INDIAN JOURNAL OF POPULATION AND DEVELOPMENT
Volume 1(2), July 2021: 209-224

A Model Based Statistical Investigation of Female Age at Marriage

Brijesh P Singh
Utpal Dhar Das
Sandeep Singh

Abstract

Age at marriage is a significant event in the life of every individual. In the Indian society, sexual activity outside the institution of marriage is a social taboo and is not accepted. Therefore, age at marriage is of an importance event due to its direct relationship with the sexual activity, family building or the level of fertility. In the present paper, a probability model has been developed to describe the distribution of the age at marriage of females in India. The model has been validated by comparing it with the actual data on female age at marriage from the National Family Health Survey. The model explains the variability in the age at marriage satisfactorily.

Introduction

The sexual activity and family building strategy in traditional societies is a very sensitive issue and discussing about sexual matters openly is almost restricted and considered as a social taboo. During past few decades, due to the increase in the level of education, common awareness and modernisation of the society, there has been an attitudinal shift in the perception about sex, especially, among the young generation. The knowledge of sex and reproductive health is also limited among both educated and uneducated females as very few rural young females could describe how body change is related to sexual intercourse and reproduction (Vlassof, 1987).

In societies where childbearing outside the institution of marriage is not socially acceptable, delay in marriage contributes significantly towards reduction in fertility by shortening the reproductive span. On the other hand, marriage at a young age is typically associated with a lower education and with lower social status also. The age at marriage varies by various socio-demographic factors such as religion/caste, geographical location, place of residence, type of family and level of income. In, India, marriage signals the beginning of the family building process. Age at marriage is also an important indicator of women's status because of its role in determining fertility level and improving health of women and children.

There are many studies on the differentials and determinants of age at marriage in India (Agarwal, 1962; Talwar, 1967; Krishnan, 1971; Malaker, 1987). In addition, probability models have been used to describe the distribution of females according to their age at marriage. Nydell (1924) has used the lognormal distribution to explain the age at marriage. Hyrenius and others (1967) have proposed logistic curve to graduate the number of females at different age at marriage. Coale (1971) has shown that the risk of first marriage for the standard population is very closely fitted by a double exponential function while Coale and McNeil (1972) have shown that the distribution of females by the age at marriage can be approximated very closely by the limiting distribution of the convolution of an infinite number of (normalised) exponentially distributed components, and that an equally good fit is obtained by the convolution of a finite number of the exponential components plus an additional Normal distribution. Malaker (1985) has observed that Coale and McNeil model failed to describe the Indian data. Hernes (1972) has fitted the Gompertz curve for US female age at marriage. According to Henry (1972), mates are not selected from country level marriage market, however, there exists a smaller circle of mates with homogeneous characteristics. McFarland (1972, 1975) has given "Iterative Adjustment Model" for marriage formation. Both Henry (1972) and McFarland (1972) have offered numerical procedures to be applied on the matrix of marriage data, but they have not given any algebraic form. Hill (1977) has used a simple polynomial to describe the marriage data. Rodriquez and Trussell (1980) have used a linear function of logarithm of standard Gamma function. Islam (1984) suggests modified exponential curve to describe the age pattern of proportion of never married females. Singh and others (1986) have refined the model proposed by Mishra (1979) and applied the refined model to describe the marriage data of Varanasi females. Diekman (1989) has used two-parameter log-logistic distribution for describing the US and German data. Mitter (1989) has used a convolution of normal and exponential distributions to understand the pattern of age at marriage. Nath and Talukdar (1992) have used Type-I extreme value distribution to describe the female age at marriage data from Assam.

Most of the models of describing the distribution of females by age at marriage discussed above are complex and contain many parameters. They have also not been found to be universally applicable. Verma and Pathak (2001) used a model for age at marriage, which was used for estimating the adolescent sterility among married females by Pathak and Prashad (1978). The model gives poor fit particularly for higher age at marriage due to the age heaping in the data. The age misreporting and digit preference complicate the model for describing the distribution of the age at marriage. To problem has been resolved to some extent by combining two consecutive age groups. Singh and others (2004) have proposed a discrete probability model (displaced negative binomial distribution) and have found that the model is suitable for describing the distribution of the age at marriage of females. Singh and others (2015) have discussed an alternative procedure to estimate the parameters of the above model. Alho (2016), on the other hand, has proposed a model in the stochastic form that has been defined in terms of rates and continuous time. On the other hand, the models proposed by Matthews and Garenne (2013a; 2013b) are deterministic, described in terms of counts and discrete time.

This paper has two objectives. The first objective is to analyse the pattern of female age at marriage in India. The second objective of the paper is to develop a suitable probability model that describes the distribution of the females by their age at marriage. The probability model so developed has been fitted using the data on female age at marriage available through the National Family Health Survey to examine the validity of the model.

Probability Distribution of the Age at Marriage

The female age at marriage is a continuous variable and it occurs after a certain age after the birth of the female. We assume that the age at marriage of a female is an independent and identically distributed (iid) random variable which follows an exponential distribution with a lower and an upper bound. There is very small probability of females marrying below a certain age. Similarly, there is very small probability of females marrying after a certain age so that these very small probabilities can be neglected. It is also well-known that the proportion of females marrying after a certain age decreases very rapidly after a certain age. Given these assumptions and considerations, the probability distribution of the female age at marriage may be modelled as

$$F(x) = 1 - exp\left[\frac{-\lambda(x-\theta_1)}{(\theta_2-x)}\right]; \theta_1 \leq x \leq \theta_2; \lambda > 0 \qquad (1)$$

Where ϑ_1 is the lower bound and ϑ_2 is the upper bound of the female age at marriage.

The hazard rate function of model (1) is given by

$$h(x) = \frac{d}{dx}\{-\,ln[\,1 - F(x)]\} \text{ therefore } h(x) = \frac{\lambda(\theta_2-\theta_1)}{(\theta_2-x)^2} \qquad (2)$$

The hazard rate is an increasing function of x during the interval, $\vartheta_1 \leq x \leq \vartheta_2$ which means that the proportion of females getting married at higher ages is small. The probability density function of the model (1) is given by

$$f(x) = \frac{\lambda(\theta_2-\theta_1)}{(\theta_2-x)^2}\left[\exp\left\{\frac{-\lambda(x-\theta_1)}{(\theta_2-x)}\right\}\right]; \theta_1 \leq x \leq \theta_2; \lambda > 0 \qquad (3)$$

The r^{th} moment of the distribution (1) is given by

$$E[X^r] = \int_{\theta_1}^{\theta_2} x^r \frac{\lambda(\theta_2-\theta_1)}{(\theta_2-x)^2} e^{\frac{-\lambda(x-\theta_1)}{(\theta_2-x)}} dx \qquad (4)$$

Let

$$\frac{\lambda(x-\theta_1)}{(\theta_2-x)} = z \Rightarrow \frac{\lambda(\theta_2-\theta_1)}{(\theta_2-x)^2}dx = dz$$

It is obvious that if $x\epsilon$ $(\vartheta_1, \vartheta_2)$ then z ϵ $(0,\infty)$.

Also, we have

$\frac{x-\theta_1}{\theta_2-x} = \frac{z}{\lambda} \Rightarrow x = \frac{z\theta_2+\lambda\theta_1}{z+\lambda}$, therefore,

$$E[X^r] = \int_0^\infty \left(\frac{z\theta_2+\lambda\theta_1}{z+\lambda}\right)^r \frac{\lambda(\theta_2-\theta_1)}{(\theta_2-x)^2} e^{\frac{-\lambda(x-\theta_1)}{(\theta_2-x)}} dx \tag{5}$$

$$\Rightarrow \int_0^\infty \left(\frac{z\theta_2+\lambda\theta_1}{z+\lambda}\right)^r e^{-z}\,dz = \int_0^\infty \left(\theta_2 + \frac{\lambda(\theta_1-\theta_2)}{z+\lambda}\right)^r e^{-z}\,dz$$

$$\Rightarrow \int_0^\infty \sum_{k=0}^r \binom{r}{k} \theta_2^{r-k} \left(\frac{\lambda(\theta_1-\theta_2)}{z+\lambda}\right)^k e^{-z}\,dz = \sum_{k=0}^r \binom{r}{k} \theta_2^{r-k} [\lambda(\theta_1-\theta_2)]^k \int_0^\infty \frac{e^{-z}}{(z+\lambda)^k} dz \tag{6}$$

Let $z+\lambda=l$ so that $z=l-\lambda$ and $dz=dl$. Now

$$\int_0^\infty \frac{e^{-z}}{(z+\lambda)^k} = \int_\lambda^\infty \frac{e^{-(l-\lambda)}}{l^k} dl = e^\lambda \int_\lambda^\infty \frac{e^{-l}}{l^k} dl \tag{7}$$

If $l = \lambda p$, then $\Rightarrow dl = \lambda dp$, and for $l = \lambda, p = 1$also for $l = \infty, p = \infty$. Therefore, equation (7) can be written as

$$\Rightarrow e^\lambda \int_1^\infty \frac{e^{-\lambda p}}{(p\lambda)^k} \lambda dp \Rightarrow \frac{e^\lambda}{\lambda^{k-1}} \int_1^\infty \frac{e^{-\lambda p}}{p^k} dp = \frac{e^\lambda}{\lambda^{k-1}} E_k(\lambda)$$

Hence

$$E[X^r] = \sum_{k=0}^r \binom{r}{k} \theta_2^{r-k} [\lambda(\theta_1-\theta_2)]^k \frac{e^\lambda}{\lambda^{k-1}} E_k(\lambda) \tag{8}$$

We know that

$$E_0(\lambda) = \frac{e^{-\lambda}}{\lambda} \tag{9}$$

and it can be easily obtained. However, $E_1(\lambda)$, $E_2(\lambda)$, etc. require extensive computations which cannot be done manually.

Putting $r=1$ in equation (8), we have first moment or arithmetic mean,

$$E[X] = \theta_2 + \lambda(\theta_1-\theta_2)e^\lambda E_1(\lambda) \tag{10}$$

Putting $r=2$ in equation (8), we get the second moment

$$E[X^2] = \theta_2^2 + 2\theta_2[\lambda(\theta_1-\theta_2)]e^\lambda E_1(\lambda) + (\theta_1-\theta_2)^2 \lambda e^\lambda E_2(\lambda) \tag{11}$$

Using (10) and (11), we can get the variance.

The parameters of the model can be obtained through the method of maximum likelihood. The likelihood function of the model is given by

$$L = \prod_{i=1}^{n} \frac{\lambda(\theta_2-\theta_1)}{(\theta_2-x_i)^2}\left[\exp\left\{\frac{-\lambda(x_i-\theta_1)}{(\theta_2-x_i)}\right\}\right] \tag{12}$$

$$= \exp\left[-\lambda \sum_{i=1}^{n}\left(\frac{x_i-\theta_1}{\theta_2-x_i}\right)\right]\lambda^n(\theta_2-\theta_1)^n \prod_{i=1}^{n}\frac{1}{(\theta_2-x_i)^2} \tag{13}$$

and the log likelihood is

$$LL = -\lambda \sum_{i=1}^{n}\left(\frac{x_i-\theta_1}{\theta_2-x_i}\right) + n\,ln\,\lambda + n\,ln(\theta_2-\theta_1) + \sum_{i=1}^{n} ln\left[\frac{1}{(\theta_2-x_i)^2}\right] \tag{14}$$

Now, differentiating with respect to λ we get

$$\frac{\partial\, ln\,\lambda}{\partial\lambda} = -\sum_{i=1}^{n}\left(\frac{x_i-\theta_1}{\theta_2-x_i}\right) + \frac{n}{\lambda} = 0 \tag{15}$$

$$\Rightarrow \frac{n}{\lambda} = \sum_{i=1}^{n}\left(\frac{x_i-\theta_1}{\theta_2-x_i}\right) \Rightarrow \hat{\lambda} = \frac{n}{\sum_{i=1}^{n}\left(\frac{x_i-\theta_1}{\theta_2-x_i}\right)} \tag{16}$$

The application of the model requires setting up lower bound (ϑ_1) and upper bound (ϑ_2) of the female age at marriage. In case of female age at marriage less than the lower bound the model will give a negative estimate of λ. Similarly, if the upper bound is lower than the maximum female age at marriage, the model will give a negative estimate of λ. It is, therefore, necessary to fix the lower and upper bounds of the female age at marriage. We assume that the lower bound of the model is the female age at menarche. On the other hand, the upper bound of the model can be set to be equal to or more than the maximum reported female age at marriage.

The goodness of fit of the model can be tested through the K-S test of goodness of fit which is a nonparametric test of the equality of one-dimensional probability distributions (Kolmogorov, 1933; Smirnov, 1933). This test is based on the empirical cumulative distribution function (ECDF). Given N ordered data points $X_1, X_2, \ldots, X_N$, the ECDF is defined as

$$E_N = \frac{n(i)}{N} \tag{17}$$

where *n*(*i*) is the number of points less than X_i and X_i are ordered from the smallest to the largest value. Equation (17) is a step function that increases by 1/N at the value of each ordered data point. The K-S test of goodness of fit is based on the maximum distance between observed and expected distributions. It is defined as absolute maximum difference between observed and expected cumulative distributions, or

$$D = \max_{1\le i\le N}\left|F(X_i) - \frac{i}{N}\right|, \tag{18}$$

where F is the theoretical cumulative distribution being tested which must be a continuous and must be fully specified (mean and standard deviation estimated from the data in a normal distribution). The hypothesis regarding the distributional form is rejected if *D* is greater than the critical value obtained from a table. At 5 per cent level of significance, the critical value of *D* large sample is obtained by $1.36/\sqrt{N}$.

Application of the Model

We have applied the above model to analyse the distribution of female age at marriage in India and its constituent states based on the data available through different rounds of the National Family Health Survey (NFHS). The NFHS programme has been initiated by the Government of India, Ministry of Health and Family Welfare. The International Institute of Population Sciences, Mumbai is responsible for organising the survey. The first round of NFHS was carried out in 1992-93 and the fifth and the latest round was carried out in 2019-21. The present analysis is based on the data available from the first and the fourth round (2015-16) of NFHS. During the first round of NFHS, information about the age at marriage was collected from 3034 currently married women aged 15-49 years. During the fourth round of NFHS, information about the age at marriage was collected from 26534 currently married women aged 15-49 years. The lower and upper bounds of the model have been taken as 12 years and 35 years respectively. The data available from NFHS suggest that the proportion of females getting married after 35 years of age was very small so that these women were excluded from the analysis. Similarly, the proportion of females getting married before 12 years of age has also been found to be very small.

The estimate of the parameter λ of the model depends upon the lower and upper bounds of the model chosen in advance. Alternative values of the lower and the upper bounds of the model can be chosen depending upon the reported data. For the given values of the lower and the upper bounds of the model, the value of λ the positive skewness in the distribution of the age at marriage. This implies that the higher the value of λ the low the mean age at marriage or the mean age at marriage is less than the median age at marriage. It is also obvious that λ will be large only when $(\vartheta_2-x_i)>(x_i-\vartheta_1)$. On the other hand, a small value of λ implies that the distribution of age at marriage is negatively skewed which implies high mean age at marriage or the mean age at marriage is higher than the median age at marriage. In other words, there is inverse relationship between λ, the parameter of the model and the mean age at marriage. A decrease in the parameter λ over time implies an increase in the mean age at marriage whereas an increase in the parameter λ implies a decrease in the mean age at marriage. In Uttar Pradesh, the parameter λ decreased from 4.002 in 1992-93 to 2.407 in 2015-16 according to NFHS which means that the female age at marriage in the state has increased over time. Based on the data available from the first round of NFHS, the female mean age at marriage in Uttar Pradesh was 15.9 years in 1992-93 which increased to 17.47 years in 2015-16.

We have first fitted the model to the distribution of female age at marriage in Uttar Pradesh to examine the suitability of the model in describing the distribution of female age at marriage using the data from the first (1992-93) and the fourth round (2015-16) of NFHS. Table 1 presents the results of fitting the model to the data from Uttar Pradesh. The maximum absolute difference (K-S test value) between the observed and expected vales of the cumulative probability of marriage up to a certain age is estimated to be 0.073 for 1992-93. Similarly, the maximum absolute difference for the period 2015-16 is estimated to be and 0.117. These values suggest that the model proposed in this paper provides a very good fit to the observed data. This is also confirmed with the mean age at marriage.

Table 1: Observed and estimated cumulative probability female marriage at specific ages in Uttar Pradesh, 1992-93 and 2015-16.

Female age at marriage	1992-93		2015-16	
(years)	Observed	Expected	Observed	Expected
12-14	0.207	0.166	0.100	0.104
14-16	0.524	0.451	0.310	0.303
16-18	0.743	0.671	0.566	0.488
18-20	0.884	0.826	0.768	0.652
20-22	0.953	0.924	0.882	0.788
22-24	0.976	0.974	0.939	0.890
24-26	0.991	0.994	0.971	0.956
26-28	0.995	0.999	0.985	0.989
28-30	0.997	1.000	0.993	0.999
30-32	0.999	1.000	0.997	1.000
32-34	1.000	1.000	1.000	1.000
Mean age at marriage	15.90	(15.87)	17.47	(17.67)
λ	4.002		2.407	
Value of K-S test	0.073		0.117	

Source: Authors' calculations

Figure 1 shows the observed and the fitted cumulative probability of distribution of age at marriage of females in age in Uttar Pradesh. The figures again confirm that the model proposed in this paper describes the distribution of the age at marriage of females quite satisfactorily. There, however, appears some deviation between the observed and fitted values of the cumulative probability of marriage by age in the middle ages of the reproductive period. This difference is comparatively narrower in 1992-93 than in 2015-16. One possible reason for the difference between the observed and estimated cumulative probability of female marriage by age may be the errors associated with the reporting of the age at marriage of females in the household surveys and in population census. It is well known that reported age at marriage of females in the household surveys as well as in population census is associated with many errors. These include, among others, recall lapse, especially in older illiterate women and errors associated with digit preference. In India, marriage of females at age less than the minimum legal age at marriage is prohibited by law and is subject to litigation. In such a scenario, there is also a strong probability that the respondents during the household surveys may not be reporting the correct age at marriage females, especially when the female has got married before reaching the minimum legal age at marriage. It is obvious from figure 1 that the model proposed here can be used for ironing out the errors associated with the reported age at marriage of females.

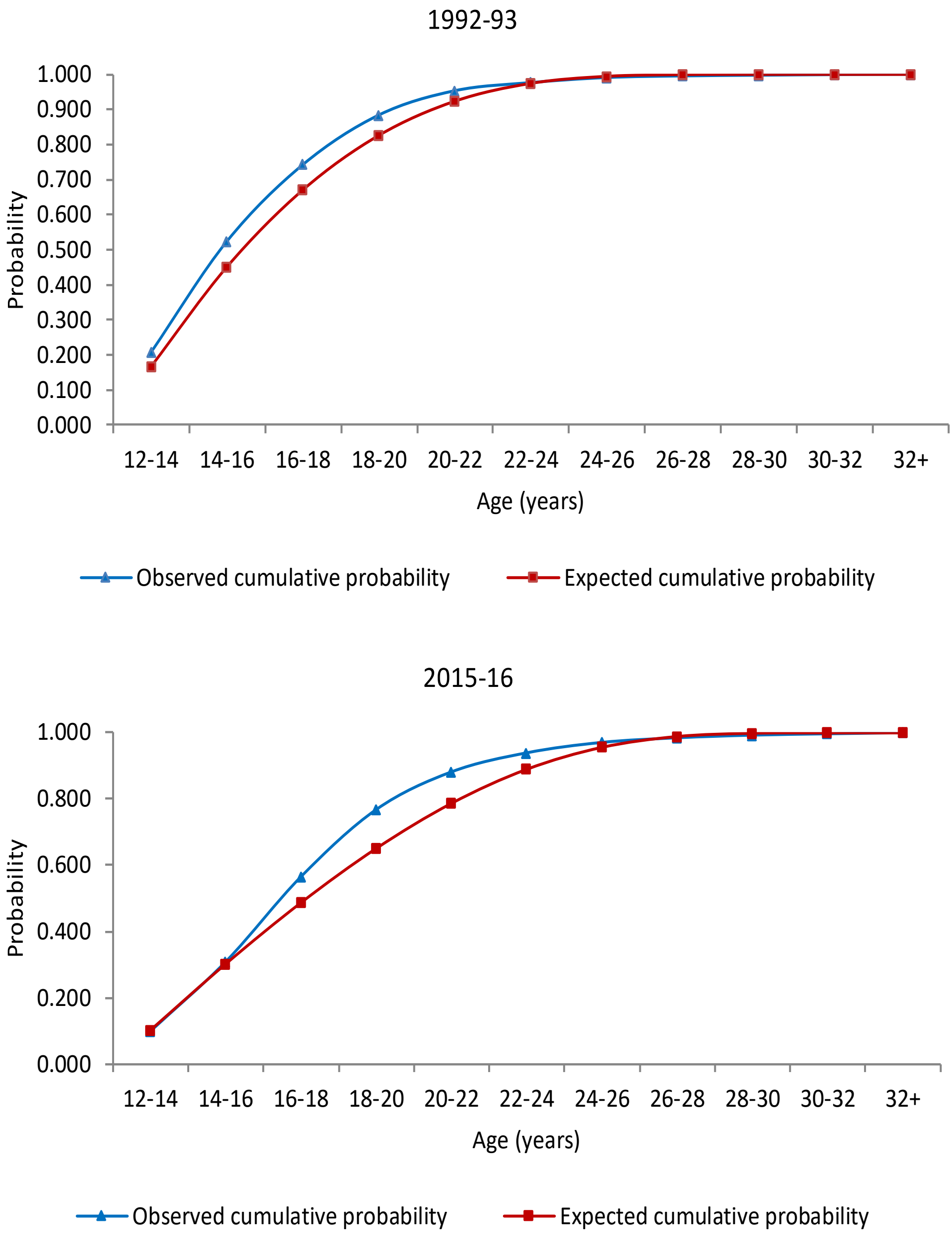

Figure 1: Observed and fitted values of cumulative probability of female age at marriage by age, Uttar Pradesh.

Table 2: Estimate of λ and implied female mean age at marriage in India and states, 1992-93 and 2015-16.

States	1992-93						2015-16					
	λ			Mean age at marriage			λ			Mean age at marriage		
	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
India	1.789	2.815	2.334	18.44	16.72	17.34	1.310	1.731	1.571	19.53	18.43	18.78
Assam	1.527	2.712	2.044	18.65	16.91	17.65	0.811	1.203	1.118	20.99	19.69	19.90
Haryana	2.261	4.448	3.219	17.76	15.86	16.61	1.634	1.980	1.838	19.40	18.58	18.88
Punjab	1.885	2.175	2.085	19.39	18.54	18.78	1.204	1.389	1.313	20.79	20.26	20.46
Bihar	3.162	4.837	4.183	16.53	14.88	15.37	2.246	2.558	2.509	17.61	17.13	17.19
Madhya Pradesh	2.598	4.830	3.812	16.69	15.13	15.62	1.703	2.703	2.261	18.45	16.94	17.44
Rajasthan	2.768	5.048	4.200	16.79	15.30	15.67	1.824	2.700	2.355	18.57	17.23	17.64
Uttar Pradesh	2.528	5.001	4.002	17.54	15.30	15.90	1.844	2.759	2.407	18.50	17.05	17.47
Gujarat	1.925	2.929	2.456	18.52	17.14	17.65	1.351	1.764	1.574	19.62	18.60	19.01
Maharashtra	2.352	4.448	3.081	17.68	15.60	16.63	1.503	2.045	1.791	19.34	18.29	18.70
Orissa	2.880	4.040	3.606	16.99	16.03	16.32	1.311	1.653	1.566	19.71	18.92	19.09
West Bengal	1.677	4.272	2.957	17.96	15.55	16.24	1.416	2.394	1.983	18.97	17.45	17.90
Andhra Pradesh	3.059	5.249	4.195	16.69	14.98	15.58	1.982	2.855	2.467	18.10	16.90	17.33
Karnataka	2.289	3.599	2.971	18.04	16.24	16.90	1.239	1.485	1.381	20.03	19.27	19.56
Kerala	1.356	1.581	1.506	20.02	19.26	19.49	1.091	1.049	1.065	20.97	21.06	21.02
Tamil Nadu	1.711	2.804	2.316	18.62	17.09	17.60	1.342	1.731	1.524	19.81	18.73	19.24

Source: Authors' calculations

Table 2 presents results of fitting the model to the data on female age at marriage for different states of the country for the total population and for rural and urban populations separately. The female mean age at marriage varies widely across the states of the country included in the analysis. The model suggests that the female mean age at marriage in 1992-93 was the lowest in Bihar followed by Andhra Pradesh, Madhya Pradesh, Rajasthan, and Uttar Pradesh. There are the states where the female mean age at marriage was less than 16 years and well below the national average. In 2015-16, the female mean age at marriage was less than the legal minimum age of female marriage in six states – Bihar followed by Andhra Pradesh, Madhya Pradesh, Uttar Pradesh, Rajasthan, and West Bengal. On the other hand, female mean age marriage was the highest in Kerala in both 1992-93 and 2015-16. In 1992-93, the female mean age at marriage was more than 18 years in only two states – Kerala and Punjab whereas, Kerala was the only state in 2015-16 where the female mean age at marriage was more than 21 years. Besides Kerala, the female mean age at marriage is estimated to be more than 20 years in Punjab in 2015-16.

The urban-rural difference in the female mean age at marriage is also revealing. In 1992-93, the female mean age at marriage in the urban areas of the country was more than 1.7 years higher than that in the rural areas. This difference has narrowed down to almost one year in 2015-16. Among different states, the urban-rural difference in the female mean age at marriage is estimated to be the widest in West Bengal but the narrowest in Kerala in 1992-93. In 2015-16 also the urban-rural difference in the female mean age at marriage was the widest in West Bengal but, the female mean at marriage in the rural areas of Kerala has been higher than that in the urban areas of the state. Kerala is the only state in the country where the female mean age at marriage appear to be higher than that in the urban areas. The decrease in the urban-rural gap in the female mean age at marriage between 1992-93 and 2015-16 has been the most rapid in Bihar followed by Haryana, Karnataka, and Maharashtra. By contrast, there has been virtually no change in the urban-rural difference in the female mean age at marriage in Madhya Pradesh during 1992 through 2016. In Rajasthan, Odisha, Punjab, and Gujarat also, there has been only a marginal decrease in the urban-rural difference in the female mean age at marriage.

The rate of increase in the female mean age at marriage in the rural areas of the country has been higher than that in the urban areas during the period under reference. In the rural areas of the country, the female mean age at marriage increased by more than 7 per cent per year between 1992-93 and 2015-16 whereas the rate of increase in the urban areas was less than 5 per cent per year. Similarly, the rate of increase in the rural female mean age at marriage has been higher than that in the urban areas in all states (Table 3). The most rapid increase in the female mean age at marriage is estimated in the rural areas of Karnataka during 1992-2016. By contrast, the rate of increase in the female mean age at marriage has been the slowest during this period in the urban areas of Kerala followed by Uttar Pradesh. In addition, the average annual rate of increase in the female mean age at marriage in the urban areas was less than 5 per cent per year in West Bengal, Bihar, and Gujarat. The decrease in the urban-rural difference in the female mean age marriage has primarily been due to the slow increase in the female mean age at marriage in the urban areas in all states.

Table 3: Proportionate change in λ and the mean age at marriage of females in India and states.

States	Proportion (per cent) change between 1992-93 and 2015-16					
	λ			Mean age at marriage		
	Urban	Rural	Total	Urban	Rural	Total
India	-26.77	-38.51	-32.69	5.91	10.23	8.30
Assam	-46.89	-55.64	-45.30	12.55	16.44	12.75
Haryana	-27.73	-55.49	-42.90	9.23	17.15	13.67
Punjab	-36.13	-36.14	-37.03	7.22	9.28	8.95
Bihar	-28.97	-47.12	-40.02	6.53	15.12	11.84
Madhya Pradesh	-34.45	-44.04	-40.69	10.55	11.96	11.65
Rajasthan	-34.10	-46.51	-43.93	10.60	12.61	12.57
Uttar Pradesh	-27.06	-44.83	-39.86	5.47	11.44	9.87
Gujarat	-29.82	-39.77	-35.91	5.94	8.52	7.71
Maharashtra	-36.10	-54.02	-41.87	9.39	17.24	12.45
Orissa	-54.48	-59.08	-56.57	16.01	18.03	16.97
West Bengal	-15.56	-43.96	-32.94	5.62	12.22	10.22
Andhra Pradesh	-35.21	-45.61	-41.19	8.45	12.82	11.23
Karnataka	-45.87	-58.74	-53.52	11.03	18.66	15.74
Kerala	-19.54	-33.65	-29.28	4.75	9.35	7.85
Tamil Nadu	-21.57	-38.27	-34.20	6.39	9.60	9.32

Source: Authors' calculations.

Table 4 presents the observed and estimated cumulative probability of marriage up to a specified age for India and for its selected states during 2015-16. In India, according to NFHS, 90 per cent of females got married by 24 years of age in 2015-16 – 87 per cent in the urban areas and 91 per cent in the rural areas. This proportion, according to our model is 93 per cent – 88 per cent in the urban areas and 96 per cent in the rural areas. In Andhra Pradesh, Bihar, Madhya Pradesh, Odisha, Rajasthan, and Uttar Pradesh, at least 90 per cent females got married by 22 years of age. By contrast, only around 80 per cent of females got married by 24 years of age in Kerala whereas only around two-third of the females got married by 22 years of age. It is also clear from the table that in the rural areas, more females got married by 18 years of age as compared to females in the urban areas in the country and in all states. In the rural areas, at least 90 per cent of females got married by 20 years of age in Andhra Pradesh, Madhya Pradesh, Uttar Pradesh, and West Bengal. By contrast, in the rural areas of Assam, Bihar, Haryana, Maharashtra, Rajasthan and Tamil Nadu, at least 90 per cent females got married by 22 years of age only. The comparison of estimated cumulative probability of marriage by age with the observed cumulative probability of marriage by age again confirms the appropriateness of the proposed model in describing the distribution of females by age at marriage.

Table 4: Observed and estimated proportion of females married by specific age in India and states, 2015-16.

Age (years)	Proportion of female married by age					
	Observed			Estimated		
	Urban	Rural	Total	Urban	Rural	Total
			Andhra Pradesh			
16	0.39	0.51	0.47	0.34	0.45	0.41
18	0.60	0.74	0.69	0.50	0.63	0.58
20	0.76	0.87	0.83	0.65	0.78	0.73
22	0.86	0.94	0.91	0.78	0.89	0.85
24	0.93	0.97	0.96	0.88	0.96	0.93
			Assam			
16	0.21	0.27	0.26	0.16	0.22	0.21
18	0.37	0.46	0.45	0.25	0.35	0.33
20	0.52	0.63	0.61	0.35	0.47	0.45
22	0.65	0.75	0.74	0.46	0.60	0.58
24	0.75	0.84	0.83	0.59	0.73	0.70
			Bihar			
16	0.43	0.49	0.48	0.38	0.42	0.41
18	0.67	0.71	0.70	0.55	0.59	0.59
20	0.81	0.85	0.85	0.70	0.74	0.74
22	0.90	0.92	0.92	0.82	0.86	0.85
24	0.95	0.96	0.96	0.91	0.94	0.94
			Gujarat			
16	0.22	0.32	0.28	0.25	0.31	0.28
18	0.43	0.54	0.50	0.38	0.46	0.43
20	0.64	0.72	0.69	0.51	0.61	0.57
22	0.79	0.85	0.83	0.65	0.74	0.70
24	0.88	0.92	0.91	0.77	0.85	0.82
			Haryana			
16	0.21	0.28	0.25	0.29	0.34	0.32
18	0.43	0.53	0.50	0.44	0.50	0.48
20	0.65	0.75	0.71	0.58	0.65	0.62
22	0.82	0.88	0.85	0.72	0.78	0.76
24	0.92	0.94	0.93	0.83	0.88	0.87
			Karnataka			
16	0.23	0.29	0.27	0.23	0.27	0.25
18	0.41	0.48	0.45	0.35	0.41	0.39
20	0.58	0.66	0.63	0.48	0.55	0.52
22	0.73	0.78	0.76	0.61	0.68	0.65
24	0.84	0.87	0.86	0.74	0.80	0.78
			Kerala			
16	0.14	0.12	0.13	0.21	0.20	0.20
18	0.28	0.28	0.28	0.32	0.31	0.31

Age (years)	Proportion of female married by age					
	Observed			Estimated		
	Urban	Rural	Total	Urban	Rural	Total
20	0.48	0.47	0.47	0.44	0.43	0.43
22	0.68	0.67	0.67	0.57	0.55	0.56
24	0.81	0.81	0.81	0.70	0.68	0.69
Madhya Pradesh						
16	0.36	0.52	0.47	0.30	0.43	0.38
18	0.58	0.74	0.69	0.45	0.61	0.55
20	0.74	0.86	0.82	0.60	0.76	0.70
22	0.84	0.93	0.90	0.73	0.87	0.82
24	0.91	0.96	0.94	0.84	0.95	0.92
Maharashtra						
16	0.26	0.33	0.31	0.27	0.35	0.31
18	0.47	0.57	0.53	0.41	0.51	0.47
20	0.65	0.76	0.72	0.55	0.66	0.62
22	0.79	0.88	0.84	0.69	0.79	0.75
24	0.89	0.94	0.92	0.81	0.89	0.86
Odisha						
16	0.25	0.29	0.28	0.24	0.29	0.28
18	0.45	0.51	0.50	0.37	0.44	0.42
20	0.63	0.71	0.70	0.50	0.59	0.57
22	0.76	0.83	0.82	0.64	0.72	0.70
24	0.86	0.91	0.90	0.76	0.84	0.82
Punjab						
16	0.10	0.10	0.10	0.22	0.25	0.24
18	0.27	0.31	0.30	0.35	0.39	0.37
20	0.50	0.58	0.55	0.47	0.52	0.50
22	0.71	0.78	0.75	0.60	0.66	0.64
24	0.85	0.90	0.88	0.73	0.78	0.76
Rajasthan						
16	0.32	0.45	0.41	0.32	0.43	0.39
18	0.54	0.70	0.65	0.47	0.61	0.56
20	0.73	0.86	0.82	0.62	0.76	0.72
22	0.85	0.94	0.91	0.75	0.87	0.84
24	0.93	0.97	0.96	0.86	0.95	0.92
Tamil Nadu						
16	0.22	0.31	0.27	0.25	0.31	0.27
18	0.41	0.53	0.47	0.38	0.46	0.42
20	0.62	0.72	0.67	0.51	0.60	0.56
22	0.76	0.84	0.80	0.64	0.74	0.69
24	0.87	0.92	0.89	0.77	0.85	0.81
Uttar Pradesh						
16	0.32	0.48	0.43	0.32	0.44	0.40

Age (years)	Proportion of female married by age					
	Observed			Estimated		
	Urban	Rural	Total	Urban	Rural	Total
18	0.56	0.73	0.68	0.48	0.62	0.57
20	0.75	0.87	0.84	0.63	0.77	0.72
22	0.86	0.94	0.92	0.76	0.88	0.84
24	0.93	0.97	0.96	0.87	0.95	0.93
			West Bengal			
16	0.32	0.44	0.40	0.26	0.40	0.34
18	0.54	0.70	0.65	0.39	0.57	0.50
20	0.70	0.86	0.81	0.53	0.72	0.65
22	0.81	0.92	0.89	0.66	0.84	0.78
24	0.88	0.96	0.93	0.79	0.93	0.89
			India			
16	0.26	0.36	0.33	0.34	0.45	0.41
18	0.46	0.58	0.54	0.50	0.63	0.58
20	0.64	0.75	0.71	0.65	0.78	0.73
22	0.78	0.85	0.83	0.78	0.89	0.85
24	0.87	0.91	0.90	0.88	0.96	0.93

Source: Authors' calculations.

Conclusions

In this paper, we have proposed a probability model that describes the distribution of the age at marriage of females. Application of the model to the data available from the National Family Health Survey suggests that the model provides a very good fit to the observed data. The model can be used for prediction purposes and for smoothing the data on female age at marriage as it is well known that the reported data on female age at marriage are associated with number of errors including errors associated with digit preference and recall lapse. The proposed model has only one parameter which depends upon the lower and upper bounds of the model. Changing the upper and lower bounds of the model leads to the change in the parameter of the model. As such, the model can be used for simulation purposes also. By changing the lower and upper bounds of the model, one can obtain the likely estimates of the mean age marriage which may be useful for formulating policies designing programmes directed towards modifying the distribution of age at marriage of females.

References

Agarwal SN (1962) *Age at marriage in India*. Allahabad, Kitab Mahal.

Alho J (2016) Modelling incidence of nuptiality. *Sociological Methodology*46(1): 283-318.

Coale A J (1971) Age patterns of marriage. *Population Studies* 25(2): 193-214.

Coale AJ, MeNeil DR (1972) The distribution by age of the frequency of first marriage in a female cohort. *Journal of American Statistical Association* 67(340): 743-749.

Dickman A (1989) Diffusion and survival models for the process of entry into marriage. *Journal of Mathematical Sociology* 14(1): 31-44.

Henry L (1972) Nuptiality. *Theoretical Population Biology* 3(2): 135-152.

Hernes G (1972) The process of entry into first marriage. *American Sociological Review* 37(2): 173-182.

Hill K (1977) Estimating adult mortality levels from information on widowhood. *Population Studies* 31(1): 75-84.

Hyrenius H, Holmberry I, Carlsson M (1967) *Demographic Methods, DM3*.Goteborg, University of Goteborg. Demographic Institute.

Islam MN (1984) Estimation of mean age at marriage: Use of a simple mathematical model. *Rural Demography* 11(1-2): 39-59.

Kolmogorov AN (1933) Sulla determinazione empirica di una legge di distribuzione. *Giornale dell' Istituto Italiano degli Attuari* 4: 83–91.

Krishnan P (1971) A note on changes in age at marriage of females and their effect on the birth rate in India. *Social biology* 18(2): 200-202.

Malaker CR (1985) Modelling of Indian nuptiality. *Sankhya* Series B: 118-127.

Malaker CR (1987) On recent developments of marriage models and their applications to Indian nuptiality. *Genus* 43(1-2): 93-102.

Matthews AP, Garenne ML (2013a) A dynamic model of the marriage market-part 1: matching algorithm based on age preference and availability. *Theoretical Population Biology* 88: 78-85.

Matthews AP, Garenne ML (2013b) A dynamic model of the marriage market-part 2: simulation of marital states and application to empirical data. *Theoretical Population Biology* 88: 86-93.

McFarland DD (1972) Comparison of alternative marriage models. In TNE Greville (Ed) *Population Dynamics*. New York, Academic Press.

McFarland DD (1975) Models of marriage and fertility. *Social Forces* 54: 66-83.

Mishra A (1979) On the fitting of a mathematical model to the statistics of age at first marriage. *Demography India* 8(1-2): 292-297.

Mitter P (1989) Compound arrival times and a search model of marriage. *Journal of Mathematical Sociology* 15(1): 1-9.

Nath DC, Talukdar PK (1992) A model that fits female age at first marriage in a traditional society. *Janasankhya* 10(1-2): 53-59.

Nydell S (1924) The construction of curves of equal frequency in case of logarithmic a correlation with application to the distribution of ages at first marriage. *Scandinavian Actuarial Journal* 1924(1): 36-53.

Pathak KB, Prasad CVS (1977) A model for estimating adolescent sterility among married women. *Demography* 14(1): 103-104.

Rodriguez G, Trussell J (1980) Maximum likelihood estimates of the parameters of Coale's nuptiality schedules from survey data. *Technical Bulletin* No. 7.

Singh KK, Singh BP (2004) On the age pattern of marriage. In DC Nath, KK Das (Eds) *Statistical Thinking: Methods and Applications*. Guwahati, Capital Publication Company.

Singh NK, Singh BP (2015) A probability model for age at marriage. *International Journal of Applied Mathematics & Statistical Sciences* (IJAMSS) 4(4): 33-38.

Singh VK, Mishra RN, Singh KK, Dwivedi SN (1986) On the pattern of age at marriage in rural India: A mathematical approach. *Rural Demography* 13(1-2): 47-54.

Smirnov NV (1933) Estimate of deviation between empirical distribution functions in two independent samples. *Bulletin Moscow University* 2: 3–16.

Talwar PP (1967) A note on changes in age at marriage of female and their effect on the birth rate in India. *Eugenics Quarterly* 14(4): 291-295.

Verma S, Pathak KB (2001) A model for female age at marriage. Paper presented at the National Conference on Emerging Methodologies of Data Analysis and Related Inferences, BHU, Varanasi, India, Feb. 5-7, 2001.

Vlassof C (1987) Awareness and access to reproductive care among adolescents. Retrieved from http://iussp2009.princeton.edu/papers/91518.