Inflated Geometric Distribution to Study the Distribution of Rural Out-Migrants

Tika Ram Aryal
Central Department of Statistics, Tribhuvan University, Nepal
Corresponding e-mail: traryal@gmail.com & traryal@rediffmail.com

Abstract: The aim of this paper is to apply inflated geometric distribution to study the distribution of rural out-migrants from a household. The inflated geometric distribution is found to be a reasonable approximation to describe this distribution. The estimated value of risk parameter $\alpha$ was found to be 0.7004229. Findings may help planners and policy makers for designing more effective and equitable rural and urban policies.

1. Introduction
Researchers have given due attention 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). Several attempts have been made to study the pattern of rural out-migration by using probability models (Aryal, 2002, Iwunor, 1995; Sharma, 1992; Yadava and Singh, 1983). Singh and Yadava (1981) applied a negative binomial distribution to describe the trend of rural out-migration. Yadava and Yadava (1988) extended the idea of a cluster and assumed that the occurrence of migration in a cluster varies from household to household. Singh (1985) developed a probability model for the total number of migrants under the assumption that there are two types of households, i.e., the households from where male members aged 15 years and above migrate singly leaving their wives and children at home, and the households where male members migrate with their wives, children and other dependent relatives.

In light of this, this paper tries to apply an inflated geometric distribution to the study of the distribution of rural out-migrants from a household level.

2. Inflated Geometric Distribution
The following assumptions are made in order to apply an inflated geometric distribution to describe the distribution of rural out-migrants from a household.

Assumption I: At any point in time, let $\alpha$ be the probability of migrating out from a household and $1-\alpha$ be the probability of not migrating from a household at survey pint.

Assumption II: If $p$ represents the probability of a single individual migrating from a household; the pattern of migration from each household follows the geometric distribution.

Let $X$ denote the number of rural out-migrants from a household while $X$ follows the inflated geometric distribution with probability density function is given below:
\[
\begin{align*}
P(X=0) &= 1 - \alpha + \alpha p \\
P(X=k) &= \alpha q^k p \text{ for } k = 1, 2, 3, \ldots \\
\end{align*}
\] (1)

provided that \( p + q = 1 \).

**Estimation of Parameters**

Let \((X_1, X_2, \ldots, X_n)\) denote a random sample of size \( n \). Each \( X_i \) counts the number of rural out-migrants from a household. Assuming that \( n_k \) \((k=0, 1, 2, \ldots, m)\) denotes the number of observations with value \( k \). The likelihood function for estimating the parameters \( \alpha \) and \( p \) can be expressed as below:

\[
L = (1 - \alpha + \alpha p)^n_0 \prod_{k=1}^{m} (\alpha p q^k)^{n_k} = (1 - \alpha + \alpha p)^n_0 \alpha n - n_0 p n - n_0 q^n
\] (2)

where, \( S = n_1 + 2n_2 + 3n_3 + \ldots + mn_m = \sum_{k=1}^{m} n_k k \) and \( n_0 + n_1 + n_2 + \ldots + n_m = n \).

Taking logarithms in Equation (2) and differentiating with respect to \( \alpha \) and \( p \) respectively and equating to zero, we get:

\[
\frac{\delta \log L}{\delta \alpha} = \frac{n_0(p-1)}{(1 - \alpha + \alpha p)} + \frac{n - n_0}{\alpha} = 0 \quad \text{and} \quad \frac{\delta \log L}{\delta p} = \frac{n - n_0}{p} - \frac{s}{1 - p} + \frac{n_0 \alpha}{(1 - \alpha + \alpha p)} = 0.
\]

On solving these two equations, we easily estimate \( \hat{\alpha} = \frac{n - n_0}{n(1 - p)} \) and \( \hat{p} = \frac{n - n_0}{\sum_{k=1}^{m} n_k k} \).

2.1 Application

The data is taken from a sample survey of the Rupandehi and Palpa districts in Nepal. Table 1 shows estimated values of the parameters, and the observed and expected number of households according to the number of out-migrants at a household level in Nepal. The estimated value of risk parameter \( \alpha \) was found to be 0.7004229 while the estimated value of \( p \) was found to be 0.6090391. The \( \chi^2 \) value was found to be 0.05, which is insignificant at a 5 per cent level of significance, which indicates that the validity of the applied inflated geometric distribution was found to be a reasonable approximation.

**Table 1** Observed and expected number of household according to the number of out-migrants

<table>
<thead>
<tr>
<th>No. of migrants per household</th>
<th>Observed</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>623</td>
<td>623.00</td>
</tr>
<tr>
<td>1</td>
<td>126</td>
<td>125.80</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>41.63</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>13.78</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>6.79</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
3. Conclusions

The inflated geometric distribution is found to be a reasonable approximation to describe the distribution of out-migrants from a household. The estimated value of risk parameter $\alpha$ was found to be 0.7004229 while the estimated value of $p$ was found to be 0.6690391. Findings may help planners and policy makers for designing more effective and equitable rural and urban policies.

REFERENCES