

Revista científica CENTROS
15 de diciembre de 2015 – Vol. 4 No. 2
ISSN: 2304-604X pp. 189-199

Recibido: 30/06/15; Aceptado: 06/12/15

Se autoriza la reproducción total o parcial de este artículo, siempre y cuando se cite la fuente completa y su dirección electrónica.

<http://www.revistacentros.com>

indexada en



http://www.latindex.unam.mx/buscador/ficPais.html?opcion=1&clave_pais=33



PRONÓSTICO DE LAS DEFUNCIONES EN LA REPÚBLICA DE PANAMÁ UTILIZANDO LA METODOLOGÍA ARIMA O BOX-JENKINS PARA LOS AÑOS 2010- 2013.

Estelina Ortega de Gómez ¹, Elena Coba ¹, José Almanza ²

¹ Departamento de Estadística, Facultad de Ciencias Naturales, Exactas y Tecnología, Universidad de Panamá, Campus Central. Email: estelinao@hotmail.com

² Departamento de Estadística, Ministerio de Educación

RESUMEN

La metodología ARIMA o Box-Jenkins se fundamenta en la extracción de los movimientos predecibles en los datos observados, descomponiendo la serie de datos en varios componentes, llamados filtros, hasta obtener residuales no predecibles cuyo comportamiento tiene poca influencia en el resultado final. La presente investigación tiene como objetivo proyectar las defunciones en la República de Panamá haciendo uso de la metodología ARIMA o Box-Jenkins. Para ello se analizó la serie histórica de las defunciones en la República de Panamá desde 1952 al 2008. Luego se aplicó un análisis de serie de tiempo aplicando toda la

metodología ARIMA o Box-Jenkins. Obteniendo como principal conclusión un modelo estadístico que permite proyectar a corto plazo y con el mínimo error de predicción las defunciones que se pueden dar en el país.

PALABRAS CLAVE: Pronóstico, Defunciones, Modelo de Predicción, Autorregresión, Interacción.

ABSTRACT

The ARIMA methodology, also called Box-Jenkins, is based on extracting the predictable movements in the observed data, decomposing the data series into several components, called filters, until unpredictable residual whose behavior has little influence on the final result. This research aims to project the deaths in the Republic of Panama using the ARIMA or Box-Jenkins methodology. The time series of death within the Republic of Panama ranged from 1952 to 2008. The analysis applied all the steps of this methodology. The main conclusion resolved in a statistical model that allowed to project at a short term and with a minimum of errors the prediction of deaths that may occur in the country.

KEYWORD: Forecast, Deaths, Prediction Model, Autoregression, Interaction

INTRODUCCIÓN

Esta investigación pretende proyectar el número de defunciones que se podrían captar en la República de Panamá utilizando la metodología ARIMA o Box-Jenkins; la cual consiste en una metodología que ha tenido un indudable éxito en la práctica profesional por varios motivos: En primer lugar por su profundidad metodológica – estadística. En segundo lugar, existe una clara y consolidada guía de aplicación empírica de la misma que permite pasar con facilidad de las situaciones de laboratorio que crea la teoría, a la praxis profesional. En tercer lugar y no menos importante, los modelos ARIMA han demostrados una gran utilidad en la predicción a corto plazo de series de alta frecuencia.

El método Box – Jenkins, es el método que se utilizará para pronosticar las defunciones, el mismo consiste en extraer los movimientos predecibles en los datos observados, descomponiendo la serie de datos observados en varios componentes, llamados filtros, hasta obtener residuales no predecibles cuyo comportamiento tiene poca influencia en el resultado final; es decir se aplica a los modelos autorregresivos de media móvil ARMA o a los modelos autorregresivos integrados de media móvil (ARIMA) para encontrar el mejor ajuste de una serie temporal de valores, a fin de que los pronósticos sean más acertados. (García, 1995).

El objetivo de esta investigación es la de proyectar las defunciones en la República de Panamá haciendo uso de la metodología ARIMA o Box-Jenkins y compararlo con los pronósticos oficiales del Estado que son calculados con otros métodos.

MATERIAL Y MÉTODO

Las estimaciones y proyecciones de la población a nivel total país, de provincia y comarca indígena, se elaboran con el Método de los Componentes, con base los análisis que se hacen por separado de cada una de las variables que intervienen en la dinámica de la población, como lo son: la fecundidad, la mortalidad y la migración interna e internacional.

Se revisó la mortalidad infantil desde 1980, tomando como punto de partida la comparación de las tasas de mortalidad infantil obtenidas de las Estadísticas Vitales, los Censos de Población y la Proyección de Población Vigente, considerando la tendencia de las tasas de mortalidad infantil del período histórico anterior.

A la tasa de mortalidad infantil se le aplicó un diferencial por sexo derivado de las Tablas Modelo Oeste de Coale & Demeny. La tendencia de este diferencial viene desde 1950. Para las defunciones de 1 a 4 años se observó el comportamiento de los registros vitales del año 1990 y del año 2000.

Como el diferencial por sexo de las estadísticas vitales resulto aceptable, se aplicó, para ambos momentos, un porcentaje de omisión estimado coherente con la omisión tanto a las defunciones totales, como a las de menores de un año estimadas de las estadísticas vitales.

La metodología Box-Jenkins es una técnica que no asume ningún patrón particular en los datos históricos de la serie a pronosticar donde aplican métodos autorregresivos y de promedio móviles. Esta técnica permite verificar el modelo elegido con los datos históricos con el objetivo de describir la serie con precisión. Hanke y Reitsch (1996).

Para la aplicación de la Metodología Box-Jenkins se consideró realizar tres filtros lineales:

Autorregresivo (p)

Integración o diferenciación (d)

Medias Móviles (q)

ARIMA(p,d,q) para la parte no estacional

Para la construcción del modelo se sugieren los siguientes pasos:

Analizar la serie (transformarla)

Obtener la autocorrelación

Verificar la estacionariedad en media y en varianza

Identificar el modelo

Estimación de los parámetros estadísticos

Análisis residuales

Proceder a Pronosticar

RESULTADOS Y DISCUSIÓN

Análisis de la Serie

El comportamiento de las defunciones en la República de Panamá desde 1952 al 2012 muestra una marcada tendencia a incrementarse, mientras que la variación no parece constante, por lo cual se observa una posible no estacionalidad en media y en varianza. De la Figura 2 se obtiene la línea de tendencia, el R^2 y la ecuación de la serie de las defunciones,

que es la que mejor se ajusta a los datos. Observando el coeficiente de determinación es del 89.7% siendo este considerado como bueno para el estudio.

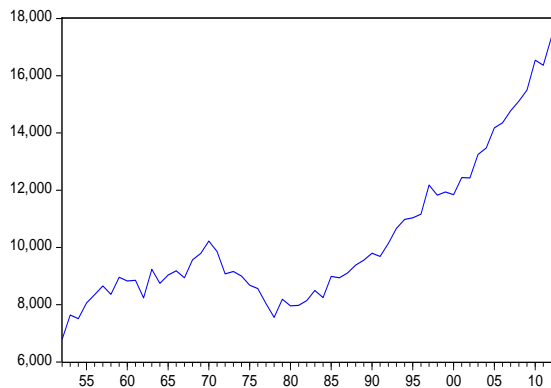


Figura 1. REPRESENTACIÓN DE LAS DEFUNCIONES DE LA REPÚBLICA DE PANAMÁ, AÑOS 1952-2012

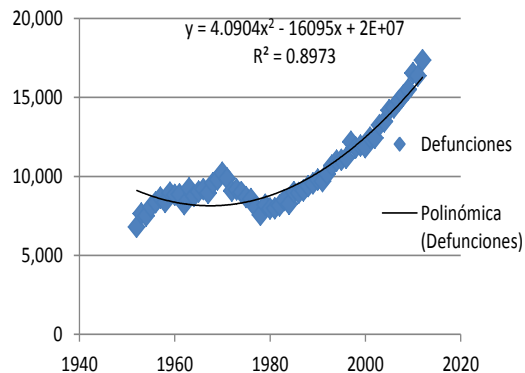


Figura 2. COMPORTAMIENTO DE LA SERIE DE DEFUNCIONES DE LA REPÚBLICA DE PANAMÁ .AÑOS 1952-2012

Al analizar el correlograma (Figura 3) se ve claramente que las defunciones (serie original) tiene tendencia, este hecho se puede ver en la función de autocorrelación, donde la serie, desde su primer retardo va decayendo lentamente a cero. Si la serie contiene una tendencia, es decir, cambia continuamente de nivel, los valores de r_k (AC, en el correlograma) no van a decrecer hacia cero rápidamente. Esto es debido a que una observación por encima de la media (o por debajo) de la media general es seguida de muchas observaciones por el mismo lado de la media debido a la tendencia. Este hecho hace a la serie de las defunciones no sea estacionaria en media, por lo que para volverla estacionaria hay que hacer uso de las diferenciaciones.

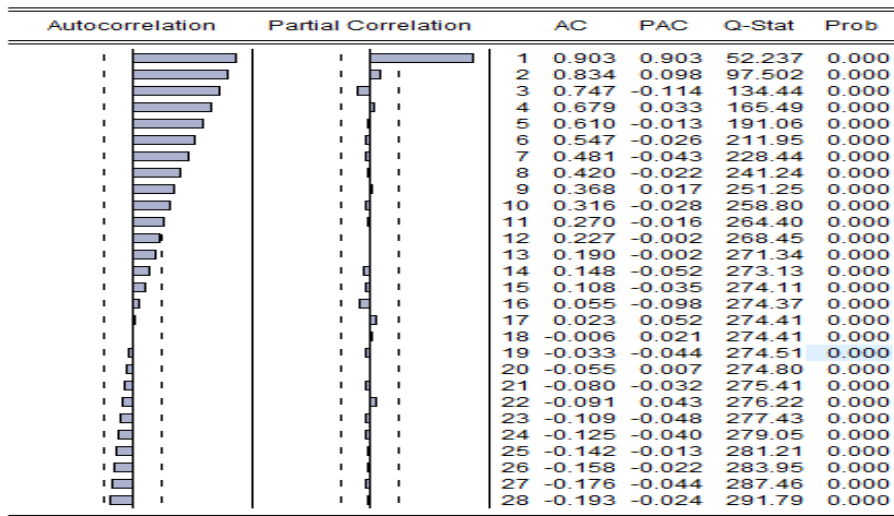


Figura 3. CORRELOGRAMA DE LA SERIE DE LAS DEFUNCIONES

Al aplicar la prueba Dicky- Fuller se comprueba que la serie de las defunciones no es estacionaria, ni en media ni en varianza, probando que el nivel de significancia alcanzado por la prueba señala que no se puede rechazar la hipótesis de la presencia de raíz unitaria al 5% de significancia.

Transformación de la Serie

Con el objetivo de estabilizar la varianza se utilizó la transformación Box-Cox la cual permitirá que las desviaciones alrededor del modelo estén normalmente distribuidas. Obteniendo de la transformación que el mínimo valor del coeficiente de variación le corresponde a igual a cero.

Una vez realizada la transformación Box-Cox se aplicó logaritmo a la serie y se aplica dos diferenciaciones logrando estabilizar la serie en media y varianza. La prueba Dicky-Fuller de las dos diferenciaciones hechas a log defunciones, se observar que mejora de un 64.4%

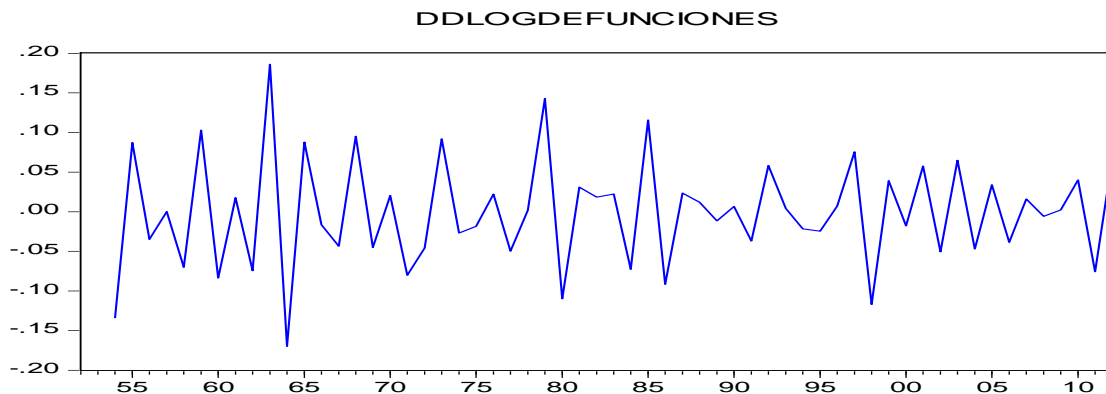


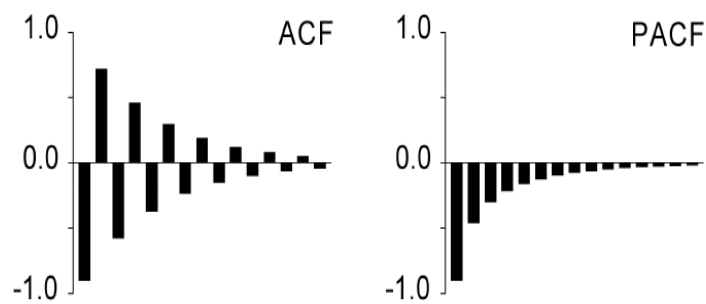
Figura 4. SERIE DDLOGDEFUNCIONES ESTACIONARIA EN MEDIA Y EN VARIANZA

ESTIMACIÓN DEL MODELO

Una vez obtenida una serie estacionaria en media y varianza, se procede a estimar los parámetros del modelo ARIMA.

Comparando con el modelo teórico

$$\text{ARMA}(1,1): \phi_1 = -0.80; \theta_1 = +0.80$$



El aspecto de las FAC y FACP se observa un decaimiento no truncado en ambos correlogramas que se ha planteado. El modelo planteado es un ARIMA (1, 2, 1) debido a las dos diferenciaciones realizadas.

Estimación

Los modelos que se probarán para evidenciar cuál de ellos es el que mejor que pronostica las defunciones en Panamá serán: ARIMA(1,2,1), ARIMA(0, 2,1) y ARIMA (1,2,0)

De ahí que las estimaciones del modelo ARIMA(1, 2,1) sin constante se puede considerar apropiado ya que presentó un R2 aceptable en comparación con los otros modelos propuestos.

Modelo 1 ARIMA(1, 2, 1) sin constante

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.415182	0.12771 2	- 3.250915	0.0019
MA(1)	-0.794897	0.08525 1	- 9.324225	0.0000
R-squared	0.641174	Mean dependent var		0.00129 5
Adjusted R-squared	0.634767	S.D. dependent var		0.06748 6
Inverted AR Roots	.42			
Inverted MA Roots	.79			

Análisis de la calidad del ajuste

La calidad del ajuste se puede medir por el coeficiente de determinación. En nuestro caso, R-squared es igual a 0.634767 lo cual indica que el modelo obtenido presenta un buen ajuste.

Validación del modelo

Análisis de los residuos.

Para la validación del modelo se analiza los residuos donde se observa que las autocorrelaciones de los residuos son significativas y entran dentro de las bandas de confianza, lo que indica que no son distintas de cero. Por su parte, el estadístico Q no muestra indicios de autocorrelación global de los residuos, puesto que el valor de Q estimado para los diferentes orden de autocorrelación que se muestran es siempre inferior al punto crítico de la χ^2 con los correspondientes grados de libertad y los niveles estándar de significatividad utilizados en el trabajo empírico, lo que nos lleva a rechazar la hipótesis nula de autocorrelación global de los residuos.

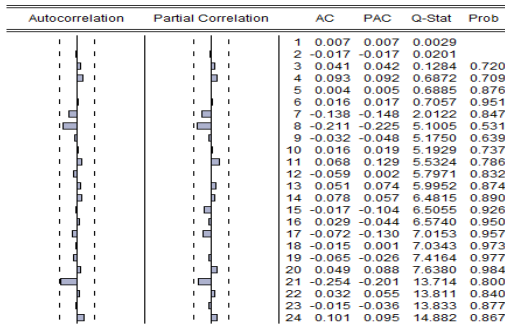


Figura 5. CORRELOGRAMA DE LOS RESIDUOS DEL MODELO ARIMA(1, 2, 1)

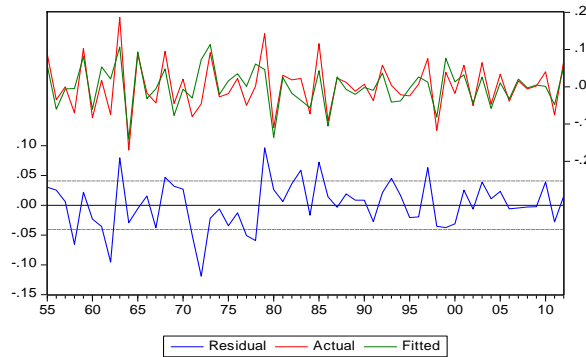


Figura 6. GRAFICO DE LOS RESIDUOS

En el gráfico de los residuos también apoya la ausencia de autocorrelación residual, puesto que la gran mayoría de los residuos están dentro de las bandas de confianza, con excepción de algún residuo anómalo. Por lo tanto, también muestra claramente que los residuos son ruido blanco.

Prueba de Durbin Watson para la independencia de los residuos.

Al realizar la prueba Durbin- Watson para probar la independencia de los residuos obtenemos un valor de 1.973972. Luego se tiene que los residuos son independiente para el modelo

Para probar la normalidad de los residuos se utiliza el estadístico de Jarque- Bera, que plantea la hipótesis de que si la curtosis y la asimetría de los datos son distintas, entonces la distribución de los residuos no es normal.

Por antes mencionado y por la tabla de los valores de los residuos, adjunta, vemos que no hay evidencia suficiente para rechazar la hipótesis nula por lo que los residuos tienen una distribución normal al 5% de significancia.

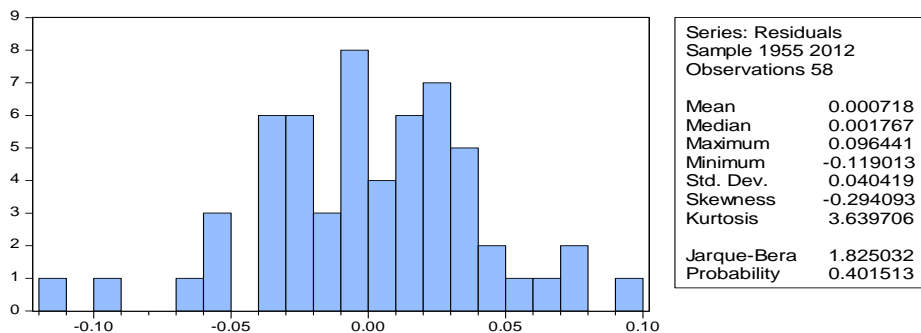


Figura.7 Prueba de normalidad para los residuos del modelo ARIMA(1, 2,1)

APLICACIÓN DEL MODELO ARIMA (1, 2,1)

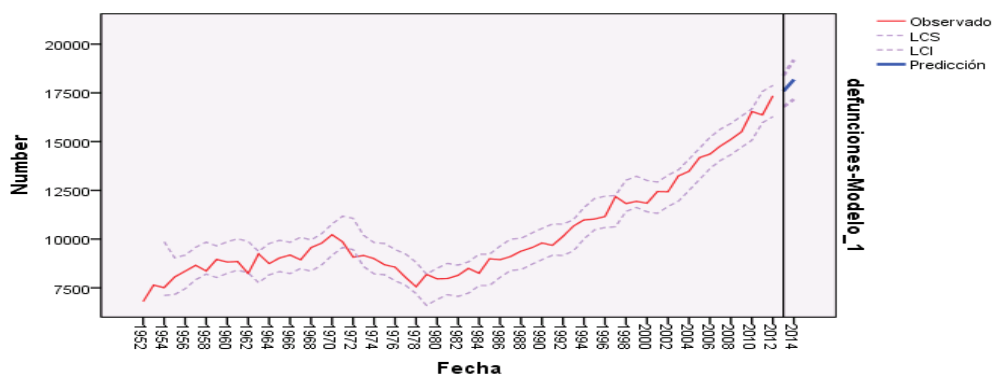


Figura 8. SERIE DE LAS DEFUNCIONES Y PRONOSTICO DEL MODELO ARIMA(1,2,1)

Las defunciones pronosticadas a través del Modelo ARIMA(1,2,1), se muestra como los pronósticos están muy cerca de las observaciones en cada años, incluso, para los últimos tres años estuvieron casi iguales. Otro hecho muy importante es ver como los valores ajustado, a través del modelo, están muy cerca de la serie original y dentro de los límites de confianza, por lo tanto se puede considerar que los pronósticos son bastante confiables.

A continuación presentamos una comparación de las defunciones oficiales registradas por Estadísticas Vitales de la Contraloría comparadas con los pronósticos publicados por la Contraloría General de la República de Panamá y las defunciones pronosticadas por el método Box – Jenkins.

Tabla 1. COMPARACIÓN DE LAS DEFUNCIONES PRONOSTICADAS POR EL INEC Y POR PRONOSTICADAS UTILIZANDO LA METOLOGÍA BOX-JENKINS

Años	Defunciones Reales	Defunciones pronosticadas		Comparación de las defunciones Reales y las Pronosticadas	
		CGR	Box- Jenkins	INEC	Box- Jenkins
1990-1995	52510	66,881	52226	-14,371	284
1995-2000	58943	72,472	59677	-13,529	-734
2000-	65773	78,450	64503	-12,677	1270

2005					
2005-2010	76288	86,054	75775	-9,766	513

CONCLUSIONES

Al comparar los resultados obtenidos entre los diferentes métodos de pronósticos, los oficiales de la Contraloría General de la República (CGR) y el Box-Jenkins durante cuatro diferentes períodos (1990-1995, 1995-2000, 2000-2005 y 2005-2010), encontramos que se reflejan grandes diferencias entre ambos métodos.

Los resultados del método de la CGR reflejan valores pronosticados por encima de los valores reales para cada período. Dichas diferencias se encuentran dentro de un rango de 9 766 a 14,400 por encima de los valores reales. Sin embargo, los resultados del método Box-Jenkins sugieren procedimientos metodológicos más acertados, ya que las diferencias son relativamente menores con valores dentro de un rango de -734 a 1270.

Lo anterior nos demuestra que la metodología Box-Jenkins es una forma de pronosticar valores y resultados con un menor error.

Siempre y cuando se siga los lineamientos de la metodología, se logra obtener series pronosticadas con un alto grado de certeza, las cuales nos pueden servir para la toma de decisiones de hoy con datos del pasado.

REFERENCIAS BIBLIOGRAFICAS

Contraloría General de la República (2012); Estimaciones y proyecciones de la población de la República, por provincia y comarca indígena, según sexo y edad: Años: 2000-30. Instituto Nacional de Estadística y Censo

Contraloría General de la República (2010); Panamá en Cifras: 2005-2009. Instituto Nacional de Estadística y Censo

Contraloría General de la República (2010): Estadísticas Vitales – Defunciones: año 2009. Instituto Nacional de Estadística y Censo

Contraloría General de la República (2009); Estadísticas Vitales –Defunciones año 2008.

Instituto Nacional de Estadística y Censo.

Contraloría General de la República; Estadística Vitales – Defunciones.

Boletín Estadístico. Año 1952-2007. Instituto Nacional de Estadística y Censo

Contraloría General de la República (2000); Proyección de Defunciones. Boletín Estadístico. Instituto Nacional de Estadística y Censo.

De La Fuente-García, D (1995) Análisis Comparativo de Cálculo de previsiones univariadas y Función de transferencia, mediante las metodologías de Box-Jenkins y redes neuronales. *Questio* 19 187-215

Villa, A; Escobedo, M; Mendez-Sanchez, N; (2004) Estimación y proyección de la prevalencia de obesidad en México a través de la mortalidad por enfermedades asociada. *Gac Médica*, (14) suplemento 2

Oderica-Mellado, M; (2004) Pronóstico de las defunciones por medio de los modelos autorregresivos integrados móviles. (10) 42. <http://www.redalyc.org/pdf/112/11204209.pdf>

Gujarati, D. (2003) *Econometría*. Editorial Mc.Graw Hill. 4ta edición México

Hanke, J (2006) *Pronóstico de los negocios*. Editorial Prentice Hall. 5ta edición

Makridakis, W (1998) *Métodos de Pronósticos*. Editorial Mc.Graw Hill. 2da edición. México

Schmidt, S. (2005) *Econometría*. Edición Mc.Graw Hill. México.