
Use of Distribution Function in Estimation of Mean of Some Demographic Duration Variables

R. C. Yadava¹, Abhay Kumar Tiwari¹, Vaishali Patel²
and Pappu Kumar Singh^{1,*}

¹*Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi-221005, Uttar Pradesh, India*

²*Department of Statistics, Meerut College, Meerut-250002, Uttar Pradesh, India*
E-mail: singhpappukumar63@gmail.com

**Corresponding Author*

Received 12 September 2024; Accepted 29 March 2025

Abstract

Life testing, Survival analysis and Reliability theory, although highly inter-related, are fascinating areas of statistical research. Here, the random variable, normally known as 'Life Time' of the equipment, is a duration variable which takes only non-negative values. Some examples have been given in the study to demonstrate the similarities between life testing and some demographic phenomena. Several procedures exist for analysing the nature of demographic duration variables such as life expectancy at birth, age at marriage, post-partum amenorrhea, breastfeeding etc. The objective of the present paper provides the procedure that how the mean of some demographic duration variables (life expectancy at birth, age at marriage, postpartum amenorrhea, breastfeeding etc.) can easily be obtained using distribution function? This method is straightforward because the required data to compute the mean is easily obtained without any misreporting, recall lapses or digit preferences of durational variables. For this study, we have considered data from UN

Journal of Reliability and Statistical Studies, Vol. 18, Issue 1 (2025), 189–212.

doi: 10.13052/jrss0974-8024.1818

© 2025 River Publishers

population division (Economic and Social Affairs) and national family health survey (NFHS). Results indicate that the estimated expectancy of life at birth, mean age at marriage, average duration of postpartum amenorrhea and breastfeeding using mentioned procedure are close to observed value.

Keywords: Reliability, distribution function, life expectancy, postpartum-amenorrhea, age at marriage, breastfeeding.

1 Introduction

Life testing, Survival analysis, and Reliability theory, although highly inter-related, are fascinating areas of statistical research. The basic concept behind the topic is to study various properties of a random variable X representing the life time of an item or equipment or any object which is subjected to failure. In most of the studies it is assumed that the item (equipment) will definitely fail in a time X which is a random variable [1–5]. The random variable X is normally known as ‘Life Time’ of the equipment. Obviously, X is a duration variable which takes only non-negative values. Although, it is usually assumed that the equipment will definitely fail after some time but this assumption may be relaxed for some situations. Of course, under the relaxed conditions, methods of analysis may require modifications.

In life testing or reliability theory, the major functions of interest are:

1. $f(x)$, the probability density function of random variable X .
2. $F(x)$, the distribution function of X , i.e. $F(x) = P[X \leq x]$.
3. $1 - F(x) = R(x)$, the survival function or reliability.
4. $\mu(x)$, the hazard rate at x representing the conditional probability that the equipment will fail in time interval $(x, x + \Delta x)$ as $\mu(x)\Delta x + O(\Delta x)$.

The results of reliability theory can also be extended to ‘Renewal Process’ and results of renewal theory can be extensively used to study the various properties of the process [6]. As an example, suppose X represent the life time of a bulb. The bulb is put for testing. After a random interval of time, the bulb will fail. Put another similar bulb for testing. It will also fail after some time. Then put another similar bulb for testing. It will also fail after some random interval of time and repeat the experiment till time T . Suppose the consecutive intervals of failures are represented as X_1, X_2, X_3, \dots then, the process become a renewal process and we can study various properties of X_1, X_2, X_3, \dots and we can also study the probability distribution of number of renewals in $(0, T)$.

Although, we have mentioned the four major functions of interest in study of reliability theory but they are not independent. In fact, these are highly related and knowing one, the other three can be easily obtained. For example,

1. If $f(x)$ is known, then
 - I. $F(x) = \int_0^x f(t)dt$
 - II. $F(x)$ is known, then $1 - F(x)$ can be easily computed.
 - III. $\mu(x) = \frac{f(x)}{1-F(x)}$ which can be easily computed.
2. If $F(x)$ is known
 - I. $1 - F(x)$ can be easily computed.
 - II. We know that $\frac{d}{dx}F(x) = f(x)$, hence $f(x)$ can be easily obtained.
 - III. $\mu(x) = \frac{f(x)}{1-F(x)}$ which can be easily obtained.
3. $1 - F(x) = S(x)$ is known
 - I. $F(x)$ can be easily found.
 - II. $f(x) = \frac{d}{dx}F(x)$ which can be easily obtained.
 - III. $\mu(x) = \frac{f(x)}{1-F(x)}$ which can be easily obtained.
4. $\mu(x)$ is known.
 - I. $1 - F(x) = e^{-\int_0^x \mu(t)dt}$ which can be easily obtained.
 - II. $1 - F(x)$ is known, hence $F(x)$ can be easily obtained.
 - III. $f(x) = \frac{d}{dx}F(x)$, which can be easily obtained.

We have mentioned different concepts of reliability theory. The mentioned results are very simple and widely known. However, we have mentioned these concepts and results only to demonstrate the similarities between reliability theory and results of different demographic processes such as mortality, morbidity and fertility.

Now, we give some examples to demonstrate the similarities between life testing and some demographic phenomena. The first example is the study of 'Life Table' in demography and life time distribution in life testing. In fact, in life testing mostly we try to find the average life time of an equipment, in demography we want to find the average life time of a person (human being) [7]. The average life time of a person is usually called as expectation of life at birth generally denoted by e_0^0 [8]. It is to be mentioned that there are some dissimilarities also between study of life time distribution in life testing and human life table study in demography. Generally, in life testing the random variable X is considered to be continuous and its probability density function $f(x)$ is assumed to be known (with unknown parameter(s)) [7].

In the context of life table, although the life time of a person (human) is definitely continuous but normally the p.d.f. of X is not available in explicit functional form. Similarly, the value of $\mu(x)$ is also not available in any explicit functional form. As a result, instead of considering X to be continuous, we consider it discrete, X taking values $0, 1, 2, \dots, \omega^+$. In this context, l_x is analogous to $1 - F(x)$, (when $l_o = 1$, since l_o is arbitrary, hence we can also take the value as l_o in that case $\frac{l_x}{l_o} = 1 - F(x)$) while q_x is analogous to $\mu(x)$. $E(X)$ are analogous to e_0^0 [9]. We will discuss later on that how $E(X)$ is obtained using the concept of Prevalence/Incidence?

The concept of life table is also useful in studying the distribution of time of first conception from marriage (or start of sexual union) [10]. The concept of both the situations are same. In case of life testing, the event of interest is the occurrence of failures of the equipment/item while in the second case, the event of interest is the occurrence of first conception [4]. In the context of morbidity, one may be interested in the study of time of cure from the time of occurrence of a particular disease. Here the concepts of life table are useful in the estimation of average duration of cure. In case of repeated occurrence of disease, the concept of renewal theory can be applied. In the context of fertility studies, repeated occurrences of birth to a female may be considered analogous to repeated renewals. One can study about the number of births in a given time interval or study about the time between consecutive births (renewals), i.e. birth intervals [11].

To obtain some specific objective(s) of a study, we need appropriate methodology. So, researchers have developed appropriate methodologies for different situations. Generally, any research which mainly deals with development of appropriate methodology comes under the purview of theoretical research while any methodology applied to real data to obtain specific results, comes under the purview of applied research. However, it is to be emphasised that it is very difficult to draw a specific line between theoretical and applied research. Whether the theoretical or applied research, the methodologies are so developed that the data requirements are easy as well as data are reasonably reliable. This is truer for the case of applied research.

It is to be mentioned that for attaining an objective, different methodologies can be used where data needs are different. We may prefer that methodology where data requirement is simple such that reliable data can easily be obtained.

Let us consider a simple example in the context of Bernoulli trials. We want to estimate the value of p , representing the probability of success in a trial. There are two well-known approaches for estimation of p . In one

case, we may note the number of successes in n trials (n fixed). Repeat the experiment N times, note the number of successes in N experiments. Find $\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_N}{N}$, equating it to its theoretical value np so that the estimate of p say \hat{p} is obtained as $\hat{p} = \frac{\bar{X}}{n}$.

Alternatively, suppose X represents the number of trials required for getting first success. Obviously, X has a geometric distribution with probability mass function (pmf)

$$P[X = x] = pq^{x-1}; \quad x = 1, 2, 3, \dots; \quad q = 1 - p.$$

obviously, $E(X) = \frac{1}{p}$.

Thus, if we have a random sample of size n (repeating the experiment n times) and we get the value of \bar{X} , then estimate of p say \hat{p} is obtained as $\hat{p} = \frac{1}{\bar{X}}$. Thus, the estimate of p can be obtained from both the approaches but the data requirements in both cases are entirely of different nature. In one case, the number of trials is fixed and number of successes is random variable while in the other case, the number of success (i.e. first success) is fixed while number of trials required is random variable. In simple hypothetical trials such as coin tossing, it is almost immaterial whether we use the first approach or second approach, but in the context of real situations such as in demography, the two approaches are analogous to number of births in a given time interval and the birth intervals between consecutive births. Obviously, data needs of the two approaches are entirely different. Researchers have developed appropriate methodologies for both the approaches according to availability of data, especially for secondary data. We may know that collecting data on number of births in a given interval of time, is perhaps easier than to collect data on birth intervals. However, both the approaches have their own advantages and disadvantages and the researchers use any one of them according to the objective of study.

2 Methodology

2.1 Estimation of Mean of Random Variable

The mean of a random variable is usually defined as

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad (1)$$

where $f(x)$ denotes the probability density function of the random variable X . This is true if the random variable is continuous (absolutely continuous).

However, if the random variable is discrete, then the $E(X)$ is defined as

$$E(X) = \sum x f(x) dx \quad (2)$$

provided the expectation exists.

This theoretical value of mean can be obtained only if the p.d.f. of random variable or p.m.f. of random variable (as the case may be) is known. In case, the p.d.f. is not available, then although the theoretical mean cannot be obtained but mean can be obtained on the basis of observed data represented by a frequency distribution of the given data.

However, if a random variable takes only non-negative values, then $E(X)$ can also be obtained as

$$E(X) = \int_0^{\infty} [1 - F(x)] dx \quad (3)$$

where $F(x)$ is the distribution function of the random variable i.e. $P[X \leq x]$.

This result is true whether the random variable is continuous or discrete [12, 13]. They clearly specify that the above result is true for non-negative continuous as well as for discrete random variables.

This is also true when X is discrete and X takes values at unequal intervals [14]. It is a known fact that the duration variables especially in demography and reliability theory the random variables take only non-negative values. So, the above result can be easily applied for the estimation of mean of a given duration variable. It is to mentioned that, many times, it is easier to obtain data on $[1 - F(x)]$ rather than the data on $f(x)$. It is also true that the data on $1 - F(x)$ are likely to be more reliable than data on $f(x)$.

The objective of the present article is to provide some examples that demonstrate how the mean of a duration random variable can be easily determined using the concept of $\int_0^{\infty} [1 - F(x)] dx = E(X)$? We now give examples one by one from different topic of demography and try to provide the mean of considered demographic duration variables.

2.2 Data

The data for this has been taken from the fifth round of the National Family Health Survey (NFHS-5) and UN Population Division, Department of Economic and Social Affairs (DESA). The fifth round of National Family Health Survey was conducted between 2019 and 2021 under the auspices of the Ministry of Health and Family Welfare (MoHFW), Government of India and International Institute of Population Sciences (IIPS) in Mumbai.

The National Family Health Survey (NFHS) is a multi-round, large-scale survey conducted in a nationally representative sample of households. The survey collected data on birth history of women, infant and child mortality, fertility, breastfeeding, postpartum amenorrhea, age at marriage, maternal and child health, reproductive health, anaemia, nutrition and family planning at the national, state and district levels in India [18]. The dataset is publicly available at <https://dhsprogram.com/data/>.

The United Nations Population Division, Department of Economic and Social Affairs (DESA), provide data on fertility, mortality, migration and urbanization at the global, regional and national levels. The essential data for computation of expectation of life at birth of women of India is obtained from World Population Prospects 2022, UN Population Division, Dept. of Economic and Social Affairs [17].

2.3 Estimation of Mean of Some Demographic Duration Random Variable

2.3.1 Expectation of life at birth (e_0^0) in Life Table (An important topic in Demography)

In fact, the expectation of life at birth in the life table essentially defines the average life time of a person. It is analogous to average life time of equipment in life testing or reliability theory.

As mentioned earlier, although the life time X is definitely continuous but because of the non-availability of the values of $\mu(x)$ (Hazard Rate) or $[1 - F(x)]$ i.e. survival function, hence for practical purposes X is taken as discrete random variable. Normally, in complete life table, the age x denotes the age in completed years or age at last birthday which normally takes values $0, 1, 2, 3, \dots, \omega^+$ [9]. In the context of abridged life table, age class intervals are more than one year. In the present case we will assume that the ages are given in completed years i.e. X takes values $0, 1, 2, \dots$

In fact, for complete life table, e_0^0 i.e. expectation of life at birth is defined as

$$e_0^0 = \frac{T_0}{l_0} \quad (4)$$

$$i.e. \quad e_0^0 = \frac{L_0 + L_1 + L_2 + \dots + L_{\omega^+}}{l_0}; \quad \text{where } T_0 = \sum_{x=0}^{\omega^+} L_x \quad (5)$$

where T_0 , is the total number of years lived by the cohort (l_0) at birth and l_0 is the total number of births.

In fact, the pivotal column of the life table is the q_x column which represents the conditional probability that a person who has reached age x (exact age x) will die before reaching age $x + 1$ (exact age $(x + 1)$) [15]. This is almost analogous the hazard rate at age x treating X to be continuous. μ_x is normally known as the force of mortality at age x . In mathematical terms, q_x is defined as,

$$q_x = \frac{l_x - l_{x+1}}{l_x} \quad \text{or} \quad = \frac{d_x}{l_x} \quad (6)$$

where d_x denotes the number of deaths between exact ages x and $x + 1$ in the life table. In fact, $e_0^0 = \frac{T_0}{l_0}$ has been obtained by using the concept of person years lived by the cohort divided by the radix l_0 . As we know, l_x denotes the number of survivors till exact age x out of l_0 births. If we take X to be continuous then $\frac{l_x}{l_0}$ denotes nothing but $1 - F(x)$ i.e. the probability of surviving from birth to age x or $P[X > x]$. Thus,

$$\frac{l_x}{l_0} = 1 - F(x) \quad (7)$$

In fact, $\frac{T_0}{l_0}$ is nothing but $\int_0^\infty [1 - F(x)]dx$ because, $\frac{T_0}{l_0}$ may be considered as $\int_0^\infty \frac{l_x}{l_0} dx$. Thus, $\frac{T_0}{l_0}$ is nothing but $\int_0^\infty [1 - F(x)]dx$. Thus, e_0^0 is nothing but $\int_0^\infty [1 - F(x)]dx = \text{Mean}$. In this context, it is to be mentioned that, L_x column of the life table gives the age distribution of the life table population (stationary population) [16]. The stationary population has the property that its rate of growth is zero and age structure remains constant over time implying that birth rate is equal to death rate.

In fact, in life table population, every year l_0 births take place while the total population is T_0 . Thus, the birth rate in life table population (stationary population) is nothing but $\frac{l_0}{T_0}$ which is reciprocal of e_0^0 i.e. birth rate = $\frac{1}{e_0^0}$. Since the rate of growth is zero, this implies as the death rate is also $\frac{1}{e_0^0}$.

In the context of prevalence and incidence, T_0 represents the prevalence while l_0 represents the incidence rate per unit of time. So, in fact, $\frac{T_0}{l_0}$ is nothing but prevalence/incidence and in this context the prevalence/incidence become the mean of X which is commonly known as the prevalence/incidence mean [14].

2.3.2 Mean age at marriage (first marriage)

The concept of mean = $\int_0^\infty [1 - F(x)]dx$ can also be applied for estimation of mean age at first marriage (we do not consider the case of repeated

marriages to avoid the complexity in the computation of $1 - F(x)$). In fact, if we consider the number of persons of age x as n_x in a population out of which n_{x_m} are married while $n_{x_{um}}$ are unmarried. If X denotes the random variable as age at marriage then, $\frac{n_{x_m}}{n_x}$ gives an estimate of $F(x)$ because here age at marriages is below x years i.e. $X < x$. Similarly, the proportion $\frac{n_{x_{um}}}{n_x}$ gives the estimate of $[1 - F(x)]$. But $\frac{n_{x_{um}}}{n_x}$ is nothing but proportion of unmarried among persons aged x .

So, if we have data on proportions of unmarried at different ages, which give the values of $[1 - F(x)]$ for different values of x , then $\int_0^\infty [1 - F(x)]dx$ gives the mean age at marriage. We illustrate the use of the methodologies using data of NFHS-5. Fortunately, the data on age at marriage of females between age group 15–49 are available in this survey. So, we can use this methodology for computing mean age at marriages of females who are currently in the age group 15–49 years.

Since the data are not available for females below age 15 years so we assume that the incidence of marriage at below age 15 years is zero (almost zero). Although X is a continuous random variable but data are only available at discrete ages. It is known fact that the number of females at each age is not equal. Hence, theoretically, the concept of prevalence/Incidence equal to mean cannot be applied. However, in practice, there is not much variation in the number of females at different ages in the age group 15–49 years. So, we can get, at least an approximate estimate of mean age at marriage by the method of prevalence/ Incidence. The incidence may be taken as the average number of females per year in the age group 15–49 years.

2.3.3 Average duration of postpartum amenorrhea (PPA)

Based on similar approach, the method has also been used for estimation of mean postpartum amenorrhea (PPA). In NFHS-5, apart from other information; data are available related to PPA status for births occurring during the last 36 months from the date of survey [18]. In this context, data are available on the number of births in each month and the number of females who are still amenorrhoeic at the date of survey out of the births in each month.

If we consider the number of women whose current age of child is x as n_x in a population out of which $n_{x_{np}}$ are those women whose period is not return i.e. still amenorrhoeic at the time of survey while n_{x_p} are those women whose period is return before the survey date. Although X is a continuous random variable but data are only available at discrete ages.

Suppose the random variable X represents the PPA period. Obviously X can take only nonnegative values. Suppose a female has given birth x period

before the survey point and she is in PPA period at the survey point, then this implies that for her $X > x$, since $P[X > x] = 1 - F(x)$ then the proportion $\frac{n_{xnp}}{n_x}$ (proportion of women continued for PPA of duration x) gives the estimate of $[1 - F(x)]$ for different values of x , then $\int_0^\infty [1 - F(x)]dx$ gives the mean postpartum amenorrhea.

2.3.4 Average duration of breastfeeding

The concept of mean $= \int_0^\infty [1 - F(x)]dx$ can also be applied for estimation of mean duration of breastfeeding. Now, if we consider the number of mothers whose current age of child is x as n_x in a population out of which n_{x_b} are those mothers who are continue to breastfeeding while $n_{x_{sb}}$ are those who stopped breastfeeding.

If X denotes the random variable as duration of breastfeeding then it can take only non-negative values. Suppose a female has given birth x period before the survey point and she is continue for breastfeeding at the survey point, then this implies that for her $X > x$, since $P[X > x] = 1 - F(x)$ then the proportion $\frac{n_{x_b}}{n_x}$ (proportion of mothers still breastfeeding of duration x) gives the estimate of $[1 - F(x)]$ for different values of x , then $\int_0^\infty [1 - F(x)]dx$ gives the mean breastfeeding. Obviously, $\frac{n_{x_{sb}}}{n_x}$ gives an estimate of $F(x)$ because here duration of breastfeeding is before x months i.e. $X < x$.

3 Results and Discussion

Table 1 represents the exact age of women, the number of women (survivors) till exact age x and proportion of survivors from birth to exact age x has given in third column of this table. The last column of table shows the smoothed of proportion of survivors from birth to exact age x . For finding the integral $\int_0^\infty [1 - F(x)]dx$ we may use an appropriate quadrature formula. For the given situation the quadrature formula, which is the simplest, seems to be appropriate so we make use of Trapezoidal rule to find the above integral. The life expectancy at birth has been determined to be 68.89 years using the trapezoidal rule to calculate the integral $\int_0^\infty \frac{l_x}{l_0} dx$, which is equivalent to $\int_0^\infty [1 - F(x)]dx$. Since $\int_0^\infty [1 - F(x)]dx$ represents the mean of the random variable X , denoting the age of survivors. This calculated value (68.89 years) is remarkably close to the observed life expectancy at birth for Indian women, which is reported as 68.9 years in the World Population Prospects, 2022 [19].

Table 2 provides essential data for computing the mean age at marriage for women in India. This data is taken from the fifth round of National Family Health Survey. The table shows the proportion of unmarried women and the smoothed proportion of unmarried women at different ages. Now the integral $\int_0^\infty [1 - F(x)]dx = \int_0^c [1 - F(x)]dx$ only for such value of c , where $[1 - F(c)] = 0$. In case of women age at marriage (regarding fertility phenomenon), the value of $c = 49$ years seems to be quite reasonable because the probability that age at marriage will be more than 49 years is almost zero. The value of $\int_0^c [1 - F(x)]dx$ has been determined to 22.02 by trapezoidal rule approximation by taking $F(x) = 0$ for $x < 15$. Thus, employing the data from NFHS-5 and utilizing the proposed method, the estimate of mean age at marriage for Indian women has been precisely calculated to be 22.02 years.

Table 3 provides essential data for Computation of Average Duration of Postpartum Amenorrhea (NFHS-5). This table represent the proportion of women whose period not return till survey date and having child of age x months before survey. Through the results, it is found that out of 3653 women surveyed, 133 women are still being in PPA period after 36 months of given the last birth before the survey date. This is perhaps due to miss-reporting in PPA data.

Now the integral $\int_0^\infty [1 - F(x)]dx$ is equivalent to $\int_0^c [1 - F(x)]dx$ only for such value of c , where $[1 - F(c)] = 0$. If we assume that the maximum possible value of X is 36 months, the value of $c = 36$ months seems to be quite reasonable because the probability that PPA will be more than 36 months is almost zero. The value of $\int_0^c [1 - F(x)]dx$ has been calculated to 6.24 months by trapezoidal rule approximation. Thus, employing the data from NFHS-5 and utilizing the proposed method, the estimate of mean PPA for Indian women has been calculated to be 6.2 months which is very close to reported mean PPA as 6.9 months in NFHS-5 report [18].

We illustrate the use of given methodologies for breastfeeding, using NFHS-5 data. Table 4 provide the relevant data for computation of mean breastfeeding. This table represent the proportion of women continue to breastfeeding till survey date and having child of age x months before survey. Through the results, it is found that out of 3653 women, 239 women are still being for breastfeeding after 60 months of given the last birth before the survey date. This anomaly is likely due to misreporting in the breastfeeding data.

Now the integral $\int_0^\infty [1 - F(x)]dx$ is equivalent to $\int_0^c [1 - F(x)]dx$ only for such value of c , where $[1 - F(c)] = 0$. Here, we assume that the

maximum possible value of X is 60 months, the value of $c = 60$ months seems to be quite reasonable because the probability that breastfeeding will be more than 60 months is almost zero. The value of $\int_0^c [1 - F(x)]dx$ has been calculated to 31.84 months by trapezoidal rule approximation. Thus, employing the data from NFHS-5 and utilizing the proposed method, the estimate of mean breastfeeding for Indian women has been calculated to be 31.8 months which is very close to observed value of median breastfeeding as 32.1 months reported in NFHS-5 report [18].

4 Conclusion

Several procedures exist for analysing the nature of demographic duration variables such as for estimation of expectation of life at birth, Asaria et al. [20] used a 'life table technique' to estimate life expectancy at birth but in this procedure, they required additional information on survival ratio, mortality rate and average size of cohort etc. whereas the proposed method require information on only proportion of survivors. For estimation of mean age at first marriage, Singh et al. [21] used 'semi-parametric' and 'non-parametric' technique but this procedure may contain some error related to recall bias which leads to age misreporting and digit preference, especially among older and less educated respondents whereas in the proposed procedure for estimation of mean age at first marriage, such type of errors may not exist. For estimation of PPA, the present methodology requires data only on proportion of women whose periods do not return whereas the methodologies proposed by Srinivasan [22] and Yadava and Bhattacharya [23] require data on open as well as closed birth intervals. In this study, we have estimated mean duration of such as life expectancy at birth, age at first marriage, postpartum amenorrhea and breastfeeding. The proposed methods differ in their data requirements and in regard to the objective of this analysis. If one is interested only in the mean of such demographic variables, a very simple method, as proposed by the author, can be used. This method is straightforward because the required data to compute the mean is easily obtained without any misreporting or recall lapses or digit preference in durational variables. In contrast, earlier approaches are based on durational data that often contain errors due to misreporting and recall lapses.

Appendix

Table 1 Computation of expectation of life at birth

Age (x)	No. of Survivors at Age x	Proportion of Survivors From Birth to Age x	Smoothed Proportion of Survivors From Birth
0	100000	1.00000	0.99061
1	97418	0.97418	0.98274
2	97117	0.97117	0.97660
3	97006	0.97006	0.97196
4	96922	0.96922	0.96864
5	96837	0.96837	0.96644
6	96755	0.96755	0.96515
7	96685	0.96685	0.96458
8	96626	0.96626	0.96452
9	96580	0.96580	0.96477
10	96542	0.96542	0.96513
11	96507	0.96507	0.96541
12	96470	0.96470	0.96543
13	96428	0.96428	0.96520
14	96378	0.96378	0.96473
15	96316	0.96316	0.96406
16	96240	0.96240	0.96321
17	96153	0.96153	0.96221
18	96056	0.96056	0.96109
19	95953	0.95953	0.95988
20	95845	0.95845	0.95860
21	95733	0.95733	0.95728
22	95616	0.95616	0.95596
23	95496	0.95496	0.95464
24	95374	0.95374	0.95334
25	95248	0.95248	0.95203
26	95119	0.95119	0.95072
27	94986	0.94986	0.94940
28	94849	0.94849	0.94806
29	94706	0.94706	0.94670
30	94559	0.94559	0.94530
31	94407	0.94407	0.94385

(Continued)

Table 1 Continued

Age (x)	No. of Survivors at Age x	Proportion of Survivors From Birth to Age x	Smoothed Proportion of Survivors From Birth
32	94251	0.94251	0.94236
33	94090	0.94090	0.94081
34	93923	0.93923	0.93919
35	93747	0.93747	0.93749
36	93559	0.93559	0.93568
37	93354	0.93354	0.93373
38	93130	0.93130	0.93163
39	92883	0.92883	0.92934
40	92613	0.92613	0.92686
41	92324	0.92324	0.92415
42	92018	0.92018	0.92119
43	91700	0.91700	0.91796
44	91369	0.91369	0.91444
45	91021	0.91021	0.91059
46	90651	0.90651	0.90638
47	90246	0.90246	0.90177
48	89794	0.89794	0.89671
49	89279	0.89279	0.89115
50	88686	0.88686	0.88506
51	88004	0.88004	0.87838
52	87231	0.87231	0.87108
53	86372	0.86372	0.86311
54	85434	0.85434	0.85442
55	84424	0.84424	0.84498
56	83348	0.83348	0.83473
57	82213	0.82213	0.82364
58	81020	0.81020	0.81166
59	79757	0.79757	0.79875
60	78410	0.78410	0.78487
61	76963	0.76963	0.76997
62	75403	0.75403	0.75401
63	73722	0.73722	0.73696
64	71917	0.71917	0.71876
65	69983	0.69983	0.69937
66	67924	0.67924	0.67876

(Continued)

Table 1 Continued

Age (x)	No. of Survivors at Age x	Proportion of Survivors From Birth to Age x	Smoothed Proportion of Survivors From Birth
67	65744	0.65744	0.65689
68	63449	0.63449	0.63383
69	61042	0.61042	0.60966
70	58523	0.58523	0.58445
71	55893	0.55893	0.55828
72	53158	0.53158	0.53122
73	50326	0.50326	0.50336
74	47414	0.47414	0.47478
75	44444	0.44444	0.44554
76	41435	0.41435	0.41574
77	38409	0.38409	0.38544
78	35385	0.35385	0.35476
79	32370	0.32370	0.32394
80	29369	0.29369	0.29327
81	26400	0.26400	0.26302
82	23483	0.23483	0.23348
83	20649	0.20649	0.20493
84	17902	0.17902	0.17765
85	15260	0.15260	0.15192
86	12781	0.12781	0.12803
87	10532	0.10532	0.10624
88	8569	0.08569	0.08685
89	6899	0.06899	0.07005
90	5492	0.05492	0.05571
91	4319	0.04319	0.04364
92	3352	0.03352	0.03362
93	2563	0.02563	0.02546
94	1929	0.01929	0.01893
95	1428	0.01428	0.01384
96	1040	0.01040	0.00997
97	745	0.00745	0.00713
98	523	0.00523	0.00511
99	361	0.00361	0.00369
100 & above	244	0.00244	0.00268

Table 2 Computation of mean age at marriage (NFHS5)

Age (x)	No. of Women	No. of Unmarried Women	Proportion of Unmarried Women	Smoothed Proportion of Unmarried Women
15	25005	24794	0.99156	0.98800
16	24714	24126	0.97621	0.99233
17	23580	22124	0.93825	0.92261
18	26631	21935	0.82366	0.81581
19	22550	15741	0.69805	0.70334
20	26061	15123	0.58029	0.59429
21	21876	10915	0.49895	0.49218
22	24912	9821	0.39423	0.40052
23	22976	7577	0.32978	0.32201
24	22875	6068	0.26527	0.25609
25	28402	5438	0.19147	0.20138
26	23254	3690	0.15868	0.15651
27	22134	2660	0.12018	0.12017
28	25502	2173	0.08521	0.09141
29	19087	1437	0.07529	0.06937
30	29221	1467	0.05020	0.05318
31	15512	705	0.04545	0.04189
32	22612	774	0.03423	0.03426
33	16716	444	0.02656	0.02896
34	16988	416	0.02449	0.02485
35	27805	626	0.02251	0.02161
36	18233	357	0.01958	0.01912
37	16073	312	0.01941	0.01725
38	21121	271	0.01283	0.01588
39	14836	195	0.01314	0.01492
40	24912	439	0.01762	0.01425
41	12227	168	0.01374	0.01379
42	17459	227	0.01300	0.01344
43	14091	161	0.01143	0.01311
44	12691	176	0.01387	0.01271
45	24333	315	0.01295	0.01215
46	13982	153	0.01094	0.01138
47	13928	132	0.00948	0.01055
48	17797	196	0.01101	0.00987
49	14019	129	0.00920	0.00954

Table 3 Computation of average duration of postpartum amenorrhea (NFHS5)

Child's Current Age in Months (x)	Women Having Child of Age x	Women Whose Period Not Return, Having Child of Age x	Proportion of Women Whose Period Not Return	Smoothed Proportion of Women Whose Period Not Return
0	3653	2883	0.78921	0.79228
1	3963	2886	0.72824	0.72145
2	4081	2582	0.63269	0.63543
3	4027	2206	0.54780	0.54911
4	4185	1979	0.47288	0.47444
5	3792	1567	0.41324	0.41169
6	3974	1458	0.36688	0.35824
7	3695	1091	0.29526	0.31144
8	3525	957	0.27149	0.26909
9	3696	879	0.23782	0.23082
10	3602	728	0.20211	0.19670
11	3739	602	0.16101	0.16679
12	4048	566	0.13982	0.14113
13	3989	475	0.11908	0.11961
14	3736	365	0.09770	0.10208
15	3537	330	0.09330	0.08840
16	3726	300	0.08052	0.07831
17	3525	245	0.06950	0.07116
18	3571	238	0.06665	0.06617
19	3252	201	0.06181	0.06258
20	3194	188	0.05886	0.05973
21	3109	174	0.05597	0.05739
22	3227	178	0.05516	0.05544
23	3024	185	0.06118	0.05377
24	3245	172	0.05300	0.05228
25	3415	155	0.04539	0.05091
26	3078	132	0.04288	0.04967
27	3118	158	0.05067	0.04853
28	3092	158	0.05110	0.04747
29	3007	146	0.04855	0.04650
30	2844	145	0.05098	0.04561
31	2793	121	0.04332	0.04481
32	2652	93	0.03507	0.04436
33	2422	105	0.04335	0.04557
34	2533	152	0.06001	0.05001
35	2389	133	0.05567	0.05922

Table 4 Computation of average duration of breastfeeding

Child's Current Age in Months (x)	Women Having Child of Age x	Women Still Breastfeeding, Having Child of Age x	Proportion of Women Still Breastfeeding	Smoothed Proportion of Women Still Breastfeeding
0	3653	3383	0.92609	0.92613
1	3963	3658	0.92304	0.92004
2	4081	3744	0.91742	0.91728
3	4027	3702	0.91929	0.91662
4	4185	3802	0.90848	0.91684
5	3792	3490	0.92036	0.91670
6	3974	3607	0.90765	0.91499
7	3695	3333	0.90203	0.91075
8	3525	3187	0.90411	0.90416
9	3696	3340	0.90368	0.89570
10	3602	3245	0.90089	0.88582
11	3739	3311	0.88553	0.87500
12	4048	3468	0.85672	0.86368
13	3989	3388	0.84934	0.85235
14	3736	3115	0.83378	0.84127
15	3537	2952	0.83461	0.83004
16	3726	3052	0.81911	0.81805
17	3525	2774	0.78695	0.80471
18	3571	2805	0.78549	0.78942
19	3252	2509	0.77153	0.77157
20	3194	2393	0.74922	0.75083
21	3109	2290	0.73657	0.72783
22	3227	2334	0.72327	0.70349
23	3024	2089	0.69081	0.67869
24	3245	2090	0.64407	0.65435
25	3415	2124	0.62196	0.63136
26	3078	1892	0.61468	0.61062
27	3118	1794	0.57537	0.59268
28	3092	1810	0.58538	0.57671
29	3007	1622	0.53941	0.56150
30	2844	1590	0.55907	0.54588
31	2793	1449	0.51880	0.52864
32	2652	1405	0.52979	0.50861
33	2422	1213	0.50083	0.48499
34	2533	1214	0.47927	0.45865

(Continued)

Table 4 Continued

Child's Current Age in Months (x)	Women Having Child of Age x	Women Still Breastfeeding, Having Child of Age x	Proportion of Women Still Breastfeeding	Smoothed Proportion of Women Still Breastfeeding
35	2389	1030	0.43114	0.43086
36	2691	981	0.36455	0.40289
37	2487	874	0.35143	0.37600
38	2662	909	0.34147	0.35147
39	2535	852	0.33609	0.33057
40	2384	765	0.32089	0.31413
41	2619	833	0.31806	0.30121
42	2362	718	0.30398	0.29046
43	2419	705	0.29144	0.28051
44	2153	595	0.27636	0.26999
45	2180	528	0.24220	0.25753
46	2094	512	0.24451	0.24220
47	2041	465	0.22783	0.22474
48	2270	421	0.18546	0.20633
49	2192	400	0.18248	0.18815
50	2205	373	0.16916	0.17138
51	2155	316	0.14664	0.15720
52	2331	363	0.15573	0.14678
53	2322	345	0.14858	0.14082
54	2234	334	0.14951	0.13808
55	2161	316	0.14623	0.13682
56	1992	274	0.13755	0.13531
57	1962	226	0.11519	0.13184
58	1945	231	0.11877	0.12467
59	2023	239	0.11814	0.11207

References

- [1] Meeker, W. Q., Escobar, L. A., and Pascual, F. G. (2022). Statistical methods for reliability data. John Wiley & Sons.
- [2] Jardine, A. K., and Tsang, A. H. (2005). Maintenance, replacement, and reliability: theory and applications. CRC press. <https://doi.org/10.1201/9781420044614>.

- [3] Klein, J.P., Moeschberger, M.L. (2003). Basic Quantities and Models. In: Survival Analysis. Statistics for Biology and Health. Springer, New York, NY. https://doi.org/10.1007/0-387-21645-6_2.
- [4] Bain, L. (1991). Statistical Analysis of Reliability and Life-Testing Models: Theory and Methods, Second Edition, (2nd ed.). Routledge. <https://doi.org/10.1201/9780203738733>.
- [5] Elsayed, E.A. (2023). Reliability, Maintainability, Safety, and Sustainability. In: Nof, S.Y. (eds) Springer Handbook of Automation. Springer Handbooks. Springer, Cham. https://doi.org/10.1007/978-3-030-96729-1_31.
- [6] Smith, W. L. (1958). Renewal theory and its ramifications. Journal of the Royal Statistical Society Series B: Statistical Methodology, 20(2), 243–284. <https://doi.org/10.1111/j.2517-6161.1958.tb00294.x>.
- [7] Gavrilov, L. A., and Gavrilova, N. S. (2005). Reliability theory of aging and longevity. *Handbook of the Biology of Aging*, 3–42. <https://doi.org/10.1016/B978-012088387-5/50004-2>.
- [8] Lancaster, H.O. (1990). Measurement of Mortality. In: Expectations of Life. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-1003-0_3.
- [9] Keyfitz, N., and Caswell, H. (2005). Reproductive Value from the Life Table. *Applied Mathematical Demography*, 183–207. https://doi.org/10.1007/0-387-27409-X_8.
- [10] Brijesh, P. S., Gunjan, S., and Singh, K. K. (2016). Application of Life Table Technique to Estimate the Fecundability through First Birth Interval Data. *Journal of Statistics Applications & Probability*, 5, 147–153. DOI: <http://dx.doi.org/10.18576/jsap/050114>.
- [11] Yadava, R. C. & Sharma, S. S. (2007). The distribution of consecutive closed birth intervals in females in uttarpradesh. *Journal of Biosocial Science* 39(2):189–199. doi:10.1017/S0021932006001404.
- [12] Mood, A. M. (1950). Introduction to the theory of statistics. McGraw-Hill. doi:10.1017/S0020269X00004552.
- [13] Williams, D. (1991). Probability with martingales. Cambridge university press. <https://doi.org/10.1017/CBO9780511813658>.
- [14] Yadava, R. C. (2018). Stochastic Modeling of Some Natural Phenomena: A Special Reference to Human Fertility. In *Handbook of statistics* (Vol. 39, pp. 187–274). Elsevier. <https://doi.org/10.1016/bs.host.2018.06.009>.

- [15] Gage, T. B. (1988). Mathematical hazard models of mortality: an alternative to model life tables. *American Journal of Physical Anthropology*, 76(4), 429–441. <https://doi.org/10.1002/ajpa.1330760403>.
- [16] Keyfitz, N. (1966). A Life Table that Agrees with the Data. *Journal of the American Statistical Association*, 61(314), 305–312. <https://doi.org/10.2307/2282820>.
- [17] UN Department of Economic and Social Affairs. <https://www.un.org/development/desa/en/>.
- [18] International Institute for Population Sciences (IIPS) & ICF. (2021). National Family Health Survey (NFHS-5), 2019–21: India. IIPS.
- [19] United Nations Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/NO. 3.
- [20] Asaria, M., Mazumdar, S., Chowdhury, S., Mazumdar, P., Mukhopadhyay, A., and Gupta, I. (2019). Socioeconomic inequality in life expectancy in India. *BMJ global health*, 4(3), e001445. <https://doi.org/10.1136/bmjgh-2019-001445>.
- [21] Singh, M., Shekhar, C., and Shri, N. (2023). Patterns in age at first marriage and its determinants in India: A historical perspective of last 30 years (1992–2021). *SSM – population health*, 22, 101363. <https://doi.org/10.1016/j.ssmph.2023.101363>.
- [22] Srinivasan, K. (1968). A Set of Analytical Models for the Study of Open Birth Intervals. *Demography*, 5(1), 34–44. <https://doi.org/10.2307/2060192>.
- [23] Yadava, R. C., and Bhattacharya, M. (1985). Estimation of parity progression ratios from closed and open birth interval data. *Technical Report Mimeo*.

Biographies



R. C. Yadava obtained his B.Sc., M.Sc. and Ph.D. degrees from Banaras Hindu University. Prof. Yadava started his teaching and research career from his parent institution and retired as Professor of Statistics. He has more than 50 years of teaching experience at Under-graduate and post-graduate levels. Nearly 20 students have obtained their Ph.D. degrees under his supervision. During his academic career, he has published more than 100 research articles in National and International Journals of repute. Some of them are Journal of American Statistical Association (JASA), Canadian Studies in Population, Journal of Biosocial Science, Journal of Data Science, Sankhya, and Demography India. During his tenure at B.H.U., he was head of the Department of Statistics, Dean Faculty of Science, Chief Proctor and many other administrative posts in B.H.U. He has also conducted many research projects as Principal Investigator/Co-investigator sponsored by Indian Council of Medical Research, New Delhi, University Grants Commission, Indian Council of Social Science Research, New Delhi. The major areas of research of Prof. Yadava are Stochastic Modeling of Human Fertility, Effects of Socio-cultural factors on Fertility and conduct of Large-Scale Sample Surveys.



Abhay Kumar Tiwari received master's degree in statistics from University of Allahabad and Ph.D. in statistics from Banaras Hindu University in 2002.

He is currently working as a Professor at the Department of Statistics, Institute of Science, Banaras Hindu University. His research areas include applied statistics, population studies and applied demography. He has more than 20 years of teaching and research experience at Under-graduate and post-graduate levels. He has published more than 80 research articles in National and International Journals of repute.



Vaishali Patel received her B.Sc. (Hons.) and M.Sc. degrees from the Institute of Science, Banaras Hindu University (BHU), Varanasi, India. She is currently an Assistant Professor at Meerut College (affiliated with Chaudhary Charan Singh University), Meerut, India. Her research focuses on Demography, Population Studies, Applied Statistics, Fertility Modelling, and Life Table Analysis. She has published various research papers in Scopus-indexed, peer-reviewed international journals.



Pappu Kumar Singh earned a B.Sc. and M.Sc. degree from University of Allahabad, Uttar Pradesh, India. He is a promising young researcher serving as a Research Scholar at the Department of Statistics, Banaras Hindu University, Varanasi, India. His research focuses on Mathematical Demography, Human Fertility Behaviour and Population Dynamics.

