Demography India (2015) ISSN: 0970-454X
Vol.44, Issue: 1&2, pp: 50-57

Research Article

A Probability Model for Sex Composition of Children in the Presence of Son Preference

Brijesh P. Singh[1], Sonam Maheshwari[2,*] & Puneet K. Gupta[3]

Abstract

In this study an attempt has been made to understand the mechanism of son preference through the modeling of the pattern of male children in Uttar Pradesh, where family size and sex composition are dominated by strong son reference. The basic binomial model is modified assuming the concept of clustering. The model has been applied to data from the state of Uttar Pradesh which is known to have strong son preference.

Introduction

The strong son preference in the society affects the sex ratio at birth through the sex selective abortion. According to census of India, sex ratio in India is defined as the number of females per 1000 males. Sex ratio in Uttar Pradesh is 912 i.e. there are 912 females for 1000 males, which is below national average of 940 as per census 2011. In 2001, the sex ratio of female was 898 per 1000 males in Uttar Pradesh. A preference for sons in society is the result of traditional religious beliefs, social customs (dowry system, lineage, familial and kinship ties etc.), economic benefits and including support of ageing parents is widespread, particularly in traditional society like Uttar Pradesh. Researchers in India and elsewhere have also noted that fertility declines over the past decade or so have intensified pressure on women to act to achieve their desired family sex composition (Das Gupta & Bhat, 1997; Gandotra, 2008; Das Gupta et al., 2009; Guilmoto, 2009). Demographers and other social scientist have awful interest in the effect of normative expectations and preferences on ‘ideal family size’, both in terms of size and composition.

Many demographers revealed that male child preferences influence reproductive outcomes extensively. They play an important role in settings, where notions of the ‘ideal family size’ are marked by a strong and persistent preference for sons (Pande, 2003; Retherford & Roy, 2003; Pande & Astone, 2007). Retherford & Roy (2003) find that sex ratios at birth differ substantially by birth order and the sex composition of children already born, particularly at higher parities. The poor availability of reproductive services is compounded by patriarchal gender and social norms that continue to restrict women’s reproductive options and in many cases dominates on women’s own reproductive preferences (Das Gupta, 1999; Barua & Kurz, 2001; Barua et al., 2004; Sheth, 2006). These norms also limit women’s ability especially when young and newly married to access the reproductive services (Jejeebhoy, 1998; Mathur et al., 2003; Jejeebhoy & Halli, 2005; MacQuarrie & Edmeades, 2011).

Many developing countries in East, South, South-east Asia and North Africa are characterized by a strong son preference i.e. a strong preference for male as opposed to female offspring (Arnold et al., 1998; Clark 2000; Jensen 2002). This strong preference is reflected in son targeting fertility behavior, also referred to in the literature as differential stopping behavior (DSB) or male-preferring stopping rules (Clark 2000). In our study, we think DSB is prevalent in majority in our society means women continue childbearing until they reach their desired number of male children as well as their

[1]Asstt. Professor (Statistics), Faculty of Commerce, Banaras Hindu University, Varanasi
[2]Lecturer (Statistics), Department of Community Medicine, SRMS-IMS, Bareilly
[3]Ph. D. Fellow, Department of Statistics, Allahabad University, Allahabad
*Corresponding author: Sonam Maheshwari, Email: maheshwarisonam2@gmail.com

ideal number of children. In Uttar Pradesh, there is a strong prevalence of two male children norm or in other words mostly women continue child bearing until they have two male children. They compare male children with the eyes, since two eyes are necessary for a good life similarly two sons are necessary. There is a popular slogan in Hindi that "*ek aankh ko aankh nahi kahte*", (one eye is not called 'eye', implying that one eye is adequate for vision)

According to National Family Health Survey (NFHS)-III conducted in 2005-06, 33.5 percent women in Uttar Pradesh, India's largest state, reported they want more sons than daughters while only 1.7 percent reported that they want more daughters. Mean ideal number of total children, sons and daughters in Uttar Pradesh is 2.5, 1.3 and 0.9 respectively, indicating a substantially high in the preference for smaller family size. A strong preference for children of one sex can be a constraint on fertility decline if couples who have achieved their preferred family size continue childbearing until they achieve their desired number of sons or daughters. This NFHS report proposes that, in Uttar Pradesh, measures of ideal family size mask a preference for sons that makes some people willing to have families larger than their stated ideal number of size. Although far from new to the demographic community, this hypothesis has renewed relevance in research settings that exhibit a strong preference for one sex, and where fertility increasingly is coming under voluntary control (Bairagi & Langsten, 1986; Chowdury & Bairagi, 1990). In the situations where effective means of family planning are available, couples can decide if and when they want to limit or space their fertility, in particular, whether such limiting should occur after the birth of a certain number of children, or a certain number of sons or daughters.

Dandekar (1955) has given two modified forms of Binomial and Poisson distributions under the specified conditions affecting the independence of successive trials which are useful for describing the variations in the number of births to a woman during a given period. Henry (1956) derived expressions for the expected number of births in a period of five years assuming that a woman had a 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. Basu (1955) reformulates Dandekar's model as a stochastic process and Dharamdhikari (1964) provides a generalized stochastic model with Dandekar's model being a special case. Dharmadhikari points out that Dandekar's implicitly assumes that the successive waiting times between live births, except for the first birth, are independent and identically distributed random variables which have an exponential distribution with the origin shifted by the (constant) length of the non susceptible period. This is then generalized by considering an arbitrary distribution of waiting times.

Singh (1963) derives a modified binomial distribution using an analogy between Neyman's (1949a) fishing model and the reproductive process. Singh adds one assumption to the set of Dandekar's that is some woman may not be exposed to risk of conception at any time during the observation period. Singh (1963) allows for heterogeneity in the model by assuming that the monthly probability of conception varies according to a Beta distribution among women. Parameter estimates are obtained by the minimum chi-square procedure of Neyman (1949b) and the model provides a good fit to the given data. Pathak (1966) extends Singh's model by allowing for an abrupt sequence of trials. Singh and Bhattacharya (1970) generalize Singh's models by allowing more than one kind of pregnancy outcome.

It is not so easy to find that the probability of birth is male (or female) with the application of probability modeling incorporating a parameter for sex preference. In this study an attempt of statistical modeling has been made to work out for the pattern of male children in Uttar Pradesh, where family size and sex composition are dominated by strong son reference. The first model is proposed under the assumptions that probability of male birth remains constant across the population of women and also across the successive births for the same women and there is no sex selective stopping of childbearing (Gokhale & Kunte, 1997). If the probability is constant among women, then the distribution of male births follows the Binomial distribution. In India mostly families want at least one daughter or two sons (International Institute for Population Sciences & ORC Macro, 2000; Pande & Astone, 2007). Taking into consideration that there is a clustering at two sons which arises due to son preference we propose an inflated Binomial distribution. Here distribution is inflated at two. This study enables us to give idea about the risk in contexts of strong sex preferences.

Construction of Model

Consider the women who delivered n children, i.e. women with fixed parity n. Let X be the number of male children out of n children; this is a random variable and follows Binomial distribution with parameter p which is the probability of male child birth to a woman. This model is proposed under the assumption that probability of male child born to a woman is constant and that sex outcomes of children at successive birth are independent. Let us define for i^{th} (i=1, 2,…. n) child born to a woman, a random variable z_i taking value 1 if the child is a male and 0 otherwise. Thus z_i's are Bernoulli variable. Now if we assume that birth of male children (born to a woman) are independent of each other having same probability p then total number of male child X to the woman is nothing but sum of independent Bernoulli variables and hence follows a binomial distribution. Therefore the distribution of X is given by:

$$P[X=x]=\binom{n}{x}p^x(1-p)^{n-x} \qquad ; \quad 0\le p\le 1,\ x=0,1,2\ldots\ldots\ldots\ldots\ldots..n \tag{1}$$

Inflated Binomial Model

In Uttar Pradesh, sex composition is dominated by son preference. Here two male children norm is popular in the society and also reflected in the distribution of the male children born. Generally women want to have at least two male children in their whole reproductive period and after that they stop childbearing if the total desired size is also attained. So we get extra frequency at two children, and hence it is reasonable to use binomial distribution inflated at two. In this model, the probability mass function is given as

$$P[X=x]=\begin{cases}(1-\alpha)+\alpha\binom{n}{2}p^2(1-p)^{n-2}\ ; & for\ x=2\\ \alpha\binom{n}{x}p^x(1-p)^{n-x}\ ; & for\ x=0,1,3,4......\end{cases} \tag{2}$$

where (1-α) is the proportion of the extra women who produce two children and p is risk of male birth.

Estimation of Parameters in the Models

In this section estimation procedure has been discussed for the models under consideration for the number of male children for fixed parity. For this purpose we consider the method of moment.

Method of Moment (MM) Estimates of Model-I

For Model-I, we know that the moment estimate of risk of birth of male children (p) is give as

$$\hat{p}=\frac{\overline{x}}{n} \tag{3}$$

where $\overline{x}$ is the mean number of male births for the woman with parity n.

Method of Moment (MM) Estimates of Model-II

The moment estimates of the parameters α and p of model-II can be obtained as follows

$$E(X) = \mu_1' = 2(1-\alpha) + \alpha np \tag{4}$$

$$E(X^2) = \mu_2' = 4(1-\alpha) + \alpha np(np+1-p) \tag{5}$$

From equation 4 we can estimate the value of α

$$\mu_1' = 2 - \alpha[2-np]$$

$$\hat{\alpha} = \frac{2-\mu_1'}{2-np} \tag{6}$$

Simplifying equation 5 to get the estimate of *p* we follow these steps

$$\mu_2' = 2(1-\alpha) + 2(1-\alpha) + \alpha np(np) + \alpha np - \alpha np^2$$

$$= 2(1-\alpha) + \alpha np(np-p) + \mu_1' \ .$$

Using equation 4 we have

$$= 2\mu_1' - \alpha np + \alpha np(np-p) = 2\mu_1' + \alpha np[np-p-1]$$

putting the value of α obtained in equation 6 the equation become

$$\mu_2' = 2\mu_1' + \left[\frac{2-\mu_1'}{2-np}\right] np[np-p-1] \tag{7}$$

$$\frac{\mu_2' - 2\mu_1'}{\mu_1' - 2} = \left[\frac{np[np-p-2+1]}{np-2}\right] = A$$

Then we get

$$A = np - \frac{np(p-1)}{np-2}$$

$$n^2p^2 - np^2 - p(nA+n) + 2A = 0$$

$$p^2(n^2-n) - p(nA+n) + 2A = 0 \tag{8}$$

This is a quadratic equation, after solving we can get the estimated value of *p*. Since *p* is the probability thus the value lies between 0 and 1 has been used for estimation of α .

Mix method of estimating the parameters

In this method we use second cell frequency and first moment. We know that

$$p(2) = 1 - \alpha + \alpha\, {}^{n}_{2}c\, p^2 (1-p)^{n-2} = \frac{n_2}{N}$$

$$= 1 - \alpha\left[1 - \frac{n(n-1)}{2} p^2 (1-p)^{n-2}\right] = \frac{n_2}{N}$$

$$\hat{\alpha} = \frac{2(N-n_2)}{2N - Nn(n-1)p^2(1-p)^{n-2}} \tag{9}$$

Here *N* is the total number of women for fixed parity *n* and n_2 is number of women with exactly two male children. After putting the value of α from equation 9 in equation 4, we can get a function of *p* but this function seems to be complicated in order to get the estimate of p. For getting the estimates of p and α, we use both the methods, i.e., method of moment estimate and mixed method.

Result and Discussion

Data for this study have been taken from NFHS-III (2005-06) Uttar Pradesh. In this Study, information based on those women who had ever been married in age group 15-49 on survey date was used. The survey collected information from ever married women on fertility, mortality, family planning and important aspects of reproductive health etc. In order to analyze the son preference, we have extracted 4233 women out of 12183 women who had completed their family size or those open birth interval is greater than five years i.e. they completed their fertility.

Since, in this study two models are proposed for fixed parity so that after obtaining the estimate of parameters for different parities, we obtained the estimated frequencies for both the models. Tables 1 to 5 show the expected frequencies along with the observed frequencies for parity 3 to 7 of Uttar Pradesh. Estimate of parameters, the value of χ^2 with degree of freedom are also given in the respective tables. The value of χ^2 shown in the tables clearly indicate that model-I is not suitable model for distribution of male birth to a woman for fixed parity for lower parities but model-II describe the data excellently, means our assumption that there are some women with strong belief of two sons are present in the society. For higher parity such as parity 7, Model-I performs well than Model-II, which indicate there is no sex preference for high fertile females. Mix method of estimation provided better fitting for all parity data because in this method for estimation of the parameter we equate second cell probability with the theoretical probability and use first moment. Thus over all we can say the model II is better model to understand the mechanism of delivering male child when son preference is widespread and possibly differential stopping behavior prevalent.

It is also observed from the tables that as parity of woman increases the risk of male birth i.e. the value of parameter p of model-II decreases, which suggest that risk of male birth decreases at higher parity and the value of α increases as parity of woman increases it resembles that value of (1-α) decreases which present the percentage of women who contribute more number of two male children decreases. After fitting of the model-I for the real data on number of male birth perhaps it is clear that the assumption that p (the risk of male birth) be same for all woman is not proper for lower parities due to the fact that two male children norm as discussed earlier. The statistical measure such as mean and variance of sample and mean and variance obtained through the proposed models are shown in 6. It may be noted that mean obtained through the use of model-I and Model- II matches with the sample mean but slight difference is observed between sample variance and variance obtained from Model-I. The variance from Model II is close to the sample variance. Table 7 represents the values of parameters obtained through method of moments for the models at a glance. The inflated binomial model at two is an appropriate model for explaining the distribution of male children to the women for fixed parity.

Table 1: Expected & Observed distributions of male birth to the women in the state of Uttar Pradesh with Parity 3

Number of male births	Observed number of women	Expected number of women		
		Model -I	Model-II	
			MM method	Mix method
0	54	69.74	64.73	65.56
1	245	273.77	212.92	227.73
2	463	358.24	495.03	463.11
3	96	156.25	85.32	101.60
Total	**858**	**858.00**	**858.00**	**858.00**
Parameters		p=0.5668	α= 0.6951 p=0.5230	α= 0.7678 p=0.5366
χ^2		60.45	18.66	3.67
df		2	1	1

Table 2: Expected & Observed distributions of male birth to the women in the state of Uttar Pradesh with Parity 4

Number of male births	Observed number of women	Expected number of women		
		Model -I	Model-II	
			MM method	Mix method
0	23	93.36	31.12	36.48
1	167	259.80	132.65	154.20
2	361	270.08	415.44	361.61
3	189	124.79	150.65	172.21
4	30	21.97	40.14	45.50
Total	**770**	**770.00**	**770.00**	**770.00**
Parameters		p=0.4094	α= 0.7359 p=0.5159	α= 0.8478 p=0.5138
χ^2		152.76	30.47	12.96
df		3	2	2

Table 3: Expected & Observed distributions of male birth to the women in the state of Uttar Pradesh with Parity 5

Number of male births	Observed number of women	Expected number of women		
		Model -I	Model-II	
			MM method	Mix method
0	13	26.59	16.13	20.32
1	96	114.68	75.85	91.65
2	255	197.83	294.82	255.38
3	164	170.63	134.21	149.21
4	58	73.58	63.12	67.31
5	10	12.69	11.87	12.15
Total	**596**	**596.00**	**596.00**	**596.00**
Parameters		p=0.4631	α= 0.7448 p=0.4847	α= 0.8490 p=0.4743
χ^2		30.64	10.02	5.97
df		4	3	3

Table 4: Expected & Observed distributions of male birth to the women in the state of Uttar Pradesh with Parity 6

Number of male births	Observed number of women	Expected number of women		
		Model -I	Model-II	
			MM method	Mix method
0	6	15.44	8.94	12.91
1	67	70.64	45.75	61.21
2	159	134.68	198.58	158.72
3	134	136.96	111.01	127.51
4	71	78.34	71.04	75.60
5+	26	26.94	27.68	27.05
Total	**463**	**463.00**	**463.00**	**463.00**
Parameters		p=0.4327	α= 0.7819 p=0.4604	α= 0.9185 p=0.4415
χ^2		11.13	23.59	4.89
df		4	3	3

Table 5: Expected & Observed distributions of male birth to the women in the state of Uttar Pradesh with Parity 7

Number of male births	Observed number of women	Expected number of women		
		Model -I	Model-II	
			MM method	Mix method
0+1	39	42.15	26.32	37.06
2	84	73.20	107.95	84.14
3	74	85.10	69.14	80.30
4	70	59.36	54.37	57.86
5	17	24.84	25.65	25.02
6+	7	6.35	7.57	6.63
Total	**291**	**291.00**	**291.00**	**291.00**
Parameters		p=0.4109	α= 0.8104 p=0.4402	α= 0.9406 p=0.4188
χ^2		7.73	19.62	5.73
df		4	3	3

Table 6: Comparison of Sample Measures with Measures Obtained Through the Models in Uttar Pradesh

Parity	Sample Mean	Sample Variance	Model I		Model II	
			Mean	Variance	Mean	Variance
3	1.7004	0.5595	1.7004	0.7366	1.7004	0.5998
4	2.0468	0.7355	1.6374	0.9671	2.0468	0.8467
5	2.3154	0.9643	2.3154	1.2432	2.3154	1.0774
6	2.5961	1.2645	2.5961	1.4728	2.5961	1.3896
7	2.8763	1.5242	2.8763	1.6944	2.8763	1.5775

*calculation is based on moment estimates

Table 7: Estimated Values of Parameters Obtained Through the Models in Uttar Pradesh

Parity	N Total Male Children	n Frequency of two male children	n/N Proportion of two male children	Model I	Model II	
				p Risk for male birth	p Risk for male birth	1-α Proportion of extra woman having two male children
3	858	463	0.5396	0.5668	0.5230	0.3049
4	770	361	0.4688	0.4094	0.5159	0.2642
5	596	255	0.4279	0.4631	0.4847	0.2553
6	463	159	0.3434	0.4327	0.4604	0.2181
7	291	84	0.2887	0.4109	0.4402	0.1896

*calculation is based on moment estimates

References

Arnold, F., Choe, M. K., & Roy, T. K. (1998) ‘Son preference, the family-building process and child mortality in India’, *Population Studies*, 523, pp. 301–315.

Bairagi, R. & Langsten R. L. (1986) ‘Sex preference for children and its implications for fertility in rural Bangladesh’, *Studies in Family Planning*, 17(6), pp. 302-307.

Barua, A. & Kurz, K. (2001) ‘Reproductive health-seeking by married adolescent girls in Maharashtra, India’, *Reproductive Health Matters*, 9, pp. 53–62.

Barua, A., Pande, R. & Macquarrie, K. (2004) ‘Caring men? Husbands’ involvement in the maternal care of young wives’, *Economic and Political Weekly*, 39, pp. 5661-5668.

Basu, D. (1955) ‘A note on the structure of a stochastic model considered by V.M. Dandekar’, *Sankhya A*, 15, pp. 251-252.

Chowdury, M. K. & Radheshyam, B. (1990) 'Son preference and fertility in Bangladesh', *Population and Development Review,* 16(4), pp. 749-757.

Clark, S. (2000) 'Son Preference and Sex Composition of Children: Evidence From India', *Demography,* 37, pp. 95-108.

Dandekar, V. M. (1955) 'Certain modified forms of binomial and Poisson distributions', *Sankhya*, 12, pp. 237-250.

Das Gupta, M. (1999) 'Lifeboat versus corporate ethic: social and demographic implications of stem and joint families', *Social Science & Medicine,* 49, pp. 173-184.

Das Gupta, M. & Bhat, P. N. M. (1997) 'Fertility decline and increased manifestation of sex bias in India', *Population Studies,* 51, pp. 307-315.

Das Gupta, M., Chung, W. & Shuzuo, L. (2009) 'Evidence for an incipient decline in numbers of missing girls in China and India', *Population and Development Review,* 35, pp. 401-416.

Dharmadhikari, S. W. (1964) 'A generalization of a stochastic model considered by V.M. Dandekar', *Sankhya A,* 26, pp. 31-38.

Gandotra, B. (2008) 'Maintaining access to safe abortion and reducing sex ratio imbalances in Asia', *Reproductive Health Matters,* 16, pp.90–98.

Gokhale, D. V. & Kunte, S. (1997) 'Probability of a male child: Is it constant over the population of couples?', *Demography India,* 26(1), pp. 1-7.

Guilmoto, C. Z. (2009) 'The sex ratio transition in Asia', *Population and Development Review,* 35, pp. 519-549.

Jejeebhoy, S. (1998) 'Adolescent sexual and reproductive behavior: a review of the evidence from India', *Social Science & Medicine,* 46, pp. 1275-1290.

Jejeebhoy, S. & Halli, S. (2005) 'Marriage patterns in rural India: influence of socio cultural context', *In Lloyd, C. B. (ed.)* Growing up Global: The Changing Transitions to Adulthood in Developing Countries: Selected Studies. National Academy Press Washington, DC, pp. 172-199.

Jensen, R. T. (2002) 'Equal Treatment, Unequal Outcomes? Generating Sex Inequality Through Fertility Behavior', *Unpublished document,* JFK School of Government, Harvard University.

MacQuarrie, K. & Edmeades, J. (2011) 'The Path Less Taken: Contraception and Abortion over the Life Course in Madhya Pradesh, India', *Population Association of America,* Washington, DC.

Mathur, S., Greene, M. & Malhotra, A. (2003) 'Too Young to Wed: The Lives, Rights, and Health of Young Married Girls', *ICRW,* Washington, DC.

Neyman, J. (1949a) 'Contributions to the theory of the chi-square test', *Proceedings of the First Berkely Symposium on Mathematical Statistics and Probability,* pp. 239-273.

Neyman, J. (1949b). 'On the problem of estimating the number of schools of fish', *University of California Publications in Statistics,* 1, pp. 21-36

Pande, R. (2003) 'Selective gender differences in childhood nutrition and immunization in rural India: the role of siblings', *Demography,* 40, pp. 395-418.

Pande, R. & Astone, N. (2007) 'Explaining son preference in rural India: the independent role of structural versus individual factors', *Population Research and Policy Review,* 26, pp. 1-29.

Pathak, K. B. (1966) 'A probability distribution for a number of conceptions', *Sankhya B,* 28, pp. 213-218.

Retherford, R. and Roy, T. K. (2003) 'Factors affecting sex-selective abortion in India and 17 major states', *International Institute for Population Sciences,* Mumbai, India.

Sheth, S. S. (2006) 'Missing female births in India', *Lancet,* 367, pp. 185-186.

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

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