

ESTIMACION DE LA ESPERANZA DE VIDA MEDIANTE MODELOS GAMLSS

LIFE EXPECTANCY ESTIMATION THROUGH GAMLSS MODELS

VALENTINA GARCÍA V^{1,a}, VALENTINA LÓPEZ M^{1,b},
VALENTINA SERRATO H^{1,c}, MARIO CESAR JARAMILLO E^{1,d}

¹ESCUELA DE ESTADÍSTICA, UNIVERSIDAD NACIONAL DE COLOMBIA, MEDELLIN, COLOMBIA

Resumen

La esperanza de vida se define como la cantidad de años que puede esperar vivir un recién nacido, suponiendo que los problemas de mortalidad más importantes a la hora de su nacimiento permanecen iguales a lo largo de su vida. Ésta, en un determinado año t , se define como el promedio de la edad de las personas fallecidas en ese año [1]. Generalmente se encuentran cantidad de estudios, artículos y publicaciones los cuales tienen como tema central la estimación de esperanza de vida, sin embargo no se consideran factores de tipo poblacionales, geográficos, ambientales, de salud e incluso económicos. Este trabajo tiene como finalidad utilizar los modelos GAMLSS (Generalized Additive Model for Location Scale and Shape) [2] para determinar los factores más influyentes en el tiempo que espera vivir una persona. Se postulan diferentes modelos para determinar cuál es el más adecuado para abordar el problema, los resultados de este estudio sirven para cuantificar los efectos de las variables en la estimación de la esperanza de vida, en particular se concluye que la región, el sexo, el ingreso y estado del país están altamente relacionadas con la esperanza de vida.

Palabras clave: Esperanza de vida, regresión, GAMLSS, índice de desarrollo humano.

Abstract

Life expectancy is defined as the amount of years that a newborn could expect to live. Assuming that all the mortality problems at his birth time remain constant. This is calculated as the average of the people who died the same year. Usually, it is easy to find a lot of investigations and articles about

^aUniversidad Nacional de Colombia. E-mail: vagarciave@unal.edu.co

^bUniversidad Nacional de Colombia. E-mail: valopezmo@unal.edu.co

^cUniversidad Nacional de Colombia. E-mail: vserrato@unal.edu.co

^dUniversidad Nacional de Colombia. E-mail: mcjarami@unal.edu.co

life expectancy but without including geographical, environmental, population and health factors. This work has the purpose of the use of GAMLSS models for make a decision about the most influent factors in the life expectancy. Several models were proposed to determine which one is the best to handle the problem. The results help to quantify the effects of the variables in the life expectancy. Specifically the conclusion is that the region, gender, income and the country development are highly related with the time a person could expect to live.

Key words: Life expectancy, lifequality, index of human development, GAMLSS, regression.

1. Introducción

Es importante estimar la esperanza de vida, ya que ésta junto con el índice de educación y el PIB componen el índice de desarrollo humano (IDH), que tiene como objetivo evaluar la calidad de vida de los países del mundo [3]. En éste trabajo se consideran varios modelos de regresión lineal para analizar si el género, la ubicación geográfica, las condiciones sociales y de salud inciden en el tiempo que se espera que viva una persona. Particularmente se concluye que el modelo GAMLSS es el más adecuado para abordar este prolema, ya que permite modelar la variabilidad de la esperanza de vida en función de los demás factores.

2. Objetivo

Definir los factores más influyentes en el tiempo que espera vivir una persona, considerando variables de tipo económicas , poblacionales, geográficos, ambientales, y de salud, esto estimando un modelo de regresión y así determinar cuáles de las variables consideradas son significativas, ver cómo y en qué cantidad éstas afectan la esperanza de vida de una persona.

3. Metodología

Inicialmente se consideran diferentes bases de datos con indicadores de desarrollo mundial provenientes del Banco Mundial (*World Bank*) [4], las cuales se escogen a consideración de los autores, con éstas se construye una base de datos completa, esto se hace mediante el programa de programación *R* [5]. Se realiza un análisis descriptivo donde específicamente se destacan el gráfico de densidades y el boxplot, ya que proporcionan gran cantidad de información relevante. Posteriormente se plantean diferentes modelos de regresión lineal múltiple, teniendo en cuenta variables económicas, poblacionales, geográficas, ambientales y de salud. La elección del mejor modelo se fundamenta en la significancia de las variables, el coeficiente de determinación R^2 , y el cumplimiento de los supuestos de los residuales.

4. Gráficas

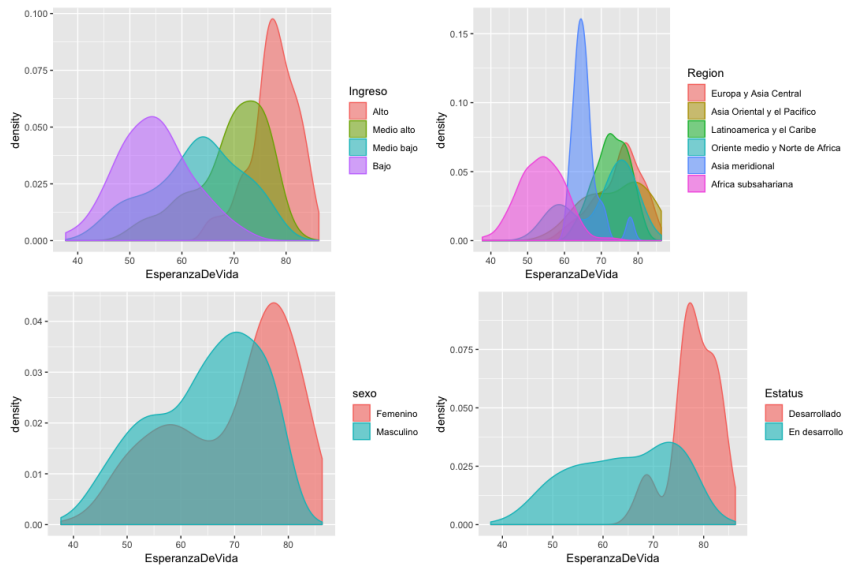


FIGURA 1: Gráfico de densidades

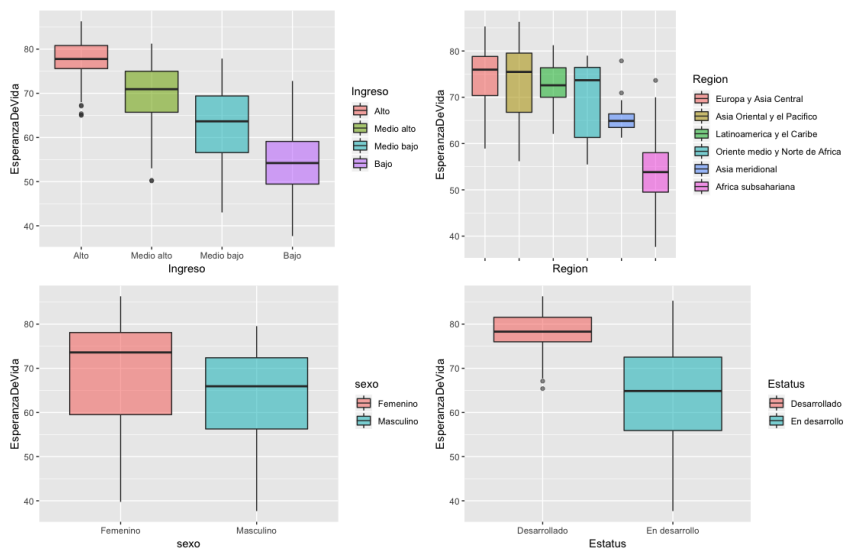


FIGURA 2: Boxplot

5. Postulación de modelos

En un principio se construyeron varios modelos con el fin de encontrar el que las variables explicaran de mejor manera la esperanza de vida, se llegó entonces al siguiente modelo:

PARÁMETRO	ESTIMACIÓN
Intercepto	93.3593
África Subsahariana	-4.9151
Asia Oriental y el Pacífico	0.5877
Europa y Asia central	-0.1677
Latinoamérica y el Caribe	-0.1421
Oriente medio y Norte de África	-0.4195
Estatus desarrollado	-0.7619
Ingreso medio bajo	-0.6821
Ingreso medio alto	1.2517
Ingreso alto	1.4777
Sexo masculino	-5.5568
Incidencia VIH	-4.5622
Nacimientos	-0.4106
Enfermedades	-0.5463

Se considera que la varianza está en función de algunas covariables y se obtuvo la siguiente estimación:

$$\widehat{\sigma^2} = -1.5797 + 0.0929Enfermedades + 0.0546Nacimientos$$

Con un pseudo $R^2 = 0.95$, lo cual indica que el modelo explica un 95 % la variabilidad de la esperanza de vida.

6. Conclusiones

- Por cada nacimiento en 1000 personas, la esperanza de vida disminuye en promedio 0.42 años aproximadamente, considerando las demás variables constantes.
- La región que más afecta la esperanza de vida es África subsahariana, la cual disminuye en casi 5 años, por otro lado nacer en Asia oriental y el pacífico aumenta en 0.58 años aproximadamente, considerando las demás variables constantes.
- Cuando los ingresos del país van de medio altos a altos, la esperanza de vida aumenta aproximadamente en 1.25 años y 1.48 años respectivamente, considerando todo lo demás constante.

- Los hombres tienen en promedio 5.5 años menos que las mujeres en su esperanza de vida.
- Cuando los casos de VIH aumentan, la esperanza de vida disminuye en promedio 4.5 años, manteniendo lo demás constante. Y en caso de que las enfermedades por cualquier problema cardiovascular, cáncncer, diabetes o problemas respiratorios crónicos aumenten, la esperanza de vida disminuye en medio año, manteniendo lo demás constante.

Referencias

- [1] Wikipedia, «Esperanza de vida,» Fundación Wikipedia.inc, 5 06 2019. [En línea]. Available: <https://es.wikipedia.org/wiki/Esperanza-de-vida>.
- [2] A. Stacinopoulos, B. Rigby y A. Akantzilioton, «Instructions on how to use the GAMLSS package in R.,» 2008.
- [3] UNDP, «Programa de las naciones unidas para el desarrollo,» 2019. [En línea]. Available: <https://www.undp.org/content/undp/es/home/>.
- [4] «The World Bank,» [En línea]. Available: <http://datatopics.worldbank.org/world-development-indicators/themes.html>. [Último acceso: Abril 2019].
- [5] R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.