Journal of Data Science 13(2015), 495-508

Modelling the Distribution of Age at Last Conception of Females

Shruti Verma[1*] and Kaushalendra K. Singh[1]

[1] *Department of Statistics, Faculty of Science, Banaras Hindu University, Varanasi,Uttar Pradesh*

Abstract: Social phenomena that are related to human beings cannot be performed under controlled conditions, making it difficult for policy planners to have an idea about the expected future conditions in the society under varying situations and forming policies. However, modelling can be really helpful to planners in these situations. The present paper attempts to find the distributions of age at last conception of women with the help of stochastic modelling for human fertility taking into consideration different parity progression behaviours among couples. This may be helpful to planners for having at least a rough idea of estimated proportion of women of different age groups who will be completing their childbearing and willing to go for sterilization after marriage under different stopping rules regarding desired family size and sex composition of children. Accordingly, these estimates will help planners to optimize the cost and service provision for sterilization programs for women.
1

Key words: Modelling, conception, stopping rules, conception interval, rate of conception, contraceptive effectiveness.

1. Introduction

In Indian context, generally the reproductive span of a woman starts with marriage and ends as soon as she experiences menopause. But it seems more logical to consider the end of the reproductive span being governed by the couples' decisions regarding maximum number of children they wish to have and stopping rules regarding sex composition of children in the family. For a woman, the start of childbearing is a crucial event of her life. Moreover, the ending point of childbearing is still very important event as it determines the effective reproductive span of a woman. Thus, it is worthwhile for the demographers and the policy makers to analyses the distribution of age at nth conception to the women, where nth conception happens to be the last conception in a woman's life. Distribution of maternal age at last conception or age at last birth (assuming one to one correspondence between a conception and a live birth) can be obtained using the survey data if the completed maternity history of each and every woman is available.

* Corresponding author

But with the information available from the retrospective surveys completed maternity history can be obtained only for those women who have completed their reproduction at the time of survey. Thus if the distribution of age at last birth is obtained from only these older women then it will totally ignore the information on the younger women who are still producing children. Thus it will be more beneficial to incorporate some indirect methods or the modelling techniques to obtain the distribution of age at last conception or birth. Hoem (1970) has obtained some probabilistic fertility models of life table type that can be extended and utilized to study the distribution of age at last conception. Further Krishnamoorthy (1979) and Suchindran and Horne (1984), have developed some methodology to obtain the distribution of age at first and last birth for a woman using the data on age-specific fertility rates (Pandey and Suchindran (1995)). In this direction, Pandey and Suchindran (1995) developed indirect technique to obtain the distribution of maternal age at births of different order using the same vital statistics.

Studying the distribution of age at last conception (or age at last birth) among various subgroups of women, having different preferences regarding stopping behaviors, also gives us an idea about the differentials in childbearing patterns among these subgroups. Age at last conception have important implications for women also, since, this will affect their involvement in other activities besides child rearing. Because if a woman stops childbearing at early ages she will have sufficient useful time to involve herself in some fruitful tasks other than child making as compared to the women who continue bearing children till higher ages. In addition, if a woman stops childbearing at early ages then there will be lesser age difference between her and her child. Further, woman’s age at last conception also has an effect on maternal and child morbidity and mortality. Since it has been observed in various studies that women who give birth in higher ages (> 35 or 40 years) are more prone to pregnancy complications and associated risks to mother and her baby (Andersen et al. (2000); Heffner (2004); Mothers35plus (nd); Paulson (nd)). In addition, the age at which a woman stops childbearing is an important factor that affects the population growth.

Study of distribution of age at last conception in various population subgroups may also help policy makers in forming various policies regarding sterilization programmes also. A very high cost is involved in the implication and operation of family planning and sterilization programmers, thus planners need to know about what proportion of women in various population subgroups and different age groups would be completing their childbearing and will require sterilization, where the sub groups may be divided according to stopping behavior of couples.

Singh et al. (1974) proposed the stochastic model to find the probability distribution of number of births in duration (0,T) after marriage. Utilizing the idea of Singh et al. (1974), this paper attempts to obtain the distributions of age at last conception for subgroups of women or couples having different contraceptive use and stopping behaviors. Thus, in the present paper it is assumed that distribution of woman’s age at last birth will be governed by not only the conception rate but also couple’s aspirations regarding number of children and sex composition of children in the family.

2. Model

The models for time interval from marriage to conceptions of different orders are derived on the basis of the following assumptions:

1. The female remains in the marital union at least up to the period of observation.
2. The probability that the first conception occurs during the time interval (t, $t + \Delta t$) is $\lambda_0 \Delta t + 0(\Delta t)$, $\lambda_0 > 0$, $t > 0$, where λ_0 represents rate of conception corresponding the first conception.
3. Each conception results in a live birth.
4. The probability of another conception after a conception is zero for a constant durartion of length h, where h is the duration of time from a conception to the start of the next menstrual cycle following delivery which is known as the 'rest period'.
5. The conditional probability of a conception during the time interval (t, $t+\Delta t$) is $\lambda_j \Delta t + o(\Delta t)$ if the j^{th} conception occurs prior t h, $t > jh$, $\lambda_j > 0$ and 0 otherwise. Where $\lambda_j's$ represents rate of conceptions corresponding to second, third … conceptions.

Now consider that successive conceptions to a women occur at times Y_0, $Y_0 + Y_1$, $Y_0 + Y_1 + Y_2$, … . Thus Y_0 is the time from marriage to first conception and Y_j ($j = 1, 2, \dots$), is the time interval between successive conceptions or the time interval between j^{th} and $(j + 1)^{th}$ conception to a women. Thus we can call them as closed conception intervals of order $0, 1, 2, \dots$.

Under the assumptions defined above it is apparent that,

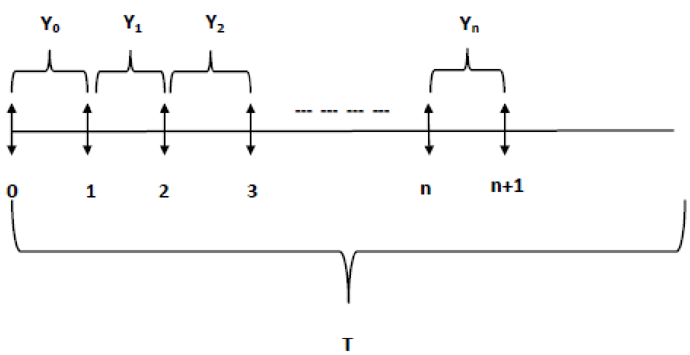

Figure 1: Pictorial representation of conception intervals

$$P(Y_o \le t) = 1 - e^{-\lambda_0 t}, \quad t > 0 \tag{1}$$

$$P(Y_j \le t) = 1 - e^{-\lambda_j (t-h)}, \quad t > h \text{ for } j = 1, 2, \cdots \tag{2}$$

We assume that the random variables Y_0, Y_1, Y_2 , … are mutually independently distributed random variables. Thus the probability density functions for Y_0 , Y_1, … , Y_n will be,

$$f_{Y_0}(y_0) = \lambda_0 . e^{-\lambda_0 y_0}, \quad y_0 > 0, \lambda > 0 \tag{3}$$

$$f_{Y_j}(y_j) = \lambda_j . e^{-\lambda_j (y_j - h)}, \quad y_j > h, \lambda_j > 0 \ \ (j = 1, 2, \cdots, n) \tag{4}$$

i.e. Y_0 follows exponential distribution with parameter λ_0 and $Y_j' s$ follow displaced exponential distribution with parameters h and $\lambda_j' s$.

For finding the distribution of age at last conception, it is required to find the distribution of sum of Y 0s which can be obtained for three major cases, viz.

- **(i)** When all λs differ
- **(ii)** When all λs are similar and
- **(iii)** When some λs are similar while some differ for different order of births.

3. Application of the Model:

Model derived in the previous section will be utilized to obtain the probabilities for timing of conceptions of different orders since marriage for a woman for some hypothetical plans where they are either using or not using contraceptives of various effectiveness. Since the assumption that all conception rates are different for each birth order of a woman doesn't seem to be so realistic so the probability calculations for this case are not done in the application. So the case that 'all λ s differ' is not considered in the plans defined below. The probabilities are estimated for the following plans, where in all the plans, $\lambda_0 = 0.65$, considering no use of contraception after marriage until first birth and $h = 1.25$ years:

Plan A: Same rate of conception for births of each order and no use of contraception; i.e. $\lambda_0 = \lambda_1 = \ldots = \lambda_n$

Plan B: Rate of conception for 2^{nd} and 3^{rd} order births is same that differs from rate of conception for rest high order births under this plan we consider following two hypothetical plans:

Plan B$_1$: Couples use contraceptives of 70% effectiveness after their first birth until their third birth and then of 92% effectiveness afterwards, i.e. $\lambda_1 = \lambda_2 = 0.3\lambda_0 = 0.2$ and $\lambda_3 = \lambda_4 = \ldots = 0.08\ \lambda_0 = 0.05$.

Plan B$_2$: Couples use contraceptives of 85% effectiveness after their first birth until their third birth and then of 92% effectiveness afterwards, i.e. $\lambda_1 = \lambda_2 = 0.15\lambda_0 = 0.1$ and $\lambda_3 = \lambda_4 = \ldots = 0.08\lambda_0 = 0.05$.

Derivations of expressions for probability that sum of closed conception intervals is less than time *t* (for plans A and B):

For Plan A:

The probability distribution of sum of closed conception intervals i.e. the distribution of $Z = Y_0 + Y_1 + Y_2 + \ldots + Y_n$ (when $\lambda_0 = \lambda_1 = \ldots = \lambda_n$ i.e. all λ's are equal) will be

$$f_Z(t) = \frac{\lambda^{n+1}}{n!} e^{-\lambda(t-nh)}(t-nh)^n; \;\; t > nh, \lambda > 0 \tag{5}$$

Which is the p.d.f. of Erlang distribution with shape parameter $n+1$ and rate or inverse scale parameter λ.

Let $u = u - nh$, so

$$fz(u) = \frac{\lambda^{n+1}}{n!} e^{-\lambda u} u^n; \;\; u > 0, \lambda > 0 \tag{6}$$

Now $P\,(Y_0 + Y_1 + Y_2 + \ldots + Yn \leq t)$ will be nothing but the cumulative distribution function (c.d.f.) of $Z = Y_0 + Y_1 + Y_2 + \ldots + Yn$ denoted by $F_z(t)$.

$$F_Z(t) = \int_0^u f_Z(x)dx$$

which ultimately simplifies to (for details see the appendix),

$$F_Z(t) = 1 - e^{-\lambda(t-nh)} \sum_{i=0}^{n} \frac{(\lambda(t-nh))^i}{i!}; \;\; t > nh, \lambda > 0 \tag{7}$$

For Plan B:

Let X and Y be continuous random variables with densities $f_X\,(x)$ and $f_Y\,(y)$ and cumulative distribution function $F_X(x)$ and $F_Y(y)$, then using the property of convolution we have

$$P(X + Y \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_X(x) f_Y(y) dx dy = \int_{-\infty}^{\infty} F_X(z-y) f_Y(y) dy \tag{8}$$

Now if $W_1, W_2, \ldots, W_n$ ~ displaced exponential (λ), then the density function for $W = W_1 + W_2 + \ldots + W_n$ is given by

$$f_W(w) = \frac{\lambda^n}{\Gamma n} e^{-\lambda(w-nh)}(w-nh)^{n-1};\ \ w > nh \tag{9}$$

which is the density of displaced gamma distribution with parameters λ, n and h.

Suppose Y_0 ~ exp (λ_0), Y_1', Y_2', ..., $Y_{n'}'$, (n' *terms*) ~ displaced exponential (λ_1) and Y_1'', Y_2'',..., $Y_{n''}''$, (n'' *terms*) ~ displaced exponential (λ_2), where $n' + n'' = n$
then,

$$P(Y_0 + Y_1' + Y_2' + \cdots + Y_{n'}' \leq t)$$

$$= \int_{n'h}^{t} (1 - e^{-\lambda_0(t-x)}).\frac{\lambda_1^{n'}}{\Gamma n'} e^{-\lambda_1(x-n'h)}(x-n'h)^{n'-1}dx;\ \ x > n'h$$

$$= I(t)\ \ (say) \tag{10}$$

Further using the relation (8), (9) and (10) we have,

$$P(Y_0 + Y_1 + \cdots + Y_n \leq t)$$

$$= P(Y_0 + Y_1' + Y_2' + \cdots + Y_{n'}' + Y_1'' + Y_2'' + \cdots + Y_{n''}'' \leq t)$$

$$= \int_{n''h}^{t} I(t-x')\frac{\lambda_2^{n''}}{\Gamma n''} e^{-\lambda_2(x'-n''h)}(x'-n''h)^{n''-1}dx';\ \ x' > n''h \tag{11}$$

Expressions (7) and (11) are nothing but the cumulative distribution functions of time from marriage to the $(n+1)^{th}$ conception for different conditions on λ·s i.e. When all are similar and when some differ but some are similar for different order of births.

The probability distribution of this random variable i.e. the time from marriage to $(n+1)^{th}$ conception can be obtained with the help of well-known relation viz.,

$$P(a < X \leq b) = F(b) - F(a)$$

Now to find the probability distribution of women's age at $(n + 1)^{th}$ conception for a cohort of women who marry at a certain age d, it is required to shift the origin of time from marriage to $(n + 1)^{th}$ conception at d.

Next to find the distribution of age at last conception of the women for this cohort, the distribution of age at $(n+1)^{th}$ conception have to be weighted with the probability of exactly *m*(=

n+1) conceptions to a woman in duration *T*, where *T* is assumed to be large such that couples complete their reproduction in duration T after marriage. It is well understood that the couples will not go on producing children indefinitely rather they will follow certain stopping rules. Thus it is reasonable to assume that the probabilities of exactly m conceptions in duration T for a woman will be affected by couple's decision regarding maximum number of children and the sex composition of the children they wish to have in the family.

Thus for finding these probabilities, we consider six hypothetical sub plans regarding couple's stopping behavior related to desired family size and sex composition of children viz.

- **Sub plan S_1 :** Produce children until one male child is born but stop childbearing at parity 2 regardless of whether a male child is born or not.
- **Sub plan S_2:** Produce children until one male child is born but stop childbearing at parity3 regardless of whether a male child is born or not.
- **Sub plan S_3:** Produce children until one male child is born but stop child bearing at parity 4 regardless of whether a male child is born or not
- **Sub plan S_4:** Produce children until one male and one female child is born but stop childbearing at parity 3 regardless of whether one male and one female child is born or not.
- **Sub plan S_5:** Produce children until one male and one female child is born but stop childbearing at parity 4 regardless of whether one male and one female child is born or not.
- **Sub plan S_6:** Produce children until at least two male and one female child is born but stop childbearing at parity 4 regardless of whether two male and one female child is born or not.

The probabilities of exactly *m* conceptions to a woman are calculated for different sub plans S_1, S_2, S_3, S_4, S_5, S_6 as follows:

Let p and $q = (1\text{-}p)$ be the probabilities of producing a male and a female child respectively. Now we illustrate how the probability of exactly *m* births (or exactly *m* conceptions) can be obtained. For an example, the probability of exactly 2 births under sub plan S_4, will be $qp + pq$. Since under sub plan S_4, couples will stop at two children iff their first child is a male and second is a female or vice versa, but if they don't achieve it i.e. if children born are both females or both males, they will move to the next conception.

Table 1: Combinations of children (by sex and number) that couples can have under various sub plans

m	Sub Plan S_1	Sub Plan S_2	Sub Plan S_3	Sub Plan S_4	Sub Plan S_5	Sub Plan S_6
1	M	M	M			
2	FM,FF	FM	FM	FM,MF	FM,MF	
3		FFM,FFF	FFM	FFM,MMF, FFF,MMM	FFM,MMF	MMF,MFM, FMM
4			FFFM,FFFF		FFFM,MMMF, FFFF,MMMM,	MFFM,FFMM,FMFM, MMMF,FFFM,FFMF, FMFF,MFFF,FFFF, MMMM,

Note: m= maximum no. of children, F= female, M= male

Table 1 shows various combinations of children that couples can have under various sub plans considering $p = 0.515$. Also the calculated probabilities of exactly m births for various sub plans have been represented in Table 2.

Now these obtained probabilities (in Table 2) are used as weights in obtaining the distribution of age at last conception from the distribution of age at $_m$ conceptions. Figures 2, 3 and 4 show the distribution of age at last conception for the cohort of women who married at age 20 years. Distributions are obtained under plans A, B_1 and B_2 for each sub plan S_1, S_2, S_3, S_4, S_5 and S_6. Age at last conception is divided into nine categories, viz. *< 22*, *22-24*, *24-26*, *26-28*, *28-30*, *30-32*, *32-34*, *34-36*, *>36*, where age is measured in years.

Table 2: Probabilities of exactly m births under various sub plans:

m	Sub Plan S_1	Sub Plan S_2	Sub Plan S_3	Sub Plan S_4	Sub Plan S_5	Sub Plan S_6
1	0.515	0.515	0.515	0.000	0.000	0.000
2	0.485	0.250	0.250	0.500	0.500	0.000
3	0.000	0.235	0.121	0.500	0.250	0.386
4	0.000	0.000	0.114	0.000	0.251	0.614
Sum	1.000	1.000	1.000	1.000	1.000	1.000

From Figure 2, representing the distribution for the group of females who don't use any form of contraception before their first conception (or first birth), it can be observed that under sub plan S_1, more than 70% women conceive their last child within 4 years of their marriage, while under sub plans S_2 and S_3 around 62 to 63% women conceive their last child within 4 years of their marriage.

Under sub plan S_4,around 31% women conceive their last child within 4 years of their marriage and around 61% within 6 years of their marriage. Under sub plans S_5 and S_6, less than 50% women conceive their last child within 6 years of their marriage. In addition under sub plan S_5, only around 3% women remain to conceive their last child after 10 years of their marriage while under sub plan S_6, around 70% women complete childbearing within 10 years of their marriage and the rest 30% women conceive their last child even after 10 years of their marriage.

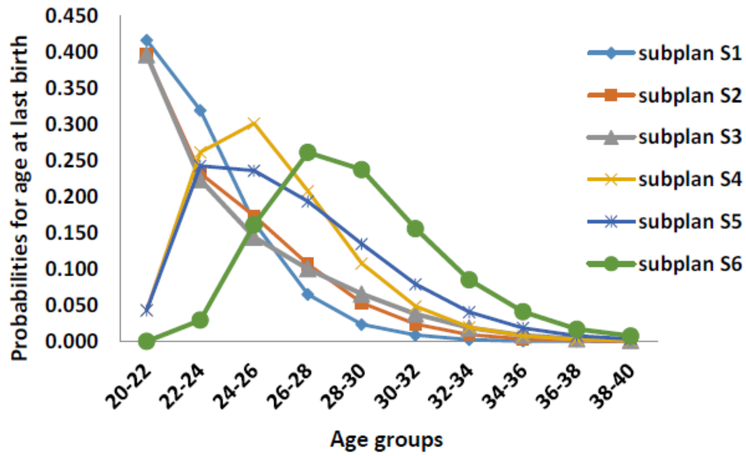

Figure 2: Distribution of age at last conception (regardless of parity) for cohort of women of marital age 20 years (Under Plan A)

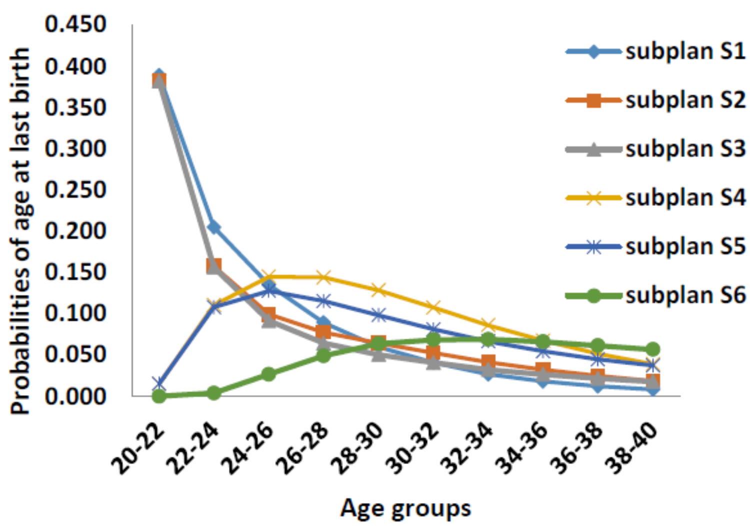

Figure 3: Distribution of age at last conception (regardless of parity) for cohort of women of marital age 20 years (Under Plan B_1)

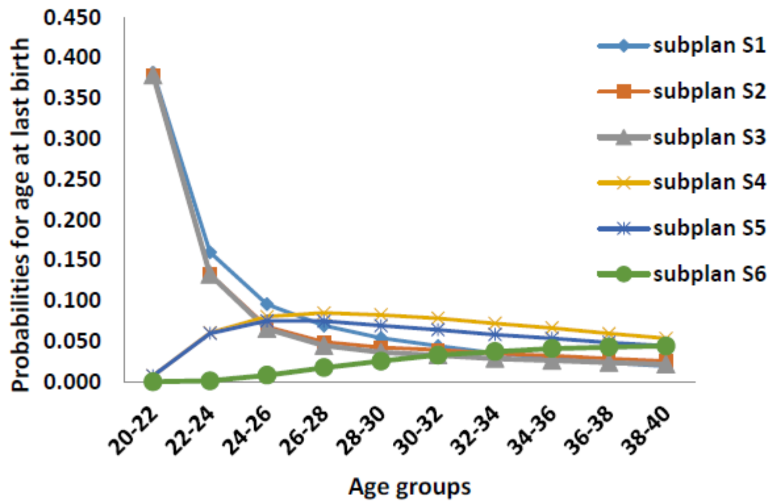

Figure 4: Distribution of age at last conception (regardless of parity) for cohort of women of marital age 20 years (Under Plan B_2)

Further, from Figure 3 it is observed that under plan B_1, for each sub plan $S_1 - S_6$ proportion of women who stop conceiving within few years of their marriage is lesser as compared to the respective proportions obtained under plan A. Under this plan B_1, since women are using contraceptives before their last conception, age at last conception is shifting towards the higher ages. Here, it is interesting to note that around 33 to 35% women still have not completed their childbearing at the age of 36 i.e. 16 years after marriage under sub plans S_5. Also under sub plan S_6, around 65% women still remain to conceive their last child even after reaching the age of 36 years.

Figure 4 represents the distribution of age at last conception for plan B_2, where women use contraceptives of higher effectiveness as compared to women following plan B1. Here also it can be observed that women's age at last conception is shifting towards the higher ages. The proportion of women remaining to conceive their last child even after 16 years of their marriage becomes still higher as compared to the women following plan B_1. This shows that as contraceptive effectiveness is increasing contraceptives have a positive effect in increasing the conception interval i.e. the spacing between two consecutive conceptions.

4. Conclusions:

The previous section clearly demonstrates that how the modelling technique can be helpful in obtaining the distribution of age at last conception of women for various stopping behaviors. In general it can be said that this modelling technique can be used in various fields where one wishes to obtain the time up to the n_{th} event or the time up to the last event. Thus modelling techniques can be very helpful in estimating and predicting the expected form of distribution under various assumptions which can be really helpful to policy makers in forming policies.

Policy makers need to know about the expected scenario that will result under various assumptions. By taking appropriate assumptions that seem to be most favorable for the present or the future, planners can get an idea about the expected estimates on the required variable and hence can make the plans and policies accordingly. Thus, this paper is an attempt to throw light on one of the useful implications of modelling techniques which can be helpful for the researchers in understanding the usefulness of modelling. Furthermore, the technique discussed can be extended to be used for obtaining the distributions under various circumstances as the research purpose may suggest.

Acknowledgment

Authors are very grateful to Prof. R. C. Yadava for his valuable suggestions and comments while working on this paper. Author (S.V.) is also thankful to University Grants Commission for financial assistance.

References

[1] Andersen, A. M. N., Wohlfahrt, J., Christens, P., Olsen, J., and Melbye, M. (2000). Maternal age and fetal loss: population based register linkage study. *British Medical Journal*, **320**:**1708–1712**.

[2] Heffner, L. J. (2004). Advanced maternal agehow old is too old? *The New England Journal of Medicine*, **351**(**19**):1927–1929.

[3] Hoem, J. M. (1970). Probabilistic fertility models of the life table type. Theoretical Population Biology, (**1**):12–38.

[4] Krishnamoorthy, S. (1979). Family formation and the life cycle. Demography, (16):**121–129**.

[5] Mothers35plus (n.d.). Getting pregnant over 35. Retrieved from http://www.mothers35plus.co.uk/pregnancyover35.htm on Dec 27, 2013.

[6] Pandey, A. and Suchindran, C. M. (1995). Some analytical models to study the distribution of maternal age at different births from the data on agespecific fertility rates. *Sankhya*, *Series B*, pages 142–150.

[7] Paulson, R. (n.d.). What are the risks of having a baby if i am 35 or older? published online by BabyCenter L.L.C.

[8] R Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

[9] Singh, S. N., Bhattacharya, B. N., and Yadava, R. C. (1974). A parity dependent model for number of births and its applications. *Sankhya, Series B*, **36**(**1**):93–102.

[10] Suchindran, C. M. and Horne, A. D. (1984). Some statistical appoaches to the modelling of selected fertility events. Proceedings of the American Statistical Association (Social Statistics Section), pages 629–634.

Received December 13th, 2014; accepted February 10th, 2015.

Shruti Verma
Department of Statistics, Faculty of Science
Banaras Hindu University
The Banaras Hindu Unversity, Varansi, Uttar Pradesh.
shrutiverma.bhu@gmail.com

Appendix

$$= \int_0^u \frac{\lambda^{n+1}}{n!} e^{-\lambda x} x^n dx$$

$$F_Z(t) = \int_0^u f_Z(x) dx$$

Let $\lambda x = k$, then

$$F_Z(t) = \frac{\lambda^{n+1}}{n!} \int_0^{\lambda u} e^{-k} \frac{k^n}{\lambda^{n+1}} dk$$

$$= \frac{1}{n!} \int_0^{\lambda u} e^{-k} k^n dk$$

$$= \frac{\gamma(n+1, \lambda u)}{n!}$$

where γ () is lower icomplete gamma function.

Let $n+1 = \alpha$

$$F_Z(t) = \frac{1}{(\alpha-1)!} \int_0^{\lambda u} e^{-k} k^{\alpha-1} dk$$

Solving the integral using integration by parts taking $k^{\alpha-1}$ as first and e^{-k} as second function,

$$F_Z(t) = -e^{-\lambda u} \sum_{i=\alpha-2}^{\alpha-1} \frac{(\lambda u)^i}{i!} + \frac{1}{(\alpha-3)!} \int_0^{\lambda u} e^{-k} k^{\alpha-3} dk$$

Solving the integral recursively using integration by parts, we get

$$F_Z(t) = -e^{-\lambda u} \sum_{i=0}^{\alpha-1} \frac{(\lambda u)^i}{i!} + 1$$

$$= 1 - e^{-\lambda u} \sum_{i=0}^{\alpha-1} \frac{(\lambda u)^i}{i!}$$

$$= 1 - e^{-\lambda u} \sum_{i=0}^{n} \frac{(\lambda u)^i}{i!}$$

$$= 1 - e^{-\lambda(t-nh)} \sum_{i=0}^{n} \frac{(\lambda(t-nh))^i}{i!} \ ; t > nh, \lambda > 0$$

which is the cumulative distribution function of Z = Y0 + Y1 + Y2 + … + Yn (where all λ_j's are equal)