On a Statistical Model Useful for Demographics: Estimating the Mean Number of Children Ever Born Through the Distribution of Male Births with an Application to Data from India

Shubhagata Roy[1], Prayas Sharma[2,3,*], K. K. Singh[4]
and Richa Srivastava[5]

[1]*ICFAI Business School, IFHE, Hyderabad, India*
[2]*Department of Decision sciences, Indian Institute of Management Sirmaur, Sirmaur, India*
[3]*Department of Statistics, Babasaheb Bhimrao Ambedkar University, Lucknow, India*
[4]*Department of Statistics, BHU, Varanasi, India*
[5]*Decision Sciences, Jaipuria Institute of Management, Lucknow, Uttar Pradesh, India*
E-mail: prayassharma02@gmail.com
[*]*Corresponding Author*

Received 01 September 2022; Accepted 13 March 2023;
Publication 15 June 2023

Abstract

The connection between male births and fertility can be easily linked with demographic transition and in defining the population distribution. In this context, it is necessary to understand the birth patterns in Indian societies which are governed by some or the other probability distributions. Although child birth is a biological process but it is very much influenced by a number of social, economic, cultural and psychological factors. Numerous demographers have proposed mathematical models to predict the number of male and female births during a given time period taking into consideration the various factors. Traditionally, estimating current levels and future trends of mean number of births is done using various life tables, cohort-component

Journal of Reliability and Statistical Studies, Vol. 16, Issue 1 (2023), 57–80.
doi: 10.13052/jrss0974-8024.1613

method, time-series analysis, micro-simulations, structural modeling, expert analysis, historical error analysis and also using an appropriate probability model and testing the model on real data. In the present study we developed a model for estimating the mean number of children ever born through the join probability distribution with its application for male births among the females of Uttar Pradesh and Bihar. The reasons of selecting these two states were their huge population and high total fertility rates. The model fits to the data of these two states, therefore it would be a good fit for the other states too, which shows the efficiency and applicability of the model. The applicability of this model has been illustrated on real data obtained from the National Family Health Survey-3 (2005–06). The various estimates of the parameters have been obtained by using the method of moments and suitability of the proposed model has been tested using the 'goodness of fit' criteria.

Keywords: Fecundity, fecundability, family planning, fertility, fertility transition, NFHS-, reproductive health.

1 Introduction

Birth of sons has always been more precious than that of daughters in Indian societies and it is deeply rooted in the patrilineal systems which have a strong belief that sons will take care of the aging parents. A major demographic outcome of son preference is that the family size depends on the birth order of sons, i.e., family size keeps on growing till the time a satisfactory number of sons are not born (Seidl, 1995). In fact, women with more sons than daughters were, in general, less likely than those with more daughters than sons to continue childbearing (Chaudhuri, 2012). The strong desire for sons, mostly, results in imbalances in the sex-ratio by family size. However, the desire for sons may demonstrate a skewed family size distribution, as the families where first child is son, would be smaller in size and the families where first child is daughter, would be larger in size (Basu & Jong, 2010). Another outcome of son preference is the occurrence of gender inequalities in health, education, moral values, employment, etc. Dandekar (1955) suggested certain modifications in Binomial and Poisson distributions which are useful in describing the birth patterns during a given period. He further modified the models for females who have entered into a conjugal relationship. Henry (1965) derived expressions for the expected number of births assuming that a woman has a constant probability of giving a birth if she had not given any birth in the preceding year and has a zero probability if she had given a live birth in the preceding year. Singh (1961, 1963, 1964, 1966, 1968) derived discrete and continuous time models for the number of complete conceptions to a female within a given time period. These models, in fact, were extensions of the models given by Feller (1948) and Neyman (1949). He further extended the models to portray a distribution regarding conception among heterogeneous group of couples assuming that fecundability (p) follows a Beta distribution in the discrete time model.

Singh et al. (1981) proposed a model for the number of complete conceptions (live births) considering fetal wastages, occurring in a couple during a specified period. In one of the recent studies, Rai et al. (2012) proposed a probability model to estimate the number of female births among the married women of seven North-East states. Similarly, another probability model for measuring fecundability has been proposed for the migrant and non-migrant couples of western Uttar Pradesh using the method of moments (Gupta et al., 2016).

In this study, a model is developed using joint probability and further used to estimate the mean number of children ever born through the distribution of male births among the females of Uttar Pradesh and Bihar. The applicability of this model has been illustrated on real data obtained from the National Family Health Survey-3 (2005–06). The various estimates of the parameters have been obtained by using the method of moments and suitability of the proposed model has been tested using the 'goodness of fit' criteria.

The reminder of this paper is organized as follows. In Section 2, Probability models for estimation is presented further the discretion about the data set along with application of proposed model is given in Section 3. Section 4, provides the discussion and results of the statistical analysis and concluding remarks are offered in Section 5.

2 Probability Model for Estimating Mean Number of Children Ever Born

Let us assume that a female gives n number of births in her reproductive span in any sequence of male or female births. Let, birth of a male child is considered to be a success and that of a female child a failure. If *X* denotes the number of births of male child and 'z' be the probability of success, then the distribution of number of male births of a given parity 'n' follows a Binomial distribution, given by,

$$P[X = x|n, z] = \binom{n}{x} z^x (1-z)^{n-x};$$
$$\text{where } 0 \leq z \leq 1;\ n > 0 \quad \text{and} \quad x = 1, 2, 3, \ldots, n. \tag{1}$$

It is assumed that the probability of male births remains constant at each birth for a given female. We further assume that the probability of male births 'z' follows Beta distribution with parameters 'a' and 'b' and is given as,

$$f(p) = \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}; \quad \text{where } 0 \leq z \leq 1 \quad \text{and} \quad a, b > 0 \tag{2}$$

Hence, the compound distribution of x and p for a given value of n will be as follows:

$$\begin{aligned} P\left[X = x \bigcap P = z|n\right] &= P[X = x|n, z] \cdot f(z) \\ &= \binom{n}{x} z^x (1-z)^{n-x} \cdot \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1} \end{aligned} \tag{3}$$

Therefore, the marginal distribution of X for a fixed value of *n* is written as,

$$P[X = x|n] = \int_0^1 \binom{n}{x} z^x (1-z)^{n-x} \cdot \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}\, dz \tag{4}$$

Further, in this model, we assume that the number of parity is a random variable and follows a Poisson distribution,

$$P[n = k] = \frac{e^{-\lambda}\lambda^k}{k!}, \quad \text{where } \lambda \text{ is the average parity and } k = 0, 1, 2 \ldots \tag{5}$$

The joint distribution of X and n is written as,

$$P\left[X = x \bigcap n = k\right] = P[X = x|n] \text{x} P[n = k] \tag{6}$$

or,

$$\begin{aligned} P[X = x] &= \sum_{k=x}^{\infty} \int_0^1 \binom{n}{x} z^x (1-z)^{n-x} \cdot \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1} dz \cdot \frac{e^{-\lambda}\lambda^k}{k!} \\ &= \frac{1}{\beta(a,b)} \int_0^1 \frac{k!}{x!(k-x)!} z^{x+a-1}(1-z)^{(k-x)+b-1} dz \cdot \frac{e^{-\lambda}\lambda^k}{k!} \end{aligned} \tag{7}$$

Let $(k - x) = y$, then (7) becomes,

$$P[X = x] = \frac{1}{\beta(a,b)x!}\int_0^1 \sum_{y=0}^{\infty} z^{a+x-1}(1-z)^{y+b-1}dz \cdot \frac{e^{-\lambda}\lambda^{x+y}}{y!}$$

$$= \frac{\lambda^x}{\beta(a,b)x!}\int_0^1 \sum_{y=0}^{\infty} z^{a+x-1}(1-z)^{y+b-1}dz \cdot \frac{e^{-\lambda}\lambda^{y}}{y!}$$

$$= \frac{\lambda^x}{\beta(a,b)x!}\int_0^1 e^{-\lambda z} z^{a+x-1}(1-z)^{b-1}dz \sum_{y=0}^{\infty} \times \frac{e^{-\lambda(1-z)}\{\lambda(1-z)\}^y}{y!} \quad (8)$$

We know that $\sum_{y=0}^{\infty} \frac{e^{-\lambda(1-z)}\{\lambda(1-z)\}^y}{y!} = 1$, hence (8) reduces to,

$$P[X = x] = \frac{\lambda^x}{\beta(a,b)x!}\int_0^1 e^{-\lambda z} z^{a+x-1}(1-z)^{b-1}dz \quad (9)^1$$

Thus, Equation (9) gives a probability mass function for the numbers of male births to a couple.

2.1 Estimation of Parameters

In this chapter, method of moments have been used to estimate the parameters $\boldsymbol{\lambda}, \boldsymbol{a}, \boldsymbol{b}$ for the proposed probability model. The first three moments for the model are as follows:

$$\mathrm{E(X)} = \frac{\lambda\beta(\mathrm{a+1,b})}{\beta(\mathrm{a,b})} = \frac{\lambda \mathrm{a}}{(\mathrm{a+b})} \quad (10)$$

$$\mathrm{E(X^2)} = \frac{\lambda^2\beta(\mathrm{a+2,b})}{\beta(\mathrm{a,b})} + \frac{\lambda\beta(\mathrm{a+1,b})}{\beta(\mathrm{a,b})} = \frac{\lambda^2(\mathrm{a+1})\mathrm{a}}{(\mathrm{a+b+1})(\mathrm{a+b})} + \frac{\lambda\mathrm{a}}{(\mathrm{a+b})} \quad (11)$$

[1] Rai, P.K., Pareek, S. and Joshi, H. 2014. "On the estimation of probability model for the number of female child births among females", Journal of Data Science, 12, pp. 137–156.

$$\mathrm{E}(\mathrm{X}^3) = \frac{\lambda^3\beta(\mathrm{a}+3,\mathrm{b})}{\beta(\mathrm{a},\mathrm{b})} + \frac{3\lambda^2\beta(\mathrm{a}+2,\mathrm{b})}{\beta(\mathrm{a},\mathrm{b})} + \frac{\lambda\beta(\mathrm{a}+1,\mathrm{b})}{\beta(\mathrm{a},\mathrm{b})}$$
$$= \frac{\lambda^3(\mathrm{a}+2)(\mathrm{a}+1)\mathrm{a}}{(\mathrm{a}+\mathrm{b}+2)(\mathrm{a}+\mathrm{b}+1)(\mathrm{a}+\mathrm{b})} + \frac{3\lambda^2(\mathrm{a}+1)\mathrm{a}}{(\mathrm{a}+\mathrm{b}+1)(\mathrm{a}+\mathrm{b})} + \frac{\lambda\mathrm{a}}{(\mathrm{a}+\mathrm{b})} \tag{12}$$

Let μ_1', μ_2' and μ_3' be the three raw moments for this distribution, and, by replacing E(X), E(X^2) and E(X^3) by μ_1', μ_2' and μ_3' respectively, we get,

$$\mu_1' = \frac{\lambda\mathrm{a}}{(\mathrm{a}+\mathrm{b})} \tag{13}$$

$$\mu_2' = \frac{\lambda^2(\mathrm{a}+1)\mathrm{a}}{(\mathrm{a}+\mathrm{b}+1)(\mathrm{a}+\mathrm{b})} + \frac{\lambda\mathrm{a}}{(\mathrm{a}+\mathrm{b})} \tag{14}$$

$$\mu_3' = \frac{\lambda^3(\mathrm{a}+2)(\mathrm{a}+1)\mathrm{a}}{(\mathrm{a}+\mathrm{b}+2)(\mathrm{a}+\mathrm{b}+1)(\mathrm{a}+\mathrm{b})} + \frac{3\lambda^2(\mathrm{a}+1)\mathrm{a}}{(\mathrm{a}+\mathrm{b}+1)(\mathrm{a}+\mathrm{b})} + \frac{\lambda\mathrm{a}}{(\mathrm{a}+\mathrm{b})} \tag{15}$$

Here λ is the mean number of children ever born to females having at least one child. So,

$$\widehat{\lambda} = \frac{B}{n - n_o} \tag{16}$$

where,

B = total number of births to females,
n = total number of females, and
n_0 = total number of females having no child.

Therefore, with the help of Equations (13), (14), (15) and (16) we can estimate the unknown parameters $\boldsymbol{\lambda}$, $\boldsymbol{a}$, $\boldsymbol{b}$.

2.2 Chi-square Test of 'Goodness of Fit'

A very powerful test for testing the inconsistency between observed and expected value is *"Chi-square test of goodness of fit"*. It enables us to find any deviation between the observed and expected values and explains whether the deviation, if any, is by chance or due to inadequacy of the theoretical model to fit into the data. The formula is given as:

$$\chi^2 = \sum_{i=1}^{n} \frac{(Oi - Ei)^2}{Ei}, \quad \text{where } (\Sigma O_i = \Sigma E_i)$$

Where, O_i is the observed frequency E_i is the expected frequency in each category. The above equation follows a '*chi-square distribution*' with $(n-k)$ degrees of freedom and, $k = 1, 2, 3 \ldots$

3 Data and Application of the Model

The proposed model has been applied on the data obtained from NFHS-3 for the states of Uttar Pradesh and Bihar. *The National Family Health Survey (NFHS) is a large-scale, multi-round survey conducted in a representative sample of households throughout India. The survey provides state and national information for India on fertility, infant and child mortality, the practice of family planning, maternal and child health, reproductive health, nutrition, anemia, utilization and quality of health and family planning services*.[2] Here, females of all parity and from different demographic background have been included in the study. The various demographic backgrounds have been taken as their residential status, educational attainment, religious beliefs, caste, working status and standard of living. In this model, information of all male births, whether alive or not at the time of survey, has been taken into consideration. Childless females have not been considered for estimating the parameters required for this study. The data set contains 12,183 (3,732 childless) females of Uttar Pradesh and 3,818 (1,075 childless) females of Bihar.

4 Results and Discussion

The initial table gives a summary of parameters involved in the probability model. The table shows that the total observed nos. of females in UP and Bihar are 12183 and 3818 respectively out of which 8451 and 2743 have given birth to at least one child. The estimated number of male births (per 1000 births) are 522.98 and 512.06 respectively in UP and Bihar which indicates that male births are more likely to take place than female births in both states. The estimated mean number of births to females who have given birth to at least one child are 3.83 and 3.98 for UP and Bihar, whereas, the same for all the females are 2.69 and 2.74 respectively. The estimated mean number of male births to females having at least one son are 3.31 and 3.42 respectively for UP and Bihar, whereas, the same for all the females are 1.82 and 1.78 births.

[2] http://rchiips.org/nfhs/

Summary of parameters involved in the probability model for Uttar Pradesh and Bihar

Particulars	Uttar Pradesh	Bihar
Total no. of females	12183	3818
Total no. of childless females	3732	1075
Total no. of females having at least one child	8451	2743
Estimated no. of male births (per 1000 births)	522.98	512.06
Estimated mean no. of births to females having at least one child (λ)	3.83	3.98
Estimated mean no. of births to all females (λ_0)	2.69	2.74
Estimated mean no. of male births to a female having at least one son (λ_1)	3.31	3.42
Estimated mean no. of male births to all females (λ_2)	1.82	1.78

Table 1 Estimated values of various parameters based on residential background of the females of Uttar Pradesh

Urban			Rural		
No. of Male Births	Observed Frequency	Expected Frequency	No. of Male Births	Observed Frequency	Expected Frequency
0	2180	2155.300	0	2551	2608.004
1	1107	1053.553	1	1355	1389.646
2	1010	1019.817	2	1396	1353.355
3	472	495.521	3	874	916.135
4	227	250.433	4	470	431.906
5	86	94.365	5	231	202.790
6	56	64.003	6	95	77.070
7+	20	25.008	7	39	34.021
			8+	14	12.073
Total	**5158**	**5158**	**Total**	**7025**	**7025**
Parameters					
λ	3.520		λ	4.121	
a	7.364		a	6.290	
b	4.124		b	3.050	
d.f.	4		d.f.	5	
χ^2 (cal.)	9.143**		χ^2 (cal.)	17.883	
$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	11.07	
$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	15.086	

* Significant at 1% level & ** Significant at 5% level.

Tables 1 & 2 describe the estimated values of various parameters as per the residential background of the females of UP and Bihar. In UP, there are 2180 childless females out of 5158 urban females and 2551 childless females out of 7025 rural females. Similarly in Bihar, there are 650 childless

Table 2 Estimated values of various parameters based on residential background of the females of Bihar

Urban			Rural		
No. of Male Births	Observed Frequency	Expected Frequency	No. of Male Births	Observed Frequency	Expected Frequency
0	650	636.102	0	789	812.898
1	297	304.098	1	505	491.493
2	302	285.507	2	471	478.902
3	143	152.423	3	308	313.577
4	62	69.073	4	139	121.927
5	30	33.799	5	61	58.201
6+	18	20.997	6	31	28.084
			7+	12	10.919
Total	**1502**	**1502**	**Total**	**2316**	**2316**
Parameters					
λ	3.690		λ	3.876	
a	6.016		a	5.727	
b	2.844		b	2.555	
d.f.	3		d.f.	4	
χ^2 (cal.)	2.943**		χ^2 (cal.)	4.239**	
$\chi^2_{0.05}$ (tab.)	7.815		$\chi^2_{0.05}$ (tab.)	9.488	
$\chi^2_{0.01}$ (tab.)	11.345		$\chi^2_{0.01}$ (tab.)	13.277	

* Significant at 1% level & ** Significant at 5% level.

females out of 1502 urban and 789 childless females out of the 2316 rural females. The estimated values of mean number of children ever born (λ), and parameters 'a' and 'b' are 3.520 and 7.364 and 4.124 respectively for the urban areas of UP. Here the calculated value of χ^2 is 9.143 at 4 d.f. and it is significant at 0.01 level. This indicates that the proposed probability model is suitable to describe the distribution of male births in urban Uttar Pradesh. In case of rural areas of UP, λ, a and b values are estimated as 4.121, 6.290 and 3.050 respectively whereas the χ^2 value is 17.883 at 5 d.f. The value is insignificant and hence the probability model does not fit to the rural data. Similarly, in urban areas of Bihar, λ, a and b are estimated at 3.690, 6.016 and 2.844 respectively, whereas the χ^2 value is 2.943 at 3 d.f. and it is significant at 5% level of significance. This shows that the model is a 'good fit' for the distribution of male births in urban areas of Bihar. In case of rural areas the estimated values of the parameter, λ, a and b are 3.876, 5.727 and 2.555 respectively and the χ^2 value is 4.239 at 4 d.f. which is acceptable at 0.05 level and is significant. Hence the model suits to the distribution of male births in rural Bihar as well.

Tables 3 & 4 show the estimated values of various parameters according to educational background of the females of UP and Bihar. The results represent that the proposed probability model does not describe the distribution of male births for uneducated females in UP ($\chi^2 = 18.649$; insignificant), whereas it proves to be a 'good fit' for the primary, secondary and highly educated females (χ^2 values are 2.991, 2.651 and 3.046, all being significant at 5% level). Similarly, in case of Bihar, the χ^2 values for all the four segments are 18.736, 8.322, 4.525 and 2.795 respectively, which means that the model does not fit well for uneducated females, it is fairly fits for primary educated females (significant at 1% level) and it is a 'good fit' for secondary and highly educated females (significant at 5% level).

Tables 5 & 6 represent the estimated values of different parameters according to religious background of the females. In UP, the λ, a and b values for Hindus are 3.749, 6.914 and 3.674 respectively. The χ^2 value is 6.160 at 4 d.f. which is highly significant at 5% level, indicating the model to be a 'good fit' for the given data. Similarly, for Muslim females, λ, a and b values are 4.457, 5.817 and 2.577 respectively. We get a very high value of χ^2 (16.540 at 4 d.f.) which makes it unbefitting model for the given data. In case of other religions, the χ^2 value is 2.829 at 1 d.f. and is highly significant at 5% level, which makes the model a 'good fit' for the distribution of male births in UP. For Bihar the χ^2 values for Hindus and Muslims are 3.635 and 10.816 respectively. Both the values are significant at 5% and 1% level and portray that the model is suitable to describe the distribution of male births among Hindus and Muslims in Bihar. The total count of females belonging to other religions was insignificant.

Tables 7 & 8 correspond to the estimated values of parameters on the basis of caste of females for UP and Bihar. The χ^2 values for SC and Gen/OBC category are 3.814 and 2.951 (both at 4 d.f.) respectively which are acceptable at 5% level of significance. This indicates the suitability of the proposed probability model for distribution of male child-births among SC and Gen/OBC category females of UP. In case of ST females, a very high value of χ^2 is obtained ($\chi^2 = 18.038$ at 2 d.f.) which signifies that the model does not fit well for the specified category of females. In Bihar, the χ^2

Table 3 Estimated values of various parameters based on educational background of the females of Uttar Pradesh

No Education			Primary			Secondary			Higher		
No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.
0	1319	1358.788	0	631	613.586	0	2111	2074.588	0	667	661.038
1	1117	1156.146	1	261	273.151	1	702	711.713	1	381	375.991
2	1366	1389.637	2	289	281.711	2	574	583.316	2	177	182.337
3	1012	958.741	3	138	147.599	3	174	181.291	3	22	29.369
4	580	565.938	4	59	61.610	4	53	58.801	4+	5	3.267
5	275	232.787	5	30	32.117	5+	17	21.292			
6	133	122.779	6+	18	16.227						
7	47	55.063									
8+	21	30.122									
Total	**5870**	**5870**	**Total**	**1426**	**1426**	**Total**	**3631**	**3631**	**Total**	**1252**	**1252**
Parameters											
λ	4.575		λ	3.624		λ	2.814		λ	2.090	
a	5.666		a	7.153		a	9.214		a	12.407	
b	2.423		b	3.913		b	5.974		b	9.167	
d.f.	5		d.f.	3		d.f.	2		d.f.	1	
χ^2 (cal.)	18.649		χ^2 (cal.)	2.291**		χ^2 (cal.)	2.651**		χ^2 (cal.)	3.046**	
$\chi^2_{0.05}$ (tab.)	11.070		$\chi^2_{0.05}$ (tab.)	7.815		$\chi^2_{0.05}$ (tab.)	5.991		$\chi^2_{0.05}$ (tab.)	3.841	
$\chi^2_{0.01}$ (tab.)	15.086		$\chi^2_{0.01}$ (tab.)	11.345		$\chi^2_{0.01}$ (tab.)	9.210		$\chi^2_{0.01}$ (tab.)	6.635	

Table 4 Estimated values of various parameters based on educational background of the females of Bihar

No Education			Primary			Secondary			Higher		
No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.
0	550	578.7093	0	144	136.153	0	653	644.318	0	89	85.8198
1	431	449.3924	1	77	75.85123	1	245	239.088	1	47	41.66841
2	469	449.0136	2	82	73.19644	2	193	189.4049	2	28	31.38502
3	336	358.8059	3	39	42.95076	3	72	80.47983	3	6	8.763489
4	172	152.5943	4	14	19.43688	4	14	18.57255	4+	5	7.363279
5	80	64.98088	5+	6	14.41173	5+	7	12.13672			
6	41	33.61734									
7+	18	9.886328									
Total	**2097**	**2097**	**Total**	**362**	**362**	**Total**	**1184**	**1184**	**Total**	**175**	**175**
Parameters											
λ	4.357		λ	3.284		λ	2.837		λ	2.105	
a	5.095		a	6.760		a	7.824		a	10.548	
b	1.924		b	3.589		b	4.653		b	7.376	
d.f.	4		d.f.	2		d.f.	2		d.f.	1	
χ^2 (cal.)	18.736		χ^2 (cal.)	8.322*		χ^2 (cal.)	4.525**		χ^2 (cal.)	2.795**	
$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	5.991		$\chi^2_{0.05}$ (tab.)	5.991		$\chi^2_{0.05}$ (tab.)	3.841	
$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	9.210		$\chi^2_{0.01}$ (tab.)	9.210		$\chi^2_{0.01}$ (tab.)	6.635	

Table 5 Estimated values of various parameters based on religious background ofthe females of Uttar Pradesh

Hindu			Muslim			Others		
No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.
0	3581	3567.941	0	1079	1096.892	0	63	60.110
1	2004	1985.920	1	407	418.392	1	45	40.698
2	1954	1945.503	2	424	447.417	2	26	28.069
3	1050	1028.512	3	283	282.836	3	7	10.203
4	501	537.187	4	193	176.695	4+	5	7.921
5	216	221.833	5	99	86.488			
6	96	103.364	6	55	41.872			
7+	47	58.739	7+	26	15.408			
Total	9449	9449	Total	2566	2566	Total	146	146
Parameters								
λ	3.749		λ	4.457		λ	2.552	
a	6.914		a	5.817		a	10.157	
b	3.674		b	2.577		b	6.916	
d.f.	4		d.f.	4		d.f.	1	
χ^2 (cal.)	6.160**		χ^2 (cal.)	16.540		χ^2 (cal.)	2.829**	
$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	3.841	
$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	6.635	

* Significant at 1% level & ** Significant at 5% level.

Table 6 Estimated values of various parameters based on religious background of the females of Bihar

Hindu			Muslim		
No. of Male Births	Observed Frequency	Expected Frequency	No. of Male Births	Observed Frequency	Expected Frequency
0	1153	1168.615	0	281	275.385
1	692	679.238	1	104	117.666
2	671	660.334	2	102	115.762
3	358	370.68	3	93	90.32
4	153	145.203	4	48	43.797
5	65	59.794	5	26	19.206
6	28	32.342	6	15	10.658
7+	9	12.794	7+	9	5.206
Total	**3129**	**3129**	**Total**	**678**	**678**
Parameters					
λ	3.648		λ	4.657	
a	6.086		a	4.767	
b	2.914		b	1.596	
d.f.	4		d.f.	4	
χ^2 (cal.)	3.635**		χ^2 (cal.)	10.816*	
$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	9.488	
$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	13.277	

* Significant at 1% level & ** Significant at 5% level.

values for SC and Gen/OBC category are 2.421 and 4.755 which are highly significant at 5% level. This indicates the model to be a 'good fit' for defining the distribution of male births among the females of given categories. The total count of ST females was insignificant for the state and hence could not be included in the analysis.

Tables 9 & 10 portray the estimated values of various parameters for working status of the females in both states. In case of non-working females of UP, the χ^2 turns out to be 14.987 whereas it is 5.721 for working females. Though both the values are significant, but a higher value is obtained for non-working category which shows that the model may be a 'good fit' at 1% level of significance. On the other hand, the model duly describes the distribution of male births among working females of the state. In case of Bihar, the

Table 7 Estimated values of various parameters based on caste of the females of Uttar Pradesh

Scheduled Caste			Scheduled Tribe			General/OBC		
No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.
0	921	941.472	0	23	32.218	0	3780	3740.310
1	428	441.512	1	11	16.777	1	2021	1991.711
2	530	519.640	2	17	16.395	2	1857	1877.965
3	362	345.136	3	14	9.173	3	969	990.691
4	205	197.550	4	10	4.747	4	481	493.704
5	83	77.203	5+	8	3.690	5	228	237.635
6	31	32.012				6	120	117.959
7+	20	25.476				7+	51	57.026
Total	**2580**	**2580**	**Total**	**83**	**83**	**Total**	**9507**	**9507**
λ	4.234		λ	4.794		λ	3.766	
a	6.122		a	5.407		a	6.882	
b	2.882		b	2.167		b	3.642	
d.f.	4		d.f.	2		d.f.	4	
χ^2 (cal.)	3.814**		χ^2 (cal.)	18.038		χ^2 (cal.)	2.951**	
$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	5.991		$\chi^2_{0.05}$ (tab.)	9.488	
$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	9.21		$\chi^2_{0.01}$ (tab.)	13.277	

* Significant at 1% level & ** Significant at 5% level.

Table 8 Estimated values of various parameters based on caste of the females of Bihar

Scheduled Caste			General/OBC		
No. of Male Births	Observed Frequency	Expected Frequency	No. of Male Births	Observed Frequency	Expected Frequency
0	194	203.804	0	1236	1216.196
1	106	119.826	1	695	681.240
2	117	106.760	2	651	653.174
3	65	56.982	3	383	381.018
4	50	45.052	4	151	170.948
5	25	23.456	5	65	76.544
6+	11	12.120	6+	50	51.880
Total	568	568	Total	3231	3231
λ	4.122		λ	3.752	
a	5.386		a	5.917	
b	2.215		b	2.745	
d.f.	3		d.f.	3	
χ^2 (cal.)	2.421**		χ^2 (cal.)	4.755**	
$\chi^2_{0.05}$ (tab.)	7.815		$\chi^2_{0.05}$ (tab.)	7.815	
$\chi^2_{0.01}$ (tab.)	11.345		$\chi^2_{0.01}$ (tab.)	11.345	

* Significant at 1% level & ** Significant at 5% level.

Table 9 Estimated values of various parameters based on working status ofthe females of Uttar Pradesh

Not working			Working		
No. of Male Births	Observed Frequency	Expected Frequency	No. of Male Births	Observed Frequency	Expected Frequency
0	3560	3506.043	0	1158	1176.957
1	1881	1852.326	1	574	602.674
2	1657	1625.508	2	747	728.492
3	894	920.267	3	450	433.733
4	431	461.738	4	264	253.262
5	204	228.128	5	112	107.872
6	98	118.289	6	52	46.711
7	35	37.540	7+	23	30.299
8+	15	25.160			
Total	**8775**	**8775**	**Total**	**3380**	**3380**
Parameters					
λ	3.672		λ	4.384	
a	7.058		a	5.912	
b	3.818		b	2.672	
d.f.	5		d.f.	4	
χ^2 (cal.)	14.987*		χ^2 (cal.)	5.721**	
$\chi^2_{0.05}$ (tab.)	11.07		$\chi^2_{0.05}$ (tab.)	9.488	
$\chi^2_{0.01}$ (tab.)	15.086		$\chi^2_{0.01}$ (tab.)	13.277	

* Significant at 1% level & ** Significant at 5% level.

Table 10 Estimated values of various parameters based on working status ofthe females of Bihar

Not working			Working		
No. of Male Births	Observed Frequency	Expected Frequency	No. of Male Births	Observed Frequency	Expected Frequency
0	1245	1221.793	0	192	200.207
1	646	618.661	1	155	171.489
2	593	601.511	2	180	167.339
3	321	336.781	3	130	129.219
4	118	129.009	4	83	76.991
5	56	61.989	5	35	29.011
6	29	34.017	6	14	13.983
7+	10	14.240	7+	8	8.760
Total	3018	3018	Total	797	797
Parameters					
λ	3.568		λ	4.577	
a	6.221		a	4.850	
b	3.050		b	1.679	
d.f.	4		d.f.	4	
χ^2 (cal.)	6.029**		χ^2 (cal.)	3.445**	
$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	9.488	
$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	13.277	

* Significant at 1% level & ** Significant at 5% level.

χ^2 values are 6.029 and 3.445 respectively for non-working and working females. Since both the values are highly significant, it can be inferred that the projected model proves to be a ‘good fit’ for the given set of data.

Tables 11 & 12 represent the estimated values of parameters according to standard of living of the females of UP and Bihar. For UP, the results show that the proposed model is not at all suitable to describe male birth patterns among the low income group ($\chi^2 = 23.022$), whereas moderately acceptable values are obtained for the middle and high income groups (χ^2 being 11.035 and 9.096 respectively). This indicates that the model may be suitable to describe the male birth patterns among middle and high income families of the state. In case of Bihar, the χ^2 values are 9.785, 6.721 and 5.958 respectively which are acceptable at 5% level of significance. Accordingly, the model is suitable for various income groups in Bihar.

Table 11 Estimated values of various parameters based on standard of living of the females of Uttar Pradesh

Low			Medium			High		
No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.
0	605	655.519	0	1325	1340.223	0	2318	2248.695
1	382	395.670	1	634	679.969	1	1211	1199.245
2	452	479.875	2	766	752.584	2	1045	1067.303
3	320	286.889	3	506	482.229	3	470	498.724
4	199	170.081	4	287	267.734	4	192	195.995
5	102	90.490	5	133	115.900	5	74	91.592
6	51	41.667	6	62	75.440	6+	48	56.447
7+	27	17.809	7+	30	28.922			
Total	**2138**	**2138**	**Total**	**3743**	**3743**	**Total**	**5358**	**5358**
Parameters								
λ	4.700		λ	4.311		λ	3.346	
a	5.515		a	6.012		a	7.746	
b	2.275		b	2.772		b	4.506	
d.f.	4		d.f.	4		d.f.	3	
χ^2 (cal.)	23.022		χ^2 (cal.)	11.035*		χ^2 (cal.)	9.096*	
$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	7.815	
$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	11.345	

* Significant at 1% level & ** Significant at 5% level.

Table 12 Estimated values of various parameters based on standard of living of the females of Bihar

Low			Medium			High		
No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.	No. of Male Births	Obs. Freq.	Exp. Freq.
0	369	387.1452	0	366	385.5782	0	506	495.2766
1	239	255.7431	1	200	220.286	1	244	233.0491
2	239	250.174	2	247	212.2078	2	223	221.54
3	202	185.6102	3	133	133.2901	3	94	100.0996
4	96	82.00647	4	65	59.96502	4	32	43.02851
5	58	47.28277	5	24	27.65226	5+	12	18.00617
6	18	15.87184	6+	22	18.02058			
7+	13	10.16637						
Total	**1234**	**1234**	**Total**	**1057**	**1057**	**Total**	**1111**	**1111**
				Parameters				
λ	4.399		λ	4.090		λ	3.267	
a	5.047		a	5.498		a	6.795	
b	1.875		b	2.256		b	3.624	
d.f.	4		d.f.	3		d.f.	2	
χ^2 (cal.)	9.785*		χ^2 (cal.)	6.721**		χ^2 (cal.)	5.958**	
$\chi^2_{0.05}$ (tab.)	9.488		$\chi^2_{0.05}$ (tab.)	7.815		$\chi^2_{0.05}$ (tab.)	5.991	
$\chi^2_{0.01}$ (tab.)	13.277		$\chi^2_{0.01}$ (tab.)	11.345		$\chi^2_{0.01}$ (tab.)	9.210	

* Significant at 1% level & ** Significant at 5% level.

5 Conclusion

From the results discussed above, it is observed that the expected frequencies obtained from the marginal distribution are very close to the observed frequencies barring a few cases. Hence it could be established that the proposed probability model fits well in most of the cases to describe the distribution of the number of male child births to females of all parity in the states of Uttar Pradesh and Bihar. The proposed model may also be generalized to other states where socio-economic status of females match with that of the above three states. The study also gives an insight about the impact of male births in the society by assessing the relationship between sex composition of children and continued childbearing. Since the probability of a specified number of male births can be figured out with the help of this model, proper policies could be framed to maintain an ideal sex ratio in those regions where the number of female births is reducing distressingly.

References

Basu, D. and Jong, R. (2010). "Son Targeting Fertility Behavior: Some Consequences and Determinants", *Demography*, 47(2), pp. 521–536.

Brass, W. (1958). "The Distribution of Births in Human Populations n RuralTaiwan", *Population Studies*, 12, pp. 51–72.

Chaudhuri, S. (2012). "The Desire for Sons and Excess Fertility: A Household level Analysis of Parity Progression in India", *International Perspectives on Sexual and Reproductive Health*, 38(4), pp. 178–186.

Gupta, C.B., Kumar, S. and Singh, B.P. (2016). "A Probability Model for Measuring Fecundability of Migrant and Non-migrant Couples", *International Journal of Statistics and Applications*, 6(3), pp. 163–167.

National Family Health Survey-III (2005–06). International Institute for Population Sciences, Mumbai.

Pathak, K.B. (1966). "A Probability Distribution for the Number of Conceptions", *Sankhya: The Indian Journal of Statistics*, 28(3/4), pp. 213–218.

Pathak, K.B. and Murty, P.K. (1985). "Socio-economic Determinants of Fertility and Mortality Decline in India", *Demography India,* 14(1), pp. 17–33.

Rai, P. K., Pareek, S. and Joshi, H. (2013). "Regression Analysis of Collinear-Data using r-k Class Estimator: Socio-economic and Demographic FactorsAffecting the Total Fertility Rate in India", *Journal of Data Science*, 11, pp. 321–340.

Rai, P.K., Pareek, S. and Joshi, H. (2014). "On the Estimation of Probability Model for the Number of Female Child", *Journal of Data Science*, 12, pp. 137–156.

Roy, S., Singh, K.K., Singh, B.P. and Gupta, K. (2015). "Study of Influence of Caste Differentials on Fertility and Contraception", *Journal of Statistics Applications & Probability Letters*, 2(2), pp. 149–160.

Roy, S., Pandey, A. and Singh, K.K. (2016). "Levels and Trends of Wanted and Unwanted Fertility in Uttar Pradesh", *Elixir International Journal of Statistics*, 94, pp. 39923–39929.

Roy, S. and Singh, K.K. (2016). "Effects of Economic Factors on Fertility in Uttar Pradesh: A Regression Analysis", *International Journal of Research in Economics and Social Sciences*, 6(7), pp. 142–153.

Seidl, C. (1995). "The Desire for a Son is the Father of Many Daughters: A Sex Ratio Paradox", *Journal of Population Economics*, 8(2), pp. 185–203.

Singh, B.P., Maheshwari, S., Pundir, P.S. and Singh, K.K. (2013). "Pattern of Human Fertility Behaviour Over Time: An Application of Bivariate Probability Model", *Journal of Rajasthan Statistical Association*, 2(1), pp. 102–107.

Singh, B.P., Gupta, K. and Singh, K.K. 2015. "Analysis of Fertility Pattern through Mathematical Curves", *American Journal of Theoretical and Applied Statistics*, 4(2), pp. 64–70.

Singh, K. K., Singh, B. P. and Singh, N. (2012). "A Probabilistic Study of Variation in Number of Child Deaths", *Journal of Rajasthan Statistical Association*, 1, pp. 54–67.

Singh, S. N. (1963). "Probability Model for the Variation in the Number of Births per Couple", *Journal of American Statistical Association*, 58, pp. 721–727.

Singh, S. N. (1964). "A Probability Model for Couple Fertility", *Sankhya Series B*, 26, pp. 89–94.

Singh, S. N. (1966). "Some Probability Distributions utilized in Human Fertility", Seminar Volume in Statistics. Banaras Hindu University, Varanasi, pp. 74–84.

Singh, S. N. (1968). A Chance Mechanism of the Variation in the Number of Births per Couple", *Journal of American Statistical Association*, 63, pp. 209–213.

Singh, S. N. and Bhattacharya, B. N. (1970). "A Generalized Probability Distribution for Couple Fertility", *Biometrics*, 26, pp. 33–40.

Singh, S. N., Bhattacharya, B. N. and Joshi, P. D. (1973). "On a Hypothetical Model for the Number of Births and its Applications", *Journal of the Indian Society of Agricultural Statistics*, 25, pp. 91–100.

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, pp. 93–102.

Singh, S. N. and Pathak, K. B. (1968). "On the Distribution of the Number of Conceptions", *Journal of Mathematical Society*, 1, pp. 41–46.

Singh, S. N. and Singh, V. K. (1981). "A Probability Model for the Number of Births in an Equilibrium Birth Process", *Journal of Biosciences*, 3, pp. 197–205.

Biographies

Shubhagata Roy is an Assistant Professor in the Department of Operations and IT at IBS Hyderabad. He has more than 20 years of experience in industry and academia. His teaching areas are Business Analytics, Business Statistics, Predictive Modeling, Operations Research, Research Methodology, Operations Management and Quality Management. His research areas include Demography, Healthcare Analytics, Insurance Analytics, Ethics and Sustainability. He holds a PhD (Statistics) from Banaras Hindu University.

Prayas Sharma is currently working as Assistant Professor in the area of Decision Sciences at Indian Institute of Management Sirmaur, Paonta Sahib, Himachal Pradesh. He has more than 10 years of academic experience, both in the domain of teaching and research. His research interest includes Survey Sampling, Estimation Procedures using Auxiliary Information and Measurement Errors, Predictive Modelling, Business Analytics and Operations Research. Dr. Sharma has published more than 40 research papers in reputed National & International journals along with one book and two chapters in book internationally published. He has more than **400 citations with H-Index 14 & I index of 15**. Dr. Sharma has a keen interest in reading, writing and publishing, he is serving 7 reputed journals as editor/associate editor and more than 30 journals as reviewer and reviewed more than 150 research papers from the prestigious.

K. K. Singh is a Professor in the Department of Statistics, Institute of Science, Banaras Hindu University Varanasi. He has more than 40 years of experience in teaching various courses in the field of Statistics. He has successfully completed several projects funded by ICMR, UGC, Rockefeller Foundation, WHO etc. He has authored more than 150 research papers and four books in collaboration with renowned Statisticians in the area of Demography and Population Science. He has been in the boards of several

Central and State Universities since past three decades. Dr. Singh holds a Ph.D. (Statistics) from Banaras Hindu University and Post-Doc from Carolina Population Center, University of North Carolina, USA.

Richa Srivastava is currently working as Assistant Professor in the area of Decision Sciences at Jaipuria Institute of Management, Lucknow, Uttar Pradesh. She is an academician and researcher having more than 8 years of experience in the field of Statistics. Her research interest includes Bayesian Statistics, Multivariate Analysis, Business Analytics, Biostatistics and Operations Research. Her research papers have been extensively published in reputed refereed journals. She has also delivered an invited talk in the International Workshop/Conference on Bayesian Theory and Applications (IWCBTA) and also won the second-best paper presentation award in IWCBTA, held in 2013. She is also associated with various academic bodies.