# Parameter Estimation of Fractional Trigonometric Polynomial Regression Model

TAIWO ABASS ISHOLA
Department of Mathematical Sciences,
Faculty of Science
Olabisi Onabanjo University, Ago-Iwoye
Ogun State, NIGERIA

OLATAYO TIMOTHY OLABISI
Department of Mathematical Sciences,
Faculty of Science
Olabisi Onabanjo University, Ago-Iwoye
Ogun State, NIGERIA

ADESANYA KAZEEM KEHINDE
Department of Health Information Management,
Ogun State College of Health Technology, Ilese-Ijebu
Ogun State, NIGERIA

*Abstract:* - Fractional Trigonometric Polynomial Regression is a form of non-linear regression in which the relationship between the outcome variable and risk variable is modelled as 1/nth degree polynomial regression by combining the function of $\sin(nx)$ and $\cos(nx)$ on the value of natural numbers. The model was used to analyse the relationship between three continuous and periodic variables. Coefficients of the model were estimated using the Maximum Likelihood Estimate (MLE) method. From the results, the model obtained indicated that an increased in body mass index will increase the level of blood pressure while age may or may not have an influence on the blood pressure level. The values of Coefficient of variation $(R^2)$ showed the variation in the dependent variable was well explained by the independent variables and the value of adjusted $(R^2)$ showed the model had a good fit with a high level of predictive power.

*Key-Words:* - Fractional regression, Polynomial, Trigonometric function, Periodic variation, Continuous variable.

## 1 Introduction

**S**tatistical models have turned into an important and critical instrument utilized to study physical marvels where biology is not excluded [1]. Long, Statistics and Mathematics have been viewed as an important avenue through which insights into procedure and processes in the mission to give answers to numerous inquiries that existing in biology [2]. Prognostic models are useful techniques that assist in decision-making and in this model patient characteristics are been used to foresee clinical results [3]. However, this helps in the administration of future patients to forestall unfavourable occasions. Nevertheless, chosen the risk variables that will be used as predictors becomes very important [4]. Moreover, in order to determine the relationship between continuous variables x and outcome variable (y), regression analysis is the initial choice. This is based on the straight line $\beta_0 + \beta_1 x$, where x is the risk variables and the level of linearity of the risk variables relied on the kind of study [5]. Conversely, challenges emerge when the assumption of linearity is observed to be illogical and an appropriate is required [6]. Two prominent and adaptable ways that allowed smooth nonlinear relationship are Splines and Fractional polynomials [7]. [8] presented Fractional polynomials as an expansion of polynomial models for deciding the functional form of a continuous predictor. These models are appropriate for nonlinear variables and these models have been utilized in numerous applications including survival and meta-relapse investigation [9].

As indicated by [10], if the infinite variable t in the straight-line model is transformed, the 1st order Fractional polynomial model is denoted by

$$Y = \beta_0 + \beta_1 X^P \quad (1)$$

The power p is selected from the accompanying set: -2. -1, -0.5, 0, 0.5, 1, 2, 3,.. with $t^o = \log t$. while the 2nd order fractional polynomial is denoted by

$$Y = \beta_0 + \beta_1 X^{P_1} + \beta_2 X^{P_2} \quad (2)$$

[11] utilized a statistical method that dependent on Fractional polynomials for the examination of potential prescient variables and deduced that the investigation of a continuous factor with Fractional polynomials gives more information about such factors, improve the statistical power by recognizing persuasive factors. Fractional polynomial models and the steps used to construct them have the fascination of effortlessness that has commanded them to several applied methodologists and clarified their use in applied research [12], [13], and [14]. In the Mathematical subfields of Numerical and Mathematical examination, a trigonometric polynomial is a finite linear combination of functions sin(nx) and cos(nx) with n assuming the values of one or more natural numbers [15]. Therefore, a trigonometric polynomial with function t is of the structure:

$$y(t) = \beta_0 + \alpha_i \cos \omega x_t + \beta_j \sin \omega x_t + e_t \quad (3)$$

where the parameters $\beta_0$, $\alpha_i$ and $\beta_j$ are real numbers, trigonometric functions, $\sin(\omega x)$ and $\cos(\omega x)$ are periodic over time with a period of $2\pi/\omega$ [16]. That is, $\sin(\omega x)$ is the same as $\sin[\omega(t + (2\pi/\omega)j)]$ $for\ j = 1,2, ....$ [17] discussed about polynomial–trigonometric regression model and call attention that the utilization of just trigonometric terms can create a poor fit at the endpoints of the interval on x when the genuine regression function on x is not periodic. Appropriately, they advance the utilization of both polynomial and trigonometric terms as a method for expelling this potential issue. [18] used polynomial terms with other nonlinear terms and get a superior model which would not have been acquired if just polynomial terms were utilized. This closely resembles what is frequently done in time series analysis when an autoregressive moving average (ARMA) model is fit with fewer terms as opposed to fitting a higher order [19]. In the event that a solitary regressor and the time plot demonstrated some proof of periodicity that is a cyclic patterned, the utilization of trigonometric terms might be progressively helpful [20]. Time series dataset that exhibits periodic variations can be found in many diverse areas and analysing this dataset involves the use of Fractional and Trigonometry Polynomial Regression that is capable of obtaining the functional forms of a continuous and periodic dataset. Therefore, this research will be used to propose a Fractional Trigonometric Polynomial regression model that can be used to analyse variables that are continuous and periodic in nature simultaneously. The parameters of the model will be obtained using Maximum Likelihood estimation method. The stability of the model will be tested using Dublin Watson Statistic while Coefficient of determination and Adjusted Coefficient of determination will be used to determine the level of variation explained and the predictive ability of the model respectively.

# 2 Methodology

## 2.1 Fractional Polynomial Trigonometric Regression

The Fractional Polynomial Trigonometric regression model is defined as

$$Y_t = \beta_0 + \beta_1 \cos \omega x_t^{p_1} + \beta_1^* \sin \omega x_t^{p_1} + \cdots + \beta_n \cos \omega x_t^{p_n} + \beta_n^* \sin \omega x_t^{p_n} + U_i \quad (4)$$

if $p_1 = \frac{1}{r_1}, p_2 = \frac{1}{r_2}, ...$ then equation (4) becomes

$$Y_t = \beta_0 + \beta_1 \cos \omega x_{xt}^{\frac{1}{r1}} + \beta_1^* \sin \omega x_t^{\frac{1}{r2}} + \cdots + \beta_n \cos \omega x_t^{\frac{1}{rn}} + \beta_n^* \sin \omega x_t^{\frac{1}{rn}} + U_i \quad (5)$$

where $\beta_0$, $\beta_j$ are the coefficients, trigonometric functions, $\sin(\omega x)$ and $\cos(\omega x)$ are periodic over time with a period of $2\pi/\omega$, $Y_t$ is the outcome variable, $x$ is the risk variable, $p_j$ are the polynomial power and $U_i$ is normally distributed with mean $(\mu)$ and variance $(\sigma^2)$.

### 2.2.1 Parameter Estimation with Maximum Likelihood Method

The likelihood function is used on equation (5) to attain

$$L(\beta_j, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{\frac{-\frac{1}{2}[y-z]^2}{\sigma^2}} \quad (6)$$

where $z = \beta_0 + \beta_1 \cos\omega x_t^{p_1} + \beta_1^* \sin\omega x_t^{p_1} + \cdots + \beta_n \cos\omega x_t^{p_n} + \beta_n^* \sin\omega x_t^{p_n}$

The following parameters $\sigma^2, \beta_0, \beta_1, \beta_1^* \ldots, \beta_n^*, \beta_n$ is estimated by taking the log-likelihood of equation (6) and this gives

$$= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}[y - z]^2 \quad (7)$$

By differentiating equation (7) with respect to $\sigma^2$ and equate to zero

$$\frac{\partial \log(L)}{\partial\sigma^2} = n\sigma^2 - \sum_{i=1}^{n}[y - z]^2$$

then, this gives

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}[y - z]^2 \quad (8)$$

Also, by differentiating equation (7) with respect to $\beta_0, \beta_1, \beta_1^*, \ldots, \beta_n^*, \beta_n$ and equate to zero will give equations 9 to 13 as

$$\sum_{i=1}^{n} y = n\beta_0 + \beta_1\sum_{i=1}^{n}\cos\omega x_t^{p_1} + \beta_1^*\sum_{i=1}^{n}\sin\omega x_t^{p_1} + \ldots + \beta_n\sum_{i=1}^{n} cos\,\omega x_t^{p_n} + \beta_n^*\sum_{i=1}^{n}\sin\omega x_t^{p_n} \quad (9)$$

$$\sum_{i=1}^{n} y\cos\omega x_t^{p_1} = \beta_0\sum_{i=1}^{n}\cos\omega x_t^{p_1} + {} + \beta_1^*\sum_{i=1}^{n}(\cos\omega x_t^{p_1})(sin\,\omega x_t^{p_1}) + \cdots + \beta_n\sum_{i=1}^{n}(cos\,\omega x_t^{p_n})(cos\,\omega x_t^{p_1}) + \beta_n^*\sum_{i=1}^{n}(\sin\omega x_t^{p_n})(\cos\omega x_t^{p_1}) \quad (10)$$

$$\sum_{i=1}^{n} y\sin\omega x_t^{p_1} = \beta_0\sum_{i=1}^{n}\sin\omega x_t^{p_1} + \beta_1\sum_{i=1}^{n}(\cos\omega x_t^{p_1})(sin\,\omega x_t^{p_1}) + \beta_1^*\sum_{i=1}^{n}(sin\,\omega x_t^{p_1})^2 + \cdots + \beta_n\sum_{i=1}^{n}(cos\,\omega x_t^{p_n})(sin\,\omega x_t^{p_1}) + \beta_n^*\sum_{i=1}^{n}(\sin\omega x_t^{p_n})(\sin\omega x_t^{p_1}) \quad 11)$$

$$\sum_{i=1}^{n} y\cos\omega x_t^{p_n} = \beta_0\sum_{i=1}^{n}\cos\omega x_t^{p_n} + \beta_1\sum_{i=1}^{n}(cos\,\omega x_t^{p_1})(cos\,\omega x_t^{p_n} + \beta_1^*\sum_{i=1}^{n}(\cos\omega x_t^{p_n})(sin\,\omega x_t^{p_1}) + \cdots + \beta_n\sum_{i=1}^{n}(\cos\omega x_t^{p_n})^2 + \beta_n^*\sum_{i=1}^{n}(\sin\omega x_t^{p_n})(\cos\omega x_t^{p_n}) \quad (12)$$

$$\sum_{i=1}^{n} y\sin\omega x_t^{p_n} = \beta_0\sum_{i=1}^{n}\sin\omega x_t^{p_n} + \beta_1\sum_{i=1}^{n}(\cos\omega x_t^{p_1})(sin\,\omega x_t^{p_n}) + \beta_1^*\sum_{i=1}^{n}(\sin\omega x_t^{p_n})(\sin\omega x_t^{p_1}) + \cdots + \beta_n\sum_{i=1}^{n}(cos\,\omega x_t^{p_n})(sin\,\omega x_t^{p_1}) + \beta_n^*\sum_{i=1}^{n}(sin\,\omega x_t^{p_n})^2 \quad (13)$$

The combination of equation (9) to (13) can be expressed in matrix form and from this, the coefficients of the Fractional Polynomial Trigonometric regression model can be obtained

## 3 Example and analysis

The time series dataset used in this research was obtained from a cohort of patients with High blood pressure (BP) in [21]. The characteristics observed for 30 months were Body Mass Index (BMI), Age (AG) and Blood Pressure (BP). The time plot of all these characteristics was displayed in Fig. 1-3. The time plot of Blood pressure, Age and Body Mass Index in Fig.1, Fig.2, Fig.3 indicate a continuous and periodic variation and this informed the use of Fractional Polynomial Trigonometric regression model.

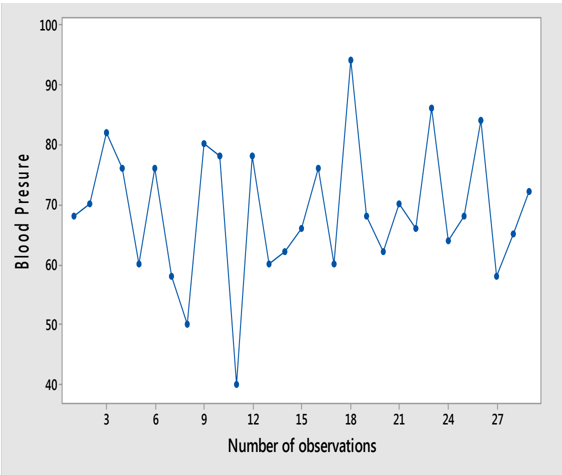


Fig.1 Time plot of blood pressure

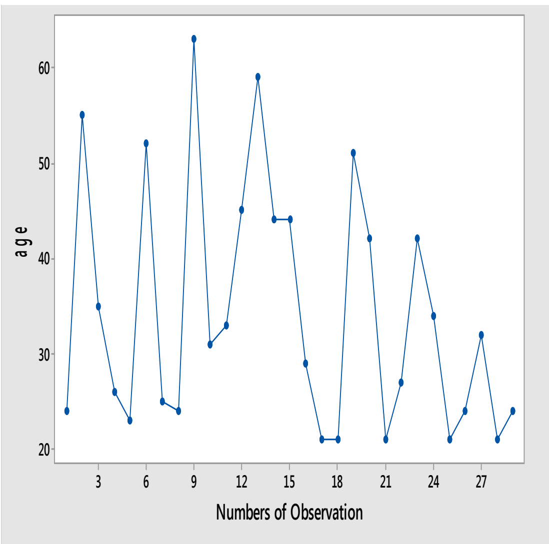


Fig.2 Time plot of age distribution

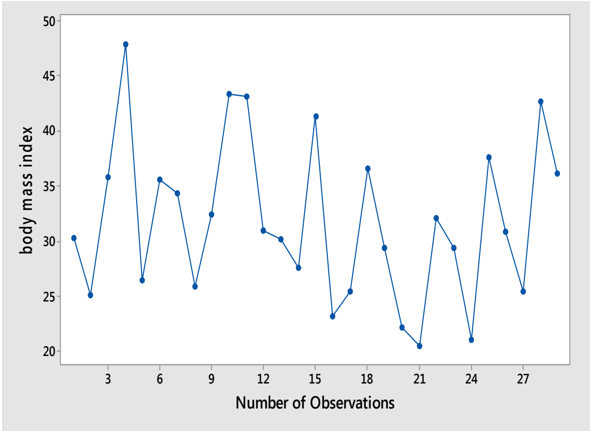


Fig.3 Time plot of body mass index

The Fractional Trigonometric polynomial regression model proposed is for analysing the continuous and periodic time series is

$$\hat{Y}_t = \beta_0 + \sum_{i=1}^{3} \beta_k cos\frac{2\pi}{12} x_t^{\;P_i} + \sum_{i=1}^{3} \beta_k{}^* sin\frac{2\pi}{12} x_t^{\;P_i}$$

$$\hat{Y}_t = \beta_0 + \sum_{i=1}^{3} \beta_k cos\frac{2\pi}{12} x_t^{\frac{1}{r_i}} + \sum_{i=1}^{3} \beta_k{}^* sin\frac{2\pi}{12} x_t^{\frac{1}{r_i}}$$

where $P_1 = \frac{1}{r_1}$, $P_2 = \frac{1}{r_2}, \dots, P_k = \frac{1}{r_k}$, $r_1 = 2, r_2 = 3$.

The proposed model is used to determine the effects of the fluctuations in body mass index and age on the level of blood pressure. The Fractional Trigonometric Polynomial regression model obtained using the Maximum Likelihood estimation method is

$$BP = 68.26472 - 4.769755\cos_{\mathrm{BMI}}$$

$$+ 0.858082\sin_{\mathrm{BMI}}$$

$$-3.815844\cos_{\mathrm{AG}}$$

$$-0.539701\sin_{\mathrm{AG}}$$

where Coefficient of Determination ($R^2$) = 0.854321, Adjusted $R^2$ 0.813375 and Durbin Watson Statistic = 1.811347.

The Fractional Trigonometric Polynomial Regression model obtained showed that the coefficient of $\sin_{\mathrm{BMI}}$, indicated that for every increase in body mass index there was a significant rise in the blood pressure level. While the coefficient of $\cos_{\mathrm{BMI}}, \cos_{\mathrm{AG}}, \sin_{\mathrm{AG}}$ indicated that a unit change may or may not indicate a reasonable reduction in Blood pressure. The values of

Coefficient of variation ($R^2$) indicated that Body Mass Index and Age explained the variations in Blood pressure up to 86% and the value of adjusted ($R^2$) at 81% showed the model is a good fit with a high-level predictive power.

# 4 Conclusion

The research article was used to propose Fractional Trigonometric Polynomial Regression model that can be used to analyse time series data that are continuous and periodic in nature. The coefficients of the model were estimated using Maximum Likelihood estimation method. The model was used to determine the nonlinear and periodic relationship between blood pressure, body mass index and Age. The results obtained indicated that an increase in body mass index will lead to a significant rise in the level of blood pressure while age may or may not have an influence on blood pressure level. The values of Coefficient of variation ($R^2$) showed the variation in the dependent variable was well explained by the independent variables and the value of adjusted ($R^2$) showed the model had a good fit with a high level of predictive power. Conclusively, there is a possibility that an uncontrolled rise in body mass index can cause a significant rise in blood pressure level while age may be a factor or not.

*References:*

[1] H. J. Cheil, J. M. McManus, R. M. Shaw, From Biology to Mathematical Models and Back: Teaching Modeling to Biology Students and Biology to Math and Engineering Students, *CBE Life Science Education*, Vol.9, No. 2, 2010, pp. 248 – 265.

[2] V. V. Ganusov, Strong Inference in Mathematical Modeling: A Method for Robust Science in the Twenty-First Century, *Frontier Microbiology*, Vol.7, 2016, pp. 1 – 10.

[3] A. M. Molinaro, M. R. Wrensch, R. B. Jenkins E. E. Passow, Statistical consideration on Prognostic for glioma, *Neuro-Oncology*, Vol. 18, 2016, pp. 609 – 623.

[4] M. R. Baneshi, F. Nakhaee, M. Law, The Use of Fractional Polynomial Models to Assess Preventive Aspect of Variables: An Example in Prevention of Mortality Following HIV Infection, *International Journal of Preventive Medicine*; Vol.4, No. 4, 2013, pp. 414 – 419.

[5] Schneider A, Hommel G, Blettner M. Linear Regression Analysis: Part 14 of a series on Evaluation of Scientific publication. *Deutsches Arztablatt International*, Vol.107, No. 44, 2010, pp. 776-782.

[6] P. Royston, G. Ambler, W. Sauerbrei, The use of Fractional Polynomials to model Continuous Risk Variables in Epidemiology, *International Journal of Epidemiology*, Vol.28, 1999, pp. 964-974.

[7] P. Royston, W. Sauerbrei, Multivariate Model building, *Biometrical Journal*, Vol.51, No. 5, 2008, pp. 874-875.

[8] P. Royston, D. G. Altman, Regression using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling, *Applied Statistics*, Vol.43, 1994 pp. 429 – 467.

[9] J. P. Jansen, Network Meta-Analysis of Survival Data with Fractional Polynomials *BMC Medical Research Methodology*, Vol. 11, No. 61,2011, pp. 11 – 61.

[10] L. Chitty, D. Altman, Charts of Fetal size, *British Medical Ultrasound Society Bulletin*, Vol.2, No. 4, 1994, pp. 9 – 19.

[11] P. Royston, W. Sauerbrei, A new measure of prognostic separation in survival data, Statistics in Medicine, Vol.23, 2004, pp. 723–748.

[12] T. P. Morris, I. R. White, J. R. Carpenter, S. J. Stanworth, P. Royston, Combining fractional polynomial model building with multiple imputation, *Statistics in Medicine*, Vol.34, No. 25, 2015, pp. 3298 – 3317.

[13] J. H. Ryoo, T. R. Konold, J. D. Long, V. J. Molfese, X. Zhou, Nonlinear Growth models with Fractional polynomials: An illustration with Early children Mathematics Ability, *Structural Equation Modeling: A Multidisciplinary Journal*, Vol.24, No. 6, 2017, pp. 879-910.

[14] R. H. Keogh, T. P. Morris, Multiple imputation in Cox regression when there are time-varying effects of Covariates, *Statistics in Medicine*, Vol.37, No. 25, 2018, pp. 3661 – 3678.

[15] M. J. D. Powell, *Approximation Theory and Methods*, Cambridge University Press, 1981.

[16] J. O. Rawlings, S. G. Pantula, D. A. Dickey, *Applied Regression Analysis: A Research tool*, 2nd ed, Springer-Verlag, *1998.*

[17] R. L. Eubank, P. Speckman, Curve fitting by polynomial-trigonometric regression, *Biometrika*, Vol.77, 1990, pp. 1 – 9.

[18] E. Ullah and T. Shah, “Trigonometric Polynomial Rings and their Factorization Properties, *MATEMATIQKI VESNIK*, Vol. 66, No. 33, 2014, pp. 301–314.

[19] W. S. Wei, *Time Series Analysis Univariate and Multivariate Methods*, 2nd ed, Pearson Addison, 2006.

[20] K. K Adesanya, Parameter Estimation of Fractional and Trigonometry Polynomial Regression of Periodic Time Series Data, PhD Thesis, Olabisi Onabanjo University, Ago -Iwoye, Nigeria, 2018.

[21] Information and Medical record office, Ijebu-Ode General Hospital, Ijebu-Ode Ogun state, Nigeria, 2018.