Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/090206

Rural out Migration at the Household Level

Niraj Kumar Singh and Mritunjay Pal Singh*

Amity Institute of Applied Sciences, Amity University, Noida, India

Received: 22 May 2019, Revised: 2 Aug. 2019, Accepted: 8 Aug. 2019
Published online: 1 Jul. 2020

Abstract: Migration is the least explored area of population change. The concept of migration involves a series of factors in terms of place of origin and destination, intervening obstacles and personal characteristics. The present paper investigates migration at household level using some statistical models. Which have been tested with the help of primary data collected by the authors in North-Eastern Bihar during the period of (2009-10).

Keywords: Migration, North-Eastern Bihar, Geometric Distribution, Log series Distribution, Inflation

1 Introduction

Migration, which is an important part of demography, is the least investigated area as compared to fertility and mortality. Based on reduced birth and death rate, migration (internal or international) has become a more important concern for demographers and other social scientists. Most of the studies in past (Bhagat 2005; Friedlander and Roshier 1966; Greenwood 1971; Isbell 1944; Lee 1966; Singh 1986; Stouffer 1940,1960 and Ziff 1946), based on the conceptualization of migration process and the scale of investigations, used macro approach by operating on highly aggregate data for countries, districts, states and the nation as a whole. These types of studies are unable to provide sufficient explanation for the tremendous regional and local heterogeneity. They also ignore the decision making process of migrating individuals (Singh & Yadava 1981). Thus a study of environment in which migration takes place and the migration decision process at the micro-level is more important. Studies of migration at the micro-level i.e. at the levels of individuals; families or households have important implications for housing policies as well as the development of other sociological models related to families and communities (Pryor, 1975; Rossi, 1955).

A migrant household (with one or more persons involved in the process of migration to do some job outside the village) may have different socio-economic and cultural characteristics, ideas, awareness and environments compared to a non-migrant household. Rural-urban migration in India is of chain type at least in the beginning of the migration process before the migrants settle in the urban areas or return back to the villages.

A useful way of addressing diversity of migration system is to treat the major contemporary cases in a general comparative framework (Simmons & Piche, 2002). More emphasis should be devoted to using models to analyse and explain important demographic issues and to predict demographic futures (Birch, 2002). One of the important property of micro level models is that the researcher can introduce individual heterogeneity- measured (accounting different rates for individual with different characteristics such as age, sex, social class, cohort, etc.) and unmeasured (choosing for each individual an adjusting factor where each on transition rate).

* Corresponding author e-mail: amritunjay.singh@gmail.com

Several Studies have been done to apply and/or formulate probability models in natural as well as social sciences (Afsar, 1995; Aryal, 2003, 2009; Wintle, 1992). Since probability models provide concise and clear representations of extensive data sets in a better way (Aryal, 2010), More attention has been paid to the proposition and derivation of probability models for the movement of human population at micro level (Yadava, 1977). Singh and yadava (1981) have introduced the negative binomial distribution to explore the pattern of rural-out migration at household level. Sharma (1984) applied this model to a different data set and found that this model is unsuitable for the total number of migrants (including females and children) from a household. Since female migrants are more likely to affect the socio-cultural characteristics of households in comparison to other females of household, it is important to study the pattern of total number of migrants from a household.

Sharma (1984) proposed a probability model under the assumption that (i) the number of male migrants aged 15 years and above follows negative binomial distribution and (ii) the distribution of living children to a couple is known. However, the distribution of living children to a couple has not yet been derived theoretically, the prior knowledge of these two distributions is difficult.

Singh (1985) proposed a probability model under the assumption that there are two types of households: first in which only males aged 15 years and above migrate and second in which males migrate with their wives and children. Several authors have proposed models in the same line to describe the distribution of household according to the total number of migrants including wife and children (Kushwaha, 1992; Sharma, 1987; Singh, 1990, 1992; Yadava, 1993; yadava & yadava, 1988; Yadava et al., 1989, 1994).

The present paper aims to investigate trends in rural-out migration at the household level through some probabilistic models and to modify in the exiting estimation procedure.

2 Data

The present study is based on primary data taken from a survey entitled “Migration and Related Characteristics- a Case Study of Northeastern Bihar” conducted during October 2009 to June 2010.

Data have been collected using a multistage random sampling procedure. The districts selected for the study are Katihar, Purnea and Bhagalpur. The sample consisted of (664) households from (8) villages.

3 Probability Models for Rural Out Migrants

3.1 Model A_1

A probability model for describing the variation in the number of rural out-migrant households has been derived on the basis of following assumptions:

– At the survey point, the household is either exposed, is not exposed to the migration risk. Let α and (1-α) be the respective probabilities.
– The probability of one male’s migration, from a household is greater than that of two males, and probability of two male’s migration is greater than that of three males from a household. Thus, the pattern of migration from a household is a decreasing function and follows a logarithmic series distribution with parameter λ.

Let x represent the number of rural out-migrants from a household. Under the assumptions (i) and (ii), the probability function of x is given by

$$\left.\begin{aligned} P(x=k) &= 1-\alpha, && k=0 \\ &= \alpha\left[\frac{-\lambda^k}{k\log(1-\lambda)}\right] \text{for} && k=1,2,3,\cdots,0<\lambda<1;0<\alpha<1 \end{aligned}\right\} \tag{1}$$

The log-series distribution has a long positive tail and the shape of the tail is similar to that of geometric distribution for large values of k. However, the log-series distribution has the advantage that it has only one parameter instead of two parameters of Negative Binomial Distribution (Chatfield et. al., 1966).

3.1.1 Estimation

Consider a sample consisting of n observations of the random variable x with probability function given by expression 1. Suppose that $n_k \quad (k=0,1,2,\cdots,m)$ represents the number of observations of k'th cell and $\sum_{k=1}^{n} n_k = n$ The likelihood function for the given sample $(x_1,x_2,\cdots,x_n)$ can be expressed as:

$$L[\alpha,\lambda|(X_1,x_2,\cdots,X_n)] = (1-\alpha)^{n_0}\prod_{k=1}^{m}\left[\alpha\left(\frac{\lambda^k}{k\log(1-\lambda)}\right)\right]^{n_k} \tag{2}$$

$$= \frac{(1-\alpha)^{n_0}(-\alpha)^{n-n_0}\lambda^{\sum_{k=1}^{m} n_k x_k}}{\left(\prod_{k=1}^{m} x_k^{n_k}\right)[\log(1-\lambda)]^{n-n_0}} \tag{3}$$

where x_k represents the value of k. Taking logarithms of equation (3), differentiating with respect to α and λ respectively and equating to zero give the following equations:

$$\frac{\delta \log L}{\delta\alpha} = -\frac{n_0}{1-\alpha} + \frac{n-n_0}{\alpha} = 0 \tag{4}$$

$$\frac{\delta \log L}{\delta\lambda} = \frac{\sum_{k=1}^{m} n_k x_k}{\lambda} + \frac{n-n_0}{(1-\lambda)(\log(1-\lambda))} = 0 \tag{5}$$

Equation (4) yields the estimator of α as $\tilde{\alpha} = \frac{n-n_0}{n}$ The estimating equation for λ is obtained by solving equation (5) as:

$$(1-\lambda)\log(1-\lambda)\sum_{k=1}^{m} n_k x_k + (n-n_0)\lambda = 0 \tag{6}$$

This equation can be solved numerically and the numerical solution of (6) is the desired maximum likelihood estimate for λ. Using the fact that $E(n_0) = n(1-\alpha), E(n-n_0) = n\alpha$

$$E(x_k) = \frac{\alpha\lambda}{(1-\lambda)\log(1-\lambda)} \text{and} \qquad E\left(\sum_{k=1}^{m} n_k x_k\right) = \frac{-n\alpha\lambda}{(1-\lambda)\log(1-\lambda)}$$

The expected values of second partial derivatives are obtained as

$$-E\left(\frac{\delta^2 \log L}{\delta\alpha^2}\right) = -\frac{E(n_0)}{(1-\alpha)^2} - \frac{E(n-n_0)}{\alpha^2} = \frac{n}{\alpha(1-\alpha)} = \phi_{11}(\text{Say}) \tag{7}$$

$$-E\left(\frac{\delta^2 \log L}{\delta\alpha^2}\right) = -\frac{E\left(\sum_{k=1}^{m} n_k x_k\right)}{\lambda^2} + \frac{[1+\log(1-\lambda)]E(n-n_0)}{[(1-\lambda)\log(1-\lambda)]^2}$$

$$= -n\alpha\left[\frac{1}{\lambda(1-\lambda)\log(1-\lambda)} + \frac{1+\log(1-\lambda)}{[(1-\lambda)\log(1-\lambda)]^2} = \phi_{22}(\text{Say})\right] \tag{8}$$

The covariance between α and λ's zero since $E\left(\frac{\delta^2 \log L}{\delta\alpha\delta\lambda}\right) = 0$, and hence the variance of α and λ can be obtained as

$$V(\tilde{\alpha}) = \frac{1}{\phi_{11}} \qquad \text{and} \qquad V(\tilde{\lambda}) = \frac{1}{\phi_{22}}$$

3.2 Model A_2

Sharma (1985) has proposed a probability model for the number of rural male out-migrants aged 15 years and above from a household under the following assumptions:

1. At any point in time, let α be the probability migration out from a household and $(1-\alpha)$ be the probability of not migrating from a household.
2. If p represents the probability of a single individual migrating from a household, the pattern of migration from each household follows the geometric distribution.

If x represents the number of migrants from a household, x follows the inflated geometric distribution with probability density function as

$$\left.\begin{array}{lll} P(x=0) = 1-\alpha+\alpha p, & & k=0 \\ P(x=k) = \alpha q^k p & \text{for} & k=1,2,3,\cdots, \end{array}\right\} \tag{9}$$

Where $p+q=1$.

3.2.1 Estimation

As mentioned above, Sharma (1985) used method of moments to estimate the parameters α and p of model (9) and obtained the asymptotic expressions for variance and covariance of the estimators using multivariate central limit theorem. Iwunor (1995) proposed an alternative estimation technique based on likelihood function and obtained the variance and covariance of the estimators.
Let $(X_1, X_2, ,..., X_n)$ denote a random sample of size n from the above-mentioned probability model. Furthermore, let $n_k(k=0,1,2,..,m)$ denote the number of observations corresponding to the kth cell. The likelihood function for estimating the parameters α and p can be expressed as:

$$L = (1-\alpha+\alpha p)^{n_0}\prod_{k=1}^{m}(\alpha p q^k)^{n_k}$$

$$= (1-\alpha+\alpha p)^{n_0}\alpha^{n-n_0}p^{n-n_0 q^s} \tag{10}$$

where, $n_0+n_1+n_2+\cdots+n_m = n$ and , $S = n_1+2n_2+3n_3+\cdots+mn_m = \sum_{k=1}^{m} n_k x_k$, x_k represents the value of k. Taking logarithms of the likelihood function (10), differentiating with respect to α and p respectively and equating to zero give the following estimating equations:

$$\frac{\delta \log L}{\delta \alpha} = \frac{n_0(p-1)}{(1-\alpha+\alpha p)} + \frac{n-n_0}{\alpha} = 0 \tag{11}$$

$$\frac{\delta \log L}{\delta p} = \frac{n-n_0}{p} - \frac{s}{1-p} + \frac{n_0\alpha}{(1-\alpha+\alpha p)} = 0 \tag{12}$$

Solution of equation (3.14) provides the estimate of α as

$$\tilde{\alpha} = \frac{n-n_0}{n(1-p)}$$

Substituting the value of α and after rearranging equation (12) yields the estimator of p as

$$\tilde{p} = \frac{n-n_0}{\sum_{k=1}^{m} n_k x_k}$$

Using the fact $E(n_0) = nP(X=0) = n(1-\alpha+\alpha p)$ and $E(n-n_0) = n\alpha(1-p)$ the expected value of second partial derivatives of $\log L$ can be obtained as

$$-E\left(\frac{\delta^2 \log L}{\delta \alpha^2}\right) = \frac{n(1-p)}{\alpha(1-\alpha+\alpha p)} = \phi_{11}(\text{say}) \tag{13}$$

$$-E\left(\frac{\delta^2 \log L}{\delta p^2}\right) = \frac{n\alpha q}{p^2}\frac{n\alpha}{pq} + \frac{n\alpha^2}{(1-\alpha+\alpha p)} = \phi_{22}(\text{say}) \quad (14)$$

$$-E\left(\frac{\delta^2 \log L}{\delta\alpha\delta p}\right) = -\frac{n}{(1-\alpha+\alpha p)} = \phi_{12}(\text{say}) \quad (15)$$

Therefore, through inverting the information matrix, the expressions for variances and covariance of the estimators can be obtained as

$$\left.\begin{aligned} V(\tilde{\alpha}) &= \frac{\phi_{22}}{\phi_{11}\phi_{22}-\phi_{12}^2} \\ V(\tilde{p}) &= \frac{\phi_{11}}{\phi_{11}\phi_{22}-\phi_{12}^2} \\ \text{and } cov(\tilde{\alpha},\tilde{p}) &= \frac{\phi_{12}}{\phi_{11}\phi_{22}-\phi_{12}^2} \end{aligned}\right\} \quad (16)$$

Table 1: Distribution of Observed and Expected Frequency of Number of Households According to the Number of Rural Out Migrants

No of Migrants	observed	Expected(Model I)	$\chi^2_{0.05}$
0	401	401	
1	147	157.5	
2	57	53.33	
3	29	24.08	
4	16	12.23	4.67
5	8	6.63	
6	5	9.22	
7	1		
8	0		
Total	664	664	
α	0.396084		
p	0.537832		
λ	0.677246		
var(α)	0.00036		
Var(λ)	0.000756		

4 Results

Estimating the parameter values for given data we had shown results in 2 different tables. Table-1 shows the observed frequencies from data and expected frequencies for Model A_1 corresponding to the estimated values of parameters and their frequencies. The fitness of data had been tested on the basis of χ^2 at 5% level of significance.

4.1 Table-1

Table 1 exhibits that for no migrants, the estimated frequency is same as observed frequency as 401. For 1 migrant the estimated frequency over estimated by 10 points to observed frequency of 147. For 2 migrants, the estimated value is 53.33 lower with 4 persons than observed value of 57. Study of 3 migrants manifested that observed frequency is 5 persons higher with respect to estimated frequency of 24.08 by value 29. For 4 migrants we found that it is 4 points underestimated than observed towards our proposed model. For value of 5 migrants we obtained the result as estimated frequency is 1 point lower to observed frequency. Since letter classes have very low frequencies so we pool data, where we get approximately similar

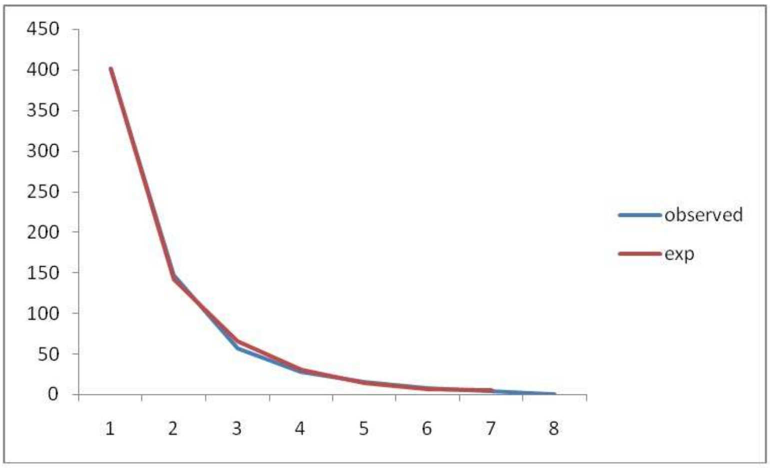


Fig. 1: Expected and Observed Frequency of Number of Migrants

Table 2: Distribution of Observed and Expected Frequency of Number of Households According to the Number of Rural Out Migrants

No of Migrants	observed	Expected(Model I)	$\chi^2_{0.05}$
0	401	401	
1	147	141.45	
2	57	65.37	
3	29	30.21	
4	16	13.96	2.109025
5	8	6.45	
6	5	5.55	
7	1		
Total	664	664	
p	0.537832		
α	0.857014		
var(α)	0.003434		
var(p)	0.000508		

total to observed frequencies. Total observed and estimated are equal. We get the estimated value of parameters for model 1 as $\alpha = 0.396$, $p = 0.538$ and $\lambda = 0.677$ with variance of parameters as $V(\tilde{\alpha}) = 0.0003$ and $var(\tilde{\lambda}) = 0.00075$. While testing the fitting of model for the data we get pooled χ^2 is 4.67, which provides that null hypothesis is being accepted or there is no significant difference in between observed and expected frequencies for model A_1 Figure-1 represents the difference in between the line of curve in observed and estimated values for model A_1.

4.2 Table-2

Table (2) reveals that for no migrants, the estimated frequency is same as observed frequency as 401. For 1 migrant the estimated frequency under estimated by 6 points to observed frequency of 147. For 2 migrants, the estimated value is 65.37 which is higher with 8 persons to observed value of 57. For 3 migrants, observed frequency is 1 person higher with respect to estimated frequency of 30.21 with respect to observed value 29. For 4 migrants we found that it is 3 points underestimated than observed towards our proposed model. For value of 5 migrants we get result as estimated frequency is 2 points lower than observed frequency. Since letter classes have very low frequencies so we pool data, where we get approximately similar

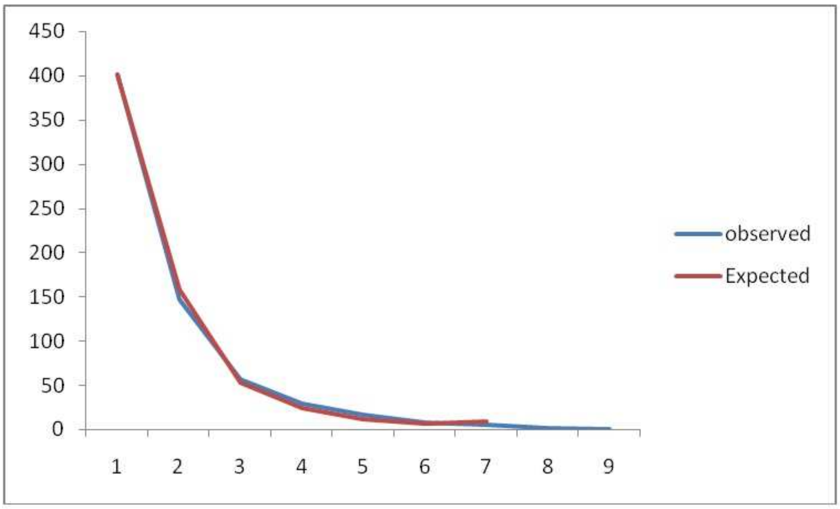


Fig. 2: Expected and Observed Frequency of Number of Migrants

total to observed frequencies. Total for observed and estimated are equal. We get the estimated value of parameters for model 1 as $\alpha = 0.857$, $p = 0.538$ with variance of parameters as $V(\tilde{\alpha}) = 0.0034$ and $var(\tilde{\lambda}) = 0.0005$. While testing the fitting of model for the data we get pooled χ^2 is 2.109, which provides that null hypothesis is being accepted or there is no significant difference in between observed and expected frequencies for model A_2. Figure-2 represents the difference in between the line of curve in observed and estimated values for model A_2

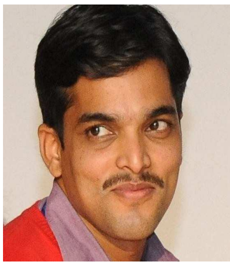
5 Conclusion

The study indicates that the proposed model (A_1) is a reasonable approximation to describe the distribution of households for the rural out migrants and at least at the micro-level. It also provides the estimates of the parameters of model (A_2) using maximum likelihood technique. The exact variance and covariance of the estimators for both the models (A_1 & A_2) have been computed.

References

[1] , Afsar, R. (1995): "Rural-urban migration and its impact on women roles and status, empowerment", A Journal of Women for Women, Vol. 2. Bangladesh.
[2] Aryal, T. R. (2003): "Probability models for the number of rural out-migrants at micro-level", The Nepali Mathematical Sciences Report, Vol. 21 (1& 2): 9-18.
[3] , Aryal, T. R. (2009): "Mathematical models to estimate the maternal mortality", Nepali Mathematical Sciences Report, Vol. 29 (1& 2) pp. 117-125.
[4] Aryal, T. R. (2010): "A new technique to construct the female marriage life-table", Nepali Mathematical Sciences Report, Vol. 30 (1& 2) pp. 111-122.
[5] Bhagat, R.B. (2005):Rural-urban classification and municipal governance in India, Singapore Journal of Tropical Geography, Wiley Online Library, Vol. 26(1): 109-114
[6] Birch, S. (2002): "Health human resource planning for the new millennium: Inputs in the production of health, illness and recovery in populations", Canadian Journal of Nursing Research, Vol. 33: 61-73.

[7] Friedlander, D., and Roshier, R. J. (1966). A study of internal migration in England and Wales: Part II. Recent internal migrants—their movements and characteristics. Population Studies, vol 20(1), 45-59.
[8] Greenwood, M. J. (1971): "A regression analysis of migration to urban areas of a less developed country: the case of India", Journal of Regional Sciences, Vol. 11, pp. 253-62.
[9] Isbell, E. C. (1944): "Internal migration in Sweden and intervening opportunities", American Sociological Review, Vol. 9, pp. 627-639.
[10] Lee, E. S. (1966): "A theory of migration", Demography, Vol. 3, pp. 47-57.
[11] Pryor, R. J. (1975): "Conceptualizing migration behavior: A problem of microdemographic analysis", In Kosinski, L. A. and Webb, J. W., New Zeland, pp.105-20.
[12] Sharma, H. L. (1987): "A probability distribution for out-migration", Janasamkhya, Vol. 5(2), pp. 95-101.
[13] Simmons, Alan and Victor, Piché. (2002): "Teaching migration and globalization", Genus, LVIII, No. 304, pp. 109-34.
[14] Singh, S. N. and Yadava, K. N. S. (1981): "Trends in rural out-migration at household level", Rural Demography, Vol. VIII, No. 1.
[15] Singh, S. N., Yadava, R. C., and Sharma, H. L. (1985). A model for rural out-migration at household level. Janasamkhya, 3(1-2), 1-7.
[16] Singh, B. P., Singh N. K., Roy, T. K. (2014): "On the Pattern of Migration in the Household: An Explanation through Binomial Law", International Journal of Statistics and Systems, Vol.9, No. 2. ISSN: 0973-2675.
[17] Singh, B. P., Singh N. K., Dixit, S. (2015): "Estimation of Parameters of Inflated Poisson-Lindley Distribution for Adult Out Migration" Journal of Institute of science & Technology, Vol. 20, Issue 2. ISSN: 2469-9062.
[18] Singh, Brijesh P., Singh Niraj K. (2016): "On the Distribution of Risk of Migration and Its Estimation" International Journal of Statistical Distributions and Applications, Vol. 2, Issue 4. ISSN: 2472-3487.
[19] Stouffer, S. A. (1940): "Intervening opportunities: a theory relating mobility and distance", American Sociological Review, Vol. 5 (6), pp. 845-867.
[20] Stouffer, S. A. (1960): "Intervening opportunities and competing migrants", Journal of Regional Science, Vol. 2, pp. 1-26.
[21] Wintle, M. (1992): "Push-factors in emigration: the case of province of Zeeland in the nineteenth Century", Population Studies, Vol. 46, pp. 523-537.
[22] Yadava, K. N. S., and Yadava, G. S. (1988): "On some probability models and their applications to the distribution of the number of migrants from a household", Janasankhya, Vol. 6(2), 137-158.
[23] Yadava, K. N. S., and Yadava, G. S. and Singh S. K. (1989): "Pattern of return migration in rural areas of eastern Uttar Pradesh, India", Journal of Institute of Economic Research, 24, No. 2, pp. 27-48.
[24] Yadava, K.N.S.,Tripathi, S. and Singh, V.S.(1994): "A model for the number of out migrants at the household level : An alternative approach", The Journal of Scientific Research, Vol. 44, pp. 125 134, 1994.
[25] Zipf, G. (1946): "The P1P2/D hypothesis: On the intercity movement of persons", American Sociological Review, Vol. 11, pp. 677-686.
Books
[26] Rossi, P. H. (1955): "Why families move: A study in the social psychology of urban residential mobility", Free Glencoe Press, New York.
[27] Aryal, T. R. (2002): "Some demographic models and their applications with reference to Nepal", unpublished thesis in Statistics, Banaras Hindu University, India.
[28] Kushwaha, S. N. S. (1992): "Some probability models for rural out-migration and migrant's fertility", unpublished Ph.D. thesis, Banaras Hindu University, India.
[29] Sharma, L. (1984): "A study of the pattern of out-migration from rural areas", unpublished Ph.D. thesis, Banaras Hindu University, India.
[30] Singh, R. B. (1986): "A study of migration pattern from rural areas", unpublished Ph.D. thesis in statistics, Banaras Hindu University, India.
[31] Yadava, K. N. S. (1977): "Some aspects of out-migration and its effects on fertility", unpublished Ph.D. thesis in statistics, Banaras Hindu University, India.
[32] Yadava, S. N. (1993): "On some migration and population growth models", Unpublished Ph.D. thesis in Statistics, Banaras Hindu University, Varanasi, India.



Niraj Kumar Singh is working as an Assistant Professor of Statistics in Amity University, Noida. His research interests span the areas of fertility, migration, and family dynamics. He is particularly interested in examining the social, structural, and economic processes that shape internal migration flows; effects of migration of fertility, marriage and urbanization. He is actively involved in teaching, research and consultancy.

Mritunjay Pal Singh is Assistant Professor of Statistics at Amity University. He received the Ph.D. degree in Statistics at Banaras Hindu University. He is working on Demography especially in Infant Mortality. He is trying to establish some fruitful approaches as they that can cover most of the realistic situations.