INDIAN JOURNAL OF POPULATION AND DEVELOPMENT
ISSN: 2583-4827; Volume 5(1), January 2025: 55-70

A Probability Model for the Analysis of Truncated Birth History

Brijesh P Singh
Kaushalendra K Singh

Abstract

Fecundability and sterility are two biological parameters that have attracted statistical demographers since 1960s. These parameters cannot be observed directly in the population but using of probability models for birth intervals and the number of births, it is possible to estimate these biological parameters. Fecundability is the monthly chance of conception and conception rate is yearly measure that can be obtain by multiplying 12 in fecundability. In this paper, we have used a probability model for inter-live birth interval to estimate these biological parameters and the probability of proceeding for the next birth. We have applied the model to two data sets which are almost one generation apart. The fitting of the model shows that the variability in the estimated parameters is very low indicates the consistency of the estimates. One of the interesting findings of the analysis is that parameters (rate of increase in conception rate and sterility) have increased over time however, the conception rate is slightly higher for younger age females and lower for the elder age females in the recent data than for the older data. To understand spatial pattern Uttar Pradesh, Bihar and Madhya Pradesh NFHS data have been used that indicates the fertility is higher in Bihar than other two states.

Introduction

The level of fertility is directly related to the proportion of females producing children and inversely related to the interval between successive live births. The study of the two aspects of fertility – spacing between births and parity progression ratio have been an interesting area of research for demographers to gain an insight of the fertility process which is a sequential time dependent process. Models for birth intervals are usually related to the timing of the first birth, inter-live birth interval (also known as closed birth interval) and open birth interval. Some other intervals like straddling, forward and interior birth intervals have also been of interest to researchers in the field of statistical demography. It is worthwhile to mention her is that the inter-live birth interval (the birth interval occurring during a specific period) is different than last or most recent closed birth interval.

For females of constant fecundability, a geometric distribution is used for the waiting time from marriage till conception, while for females of heterogeneous fecundability, the resulting distribution for the waiting time till conception is the beta geometric (Gini, 1924; Henery, 1958; Srinivasan, 1966; Singh et al, 2018). Singh (1964) has developed a probability model of waiting time up to the first conception by using the exponential distribution with a fecundability parameter following a Pearson type-III distribution.

Different mathematical models have been proposed to explain the nature of birth intervals and have been applied to the real time data to estimate fecundability and sterility (Bhattacharya, 1971; Sheps and Menken, 1973; Leridon, 1977; Mode, 1985). It is usually assumed in most these models that, all females are fecund at the time of marriage and fecundability is constant for a female till the first conception. These models also assume that fecundability may vary across females. However, these models, often, do not describe the data satisfactorily, especially when the age at the start of cohabitation is low (Singh 1964; Suchindran and Koo, 1999). Bhattacharya et al (1989) has described a model for the time of first birth which takes fecundability as the time dependent variable during the early period of married life and thus indirectly incorporates biological as well as socio-cultural factors responsible for low fecundability. Suchindran and Lachenbruch (1974) have also estimated parameters of a model for the first live birth interval. Sterility, which is biologically important, has also been studied using the birth interval data (Pathak and Prasad, 1977; Nair, 1983). Moreover, Singh et al (2002) have used the Singh (1968) model for the number of births and have obtained maximum likelihood estimate of fecundability and sterility over time and found that both are increasing.

Under the natural fertility conditions, usually, a closed birth interval is decomposed into some components viz. post-partum amenorrhea period, menstruating interval, time added by foetal wastages/temporary separation (due to short visit of the female to her maternal home) and gestational period and the distribution of the sum of these components is derived using the theory of semi-Markov process with stationary transition densities. The closed birth intervals are useful in studying the pattern of reproduction and estimation of certain parameters underlying the reproductive process of those females who continue to reproduce. Considerable attention has been paid towards the formulation of probability models for inter-live birth intervals under various sets of assumptions, especially for explaining data collected under different sampling frames. The details of this work till 1972 is given in Sheps and Menken (1973). An excellent survey and critical review of the work can also be found in Mode (1985). Braun (1977) extended D'Souza (1973, 1974) work and developed models for inter-live birth intervals capturing some salient features of the data for describing the whole reproduction process. Braun and Hoem (1979), Heckman and Singer (1982) proposed models incorporating co-variates information. George (1973) proposed a simple probability model and then generalized this (George and Mathai, 1975). Further, Bhattacharya et al (1986, 1988) derived a probability model for inter-live birth intervals which is applicable in situation where practice of abstinence following a child birth and taboos regulating coital frequency during the early part of the interval are widespread. Suchindran and Horne (1984) and Horne et al (1990) modelled some selected aspects of childbearing process and explained the parameters involved. Some probability models have also been developed considering various socio-demographic setups for the closed birth interval (Pandey et al, 1998; Mukherjee et al, 1991; Singh, 2002).

Several techniques are available in the literature for analysing the birth interval data from cross-sectional fertility surveys containing retrospective birth history of females of childbearing age. However, the completeness with which retrospective surveys collect information for the number and the timing of births is a controversial issue, even in case of fully designed surveys. Those children who were born many years in the past and have either died at an early age or living separately or working elsewhere for a long time (married daughters, son living separately or working elsewhere) are often under reported by the respondents. Thus, omission of events lengthens the birth interval and lowers the order of the subsequent birth interval. Further, error in dating of events that occurred deep in the past also influences the length of the birth interval.

In recent years, many retrospective fertility surveys have collected truncated birth histories (births in the last five years plus births earlier than five years and after the first birth) of females. The rationale for including a truncated birth history is based on economy and is preferable for areas where past demographic rates are well documented and where interest is in recent experience. There is, therefore, a need of modelling the truncated birth history data. In this paper, we apply a probability model that can be used to estimate biological parameters related to sterility and fecundability from the distribution of inter-live birth interval from the truncated birth history data. We also estimate the rate of change in conception rate as the chance of conception in a calendar year. The conception rate is different from fecundability which is the chance of conception during a lunar month. The model applied in this paper is a linear function of time. Mukherjee et al (1991) has analysed the conception in terms of polynomial of degree one and degree and as constant in the analysis of the last closed birth interval. The present approach is simple than the approach adopted by Mukherjee et al (1991) but the results are very similar. The information on number of females with no birth during the truncated period, distribution of females with exactly one birth during the truncated period by the time between start of the observation period and the birth, and the distribution of females with two or more births during the truncated period by length of the last closed interval have been used to obtain estimates. To test the suitability, the model has been applied to two real data sets having a gap of 25 years or almost one generation to explain the variation in the parameters over time. The R software has been used to obtain estimates of the parameters of the model.

The Model

This section gives a brief description of the model. Suppose married females of current age "b" are sampled at some time T_2 and the data of the two most recent births if they occurred during the preceding T years of the survey are recorded. The proportion of females with no birth during (T_1, T_2) where $T_1 = (T_2 - T)$, distribution of proportion of females with exactly one birth according to the time between T_1 and the birth, and distribution of proportion of females with two or more births during (T_1, T_2) according to the length of the closed interval are obtained under the following assumptions:

(i) T_1 is a distant point since marriage and the parameters of reproduction has been constant for the period considerably prior to female's age "a" so that the equilibrium is attained at T_1.

(ii) Let α is the probability that a female was fecund at T_1 so that $(1-\alpha)$ is the probability that a female was sterile at T_1 and remains sterile during the period (T_1, T_2).

(iii) After a live birth, the female proceeds to the next birth with probability β and for such a female, the probability that the next birth will occur on or before time t is $K(t)$.

Let us consider a female fecund at time T_1. Let $S_1, S_2, \ldots\ldots$ is the time of successive births after T_1, and let $X_1 = S_1$ and $X_i = S_i - S_{i-1} (i = 2,3\ldots)$. Since the reproduction process was in equilibrium at time T_1, the probability density function of the time of the first live birth is

$$k^*(t) = \int_0^t \left[\frac{1-K(x)}{\mu}\right] dx \quad (1)$$

where $K(x)$ is the distribution function of inter-live birth interval for a female who proceeds to the next birth and μ is the mean length of the inter-live birth interval and is given by

$$\mu = \int_0^\infty [1-K(y)]dy \quad (2)$$

$X_2, X_3 \ldots\ldots$ are identically and independently distributed random variables each having distribution function $\beta K(x)$.

The proportion of couples with no birth during (T_1, T_2) is given by

$$B_0^*(T) = (1-\alpha) + \alpha[1-k^*(T)] \quad (3)$$

The proportion of females fecund at T_1 and will deliver exactly one birth during (T_1, T_2) which will occur during $(t, t+dt)$, $(T_1 < t < T_2)$ is

$$B_1^*(t)dt = \alpha P[t < S_1 < t+dt, S_2 > T]$$
$$= \alpha P[t < S_1 < t+dt, X_2 > T-t]$$
$$= \alpha k^*(t)[1-\beta K(T-t)]dt \quad (4)$$

The inter-live birth interval within the period (T_1, T_2) could be observed only for those females who have given two or more live births during the period. The proportion of couples fecund at T_1 and who have given exactly $i (i \geq 2)$ births during (T_1,T_2) and the length of the interval between $(i-1)^{th}$ and i^{th} which is smaller or equal to t is

$$\alpha\beta^i P[(X_i \leq t) \cap (S_1 \leq T) \cap (S_{i+1} \leq T)] + \alpha\beta^{i-1}(1-\alpha)P[(X_i \leq t) \cap (S_i \leq T))]$$

Thus, the proportion of couples fecund at T_1 with the length of the inter-live interval lying between t and $(t+dt)$ is

$$B_2^*(t)dt = \sum_{i \geq 2} \{\alpha\beta^i P[(S_{i-1} \leq T-t) \cap (S_{i-1} + X_{i+1} > T-t)]k(t)dt$$
$$+\alpha\beta^{i-1}(1-\alpha)P[S_{i-1} \leq T-t]k(t)dt\}$$
$$= \sum_{i \geq 2} \left\{ \begin{matrix} \alpha\beta^i [K^*\$K^{(i-2)}(T-t) - [K^*\$K^{(i-1)}(T-t)]k(t)dt + \\ \alpha\beta^{i-1}(1-\alpha)K^*\$K^{(i-2)}(T-t)k(t)dt \end{matrix} \right\}$$
$$= \alpha\beta K(T-t)k(t)dt \quad (5)$$

where $K^{(i-1)}(t)$ is the *n*-fold convolution of $K(t)$ with itself and the symbol \$ stands for the convolution and $k(t)$ is the density function of closed birth interval $K(t)$.

Bhattacharya et al (1988) proposed a probability model for inter-live birth interval which is applicable in situations where practice of abstinences following a child birth and taboos regulating coital frequency during the early part of interval are widespread. The model assumes that coitus starts after abstinence and increases with time up to a certain point and then it becomes constant till the next birth. The distribution, *K(t)* of the length of the inter-live birth interval is derived under the following assumptions:

(1) The duration of post-partum amenorrhea (PPA), say U, and the period of sexual abstinence, say V, following a live birth are independently distributed, non-negative random variables with corresponding distribution functions $G_1(t)$ and $G_2(t)$ respectively. The distribution of the non-susceptible period say, Z, associated with a live birth is given by $Z = max(U, V)$

(2) For a female with $Z = z$ and $V = v$, the conditional instantaneous risk of conception following a live birth after time t is

$m(t|v); t > z$ and $m_0 = \lim_{t\to\infty}[m(t|v)]$; for all v

Since coitus resumes after the period of abstinence, its frequency and consequently $m(t|v)$ is assumed to depend on the duration of the postpartum abstinence and t, until a conception occurs or the normal level is attained, whichever is earlier.

(3) θ is the probability that a conception results in a foetal loss, $0 \leq \theta < 1$.

(4) The length of the non-susceptible period comprising the duration of pregnancy and PPA associated with foetal loss is an exponentially distributed random variable with mean $\frac{1}{c}, c > 0$. Where *c* is the parameter of the exponential distribution.

(5) The conditional instantaneous risk of conception following the termination of the non-susceptible period following a foetal loss is m_0. Functional forms of $G_1(t)$,$G_2(t)$,$m(t|v)$ and constants involved therein and the parameters θ and *c* do not change with age and parity in the interval (T_1, T_2).

Under the assumption (1), model given by equation (5) can further be formulated as follows:

The duration of postpartum abstinence will be the period of non-susceptibility when PPA is less than or equal to the period of abstinence. In this case the probability that Z will lie in the interval $(z, z + dz)$ is $G_1(z)dG_2(z)$.

Again, the duration of non-susceptibility will be the duration of PPA when the period of abstinence is less than the duration of PPA. Thus, the probability that Z and V will lie in the interval $(z, z + dz)$ and $(v, v + dv)$ respectively $(0 < v < z)$ is $dG_1(z)dG_2(v)$.

Therefore, the proportion of females with two or more live births with the length of inter-live interval lying between t and $(t + dt)$ is

$$B_2^*(t)dt = \int_{[0,t)} G1(z)dG2(z)\{\alpha\beta K*(T-t|z,z)k(t|z,z)dt\}$$
$$+\int_{[0,t]} dG1(z)\int_{[0,t]} dG2(v)\{\alpha\beta K*(T-t|v,z)k(t|v,z)dt\} \quad (6)$$

The proportion of females with exactly one live birth which occurred during t and $(t+dt)$ is

$$B_1^*(t)dt = \int_{[0,t)} G1(z)dG1(z)\{\alpha k*(t|z,z)[1-\beta K(T-t|z,z)]dt\}$$
$$+\int_{[0,t]} dG1(z)\int_{[0,z)} dG2(v)\{\alpha k*(t|v,z)[1-\beta K(T-t|v,z)]dt\} \quad (7)$$

The proportion of females in the population with no birth during (T_1,T_2) is

$$B_0^*(T) = \left\{1-\int_{(0,T]}\int_{[0,t)} \alpha G1(z)\,dG2(z)k*(t|z,z)dt\right\}$$
$$+\left\{1-\int_{[0,T]}\int_{[0,t)} \alpha dG1(z)\int_{[0,z)} dG2(v)\,k*(t|v,z)dt\right\} \quad (8)$$

The proportion of females with exactly one live birth during (T_1, T_2) and time between T_1 and time of the birth smaller than or equal to t, say $B_1(t)$, the proportion of females with two or more live births during (T_1, T_2), and length of inter-live birth interval smaller than t, say $B_2(t)$, are given by

$$\left.\begin{aligned} B_1(t) &= \int_0^t B_1^*(t)\,dt, 0<t<T \\ B_2(t) &= \int_0^t B_2^*(t)dt, 0<t<T \\ \text{and } B_0(T) &+ B_1(T) + B_2(T) = 1 \end{aligned}\right\} \quad (9)$$

Illustration

We have applied the above model to analyse the temporal and regional variation in selected indicators of reproductivity. The analysis of the temporal variation is based on two datasets available from two surveys carried out in district Varanasi of Uttar Pradesh. The first dataset is based on the survey "Status of Women and Fertility in Eastern Uttar Pradesh" which was conducted in 1996 (Singh, 1998). This survey covered 1432 eligible females aged 15-49 years in district Varanasi of Uttar Pradesh, India. The second dataset is based on the data available from the fifth round of the National Family Health Survey (NFHS-5) which is a large-scale, multi-round survey conducted across India that provides essential information on population, health, and nutrition indicators (Government of India, 2022). This survey covered 1403 eligible females aged 15-49 years in district Varanasi.

The regional analysis, on the other hand is based on the data available from NFHS-5 for three states of India – Bihar, Madhya Pradesh, and Uttar Pradesh – which are among the high fertility states of the country. A comparison of the results from the temporal perspective based on the two surveys about 25 years apart and from the regional perspective helps in understanding the change and the variation in fecundability and sterility among females.

The analysis is confined to females aged 25-45 years with an effective marriage duration of at least 12 years who did not produce a live birth during the interval (T_1, T_2), women who produced exactly one live birth during the interval (T_1, T_2), and women who produced two or more live births during the interval (T_1, T_2). Only those females who had at least 12 years of effective marriage duration (married for five years or more, 7 years preceding the survey) were considered because the reproductive process may be assumed to have attained equilibrium by that time. Among the eligible female, only those who or whose husband did not adopt any terminal method of family planning program and who were the usual residents of the village were included in the study. Information on the current reproductive status of females (menstruating, pregnant, amenorrhoeic, menopaused) on the reference date of the survey was collected and those females who had reached menopause were excluded.

The model requires information on the distribution of PPA the period of abstinence τ (the time beyond abstinence during which coital frequency depends on time), θ, the incidence of foetal loss. The mean duration of the non-susceptible period associated with the foetal loss was taken as 3 months and 0.15 is the incidence of foetal loss. Analysis of the data on PPA from surveys in Eastern Uttar Pradesh and in rural areas of India and Bangladesh, where extended breastfeeding is the norm reveals that the distribution of PPA has two modes, one within few months after birth and other many months latter (Misra et al, 2021). I have, therefore, considered two groups of females whose PPA takes two values t_1 and t_2 with probability p_1 and p_2 respectively, $0 < p_1 \leq 1$, $p_1 + p_2 = 1$. Empirical data available from the survey suggests t_1=0.1 years and t_2=1.00 year and the values of p_1 are taken as 0.60, 0.50 and 0.45 for females aged 25-30 years, 30-35 years, and 35-45 years, respectively.

Now let us consider that –

1. The duration of postpartum abstinence is same for all females, and its length is τ_1.
2. The conception rate $m(t|\tau_1)$ is a polynomial of degree r in t for $\tau_1 < t \leq \tau_2, (\tau_2 = \tau_1 + \tau)$ and constant thereafter and is of the form

$$m(t|\tau_1) = \begin{cases} \sum_{j=0}^{r} q_j (t-\tau_1)^j & for\ z < t \leq \tau_2 \\ \sum_{j=0}^{r} q_j \tau^j & for\ t > \tau_2 \end{cases} \tag{10}$$

When we assume the conception rate $m(t|\tau_1)$ to be constant then the distribution involves only one parameter q_0 (conception rate at the start of cohabitation). When we assume the conception rate $m(t|\tau_1)$ to be a polynomial of degree one, then the distribution involves two parameters q_0 (conception rate at the start of cohabitation) and q_1 (a measure of rate of increase).

Under these assumptions the expressions for $B_2^*(t)dt, B_1^*(t)dt$ and $B_0^*(T)$ in (6), (7) and (8) reduce to

$$B_2^*(t)dt = \sum_{i=1}^{2} p_i [\alpha\beta K * (T-t|i) k(t|i) dt] \tag{11}$$

$B_1^*(t)dt = \sum_{i=1}^{2} p_i\{\alpha k * (t|i)[1 - \beta K(T - t|i)dt]\}$ (12)

and $B_0^*(t)dt = p_i\{1 - \int_0^T \alpha k * (t|i)dt\}$ (13)

where $K * (t|i)$ and $K(t|i)$ are the distribution function of time of first recording and between successive recordings of births for a female with $V = \tau_1$and $U = t_i (i = 1,2)$. $k * (t|i)$ and $k(t|i)$ are the corresponding density function of $K * (t|i)$ and $K(t|i)$ where

$K(t|i) = \sum_{j=0}^{2} A_j\, K_j(t|i)$ (14)

Denote by $M(t|i) = \int_{z_i}^{t-g} m(x|\tau_1)dx, z_i = max(\, t_i, \tau_1)(0 \le z_i < t - g)$

$h_i = t_i + g, i = 1,2$ and $h' = \tau_1 + g$ then

$K_0(t|i) = 1 - exp\{-M(t|i)\}$ (15)

The function $M(t|i)$ reduces to the following for $h' \ge hi$

$$M(t|i) = \begin{cases} \varphi(t) & ;\, if\ h' < t \le h' + \tau \\ \varphi(\tau + h') + \sum_{j=0}^{r}(q_j\tau^j)\,(t - h' - \tau) & ;\, if\ t > h' + \tau \end{cases} \quad (16)$$

For $h' < h \le h' + \tau$

$$M(t|i) = \begin{cases} \varphi(t) - \varphi(h_i) & ;\, if\ h' < t \le h' + \tau \\ \varphi(\tau + h') - \varphi(h_i) + \sum_{j=0}^{r}(q_j\tau^j)\,(t - h' - \tau) & ;\, if\ t > h' + \tau \end{cases} \quad (17)$$

For $h_i > h' + \tau$

$$M(t|i) = \sum_{j=0}^{r} q_j\, \tau^j (t - h_i) if t > h_i$$

and $\varphi(t) = \sum_{j=0}^{r} \frac{q_j}{j+1}(t - h')^{j+1}$ (18)

Similarly, $K_j(t|i), (j = 1,2)$ can be obtained. Scoring method was used to obtain maximum likelihood estimates of the parameters of the model (11, 12 and 13) (Singh, 2002). In the estimation of parameters, it was assumed that θ=0.15, $\frac{1}{c}$=0.5 years, τ=0.25 years and τ=2 years. When $m(t|\tau_1)$ is linear, the distribution involves four parameters q_0, q_1, α and β. The initial value of the q_0, q_1 can be obtained by equating the observed mean of the inter-live birth interval of females with two or more births to its theoretical values. The initial estimates of β and α were obtained as a solution of the equation (11, 12 and 13)

$$\hat{\beta} = \frac{1 - \frac{(N-n-n')}{N}}{1 - [1 - \hat{K}*(T)]} = \frac{n+n'}{N\hat{K}*(T)} \quad (19)$$

and $$\hat{\alpha} = \frac{\beta\left(1 - \frac{(N-n-n')}{N} - \frac{n'}{N}\right)}{1 - [1 - \hat{K}*(T)] - \int_0^T \hat{B}_1(t)dt} = \frac{\beta n}{N[\hat{K}*(T) - \int_0^T \hat{B}_1(t)dt]} \quad (20)$$

where N, n and $n'$are respectively the total number of females under study, number of females with two or more births, and number of females with exactly one birth during (T_1, T_2).

Table 1: Distribution of the females according to zero birth, exactly one birth and two or more births during 7 years preceding the survey date by current age (1996)

Number of births	Intervals (in years)	Current age of females in years whose effective marriage duration is ≥ 12											
		25-30		30-35		35-45		25-30		30-35		35-45	
		O	E	O	E	O	E	O	E	O	E	O	E
		1996 Survey						2021 Survey					
Females with zero birth	-	9	9.4	35	35.7	140	141.2	7	6.8	31	33.9	128	126.4
Females with one birth	Interval between T_1 and first birth during 7 years before T_2												
	0.0 – 2.50	6	6.2	21	20.7	32	31.6	5	5.7	19	17.2	27	25.1
	2.50 – 4.50	5		14		15	13.8	4		12		11	8.1
	4.50 – 7.00	1	5.8	3	15.2	8	6.6	1	4.7	2	14.7	5	5.4
Females with two or more births	Inter-live birth interval												
	0.0 – 1.25	5	6.3	14	9.4	9	6.4	5	5.2	14	15.3	8	6.9
	1.25 – 1.75	7	8.5	12	17.6	12	13.3	7	7.3	10	9.1	12	10.1
	1.75 – 2.50	45	46.6	69	62.7	54	49.6	37	35.2	61	59.2	48	55.3
	2.50 – 3.25	33	30.7	33	36.5	29	32.4	29	30.7	28	30.1	27	33.6
	3.25 – 4.00	11	8.8	17	19.3	16	18.7	9	8.2	13	11.2	14	10.7
	4.00 – 4.75	3		5		9	11.1	3		5		7	
	4.75 – 7.00	2	4.7	3	8.9	6	5.3	1	4.2	2	6.3	4	9.4
	Total	127	127.0	226	226.0	330	330.0	108	108.0	197	197.0	291	291.0
	χ^2	1.36		5.60		3.23		0.405		1.238		5.315	
	df	4		4		6		4		4		5	
	p-value	0.851		0.231		0.779		0.982		0.871		0.379	

O-Observed; E-Expected

Source: Authors

Table 2: Estimates of the parameters of the model, variance, and correlation coefficients between estimates of the parameters (1996)

Particulars	Current age of those females whose marriage duration is ≥12 years (years)					
	25-30	30-35	35-45	25-30	30-35	35-45
	1996 Survey			2021 Survey		
q_0 (Conception rate)	0.453	0.344	0.294	0.461	0.338	0.285
q_1 (Measure of rate of increase in conception rate)	0.161	0.173	0.178	0.169	0.180	0.183
α (Probability that a female is fecund)	0.998	0.950	0.623	0.986	0.889	0.568
1-α (Probability that a female is sterile)	0.002	0.050	0.377	0.014	0.111	0.432
β (Probability that female proceeds for next birth)	0.985	0.974	0.811	0.980	0.949	0.782
Variance(q_0)	1.145	3.545	3.751	1.991	1.823	3.201
Variance(q_1)	0.743	1.094	2.133	0.736	0.948	1.661
Variance(α)	0.144	0.253	0.296	0.133	0.223	0.329
Variance(β)	0.193	0.304	0.585	0.203	0.281	0.419
Correlation$(q_0, q_1)(\times -1)$	0.943	0.875	0.799	0.911	0.880	0.734
Correlation$(q_0, \alpha)(\times -1)$	0.744	0.741	0.394	0.711	0.745	0.350
Correlation(q_0, β)	0.065	0.085	0.007	0.051	0.087	0.006
Correlation(q_1, α)	0.641	0.632	0.298	0.668	0.684	0.272
Correlation$(q_1, \beta)(\times -1)$	0.393	0.561	0.317	0.363	0.589	0.297
Correlation(α, β)	0.049	0.038	0.013	0.041	0.036	0.010

Source: Authors

Table 3: Estimates of the parameters, variances, and correlation coefficient between estimators for Uttar Pradesh, Bihar, and Madhya Pradesh NFHS-V data.

Particulars	Current age of those females whose marriage duration is ≥12 years (years)								
	Uttar Pradesh			Bihar			Madhya Pradesh		
	25-30	30-35	35-45	25-30	30-35	35-45	25-30	30-35	35-45
q_0 (Conception rate)	0.452	0.323	0.281	0.473	0.343	0.289	0.443	0.321	0.273
q_1 (Measure of rate of increase in conception rate)	0.163	0.179	0.183	0.171	0.183	0.186	0.159	0.176	0.178
α (Probability that a female is fecund)	0.987	0.891	0.564	0.991	0.902	0.582	0.985	0.889	0.553
1-α (Probability that a female is sterile)	0.013	0.109	0.436	0.009	0.098	0.418	0.015	0.111	0.447
β (Probability that female proceeds for next birth)	0.983	0.939	0.751	0.988	0.953	0.797	0.981	0.929	0.722
Variance(q_0)	1.691	1.859	3.413	1.892	1.823	3.226	2.012	1.623	3.221
Variance(q_1)	0.735	0.941	1.653	0.783	0.938	1.651	0.766	0.938	1.631
Variance(α)	0.130	0.233	0.341	0.131	0.203	0.349	0.123	0.212	0.301
Variance(β)	0.201	0.281	0.423	0.223	0.261	0.439	0.198	0.241	0.403
Correlation(q_0, q_1)($\times -1$)	0.911	0.883	0.727	0.901	0.877	0.714	0.901	0.892	0.763
Correlation(q_0, α)($\times -1$)	0.701	0.748	0.359	0.722	0.741	0.355	0.704	0.739	0.353
Correlation(q_0, β)	0.051	0.081	0.005	0.058	0.078	0.003	0.053	0.091	0.005
Correlation(q_1, α)	0.661	0.696	0.263	0.643	0.681	0.252	0.666	0.666	0.261
Correlation(q_1, β)($\times -1$)	0.343	0.581	0.291	0.323	0.566	0.293	0.367	0.579	0.281
Correlation(α, β)	0.038	0.033	0.009	0.042	0.031	0.011	0.048	0.037	0.013

Source: Authors

Discussion and Conclusion

The constant form of the hazard function gave a poor fit (using χ^2test) for all datasets so the results of the model are not presented here. However, the linear form of the hazard function $m(t|\tau_1)$ gave an adequate fit (using χ^2 test) for all age-groups in both 1996 survey and 2011 survey. The expected number of females with zero live birth, one live birth and two or more live births obtained from the model are presented in table 1 along with the observed number of women with zero live birth, one live birth and two and more live births. The difference between the observed and the estimated number of women has been found to be statistically insignificant as may be seen from the table.

On the other hand, estimates of the parameters of the model are presented in table 2. It can be observed from the table the estimate of conception rate at the start of cohabitation, $\hat{q}_0$, is the highest for females whose current age is 25-30 years while $\hat{q}_1$, a measure of the rate of increase in the conception rate, is the lowest. Thus, the age group 25-30 shows high conception rate in the beginning of the interval and subsequently with time, falls below the other two groups. Rate of increase of conception rate $\hat{q}_1$ increases with the increase in the age of the female. This pattern is observed in both the surveys. The values of $\hat{q}_0$ and $\hat{q}_1$ is, however, found to be higher in 2021 compared to those in 1996 in females aged 25-30 years but lower in females of other age groups. This may be due to the fact that females are delaying the birth and, are attaining maximum fecundability in the age group 25-30 years and after that due to increase in secondary sterility, lower sexual activity and use of contraceptive, the low value of estimates of $\hat{q}_0$ and $\hat{q}_1$ are observed in age group 30-35 and 35-45. The instantaneous risk of pregnancy following the previous live birth is assumed to increase during (τ_1, τ_2) and attains a plateau thereafter. In fact, both τ_1 and τ_2 may depend on the survival status of the child, breastfeeding practices, other demographic characteristics (such as age, marital duration, parity number of surviving children, and their age and sex) and cultural characteristics and they may vary from female to female. The variation in $m(t|\tau_1)$ may be incorporated in the present model in a manner similar to that discussed in Bhattacharya et al (1988). The measure of sterility $(1-\alpha)$ is observed more in all age groups in 2021 in comparison of 1996 due to the increasing behaviour of contraceptive use. The probability that the female proceeds for the next birth (β) is found lower in 2021 than in 1996 for all females due to voluntary sterility (vasectomy, tubectomy and other contraceptive use) and aversion of more child. The variance of all parameters in both datasets is quite low which indicates the consistency and the efficiency of the model. The correlation coefficient between $\hat{q}_0$ and $\hat{q}_1$ is very high and negative but decreases with the increase in the age of the female. It is observed that at the elder age the risk of conception is lower in comparison to the younger age. Similar pattern has been observed for correlation between conception rate and probability of sterility. The correlation between α and β is positive but very low for females of all ages in both surveys which means that females with higher fecundability proceed for the next birth.

We have also fitted the model to the data for Bihar, Madhya Pradesh, and Uttar Pradesh for the year 2021 based on the data available from NFHS-5. Table 3 presents the estimates of the parameter of the model for the three states. Among the three states, the conception rate, $\hat{q}_0$, and the rate of increase in the conception rate, $\hat{q}_1$, is found to be the highest in Bihar followed by Uttar Pradesh and Madhya Pradesh. The data available from

NFHS-5 also reveals that the fertility (measured in terms of the total fertility rate) is also the highest in Bihar (2.98 births per woman of childbearing age). followed by Uttar Pradesh (2.35 births per woman of child bearing age) and Madhya Pradesh (1.99 births per woman of childbearing age) (Government of India, 2022). The fecundability is also higher in Bihar compared to Uttar Pradesh and Madhya Pradesh in women of all age groups. It appears that relatively high fecundability thus high conception rate, along with relatively higher rate of increase in the conception rate with the increase in age in Bihar, are contributing factors to high current fertility in Bihar relative to Uttar Pradesh and Madhya Pradesh. The probability of proceeding to the next birth after a birth is also relatively higher in Bihar than that in Uttar Pradesh and Madhya Pradesh.

Factors that influence fecundability of women include demographic factors such as age, life style choices such as smoking and use of alcohol, health and nutritional status, particularly, sexually transmitted infections and hormonal imbalances, and a host of social and cultural factors such as the desire or the demand for children, social support system, and religious beliefs and cultural traditions. A discussion on how these and many other factors influence fecundability differentially in Bihar, Uttar Pradesh and Madhya Pradesh is out of the scope of this paper. However, understanding the determinants of fecundability in the three states may provide an insight about the variation in the current fertility across the three states.

On the other hand, the probability of a woman being sterile is estimated to be, the highest in Madhya Pradesh but the lowest in Bihar. Sterility in women can be classified into primary and secondary sterility and an important factor in secondary sterility is the sickle cell disease which has comparatively high prevalence in Madhya Pradesh. The prevalence of sickle cell disease is particularly high in the tribal population and the proportion of the tribal population to the total population is the highest in Madhya Pradesh among the three states. Madhya Pradesh has the highest proportion of the tribal population.

The fitting of the model reveals that cross-sectional (across states) data have shown a decrease in the conception rate with the increase in the age of the female. The decrease in the conception rate may be attributed to various biosocial factors including age dependent infertility, use of contraceptive methods to regulate fertility and temporary separation between the partners which are normally absent during the early stages of the reproductive life. Conception rate depends upon fecundability which is affected by the coital frequency, and which varies from female to female, being a matter of individual choice and because of various other factors. Traditional intercourse taboos of one form or the other that vary in the degree of restriction affect the coital frequency. In the joint family system in which there is a custom of occupying a separate physical space by the partners even in the night and meeting each other only when the situation permits such as when small kids and elders in the family are sleeping seriously compromises coital frequency. There is also, shortage of space in the house necessary for the privacy of the intimate act. Postpartum taboos on sexual activity to avoid pregnancy to prolong the duration of breast-feeding and other taboos also force couples to abstain from intimate relationship for a substantial period of time. Abstaining from the intimate act in the presence of a daughter-in-law or grown-up children in the family, and also possibly due to the poor nutritional status also has a telling impact on coital frequency. A comparison of the conception rate estimated from the data from 1996 survey with the data from 2021 survey suggests that conception rate in younger women has increased but has decreased in older women. Further, it is observed that temporal data reveals the measure of rate of increase in conception rate and sterility are increasing over time for all age groups due to the weakening of social taboos, increasing protected coital frequency. In any case, this is an issue which needs to be investigated further.

The analysis also reveals difference in fecundability and sterility across the three high fertility states of India. All the three states are from the same social and cultural zone of India and are amongst the poorest states of the country. There is a need to investigate the reasons behind the variation in the fecundability and sterility across the three states as both fecundability and sterility influences the level of fertility.

References

Bhattacharya BN (1971) On Some Stochastic Models for Couple Fertility. Unpublished Ph.D. Thesis, Banaras Hindu University, Varanasi, India.

Bhattacharya BN, Singh KK (1986) A probability model for interior birth interval and its application. *Canadian Studies in Population* 13(2): 167-180.

Bhattacharya BN, Pandey CM, Singh KK (1988) Model for interlive birth interval and some social factors. *Janasamkhya* 6(1): 57-77.

Bhattacharya BN, Singh KK, Singh U, Pandey CM (1989) An extension of model for first birth interval and some social factors. *Sankhya*, *Series B* 51: 115-124.

Braun HI (1977) Stochastic model for birth histories with application to the study of fertility patterns. Princeton, New Jersey, USA, Princeton University, Department of Statistics, Technical Report No. 124, Series 2.

Braun HI, Hoem JM (1979) Modeling cohabitation birth intervals in the current Danish population. Copenhagen, University of Copenhagen, Laboratory of Actuarial Mathematics. Working paper No. 24.

D'Souza S (1973) Interlive birth intervals of non-contraceptive population: A data analytic study. *Social Action* 23: 404-425.

D'Souza S (1974) *Closed Birth Intervals – A Data Analytic Study*. New Delhi, Sterling Publishers Private Limited.

George A (1973) A probability model for inter-birth interval. *The Canadian Journal of Statistics, Section C: Applications* 1(2): 201-209.

George A, Mathai AM (1975) A generalized distribution for interlive birth interval. *Sankhya, Series B* 37: 332-342.

Gini C (1924) First investigations on the fecundability of a woman. In *Proceedings of the International Mathematics Congress, Toronto* 2: 889-892.

Government of India (2022) *National Family Health Survey (NFHS-5) 2019-2021. India*. Mumbai, International Institute for Population Sciences.

Heckman JJ, Singer B (1982) Population heterogeneity in demographic models. In KC Land, A Rogers (Eds) *Multidimensional Mathematical Demography*, New York, Academic Press.

Henry L (1958) Intervals between confinements in the absence of birth control. *Eugenics Quarterly* 5(4): 200-211.

Horne AD, El-Khorazaty MN, Suchindran CM (1990) Statistical modeling of selected aspects of the childbearing process with application to World Fertility Survey countries. *Mathematical Population Studies* 2(3): 183-207.

Leridon H (1977) *Human fertility – The Basic Components*. Chicago, University of Chicago Press.

Mishra R, Singh KK, Singh Brijesh P (2021) Duration of post-partum amenorrhoea: a model-based approach. *Indian Journal of Population and Development* 1(1): 41-50.

Mode CJ (1985) *Stochastic Processes in Demography and their Computer Implementation*. Berlin, Springer-Verlog.

Mukherjee S, Singh KK, Suchindran CM, Bhattacharya BN (1991) A probability distribution for last closed birth interval. *Demography India* 20(2): 259-272.

Nair NU (1983) A stochastic model for estimating adolescent sterility. *Biometrical Journal* 25(6): 557-561.

Pandey A, Dwivedi SN, Mishra RN (1998) A stochastic model for the study of last closed birth interval with some biosocial components. *Mathematical Population Studies* 7(1): 1-27.

Pathak KB, Prasad CVS (1977) A model for estimating adolescent sterility among married women. *Demography* 14: 103-104.

Singh Brijesh P (2002) Fertility Analysis Through Application of Models. Unpublished Ph.D. Thesis, Banaras Hindu University, Varanasi, India.

Singh Brijesh P, Singh G, Singh KK (2018) On the number of menstrual cycles required for first conception: An insight of chance mechanism. *Demography India* 47(2): 01-15.

Singh KK (1998) Status of Women and Fertility in Eastern Uttar Pradesh. Report of a Research Project. Varanasi, Banaras Hindu University.

Singh KK, Singh Brijesh P, Singh U, Singh K (2002) Trends in fecundability and sterility over time. *Progress of Mathematics* 36(1 and 2): 111-129.

Singh SN (1964) On the time of first birth. *Sankhya* 26(1 and 2) B: 95-102.

Singh SN (1968) A chance mechanism of variation in number of births per couple. *Journal of American Statistical Association* 63: 209-213.

Sheps MC, Menken JA (1973) *Mathematical Models of Conceptions and Birth*. Chicago, University of Chicago Press.

Srinivasan K (1966) An application of a probability model to the study of interlive birth intervals. *Sankhya, Series B* 28: 175-182.

Suchindran CM, Lachenbruch PA (1974) Estimates of parameters in a probability model for first live birth interval. *Journal of American Statistical Association* 69: 507-513.

Suchindran CM, Horne AD (1984) Some statistical approaches to the modelling of selected fertility events. *Proceedings of American Statistical Association (Social Statistics Section)* Washington, DC, 629-634.

Suchindran CM, Koo HP (1999) Statistical models for human reproduction. In CR Rao, PK Sen (Eds) *Handbook in Statistics: Environmental and Public Health Statistics*. New York, Elsevier.