Journal of Data Science **12**(2014), 137-156

On the Estimation of Probability Model for the Number of Female Child Births among Females

Piyush Kant Rai*, Sarla Pareek and Hemlata Joshi
Banasthali University

Abstract: The study of pattern of female child birth is one of the most crucial area of human demography because it plays very important role in the building of a nation. In the present study, an attempt has been made to work-out the pattern of female child births among females belongs to different subdomains of population through the probability model and the parameters involved in the probability model under consideration has also been estimated. The suggested model, for illustration has been applied to an observed set of data taken from NFHS-III (2005-06) for the seven North-East states of India known as Seven Sisters.

Key words: Fecundability, female child birth, fertility, probability distribution, Seven Sisters.

1. Introduction

Fertility has central importance in demographic analysis as births are vital component of population dynamics. Level and trend of fertility indicates the standard of development of nation. Fertility is defined as the actual child bearing performance of woman or a group of women measured in terms of the actual number of children born to them. Female child plays most important role in fertility dynamics and in the nation development. She establishes the institution of family life, brings up the children and makes them good citizens. Her strength in totality contributes in the making of an ideal society and thus ideal country.

Birth of female child is of major concern in research in the field of demography because of its apparent relationship with the level of fertility. The connection between birth of females and fertility or vice versa are the root of many explanations of demographic transition and in shaping of population distribution. Here, sex ratio is a great source to find the equality of males and females in a society

*Corresponding author.

at a given period of time. Sex ratio is a term used to define number of females per 1000 males (census of India). In India, sex ratio was balanced till the time of Independence, thereafter it has declined regularly (Figures 1 and 2). According to the Census of India 2011, sex ratio has shown some improvement in the last 10 years. It has gone up from 933 in 2001 to 940 in 2011. There has been some improvement in the sex ratio of India but child sex ratio (sex ratio of children under 6 years of age) is even lower than overall female sex ratio (as shown in Table 1 and Figure 2). Child sex ratio has gone down during all past decades in India. According to the 2011 census data, in the last decade alone, the child sex ratio has dipped from 927 girls to 914 girls per 1000 boys which is decreased by 1.40%, the lowest since independence (1947). This is not only the single case, because the total female population in general is very low in comparison to male population in all the last decades. This shows that there is a biggest disparity between the choice of male and female child. The government of India faces the problem of continuing decline in the number of female child during past decades (as given in Table 1 and Figure 1). So, it is necessary to understand the behaviour and trend of female child births to the mothers.

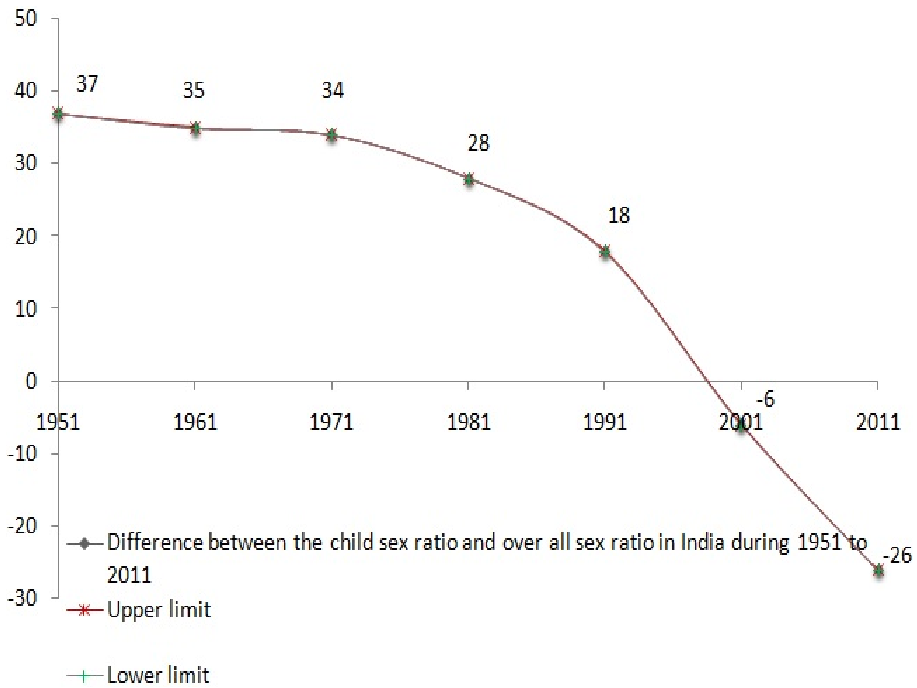

Figure 1: Difference between child sex ratio and over all sex ratio in India during 1951 to 2011

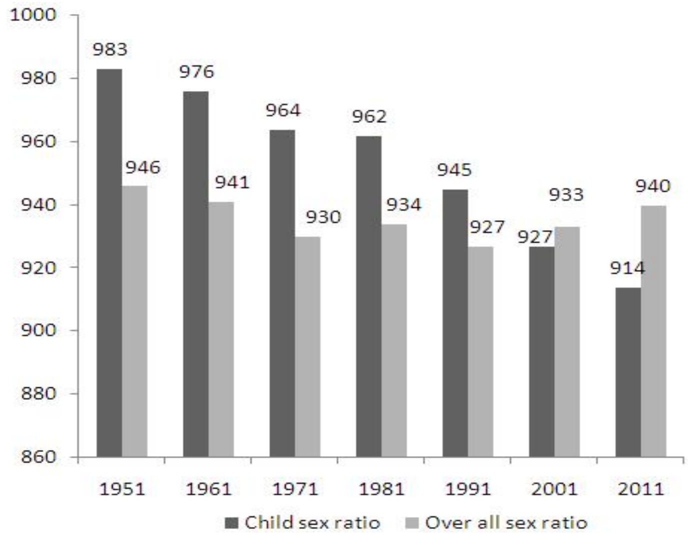

Figure 2: Child and over all sex ratio in India during 1951 to 2011

Table 1: Child sex ratio and over all sex ratio of India during 1951 to 2011

Year	1951	1961	1971	1981	1991	2001	2011
Child Sex Ratio	983	976(-7)	964(-12)	962(-2)	945(-17)	927(-18)	914(-13)
Sex Ratio	946	941(-5)	930(-11)	934(+4)	927(-7)	933(+6)	940(+7)
Difference	37	35	34	28	18	-6	-26
Confidence limit(UL)	37.02	35.02	34.02	28.02	18.02	-5.98	-25.98
Confidence limit(LL)	36.98	34.98	33.98	27.98	17.98	-6.02	-26.02

Henry (1956) derived expressions for the expected number of births in a period of five years assuming that a woman had a constant probability of giving a live birth if she had given no birth in the preceding year and has a zero probability if she had given a birth in the preceding year. Dandekar (1955) has given certain modifications of binomial and poisson distributions which are useful for describing the variation in the number of births to a female during a given period (O, T) under the assumptions (i) each delivery ends in a live birth and (ii) a constant period of non-susceptibility is associated with each delivery. He also obtained their modified forms for the case when the start of the observational period is at a very distant point since marriage. Brass (1958) introduced some modifications into Dandekar's model and applied it to different problems.

Singh (1961, 1963, 1964, 1966, 1968) has derived the discrete and continuous time models for the number of complete conceptions (a conception is complete if it results in a live birth) to a female within time interval (O, T) under different set of assumptions. These distributions are extension of the distributions given by Feller (1948) for counter problems, Neyman (1949) for the number of schools of fish and Dandekar's modifications of Binomial and Poisson distributions. Singh (1963,

1966) has also extended the models (Singh, 1963, 1966) to describe distribution of complete conceptions for a heterogeneous group of couples and has assumed that fecundability (p) follows beta-distribution in discrete time model.

Pathak (1966) introduced one more refinement in Singh's discrete time model. He assumed that a certain proportion of fecund women are not susceptible at the beginning of the period because either they are pregnant or are in the post-partum amenorrhoea. Singh and Pathak (1968) introduced the above refinement in Singh's continuous time model. These models gave more or less same fit as those of Singh (1963, 1968). Sheps and Perrin (1966) derived a probability distribution for the number of complete conceptions incorporating foetal losses which is an extension of the model given in Singh (1963). Following Sheps and Perrin, Singh *et al.* (1973) have generalized the probability distribution of Singh (1968) for the number of complete conceptions to a couple during a given period of time under the assumptions of Singh and Bhattacharya (1970). Singh *et al.* (1974), derived probability models as a means of describing the variation in the number of births to a couple in a given period of time and Singh *et al.* introduced probability model for the number of complete conceptions taking into account foetal wastages, occurring in a couple during a specified period of time. Singh *et al.* (1981) also derived a probability model for number of births in an equilibrium birth process and Singh *et al.* (2012) developed a probability model for the number of child deaths among females.

In this paper an attempt has been made to develop a probability model to explain the pattern of births of female child for all females in the society. The applications of this model are illustrated through the real data taken from National Family Health Survey-III (NFHS-III) for seven North-East states (Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland and Tripura) of India. The estimates of the parameters are also obtained and the suitability of proposed discrete probability model are checked on the basis of goodness of fit test using observed data.

2. Probability Model for the Number of Birth of Female Child

Let, the woman having n number of children out of which there may be any sequence of births of male and female child and the occurrence of birth of female child is said to be success and its non occurrence (i.e., birth of male child) is failure. Thus, X denote the number of births of female child to a female and p be the probability of success (i.e., giving birth to the female child) then the distribution of number of female child births to the females of given parity n follows the binomial distribution given as

$$P[X = x|n, p] = \binom{n}{x} p^x (1-p)^{n-x}, \quad 0 \le p \le 1,\ n > 0,$$

where $x = 0, 1, 2, \cdots, n$.

It is assumed that the probability of female births remain constant at each birth for a given female. We assume that probability of giving birth to the female child p follows beta distribution (due to its flexibility) with parameters a and b, which is given as

$$f(p) = \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1}, \quad 0 \leq p \leq 1,\ a, b > 0.$$

Therefore, the joint distribution of x and p for given n is given by

$$\begin{aligned} P[X = x \cap P = p|n] &= P[X = x|n, p] \times f(p) \\ &= \binom{n}{x} p^x (1-p)^{n-x} \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1}. \end{aligned}$$

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

$$P[X = x|n] = \int_0^1 \binom{n}{x} p^x (1-p)^{n-x} \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1} dp. \tag{2.1}$$

Further in this model, we assume that the number of parity (number of children ever born to a female) is also a random variable and follows a Poisson distribution

$$P[n = k] = \frac{e^{-\lambda}\lambda^k}{k!},$$

where $k = 0, 1, 2, \cdots, n$ and λ is the average parity. The joint distribution of X and n is, written as

$$P[X = x \cap n = k] = P[X = x|n] \times P[n = k].$$

Hence, Marginal distribution of X is given as

$$P[X = x] = \sum_{k=x}^{\infty} \int_0^1 \binom{k}{x} p^x (1-p)^{n-x} \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1} dp \times \frac{e^{-\lambda}\lambda^k}{k!}. \tag{2.2}$$

After simplification, the (2.2) reduces to

$$P[X = x] = \frac{\lambda^x}{\beta(a,b)x!} \int_0^1 e^{-\lambda p} p^{a+x-1}(1-p)^{b-1} dp. \tag{2.3}$$

It is easy to verify that

$$\sum_{x=0}^{\infty} P[X = x] = 1.$$

Thus, $P[X = x]$ is a probability mass function for the birth of females to females.

We may modify the proposed model if the probability of giving birth to the female child p follows inverted beta distribution with parameters a and b,

$$f(p) = \frac{1}{\beta(a,b)} \frac{p^{b-1}}{(1+p)^{a+b}}, \quad p > 0,\ a, b > 0.$$

Therefore, the marginal distribution of X at fixed n is written as

$$P[X = x|n] = \int_0^\infty \binom{n}{x} p^x (1-p)^{n-x} \frac{1}{\beta(a,b)} \frac{p^{b-1}}{(1+p)^{a+b}} dp,$$

and by using poisson density, the marginal distribution of X can be written as

$$P[X = x] = \frac{\lambda^x}{\beta(a,b)x!} \int_0^\infty e^{-\lambda p} \frac{p^{b+x-1}}{(1+p)^{a+b}} dp. \tag{2.4}$$

Also, the proposed model gives the marginal density of number of female child birth of given parity n by taking p as $U(a, b)$, i.e., $f(p) = 1/(b - a)$, $a < p < b$.

3. Estimation Procedure

In this paper, the method of moments is used to estimate the unknown parameters (λ, a and b) of the model given in (2.3) considered for number of female child births to females of all parity. The method of moments provides estimates which are consistent but not as efficient as maximum likelihood estimator. This method of moments is often used because it is very simple in computation than the other methods. Therefore, the first three moments of the probability model given in (2.3) can be carried out as follows

$$\begin{aligned}
E(X) &= \frac{\lambda a}{a+b}, \\
E(X^2) &= \frac{\lambda^2 (a+1)a}{(a+b+1)(a+b)} + \frac{\lambda a}{a+b}, \\
E(X^3) &= \frac{\lambda^3 (a+2)(a+1)a}{(a+b+2)(a+b+1)(a+b)} + \frac{3\lambda^2 (a+1)a}{(a+b+1)(a+b)} + \frac{\lambda a}{a+b}.
\end{aligned}$$

Let μ_1', μ_2' and μ_3' denotes the first, second and third raw moments about origin for the data, therefore we can replace $E(X)$, $E(X^2)$ and $E(X^3)$ by μ_1', μ_2' and μ_3' respectively in the above equations. Hence, we can get the three equations with unknown parameters λ, a and b as given below

$$\mu_1' = \frac{\lambda a}{a+b}, \tag{3.1}$$

$$\mu_2' = \frac{\lambda^2(a+1)a}{(a+b+1)(a+b)} + \frac{\lambda a}{a+b}, \tag{3.2}$$

$$\mu_3' = \frac{\lambda^3(a+2)(a+1)a}{(a+b+2)(a+b+1)(a+b)} + \frac{3\lambda^2(a+1)a}{(a+b+1)(a+b)} + \frac{\lambda a}{a+b}, \tag{3.3}$$

where λ is the mean number of children ever born to females having at least one child ever born. In order to calculate the mean number of children ever born to females having at least one child ever born (λ), we should have the following information

1. E, the total number of births and,
2. $N - n_0$, i.e., the number of females having at least one child ever born, where N is the total number of females considered for study and n_0 is the total number of childless females in a particular category.

Hence,

$$\lambda = \frac{E}{N - n_0}. \tag{3.4}$$

And by solving the (3.1) to (3.4), we can obtain the values of the unknown parameters a, b and λ.

4. Source of Data

We apply the above proposed model for the number of female child births to females of all parity to the data obtained form the National Family Health Survey (2005-06) for the states known as Seven Sisters of India. The proposed model is also applied for urban and rural background, hindu, muslim and christian religions and wealth index wise in these states. Only those females have been considered in the study who have given birth to at least one child. The frequency of the women have been taken on the basis of 0 female child birth, 1 female child birth, 2 female child births out of all births and so on.

The 2005-06 National Family Health Survey (NFHS-III) is the third in the NFHS series of surveys of India. The first NFHS was conducted in 1992-93, and the second (NFHS-II) was conducted in 1998-99. All three NFHS surveys were conducted under the stewardship of the Ministry of Health and Family Welfare (MOHFW), Government of India. MOHFW designated the International Institute for Population Sciences (IIPS), Mumbai, as the nodal agency for the surveys.

Funding for NFHS-III was provided by the United States Agency for International Development (USAID), the United Kingdom Department for International Development (DFID), the Bill and Melinda Gates Foundation, UNICEF, UNFPA, and the Government of India. Technical assistance for NFHS-III was provided by Macro International, Maryland, USA. Assistance for the HIV component of the survey was provided by the National AIDS Control Organization (NACO) and the National AIDS Research Institute (NARI), Pune.

National family health survey is the nationwide sample survey which consider the following sampling design and techniques of data collection.

Sample Design: The urban and rural samples within each state were drawn separately and, to the extent possible, the sample within each state was allocated proportionally to the size of the state's urban and rural populations. A uniform sample design was adopted in all the states. In each state, the rural sample was selected in two stages: the selection of primary sampling units (PSUs), which are villages, with probability proportional to population size (PPS) at the first stage, followed by the random selection of households within each PSU in the second stage. In urban areas, a three-stage procedure was followed. In the first stage, wards were selected with PPS sampling. In the next stage, one census enumeration block (CEB) was randomly selected from each sample ward. Each ward comprises several enumeration blocks (CEB) created for the census. A list of all the CEBs in a selected ward formed the sampling frame at the second stage. Such lists of CEBs in the selected wards were made available for use for NFHS-III by the census office on request. Each CEB is comprised of about 150-200 households. In the final stage, households were randomly selected within each sample CEB.

Sample Selection: In rural areas, the 2001 census list of villages served as the sampling frame. The list was stratified by a number of variables. The first level of stratification was geographic, with districts being subdivided into contiguous regions. Within each of these region, villages were further stratified using selected variables from the following list: village size, percentage of males working in the non-agricultural sector, percentage of the population belonging to scheduled castes or scheduled tribes, and female literacy. In addition to these variables, HIV prevalence status, i.e., "High", "Medium" and "Low" as estimated for all the districts in high HIV prevalence states, was used for stratification in the high HIV prevalence states. Female literacy was used for implicit stratification (i.e., the villages were ordered prior to selection according to the proportion of females who were literate) in most states although it may be an explicit stratification variable in few states.

Number of households included in the sample by NFHS-III are given in the Table 2 for the seven states of India. NFHS-III interviewed women of age group

15-49 and men of age group 15-54 from selected households to obtain information on population, health, and nutrition in the states. The household response rate (%) for the seven states are also shown in the table as a whole, the individual response rates for eligible women and for eligible men.

Table 2: Household sample selected by NFHS-III for Seven Sisters

States	Number of household included in the sample	Women between 15-49 age group	Men between 15-54 age group	Household response rate (%)		
				Whole	Eligible women	Eligible men
Assam	3437	3840	1394	98	95	86
Arunachal Pradesh	1526	1647	711	98.8	96.9	94.7
Manipur	3498	4512	3915	98.7	94.7	88.4
Meghalaya	1900	2124	720	98	90	78
Mizoram	1513	1791	665	99.7	98.3	97.4
Nagaland	3866	3896	3971	98	95	92
Tripura	1574	1906	711	98	97	92

Fieldwork for NFHS-III was conducted in two phases from November 2005 to August 2006. Eighteen research organizations, including six Population Research Centres, collected the data and conducted data entry and editing operations. The HIV testing of blood samples was done by SRL Ranbaxy, Mumbai. External quality control for the HIV testing of blood samples was done by the National AIDS Research Institute (NARI), Pune. The authorities conducting the NFHS claim that several measures and procedures were used to obtain complete and accurate reporting, but despite this NFHS is still subject to errors and biases that are inherent to all retrospective surveys. The estimates from sample surveys are affected by two type of errors: (1) non-sampling errors and (2) sampling errors. Failure to locate and interview the correct household, misunderstanding of the questions on the part of either interviewer or the respondent and data entry errors are the causes of non sampling errors. Although numerous efforts were made during the implementation of the NFHS-III survey to minimize this type of errors. Non-sampling errors are impossible to avoid and difficult to evaluate statistically. Sampling errors, on the other hand can be evaluated statistically. Non availability of information regarding some variables in NFHS data affected the study.

5. Result and Discussion

The observed and expected frequency curves of females according to the number of female child births of all parities for different domains of the North-East states of India are shown by the Figures 3 to 10 in the Appendix. Solid lines represent the observed frequency curves and dotted lines show the expected frequencies. The different domains of the states of North-East zone of India namely Seven Sisters (Assam, Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland and Tripura) which are considered as region, religion and wealth index categorization of the population. Further region is divided into urban and rural background and hindu, muslim and christian religions of population are considered in the study. Also, three categories of wealth index are considered as: poor, middle and rich. Table 4 provides the estimated values of parameters of the model as derived in Section 3 under different domains of North-East states. The values of λ (i.e., the mean number of children ever born to females having at least one child ever born) of the model vary from 1.754 to 4.466 while the values of a and b vary as 1.271 to 55.451 and 1.019 to 55.132 respectively.

Table 3 shows the values of chi-square test statistic obtained from the data of North-East states and the test of significance values are also mentioned at 0.01 as well as 0.05 levels. It can be observed from the table that calculated values of chi-square is smaller than the corresponding tabulated values at 1% or 5% level of significance for approximately more than 75% of the sub-domains of Seven Sisters (details are provided in Appendix).

Table 3: Calculated values of chi-square for different domains of Seven Sisters

Domain →	Region		Religion			Wealth Index		
State ↓	Urban	Rural	Hindu	Muslim	Christian	Poor	Middle	Rich
Assam	7.230	7.309	44.534*	8.547	3.372	15.296*	1.722	0.597
Arunachal Pradesh	3.2773	11.1431	4.6340	3.7870	9.768	3.419	9.310	4.538
Manipur	10.363	28.416*	56.567*	10.964	10.328	19.349*	9.609	31.487*
Meghalaya	9.822	12.847	5.062	4.066	18.397*	10.570	1.421	12.769*
Mizoram	6.279	20.128*	-	-	22.944*	8.808	5.343	36.645*
Nagaland	7.565	24.577*	5.613	3.927	25.612*	13.024	17.115*	9.787
Tripura	7.860	14.005*	36.883*	10.839*	-	16.843*	18.410*	5.103

* Significant at 0.01.
- Non availability of unit in the domain.

The most notable point from the Table 3 is that the proposed model is good fitted for the urban population (region-wise) throughout the Seven Sisters. And moderately for the muslim religion as well as middle wealth index population of the seven states have positive support with the model. Also the proposed model is applied on the data obtained through NFHS-III according to education attainment in Nagaland state of India (Table 4 and Figure 10). Here, we find that for the education of the mother whether they are illiterate or having higher education, the model is good fitted for Nagaland. The one strongest reason of non-suitability of the proposed model for some of sub-domains in the Seven Sisters are because of the deviation in model assumptions.

From the results obtained, we may reach on the conclusion that the proposed probability model may be considered to be suitable to describe the distribution of the number of female child births to females of all parity not only in the given Seven Sisters but in general, for those regions where model assumptions meet out. The remarkable utility of this model is to provide motivational aspects to measure the female child births and the mean number of children ever born to females through the distribution of females according to parity. This model also gives the new dimension in the field of demography to study the pattern of female child births among the females of all parity in the different domains of the population.

The detailed information in terms of said frequency curves can be visualized through the table and list of figures (i.e., Table 4 and Figures 3 to 10 respectively) provided in the Appendix which show the obtained results for seven states of the country. The model works for different sub-domains of these states for getting the distribution of number of female births to the females of reproductive period yet some study is required for other states of the country socio-economically sub-domains of the population.

Appendix

Table 4: Estimated values of parameters for different domains of Seven Sisters

Arunachal Pradesh						
domains	parameters			chi-square calculated	degree of freedom	chi-square tabulated
	λ	a	b			
Urban	2.988	4.975	5.466	3.2773**	3	7.815
Rural	3.754	3.740	4.283	11.1431*	4	11.668
Hindu	2.449	4.966	4.006	4.634**	3	7.815
Muslim	3.128	9.936	3.255	3.787**	2	5.991
Christian	3.936	3.867	4.348	9.768*	4	13.277
Poor	4.104	8.477	8.831	3.419**	4	9.488
Middle	3.531	5.532	6.828	9.31*	3	11.345
Rich	2.802	3.048	3.573	4.538**	3	7.815

Table 4 (continued): Estimated values of parameters for different domains of Seven Sisters

domains	parameters			chi-square calculated	degree of freedom	chi-square tabulated
	λ	a	b			
Assam						
Urban	3.649	34.460	81.680	7.23**	3	7.815
Rural	2.623	1.966	1.445	7.309**	5	11.070
Hindu	2.661	38.730	41.450	44.534*	4	13.277
Muslim	3.648	3.645	3.809	8.547**	4	9.488
Christian	3.344	2.545	3.047	3.372**	2	5.991
Poor	3.396	3.478	3.624	15.296*	5	15.086
Middle	3.877	4.439	8.193	1.722**	3	7.815
Rich	1.754	5.339	3.912	0.597**	3	7.815
Manipur						
Urban	2.889	2.660	3.110	10.363*	4	13.277
Rural	3.400	3.228	3.506	28.416*	5	15.086
Hindu	2.661	6.382	6.479	56.567*	4	13.277
Muslim	4.248	1.530	1.606	10.964*	4	13.277
Christian	3.624	7.526	8.042	10.328*	4	13.277
Poor	3.754	3.304	3.429	19.349*	4	13.277
Middle	3.506	3.865	4.298	9.6096*	4	13.277
Rich	2.778	2.644	3.033	31.487*	4	13.277
Meghalaya						
Urban	3.004	40.414	43.894	9.822*	3	11.345
Rural	3.868	3.978	4.104	12.847*	5	15.086
Hindu	2.242	55.451	43.129	5.062**	2	5.991
Muslim	3.568	6.772	6.331	4.066**	2	5.991
Christian	3.720	3.661	3.822	18.397*	5	15.086
Poor	3.981	6.604	6.538	10.5703*	4	13.277
Middle	4.020	5.676	6.398	1.4206**	4	9.488
Rich	2.991	14.737	16.294	12.769*	3	11.345
Mizoram						
Urban	2.278	2.901	2.410	6.279**	3	7.815
Rural	3.267	1.925	1.901	20.128*	5	15.086
Christian	3.052	28.725	30.471	22.944*	4	13.277
Poor	3.561	5.674	5.673	8.808*	3	11.345
Middle	3.343	6.848	8.638	5.343**	2	5.991
Rich	2.458	1.271	1.019	36.645*	5	15.086

Table 4 (continued): Estimated values of parameters for different domains of Seven Sisters

domains	parameters			chi-square calculated	degree of freedom	chi-square tabulated
	λ	a	b			
Nagaland						
Urban	3.091	44.076	46.660	7.565**	4	9.488
Rural	4.466	3.536	4.271	24.577*	6	16.812
Hindu	2.698	5.259	5.708	5.613**	3	7.815
Muslim	3.186	12.472	13.282	3.927**	2	5.991
Christian	3.707	4.387	4.597	25.612*	5	15.086
Poor	4.173	5.251	5.807	13.024*	5	15.086
Middle	3.861	5.615	6.294	17.115*	4	13.277
Rich	3.919	44.667	55.132	9.787*	4	13.277
Illiterate	4.258	6.6406	6.907	4.806**	5	11.070
Primary education	3.928	7.802	7.621	18.2897*	5	15.086
Secondary education	3.1515	4.384	4.382	22.1027*	5	15.086
Higher education	1.324	6.87361	0.537	3.5121**	1	3.841
Tripura						
Urban	2.120	27.061	29.183	7.86*	3	11.345
Rural	2.916	4.002	4.561	14.005*	3	11.345
Hindu	2.603	8.381	8.932	36.883*	3	11.345
Muslim	3.535	19.704	22.707	10.839*	2	9.210
Poor	3.266	10.249	11.307	16.843*	3	11.345
Middle	2.801	3.297	3.442	18.410*	4	13.277
Rich	1.909	26.998	34.359	5.103**	2	5.991

Note: * indicates that values are compared at 0.01 level of significance, ** indicates that values are compared at 0.05 level of significance.

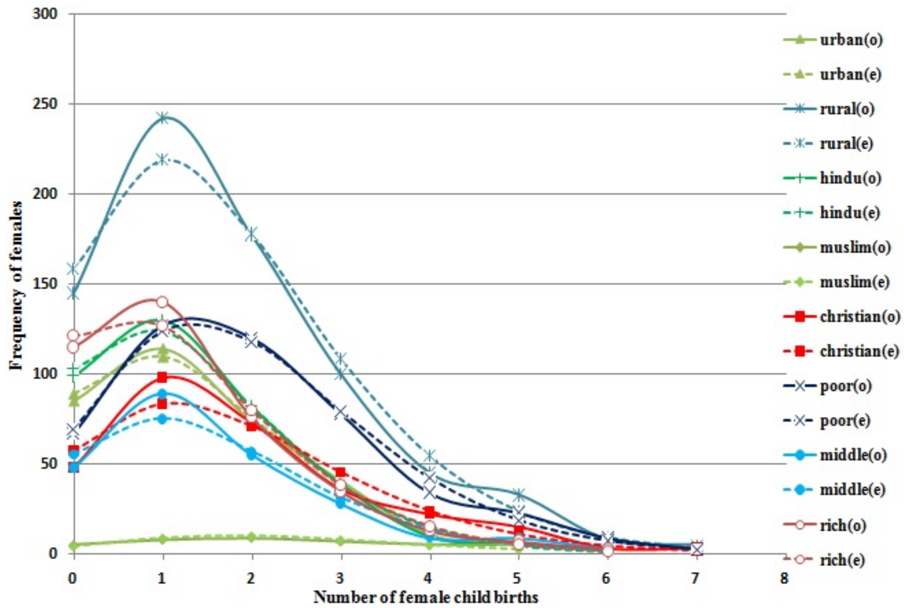

Figure 3: Observed and expected frequency curves for different domains of Arunachal Pradesh

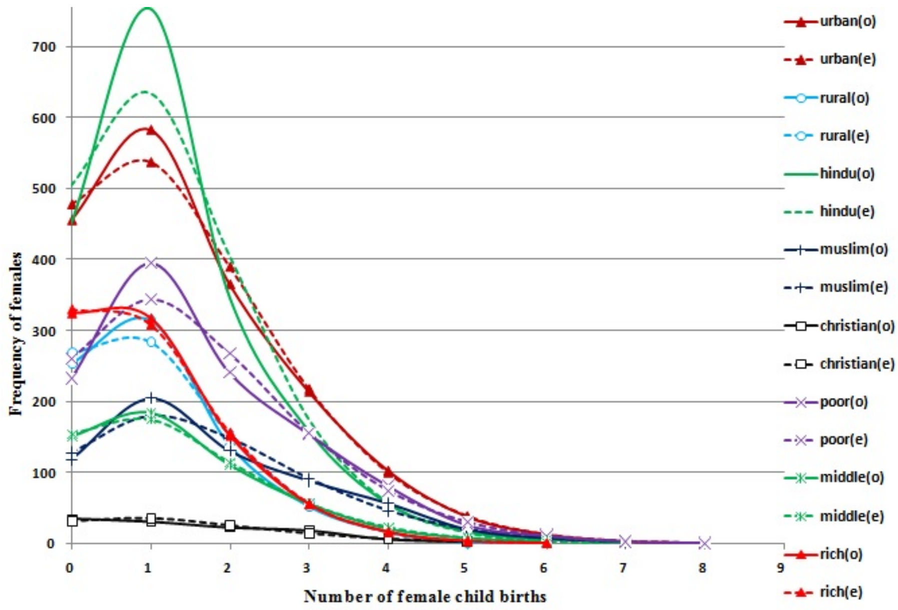

Figure 4: Observed and expected frequency curves for different domains of Assam

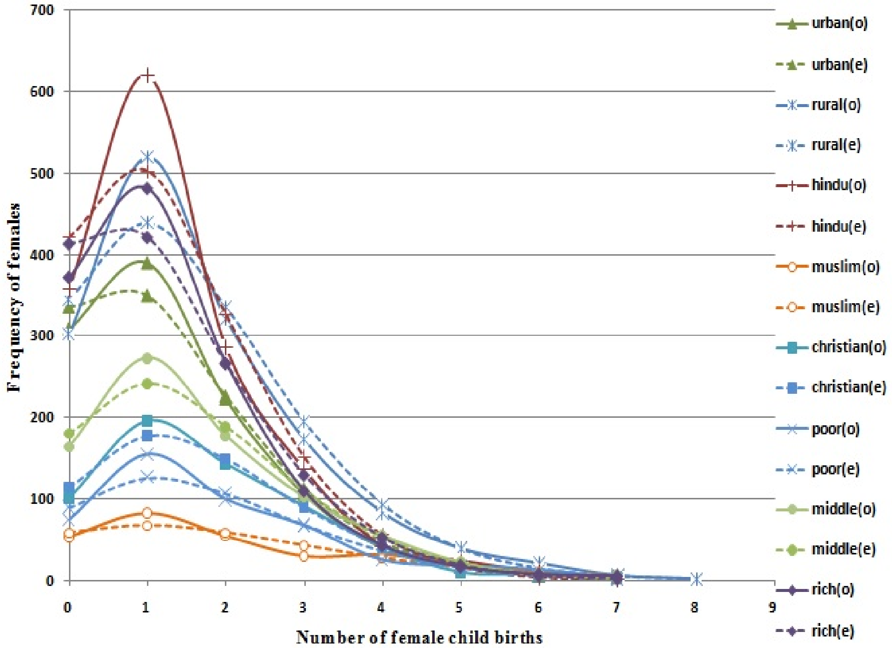

Figure 5: Observed and expected frequency curves for different domains of Manipur

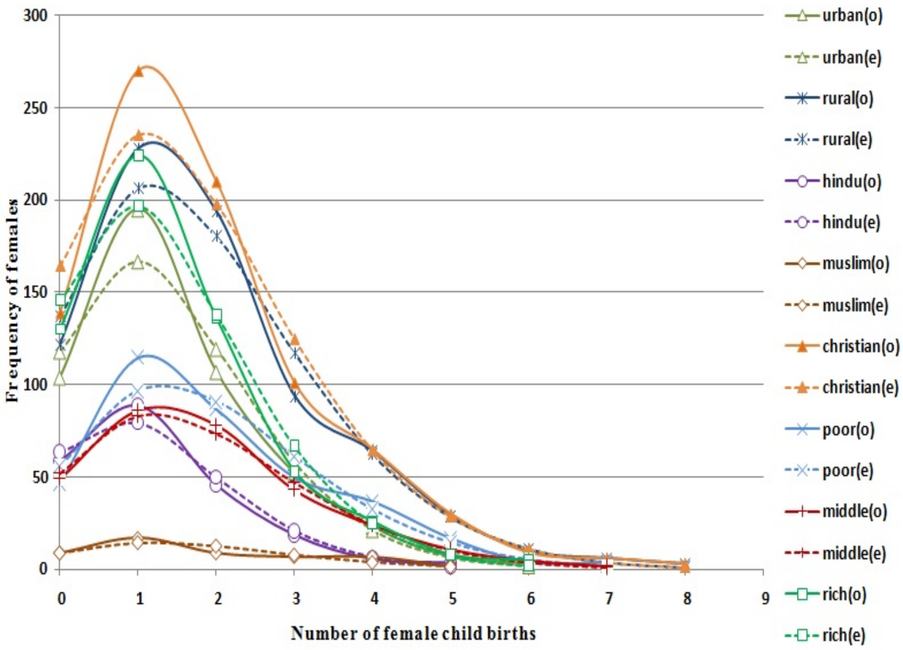

Figure 6: Observed and expected frequency curves for different domains of Meghalaya

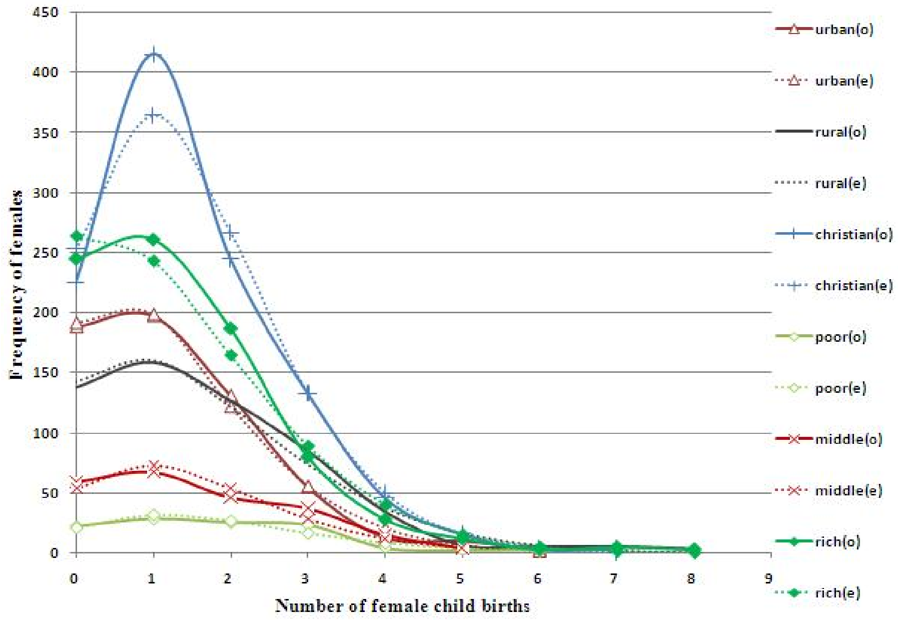

Figure 7: Observed and expected frequency curves for different domains of Mizoram

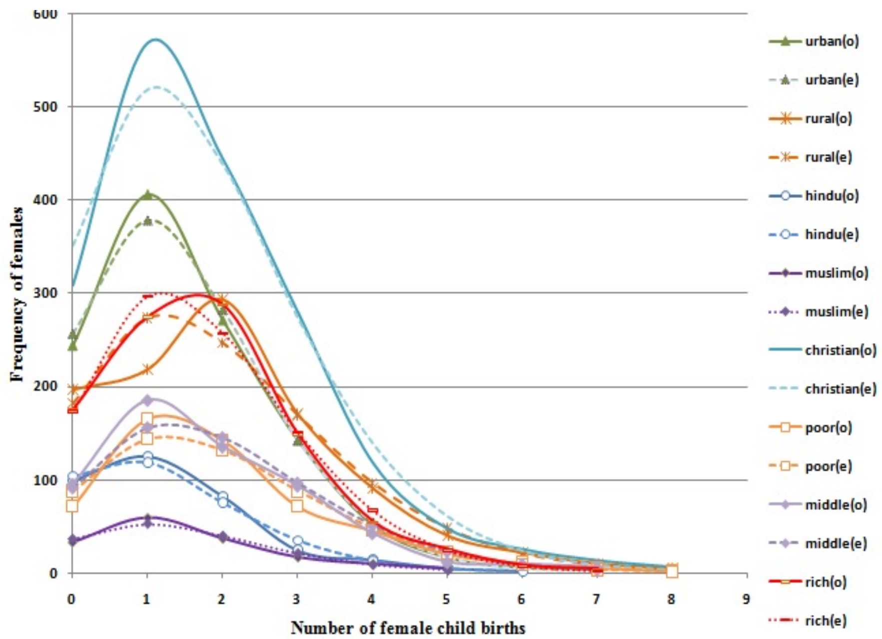

Figure 8: Observed and expected frequency curves for different domains of Nagaland

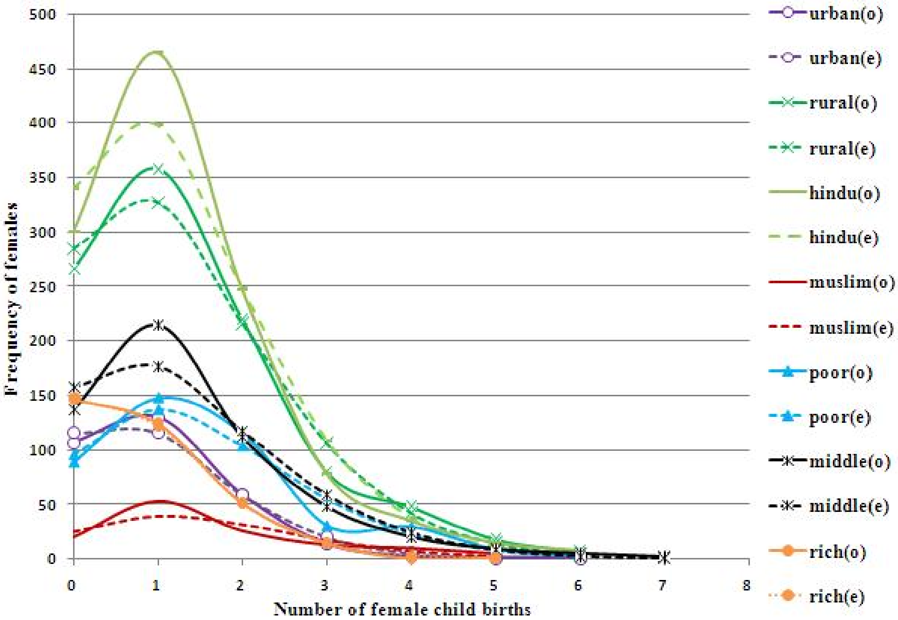

Figure 9: Observed and expected frequency curves for different domains of Tripura

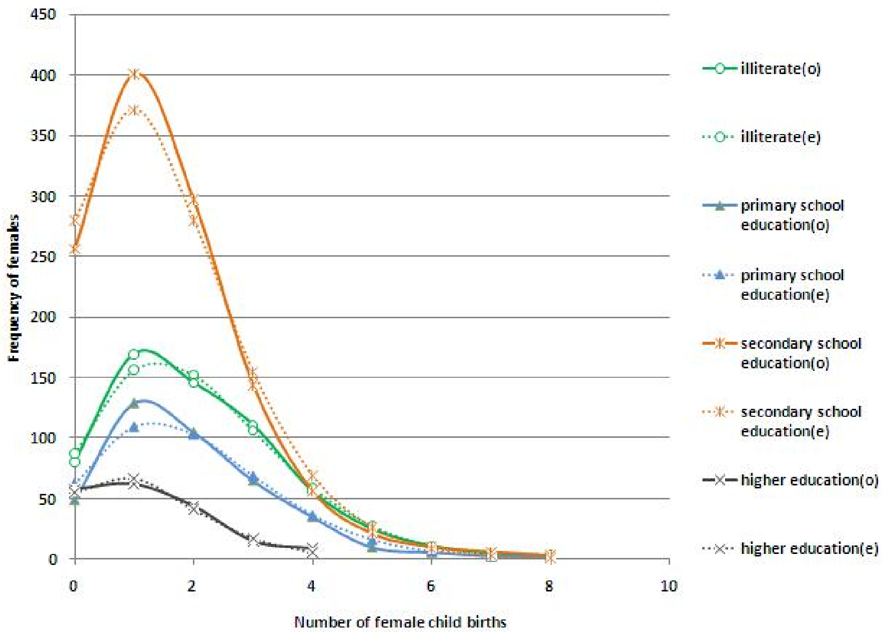

Figure 10: Observed and expected frequency curves for education attainment in Nagaland

References

Alauddin Chowdhury, A. K. M., Khan, A. R. and Chen, L. (1976). The effect of child mortality experience on subsequent fertility: in Pakistan and Bangladesh. *Population Studies* **30**, 249-261.

Bhuyan, K. C. (1996). Fertility differential according to socio-economic status and family planning adoption in rural Bangladesh. *Shankhya, Series B* **58**, 302-322.

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

Dandekar, V. M. (1955). Certain modified forms of binomial and poisson distributions. *Shankhya, Series B* **12** 237-250.

Feller, W. (1948). On probability problems in theory of counters. In *Studies and Essays Presents to R. Courant on His 60th Birthday* (Edited by K. O. Friedrichs, O. E. Neugebauer and J. J. Stoker), 105-115. Interscience, New York.

Henry, L. (1956). *Aneienness Familles Genevoises*: *É'tude Démographique, XVIe-XXe Siècle*, 227-229. Presses Universitaires de France, Paris.

Neyman, J. (1949). *On the Problem of Estimating the Number of Schools of Fish.* Publications in Statistics, Volume 1, 21-36. University of California press, Berkelry.

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

Pathak, K. B. (1966). A probability distribution for the number of conceptions. *Shankhya, Series B* **28**, 213-218.

Rai, P. K., Pareek, S. and Joshi, H. (2013). Regression analysis of collinear data using r-k class estimator: socio-economic and demographic factors affecting the Total Fertility Rate (TFR) in India. *Journal of Data Science* **11**, 321-340.

Rutstein, S. and Medica, V. (1975). The effects of infant and child mortality on fertility in Latin America. In *Committee for International Coordination of National Research in Demography* (*CICRED*), 225-246. Seminar on Infant Mortality in Relation to the Level of Fertility. Paris, CICRED, Bangkok.

Sheps, M. C. and Perrin, E. B. (1966). Further results from a human fertility model with a variety of pregnancy outcomes. *Human Biology* **38**, 180-193.

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**, 54-67.

Singh, S. N. (1961). *A Hypothetical Chance Mechanism of Variation in Number of Births Per Couple.* Ph.D. Thesis, University of California, Berkeley.

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

Singh, S. N. (1964). A probability model for couple fertility. *Shankhya, Series B* **26**, 89-94.

Singh, S. N. (1966). Some probability distributions utilized in human fertility, 74-84. Seminar Volume in Statistics. Banaras Hindu University, Varanasi.

Singh, S. N. (1968). A chance mechanism of the variation in the number of births per couple. *Journal of American Statistical Association* **63**, 209-213.

Singh, S. N. and Pathak, K. B. (1968). On the distribution of the number of conceptions. *Journal of Mathematical Society* **1**, 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**, 197-205.

Singh, S. N. and Bhattacharya, B. N. (1970). A generalized probability distribution for couple fertility. *Biometrics* **26**, 33-40.

Singh, S. N., Bhattacharya, B. N. and Joshi, P. D. (1973). On a hypothetical model for the number of births and its application. *Journal of the Indian Society of Agricultural Statistics* **25**, 91-100.

Singh, S. N., Bhattacharya, B. N. and Yadava, R. C. (1974). A parity dependent model for number of births and its applications. *Shankhya, Series B* **36**, 93-102.

Received May 22, 2013; accepted July 12, 2013.

Piyush Kant Rai
Department of Mathematics and Statistics
Banasthali University
Banasthali Vidyapith, P.O. Banasthali Vidyapith, Rajasthan - 304022, India
raipiyush5@gmail.com

Sarla Pareek
Department of Mathematics and Statistics
Banasthali University
Banasthali Vidyapith, P.O. Banasthali Vidyapith, Rajasthan - 304022, India
psarla13@gmail.com

Hemlata Joshi
Department of Mathematics and Statistics
Banasthali University
Banasthali Vidyapith, P.O. Banasthali Vidyapith, Rajasthan - 304022, India
kirtidbest@gmail.com