Uses and Importance of Models in Analysing Real Data for First Birth Interval

Shilpi Tanti

**Senior Research Fellow, DST-Centre for Interdisciplinary Mathematical Sciences, Department of Statistics, Faculty of Science, Banaras Hindu University, Varanasi-221005, E-mail: shilpi.tanti@gmail.com.*

K. K. Singh

***Professor, Department of Statistics, Faculty of Science, Banaras Hindu University, Varanasi-221005, E-mail:*

.

Acknowledgments:

I would like to acknowledge Prof. R. C. Yadava for his valuable suggestions & comments. I am also thankful to DST-Centre for Interdisciplinary Mathematical Sciences for financial assistance to carry out my research work

Note: This paper is conferred with Chandrasekaran Young Scientist Award in the National conference of Indian Association for the Study of Population held in GIPE, Pune during 13-15 Dec’ 2012.

Uses and Importance of Models in Analysing Real Data for First Birth Interval

Abstract:

The levels and trends of human fertility can be obtained from the descriptive study of the real data. But some inherent characteristics of this phenomenon namely fecundability, sterility, etc. can be estimated only by applying appropriate models to the real data. Here, an attempt has been made to explore the uses and importance of the probability models relating to human fertility through real data. In this study, different probability models have been considered to estimate the parameters by applying them to the data of the First Birth Interval in different major states of India obtained from National Family Health Survey - III.

1. Introduction:

A model is a mathematical abstract form of a real phenomenon, which includes the variables that may account for the explanation of the aspects taken into account. Modeling is one of the possible forms of scientific approach, often used in social sciences and particularly in demography to understand fertility, mortality, migration, nuptiality and other demographic measures. Broadly, a mathematical model is classified into two categories namely, Deterministic model and Stochastic model.

A model for a phenomenon is formulated keeping in mind that it embodies the essential features of the process. For this, a model builder generally makes certain assumptions about the process based on his/her experience and intuition and tries to describe the behavior of the process in terms of mathematical equations. The three important uses of the models are:

i) It can be used for prediction purposes. In fact, the phenomenon may have different components which may be inter-related and the model incorporates all these relationships in terms of a set of mathematical equations. Thus, a model provides a method for investigating the possible consequences in the process due to various alterations in the determinants of the process.

ii) It can be used for estimation of the parameters of the process by applying it to observed data relating to the process. These parameters will provide information about some unobserved characteristics of the process.

iii) It can be used for explaining certain apparent inconsistencies in the observed data relating to the phenomenon under consideration.

Models in the fertility analysis were initiated by Henry (1953) by using the concept of fecundability first put forth by Gini (1924). Analysis of the waiting time for first conception signifies couple's fertility at early stages of married life. That is why; this duration is very much influenced by age at marriage. The mean first birth interval (FBI), in case of lower ages at marriage, is higher due to adolescent sterility, short visits to parents, restrictions on frequent sexual union and other social norms and taboos, whereas, the mean FBI for higher ages at marriage becomes lower as the strictness of the social norms and taboos decreases with increasing age and also due to some other personal factors. Hence, the study of FBI ascertained according to different ages at marriage is more appropriate and logical.

FBI has some unique features to investigate since usually females don't like to use contraception to postpone the first birth and there is lower chance of recall lapse in reporting the time of first birth as it is the most important event in the life of the female. The nature of FBI is again somewhat different from other birth intervals as it is not influenced by the post-partum amenorrhea (PPA) period and thus it is generally studied separately from birth intervals of higher order.

The length of FBI depends on the conception rate or fecundability of the females. The terms fecundability and conception rate are dependent on time, whether it is taken as discrete or continuous. If unit of time is taken as one month then the conception rate may be interpreted as fecundability. If the unit of time is taken as one year then it is known as yearly conception rate. Conception rate is analogous to hazard rate used in life testing problem. Conception rate is the risk of conception in time $(t, t + \Delta t)$ under the condition that conception has not occurred in time$(0, t)$. The probability that an event will occur during a time interval is proportional to the length of that time interval.

Keeping the primacy of the models, the present paper focuses on the uses and importance as well as comparison of the models of first conceptive delay followed by the distribution of first birth interval (FBI), in particular, to the females of different ages at marriage of specific marital duration.

2. Models & the assumptions:

Conception rate is estimated many times from the data on time for first conception through the technique of probability modeling where the event of occurrence of conception is assumed as random. Generally, data on first conception time are not available and these are obtained from the data on FBI on the assumption that there is one to one correspondence between conception and live birth. Hence, subtracting gestation period (9 months or 0.75 years) from the duration of FBI, one may have data of first conception. In literature, there are many crucial assumptions for the indirect estimation of conception rate. These assumptions are broadly classified into three categories: (Pratap, 2011)

I. Conception rate of each female is constant till the time of first conception and population is homogeneous with respect to conception rate.

II. Conception rate of each female is constant till the time of first conception but population is heterogeneous with respect to conception rate.

III. Conception rate is time dependent (time being measured from the time of marriage).

It is expected that assumptions (I) and (II) may be more appropriate for females of higher ages at marriage depending on homogeneity and heterogeneity in the population while the assumption (III) may be more appropriate for females of lower ages at marriage, say less than 15-16 years, as it indirectly incorporates the adolescent sterility and other social norms and taboos associated with it (Pathak, 1977; Pathak, 1978; Pathak, 1981; Nair, 1983; Nair, 1983a; Bhattacharya, 1988). The fertility behavior of a female who married in her adolescent ages becomes quite different than that of a female who married at later ages on account of biological phenomena. Under this situation, conception rate may be assumed to have an increasing trend over time or increasing up to some level and then remaining constant.

Here four probability models of FBI are discussed and then applied to the observed data of FBI to check the adequacy of the models. Model I and II are derived for females of higher ages at marriage whereas the assumptions of Model III and IV are more suitable for females of lower ages at marriage.

Model I:

Model I is derived on the assumption that conception rate is constant for each female from marriage to first conception. Let X denote the time between marriage and the first conception. If the time is treated as continuous then the assumption (I) implies that the chance of conception between time t and $(t + \Delta t)$ is $\lambda.\Delta t + O(\Delta t)$ with p.d.f of X as

$$f_1(x) = \begin{cases} \lambda e^{-\lambda x} \\ 0 \;;\; otherwise \end{cases}; \qquad x > 0, \qquad \lambda > 0, \quad \dots (2.1)$$

Here λ represents the conception rate per unit of time.

Model II:

The females under study are usually coming from various socio-economic, demographic and biological backgrounds. Hence, the assumption of constant conception rate may not be reasonable. In a study, the conception rate showed a declining trend with increasing time for females of higher ages at marriage (Pratap, 2011). This feature cannot be completely removed by disaggregation and is usually viewed as a selection effect in which the more fecund females tend to conceive first. To capture this selection effect, it may be assumed that for a female X (duration from marriage to first conception) follows the density given in Equation (2.1), where λ varies in the population from female to female and λ follows a probability distribution with p.d.f. $f(\lambda)$. Hence, if a female under study is randomly selected from the population, then the unconditional distribution of X is given as

$$f_2(x) = \int_0^\infty \lambda e^{-\lambda x} f(\lambda) d\lambda \;; \quad \cdots \; (2.2)$$

This mixture can take a parametric form or be left arbitrary. The most widely published model for incorporating heterogeneity in conception rate is Pearson type III distribution. The choice of this distribution is due to its flexibility, mathematical applicability and interpretation.

$$f(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{(\alpha-1)} e^{-\lambda\beta} \;; \;\; \alpha > 0, \qquad \beta > 0, \cdots \; (2.3)$$

where α and β are positive constants and $\Gamma\,(.)$ is the gamma function.

Under this situation, $f_2(x)$ becomes

$$f_2(x) = \frac{\alpha\beta^\alpha}{(\beta + x)^{(\alpha+1)}} \;; \qquad \alpha > 0, \qquad \beta > 0, \qquad x > 0, \quad \cdots \; (2.4)$$

It should be noted that λ differs from female to female and it is constant over time for a fixed female.

Model III:

Model III is derived on the basis of the assumption that conception rate is $\lambda(t)$, which is a function of time *t*. Under this situation;

$$f_3(x) = \begin{cases} \lambda(t)e^{-\int_0^x \lambda)(t)dt} & ; \quad x > 0, \\ 0 \quad ; \quad otherwise \end{cases} \qquad \dots (2.5)$$

If $\lambda(t)$ is assumed as a linear function of time, i.e., $\lambda(t) = a + bt$, then the probability density function of X is given by

$$f_3(x) = \begin{cases} (a+bx)e^{-(ax+b\frac{x^2}{2})}; & x > 0, \\ 0 \quad ; \quad otherwise \end{cases} \qquad \dots (2.6)$$

Model IV:

Model IV is derived under the following assumptions (see Nath et. al, 1995):

i) The cohort of females is a mixture of two groups - (a) the adolescent sterile group (those who are not biologically mature at the time of marriage but are exposed to the risk of Ovulation) and (b) the ovulation group (those who are biologically mature at the time of marriage and are exposed to the risk of conception). Let θ and $1-\theta$ be the proportions of two types of females respectively.

ii) For group (a) females, the interval between marriage and the time of ovulatory menstruation follows a negative exponential distribution with parameter μ and the duration of waiting time to conception from ovulatory menstruating state, follows a negative exponential distribution with parameter λ.

iii) Group (b) female moves to the state of conception according to a negative exponential distribution with parameter λ.

The probability density function of X is given by;

$$f_4(x) = \theta f(x_1 + x_2) + (1-\theta) f(x_2) \ ; \quad 0 \le \theta \le 1 \ , \quad \dots (2.7)$$

where X_1 is the waiting time required for a female to move to the state of ovulation from the adolescent sterile state and X_2 is the waiting time for a female to move to the state of first conception from the start of ovulation state. By solving the above equation, we get;

$$f_4(x) = \frac{\theta\mu\lambda}{\lambda - \mu}\left(e^{-\mu x} - e^{-\lambda x}\right) + (1-\theta)\lambda e^{-\lambda x}; \quad 0 \le \theta \le 1, \qquad \mu > 0, \qquad \lambda > 0$$

$$= \alpha\mu e^{-\mu x} + \beta\lambda e^{-\lambda x} \ ; \qquad \dots (2.8)$$

where $\alpha = \frac{\theta\lambda}{\lambda-\mu}$ and $\beta = 1 - \alpha$.

The procedures of the estimation of the parameters involved in all the four models are briefly described in the following section.

3. Estimation of the parameters of the models:

In Model I, the maximum likelihood estimator as well as moment estimator of λ is $\frac{1}{\bar{X}}$, where $\bar{X}$ is sample mean. In this paper, the method of maximum likelihood (M.L.) is being proposed for estimation of parameters involved in Models II, III and IV. The moment estimators of the parameters involved in Model II have disadvantage as the moment estimator does not exist for $\alpha \le 2$. Hence, maximum likelihood (M.L.) estimator for the parameters of this continuous time model for first conception is preferable. It may be noted that Nath et. al, 1995 proposed the method of moments to estimate the parameters of the model but here the method of M. L. is being proposed. The procedure can be briefly described as below:

Let $x_1, x_2, \dots, x_n$ be a random sample of size n from the population with density function $f_i(x)$; $i = 2, 3, 4$. The logarithm of the likelihood functions for Model II, III and IV are given in Expressions 3.1-3.3.

$$\log_2 L = n log\alpha + n\alpha log\beta - (\alpha+1)\sum_{i=1}^{n} log(x_i + \beta), \ \cdots \ (3.1)$$

$$log_3 L = \sum_{i=1}^{n} \log(a + bx_i) - a\sum_{i=1}^{n} x_i - b\sum_{i=1}^{n} \frac{x_i^2}{2} \ , \qquad \dots \qquad (3.2)$$

$$log_4 L = \sum_{i=1}^{2} \log\left(\alpha\mu e^{-\mu x_i} + \beta\lambda e^{-\lambda x_i}\right) \ , \qquad \dots \ (3.3)$$

The above likelihoods are used to estimate the parameters of the models when marital duration is infinite. But in this study specific finite marital duration say, T, is considered, hence the M.L. estimates of the parameters are obtained by fitting the truncated form of the distributions and the form of the truncated distributions are as follows:

$$f_2^*(x) = \frac{f_2(x)}{F_2(T)} \ ; 0 \le x \le T, \ \dots \ (3.4)$$

$$f_3^*(x) = \frac{f_3(x)}{F_3(T)} \ ; 0 \le x \le T, \quad \ldots \quad (3.5)$$

$$f_4^*(x) = \frac{f_4(x)}{F_4(T)} \ ; 0 \le x \le T, \quad \ldots \quad (3.6)$$

Now, the log likelihood functions of the truncated distributions having respective densities $f_i^*(x); i = 2, 3, 4$ are as follows:

$$log_2^* L = \ nlog\alpha \ + \ n\alpha log\beta - (\alpha + 1)\sum_{i=1}^{n} \log(x_i \ + \ \beta) - nlog(1 - \frac{\beta^\alpha}{(T + \beta)^\alpha}), \quad \cdots \qquad (3.7)$$

$$log_3^* L = \ \sum_{i=1}^{n} \log(a + bx_i) - a\sum_{i=1}^{n} x_i - b\sum_{i=1}^{n} \frac{x_i^2}{2} - \ nlog(1 - e^{-\left(aT + b\frac{T^2}{2}\right)}) \ , \quad \ldots \ (3.8)$$

$$log_4^* L = \sum_{i=1}^{n} \log\left(\alpha\mu e^{-\mu x_i} + \beta\lambda e^{-\lambda x_i}\right) - nlog(1 - \ \alpha e^{-\mu T} - \beta e^{-\lambda T}) \ , \qquad \ldots \ (3.9)$$

The M. L. estimates of the parameters $\hat{\alpha}$ and $\hat{\beta}$ involved in Model II are obtained from the Equations 3.10-3.11 given below

$$S_{21} = \ \frac{\partial logL}{\partial \alpha} \ = \ 0 \ ; \ \cdots \quad (3.10)$$

$$S_{22} = \ \frac{\partial logL}{\partial \beta} \ = \ 0 \ ; \ \cdots \quad (3.11)$$

Solving the above equations, we obtain

$$\frac{n}{\alpha} + \ nlog\beta - \sum_{i=1}^{n} \log(x_i + \beta) = n\frac{\partial}{\partial \alpha} log\left(1 - \frac{\beta^\alpha}{(T + \beta)^\alpha}\right) \ ; \ \cdots \ (3.12)$$

$$\frac{n\alpha}{\beta} - (\alpha + 1)\sum_{i=1}^{n} \frac{1}{(x_i + \beta)} = n\frac{\partial}{\partial \alpha} log\left(1 - \frac{\beta^\alpha}{(T + \beta)^\alpha}\right) \ ; \quad \cdots \quad (3.13)$$

The M. L. estimates of the parameters a and b of Model III are obtained from the following equations:

$$S_{31} = \frac{\partial logL}{\partial a} = 0 \ , \ldots \ (3.14)$$

$$S_{32} = \frac{\partial logL}{\partial b} = 0 \ , \ldots \ (3.15)$$

i.e.;

$$\sum_{i=1}^{n}\frac{1}{(a+bx_i)}=\sum_{i=1}^{n}x_i+\frac{nTe^{-aT}}{1-e^{-\left(aT+b\frac{T^2}{2}\right)}};\ \ldots(3.16)$$

$$\sum_{i=1}^{n}\frac{x_i}{(a+bx_i)}=\sum_{i=1}^{n}\frac{x_i^2}{2}+\frac{nT^2e^{-b\frac{T^2}{2}}}{2(1-e^{-\left(aT+b\frac{T^2}{2}\right)})};\ \ldots(3.17)$$

Again, the estimates of the parameters θ, μ and λ involved in Model IV are obtained by solving the following equations:

$$S_{31}=\frac{\partial logL}{\partial\theta}=0\,,\ \ldots\ (3.18)$$

$$S_{32}=\frac{\partial logL}{\partial\mu}=0\,,\ \ldots\ (3.19)$$

$$S_{33}=\frac{\partial logL}{\partial\lambda}=0\,,\ \ldots\ (3.20)$$

i.e.;

$$\sum_{i=1}^{n}\left(\frac{\mu e^{-\mu x_i}-\lambda e^{-\lambda x_i}}{\alpha\mu e^{-\mu x_i}+\beta\lambda e^{-\lambda x_i}}\right)=\frac{n(e^{-\lambda T}-e^{-\mu T})}{1-\ \alpha e^{-\mu T}-\beta e^{-\lambda T}};\ \ldots \qquad (3.21)$$

$$\sum_{i=1}^{n}\left(\frac{\theta\lambda(\mu-\lambda)x_i e^{-\mu x_i}}{\alpha\mu e^{-\mu x_i}+\beta\lambda e^{-\lambda x_i}}\right)=\frac{n\{(\lambda-\mu)\theta\lambda Te^{-\mu T}+\theta\lambda\left(e^{-\lambda T}-e^{-\mu T}\right)\}}{1-\ \alpha e^{-\mu T}-\beta e^{-\lambda T}};\ \ldots \qquad (3.22)$$

$$\sum_{i=1}^{n}\left(\frac{[\mu^2-\theta\lambda^2-2\lambda\mu+2\theta\lambda\mu-\{\lambda^3(1-\theta)+\mu^2\lambda+\theta\mu\lambda^2\}x_i]e^{-\lambda x_i}-\theta\mu^2e^{-\mu x_i}}{\alpha\mu e^{-\mu x_i}+\beta\lambda e^{-\lambda x_i}}\right)$$

$$=\frac{n\{(\lambda-\mu-\theta\lambda)T(\lambda-\mu)e^{-\lambda T}+\theta\mu\left(e^{-\mu T}-e^{-\lambda T}\right)\}}{1-\ \alpha e^{-\mu T}-\beta e^{-\lambda T}};\ \ldots \qquad (3.23)$$

These sets of expressions (Expressions 3.12-3.13, 3.16-3.17 and 3.21-3.23) are quite complicated and no explicit solution exists. However, M.L. estimates of the parameters are computed by Newton Raphson method taking certain guess values of the parameters through the expression of log likelihood function using R-software.

4. Application:

The data utilized for the present study have been taken from National Family Health Survey III (NFHS-III). NFHS provides the data on marriage to FBI (in months), age at marriage (in years), date of marriage (in CMC) and date of the survey (in CMC), etc. The data on waiting time to first conception are obtained by subtracting nine months from the interval from marriage to first birth assuming one to one correspondence between conception and live birth.

Here, only those females are considered whose marriage took place at least seven years prior to the reference date of the survey. This is done to take account the truncation effect, as discussed in Sheps et. al (1970). It is well known that conception outside wedlock is generally not accepted in Indian society; hence the negative FBI durations are excluded from the study. Again, the FBI durations of less than nine months are also excluded from the study since in this study the gestation period is taken as nine months. Females are divided into two groups according to their age at marriage (lower ages at marriage (<16 years) & higher ages at marriage (>=16 years)). Models are then applied to the data of first conception for different major states viz., Bihar, Uttar Pradesh, Madhya Pradesh, Rajasthan, Assam, Orissa, Maharashtra, West Bengal, Andhra Pradesh, Karnataka, Kerala and Tamil Nadu situated in different regions of the country. Models I and II are applied to the data of first conception for the females whose age at marriage >=16 years, whereas, Model III and IV are applied to the data of first conception for the females whose age at marriage is <16 years.

5. Discussion & Conclusions:

Tables 1-6 present the observed and expected frequencies of waiting time to first conception, the estimated parameters and the respective chi-square values as well as AIC values under Model I and Model II for females of higher ages at marriage, i.e. for the females whose age at marriage is >=16 years. From the tables, it is observed that Model I is not fitting well in most of the considered states while Model II seems to give reasonably good fit to all the data sets. This gives a clue that populations of the females of different states are not homogeneous within themselves. Thus, the assumption of heterogeneity with respect to conception rate seems to be more appropriate for describing the phenomenon of time of first conception for females of higher ages at marriage.

However, one natural question arises; “Why is Model I fitting well to some of the states?” The answer of this question may be that females of those states may have same social status and each of them may have gone through the similar socio-cultural norms and taboos after marriage i.e., population is more or less homogeneous.

Model II explains the variation in conception rate among the females in different states of the country. It has been observed that females of the states viz., Assam, Maharashtra, Andhra Pradesh, Karnataka, Kerala, and Tamil Nadu are more heterogeneous than the rest considered parts (see Fig 1). It may be due to the socio-cultural variations between the states. Thus, with the help of Model II, it has become possible to observe the heterogeneity in the population.

Tables 7-11 present the observed and expected frequencies of waiting time to first conception, the estimated parameters and the respective value of chi-square as well as Akaike Information Criterion (AIC and AIC corrected) values under Model III and Model IV for females of lower ages at marriage. These tables show that there is no significant difference between the chi-square values of Model III and Model IV, both are fitting quite well. Model III is derived on the assumption that conception rate is a linear function of time t. As it is mentioned earlier that FBI signifies couple's fertility at early stages of married life and in traditional society like India, the early part of married life is governed by large number of socio-cultural norms and taboos. These social norms and taboos decrease with the passage of time of marriage. Along with these social norms and taboos, when the age at marriage is low, there is one most important biological factor that influences this duration variable, is referred to as adolescent sterility. Model III indirectly incorporates all these factors together by considering conception rate as time dependent.

Model IV is based on the assumption that the females of lower ages at marriage are a mixture of two groups: the adolescent sterile group and the ovulation group. In case of lower ages at marriage, the most important chance mechanism which influences the fertility behavior of female is adolescent sterility. It is worthwhile to mention that the extent of all social norms and taboos can be reduced by one's effort but the extent of adolescent sterility cannot be reduced by one’s effort. Model IV ascertains the extent of adolescent sterility. May be due to this possible reason, Model IV performs well than Model III when both of them are applied to the real data. But the beauty of Model III is that in addition to its mathematical simplicity; it explains the phenomena very well.

In Model III, an indirect approximation of adolescent sterility is obtained through the parameter a, whereas, in Model IV, proportion of adolescent sterility at the time of marriage is estimated through the parameter θ. The smaller values of a may be attributed to adolescent sub-fecundability, strict traditional coitus regulation, etc., whereas the higher values of b may be responsible for the attainment of fecundable state, gradual withdrawal of sexual restrictions, etc., with the passage of time. It is observed that the value of α is smaller for the states Bihar, Uttar Pradesh, Madhya Pradesh and Rajasthan whereas the value of θ is higher for these states. Again, with the help of estimated parameters μ and λ, one can get an idea about the average time, a female will take to reach at ovulation state from adolescent sterile state and to have her first conception after having exposed to the risk of conception. Model IV is complicated as compared to Model III but it provides more inherent information from the phenomenon.

Here, all the four models have been applied on various data sets of waiting time to first conception and on the basis of the values of chi-square, it can be concluded that whether a model is appropriate or not. Appropriateness of a model depends upon the assumptions under consideration. If a model is appropriate for a process, then on an average it will fit the data of that process. Sometimes, there may be many models for the same phenomenon and they fit the data well. But, it does not mean that all the models are correct. Therefore, one must have logical interpretation of the parameters involved in the probability models. Consequently these parameters are used as alternative measures of various aspects of the process under consideration. These estimates can be easily compared so that valid and informative conclusions can be drawn.

Thus, it may be said that probability models are very useful and play an important role for explaining the real phenomenon. The variability, uncertainty and complexity of the phenomenon under study can be deeply understood with the help of the models and on the basis of these probability models, decision makings are validated. Some unobserved characteristics can be estimated with the applications of the models to the real data. For example; the extent of heterogeneity as well as homogeneity in the population, how conception rate varies over time and the proportion of adolescent sterility can be estimated. These findings may be helpful for policy makers to frame appropriate policies and their implementation.

6. **Tables & Figure:**

Table 1: Observed and expected frequencies of waiting time to conception for Bihar and Uttar Pradesh under Model I and Model II for females of higher ages at marriage (>=16 years)

C. I.* (in years)	States					
	Bihar			Uttar Pradesh		
	Observed	Model I	Model II	Observed	Model I	Model II
0-1	334	332.9	335.7	1446	1426.5	1436.3
1-2	207	195.5	193.7	824	801.7	795.1
2-3	106	114.8	113.0	399	450.6	444.5
3-4	70	67.4	66.6	271	253.2	250.8
4-5	32	39.6	39.7	147	142.3	142.8
5-6	25	23.2	23.8	73	89.9	82.1
6-7	13	13.6	14.5	39	44.9	47.5
Total	787	787.0	787.0	3199	3199.0	3199.0
$\chi^2_{(cal)}$		$\chi^2_5 = 3.068$	$\chi^2_4 = 3.210$		$\chi^2_5 = 9.583$	$\chi^2_4 = 10.049$
Estimated parameters		$\hat{\lambda} = 0.532$	$\hat{\alpha} = 26.182$ $\hat{\beta} = 48.472$ $E(\lambda) = 0.540$ $V(\lambda) = 0.011$		$\hat{\lambda} = 0.576$	$\hat{\alpha} = 34.363$ $\hat{\beta} = 58.854$ $E(\lambda) = 0.584$ $V(\lambda) = 0.010$
AIC		2385.066	2386.900		9346.945	9348.513
AICc		2385.063	2386.884		9346.940	9348.509

* Class Interval denoting waiting time to first conception

Table 2: Observed and expected frequencies of waiting time to conception for Madhya Pradesh and Rajasthan under Model I and Model II for females of higher ages at marriage (>=16 years)

C. I. (in years)	States					
	Madhya Pradesh			Rajasthan		
	Observed	Model I	Model II	Observed	Model I	Model II
0-1	864	908.2	918.5	386	426.0	430.9
1-2	591	509.2	508.6	312	264.0	264.3
2-3	284	285.4	282.4	191	163.6	162.4
3-4	148	160.0	157.2	87	101.4	100.0
4-5	89	89.7	87.7	51	62.9	61.7
5-6	33	50.3	49.1	28	39.0	38.1
6-7	22	28.2	27.5	26	24.1	23.6
Total	2031	2031.0	2031.0	1081	1081.0	1081.0
$\chi^2_{(cal)}$		$\chi^2_5 = 23.529$	$\chi^2_4 = 23.530$		$\chi^2_5 = 24.559$	$\chi^2_4 = 24.779$
Estimated parameters		$\hat{\lambda} = 0.579$	$\hat{\alpha} = 130.649$ $\hat{\beta} = 221.773$ $E(\lambda) = 0.589$ $V(\lambda) = 0.003$		$\hat{\lambda} = 0.478$	$\hat{\alpha} = 126.965$ $\hat{\beta} = 260.841$ $E(\lambda) = 0.487$ $V(\lambda) = 0.002$
AIC		5922.795	5926.361		3417.122	3419.769
AICc		5922.794	5926.355		3417.120	3419.758

Table 3 Observed and expected frequencies of waiting time to conception for Assam and Orissa under Model I and Model II for females of higher ages at marriage (>=16 years)

C. I. (in years)	States					
	Assam			Orissa		
	Observed	Model I	Model II	Observed	Model I	Model II
0-1	569	562.3	335.7	672	648.9	665.1
1-2	246	243.9	193.7	328	329.8	316.9
2-3	98	105.8	113.0	149	167.6	158.5
3-4	34	45.9	66.6	75	85.2	82.8
4-5	21	19.9	39.7	49	43.3	44.9
5-6	16	8.6	23.8	19	22.0	25.2
6-7	6	3.7	14.5	16	11.2	14.6
Total	990	990.0	787.0	1308	1308.0	1308.0
$\chi^2_{(cal)}$		$\chi^2_5 = 11.464$	$\chi^2_4 = 7.526$		$\chi^2_5 = 7.353$	$\chi^2_4 = 3.810$
Estimated parameters		$\hat{\lambda} = 0.835$	$\hat{\alpha} = 5.300$ $\hat{\beta} = 5.562$ $E(\lambda) = 0.877$ $V(\lambda) = 0.985$		$\hat{\lambda} = 0.677$	$\hat{\alpha} = 9.532$ $\hat{\beta} = 13.278$ $E(\lambda) = 0.718$ $V(\lambda) = 0.054$
AIC		2298.851	2289.507		3506.633	3505.111
AICc		2298.849	2289.495		3506.631	3505.102

Table 4: Observed and expected frequencies of waiting time to conception for Maharashtra and West Bengal under Model I and Model II for females of higher ages at marriage (>=16 years)

C. I. (in years)	States					
	Maharashtra			West Bengal		
	Observed	Model I	Model II	Observed	Model I	Model II
0-1	1751	1617.0	1732.5	976	955.0	978.4
1-2	606	713.0	595.5	455	453.0	432.6
2-3	230	314.4	263.0	205	214.9	202.6
3-4	152	138.6	135.5	84	101.9	99.7
4-5	84	61.1	77.5	43	48.4	51.2
5-6	33	27.0	47.8	26	22.9	27.4
6-7	27	11.9	31.3	18	10.9	15.1
Total	2883	2883.0	2883.0	1807	1807.0	1807.0
$\chi^2_{(cal)}$		$\chi^2_5 = 80.240$	$\chi^2_4 = 12.260$		$\chi^2_5 = 9.744$	$\chi^2_4 = 5.595$
Estimated parameters		$\hat{\lambda} = 0.835$	$\hat{\alpha} = 2.587$ $\hat{\beta} = 2.494$ $E(\lambda) = 1.037$ $V(\lambda) = 0.416$		$\hat{\lambda} = 0.746$	$\hat{\alpha} = 9.884$ $\hat{\beta} = 12.418$ $E(\lambda) = 0.796$ $V(\lambda) = 0.064$
AIC		6794.443	6704.183		4554.119	4550.142
AICc		6794.442	6704.179		4554.117	4550.135

Table 5: Observed and expected frequencies of waiting time to conception for Andhra Pradesh and Tamil Nadu under Model I and Model II for females of higher ages at marriage (>=16 years)

C. I. (in years)	States					
	Andhra Pradesh			Tamil Nadu		
	Observed	Model I	Model II	Observed	Model I	Model II
0-1	964	926.7	974.1	1547	1406.4	1529.1
1-2	449	454.7	413.2	364	525.1	377.1
2-3	167	223.1	198.4	159	196.0	155.2
3-4	111	109.5	104.5	77	73.2	80.5
4-5	71	53.7	59.1	47	27.3	47.8
5-6	22	26.4	35.4	26	10.2	30.9
6-7	23	12.9	22.2	22	3.8	21.4
Total	1807	1807.0	1807.0	2242	2242.0	2242.0
$\chi^2_{(cal)}$		$\chi^2_5 = 29.822$	$\chi^2_4 = 16.078$		$\chi^2_5 = 196.189$	$\chi^2_4 = 1.740$
Estimated parameters		$\hat{\lambda} = 0.712$	$\hat{\alpha} = 4.143$ $\hat{\beta} = 5.096$ $E(\lambda) = 0.904$ $V(\lambda) = 1.208$		$\hat{\lambda} = 0.985$	$\hat{\alpha} = 1.598$ $\hat{\beta} = 1.054$ $E(\lambda) = 1.515$ $V(\lambda) = 1.437$
AIC		4694.871	4675.915		4516.769	4296.179
AICc		4694.869	4675.908		4516.768	4296.174

Table 6: Observed and expected frequencies of waiting time to conception for Karnataka and Kerala under Model I and Model II for females of higher ages at marriage (>=16 years)

C. I. (in years)	States					
	Karnataka			Kerala		
	Observed	Model I	Model II	Observed	Model I	Model II
0-1	932	885.5	942.0	1051	1010.5	1080.2
1-2	342	387.3	330.0	293	341.9	249.1
2-3	155	169.4	144.1	97	115.7	95.8
3-4	61	74.1	72.6	37	39.2	46.9
4-5	39	32.4	40.5	21	13.3	26.5
5-6	29	14.2	24.3	14	4.5	16.4
6-7	11	6.2	15.4	13	1.5	10.9
Total	1569	1569.0	1569.0	1526	1526.0	1526.0
$\chi^2_{(cal)}$		$\chi^2_5 = 31.858$	$\chi^2_4 = 5.469$		$\chi^2_5 = 123.306$	$\chi^2_4 = 12.534$
Estimated parameters		$\hat{\lambda} = 0.827$	$\hat{\alpha} = 3.040$ $\hat{\beta} = 2.979$ $E(\lambda) = 1.021$ $V(\lambda) = 0.343$		$\hat{\lambda} = 1.083$	$\hat{\alpha} = 1.918$ $\hat{\beta} = 1.186$ $E(\lambda) = 1.617$ $V(\lambda) = 1.363$
AIC		3670.536	3631.328		2796.351	2670.683
AICc		3670.535	3631.321		2796.349	2670.675

Table 7: Observed and expected frequencies of waiting time to conception for Bihar and Uttar Pradesh under Model III and Model IV for females of lower ages at marriage (<16 years)

C. I. (in years)	States					
	Bihar			Uttar Pradesh		
	Observed	Model III	Model IV	Observed	Model III	Model IV
0-1	199	206.0	213.5	474	492.2	504.3
1-2	196	195.1	188.1	432	424.1	407.5
2-3	133	154.3	140.7	308	316.3	291.9
3-4	101	104.8	97.1	184	207.5	195.9
4-5	85	61.9	63.8	127	120.8	126.1
5-6	35	32.1	40.6	82	62.8	78.9
6-7	20	14.7	25.2	46	29.3	48.4
Total	769	769.0	769.0	1653	1653.0	1653.0
$\chi^2_{(cal)}$		$\chi^2_4 = 14.071$	$\chi^2_3 = 10.783$		$\chi^2_4 = 19.402$	$\chi^2_3 = 5.151$
Estimated parameters		$\hat{a} = 0.253$ $\hat{b} = 0.109$	$\hat{\mu} = 0.617$ $\hat{\lambda} = 0.617$ $\hat{\theta} = 0.589$		$\hat{a} = 0.300$ $\hat{b} = 0.097$	$\hat{\mu} = 0.623$ $\hat{\lambda} = 0.624$ $\hat{\theta} = 0.513$
AIC		2703.188	2683.777		5698.941	5647.489
AICc		2703.172	2683.745		5698.948	5647.504

Table 8: Observed and expected frequencies of waiting time to conception for Madhya Pradesh and Rajasthan under Model III and Model IV for females of lower ages at marriage (<16 years)

C. I. (in years)	States					
	Madhya Pradesh			Rajasthan		
	Observed	Model III	Model IV	Observed	Model III	Model IV
0-1	357	394.3	396.6	207	222.0	226.7
1-2	388	353.7	358.7	238	226.1	223.8
2-3	273	263.0	248.3	176	186.5	172.5
3-4	141	166.4	154.4	119	129.5	119.5
4-5	80	90.8	90.7	84	77.2	77.9
5-6	58	43.0	51.6	39	39.8	48.8
6-7	32	17.8	28.7	36	17.9	29.8
Total	1329	1329.0	787.0	899	899.0	899.0
$\chi^2_{(cal)}$		$\chi^2_4 = 28.942$	$\chi^2_3 = 12.415$		$\chi^2_4 = 21.862$	$\chi^2_3 = 6.423$
Estimated parameters		$\hat{a} = 0.289$ $\hat{b} = 0.120$	$\hat{\mu} = 0.635$ $\hat{\lambda} = 0.938$ $\hat{\theta} = 0.617$		$\hat{a} = 0.220$ $\hat{b} = 0.119$	$\hat{\mu} = 0.644$ $\hat{\lambda} = 0.645$ $\hat{\theta} = 0.693$
AIC		4457.000	4426.812		3191.852	3163.038
AICc		4457.009	4426.830		3191.838	3163.012

Table 9: Observed and expected frequencies of waiting time to conception for Maharashtra and West Bengal under Model III and Model IV for females of lower ages at marriage (<16 years)

C. I. (in years)	States					
	Maharashtra			West Bengal		
	Observed	Model III	Model IV	Observed	Model III	Model IV
0-1	418	430.1	435.7	379	381.5	389.7
1-2	295	283.7	273.6	236	249.8	236.4
2-3	138	171.9	163.2	141	152.0	141.8
3-4	118	95.9	94.2	91	86.1	84.3
4-5	57	49.4	53.0	60	45.6	49.7
5-6	23	23.6	29.3	22	22.5	29.1
6-7	16	10.4	16.0	19	10.4	17.0
Total	1065	1065.0	1065.0	948	948.0	948.0
$\chi^2_{(cal)}$		$\chi^2_4 = 16.745$	$\chi^2_3 = 14.003$		$\chi^2_4 = 13.507$	$\chi^2_3 = 4.960$
Estimated parameters		$\hat{a} = 0.502$ $\hat{b} = 0.031$	$\hat{\mu} = 0.717$ $\hat{\lambda} = 0.718$ $\hat{\theta} = 0.314$		$\hat{a} = 0.479$ $\hat{b} = 0.062$	$\hat{\mu} = 0.612$ $\hat{\lambda} = 0.612$ $\hat{\theta} = 0.170$
AIC		3255.406	3238.886		3053.203	2914.378
AICc		3255.394	3238.863		3053.191	2914.352

Table 10: Observed and expected frequencies of waiting time to conception for Andhra Pradesh and Karnataka under Model III and Model IV for females of lower ages at marriage (<16 years)

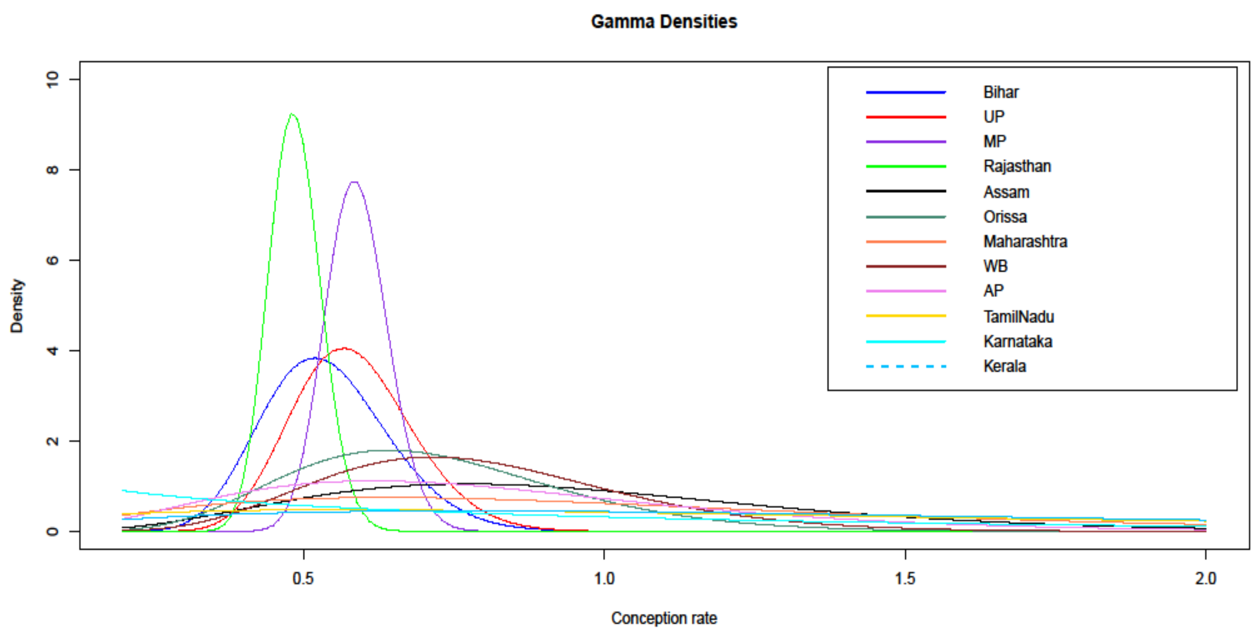
C. I. (in years)	States					
	Andhra Pradesh			Karnataka		
	Observed	Model III	Model IV	Observed	Model III	Model IV
0-1	500	526.2	536.6	389	418.6	427.9
1-2	408	384.4	373.6	256	244.2	229.0
2-3	224	254.2	233.6	136	138.1	126.0
3-4	134	153.0	143.9	74	75.7	72.2
4-5	108	84.2	88.5	37	40.3	43.8
5-6	63	42.4	54.4	35	20.8	28.7
6-7	27	19.6	33.4	21	10.4	20.4
Total	1464	1464.0	1464.0	948	948.0	948.0
$\chi^2_{(cal)}$		$\chi^2_4 = 28.186$	$\chi^2_3 = 13.654$		$\chi^2_4 = 23.618$	$\chi^2_3 = 10.032$
Estimated parameters		$\hat{a} = 0.402$ $\hat{b} = 0.077$	$\hat{\mu} = 0.487$ $\hat{\lambda} = 2.390$ $\hat{\theta} = 0.237$		$\hat{a} = 0.561$ $\hat{b} = 0.029$	$\hat{\mu} = 0.665$ $\hat{\lambda} = 0.095$ $\hat{\theta} = 0.198$
AIC		4754.744	4716.744		3409.211	3374.752
AICc		4754.735	4716.727		3409.197	3374.725

Table 11: Observed and expected frequencies of waiting time to conception for Kerala and Tamil Nadu under Model III and Model IV for females of lower ages at marriage (<16 years)

C. I. (in years)	States					
	Kerala			Tamil Nadu		
	Observed	Model III	Model IV	Observed	Model III	Model IV
0-1	82	90.4	92.8	350	357.1	365.0
1-2	60	52.5	48.6	151	149.4	137.8
2-3	28	29.6	26.7	60	64.8	57.6
3-4	15	16.2	15.6	25	29.2	28.1
4-5				15	13.6	16.4
5-6	} 19	15.3	20.3	15	6.6	11.0
6-7				8	3.3	8.1
Total	204	787.0	204.0	624	624.0	624.0
$\chi^2_{(cal)}$		$\chi^2_2 = 2.906$	$\chi^2_3 = 4.062$		$\chi^2_4 = 18.798$	$\chi^2_3 = 3.902$
Estimated parameters		$\hat{a} = 0.564$ $\hat{b} = 0.027$	$\hat{\mu} = 0.755$ $\hat{\lambda} = 0.259$ $\hat{\theta} = 0.221$		$\hat{a} = 0.860$ $\hat{b} = -0.039$	$\hat{\mu} = 1.076$ $\hat{\lambda} = 0.225$ $\hat{\theta} = 0.175$
AIC		615.401	610.186		1513.673	1497.314
AICc		615.341	610.066		1513.654	1497.276

$\boldsymbol{\chi^2_{tab} = 3.841\ (1\ d.f.), 5.991\ (2\ d.f.), 7.815\ (3\ d.f.), 9.488\ (4\ d.f.), 11.070\ (5\ d.f.)}$

Figure 1: Graph showing the heterogeneity in conception rate of female in some states of India:



References:

Bhattacharya, B.N., Pandey, C.M., Singh, K.K.; Model for first birth interval and some social factors. *Journal of Mathematical Biosciences*, 92,17-28 (1988)

Gondotra, M.M., Das, N.; Age at menarche in an Indian population. *Health and Population-Perspectives and Issues*, 5(3), 168-181 (1982)

Nair, N.U.; On a distribution of first conception delays in the presence of adolescent sterility. *Demography India*, 12, 209 (1983)

Nair, N.U.; A stochastic model for estimating adolescent sterility. *Biometrical Journal*, 25(5) (1983)

Nath, D.C., Land, K.C., Singh, K.K.; A waiting time distribution for the first conception and its application to a non-contracepting traditional society. *Genus* 51(1/2), 95{103 (1995)

Pathak, K.B.; An extension of the waiting time distribution of first conception. *Journal of Biosciences*, 10, 231{234 (1978)

Pathak, K.B., Pandey, A.; Analytical model of human fertility with provision for adolescent sterility. *Health and Population-Perspectives and Issues*, 7, 171{180 (1981)

Pathak, K.B., Prasad, C.V.S.; A model for estimating adolescent among married women. *Demography* 14, 103-104 (1977)

Pratap, M.; A study on fertility changes through birth interval approach. *Ph.D. thesis*, Banaras Hindu University (2011)

Sheps, M.C., Menken, J.A., Ridley, J.C., Lingler, J.W.; Truncation effect in closed and open birth interval data. *Journal of the American Statistical Association*, 65(330), 678-693 (1970)