Current Journal of Applied Science and Technology

39(41): 78-91, 2020; Article no.CJAST.63311
ISSN: 2457-1024
(Past name: British Journal of Applied Science & Technology, Past ISSN: 2231-0843, NLM ID: 101664541)

Some Size Biased Probability Models for Adult out Migration Pattern

Brijesh P. Singh[1], Sandeep Singh[1*], Utpal Dhar Das[1] and Gunjan Singh[2]

[1]*Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi, India.*
[2]*Department of Statistics, Amity School of Applied Sciences, Amity University, Lucknow, India.*

Authors' contributions

This work was carried out in collaboration among all authors. Authors BPS and GS designed the study, developed the probability models, wrote the protocol and first draft of the manuscript. Authors SS and UDD managed the statistical analyses of the study. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/CJAST/2020/v39i4131122
Editor(s):
(1) Prof. Qing-Wen Wang, Shanghai University, China.
Reviewers:
(1) Wankhede Vishal Ashok, SHHJB Polytechnic, MSBTE, India.
(2) Jaffry Zakaria, Sultan Idris Education University, Malaysia.
Complete Peer review History: http://www.sdiarticle4.com/review-history/63311

Original Research Article

Received 05 October 2020
Accepted 11 December 2020
Published 21 December 2020

ABSTRACT

In this paper an attempt has been made to inspect the distribution of the number of adult migrants from household through size biased probability models based on certain assumptions. Size Biased Poisson distribution compounded with various forms of Gamma distribution i.e. Gamma $(1,\theta)$, Gamma $(2,\theta)$ and mixture of Gamma $(1,\theta)$ and Gamma $(2,\theta)$ has been examined for some real data set of adult migration. The parameters of the proposed model have been estimated by method of maximum likelihood. χ^2 test indicates that the distributions proposed here are quite satisfactory to explain the pattern of adult out migration.

Keywords: Probability models; migration pattern; adult out migration; Poisson distribution; Gamma distribution.

**Corresponding author: E-mail: sandeepsingh.stats@gmail.com;*

1. INTRODUCTION

Migration is a basic phenomenon that affects population composition of a society. It is a process that includes several factors like age, sex, marital status, education and some other population event, affecting the movements of an individual of a household. Adults are more prone to migrate than other persons of the society. Lack of social amenities, job opportunities, and education facilities compels an adult to migrate from a place to other. Many attempts have been made for study of migration at micro level. Migration believed as an event, is highly selective with regard to age, with young adults generally being the most mobile group in any population. The probability models can also be adopted for explanation of any population event very efficiently. Several studies have been done to explain rural out migration [1,2,3,4,5,6,7,8,9,10] with the help of different probability models.

First attempt of probabilistic model building in this direction was initiated by Singh & Yadava [1], explained rural out migration with the help of probability model using negative binomial distribution. Yadava & Singh [2] introduced an idea of cluster for the number of migrants from a household and proposed a model assuming that migrants from a household occur in clusters. Sharma [4] assumed that the number of migrants per cluster follows the displaced geometric distribution and got the probability distribution for the number of migrants as Polya-Aeppli distribution. Further, Singh et al. [11] has applied a mixture of negative binomial and thomas distribution to describe the pattern of total number of migrants from a household. Sharma (1985) introduced a new parameter for capturing inflated nature of pattern of migration in model building. He applied the inflated geometric distribution as well as the inflated generalized Poisson distribution for probability modelling to describe the trends in rural out-migration at the micro level. Yadava & Yadava [6] proposed a model with displaced geometric distribution instead of taking poisson distribution for the occurrence of number of migrants and truncated a truncated Polya-Aeppli distribution.

An alternative estimation technique using likelihood function for inflated geometric distribution is proposed by Iwunor [12] and also obtained the variance and covariance for the estimators. The likelihood function using multinomial combination is only derived, but finally estimates of the parameters obtained by mean-zero frequency method by Iwunor [12]. Hossain [13] and Aryal [14] used maximum likelihood method to estimate the parameters of the model considered and applied it to different data set. Hossain [13] has also used the geometric model for describing the pattern of migration in Bangladesh. Again, Aryal [9] has used the same model to examine the pattern of migration in Nepal.

Recently, Singh et al. [15] developed a model for adult migration for the fixed household size to know the effect of size of household on the adult migration and used inflated binomial and beta binomial distribution. Singh et al. [16] introduced inflated Poisson-Lindley distribution for the pattern of adult migration in the household. Further Singh et al. [10] has applied inflated geometric and beta-geometric based models for explaining the pattern of migration. Singh et al. [17] has applied some zero adjusted distribution for exploring the pattern of adult out migration and found a good fit.

In this article an attempt has made to examine the suitability of the proposed models applying them to various datasets to explore pattern of adult out migration of some regions with probability models in aggregate. The probability models are employed to explore the real data set of adult migration to examine the suitability of models.

2. MODEL

For development of model we assume, some of the households have varying number of adult migrants and some household have no adult migrants. Thus number of household with no migration becomes truncated and the data can be better explained with size biased models. Keeping this fact into consideration an attempt has been made with such assumptions we propose a probability model for the number of rural adult out migrants from a household:

(i) At any point of time, let α be the probability of household in which adult migration occurs and the probability of household having no migration is $(1-\alpha)$.

(ii) If X be the number of adult migrants from a household and follows size biased Poisson distribution with parameter λ and it is assumed that this parameter varies according to

1. Gamma $(1,\theta)$ distribution i.e. exponential distribution.
2. Gamma $(2,\theta)$ distribution i.e. size biased exponential distribution.
3. A mixture of Gamma $(1,\theta)$ and Gamma $(2,\theta)$ distribution i.e. Lindley distribution.

It is assumed because huge disparity in terms of social standard of the household present in the society which affects the amount of migration. Now, for size-biased distribution, let X be a non negative random variable having probability function $p^*(x,\theta)$, with unknown parameter θ, where $\theta \in \Theta$, the parameter space and the $E(X)$ expected value. Then $p(x,\theta)$ is known as size biased distribution with probability function as

$$p(x,\theta) = \frac{x.p^*(x,\theta)}{E(x)} \tag{2.1}$$

So, let X be a non negative random variable having probability function $p^*(x)$,

$$p^*(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad ; \lambda > 0;\ x = 0,1,2,..... \tag{2.2}$$

with mean $E(X) = \lambda$ then size-biased Poisson distribution will be

$$p(x) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \quad ; \quad \lambda > 0;\ x = 1,2,.... \tag{2.3}$$

with mean $E(X) = 1+\lambda$.

2.1 Model I

Under above two assumptions of model, the probability distribution for the number of adult migrants at household level can be established as size-biased Poisson with gamma $(1,\theta)$ distribution which can be given as

$$p(X=x) = p_x(\lambda) = \int_0^\infty P(X = x|\lambda) f(\lambda)\, d\lambda = \frac{\theta}{(x-1)!}\int_0^\infty e^{-\lambda}\lambda^{x-1}e^{-\theta\lambda}\, d\lambda \tag{2.1.1}$$

where $P(X = x|\lambda) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}$ and $f(\lambda) = \theta e^{-\theta\lambda}$. After integration, we get,

$$p(X=x) = \frac{\theta}{(\theta+1)^x} \quad ; \ \theta > 0,\ x = 1,2,..... \tag{2.1.2}$$

with mean $E(X) = 1+\frac{1}{\theta}$. Here, if we assume $p = \frac{\theta}{(\theta+1)}$ and $q = 1-p = \frac{1}{(\theta+1)}$ the above distribution

reduces to $p(x) = \left(\frac{\theta}{(\theta+1)}\right)\left(\frac{1}{(\theta+1)}\right)^{x-1} = pq^{x-1} \quad ; \ \theta > 0,\ x = 1,2,.....$ (2.1.3)

which is the truncated form of geometric distribution i.e. the marginal distribution of X is a zero truncated geometric distribution with parameter (p).

Hence, finally the first proposed model is

$$p(x) = \begin{cases} 1-\alpha & ; \ x = 0 \\ \alpha\frac{\theta}{(\theta+1)^x} & ; \ x = 1,2,..... \end{cases} \tag{2.1.4}$$

2.1.1 Moments and related measures of Model I

The r^{th} factorial moment about origin of above proposed model can be obtained as

$$\mu_{(r)}' = E\left[E\left(X^{(r)}|\lambda\right)\right], \quad \text{where } X^{(r)} = X(X-1)(X-2)...(X-r+1) \tag{2.1.5}$$

$$= \alpha\int_0^\infty\left[\sum_{x=1}^\infty X^{(r)}\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\right].\theta e^{-\theta\lambda}d\lambda = \alpha\int_0^\infty\left[\lambda^{r-1}\left\{\sum_{x=r}^\infty x\frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!}\right\}\right].\theta e^{-\theta\lambda}d\lambda$$

Let, $y = x - r$ then

$$= \alpha\int_0^\infty\left[\lambda^{r-1}\left\{\sum_{y=0}^\infty (y+r)\frac{e^{-\lambda}\lambda^{y}}{y!}\right\}\right].\theta e^{-\theta\lambda}d\lambda = \alpha\int_0^\infty \lambda^{r-1}(\lambda+r)\theta e^{-\theta\lambda}d\lambda$$

Solving the integration, we get,

$$\mu_{(r)}' = \alpha\frac{\underline{|r}(1+\theta)}{\theta^r} \tag{2.1.6}$$

$$\mu_1' = \frac{\alpha(1+\theta)}{\theta}, \quad \mu_2' = \frac{2\alpha(1+\theta)}{\theta^2}, \quad \mu_3' = \frac{6\alpha(1+\theta)}{\theta^3}, \quad \mu_4' = \frac{24\alpha(1+\theta)}{\theta^4}$$

Now the central moments of the model are given by

$$\mu_2 = \mu_2' - (\mu_1')^2 = \frac{\alpha(1+\theta)(2-\alpha(1+\theta))}{\theta^2} \tag{2.1.7}$$

$$\mu_3 = \frac{2\alpha(1+\theta)\left(2\alpha^2(1+\theta)^2 + 3\alpha(1+\theta) + 3\right)}{\theta^3} \tag{2.1.8}$$

$$\mu_4 = \frac{3\alpha(1+\theta)\left(8 - 8\alpha(1+\theta) + 4\alpha^2(1+\theta)^2 - \alpha^3(1+\theta)^3\right)}{\theta^4} \tag{2.1.9}$$

The expressions for coefficient of variation ($C.V.$), coefficient of Skewness ($\sqrt{\beta_1}$) and coefficient of Kurtosis (β_2) of the model I are, thus obtained as

$$C.V. = \frac{\sqrt{\alpha(1+\theta)(2-\alpha(1+\theta))}}{\alpha(1+\theta)} \tag{2.1.10}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2\alpha(1+\theta)\left(2\alpha^2(1+\theta)^2 + 3\alpha(1+\theta) + 3\right)}{\left[\alpha(1+\theta)(2-\alpha(1+\theta))\right]^{3/2}} \tag{2.1.11}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\left(8-8\alpha(1+\theta)+4\alpha^2(1+\theta)^2-\alpha^3(1+\theta)^3\right)}{\left[\alpha(1+\theta)\left(2-\alpha(1+\theta)\right)\right]^2} \tag{2.1.12}$$

2.2 Model II

Under above two assumptions of model building, the second model for the probability distribution for the number of male migrants at household level can be proposed assuming size-biased Poisson with gamma $(\theta,2)$ distribution (i.e. size-biased exponential distribution) that can be modelled as

The size-biased Poisson distribution

$$P(X=x|\lambda) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \quad ; \quad \lambda>0;\ x=1,2,.... \tag{2.2.1}$$

with mean $E(X)=1+\lambda$

Here, the parameter λ follows gamma $(\theta,2)$ distribution.

$$f(\lambda)=\lambda\theta^2e^{-\theta\lambda} \quad ; \quad \lambda>1,\ \theta>0 \tag{2.2.2}$$

Therefore, size-biased Poisson with gamma $(\theta,2)$ distribution can be written as

$$p(X=x)=p_x(\lambda)=\int_0^\infty P\left(X=x|\lambda\right)f(\lambda)d\lambda=\frac{\theta^2}{(x-1)!}\int_0^\infty e^{-\lambda}\lambda^{x-1}\lambda e^{-\theta\lambda}d\lambda$$

After integration, we get,

$$p(x)=\theta x\left(\frac{\theta}{\theta+1}\right)\left(\frac{1}{\theta+1}\right)^x \quad ;\ \theta>0,\ x=1,2,...$$

with mean $E(X)=1+\frac{2}{\theta}$. Here if we assume $p=\frac{\theta}{(\theta+1)}$ and $q=1-p=\frac{1}{(\theta+1)}$ the above distribution reduces to

$$p(x)=x\left(\frac{\theta}{(\theta+1)}\right)^2\left(\frac{1}{(\theta+1)}\right)^{x-1}=xp^2q^{x-1} \quad ; \quad \theta>0,\ x=1,2,..... \tag{2.2.3}$$

which is the size biased form of geometric distribution i.e. the marginal distribution of X is a size-biased geometric distribution with parameter (p). Hence, finally the proposed second model is

$$p(x)=\begin{cases}1-\alpha & ;\ x=0\\ \alpha\theta x\left(\frac{\theta}{\theta+1}\right)\left(\frac{1}{\theta+1}\right)^x & ;\ x=1,2,.....\end{cases} \tag{2.2.4}$$

2.2.1 Moments and related measures of Model II

The r^{th} factorial moment about origin of above proposed model can be obtained as

$\mu_{(r)}' = E\left[E\left(X^{(r)}|\lambda\right)\right]$, where $X^{(r)} = X(X-1)(X-2)...(X-r+1)$

$$\mu_{(r)}' = \int_0^\infty \alpha\left[\sum_{x=1}^{\infty} X^{(r)} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\right].\lambda\theta^2 e^{-\theta\lambda} d\lambda = \alpha\int_0^\infty\left[\lambda^{r-1}\left\{\sum_{x=r}^{\infty} x\frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!}\right\}\right].\lambda\theta^2 e^{-\theta\lambda} d\lambda$$

Let, $y = x - r$ then

$$= \alpha\int_0^\infty\left[\lambda^{r-1}\left\{\sum_{y=0}^{\infty}(y+r)\frac{e^{-\lambda}\lambda^{y}}{y!}\right\}\right].\lambda\theta^2 e^{-\theta\lambda} d\lambda = \alpha\int_0^\infty \lambda^{r-1}(\lambda+r)\lambda\theta^2 e^{-\theta\lambda} d\lambda$$

Solving the integration, we get,

$$\mu_{(r)}' = \alpha\frac{\underline{|r}\left((r+1)+\theta r\right)}{\theta^r} \tag{2.2.5}$$

$$\mu_1' = \alpha\frac{(\theta+2)}{\theta} \quad \mu_2' = \frac{2\alpha(2\theta+3)}{\theta^2} \quad \mu_3' = \frac{6\alpha(3\theta+4)}{\theta^3} \quad \mu_4' = \frac{24\alpha(4\theta+5)}{\theta^4}$$

Now the central moments of the model II are given by

$$\mu_2 = \mu_2' - (\mu_1')^2 = \frac{\alpha}{\theta^2}\left[2(2\theta+3) - \alpha(\theta+2)^2\right] \tag{2.2.6}$$

$$\mu_3 = \frac{2\alpha}{\theta^3}\left[\alpha(\theta+2)^2 - 3\alpha(\theta+2)(2\theta+3) + 3(3\theta+4)\right] \tag{2.2.7}$$

$$\mu_4 = \frac{3\alpha}{\theta^4}\left[-\alpha^3(\theta+2)^4 + 4\alpha\left(2\theta^3+5\theta^2-4\right) - 8(4\theta+5)\right] \tag{2.2.8}$$

The expressions for coefficient of variation ($C.V.$), coefficient of Skewness ($\sqrt{\beta_1}$) and coefficient of Kurtosis (β_2) of the model II are thus obtained as

$$C.V. = \frac{\sqrt{\alpha\left[2(2\theta+3) - \alpha(\theta+2)^2\right]}}{\alpha(\theta+2)} \tag{2.2.9}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2\alpha(\theta+2)^2 - 6\alpha(\theta+2)(2\theta+3) + 6(3\theta+4)}{\sqrt{\alpha}\left[2(2\theta+3) - \alpha(\theta+2)^2\right]^{3/2}} \tag{2.2.10}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\left[-3\alpha^3(\theta+2)^4 + 12\alpha(\theta+2)^2 - 24\alpha(\theta+2)(3\theta+4) + 24(4\theta+5)\right]}{\alpha\left[2(2\theta+3) - \alpha(1+\theta)^2\right]^2} \tag{2.2.11}$$

2.3 Model III

According to the above two assumptions, the third model for the probability distribution for the number of male migrants at household level can be proposed assuming the size biased Poisson with Lindley distribution. The size-biased Poisson distribution

$$P(X = x|\lambda) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \quad ; \quad \lambda > 0; \; x = 1,2,.... \tag{2.3.1}$$

with mean $E(X) = 1 + \lambda$

Here, the parameter λ follows Lindley distribution. We know that, Lindley distribution with the probability density function (pdf):

$$f(\lambda) = \frac{\theta^2(1+\lambda)e^{-\theta\lambda}}{(\theta+1)} \quad ; \theta > 0 \; , \; x = 0,1,2,........., \tag{2.3.2}$$

with mean $\frac{\theta+2}{\theta(\theta+2)}$, which is mixture of Gamma $(1,\theta)$ and Gamma $(2,\theta)$ distributions with mixing parameter $\frac{\theta}{\theta+1}$ and $\frac{1}{\theta+1}$. Therefore, size-biased poisson with lindley distribution can be written as

$$p(X = x) = p_x(\lambda) = \int_0^\infty P(X = x|\lambda) f(\lambda) d\lambda = \frac{\theta^2}{(\theta+1)(x-1)!}\int_0^\infty e^{-\lambda(\theta+1)}(1+\lambda)\lambda^{x-1} d\lambda$$

After integration, we get,

$$p(x) = \frac{\theta^2(\theta+x+1)}{(\theta+1)^{x+2}} \quad ; \; \theta > 0, \; x = 1,2,... \tag{2.3.3}$$

with mean $E(X) = \frac{\theta^2+2\theta+2}{\theta(\theta+1)}$. This probability distribution can also be obtained by mixing model I and model II with mixing parameter $\frac{\theta}{\theta+1}$ and $\frac{1}{\theta+1}$.

$$p(x) = \frac{\theta}{(\theta+1)} p_1(x) + \frac{1}{(\theta+1)} p_2(x) \tag{2.3.4}$$

where $p_1(x) = \frac{\theta}{(\theta+1)^x}$ (i.e. model I) and $p_2(x) = x\left(\frac{\theta^2}{\theta+1}\right)\left(\frac{1}{\theta+1}\right)^x$ (i.e. model II). Therefore

$$p(x)=\frac{\theta}{(\theta+1)}\left(\frac{\theta}{(\theta+1)^{x}}\right)+\frac{1}{(\theta+1)}\left(x\left(\frac{\theta^2}{\theta+1}\right)\left(\frac{1}{\theta+1}\right)^{x}\right)=\frac{\theta^2}{(\theta+1)^{x+1}}+\frac{x\theta^2}{(\theta+1)^{x+2}}$$

$$=\frac{\theta^2}{(\theta+1)^{x+1}}\left(1+\frac{x}{(\theta+1)}\right)=\frac{\theta^2(\theta+x+1)}{(\theta+1)^{x+2}} \tag{2.3.5}$$

This is same probability distribution as above. Hence, finally the proposed third model is

$$p(x)=\begin{cases}1-\alpha & ;\ x=0\\ \alpha\dfrac{\theta^2(\theta+x+1)}{(\theta+1)^{x+2}} & ;\ x=1,2,.....\end{cases} \tag{2.3.6}$$

2.3.1 Moments and related measures of Model III

The r^{th} factorial moment about origin of above proposed model can be obtained as

$$\mu_{(r)}'=E\left[E\left(X^{(r)}|\lambda\right)\right],$$

where $X^{(r)}=X(X-1)(X-2)...(X-r+1)$

$$\mu_{(r)}'=\int_0^\infty \alpha\left[\sum_{x=1}^{\infty}X^{(r)}\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\right].\lambda\theta^2 e^{-\theta\lambda}d\lambda=\alpha\int_0^\infty\left[\lambda^{r-1}\left\{\sum_{x=r}^{\infty}x\frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!}\right\}\right].(1+\lambda)\frac{\theta^2}{(\theta+1)}e^{-\theta\lambda}d\lambda$$

Let, $y=x-r$ then

$$=\alpha\int_0^\infty\left[\lambda^{r-1}\left\{\sum_{y=0}^{\infty}(y+r)\frac{e^{-\lambda}\lambda^{y}}{y!}\right\}\right](1+\lambda)\frac{\theta^2}{(\theta+1)}e^{-\theta\lambda}d\lambda=\alpha\int_0^\infty\lambda^{r-1}(\lambda+r)(1+\lambda)\frac{\theta^2}{(\theta+1)}e^{-\theta\lambda}d\lambda$$

Solving the integration, we get,

$$\mu_{(r)}'=\alpha\frac{\lfloor r\,\theta^2\left(\theta(\theta+r+1)+r+1\right)}{(\theta+1)\theta^{r+2}} \tag{2.3.7}$$

$$\mu_1'=\frac{\alpha(\theta^2+2\theta+2)}{\theta(\theta+1)}\qquad \mu_2'=\frac{\alpha(2\theta^2+6\theta+6)}{\theta^2(\theta+1)}\qquad \mu_3'=\frac{\alpha(6\theta^2+24\theta+24)}{\theta^3(\theta+1)}\qquad \mu_4'=\frac{\alpha(24\theta^2+120\theta+120)}{\theta^4(\theta+1)}$$

Now, the central moments of the model are given by

$$\mu_2=\mu_2'-(\mu_1')^2=\frac{\alpha}{\theta^2(\theta+1)^2}\left[-\alpha\theta^4+2\theta^3(1-4\alpha)+4\theta^2(3-8\alpha)+4\theta(3-2\alpha)-4\alpha+6\right] \tag{2.3.8}$$

$$\mu_3=\frac{\alpha\left[2\alpha^2(\theta^2+2\theta+2)^3-3\alpha(2\theta^2+6\theta+6)(\theta^2+2\theta+2)(\theta+1)+(6\theta^2+24\theta+24)(\theta+1)^2\right]}{\theta^3(\theta+1)^3} \tag{2.3.9}$$

$$\mu_4 = \frac{\alpha\begin{bmatrix} -3\alpha^3\left(\theta^2+2\theta+2\right)^4+6\alpha^2\left(2\theta^2+6\theta+6\right)\left(\theta^2+2\theta+2\right)\left(\theta+1\right) \\ -4\alpha\left(6\theta^2+24\theta+24\right)\left(\theta^2+2\theta+2\right)\left(\theta+1\right)^2+\left(24\theta^2+120\theta+120\right)\left(\theta+1\right)^3 \end{bmatrix}}{\theta^4\left(\theta+1\right)^4} \tag{2.3.10}$$

The expressions for coefficient of variation (*C.V.*), coefficient of Skewness ($\sqrt{\beta_1}$) and coefficient of Kurtosis (β_2) of the model III are thus obtained as

$$C.V. = \frac{\sqrt{\alpha\left(2\theta^2+6\theta+6\right)\left(\theta+1\right)}}{\alpha\left(\theta^2+2\theta+2\right)} \tag{2.3.11}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2\alpha^2\left(\theta^2+2\theta+2\right)^3+3\alpha\left(2\theta^2+6\theta+6\right)\left(\theta^2+2\theta+2\right)\left(\theta+1\right)+\left(6\theta^2+24\theta+24\right)\left(\theta+1\right)^2}{\sqrt{\alpha}\left[-\alpha\theta^4+2\theta^3\left(1-4\alpha\right)+4\theta^2\left(3-8\alpha\right)+4\theta\left(3-2\alpha\right)-4\alpha+6\right]^{3/2}} \tag{2.3.12}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\begin{bmatrix} -3\alpha^3\left(\theta+2\right)^4+6\alpha^2\left(2\theta^2+6\theta+6\right)\left(\theta^2+2\theta+2\right)\left(\theta+1\right) \\ -4\alpha\left(6\theta^2+24\theta+24\right)\left(\theta^2+2\theta+2\right)\left(\theta+1\right)^2+\left(24\theta^2+120\theta+120\right)\left(1+\theta\right)^3 \end{bmatrix}}{\alpha\left[-\alpha\theta^4+2\theta^3\left(1-4\alpha\right)+4\theta^2\left(3-8\alpha\right)+4\theta\left(3-2\alpha\right)-4\alpha+6\right]^2} \tag{2.3.13}$$

3. ESTIMATION OF PARAMETERS

3.1 Model I

Suppose $x_1, x_2, \ldots, x_n$ be a random sample and the probability distribution function of the sample. Here, Assume that n_k (k=0, 1, 2,…, m) are the observation of the k^{th} cell and $\sum_{k=0}^{m} n_k = n$. The mle for the given sample ($x_1, x_2, \ldots, x_n$) can be given as:

$$L\left[\alpha,\theta \mid (x_1,x_2,\ldots x_n)\right] = \left(1-\alpha\right)^{n_0}\prod_{k=1}^{m}\left[\frac{\alpha\theta}{\left(\theta+1\right)^{x_k}}\right]^{n_k}$$
$$= \left(1-\alpha\right)^{n_0}\alpha^{n-n_0}\theta^{n-n_0}\prod_{k=1}^{m}\frac{1}{\left(\theta+1\right)^{x_k n_k}} \tag{3.1.1}$$

and so its log-likelihood function is thus obtained as

$$\log L = n_0\log\left(1-\alpha\right)+\left(n-n_0\right)\log\alpha+\left(n-n_0\right)\log\theta-\left(\sum_{k=1}^{m}x_k n_k\right)\log\left(\theta+1\right) \tag{3.1.2}$$

After differentiation of the above equation and equating it with zero, we get the following equations

$$\frac{\partial \log L}{\partial\alpha} = \frac{-n_0}{\left(1-\alpha\right)}+\frac{n-n_0}{\alpha} = 0 \tag{3.1.3}$$

$$\frac{\partial \log L}{\partial \theta} = \frac{\left(n-n_0\right)}{\theta} - \frac{S}{\left(\theta+1\right)} = 0 \qquad \textit{where } S = \sum_{k=1}^{m} x_k n_k \tag{3.1.4}$$

The ML estimates of α and θ from the above equations are given as

$$\hat{\alpha} = \frac{n-n_0}{n} = 1 - \frac{n_0}{n} \text{ and } \hat{\theta} = \frac{\alpha}{\bar{x}-\alpha}$$

3.2 Model II

Suppose $x_1, x_2, \ldots, x_n$ be a random sample and the probability distribution function of the sample. Here, Assume that n_k (k=0, 1, 2,…, m) are the observation of the k^{th} cell and $\sum_{k=0}^{m} n_k = n$. The mle for the given sample ($x_1, x_2, \ldots, x_n$) can be given as:

$$L\left[\alpha,\theta \mid (x_1, x_2, \ldots x_n)\right] = (1-\alpha)^{n_0} \prod_{k=1}^{m} \left[\alpha\theta x_k \left(\frac{\theta}{\theta+1}\right)\left(\frac{1}{\theta+1}\right)^{x_k}\right]^{n_k}$$

$$= (1-\alpha)^{n_0} \alpha^{n-n_0} \theta^{2(n-n_0)} \left(\frac{1}{\theta+1}\right)^{n-n_0} \left(\frac{1}{\theta+1}\right)^{\sum_{k=1}^{m} x_k n_k} \prod_{k=1}^{m} x_k^{n_k} \tag{3.2.1}$$

and its log likelihood function is thus obtained as

$$\log L = n_0 \log(1-\alpha) + (n-n_0)\log\alpha + 2(n-n_0)\log\theta - (n-n_0)\log(\theta+1) - \left(\sum_{k=1}^{m} x_k n_k\right)\log(\theta+1) - \sum_{k=1}^{m} n_k \log x_k \tag{3.2.2}$$

After differentiation of the above equation and equating it with zero, we get the following equations:

$$\frac{\partial \log L}{\partial \alpha} = \frac{-n_0}{(1-\alpha)} + \frac{n-n_0}{\alpha} = 0 \tag{3.2.3}$$

$$\frac{\partial \log L}{\partial \theta} = \frac{2(n-n_0)}{\theta} - \frac{n-n_0}{(\theta+1)} - \frac{S}{(\theta+1)} = 0 \qquad \textit{where } \sum_{k=1}^{m} x_k n_k \tag{3.2.4}$$

The ML estimates of α and θ from the above equations are given as

$$\hat{\alpha} = \frac{n-n_0}{n} = 1 - \frac{n_0}{n} \text{ and } \hat{\theta} = \frac{2\alpha}{\bar{x}-\alpha}$$

3.3 Model III

Suppose $x_1, x_2, \ldots, x_n$ be a random sample and the probability distribution function of the sample. Here, Assume that n_k (k=0, 1, 2,…, m) are the observation of the k^{th} cell and $\sum_{k=0}^{m} n_k = n$. The mle for the given sample ($x_1, x_2, \ldots, x_n$) can be given as

$$L\left[\alpha,\theta \mid (x_1,x_2,...x_n)\right]=\left(1-\alpha\right)^{n_0}\prod_{k=1}^{m}\left[\alpha\theta^2\frac{\left(\theta+x_k+1\right)}{\left(1+\theta\right)^{x_k+2}}\right]^{n_k}$$

$$=\left(1-\alpha\right)^{n_0}\alpha^{n-n_0}\theta^{2(n-n_0)}\left(\frac{1}{\theta+1}\right)^{2(n-n_0)}\left(\frac{1}{\theta+1}\right)^{\sum_{k=1}^{m}x_k n_k}\prod_{k=1}^{m}\left(x_k+\theta+1\right)^{n_k} \tag{3.3.1}$$

and its log likelihood function is thus obtained as

$$\log L=n_0\log\left(1-\alpha\right)+\left(n-n_0\right)\log\alpha+2\left(n-n_0\right)\log\theta-2\left(n-n_0\right)\log\left(\theta+1\right)$$
$$-\left(\sum_{k=1}^{m}x_k n_k\right)\log\left(\theta+1\right)+\sum_{k=1}^{m}n_k\log\left(x_k+\theta+1\right) \tag{3.3.2}$$

After differentiation of the above equation and equating it with zero, we get the following equations:

$$\frac{\partial\log L}{\partial\alpha}=\frac{-n_0}{\left(1-\alpha\right)}+\frac{n-n_0}{\alpha}=0 \tag{3.3.3}$$

$$\frac{\partial\log L}{\partial\theta}=\frac{2\left(n-n_0\right)}{\theta}-\frac{2\left(n-n_0\right)}{\left(\theta+1\right)}-\frac{S}{\left(\theta+1\right)}+\sum_{i=1}^{m}\frac{n_k}{x_k+\theta+1}=0 \tag{3.3.4}$$

The ML estimate of from the above equation is given as

$$\hat{\alpha}=\frac{n-n_0}{n}=1-\frac{n_0}{n}$$

and the estimate of θ cannot be estimated analytically. Therefore, we can get the estimates numerically.

4. RESULTS AND DISCUSSION

The parameters estimated for the proposed models with maximum likelihood. Further to check the suitability, when the model has been applied to the data set collected from a survey entitled “Migration and Related Characteristics-a Case Study of North-Eastern Bihar” conducted during October 2009 to June 2010 (flooded area of Koshi river), Varanasi data collected under a sample survey “Rural development and population growth (RDPG) survey” conducted in 1978 in Varanasi district and used by Sharma [5] and Iwunor [12], Nepal [14] and Comilla district of Bangladesh [13]. It is shown in Tables 1, 2, 3 and 4 respectively.

Tables 1, 2, 3 and 4 provides the fitting, parameter estimates and χ^2 with *p*-value for every set of data for which all models have been applied. Table 1 show that the proposed model I fit the data excellently whereas model II and III comparatively less efficient for the data considered here in this study. Estimate of first parameter is same for all models, since this parameter α is the proportion of the household with no adult out migration. The second parameter for model I, II and III are 1.1671, 2.3342 and 1.6095 respectively. The values of χ^2 with degrees of freedom and *p*-value are also given in the tables. From Table 3, it is obvious that all the three models perform well for data of Nepal adult out migration and explains it very well as the chi square values are very less and *p*-value for model I and III is more than 0.9 percent and about 0.35 percent for model II. The value of χ^2 and its *p*-value shown in the tables clearly indicate that the model I is best, model III is worse than model I but better than model II and model II is worst among all three models. Comparing these models to each other, model I perform better as *p*-value is highest.

Table 1. Observed and expected frequency of the number of households according to the adult migrants in flooded area of Kosi river

Number of migrants	Observed number of households	Model I	Model II	Model III
0	401	401.00	401.00	401.00
1	147	141.65	128.90	138.40
2	57	65.36	77.32	67.73
3	29	30.16	34.79	31.59
4	16	13.92	13.91	14.26
5	8	6.42	5.22	6.29
6	5	2.96	1.88	2.73
7	1	2.53	0.99	2.13
Total	664	664.00	664.00	664.00
Mean= 0.7365		χ^2=1.99 (after pooling) *p*-value=0.5737 (d.f.=3)	χ^2=13.49 (after pooling) *p*-value=0.004 (d.f.=3)	χ^2=3.38 (after pooling) *p*-value=0.3356 (d.f.=3)
Estimated value of parameters		α=0.3961 θ=1.1671	α=0.3961 θ=2.3342	α=0.3961 θ=1.6095

Table 2. Observed and expected frequency of the number of households according to the adult migrants in Varanasi District

Number of migrants	Observed number of households	Model I	Model II	Model III
0	1032	1032.00	1032.00	1032.00
1	95	88.51	85.45	88.07
2	19	27.78	31.81	28.28
3	10	8.72	8.88	8.81
4	2	2.74	2.20	2.68
5	2	0.86	0.51	0.80
6	0	0.27	0.11	0.24
7	1	0.14	0.03	0.11
Total	1161	1161.00	1161.00	1161.00
Mean= 0.1619		χ^2=3.66 (after pooling) *p*-value=0.056 (d.f.=1)	χ^2=7.13 (after pooling) *p*-value=0.008 (d.f.=1)	χ^2=4.02 (after pooling) *p*-value=0.044 (d.f.=1)
Estimated value of parameters		α=0.1111 θ=2.1864	α=0.1111 θ=4.3728	α=0.1111 θ=2.7669

Data source: Varanasi (1978)

Table 3. Observed and expected frequency of the number of households according to adult migrants in Nepal

Number of migrants	Observed number of households	Model I	Model II	Model III
0	623	623.00	623.00	623.00
1	126	125.78	120.83	125.02
2	42	41.63	47.92	42.47
3	13	13.78	14.25	13.97
4	4	4.56	3.77	4.49
5	2	1.51	0.93	1.42
6	1	0.75	0.29	0.64
Total	811	811.00	811.00	811.00

Mean= 0.3465	χ^2 =0.0526 (after pooling) *p*-value=0.9740 (d.f.=2)	χ^2 =1.87 (after pooling) *p*-value=0.392 (d.f.=2)	χ^2 =0.11 (after pooling) *p*-value=0.9452 (d.f.=2)
Estimated value of parameters	α =0.2318 θ =2.0215	α =0.2318 θ =4.0430	α =0.2318 θ =2.5853

Data source: Aryal (2002)

Table 4. Observed and expected frequency of the number of households according to adult migrants in Comilla district of Bangladesh

Number of migrants	Observed number of households	Model I	Model II	Model III
0	1941	1941.00	1941.00	1941.00
1	542	529.47	515.82	527.82
2	124	146.49	165.63	148.56
3	48	40.53	39.89	40.74
4	13	11.21	8.54	10.96
5	4	3.10	1.71	2.90
6	1	1.19	0.40	1.02
Total	2673	2673.00	2673.00	2673.00
Mean= 0.3786		χ^2 =5.53 (after pooling) *p*-value=0.063 (d.f.=2)	χ^2 =18.495 (after pooling) *p*-value=0.000 (d.f.=2)	χ^2 =6.38 (after pooling) *p*-value=0.041 (d.f.=2)
Estimated value of parameters		α =0.2739 θ =2.6143	α =0.2739 θ =5.2286	α =0.2739 θ =3.23203

Data source: Hossain [12]

5. CONCLUSIONS

This paper represent the probability models applied to explain the pattern of adult out migration and explains the distribution of households according to number of adult migrants (age fifteen years and above) in different sets of the observed data. The suitability of proposed model is tested with primary data from four different time and space. On the basis χ^2 and its *p*-value we can reach at the conclusion that model I is found best among the considered models here for all real data sets on adult out migration.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Singh SN, Yadava KNS. Trends in rural out-migration at household level. Rural Demography. 1981;8(1):53-56.
2. Yadava KNS and Singh SRJ. A model for the number of rural out-migrants at household level. Rural Demogr. 1983; 10(1-2):23-33
3. Singh SN, Yadava RC, Sharma HL. A model for rural out-migration at household level. Janasamkhya. 1985;3(1-2):1-7.
4. Sharma L. A study of the pattern of out migration from rural areas. Unpublished Ph.D. Thesis in Statistics, Banaras Hindu University, Varanasi, India; 1984.
5. Sharma HL. A probability distribution for out-migration. Janasamkhya. 1987;5(2):95-101.
6. Yadava KNS, Yadava GS. On some probability models and their applications to the distribution of the number of migrants from a household. Janasankhya. 1988;6(2):137-158.
7. Yadava KNS, Yadava GS, Singh SK. Pattern of return migration in rural areas of Eastern Uttar Pradesh, India. Journal of Institute of Economic Research. 1989;24(2):27-48.
8. Yadava KNS, Tripathi S, Singh VS. A model for the number of out migrants at the household level: An alternative approach. The Journal of Scientific Research. 1994;44:125-134.

9. Aryal TR. Probability models for the number of rural out-migrants at micro-level. The Nepali Mathematical Sciences Report. 2003;21(1&2):9-18.
10. Singh Brijesh P, Singh Niraj K. On the distribution of risk of migration and its estimation. International Journal of Statistical Distributions and Applications. 2016;2(4):67-71.
11. Singh SN, Yadava RC, Sharma HL. A model for rural out-migration at household level. Janasamkhya. 1985; 3(1-2):1-7.
12. Iwunor CC. Estimation of parameters of the inflated geometric distribution for rural out-migration. Genus. 1995;51(3/4):253-260.
13. Hossain M. Some demographic models and their applications with special reference to Bangladesh. An Unpublished Thesis Submitted in Banaras Hindu University, Varanasi, India; 2000.
14. Aryal TR. Inflated geometric distribution to study the distribution of rural out-migrants. Journal of the Institute of Engineering. 2011;8(1-2):266-268.
15. Singh Brijesh P, Singh Niraj K, Roy Tapan K. On the pattern of migration in the household: An explanation through binomial law. International Journal of Statistics and Systems. 2014;9:2.
16. Singh Brijesh P, Singh Niraj K, Dixit Shweta. Estimation of parameters of inflated Poisson-Lindley distribution for adult out migration. Journal of Institute of Science & Technology. 2015;20(2):6-10.
17. Singh Brijesh P, Singh S, Das UD. Some zero adjusted probability model for adult out migration. Bulletin of Mathematics and Statistics Research. 2020;8(3):14-19.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/63311