

R. C. Yadava* and Meenakshi Srivastava**

On the Distribution of Births Over Time in an Equilibrium Birth Process for a Female Giving Specified Number of Children in a Given Period

Introduction

DURING the last few decades a large number of probability models have been developed to study the variation in the number of births to a female in a given period of time, under varying sets of assumptions (Brass 1958; Dandekar 1955; PerrinandSheps 1964; Singh 1961, 1964, 1968; Singh and Singh, 1981; SinghandYadaval977, and others). The starting point of the interval in these models is either the marriage or a distant point since marriage, such that the process is in the state of equilibrium. Models derived by Dandekar (1955) (for abrupt sequence of births). Singh and Singh (1981) and Singh and Yadava (1977) come under the second category. One of the main characteristics of such a process is that the births are uniformly distributed over time. It becomes logical to investigate whether or not the births are uniformly distributed over time for a female giving specified number of children, in a given period in an equilibrium birth process.

This distribution may be quite useful in finding out the distribution of survivors among these births which may be of immense use. In this respect, if the theoretical expressions for timings of these births are known, then the corresponding exposure periods for their survival are also known from which probabilities of survival of these births (occurring at different points of time) could be easily obtained knowing the level of mortality. These survival probabilities along with the distribution of births over time may be utilized to find the distribution of survivors among these births.

The purpose of the present paper is to find out the distribution of births over time for a female giving specified number of children in an interval of length T in an equilibrium birth process, and also to examine the nature of such a distribution. In order to find this distribution, we have developed a simple technique in which we have divided the length of observational period T into a number of segments and have obtained the distribution of births in these segments by utilizing simple laws of probability.

Notations and Assumptions

Let $X(T_0, T)$ denote the number of conceptions to a female during the period $(T_0, T_0 + T)$, where the starting point T_0 is a distant point since marriage, such that the process is in

* Professor, Department of Statistics, Faculty of Science, Banaras Hindu University, Varanasi-221 005.

** Lecturer, Department of Statistics, Institute of Social Sciences, Agra University, Agra-282 004

equilibrium at T_0 . Obviously, T represents the length of the interval. The distribution is based on the same assumptions as considered by Dandekar (1955) (for abrupt sequence) and Singh and Yadava (1977) (ignoring the foetal losses).

These models assume constant values of conception rate λ and non-susceptible period (h), which is the sum of the period of gestation and period of post-partum amenorrhoea. This implies that if there is a conception at a point of time, then there is no possibility of any other conception in the next h unit of time.

It is known that when T_0 is a distant point since marriage then under these assumptions the distribution of $X(T_0, T)$ is independent of T_0 . Obviously, $X(T_0, T)$ can take values $0, 1, 2, \dots, n$ where n is the maximum possible number of conceptions in the period $(T_0, T_0 + T)$. Here $n = T/h$, when T is a multiple of h and $n = [T + h/h]$ when T is not a multiple of h ; $[T + h/h]$ stands for the greatest integer not exceeding $T + h/h$.

Our main objective here is to investigate that given $X(T_0, T) = i$ ($i \geq 1$), how these i conceptions are distributed over period $(T_0, T_0 + T)$ of length T .

It seems very difficult to obtain the exact distribution of these i conceptions over the time interval $(T_0, T_0 + T)$. Since there will be a large number of combinations for the timing of occurrence of these i conceptions over time interval T , it becomes very difficult to compute the probabilities of occurrences of all these combinations. However, the problem can be simplified to some extent by dividing the interval $(T_0, T_0 + T)$ into a number of segments and deriving the distribution of these i conceptions in these segments. This helps in tackling the problem to a large extent. The lengths of these segments are so chosen that there can be utmost one conception in each segment. For this purpose, the observational period $(T_0, T_0 + T)$ of length T is divided into segments of length h , such that either all the segments are of length h (if T is a multiple of h) or all, except the last segment, are of length h and the length of the last segment is say $y < h$ (if T is not a multiple of h). Obviously, there can be either one conception or zero conception in each segment. Further, if there is no conception in any segment of length h , then at the starting point of the next segment the female will definitely be exposed to the risk of conception. And, as such, the models incorporating the assumption that the female is exposed to the risk of conception at the beginning of the interval, can be utilized. This is the main advantage of such a division.

In the present case, two models given by Singh (1968) and Singh and Yadava (1977) are used for computing the probability distribution under consideration. In this context, it is worthwhile to mention that the model proposed by Singh (1968) is useful for finding the distribution of number of conceptions in an interval of length T where the female is exposed to the risk of conception at the beginning of the interval, whereas the model proposed by Singh and Yadava (1977) is useful for finding the distribution of number of conceptions in an interval of length T in an equilibrium birth process.

Let P_{x_1, x_2, \dots, x_n} denote the probability that there will be x_1 conception in the first interval of length h , x_2 in the second interval of length h , x_n in the n th (last) interval; obviously, then, x_i ($i = 1, 2, \dots, n$) can take value either zero or one, i.e., either there can be no conception or one conception in any of the interval. Let $P_m(T)$ denote the probability that there will be m conceptions in $(T_0, T_0 + T)$.

Thus $P_m(T) = \sum_{x_1, x_2, \dots, x_n} P_{x_1, x_2, \dots, x_n}$ such that $\sum x_i = m, m = 0, 1, 2, \dots, n$. It is easily seen that there will be in all 2^n probabilities of type P_{x_1, x_2, \dots, x_n} ($x_i = 0$ or $1, i = 1, 2, \dots, n$).

For finding all these 2^n probabilities we proceed as follows. Let us assume that all the probabilities of type P_{x_1, x_2, \dots, x_r} ($r = 1, 2, \dots, n-1$), are known. Thus 2^n probabilities of type P_{x_1, x_2, \dots, x_n} for different values of x_1, x_2, \dots, x_n can be computed considering two cases viz. $x_{n-1} = 0$ and $x_{n-1} = 1$.

Case 1: When $x_{n-1} = 0$

When $x_{n-1} = 0$, then again there can be two cases for the value of x_n :

1(a): x_n takes value zero.

1(b): x_n takes value one.

Obviously, when $x_{n-1} = 0$, then female will be in the susceptible state at the beginning of the n th interval and the conditional probability of a conception or no conception in the last segment for this case can directly be obtained from the model given by Singh (1968). In fact, the conditional probabilities of no conception and one conception in the last interval will be $e^{-\lambda y}$ and $(1 - e^{-\lambda y})$ respectively, where y is the length of the n th interval. Obviously, y will be equal to h if T is a multiple of h , otherwise it will be smaller than h .

Thus, for specified values of x_1, x_2, \dots, x_{n-1} , P_{x_1, x_2, \dots, x_n} for $x_n = 0$ and $x_n = 1$ are given as

$$P_{x_1, x_2, \dots, x_n} = P_{x_1, x_2, \dots, x_{n-1}} \times e^{-\lambda T} \quad (\text{for } x_n = 0) \quad (1)$$

$$P_{x_1, x_2, \dots, x_n} = P_{x_1, x_2, \dots, x_{n-1}} \times (1 - e^{-\lambda T}) \quad (\text{for } x_n = 1) \quad (2)$$

as in general for two events A and B , $P(AB) = P(A) \times P(B/A)$.

Since it is assumed that $P_{x_1, x_2, \dots, x_{n-1}}$ are known for all possible values of x_1, x_2, \dots, x_{n-1} , all the probabilities for the case $x_{n-1} = 0$ are easily known.

CASE 2: When $x_{n-1} = 1$

Here also there will be two cases according to the values of x_n :

2(a): x_n takes value zero

2(b): x_n takes value one.

Let us first consider the situation when $x_n = 1$. For this situation also, there may arise two cases: 2 (b(i)) either all x_i 's are one or 2 (b(ii)) there is atleast one x_i , ($i = 1, 2, \dots, n-2$) which is zero. Thus, when all x_i 's are one, P_{x_1, x_2, \dots, x_n} denotes the probability of exactly n conceptions in the interval $(T_0, T_0 + T)$, which can be easily obtained from the model of Singh and Yadava (1977). However, for the case 2(b(ii)), i.e., when atleast one x_i ($i = 1, 2, \dots, n-2$) is zero, the probability can be obtained in the following manner. Let the r th position be the last position such the $x_r = 0$ and after that no x_i is zero, ($r = 1, 2, \dots, n-2$). Obviously, the female is exposed to the risk of conception at the beginning of the $(r+1)$ th interval and hence the model of Singh (1968) can be utilized to compute the probability of exactly $(n-r)$ conceptions in the interval of length $(T-rh)$. Thus,

$$P_{x_1 x_2, \dots, x_n} = P_{x_1 x_2, \dots, x_r} \times P_{n-r}^* (T-rh)$$

(since $P(AB) = P(A) \times P(B|A)$)

From our assumption, $P_{x_1 x_2, \dots, x_r}$ are known for all possible values of x_1, x_2, \dots, x_r , hence $P_{x_1 x_2, \dots, x_n}$ can easily be computed from (3). For the case when $x_n = 0$, the probability $P_{x_1 x_2, \dots, x_n}$ can be computed as follows. The event corresponding to the specified values of x_1, x_2, \dots, x_{n-1} can be considered as the union of two mutually exclusive events:

1. x_1, x_2, \dots, x_{n-1} take the above specified values and x_n is zero.
2. x_1, x_2, \dots, x_{n-1} take the above specified values and x_n is one.

Thus we have the relation

$$P_{x_1 x_2, \dots, x_{n-1}} = P_{x_1 x_2, \dots, x_{n-1} x_n; x_n=0} + P_{x_1 x_2, \dots, x_{n-1} x_n; x_n=1} \quad (4)$$

We have already assumed that $P_{x_1 x_2, \dots, x_{n-1}}$ is known and $P_{x_1 x_2, \dots, x_n}$ when $x_n = 1$ can be computed by the method explained above. Hence $P_{x_1 x_2, \dots, x_{n-1} x_n; x_n=0}$ is known by the help of equation (4).

Therefore, for all the 2^n possible permutations of x_1, x_2, \dots, x_n , the probabilities $P_{x_1 x_2, \dots, x_n}$ can be computed easily provided that all the probabilities $P_{x_1 x_2, \dots, x_r}$ ($r = 1, 2, \dots, n-1$) are known. This is equivalent to saying that for 2^n probabilities of the type $P_{x_1 x_2, \dots, x_n}$ ($x_i = 0$ or 1), all the probabilities of this type for $T = h, T = 2h, T = 3h, \dots, T = (n-1)h$ should be known. The above argument is true for any specified value of n .

Thus, for example, if the value of T is such that $n = 4$, then in all there will be 2^4 possible probability combinations of the type $P_{x_1 x_2 x_3 x_4}$ ($x_i = 0$ or $1, i = 1, 2, 3, 4$) such as $P_{0000}, P_{1000}, P_{0100}, P_{0010}, P_{0001}, P_{1100}, P_{1010}, P_{1001}, P_{0101}, P_{0110}, P_{0011}, P_{1110}, P_{1011}, P_{1101}, P_{0111}$, and P_{1111} . And for the computation of all these probabilities, it is assumed that,

- (i) 2^1 probabilities P_0 and P_1 when $T = h$;
- (ii) 2^2 probabilities P_{00}, P_{10}, P_{01} , and P_{11} when $T = 2h$;
- (iii) 2^3 probabilities $P_{000}, P_{100}, P_{010}, P_{001}, P_{110}, P_{101}, P_{011}$ and P_{111} when $T = 3h$.

are known.

Now, P_0 and P_1 denote the probability of zero and one conception in the first segment which can be easily obtained from the model of Singh and Yadava (1977) considering $T = h$. Similarly, P_{00} and P_{11} denote the probabilities of zero and two conceptions in an equilibrium birth process with $T = 2h$, and they can easily be obtained from Singh and Yadava (1977).

Now computation of P_{10} comes under the case 2(b) of the suggested procedure and for this we have the relation

$$P_1 = P_{10} + P_{11} \quad (5)$$

where P_1 denotes the probability of one conception in the first interval of length h in an equilibrium birth process and this can also be easily obtained from Singh and Yadava (1977). It has already been mentioned that P_{11} can be obtained from the above model: hence P_{10} is known.

Now computation of P_{01} comes under the case 1 (b). Therefore,

$$P_{01} = P_0 \times P_1^*$$

where P_1^* is the conditional probability of one conception in an interval of length h given that female is exposed to the risk of conception at the beginning of the interval which can be obtained from the model of Singh (1968). Of course, P_0 is the probability of zero conception in an equilibrium birth process in an interval of length h which can be computed from the model of Singh and Yadava (1977). Thus from (6), P_{01} is known.

Thus all the probabilities for $n=2$, viz., P_{00} , P_{10} , P_{01} , and P_{11} , are known, and hence, by the method explained earlier all the 2^3 probabilities for $n=3$ can be computed easily. Thus by the method of induction all the 2^n probabilities for general n can be computed. It must be pointed out here that the calculation of actual probabilities can be done only when the parameters λ and h are known.

It is already assumed that there is one-to-one correspondence between a conception and a birth. The conceptions in the interval $(T_0, T_0 + T)$ correspond to births in the interval $(T_0 + g, T_0 + T + g)$ where g is the gestation period for a live birth. In the equilibrium birth process, the probabilities $P_{x_1 x_2 \dots x_n}$ do not depend on T_0 and hence these are corresponding probabilities for births also in an interval of length T in the equilibrium birth process. Thus, without loss of generality, we can take the distribution of $X(T_0, T)$ as the distribution of number of births also in an interval of length T given that the starting point of the interval is a distant point since marriage.

Patterns of the Distribution of Births over Time

In the following paragraphs an investigation is made to study the pattern of the distribution of births over time. The pattern of the distribution of births over time has been investigated for the case when T is a multiple of h . In this situation, the following characteristics have been observed, being based on probability expressions as well as certain numerical computations.

For illustration the distribution of births has been derived for $T=4h$. The values of h and λ are taken as 1.5 years and 0.5 respectively. All the 2^4 probabilities for the case $n=4$ with $T=4h$ are presented in Table 1.

A review of the values of the probabilities of the type $P_{x_1 x_2 \dots x_n}$ reveals that:

1. The probabilities are symmetric in $x_1 x_2 \dots x_n$, i.e., for example,

$$P_{001} = P_{100}, P_{1101} = P_{1011} \dots$$

2. Given that r births have occurred in the interval $(T_0, T_0 + T)$, these r births are not uniformly distributed over time interval $(T_0, T_0 + T)$. It may be recalled here that the births are uniformly distributed over time in the equilibrium birth process. For example, in this case the probability of three births in the interval is 0.1506. Thus, the number of births of this type becomes $3 \times 0.1506 = 0.4518$ (per female). These births are distributed in the four segments as 0.1275, 0.0984, 0.0984, and 0.1275 respectively. This shows that these births are not uniformly distributed over time.

TABLE 1 : COMPUTED VALUES OF PROBABILITIES OF THE TYPE $P_{x_1 x_2 \dots x_n}$ FOR NUMERICAL ILLUSTRATION FOR SPECIFIED VALUES OF $h = 1.5$ YEARS, $\lambda = 0.5$ AND $T = 4h$

<i>Probabilities</i>	
<i>Type</i>	<i>Value</i>
P_{0000}	0.0602
P_{1000}	0.0673
P_{0001}	0.0673
P_{0100}	0.0956
P_{0010}	0.0956
P_{1100}	0.0468
P_{0011}	0.0468
P_{1010}	0.1069
P_{0101}	0.1069
P_{1001}	0.0751
P_{0110}	0.0759
P_{1110}	0.0231
P_{0111}	0.0231
P_{1101}	0.0522
P_{1011}	0.0522
P_{1111}	0.0050

References

- Brass, W., 1958, The distribution of births in human population. *Population Studies*, **12**: 51-72.
- Dandekar, V. M., 1955, Certain modified forms of binomial and Poisson distribution. *Sankhya*, **15**: 237-250.
- Perrin, E. B. and Sheps, M. C., 1964, Human reproduction — A stochastic process. *Biometrics*, **20**: 28-45.
- Singh, S. N., 1961, A hypothetical chance mechanism of variation in number of births per couple. *Ph. D. thesis*, University of California, Berkeley, USA.
- _____, 1964, A probability model for couple fertility. *Sankhya, ser. B*, **26**: 89-94.
- _____, 1968, A chance mechanism of the variation in the number of births per couple, *J. Amer. Statist. Assoc.*, **63**: 209-213.
- _____, and Singh, V.K., 1981, A probability model for the number of births in an equilibrium birth process. *J. Biosci.*, **3**: 197-205.
- _____, and Yadava, R. C., 1977, A generalized probability model for an equilibrium birth process. *Demography India*, **6**: 163-173.