Modeling Fertility Rate from Life Expectancy

Olivia Rueschhoff and Dr. Amy Ekanayake

Department of Mathematics and Philosophy, Western Illinois University

Abstract

In this research project, several empirical models were examined to best describe the relationship between the average life expectancy and the average fertility rate in the United States from 1950 – 2023. Both data sets were from Macrotrends.net. The models developed to reflect the data include polynomials, logarithmic, exponential, and power models. We used a 3:1 train-test split to assess the model predictions using the smallest maximum absolute error and the smallest squared sum of residuals. In addition, the trend of the data and the R-squared values were utilized to evaluate the best fit model for the training set. We observed that the best model to represent the data is the sixth-order polynomial. It performed well on the test data: it had the smallest maximum absolute error and the smallest squared sum of errors. It also closely follows the shape of the training data and has the highest R-squared value for the training data. Based on these findings, we conclude that polynomial models are well-suited to describe the average fertility rate per year in terms of the average life expectancy per year in the United States.

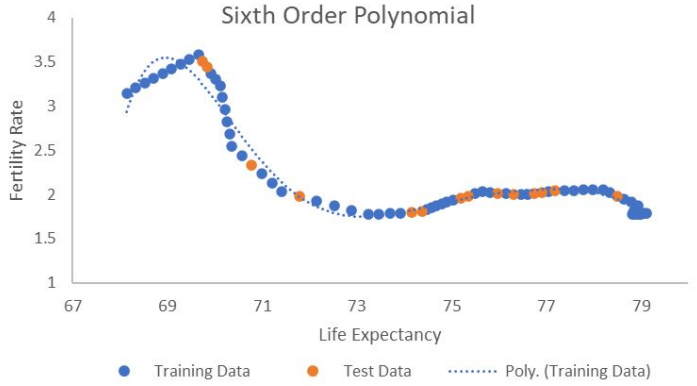
Introduction

An empirical model is an equation created from the results of known data with the goal of determining which model captures the most data points. This research project used empirical models to describe the relationship between the United States Average Life Expectancy compared to the Average Fertility Rate from 1950 – 2023. The empirical models that have been developed and compared to the data were discovered using Microsoft Excel. The 9 models found were then tested for the best model using the lowest maximum absolute value of the residuals and the lowest squared sum of residuals as criteria using a 3:1 train-test split where we removed 25% of the data and tried to see which model could most accurately fill in the missing data. Finding a relationship between these two data sets was interesting because it is not something we would normally associate with each other, but they clearly have a relationship.

Test Table 1

Model Type	m	s
1st Order Polynomial	0.64745	1.92337
2nd Order Polynomial	0.72002	2.39679
3rd Order Polynomial	0.57479	0.93255
4th Order Polynomial	0.42443	0.73966
5th Order Polynomial	0.30738	0.40195
6th Order Polynomial	0.2675324	0.17725
Logarithmic	0.635903	1.871467
Exponential	0.712443	1.77844
Power	0.702693	1.728919

Table 1: Table 1: the maximums and sums of each test set found

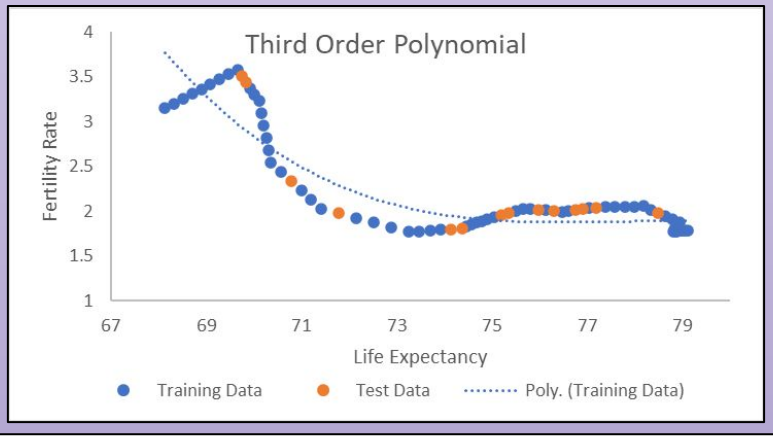


Acknowledgements

Special thanks to Dr Amy Ekanayake and Dr Dinesh Ekanayake, the Department of Mathematics and Philosophy, and the Centennial Honors College for making this project possible!

Equation for Sixth Order Polynomial, Best Empirical Model for the Relationship:

$$F = -0.00009105857152L^6 + 0.04083087920548L^5 - 7.62545555706768L^4 + 759.196917911132L^3 - 42{,}498.2172505625L^2 + 1{,}268{,}185.83180976L - 15{,}760{,}604.3434054$$



Results/Conclusion

Of all the models tested to find the relationship between the data, one model was found to describe the data the best. The best model is the sixth order polynomial model because it has the lowest errors from the test data's results both times, the largest R-Squared value, and captures the shape of the data most closely. The sixth order polynomial makes sense logically for being the best model as well; since the more parameters that are able to be adjusted in an equation, the better the model can more closely capture all of the data. With this logic it is very likely that a higher degree polynomial would be an even better model for the data and would correct the 3.33% unexplained data within the sixth order polynomial. Overall, out of all nine models tested, the sixth order polynomial best captures the relationship between the average fertility rate and average life expectancy within the United States from 1950 to 2023.

Reference

U.S. Fertility Rate 1950-2023. (2023). Retrieved from Macrotrends.net: https://www.macrotrends.net/countries/USA/united-states/fertility-rate

U.S. Life Expectancy 1950-2023. (2023). Retrieved from Macrotrends.net: https://www.macrotrends.net/countries/USA/united-states/life-expectancy