**Universidad de Oriente**

Artículo original

**Modelación matemática de los casos autóctonos de Covid-19. Municipio Bayamo. Febrero- Agosto 2021**

Mathematical modeling of the autochthonous cases of Covid-19. Bayamo Municipality. February- August 2021

Modelagem matemática dos casos autóctones de Covid-19. Municipio de Bayamo. Fevereiro a agosto de 2021

Ricardo Roberto Fonseca Pantoja I\* <https://orcid.org/0000-0002-4816-7518>

Zulin Fonseca González II  [https://orcid.org/0000-0001-8995-9661](https://orcid.org/0000-0001-8995-9661%20)

I Universidad de Oriente. Facultad de Matemática y Computación. Santiago de Cuba, Cuba.

II Centro Provincial de Higiene, Epidemiología y Microbiología. Bayamo. Granma, Cuba.

\* Autor para la correspondencia. Email: [zfonsecag@infomed.sld.cu](mailto:zfonsecag@infomed.sld.cu)

**RESUMEN**

El municipio Bayamo acumuló, 8162 casos positivos autóctonos de febrero a agosto en el año 2021, es el centro de la epidemia en la provincia de COVID-19 provocada por el SARS –CoV-2 determinado por el test de Proteína C Reactiva, representa el 53,2 % del total de los casos en ese periodo en Granma, muy diferente a lo ocurrido en el año 2020 en el cual la provincia acumuló solamente 185 personas contagiadas en nueve meses, con una tasa de 22.6 la más baja de Cuba. La provincia Granma acumuló 119 fallecidos en agosto/2021 que representa el 62,9 % de todos los muertos desde que comenzó la pandemia hasta agosto, lo que indica la alta incidencia de la epidemia que hay en estos momentos. Para la modelación matemática y el análisis de los casos positivos autóctonos de todos los ocurridos durante los meses de febrero a agosto en el año 2021 en Bayamo se obtuvieron polinomios de grado tres y cuatro que modelan el comportamiento de la epidemia durante los siete meses analizados, así como el de los fallecidos durante el mes de agosto en Granma con un carácter predictivo mayor al 98 % en todos los modelos.

**Palabras clave**: Epidemia; COVID-19; Modelo estadístico-matemático; Ecuación de regresión; Media móvil; Pronóstico; Pendiente epidemiológica.

**ABSTRACT**

The Bayamo municipality accumulated 8162 autochthonous positive cases from February to August in 2021, it is the center of the epidemic in the province of COVID-19 caused by SARS-CoV-2 determined by the C-Reactive Protein test, represents the 53.2% of the total cases in that period in Granma, very different from what happened in 2020 in which the province accumulated only 185 infected people in nine months, with a rate of 22.6, the lowest in Cuba. Granma province accumulated 119 deaths in August / 2021, which represents 62.9% of all deaths since the pandemic began until August, which indicates the high incidence of the epidemic that exists at the moment. For the mathematical modeling and analysis of the autochthonous positive cases of all those that occurred during the months of February to August in 2021 in Bayamo, polynomials of degree three and four were obtained that model the behavior of the epidemic during the seven months analyzed. as well as that of the deceased during the month of August in Granma with a predictive character greater than 98% in all models.

**Keywords:** Epidemic; COVID-19; Statistical-mathematical model; Regression equation; Moving average; Forecast; Epidemiological pending.

**RESUMO**

O município de Bayamo acumulou 8.162 casos autóctones positivos de fevereiro a agosto de 2021, é o centro da epidemia na província de COVID-19 causada pelo SARS-CoV-2 determinado pelo teste da Proteína C Reativa, representa 53,2% de o total de casos nesse período no Granma, muito diferente do que aconteceu em 2020 em que a província acumulou apenas 185 pessoas infectadas em nove meses, com uma taxa de 22,6, a mais baixa de Cuba. A província do Granma acumulou 119 mortes em agosto / 2021, o que representa 62,9% de todas as mortes desde o início da pandemia até agosto, o que indica a alta incidência da epidemia que existe no momento. Para a modelagem matemática e análise dos casos positivos autóctones de todos os ocorridos durante os meses de fevereiro a agosto de 2021 em Bayamo, foram obtidos polinômios de grau três e quatro que modelam o comportamento da epidemia durante os sete meses analisados. bem como o dos falecidos durante o mês de agosto no Granma com caráter preditivo superior a 98% em todos os modelos.

**Palavras-chave:** Epidemia; COVID-19; Modelo estatístico-matemático; Equação de regressão; Média móvel; Previsão; Epidemiológica pendente.

Recibido:

Aprobado:

**Introducción**

El 31 de diciembre de 2019, las autoridades de la ciudad de Wuhan, en la provincia de Hubei, China, reportaron los primeros casos de síndrome respiratorio agudo de etiología desconocida. El 7 de enero de 2020, las autoridades chinas informaron que un nuevo coronavirus (2019-nCoV). El 11 de marzo de 2020 la COVID-19 fue considerada como una pandemia por la OMS. En Cuba se diseñó, en enero de 2020, el Plan de Medidas para el enfrentamiento a la COVID-19, que involucra a todos los Organismos de la Administración Central del Estado, las empresas, el sector no estatal y la población en general. El 11 de marzo de 2020 fueron confirmados los primeros casos de la enfermedad en el país: tres turistas de la región italiana de Lombardía, que fueron hospitalizados de manera inmediata.

La provincia Granma, cuarta del país por el número de habitantes y quinta en extensión territorial ha visto aumentado considerablemente el número de casos de Covid-19 en el año 2021, (1,2) de febrero a agosto del 2021, el municipio Bayamo es el epicentro de la epidemia en Granma. Para realizar este trabajo se consideró como casos confirmados los pacientes que resultaron positivos al estudio virológico para la COVID-19, con o sin sintomatología que hayan enfermado a partir de otro caso ocurrido en el territorio nacional y se define; como caso autóctono.

La medición es vital en el estudio de las enfermedades en la población, (3) en el estudio de una epidemia tiene gran importancia modelar su comportamiento y si es posible predecir su desarrollo, por lo que la aplicación de modelos estadístico-matemáticos pueden servir para analizar tendencias, realizar pronósticos y tomar decisiones para evitar el aumento de la epidemia. La Matemática y la Estadística, (4) permiten avizorar el futuro desde el presente y tener un mejor conocimiento del desarrollo de esta enfermedad para disminuir sus efectos y modelar el comportamiento de esta enfermedad hasta eliminarlos, por tanto, el objetivo de este trabajo es modelar los casos autóctonos y su comportamiento en el municipio Bayamo de la COVID-19 durante los meses de febrero a agosto en el año 2021 y de los fallecidos en Granma durante el mes de agosto.

**Métodos**

Los datos analizados corresponden a los informados diariamente por la Dirección Nacional de Higiene y Epidemiología del Ministerio de Salud Pública de Cuba por los medios de prensa nacionales. (1)

Para modelar el comportamiento de la epidemia en el municipio Bayamo se consideró como variable independiente el día del mes en que se obtuvieron los resultados positivos de la cantidad de casos que hubo ese día y variable dependiente el número de casos positivos autóctonos, se probaron veintisiete modelos matemáticos; lineales, potenciales exponenciales, inversos, recíprocos, logarítmicos, intrínsecamente lineales y polinómicos (4), hasta de grado 8. Se hizo un análisis de correlación y regresión utilizando los métodos de los mínimos cuadrados y el de Gauss-Newton, con un máximo de 50 iteraciones y como criterio de convergencia de los parámetros y 10-6 de precisión, así como se hizo un análisis de residuales, se utilizó la función de pérdida:

FP = Valor absoluto (Yobservado – Ycalculado)

se utilizaron los paquetes estadísticos profesionales SPSS v25 de la IBM, (5) Statistica v12 de StatSoft, (6) InfoStat v2013. (7) También se hizo una corrida de un modelo matemático polinómico de los siete meses considerando un valor consecutivo de los días para día 1 (1ro. de febrero) hasta el día 212 que corresponde al 31 de agosto con el acumulado de los casos autóctonos de la Covid-19 en Bayamo.

Teniendo en cuenta que en el mes de febrero hubo algunas anomalías con el tiempo de realización de los PCR, que demoraban por capacidad de laboratorios para realizarlos hasta cuatro días, se corrió un modelo de regresión mínimo cuadrático polinómico de grado tres con la variable dependiente con las **medias móviles simples**, en este caso se utilizó la media móvil previa teniendo en cuenta el valor de datos antiguos, es la media aritmética de los datos anteriores, es una técnica elemental para predicciones:

n = 4 de los casos autóctonos en ese mes con 25 medias y fórmula: m = 4; … ; 28 para mejorar el pronóstico del posible modelo a obtener y probar una posibilidad más ante un inconveniente que puede variar los resultados estadísticos.

Teniendo en cuenta que el mes de agosto ha sido para Granma hasta ahora el de mayor letalidad por la alta virulencia del SAR-Cov 2 se analizaron el número de muertes diarias acumuladas en la provincia; 119 fallecidos, (1) fueran o no residentes fijos, se aplicó un modelo de mejor ajuste polinómico de grado 3.

Los modelos obtenidos por ser funciones polinómicas, son derivables y aplicando la definición geométrica de la primera derivada en un punto que representa la recta tangente a la curva en un ese punto, por lo tanto, si se calcula la primera derivada del modelo obtenido se puede conocer la pendiente en un punto (un día del mes) y da como resultado un valor numérico de la pendiente de la recta tangente a la curva y su interpretación epidemiológica.

**Resultados**

Mediante la modelación estadístico-matemática se obtuvo el polinomio segmentado (compuesto) de mejor ajuste de grado 3 con variable independiente (X) el día epidemiológico correspondiente al mes analizado y la dependiente correspondiente (Y) a los casos autóctonos acumulados mensualmente. Fig.1.

**Fig.1.** Modelación estadístico-matemática de casos autóctonos acumulados mensualmente.

Este modelo presenta como dominio (Dom) para cada rama los días del mes correspon-diente, se obtuvieron altos coeficientes de correlación (r): febrero **(0,999)**, marzo (0,994), abril (0,998), mayo (0,996), junio (0,999), julio (0,999) y agosto (0,998) respectivamente tienen un coeficiente de determinación de: 99,8, 98,8 %; 99,6 %; 99,2 %, 99,8 %, 99,8 % y 99,6 % lo que indica que como modelos de pronósticos tienen muchas posibilidades, estadísticamente fueron altamente significativos con una probabilidad de error en los análisis de varianza mayor al 99.9 % de probabilidad de error, con la característica de que son de grado tres.

Las ecuaciones obtenidas son monótonas crecientes en todo su dominio, por la característica de las gráficas que modelan, una parte son cóncavas y otros convexas, pueden tener varios puntos de inflexión, y por ser valores epidemiológicos acumulativos carecen de ceros reales, si en el tiempo se controla la epidemia cuando no haya casos se modela por una recta constante, paralela al eje X de las abscisas.

Del análisis del acumulado consecutivo de todos los casos autóctonos durante los siete meses se obtuvo el modelo representado por el polinomio de grado 4:

Y = 22,75568 – 14,68493·X + 1,24617·X2 – 0,010564·X3 + 0,0000278·X4

con un valor alto de r = 0,999, representado en el gráfico 1.



**Gráf. 1.** Modelo correspondiente a los casos autóctonos de Covid-19 en el municipio Bayamo, desde febrero a agosto, año 2021.

Coeficiente de correlación r = 0,999

Esta ecuación modela con mucha fiabilidad la epidemia durante el tiempo analizado. Si se fuera a realizar algún pronóstico para el mes de septiembre es mejor utilizar el polinomio de grado 3 del modelo compuesto el correspondiente a la rama de agosto, por no estar influenciado por valores muy distantes del día que se quiere pronosticar. Teniendo en cuenta que una epidemia es un fenómeno contra la que se combinan todos los recursos humanos y materiales para controlarla, que es hacer cero el número de casos positivos en el tiempo, y los controles deben dar resultados en un periodo no extenso en el tiempo.

Se realizó una corrida con el acumulado de las medias móviles de los casos autóctonos correspondiente al mes de febrero y se obtuvo la ecuación:

Y = - 39,062916 + 12,721471·X - 0,593686·X2 + 0,0179437·X3

Esta ecuación tiene muy buen ajuste con coeficiente de correlación muy alto r = 0,99958945, casi uno, con dominio día 4 a 28 de febrero y supera a febrero en la rama correspondiente del modelo compuesto. Las ventajas de las medias móviles en epidemiologia permiten realizar mejores pronósticos correspondiente a febrero), corregir las variaciones de analizar muchas pruebas de PCR un día y pocas al siguiente, situación que aumenta las variaciones y los errores standard de los cálculos que se hacen. (Graf.2 y Graf.3)



**Gráf.2.** Acumulado medias móviles febrero.

r = 0,9996

Y = - 39,062916 + 12,721471·X - 0,593686·X2 + 0,0179437·X3



**Gráf. 3.** Acumulado diario de febrero.

r = 0.9988

Y = - 9,8146535 + 10,6631343·X – 0,5113412·X2 + 0,0183338·X3

Observando los gráficos 2 y 3 se aprecia que el de mejor ajuste está representado por el correspondiente a las medias móviles.

Todos los modelos permiten calcular:

El promedio diario de los casos autóctonos acumulados en cualquier día del mes, dividiendo el resultado obtenido entre el número de días (#días):

La pendiente epidemiológica. Derivando el modelo y sustituyendo por el número de días (#días) en la primera derivada se obtiene el valor de la pendiente de la recta tangente a la curva en ese punto y se interpreta como un indicador del grado de contagiosidad de la epidemia y realizar comparaciones. Analicemos el comportamiento de la epidemia los días 10 de marzo y abril

Para marzo la primera derivada del modelo para X = 10 es:

en abril la primera derivada del modelo es:

Respectivamente para X = 10 los resultados son:  38,24842 y

50.0353‬. Comparando esos dos valores indica que la velocidad de crecimiento de la epidemia el día 10 de abril fue superior al 10 de marzo. Matemáticamente esos valores representan las pendientes de las rectas tangentes a las respectivas curvas en el punto de abscisa igual a 10. Estos valores representan también los ángulos de inclinación de las respectivas curvas puntos tangenciales. Conociendo que la tangente (Tan) de un ángulo alfa α es Tan(α) = m (pendiente) y m, es posible conocer el ángulo de inclinación (α) mediante la expresión ArcTan(m) = α por definición trigonométrica, realizando las operaciones matemáticas correspondientes se calculan los respectivos ángulos; 88,50 grados y 88,85 grados lo que significa que las pendientes son extremadamente altas y no hay control de la epidemia.

Es importante tener presente que el control de la epidemia se logra en el tiempo cuando la variable de los enfermos activos se hace cero y para ello se puede formular esta variable como: **casos activos = casos positivos – (altas epidemiológicas + fallecidos + traslados al extranjero + evacuados al extranjero)**. En el caso de este trabajo se escogió como variable independiente el tiempo expresados mediante los días del mes, porque el día a día resume todos los factores positivos y negativos (las indisciplinas de la población, la falta de percepción del riesgo) que permiten el aumento o disminución de la epidemia y que harían demasiado complejo el pronóstico y comportamiento de la epidemia.

Teniendo en cuenta la situación epidemiológica de la provincia Granma y el número de fallecidos que hubo en agosto 2021 (donde hay un pico pandémico) se modeló el acumulado de fallecidos para este mes, obteniéndose el polinomio de mejor ajuste de grado 3 con r = 0,997. (Gráf. 4)



**Gráf. 4.** Acumulado diario de fallecidos de Covid-19 en la provincia Granma en agosto 2021.

Coeficiente de correlación r = 0,997

Y = -1,90434 + 3,67937X – 0,107188X2 + 0,00390028X3

Si se quisiera hacer un pronóstico de cuántos fallecidos habría durante el mes de septiembre el cálculo es para X1 = 31 último día del mes de agosto y X2 = X1 + 30 = 61 (30 días del mes de agosto) que nos daría los fallecidos del mes de septiembre en Granma = 584 (Y2 - Y1)) personas muertas. Indicando que si no se toman todas las medidas necesarias para el control de la epidemia morirían como promedio diario aproximadamente 19,5 personas por causa de esta enfermedad.

**Discusión**

Una epidemia depende de un conjunto de variables, entre ellas están; las características de la epidemia, grupos de susceptibles, infectados, recuperados, las condiciones edafo-climáticas, la robustez del sistema de salud, las condiciones nutricionales y socioeconómicas de las personas, su disciplina, cultura de la sociedad, las indisciplinas epidemiológicas y otros factores. Controlar todas las variables se hace muy complejo, pero todas ellas están incluidas en el día a día, es decir, en el día epidemiológico (X) de la ocurrencia de la epidemia, por ello fue escogido el día del mes para modelar la epidemia y el acumulado de los casos autóctonos (Y) un modelo extremadamente simple que permite modelar y pronosticar crudamente el comportamiento en cuanto al número de personas infectadas o contagiadas en la epidemia.

La modelación matemática de una epidemia reviste gran importancia y así lo corroboran; Casals, Guzmán y Caylá. (8) que identificaron 617 artículos científicos, de los cuales seleccionaron un total de 162 artículos relacionados con modelos matemáticos utilizados en el estudio de enfermedades transmisibles. Díaz-Quijano, (9) hizo un estudio relacionado con los aspectos básicos de la utilización de los análisis de regresión en los estudios epidemiológicos, siendo su objetivo la aplicación de las regresiones lineales múltiples. Abelló, Guinovart y Morales, (10) realizaron una investigación aplicando el modelo SIR básico y políticas antiepidémicas de salud pública para la COVID-19 en Cuba resaltando la importancia de la epidemiología matemática en la propagación de enfermedades teniendo muy en cuenta la variable número de infectados reafirmando la importancia del control del número de pacientes infectados autóctonos en el tiempo.

Kyurkchiev N, Illiev A, Rahnev A, Terzieva, (11) plantean que las variables polinómicas proporcionan una muy buena aproximación de los datos específicos sobre la pandemia de Covid-19 en Cuba y obtienen varios modelos de crecimiento modelando la epidemia en Cuba desde el 10/marzo/2020 hasta el 05/junio/2020. Pérez, Curra y Almaguer, (12) investigaron en la modelación de la epidemia de Covid-19 en la provincia Holguín, analizando las personas susceptibles, infectadas, recuperadas y fallecidas mediante un modelo SIRD, el trabajo fue avalado por ocho publicaciones sobre modelación matemática que hicieron otros autores algunos de ellos de la República Popular China.

Zenteno, et. al. (13) modelaron en México el número de casos diarios de coronavirus mediante polinomios de grados; tres, cuatro y cinco obteniendo buenos ajustes. Vega y Pérez, (14) en la provincia Holguín obtuvieron un modelo polinómico de grado cinco en el análisis de la efectividad del control de la epidemia demostrando la posibilidad de emplear modelos polinómicos para el estudio de esta pandemia en Cuba.

Los anteriores autores demuestran que es posible utilizar una gran variedad de modelos matemáticos diferentes en el análisis de una epidemia, por lo que este trabajo corrobora lo anterior mediante la utilización y obtención de modelos polinómicos utilizando dos variables; el tiempo epidemiológico y la otra que utilizan los modelos SIR, el número de infectados, así como, es posible utilizar el cálculo diferencial en el análisis del comportamiento de la epidemia cuando se modela mediante polinomios su comportamiento

**Conclusiones**

Teniendo en cuenta los resultados obtenidos se concluye que, es posible utilizar como variable de pronóstico el día epidemiológico para la modelación de los casos autóctonos acumulados diarios. Los mejores modelos matemáticos mensuales obtenidos fueron polinomios de grado tres para analizar el comportamiento de la epidemia, con coeficiente de correlación mayor a 0,99. Para un pronóstico del comportamiento de la epidemia los mejores ajustes se obtuvieron mensualmente. El mejor ajuste para los siete meses analizados del acumulado de casos diarios se logró mediante un polinomio de grado cuatro. Es posible utilizar las medias armónicas para modelar la epidemia de Covid-19 cuando se presentan intermitencias en los análisis de PCR.

**Referencias bibliográficas**

1. Covid19CubaData. Tablero estadístico interactivo sobre la COVID-19 en Cuba 2020 [citado 15/02/2021 a 10/09/2021]. Utilizable en: https://covid19cubadata.github.io/#cuba.

2. Ministerio de Salud Pública. Anuario Estadístico de Salud 2020 Editado 2021, Dirección de Registros Médicos y Estadísticas de Salud, La Habana, Cuba. Correo electrónico: drmes@msp.sld.cu. Sitio: https://temas.sld.cu/estadisticassalud/ y http://bvscuba.sld.cu/anuario-estadisticode-cuba/

3. Fajardo-Gutiérrez A. Medición en epidemiología: prevalencia, incidencia, riesgo, medidas de impacto. *Rev Alerg Mex*. 2017;64(1):109-120

4. R Ellison SL, Barwi VJ, Duguid Farrant TJ. Chapter 7. Regression. In: Practical Statistics for the Analytical Scientist. 2ed. Published by The Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge CB4 0WF, UK; 2009. p92-113.

5. SPSS® Statistics, IBM® Company Statistical Package for The Social Sciences for Windows (programa de computación) ver. 25.0.0, Chicago, Illinois, United State of America. 2017, is available on the Web at, http://www.ibm.com/legal/copytrade.shtml.

6. StatSoft, Inc. (2014): SOFTWARE STATISTICA (data analysis software system), version 12. [www.statsoft.com](http://www.statsoft.com).

7. Di Rienzo JA, Casanoves F, Balzarini MG, Gonzalez L, Tablada M, Robledo CW. InfoStat (programa de computación) versión 2013. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina. URL <http://www.infostat.com.ar>

8. Casals M, Guzmán Katty, Caylà JA. MODELOS MATEMÁTICOS UTILIZADOS EN EL ESTUDIO DE LAS ENFERMEDADES TRANSMISIBLES. Barcelona, España. Rev Esp Salud Pública 2009; 83: 689-695 N.° 5 - Septiembre-Octubre 2009. Email: [mcasals@aspb.cat](mailto:mcasals@aspb.cat)

9. Diaz-Quijano FA. Regresiones aplicadas al estudio de eventos discretos en epidemiología. Rev Univ Ind Santander Salud. 2016; 48(1): 9-15. DOI: http://dx.doi.org/10.18273/revsal.v48n1-2016001

10. Abelló Ugalde IA, Guinovart Díaz R, Morales Lezca W. El modelo SIR básico y políticas antiepidémicas de salud pública para la COVID-19 en Cuba. Revista Cubana de Salud Pública. 2020;46(Supl. especial):e2597.

11. Kyurkchiev N, Illiev A, Rahnev A, Terzieva T. Another look at a good approximation of data for the distribution of COVID-19 in Cuba. Rev haban cienc méd. 2020; 19(3): e3445. Available from: http://www.revhabanera.sld.cu/index.php/rhab/article/view/3445

12. Roberto Pérez R., R., Curra S. D. y Almaguer M., L. (2020): Análisis preliminar de modelos SIRD para la predicción de la COVID-19: caso de la provincia de Holguín. Anales de la Academia de Ciencias. Ciencias Naturales y Exactas. <https://orcid.org/0000-0001-5741-5168>.

13. Zenteno AC, Santiago MC, Romero Y, Pérez J. Rubín GT, Álvarez AE, et al. Optimización de los coeficientes del modelo predictivo del número de casos diarios de coronavirus Covid-19 en México, ISSN 1870-4069, pp. 293–302; acc. 19-10-2020, Research in Computing Science 149(11), 2020, email { ana.zenteno, marycarmen.santiago, yeiny.romero, judith.perez, gustavo.rubin }@correo.buap.mx

14. Vega de la Cruz LO, Pérez Pravia MC. Efectividad de las medidas ante la COVID-19 basada en el análisis del curso pandémico. *Arch méd Camagüey*. 2020;24(5):e7645 Artículo original ISSN 1025-0255, Autor de correspondencia (email): leovega@uho.edu.cu

**Conflicto de intereses**

Los autores declaran que no hay conflictos de intereses.

**Contribución de autoría**

Ricardo Roberto Fonseca Pantoja: Conceptualización, Análisis formal, Metodología, Investigación, Supervisión y Redacción –revisión y edición.

Zulin Fonseca González: Conceptualización, Curación de datos, Visualización, Redacción –borrador original.

Yo, Ricardo Roberto Fonseca Pantoja, en nombre de los coautores, declaro la veracidad del contenido de este artículo.