**Demography India**
A Journal of Indian Association of Study of Population
Journal Homepage: https://demographyindia.iasp.ac.in/

# A Revisit to "Chance Mechanism of the Variation in the Number of Births Per Couple"

R. C. Yadava[1] and Piyush Kant Rai[2]*

***Keywords***

Stochastic modelling, Exponential Distribution, Poisson distribution, Cohabitation

**Abstract**

Study on human fertility through mathematical modelling has been an innovative process to understand the phenomenon in a better way. In this direction, S. N. Singh (1968) proposed a Stochastic Model for describing the variation in the number of conceptions to a female in a given interval of time (0, T) under some simplifying assumptions. However, the proof of derived distribution was not given in the paper, although some references were cited to derive the given distribution. In the present article, we have proposed two simple proofs for obtaining the results given in the above-mentioned paper. The given proofs are very interesting and easily understandable. It is also hope that the proposed proofs may felicitate the researchers, especially young researchers to understand the whole mechanism of the derived distribution.

*Corresponding Author
[1]Dept. of Statistics, Institute of Science, Banaras Hindu University, (B.H.U.), Varanasi- 221005, India. E-Mail: rcyadava66@yahoo.co.in
[2]Dept. of Statistics, Institute of Science, Banaras Hindu University, (B.H.U.), Varanasi- 221005, India. E-Mail: raipiyush5@gmail.com, piyush.rai01@bhu.ac.in

**Introduction**

Human reproduction is a complex process. Although, essentially it is a biological process but many times highly affected by large number of socio-cultural factors prevailing in the society. It is one of the most crucial factors in determining the demographic features of a country. Precise estimates of human fertility are very much needed for formulation, implementation and execution of various developmental plans of a country.

In order to have better understanding of the phenomenon, demographers have used various methodologies and techniques to analyse various aspects of human fertility. The development of mathematical models is one of them. A 'model' is an abstraction of a real phenomenon. If this abstraction is done in terms of mathematical relationship(s), it is called a mathematical model for the phenomenon under study. These models have novel interpretative, predictive and communicative values which enhance their utility considerably.

Depending upon the nature of the phenomena under consideration, a mathematical model falls broadly in either of the two categories: (i) Deterministic and (ii) Stochastic (i.e. probabilistic). A deterministic model has an element of certainty where the end result is certain while stochastic models are more appropriate for the situations where phenomenon under consideration is of random nature.

Human fertility is mainly measured through 'births' and 'birth' is essentially a phenomenon of random nature. Thus, the stochastic models have been considered to be more appropriate for the study of human fertility.

A large number of stochastic models have been proposed for various aspects of human fertility. These mainly come into the following two categories: (i) How frequently females give births? (ii) How these births are spaced? These are mainly referred as (i) Models for number of births/conceptions in a given interval of time and (ii) Models for birth intervals under various sampling frames.

Some of the initial models on birth intervals are given by Potter and Parker (1964), Sheps (1964), Singh (1964), Srinivasan (1966,1967,1968) and many others. Similarly, some initial model for the number of birth/conceptions are given by Dandekar (1955), Brass (1958), Singh (1961,1963), Sheps and Menken (1973) etc.

In this direction Singh (1968) also published a paper entitled "A chance mechanism of variation in the number of births per couple" in the 'Journal of American Statistical Association'. In this paper Singh has obtained the probability distribution of a random variable representing the number of conceptions to female in a given interval of time (0, T) along with its assumptions. Although, the probability distribution of the random variable is given in the paper but no proof for the same is provided in the same. This may perhaps mainly be due to paucity of space for publication of the results in detail at that time. Although, some references are given to prove the results which are quite old and cumbersome also (Feller, 1948; Blyth, 1949; Neyman, 1949).

This paper is considered to be a basic paper on the topic. However, perhaps due to non-availability of the methodology to prove the results, the researchers especially young researchers face difficulties to understand the mathematical derivation of the given results and many times either leave it or postpone for future time for understanding the given results.

The objective of the present paper is to give a methodology to prove the given results in Singh (1968) which may be helpful in deriving the results and giving motivation for further research in the field.

The present article is highly connected with the paper of Singh (1968), hence it is more desirable to give a discerption of the paper here also.

So, we present the derived distribution along with its assumptions in the format published in the Journal for ready reference and continuity.

Let X denote the number of conceptions to a female in a given interval of time (0, T). Though explicitly not mentioned, the derived probability distribution for X is essentially a Stochastic model for human fertility describing the variation in the number of conceptions to a female in a given interval of time (0, T). It is also assumed that every conception result in live birth.

The probability distribution of X was derived under the following assumptions:

***Assumptions 1(a).*** The number of cohabitations of a couple during a given time interval (0, T) of length T is a random variable and follows a Poisson distribution

$$P\,[Z = k] = e^{-\lambda_1 T}\,\frac{(\lambda_1 T)^k}{k!}\; k=0,1,2\ldots;\; \lambda_1 > 0;\; T > 0 \quad (1)$$

where, Z denotes the number of co-habitations during the time (0, T) and $\lambda_1$ is a constant.

***Assumption1(b).*** The cohabitations are mutually independent and $p_1$, the probability that a co-habitation results in a conception, is constant. It can easily be seen that the number of such conceptions during the time interval (0, T) follows a Poisson distribution with parameter $\lambda T = \lambda_1 p_1 T$ under the assumptions 1(a) and 1(b). These assumptions are strong, but in the absence of any empirical evidence on the distribution of the number of co-habitations, we have assumed that the Poisson distribution applies because of its simplicity and range of variability.

***Assumption2.*** After each conception, there is no possibility of another conception for a constant time h, h is the duration of time from a conception to the start of the next menstrual cycle following delivery. For a given female the variation of h is small, so the assumption of h being a constant is reasonable as a first approximation.

***Assumption3.*** Either the female is exposed to the risk of a conception throughout the interval (0, T) or she is not exposed to this risk at any time during the interval (0, T). Let $\alpha$ and $(1-\alpha)$ be the respective probabilities. According to Singh (1963) the total number of conceptions during the time interval (0, T) cannot be more than n, where n = [T/h] +1 and [T/h] stands for the greatest integer not exceeding T/h.

Under the assumptions (1) and (2), from the equations (6) of Neyman (1949), the distribution of the number of conceptions during the time interval (0, T) is

$$P[X = 0] = e^{-\lambda T} \quad (2)$$

$$P[X = i] = \sum_{m=0}^{i} e^{-\lambda(T-ih)} \frac{[\lambda(T-ih)]^m}{m!} - \sum_{m=0}^{i-1} e^{-\lambda(T-ih+h)} \frac{[\lambda(T-ih+h)]^m}{m!} \quad (3)$$

for i=0,1, 2,…,n-1

$$P[X = n] = 1 - P(X \le n-1) \quad (4)$$

These equations are also connected with the Geiger counter problem with a finite resolving time equal to h. The problem of Geiger counter has been discussed by Feller (1948) and many others. Dandekar (1955) has derived this distribution as a modification of the Poisson distribution. Under the assumptions (1), (2) and (3) we get the following trivial extension of the above distribution.

The probability expressions for the distribution of X under the assumptions (1), (2) and (3) are given as

$$P[X=0]=1-\alpha+\alpha e^{-\lambda T} \quad (5)$$

$$P[X=i]=\alpha\left[\sum_{m=0}^{i} e^{-\lambda(T-ih)} \frac{[\lambda(T-ih)]^m}{m!} - \sum_{m=0}^{i-1} e^{-\lambda(T-ih+h)} \frac{[\lambda(T-ih+h)]^m}{m!}\right] \quad (6)$$

for i=0, 1, 2,…, n-1

Thus, $P[X = n]$ can be obtained as

$$P[X = n] = 1 - P(X \le n-1) \quad (7)$$

where n is the maximum possible number of conceptions in (0,T)

and n $=\left[\frac{T+h}{h}\right]$, is the largest integer not exceeding $\frac{T+h}{h}$.

The model was applied to a real set of data representing the number of children born in 5 years assuming that every conception results in a live birth and the fit was found to be satisfactory according to chi-square criterion.

Here, it is to be mentioned that although the probability expressions were clearly specified but- no mathematical proof was provided for the same.

Perhaps, this might have been due to paucity of space for publication in the Journal at that time, although certain references are given which are now too old (more than 50 years) and cumbersome as well. Now-a-days, many of the researchers, especially young researchers face difficulties in deriving the probability expressions given in the paper Singh (1968). Although, there can be different ways of proving the results given in Singh (1968) but it is not known that which method was applied by Singh (1968) to obtain the given results. Alternatively, in this article we give two simple proofs to derive the same.

**Methodology**

The proposed methodology for proving the results given in Singh (1968) is mainly based on the utilisation of interrelationship between Poisson process and exponential distribution. In this context we proposed two methods to prove the same results on two concepts (i)based on distribution function and (ii) direct derivation of probability distribution of X(t).

**The Proposed Proofs**

The first proof is as follows:

As per assumptions (1a) and (1b), the number of conceptions in (0, T), follows a Poisson distribution with parameter λT.

Under the above assumptions, the intervals between consecutive conceptions will follow exponential distribution which will be independent and identically distributed (i.i.d.) with p.d.f.

$$f(t)=\lambda e^{-\lambda t}\ ,\lambda > 0, t > 0 \qquad (8)$$

Now let us define, $T_0$ as time between marriage to first conception, $T_1$ as time between first and second conceptions, $T_2$ as time between second and third conceptions and so on.

Incorporating assumption (2), we see that $T_0$, $T_1$, $T_2$,.... are independently distributed where $T_1$, $T_2$,... are i.i.d. random variables each having a displaced exponential distribution with p.d.f.

$$f(t)=\lambda e^{-\lambda(t-h)}\ , t > h, \lambda > 0. \qquad (9)$$

and $T_0$ follows exponential distribution with p.d.f. given as,

$$f(t)=\lambda e^{-\lambda t}\quad , t > 0, \lambda > 0 \qquad (10)$$

Now, let us define

$X_0=T_0$

$X_i=T_i-h$ , i=1,2,…

then, clearly $X_0$, $X_1$, $X_2$,… are i.i.d. random variables each with p.d.f.

$$f(x)=\lambda e^{-\lambda x}\ ,\ x > 0,\ \lambda > 0 \qquad (11)$$

Clearly, $Z_i=X_0+X_1+X_2+\ldots+X_i$, which is the sum of (i+1), i.i.d. exponential distributions will follow a Gamma distribution with p.d.f.

$$f(z_i)=\lambda^{i+1}e^{-\lambda z_i}\frac{z_i^{\,i}}{i!}\quad i=0,1,2,\ldots\ \lambda > 0,\ z_i > 0 \qquad (12)$$

Clearly, the event $T_0+T_1+\ldots+T_i<T$ is equivalent to the event that at least (i+1) conceptions occur in (0,T).

Similarly, the event $T_0+T_1+\ldots+T_i<T$ is equivalent to the event that $X_0+X_1+X_2+\ldots+X_i< T-ih$

Thus, $P[T_0+T_1+\ldots+T_i< T] = P[X_0+X_1+X_2+\ldots+X_i< T-ih]=P[Z_i<T-ih]= \int_0^{T-ih}\lambda^{i+1}e^{-\lambda z_i}\frac{z_i^{\,i}}{i!}\,dz_i$ (13)

Integration by parts gives,

$$P[Z_i<T-ih]=1-e^{-\lambda(T-ih)}\sum_{m=0}^{i}\frac{\lambda^m(T-ih)^m}{m!} \qquad (14)$$

Similarly,

$$P[z_{i-1}<T-(i-1)h]=1-e^{-\lambda[T-(i-1)h]}\sum_{m=0}^{i-1}\frac{\lambda^m[T-(i-1)h]^m}{m!} \qquad (15)$$

The event $z_{i-1}< T$ is nothing but the probability that at least i conceptions occur in (0, T). Thus, the probability that exactly i conceptions occur in (0, T) is given as

$$P[X=i]=[1e^{-\lambda[T-(i-1)h]}\sum_{m=0}^{i-1}\frac{\lambda^m[T-(i-1)h]^m}{m!}]-[1-e^{-\lambda(T-ih)}\sum_{m=0}^{i}\frac{\lambda^m(T-ih)^m}{m!}]=e^{-\lambda(T-ih)}\sum_{m=0}^{i}\frac{\lambda^m(T-ih)^m}{m!}\ e^{-\lambda[T-(i-1)h]}\sum_{m=0}^{i-1}\frac{\lambda^m[T-(i-1)h]^m}{m!}]\quad i=0,1,2\ldots n-1 \qquad (16)$$

Here also,

$$P[X=n]=1-P[X\leq n-1]$$

In fact

$$P[X=n]=1-[1-e^{-\lambda[T-(i-1)h]}\sum_{m=0}^{i-1}\frac{\lambda^m[T-(i-1)h]^m}{m!}] \qquad (17)$$

such that P[X=0]+P[X=1]+…+P[X=n]=1

Incorporating the assumption (3), we see that

$$P[X=0]=1-\alpha+\alpha e^{-\lambda T} \qquad (18)$$

$$P[X=i]=\alpha[\sum_{m=0}^{i} e^{-\lambda(T-ih)} \frac{[\lambda(T-ih)]^m}{m!} - \sum_{m=0}^{i-1} e^{-\lambda(T-ih+h)} \frac{[\lambda(T-ih+h)]^m}{m!}] \quad i=1,2,\ldots,n-1 \tag{19}$$

and, $P[X=n]=1-P[X\leq n-1]$.

This completes the proof for the given results.

We now give an alternative proof for the above probability distribution.

**An Alternative Proof**

We notice that if h=0, then the number of conceptions to a fecund female will follow Poisson distribution with parameter $\lambda T$. The alternation in the probability distribution is mainly due to incorporation of period h>0 (which we may call the rest period).

In fact, suppose a female has i conceptions in (0,T), then the rest periods h (for each conception) may

1: entirely lie in (0,T) or

2: rest periods for (i-1) conceptions entirely lie in (0,T) but some part of h of the ith conception may lie in (0,T) while some part may lie beyond T.

Considering the above two situations (events), Singh (1963) proposed a model for number of conceptions in (0,T) based on similar assumptions as of Singh (1968) but treating time to be discrete.(unit of time taken as one month)

Singh (1963) utilized the above concept and computed the probability expressions for the above mentioned two situations (events). These two events are mutually exclusive, hence the probability of i conceptions can be computed as the sum of two probabilities.

We now attempt to obtain the expressions for the model of Singh (1968) under assumptions (1), (2), and (3) of Singh (1968) treating time to be continuous.

Let us consider the first situation (event) that exactly i conceptions occur in (0, T) when the rest periods (h) associated with each of the i conceptions entirely lie in (0, T).

The probability of this event is equal to

$$e^{-\lambda[T-ih]} \frac{[\lambda(T-ih)]^i}{i!} \tag{20}$$

i.e. probability of i events in (0, T-ih) in Poisson process with parameter $\lambda(T-ih)$

Now, we consider the second situation (event) i.e. when the rest period (h) related to ith conception partially lies in (0,T) while some part of it is beyond T [of course rest periods associated with (i-1) conceptions will definitely lie in (0,T)].

Obviously, this can happen only if the $i^{th}$ conception occurs in the interval (T-h, T). For this we have to obtain

$$P[T-h\leq T_0+T_1,+\ldots+T_{i-1}\leq T].$$

This is equivalent to finding $P[T-ih \leq z_{i-1}\leq T-(i-1)h]$.

Obviously $X_0$, $X_1\ldots$ $X_{i-1}$ are independent and identically distributed exponential variates with mean $1/\lambda$

Thus, $z_{i-1}=X_0+X_{1+}\ldots,+X_{i-1}$ will be a gamma variate.

For simplicity of notation, let us put

$Y=Z_{i-1}$, then p.d.f. of Y will be

$$f(Z_{i-1})=\lambda^i e^{-\lambda Z_{i-1}} \frac{(Z_{i-1})^{i-1}}{i-1!},\ Z_{i-1}>0,\ \lambda>0 \tag{21}$$

Thus, $P[T-ih\leq Z_{i-1} \leq T-(i-1)h]=$

$$\int_{T-ih}^{T-(i-1)h} \lambda^i e^{-\lambda Z_{i-1}} \frac{(Z_{i-1})^{i-1}}{(i-1)!} d Z_{i-1}$$

$$= e^{-\lambda(T-ih)} \sum_{m=0}^{i-1} \frac{\lambda(T-ih)^m}{m!} - e^{-\lambda[T-(i-1)h]} \sum_{m=0}^{i-1} \frac{\lambda[T-(i-1)h]^m}{m!} \tag{22}$$

Adding the two probabilities of two mutually exclusive events, we get

$$P[X=i]= e^{-\lambda(T-ih)} \frac{\lambda(T-ih)^i}{i!} + e^{-\lambda(T-ih)} \sum_{m=0}^{i-1} \frac{\lambda(T-ih)^m}{m!} - e^{-\lambda[T-(i-1)h]} \sum_{m=0}^{i-1} \frac{\lambda[T-(i-1)h]^m}{m!}$$
$$= e^{-\lambda(T-ih)} \sum_{m=0}^{i} \frac{\lambda(T-ih)^m}{m!} - e^{-\lambda[T+(i-1)h]} \sum_{m=0}^{i-1} \frac{\lambda[T-(i-1)h]^m}{m!} \tag{23}$$

Incorporating the assumption (3), we get

$$P[X=0]=1-\alpha+\alpha e^{-\lambda T} \tag{24}$$

$$P[X=i]=\alpha[ e^{-\lambda(T-ih)} \sum_{m=0}^{i} \frac{\lambda(T-ih)^m}{m!} - e^{-\lambda[T+(i-1)h]} \sum_{m=0}^{i-1} \frac{\lambda[T-(i-1)h]^m}{m!}] \tag{25}$$

$$P[X=n]=1-P[X\leq n-1] \tag{26}$$

Thus, the result is proved in an alternative way also.

**Discussion and Conclusion**

It is already mentioned that there can be many methods to prove the same results in different ways. However, we feel that the methods proposed by us are quite simple and easily understandable. We hope that the proofs given in the present article will facilitate the young researchers to understand the basic concepts of Stochastic Modeling as well as the related Mathematical derivations.

**Conflict s of Interests.** "The authors declare no conflict of interests" for this manuscript.

**Acknowledgement:** We are very thankful to chief editor of the Journal and the reviewers of the paper for proving fruitful suggestions to improve the quality of the paper.

**References**

Blyth, C. R. (1949). Contribution to the Statistical theory of the Geiger Muller counter. Technical Report No. 8 Statistical Laboratory, University of California, (unpublished).

Brass, W. (1958). The distribution of births in human populations in rural Taiwan. Population Studies, 12(1), 51-72.

Dandekar, V. M. (1955). Certain modified forms of binomial and Poisson distributions. Sankhya, Vol.15, pp. 237-51.

Feller, W. (1948). On probability problems in the theory of counters. Studies and Essays (Presented to R. Courant on his 60th birthday), pp.105-115.

Neyman, J. (1949). On the problems of estimating the number of schools of fish, University of California Publications in Statistics; Vol. 1, pp. 21-36.

Neyman, J. (1949). Contribution to the theory of $\chi^2$ test. In Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, pp. 239-73.

Potter, R. G., & Parker, M. P. (1964). Predicting the time required to conceive. Population studies, 18(1), 99-116.

Sheps, M. C. (1964). On the time required for conception. Population Studies, 18(1), 85-97.

Sheps, M. C., Menken, J., & Radick, A. P. (1973). Mathematical models of conception and birth. University of Chicago Press.

Singh, S. N. (1961). A hypothetical chance mechanism of variation in number of births per couple. University of California, Berkeley.

Singh, S. N. (1968). A Chance Mechanism of the variation in the number of births per couple. The Journal of the American Statistical Association, Vol. 63, pp. 209-213.

Singh, S. N. (1963). Probability models for the variation in the number of births per couple, The Journal of the American Statistical Association, Vol. 58, pp. 721-727.

Srinivasan, K. (1966). An application of a probability model to the study of interlive birth intervals. Sankhyā: The Indian Journal of Statistics, Series B, 175-182.

Srinivasan, K. (1967). A probability model applicable to the study of inter-live birth intervals and random segments of the same. Population Studies, 21(1), 63-70.

Srinivasan, K. (1968). A set of analytical models for the study of open birth intervals. Demography, 5, 34-44