Rama Shanker – Kamlesh Kumar Shukla

A generalization of zero-truncated Poisson-Sujatha distribution*

RAMA SHANKER
Associate Professor
Department of Statistics, Assam University
Silchar,
India
Email: shankerrama2009@gmail.com

KAMLESH KUMAR SHUKLA (corresponding author)
Associate Professor
Department of Statistics, College of Science
Mainefhi, Asmara,
Eritrea
Email: kkshukla22@gmail.com

In this study, the authors propose a generalization of the zero-truncated Poisson-Sujatha distribution that includes both the zero-truncated Poisson-Lindley and the zero-truncated Poisson-Sujatha distributions as special cases. The statistical properties based on moments, including behaviours due to the coefficient of variation, skewness, kurtosis, and the index of dispersion are studied. The estimation of parameters using maximum likelihood is carried out. Finally, the goodness of fit of the proposed distribution is presented with four datasets.

KEYWORDS: moments-based measures, maximum likelihood estimation, goodness of fit

Assuming $P_0(x;\,\theta)$, where θ is the scale parameter, as the pmf (probability mass function) of the original distribution, the zero-truncated version of $P_0(x;\,\theta)$ can be defined as

$$P(x;\,\theta) = \frac{P_0(x;\,\theta)}{1 - P_0(0;\,\theta)};\ x = 1,\ 2,\ 3,\ldots \qquad /1/$$

Shanker [2016b] proposed PSD (Poisson-Sujatha distribution) defined by its pmf:

$$P_0(x;\,\theta) = \frac{\theta^3}{\theta^2 + \theta + 2} \cdot \frac{x^2 + (\theta + 4)x + (\theta^2 + 3\theta + 4)}{(\theta + 1)^{x+3}};\ x = 0,\ 1,\ 2,\ 3,\ldots,\ \theta > 0. \qquad /2/$$

* The authors are thankful to the Editor-in-Chief as well as the anonymous reviewers for the fruitful comments given to improve the quality of the paper.

Statistical properties, the estimation of parameters using both the method of moments and the method of maximum likelihood, and the various applications of PSD are explained in *Shanker* [2016b]. The PSD is a Poisson mixture of the Sujatha distribution, wherein the mean of the Poisson distribution λ follows the Sujatha distribution introduced by *Shanker* [2016a], and is defined by its pdf (probability density function):

$$f_1(\lambda;\,\theta) = \frac{\theta^3}{\theta^2+\theta+2}\left(1+\lambda+\lambda^2\right)e^{-\theta\lambda};\ \lambda>0,\ \theta>0. \qquad /3/$$

Shanker [2016a] discussed various statistical properties, the estimation of the parameter, and the applications of the Sujatha distribution for modelling lifetime data from various fields of knowledge. It has been observed that the Sujatha distribution yields a much closer fit than both the exponential distribution and the Lindley distribution, introduced by *Lindley* [1958] and studied by *Ghitany–Atieh–Nadarajah* [2008a].

A ZTPSD (zero-truncated Poisson-Sujatha distribution) using /1/ and /2/, introduced by *Shanker* and *Hagos* [2015a], is defined by its pmf:

$$P_1(x;\,\theta) = \frac{\theta^3\left\{x^2+(\theta+4)x+\left(\theta^2+3\theta+4\right)\right\}}{\left(\theta^4+4\theta^3+10\theta^2+7\theta+2\right)(\theta+1)^x};\ x=1,\,2,\,3,\ldots,\ \theta>0. \qquad /4/$$

Shanker–Shukla–Hagos [2017] have introduced AGSD (a generalization of the Sujatha distribution) having scale parameter θ, shape parameter α, and pdf described by

$$f_2(x;\,\theta,\,\alpha) = \frac{\theta^3}{\theta^2+\theta+2\alpha}\left(1+x+\alpha x^2\right)e^{-\theta x};\ x>0,\ \theta>0,\ \alpha>0. \qquad /5/$$

We can easily demonstrate that the Sujatha distribution and the Lindley distribution introduced by *Lindley* [1958] are particular cases of AGSD in /5/ for $\alpha = 1$ and $\alpha = 0$, respectively.

Assuming that the parameter λ of the Poisson distribution follows AGSD /5/, the pmf of AGPSD (a generalization of the Poisson-Sujatha distribution) obtained by *Shanker* and *Shukla* [2019] is described by

$$P_2(x;\,\theta,\alpha) = \frac{\theta^3\left\{\alpha x^2+(\theta+3\alpha+1)x+\left(\theta^2+3\theta+2\alpha+2\right)\right\}}{\left(\theta^2+\theta+2\alpha\right)(\theta+1)^{x+3}}; \qquad /6/$$

$$x=1,\,2,\,3,\ldots,\ \theta>0,\ \alpha>0.$$

An ZTPLD (zero-truncated Poisson-Lindley distribution) introduced by *Ghitany–Al-Mutairi–Nadarajah* [2008b] is defined by its pmf:

$$P_3(x;\ \theta) = \frac{\theta^2}{\theta^2 + 3\theta + 1} \frac{x + \theta + 2}{(\theta + 1)^x};\ x = 1,\ 2,\ 3,....,\ \theta > 0. \qquad /7/$$

Note that the ZTPLD is a zero-truncated version of the PLD (Poisson-Lindley distribution) proposed by *Sankaran* [1970]. Sankaran obtained the PLD by compounding the Poisson distribution with the Lindley distribution. A PLD is defined by its pmf:

$$P_4(x;\ \theta) = \frac{\theta^2 (x + \theta + 2)}{(\theta + 1)^{x+3}};\ x = 0,\ 1,\ 2,...,\ \theta > 0. \qquad /8/$$

Shanker and *Hagos* [2015b] have discussed various applications of PLDs in the biological sciences.

In this research work, AGZTPSD (a generalization of the zero-truncated Poisson-Sujatha distribution), of which ZTPLD and ZTPSD are particular cases, has been obtained using the definition of zero-truncation, and by compounding the SBPD (size-biased Poisson distribution) with an assumed continuous distribution. Its raw moments and central moments have been provided. The expressions for the CV (coefficient of variation), skewness, kurtosis, and the index of dispersion were obtained and their behaviours are discussed graphically. Maximum likelihood estimation is discussed as a means to estimate the parameters of AGZTPSD. The goodness of fit of AGZTPSD is discussed with four datasets and the fit is compared with the ZTPD (zero-truncated Poisson distribution), ZTPLD, and the ZTPSD.

1. A generalization of zero-truncated Poisson-Sujatha distribution

Using /1/ and /6/, the pmf of AGZTPSD can be obtained by

$$P(x;\ \theta,\ \alpha) = \frac{\theta^3 \left\{\alpha x^2 + (\theta + 3\alpha + 1) x + \left(\theta^2 + 3\theta + 2\alpha + 2\right)\right\}}{\left\{\theta^4 + 4\theta^3 + (6\alpha + 4)\theta^2 + (6\alpha + 1)\theta + 2\alpha\right\}(\theta + 1)^x}; \qquad /9/$$

$$x = 1,\ 2,\ 3,...,\ \theta > 0,\ \alpha > 0.$$

It can be easily verified that both the ZTPLD and ZTPSD are particular cases of AGZTPSD for $\alpha = 0$ and $\alpha = 1$, respectively. It should be noted that many statistical properties of a distribution are based on moments. However, it is very difficult and complicated to obtain the moments of AGZTPSD directly. Therefore, to obtain the moments of AGZTPSD, we attempt to derive the pmf of AGZTPSD in a very simple way. AGZTPSD can also be obtained by compounding the SBPD that has a pmf described by

$$g(x|\lambda) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!};\ x = 1, 2, 3, \ldots, \lambda > 0. \qquad /10/$$

Suppose in Equation /10/ the parameter λ of the SBPD follows a continuous distribution with a pdf described by

$$h(\lambda|\theta, \alpha) = \frac{\theta^3\left\{\alpha(\theta+1)^2\lambda^2 + (\theta+1)(\theta+2\alpha+1)\lambda + \left(\theta^2+3\theta+2\alpha+2\right)\right\}e^{-\theta\lambda}}{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha}; \qquad /11/$$

$\lambda > 0,\ \theta > 0,\ \alpha > 0.$

The pmf of AGZTPSD is thus obtained as

$$P(x;\theta,\alpha) = \int_0^\infty g(x|\lambda)\cdot h(\lambda|\theta,\alpha)\,d\lambda$$

$$= \int_0^\infty \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\cdot\frac{\theta^3\left\{\alpha(\theta+1)^2\lambda^2 + (\theta+1)(\theta+2\alpha+1)\lambda + \left(\theta^2+3\theta+2\alpha+2\right)\right\}e^{-\theta\lambda}}{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha}d\lambda \qquad /12/$$

$$= \frac{\theta^3}{\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}(x-1)!}\times$$

$$\times\int_0^\infty e^{-(\theta+1)\lambda}\lambda^{x-1}\cdot\left\{\alpha(\theta+1)^2\lambda^2+(\theta+1)(\theta+2\alpha+1)\lambda+\left(\theta^2+3\theta+2\alpha+2\right)\right\}d\lambda$$

$$= \frac{\theta^3}{\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}(x-1)!}\times$$

$$\times\left[\begin{array}{l}\alpha(\theta+1)^2\int_0^\infty e^{-(\theta+1)\lambda}\lambda^{x+2-1}d\lambda + (\theta+1)(\theta+2\alpha+1)\int_0^\infty e^{-(\theta+1)\lambda}\lambda^{x+1-1}d\lambda + \\ +\left(\theta^2+3\theta+2\alpha+2\right)\int_0^\infty e^{-(\theta+1)\lambda}\lambda^{x-1}d\lambda\end{array}\right]$$

$$= \frac{\theta^3}{\left\{\theta^4 + 4\theta^3 + (6\alpha + 4)\theta^2 + (6\alpha + 1)\theta + 2\alpha\right\}(x-1)!} \times$$

$$\times \left[\frac{\alpha(\theta+1)^2 \Gamma(x+2)}{(\theta+1)^{x+2}} + (\theta+1)\frac{(\theta+2\alpha+1)\Gamma(x+2)}{(\theta+1)^{x+1}} + \frac{(\theta^2+3\theta+2\alpha+2)\Gamma(x)}{(\theta+1)^x} \right]^1$$

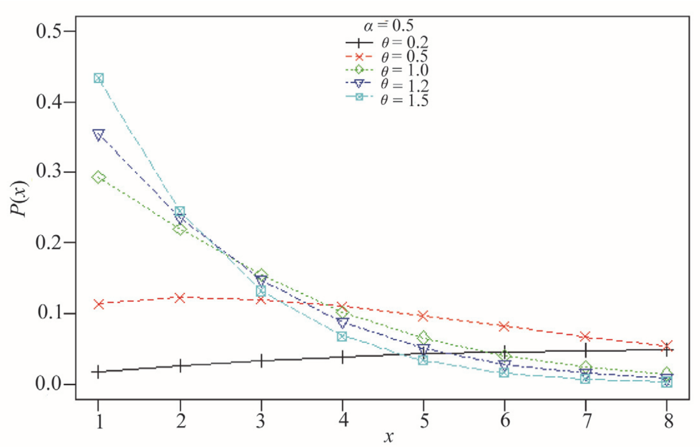
$$= \frac{\theta^3}{\left\{\theta^4 + 4\theta^3 + (6\alpha + 4)\theta^2 + (6\alpha + 1)\theta + 2\alpha\right\}} \times$$

$$\times \left[\frac{\alpha(x+1)x}{(\theta+1)^x} + \frac{(\theta+2\alpha+1)x}{(\theta+1)^x} + \frac{(\theta^2+3\theta+2\alpha+2)}{(\theta+1)^x} \right]$$

$$= \frac{\theta^3\left\{\alpha x^2 + (\theta + 3\alpha + 1)x + (\theta^2 + 3\theta + 2\alpha + 2)\right\}}{\left\{\theta^4 + 4\theta^3 + (6\alpha + 4)\theta^2 + (6\alpha + 1)\theta + 2\alpha\right\}(\theta+1)^x};\ x = 1, 2, 3, ..., \theta > 0, \alpha > 0.$$

The former equation describes the pmf of AGZTPSD with parameters θ and α, as obtained in /9/. The behaviour of the pmf of AGZTPSD for varying values of parameters is shown graphically in Figure 1.

Figure 1. Probability mass function, $P(x)$ of AGZTPSD for values of parameters θ and α



(Continued on the next page)

[1] $\Gamma(x)$ is a complete gamma function.

(Continued)

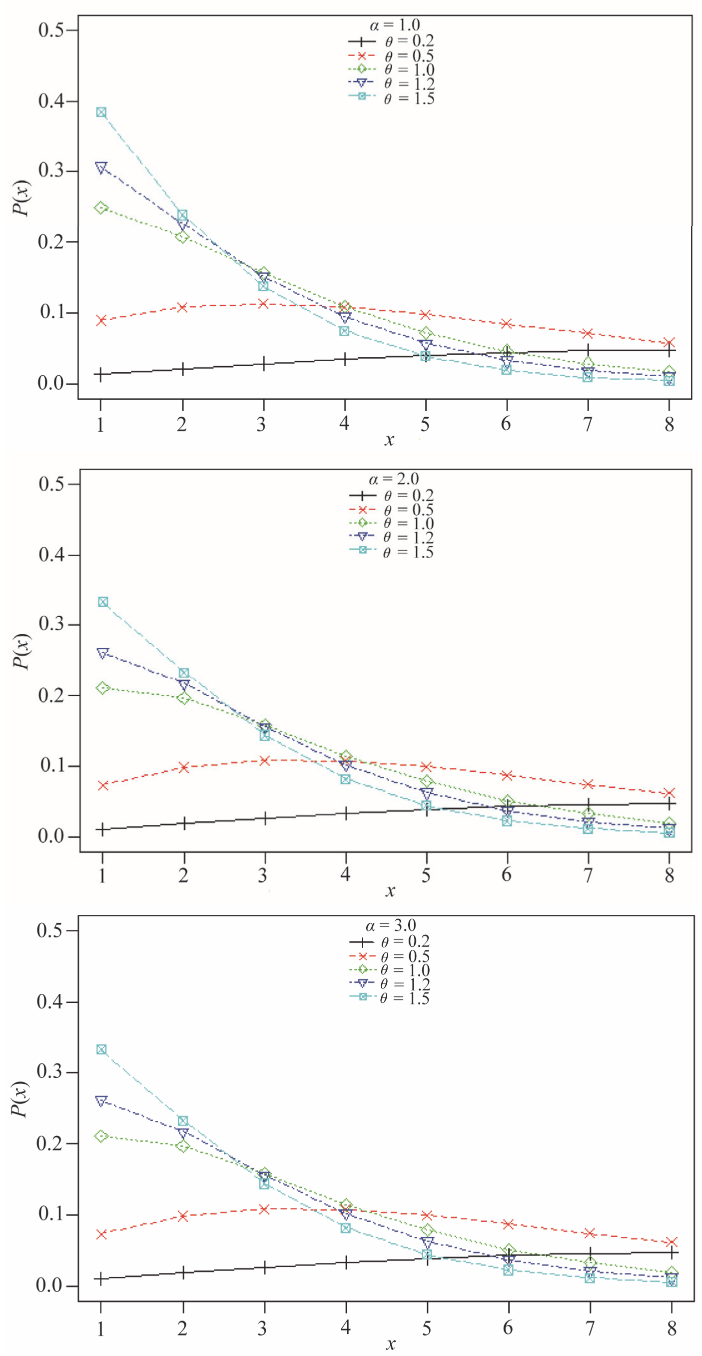

α = 1.0
θ = 0.2
θ = 0.5
θ = 1.0
θ = 1.2
θ = 1.5
P(x)
x
α = 2.0
θ = 0.2
θ = 0.5
θ = 1.0
θ = 1.2
θ = 1.5
P(x)
x
α = 3.0
θ = 0.2
θ = 0.5
θ = 1.0
θ = 1.2
θ = 1.5
P(x)
x

Figure 1 shows that pmf of AGZTPSD increases as the value of θ increases, whereas it slightly decreases as the value of α increases.

2. Moments

Using /12/ and expectation operator E, the r^{th} factorial moment about the origin $\mu_{(r)}'$ of AGZTPSD /9/ can be obtained as

$$\mu_{(r)}' = E\left[E\left(X^{(r)} \mid \lambda \right) \right],$$

where $X^{(r)} = X(X-1)(X-2)...(X-r+1)$.

$$\mu_{(r)}' = \int_0^\infty \left[\sum_{x=1}^\infty x^{(r)} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \right] \cdot$$

$$\cdot \frac{\theta^3 \left\{ \alpha(\theta+1)^2 \lambda^2 + (\theta+1)(\theta+2\alpha+1)\lambda + \left(\theta^2 + 3\theta + 2\alpha + 2\right) \right\} e^{-\theta\lambda}}{\theta^4 + 4\theta^3 + (6\alpha+4)\theta^2 + (6\alpha+1)\theta + 2\alpha} d\lambda$$

$$= \int_0^\infty \left[\lambda^{r-1} \sum_{x=r}^\infty x \frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!} \right] \cdot$$

$$\cdot \frac{\theta^3 \left\{ \alpha(\theta+1)^2 \lambda^2 + (\theta+1)(\theta+2\alpha+1)\lambda + \left(\theta^2 + 3\theta + 2\alpha + 2\right) \right\} e^{-\theta\lambda}}{\theta^4 + 4\theta^3 + (6\alpha+4)\theta^2 + (6\alpha+1)\theta + 2\alpha} d\lambda.$$

Assuming $y = x - r$, we get

$$\mu_{(r)}' = \int_0^\infty \left[\lambda^{r-1} \sum_{y=0}^\infty (y+r) \frac{e^{-\lambda}\lambda^{y}}{y!} \right] \cdot$$

$$\cdot \frac{\theta^3 \left\{ \alpha(\theta+1)^2 \lambda^2 + (\theta+1)(\theta+2\alpha+1)\lambda + \left(\theta^2 + 3\theta + 2\alpha + 2\right) \right\} e^{-\theta\lambda}}{\theta^4 + 4\theta^3 + (6\alpha+4)\theta^2 + (6\alpha+1)\theta + 2\alpha} d\lambda$$

$$= \int_0^\infty \lambda^{r-1}(\lambda + r)\frac{\theta^3\left\{\alpha(\theta+1)^2\lambda^2 + (\theta+1)(\theta+2\alpha+1)\lambda + \left(\theta^2+3\theta+2\alpha+2\right)\right\}e^{-\theta\lambda}}{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha}d\lambda.$$

By using the gamma integral and some algebraic simplification, we obtain the expression for the r^{th} factorial moment about the origin of AGZTPSD as

$$\mu_{(r)}' = \frac{r!(\theta+1)\left\{\begin{array}{l}\theta^4+(r+3)\theta^3+\left(\alpha r^2+3\alpha r+2\alpha+2r+3\right)\theta^2+\left(2\alpha r^2+6\alpha r+4\alpha+r+1\right)\theta \\ +\left(\alpha r^2+3\alpha r+2\alpha\right)\end{array}\right\}}{\theta^r\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}}; \quad /13/$$

$r = 1, 2, 3, \ldots$

By substituting $r = 1, 2, 3, \text{ and } 4$ in Equation /13/, the first four factorial moments about the origin are obtained as

$$\mu_{(1)}' = \frac{(\theta+1)\left\{\theta^4+4\theta^3+(6\alpha+5)\theta^2+(12\alpha+2)\theta+6\alpha\right\}}{\theta\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}},$$

$$\mu_{(2)}' = \frac{2(\theta+1)\left\{\theta^4+5\theta^3+(12\alpha+7)\theta^2+(24\alpha+3)\theta+12\alpha\right\}}{\theta^2\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}},$$

$$\mu_{(3)}' = \frac{6(\theta+1)\left\{\theta^4+6\theta^3+(20\alpha+9)\theta^2+(40\alpha+4)\theta+20\alpha\right\}}{\theta^3\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}},$$

$$\mu_{(4)}' = \frac{24(\theta+1)\left\{\theta^4+7\theta^3+(40\alpha+11)\theta^2+(60\alpha+5)\theta+30\alpha\right\}}{\theta^4\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}}.$$

By using the relationship between the moments about the origin and the factorial moments about the origin, the first four moments about the origin of the AGZTPSD can be obtained as

$$\mu_1' = \frac{(\theta+1)\left\{\theta^4+4\theta^3+(6\alpha+5)\theta^2+(12\alpha+2)\theta+6\alpha\right\}}{\theta\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}},$$

$$\mu_2' = \frac{(\theta+1)\left\{\theta^5+6\theta^4+(6\alpha+15)\theta^3+(36\alpha+16)\theta^2+(54\alpha+6)\theta+24\alpha\right\}}{\theta^2\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}},$$

$$\mu_3' = \\ = \frac{(\theta+1)\left\{\theta^6+10\theta^5+(6\alpha+41)\theta^4+(84\alpha+80)\theta^3+(270\alpha+72)\theta^2+(312\alpha+24)\theta+120\alpha\right\}}{\theta^3\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}},$$

$$\mu_4' = \\ = \frac{(\theta+1)\left\{\begin{matrix}\theta^7+18\theta^6+(6\alpha+111)\theta^5+(180\alpha+340)\theta^4+(1062\alpha+534)\theta^3+(2568\alpha+408)\theta^2\\ +(2160\alpha+120)\theta+720\alpha\end{matrix}\right\}}{\theta^4\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}}.$$

Again, using the relationship between the moments about the mean and moments about the origin, the moments about the mean of AGZTPSD are obtained as

$$\mu_2 = \sigma^2 = \\ = \frac{(\theta+1)\left\{\begin{matrix}\theta^8+9\theta^7+(18\alpha+31)\theta^6+(110\alpha+51)\theta^5+(72\alpha^2+228\alpha+42)\theta^4\\ +(192\alpha^2+210\alpha+16)\theta^3+(180\alpha^2+86\alpha+2)\theta^2+(72\alpha^2+12\alpha)\theta+12\alpha^2\end{matrix}\right\}}{\theta^2\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}^2},$$

$$\mu_3 = \frac{(\theta+1)\left\{\begin{array}{l}\theta^{13}+15\theta^{12}+(24\alpha+99)\theta^{11}+(310\alpha+372)\theta^{10}+\left(180\alpha^2+1670\alpha+867\right)\theta^9\\ +\left(1776\alpha^2+4846\alpha+1293\right)\theta^8+\left(432\alpha^3+6780\alpha^2+8292\alpha+1237\right)\theta^7\\ +\left(2736\alpha^3+13144\alpha^2+8666\alpha+742\right)\theta^6+\left(6264\alpha^3+14380\alpha^2+5530\alpha+264\right)\theta^5\\ +\left(7224\alpha^3+9260\alpha^2+2078\alpha+50\right)\theta^4+\left(4800\alpha^3+3524\alpha^2+420\alpha+4\right)\theta^3\\ +\left(1944\alpha^3+756\alpha^2+36\alpha\right)\theta^2+\left(456\alpha^3+72\alpha^2\right)\theta+48\alpha^3\end{array}\right\}}{\theta^3\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}^3},$$

$$\mu_4 = \frac{(\theta+1)\left\{\begin{array}{l}\theta^{18}+26\theta^{17}+(30\alpha+298)\theta^{16}+(770\alpha+1986)\theta^{15}+\left(324\alpha^2+8276\alpha+857\right)\theta^{14}\\ +\left(7776\alpha^2+49174\alpha+25297\right)\theta^{13}+\left(1512\alpha^3+73104\alpha^2+180846\alpha+52622\right)\theta^{12}\\ +\left(32616\alpha^3+358076\alpha^2+437004\alpha+78360\right)\theta^{11}\\ +\left(2592\alpha^4+253224\alpha^3+1034472\alpha^2+718238\alpha+83924\right)\theta^{10}\\ +\left(48384\alpha^4+946176\alpha^3+1891400\alpha^2+818562\alpha+64295\right)\theta^9\\ +\left(289008\alpha^4+1979376\alpha^3+2289980\alpha^2+651612\alpha+34646\right)\theta^8\\ +\left(725184\alpha^4+2569904\alpha^3+1886072\alpha^2+360406\alpha+12716\right)\theta^7\\ +\left(1021824\alpha^4+2200632\alpha^3+1066976\alpha^2+135494\alpha+3002\right)\theta^6\\ +\left(912432\alpha^4+1278744\alpha^3+410292\alpha^2+32948\alpha+408\right)\theta^5\\ +\left(544176\alpha^4+504680\alpha^3+103016\alpha^2+4656\alpha+24\right)\theta^4\\ +\left(219168\alpha^4+130848\alpha^3+15264\alpha^2+288\alpha\right)\theta^3\\ +\left(58128\alpha^4+20304\alpha^3+1008\alpha^2\right)\theta^2\\ +\left(9360\alpha^4+1440\alpha^3\right)\theta+720\alpha^4\end{array}\right\}}{\theta^4\left\{\theta^4+4\theta^3+(6\alpha+4)\theta^2+(6\alpha+1)\theta+2\alpha\right\}^4}.$$

The CV, coefficient of skewness $\left(\sqrt{\beta_1}\right)$, coefficient of kurtosis $\left(\beta_2\right)$, and the index of dispersion $\left(\gamma\right)$ of AGZTPSD are obtained as

$$CV = \frac{\sigma}{\mu_1'} = $$

$$= \frac{\sqrt{\left\{\begin{array}{l}\theta^8 + 9\theta^7 + \left(18\alpha + 31\right)\theta^6 + \left(110\alpha + 51\right)\theta^5 + \left(72\alpha^2 + 228\alpha + 42\right)\theta^4 \\ + \left(192\alpha^2 + 210\alpha + 16\right)\theta^3 + \left(180\alpha^2 + 86\alpha + 2\right)\theta^2 + \left(72\alpha^2 + 12\alpha\right)\theta + 12\alpha^2\end{array}\right\}}}{\sqrt{\left(\theta + 1\right)}\left\{\theta^4 + 4\theta^3 + \left(6\alpha + 5\right)\theta^2 + \left(12\alpha + 2\right)\theta + 6\alpha\right\}},$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\left(\mu_2\right)^{3/2}} = $$

$$= \frac{\left\{\begin{array}{l}\theta^{13} + 15\theta^{12} + \left(24\alpha + 99\right)\theta^{11} + \left(310\alpha + 372\right)\theta^{10} + \left(180\alpha^2 + 1670\alpha + 867\right)\theta^9 \\ + \left(1776\alpha^2 + 4846\alpha + 1293\right)\theta^8 + \left(432\alpha^3 + 6780\alpha^2 + 8292\alpha + 1237\right)\theta^7 \\ + \left(2736\alpha^3 + 13144\alpha^2 + 8666\alpha + 742\right)\theta^6 + \left(6264\alpha^3 + 14380\alpha^2 + 5530\alpha + 264\right)\theta^5 \\ + \left(7224\alpha^3 + 9260\alpha^2 + 2078\alpha + 50\right)\theta^4 + \left(4800\alpha^3 + 3524\alpha^2 + 420\alpha + 4\right)\theta^3 \\ + \left(1944\alpha^3 + 756\alpha^2 + 36\alpha\right)\theta^2 + \left(456\alpha^3 + 72\alpha^2\right)\theta + 48\alpha^3\end{array}\right\}}{\sqrt{\left(\theta + 1\right)}\left\{\begin{array}{l}\theta^8 + 9\theta^7 + \left(18\alpha + 31\right)\theta^6 + \left(110\alpha + 51\right)\theta^5 + \left(72\alpha^2 + 228\alpha + 42\right)\theta^4 \\ + \left(192\alpha^2 + 210\alpha + 16\right)\theta^3 + \left(180\alpha^2 + 86\alpha + 2\right)\theta^2 + \left(72\alpha^2 + 12\alpha\right)\theta + 12\alpha^2\end{array}\right\}^{3/2}},$$

$$\beta_2 = \frac{\mu_4}{\mu_2^{\,2}} =$$

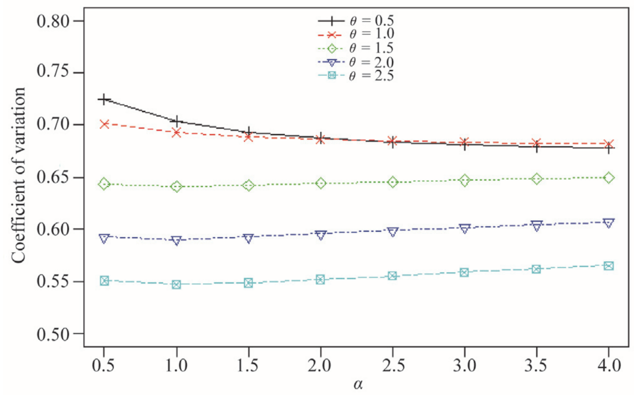
$$= \frac{\left\{\begin{array}{l} \theta^{18} + 26\theta^{17} + \left(30\alpha + 298\right)\theta^{16} + \left(770\alpha + 1986\right)\theta^{15} + \left(324\alpha^2 + 8276\alpha + 857\right)\theta^{14} \\ + \left(7776\alpha^2 + 49174\alpha + 25297\right)\theta^{13} + \left(1512\alpha^3 + 73104\alpha^2 + 180846\alpha + 52622\right)\theta^{12} \\ + \left(32616\alpha^3 + 358076\alpha^2 + 437004\alpha + 78360\right)\theta^{11} \\ + \left(2592\alpha^4 + 253224\alpha^3 + 1034472\alpha^2 + 718238\alpha + 83924\right)\theta^{10} \\ + \left(48384\alpha^4 + 946176\alpha^3 + 1891400\alpha^2 + 818562\alpha + 64295\right)\theta^{9} \\ + \left(289008\alpha^4 + 1979376\alpha^3 + 2289980\alpha^2 + 651612\alpha + 34646\right)\theta^{8} \\ + \left(725184\alpha^4 + 2569904\alpha^3 + 1886072\alpha^2 + 360406\alpha + 12716\right)\theta^{7} \\ + \left(1021824\alpha^4 + 2200632\alpha^3 + 1066976\alpha^2 + 135494\alpha + 3002\right)\theta^{6} \\ + \left(912432\alpha^4 + 1278744\alpha^3 + 410292\alpha^2 + 32948\alpha + 408\right)\theta^{5} \\ + \left(544176\alpha^4 + 504680\alpha^3 + 103016\alpha^2 + 4656\alpha + 24\right)\theta^{4} \\ + \left(219168\alpha^4 + 130848\alpha^3 + 15264\alpha^2 + 288\alpha\right)\theta^{3} \\ + \left(58128\alpha^4 + 20304\alpha^3 + 1008\alpha^2\right)\theta^{2} \\ + \left(9360\alpha^4 + 1440\alpha^3\right)\theta + 720\alpha^4 \end{array}\right\}}{\left(\theta + 1\right)\left\{\begin{array}{l} \theta^8 + 9\theta^7 + \left(18\alpha + 31\right)\theta^6 + \left(110\alpha + 51\right)\theta^5 + \left(72\alpha^2 + 228\alpha + 42\right)\theta^4 \\ + \left(192\alpha^2 + 210\alpha + 16\right)\theta^3 + \left(180\alpha^2 + 86\alpha + 2\right)\theta^2 + \left(72\alpha^2 + 12\alpha\right)\theta + 12\alpha^2 \end{array}\right\}^2},$$

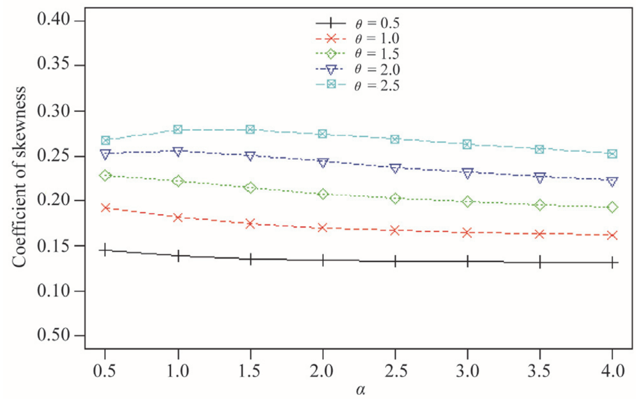
$$\gamma = \frac{\sigma^2}{\mu} =$$

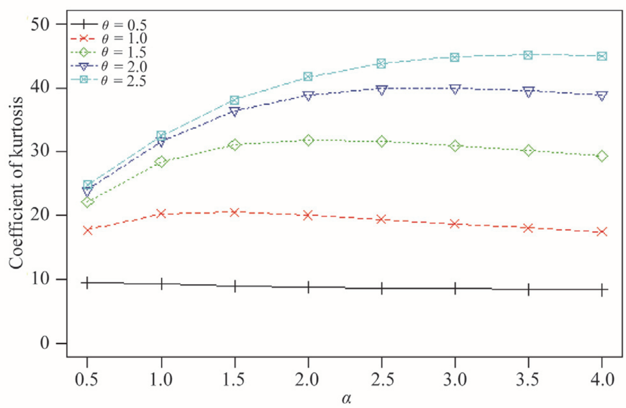
$$= \frac{\left(\theta + 1\right)\left\{\begin{array}{l} \theta^8 + 9\theta^7 + \left(18\alpha + 31\right)\theta^6 + \left(110\alpha + 51\right)\theta^5 + \left(72\alpha^2 + 228\alpha + 42\right)\theta^4 \\ + \left(192\alpha^2 + 210\alpha + 16\right)\theta^3 + \left(180\alpha^2 + 86\alpha + 2\right)\theta^2 + \left(72\alpha^2 + 12\alpha\right)\theta + 12\alpha^2 \end{array}\right\}}{\theta\left(\theta + 1\right)\left\{\theta^4 + 4\theta^3 + \left(6\alpha + 4\right)\theta^2 + \left(6\alpha + 1\right)\theta + 2\alpha\right\}\left\{\theta^4 + 4\theta^3 + \left(6\alpha + 5\right)\theta^2 + \left(12\alpha + 2\right)\theta + 6\alpha\right\}}.$$

The behaviours of the CV, skewness, kurtosis, and the index of dispersion of AGZTPSD for varying values of parameters are shown graphically in Figure 2.

Figure 2. Behaviours of CV, skewness, kurtosis and index of dispersion of AGZTPSD for values of θ and α

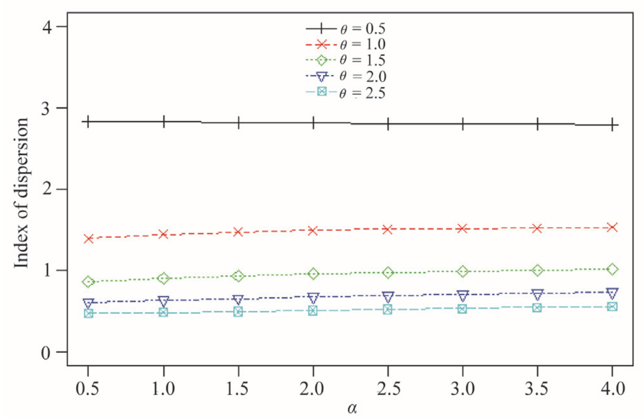






(Continued on the next page)

(Continued)



3. Maximum likelihood estimation

Let $x_1, x_2,...,x_n$ be the observed values of a random sample $(X_1, X_2,...,X_n)$ from AGZTPSD(θ,α) /9/, and let f_x be the observed frequency in the sample corresponding to $X = x(x = 1, 2, 3,...,k)$ such that $\sum_{x=1}^{k} f_x = n$, where k is the largest observed value having a non-zero frequency. The likelihood function L of AGZTPSD is described by

$$L = \left(\frac{\theta^3}{\theta^4 + 4\theta^3 + (6\alpha + 4)\theta^2 + (6\alpha + 1)\theta + 2\alpha} \right)^n \times$$

$$\times \frac{1}{(\theta + 1)^{\sum_{x=1}^{k} x f_x}} \prod_{x=1}^{k} \left\{ \alpha x^2 + (\theta + 3\alpha + 1)x + (\theta^2 + 3\theta + 2\alpha + 2) \right\}^{f_x}.$$

The log likelihood function is given by

$$\log L = 3n \log\theta - n \log\left\{ \theta^4 + 4\theta^3 + (6\alpha + 4)\theta^2 + (6\alpha + 1)\theta + 2\alpha \right\} - \sum_{x=1}^{k} x f_x \log(\theta + 1)$$

$$+ \sum_{x=1}^{k} f_x \log\left\{ \alpha x^2 + (\theta + 3\alpha + 1)x + (\theta^2 + 3\theta + 2\alpha + 2) \right\}.$$

The MLEs (maximum likelihood estimate) $\left(\hat{\theta}, \hat{\alpha}\right)$ of parameters $\left(\theta, \alpha\right)$ of AGZTPSD /9/ can be obtained from the solutions of the following log likelihood equations:

$$\frac{\partial \log L}{\partial \theta} = \frac{3n}{\theta} - \frac{n\left\{4\theta^3 + 12\theta^2 + 2\left(6\alpha + 4\right)\theta + \left(6\alpha + 1\right)\right\}}{\theta^4 + 4\theta^3 + \left(6\alpha + 4\right)\theta^2 + \left(6\alpha + 1\right)\theta + 2\alpha} - \frac{n\bar{x}}{\theta + 1}$$
$$+ \sum_{x=1}^{k} \frac{\left(x + 2\theta + 3\right) f_x}{\alpha x^2 + \left(\theta + 3\alpha + 1\right)x + \left(\theta^2 + 3\theta + 2\alpha + 2\right)} = 0,$$

$$\frac{\partial \log L}{\partial \alpha} = \frac{-n\left(6\theta^2 + 6\theta + 2\right)}{\theta^4 + 4\theta^3 + \left(6\alpha + 4\right)\theta^2 + \left(6\alpha + 1\right)\theta + 2\alpha}$$
$$+ \sum_{x=1}^{k} \frac{\left(x^2 + 3x + 2\right) f_x}{\alpha x^2 + \left(\theta + 3\alpha + 1\right)x + \left(\theta^2 + 3\theta + 2\alpha + 2\right)} = 0.$$

The former two log likelihood equations cannot be expressed in their closed forms and seem difficult to solve directly. The MLE $\left(\hat{\theta},\hat{\alpha}\right)$ of parameters $\left(\theta,\alpha\right)$ of AGZTPSD /9/ can be computed directly by solving the log likelihood equations using the Newton-Raphson iteration available in R-software, until sufficiently close values of $\hat{\theta}$ and $\hat{\alpha}$ are obtained. The initial values for the parameters θ and α are generally assumed to be $\theta = 0.5$ and $\alpha = 1$.

4. Goodness of fit

AGZTPSD was fitted to a number of real count datasets to test its goodness of fit over the ZTPD, ZTPLD, and the ZTPSD. It was observed that in most cases it produces a better fit. MLEs of the parameter were used to fit the ZTPD, ZTPLD, ZTPSD, and AGZTPSD. In this study, four examples of real data sets are presented. The first dataset comprises the data regarding the number of households having at least one migrant as reported by *Singh* and *Yadav* [1971]. The second dataset is the data regarding the number of counts of sites with particles from Immunogold data, reported by *Mathews* and *Appleton* [1993]. The third dataset is the data regarding the number of flower heads with fly eggs, reported by *Finney* and *Varley* [1955], and the fourth dataset is the data regarding the number of snowshoe hares captured over 7 days, reported by *Keith* and *Meslow* [1968]. It is obvious from the values of Chi-square $\left(\chi^2\right)$ and *p*-values that the ZTPSD produces a much closer fit than the ZTPD and ZTPLD. Therefore, the ZTPSD can be considered a more suitable tool for modelling count data (excluding zero-count) than the ZTPD and ZTPLD.

Table 1

Number of households having at least one migrant, by number of migrants

Number of migrants	Observed frequency	Expected frequency			
		ZTPD	ZTPLD	ZTPSD	AGZTPSD
1	375	354.0	379.0	378.3	377.0
2	143	167.7	137.2	137.8	139.2
3	49	52.9	48.4	48.7	49.1
4	17	12.5 }	16.7	16.8	16.6
5	2	2.4 }	5.7 }	5.6 }	5.4
6	2	0.4 }	1.9 }	1.8 }	1.7
7	1	0.1 }	0.6 }	0.6 }	0.5
8	1	0.0 }	0.5 }	0.4 }	0.5
Total	*590*	*590.0*	*590.0*	*590.0*	*590.0*
MLE		$\hat{\theta}$= 0.947486	$\hat{\theta}$= 2.284782	$\hat{\theta}$= 2.722929	$\hat{\theta}$= 2.9556 $\hat{\alpha}$= 2.1327
χ^2		8.92	1.14	0.91	0.66
df		2	3	3	2
p-value		0.011	0.769	0.822	0.718
$-2\log L$		1,203.3	1,186.1	1,186.2	1,186.1
AIC		1,205.3	1,188.1	1,188.2	1,190.1

Note. Here and in the following tables, ZTPD: zero-truncated Poisson distribution; ZTPLD: zero-truncated Poisson-Lindley distribution; ZTPSD: zero-truncated Poisson-Sujatha distribution; AGZTPSD: a generalization of the zero-truncated Poisson-Sujatha distribution; MLE: maximum likelihood estimate; *L*: likelihood function; AIC: Akaike information criterion.

Source: Own calculation based on *Singh–Yadav* [1971].

Table 2

Number of counts of sites with particles from Immunogold data

Number of sites with particles	Observed frequency	Expected frequency			
		ZTPD	*ZTPLD*	*ZTPSD*	*AGZTPSD*
1	122	115.9	124.8	124.4	122.3
2	50	57.4	46.8	47.0	49.4
3	18	18.9	17.1	17.2	17.7
4	4	4.7 }	6.1 }	6.1 }	5.8
5	4	1.1 }	3.2 }	3.3 }	2.8
Total	*198*	*198.0*	*198.0*	*198.0*	*198.0*
MLE		$\hat{\theta}$= 0.990586	$\hat{\theta}$= 2.18307	$\hat{\theta}$= 2.614691	$\hat{\theta}$= 3.3839 $\hat{\alpha}$= 12.5364
χ^2		2.14	0.51	0.46	0.05
df		2	2	2	1
p-value		0.343	0.775	0.794	0.816
$-2\log L$		413.8	409.2	409.1	408.6
AIC		415.8	411.2	411.1	412.6

Source: Own calculation based on *Mathews–Appleton* [1993].

Table 3

Number of flower heads with fly eggs

Number of fly eggs	Number of flowers	Expected frequency			
		ZTPD	ZTPLD	ZTPSD	AGZTPSD
1	22	15.3	26.8	25.1	22.9
2	18	21.8	19.8	19.9	20.7
3	18	20.8	14.0	14.6	15.8
4	11	14.9	9.5	10.2	10.9
5	9	8.5	6.3	6.7	7.0
6	6	4.0 }	4.2 }	4.3	4.3
7	3	1.7 }	2.7 }	2.7	2.6
8	0	0.6 }	1.7 }	1.6	1.5
9	1	0.4 }	3.0 }	2.9	2.3
Total	*88*	*88.0*	*88.0*	*88.0*	*88.0*
MLE		$\hat{\theta}$= 2.8604	$\hat{\theta}$= 0.7186	$\hat{\theta}$= 0.9813	$\hat{\theta}$= 1.1482 $\hat{\alpha}$= 16.1963
χ^2		6.65	3.78	2.40	1.31
df		4	4	4	3
p-value		0.155	0.436	0.662	0.726
$-2\log L$		333.1	334.7	332.3	330.4
AIC		335.1	336.7	334.3	334.4

Source: Own calculation based on *Finney–Varley* [1955].

Table 4

Number of snowshoe hares captured over 7 days

Number of times hares were caught	Observed frequency	Expected frequency			
		ZTPD	ZTPLD	ZTPSD	AGZTPSD
1	184	174.6	182.6	182.5	182.6
2	55	66.0	55.3	55.3	55.2
3	14	16.6	16.4	16.4	16.3
4	4	3.2	4.8	4.7	4.8
5	4	0.6	1.9	2.1	2.1
Total	*261*	*261.0*	*261.0*	*261.0*	*261.0*
MLE		$\hat{\theta}=0.7563$	$\hat{\theta}=2.8639$	$\hat{\theta}=3.320$	$\hat{\theta}=3.032$ $\hat{\alpha}=0.277$
χ^2		2.46	0.61	0.57	0.51
df		1	2	2	1
p-value		0.116	0.735	0.752	0.475
$-2\log L$		461.9	453.6	453.6	453.6
AIC		463.9	455.6	455.6	457.6

Source: Own calculation based on *Keith–Meslow* [1968].

5. Conclusions

In this study AGZTPSD, a generalization of the ZTPSD, which includes both the ZTPLD and the ZTPSD as particular cases, was introduced. The moments of the distribution were derived. The natures of the CV, skewness, kurtosis, and the index of dispersion were studied graphically. The maximum likelihood estimation was discussed for the estimation of the parameters of the distribution. The goodness of fit of the distribution was explained with four datasets, and it shows quite a satisfactory fit over other zero-truncated distributions including the ZTPD, ZTPLD, and the ZTPSD.

References

FINNEY, D. J. – VARLEY, G. C. [1955]: An example of the truncated Poisson distribution. *Biometrics*. Vol. 11. No. 3. pp. 387–394. https://doi.org/10.2307/3001776

GHITANY, M. E. – ATIEH, B. – NADARAJAH, S. [2008a]: Lindley distribution and its applications. *Mathematics and Computers in Simulation*. Vol. 78. No. 4. pp. 493–506. https://doi.org/10.1016/j.matcom.2007.06.007

GHITANY, M. E. – AL-MUTAIRI, D. K. – NADARAJAH, S. [2008b]: Zero-truncated Poisson-Lindley distribution and its applications. *Mathematics and Computers in Simulation*. Vol. 79. No. 3. pp. 279–287. https://doi.org/10.1016/j.matcom.2007.11.021

KEITH, L. B. – MESLOW, E. C. [1968]: Trap response by snowshoe hares. *Journal of Wildlife Management*. Vol. 32. No. 4. pp. 795–801. https://doi.org/10.2307/3799555

LINDLEY, D. V. [1958]: Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society, Series B*. Vol. 20. No. 1. pp. 102–107. https://doi.org/10.1111/j.2517-6161.1958.tb00278.x

MATHEWS, J. N. S. – APPLETON, D. R. [1993]: An application of the truncated Poisson distribution to Immunogold assay. *Biometrics*. Vol. 49. No. 2. pp. 617–621. https://doi.org/10.2307/2532574

SANKARAN, M. [1970]: 275. Note: The discrete Poisson-Lindley distribution. *Biometrics*. Vol. 26. No. 1. pp. 145–149. https://doi.org/10.2307/2529053

SHANKER, R. – HAGOS, F. [2015a]: Zero-truncated Poisson-Sujatha distribution with applications. *Journal of Ethiopian Statistical Association*. Vol. 24. pp. 55–63.

SHANKER, R. – HAGOS, F. [2015b]: On Poisson-Lindley distribution and its applications to biological sciences. *Biometrics & Biostatistics International Journal*. Vol. 2. No. 4. pp. 1–5. https://doi.org/10.15406/bbij.2015.02.00036

SHANKER, R. [2016a]: Sujatha distribution and its applications. *Statistics in Transition* (New Series). Vol. 17. No. 3. pp. 391–410. https://doi.org/10.21307/stattrans-2016-029

SHANKER, R. [2016b]: The discrete Poisson-Sujatha distribution. *International Journal of Probability and Statistics*. Vol. 5. No. 1. pp. 1–9. https://doi.org/10.5923/j.ijps.20160501.01

SHANKER, R. – SHUKLA, K. K. – HAGOS, F. [2017]: A generalization of Sujatha distribution and its applications to real lifetime data. *Journal of Institute of Science and Technology*. Vol. 22. No. 1. pp. 77–94. https://doi.org/10.3126/jist.v22i1.17742

SHANKER, R. – SHUKLA, K. K. [2019]: A generalization of Poisson-Sujatha distribution and its applications to ecology. *International Journal of Biomathematics*. Vol. 12. No. 2. pp. 1–11. https://doi.org/10.1142/S179352451950013X

SINGH, S. N. – YADAV, R. C. [1971]: Trends in rural out-migration at household level. *Rural Demography*. Vol. 8. Vol. 1. pp. 53–61.