisión como procesos de muestreo	23
3.1.2. Variables aleatorias y reglas de aleatoriedad	24
3.2. El modelo PSIR	28
3.2.1. Los modelos de Greenwood y Reed-Frost	35
4. Resultados	41
4.1. PSIR en poblaciones grandes y determinismo	41
4.2. Convergencia del PSIR a un Proceso Poisson no Homógeneo en el tiempo	44
4.3. Análisis de la probabilidad de infección en el PSIR	46
4.4. Modelo aplicado a influenza	52
4.5. Modelo aplicado a VIH	54
4.5.1. Extensión para considerar grupos de riesgo	54
5. Discusión	55
5.1. Ventajas	55
5.2. Desventajas	56
5.3. Posibles extensiones del PSIR	56

Capítulo 1

Resumen general

En este trabajo desarrollo un modelo epidemiológico basado en un muestreo probabilista cuyo interés es describir la cadena de transmisión de enfermedades infecciosas, principalmente en poblaciones pequeñas.

La motivación para realizar esto es que el modelo SIR y sus modificaciones, entre los modelos deterministas más comúnmente usados, se basan en supuestos matemáticos que rara vez concuerdan con la realidad. El SIR describe la cadena de transmisión de una enfermedad infecciosa mediante un sistema de ecuaciones diferenciales, basado en supuestos como el de un tamaño poblacional grande que permita usar densidades poblacionales así como una mezcla homogénea entre los individuos de la población. Esto hace que no se pueda usar en poblaciones pequeñas o con mezcla heterogénea, lo cual casi siempre ocurre.

Es importante plantear un modelo que sea adaptable a cualquier tamaño poblacional, siendo discreto en los valores que toma y cuyo desarrollo matemático este basado en el sentido común. Una manera de hacerlo es partir la población en individuos susceptibles, infectados y removidos como en el SIR y considerar que el número de nuevos infectados y de nuevos removidos por unidad de tiempo son muestreos binomiales que generan variables aleatorias con ciertos parámetros, para luego describir la cadena de transmisión mediante un sistema de ecuaciones parecido en su estructura al SIR pero en términos de dichas variables aleatorias. El sistema de ecuaciones será llamado modelo PSIR.

Se comenzará por revisar el modelo SIR determinista, y algunas modificaciones del modelo tales como el SI y SIR metapoblacional, esto con la intención de plantear ideas generales con las que se trabajará para generar el pSIR para luego extenderlo al caso de los grupos de riesgo. Por lo tanto, los resultados generados a partir de los modelos deterministas forman parte de la literatura conocida en textos sobre modelación epidemiológica.

El interés en los grupos de riesgo recae en como afecta su tamaño y sus contactos infecciosos a la cadena de transmisión. Por ejemplo, en el VIH existen grupos poblacionales en los que la probabilidad de infección es mayor, como en el caso de los hombres que tienen sexo con hombres (HSH) y los trabajadores sexuales (TS) pues la principal forma de contagio

es mediante el contacto sexual.

El planteamiento del modelo PSIR, PSI, la extensión metapoblacional, el coeficiente análogo al R_0 del modelo, el análisis a la probabilidad de infección así como, en general, los resultados generados a partir del modelo son un trabajo en conjunto entre mi asesor Marco Arieli Herrera Valdez, el revisor de mi trabajo, Sergio Iván López Ortega, en el planteamiento de los Procesos Poisson no Homogéneos como proceso límite del PSIR, y yo. Sin embargo se compara el modelo generado por nosotros con modelos parecidos de tipo estocástico basados en cadenas Binomiales como el modelo de Greenwood y Reed-Frost.

Una vez planteado el modelo PSIR se plantea su uso para influenza.

Por último se discuten las ventajas y desventajas del uso del PSIR, así como posibles modificaciones de este.

Capítulo 2

Modelos epidemiológicos tipo SIR

2.1. Modelación epidemiológica

La epidemiología es una rama de la ciencia médica dedicada al estudio de la distribución de enfermedades en las poblaciones humanas con el propósito de implementar medidas que reduzcan tanto la proporción de individuos que presentan la enfermedad (prevalencia) como la cantidad de nuevos individuos infectados por unidad de tiempo (incidencia).

Por otro lado, los modelos matemáticos que describen la propagación de las enfermedades infecciosas en una población permiten contar con herramientas que ayudan a predecir el comportamiento de la enfermedad en la población a largo plazo. Uno de estos modelos es el modelo SIR, que describe la dinámica de enfermedades infecciosas clasificando a los individuos como susceptibles, infectados y removidos, siendo los susceptibles aquellos propensos a adquirir la enfermedad, los infectados aquellos que participan como propagadores de la enfermedad en la cadena de transmisión y removidos aquellos individuos que no participan en la cadena.

2.2. Modelos SIR, SI y Metapoblacionales

No todas las enfermedades pueden ser modeladas por una dinámica del tipo susceptible-infectado-removido. Por ejemplo, para el virus de inmunodeficiencia humana (VIH), la dinámica poblacional es del tipo susceptible-infectado (modelo SI). También hay otras enfermedades infecciosas cuya dinámica es del tipo susceptible-infectado-susceptible (modelo SIS), etc. De interés particular, estos modelos se pueden extender considerando subpoblaciones en donde las enfermedades infecciosas tengan una mayor incidencia que en el resto de la población (modelos metapoblacionales).

En este trabajo se consideran tres modelos, SIR, SI, y una extensión metapoblacional del SIR, por lo que ahora nos corresponde plantear cada uno de estos.

2.2.1. El modelo SIR

Para el modelo SIR se supone que hay una mezcla homogénea en la población, es decir, cualesquiera dos individuos pueden tener contacto entre sí y en el caso de que uno de ellos sea susceptible y otro infectado puede ocurrir un contagio. Además, no considera factores como las tasas de natalidad y mortalidad o migración por lo que el tamaño poblacional permanece constante y dicho tamaño es lo suficientemente grande como para plantear una dinámica de propagación en términos de densidades poblacionales y por tanto las variables sean continuas.

Se comienza considerando una población de individuos de tamaño constante N . Dada una enfermedad infecciosa los individuos se clasifican de acuerdo a la condición clínica en la que se encuentran, siendo $S(t)$, $I(t)$, $R(t)$ la cantidad de susceptibles, infectados y removidos al tiempo t respectivamente.

La cadena de transmisión comienza cuando un individuo susceptible tiene contacto con un infectado, pasando al estado infectado y volviéndose propagador de la enfermedad para luego de un tiempo pasar al estado de removido en donde ya no participa en la cadena de transmisión de manera activa. La población constante en el tiempo permite la relación

$$S(t) + I(t) + R(t) = N. \quad (2.1)$$

Al normalizar con respecto a N se obtiene

$$\frac{S(t)}{N} + \frac{I(t)}{N} + \frac{R(t)}{N} = 1. \quad (2.2)$$

Los cocientes resultan ser las proporciones de la población que se encuentran en dichos estados. Con la finalidad de no sobrecargar la notación nos referiremos a dichas proporciones como $x(t)$, $y(t)$ y $z(t)$. De nuevo, éstas son las proporciones de individuos susceptibles, infectados y removidos dentro de la población que con base en la suposición de homogeneidad en los contactos entre individuos se puede considerar a $x(t)$, $y(t)$, $z(t)$ como las probabilidades encontrar a individuos susceptibles, infectados o removidos en la población. Así, la ecuación (2.2) toma la forma

$$x(t) + y(t) + z(t) = 1. \quad (2.3)$$

Para un intervalo de tiempo $h > 0$ los cambios en las proporciones $x(t)$, $y(t)$ y $z(t)$ con respecto t son

$$\frac{\Delta x}{h} = \frac{x(t+h) - x(t)}{h} \quad (2.4)$$

$$\frac{\Delta y}{h} = \frac{y(t+h) - y(t)}{h} \quad (2.5)$$

$$\frac{\Delta z}{h} = \frac{z(t+h) - z(t)}{\Delta t}. \quad (2.6)$$

Recordando que un individuo pasa de susceptible a infectado una vez que haya contraído la enfermedad, y luego pasa a ser removido cuando ya no forme parte de la cadena de transmisión.

Si la tasa con la que un individuo susceptible pasa a ser infectado es un valor $\lambda > 0$ entonces λx es la tasa de infección dentro de la población. Además λ puede ser escrita como $\lambda = \beta y$ con $\beta > 0$ siendo el producto entre la tasa de contacto entre individuos y la probabilidad de transmisión de la enfermedad. Es decir, βxy es la tasa de infección dentro de la población. Por otro lado, sea $\gamma > 0$, donde γ es la tasa con la que se remueve un individuo del sistema, entonces γy es la tasa de remoción dentro de la población.

Igualando las tasas de infección y remoción a los cambios en las proporciones con respecto al tiempo tenemos el sistema

$$\frac{\Delta x}{h} = -\beta xy \quad (2.7)$$

$$\frac{\Delta y}{h} = \beta xy - \gamma y \quad (2.8)$$

$$\frac{\Delta z}{h} = \gamma y. \quad (2.9)$$

El cual es un sistema de ecuaciones en diferencias que describen como se da la propagación de la enfermedad en la población.

Al hacer el tamaño poblacional lo suficientemente grande como para que las proporciones de individuos puedan ser tomadas como densidades poblacionales y tomando el límite cuando $h \rightarrow 0$, los cambios en las densidades se convierten en cambios instantáneos en el tiempo

$$\lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h} = -\beta xy \quad (2.10)$$

$$\lim_{h \rightarrow 0} \frac{y(t+h) - y(t)}{h} = \beta xy - \gamma y \quad (2.11)$$

$$\lim_{h \rightarrow 0} \frac{z(t+h) - z(t)}{h} = \gamma y. \quad (2.12)$$

El lado izquierdo de las ecuaciones son las derivadas con respecto al tiempo de las densidades poblacionales

$$\partial_t x = -\beta xy \quad (2.13)$$

$$\partial_t y = \beta xy - \gamma y \quad (2.14)$$

$$\partial_t z = \gamma y. \quad (2.15)$$

Para las curvas de susceptibles, infectados y removidos generadas por el modelo SIR (Figura 2.1), el signo en la ecuación (2.13) es debido a que la densidad de susceptibles baja conforme se van infectando por lo que la tasa de cambio es negativa.

La ecuación (2.14) describe como aumenta la densidad de infectados debido a la tasa de

infección y luego este crecimiento es contrarrestado por la tasa de remoción. El cambio en la densidad de removidos está descrito por la tasa de remoción en la ecuación (2.15).

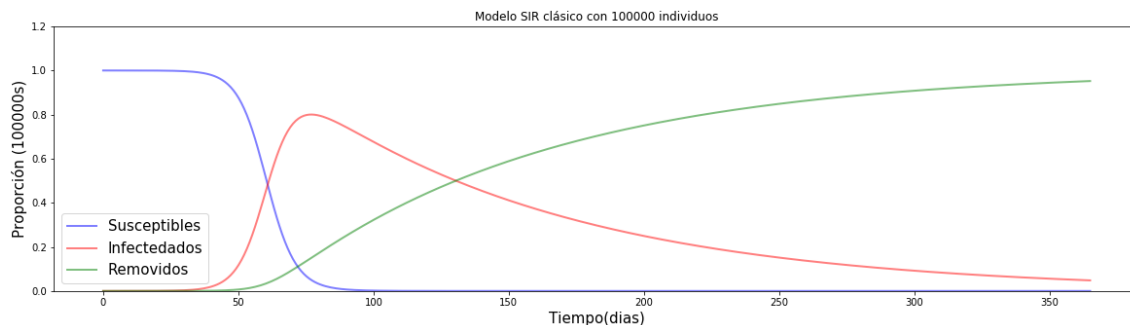


Figura 2.1: Dinámica del modelo SIR para una población de tamaño $N = 100000$, $\beta = 0.2$, $\gamma = 1/100$.

Este sistema de ecuaciones diferenciales es parecido al que plantearon Kermack y McKendrick (Kermack and McKendrick, 1927)

$$\partial_t x = -kxy \quad (2.16)$$

$$\partial_t y = kxy - ly \quad (2.17)$$

$$\partial_t z = lz. \quad (2.18)$$

para describir la propagación de las enfermedades infecciosas con tasas de infección y remoción constantes para poblaciones grandes, de población constante y con mezcla homogénea, con la diferencia que $x + y + z = N$.

Análisis cualitativo y propiedades del SIR

Como condiciones iniciales del modelo suponemos que al tiempo $t = 0$ las densidades de susceptibles e infectados son mayores que cero mientras que la proporción de removidos es cero. Es decir $x(0) > 0$, $y(0) > 0$ y $z(0) = 0$. Por otro lado, dentro del modelo, la tasa de remoción γ puede ser tomado como el inverso multiplicativo del periodo promedio de infección. Es decir, $\gamma = 1/\text{tiempo promedio de remoción}$

Un aspecto importante del SIR es que el sistema de ecuaciones pueden ser utilizado para determinar el comportamiento de la cadena de infección cuando el tiempo tiende a infinito, así como dar criterios que determinen cuando hay una epidemia. Si inicialmente $x(0)$ se aproxima a 1, para un individuo infectado se define el *número básico de reproducción*, como el número de nuevos casos de infección producidos por un infectado dentro de una población casi totalmente susceptible ($S \approx N$), $R_0 = \frac{S\beta}{N\gamma} \approx \frac{\beta}{\gamma}$. El número básico de reproducción

es muy importante en el análisis epidemiológico pues determina que un brote epidémico puede ocurrir cuando $R_0 > 1$ (Brauer et al., 2001). El radio básico de reproducción y la condición de epidemia pueden deducirse cuando el cambio en los infectados sea positivo, reescribiendo la ecuación (2.17) como

$$\partial_t y = y(\beta x - \gamma). \quad (2.19)$$

Si $\partial y(t) > 0$ entonces $y(\beta x - \gamma) > 0$, dado que en $t = 0$ tenemos $y(0) > 0$ se tiene que $\beta x(0) - \gamma > 0$ de donde $x(0) > \frac{\gamma}{\beta}$ de donde $x(0) > \frac{1}{R_0}$. Si la densidad de susceptibles aproximada a 1, entonces la condición inicial para que haya una epidemia esta determinado por $1 < R_0$.

Se pueden encontrar ecuaciones para el cambio en alguna de las variables en términos de otra, por ejemplo, tomando el cambio de los infectados con respecto a los susceptibles se tiene $\partial y(x) = \frac{\beta xy - \gamma y}{-\beta xy} = \frac{\gamma}{\beta x} - 1$, de donde

$$\partial_y x = \frac{1}{R_0 x} - 1. \quad (2.20)$$

Para el cambio en los susceptibles con respecto a los removidos, $\partial_x z = \frac{-\beta xy}{\gamma y} = -\frac{\beta}{\gamma}x$, la cual resulta ser una ecuación diferencial de variables separables $\partial x(z) = -R_0 x$ cuya solución es

$$x(t) = x(0)e^{-R_0 z(t)}. \quad (2.21)$$

que describe un decaimiento exponencial en los susceptibles con una tasa proporcional al producto entre la proporción de removidos en el sistema $z(t)$ y el cociente de la tasa de infección entre la tasa de remoción.

El número básico de reproducción, R_0 , se puede pensar como

$$\frac{\beta}{\gamma} = \frac{(\text{Tasa de contacto})(\text{Probabilidad de transmisión})}{\text{Tasa de retiro}}$$

o bien

$$\beta \left(\frac{1}{\gamma} \right) = (\text{Tasa de contacto})(\text{Probabilidad de transmisión})(\text{Periodo de infección})$$

Dado que β y γ son constantes, R_0 es una constante por lo que $R_0 x(t)$ es una función decreciente pues la densidad de susceptibles siempre baja debido a que pasan a ser infectados y estos no regresan al estado susceptible, debido a que el brote epidémico en $t = 0$ se da si $x(0) > \frac{\gamma}{\beta} = \frac{1}{R_0}$ es equivalente decir que $R_0 x(0) > 1$. Extendiendo el resultado a cualquier tiempo t , la desigualdad $R_0 x(t) > 1$ da un criterio para determinar hasta que momento continua la epidemia en la población.

Comportamiento límite La población está libre de la enfermedad si para algún tiempo t la densidad de infectados es igual a cero. Llamemos a esto $y(\infty) = 0$, de donde $x(\infty) + z(\infty) = 1$. Combinando esto con la ecuación (2.21) tenemos $x(\infty) = 1 - z(\infty) = x(0)e^{-R_0z(\infty)}$. Despejando se obtiene que

$$1 - z(\infty) - x(0)e^{-R_0z(\infty)} = 0. \quad (2.22)$$

La ecuación no tiene solución exacta pero es posible aproximarla mediante métodos numéricos.

Si t es lo suficientemente grande para que el lado derecho de la ecuación aproxime mejor a cero puede entonces ser tomado como el tiempo total de la epidemia.

El modelo SIR proporciona una forma para describir la propagación de una enfermedad infecciosa sobre una población, siendo de gran utilidad para estimar la proporción de individuos susceptibles que son necesarios para comenzar una epidemia así como el tamaño final de esta.

2.2.2. El modelo SI

El modelo SI surge cuando la enfermedad no admite estado removido por lo que los individuos únicamente pueden estar en los estados susceptibles e infectados.

La ecuación para las densidades toma la forma $x(t) + y(t) = 1$ y por lo tanto el sistema de ecuaciones diferenciales asociado a la dinámica es (Figura 2.2)

$$\partial_t x = -\beta xy \quad (2.23)$$

$$\partial_t y = \beta xy. \quad (2.24)$$

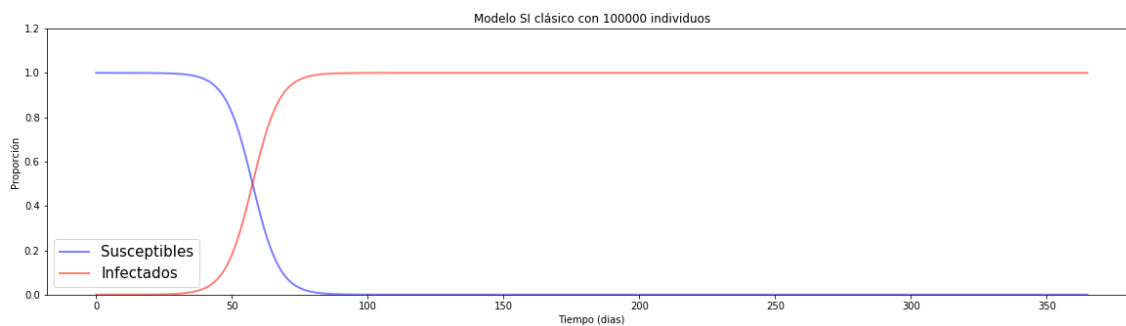


Figura 2.2: Dinámica del modelo SI para una población de tamaño $N = 100000$, $\beta = 0.2$, $\gamma = 1/100$.

Dado que $y = 1 - x$, sustituyendo se tiene que

$$\partial_t x = -\beta x(1 - x) \quad (2.25)$$

$$\partial_t y = \beta x(1 - x). \quad (2.26)$$

Se tiene entonces una descripción completa del sistema únicamente con base en el cambio en los susceptibles, la cual es una ecuación diferencial logística cuya solución está dada por

$$x(t) = \frac{x(0)}{e^{\beta t} + 1}. \quad (2.27)$$

2.2.3. Extensión metapoblacional del SIR

La extensión metapoblacional del SIR considera la existencia de subpoblacionales que tienen contacto entre sí. Estas subpoblaciones pueden hacer que la cadena de transmisión se vea afectada dependiendo de que tan grandes son dichos grupos, sus respectivas tasas de infección y los contactos entre individuos de diferentes grupos.

Una forma de visualizar el contacto entre subpoblaciones es considerar un grafo con una cantidad n de nodos, cada nodo representando una subpoblación que se conecta con el resto de ellos incluyéndose a si mismo y tiene asociado una cantidad relativa de individuos S_i, I_i, R_i para $i \in \{1, \dots, n\}$.

Otra manera de visualizarlo es pensar a la población como un conjunto y realizar una partición del conjunto en n subconjuntos ajenos, representando cada uno de ellos las subpoblaciones, y a su vez, cada subconjunto está dividido en susceptibles, infectados y removidos. Se considera una población de tamaño N constante, S la cantidad de individuos susceptibles en la población y n subpoblaciones con diferentes tasas de infección.

Sea $S_i(t)$ la cantidad de individuos susceptibles en el grupo i al tiempo t , con $i \in \{1, \dots, n\}$. Los susceptibles totales en la población se pueden descomponer en términos de los $S_i(t)$ como $S(t) = \sum_{i=1}^n S_i(t)$. Normalizando con respecto al tamaño poblacional N se tiene $x(t) = \frac{1}{N} \sum_{i=1}^n S_i(t) = \sum_{i=1}^n \frac{S_i(t)}{N} = \sum_{i=1}^n x_i(t)$. Los términos $x_i(t)$, cuando N es muy grande, son las densidades poblacionales relativas de cada uno de las n subpoblaciones existentes.

Análogamente para cada subpoblación i con $i \in \{1, \dots, n\}$, las densidades relativas para los infectados y removidos son $y_i(t)$ y $z_i(t)$.

Para $i \in \{1, \dots, n\}$, el cambio en la densidad de susceptibles $x_i(t)$ es igual a la tasa con la que se pasa a ser de ser susceptible a infectado dentro de la misma subpoblación, dicha tasa es la suma de los productos $\beta_{ij} x_i y_j$, donde las β_{ij} son las tasas de contacto entre la i -ésima subpoblación susceptible y cada una de las j con $j \in \{1, \dots, n\}$ subpoblaciones infectadas y $x_i y_j$ es la respectiva tasa de contagio entre la subpoblación i susceptible y la subpoblación j infectada.

El cambio en la densidad de infectados y_i es la tasa con la que llegan infectados a la subpoblación i menos la tasa con la que estos son removidos.

Por último, el cambio en la densidad de removidos z_i es la tasa con la que en la subpoblación i los individuos se remueven.

La extensión metapoblacional del SIR queda determinada por el sistema de ecuaciones

$$\partial_t x_i = - \sum_{j=1}^n \beta_{ij} x_i y_j \quad (2.28)$$

$$\partial_t y_i = \sum_{j=1}^n \beta_{ij} x_i y_j - \gamma_i y_i \quad (2.29)$$

$$\partial_t z_i = \gamma_i y_i. \quad (2.30)$$

El caso más sencillo del modelo es considerar dos subpoblaciones, en donde las ecuaciones generadas son

$$\partial_t x_1 = -(\beta_{11} x_1 y_1 + \beta_{12} x_1 y_2) \quad (2.31)$$

$$\partial_t y_1 = \beta_{11} x_1 y_1 + \beta_{12} x_1 y_2 - \gamma_1 y_1 \quad (2.32)$$

$$\partial_t z_1 = \gamma_1 y_1 \quad (2.33)$$

$$\partial_t x_2 = -(\beta_{21} x_2 y_1 + \beta_{22} x_2 y_2) \quad (2.34)$$

$$\partial_t y_2 = \beta_{21} x_2 y_1 + \beta_{22} x_2 y_2 - \gamma_2 y_2 \quad (2.35)$$

$$\partial_t z_2 = \gamma_2 y_2. \quad (2.36)$$

Con la finalidad de hacer más explícita la extensión del SIR al modelo metapoblacional, se definen los vectores de densidades de susceptibles, infectados y removidos $x(t)$, $y(t)$, $z(t)$ y sus vectores de derivadas $\partial x(t)$, $\partial y(t)$, $\partial z(t)$ respectivamente como

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}, \quad z(t) = \begin{pmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ z_n(t) \end{pmatrix},$$

$$\partial_t x = \begin{pmatrix} \partial_t x_1 \\ \partial_t x_2 \\ \vdots \\ \partial_t x_n \end{pmatrix}, \quad \partial_t y = \begin{pmatrix} \partial_t y_1 \\ \partial_t y_2 \\ \vdots \\ \partial_t y_n \end{pmatrix}, \quad \partial_t z = \begin{pmatrix} \partial_t z_1 \\ \partial_t z_2 \\ \vdots \\ \partial_t z_n \end{pmatrix}$$

Por otro lado, sea

$$\beta = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{pmatrix}$$

la matriz de tasas de transmisión debido al contacto entre individuos de las diferentes subpoblaciones, donde los coeficientes β_{ij} son las tasas de transmisión entre la subpoblación

i y la subpoblación j , en el caso de los β_{ii} , estos son las tasas de transmisión dentro de una

misma subpoblación y $\gamma = \begin{pmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & \gamma_n \end{pmatrix}$ la matriz de tasas de remoción para cada

grupo.

Sea X una matriz diagonal cuyos elementos X_{ii} sean justamente las entradas del vector x ,

es decir $X_{ii} = x_i$. Entonces $X = \begin{pmatrix} X_{11} & 0 & \cdots & 0 \\ 0 & X_{22} & \cdots & 0 \\ \vdots & \ddots & \cdots & 0 \\ 0 & \cdots & \cdots & X_{nn} \end{pmatrix} = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \ddots & \cdots & 0 \\ 0 & \cdots & \cdots & x_n \end{pmatrix}$

Con los elementos definidos es posible dar un sistema de ecuaciones algebraicas para la extensión metapoblacional del SIR de la forma (Figura 2.3)

$$\partial_t x = -X\beta y \quad (2.37)$$

$$\partial_t y = X\beta y - \gamma y \quad (2.38)$$

$$\partial_t z = \gamma y \quad (2.39)$$

Mostrar que el sistema de ecuaciones algebraico es congruente con las ecuaciones de la extensión metapoblacional es sencillo, del lado derecho de la ecuación (2.37) se tiene $-X\beta y =$

$$\begin{aligned} & - \begin{pmatrix} x_{11} & 0 & \cdots & 0 \\ 0 & x_{22} & \cdots & 0 \\ \vdots & \ddots & \cdots & 0 \\ 0 & \cdots & \cdots & x_{nn} \end{pmatrix} \begin{pmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & \cdots & \cdots \\ \vdots & \ddots & \cdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = - \begin{pmatrix} x_{11} & 0 & \cdots & 0 \\ 0 & x_{22} & \cdots & 0 \\ \vdots & \ddots & \cdots & 0 \\ 0 & \cdots & \cdots & x_{nn} \end{pmatrix} \begin{pmatrix} \sum_{j=1}^n \beta_{1j} y_j \\ \sum_{j=1}^n \beta_{2j} y_j \\ \vdots \\ \sum_{j=1}^n \beta_{nj} y_j \end{pmatrix} = \\ & - \begin{pmatrix} x_1 (\sum_{j=1}^n \beta_{1j} y_j) \\ x_2 (\sum_{j=1}^n \beta_{2j} y_j) \\ \vdots \\ x_n (\sum_{j=1}^n \beta_{nj} y_j) \end{pmatrix} = \begin{pmatrix} -(\sum_{j=1}^k \beta_{1j} x_1 y_j) \\ -(\sum_{j=1}^k \beta_{2j} x_2 y_j) \\ \vdots \\ -(\sum_{j=1}^k \beta_{nj} x_n y_j) \end{pmatrix} = \begin{pmatrix} \partial x_1 \\ \partial x_2 \\ \vdots \\ \partial x_n \end{pmatrix} = \partial_t x \end{aligned}$$

Para la ecuación (2.38), el desarrollo algebraico del lado derecho de la ecuación es

$$\begin{aligned} X\beta y - \gamma y &= \begin{pmatrix} \sum_{j=1}^n \beta_{1j} x_1 y_j \\ \sum_{j=1}^n \beta_{2j} x_2 y_j \\ \vdots \\ \sum_{j=1}^n \beta_{kj} x_n y_j \end{pmatrix} - \begin{pmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \ddots & \cdots & 0 \\ 0 & \cdots & \cdots & \gamma_k \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n \beta_{1j} x_1 y_j \\ \sum_{j=1}^n \beta_{2j} x_2 y_j \\ \vdots \\ \sum_{j=1}^n \beta_{nj} x_n y_j \end{pmatrix} - \begin{pmatrix} \gamma_1 y_1 \\ \gamma_2 y_2 \\ \vdots \\ \gamma_n y_n \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^n \beta_{1j} x_1 y_j - \gamma_1 y_1 \\ \sum_{j=1}^k \beta_{2j} x_2 y_j - \gamma_2 y_2 \\ \vdots \\ \sum_{j=1}^n \beta_{nj} x_n y_j - \gamma_n y_n \end{pmatrix} = \begin{pmatrix} \partial y_1 \\ \partial y_2 \\ \vdots \\ \partial y_n \end{pmatrix} = \partial_t y \end{aligned}$$

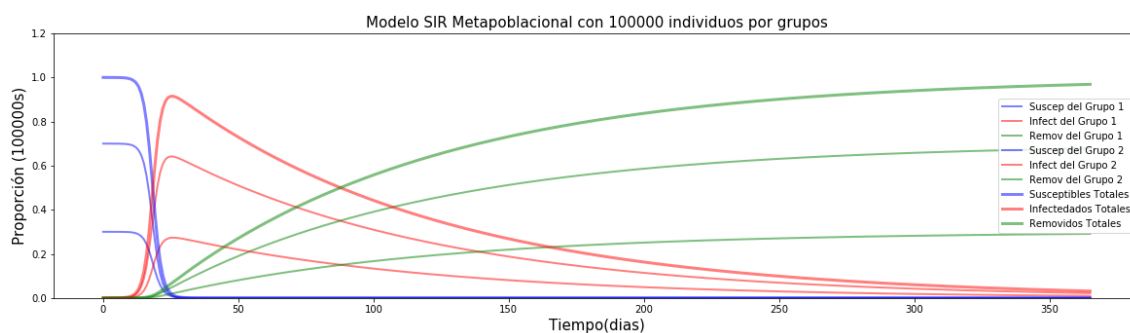


Figura 2.3: Modelo SIR Metapoblacional para una población de 100000 individuos.

Los grupos de riesgo en las poblaciones Los grupos de riesgo son un tipo de subpoblaciones formadas por individuos que son más propensos a adquirir una enfermedad infecciosa, por ejemplo, para el VIH los grupos se conforman, en general, por el grupo de los hombres que tienen sexo con hombres (HSH), trabajadores sexuales (TS), usuarios de drogas inyectables (UDI), en estos se ha observado una mayor prevalencia del virus que en el resto de la población, es decir la proporción de individuos de estos grupos que tienen VIH es mayor que en el resto de la población.

Otro ejemplo es el de la influenza tipo A, en general para todas sus cepas, en donde el grupo vulnerable se conforma de adultos mayores, niños, mujeres embarazadas, pacientes de cáncer, individuos con enfermedades cardiovasculares, con afecciones neurológicas e individuos inmunocomprometidos.

La importancia de los grupos de riesgo radica tanto en su tamaño como en las probabilidades de transmisión tanto dentro del grupo como entre grupos, estos parámetros deben ser capturados dentro del modelo metapoblacional mediante las tasas β_{ij} y γ_i por lo que simplemente intentar hacer estimaciones de estos valores puede ser un reto.

2.2.4. El problema del SIR y sus derivados

Los modelos revisados son de corte determinista y se expresan en sistemas de ecuaciones diferenciales, estos trabajan con base en supuestos sobre la población tales como la mezcla homogénea, el cual permite suponer que todos los individuos pueden tener contacto entre sí, así como el que todos interactúen de la misma forma, y el tamaño poblacional grande, que permite generar densidades poblacionales que se encuentren dentro del conjunto de los racionales y se pueda considerar una noción de continuidad en las funciones. Sin embargo dichos supuestos en general no se cumplen debido a que en comunidades pequeñas el tamaño poblacional no es lo suficientemente grande para considerar válidos los sistemas de ecuaciones diferenciales, en el mejor de los casos podríamos trabajar con un sistema de ecuaciones en diferencia. Luego, si el tamaño de la población es pequeño podría darse el

caso en que todos los individuos tienen contacto entre sí pero si el tamaño de la población es muy grande, lo suficiente para hacer válidas las ecuaciones del SIR, entonces no es realista pensar que hay un contacto de ese estilo, y por tanto no es válido pensar en mezcla homogénea.

El interés ahora será generar un modelo que sea flexible con el tamaño poblacional por lo que se considera un modelo discreto y además no deberá suponer de manera estricta la mezcla homogénea.

Capítulo 3

Modelos epidemiológicos discretos para poblaciones pequeñas

3.1. Visión probabilista de procesos epidemiológicos

Los procesos de infección y remoción, para la cadena de transmisión de una enfermedad infecciosa, pueden ser considerados de carácter aleatorio pues para un tiempo t y un intervalo de tiempo $\delta > 0$ no se conoce con exactitud la cantidad de susceptibles que serán infectados así como cuantos infectados serán removidos de la cadena de transmisión durante el intervalo de tiempo $(t + \delta)$.

Por otro lado, dentro de la cadena de transmisión puede considerarse una probabilidad de contagio, determinada por factores intrínsecos de la población, la forma en la que se dan los contactos infecciosos necesarios para que haya transmisión y aspectos biológicos de la población que faciliten o no el contagio; además de una probabilidad de remoción determinada por factores de la enfermedad.

La descripción de los procesos de infección y remoción en términos de los factores antes mencionados genera una serie de reglas con las cuales se puede plantear la dinámica de la cadena de transmisión en un sentido probabilista pues se puede considerar que las cantidades de infectados y removidos por unidad de tiempo son variables aleatorias, de naturaleza discreta, específicamente con un soporte finito, y cuyo muestreo en el tiempo se puede tomar discreto o continuo.

3.1.1. Cadenas de transmisión como procesos de muestreo

Identificar los elementos que participan en la cadena de transmisión de una enfermedad infecciosa permite caracterizar el proceso epidémico. Por ejemplo, en la influenza, en general para sus cepas, la transmisión se da entre un infectado y uno o varios susceptibles, los cuales tienen contacto con el infectado. Este contacto se produce en general mediante la dispersión de aerosoles generados por el infectado. En el VIH la transmisión se da princi-

palmente mediante contacto sexual aunque también puede darse por transfusión sanguínea, en el parto o compartiendo jeringas con alguien infectado. En el dengue, es necesaria la participación del mosquito transmisor del virus pero también que hayan personas infectadas a las cuales el mosquito pique para que este pueda propagar el virus a individuos susceptibles.

El tipo de contacto necesario para la transmisión, entonces, varía entre las enfermedades y determina en parte como se da la cadena de transmisión.

Por otro lado, así como el contacto infectado-susceptible es necesario para que haya un contagio no es necesariamente suficiente pues pueden haber factores del tipo fisiológicos a nivel individuo que lo protejan de contagiarse aún habiendo un contacto, caso de la inmunidad dada por la vacunación o el buen funcionamiento del sistema inmunológico de cada individuo. Podemos hablar entonces de una probabilidad de contagio para cada uno de los susceptibles en términos de los contactos y el que efectivamente se contagie el susceptible. Otro factor en la propagación es la noción de independencia entre los casos de nuevos infectados así como entre los casos de nuevos removidos. Es decir, el hecho de que, los eventos de infección para los susceptibles y de remoción para los infectados son independientes.

En este trabajo se considera que efectivamente hay independencia entre los casos de susceptibles que se infectaron así como entre las remociones, lo que genera otro factor a considerar dentro de la cadena de transmisión.

De esta manera, en principio, se pueden considerar como factores que caracterizan la cadena de transmisión de una enfermedad al contacto entre susceptibles e infectados (como condición necesaria pero no suficiente para que haya transmisión de una enfermedad infecciosa), la probabilidad efectiva de transmisión y la independencia entre los casos de susceptibles que efectivamente se infectan, así como en los casos de nuevos removidos.

Estas condiciones permiten plantear una serie de reglas para la incidencia de la enfermedad, es decir, el número de nuevos infectados por unidad de tiempo, así como para el número de casos que son removidos de la cadena por unidad de tiempo con base en el muestreo de estas. Esto se sintetiza en un conjunto de variables aleatorias generando el modelo PSIR.

3.1.2. Variables aleatorias y reglas de aleatoriedad

Fundamentos de teoría de probabilidad Se comienza revisando conceptos básicos necesarios en teoría de probabilidad que permitan comprender de mejor manera el planteamiento del modelo PSIR.

La teoría de la probabilidad es la rama teórica de las matemáticas encargada de realizar mediciones de factibilidad en la ocurrencia de eventos en un experimento o fenómeno aleatorio, no determinista, los cuales están sujetos a ciertas reglas en su comportamiento.

Decimos que un espacio de probabilidad es una terna (Ω, Σ, P) donde Ω es un espacio muestra, es decir, el conjunto formado por todas las posibles salidas de un experimento o fenómeno, $\Omega = \{\omega_i : \omega_i \text{ es una salida del experimento}\}$, Σ es un álgebra de conjuntos o una σ -álgebra de conjuntos en el caso en que Ω no sea finito asociada a Ω , es decir, es

una familia de subconjuntos del conjunto Ω que satisface Ω y $\phi \in \Sigma$, si $B \subset \Omega$ tal que $B \in \Sigma$ entonces $B^c \in \Sigma$ y si $(B_i)_{i=1}^n \in \Sigma$ es una colección a lo más numerable entonces $\cup_{i=1}^n B_i \in \Sigma$. A los elementos B de la σ -álgebra se conocen como eventos. Además, P es una medida de probabilidad, es decir, es un mapeo $P : \Omega \rightarrow [0, 1]$ que satisface que si $A \subset \Omega$ entonces $P(A) \geq 0$, si $(B_i)_{i=1}^n$ es una colección a lo más numerable de subconjuntos ajenos de Ω entonces $P(\cup_{i=1}^n B_i) = \sum_{i=1}^n P(B_i)$ y $P(\Omega) = 1$.

Decimos que una variable aleatoria (v.a.) es una función $X : \Omega \rightarrow A \subset R$ tal que $X^{-1}(x_i) = \{\omega_i \in \Omega : X(\omega_i) = x_i\} \in \Sigma$. Es decir, una variable aleatoria es una función que toma subconjuntos de Ω y les asigna un número real, sin embargo, la condición que debe cumplirse es que la imagen inversa del ese valor realmente caiga dentro de la σ -álgebra. Esto crea una relación directa entre los elementos $B \in \Sigma$ y los $\{X(\omega_i) = x_i : \omega_i \in B\}$ e induce una medida de probabilidad de la forma $p(B) = P(\{X(\omega_i) = x_i : \omega_i \in B\})$ en donde $p(B)$ es justamente el valor asignado por la medida de probabilidad original P por lo que es equivalente trabajar con los elementos asociados por la variable aleatoria y los elementos de la σ -álgebra a la hora de realizar el cálculo de probabilidades.

Así, las reglas de cálculo de probabilidades con variables aleatorias, $P(X = x)$ y $f_X(x)$, son no negativas para x en el rango de X , en el caso discreto $\sum_{x_i} P(X = x_i) = 1$ para x_i en el rango de X y en el caso continuo $\int_R f_X(x) dx = 1$. A $P(X = x)$ y $f_X(x)$ se les conoce como la función de masa o de densidad de la variable aleatoria X .

Se dice que dos eventos A y $B \in \Sigma$ son independientes si $P(A \cap B) = P(A)P(B)$, esto se refiere a que la probabilidad de que ocurran dos eventos de manera conjunta se puede calcular como el producto de cada una de las probabilidades y en esencia habla acerca de que el evento A no afecta la ocurrencia de B y viceversa. La versión de independencia para v.a.'s es $P(X = x, Y = y) = P(X = x)P(Y = y)$, esto puede ser extendido a una cantidad n de v.a.'s, en donde la colección $\{X_i\}_{i=1}^n$ es independiente si $P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i)$.

Debido a la naturaleza discreta en el modelo que nos interesa desarrollar nos centramos en las variables aleatorias discretas, en particular, las v.a.'s Bernoulli y Binomial.

La construcción de la Bernoulli comienza considerando un experimento aleatorio con únicamente dos salidas, asociadas a éxito y fracaso respectivamente. Para fines prácticos se considera el lanzar una moneda al aire, aislada de factores que alteren su resultado, y ver cual es el resultado. En este caso nuestro conjunto $\Omega = \{\text{Sol}, \text{Águila}\}$; puesto que solo hay dos posibles resultados, una σ -álgebra que nos sirve es 2^Ω , es decir el conjunto potencia de Ω . Ahora, se considera que existe un valor $p \in [0, 1]$ tal que $P(\{\text{Sol}\}) = p$ y por tanto $P(\{\text{Águila}\}) = 1 - p$. Nótese que sol y águila están intencionalmente entre corchetes debido a que son elementos de la σ -álgebra y por tanto son los posibles eventos que pueden ocurrir. Además dichos eventos son excluyentes entre sí, no pueden ocurrir ambos al mismo tiempo. Para este experimento, la variable aleatoria que modela el fenómeno es

$$X(\omega) = \begin{cases} 1 & \text{si } \omega = \{\text{Águila}\}, \\ 0 & \text{si } \omega = \{\text{Sol}\}. \end{cases}$$

Esto implica que $P(\{\text{Águila}\}) = P(X = 1)$ y $P(\{\text{Sol}\}) = P(X = 0)$. A este tipo de variables aleatorias se les conoce como Bernoulli(p) donde p es la probabilidad de éxito o fracaso de nuestro experimento, dependiendo de como se defina. La función de masa de probabilidad queda definida de la siguiente manera

$$P(X = x) = \begin{cases} p & \text{si } x = 1 \\ 1 - p & \text{si } x = 0 \\ 0 & \text{e.o.c.} \end{cases}$$

Las v.a.'s Binomiales surgen al preguntar por la cantidad de éxitos en una sucesión de n ensayos Bernoulli del mismo parámetro e independientes entre sí, por lo que se tiene una cantidad n de variables aleatorias independientes e idénticamente distribuidas.

La descripción del fenómeno visto como un conjunto de v.a. se le conoce como muestra aleatoria \bar{X} y puede visualizarse como un vector de n entradas en donde cada entrada es una variable, es decir, $\bar{X} = (X_1 = x_1, \dots, X_n = x_n)$. La independencia entre los ensayos Bernoulli genera que la función de masa de probabilidad para la realización completa del experimento sea entonces $P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1)P(X_2 = x_2)\dots P(X_n = x_n)$. El número de éxitos puede calcularse al contar la cantidad de entradas del vector \bar{X} con valor 1 mediante una v.a. definida como $Y = \text{número de éxitos en la muestra}$. Notemos que Y puede ser expresada en términos de las X_i como $Y = \sum_{i=1}^n X_i$. Si $Y = k$ para $0 \leq k \leq n$ entonces habrán k éxitos en la muestra de todas las maneras posibles en las que puedan ordenarse, por lo que la función de masa de probabilidad para Y queda definida como

$$P(Y = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{si } k \in \{0, \dots, n\}, \\ 0 & \text{e.o.c.} \end{cases}$$

La v.a. que cuenta el número de éxitos (o fracasos) en una cantidad n de realizaciones Bernoulli de parámetro p recibe el nombre de Binomial(n, p). (Figura 3.1)

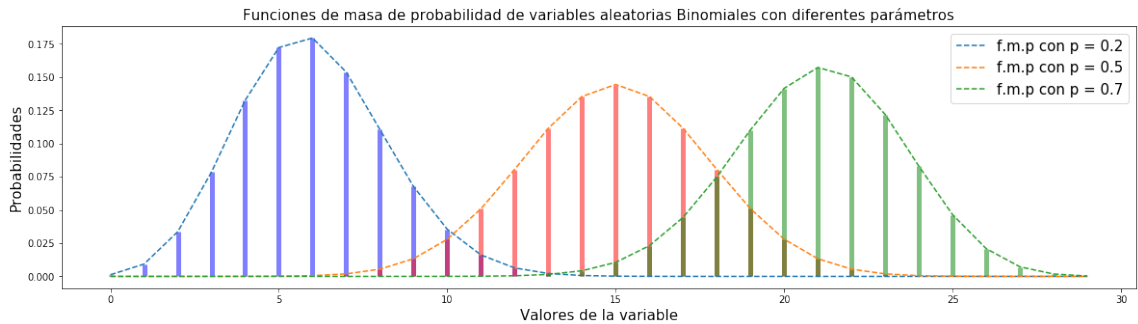


Figura 3.1: Función de masa de probabilidad para variables aleatorias binomiales.

El cálculo de probabilidades utilizando variables aleatorias se complementa con la esperanza de la variable como medida de centralidad y su varianza como medida de dispersión. La esperanza es un promedio ponderado de los valores que toma la v.a. y nos dice en entorno a que valor real se centra nuestra distribución, está se define como $E(X) = \sum_{x_i} x_i P(X = x_i)$. La varianza es una medida de que tanta distancia se espera que haya entre los valores que toma la variable y su media y se define como $Var(X) = E((X - \mu)^2) = E(X^2) - \mu^2$ donde $\mu = E(X)$.

Para $X \sim Bernoulli(p)$ su valor esperando es $E(X) = p$ y $Var(X) = p(1 - p)$ mientras que para $X \sim Bin(n, p)$ su esperanza y varianza son $E(X) = np$ y $Var(X) = np(1 - p)$ respectivamente.

Por último, un concepto muy importante en probabilidad es el de probabilidad condicional. Dados dos eventos A y B en Σ tal que $P(B) > 0$, se define la probabilidad condicional de A dado B como $P(A|B) := \frac{P(A \cap B)}{P(B)}$. La probabilidad condicional es una nueva medida de probabilidad, la cual es el resultado de reestructurar la σ -álgebra del espacio de probabilidad original al saber que el evento B ya ha ocurrido, lo que se interpreta como información y restringe es espacio de probabilidad original a uno en donde solo ocurran aquellos eventos relacionados con B .

Cadenas de Markov en tiempo discreto. Los procesos estocásticos son representaciones matemáticas de la evolución en el tiempo de un fenómeno aleatorio. Así, el fenómeno puede ser expresado como una sucesión de variables aleatorias $\{X_t\}_{t \in T}$, en donde cada X_t tiene asociada una distribución de probabilidad y los valores x que toman las variables se conocen como los estados del proceso.

Las cadenas de Markov en tiempo discreto son un tipo de proceso estocástico en el que los valores n en T junto con los estados i del proceso son discretos y en el que se satisface la propiedad de Markov $P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)$, esta propiedad se interpreta como que el estado de la cadena en el siguiente paso en el tiempo solo depende del estado de la cadena en el tiempo actual. Además, para $X_n = i$ existe una probabilidad fija p_{ij} de que $X_{n+1} = j$, es decir, la probabilidad condicional $P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = p_{ij}$, dicha probabilidad no depende de n . Los valores p_{ij} se conocen como probabilidades de transición del estado i al estado j , y satisfacen que $p_{ij} \geq 0$ y $\sum_j p_{ij} = 1$. Dichas probabilidades de transición pueden ordenarse en una matriz de tamaño $n \times n$ en donde n es la cantidad de estados que pueden tomar las variables aleatorias. Si en particular la cantidad de estados disponibles es finita, dígame los estados $\{0, 1, \dots, N\}$, la matriz de probabilidades de transición tiene la forma

$$\begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0N} \\ p_{10} & p_{11} & \cdots & p_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N0} & p_{N1} & \cdots & p_{NN} \end{pmatrix}$$

Esta matriz caracteriza la transición del fenómeno en el siguiente paso del tiempo a cualquiera de los estados que pueden tomar las v.a.'s.

3.2. El modelo PSIR

El PSIR visualiza el proceso de infección y remoción como un proceso de muestreo aleatorio en tiempo discreto, en particular, el conteo en la cantidad de susceptibles que pasan a ser infectados desde un tiempo t hasta un tiempo $(t+\delta)$ con $\delta > 0$ puede modelarse mediante una v.a. Binomial en donde el éxito se interpreta como una infección exitosa, así mismo, la remoción se considera también una v.a. binomial en donde un éxito es una remoción.

Sea una población de tamaño constante N , y $S(t)$, $I(t)$ y $R(t)$ la cantidad de susceptibles, infectados y removidos al tiempo t tales que $S(t) + I(t) + R(t) = N$.

Esto implica que $\frac{S(t)}{N}$, $\frac{I(t)}{N}$ y $\frac{R(t)}{N}$ son las proporciones de susceptibles, infectados y removidos en la población al tiempo t .

Para una enfermedad infecciosa, se considera que la transmisión se da debido a un contacto susceptible-infectado para luego el susceptible efectivamente enfermarse.

Así, la probabilidad de infección para un individuo susceptible fijo está dada como $P(\{\text{Infección para un individuo susceptible fijo}\}) = P(\{\text{Contacto infeccioso}\}) \times P(\{\text{Transmisión efectiva}\}|\{\text{Contacto infeccioso}\})$

Se puede pensar $P(\{\text{Contacto infeccioso}\})$ como la proporción de individuos infectados dentro de la población, es decir $\frac{I(t)}{N}$, mientras que $P(\{\text{Transmisión efectiva}\}|\{\text{Contacto infeccioso}\})$ se puede tomar como un parámetro $\beta \in (0, 1)$. Por lo tanto $P(\text{Infección para un individuo susceptible}) = \beta \frac{I(t)}{N}$.

Para un tiempo t la probabilidad de infección para un individuos está determinada como $P(\text{Infección entre } t \text{ y } (t + \delta)) = \beta \frac{I(t)}{N}$.

Para cada individuo susceptible i , se define la v.a.

$$X_i(t) = \begin{cases} 1 & \text{Si el } i\text{-ésimo susceptible es infectado entre el tiempo } t \text{ y } (t + \delta) \\ 0 & \text{Si no es infectado.} \end{cases}$$

Donde $P(X_i(t) = 1) = \beta \frac{I(t)}{N}$ y $P(X_i(t) = 0) = 1 - \beta \frac{I(t)}{N}$.

La cantidad total de nuevos infectados en el tiempo t es una v.a. denotada por $X(t)$.

Por otro lado, sea $0 < \epsilon < 1$ la proporción de individuos susceptibles en contacto con los infectados por unidad de tiempo, entonces, $\epsilon S(t)$ es la cantidad de susceptibles en contacto con los infectados por lo que para una cantidad $\lfloor \epsilon S(t) \rfloor$ de susceptibles disponibles en la cadena de transmisión, $X(t) = \sum_{i=1}^{\lfloor \epsilon S(t) \rfloor} X_i$. Si suponemos que cada persona se infecta de manera independiente de las demás, lo que implica que las variables son independientes entre sí, debido a que $X(t)$ es una suma de variables Bernoulli entonces durante un intervalo

de tiempo $\delta > 0$ se tiene que $X \sim \text{Bin}(\lfloor \epsilon S(t) \rfloor, \beta \frac{I(t)}{N} \delta)$ y su función de masa de probabilidad es

$$f_{X(t)}(x) = \begin{cases} \binom{\lfloor \epsilon S(t) \rfloor}{x} (\beta \frac{I(t)}{N} \delta)^x (1 - \beta \frac{I(t)}{N} \delta)^{\lfloor \epsilon S(t) \rfloor - x}, & x \in \{0, \dots, \lfloor \epsilon S(t) \rfloor\} \\ 0 & e.o.c \end{cases}$$

El término $\lfloor \epsilon S(t) \rfloor$ se refiere al piso de $\epsilon S(t)$ y es importante aclarar que la finalidad de plantear un número específico de susceptibles que participen en la cadena de transmisión es expresar que no puede haber contacto entre cualesquiera dos individuos en la población y por tanto considerar que no hay mezcla homogénea.

El número esperado de nuevos infectados por unidad de tiempo δ para la v.a. X es $E(X(t)) = \lfloor \epsilon S(t) \rfloor \beta \frac{I(t)}{N} \delta$.

Para los nuevos removidos, se considera que un individuo infectado i pasa a ser removido luego de un tiempo T_i . Suponemos que se tiene una cantidad n de infectados con tiempos de remoción de la cadena de transmisión T_i para $i \in \{1, \dots, n\}$. El tiempo \bar{T} definido como $\bar{T} = \frac{1}{n} \sum_{i=1}^n T_i$ aproxima al tiempo promedio de remoción por lo que para n suficientemente grande el inverso de \bar{T} , $\gamma = \frac{1}{\bar{T}}$ aproxima a la tasa promedio de remoción.

Dado un intervalo de tiempo δ , $\gamma \delta$ es $P(\text{remoción del sistema en el intervalo de tiempo } (t, t + \delta))$.

Si suponemos remociones independientes entre los infectados con probabilidad $p = \gamma \delta$, el número de nuevos removidos por unidad de tiempo δ es una variable aleatoria Y , para una cantidad $I(t)$ de infectados disponibles en la población $Y(t) \sim \text{Bin}(I(t), \gamma \delta)$. El modelo PSIR queda determinado de la siguiente manera, para un tiempo t y un intervalo de tiempo $\delta > 0$, los muestreos binomiales sobre los susceptibles y los infectados determinan la cantidad de susceptibles, infectados y removidos en el tiempo $(t + \delta)$ mediante las

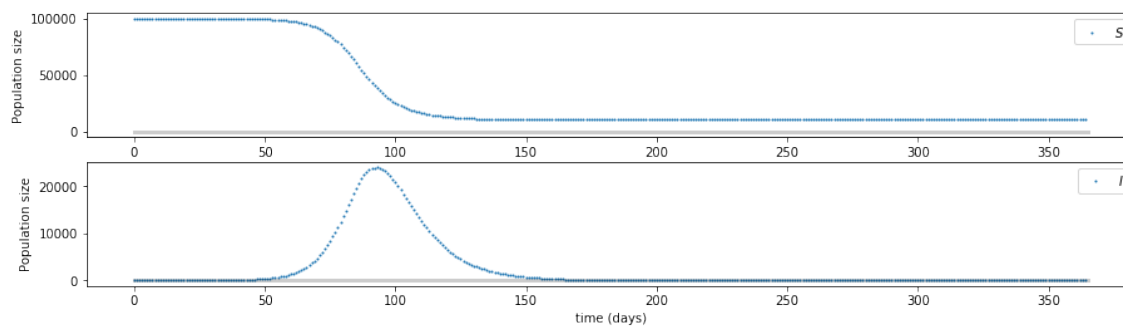


Figura 3.2: Curva de susceptibles e infectados generados por el modelo PSIR para una población de tamaño $N = 100000$, $\beta = 0.5$ y $\gamma = 0.1$

ecuaciones

$$S(t + \delta) = S(t) - X(t) \quad (3.1)$$

$$I(t + \delta) = I(t) + X(t) - Y(t) \quad (3.2)$$

$$R(t + \delta) = R(t) + Y(t) \quad (3.3)$$

Donde $X(t) \sim \text{Bin}(\lfloor \epsilon S(t) \rfloor, \beta \frac{I(t)}{N} \delta)$ y $Y(t) \sim \text{Bin}(I(t), \gamma \delta)$.

El muestreo generado por $X(t)$ y $Y(t)$ implica que la cantidad de susceptibles baja conforme pasa el tiempo, mientras que la cantidad de infectados se incrementa pues están llegando nuevos infectados para luego comenzar a descender (Figura 3.2).

El PSIR es un modelo estocástico que tanto el tiempo como el conjunto de estados de las variables son discretas, es markoviano pues, por construcción, el estado de la variable en un tiempo $(t + \delta)$ solo depende del paso anterior, es un proceso estocástico en conjunto pues se tiene una sucesión ordenada de variables aleatorias sin embargo el conjunto de variables aleatorias no es idénticamente distribuido. Las v.a. no son idénticamente distribuidas pues los parámetros tanto de $X(t)$ como de $Y(t)$ varían con respecto al tiempo, específicamente, $X(t) \sim \text{Bin}(\lfloor \epsilon S(t) \rfloor, \beta \frac{I(t)}{N} \delta)$ y conforme pasa el tiempo el número de susceptibles baja mientras que el número de infectados aumenta o disminuye dependiendo de cuántos serán removidos en el siguiente paso por lo que los valores en los parámetros de las distribuciones cambia, mismo caso es el de $Y(t) \sim \text{Bin}(I(t), \gamma \delta)$. Un análisis minucioso de como cambian estas variables aleatorias se lleva a cabo en la sección de resultados.

Para el PSIR se puede calcular un análogo al número básico de reproducción, R_0 . En el modelo determinista nos fijamos cuando $\partial_t y > 0$ lo que termina estableciendo que $R_0 = \beta/\gamma$. La condición en el PSIR es que la cantidad de infectados menos la de removidos sea mayor que cero, esto es $X(t) - Y(t) > 0$ implica que $X(t) > Y(t)$, esto es, nos interesa que la cantidad de individuos que se infectan sea mayor que la cantidad que se remueven. Esta condición no se cumple de manera determinista pero podemos aplicar la esperanza de ambos lados. Tenemos $X(t) > Y(t)$ implica que $E(X(t)) > E(Y(t))$ a su vez quiere decir que $\lfloor \epsilon S(t) \rfloor \beta \frac{I(t)}{N} \delta > I(t) \gamma \delta$ de donde $\epsilon S(t) > N \frac{\gamma}{\beta}$, el cual es análogo al SIR determinista. Un factor a tomar en cuenta dentro del PSIR es la probabilidad de infección dado el contacto β de la variable $X(t)$. Una forma de caracterizar dicho parámetro es considerar como la enfermedad actúa en diferentes grupos de la población. Por ejemplo, al condicionar sobre grupos de edad (GE) se obtiene mediante la ley total de probabilidad que

$$P(\text{Infección}|\text{Contacto}) = \sum_{\text{GE}} P(\text{Infección}|\text{Contacto}|\text{Pertenece al GE})P(\text{Pertenece al GE})$$

Donde $P(\text{Pertenece al GE})$ puede pensarse como la proporción de individuos en la población que pertenece a cada grupo, mientras que $P(\text{Infección}|\text{Contacto}|\text{Pertenece al GE})$ puede tomarse como las proporciones de individuos pertenecientes a cada grupo de edad que tienen la enfermedad, en donde un buen registro de los casos de infección es necesario para poder tener una buena aproximación de dicha probabilidad.

Modelo PSI y PSIR Metapoblacional

El modelo PSI La versión PSI del modelo, bajo el esquema que se ha planteado resulta ser idéntica construcción al PSIR, de hecho es más sencillo pues solo es necesaria $X(t)$ debido a que no hay remoción en el sistema.

En este caso, $S(t) + I(t) = N$ por lo que las ecuaciones del PSI son

$$S(t + \delta) = S(t) - X(t) \quad (3.4)$$

$$I(t + \delta) = I(t) + X(t). \quad (3.5)$$

con $X(t)$ como el número de nuevos infectados y cuya distribución es binomial con parámetros $S(t)$ y $\beta \frac{I(t)}{N} \delta$.

La dinámica del PSI es muy fácil de entender debido que solo se tiene la v.a. $X(t)$ llevando los elementos muestreados del estado susceptible a infectado y dado que el tamaño poblacional permanece constante se tiene que la cantidad de infectados será el complemento de la cantidad de infectados dentro de la población, es decir $S(t) + I(t) = N$, esto también implica que la cantidad de infectados puede verse como la ecuación $I(t) = N - S(t)$ (Figura 3.3).

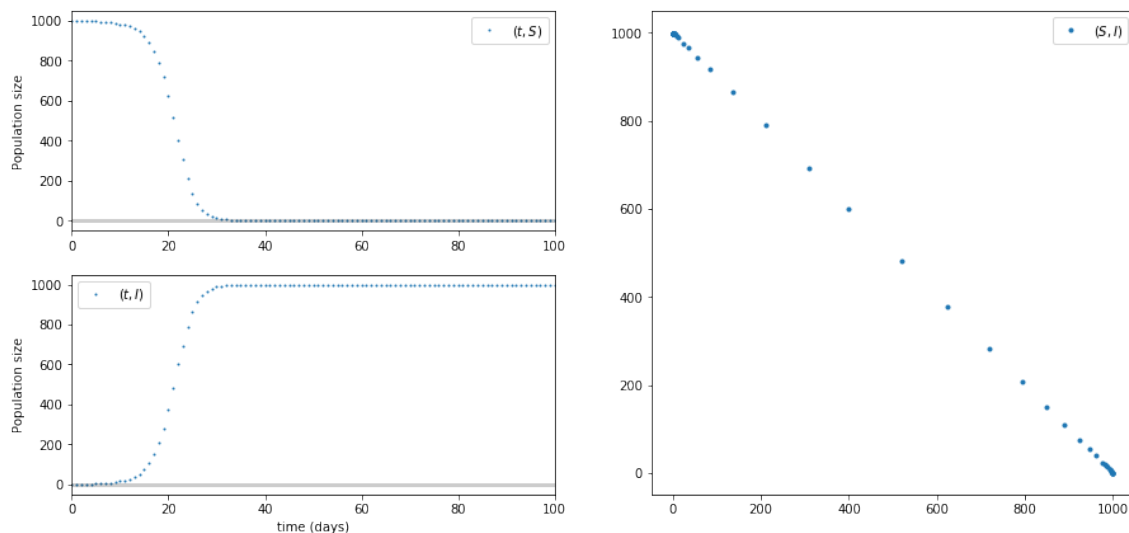


Figura 3.3: Del lado izquierdo, las curvas de susceptibles e infectados del PSI, del lado derecho la gráfica de los puntos (S, I) .

Un hecho a destacar es que en el calculo de $E(X(t))$ se tiene $E(X(t)) = \lfloor \epsilon S(t) \rfloor \beta \frac{I(t)}{N} \delta = \lfloor \epsilon S(t) \rfloor \beta \frac{N - S(t)}{N} \delta = \lfloor \epsilon S(t) \rfloor \beta (1 - \frac{S(t)}{N}) \delta$ o equivalentemente $E(X(t)) = \epsilon N \beta x(t) (1 - x(t)) \delta$

con $x(t) = \frac{S(t)}{N}$ el cual es análogo a la ecuación en el cambio en los infectados en el SI determinista.

Extensión a subgrupos poblacionales. El pSIR puede extenderse al caso de los subgrupos poblacionales. Se empieza considerando el caso en el que la población está dividida en dos subgrupos para luego generalizarlo a una cantidad n .

Sea una población de tamaño N constante, conformada de dos grupos, G_1 y G_2 , cada uno con cantidades S_1, I_1, R_1 y S_2, I_2, R_2 de susceptibles, infectados y removidos respectivamente. En la población en general se tiene que $S(t) = S_1(t) + S_2(t)$, $I(t) = I_1(t) + I_2(t)$ y $R(t) = R_1(t) + R_2(t)$. Además, debido a que la población permanece constante, $S(t) + I(t) + R(t) = N$ por lo que $S_1(t) + S_2(t) + I_1(t) + I_2(t) + R_1(t) + R_2(t) = N$. En la población, se puede considerar que los contactos infecciosos se dan entre individuos del mismo grupo, es decir entre individuos de G_1 con G_1 y G_2 con G_2 así como entre individuos de grupos distintos, es decir entre individuos de G_1 con G_2 .

La dinámica de transmisión puede describirse extendiendo las nociones del PSIR de la siguiente manera.

Sea $0 < \epsilon_1 < 1$ donde ϵ_1 es la proporción de individuos de S_1 que tienen contacto con individuos de I_1 , es decir la proporción de susceptibles del grupo 1 que tienen contacto con infectados de su mismo grupo, entonces $\lfloor \epsilon_1 S_1(t) \rfloor$ es la cantidad de susceptibles del grupo 1 que tienen contacto con infectados del grupo 1 al tiempo t , en donde $\lfloor a \rfloor$ denota el piso de $a \in R$. El valor $\lfloor (1 - \epsilon_1) S_1(t) \rfloor$ denota entonces la cantidad de susceptibles del grupo 1 que tienen contacto con infectados del grupo 2.

El número de nuevos infectados, al tiempo t , generados por el contacto entre susceptibles e infectados de G_1 se puede pensar como un muestreo binomial, la cual genera una variable aleatoria $X_{1,1}(t) =$ número de nuevos infectados debido al contacto entre individuos de S_1 con I_1 , donde $X_{1,1}(t) \sim Bin(\lfloor \epsilon_1 S_1(t) \rfloor, \beta \frac{I_1(t)}{n} \delta)$.

Por otro lado, el número de nuevos infectados dados por el contacto entre susceptibles de G_1 con infectados de G_2 es una variable aleatoria $X_{1,2} =$ Nuevos infectados debido al contacto entre individuos de S_1 con I_2 Con $X_{1,2}(t) \sim Bin(\lfloor (1 - \epsilon_1) S_1(t) \rfloor, \beta \frac{I_2(t)}{n} \delta)$.

De manera análoga, sea $0 < \epsilon_2 < 1$ la proporción de susceptibles del grupo 2 que tienen contacto con infectados del grupo 2. El número de nuevos infectados debido al contacto entre susceptibles e infectados del grupo 2 es una variable aleatoria que corresponde a un muestreo binomial, $X_{2,2} =$ Nuevos infectados debido al contacto entre individuos de S_2 con I_2 donde $X_{2,2}(t) \sim Bin(\lfloor \epsilon_2 S_2(t) \rfloor, \beta \frac{I_2(t)}{N} \delta)$. El número de nuevos infectados debido al contacto entre susceptibles de G_2 con infectados de G_1 es una variable aleatoria binomial $X_{2,1} =$ Nuevos infectados debido al contacto entre individuos de S_2 con I_1 por lo que $X_{2,1}(t) \sim Bin(\lfloor (1 - \epsilon_2) S_2(t) \rfloor, \beta \frac{I_1(t)}{N} \delta)$

Una vez que los individuos se infectan, tanto para G_1 como para G_2 , el número de nuevos removidos en cada grupo son muestreos binomiales de la forma $Y_1(t) \sim Bin(I_1(t), \gamma_1 \delta)$ y $Y_2(t) \sim Bin(I_2(t), \gamma_2 \delta)$ (puede considerarse que $\gamma_1 = \gamma_2$ si las tasas de remoción son

iguales en ambos grupos). Notese también que $\lfloor \epsilon_1 S_1(t) \rfloor + \lfloor (1 - \epsilon_1) S_1(t) \rfloor \leq \epsilon_1 S_1(t) + (1 - \epsilon_1) S_1(t) \leq S_1(t)$ y $\lfloor \epsilon_2 S_2(t) \rfloor + \lfloor (1 - \epsilon_2) S_2(t) \rfloor \leq \epsilon_2 S_2(t) + (1 - \epsilon_2) S_2(t) \leq S_2(t)$ por lo que $X_{1,1}, X_{1,2}, X_{2,1}, X_{2,2}$ están bien definidas.

Las ecuaciones para la extensión de PSIR con dos subgrupos poblacionales son

$$S_1(t + \delta) = S_1(t) - (X_{1,1}(t) + X_{1,2}(t)) \quad (3.6)$$

$$S_2(t + \delta) = S_2(t) - (X_{2,1}(t) + X_{2,2}(t)) \quad (3.7)$$

$$I_1(t + \delta) = I_1(t) + (X_{1,1}(t) + X_{1,2}(t)) - Y_1(t) \quad (3.8)$$

$$I_2(t + \delta) = I_2(t) + (X_{2,1}(t) + X_{2,2}(t)) - Y_2(t) \quad (3.9)$$

$$R_1(t + \delta) = R_1(t) + Y_1(t) \quad (3.10)$$

$$R_2(t + \delta) = R_2(t) + Y_2(t). \quad (3.11)$$

El comportamiento de la extensión metapoblacional del PSIR dependerá del tamaño de las poblaciones susceptibles S_1 y S_2 y de las proporciones en los que estos estén en contacto con sus respectivos grupos de infectados I_1 e I_2 , proporciones bajas de individuos en contacto con sus respectivos grupos implica un contacto cruzado grande por lo que habrá un comportamiento caótico en la curva de infectados (Figura 3.4), mientras que una proporción grande de susceptibles en contacto con sus respectivos infectados generará un comportamiento muy parecido al PSIR sin considerar grupos metapoblacionales pues es como considerar dos poblaciones separadas e independientes que al final se están sumando, (Figura 3.5)

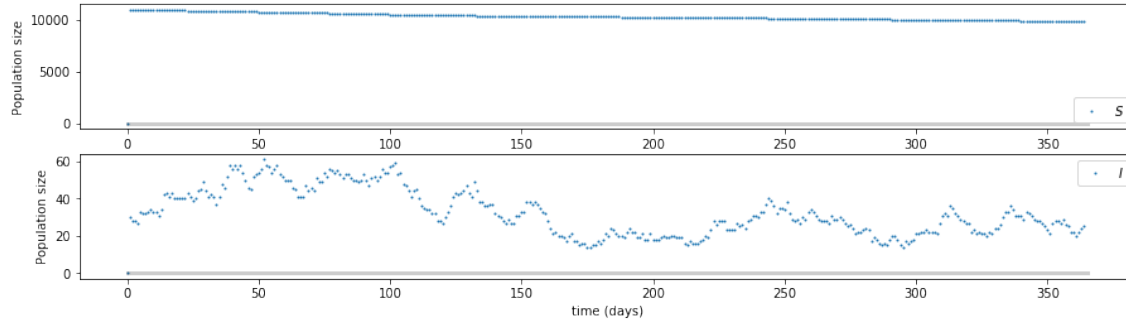


Figura 3.4: Comportamiento del PSIR Metapoblacional cuando 2/10 de la población susceptible 1 está en contacto con la población infectada 1 y 3/10 de la población susceptible 2 está en contacto con la población infectada 2.

La suma de variables aleatorias binomiales en este caso no es binomial debido a la diferencia entre los parámetros de éxito de cada distribución, sin embargo, suponiendo que las variables son independientes, debido a que para ambos grupos se tienen bien definidos a los individuos que tienen contacto entre sí y que no está el caso en el que un individuo

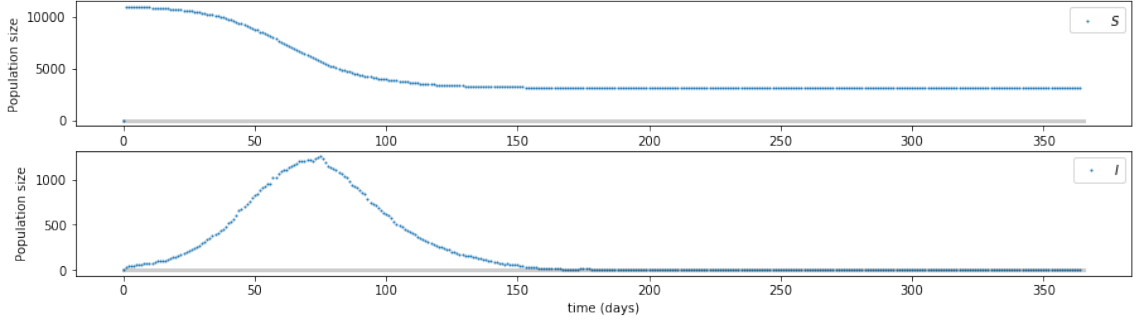


Figura 3.5: Comportamiento del PSIR Metapoblacional cuando $7/10$ de la población susceptible 1 está en contacto con la población infectada 1 y $7/10$ de la población susceptible 2 está en contacto con la población infectada 2.

susceptible tenga un doble contacto infeccioso, se puede trabajar con la ley de probabilidad de la suma. Nombrando $p_{1,1}, p_{1,2}, p_{2,1}, p_{2,2}$ a los respectivos parámetros de probabilidad éxito de las y $[\epsilon_1 S_1] = r_1$, $[(1 - \epsilon_1) S_1] = r_2$, $[\epsilon S_2] = r_3$, $[(1 - \epsilon_2) S_2] = r_4$, distribuciones de $X_{1,1}, X_{1,2}, X_{2,1}, X_{2,2}$. Como resultado, los nuevos infectados del grupo 1 son variables aleatorias binomiales con distribuciones $X_{1,1} \sim \text{Bin}(r_1, p_{1,1})$ y $X_{1,2} \sim \text{Bin}(r_2, p_{1,2})$. que en su suma se tiene

$$\begin{aligned} P(X_{1,1} + X_{1,2} = n) &= \sum_{x=0}^n P(X_{1,1} = x, X_{1,2} = n - x) = \sum_{x=0}^n P(X_{1,1} = x)P(X_{1,2} = n - x) \\ &= \sum_{x=0}^n \binom{r_1}{x} (p_{1,1})^x (1 - p_{1,1})^{r_1 - x} \binom{r_2}{n - x} (p_{1,2})^{n - x} (1 - p_{1,2})^{r_2 - (n - x)}. \end{aligned}$$

De manera análoga, $X_{2,1} \sim \text{Bin}(r_3, p_{2,1})$ y $X_{2,2} \sim \text{Bin}(r_4, p_{2,2})$ y

$$\begin{aligned} P(X_{2,1} + X_{2,2} = n) &= \sum_{x=0}^n P(X_{2,1} = x, X_{2,2} = n - x) = \sum_{x=0}^n P(X_{2,1} = x)P(X_{2,2} = n - x) \\ &= \sum_{x=0}^n \binom{r_3}{x} (p_{2,1})^x (1 - p_{2,1})^{r_3 - x} \binom{r_4}{n - x} (p_{2,2})^{n - x} (1 - p_{2,2})^{r_4 - (n - x)}. \end{aligned}$$

La esperanza y varianza para ambas sumas están dadas como

$$E(X_{1,1} + X_{1,2}) = r_1 p_{1,1} + r_2 p_{1,2} \quad (3.12)$$

$$E(X_{2,1} + X_{2,2}) = r_3 p_{2,1} + r_4 p_{2,2} \quad (3.13)$$

$$\text{Var}(X_{1,1} + X_{1,2}) = r_1 p_{1,1} (1 - p_{1,1}) + r_2 p_{1,2} (1 - p_{1,2}) \quad (3.14)$$

$$\text{Var}(X_{2,1} + X_{2,2}) = r_3 p_{2,1} (1 - p_{2,1}) + r_4 p_{2,2} (1 - p_{2,2}). \quad (3.15)$$

A continuación, generalizamos a una cantidad n de subgrupos dentro de la población. Sean S_i, I_i, R_i la cantidad de susceptibles, infectados y removidos del i -ésimo grupo con $i \in \{1, \dots, n\}$. Donde los susceptibles totales son $\sum_{i=1}^n S_i = S$, los infectados totales son $\sum_{i=1}^n I_i = I$ y los removidos totales son $\sum_{i=1}^n R_i = R$, luego sea ϵ_{ij} la proporción de susceptibles del grupo G_i en contacto con infectados del grupo G_j , donde $\epsilon_{ij} \geq 0, \forall i, j \in \{1, \dots, n\}$ y $\sum_{j=1}^n \epsilon_{ij} = 1$.

Así, $\epsilon_{ij} S_i(t)$ es el número de susceptibles del grupo G_i en contacto con infectados del grupo G_j , por lo que la variable aleatoria $X_{i,j} =$ Nuevos infectados debido al contacto entre individuos de S_i con I_j es una v.a con distribución binomial $X_{i,j}(t) \sim Bin(\lfloor \epsilon_{ij} S_i(t) \rfloor, \beta \frac{I_j(t)}{N} \delta)$. Luego, $Y_i =$ Nuevos removidos de G_i es una v.a. binomial de la forma $Y_i(t) \sim Bin(I_i(t), \gamma_i \delta)$. Las ecuaciones de la cadena de transmisión quedan definidas como

$$S_i(t + \delta) = S_i(t) - \sum_{j=1}^n X_{i,j}(t) \quad (3.16)$$

$$I_i(t + \delta) = I_i(t) + \left(\sum_{j=1}^n X_{i,j}(t) \right) - Y_i(t) \quad (3.17)$$

$$R_i(t + \delta) = R_i(t) + Y_i(t). \quad (3.18)$$

3.2.1. Otros modelos estocásticos para cadenas de transmisión

En la literatura se pueden encontrar modelos estocásticos que describen la dinámica de transmisión de una enfermedad infecciosa, (Brauer et al., 2001), Brauer et al. (2008), (Allen et al., 2008). Debido a que el pSIR tiene un comportamiento markoviano, pues la transición entre los valores de las variables aleatorias solo dependen del paso anterior, nos enfocaremos en revisar algunos modelos estocásticos basados en cadena de Markov a tiempo discreto (CMTD) y en particular revisaremos los modelos de Greenwood y Reed-Frost. Para el modelo SIR-CMTD, sean $S(t), I(t)$ y $R(t)$ v.a. discretas para el número de susceptibles, infectados y removidos al tiempo t respectivamente. El proceso de infección es un proceso bivariado $\{(S(t), I(t))\}$ en donde $R(t) = N - S(t) - I(t)$ cuya función de masa de probabilidad es una función conjunta $p_{(s,i)}(t) = P(S(t) = s, I(t) = i)$. Además, el proceso tiene la propiedad de Markov y es homogéneo en el tiempo, por lo que las probabilidades de transición de la cadena se pueden determinar de la siguiente manera

Sea $\Delta t > 0$ un intervalo de tiempo tal que $\Delta t > 0$ pueda ser escogido lo suficientemente pequeño, entonces

$$p_{(s,i)(s+k,i+j)}(\Delta t) = P((\Delta S, \Delta I) = (k, j) | (S(t) = s, I(t) = i)) \quad (3.19)$$

Donde $\Delta S = -S(t)$, $\Delta I = -I(t)$ y cuyas probabilidades de transición son

$$p_{(s,i)(s+k,i+j)}(\Delta t) = \begin{cases} \frac{\beta si}{N} \Delta t & si & (k, j) = (-1, 1) \\ \gamma i \Delta t & si & (k, j) = (0, 1) \\ i \Delta t & si & (k, j) = (-1, 1) \\ (N - s - i) \Delta t & si & (k, j) = (0, 0) \\ 1 - \frac{\beta si}{N} \Delta t - (\gamma i + (N - s)) \Delta t & si & (k, j) = (0, 0) \\ 0 & e.o.c & \end{cases} \quad (3.20)$$

donde β es la tasa de contacto y γ es la tasa de recuperación.

Aplicando la propiedad de Markov se puede escribir $p_{(s,i)}(t + \Delta t)$ en términos de las probabilidades de transición como

$$\begin{aligned} p_{(s,i)}(t + \Delta t) &= p_{(s+1,i-1)}(t) \frac{\beta}{N} (i-1)(s+1) \Delta t + p_{(s,i+1)}(t) \gamma (i+1) \Delta t \\ &+ p_{(s-1,i+1)}(t) (i+1) \Delta t + p_{(s-1,i)}(t) (N - s + 1 - i) \Delta t \\ &+ p_{(s,i)}(t) \left(1 - \frac{\beta}{N} si + \gamma i + N - s \right) \Delta t \end{aligned}$$

Además, la esperanza de $S(t)$ e $I(t)$ son respectivamente $E(S(t)) = \sum_{s=0}^N s p_{(s,i)}(t)$ y $E(I(t)) = \sum_{i=0}^N i p_{(s,i)}(t)$

Los modelos de Greenwood y Reed-Frost Dentro de las CMTD se encuentran los modelos de Greenwood y Reed-Frost, el interés es revisarlos, aunque no en profundidad, para luego compararlos con el PSIR diferenciando ambas visiones.

Sean X_t y Y_t el número susceptibles e infectados al tiempo t , $t \in \{0, 1, \dots\}$ tal que $X_0 = x_0$ y $Y_0 = y_0 \geq 0$. Sea $p \in (0, 1)$ la probabilidad de contacto entre un susceptible y un infectado, el contacto resulta en una infección con probabilidad $\beta \in (0, 1)$.

Como $q = 1 - p$ es la probabilidad de que no haya contacto entonces la probabilidad de no haya una infección debido a ningún infectado es $1 - p + p(1 - \beta) = 1 - p\beta = \alpha$.

El modelo de Greenwood establece que en el tiempo $(t + 1)$, la cantidad de susceptibles más la cantidad de infectados debe ser igual a la cantidad de susceptibles en el tiempo t , es decir $X_{(t+1)} + Y_{(t+1)} = X_t$ lo que equivale a pensar que la cantidad de susceptibles en el tiempo $(t + 1)$ es igual a la cantidad de susceptibles disponibles en el tiempo t menos los que se infectaron en el tiempo $(t + 1)$, es decir $X_{(t+1)} = X_t - Y_{(t+1)}$. Se tiene entonces que

$$p_{(x,y)_t(x',y')_{t+1}} \equiv P((X, Y)_{t+1} = (x', y')_t | (X, Y)_t = (x, y)_{t+1}) = \binom{x_t}{x_{t+1}} \alpha^{x_{t+1}} (1 - \alpha)^{x_t - x_{t+1}} \quad (3.21)$$

Es decir, la cantidad de susceptibles que quedan en el sistema al tiempo $(t + 1)$ es un muestreo binomial de parámetros (x_t, α) , $X_{t+1} \sim \text{Bin}(X_t, \alpha)$. Nótese la equivalencia de la ecuación (3.21) con el muestreo dado por el PSIR. En el modelo de Greenwood α puede

ser considerada como la probabilidad de no infección, por lo que $\alpha^{x_{t+1}}$ es la probabilidad de que x_{t+1} susceptibles no sean infectados, esto implica que se toman a los susceptibles restantes como el muestreo, es decir, X_{t+1} es el número de no infectados en el tiempo $(t + 1)$ mientras que en el PSIR se muestrean a los nuevos infectados. Por tanto, en el modelo de Greenwood $\{X_t\}$ es una cadena de Markov en donde $P(X_{t+1} = x_{t+1} | X_t = x_t) = \binom{x_t}{x_{t+1}} \alpha^{x_{t+1}} (1 - \alpha)^{x_t - x_{t+1}}$.

El modelo de Greenwood supone que la causa de la infección no está relacionada con la cantidad de infectados por lo que α es un parámetro que no depende de Y por lo que el modelo de Reed-Frost es una modificación al modelo de Greenwood en donde la probabilidad de infección en el tiempo $(t + 1)$ es α^{Y_t} , esto es debido a que se considera que un susceptible en t sigue siendo susceptible en $(t + 1)$ si no hay contacto con ningún infectado en t por lo que la probabilidad de no infección es α^{Y_t} , así $X_{t+1} \sim \text{Bin}(X_t, \alpha^{Y_t})$ y sus probabilidades de transición en un paso están dadas por

$$P_{(x,y)_t(x',y')_{t+1}} = \binom{x_t}{x_{t+1}} (\alpha^{y_t})^{x_{t+1}} (1 - \alpha^{y_t})^{x_t - x_{t+1}}. \quad (3.22)$$

Así, el proceso $\{(X, Y)_t\}$ un proceso bivariado de Markov.

Los modelos de Greenwood y Reed-Frost vs el pSIR Los modelos de Greenwood y Reed-Frost son modelos epidemiológicos estocásticos para poblaciones pequeñas cuyos tiempos y estados son discretos, están basados en cadenas binomiales y en los que se plantea la dinámica de propagación de una enfermedad en términos de un muestreo binomial sobre los susceptibles, sin embargo, existen diferencias entre estos modelos y el pSIR que se ha planteado en este trabajo.

El modelo de Greenwood es un proceso estocástico homogéneo en el tiempo pues la distribución de sus v.a.'s no depende del instante en el que se encuentran, así podemos pensar en un proceso con incrementos independientes y estacionarios cuya matriz de transición de probabilidad para una cantidad x_0 de susceptibles y una probabilidad α de no contagio es

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 - \alpha & \alpha & 0 & 0 & \cdots & 0 \\ (1 - \alpha)^2 & 2(1 - \alpha)\alpha & \alpha^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ (1 - \alpha)^{x_0} & x_0(1 - \alpha)^{x_0 - 1} & \binom{x_0}{2}(1 - \alpha)^{x_0 - 2}\alpha^2 & \cdots & \cdots & \alpha^{x_0} \end{pmatrix}. \quad (3.23)$$

Para el modelo de Reed-Frost la matriz de transiciones de probabilidad es una matriz de

$(x_0 + 1)^2 x(x_0 + 1)^2$ dada por

$$P = \begin{pmatrix} P_{00} & 0 & 0 & 0 & \cdots & 0 \\ p_{10} & p_{11} & \cdots & \cdots & \cdots & P_{1x_0} \\ \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ P_{x_0 0} & P_{x_0 1} & \cdots & \cdots & \cdots & P_{x_0 x_0} \end{pmatrix}. \quad (3.24)$$

Donde cada P_{ij} es una matriz cuadrada de rango $(x_0 + 1)$ con $P_{ij} = (p_{(k,i)(l,j)})$, $k, l \in \{0, \dots, x_0\}$, $j + l = k$ y $p_{(k,i)(l,j)} = P((X, Y)_{t+1} = (l, j) | (X, Y)_t = (k, i)) = \binom{k}{l} (1 - \alpha^i)^j \alpha^{jl}$ y $p_{(k,i)(l,j)} = 0$ en otro caso, es decir P_{ij} registra las probabilidades de transición de $Y_t = i$ a $Y_{t+1} = j$.

Esto difiere a lo que ocurre en el PSIR, en el cual, no hay necesidad de fijarse en la matriz de transiciones de probabilidad, pues para cada instante en el tiempo las probabilidades dadas por $X(t)$ y $Y(t)$ caracterizan al sistema únicamente en ese instante, una forma de pensar la acción del PSIR es imaginar que en cada tiempo t_i el proceso de muestreo se reinicia con diferentes parámetros en sus distribuciones, esto es, si para el tiempo t , $S(t) = k$ entonces $X(t) \in \{0, 1, \dots, k\}$ por lo que en el tiempo $(t+1)$, $S(t+1) = i$ con i fijo, $0 \leq i \leq k$ y $X(t+1) \in \{0, 1, \dots, i\}$, reduciendo o igualando en el mejor de los casos el espacio de estados del proceso.

Una modificación del modelo de Reed-Frost tipo SIR El modelo de Reed-Frost tiene una modificación muy parecida al PSIR (Fresán-Figueroa et al.). Para una población de tamaño constante, contactos entre cualesquiera dos individuos por unidad de tiempo y una enfermedad sin tiempo de incubación en donde el individuo adquiere inmunidad después de enfermarse se plantea el sistema de ecuaciones

$$S_n = S_{n-1} - V_n \quad (3.25)$$

$$I_n = I_{n-1} + V_n - W_n \quad (3.26)$$

$$R_n = R_{n-1} + W_n \quad (3.27)$$

Donde S_n , I_n y R_n son las cantidades de susceptibles, infectados y removidos en la población al tiempo n , $V_n \sim \text{Bin}(S_{n-1}, \theta_n)$ y $W_n \sim \text{Bin}(I_{n-1}, \gamma)$. En este caso $\theta_n = 1 - (1 - \zeta)^{I_{n-1}}$ siendo ζ la probabilidad de que en un contacto susceptible-infectado, el susceptible adquiera la enfermedad por lo que θ_n es la probabilidad de que para un susceptible, adquiera la enfermedad debido a algún contacto con uno de los I_{n-1} infectados al tiempo $n - 1$. Por otro lado, γ es la probabilidad de que un infectado pase a ser removido, con γ constante que depende de la enfermedad.

En este caso, las probabilidades de transición dadas por (Fresán-Figueroaa et al.) son

$$\begin{aligned} P(S_n = s_n, I_n = i_n | S_{n-1} = s_{n-1}, I_{n-1} = i_{n-1}) &= P(W_n = \Delta s_n - \Delta I_n) P(V_n = \Delta s_n) \\ &= \binom{s_{n-1}}{\Delta s_n} (1 - (1 - \zeta)^{i_{n-1}})^{\Delta s_n} (1 - \zeta)^{i_{n-1} \Delta s_n} \binom{i_{n-1}}{\Delta s_n - \Delta i_n} \gamma^{\Delta s_n - \Delta i_n} (1 - \gamma)^{i_n \Delta s_n} \end{aligned}$$

La modificación del modelo de Reed-Frost considera que la probabilidad de remoción, aunque constante, es diferente para cada enfermedad. Esto junto con la dinámica de transmisión planteada en términos de muestreos Binomiales es similar a la dinámica planteada por el PSIR, sin embargo, la forma de considerar la probabilidad de infección así como el supuesto de mezcla homogénea en la población diferencian cualitativamente ambos modelos.

Capítulo 4

Resultados

Planteado el PSIR es necesario analizar que ocurre con el modelo cuando cambian las condiciones del tamaño poblacional, los parámetros en la probabilidad de infección y si realmente existe algún tipo de convergencia en el modelo al SIR determinista más allá de las similitudes entre sus ecuaciones.

Se comienza analizando $X(t)$ cuando el tamaño poblacional aumenta junto con una cantidad de susceptibles grande para luego centrarnos en los efectos que tienen tanto el parámetro β así como el cociente $\frac{I(t)}{N}$ en la probabilidad de infección.

4.1. PSIR para poblaciones grandes y su relación con el SIR determinista

En el PSIR, el número de nuevos infectados y el número de nuevos removidos en un intervalo de tiempo de la forma $(t, t + \delta)$ son variables aleatorias binomiales $X(t) \sim \text{Bin}(\lfloor \epsilon S(t) \rfloor, \beta \frac{I(t)}{N} \delta)$ y $Y(t) \sim \text{Bin}(I(t), \gamma \delta)$ por lo que la cantidad de susceptibles en el tiempo $(t + \delta)$ es $S(t + \delta) = S(t) - X(t)$ y la cantidad de infectados en $(t + \delta)$ es $I(t + \delta) = I(t) + X(t) - Y(t)$ de esta manera, para cada tiempo del muestreo t_i con $i \in \{0, 1, 2, \dots\}$ hay una v.a. binomial $X(t_i)$ y $Y(t_i)$, en donde las colecciones $\{X_{t_i}\}$ y $\{Y_{t_i}\}$ de todas las variables aleatorias generan el proceso estocástico de la cadena de transmisión.

Nótese que en cada X_{t_i} los parámetros de las distribuciones cambian conforme pasa el tiempo, esto es ocurre pues la cantidad de susceptibles baja y la de infectados sube. Esto implica que en los tiempos de muestreo t_i y t_j , con $i \neq j$, se tenga que $S(t_i) \neq S(t_j)$ y también $\frac{I(t_i)}{N} \neq \frac{I(t_j)}{N}$.

El modelo está pensado para capturar la naturaleza discreta del muestreo y de esta manera el tamaño poblacional no sea un problema. El interés es entonces identificar que ocurre con el modelo cuando la población es muy grande.

Considerando una población de tamaño N , la ecuación para la cantidad de susceptibles

$S(t+\delta) = S(t) - X(t)$ es equivalente a $S(t+\delta) - S(t) = -X(t)$ por lo que $\frac{S(t+\delta) - S(t)}{N} = \frac{-X(t)}{N}$, donde $\frac{X(t)}{N} \in \{0, 1/N, 2/N, \dots, \lfloor \epsilon S(t) \rfloor / N\}$ es la proporción de individuos infectados durante el intervalo $(t, t+\delta)$. Si nombramos $x(t) = \frac{\lfloor \epsilon S(t) \rfloor}{N}$ se tiene que $x(t+\delta) - x(t) = -\frac{X(t)}{N}$. Nótese que del lado derecho se tiene una v.a. mientras que del izquierdo son valores reales fijos, por lo que al calcular la esperanza de ambos lados se tiene que $x(t+\delta) - x(t) = E(-\frac{X(t)}{N})$. Del lado derecho $E(-\frac{X(t)}{N}) = -\frac{1}{N}E(X(t)) = -\frac{1}{N}(\lfloor \epsilon S(t) \rfloor \beta \frac{I(t)}{N} \delta) = -\beta \frac{\lfloor \epsilon S(t) \rfloor}{N} \frac{I(t)}{N} \delta$. Para un nivel de exposición $\epsilon = 1$, lo que implica la mezcla homogénea, la última parte de la igualdad es $-\beta x(t)y(t)\delta$, donde $y(t) = \frac{I(t)}{N}$. Se concluye entonces que la proporción esperada de susceptibles en el intervalo δ , $x(t+\delta) - x(t) = -\beta x(t)y(t)\delta$. Si hacemos tender N a infinito los valores de la ecuación son densidades poblacionales por lo que $\frac{x(t+\delta) - x(t)}{\delta} = -\beta x(t)y(t)$ es el cambio en la densidad de susceptibles con respecto al tiempo, tomando el límite cuando δ tiende a 0 se tiene que $\partial_t x = -\beta x(t)y(t)$. Esto nos indica que la ecuación (2.13) del modelo SIR determinista se alcanza como caso límite de la proporción esperada de susceptibles de la ecuación (3.1) cuando la cantidad de individuos en la población es grande, $N \rightarrow \infty$, la exposición al contacto $\epsilon = 1$ y el intervalo del muestreo es muy pequeño, $\delta \rightarrow 0$.

Para la ecuación de la cantidad de infectados se tiene $I(t+\delta) = I(t) + X(t) - Y(t)$ lo que implica que $I(t+\delta) - I(t) = X(t) - Y(t)$, entonces $\frac{I(t+\delta) - I(t)}{N} = \frac{X(t) - Y(t)}{N}$ lo que implica que $y(t+\delta) - y(t) = \frac{X(t) - Y(t)}{N}$. Del lado izquierdo tenemos que $E(y(t+\delta) - y(t)) = y(t+\delta) - y(t)$ mientras que del lado derecho, $E(\frac{X(t) - Y(t)}{N}) = \frac{1}{N}E(X(t) - Y(t)) = \frac{1}{N}(\lfloor \epsilon S(t) \rfloor \beta \frac{I(t)}{N} \delta - I(t)\gamma\delta) = \beta \frac{\lfloor \epsilon S(t) \rfloor}{N} \frac{I(t)}{N} \delta - \gamma \frac{I(t)}{N} \delta = \beta x(t)y(t)\delta - \gamma y(t)\delta$, cuando $\epsilon = 1$, se concluye entonces que la proporción esperada de infectados es $y(t+\delta) - y(t) = \beta x(t)y(t)\delta - \gamma y(t)\delta$. Haciendo el cociente entre δ se tiene $\frac{y(t+\delta) - y(t)}{\delta} = \beta x(t)y(t)\delta - \gamma y(t)$ y tomando el límite cuando $\delta \rightarrow 0$ se tiene $\partial_t y(t) = \beta x(t)y(t)\delta - \gamma y(t)$. Esto implica que el cambio en los infectados del SIR dado por la ecuación (2.14) es la proporción esperada de infectados del PSIR cuando el tamaño poblacional es grande, $N \rightarrow \infty$, la exposición al contacto $\epsilon = 1$, y el tiempo entre muestreos es pequeño $\delta \rightarrow 0$.

Análogamente, al realizar el mismo análisis que con las dos ecuaciones anteriores para la ecuación para los removidos del PSIR dada como $R(t+\delta) = R(t) + Y(t)$, se tiene que la ecuación (2.15) del SIR determinista se alcanza como caso límite de la proporción esperada de removidos en el PSIR cuando $N \rightarrow \infty$ y $\delta \rightarrow 0$.

Resumiendo, se puede argumentar que el modelo SIR determinista es un caso límite del comportamiento esperado del PSIR cuando el tamaño poblacional es muy grande, lo suficiente como para que las proporciones de susceptibles, infectados y removidos puedan ser tomadas como densidades poblacionales y cuando el tiempo entre los muestreos es muy chico. (Figura 4.1)

Otro factor a tomar en cuenta es cuando $N \rightarrow \infty$ y la cantidad de susceptibles $S(t) \approx N$ y $\epsilon = 1$, es decir, una población grande con cantidad de susceptibles parecida al total y mezcla homogénea. Se tiene entonces que $\frac{S(t)}{N} \rightarrow 1$ mientras que $\frac{I(t)}{N} \rightarrow 0$, en principio entonces, se tiene una probabilidad de contacto con un infectado baja por el simple hecho

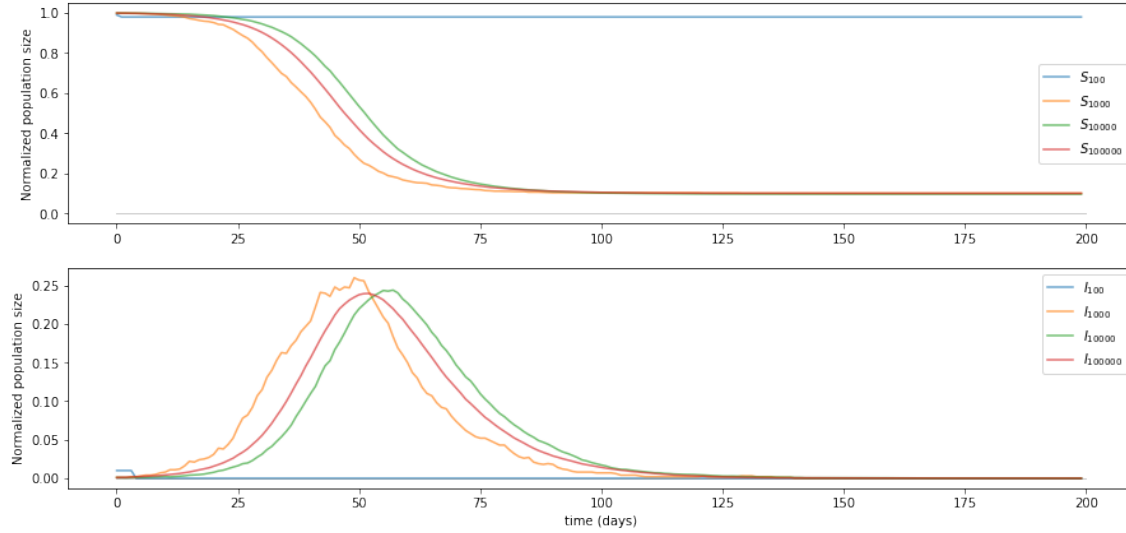


Figura 4.1: Comportamiento del PSIR al aumentar el tamaño poblacional de susceptibles.

de que la cantidad de susceptibles es casi el total. De la ecuación para los susceptibles de PSIR se tiene que $E(S(t + \delta) = E(S(t)) - X(t))$, lo que implica que en términos de esperanzas $S(t + \delta) = S(t) - E(X(t))$ y por tanto $S(t + \delta) = S(t) - S(t)\beta\frac{I(t)}{N}\delta$, es decir, en promedio, $S(t + \delta) = S(t)(1 - \frac{I(t)}{N}\delta)$, luego $(1 - \frac{I(t)}{N}\delta) \approx 1$ pues $\frac{I(t)}{N} \rightarrow 0$ lo que implica que en promedio $S(t + \delta) \approx S(t)$ cuando el tamaño poblacional es muy grande y la cantidad de susceptibles es parecida al total. El que la cantidad esperada de susceptibles en el siguiente tiempo sea aproximadamente la cantidad de susceptibles en el tiempo anterior, bajo las condiciones dichas, implica que no hay una disminución drástica en el tamaño de los susceptibles por lo que no se infectarán tantos individuos en el siguiente muestreo a nivel poblacional.

Por otro lado, al calcular la esperanza de ambos lados en la ecuación para los infectados del PSIR se tiene $E(I(t + \delta) = E(I(t) + X(t) - Y(t))$, nuevamente la esperanza de $I(t)$ es ella misma, por lo que en promedio $I(t + \delta) - I(t) = E(X(t)) - E(Y(t))$. Así la cantidad promedio de infectados en $(t + \delta)$ es $I(t + \delta) - I(t) = +S(t)\beta\frac{I(t)}{N}\delta - I(t)\gamma\delta$. Al factorizar $I(t)$ del lado derecho se tiene $I(t + \delta) - I(t) = I(t)(\beta\frac{S(t)}{N}\delta - \gamma\delta)$, dado que $\frac{S(t)}{N} \approx 1$ la expresión termina siendo $I(t + \delta) - I(t) \approx I(t)(\beta - \gamma)\delta$. Esta ecuación tiene sentido solo cuando $I(t)(\beta - \gamma)\delta \geq 0$ lo que implica que $(1 + (\beta - \gamma)\delta) \geq 0$, luego $(\beta - \gamma)\delta \geq 0$, dado que $\delta \geq 0$ se tiene que $\beta - \gamma \geq 0$ lo que implica que $\beta \geq \gamma$ y por tanto $\frac{\beta}{\gamma} \geq 1$ o $R_0 \geq 1$. Esto es una fuerte conexión entre el PSIR y el SIR determinista pues el número esperado de infectados depende en el fondo del radio básico de reproducción.

Obsérvese también que la varianza de $X(t)$ está dada como $Var(X(t)) = S(t)\beta\frac{I(t)}{N}\delta(1 -$

$\beta \frac{I(t)}{N} \delta$) bajo el supuesto de población grande y una cantidad de susceptibles parecida al total se tiene que $Var(X(t)) \approx \beta I(t) \delta$ pues $\frac{I(t)}{N} \rightarrow 0$ cuando $N \rightarrow \infty$ mientras que $\frac{S(t)}{N} \rightarrow 1$. Por otro lado $E(X(t)) = S(t) \beta \frac{I(t)}{N} \delta \approx \beta I(t) \delta$, esto implica que tanto la esperanza y la varianza del número de nuevos infectados es muy parecida para poblaciones grandes y con casi la misma cantidad de susceptibles como el total y ambas se pueden ver como una función lineal de los infectados justo en ese momento. Esto ayuda a entender porque el número de nuevos casos infectados no se dispara abruptamente para poblaciones muy grandes.

En los casos del PSI y la extensión metapoblacional del PSIR, se tienen comportamientos similares. Para el PSI, las curvas generadas por el modelo para poblaciones al aumentar la población tienden a suavizarse (Figura 4.2).

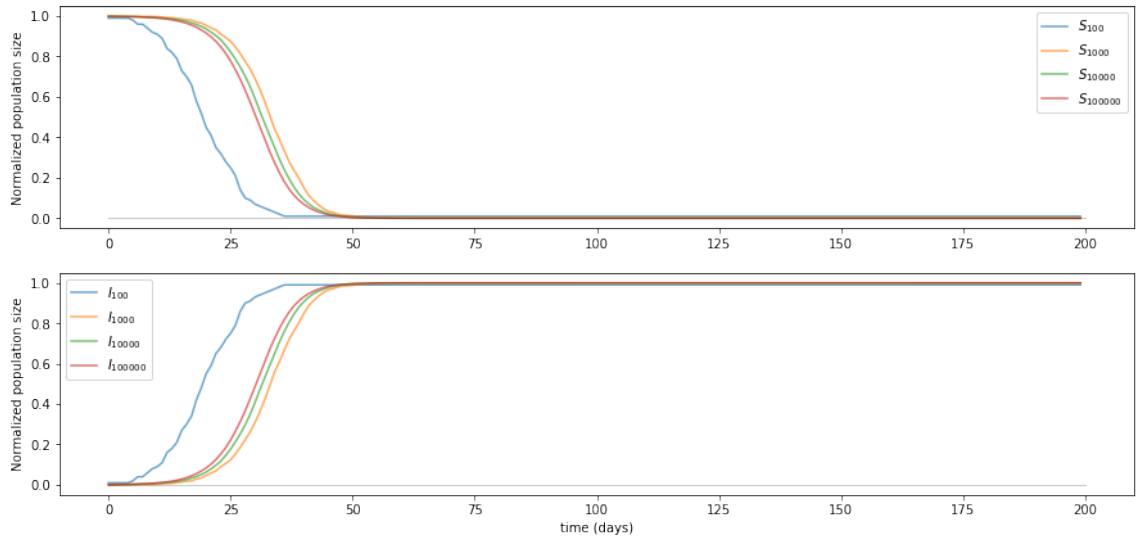


Figura 4.2: Comportamiento del PSI al aumentar el tamaño poblacional de susceptibles con $S = 100, 1000, 10\ 000, 100\ 00, 1\ 000\ 000$.

Mientras que para la extensión metapoblacional se observa el mismo comportamiento cuando la cantidad de susceptibles incrementa (Figura 4.3).

4.2. Convergencia del PSIR a un Proceso Poisson no Homógeno en el tiempo

Consideramos $0 < \delta \ll 1$ fijo y el intervalo $(t, t + \delta)$. Definimos la malla $\{t_i\}_{i=1}^n = \{t + \delta(\frac{\lambda}{n})\}_{i=1}^n$ en $(t, t + \delta)$. Dado que $\delta \ll 1$, al aproximar la distribución de cada variable se tiene $X(t_i) \underset{aprox}{\sim} \text{Bin}(\lfloor \epsilon S(t) \rfloor, \beta \frac{I(t) \delta}{N} (\frac{1}{n}))$ y $Y(t_i) \underset{aprox}{\sim}$

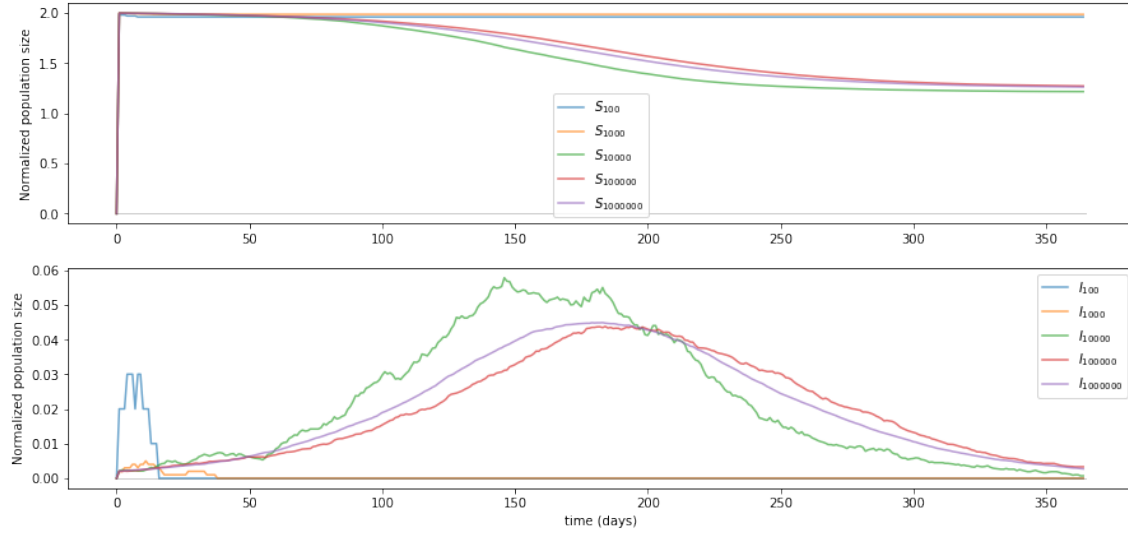


Figura 4.3: Comportamiento del PSIR Metapoblacional al aumentar el tamaño poblacional de susceptibles con $S = 100, 1000, 10000, 100000, 1000000$.

$\text{Bin}(I(t), \gamma\delta(\frac{1}{n}))$), donde los procesos $\{X(t_i)\}, \{(t_i)\}$ son independientes. Así,

$$X^n(t) := \sum_{i=1}^n X(t_i) \sim \text{Bin}(n \lfloor \epsilon S(t) \rfloor, \beta(\frac{I(t)}{N})\delta(\frac{1}{n}))$$

$$Y^n(t) := \sum_{i=1}^n Y(t_i) \sim \text{Bin}(nI(t), \gamma\delta(\frac{1}{n}))$$

Por lo que $X^n(t) \xrightarrow{n \rightarrow \infty} \text{Poisson}(\lfloor \epsilon S(t) \rfloor \beta \frac{I(t)}{N} \delta)$ y $Y^n(t) \xrightarrow{n \rightarrow \infty} \text{Poisson}(I(t)\gamma\delta)$. Esto representa el límite cuando la malla discreta tiene norma tendiendo a cero, de los cambios en el intervalo $(t, t + \delta)$.

Por la propiedad de independencia de los incrementos, se sigue que el proceso a tiempo continuo ($\delta \rightarrow 0$) está dado por el sistema acoplado de Procesos Poisson no homogéneos

$$S(t) = S(0) - N_1(t) \tag{4.1}$$

$$I(t) = I(0) + N_1(t) - N_2(t) \tag{4.2}$$

$$R(t) = R(0) + N_2(t). \tag{4.3}$$

Donde $N_1(t) \sim \text{P.P. no Homogéneo}$ con intensidad $S(t)\beta\frac{I(t)}{N}$ y $N_2(t) \sim \text{P.P. no Homogéneo}$ con intensidad $I(t)\gamma$.

4.3. Efecto del parámetro p de $X(t)$ en el comportamiento del PSIR

En la v.a. $X(t)$, la probabilidad de contagio es $p = \beta \frac{I(t)}{N} \delta$ donde β es la probabilidad de contagio efectivo, $\frac{I(t)}{N}$ la probabilidad de contacto con un individuo infectado y δ la unidad de tiempo en el que se realiza el muestreo, por tanto p depende de $I(t)$ y β .

En las distribuciones binomiales las probabilidades de éxito p pequeñas generan el efecto de decantar la función de masa hacia la izquierda, mientras que para p grande la función de masa se decanta hacia la derecha, llegando a una simetría de la función de masa cuando $p = \frac{1}{2}$. Es por esto que el valor de p juega un factor en la dispersión de la enfermedad. Valores de p pequeños al tiempo t hacen que sea más probable que un grupo reducido de susceptibles sea infectado en el tiempo $t + \delta$ y recíprocamente p grandes generarán que haya un mayor número de infectados en el siguiente tiempo.

Dado que δ es la unidad de tiempo del muestreo, esta debe ser consistente siempre, las unidades del tiempo deben ser las mismas durante todo el proceso y no es válido comenzar con un muestreo en días para luego cambiar a semanas.

Tomando β e $\frac{I(t)}{N}$ y considerando δ fija, la probabilidad de contagio p es pequeña si $\beta \approx 0$ y $\frac{I(t)}{N} \approx 0$ o alguna de las dos es pequeña.

Una cantidad $\frac{I(t)}{N} \approx 0$ implica que hay muy pocos infectados al tiempo t en comparación con el tamaño de la población por lo tanto la probabilidad de contagio será baja, del otro lado $\beta \approx 0$ implica que la probabilidad de transmisión efectiva es muy chica, lo que se interpreta como una enfermedad difícil de contagiar aún entrando en contacto con un infectado.

Caso contrario es cuando $\frac{I(t)}{N} \approx 1$ y cuando $\beta \approx 1$, el primero nos dice que el número de infectados es muy parecido al total de la población por lo que la enfermedad esta propagada en casi todos los individuos, mientras que $\beta \approx 1$ nos indica una probabilidad de contagio efectiva muy alta por lo que estaríamos ante el caso de una enfermedad infecciosa muy fácil de contagiar.

Si tanto $\frac{I(t)}{N}$ como β son ambos aproximados a cero o son muy pequeños entonces $p = \beta \frac{I(t)}{N} \delta$ será en general pequeño por lo que nos encontraríamos con una enfermedad difícil de contagiar con una cantidad baja de infectados por lo que su dispersión se intuye debería de ser lenta y con una propagación pequeña a través de la población.

El caso extremo es cuando $\frac{I(t)}{N}$ y β son aproximadamente 1 en donde tenemos una gran cantidad de infectados y una enfermedad fácil de contagiar por lo que el contagio debe de ser rápido y de gran alcance.

Si $\frac{I(t)}{N} \approx 1$ y β no, entonces $\beta \frac{I(t)}{N} \approx \beta$ por lo que la probabilidad de contagio general será en mayor medida determinado por el tamaño de β y por lo tanto de la facilidad con la que la enfermedad se contagia.

Por otro lado, si $\beta \approx 1$ e $\frac{I(t)}{N}$ no, entonces $\beta \frac{I(t)}{N} \approx \frac{I(t)}{N}$ y por lo tanto la probabilidad de contagio dependerá en mayor medida de la cantidad de infectados.

El efecto que tienen ambas cantidades en la probabilidad de contagio crea un efecto similar al que se tendría cuando se contrae y expande un acordeón pues una puede subir mientras que la otra baja logrando una mayor flexibilidad a la hora de intentar ver como es dicha probabilidad de contagio (Figura 4.4).

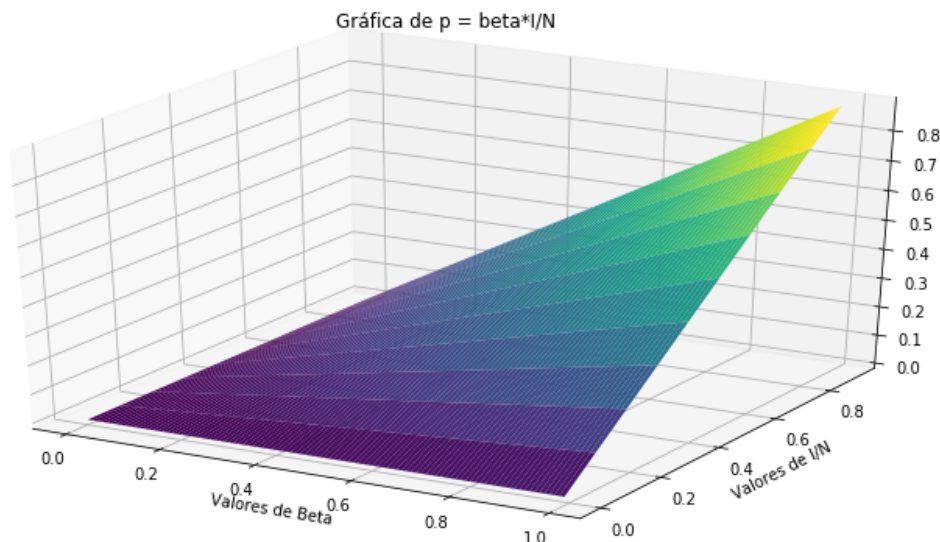


Figura 4.4: Gráfica del producto $\beta \frac{I}{N}$.

Ejemplos de los efectos que tiene el parámetro p de la distribución de $X(t)$ en la dinámica del modelo aparecen en las gráficas de las Figuras 4.5 y 4.6. En estas, las realizaciones del PSIR tienen como parámetro de transmisión efectiva $\beta = 0.5$ y $\beta = 0.7$ respectivamente. El mismo análisis realizado en los parámetros de las distribuciones el PSIR es válido para el PSIR y la extensión metapoblacional del PSIR, siendo la probabilidad de contagio un factor cuyo efecto flexibiliza las condiciones en las actúa la cadena de transmisión. Para el caso del PSI, las gráficas de las realizaciones del modelo cuando $\beta = 0.5$ puede observarse en la Figura 4.7 y con $\beta = 0.7$ en la Figura 4.8.

Para la extensión metapoblacional, las figuras resultan de las realizaciones del modelo con $\beta = 0.5$ y $\beta = 0.7$ son 4.9 y 4.10 respectivamente.

El cambio en las probabilidades de infección en cada muestreo implica un cambio en las funciones de masa de probabilidad para cada una de las v.a. $X(t)$ (Figura 4.11) y también implica que diferentes factores en la cadena de transmisión terminan ocasionando resultados similares.

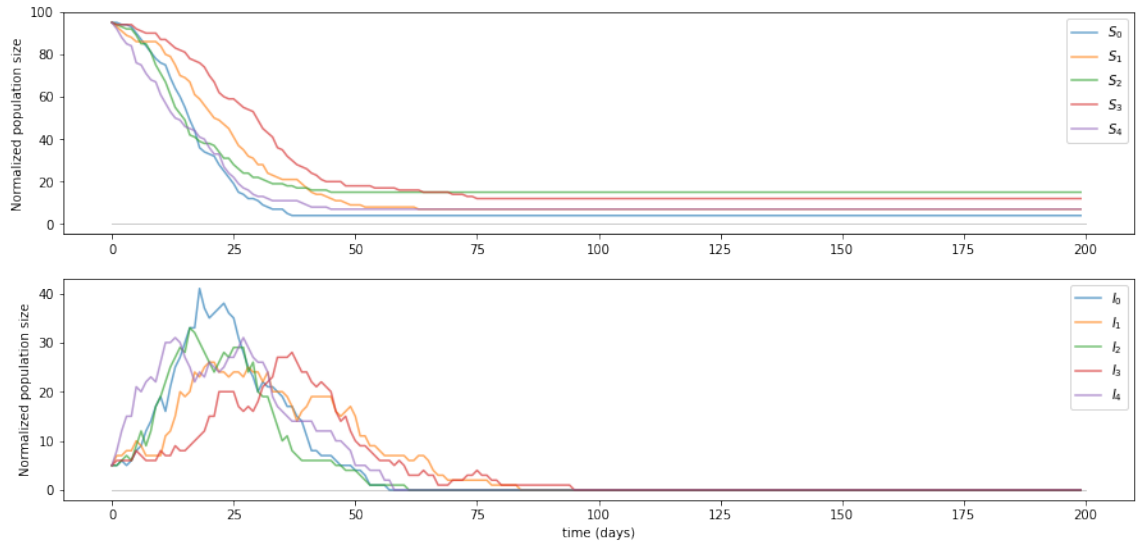


Figura 4.5: Comportamiento del PSIR considerando que la mitad de la población es expuesta al contagio por unidad de tiempo.

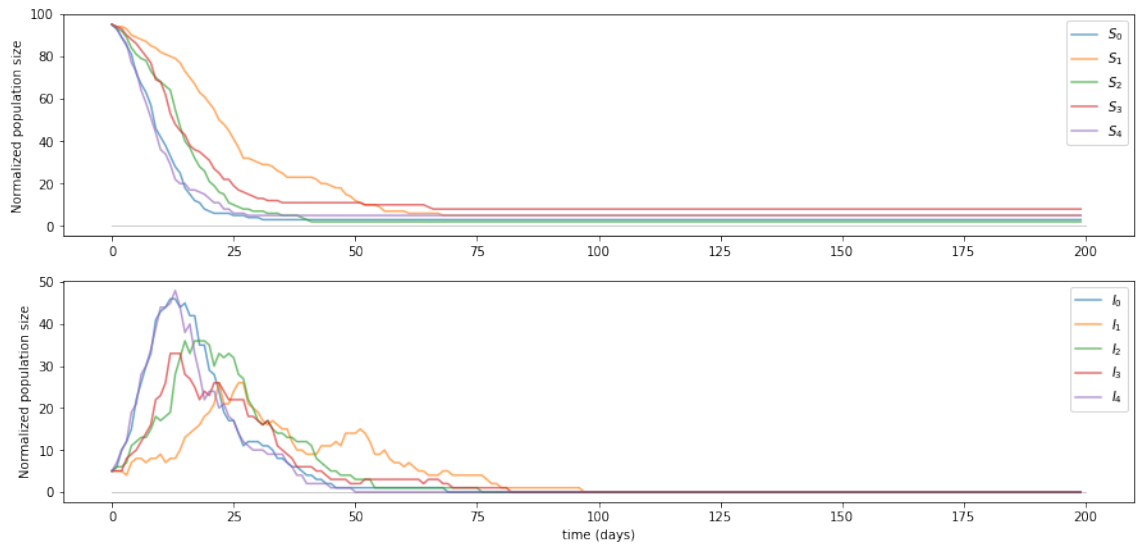


Figura 4.6: Comportamiento del PSIR considerando que $7/10$ de la población susceptible es expuesta al contagio por unidad de tiempo.

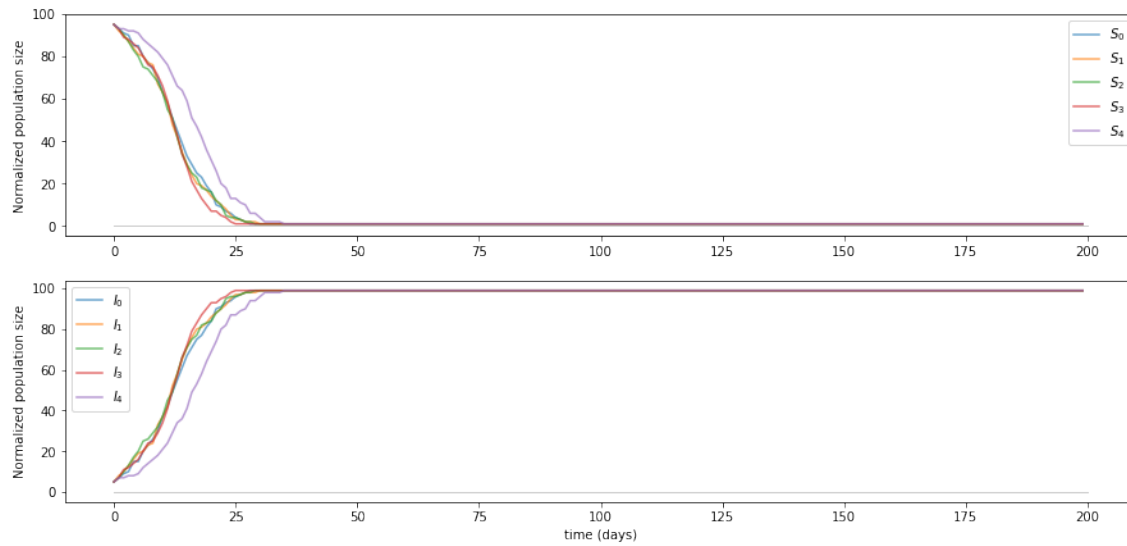


Figura 4.7: Comportamiento del PSI considerando que un medio de la población susceptible es expuesta al contagio por unidad de tiempo.

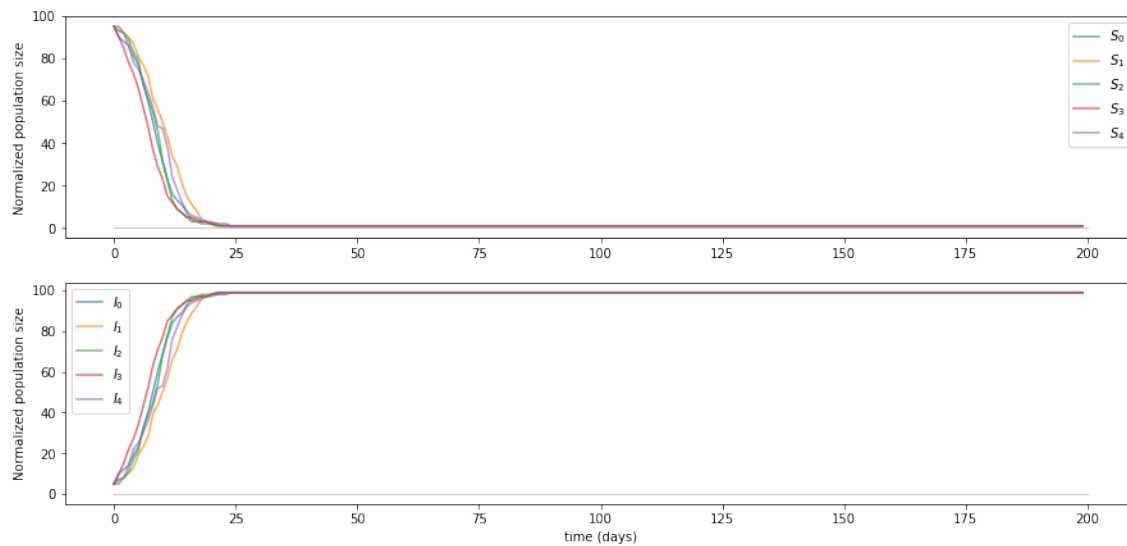


Figura 4.8: Comportamiento del PSI considerando que $7/10$ de la población susceptible es expuesta al contagio por unidad de tiempo.

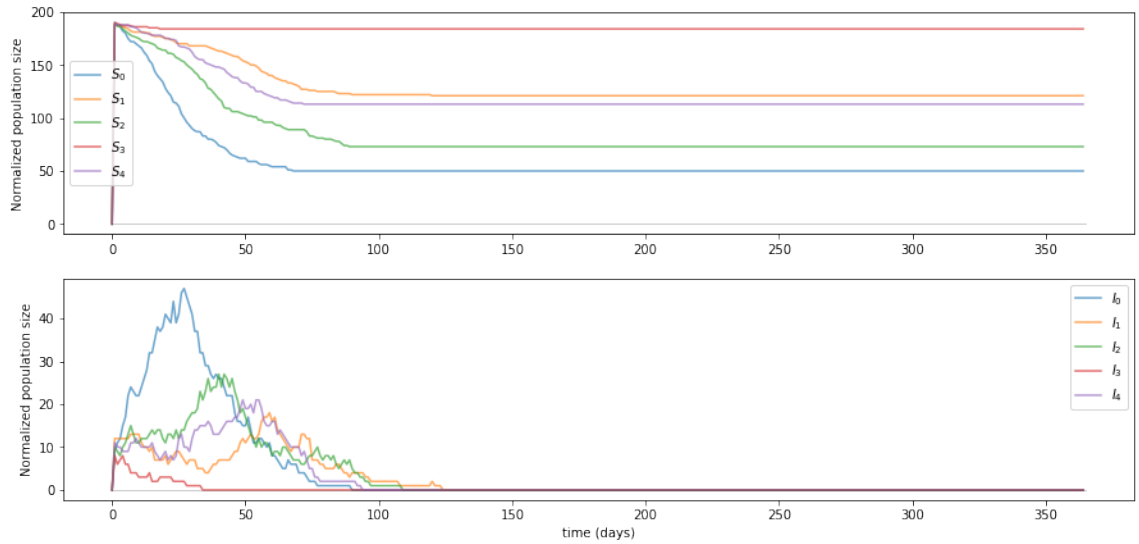


Figura 4.9: Comportamiento del PSIR Metapoblacional considerando que un medio de la población susceptible es expuesta al contagio por unidad de tiempo.

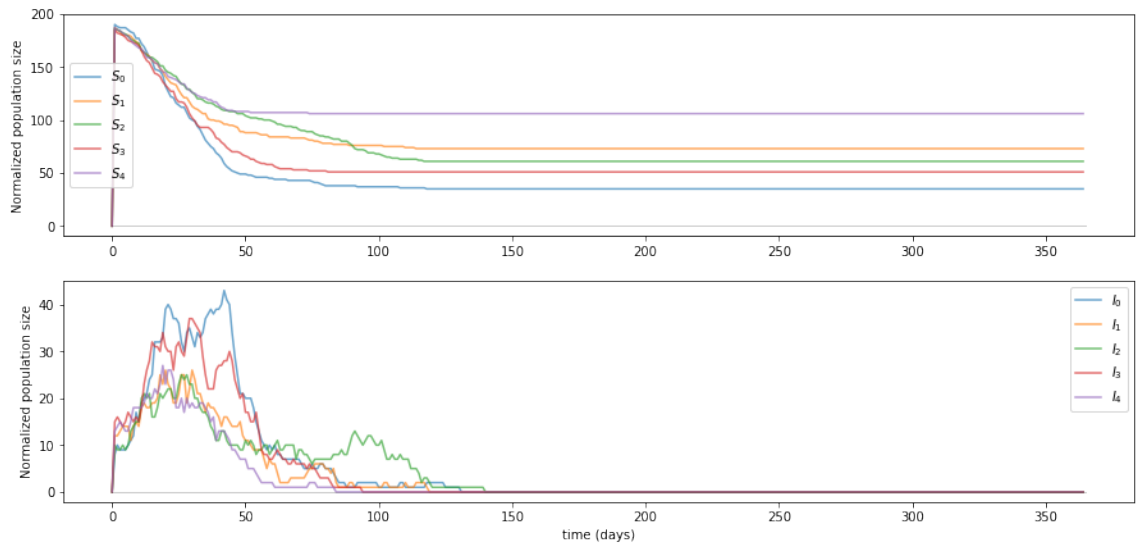


Figura 4.10: Comportamiento del PSIR Metapoblacional considerando que 7/10 de la población susceptible es expuesta al contagio por unidad de tiempo.

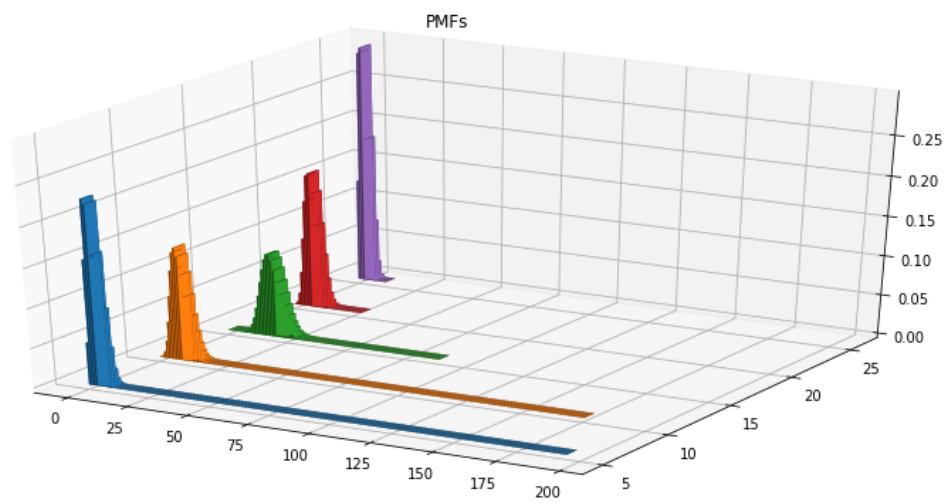


Figura 4.11: Evolución en las funciones de masa de probabilidad para diferentes tiempos del muestreo.

4.4. Modelo aplicado a influenza

La influenza es un virus extendido a nivel mundial, el cual forma parte de las infecciones respiratorias agudas, además, hay tres tipos de virus de influenza, A, B y C, dividiéndose el tipo A en los subtipos que se generan debido a las proteínas Hemaglutinina y Neuraminidasa, habiendo 16 y 9 tipos de estas proteínas respectivamente. Para este caso en específico del tipo A, el virus se propaga mediante epidemias estacionales que se repiten cada año, lo que genera que anualmente los gobiernos gasten grandes cantidades de dinero en el tratamiento y hospitalización de individuos infectados con este virus.

Las epidemias más severas registradas se presentaron en 1918, ocasionada por el A-H1N1, 1957 por el A-H2N2, 1968 por el A-H3N2 y 2009 ocasionada por el A-H1N1.

La transmisión es de persona a persona, principalmente mediante el aerosol que se genera por toser o estornudar. El tiempo de recuperación en los infectados pueden pasar de las dos semanas hasta que comienzan signos de mejoría en su condición.

Para la aplicación del PSIR a la influenza se considera $R_0 = 2.5$, mientras que $\gamma = \frac{1}{15} \approx 0.066$ por lo que $\beta \approx 0.166$, las gráficas de las realizaciones del PSIR muestran el comportamiento del modelo a lo largo de un año cuando la cantidad de susceptibles aumenta. En el caso de la Figura ?? se tiene que un medio de la población susceptible está expuesta a la infección, mientras que en la Figura 4.13, 7/10 de la población susceptible está expuesta.

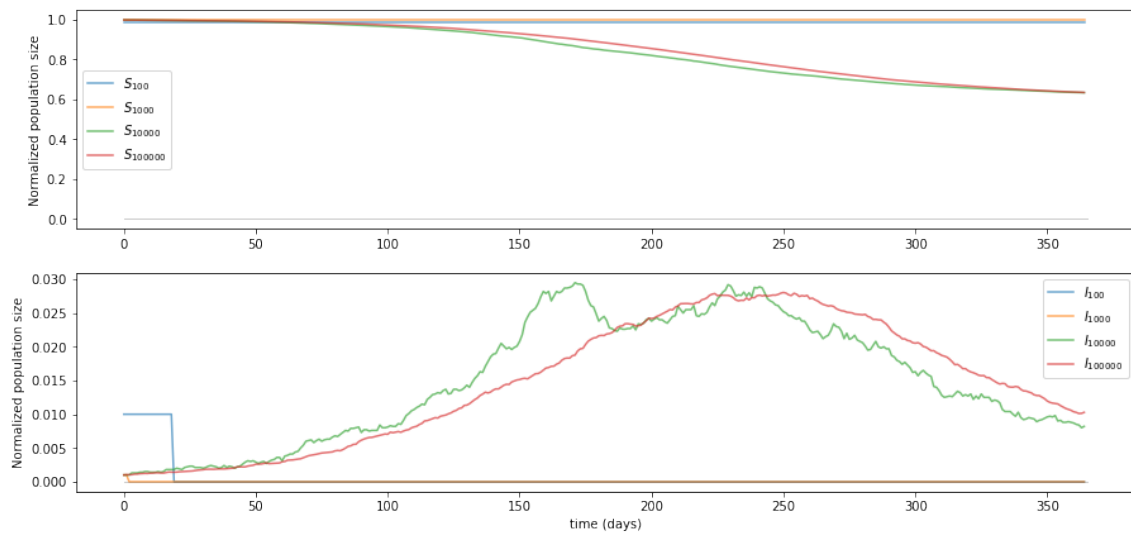


Figura 4.12: Aplicación del PSIR a influenza a 1/2 exposición.

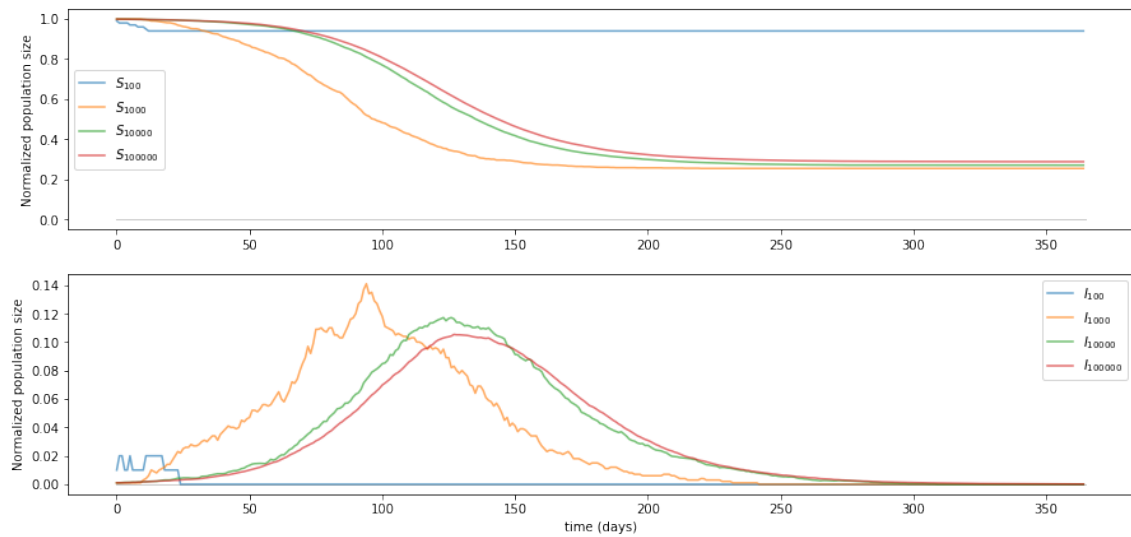


Figura 4.13: Aplicación del PSIR a influenza con 7/10 de exposición.

4.5. Modelo aplicado a VIH

El VIH tomó por sorpresa al mundo durante la década 1980, es un lentivirus de la familia retroviridae que causa el síndrome de inmunodeficiencia adquirida (SIDA). Hay dos tipos, el VIH-1 y el VIH-2 siendo el primero el que se ha extendido por todo el mundo mientras que el segundo se presenta casi exclusivamente en África.

La principal forma de contagio es mediante contacto sexual sin protección, sin embargo, otras formas de contagio son la transfusión sanguínea y el uso compartido de jeringas-agujas por parte de los infectados.

Debido a que el virus causa un deterioro gradual en el sistema inmune, atacando los linfocitos T, el individuo poco a poco ve comprometida su salud por lo que su remoción del sistema tardará un tiempo considerable. Además el uso de medicamentos antirretrovirales hace que baje el índice de viremia en sangre por lo que este tiempo de remoción puede extenderse todavía más.

4.5.1. Extensión para considerar grupos de riesgo

Los grupos de riesgo aparecen debido a que en determinadas enfermedades hay subpoblaciones más propensas a adquirir la enfermedad, esto dependiendo de diferentes factores. En el caso de la influenza los grupos pueden situarse por edad de los individuos, siendo los jóvenes y los ancianos más propensos a contraer el virus u enfermarse. Mientras, para el VIH los grupos de riesgo son los hombres que tienen sexo con hombres, individuos transgeneros, comerciantes sexuales y los usuarios de drogas inyectables.

Considerando esto, para utilizar la extensión metapoblacional del PSIR habría que tener en claro la cantidad de individuos que pertenecen a estos grupos, la proporción de individuos de cada grupo en contacto con los demás, las probabilidades de transmisión de cada grupo y las probabilidades de remoción.

Capítulo 5

Discusión

5.1. Ventajas

En el Capítulo 2 de este trabajo se planteó el modelo SIR determinista bajo los supuestos importantes de poblaciones constantes grandes y mezcladas de manera homogénea. Dichos supuestos son un problema, no en el sentido matemático, pues el modelo está bien definido y funciona, sino en el sentido práctico al querer aplicarlo en poblaciones pequeñas. Siendo realistas no importa que tan grande sea una población, esta seguirá siendo finita y de manera estricta no se alcanzará el tamaño suficiente como para que las densidades poblacionales de susceptibles, infectados y removidos puedan ser efectivamente continuas en los reales, y aún relajando la continuidad de las variables, hay enfermedades que están sujetas a regiones particulares del mundo, en donde el tamaño de las poblaciones no alcanza a ser suficiente.

Una justificación del por qué se puede tomar una población constante en el SIR es debido a que el periodo en el que se desarrolla la epidemia es corto por lo que no habrá un cambio significativo en el tamaño poblacional, esto tampoco es cierto debido a que en particular en poblaciones grandes hay nacimientos y muertes constantemente.

Por otro lado la mezcla homogénea supone que todos los individuos puedan tener contacto entre sí. Esto es criticable pues en poblaciones grandes, como la Ciudad de México que tiene un aproximado de 8.9 millones de habitantes hasta el 2015, es poco realista que todos puedan tener contacto con todos. Para poblaciones pequeñas, pueblos pequeños y comunidades chicas de a lo más un par de miles de habitantes posiblemente si se pueda considerar como válida la mezcla homogénea, sin embargo se vuelve a caer en el problema del tamaño poblacional.

Retomar estos dos factores es necesario para puntualizar una ventaja del PSIR, pues al tener un modelo basado en un muestreo discreto de infectados y removidos permite tomar cualquier tamaño poblacional y trabajar con este. En segundo lugar no supone una mezcla homogénea en la población debido al factor de exposición ϵ por parte los susceptibles

haciendo que no todos entren en la cadena de transmisión, por lo que tampoco se tiene ese problema.

Todo el modelo trabaja con base en el sentido común del muestreo y a pesar de ser un modelo estocástico basado en una cadena de Markov no es necesario el uso de la matriz de transición de probabilidad, lo cual puede ser útil si se compara la extensión metapoblacional del PSIR con el modelo de Reed-Frost en el cual dicha matriz podría traer problemas al querer programarlo debido a su tamaño.

La naturaleza discreta del modelo permite tener una buena cuantificación de los nuevos casos pues se todo el tiempo se trabaja con valores enteros. En cuanto a la parte computacional, el modelo resulta ser sencillo y tampoco trae consigo grandes complicaciones en su implementación.

Por último, el PSIR tiene una conexión con el SIR determinista al tratarse este último como un caso límite cuando justo la población es muy grande y el tiempo entre los muestreos es muy chico.

5.2. Desventajas

El PSIR es un modelo planteado en este trabajo por lo que todavía falta analizar con mayor profundidad su efectividad real.

Se desarrolló el modelo planteando una serie de supuestos, que al observar sus características y ponerlos juntos en una estructura matemática resulto equivalente a considerar un muestreo binomial, sin embargo, eso no quiere decir que este sea el único muestreo válido. Para plantear un modelo de este tipo pueden haber enfermedades cuyo muestreo tenga otro tipo de distribución. En ese sentido, el PSIR no es necesariamente un modelo que capture la cadena de transmisión de todas las enfermedades infecciosas.

Esto está relacionado con la extensión metapoblacional del PSIR en donde la forma en la que se distribuyen los individuos cambie dependiendo del grupo al que pertenezca y no sea necesariamente de comportamiento binomial.

También resulta importante en la extensión del PSIR que al considerar dos o tres subpoblaciones no se tiene tanto problema para realizar el computo del modelo, sin embargo, no se ha hecho el programa para más poblaciones y no se sabe cual será el extensivo computacional.

5.3. Posibles extensiones del PSIR

Dado que identificar el como se da la dinámica de transmisión, cuál es el tipo de contacto infeccioso y cómo generar un muestreo a partir de dicha información es clave para generar el modelo PSIR, lo primero que se puede plantear es si hay enfermedades cuya dinámica de muestreo se ajuste a otros tipos de distribución de probabilidad en lugar de la binomial usada.

Otra punto de interés es ver si se puede ampliar el modelo tomando en cuenta los mo-

vimientos humanos dentro de una región geográfica como un factor de importancia en la propagación de las enfermedades.

También sería interesante ver si puede incorporar dentro del modelo los lugares públicos como focos de infección.

Por último, una modificación importante puede ser el cambiar el estado del tiempo del muestreo de discreto a continuo por lo que las v.a. tendrían dentro de sus parámetros densidades de probabilidad que dependan del tiempo y ver que condiciones se necesitan para seguir preservando la markovianidad del modelo.

Bibliografía

Linda JS Allen, Fred Brauer, Pauline Van den Driessche, and Jianhong Wu. *Mathematical epidemiology*, volume 1945. Springer, 2008.

Fred Brauer, Carlos Castillo-Chavez, and Carlos Castillo-Chavez. *Mathematical models in population biology and epidemiology*, volume 1. Springer, 2001.

Fred Brauer, PV den DRIESSCHE, and Jianhong Wu. *Lecture notes in mathematical epidemiology*. Berlin, Germany: Springer, 2008.

Julian Alberto Fresán-Figueroaa, Arturo Erdely Ruiza, and Ismael Velázquez Ramireza. Un modelo sir probabilístico.

WO Kermack and AG McKendrick. Contributions to the mathematical theory of epidemics. *Proc Roy Soc Lond*, 115:700–721, 1927.