Demography India
Vol. 48, No. 1 (2019), pp. 49-62 ISSN 0970-454X

A Study of Timing of First Sexual Initiation in India: Application of Some Potential Probability Distributions

Abhishek Gautam[*1,2], Kaushalendra Kumar Singh[1], Brijesh P. Singh[1] and Ravi Verma[2]

Abstract: Age at first sexual initiation is a very sensitive and important issue and discussing about these issues openly with anybody is almost restricted and treated as inviolable in India. Recently, with the increase in the level of education and modernization of the society, a paradigm shift is observed in attitude and perception of the young generation about sexual activity, that is new generation boys and girls have started experimenting with sexual activity in the pre marriage period. In this study an attempt has been made to understand, age and residence wise differences in age at first sexual initiation or sexual debut. In this paper an attempt has been made to develop a model for the pattern of sexual debut with a suitable probability distribution. Findings reveal that the average age at sexual debut and intensity of initiation of first sex is significantly different in urban than rural area. It is also observed that early sexual debut is seen among the younger generation.

Keywords: Age of Sexual Initiation, Weibull Distribution, Probability Modelling, Maximum Likelihood Estimate, Hazard Rate, Kolmogorov-Smirnov test.

Introduction

The age at first sexual initiation is a significant event in one's personal life. In Indian traditional society, it is very sensitive issue and discussing about sexual matters openly with anybody is nearly completely restricted and considered as serious taboo. Because of this, it remained a closed-door issue for the people living in traditional societies. It is because of the prevailing extreme authoritarian view in such matters which is opposite to the anarchic view. The believers of the extreme authoritarian view claim that sexual problems can be disappeared if the 'sexual mores' are observed because sexual mores are developed on the basis of the wisdom of cumulative human experience over thousands of years. On the other hand, believers of anarchic views argue that sex is natural and private; hence attempts to regulate it are contrary to nature and an invasion of privacy. According to this view, the effort to suppress and control sexual expression creates hypocrisy; subject individuals to the risk of public disgrace, and often lead to excessive guilt and neurosis. The discussion over sexual regulation thus boils down in practice to a conflict between 'conservative' and 'liberal' points of view, neither of which can give an absolute solution. With the increase in the level of education and modernization of the society, a paradigm shift is observed from extreme authoritarian view to anarchic view, in the attitude and perception of the young generation about sexual activity, that is new generation boys and girls have started experimenting with sexual activity in the pre marriage period.

In India law legalizes the sexual union only through marriage; outside marriage as well as pre-marital sex is a social taboo and not permitted at all. In the wake of the HIV pandemic, sexual initiation before marriage or beyond marriage has become a focus of attention. In various studies, it is reported that the age at first sex increases the risk factor for the acquisition

[*] Corresponding author, abhishekiips@gmail.com
[1] Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi-221005, Uttar Pradesh.
[2] International Centre for Research on Women, Asia Regional Office, C-59 south Extension Part-2, New Delhi-110049.

of HIV. Persons starting their sexual activities at young age are likely to report higher rates of STIs, drug and alcohol abuse, multiple sexual partners and to engage in unprotected sex (Dickson *et al.*, 1998; Duncan *et al.*, 1990; Greenberg *et al.*, 1992; Manning *et al.*, 2000; Musie and Matthew, 2013). Several studies have reported a negative association between delaying sexual debut and HIV prevalence (Gregson *et al.*, 2006; Stoneburner and Low-Beer, 2004). It has also been found that sexual initiation at earlier ages increases the risk of major depressive episode (Gonçalves, 2017). From an intervention point of view, the proportion of young people who are sexually active, especially before they form stable partnerships, is an important area of concern. Early age at sexual debut and the number of pre-marital partners have been shown to be correlated with risk behaviour later in the life (White *et al.*, 2000). Contextual data on sexual partners and circumstances are needed to understand and assess the risk associated with early sexual experiences, and thereby design appropriate policies and programmes. Systematic information on sexuality and sexual behaviour in India is scarce. Despite a rapid increase in intervention research on sexual behaviour and health in the second half of the 1990s, most studies covered groups with high risk behaviour, and little is known about what happens in the general population (Nag, 1995; Pelto, 1999). An exception is a study of married men in the state of Uttar Pradesh, in which 14.5 per cent of the men were reported to have had sexual intercourse before marriage (Singh *et al.*, 1998). Only a few studies of Indian young people report on their sexual behaviour (Jejeebhoy, 1998). Studies on sexual behaviour is generally difficult due to various social taboos associated with it and due to this data on sexual behaviour in India is too scare to draw any meaningful conclusion. Few studies are carried out form diverse perspective and among diverse groups of people (Nag, 1995).

The reason behind existence of sexual norms can be understood in terms of their application in human society. Like other forms of behaviour, sexual activity related behaviour must be learned. Without socialization, human beings cannot know even how to copulate. The sex norms are like other norms in society, which help the business of society accomplished, and contribute in smooth functioning of the system as well. As a powerful drive, sex can be utilized to motivate people to perform in ways that benefit the community at large, while lack of regulation may bring 'confliction and disruption'. Sexual gratification cannot be entirely suppressed after puberty. It is not required for individual survival as release from hunger, thirst or fatigue. But the sex drive is capable of an extra ordinary amount of both situational and emotional, conditioning. This gives the sex norms an amazing variety of possible behaviours to regulate. However, the knowledge of sex and reproductive health are limited among both educated and uneducated adolescent, very few rural adolescent females could describe how body change related to sexual intercourse and reproduction (Vlassof, 1987).

In India, information on the sexual activity is scarcely available because of socio-cultural taboos associated with the sexual issues; also, data quality is poor and is under reported. The early sexual initiation for prolong period of exposure leads to more pregnancy and more chance of STIs (O'Donnell *et al.*, 2001). In urban area, free lifestyle and friendship among young males and females also enhance the sexual activities (Jejeebhoy, 1998). Sexual initiation is a significant transitional point for adolescents (Carpenter, 2005) and its timing can impact their health and well-being. Studies have demonstrated that the timing of sexual debut is linked to a wide range of individual socio-demographic and psychosocial factors, as well as broad structural level factors (Luke, 2003; Maticka-Tyndale *et al.*, 2005; Tenkorang and Maticka-Tyndale, 2008). The timing and situation of first sexual debut among youngsters continues to attract much interest because this marks the beginning of their exposure to a range of sexual and reproductive health outcomes (Yode and LeGrand, 2012). The situation under which the

event occurs may also have implications for their future sexual behaviours and health (Ghebremichael *et al.*, 2009) but the mechanism through which age at sexual debut in men who have sex with women may influence sexual activity is unknown. In the United States, the average age for first vaginal intercourse is 17.1 years among men, with 90 percent of youth sexually active by age 19 (Kaestle *et al.*, 2005).

Drawing on in-depth sexual case histories, this article provides information in the context to young men's first sexual encounters, describing the range of sexual partners. These quantitative findings are supplemented by data from a general population sample giving numerical estimates of the parameters of the age at sexual debut, which allow for comparison between various groups of the society. The present study aims to obtain an appropriate statistical probability model for the age pattern of the sexual debut and to test how well the chosen statistical probability model fits to the data. Also, this study makes basic demographic information of the males and to compare the mean ages of the occurrence of sexual debut between two different places of residence, viz., rural and urban. This study will help to draw the attention of demographers to the most prevalent factor in the study region as probability estimation of intensity of sexual debut for a particular age group and place of residence can be useful tool in the development of a risk prediction model using covariates.

Application of Some Potential Probability Distributions for Age at Sexual Initiation

Model-I

Age at first sexual initiation is a continuous variable and occurs at a certain age after the birth of the respondent. Considering this as an event and assuming that it is a subjected to a constant risk, we can think that it must be exponentially distributed. Thus, we can model age at first sexual initiation as a random variable *X* following an exponential distribution with Probability Density Function (PDF) given as follows

$$f(x)=\begin{cases}\lambda\exp(-\lambda x), x\geq 0\\ 0, otherwise\end{cases} \quad (1)$$

Here λ is the risk of sexual initiation. Then, the proportion of male with sexual initiation less than or equal to *x* is given by

$$F(x)=\begin{cases}1-\exp(-\lambda x), x\geq 0\\ 0, otherwise\end{cases} \quad (2)$$

The suitability of the above model can be challenged on the ground that the age of sexual initiation starts at some point of time which is certainly other than zero whereas the assumed model starts from zero. In order to incorporate this fact in the model, we propose the use of displaced exponential model described below and called hereafter as model- II.

Model-II

The probability density function of *X* following displaced exponential distribution, which is also known as two parameters exponential distribution *Exp* (λ , θ), is as follows:

$$f(x)=\lambda\exp[-\lambda(x-\theta)],\ where\ x\geq\theta\ and\ \lambda>0 \quad (3)$$

and the cumulative distribution function is

$$F(x)=1-\exp[-\lambda(x-\theta)] \quad (4)$$

Although the above model is simple but one can question that the inherent assumption of constant risk for the event of sexual initiation for everybody seems to be too restrictive to be true for its applicability because the population does consist two types of individuals firstly believers of extreme authoritarian view and secondly believers of anarchic views. Naturally, the first group has lesser risk to the exposure of event as compared to the other group. This motivated us to think of a mixture of exponential distributions as a suitable model. We will call this model as model –III and it is described below:

Model-III

Model II is a simple model that is perhaps not suitable for the age at sexual initiation because it is not homogeneous in the population. It is controlled by several character of the respondent. Thus, in order to capture the heterogeneity, we assume age at sexual initiation follows a mixture of two displaced exponential distributions, i.e.

$$f(x) = \alpha\theta_1 e^{-\theta_1(x-k)} + (1-\alpha)\theta_2 e^{-\theta_2(x-k)} \text{ where } x \geq k, 0 < \alpha < 1 \text{ and } \theta_1, \theta_2 > 0 \quad (5)$$

and the cumulative distribution function is

$$F(x) = 1 - \alpha e^{-\theta_1(x-k)} - (1-\alpha)e^{-\theta_2(x-k)} \quad (6)$$

This distribution has four parameters. One parameter among these i.e. *k* can be assumed to be the age from which one is exposed to the risk of sexual initiation.

Estimation of the parameters of the models

Estimation of parameter of Model-I

This model contains only one parameter to be estimated from the observed distribution. The estimated value by the method of maximum likelihood is given as

$\hat{\lambda} = \frac{1}{\bar{x}}$ here $\bar{x}$ is the mean age of the first sexual initiation.

Estimation of parameters of Model-II

The Method of Maximum Likelihood (ML) for the parameters of the 2-parameter Exponential Distribution or displaced exponential distribution is easy to obtain.

If X_1 , X_2 , ..., X_n is a random sample from Exp (λ, θ), the likelihood function can be written as

$$L(\theta, \lambda) = \lambda^n \exp\left[-\lambda \sum_{i=1}^{n} (x_i - \theta)\right] \text{ provided each } X_i \text{ is } \leq \theta. \quad (7)$$

and the log likelihood will be

$$\log L(\theta, \lambda) = l = n \log \lambda - \lambda \sum_{i=1}^{n} (x_i - \theta) \ . \quad (8)$$

It is easy to see that the likelihood is maximized for the value of θ that minimizes $\sum_{i=1}^{n}(x_i - \theta)$. Hence $X_{(1)}$ = min.(X_1 , X_2 , ..., X_n) is MLE $\hat{\theta}$ of θ. To find the MLE of λ , we have to solve the following equation

$$\frac{\partial L}{\partial \lambda} = \frac{n}{\hat{\lambda}} + n\hat{\theta} - \sum_{i=n}^{n} x_i = 0 \quad (9)$$

which provides the solution as follows:

$$\hat{\lambda} = \frac{n}{(\sum_{i=1}^{n} x_i) - n\hat{\theta}} \tag{10}$$

Estimation of parameters of Model-III

The maximum likelihood estimates of the parameters are quite complicated because the likelihood function and the resulting likelihood equations are not analytically solvable. Therefore, for the sake of ease and simplicity, we propose the use of method of moments for the estimation of the parameters.

MM Estimate of Mixture of displaced exponential distribution

The expressions for the first three raw moments of the distribution can be easily evaluated as given below:

$$\mu_1' = \alpha\left(\frac{1}{\theta_1} + k\right) + (1-\alpha)\left(\frac{1}{\theta_2} + k\right) \tag{11}$$

$$\mu_2' = \alpha\left[\left(\frac{1}{\theta_1} + k\right)^2 + \frac{1}{\theta_1^2}\right] + (1-\alpha)\left[\left(\frac{1}{\theta_2} + k\right)^2 + \frac{1}{\theta_2^2}\right] \tag{12}$$

$$\mu_3' = \alpha\left[\left(\frac{1}{\theta_1} + k\right)^3 + \frac{5}{\theta_1^2}\left(\frac{1}{\theta_1} + k\right) - \frac{2k}{\theta_1^2}\right] + (1-\alpha)\left[\left(\frac{1}{\theta_1} + k\right)^3 + \frac{5}{\theta_1^2}\left(\frac{1}{\theta_1} + k\right) - \frac{2k}{\theta_1^2}\right] \tag{13}$$

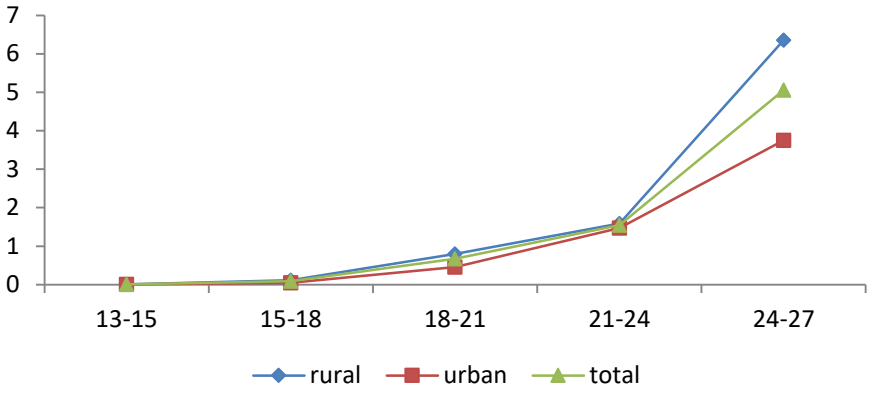
We tried to fit the above distributions to the various set of data on age at sexual initiation and noted that these do not fit well. Even the model III which is a mixture of two exponential distributions failed to explain satisfactorily the phenomenon (age at sexual initiation) considered here in this study. Therefore, it is thought that the model choice can be made on the basis of a study the nature of hazard rate.

The phenomenon of timing of sexual initiation i.e. the chance of experiencing sex increases with the age, Thus, it seems natural to expect that it may be considered to be a phenomenon of increasing type of hazard rate. Table 1 and figure 1 supports that the age at sexual debut is a phenomenon of increasing type of hazard rate.

Table 1: Hazard Rate of age at First Sexual Initiation

Age	Place of residence		Total
	Rural	Urban	
<15	0.00747	0.003976	0.006353
15-18	0.115625	0.047917	0.093056
18-21	0.801126	0.458967	0.670534
21-24	1.587379	1.473684	1.542773
24-27	6.357143	3.75	5.053571

Figure 1: Plot of Hazard Rate of age at first sexual initiation in India



The rate of change of hazard rate of sexual initiation in rural area is higher than the rate of change of hazard rate in urban areas. Also the hazard rate at every age is more in rural area than the urban area. From the above discussions, it is clear that for modelling first sexual initiation in India, we are motivated to consider a distribution which possesses monotone increasing hazard rate and hence propose the following model:

Model-IV

Let X be a random variable denoting the age at sexual initiation. Naturally, it shall take only non-negative values starting from some value other than zero. Further, we consider hazard rate to be the power function of X given by:

$$h(x;\alpha,\beta,\gamma)=\frac{\beta}{\alpha}\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1},\ where\ \alpha,\beta>0, \qquad 0<\gamma<x<\infty \qquad (14)$$

Depending on the choice of different values of β, its hazard rate can be increasing, decreasing or constant. It may be noted that $h(x)$ is increasing if $\beta>1$, decreasing if $\beta<1$, and is constant if $\beta=1$. For a given hazard rate, the corresponding Probability Density Function can be obtained from the following expression:

$$f(x)=h(x)\exp\left\{-\int_0^x h(w)dw\right\}.$$

Substituting h(.) from (14) in the above expression, we get after simplification, the corresponding p.d.f. as,

$$f(x;\alpha,\beta,\gamma)=\left(\frac{\beta}{\alpha}\right)\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1}\exp\left\{-\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right\}\ where\,\alpha,\beta>0\,and\,\gamma\le x \qquad (15)$$

This is the Probability Density Function of well-known three-parameter Weibull distribution which is one of the most popular models possessing monotone hazard rate. Since, β controls the shape of the distribution so it is called the shape parameter, while α and γ are usually referred to as the scale and location parameters. The value of β varies from data to data, but in many situations value of β varies from 1 to 3 which seem to be appropriate for good fit (Lawless, 2002). It is perhaps the most widely used lifetime distribution due to its broad range of shapes and has been used to describe events and processes which are quite common and

frequently used in biometrical sciences (Lawless, 2002) and may be a suitable model for the age at sexual initiation; where, α can be interpreted as propensity of sexuality and β as shape of sexual pattern. The location parameter γ gives the idea about the threshold value i.e. initial age of sexual initiation. Thus the unique feature of the model is that this provides an idea about the threshold value (theoretical value of starting point) and propensity or intensity of age at sexual initiation in different group of population. Singh *et al.* (2010) and Singh and Roy (2015) have used this distribution to model the data on age at sterilization from India and Bangladesh. The mean and variance of the distribution is given by:

$$E(x)=\int_0^{\infty} x\, f(x)\, dx=\gamma+\alpha\ \Gamma\left(1+\frac{1}{\beta}\right) \tag{16}$$

$$Var(x)=\alpha^2\left[\Gamma\left(1+\frac{2}{\beta}\right)-\left\{\Gamma\left(1+\frac{1}{\beta}\right)\right\}^2\right] \tag{17}$$

Parameter Estimation

For applying the above proposed distributions, it is needed to estimate the value of its parameters which we propose to be obtained by using maximum likelihood estimation method. The method of maximum likelihood (Harter and Moore, 1965a; Harter and Moore, 1965b, and Cohen, 1965) is a commonly used procedure because it possesses various desirable properties. Let $x_1, x_2, \ldots\ldots\ldots, x_n$ be a random sample of size n drawn a probability density function $f(x;\alpha,\beta,\gamma)$ where α, β and γ are unknown parameters. The likelihood function is joint density of these *n* random observations treated as a function of the unknown parameter. Thus**,** the likelihood function is multiplication of individual density function given as,

$$L(\alpha,\beta,\gamma)=\prod_{i=1}^{n} f(x_i;\alpha,\beta,\gamma)$$

Using (15), we get

$$L(\alpha,\beta,\gamma)=\frac{\beta^n}{\alpha^{\beta n}}\left[\prod_{i=1}^{n}\left(x_i-\gamma\right)^{\beta-1}\right]e^{-\frac{1}{\alpha^{\beta}}\sum_{i=1}^{n}\left(x_i-\gamma\right)^{\beta}} \tag{18}$$

The maximum likelihood estimator (MLE) of the parameters is that value which maximizes the likelihood function. Thus, these can be obtained by solving the equations resulting from setting the three partial derivatives of log *L* (α, β, γ) with respect to α, β and γ to zero. The resulting likelihood equations are given below:

$$\left(\hat{\alpha}\right)^{\hat{\beta}}-\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\hat{\gamma}\right)^{\hat{\beta}}=0 \tag{19}$$

$$\frac{\sum_{i=1}^{n}\left(x_i-\hat{\gamma}\right)^{\hat{\beta}}\ln\left(x_i-\hat{\gamma}\right)}{\sum_{i=1}^{n}\left(x_i-\hat{\gamma}\right)^{\hat{\beta}}}-\frac{1}{\hat{\beta}}-\frac{1}{n}\sum_{i=1}^{n}\ln\left(x_i-\hat{\gamma}\right)=0 \tag{20}$$

$$\left(\hat{\beta}-1\right)\sum_{i=1}^{n}\left(x_i-\hat{\gamma}\right)^{-1}-\hat{\beta}\hat{\alpha}^{-\hat{\beta}}\sum_{i=1}^{n}\left(x_i-\hat{\gamma}\right)^{\hat{\beta}-1}=0 \qquad (21)$$

The above equations are not solvable analytically and hence one has to opt for some numerical iterative technique. Marks (2005) have suggested the use of Newton-Raphson algorithm for obtaining the maximum likelihood estimation of three-parameter Weibull distribution and the same has been used here also.

Goodness-of-fit test

Goodness of fit test procedures is intended to detect the existence of a significant difference between the observed frequency of occurrence and the theoretical pattern of occurrence of first sexual initiation. For testing the goodness of fit of three parameter Weibull distribution we have used Kolmogorov-Smirnov (K-S) test for testing the hypothesis:

H_0: Data follow a three parameter Weibull distribution

against the alternative hypothesis

H_a: Data do not follow the three parameter Weibull distribution

A brief description of the test is given below:

Kolmogorov-Smirnov test of goodness of fit

The Kolmogorov-Smirnov (K-S) test is based on the empirical cumulative distribution function (ECDF). Given *N* ordered data points Y_1, Y_2, ..., Y_N, the ECDF is defined as:

$$E_N = n(i)/N$$

where *n(i)* is the number of points less than Y_i (the ith ordered observation, ordered from smallest to largest). This is a step function that increases by 1/*N* at each ordered data point. The K-S test is based on the maximum distance between the fitted theoretical cumulative distribution and the ECDF. The Kolmogorov-Smirnov test statistic is defined as:

$$D=\max_{1\le i\le N}\left|F(Y_i)-\frac{i}{N}\right|$$

Where *F* is the theoretical cumulative distribution. The hypothesis regarding the distributional form is rejected if the test statistic *D* is greater than the critical value obtained from the table. At 5% level of significance the critical value of K-S test for large sample can be approximated by $1.36/\sqrt{N}$, where N is the sample size.

Data

The data on age at first sexual initiation considered in this study has been taken from a study, conducted by International Center for Research on Women (ICRW) with support from UNFPA in 2013-14, entitled "Masculinity, Intimate Partner Violence and Son Preference in India Study" (Nanda *et al.,* 2014). The study was conducted in seven states namely Uttar Pradesh, Rajasthan, Punjab and Haryana, Odisha, Maharashtra and Madhya Pradesh. The study was designed to gather information from men about their attitude towards a preference for son, intimate partner violence and masculine behaviour. This is the first-ever study of its kind to be

carried out in India on such a large scale. Out of the total 9,135 men covered in the original survey, the current analysis is focused on 6069 currently married men aged 18-49. Several questions were asked from the respondents including some questions on the timing of sexual initiation. In this study our aim is to model the pattern of age at sexual initiation. Due to recall lapse, it is often seen that a response, regarding an event which has taken place a long time ago, is often potentially erroneous. Therefore, it is decided to include in the present study only those males whose age at the time of collection of the data was less than 30 years for modelling the age at sexual initiation. Thus, the sample size for study is 1584 among these 1079 males from rural background and 505 males from urban background. Further we categorise the above data according to the age of respondent in two groups i.e. age of respondents between 25-30 years and less than 25 years for comparison purpose.

Results

Table 2 shows the average and standard deviation of the age at sexual initiation according to the age of the respondent on survey date and the place of their residence. Average age of sexual initiation of the younger respondents (age <30 years) is 21.08 years however for elder respondent's (age ≥30 years) sexual debut is about 23 years. The variation in age at sexual debut for younger respondents is more than that of elder respondents. The average age at sexual initiation for urban respondents is 23 years which is more than that of rural respondents (21.76 years).

Table 2: Summary measure of age at first Sexual Initiation

Characteristics	Mean	Standard Deviation	Sample size N
Age of the respondent			
<30	21.08	2.843	1584
30-40	22.75	3.827	2461
≥40	22.52	4.096	2024
Place of residence			
Rural	21.76	3.608	3713
Urban	22.99	3.863	2356
Total	22.24	3.757	6069

Table 3: ANOVA for age at first Sexual Initiation

Source of variation	Sum of Squares	df	Mean Square	F	p-value
Between age groups	2901.132	2	1450.566	106.322	0.000
Within age groups	82758.983	6066	13.643		
Total	85660.115	6068			

Table 4: Post Hoc Tests for Multiple Comparisons of age at first Sexual Initiation

Age group of respondents		Mean Difference From I^{th} to J^{th} group	Std. Error	p-value	95% Confidence Interval	
I	J				Lower	Upper
<30	30-40	-1.663	0.119	0.000	-1.95	-1.38
	≥40	-1.431	0.124	0.000	-1.73	-1.13
30-40	≥40	0.232	0.111	0.109	-0.03	0.50

Table 5: Summary measure of age at first Sexual Initiation in India according to place of residence and age of respondents

Age of respondents	Place of residence	Sample size N	Mean	Std. Deviation	Std. Error Mean
<30	rural	1079	20.82	2.845	0.087
	urban	505	21.66	2.755	0.123
30-40	rural	1444	22.18	3.670	0.097
	urban	1017	23.55	3.902	0.122
≥40	rural	1190	22.09	3.978	0.115
	urban	834	23.12	4.189	0.145
All	rural	3713	21.76	3.608	0.059
	urban	2356	22.99	3.863	0.080

From the above discussions, we see that the mean age at first sexual initiation varies from one group of respondents to other. In order to test whether this difference is statistically significant are not; we carried out an analysis of variance, assuming a model $y_{ij} = \mu + \alpha i + \varepsilon_{ij}$ where y_{ij} is age at first sexual initiation for j^{th} respondent belonging to ith (i=1, 2, 3) group of respondents, μ is general mean age, αi is the mean effect due to i^{th} group of respondents and ε_{ij} is the error term. We wish to test the hypothesis H0: $\alpha 1 = \alpha 2 = \alpha 3$ against the alternative that any two differ. The results obtained are summarized in Table 3 and shows that the average age at sexual initiation is significantly different for the various age groups of respondents. Thus, we carried out the post hoc tests for multiple comparisons and the results are presented in table 4, which reveals that younger respondents (age <30 years) are having significantly less mean age at first sexual initiation than others.

Table 5 and 6 present the urban/rural differential in age at sexual initiation according to the age of the respondents on survey date. In all the age categories of the respondents, age at sexual initiation in rural areas is found to be smaller than that of urban areas. The finding of two independent sample tests indicates that the differences in the age at sexual initiation are highly significant for the urban/rural respondents in the various age categories of age of respondents on survey date.

Table 6: Test for Equality of Mean age at Sexual Initiation for Rural versus Urban

Age of respondents	Absolute value of the statistics	df	p-value	Mean Difference	Std. Error of Difference	95% Confidence Interval of the Difference	
						Lower	Upper
<30	5.519	1582	0.000	-0.838	0.152	-1.136	-0.540
30-40	8.878	2459	0.000	-1.369	0.154	-1.672	-1.067
≥40	5.589	2022	0.000	-1.026	0.184	-1.386	-0.666
All	12.641	6067	0.000	-1.235	0.098	-1.426	-1.043

Table 7, 8 and 9 provides empirical and theoretical distribution of the age at sexual initiation for all, urban and rural males respectively. The calculated value of K-S test statistics and corresponding critical value at 5 percent of level of significance are also provided in the respective tables, which show that the considered distribution i.e. three parameter Weibull distribution is an appropriate choice for modelling of age at sexual initiation.

Table 7: Age at Sexual Initiation for all considered males

Age at sexual initiation	Age of the respondents on survey date					
	Less than 25 years		Between 25 and 30 years		Less than 30 years	
	Observed number of males	Expected number of males	Observed number of males	Expected number of males	Observed number of males	Expected number of males
13-15	5	9.88	5	9.13	10	18.72
15-18	78	68.98	56	71.30	134	216.04
18-21	250	250.59	328	334.37	578	543.79
21-24	178	183.37	345	342.10	523	553.39
24-27	13	11.18	270	268.76	283	221.79
27-30	-	-	56	34.34	56	30.27
Total	524	524.00	1060	1060.00	1584	1584.00
Value of K-S test statistics	0.031		0.024		0.057	
critical value at 5%	0.059		0.042		0.034	

Table 8: Age at Sexual Initiation for Urban males

Age at sexual initiation	Age of the respondents on survey date					
	Less than 25 years		Between 25 and 30 years		Less than 30 years	
	Observed number of males	Expected number of males	Observed number of males	Expected number of males	Observed number of males	Expected number of males
13-15	0	0.45	2	2.04	2	3.46
15-18	13	21.48	10	15.13	23	48.42
18-21	77	69.42	74	68.92	151	153.49
21-24	65	65.83	131	134.42	196	195.07
24-27	4	1.82	101	103.50	105	91.98
27-30	-	-	28	21.99	28	12.58
Total	159	159.00	346	346.00	505	505.00
Value of K-S test statistics	0.056		0.017		0.058	
critical value at 5%	0.108		0.073		0.061	

Table 9: Age at Sexual Initiation for Rural males

Age at sexual initiation	Age of the respondents on survey date					
	Less than 25 years		Between 25 and 30 years		Less than 30 years	
	Observed number of males	Expected number of males	Observed number of males	Expected number of males	Observed number of males	Expected number of males
13-15	5	9.25	3	6.20	8	14.39
15-18	65	54.47	46	57.42	111	169.99
18-21	173	175.35	254	247.93	427	389.77
21-24	113	122.65	214	223.74	327	356.08
24-27	9	3.28	169	158.49	178	131.09
27-30	-	-	28	20.22	28	17.68
Total	365	365.00	714	714.00	1079	1079.00
K-S test statistics	0.017		0.026		0.061	
critical value at 5%	0.071		0.051		0.041	

The estimates of the parameters involved in the considered distribution in this study to model age at sexual initiation are provided in Table 10. The scale parameter α is propensity or tendency of the phenomena and γ is threshold or location parameter which indicates the theoretical starting value of the phenomena considered. On fitting the three parameter Weibull distribution on age at sexual initiation for the respondents whose age on survey date was greater than 30 years, we observe that the propensity parameter α is more for urban respondents (10.42) than that for rural respondents (8.99), indicating that the tendency of sexual initiation age in urban areas was earlier than that in the rural areas.

Table 10: Estimates of parameters for the different groups

Parameters	Age of the respondents on survey date								
	Less than 25 years			Between 25 and 30 years			Less than 30 years		
	All	Urban	Rural	All	Urban	Rural	All	Urban	Rural
α (Propensity)	10.759	7.789	10.922	11.072	12.858	10.002	9.581	10.421	8.986
β (Shape)	5.508	4.325	5.407	3.861	4.779	3.422	3.312	3.750	3.082
γ (Threshold)	9.759	12.989	9.452	11.764	10.604	12.499	12.487	12.238	12.781

This finding is supported by the theoretical value of γ for the urban (12.24) and for rural (12.78). A similar result has also been observed when the age of respondents on survey date was 25-30 years; but in case of respondent who belonged to age group less than 25 years on survey date, the propensity is more for rural males than urban males. The estimated values of the threshold parameter support the above findings. It may also be noted that overall, value of threshold parameter is least for the age group of respondents less than 25 years and increases as the age of the respondent group increases. A similar trend can be noted for rural respondent also; but, for urban respondents, the least value of threshold parameter is noted for the age group 25-30 years. The estimate of propensity parameter for rural male is smallest for the age group more than 30 years and maximum for age group less than 25; but for urban males the highest estimate for the propensity is observed for the respondents' age group 25 to 30 years and the lowest for the age group less than 25 years. Hence one may infer from the present study that there is an indication of transition of Indian society from authoritarian to anarchic viewpoint and interestingly it is more for rural as compared to urban.

Conclusions

It is well known that variation in the age at first sexual initiation is affected by many socio-demographic factors. This study proves that the age at first sexual initiation is lowering down with the age of respondent at the time of survey. The age at sexual initiation is noted to be slightly lower in rural area than urban in India. These results are interesting because the education level is low in rural area in comparison of urban area. The age at first sexual initiation is observed earlier in younger men included in the study than older men. This shows that the attitude towards sexual matters is changing in India. Perhaps, people are adopting more liberal view towards it. It may further be concluded from the present study that the age at sexual debut can be modelled by Weibull probability law provided the data relates to the people whose ages are close to each other i.e. lie in short period of time. However, if the data is collected from people having their age spread over a large period of time, a single Weibull distribution may not be able to explain the phenomenon; thus, one can try some compound distributions assuming appropriate distribution for each of the parameters of the distribution. A future research is further needed to explore the impact of some of the covariates on age at sexual initiation and to know the possible factors responsible for early sexual initiation in the society.

References

Carpenter, L., 2005, *Virginity Lost: An intimate portrait of first sexual experiences*. New York: New York University Press.

Cohen, A. C., 1965, Maximum likelihood estimation in the Weibull distribution based on complete and censored samples. *Technometrics*, 7, 579-588.

Dickson N., Paul C., Herbison P., Silva P., 1998, First sexual intercourse: age, coercion, and later regrets reported by a birth cohort. *BMJ*, 316: 29–33.

Duncan M. E., Tibaux G., Pelzer A., Reimann K., Peutherer J. F., Simmonds P., Young H., Jamil Y., Daroughar S., 1990, First coitus before menarche and risk of sexually transmitted disease. *Lancet*, 335: 338–340.

Ghebremichael, M., Larsen, U. & Paintsil, E., 2009, Association of age at first sex with HIV-1, HSV-2 and other sexual transmitted infections among women in northern Tanzania. *Sexually Transmitted Diseases*, 36(9): 570.

Gonçalves H., Gonçalves Soares A. L., Bierhals I. O., Machado A. K. F., Fernandes M. P., Hirschmann R., Da Silva T. M., Wehrmeister F. C., Menezes A. M.B., 2017, Age of sexual initiation and depression in adolescents: Data from the 1993 Pelotas (Brazil) Birth Cohort. *Journal of Affective Disorder, 15:221*, 259-266.

Greenberg J., Magder L., Aral S., 1992, Age at first coitus. A marker for risky sexual behavior in women. *Sexually Transmitted Disease*, 19(6): 331-4.

Gregson S, Garnett G. P., Nyamukapa C. A., Hallett T. B., Lewis J. C. J., Mason P. R., Chandiwana S. K. Anderson M. R., 2006, HIV decline associated with behavior change in eastern Zimbabwe. *Science*, 311: 664–66.

Harter, H. L. and Moore A. H., 1965a, Point and interval estimators based on order statistics, for the scale parameter of a Weibull population with known shape parameter. *Technometrics, Vol. 7*.

Harter, H. L. and A. H. Moore, 1965b, Maximum likelihood estimation of the parameters of Gamma and Weibull populations from complete and from censored samples, *Technometrics*, Vol. 7, No. 4.

Jejeebhoy, S.J., 1998, Adolescent sexual and reproductive behaviour: a review of the evidence from India. *Social Science & Medicine,* 46(10): 1275-1290.

Kaestle, C., Halpern, C., Miller, W., & Ford, C., 2005, Young age at first sexual intercourse and sexually transmitted infections in adolescents and young adults. *American Journal of Epidemiology,* 161: 774-780.

Lawless J.F., 2002, *Statistical models and methods for lifetime data*. Wiley series in probability and statistics, second edition, New York: John Wiley and sons Inc.

Luke, N., 2003, Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa. *Studies in Family Planning,* 34(2): 67–86.

Manning W. D, Longmore M.A., Giordano P. C., 2000. The relationship context of contraceptive use at first intercourse. *Family Planning Perspective,* 32(3): 104-110.

Marks N.B., 2005, Estimation of Weibull parameters from common percentiles. *Journal of Applied Statistics,* 32 (1): 17-24.

Maticka-Tyndale, E., Gallant, M., Brouillard-Coyle, C., Holland, D., Metcalfe, K., Wildish, J. and Gichuru, M., 2005, The sexual scripts of Kenyan young people and HIV prevention. *Culture, Health & Sexuality,* 7(1): 27–41.

Musie S. Ghebremichael, and Matthew D. Finkelman, 2013, The Effect of premarital sex on sexually transmitted infections (STIs) and high-risk behaviors in women. *Journal of AIDS and HIV Research,* 5(2): 59–64.

Nag, M., 1995, Sexual behaviour in India with risk of HIV/AIDS transmission. *Health Transition Review,* 5: 293-305.

Nanda Priya, Gautam Abhishek, Verma Ravi, Khanna Aarushi, Khan Nizamuddin, Brahme Dhanashri, Boyle Shobhana and Kumar Sanjay, 2014, *Study on Masculinity, Intimate Partner Violence and Son Preference in India*. New Delhi: International Center for Research on Women.

O'Donnell L, O'Donnell C. R and Stueve A., 2001, Early sexual initiation and subsequent sex-related risks among urban minority youth: the reach for health study. *Family Planning Perspective,* 33(6): 268–275.

Pelto, P.J., 1999, *Sexuality and sexual behaviour: the current discourse,* in S. Pachauri (ed.) Implementing a Reproductive Health Agenda in India: The Beginning, New Delhi: Population Council.

Singh, Brijesh P. and Tapan Kumar Roy, 2015, Timing of Sterilization: An Application of Weibull Model Under MCMC Approach in *Indian Journal of Statistics and Application*, 4(1): 6-17.

Singh, K. K, Brijesh P. Singh & Tapan K. Roy, 2010, A Probability Model for Age at Sterilization in Bangladesh. *Journal of Empirical Research in Social Science,* 5(1-2): 58-70.

Singh, K. K., Bloom S. S., and Tsui A. O., 1998, Husbands' reproductive health knowledge, attitudes, and behavior in Uttar Pradesh, India. *Studies in Family Planning,* 29(4): 388-399.

Stoneburner R., Low-Beer D., 2004, Population-level HIV declines and behavioral risk avoidance in Uganda. *Science.* 304: 14–18.

Tenkorang, E. Y. & Maticka-Tyndale, E., 2008, Factors influencing the timing of first sexual intercourse among young people in Nyanza, Kenya. *International Family Planning Perspectives,* 34(4): 177–188.

Vlassof, C., 1987, *Awareness and assess to reproductive care among adolescents*. Retrieved from http://iussp2009.princeton.edu/papers/91518.

White, R., J. Cleland and Carael M., 2000, Links between premarital sexual behaviour and extramarital intercourse: a multi-site analysis. *AIDS, 14:* 2323-2331.

Yode, M. & LeGrand, T., 2012, Association between age at first sexual relation and some indicators of sexual behaviour among adolescents. *African Journal of Reproductive Health,* 16(2): 173–188.