Publications

Available Online

J. Sci. Res. **16** (1), 89-96 (2024)

JOURNAL OF SCIENTIFIC RESEARCH

www.banglajol.info/index.php/JSR

An Inflated Modelling of Zero Truncated Poisson Ailamujia Distribution and Its Application to Child Mortality and Genetics Count Data

A. Agarwal*, H. Pandey

Department of Mathematics and Statistics, D. D. U Gorakhpur University, Gorakhpur, U.P, India

Received 31 March 2023, accepted in final revised form 15 September 2023

Abstract

In this paper, an attempt has been made to develop a model as Inflated form of zero truncated Poisson Ailamujia distribution. The proposed model involved two parameters, estimated by the maximum likelihood estimation method and derived as an expression of the Fisher Information Matrix. To evaluate the adequacy and applicability of the proposed model, it is applied to real data sets on child mortality and genetic count data.

Keywords: Poisson ailamujia distribution; Inflated distribution; Method of maximum likelihood; Fisher information matrix; Child mortality; Genetics count data.

doi: http://dx.doi.org/10.3329/jsr.v16i1.65361 J. Sci. Res. **16** (1), 89-96 (2024)

1. Introduction

In recent years, a number of discrete distributions have been presented for the actual lifespan data in probability distribution theory to disclose diminishing or growing failure rate functions. Pandey and Kishun [1], Pandey and Tiwari [2], Pandey *et al.* [3], and Agarwal and Pandey [4,5] have introduced various discrete distributions for different types of vital events.

A random variable X has a Poisson distribution, and its probability mass function is-

$$P(\lambda) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}; \begin{matrix} x = 0, 1, 2, \dots \\ \lambda > 0 \end{matrix} \tag{1.1}$$

whereλ is the parameter of the distribution.

LV *et al.* [6] have formulated Ailamujia distribution with probability density function as follows-

$$AD(\alpha) = f(X = x, \alpha) = 4x\,\alpha^2 e^{-2\alpha x}; \begin{matrix} x \geq 0 \\ \alpha > 0 \end{matrix} \tag{1.2}$$

Then Hassan *et al.* [7] formulated the probability mass function of the compounded form of $P(\lambda), A(\alpha)$is the following way–

$$p(X = x, \alpha) = \frac{4\alpha^2(1+x)}{(1+2\alpha)^{x+2}}; \quad \begin{matrix} x = 0, 1, 2, \dots \\ \alpha > 0 \end{matrix} \tag{1.3}$$

* *Corresponding author*: abhishek.agarwal014@gmail.com

A zero truncated distribution with discrete distributional characteristics supports the set of positive integers. These distributions can be used extensively when dealing with real-world data generated via a technique that excludes zero counts. Mathematically, it is defined as:

Let $p_0(x;\theta)$ is the original distribution with support for non-negative positive integers, then the zero truncated form of $p_0(x;\theta)$ with a support set of positive integers given by

$$p(x;\theta) = \frac{p_0(x;\theta)}{1-p_0(0;\theta)};\ x = 1,2,3,\ldots \tag{1.4}$$

In this way, Agarwal and Pandey[8] have suggested a zero truncated model of Poisson Ailamujia distribution (1.3) in the following way-

$$p(x;\alpha) = \frac{\frac{4\alpha^2(1+x)}{(1+2\alpha)^{x+2}}}{\left\{1-\frac{4\alpha^2}{(1+2\alpha)^2}\right\}}$$

$$\Rightarrow p(x;\alpha) = \frac{4\alpha^2(1+x)}{(1+4\alpha)(1+2\alpha)^x};\ x = 1,2,3,\ldots \tag{1.5}$$

2. Proposed Probability Model

When dealing with real-world data, it sometimes happens that conventional distribution-based models cannot manage a large number of zeros and must be extended to an inflated form. The application of the inflated Poisson and Negative Binomial distribution models has been covered extensively in many literatures.

Let x be the number of events that occurred at the survey area, with assumption $\beta$and $(1-\beta)$be the probability of being exposed to the risk of the event and not exposed to the risk of the event, respectively.

Let X follow zero truncated Poisson Ailamujia distribution (1.5), and incorporating the above assumptions, we get an inflated model of (1.5) in the following form –

$$p(X=x) = \begin{cases} 1-\beta\ldots\ldots\ldots\ldots\ldots\ldots\ldots;x=0 \\ \beta.\ \frac{4\alpha^2}{(1+4\alpha)}\left\{\frac{1+x}{(1+2\alpha)^x}\right\} \quad ;\ x=1,2,3,\ldots \end{cases} \tag{2.1}$$

Where α and β are the parameters.

with Mean $= \beta.\frac{(1+2\alpha)^2}{\alpha(1+4\alpha)}$

3. Estimation of Parameters and Fisher Information Matrix

The parameters of the proposed Inflated probability model are estimated by the method of maximum likelihood in the following way –

$$L = (1-\beta)^{f_0}\left\{\frac{8\alpha^2\beta}{(1+2\alpha)(1+4\alpha)}\right\}^{f_1}\left[1-\left\{(1-\beta)+\frac{8\alpha^2\beta}{(1+2\alpha)(1+4\alpha)}\right\}\right]^{f-f_0-f_1}$$

$$\log L = f_0\log(1-\beta) + f_1\log\left\{\tfrac{8\alpha^2\beta}{(1+2\alpha)(1+4\alpha)}\right\} + (f-f_0-f_1)\log\left\{\beta-\tfrac{8\alpha^2\beta}{(1+2\alpha)(1+4\alpha)}\right\} \quad (3.1)$$

Now, by partially differentiating (3.1) with respect to $\beta$and equating to zero, we get-

$$\hat{\beta} = 1-\frac{f_0}{f}$$

And partially differentiate (3.1) with respect to α and equating to zero, we get –

$$\hat{\alpha} = \frac{3f_1 \pm \sqrt{8ff_1 - 8f_0f_1 + {f_1}^2}}{8(f-f_0-f_1)}$$

we consider only $\hat{\alpha} = \dfrac{3f_1 + \sqrt{8ff_1 - 8f_0f_1 + {f_1}^2}}{8(f-f_0-f_1)}$ $\because \alpha > 0$

Where $\hat{\alpha}$ and $\widehat{\beta}$ are maximum likelihood estimations of α and β, respectively.

Now $$\frac{E\left(\frac{-\partial^2}{\partial\beta^2}\log L\right)}{f} = \left[\frac{1}{1-\beta}+\frac{1}{\beta}\right]$$

and $$\frac{E\left(\frac{-\partial^2}{\partial\alpha^2}\log L\right)}{f} = \frac{32\beta(1+12\alpha+53\alpha^2+102\alpha^3+72\alpha^4)}{(1+6\alpha+8\alpha^2)(1+16\alpha^2+8\alpha)(1+6\alpha)(1+2\alpha)^2}$$

Ultimately we get –

$$E(f_0) = f(1-\beta)$$

$$E(f_1) = f\left(\frac{8\alpha^2\beta}{1+6\alpha+8\alpha^2}\right)$$

After estimating all the parameters involved in the proposed model by M.L.E., we are interested in calculating the variance of each estimated parameter.

$$FIM = \begin{bmatrix} \frac{1}{1-\beta}+\frac{1}{\beta} & 0 \\ 0 & \frac{32\beta(1+12\alpha+53\alpha^2+102\alpha^3+72\alpha^4)}{(1+6\alpha+8\alpha^2)(1+16\alpha^2+8\alpha)(1+6\alpha)(1+2\alpha)^2} \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$$

$$V(\hat{\alpha}) = \frac{1}{f}\left[\frac{\Phi_{11}}{\Phi_{11}\Phi_{22}}\right] = \frac{1}{f}\left\{\frac{(1+6\alpha+8\alpha^2)(1+16\alpha^2+8\alpha)(1+6\alpha)(1+2\alpha)^2}{32\beta(1+12\alpha+53\alpha^2+102\alpha^3+72\alpha^4)}\right\}$$

$$V(\hat{\beta}) = \frac{1}{f}\left[\frac{\Phi_{22}}{\Phi_{11}\Phi_{22}}\right] = \frac{\beta(1-\beta)}{f}$$

4. Application of the Proposed Inflated Model

The involved parameters of the proposed inflated model are estimated with the help of the maximum likelihood method and derive the Fisher Information Matrix. The suggested inflated model's applicability is tested using data from Pandey and Kishun [1] for child mortality in various nations and genetic count data from Hassan *et al*. [7] on streptonigrin treatment in rabbit lymphoblast.

Table 1. Distribution of observed and expected number of families according to the number of child deaths in Eastern Uttar Pradesh, India.

Number of child deaths	Observed no. of families	Expected no. of families	
		PAD	IMZTPAD
0	506	467.4	506
1	178	224.8	178.2
2	76	81.14	78.42
3	32	26.02	30.66
4	8	7.82	11.24
5			
6	6	2.82	5.48
7	3 } 10		
	1		
Total	810	810	810
Mean = 0.6333			
ML Estimate		$\hat{\alpha} = 1.5789$ –	$\hat{\alpha} = 1.205, \hat{\beta} = 0.3753$ $V(\hat{\alpha}) = 0.0156, V(\hat{\beta}) = 2.89 \times 10^{-4}$
χ^2		32.90	4.79
d.f		4	3
p-value		3.10×10^{-3}	0.1878

Table 2. Distribution of observed and expected number of families according to the number of child deaths in North East Brazil.

Number of child deaths	Observed no. of families	Expected no. of families	
		PAD	IMZTPAD
0	769	724.35	769
1	185	246.10	185
2	60	62.69	66.67
3	26	14.19	21.35
4			
5	9	3.67	8.98
6	1 } 11		
7	1		
	0		
Total	1051	1051	1051
Mean = 0.4091			
ML Estimate		$\hat{\alpha} = 2.444$ –	$\hat{\alpha} = 1.581, \hat{\beta} = 0.2683$ $V(\hat{\alpha}) = 0.0327, V(\hat{\beta}) = 1.86 \times 10^{-4}$

χ^2	42.48	2.12
d.f	3	2
p-value	3.916×10^{-4}	0.3464

Table 3. Distribution of observed and expected number of families according to the number of child deaths in North East Libya.

Number of child deaths	Observed no. of families	Expected no. of families	
		PAD	IMZTPAD
0	805	784.91	805
1	306	326.83	305.97
2	93	102.02	100.64
3	36	28.31	29.42
4			
5	7	9.93	10.97
6	2 } 12		
7	1		
	2		
Total	1252	1252	1252
Mean = 0.5255			
ML Estimate		$\hat{\alpha} = 1.902$	$\hat{\alpha} = 1.780, \hat{\beta} = 0.3570$
		–	$V(\hat{\alpha}) = 0.0278, V(\hat{\beta}) = 1.83 \times 10^{-4}$
χ^2		5.13	2.14
d.f		3	2
p-value		0.1625	0.3430

Table 4. Mammalian cytogenetic dosimetry lesions in rabbit lymphoblast induced by Streptonigrin (NSC-45383) Exposure -70 μg/kg.

Class/Exposure	Observed Frequency	Expected Frequency	
		PAD	IMZTPAD
0	200	184.03	200
1	57	79.77	57.03
2	30	25.93	26.09
3	7	7.49	10.61
4			
5	4	2.78	6.27
6	0 } 6		
	2		
Total	300		300
Mean = 0.5533			
ML Estimate		$\hat{\alpha} = 1.807$	$\hat{\alpha} = 1.139, \hat{\beta} = 0.333$
		–	$V(\hat{\alpha}) = 0.0416, V(\hat{\beta}) = 7.4 \times 10^{-4}$
χ^2		12.26	1.81
d.f		3	2
p-value		0.0065	0.4045

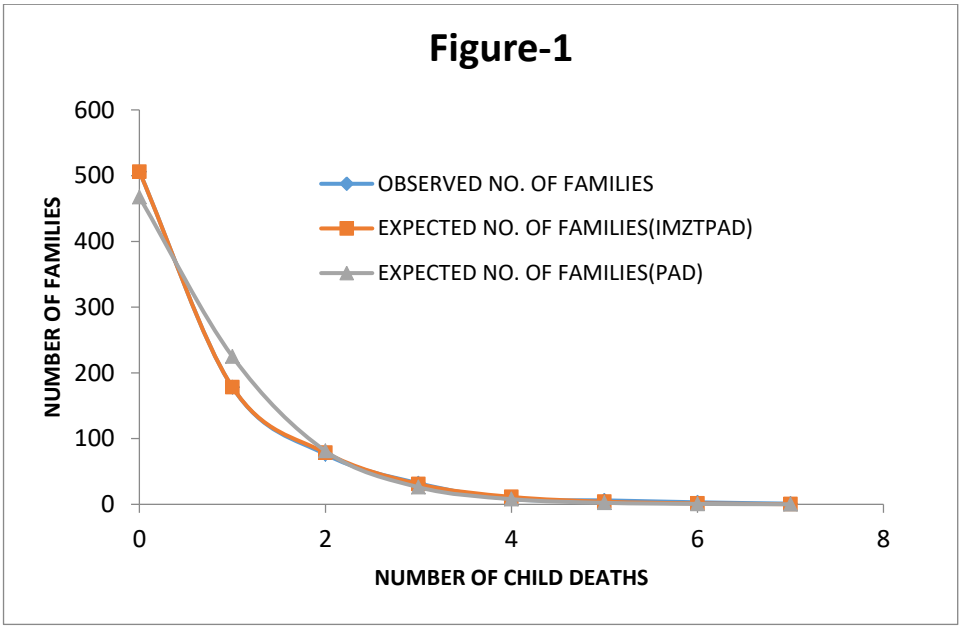

Fig. 1. Graphical presentation showing the observed and Expected number of families according to the number of child deaths in Eastern Uttar Pradesh (India).

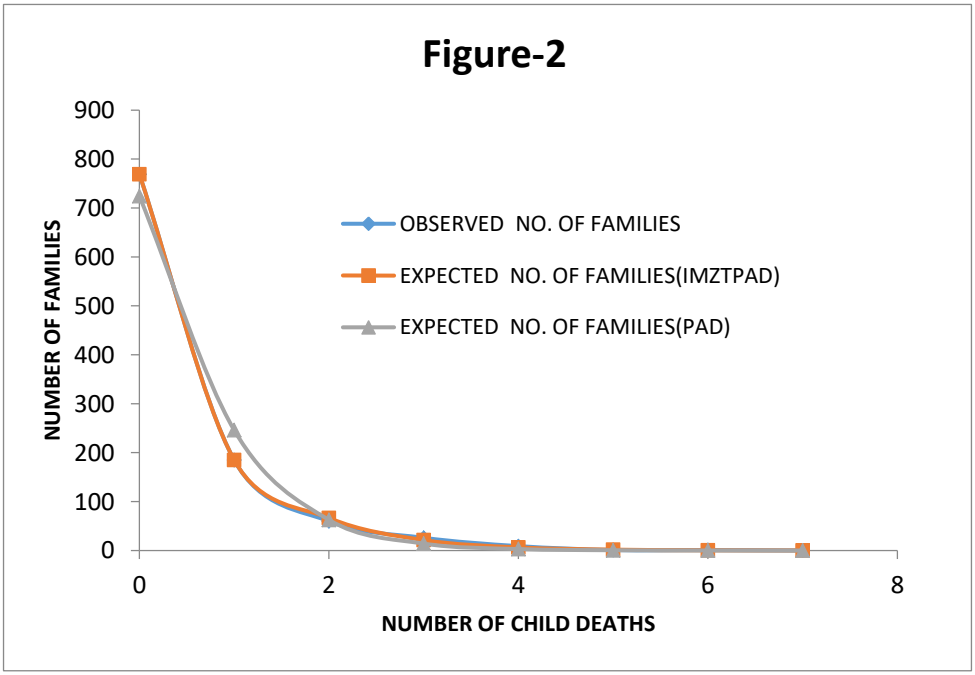

Fig. 2. Graphical presentation showing observed and expected number of families according to the number of child deaths in North East Brazil.

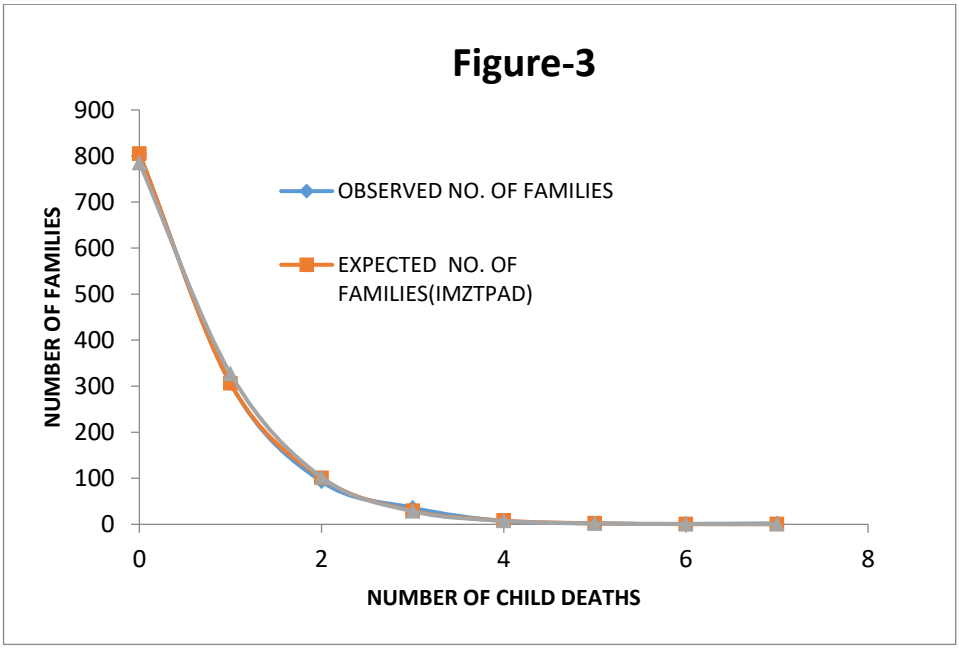

Fig. 3. Graphical presentation showing observed and expected number of families according to the number of child deaths in North East Libya.

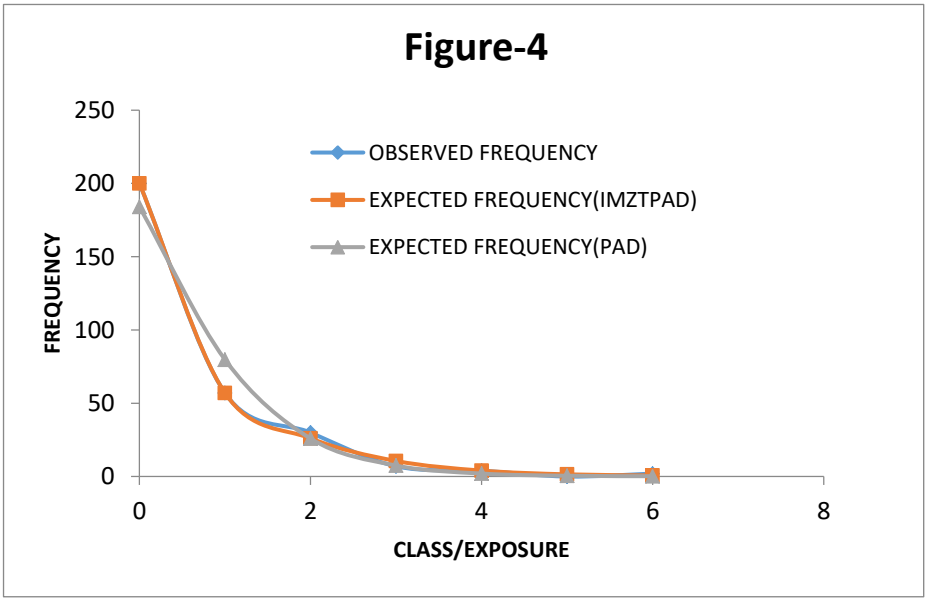

Fig. 4. Graphical presentation showing observed and expected frequency of mammalian cytogenetic dosimetry lesions in rabbit lymphoblast induced by streptonigrin (NSC-45383) Exposure -70 µg/kg.

5. Conclusion

This study introduces the inflated version of the zero truncated Poisson Ailamujia distribution (IMZTPAD). Additionally, we have computed a number of characteristics, including mean, variance, and fisher information. The proposed model included two parameters that were used to describe the pattern of mortality and genetic count data. These parameters were computed using the maximum likelihood technique. A chi-square goodness of fit test determines whether sample data are representative of the population. A significant correlation between two data sets is indicated by a low chi-square score. There is sufficient evidence to determine that the observed distribution doesn't differ from the expected distribution if the p-value for a Chi-square test is greater than your significance threshold. On the basis of chi-square and its p-value, the suggested model's applicability was evaluated using the set of observed data. We reached the conclusion that the proposed model is suitable for the observed set of mortality and genetic count data, and it was demonstrated that the suggested IMZTPAD model is superior to the already employed PAD model, proving that inflated models are better than simple models. Overall research shows that the proposed inflated version of zero truncated Poisson Ailamujia model may also be helpful in policy making, rural development, fresh environment, and medical facilities for the development of society. The work done by Rao and Pandey may also be examined in relation to the Bayesian analysis of the suggested inflated model.

References

1. H. Pandey and J. Kishun, Revista Colombiana de Estadistica. **33**, 1 (2010).
2. H. Pandey and R. Tiwari, Recent Res. Sci. Technol. **3**, 100 (2011).
3. A. Pandey, H. Pandey, and V. K. Shukla, J. Comput. Math. Sci. **6**, 702 (2015).
4. A. Agarwal and H. Pandey, Int. J. Stat. Appl. Math. **7**, 48 (2022). https://doi.org/10.22271/maths.2022.v7.i6a.899
5. A. Agarwal and H. Pandey, Asian J. Probab. Stat. **20**, 57 (2022). https://doi.org/10.9734/AJPAS/2022/v20i4439
6. H. Q. Lv, L. H Gao, and C. L Ches, J. Acad. Armored Force Eng. **3**, 48 (2002).
7. A. Hassan, S. Bilal, A. Rashid, and G. Ahmad, J. Stat. Theory Appl. **19**, 102 (2020). https://doi.org/10.2991/jsta.d.200224.006
8. A. Agarwal and H. Pandey, South E. Asian J. Math. And Math Sci. **19**, 417 (2023). https://doi.org/10.56827/SEAJMMS.2023.1902.31
9. A. K. Rao and H. Pandey, Lithuanian J. Stat. **56**, 88 (2017). http://dx.doi.org/10.15388/LJS.2017.13675
10. A. K. Rao and H. Pandey, J. Econometrics Stat. **2**, 127 (2022). https://doi.org/10.47509/IJMSOR.2022.v02i01.05