ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2017-54834

The Discrete Poisson-Aradhana Distribution

Kesikli Poisson-Aradhana Dağılımı

Rama SHANKER[a]

[a]Department of Statistics,
Eritrea Institute of Technology,
Asmara, Eritrea

Geliş Tarihi/*Received:* 25.01.2017
Kabul Tarihi/*Accepted:* 22.02.2017

Yazışma Adresi/*Correspondence:*
Rama SHANKER
Eritrea Institute of Technology,
Department of Statistics, Asmara,
ERITREA/ERİTRE
shankerrama2009@gmail.com

ABSTRACT In this paper, a discrete Poisson-Aradhana distribution has been obtained by compounding the Poisson distribution with the Aradhana distribution. The r th factorial moment about origin of the proposed distribution has been derived and hence the first four moments about origin and the central moments have been obtained. Its statistical properties including coefficient of variation, skewness, kurtosis, moment generating function, increasing hazard rate, unimodality and over-dispersion have been discussed. Estimation of the parameter has been discussed using method of maximum likelihood and that of method of moments. The goodness of fit of the proposed distribution has been discussed with two examples of count data from biological science and the fit has been compared with one parameter Poisson and Poisson-Lindley distributions.

Keywords: Aradhana distribution; Poisson-Lindley distribution; statistical properties; estimation of parameter; goodness of fit

ÖZET Bu makalede, kesikli Poisson-Aradhana dağılımı Poisson dağılımının Aradhana dağılımı ile birleştirilmesinden elde edilmiştir. Önerilen dağılımın orijine bağlı r.inci faktör momenti türetilmiştir dolayısıyla orijine bağlı ilk dört moment ve merkezi momentler elde edilmiştir. Bu momentlerin değişim katsayısını, çarpıklığını, basıklığını, moment üreten fonksiyonunu, artan hazard hızını, tek şekillilik ve fazla yayılımı içeren özellikleri tartışılmıştır. Parametre tahminleri maksimum olabilirlik yöntemini ve moment yöntemini kullanarak tartışılmıştır. Önerilen dağılımın uyum iyiliği biyoloji biliminden elde edilen iki örnekle tartışılmıştır ve uygunluğu tek parametreli Poisson ve Poisson-Lindley dağılımları ile karşılaştırılmıştır.

Anahtar Kelimeler: Aradhana dağılımı; Poisson-Lindley dağılımı; istatistiksel özellikler; parametre tahmini; uyum iyiliği

Shanker (2016) has introduced lifetime distribution named 'Aradhana distribution' having probability density function (p.d.f) and cumulative distribution function (c.d.f).[1]

$$f(x;\theta)=\frac{\theta^3}{\theta^2+2\theta+2}(1+x)^2 e^{-\theta x} \quad ; x>0,\ \theta>0 \tag{1.1}$$

$$F(x;\theta)=1-\left[1+\frac{\theta x(\theta x+2\theta+2)}{\theta^2+2\theta+2}\right]e^{-\theta x} \quad ; x>0, \theta>0 \tag{1.2}$$

for modeling lifetime data from engineering and biomedical science. It has been shown by Shanker (2016) that Aradhana distribution is a three-component mixture of an exponential distribution having scale parameter

θ, a gamma distribution having shape parameter 2 and scale parameter θ, and a gamma distribution with shape parameter 3 and scale parameter θ with their mixing proportions $\frac{\theta^2}{\theta^2+2\theta+2}$, $\frac{2\theta}{\theta^2+2\theta+2}$ and $\frac{2}{\theta^2+2\theta+2}$, respectively.[1] Its statistical properties including its shape, moments, skewness , kurtosis, hazard rate function, mean residual life function, stochastic ordering, mean deviations,[1] Bonferroni and Lorenz curves, and stress-strength reliability have been discussed by Shanker (2016). Further, Shanker (2016) has shown that both the method of maximum likelihood estimation and the method of moment give the same estimate of the parameter of Aradhana distribution.[1] The applicability and goodness of fit of Aradhana distribution has been established through some lifetime data – sets from medical science and engineering and observed by Shanker (2016) that in most data-sets it gives better fit than Akash and Shanker distributions introduced by Shanker (2015 a, 2015 b), Lindley (1958) distribution and exponential distribution.[1-4] Recently, Shanker et al (2016) have done extensive study on the applications of Akash distribution introduced by Shanker (2015 a), Lindley (1958) distribution and exponential distribution for modeling lifetime data from biomedical science and engineering and observed that Akash distribution provides better fit than Lindley and exponential distributions in many data sets.[2,3,5] Further, Shanker and Hagos (2016 a, 2016 b) have detailed and critical comparative study on applications of Akash, Shanker, Lindley, and exponential distributions and that of Aradhana, Sujatha, Lindley and exponential distributions.[6,7]

In the present paper, a Poisson mixture of Aradhana distribution introduced by Shanker (2016) named, "Poisson-Aradhana distribution (PAD)" has been suggested.[1] Its mathematical and statistical properties including its shape, moments, coefficient of variation, skewness, kurtosis etc have been discussed. The estimation of its parameter has been discussed using method of maximum likelihood estimation and the method of moments. The goodness of fit of PAD along with Poisson distribution and Poisson-Lindley distribution (PLD), a Poisson mixture of Lindley (1958) distribution and introduced by Sankaran (1970), has been given for two count data sets.[2,8]

POISSON-ARADHANA DISTRIBUTION

Suppose the parameter λ of the Poisson distribution follows Aradhana distribution (1.1). Then the Poisson mixture of Aradhana distribution (1.1) can be obtained as

$$P(X=x)=\int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}\cdot\frac{\theta^3}{\theta^2+2\theta+2}(1+\lambda)^2 e^{-\theta\lambda}d\lambda \tag{2.1}$$

$$=\frac{\theta^3}{(\theta^2+2\theta+2)x!}\int_0^\infty e^{-(\theta+1)\lambda}\left[\lambda^x+2\lambda^{x+1}+\lambda^{x+2}\right]d\lambda$$

$$=\frac{\theta^3}{\theta^2+2\theta+2}\cdot\frac{x^2+(2\theta+5)x+(\theta^2+4\theta+5)}{(\theta+1)^{x+3}}\;;x=0,1,2,...,\theta>0 \tag{2.2}$$

We name this distribution "Poisson-Aradhana distribution (PAD)".

Recall that the Poisson-Lindley distribution (PLD) having probability mass function (p.m.f)

$$P(X=x)=\frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}}\;;x=0,1,2,...,\theta>0 \tag{2.3}$$

has been introduced by Sankaran (1970).[8] The PLD has been obtained by compounding Poisson distribution with Lindley distribution, introduced by Lindley (1958) having p.d.f.[2]

$$f(x,\theta)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}\ ;x>0,\theta>0 \tag{2.4}$$

The main motivation for considering Poisson-Aradhana distribution (PAD) is two folds: firstly Aradhana distribution gives better fit than Lindley distribution and secondly Poisson-Lindley distribution (PLD), a Poisson mixture of Lindley (1958) distribution gives better fit than Poisson distribution, the Poisson-Aradhana distribution is expected to gives better fit than both Poisson and Poisson–Lindley distribution.[2]

The graphs of the p.m.f of PAD and PLD for different values of the parameter are shown in the Figure 1.

MOMENTS AND ASSOCIATED MEASURES

The r th factorial moment about origin of PAD (2.2) can be obtained as

$$\mu_{(r)}{}' = E\left[E\left(X^{(r)}\,|\,\lambda\right)\right],\quad \text{where}\quad X^{(r)}=X(X-1)(X-2)...(X-r+1).$$

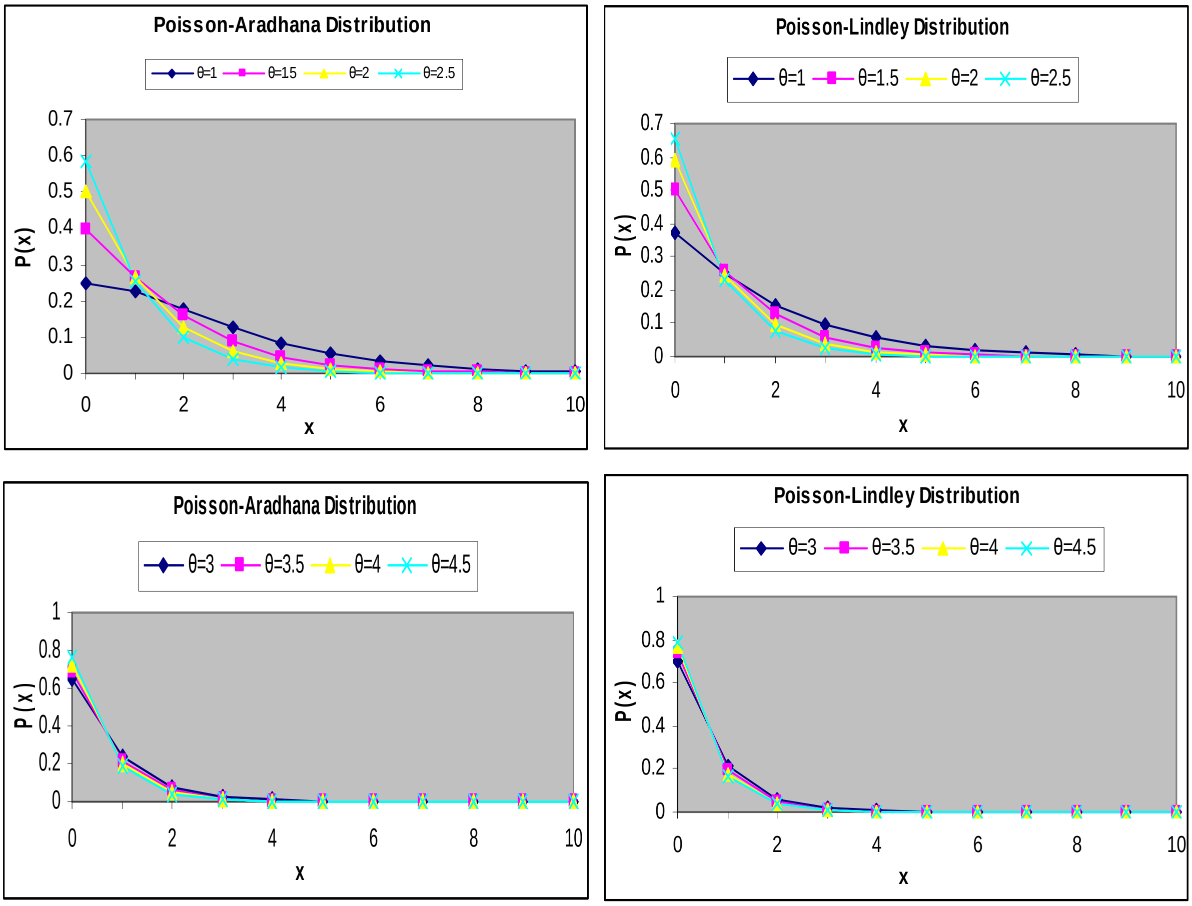


FIGURE 1: Graphs of probability mass function of PAD and PLD for different values of the parameter θ.

Using (2.1), the r th factorial moment about origin of PAD (2.2) can be obtained as

$$\mu_{(r)}' = E\left[E\left(X^{(r)} \mid \lambda\right)\right] = \frac{\theta^3}{\theta^2+2\theta+2}\int_0^\infty \left[\sum_{x=0}^{\infty} x^{(r)} \frac{e^{-\lambda}\lambda^x}{x!}\right](1+\lambda)^2 e^{-\theta\lambda} d\lambda$$

$$= \frac{\theta^3}{\theta^2+2\theta+2}\int_0^\infty \lambda^r \left[\sum_{x=r}^{\infty} \frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!}\right](1+\lambda)^2 e^{-\theta\lambda} d\lambda$$

Taking $x+r$ in place of x within the bracket, we get

$$\mu_{(r)}' = \frac{\theta^3}{\theta^2+2\theta+2}\int_0^\infty \lambda^r \left[\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{x!}\right](1+\lambda)^2 e^{-\theta\lambda} d\lambda$$

The expression within the bracket is clearly unity and hence we have

$$\mu_{(r)}' = \frac{\theta^3}{\theta^2+2\theta+2}\int_0^\infty \lambda^r (1+\lambda)^2 e^{-\theta\lambda} d\lambda$$

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

$$= \frac{\theta^3}{\theta^2+2\theta+2}\left[\int_0^\infty e^{-\theta\lambda}\lambda^{r+1-1} d\lambda + 2\int_0^\infty e^{-\theta\lambda}\lambda^{r+2-1} d\lambda + \int_0^\infty e^{-\theta\lambda}\lambda^{r+3-1} d\lambda\right]$$

$$= \frac{\theta^3}{\theta^2+2\theta+2}\left[\frac{\Gamma(r+1)}{\theta^{r+1}} + \frac{2\Gamma(r+2)}{\theta^{r+2}} + \frac{\Gamma(r+3)}{\theta^{r+3}}\right]$$

$$= \frac{\theta^3\ r!}{\theta^2+2\theta+2}\frac{\theta^2+2(r+1)\theta+(r+1)(r+2)}{\theta^{r+3}}$$

Thus a general expression for the r th factorial moment of PAD (2.2) are obtained as

$$\mu_{(r)}' = \frac{r!\left[\theta^2+2(r+1)\theta+(r+1)(r+2)\right]}{\theta^r\left(\theta^2+2\theta+2\right)}\ ; r=1,2,3,.... \tag{3.1}$$

Substituting $r = 1, 2, 3, \text{and } 4$ in (3.1), the first four factorial moments about origin can be obtained and using the relationship between factorial moments about origin and moments about origin, the first four moment about origin of the PAD (2.2) are obtained as

$$\mu_1' = \frac{\theta^2+4\theta+6}{\theta\left(\theta^2+2\theta+2\right)}$$

$$\mu_2' = \frac{\theta^3+6\theta^2+18\theta+24}{\theta^2\left(\theta^2+2\theta+2\right)}$$

$$\mu_3' = \frac{\theta^4+10\theta^3+48\theta^2+96\theta+120}{\theta^3\left(\theta^2+2\theta+2\right)}$$

$$\mu_4' = \frac{\theta^5+18\theta^4+126\theta^3+480\theta^2+960\theta+720}{\theta^4\left(\theta^2+2\theta+2\right)}$$

Again using the relationship between moments about mean and the moments about origin, the moments about mean of the PAD (2.2) are obtained as

$$\mu_2=\sigma^2=\frac{\theta^5+7\theta^4+24\theta^3+44\theta^2+36\theta+12}{\theta^2\left(\theta^2+2\theta+2\right)^2}$$

$$\mu_3=\frac{\theta^8+11\theta^7+62\theta^6+190\theta^5+360\theta^4+396\theta^3+264\theta^2+120\theta+48}{\theta^3\left(\theta^2+2\theta+2\right)^3}$$

$$\mu_4=\frac{\left(\begin{array}{l}\theta^{11}+20\theta^{10}+186\theta^9+1053\theta^8+4040\theta^7+10856\theta^6+20896\theta^5\\+28856\theta^4+28176\theta^3+18144\theta^2+6624\theta+720\end{array}\right)}{\theta^4\left(\theta^2+2\theta+2\right)^4}$$

The coefficient of variation $(C.V)$, coefficient of Skewness $\left(\sqrt{\beta_1}\right)$, coefficient of Kurtosis (β_2), and index of dispersion (γ) of the PAD (2.2) are thus obtained as

$$C.V=\frac{\sigma}{\mu_1'}=\frac{\sqrt{\theta^5+7\theta^4+24\theta^3+44\theta^2+36\theta+12}}{\theta^2+4\theta+6}$$

$$\sqrt{\beta_1}=\frac{\mu_3}{\mu_2^{3/2}}=\frac{\theta^8+11\theta^7+62\theta^6+190\theta^5+360\theta^4+396\theta^3+264\theta^2+120\theta+48}{\left(\theta^5+7\theta^4+24\theta^3+44\theta^2+36\theta+12\right)^{3/2}}$$

$$\beta_2=\frac{\mu_4}{\mu_2^2}=\frac{\left(\begin{array}{l}\theta^{11}+20\theta^{10}+186\theta^9+1053\theta^8+4040\theta^7+10856\theta^6+20896\theta^5\\+28856\theta^4+28176\theta^3+18144\theta^2+6624\theta+720\end{array}\right)}{\left(\theta^5+7\theta^4+24\theta^3+44\theta^2+36\theta+12\right)^2}$$

$$\gamma=\frac{\sigma^2}{\mu_1'}=\frac{\theta^5+7\theta^4+24\theta^3+44\theta^2+36\theta+12}{\theta\left(\theta^2+2\theta+2\right)\left(\theta^2+4\theta+6\right)}$$

The expressions for μ_1', μ_2, C.V, $\sqrt{\beta_1}$, β_2 and γ of PLD (2.3) obtained by Sankaran (1970) and Ghitany and Al-Mutairi (2009) are given by[8,9]

$$\mu_1'=\frac{\theta+2}{\theta(\theta+1)},\qquad \mu_2=\frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2}$$

$$C.V=\frac{\sigma}{\mu_1'}=\frac{\sqrt{\theta^3+4\theta^2+6\theta+2}}{\theta+2}$$

$$\sqrt{\beta_1}=\frac{\mu_3}{\mu_2^{3/2}}=\frac{\theta^5+7\theta^4+22\theta^3+32\theta^2+18\theta+4}{\left(\theta^3+4\theta^2+6\theta+2\right)^{3/2}}$$

$$\beta_2=\frac{\mu_4}{\mu_2^2}=\frac{\left(\theta^7+15\theta^6+87\theta^5+258\theta^4+406\theta^3+338\theta^2+144\theta+24\right)}{\left(\theta^3+4\theta^2+6\theta+2\right)^2}$$

$$\gamma=\frac{\sigma^2}{\mu_1'}=\frac{\theta^3+4\theta^2+6\theta+2}{\theta(\theta+1)(\theta+2)}$$

To study the behavior of μ_1', μ_2, C.V, $\sqrt{\beta_1}$, β_2 and γ of PAD and PLD, values of these characteristics for different values of parameter θ have been computed and presented in Table 1.

The graphs of coefficient of variation (C.V), coefficient of skewness $\left(\sqrt{\beta_1}\right)$, coefficient of kurtosis (β_2), and Index of dispersion (γ) of PAD for different values of the parameter θ have been shown in the Figure 2

It is obvious from the graphs of C.V. of PAD and PLD that C.V. of both PAD and PLD are increasing function of the parameter θ. Further, the graph of C.V. of PLD is slightly higher than the graph of C.V. of PAD.

The graphs of coefficient of skewness for both PAD and PLD are increasing function of the parameter θ and as the value of the parameter θ increases, the distance between the graphs getting narrower.

It is clear from the graphs of the coefficient of kurtosis that graphs of coefficient of kurtosis for both PAD and PLD are increasing function of the parameter θ and the distance between the graphs getting wider for increasing values of the parameter θ

The graphs of the index of dispersion of PAD and PLD are decreasing for increasing value of the parameter θ and both are becoming identical for the value of $\theta > 6$.

MATHEMATICAL AND STATISTICAL PROPERTIES

INCREASING HAZARD RATE AND UNIMODALITY

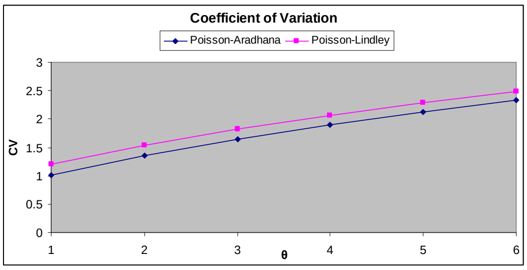
The PAD (2.2) has an increasing hazard rate (IHR) and unimodal. Since

$$\frac{P(x+1;\theta)}{P(x;\theta)} = \frac{1}{\theta+1}\left[1+\frac{2(x+\theta+2)}{x^2+(2\theta+5)x+(\theta^2+4\theta+5)}\right]$$

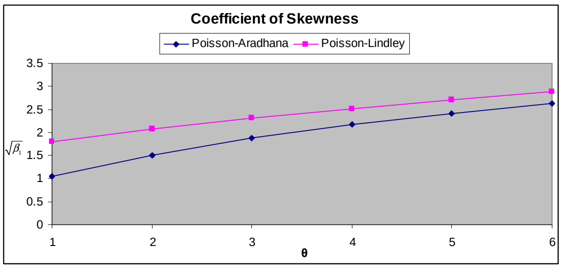
TABLE 1: Values of $\mu_1', \mu_2, \mathrm{C.V}, \sqrt{\beta_1}, \beta_2$ and γ of PAD for different values of θ.

	Values of θ for Poisson-Aradhana Distribution					
	1	2	3	4	5	6
μ_1'	2.2	0.9	0.529412	0.365385	0.275676	0.22
μ_2	4.96	1.49	0.758939	0.481879	0.344543	0.264933
CV	1.012321	1.356284	1.645545	1.899847	2.129235	2.339622
$\sqrt{\beta_1}$	1.05156	1.510906	1.877012	2.170024	2.416428	2.632192
β_2	7.776535	8.562542	9.504795	10.47295	11.4443	12.41664
γ	2.254545	1.655556	1.433551	1.318826	1.249815	1.204242

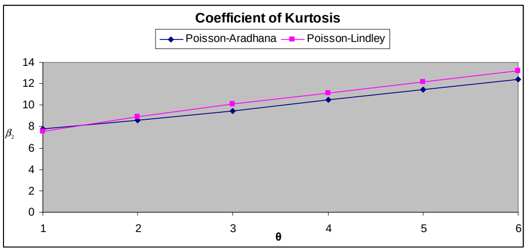
	Values of θ for Poisson-Lindley Distribution					
	1	2	3	4	5	6
μ_1'	1.5	0.666667	0.416667	0.3	0.233333	0.190476
μ_2	3.25	1.055556	0.576389	0.385	0.285556	0.225624
CV	1.20185	1.541104	1.822087	2.068279	2.290174	2.493742
$\sqrt{\beta_1}$	1.792108	2.083265	2.314307	2.517935	2.704839	2.87957
β_2	7.532544	8.941828	10.10611	11.17187	12.19654	13.203
γ	2.166667	1.583333	1.383333	1.283333	1.22381	1.184524



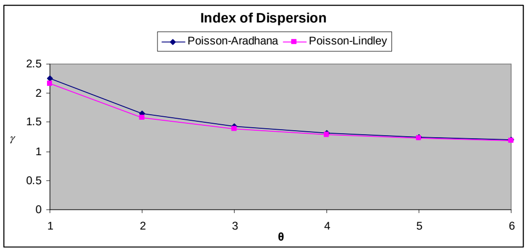
It is obvious from the graphs of C.V. of PAD and PLD that C.V. of both PAD and PLD are increasing function of the parameter θ. Further, the graph of C.V. of PLD is slightly higher than the graph of C.V. of PAD.



The graphs of coefficient of skewness for both PAD and PLD are increasing function of the parameter θ and as the value of the parameter θ increases, the distance between the graphs getting narrower.



It is clear from the graphs of the coefficient of kurtosis that graphs of coefficient of kurtosis for both PAD and PLD are increasing function of the parameter θ and the distance between the graphs getting wider for increasing values of the parameter θ.



The graphs of the index of dispersion of PAD and PLD are decreasing for increasing value of the parameter θ and both are becoming identical for the value of $\theta > 6$.

FIGURE 2: Graphs of coefficient of variation (C.V), coefficient of skewness $\left(\sqrt{\beta_1}\right)$, coefficient of kurtosis $\left(\beta_2\right)$, and Index of dispersion $\left(\gamma\right)$ of PAD and PLD for different values of the parameter θ

is a decreasing function in x, $P(x;\theta)$ is log-concave. Therefore, the PAD has an increasing hazard rate and unimodal. A detailed discussion about interrelationship between log-concavity, unimodality and increasing hazard rate (IHR) of discrete distributions can be seen in Grandell (1997).[10]

GENERATING FUNCTIONS

The probability generating function of PAD (2.2) can be obtained as

$$P_X(t) = \frac{\theta^3}{\left(\theta^2+2\theta+2\right)\left(\theta+1\right)^3}\left[\sum_{x=0}^{\infty} x^2\left(\frac{t}{\theta+1}\right)^x + (2\theta+5)\sum_{x=0}^{\infty} x\left(\frac{t}{\theta+1}\right)^x + \left(\theta^2+4\theta+5\right)\sum_{x=0}^{\infty}\left(\frac{t}{\theta+1}\right)^x\right]$$

$$= \frac{\theta^3}{\left(\theta^2+2\theta+2\right)\left(\theta+1\right)^2}\left[\frac{2(\theta+1)t}{(\theta+1-t)^3} + \frac{2(\theta+3)t}{(\theta+1-t)^2} + \frac{\theta^2+4\theta+5}{(\theta+1-t)}\right]$$

The moment generating function of the PAD (2.2) is thus obtained as

$$M_X(t) = \frac{\theta^3}{\left(\theta^2+2\theta+2\right)\left(\theta+1\right)^2}\left[\frac{2(\theta+1)e^t}{\left(\theta+1-e^t\right)^3} + \frac{2(\theta+3)e^t}{\left(\theta+1-e^t\right)^2} + \frac{\theta^2+4\theta+5}{\left(\theta+1-e^t\right)}\right]$$

OVER-DISPERSION

The PAD (2.2) is always over-dispersed $\left(\sigma^2 > \mu\right)$. We have

$$\sigma^2 = \frac{\theta^5+7\theta^4+24\theta^3+44\theta^2+36\theta+12}{\theta^2\left(\theta^2+2\theta+2\right)^2}$$

$$= \frac{\theta^2+4\theta+6}{\theta\left(\theta^2+2\theta+2\right)}\left[\frac{\theta^5+\theta^4+8\theta^3+16\theta^2+12\theta+12}{\theta\left(\theta^2+2\theta+2\right)\left(\theta^2+4\theta+6\right)}\right]$$

$$= \frac{\theta^2+4\theta+6}{\theta\left(\theta^2+2\theta+2\right)}\left[1+\frac{\theta^4+8\theta^3+24\theta^2+24\theta+12}{\theta\left(\theta^2+2\theta+2\right)\left(\theta^2+4\theta+6\right)}\right]$$

$$= \mu\left[1+\frac{\theta^4+8\theta^3+24\theta^2+24\theta+12}{\theta\left(\theta^2+2\theta+2\right)\left(\theta^2+4\theta+6\right)}\right] > \mu \cdot$$

This shows that PAD (2.2) is always over dispersed. This is also obvious from the values of μ_2 and μ_1' of PAD in Table 1.

ESTIMATION OF THE PARAMETER

Maximum Likelihood Estimate (MLE)

Let $\left(x_1, x_2, ..., x_n\right)$ be a random sample of size n from the PAD (2.2) 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 non-zero frequency. The likelihood function L of the PAD (2.2) is given by

$$L = \left(\frac{\theta^3}{\theta^2+2\theta+2}\right)^n \frac{1}{\left(\theta+1\right)^{\sum_{x=1}^{k}(x+3)f_x}} \prod_{x=1}^{k}\left[x^2+\left(2\theta+5\right)x+\left(\theta^2+4\theta+5\right)\right]^{f_x}$$

The log likelihood function is thus obtained as

$$\log L = n\log\left(\frac{\theta^3}{\theta^2+2\theta+2}\right) - \sum_{x=1}^{k}\left(x+3\right)f_x\log\left(\theta+1\right) + \sum_{x=1}^{k} f_x \log\left[x^2+\left(2\theta+5\right)x+\left(\theta^2+4\theta+5\right)\right]$$

The first derivative of the log likelihood function is given by

$$\frac{d\log L}{d\theta} = \frac{n\left(\theta^2+4\theta+6\right)}{\theta\left(\theta^2+2\theta+2\right)} - \frac{n\left(\overline{x}+3\right)}{\theta+1} + \sum_{x=1}^{k}\frac{2\left(x+\theta+2\right)f_x}{\left[x^2+\left(2\theta+5\right)x+\left(\theta^2+4\theta+5\right)\right]}$$

where $\overline{x}$ is the sample mean.

The maximum likelihood estimate (MLE), $\hat{\theta}$ of θ of PAD (2.2) is the solution of the equation $\frac{d\log L}{d\theta} = 0$ and is given by the solution of the following non-linear equation

$$\frac{n\left(\theta^2+4\theta+6\right)}{\theta\left(\theta^2+2\theta+2\right)} - \frac{n\left(\overline{x}+3\right)}{\theta+1} + \sum_{x=1}^{k}\frac{2\left(x+\theta+2\right)f_x}{\left[x^2+\left(2\theta+5\right)x+\left(\theta^2+4\theta+5\right)\right]} = 0$$

This non-linear equation can be solved using any numerical iteration methods such as Newton- Raphson method, Bisection method, Regula-Falsi method etc. In this paper, Newton-Raphson method has been used to solve the above equation for the MLE of the parameter θ.

METHOD OF MOMENT ESTIMATE (MOME)

Let $(x_1, x_2, ..., x_n)$ be a random sample of size n from the PAD (2.2). Equating the population mean to the corresponding sample mean, the MOME, $\tilde{\theta}$ of θ of PAD (2.2) is the solution of the following cubic equation

$$\bar{x}\,\theta^3 + (2\bar{x} - 1)\theta^2 + 2(\bar{x} - 2)\theta - 6 = 0$$

where $\bar{x}$ is the sample mean.

APPLICATIONS AND GOODNESS OF FIT

The PAD has been fitted to a number of data sets to test its goodness of fit over Poisson distribution (PD) and Poisson-Lindley distribution (PLD. The maximum likelihood estimate (MLE) has been used to fit the PAD. In the present paper, two examples of observed count data sets, for which the PAD, PD, and PLD has been fitted, are presented. The data set in Table 2 is the distribution of European corn-borer larvae pyrausta in field corn due to Mc. Guire *et al* (1957) and the data set in Table 3 is due to Kemp and Kemp (1965) regarding the distribution of mistakes in copying groups of random digits.[11,12] The MLE of parameter, standard error of parameter, value of χ^2, degrees of freedom (d.f.) along with p-values have been given in the respective tables. In order to compare distributions, $-2\ln L$, AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) for data set in Table 2 and 3 have also been computed and presented in the tables. The formulae for computing AIC and BIC are as follows:

$AIC = -2\ln L + 2k$, $BIC = -2\ln L + k\ln n$, where $k =$ number of parameters and $n =$ sample size. The best distribution is the distribution which corresponds to lower values of $-2\ln L$, AIC, and BIC.

It is clear from the goodness of fit based on chi-square (χ^2) of PAD, PLD and PD that PAD gives much closer fit than PLD and PD in data set in Table 2 and 3. Further, on the basis of values of $-2\log L$, AIC and BIC, PAD is the best distribution than PD and PLD for data set in Table 2 and PLD is slightly better than PD and PAD for the data set in Table 3. Thus PAD and PLD are competing for modeling discrete data from biological sciences.

The 95% confidence intervals for parameter θ of PD, PLD, and PAD for data set in Table 2 and Table 3 are presented in Table 4.

The profile of likelihood estimation for parameter $\hat{\theta}$ of PD, PLD and PAD for data sets in Table 2 and 3 are presented in Figure 3. The first row contains the profile of likelihood estimation for parameter $\hat{\theta}$ of PD, PLD and PAD for data sets in Table 2 and the second row contains the profile of likelihood estimation for parameter $\hat{\theta}$ of PD, PLD and PAD for data sets in Table 3.

CONCLUDING REMARKS

In this paper, Poisson-Aradhana distribution (PAD) has been obtained by compounding Poisson distribution with Aradhana distribution introduced by Shanker (2016).[1] The expression for the r th factorial moment has been derived and hence its raw moments and central moments have been given. The expression for coefficient of variation, skewness and kurtosis has been obtained. The maximum likelihood estimation and the method of moments for estimating its parameter have been discussed. The distribution has been fitted using maximum likelihood estimate to two real data set to test its goodness of fit over Poisson distribution (PD) and Poisson-Lindley distribution (PLD) and observed that PAD gives much closer fit than PD and PLD in Table 2 while in Table 3 PLD gives slightly better fit than PD and PAD.

TABLE 2: Observed and expected number of European corn- borer of Mc. Guire *et al* (1957).[12]

Number of bores per plant	Observed frequency	Expected frequency		
		PD	PLD	PAD
0	188	169.4	194.0	192.0
1	83	109.8	79.5	81.5
2	36	35.6	31.3	32.0
3	14	7.8	12.0	11.9
4	2	1.2	4.5	4.3
5	1	0.2	2.7	2.3
Total	324	324.0	324.0	324.0
ML Estimate		$\hat{\theta} = 0.648148$	$\hat{\theta} = 2.043252$	$\hat{\theta} = 2.560674$
S.E of parameter		0.04472	0.15131	0.16557
χ^2		15.19	1.29	0.73
d.f.		2	2	2
p-value		0.0005	0.5247	0.6942
$-2\log L$		724.49	714.09	713.00
AIC		726.49	716.09	715.00
BIC		727.00	716.60	715.51

TABLE 3: Distribution of mistakes in copying groups of random digits.

No. of errors per group	Observed Frequency	Expected Frequency		
		PD	PLD	PAD
0	35	27.4	33.0	32.6
1	11	21.5	15.3	15.7
2	8	8.4	6.8	6.9
3	4	2.2	2.9	2.9
4	2	0.5	2.0	1.9
Total	60	60.0	60.0	60.0
ML estimate		$\hat{\theta} = 0.7833$	$\hat{\theta} = 1.7434$	$\hat{\theta} = 2.22937$
S.E. of parameter		0.11426	0.28093	0.31240
χ^2		7.98	2.20	2.03
d.f.		1	1	1
p-value		0.0047	0.1380	0.1542
$-2\log L$		155.09	146.70	146.78
AIC		157.09	148.70	148.78
BIC		156.89	148.48	148.56

TABLE 4: The 95% confidence intervals (C.Is) for parameter of PD, PLD, and PAD for data set in Table 2 3.

Data set	PD		PLD		PAD	
	Lower	Upper	Lower	Upper	Lower	Upper
Table 2	0.56439	0.73981	1.77031	2.36714	2.26010	2.91280
Table 3	0.58021	1.02912	1.27954	2.41068	1.70661	2.96095

Acknowledgements

The author is grateful to the editor-in-chief of the journal and the two anonymous reviewers fot their consttuctive comments which impreved the quality of the paper.

Conflict of Interest

No conflicts of interest between the authors and/or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Financial Support

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Authorship Contributions

While preparing the study; there is no area of conflict of interest during the data collection, interpretation of the results and writing of the article.

REFERENCES

1. Shanker R. Aradhana distribution and its applications. International Journal of Statistics and Applications 2016;6(1):23-34.
2. Lindley DV. Fiducial distributions and Bayes' theorem. J R Stat Soc B 1958;20(1):102-7.
3. Shanker R. Akash distribution and Its Applications. International Journal of Probability and Statistics 2015a;4(3):65-75.
4. Shanker R. Shanker distribution and its applications. International Journal of Statistics and Applications 2015b;5(6):338-48.
5. Shanker R, Hagos F, Sujatha S. On modeling of lifetime data using one parameter Akash, Lindley and exponential distributions. Biom Biostat Int J 2016;3(2):1-10.
6. Shanker R, Fesshaye H. On modeling of lifetime data using Akash, Shanker, Lindley and exponential distributions. Biom Biostat Int J 2016a;3(6):1-9.
7. Shanker R, Fesshaye H. On modeling of lifetime data using Aradhana, Sujatha, Lindley and exponential distributions. Biom Biostat Int J 2016b;4(1):1-12.
8. Sankaran M. The discrete Poisson-Lindley distribution. Biometrics 1970;26:145-9.
9. Ghitany ME, Al-Mutairi DK. Estimation method for the discrete Poisson-Lindley distribution. J Stat Comput Simul 2009;79(1):1-9.
10. Grandell J. Mixed Poisson Processes. 1st ed. London: Chapman & Hall; 1997. p.268.
11. Kemp CD, Kemp AW. Some properties of the 'hermite' distribution. Biometrika 1965;52(3):381-94.
12. McGuire JU, Brindley TA, Bancroft TA. The distribution of European corn-borer larvae pyrausta in field corn. Biometrics 1967;13(1):65-78.