

75.26 Simulación

TP final:

**Estimating the number of
SARS-CoV-2 infections and
the
impact of social distancing in
the United States**

Santiago Nicolás Risaro Sesar - 84623

1º cuatrimestre 2020

Índice

Resumen	3
Introducción	4
Modelo	6
SIR con cuarentena	6
Conteo de muertes por día	6
Definición de θ y p	7
Interacción entre el modelo SIR y la estimación de muertes	9
Simulación	11
Metrópolis-Hastings	11
Metrópolis adaptativo	11
Valores a priori de los parámetros	12
Resultados	12
Nota	16
Dificultades de la implementación	17
Trabajo a futuro	18
Código fuente	19
Bibliografía	20

Resumen

Conocer la cantidad de personas infectadas por SARS-CoV-2 es muy importante para diseñar políticas de cuidado que permitan frenar el avance de la enfermedad y evitar la saturación del sistema hospitalario.

En el trabajo analizado [1] se muestra que, en Estados Unidos, el número de personas infectadas reportado por las autoridades es inferior al de personas realmente infectadas y se estudia cuán efectivas resultan las medidas de aislamiento implementadas.

Se utiliza un modelo bayesiano de las muertes producidas por esta enfermedad para determinar los parámetros de las ecuaciones diferenciales que modelan la epidemia y así poder determinar la cantidad de infectadas real.

Adicionalmente este modelo permite validar la efectividad de las medidas de aislamiento y estimar el momento en el que se produjo el primer contagio en la región estudiada.

Introducción

A fines del año 2019 se detectó en Wuhan, China, un brote de una nueva cepa del virus SARS-CoV que afectaba a los humanos, el SARS-CoV 2. A la enfermedad causada por este virus se la denominó COVID-19 (Corona Virus Disease 2019).

En enero del 2020 se confirmó que este virus puede transmitirse entre humanos [2], pocos días después la Organización Mundial de la Salud (OMS) declaró la emergencia sanitaria internacional. Rápidamente el virus se esparció por el mundo y el 11 de marzo se categorizó a esta enfermedad como una pandemia por su alcance global; en ese momento había contagios reportados en 114 países ¹

En Estados Unidos el primer caso se reportó el 20/01/2020, en Washington, a partir de ese momento comenzó a esparcirse por todo el país.

La velocidad con la que se incrementan los casos, la escasez de recursos para realizar tests y el tiempo que lleva obtener los resultados pueden llevar a que el número de casos reportados por cada estado no se corresponda con el número real de personas infectadas. Si, además, se considera que muchas personas no presentan síntomas y que por ese motivo no se realizan un examen hay más razones para creer que el número de casos está subestimado.

Por otro lado, debido a la alta tasa de contagios, y a la imposibilidad de ofrecer una solución mejor para controlar la enfermedad, la mayoría de las ciudades comenzaron a aplicar medidas de distanciamiento social o cuarentenas.

En este paper se intentan determinar dos cosas (1) cuántas personas infectadas hay realmente y (2) cuán efectivas son las medidas de aislamiento para frenar el avance de la enfermedad.

Asumiendo que la información de personas contagiadas en cada estado no es confiable esta tarea es bastante compleja. Sin embargo, el número de personas fallecidas a causa de esta enfermedad es un dato más confiable pues son las que presentan mayores síntomas y, en general, fueron hospitalizadas por lo tanto hay más chances de que les hayan realizado un test para detectar el virus.

Los y las autoras del paper partieron de un modelo SIR (cantidad de personas susceptibles, infectadas y recuperadas a lo largo del tiempo) clásico y lo modificaron de manera tal que permitiera mostrar el efecto de la cuarentena. Adicionalmente implementaron una manera de relacionar la cantidad de muertes a lo largo del tiempo con la cantidad de personas infectadas utilizando como base la información de Wuhan, Corea del Sur y el crucero Diamond Princess por ser lugares en los que se llevó a cabo un testeo masivo y un registro exhaustivo de la

¹WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020 [<https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>]

información. En base a esto se estima una tasa de mortalidad con respecto a los contagios y el tiempo que pasa desde que una persona se contagia hasta que muere.

Utilizando la información oficial de muertes por día en distintos estados se busca qué parámetros del modelo SIR mejor ajustan a la cantidad de infectadas por día esperadas para que se produzcan las muertes observadas.

Modelo

SIR con cuarentena

El modelo de este trabajo posee 4 parámetros principales β , γ , ϕ , y T_0 .

T_0 indica el día en el que se infectó la primera persona en la región.

$\{\beta, \gamma\}$ son los parámetros de un modelo SIR básico (sin variaciones demográficas y en el que las personas recuperadas se inmunizan) y representan la tasa de contagios y de recuperación.

El parámetro ϕ indica cuán efectiva es la medida de aislamiento adoptada; $0 < \phi < 1$, si $\phi = 0$ la medida es 100% efectiva y nadie más se contagia, por lo tanto la enfermedad desaparece, por el contrario $\phi = 1$ indica que el aislamiento no afecta para nada el curso de la enfermedad.

Las ecuaciones diferenciales de este modelo SIR propuestas son

$$\begin{aligned}\frac{\partial S_t}{\partial t} &= \begin{cases} -\beta S_t I_t N^{-1} & t < T_1 \\ -\phi \beta S_t I_t N^{-1} & t \geq T_1 \end{cases} \\ \frac{\partial I_t}{\partial t} &= \begin{cases} \beta S_t I_t N^{-1} - \gamma I_t & t < T_1 \\ \phi \beta S_t I_t N^{-1} - \gamma I_t & t \geq T_1 \end{cases} \\ \frac{\partial R_t}{\partial t} &= \gamma I_t,\end{aligned}$$

En las que S es la cantidad de personas Susceptibles, I las Infectadas, R las recuperadas y N la población total de la zona estudiada.

T_1 es una constante que indica en qué momento se adoptaron medidas de distanciamiento. Para este estudio se considera que el aislamiento comenzó el día en que tanto restaurantes como escuelas cerraron²

Conteo de muertes por día

Para este trabajo es imprescindible modelar de la mejor manera posible las muertes provocadas por la enfermedad dado que ese es el dato que se toma como fuente de verdad. El modelo propuesto depende de la tasa de fatalidad, la cantidad de días que pasan desde que la

² <https://github.com/COVID19StatePolicy/SocialDistancing>

persona se contagia hasta que fallece (dado que la persona falleció) y la cantidad de personas que se infectan por día.

Se define la variable p como la tasa de fatalidad de la enfermedad. ν_t indica la cantidad de nuevas infectadas en el día t . Dado $\theta = \{\theta_s : s = 0, 1, \dots, m\}$ el conjunto de probabilidades que definen el tiempo desde el contagio hasta la muerte, θ_s es la probabilidad de morir S días después del contagio dado que la persona falleció. Definimos $X(t, t')$ como la cantidad de infectadas en el día t que mueren el día t' . El modelo se comporta de la siguiente manera

$$X(t, t') \mid p, \theta \sim \text{Poisson}(p\nu_t\theta_{(t'-t)}).$$

Y la cantidad de muertes observadas en el día r es

$$D(r) = \sum_{t=1}^r X(t, r),$$

Que tiene la siguiente distribución

$$D(r) \sim \text{Poisson} \left(p \sum_{t=1}^r \nu_t \theta_{(r-t)} \right).$$

Se eligió una distribución de Poisson en vez de una binomial pues permite valores reales en vez de naturales lo que simplifica algunos cálculos y porque para este conjunto de parámetros la distribución de Poisson es una buena aproximación de la binomial (n es “grande” y p es “pequeño”).

Definición de θ y p

Estos dos son los parámetros más importantes del sistema pues se utilizan para validar cuán certera es la simulación con respecto a las muertes observadas. Cabe recordar que para este modelo la fuente de verdad son las muertes por día y se estima que cantidad de contagiados debería haber para llegar a esa cantidad de fallecimientos por día.

El parámetro p indica cuantas personas infectadas de COVID-19 fallecen. En el paper indican que es el más complejo de estimar y que no lograron encontrar un único valor que los satisfaga por eso tomaron los 3 valores más probables de acuerdo a varios estudios. Según Russell et al [3] que tomaron datos del crucero Diamond Princess estimaron este valor en 1.2%. En base a datos de Corea del Sur (que se consideran más confiables por su política de testeos masivos) se obtiene $p=1,4\%$. Otro estudio [4] calcula que, para China, esta tasa es del 0.4% . Un paper [5] reporta que en Wuhan es del 1.4%. Hay que tener en cuenta que diferencias demográficas y

en la infraestructura hospitalaria de las regiones sumado a otros factores como la interacción de la COVID-19 con otras enfermedades pueden alterar este número. Por estos motivos los y las autoras del paper decidieron tomar 3 valores posibles para p 0.5%, 1% y 1.5% y realizar las simulaciones 3 veces variando este parámetro.

Para calcular las probabilidades de θ se suma el tiempo de incubación de la enfermedad y el tiempo que pasa desde los primeros síntomas hasta la muerte y eso da el tiempo aproximado más probable desde el contagio hasta la muerte (siempre tomando en cuenta los casos de la gente que fallece por esta enfermedad).

El tiempo de incubación fue calculado por Lauer et al. [6] con una mediana de 5.1 días con un percentil 99 de 14 días. Una distribución Poisson-Gamma de parámetros $\{\alpha, \beta\} = \{5.5, 1.1\}$ se ajusta bien a estos valores.

Para el tiempo desde los síntomas hasta la muerte se basaron en el estudio de Zhou et al. [7] que calcularon una mediana de 18.5 días con un rango intercuantil de 15 a 22 días que se asemeja bastante a una Poisson-Gamma $\{\alpha, \beta\} = \{27.75, 1.5\}$.

Para obtener las probabilidades de muerte a los s días del contagio θ_s se toman 100000 muestras de cada una de estas distribuciones, se las suma y se cuenta la cantidad de ocurrencias de cada valor; la frecuencia de cada valor obtenido se divide por el total de muestras y se obtiene la probabilidad asociada a dicho valor.

En la figura 1 se muestra un histograma normalizado de una realización de dicha simulación, el valor promedio es de 23 días, bastante cercano a los 22.6 días que resultan de sumar las medianas informadas por los dos estudios mencionados, con una probabilidad asociada de un poco más del 6%. Esto se lee “Si una persona fallece por COVID-19 tiene un 6% de probabilidad de hacerlo 23 días después de contagiarse”.

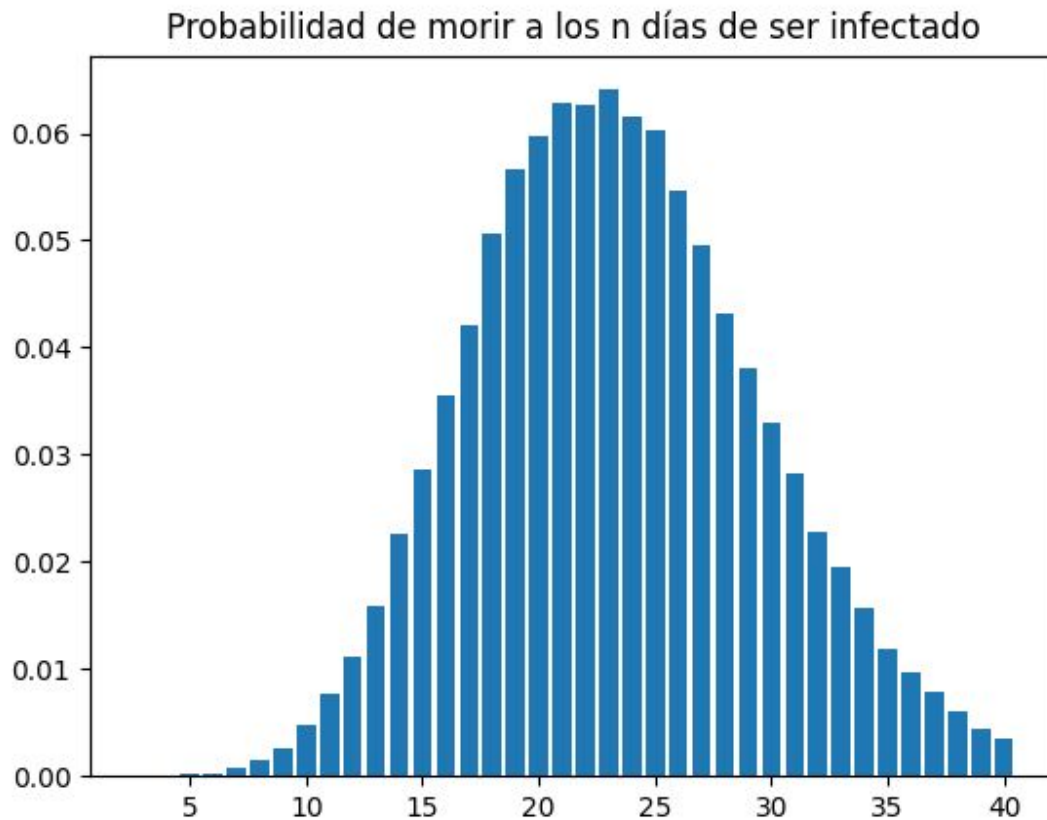
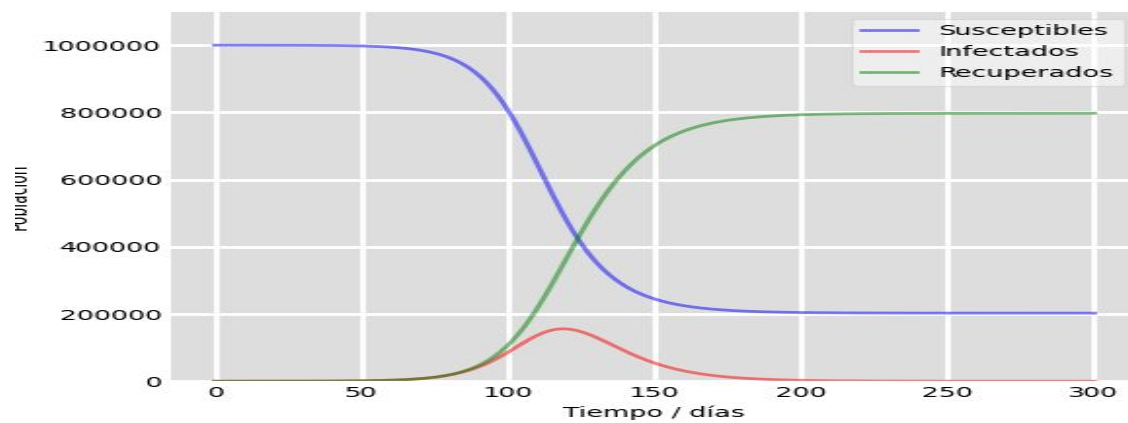
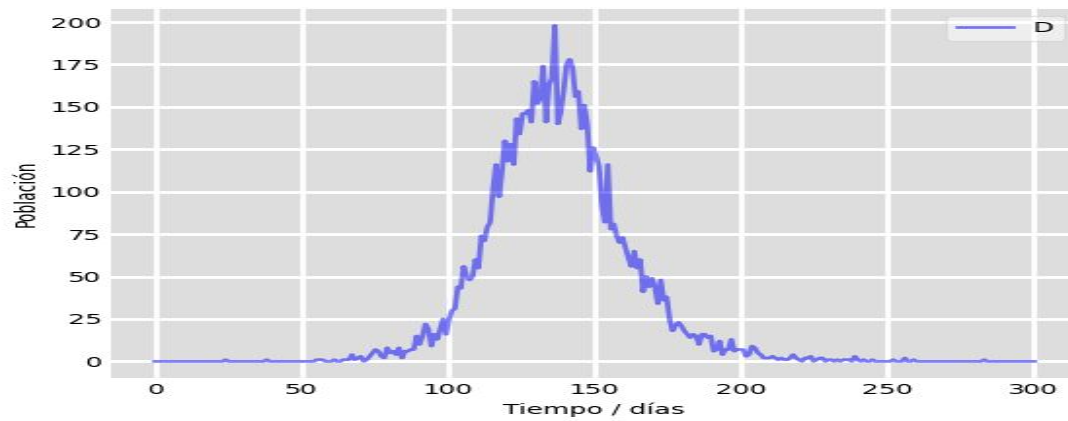
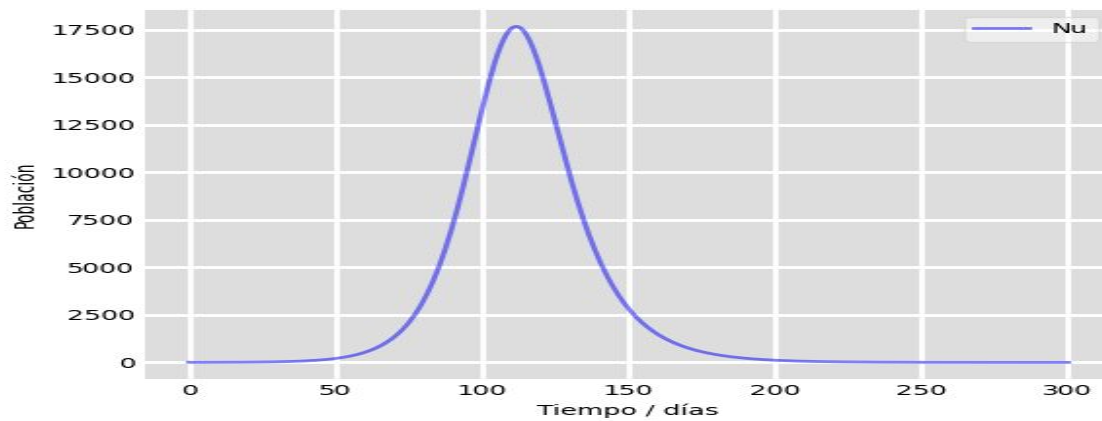


Figura 1. Tomando en cuenta que la personas fallece por COVID-19, qué probabilidad tiene de hacerlo s días después de contagiarse.

Interacción entre el modelo SIR y la estimación de muertes

Para comprobar el modelo y validar cómo interactúan los distintos parámetros se realiza un modelo SIR iterativo en el que para cada día se van contando cuántas muertes se producirán a futuro y se las incluye en el conteo de los recuperados. También se contabiliza día a día la cantidad de nuevas personas infectadas.

Una realización de esta simulación genera los siguientes gráficos



Gráficas de S, I, R, Nu y D respecto al tiempo

Lo más destacable de este gráfico es observar que la curva de D (muertes por día) tiene la misma forma que la de Nu (contagios por día) pero con un retraso.

Simulación

La intención es hallar los valores que parametrizan el comportamiento de la enfermedad a partir de uno de sus efectos secundarios (en este caso los fallecimientos). Para ello se realizan simulaciones con distintos valores para dichos parámetros; luego se calculan las muertes que se producirían con esos contagios simulados, utilizando las probabilidades de θ y p descritas en la sección anterior y se rechaza o acepta la propuesta si la verosimilitud (likelihood) calculada es mayor a una variable aleatoria uniforme entre 0 y 1.

En este caso se utilizó un algoritmo Metrópolis adaptativo para realizar las simulaciones y la aceptación o rechazo de los parámetros.

Metrópolis-Hastings

Metropolis-Hastings es un algoritmo de la familia de algoritmos MCMC (Markov Chain Monte-Carlo). Es una cadena de Markov pues el estado en cada iteración de la simulación depende del estado en la iteración anterior.

Se parte de un valor a priori para cada parámetro y se calcula la verosimilitud para esos parámetros; luego se propone un nuevo conjunto de parámetros se recalcula la verosimilitud y se los acepta o rechaza si la verosimilitud nueva es mayor a la anterior o si la verosimilitud nueva es mayor a un valor aleatorio. Si se acepta se actualizan los parámetros si no se mantienen los parámetros anteriores. [8]

Metrópolis adaptativo

El algoritmo Metrópolis adaptativo es una variación al algoritmo anterior propuesta por Hario et al [9]. Este algoritmo calcula la covarianza de la distribución propuesta utilizando la memoria de todas las propuestas previamente realizadas. En este sentido no es estrictamente una cadena de Markov, pero los autores demuestran en su paper que se mantienen propiedades como la ergodicidad.

Este algoritmo propone una fase de adaptación en la que la covarianza se calcula sólo con unos valores iniciales, es una constante. Luego se comienza a calcular con los valores acumulados.

Para proponer el nuevo set de parámetros se calcula una normal multivariada con media en los parámetros de la iteración anterior y con la covarianza calculada en base a todas las iteraciones anteriores.

Una de las diferencias de este algoritmo con M-H es que en caso de rechazo de una propuesta se guarda la propuesta anterior como propuesta actual. Esto lleva a que en el resultado final haya tantas propuestas como iteraciones y algunas de ellas con valores repetidos.

En el paper estudiado se implementa este algoritmo y se realizan 50000 iteraciones para cada estado estudiado. En cada iteración la nueva propuesta es evaluada utilizando una densidad proporcional a

$$\ell(\nu(\eta), \eta) = \sum_{t=1}^T p(D(t) | \nu) + \log(\pi(\eta)),$$

Donde $p(D(t)|\nu)$ es la probabilidad de que exista una cantidad determinada de muertes para la cantidad de infecciones diarias y $\pi()$ representa la densidad a priori.

Valores a priori de los parámetros

Para este tipo de algoritmos se debe proponer un estado inicial. Cuanto más cercanos sean estos valores a los valores reales de los parámetros mejor ya que el algoritmo puede converger más rápidamente.

Adicionalmente debe indicarse que distribución posee cada parámetro para acotar los valores que propone el algoritmo.

Basándose en los estudios de Ferguson et al. [10] y Prem et al. [11] se propone para γ un valor inicial de 6 días y la distribución propuesta es una normal truncada entre 3.4 y 9.4 con media 6.4 y desvío de 1.5. Para β se propone un valor inicial de $1/\gamma \cdot R_0$ con $R_0 = 2.3$ y la distribución, que depende de γ es una normal truncada entre γ y 4γ con media 2.5γ y desvío 1.5γ .

T_0 , el día del primer contagio, se distribuye uniformemente en el rango del 1/1/2020 al 20/2/2020 con valor inicial 30/1/2020.

ϕ , la efectividad de las medidas de aislamiento, comienza en 0.4 y tiene una distribución uniforme entre 0.01 y 0.99

Resultados

Replicando lo realizado por los y las autoras del paper, tomamos la información de casos provista por el New York Times al 18/04/2020³.

³ <https://github.com/nytimes/covid-19-data>

Se ejecutaron 50000 iteraciones de la simulación, comenzando con los valores indicados, y recolectamos los valores propuestos para cada parámetro en cada iteración.

Para determinar cuál es el valor más probable de cada parámetro descartamos los primeros 10000 valores calculados (corresponden al período de adaptación del algoritmo) y calculamos el promedio de cada uno. Se obtuvieron los siguientes resultados

	$p=0.015$	$p=0.01$	$p=0.005$
β	0.3573	0.36874	0.4082
γ	0.1623	0.1656	0.1654
T_0	26	26	35
φ	0.6174	0.6146	0.5526

Esto da, por ejemplo para la primera columna, un R_0 de 2.20 mientras que en el paper indican que obtuvieron un R_0 promedio de 2.28.

Un detalle a notar es que esta simulación calcula T_0 (la fecha del primer contagio) en 26, es decir el 26/1/2020, mientras que, según la información oficial, el primer caso reportado fue el 25; esto es interesante porque, si bien las fechas son cercanas, es poco probable que la persona se haya realizado un testeo al poco tiempo de infectarse ya que los síntomas tardan varios días en presentarse y en etapas iniciales de la enfermedad no se realizaban muchos análisis. Sin embargo, este valor es consistente con la primera muerte reportada que fue el 4/3/2020. Recordemos que este modelo calcula que el tiempo desde el contagio hasta la muerte es de 23/24 días en promedio con valores máximos de 40 días.

Otro resultado destacable es que el valor de φ obtenido es 0.6, de acuerdo al modelo propuesto un valor de 1 indica que la cuarentena no es efectiva para nada. Según esto la cuarentena está más cerca de ser inútil que una buena herramienta para prevenir contagios. Esto podría deberse a los pocos datos con los que se realizaron las simulaciones; el 18/4/2020 iban 35 días de cuarentena en el estado estudiado.

Para validar el modelo realizamos una comparación entre los valores de contagios reportados y los que se obtienen de las iteraciones de la simulación, utilizando la misma mecánica para

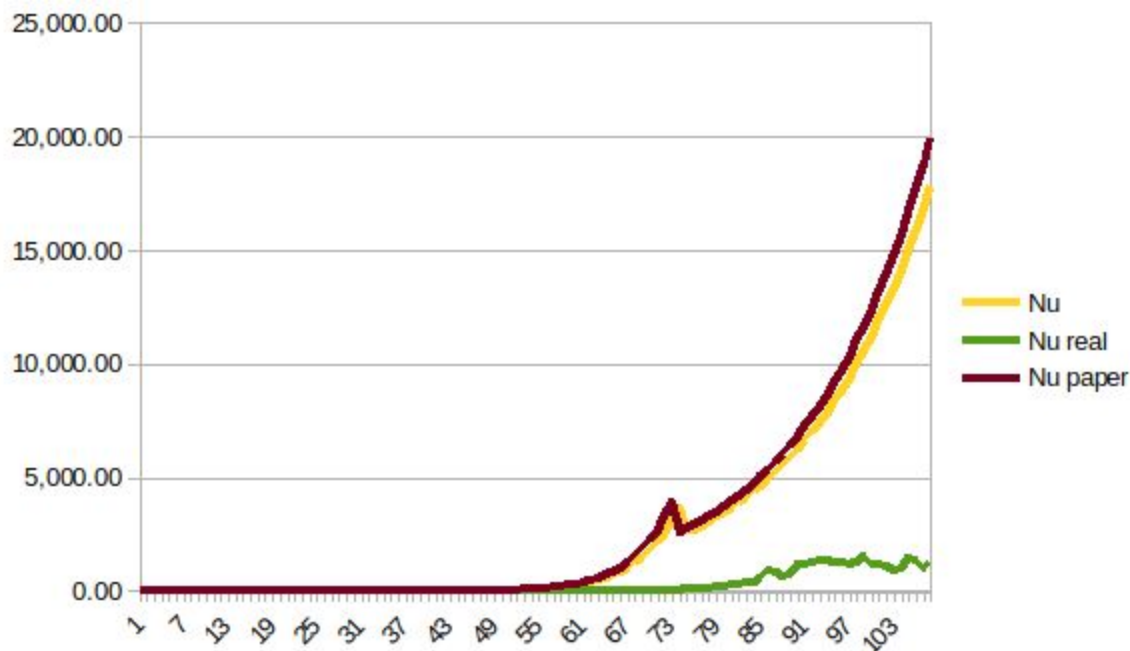
obtener los promedios de nuevas infecciones por día que la que se utilizó para calcular los parámetros, y obtuvimos la siguiente gráfica.



Casos reportados vs simulados

Si bien una diferencia era esperable dadas las premisas en las que se basa el trabajo, una diferencia tan grande no parece correcta. En uno de los gráficos del paper se muestra el promedio de “subconteo” por día y está cerca de 10. Cuando aquí, en su pico, es varios órdenes de magnitud mayor.

Esto nos hace sospechar que hay un error en la implementación de nuestro algoritmo. Revisamos el código, encontramos algunos bugs, volvimos a correr las simulaciones luego de corregirlos... pero los resultados fueron similares. Comenzamos a revisar el código provisto junto al paper y observamos una curva similar a la nuestra.

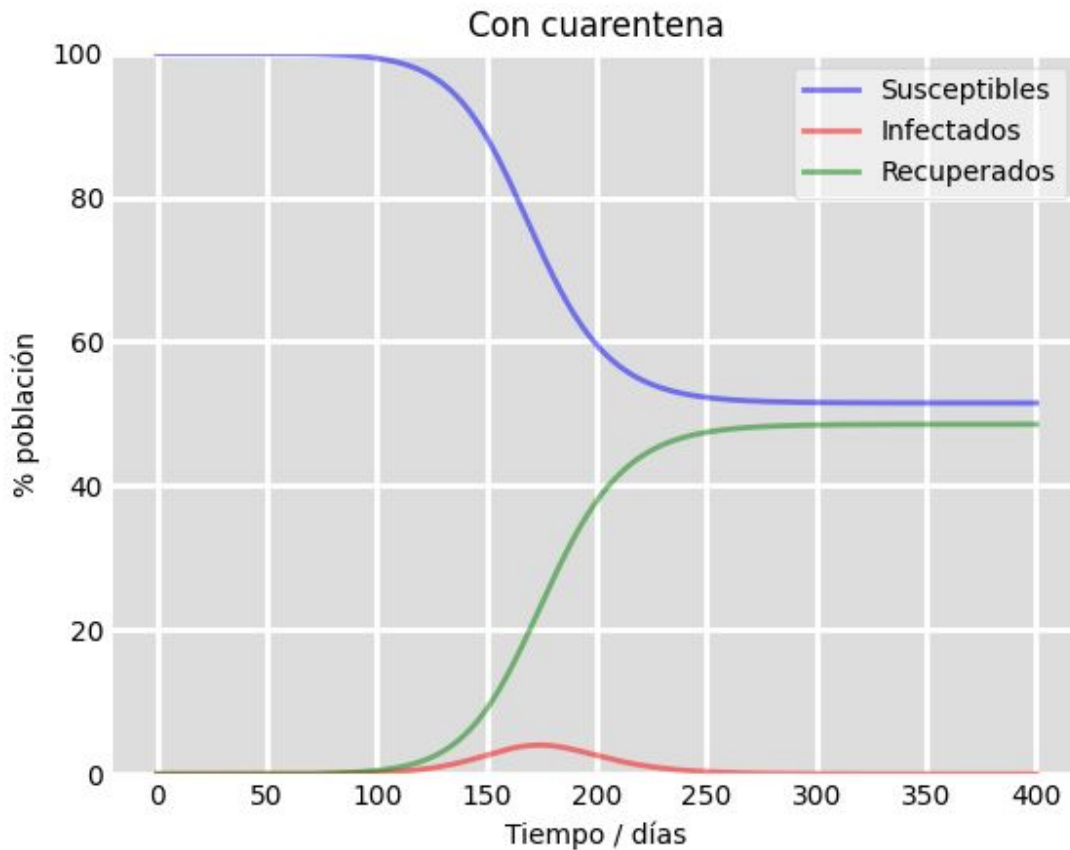


Comparación entre el Nu calculado en el paper, el real y el de nuestras simulaciones para una de las iteraciones

En el código del paper, para obtener el subconteo, realizan algunos cálculos que nuestro desconocimiento de R no nos permitió comprender. Probablemente algo de eso explique la diferencia pero no logramos entender bien esa parte del código y en el texto del paper no se dan mayores explicaciones.

Otra cosa que llama la atención es la caída abrupta en la cantidad de contagios al iniciarse la cuarentena (día 73), lo intuitivo sería ver que la curva baja con un poco más de suavidad y que la pendiente de la curva de contagios luego de implementada la cuarentena es un poco menos pronunciada. Sin hacer los cálculos parece ser muy parecida antes y después de la implementación de las medidas de aislamiento.

Luego de obtener estos valores hicimos una simulación del modelo SIR propuesto por las y los autores utilizando los parámetros calculados por el algoritmo MCMC y obtuvimos la siguiente gráfica.



Simulación de SIR con los parámetros obtenidos de la simulación Metrópolis Adaptativa

Aquí se observa que el porcentaje de personas infectadas se acerca a cero alrededor del día 250 del año (6/9/2020). El 2 de septiembre el estado de California informa 5407 casos nuevos de coronavirus mientras que esta simulación “predice” unos 18000 casos diarios para esa fecha.

Nota

En el paper hicieron las simulaciones para cuatro estados (California, Nueva York, Florida y Washington) y para cada uno de los 3 valores de p propuestos. Aquí sólo se muestran los resultados para la simulación de California por brevedad, tiempo y por creer que sumarlos no aporta mucho al análisis realizado.

Dificultades de la implementación

La primera dificultad con la que nos topamos fue la implementación del algoritmo Metrópolis adaptativo. Encontramos un paquete para python⁴ que lo implementa pero su complejidad, la falta de documentación y las particularidades de la simulación nos hicieron desistir de su uso y optar por programarlo completo.

Lo cual nos lleva al siguiente problema, si bien el algoritmo tiene algunas sutilezas no es particularmente complejo, pero había algunos detalles mencionados en el paper que no nos cuadraban del todo y que con la sola lectura del texto no lográbamos descifrar. Por suerte en el paper se provee el link a un repositorio de código en R con el que se hicieron las simulaciones. Ver este código ayudó a lograr que los valores obtenidos se acercaran más a los mencionados en el paper pero también nos sesgó y puede haber hecho que no exploremos otras alternativas.

Por otro lado, comparando algunos valores obtenidos con los del paper se observan pequeñas diferencias que surgen de distintas implementaciones de R y de las bibliotecas de python utilizadas. Por ejemplo, en el cálculo de las probabilidades de theta, las funciones de random de python no retornan 0, 1 y 2 pero en R sí lo hacen, aunque con una probabilidad muy baja. Esto no es un gran problema pero explica algunas discrepancias en los resultados que obtuvimos con respecto a los del trabajo estudiado.

⁴ <https://pypi.org/project/pymcmcstat/>

Trabajo a futuro

Consideramos que el punto planteado por los y las autoras en el paper es interesante y que merece una segunda revisión. Por un lado para comprender por qué el conteo de casos por día nos da una diferencia tan grande y por otro lado para probar algunas variaciones como por ejemplo:

- Ejecutar la simulación con información de casos más actualizada.
- Buscar valores actualizados de p y θ . Ahora que pasó más tiempo probablemente haya mejor información y más confiable sobre la tasa de fatalidad y el tiempo hasta la muerte de las personas contagiadas.
- Buscar valores actualizados para los límites de β y γ . Lo mismo aplica para los parámetros a priori del modelo SIR. Seguramente al día de hoy se cuenta con más información que la citada en el paper.
- Incluir en el modelo información sobre las reaperturas. Ahora que ha pasado mucho tiempo desde iniciados los aislamientos varias ciudades comenzaron un esquema de reaperturas. Si en la simulación se consideran más días probablemente esto afecte a los resultados.
- Otro camino que sería interesante explorar es aplicar este algoritmo para datos de Argentina o de la Ciudad de Buenos Aires. Existe un repositorio⁵ de un grupo de estudios de la UTN con bastante información. Por cuestiones de tiempo no hicimos pruebas con esos datos, pero es técnicamente factible.

⁵ <https://github.com/GIBD/covid>

Código fuente

El código para estas simulaciones se encuentra en el siguiente repositorio

<https://github.com/nickrisaro/simulacion-tp-final>

Está programado y probado con python 3.6.9 y requiere de las bibliotecas **scipy** y **statsmodels**. Todas las dependencias necesarias están listadas en el archivo requirements.txt, pueden instalarse utilizando `pip install -r requirements.txt`

Para ejecutar la simulación se debe correr el script `MCMC.py`. La ejecución demora unos 3 minutos (en un i5 con 16GB de RAM) y va informando su avance durante el proceso. Al finalizar guarda los resultados en la carpeta `salida/mcmc` e informa algunos valores por consola.

Para la gráfica de theta se puede ejecutar el script `Theta.py`. El gráfico se guarda en la carpeta `salida`.

La simulación del modelo SIR con el cálculo de muertes por día se puede realizar ejecutando `SIR.py`.

La comparativa de casos por día se logra con el script `Comparar_casos_por_dia.py`.

La simulación del modelo SIR propuestos en el paper se puede realizar con el script `SIR_Cuarentena.py`, si detecta que se ejecutó `MCMC.py` toma los valores de entrada del resultado de dicha simulación.

Bibliografía

[1] Johndrow, J., Lum, K., Gargiulo, M., Ball, P. (2020). Estimating the number of SARS-CoV-2 infections and the impact of social distancing in the United States

[<https://arxiv.org/abs/2004.02605>]

[2] Chan J., et al (2020) A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster

[[https://doi.org/10.1016/S0140-6736\(20\)30154-9](https://doi.org/10.1016/S0140-6736(20)30154-9)]

[3] Russell, T. W., Hellewell, J., Jarvis, C. I., van Zandvoort, K., Abbott, S., Ratnayake, R., Flasche, S., Eggo, R. M., Kucharski, A. J., and nCov working group, C. (2020). Estimating the infection and case fatality ratio for covid-19 using age-adjusted data from the outbreak on the diamond princess cruise ship. medRxiv.

[4] Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., and Shaman, J. (2020b). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov2). Science.

[5] Wu, J. T., Leung, K., Bushman, M., Kishore, N., Niehus, R., de Salazar, P. M., Cowling, B. J., Lipsitch, M., and Leung, G. M. (2020). Estimating clinical severity of covid-19 from the transmission dynamics in wuhan, china. Nature medicine, pages 1–5.

[6] Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., Azman, A. S., Reich, N. G., and Lessler, J. (2020). The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: Estimation and application. Annals of Internal Medicine.

[7] Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., et al. (2020). Clinical course and risk factors for mortality of adult inpatients with covid-19 in wuhan, china: a retrospective cohort study. The Lancet.

[8] MacKay, David (2003). Information Theory, Inference, and Learning Algorithms. Cambridge University Press. [Caps 29 y 30] Accedido online en

[<http://www.inference.org.uk/itprnn/book.pdf>]

[9] Haario, H., Saksman, E., Tamminen, J., et al. (2001). An adaptive metropolis algorithm. Bernoulli, 7(2):223–242

[10] Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., et al. (2020). Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand. London: Imperial College COVID-19 Response Team, March, 16.

[11] Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., Davies, N., Flasche, S., Clifford, S., Pearson, C. A., Munday, J. D., et al. (2020). The effect of control strategies to reduce social mixing on outcomes of the covid-19 epidemic in wuhan, china: a modelling study. *The Lancet Public Health*.