J. Biosci., Vol. 3 Number 2, June 1981, pp. 197–205. © Printed in India.

A probability model for the number of births in an equilibrium birth process

S. N. SINGH and V. K. SINGH*
Demographic Research Centre, Banaras Hindu University, Varanasi 221 005

* Department of Statistics

MS received 18 February 1980

Abstract. A probability model for the number of complete conceptions (that is, live births) taking into account foetal wastages, occurring in a couple during a specified period of time (T_0, T_0+T) is developed assuming that the data was collected starting a long time after marriage. A method of estimating some of the underlying parameters is given. The model is applied to data obtained in a Varanasi Survey in 1969-70.

Keywords. Equilibrium birth process; estimation of fecundity; sterility; method of moments.

Introduction

The majority of the probability models (Brass, 1958; Pathak, 1966,1970; Perrin and Sheps, 1964; Singh, 1961,1964,1968) were developed using the renewal theory and dealt with a sequence of occurrences of a defined event such as conception or birth. These models assumed that the female was exposed to the risk of conception at the beginning of the observational period. This assumption is not always true, especially when the data on the number of births per female were collected in a retrospective survey. The information in such situations may relate to the number of births observed for any segment of the reproductive life of the female. Obviously for the same reason, models dealing with the number of births during the period (O, T), where O refers to the time of marriage, are not applicable in such situations.

Sheps *et al.* (1969) have pointed out that if the observations, instead of starting at marriage for each woman, begin at a later date and the process has been in effect for a long time, the distribution of births becomes an equilibrium birth process. Recently, Singh and Yadava (1977) obtained a model for the number of conceptions in an equilibrium birth process. Since data regarding the number of live births are readily available and involve no memory bias, it is better to develop a probability model for number of births in order to estimate fecundity and other fertility parameters.

In the present paper, a model describing the variation in the number of births in an equilibrium birth process has been derived. Since a significant proportion of conceptions results in foetal wastage, it has been taken into account. Though fertility parameters vary at the individual level, like the female's age and pregnancy history and also between females, they are assumed to be constant here at both the levels due to the complex nature of the model. The suggested model and the underlying assumptions are given below.

The model

Let $X(T_0, T)$ denote the number of complete conceptions (that is, live births) to a female during the period (T_0, $T_0 + T$), where T_0 is measured from marriage, that is, the observation starts after a time T_0 from marriage. $X(T_0, T)$ can assume values 0, 1,2, …… The probability distribution of $X(T_0, T)$ is derived under the following assumptions:

(i) The female has led a married life from marriage to the end of the observational period.
(ii) A conception may result in a live birth or in a foetal wastage. If a conception results in a live birth it is called complete, failing which it is known as incomplete. Let θ and $(1-\theta)$ be the respective proabablilities $(0<\theta<1)$.
(iii) The waiting time for first conception since marriage follows an exponential distribution with a probability density function (PDF)

$$f(t) = \lambda e^{-\lambda t}, \quad t>0, \quad \lambda>0 \qquad (1)$$

and the time between the rth and $(r+1)$th conceptions (r=1, 2, ...) follows displaced exponential distribution with PDF

$$f(t) = \lambda e^{-\lambda (t-h_1)}, \quad t>h_1, \quad \lambda>0, \qquad (2)$$

or

$$f(t) = \lambda e^{-\lambda (t-h_2)}, \quad t>h_2, \quad \lambda>0, \qquad (3)$$

depending on whether rth conception is incomplete or complete.

The above assumption makes it clear that if there is a conception at a point of time, no other conception is possible either in h_1 or h_2 units of time according as the conception is incomplete or complete. These periods of temporary sterility are known as 'rest periods' $(h_1<h_2)$.

(iv) Either a female is exposed (except during a pregnancy or post-partum amenorrhoea period) to the risk of conception throughout the observational period or she is not exposed to this risk at any time during the interval (T_0, $T_0 + T$). Let α and $(1-\alpha)$ be the respective probabilities.

(v) T_0 is at a considerable distance from marriage.

Under the above assumptions the distribution of $X(T_0, T)$ is given by

$$P_0 = 1 - \alpha + \alpha\left[H_0(T) - H_1(T)\right] \tag{4}$$

$$P_r = \alpha\left[H_r(T) - H_{r+1}(T)\right], \quad r = 1, 2, \ldots\ldots, n-1 \tag{5}$$

$$P_n = \alpha H_n(T), \tag{6}$$

where $H_0(T) = 1$ and

$$H_{r+1}(T) = \sum_{\text{over } j} \binom{r+j}{j} \frac{(1-\theta)^j \theta^{r+1}}{1+\lambda\bar{h}} \Big[G_1(T,\lambda) - (1-\theta) G_2(T,\lambda,h_1) - \theta G_2(T,\lambda,h_2)\Big]$$

for $T > jh_1 + rh_2 + h_2$

$$= \sum_{\text{over } j} \binom{r+j}{j} \frac{(1-\theta)^j \theta^{r+1}}{1+\lambda\bar{h}} \left[G_1(T,\lambda) - (1-\theta) G_2(T,\lambda,h_1)\right]$$

for $jh_1 + rh_2 + h_1 < T \leq jh_1 + rh_2 + h_2$

$$= \sum_{\text{over } j} \binom{r+j}{j} \frac{(1-\theta)^j \theta^{r+1}}{1+\lambda\bar{h}} G_1(T,\lambda)$$

for $jh_1 + rh_2 < T \leq jh_1 + rh_2 + h_1$, ...
$j = 0, 1, 2, \ldots\ldots;\ r = 0, 1, 2, \ldots\ldots, n-1,$

where (7)

$$G_1(T,\lambda) = \lambda(T - jh_1 - rh_2) - (j+r) + e^{-\lambda(T-jh_1-rh_2)} \sum_{s=0}^{j+r-1} \sum_{k=0}^{s} \frac{[\lambda(T-jh_1-rh_2)]^k}{\lfloor k}$$

and

$$G_2(T,\lambda,h_i) = \lambda(T - jh_1 - rh_2 - h_i) - (j+r+1) + e^{-\lambda(T-jh_1-rh_2-h_i)} \sum_{s=0}^{j+r} \sum_{k=0}^{s} \frac{[\lambda(T-jh_1-rh_2-h_i)]^k}{\lfloor k}, \quad i = 1, 2.$$

The derivation of the probability distribution for $X(T_0, T)$ is given below in a separate section.

Estimation

The model consists of five parameters, .h_1, h_2, θ, α and λ. Although with suitable statistical techniques all the model parameters can be estimated simultaneously, it is not advisable using some observed distribution containing only a few cells with significant number of observations. Since the values of some of the fertility parameters are more or less constant from sample to sample, their values can be assumed beforehand on an empirical basis. The procedure of estimation of other parameters can then be developed. In this context, assuming the values of outcome-dependent rest periods h_1 and h_2 and the probability of an incomplete conception θ, a procedure for estimating rest of the parameters λ and α is given below:

If $X_1, X_2 \ldots, X_N$ is a random sample of size N from a population, then

$$E(X) = \sum_{r=0}^{n} r P_r = \alpha U(\lambda), \tag{8}$$

where $U(\lambda)$ denotes a function of λ. Since the samples are taken from the same population, we have

$$E(\bar{x}) = \alpha U(\lambda), \tag{9}$$

where is x the sample mean. Following Cox and Miller (1967) and Singh and Yadava (1977) an approximate expression of $U(\lambda)$ can be written as

$$U(\lambda) = \frac{\theta \lambda T}{1 + \lambda \bar{h}} \; ; \; \bar{h} = (1-\theta) h_1 + \theta h_2, \tag{10}$$

so that equation (9) can be written as

$$\bar{x} = \frac{\alpha \theta \lambda T}{1 + \lambda \bar{h}}. \tag{11}$$

Further, let $N_r (r=0, 1, \ldots, n)$ denote the number of couples who have r complete conceptions in the sample. Equating the observed proportion of couples with $r = 0$ to expression (4) we get

$$\alpha = \frac{N - N_0}{N H_1(T)} \tag{12}$$

Using equations (11) and (12), α and λ, the estimates of α and λ respectively, can be obtained.

Derivation of the model

Under the assumptions stated previously, the distribution of $X(T_0, T)$ becomes independent of the starting point T_0 (Cox and Miller, 1967); henceforth we can write $X(T)$ for $X(T_0, T)$.

Let the conceptions be counted from the point T_0 and let the successive concepttions occur at times $T_0 + T_1$, $T_0 + T_1 + T_2, \ldots,$ so that T_1, be the time of first conception from T_0 and T_r $(r>1)$ be the time between $(r^{-}1)$th and rth conceptions. Since T_0 is at a considerable distance from marriage, the above birth process becomes an equilibrium birth process. Hence, the waiting times $T_2, T_3, \ldots,$ are independently and identically distributed but T_1 has a different distribution given by

$$f_1(t) = \frac{1-F(t)}{\mu}, \tag{13}$$

where $F(t)$ and μ are respectively, the distribution function and mean corresponding to the PDF of T_r ($r > 1$) (Cox and Miller, 1967). From assumptions (ii) and (iii), it is easy to verify that

$$\mu = \frac{1+\lambda\bar{h}}{\lambda} \; ; \; \bar{h} = (1-\theta)h_1 + \theta h_2 \tag{14}$$

and the P.D.F of T_1 is T_1 is

$$f_1(t) = \begin{cases} 1/\mu , & 0<t<h_1 \\ \dfrac{1-(1-\theta)\left[1-e^{-\lambda(t-h_1)}\right]}{\mu} & , h_1<t<h_2 \\ \dfrac{1-(1-\theta)\left[1-e^{-\lambda(t-h_1)}\right]-\theta\left[1-e^{-\lambda(t-h_2)}\right]}{\mu} & t>h_2 \end{cases} \tag{15}$$

Let $F_{j+1}(t/j)$ be the distribution function of X_0, which is the waiting time of (j+l)th conception on the assumption that preceding j conceptions are incomplete. Then, clearly

$$X_0 = \sum_{r=1}^{j+1} T_r \tag{16}$$

where T_1 has PDF (13) and $T_2, T_3, \ldots, T_{j+1}$ each has PDF

$$f(t) = \lambda e^{-\lambda(t-h_1)}, \; t>h_1, \; \lambda>0 . \tag{17}$$

Since T_1 $T_2, \ldots, T_{j+1}$ are mutually independent random variables, the distribution function $F_{j+1}(t/j)$ is given by

$$F_{j+1}(t/j) = \int_0^t F(t-u)\, dF^j(u) \tag{18}$$

where $F(t-us)$ is the distribution function of T_1 and Fj (u) is the j-fold convolution of the distribution function associated with the PDF (17). It can be seen that

$$\begin{aligned} F_{j+1}(t/j) &= \frac{A_1(t,\lambda)}{1+\lambda\bar{h}} \quad \text{for } jh_1<t\leq jh_1+h_1 \\ &= \frac{1}{1+\lambda\bar{h}}\left[A_1(t,\lambda)-(1-\theta)A_2(t,\lambda,h_1)\right] \\ &\qquad \text{for } jh_1+h_1<t\leq jh_1+h_2 \\ &= \frac{1}{1+\lambda\bar{h}}\left[A_1(t,\lambda)-(1-\theta)A_2(t,\lambda,h_1)-\theta A_2(t,\lambda,h_2)\right] \\ &\qquad \text{for } t>jh_1+h_2 , \end{aligned} \tag{19}$$

where

$$A_1(t,\lambda) = \lambda(t-jh_1) - j + e^{-\lambda(t-jh_1)} \sum_{s=0}^{j-1} \sum_{k=0}^{s} \frac{[\lambda(t-jh_1)]^k}{\lfloor k} \tag{20}$$

and

$$A_2(t,\lambda,h_i) = \lambda(t-jh_1-h_i) - (j+1) + e^{-\lambda(t-jh_1-h_i)} \sum_{s=0}^{j} \sum_{k=0}^{s} \frac{[\lambda(t-jh_1-h_i)]^k}{\lfloor k}, \quad i = 1,2. \tag{21}$$

Since $\theta(1-\theta)^{j}$ is the probability that there are j incomplete conceptions preceding a complete conception, the distribution function of the waiting time of the first complete conception from T_0 is

$$P[X_0 \leq T] = H_1(T) = \sum_{\text{over } j} \theta(1-\theta)^j \, F_{j+1}(t/j); \quad j = 0,1,2,\ldots\ldots,[T/h_1] \tag{22}$$

If $\phi_1(s)$ be the Laplace transform of $H_1(T)$, then

$$\phi_1(s) = \left[\frac{\theta}{\mu s} - \theta(1-\theta)\frac{e^{-sh_1}}{\mu s}\frac{\lambda}{\lambda+s} - \theta^2 \frac{e^{-sh_2}}{\mu s}\frac{\lambda}{\lambda+s}\right] \left[1-(1-\theta)e^{-sh_1}\frac{\lambda}{\lambda+s}\right]^{-1} \tag{23}$$

Now let X_i be the time between ith and $(i+1)$th complete conception ($i>0$). It is clear that the distribution function of X_i ($i>0$) will differ from X_0 only due to the distribution of T_1.

Here

$$X_i = \sum_{r=i(j+1)+1}^{(j+1)(i+1)} T_r, \quad (i>0) \tag{24}$$

where $T_{ij+(i+1)}$'s ($i = 1,2,3\ldots$) have the displaced exponential distribution given by (3). Hence the Laplace transform $\varphi_2(s)$, corresponding to the PDF OF X_i ($i = 1,2,\ldots$) is the same as given by Singh *et al* (1973),

$$\phi_2(s) = \left(\frac{\lambda}{\lambda+s}\right) e^{-sh_2} \left[1-(1-\theta)\frac{\lambda}{\lambda+s} e^{-sh_1}\right]^{-1} \tag{25}$$

Since X_0, X_1, . . . are independent random variables, the Laplace transform of the waiting time of $(r + 1)$th complete conception from T_0 is

$$\begin{aligned}\phi_{r+1}(s) &= \phi_1(s)\left[\phi_2(s)\right]^r \\ &= \sum_{j=0}^{\infty}\binom{r+j}{j}(1-\theta)\,\theta^{r+1}\left(\frac{\lambda}{\lambda+s}\right)^{j+r}\frac{e^{-s(jh_1+rh_2)}}{\mu s} \\ &\quad - \sum_{j=0}^{\infty}\binom{r+j}{j}(1-\theta)^{j+1}\theta^{r+1}\left(\frac{\lambda}{\lambda+s}\right)^{j+r+1}\frac{e^{-s(jh_1+rh_2+h_1)}}{\mu s} \\ &\quad - \sum_{j=0}^{\infty}\binom{r+j}{j}(1-\theta)^{j}\theta^{r+2}\left(\frac{\lambda}{\lambda+s}\right)^{j+r+1}\frac{e^{-s(jh_1+rh_2+h_2)}}{\mu s}\end{aligned} \qquad (26)$$

Now if $H_{r+1}(T)$ be the inverse of (26), we have expression (7) and hence under the given assumptions the model (4), (5) and (6) follow.

Application and results

To test the validity of the suggested model, it has been applied to the data on number of births recorded in a Varanasi Survey of 1969-70. The survey was conducted by the Demographic Research Centre, Banaras Hindu University in about 2200 households scattered in 52 villages of Varanasi (Rural). The reference date of the survey was October 1969. The details of the survey methodology are given in Singh *et al* (1970).

Table 1 presents the observed distribution of the number of live births in a fixed period of seven years before the reference date to women whose present age was in the range 25-30 years. As the average age at marriage for women in the surveyed area is nearly 15 years and the data relates to women aged 25-30 years, the start of the observational period may be assumed to be at a considerable distance from marriage.

Table 1. Distribution of the observed and expected number of births in the last seven years to women aged 25-30 years in the Varanasi survey.

Number of births	Observed number of women	Expected number of women
0	22	22.0
1	58	63.7
2	174	171.7
3	142	132.0
4	32	35.4
5	3	6.2
Total	431	431.0

$x^2 = 3.3$ (degrees of freedom = 3)

Before applying the model to the data, it is necessary to point out the following. An interesting feature of the distribution of post-partum amennorrhoea (p.p.a) associated with a live birth, in the survey, is that it has two peaks. Thus, if it is assumed that the rest period associated with a live birth h_2 takes two values and h'_2 and h''_2 and π and $(1\ \pi)$ are the proportion of women with the value of h_2 is as and h'_2 and h''_2 respectively, the model given by (4), (5), (6) and expressions for estimating λ and α, (11) and (12) respectively are slightly modified as

$$P_r^* = \pi P_r(h_2 = h'_2) + (1-\pi) P_r(h_2 = h''_2)$$

$$r = 0, 1, 2, \ldots\ldots, n \qquad (27)$$

$$\bar{X} = \alpha\theta\left[\frac{\pi\lambda T}{1+\lambda \bar{h}_1} + \frac{(1-\pi)\lambda T}{1+\lambda \bar{h}_2}\right] \qquad (28)$$

$$\hat{\alpha} = \frac{N-N_0}{N[\pi H_1(T/h_2 = h'_2) + (1-\pi) H_1(T/h_2 = h''_2)]} \qquad (29)$$

where $h_1 = (1 - \theta)\ h_1 + \theta h'_2$ and $h_2 = (1 - \theta)\ h_1 + \theta h''_2$ in the expression (14). $P_r(h_2 = h'_2)$ and $P_r(h_2 = h''_2)$, $r = 0,1,2, \ldots, n$ in the expression (27) above are given by equations (4), (5) and (6) respectively by substituting $h_2 = h'_2$ and h''_2. The meaning for $H_1\ (T/h_2 = h'_2)$ and $H_1\ (T/h_2 = h''_2)$ is similar to that in expression (29).

In order to apply the modified model given by (27), we take T= 7 years. Singh and Bhaduri (1971) and Singh (1978) have shown that the two peaks of h_2 occur at 3 months and 12 months respectively. Hence h'_2= 1.00 year and h''_2= 1.75 years. Further, we assume h_1 = 0.50 years, θ = 0.90 and π= 0.40. These results also agree with those given in Singh and Bhaduri (1971). Using these values, the estimates of α and λ are obtained, after a few iterations, as

$$\alpha = 0.9644, \qquad \lambda = 0.73.$$

Once the estimates of α and λ are obtained, the expected frequencies are easily calculated. The observed and expected frequencies are presented in the second and third columns respectively in table 1. The insignificant value of λ^2-statistic at 5% level of significance suggests that the model under consideration is a better approximation to the situation.

Conclusions

The present model may be utilized to analyse the data on the number of births in different age segments of a woman's reproductive life, as it considers the situation that some of them may not be susceptible for conception at the time when the investigator started to collect the data. The model may repeatedly be used with some further refinement for finding out the fertility performance of a woman as she passes through different age groups.

References

Brass, W. (1958) *Popul. Stud.,* **12,** 51.
Cox, D. R. and Miller, H. D. (1967) *The theory of stochastic process,* (London: Methuen and Co.)
Pathak, K. B. (1966) *Sankhya,* **B28,** 213.
Pathak, K. B. (1970) *Artha Vijnana,* **12,** 429.
Perrin, E. B. and Sheps, M. C. (1964) *Biometrics,* **20,** 28.
Sheps, M. C., Menken, J. A. and Radick, A. P. (1969) *Demography,* **6,** 161.
Singh, S.N. (1961) *A hypothetical chance mechanism of variation in number of births per couple,* Ph.D. thesis, University of California, Berkeley, U.S.A.
Singh, S.N. (1964) *Sankhya,* **B26,** 89.
Singh, S. N. (1968) *J. Am. Stat. Assoc,* **63,** 209.
Singh, V. K. (1978) *Some mathematical models for couple fertility and their applications,* Ph. D. thesis, Banaras Hindu University, Varanasi.
Singh, S. N. and Bhaduri, T. (1971) *Proceedings of All India Seminar on Demography and Statistics,* Banaras Hindu University, Varanasi, p. 284.
Singh, S. N., Bhattacharya, B. N. and Joshi, P. D. (1973) *J.Indian Soc. Agr. Stat.,* **25,** 91.
Singh, S. N. and Yadava, R. C. (1977) *Demography India,* **6,**163.
Singh, S. N., Yadava, R. C. and Bhaduri, T. (1970) *Prajna,* **16,**176.