Vol.10.Issue.4.2022 (Oct-Dec)

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

http://www.bomsr.com

Email:editorbomsr@gmail.com

INTERNATIONAL STANDARD SERIAL NUMBER 2348-0580

RESEARCH ARTICLE

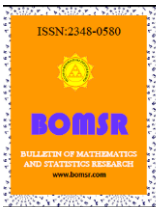
HIMANSHU DISTRIBUTION AND ITS APPLICATIONS

ABHISHEK AGARWAL[1], HIMANSHU PANDEY[2]

[1,2]Department of Mathematics And Statistics,D.D.U Gorakhpur University
Gorakhpur, U.P., India

Email: abhishek.agarwal014@gmail.com; himanshu_pandey62@yahoo.com

DOI:10.33329/bomsr.10.4.24



ABSTRACT

In this paper, a new discrete distribution named "Himanshu distribution" for modelling the real life data sets from Biomedical, Demography has been suggested. Some important Mathematical Properties of the proposed distribution including moments, coefficient of variation, index of dispersion, skewness and kurtosis have been discussed. The Maximum likelihood estimation and Method of moments for estimating the parameters have been discussed. The proposed distribution fitted to observed sets of data for satisfying its suitability and draw some conclusion.

Keywords - Discrete distribution, Moments, Skewness, Kurtosis, Estimation of parameters.

1. Introduction

A number of discrete distribution for modelling and analysing the real life data such as Geometric distribution, Poission distribution, Hyper geometric distribution, Negative Binomial distribution etc. are available in the statistical literature.

Recently the number of distributions such as Shanker distribution (2015), Akash distribution (2015), Sujatha distribution (2016), Suja distribution (2017), Ishita distribution (2017), Pranav distribution (2018) in continuous nature and Poission Lindley distribution (2015), Poission Sujatha distribution (2016), Poission Akash distribution (2017) in discrete nature have been developed.

In this paper a new distribution in discrete nature having its probability mass function -

$$P(X=x)=p^n(1-p^n)^x \quad ; \qquad \left.\begin{array}{l} x=0,1,2,\ldots \\ n\in I^+ \\ 0<p<1 \end{array}\right\} \qquad (1.1)$$

has been suggested for modelling and analysing various type of real life data sets. We would call this distribution " Himanshu distribution". Equation (1.1) can be shown as generalised form of Geometric distribution.

2. Properties Of Proposed Distribution

(i) Moment Generating Function – M.G.F of the distribution can be given as –

$$M_X(t)=\sum_{x\in R(x)} e^{tx}\,p(x)$$

$$=\sum_{x=0}^{\infty} e^{tx}p^n(1-p^n)^x$$

$$M_X(t)=\frac{p^n}{1-e^t(1-p^n)}$$

(ii) Probability Generating Function – P.G.F of the distribution can be given as –

$$P_X(s)=\frac{p^n}{1-s(1-p^n)}$$

(iii) Characteristic function - Characteristic function of the distribution can be given as-

$$\varphi_X(t)=\sum_{x\in R(x)} e^{itx}\,p(x)$$

$$=\sum_{x=0}^{\infty} e^{itx}p^n(1-p^n)^x$$

$$\varphi_X(t)=\frac{p^n}{1-e^{it}(1-p^n)}$$

(iv) Moments – The r^{th} moment about origin of Himanshu distribution can be obtained as –

$$\mu_r'=\sum_{x=0}^{\infty} x^r p(x)$$

$$=\sum_{x=0}^{\infty} x^r p^n(1-p^n)^x$$

$$\mu_r'=p^n\{1^r(1-p^n)+2^r(1-p^n)^2+3^r(1-p^n)^3+\cdots\}$$

By putting r = 1, 2, 3, 4 in above equation we get

$$\text{Mean} \;=\mu_1'=\frac{1-p^n}{p^n}, \qquad \mu_2'=\frac{(1-p^n)(2-p^n)}{p^{2n}}$$

$$\mu_3'=\frac{(1-p^n)(p^{2n}-6p^n+6)}{p^{3n}}, \qquad {\mu'}_4=\frac{-(1-p^n)(p^{3n}-14p^{2n}+36p^n-24)}{p^{4n}}$$

Now central moments of Himanshu distribution can be given as –

$$\mu_2 = \frac{1-p^n}{p^{2n}}, \mu_3 = \frac{(1-p^n)(2-p^n)}{p^{3n}} \ and \ \mu_4 = \frac{(1-p^n)(9-9p^n+p^{2n})}{p^{4n}}$$

(v) **Median -** Median of the distribution can be obtained as –

$$\sum_{x=0}^{M} p^n(1-p^n)^x = \frac{1}{2}$$

$$1-(1-p^n)^{M+1} = \frac{1}{2}$$

$$\Rightarrow \ M = \left[\frac{-1}{\log_2(1-p^n)}\right] - 1$$

(vi) **Mode –** Mode of the Himanshu distribution can be obtained as –

$$P(X=x) = p^n(1-p^n)^x \ ; \qquad \begin{array}{l} x = 0,1,2,\dots \\ 0 < p < 1 \\ n \in I^+ \end{array}$$

$$P(X=0) = p^n$$

$$P(X=1) = p^n(1-p^n)$$

$$P(X=2) = p^n(1-p^n)^2$$

$$\vdots$$

$$\text{Hence X} = 0 \text{ is mode}$$

(vii) Now coefficient of skewness can be given as –

$$\gamma_1 = \sqrt{\beta_1} = \frac{2-p^n}{\sqrt{1-p^n}} > 0$$

⇒ Distribution is positively skewed

(viii) Coefficient of kurtosis of Himanshu distribution can be obtained as –

$$\beta_2 = \frac{9-9p^n+p^{2n}}{1-p^n} \text{and } \gamma_2 = \frac{6(1-p^n)+p^{2n}}{1-p^n} > 0$$

⇒ Distribution is Leptokurtic

(ix) Coefficient of variation and Index of dispersion can be obtained as –

$$C.V = \frac{\sigma}{\mu_1'} = \frac{1}{\sqrt{1-p^n}}$$

$$\text{Index of dispersion} = \frac{\sigma^2}{\mu_1'} = \frac{1}{p^n}$$

3. Parameter Estimation

The parameter of the proposed distribution can be estimated by method of moments and method of maximum likelihood.

Method of moments – Let $x_1, x_2, \dots, x_k$ be a random sample of size k then

$$\bar{x} = \mu_1'$$

$$\Rightarrow \hat{p} = \left(\frac{1}{\bar{x}+1}\right)^{\frac{1}{n}}$$

Method Of Maximum Likelihood – Let $x_1, x_2, \dots, x_k$ be a random sample of size k, then its likelihood function can be defined as –

$$L = \prod_{i=1}^{k} f(x_i\,; p)$$

$$L = p^{nk}(1-p^n)^{\sum_{i=1}^{k} x_i}$$

$$\text{Now} \frac{\partial}{\partial p} \log L = 0$$

$$\Rightarrow \hat{p} = \left(\frac{1}{\bar{x}+1}\right)^{\frac{1}{n}}$$

Particular Cases of Proposed Distribution – In the proposed Himanshu distribution if we put n=1, 2, 3,... we get different type of discrete distributions including Geometric distribution.

Case 1. If we put n = 2 then we have the p.m.f in the following form –

$$P(X=x) = p^2(1-p^2)^x \ ; \qquad \begin{matrix} x = 0,1,2,\dots \\ 0 < p < 1 \\ n \in I^+ \end{matrix}$$

with properties –

(i) M.G.F –

$$M_X(t) = \frac{p^2}{1-e^t(1-p^2)}$$

(ii) P.G.F –

$$P_X(s) = \frac{p^2}{1-s(1-p^2)}$$

(iii) Characteristic Function –

$$\varphi_X(t) = \frac{p^2}{1-e^{it}(1-p^2)}$$

(iv) Moments –

$$\mu_1' = \frac{1-p^2}{p^2} \qquad \mu_2' = \frac{(1-p^2)(2-p^2)}{p^4}$$

$$\mu_3' = \frac{(1-p^2)(p^4-6p^2+6)}{p^6}, \qquad \mu_4' = \frac{-(1-p^2)(p^6-14p^4+36p^2-24)}{p^8}$$

and central moments are –

$$\mu_1 = 0 \qquad \mu_2 = \frac{1-p^2}{p^4}$$

$$\mu_3 = \frac{(1-p^2)(2-p^2)}{p^6} \ and \ \mu_4 = \frac{(1-p^2)(9-9p^2+p^4)}{p^8}$$

(v) Median –

$$M = \left[\frac{-1}{\log_2(1-p^2)}\right] - 1$$

(vi) Mode –

$$P(X = x) = p^2(1-p^2)^x \ ; \qquad \begin{array}{l} x = 0,1,2,\dots \\ 0 < p < 1 \\ n \in I^+ \end{array}$$

$$P(X = 0) = p^2$$

$$P(X = 1) = p^2(1-p^2)$$

$$P(X = 2) = p^2(1-p^2)^2$$

$$\vdots$$

Hence $\mathrm{X} = 0$ is mode

(vii)$\gamma_1 = \sqrt{\beta_1} = \frac{2-p^2}{\sqrt{1-p^2}} > 0$

⇒ Distribution is positively skewed

(viii)$\beta_2 = \frac{9-9p^2+p^4}{1-p^2}$ and $\gamma_2 = \frac{6(1-p^2)+p^4}{1-p^2}$

(ix) $\mathrm{C.V} = \frac{1}{\sqrt{1-p^2}}$ and Index of dispersion $= \frac{1}{p^2}$

Parameter Estimation –

Method of moment estimator can be given by –

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

Maximum Likelihood estimator can be given by –

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

Case 2. It was put n=3 then we have the p.m.f in the following form –

$$P(X = x) = p^3(1-p^3)^x \ ; \qquad \begin{array}{l} x = 0,1,2,\dots \\ 0 < p < 1 \\ n \in I^+ \end{array}$$

with properties –

(i) M.G.F –

$$M_X(t) = \frac{p^3}{1 - e^t(1 - p^3)}$$

(ii) P.G.F –

$$P_X(s) = \frac{p^3}{1 - s(1 - p^3)}$$

(iii) Characteristic Function –

$$\varphi_X(t) = \frac{p^3}{1 - e^{it}(1 - p^3)}$$

(iv) Moments –

$$\mu_1' = \frac{1 - p^3}{p^3} \qquad \mu_2' = \frac{(1 - p^3)(2 - p^3)}{p^6}$$

$$\mu_3' = \frac{(1 - p^3)(p^6 - 6p^3 + 6)}{p^9} \qquad \mu_4' = \frac{-(1 - p^3)(p^9 - 14p^6 + 36p^3 - 24)}{p^{12}}$$

and central moments are –

$$\mu_2 = \frac{1 - p^3}{p^6}, \qquad \mu_3 = \frac{(1 - p^3)(2 - p^3)}{p^9}$$

$$\mu_4 = \frac{(1 - p^3)(9 - 9p^3 + p^6)}{p^{12}}$$

(v) Median –

$$M = \left[\frac{-1}{\log_2(1 - p^3)}\right] - 1$$

(vi) Mode –

$$\text{X} = 0 \text{ is always mode}$$

(vii)

$$\gamma_1 = \sqrt{\beta_1} = \frac{2 - p^3}{\sqrt{1 - p^3}} > 0$$

⇒ Distribution is positively skewed

(viii)

$$\beta_2 = \frac{9 - 9p^3 + p^6}{1 - p^3} \text{ and } \gamma_2 = \frac{6(1 - p^3) + p^6}{1 - p^3} > 0$$

⇒ Distribution is Leptokurtic

(ix)

$$\text{C.V} = \frac{1}{\sqrt{1 - p^3}} \text{ and Index of dispersion} = \frac{1}{p^3}$$

Parameter Estimation –

Method of Moment estimator can be given by –

$$\hat{p} = \left(\frac{1}{\bar{x} + 1}\right)^{\frac{1}{3}}$$

Maximum likelihood estimator can be given by –

$$\hat{p} = \left(\frac{1}{\bar{x} + 1}\right)^{\frac{1}{3}}$$

4. Application

The proposed distribution after assuming n=2 have been applied to the survey data on migration pattern taken from various surveys entitled Migration and related characteristics –"A case study of North Eastern Bihar 2009-2010" taken from Dubey etal(2022), Sample survey of Rupandehi and Palpa district in Nepal - 2011" used by Aryal (2011) and sample survey "Impact of Migration on fertility in Bangladesh: A study of Comilla district – 1997" used by Hossain (2000) and biological data reported as genetics count data sets of Catchesideetal taken from Hassan(2020).

Table 1. Observed and Expected frequency of the number of households according to the migrants in flooded area of Kosi River

Number of migrants	Observed no. of households	Expected no. of households
0	401	382.39
1	147	162.13
2	57	68.74
3	29	29.14
4	16	12.35
5	8	5.24
6	$\left.\begin{matrix}5\\1\end{matrix}\right\}6$	4.01
7		
Total	**664**	**664**
Mean=0.7364 Variance=1.2789 $\hat{p} = 0.7588$	$\chi^2 = 7.83$ (after pooling) p-value=0.1658 $\chi^2_{(5)} = 11.07$ at 5% level of significance	

Table 2. Observed and Expected frequency of the number of households according to migrants in Nepal

No. of migrants	Observed no. of households	Expected no. of households
0	623	602.49
1	126	154.83
2	42	39.79
3	13	10.22
4	$\left.\begin{matrix}4\\2\\1\end{matrix}\right\}7$	3.67
5		
6		
Total	**811**	**811**
Mean=0.3465 Variance=0.4667 $\hat{p} = 0.8617$	$\chi^2 = 9.93$ (after pooling) p-value=0.0191 $\chi^2_{(3)} = 11.345$ at 1% level of significance	

Table 3. Observed and Expected frequency of the number of households according to the migrants in Comilla district of Bangladesh

No. of migrants	Observed no. of households	Expected no. of households
0	1941	1938.72
1	542	532.37
2	124	146.19
3	48	40.14
4	$\left.\begin{matrix}13\\4\\1\end{matrix}\right\}18$	15.58
5		
6		
Total	**2673**	**2673**
Mean=0.3786 Variance=0.5219 $\hat{p} = 0.8516$	$\chi^2 = 5.44$ (after pooling) p-value=0.1422 $\chi^2_{(3)} = 7.815$ at 5% level of significance	

Table 4. Distribution of Number of chromatidabberrations (0.2g chinonl, 24 hours)		
No. of Aberrations	**Observed frequency**	**Expected frequency**
0	268	258.48
1	87	91.45
2	26	32.35
3	9	11.44
4	$\left.\begin{matrix}4\\2\\1\\3\end{matrix}\right\}10$	6.28
5		
6		
7		
Total	**400**	**400**
Mean=0.5475 Variance=0.8510 $\hat{p} = 0.8030$	$\chi^2 = 4.52$ (after pooling) p-value=0.2105 $\chi^2_{(3)} = 7.815$ at 5% level of significance	

Table 5. Mammalian cytogenetic dosimetry lesions in Rabbit lymphoblast induced by streptonigrin (NSC-45383) Exposure-70 µg/Kg		
Class/Exposure µg/Kg	**Observed frequency**	**Expected frequency**
0	200	193.11
1	57	68.8
2	30	24.5
3	7	8.7
4	$\left.\begin{matrix}4\\0\\2\end{matrix}\right\}6$	4.85
5		
6		
Total	**300**	**300**
Mean=0.5533 Variance=0.8594 $\hat{p} = 0.8023$	$\chi^2 = 4.09$ (after pooling) p-value=0.2519 $\chi^2_{(3)} = 7.815$ at 5% level of significance	

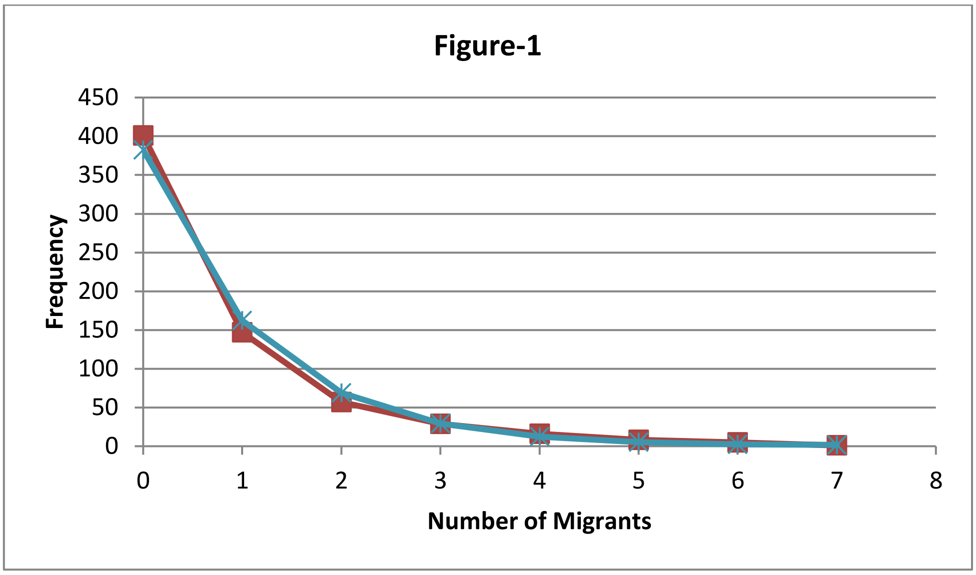


Figure-1
Frequency
450
400
350
300
250
200
150
100
50
0
0 1 2 3 4 5 6 7 8
Number of Migrants

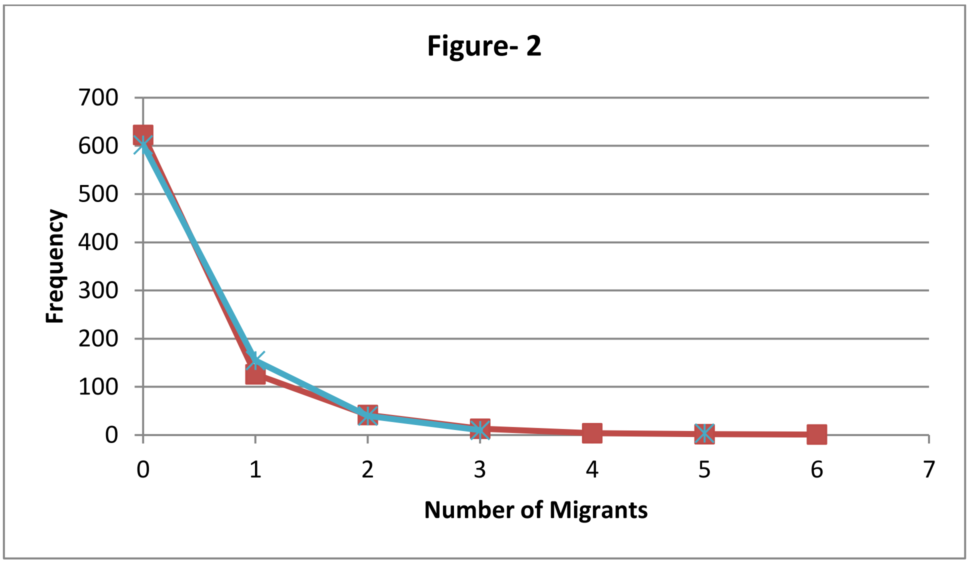


Figure- 2
Frequency
700
600
500
400
300
200
100
0
0 1 2 3 4 5 6 7
Number of Migrants

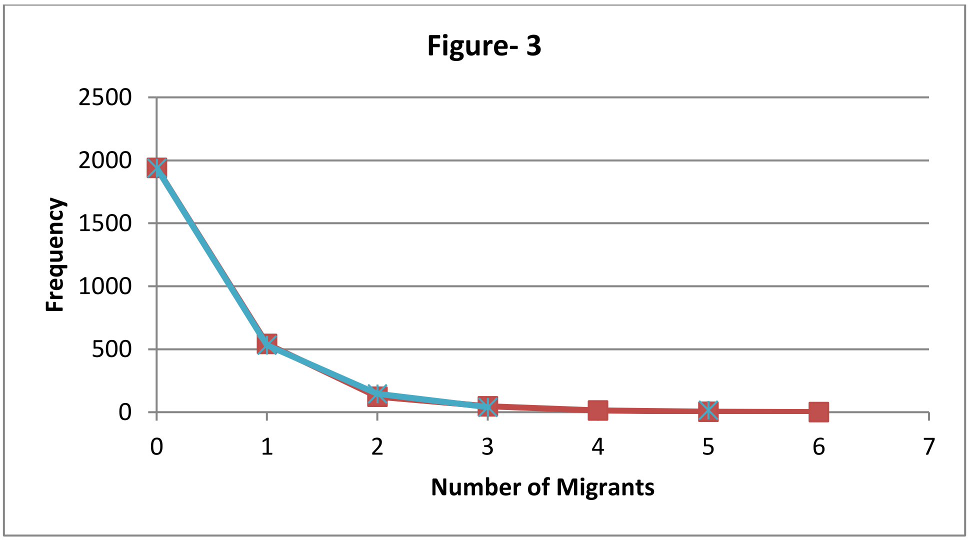
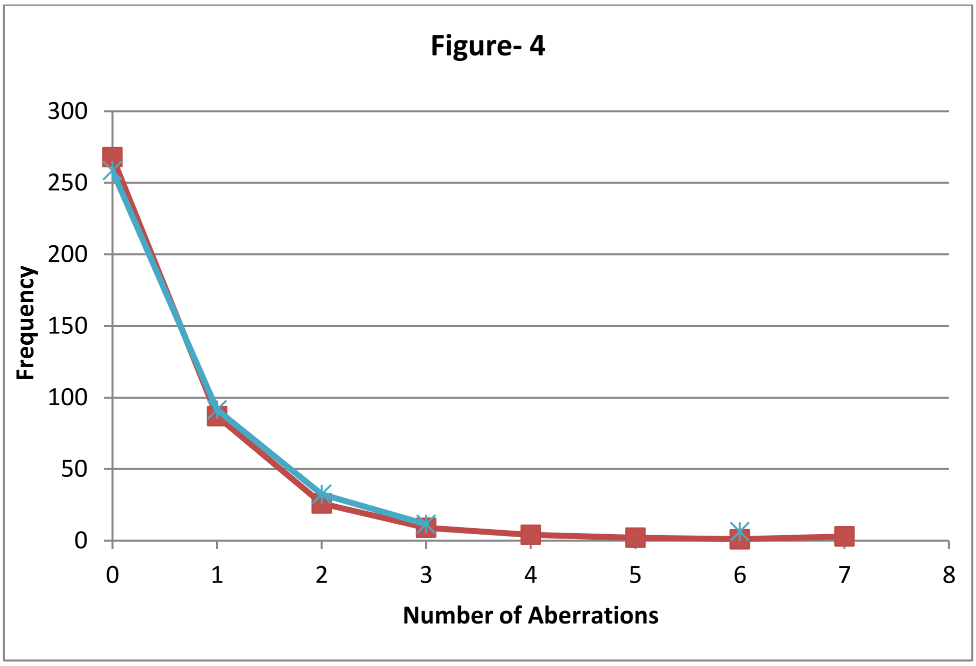
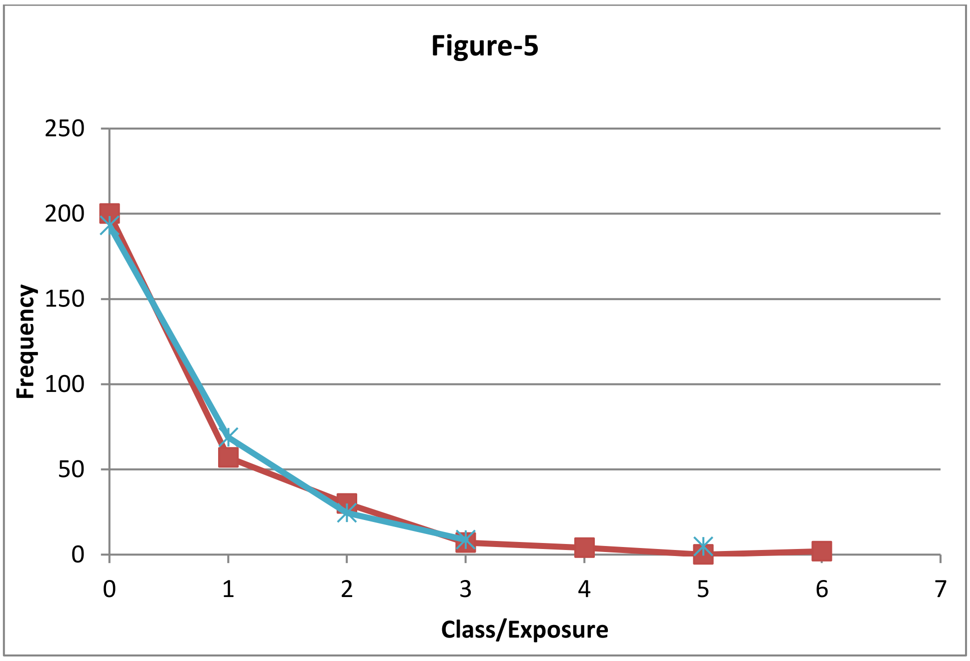


Figure- 3
Frequency
2500
2000
1500
1000
500
0
0 1 2 3 4 5 6 7
Number of Migrants





5. Conclusion

The study with Respect to above tables and graphs shows that the proposed distribution is a good approximation for migration pattern in three different countries as well as biological science data. The estimated value of parameter, the value of chi-square with degree of freedom and p-values are also given in table. The value of chi-square shown in the tables clearly indicate the proposed distribution describes the distribution of Migration pattern of Adult migrants from households and as well as for Genetics count data. We observe that on the basis of χ^2 and p-value the agreement between observed and expected frequencies of proposed distribution is fairly good. The importance of this study is that whenever the same conditions prevails in any part of the society, we should not do same exercise again. We simply apply the proposed distribution as a model and get the desired estimates. It not only saves our time but also saves money and other resources.

The author believes that the proposed distribution as a model used for further studies related with the real life problems and it could be helpful in policy making, rural development, medical and environmental conditions of the society.

References

[1]. Aryal TR. Inflated Geometric Distribution to Study the Distribution of Rural Out-migrants. Journal of the Institute of Engineering. 2011;8(1):266-268.

[2]. Dubey A., Pandey H., An inflated probability models for adult out migration, International Journal of Statistics and Applied Mathematics.2022;7(5):80-82.

[3]. Hassan A.,BilalS.,RashidA.,Ahmad G.,A new flexible discrete distribution with applications to count data,Journal of Statistical Theory of Applications.2020;19(1):102-108.

[4]. Hossain MZ. Some demographic models and their application with special reference to Bangladesh. Unpublished Ph.D. thesis, Banaras Hindu University, India; c2000.

[5]. Sankaran M. The discrete Poisson-Lindley distribution, Biometrics. 1970; 26:145-149.

[6]. Shankar, R., (2015), Shanker Distribution and its Applications, International Journal of Statistic and Applications, 5(6):338-348

[7]. Shanker R. & Shukla, K. K.Ishita distribution and Its Applications,BBIJ.2017: 5(2):1-9.

[8]. Shanker R. Akash distribution and its Applications, International Journal of Probability and Statistics. 2015; 4(3):65-75.

[9]. Shanker R. Suja distribution and its Applications, International Journal of Probability and Statistics. 2017; 6(2):11-19.

[10]. Shanker R. The discrete Poisson-Akash distribution, International Journal of Probability and Statistics. 2017a; 5(1):1-9.

[11]. Shanker R.(2016h): The discrete Poisson-Sujatha distribution, International Journal of Probability and Statistics, 5(1):1-9

[12]. Shanker, R. (2016), Sujatha distribution. Statistics In Transition, 17(3):391-410.

[13]. Shukla, K.K.Pranav distribution with properties and Its Applications,BBIJ.2018: 7(3):244-254.