

A Probability Distribution for First Birth Interval

O. P. Singh

Department of Statistics
Udai Pratap College
Varanasi - 221002, India

V. K. Singh

Department of Statistics
Banaras Hindu University
Varanasi - 221005, India

D. S. Pathak

Department of Statistics
Faculty of Science and Technology
M. G. Kashi Vidyapeeth
Varanasi - 221002, India

Abstract

The paper attempts to develop a probability model for first birth interval incorporating incidence of foetal wastages prior to live birth and the phenomenon of physical separation which are still prevalent in many developing societies. The fact that fecundability varies considerably over a random group of females is also taken into account. Estimates of certain parameters of the model have been obtained by utilizing a real set of data on the time of first complete conception.

Key Words: First birth interval, stochastic model, fecundability, physical separation.

Résumé

Cet article tente de développer un modèle de probabilité pour l'intervalle de première naissance en incorporant les incidences de pertes fœtales volontaires avant une naissance vivante ainsi que le phénomène de séparation physique; conditions qui prévalent encore dans beaucoup de sociétés en voie de développement. Le fait que le niveau de fécondabilité varie considérablement parmi un groupe aléatoire de femmes est aussi pris en compte. Les estimations de certains paramètres du modèle ont été obtenus en utilisant un ensemble de données réel sur le moment de la première conception portée à terme.

Mots-clés: l'intervalle de première naissance, le modèle stochastique, la fécondabilité, la séparation physique

Introduction

For obvious reasons, in the study of human reproduction, measurement of fecundability has got an important place. Among the various types of fertility data which have been used for the purpose, data on first birth interval have an upper hand over all other types of birth interval due to certain reasons. First, being the earliest and first event of the married life of a female, it hardly suffers from recall lapse; second, it is free from the period of post-partum amenorrhea (p.p.a.) associated with a live birth. Other birth intervals are heavily affected by the erratic fluctuations of this period.

Attempts to analyze first birth interval for the estimation of natural conception rate (fecundability) dates back to Gini (1924). Since then several attempts have been made to investigate this interval and a number of demographers have formulated several stochastic models to describe this interval under various sets of assumptions for the related situations. For instance, frequent occurrence of foetal wastages prior to first live birth considerably increase the length of first birth interval consequently, Srinivasan (1966) and Chakraborty (1976) developed discrete and continuous time models respectively for first birth interval in presence of foetal wastages.

Apart from biological components of fertility, the first birth interval is often prolonged due to some social customs prevalent in certain societies. For instance, in rural parts of India, married females stay in their parents' home soon after marriage for a considerable period of time which ultimately delays the first birth. This period of physical separation of Indian females may be compared

with the situation of some western countries where females delay their first birth voluntarily by adopting cent percent effective contraceptives just after marriage. Singh and Singh (1983), Singh et al. (1985 b), Singh (1987) and others have propounded stochastic models for estimating fecundability through first birth interval in presence of physical separation in the Indian context.

This paper aims at devising a continuous time stochastic model for first birth interval under the assumption that (i) foetal wastages occur prior to first live birth (ii) the interval consists of a period of physical separation and (iii) the rate of conception is constant for a woman during her entire period of reproduction but varies over a random group of women. With a view to illustrate the use of proposed model for analyzing first birth interval and thereby to get reliable estimates of some parameters of the model, particularly of fecundability, The model has been applied to one set of empirical data taken from Singh et al. (1985b), assuming values of other parameters to be known from other results obtained over the same area.

Underlying Assumption and the Model

Let T denotes the time of first complete conception, that is, the waiting time of the first conception since the marriage which results into a live birth. Obviously, in the presence of physical separation and foetal wastages prior to first complete conception, T comprises of the following random components:

- (i) The period of physical separation just after marriage,
- (ii) Waiting time for conceptions occurring before first complete conception, and
- (iii) Periods of non-susceptibility (gestation plus p.p.a.) associated with conceptions resulting in foetal wastages prior to first complete conception.

Thus, given that n incomplete conception (foetal wastages) occur prior to first complete conception, we have

$$T = Z + nh + \sum_{k=1}^{n+1} X_k \quad (1)$$

where Z is the period of physical separation; h is the period of non-susceptibility associated with an incomplete conception and $X_1, X_2, \dots, X_{n+1}$ are waiting times for $(n+1)$ conceptions. As regards components of T , we make the following assumptions:

- (i) Z is random variable having p values $z_1, z_2, \dots, z_p$ ($z_{i-1} < z_i$; $i = 2, 3, \dots, p$) with probabilities α_i ; $0 < \alpha_i < 1$, $\sum_{i=1}^p \alpha_i = 1$ respectively.

- (ii) The waiting time from marriage to first conception, X_1 , follows an exponential distribution, with parameter m, as

$$g_1(x) = me^{-mx}; \quad x > 0, \quad m > 0 \quad (2)$$

Where, X_r , the time elapsed between the $(r-1)^{th}$ incomplete conception to next conception ($r = 2, 3, \dots, n+1$) follows a displaced exponential distribution given by–

$$g_r(x) = me^{-m(x-h)}; \quad x > h, \quad m > 0 \quad (3)$$

In (2) and (3) above, m is the rate of occurring conception (conception rate) that is a measure of fecundability.

- (iii) m varies over a randomly selected group of women with type III distribution given by

$$g(m) = \frac{b^a}{a!} e^{-bm} \frac{m^{a-1}}{a!}; \quad m > 0, \quad a, b > 0 \quad (4)$$

- (iv) the number of incomplete conceptions (N) prior to the first complete conception follows a geometric distribution with probability mass function

$$P[N = n] = \theta(1-\theta)^n; \quad n = 0, 1, 2, \dots; \quad 0 < \theta < 1 \quad (5)$$

where θ stands for the probability of a conception resulting to complete conception.

- (v) Waiting times for conceptions are mutually independent.

Remark 1

It is remarkable here that while assumptions (i) and (iv) are based on empirical evidences, assumption (ii) and (iii) are of quite theoretical nature and need support from empirical data. Needless to say that validity of distribution $g_1(x)$, $g_r(x)$ and $g(m)$ has been tested in a number of previous works of similar kind, such as Singh (1964), Singh and Singh (1983), Singh et al. (1976), Singh et al. (1985b) and others.

Theorem 1

The distribution function, $F(t)$ of T under the assumptions (i) to (v) is

$$F(t) = \begin{cases} 0 & t < z_1 \\ \sum_{s=1}^{i-1} \alpha_s \sum_{r=0}^{Q_s} \theta(1-\theta)^r [1 - L_{s(r,t)}(a, r+1)]; & z_{i-1} \leq t < z_i; \text{ for } i = 2, 3, \dots, p \\ \sum_{s=1}^i \alpha_s \sum_{r=0}^{Q_s} \theta(1-\theta)^r [1 - L_{s(r,t)}(a, r+1)]; & t \geq z_p \end{cases} \quad (6)$$

where $L_i(r, t) = b / (b + t - rh - z_i)$; $i = 1, 2, \dots, p$

$$I_y(p, q) = \int_0^y \frac{1}{B(p, q)} x^{p-1} (1-x)^{q-1} dx; \quad B(p, q) = \frac{\Gamma(p) \Gamma(q)}{\Gamma(p+q)}$$

and $Q_s = \left\lfloor \frac{t - z_s}{h} \right\rfloor$; $[u]$ stands for the greatest integer not exceeding u .

The derivation of $F(t)$ is presented in Appendix A.

Moments of the Distribution

Let us denote the Laplace transform corresponding to conditional probability density function (p.d.f.) of T, conditional on m, by $\Phi(s/m)$. Then, we have

$$\Phi(s/m) = \int_0^\infty e^{-st} f(t/m) dt \quad (7)$$

where, $f(t/m)$ is given by (A3) in appendix A.

It is easy to see that

$$\Phi(s/m) = \sum_{i=1}^p \alpha_i \theta m \frac{e^{-sz_i}}{m + s - (1-\theta)me^{-sh}} \quad (8)$$

Hence, the first two moments about origin of conditional p. d. f. of t, $f(t/m)$ are obtained as

$$E(T/m) = \bar{z} + \frac{\left(\frac{1}{m}\right) + h(1-\theta)}{\theta} \quad (9)$$

$$E(T^2/m) = \frac{\bar{z}^2 + 2\bar{z} \left(\frac{1}{m}\right) + h(1-\theta)}{\theta} + \frac{\left(\frac{2}{m^2}\right) + \left(\frac{4h}{m}\right)(1-\theta) + h^2(1-\theta)^2 + h^2(1-\theta)}{\theta^2} \quad (10)$$

$$\text{where } \bar{z} = \sum_{i=1}^p \alpha_i z_i ; \quad \bar{z}^2 = \sum_{i=1}^p \alpha_i z_i^2$$

A Probability Distribution for First Birth Interval

Now considering the variation in fecundability among women according to the assumption (iii), the unconditional moments of T about origin are obtained as

$$E(T) = \int_0^{\infty} E(T/M) g(m) dm = \bar{z} + \frac{\frac{b}{a-1} + h(1-\theta)}{\theta} \quad (11)$$

$$\begin{aligned} E(T^2) &= \int_0^{\infty} E(T^2/m) g(m) dm \\ &= \bar{z}^2 + \frac{2\bar{z} \left\{ \frac{b}{a-1} + h(1-\theta) \right\}}{\theta} + \frac{\frac{2b^2}{(a-1)(a-2)} + \frac{4bh(1-\theta)}{a-1} + (1-\theta)^2 h^2 + (1-\theta)h^2}{\theta^2} \end{aligned} \quad (12)$$

where variance of T is given by

$$\begin{aligned} V(T) &= E(T^2) - \{E(T)\}^2 \\ &= \bar{z}^2 - \bar{z}^2 + \frac{\frac{b^2}{(a-1)^2(a-2)} + \frac{2b(1-\theta)h}{(a-1)} + (1-\theta)h^2}{\theta^2} \end{aligned} \quad (13)$$

Application

The suggested model can be applied to a variety of situation due to its flexibility with a view to estimate the rate of conception and/or other parameters of the model. It is remarkable here that while applying a model to a practical situation, the estimation of all or a large number of parameters is not advisable, as, it involves generally higher order moments which give rise to serious calculation error. It is, therefore, advisable to estimate the parameters which are not directly observable and to assume the values of other parameters which can be approximated through past experiences or past studies.

As for as the present model is concerned, it consists of a large number of parameters - z_i ($i = 1, 2, \dots, p$), α_i ($i = 1, 2, \dots, p$), θ , a , b and h . In the absence of an appropriate estimation technique for estimating all the parameters, we have shown the application of the model considering only two values of z in order to reduce the complexity in the calculation. It is just an illustration to show how the model can be utilized to estimate the conception rate in a particular region where the cultural dimension of physical separation of spouses is still prevalent.

The data on first complete conception time as given in Singh et al. (1985b) has been taken into consideration. These data relates to "A Demographic Survey of Varanasi (Rural)" conducted by Demographic Research Centre, Banaras Hindu University, Varanasi in the year 1969-70. A sample of about 2200 households scattered in 52 villages of Varanasi Tehsil (an administrative subunit of a district) was selected and a complete record of birth history of all eligible couples in a household was taken with greater emphasis on the fertility performance in last seven years. Because of some prevailing social taboos in this rural part of India, the phenomenon of physical separation of spouses was found quite common. Earlier studies on the data on first birth interval collected through this survey in the region, like Chakrabarty (1976); Singh et al. (1985 a); Dwivedi (1985); Singh and Singh (1983) have shown that the physical separation of spouses heaps mainly at two points; $z_1=0$ month and $z_2=12$ months with approximate probabilities 0.65 and 0.35. Moreover, the length of non-susceptible period associated with an incomplete conception may safely be assumed to be 6 months. With these known values of parameters z , α , θ and h , it can be possible to estimate the parameters a and b with appropriate data.

Table 1 presents the observed distribution of the time of first complete conception for females whose marriage duration was more than 10 years on the reference date of the survey. The mean and variance of the distribution were obtained as 2.7565 years and 6.0515 years square respectively. Using the expression (11) and (13), the moment estimates of a and b are obtained as –

$$\hat{a} = 280.41966, \quad \hat{b} = 7094.5212$$

Thus, we get $E(m) = a / b = 0.0395$, which is an estimate of conception rate (m). This may be considered as an estimate of fecundability in the region. Dwivedi (1985), Singh and Singh (1983) and Singh et al. (1985a) have also obtained the estimate of fecundability in the region as 0.0391, 0.0390 and 0.0392 respectively. This shows the validity of the model for the purpose of estimating fecundability in such a population which observes physical separation of the spouses just after their marriage.

Table 1 depicts the expected frequencies as obtained through the present model and as reported in Singh et al. (1985 b). If increased proximity of expected frequencies with the observed ones in a measure of improvement of the model, the suggested model may be considered as an improvement over Singh et al. (1985 b) model.

Concluding Remarks

The salient feature of the suggested model is that it takes into account one of the important causes of delay in first complete conception, namely, physical separation that is generally a common phenomenon in some of the developing societies. Moreover, it also takes into account the variability of fecundability over a heterogeneous group of married women. Thus, the model has a wider application as regards to the problem of estimation of fecundability or other characteristics in a specific population of women. It is remarkable here that the model is developed with a probability mass function for the random variable 'physical separation', thus, giving scope to consider a series of probability distributions to the variable z , probably using simulation models. However, as far as the practical application of the model is concerned, it would be always better to keep in mind the previous results so as to select an appropriate distribution of z in order to make the model simple for application purpose.

Acknowledgements:

The authors wish to express their thanks to the two anonymous referees for their valuable comments and suggestions in the earlier draft of this paper.

Table 1
Distribution of the Time for First Complete Conception

Time in Months	Observed Frequency	Expected Frequency under	
		Suggested Model	Singh et al. (1985b) model
0-15	108	115	119
15-27	97	91	88
27-39	63	60	59
39-51	44	40	39
51-63	26	26	26
63-75	16	17	17
75-87	11	11	11
87-99	7	8	9
99+	11	15	15
Total	383	383	383
X^2		2.62	4.41
Degrees of freedom		6	6

References:

- Chakrabarty, K. C. 1976. Some Probability distribution for Birth Interval, Unpublished Ph.D. Thesis, Banaras Hindu University, Varanasi, India.
- Dwivedi, S. N. 1985. Some Probability Models for Analysis of Birth Interval Data, Unpublished Ph.D. Thesis, Banaras Hindu University, Varanasi, India.
- Gini, C. 1924. Premieres Researches sur la Fecondabilite de la femme, *Proceedings of the International Mathematics Congress*, Toronto, 889.
- Singh, O. P. 1987. A Study of Certain Stochastic Models for Birth Intervals, Unpublished Ph.D. Thesis, University of Gorakhpur, Gorakhpur, India.
- Singh, S. N. 1964. On the time of first birth, *Sankhya*, 26 B: 95
- Singh, S. N., Chakrabarty, K. C. and Singh, V. K. 1976. A modification of a continuous time model for first conception, *Demography*, 12: 37.
- Singh, V. K. and Singh, U. N. 1983. Period of temporary seperation and its effect on first birth interval, *Demography India*, 12: 282.
- Singh, S. N., Singh, V. K. and Buman, D. 1985a. A continuous time model for first birth interval in rural environment of India, *Rural Demography*, 12: 55.
- Singh, U. P. Singh, V. K. and Singh, O. P. 1985b. On a distribution of first birth interval with random components, *Janasamkhya*, 3(2): 91.
- Srinivasan, K. 1966. An application of a probability model to study the interlive intervals, *Sankhya*, 28 B: 175.

Appendix

Under the assumptions (ii) and (iv) the p.d.f. of $\sum_{k=1}^{n+1} X_k$ for given value of n will be–

$$f_1(t/n) = \frac{m^{n+1} t^n e^{-mt}}{n!} ; t > 0, m > 0 \quad (A1)$$

So that the unconditional p.d.f. of $\sum_{k=1}^{n+1} X_k$ under assumption (iv) is

$$f_2(t) = \sum_{n=0}^{\lfloor t/n \rfloor} \theta (1-\theta)^n f_1(t/n) \quad (A2)$$

Now combining the assumptions (i), (ii), (iv) and (v), the conditional distribution of T (for a particular woman among a group of women) for given m can be obtained as–

$$f(t/m) = \begin{cases} 0 & ; 0 \leq t \leq z_1 \\ \sum_{s=1}^{i-1} \alpha_s \sum_{r=0}^{Q_s} \frac{\theta (1-\theta)^r m^{r+1} (t-rh-z_i)^r e^{-m(t-rh-z_i)}}{r!} ; & z_{i-1} \leq t \leq z_i; \text{ for } i = 2, 3, \dots, p \\ \sum_{s=1}^i \alpha_s \sum_{r=0}^{Q_s} \frac{\theta (1-\theta)^r m^{r+1} (t-rh-z_i)^r e^{-m(t-rh-z_i)}}{r!} ; & t \geq z_p \end{cases} \quad (A3)$$

A Probability Distribution for First Birth Interval

If the variability in the conception rate, m among women as given under the assumption (iii) is considered, the unconditional distribution of T is given as

$$f(t) = \int_0^\infty f(t/m)g(m)dm \quad (A\ 4)$$

Using the form of $f(t/m)$ as obtained in (A3), the expression of $f(t)$ can be obtained as

$$f(t) = \begin{cases} 0 & ; \quad 0 \leq t \leq z_1 \\ \sum_{s=1}^{i-1} \alpha_s \sum_{r=0}^{Q_s} \theta (1-\theta)^r \frac{b^a (t-rh-z_i)^r}{B(a, r+1) (b+t-rh-z_i)^{a+r+1}}; & z_{i-1} \leq t \leq z_i; \text{ for } i = 2, 3, \dots, p \\ \sum_{s=1}^i \alpha_s \sum_{r=0}^{Q_s} \theta (1-\theta)^r \frac{b^a (t-rh-z_i)^r}{B(a, r+1) (b+t-rh-z_i)^{a+r+1}}; & t \geq z_p \end{cases} \quad (A\ 5)$$

Integrating $f(t)$ in the appropriate ranges of t , the distribution function $F(t)$ of T can be deduced from (A5). Hence the expression (6) follows.