

Estudio de la transmisión de COVID-19 en Bolivia a través del proceso de Galton Watson

Study of the transmission of COVID-19 in Bolivia through the Galton Watson process

Dindo Valdez Blanco¹

Instituto de Estadística Teórica y Aplicada, Universidad Mayor de San Andrés, La Paz-Bolivia

✉ dvaldez@fcpn.edu.bo

Artículo recibido: 2022-02-22

Artículo aceptado: 2022-03-22

Resumen

En la actualidad la pandemia del COVID-19 ha sido tema de estudio desde muchos enfoques, uno de ellos es el análisis de la evolución del virus y los casos nuevos reportados, específicamente para la determinación del promedio de la transmisión de la enfermedad y la probabilidad de extinción del mismo, la metodología del presente trabajo analiza el comportamiento del virus en Bolivia desde su aparición en el mes de abril 2020 hasta el 30 de enero del 2022, como un proceso estocástico de Galton-Watson para determinar el promedio de reproducción en cada mes y la probabilidad de extinción del COVID-19 para cada mes, de manera que se obtenga algún patrón o comportamiento sistemático que sea de utilidad para el monitoreo del virus.

Palabras clave: Procesos de ramificación, promedio de reproducción, probabilidad de extinción, procesos estocásticos

Abstract

Currently, the COVID-19 pandemic has been the subject of study from many approaches, one of them is the analysis of the evolution of the virus and the new cases reported, specifically for the determination of the average transmission of the disease and the probability extinction of the same, the methodology of this work analyzes the behavior of the virus in Bolivia from its appearance in the month of April 2020 to January 30, 2022 as a Galton-Watson stochastic process to determine the average reproduction in each month and the probability of extinction of COVID-19 for each month, so that some pattern or systematic behavior is obtained that is useful for monitoring the virus.

Keywords: Branching processes, reproduction rate, extinction probability, stochastic processes

1. Introducción

Los procesos de ramificación son una clase de procesos estocásticos que modelan el crecimiento de una población. Se utilizan ampliamente en biología y epidemiología para estudiar la propagación de enfermedades infecciosas y epidemias. En 1873, Sir Francois Galton se plantea una pregunta, respecto a los apellidos paternos ¿qué proporción de sus apellidos paternos se habrán extinguido después de r generaciones?. Su motivación original era estudiar la extinción de los apellidos familiares, un tema que preocupaba a la aristocracia victoriana en el siglo XIX. El reverendo Henry William Watson

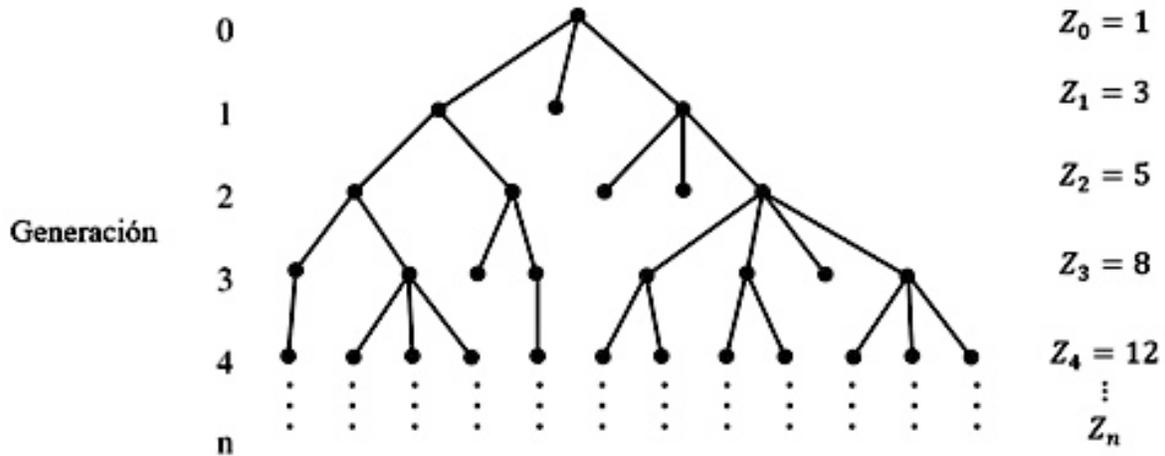
¹ Maestría en Ciencias Estadísticas, Licenciado en Estadística. Profesor de Estadística - Universidad Mayor de San Andrés. ORCID: 0000-0003-0704-0980

respondió con una solución. El estudio de la ramificación. Los procesos surgieron de la colaboración de Watson y Galton. Sus resultados fueron descubiertos independientemente por el estadístico francés Irénée-Jules Bienaymé. Por tal razón el modelo de proceso de ramificación a veces se denomina proceso Bienaymé-Galton-Watson.

La idea general de los procesos de ramificación es la siguiente: en un instante de tiempo se identifica un conjunto de individuos que forman la generación 0, esta generación produce descendientes, cuya colección forma la generación 1. Los descendientes de la generación 1 constituyen la generación 2, y así sucesivamente. Estos procesos modelan la evolución en el tiempo de procesos con estructuras tipo árbol.

De esta manera, si se define el proceso estocástico $\{Z_n\}$ como el número de individuos en la n-ésima generación, entonces la Figura 1 muestra la evolución del proceso de ramificación de Galton-Watson con $Z_0=1$ (Athreya ,1972).

Figura 1.
Esquema de la evolución del proceso de ramificación de Galton-Watson



Fuente: Athreya, 1972

La pandemia del COVID-19 en la actualidad requiere de su estudio y modelación desde muchos aspectos, uno de ellos es el análisis de la propagación del virus, por ello el objetivo del presente trabajo es la aplicación del proceso de Galton-Watson para estimar el promedio de la transmisión de la enfermedad y la probabilidad de extinción del mismo en Bolivia.

2. Metodología

La metodología del trabajo se basa en el procedimiento de un estudio de propagación del COVID-19 en Bolivia a partir del proceso de Galton-Watson, tomando en cuenta los casos reportados desde abril de 2020 hasta enero de 2022, por tanto, es un método de estudio prospectivo. Los datos para el estudio corresponden a los reportes diarios de casos confirmados de COVID-19 en Bolivia realizados por el Ministerio de Salud y que se encuentran disponibles en el repositorio de datos COVID-19, del Centro de Ciencias e Ingeniería de Sistemas (Johns Hopkins University, 2021) situada en Baltimore, Estados Unidos.

3. El proceso de Galton-Watson

Sea $\{Z_n; n \geq 0\}$ una secuencia de variables aleatorias definidas como:

$$Z_n = \sum_{k=1}^{Z_{n-1}} X_{n,k} ; n \geq 1 \quad (1)$$

Donde $\{X_{n,k}; n, k \geq 1\}$ es una familia de variables aleatorias independientes e idénticamente distribuidas con una función de probabilidad $\{p_n; n \geq 0\}$ de tal forma que es independiente de Z_0 . Para cada $n \geq 0$, $X_{n,k}$ representa el número de hijos del i -ésimo individuo en la n -ésima generación, la variable aleatoria Z_n representa el tamaño de la n -ésima generación de una población. Entonces $\{Z_n; n \geq 0\}$ es conocido como el proceso de Galton-Watson.

Por la definición anterior, la distribución de Z_n queda completamente determinada por dos parámetros de entrada, la distribución p_n y Z_0 . De tal forma que $\{Z_n; n \geq 0\}$ constituye una cadena de Markov homogénea con matriz de transición $[p_{ij}]; i, j \geq 0$, dado por:

$$p_{ij} = P[Z_{n+1} = j \mid Z_n = i] = p_j^{*i}$$

Donde p_j^{*i} denota la i -ésima convolución de p_n .

Si se define el evento de extinción de la población como:

$$\text{Extinción} = \{\exists n \in \mathbb{N} \text{ tal que } Z_n = 0\} = \bigcup_{n \in \mathbb{N}} \{Z_n = 0\}$$

Entonces hallar la probabilidad de extinción es:

$$\eta = P(\text{Extinción}) = P(Z_n = 0) ; \text{ para algún } n \in \mathbb{N}$$

De tal forma que:

Si $P(X=0) = 0$ entonces $\eta = 0$.

Si $P(X=0) > 0$ entonces $\eta > 0 \rightarrow \eta \geq P(X=0) > 0$. Indica que el individuo original tiene 0 hijos.

Watson demostró en su artículo original con Galton que, si $P(X=0) > 0$ entonces de hecho $\eta = 1$ para cualquier X . Sin embargo, su documento original tenía un error en vista que es posible tener $\eta \in (0,1)$, dependiendo del valor de $\mu = E[X]$. Se tienen tres casos posibles:

Si $\mu < 1 \rightarrow \eta = 1$, entonces la población siempre se extingue.

Si $\mu > 1 \rightarrow \eta < 1$, entonces la población tiene la probabilidad de sobrevivir.

$$\text{Si } \mu = 1 \rightarrow \begin{cases} \eta = 1 & \text{si } P(X = 1) < 1 \\ \eta = 0 & \text{si } P(X = 1) = 1 \end{cases}$$

Según lo anterior, se puede clasificar a los procesos de Galton-Watson en 3 tipos:

Tipo sub crítico, si $\mu < 1$.

Tipo crítico, si $\mu = 1$.

Tipo supercrítico, si $\mu > 1$.

4. Tamaño medio de la n-ésima generación

La ecuación (1) se representa como una suma de variables aleatorias i.i.d (Athreya, 1972). Estos resultados pueden ser utilizados para hallar los momentos de la variable aleatoria Z_n .

Sea $\mu = \sum_{k=0}^{\infty} k p_k$ la media de la distribución de la descendencia, para encontrar el tamaño medio de la n-ésima generación $E[Z_n]$ condicionado en Z_{n-1} , por la ley de la esperanza total

$$\begin{aligned} E[Z_n] &= \sum_{k=0}^{\infty} E[Z_n | Z_{n-1} = k] P[Z_{n-1} = k] = \sum_{k=0}^{\infty} E \left[\sum_{i=1}^{Z_{n-1}} X_{n,i} | Z_{n-1} = k \right] P[Z_{n-1} = k] \\ &= \sum_{k=0}^{\infty} E \left[\sum_{i=1}^k X_{n,i} | Z_{n-1} = k \right] P[Z_{n-1} = k] = \sum_{k=0}^{\infty} E \left[\sum_{i=1}^k X_{n,i} \right] P[Z_{n-1} = k] \\ &= \sum_{k=0}^{\infty} k \mu P[Z_{n-1} = k] = \mu E[Z_{n-1}] \end{aligned}$$

Luego aplicando la fórmula de recurrencia, se obtiene:

$$E[Z_n] = \mu E[Z_{n-1}] = \mu^2 E[Z_{n-2}] = \dots = \mu^n ; n \geq 0, Z_0 = 1 \quad (2)$$

De esta manera se obtienen los tres casos mencionados anteriormente

$$\lim_{n \rightarrow \infty} E[Z_n] = \lim_{n \rightarrow \infty} \mu^n = \begin{cases} 0 & ; \text{si } \mu < 1 \\ 1 & ; \text{si } \mu = 1 \\ \infty & ; \text{si } \mu > 1 \end{cases} \quad (3)$$

El proceso de ramificación es sub crítico si $\mu < 1$, crítico si $\mu = 1$ y supercrítico si $\mu > 1$.

En el caso que $Z_0 \neq 1$, el número medio de reproducción de la generación n se calcula por la fórmula:

$$E [Z_n] = Z_0 \mu^n ; n \geq 0$$

5. Varianza del tamaño de la n-ésima generación

Para explorar el proceso de extinción en los casos crítico y supercrítico ($\mu \geq 1$), se considera la varianza del tamaño de la enésima generación $V[Z_n]$. Sea σ^2 la varianza de la distribución de la descendencia $X_{n,k}$. Por el teorema de la varianza total,

$$V[Z_n] = V[E(Z_n | Z_{n-1})] + E[V(Z_n | Z_{n-1})]$$

Como ya se demostró en 2,

$$E(Z_n | Z_{n-1} = k) = E\left(\sum_{i=1}^k X_{n,i}\right) = \sum_{i=1}^k E[X_{n,i}] = k$$

Entonces

$$E(Z_n | Z_{n-1}) = \mu * Z_{n-1}$$

De manera similar

$$V(Z_n | Z_{n-1} = k) = V\left(\sum_{i=1}^k X_{n,i}\right) = \sum_{i=1}^k V[X_{n,i}] = k \sigma^2$$

De esta manera se tiene

$$V(Z_n | Z_{n-1}) = \sigma^2 Z_{n-1}$$

Bajo la suposición de independencia de los $X_{n,i}$ y aplicando el teorema de la varianza total (Ospina, 2021), se tiene

$$V[Z_n] = V[\mu Z_{n-1}] + E[\sigma^2 Z_{n-1}] = \mu^2 V[Z_{n-1}] + \sigma^2 \mu^{n-1}; n \geq 1$$

Como $V[Z_0]=0$, la ecuación anterior resulta:

$$V[Z_1] = \mu^2 V[Z_0] + \sigma^2 = \sigma^2$$

$$V[Z_2] = \mu^2 V[Z_1] + \sigma^2 \mu = \sigma^2 \mu (1 + \mu)$$

$$V[Z_3] = \mu^2 V[Z_2] + \sigma^2 \mu^2 = \sigma^2 \mu^2 (1 + \mu + \mu^2)$$

De manera general el patrón anterior provee el resultado:

$$V[Z_n] = \sigma^2 * \mu^{n-1} \sum_{k=0}^{n-1} \mu^k = \begin{cases} n\sigma^2 & ; \text{ si } \mu = 1 \\ \frac{\sigma^2 \mu^{n-1} (\mu^n - 1)}{\mu - 1} & ; \text{ si } \mu \neq 1 \end{cases} \quad (4)$$

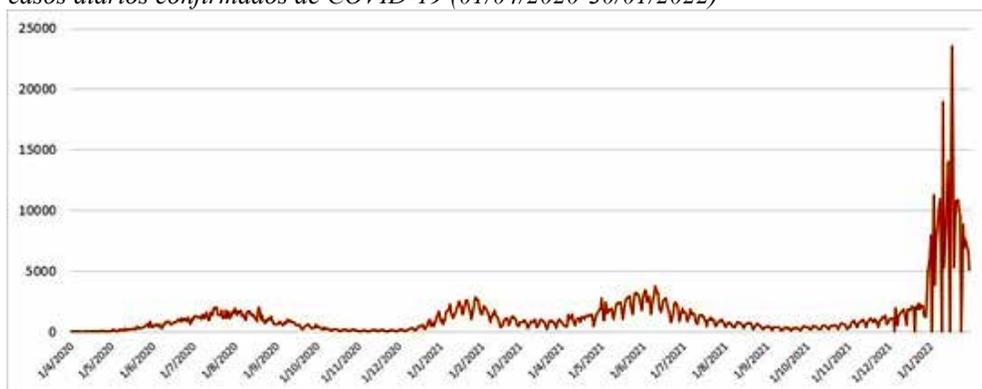
En el caso subcrítico ($\mu < 1$), tanto la media como la varianza del tamaño de la generación tienden a 0. En el caso crítico ($\mu = 1$), el tamaño medio de cada generación es uno, pero la varianza es una función de crecimiento lineal de n.

En el caso supercrítico ($\mu > 1$), la varianza crece exponencialmente. La gran diferencia entre la media μ^n y la varianza sugiere que en algunos casos tanto la extinción y el crecimiento ilimitado son resultados posibles.

6. Análisis de la reproducción y la probabilidad de extinción del COVID-19 en Bolivia

Figura 2.

Bolivia: casos diarios confirmados de COVID-19 (01/04/2020-30/01/2022)



Fuente: Elaboración propia en base a datos del Centro de Ciencias e Ingeniería de Sistemas de la Universidad Johns Hopkins

Aplicando el modelo de Galton-Watson con una distribución de Poisson se estiman el promedio de reproducción y la probabilidad de extinción para cada mes, los resultados se muestran en la Tabla 1.

Tabla 1:

Bolivia: Promedio estimado de reproducción del COVID-19 y probabilidad estimada de extinción del virus por mes

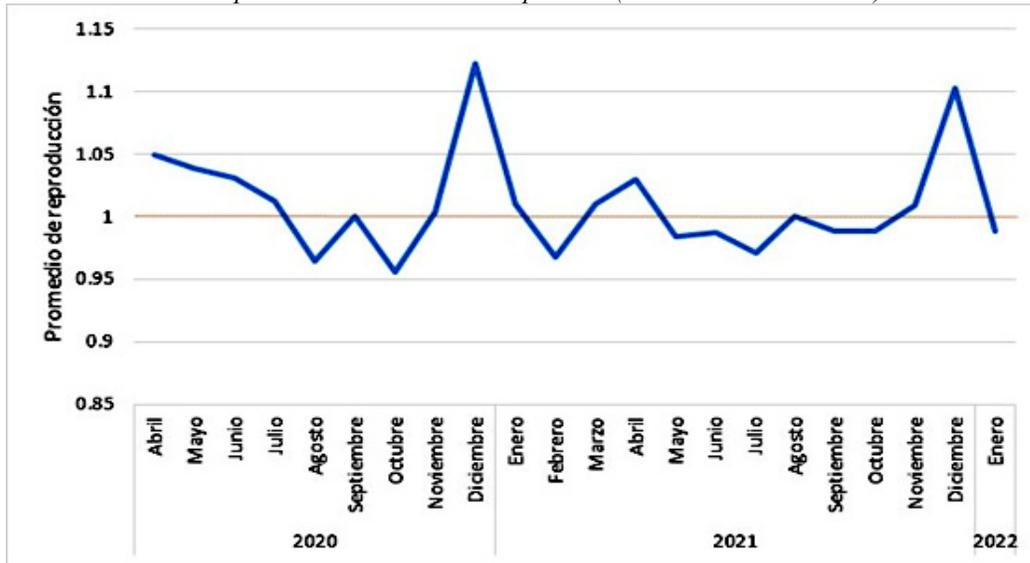
Año	Mes	Promedio	Probabilidad de extinción	
2020	Abril	1.0492	0.9077	
	Mayo	1.0387	0.9264	
	Junio	1.0312	0.9401	
	Julio	1.0130	0.9744	
	Agosto	0.9649	1.0000	
	Septiembre	1.0001	0.9998	
	Octubre	0.9559	1.0000	
	Noviembre	1.0035	0.9930	
	Diciembre	1.1226	0.7898	
	2021	Enero	1.0104	0.9795
		Febrero	0.9684	1.0000
		Marzo	1.0105	0.9793
Abril		1.0297	0.9429	
Mayo		0.9845	1.0000	
Junio		0.9880	1.0000	
Julio		0.9708	1.0000	
Agosto		1.0011	0.9978	
Septiembre		0.9884	1.0000	
Octubre		0.9884	1.0000	
Noviembre		1.0098	0.9807	
Diciembre		1.1030	0.8192	
2022	Enero	0.9881	1.0000	

Estudio de la transmisión de COVID-19 en Bolivia a través del proceso de Galton Watson

La Figura 3 muestra la variación de la tasa promedio de reproducción durante cada mes entre abril de 2020 y enero de 2022, se puede apreciar que en el 50% de los meses el promedio de reproducción se sitúa por encima de 1, y el mes con más incidencia es diciembre de 2020 y 2021, dando una pauta del comportamiento estacional de la enfermedad. De acuerdo a la ecuación 2, esto indica que la enfermedad está a punto de extinguirse en esos meses.

Figura 3.

Bolivia: Promedio de reproducción del COVID-19 por mes (01/04/2020-30/01/2022)

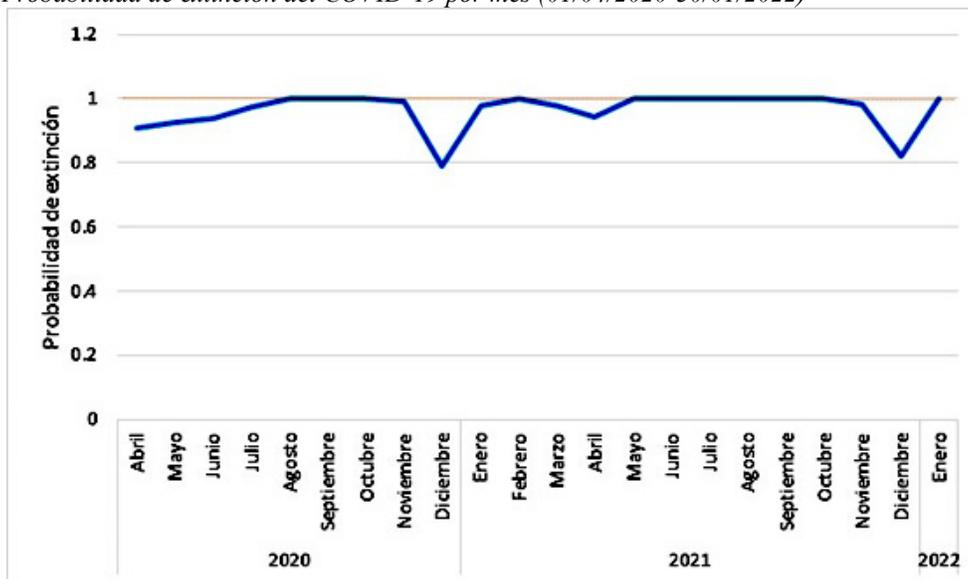


Fuente: Elaboración propia en base a datos del Centro de Ciencias e Ingeniería de Sistemas de la universidad Johns Hopkins

La Figura 3 muestra que el promedio de reproducción es mayor los meses de diciembre, indicando un patrón cíclico cada año. La Figura 4 muestra que la probabilidad de extinción del virus ha estado por encima de 1 en la mayoría de los meses, de acuerdo a la ecuación 2, esto indica que la enfermedad está a punto de extinguirse en esos meses.

Figura 4.

Bolivia: Probabilidad de extinción del COVID-19 por mes (01/04/2020-30/01/2022)



Fuente: Elaboración propia en base a datos del Centro de Ciencias e Ingeniería de Sistemas de la universidad Johns Hopkins

La Figura 4 confirma que la propagación del virus COVID-19 tiene un efecto estacional y que durante el mes de diciembre de las gestiones 2020 y 2021 se tienen la menor probabilidad de extinción del virus.

7. Discusión

El análisis de la propagación del virus en Bolivia a partir del promedio de reproducción y la probabilidad de extinción durante cada mes a partir de abril del 2020 a enero de 2022. El presente trabajo sugiere un efecto estacional de la propagación del virus durante el mes de diciembre de cada año, sin embargo, aún se necesita analizar el comportamiento en los meses posteriores.

El trabajo realizado por Crespo (2020) es un buen punto de partida de tales estudios, ante la carencia de datos se basa en simulaciones. Es importante contrastar estos resultados con otros países para tener una mayor comprensión del fenómeno de propagación del virus.

8. Conclusión

Los resultados obtenidos indican que el proceso de Galton-Watson puede ser utilizado en apoyo a la toma de decisiones y la comprensión de la evolución en la transmisión del COVID-19. La estimación del promedio de reproducción y de su varianza se realizó por máxima verosimilitud. En Bolivia, los resultados muestran que el promedio de reproducción mensual del virus COVID-19 varía entre 0.95 y 1.12, este aspecto revela que el país no está tan cerca de mitigar el virus. El comportamiento del proceso de Galton Watson se fundamenta en la estimación del promedio de reproducción μ , el mismo puede ser estimado a partir de una distribución de Poisson.

Los meses con mayor promedio de reproducción ha sido diciembre del 2020 y diciembre del 2021, aspecto que refleja la necesidad de tomar acciones para aminorar la propagación del virus durante cada fin de año.

Se recomienda el estudio del proceso de Galton-Watson con estimadores no paramétricos. La distribución de probabilidad supuesta para la aplicación es la ley de Poisson, se recomienda analizar el uso de otras distribuciones de probabilidad como la distribución geométrica.

Referencias Bibliográficas

Athreya, P. (1972). *Branching Processes*, New York: Springer Verlag.

Abbott, S. (2020). The transmissibility of novel Coronavirus in the early stages of the 2019-20 outbreak in Wuhan: Exploring initial point-source exposure sizes and durations using scenario analysis [version 1]. *Wellcome Open Research* 2020, 5(17), 1-12. <https://wellcomeopenresearch.org/articles/5-17>

Blanco, L. (1996). ¿Qué es un Proceso de Ramificación? *Boletín de Matemáticas*, Nueva Serie, Santa fe de Bogotá, III, 43-50. https://www.academia.edu/33989688/Proceso_de_Galton_Watson?auto=download

Boldog, P., Tekeli T., Vizi Z., Dénes A., Bartha F. A. y Rost G. (2020) *Risk assessment of novel*

- coronavirus COVID-19 outbreaks outside China. Journal of Clinical Medicine.* 9(2), 15-37. <https://www.mdpi.com/2077-0383/9/2/571>
- Johns Hopkins University (diciembre, 2021). *Coronavirus Resource Center: COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.* <https://coronavirus.jhu.edu/map.html>
- Crespo, M. De la Cruz, P. Quevedo, E. Valdés, J (2020). Los Procesos de Ramificación y su Aplicación al Estudio de la Transmisión de la COVID-19 en Cuba. *Revista de Ciencias Matemáticas*, 34(1), 89-97. <http://www.revinfodir.sld.cu/index.php/infodir/article/download/1042/1244>
- Guttorp, P. y Perlman, M. D. (2022). Predicting Extinction or Explosion in a Galton-Watson Branching Process with Power Series Offspring Distribution. University of Washinton. Technical Report No.626. <https://stat.uw.edu/research/tech-reports/predicting-extinction-or-explosion-galton-watson-branching-process-power-series-offspring>
- Harris, T.E. (1963). *The Theory of Branching processes.* Springer Verlag, Berlín.
- Heyde, C. C.(1970). Extension of a result of Seneta for the Super Critical. Galton-Watson Process. *The Annals of Mathematical Statistics.* Vol. 41, No.2, 739-742. <https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-41/issue-2/Extension-of-a-Result-of-Seneta-for-the-Super-Critical/10.1214/aoms/1177697127.full>
- Imai, N., Cori, I. A., Dorigatti, M. Baguelin, C. A. Donnelly, S. Riley y N. M. Ferguson (2020) Report 3: Transmissibility of COVID-19. Imperial College London COVID-19 Response Team. <https://www.jstor.org/stable/2239385>
- Ospina, R.(2001). Proceso de Galton Watson. *Revista Colombiana de Estadística*, 24(1), 13-26. file:///C:/Users/xiome/Downloads/Proceso_de_Galton_Watson.pdf
- Seneta, E. (1970). An explicit-limit theorem for the critical Galton-Watson process with immigration. *Journal Royal Statistics Society. Series B* (32), 149-152. <https://www.jstor.org/stable/2984413>
- Scott, D. (1987). On posterior asymptotic normality and asymptotic normality of estimators for the Galton-Watson process. *Journal Royal Statistics Society, Series B* (49), 209-214. <https://zh.booksc.eu/book/61592903/7f6ded>
- Stigler, S. M. (1971). The estimation of the probability of extinction and other parameters associated with branching processes. *Biometrika*, 58, 499-508. <https://www.jstor.org/stable/2334385>
- Waugh, W.A. O’N. (1958). Conditioned Markov processes. *Biometrika* 45, 241-249. <https://academic.oup.com/biomet/article-abstract/45/1-2/241/264700>