



DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

28 de junio de 2021

2do Cuatrimestre de 2019

Integrante	LU	Correo electrónico
Julian Len	467/14	julianlen@gmail.com
Tomas Jaratz	59/15	tjaratz@gmail.com



Facultad de Ciencias Exactas y Naturales

Universidad de Buenos Aires

Ciudad Universitaria - (Pabellón I/Planta Baja)

Intendente Güiraldes 2160 - C1428EGA

Ciudad Autónoma de Buenos Aires - Rep. Argentina

Tel/Fax: (54 11) 4576-3359

<http://www.fcen.uba.ar>

Índice

1. Introducción y motivación	2
2. Preliminares	3
2.1. Grafo	3
2.2. Modelos de epidemias	3
2.3. Conceptos estadísticos	5
3. Software	6
3.1. Topologías de Redes Implementadas	8
3.1.1. Random regular network	8
3.1.2. Lattices bidimensionales	8
3.1.3. Small World network	9
3.1.4. Scale Free network	9
3.1.5. Balanced Tree network	10
4. Experimentos	12
4.1. Experimento 1: El modelo de epidemia es determinante	12
4.2. Experimento 2	20
4.2.1. GRID	20
4.2.2. Random	23
4.2.3. Scale Free	26
4.2.4. Small World	30
4.2.5. Balanced Tree	32
4.3. Experimento 3	34
4.4. La cantidad de contagios y muertes efectivamente es diferente	35
5. Trabajos Futuros	37
6. Conclusión	38
6.1. Apéndice 1	40

1. Introducción y motivación

En el estudio de epidemias, múltiples temas asociados a la biología se entremezclan con temas sociales. Cuando hablamos de epidemias, nos referimos a enfermedades contagiosas causadas por patógenos biológicos (gripe, lepra, ETS) que se propagan de persona a persona. Las epidemias pueden moverse rápidamente a través de la población o pueden persistir durante largo tiempo a bajo nivel, de un día para otro la epidemia puede crecer en números de casos o tener olas con patrones cíclicos de crecimiento y decrecimiento de prevalencia. La expansión de una epidemia a través de grupos de personas, es determinada no sólo por las características intrínsecas del virus, si no también por la estructura de la red de personas por la cuál se expande, ya que la oportunidad de expansión de un virus es a través del contacto entre personas. Esto nos lleva a pensar que la red debajo de todo esto es crucial para entender cómo se comportan las epidemias.

El patógeno y la red están constantemente interrelacionadas, inclusive con la misma población, para dos epidemias distintas, la red de contactos puede tener estructuras completamente diferentes, dependiendo el modo de transmisión del virus. Por ejemplo para una enfermedad que se contagia a través del aire, por cada persona, los cantidad de contactos se consideran mucho mayor que si fuera solo a través de un contacto físico. [9]

En Enero del 2019 se presenta una nueva enfermedad en Wuhan ciudad de China, causada por un nuevo tipo de coronavirus Cov-SARS-19 convirtiéndose rápidamente en pandemia. Poco se sabía del virus, y en el correr del tiempo se lo conoció cada vez más en cuestiones como, su índice de reproducción, índice de fatalidad y lo más importante, su alto nivel de contagiosidad causada en mayor medida por asintomáticos. Muchas estrategias se implementaron alrededor del Mundo, cuarentenas estrictas, cuarentenas segmentadas por edad (pues el ratio de mortalidad era bajo para jóvenes), etc. En paralelo se presentaron dos estrategias para contener el virus, (1) aplanar la curva de contagios permitiendo responder al Sistema de Salud, (2) cortar la curva de nuevos casos haciendo desaparecer la circulación del virus del territorio. Para lograr esto se implementaron pruebas o tests para conocer si una persona portaba el virus, y si esta era portadora se buscaba y aislaba los contactos estrechos de hasta 48 a 72 horas previo a presentar síntomas. En la carrera por combatir al virus, para anticiparse a su comportamiento o tomar decisiones acertadas, se utilizaron diversos sistemas de simulación complejos [17]. Los mismos permiten, utilizando modelos de epidemia, y parámetros acorde a la población en estudio, mostrar el comportamiento del virus a lo largo de la epidemia, mejorando la precisión a medida que se conoce más información como por ejemplo, ratio de hospitalización, población de riesgo, etc.

En el siguiente trabajo presentaremos un simulador de epidemias, el cual dado una una red de personas y una cantidad de tiempo permite simular el desarrollo de una epidemia. En particular utilizaremos los parámetros conocidos del Cov-SARS-19. De esta manera, buscamos responder cuánto modifica la transmisión y velocidad de propagación de un virus factores como, la conectividad del grafo, el grupo de personas inicial, el modelo de epidemia o la probabilidad de contagio de un nodo (o persona) a otro.

El resto del trabajo práctico estará dividido en las siguientes secciones, en la Sección Preliminares 2 abordaremos la teoría detrás de conceptos que utilizaremos a lo largo del trabajo práctico. Luego en la sección de Software 3 contaremos acerca de la implementación del mismo, qué funciones provee y cómo genera la simulación. Finalmente en la sección de Experimentos 4 nos plantearemos ciertas hipótesis para simular, utilizando los parámetros mencionados, el comportamiento del virus en diferentes redes con distintas características.

2. Preliminares

2.1. Grafo

Un *grafo* es una estructura matemática que puede representar relaciones entre un conjunto de elementos. Un grafo se representa como $G = (V, E)$ donde V es un conjunto de nodos o vértices y E un de ejes o relaciones entre nodos. Si v_1, v_2 son dos vértices, de haber un eje e_1 que los conecte, se denota como $e_1 = (v_1, v_2)$ o (v_2, v_1) . En particular si existe un eje e_k tal que $e_k = (v_i, v_j)$, entonces v_i y v_j se consideran *vecinos*. Existen dos tipos de grafo *dirigido* o *dígrafo* y *no-dirigido* o *grafo*. En el dígrafo los ejes pasan a tener dirección, por lo tanto entre dos vértices v_1 y v_2 pueden existir dos ejes $e_1 = (v_1, v_2)$ y $e_2 = (v_2, v_1)$. Mientras que en el grafo, $e_1 == e_2$. Por último, definimos el *grado* de un nodo, como la cantidad de vecinos. En el caso de los digrafos, se divide el grado en *in-degree* la cantidad de ejes que llegan a un nodo y *out-degree* los ejes que salen de un nodo.

Luego en el mundo de los grafos podemos diferenciar ciertas clases de grafos distinguidas por sus propiedades. Para esto primero definimos un *camino* en un grafo como una secuencia de nodos vecinos. Dos vértices están *conectados* si existe un camino de uno al otro. En particular se llama *camino simple*, al camino del grafo que no repite nodos en su secuencia. Luego definimos *ciclo* al camino simple que empieza y termina en el mismo nodo. Por otro lado, se puede distinguir la conectividad de un grafo. Se dice que el grafo es *conexo* si existe un camino entre todo par de nodos. Podemos hablar de grafos más (o menos) conexos según la cantidad de caminos entre dos pares de vértices cualquiera. Por lo tanto, un grafo es de tipo *árbol* si es conexo y no admite ciclos. El árbol tiene la propiedad de contar con un nodo *raíz* y por cada vecino otro subárbol. Finalmente, otra propiedad que puede caracterizar un grafo es el coeficiente de *clustering*, indica la probabilidad que dado un nodo, dos vecinos seleccionados aleatoriamente también sean vecinos (por ejemplo, 1 indica que todos los vecinos del nodo son vecinos). Se dice *cluster* a aquellos nodos más conectados entre sí que con el resto, en particular estos nodos presentarán un coeficiente de clustering superior.

Tal como explican en [9], un grafo es útil para representar una red. Por ejemplo, un grafo puede representar una red de computadoras, una red de sitios web conectados entre sí a través de hipervínculos. También se puede utilizar un grafo para representar una abstracción de una red de contactos entre personas, donde cada nodo representa a un individuo o a un grupo de personas, mientras que un eje puede representar el contacto entre ellas.

2.2. Modelos de epidemias

Para los modelos de epidemias, tomamos un grafo como red subyacente por donde se expande el virus. En esta red, sea A y B nodos del grafo tal que B es vecino de A, A está infectado por el virus y sea p la probabilidad de que A pueda contagiar a B. Esta probabilidad p se representa en el grafo como el *peso* de la arista que conecta los nodos A y B. Luego todo nodo en el gráfico que esté infectado, contagiará a sus vecinos con la probabilidad p representada en el eje que conecta a este con cada uno de sus vecinos.

Para empezar a hablar de modelos de epidemia, vamos a introducir primero el caso más simple. El tipo de grafo que usaremos para este caso es de tipo Árbol y el modelo de contagio se denomina *Branching Process* y funciona de la siguiente manera:

1. Primer ola: Suponemos un nodo contagiado, denominado *nodo raíz*, que transmite la enfermedad con probabilidad p a cada nodo vecino. Luego suponemos que la raíz se conecta con k nodos. Estos k nodos pertenecen a la primer ola de la epidemia, algunos se contagiarán y otros no.
2. Segunda ola: Luego si un nodo de la primer ola se conecta con otros k nodos distintos, resulta en una segunda ola de $k * k = k^2$ nodos. Cada infectado de la primer ola podrá contagiar a sus k vecinos de la segunda ola con probabilidad p .
3. Ola subsecuente: Luego cada nodo de cada ola contagiará otros k nodos nuevos (con nuevos nos referimos a no perteneciente a ninguna ola anterior) y distintos (dos nodos de la misma ola no contagian al mismo nodo nuevo).

Un ejemplo que muestra el Branching Process tomado de [9] con $k=3$

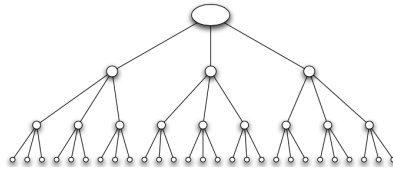


Figura 1: Grafo sobre el cual se hace branching process

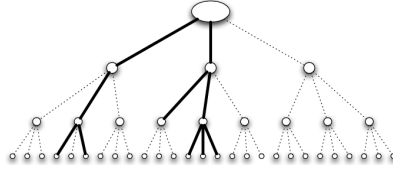


Figura 2: Branching process con la probabilidad de contagio p alta

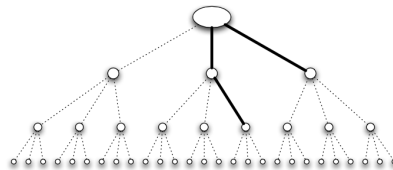


Figura 3: Branching process con la probabilidad de contagio p baja

Luego observamos que el comportamiento de la epidemia en este tipo de modelos, puede por un lado ser agresiva, como en el segundo caso, donde las olas se siguen expandiendo, aunque también pueden ser epidemias leves como en el tercer caso, donde no pasa de la segunda ola y la epidemia se termina.

Generalizando, podemos definir modelos de epidemias que pueden ser aplicados a cualquier tipo de grafo y no únicamente a un árbol. Además se le agrega el factor tiempo, pues cada uno de los nodos puede evolucionar en distintos estados en el transcurso del tiempo de la epidemia.

Susceptible-Infectado-Removido o **SIR** es uno de los modelos de epidemia. En este caso, las variables que entran en juego son el tiempo (t_i) y la probabilidad de infección p . Luego las condiciones del modelo son:

- Inicialmente algunos nodos están en estado infectado (I) mientras que otros están susceptibles (S).
- Cada nodo queda en I por cierto tiempo t_i , pudiendo contagiar con cierta probabilidad p .
- Cuando transcurre t_i tiempo, el nodo queda en estado recuperado (R) el cual no puede contagiar ni infectarse nuevamente.

Luego existen dos modelos que surgen del modelo SIR. El modelo **Susceptible-Infectado-Recuperado-Susceptible** o **SIRS** y **Susceptible-Infectado-Susceptible** o **SIS**. Para el primero existe un tiempo t_r en el cual una vez transcurrido, el nodo pasa de estado R a estado S. Mientras que en el modelo SIS, no genera ningún tipo de inmunidad en el nodo, permitiendo que el nodo luego de t_i vuelva a estar susceptible de infectarse.

2.3. Conceptos estadísticos

Dentro de las dificultades que se nos presentaron a la hora de modelar la epidemia, nos encontramos primero con el interrogante de cómo calcular la probabilidad de contagio de un nodo susceptible en base al estado de sus nodos vecinos. Con esto nos referimos a que si por ejemplo, el nodo A se encuentra susceptible y tiene como vecinos a los nodos B, C, D y E de los cuales solo B y C están infectados, debemos determinar qué probabilidad de contagiarse el virus tiene A en cada iteración. Para esto decidimos utilizar el ratio de nodos infectados sobre el total de vecinos para estimar dicha probabilidad de contagio, por lo que en nuestro ejemplo anterior nuestro nodo A va a tener un 50 % de probabilidad de contagiarse el virus en cada iteración hasta que B y/o C se recuperen o D y/o E se contagien, aumentando o disminuyendo dicha probabilidad. Mientras que si ningún vecino se encuentra infectado el nodo A no se contagiará, pero si todos se encuentran infectados lo hará con 100 % de probabilidad. Por lo que podemos definir la probabilidad de contagiarse de un nodo n en la iteración i como:

$$P(n)_i = (\#vecinos_infectados)_i / (\#vecinos_totales)_i$$

Donde $\#vecinos_totales_i$ equivale a la cantidad de los vecinos del nodo n que se encuentran vivos en la iteración i , por lo que si un nodo muere no será tomado en cuenta como vecino de ningún otro en las siguientes iteraciones.

Por otro lado, podemos conocer cual es la probabilidad de una persona, una vez que contrajo el virus, de fallecer. Supongamos por ejemplo que dicha probabilidad es p , esta p representa la probabilidad de que una persona muera en el transcurso de toda la infección, no en un día en particular. Es por esto que si en cada iteración calculamos la probabilidad de morir de un nodo infectado como p , la probabilidad acumulada en el transcurso de 14 o 21 iteraciones sería mucho mayor, ya que el nodo puede morir en cualquier iteración mientras esté infectado.

Para esto necesitamos encontrar una distribución de probabilidad adecuada a este fenómeno que nos permita encontrar, conociendo la probabilidad p de morir en el transcurso de los n días que dura la enfermedad, la probabilidad p' de morir en cada día que el contagiado está en estado grave. Luego, por cada día si nuestra probabilidad de morir es p , el nodo muere o no, es decir, el nodo murió en el primer día, o bien no murió en el primero y si en el segundo, o bien no murió en los primeros dos días y sí en el tercero y así sucesivamente hasta el día n . Matemáticamente lo podemos representar con la siguiente sumatoria:

$$p = \sum_{i=0}^{n-1} p'(1-p')^i$$

Donde p representa la probabilidad de muerte dado que el nodo se infectó, n representa la cantidad de días que dura la enfermedad y p' la probabilidad de morir en un día cualquiera.

De esta forma conociendo p y n , el único paso restante para encontrar p' consiste en despejar el valor de la ecuación, lo cual se puede hacer de la siguiente forma:

$$p' = 1 - e^{\frac{\ln(1-p)}{n}}$$

Cuya demostración estará en la Demostración 1 en el apéndice 16.1. Luego utilizando esta metodología aproximamos la probabilidad de morir en un día aislado si el nodo se infecta para así poder ver la evolución en el transcurso de cada iteración del modelo.

3. Software

Para implementar los distintos modelos de redes [8] como estructuras basadas en grafos, utilizamos la librería NetworkX [networkx], la cual brinda una interfaz para instanciar dichas redes variando los atributos necesarios para su configuración. Luego utilizamos estos grafos como base de nuestro modelo, donde los nodos representan personas y los ejes que los unen representan contacto directo entre dichas personas. Una particularidad que es necesario destacar es que las redes que creamos son estáticas y su estructura no se ve modificada durante la simulación, lo que implica que en el grafo si dos nodos son vecinos, lo serán siempre.

La librería NetworkX permite agregarle atributos a los nodos. Para representar los modelos SIS, SIR y SIRS implementamos los estados que representan el momento de la infección en que se encuentra cada nodo y su potencial transición hacia los siguientes estados, con la salvedad de que introducimos un estado nuevo para representar el hecho de que un nodo pueda morir y subdividimos el estado infectado en tres subestados, *incubando*, *infectado leve* e *infectado grave*. Esto último con la intención de obtener más detalle en el análisis y poder diferenciar entre las distintas formas de evolución y manifestación de la enfermedad. Por otro lado, un nodo puede pertenecer a la población de riesgo. Los nodos de riesgo tendrán mas posibilidades de desarrollar un cuadro clínico grave, y solo los nodos que manifiesten este cuadro estarán en riesgo de morir, mientras que aquellos que transiten la infección levemente no van a correr este riesgo. Otra de las consideraciones que tomamos es que de los nodos que se infecten gravemente que sean de riesgo tendrán mayor probabilidad de morir que los demás. A continuación presentamos dichos estados y sus transiciones.

- **Susceptible:** el nodo no se encuentra infectado y puede contraer la enfermedad. En cada iteración se calcula la probabilidad de contagiarse como $\frac{\#vecinos_infectados}{\#vecinos_totales}$ y en caso de contagio el nodo pasa a incubar el virus.
- **Incubando:** el nodo puede contagiar a otros nodos pero no presenta síntomas, una vez que el tiempo de incubación culmina, el nodo puede pasar a estado *infectado grave* o *infectado leve* con cierta probabilidad que depende de si es un nodo de riesgo o no.
- **Infectado leve:** el nodo podrá contagiar a otros y al cabo de de un tiempo pasará al estado *recuperado* o *susceptible*, dependiendo del modelo. No incluimos en el simulador la opción de evolucionar a una infección grave a modo de simplificación.
- **Infectado grave:** el nodo transitará por cierta cantidad de tiempo en este estado con la posibilidad de morir en cada iteración de la simulación. El nodo luego podrá transitar al estado *muerto*, *recuperado* o *susceptible* dependiendo del modelo.
- **Muerto:** el nodo permanece en este estado hasta terminar la simulación y no se lo tiene más en cuenta como vecino de ningún otro nodo.
- **Recuperado:** el nodo queda en este estado hasta terminar la simulación o por un tiempo determinado y luego pasa a *susceptible*, dependiendo del modelo, mientras se encuentra en este estado no se puede contagiar nuevamente.

Figura 4: En la figura podemos observar las máquinas de estado que representan la transición de estados en cada modelo. De izquierda a derecha: SIS, SIR y SIRS

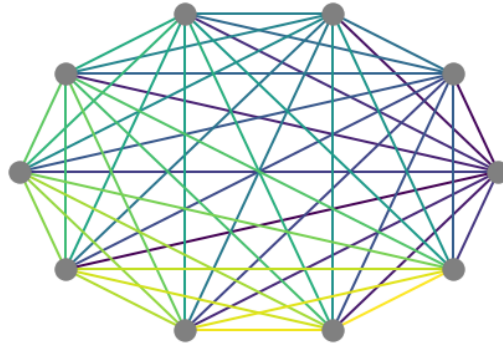


La simulación comienza siempre con un porcentaje de nodos elegidos aleatoriamente incubando y el resto en estado susceptible, luego, en cada iteración, que representa el paso de una unidad de tiempo(día, mes, año, etc.), se vuelve a computar el estado de cada nodo y se lo actualiza, escribiendo finalmente a un archivo *csv* el estado en que se encuentra la red. Con estado de la red nos referimos particularmente al porcentaje de nodos que se encuentran en cada estado y la cantidad de nodos que fueron infectados en total. La cantidad de iteraciones totales es un parámetro que recibe el programa antes de correr la simulación. Finalmente podemos utilizar el archivo de salida para realizar los análisis que consideremos pertinentes.

3.1. Topologías de Redes Implementadas

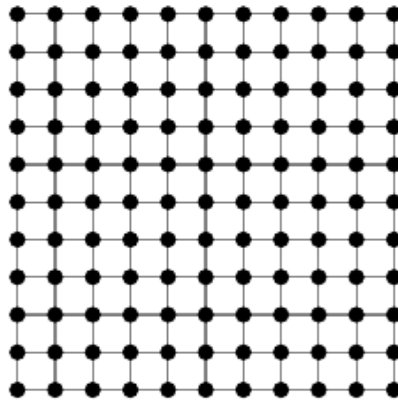
Dado que existen una gran variedad de grafos diferentes, en la literatura [12] existen algunos en particular que representan ciertas configuraciones de una red (por ejemplo, la red de referencias entre páginas web, la red de cómo las personas se conocen entre sí, etc.). Cada una de estas configuraciones se construyen a partir de ciertos datos como el grado, la conectividad entre nodos, si hay nodos con mayor grado que otros o coeficientes como el de clustering. A continuación presentamos las características esenciales de cada una de ellas junto con los parámetros necesarios para su generación

3.1.1. Random regular network



Se caracteriza por estar formada por nodos conectados aleatoriamente entre sí siguiendo una distribución uniforme. Todos los nodos tienen exactamente el mismo grado y se destacan por no presentar propiedades de clustering, es decir cada vértice cuenta con un bajo índice de clustering por lo tanto, no se generan grupos distinguibles de nodos conectados entre sí y desconectados del resto, por lo que la distribución suele ser altamente homogénea a nivel individual, por otro lado al ser aleatorias las conexiones, suele haber caminos cortos entre pares de nodos elegidos al azar. Luego, dado los parámetros n que representa cantidad de nodos del grafo y d el grado de cada uno de estos nodos, construimos el grafo siguiendo el algoritmo presentado en [16] resultando en un d -grafo regular, es decir un grafo donde todos los nodos tienen d vecinos.

3.1.2. Lattices bidimensionales

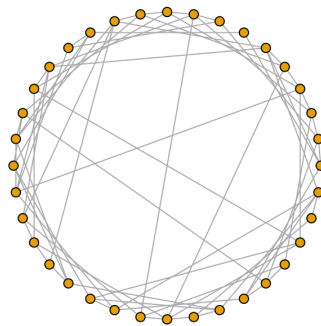


También conocida como Grid o Mesh, se caracteriza por disponer a los nodos en forma de grilla, usualmente de dos dimensiones y luego crear ejes solamente entre los nodos adyacentes. Debido a la manera en que se conectan los nodos, en modo de cuadrilla, el coeficiente de clustering es alto. Otras características de este tipo de grafos, la distancia de un nodo a otro suele ser mayor que en las redes

aleatorias debido a que no existen saltos grandes entre nodos, pues cada nodo solo está conectado a sus nodos adyacentes. Luego, a la hora de medir una epidemia, el crecimiento inicial será más lento que en otros tipos de redes. Otra característica de las redes en forma de cuadrilla o Lattices, es la forma de expansión de una epidemia, empezando por un nodo semilla o 'seed' y luego, expandiéndose en olas circulares a lo largo de la cuadrilla.

Este tipo de redes se suele utilizar en simulaciones por computadora de procesos de contactos e incendio forestales. El primero suele ser una abstracción del modelo SIS (caracterizado por β y γ) y el segundo por modelo SIR (el árbol se quema, deja un sitio vacío y puede renacer otro nuevo árbol en ese sitio). Sin embargo para el modelo de epidemias, este tipo de redes pueden simular por ejemplo, zonas urbanas altamente conectadas o pobladas. Finalmente, dado los parámetros b y h creamos una cuadrilla con $b * h$ nodos.

3.1.3. Small World network



Si uno piensa a las relaciones personales entre las personas como un grafo donde los nodos son personas y los ejes indican que se conocen entre ellas, el fenómeno de small-world describe que entre dos nodos cualesquiera existe un camino de a lo sumo 6 nodos de largo. Es decir, para llegar de una persona a otra cualquiera, se puede llegar en a lo sumo 6 personas como intermedias.

En particular, las redes de mundo pequeño tienen la propiedad de presentar altos niveles de clustering, como las redes Lattice, pero a diferencia de estas, la distancia promedio entre los nodos suele ser mucho menor. En general si tenemos una red con n nodos, la distancia promedio l entre todo par de nodos, suele estar acotada por el logaritmo de n , por lo que $l < \log(n)$.

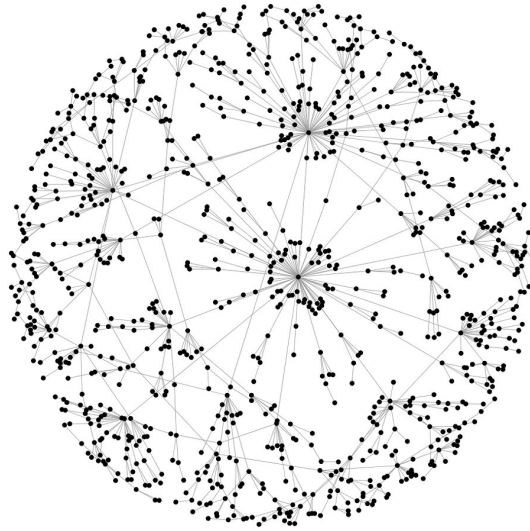
Este tipo de redes se obtiene formando un anillo de n nodos y conectando a cada uno con sus k vecinos más cercanos dentro del anillo, luego se crean "atajos", es decir, sean u y v nodos en el anillo, se reemplaza el eje (u,v) con probabilidad p por el eje (u,w) con w elegido al azar de manera uniforme.

Luego debido a su alto clustering, en las epidemias los contagios suelen ser mayormente locales dentro del cluster, aunque la transmisión entre distintos clusters también es posible. Pero debido a la gran cantidad de caminos cortos dentro de cada cluster, es que las epidemias en él suelen expandirse rápidamente.

Para crear este tipo de redes, dado n la cantidad de nodos, k para los k -vecinos más cercanos y p , la probabilidad de reemplazar un eje por otro, es que se construye la red de Small World.

3.1.4. Scale Free network

Uno de los parámetros interesantes en las redes, es la distribución del grado en los nodos. En el caso de redes Small-world, grid o random los grados suelen ser homogéneos, pero esto difiere de la realidad en la medida que, en las redes de personas suelen haber muchos nodos con pocas conexiones y unos pocos con muchas más conexiones que el resto, también denominados super-spreaders.

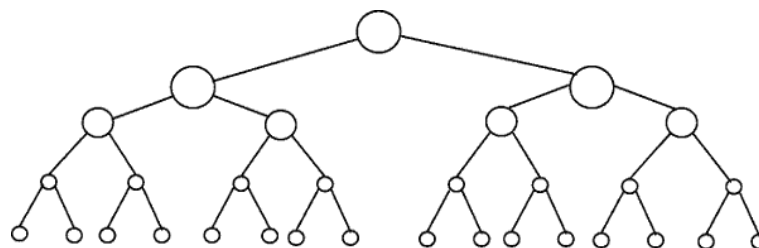


Las redes scale-free pueden construirse dinámicamente de modo que, al igual que hacen las personas, un nodo al ingresar al grafo tiene más probabilidad de conectarse con un nodo de mayor grado. Esto genera que la distribución de conexiones entre personas siga una power-law, es decir, una función que decrece como K para alguna potencia fija C siguiendo la función $1/k^c$.

Estudiar este tipo de redes suele ser muy interesante, ya que los super-spreaders tienen una probabilidad muy alta de contagiarse y suelen ser grandes focos de contagio. Esta es una de las razones por las que, por ejemplo, existen enfermedades sexuales que continúan manifestándose en poblaciones cuya gran mayoría de miembros son monógamos en relaciones a largo plazo, pero aún así existen grupos de personas con múltiples contactos sexuales que ayudan a mantener la epidemia activa. Lo mismo para enfermedades de transmisión por aire, gran parte de los contagios se debe a estos super-spreaders. Además, en este clase de redes se ha demostrado que la vacunación de individuos aleatoria no es útil para combatir la expansión de una enfermedad pero en cambio una vacunación específica sobre los super-spreaders puede ser de gran utilidad para detener la propagación de la enfermedad.[12]

Para la construcción de la red Scale-free se necesita n la cantidad de nodos, α la probabilidad de agregar un nodo conectado a otro nodo existente elegido aleatoriamente acorde a la distribución de los in_degrees, β la probabilidad de agregar un nuevo eje entre dos nodos existentes, al igual que α , se elige aleatoriamente según la distribución de los in_grados y γ la probabilidad de agregar un nuevo nodo conectado a otro existente elegido aleatoriamente según la distribución de out_degrees de los nodos. Luego el grafo se construye siguiendo el algoritmo presentado en [4]

3.1.5. Balanced Tree network



Los árboles por otro lado son estructuras ampliamente utilizadas en el ámbito de las ciencias de la computación y la matemática discreta y se los define como grafos conexos, acíclicos y no dirigidos. En nuestra implementación consisten de un nodo raíz del cual nace una cantidad fija r de nodos llamados "hijos", no conectados entre sí, que pueden tener r nodos como hijos (tampoco conectados entre sí) o ninguno, y así sucesivamente. La cantidad de hijos r por nodo se conoce como **branching factor**, mientras que a la cantidad de niveles de nodos con hijos se la conoce como la **altura** del árbol. Estos dos parámetros serán los que utilizemos para definir nuestra red. Si pensamos en la estructura de un árbol

genealógico(con la salvedad de que los hijos que tiene un progenitor pueden variar) podemos relacionarla con la estructura de árbol que planteamos recién, y estudiar como ciertas enfermedades hereditarias se propagan según las generaciones.

4. Experimentos

Para analizar diferentes comportamientos de la epidemia, nos basamos fuertemente en el nuevo virus **Sars-Cov-2** y en sus propiedades reportadas por fuentes oficiales. Los datos obtenidos fueron los siguientes:

Estado	Duración (días)	Usado en la simulación (días)
Incubación	7 - 14	7
Infección mild (leve) - paciente no riesgo	14	10
Infección grave	21-42	21
Infección grave - paciente no riesgo	14	21
Tiempo de Recuperación	-	14

Cuadro 1: En la siguiente tabla podremos observar los tiempos utilizados para el experimento. Estos se basaron en los tiempos publicados por estudios de la OMS [15]. En algunos casos asumimos la posibilidad de que el COVID-19 pueda ser un modelo de epidemia de tipo SIRS y asumimos un tiempo de recuperación de 14 días donde el contagiado vuelve a ser susceptible, siendo esto meramente suposición. Por último, utilizamos 10 días de infección leve como la media de casos. [13]

Luego, salvo experimentos donde se especifique lo contrario, utilizamos una muestra de los datos de Argentina. Para esto, para representar su población (actualmente de 45 millones), modelamos un total de 50000 nodos. Para determinar la probabilidad de ser población de riesgo, tomamos en cuenta la distribución etárea de Argentina y consideramos aquellos mayores de 65 años. Luego tomando en cuenta los datos registrados en [1], 25 % de los nodos será población de riesgo.

Finalmente, el pasaje de estados de cada uno de los nodos, tal como fue explicado en la sección 3 se da con cierta probabilidad, tal pasaje es representado en la siguiente tabla.

Estado Actual (EA)	Estado Siguierte (ES)	Probabilidad (P')	Probabilidad (P)
Susceptible (paciente riesgo)	Infectado Grave	-	0.57
Susceptible (paciente no riesgo)	Infección Grave	-	0.2
Infectado Grave (paciente riesgo)	Muerto	0.0053	0.1
Infectado Grave (paciente no riesgo)	Muerto	0.0009	0.02

Cuadro 2: En la siguiente tabla observamos para la población de riesgo o no riesgo, la probabilidad de pasar de un estado al otro, siendo la primer columna el estado actual (EA) y la segunda el potencial siguiente estado (ES). La probabilidad P' se calculó de [14] tomando la cota superior y promediando, por un lado, los pacientes de edad ≤ 65 y por el otro, el resto. Luego, para calcular la probabilidad P de pasar al estado Muerto de un nodo, utilizamos la distribución explicada en la sección 2 utilizando P'. Finalmente, la probabilidad P' no se utilizó en la transición de estados de Susceptible a Infectado Grave ya que sucede una única vez.

Es decir, siguiendo la tabla de transición de estados Susceptible-Infectado (grave o leve)-Recuperado, el nodo en estado EA pasará con probabilidad P al estado ES.

4.1. Experimento 1: El modelo de epidemia es determinante

Si bien el ideal de una simulación sería modelar una red de personas “real”, tal como explican en [12], determinar una red completa de contactos requiere conocer a cada individuo de la población y sus contactos, tarea que tomaría un tiempo enorme, además que requeriría un volumen de datos muy grande, entre otros problemas. Además el tipo de contacto que se modela, está fuertemente relacionado con aquellos que influyen con la expansión del virus o la epidemia, ya que no es lo mismo modelar una epidemia ETS que un virus que se expande por el aire. En conclusión, el tipo de red que se utiliza es simulada, y se configura para que se adapte a un tipo de población, variando ciertos parámetros para resolver cuestiones del tipo “..qué pasaría si..: Finalmente, las redes se modelan en términos de la distribución física de las personas y cómo son sus conexiones, simplificando los procesos complejos envueltos en las redes de personas “reales”.

Para determinar acerca de cuánto varía el ratio de letalidad, la velocidad de contagios y la cantidad de nodos infectados totales según el tipo de epidemia (SIS, SIRS o SIS), tomamos el tipo de grafo **Random Graph** explicada en la sección 3. En esta red, la posición de los individuos es irrelevante, y las conexiones son formadas de manera aleatoria. De esta manera podemos olvidarnos de cómo esto influye y concentrarnos en los demás parámetros.

Retomando, utilizando Random Graph corrimos el simulador [10] con 50000 nodos, un total de 200 iteraciones (se puede pensar una iteración como cierta unidad de tiempo fija). Por cada iteración, repetimos la ejecución de la misma 5 veces. Como cada ejecución de una iteración devuelve el porcentaje de nodos en cada estado (susceptible infectado, y/o recuperado), al correr repetidas veces una misma iteración promediamos los 5 porcentajes resultantes. Finalmente tendremos 200 valores, uno por promedio de cada iteración.

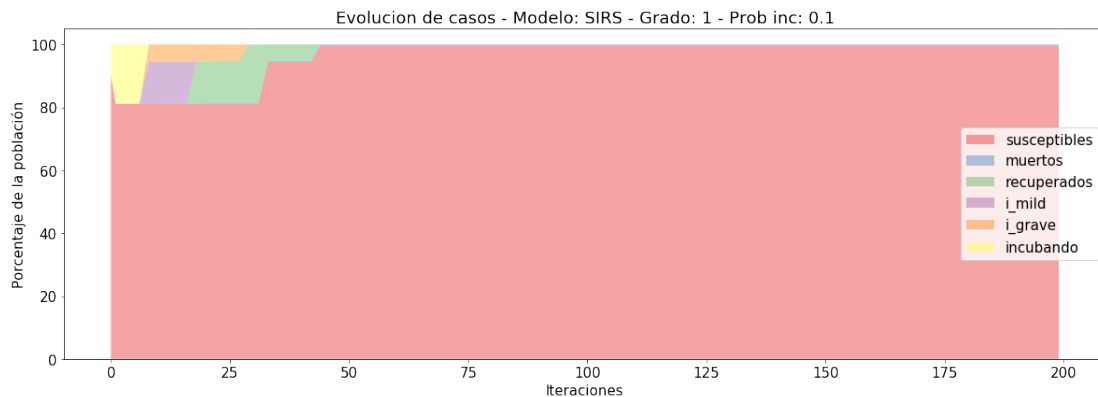
Los parámetros que variaron fueron, la probabilidad de que un nodo empiece incubando (π), y el grado de cada nodo (gr). Luego los (π, gr) posibles fueron:

Variable	Valores
PI	0.01, 0.02, 0.1
GR	1, 10, 100, 1000

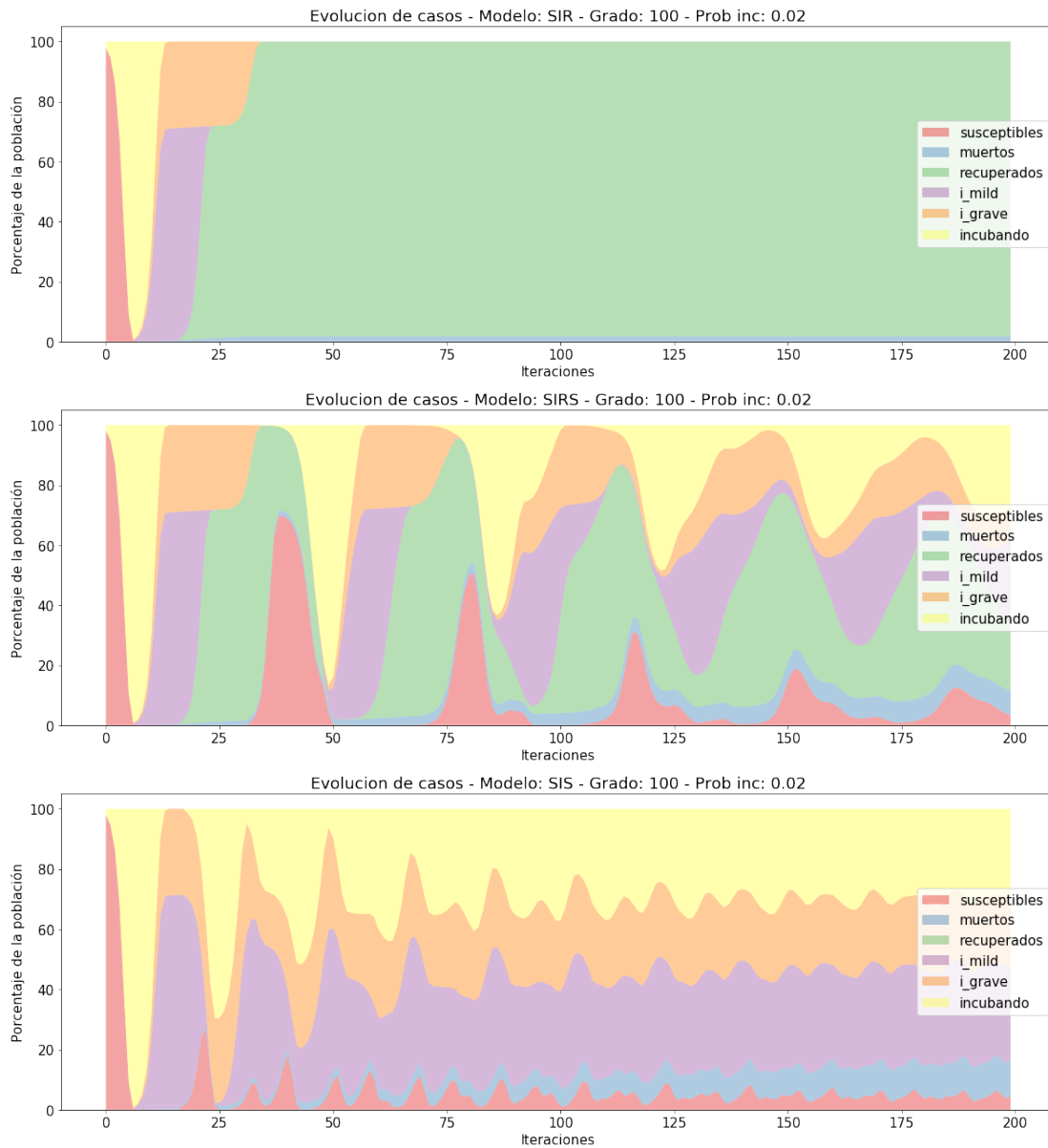
Luego de correr la simulación, podemos observar en el siguiente gráfico un panorama general de cómo se comporta cada uno de los modelos de epidemia. Para esto tomamos para SIRS SIS Y SIR, con $gr = 100$ y $\pi = 0.02$.

En los tres gráficos se puede observar un panorama general del comportamiento de los modelos de epidemias. Por un lado, podemos observar que en el modelo SIR se alcanza una curva de contagios graves y mild, luego la mayoría de la población se contagia, finalmente un porcentaje termina en estado grave mientras que la mayoría en estado leve o mild. Proporcionalmente a esto último, algunos nodos se recuperan (la gran mayoría), mientras que algunos fallecen. Por otro lado, dada la dinámica de las epidemias SIRS, podemos ver rebrotes o ciclos de epidemia, donde luego de tener un gran porcentaje recuperado, se inmuniza por cierta ventana de tiempo. Pero existen los “rebotes” siempre y cuando haya una cantidad (no necesariamente muy grande) infectada y por lo tanto contagiosa. De hecho, la dinámica SIRS resulta en una tasa de letalidad mayor, dado la cantidad de veces que un nodo tiene la posibilidad de fallecer. Finalmente en la dinámica SIS, rápidamente los nodos en estado Susceptible casi que desaparecen, para darle lugar a dos estados posibles, infectado o muerto.

Además, no es únicamente el modelo de epidemia el que define como se comporta el movimiento del virus, en los casos donde $gr = 1$, la estructura de la red termina siendo un árbol. Luego si un nodo no contagia a nadie en la ventana de tiempo infectado, el virus no se sigue propagando entre sus vecinos tal como explicamos en el modelo de epidemia, branching process en la sección 2.



Otra de las diferencias importantes del modelo de epidemia es sobre la **Tasa de Letalidad**. Utilizando el resultado de la simulación, comparamos el crecimiento de cantidad de muertes según los parámetros y el modelo.



Observando los resultados, podemos concluir dos cosas. La primera es que para modelos donde existe el estado Recuperado y el grado de los nodos es lo mínimo posible, la tasa de letalidad pasa a ser realmente baja. Esto cambia cuando $gr > 1$. Y a partir de grados mayores sin importar el π , lo que más peso tiene en la tasa de letalidad es el modelo de epidemia, siendo SIS el modelo que mayor cantidad de veces hace pasar a un nodo por infectados, y por otro lado SIR el modelo con menor cantidad de muertos, ya que una vez generada inmunidad, es decir, cierta cantidad de nodos inmunizados, la epidemia se detiene.

Otro parámetro importante para evaluar la evolución de una epidemia, es tomar el porcentaje de infectados a lo largo del tiempo y cómo es su evolución. Para esto fijamos gr en 100, basándonos en la cantidad de interacciones que una persona promedio tiene a lo largo del día (para esto no solo consideramos familiares o amigos/as, también se puede pensar en personal de atención en locales, colegas de trabajo, conductores de transporte público o privado, entre otros), y luego fijamos $\pi = 0.02$ variando el modelo de epidemia (en la práctica, iteramos π con 0.1, 0.01 y 0.02, no encontrando diferencias significativas entre estos para el mismo modelo de epidemia, por lo tanto únicamente mostramos un representante de los resultados por modelo). Los resultados son los siguientes:

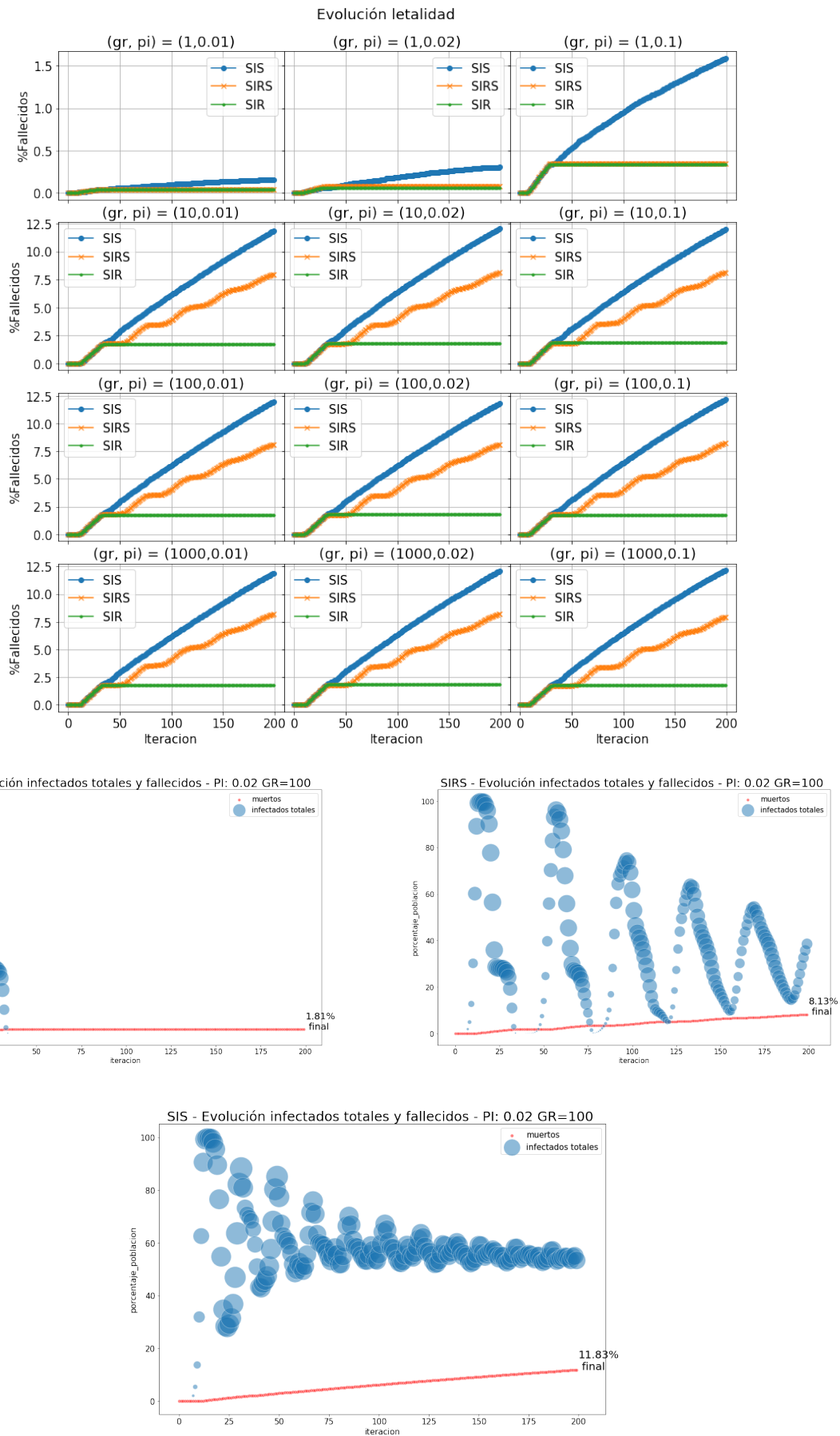


Figura 6: En las siguientes figuras se puede observar en el final de la curva de fallecidos el porcentaje final, luego el tamaño de cada punto indica la parte de infectados graves dentro de los infectados totales

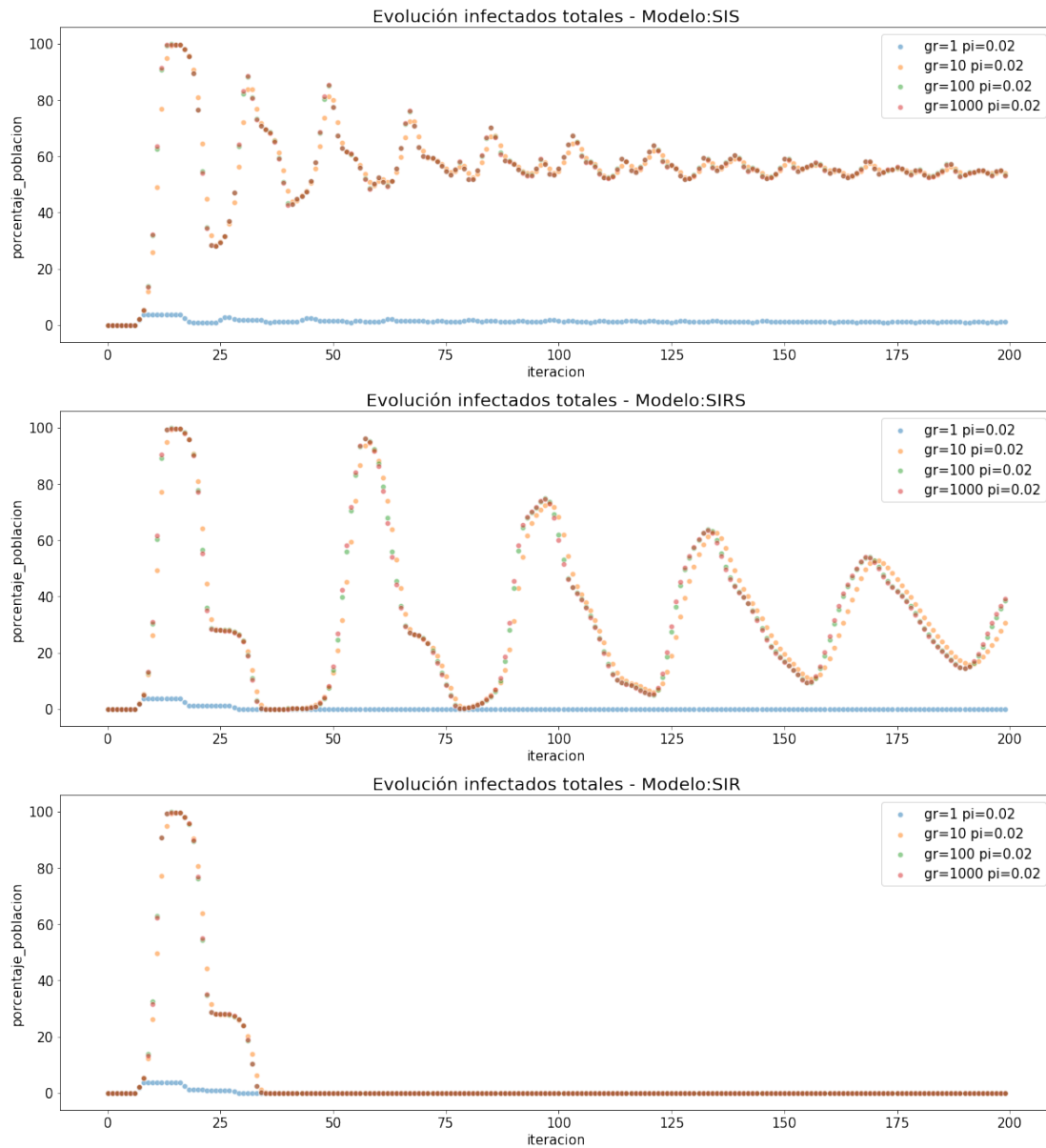
A partir de estos resultados, respaldamos el patrón de evolución de cada epidemia visto en resultados anteriores, es decir, en SIR existe un “pico de la curva” generada por el estado Recuperado de los nodos, mientras que en SIS hay “olas de contagio” mucho más cortas de Susceptibles-Infectados, pero en ningún momento se estaciona. Por último SIRS, tiene ciclos y cada ciclo con menor amplitud.

Luego en SIRS se puede observar el fenómeno de “rebrote” de la epidemia. Esto se observa en aquellas iteraciones donde el porcentaje de población infectada es muy bajo pero al ser distinto a cero, aquellos nodos en estado recuperado que regresan a susceptible, vuelven a re infectarse y generando una nueva ola de contagios (esto se observa con los diferentes picos).

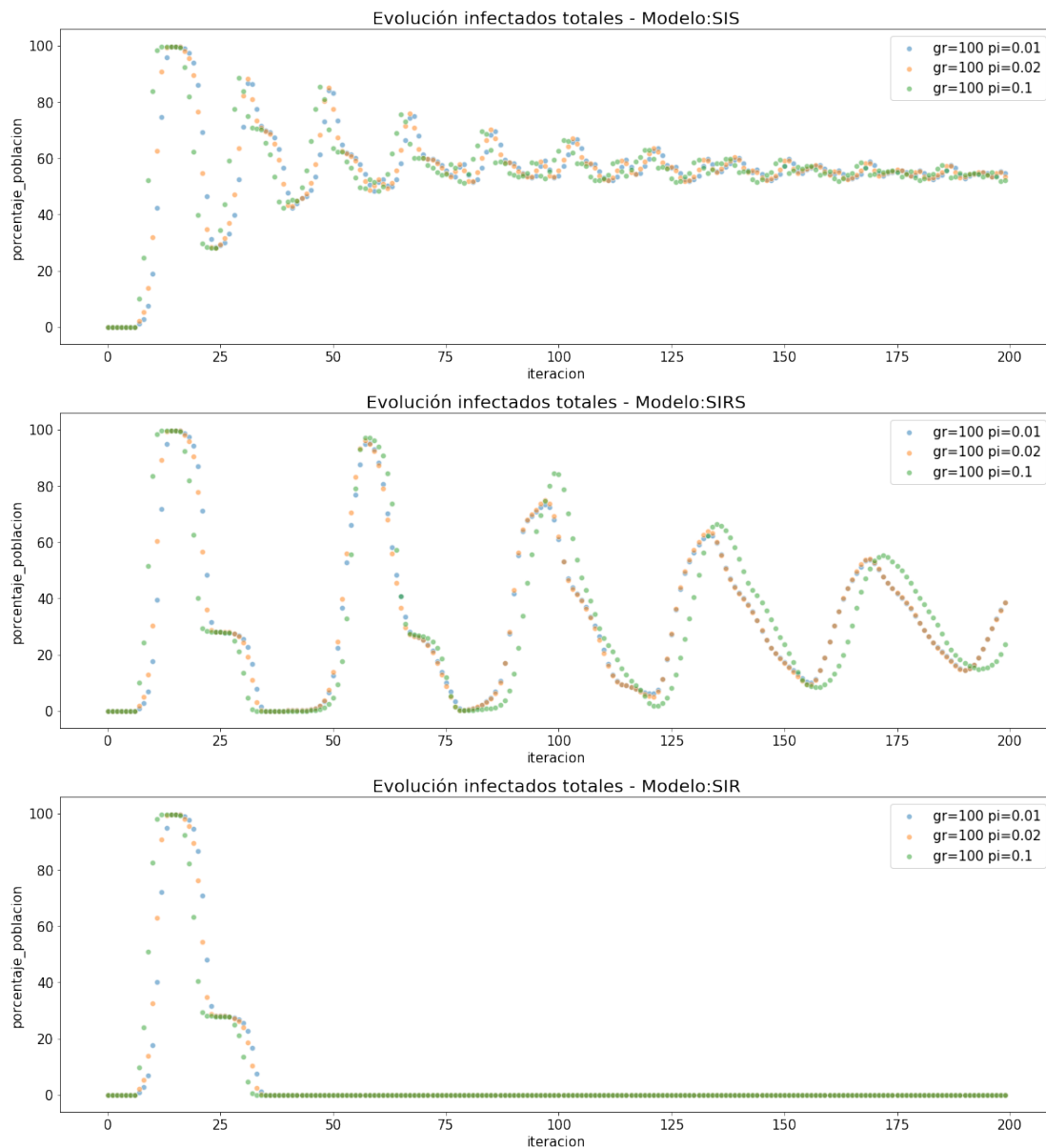
Otro fenómeno que se puede observar en los tres modelos, es el rol de los contagiados graves. Si bien en general a mayor cantidad de infectados, mayor va a ser la cantidad de infectados graves (esto se observa en los tres gráficos), vemos que a medida que la epidemia se va “estabilizando”, el porcentaje de infectados graves por sobre infectados totales va creciendo. Esto se puede deber a lo siguiente, en las primeras olas de la epidemia aquellas personas de riesgo son las que fallecen. Luego, los nodos que sobreviven pertenecen a población de no riesgo. Esto último permite que haya menor letalidad, y sumado a que las infecciones graves duran una ventana de tiempo mayor, es que vemos la concentración de infectados graves al final de las iteraciones (esto último también se puede ver en la iteración 25 del modelo SIR donde se observa una meseta en la curva de contagios). Un experimento para afirmar esto último podría ser distinguir el número de fallecidos e infectados para la población de riesgo diferenciada de la población de no riesgo.

Finalmente y relacionado a esto último, el porcentaje de fallecidos final, en la epidemia SIS (como vimos en experimentos anteriores) es la mayor de las tres epidemias, mientras que en SIR la menor. Además, se puede observar una correlación entre cantidad de infectados graves y crecimiento de la curva de fallecidos. Sumado a esto, los picos también generan crecimientos en la letalidad.

Otra variante que sumamos a los experimentos es para cada modelo observar cuál es el peso en el crecimiento de una epidemia tanto del grado de cada nodo como el porcentaje inicial de infectados. Para esto primero comparamos por cada modelo de epidemia, fijando π_i y variando γ_r (al igual que antes, los resultados para diferentes π_i son similares, por lo tanto utilizamos $\pi_i=0.02$ siendo este representativo).

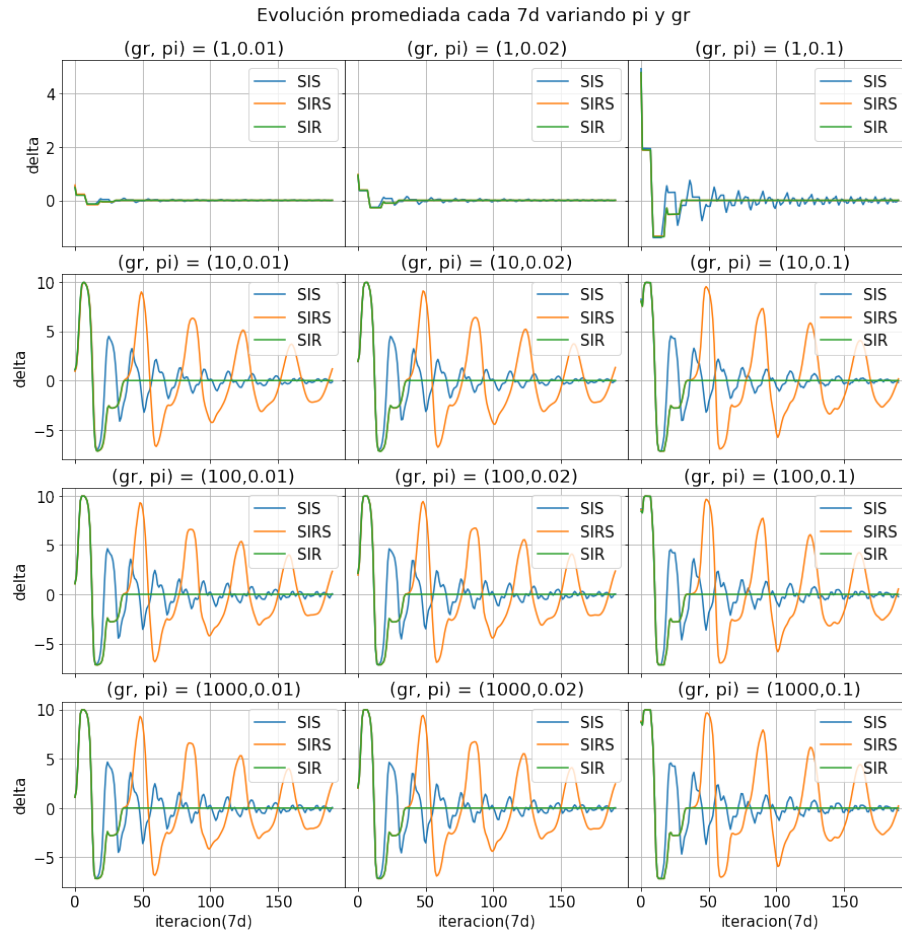


Como conclusión podemos observar que $gr = 1$ es determinante para cualquier tipo de epidemia. Mientras que en el resto de los grados, el determinante es el modelo de epidemia. Luego vamos a observar el peso del porcentaje inicial de infectados en la evolución de la epidemia. Para esto, dado que sabemos que el grado en 1 determina la evolución, vamos a utilizar $gr = 100$ con la misma justificación que en experimentos anteriores.



Nuevamente se confirma que el modelo es lo más determinante de una epidemia. Sin embargo, podemos observar en modelos como SIRS que aquel corrida con mayor π (0.1), es aquel que en todo momento resulta en una mayor amplitud. Esto indica que, es importante a la hora de medir una epidemia aproximar el porcentaje inicial que comenzó con el virus.

Finalmente, buscamos medir cuánto es la estabilidad por cada modelo, dependiendo de π y gr . Para esto sacamos un promedio móvil de infectados cada 7 días, y medimos con respecto a los 7 días anteriores. El resultado fue el siguiente:



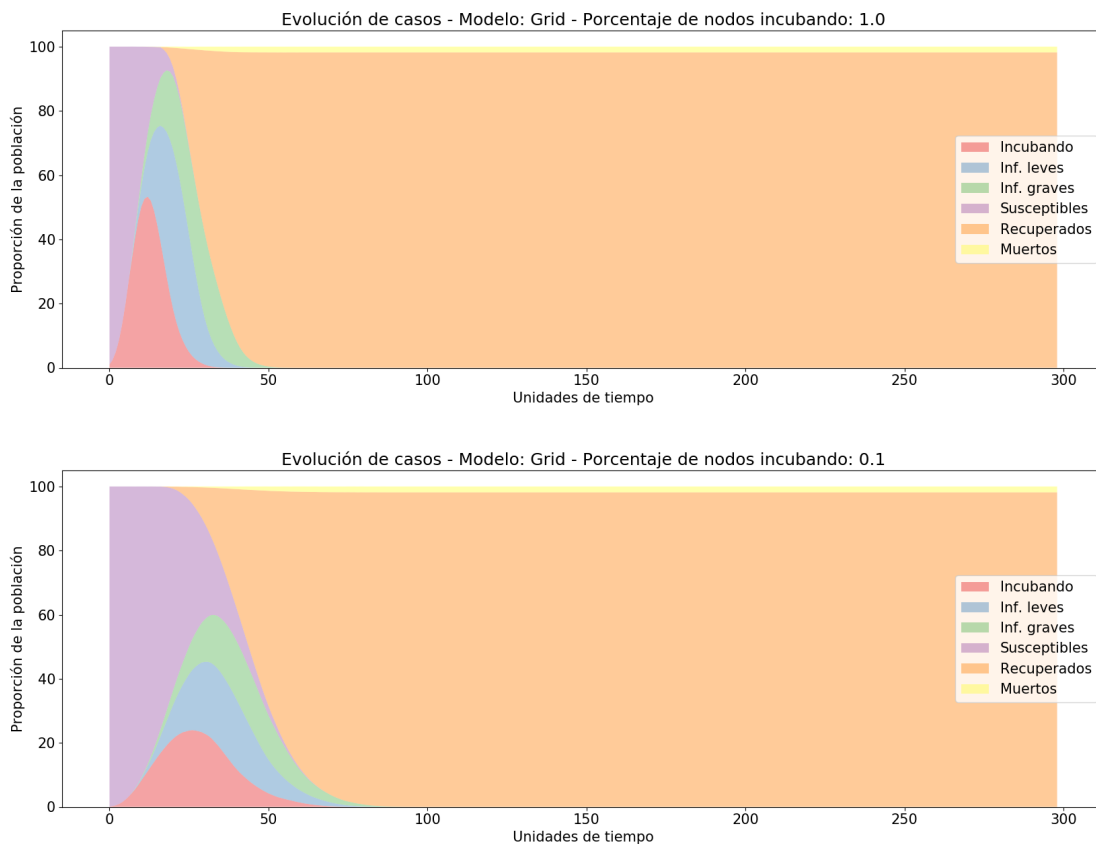
Como bien vimos a lo largo de esta sección, $g_r=1$ es determinante. Sin embargo, para $g_r=1$ y un mayor porcentaje inicial de infectados, deja de haber cierta estabilidad. Para grados mayor a 1, lo que determina la estabilidad y su comportamiento, es el modelo. En particular, podemos observar comportamientos particulares del modelo, como por ejemplo, el descenso de amplitud para SIRS, la pronta estabilización luego del "pico" de contagios para SIR y finalmente para SIS los bajos ciclos entre Susceptible e Infectados pero duradero.

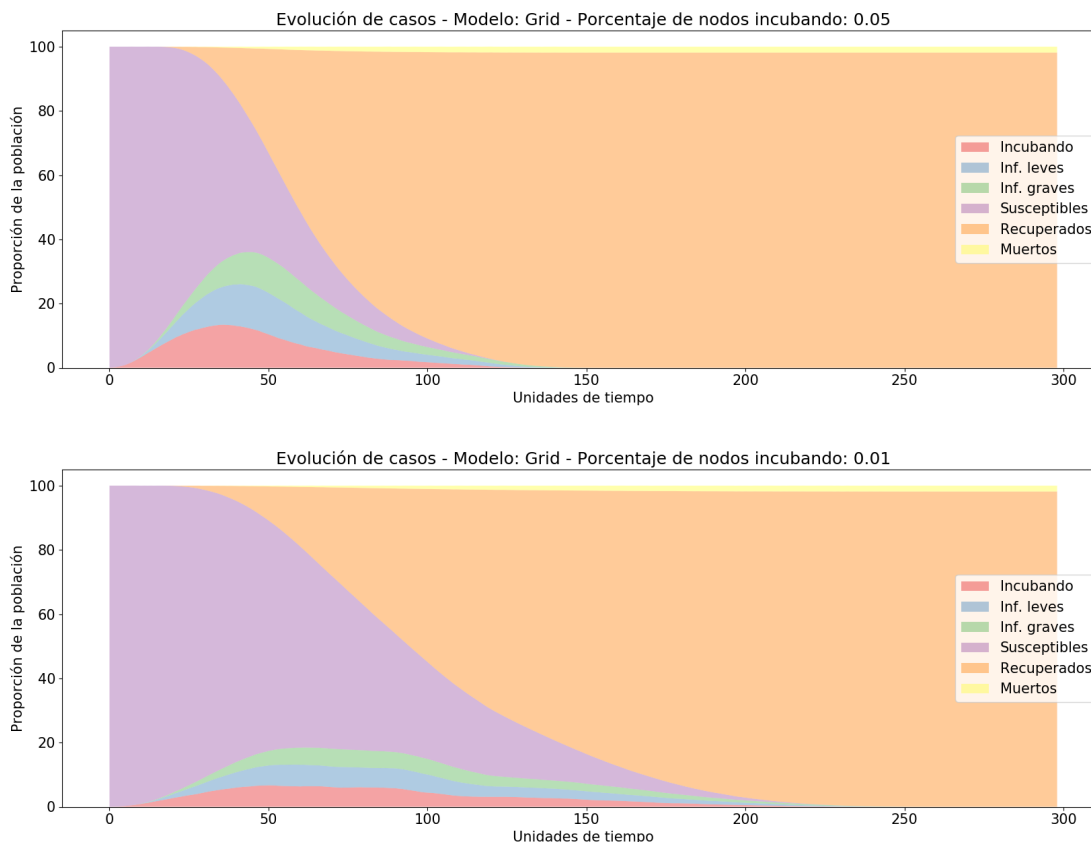
4.2. Experimento 2

Uno de los grandes interrogantes aún sin responder es a qué tipo de modelo epidemiológico se corresponde la epidemia causada por el Covid, en este experimento decidimos estudiar cómo evoluciona la infección en el caso de que esta se comporte de acuerdo al modelo SIR y poder responder a ciertos interrogantes. ¿Se logra alcanzar la inmunidad a la enfermedad en algún momento? ¿Cuándo se llega al pico de contagios? ¿Con qué velocidad evoluciona la infección? ¿Se llegan a contagiar todos los nodos? En caso contrario ¿Qué porcentaje del total queda susceptible cuando se estabiliza la curva de contagios?. Para llevar a cabo las pruebas generamos redes de los 5 tipos implementados en el framework, Random, Grid, Sale-Free, Small-World y Balanced Tree y variamos los parámetros necesarios para generarlas, por ejemplo la altura y la anchura en los árboles, la cantidad de nodos que inician infectados, etc, con el fin de entender cuánto y cómo dichas características alteran el comportamiento de la infección en dichas redes. Para cada configuración generada realizamos siempre 5 ejecuciones y promediamos los resultados de cada una con el objetivo de tener un resultado consistente.

4.2.1. GRID

Para esta topología utilizamos redes de 50.000 nodos ubicados en una grilla de 200x250 y realizamos la simulación del modelo SIR variando el porcentaje de nodos infectados iniciales con valores de 0.01, 0.05, 0.1 y 1 % durante 300 unidades de tiempo, los resultados de la evolución obtenidos para cada una fueron los siguientes.





Resulta evidente que al disminuir la cantidad de nodos que inician incubando, la epidemia se expande más lentamente y la altura de los picos disminuye pronunciadamente, por otro lado si observamos la proporción de nodos muertos (el segmento amarillo arriba de todo), ésta es prácticamente idéntica en todos los casos, lo cual resulta esperable dado que en todas las configuraciones se infectó el 100 % de los nodos.

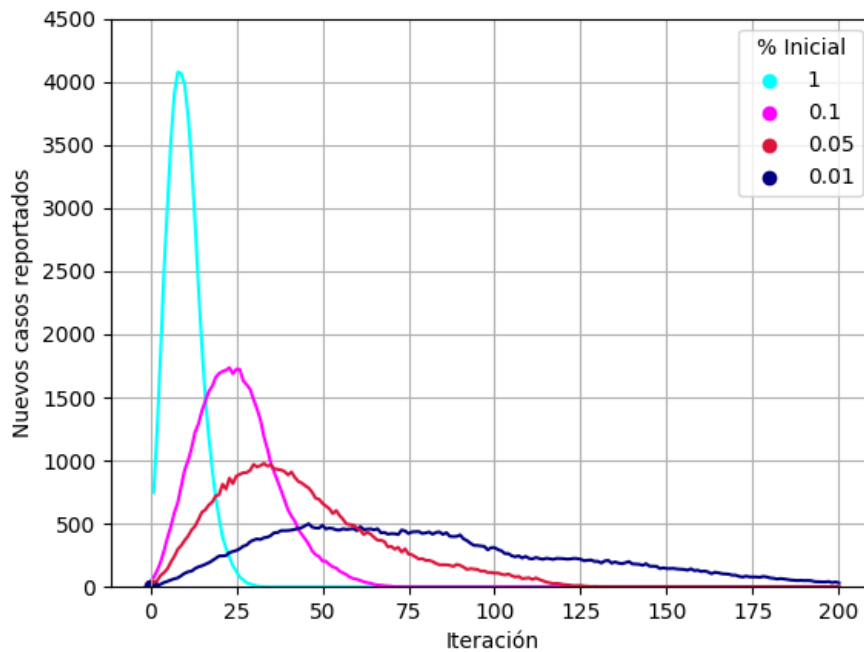
En la siguiente tabla podemos observar para cada configuración cuando fue el pico de contagios, o sea, en qué iteración se infectaron más nodos, el porcentaje total de nodos infectados y la última iteración en que se detectaron nuevos contagios, que no representa otra cosa mas que el fin de la epidemia.

% Inicial de infectados	Pico	Nodos infectados	Iteración último contagio
1	9	100 %	34
0.1	24	100 %	73
0.05	34	100 %	132
0.01	47	100 %	237

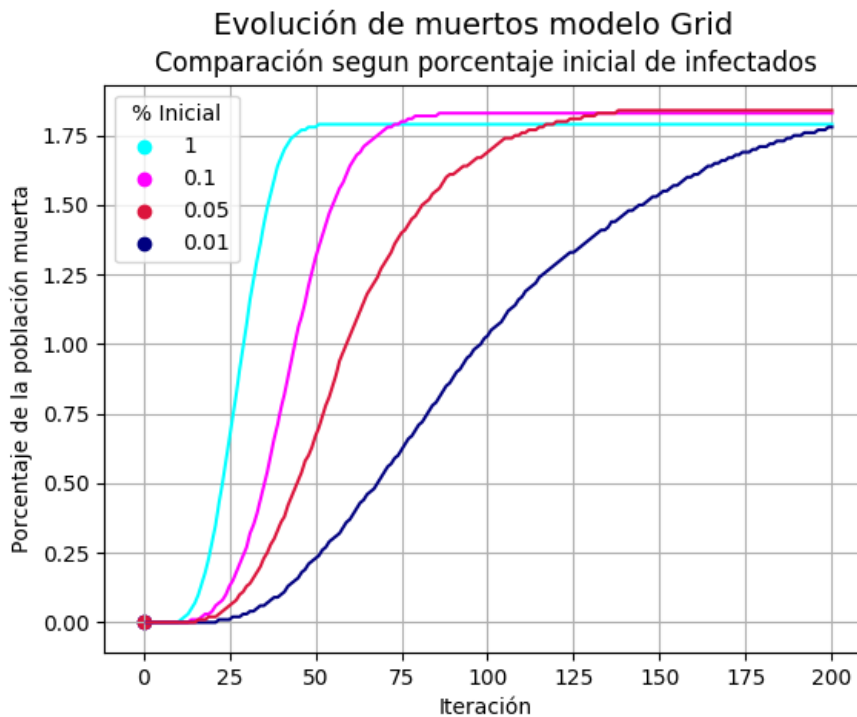
Cuadro 3: Tabla comparativa modelo grid.

Unificando los datos de las cuatro configuraciones podemos mirar por ejemplo cómo evoluciona la curva de nuevos casos detectados por iteración en cada variación del modelo.

Evolución de nuevos casos por día modelo Grid



Por último podemos hacer el mismo análisis para ver como evoluciona el porcentaje de muertos por iteración para cada configuración:



Estos dos gráficos resumen la información de los cuatro mostrados anteriormente respecto de la velocidad en que evoluciona la infección y las muertes reportadas. Es notable el hecho de que partiendo con un porcentaje de nodos infectados del 0.01, nunca se haya alcanzado un pico abrupto en la cantidad de infectados reportados por día, manteniéndose este siempre por debajo de 500. Por otro lado si miramos la evolución en el porcentaje de muertos, se puede apreciar como todas las curvas convergen al mismo valor aunque se acerquen a distintas velocidades.

4.2.2. Random

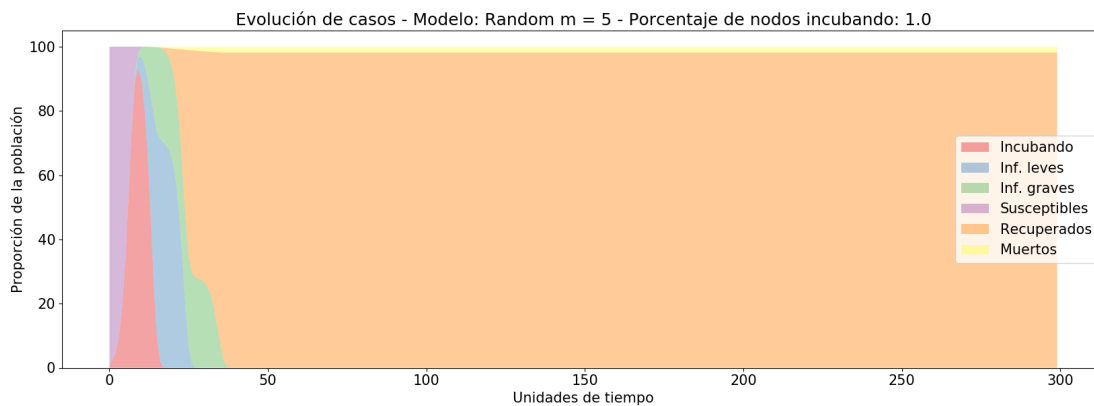
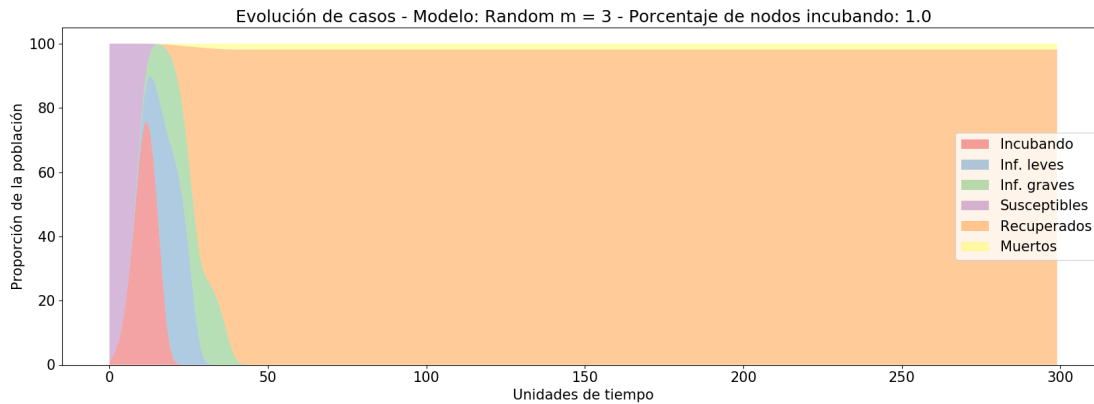
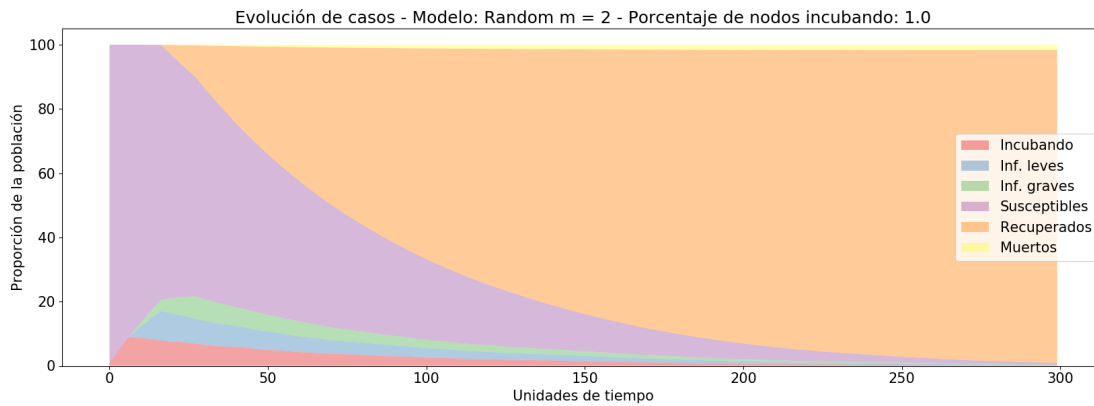
Como mencionamos en la sección 3.1 sobre las topologías de redes implementadas, las redes random se forman especificando la cantidad total de nodos y el grado de cada uno de ellos. Para este experimento decidimos crear redes de tamaño fijo, con 50.000 nodos, y variar el porcentaje de nodos que comienzan incubando así como el grado de estos. Los porcentajes de nodos que inician incubando los situamos en 1 %, 0.1 % y 0.01 % mientras que el grado de los nodos en 2, 3 y 5, dándonos un total de 9 configuraciones posibles para correr el experimento. A diferencia de las redes grid, tenemos un parámetro más para analizar y ver como se comporta la epidemia que es el grado de los nodos. En la siguiente tabla comparativa podemos ver los resultados obtenidos para las simulaciones durante 300 unidades de tiempo, por lo que en aquellas cuya iteración de último contagio equivale a 300, la epidemia no habría cesado al culminar la simulación:

% Inicial de infectados	Grado	Pico	Nodos infectados	Iteración último contagio
1	2	1	99.48 %	300
1	3	9	100 %	18
1	5	7	100 %	12
0.1	2	7	33.98 %	300
0.1	3	15	100 %	26
0.1	5	10	100 %	16
0.01	2	127	4.72 %	300
0.01	3	22	100 %	32
0.01	5	14	100 %	22

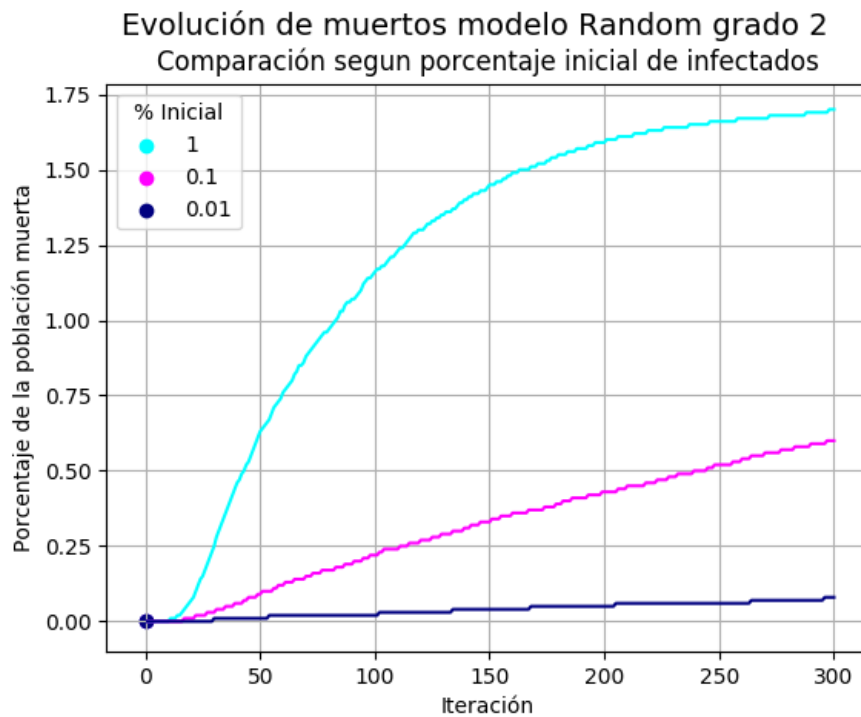
Cuadro 4: Tabla comparativa modelo random.

El análisis que podemos hacer sugiere que no hay mucha diferencia entre las configuraciones cuyos nodos tienen grado 3 o 5, en todas se infecta el 100 % de los nodos mientras que el pico de contagios y el cese de la propagación de la epidemia culminan rápidamente, de modo que aumentar la cantidad de contagiados inicial lo único que logra es adelantar el pico y la velocidad con que todos los nodos se infectan. Por otro lado pasar de grado 3 a 5 tampoco parece revelar alguna diferencia notable en la evolución de la infección, pero si miramos los resultados de las configuraciones donde el grado de los nodos es 2, si hay una diferencia notable con las demás. En primer lugar, variar el porcentaje inicial de contagiados determina en gran medida la velocidad de propagación, pasando de un 99.48 % de nodos infectados con 1 % de población inicial infectada a 33.98 % con 0.1 % hasta llegar a 4.72 % de infectados con 0.001 % de contagiados iniciales. Además de esto el pico de infectados pasa de ser la primera iteración con 1 % de nodos incubando a darse en la 127 con 0.001 % de nodos incubando

A continuación podemos ver la comparación entre las configuraciones cuyo porcentaje inicial de nodos infectados es 1 % variando entre los tres grados posibles quedando en manifiesto el hecho de que aumentar de grado 3 a 5 no representa prácticamente ninguna diferencia.

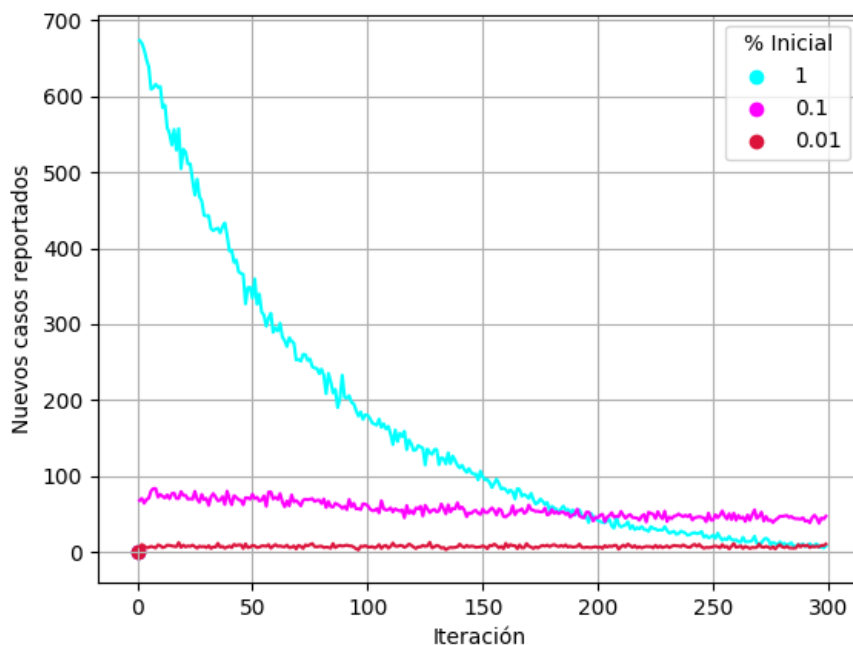


Nos resultó interesante por otro lado comparar la evolución de muertos e infectados entre las configuraciones con grado 2 ya que fueron las que mas diferencias presentaron entre si, de esta forma podemos ver como cambia la velocidad de contagio y la curva de decesos en función de la proporción de contagiados inicial.



Podemos observar como la curva celeste, que representa la red que comenzó con 1% de nodos contagiados crece mucho más rápido y una vez que se acerca a la última iteración comienza a estabilizarse, lo cual tiene sentido dado que para ese entonces casi todos los nodos se infectaron. Por otro lado en aquellas configuraciones que comienzan con 0.1% y 0.01% de infectados la cantidad de muertes continúa aumentando a un ritmo muy lento dada la baja velocidad de contagios que presentan, como podemos observar en el gráfico de abajo.

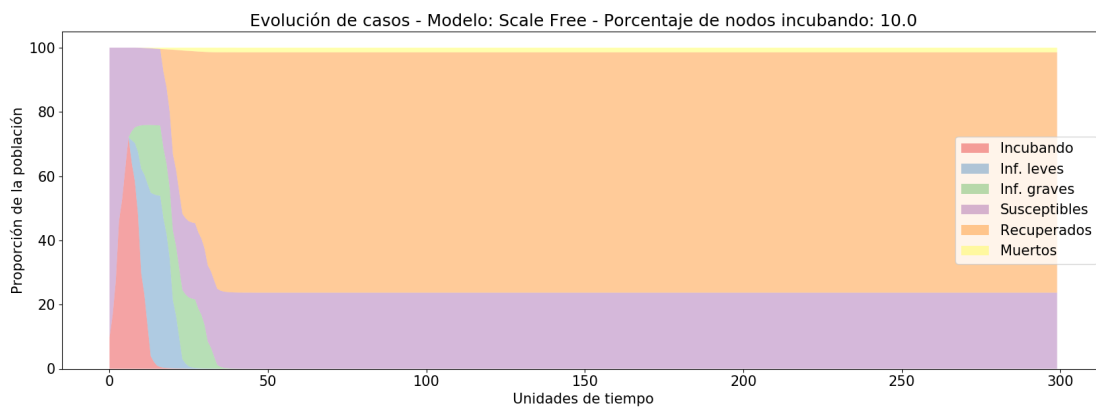
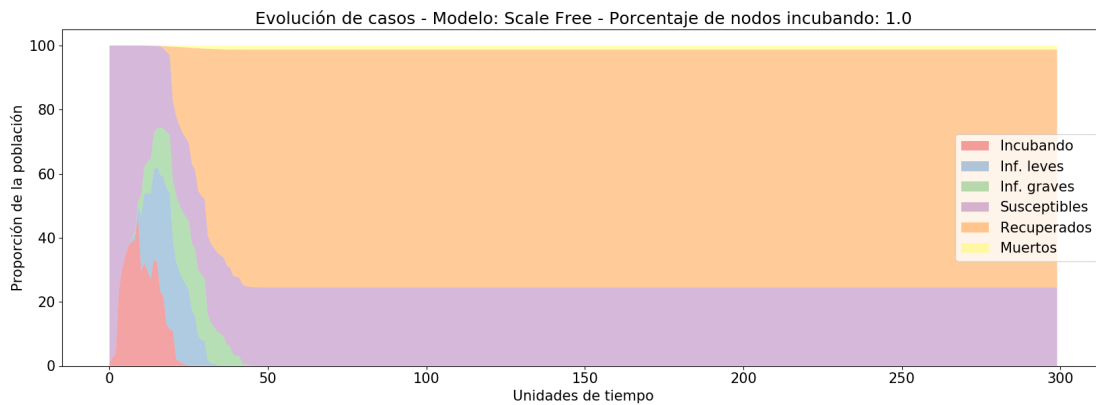
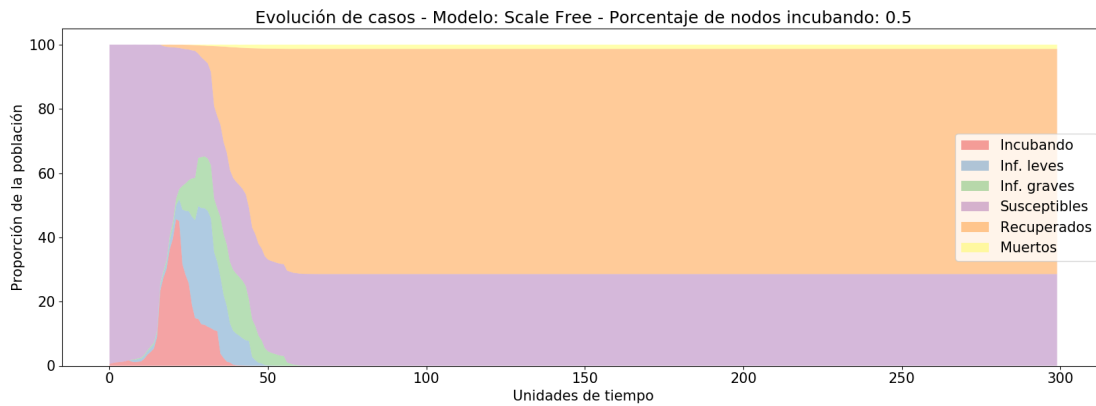
Evolución de nuevos casos por día modelo Random grado 2



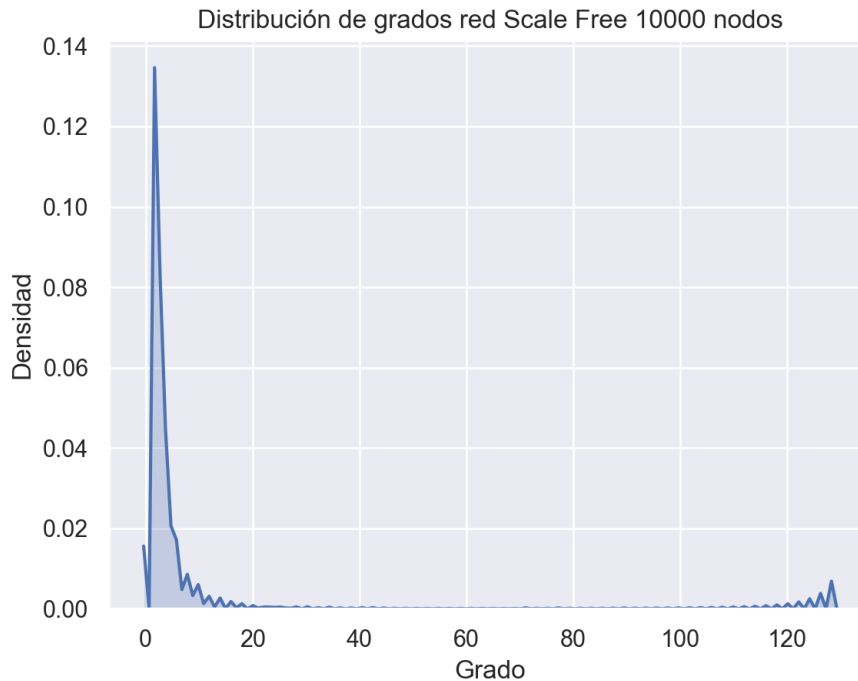
En conclusión, las diferencias mas notables las obtuvimos al pasar de 2 a 3 o 5 el grado de los nodos y en las redes de grado 2 variando la cantidad de nodos infectados, mientras que en las redes de grado 3 y 5 pareciera que cambiar la proporción inicial de infectados solamente desplaza las curvas levemente.

4.2.3. Scale Free

Como mencionamos en el apartado sobre topologías de redes, las libres de escala tienen la particularidad de seguir una distribución exponencial respecto del grado de los nodos, por lo que un alto número de estos va a tener muy pocas o cero conexiones, mientras que un pequeño grupo va a estar altamente relacionado con otros nodos. Para llevar a cabo estas pruebas decidimos generar grafos de 10000 nodos, puesto que generar grafos mas grandes con estas características consume mucho tiempo de procesamiento, y luego variar la proporción de nodos que comienzan infectados. Corrimos el experimento iniciando con 0.1, 0.5, 1, 10, 20 y 30 % de nodos infectados y obtuvimos resultados un tanto interesantes, en prácticamente todas las ejecuciones la evolución se manifestó de la misma forma, alcanzando rápidamente un pico de contagios el cual baja rápidamente pero que nunca alcanza la totalidad de los nodos en la red, nos limitamos a mostrar solo los resultados con 0.5, 1 y 10 ya que los resultados con más de 10% de nodos incubando no difieren mucho con este último:



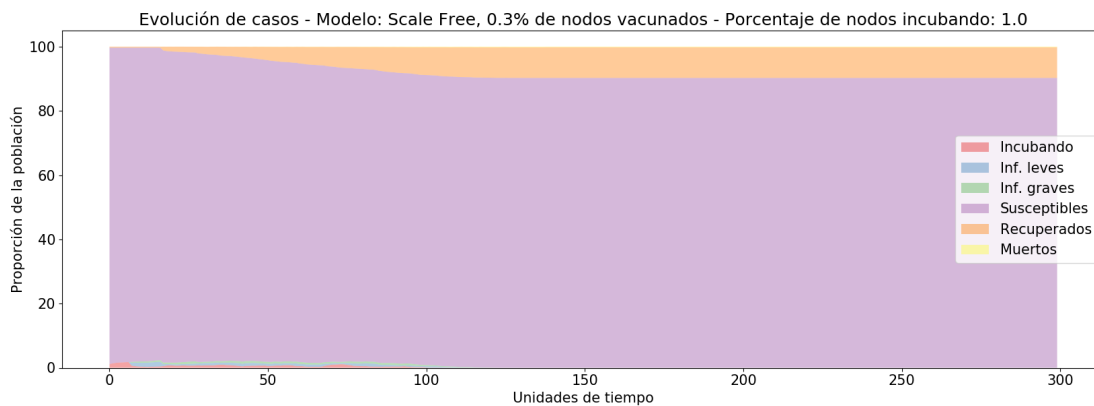
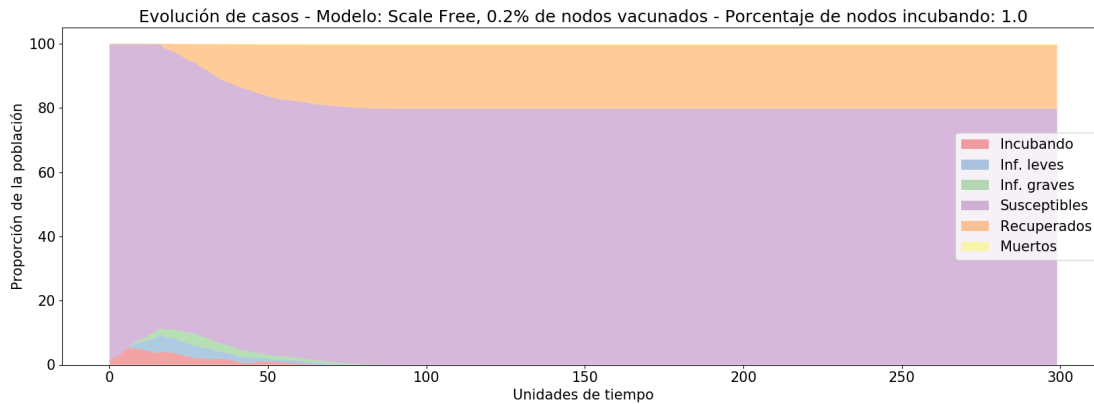
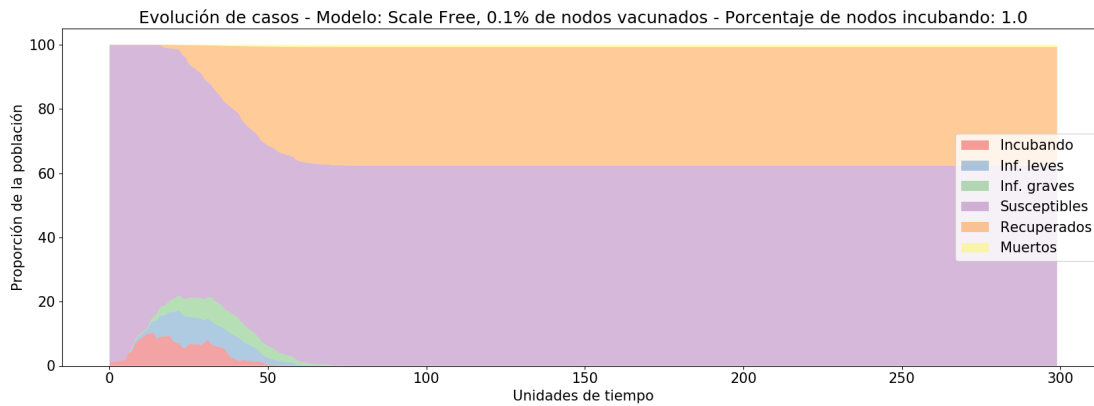
Como es evidente en todas las configuraciones el pico se alcanza rápidamente y los contagios se estabilizan alrededor de la iteración 50. Algo particular de estas redes es que siempre quedó un gran porcentaje de los nodos sin infectar, independientemente de la proporción de infectados inicial. Nos preguntamos cual podría ser la causa principal de esto y para ello generamos nuevamente grafos de 10000 nodos Scale Free con la misma librería, **networkX**, y analizamos la distribución de grados.



Este gráfico explica el comportamiento que vemos en la evolución, ya que al distribuirse de esta forma los grados de los nodos, vamos a tener miles de ellos aislados o conectados solamente a otro nodo lo que nos sugiere que nuestra red contiene una alta cantidad de componentes conexas (grupos de nodos aislados de los demás). Si la red que generamos contiene muchos grupos de nodos aislados del resto, es evidente el hecho de que la infección no puede llegar a ellos salvo que alguno de los nodos pertenecientes a dicha componente conexa se encuentre infectado.

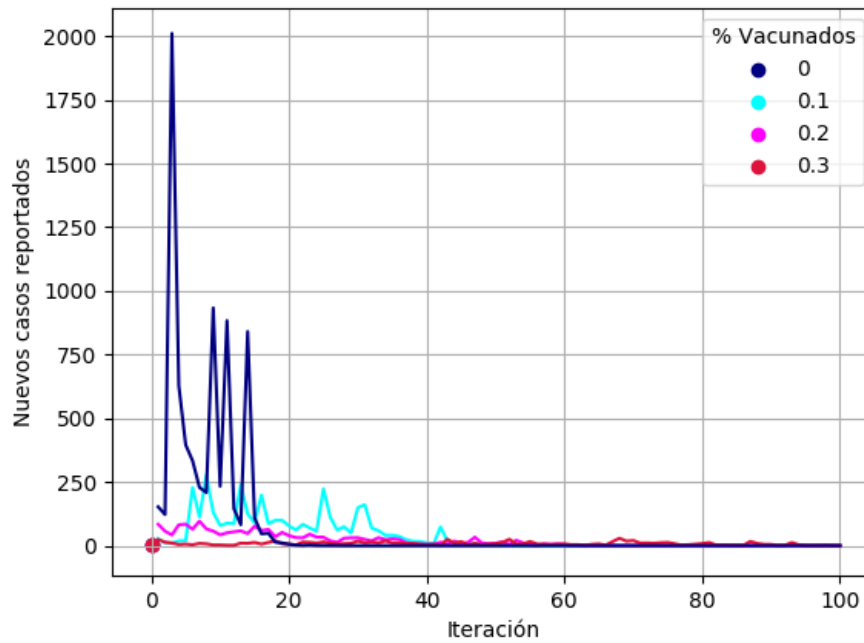
Si pensamos en las infecciones de transmisión sexual, este tipo de redes podrían ser útiles para modelarlas, dado que se puede pensar a las componentes conexas como parejas monogámicas o personas individuales que no mantienen vínculos sexuales (por ejemplo niños pequeños) y que por esta razón no presentan riesgo de contagio. Mientras que otro grupo de nodos mas pequeño se relaciona activamente y esta en contacto permanente con otros nodos de la red, por lo que la infección puede alcanzarlos eventualmente.

Para finalizar nos preguntamos cómo evolucionaría la enfermedad si tomamos un porcentaje de los nodos de mayor grado y simulamos una vacunación de dichos nodos (en nuestro caso directamente iniciándolos en estado recuperado). ¿Será suficiente vacunar una pequeña proporción de nodos en la red para evitar la propagación de la enfermedad? En dicho caso ¿Qué porcentaje necesita recibir la vacuna? Para esto configuramos una red con 1 % de nodos infectados y procedimos a vacunar el 0.1, 0.2 y 0.3 % de los nodos, pero tomando aquellos con mayor grado en vez de vacunarlos aleatoriamente y obtuvimos estos resultados:



Esto nos muestra que vacunando un 0.3 % (solo 30 de 10000) de los nodos, pero principalmente aquellos que presentaban mayor grado, evita que la infección se propague casi por completo, además que con vacunar al 0.1 % o 0.2 % alcanza para ver efectos cuantitativos. Comparando la velocidad de propagación esto queda en evidencia.

Evolución de casos modelo Scale Free 1% infectados



También lo podemos observar en el porcentaje final de nodos infectados.

% Vacunados	%Infectados Totales
0	75.4
0.1	37.5
0.2	20
0.3	9.4

Cuadro 5: Tabla comparativa Scale Free 1 % de infectados iniciales.

4.2.4. Small World

Las redes de mundo pequeño, como explicamos anteriormente, se caracterizan por presentar un alto nivel de clustering y además por tener en promedio una distancia corta entre todo par de nodos. Las redes de este tipo se configuran ubicando a los nodos en forma de anillo, luego se conecta a cada nodo con sus k vecinos mas cercanos y se generan ejes con una proporción p de los demás nodos. Es por esto que k y p son dos variables que nos interesan a la hora de ver como evolucionan los contagios, en este experimento decidimos fijar una proporción inicial de nodos infectados de 0.1 estática y asignar distintos valores a k y p para ver como afectan la evolución de la enfermedad. Para esto generamos redes situando k en 2, 3 y 4 y para cada configuración probamos con valores de p en 0.1, 0.3 y 0.5.

Incluimos los resultados de la evolución resumidos en una tabla dado que son demasiados gráficos para mostrarlos todos. En este experimento realizamos simulaciones de 365 unidades de tiempo, si en la tabla vemos que la iteración de último contagio es 365 esto significa que al finalizar la simulación todavía había nodos contagiándose, por lo que la epidemia no habría cesado al cabo de dicho lapso de tiempo.

k	p	Pico	Nodos infectados	Iteración último contagio
2	0.1	89	80.1 %	365
2	0.3	75	87.7 %	365
2	0.5	51	96.2 %	365
3	0.1	100	88.5 %	365
3	0.3	80	89.8 %	365
3	0.5	67	95.1 %	339
4	0.1	17	100 %	34
4	0.3	12	100 %	23
4	0.5	12	100 %	20

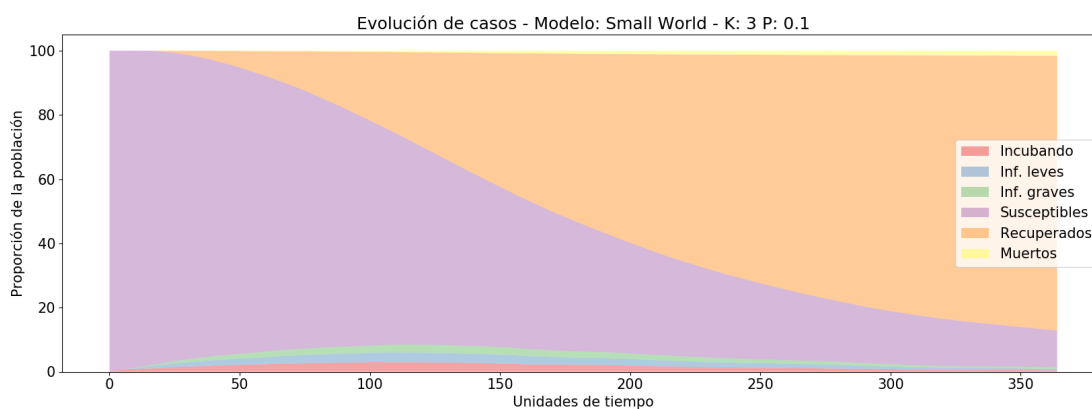
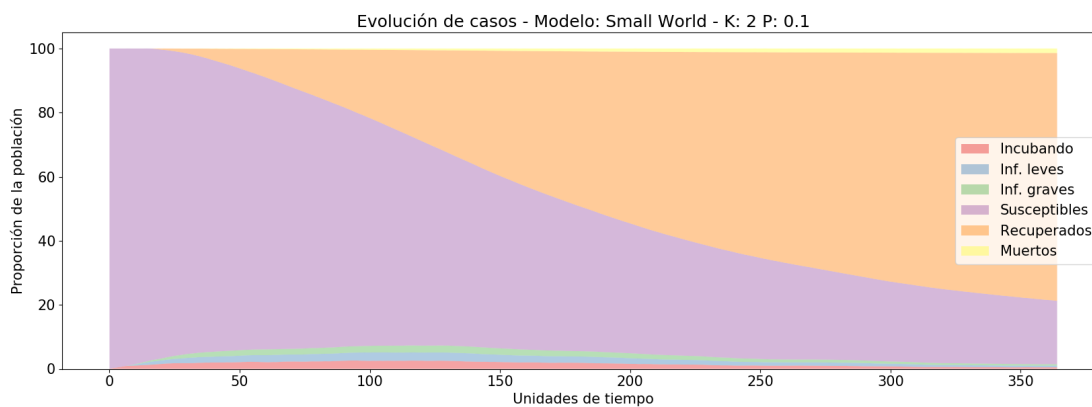
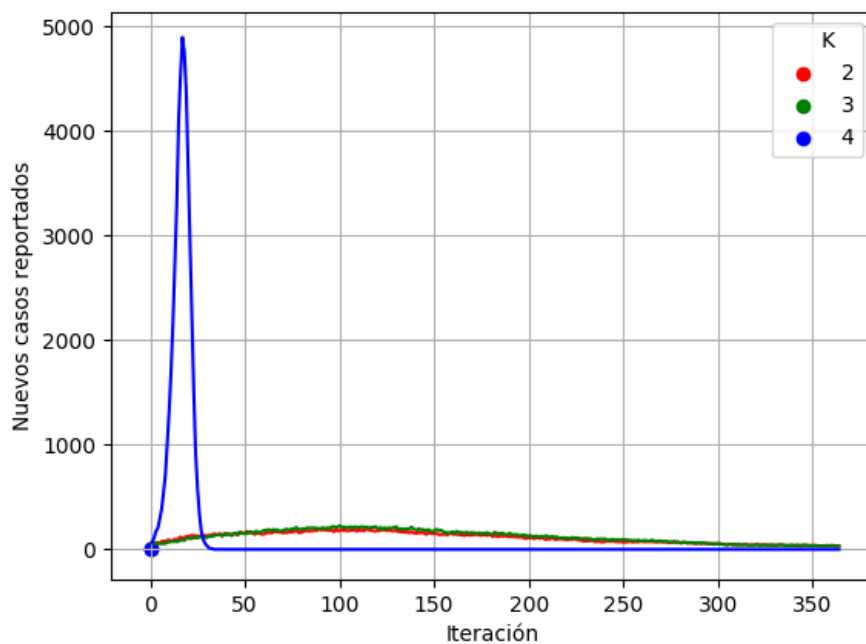
Cuadro 6: Tabla comparativa modelo Small World.

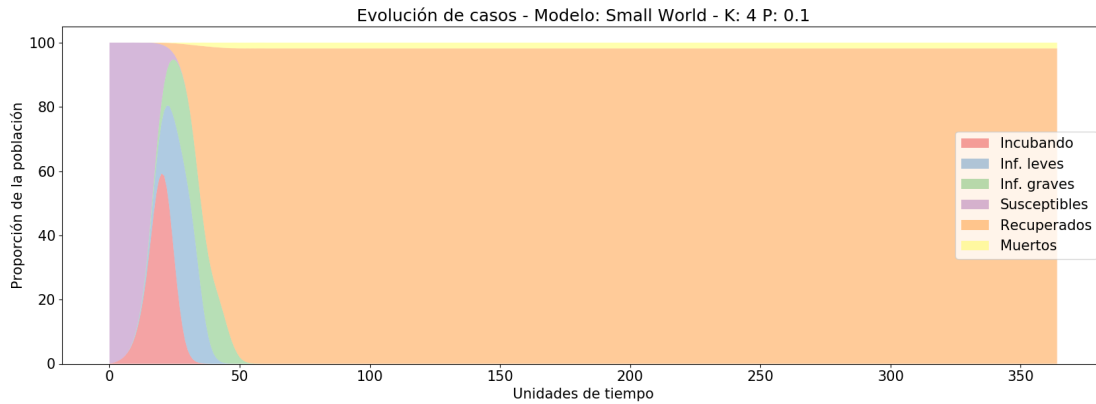
De estos resultados podemos concluir los siguientes puntos:

- Fijar k en 2 o 3 no genera cambios significativos en el porcentaje de infectados al final de la simulación.
- Fijado k , aumentar p implica mas contagios y un pico mas temprano.
- Al fijar k en 4 la infección alcanza a todos los nodos y finaliza abruptamente(en las iteraciones 34, 23 y 20)
- Aún quintuplicando los valores de p para k 2 y 3 no se llegaron a infectar todos los nodos por lo que parece menos significativo que k en cuanto a la evolución respecta.

Podemos ver la diferencia que hay en la velocidad de contagios y en la evolución en aquellas configuraciones con $p = 0.1$

Evolución de nuevos casos por día modelo Small World P: 0.1

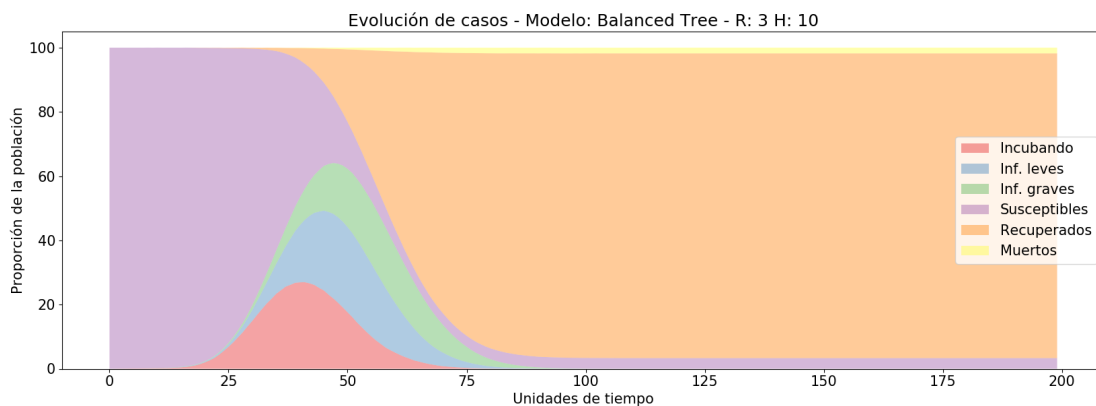
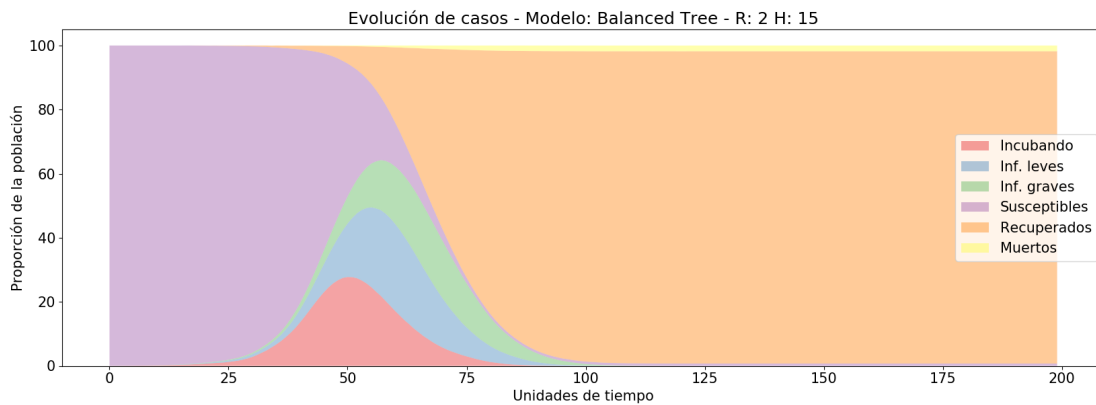




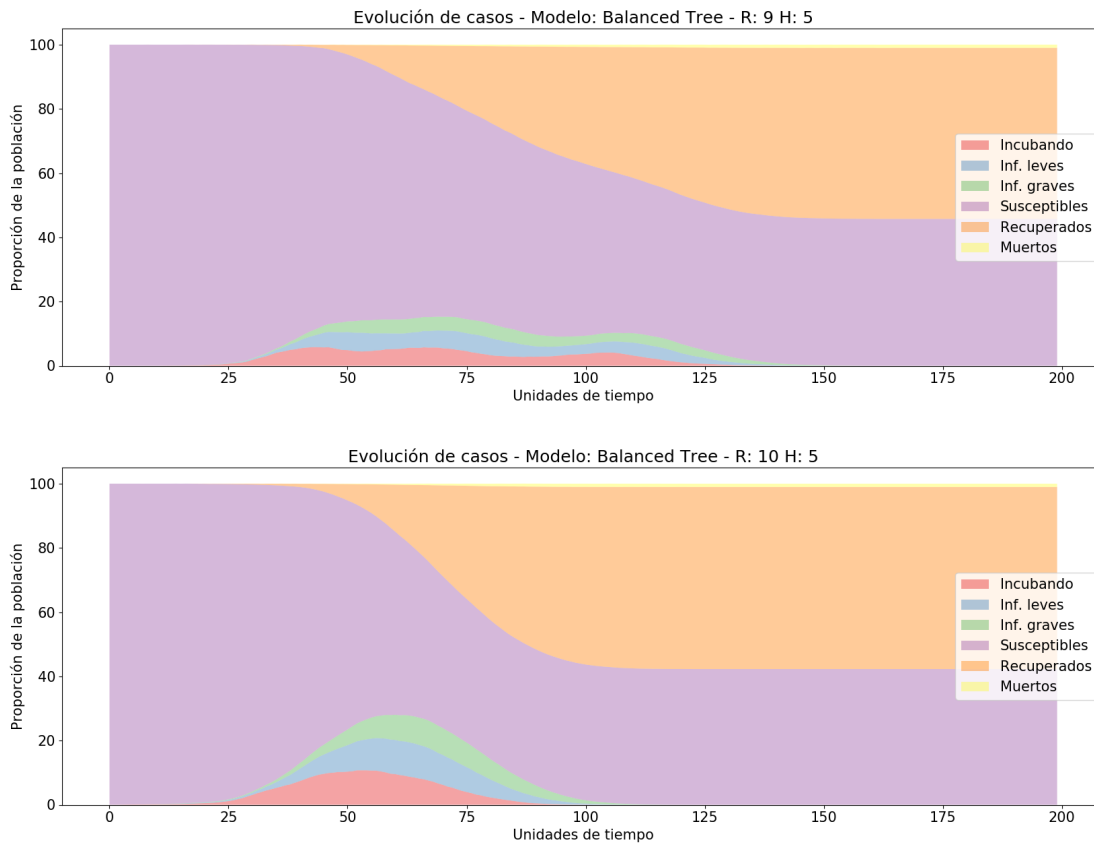
4.2.5. Balanced Tree

En el último apartado de este experimento, realizamos pruebas sobre árboles balanceados alterando tanto el branching factor r como la altura h del árbol utilizando al igual que en las redes Small World, un porcentaje fijo de nodos iniciales contagiados ubicado en 0.01.

Luego generamos dos tipos de árboles distintos, por un lado árboles con un branching factor de 2 y 3 hijos por nodo pero de una altura considerable, 15 y 10 respectivamente, y por el otro, árboles con un branching factor de 9 y 10 hijos por nodo pero de altura 5. Dado el crecimiento exponencial relacionado con r y h estos fueron los parámetros que nos parecieron razonables para diferenciar lo más posible los árboles "altos" respecto de aquellos "anchos". Como la cantidad de nodos de un árbol depende directamente de su altura y su aridad, las redes generadas no contaron con la misma cantidad de nodos. A continuación podemos ver la evolución de la infección en cada una:



Podemos ver como en ambos modelos la enfermedad evoluciona casi idénticamente, salvo por el hecho de que queda un pequeño porcentaje mayor de nodos sin infectar cuando el branching factor es 3.







Por otro lado al aumentar al branching factor considerablemente, 9 y 10 versus 2 y 3, obtenemos que la infección no solo se propaga mas lentamente si no que además, la proporción de nodos remanentes sin infectar es casi del 50 %, mientras que en las redes con branching factor menor se infectó casi el 100 % de los nodos.

En general no hubo muchos mas datos que nos parezcan relevantes de mostrar en este análisis.

4.3. Experimento 3

A lo largo de la historia, existieron diferentes epidemias generadas por diferentes virus, cada uno de estos con características completamente diferentes, tanto su velocidad de reproducción, sus formas de transmisión y también su modelo de epidemia. Algunas características de estas se ven resumidas en la siguiente tabla:

Enfermedad	Gripe	COVID-19*	SARS	MERS
Patógeno	 Influenza	 SARS-CoV-2	 SARS-CoV	 MERS-CoV
Reproducción base	1.3	2.0 - 2.5	3	0.3 - 0.8
Letalidad	0.05 - 0.1%	~3.4%	9.6 - 11%	34.4%
Tiempo de incubación	1 - 4 días	4 - 14 días	2 - 7 días	6 días
Tasa de hospitalización	2%	~19%	Mayoría de casos	Mayoría de casos
Infecciones anuales (globales)	1.000M	174,779	9098 (2003)	420
Muertes anuales (globales)	10K - 61K	6,686	0 (desde 2003)	0 (desde 2014)

* Datos actualizados al 16/3 y en progreso

Figura 7: Imágen tomada de [2]

En particular, la Gripe o Influenza es un virus estacional (generalmente aumenta su aparición en fechas invernales). Si bien la gripe de manera descontrolada provocaría una gran cantidad de muertes o enfermos graves, cada año se fabrica una vacuna que impide una gran expansión del virus.

Sin embargo, en el transcurso de la pandemia producida por el virus Sars-Cov-2, una de las frases más escuchada fue “es sólo una gripe”. Pero a medida que se conocía más de la epidemia, se hizo clara la diferenciación entre las enfermedades.

Para esto, [11] buscamos utilizar el simulador para comparar en un contagio “descontrolado” de una misma población (dadas las características de cada virus), el impacto del Sars-Cov-2 y la influenza en ella.

Las características que utilizamos para modelar el virus de la Influenza son:

Casos	Tiempo hasta recuperación (días)	Usado (días)
Incubación	2	2
Infección mild (leve) - paciente no riesgo	5	5
Infección grave	14	14
Infección grave - paciente no riesgo	14	14
Tiempo de Recuperación	200	200

Cuadro 7: En la siguiente tabla podremos observar los tiempos utilizados para el experimento. Estos se basaron en los tiempos publicados por los siguiente reportes [3] [6] [7]. Otro dato a tener en cuenta, es que se conoce inmunidad de al menos un año, donde el virus muta lo suficiente para que esta inmunidad no sirva del todo

Además, buscamos qué datos se conocen sobre la probabilidad de cada población de pasar una enfermedad grave o no según el grupo etario donde se pertenezca. Lo mismo buscamos para la probabilidad de deceso (para esta probabilidad tomamos en cuenta que el curso de la enfermedad estaría siendo grave).

Estado	Probabilidad
Infectado Grave - paciente riesgo	0.54
Infección Grave - paciente no riesgo	0.00074
Muerto - paciente riesgo	0.000001
Muerto - paciente no riesgo	0.00013238475

Cuadro 8: En la siguiente tabla observamos para la población de riesgo o no riesgo, la probabilidad de pasar de un estado al otro, siendo la primer columna el potencial siguiente estado. La probabilidad se calculó de [5]. Luego para calcular la probabilidad de pasar al estado Muerto de un nodo perteneciente a población no riesgo, utilizamos la distribución binomial despejando y luego utilizando la probabilidad de cada experimento, mientras que para calcular la probabilidad de deceso de un nodo de población riesgo hicimos una equivalencia entre las tasas de letalidades conocidas de Covid-19 e Influenza.

Finalmente, utilizando grafos Random que mantengan variando características como el grado de los nodos y la cantidad de nodos iniciales incubando, asumiendo para ambas enfermedades un modelo de epidemia SIR, medimos el crecimiento de nuevos contagios y muertes. Esto lo logramos corriendo 200 iteraciones por enfermedad, en particular para la Influenza variamos $gr = [100, 1000]$, mientras que en Covid-19 fijamos $gr = 100$. Por otro lado en ambos fijamos pi en 0.02. Cada iteración la corrimos 5 veces y sacamos un promedio del resultado.

4.4. La cantidad de contagios y muertes efectivamente es diferente

Uno de los primeros resultados que mostramos, es el crecimiento de nuevos contagios por iteración.

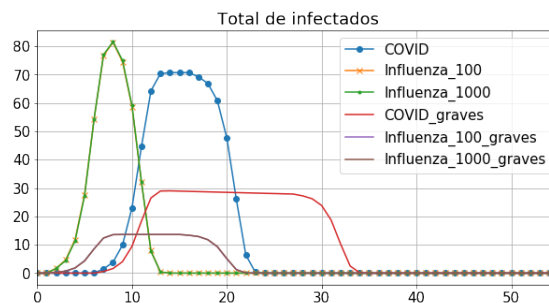


Figura 8: En este gráfico se muestran los casos graves como (X grave) con X el virus, además en el mismo gráfico mostramos la evolución de la influenza tanto para $gr = 100$ como para $gr = 1000$

Una de las primeras conclusiones que podemos obtener a partir de la simulación es, la tasa de contagio es parecida entre ambos virus. El cambio importante es, por un lado la duración de los enfermos graves

de COVID (esto genera ocupación en el sistema sanitario, el cuál puede generar su colapso), por otro su porcentaje. Esto último impacta directamente en la letalidad como veremos más adelante. Además, podemos ver que es necesario una dinámica de contactos mucho mayor para que la gripe tenga un nivel de contagio similar al Covid-19.

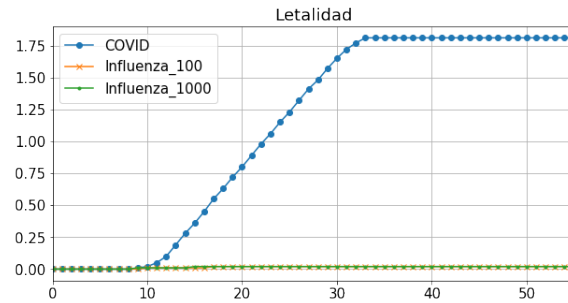


Figura 9: En este gráfico se muestran la cantidad de muertes siguiendo la nomenclatura del gráfico anterior

Finalmente, podemos observar que el ratio de letalidad del Covid vs Influenza es totalmente diferente. Si bien en el gráfico se observa que la Influenza es aproximadamente 0, esta impresión se da debido a la escala. Aunque a comparación del COVID, es realmente baja.

5. Trabajos Futuros

A partir del trabajo realizado, podemos distinguir diferentes puntos que permiten ampliar lo desarrollado.

- Expansión del simulador.

Poder desarrollar grafos dinámicos, donde los contactos no sean siempre los mismos entre los nodos. Esto permitiría simular por ejemplo, diferencia de movilidad, y como esto impacta en la expansión de una pandemia.

Vacunación, es decir, modelar la existencia de una vacuna, que automáticamente genera una respuesta de inmunidad (estado Recuperado), en el nodo.

Concurrencia. Es decir, poder modelar el tiempo de contacto entre cada uno de los nodos. Esto permitiría modelar la ventana de tiempo de contagio de una persona que está incubando el virus.

Asintomáticos. Modelar asintomáticos podría aumentar la precisión del simulador. Ya que uno de los mayores problemas del Sars-Cov-2 es la presencia de asintomáticos, pues estos últimos siguen transmitiendo el virus.

- Análisis utilizando el simulador.

Ajustar con mayor precisión los datos a una población e intentar comparar los datos con una muestra real, y poder predecir según variables como evolucionaría la epidemia.

6. Conclusión

A lo largo del trabajo práctico tratamos el desarrollo de un simulador de epidemias sobre redes de personas representadas por grafos y la utilización del mismo para el análisis de diferentes tipos de epidemias. En particular, modelamos la epidemia producida por el Sars-Cov-2, para esto buscamos los parámetros necesarios para que el programa simule tal virus. Además, buscamos variar aspectos como el modelo de epidemia, el porcentaje inicial de personas o nodos incubando, el número de contactos, para luego variar también el tipo de grafo y los diferentes parámetros que modelan el grafo.

Del primer experimento 4.1, pudimos concluir que el factor con mayor peso en una epidemia es su tipo, siendo SIS y SIRS aquellos con mayor tasa de mortalidad y contagios debido a las reiteradas veces donde un nodo puede estar en el estado Infectado. Además pudimos observar la dinámica de cada tipo de epidemia, siendo SIR representada por un pico de contagios, SIRS con ciclos disminuyendo su amplitud y finalmente SIS convergiendo a un porcentaje grande de infectados graves. Otra de las conclusiones que pudimos observar a partir del experimento, fue el rol de los contagios graves, que muestran una parte de la dinámica de la epidemia. Es decir, cómo luego de los picos de contagios los infectados leves se recuperan y los que restan son los infectados graves. Este último grupo por un lado aumenta la cantidad de nodos muertos pero también son los nodos que, debido a su larga duración en estado infectado, siguen infectando a los nodos en estado Susceptible y por lo tanto comienzan nuevas olas de la epidemia. Luego, buscamos ver para cada tipo de epidemia, como varía la evolución de esta modificando por un lado el grado de los nodos y por otro las probabilidades de que un nodo empiece infectado. Para los grados, concluimos que situaciones extremas donde el grado es 1 (esto podría ser equivalente a medidas como cuarentenas estrictas) afectan directamente a la evolución de la epidemia, pudiendo verse reducida a 0 contagios en muy poco tiempo. Para grados mayores a 1 ya la epidemia evoluciona similar, con pequeñas diferencias de amplitud en la cantidad de infectados por unidad de tiempo. Finalmente, obtuvimos que el porcentaje inicial de infectados no es determinante en la evolución de la epidemia, sí lo son medidas como el distanciamiento social (disminuyendo el grado), o características del virus que modelan la epidemia.

En lo que respecta a los resultados del experimento 2, en donde estudiamos la evolución de la enfermedad para los distintos tipos de redes asumiendo que esta se comporta de acuerdo al modelo SIR podemos destacar los siguientes puntos:

- **Scale Free:** Se llega rápidamente a un pico de contagios independientemente del porcentaje inicial de nodos infectados, además un porcentaje de la población no llega a infectarse debido a la baja conectividad de una gran porción de los nodos. Por último se comprobó que con vacunar a un 0,3 % de los nodos, eligiendo aquellos de mayor grado, se evita la infección casi por completo.
- **Small World:** Valores de k (cantidad de vecinos) menores a 4 generan una curva de contagios mas lenta o aplanada y luego de la última iteración aún continúan apareciendo nuevos contagios. Con $k = 4$ la velocidad de contagios crece exponencialmente, por otro lado aumentar p (probabilidad de conectar a los nodos aleatoriamente) solo desplaza la curva mas hacia la derecha.
- **Balanced Tree:** Aumentar el branching factor impacta inversamente en la velocidad de propagación de la infección y en el porcentaje de nodos infectados.
- **Random:** Para m (grado de los nodos) ≥ 3 el 100 % de los nodos se infectaron mientras que para $m = 2$ y un porcentaje inicial de infectados de 0.1 y 0.01 se contagiaron respectivamente el 34 y el 5 % del total, con picos mas tardíos.
- **Grid:** El porcentaje de muertos e infectados no varió al aumentar el porcentaje de nodos infectados inicial pero esto si impactó en la velocidad y en el pico de los contagios creciendo proporcionalmente a medida que aumentamos el porcentaje inicial de infectados.

Finalmente, en el tercer experimento 4.3 utilizando de la misma manera los parámetros del virus de la Gripe y del Covid-19, buscando rechazar el dicho escuchado durante la epidemia del Coronavirus de que el Covid-19 es como una gripe. Si bien como primer resultado obtuvimos que la cantidad de infectados en ambas es similar, la mayor diferencia (y la más preocupante) es la duración de los infectados graves. Pensando en una población sin ninguna medida sanitaria, los infectados graves por Covid-19, debido a su contagiosidad durante mayor tiempo provocarían más olas de contagio y por ende, las probabilidades

que algún nodo se contagie aumenta. Esto último tiene como consecuencia que las chances de que un nodo evolucione al estado infectado grave aumente y por lo tanto, aumenta la tasa de letalidad, siendo el Coronavirus más letal que el virus de la Influenza.

6.1. Apéndice 1

Demostración 1

$$p = \sum_{i=0}^{n-1} p' (1-p')^i$$

$$\frac{p}{p'} = \sum_{i=0}^{n-1} (1-p')^i$$

$$\frac{p}{p'} = \frac{1 - (1-p')^n}{1 - (1-p')}$$

$$\frac{p}{p'} = \frac{1 - (1-p')^n}{p'}$$

$$p = 1 - (1-p')^n$$

$$1-p = (1-p')^n$$

$$\ln(1-p) = n \ln(1-p')$$

$$\frac{\ln(1-p)}{n} = \ln(1-p')$$

$$1 - e^{\frac{\ln(1-p)}{n}} = p'$$

Referencias

- [1] *Age Structure*. URL: <https://www.cia.gov/library/publications/the-world-factbook/fields/341.html>.
- [2] El Gato y la Caja. *Coronavirus: lo que tenés que saber*. URL: <https://elgatoylacaja.com/noticias/coronavirus-lo-que-tenes-que-saber/>.
- [3] *Clinical Signs and Symptoms of Influenza*. URL: <https://www.cdc.gov/flu/professionals/acip/clinical.htm>.
- [4] *Directed Scale-Free Graphs*. URL: <https://www.microsoft.com/en-us/research/uploads/prod/2016/11/Directed-scale-free-graphs.pdf>.
- [5] *Estimating Influenza Disease Burden from Population-Based Surveillance Data in the United States*. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0118369>.
- [6] *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs [Internet]*. URL: <https://www.ncbi.nlm.nih.gov/books/NBK63484/>.
- [7] *Influenza*. URL: <https://medbroadcast.com/condition/getcondition/influenza>.
- [8] Tomas Jaratz Julian Len. *Simulador*. URL: <https://github.com/tomyjara/tprse>.
- [9] David Easley. Jon Kleinberg. «Networks, Crowds, and Markets: Reasoning about a Highly Connected World». En: ().
- [10] Julian Len. *Simulador*. URL: https://github.com/tomyjara/tprse/tree/exp_1.
- [11] Julian Len. *Simulador*. URL: <https://github.com/tomyjara/tprse/tree/exp3>.
- [12] Ken T. D. Eames Matt J. Keeling. *Networks and Epidemic Models*. DOI: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsif.2005.0051>.
- [13] *Novel Coronavirus(2019-nCoV)*. URL: <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200127-sitrep-7-2019--ncov.pdf>.
- [14] *Percentage of COVID-19 cases in the United States from February 12 to March 16, 2020 that resulted in hospitalization, by age group*. URL: <https://www.statista.com/statistics/1105402/covid-hospitalization-rates-us-by-age-group/>.
- [15] *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. OMS. URL: <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>.
- [16] A. Steger y N.C. Wormald. *Generating random regular graphs quickly*. URL: <https://pdfs.semanticscholar.org/a11c/daa94e777b0d9752a326224f742aa3f71c3b.pdf>.
- [17] *Trabajo simulacion*. URL: <https://enys.conicet.gov.ar/coronavirus-que-predicen-los-cientificos-sobre-el-pico-de-la-pandemia-en-la-argentina/>.