

Research Article

Using Skew-Logistic Probability Density Function as a Model for Age-Specific Fertility Rate Pattern

Sahar Asili,¹ Sadegh Rezaei,¹ and Lotfollah Najjar²

¹ Department of Statistics, Amirkabir University of Technology, Tehran 15875-4413, Iran

² College of Information Science Technology, University of Nebraska-Omaha, Omaha, NE 68182, USA

Correspondence should be addressed to Sadegh Rezaei; srezaei@aut.ac.ir

Received 4 February 2014; Revised 13 April 2014; Accepted 27 April 2014; Published 21 May 2014

Academic Editor: Stefano Curcio

Copyright © 2014 Sahar Asili et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fertility rate is one of the most important global indexes. Past researchers found models which fit to age-specific fertility rates. For example, mixture probability density functions have been proposed for situations with bi-modal fertility patterns. This model is less useful for unimodal age-specific fertility rate patterns, so a model based on skew-symmetric (skew-normal) pdf was proposed by Mazzuco and Scarpa (2011) which was flexible for unimodal and bimodal fertility patterns. In this paper, we introduce skew-logistic probability density function as a better model: its residuals are less than those of the skew-normal model and it can more precisely estimate the parameters of the model.

1. Introduction

Some countries have a bimodal shape of age-specific fertility rates that classical bell shaped models [1] cannot fit suitably. It is possible to fit these patterns by means of a mixture model. Past researchers have proposed some models for these patterns, which we briefly mention as follows.

Chandola et al. [2] have proposed a Hadwiger mixture model which has seven parameters. Peristera and Kostaki in their paper [3] have developed a model, based on normal distribution which has six parameters. Schmertmann [4] has proposed a piecewise quadratic spline function which has 13 parameters. Mazzuco and Scarpa [5] have introduced a different model based on skew-normal density function which has 4 parameters. Because of the skewness parameter, that model has been suitable for most types of fertility patterns, including bimodal fertility patterns.

In this paper, as a developed by Mazzuco and Scarpa (2011) [5], a skew-logistic model is proposed. This model is fitted to the age-specific fertility rate data, and it has been shown that the residuals of this model are less than the residuals of the skew-normal model. The researchers have determined that this model estimates the parameters more precisely.

Section 2 contains a brief review of existing models of age-specific fertility rates. Section 3 is an introduction to our proposed model for age-specific fertility rates, based upon a skew-logistic probability density function. Fitting fertility models to real data is done in Section 4. Conclusions and future research are explained in Section 5.

2. Summary of Some Age-Specific Fertility Models

The general form of a fertility curve is as follows:

$$g(x; R, \theta_2, \dots, \theta_r) = R \cdot h(x; \theta_2, \dots, \theta_r), \quad (1)$$

where $h(x; \theta_2, \dots, \theta_r)$ is a probability density function (pdf) on the real line with $r-1$ parameters and R is the r th parameter representing the total fertility rate. In different models of fertility rate, function $h(x; \theta_2, \dots, \theta_r)$ is also different; for example, it may be inverse Gaussian, the Gamma, the Beta, the Coale-Trussell, the Brass, or the Gompertz pdfs. Using the inverse Gaussian instead of $h(x; \theta_2, \dots, \theta_r)$ in the Hadwiger

model was useful for bimodal models. The Hadwiger function is as following:

$$g(x; a, b, c) = \frac{ab}{c} \left(\frac{c}{x}\right)^{3/2} \exp\left\{-b^2\left(\frac{c}{x} + \frac{x}{c} - 2\right)\right\}. \quad (2)$$

In 1999, Chandola et al. introduced the ‘‘Hadwiger mixture model’’ [2] which has the following formula:

$$\begin{aligned} g(x; a, m, b_1, b_2, c_1, c_2) \\ = m \frac{ab_1}{c_1} \left(\frac{c_1}{x}\right)^{3/2} \exp\left\{-b_1^2\left(\frac{c_1}{x} + \frac{x}{c_1} - 2\right)\right\} \\ + (1 - m) \frac{ab_2}{c_2} \left(\frac{c_2}{x}\right)^{3/2} \exp\left\{-b_2^2\left(\frac{c_2}{x} + \frac{x}{c_2} - 2\right)\right\}. \end{aligned} \quad (3)$$

Here, $0 \leq m \leq 1$ is the mixture parameter. Another model was proposed in [4] with the following formula:

$$\begin{aligned} g(x; R, \alpha, \beta, \theta_0, \theta_4, t_0, t_4) \\ = R \cdot I(\alpha \leq x \leq \beta) \cdot \sum_{k=0}^4 \theta_k (x - t_k)^2. \end{aligned} \quad (4)$$

Here, $I(\cdot)$ is the indicator function, α and β are the age limits, t_k are the spline knots, and also θ_k are the parameters. Then Peristera and Kostaki [3] proposed another model for fertility patterns based on normal mixture model which has the following formula:

$$\begin{aligned} g(x; c_1, c_2, \mu_1, \mu_2, \sigma_1, \sigma_2) \\ = c_1 \exp\left\{-\left(\frac{x - \mu_1}{\sigma_1}\right)^2\right\} + c_2 \exp\left\{-\left(\frac{x - \mu_2}{\sigma_2}\right)^2\right\}. \end{aligned} \quad (5)$$

In [5], a new model was proposed to fit age-specific data based on skew-symmetric (skew-normal) density function. This model is flexible for almost all types of fertility patterns. We briefly explain their proposed model as follows: at the beginning, we review the skew-normal pdf, which was studied in

$$f(x; \xi, \omega^2, \alpha) = 2\omega^{-1} \phi\left(\frac{x - \xi}{\omega}\right) \Phi\left\{\alpha\left(\frac{x - \xi}{\omega}\right)\right\}. \quad (6)$$

Ma and Genton [6] showed that the skew-normal pdf is unimodal, so this model is useful for fertility patterns such as USA age-specific fertility rates (1963); compare [5]. To improve this model, Mazzuco and Scarpa proved that model (1) may be generalized using the results shown in [7–10]. There is more information about this as follows.

For any symmetric pdf f_0 and distribution function G with a symmetric density, function (2) is a density function for any odd function $\omega(\cdot)$. Consider

$$f(x) = 2f_0(x)G\{\omega(x)\}. \quad (7)$$

Now if pdf and cdf of standard normal distribution replace f_0 and G , respectively, and insert $\omega(x) = \alpha x$,

the Flexible Generalized Skew-Normal (FGSN) distributions formed as follows:

$$\begin{aligned} f(x; \xi, \omega^2, \alpha, \beta) \\ = 2\omega^{-1} \phi\left(\frac{x - \xi}{\omega}\right) \Phi\left\{\alpha\left(\frac{x - \xi}{\omega}\right) + \beta\left(\frac{x - \xi}{\omega}\right)^3\right\}. \end{aligned} \quad (8)$$

According to [6], the pdf in (8) has at almost two modes, so this model is adequate for bimodal fertility patterns. As you can see, there are now 4 parameters α, β, ξ , and ω in this model which will be interpreted as the following.

Parameter ξ is location parameter, and ω is scale parameter. Parameter α is the skewness parameter in the skew-normal distribution when $\beta = 0$, and, in the FGSN distribution, β is the skewness parameter. The two parameters α and β are related to the location of two modes fertility patterns. We do not have the exact value of them, so we assumed that they vary between -5 and $+5$. According to Mazzuco and Scarpa, there is different plots for the situations in which ξ, ω , and α are fixed, and the parameter β has various values. After interpreting the parameters of Mazzuco and Scarpa [5], their model was fitted to real data. Notice that parameters of fertility patterns are estimated through nonlinear least squares, by minimizing the following term:

$$S(R, \theta_2, \dots, \theta_r) = \sum_{x=b}^e \{g(x; R, \theta_2, \dots, \theta_r) - f_x\}^2, \quad (9)$$

where f_x is the real age-specific fertility rate, $g(x; R, \theta_2, \dots, \theta_r)$ is the fertility rate at age x given by the fertility model used, and b and e are the ages at the beginning and at the end of the fertile period, respectively; compare [5]. In that paper, you can find two figures, the Italian fertility model and the USA fertility model. In these figures, the residuals of the skew-normal model are somewhat less than those of the other models. Although there are some models in these figures which do not have smaller residuals, because these models have more parameters, again we can claim that skew-normal is preferred.

In the next section, we show that the skew-logistic model as a new fertility model has 4 parameters, just as the skew-normal model. However, its residuals are less than skew-normal, so it is the preferred model so far.

3. Skew-Logistic Distribution as a Fertility Model

In this paper, our aim is to introduce and to fit a skew-logistic model for age-specific fertility patterns.

Equations (10) and (11) show cdf and pdf of the logistic respectively:

$$G(x) = \frac{1}{1 + e^{-\lambda x}}, \quad (10)$$

$$g(x) = \frac{\lambda e^{-\lambda x}}{(1 + e^{-\lambda x})^2}. \quad (11)$$

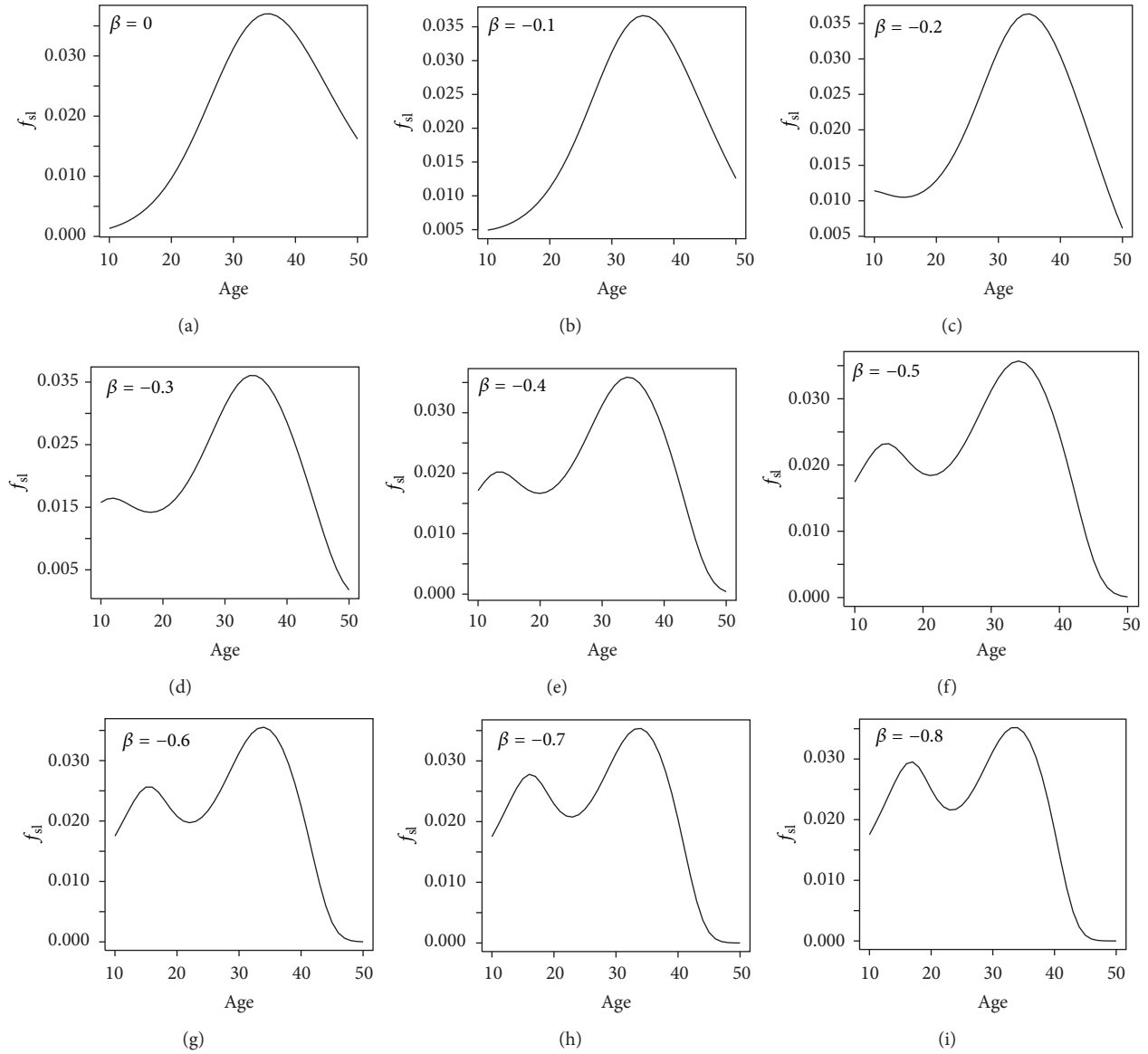


FIGURE 1: Skew-logistic probability function when $\xi = 30, \sigma = 8, \alpha = 1$, and β has different value.

The following formula was introduced by Azzalini for skewing a symmetric distribution:

$$f(x) = 2f_0(x)G\{\omega(x)\}. \tag{12}$$

If we replace (10) and (11) instead of G and f_0 at (12) and put $\omega = \alpha x$, we will have skew-logistic distribution function which is as follows:

$$f_{sl}(x; \alpha) = \frac{2e^{-x}}{(1 + e^{-x})^2(1 + e^{-\alpha x})}. \tag{13}$$

Here, α is the skewness parameter and $x \in \mathbb{R}$. Now we transform x to $(y - \mu)/\sigma$, in which μ is the location parameter and σ is the scale parameter:

$$f_{sl}(y; \alpha, \mu, \sigma) = \frac{2e^{-(y-\mu)/\sigma}}{\sigma(1 + e^{-(y-\mu)/\sigma})^2(1 + e^{-\alpha((y-\mu)/\sigma)})}; \tag{14}$$

$y \in \mathbb{R}.$

If we rewrite (14) based on the odd power, we will have

$$f_{sl}(y; \alpha, \beta, \mu, \sigma) = 2\sigma^{-1} \cdot \frac{e^{-(y-\mu)/\sigma}}{(1 + e^{-(y-\mu)/\sigma})^2(1 + e^{-\alpha((y-\mu)/\sigma) - \beta((y-\mu)/\sigma)^3})}. \tag{15}$$

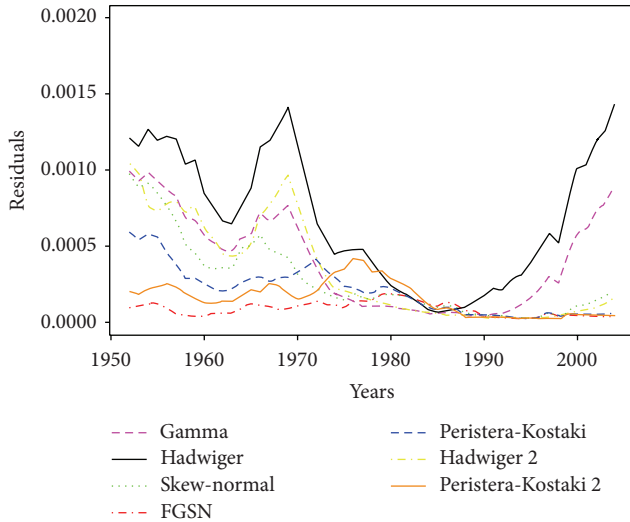


FIGURE 2: Residuals of fertility models fitted to Italy data.

Equation (15) is the model which we propose for age-specific fertility patterns. Examples which show that this pdf can have two modes are in Figure 1.

4. Fitting Fertility Models to Real Data

As it is shown in [5] and Figure 2, which is for an Italian data set, the skew-normal model is generally a better fit, with lower residuals.

In this section, the skew-normal and skew-logistic fertility models will be fitted to real data and their quality will be assessed. The results show that the skew-logistic's fit is better than that of skew-normal, so, with respect to Mazzuco and Scarpa's findings, we conclude that skew-logistic has the best fit of the models mentioned above.

To compare, we will use the fertility data of Ireland in the same years. The data of Greece will be surveyed in Figure 3. Data are taken from Human Fertility Database (HFD) [11] and Eurostat (<http://epp.eurostat.ec.europa.eu/>). Notice that the parameters of the model will be estimated through nonlinear least squares, just as skew-normal; it means by minimizing (9) that here $g(x; R, \theta_2, \dots, \theta_r)$ is the fertility rate at age x given by the skew-logistic fertility model. This model is fitted to the age-specific fertility rate data of Ireland in Figure 3. Looking at this figure, it appears that skew-normal and skew-logistic have a similar pattern but it is clear that skew-logistic is preferred to skew-normal because its sum of square of residuals is much lower than that of skew-normal. It is also superior to other models introduced later (there is more information in Section 5). Also, fitting the model to age-specific fertility rate data of Greece in Figure 4, it is again shown that the skew-logistic has a lower sum of square of residuals.

5. Conclusions

A new fertility model has been proposed which is based on skew-logistic probability density function. Previously, other

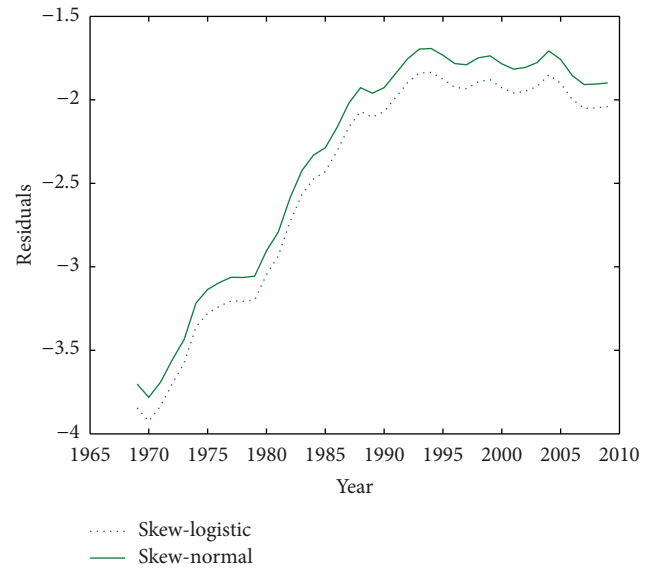


FIGURE 3: Residuals of fertility models fitted to Ireland data.

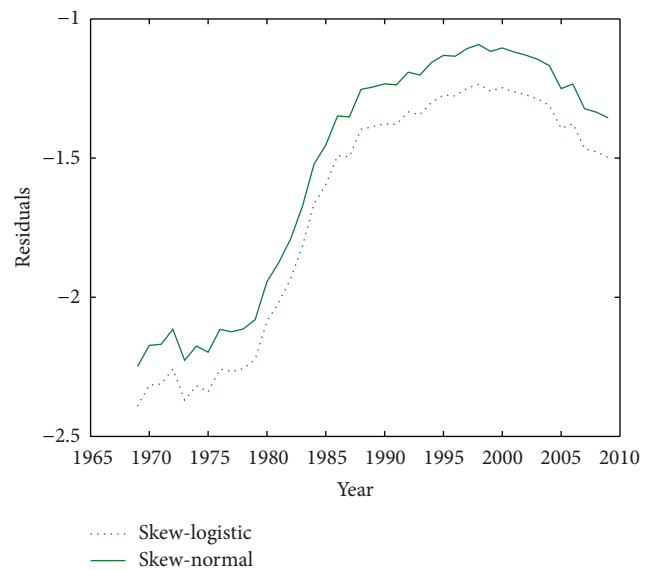


FIGURE 4: Residuals of fertility models fitted to Greece data.

models were used which had multiple parameters and so were less ideal for our aim. Recently, Mazzuco and Scarpa (2011) [5] proposed a model based on generalization of the skew-normal distribution (FGSN). The advantage of the FGSN model is its flexibility for complex fertility patterns and also the number of its parameters, which is lower than the later models. We showed that the model based on the skew-logistic distribution function is better than skew-normal not because of having lower number of parameters but because of having a better fit to data. For future research, it is proposed that student's t -distribution is used instead of logistic.

Because it may be possible that a model based on student's t -distribution will have fewer parameters. It is also possible that, skewing by Fernandez and Steel [12] method instead of Azzalini [13], better fit to real data may be produced.

Finally it is necessary to discover whether the MATLAB program could be used for calculating parameters and drawing graphs. Outputs of this program show that, with the same data, the norm of the residual (resnorm) of the skew-normal model is 148.7917 and the resnorm of the skew-logistic model is 143.5909. This shows that the accuracy of the latter is greater than the former.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] J. M. Hoem, D. Madsen, J. L. Nielsen, E.-M. Ohlsen, H. O. Hansen, and B. Ren-nermalm, "Experiments in modelling recent Danish fertility curves," *Demography*, vol. 18, no. 2, pp. 231–244, 1981.
- [2] T. Chandola, D. Coleman, and R. W. Hiorns, "Recent European fertility patterns: fitting curves to "distorted" distributions," *Population Studies*, vol. 53, no. 3, pp. 317–329, 1999.
- [3] P. Peristera and A. Kostaki, "Modeling fertility in modern populations," *Demographic Research*, vol. 16, pp. 141–194, 2007.
- [4] C. P. Schmertmann, "A system of model fertility schedules with graphically intuitive parameters," *Demographic Research*, vol. 9, pp. 82–110, 2003.
- [5] S. Mazzuco and B. Scarpa, "Fitting age-specific fertility rates by a skew-symmetric probability density function," 2011, <http://www2.lse.ac.uk/socialPolicy/BSPS/pdfs/2011.Mazzuco.pdf>.
- [6] Y. Ma and M. G. Genton, "Flexible class of skew-symmetric distributions," *Scandinavian Journal of Statistics*, vol. 31, pp. 459–468, 2004.
- [7] A. Azzalini and A. Capitanio, "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution," *Journal of the Royal Statistical Society B*, vol. 65, no. 2, pp. 367–389, 2003.
- [8] A. Azzalini, "The skew-normal distribution and related multivariate families (with distribution)," *Scandinavian Journal of Statistics*, vol. 32, no. 2, pp. 159–188, 2005.
- [9] M. G. Genton, "The Skew-normal," *Scandinavian Journal of Statistics*, vol. 32, no. 2, pp. 189–198, 2005.
- [10] A. Azzalini, "Rejoinder," *Scandinavian Journal of Statistics*, vol. 32, no. 2, pp. 199–200, 2005.
- [11] Human Fertility Database, "Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria)," <http://www.humanfertility.org/>.
- [12] C. Fernandez and M. F. J. Steel, "On Bayesian modeling of fat tails and skewness," *Journal of the American Statistical Association*, vol. 93, no. 441, pp. 359–371, 1998.
- [13] A. Azzalini, "A class of distributions which includes the normal ones," *Scandinavian Journal of Statistics*, vol. 12, pp. 171–178, 1985.