Jour. Ind. Soc. Ag. Statistics
Vol. XXXVII, No. 2 (1985), pp. 145-153

A PROBABILITY MODEL FOR NUMBER OF CONCEPTIONS IN EQUILIBRIUM BIRTH PROCESS

S. N. SINGH, V. K. SINGH and R. N. SINGH
Department of Statistics, Banaras Hindu University, Varanasi 221005

(Received: December, 1981)

SUMMARY

In an equilibrium birth process births are counted from an abrupt point T_0. In this paper, a model for number of conceptions occurred during the period $(T_0, T_0 + T)$ has been derived assuming that a conception results in either of the k types of pregnancy outcomes. The mean and variance of the distribution have been obtained, which are further used to estimate the fecundability and abortion rate applying the distribution to two sets of observed data taken from Singh and Yadava [4]. The suggested model may be considered as an extension of that of Singh and Yadava [4].

Keywords: Equilibrium birth process; constant risk conception; equilibrium renewal process.

1. Introduction

An equilibrium birth process can be regarded as an ordinary birth process in which the system has been running for a long time before it is first observed. For instance, if a sequence of conceptions or births occurring to a female is observed after a long time since her marriage, the above birth process is said to be in equilibrium (Sheps *et al.* [3]).

While dealing with the modification of ordinary binomial and Poisson distributions for applying them to the observed distributions of number of births, Dandekar [2] obtained a probability distribution for number of children born to a specified age cohort of women when the start of the observational period was at a considerable distance since marriage. The model was derived under the assumption of a constant risk of conception, a constant period of non-susceptibility associated with each delivery and one to one correspondence between conception and live birth. Recently, Singh and Yadava [4] generalised the above model for two types of conceptions called 'complete' and 'incomplete'.

Bearing in mind the fact that the period of non-susceptibility is a key factor in birth spacing and varies considerably depending upon the mode of termination of a pregnancy, there appears to be consensus on the need of developing a probability model which accounts for the different outcomes of a conception. In the present paper the model given by Singh and Yadava [4], has been extended to k types of pregnancy outcomes. For the sake of simplicity, the risk parameter is assumed to be constant throughout the period of investigation. Following the results of renewal theory, expressions for the mean and variance of the concerned distribution have been obtained. The application of the model is illustrated on the data presented in Singh and Yadava [4], assuming that a conception leads to an abortion or in a live birth which is further classified according as the child dies in infancy or survives more than a year, that is, assuming $k = 3$. Consequently, the rate of conception as well as the rate of abortion have been estimated while other parameters are assumed to be known.

2. The Model

Let the period of observation is $(T_0, T_0 + T)$ of length T where T_0 is measured from marriage. Further, let us denote the number of conceptions occurred to a female during the period of investigation by $X(T_0, T)$. The probability model is derived under the assumptions:

A1 : The female has led a married life during the observational period.

A2 : θ_i $(i = 1, 2, \ldots, k)$ is the probability that a conception leads to the ith type of pregnancy outcome $\left(\sum_{i=1}^{k} \theta_i = 1\right)$.

A3 : The waiting time of first conception from marriage follows an exponential distribution with probability density function (p.d.f.)

$$f_0(t) = \lambda e^{-\lambda t}, \quad t > 0; \lambda > 0, \tag{2.1}$$

whereas the time between rth and $(r + 1)$th conception $(r = 1, 2, \ldots)$ follows displaced exponential distribution,

$$f_r(t)\ \lambda e^{-\lambda(t-h_i)},\ t > h_i, \tag{2.2}$$

if the rth conception ends in the ith type of pregnancy outcome. The assumption A3 states that if there is a conception at a point of time then there is no other conception in h_i units of time, if the conception is ith type. We assume that $h_{i-1} < h_i$ for $i = 1, 2, \ldots, k$ and $h_0 = 0$.

A4 : T_0 is at a considerable distance marriage.

As pointed out by Singh and Yadava [4] the total number of conceptions cannot be more than n where $n = T/h_1$ or $[(T + h_1)/h_1]$, according as T is a multiple of h_1 or not. $[x]$ stands for the greatest integer not exceeding x. The probability function of $X(T_0, T)$ is given by

$$P_0 = 1 - H_0(T) + H_1(T) \tag{2.3}$$

$$P_r = H_{r-1}(T) - 2H_r(T) + H_{r+1}(T),\ r = 1, 2, \ldots, n, \tag{2.4}$$

where

$$H_r(T) = \frac{\lambda}{1 + \lambda\bar{h}} \Bigg[\sum_{a=0}^{t_r} \sum_{b=0}^{t_{r-a}} \cdots \sum_{u=0}^{t_{r-a-b-\cdots-j}} \sum_{v=0}^{t_{r-a-b-\cdots-j-u}} \binom{r}{a}\binom{r-a}{b}\binom{r-a-b-\cdots-j}{u}\binom{r-a-b-\cdots-j-u}{v} \theta_k^a \theta_{k-1}^b \cdots \theta_3^u \theta_2^v \theta_1^{r-a-b-\cdots-u-v} \int_{m'}^{T} \Bigg\{ 1 - e^{-\lambda(t-m')} \sum_{s=0}^{r-1} \frac{\lambda^s (t-m')^s}{s!} \Bigg\} dt \Bigg]. \tag{2.5}$$

In the above expression

$$t_r = \min\left[r, \frac{T - rh_1}{h_k - h_1} \right]$$

and

$$m' = ah_k + bh_{k-1} + \ldots + (r - a - b \ldots - v)\, h_1.$$

It is easy to see that

$$H_0(T) = \frac{\lambda T}{1 + \lambda\bar{h}}, \quad \text{where } \bar{h} = \sum_{i=1}^{k} \theta_i h_i.$$

The derivation of the model is given in the following paragraphs:

Let the conceptions be counted from the point T_0 and let the time of rth conception $(r = 1, 2, \ldots)$ be $Z_r = T_1 + T_2 + \ldots + T_r$ where T_1 is the time of first conception measured from T_0 and $T_r (r > 1)$ is the time between $(r - 1)$th and rth conception. If T_0 is at a considerable distance since marriage, the above birth process becomes an equilibrium renewal process where the intervals $T_2, T_3, \ldots$, are independently and identically distributed but T_1 has a different distribution (Cox and Miller, 1965).

The probability density function (p.d.f.) of $T_r (r > 1)$ is given by

$$
\begin{aligned}
f(t) &= 0, && 0 < t < h_1 \\
&= \sum_{i=1}^{j-1} \theta_i \lambda e^{-\lambda(t-h_i)}, && h_{j-1} < t < h_j, \\
& && (j = 2, 3, \ldots, k) \\
&= \sum_{i=1}^{k} \theta_i \lambda e^{-\lambda(t-h_i)}, && t > h_k,
\end{aligned}
\tag{2.6}
$$

and that of T_1 is

$$
f_1(t) = \frac{1 - F(t)}{\mu}, \tag{2.7}
$$

where $F(t)$ and μ are respectively the distribution function and mean associated with $f(t)$.

Let $\phi(s)$ be the Laplace transform of $f(t)$, then the same for Z_r will be

$$
\frac{\{1 - \phi(s)\} \{\phi(s)\}^{r-1}}{\mu s} = \frac{\{\phi(s)\}^{r-1}}{\mu s} - \frac{\{\phi(s)\}^{r}}{\mu s}, \tag{2.8}
$$

for $T_1, T_2, \ldots$ are mutually independent random variables. Inverting (2.8) the p.d.f. $f_r^*(t)$ of Z_r can be obtained as

$$
f_r^*(t) = \frac{1}{\mu} F^{(r-1)}(t) - \frac{1}{\mu} F^{(r)}(t), \tag{2.9}
$$

where $F^{(r)}(t)$ is the distribution function corresponding to r-fold convolution of $f(t)$ with itself.

The expression for $F^{(r)}(t)$ is given by

$$
\begin{aligned}
F^{(r)}(t) = \sum_{a=0}^{t_r} \sum_{b=0}^{t_{r-a}} \sum_{u=0}^{t_{r-a-b-\cdots-j}} \sum_{v=0}^{t_{r-a-b-\cdots-j-u}} & \binom{r}{a} \binom{r-a}{b} \cdots \binom{r-a-\cdots-j}{u} \\
& \binom{r-a-b-\cdots-j-u}{v} \theta_k^a \theta_{k-1}^b \ldots \theta_3^u \theta_2 \theta_1^{r-a-b-\cdots-u-v} \\
& \left[1 - e^{-\lambda(t-m')} \sum_{s=0}^{r-1} \frac{\lambda^s (t-m')^s}{s!} \right].
\end{aligned}
\tag{2.10}
$$

Now, following Cox and Miller [1] and letting

$$
H_r(T) = \int_0^T \frac{F^{(r)}(t)}{\mu} \, dt,
$$

we have

$$P_0 = 1 - H_0(T) + H_1(T) \tag{2.11}$$

and $P_r = H_{r-1}(T) - 2H_r(T) + H_{r+1}(T), \quad r = 1, 2, \ldots, n.$ (2.12)

3. Mean and Variance of the Distribution

Since $X(T_0, T)$ assumes values $0, 1, \ldots, n$ with probabilities $P_0, P_1, \ldots, P_n$ respectively, we have

$$E[X(T_0, T)] = \sum_{r=0}^{n} rP_r = H_0(T) = \frac{\lambda T}{1 + \lambda \bar{h}}\cdot \tag{3.1}$$

Similarly, the variance of the distribution is given by

$$V[X(T_0, T)] = \sum_{r=0}^{n} r^2 P_r - E[X(T_0, T)]^2. \tag{3.2}$$

Due to the complex nature of the model, a simple expression for $V[X(T_0, T)]$ is difficult to obtain. However, based on renewal theory, Cox and Miller [1] have presented an approximate expression for the variance of an equilibrium birth process, for large T, as

$$V[X(T_0, T)] = \frac{\sigma^2 T}{\mu^3} + \left(\frac{1}{6} + \frac{\sigma^4}{2\mu^4} - \frac{\mu_3}{3\mu^3}\right) + 0\,(1), \tag{3.3}$$

where σ^2 and μ_3 have their usual meaning.

Using the expression of the p.d.f. of T_r, we have

$$\mu = E(T_r) = \int tf(t)\,dt$$

$$= \sum_{i=1}^{k} \theta_i \left(\frac{1}{\lambda} + h_i\right) = \frac{1}{\lambda} + \bar{h}, \tag{3.4}$$

$$\mu_2' = \sum_{i=1}^{k} \left\{ \int_{h_i}^{\infty} t^2\, \theta_i \lambda e^{-\lambda(t-h_i)}\, dt \right\}$$

$$= \frac{2}{\lambda^2} + \frac{2}{\lambda}\, \bar{h} + \sum_{i=1}^{k} \theta_i\, h^2, \tag{3.5}$$

and

$$\mu_3' = \sum_{i=1}^{k} \left\{ \int_{h_i}^{\infty} t^3\, \theta_i \lambda e^{-\lambda(t-h_i)}\, dt \right\}$$

$$= \frac{6}{\lambda^3} + \frac{6}{\lambda^2}\, \bar{h} + \frac{3}{\lambda} \Sigma\, \theta_i h_i^2 + \Sigma\, \theta_i h_i^3, \tag{3.6}$$

so that

$$\sigma^2 = \mu_2' - \mu^2 \tag{3.7}$$

and

$$\mu_3 = \mu_3' - 3\mu_2'\mu + 2\mu^3. \tag{3.8}$$

4. The Model for $k = 3$

It is apparent from the expression (2.5) that if only two distinct outcomes of a conception are assumed (that is, $k = 2$), the model reduces to that of Singh and Yadava [4]. However, the case of $k = 3$ may be interesting in some situations, for which the expression of $H_r(T)$ is given as follows:

$$H_r(T) = \frac{\lambda T}{1+\lambda \bar{h}} - r + \frac{1}{1+\lambda \bar{h}} \sum_{j=0}^{r} \sum_{i=0}^{r-j} \binom{r}{j} \binom{r-j}{i} \theta_1^{r-j-i}\, \theta_2^i \theta_3^j$$

$$e^{-\lambda(T - jh_3 - ih_2 - \overline{r-j-i}h_1)} \sum_{s=0}^{r-1} \sum_{k=0}^{s} \frac{\lambda^k (T - jh_3 - ih_2 - \overline{r-j-i}h_1)^k}{k!}$$

if $t_r = r$ (4.1)

$$= \sum_{j=0}^{t^r} \sum_{i=0}^{t_{r-j}} \binom{r}{j} \binom{r-j}{i} \theta_3^j \theta_2^i \theta^{r-j-i}$$

$$\left[\frac{\lambda (T - jh_3 - ih_2 - \overline{r-j-i}h_1)}{1+\lambda\bar{h}} - \frac{1}{1+\lambda\bar{h}} + \frac{1}{1+\lambda\bar{h}} \right.$$

$$e^{-\lambda(T - jh_3 - ih_2 - \overline{r-j-i}h_1)} \sum_{s=0}^{r-1} \sum_{k=0}^{s} \frac{\lambda^k (T - jh_3 - ih_2 - \overline{r-ji-h_1})^k}{k!}$$

if $t_r < r$. (4.2)

Accordingly expressions for P_0 and P_r ($r = 1, 2, \ldots n$) can be obtained.

5. Application

Although the model takes into account k types of pregnancy outcomes, it is generally impossible to get accurate information on all the variables involved and to obtain suitable estimation procedures with the existing statistical techniques. In this context, it is to be noted here that in the absence of reliable data on number of conceptions, Singh and Yadava [4] successfully applied the model to the data on number of live birins assuming $k = 2$. Since the period of non-susceptibility associated with an infant death is relatively shorter than that associated with a child surviving more than a year, they considered these two types of live births separately for applying the suggested model. However, if in addition of these two categories of live births, one also considers abortions, the application of the model may be seen for $k = 3$.

For the application purpose, the same two sets of observed distributions, as presented in Singh and Yadava [4] have been considered. These distributions are taken from 'A Demographic Survey of Varanasi (Rural)' and relate to the number of births in the last 5 and 7 years to eligible females of age group 25-29 having at least 5 and 7 years of marriage durations. For detailed account of the data and survey methodology see Singh and Yadava [4] and Singh *et al.* [5].

In the surveyed area, the average age at marriage for females is nearly 15 years. It is to be noted here that the fertility performance of females aged 25-29 years in the last 5 and 7 years respectivey have been considered. Thus, in the first case, the observed data relates to the fertility performance of the females when they were in the age range (22.5-27.5) years. Similarly, in the second case, the data relates, on the average, to the fertility performance of females in the age range (20.5-27.5) years. The start of the observational period in both the sets, therefore, may practically be considered at a considerable distance from marriage.

While applying the model for $k = 3$, three types of pregnancy outcomes have been considered in respect of associated rest periods, namely, (i) abortions with associated rest period h_1, (ii) those live births ending in infant deaths with non-susceptible period h_2 and (iii) live births surviving more than a year with attached rest period h_3. Let θ_1, θ_2 and θ_3 are the respective probabilities of (i), (ii) and (iii). Thus, in all the model has 7 parameters, viz., θ_1, θ_2, θ_3, h_1, h_2, h_3 and λ for a given T. Since explicit expressions for higher moments of the model are difficult to obtain, it has been tried to build up an estimation procedure for estimating two of the parameters θ_1 and λ with the help of expressions for mean and variance of the distribution assuming others to be known. It should be emphasized here that, generally, the rate of conception as well as the rate of abortion are not directly measurable, hence they should be estimated with the help of certain statistics based on empirical data.

For the surveyed area, Singh and Bhaduri have shown that $h_2 = 1.00$ year and $h_3 = 1.75$ years. Here it is assumed $h_1 = 0.5$ year. Further, the infant mortality rate was reported 160 per thousand. Using these values and equations (3.1) and (3.3) with $k = 3$, the estimates of θ_1 and λ are obtained as follows :

Firstly, some suitable value of θ_1 was chosen. With this value of θ_1, the values of θ_2 and θ_3 were obtained using the equations

$$\theta_2 = 0.16 \{1 - \theta_1\}, \tag{5.1}$$

and

$$\theta_1 + \theta_2 + \theta_3 = 1 \tag{5.2}$$

Now these values were used in (3.1) to obtain an estimate of λ and lastly these estimated values were used to get the expected variance. This iterative procedure was used until the equations (3.1) and (3.3) were satisfied.

Tables 1 and 2 present respectively the distributions for $T = 5$ and 7 years along with the estimates of θ_1 and λ. The expected frequencies

TABLE 1—DISTRIBUTIONS OF ELIGIBLE COUPLES AGED 25-29 YEARS ACCORDING TO THE NUMBER OF BIRTHS DURING LAST 5 YEARS

Number of births	*Observed number of women*	*Expected number of women ($k = 3$) $\hat{\lambda} = 0.65$ $\hat{\theta}_1 = 0.15$*	*Expected number of women ($k = 2$) $\hat{\lambda} = 0.605$*
(1)	(2)	*(3)*	*(4)*
0	20	22.8	25.8
1	155	150.6	150.4
2	200	199.7	189.2
3	45	49.2	54.9
4 and above	5	2.7	4.7
Total	425	425.0	425.0
χ^2		0.542	3.865
d. f.		1	3

TABLE 2—DISTRIBUTIONS OF ELIGIBLE FEMALES AGED 25-29 YEARS ACCORDING TO THE NUMBER OF BIRTHS DURING THE LAST 7 YEARS

Number of births	*Observed number of women*	*Expected number of women ($k = 3$) $\hat{\lambda} = 0.705$, $\hat{\theta}_1 = 0.09$*	*Expected number of women ($k = 2$) $\hat{\lambda} = 0.620$*
(1)	(2)	*(3)*	*(4)*
0	3	4.3	6.9
1	58	52.8	61.7
2	166	171.8	162.7
3	151	151.0	135.8
4 and above	34	28.1	44.9
Total	412	412.0	412.0
χ^2		1.804	6.840
d. f.		1	3

obtained from the suggested model are given in Column 3. The estimate of λ and expected frequencies obtained in Singh and Yadava [4] have been given under Column 4. In both the cases, the calculated values of χ^2 are insignificant at 5 per cent level of significance, showing that there is an agreement between the observed and expected frequencies.

6. Conclusions

A close look on both the tables shows that as far as the fitting of models is concerned, both the models (for $k = 2$ and 3) fit the data well. However, the most striking feature of the suggested model is that it provides an estimate of rate of foetal wastage using the data on number of births, which is otherwise difficult to ascertain inspite of sincere efforts of the investigator.

The estimates of the conception rate, λ, for both the sets of data are comparatively higher than those of Singh and Yadava [4] because of the inclusion of the chance of foetal wastages in the model. It is to be noticed that the estimate of λ for the the first set of data corresponds to the age 24 years, while the same for second set corresponds to the age 25 years of the females.

The abortion rates for the two sets of data have been estimated as 15 and 9 per cent respectively. The low estimate in the latter case may be due to under-reporting of abortions as they were the events of remote past.

The major limitation of the proposed model is its dependency on a large number of parameters. However, while applying it to an observed data, value of some of the parameters can be guessed through empirical investigations of similar kind.

ACKNOWLEDGEMENT

The authors are greatful to the referee for his valuable suggestions.

REFERENCES

[1] Cox, D. R. and Miller, H. D. (1965) : *The Theory of Stochastic Processes*, Methuen and Co., London.

[2] Dandekar, V. M. (1955) : Certain modified forms of binomial and Poisson distributions, *Sankhya* **15A** : 237-250.

[3] Sheps, M. C., Menken, J. A. and Radick, A. P. (1969) : Probability models for family building : An analytical review, *Demography* **6** : 161-183.

[4] Singh, S. N., Yadava, R. C. (1977) : A generalised probability model for an equilibrium birth process, *Demography India*, **6** : 163-173.

[5] Singh, S. N., Yadava, R. C. and Bhaduri, T. (1970) : A demographic Survey of Varanasi (Rural), *Prajna* **16** (1) : 176-189.