

# Impacto de la cuarentena en la disminución del índice de crecimiento potencial de muertes y positividad en zonas de riesgo extremo en Perú

Tania Incio Gomez  
Lima, Perú

**Abstract—El coronavirus COVID-19 ha sido clasificado por la Organización Mundial de la Salud como una emergencia en salud pública de importancia internacional, nuestro estudio se encargará de analizar el impacto que ha tenido el coronavirus desde que llegó al Perú así como la aplicación de un modelo de riesgo epidemiológico y cálculo del indicador estadístico EPG que medirá el índice de crecimiento potencial de casos positivos por COVID-19 comparando este indicador antes de iniciar la cuarentena obligatoria, durante y después de ella.**

**Keywords—COVID, Python, Análisis, estadística, probabilidad, Matplotlib, Pandas.**

## I. INTRODUCCIÓN

La cuarentena ha sido un método ampliamente implicado durante muchas décadas como uno de los métodos más antiguos y eficaces para controlar los brotes de enfermedades transmisibles. La cuarentena se define como la restricción de personas que han estado expuestas a una enfermedad contagiosa y el aislamiento se define como la separación de las personas enfermas. Sin embargo, estas dos palabras se usan indistintamente hoy en día. Un nuevo coronavirus llamado coronavirus del síndrome respiratorio agudo severo (SARS-Cov2) surgió como la causa de neumonía viral en Wuhan, China, en diciembre de 2019. La Organización Mundial de la Salud (OMS) la denominó enfermedad por coronavirus 2019 o “COVID-19” y posteriormente la declaró estado pandémico el 11 de marzo de 2020. Al 27 de marzo de 2020, más de 500 000 casos de COVID-19, incluidas 23 000 muertes, se han reportado en más de 180 países y territorios a nivel mundial. China fue el primer gobierno en imponer cuarentenas y prohibiciones de viaje en Wuhan el 23 de enero de 2020 y luego en todo el país. Notaron que la enfermedad difícilmente se controlaría sin restricciones tan estrictas. La eficacia de las políticas de salud siempre ha sido de importancia, especialmente durante los brotes de enfermedades contagiosas, y merece una investigación.

La cuarentena se ha utilizado como una medida temprana de contención de COVID-19 en muchos países, con medidas adicionales de distanciamiento físico también introducidas a medida que los brotes han crecido. Esto se planteó para mantener el control de la infección y, al mismo tiempo, reducir la alteración de las poblaciones, es necesario comprender qué combinación de medidas, incluidos nuevos enfoques de rastreo digital y distanciamiento físico menos intensivo, podría ser necesaria para reducir la transmisión.

Nuestro objetivo fue analizar si efectivamente la cuarentena genera un impacto positivo en la reducción del índice de crecimiento potencial (EPG) de casos positivos y muertes en las zonas de riesgo extremo establecidas en Perú en Enero 2021, las cuales fueron catalogadas como zonas de riesgo extremo a causa del alto nivel de EPG que mantenía un crecimiento progresivo semana a semana sin indicios de disminución.

## II. TRABAJOS RELACIONADOS

- A. *Efficacy of Mass Quarantine as Leverage of Health System Governance During COVID-19 Outbreak (Eficacia de la cuarentena masiva como impulsor de la gobernanza del sistema de salud durante el brote de COVID-19)*  
Irán - Abril 2020

El análisis explica que la cuarentena fue y es uno de los métodos más temidos e incomprensidos para controlar las epidemias de enfermedades infecciosas, la cual puede tener importantes implicaciones psicológicas, emocionales y financieras para las familias y los pacientes en cuarentena. Considerando la propagación del COVID-19 en la mayoría de los países, los formuladores de políticas de salud necesitan evidencia convincente de que las políticas implementadas previamente, como la cuarentena, fueron medidas efectivas. Por ello el estudio, tuvo como objetivo el investigar la eficacia de la política de cuarentena masiva aplicada en China durante el brote de COVID-19.

A nivel de resultados, los estudios confirmaron el efecto positivo de implementar cuarentena masiva y limitaciones de movimiento en China durante la epidemia de COVID-19. Sin embargo, dos estudios mostraron que la restricción de viajes y el bloqueo de la ciudad no eran muy eficientes. Mientras tanto, dos estudios también sugieren métodos efectivos distintos de la cuarentena sin mencionar el éxito o el fracaso de la cuarentena. En comparación con otras enfermedades infecciosas emergentes, por ejemplo, el síndrome respiratorio agudo severo (SARS) en 2003, Canadá, China y Taiwán fueron los países más afectados por SARS. Se cree que el uso excesivo de la cuarentena ha provocado temor en el público y, por lo tanto, ha resultado contraproducente.

- B. *Cuarentena sola o en combinación con otras medidas de salud pública para controlar la covid-19*  
Chile - Setiembre 2020

El análisis se enfoca en la evaluación de la efectividad de la cuarentena durante los brotes graves de coronavirus. Se definieron regiones con alta transmisión de la enfermedad para realizar el análisis. Fueron aplicados estudios de cohortes, estudios de casos y controles, series de tiempo, series de tiempo interrumpido, series de casos y estudios de modelización matemática que evaluaron el efecto de la cuarentena para controlar la covid-19.

Se incluyeron 51 estudios; cuatro estudios observacionales y 28 estudios de modelización sobre covid-19, un estudio observacional y un estudio de modelización sobre el MERS, tres estudios observacionales y 11 de modelización sobre el SARS.

A nivel de resultados, los estudios de modelización informaron de forma uniforme de un efecto beneficioso de las medidas de cuarentena simuladas. Por ejemplo, la cuarentena de las personas expuestas a casos confirmados o

sospechosos podría evitar entre el 44% y el 96% de los casos incidentes y entre el 31% y el 76% de las muertes. En conclusión, los resultados indican de forma consistente que la cuarentena es importante para reducir la incidencia y la es importante para garantizarla efectividad.

III. MÉTODO

Dado que nuestra investigación se basa en el análisis de casos positivos y muertes por COVID-19, optamos por consumir información que brinda el estado peruano desde su Plataforma Nacional de Datos Abiertos (PNDA). Estas fuentes nos proporcionaban la data necesaria para poder atacar la propuesta de investigación y validar o rechazar nuestra hipótesis, por ello, no se requirió consumir otra información que no sea de la PNDA. A continuación, detallaremos de manera específica qué data sets fueron usados, cómo están compuestos y de que manera fueron analizados.

A. Set de datos

Utilizamos dos data sets para cubrir con la data de casos positivos de en Perú y las defunciones totales con la posibilidad de identificar las causadas por COVID-19. Se optó por considerar una fuente con el total de defunciones dado que, en entrevistas previas con especialistas, se señaló que el sistema salud muchas veces no etiqueta de manera correcta una defunción. Existen casos de personas que cuentan con una enfermedad pre existente y que con el COVID se agrava y termina falleciendo la persona, sin embargo, en la data de defunción se etiqueta a la persona con la enfermedad pre existente como causa de defunción a pesar de que el COVID realmente fue el causante.

1. Casos positivos por COVID-19 – Ministerio de Salud (MINSA)

- Los datos publicados corresponden al total de casos reportados que dieron positivo al COVID – 19, por departamento, provincia y distrito. Además, dentro del data set se encuentran datos que permiten identificar características principales del paciente tales como edad, sexo y fecha de obtención del resultado positivo. Estas variables nos permitirán perfilar al tipo de paciente positivo en las regiones de alto riesgo que se delimitan en esta investigación. A nivel de frecuencia de actualización del reporte, el Portal de Datos Abiertos indica que el MINSA realiza una actualización diaria de la información.

Columna	Tipo dato
Fecha corte	Date
Uuid	Varchar(300)
Departamento	Varchar(300)
Provincia	Varchar(300)
Distrito	Varchar(300)
Metododx	Varchar(300)
Edad	Int
Sexo	Varchar(300)
Fecha resultado	Date

Tabla 1: Lista de campos disponibles - Casos Positivos COVID-19

2. Información de Fallecidos - Sistema Informático Nacional de Defunciones (SINADEF)

- La data que disponibiliza el SINADEF corresponde al total de fallecimientos que incluye defunciones fetales, generales y los diagnósticos y antecedentes de enfermedades causantes de la defunción. Estas variables son relevantes para nuestro análisis dado que, según lo explicado al inicio de este apartado, las defunciones no tienen un 100% de precisión al momento de etiquetar la

mortalidad durante la pandemia de covid-19, aunque no existe certeza sobre la magnitud del efecto. La aplicación temprana de la cuarentena y la combinación de la cuarentena con otras medidas de salud pública

causa del fallecimiento, por ello optamos por trabajar con toda la data y evaluar los incrementos y decrementos de defunciones. A nivel de frecuencia de actualización del reporte, el Portal de Datos Abiertos indica que el SINADEF realiza una actualización diaria de la información.

Columna	Tipo dato
Tipo seguro	Varchar(300)
Sexo	Varchar(300)
Edad	Varchar(300)
Estado civil	Varchar(300)
Nivel de instrucción	Varchar(300)
Pais domicilio	Varchar(300)
Departamento	Varchar(300)
Provincia	Varchar(300)
Distrito	Varchar(300)
Fecha	Date
Año	Varchar(300)
Mes	Varchar(300)
Tipo lugar	Varchar(300)
Institucion	Varchar(300)
Muerte violenta	Varchar(300)
Necropsia	Varchar(300)
Debido a (causa a)	Varchar(300)
Causa a (cie-x)	Varchar(300)

Tabla 2: Lista de campos disponibles – Defunciones SINADEF

B. Extracción, uso y limpieza

Desde el Portal de Datos Abiertos, nos dirigimos a la URL con la información en formato .csv proveniente directamente desde el MINSA, se visualizó en hojas de cálculo para luego proceder con la subida a un notebook en Google Colab utilizando la librería Pandas en la importación de los datos.

```
import requests
# Download csv from MINSA
url="https://cloud.minsa.gob.pe/s/Y8w3wHsEdYQS2Rp/download"
s=requests.get(url).content

df_pos = pd.read_csv(io.StringIO(s.decode('utf-8')), sep=';')
```

Fig. 1: Código utilizado para la carga de datos MINSA

```
# Download csv from SINADEF
url="https://cloud.minsa.gob.pe/s/ngF2irNbFomCLaa/download"
s=requests.get(url).content

df_fall = pd.read_csv(io.StringIO(s.decode('utf-8')), sep=';')
```

Fig. 2: Código utilizado para la carga de datos SINADEF

Como siguiente paso, optamos por realizar la limpieza de los datos:

Para los casos positivos (MINSA): Eliminamos todas los registros de la columna DEPARTAMENTO que tengan como valor ‘SIN REGISTRO’, al mismo tiempo que eliminamos todos los registros que tengan como valores NULOS en las columnas FECHA\_RESULTADO y DEPARTAMENTO usando el método dropna()

```
[ ] data.isnull().values.sum(axis = 0)

array([ 0,  0,  0,  0,  0,  0,  0, 167,  0, 3818])

[ ] data.isnull().sum()

[ ] # Remove the last 4 columns Unnamed
data_clean = data.drop(columns=['UUID'], axis=1)

[ ] data_clean.isna().sum()
```

Fig. 1: Limpieza de datos - Casos Positivos COVID-19

Para el número de fallecidos - SINADEF: Eliminamos todas los registros de la columna ‘DEPARTAMENTO DOMICILIO’ que tengan como valor ‘SIN REGISTRO’, al mismo tiempo que eliminamos todos los registros que tengan como valores NULOS en las columnas anteriormente mencionadas con el método dropna().

```
data.isnull().values.sum(axis = 0)

data.isnull().sum()

# Remove the last 4 columns Unnamed
data = data.drop(columns=['Unnamed: 31','Unnamed: 32','Unnamed: 33','Unnamed: 34'], axis=1)

data.isna().sum()
```

Fig. 2: Limpieza de datos – Defunciones SINADEF

#### IV. HIPÓTESIS

La hipótesis del presente trabajo de investigación se diseña como una relación causal y se enuncia de la siguiente forma:

##### A. Hipótesis Alternativa(Ha)

El índice de crecimiento potencial(EPG) de positividad por COVID-19 disminuye al aplicar la cuarentena obligatoria en Perú.

##### B. Hipótesis Nula(H0)

El índice de crecimiento potencial(EPG) de positividad por COVID-19 no disminuye al aplicar la cuarentena obligatoria en Perú.

#### V. ANÁLISIS DE RESULTADOS

En este estudio empleamos un método empírico de riesgo epidemiológico, el cual calcula el índice de crecimiento potencial(EPG) de casos positivos que nos permitirá validar la hipótesis, este método ha sido verificado con la evolución del número de casos confirmados en otros países.

##### A. Modelo Empírico de Riesgo Epidemiológico desarrollado por BioComSc(Computational Biology and Complex Systems Research Group) - Universidad de Cataluña

Este modelo está basado en el modelo de Gompertz. Estudios anteriores han demostrado que el referido modelo de Gompertz describe correctamente la epidemia de COVID-19 en todos los países analizados con el mismo. Es un modelo empírico que parte de un crecimiento exponencial de casos, los cuales luego disminuyen gradualmente en su tasa de crecimiento específica. Por lo tanto, este modelo es adecuado para describir una epidemia como la actual que se caracteriza por un crecimiento exponencial inicial, y posteriormente por una disminución progresiva en la velocidad de propagación.

###### a. Modelo de Gompertz

Es un modelo empírico que parte de un crecimiento exponencial pero que disminuye gradualmente su tasa de crecimiento específica. Por lo tanto, es adecuado para describir una epidemia que se caracteriza por un crecimiento exponencial pero una disminución progresiva en la velocidad de esparsimiento siempre que un control apropiado sea aplicado.

El modelo de Gompertz se describe mediante la ecuación:

$$N(t) = K e^{-ln(\frac{K}{N_0})} \cdot e^{-a \cdot (t-t_0)}$$

Donde N(t) es el número acumulado de casos confirmados en t (en días), y N0 es el número de casos acumulados en el día t0.

Este modelo tiene dos parámetros:

- a = velocidad a la que se ralentiza la tasa específica de propagación.
- K = Número de casos acumulados esperados a la fecha de la última actualización

Teniendo como base el modelo de Gompertz, el Grupo de Investigación de la Universidad de Cataluña (BioComSc) ha calculado los siguientes indicadores:

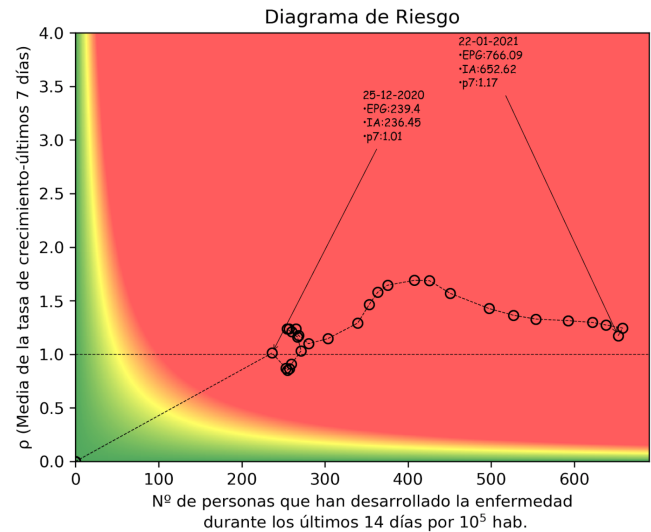
- ρ: esta variable está relacionada con el número de reproducción, es decir, con el número de nuevas infecciones causado por un solo caso. Se evalúa de la siguiente manera para el día anterior al último informe (t-1):

$$\rho(t-1) = \frac{N_{new}(t) + N_{new}(t-1) + N_{new}(t-2)}{N_{new}(t-5) + N_{new}(t-6) + N_{new}(t-7)}$$

Donde N(t) es el número de nuevos casos confirmados en el día t.

- IA(Incidencia Acumulada): Número de casos o muertes acumuladas dividido por el número de casos o muertes confirmados acumulados y notificados como un porcentaje, es un indicador indirecto del nivel de diagnóstico.

Basado en estos dos indicadores podemos crear Mapas de Riesgo Epidemiológicos, el siguiente gráfico es un ejemplo de Mapa de Riesgo de la población de Perú a partir del día 25/12/2020 al 25/01/2021, donde claramente podemos notar la evolución de la pandemia:



##### a. Interpretación del mapa de riesgo

- La escala de color de fondo se establece de acuerdo con el valor del índice EPG (verde, EPG <30, rojo, EPG> 100)
- El EPG o índice de crecimiento potencial es calculado por la multiplicación de las variables  $p$  e  $IA$  es decir del producto del eje X con el eje Y.

- El eje x muestra el indicador de potencial de contagio (IA14, incidencia acumulada de los últimos 14 días, por 100.000 habitantes).
- El eje y muestra el indicador de tasa de contagio, el objetivo es mantener la tasa por debajo de 1 lo que indicaría que el número de casos nuevos estaría disminuyendo (p7, número reproductivo empírico).
- Cada punto en el diagrama representa un día y nos indica si una región determinada se encuentra dentro o fuera de la zona de bajo riesgo epidemiológico de COVID-19.
- El color rojo significa zona de Alto Riesgo de COVID-19, el color amarillo zona de Riesgo Medio y el color verde es la zona de Bajo Riesgo, en el ejemplo muestra que el Peru esta en la zona de Alto Riesgo Epidemiologico.

B. Variable analizada: EPG

El EPG es el índice de crecimiento potencial, mide las posibles muertes o casos nuevos diagnosticados en los próximos 14 días, podemos decir que está relacionado con la probabilidad de nuevos brotes epidémicos.

Es calculado por el producto entre la incidencia acumulada de los últimos 14 días con la variable p7 que es la variable que indica la media de personas que se infectan de un mismo caso.

EPG = IA14 \* p7

Usaremos este indicador para calcular el índice de crecimiento potencial de muertes y casos positivos antes de la cuarentena y después de ella para verificar si la cuarentena produjo algún efecto significativo en el crecimiento del EPG.

C. Análisis y aplicación del modelo de riesgo epidemiológico a nuestros datos

Para la aplicación del modelo de riesgo a nuestros datos se realizó los siguientes pasos:

- 1. Revisión de columnas del dataset
  - Para los casos positivos - MINSA

La base de datos del MINSA tiene 9 columnas

Data columns (total 9 columns):		
#	Column	Non-Null Count
0	FECHA_CORTE	1186698 non-null
1	UUID	1186698 non-null
2	DEPARTAMENTO	1186698 non-null
3	PROVINCIA	1186698 non-null
4	DISTRITO	1186698 non-null
5	METODODX	1186698 non-null
6	EDAD	1186408 non-null
7	SEXO	1186698 non-null
8	FECHA_RESULTADO	1184675 non-null

De las cuales usaremos FECHA\_RESULTADO, DEPARTAMENTO, PROVINCIA, DISTRITO y METODODX

- Para el número de fallecidos - SINADEF

La base de datos de SINADEF tiene 35 columnas

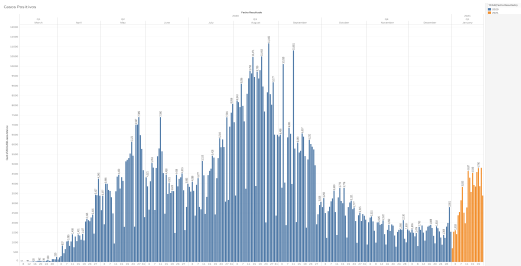
Data columns (total 35 columns):		
#	Column	Non-Null Count
0	N	553362 non-null
1	TIPO SEGURO	553362 non-null
2	SEXO	553362 non-null
3	EDAD	553362 non-null
4	TIEMPO EDAD	553362 non-null
5	ESTADO CIVIL	553362 non-null
6	NIVEL DE INSTRUCCIÓN	553362 non-null
7	COD# UBIGEO DOMICILIO	553362 non-null
8	PAIS DOMICILIO	553362 non-null
9	DEPARTAMENTO DOMICILIO	553362 non-null
10	PROVINCIA DOMICILIO	553362 non-null
11	DISTRITO DOMICILIO	553362 non-null
12	FECHA	553362 non-null
13	AÑO	553362 non-null
14	MES	553362 non-null
15	TIPO LUGAR	553362 non-null
16	INSTITUCION	553362 non-null
17	MUERTE VIOLENTA	553362 non-null
18	NECROPSIA	553362 non-null
19	DEBIDO A (CAUSA A)	553362 non-null
20	CAUSA A (CIE-X)	553362 non-null
21	DEBIDO A (CAUSA B)	553362 non-null
22	CAUSA B (CIE-X)	553362 non-null
23	DEBIDO A (CAUSA C)	553362 non-null
24	CAUSA C (CIE-X)	553362 non-null
25	DEBIDO A (CAUSA D)	553362 non-null
26	CAUSA D (CIE-X)	553362 non-null
27	DEBIDO A (CAUSA E)	553362 non-null
28	CAUSA E (CIE-X)	553362 non-null
29	DEBIDO A (CAUSA F)	553362 non-null
30	CAUSA F (CIE-X)	553362 non-null
31	Unnamed: 31	0 non-null
32	Unnamed: 32	0 non-null
33	Unnamed: 33	0 non-null
34	Unnamed: 34	0 non-null

De las cuales usaremos FECHA, ‘DEPARTAMENTO DOMICILIO’, ‘DEBIDO A (CAUSA A)’, ‘DEBIDO B (CAUSA B)’, ‘DEBIDO C (CAUSA C)’, ‘DEBIDO D (CAUSA D)’, ‘DEBIDO E (CAUSA E)’, ‘DEBIDO F (CAUSA F)’

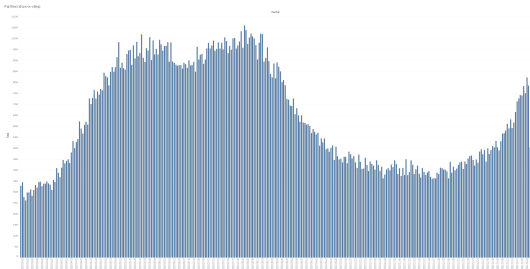
2. Análisis descriptivo

Una vez cargados los datos se graficaron las frecuencias de los siguientes indicadores

- Casos Positivos: Desde el 06/03/2020 al 15/01/2021: En el eje x se muestran las fechas y en el eje y se muestra el número de casos positivos. Podemos ver claramente la evolución del virus en la primera ola (del 06/03/2020 hasta 15/12/2020) y el inicio de la segunda ola (del 15/12/2020 al 15/01/2021)

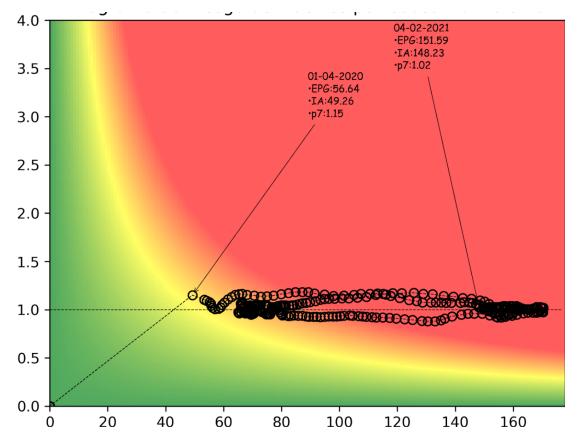


- Número de Muertes: Desde el 06/03/2020 al 15/01/2021: En el eje x se muestran las fechas y en el eje y se muestra el número de muertes causadas por el COVID-19. Podemos ver claramente la curva del crecimiento de muertes en la primera ola (del 06/03/2020 hasta 15/12/2020) y al inicio de la segunda ola (del 15/12/2020 al 15/01/2021)



- Indicador EPG de mortalidad: Desde el 14/04/2020 al 05/02/2021: En el eje x se muestran la densidad de muertes(IA) por 10,000 habitantes y en el eje y el ritmo de muertes(p). Podemos ver claramente la evolución del virus desde el inicio de la pandemia.



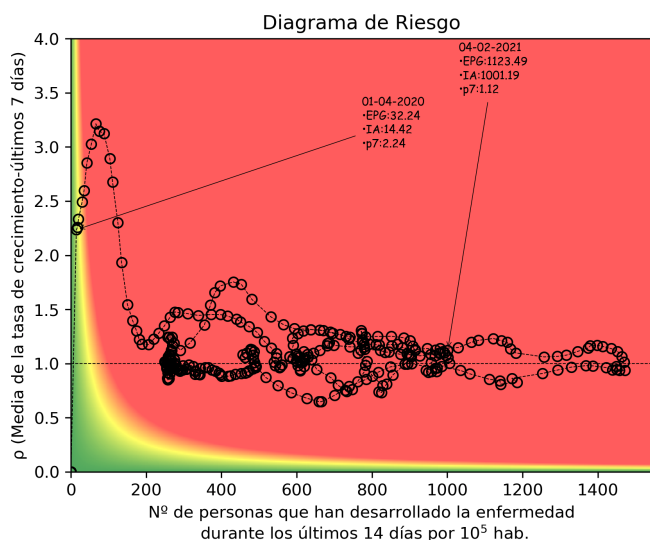


EPG - día 01/04/2020 = 56.64

EPG - día 05/02/2021 = 151.59

Al revisar el EPG en ambos extremos nos damos cuenta que para el mes de abril estábamos cerca a la zona de bajo riesgo con un EPG cerca a 57 pero para el día 05/02/2021 vemos un incremento significativo del EPG a 152 pasando a la zona de riesgo alto, lo que nos hace pensar que la decisión del estado peruano de decretar cuarentena obligatoria en algunas regiones tiene sentido.

- Indicador EPG de positividad: Desde el 01/04/2020 al 04/02/2021: En el eje x se muestran el número de personas infectadas durante los últimos 14 días (IA) por 10,000 habitantes y en el eje y el ritmo de crecimiento de infectados(p).



-EPG en el día 01/04/2020 = 32.24

-EPG en el día 05/02/2021 = 1123.49

Aproximadamente después de 3 semanas de iniciada la pandemia en Perú(día 01/04/2020) el gráfico nos muestra que estábamos en la zona de bajo riesgo con un EPG de 32.24 pero que ha tenido un crecimiento exponencial al día 04/02/2021 con un EPG de 1123.49 lo que nos ubica en la zona de riesgo alto, en estos casos se recomienda tomar medidas de control al estado peruano.

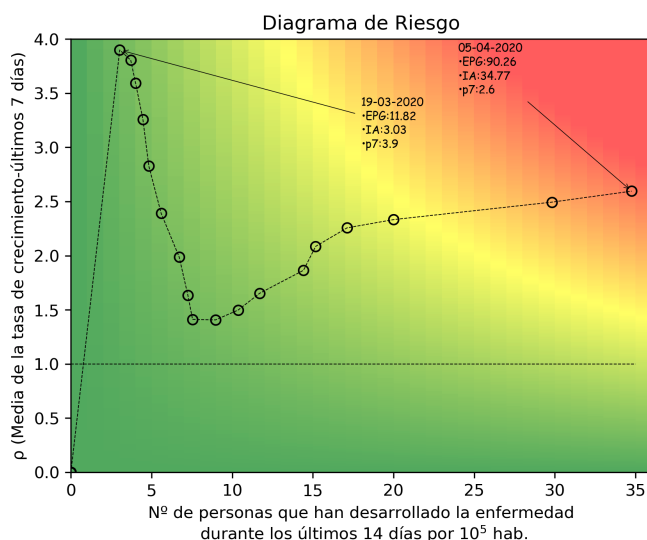
Después de realizar un análisis a nivel nacional de los efectos del COVID-19 procederemos hacer un análisis para comprobar o rechazar nuestra hipótesis.

## 1. Análisis de EPG - Antes de la Primera Cuarentena

Después de analizar los datos de muertes y positividad a nivel nacional procederemos a dividir la data en rangos de tiempo con la intención de revisar la evolución del indicador EPG antes de la cuarentena y después de ella.

La pandemia inició el 06/03/2020 en Perú pero al necesitar mínimo 14 datos acumulados para calcular el EPG procederemos a tomar como fecha de inicio del análisis a

partir del día 19/03/2020. La cuarentena inició el 16/03/2020 pero para visualizar los efectos de la cuarentena se debe esperar mínimo 10 días, así que decidimos analizar el EPG de todo el mes de Marzo, tiempo en que la cuarentena no tendría efecto aún.



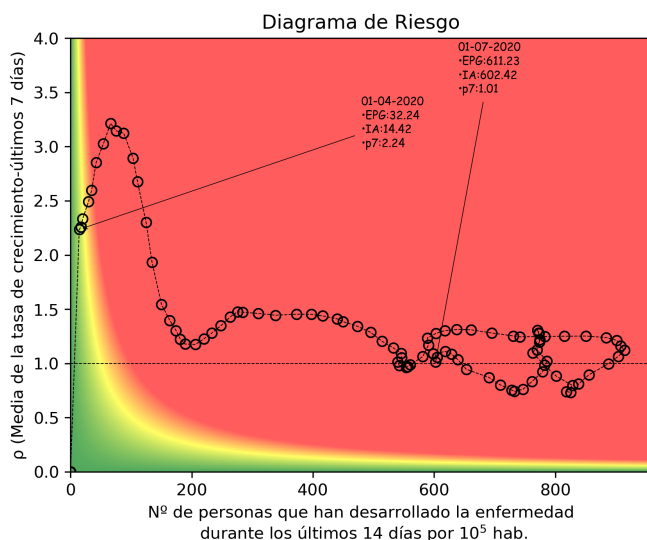
-EPG el día 19/03/2020: 11.82, nos encontramos en la zona de bajo riesgo

- EPG el día 05/04/2020: 90.26 estamos ingresando a la zona de riesgo alto

Durante el mes de Marzo visualizamos el incremento del EPG de 11.82 a 90.26 pasando del nivel bajo al nivel alto de riesgo epidemiológico.

## 2. Análisis de EPG - Durante la Primera Cuarentena

La cuarentena en Perú duró más de 100 días iniciando el día 16/03/2020 al 01/08/2020, los efectos de la cuarentena se pueden visualizar a partir del mes de abril, a continuación mostramos la evolución del EPG del 01/04/2020 al 01/07/2020 .



-EPG el día 01/04/2020: 32.24, nos encontramos en la zona de riesgo medio.

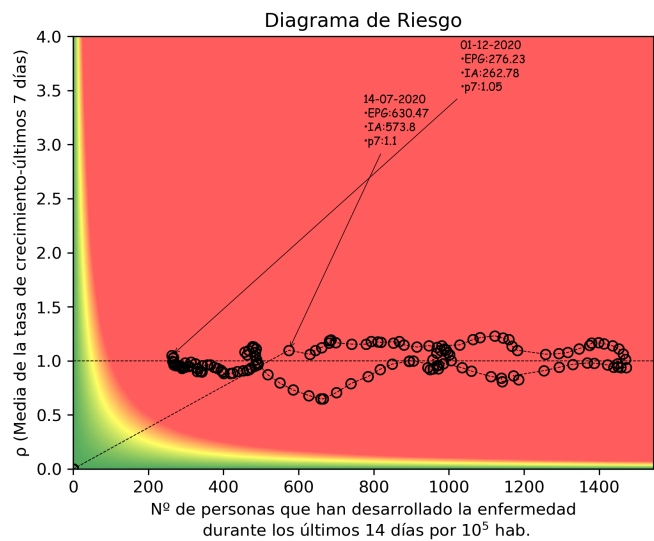
-EPG el día 11/04/2020: 298.76 pico máximo alcanzado.

-EPG el día 01/07/2020: 611.23 estamos en la zona de riesgo alto.

El incremento potencial de positividad o EPG ha tenido un crecimiento significativo durante la cuarentena, inició con 32.24, también podemos ver en la gráfica un pico de EPG de 298.76 que luego baja significativamente y se mantiene con un ritmo de crecimiento(p7) alrededor de 1.5 y nos damos cuenta que a partir de junio a abril la curva regresa y termina la cuarentena con un EPG de 611.23.

### 3. Análisis de EPG - Después de la Primera Cuarentena

La cuarentena en Perú terminó el 01/07/2020, a continuación analizaremos como ha evolucionado el EPG después que el estado peruano finalizó la cuarentena, hemos generado el mapa de riesgo a partir del 14/07/2020 al 01/12/2020.



-EPG el día 14/07/2020: 630.47

-EPG el día 14/08/2020: 1497.13

-EPG el día 01/12/2020: 276.23

Claramente podemos notar el incremento del EPG después de finalizada la cuarentena pero notamos también que el día 14/08/2020 llegó el EPG tomó su valor máximo de 1497.47 (punto extremo derecho del gráfico) y luego inicia un claro descenso de EPG hasta llegar a 276.23 en el mes de diciembre, esto se puede deber a factores como inmunidad de rebaño o zero prevalencia.

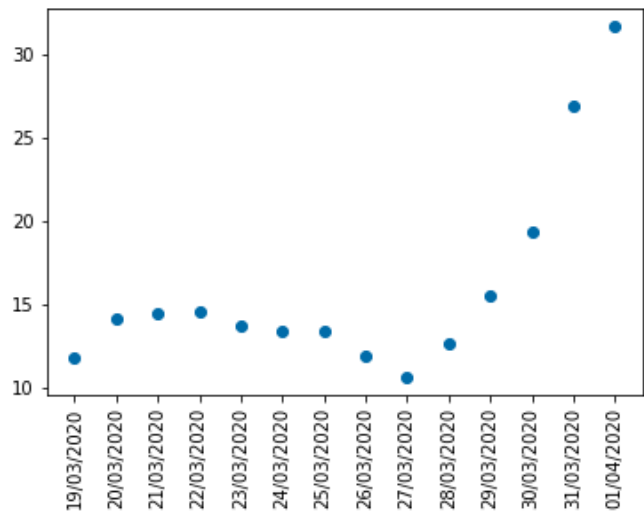
### 4. Conclusión del Análisis de EPG

Nuestro objetivo era determinar si la cuarentena dictada por el estado peruano disminuiría el incremento potencial de casos positivos o EPG, como revisamos al inicio de la cuarentena teníamos un EPG de 34.77 y al finalizar la cuarentena obtuvimos un EPG de 602.42 claramente la cuarentena no disminuyó el indicador EPG, pero es probable que si no hubiera sido aplicada el indicador EPG hubiera sido mayor, eso podría ser tema de otro estudio.

#### A. Prueba de Hipótesis

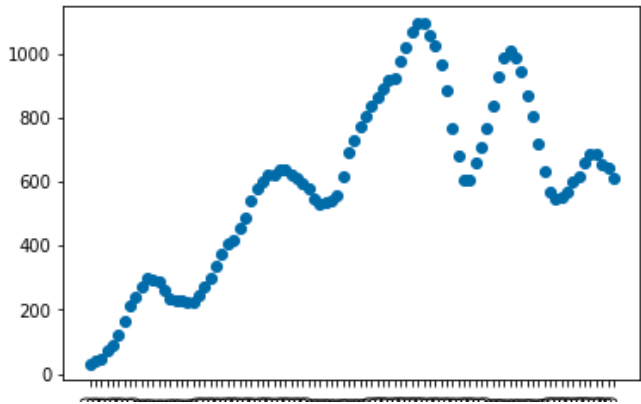
Puesto que nuestro indicador estadístico es el índice de crecimiento potencial o EPG, procedemos a graficar su evolución en cada etapa de la pandemia y revisaremos qué distribución sigue:

- EPG antes de la cuarentena



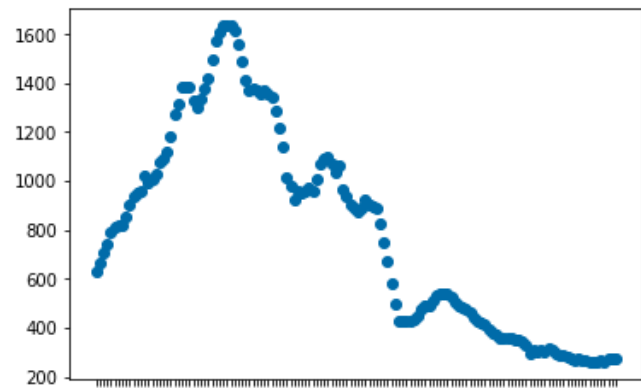
Podemos ver que a partir del 27/03/2020 inicia una distribución exponencial, pero tenemos pocos datos para determinar si es una distribución exponencial, lo confirmaremos en los siguientes gráficos.

- EPG durante la cuarentena



A partir del 01/04/2020 tiende a seguir una distribución exponencial pero luego disminuye a partir del 30/06/2020.

- EPG después de la cuarentena



A partir del 01/07/2020 hasta aproximadamente el 01/11/2020 parece seguir una distribución normal, pero luego tiende a disminuir

Podemos concluir que en las tres etapas analizadas de la pandemia (antes, durante y después de la cuarentena) los datos parecen seguir una distribución exponencial lo cual tendría sentido puesto que analizamos la evolución de la epidemia, podríamos usar una prueba no paramétrica para validar o rechazar nuestra hipótesis, pero dado que nuestra hipótesis es “El EPG disminuye a causa de la cuarentena” claramente vemos que en los gráficos y valores que el EPG no ha disminuido a causa de la cuarentena, por lo tanto no tenemos evidencia para rechazar nuestra hipótesis nula, así que la aceptamos.

Conclusión: La cuarentena no disminuye el índice de crecimiento potencial de positividad de COVID-19 en el Perú.

### VI. DISEÑO DE EXPERIMENTO

#### A. Población y muestra:

Nuestra población estará compuesta por todos los habitantes de Lima en el periodo de inicio en la segunda ola de COVID-19 en el 2021. En base a esta población hemos propuesta trabajar con un muestreo probabilístico de múltiples etapas con el fin de dividir nuestra población en conglomerados definidos por ciudades y distritos, consideramos que de esta manera obtendremos una muestra representativa de toda la región

## B. Tipo de Experimento:

De acuerdo a la hipótesis a validar optamos por trabajar con un experimento controlado como el de grupo de control y grupo experimental, el cual, con su estructura de aplicar una variable independiente y dependiente, nos permitirá comparar los resultados de manera confiable y llegar a una conclusión.

El grupo de control, en nuestro experimento, estará compuesto por las personas de la región Lima que opten por no acatar la cuarentena continuando con sus labores diarias sin ningún tipo de restricción. Evaluaremos en este grupo la evolución del EPG en el mismo periodo de tiempo que el grupo experimental para no alterar los resultados por analizar los datos en diferentes fechas.

El grupo experimental, estará compuesto por las personas de la región Lima que opten por acatar la cuarentena junto con las medidas de restricción propuestas por el gobierno. Se evaluará de igual manera en este grupo la evolución del EPG en el mismo periodo de tiempo con el fin de validar las conclusiones de nuestro análisis previo.

## C. Instrumento de medicion:

Para el desarrollo del experimento hemos seleccionado dos instrumentos de medición como la encuesta para identificar las personas que acataran la cuarentena y la evidencia física para analizar los resultados de contagios y muertes en nuestra muestra seleccionada.

## Sustento confiabilidad, validez y objetividad:

Haciendo un breve resumen de los tres requisitos básicos que debe cumplir un instrumento de medición, decimos que la confiabilidad determina si un instrumento es consistente en base a los resultados que muestra, la validez de contenido refleja el grado en el que un instrumento refleja el dominio específico

de contenido de lo que se mide y finalmente la objetividad hace referencia a qué tan permeable es el instrumento a la influencia de los sesgos y tendencias del investigador o investigadores que lo administran, califican o interpretan. En base a estos conceptos, consideramos que la evidencia física que usaremos, la cual será información directa de instituciones como Minsa y Sinadef, tendrá validez de contenido al ser información regulada proveniente de una institución del estado, será también confiable por ser consistente en sus datos mostrados y finalmente objetivo dado que esta completamente limpia de sesgo de investigador o tendencias externas al ser información automatizada proveniente de los mismos sistemas de origen.

## D. Propósito del experimento

Este análisis y aplicación del modelo epidemiológico no pretende interpretar las causas de la evolución de los casos de COVID-19 sino permitir una predicción de las tendencias de los niveles de riesgo existentes en un área geográfica determinada y en el corto plazo. Además, se debe de tener en cuenta que los efectos de las medidas de control cuya aplicación comience en un día determinado se observan recién hasta 7-10 días después aproximadamente.

Nuestro fin fue entender la evolución de la pandemia en Perú a partir de modelos epidemiológicos y de ser posible entender porque se aplican algunas medidas de control como la cuarentena.

## E. Recursos

A continuación, detallamos los datos involucrados en nuestro estudio.

- Personas: Se necesitó dos analistas para la extracción de la data, limpieza y cálculo de indicadores estadísticos, así como para la creación e implementación de la encuesta.
- Tiempo: Al trabajar con datos abiertos dependemos de entidades como el MINSA o SINADEF sean los encargados de subir los datos diariamente para poder analizarlos.

Para el análisis de la primera cuarentena ya contábamos con datos históricos y pudimos realizar el estudio en un corto plazo (3 semanas).

- Costo: Al utilizar herramientas open-source como Google Colab, python y la versión gratuita de Tableau así como los datos abiertos de MINSA y SINADEF no realizamos gastos económicos

## VII. CONCLUSIONES

A partir del análisis estadístico de los datos y aplicando el modelo epidemiológico de riesgo se puede observar que hay evidencia para no rechazar la hipótesis nula (la cuarentena no reduce el EPG), por lo tanto, el índice de crecimiento potencial o EPG para los casos positivos aumento al finalizar la cuarentena.

Se apreciaron diferencias significativas de crecimiento del EPG antes, durante y después del aislamiento preventivo obligatorio, pudimos notar que en el último mes de la cuarentena (junio) el ritmo de contagios empezó a disminuir y que al finalizar la cuarentena (julio) nuevamente empezó a subir la curva de contagios, esto demuestra que la cuarentena ayuda a disminuir el ritmo de crecimiento potencial pero en nuestro caso se tardó más de dos meses en notarse los resultados. Esto no significa que la cuarentena no haya sido una buena medida de control tomada por el estado peruano pero es muy probable que no se haya cumplido estrictamente.

De igual forma, nos dimos cuenta que las medidas de mitigación como la cuarentena tienen como finalidad reducir la velocidad de transmisión(reducir la velocidad del EPG) no exactamente esperar una disminución de este.

La cuarentena fue analizada a nivel de norma regulatoria que impone el gobierno, mas no a nivel de cumplimiento por persona, señalamos este punto dado que existen factores exógenos como cultura, estado socioeconómico, valores, entre otros, que pueden haber influido en el resultado del análisis. Por ello, es pertinente mencionar que nuestro resultado esta acotado solo la efectividad de la implantación de la cuarentena en una población sin asegurar el cumplimiento de la misma.

## REFERENCIAS

- [1] Akbarialiabad, H., & Ahmadi, M. (2020). Efficacy of Mass Quarantine as Leverage of Health System Governance During COVID-19 Outbreak. *Iranian Medicine*, 266-267.
- [2] Dobrescu, A., Chapman, A., & Persad, E. (2020). B.Quarantine alone or in combination with other public health measures to control COVID-19. *Cochrane Library*, 55-57.