Tracing Mathematical Function of Age Specific Fertility Rate in Peninsular Malaysia

Norsyela Muhammad Noor Mathivanan, Puzziawati Ab Ghani, Nor Azura Md.Ghani
Center for Statistical and Decision Sciences Studies, Faculty of Computer & Mathematical Sciences
Universiti Teknologi MARA, Selangor, Malaysia

Article Info	ABSTRACT
	The size, structure, and composition of a population are affected by the fertility rates at any point of time. Many researchers took the opportunity to exploit the fertility rates in obtaining better fertility patterns for their country. The curve for the age specific fertility rate is consistent, and this feature allows the curve to be matched with a mathematical model. This paper aimed to identify the best mathematical model that fits the recent age specific fertility rate in Peninsular Malaysia. This study fitted the fertility data of Peninsular Malaysia from 1996 to 2014 to the four mathematical models, which were Hadwiger, Gamma, Beta, and Gompertz models. From the comparisons of the four models, it was found that the best fitted mathematical model is Hadwiger model. In relation to the data of early 21st century, there was an inclination for the best fitted mathematical model from Hadwiger model to Beta model. Hence, the best mathematical model for each year can be used to convert a fertility schedule classified in a five-year age group into a fertility schedule for a single-year of age in Peninsular Malaysia. This model also can be helpful for population projections by using limited and defective data.

Corresponding Author:

Nor Azura Md.Ghani,
Center for Statistical and Decision Sciences Studies
Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA
40450 Shah Alam, Selangor Malaysia
Email: azura@tmsk.uitm.edu.my

1.   INTRODUCTION

Fertility plays an important role in the growth of a population. The size, structure, and composition of a population are affected by the birth rate at any point of time. Therefore, researchers took the opportunity to exploit the fertility rates in the effort to obtain the better fertility patterns for their country. Fertility can be defined as the product or output of reproduction instead of the ability to have children [1]. The common measure of fertility rates includes the crude birth rate (CBR), general fertility rate (GFR), total fertility rate (TFR), gross reproduction rate (GRR), net reproduction rate (NRR), and age-specific fertility rate (ASFR). Fertility rates can either be period based or cohort based and are referred as a benchmark or indicator to describe the behaviour of fertility in a population.

Most of the time, CBR is useful in determining the growth rate for the population of a country. However, TFR is commonly used as a national indicator for the family size in a country, and in many cases, age-specific fertility rates (ASFRs) are the most favoured measurement due to the impact of women's

reproductive age towards fertility behaviours. The baseline for the reproductive age of women is in the range of 15 to 49 years old. This measure is more suitable to describe fertility because it is more geared to those who were directly involved in the birth process [2]. In Malaysia, the TFR has declined quite rapidly from 4.0 children per woman in 1980 to 3.0 children per woman in 2000, and the latest statistics showed that in 2014, the rate reached the population replacement level of 2.1 children [3]. These declining fertility rates have raised concerns regarding the potential implications on the population structure. Besides that, Malaysia is also facing population ageing as there are more people within the age range of 60 years old and above, rather than children who are younger than five years old. This situation could lead to the phenomenon of ageing population. Hence, the TFR is an important fertility measurement in monitoring the fertility behaviors around the globe.

Although most studies have provided analysis on the TFR in understanding fertility transition, this measurement is actually built from the ASFRs. Moreover, the distribution of the TFR is sensitive to the timing effects. A closer look at the timing of births, e.g., mother's age is needed to obtain a more comprehensive view of the fertility behavior and changes over time. The patterns of fertility across different age groups may vary although the same TFRs are produced. There is a possibility that different countries have similar TFRs but with different compositions of ASFRs such as lower fertility for younger age range rather than older age range and vice versa. Normally, the pattern of fertility is hypothesized to decline in the older age range first, followed by the decline in the youngest age range [4].

Most of the countries have recorded the ASFRs for seven groups of five-year interval age of ASFRs, i.e., 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, and 45-49. In visualizing these ASFRs, the value point for each group of age interval is combined with one another in a crude way. One of the alternatives to smoothen this curve is by fitting a specific mathematical function that suits the ASFRs. This idea is possible and relevant to be executed as the fertility pattern for any countries is a typical bell shaped with the right skewed deviation over the time. This fact has made it possible to convert a fertility schedule classified in a five-year age group into a fertility schedule for a single-year of age. Besides that, a suitable mathematical function can be helpful for population projections and various approaches in demographic estimation by using limited and defective data. The common fertility shape implies that fertility has a zero value before entering the reproductive age, i.e., 14 years old. Then, the positive values increase in the range of 15-19 years old, and peak at the age group of 20-29 years old. The fertility declines slowly until the value reaches almost zero by the age of 50 and above.

A variety of mathematical functions had been proposed by previous researchers in fitting the ASFRs curve throughout the world. The mathematical functions include the Hadwiger function [5, 6], the Beta and Gamma functions which are similar to the Pearson type I and III curves [7], Brass procedure [8], Gompertz curve [9], Polynomial models [10], Coale-Trussell model [11], and cubic spline [12]. The most frequently used mathematical functions are the Hadwiger, Gamma, Beta, Gompertz and Coale-Trussell functions [2]. However, the Coale-Trussel function is widely used to fit the age-specific to the marital fertility rate (ASMFR). This function describes the combination of three aspects, i.e., fertility, contraception, and age at marriage [15]. These various mathematical functions were applied to different ASFRs all around the world, where the best mathematical function that suited the ASFRs for the particular population was dependent on the ASFRs distribution. In short, different populations may have different mathematical functions that fit well with the ASFRs data.

Previous study took the initiative to study the mathematical functions that best suit ASFRs in Peninsular Malaysia from 1960 until 1995 in providing a clearer view of the fertility pattern in Peninsular Malaysia [2]. Several mathematical functions with parameters, i.e., the Hadwiger, Gamma, Beta, Gompertz and Coale-Trussell functions were compared in his study. Even though the study found that the Hadwiger function was the best mathematical function that fit ASFRs for the last fifteen years, the rationality of the best function over the time has often been disputed by demographers [13, 14]. Therefore, the aim of this study is to identify the mathematical function that best fits the recent fertility data in Peninsular Malaysia.

2. RESEARCH METHOD

The data used in the study included the TFRs and ASFRs in Peninsular Malaysia from 1996 to 2014. The ASFRs data for aged 15 years to 49 years are recorded by a single age interval i.e. 15, 16, ... , 49 or five-year age interval i.e. 15-19, 20-24, … , 45-49. This study used the mid-point of the age interval i.e. 17.5, 22.5, 27.5, 32.5, 37.5, 42.5 and 47.5 for fitting the data with the mathematical functions. Let assume that represents the ASFR retrieved from vital statistics. Meanwhile, represents the mathematical function that will be fitted with the ASFRs data. Then, the estimation of the can be obtained with the that minimizes the function of,

$$v = \sum_{u=17.5} \left(f(u;\theta) - f_u \right)^2 \qquad (1)$$

This function solution can be obtained with the help of statistical packages such as SPSS, S-Plus, STATISTICA, SAS and R-programming through the nonlinear estimation methods. The non-linear estimation model has been widely used in various field of studies include in study of population growth [1], predictive control [15, 16], pattern recognition [17] etc. In this study, the parameter was estimated by the R-programming software. This study used the coefficient of determination, R^2 to identify the best mathematical functions that fit well with the observed ASFR data. The mathematical function for the ASFRs consist of two important parts such as,

$$f(u:\theta) = \theta \, h(u:\theta_2,...,\theta_r) \qquad (2)$$

where u represents the age of the mother in all models that are used in the study. Then, $h(u:\theta_2,....\theta_r)$ is the probability density function with $(r-1)$ as the parameter of $\theta_2,.....\theta_r$ and this function are represented differently according to the four mathematical functions i.e. Gompertz, Hadwiger, Beta and Gamma models. Meanwhile, the θ_1 is the parameter that represents the TFR.

2.1. Gompertz Model

The Gompertz function [9] is a sigmoid function. It is a special case of the generalised logistic function. The double exponential function will become the Gompertz function when the parameter of $\alpha=\lambda$. Meanwhile, the Gompertz function will become the Gumbel function if the $\beta=1$. The Gompertz function can be expressed by,

$$h(u) = \frac{\beta}{\alpha} \exp\left(-\frac{u-m}{\alpha} - \beta \exp\left(-\frac{u-m}{\alpha} \right) \right) \text{ where } m = 14 \qquad (3)$$

The m value can be defined as the lowest marriage age of the population where for this function, the m was 14.

2.2. Hadwiger Model

The Hadwiger function [5] is expressed by,

$$h(u) = \frac{\alpha}{\beta\sqrt{\pi}} \left(\frac{\beta}{u-m} \right)^{3/2} \exp\left[-\alpha^2 \left(\frac{\beta}{u-m} \right) + \frac{u-m}{\beta} - 2 \right] \text{ where } u > m \qquad (4)$$

u - m is the age of the mother at birth and m is a constant and normally equals to zero. The parameters of the model may have a demographic interpretation as the parameter is associated with total fertility, α determines the height of the curve and β is related to the mean age of motherhood, while the term is associated with the maximum ASFR [17].

2.3. Gamma Model

The Gamma function [7] is given by,

$$h(u) = \frac{1}{r(\alpha)\beta^\alpha} (u-m)^{\alpha-1} \exp\left(-\frac{u-m}{\beta} \right) \text{ where } u > m \qquad (5)$$

The m value is the youngest age of the mother to give birth, which was previously 14. The parameters α and β are related to the mode, mean, and variance of the function but not in a simple linear way and thus, they do not have direct demographic interpretations [7].

2.4. Beta Model

The Beta function [7] is given by,

$$h(u) = B(\alpha, \beta)\left(\frac{u-a}{b-a}\right)^{\alpha-1}\left(1 - \frac{u-a}{b-a}\right)^{\beta-1}(b-a)^{-1} \text{ where } a < u < b \qquad (6)$$

The value for *a* represents the youngest age of the mother to give birth which is 15, meanwhile, *b* represents the oldest age of the mother to give birth which is 50. The parameters, *α* and *β* represent the lower and upper age limits of fertility but showed that in several cases the value of *β* far exceeds the maximum age.

3.   RESULTS AND ANALYSIS

Table 1 shows parameter *θ* where represents the estimated TFR based on the four mathematical models with the observed TFR. It is found that all mathematical models able to estimate closely the value of TFR through the years. Whereas parameters *α* and *β* do not have any demographic interpretations except for parameter α of the Hadwiger model which is related to the mean age of motherhood.

Table 1. Parameter Estimation for Each of the Mathematical Functions from 1996-2014

Year	Observed TFR	Mathematical function of ASFRs			
		Gompertz	Hadwiger	Gamma	Beta
1996	3.22	3.349	3.285	3.270	3.197
1997	3.14	3.264	3.202	3.186	3.117
1998	3.02	3.151	3.089	3.076	3.005
1999	2.94	3.057	2.996	2.981	2.915
2000	3.01	3.118	3.056	3.042	2.977
2001	2.76	2.862	2.805	2.791	2.731
2002	2.67	2.762	2.706	2.694	2.636
2003	2.56	2.640	2.586	2.574	2.520
2004	2.51	2.585	2.531	2.520	2.467
2005	2.41	2.483	2.428	2.419	2.367
2006	2.35	2.425	2.369	2.360	2.309
2007	2.33	2.408	2.350	2.342	2.292
2008	2.33	2.404	2.342	2.334	2.284
2009	2.30	2.370	2.308	2.299	2.252
2010	2.18	2.248	2.189	2.181	2.136
2011	2.23	2.298	2.234	2.228	2.183
2012	2.25	2.317	2.253	2.246	2.202
2013	2.09	2.149	2.094	2.087	2.046
2014	2.10	2.17	2.113	2.107	2.066

Based on Table 2, Hadwiger model was the best mathematical model that fits ASFRs with 99.46% matched the observed data. Unlike in 1996 to 2007 where the model had shown the best performance among the four mathematical models, Hadwiger and Beta models had a similar R^2 value in 2008. Then, the Beta model seemed to fit the data from 2009 to 2014 better than the other models. The performance of the Hadwiger model slightly decreased over the time. It seemed that there was an inclination of a shift from the Hadwiger to the Beta model. The third mathematical model that fits well with the data was the Gamma model. The Gompertz model had the lowest R^2 values which indicate it was the worst among the mathematical models used in the study.

The inclination of a shift observed from Hadwiger model to Beta model is due to the change in the structure of fertility based on the women's age in the 21st century. This interesting finding matched those observed in the earlier study where the Beta model provided the best fit for most of the ASFRs in countries with the non-enhanced early-age fertility [17]. The population in Malaysia also experiences the same situation. The country experiences low fertility rates for women in their early reproductive ages. This situation indicates that women either delayed their marriage or practice family planning. The formulation of the Beta model with inclusion of age limitation at marriage suited with the changes in the current fertility behaviours highly related to the postponement of marriage among women in Peninsular Malaysia. Hence, it was proven that there is a high possibility that the Beta model will be the best mathematical model that fits the ASFR in Peninsular Malaysia compared to other three mathematical models in the future.

Table 2. Coefficient of Determination, R^2 for Each of the Mathematical Functions from 1996-2014

Year	Mathematical function of ASFRs			
	Gompertz	Hadwiger	Gamma	Beta
1996	0.9861	0.9960	0.9920	0.9953
1997	0.9860	0.9960	0.9922	0.9953
1998	0.9846	0.9974	0.9911	0.9954
1999	0.9848	0.9955	0.9919	0.9933
2000	0.9851	0.9954	0.9926	0.9915
2001	0.9840	0.9945	0.9913	0.9910
2002	0.9853	0.9959	0.9930	0.9907
2003	0.9844	0.9954	0.9925	0.9905
2004	0.9849	0.9958	0.9931	0.9897
2005	0.9837	0.9962	0.9931	0.9915
2006	0.9834	0.9966	0.9934	0.9921
2007	0.9809	0.9961	0.9924	0.9932
2008	0.9765	0.9949	0.9901	0.9949
2009	0.9739	0.9936	0.9884	0.9950
2010	0.9732	0.9933	0.9881	0.9944
2011	0.9713	0.9925	0.9874	0.9945
2012	0.9705	0.9919	0.9870	0.9942
2013	0.9742	0.9921	0.9884	0.9981
2014	0.9715	0.9910	0.9870	0.9929
Average of R^2	0.9802	0.9946	0.9908	0.993

4.   CONCLUSION

The ASFR in Peninsular Malaysia has a typical bell shape over time and it is possible to match the ASFR curves with specific mathematical functions. Four common mathematical functions i.e. the Hadwiger, Gamma, Beta and Gompertz models are used to trace the best mathematical function that fits the ASFRs of overall Peninsular Malaysia population. The Hadwiger model fits the best from 1996 to 2007. However, there was an inclination of a shift between the Hadwiger and Beta models where the latter model fits the best from 2009 to 2014. Hence, there was a high possibility that the Beta model will be the best mathematical model that fits the ASFRs in Peninsular Malaysia in the 21st century. Even though this study proved that the four common mathematical functions fit well with the data, other researchers can take opportunities to use other distributions e.g. cubic spline, polynomial models etc. that may also fit the data.

REFERENCES

[1]   Zakria M, Yaseen M. Modelling Fertility Patterns of Pakistan during 1984 to 2007. *Pakistan Journal of Statistics*, 2013; 29(4), 479–486.
[2]   Jemain A. A. Menjejak Fungsi Matematik Kadar Fertiliti Umur Tertentu untuk Kelahiran di Semenanjung Malaysia. *Bulletin of the Malaysian Mathematical Sciences Society*; 2001; 24(2), 137–148.
[3]   Subramaniam G, Saleh N M. Does Work Environment Impact Fertility Rate? A Comparison between Formal and Informal Sectors in Malaysia. *Journal of Emerging Economies and Islamic Research*, 2016; 4(1),2-16.
[4]   Pantazis A. Age-Specific Fertility Dynamics: Sub-Saharan African Fertility in a Global Cont ext. PhD Thesis. Seattle: University of Washington; 2016.
[5]   Hadwiger H. (1940). Eine analytische Reprodutionsfunktion für biologische Gesamtheiten. *Skandinavisk Aktuarietidskrift*, 1940; 23, 101-113.
[6]   Gilje E, Yntema L. *The shifted Hadwiger fertility function*. Central Bureau of Statistics of Norway. Report number: IO70/14. 1970.
[7]   Hoem J M, Madsen D, Nielsen J L, Ohlsen E M, Hansen H O, Rennermalm B. Experiments in modelling recent Danish fertility curves. *Demography*, 1981; 18(2).231–244.
[8]   Brass W. Perspectives in population prediction: Illustrated by the statistics of England and Wales. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 1974; 137(4), 532–583.
[9]   Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 1825; 182:513–85.
[10] Brass, W. The graduation of fertility distributions by polynomial functions. *Population Studies*, 1960; 14: 148-162.
[11] Coale A J, Trussell T J. Model Fertility Schedules: Variations in the Age Structure of Childbearing in Human Populations. *Population Index*, 1974; 40(2), 185-258.
[12] Hoem J M, Rennermalm B. *On the statistical theory of graduation by splines*. University of Copenhagen, Laboratory of Actuarial Mathematics. Working paper: 14. 1978.
[13] Jemain A A, Puzziawati A G. Menjejak Perubahan Fertiliti dan Umur Perkahwinan di Semenanjung Malaysia. *Jurnal Teknologi*, 2003; 39(C), 75–89.

[14] Puzziawati A G. (2006). *Modelling of cohort fertility changes among major ethnics in Peninsular Malaysia.* Proceedings of the National Statistics Conference. Kuala Lumpur. 2006: 1-10.

[15] Lu J, You J, Yang Q. Nonlinear Model Predictive Controller Design for Identified Nonlinear Parameter Varying Model. *TELKOMNIKA*, 2012; 10(3), 514-523.

[16] Al-Omari Z, Hamzeh A, Hamed S A, Sandouk A, Aldahim G. A Mathematical Model for Minimizing Add-On Operational Cost in Electrical Power Systems Using Design of Experiments Approach. *International Journal of Electrical and Computer Engineering (IJECE)*, 2015; 5(5), 948-956.

[17] Zhixia J, Jianjun Y, Haifeng L, Zhengqiang W. Mathematical Programming Model on Joint and Recovery of Paper Scrap. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 2014; 12(7), 5522-5528.

[18] Chandola T, Coleman D A, Hiorns R W. Distinctive Features of Age-Specific Fertility Profiles in the English-Speaking World: Common Patterns in Australia, Canada, New Zealand and the United States 1970-1998. *Population Studies*, 2002; 53, 317-329.

[19] Peristera P, Kostaki A. Modeling fertility in modern populations. *Demographic Research*, 2007; 16, 141–194.