S. Mitra and S. N. Banerjee

An Analysis of Intervals Between Successive Orders of Birth

The Problem

DEMOGRAPHIC literature abounds in studies aimed at measuring the probabilites of conception allowing for the nonfecund period of pregnancy (Dandekar, 1955; Basu, 1955; Perrin and Sheps, 1963; Chiang, 1971; Srinivasan, 1972; Das Gupta, 1973; Bongaarts, 1975, to name a few). Keyfitz (1977) has provided an excellent summary of some of these works and has made significant contributions of his own. Das Gupta and Sheps took the fetal loss in consideration, in developing their models to generate the estimates of these probabilities. The distributions of birth intervals between successive birth orders provide the usual data source for some of these studies. In this investigation, we have also analyzed the same distributions somewhat differently, in order to estimate the lengths of the nonsusceptible periods that begin from the time of conception. A conception may end either in a live birth or in a fetal loss and accordingly, mathematical formulations of the estimates of the lengths of these two non-susceptible periods followed by those of the associated probabilities of these two types of conceptions will be attempted in this paper. We shall assume that, (1) the conditions remain favorable for a woman to conceive if she is in the susceptible state, (2) the non-susceptible state following a conception ending in a live birth remains the same for all women for a given order of birth, and (3) a constant non-susceptible period applies also to all conceptions ending in an abortion or in a still birth.

The Probability Distribution

The interval J between two successive live births can be decomposed into three additive components, namely, (1) the non-susceptible period following the first of the two live births plus the length of pregnancy of the second that ended in a live birth, say m, (2) a period of length kt_2 where k is the length of the non-susceptible period following a conception ending in a fetal loss (abortion or still birth) and t_2 is the number of such cases between the two successive live births, (3) the period t_1 in which the woman was in the susceptible state but did not conceive and finally (4) the unit of time where she conceived to give a live birth. The unit of time is traditionally taken to be a month, in which case, we can write J as a function of t_1 and t_2 (since m and k are constants) as

$$J(t_1, t_2) = m + t_1 + kt_2 + 1. \tag{1}$$

If in a given month during the susceptible period, the probability of a conception ending in a live birth is denoted by p and that ending in a wastage by r, then the probability of $J\,(t_1, t_2)$ can be expressed as

$$P\,[J\,(t_1, t_2)] = p \binom{t_1 + t_2}{t_1} q^{t_1} r^{t_2} \tag{2}$$

where,

$$q = 1 - p - r \tag{3}$$

is the monthly probability that no conception would take place. To summarize, the interval between the two live births is given by (1) in which given m and k, conception was avoided a total of t_1 times, took place $t_2 + 1$ times ending in wastage the first t_2 times and in a live birth the $(t_2 + 1)$th times and (2) provides a measure of its probability.

The Mean Length of the Interval

The sum of the probabilities for all values of t_1 and t_2 is

$$\sum_{t_1, t_2 = 0}^{\infty} P[J(t_1, t_2)] = p \sum_{n = 0}^{\infty} (q + r)^n = \frac{p}{1 - q - r} = 1 \tag{4}$$

as it should, where $n = t_1 + t_2$. The excepted value of

$$t = t_1 + kt_2 + 1 \tag{5}$$

can be expressed as

$$E(t) = E[t_1 + k(n - t_1)] + 1$$

$$= p \sum_{n=0}^{\infty} \sum_{t_1=0}^{\infty} \left[t_1 \binom{n}{t_1} q^{t_1} r^{n-t_1} + k(n - t_1) \binom{n}{t_1} q^{t_1} r^{n-t_1} \right] + 1.$$

$$= p(q + kr) \sum_{n=1}^{\infty} n(q + r)^{n-1} + 1 = \frac{p(q + kr)}{(1 - q - r)^2} + 1 = \frac{q + kr}{p} + 1 \quad (6)$$

Thus, the average length of the interval (see equations 1 and 5)

$$J(t_1, t_2) = m + t \quad (7)$$

can be expressed as

$$\mu_1 = E(m + t) = m + E(t) = m + \frac{q + kr}{p} + 1 \quad (8)$$

since m is a constant.

The Moment Generating Function Around the Mean

The moment generating function $\phi(z)$ of t can be written as

$$\phi(z) = p \sum_{n=0}^{\infty} \sum_{t_1=0}^{\infty} e^{z[t_1+k(n-t_1) - q + kr/p]} \binom{n}{t_1} q^{t_1} r^{n-t_1} \quad (9)$$

where the moments are measured around the mean. On simplification, (9) can be written as

$$\phi(z) = pe^{-z(q + rk)/p} \sum_{n=0}^{\infty} (qe^z + re^{kz})^n = \frac{pe^{-z(q+rk)/p}}{1 - qe^z - re^{kz}}. \quad (10)$$

The ith moment around the mean, i.e. μ_i can then be obtained as the ith derivative of $\phi(z)$ with respect to z at $z = 0$, i.e.

$$\mu_i = \phi^{(i)}(0). \quad (11)$$

Accordingly, we obtain

$$\mu_2 = \frac{q + rk^2}{p} + \left(\frac{q + rk}{p}\right)^2 \tag{12}$$

$$\mu_3 = \frac{q + rk^3}{p} + \frac{3(q + rk^2)(q + rk)}{p^2} + 2\left(\frac{q + rk)}{p}\right)^3 \tag{13}$$

$$\mu_4 = \frac{q + rk^4}{p} + \frac{4(q + rk^3)(q + rk) + 6(q + rk^2)^2}{p^2}$$

$$+ \frac{18(q + rk^2)(q + rk)^2}{p^3} + \left(\frac{9(q + rk)}{p}\right)^4. \tag{14}$$

Note that these are the moments of t around the mean, and therefore, are also the moments of $J(t_1, t_2)$ because of (7).

Estimates of the Parameters

Observe that the proposed model is based on four independent parameters, namely, m, k, p and r. The remaining parameter q can be obtained fiom (3). We propose to estimate the parameters from the equations generated by the first four moments. The equations are non linear in nature, but their forms are amenable to an iterative approach which may be attempted by rewriting the equations (8), (12), (13) and (14), as

$$\mu_1 = m + 1 + A \tag{15}$$

$$\mu_2 = A^2 + B \tag{16}$$

$$\mu_3 = 2A^3 + 3AB + C \tag{17}$$

$$\mu_4 = 9A^4 + 18A^2B + 6B^2 + 4\,AC + D \tag{18}$$

where

$$A = \frac{q + rk}{p} \tag{19}$$

$$B = \frac{q + rk^2}{p} \tag{20}$$

$$C = \frac{q + rk^3}{p} \tag{21}$$

and

$$D = \frac{q + rk^4}{p}. \tag{22}$$

Note that a trial solution of m is all that is needed to generate the values of A, B, C and D successively from equations (15) through (18). Thereafter, k can be solved either from

$$k = \frac{C - B}{B - A} \tag{23}$$

or from

$$k = \frac{D - C}{C - B}. \tag{24}$$

For the sake of consistency, the solutions obtained from (23) and (24) should be equal. Consequently, the condition

$$k = \frac{D - C}{C - B} = \frac{C - B}{B - A} \tag{25}$$

must be met. The trial solution of m can be suitably adjusted to that effect and once k is known r/p can be solved from, say, (19) and (20) as

$$\frac{r}{p} = \frac{B - A}{k(k - 1)} \tag{26}$$

and, thereafter, q/p can be obtained from (8) as

$$\frac{q}{p} = \mu_1 - m - 1 - \frac{kr}{p} \tag{27}$$

As the left hand sides of (26) and (27) add up to

$$\frac{q + r}{p} = \frac{1}{p} - 1 \tag{28}$$

because of (3), all the parameters are now known.

No Pregnancy Wastage

For the special case when all conceptions end in live births, that is, when $r = 0$, and the interval between successive births is reduced to

$$J = m + t_1 + 1 \tag{29}$$

the moments can be simplified as

$$\mu_1 = m + 1 + \frac{q}{p} = m + \frac{1}{p} \tag{30}$$

$$\mu_2 = \frac{q}{p} + \frac{q^2}{p^2} = \frac{q}{p^2} \tag{31}$$

$$\mu_3 = \frac{q}{p} + \frac{3q^2}{p^2} + \frac{2q^3}{p^3} = \frac{q + q^2}{p^3} \tag{32}$$

$$\mu_4 = \frac{q}{p} + \frac{10q^2}{p^2} + \frac{18q^3}{p^3} + \frac{9q^4}{p^4} = \frac{q + 7q^2 + q^3}{p_4} \tag{33}$$

as

$$p + q = 1. \tag{34}$$

In this example, however, there are only two unknown parameters, which may be solved from the first two equations, namely, (30) and (31).

Rewriting (31) as

$$\frac{1}{p^2} - \frac{1}{p} - \mu_2 = 0 \tag{35}$$

we get

$$\frac{1}{p} = \frac{1 + \sqrt{1 + 4\mu_2}}{2} \tag{36}$$

as the negative solution is inadmissible. Substitution $1/p$ in (30) produces the solution for m. [The assumption of $r = 0$ may also be validated by substituting the estimate of p from (36) in the expressions for higher order moments in (32) and (33).]

Fertility Control

So far, we have examined the distributions of birth interval that are not subject to the influences of birth control practices. But where fertility is planned, couples may decide to wait a certain number of months, say w, before trying to have another child. In fact, w may even include the abortions that are performed to limit the family size. The interval between two successive births may then be expressed as

$$J(w, t) = m + w + t + 1 \tag{37}$$

where for reasons of simplicity, we have excluded other cases of fetal losses.

That is why, the term kt_2 of equation (1) is absent in (37) and because of that, t_1 has been placed by t. It is evident that like t, w is a random variable and its distribution is independent of t. If we assume that w is normally distributed with a mean n and a variance v, the moments of $J(w, t)$ can be written as

$$\mu_1' = m + n + \frac{1}{p} \tag{38}$$

$$\mu_2 = v + \frac{q}{p^2} \tag{39}$$

$$\mu_3 = \frac{q + q^2}{p^3} \tag{40}$$

following the same principle that was used to derive (30)-(32). Note that the third moment of w around its mean is zero because of the normality assumption resulting in a simple expression for the third moment of $J(w, t)$ in (40). The parameters p and $q = 1 - p$, therefore can be solved from (40) and substitution of their values in (38) and (39) will then generate estimates of v and $m + n$. Unfortunately, independent estimates m and n will not be available for this model. However, $m + n$ can still be interpreted as the average length of the non-susceptible period between successive live births part of which is natural while the remainder is artificially introduced.

The consistency of the model can be checked by computing the fourth moment, namely,

$$\mu_4 = \frac{q + 7q^2 + q^3}{p^4} + 3v^2 + \frac{6vq}{p^2} \tag{41}$$

remembering that the fourth moment of w is equal to $3v^2$ due to the assumption of normality and further that the fourth moment of the sum of two independent variables, say x and y can be expressed as

$$\mu_4(x + y) = \mu_4(x) + \mu_4(y) + 6\mu_2(x)\,\mu_2(y). \tag{42}$$

Applications

The model developed on the assumption of no pregnancy wastage is the simplest of all, and we have attempted to measure its fit on an Indian example. The data were obtained from a sample of women in South India (Srinivasan, 1972) in 1965. Table 1 provides the estimates of the parameters and the goodness of fit.

The expected distributions were obtained by using the estimates of the starting points of the curves provided by m in Table 1.

TABLE 1—MONTHLY PROBABILITY OF CONCEPTION AND OTHER PARAMETERS OF INDIAN WOMEN

Parameters	*Interval between births*		
	first and second	*second and third*	*third and fourth*
Sample size	297	306	245
Average interval (months)	36.1	38.9	38.6
Variance μ_2	345.8	407.8	380.1
Monthly probability of conception ρ	.052	.048	.050
Nonsusceptible period (months) m	17.0	18.2	18.6
χ^2 4 d.f. (goodness of fit)	1.04	9.34	7.00

The fit was highly satisfactory for samples of such small sizes. The differences between the average intervals and for different orders of birth are not statistically significant. We did not go through the trouble of estimating the standard errors of the non-susceptible periods and those of the monthly probabilities of conception as it is quite apparent that these are also independent of the order of birth. For all of these orders of birth, the monthly probability of conception of .05 can be regarded as a reasonable approximation.

We would like to note that the validity of the aforementioned model is based on the assumptions of the (1) absence of fertility control practices as well as of the (2) absence of pregnancy wastage. It may be pointed out that the former assumption is quite appropriate for India in general and for the sample in particular during the period concerned. we would like to argue in favor of the latter in view of the goodness of fit provided by the model.

For reasons of simplicity, we have chosen the model based on fertility control to experiment with an US example. For the year 1975, a distribution of birth intervals based on a fifty percent sample of all births (NCHS, 1978) generated Table 2.

TABLE 2—MONTHLY PROBABILITY OF CONCEPTION AND OTHER PARAMETERS OF US WOMEN

Parameters	*Interval between births*		
	first and second	*second and third*	*third and fourth*
Average interval (months)	40.2	44.4	45.6
Variance μ_2	422.5	557.1	616.1
Third moment μ_3	6777.3	6121.8	5947.0
Monthly probability of conception	.064	.067	.067
Average period of nonconception $m + n$	24.7	29.4	30.8
Variance of $m + n$ or v	197.0	340.4	409.4
Index of dissimilarity Δ	9.7	9.3	9.8

The model seems to be quite satisfactory and as in the Indian example, the probability of conception, although somewhat higher, seems to remain stable for all birth orders. It may be noted that in this model, the non-susceptible period is confounded with a part of the susceptible period during which the couples deliberately avoided conception. Assuming that the former component is of the order of 15 months (around 18 months in the Indian example), the average couple seems to have delayed the time of birth of their second child by about 10 months, their third and fourth child by about 15 months.

Note that in this example we have chosen the index of dissimilarity A as a measure of goodness of fit, rather than X^2 because the sample was too large for the latter. The values of A, obtained by summing the positive differences between the observed and the model distributions, converted into their percentage forms, are not as small as we expected. It is possible that the normality assumption of *w* is somewhat questionable, although the departure from normality does not seem to be large enough to justify other models and sacrifice the simplicity achieved in the derivation of these results.

References

1. Basu, D., 1955, A *note* on this structure of stochastic model considered by V. M. Dandekar, *Sankhya, 15, 251-252.*
2. Bongaarts, John, 1975, A method for the estimation of fecundubilily, *Demography,* 12, 645-660.
3. Chiang, Chin Long, 1971, A stochastic model of human fertility, *Biometrics,* 27, 345-56.
4. Dandekar, V. M., 1955, Certain modified forms of Binomial and Poisson distributions, *Sankhya,* 15, 237-50.
5. Das Gupta, Prithwis, 1973, *A Stochastic Model of Human Reproduction,* Institute of International Studies, University of California, Berkeley.
6. Keyfitz, Nathan, 1977, *Introduction to the Mathematics of Population,* Addison-Wesley Publishing Company.
7. Perrin, E. B. and Sheps, M. C., 1963, Human reproduction : a stochastic process, *Biometrics,* 20, 28-45.
8. Srinivasan, K., 1972, *Analytical Model for Two Types of Birth Intervals with Applications to Indian Population,* St. Joseph's Press, Trivandrum-695014, Kerala, India.