Modelo SEIR con compartimientos para la propagación del Covid-19 en el Pacífico Nariñense

Andrés Felipe Rincón Prieto

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Estadística
Bogotá, Colombia
2023

Modelo SEIR con compartimientos para la propagación del Covid-19 en el Pacífico Nariñense

Andrés Felipe Rincón Prieto

Tesis o trabajo de grado presentado como requisito parcial para optar al título de:
Maestría en Ciencias Estadística

Director(a):
PnD Viswanathan Arunachalam

Línea de Investigación:
Procesos Estocásticos
Grupo de Investigación:
Procesos Estocásticos

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Estadística
Bogotá, Colombia
2023

Agradecimientos

Agradezco a mis padres Fabian Rincón Guerreo y Patricia Prieto Monroy quienes han sido mi pilar fundamental y apoyo en mi formación personal y académica y que me han dado todo lo que soy como persona, es decir, mis valores, mis principios, mi empeño y todo ello de una manera desinteresada y llena de amor. Además de todas las personas que me ayudaron de forma directa e indirecta en la construcción de este documento, como mis compañeros del grupo de investigación de procesos estocásticos como Andrés Ríos y Erika Martínez, así como a mi director de Tesis Viswanathan Arunachalam, sin ayuda de todos ustedes este trabajo no habría podido realizarse. Gracias....

Índice general

Resumen ix

Abstract x

Lista de figuras xi

Lista de tablas 1

1. Introducción 2

2. Preliminares 4
2.1. Conceptos básicos estadísticos 4
2.2. Método numérico de resolver ecuaciones diferenciales acopladas 9
2.3. Método de calibración del modelo 11
2.4. Método numérico para ecuaciones diferenciales estocásticas 14

3. Modelo SEIR 16
3.1. Modelo epidemiológico SIR determinista 16
3.2. Modelo epidemiológico SEIR determinista 18
3.2.1. Número reproductivo básico 19
3.3. Modelo epidemiológico SEIR estocástico 21

4. Análisis de datos 23

5. Modelo SEIR modificado 27
5.1. Modelo SEIR modificado sin vacunación 27
5.2. Modelo SEIR modificado con vacunación 35
5.3. Modelo SEIR modificado estocástico con vacunación 40
5.4. Número reproductivo básico del modelo estocástico con vacunación 43

A. Códigos usados para modelar el comportamiento del COVID-19 50
A.1. Tratamiento de datos 50
A.2. Modelo SEIR modificado sin vacunación 53
A.3. Modelo SEIR modificado con vacunación 56
A.4. Modelo SEIR modificado estocastico con vacunación 59

Bibliografía **62**

Modelo SEIR con compartimientos para la propagación del Covid-19 en el Pacífico Nariñense

Resumen

El SARS-COV2 es un virus que se ha propagado por todo el mundo y ha generado una crisis de sanitaria global. El objetivo de este trabajo es proponer un modelo SEIR modificado para describir la dinámica de la propagación del COVID-19 en el Pacífico Nariñense a partir de los datos suministrados por el Instituto Nacional de Salud (INS). Computacionalmente se realizan experimentos para el modelo SEIR modificado con lo que se espera estimar algunos parámetros del modelo y proponer la modelación de posibles escenarios de la propagación del virus.

Palabras clave: (COVID-19, SEIR - modificado, Estimación de parámetros, Pacífico Nariñense, Simulación estocástica).

SEIR model with compartments for propagation of the Covid-19 in the Pacific of Nariño

Abstract

SARS-COV2 is a virus that has spread worldwide and has generated a global health crisis. The objective of this work is to propose a modified SEIR model to describe the dynamics of the spread of COVID-19 in the Pacific of Nariño based on data provided by the National Institute of Health (INS). Computational experiments are performed for the modified SEIR model with which it is expected to estimate some parameters of the model and propose the modeling of possible scenarios of virus spread.

Key words: (COVID-19, SEIR - modified, Parameter estimation, Pacifico Nariñense, Stochastic simulation).

Lista de Figuras

4-1. Datos de infectados por día (a), recuperados por día (b) y fallecidos por día (c) en el Pacífico Nariñense entre el 24 de marzo de 2020 y el 6 de marzo de 2022. 25
4-2. Infectados 26

5-1. Modelo SEIR modificado sin vacunación 28
5-2. Simulaciones sin vacunación variando valores de la probabilidad de tener un resultado falso negativo (f). 34
5-3. Modelo SEIR modificado con vacunación. 35
5-4. Simulaciones con vacunación variando valores de la probabilidad de tener un resultado falso negativo (f). 39
5-5. Simulaciones estocásticas tomando diferentes valores de la probabilidad de tener un resultado falso negativo (f) 43

Lista de Tablas

2-1. Tabla de error de ejemplo de simulación del método de Euler 11

4-1. Estadística descriptiva de los infectados registrados en el INS para el departamento del Nariño 24

5-1. Descripción de parámetros para el modelo SEIR modificado 29
5-2. Descripción de individuos para los modelos SEIR modificados. 30
5-3. Los parámetros δ_1, δ_2, γ_1, γ_2, α_u y α_p son cruciales para modelar la transmisión del SARS-CoV-2 en individuos no testeados, incluyendo falsos negativos. Dado que en Colombia no se ha llevado a cabo una prueba de antígenos masiva para estimarlos, se utilizaron los mismos parámetros del estudio Incorporating false negative tests in epidemiological models for SARS-CoV-2 transmission and reconciling with seroprevalence estimates [Bhattacharyya et al., 2021]. Por otro lado, los parámetros β y r se estiman utilizando el método propuesto en la sección 2.3 y se asumen $\mu = 0$ y $\lambda = 0$. 33
5-4. Tabla de parámetros estimados β y r como su respectivo número reproductivo básico (R_0^S) para diferentes valores de falsos negativos: $f = 0$, $f = 0{,}048$, $f = 0{,}1$ y $f = 0{,}3$ sin vacunación 35
5-5. Tabla de parámetros estimados β y r como su respectivo número reproductivo básico (R_0^C) para diferentes valores de falsos negativos: $f = 0$, $f = 0{,}048$, $f = 0{,}1$ y $f = 0{,}3$ con vacunación 40
5-6. Valores de $R_{0,E}^C$ para el modelo SEIR modificado estocástico 47

1. Introducción

El virus SARS COV-2, causante de la COVID-19, se ha propagado por todo el mundo desde China causando muertes en la mayoría de los países. El número de personas afectadas por el virus aumenta fácilmente debido a que el coronavirus es de fácil transmisión [Rivera-Rodriguez and Urdinola, 2020]. Con el fin de estudiar la propagación del virus, se han propuesto diferentes modelos matemáticos entre ellos modelos epidemiológicos deterministas y también modelos estocásticos. [Rivera-Rodriguez and Urdinola, 2020].

Los *modelos deterministas epidemiológicos* utilizan *ecuaciones diferenciales ordinarias* para modelar la dinámica de la dispersión de un agente infeccioso. Algunos de ellos son los modelos **SIR**, (por sus siglas, Susceptibles, Infectados y Removibles) [Cooper et al., 2020], **SIS** (por sus siglas,Susceptibles e Infectados) y **SEIR** (por sus siglas, susceptibles, expuestos, infectados y recuperados). El primer modelo matemático epidemiológico (SIR) fue desarrollado en 1927 por Kermack y McKendrick [Velasco-Hernandez, 2021]. En este modelo se distinguen tres tipos de individuos: los susceptibles, los infectados y los removibles. Los individuos Susceptibles son aquellos que están propensos a contagiarse de una enfermedad al tener contacto infeccioso con los Infectados, los cuales además pueden presentar o no síntomas de la enfermedad. Los individuos Removibles son aquellos que dejan de ser infectados, en otras palabras, son aquellos individuos que fallecen a causa de la enfermedad, o son aislados por sus síntomas o presentan cierta inmunidad frente a la enfermedad; lo cual se puede deber, a que recibieron un tratamiento médico adecuado o por acción de su sistema inmunológico.

Se ha observado que la enfermedad del COVID-19 en general tiene un largo periodo de incubación, que varia entre 2 a 14 días, en comparación con el virus de la gripa que tiene un periodo entre 1 y 4 días el que los individuos contagiados permanecen asintomáticos [Bhattacharyya et al., 2021]. Esto, en un país como Colombia, con limitaciones económicas y logísticas para realizar pruebas masivas de detección del virus, hace que haya una rápida propagación de la epidemia pues los individuos contagiados que se encuentran en fase de incubación pueden transmitir el virus. En este trabajo se busca desarrollar un modelo que describa la dinámica de la propagación de COVID-19 en el departamento de Nariño, ajustándose adecuadamente a los datos del número de casos reportados de SARS-CoV-2, así como de los recuperados y fallecidos. Esto permitirá obtener estimaciones más precisas y determinar cuándo se espera que el virus deje de ser una amenaza para este departamento.

Puesto que el modelo además de ajustarse a los datos busca considerar tanto la tasa de contacto como la proporción de la población aislada en la ciudad y espera establecer una diferencia entre fallecimientos para la población joven y adulta, se propone desarrollar un modelo estocástico SEIR modificado en el que además de considerar los compartimientos de Susceptibles, Expuestos, Infectados y Recuperados se tendrán en cuenta compartimientos en el grupo de los falsos negativos y asintomáticos. El modelo permitirá describir la distribución de Susceptibles, Infectados, Recuperados, Falsos Negativos e Infectados no Testeados, así como establecer su variabilidad. Finalmente se realizará un ajuste del modelo propuesto con el fin de establecer el número reproductivo básico, que corresponde al número de personas que puede infectar un individuo.

En este trabajo he desarrollado un código utilizando la herramienta de programación de Python que permite modelar la dinámica de propagación del COVID-19 a través un modelo SEIR con prueba diagnóstica, el cual se encuentra en Anexos (A)

Nuestro objetivo es proporcionar a los tomadores de decisiones y a los investigadores, estimaciones de probabilidad para cada escenario considerado, junto con estimaciones del número reproductivo básico, R_0 en los escenarios con vacunación y sin vacunación. La información obtenida a partir del modelo permitirá crear medidas óptimas y eficientes para que con un gasto mínimo por parte de los entes gubernamentales, se creen medidas de choque efectivas frente a la propagación y control de la enfermedad.

2. Preliminares

En este capítulo se presentan algunos conceptos básicos de procesos estocásticos, movimiento browniano, ecuaciones diferenciales estocásticas, los cuales son necesarios e indispensables para comprender los capítulos posteriores.

2.1. Conceptos básicos estadísticos

Definición: [**Blanco, 2019**] El conjunto Ω de todos los posibles resultados de un experimento aleatorio se llama espacio muestral. Los elementos $\omega \in \Omega$ son llamados puntos muestrales.

Definición: [**Blanco, 2019**] Sea $\Omega \neq \emptyset$. Una colección de $\Im$ de subconjuntos de Ω es una σ-álgebra sobre Ω, si y solo si,

- $\Omega \in \Im$
- Si A $\in \Im$ entonces $A^c \in \Im$
- Si A_1,A_2,... $\in \Im$ entonces $\bigcup_{i=1}^{\infty} A_i \in \Im$

Los elementos de $\Im$ se llaman eventos.

Definición: [**Blanco, 2019**] Sean $(\Omega, \Im)$ un espacio medible. Una función P definida sobre $\Im$ y de valor real que satisface las siguientes condiciones:

- $P(A) \geq 0$ para todo $A \in \Im$
- $P(\Omega) = 1$
- Si $A_1, A_2, ...$ son elementos de $\Im$ mutuamente excluyentes, esto es $A_i \cap A_j = \emptyset$ para todo $i \neq j$, entonces $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

se llama medida de probabilidad sobre $(\Omega, \Im)$. La tripla $(\Omega, \Im, P)$ se llama espacio de probabilidad.

Definición: [**Blanco, 2019**] Sea $(\Omega, \Im, P)$ un espacio de probabilidad. Si $A, B \in \Im$ con $P(A) > 0$, entonces se define la probabilidad del evento B bajo la condición A como sigue:

$$P(B|A) := \frac{P(A \cap B)}{P(A)}$$

Teorema: [**Blanco, 2019**] Sea $A_1, A_2, ...$ una partición finita o numerable de Ω, es decir, $A_i \cap A_j = \emptyset$ para todo $i \neq j$, $A_i \in \Im$, $\bigcup_{i=1}^{\infty} A_i = \Omega$ y $P(A_i) > 0$ para todo i. Entonces para cualquier $B \in \Im$ se satisface:

$$P(B) = \sum_n P(B|A_n)P(A_n)$$

Definición: [**Blanco, 2019**] Dos eventos A y B son independientes si y sólo si,

$$P(A \cap B) = P(A)P(B)$$

En caso contrario se dice que los eventos son dependientes.

Con el fin de comprender los cálculos estocásticos que se llevan a acabo en la sección 5.3 se presentan las siguientes definiciones básicas correspondientes a procesos estocásticos:

Definición: [**Blanco, 2019**] Un proceso estocástico es una familia de variables aleatorias $(X_t)_{t \in T}$, definidas sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$ y con valores en un espacio medible (S,R). T se llama conjunto de índices del proceso y S se llama espacio de estados. Los elementos de S se llaman estados.

Definición: [**Blanco, 2019**] Sea $(X_t)_{t \in T}$ un proceso estocástico. Si:

- T es un subconjunto de los números naturales, entonces $(X_t)_{t \in T}$ recibe el nombre de sucesión aleatoria o proceso de parámetro discreto.
- T es un intervalo de los reales, entonces $(X_t)_{t \in T}$ recibe el nombre de proceso de parámetro continuo.
- El conjunto S de estados es discreto, es decir finito o numerable, entonces $(X_t)_{t \in T}$ recibe el nombre de proceso de estados discreto.
- El conjunto S de estados es no numerable, entonces $(X_t)_{t \in T}$ recibe el nombre de proceso de estado continuo.

Definición: [**Karatzas and Shreve, 2012**] Sea $\{X_t\}_{t \in T}$ un proceso estocástico definido sobre un espacio de probabilidad $(\Omega, \Im, P)$ y con valores en $(S, \mathcal{G})$. La función $(t, \omega) \to X_t(w)$ para un $\omega \in \Omega$ se llama la trayectoria de $\{X_t\}_{t \in T}$ asociado a ω.

Definición: [**Kuo, 2006**] Un proceso estocástico $(B_t)_{t\geq 0}$ es llamado un movimiento browniano estándar si satisface las siguientes condiciones:

- $P\{B_0 = 0\} = 1$
- Para cualquier $0 \leq s \leq t$ la variable aleatoria $B(t) - B(s)$ es distribuida normalmente con media 0 y varianza t-s. Esto es, para cualquier $a < b$ se tiene,

$$P\{a \leq B(t) - B(s) \leq b\} = \frac{1}{\sqrt{2\pi(t-s)}} \int_a^b e^{-x^2/2(t-s)} dx$$

- B_t tiene incrementos independientes, es decir, para cualquier $0 \leq t_1 \leq t_2 \leq ... \leq t_n$, las variables aleatorias

$$B(t_1), B(t_2) - B(t_1), ..., B(t_n) - B(t_{n-1})$$

  son independientes

- Las trayectorias son continuas con probabilidad 1

Definición: [**Karatzas and Shreve, 2012**] Sean $\{\Im_t\}_{t\geq 0}$ una colección de σ-álgebras sobre el espacio de probabilidad $(\Omega, \Im, P)$ y $\{X_t\}_{t\geq 0}$ un proceso estocástico. Se dice que $\{\Im_t\}_{t\geq 0}$ es una **filtración** de $\Im$ si para todo $0 \leq s \leq t < +\infty$ se tiene que $\Im_s \subseteq \Im_t \subseteq \Im$.
Se dice que el proceso $\{X_t\}_{t\geq 0}$ es $\Im_t$- **adaptado**, si y solo si, X_t es $\Im_t$-medible para todo $t \geq 0$.

Definición: [**Øksendal, 2003**] Sean $(\Omega, \Im, P)$ un espacio de probabilidad y $\{\Im_t\}_{t\geq 0}$ una filtración de $\Im$ tal que

1. Un movimiento browniano estándar $\{B_t\}_{t\geq 0}$ está definido sobre el espacio de probabilidad $(\Omega, \Im, P)$
2. $\{\Im_t\}_{t\geq 0}$ satisface que $\Im_t = \sigma(\{B_s : 0 \leq s \leq t\})$ para todo $t \geq 0$

La filtración $\{\Im_t\}_{t\geq 0}$ se conoce como la filtración browniana.

Definición: [**Øksendal, 2003**] Sea $X = \{X_t\}_{t\geq 0}$ y $Y = \{Y_t\}_{t\geq 0}$ dos procesos estocásticos sobre el espacio de probabilidad $(\Omega, \Im, P)$. Se define el **producto interno** en $L^2(\Omega, \Im, P)$ como $\langle X, X\rangle_t := E(X_t^2)$ y $< X, Y >_t := E(X_t Y_t)$

Definición: [**Øksendal, 2003**] Sea $\{N_t\}$t ≥ 0, una familia creciente de $\sigma-$ álgebras de subconjunto de Ω. Un proceso $g(t, w) : [0, \infty) \times \Omega \to R^n$ es llamada N_t adaptado si para cada $t \geq 0$ la función

$$\omega \to g(t, \omega)$$

es N_t-medible.

El conjunto de las funciones $\nu(S,T)$ son las funciones integrables en el intervalo [S,T]. Formalmente se definen como

Definición: [Øksendal, 2003] Sea $(\Omega, \Im, (\Im_t)_{t\geq 0}, P)$ un espacio de probabilidad filtrado con $(\Im_t)_{t\geq 0}$, la filtración browniana $\{\Im_t\}_{t\geq 0}$ y sea la clase de las funciones f definidas por

$$f : [0,\infty) \times \Omega \to R$$

$$(t,\omega) \to f(t,\omega)$$

se tiene que $\{f(t)\}_{t\geq 0}$ es un proceso estocástico. Se dice que $f \in \nu(S,T)$ (mas adelante, se llaman funciones integrables en el sentido de Itô) con $0 < S < T$, sí y sólo si, $\{f(t)\}_{t\geq 0}$ cumple las siguientes condiciones:

1. $(t,\omega) \to f(t,\omega)$ es $\mathcal{B}([0,+\infty)) \otimes \Im$-medible, para todo $t \geq 0$ y para todo $\omega \in \Omega$
2. $f(t,w)$ es $\Im_t$ - adaptado
3. $E[\int_S^T f(t,\omega)^2 dt] < \infty$ para todo $\omega \in \Omega$

Definición: [Øksendal, 2003] Sea $(\Omega, \Im, (\Im_t)_{t\geq 0}, P)$ un espacio de probabilidad filtrado con $\{\Im_t\}_{t\geq 0}$ la filtración browniana. Sea una función $\phi \in \nu(S,T)$ es llamada función elemental si y sólo si, es de la forma

$$\phi(t,w) = \sum_j e_j(\omega) I_{[t_j,t_{j+1})}(t)$$

donde e_j son funciones $(t,\omega) \to e_j(t,\omega)$ tales que, e_j es $\Im_{t_j}$- medible para todo $j = 0,1,2,..$ y

$$I_{[t_j,t_{j+1})} = \begin{cases} 1 \text{ si t} \in [t_j,t_{j+1}) \\ 0 \text{ si t} \notin [t_j,t_{j+1}) \end{cases} \quad (2\text{-}1)$$

Toda función $f \in \nu(S,T)$ se puede aproximar por medio de sucesiones de funciones elementales $\{\phi_n\}$.

Se define la integral de Itô en el intervalo $[S,T]$ con $S,T > 0$ para las funciones elementales ϕ como sigue:

Definición: [Øksendal, 2003] Sea ϕ una función elemental. La integral de Itô de la función ϕ en el intervalo $[S,T]$ con $0 < S < T$ se define como

$$\int_S^T \phi(t,\omega)dB(t,\omega) = \sum_{j\geq 0} e_j(w)[B_{tj+1} - B_{t_j}](\omega)$$

De forma mas general para definir la integral de Itô para toda $f \in \nu(S,T)$ se demuestran las siguientes afirmaciones (Pasos):

Paso 1: Sea $g \in \nu(S,T)$ acotada y $g(.,\omega)$ continua para todo $\omega \in \Omega$, entonces existe una sucesión de funciones elementales $\{\phi_n\}_{n\in N} \subseteq \nu(S,T)$ tal que para todo $\omega \in \Omega$

$$E\left[\int_0^T (g(t,\omega) - \phi_n(t,\omega))^2 dt\right] \to 0 \text{ cuando } n \to \infty$$

Prueba: Se define $\phi_n(t,\omega) = \sum_j g(t_j,\omega) I_{[t_j,t_{j+1})}(t)$. Entonces ϕ_n es elemental desde que $g \in \nu(S,T)$, y

$$\int_s^T (g - \phi_n)^2 dt \to 0 \text{ as } n \to \infty, \text{ para cada } \omega$$

Como g es continua para cada ω, se tiene que $E\left[\int_S^T (g-\phi_n)^2 dt\right] \to 0$ cuando $n \to \infty$.

Paso 2: Sea $h \in \nu(S,T)$ acotada, entonces existe una sucesión de funciones acotadas $\{g_n\}_{n\in N} \subseteq \nu(S,T)$ y $g(.,\omega)$ continua para todo $\omega \in \Omega$ y para todo $\omega \in \Omega$

$$E\left[\int_0^T (h(t,\omega) - g_n(t,\omega))^2 dt\right] \to 0 \text{ cuando } n \to \infty$$

Paso 3: Sea $f \in \nu(S,T)$ acotada, entonces existe una sucesión de funciones acotadas $\{h_n\}_{n\in N} \subseteq \nu(S,T)$ tal que para todo $\omega \in \Omega$

$$E\left[\int_0^T (f(t,\omega) - h_n(t,\omega))^2 dt\right] \to 0 \text{ cuando } n \to \infty$$

Definición: [Øksendal, 2003] Sea $f \in \nu(S,T)$. Entonces la integral de Itô de f (desde S hasta T) es definida como

$$\int_S^T f(t,\omega)dB_t(\omega) = \lim_{n\to\infty} \int_S^T \phi_n(t,\omega)dB_t(\omega) \text{ limite en } L^2(P)$$

donde $\{\phi_n\}$ es una secuencia de funciones elementales tales que:

$$E\left[\int_S^T (f(t,\omega) - \phi_n(t,\omega))^2 dt\right] \to 0 \text{ cuando } n \to \infty$$

Teorema: [Øksendal, 2003] (La isometria de Itô) Si $f \in \nu(S,T)$ entonces

$$E\left[\left(\int_S^T f(t,\omega)dB_t\right)^2\right] = E\left[\int_S^T f^2(t,\omega)dt\right]$$

Corolario: [Øksendal, 2003] Si $f(t,\omega) \in \nu(S,T)$ y $f_n(t,\omega) \in \nu(S,T)$ para $n = 1, 2, ...$ y $E\left[\int_S^T (f_n(t,\omega) - f(t,\omega))^2\right] \to 0$ cuando $n \to \infty$ Entonces

$$\int_S^T f_n(t,\omega)dB_t(\omega) \to \int_S^T f(t,\omega)dB_t(\omega) \text{ en } L^2(P) \text{ cuando n } \to \infty$$

A continuación se enuncian algunas propiedades de la integral de Itô:

Teorema: [Øksendal, 2003] Sean $(\Omega, \Im, P)$ un espacio de probabilidad sobre el cual está definida la filtración browniana $\{\Im_t\}_{t\geq 0}$, $f, g \in \nu(0,T)$, $c \in R$ y sea $0 \leq S < U < T$. Entonces

- $\int_S^T f dB_t = \int_S^U f dB_t + \int_U^T f dB_t$ para casi todo $\omega \in \Omega$
- $\int_S^T (cf + g)dB_t = c\int_S^T f dB_t + \int_S^T g dB_t$ para casi todo $\omega \in \Omega$.
- $E\left[\int_S^T f dB_t\right] = 0$
- $\int_S^T f dB_t$ es $\Im_t$ - medible

Definición: [Blanco et al., 2012]: Sea $\{X_n; n > 0\}$ una secuencia de variables aleatorias definida en el espacio de probabilidad $(\Omega, \Im, P)$ y sea $(\Im_n)_{n\geq 0}$ una filtración en $\Im$. Supóngase que $\{X_n; n > 0\}$ es adaptado a la filtración $(\Im_n)_{n\geq 0}$ y $E(X_n)$ existe para toda n. Se dice que $\{X_n; n > 0\}$ es una $(\Im_n)_n$- martingala si y sólo si $E(X_n|\Im_m) = X_m$ casi siempre para todo $m \leq n$

2.2. Método numérico de resolver ecuaciones diferenciales acopladas

Con el fin de estimar el comportamiento de los diferentes grupos propuestos para el modelo SEIR con compartimientos para la propagación del COVID-19 en el Pacífico Nariñense, se resuelven las ecuaciones diferenciales propuestas en las secciones 5.1 y 5.2 utilizando herramientas del lenguaje de programación Python tales como la librería *scipy* y el paquete *integrate.odeint* que utiliza la función *lsoda* proveniente del paquete *odepack* de FORTRAN, el cual está desarrollado para utilizar el método lineal de varios pasos [Hairer et al., 1993], en particular el método de Euler que resuelve ecuaciones diferenciales ordinarias; el cual se

describe a continuación:

Metodo de Euler: [Atkinson et al., 2011]
Se considera un sistema de ecuaciones diferenciales dado por:

$$\frac{dy_i}{dt} = f(t, \mathbf{y}) \ , \ \mathbf{y}(t_0) = \mathbf{y}_0$$

en donde $\mathbf{y}$ representa el vector $\mathbf{y} = (y_1, y_2, ..., y_n)^T$ y $i = 1, ..., n$. Además se tiene que $t_0 < t_1 < ... < t_N$. Por simplicidad se toman $t_n = t_0 + nh$, con $n = 0, 1, ..., N$. El método de Euler utiliza la derivación numérica hacia adelante en donde se realiza la siguiente aproximación

$$\frac{dy_i}{dt} \approx \frac{y_i(t+h) - y_i(t)}{h}$$

Aplicando $t = t_n$ se tiene que $y_i'(t_n) = f(t_n, \mathbf{y}(t_n))$, de donde se obtiene que

$$\frac{y_i(t_{n+1}) - y_i(t_n)}{h} \approx f(t_n, \mathbf{y}(t_n))$$

$$y_i(t_{n+1}) \approx y_i(t_n) + hf(t, \mathbf{y})$$

El método de Euler por lo tanto puede describirse como:

$$y_{i_{n+1}} \approx y_{i_n} + hf(t, \mathbf{y}), \qquad 0 \leq n \leq N-1$$

Ejemplo: [Atkinson et al., 2011]
Considerar la ecuación diferencial $\frac{dy}{dt} = -y, \quad y(0) = 1$.
Se sabe que la solución exacta de esta ecuación es es $y(t) = e^{-t}$. Por otro lado aplicando el método de Euler y tomando $h = 0{,}05$ se tiene que por ejemplo que $t_0 = 0, t_1 = 0{,}05, t_2 = 0{,}1, t_3 = 0{,}15, t_4 = 0{,}2$

$$\begin{aligned}
&y_0 = 1 \\
&y_1 = y_0 + hf(t_0, y_0) = 1 - 0{,}05 \cdot 1 = 0{,}95 \\
&y_2 = y_1 + hf(t_1, y_1) = 0{,}95 - 0{,}05 \cdot 0{,}95 = 0{,}9025 \\
&y_3 = y_2 + hf(t_2, y_2) = 0{,}9025 - 0{,}05 \cdot 0{,}9025 = 0{,}857375 \\
&y_4 = y_3 + hf(t_3, y_3) = 0{,}857375 - 0{,}05 \cdot 0{,}857375 = 0{,}81450625
\end{aligned}$$

Tomando la solución analítica $y(t) = e^{-t}$ se tiene que

$$\begin{aligned}
&y(0{,}00) = 1 \\
&y(0{,}05) = 0{,}9512294 \\
&y(0{,}10) = 0{,}9048374 \\
&y(0{,}15) = 0{,}8607079 \\
&y(0{,}20) = 0{,}8187308
\end{aligned}$$

Se puede calcular el error restando el valor de la solución del método de Euler a la solución analítica de donde se muestra en la tabla **2-1**, en donde se evidencia que medida que h disminuye el error también disminuye.

t_n	Error
t_1	0.0012294
t_2	0.0023374
t_3	0.0033329
t_4	0.0042246

Tabla 2-1.: Tabla de error de ejemplo de simulación del método de Euler

2.3. Método de calibración del modelo

Para encontrar los valores óptimos de los parámetros del modelo SEIR con compartimentos para la propagación de COVID-19 en el Pacífico Nariñense, se utiliza el **método de mínimos cuadrados no lineales**. Para esto, se emplean herramientas de programación en Python, incluyendo la librería *lmfit* y el paquete minimize, que implementan el método no linear least square. Los estimadores basados en el procedimiento de mínimos cuadrados no lineales presentan sesgo, a diferencia de los estimadores basados en mínimos cuadrados lineales.Se puede utilizar una región de tolerancia para reducir el sesgo, la elección de la región de tolerancia puede tener un gran impacto en la precisión y el éxito del ajuste del modelo. En este modelo se utiliza una región de tolerancia definido previamente al ajuste iterativo del modelo a los datos. Durante este proceso, se calculan las derivadas parciales de la función modelo con respecto a los parámetros para encontrar la dirección en la que deben actualizarse los valores de los mismos.

La región de tolerancia es un intervalo de valores posibles para cada parámetro que se considera plausible en función de la información disponible. La elección de la región de tolerancia puede tener un gran impacto en la precisión y el éxito del ajuste del modelo. En general, la región de tolerancia se define en términos de límites inferiores y superiores para cada parámetro. En este este estudio se asigna una región de tolerancia para r, debido a que r es una probabilidad se debe tener que $0 \leq r \leq 1$

La información para realizar este proceso se tomó del Instituto Nacional de Salud siguiendo el tratamiento de datos presentado en la sección 4 de este documento, posteriormente el modelo porpuesto se ajusta a los datos. En este método se emplea una región de tolerancia para el ajuste de los parámetros el cual se describe a continuación:

Sea $\theta = (\theta_1, ..., \theta_m)$ un parámetro multidimensional. Considere $f : R^n \to R$ una función que depende del parámetro θ. Dados unos puntos $(x_1, y_1) \cdots (x_n, y_n)$ se define la suma de cuadrados residual como:

$$g(\theta) = ||y - f(x; \theta)||_2^2 = ||R(\theta)||_2^2$$

donde $g(\theta)$ es la función objetivo, que optimiza minimizando el resultado. Se define la región de tolerancia Δ_k que el intervalo en que un determinado parámetro θ_k le esta permitido estar y según el modelo y es determinado por quienes están implementando el modelo. Por ejemplo, si se tiene que estimar θ_j como el valor de una probabilidad, este valor esta restringido a estar entre 0 y 1.

Luego mediante el método del gradiente se actualizan cada uno de los parámetros teniendo en cuenta que:

$$\theta_{k+1} \to \theta_k + \eta \nabla g(\theta)$$

En donde η es una constante llamada tasa de aprendizaje y siempre teniendo en cuenta que $\theta_k \in \Delta_k$

Se puede elegir η de diferentes maneras. La mas popular consiste en elegir η como una contante muy pequeña aunque esto tener un costo computacional muy alto. Otra forma consiste en utilizar método llamado "line search", que consiste en evaluar $f(x - \eta \nabla f(x))$ muchas veces para diferentes valores de η y escoger el valor de η que disminuye la función objetivo. El algoritmo en general consiste en:

Algoritmo: Procedimiento del método del gradiente
Definir:
θ_0 / Valor de θ inicial
f / Función a minimizar
δ / Valor de convergencia
$t \leftarrow 1$
do
$\quad \theta_{t+1} \leftarrow \theta_t - \eta \nabla f(\theta_t)$
$\quad t \leftarrow t + 1$
while $\|\|\theta_{t+1} - \theta_t\|\| > \delta$
return (θ_t)

Ejemplo del método del gradiente:

Supongamos que se quiere minimizar la función $h(x) = x^2 + 1$, supongamos un valor inicial $x_0 = 5$ y supongamos una tasa de aprendizaje $\eta = 0{,}4$.Primero hallamos la derivada $h'(x) = 2x$, por lo tanto tenemos que

- $x_0 = 5$,
- $\nabla h(x) = 2x = 10$,
- $x_1 = x_0 - \eta \nabla h(x) = 5 - 0{,}4 * 10 = 1$,

Repitiendo 5 veces el procedimiento anterior tenemos que:

- $x_1 = 1{,}0$,
- $x_2 = 0{,}1$,
- $x_3 = 0{,}04$,
- $x_4 = 0{,}008$,
- $x_5 = 0{,}0016$,

Si el computador pudiera iterar infinitas veces llegaríamos al resultado de que $x = 0$, que corresponde al valor de x que alcanza el mínimo de la función, dado que:

$h'(x) = 2x$,
$h'(x_0) = 0$,
$2x_0 = 0$,
$h''(x) = 2$.

Por lo tanto para $x = 0$ hay un mínimo.

Las desviaciones estándar de los parámetros estimados se pueden calcular a partir de la matriz de covarianza inversa. Esta matriz contiene información sobre la curvatura de la función de ajuste en las proximidades del mínimo, y su inversa proporciona información sobre la varianza y covarianza de los parámetros estimados.

El método de diferencia central se utiliza para aproximar la matriz hessiana de una función. Dado un punto x en el espacio de parámetros y una función $f(\theta)$, la matriz hessiana se puede aproximar utilizando las segundas derivadas de la función evaluadas en x.

El método de diferencia central aproxima las segundas derivadas de la función mediante la evaluación de la función en puntos cercanos a x. Para un vector de parámetros $\theta = (\theta_1, \theta_2, ..., \theta_n)$, la matriz hessiana se puede aproximar mediante la fórmula:

$$H_{i,j} = \frac{\partial^2 f}{\partial \theta_i \partial \theta_j} \approx \frac{f(\theta_i + \epsilon_j, \theta_j + \epsilon_j) - f(\theta_i + \epsilon_j, \theta_j - \epsilon_j) - f(\theta_i - \epsilon_j, \theta_j + \epsilon_j) + f(\theta_i - \epsilon_j, \theta_j - \epsilon_j)}{4\epsilon_i \epsilon_j},$$

Donde ϵ_i y ϵ_j son pequeños valores que se utilizan para calcular las diferencias centrales. La matriz hessiana aproximada H tiene dimensiones $n \times n$, donde n es el número de parámetros en el modelo.

Esta aproximación de la matriz hessiana se utiliza en el cálculo de la desviación estándar de los parámetros en el método de mínimos cuadrados no lineales, ya que la inversa de la matriz hessiana se utiliza para calcular la matriz de covarianza de los parámetros.

Matemáticamente, si $\hat{\theta}$ es el vector de parámetros estimados y $H(\hat{\theta})$ es la matriz Hessian evaluada en $\hat{\theta}$, entonces la matriz de covarianza de los parámetros es dada por:

$$Cov(\hat{\theta}) = H(\hat{\theta})^{-1}.$$

Las desviaciones estándar de los parámetros se calculan a partir de los elementos diagonales de esta matriz de covarianza. Si diag(cov($\hat{\theta}$)) denota el vector de elementos diagonales de la matriz de covarianza, entonces las desviaciones estándar de los parámetros son:

$$std(\hat{\theta}) = \sqrt{\text{diag}(\text{cov}(\hat{\theta}))}.$$

2.4. Método numérico para ecuaciones diferenciales estocásticas

Con el fin de realizar los cálculos estocásticos propuestos en este documento, se utiliza la herramienta de programación de python; libreria *sdeint*, que resuelve ecuaciones diferenciales estocásticas numéricamente. Dentro de los métodos numéricos de esta librería se encuentran los algoritmos de Stratonovich-Heun, Runge-Kutta estocástico y el método de Euler-Maruyama. En esta tesis se utilizó este último método para realizar las estimaciones .

El método de Euler-Maruyama, introducido por el matemático japonés Gisiro Maruyama (1916,1986) como una extensión del método de Euler [Allen, 2010], es una técnica de integración numérica para obtener soluciones aproximadas de un sistema de ecuaciones diferenciales estocásticas a partir de un valor inicial dado $X_0 = x_0$. Sea $0 = t_0 < t_1 < ... < t_{k-1} < t_k = T$ una partición del intervalo [0, T], donde la longitud de cada subintervalo es $\Delta t = t_{i+1} - t_i = T/k$, donde se tiene que $t_{i+1} = t_i + \Delta t$ y $\Delta B_i = \Delta B(t_i) = B(t_i + \Delta t) - B(t_i)$. Para cada trayectoria del proceso estocástico, el valor de X_{t_i+1} se aproxima utilizando únicamente el valor del paso de tiempo anterior, X_{t_i}. Entonces, para encontrar las trayectoria del proceso estocástico por el método de Euler-Maruyama se implementa la siguiente ecuación:

$$X_{t_{i+1}} = X_{t_i} + \mu\left(t_i, X_{t_i}\right)\Delta t + \sigma\left(t_i, X_{t_i}\right)\Delta B_i$$

para todo $i = 0, 1, ...k - 1$. Como ΔB_i es un movimiento browniano, éste se distribuye de forma $\Delta B_i \sim N(0, \Delta t)$ La tasa de deriva μ y la tasa de volatilidad σ pueden depender tanto del tiempo t como del valor actual de la variable estocástica X_t. La tasa de deriva μ determina la tendencia o dirección esperada del proceso estocástico en el tiempo, mientras que la tasa de volatilidad σ describe la magnitud esperada de las fluctuaciones aleatorias en el proceso estocástico.[Allen, 2010]. Sea $\sqrt{\Delta t}\eta$ tal que $\eta \sim N(0, 1)$. Entonces tenemos que $\sqrt{\Delta t}\eta \sim N(0, \Delta t)$. Teniendo en cuenta estas consideraciones se describe el algoritmo de Euler Maruyama:

Algoritmo: Método de Euler- Maruyama
Define: $\mu(t, X_t), \sigma(t, X_t)$;
Input: t_0,X_{t_0}, Δt, k;
for $i := 0$ **to** k-1 **do**
$\eta_i \sim N(0, 1)$
$X_{t_{i+1}} = X_{t_i} + \mu(t, X_t)\Delta t + \sigma(t, X_t)\sqrt{\Delta t}\eta_i$
$t_{i+1} = t_i + \Delta t$
print t_{i+1} and $X_{t_{i+1}}$

En el siguiente capítulo, continuaremos explorando el mundo de la epidemiología y profundizaremos en la teoría detrás de los modelos matemáticos SEIR determinista y estocástico, a través de una explicación detallada de los fundamentos teóricos. Al finalizar la lectura del capitulo estaremos en capacidad de entender la teoría y de aplicarla a situaciones reales en capítulos posteriores. Así, estaremos mejor preparados para comprender y abordar la propagación de enfermedades infecciosas en el mundo real.

3. Modelo SEIR

En este capítulo se introducen los modelos matemáticos epidemiológicos SEIR determinista y estocástico. Primero se presentan los modelos SIR que son los que conllevan al modelo SEIR y se define el número reproductivo básico, que se refiere a la cantidad de individuos que pueden ser infectados por una persona que adquirió el virus.

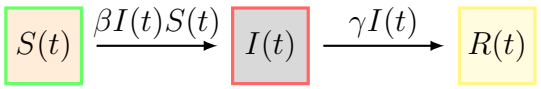
3.1. Modelo epidemiológico SIR determinista

Este modelo utiliza ecuaciones diferenciales ordinarias que fueron descritas originalmente por el bioquímico Kermack (1898, 1970) y el médico militar y epidemiológico McKendrick(1876, 1943) en 1927 [Uribarri et al., 2013], y considera tres tipos de individuos: **Susceptibles**, **Infectados**, y **Removibles**, los cuales son mutuamente excluyentes, sobre una población que es considerada cerrada (población que permanece constante) bajo un número total N. Es decir

$$S(t) + I(t) + R(t) = N, \text{ para todo } t \geq 0 \tag{3-1}$$

en donde $S(t)$, $I(t)$ y $R(t)$ denotan los individuos Susceptibles, Infectados y Removibles respectivamente para un tiempo t. Los **Individuos Susceptibles** son aquellos que están propensos a contagiarse. Los **Individuos Infectados** son los que transmiten el virus y pueden presentar o no sintomatología. Los **Individuos Removibles** son aquellos que logran recuperarse del virus, por lo que desarrollan inmunidad frente a la enfermedad o fallecen, es decir dejan de ser infecciosos.

Uno de los importantes supuestos de este modelo es el flujo de la población: una vez que un individuo susceptible se contagia se convierte en un individuo infectado, el cual a su vez en un tiempo determinado, se convierte en un individuo removible. Una población que cumpla este flujo, en una enfermedad determinada es considerada una **población de tipo SIR**, como se ilustra a continuación.



En este diagrama β representan la **tasa de infectividad** y $1/\gamma$ el **tiempo de remoción**. La **tasa de infectividad** se expresa en términos del número de personas /unidad de tiempo. Entonces si tenemos en el tiempo t, $S(t)$ personas susceptibles y $I(t)$ personas infectadas, se tiene por lo tanto, $\beta I(t)S(t)$ nuevos infectados por unidad de tiempo. Entonces se deduce que el cambio de susceptibles por unidad de tiempo es $(\frac{dS(t)}{dt})$ corresponde a $-\beta I(t)S(t)$ y el cambio respecto a infectados por unidad de tiempo $(\frac{dI(t)}{dt})$ corresponde a los nuevos infectados $\beta I(t)S(t)$. De otro lado el tiempo de remoción $\frac{1}{\gamma}$, corresponde a lo que tardaría, en promedio, una persona infectada a convertirse en removible, por lo que al tener I(t) personas infectadas, se tendrá $\gamma I(t)$ nuevos removibles por unidad de tiempo, entonces se tiene que el cambio de infectados por unidad de tiempo $(\frac{dI(t)}{dt} = \beta I(t)S(t) - \gamma I(t)$ y el cambio removibles por unidad de tiempo $(\frac{dR(t)}{dt})$ es $\gamma I(t)$.

El flujo de individuos entre los diferentes comportamientos queda descrito por el siguiente sistema de ecuaciones diferenciales

$$\begin{aligned} \frac{dS(t)}{dt} &= -\beta I(t)S(t), \\ \frac{dI(t)}{dt} &= \beta I(t)S(t) - \gamma I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t) \end{aligned} \tag{3-2}$$

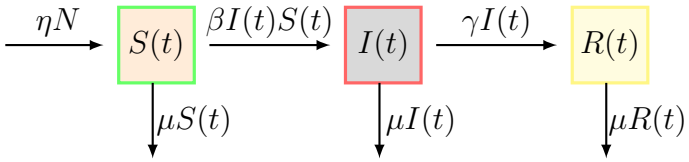
Este modelo SIR no diferencia a los individuos recuperados de los individuos fallecidos.

Con el fin de considerar esta diferencia existe el modelo SIR con demografía. En éste [Bastin, 2018], se toma en cuenta la tasa de influencia, que es, el número promedio de nuevos susceptibles por unidad de tiempo (inmigración de susceptibles), por lo general se representa como ηN , donde N es el número inicial de susceptibles y η es una constante por unidad de tiempo llamada tasa de inmigración [Bastin, 2018]. Se debe tener en cuenta que $1/\gamma$ ya no es el tiempo de remoción, sino tiempo de recuperación.

Se considera además la tasa de emigración μ, en la cual se considera únicamente los individuos removidos. Por ejemplo si fallecen 50 individuos se tiene que $\mu = 50$ fallecidos /año.A este modelo también se le conoce como el modelo SIR con nacimientos y muertes y su sistema de ecuaciones está descrito como se muestra a continuación:

$$\begin{aligned} \frac{dS(t)}{dt} &= \eta N - \beta I(t)S(t) - \mu S(t), \\ \frac{dI(t)}{dt} &= \beta I(t)S(t) - \gamma I(t) - \mu I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t) - \mu R(t). \end{aligned} \tag{3-3}$$

Y el diagrama de flujo viene dado por:



Si se tiene que $\mu = \eta$ se tiene la ecuacion 3-1, en caso contrario N dependería de t es decir $N(t)$, por lo que se tiene:

$$S(t) + I(t) + R(t) = N(t), \text{ para todo } t \geq 0 \tag{3-4}$$

3.2. Modelo epidemiológico SEIR determinista

El modelo epidemiológico SEIR es una adaptación del modelo SIR, en el que un individuo puede asumir de manera excluyente uno de los siguientes estados: Susceptible, Expuesto, Infectado y Removido, se asume una población invariable, $N = S + E + I + R$ [Culqui Sánchez et al., 2020]. En general los individuos Susceptibles, que son infectados se convierten en individuos Expuestos y después del periodo de incubación se transforman en Infectados y adquieren la capacidad de infectar a otros individuos. Los individuos Removidos, son los infectados que han sido inmunizados o fallecieron a causa de la enfermedad.

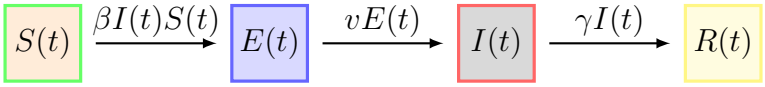
Específicamente la población susceptible es el número de individuos que pueden ser contagiados por el virus en estudio, se debe considerar que algunas enfermedades requieren primero de un periodo de incubación durante el cual la enfermedad está presente en el organismo, pero no es contagiosa para los demás, hasta que los primeros síntomas son apreciables, lo que sucede tras un periodo de incubación (estado latente) [Culqui Sánchez et al., 2020]. La población expuesta está conformada por los individuos que han mantenido contacto con infectados (cercos epidemiológicos), pero que todavía no han desarrollado la enfermedad y todavía no son contagiosos.

En el modelo SEIR se tiene un tiempo promedio de incubación $1/v$, este es el tiempo que trascurre desde el momento en el que el individuo fue infectado y la aparición de los síntomas, representa la velocidad con que los individuos pasan del estado latente al infectado.

Entonces $vE(t)$ representa el número de personas expuestas por unidad de tiempo que pueden transmitir el virus, puesto que ha pasado el tiempo $1/v$ suficiente para la incubación del agente infeccioso. Entonces de manera análoga al modelo SIR, el sistema de ecuaciones que se tiene para el modelo SEIR es:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\beta I(t)S(t), \\
\frac{dE(t)}{dt} &= \beta I(t)S(t) - vE(t), \\
\frac{dI(t)}{dt} &= vE(t) - \gamma I(t), \\
\frac{dR(t)}{dt} &= \gamma I(t).
\end{aligned} \tag{3-5}$$

El diagrama de flujo de este modelo viene dado por:



3.2.1. Número reproductivo básico

El número reproductivo básico (R_0) representa el número de nuevos infectados producidos por un solo infectado, si toda la población es susceptible y se utiliza para estimar la velocidad con la que una enfermedad puede propagarse en un población, por lo tanto es una información importante en la epidemiología de enfermedades infecciosas [Ridenhour et al., 2018].

Para calcular el número reproductivo básico R_0 se utilizara la matriz denominada "next-generation matrix (NGM)"por su nombre en inglés. Dicha matriz por lo general se denota como $\mathbf{K}$ y fue introducida por Odo Diekmann [Diekmann et al., 2010]. A continuación se describe la construcción de NGM para cualquier sistema de ecuaciones diferenciales ordinarias.

Para calcular R_0 se parte se toma el sistema de ecuaciones diferenciales ordinarias que describen la producción de nuevas infecciones y cambios de estado entre los individuos infectados, este conjunto ecuaciones diferenciales se conocen con el nombre de subsistema infectado. El primer paso es linealizar el subsistema infectado de ecuaciones diferenciales ordinarias. Epidemiológicamente, la linealización refleja que R_0 caracteriza el potencial de propagación inicial de un agente infeccioso cuando se introduce en una población completamente susceptible, y que suponemos que el cambio en la población susceptible es insignificante durante la propagación inicial. Este subsistema infectado linealizado es el punto de partida de nuestros cálculos.

Cualquier sistema de ecuaciones diferenciales lineales, puede ser descrita por una matriz Jacobiana.La estructura de esta matriz se puede relacionar con la interpretación epidemiológica. En particular, se puede determinar R_0 descomponiendo primero la matriz como $T+\Sigma$, donde

T es la parte de transmisión, que describe la producción de nuevas infecciones, y Σ es la parte de transición, que describe los cambios de estado (incluida la eliminación por muerte). o la adquisición de inmunidad). Luego se calcula el eigenvalor dominante o más precisamente el radio espectral ρ de la matriz -$\mathbf{T}\Sigma^{-1}$ [Diekmann et al., 2010]

Ejemplo: [Diekmann et al., 2010] Si se considera un sistema con los siguientes estados: S (Susceptibles); E_1 (Expuestos categoría 1); E_2 (Expuestos categoría 2); I (Infectados) y R (Recuperados/removidos/inmunes). La idea de tener dos comportamientos de expuestos que representan dos categorías, en las cuales cada grupo de expuestos se infectan del virus a diferentes velocidades. Sea β la tasa de transmisión del virus, μ la tasa de nacimientos y muertes, v_1 y v_2 las tasas de salida del estado de expuesto, γ la tasa a la que se deja el estado infeccioso, se supone que hay una proporción fija de las dos categorías en la población, $\rho : \rho$ -1 . Por lo que se tiene el siguiente sistema de ecuaciones diferenciales

$$\begin{aligned}
\frac{dS}{dt} &= \mu N - \beta\frac{SI}{N} - \mu S \\
\frac{dE_1}{dt} &= \rho\beta\frac{SI}{N} - (v_1+\mu)E_1, \\
\frac{dE_2}{dt} &= (1-\rho)\beta\frac{SI}{N} - (v_2+\mu)E_2, \\
\frac{dI}{dt} &= v_1E_1 + v_2E_2 - (\gamma+\mu)I, \\
\frac{dR}{dt} &= \gamma I - \mu R.
\end{aligned} \tag{3-6}$$

con $N = S + E_1 + E_2 + I + R$. El sistema tiene tres estados infecciosos E_1, E_2 y I, y dos estados no infectados S y R. En el estado de la preinfección se tiene que $E_1 = E_2 = I = R = 0$ de donde se deduce que $S = N$. Tomando esta ecuación y reemplazando en el sistema de ecuaciones diferenciales se tiene, tomando solo las ecuaciones correspondientes a la transmisión del virus, que:

$$\begin{aligned}
\frac{dE_1}{dt} &= \rho\beta I - (v_1+\mu)E_1, \\
\frac{dE_2}{dt} &= (1-\rho)\beta I - (v_2+\mu)E_2, \\
\frac{dI}{dt} &= v_1E_1 + v_2E_2 - (\gamma+\mu)I.
\end{aligned} \tag{3-7}$$

Si tomamos $\mathbf{x} = (E_1, E_2, I)'$ donde ' denota transpuesta y escribimos el sistema de ecuaciones diferenciales tenemos que

$$\dot{\mathbf{x}} = (\mathbf{T} + \Sigma)\mathbf{x},$$

La matriz $\mathbf{T}$ corresponde a las componentes transmisores y la matriz Σ a las transitoras. Linealizando el sistema de ecuaciones diferenciales tenemos que

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & \rho\beta \\ 0 & 0 & (1-\rho)\beta \\ 0 & 0 & 0 \end{pmatrix},$$

y además

$$\Sigma = \begin{pmatrix} -(v_1+\mu) & 0 & 0 \\ 0 & -(v_2+\mu) & 0 \\ v_1 & v_2 & -(\gamma+\mu) \end{pmatrix},$$

Entonces tenemos que

$$\begin{aligned} \mathbf{K} &= -\mathbf{T}\Sigma^{-1} \\ &= \begin{pmatrix} 0 & 0 & \rho\beta \\ 0 & 0 & (1-\rho)\beta \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{v_1+\mu} & 0 & 0 \\ 0 & \frac{1}{v_2+\mu} & 0 \\ \frac{v_1}{(v_1+\mu)(\gamma+\mu)} & \frac{v_2}{(v_2+\mu)(\gamma+\mu)} & \frac{1}{\gamma+\mu} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\rho\beta v_1}{(v_1+\mu)(\gamma+\mu)} & \frac{\rho\beta v_2}{(v_2+\mu)(\gamma+\mu)} & \frac{\rho\beta}{\gamma+\mu} \\ \frac{(1-\rho)\beta v_1}{(v_1+\mu)(\gamma+\mu)} & \frac{(1-\rho)\beta v_2}{(v_2+\mu)(\gamma+\mu)} & \frac{(1-\rho)\beta}{\gamma+\mu} \\ 0 & 0 & 0 \end{pmatrix}, \end{aligned} \tag{3-8}$$

Ahora se halla R_0 hallando el eigenvalor dominantes que puede ser obtenido mediante la fórmula

$$\begin{aligned} R_0 &= \rho(K), \\ R_0 &= \frac{1}{2}\left(trace(\mathbf{K}) + \sqrt{trace(\mathbf{K})^2 - 4det(\mathbf{K})}\right), \end{aligned} \tag{3-9}$$

Lo cual es equivalente a

$$R_0 = \left(\frac{\rho v_1}{v_1+\mu} + \frac{(1-\rho)v_2}{v_2+\mu}\right)\frac{\beta}{\gamma+\mu}. \tag{3-10}$$

El número reproductivo básico también puede ser calculado por el método de la integral que se describe a continuación:

3.3. Modelo epidemiológico SEIR estocástico

A partir de los modelos epidemiológicos deterministas se construyen los modelos epidemiológicos estocásticos. Recordando que β es la tasa de infectividad, entonces si tomamos dt tenemos que βdt denota el número de personas infectadas durante el periodo de tiempo dt [Ríos Gutiérrez, 2019]. Si asumimos β como la variable aleatoria real que toma valor constante β, entonces tomando $dt \to 0$ se tiene que $\beta dt \to 0$, así,

$$E(\beta dt) = \beta dt \text{ y } V(\beta dt) = 0$$

Ahora si bajo la hipótesis de el número de personas infectadas durante el periodo dt bajo perturbaciones (que por ejemplo pueden ser climáticas), que se denota como $\tilde{\beta}dt$ y se puede escribir como:

$$\tilde{\beta}dt = \beta + \sigma B(t)$$

donde β es la constante positiva que denota la tasa de infectividad y σ la constante que determina cuánto varia la población acorde a $B(t)$ que es el movimiento browniano. El diferencial puede ser denotado como:

$$d(\beta + \sigma B(t)) = \beta dt + \sigma dB(t)$$

En donde:

$$B(T) - B(t) \sim N(0, T - t)$$

Entonces $dB(t) \sim N(0, dt)$ (tomando $dt = [t, T)$). Entonces cuando $dt \to 0$ entonces

$$E(\beta dt + \sigma dB(t)) = \beta dt \to 0 \text{ y } V(\beta dt + \sigma dB(t)) = \sigma^2 dt \to 0$$

En la ecuación anterior β corresponde a la componente determinista y σ la constante de componente estocástica. Reemplazando estos resultados en la ecuación SEIR determinista tenemos que:

$$\begin{aligned} dS(t) &= (\eta N - \beta I(t)S(t) - \mu S(t))dt - \sigma I(t)S(t)dB(t) \\ dE(t) &= (\beta I(t)S(t) - vE(t) - \mu E(t))dt + \sigma I(t)S(t)dB(t) \\ dI(t) &= (vE(t) - \gamma I(t) - \mu I(t))dt \\ dR(t) &= (\gamma I(t) - \mu R(t))dt \end{aligned} \tag{3-11}$$

En el siguiente capítulo se describe la fuente de datos utilizada en este estudio, que es el Instituto Nacional de Salud de Colombia (INS). Se explica el objetivo de esta institución y la información que se encuentra en su base de datos de casos positivos de COVID-19. A continuación, se describe cómo se filtran y seleccionan los datos correspondientes al departamento de Nariño para su análisis.

4. Análisis de datos

Los datos utilizados para realizar esta tesis fueron tomados del Instituto Nacional de Salud de Colombia (INS) [INS, 2022a]. Esta institución es adscrita al ministerio de Salud de Colombia y tiene como objetivo el proponer políticas y normas, promover, desarrollar y difundir la investigación científica-tecnológica y brindar servicios de salud en los campos de salud pública, control de enfermedades transmisibles y no transmisibles, alimentación y nutrición, producción de biológicos, control de calidad de alimentos, productos farmacéuticos y afines, salud ocupacional y protección del medio ambiente [INS, 2022b].

La base de datos utilizada en este estudio puede ser encontrada en los datos abiertos del INS de los casos positivos de COVID-19. Esta base de datos contiene información relacionada con fechas de diagnostico, recuperación y fallecimiento de cada ciudadano Colombiano que fue reportado como positivo, así como de la edad de cada individuo, su sexo, su cédula de ciudadanía y el departamento en donde se encuentra, entre otras [Datos Abiertos, 2022]. Con base en estos datos se filtran y seleccionan solo aquellos ciudadanos correspondientes al departamento de Nariño para hacer el análisis propuesto en esta tesis.

Uno de los principales factores que influyen en el calculo del R_0 es el comportamiento humano. Las medidas de distanciamiento social, el uso de mascarillas y el lavado de manos son esenciales para reducir la propagación de enfermedades. Sin embargo, las diferentes culturas y hábitos de diferentes países pueden influir en la aceptación y cumplimiento de estas medidas. Por ejemplo, en algunos países, como Japón, el uso de mascarillas y el distanciamiento social son prácticas comunes en la vida cotidiana, lo que puede reducir el R_0 en caso de un brote. Por otro lado, en otros países donde las normas culturales no favorecen el distanciamiento social, el R_0 puede ser más alto por lo que cada región particular del mundo puede tener una dinámica epidemiológica diferencial.

Con base en estos datos encontrados se calcula el número de de infectados, recuperados y fallecidos por día del COVID-19 en el departamento del Nariño entre el 24 de marzo de 2020 y el 6 de marzo de 2022 utilizando la librería **pandas** del lenguaje de programación de Python (**Anexo A.1**) y se grafica (ver figura **4-1**). Con estos datos, se calcula la frecuencia acumulada de infectados, recuperados y fallecidos para así calcular la cantidad de infectados reales en un punto de tiempo. Este valor se obtiene restando a la frecuencia acumulada de infectados la frecuencia acumulada de recuperados y de fallecidos como se muestra a continuación:

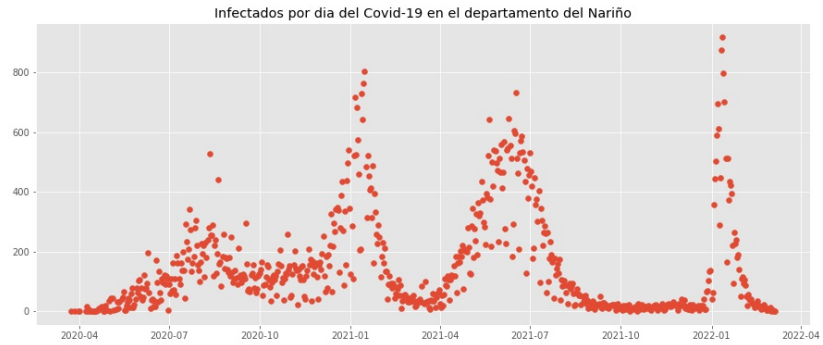
$$\begin{aligned}\text{Infectados} =&\text{Frecuencia acumulada de infectados} - \text{Frecuencia acumulada de recuperados} \\ &- \text{Frecuencia acumulada de fallecidos}\end{aligned}$$

Se calcula el número de Infectados y se grafica (ver figura **4-2**), en donde se observan diferentes picos a lo largo de los años 2020 y 2021. El primero a mediados de agosto del 2020, el segundo a mediados de enero de 2021, el tercero a finales de junio de 2021 y un cuarto pico a mediados de enero de 2022.

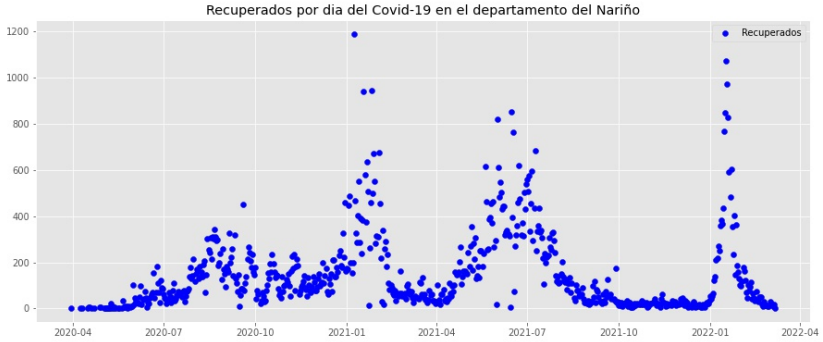
La información obtenida sobre el número de infectados es en la que nos basamos para estimar algunos parámetros del modelo SEIR modificado descrito en la sección **5**. Adicionalmente con base en los datos obtenidos se procede a realizar estadísticas descriptivas tal y como se muestran en la tabla **4-1**. El promedio de nuevos infectados es de 146.14 por día, y el acumulado de infectados es de 50677.98 hasta la fecha. El promedio de nuevos recuperados es de 140.71 por día, y el acumulado de recuperados es de 46567.55. El promedio de nuevos muertos es de 5.2945 por día, y el acumulado de muertos es de 1865.36. La desviación estándar indica la variabilidad de los datos alrededor del promedio. En este caso, las desviaciones estándar, indican una gran variabilidad en los datos.

	Nuevos infectados	Acumulados Infectados	Nuevos recuperados	Acumulados recuperados	Nuevos muertos	Acumulados muertos
Promedio	146.14	50677.98	140.71	46567.55	5.2945	1865.36
Std	166.91	1316.87	170.23	34601.99	5.47	1316.87
Minimo	0	1	0	0	0	0
25 %	22	17464	26	13526	1	700
50 %	91	49836	77	46265	4	1862
75 %	194	89009	185	84410	8	3346
Maximo	920	104200	1189	100327	32	3775

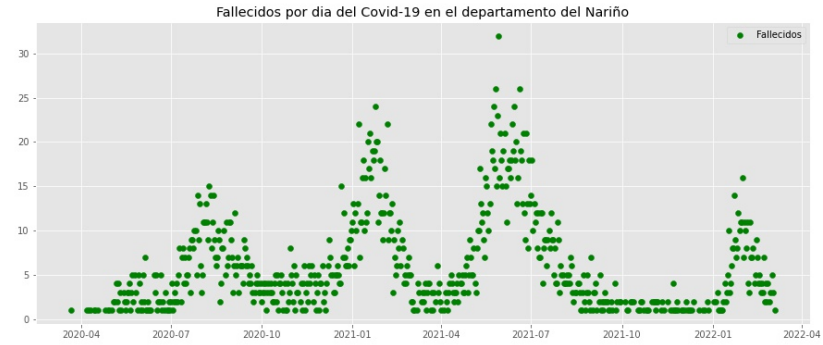
Tabla 4-1.: Estadística descriptiva de los infectados registrados en el INS para el departamento del Nariño



(a) Diagnoticados positivos de Covid-19 por día en el departamento de Nariño



(b) Recuperados por día del Covid-19 en el departamento de Nariño



(c) Fallecidos por día del Covid-19 en el departamento de Nariño

Figura 4-1.: Datos de infectados por día (a), recuperados por día (b) y fallecidos por día (c) en el Pacífico Nariñense entre el 24 de marzo de 2020 y el 6 de marzo de 2022.

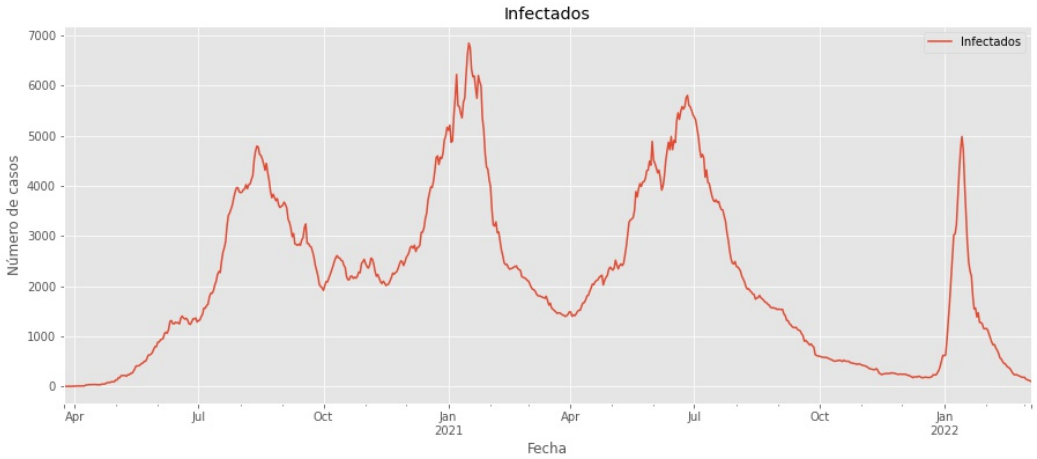


Figura 4-2.: Infectados

En el siguiente capitulo se proponen tres modelos SEIR modificados. El primer modelo se presenta como un modelo determinístico que incorpora diferentes comportamientos de individuos infectados, tales como aquellos que no han sido sometidos a pruebas diagnósticas, falsos negativos y aquellos individuos que han sido sometidos a pruebas. El segundo modelo, también determinístico, integra los mismos compartimentos que el primer modelo, pero añade uno adicional para incluir el proceso de vacunación. Finalmente, el tercer modelo es de naturaleza estocástica, y sus compartimentos son idénticos a los del segundo modelo determinístico.

5. Modelo SEIR modificado

El modelo SEIR puede describir con precisión la dinámica de infección de los virus en poblaciones que siempre presentan síntomas al contraer un patógeno y en donde se realice un número adecuado de pruebas diagnósticas. Sin embargo, muchas personas que contraen la enfermedad provocada por el virus del COVID-19, pueden o no presentar síntomas y transmitir la enfermedad, por lo que muchas personas infectadas pueden estarlo sin saberlo, pues al no desarrollar síntomas, no se realizan una prueba diagnóstica [Woloshin et al., 2020]. A su vez las pruebas diagnósticas que se realizan a las personas que deciden realizárselas no tienen un sensibilidad igual a 1 por lo que pueden existir pruebas con resultado falso negativo, por lo tanto se propone que para modelar la propagación del COVID-19, se incluyan compartimentos que incluyan los individuos no testados que están infectados y los individuos con resultado de prueba diagnóstica de falso negativo.

Con el fin de proponer un modelo SEIR modificado, se tiene en cuenta un modelo sin vacunación y otro con vacunación. El primero teniendo en cuenta que los datos obtenidos para este estudio están desde el día 24 de marzo de 2020 (día en el que se reportó el primer caso de COVID-19 en el Pacifico Nariñense) cuando aún no existía vacuna y el 1 de julio de 2020. Para el modelo con vacunación se tendrán en cuenta los datos obtenidos , desde abril de 2021, mes en el cual se avanzaba en la vacunación en Colombia, cabe aclara que para el mes de Febrero de 2021 se inicio con la vacunación.

5.1. Modelo SEIR modificado sin vacunación

Con el fin de estudiar la dinámica de propagación del virus antes y después de la vacunación es importante proponer el modelo SEIR modificado sin vacunación, en este además se tiene en cuenta la presencia de individuos infectados no detectados y aquellos que dan falsos negativos en la prueba diagnóstica del COVID-19.

Siguiendo el modelo propuesto por [Bhattacharyya et al., 2021] se propone un modelo SEIR modificado en el que se tiene en cuenta que hay nueve tipos de individuos: susceptibles (S), expuestos (E), con resultado positivo en la prueba de detección de SARS-CoV-2 (P), no testeados que tienen SARS-CoV-2 (U), que dan falsos negativos en la prueba de detección de SARS-CoV-2 (F), recuperados y no testeados o con resultado falso negativo (A), fallecidos y no testeados o con resultado falso negativo (B), con resultado positivo y posteriormente

recuperados (L), con resultado positivo y que son reportados fallecidos a causa del virus (D). Los supuestos del modelo SEIR modificado sin vacunación son:

- Homogeneidad de la población: El modelo asume que todos los individuos en la población tienen el mismo riesgo de infección y que interactúan de manera homogénea. Este supuesto puede no ser válido en poblaciones donde hay diferentes grupos de edad, género, niveles socioeconómicos, etc.
- Contactos aleatorios: El modelo asume que los contactos entre individuos son aleatorios y que la probabilidad de infección es la misma para todos los contactos. En la realidad, las personas pueden tener diferentes niveles de exposición al virus dependiendo de su trabajo, estilo de vida y otros factores.
- Tiempo de incubación fijo: El modelo asume que el tiempo de incubación de la enfermedad es constante y conocido. En la realidad, el tiempo de incubación puede variar entre individuos y puede no ser conocido con precisión.
- Inmunidad permanente: El modelo asume que una vez que un individuo se recupera de la enfermedad, se vuelve inmune de por vida y no puede volver a infectarse. En la realidad, la duración de la inmunidad puede variar y en algunos casos puede no ser permanente.
- Ausencia de vacunas o tratamientos: El modelo no considera la existencia de vacunas o tratamientos efectivos para la enfermedad. En la realidad, la disponibilidad de una vacuna o tratamiento puede cambiar significativamente la propagación de la enfermedad.

El modelo SEIR modificado propuesto para analizar los falsos negativos y no testeados que tienen el virus se muestran en la figura **5-1**. Los parámetros expuestos en esta figura, junto con algunos factores de escala se encuentran descritos en la tabla **5-1**:

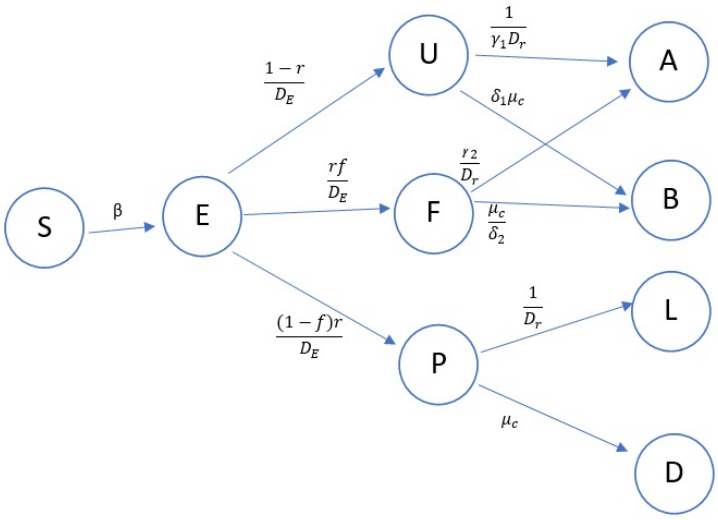


Figura 5-1.: Modelo SEIR modificado sin vacunación

Parámetros	Descripción
D_E	Periodo de incubación promedio (en días).
D_r	Número promedio de días que pasan hasta la recuperación de las personas infectadas
D_t	Número promedio de días para que se obtenga el resultado de la prueba después de que se realice la prueba a una persona.
β	Tasa de transmisión del virus.
μ_c	Tasa de mortalidad atribuible a COVID-19, la cual es equivalente al inverso del número promedio de días para la muerte a partir del inicio de la enfermedad, multiplicado por la probabilidad de muerte de un individuo infectado.
γ_1, γ_2	Factores de escala para la tasa de recuperación de individuos no detectados y falsos negativos, respectivamente.
δ_1, δ_2	Factores de escala para la tasa de mortalidad de individuos no detectados y falsos negativos, respectivamente.
r	Probabilidad de hacerse la prueba para detectar individuos infecciosos.
f	Probabilidad de un resultado de prueba de diagnóstico de RT-PCR falso negativo.
α_u	Factor de escala para la tasa de propagación de la infección por individuos no sometidos a prueba. Se supone que α_u es menor que 1 ya que la mayoría de U consiste en casos asintomáticos o con síntomas leves que se sabe que propagan la enfermedad a una tasa mucho menor.
α_p	La tasa de propagación de la infección entre pacientes con resultado positivo en la prueba, en comparación con los resultados falsos negativos.

Tabla 5-1.: Descripción de parámetros para el modelo SEIR modificado

Mediante la modificación del modelo SEIR se busca analizar los parámetros de acuerdo con la información que se tiene para realizar la modelación de la epidemia de acuerdo a la dinámica de la población. El tiempo está determinado por días ya que la información se encuentra dada de esta manera. Para la estimación de los parámetros se utilizan los boletines del Instituto Nacional de Salud (INS) donde se reportan los nuevos casos del COVID-19 en Colombia, así como también el número de recuperados y pacientes fallecidos.

En el modelo SEIR modificado sin vacuna, el individuo susceptible (S) se convierte en individuo expuesto (E), con tasa de transmisión β. A su vez el individuo expuesto (E), pasa a ser un individuo infectado sin nunca hacerse la prueba (U), con una tasa de transmisión de $\frac{1-r}{D_E}$, donde r representa la probabilidad de hacerse la prueba para detectar el virus y D_E el periodo de incubación en días. El individuo expuesto (E) puede hacer la prueba y dar un falso negativo (F) con una tasa de transmisión $\frac{rf}{D_E}$, donde f representa la probabilidad de un resultado de prueba de diagnóstico de RT-PCR falso negativo. A su vez el individuo infectado no testeado (U) pueden recuperarse (A) o fallecer (B) con tasas de transmisión de $\frac{1}{\gamma_1 D_r}$ y $\delta_1 \mu_c$ respectivamente. Mientras que el individuo con resultado falso negativo (F) también puede recuperarse (A) o fallecer (B) con tasas de transmisión de $\frac{\gamma_2}{D_r}$ y $\frac{\mu_c}{\delta_2}$. Por otro

lado los individuos expuestos (E) pueden hacerse la prueba y dar positivo (P) con una tasa de transmisión de $\frac{(1-f)r}{D_E}$. Estos a su vez pueden recuperarse (L) o fallecer (D) con tasas de transmisión de $\frac{1}{D_r}$ y μ_c.

La descripción de los compartimientos de este modelo se encuentran en la tabla **5-2**

Variable	**Descripción**
S	Individuo susceptible
E	Individuo expuesto
U	Individuo sin realizar prueba diagnóstica
F	Individuo testeado como falso negativo
P	Individuo testeado como positivo
A	Recuperados y no testeados
B	Muertos y no testeados
L	Recuperado reportado como positivo
D	Muerto reportado como positivo

Tabla 5-2.: Descripción de individuos para los modelos SEIR modificados.

Las ecuaciones que describen el modelo SEIR modificado determinista planteado se definen como:

$$\begin{aligned}
\frac{dS}{dt} &= \lambda - \beta(\alpha_u U + F + \alpha_p P)\frac{S}{N} - \mu S,\\
\frac{dE}{dt} &= \beta\frac{S}{N}(\alpha_u U + F + \alpha_p P) - \frac{E}{D_E} - \mu E,\\
\frac{dU}{dt} &= \frac{(1-r)E}{D_E} - \frac{U}{\gamma_1 D_r} - \delta_1\mu_c U - \mu U,\\
\frac{dF}{dt} &= \frac{rfE}{D_E} - \frac{\gamma_2}{D_r}F - \frac{\mu_c}{\delta_2}F - \mu F,\\
\frac{dP}{dt} &= \frac{(1-f)r}{D_E}E - \frac{P}{D_r} - \mu_c P - \mu P,\\
\frac{dL}{dt} &= \frac{P}{D_r} - \mu L,\\
\frac{dD}{dt} &= \mu_c P,\\
\frac{dA}{dt} &= \frac{U}{\gamma_1 D_r} + \frac{\gamma_2 F}{D_r} - \mu A,\\
\frac{dB}{dt} &= \delta_1\mu_c U + \frac{\mu_c F}{\delta_2}.
\end{aligned} \tag{5-1}$$

Con el fin de calcular el número reproductivo básico (R_0). Se siguen los pasos de la sección

(3,2,1). Se escribe de forma matricial los estados que pueden transmitir el virus, por lo que se tiene:

$$\begin{pmatrix} \frac{dE}{dt} \\ \frac{dU}{dt} \\ \frac{dF}{dt} \\ \frac{dP}{dt} \end{pmatrix} = \begin{pmatrix} -\left(\frac{1}{D_E}+\mu\right) & \beta\alpha_u & \beta & \beta\alpha_p \\ \frac{1-r}{D_E} & -\left(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+\mu\right) & 0 & 0 \\ \frac{rf}{D_E} & 0 & -\left(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu\right) & 0 \\ \frac{(1-f)r}{D_E} & 0 & 0 & -\left(\frac{1}{D_r}+\mu_c+\mu\right) \end{pmatrix} \cdot \begin{pmatrix} E \\ U \\ F \\ P \end{pmatrix}.$$

Donde se tiene la matriz transmisora Σ:

$$\Sigma = \begin{pmatrix} -\left(\frac{1}{D_E}+\mu\right) & 0 & 0 & 0 \\ \frac{1-r}{D_E} & -\left(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+\mu\right) & 0 & 0 \\ \frac{rf}{D_E} & 0 & -\left(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu\right) & 0 \\ \frac{(1-f)r}{D_E} & 0 & 0 & -\left(\frac{1}{D_r}+\mu_c+\mu\right) \end{pmatrix},$$

Y la matriz transmisora $\mathbf{T}$:

$$\mathbf{T} = \begin{pmatrix} 0 & \beta\alpha_u & \beta & \beta\alpha_p \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\Sigma^{-1} = \begin{pmatrix} \frac{D_E}{D_E\mu+1} & 0 & 0 & 0 \\ -\frac{r-1}{(D_E\mu+1)(\delta_1+\mu_c+\mu+\frac{1}{D_r\delta_1})} & \frac{1}{\left(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+\mu\right)} & 0 & 0 \\ -\frac{rf}{(D_E\mu+1)\left(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu\right)} & 0 & \frac{1}{\left(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu\right)} & 0 \\ -\frac{(1-f)r}{(D_E\mu+1)(\mu_c+\mu+\frac{1}{D_r})} & 0 & 0 & \frac{1}{\left(\frac{1}{D_r}+\mu_c+\mu\right)} \end{pmatrix},$$

Se calculan los valores propios siguiendo los pasos de la sección (3.2.1) de donde se puede hallar el número reproductivo básico:

$$R_0^S = \frac{1}{2}\left(trace(\text{-}\mathbf{T}\Sigma) + \sqrt{trace(\text{-}\mathbf{T}\Sigma)^2 - 4det(\text{-}\mathbf{T}\Sigma)}\right),$$

Entonces

$$R_0^S = \frac{\beta}{\mu D_e + 1}\left(\frac{\alpha_u(1-r)}{\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+\mu} + \frac{\alpha_p r(1-f)}{\frac{1}{D_r}+\mu_c+\mu} + \frac{rf}{\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu}\right). \tag{5-2}$$

Con el fin de desarrollar los cálculos propuestos en el modelo se utilizan los datos disponibles en el Instituto Nacional de Salud. En esta base de datos se encuentra la información acerca de las personas infectadas día a día, el número de muertes y se analiza la información en el lenguaje de programación de Python.

La tasa de recuperación $1/Dr$ y muerte μ_c se calculan con base en los datos suministrados por el INS. Dr corresponde al tiempo de recuperación de la enfermedad, entonces para calcularlo se toma en cada registro los datos de inicio de síntomas y fecha de recuperación para cada persona, se calcula el número de días entre estas dos fechas y posteriormente se toma un promedio de todos los registros y se divide por la probabilidad de que una persona se recupere de la enfermedad, la cual se obtiene restando a 1 la probabilidad de que fallezca, la cual se determina tomando $\frac{\text{total de fallecidos}}{\text{total de fallecidos + total de recuperados}}$.El valor de μ_c se halla de forma análoga pero con los datos de los fallecimientos y multiplicando por la probabilidad de que fallezca el individuo. De donde se tiene que $Dr = 16{,}6269063/(1-0{,}03634)$ y $\mu_c = 0{,}03634 * (1/20{,}18022)$

Se toman el número de infectados día a día, desde el día 24 de marzo de 2020 hasta el 1 de julio de 2020 y se ajustan los parámetros r y β de acuerdo al método de calibración expuesto en la sección 2.3 de este documento. Los demás parámetros son tomados de acuerdo al estudio [Bhattacharyya et al., 2021]; $D_E = 5{,}2$, $D_r = 17{,}8$, $\mu_c = 0{,}0562$, $\delta_1 = 0{,}3$, $\delta_2 = 0{,}7$, $\gamma_1 = 0{,}6$, $\gamma_2 = 0{,}7$, $\alpha_p = 0{,}5$, $\alpha_u = 0{,}7$, f puede ser $0, 0{,}048, 0{,}1$ y $0{,}3$. Y se asumen las siguientes condiciones iniciales para el dia 24 de marzo de: $E(0) = 15$ (Número inicial de personas expuestas), $U(0) = 0$ (Número inicial de personas infectadas no testeadas), $P(0) = 1$ (Número inicial de personas testeadas positivas en la prueba de diagnóstico del COVID-19), $F(0) = 0$ (Número inicial de personas con resultado falso negativo en la prueba de COVID-19), $A(0) = 0$ (Recuperados que fueron no testeados o falsos negativos), $B(0) = 0$ (Fallecidos que fueron no testeados o falsos negativos), $L(0) = 0$ (Recuperados que fueron testeados como positivos en el diagnostico de la prueba del COVID-19), $D(0) = 0$ (Fallecidos que fueron testeados como positivos en el diagnostico de la prueba del COVID-19). Los datos se muestran en la siguiente tabla y algunos son tomados del estudio previo de [Bhattacharyya et al., 2021] realizado en la India que son descritos en la tabla **5-3**

Posteriormente mediante el método de calibración expuesto en la sección 2.3 se estiman r y β, teniendo en cuenta que $0 \leq r \leq 1$, pues r es una probabilidad y modelando f para diferentes valores. Los parámetros r y β se estiman minimizando la función de pérdida únicamente sobre los datos de infectados testeados (y_i). Por lo que se tiene la función de pérdida $g(\theta) = ||y_i - P(\beta, r; t)||_2^2 = ||R(\beta, r)||_2^2$ y se minimiza utilizando el método de calibración (sección 2.3).

Una vez encontrados todos los parámetros se simulan las trayectorias para los individuos no testeados positivos y los falsos negativos para los diferentes valores de f como se observa

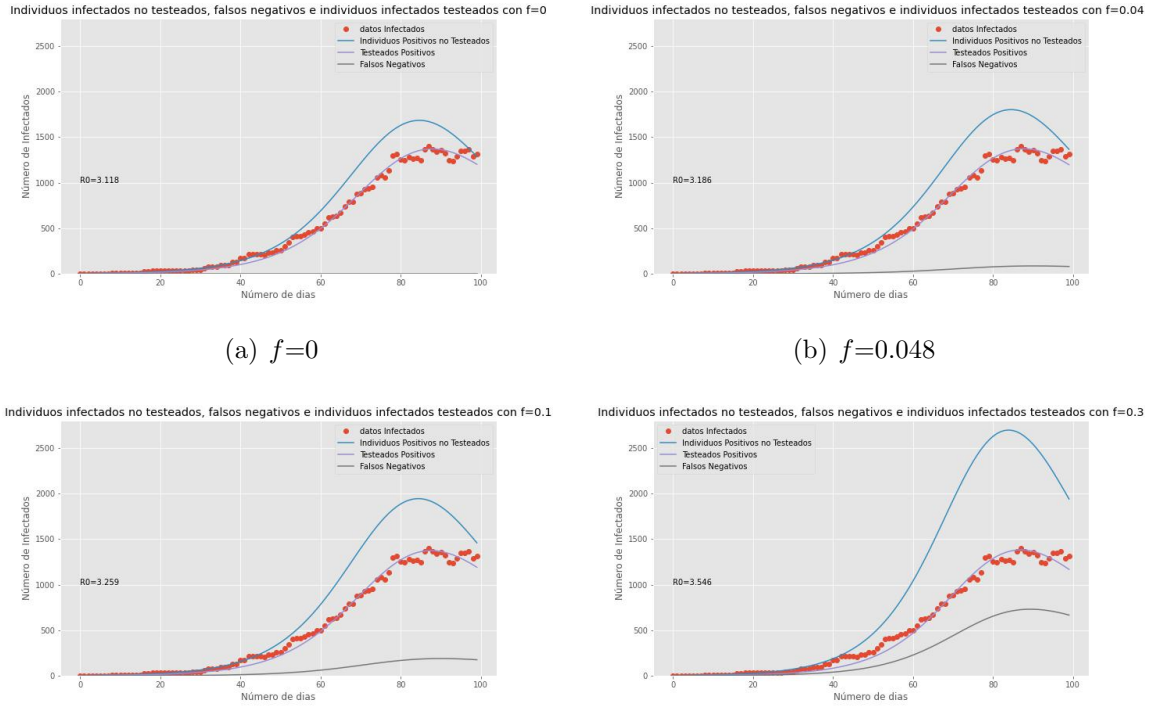
Parámetro	Valor
D_r	17.8
μ_c	0.0562
D_E	5.2
f	0.1, 0.3, 0.048, 0
δ_1	0.3
δ_2	0.7
γ_1	0.6
γ_2	0.7
α_u	0.7
α_p	0.5

Tabla 5-3.: Los parámetros δ_1, δ_2, γ_1, γ_2, α_u y α_p son cruciales para modelar la transmisión del SARS-CoV-2 en individuos no testeados, incluyendo falsos negativos. Dado que en Colombia no se ha llevado a cabo una prueba de antígenos masiva para estimarlos, se utilizaron los mismos parámetros del estudio Incorporating false negative tests in epidemiological models for SARS-CoV-2 transmission and reconciling with seroprevalence estimates [Bhattacharyya et al., 2021]. Por otro lado, los parámetros β y r se estiman utilizando el método propuesto en la sección 2.3 y se asumen $\mu = 0$ y $\lambda = 0$.

en la figura **5-2** que presentan un error cuadrático medio menor a 100 para los individuos infectados que se realizaron la prueba diagnostica.Las ecuaciones descritas anteriormente se simulan utilizando el método de Euler descrito en la sección (2.2) y pueden ser escritas como:

$$\begin{aligned}
S_{i+1} &= S_i + h\left(\lambda - \beta(\alpha_u U_i + F_i + \alpha_p P_i)\frac{S_i}{N} - \mu S_i\right), \\
E_{i+1} &= E_i + h\left(\beta\frac{S_i}{N}(\alpha_u U_i + F_i + \alpha_p P_i) - \frac{E_i}{D_E} - \mu E_i\right), \\
U_{i+1} &= U_i + h\left(\frac{(1-r)E_i}{D_E} - \frac{U_i}{\gamma_1 D_r} - \delta_1\mu_c U_i - \mu U_i\right), \\
F_{i+1} &= F_i + h\left(\frac{rfE_i}{D_E} - \frac{\gamma_2}{D_r}F_i - \frac{\mu_c}{\delta_2}F_i - \mu F_i\right), \\
P_{i+1} &= P_i + h\left(\frac{(1-f)r}{D_E}E_i - \frac{P_i}{D_r} - \mu_c P_i - \mu P_i\right), \\
L_{i+1} &= L_i + h\left(\frac{P_i}{D_r} - \mu L_i\right), \\
D_{i+1} &= D_i + h\left(\mu_c P_i\right),
\end{aligned} \tag{5-3}$$

$$A_{i+1} = A_i + h\left(\frac{U_i}{\gamma_1 D_r} + \frac{\gamma_2 F_i}{D_r} - \mu A_i\right),$$
$$B_{i+1} = B_i + h\left(\delta_1 \mu_c U_i + \frac{\mu_c F_i}{\delta_2}\right).$$



(a) f=0 (b) f=0.048

(c) f=0.1 (d) f=0.3

Figura 5-2.: Simulaciones sin vacunación variando valores de la probabilidad de tener un resultado falso negativo (f).

Los parámetros obtenidos mediante el método de calibración se muestran en las tablas **5-4** para diferentes simulaciones variando los valores de f.La estimación del R_0, que se observa en los diferentes casos se puede realizar mediante la formula 5-2.

Los valores del número reproductivo básico sin vacunación, R_0^S, para los diferentes valores de f se encuentran en la tabla **5-4**.Se puede deducir de la tabla **5-4** que a medida que la probabilidad de ser falso negativo aumenta, el número reproductivo (R_0) también aumenta.

f	Parámetro	Valor	Desviación estándar	R_0^S
0	β	0.40902	0.00139	3.118
	r	0.35976	0.00322	
0.048	β	0.40472	0.00147	3.186
	r	0.32700	0.00301	
0.1	β	0.40059	0.00155	3.259
	r	0.35213	0.00331	
0.3	β	0.38969	0.00189	3.546
	r	0.33761	0.00354	

Tabla 5-4.: Tabla de parámetros estimados β y r como su respectivo número reproductivo básico (R_0^S) para diferentes valores de falsos negativos: $f = 0$, $f = 0{,}048$, $f = 0{,}1$ y $f = 0{,}3$ sin vacunación

5.2. Modelo SEIR modificado con vacunación

Para este modelo se tiene en cuenta los 9 tipos de individuos del modelo SEIR modificado sin vacunación y se incluye el individuo vacunado (V). Este modelo añade los parámetros α_v y ρ como se muestra en la figura **5-3**

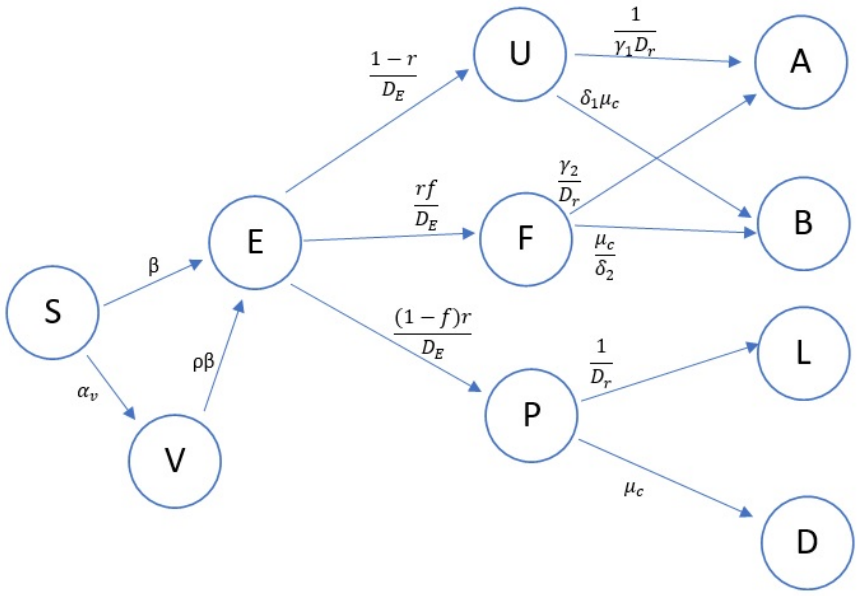


Figura 5-3.: Modelo SEIR modificado con vacunación.

En donde α_v representa la fracción de la clase susceptible (S) que está siendo vacunada en un determinado momento. Además, la vacunación puede reducir, pero no eliminar por completo, la susceptibilidad a la enfermedad. Por lo tanto, el modelo incluye un factor ρ en la tasa de contacto de los miembros vacunados, tal que $0 \leq \rho \leq 1$. Si $\rho = 0$ la vacunación

es perfectamente efectiva mientras $\rho = 1$ significa que la vacunación no tiene ningún efecto. El nuevo sistema de ecuaciones diferenciales ordinarias para este modelo se describe a continuación.

$$\begin{aligned}
\frac{dS}{dt} &= \lambda - \beta(\alpha_u U + F + \alpha_p P)\frac{S}{N} - \mu S - \alpha_v \frac{S}{N}, \\
\frac{dV}{dt} &= \alpha_v \frac{S}{N} - \rho\beta(\alpha_u U + F + \alpha_p P)\frac{V}{N} - \mu V, \\
\frac{dE}{dt} &= \beta\frac{S}{N}(\alpha_u U + F + \alpha_p P) + \rho\beta(\alpha_u U + F + \alpha_p P)\frac{V}{N} - \frac{E}{D_E} - \mu E, \\
\frac{dU}{dt} &= \frac{(1-r)E}{D_E} - \frac{U}{\gamma_1 D_r} - \delta_1\mu_c U - \mu U, \\
\frac{dF}{dt} &= \frac{rfE}{D_E} - \frac{\gamma_2}{D_r}F - \frac{\mu_c}{\delta_2}F - \mu F, \\
\frac{dP}{dt} &= \frac{(1-f)r}{D_E}E - \frac{P}{D_r} - \mu_c P - \mu P, \\
\frac{dL}{dt} &= \frac{P}{D_r} - \mu L, \\
\frac{dD}{dt} &= \mu_c P, \\
\frac{dA}{dt} &= \frac{U}{\gamma_1 D_r} + \frac{\gamma_2 F}{D_r} - \mu A, \\
\frac{dB}{dt} &= \delta_1\mu_c U + \frac{\mu_c F}{\delta_2}.
\end{aligned} \tag{5-4}$$

Al igual que en el modelo propuesto en la sección 5.1 en el modelo SEIR con vacunación se asumen los siguientes supuestos sobre la población :

- Homogeneidad de la población: El modelo asume que todos los individuos en la población tienen el mismo riesgo de infección y que interactúan de manera homogénea. Este supuesto puede no ser válido en poblaciones donde hay diferentes grupos de edad, género, niveles socioeconómicos, etc.
- Contactos aleatorios: El modelo asume que los contactos entre individuos son aleatorios y que la probabilidad de infección es la misma para todos los contactos. En la realidad, las personas pueden tener diferentes niveles de exposición al virus dependiendo de su trabajo, estilo de vida y otros factores.
- Tiempo de incubación fijo: El modelo asume que el tiempo de incubación de la enfermedad es constante y conocido. En la realidad, el tiempo de incubación puede variar entre individuos y puede no ser conocido con precisión.
- Inmunidad permanente: El modelo asume que una vez que un individuo se recupera de la enfermedad, se vuelve inmune de por vida y no puede volver a infectarse. En la

realidad, la duración de la inmunidad puede variar y en algunos casos puede no ser permanente.

El cálculo del número reproductivo básico (R_0) se realiza de forma similar al de la sección (3.2.1) de donde se tiene que:

$$\begin{pmatrix} \frac{dE}{dt} \\ \frac{dU}{dt} \\ \frac{dF}{dt} \\ \frac{dP}{dt} \end{pmatrix} = \begin{pmatrix} -\left(\frac{1}{D_E}+\mu\right) & \beta\alpha_u(1+\alpha_v\rho) & \beta(1+\alpha_v\rho) & \beta\alpha_p(1+\alpha_v\rho) \\ \frac{1-r}{D_E} & -\left(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+\mu\right) & 0 & 0 \\ \frac{rf}{D_E} & 0 & -\left(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu\right) & 0 \\ \frac{(1-f)r}{D_E} & 0 & 0 & -\left(\frac{1}{D_r}+\mu_c+\mu\right) \end{pmatrix} \cdot \begin{pmatrix} E \\ U \\ F \\ P \end{pmatrix},$$

entonces se tiene que:

$$\Sigma = \begin{pmatrix} -\left(\frac{1}{D_E}+\mu\right) & 0 & 0 & 0 \\ \frac{1-r}{D_E} & -\left(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+\mu\right) & 0 & 0 \\ \frac{rf}{D_E} & 0 & -\left(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu\right) & 0 \\ \frac{(1-f)r}{D_E} & 0 & 0 & -\left(\frac{1}{D_r}+\mu_c+\mu\right) \end{pmatrix},$$

$$\mathbf{T} = \begin{pmatrix} 0 & \beta(\alpha_u+\alpha_v) & \beta\alpha_v & \beta(\alpha_p+\alpha_v) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\Sigma^{-1} = \begin{pmatrix} \frac{D_E}{D_E\mu+1} & 0 & 0 & 0 \\ -\frac{r-1}{(D_E\mu+1)(\delta_1+\mu_c+\mu+\frac{1}{D_r\delta_1})} & \frac{1}{\left(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+\mu\right)} & 0 & 0 \\ -\frac{rf}{(D_E\mu+1)\left(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu\right)} & 0 & \frac{1}{\left(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu\right)} & 0 \\ -\frac{(1-f)r}{(D_E\mu+1)(\mu_c+\mu+\frac{1}{D_r})} & 0 & 0 & \frac{1}{\left(\frac{1}{D_r}+\mu_c+\mu\right)} \end{pmatrix},$$

Siguiendo los pasos de la sección (3.2.1) como

$$R_0^C = \frac{1}{2}\left(trace(\text{-}\mathbf{T}\Sigma) + \sqrt{trace(\text{-}\mathbf{T}\Sigma)^2 - 4det(\text{-}\mathbf{T}\Sigma)}\right),$$

Entonces

$$R_0^C = \frac{\beta}{\mu D_e+1}\left(\frac{\alpha_u(1+\alpha_v\rho)(1-r)}{\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+\mu} + \frac{\alpha_p(1+\alpha_v\rho)r(1-f)}{\frac{1}{D_r}+\mu_c+\mu} + \frac{rf(1+\alpha_v\rho)}{\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu}\right), \qquad (5\text{-}5)$$

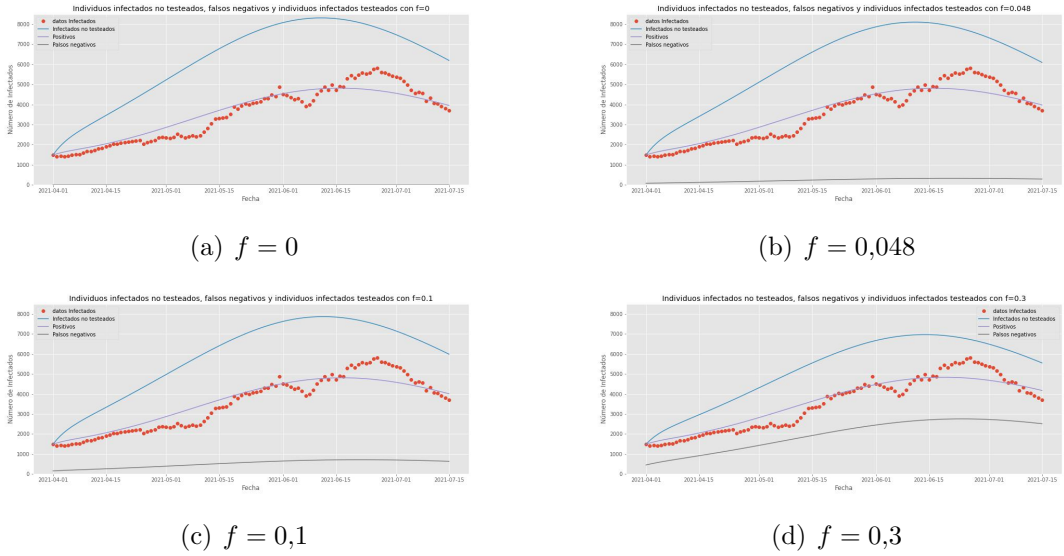
donde R_0^C es el número reproductivo básico para el modelo con vacunación.

Para este modelo se tomaron los datos desde 1 de abril de 2021 hasta el 15 de julio de 2021 donde ocurren uno de los picos de la pandemia, mediante el método de calibración expuesto en la sección 2.3 se estiman r y β, teniendo en cuenta que $0 \leq r \leq 1$, pues r es una probabilidad y modelando f para diferentes valores. Se utiliza para modelar la función de pérdida únicamente los datos sobre los infectados y_i. Por lo que se tiene la función de pérdida $g(\theta) = ||y_i - P(\beta, r; t)||_2^2 = ||R(\beta, r)||_2^2$ y se minimiza utilizando el método de calibración (sección 2.3).

Una vez encontrados todos los parámetros se simulan las trayectorias para los individuos no testeados positivos y los falsos negativos para los diferentes valores de f como se observa en la figura **5-2**. Se varia el valor de f en el porque este parámetro puede tener un impacto en las trayectorias simuladas de los individuos no testeados positivos y los falsos negativos. Las ecuaciones descritas anteriormente se simulan utilizando el método de Euler descrito en la sección (2.2) y pueden ser escritas como:

$$
\begin{aligned}
S_{i+1} &= S_i + h\left(\lambda - \beta(\alpha_u U_i + F_i + \alpha_p P_i)\frac{S_i}{N} - \mu S_i - \alpha_v \frac{S_i}{N}\right), \\
V_{i+1} &= V_i + h\left(\alpha_v \frac{S}{N} - \rho\beta(\alpha_u U + F + \alpha_p P)\frac{V}{N} - \mu V\right), \\
E_{i+1} &= E_i + h\left(\beta\frac{S_i}{N}(\alpha_u U_i + F_i + \alpha_p P_i) + \rho\beta(\alpha_u U_i + F_i + \alpha_p P_i)\frac{V_i}{N} - \frac{E_i}{D_E} - \mu E_i\right), \\
U_{i+1} &= U_i + h\left(\frac{(1-r)E_i}{D_E} - \frac{U_i}{\gamma_1 D_r} - \delta_1\mu_c U_i - \mu U_i\right), \\
F_{i+1} &= F_i + h\left(\frac{rfE_i}{D_E} - \frac{\gamma_2}{D_r}F_i - \frac{\mu_c}{\delta_2}F_i - \mu F_i\right), \\
P_{i+1} &= P_i + h\left(\frac{(1-f)r}{D_E}E_i - \frac{P_i}{D_r} - \mu_c P_i - \mu P_i\right), \\
L_{i+1} &= L_i + h\left(\frac{P_i}{D_r} - \mu L_i\right), \\
D_{i+1} &= D_i + h\left(\mu_c P_i\right), \\
A_{i+1} &= A_i + h\left(\frac{U_i}{\gamma_1 D_r} + \frac{\gamma_2 F_i}{D_r} - \mu A_i\right), \\
B_{i+1} &= B_i + h\left(\delta_1\mu_c U_i + \frac{\mu_c F_i}{\delta_2}\right).
\end{aligned}
\tag{5-6}
$$

La figura (5-4) presenta los resultados de las trayectorias de las simulaciones de individuos con resultado falso negativo, infectados no testeados, e individuos infectados con prueba diagnóstica. Para este último grupo, se obtiene un error cuadrático medio de alrededor de



(a) $f = 0$ (b) $f = 0{,}048$

(c) $f = 0{,}1$ (d) $f = 0{,}3$

Figura 5-4.: Simulaciones con vacunación variando valores de la probabilidad de tener un resultado falso negativo (f).

1000. La Figura 5.4(a) muestra el escenario en el que la probabilidad que la prueba diagnostica sea falsa negativo es 0 ($f = 0$). En este caso, el número de individuos infectados aumenta hasta alcanzar un pico. La Figura 5.4(b) muestra el resultado de una simulación en la que la probabilidad de obtener un resultado falso negativo es baja ($f = 0{,}048$). Aunque el número de individuos infectados sigue siendo alto, se observa una disminución gradual en el número de infectados después de un período de rápido aumento. En la Figura 5.4(c), la probabilidad de obtener un resultado falso negativo se ha incrementado ($f = 0{,}1$). El número de individuos infectados aumenta rápidamente al principio y luego disminuye lentamente. En la Figura 5.4(d), se ha incrementado la probabilidad de obtener un resultado falso negativo ($f = 0{,}3$). En este escenario, el número de individuos infectados no testeados es menor que en los otros modelos utilizando los otros valores de probabilidades, esto se debe a que la probabilidad de tener falso negativo altera los compartimientos del modelo de todos los compartimientos que transmiten el virus.

Los parámetros obtenidos mediante el método de calibración se muestran en la tabla **5-5** para diferentes simulaciones variando los valores de f. La estimación del número reproductivo básico con vacunación, R_0^C, que se observa en los diferentes casos se puede realizar mediante la ecuación 5-5. Los valores de R_0^C para los diferentes valores de f se encuentran en la tabla **5-5**

f	Parámetro	Valor	Desviación estándar	R_0^C
0	β	0.194894	0.01434	1.891
	r	0.29102	0.01943	
0.048	β	0.18442376	0.02137664	2.042
	r	0.30011677	0.03052771	
0.1	β	0.18046702	0.02099582	2.078
	r	0.28313102	0.002419797	
0.3	β	0.17154699	0.01491504	2.108
	r	0.26872638	0.01647014	

Tabla 5-5.: Tabla de parámetros estimados β y r como su respectivo número reproductivo básico (R_0^C) para diferentes valores de falsos negativos: $f = 0$, $f = 0{,}048$, $f = 0{,}1$ y $f = 0{,}3$ con vacunación

5.3. Modelo SEIR modificado estocástico con vacunación

En esta sección se introduce el modelo estocástico epidemiológico con perturbaciones aleatorias. En este modelo se asume $(\Omega, \Im, \{\Im_t\}_{t\geq 0}, P)$ es un espacio de probabilidad completo con una filtracion $\{\Im_t\}_{t\geq 0}$. Se define:

$$\tilde{\beta} = \beta + \sigma W(t)$$

donde β y σ son constantes positivas, y $\{W(t)\}_{t\geq 0}$ es el movimiento browniano estandar con $W(T) - W(t) \sim N(0, T-t)$. Donde β representa la tasa de infección y σ un parámetro de perturbación.

El modelo SEIR con vacunación es una herramienta útil para predecir la propagación del COVID-19 y evaluar el impacto de las medidas de vacunación. Sin embargo, este modelo se basa en una serie de supuestos y simplificaciones, que pueden no reflejar con precisión la realidad de la situación epidemiológica [Tornatore et al., 2006]. En particular, el modelo SEIR determinista asume que la transmisión del virus es constante y predecible en el tiempo.

Sin embargo, la propagación del COVID-19 está sujeta a fluctuaciones aleatorias debido a factores estocásticos, como la variabilidad en el comportamiento humano, la capacidad del sistema de salud para detectar y tratar casos, y la eficacia de la vacuna en la población. Estas fluctuaciones pueden afectar significativamente la propagación del virus y la eficacia de las medidas de vacunación.

Por lo tanto, el propósito de considerar una perturbación estocástica en el modelo SEIR con vacunación es tener en cuenta la variabilidad aleatoria en la propagación del virus y la eficacia de la vacuna. Esto permite obtener una comprensión más realista de la dinámica de

la epidemia y proporciona una mejor base para la toma de decisiones sobre medidas de salud pública y políticas de vacunación.

La inclusión de perturbaciones estocásticas en el modelo SEIR con vacunación permite simular diferentes escenarios epidemiológicos y evaluar la probabilidad de ocurrencia de eventos como brotes epidémicos y la propagación del virus en poblaciones con diferentes tasas de vacunación. También permite estimar la eficacia de las medidas de vacunación en diferentes condiciones, lo que es fundamental para la toma de decisiones informadas sobre la gestión de la pandemia del COVID-19. En resumen, la inclusión de perturbaciones estocásticas en el modelo SEIR con vacunación ayuda a mejorar la precisión y la utilidad del modelo para predecir la propagación del virus y diseñar estrategias efectivas de salud pública y vacunación.

Con el fin de incluir la aleatoriedad en un modelo epidémico compartimental, varios autores han propuesto modelos con perturbación estocástica [Tornatore et al., 2006]. Esto significa que se modifica el sistema de ecuaciones diferenciales ordinarias (EDO) agregando ruido estocástico o perturbaciones estocásticas, dando lugar a un sistema de ecuaciones diferenciales estocásticas (EDS).

Con el fin de considerar la aleatoriedad en el modelo SEIR modificado con vacunación se propone el modelo SEIR modificado estocástico con vacunación [Witbooi et al., 2015], en el que se introduce un ruido estocástico que varia el sistema de ecuaciones diferenciales. De allí se tiene el siguiente sistema de ecuaciones diferenciales estocásticas:

$$
\begin{aligned}
dS &= \left(\lambda - \beta(\alpha_u U + F + \alpha_p P)\frac{S}{N} - \mu S - \alpha_v \frac{S}{N}\left(\alpha_u U + F + \alpha_p P\right)\right)dt - \sigma S\frac{(\alpha_u U + F + \alpha_p P)}{N}dW,\\
dV &= \left(\alpha_v \frac{S}{N}\left(\alpha_u U + F + \alpha_p P\right) - (\rho\beta + \mu)V\right)dt - \sigma\rho V\frac{(\alpha_u U + F + \alpha_p P)}{N}dW,\\
dE &= \left(\beta\frac{S}{N}(\alpha_u U + F + \alpha_p P) + V\rho\beta - \frac{E}{D_E} - \mu E\right)dt + \sigma(S + \rho V)\frac{(\alpha_u U + F + \alpha_p P)}{N}dW,\\
dU &= \left(\frac{(1-r)E}{D_E} - \frac{U}{\gamma_1 D_r} - \delta_1\mu_c U - \mu U\right)dt,\\
dF &= \left(\frac{rfE}{D_E} - \frac{\gamma_2}{D_r}F - \frac{\mu_c}{\delta_2}F - \mu F\right)dt,\\
dP &= \left(\frac{(1-f)r}{D_E}E - \frac{P}{D_r} - \mu_c P - \mu P\right)dt,\\
dL &= \left(\frac{P}{D_r} - \mu L\right)dt,\\
dD &= \mu_c P dt,\\
dA &= \left(\frac{U}{\gamma_1 D_r} + \frac{\gamma_2 F}{D_r} - \mu A\right)dt,\\
dB &= \left(\delta_1\mu_c U + \frac{\mu_c F}{\delta_2}\right)dt.
\end{aligned}
\tag{5-7}
$$

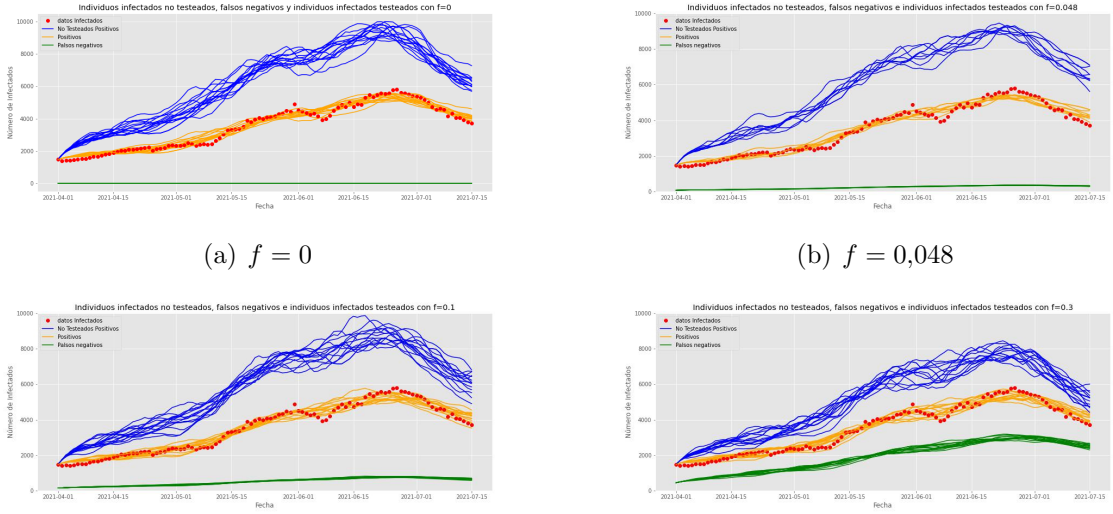
Donde $\{W_t\}_{t\geq 0}$ es un movimiento browniano. Para obtener la soluciones aproximadas del

anterior sistema de ecuaciones diferenciales estocásticas se implementó el método de Euler-Maruyama. Se utilizó el lenguaje de programación Python 3.7 para llevar a cabo el algoritmo computacional, previamente mencionado con la respectiva discretización del sistema. Se simulan los datos utilizados los mismos parámetros encontrados en el modelo con vacunación y con un valor $\sigma = 0{,}2$ [Witbooi et al., 2015]. Dicha discretización está determinada por:

$$
\begin{aligned}
S_{i+1} &= S_i \left(\lambda - (\beta + \alpha_v)\left(\alpha_u U_i + F_i + \alpha_p P_i\right)\frac{S_i}{N} - \mu S_i\right)\Delta t - \sigma S_i \frac{(\alpha_u U_i + F_i + \alpha_p P_i)}{N}\sqrt{\Delta t}\eta,\\
V_{i+1} &= V_i + \left(\alpha_v \frac{S_i}{N}\left(\alpha_u U_i + F_i + \alpha_p P_i\right) - (\rho\beta + \mu)V_i\right)\Delta t - \sigma\rho V_i \frac{(\alpha_u U_i + F_i + \alpha_p P_i)}{N}\sqrt{\Delta t}\eta,\\
E_{i+1} &= E_i + \left(\frac{\beta S_i}{N}(\alpha_u U_i + F_i + \alpha_p P_i) + V_i\rho\beta - \frac{E_i}{D_E} - \mu E_i\right)\Delta t + \sigma(S_i + \rho V_i)\frac{(\alpha_u U_i + F_i + \alpha_p P_i)}{N}\sqrt{\Delta t}\eta,\\
U_{i+1} &= U_i + \left(\frac{(1-r)E}{D_E} - \frac{U_i}{\gamma_1 D_r} - \delta_1\mu_c U_i - \mu U_i\right)\Delta t,\\
F_{i+1} &= F_i + \left(\frac{rfE_i}{D_E} - \frac{\gamma_2}{D_r}F_i - \frac{\mu_c}{\delta_2}F_i - \mu F_i\right)\Delta t,\\
P_{i+1} &= P_i + \left(\frac{(1-f)r}{D_E}E_i - \frac{P_i}{D_r} - \mu_c P_i - \mu P_i\right)\Delta t,\\
L_{i+1} &= L_i + \left(\frac{P_i}{D_r} - \mu L_i\right)\Delta t,\\
D_{i+1} &= D_i + \mu_c P_i \Delta t,\\
A_{i+1} &= A_i\left(\frac{U_i}{\gamma_1 D_r} + \frac{\gamma_2 F_i}{D_r} - \mu A_i\right)\Delta t,\\
B_{i+1} &= B_i + \left(\delta_1\mu_c U_i + \frac{\mu_c F_i}{\delta_2}\right)\Delta t.
\end{aligned}
\tag{5-8}
$$

Se obtiene entonces un modelo epidemiológico en donde las trayectorias de los individuos no testeados, positivos y falsos negativos son variables aleatorias. Los resultados de las trayectorias estocasticas mostradas en la figura **5-5** tienen un error cuadrático medio menor a 230. En general los modelos SEIR estocásticos permiten representar de forma mas natural la propagación de la enfermedad, pues los parámetros usados para caracterizar el modelo son variables aleatorias que tienen unos comportamientos estimados pero no se conoce con certidumbre previamente cuál será el valor que tomen. De allí la importancia de la modelación estocástica. En la siguiente sección se calcula el número reproductivo básico estocástico asociado al modelo propuesto.

Todas las simulaciones presentadas en las Figuras 5.5(a), 5.5(b), 5.5(c), y 5.5(d) fueron realizadas utilizando ecuaciones diferenciales estocásticas. Es importante destacar que todas las variables que influyen en la propagación del virus fueron modeladas teniendo en cuenta la probabilidad de obtener un resultado falso negativo en las pruebas diagnósticas. En la Figura 5.5(a), se muestra el escenario en el que la probabilidad de obtener un resultado falso negativo es 0, y se observa un aumento en el número de individuos infectados hasta alcanzar



(a) $f = 0$ (b) $f = 0{,}048$

(c) $f = 0{,}1$ (d) $f = 0{,}3$

Figura 5-5.: Simulaciones estocásticas tomando diferentes valores de la probabilidad de tener un resultado falso negativo (f)

un pico. En la Figura 5.5(b), se utilizó una probabilidad baja de falso negativo ($f = 0{,}048$), lo que resultó en una disminución gradual del número de infectados después de un período de rápido aumento. En la Figura 5.5(c), se incrementó la probabilidad de falso negativo ($f = 0{,}1$), lo que provocó un rápido aumento en el número de infectados seguido de una disminución lenta. Finalmente, en la Figura 5.5(d), se utilizó una probabilidad aún mayor de falso negativo ($f = 0{,}3$), lo que resultó en una disminución en el número de individuos infectados no testeados en comparación con los otros escenarios. Es importante destacar que la probabilidad de obtener un resultado falso negativo afecta todos los compartimientos del modelo que transmiten el virus y, por lo tanto, es un factor crucial a tener en cuenta en la simulación de la propagación del virus.

5.4. Número reproductivo básico del modelo estocástico con vacunación

En esta sección se aborda el tema del número reproductivo básico en un modelo que incluye la vacunación. El procedimiento para determinarlo se basa en los trabajos previos de [Niño-Torres et al., 2022] y [Ríos-Gutiérrez et al., 2021].

Como se describió en la sección 3.2.1 el número reproductivo básico R_0 puede ser calculado

mediante el uso de la integral:

$$R_0 := \int_0^{+\infty} b\,(a)\,F\,(a)\,da,$$

donde $b\,(a)$ es el número promedio de individuos infectados, y $F\,(a)$ es la probabilidad de que un nuevo individuo infectado infecte continuamente durante el intervalo de tiempo entre 0 y a.

Considere el modelo SEIR modificado con perturbación aleatoria, el número reproductivo estocástico

$$R_{0,v}^{SEIR} := \int_0^{+\infty} \left(\beta + \sigma B\,(a)\right) N e^{-(\mu+\gamma)a} da\ ,$$

En el modelo SEIR modificado con perturbación aleatoria y vacunación, el número reproductivo estocástico puede ser modelado por la ecuación anterior, donde $R_{0,v}^{SEIR}$ es una variable aleatoria normalmente distribuida. En esta ecuación, β es la tasa de transmisión efectiva, σ es la tasa de transmisión efectiva entre individuos vacunados, $B(a)$ es la función de densidad de probabilidad de la perturbación aleatoria que es nuestro estudio es el movimiento browniano, N es el tamaño de la población, μ es la tasa de mortalidad natural y γ es la tasa de recuperación, donde $R_{0,v}^{SEIR}$ es una variable aleatoria normalmente distribuida.

Como se describió en la sección 5.2 el número reproductivo básico puede ser calculado mediante la ecuación 5-5:

$$R^C = \frac{\beta}{D_e+1}\left(\frac{\alpha_u(1+\alpha_v\rho)(1-r)}{\frac{1}{\gamma_1 D_r}+\delta_1\mu_c} + \frac{\alpha_p(1+\alpha_v\rho)r(1-f)}{\frac{1}{D_r}+\mu_c+\mu} + \frac{rf(1+\alpha_v\rho)}{\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}}\right)$$

El número reproductivo básico mostrado en la ecuación anterior, puede interpretarse como la suma de tres números reproductivos básicos; teniendo en cuenta el número promedio de nuevos infectados que no se realizarán las pruebas, el número promedio de nuevos infectados que se realizarán las pruebas diagnósticas y cuyo resultado no se considerará como falso negativo y el número promedio de nuevos infectados que se considerarán como falsos negativos.

$$R^C = \sum_{i=1}^{3} R_i^C$$

Donde se tiene que:

$$R^C = \frac{\beta}{D_e+1}\cdot\frac{\alpha_u(1-r)(\alpha_v\rho+1)}{\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+} + \frac{\beta}{D_e+1}\cdot\frac{\alpha_p r(1-f)(\alpha_v\rho+1)}{\frac{1}{D_r}+\mu_c} + \frac{\beta}{D_e+1}\cdot\frac{rf(\alpha_v\rho+1)}{\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}}$$

Y por lo tanto:

A. Teniendo en consideración el número reproductivo básico el número promedio de nuevos infectados que no se realizarán las pruebas, al tener contacto con aquellos individuos infecciosos, corresponde a:

$$R_1^C = \frac{\beta}{D_e + 1} \cdot \frac{\alpha_u(1-r)(\alpha_v\rho + 1)}{\frac{1}{\gamma_1 D_r} + \delta_1\mu_c}.$$

B. El número promedio de nuevos infectados que se realizarán las pruebas diagnósticas y cuyo resultado no se considerará como falso negativo, al tener contacto infeccioso con aquellos individuos que son infecciosos corresponde a:

$$R_2^C = \frac{\beta}{D_e + 1} \cdot \frac{\alpha_p r(1-f)(\alpha_v\rho + 1)}{\frac{1}{D_r} + \mu_c}.$$

C. El número promedio de nuevos infectados que se considerarán como falsos negativos, corresponde a:

$$R_3^C = \frac{\beta}{D_e + 1} \cdot \frac{rf(\alpha_v\rho + 1)}{\frac{\gamma_2}{D_r} + \frac{\mu_c}{\delta_2}}.$$

A continuación se procede a establecer las funciones $b(a)$ y $F(a)$ [Heffernan et al., 2005] correspondientes a

- $b(a)$ que corresponde al número promedio de nuevos individuos infectados (de una población completamente susceptible) por un individuo infectado si éste permanece infeccioso durante un tiempo a, y,
- $F(a)$ que es la probabilidad de que un nuevo infectado continúe infectando durante el tiempo a

Las funciones $b(a)$ para los ítem A, B y C están dadas respectivamente por $b_A(a) = \beta\alpha_u(1-r)(\alpha_v\rho + 1)$, $b_B(a) = \beta\alpha_p r(1-f)(\alpha_v\rho + 1)$, $b_C(a) = \beta(\alpha_v\rho + 1)$.

De otra parte, siguiendo un razonamiento similar a [Ríos-Gutiérrez et al., 2021], se establece que las funciones $F(a)$ para los ítem A, B y C están dadas respectivamente por

$F_A(a) = \exp\left\{-(D_e + 1)(\frac{1}{\gamma_1 D_r} + \delta_1\mu_c)a\right\}$,
$F_B(a) = \exp\left\{-(D_e + 1)(\frac{1}{D_r} + \mu_c)a\right\}$ y

$F_C(a) = \exp\left\{-(D_e+1)(\frac{\gamma_2}{D_r} + \frac{\mu_c}{\delta_2})a\right\}$.
N
Obsérvese que

$$\begin{aligned}\int_0^{+\infty} b_A(a)F_A(a)da &= (\beta\alpha_u(1+\alpha_v\rho)(1-r)) \int_0^{+\infty} e^{-(D_e+1)(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c)a}da \\ &= \frac{\beta\alpha_u(1+\alpha_v\rho)(1-r)}{(D_e+1)(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c)}\end{aligned}$$

obteniendo la expresión correspondiente a A. De manera análoga se obtienen las expresiones para B y para C . Para el modelo estocástico se definen las funciones $F(a)$ de manera análoga al caso determinista y $b(a)$ como sigue:

$$\begin{aligned}b_{A,e}(a) &= (\beta\rho + \sigma\rho W(a))\alpha_u\alpha_v(1-r) + (\beta + \sigma W(a))\alpha_u(1-r), \\ b_{B,e}(a) &= (\beta\rho + \sigma\rho W(a))\alpha_p\alpha_v r(1-f) + (\beta + \sigma W(a))\alpha_p r(1-f), \\ b_{C,e}(a) &= (\beta\rho + \sigma\rho W(a))rf\alpha_v + (\beta + \sigma W(a))rf,\end{aligned}$$

Por ejemplo, la integral del ítem A, se tiene la siguiente expresión para establecer el correspondiente promedio, tomando $\psi = (D_e+1)(\frac{1}{\gamma_1 D_r} + \delta_1\mu_c)$

$$\int_0^{+\infty} b_{A,e}(a)F_{A,e}(a)da = \int_0^{+\infty} \beta\alpha_u(1+\alpha_v\rho)(1-r)e^{-\psi a}da + \int_0^{+\infty} \sigma\alpha_u(1+\alpha_v\rho)(1-r)W(a)e^{-\psi a}da$$
$$= \frac{\beta\alpha_u(1+\alpha_v\rho)(1-r)}{(D_e+1)(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c)} + \int_0^{+\infty} \sigma\alpha_u(1+\alpha_v\rho)(1-r)W(a)e^{-\psi a}da$$

donde

$$\int_0^{+\infty} \sigma\alpha_u(1+\alpha_v\rho)(1-r)W(a)e^{-\psi a}da \sim N\left(0, \frac{\sigma^2\alpha_u^2(1+\alpha_v\rho)^2(1-r)^2}{2(D_e+1)^3(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c)^3}\right),$$

usando el procedimiento utilizado en [Ríos-Gutiérrez et al., 2021]; de manera que

$$\int_0^{+\infty} b_{A,e}(a)F_{A,e}(a)da \sim N\left(\frac{\beta\alpha_u(1+\alpha_v\rho)(1-r)}{(D_e+1)(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c)}, \frac{\sigma^2\alpha_u^2(1+\alpha_v\rho)^2(1-r)^2}{2(D_e+1)^3(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c)^3}\right)$$

De manera análoga, se tienen las siguientes distribuciones:

$$\int_0^{+\infty} b_{B,e}(a)F_{B,e}(a)da \sim N\left(\frac{\beta\alpha_p(1+\alpha_v\rho)r(1-f)}{(D_e+1)(\frac{1}{D_r}+\mu_c)}, \frac{\sigma^2\alpha_p^2(1+\alpha_v\rho)^2r^2(1-f)^2}{2(D_e+1)^3(\frac{1}{D_r}+\mu_c)^3}\right)$$

$$\int_0^{+\infty} b_{C,e}(a)F_{C,e}(a)da \sim \quad N\left(\frac{\beta r f(1+\alpha_v\rho)}{(D_e+1)(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2})}, \frac{\sigma^2 r^2 f^2(1+\alpha_v\rho)^2}{2(D_e+1)^3(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2})^3}\right).$$

Bajo el supuesto de que los eventos determinados por el hecho de que un individuo infectado no se realice una prueba diagnóstica, se realice la prueba y su resultado sea considerado como falso negativo, o el resultado no se considere como falso negativo; son eventos independientes, entonces se tiene que las variables aleatorias $\int_0^{+\infty} b_{A,e}(a)F_{A,e}(a)da$, $\int_0^{+\infty} b_{B,e}(a)F_{B,e}(a)da$ y $\int_0^{+\infty} b_{C,e}(a)F_{C,e}(a)$ son independientes, por consiguiente:

$$\begin{aligned} R_{0,E}^C &= \int_0^{+\infty} b_{A,e}(a)F_{A,e}(a)da + \int_0^{+\infty} b_{B,e}(a)F_{B,e}(a)da + \int_0^{+\infty} b_{C,e}(a)F_{C,e}(a)da \\ &\sim N\left(R_0^C, \frac{\sigma^2\alpha_u^2(1+\alpha_v\rho)^2(1-r)^2}{2(\mu D_e+1)^3(\frac{1}{\gamma_1 D_r}+\delta_1\mu_c+\mu)^3} + \frac{\sigma^2\alpha_p^2(1+\alpha_v\rho)^2 r^2(1-f)^2}{2(\mu D_e+1)^3(\frac{1}{D_r}+\mu_c+\mu)^3}\right. \\ &\quad \left. + \frac{\sigma^2 r^2 f^2(1+\alpha_v\rho)^2}{2(\mu D_e+1)^3(\frac{\gamma_2}{D_r}+\frac{\mu_c}{\delta_2}+\mu)^3}\right), \end{aligned}$$

donde $R_{0,E}^C$ es la variable aleatoria reproductiva básica [Ríos Gutiérrez, 2019] del modelo estocástico con vacunación. Se estima el número reproductivo para los diferentes valores de f (ver tabla **5-6**):

f	$R_{0,E}^C$
0	N (1,891, 3,892)
0,048	N (2,042, 3,794)
0,1	N (2,078, 3,955)
0,3	N (2,108, 4,135)

Tabla 5-6.: Valores de $R_{0,E}^C$ para el modelo SEIR modificado estocástico

En particular, para el modelo SEIR estocástico, el número reproductivo básico es una variable aleatoria que se distribuye normal y aumenta su media, a medida que la probabilidad de ser falso negativo (f) aumenta, como se observa en la tabla **5-6** para los valores de f entre 0 y 0.3 la media se ubica entre 1,819 y 2,108, mientras que su varianza se encuentra entre 3,794 y 4,135. Se sugiere en futuros trabajos comparar el número reproductivo encontrado en esta Tesis con el de otras ciudades en Colombia aplicando el modelo propuesto en este trabajo o utilizando otros modelos ya que dicho número puede variar de región en región.

Por último, en el siguiente capítulo se encuentran algunas de las conclusiones y recomendaciones propuestas de esta Tesis, que pueden ser utilizadas para futuras investigaciones en modelos SEIR deterministas y estocásticos como en otras investigaciones epidemiológicas.

Conclusiones y Recomendaciones

El modelo epidemiológico estocástico propuesto se ve potencialmente limitado porque se realizan suposiciones sobre los datos iniciales de individuos no testeados que tienen el virus y falsos negativos tanto para los modelos de vacunación como los modelos sin vacunación, dado que no se cuenta con información de estos grupos y que además varían en el tiempo. Además de que no se tiene en cuenta migraciones y efectos demográficos que pueden ser fundamentales en la propagación de una pandemia.

Por medio de este estudio se ha podido concluir que cuando se utiliza el modelo SEIR modificado con vacunación se obtiene un número reproductivo básico menor que indica que un individuo infectado bajo este modelo contagia menos individuos que si éste estuviera en el modelo sin vacunación. A medida que aumenta la probabilidad de que la prueba de resultado falso negativo, así mismo aumenta el número reproductivo básico (R_0) estimado tanto para el modelamiento con vacunación como para el modelamiento sin vacunación.

En términos generales, se puede decir que cuanto más complejo es el modelo epidemiológico utilizado, mayor será su capacidad para describir con precisión la propagación del virus en una región determinada. El modelo estocástico SEIR modificado es el más complejo de todos los presentados y tiene en cuenta varios factores importantes como los falsos negativos y asintomáticos. Sin embargo, se necesitan más investigaciones y datos para determinar cuál es el mejor modelo propuesto en este trabajo.

Sin embargo, a pesar de los diferentes problemas mencionados, el modelo le permite al investigador obtener una perspectiva del posible alcance que puede traer una enfermedad infecciosa, en este caso la COVID-19, con comportamientos aleatorios en la población. Este trabajo abre un camino para explorar modelos de compartimentos más novedosos que tenga en cuenta las migraciones, efectos demográficos y espaciales. El modelo SEIR estocástico modificado presentado en el documento es una buena aproximación para describir la propagación del Covid-19 en el Pacífico Nariñense. Sin embargo, se necesitan más investigaciones y análisis adicionales para mejorar su precisión y utilidad. En particular, se sugiere realizar estudios de sensibilidad para evaluar cómo cambian las predicciones del modelo cuando se varían los parámetros clave.

A. Códigos usados para modelar el comportamiento del COVID-19

A.1. Tratamiento de datos

Listing A.1: Tratamiento de datos

```
import os
import sys
from scipy.integrate import odeint
import sdeint
import plotly.graph_objects as go
import plotly.io as pio
import requests
from lmfit import minimize, Parameters, Parameter, report_fit
import numpy as np
import pandas as pd
import scipy as sp
import matplotlib.pylab as plt
from matplotlib.pylab import rcParams
from scipy.signal import savgol_filter
rcParams['figure.figsize'] = 15, 6
pio.renderers.default = "notebook"
%matplotlib inline
plt.style.use('ggplot')
## Datos de covid https://www.datos.gov.co/Salud-y-Protecci-n-Social/
Casos-positivos-de-COVID-19-en-Colombia/gt2j-8ykr/data  ##
data=pd.read_csv("gdrive/MyDrive/Colab_Notebooks/DatosNari o.csv")

#Infectados por Dia

data['Fecha_de_diagn stico']=pd.to_datetime(data
['Fecha_de_diagn stico'])
#convert to time series:
```

```
Infectados=data["Fecha de diagn stico"].value_counts()
Infectados=Infectados.sort_index()
Infectados.index = pd.to_datetime(Infectados.index)
Infectados=Infectados.to_frame(name="Infectados")

####Graficar la Serie#####
rcParams['figure.figsize'] = 15, 6
plt.plot(Infectados)
plt.title('Casos Infectados por dia del Covid-19 en el departamento
del Nari o')
plt.scatter(Infectados.index,Infectados, label="Infectados")
plt.savefig("gdrive/MyDrive/ImagenesTesis/Infectadosdia.jpg")
plt.legend(loc='best')
plt.show()

#Recuperados por dia
Recuperados=data[data["Recuperado"]=="Recuperado"]
Recuperados['Fecha de recuperaci n']=pd.to_datetime(Recuperados
['Fecha de recuperaci n'])
Recuperados=Recuperados["Fecha de recuperaci n"].value_counts()
Recuperados=Recuperados.sort_index()
Recuperados.index = pd.to_datetime(Recuperados.index)
Recuperados=Recuperados.to_frame(name="Recuperados")


####Graficar la Serie#####
rcParams['figure.figsize'] = 15, 6
plt.scatter(Recuperados.index,Recuperados, label="Recuperados")
#plt.scatter(Infectados.index, Infectados, label="Infectados")
plt.title('Casos de infectados y recuperados Covid-19 en Nari o
por dia')
plt.legend(loc='best')
plt.show()

Fallecidos=data[(data["Recuperado"]=="Fallecido")|(data["Recuperado"]==
"fallecido")|(data['Recuperado'].isnull())]
Fallecidos['Fecha de muerte']=pd.to_datetime(Fallecidos['Fecha
de muerte'])
Fallecidos=Fallecidos["Fecha de muerte"].value_counts()
Fallecidos=Fallecidos.sort_index()
```

```
Fallecidos.index = pd.to_datetime(Fallecidos.index)
Fallecidos=Fallecidos.to_frame(name="Fallecidos")
####Graficar la Serie#####
plt.scatter(Fallecidos.index,Fallecidos, label="Fallecidos")
plt.scatter(Recuperados.index,Recuperados, label="Recuperados")
plt.scatter(Infectados.index,Infectados, label="Infectados")
plt.title('Casos de infectados, recuperados y muertos Covid-19
en Nari o por dia')
plt.legend(loc='best')
plt.show()

#Union de datos

Datos1=Infectados.join(Recuperados)
Datos=Datos1.join(Fallecidos)
Datos=Datos.fillna(0)

s = pd.Series(pd.date_range("2020-3-24", periods=365*2-17, freq="D"))
td = pd.Series([pd.Timedelta(days=i) for i in range(365*2-17)])
df = pd.DataFrame({"A": s, "B": td})
df.index = pd.to_datetime(df["A"])
Datosfinal=df.join(Datos)
Datosfinal = Datosfinal.replace(np.nan, 0)

def Totales(NombreColumnaDataframe,NombreTotales):
    total=Datosfinal[NombreColumnaDataframe].values
    Lista=[]
    for i in range(len(total)):
        if i==0:
            save=total[i]
            Lista.append(save)
        else:
            save=total[i]+Lista[-1]
            Lista.append(save)
    Datosfinal[NombreTotales]=Lista

Totales("Infectados","Total")
Totales("Recuperados","Total Recuperados")
Totales("Fallecidos","Total Fallecidos")
```

```
Datosfinal['Real_infected'] = Datosfinal['Total']
- Datosfinal['Total_Recuperados']
- Datosfinal['Total_Fallecidos']
Datosfinal['total_recovered_or_dead'] = Datosfinal['Total_Recuperados'] +
Datosfinal['Total_Fallecidos']
```

A.2. Modelo SEIR modificado sin vacunación

```
def ode_model(z, t, beta, r):
    """
    Reference https://www.idmod.org/docs/hiv/model-seir.html
    """
    S, E, U, P, F, RU,RR,DU,DR = z
    N=S+E+U+P+F+RU+RR+DU+DR
    dSdt =  -beta*S/initN*(alphap*P+alphau*U+F)+Lambda-mu*S
    dEdt =  beta*S/initN*(alphap*P+alphau*U+F)-E/De - mu*E
    dUdt =  (1-r)*E/De-U/(beta1*Dr)-delta1*muc*U-mu*U
    dPdt =  r*(1-f)*E/De-P/Dr-muc*P-mu*P
    dFdt =  r*f*E/De-beta2*F/Dr-muc*F/delta2-mu*F
    dRUdt = U/(beta1*Dr)+beta2*F/Dr-mu*RU
    dRRdt = P/Dr-mu*RR
    dDUdt = delta1*muc*U+muc*F/delta2
    dDRdt = muc*P
    return [dSdt,dEdt, dUdt, dPdt, dFdt, dRUdt, dRRdt, dDUdt, dDRdt]

def ode_solver(t, initial_conditions, params):
    initE, initU, initP, initF, initRU, intRR,
    intDU, intDR, initN   = initial_conditions
    beta, r = params['beta'].value, params['r'].value
    initS = initN - (initE+initU+initP+initF+initRU+intRR+intDU+intDR)
    res = odeint(ode_model, [initS,initE, initU, initP, initF, initRU,
    initRR, initDU, initDR], t, args=(beta, r))
    return res

def error(params, initial_conditions, tspan, data):
    sol = ode_solver(tspan, initial_conditions, params)
    #return (sol[:, [3,6,8]] - data).ravel()
    return (sol[:, [3]] - data).ravel()
```

```
#data = Datosfinal.loc["2020-03-24 ":"2020-08-14 ",
['Real infected', 'Total Recuperados',
'Total Fallecidos']].values
data = Datosfinal.loc["2020-03-24 ":"2020-07-01 ",
['Real infected']].values
days = data.shape[0]
tspan = np.arange(0, days, 1)
tspan
plt.plot(tspan, data[:, 0], label='Infectados')
#plt.plot(tspan, data[:, 1], label='Recuperados')
#plt.plot(tspan, data[:, 2], label='Muertos')
plt.legend()
plt.show()


f=0.1
initN = Datosfinal["Total"].loc["2020-07-01"]*4/(0.9*(1-f))
print(initN)
# S0 = 966000000
initE = 15
initU = 0
initP = 1
initF = 0
initRU =0
initRR =0
initDU =0
initDR =0
initial_conditions = [initE, initU, initP, initF, initRU,
initRR, initDU, initDR, initN]

alphap = 0.5
alphau = 0.7
Lambda =0
mu=0
muc= 0.0562
De= 5.2
beta1= 0.6
beta2=0.7
Dr= 17.8
delta1= 0.3
```

```
delta2= 0.7
R0=4
r=0.2

params = Parameters()
params.add('beta', min=0.01, max=10)
params.add('r',min=0.01, max=0.999)

# fit model and find predicted values
result = minimize(error, params, args=(initial_conditions, tspan, data),
method= 'least_square' )
# result = minimize(error, params, args=(initial_conditions, tspan, data),
method='leastsq', \
# **{'xtol':1.e-15, 'ftol':1.e-15})
result.params

Nombre_parametros=[]
Parametros_estimados=[]
for i in result.params.items():
  Parametros_estimados.append(i[1].value)
  Nombre_parametros.append(i[0])
tspan2 = np.arange(0, 1*days, 1)

initS = initN     - (initE + initU + initP + initF + initRU + initRR
+ initDU + initDR)
Estimacion = odeint(ode_model, [ initS,initE, initU, initP, initF,
initRU,initRR, initDU, initDR], tspan2, args=(Parametros_estimados[0],
Parametros_estimados[1]))

#plot
plt.plot(tspan, data[:, 0],'o',label="datos Infectados" )
#plt.plot(tspan, Recuepradosmodelo1[:,0],'o',label="datos Recuperados" )
#plt.plot(tspan, data[:, 2],'o',label="datos Muertos" )
#plt.plot(tspan2, Estimacion[:, 0], label="Susceptibles" )
#plt.plot(tspan2, Estimacion[:, 2], label="Expuestos" )
plt.plot(tspan2, Estimacion[:, 2], label="Individuos Positivos
no Testeados" )
plt.plot(tspan2, Estimacion[:, 3], label="Testeados Positivos" )
plt.plot(tspan2, Estimacion[:, 4], label="Falsos Negativos" )
#plt.plot(tspan2, Estimacion[:, 6], label="Recuperados testeados
```

```
positivos" )
plt.title("Individuos infectados no testeados, falsos negativos
y individuos infectados testeados")
plt.xlabel("N mero de dias")
plt.ylabel("N mero de Infectados")
plt.legend()
plt.savefig("gdrive/MyDrive/ImagenesTesis/Modelosinvacuna.jpg")
plt.show()

print("Fitted MAE")
print('Infected: ', np.mean(np.abs(Estimacion[:, 3] - data[:, 0])))


print("\nFitted RMSE")
print('Infected: ', np.sqrt(np.mean((Estimacion[:, 3] - data[:, 0])**2)))

#calculo R_0

beta=Parametros_estimados[0]
r=Parametros_estimados[1]
print(beta)
print(r)
Fac1=alphau*(1-r)/((1/(beta1*Dr))+delta1*muc+ mu)
Fac2=alphap*r*(1-f)/((1/Dr)+muc+mu)
Fac3=r*f/((beta2/Dr)+muc/delta2+mu)
R0=beta/(mu*De+1)*(Fac1+Fac2+Fac3)
R0
```

A.3. Modelo SEIR modificado con vacunación

```
def ode_model_con_vacuna(z, t, beta,p, r):
    S, V, E, U, P, F, RU,RR,DU,DR = z
    N=S+V+E+U+P+F+RU+RR+DU+DR
    dSdt = -beta*S/initN*(alphap*P+alphau*U+F)+Lambda-mu*S-
    alpha_vac*S/initN*(alphap*P+alphau*U+F)
    dVdt = alpha_vac*S/initN*(alphap*P+alphau*U+F)-V*p*beta-mu*V
    dEdt = beta*S/initN*(alphap*P+alphau*U+F)+V*p*beta-E/De - mu*E
    dUdt = (1-r)*E/De-U/(beta1*Dr)-delta1*muc*U-mu*U
    dPdt = r*(1-f)*E/De-P/Dr-muc*P-mu*P
```

```
    dFdt =  r*f*E/De-beta2*F/Dr-muc*F/delta2-mu*F
    dRUdt = U/(beta1*Dr)+beta2*F/Dr-mu*RU
    dRRdt = P/Dr-mu*RR
    dDUdt = delta1*muc*U+muc*F/delta2
    dDRdt = muc*P
    return [dSdt,dVdt,dEdt, dUdt, dPdt, dFdt, dRUdt, dRRdt, dDUdt, dDRdt]

def ode_solver_vacuna(t, initial_conditions, params):
    initV, initE, initU, initP, initF, initRU, intRR, intDU,
    intDR, initN   = initial_conditions
    beta, p,r = params['beta'].value, params['p'].value,
    params['r'].value
    initS = initN - (initE+initU+initP+initF+initRU+intRR+intDU
    +intDR+initV)
    res = odeint(ode_model_con_vacuna, [initS, initV ,initE, initU,
     initP,initF, initRU, initRR, initDU, initDR], t, args=(beta, p,r))
    return res

def error_vacuna(params, initial_conditions, tspan, data):
    sol = ode_solver_vacuna(tspan, initial_conditions, params)
    #return (sol[:, [3,6,8]] - data).ravel()
    return (sol[:, [4]] - data).ravel()

#data = Datosfinal.loc["2020-03-24  ":"2020-08-14   ", ['Real infected',
'Total Recuperados', 'Total Fallecidos']].values
desde,hasta="2021-04-01","2021-07-15"
data = Datosfinal.loc[desde:hasta, ['Real infected']].values
days = data.shape[0]
tspan = np.arange(0, days, 1)
tspan
plt.plot(tspan, data[:, 0], label='Infectados')
#plt.plot(tspan, data[:, 1], label='Recuperados')
#plt.plot(tspan, data[:, 2], label='Muertos')
plt.legend()
plt.show()


f=0.1
initN = (Datosfinal["Total"].loc[hasta]-Datosfinal["Total"].loc[desde])
*4/(0.5*(1-f))
```

```
initV = 0.1*initN
print(initN)
# S0 = 966000000
initE = 150#*Datosfinal['Real infected'].loc[desde]
initU = Datosfinal['Real infected'].loc[desde]
initP = Datosfinal['Real infected'].loc[desde]
initF = Datosfinal['Real infected'].loc[desde]*0.10
initRU =0
initRR =0#Datosfinal['Total Recuperados'].loc[desde]
initDU =0
initDR =Datosfinal['Total Fallecidos'].loc[desde]
initial_conditions = [initV, initE, initU, initP, initF, initRU,
initRR, initDU, initDR, initN]

alpha_vac=0.15
#p=0.1
alphap = 0.5
alphau = 0.7
Lambda = 3.95*10**(-5)
mu= 3.95*10**(-5)
Lambda =0
mu=0
muc= 0.0562
De= 5.2
beta1= 0.6
beta2=0.7
Dr= 17.8
delta1= 0.3
delta2= 0.7
R0=4
r=0.2

params = Parameters()
params.add('beta', min=0.01, max=10)
params.add('p', min=0.01, max=0.5)
params.add('r',min=0.001, max=0.999)

# fit model and find predicted values
result_vacuna = minimize(error_vacuna, params,
    args=(initial_conditions, tspan, data), method= 'least_square' )
```

```
result_vacuna.params

Nombre_parametros=[]
Parametros_estimados=[]
for i in result_vacuna.params.items():
  Parametros_estimados.append(i[1].value)
  Nombre_parametros.append(i[0])
tspan2 = np.arange(0, 1*days, 1)

initS = initN     - (initE + initU + initP + initF + initRU + initRR
+ initDU + initDR+initV)
Estimacion = odeint(ode_model_con_vacuna, [ initS,initV,initE, initU,
initP, initF, initRU, initRR, initDU, initDR], tspan2,
args=(Parametros_estimados[0], Parametros_estimados[1],
Parametros_estimados[2]))

#Plot
plt.plot(indef, data[:, 0],'o',label="datos Infectados" )
plt.plot(indef, Estimacion[:, 3], label="No Testeados Positivos" )
plt.plot(indef, Estimacion[:, 4], label="Positivos" )
plt.plot(indef, Estimacion[:, 5], label="Palsos negativos" )
plt.title("Individuos infectados no testeados, falsos negativos
y individuos infectados testeados entre 01-04-2021 y 15-07-2021")
plt.xlabel("Fecha")
plt.ylabel("N mero de Infectados")
plt.legend()
plt.savefig("gdrive/MyDrive/ImagenesTesis/Modelodeterministicovacuna.jpg")
plt.show()
```

A.4. Modelo SEIR modificado estocastico con vacunación

```
import matplotlib.pyplot as plt
import numpy as np
import sdeint


beta=Parametros_estimados[0]
p=Parametros_estimados[1]
r=Parametros_estimados[2]
```

```
def ff(y, t):
  S, V, E, U, P, F, RU,RR,DU,DR = y
  N=S+V+E+U+P+F+RU+RR+DU+DR
  dSdt =  -beta*S/initN*(alphap*P+alphau*U+F)+Lambda-mu*S-
            alpha_vac*S/initN
  dVdt = alpha_vac*S/initN-V*p*beta*(alphap*P+alphau*U+F)/initN-mu*V
  dEdt =  beta*S/initN*(alphap*P+alphau*U+F)+V*p*beta*(alphap*P+
       alphau*U+F)/initN-E/De - mu*E
  dUdt =  (1-r)*E/De-U/(beta1*Dr)-delta1*muc*U-mu*U
  dPdt =  r*(1-f)*E/De-P/Dr-muc*P-mu*P
  dFdt =  r*f*E/De-beta2*F/Dr-muc*F/delta2-mu*F
  dRUdt = U/(beta1*Dr)+beta2*F/Dr-mu*RU
  dRRdt = P/Dr-mu*RR
  dDUdt = delta1*muc*U+muc*F/delta2
  dDRdt = muc*P
  return np.array([dSdt,dVdt,dEdt, dUdt, dPdt, dFdt, dRUdt, dRRdt,
                   dDUdt, dDRdt])


def GG(y, t):
  S, V, E, U, P, F, RU,RR,DU,DR = y
  return np.array([[-0.15*S/initN*(alphap*P+alphau*U+F), 0, 0,0,0,0,
  0,0,0,0],
                [-0.15*p*V/initN*(alphap*P+alphau*U+F), 0, 0,0,0,0,0,
                0,0,0],
                [0.15*(p*V+S)/initN*(alphap*P+alphau*U+F),0,0,0,0,0,0,
                0,0,0],
                [0,0,0,0,0,0,0,0,0,0],
                [0,0,0,0,0,0,0,0,0,0],
                [0,0,0,0,0,0,0,0,0,0],
                [0,0,0,0,0,0,0,0,0,0],
                [0,0,0,0,0,0,0,0,0,0],
                [0,0,0,0,0,0,0,0,0,0],
                [0,0,0,0,0,0,0,0,0,0]])

result = sdeint.itoEuler(ff, GG, np.array([ initS,initV,initE, initU,
initP, initF, initRU, initRR, initDU, initDR]), tspan2)
print(result)
plt.plot(tspan, data[:, 0],'o',label="datos Infectados" )
plt.plot(tspan2,result[:,3], label="No Testeados Positivos" )
```

```
plt.plot(tspan2,result[:,4], label="Positivos" )
plt.plot(tspan2,result[:,5],label="Palsos negativos"  )
#plt.plot(tspan,result[:,3])
plt.show()

print("Fitted MAE")
print('Infected: ', np.mean(np.abs(result[:,4] - data[:, 0])))


print("\nFitted RMSE")
print('Infected: ', np.sqrt(np.mean((result[:,4] - data[:, 0])**2)))

Graficas_estocasticas=[]
for i in range(10000):
  result2 = sdeint.itoEuler(ff, GG, np.array([ initS,initV,initE,
  initU, initP, initF, initRU, initRR, initDU, initDR]), tspan2)
  error=np.mean(np.abs(result2[:,4] - data[:, 0]))
  if error < 230:
    Graficas_estocasticas.append(result2)

print(len(Graficas_estocasticas))

indef=Datosfinal.loc[desde:hasta, ['Real infected']].index
for i in Graficas_estocasticas:
  plt.plot(indef,i[:,3],color="blue" )
  plt.plot(indef,i[:,4],color="orange" )
  plt.plot(indef,i[:,5],color="green"  )
plt.plot(indef, data[:, 0],'o',label="datos Infectados",color="red")
plt.plot(indef,i[:,3], label="No Testeados Positivos",color="blue" )
plt.plot(indef,i[:,4], label="Positivos",color="orange" )
plt.plot(indef,i[:,5],label="Palsos negativos",color="green"  )
plt.title("Individuos infectados no testeados, falsos negativos
y individuos infectados testeados con f={}".format(f))
plt.xlabel("Fecha")
plt.ylabel("N mero de Infectados")
plt.legend()
plt.savefig("gdrive/MyDrive/ImagenesTesis/Modeloestocasticovacuna
conf={}.jpg".format(f))
plt.show()
```

Bibliografía

[Allen, 2010] Allen, L. J. (2010). *An introduction to stochastic processes with applications to biology*. CRC press.

[Atkinson et al., 2011] Atkinson, K., Han, W., and Stewart, D. E. (2011). *Numerical solution of ordinary differential equations*. John Wiley & Sons.

[Bastin, 2018] Bastin, G. (2018). Lectures on mathematical modelling of biological systems. *GBIO2060*.

[Bhattacharyya et al., 2021] Bhattacharyya, R., Kundu, R., Bhaduri, R., Ray, D., Beesley, L. J., Salvatore, M., and Mukherjee, B. (2021). Incorporating false negative tests in epidemiological models for SARS-Cov-2 transmission and reconciling with seroprevalence estimates. *Scientific reports*, 11(1):1–14.

[Blanco, 2019] Blanco, L. (2019). *Probabilidad*. Univ. Nacional de Colombia.

[Blanco et al., 2012] Blanco, L., Arunachalam, V., and Dharmaraja, S. (2012). *Introduction to probability and stochastic processes with applications*. John Wiley & Sons.

[Cooper et al., 2020] Cooper, I., Mondal, A., and Antonopoulos, C. G. (2020). A SIR model assumption for the spread of covid-19 in different communities. *Chaos, Solitons & Fractals*, 139:110057.

[Culqui Sánchez et al., 2020] Culqui Sánchez, M., Nasimba Quinatoa, J., and Chiliquinga Calderón, E. (2020). Aplicación del modelo matemático SEIR en la pandemia por Covid-19, relevancia en salud pública. *Vive Revista de Salud*, 3(9):275–290.

[Datos Abiertos, 2022] Datos Abiertos, I. (2022). Casos positivos de Covid-19 en colombia.

[Diekmann et al., 2010] Diekmann, O., Heesterbeek, J., and Roberts, M. G. (2010). The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society Interface*, 7(47):873–885.

[Hairer et al., 1993] Hairer, E., Wanner, G., and Nørsett, S. P. (1993). Classical Mathematical Theory. *Solving Ordinary Differential Equations I: Nonstiff Problems*, pages 1–128.

[Heffernan et al., 2005] Heffernan, J. M., Smith, R. J., and Wahl, L. M. (2005). Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*, 2(4):281–293.

[INS, 2022a] INS (2022a). Covid-19 en colombia.

[INS, 2022b] INS (2022b). Reseña histórica.

[Karatzas and Shreve, 2012] Karatzas, I. and Shreve, S. (2012). *Brownian Motion and Stochastic Calculus*, volume 113. Springer Science & Business Media.

[Kuo, 2006] Kuo, H.-H. (2006). *Introduction to Stochastic Integration*. Springer.

[Niño-Torres et al., 2022] Niño-Torres, D., Ríos-Gutiérrez, A., Arunachalam, V., Ohajunwa, C., and Seshaiyer, P. (2022). Stochastic modeling, analysis, and simulation of the covid-19 pandemic with explicit behavioral changes in bogotá: A case study. *Infectious Disease Modelling*, 7(1):199–211.

[Øksendal, 2003] Øksendal, B. (2003). Stochastic Differential Equations. In *Stochastic differential equations*, pages 65–84. Springer.

[Ridenhour et al., 2018] Ridenhour, B., Kowalik, J. M., and Shay, D. K. (2018). El número reproductivo básico (R0) consideraciones para su aplicación en la salud póblica. *American Journal of Public Health*, 108(S6):S455–S465.

[Ríos-Gutiérrez et al., 2021] Ríos-Gutiérrez, A., Torres, S., and Arunachalam, V. (2021). Studies on the basic reproduction number in stochastic epidemic models with random perturbations. *Advances in difference equations*, 2021(1):1–24.

[Ríos Gutiérrez, 2019] Ríos Gutiérrez, A. S. (2019). Modelos epidemiológicos estocásticos y su inferencia: casos sis y seir. *Universidad Nacional de Colombia*.

[Rivera-Rodriguez and Urdinola, 2020] Rivera-Rodriguez, C. and Urdinola, B. P. (2020). Predicting hospital demand during the Covid-19 outbreak in Bogota, Colombia. *Frontiers in Public Health*, 8:710.

[Tornatore et al., 2006] Tornatore, E., Buccellato, S. M., and Vetro, P. (2006). On a stochastic disease model with vaccination. *Rendiconti del Circolo Matematico di Palermo*, 55(2):223–240.

[Uribarri et al., 2013] Uribarri, S. G., Meza, M. A. R., and Cota, J. L. C. (2013). Las matemáticas de las epidemias: caso México 2009 y otros. *CIENCIA ergo-sum, Revista Científica Multidisciplinaria de Prospectiva*, 20(3):238–246.

[Velasco-Hernandez, 2021] Velasco-Hernandez, J. X. (2021). Modelling epidemics: a perspective on mathematical models and their use in the present SARS-Cov-2 epidemic. *arXiv preprint arXiv:2104.07793*.

[Witbooi et al., 2015] Witbooi, P. J., Muller, G. E., and Van Schalkwyk, G. J. (2015). Vaccination control in a stochastic SVIR epidemic model. *Computational and mathematical methods in medicine*, 2015.

[Woloshin et al., 2020] Woloshin, S., Patel, N., and Kesselheim, A. S. (2020). False negative tests for SARS-Cov-2 infection—challenges and implications. *New England Journal of Medicine*, 383(6):e38.