Istituto di Matematica Guido Castelnuovo and Istituto di Calcolo delle Probabilità,
University of Rome, Italy

FERTILITY AS A FUNCTION OF THE WOMAN'S AGE AND YEAR OF BIRTH IN ITALY *

FLAVIA M. MASCIOLI, CARLA ROSSI

SUMMARY

A mathematical model is suggested to represent the trend of the fertility coefficient in the Italian woman as a function of her age and year of birth. The study has been based on data from the fertility tables by Livi Bacci-Santini interpreted by one single formula.

INTRODUCTION

The object of the present work is to suggest a stochastic model of a growing population, alternative to the deterministic model by Lotka (1939).

Under the assumption of a growing population with constant rate of increase, r, and constant birth rate, b, Lotka analyzes the fraction $c(x)dx$ of individuals of the present population now alive and aged x to $x + dx$.

$$c(x)dx = be^{-rx}l(x)dx, \qquad [1]$$

where $l(x)$ indicates the chance for a child born now, to live up to age x. Making use of [1], Keyfitz (1968) mathematically derives some interesting results.

It should be noted that $c(x)dx$ is only the expected value of the number $\bar{c}(x)dx$ of living subjects between age x and $x + dx$. This quantity is a random number, subject to fluctuations of several types. The number, $\bar{c}(x)dx$, as obtained from the observations, may be written as a sum of several terms: the first one is $c(x)dx$, while the others denote the fluctuations that may be identified with the errors ε_i. They have a distribution with mean 0 and variance σ_i^2. Therefore,

$$\bar{c}(x)dx = c(x)dx + \Sigma_i \varepsilon_i dx.$$

Since one makes use of quantities affected by errors as data, we propose to study the way in which these errors spread over the results. Because of the serious difficulties encountered in the collection of some statistical data, it has not yet been pos-

* The present work has been carried out within the program of the G.N.A.F.A. of the Italian National Research Council (CNR).

Paragraphs 1 and 2 are due to C. Rossi (Istituto di Calcolo delle Probabilità); paragraphs 3, 4, and 5, to F.M. Mascioli (Istituto di Matematica Guido Castelnuovo); while introduction and conclusions are due to both. Both authors thank Professor B. de Finetti for suggesting and guiding the present work.

Rec. 16. IX. 1972

Acta Genet. Med. Gemellol. (1972), 21: 257-264

sible to complete the study of the stochastic model. To develop this model, an experimental analysis of the woman's fertility coefficient was to be performed. We confined ourselves to Italian data, and now report the results obtained in this preliminary work, as they are necessary to continue the study of this model.

The fertility coefficient is an index with a variability determined by biological and social factors; it represents the probability for a woman of a given age to bear a child within a year. This coefficient is computed with rates which are sensitive to short and long factors. These factors generate and perturb the above-mentioned rates.

In order to eliminate the perturbations, we have studied in detail the fertility tables, as calculated by Livi Bacci-Santini (1969). This allows to identify the general trend that underlies the apparently irregular data given in these tables. In this analysis, the serious alteration of the rates caused by World War II have not been taken into account. In fact, the fertility trend is only determined by biological, social, and economical factors, and not by random external ones, such as wars, epidemics, etc.

The fertility coefficients, computed for effective generations from 1914 to 1935, have been studied as a function of the woman's age and year of birth.

1. THE FERTILITY COEFFICIENT AS A FUNCTION OF THE WOMAN'S AGE

As a first step, the problem has been dealt with graphically: for every generation, woman's ages, from 13 to 50, have been indicated on the abscissa, while fertility coefficients have been indicated on the ordinate. (See Fig. 1).

a: *Year 1915* *b*: *Year 1935*

FIG. 1. *Fertility coefficient as a function of the woman's age*

In this way, for the generations from 1914 to 1935, a family of curves has been obtained; since they correspond to rates, these show some irregularities, the largest one being due to the influence of World War II.

https://doi.org/10.1017/S1120962300010969 Published online by Cambridge University Press

The comparison of the above curves with some theoretic curves shows that they may be rather well approximated by a gamma curve of the form,

$$\gamma(x) = K(x - 13)^{\alpha}\, e^{-\lambda(x-13)}. \qquad [2]$$

In order to localize the flex points of $\gamma(x)$, it may be noted that they are symmetric with respect to the location of the maximum.

The behavior of the first finite differences of the data concerning the above mentioned years has then been studied. The flex points (where the first differences reach a maximum or a minimum) and the location of the maximum (where the first differences are close to zero) coincide almost exactly for all the curves. In particular, the first flex point is found between the ages of 20 and 21, and shall therefore be taken as $x = 20.5$. The maximum is attained at $x = 26.5$. The second flex point is indeed symmetric to the first with respect to the maximum.

In conclusion, we can see that both shape of the curves and the check on the position of the flex points justify the interpolation of the data with a curve $\gamma(x)$.

2. EVALUATION OF THE CURVE $\gamma(x)$

We shall make use of the location of the maximum and of the first flex point in order to determine the values of the parameters α and λ. Then, the function's maximum fixes the scaling factor K.

Let us consider [2] concerning the year 1914. The first derivative may be made equal to zero in correspondence to the maximum, thus obtaining:

$$13.5\lambda - \alpha = 0; \qquad [3]$$

hence, $\alpha = 13.5\lambda$.

Making then the second derivative equal to zero in correspondence to the first flex point, we obtain:

$$7.5 = \alpha - \sqrt{\alpha}. \qquad [4]$$

Let us consider the system of equations [3] and [4]; we obtain:

$$\alpha = 5.02 \text{ and } \lambda = 0.375.$$

In order to determine the scaling factor K, let us consider the maximum of the function [2]:

$$\gamma(26.5) = 180 = K(13.5)^{5.02} \exp\,[(-0.375)\ (13.5)]$$

and then

$$K = 180(13.5)^{-5.02} \exp\,[(0.375)\ (13.5)] = 0.063.$$

We have thus completely determined the $\gamma(x)$ function, which gives us the Italian woman's fertility coefficient, relative to the year 1914, as a function of the age:

$$\gamma(x) = 0.063(x - 13)^{5.02} \exp\,[-0.375\ (x - 13)]. \qquad [5]$$

3. THE FERTILITY COEFFICIENT AS A FUNCTION OF THE YEAR OF BIRTH

Also in this case, the problem shall be dealt with graphically.

The year of birth of the woman, from 1914 ($t = 0$) to 1935, has been plotted on the abscissa, while the fertility coefficients have been plotted on the ordinate, so that a curve has been obtained for each age from 24 to 35 (Fig. 2). However, the interpretation of these graphs is not so straightforward as in the previous case.

Our data being insufficient, we had recourse to the Vital Statistics 1962 Yearbook of the Italian Institute of Statistics (ISTAT). There we found graphs representing

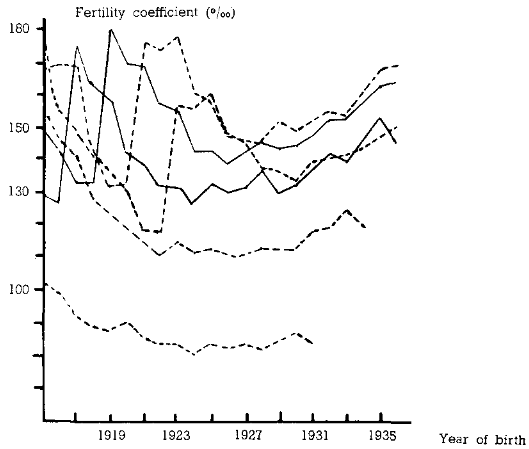


FIG. 2. *Fertility coefficient as a function of the year of birth (Age 20-35)*

the evolution of birth rates in Italy from the beginning of the century. A substantial agreement could thus be observed between birth rate and fertility coefficient, when our data allowed the comparison. Because of the lack of the data required for a direct study, it seemed correct to assume, for the fertility coefficient, a similar behaviour to that of the birth rate.

Neglecting, as already explained, the effects of World War II, and taking into account the information drawn from the ISTAT Yearbook (Fig. 3), the trend of the fertility coefficient as a function of the year of birth has been approximated by an exponential decreasing curve like

$$f(t) = \exp\,(-\mu t). \qquad [6]$$

It may be seen that the curve [6] does often not really fit the observed rates. The errors are due to random events of several kinds (economic, climatic, etc.) that may influence the birth rate. This may be taken into account by adding to [6] some random variables with a well determined probability distribution. For our purposes, however, it is enough to know the function given by [6].

https://doi.org/10.1017/S1120962300010969 Published online by Cambridge University Press

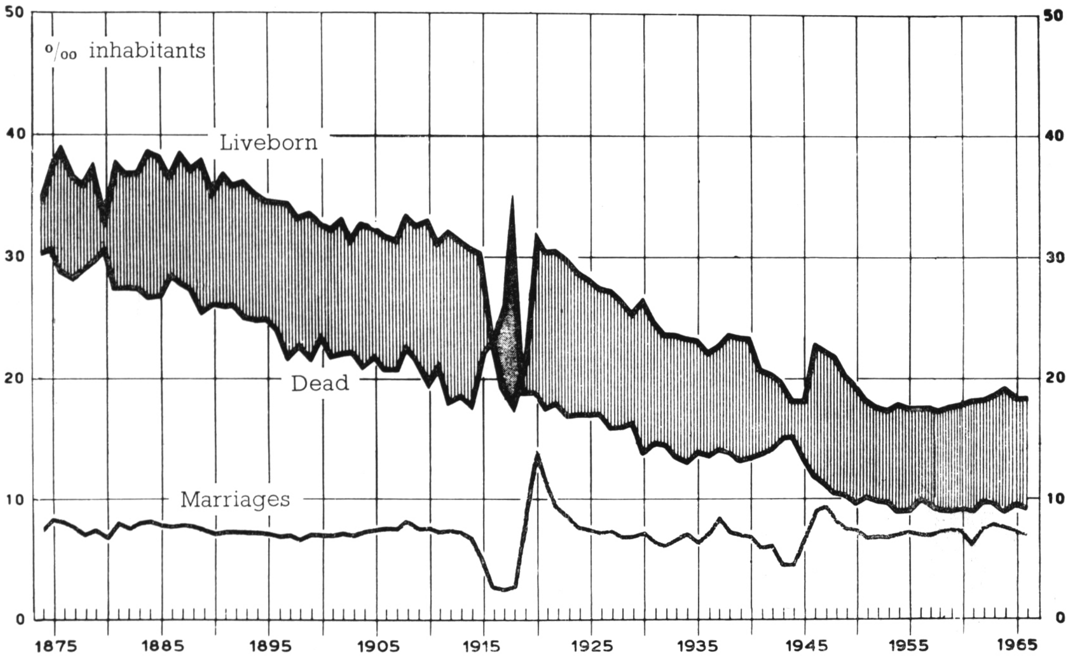


FIG. 3. *Evolution of birth rates in Italy* [ISTAT Yearbook for 1962].

4. EVALUATION OF THE EXPONENTIAL PARAMETER μ

We have considered the family of curves representing the most fertile ages for the woman, 24 to 35, and observed to what extent, in the average, the fertility coefficients for those ages decreased from 1914 to 1935. A difference of 20/1,000 was found between the coefficient of the first year and that of the last one. Taking as a unity on the ordinate the ordinate of the above curves in the point, $t = 0$, the difference between the values in the first and the last year is 0.12.

This allows us to evaluate μ:

$$\mu = 0.00656.$$

5. THE FERTILITY COEFFICIENT AS A FUNCTION OF THE WOMAN'S AGE AND YEAR OF BIRTH

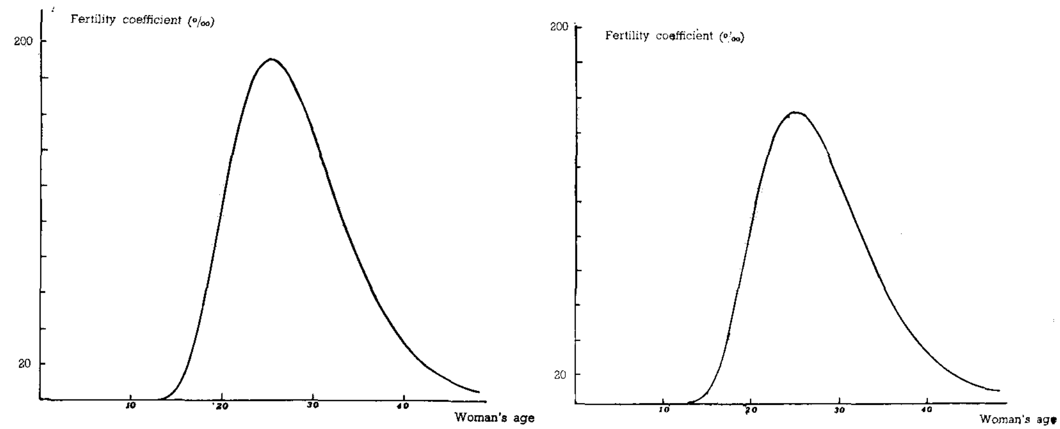
The trend of the Italian woman's fertility coefficient, as a function of her age and year of birth, may be expressed by the product [1] of the two functions [5] and [6]:

$$g(x, t) = \gamma(x)\, f(t) = 0.063\ (x - 13)^{5.02} \exp \{- [0.375\ (x - 13) + 0.00656t]\}. \quad [7]$$

[1] From this product, we may have an account of the scaling factor k in [2] as a function of time.

https://doi.org/10.1017/S1120962300010969 Published online by Cambridge University Press

In order to verify that, and making use of the IBM 1130 computer, we have, obtained the graphs of $g(x, t)$, where the time appears as a parameter. These are a family of gamma curves ranging from 1914 to 1935 (Fig. 4). The envelope of the



a: *Year* 1915 *b*: *Year* 1935

FIG. 4. *Interpolating function of the fertility coefficient*

maximum points of those curves has the decreasing exponential behaviour given by [6] (Fig. 5).

A comparison with the curves obtained from the tables by Livi Bacci-Santini has shown that the function $g(x, t)$ yields a good description of reality (Fig. 6).

From [7] it is possible to obtain two other functions that are useful for the calculations to be now developed.

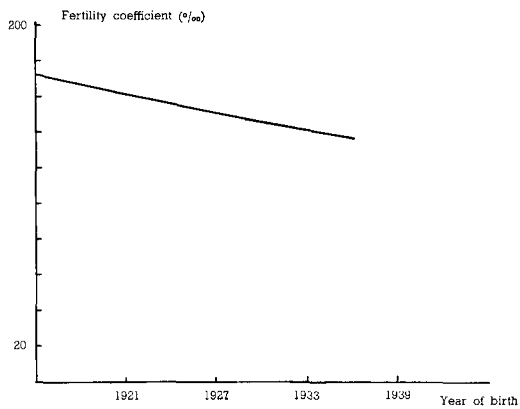
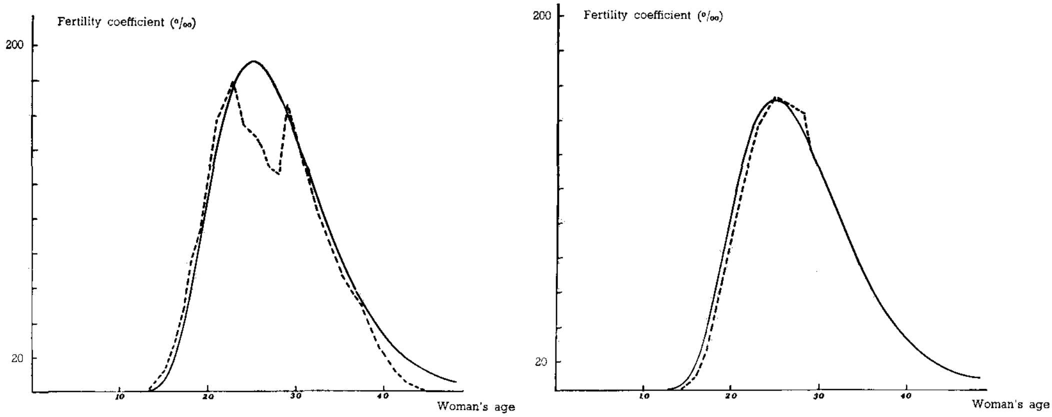


FIG. 5. *Trend function of the fertility coefficient*



a: *Year 1915* *b*: *Year 1935*

FIG. 6*a*. *Fertility coefficient as a function of the woman's age and its interpolating function*

Calling $z = t + x$ the actual time, we may obtain from $g(x, t)$, with a change of variable, the function $h(x, z)$ that yields the fertility coefficient of the Italian woman as a function of age and time:

$$h(x, z) = 0.063\ (x - 13)^{5.02} \exp \{- [0.375\ (x - 13) + 0.00656(z - x)]\}. \quad [8]$$

Similarly, we may obtain the function $P(t, z)$, that yields the fertility coefficient of the Italian woman as a function of year of birth and time:

$$p(t, z) = 0.063\ (z - t - 13)^{5.02} \exp \{- [0.375\ (z - t - 13) + 0.00656t]\}. \quad [9]$$

The function $g(x, t)$ may be used to evaluate the average total number of offsprings of women of a given generation. In fact, calling $D(t)$ the number of women born in the year t, the average number of women born in the year t, and aged between x and $x + 1$, will be $L(x, t) = D(t)l(x)$.

Thus, we have

$$N(t) = \sum_{13}^{50} {}_x\ L(x, t)\ g(x, t) \quad [10]$$

that describes the average total number of births from a woman of the generation of the year t.

https://doi.org/10.1017/S1120962300010969 Published online by Cambridge University Press

REFERENCES

D'Addario R. 1964. Rappresentazione analitica delle funzioni statistiche. In: Lezioni di Metodologia Statistica per Ricercatori. Ediz. dell'Istituto di Calcolo delle Probabilità e di Statistica, Roma.

Isaacson F., Keller H.B. 1966. Analysis of Numerical Methods. Wiley, New York.

ISTAT 1962. Annuario di Statistiche Demografiche. Roma.

Joshi D.D. 1954. Les Processus Stochastiques en Démographie. Publication de l'Institut de Statistique de l'Université de Paris.

Keyfitz N. 1968. Introduction to Mathematics of Population. Addison-Wesley, New York.

Lotka A.J. 1939. Théorie analytique des associations biologiques. II. Actualités Scientifiques et Industrielles, 780. Paris.

Lasorsa G. 1965. Metodologia Statistica. Liguori, Napoli.

Livi Bacci M., Santini A. 1969. Tavole di Fecondità della Donna Italiana Secondo le Generazioni di Appartenenza. Dipartimento di Statistica Matematica, Università di Firenze.

Ralston A. 1965. A First Course in Numerical Analysis. McGraw Hill, New York.

RIASSUNTO

Viene proposto un modello matematico rappresentativo dell'andamento del coefficiente di fecondità della donna italiana in funzione dell'età e dell'anno di nascita. Lo studio è stato basato sui dati delle tavole di Livi Bacci-Santini, che vengono interpretati da un'unica formula.

RÉSUMÉ

Un modèle mathématique est proposé, qui représente les variations du coefficient de fertilité de la femme italienne d'après l'age et l'année de naissance. L'étude se base sur les données des tables de Livi Bacci-Santini, qui sont interprétées par une formule unique.

ZUSAMMENFASSUNG

Es wird ein mathematisches Modell vorgeschlagen, das den Trend des Fruchtbarkeitskoeffizienten bei den Italienerinnen nach Alter und Geburtsjahr darstellt. Die Untersuchung stützt sich auf die Daten der Tafeln nach Livi Bacci-Santini, die in einer einzigen Formel ausgedrückt werden.

Dr.ssa Carla Rossi, Istituto di Calcolo delle Probabilità, Facoltà di Scienze Statistiche dell'Università, Rome, Italy.