

Estimación del número de reproducción (R) en tiempo real PARA EL PERÚ

MARCO VEGA* Y ERICK LAHURA**

Este artículo resume una metodología de estimación del número de reproducción R del COVID-19, que permite cuantificar la capacidad de transmisión de dicho virus. Para ello, los autores presentan las diversas versiones que se estiman en el BCRP con información en tiempo real.



* Subgerente de Investigación Económica del BCRP
marco.vega@bcrp.gob.pe



** Especialista en Investigación, Subgerencia de Investigación Económica del BCRP
erick.lahura@bcrp.gob.pe

Desde el inicio de la pandemia del COVID-19, el BCRP hace el seguimiento diario y análisis de las estadísticas asociadas a este virus, pues este ha tenido un impacto sin precedentes en la situación económica presente y futura del Perú. Uno de los indicadores más importantes respecto a la evolución de la pandemia es el número de reproducción R del COVID-19, el cual permite cuantificar la capacidad de transmisión del virus. Este número es estimado a nivel mundial y se actualiza permanentemente. En este artículo, se resume una metodología de estimación de este número y se presentan las diversas versiones que se estiman en el BCRP con información en tiempo real.

EL NÚMERO DE REPRODUCCIÓN R

El número de reproducción R mide la capacidad de transmisión de una enfermedad, es decir, qué tan infecciosa es una enfermedad. Usualmente, se analizan dos números de reproducción: R_0 o R básico y R_t o R efectivo. El primero mide la capacidad de transmisión de una enfermedad al inicio de esta, cuando todos los individuos de la población son susceptibles, es decir, mide el grado de contagio natural de una enfermedad. Por ejemplo, si se asume que el R básico del COVID-19 es 3, significa que, en una población donde nadie es inmune y no se toman acciones para controlar el brote de la enfermedad, cada individuo infectado transmite el virus a tres personas en promedio. El R_0 excluye los nuevos casos producidos por los casos secundarios.

El R efectivo, por su parte, permite medir el número promedio de personas que en un momento de tiempo t pueden contagiarse de casos secundarios de contagios. Evidentemente, las ac-

ciones implementadas para evitar la propagación de una enfermedad disminuyen su capacidad de transmisión y, por ende, el R efectivo. En el caso del COVID-19, se espera que el distanciamiento social y la utilización de mascarillas disminuyan la transmisibilidad del virus y, en consecuencia, que el R efectivo disminuya.

Los epidemiólogos sostienen que la única forma de combatir al COVID-19 es entender y gestionar en función al R efectivo. Los valores de R_t tiene diferentes implicancias: (i) si $R > 1$, el número de nuevos casos crece en el tiempo y la enfermedad prolifera; (ii) si $R = 1$, el número de casos nuevos se mantiene constante; y (iii) si $R < 1$, el número de casos nuevos disminuye en el tiempo. Durante la pandemia del COVID-19, los países empezaron a reducir las restricciones de aislamiento cuando R_t registró niveles menores a 1. Si R_t se reduce por debajo de 1 y también se reduce el número de contagios en general, el virus se vuelve manejable.

METODOLOGÍA DE ESTIMACIÓN

Para estimar el número de reproducción R_t en tiempo real, se utiliza un modelo de espacio de estados similar a Marioli y otros (2020), pero más cercano a Vladeck (2020). La diferencia entre ambos modelos es que este último se basa en un modelo Poisson para la ocurrencia de nuevos casos:

$$\begin{aligned} \text{Ecuación de observación: } I_{t+1} &\sim \text{Poisson}(I_t e^{\theta_t}) \\ \text{Ecuación de estado: } \theta_t &\sim N(\theta_{t-1}, \sigma) \end{aligned}$$

donde t representa días. I_t es el número total de personas contagiadas en los últimos 20 días¹, es decir, es un indicador suavizado del ritmo de contagios. La variable θ_t es la tasa esperada de crecimiento en el ritmo de los contagios. Esta variable puede ser expresada como $\theta_t = \gamma(R_t - 1)$, donde γ es la inversa del número de días que demora la aparición de síntomas en una persona infectada (intervalo serial) y R_t es el número de reproducción efectivo. En este modelo, θ_t es la variable de estado no observable, pues depende de R_t que no es observable:

$$R_t = \frac{\theta_t}{\gamma} + 1$$

El modelo de espacio de estados se estima utilizando el filtro de Kalman. De esta forma, R_t se obtiene a partir de la estimación de $\theta_t = \gamma(R_t - 1)$.

Se puede apreciar que si el ritmo de contagios crece $\theta_t > 0$, entonces $R_t > 1$. Cuando el ritmo de contagios deja de crecer, se tiene $\theta_t = 0$ y, por tanto, $R_t = 1$, mientras que si el ritmo de contagio está en franco descenso $\theta_t < 0$, se obtiene $R_t < 1$.



Uno de los indicadores más importantes respecto a la evolución de la pandemia es el número de reproducción R del COVID-19, el cual permite cuantificar la capacidad de transmisión del virus.



¹ Si I_t fuera el número acumulado de personas contagiadas en t , entonces R nunca sería menor a 1. Para permitir la posibilidad de que $R < 1$, se asume que I_t representa el número total de personas contagiadas en los últimos N días. En este caso, $N=20$.

CUADRO 1 ■ Versiones de nuevos infectados diarios utilizados para las estimaciones de versiones correspondientes del número R

Versión	Definición	Fuente
I1	Número oficial de nuevos infectados que fueron identificados con cualquier tipo de prueba.	Sala Situacional MINSA
I1*	Número oficial de nuevos infectados que fueron identificados solo con pruebas PCR.	Datos Abiertos MINSA
I2	Número de nuevos infectados estimados a partir de las cifras oficiales de muertes.	Sala Situacional MINSA
I3	Número de nuevos infectados estimados a partir de los estimados de excesos de muertes.	SINADEF
I3*	Número de nuevos infectados estimados a partir de los estimados de excesos de muertes ajustados por corrección de sesgo de fin de muestra.	SINADEF

1/ PCR HACE REFERENCIA A LAS SIGLAS EN INGLÉS DE "REACCIÓN EN CADENA DE LA POLIMERASA". LA PRUEBA PCR ES UNA PRUEBA DE DIAGNÓSTICO QUE PERMITE DETECTAR UN FRAGMENTO DEL MATERIAL GENÉTICO DE UN PATÓGENO O MICROORGANISMO.

En el caso del COVID-19, se estima que el intervalo serial se encuentra entre 4 y 8 días. Las estimaciones que se presentan asumen que el intervalo serial es 6 días², de acuerdo con Nishiura, Linton y Akhmetzhanov (2020), es decir, $1/\gamma = 6$.

ESTIMADOS DE R

El insumo principal en la estimación de R_t es el número de nuevos infectados, a partir del cual se calcula I_t . Sin embargo, esta cifra no solo refleja la transmisibilidad del virus, sino también la intensidad en la realización de pruebas aplicadas para detectar infectados. Además, las cifras oficiales de infectados sólo se computan a partir del día de detección y no a partir del inicio de síntomas. Es decir, la variable I_t observada tiene errores de medición y debe usarse con cautela en la medición de R .

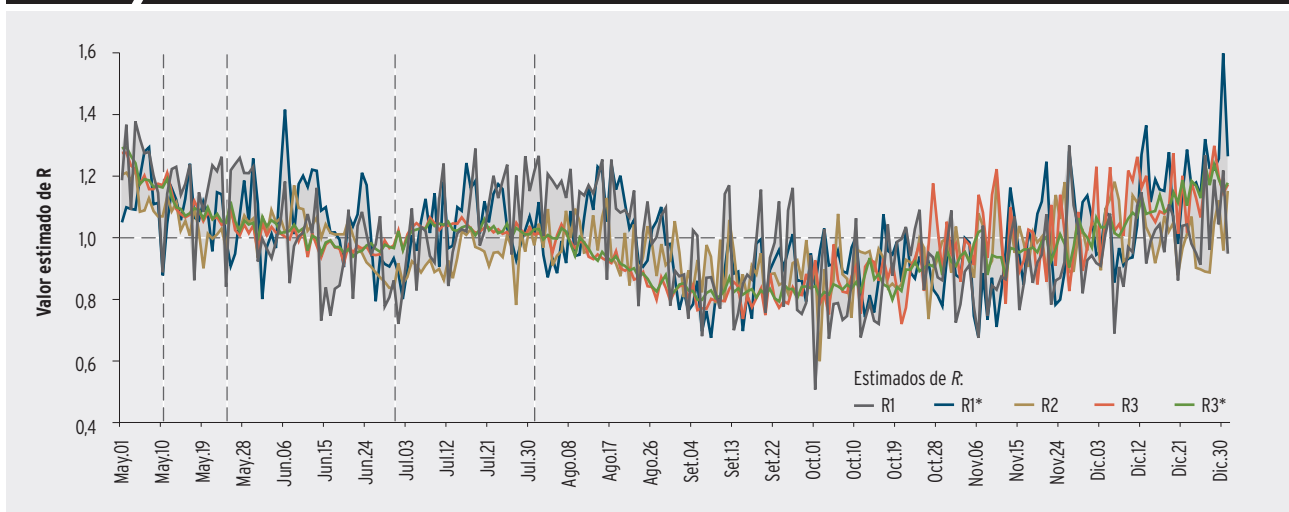
Tomando en cuenta dicho problema de medición, se estiman diferentes versiones de R que se diferencian en la aproximación de la cifra de nuevos infectados, de acuerdo con el Cuadro 1.

Las versiones I2, I3 e I3* son aproximaciones de contagios estimados a partir de cifras de muertes. La versión I2 corresponde a la cifra oficial de

muertes publicada diariamente en la Sala Situacional COVID-19. En cambio, I3 e I3* corresponden a las cifras de fallecimientos diarios con certificados de defunción que atribuyen COVID-19, coronavirus o SARS-CoV-2 como posible causa de muerte. Esta última información se obtiene del Sistema Informático Nacional de Defunciones (SINADEF), que publica estas estadísticas conforme va llegando información al sistema y, por ello, son posibles de corregirse hacia atrás.

Las cifras de muertes del día anterior suelen ser relativamente bajas respecto a los demás días porque aún no llega toda la información de fallecidos del día anterior. Con el paso de los días, la información se estabiliza y refleja la verdadera cifra de ese día. Por ello, las cifras en tiempo real tienen un sesgo hacia abajo por publicación temprana. Para controlarlo, nuestro cálculo I3 corta la serie de fallecidos hasta unos días atrás para obviar las cifras sesgadas, mientras que I3* considera una corrección estadística del sesgo sobre la base de las correcciones históricas que se tienen que hacer con las cifras hasta que se estabilicen.

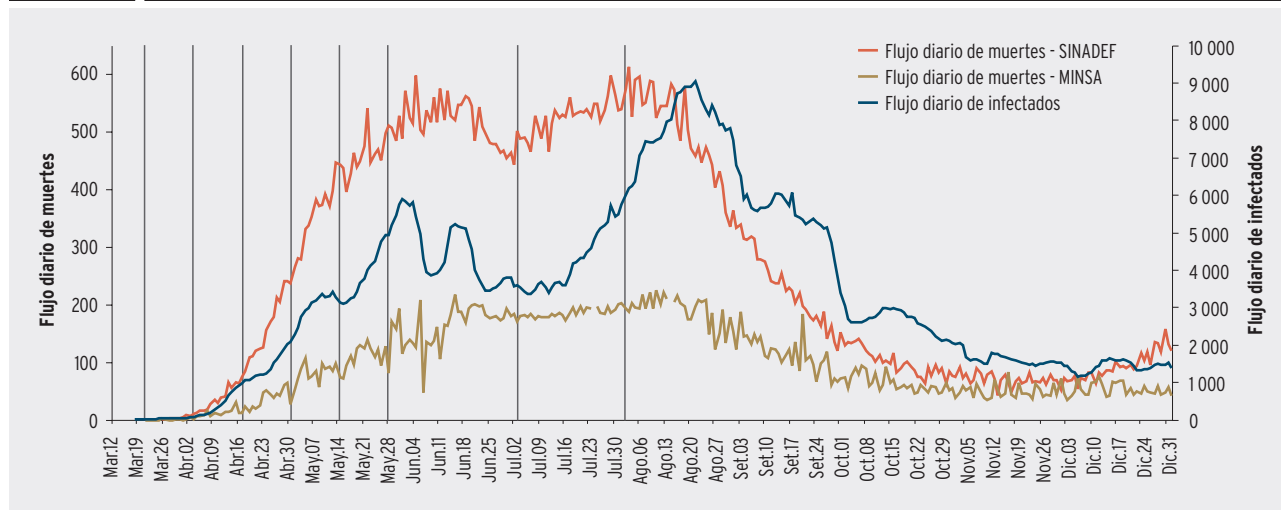
GRÁFICO 1 ■ Estimados del número de reproducción R



FUENTE: MINSA Y SINADEF.

² Nishiura, Linton y Akhmetzhanov (2020) estiman que es 4,6 días con un intervalo de confianza al 95 % entre 3,5 y 5,9.

GRÁFICO 2 ■ Evolución del flujo diario de infectados y muertes por COVID-19



NOTA: EL FLUJO DE INFECTADOS ES UN PROMEDIO MÓVIL DE LOS NUEVOS CASOS DETECTADOS EN LOS ÚLTIMOS 7 DÍAS.
FUENTE: MINSa Y SINADEF.

Para estimar el número de infectados a partir del número de muertes, se asume que (i) entre la infección y la muerte de una persona (si sucede) transcurren aproximadamente $H=7$ días y (ii) la tasa de fatalidad de los casos de COVID-19 es $\phi = 3,4$ (según la Organización Mundial de la Salud). Bajo estos supuestos, el estimado del número de infectados el día t , I_t , se obtiene según la siguiente fórmula:

$$I_t = \frac{M_{t+H}}{\phi}$$

Dado que la estimación I_t está en función del adelanto de las muertes, la estimación de los 7 días más recientes de I_t requiere predecir el número de muertes hasta $H=7$ días en adelante. Entre ambos supuestos sobre H y sobre ϕ , el más relevante para el cálculo de R_t resulta ser el número de días que pasan entre la detección del virus y la muerte (en caso ocurra).

En total, se estiman cinco R alternativos: R_1 basado en I_1 , R_1^* basado en I_1^* , R_2 basado en I_2 , R_3 basado en I_3 y R_3^* basado en I_3^* . El Gráfico 1 (pág. 41) muestra la evolución de los diferentes estimados de R efectivo basados en la información disponible al 31 de diciembre del 2020 (fecha de corte de este artículo). Las líneas verticales punteadas indican las fechas en las que se extendió el período de aislamiento social obligatorio. Es importante mencionar que, en la práctica, la estimación de cada R alternativo se realiza en tiempo real tomando en cuenta la información diaria más actualizada sobre los nuevos casos de infectados

y muertes, lo cual puede generar algunos errores en la estimación.

Se observa que, entre fines de agosto y fines de octubre, los diferentes estimados del número de reproducción se ubicaron por debajo de 1, es decir, el contagio se encontraba en una fase decreciente. Sin embargo, desde noviembre 2020, se registra una tendencia creciente del número de reproducción.

Este resultado es consistente con la evolución del número de infectados y fallecidos reportados al 31 de diciembre del 2020. El ritmo de contagios detectados se desaceleró desde fines de agosto. Como se observa en el Gráfico 2, el flujo diario de infectados (promedio móvil de los últimos 7 días) se redujo de casi 9 000 a finales de agosto a menos de 2 000 a inicios de noviembre. Durante el mismo período, el flujo de muertes se redujo aproximadamente de 600 a 60 casos según SINADEF y de 300 a 40 casos según MINSa. Sin embargo, desde inicios de diciembre se observa una tendencia creciente en el flujo de muertes según SINADEF.

CONCLUSIÓN

Los diferentes estimados del número de reproducción R del COVID-19 han mostrado una dinámica similar desde el inicio de la pandemia, lo que aporta a la confiabilidad de su uso como herramientas para guiar las decisiones de política económica y de salud. Con información al 4 de enero, los estimados de R muestran una tendencia creciente en la transmisión del COVID-19.

REFERENCIAS:

- Marioli, F. A., Bullano, F., Kučinskis, S., y Rondón-Moreno, C. (2020). Tracking R of COVID-19: A New Real-Time Estimation Using the Kalman Filter. medRxiv. DOI: 10.1101/2020.04.19.20071886.
- Nishiura, H., Linton, N. M., y Akhmetzhanov, A. R. (2020). Serial interval of novel coronavirus (COVID-19) infections. *International journal of infectious diseases*, 93 (2020), 284-286. DOI: 10.1016/j.ijid.2020.02.060.
- Vladeck, T. (2020). *Using State Space Models to Estimate Rt*. Extraído de la página web: <https://tomvladeck.com/2020/04/15/rt-ssm.html>.