

75.26 - Simulación
Primer Cuatrimestre de 2021

Trabajo Práctico 1

Fernando Sinisi	99 139	fsinisi@fi.uba.ar
Tito Lasanta	99 162	tilasanta@fi.uba.ar
Santiago Almejún	99 441	salmejun@fi.uba.ar
Esteban Djeordjian	100 701	edjeordjian@fi.uba.ar

Índice

Ejercicio 1	2
Ejercicio 2	4
2.1 Test Chi cuadrado	4
2.2 Test espectral	5
Ejercicio 3	7
Ejercicio 4	10
Ejercicio 5	14
Ejercicio 6	17
Ejercicio 7	18
Ejercicio 8	26
Ejercicio 9	32
9.1. Sin inmunidad	32
9.2. Con inmunidad	41
Ejercicio 10	48
Anexo: código fuente	57

Ejercicio 1

Un Generador Congruencial Lineal (GCL) es un algoritmo que nos permite obtener una secuencia de números pseudoaleatorios, calculados con una función lineal de la siguiente forma:

$$U_{n+1} = (a * U_n + c) \bmod m$$

Donde a es un multiplicador, c es un incremento, U_n es el número aleatorio generado en la última iteración (en caso de ser la 1ra iteración es la "semilla") y m es el módulo.

Sabemos que gcl_a nos da números pseudoaleatorios entre 0 y $m-1$ por lo tanto si quiero que me devuelva números entre 0 y 1 manteniendo la distribución lo que se hace es dividir todos estos números por m .

Un posible gráfico para visualizar los números es un histograma, y para el mismo necesitamos definir el ancho de los intervalos. Para esto utilizamos la regla de Sturges:

$$a = \frac{(X_{max} - X_{min})}{1 + \log_s(N)}$$

Siendo X_{max} el valor máximo de los números generados, X_{min} el valor mínimo de los números generados y N la cantidad de números generados.

```
Ancho del intervalo para el histograma del punto a: 215485318.88304016
Ancho del intervalo para el histograma del punto b: 0.05017158549349759
Cantidad de contenedores para el histograma del punto a: 19.931568569324174
Cantidad de contenedores para el histograma del punto a: 19.931568569324174
Cantidad de contenedores para la cantidad de muestras dada: 19.931568569324174
```

Como podemos observar, por más que el ancho de los intervalos sean distintos, la regla de sturges nos da la misma cantidad de contenedores para nuestros números, tanto con gcl_a como con gcl_b . Esto se comprueba ya que la regla de sturges definida para la cantidad de contenedores solo depende de la cantidad de elementos en la muestra y en nuestro caso es la misma.

Para los siguientes histogramas usaremos esta regla para calcular los bins (aunque también se puede definir en la función hist sin tener que calcularlo manualmente).

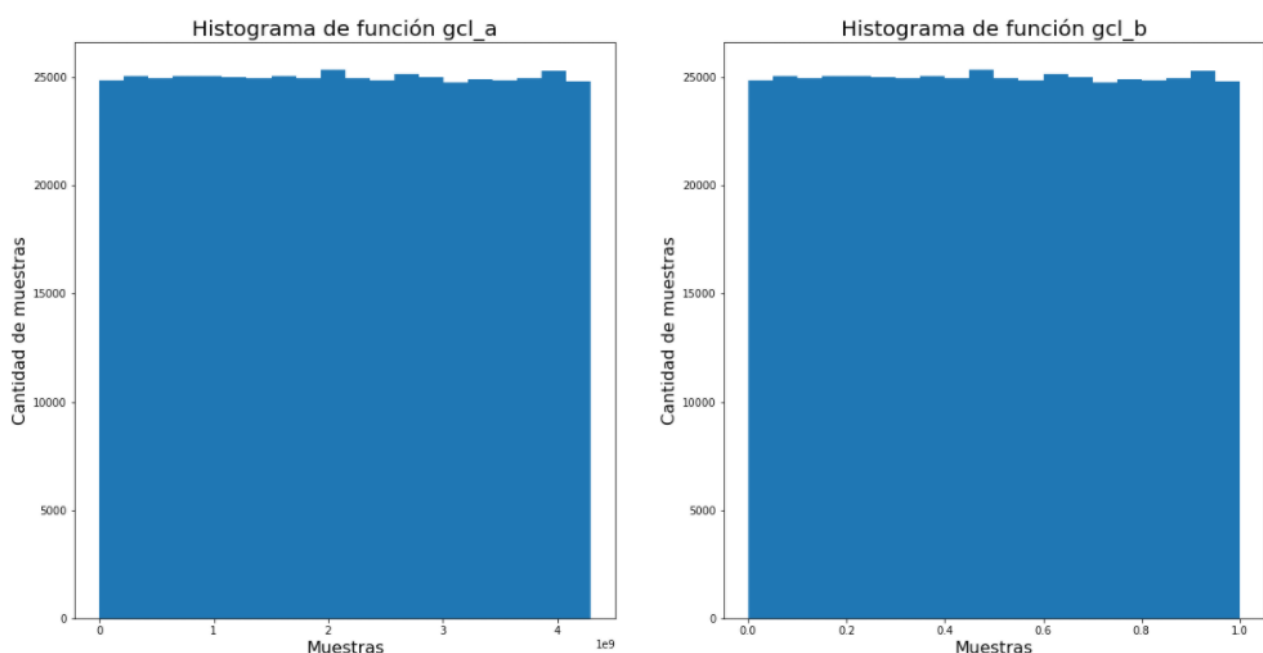
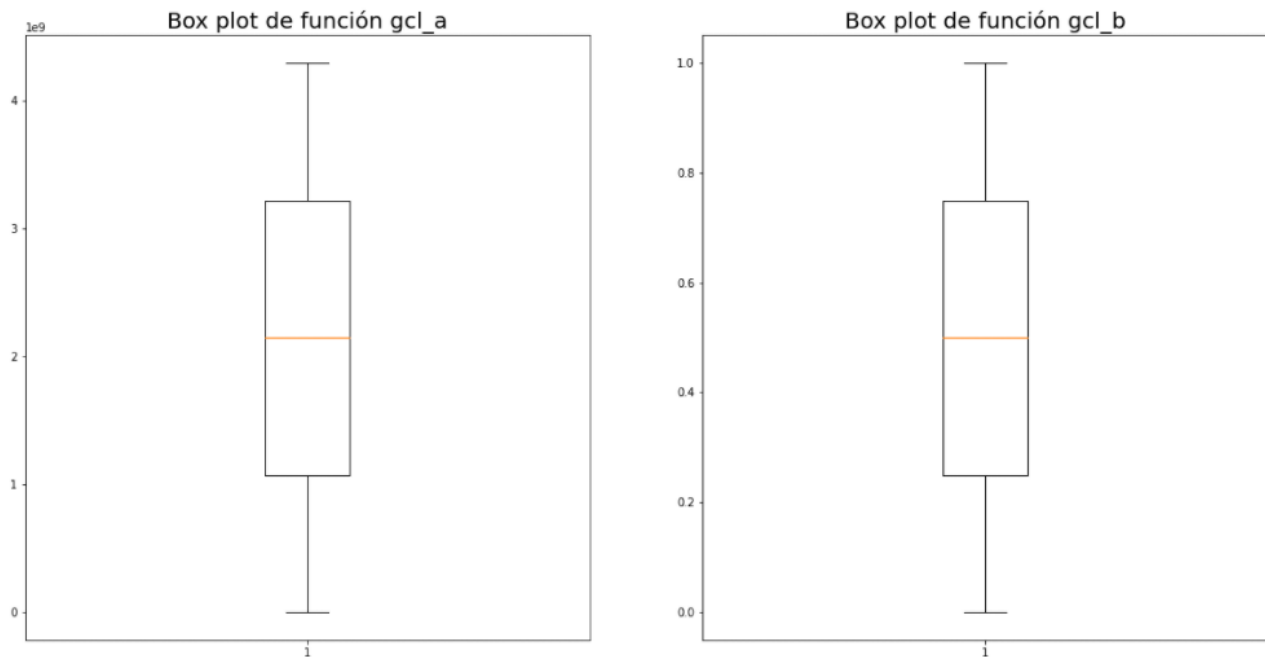


Fig. 1: histogramas de gcl_a y gcl_b .

Fig. 2: boxplots de gcl_a y gcl_b .

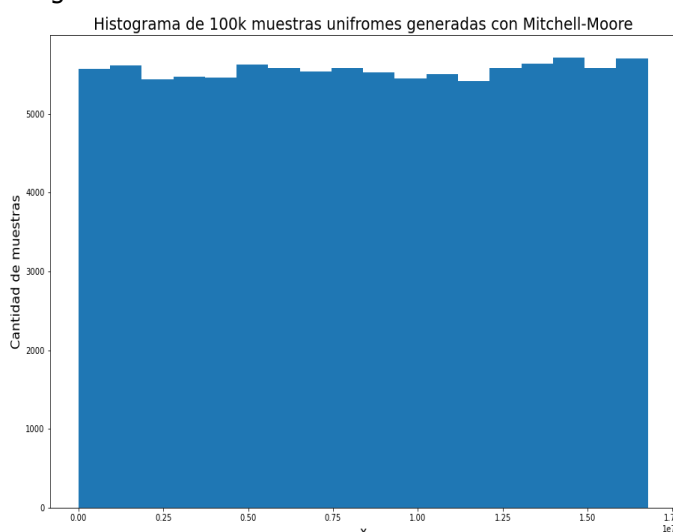
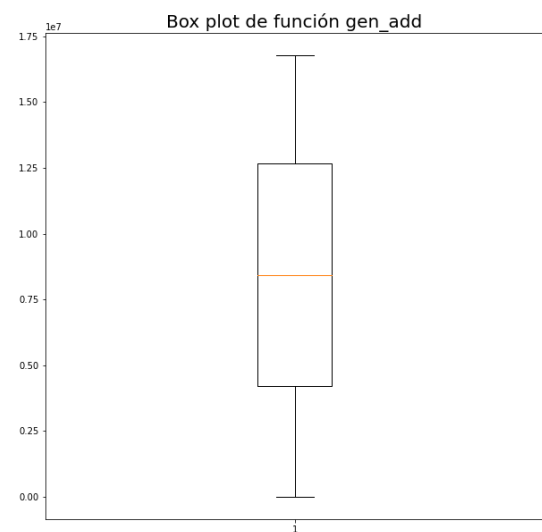
Como se puede observar tanto en los dos histogramas como en los box plots se puede ver que los números de la muestra tienen distribución uniforme. Además, tanto los histogramas como los boxplots son iguales. Esto tiene sentido, ya que la operación de dividir por m para ir al intervalo $(0,1)$ no altera las características de muestras, sino que simplemente se trata de un cambio de escala.

Luego, se pide utilizar el generador aditivo propuesto por Mitchel y Moore:

$$X_n = (X_{n-24} + X_{n-55}) \bmod m$$

donde x_0, x_1, \dots, x_{54} son números arbitrarios no todos iguales, $n \geq 55$ y m es par ($m = 2^{24}$)

Para obtener los 55 números iniciales requeridos como semilla, se utilizó el generador gcl_a . Se utilizó como módulo el valor 2^{24} como lo indica el enunciado. Veamos un boxplot y un histograma de muestras de este generador:

Fig. 3: histograma de gen_add .Fig. 4: boxplot de gen_add .

Como se puede observar, este generador también proporciona números aleatorios uniformes, esta vez entre 0 y 2^{24} ($2^{24} = 16\,777\,216$).

Conclusión

Los generadores de números vistos basados en la operación módulo son pseudoaleatorios, dada la predictibilidad que se tiene en una secuencia generada sobre los valores siguientes a partir del primero (incluso sin considerar la semilla). Por otro lado, devuelven valores con una distribución uniforme, debido a que el resto de la división de la serie de números naturales con respecto a un número siempre sigue la secuencia entre 0 y el dividendo menos uno (donde el resto cero se obtiene a partir de la división con un múltiplo, y el resto de los valores a partir de continuar la división con los naturales siguientes).

Ejercicio 2

Se decidió realizar, para cada generador del ejercicio 1 (*gcl_a*, *gcl_b* y *gen_add*), un test estadístico chi cuadrado y un test espectral.

2.1 Test Chi cuadrado

Se realizaron tests chi cuadrado teniendo en cuenta la discretización del espacio muestral para el caso continuo. Las siguientes tablas muestran los resultados; las filas representan un nivel de significación alfa y las columnas representan la cantidad de muestras utilizadas para el test. Un resultado de 1 representa aceptación y un resultado de 0 representa rechazo.

	10k	50k	100k	150k	200k
0.1%	1.0	1.0	1.0	1.0	1.0
0.5%	1.0	1.0	1.0	1.0	1.0
1.0%	1.0	1.0	1.0	1.0	1.0
3.0%	1.0	1.0	1.0	1.0	1.0
5.0%	1.0	1.0	1.0	1.0	1.0
7.0%	1.0	1.0	1.0	1.0	1.0

Tabla. 1: resultados de test chi2 para el generador *gcl_a*.

Como podemos observar no tenemos evidencia para rechazar la hipótesis de que la distribución es uniforme.

	10k	50k	100k	150k	200k
0.1%	1.0	1.0	1.0	1.0	1.0
0.5%	1.0	1.0	1.0	1.0	1.0
1.0%	1.0	1.0	1.0	1.0	1.0
3.0%	1.0	1.0	1.0	1.0	1.0
5.0%	1.0	1.0	1.0	1.0	1.0
7.0%	1.0	1.0	1.0	1.0	1.0

Tabla. 2: resultados de test chi2 para el generador *gcl_b*.

Luego de aplicar el test variando la cantidad de muestras y el nivel de significación puedes decir que al igual que con *gcl_a* no tenemos evidencia suficiente para rechazar la hipótesis que la distribución de los números aleatorios generados es uniforme.

	10k	50k	100k	150k	200k
0.1%	1.0	1.0	1.0	1.0	1.0
0.5%	1.0	1.0	1.0	1.0	1.0
1.0%	1.0	1.0	1.0	1.0	1.0
3.0%	1.0	1.0	1.0	1.0	1.0
5.0%	1.0	1.0	1.0	1.0	1.0
7.0%	1.0	0.0	1.0	1.0	1.0

Tabla 3: resultados de test chi2 para el generador *gen_add*.

En este caso de las 30 combinaciones dadas por los distintos parámetros tuvimos 1 caso donde se rechazó la hipótesis, pero por amplia mayoría podemos concluir que los números tienden a tener distribución uniforme.

2.2 Test espectral

Este test consiste en graficar los 2 números consecutivos generados como coordenadas en un scatter plot. Si los mismos no tienen ningún patrón podemos decir que tienen distribución uniforme.

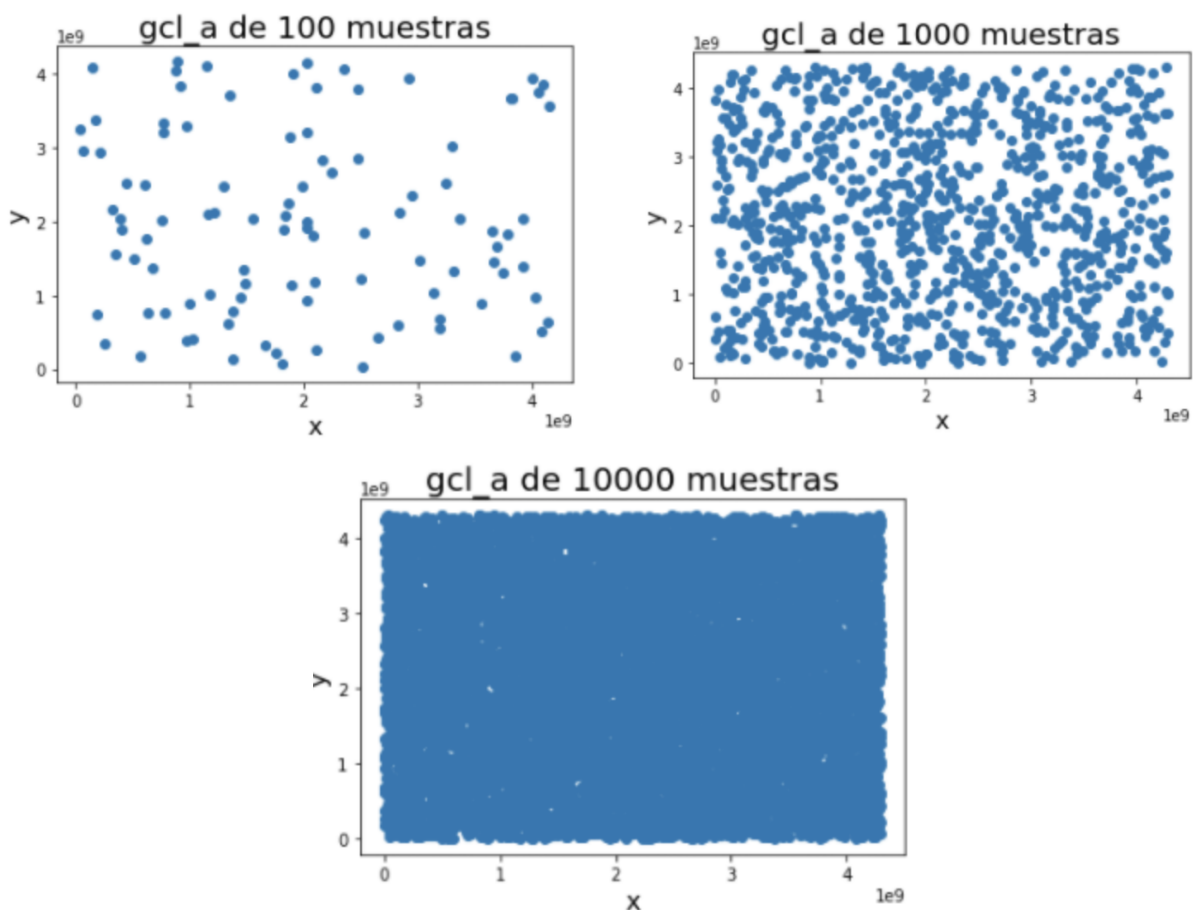


Fig. 5: scatterplots de *gcl_a* de 100, 1000 y 10000 muestras.

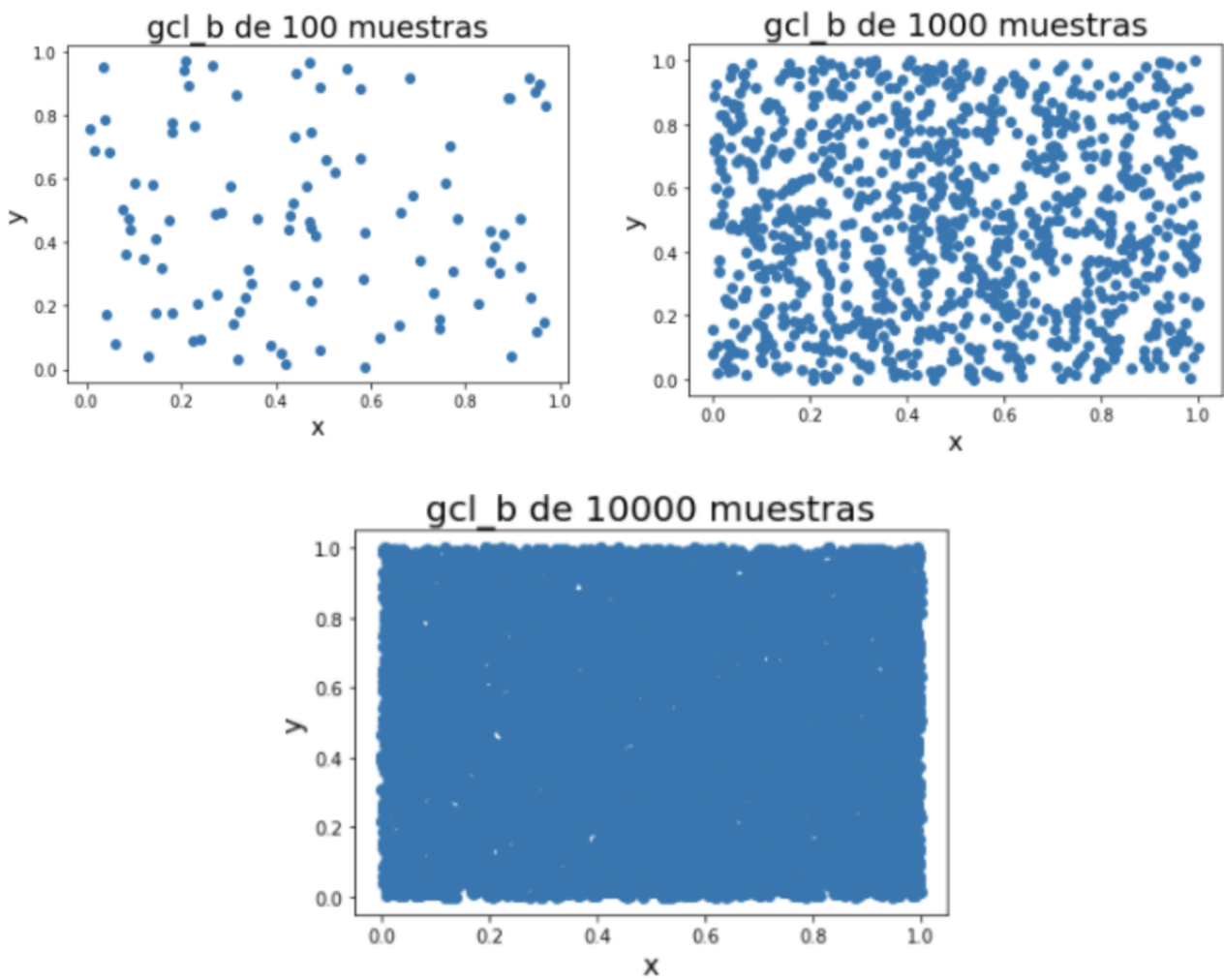
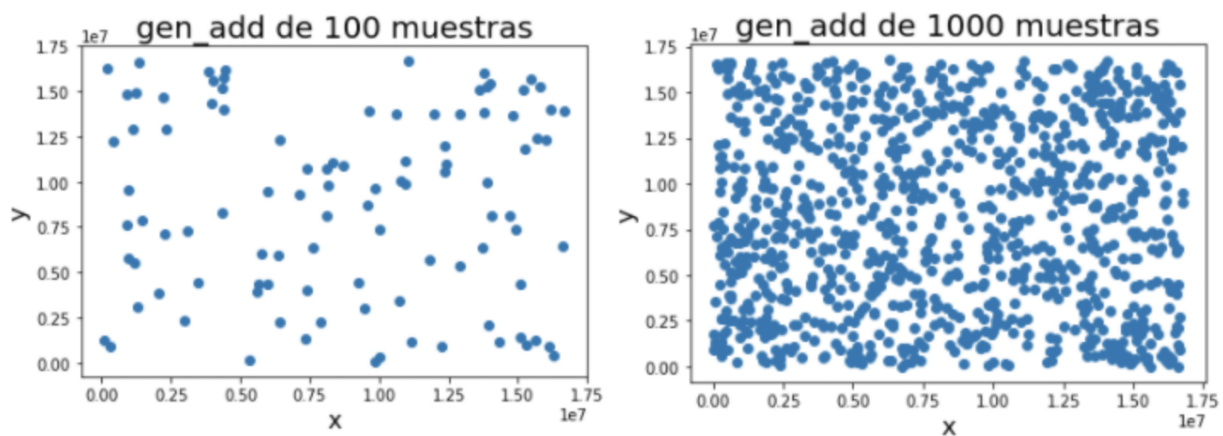


Fig. 6: scatterplots de gcl_b de 100, 1000 y 10000 muestras.



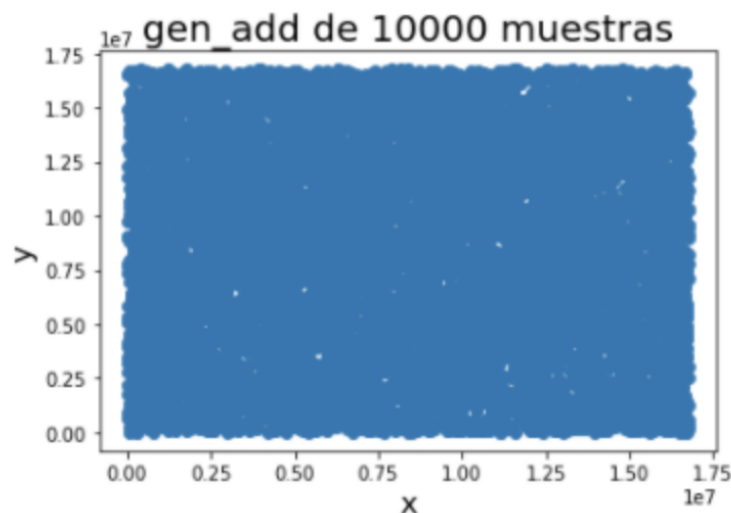


Fig. 7: scatterplots de *gen_add* de 100, 1000 y 10000 muestras.

Finalmente podemos observar que en los 9 gráficos no hay un patrón aparente en los números y que los mismos se distribuyen aleatoriamente en el plano, por ende tienen distribución uniforme.

Conclusión

Tanto los tests analíticos como la comprobación gráfica de la falta de patrones en los números generados confirman la distribución anticipada por los histogramas realizados en el punto anterior. Esta correspondencia es particularmente coherente dado que la semilla utilizada en ambos ejercicios es la misma.

Ejercicio 3

Se desea generar pares ordenados aleatorios en el círculo de centro (10, 10) con radio 5 como en la figura.

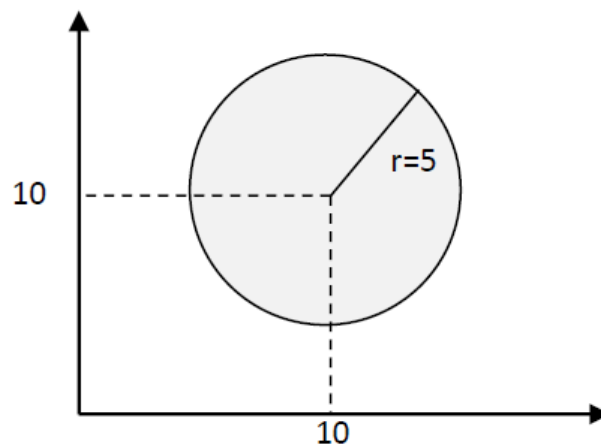


Fig. 8: esquema del ejercicio 3.

Para esto, se siguen los siguientes pasos:

1. Generar números aleatorios u_i entre 0 y 1. En el punto a) se utiliza un generador provisto por el lenguaje mientras que en el b) se utilizó el generador congruencial lineal del punto 1b.
2. Aplicar la siguiente transformación a cada número: $(b - a) * u_i + a$. De esta forma, es posible obtener números aleatorios en el intervalo $[a, b]$. En este caso, se tiene que $a = 5$, $b = 10$.

3. Repetir pasos 1 y 2 para obtener otra secuencia de números aleatorios entre 5 y 10. Se tiene la precaución de que, como se tratan de números pseudoaleatorios y dos corridas del mismo generador con los mismos parámetros resultarán en la misma secuencia de números, es necesario mezclar una de las secuencias.
4. Tomar un número de cada secuencia y formar un par (x_i, y_i) .
5. Calcular la distancia contra el centro del círculo $(10, 10)$ y, si la misma es menor al radio 5, entonces lo guardamos. En caso contrario, se descarta.

El siguiente es un scatterplot de los pares obtenidos con el generador del lenguaje para el punto a):

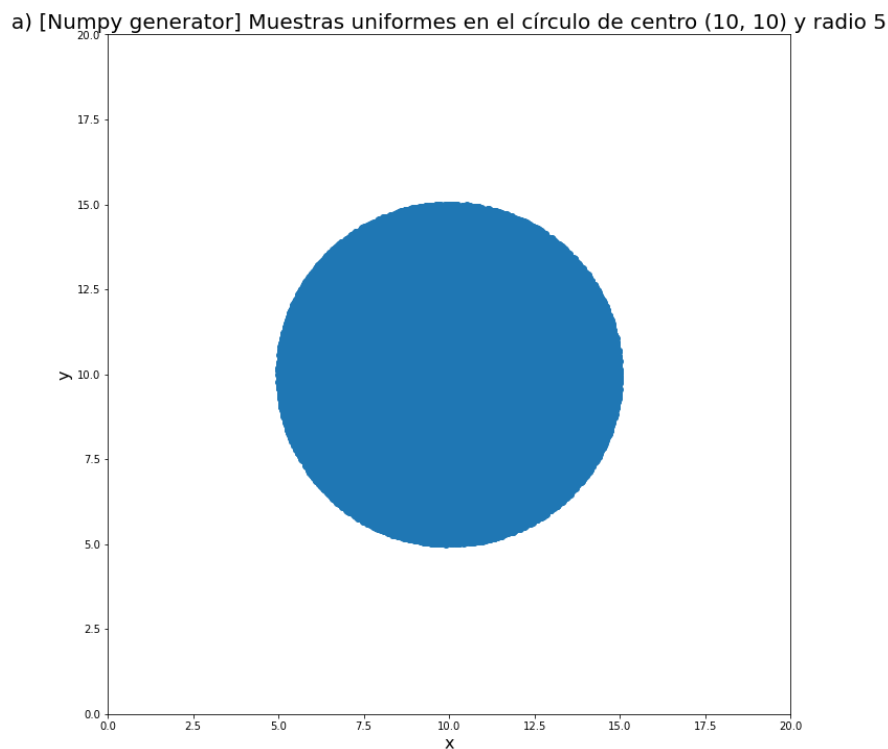


Fig. 9: muestras aleatorias en el área pedida utilizando generador de numpy.

El siguiente es un scatterplot de los pares obtenidos con el generador congruencial lineal del 1b para el punto b):

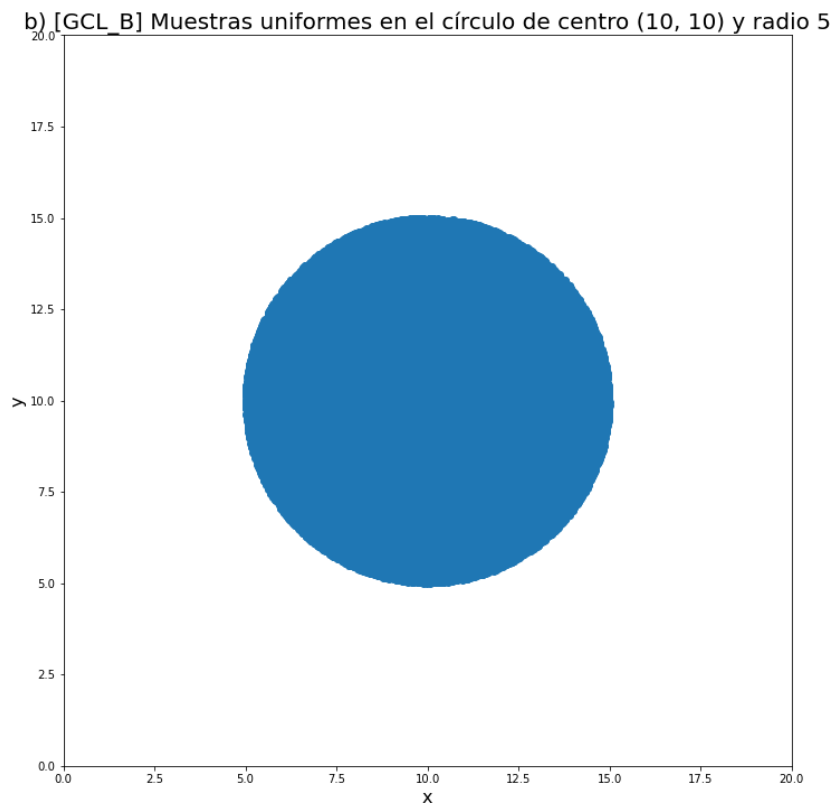


Fig. 10: muestras aleatorias en el área pedida utilizando generador *gcl_b*.

Como se puede observar en ambos gráficos, los pares aleatorios se ubican dentro del círculo de centro (10, 10) y radio 5.

Para el cálculo del factor de rendimiento, se tomaron dos enfoques. Por un lado, se calcula de manera teórica como el cociente entre el área del círculo y el área del cuadrado de lado 10 que lo encierra.

$$\text{Factor de rendimiento} = \frac{\text{área del círculo}}{\text{área del cuadrado}} = \frac{\pi \cdot r^2}{\text{base} \cdot \text{altura}} = \frac{\pi \cdot 5^2}{10 \cdot 10} = \frac{25\pi}{100} = \frac{1}{4}\pi = 0,785398\dots$$

Por otra parte, se calcula de manera empírica como la cantidad de muestras válidas sobre cantidad de muestras totales. Para una corrida con 100.000 muestras se observó lo siguiente:

$$a) \text{ Factor de rendimiento} = \frac{\text{Muestras válidas}}{\text{Muestras totales}} = \frac{78\,256}{100\,000} = 0,78256$$

$$b) \text{ Factor de rendimiento} = \frac{\text{Muestras válidas}}{\text{Muestras totales}} = \frac{78\,549}{100\,000} = 0,78549$$

Vemos que los factores de rendimiento empíricos son muy parecidos al factor de rendimiento teórico, con el del punto b) siendo un poco más preciso.

Conclusión

El hecho de generar números aleatorios en un espacio acotado con coordenadas cartesianas y después limitar una parte del mismo para lograr la distribución pedida hace que el proceso de generación de números con la distribución pedida tenga un "remanente", el cual se estima porcentualmente en el punto c. Muchas veces se pueden estudiar formas de aumentar el factor de rendimiento del proceso para mejorar el desempeño del mismo (disminuyendo el "desperdicio"). Para este caso, una opción posible podría haber sido considerar valores aleatorios de radio y ángulo en coordenadas polares con el objetivo de lograr la distribución pedida.

Ejercicio 4

Se pide generar números aleatorios que sigan la siguiente pdf:

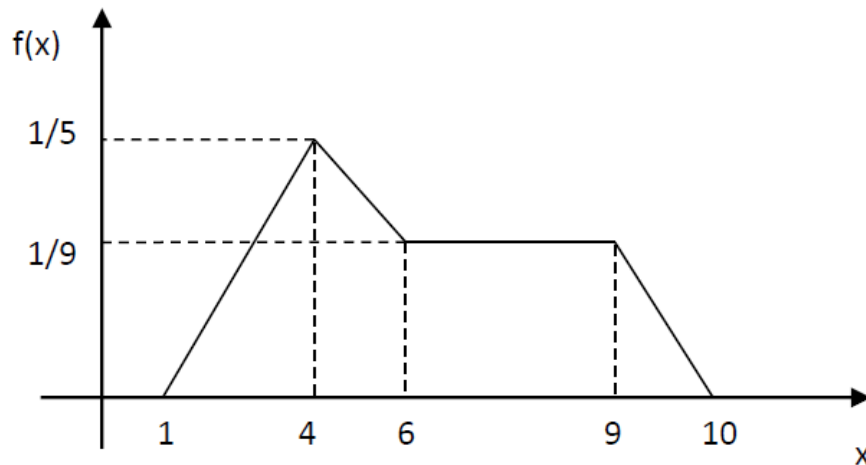


Fig. 11: esquema del ejercicio 4.

Primero, para el punto a), se define esta función:

$f_X(x) = 0$	<i>si</i>	$x < 1$
$(1/15)x - (1/15)$	<i>si</i>	$1 \leq x < 4$
$-(2/45)x + (17/45)$	<i>si</i>	$4 \leq x < 6$
$(1/9)$	<i>si</i>	$6 \leq x < 9$
$-(1/9)x + (10/9)$	<i>si</i>	$9 \leq x < 10$
0	<i>si</i>	$x \geq 10$

Al graficar, se observa la misma función que plantea el enunciado:

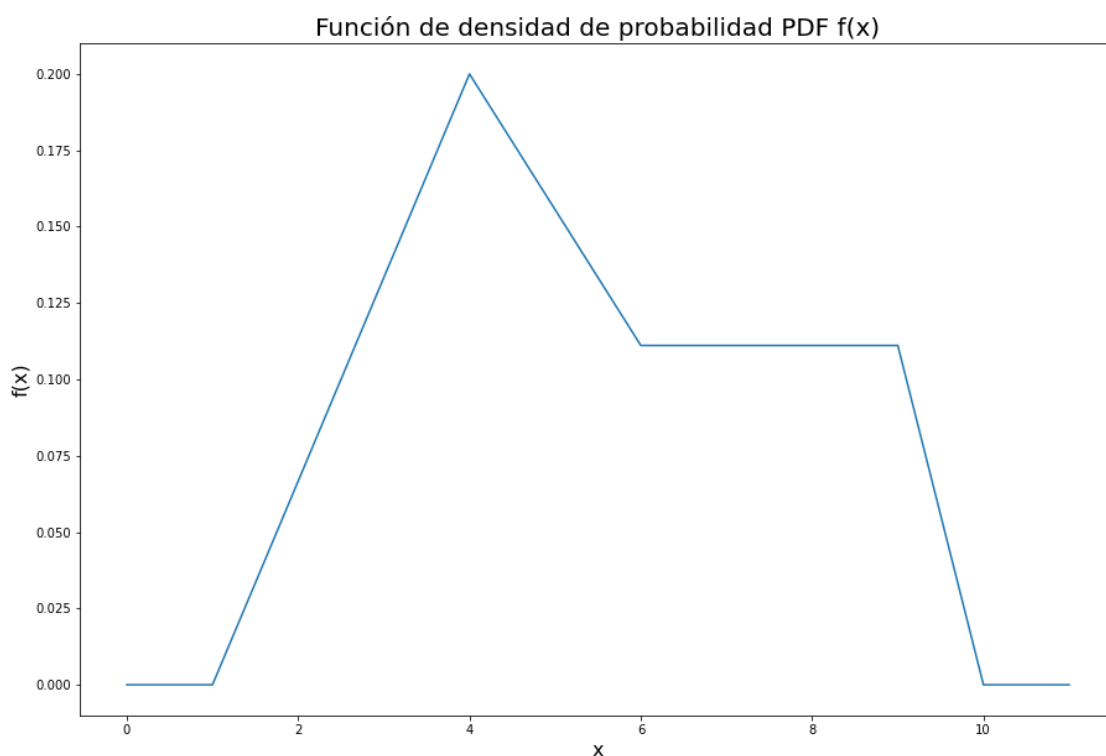


Fig. 12: pdf del ejercicio 4 generada.

Para el punto b) se pide hallar la función acumulada y su inversa. Primero, obtenemos la función de probabilidad acumulada cdf integrando la pdf por tramos.

$$\begin{aligned}
 F_X(x) &= 0 & \text{si } x < 1 \\
 &\int_1^x ((1/15)x + (-1/15)) dx & \text{si } 1 \leq x < 4 \\
 &\int_1^4 ((1/15)x + (-1/15)) dx + \int_4^x ((-2/45)x + (17/45)) dx & \text{si } 4 \leq x < 6 \\
 &\int_1^4 ((1/15)x + (-1/15)) dx + \int_4^6 ((-2/45)x + (17/45)) dx + \\
 &\int_6^x 1/9 dx & \text{si } 6 \leq x < 9 \\
 &\int_1^4 ((1/15)x + (-1/15)) dx + \int_4^6 ((-2/45)x + (17/45)) dx + \\
 &\int_6^9 1/9 dx + \int_9^x ((-1/9)x + (10/9)) dx & \text{si } 9 \leq x < 10 \\
 &\int_1^4 ((1/15)x + (-1/15)) dx + \int_4^6 ((-2/45)x + (17/45)) dx + \\
 &\int_6^9 1/9 dx + \int_9^{10} ((-1/9)x + (10/9)) dx & \text{si } x \geq 10
 \end{aligned}$$

$$\begin{aligned}
 F_X(x) &= 0 & \text{si } x < 1 \\
 &(1/30)x^2 + (-1/15)x + (1/30) & \text{si } 1 \leq x < 4 \\
 &3/10 + (-1/45)x^2 + (17/45)x + (-52/45) & \text{si } 4 \leq x < 6 \\
 &3/10 + 14/45 + (1/9)x + (-2/3) & \text{si } 6 \leq x < 9 \\
 &3/10 + 14/45 + 1/3 + (-1/18)x^2 + (10/9)x + (-11/2) & \text{si } 9 \leq x < 10 \\
 &3/10 + 14/45 + 1/3 + 1/18 & \text{si } x \geq 10
 \end{aligned}$$

$$\begin{aligned}
 F_X(x) &= 0 & \text{si } x < 1 \\
 &(1/30)x^2 + (-1/15)x + (1/30) & \text{si } 1 \leq x < 4 \\
 &3/10 + (-1/45)x^2 + (17/45)x + (-52/45) & \text{si } 4 \leq x < 6 \\
 &11/18 + (1/9)x + (-2/3) & \text{si } 6 \leq x < 9 \\
 &17/18 + (-1/18)x^2 + (10/9)x + (-11/2) & \text{si } 9 \leq x < 10 \\
 &1 & \text{si } x \geq 10
 \end{aligned}$$

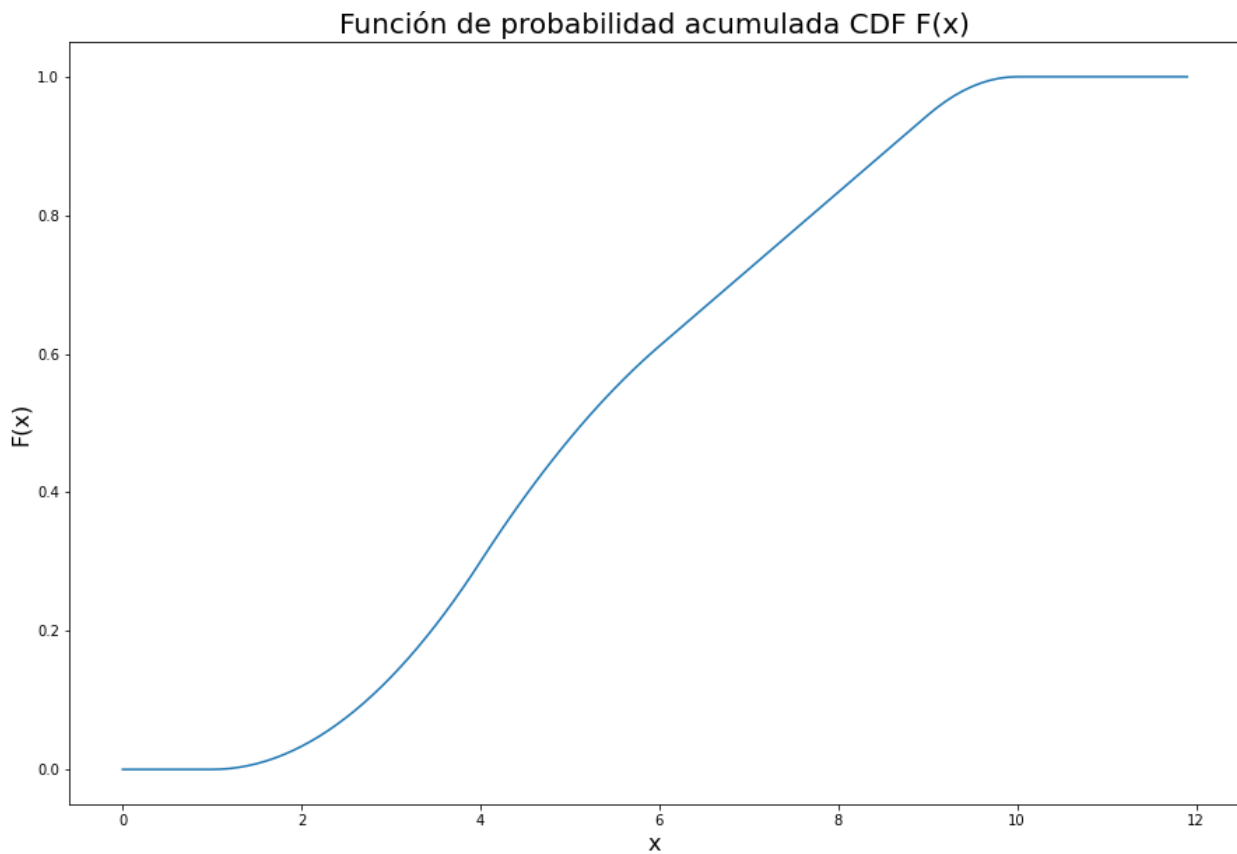


Fig. 13: cdf del ejercicio 4; integral de pdf.

Ahora calculamos su inversa:

$$F_X^{-1}(u) = \begin{cases} 1 + \sqrt{30u} & \text{si } 0 \leq u < 3/10 \\ 17/2 - \sqrt{-45(u - 3/4)} & \text{si } 3/10 \leq u < 11/18 \\ 9(u + 1/18) & \text{si } 11/18 \leq u < 17/18 \\ 10 - \sqrt{-18(u - 1)} & \text{si } 17/18 \leq u < 1 \end{cases}$$

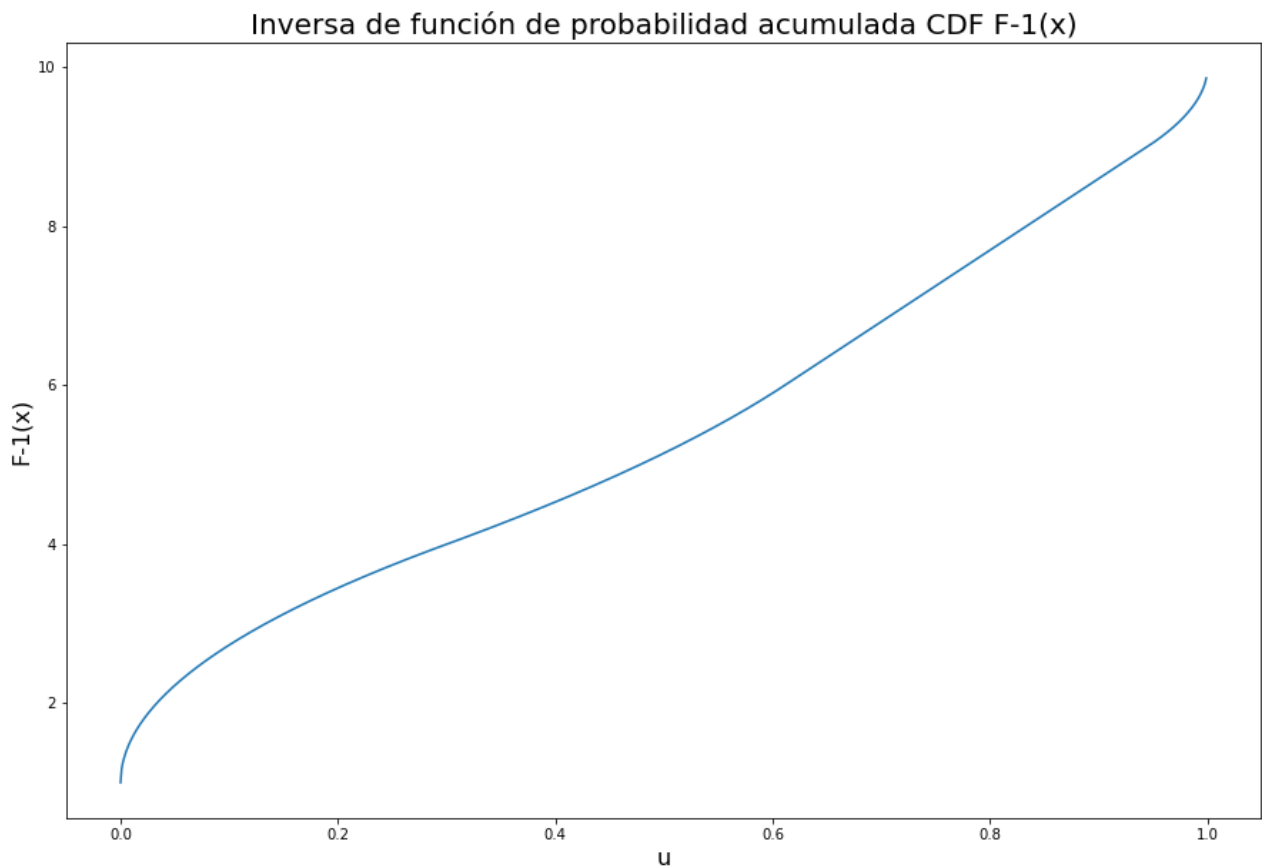


Fig. 14: cdf inversa del ejercicio 4.

Luego, para obtener números aleatorios con la pdf pedida, simplemente se generan números aleatorios entre 0 y 1 y luego se aplica esta función de probabilidad acumulada inversa. Veamos un histograma de muestras de esta variable:

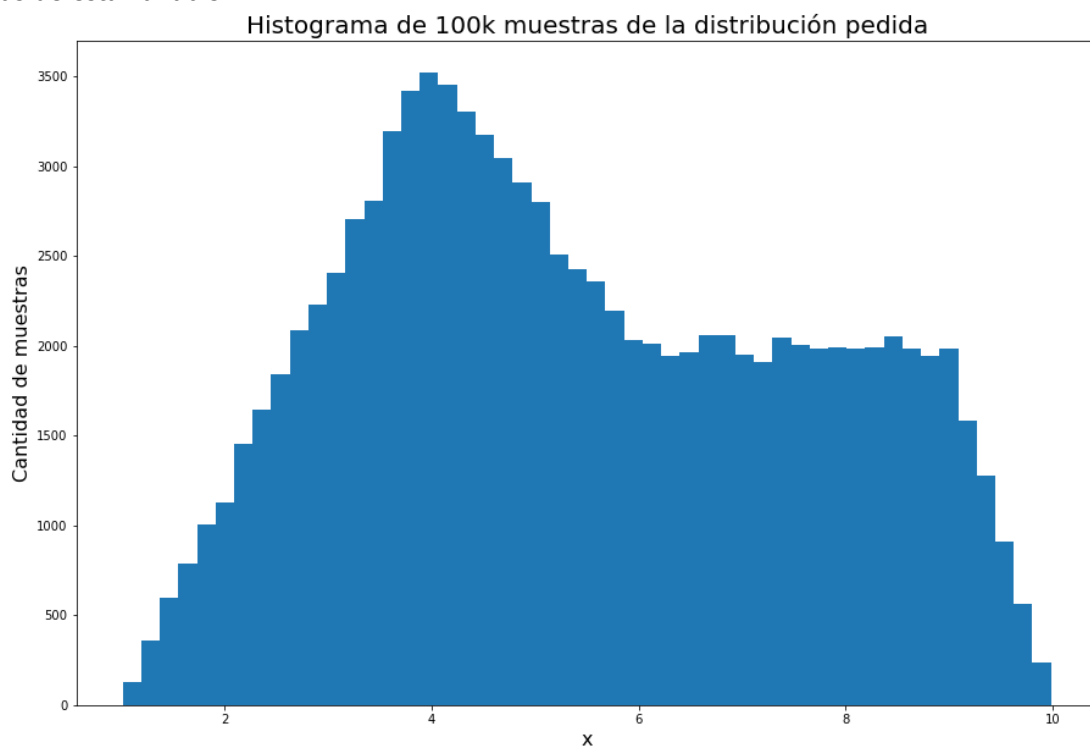


Fig. 15: histograma de muestras generadas que sigan pdf pedida.

Como se puede notar, este histograma tiene la misma forma que la función de densidad de probabilidad pedida, por lo cual se puede concluir que los números aleatorios generados exitosamente responden a esta ley de probabilidad.

Conclusión

Es particularmente cómodo el uso de herramientas informáticas para facilitar el tratamiento analítico del problema de hallar la inversa de una función. Adicionalmente, el histograma es una sencilla aproximación para comprobar que la inversa hallada es correcta, así como también puede ser útil para aproximar distribuciones en general.

Ejercicio 5

Para implementar un generador de números aleatorios con distribución normal decidimos utilizar el método de aceptación y rechazo. Primero generamos números con la distribución normal 0,1 y una vez que obtuvimos los mismos los "destandarizamos", es decir multiplicamos por el desvío esperado y media de la distribución que queríamos obtener.

En este caso teníamos que generar una distribución normal 15,3.

Por último como teníamos que generar para el método de aceptación y rechazo números con distribución uniforme y exponencial, realizamos los cálculos de 3 maneras distintas:

- Utilizando el generador de números random con distribución uniforme y exponencial dado por el lenguaje.
- Utilizando el generador de números random con distribución uniforme dado por el lenguaje y con el mismo mediante el método de la transformada inversa lograr obtener los números aleatorios con distribución exponencial.
- Utilizando el generador de números random con distribución uniforme del punto 1 y con el mismo mediante el método de la transformada inversa lograr obtener los números aleatorios con distribución exponencial.

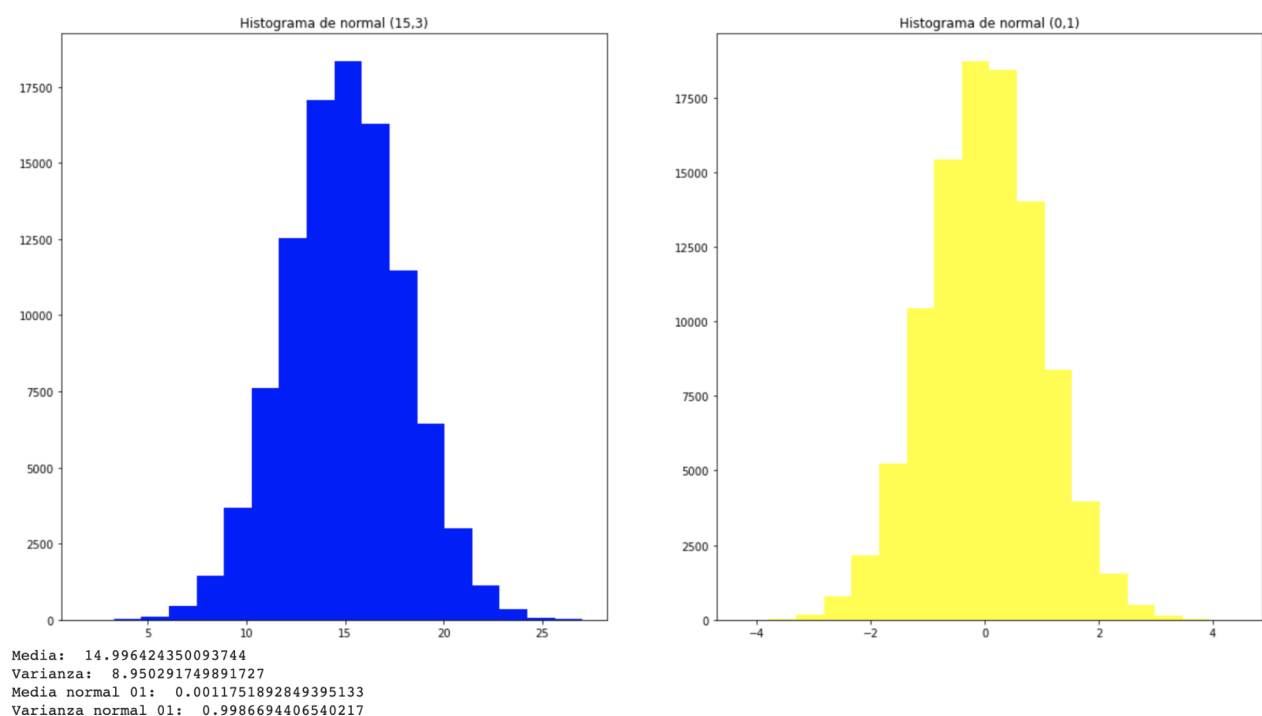


Fig. 16: histogramas de muestras generadas mediante la primera opción.

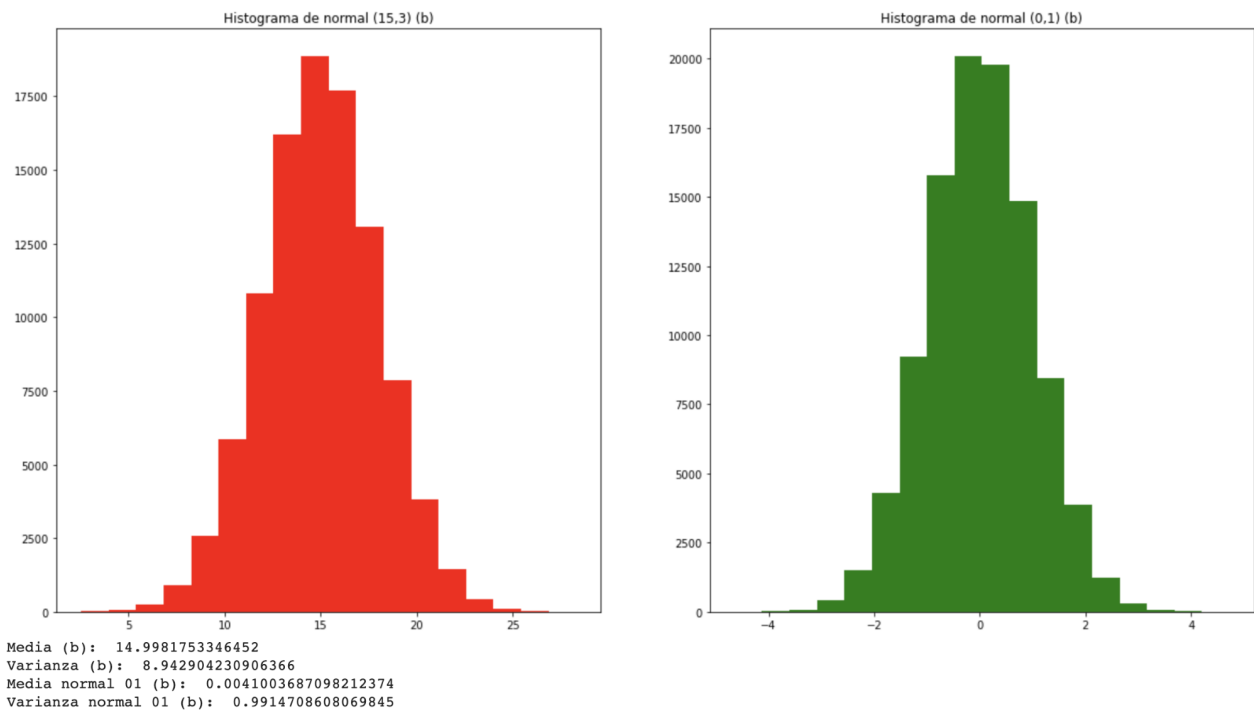


Fig. 17: histograma de muestras generadas mediante la segunda opción.

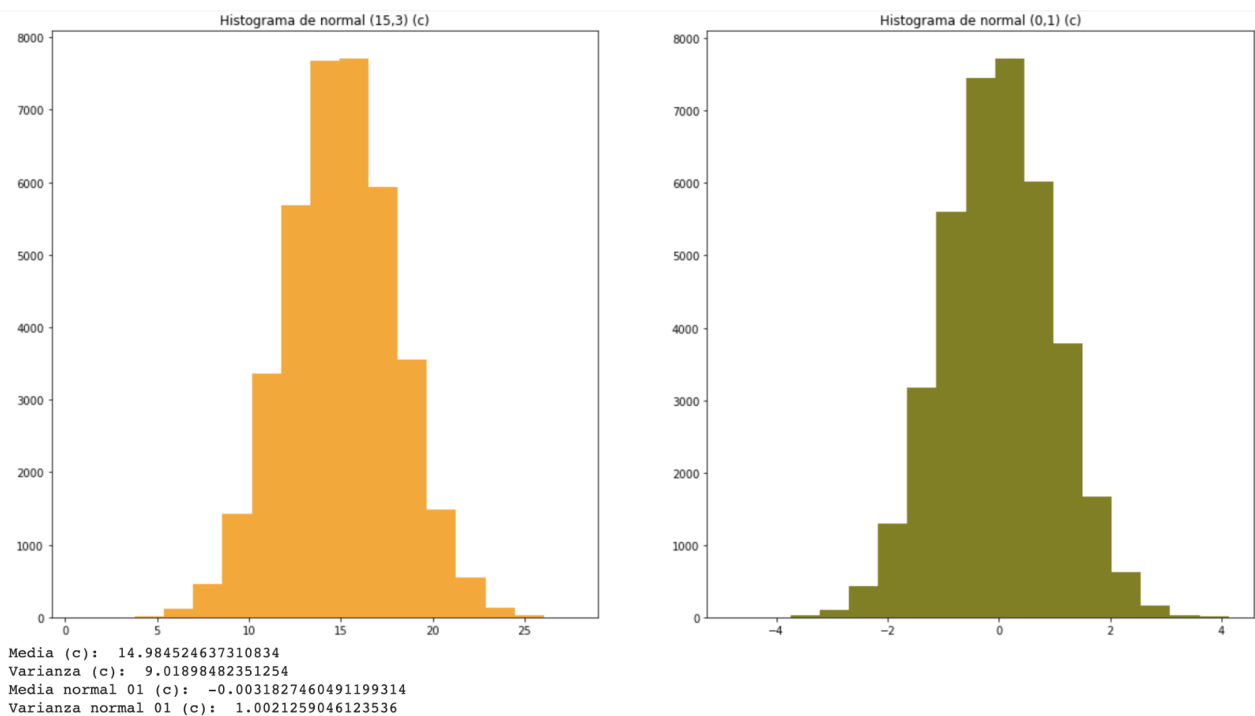


Fig. 18: histograma de muestras generadas mediante la tercera opción.

Como podemos observar sin realizar ninguna comprobación empírica ni test podemos decir que en los 3 casos el histograma se parece mucho a como se graficaría una distribución normal.

Finalmente realizamos tests χ^2 y Kolmogorov Smirnov con variación de parámetros (α y N):

chi2:

Estadístico: 15.930607161456791	Estadístico: 17.126834014521613
Alpha: 0.01	Alpha: 0.05
P-valor: 0.6619028550371833	P-valor: 0.5812753843718971
N: 1000	N: 1000
No hay evidencia suficiente para rechazar H0	No hay evidencia suficiente para rechazar H0
Estadístico: 12.098483405782202	Estadístico: 31.726709460721757
Alpha: 0.1	Alpha: 0.01
P-valor: 0.8813726764146378	P-valor: 0.03354148224242036
N: 1000	N: 10000
No hay evidencia suficiente para rechazar H0	No hay evidencia suficiente para rechazar H0
Estadístico: 21.94387991721686	Estadístico: 28.62223320065495
Alpha: 0.05	Alpha: 0.1
P-valor: 0.2870581762649132	P-valor: 0.07216406888858241
N: 10000	N: 10000
No hay evidencia suficiente para rechazar H0	Se rechaza la H0
Estadístico: 62.61813928588678	Estadístico: 40.02902713935863
Alpha: 0.01	Alpha: 0.05
P-valor: 1.4806525896082429e-06	P-valor: 0.003243727040581364
N: 100000	N: 100000
Se rechaza la H0	Se rechaza la H0
Estadístico: 45.55627141035381	
Alpha: 0.1	
P-valor: 0.0005725621597124229	
N: 100000	
Se rechaza la H0	

Kolmogorov Smirnov:

Estadístico: 0.007796112569292135	Estadístico: 0.017718618941632402
Alpha: 0.01	Alpha: 0.05
P-valor: 0.8809713319642719	P-valor: 0.5274953453076602
N: 1000	N: 1000
No hay evidencia suficiente para rechazar H0	No hay evidencia suficiente para rechazar H0
Estadístico: 0.0328265724001815	Estadístico: 0.011809593589492962
Alpha: 0.1	Alpha: 0.01
P-valor: 0.11337050195379442	P-valor: 0.06097887445692565
N: 1000	N: 10000
No hay evidencia suficiente para rechazar H0	No hay evidencia suficiente para rechazar H0
Estadístico: 0.007942567865285327	Estadístico: 0.0015222976295692359
Alpha: 0.05	Alpha: 0.1
P-valor: 0.28168315062197713	P-valor: 0.9537424164459202
N: 10000	N: 10000
No hay evidencia suficiente para rechazar H0	No hay evidencia suficiente para rechazar H0
Estadístico: 0.0017414606588768144	Estadístico: 0.0025550398903096405
Alpha: 0.01	Alpha: 0.05
P-valor: 0.5446036048191528	P-valor: 0.27053653579516906
N: 100000	N: 100000
No hay evidencia suficiente para rechazar H0	No hay evidencia suficiente para rechazar H0
Estadístico: 0.0019739411697390907	
Alpha: 0.1	
P-valor: 0.45812937708991214	
N: 100000	
No hay evidencia suficiente para rechazar H0	

La mayoría de los test no rechazan la hipótesis de que la distribución es normal, con esto podemos concluir que los números generados por el generador implementado cumple con una distribución normal con media 15 y desvío 3.

Por último utilizando el ejemplo brindado en clase de un test Anderson Darling obtuvimos los mismos resultados:

```
Alpha: 0.15
Estadistico: 0.24929997375011226
Umbral: 0.574
Datos normales (No rechazo H0)
Alpha: 0.1
Estadistico: 0.24929997375011226
Umbral: 0.653
Datos normales (No rechazo H0)
Alpha: 0.05
Estadistico: 0.24929997375011226
Umbral: 0.784
Datos normales (No rechazo H0)
Alpha: 0.025
Estadistico: 0.24929997375011226
Umbral: 0.914
Datos normales (No rechazo H0)
Alpha: 0.01
Estadistico: 0.24929997375011226
Umbral: 1.088
Datos normales (No rechazo H0)
```

Conclusión

Resulta interesante cómo a partir de la capacidad de generación de números aleatorios uniformes los mismos pueden ser transformados para alcanzar una distribución deseada. Tanto el p-valor como los histogramas son versátilmente funcionales a la comprobación de que la transformación efectivamente logra la distribución que se buscaba, y usarlos en conjunto es útil como un doble chequeo.

Ejercicio 6

Se obtienen los 100.000 números aleatorios del archivo provisto y se grafica un histograma para tener una idea inicial de qué distribución siguen.

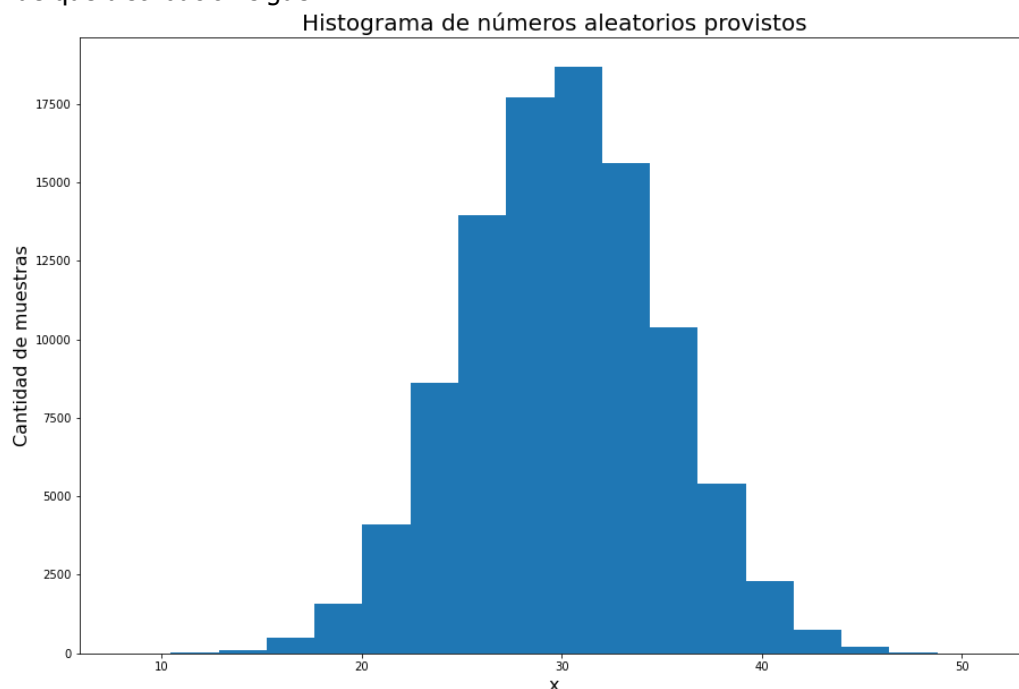


Fig. 19: histograma de muestras aleatorias del archivo *random_numbers.txt*.

A simple vista, podemos sospechar de que estos números siguen una distribución normal dada la forma de este histograma.

Se realizó un test estadístico Kolmogorov-Smirnov. La idea es ver si las muestras provistas se ajustan a una determinada distribución de probabilidad. En este caso, se realizó el test para las tres distribuciones presentadas en el enunciado: uniforme, exponencial y normal. Los resultados fueron los siguientes:

```
Test KS: uniforme entre 25 y 35
Estadístico: 0.15790999999999988
P valor: 0.0
```

```
Test KS: exponencial de lambda 1/30
Estadístico: 0.241456324423321
P valor: 0.0
```

```
Test KS: normal de media 30 y desvío 5
Estadístico: 0.0008726227711003398
P valor: 0.8582365352730704
```

Para las distribuciones uniforme y exponencial, se obtiene un p-valor de 0. Esto quiere decir que, sin importar el valor de significancia alfa que se seleccione, siempre será mayor al p-valor obtenido. Es decir, siempre se rechazará la hipótesis nula. En cambio, para el caso de la normal, se obtiene un p-valor grande de 0,85. Uno puede seleccionar un nivel de significación alfa tan pequeño como se desee y nunca habrá suficiente información para rechazarla.

Así, se concluye que los números aleatorios provistos se ajustan mejor a una distribución normal de media 30 y desvío 5, tal como fue intuido con el histograma.

Conclusión

Nuevamente, el histograma funciona como una comprobación rápida de que una muestra se corresponde con una distribución bien conocida, mientras que el test computarizado da una comprobación analítica para reafirmar lo visualizado de forma gráfica.

Ejercicio 7

Esquema de solución

Dada la discretización propuesta en la consigna, se modelan los estados del sistema en una matriz de 200x200 (donde cada unidad es una celda que puede contener un agente). Los valores que pueden tomar las celdas corresponden a strings que contendrán la información de cada agente, donde un número especifica su identificador único, o "id", una letra especifica su tipo ("A", "B" o "C"), "X" representa que está infectado, y el caso de la celda vacía se representa con un punto ".").

En cuanto al movimiento de los agentes según el modelo de Random Walking, el mismo permite que los desplazamientos sean a cualquiera de las "cuatro direcciones" (norte, sur, este u oeste). Para la medición de distancias se utiliza la distancia Manhattan, por lo que las celdas posibles a una distancia dada para un par de coordenadas (x, y) dentro de la matriz serán conocidas a la hora de "enfermar" agentes.

Para lograr la lógica del 2% de infectados iniciales, se eligen de forma uniformemente aleatoria y sin repetición una cantidad proporcional de números del 0 al número de agentes menos uno, es decir, una lista de identificadores de agentes que estarán enfermos desde el comienzo.

Con respecto a la distribución uniforme de los agentes, se eligen números aleatorios de forma equiprobable sobre el espacio de posibilidades para las coordenadas de la matriz. En el caso de que la celda se encuentre ocupada, se repite el proceso hasta encontrar un par (x, y) tal que la celda esté libre. Por supuesto, es una hipótesis que el número de agentes elegido es menor a la cantidad de celdas disponibles.

Se definen a través de arreglos de variables posiciones cercanas (con distancia menor o igual a 3 celdas) y lejanas (con distancia entre 3 y 6 celdas) dada la distinción hecha en el enunciado sobre la probabilidades de enfermar, consideran cada uno de los casos cercanos y lejanos según la distancia Manhattan. Para cada una de las celdas con posibilidad de ser infectada, se toma una variable aleatoria uniforme entre 0 y 1 y se la compara con la probabilidad de infectar. En caso de que el agente esté infectado, se agrega a la celda una "X" (separada con "_") para indicar esto.

Un punto importante es que la infección del agente se hace en una matriz auxiliar, ya que no se debe modificar el estado actual del sistema que se está utilizando para el análisis de las infecciones. De otra forma, los infectados recientes podrían considerarse como infectados ya existentes en la misma iteración, dando lugar a una evolución incorrecta. Es importante considerar que los infectados nuevos solo deben poder infectar a partir de la siguiente iteración. Durante la simulación, una vez que se hayan analizado las probabilidades de infección para cada agente, el estado del sistema será actualizado.

Para la animación y los gráficos pedidos, se mantienen dos listas: una de agentes infectados y otra de sanos, las cuales contiene tuplas con 3 valores: el identificador del agente, su posición en X y su posición en Y. Se empieza agregando a dichas listas el estado inicial del sistema, y al final de cada iteración se agregarán los estados actualizados de agentes enfermos y sanos.

Algoritmo de simulación

El algoritmo de simulación tiene un ciclo principal que pretende continuar la evolución del sistema según la cantidad de iteraciones pedidas, o bien, hasta que todos los agentes se enfermen. En cada iteración hay tres partes: la infección, el movimiento y la actualización.

Durante la infección, se recorre el estado actual de la matriz intentando infectar agentes en caso de que la celda actual contenga un infectado. Como se comentó anteriormente, las infecciones se realizan en una copia del estado del sistema para que los nuevos enfermos no infecten al mismo tiempo que los ya existentes.

Al finalizar la etapa de infección, se actualiza el estado del sistema. A continuación, durante la etapa de movimiento se consideran las distintas "velocidades" de los agentes, utilizando la función módulo con el número de iteración para determinar si el agente debe moverse o no. Para el movimiento de los agentes en sí, se debe considerar a priori el caso en donde no hay ninguna celda libre al que el agente pueda moverse. Esta verificación se hace a medida que los agentes se mueven. En consecuencia, la actualización de la posición de los agentes también se hace en la misma matriz auxiliar contra la cual se verifica si el movimiento es posible. Superada la verificación para garantizar la posibilidad de movimiento, se elige una dirección equiprobablemente al azar, y se verifica si dicho movimiento es posible (es decir, si la celda destino no está ocupada). Esto se repite hasta encontrar una posición libre, ya que en este punto se sabe que existe un lugar al cual el agente puede moverse.

Finalmente, con el estado actualizado del sistema luego de las infecciones y los movimientos de cada agente, se guardan listas de los agentes sanos e infectados correspondientes a la última iteración para poder realizar las visualizaciones.

Visualizaciones

Caso con 500 agentes

La animación de ejemplo se encuentra en el siguiente enlace:
<https://drive.google.com/file/d/1OnnPYaMzxhdZSt5umPgyJmiaKqcSwEBU>

En la misma, así como en todas las animaciones siguientes, se identifican los agentes infectados como puntos rojos, y los sanos como puntos verdes. A continuación, se muestran las estadísticas correspondientes a dicho escenario.

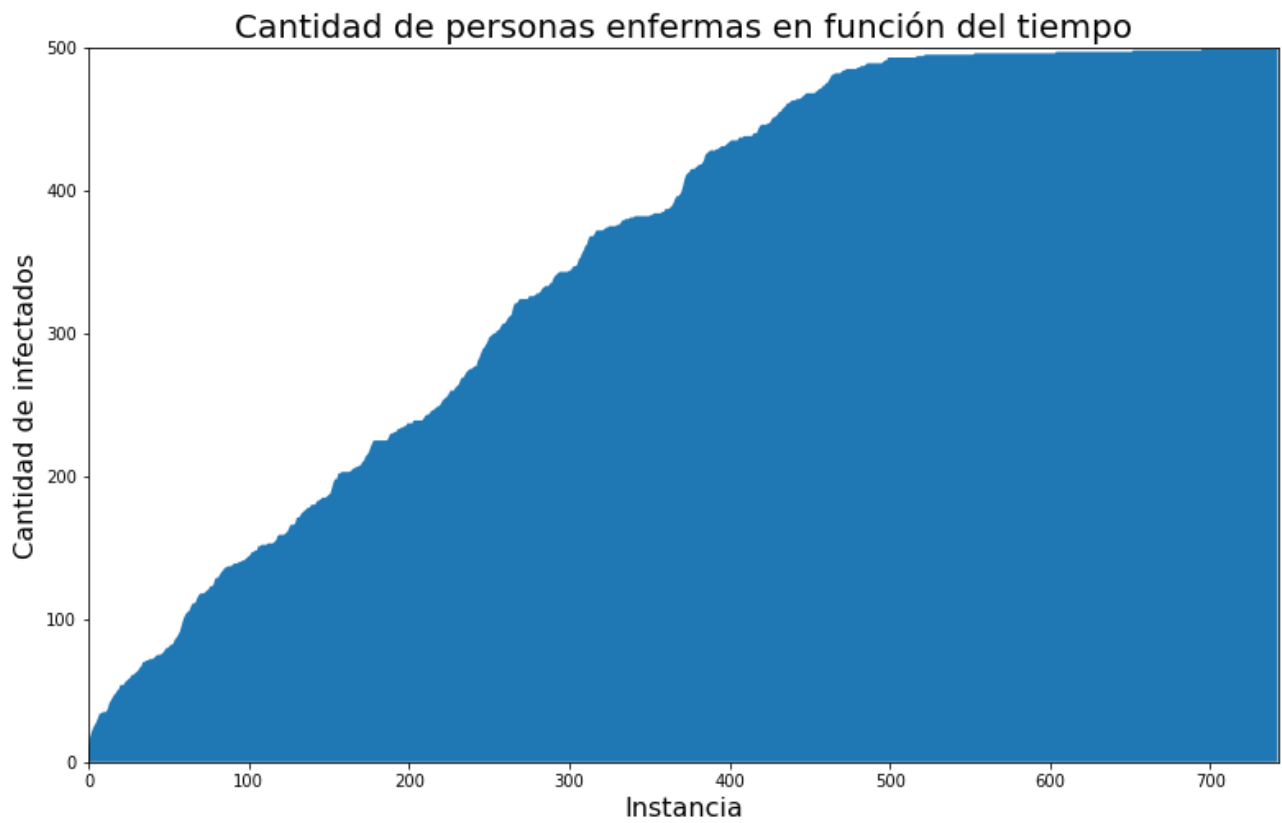


Fig. 20: enfermos en cada instante para 500 agentes.

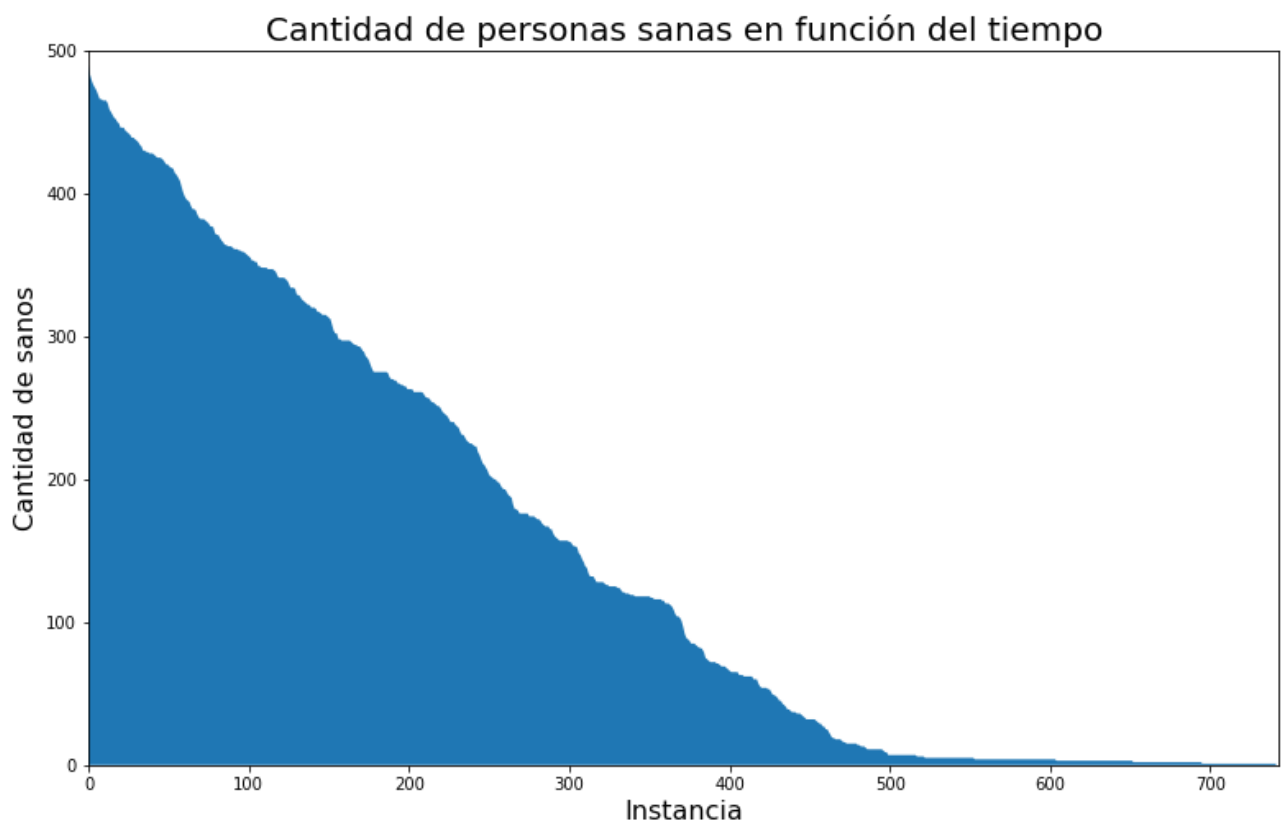


Fig. 21: sanos en cada instante para 500 agentes.

Como era esperable, se observa que la cantidad de sanos e infectados son complementarios. Adicionalmente se puede ver que la cantidad de infectados aumenta más rápido al comienzo. Esto se puede adjudicar al tipo de Random Walking utilizado, ya que al tener un movimiento equiprobable entre (cuando mucho) cuatro direcciones, el mismo parece hacer tender a que el agente termine cerca de donde empezó. En consecuencia, los contagios tienden a ser solo con los vecinos próximos, y la infección generalizada solo se demora por agentes que están lo suficientemente alejados como para no tener un contagio inmediato.

A continuación se muestran estadísticas asociadas a los tiempos de contagio, tanto en total como para cada tipo de agente por separado.

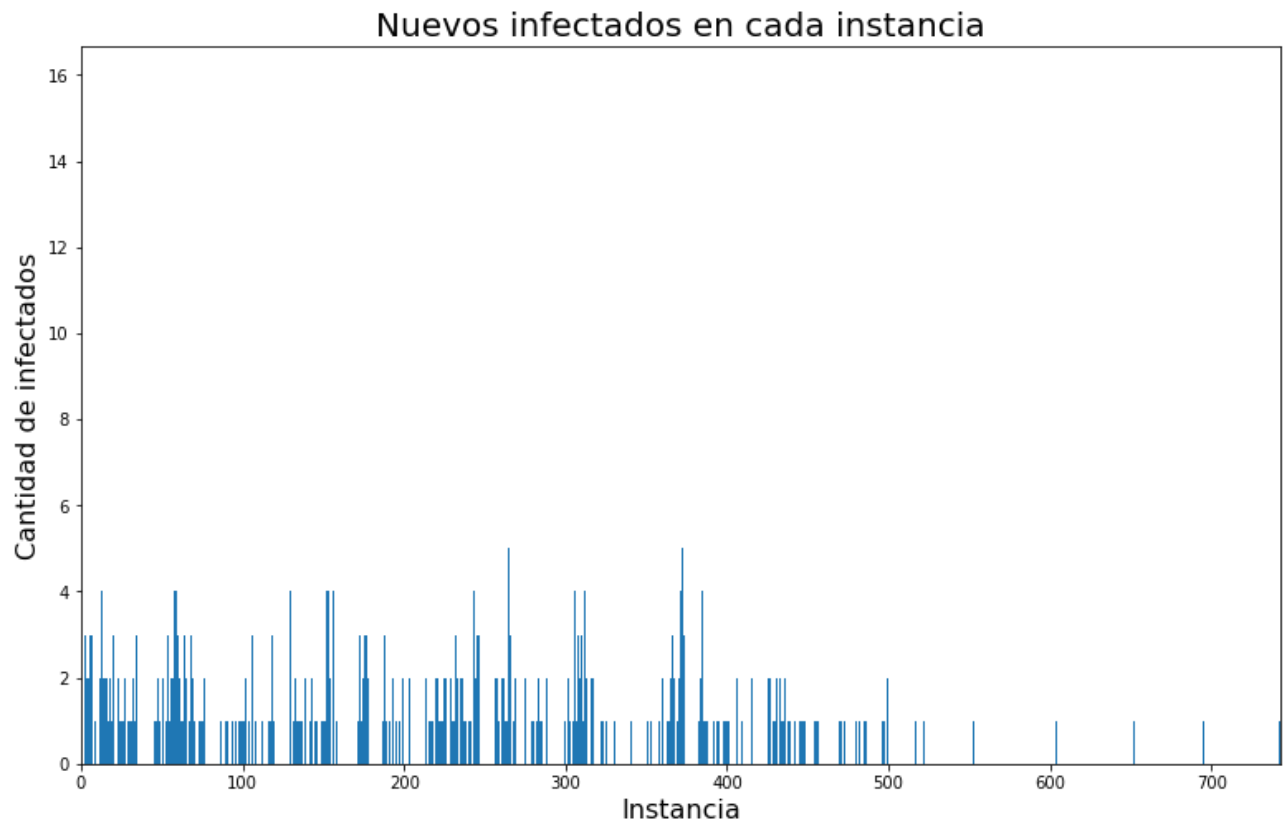


Fig. 22: nuevos infectados en cada instancia para 500 agentes.

La distribución de la cantidad de agentes infectados muestra también una cierta concentración al principio que a largo plazo (pasada la mitad de las iteraciones) decae hasta ser ínfima, pero cuyo pico no está exactamente al inicio.

Caso con 1 000 agentes

La animación de ejemplo se encuentra en el siguiente enlace:

<https://drive.google.com/file/d/15aO0r66LKBaVOsqXGHFnUeVd8SqRAMZc>

A continuación, se muestran las estadísticas correspondientes a dicho escenario.

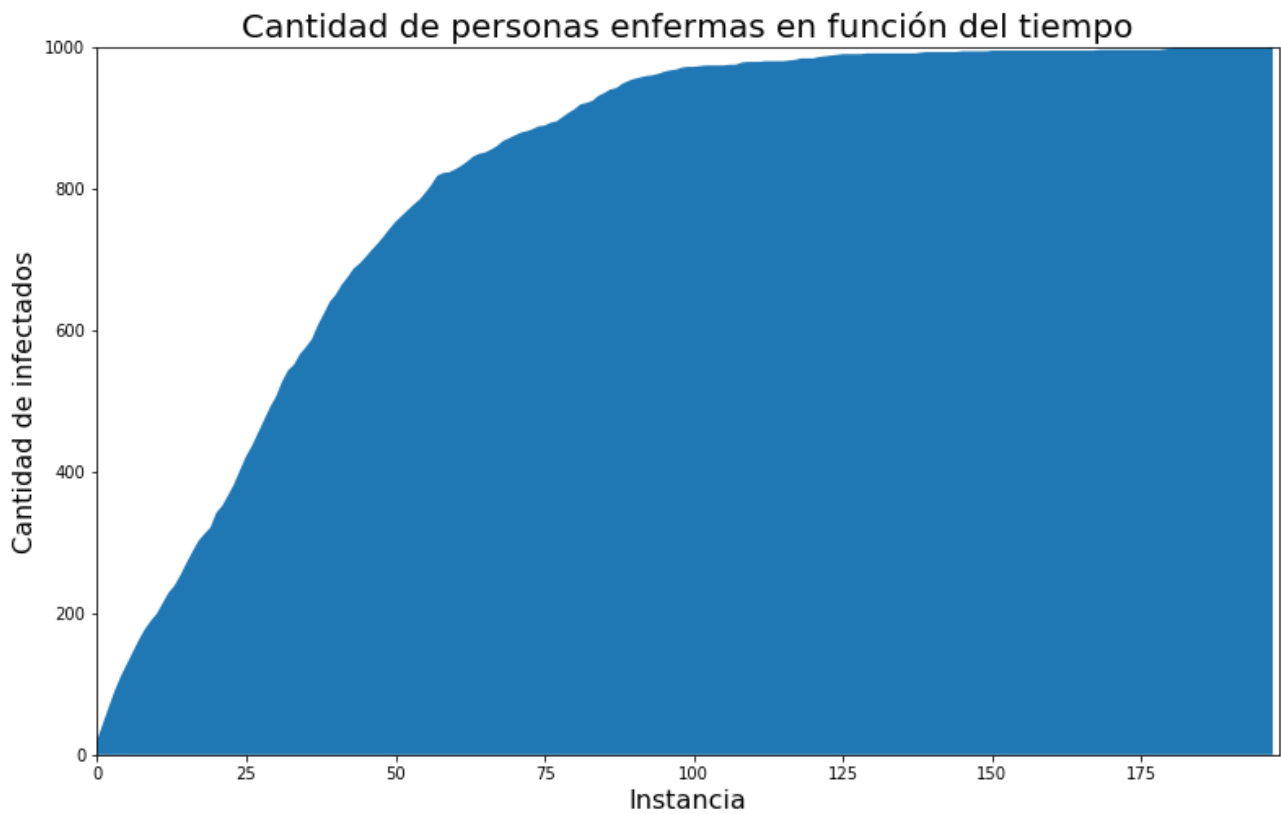


Fig. 23: enfermos en cada instante para 1 000 agentes.

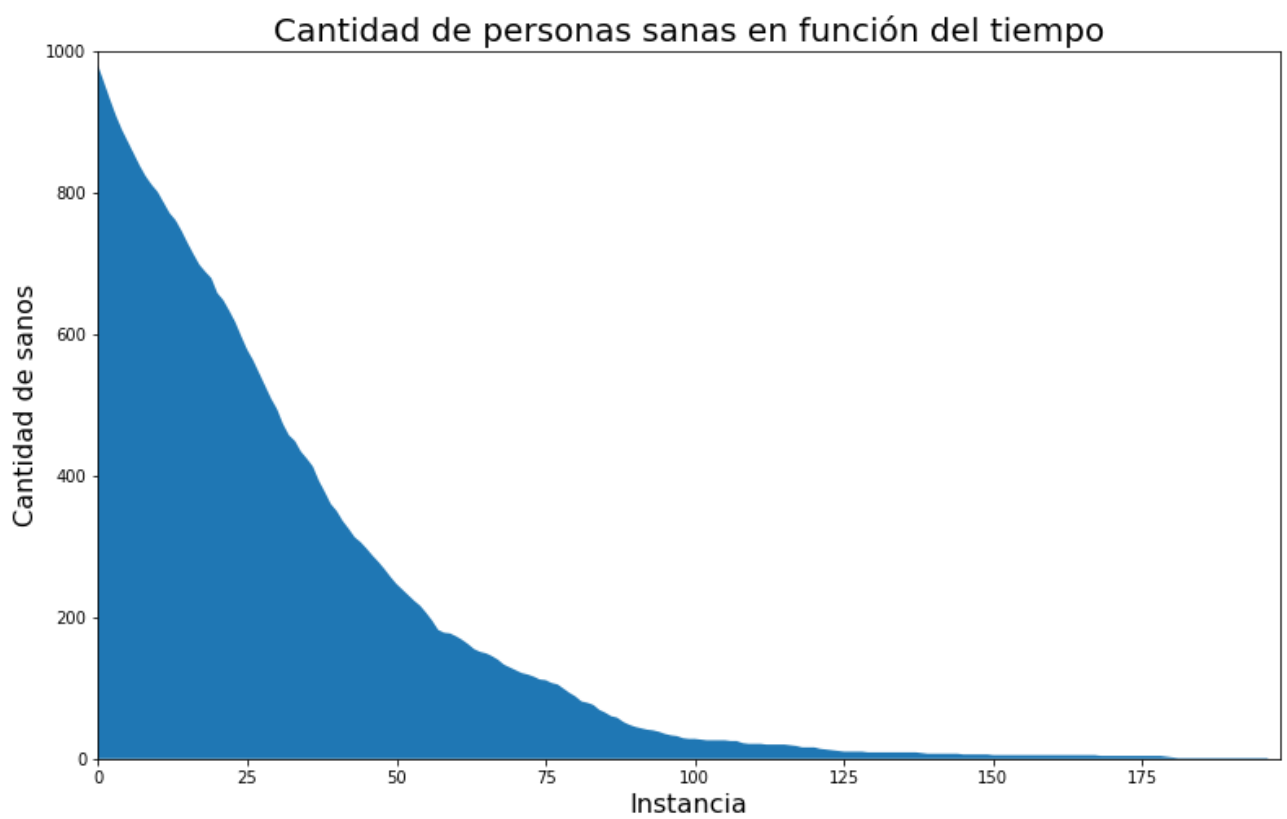


Fig. 24: sanos en cada instante para 1 000 agentes.

Los gráficos anteriores se ven similares al caso de 500 agentes. Se observa como se repite la mecánica anterior sobre la mayoría de los agentes infectados en poco tiempo, en contraste con la cantidad total de iteraciones. Por otro lado, se ve que al duplicar la cantidad de individuos se acortó a casi un cuarto el tiempo de contagio generalizado.

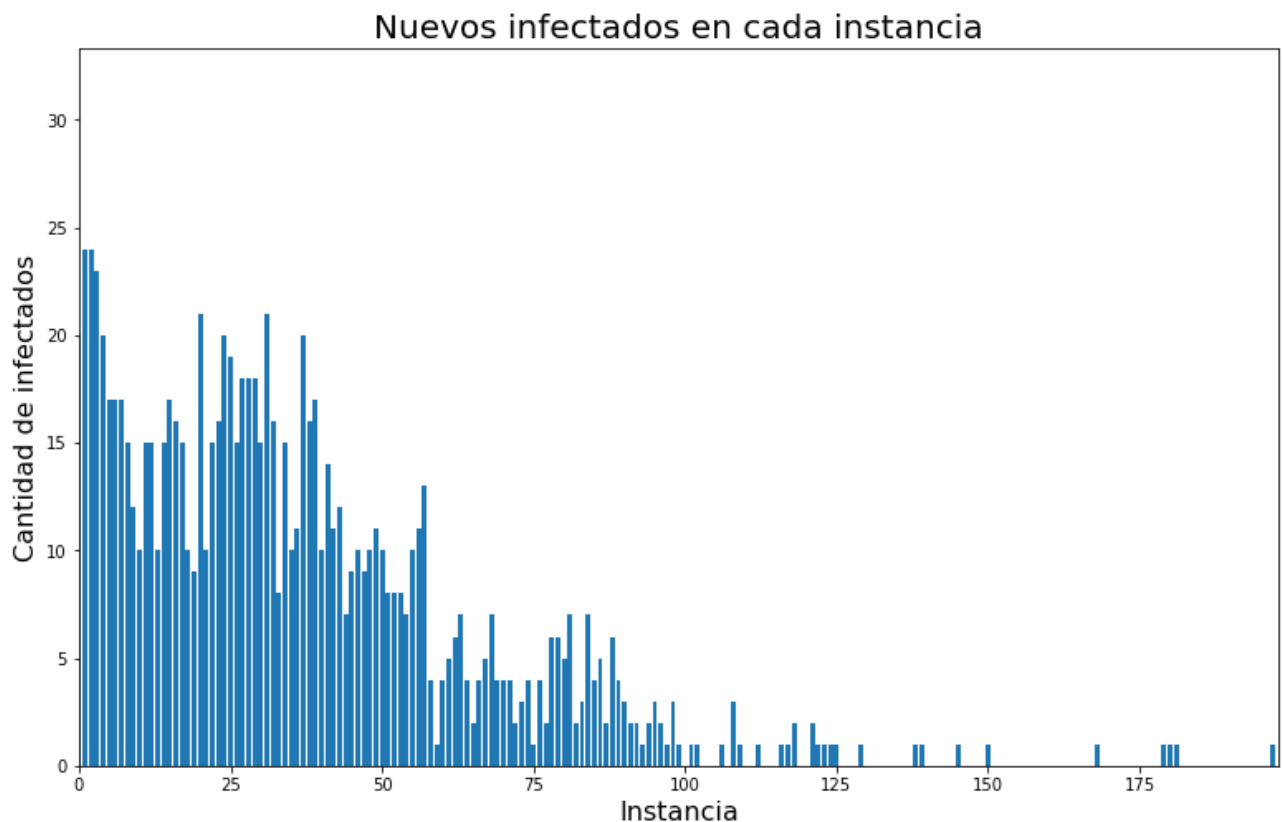


Fig. 25: nuevos infectados en cada instancia para 1 000 agentes.

Para las estadísticas sobre tiempo de contagios, el aumento de agentes parece haber seguido parcialmente la tendencia anterior. Es imaginable un pico aproximadamente en las 30 iteraciones, que excluyendo la condición inicial parecería asemejarse levemente a una distribución Poisson (dado su desplazamiento de la posición inicial). Sin embargo, la distribución en este caso parecería asemejarse en cierta forma a una exponencial debido a la fuerte alza inicial de casos. Posiblemente ninguna de las dos distribuciones sea del todo representativa para el caso general de la simulación, dada la gran variación que puede haber en la cantidad de agentes.

Caso con 2 000 agentes

La animación de ejemplo se encuentra en el siguiente enlace:

<https://drive.google.com/file/d/1KmRv0kEfmTv-QujYVhwQkLNMg9XMezw/view?usp=sharing>

Esta última animación es bastante breve dado el número de agentes involucrados. Así como en el ejemplo anterior el aumento en la cantidad de agentes redujo considerablemente el tiempo de contagio generalizado, tener cuatro veces más agentes que en el caso inicial hizo que el tiempo de contagio de todos los agentes se reduzca a un valor cercano al 5% del original.

A continuación, se muestran las estadísticas correspondientes al escenario de la simulación.

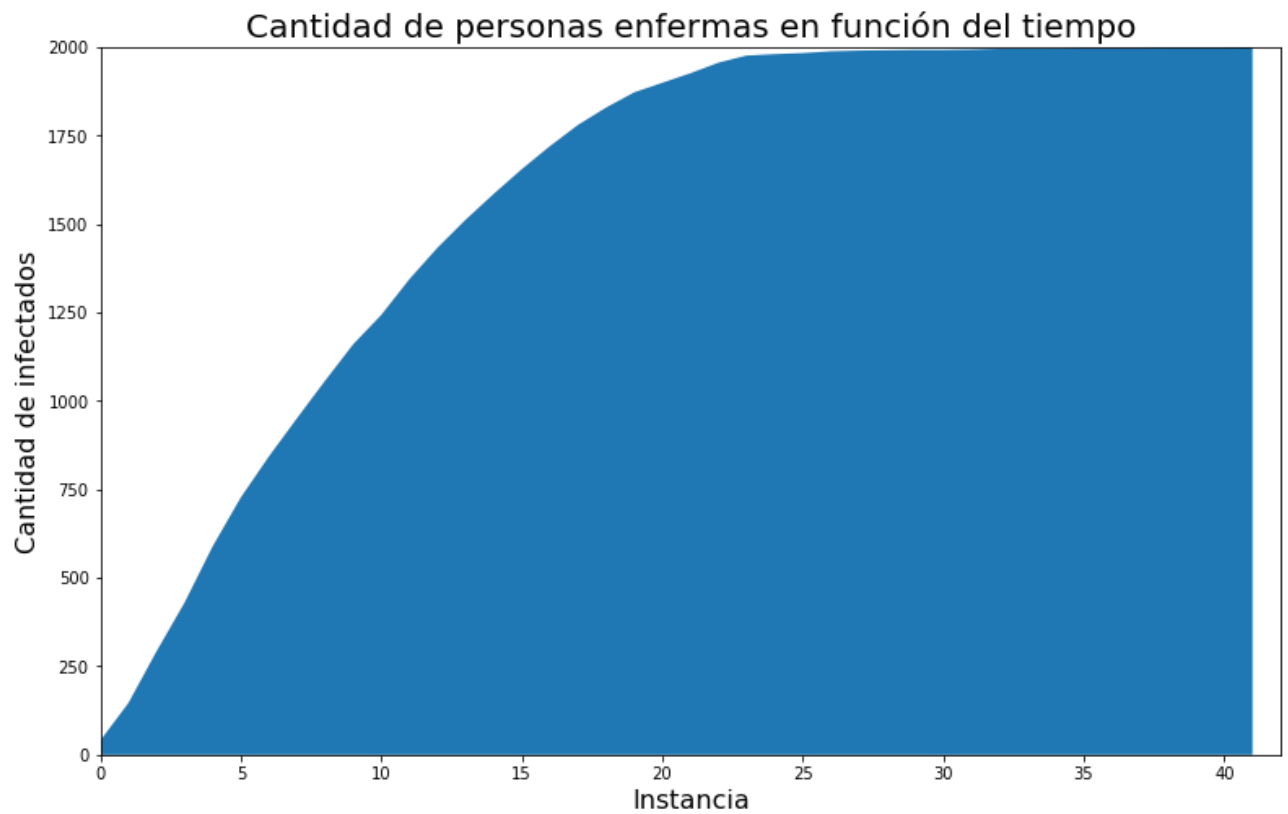


Fig. 26: enfermos en cada instante para 2 000 agentes.

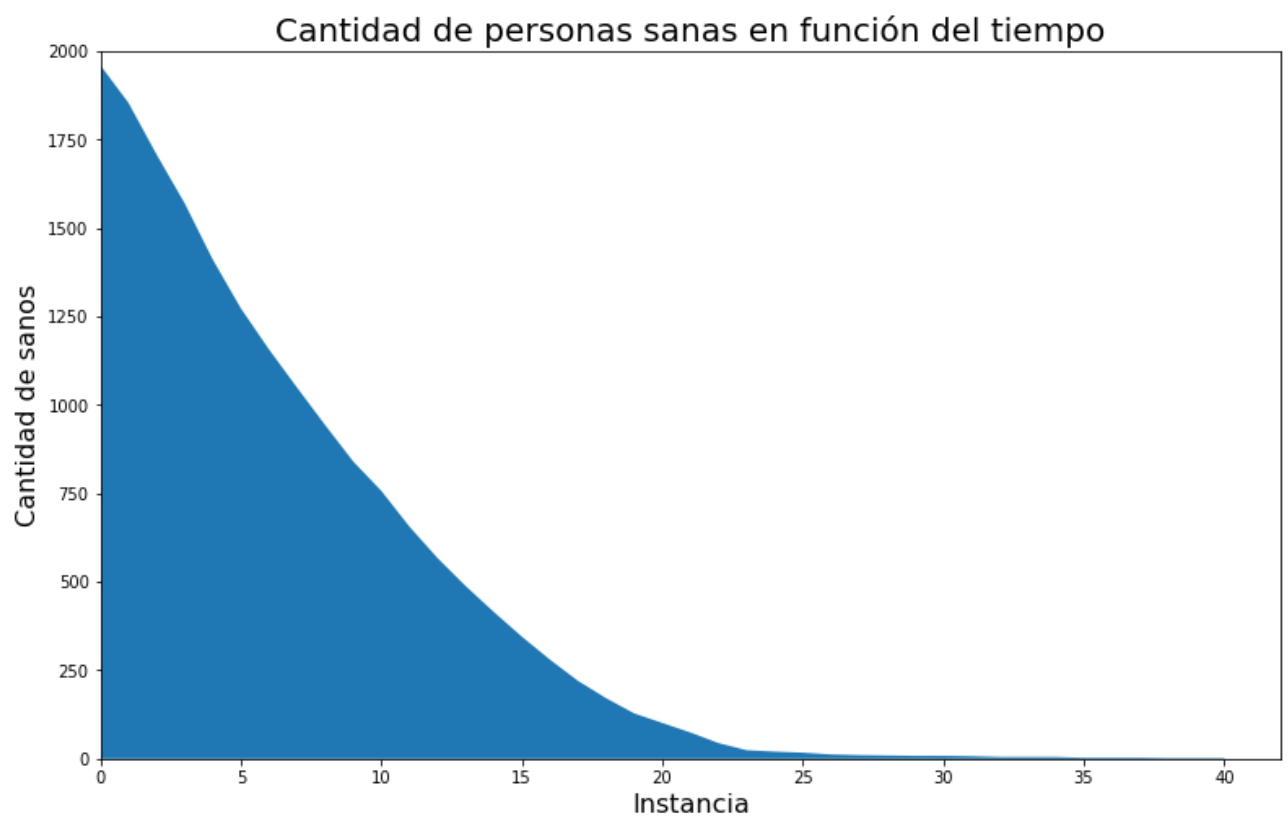


Fig. 27: sanos en cada instante para 2 000 agentes.

Más allá de que las curvas estén estilizadas debido a que los saltos de un valor a otro en las cantidad de sanos o infectados es menor, se repiten las mismas tendencias que se observaron anteriormente: un rápido aumento de casos debido al contagio de vecinos (que en este escenario es particularmente veloz debido a la cantidad de agentes).

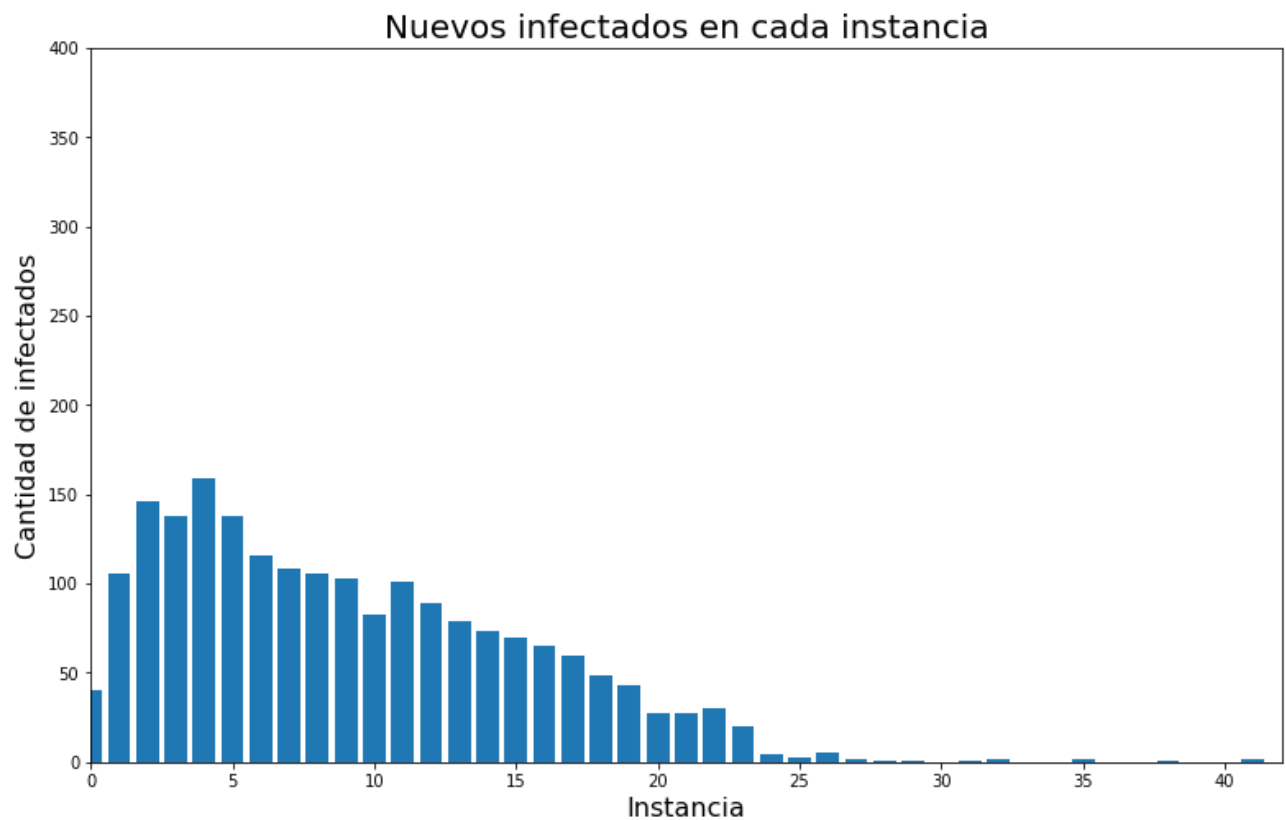


Fig. 28: nuevos infectados en cada instancia para 2 000 agentes.

En este último caso de análisis de tiempos de contagio es en donde más claramente se ve que la similitud de los gráficos con una distribución de Poisson. Quizás, al inicio del ejercicio, hubiera habido una tentación de afirmar que debido a la "falta de memoria" en la actualización de los estados del sistema, la propagación de la enfermedad debería seguir una distribución aproximadamente exponencial. Sin embargo, como ya se afirmó antes, no debe haber una distribución bien conocida que dé predictibilidad analítica sobre el modelo simulado para el caso general, ya que la dispersión de los datos varía mucho con la cantidad de agentes, y por sobre todo, el sistema presenta múltiples variables aleatorias, por lo que determinar la distribución de la mezcla no es inmediato, y cualquier apreciación a partir de una "curva de nivel" dada por una determinada cantidad de agentes está sesgada por dicho valor. Basta considerar los casos de un número de agentes suficientemente bajo (donde el contagio podría nunca propagarse) y de un número suficientemente alto (donde el contagio generalizado es instantáneo) para ver que el sistema modelado no puede describirse de forma inmediata solamente utilizando una distribución clásica, y esto es sin hablar de detalles de implementación para la forma en que la simulación es efectivamente procesada.

Conclusión

El hecho de que los agentes no puedan curarse parece tender a hacer que todos los agentes terminen infectados en bastante menos de 5 000 iteraciones (al menos un orden de magnitud menos para las cantidades simuladas de agentes, las cuáles se entienden como representativas del problema para estudiar). Por otro lado, el hecho de que el Random Walking utilizado permita ir a cualquier dirección de forma equiprobable parece llevar a los agentes a mantenerse en una posición cercana a la original, lo cual concentra la mayoría de los contagios en los instantes iniciales.

Ejercicio 8

Esquema de solución

En este caso, solo hay una modificación localizada en la forma de elegir la dirección para el Random Walking. Para cumplir con el cambio de probabilidad mencionado en la consigna, se simula una variable mezcladora con probabilidad 0,2. En el caso de que una uniforme entre 0 y 1 que se utiliza para la simulación sea menor a dicho valor, se mantiene la dirección anterior del agente en cuestión (la cual es persistida a medida que se realiza el movimiento). Si dicha variable auxiliar es mayor al valor de la variable mezcladora, entonces se elige una dirección uniformemente al azar entre las cuatro posibles, logrando la probabilidad indicada de mantener la dirección previa. No hay aclaraciones adicionales para hacer desde el lado del algoritmo de simulación ya que es el mismo.

Visualizaciones

Caso con 500 agentes

La animación de ejemplo se encuentra en el siguiente enlace:

<https://drive.google.com/file/d/1e4b9i13DUb8Bx277Tz7ugsIJYJQY9-qm/view?usp=sharing>

A continuación, se muestran las estadísticas correspondientes a dicho escenario.

Además de exhibir curvas con una pendiente mucho más definida, en este escenario la cantidad de iteraciones en que la población total se enferma disminuye. Esto es intuitivamente esperable si se entiende que aumentar la probabilidad de desplazarse en la misma dirección significa tener una mayor tendencia a alejarse de la posición inicial, y así, aumentar la propensión a estar en contacto con agentes enfermos siendo sano, o bien, con agentes sanos si se está infectado. A continuación se presentan estadísticas de los tiempos de infección.

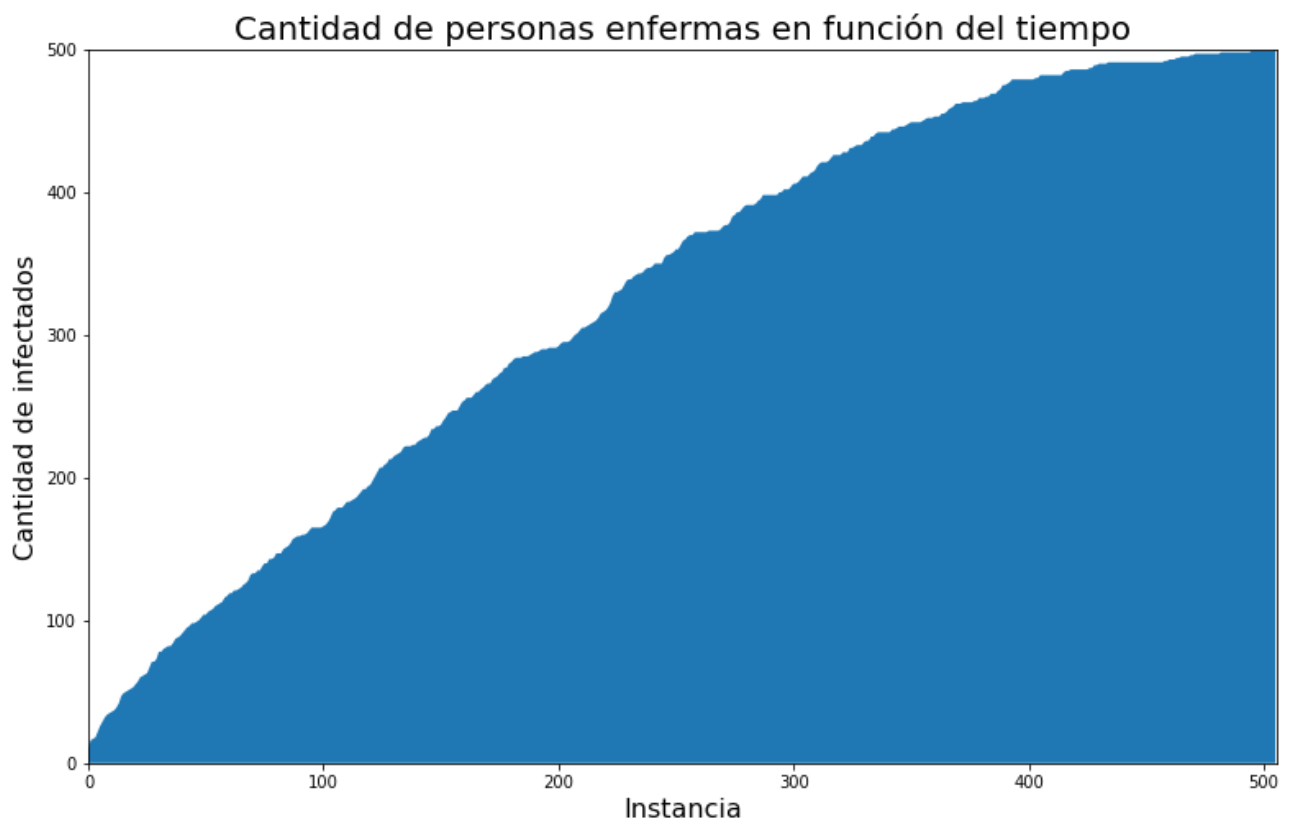


Fig. 29: enfermos en cada instante para 500 agentes con Correlated Random Walking.

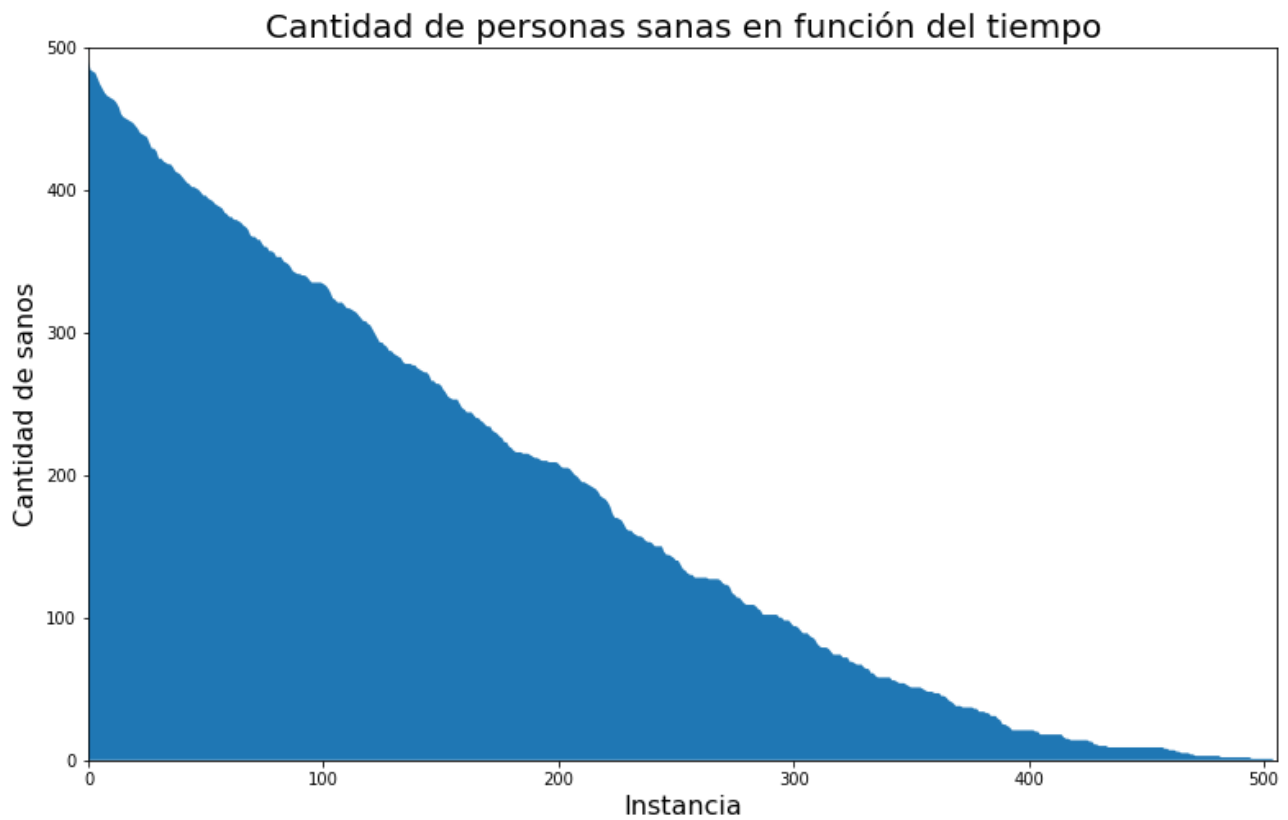


Fig. 30: sanos en cada instante para 500 agentes con Correlated Random Walking.

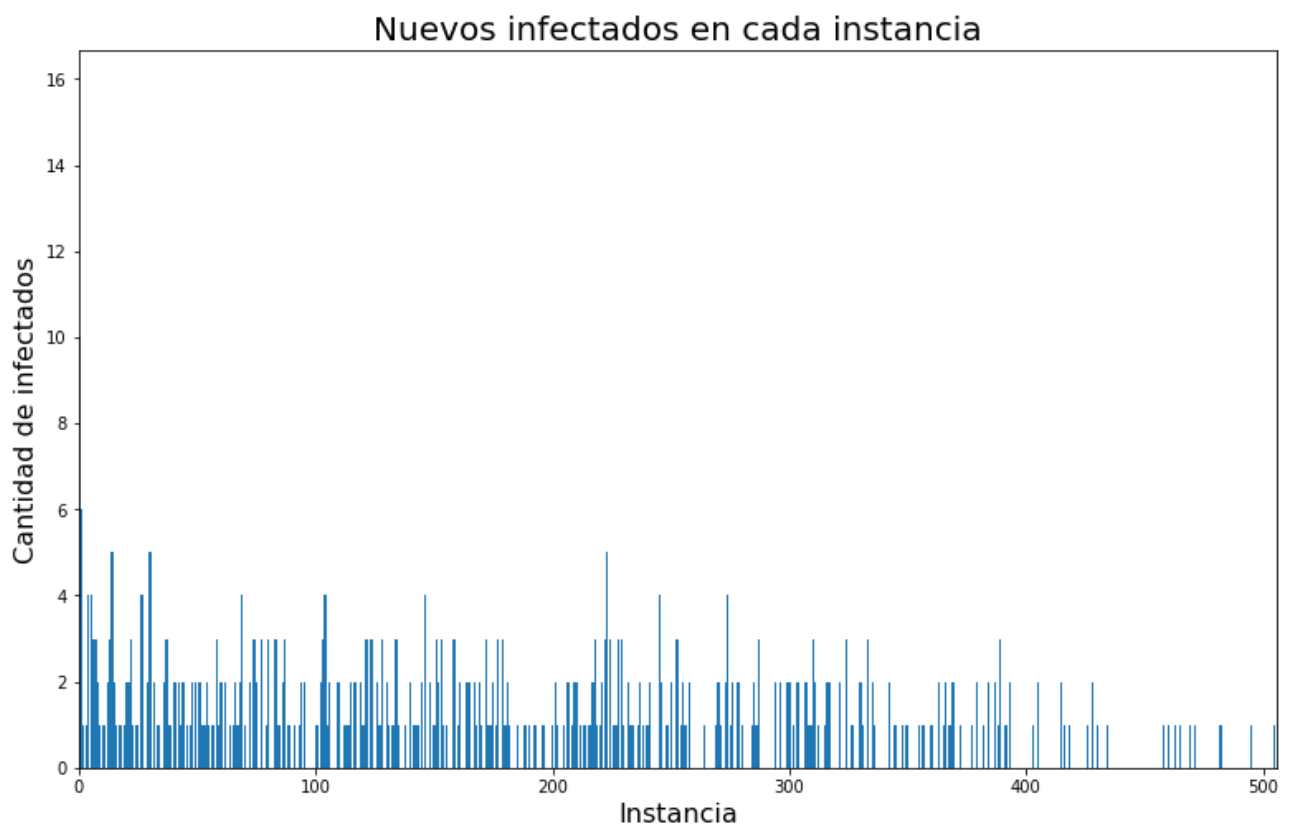


Fig. 31: nuevos infectados en cada instancia para 500 agentes con Correlated Random Walking.

Se puede observar que la mínima noción aproximada que se podía apreciar de la distribución de Poisson ahora se ve reemplazada por una tendencia más uniforme. Lo que se podría entender a partir de esto es que el Correlated Random Walking desconcentró los contagios de los instantes iniciales. Para ver si esto realmente es una tendencia (ya que tanto en este caso como en el ejercicio anterior el caso de 500 agentes es el que más dificulta ver patrones) se evalúan las evoluciones para el doble y el cuádruple de agentes, análogamente al caso anterior.

Caso con 1 000 agentes

La animación de ejemplo se encuentra en el siguiente enlace:

https://drive.google.com/file/d/1lQ91K3acBqkM8NiYwX_3jYPtHqYM8Wjq/view?usp=sharing

En este caso parece haber mucha mayor correlación con lo observado en el ejercicio anterior, más allá de la leve disminución en la cantidad de iteraciones para llegar al contagio de la totalidad de los agentes. De hecho, esto muestra que el contagio durante la fase inicial se dio de forma algo más abrupta. A continuación, se muestran las estadísticas correspondientes a dicho escenario.

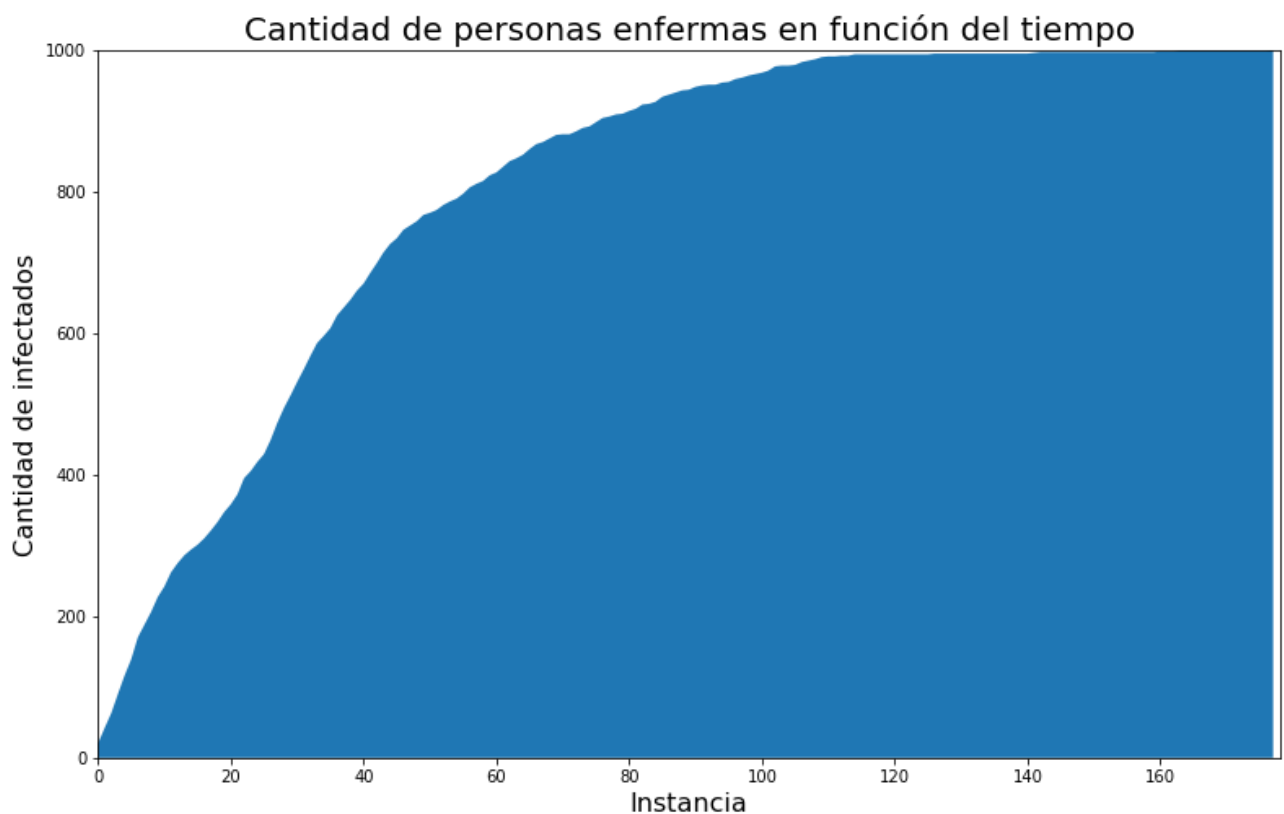


Fig. 32: enfermos en cada instante para 1 000 agentes con Correlated Random Walking.

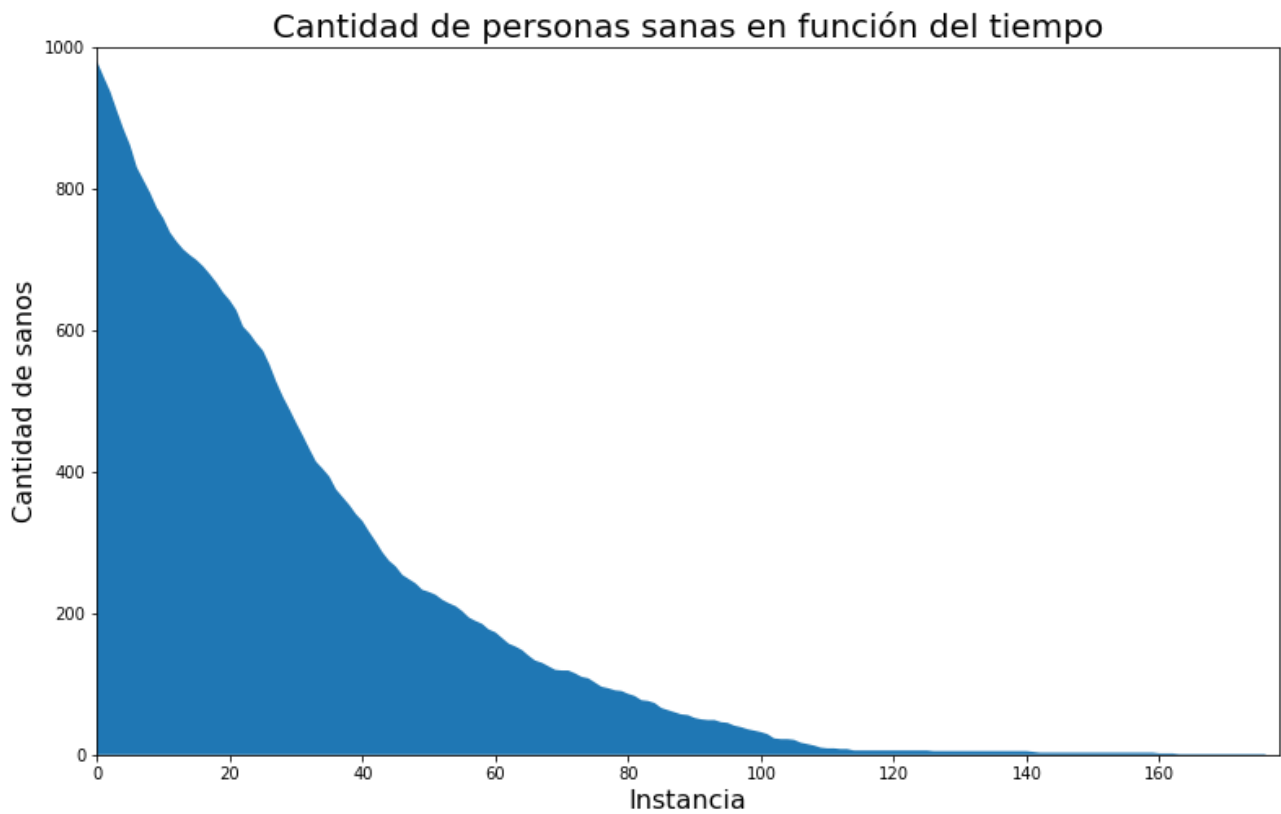


Fig. 33: sanos en cada instante para 1 000 agentes con Correlated Random Walking.

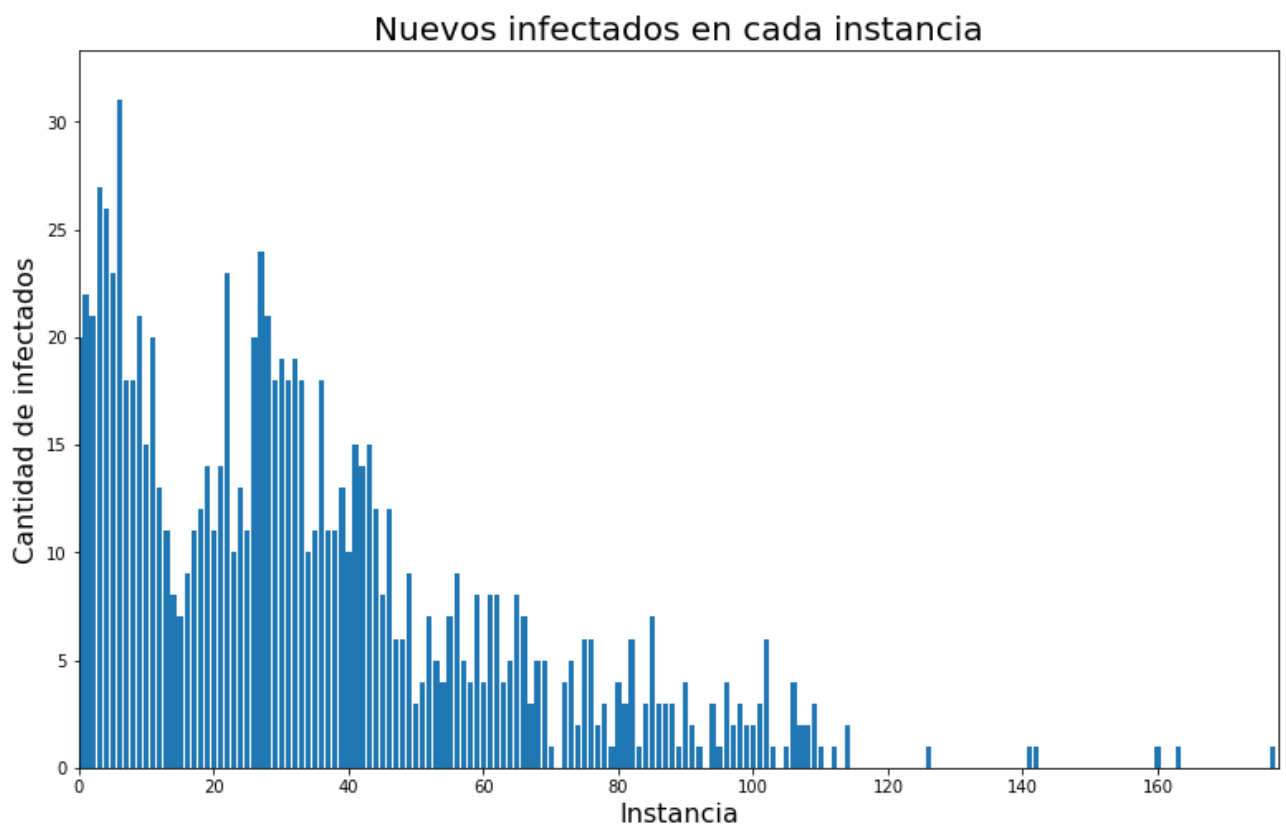


Fig. 34: nuevos infectados en cada instancia para 1 000 agentes con Correlated Random Walking.

Análogamente al ejercicio anterior, aparecen picos secundarios luego del inicio, todos ellos concentrados en los instantes iniciales. Para este caso con un mayor número de agentes, el Correlated Random Walking no parece tener un efecto marcado. Sería entendible que sucediera lo mismo al aumentar aún más los agentes, cosa que se verifica para confirmar la observación del acotado al alcance que tiene la variante aplicada para el movimiento de los agentes.

Caso con 2 000 agentes

La animación de ejemplo se encuentra en el siguiente enlace:

https://drive.google.com/file/d/1AFIAvrz-4LCTJO89FaKbBb_ktWRqER4Z/view?usp=sharing

A continuación, se muestran las estadísticas correspondientes a dicho escenario.

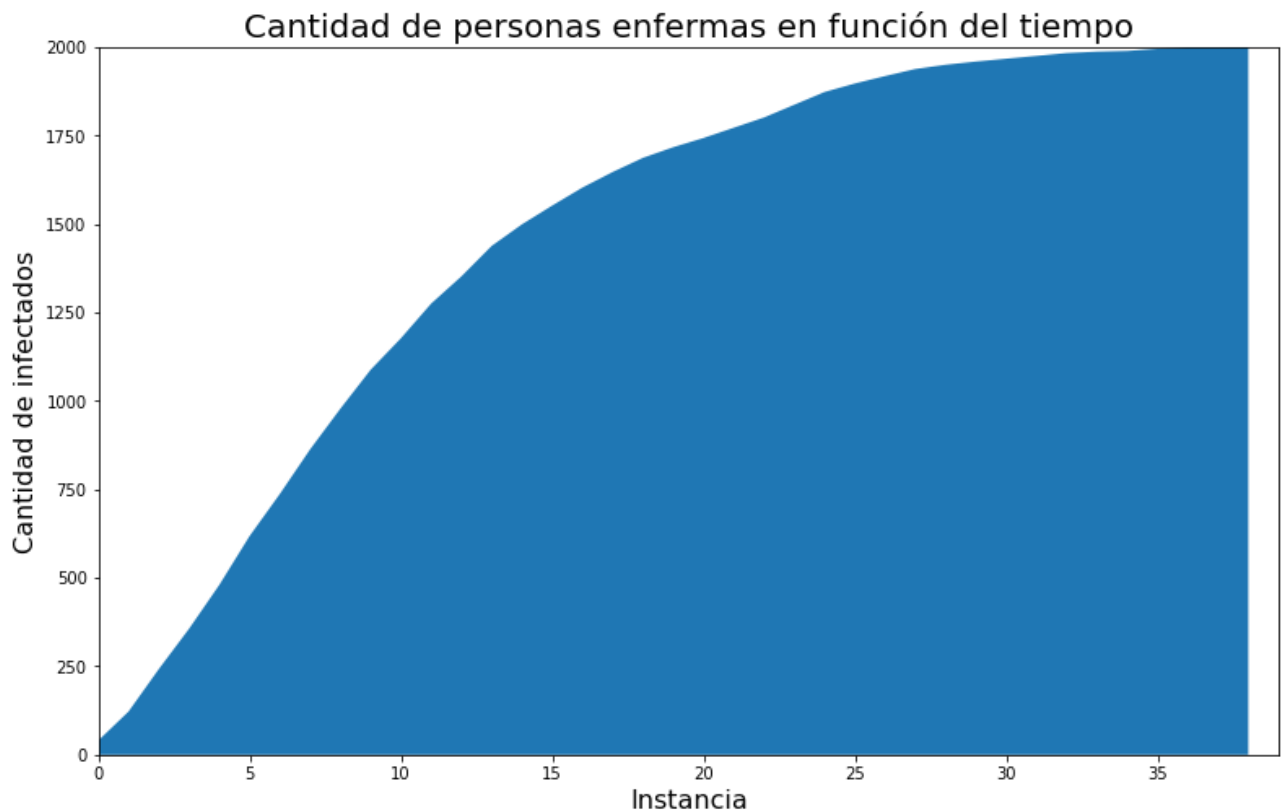


Fig. 35: enfermos en cada instante para 2 000 agentes con Correlated Random Walking.

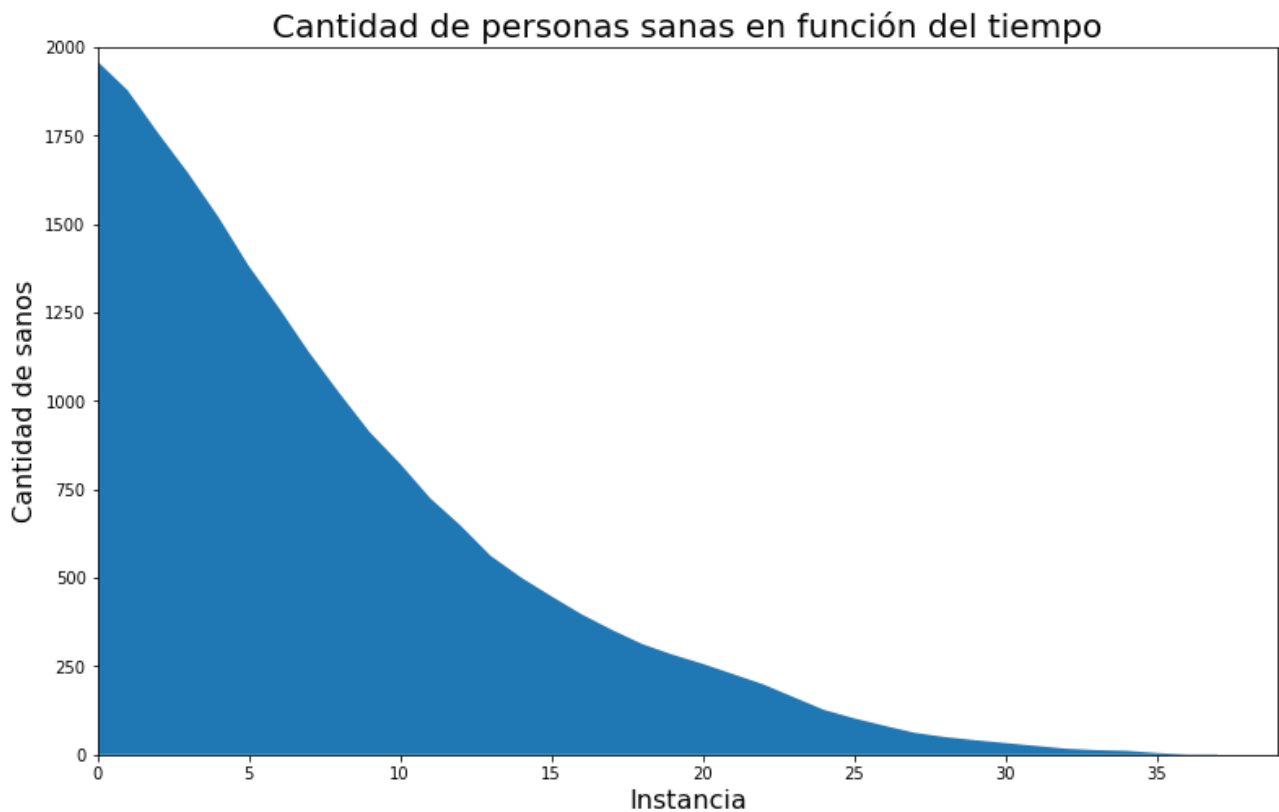


Fig. 36: sanos en cada instante para 2 000 agentes con Correlated Random Walking.

Se confirma entonces la tendencia para cantidades altas de agentes: la variante de Random Walking no es suficiente como para modificar la mecánica de contagios. Se ve que los gráficos son similares a los análogos del punto anterior.

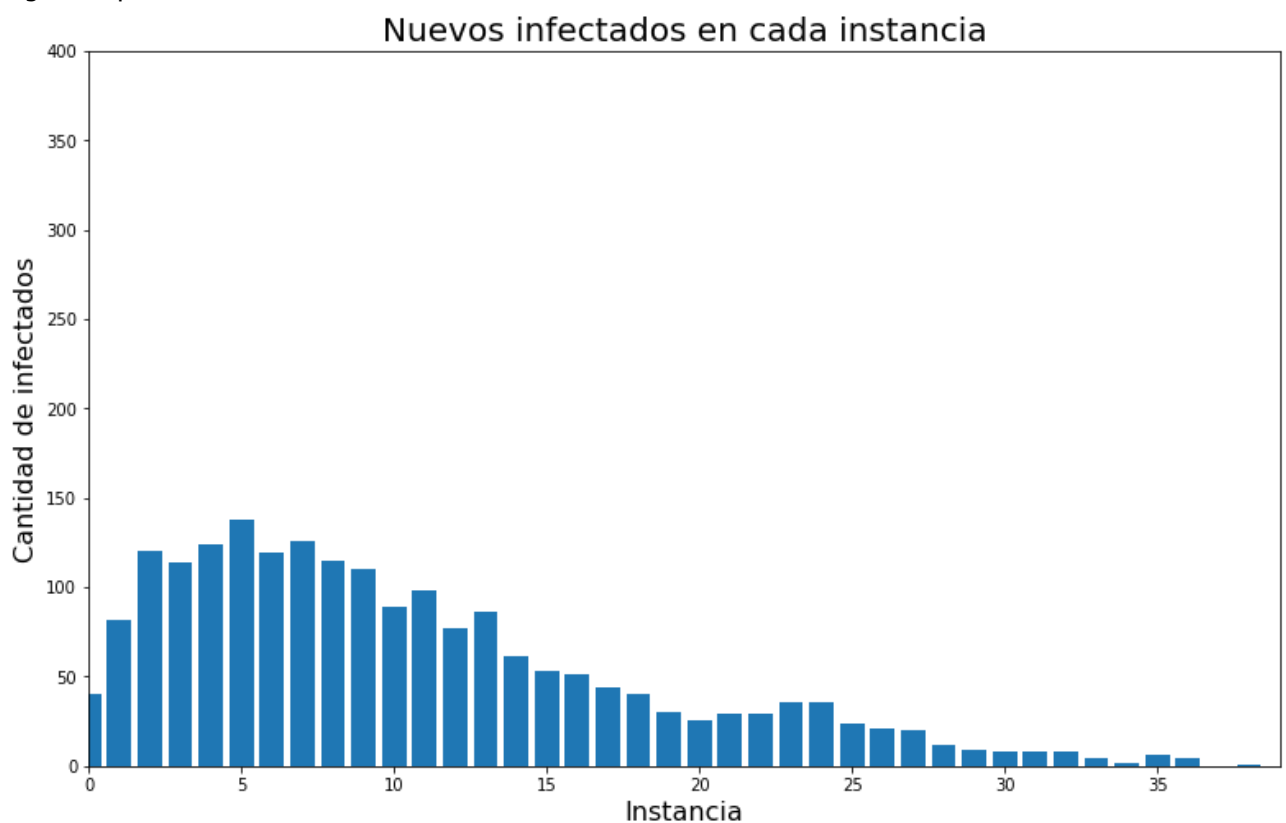


Fig. 37: nuevos infectados en cada instancia para 2 000 agentes con Correlated Random Walking.

Abandonando la idea de la tendencia más uniforme que parecía mostrarse al principio del ejercicio, y viendo los picos "movidos" con mayor cantidad de agentes, se confirma que la variante de movimiento solo tiene injerencia para casos con pocos agentes, donde los mismos no están lo suficientemente cerca desde el inicio.

Conclusión

El patrón de movimiento dado por Correlated Random Walking tiende a aumentar la velocidad de los contagios dado que propicia a que los agentes se alejen más de su posición inicial, pero esto no es suficiente como para modificar la distribución de los contagios si existen muchos agentes, ya que la sustancial mayoría ocurren al principio debido a la cercanía de los mismos.

Ejercicio 9

Esquema de solución

El esquema con el que se realizan las simulaciones en este caso es partiendo del escenario del ejercicio 8, es decir, con Correlated Random Walking. Otra cuestión de estructura básica es que ahora no solo interesa si los agentes están infectados, sino desde cuándo lo están. Es por eso que ahora la condición de infección en cada celda deja de ser la presencia de una "X" para convertirse en un número que representa hace cuántas iteraciones dicho agente está infectado. Correspondientemente, se debe cambiar la forma en que se guardan los infectados nuevos en cada iteración, y ya que se agrega la posibilidad de que los mismos se curen, también va a mantenerse un arreglo de los curados en cada instancia como estadística adicional.

Algoritmo de simulación

Se agrega un ciclo más al inicio para curar a los agentes enfermos cuando las condiciones de α y β lo permitan, lo cual es posible de verificar gracias a que se mantienen la cantidad de iteraciones desde que el agente está infectado. Para curar a un agente, simplemente se establece en 0 el valor correspondiente al tiempo de infección.

Adicionalmente, para este ejercicio se distinguen el caso donde existe inmunidad, es decir, la imposibilidad de volver a contagiarse por parte de los agentes que ya se enfermaron, y el caso en donde no. Para esto, simplemente se mantiene una variable que indica si el sistema presenta inmunidad posterior al contagio, y un diccionario que almacena a los agentes inmunes (es decir, curados). El mismo se consulta a la hora de infectar en caso de que exista inmunidad. En otro caso, la forma de infección es similar.

Visualizaciones

9.1. Sin inmunidad

2 000 agentes, $\alpha = 20$, $\beta = 0,1$

La animación de ejemplo se encuentra en el siguiente enlace:

<https://drive.google.com/file/d/1WvzeJ47o1ux0-uePEErCOGtKnpiVtaDO>

El hecho de que hubiera inmunidad hizo que el sistema pudiera evolucionar con un equilibrio dinámico hasta las 5 000 iteraciones, pero el enlace anterior contiene una animación solo con las primeras 500. Esto es debido a que, por un lado, animar 5 000 iteraciones parece requerir más memoria de la que Google Colab puede ofrecer sin que se caiga la sesión (al menos, con la herramienta elegida: plotly), y por otro, el sistema parece entrar en régimen rápidamente, con lo que se tiene que incluso una décima parte de la evolución del sistema muestra bastante redundancia en cuanto a su comportamiento, como se verá en los gráficos de estadísticas posteriores.

Esto es decir que si bien existe la posibilidad de recuperarse de la enfermedad, la imposibilidad de inmunización logra un equilibrio entre la tasa de contagios y recuperaciones que mantiene al sistema en un estado cuasi estacionario con leves perturbaciones.

Se atribuye el hecho de que el sistema haya encontrado el equilibrio tan rápidamente a la gran cantidad de agentes, la cual posiblemente incluso para valores un poco más bajos α y algo más altos de β hubieran causado el mismo "efecto de ahogamiento" a la minoría de agentes que se recupera. Dicho de otra forma, si existe gran cantidad de agentes, y por ende, gran cantidad de vecinos cercanos para cada persona, la tasa de recuperación debe ser muy rápida para intentar detener el contagio dado por el elevado número de personas en un espacio acotado. A continuación, se muestran las estadísticas correspondientes a dicho escenario.

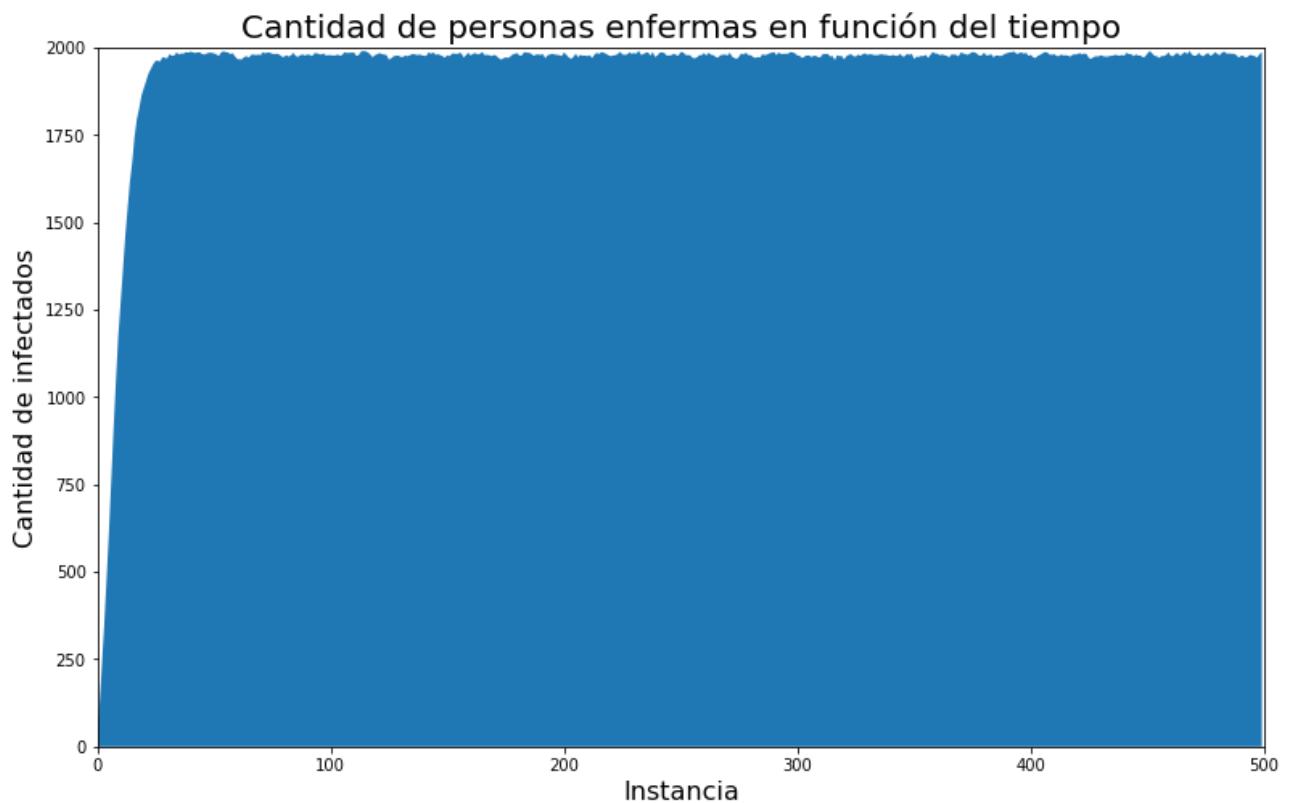


Fig. 38: enfermos en cada instante para $N = 2\,000$, $\alpha = 20$, $\beta = 0,1$.

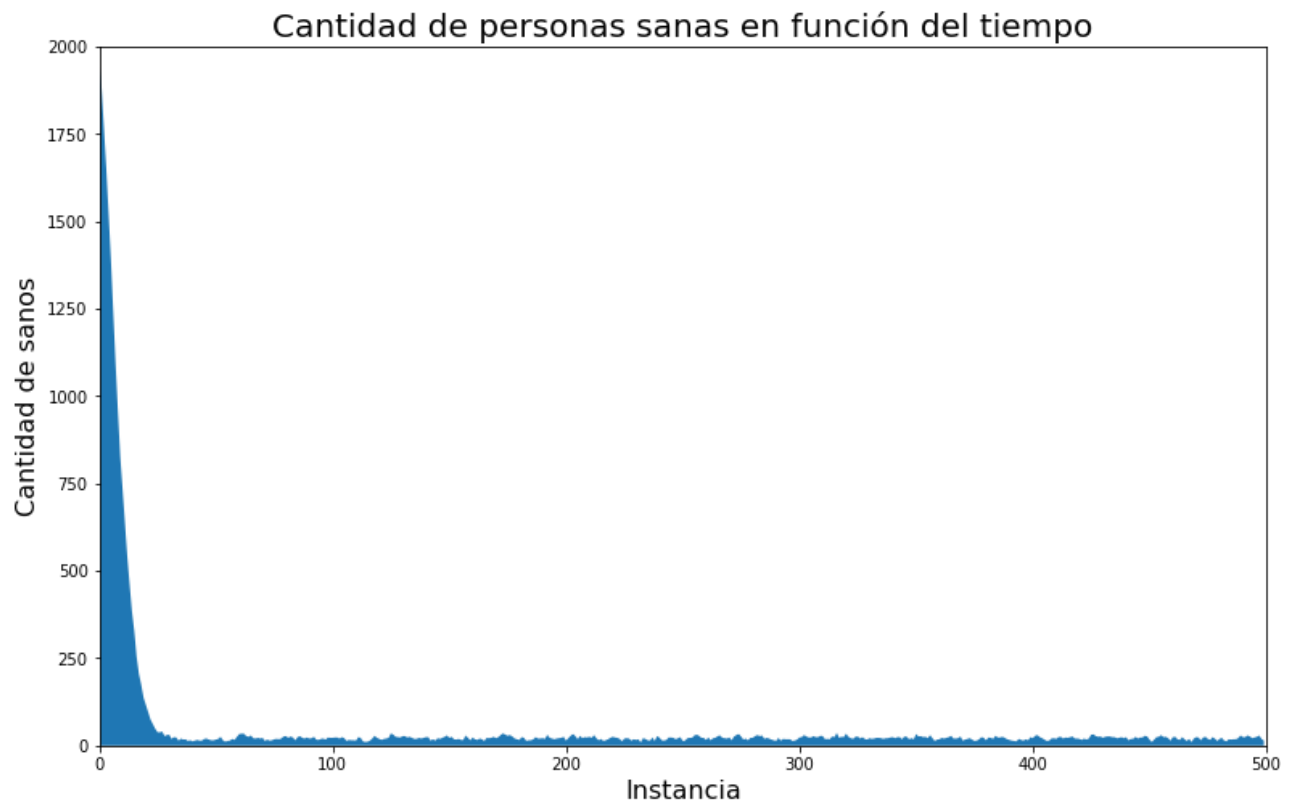


Fig. 39: sanos en cada instante para para $N = 2\,000$, $\alpha = 20$, $\beta = 0,1$.

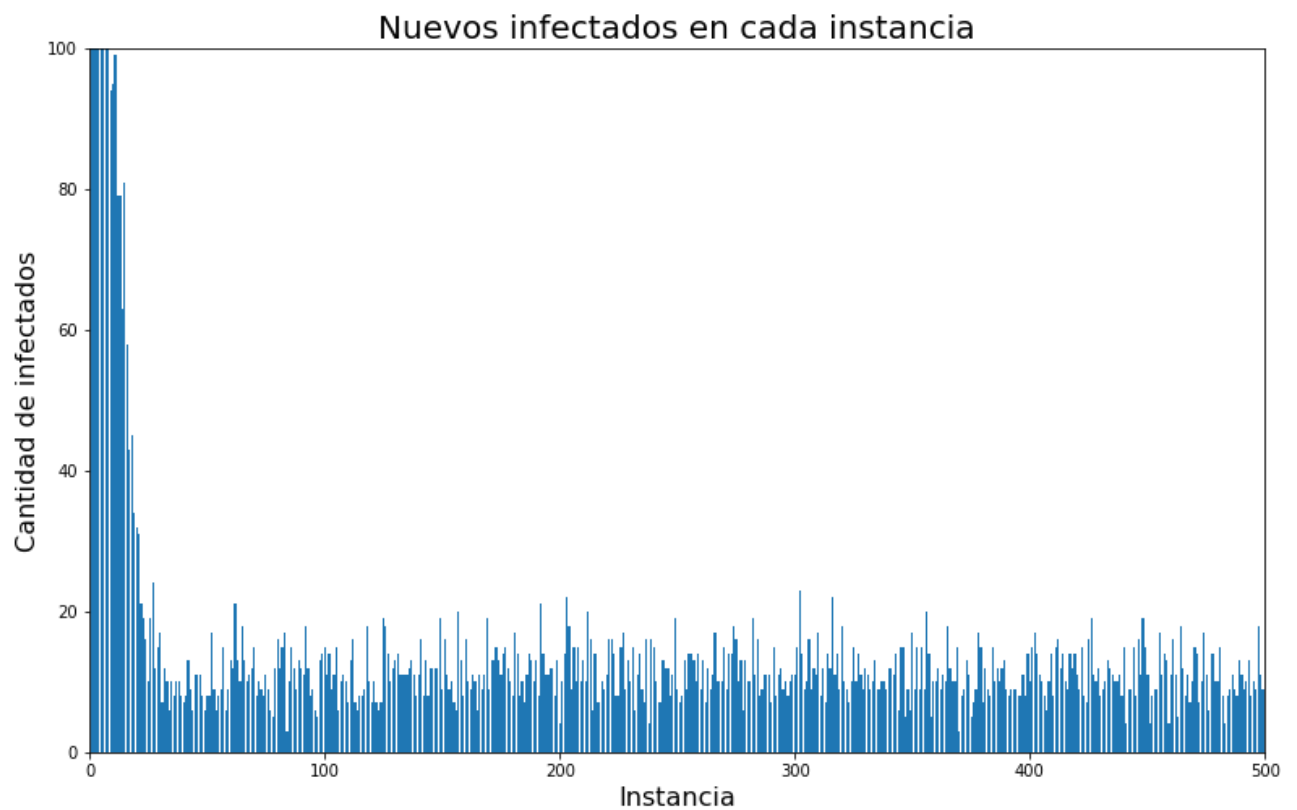


Fig. 40: nuevos infectados en cada instancia para $N = 2\,000$, $\alpha = 20$, $\beta = 0,1$.

Para el caso de nuevos infectados se ve una elevada cantidad de casos iniciales que luego entran en el equilibrio apreciado más arriba. Al igual que en casos de ejercicios anteriores, al haber más agentes de un cierto tipo, sus distribuciones de tiempos de infección se condicen más con el gráfico asociado al que tiene la totalidad de los agentes, pero en todos los casos las tendencias son similares. Un equilibrio dinámico similar puede verse en cuanto a las distribuciones de tiempos de recuperación.

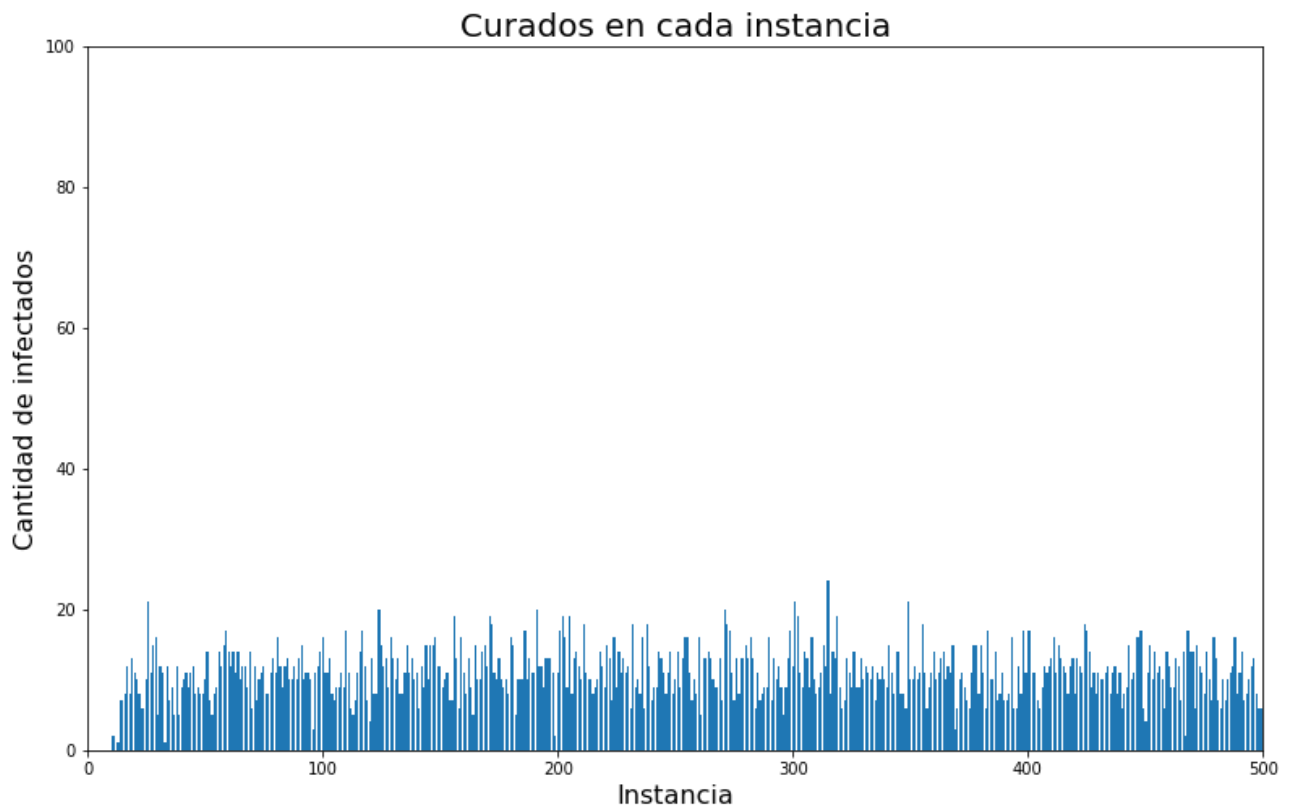


Fig. 41: nuevos recuperados en cada instancia para $N = 2\,000$, $\alpha = 20$, $\beta = 0,1$.

Análogamente para el caso de los agentes recuperados, se ve el equilibrio alcanzado y la correspondencia con la distribución general dependiendo del número de agentes. Como comentario complementario, se visualiza que al inicio no hay recuperados, lo cual es coherente ya que coincide con el valor de la constante dada para empezar a tener posibilidad de recuperarse.

1 000 agentes, $\alpha = 10$, $\beta = 0,5$

La animación de ejemplo se encuentra en el siguiente enlace:

https://drive.google.com/file/d/12cCGbyBVLShABX4J5ExppYNvOO8_XXpn/edit

En la misma tampoco se incluyen la totalidad de iteraciones ya que el sistema rápidamente llega a un punto de equilibrio dinámico.

A continuación, se muestran las estadísticas correspondientes a dicho escenario.

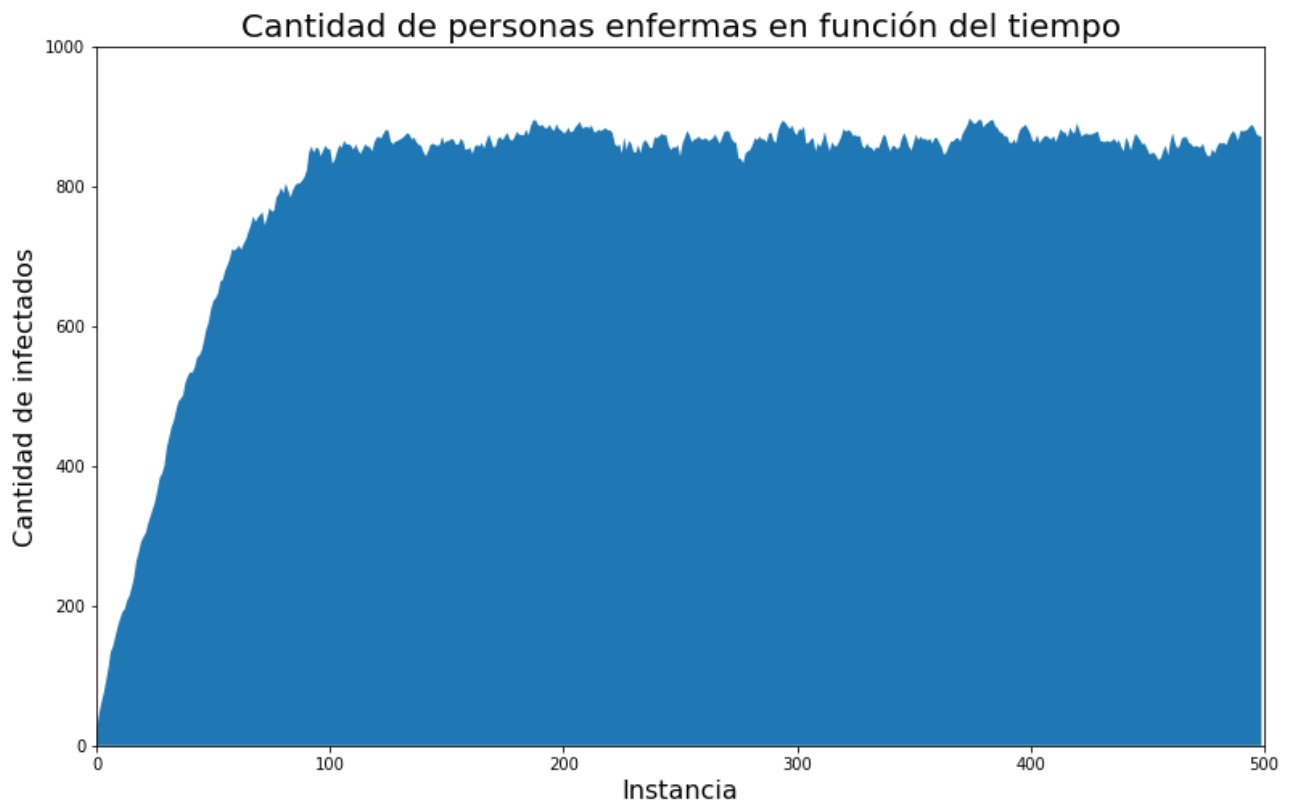


Fig. 42: enfermos en cada instante para $N = 1\,000$, $\alpha = 10$, $\beta = 0,5$.

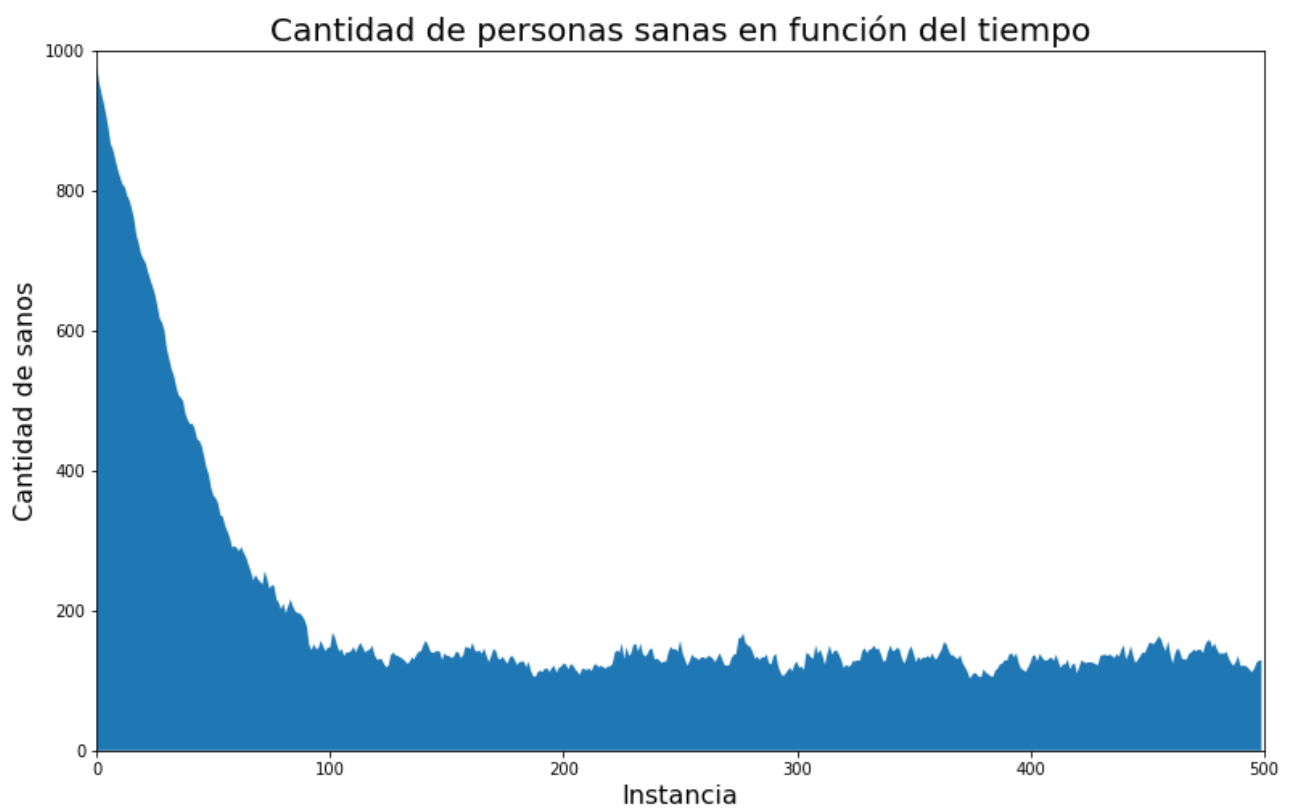


Fig. 43: sanos en cada instante para para $N = 1\,000$, $\alpha = 10$, $\beta = 0,5$.

El sistema tarda más en volverse "pseudo estacionario", y el equilibrio alcanzado es con un nivel de infectados total menor (esto último se puede visualizar en la animación como "más puntos verdes" al entrar en régimen). Incluso para la mitad de los agentes, en contraste con el ejercicio anterior, y con una mucho mayor chance de recuperación y en menos tiempo, la tasa de infección y recuperación parecen equipararse rápidamente.

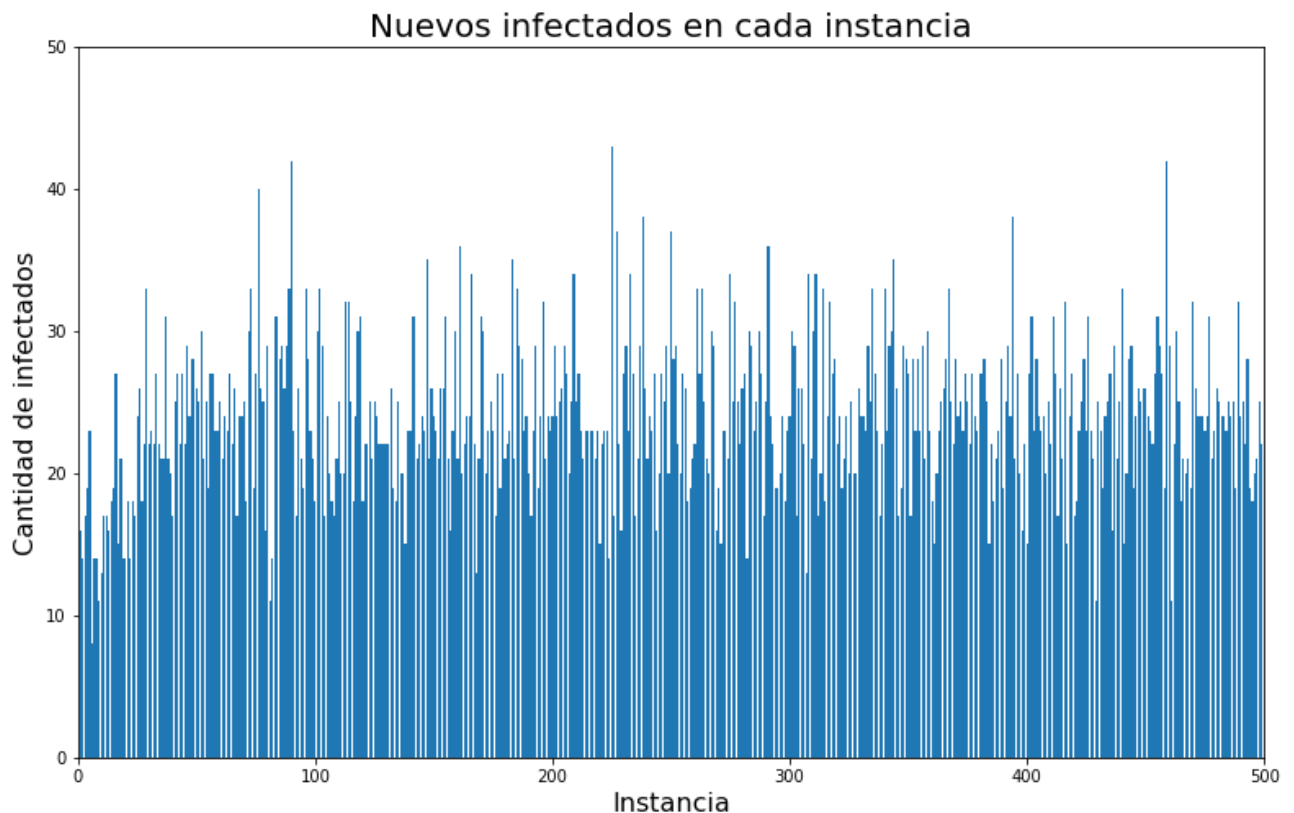


Fig. 44: nuevos infectados en cada instancia para $N = 1\,000$, $\alpha = 10$, $\beta = 0,5$.

Las distribuciones anteriores muestran tendencias diferentes al caso anterior ya que al haber más agentes sanos en un instante dado, la cantidad que puede infectarse es correspondientemente mayor, lo cual se puede ver traducido en un régimen regular de contagios más alto. Esto efectivamente se corresponde con los gráficos a continuación.

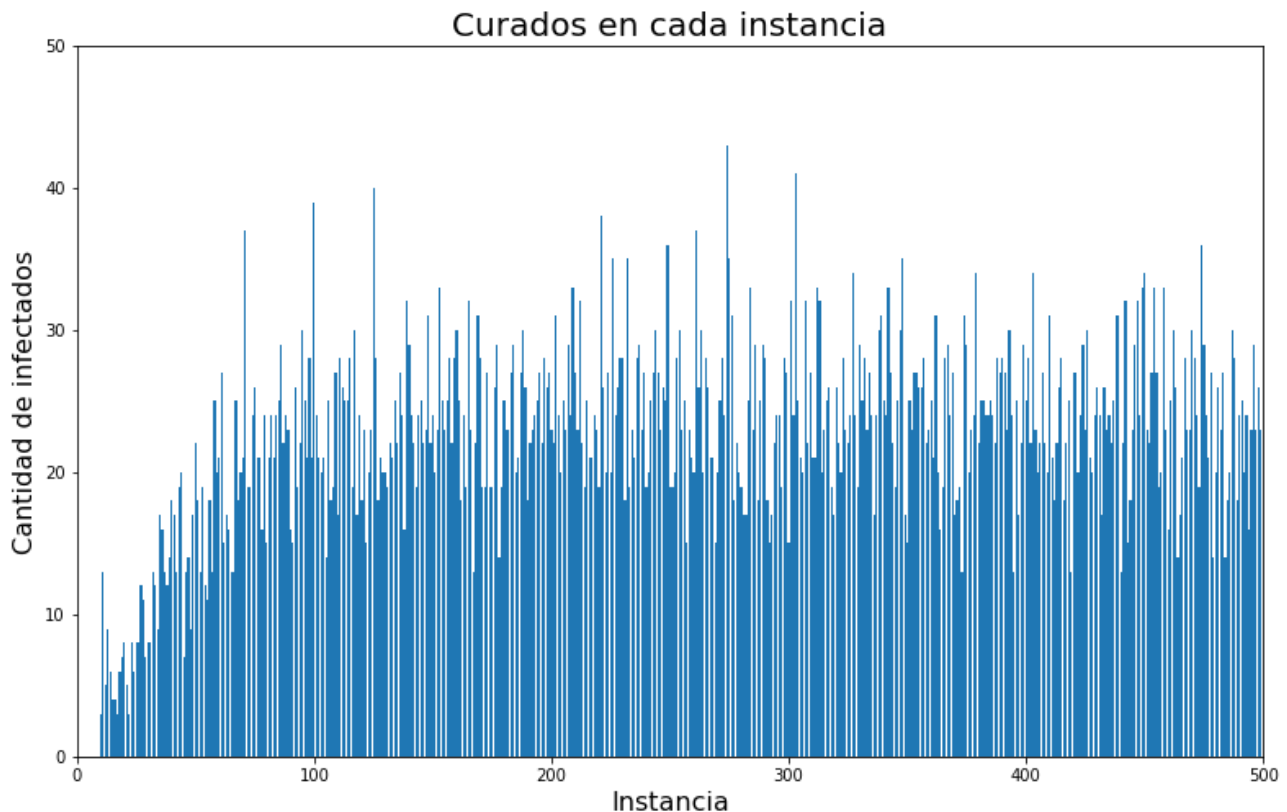


Fig. 45: nuevos recuperados en cada instancia para $N = 1\,000$, $\alpha = 10$, $\beta = 0,5$.

1 000 agentes, $\alpha = 10$, $\beta = 0,7$

La animación de ejemplo se encuentra en el siguiente enlace:

<https://drive.google.com/file/d/1OT3QH0ok3uFMT8j6Yy9DSsDenvXte-E/view?usp=sharing>

Se observa que aumentar aún más la probabilidad de recuperación luego del contagio no tiene una inferencia sustancial, al igual que en el caso anterior. Las curvas se ven muy similares en ambos casos, por lo que la cantidad de agentes parece obligar a un equilibrio dinámico más allá del valor de β . Esta similitud también se da en los gráficos a continuación, por lo que no hay cuestiones importantes a resaltar. A continuación, se muestran las estadísticas correspondientes a dicho escenario.

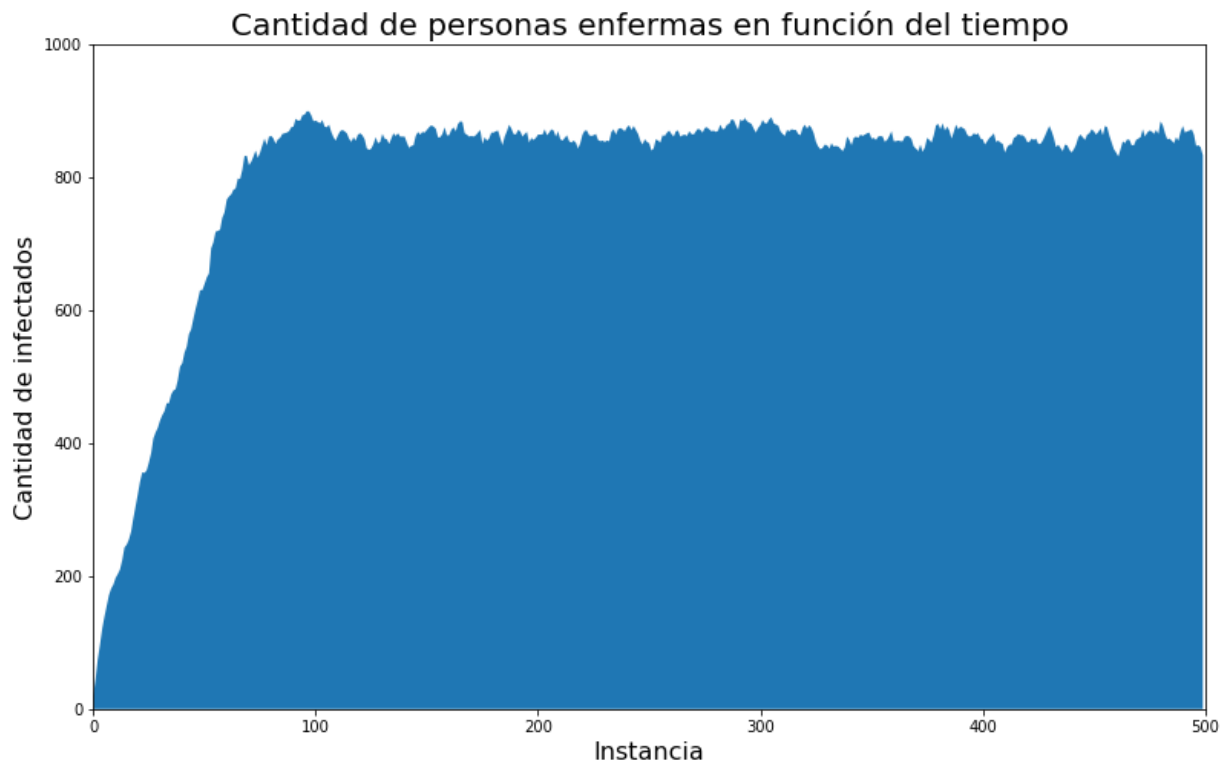


Fig. 46: enfermos en cada instante para $N = 1\,000$, $\alpha = 10$, $\beta = 0,7$.

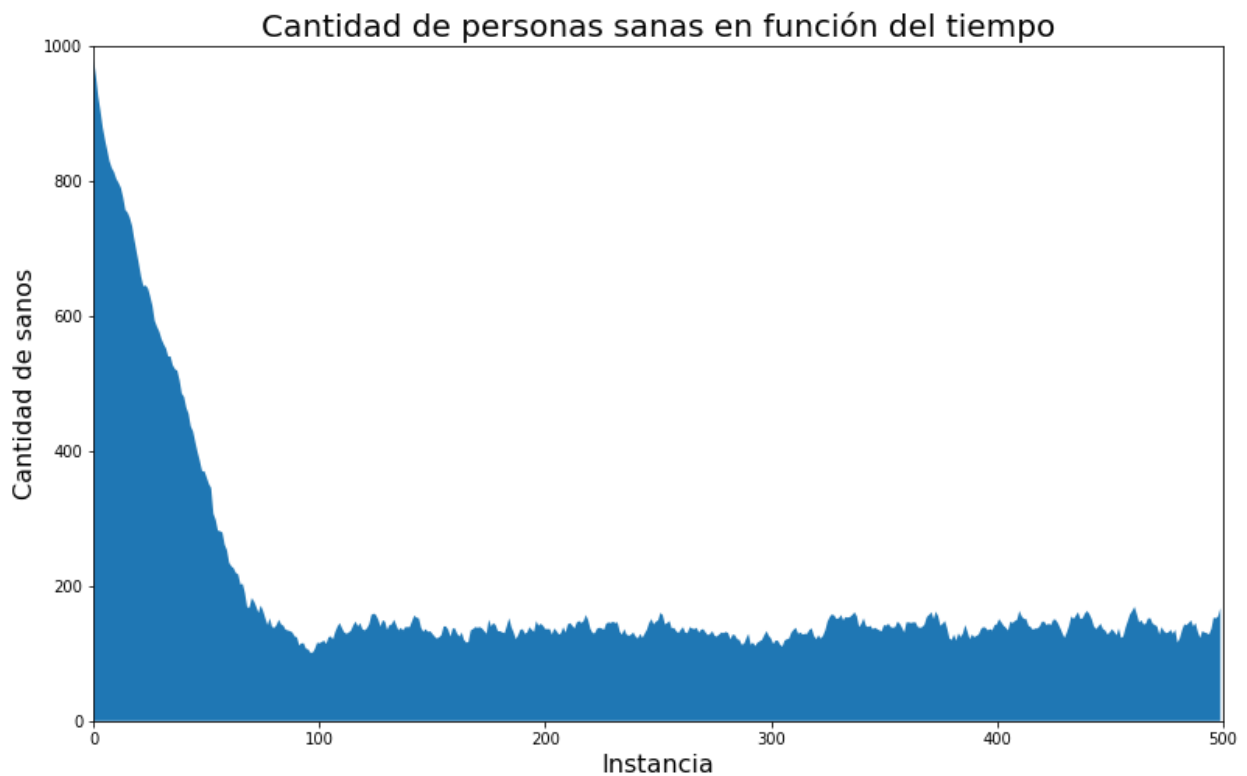


Fig. 47: sanos en cada instante para para $N = 1\,000$, $\alpha = 10$, $\beta = 0,7$.

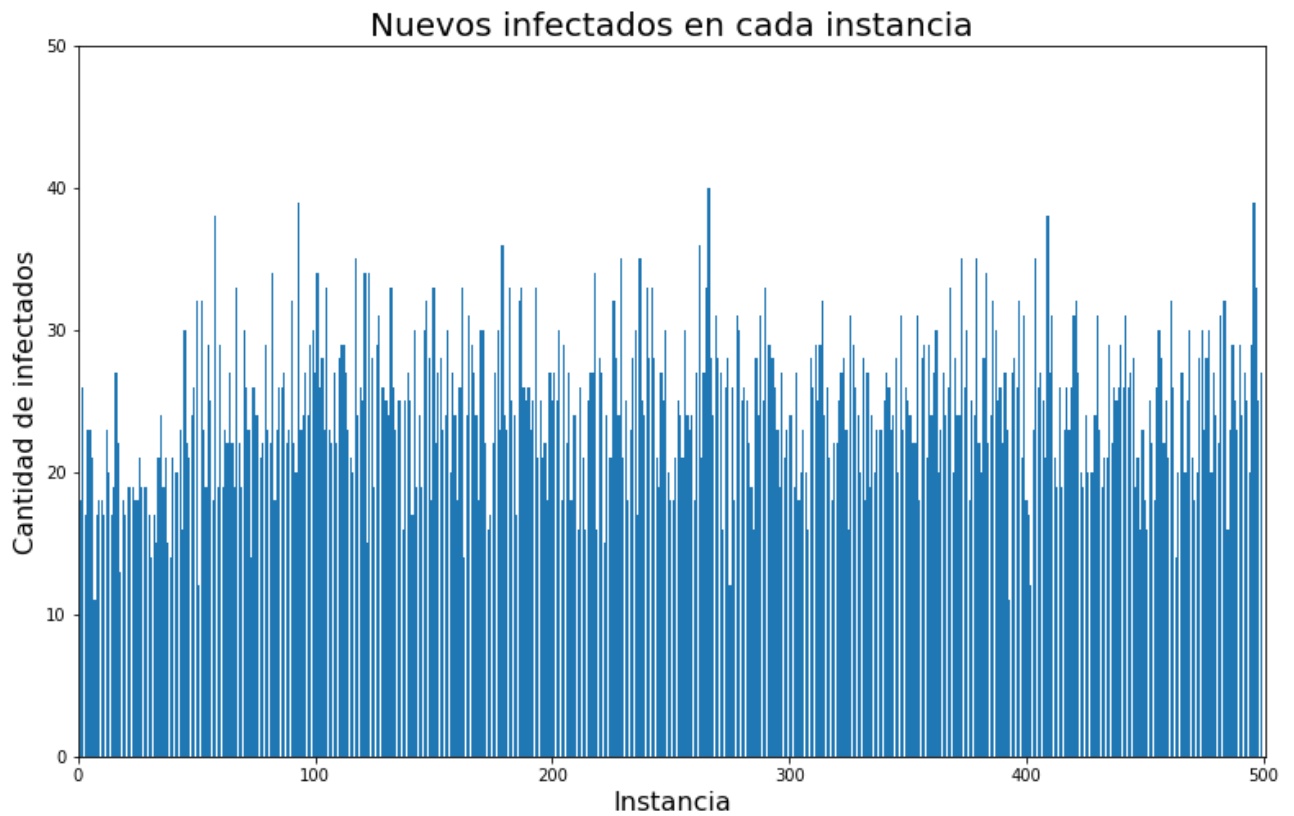


Fig. 48: nuevos infectados en cada instancia para $N = 1\,000$, $\alpha = 10$, $\beta = 0,7$.

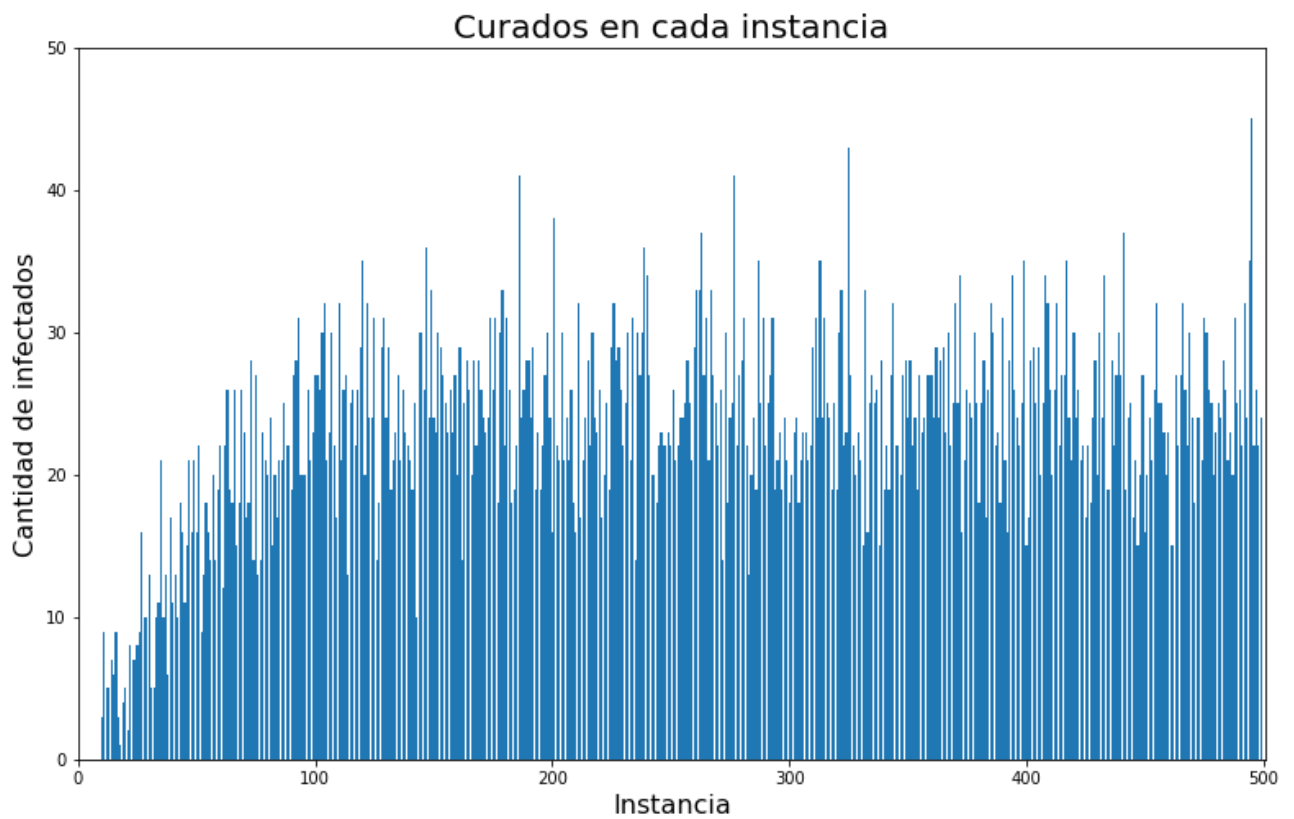


Fig. 49: nuevos recuperados en cada instancia para $N = 1\,000$, $\alpha = 10$, $\beta = 0,7$.

9.2. Con inmunidad

2 000 agentes, $\alpha = 10$, $\beta = 0,1$

La animación de ejemplo se encuentra en el siguiente enlace:

https://drive.google.com/file/d/1ALOiVgDOKIPkIB7WocJjy_9lWYQrI4yw/view?usp=sharing

La curva de contagios parece mostrar un comportamiento similar a los ejercicios 7 y 8 hasta su correspondiente pico, en cuanto al rápido aumento de contagios debido a la cercanía de agentes para un valor alto de N . Cuando el efecto de la inmunidad empieza a ser notorio, la cantidad de infectados cae abruptamente hasta terminar con la epidemia. Se entiende que la curva no se ve del todo simétrica debido a que la probabilidad de recuperación luego de las 10 iteraciones estando infectado no es tan alta. Si lo fuera, una vez pasado la cantidad de iteraciones dado por α , sería esperable que los agentes se curaran tan rápido como se contagiaron. En cualquier caso, se ve que la cantidad de infectados que no contribuye a la simetría integra una "long tail" de minoría de agentes, ya que a cada iteración que pasa luego de la infección, la recuperación es más probable. A continuación, se muestran las estadísticas correspondientes a dicho escenario.

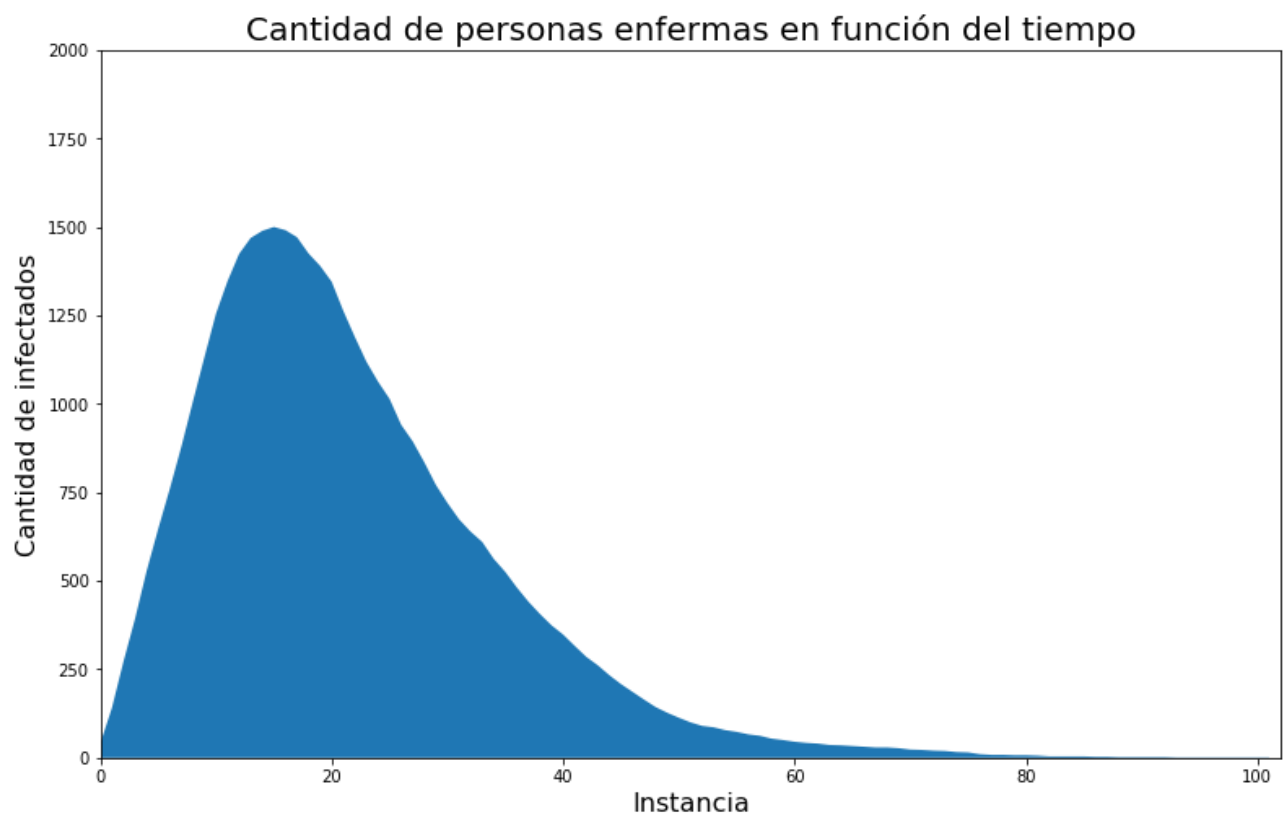


Fig. 50: enfermos en cada instante para $N = 2\,000$, $\alpha = 10$, $\beta = 0,1$ con inmunidad.

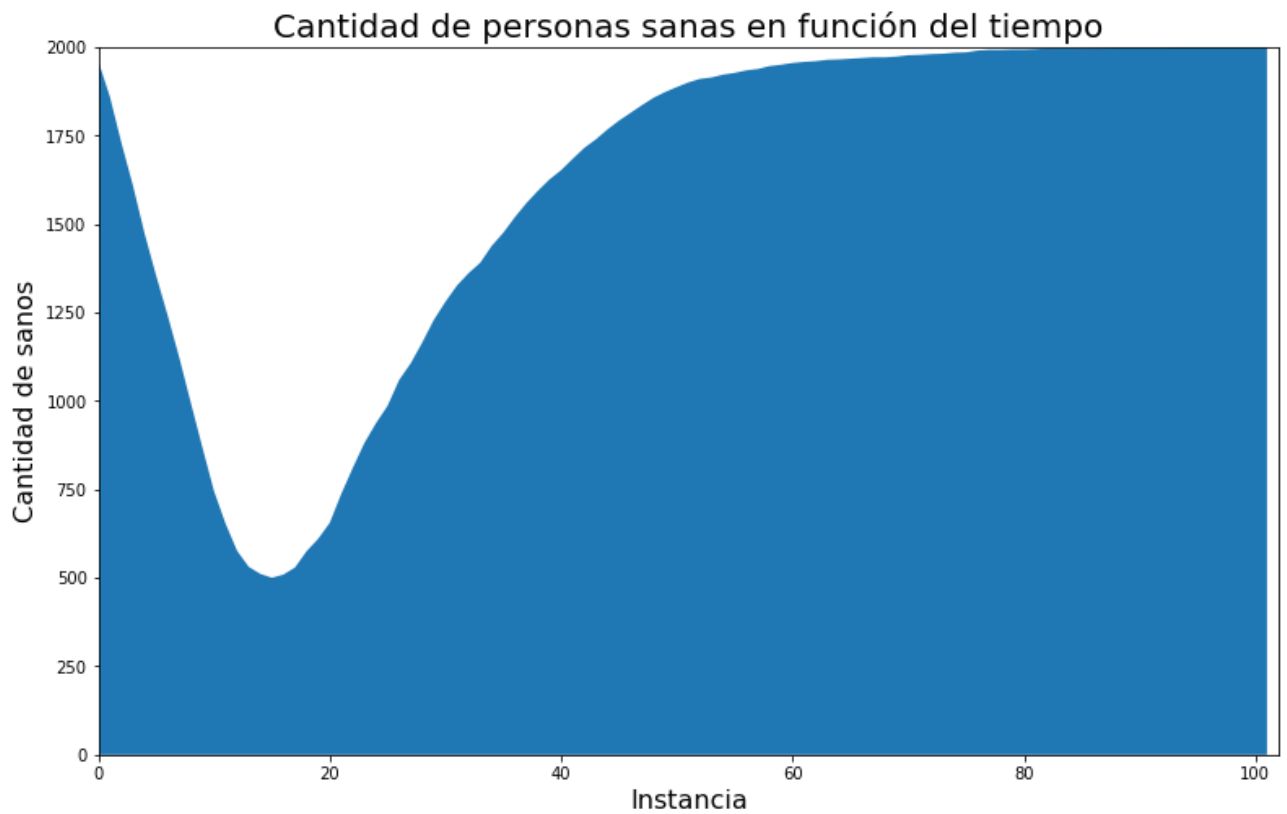


Fig. 51: sanos en cada instante para para $N = 2\,000$, $\alpha = 10$, $\beta = 0,1$ con inmunidad.

A continuación se presentan las estadísticas de los tiempos de contagio.

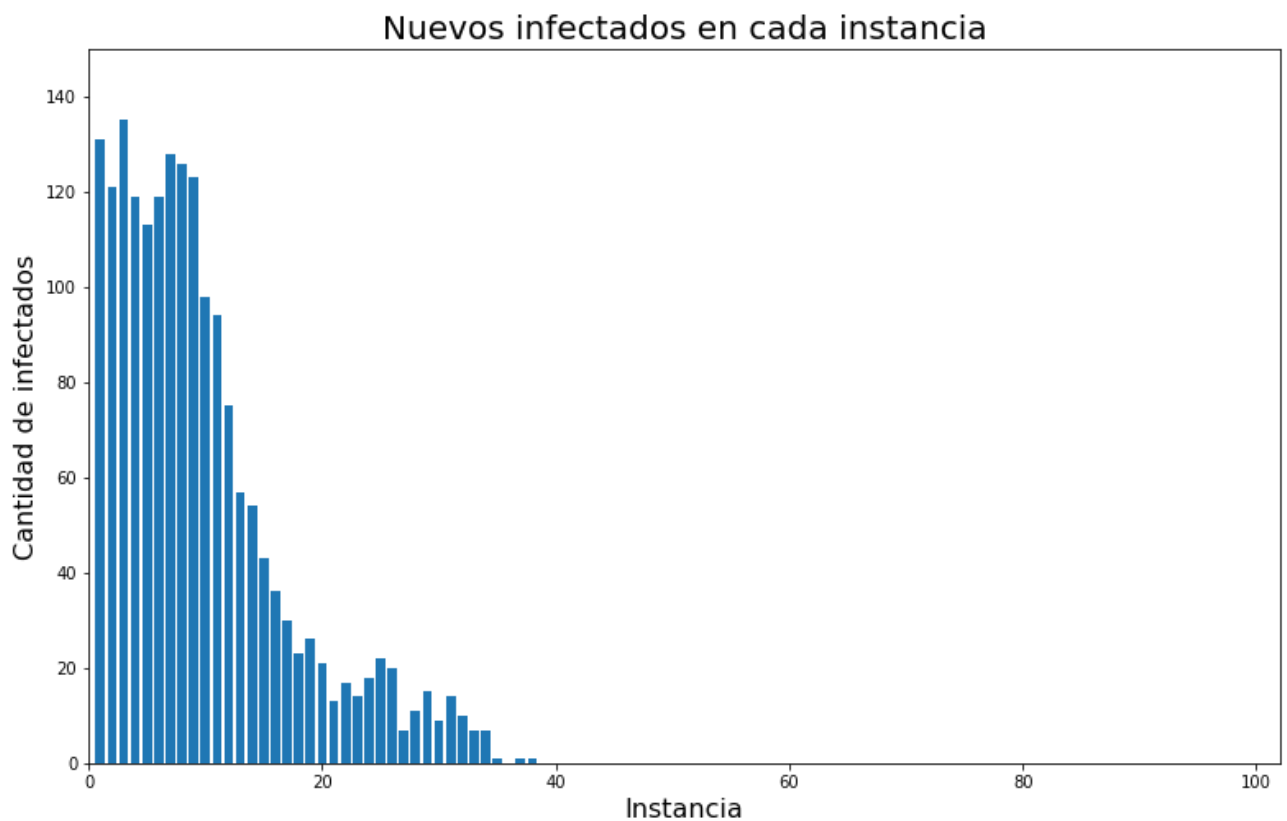


Fig. 52: nuevos infectados en cada instancia para $N = 2\,000$, $\alpha = 10$, $\beta = 0,1$ con inmunidad.

Correspondientemente con lo visto en otros ejercicios, la analogía entre las distribuciones asociadas a los tipos de agente particulares y la distribución de la totalidad de agentes se da en mayor medida si la cantidad de agentes es mayor. En todos los casos, se ve un reflejo de lo observado las curvas de infectados y sanos en cuanto al elevado contagio inicial.

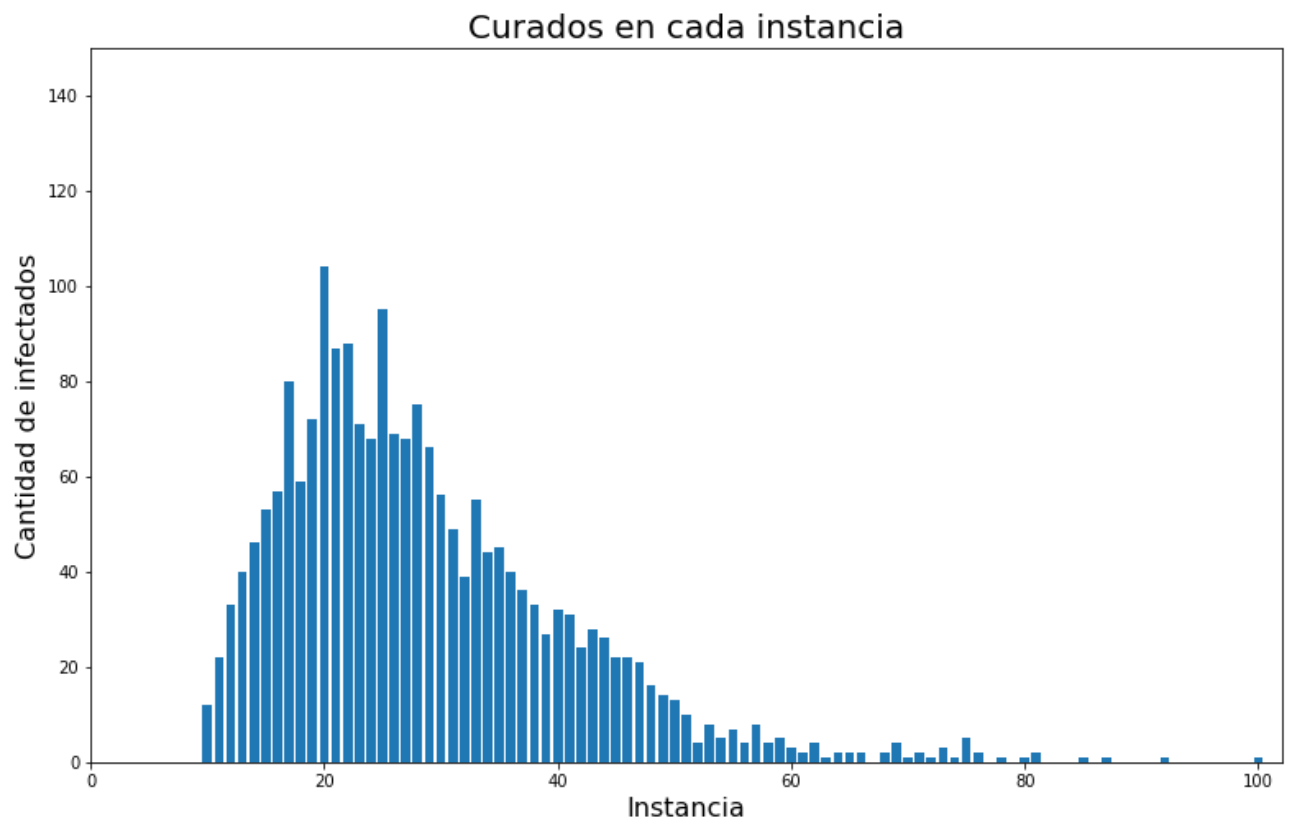


Fig. 53: nuevos recuperados en cada instancia para $N = 2\,000$, $\alpha = 10$, $\beta = 0,1$ con inmunidad.

Las curvas de sanados son análogas a las ya observadas, sin presentar tendencias particulares.

1 000 agentes, $\alpha = 20$, $\beta = 0,1$

La animación de ejemplo se encuentra en el siguiente enlace:

<https://drive.google.com/file/d/1X5bpr5aCrExiH3aynMBPzfAtgY86VxUc/view?usp=sharing>

Haber disminuido a la mitad la cantidad de agentes involucrados en la simulación parece haber reducido el pico de contagios. Esto es plausible: al haber en promedio un 50% de agentes vecinos, la proporción de contagios iniciales debe disminuir considerablemente. Fuera del valor absoluto, la mecánica de la epidemia parece ser la misma que en el caso anterior. Cabe destacar que el aumento en la cantidad de iteraciones totales también se corresponde con lo ocurrido en otros ejercicios al contrastar los casos con menos agentes. A continuación, se muestran las estadísticas correspondientes a dicho escenario.

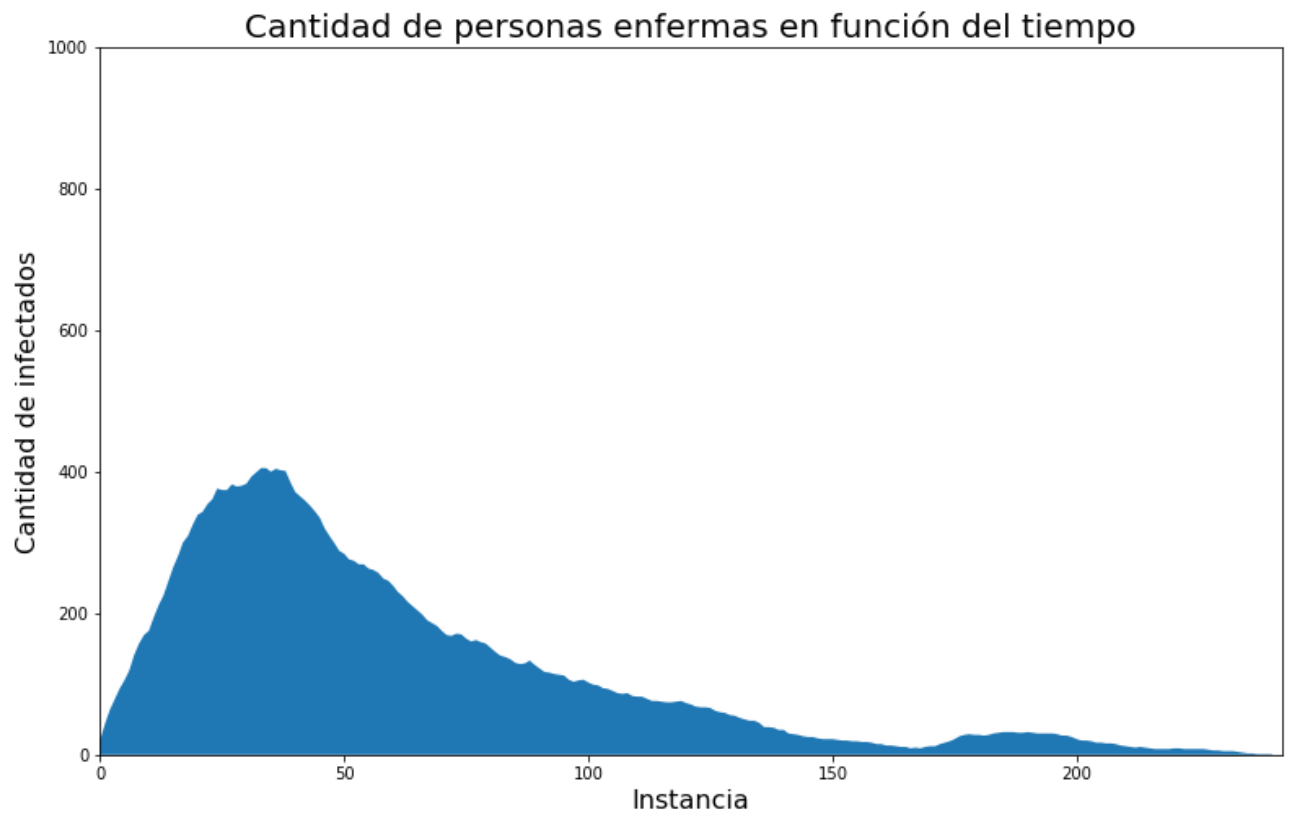


Fig. 54: enfermos en cada instante para $N = 1\,000$, $\alpha = 20$, $\beta = 0,1$ con inmunidad.

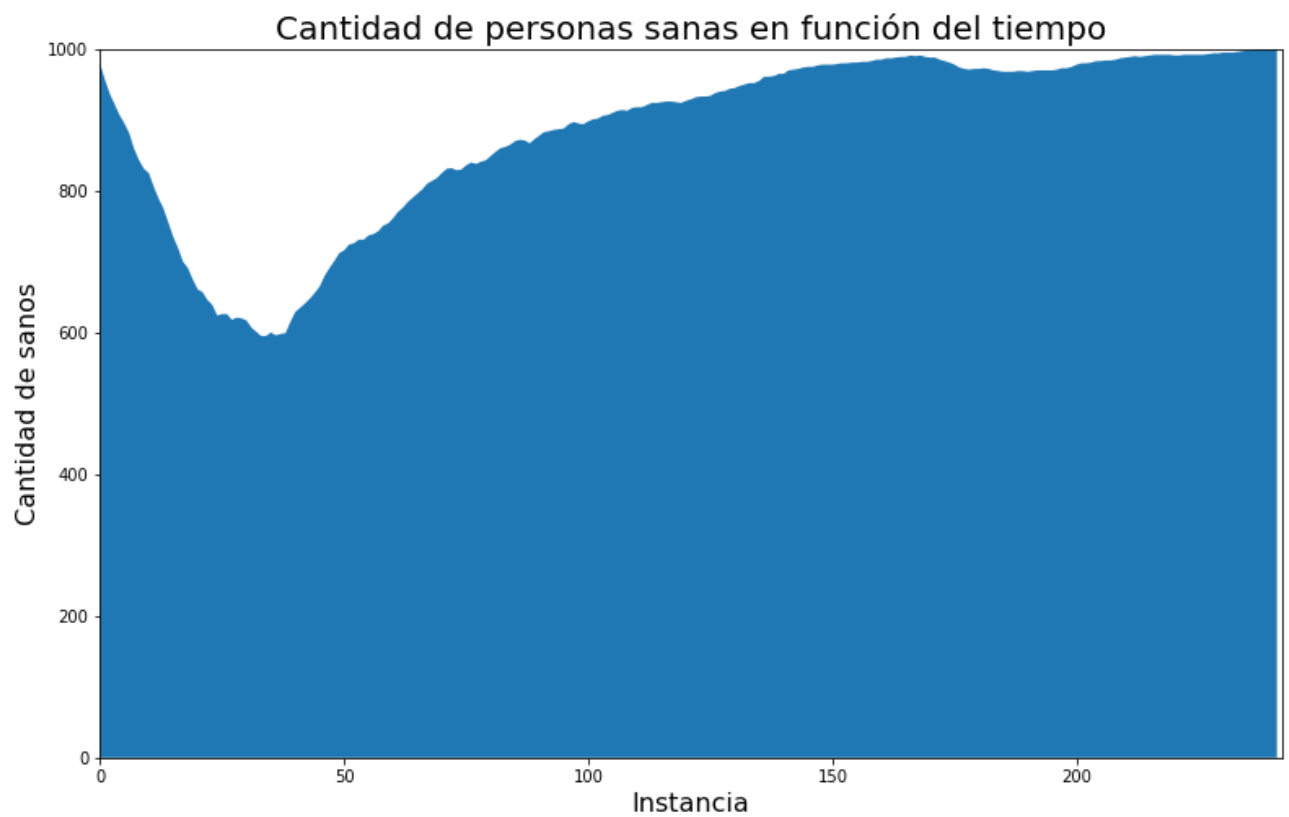


Fig. 55: sanos en cada instante para para $N = 1\,000$, $\alpha = 20$, $\beta = 0,1$ con inmunidad.

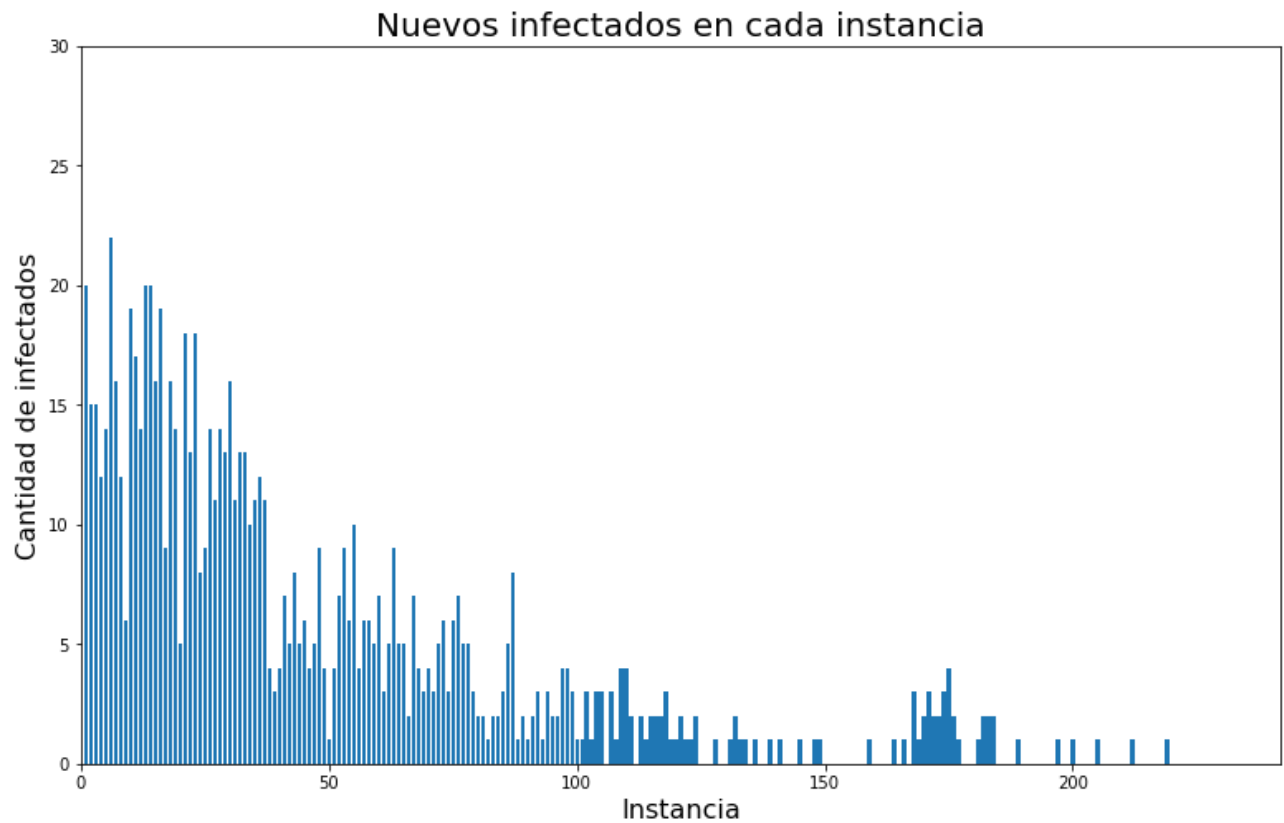


Fig. 56: nuevos infectados en cada instancia para $N = 1\,000$, $\alpha = 20$, $\beta = 0,1$ con inmunidad.

Los tiempos de contagio se condicen con las curvas de sanos y enfermos, guardando una analogía similar a lo visto en casos anteriores.

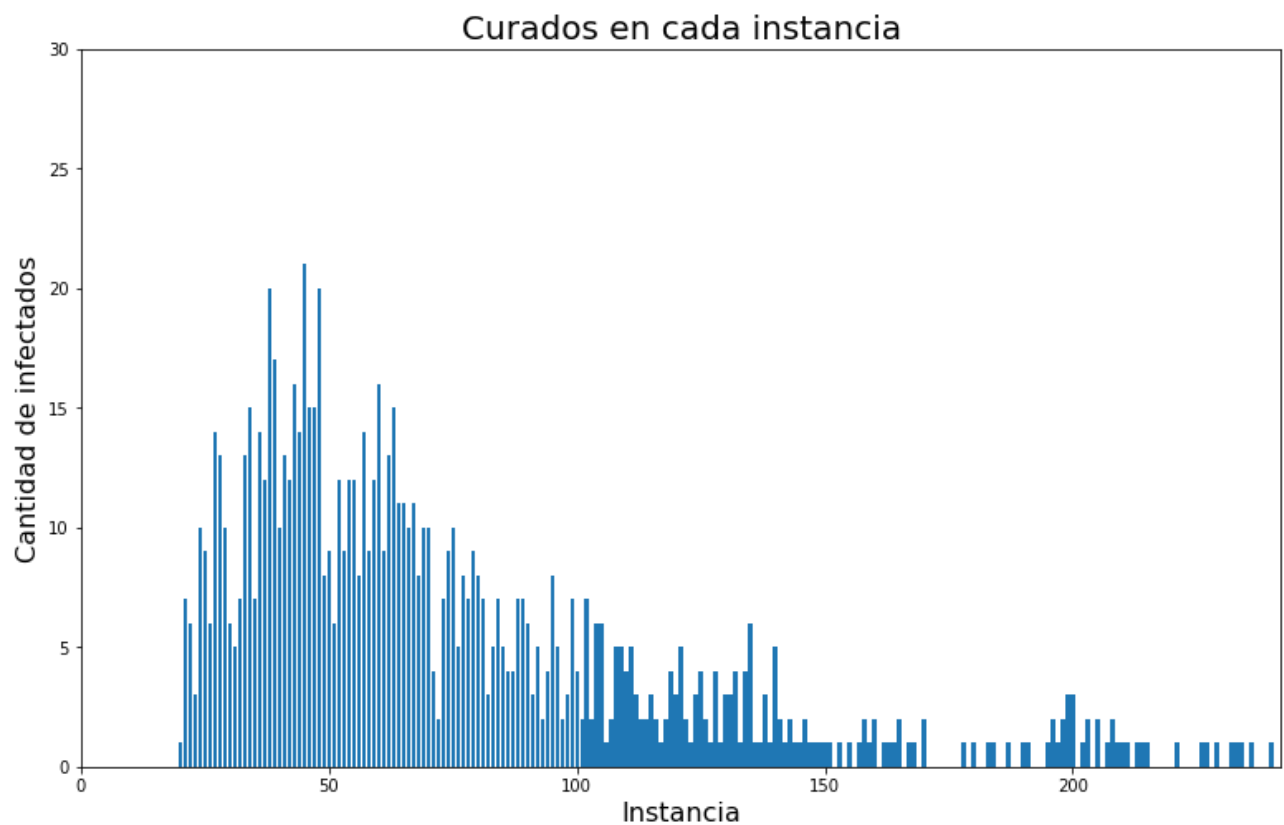


Fig. 57: nuevos recuperados en cada instancia para $N = 1\,000$, $\alpha = 20$, $\beta = 0,1$ con inmunidad.

1 000 agentes, $\alpha = 10$, $\beta = 0,2$

La animación de ejemplo se encuentra en el siguiente enlace:

https://drive.google.com/file/d/1artewTCpVIJbAjPb_0J1KthK4aHexYkH/view?usp=sharing

La combinación de disminuir a la mitad la cantidad de iteraciones hasta recuperar la salud y la duplicación de la probabilidad de recuperarse disminuyen aún más la curva de contagios, manteniendo el ya conocido patrón de concentrar la mayoría de los nuevos infectados en las iteraciones iniciales. A continuación, se muestran las estadísticas correspondientes a dicho escenario.

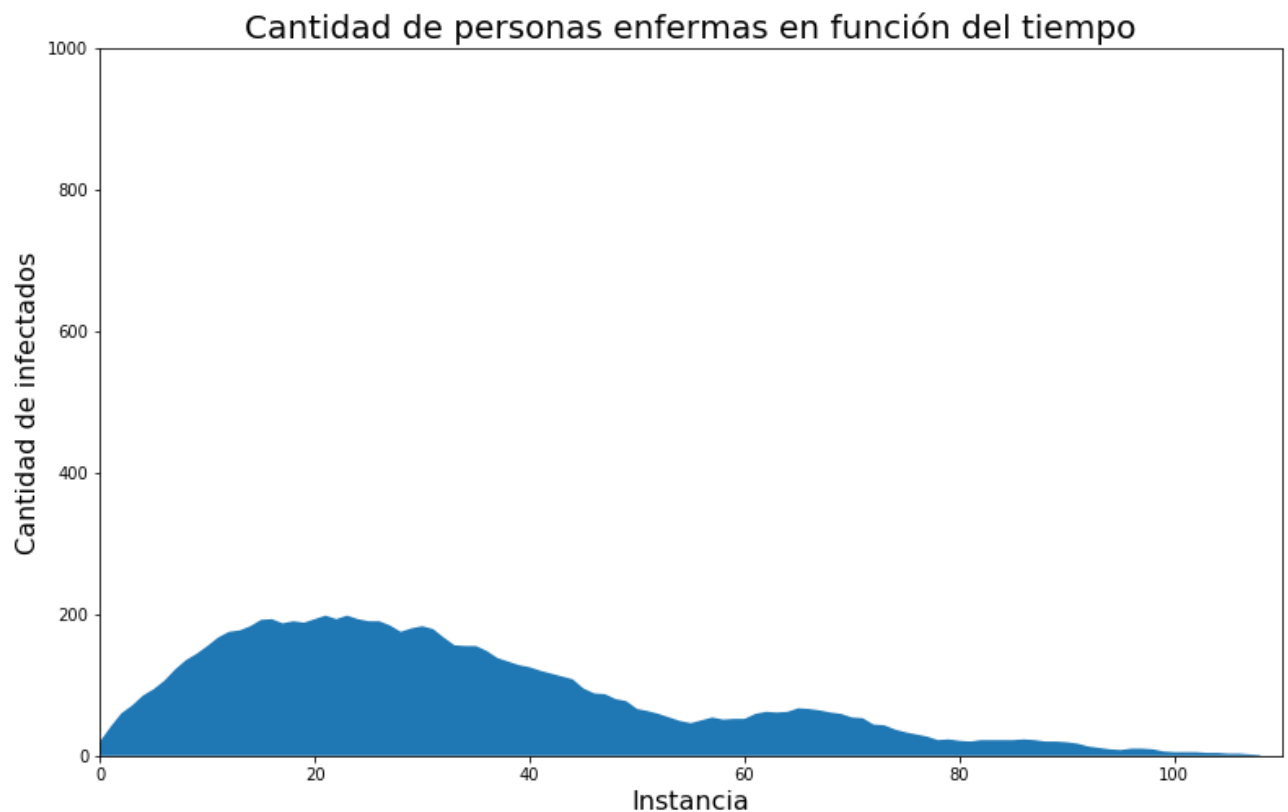


Fig. 58: enfermos en cada instante para $N = 1\,000$, $\alpha = 10$, $\beta = 0,2$ con inmunidad.

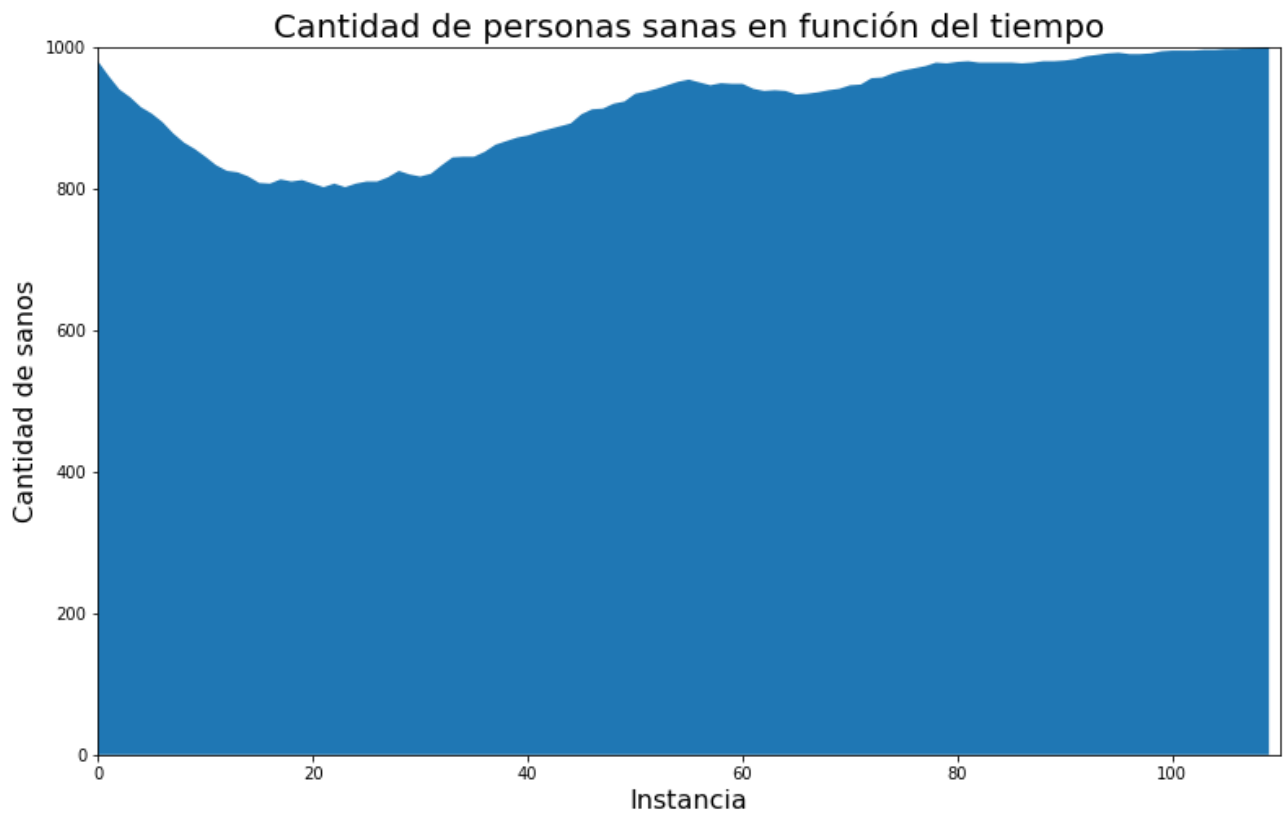


Fig. 59: sanos en cada instante para para $N = 1\,000$, $\alpha = 10$, $\beta = 0,2$ con inmunidad.

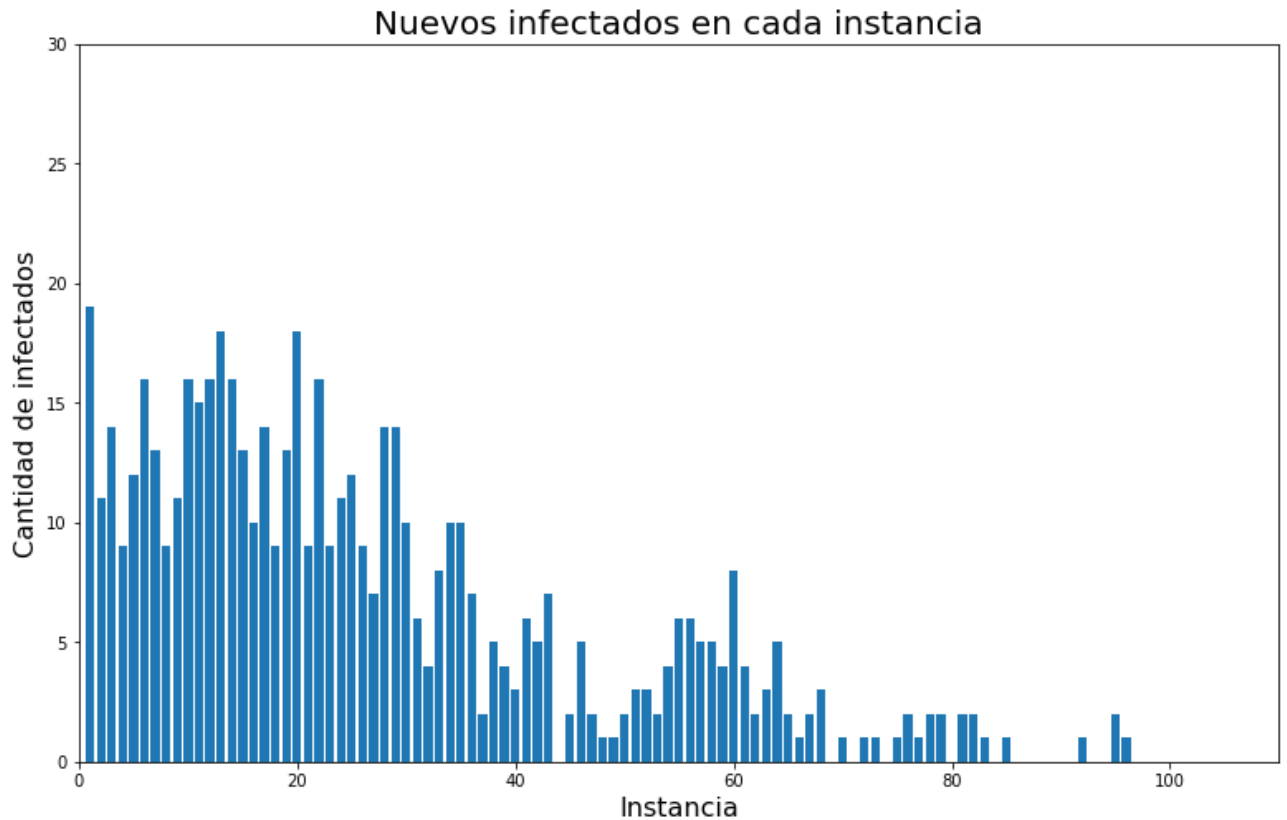


Fig. 60: nuevos infectados en cada instancia para $N = 1\,000$, $\alpha = 10$, $\beta = 0,2$ con inmunidad.

Dado como se repiten tendencias anteriormente vistas entre la correspondencia de los agentes individuales con la distribución total, y la concentración inicial de contagios, no hay detalles particulares que remarcar en este caso.

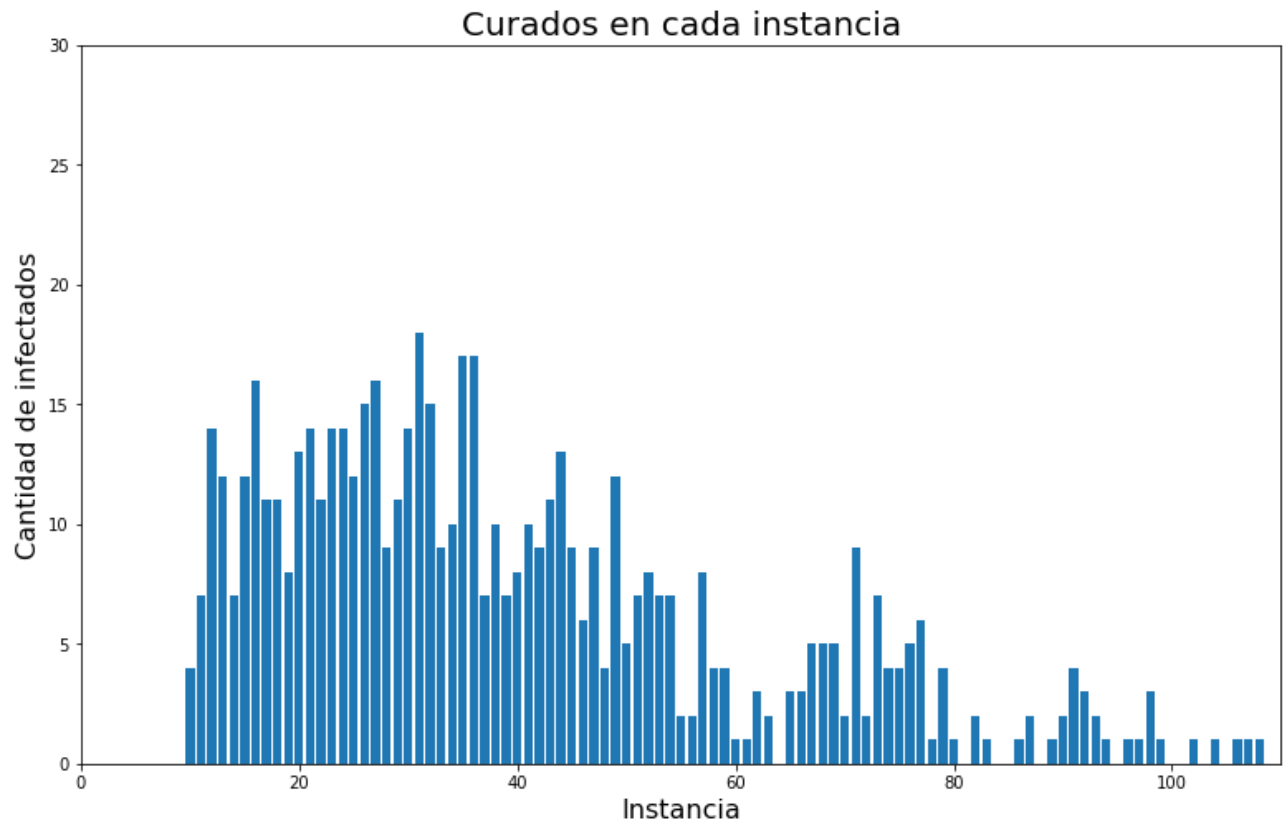


Fig. 61: nuevos recuperados en cada instancia para $N = 1\,000$, $\alpha = 10$, $\beta = 0,2$ con inmunidad.

Pasadas las 10 iteraciones iniciales, se la mayoría de recuperados iniciales se debe a la concentración de infectados, la cual baja correspondientemente luego de la mitad de las iteraciones que el sistema realiza en su evolución hasta llegar

Conclusión

Se puede observar que la capacidad de recuperación de la enfermedad sin haber inmunidad para una densidad de población representativa (como la dada por 1 000 agentes) no conlleva al sanado de la totalidad de los agentes, sino que se establece rápidamente un equilibrio dinámico entre infectados y recuperados recientes. En el mismo, siempre hay agentes curados que en breve se vuelven a infectar. Sin embargo, incrementar la probabilidad de recuperación tendiendo a un valor cercano a 1 causa que la cantidad de infectados en régimen sea menor.

La recuperación de la enfermedad con inmunidad causa que la simulación termine con todos sanos, en vez de todos enfermos. Aumentar la densidad de agentes causa que el pico de infectados ocurra más rápidamente, y que el mismo sea más acentuado.

Ejercicio 10

Esquema de solución

El escenario con el que se realizan las simulaciones en este caso es a partir de 1 000 agentes, $\alpha = 20$, $\beta = 0,1$ con Correlated Random Walking e inmunidad luego del contagio. Para simular los agentes "autorizados" a moverse, simplemente se eligen uniformemente al azar y sin repetición una cantidad de agentes asociada al λ escogido (o bien, porcentaje de movilidad) según sus identificadores, de forma

análoga a como se eligen los infectados iniciales. Entonces, durante el bucle del movimiento simplemente se verifica si el agente en cuestión está dentro del diccionario de agentes a los que se les permite moverse.

Con respecto a los cambios de probabilidades mencionados en la consigna, esto solo se traduce en un cambio de variables. No hay otra aclaración para hacer en términos del algoritmo de simulación ya que es el mismo, dadas las condiciones elegidas que se describieron anteriormente.

Visualizaciones

75% de movilidad

La animación de ejemplo se encuentra en el siguiente enlace:

https://drive.google.com/file/d/1FAdnm_Hg_-HQpZeyiGZmgrVNI5cFH7uD/view?usp=sharing

A continuación, se muestran las estadísticas correspondientes a dicho escenario.

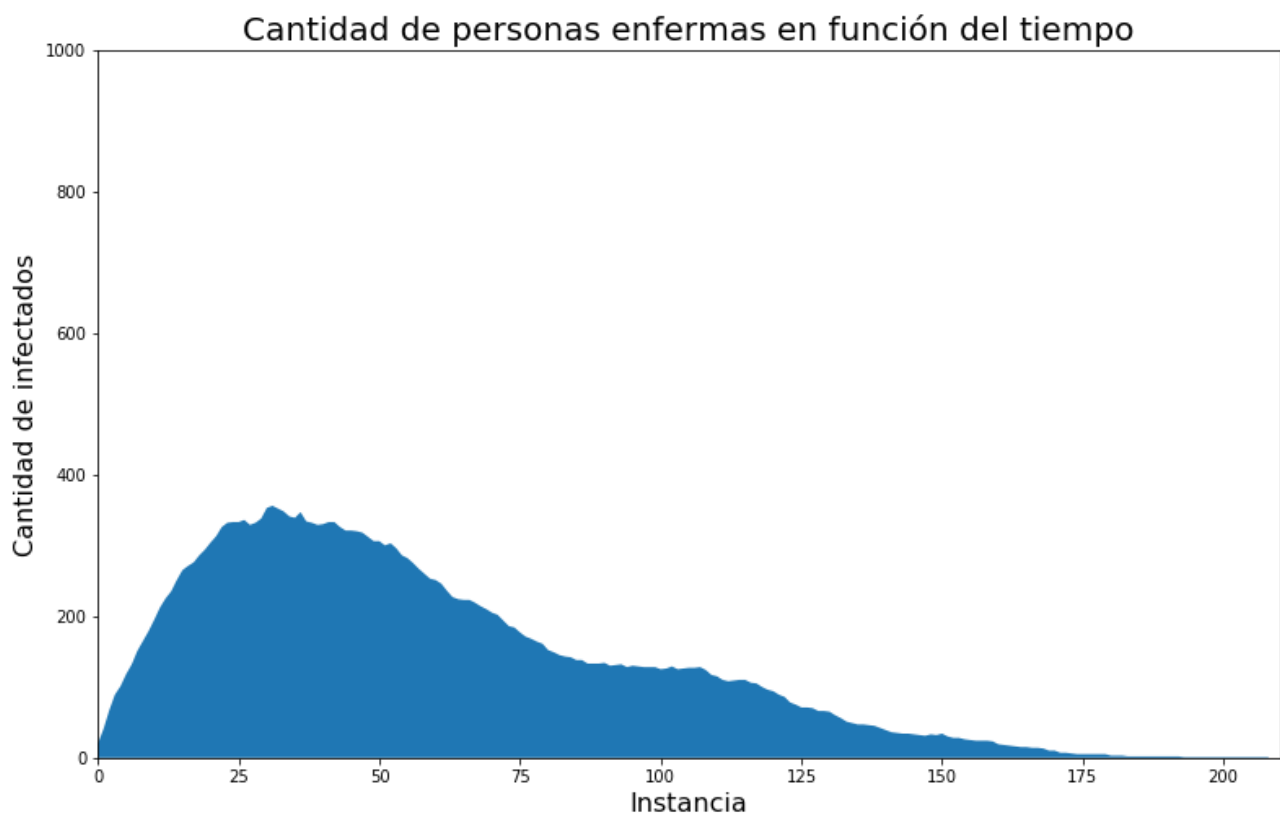


Fig. 62: enfermos en cada instante con 75% de movilidad.

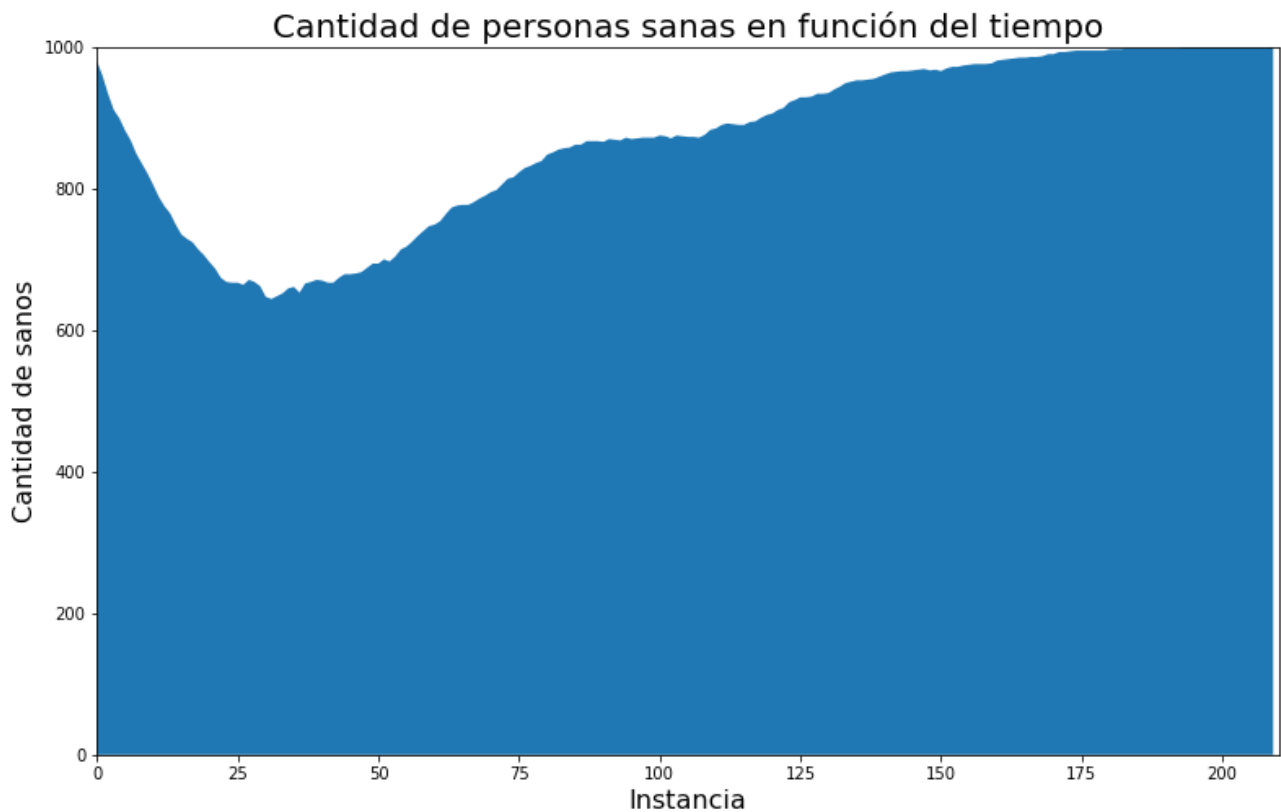


Fig. 63: sanos en cada instante con 75% de movilidad

Ambas curvas se ven similares a sus casos análogos del ejercicio 9. Por otra parte, la cantidad de iteraciones en las que el sistema alcanza la totalidad de agentes sanos es similar.

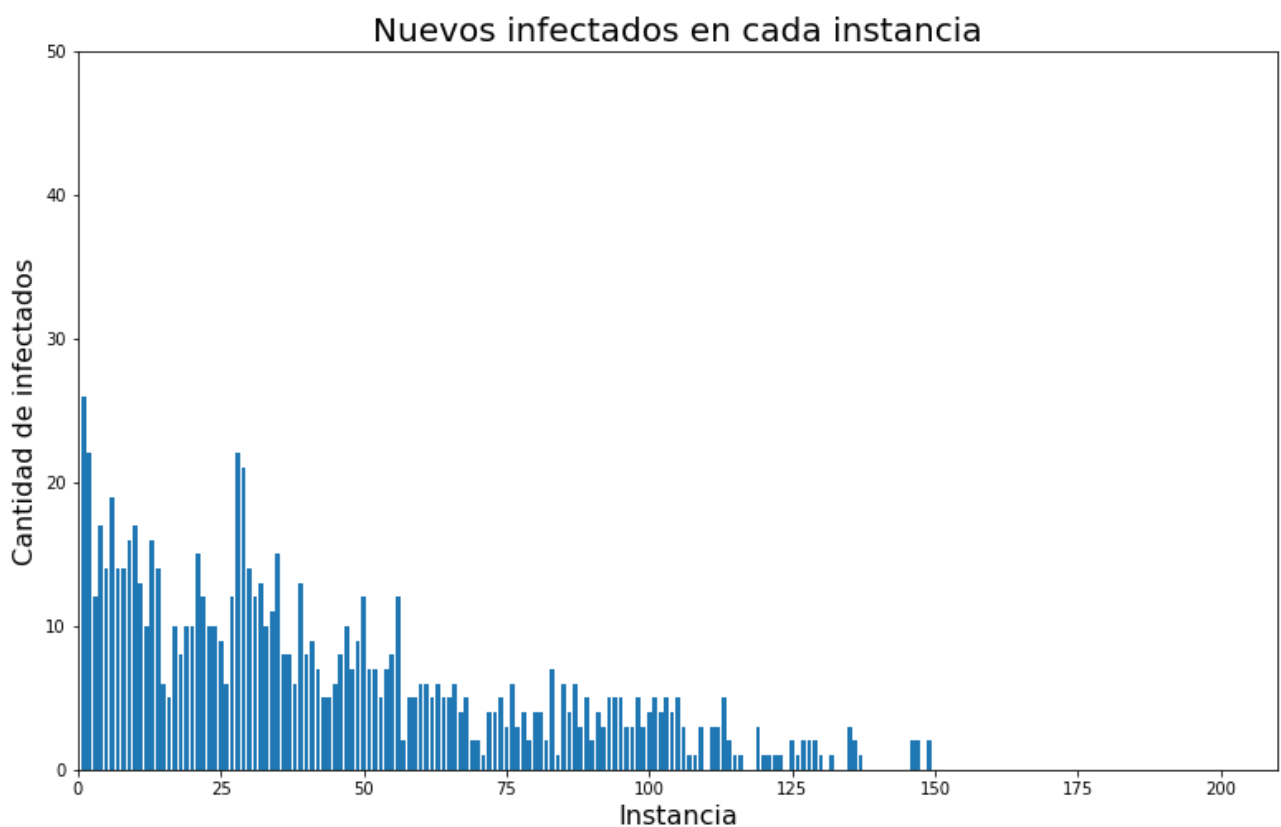


Fig. 64: nuevos infectados en cada instante con 75% de movilidad.

En cuanto a los tiempos de contagios, la diferencia que se ve con el ejercicio 9 es que la cantidad de enfermos iniciales es bastante menor. Esto se atribuye al hecho de que el menor movimiento de los agentes causa una menor tasa de infección entre los vecinos cercanos a los agentes enfermos. Fuera de eso, las distribuciones se ven similares con respecto al caso análogo del ejercicio anterior.

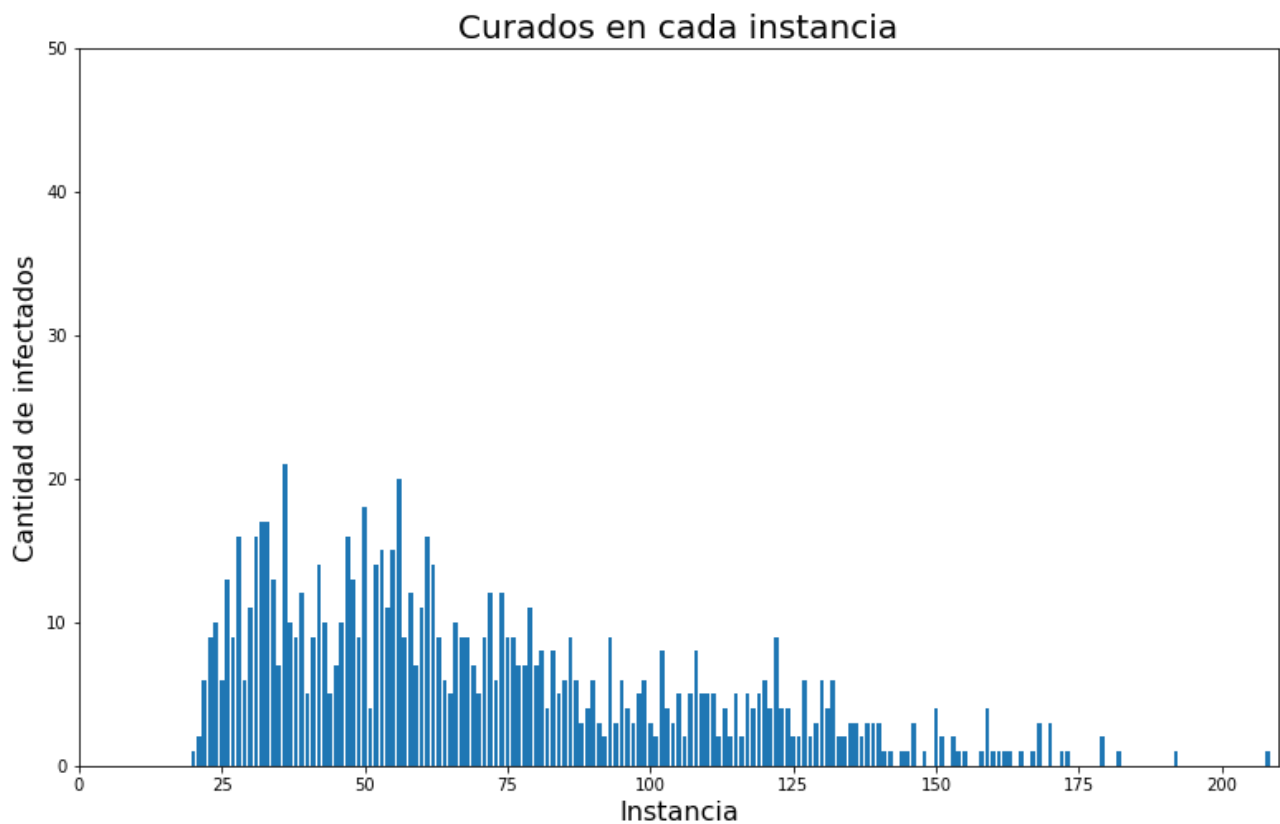


Fig. 65: nuevos agentes curados en cada instante con 75% de movilidad.

Correspondientemente, la cantidad de agentes curados inicialmente es menor al caso análogo del ejercicio 9. Más allá de eso, tanto las distribuciones en sí como la correspondencia entre los gráficos de los agentes individuales con la visualización de tiempos de contagios general se mantiene similar al ejercicio anterior.

50% de movilidad

La animación de ejemplo se encuentra en el siguiente enlace:

<https://drive.google.com/file/d/1wAFkF-hAqdt6E2R3MvEchdseSdJ68fa2/view?usp=sharing>

A continuación, se muestran las estadísticas correspondientes a dicho escenario.

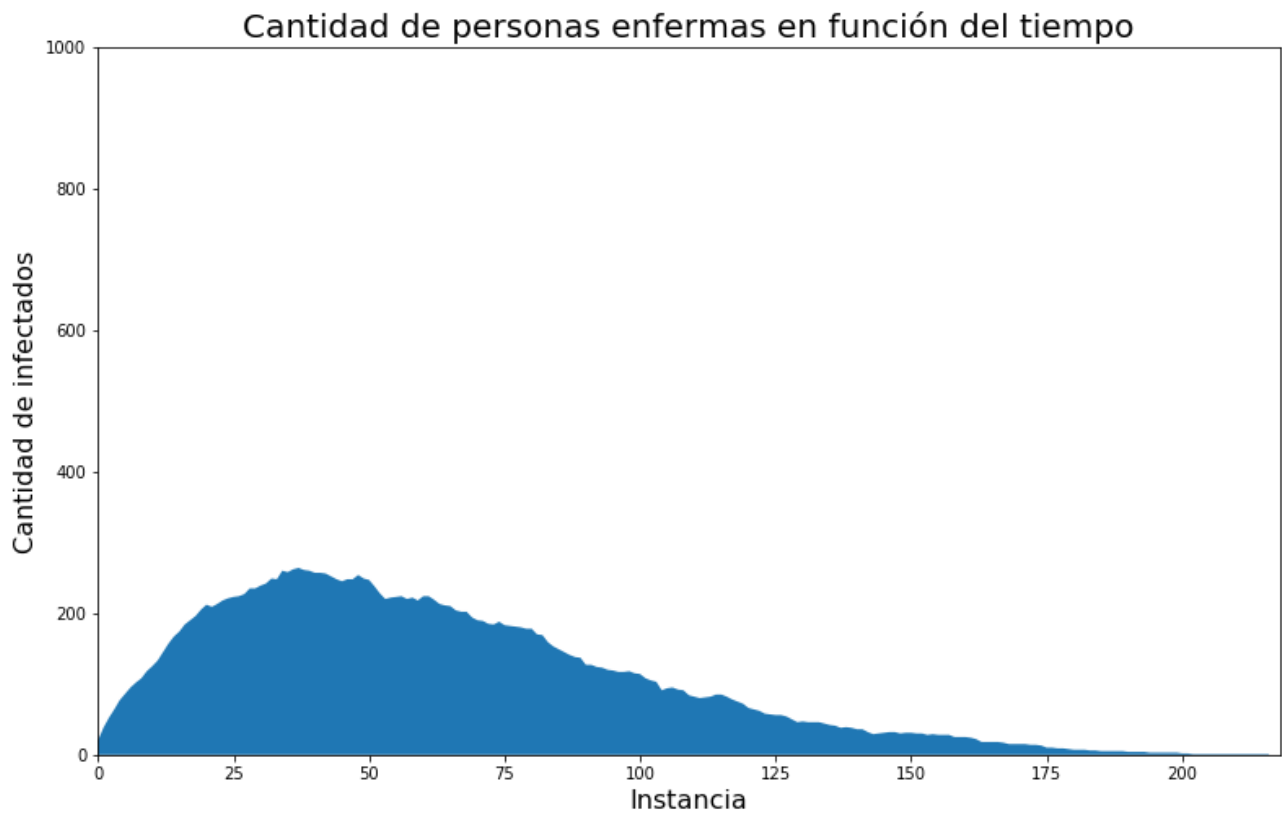


Fig. 66: enfermos en cada instante con 50% de movilidad.

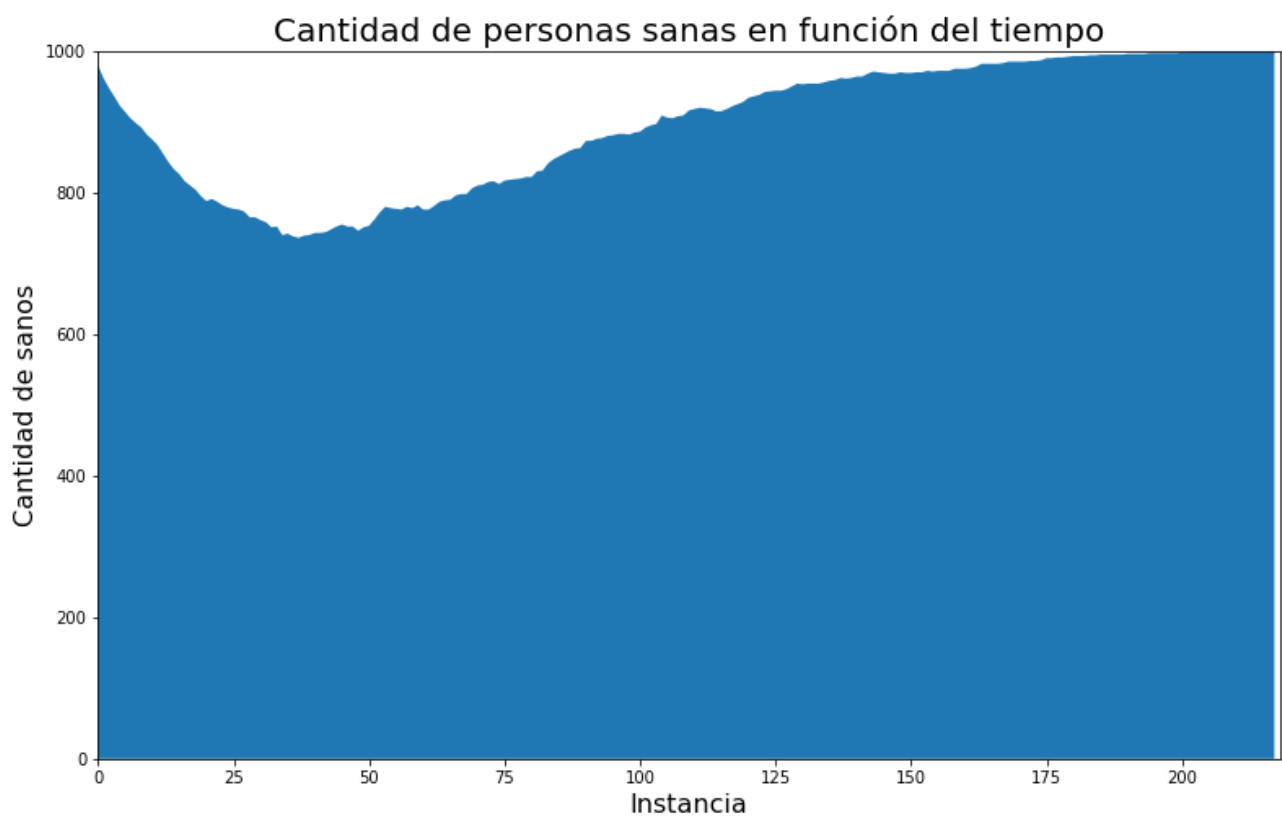


Fig. 67: sanos en cada instante con 50% de movilidad.

Nuevamente, no se ven diferencias sustanciales con el caso análogo al ejercicio 9 en cuanto a las distribuciones en sí más allá de la disminución en la magnitud del pico de contagios. Una hipótesis es que la inmunidad, sumada a la rápida posibilidad de recuperación dada por α , restringen las posibilidades de modificar la evolución del sistema si solamente se establece la reducción del movimiento. La cuestión es que la mayoría de los contagios siempre se dan al inicio para las cantidad de agentes elegidas, y como ya se analizó anteriormente, se ve que el hecho de que haya inmunidad hace que el sistema supere la epidemia con relativa rapidez. En ese contexto, la limitación del movimiento solo disminuye la cantidad de contagios iniciales, pero la evolución del sistema es similar dado que el tipo de movimiento es el mismo (Correlated Random Walking), así como la proporción en que los agentes se recuperan. Una observación interesante, entonces, es que para suficientes agentes distribuidos uniformemente en un espacio "cerrado", limitar el movimiento no es suficiente para cambiar la forma en que la epidemia se desarrolla, pero permite que la cantidad de gente infectada sea menor, lo cual en la práctica podría salvar vidas.

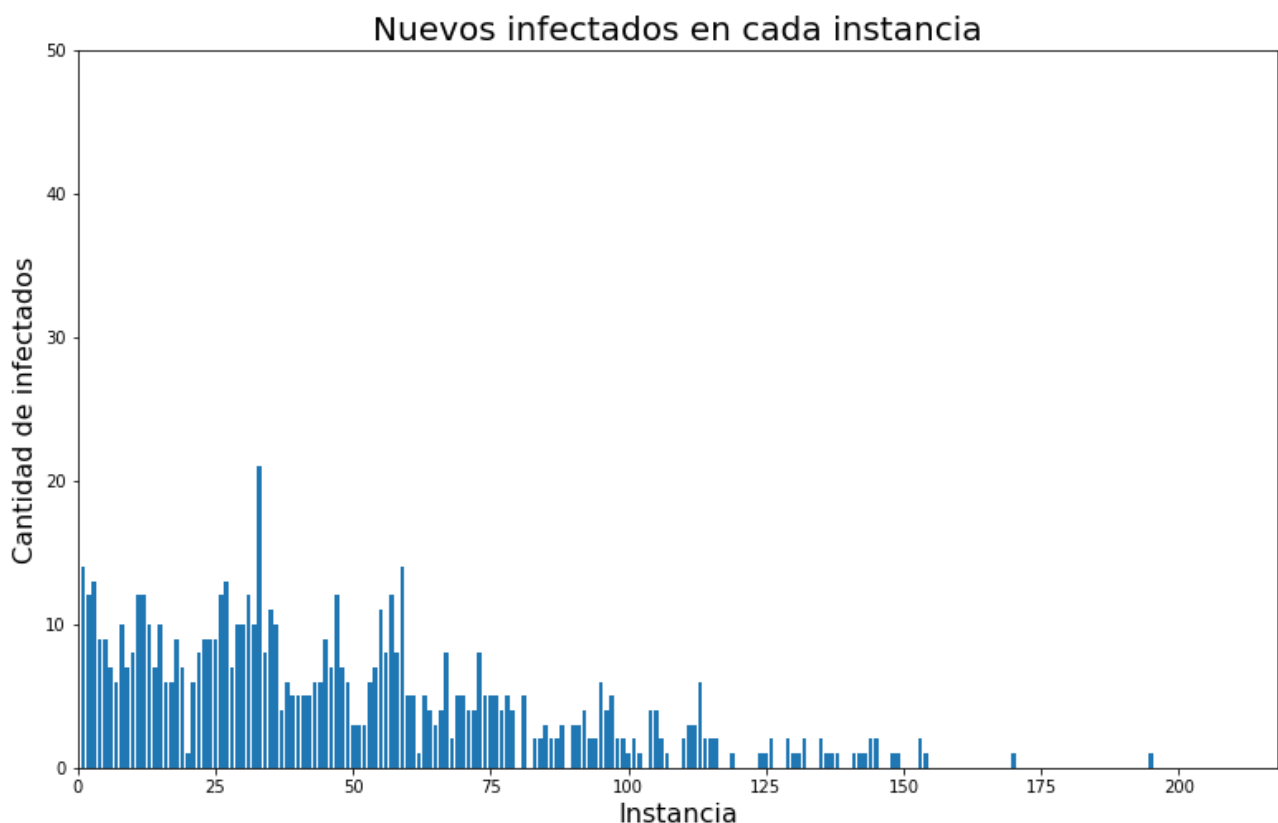


Fig. 68: nuevos infectados en cada instante con 50% de movilidad.

Asociado a lo mencionado anteriormente, se ve la menor proporción de agentes infectados iniciales, que a su vez se traduce en una menor cantidad de infectados totales. En líneas generales esa es la única diferencia que puede entenderse entre este caso y el ejercicio anterior.

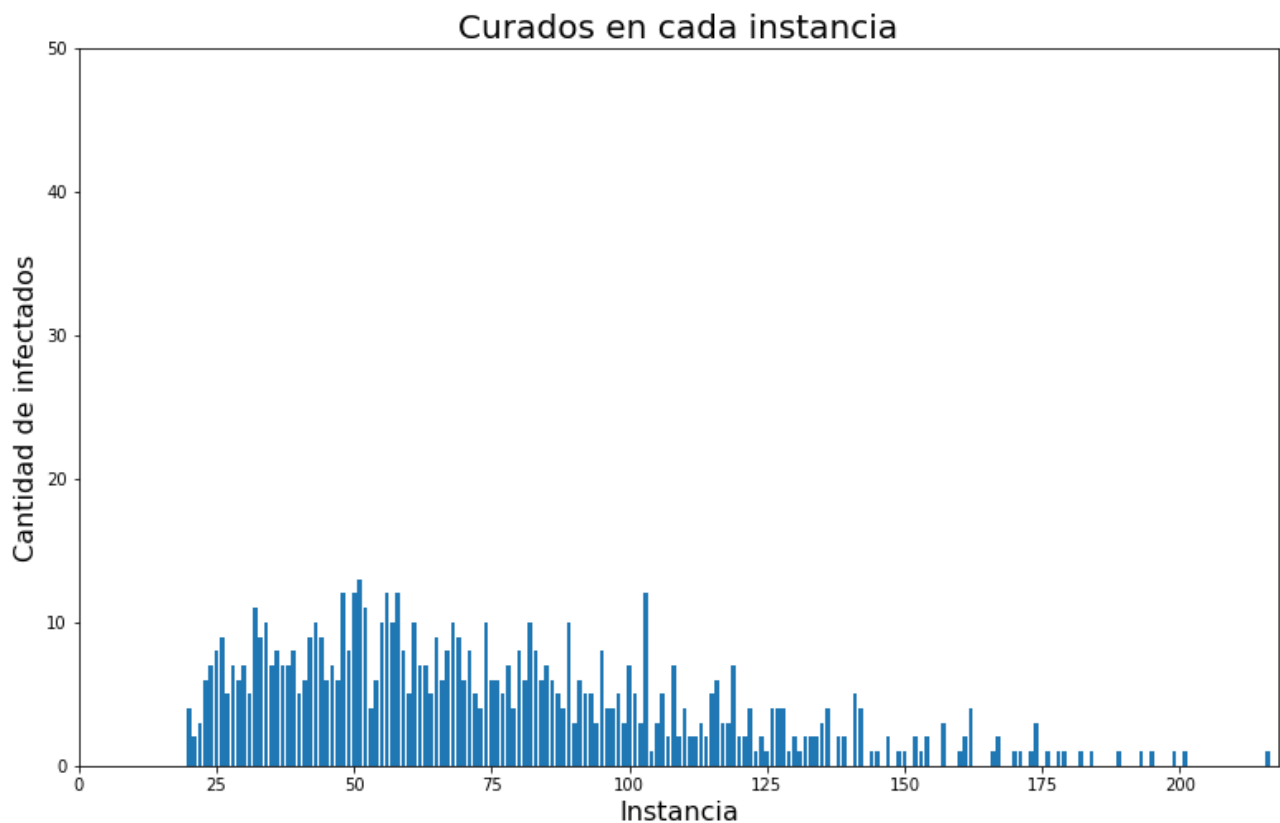


Fig. 69: agentes curados en cada instante con 50% de movilidad.

Análogamente al caso anterior, la diferencia con el ejercicio 9 es la menor proporción de curados, asociada a una menor cantidad de infectados.

25% de movilidad

La animación de ejemplo se encuentra en el siguiente enlace:

<https://drive.google.com/file/d/18T5r5d-NqFyNtLNEqPQ9rToORHBkah6v/view?usp=sharing>

En este caso, la restricción en el movimiento de los agentes no solo disminuye el pico de contagios con respecto al ejercicio 9, sino que también logra que la epidemia termine en menos iteraciones, lo cual es un factor adicional para disminuir la cantidad total afectada por la enfermedad. A continuación, se muestran las estadísticas correspondientes a dicho escenario.

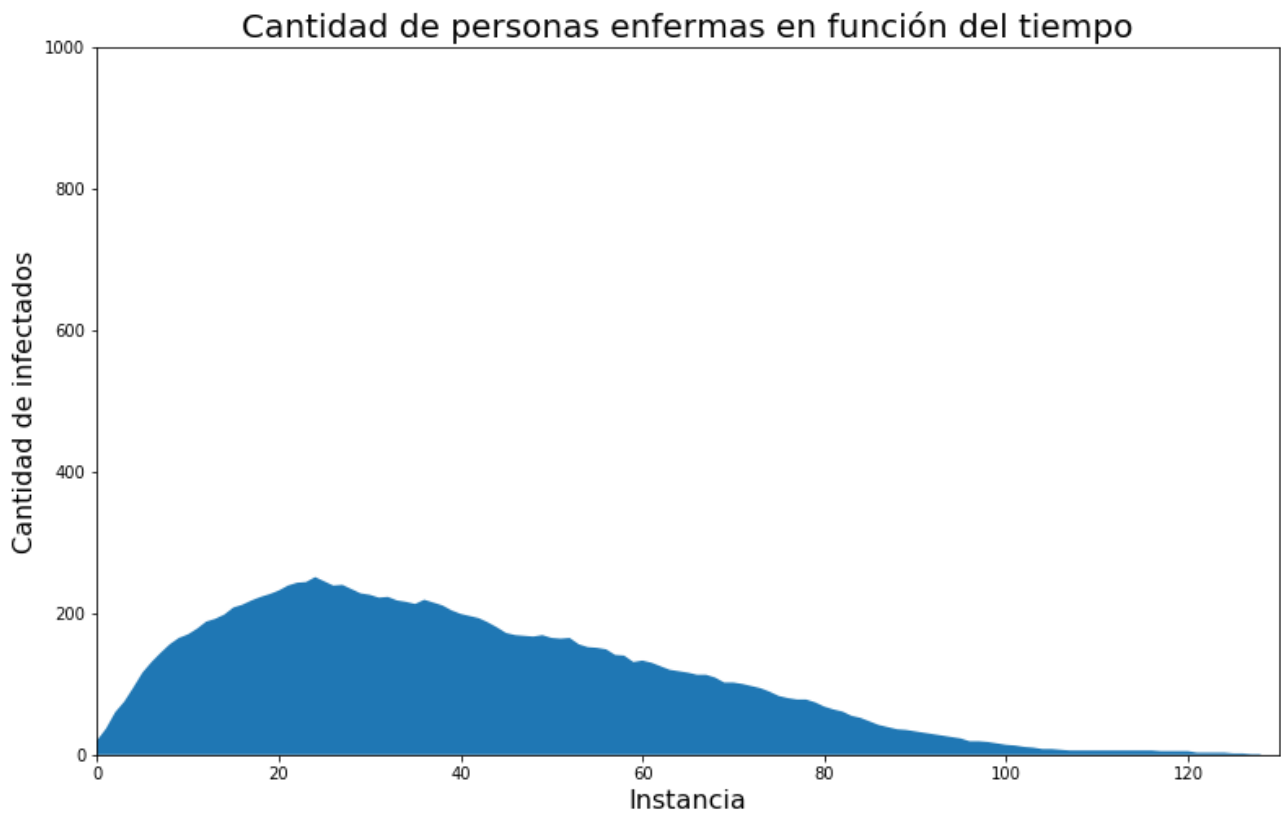


Fig. 70: enfermos en cada instante con 25% de movilidad.

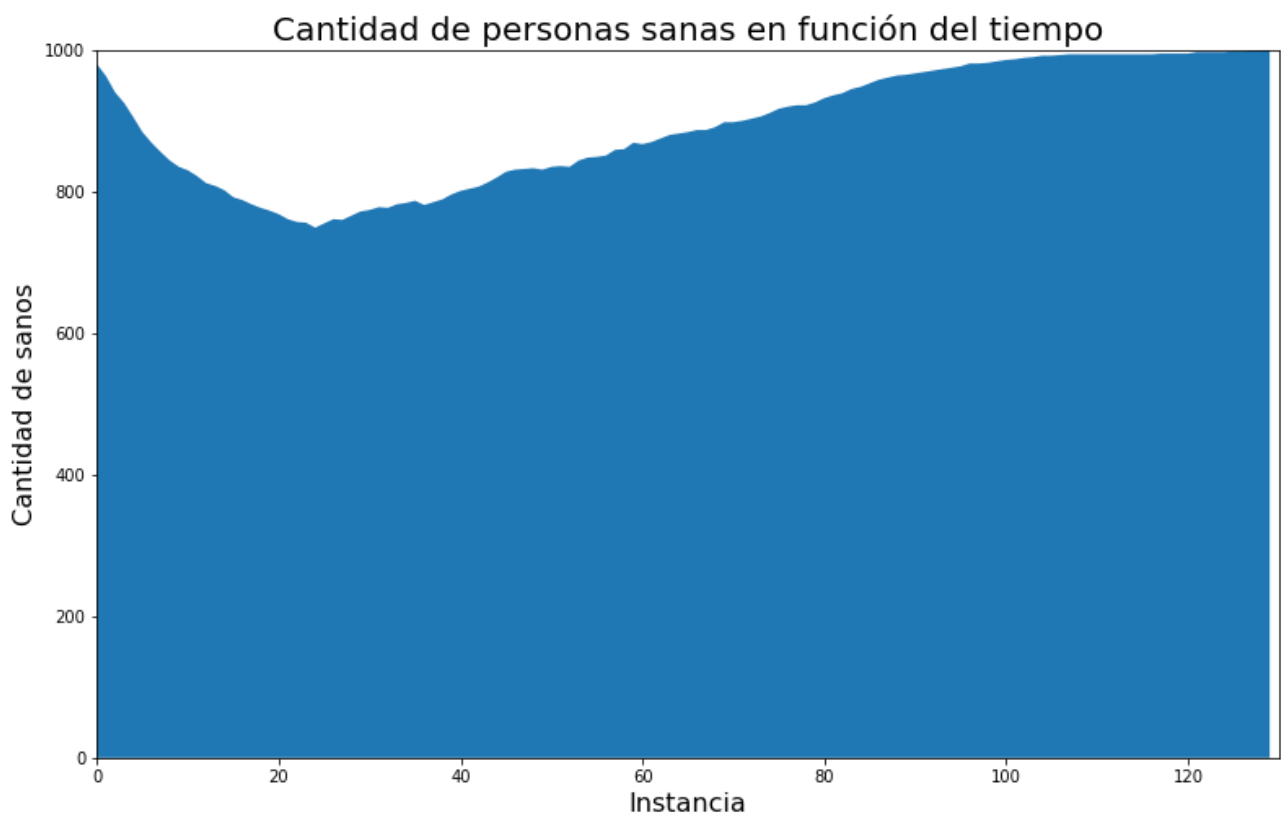


Fig. 71: sanos en cada instante con 25% de movilidad.

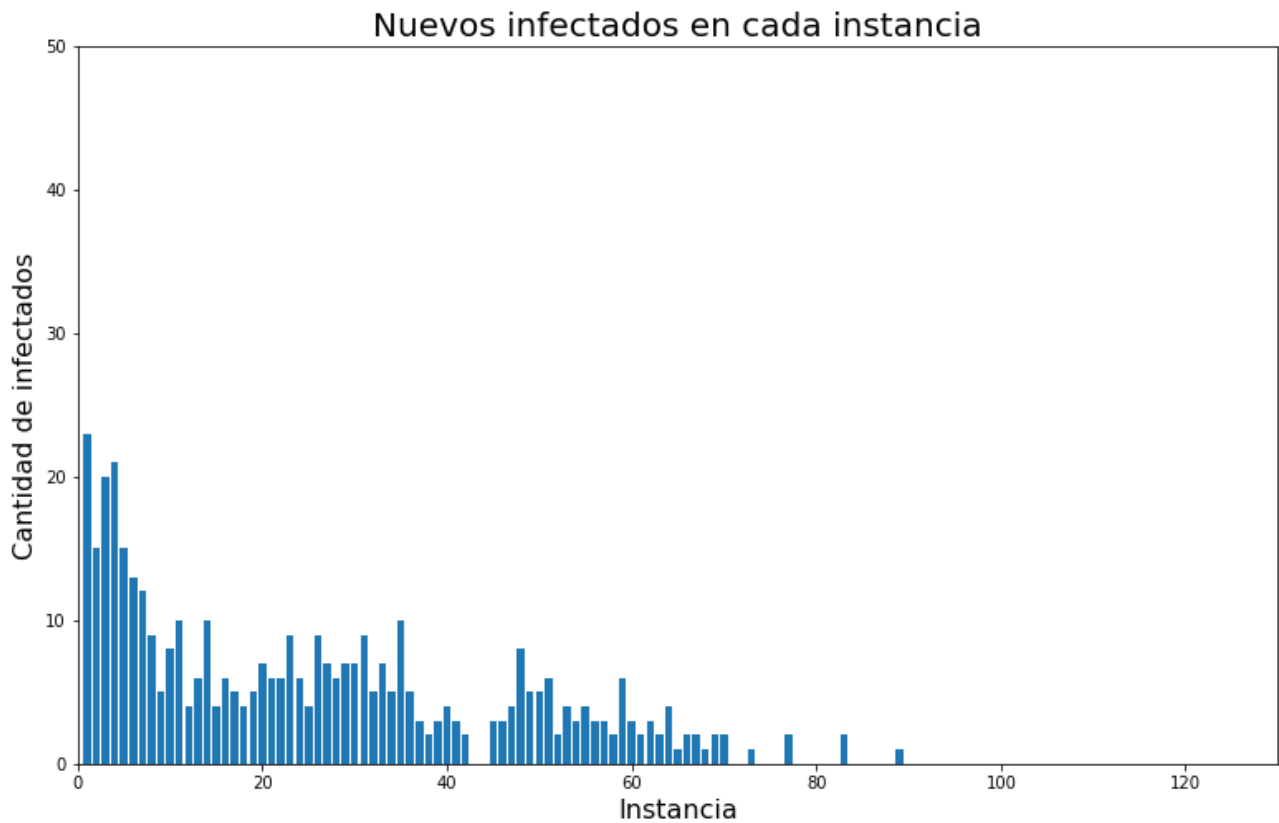


Fig. 72: nuevos agentes enfermos en cada instante con 25% de movilidad.

Análogamente a los otros casos de este ejercicio, la correspondencia de los tiempos de infección y la curva de contagios es clara. Se mantiene el patrón de la mayor concentración de casos al inicio igual que en el resto de las simulaciones.

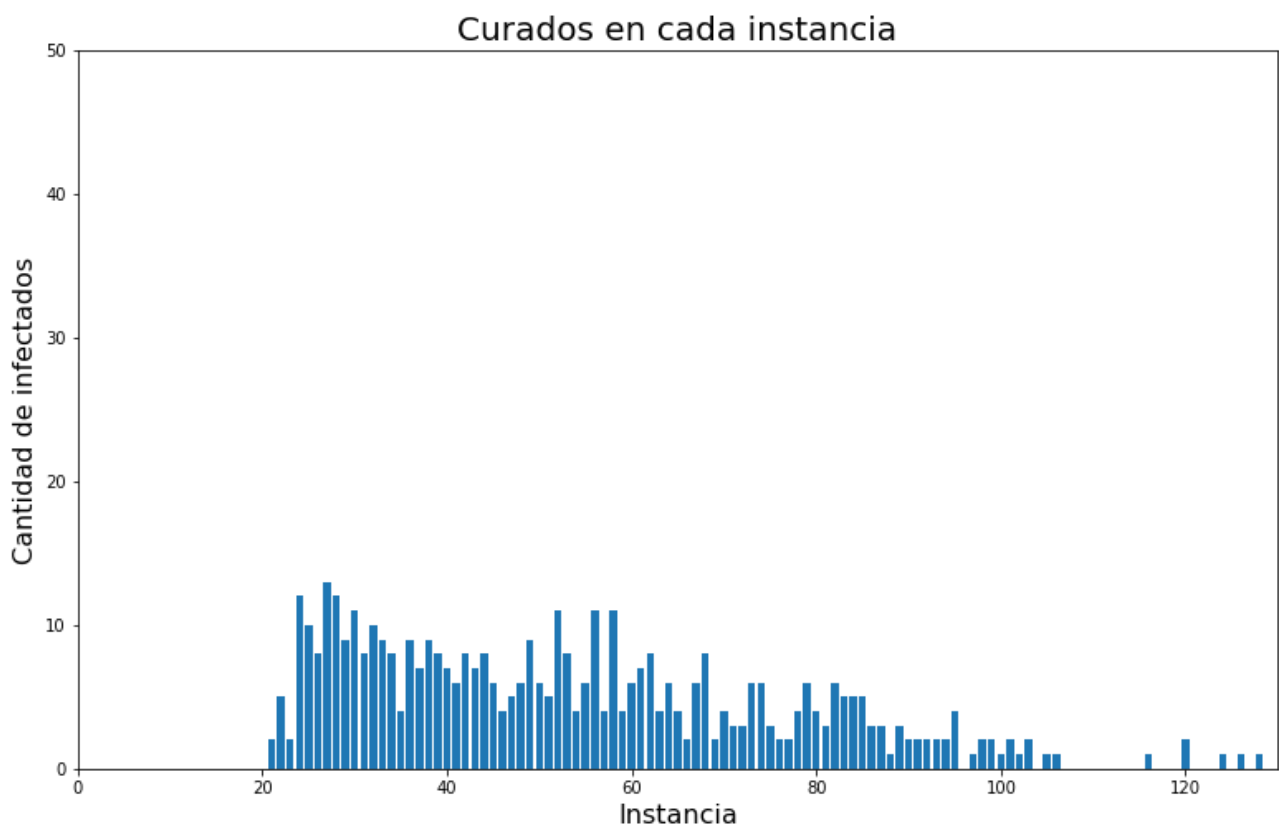


Fig. 73: nuevos agentes curados en cada instante con 25% de movilidad.

Conclusión

El inmovilizar a un porcentaje de los agentes parece haber causado que el pico de infecciones sea más leve y bajo. Adicionalmente, se puede observar que para el caso del 25% de movilidad se restringe lo suficiente el contagio entre agentes como para que el tiempo que dura la simulación se reduzca a casi la mitad. Ambas cuestiones llevadas a la práctica contribuirían a que la cantidad de personas que se enferma sea menor con respecto al caso de libre movilidad, lo cual además de proteger vidas podría simplificar el manejo de la epidemia en la sociedad donde ocurra, que en última instancia contribuye al ahorro de recursos.

Conclusión general sobre las simulaciones de infecciones

Las simulaciones realizadas han sido acumulativas con respecto a la cantidad de factores tenidos en cuenta para analizar la evolución del sistema de agentes en una situación epidémica. Dentro de lo observado, se puede considerar que para entender mejor el desenvolvimiento de un sistema real en este tipo de situaciones siempre hay más variables a considerar, pero muchas de ellas pueden ser emuladas a partir de parámetros representativos, al igual que la mecánica de la simulación es solo una aproximación de las interacciones que ocurren en la práctica. La misma, sin embargo, puede dar elementos de predictibilidad, como el hecho de que la inmunidad garantiza el fin de la epidemia de forma mucho más veloz que la infección generalizada en caso de no haberla, o que la restricción del movimiento de agentes permite totalizar menos enfermos, y una recuperación total de la población más rápida.

Por otra parte, la simulación permite hacer pequeñas reproducciones de comportamientos. Se puede ver, por ejemplo, que al aumentar la densidad de los agentes el contagio se realiza de forma vertiginosa dada la interacción de los individuos con sus vecinos cercanos. Esto es representativo de lo que ocurre en la práctica: si hay más personas en contacto entre sí, y no existen grupos aislados (como en una ciudad), entonces la propagación se acelera.

Anexo: código fuente

Se adjuntan los notebooks utilizados para lograr la solución descrita.