



Informe sobre la evolución de la epidemia de covid-19 en Chile

Camila Arroyo, Eduardo Engel, Diego Pardow y Pablo Simonetti

5 de junio de 2020

Introducción

Queremos entregar las herramientas para poder replicar la estimación de muertes en exceso realizada en el [informe](#) del 3 de mayo. Para esto se pondrá a disposición 6 archivos disponibles en el siguiente [link](#).

1. Este manual que permitirá replicar las estimaciones realizadas paso a paso.
2. ‘Muertes en exceso.do’: Archivo dofile del programa estadístico STATA que permitirá replicar las estimaciones, y que tendrá como output un archivo excel con los datos mostrados en el informe.
3. Documento .dta (STATA) denominado ‘fecha auxiliar’ que permitirá crear años de 52 semanas de 7 días cada una.
4. Base de datos utilizada descargada el 2 de junio. Cabe destacar que en el archivo dofile vendrá una forma de poder descargar las bases de datos directamente del github del Ministerio de Ciencias, sin embargo, las bases de defunciones del Registro Civil presentan cambios día a día, por ende aunque se comparen las mismas fechas, pueden observarse cambios producto de esta actualización. La base de datos utilizada en el informe del 3 de junio fue descargada el día 2 de junio, y se pondrá a la disposición para que los resultados obtenidos sean equivalentes a los presentados en el informe.
5. Excel denominado ‘Estimaciones’, en este excel se exportarán los resultados una vez se haya hecho la estimación. Antes de hacer la estimación el excel se encuentra vacío

en los campos de los datos de interés, pero contiene los nombres de las columnas y filas.

6. 2 documentos en R que permiten estimar los modelos de series de tiempo (modelos ARIMA), puestos a disposición por Juan Díaz, académico de la Universidad de Chile.

A continuación se presentarán los pasos para poder replicar las estimaciones:

Pasos a seguir:

1. Toda la información utilizada proviene de las bases de datos disponibles en la página del Ministerio de Ciencias <http://www.minciencia.gob.cl/covid19>.
2. Lo primero que se debe hacer, es ir a las bases de datos que se utilizarán, en este caso: (i) [Defunciones en Chile](#), (ii) [Casos totales por región incremental](#), y (iii) [fallecidos por región incremental](#). Es importante que el dofile puesto a disposición carga directamente estas bases de datos en la web al programa estadístico STATA (se puede ver en el dofile en los puntos 1,2 y 7), pero este paso sirve para identificarlas y tener un punto de referencia.

Lo que se explica a continuación se realiza en el dofile ‘Muertes en exceso.do’

3. Para poder correr el dofile en el programa STATA se debe elegir un directorio donde se guarden las bases de datos. Al principio del dofile hay un ejemplo que se debe cambiar por el directorio de la carpeta escogida. Además se deben dejar los documentos puestos a disposición por Espacio Público en esa carpeta.
4. Una vez cargada la base de datos, se procede a ordenar la información de contagios y fallecidos. En concreto se ponen los números correspondientes a cada región, se cambia la disposición de la base de datos, y se ponen las fechas correspondientes a los datos.
5. Se descarga descarga de la misma forma la base de datos de defunciones diarias (desagregada por comuna), que está disponible desde el 2010. Dado el tamaño de la base de datos esto puede tomar varios minutos.
6. Se unen todas las bases de datos. Como se mencionó la base descargada hoy (4 de junio) será distinta a la descargada el 2 de junio.
7. Se crea una base de datos con años ficticios con 364 días. Se eliminan los 29 de febrero y un día extra en todos los años. Así se dispone de una base de datos con años ficticios de 364 días, comparables entre sí (ítem 11 del dofile).

Para las estimaciones a continuación se utilizará la base descargada el 2 de junio, que hemos puesto a disposición.

8. Se une la serie final del 2 de junio con la base de semanas creada anteriormente. Se hace una base que contiene información por año, semana y región (punto 11.4 del dofile), esta es la base que se utilizará para la regresión.
9. Se crean variables de interés para la regresión, y se realiza la estimación del siguiente modelo para el periodo previo a covid (antes de la semana 45 del 2020, es decir, hasta el 1 de abril):

$$\log(fallecidos)_s = \alpha + \beta_s + \gamma((T - 1) * 52 + s)$$

Donde α corresponde a la constante del modelo, β_s a efectos fijos por semana, y γ a un parámetro que acompaña a una variable que mide tendencia en la serie de tiempo. Los resultados de la estimación permiten proyectar los fallecidos para las semanas donde no se realizó la estimación (del 2 de abril al 27 de mayo). Posteriormente se comparan los fallecidos reales con la exponencial de los fallecidos proyectados (que estaban en logaritmo), y se obtiene lo que se denomina muertes en exceso.

10. Se arma una base nacional, por semanas y se realiza la misma estimación.
11. Los resultados de muertes en exceso por semana y región se exportan al Excel ‘Estimaciones’, donde cada hoja corresponde a una región. En la hoja ‘pais’ se exportan los resultados del modelo agregado para todo el país. Luego, en la hoja ‘Consolidado’ se presentan los resultados consolidados.