

UNIVERSIDAD DE COSTA RICA

“Análisis y simulación espacial de la pandemia COVID-19 a nivel cantonal, para el caso de Costa Rica, CIOdD-UCR”

Elaborado por: MSI. Agustín Gómez Meléndez

Coordinador Unidad de Estadísticas

Centro de Investigación Observatorio del Desarrollo (CIOdD)

El presente informe es un aporte más que el Centro de Investigación Observatorio del Desarrollo (CIOdD) de la Universidad de Costa Rica realiza de manera sistemática reuniendo, organizando e interpretando la dinámica de los casos de la enfermedad COVID-19 en Costa Rica tomando siempre como insumo la información oficial suministrada por el Ministerio de Salud. En este caso el valor agregado consiste en analizar e integrar la dinámica de los tres años de pandemia por la Covid-19 en Costa Rica con e interés de generar una visión prospectiva.

Costa Rica al 3 de marzo de 2023 en un transcurrir de 156 semanas, que inician el 6 de marzo del 2020 reporta un total de 1.207.545 casos de Covid-19, en esos tres años han fallecido un total 9.265 personas ya sea por Covid-19 o por causas asociadas al virus. Durante este tiempo se han podido caracterizar siete grandes eventos, de los cuales seis pueden clasificarse como olas pandémicas.

Tabla 1. Detalle de los eventos acontecidos en Costa Rica por Covid-19

NOMBRE DEL EVENTO	FECHAS	VARIANTE
EVENTO 0 -84 DÍAS-	6 de marzo y terminó el 24 de mayo del 2020	Original
PRIMERA -280 DÍAS-	25 de mayo del 2020 y terminó el 20 de febrero del 2021	Original
SEGUNDA -154 DÍAS-	21 de febrero del 2021 y hasta el 28 de julio del 2021	Alpha
TERCERA -147 DÍAS-	29 de julio hasta el 16 de diciembre del 2021	Delta
CUARTA -105 DÍAS-	17 de diciembre del 2021 a 15 de abril 2022	Ómicron
QUINTA -217 DÍAS-*	16 abril hasta 12 noviembre 2022	Subvariantes BA.1 BA.2 y BA.4 y BA.5 y B.Q1 y BQ1.1
SEXTA -105 DÍAS-	13 noviembre al 3 marzo 2023	B.Q1 y BQ1.1 y XBB. 1.5 Kraken



* A partir del 30 de mayo del 2022 se perdió la trazabilidad de la pandemia debido al Hackeo sufrido a la CCSS y hasta la fecha no se ha evidenciado interés o voluntad de las instancias oficiales por recuperar el proceso estadístico previo al hackeo.

A continuación se presenta un cuadro resumen con las principales estadísticas para cada evento a analizar.

Tabla 1. Principales estadísticas por Covid-19 para Costa Rica, según evento ola pandémica.

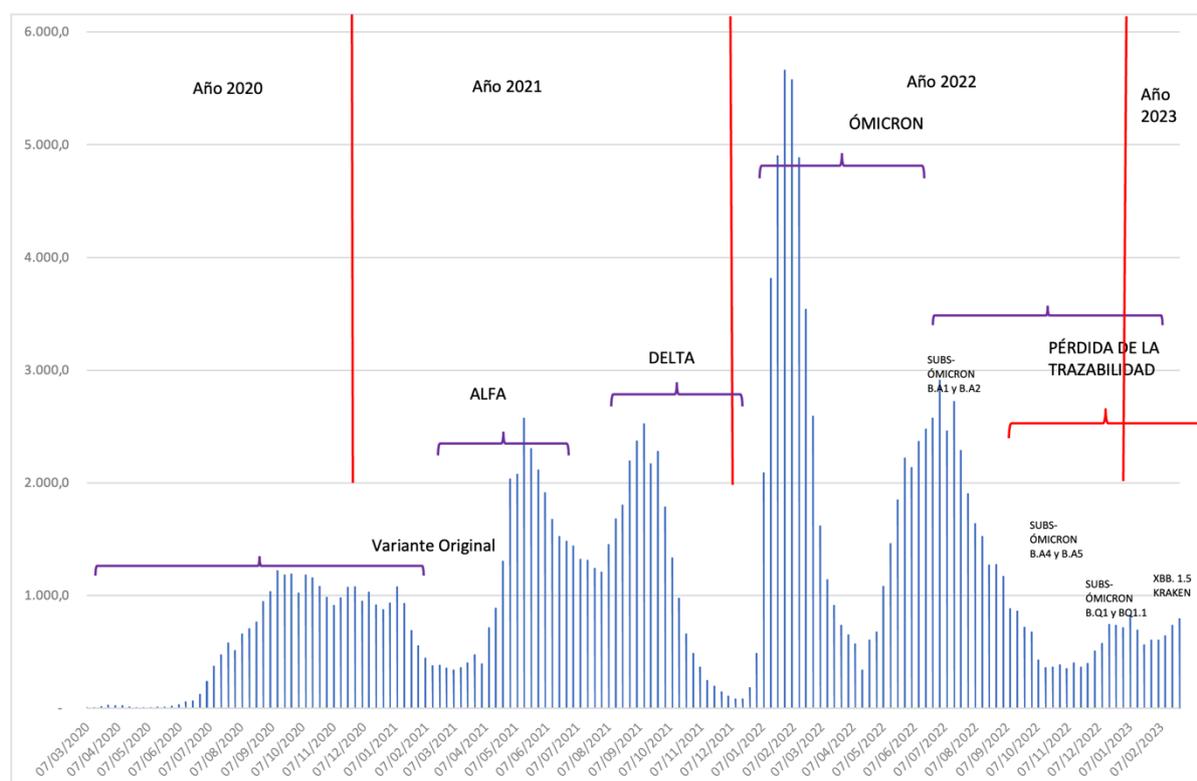
EVENTOS	0. EVENTO 0	1. PRIMERA	2. SEGUNDA	3. TERCERA	4. CUARTA	5. QUINTA	6. SEXTA	PANDEMIA EN CR
CASOS ACUMULADOS	918	203.788	203.494	161.545	274.084	296.388	66.328	1.206.545
FALLECIMIENTOS POR EVENTO	10	2.793	2.239	2.310	996	703	214	9.265
SEMANAS	12	40	22	21	15	31	15	156
PROMEDIO DE FALLECIMIENTOS POR EVENTO	12	40	22	21	15	31	15	464
PROMEDIO DE INTERNAMIENTOS POR EVENTO	18	399	832	781	612	271	142	314
PROMEDIO DE INTERNAMIENTOS EN SALÓN POR EVENTO	9	248	489	481	519	231	128	150
PROMEDIO DE INTERNAMIENTOS EN UCI POR EVENTO	9	151	343	300	93	40	14	2
PROMEDIO DE LA TASA DE REPRODUCIBILIDAD POR EVENTO	1,53	3,91	1,74	0,79	6,54	1,05	1,14	2,48

De esta tabla es interesante rescatar algunos hechos:

- Al analizar el comportamiento de la pandemia en Costa Rica un 50 % del total de los casos reportados se concentran durante el evento Cuarto y Quinto; es decir con el inicio de la afectación por Ómicron y sus variantes.
- Por otro lado, al analizar el comportamiento de los fallecimientos el 80 % de los mismos se reportaron durante la primera, segunda y tercera ola, lo cual es asociado directamente a la afectación de la variante original, Alpha y Delta, mientras que a partir de la cuarta ola esto con la presencia de los procesos de vacunación (de diciembre de 2022 a la fecha) las estadísticas reflejan una mortalidad del 20% de los casos.
- Respecto al comportamiento durante los tres años de pandemia, Costa Rica ha afrontado varias olas pandémicas; unas caracterizadas por variantes altamente contagiosas, sumadas en la etapa inicial a la ausencia de las vacunas.

- Durante el segundo año las principales características estuvieron asociadas dos variantes altamente contagiosas y letales, como lo fueron Alpha y Delta, en donde la estrategia de vacunación impulsada durante esos meses, contuvo lo que pudo ser una alta mortalidad
- Para inicios de 2022 se reportó la afectación por Ómicron una de las variantes de mayor contagio; pero no con una alta tasa de letalidad, la cual a su vez ha ocasiono un gran número de subvariantes que han afectado el país y al mundo, a partir de esa fecha.

Gráfico 1. Evolución del total de casos acumulados de Covid-19 por semana epidemiológica en Costa Rica.



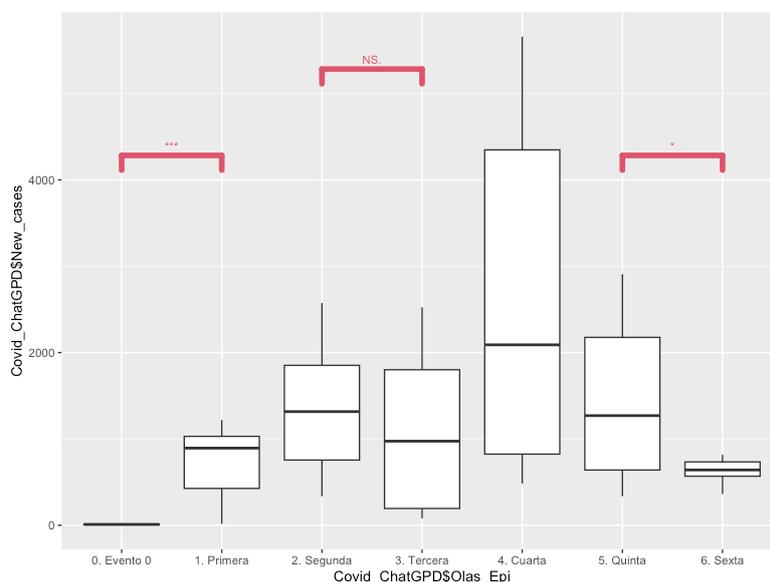
El gráfico anterior permite comprender el comportamiento mencionado de una forma más visual y es así como se puede apreciar que una de las olas pandémica se ve representada por un aumento exponencial en el contagio seguido por un descenso relativamente rápido. Durante el año 2022 se observa el amplio contagio sufrido por la variante Ómicron, durante los meses de enero y febrero, el cual volvió a repuntar a inicios de marzo, cuando el comportamiento y reporte de los casos se vio afectado por el hackeo informático de la CCSS, en donde cambia la tendencia y por casi más de dos meses no se reportó información asociada a Covid-19, ni a ninguna otra sintomatología.

Ya para finales de 2022 el país inicia otro proceso de repunte en la cantidad de contagios el cual ha sido sostenido ya por más de 15 semanas incluyendo las primeras del 2023.

Análisis de casos Covid-19 según ola pandémica

La información que arroja el Gráfico 2 evidencia que el evento cero es significativamente diferente al resto de los eventos, lo que apunta a que la evolución de la pandemia durante esos primeros meses tuvo debido a las medidas sanitarias implementadas, un efecto totalmente deferente y controlado, respecto a los demás eventos. En cuanto al comportamiento de la segunda a la cuarta analizando el estadístico de la mediana (la cual se interpreta como el punto medio de las observaciones) no tiende a ser diferente entre las olas; sobre la quinta ola y respecto al evento seis se puede concluir que la mediana es mucho más alta que lo reportado en los eventos previos a lo largo de los 102 días. En cuanto a la variabilidad durante la cuarta ola se presentaron los casos promedios diarios más altos de toda la pandemia.

Gráfico 2. Diagrama de cajas* del comportamiento promedio de casos semanales por Covid-19 en Costa Rica según evento pandémico

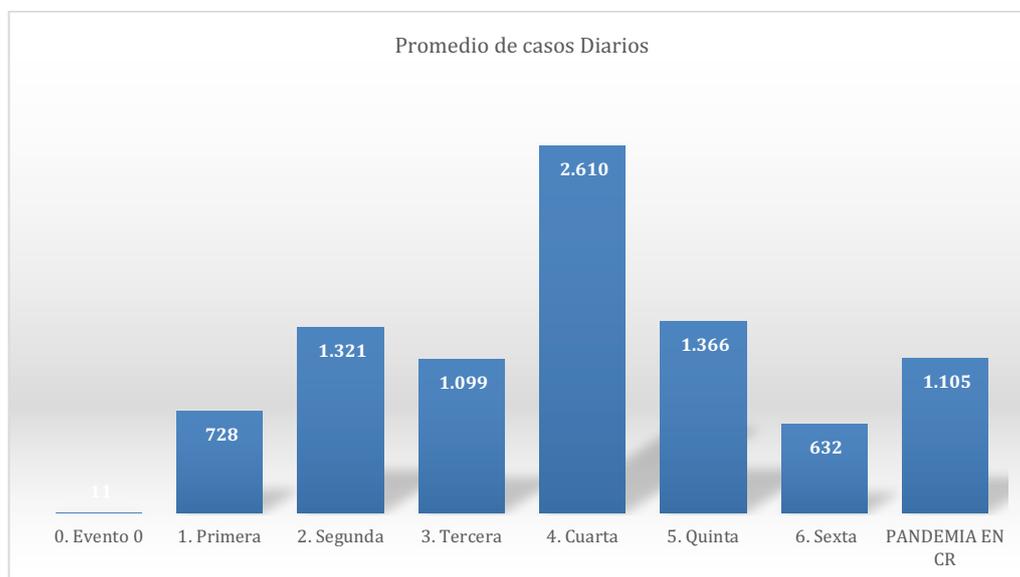


*Para mayor detalle ver Ficha Técnica 1.

Analizando el comportamiento de los casos promedios diarios por Covid-19 durante los eventos de la cuarta ola, el promedio diario fue de 2.610 casos, siendo casi el doble que los reportados durante la

segunda, tercera y quinta ola. Por otro lado, el comportamiento de casos diarios promedios de la sexta ola es muy similar al acontecido durante la primer ola en el año 2020.

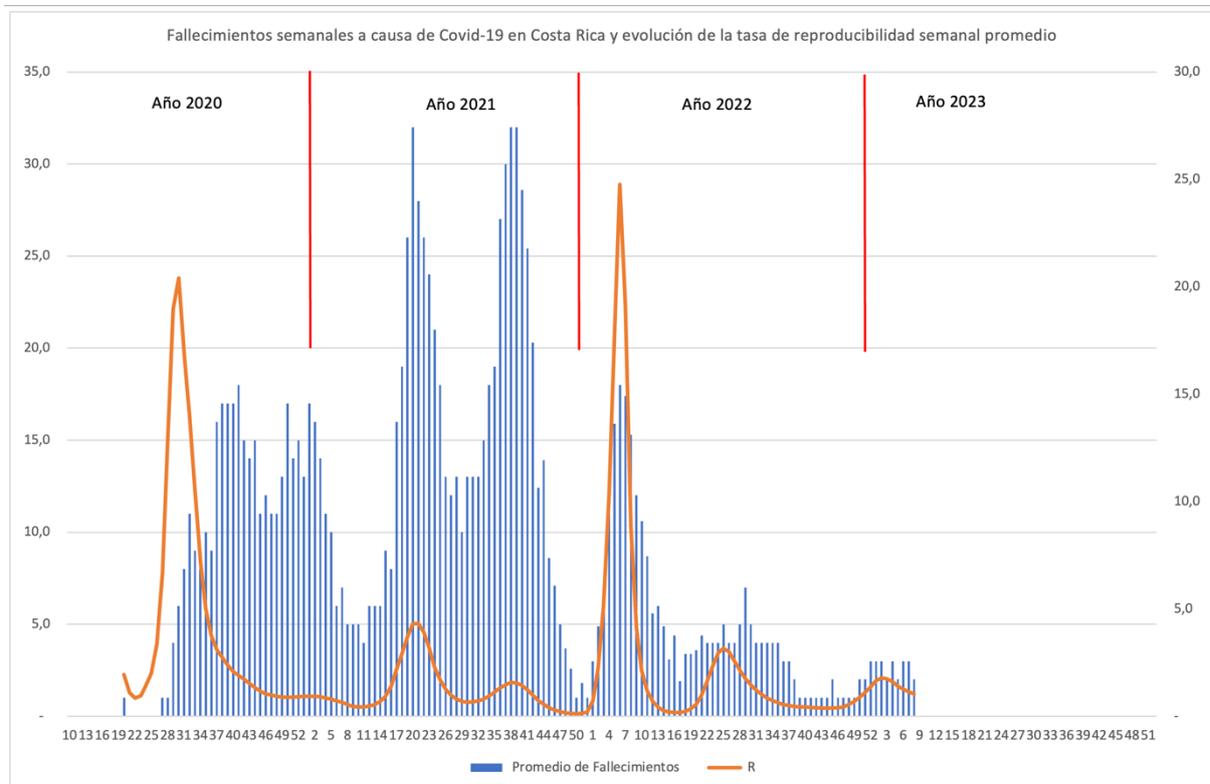
Gráfico 3. Promedio de casos diarios de Covid-19 según evento.



Análisis de fallecimientos por Covid-19

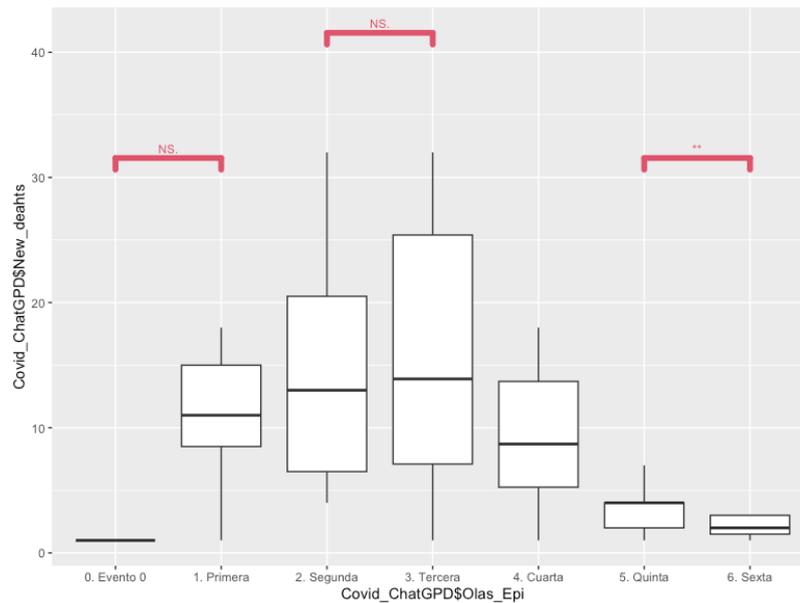
De acuerdo con el gráfico 4, el comportamiento de los fallecimientos semanales se puede analizar en términos de que la tendencia es congruente con las olas pandémicas acontecidas en el país, con excepción del total de fallecimientos durante la quinta ola. Este efecto está directamente asociado al alto proceso de vacunación en segunda dosis que el país tenía a la fecha.

Gráfico 4. Evolución del total de fallecimientos semanales a causa de Covid-19 en Costa Rica y evolución de la Tasa de Reproducibilidad -Rt-



Sobre el comportamiento de los fallecimientos por Covid-19 se aprecia que el evento 0 y la ola quinta junto con el evento 6 son significativamente diferentes a los eventos generados por las olas, primera a la cuarta. Se observa una tendencia clara de decrecimiento en la cantidad de muertes promedio por semana epidemiológico a partir de la ola cuarta, lo cual está asociado con los altos niveles de vacunación mencionados al inicio del reporte.

Gráfico 5. Diagrama de cajas del comportamiento de fallecimientos semanales por Covid-19 en Costa Rica según evento pandémico

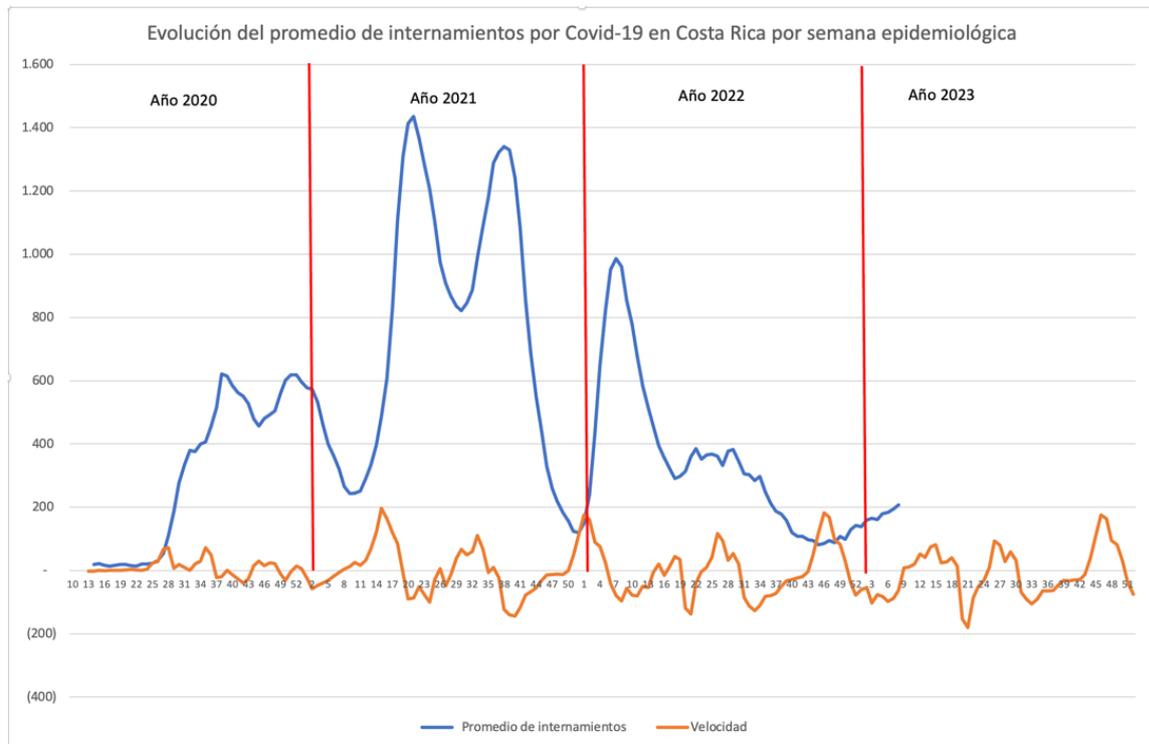


En cuanto al análisis de las muertes reportadas durante la quinta ola y el evento sexto se puede que si existen diferencias significativas en el comportamiento de la pandemia; es decir durante la quinta ola se dio un mayor promedio de fallecimientos por covid-19 que los reportados hasta el momento en la sexta ola.

Análisis de hospitalizaciones por Covid-19

Con respecto a las hospitalizaciones promedio por COVID-19 durante estos tres años, se evidencia que el comportamiento ha sido muy diferente a lo largo de este período. Durante el primer año, se caracterizó por altas medidas sanitarias que ayudaron a la contención y propagación del virus, por tanto a un menor proceso de hospitalización. En el segundo año, se observó una disminución en la aplicación de estas medidas sanitarias, lo que resultó en un aumento en los procesos de hospitalización, sumado a un proceso muy lento en la vacunación contra la Covid-19. Al inicio de 2022, se dio un efecto comprometedor en el sistema de salud, debido al alto volumen de personas contagiadas, es decir la afectación por Ómicron impactó el sistema sanitario, pero la estancia promedio durante esas fechas no fue tan alta, debido a los procesos de vacunación alcanzados a la fecha. Sin embargo, para el inicio de 2023, se puede observar una tendencia al crecimiento importante y sostenido a lo largo de la sexta ola, aumentando a una tasa constante del 4 % de una semana a otra.

Gráfico 6. Evolución promedio de internamientos por Covid-19 en Costa Rica por semana epidemiológica.



Escenario para el 2023

Tomando en cuenta el comportamiento de los casos reportados durante las últimas tres semanas, es claro que el país se nuevamente está en un proceso creciente en el reporte de casos y en el reporte de las hospitalizaciones y se encuentra estacionado en un promedio de fallecimientos de casi tres personas al día. Esto refleja un tipo de crecimiento de carácter exponencial; el cual, a diferencia de los otros eventos, puede impactar al país debido a la baja cobertura de vacunación en tercera y cuarta dosis; así como al relajamiento total de toda medida sanitaria.

Pronóstico para el 2023 y 2024.

Existen diferentes métodos para analizar y proyectar el comportamiento de la pandemia por Covid-19, al cabo de tres años de la misma se decidió utilizar un método de estimación o pronóstico desde el inicio de la pandemia al 31 de diciembre del 2022, para estimar el comportamiento de 2023 y analizar su evolución. El método consiste en una técnica de pronóstico de series de tiempo que

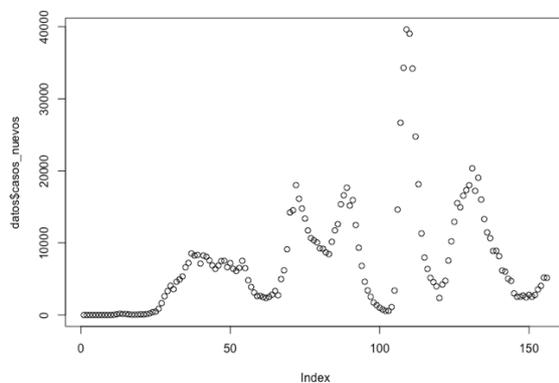
combina el enfoque de STL (*Seasonal-Trend decomposition using Loess*) y el modelo ARIMA (*AutoRegressive Integrated Moving Average*). El primero es un método que separa una serie de tiempo en tres componentes: la tendencia, la estacionalidad y el residuo. La tendencia es la evolución a largo plazo de la serie de tiempo; mientras que la estacionalidad es la variación cíclica que se repite en períodos regulares, y el residuo es la parte de la serie de tiempo que no puede explicarse por la tendencia o la estacionalidad.

El modelo ARIMA, por otro lado, es un modelo estadístico que se utiliza para analizar y pronosticar series de tiempo estacionarias. Se basa en tres parámetros: p , d y q , que representan el número de términos autoregresivos, el número de veces que se diferencia la serie de tiempo para hacerla estacionaria, y el número de términos de media móvil, respectivamente. El método empleado utiliza ambos enfoques y genera un pronóstico más confiable y preciso. Los resultados del modelo no solo arrojan una estimación altamente confiable sino que replican la tendencia que reportada por el país para estas primeras ocho semanas de reporte.

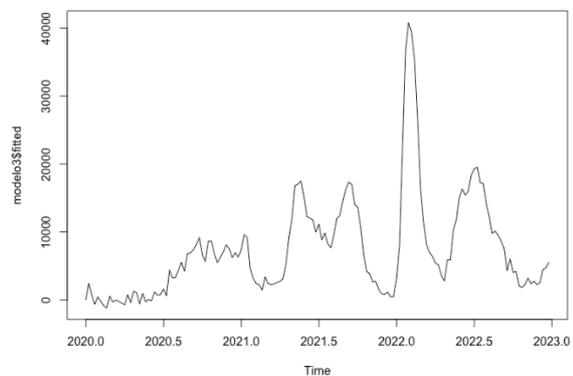
El porque es importante estimar la cantidad posibles de casos radica en la estimación del total de personas que podrían estar ingresando en los procesos de hospitalización, así como los posibles picos u olas futuras, información que es y debería ser utilizada en la formulación de políticas públicas en materia de Salud. En la ficha técnica 2 se presentan el detalle del método así como los resultados estadísticos.

Gráfico 7. Modelo de estimación de casos para el 2023 utilizando método STL + ARIMA

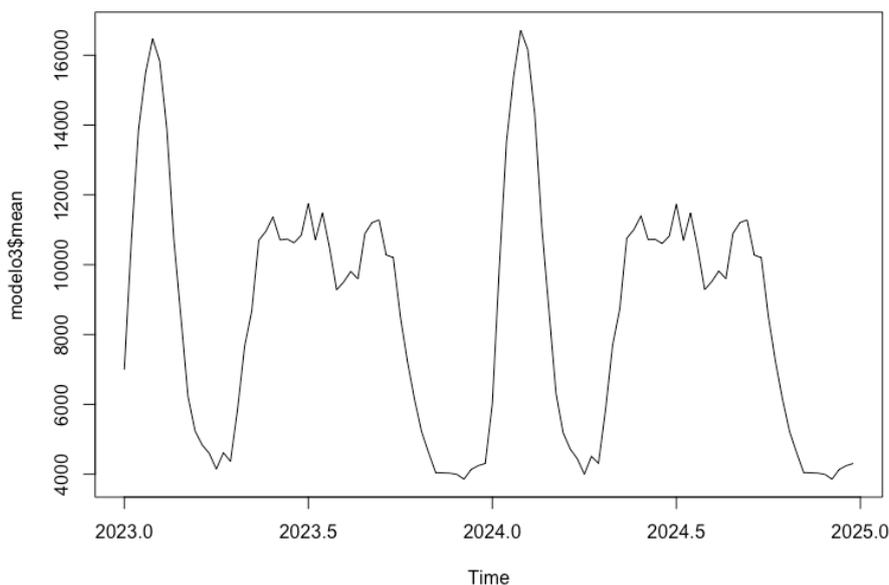
Tendencia real



Estimación



Estimación para el 2023



Para comprender la tendencia en el comportamiento de la pandemia, es necesario asumir ciertos supuestos relacionados con la estacionalidad, linealidad, ausencia de autocorrelación, homocedasticidad e independencia de las observaciones. Sin embargo, estos supuestos no pueden cumplirse en su totalidad, debido a que el reporte de casos están influenciadas por su comportamiento previo; es decir, los casos de una semana afectarán los casos de la siguiente semana debido al aspectos propios del contagio. No obstante, la estimación realizada por el modelo es de momento lo más aproximado y útil para visualizar una posible tendencia de la pandemia desde 2022 hasta 2025.

Teniendo en cuenta estos supuestos, el modelo indica que, si se mantiene la tendencia actual, el país continuará experimentando los efectos de la sexta ola hasta finales de marzo; para luego experimentar un posible repunte, la séptima ola, durante el resto del año. Esta séptima ola estaría caracterizada por una meseta amplia con picos de casos que alcanzarían un promedio de 10 000 casos semanales. Si la tendencia continúa, este fenómeno se repetiría en 2024.

Aunque esta es una estimación de lo que podría estar afectando al país, los casos reales reportados en las últimas dos semanas epidemiológicas, la 7 y la 8, rondan los 5 300 casos en promedio. Si se considera que por cada caso reportado pueden existir al menos dos o tres casos adicionales y el sistema de salud captura todos los casos, el promedio de reporte podría rondar entre 10 600 y 17 000 casos semanales esto no difiere mucho de la estimación del modelo.



En cuanto a las hospitalizaciones por Covid-19, se estima que entre 250 y 390 personas podrían estar siendo internadas semanalmente. Aunque esta cifra puede ser alarmante, el sistema hospitalario reportó un total de 207 personas internadas durante la semana epidemiológica 8. Sin embargo, el crecimiento del 4 % indica que se podría llegar a un total de 300 personas hospitalizadas al cabo de 8 semanas, es decir, a mediados de abril de 2023.

Conclusiones a tres años de pandemia.

1. Si bien es cierto el escenario era completamente incierto, el modelo costarricense y las decisiones de política pública le permitieron a Costa Rica tomar acciones inmediatas para evitar en primera instancia un contagio masivo y por lo tanto un escenario de letalidad que pudo haberse salido de control.
2. La generación de protocolos y medidas sanitarias que científicamente fueron evaluadas y comprobada por parte de la Universidad de Costa Rica, fueron insumos claves para el manejo y la contención de la pandemia en Costa Rica.
3. El sistema de Salud logro materializar su capacidad de respuesta construyendo una estrategia de atención que le permitió no solo atender a casi 1500 personas diarias internadas por una misma enfermedad en las diferentes modalidades de atención sino también lo concerniente a las unidades de cuidados intensivos que se constituyeron en el elemento vital para salvaguardar vidas sin generar un colapso hospitalario.
4. Como efecto colateral de las medidas sanitarias las brechas en salud, economía, empleo y educación hicieron que el país esté en una transición muy complicada de recuperación post-pandemia, y adicionalmente consecuencias como el apagón educativo, pero a esto se le suma la crisis por los contenedores, la guerra entre Rusia y Ucrania que siguen afectando la estabilidad económica del país y del mundo en general.
5. Es importante tener en cuenta que la pandemia no ha terminado aún ni estará pronta a terminar, alrededor del mundo se siguen registrando casos y fallecimientos, en donde las lecciones aprendidas nos indican que las herramientas para frenar esta tendencia es la vacunación y las medidas sanitarias.
6. También es importante resaltar que el país enfrenta actualmente una ola, la cual está derivando en un crecimiento constante semana a semana de personas en procesos de hospitalización y un aumento importante en los fallecimientos, los cuales requieren de una atención inmediata para evitar escenarios complicados o de posible colapso del sistema sanitario.

NOTA TÉCNICA 1.

El gráfico de caja (“box-plot” en inglés) es una forma de presentación estadística destinada, fundamentalmente, a resaltar aspectos de la distribución de las observaciones en una o más series de datos cuantitativos. Fue ideado por John Tukey, de la Universidad de Princeton (EUA.) en 1977 y los detalles que siguen corresponden a la descripción dada por este autor.

ELEMENTOS CONSTRUCTIVOS. Este gráfico utiliza una sola escala: la correspondiente a la variable de los datos que se presentan. Es decir, no utiliza escala de frecuencias. Por lo tanto, no corresponde asociarlo a los que utilizan el sistema de coordenadas cartesianas. Los elementos que los constituyen son (ver figura 1):

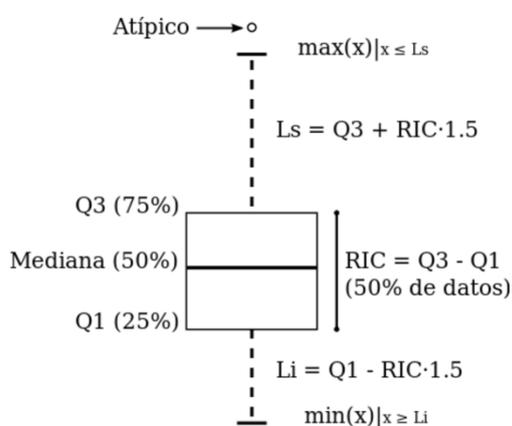


Diagrama de Caja (Fuente: Wikipedia)

- La caja: Es un rectángulo que abarca el recorrido (o rango, o intervalo) intercuartílico (RIC) de la distribución; o sea, el tramo de la escala que va desde el primer cuartil (C1) al tercer cuartil (C3). Esto incluye el 50 % de las observaciones centrales.
- Mediana: Se dibuja mediante una línea (algunos lo marcan con un asterisco, otros con una cruz) dentro de la caja y a la altura de la escala que corresponde al valor de esa medida.
- Bigotes: Son líneas que salen a los costados de la caja y que sirven como referencia para ubicar las observaciones que están por fuera del 50 % central de la distribución.
- Cercados interiores: Indica la finalización de los bigotes. A veces no se dibujan.
- Cercados exteriores: Ubicados más periféricamente en la distribución. Casi nunca se dibujan.
- Periféricos (o periféricos próximos): Señalamiento de las observaciones que se encuentran entre el cercado interior y el cercado exterior.
- Periféricos lejanos (o periféricos extremos): Señalamiento de las observaciones que se encuentran fuera del cercado exterior.

(Palladino, 2011)



Nota técnica 2.

El método de *forecast* STL + ARIMA (4,1,1) es una técnica que se utiliza para hacer pronósticos de series de tiempo, que combina dos enfoques diferentes: el método STL (*Seasonal-Trend decomposition using Loess*) y el modelo ARIMA (*AutoRegressive Integrated Moving Average*).

El enfoque STL es un método de descomposición que separa una serie de tiempo en tres componentes: la tendencia, la estacionalidad y el residuo. La tendencia es la evolución a largo plazo de la serie de tiempo, mientras que la estacionalidad es la variación cíclica que se repite en períodos regulares, y el residuo es la parte de la serie de tiempo que no puede explicarse por la tendencia o la estacionalidad.

El modelo ARIMA, por otro lado, es un modelo estadístico que se utiliza para analizar y pronosticar series de tiempo estacionarias. Se basa en tres parámetros: p , d y q , que representan el número de términos autoregresivos, el número de veces que se diferencia la serie de tiempo para hacerla estacionaria, y el número de términos de media móvil, respectivamente.

El método de *forecast* STL + ARIMA (4,1,1) combina estos dos enfoques para producir un pronóstico más preciso y confiable. Primero, se aplica el método STL para descomponer la serie de tiempo en sus componentes de tendencia, estacionalidad y residuo. A continuación, se aplica el modelo ARIMA (4,1,1) a los datos residuales para modelar cualquier patrón remanente que no se haya explicado en la fase anterior. Finalmente, se suman los resultados de la predicción de ambos enfoques para obtener el pronóstico final de la serie de tiempo.

Supuestos del modelo

El modelo STL + ARIMA (4,1,1) para pronósticos de series de tiempo se basa en varios supuestos. A continuación, se detallan los más importantes:

- Estacionariedad: El modelo ARIMA se basa en el supuesto de que la serie de tiempo es estacionaria, lo que significa que su media y varianza no cambian con el tiempo. Para lograr la estacionariedad, se puede aplicar una transformación a los datos, como la diferenciación o la transformación logarítmica.
- Linealidad: El modelo ARIMA asume que la relación entre los valores pasados y futuros de la serie de tiempo es lineal. Es decir, supone que la serie de tiempo sigue un patrón predecible que puede ser modelado mediante una combinación lineal de términos autoregresivos y de media móvil.
- Ausencia de autocorrelación: El modelo ARIMA también supone que no hay autocorrelación entre los residuos, es decir, que los errores de predicción son aleatorios y no están correlacionados con los errores pasados.
- Homocedasticidad: El modelo ARIMA también supone que los errores tienen una varianza constante a lo largo del tiempo. Es decir, que la magnitud de los errores no cambia a medida que se avanza en el tiempo.
- Independencia: El modelo ARIMA supone que las observaciones de la serie de tiempo son independientes unas de otras, lo que significa que no hay una relación sistemática entre ellas. Esto implica que la serie de tiempo no se ve afectada por factores externos no incluidos en el modelo.



En resumen, el método de forecast STL + ARIMA (4,1,1) combina la descomposición de la serie de tiempo en tendencia, estacionalidad y residuo con un modelo ARIMA para producir pronósticos más precisos y confiables de una serie de tiempo.

Proceso automatizado con ChatGPT

Forecast method: STL + ARIMA(4,1,1)

Model Information:

Series: x

ARIMA(4,1,1)

Coefficients:

	ar1	ar2	ar3	ar4	ma1
	0.2219	0.6814	-0.1078	-0.4283	0.3652
s.e.	0.1935	0.1406	0.0906	0.0917	0.2097

sigma² = 1415817: log likelihood = -1315.88

AIC=2643.77 AICc=2644.34 BIC=2662.03

Error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	18.37707	1166.774	855.5994	NaN	Inf	0.1170491	-0.00985087

```
# Convertir el vector de fecha en una serie temporal
datos_ts <- ts(datos$casos_nuevos, start = c(2020, 1), frequency = 52)
```

```
# Ajustar modelo
```

```
library(forecast)
```

```
modelo3 <- stlf(datos_ts, method = "arima", s.window = "periodic")
```

```
mae3 <- mae(fitted(modelo3), datos_ts)
```

```
print(paste0("MAE Modelo 3: ", mae3))
```

```
plot(datos$casos_nuevos)
```

```
plot(modelo3$fitted) #mejor ajuste
```

```
plot(modelo3$lower)
```

```
plot(modelo3$mean)
```

```
plot(modelo3$upper)
```

```
summary(modelo3)
```