

ACTUALIZACIÓN: ESTIMACIÓN DE PUNTOS DE CAMBIO USANDO REGRESIÓN SEGMENTADA PARA DATOS DEL BROTE DE COVID-19 EN CHILE

FELIPE OSORIO

RESUMEN. En el presente reporte se utiliza la metodología de regresión segmentada para datos en la familia exponencial propuesto por Muggeo [Stat. Med. 22: 3055-3071, 2003]. El objetivo ha sido actualizar el reporte realizado el día 15 de Abril con la información más actual suministrada por el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. El principal hallazgo en este reporte corresponde a que los *tiempos requeridos para la duplicación* del número de casos de COVID-19 *se han reducido* para las regiones Metropolitana, Coquimbo, Valparaíso, O'Higgins y Biobío, así como a nivel nacional.

1. INTRODUCCIÓN

Para evaluar las estrategias de cuarentena en casos de COVID-19 en Chile se desarrolló la estimación de puntos de cambio usando el procedimiento de regresión segmentada propuesto por Muggeo (2003). Como en Muggeo et al. (2020), se asume un modelo de crecimiento exponencial para el número acumulado de casos infectados desde que ha iniciado el brote de COVID-19. El modelo puede ser usado para determinar si las estrategias de confinamiento llevan a desacelerar la tendencia observada. Para fijar ideas, se consideró el modelo

$$\log E(Y_t) = \beta_0 + \beta_1 t + \delta_1(t - \psi_1)_+ + \dots + \delta_K(t - \psi_K)_+, \quad (1)$$

donde Y_t , representa el número de acumulado de casos infectados para el día $t = 1, 2, \dots$, ψ_1, \dots, ψ_K denota los puntos de cambio, mientras que las pendientes asociadas a cada uno de los regímenes son $\beta_1, \beta_2 = \beta_1 + \delta_1, \dots, \beta_{K+1} = \beta_1 + \sum_{r=1}^K \delta_r$. Debemos destacar que los *tiempos de duplicación* se pueden obtener desde el modelo (1) como la siguiente reparametrización $\phi_r = \log(2)/\beta_r$ para cada régimen $r = 1, 2, \dots, K + 1$. Se llevó a cabo el ajuste del modelo de regresión para datos Poisson y la selección de los puntos de cambio se realiza basado en el criterio de información de Schwarz (SIC). La metodología se encuentra implementada en el paquete para R, *segmented* (Muggeo, 2008). Adicionalmente, en Muggeo et al. (2020) se presenta un conjunto de rutinas para simplificar el proceso de estimación.

2. RESULTADOS PARA LOS DATOS DEL BROTE DE COVID-19 EN CHILE

Los datos disponibles corresponden al número de contagios acumulado desde el inicio del brote de COVID-19 para cada una de las regiones en Chile, descargados el día 15 de Mayo de 2020 desde la página web: <https://github.com/MinCiencia/Datos-COVID19/tree/master/output/producto3>.

A continuación se presentan los resultados de estimación y gráficos del ajuste del modelo de regresión segmentada para cada una de las regiones de Chile. Un aspecto interesante es llevar a cabo una comparación con el análisis del regresión segmentada con fecha 28 de Abril. En aquél informe podíamos apreciar como los

Fecha: 16 de Mayo, 2020.

Reporte del 15 de Abril, disponible en: <http://matematica.usm.cl/covid-19-en-chile/>.

tiempos de duplicación, entendido como el número de días necesario para duplicar el número de casos que presentan la enfermedad, habían incrementado. Mientras que, desde la Tabla 1 podemos notar que tanto a nivel nacional como en las regiones de Coquimbo, Valparaíso, O'Higgins y Biobío los tiempos de duplicación se han reducido, hecho que ocurre en torno del final del mes de Abril inicios de Mayo (ver Tabla 2). Adicionalmente, se presenta la estimación e intervalos de confianza para los tiempos de duplicación en algunas regiones seleccionadas. Es notable la aceleración en la tendencia (con la consecuente disminución en los tiempos de duplicación) que se ha producido en las regiones de Coquimbo (desde el 3 de Mayo) y de O'Higgins (a partir del 20 de Abril) (consulte también las Tablas 3 a 8 así como las Figuras 1, 2 y 3). Por otro lado, debemos destacar la regresión ajustada para la región del Maule (ver Tabla 1 y Figura 4(e) en el Anexo), donde se aprecia una fuerte desaceleración de la curva estimada.

El código en R para replicar los resultados presentados así como este reporte se encuentran disponibles en: <https://github.com/faosorios/COVID-19/tree/master/segmented>

AGRADECIMIENTOS

El autor agradece a Alonso Ogueda por proveer de un script en Jupyter notebook para ejecutar interactivamente (y actualizar) el reporte presentado en este documento.

REFERENCIAS

- Muggeo, V.M.R. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine* **22**, 3055-3071.
- Muggeo, V.M.R. (2008). Segmented: an R package to fit regression models with broken-line relationships. *Rnews* **8**, 20-25.
- Muggeo, V.M.R., Sottile, G., Porcu, M. (2020). Modelling COVID-19 outbreak: segmented regression to assess lockdown effectiveness. doi: [10.13140/RG.2.2.32798.28485](https://doi.org/10.13140/RG.2.2.32798.28485)

TABLA 1. Estimación de los tiempos de duplicación para los casos de COVID-19 en Chile.

Región	# de regímenes	Tiempos de duplicación								
		2.65	3.77	6.30	9.39	14.14	17.31	9.95	12.96	9.58
Chile	9	2.65	3.77	6.30	9.39	14.14	17.31	9.95	12.96	9.58
Santiago	7	3.99	8.60	10.78	14.57	8.04	11.09	8.28		
Arica	3	2.62	11.02	39.19						
Tarapaca	3	10.86	5.94	9.53						
Antofagasta	2	8.86	14.03							
Atacama	2	7.29	34.41							
Coquimbo	4	0.02	4.60	56.32	11.17					
Valparaíso	3	5.37	16.76	10.73						
O'Higgins	5	0.02	5.68	26.91	12.32	6.69				
Maule	3	18.52	2.79	37.03						
Ñuble	3	5.73	39.01	83.89						
Biobío	4	6.09	21.49	56.59	23.81					
Araucanía	4	3.62	7.44	18.80	55.65					
Los Ríos	3	1.93	10.34	55.00						
Los Lagos	3	4.41	13.51	44.34						
Aysén	2	5.10	191.28							
Magallanes	3	5.97	14.98	34.83						

TABLA 2. Estimación de los puntos de cambio para los casos de COVID-19 en Chile.

Región	# de regímenes	Puntos de cambios (fechas)							
Chile	9	03-19	03-27	04-02	04-09	04-17	04-28	05-02	05-11
Santiago	7	03-26	04-10	04-15	04-27	05-03	05-10		
Arica	3	04-10	04-29						
Tarapaca	3	05-01	05-07						
Antofagasta	2	05-03							
Atacama	2	05-12							
Coquimbo	4	03-20	04-07	05-03					
Valparaíso	3	04-03	05-05						
O'Higgins	5	03-20	04-05	04-20	05-11				
Maule	3	04-15	04-18						
Ñuble	3	04-06	04-20						
Biobío	4	04-05	04-20	05-05					
Araucanía	4	03-31	04-07	04-23					
Los Ríos	3	03-31	04-13						
Los Lagos	3	04-03	04-11						
Aysén	2	04-06							
Magallanes	3	04-11	04-21						

TABLA 3. Estimación y intervalos de confianza para el tiempos de duplicación datos de COVID-19 en Chile.

Segmento	Estimación	Intervalo de confianza (95 %)	
1	2.6491	2.2961	3.1303
2	3.7726	3.5798	3.9873
3	6.2996	5.8218	6.8622
4	9.3911	8.7610	10.1188
5	14.1375	13.0032	15.4886
6	17.3131	16.6726	18.0043
7	9.9479	9.0660	11.0197
8	12.9650	12.5830	13.3712
9	9.5807	8.6892	10.6763

TABLA 4. Estimación y intervalos de confianza para el tiempos de duplicación: datos de COVID-19 en Santiago.

Segmento	Estimación	Intervalo de confianza (95 %)	
1	3.9923	3.7952	4.2111
2	8.6005	8.2793	8.9475
3	10.7760	9.6321	12.2283
4	14.5745	13.9534	15.2534
5	8.0428	7.6365	8.4949
6	11.0936	10.6095	11.6239
7	8.2762	7.7582	8.8683

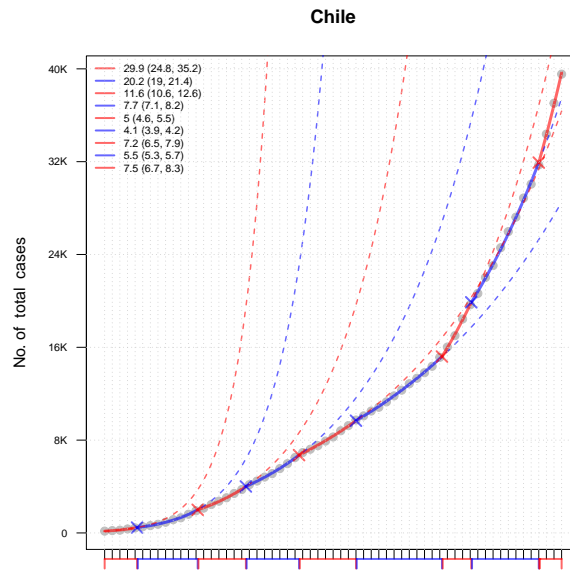


FIGURA 1. Regresión segmentada y puntos de cambio para datos de COVID-19 a nivel nacional.

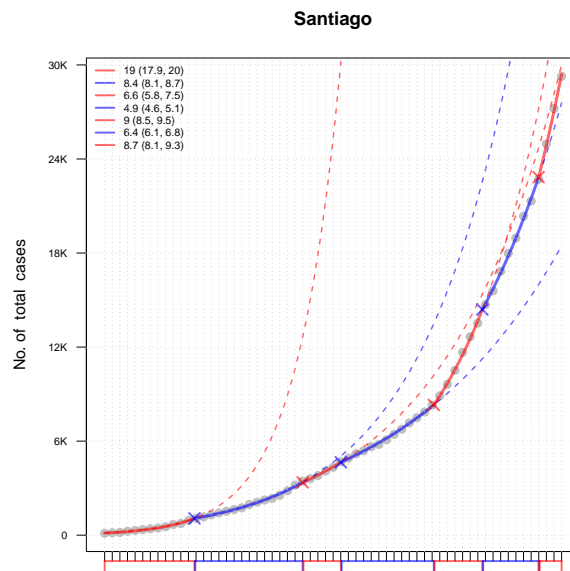


FIGURA 2. Regresión segmentada y puntos de cambio para datos de COVID-19 en Región Metropolitana

TABLA 5. Estimación y intervalos de confianza para el tiempos de duplicación: datos de COVID-19 en Coquimbo.

Segmento	Estimación	Intervalo de confianza (95 %)	
1	0.0221	0.0001	0.0286
2	4.5995	3.9363	5.5319
3	56.3260	37.8521	110.0129
4	11.1742	9.1872	14.2582

TABLA 6. Estimación y intervalos de confianza para el tiempos de duplicación: datos de COVID-19 en Valparaíso.

Segmento	Estimación	Intervalo de confianza (95 %)	
1	5.3732	3.7689	9.3562
2	16.7589	15.9697	17.6297
3	10.7322	9.6579	12.0751

TABLA 7. Estimación y intervalos de confianza para el tiempos de duplicación: datos de COVID-19 en O'Higgins.

Segmento	Estimación	Intervalo de confianza (95 %)	
1	0.0222	0.0001	0.0286
2	5.6760	4.5159	7.6389
3	26.9110	16.0250	83.9263
4	12.3204	10.9167	14.1380
5	6.6925	5.1704	9.4854

TABLA 8. Estimación y intervalos de confianza para el tiempos de duplicación: datos de COVID-19 en Biobío.

Segmento	Estimación	Intervalo de confianza (95 %)	
1	6.0947	5.4613	6.8942
2	21.4936	18.4387	25.7618
3	56.5927	41.3400	89.6814
4	23.8089	18.9971	31.8835

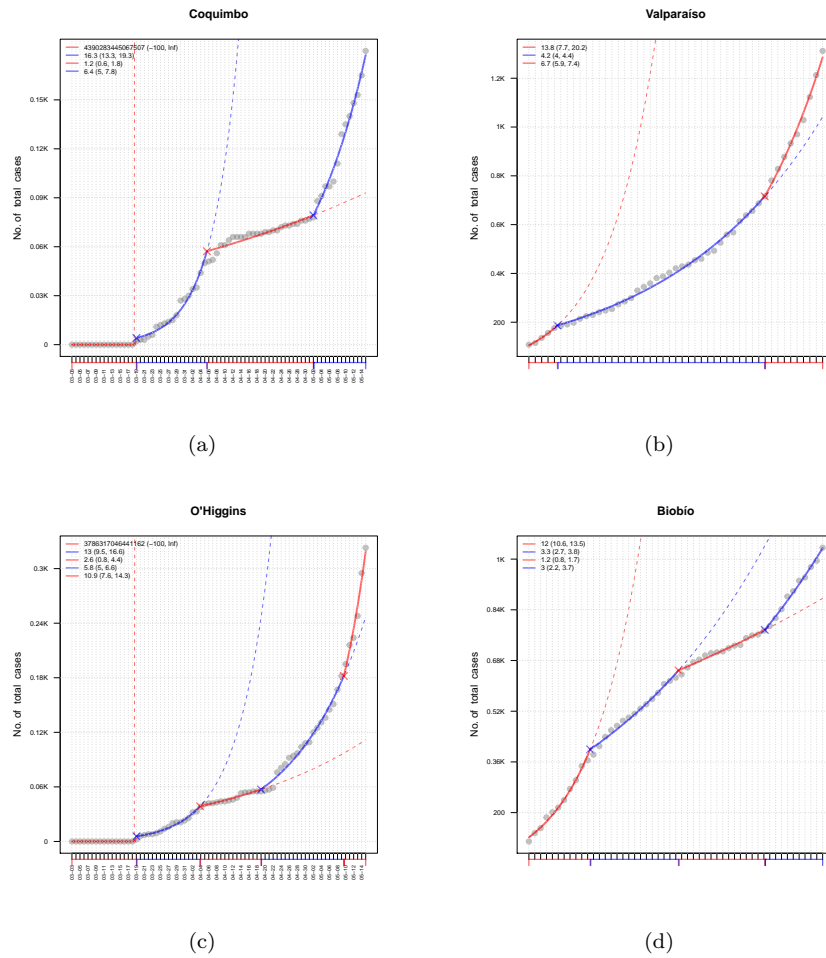


FIGURA 3. Regresión segmentada y puntos de cambio para regiones de Coquimbo, Valparaíso, O'Higgins y del Biobío.

ANEXOS

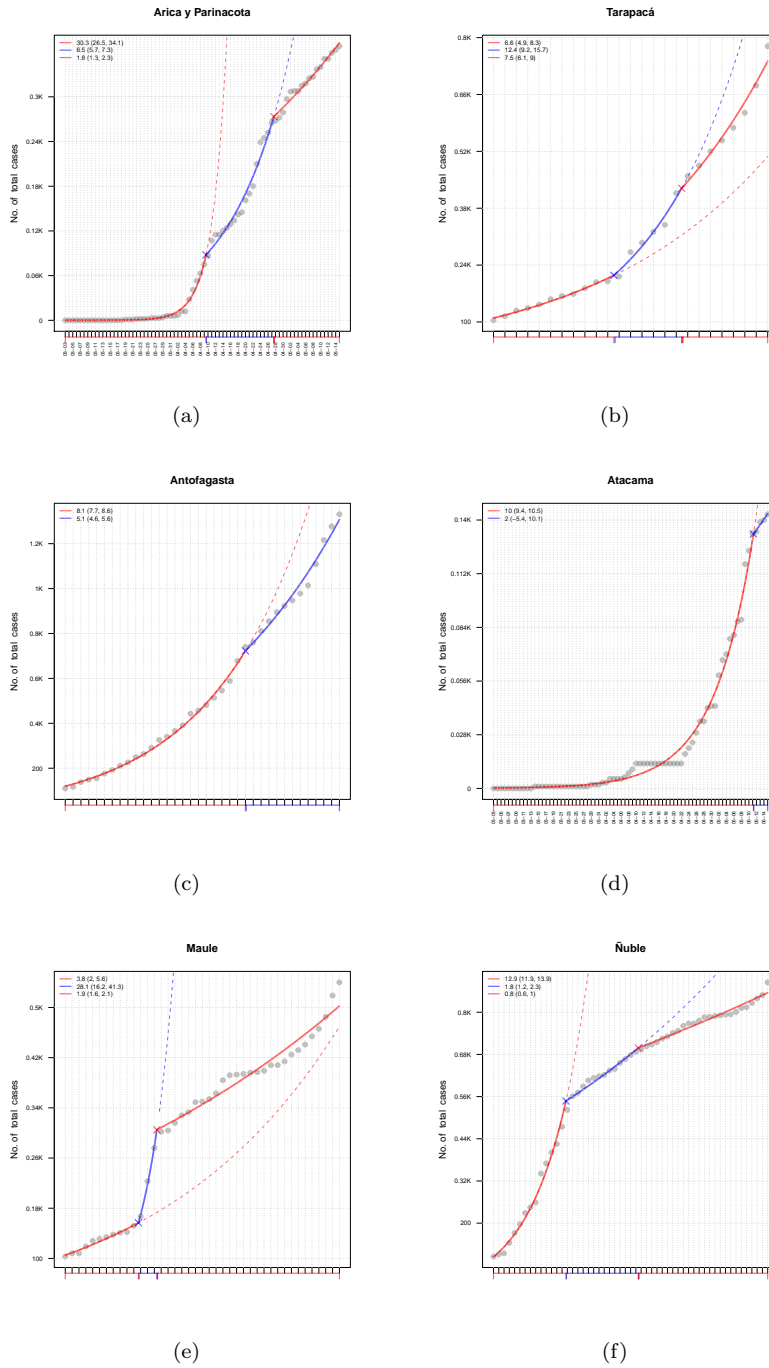


FIGURA 4. Regresión segmentada y puntos de cambio para regiones de Arica a Atacama, el Maule y Ñuble.

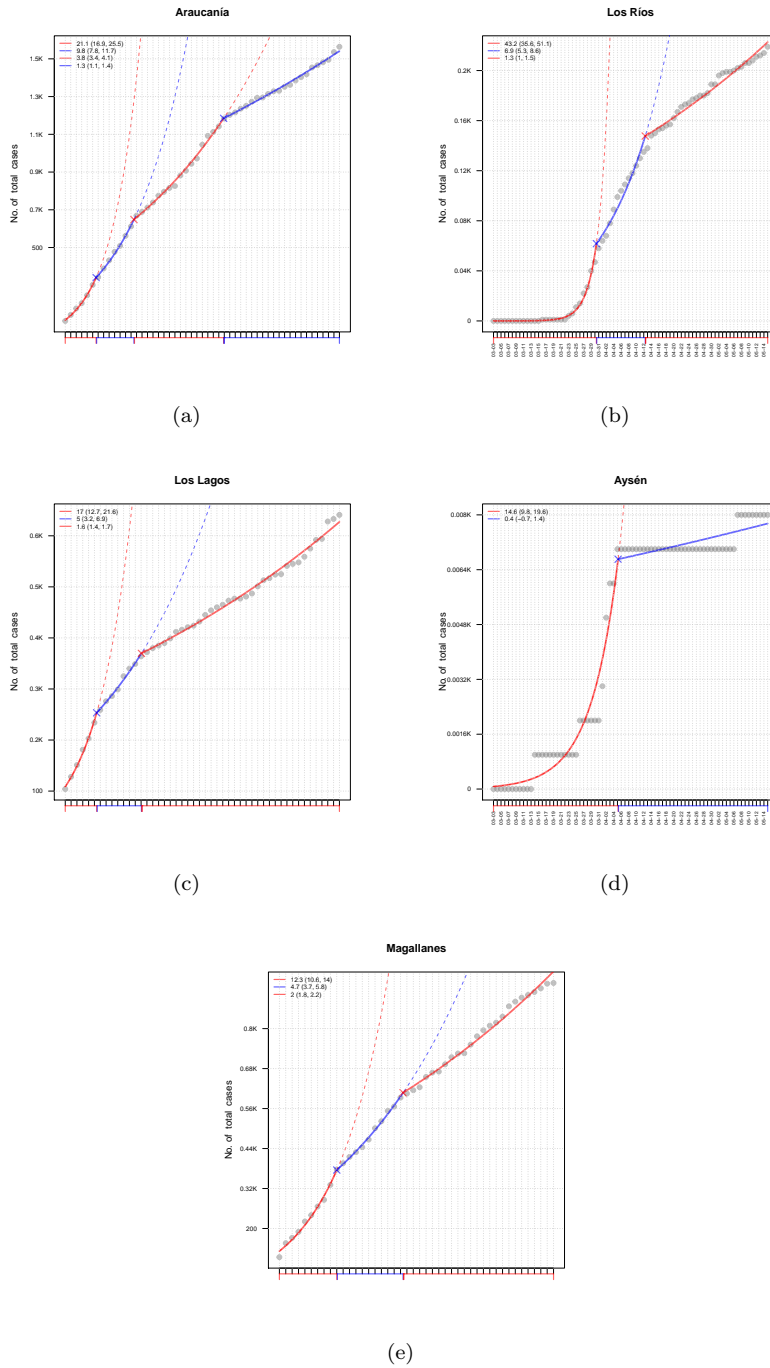


FIGURA 5. Regresión segmentada y puntos de cambio para regiones de La Araucanía a Magallanes.

DEPARTAMENTO DE MATEMÁTICA, UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA, CHILE
Current address: Departamento de Matemática, Universidad Técnica Federico Santa María,
Av. España 1680, Casilla 110-V, Valparaíso, Chile
Email address: felipe.osorios@usm.cl