WEP Warwick Evans Publishing

Transactions on Materials, Biotechnology and Life Sciences
ISSN: 2960-1932, eISSN: 2960-2319 | Volume 1
FBB 2023

Investigating Human Expectancy by Linear Regression

Linhui Gao

School of Jinan New Channel school, Jinan, China

Abstract. In recent years, many people have studied the human lifespan, some people study how to extend the human lifespan, explore the limits of the human lifespan, and so on. But what are the factors that affect human lifespan, and will they lead to an increase or decrease in lifespan? This paper was analyzed by using linear regression, where specific analyses were performed using R-square and multiple R, mainly studies the influence of three factors, BMI, Schooling, and HIV/AIDS, on human life expectancy between developing and developed countries through linear regression methods.

Keywords: life expectancy, linear regression, status, alcohol, BMI, HIV/AIDS, Schooling.

1. Introduction

Driven by technological advances, human life expectancy has increased dramatically since the 19th century. Demographic evidence suggests that mortality rates in old age continue to decline and the maximum age of death rises, which may gradually extend human life spans. There are five factors that could have an impact on human longevity. Firstly, AIDS can have a negative impact on longevity by compromising the immune system and making it difficult for the body to fight off other diseases [1]. In addition, there are some positive factors.

Firstly, BMI Body mass index (BMI) is a person's weight in kilograms (or pounds) divided by the square of their height in meters (or feet). The National Health Department recommends that the BMI of adults in China should be maintained between 18.5 (kg/㎡) and 25 (kg/㎡) [2]. It is often used as a quick indicator of physical health and, importantly, the risk of chronic disease. People with a very low or very high BMI are at higher risk [3]. The status indicator refers to the comparison of developed countries and developing countries. Differences in life expectancy due to income inequality, differences in public health systems, and levels of education in developing and developed countries [4]. Schooling refers to the level of education of an individual. A major synthesis of the existing literature.

2. Literature review

Md. Nazrul Islam MONDAL and Mahendran SHITAN(2015) research the impact of socio-health factors on life expectancy in low and lower middle-income countries [8].Joshua E. Seifarth BHK, Cheri L. McGowan PhD, Kevin J. Milne PhD (2012) used PubMed, ISI Web of Knowledge, and Google Scholar to demonstrate that there are indeed biological differences between the sexes in both sociology and biology, including differences in genetic and physiological factors. biological factors that may explain, at least in part, the female advantage in human life expectancy[9]. Kenneth G. Manton, Eric Stallard and H. Dennis Tolley (2012) research evidence the life expectancy of selected populations with good health behaviours was examined and multivariate risk factor models were applied to longitudinal data. The risk factor models and the population data provided consistent estimates of the lower bound of the theoretical limits of human life expectancy[10]. Uma Murti,Mohd Shahidan Shari - Paul Anthony, Mariadas Norazella Zainol Abidin (2021) A panel ARDL approach was used and PMG estimator was selected for the study. The results show that economic growth, population growth and healthcare expenditures can have a significant positive impact on life expectancy, but CO_2 emissions can have a significant negative impact on life expectancy[11]. Tadele Girum, Mulugeta Shegaze, and Ebrahim Muktar(2018):Life expectancy and socioeconomic, medical,

and morbidity indicator factors were shown to be significantly correlated. Life expectancy was positively correlated with the human development index, adult literacy rate, contraceptive prevalence rate, HIV incidence rate, and tuberculosis incidence rate, and negatively correlated with the underage dependency ratio, total fertility rate, and child mortality rate.[13].

3. Significant

This paper calculates the effects and correlations of status and the factors BMI, Schooling, and HIV/AIDS on life expectancy using simple linear regression and multiple linear regression, respectively, based on time-sensitive data.Readers can adjust life expectancy according to this article by adjusting for four factors: STATUS, SCHOOLING, BMI, and HIV/AIDS.

4. Innovations and shortcomings:

Literature data is weak in timeliness, this paper uses the latest data in recent years to analyse the impact of the four factors on the expected life more comprehensively. Many domestic and international studies the limits or trends of life expectancy, the impact of the economy on life expectancy or the impact of disease on life expectancy. But in this paper, the impact of AIDS on life expectancy is analysed in a more detailed way, and simple linear regression and multiple linear regression are used to study BMI, schooling and status respectively.

The disadvantage is that some of the data analyses are incomplete due to my limited academic skills and still shallow analytical skills.

5. Methodology:

Comparison and induction method: this paper analyses the data changes of different forms of changes and trends of icons, R-square and multiple square, etc. inductively to find out the differences in the changes of these variables.

simple linear regression：Simple linear regression is a model with a single independent variable, X, which has a relationship with the dependent variable, Y. Mainly used in this paper to analyse the relationship between status and life expectancy

multiple linear regression：a statistical method used to describe the simultaneous relationships of numerous factors with one continuous outcome[12].This paper focuses on the use of this method to analyse the impact of status, Schooling, BMI and HIV on life expectancy.

6. Data Analysis:

Simple linear regression is a model with a single independent variable, X, which has a relationship with the dependent variable, Y. The relationship is:

$$y = \beta 0 + \beta 1\, x + \varepsilon$$

this formula is a type of linear regression equation, where y within the equation represents the dependent variable that you want to predict, x is the independent variable, and β0 and β1 are the coefficients to be estimated, which will be calculated.

The first study is about will investigate how the status variable affects the dependent variable with regard to human life expectancy. The β0 and β1 were calculated by the ordinary least square method and the following equation was obtained.

$$Life\ Expectancy = 68.09 + 10.67Status \quad (1)$$

In particular, country status is a dummy variable that is set to 1 if the country is a developed country and 0 otherwise. This regression equation illustrates that the more developed the country, the longer the human life expectancy, which is expected to increase by 10.67 years on average. when the country is underdeveloped and the independent variable is not available, the expected human life expectancy is 68.09 on average.

Table 1. Data on the correlation between human longevity and status by simple linear regression]

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.56960351
R Square	0.32444816
Adjusted R Sc	0.32101896
Standard Erro	7.73391414
Observations	199

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	5659.1545	5659.154497	94.61344539	1.6426E-18
Residual	197	11783.2453	59.81342793		
Total	198	17442.3998			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	68.093	0.77339141	88.04468057	3.8407E-160	66.567811	69.618189	66.567811	69.618189
Status	10.6655859	1.09649912	9.726944299	1.6426E-18	8.50320293	12.8279688	8.50320293	12.8279688

From table 1 we can see that the R square is 0.32, meaning that 32% of the difference in life expectancy can be explained by the status of the country. In addition, multiple R represents a correlation of 0.57 between status and human lifespan.

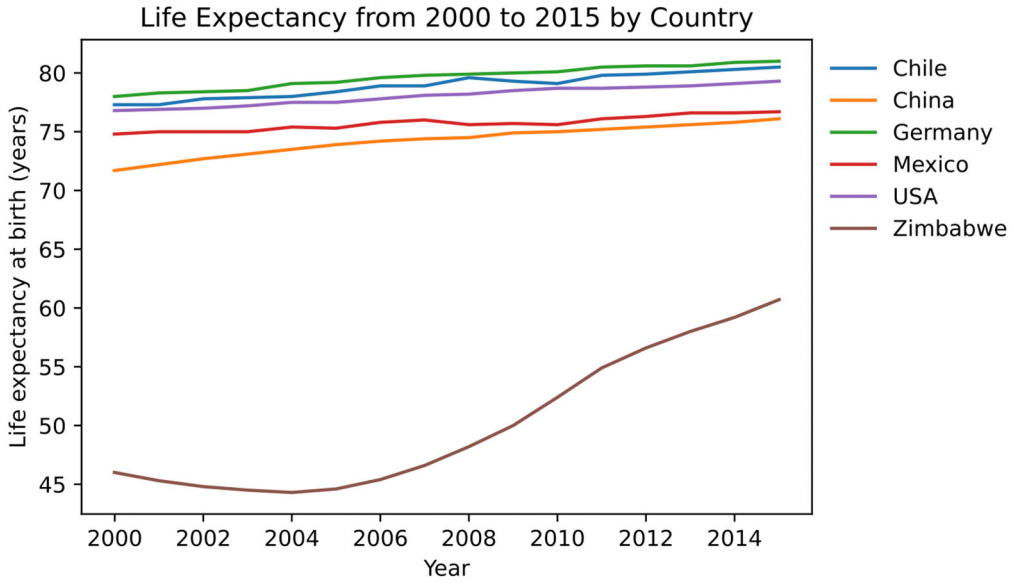

Figure 1. Trends in human life expectancy in six countries between 2000-2015 [5].

Over time, we can see from graph 1 that the average human life expectancy in the three developed countries of intelligence, Chile, Germany and the United States is around 77 to 81 years. The two developing countries Mexico, and Zimbabwe, have an average human life expectancy of 75 years and under. Developed countries as a whole have a higher life expectancy than developing countries.

Observational data suggest that life expectancy is also influenced by other factors at the same time. It is therefore important to control for variables that may affect human life expectancy. Model 1 is

extended by including additional regression variables: number of deaths from AIDS per 1,000 population over last 4 years old born alive (HIV/AIDS); years of schooling attended (Schooling); Average body mass index of the total population (BMI).

$$Life\ Expectancy = 61.54 + 3.71Status - \frac{7.86HIV}{AIDS} + 0.61Schooling + 0.09BMI \quad (2)$$

Table 2. Data on the correlation between human longevity, Status, AIDS, Schooling, and BMI by multiple regression analysis.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.88986701
R Square	0.7918633
Adjusted R Sc	0.78757182
Standard Err	4.32590108
Observations	199

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	13811.9963	3452.999073	184.5199353	5.9001E-65
Residual	194	3630.40351	18.71342014		
Total	198	17442.3998			

	Coefficients	*Standard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	61.5385532	1.25661369	48.97173545	3.6695E-111	59.0601749	64.0169316	59.0601749	64.0169316
HIV/AIDS	-7.858331	0.56229238	-13.97552476	3.6527E-31	-8.967322	-6.74934	-8.967322	-6.74934
Schooling	0.60583554	0.0996265	6.08106828	6.23032E-09	0.40934543	0.80232565	0.40934543	0.80232565
Status	3.71092379	0.74797856	4.96127029	1.52649E-06	2.23570997	5.18613762	2.23570997	5.18613762
BMI	0.09066871	0.01959835	4.626344769	6.79233E-06	0.05201553	0.12932189	0.05201553	0.12932189

From Table 2, the R square is 0.79, meaning that the 79% difference in life expectancy can be explained by the status, AIDS, Schooling, and BMI of the population. In addition, multiple R represents a correlation of 0.89 between status, human lifespan, AIDS, Schooling, and BMI. In addition, it can be seen that the R square of the multiple regression model increased from 0.57 to 0.89 compared to the previous simple regression model, which means that the multiple regression model is more suitable for this study.

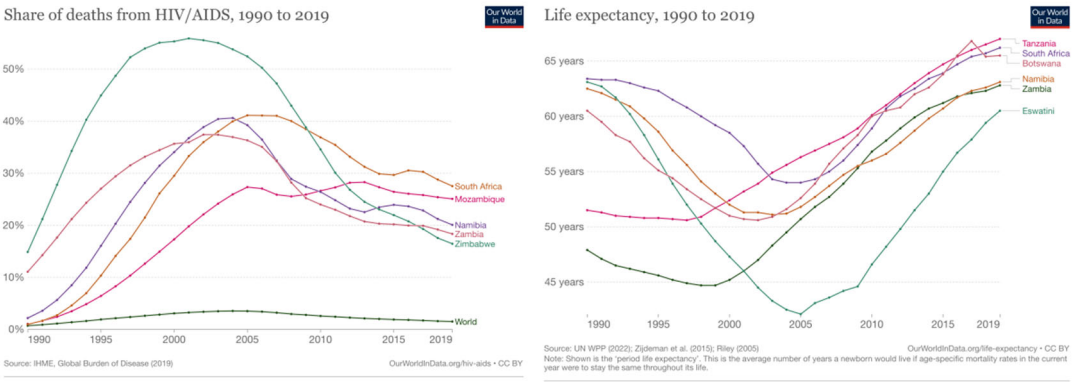

Figure 2. Share of deaths from HIV/AIDS, from 1990 to 2019 on the left, and Life expectancy, from 1990 to 2019 on the right [6].

The first chart shows the proportion of AIDS-related deaths in each country, and the second chart shows the projected human life expectancy for each country from 1990 to 2019 [6]. Comparing the two charts, it is clear that while the proportion of AIDS-related deaths is increasing, the human life expectancy in the country is decreasing, which is inversely proportional.

7. Conclusion:

Calculations show that among the four variables, three of them have a positive effect on human longevity: status, BMI, and schooling. Status has the having the greatest impact on longevity (3.71), followed by schooling (0.61) and BMI (0.09). The last variable, AIDS, has a negative effect on human longevity.

References

[1] https://www.healthline.com/health/hiv-aids/life-expectancy. Madell, Robin. "Facts about HIV: Life Expectancy and Long-Term Outlook." Healthline, Healthline Media, 27 Apr. 2018, www.healthline.com/health/hiv-aids/life-expectancy.

[2] https://www.cdc.gov/obesity/basics/adult-defining.html. Centers for Disease Control and Prevention. "Defining Adult Overweight and Obesity." Centers for Disease Control and Prevention, 2021, www.cdc.gov/obesity/basics/adult-defining.html.

[3] https://www.cdc.gov/healthyweight/assessing/bmi/index.html. CDC. "Body Mass Index (BMI)." Centers for Disease Control and Prevention, 2019, www.cdc.gov/healthyweight/assessing/bmi/index.html.

[4] https://gradebuddy.com/doc/3139338/life-expectancies-in-developing-and-developed-countries/#:~:text=This%20essay%20provides%20a%20survey%20of%20three%20developmental,public%20health%20systems%2C%20income%20inequalities%20and%20education%20levels. "Life Expectancies in Developing and Developed Countries - D3139338 - Gradebuddy." Gradebuddy.com, gradebuddy.com/doc/3139338/life-expectancies-in-developing-and-developed-countries/#:~:text=This%20essay%20provides%20a%20survey%20of%20three%20developmental. Accessed 1 Aug. 2023.

[5] https://max-torch.github.io/2021/05/14/GDP.html. "Life Expectancy and GDP - Alexander Lacson." Max-Torch.github.io, max-torch.github.io/2021/05/14/GDP.html. Accessed 1 Aug. 2023.

[6] https://ourworldindata.org/hiv-aids#citation. Roser, Max, and Hannah Ritchie. "HIV / AIDS." Our World in Data, 3 Apr. 2018, ourworldindata.org/hiv-aids#citation.

[7] https://www.kaggle.com/datasets/amirhosseinmirzaie/countries-life-expectancy. "Countries Life Expectancy." Www.kaggle.com, www.kaggle.com/datasets/amirhosseinmirzaie/countries-life-expectancy. Accessed 1 Aug. 2023.

[8] MONDAL, Md. Nazrul Islam, and Mahendran SHITAN. "Impact of Socio-Health Factors on Life Expectancy in the Low and Lower Middle Income Countries." Iranian Journal of Public Health, vol. 42, no. 12, 1 Dec. 2013, pp. 1354–1362, www.ncbi.nlm.nih.gov/pmc/articles/PMC4441932/.

[9] Seifarth, Joshua E., et al. "Sex and Life Expectancy." Gender Medicine, vol. 9, no. 6, Dec. 2012, pp. 390–401, https://doi.org/10.1016/j.genm.2012.10.001. Accessed 25 Nov. 2019. https://doi.org/10.1016/j.genm.2012.10.001.

[10] Manton, Kenneth G., et al. "Limits to Human Life Expectancy: Evidence, Prospects, and Implications." Population and Development Review, vol. 17, no. 4, 1991, pp. 603–37. JSTOR, https://doi.org/10.2307/1973599. Accessed 8 Aug. 2023.https://muse.jhu.edu/article/485100.

[11] Uma, Murthy, et al. "The Relationships between CO 2 Emissions, Economic Growth and Life Expectancy." Noorazeela Zainol ABIDIN / Journal of Asian Finance, vol. 8, no. 2, 2021, pp. 801-0808, www.koreascience.or.kr/article/JAKO202104142246630.pdf, https://scholar.kyobobook.co.kr/article/detail/4010028156025

[12] Eberly, Lynn E. "Multiple Linear Regression." Methods in Molecular Biology (Clifton, N.J.), vol. 404, 2007, pp. 165–187, pubmed.ncbi.nlm.nih.gov/18450050/#:~:text=This%20chapter%20describes%20multiple%20linear, https://doi.org/10.1007/978-1-59745-530-5_9.

[13] Girum, T., Muktar, E., and Shegaze, M. (2018). Determinants of life expectancy in low and medium human development index countries. Medical Studies/Studia Medyczne, 34(3), pp.218-225. https://doi.org/10.5114/ms.2018.78685. https://www.termedia.pl/Determinants-of-life-expectancy-in-low-and-medium-human-development-index-countries,67,33871, 0, 1.html.