ON A HYPOTHETICAL MODEL FOR THE NUMBER OF BIRTHS AND ITS APPLICATION

S.N. SINGH, B.N. BHATTACHARYA AND P.D. JOSHI
Demographic Research Centre, Banaras Hindu University
(Received on 9-1-1971)

1. INTRODUCTION

Problems concerning the variation in the number of births to a couple, during a given period of time, have been of interest to planners, demographers and other social scientists. Of late, considerable stress has been laid on the study of the above phenomenon with the help of probability models. Brass (1958), Dandekar (1955), Henry (1956), Pathak (1956), Singh (1963, 1964, 1968) and others have derived probability distributions for the number of complete conceptions (a conception is complete if it results in a live birth otherwise it is incomplete) to a couple on varying sets of assumptions. These distributions are useful in getting the estimates of parameters involved in human reproduction. They are also useful in the evaluation of different family planning programmes. The above mentioned distributions are based, among other things on the assumption that there is one to one correspondence between a conception and a birth. A probability model making allowances for foetal losses would be more realistic to describe the phenomenon and would provide better estimate of fecundability. Sheps and Perrin (1966) derived a probability distribution for the number of complete conceptions incorporating foetal losses which is an extension of the model given in Singh (1963). The purpose of this paper is to generalize the probability distribution of Singh (1968) for the number of complete conceptions to a couple during a given period of time under the assumptions of Singh and Bhattacharya (1970), which accounts for incomplete conceptions. Though the present model may be derived as a marginal distribution from the model of Singh and Bhattacharya (1971), an alternative derivation is given in Section 2, keeping in view the utility of the new procedure in other situations. The estimates of conception rate and of proportion of fecund couples have been obtained following the procedure cited in Singh (1964) and also Singh and Bhattacharya (1970). Expressions for asymptotic variances and covariances are also given. For illustration, the proposed distribution has been applied to observed data of Dandekar (1955). When any distribution is applied to observed data, reasonable values of the parameters assumed known at the time of estimation, are used. For the present distribution the value of three parameters are assumed known. Some discussions about the estimates of the parameters are presented in section 3.

SECTION 2

2.1. This section deals with the derivation of a continuous time probability distribution, based on the assumptions of Singh and Bhattacharya (1970), for describing the variation in number of complete conceptions to a female during the period (O,T). For completeness the assumptions are reproduced below.

Assumption 1: The female is susceptible to conception at the beginning of the observational period and has led a married life throughout the period of observation.

Assumption 2(*a*): The number of coitions of a couple during any arbitrary time interval (t_1, t_2), $0<t_1<t_2<T$ is a random variable and follows is Poisson distribution with the parameter $m_1 (t_2-t_1)$, $m_1>0$.

Assumption 2(b): The coitions are mutually independent and p_1, the probability that a coition results in a conception, is constant.

It is easy to see that the waiting time to the first conception is distributed exponentially. If X_0 denotes the waiting time for a female to conceive,

$$P(X_0 \leqslant t)=F_0(t)=1-e^{-mt};\ (t>0),\ m=m_1p_1$$

Assumption 3: Let θ be the probability that a conception is complete, so that $(1-\theta)$ is the probability of an incomplete conception.

Assumption 4: Every conception involves a period of temporary sterility, called the 'rest period', comprising the duration of pregenancy and post-partum amenorrhoea. Let h_1 and h_2 be the rest periods associated with incomplete and complete conceptions respectively.

Under this assumption the maximum number of complete conceptions to a female during the period (O,T) cannot be more than

n, where n-$[T/h_2]$ if $[T/h_2]=T/h_2$ otherwise it is equal to $[T/h_2]+1$; $[T/h_2]$ stands for the greatest integer not exceeding T/h_2.

Assumption 5: Either the female is exposed (except during a pregnancy or post-partum period) to the risk of a conception throughout the period (O,T) or she is not exposed to this risk at any time during the interval (O,T). Let a and $(1-a)$ be the respective probabilities.

Let X denote the number of complete conceptions to a couple during the time interval (O,T). X can take values $0, 1, 2, \ldots, n$. Under the assumptions 1, 2, 3, 4 and 5 the probability function, P_x of X is given by

$$P_o=1-a\left[\sum_{j=0}^{k_o}\theta\,(1-\theta)^j\right.$$

$$\left.\left\{1-e^{-m\,(T-jh_1)}\sum_{s=0}^{j}\frac{\{m\,(t-jh_1)\}^s}{s!}\right\}\right] \tag{2.1}$$

$$P_r=a\left[\sum_{j=0}^{k_{r-1}}\binom{r+j-1}{j}\theta^r(1-\theta)^j.\right.$$

$$\left\{1-e^{-m\,(T-jh_1-\overline{r-1}h_2)}\sum_{s=0}^{r+j-1}\frac{\{m\,(T-jh_1-\overline{r-1}h_2)\}^s}{s!}\right\}$$

$$-\sum_{j=0}^{k_r}\binom{r+j}{j}\theta^{r+1}\,(1-\theta)^j.$$

$$\left.\left\{1-e^{-m(T-jh_1-rh_2)}\sum_{s=0}^{r+j}\frac{\{m(T-jh_1-rh_2)\}^s}{s!}\right\}\right]$$

$$r=1, 2, \ldots, (n-1). \tag{2.2}$$

$$P_n=1-P\,(X\leqslant n-1)$$

where $k_r=[(T-rh_2)/h_1]$. (2.3)

Proof. According to Assumption 5, 'a' is the probability that a female may have 0, 1, 2,..., n complete conceptions during the period (O,T) while $(1-a)$ is the probability that she is not exposed to the risk of a conception during this period.

Let the successive complete conceptions occur at times Z_r, $r=1, 2,...,$; and let $T_r=Z_{r+1}-Z_r$ $(r\geqslant 1)$ denote the time between the $(r+1)$-th and r^{th} complete conceptions. Denote by T_o the time of first complete conception from the start of the observational period. From assumptions of the model it is obvious that the distribution of T_o will be different from those of other T_r's $(r\geqslant 1)$ since the starting point of T_o is not a rest period as in the case of other T_r's. It can easily be seen that T_r's $(r\geqslant 0)$ are mutually independent and $T_1, T_2,...$ (but not T_o) have a common distribution.

Now we derive the distribution function $F_{j+1}(t/j)$ of T_o/j, the waiting time of the $(j+1)$-th conception on the assumption that the preceding j conceptions are incomplete. Here

$$T_o/j=\sum_{i=0}^{j} X_i \tag{2.4}$$

where X_i $(1\leqslant i\leqslant j)$ is the waiting time between the i^{th} incomplete conception and the time of $(i+1)$-th conception. X_o is the time of first conception. By our assumptions the X_i's $(i=1, 2,...,j)$ are independently and identically distributed random variables with the distribution function

$$\left.\begin{aligned} F_1(t)=P(X_i\leqslant t)&=1-e^{-m(t-h_1)} && \text{for } t>h_1 \\ &=0 && \text{for } t\leqslant h_1 \end{aligned}\right\} \tag{2.5}$$

Following the lines of derivation of equation (6.49) in (Ch. 6) Bharucha-Reid (1960), $F_{j+1}(t/j)$ is given by

$$\left.\begin{aligned} F_{j+1}(t/j)&=1-e^{-m(t-jh_1)}\sum_{s=0}^{j}\frac{\{m(t-jh_1)\}^s}{s\,!} && \text{for } t>jh_1 \\ &=0 && \text{for } t\leqslant jh_1 \end{aligned}\right\} \tag{2.6}$$

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

$$P(T_o\leqslant t)=H_1(t)$$

$$= \sum_{j=0}^{[t/h_1]} \theta(1-\theta)^j \left[1 - e^{-m(t-jh_1)} \sum_{s=0}^{j} \frac{\{m(t-jh_1)\}^s}{s\,!} \right]$$

for $t>0$ (2.7)

Let $\phi_0(s)$ be the Laplace transform of $H_1(t)$.

$$\phi_0(s) = \int_0^{\infty} e^{-st} dH_1(t) = \frac{\theta m}{(m+s)} \left[1 - \frac{(1-\theta)e^{-sh_1}m}{(m+s)} \right]^{-1} \quad .. (2.8)$$

The expressions for distribution function as well as the Laplace transformation for the first complete conception differ from other intervals between consecutive complete conceptions only in the component of rest period following a complete conception (h_2). Hence in the case of $T_r (r \geqslant 1)$ we have the distribution function and Laplace transformation given by

$$H_1^*(t) = P(T_r \leqslant t)$$

$$= \sum_{j=0}^{[(t-h_2)/h_1]} \theta(1-\theta)^j \left[1 - e^{-m(t-jh_1-h_2)} \cdot \sum_{s=0}^{j} \frac{\{m(t-jh_1-h_2)\}^s}{s!} \right] \text{ for } t > h_2$$

$= 0$ otherwise. ...(2.9)

and

$$\phi_1(s) = \theta \left(\frac{m}{m+s} \right) e^{-sh_2} \left[1 - \frac{(1-\theta)e^{-sh_1}m}{m+s} \right]^{-1} \quad ...(2.10)$$

If $\phi_{r+1}(s)$ denotes the Laplace transform of Z_{r+1}, then

$$\phi_{r+1}(s) = \phi_0(s)[\phi_1(s)]^r$$

$$= \sum_{j=0}^{\infty} \binom{r+j}{j} \theta^{r+1} (1-\theta)^j e^{-s(rh_2+jh_1)} \cdot \left(\frac{m}{m+s} \right)^{r+1+j} \quad ...(2.11)$$

If $H_{r+1}(t)$ is the inverse of (2.11), then

$$H_{r+1}(t)=\sum_{j=0}^{[(t-rh_2)/h_1]}\binom{r+j}{j}\theta^{r+1}(1-\theta)^j.$$

$$\left[1-e^{-m(t-rh_2-jh_1)}\sum_{s=0}^{r+j}\frac{\{m(t-rh_2-jh_1)\}^s}{s!}\right] \qquad \text{...(2.12)}$$

Under Assumptions 1, 2(*a*), 2(*b*), 3, 4 and 5 the probability, K_r, of at most r complete conceptions during (O, T) is

$$K_r=1-aH_{r+1}(T)\ ;\ r=0,\ 1,\ 2,\ldots,\ n-1 \qquad \text{...(2.13)}$$

$$K_n=1. \qquad \text{...(2.14)}$$

Hence the probability expressions (2.1), (2.2) and (2.3) follow.

For $\theta=1$, the probability expressions reduce to one obtained by Singh (1968).

2.2. This model involves five parameters a, m, h_1, h_2 and θ. For given values of h_1, h_2 and θ the estimates $\hat{a}$ and $\hat{m}$ of a and m respectively can be derived from the equations obtained by equating the observed proportion of couples with zero complete conception and the observed mean number of complete conceptions to their theoretical values.

It can be shown that for large N, $\sqrt{N}\,(\hat{a}-a)$ tends to $N(0,\ \sigma_3^2)$ and $\sqrt{N}(\hat{m}-m)$ tends to $N(0,\sigma_4^2)$ where

$$\triangle^2\sigma_3^2=[a\ H_1'(T)]^2\ \sigma_1^2+[a\ F'\ (m)]^2\ [aH_1(T)]\ [1-aH_1(T)]$$
$$-2[aH'_1(T)]\ [aF'(m)]\ [aF(m)]\ [1-aH_1(T)] \qquad \text{...(2.15)}$$

$$\triangle^2\sigma_4^2=[F(m)]^2\ [aH_1(T)]\ [1-aH_1(T)]+[H_1(T)]^2\sigma_1^2$$
$$-2[F(m)]^2\ [aH_1(T)]\ [1-aH_1(T)] \qquad \text{...(2.16)}$$

$$\triangle^2\sigma_{34}=[aF(m)]^2\ [H'_1(T)]\ [1-aH_1(T)]$$
$$-[aH_1(T)]\ [H'_1(T)]\sigma_1^2 \qquad \text{...(2.17)}$$

where σ_1^2=population variance,

$$\triangle^2=\{a[F(m)]\ [H'_1(T)]-a[F'(m)]\ [H_1(T)]\}^2$$

σ_{34}=Covariance between $\sqrt{N}(\hat{a}-a)$ and $\sqrt{N}(\hat{m}-m)$ and $F'(m)$ and $H'_1(T)$ denote the derivatives of $F(m)$ and $H_1(T)$ respectively with respect to m.

2.3. The observed distribution on the number of complete conceptions are not available, however, the model can be applied to the number of births per woman during interval (O, T) assuming one to one correspondence between a complete conception and single live birth. We apply the model to the data given in Table I in Dandekar (1955). The table given below shows the observed distribution of the number of children born to women in the age group 26-30, during a five year period. In the present case $T=5$ years. The justification for taking T to be 5 years has been discussed in Singh (1967). In order to apply the Model B_1 to the data, we take $h_1=0\cdot5$ years, $h_2=1\cdot25$ years and $O=0\cdot80$. The aforesaid assumptions affix the value of n to be 4. For the data $N_o=105$, $N=369$ and $\bar{x}=1\cdot2168$. Following the procedure given in Section 2.2, the estimates of a and m are obtained as

$$\hat{a}=0\cdot79596 \quad \text{and} \quad \hat{m}=0\cdot61.$$

The frequencies in column 3 of the Table 1 are based on these values. From expressions (2.15), (2·16) and (2.17) we get

$$\hat{V}(\hat{a})=0\cdot0008425, \quad \hat{V}(\hat{m})=0\cdot001622 \quad \text{and}$$

$$\hat{Cov}(\hat{a}, \hat{m})=-0\cdot000892.$$

TABLE 1

Distribution of the observed and the expected number of children for women in the age group 26-30

Number of children	*Observed frequencies*	*Expected frequencies*
0	105	105·00
1	114	112·70
2	118	118·31
3	29	32·05
4 or more	3	0·94
Total	369	369·00

SECTION 3

3.1. Singh (1963, 1964, 1968) assumed the mean duration of gestation plus post-partom amenorrhoea associated with a complete conception, h_2, to be 12 months and obtained the estimates of proporation of fecund couples and fecundability from Dandekar's data. Sheps (1966) has remarked "Examination of Singh's model together with the data shows, however, that a higher value (of h) would produce inadmissible estimates for the other parameters". For this purpose the data given in Table 1 in Dandekar (1955) (reproduced in Table 1) has been used and the same procedure for estimation as given in Section 2.2. has been applied. The estimates of a and m have been obtained for different values of h_2 and θ, taking h_1 to be 0·5 year in each case, as shown in the Table.

TABLE 2

Estimates of a and m for different values of h_2 and θ

h_2 (*in years*)	θ	$\hat{m}$	$\hat{a}$
1·00	1·00	0·400	0·827
,,	0·85	0·485	0·829
,,	0·80	0·524	0·829
,,	0·75	0·575	0·830
,,	0·70	0·610	0·845
1·25	1·00	0·460	0·795
,,	0·85	0·560	0·795
,,	0·80	0·610	0·795
,,	0·75	0·660	0·797
,,	0·70	0·710	0·801
1·50	1·00	0·535	0·768
,,	0·85	0·600	0·784
,,	0·80	0.650	0·784
,,	0·75	0·700	0·786
,,	0·70	0·750	0·791
2·00	1·00	0·710	0·736
,,	0·85	0·900	0·736
,,	0·80	1·000	0·736
,,	0·75	1·100	0·736
,,	0·70	1·200	0·738

It is observed that even for larger values of h_2 the estimates are not inconsistent but the Chi square will not give a good fit. It should be noted that for the same data still larger value of Chi square has been reported by Dandekar. For a given h_2, as is evident from the Table 2, the variations in 'a' negligible for different values of θ but decreases gradually with increasing values of h_2. Also for a given h_2, m decreases with increase in θ but it increases with increase in h_2. In short the results show that the model is usefu l as a first approximation and pilot estimates can be obtained from the observed data with the help of the procedure cited in Section 2.2.

3.2. For simplicity of derivation of the model the parameters are assumed to be homogeneous in time. This assumption is approximately true during any short period of observation of the system. Empirical studies suggest that at least some parameters of reproduction are age-dependent. In order to apply the model to the data for longer durations of observational period, the variation in the parameters over time should be taken into account. Singh (1961, 1966) derived a probability distribution for the number of complete conceptions to a couple wherein he assumed the conception rate ($m(t)$) to be a function of time with no foetal wastages and a constant period of nonsusceptibility associated with each conception. Pathak (1970) has applied this distribution to the data of Saxena (1966) taking $m(t) = a+bt$, and come to the conclusion that the variation in $m(t)$ within five years for the data is small. The distribution of Singh has been extended in Bhattacharya (1970) which makes allowances for foetal deaths. Due to the coplexity of expressions the distribution is not presented here.

Summary

The probability of distribution for the number of complete conceptions to a couple during a given period of time has been generalised on certain assumptions. The estimates of conception rate of proportion of fecund couples have been obtained. Expressions for asymptotic variances and covariance are given. The proposed distribution have been applied to observed data for the sake of illustrations.

References

Bharucha-Reid, A.T. (1960) : Elements of the theory of Markov processes and their applications. McGraw-Hill Book Company, Inc., New York.

Bhattacharya, B.N. ((1970) : On some stochastic models for couple fertility. Thesis submitted for the degree of Ph. D. at Banaras Hindu University.

Brass, W. (1958) : The distribution of births in human populations. Population Studies, 12, 51-72.

Dandekar, V.M. (1955) : Certain modified forms of binomial and Poisson distributions. Sankhya 15, 237-51.

Henry, L. (1956) : Ancienness Familles Genevoises E'tude Demographique XVI-XX. siecle (presses Universitaires de France—Paris, 1956), 227-29.

Phathak, K.B. (1966) : A probability distribution for the number of conceptions. Sankhya 28-B, 213-18.

Pathak, K.B. (1970) : A time-dependent distribution for the number of conceptions. Artha Vijnana, Vol. 12, No. 3. 429-35.

Saxena, P.C. (1966) : A demographic Survey of Banaras Hindu University (unpublished).

Sheps, M.C. and Edward, B.P. (1966) : Further results from a Human fertility model with a variety of pregnancy outcomes. Human Biology, 38, 180-93.

Sheps, M.C. (1966) : Applications of probability models to the study of patterns of human reproduction. In M.C. Sheps and J.C. Ridley (eds.), Public Health and Population Change. Pittsburgh University of Pittsburgh Press.

Singh, R.P. (1967) : On a model of couple fertility. Sankhya, Series B, 29, 319-20.

Singh, S.N. (1961) : Thesis submitted for the degree of Ph.D. at the University of California, Berkeley, U.S.A.

Singh, S.N. (1963) : Probability models for the variation in number of births per couple. J. Amer. Stat. Assoc., 58 721-27.

Singh, S.N. (1964) : A Probability model for couple fertility. Sankhya Series, B, 26, 89-94.

Singh, S.N. (1966) : Some probability distributions utilized in human fertility. Seminar Volume in Statistics, B.H.U., 74-84.

Singh, S.N. (1968) : A change mechanism of variation in the number of births per couple. J. Amer. Stat. Assoc., 63, 209-13.

Singh, S.N. and Bhattacharya, B.N. (1970) : A generalized probability distribution for couple fertility. Biometrics, 26, 33-40.

Singh, S.N. and Bhattacharya, B.N. (1971) : On some probability distributions for couple fertility (to appear in Sankhya).