

# Probability Model Based on Zero Truncation of Himanshu Distribution and Their Applications.

**Abstract-** In this paper we have developed a probability model to study the distribution of the number of male migrants from households and patterns of child mortality through zero truncated Himanshu distribution (ZTHD). The parameter of the proposed model is estimated by method of moment and method of maximum likelihood and its suitability is tested by fitting to observed migration data, collected from various surveys.

**Keywords -** Migration, Himanshu distribution, Zero truncated distribution, Estimation technique.

## 1. Introduction

Probability models based on distributions play an important role in the various fields of Social Sciences, Medical Sciences, Environmental Sciences etc. Recently Pandey and Jai kishun(2010), Pandey et al.(2014), Shankar(2017), Shukla and Shankar(2017,2019,2020), Rashid et al.(2022), Dubey and Pandey(2022) have suggested probability models for vital events and biological science data.

“Zero truncated distributions are certain discrete distributions having support on the set of positive integers. These distributions are applicable for situations when data to be modelled originate from a mechanism that generates data excluding zero-counts”. [14]

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

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

Agarwal and Pandey (2022) have introduced Himanshu distribution having probability mass function as-

$$P(X = x) = p^n(1 - p^n)^x; \quad \left. \begin{array}{l} x = 0, 1, 2, \dots \\ 0 < p < 1 \\ n \in I^+ \end{array} \right\} \quad (1.2)$$

with Mean =  $\frac{1 - p^n}{p^n}$  and Variance =  $\frac{1 - p^n}{p^{2n}}$

## 2. Proposed Probability Model -

Using (1.1) and (1.2) the probability mass function of zero truncated Himanshu distribution (ZTHD) is given by –

$$P(X = x) = p^n(1 - p^n)^{x-1}; \quad \left. \begin{array}{l} x = 1, 2, 3, \dots \\ 0 < p < 1 \\ n \in I^+ \end{array} \right\} \quad (2.1)$$

The moment generating function of (2.1) is given as-

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

$$M_X(t) = \frac{p^n e^t}{1 - e^t(1-p^n)} \quad (2.2)$$

Then the first four moments (about origin) are as follows-

$$\text{Mean} = \mu'_1 = \frac{1}{p^n} \quad \mu'_2 = \frac{2-p^n}{p^{2n}}$$

$$\mu'_3 = \frac{(p^{2n} - 6p^n + 6)}{p^{3n}} \quad \text{and} \quad \mu'_4 = \frac{-(p^{3n} - 14p^{2n} + 36p^n - 24)}{p^{4n}} \quad (2.3)$$

Now the first four (central moments) are given by-

$$\mu_1 = 0, \quad \mu_2 = \frac{1-p^n}{p^{2n}}$$

$$\mu_3 = \frac{(1-p^n)(2-p^n)}{p^{3n}} \quad \text{and} \quad \mu_4 = \frac{(1-p^n)(p^{2n} - 9p^n + 9)}{p^{4n}} \quad (2.4)$$

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

⇒ Distribution is positively skewed

$$\gamma_2 = \frac{p^{2n} - 6p^n + 6}{1-p^n}$$

⇒ Distribution is leptokurtic

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

### 3. Estimation of parameter of ZTHD –

The parameter  $p$  of the proposed model (2.1) is estimated by maximum likelihood estimation method in the following way –

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

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

Now taking log and upon differentiating w.r.t  $p$  and equating to zero we get -

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

Again the parameter  $p$  estimated by method of moments after using (2.3) we get –

$$\bar{x} = \mu'_1$$

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

### 4. Application

The proposed probability model at  $n=2$  is fitted using some real data sets collected from different sample surveys entitled “Demographic survey of Chandauli district (Rural Area- 2001-2002)”- (2015); Rural development and population growth survey 1978-PRC, BHU” (2015); and “the survey collected by the researcher in Varanasi district (2018) for the single adult male migrant of 15 years and above.

“Mortality does not depend only on the biological and epidemiological factors, it also depends upon the prevailing health conditions, Medical facilities, environmental conditions and some other socio- economic and cultural factors. In developing and under developing countries, the mortality among infants and children is found much higher than that among youngsters above 15 years. The high infant mortality has thrown a serious challenge to the doctors, medical personnel’s and is considered as one of the sensitive position of existing medical and health facilities in the population”. [14]. Therefore, we attempt to study the infant mortality pattern through our proposed model. In this respect the real data set of Sri Lanka taken from the survey by Meegama (1980) and the real data set of India taken from Lal (1955) are used.

**Table 1. Observed and expected number of household with at least one male migrant according to the number of male migrants aged 15 years and above (survey 2001 data)**

Number of migrants	Observed no. of households	Expected no. of households
1	97	97.37
2	35	38.11
3	19	14.91
4	6 } 3 } 9	9.61
5		
<b>Total</b>	<b>160</b>	<b>160</b>
Mean=1.643 Variance=1.057 $\hat{p} = 0.7801$	$\chi^2 = 1.41$ (after pooling) p-value=0.4941 $\chi^2_{(2)} = 5.99$ at 5% level of significance	

**Table 2. Observed and expected number of households with at least one male migrant according to the number of male migrants aged 15 years and above (survey 1978 data)**

Number of migrants	Observe no. of households	Expected no. of households
1	375	384.32
2	143	133.97
3	49	46.70
4	17	16.28
5+	6	8.72
<b>Total</b>	<b>590</b>	<b>590</b>
Mean=1.53 Variance=0.8213 $\hat{p} = 0.8084$	$\chi^2 = 1.82$ (after pooling) p-value=0.6105 $\chi^2_{(3)} = 7.815$ at 5% level of significance	

**Table 3. Observed and expected number of households with at least one male migrant according to the number of male migrants aged 15 years and above (survey 1978 data) in three types of households.**

Number of migrants	Type of households						
	Semi urban		Remote		Growth centre		
	Observed	Expected	Observed	Expected	Observed	Expected	
1	95	89.4	176	176.4	154	152.3	
2	19	27.4	59	58.0	47	51.4	
3	10 } 2 } 3 } 15	12.2	18	19.0	18	17.3	
4			6 } 4 } 10	9.6	9	11	8.9
5					2		
<b>Total</b>	<b>129</b>	<b>129</b>	<b>263</b>	<b>263</b>	<b>230</b>	<b>230</b>	
Mean= Variance= $\chi^2$ (after pooling) = $\hat{p}$ = p-value= d.f.=	1.44 0.6374 3.56 0.8327 0.0591 1		1.49 0.7312 0.08 0.8190 0.9607 2		1.51 0.7542 0.91 0.8180 0.6344 2		

**Table 4. Observed and Expected number of households having adult Male Migrants aged 15 years and above.**

Number of migrants	Observed no. of households	Expected no. of households
1	97	98.90
2	42	40.33
3	16	16.44
4	7 } 5 } 12	11.33
5+		
<b>Total</b>	<b>167</b>	<b>167</b>
Mean=1.68 Variance=1.142 $\hat{p} = 0.7695$	$\chi^2 = 0.13$ (after pooling) p-value=0.9370 $\chi^2_{(2)} = 5.99$ at 5% level of significance	

**Table 5. 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
1	409	408.7
2	88	88.65
3	19	19.22
4	5 } 1 } 6	5.42
5		
<b>Total</b>	<b>522</b>	<b>522</b>
Mean=1.27 Variance=0.3538 $\hat{p} = 0.8849$	$\chi^2 = 0.067$ (after pooling) p-value=0.9670 $\chi^2_{(2)} = 5.99$ at 5% level of significance	

**Table 6. 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
1	71	72.90
2	32	27.86
3	7	10.64
4	5 } 3 } 8	6.59
5		
<b>Total</b>	<b>118</b>	<b>118</b>
Mean=1.618 Variance=1.001 $\hat{p} = 0.7860$	$\chi^2 = 2.19$ (after pooling) p-value=0.3345 $\chi^2_{(2)} = 5.99$ at 5% level of significance	

**Table 7. The number of mothers of the Urban Area with at least two live births by the number of infant and child deaths.**

No. of Infant and child deaths	Observed no. of mothers	Expected no. of mothers
1	176	172.06
2	44	50.72
3	16	14.95
4	6 } 8 2 }	6.27
5		
<b>Total</b>	<b>244</b>	<b>244</b>
Mean=1.41 Variance=0.5930 $\hat{p} = 0.8397$	$\chi^2 = 1.53$ (after pooling) p-value=0.4653 $\chi^2_{(2)} = 5.991$ at 5% level of significance	

**Table 8. The number of mothers of the Rural Area with at least two live births by the number of infant and child deaths.**

No. of Neonatal deaths	Observed no. of mothers	Expected no. of mothers
1	745	739.99
2	212	212.45
3	50	60.99
4	21	17.51
5	7 } 10 3 }	7.06
6		
<b>Total</b>	<b>1038</b>	<b>1038</b>
Mean=1.40 Variance=0.5651 $\hat{p} = 0.8443$	$\chi^2 = 3.92$ (after pooling) p-value=0.2702 $\chi^2_{(3)} = 7.815$ at 5% level of significance	

**Table 9. The number of literate mothers with at least one live birth by the number of infant deaths.**

No. of Neonatal deaths	Observed no. of mothers	Expected no. of mothers
1	683	675
2	145	153
3	29	34.7
4	11 } 16 5 }	10.3
5		
<b>Total</b>	<b>873</b>	<b>873</b>
Mean=1.29 Variance=0.3794 $\hat{p} = 0.8793$	$\chi^2 = 4.72$ (after pooling) p-value=0.0944 $\chi^2_{(3)} = 5.99$ at 5% level of significance	

**Table 10. The number of mothers of the completed fertility having experienced at least one child death.**

No. of child deaths	Observed no. of mothers	Expected no. of mothers
1	89	83.98
2	25	31.73
3	11	11.99
4	6 3 1 } 10	7.30
5		
6		
<b>Total</b>	<b>135</b>	<b>135</b>
Mean=1.60 Variance=0.9767 $\hat{p} = 0.7887$	$\chi^2 = 2.80$ (after pooling) p-value=0.2465 $\chi^2_{(2)} = 5.99$ at 5% level of significance	

**Table 11. The number of mothers having at least one neonatal deaths.**

No. of Neonatal deaths	Observed no. of mothers	Expected no. of mothers
1	567	562.13
2	135	138.56
3	28	34.15
4	11 5 } 16	11.15
5		
<b>Total</b>	<b>746</b>	<b>746</b>
Mean=1.32 Variance=0.4342 $\hat{p} = 0.8680$	$\chi^2 = 3.33$ (after pooling) p-value=0.1891 $\chi^2_{(2)} = 5.99$ at 5% level of significance	

**Discussion-**From table 1 to 11 it clearly indicates that the calculated value of  $\chi^2$  is less than tabulated value of  $\chi^2$ , Hence we conclude that there is no significant difference between the observed and expected value of data (i.e. Our proposed model found suitable for migration and mortality data.)

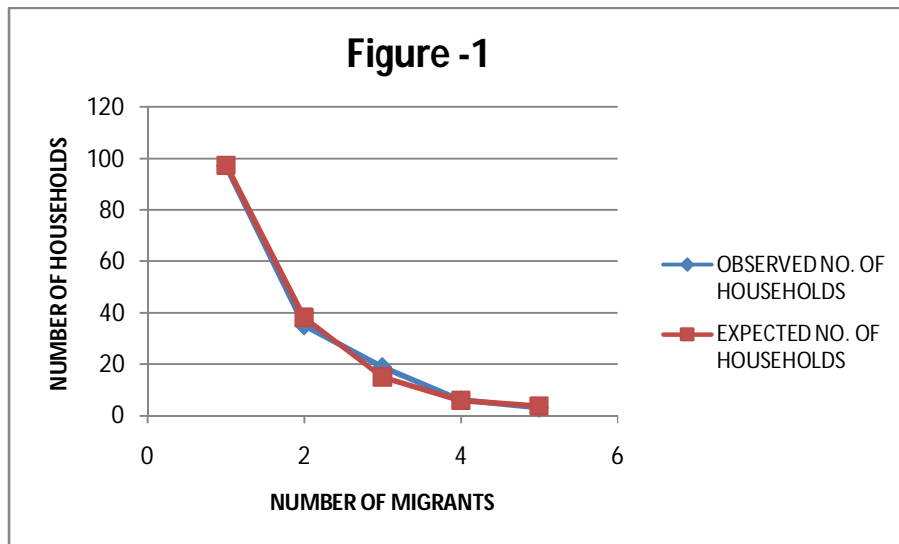


Fig.1 Graphical presentation showing expected and observed number of household aged 15 years and above (survey 2001 data)

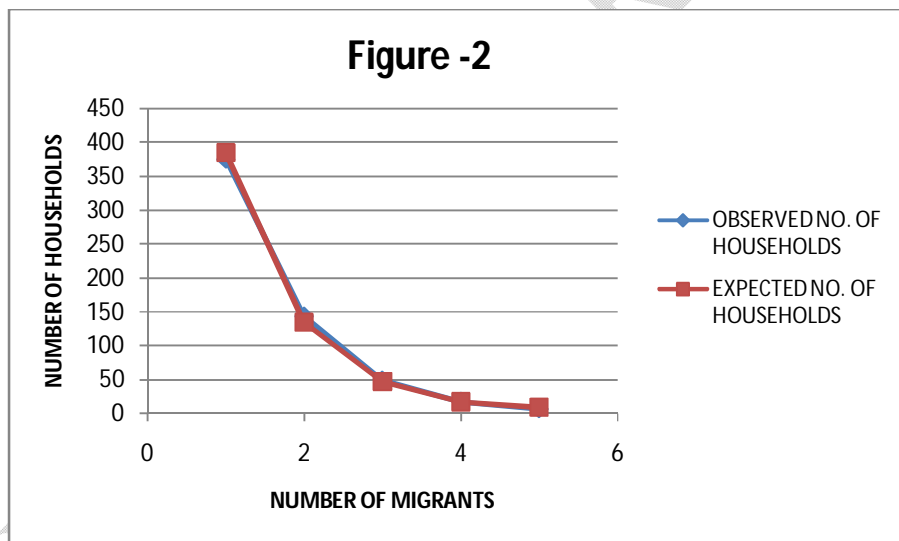


Fig.2 Graphical presentation showing expected and observed number of household aged 15 years and above (survey 1978 data)

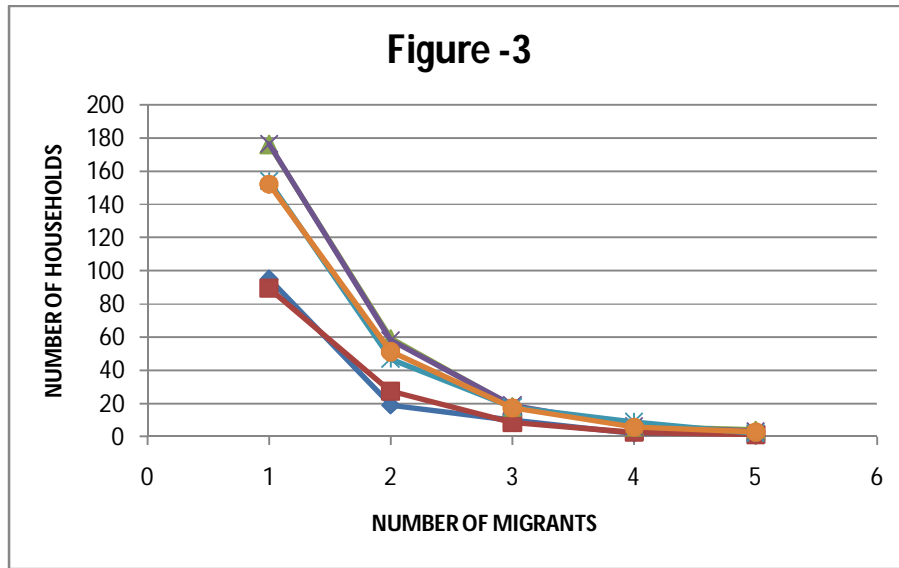


Fig.3 Graphical presentation showing expected and observed number of household aged 15 years and above (survey 1978 data) in three types of households.

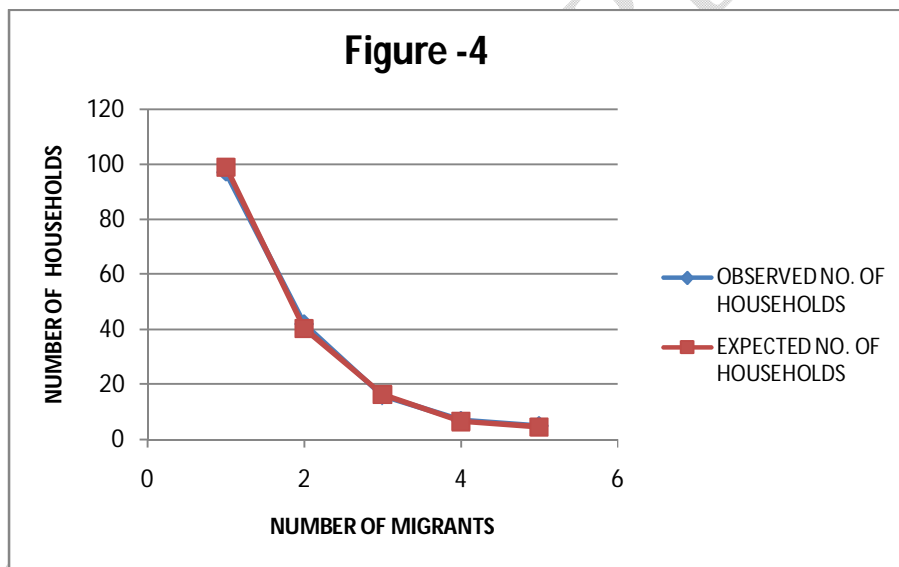


Fig.4 Graphical presentation showing Observed and Expected number of households having adult Male Migrants aged 15 years and above

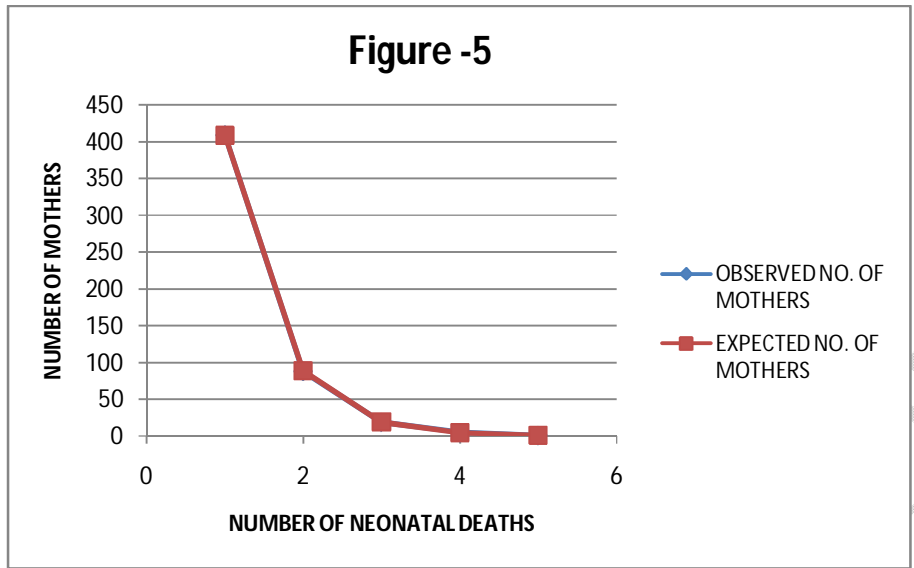


Fig.5 Graphical presentation showing Rural Area having at least one live birth and one neonatal death.

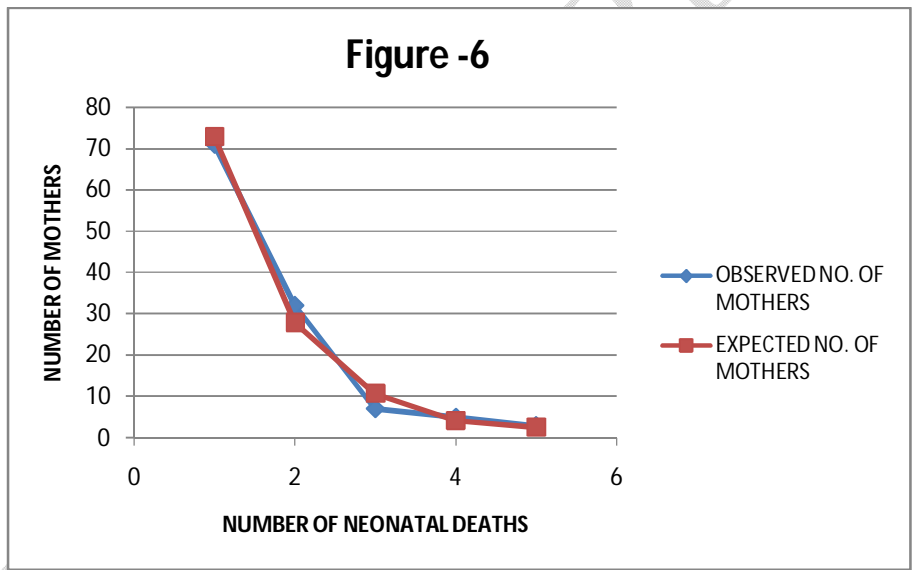


Fig.6 Graphical presentation showing Estate Area having at least one live birth and one neonatal death

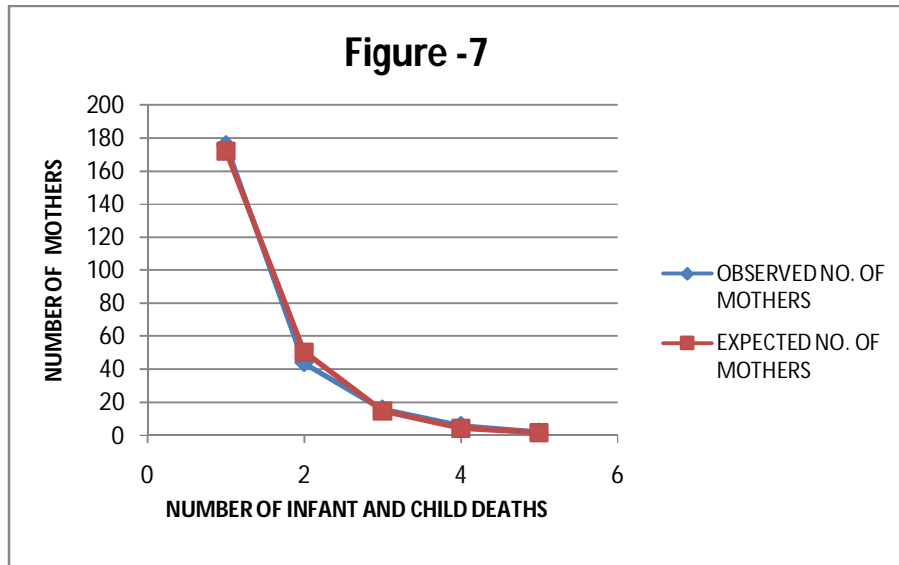


Fig.7 Graphical presentation showing Urban Area with at least two live births by the number of infant and child deaths

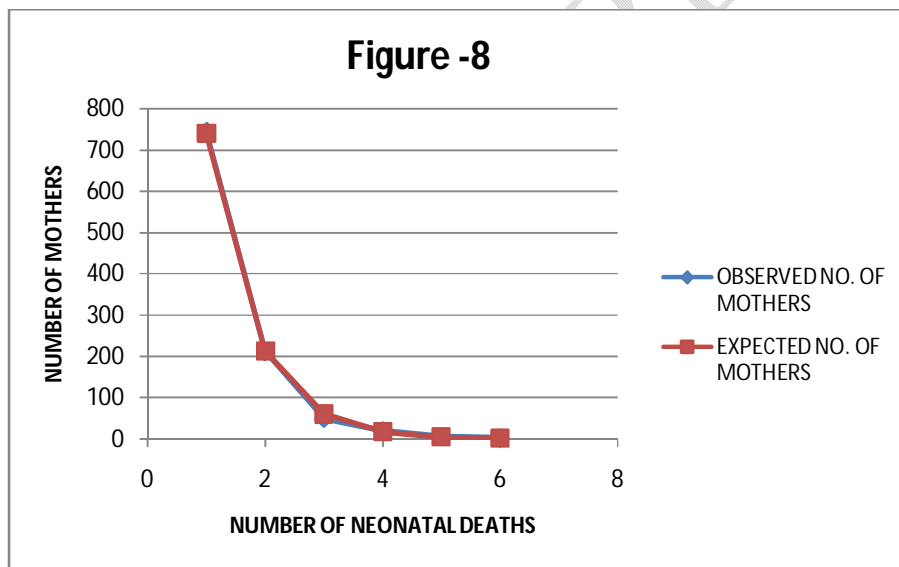


Fig 8 Graphical presentation showing Rural Area with at least two live births by the number of infant and child deaths.

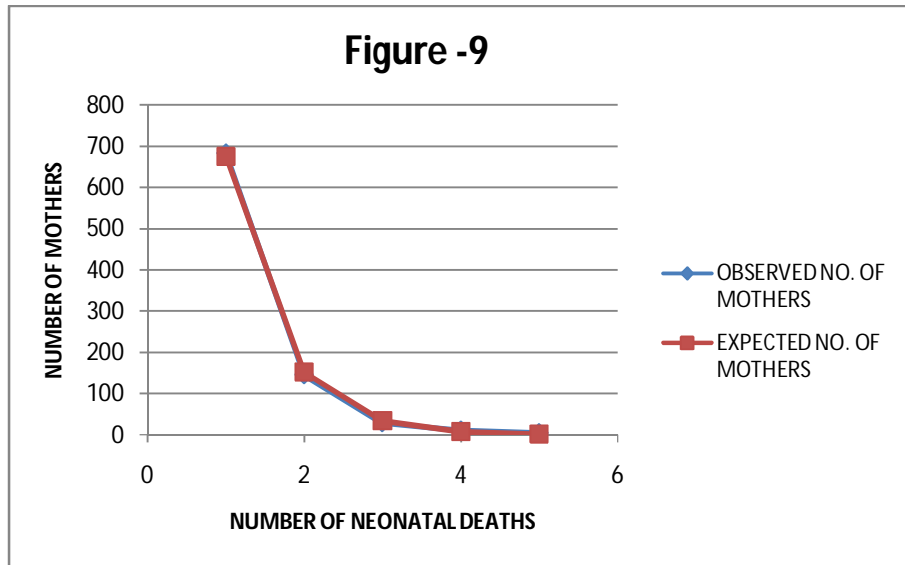


Fig.9 Graphical presentation showing literate mothers with at least one live birth by the number of infant deaths.

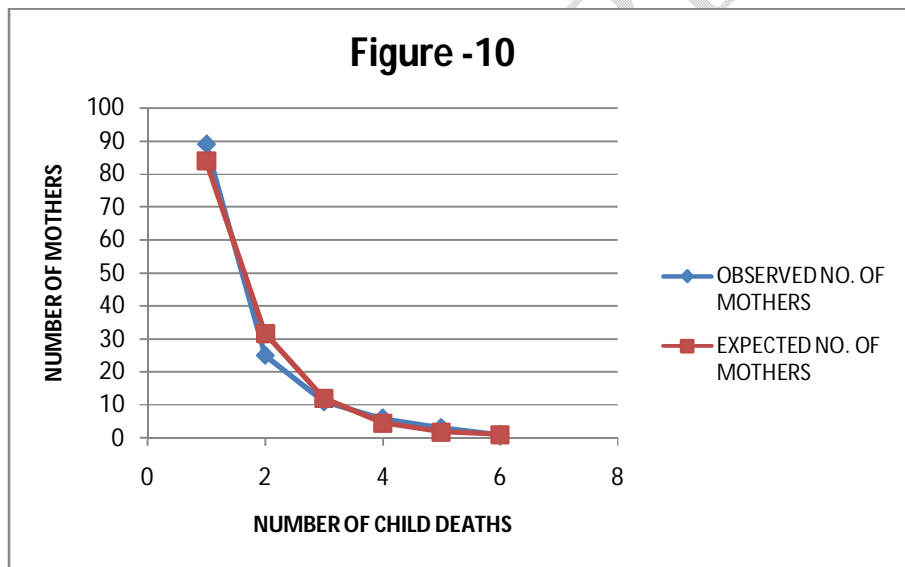


Fig.10 Graphical presentation showing mothers of the completed fertility having experienced at least one child death

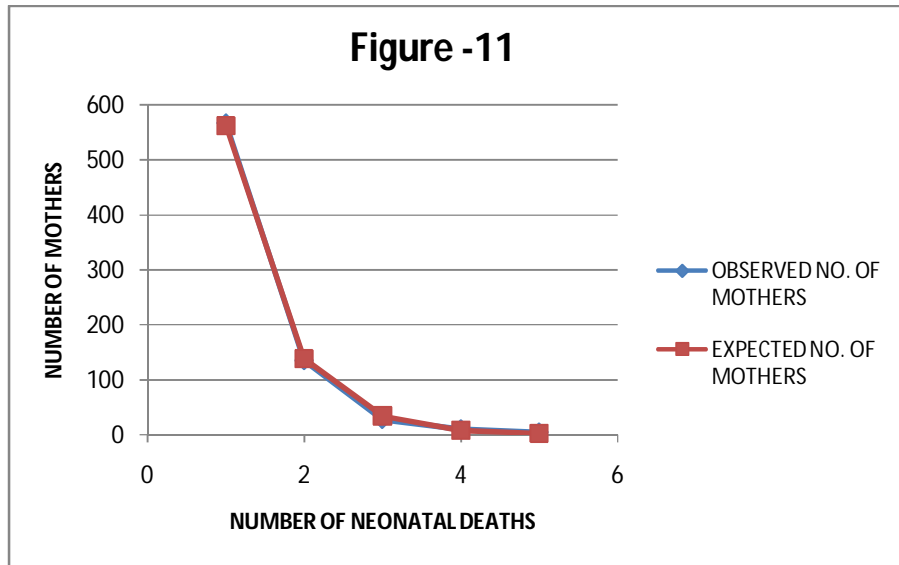


Fig.11 Graphical presentation showing mothers having at least one neonatal deaths

UNDER PEER REVIEW

## 5. Conclusion

According to the value of  $\chi^2$  and p-value From (table 1,2,3,4,5,6,7,8,9,10,11) and graphical representation between  $O_i$  and  $E_i$ , the nature and behavior of proposed ZTHD Model found suitable for the migration pattern as well as infant mortality pattern of different Regions.

The overall studies shows that the proposed ZTHD Model could also be helpful in policy making, Rural development, Fresh Environment, Medical Facilities for the betterment of the society. In the light of work done by Rao and Pandey (2017,2019), it is also possible for the Bayesian Analysis of the proposed model.

## Reference

- [1].Agarwal A.,Pandey H.,Himanshu distribution and its applications,Bulletin of Mathematics and Statistics Research.2022;10(4):24-35.
- [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].Lal DN (1955) Patna in 1955, A Demographic Sample Survey, Demographic Research Centre, Department of Statistics, Patna University, India.
- [4].Meegama SA (1980) Socio-economic determinants of infant and child mortality in Sri Lanka, An analysis of post war experience, International Statistical Institute (World Fertility Survey), Netherland.
- [5].Pandey H. etal., A displaced probability model for the child mortality in a family, International Journal of Statistics and Applications.2014;4(2):89-95.
- [6].Pandey H.,Jai kishun, A probability model for the child mortality in a family, Revista Colombiana de Estadística.2010;33(1):1-11.
- [7].Rao A.K etal., Bayesian estimation of the parameter of the p-dimensional size biased Rayleigh distribution, Lietuvos Statistikos Darbai.2017;56(1):88-91.
- [8].Rao, Pandey H., Parameter estimation of Nakagami distribution under precautionary loss function, South East Asian journal of Mathematics and Mathematical Sciences.2019;16(2):12-18.
- [9].Shukla K.K, Shankar R., Tiwari M., Zero truncated poisson Ishita distribution and its application, Journal of Scientific Research.2020;64(2):287-294.
- [10]. Shankar R., Ishita distribution and its application to model lifetime data, Biometrics and Biostatistics International Journal.2017;5(2):1-9.

- [11]. Shukla K.K,Shankar R.,A discrete poisson Ishita distribution and its application,International Journal of Statistics and Economics.2019;20(2):109-122.
- [12]. Rashid M.,Chesneau C.,Sibu D.,Monisha M.,Maya R.,Lagrangian Zero truncated poisson distribution: Properties,Regression Model and Applications.Symmetry.2022;14(9):1775-1785.
- [13]. Shukla K.K,Shankar R.,A zero truncated two parameter poisson Lindley distribution with an application to biological sciences.Turkiye Klinikleri journal of Biostatistics2017;9(2):85-95.
- [14]. Shanker R, Hagos F, Sujatha S, Abrehe Y. On zero-truncation of Poisson and Poisson-Lindley distributions and their applications. Biometrics & Biostatistics International Journal. 2015;2(6):1-4.