South East Asian J. of Mathematics and Mathematical Sciences
Vol. 19, No. 2 (2023), pp. 417-430

DOI: 10.56827/SEAJMMS.2023.1902.31 **ISSN (Online): 2582-0850**

ISSN (Print): 0972-7752

ON ZERO TRUNCATION MODELLING OF POISSON AILAMUJIA DISTRIBUTION AND ITS APPLICATIONS

Abhishek Agarwal and Himanshu Pandey

Department of Mathematics and Statistics,
D.D.U. Gorakhpur University,
Gorakhpur, Uttar Pradesh, INDIA

E-mail : abhishek.agarwal014@gmail.com, himanshu_pandey62@yahoo.com

(**Received:** *Feb. 21, 2023* **Accepted:** *Jul. 13, 2023* **Published:** *Aug. 30, 2023)*

Abstract: In present paper we have developed zero truncated Poisson Ailamujia distribution. The suggested model's parameter is calculated using the maximum likelihood technique. The proposed model fitted to observed set of data related to the events based on mortality and drawn some conclusions.

Keywords and Phrases: Poisson Ailamujia distribution, Zero truncated distribution, Method of maximum likelihood, Mortality, Goodness of fit.

2020 Mathematics Subject Classification: 62E10, 62F10, 62P25, 62-08, 62-11.

1. Introduction

In probability theory zero truncated distribution play an important role to draw the valuable inferences in respect of real life problems. When the data come from an environment that doesn't provide data with zero counts, zero truncated distributions are appropriate. Following (Shanker et al., 2015), Shanker R. (2017), Agarwal and Pandey (2022), let $p_0(x;\theta)$ is the original discrete distribution with support non negative positive integers, then the zero truncated form of $p_0(x;\theta)$ with support set of positive integers is given by-

$$p(x,\theta) = \frac{p_0(x;\theta)}{1-p_0(0;\theta)}; \qquad x = 1, 2, 3, ... \tag{1.1}$$

Considering that despite a notable drop in the majority of developed countries, infant and child mortality rates are still alarmingly high in underdeveloped nations. In demographic science infant mortality and child mortality are sensitive index of a national health condition and effect of socio- cultural, economic and psychological factors of the country. In this respect Hill and David (1989), Kabir and Amir (1993), Krishnan and Jin (1993), Sastry N. (1997), Srivastava (2001), Pandey and Kishun (2010), Agarwal and Pandey (2022) drawn various inferences through the model building with the help of the distribution related to mortality data.

The primary goal of the current study is to create a zero-truncated version of the Poisson Ailamujia distribution and use maximum likelihood to estimate the distribution's parameter. With the observed set of data, the applicability of the zero truncated distribution model was examined, and several inferences were made.

2. Zero Truncated Model of the Poisson Ailamujia Distribution

Let Random variable X has a Poisson distribution with $\lambda > 0$ with probability mass function

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x = 0, 1, 2, \dots . \tag{2.1}$$

And the probability density function of Ailamujia distribution with parameter α is given by

$$f(x) = 4x\alpha^2 e^{-2\alpha x}; \quad x \geq 0, \alpha > 0. \tag{2.2}$$

Then the probability mass function of a compound of $P(\lambda)$ with $AD(\alpha)$ is given by Hassan etal (2020)

$$PAD(X = x, \alpha) = \frac{4\alpha^2(1+x)}{(1+2\alpha)^{x+2}}; \quad x = 0, 1, 2, ...; \quad \alpha > 0. \tag{2.3}$$

The form of zero truncated Modelling of Poisson Ailamujia distribution (ZTPAD) form equation (2.3) is as follows
By the definition of zero truncation (1.1),

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

$$\Rightarrow p(x;\alpha) = \frac{4\alpha^2(1+x)}{(1+4\alpha)(1+2\alpha)^x}; \quad x = 1, 2, 3, ...; \quad \alpha > 0 \tag{2.4}$$

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

3. Estimation of the parameter (α)

The proposed zero truncation modelling on Poisson Ailamujia distribution (2.4) has a parameter $\alpha > 0$ which is estimated by method of maximum likelihood in the following way.
Let $x_1, x_2, ..., x_n$ be a random sample of size n drawn from the population which follows zero truncated Poisson Ailamujia distribution (2.4), then its likelihood function can be defined as

$$L = \prod_{i=1}^{n} p(x_i;\alpha)$$
$$L = \left(\frac{4\alpha^2}{(1+4\alpha)}\right)^n \frac{(1+x_1)^{f_1}(1+x_2)^{f_2}...(1+x_n)^{f_n}}{(1+2\alpha)^{\sum_{i=1}^{n} f_i x_1}} \tag{3.1}$$

Now

$$\log L = n\log\left(\frac{4\alpha^2}{(1+4\alpha)}\right) + \sum_{i=1}^{n} f_i \log(1+x_i) - \sum_{i=1}^{n} f_i x_i \log(1+2\alpha) \tag{3.2}$$

Put $\dfrac{d}{d\alpha}\log L = 0$

$$\Rightarrow \frac{n}{\left(\frac{4\alpha^2}{(1+4\alpha)}\right)}\left\{\frac{(1+4\alpha)8\alpha - 16\alpha^2}{(1+4\alpha)^2}\right\} - \frac{2\sum f_i x_i}{(1+2\alpha)} = 0$$
$$\Rightarrow \frac{n}{\alpha}\left(\frac{2+4\alpha}{1+4\alpha}\right) = \frac{2\sum f_i x_i}{(1+2\alpha)}$$
$$\Rightarrow \bar{x} = \frac{(1+2\alpha)^2}{\alpha(1+4\alpha)}$$
$$\Rightarrow 1 + 4\alpha^2 + 4\alpha = \bar{x}(\alpha + 4\alpha^2)$$
$$\Rightarrow 4\alpha^2(\bar{x}-1) + \alpha(\bar{x}-4) - 1 = 0$$
$$\Rightarrow \widehat{\alpha} = \frac{-(\bar{x}-4) \pm \sqrt{\bar{x}^2 + 8\bar{x}}}{8(\bar{x}-1)}$$

we consider only positive value of $\widehat{\alpha} = \dfrac{-(\bar{x}-4) + \sqrt{\bar{x}^2+8\bar{x}}}{8(\bar{x}-1)}$ as $\alpha > 0 \ \because \bar{x} > 0$.

Now

$$\frac{d^2}{d\alpha^2}\log L = \left[\frac{4n(\alpha+4\alpha^2)-2n(1+2\alpha)(1+8\alpha)}{\{\alpha(1+4\alpha)\}^2}\right] + \frac{4\sum_i f_i x_i}{(1+2\alpha)^2}$$

$$\Rightarrow \frac{1}{n}E\left(-\frac{d^2}{d\alpha^2}\log L\right) = \frac{16\alpha^2+16\alpha+2}{\alpha^2(1+4\alpha)^2} - \frac{4\bar{x}}{(1+2\alpha)^2}$$

$$= \frac{(1+2\alpha)^2(16\alpha^2+16\alpha+2)-4\bar{x}\alpha^2(1+4\alpha)^2}{\alpha^2(1+4\alpha)^2(1+2\alpha)^2}$$

$$\text{Hence} \quad V(\widehat{\alpha}) = \frac{1}{n}\left[\frac{\alpha^2(1+4\alpha)^2(1+2\alpha)^2}{(1+2\alpha)^2(16\alpha^2+16\alpha+2)-4\bar{x}\alpha^2(1+4\alpha)^2}\right]$$

4. Application

In addition to biological and epidemiological variables, other socio-economic and cultural elements that affect mortality include current medical facilities, environmental circumstances etc. Infant and child mortality rates are substantially higher than those for children over 15 years old in developing and underdeveloped nations. The high infant mortality has presented a significant challenge to the medical professionals, and it is regarded as one of the critical situations of the population's current medical and health facilities. Hence, using our suggested model, we attempt to examine the newborn mortality trend. The data set of Meegama (1980) was applied to represent Sri Lanka, while the data from Lal's and Kadam Kuan, Patna's (1975) surveys—both of which were cited by Mishra (1979)—were employed to represent India. The maximum likelihood estimates, chi-squares, values of $-2\log L$, Akaike information criterion $(AIC) = -2\log L + 2k$, Bayesian information criterion $(BIC) = -2\log L + k\log n$, Akaike information criterion corrected $(AICC) = AIC + \frac{2k(k+1)}{n-k-1}$; where k is the number of parameters involved in the distribution; have been computed for the following datasets.

The proposed Model (2.4) is compared with One-Dimensional Biased probability Model based on Himanshu distribution (ODBHD) given by Agarwal and Pandey (2023) follows as

$$P(X=x) = xp^{2n}(1-p^n)^{x-1}; \quad x = 1,2,3,...,; 0 < p < 1, \quad n \in I^+$$

with mean $= \frac{2}{p^n} - 1$ and variance $= 2\left(\frac{1-p^n}{p^{2n}}\right)$

Table 1. The number of mothers of the Rural Area having at least one live birth and one neonatal death.

Number of neonatal deaths	Observed no. of mothers	Expected no. of mothers	
		ZTPAD	ODBHD
1	409	405.82	402.67
2	88	92.79	97.92
3	19	18.86	17.86
4	$\left.\begin{matrix}5\\1\end{matrix}\right\}6$	4.52	3.54
5			
Total	**522**	**522**	**522**
ML Estimates		$\hat{\alpha} = 2.78$	$\hat{p} = 0.9372$
Standard error of M.L.E.		0.2479	5.056×10^{-3}
χ^2		0.75	2.86
d.f.		2	2
p-value		0.6872	0.2393
-2 log L		303.680	303.808
AIC		305.680	305.808
BIC		306.392	306.525
AICC		305.687	305.815

Table 2. The number of mothers of the Estate Area having at least one live birth and one neonatal death.

Number of neonatal deaths	Observed no. of mothers	Expected no. of mothers	
		ZTPAD	ODBHD
1	71	70.90	69.27
2	32	29.89	32.38
3	7	11.20	11.34
4	$\left.\begin{matrix}5\\3\end{matrix}\right\}8$	6.01	5.01
5			
Total	**118**	**118**	**118**
ML Estimates		$\hat{\alpha} = 1.279$	$\hat{p} = 0.8753$
Standard error of M.L.E.		0.1732	0.0137
χ^2		2.38	3.48
d.f.		2	2
p-value		0.3042	0.1755
-2 log L		110.354	110.600
AIC		112.354	112.600
BIC		112.425	112.671
AICC		112.388	112.634

Table 3. The number of mothers of the Urban Area with at least two live births by the number of infant and child deaths.

Number of Infant and child deaths	Observed no. of mothers	Expected no. of mothers	
		ZTPAD	ODBHD
1	176	169.58	168.01
2	44	53.80	57.19
3	16	15.17	14.60
4	$\left.\begin{matrix}6\\2\end{matrix}\right\}8$	5.45	4.20
5			
Total	**244**	**244**	**244**
ML Estimates		$\hat{\alpha} = 1.864$	$\hat{p} = 0.9109$
Standard error of M.L.E.		0.2046	8.507×10^{-3}
χ^2		3.26	6.97
d.f.		2	2
p-value		0.1959	0.0306
-2 log L		182.932	184.179
AIC		184.932	186.179
BIC		185.319	186.566
AICC		184.948	186.195

Table 4. The number of mothers of the Rural Area with at least two live births by the number of infant and child deaths.

Number of Infant and child deaths	Observed no. of mothers	Expected no. of mothers	
		ZTPAD	ODBHD
1	745	730.47	720.83
2	212	244.80	230.66
3	50	61.49	55.35
4	21	15.77	11.81
5	$\left.\begin{matrix}7\\3\end{matrix}\right\}10$	5.47	19.34
6			
Total	**1038**	**1038**	**1038**
ML Estimates		$\hat{\alpha} = 1.937$	$\hat{p} = 0.9128$
Standard error of M.L.E.		0.1048	4.093×10^{-3}
χ^2		8.64	14.48
d.f.		3	3
p-value		0.0344	0.0023
-2 log L		760.430	763.547
AIC		762.430	765.547
BIC		763.446	766.563
AICC		762.433	765.550

Table 5. The number of Literate mothers with at least one live birth by the number of infant and child deaths.

Number of Infant and child deaths	Observed no. of mothers	Expected no. of mothers	
		ZTPAD	**ODBHD**
1	683	669.86	665.89
2	145	157.33	168.60
3	29	34.12	32.01
4	11 } 16	11.69	6.49
5	5		
Total	**873**	**873**	**873**
ML Estimates		$\hat{\alpha} = 2.633$	$\hat{p} = 0.9345$
Standard error of M.L.E.		0.1777	3.977×10^{-3}
χ^2		3.58	17.94
d.f.		2	2
p-value		0.1669	0.0001
-2 log L		526.844	530.158
AIC		528.844	532.158
BIC		529.785	533.099
AICC		528.848	532.162

Table 6. The number of mothers of the completed fertility having experienced at least one child death.

Number of child deaths	Observed no. of mothers	Expected no. of mothers	
		ZTPAD	**ODBHD**
1	89	81.77	79.88
2	25	34.05	36.85
3	11	12.60	12.75
4	6	6.58	5.52
5	3 } 10		
6	1		
Total	**135**	**135**	**135**
ML Estimates		$\hat{\alpha} = 1.301$	$\hat{p} = 0.8770$
Standard error of M.L.E.		0.1656	0.01281
χ^2		5.024	8.72
d.f.		2	2
p-value		0.0811	0.0127
-2 log L		125.897	127.475
AIC		127.897	129.475
BIC		128.027	129.605
AICC		127.927	129.505

Table 7. The number of mothers having at least one neonatal death.

Number of Neonatal Deaths	Observed no. of mothers	Expected no. of mothers	
		ZTPAD	ODBHD
1	567	556.76	554.39
2	135	145.69	153.01
3	28	33.89	31.67
4	11 } 16	9.65	6.92
5	5		
Total	**746**	**746**	**746**
ML Estimates		$\hat{\alpha} = 2.366$	$\hat{p} = 0.9284$
Standard error of M.L.E.		0.1647	4.465×10^{-3}
χ^2		6.173	14.72
d.f.		2	2
p-value		0.0456	0.0006
-2 log L		481.708	483.907
AIC		483.708	485.907
BIC		484.580	486.779
AICC		483.713	485.912

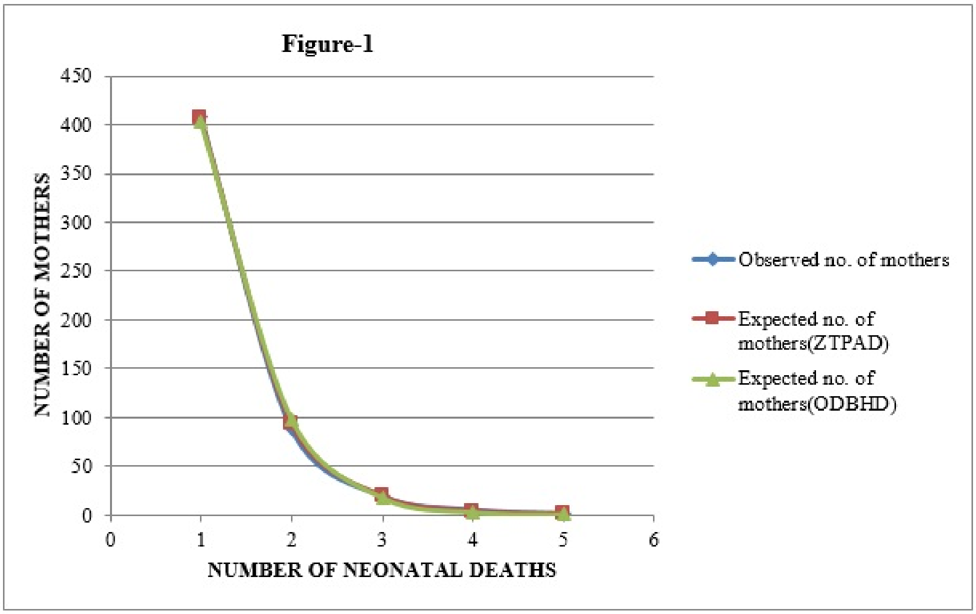

Figure 1: Graphical presentation showing observed and expected number of mothers of the Rural Area having at least one live birth and one neonatal death.

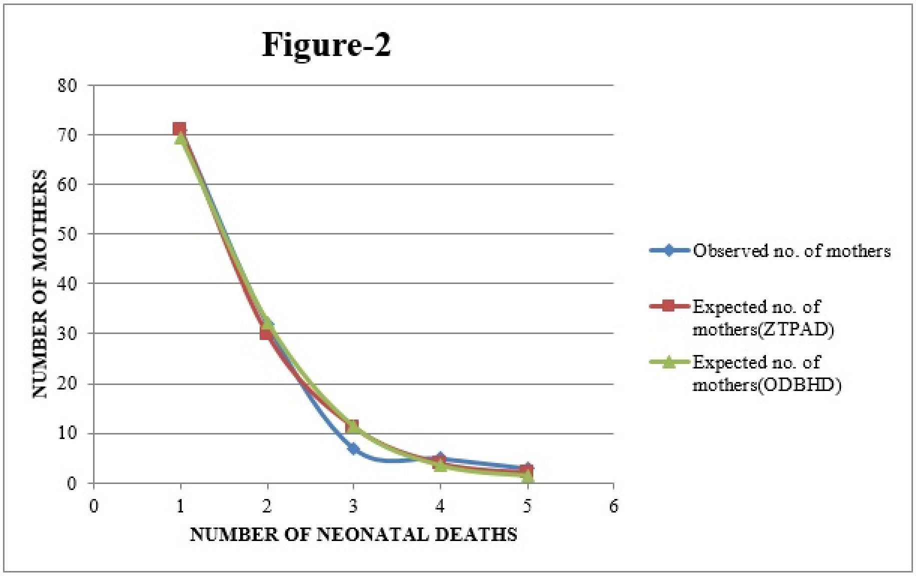

Figure 2: Graphical presentation showing observed and expected number of mothers of the Estate Area having at least one live birth and one neonatal death.

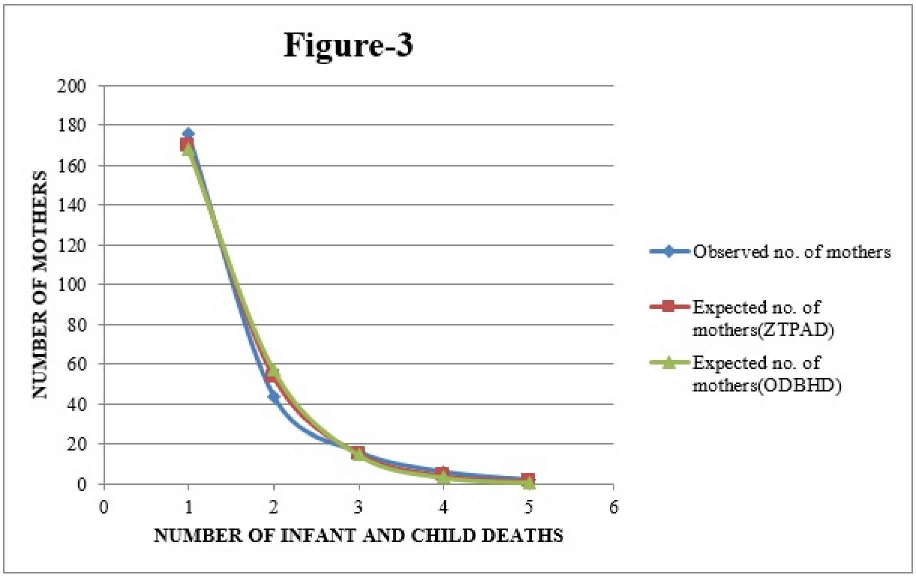

Figure 3: Graphical presentation showing observed and expected number of mothers of the Urban Area with at least two live births by the number of infant and child deaths.

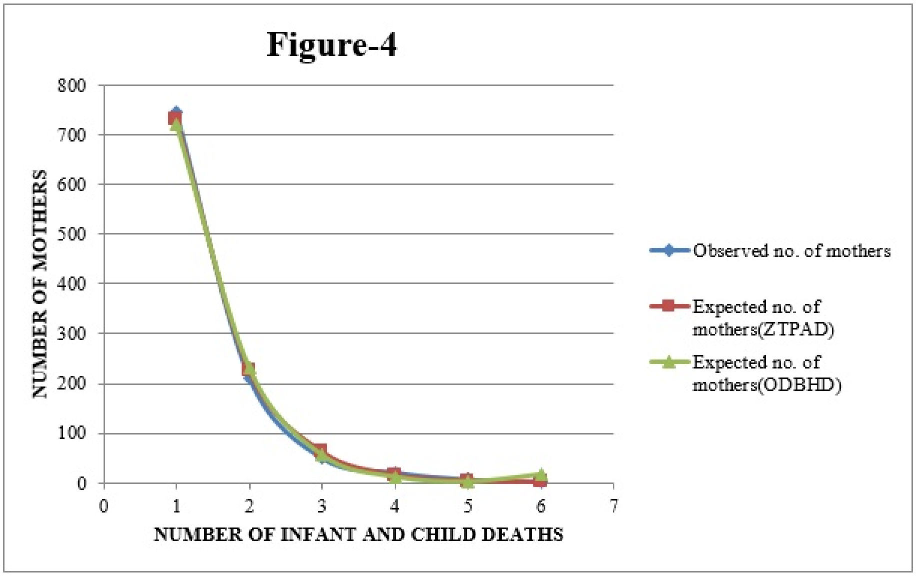

Figure 4: Graphical presentation showing observed and expected number of mothers of the Rural Area with at least two live births by the number of infant and child deaths.

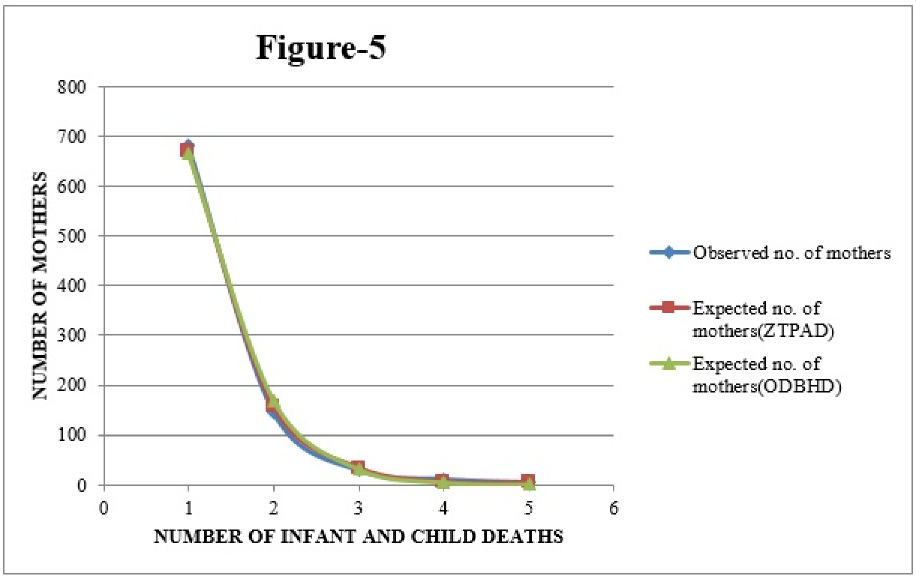

Figure 5: Graphical presentation showing observed and expected number of Literate mothers with at least one live birth by the number of infant and child deaths.

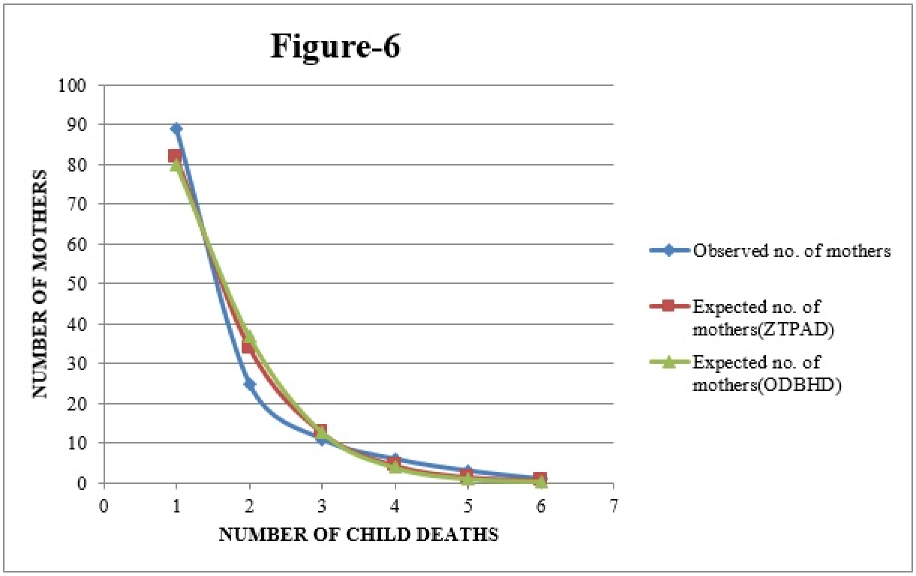

Figure 6: Graphical presentation showing observed and expected number of mothers of the completed fertility having experienced at least one child death.

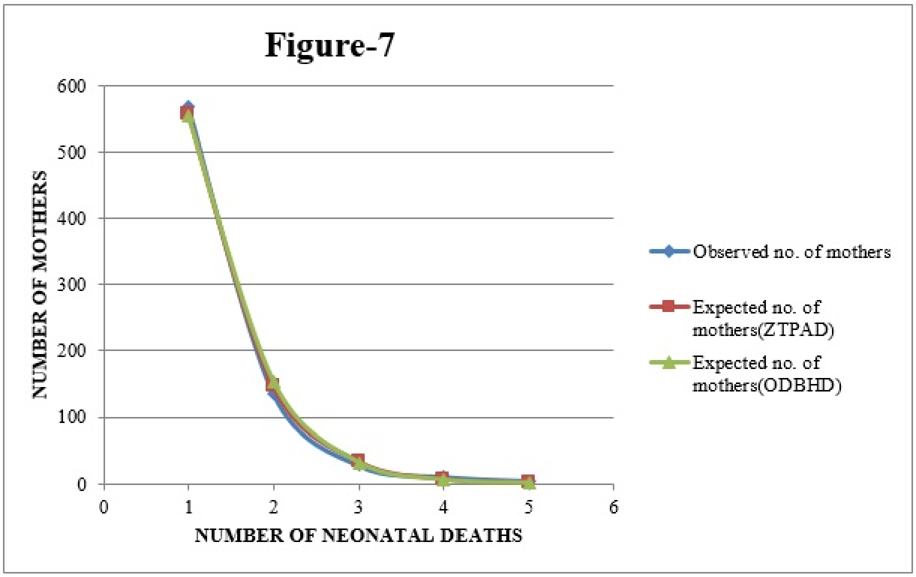

Figure 7: Graphical presentation showing observed and number of mothers having at least one neonatal death.

5. Conclusion

A chi-square goodness of fit test determines if sample data matches a population. By computing several metrics, including chi square, p-value, AIC, BIC, and AICC, an attempt has been made to assess the goodness of fit of ZTPAD over ODBHD. For demographic data sets associated to mortality, the model (ZTPAD) with the lowest chi square, highest p-value, and lowest AIC, BIC, and AICC values was determined to have the greatest fit over ODBHD. A low value for chi-square means there is a high correlation between your two sets of data. From tables (1 to 7), it clearly indicates calculated χ^2 is less than the critical χ^2 value at 1% and 5% level of significance, Hence we conclude there is no significant difference between the observed and expected value of the given data set. For a Chi-square test, a p-value that is greater than to your significance level indicates there is sufficient evidence to conclude that the observed distribution is same as the expected distribution. According to the values of χ^2, AIC, BIC, AICC and p-value from (Tables 1-7) and graphical representation between Oi and Ei, the nature and behavior of proposed zero truncated Poisson Ailamujia model found best suitable for the mortality pattern of different regions.

Overall research demonstrates that the suggested zero truncated Poisson Ailamujia model might also be useful in policy making, rural development, fresh environment, and medical facilities for the development of society. Probability is used in Bayesian analysis for both data and hypotheses. It pertains to a subjective assessment of the veracity of an occurrence. A different approach to traditional statistics is provided by Bayesian statistics. It stands out for its capacity to characterize uncertain values using probability distributions, which leads to elegant solutions to several challenging statistical problems and is extensively useful in the fields of demography, medicine, and insurance. Now that these viewpoints and the work of Rao and Pandey (2021) have been taken into consideration, it is possible to employ the Bayesian Analysis of the suggested model by figuring out various loss functions.

Acknowledgement

The Authors are thankful to the Referees for their valuable suggestions.

References

[1] Agarwal, A., & Pandey, H., Probability Model based on Zero Truncation of Himanshu distribution and their Applications, Asian Journal of Probability and Statistics, 20(4) (2022), 57-67.

[2] Agarwal, A., & Pandey, H., A One-Dimensional Biased Probability Model

based on Himanshu distribution for vital events related to Migration and Mortality data, Journal of Statistical Research, 13(1) (2023), 48-60.

[3] Hassan, A; Bilal, S., Rashid, A., & Ahmad, G., A new flexible discrete distribution with Application to count data, Journal of Statistical Theory and Application, 19(1) (2020), 102-108.

[4] Hill, A. G., & David, H. P., Measuring Child Mortality in the Third World: Neglected Sources and Novel Approaches, IUSSP Proceeding of Interactional Conference, New Delhi, India, 1989.

[5] Meegama, S. A., Socio-economic determinants of infant and child mortality in Sri Lanka, An analysis of post war experience, International Statistical Institute (World Fertility Survey), Netherland, 1980.

[6] Mishra, A., Generalizations of some discrete distributions, Unpublished Ph.D. thesis, Patna University, Patna, India, 1979.

[7] Kabir, M., & Amir, R., Factors influencing child mortality in Bangladesh and their implications for the national health programme, Asia-Pacific population Journal, 8(3) (1993), 31-46.

[8] Krishnan, P., & Jin, Y., A Statistical Model of Infant Mortality, Janasamkhya, 11(2) (1993), 67-71.

[9] Lal, D. N., A Demographic Sample Survey, Demographic Research Centre, Department of Statistics, Patna University, India, 1975.

[10] Pandey, H., & Jai Kishun., A Probability Model for the Child Mortality in a Family, Revista colombiana de Estadistica, 33(1) (2010), 1-11.

[11] Rao, A. K., & Pandey, H., Reliability estimation of Weibull Exponential Distribution via Bayesian Approach, Ratio Mathematica, 40(1) (2021), 123-137.

[12] Rao, A. K., & Pandey, H., Bayesian Estimation of Shape Parameter of Power Lomax function under different loss function, Journal of Mathematical Science and Computational Mathematics, 2(2) (2021), 227-235.

[13] Sastry, N., Family level clustering of childhood mortality risk in Northeast Brazil, Population studies, 51(1) (1997), 245-261.

[14] Shanker, R., Fesshaye, H., Selvaraj, S., On Zero truncation of Poisson and Poisson-Lindley Distributions and their Applications, Biometrics and Biostatistics International Journal, 2(6) (2015), 168-181.

[15] Shanker, R., A Zero Truncated Poisson-Amarendra distribution and its Application, International Journal of Probability and Statistics, 6(4) (2017), 82-92.

[16] Srivastava, S., Some Mathematical Models in Demography and their Application, Unpublished Ph.D. thesis in Statistics, BHU (Varanasi) India, 2001.