MedCrave Step into the World of Research
Biometrics & Biostatistics International Journal

Research Article

Open Access

On discrete poisson-shanker distribution and its applications

Abstract

A simple method for obtaining moments of Poisson-Shanker distribution (PSD) introduced by Shanker[1] has been proposed. The first four moments about origin and the variance have been obtained. The goodness of fit and the applications of the PSD have been discussed with count data from ecology, genetics and thunderstorms and the fit is compared with one parameter Poisson distribution (PD) and Poisson-Lindley distribution (PLD) introduced by Sankaran.[2]

Keywords: shanker distribution, poisson-shanker distribution, poisson-lindley distribution; Moments; Estimation of parameter; Applications

Volume 5 Issue 1 - 2017

Rama Shanker,[1] Hagos Fesshaye,[2] Ravi Shanker,[3] Tekie Asehun Leonida,[4] Simon Sium[1]

1Department of Statistics, Eritrea Institute of Technology, Eritrea
2Department of Economics, College of Business and Economics, Eritrea
3Department of Mathematics, GLA College NP University, India
4Department of Applied Mathematics, University of Twente, Netherlands

Correspondence: Rama Shanker, Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea, Email shankerrama2009@gmail.com

Received: December 13, 2016 | **Published:** January 18, 2017

Introduction

The Poisson-Shanker distribution (PSD) defined by its probability mass function

$$P(X=x)=\frac{\theta^2}{\theta^2+1}\frac{x+\left(\theta^2+\theta+1\right)}{(\theta+1)^{x+2}}\;;x=0,1,2,...,\theta>0 \qquad (1.1)$$

has been introduced by Shanker[1] for modeling count data-sets. Shanker[1] has shown that PSD is a Poisson mixture of Shanker distribution introduced by Shanker[3] when the parameter λ of the Poisson distribution follows Shanker distribution of Shanker[3] having probability density function

$$f(\lambda;\theta)=\frac{\theta^2}{\theta^2+1}(\theta+\lambda)e^{-\theta\lambda}\;;\lambda>0,\,\theta>0 \qquad (1.2)$$

We have

$$P(X=x)=\int_0^\infty\frac{e^{-\lambda}\lambda^x}{x!}\cdot\frac{\theta^2}{\theta^2+1}(\theta+\lambda)e^{-\theta\lambda}d\lambda \qquad (1.3)$$

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

$$=\frac{\theta^2}{\theta^2+1}\frac{x+\left(\theta^2+\theta+1\right)}{(\theta+1)^{x+2}}\;;x=0,1,2,...,\theta>0\,. \qquad (1.4)$$

Which is the Poisson-Shanker distribution (PSD), as given in (1.1).

Shanker[3] has shown that the Shanker distribution (1.2) is a two component mixture of an exponential (θ) distribution, a gamma (2, θ) distribution with their mixing proportions $\frac{\theta^2}{\theta^2+1}$ and $\frac{1}{\theta^2+1}$ respectively. Shanker[3] has discussed its various mathematical and statistical properties including its shape, moment generating function, moments, skewness, kurtosis, hazard rate function, mean residual life function, stochastic orderings, mean deviations, distribution of order statistics, Bonferroni and Lorenz curves, Renyi entropy measure, stress-strength reliability , amongst others along with estimation of parameter and applications. Shanker & Hagos[4] have detailed study on modeling lifetime data using one parameter Akash distribution introduced by Shanker,[5] Shanker distribution of Shanker,[3] Lindley[6] distribution and exponential distribution.

The probability mass function of Poisson-Lindley distribution (PLD) given by

$$P(X=x)=\frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}}\;;\;x=0,\,1,\,2,\ldots,\;\theta>0\,. \qquad (1.5)$$

has been introduced by Sankaran[2] to model count data. The distribution arises from the Poisson distribution when its parameter λ follows Lindley[6] distribution with its probability density function

$$f(\lambda,\theta)=\frac{\theta^2}{\theta+1}(1+\lambda)\,e^{-\theta\lambda}\;;\quad x>0,\theta>0 \qquad (1.6)$$

Shanker et al.,[7] have critical study on modeling of lifetime data using exponential and Lindley[6] distributions and observed that in some data sets Lindley distribution gives better fit than exponential distribution while in some data sets exponential distribution gives better fit than Lindley distribution. Shanker & Hagos[8] have detailed study on Poisson-Lindley distribution and its applications to model count data from biological sciences.

In this paper a simple method of finding moments of Poisson-Shanker distribution (PSD) introduced by Shanker[1] has been suggested and hence the first four moments about origin and the variance have been presented. It seems that not much work has been done on the applications of PSD so far. The PSD has been fitted to the some data sets relating to ecology, genetics and thunderstorms and the fit is compared with Poisson distribution (PD), and the Poisson-Lindley distribution (PLD). The goodness of fit of PSD shows satisfactory fit in majority of data sets.

Moments

Using (1.3) the r^{th} moment about origin of PSD (1.1) can be obtained as

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

It is clear that the expression under the bracket in (2.1) is the r^{th} moment about origin of the Poisson distribution. Taking $r=1$ in (2.1) and using the first moment about origin of the Poisson distribution, the first moment about origin of the PSD (1.1) can be obtained as

$$\mu_1' = \frac{\theta^2}{\theta^2+1}\int_0^\infty\lambda(\theta+\lambda)e^{-\theta\lambda}d\lambda = \frac{\theta^2+2}{\theta\left(\theta^2+1\right)}$$

Again taking $r=2$ in (2.1) and using the second moment about origin of the Poisson distribution, the second moment about origin of the PSD (1.1) can be obtained as

$$\mu_2' = \frac{\theta^2}{\theta^2+1}\int_0^\infty\left(\lambda^2+\lambda\right)(\theta+\lambda)e^{-\theta\lambda}d\lambda = \frac{\theta^3+2\theta^2+2\theta+6}{\theta^2\left(\theta^2+1\right)}$$

Similarly, taking $r=3$ and 4 in (2.1) and using the third and fourth moments about origin of the Poisson distribution, the third and the fourth moments about origin of the PSD (1.1) are obtained as

$$\mu_3' = \frac{\theta^4+6\theta^3+8\theta^2+18\theta+24}{\theta^3\left(\theta^2+1\right)}$$

$$\mu_4' = \frac{\theta^5+14\theta^4+38\theta^3+66\theta^2+144\theta+120}{\theta^4\left(\theta^2+1\right)}$$

The variance of Poisson-Shanker distribution can thus be obtained as

$$\mu_2 = \sigma^2 = \frac{\theta^5+\theta^4+3\theta^3+4\theta^2+2\theta+2}{\theta^2\left(\theta^2+1\right)^2}$$

Estimation of parameter

Maximum likelihood estimate (MLE) of the parameter: Suppose $(x_1, x_2, ..., x_n)$ is a random sample of size n from the PSD (1.1) and suppose 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 PSD (1.1) is given by

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

The log likelihood function is thus obtained as

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

The first derivative of the log likelihood function is given by

$$\frac{d\log L}{d\theta} = \frac{2n}{\theta\left(\theta^2+1\right)} - \frac{n(\bar{x}+2)}{\theta+1} + \sum_{x=1}^{k}\frac{(2\theta+1)f_x}{x+\left(\theta^2+\theta+1\right)}$$

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

The maximum likelihood estimate (MLE), $\hat{\theta}$ of θ of PSD (1.1) is the solution of the following non-linear equation

$$\frac{2n}{\theta\left(\theta^2+1\right)} - \frac{n(\bar{x}+2)}{\theta+1} + \sum_{x=1}^{k}\frac{(2\theta+1)f_x}{x+\left(\theta^2+\theta+1\right)} = 0$$

This non-linear equation can be solved by any numerical iteration methods such as Newton-Raphson method, Bisection method, Regula-Falsi method etc. In this paper, Newton-Raphson method has been used for estimating the parameter.

Shanker[1] has showed that the MLE of θ of PSD (1.1) is consistent and asymptotically normal.

Method of moment estimate (MOME) of the parameter: Equating the population mean to the corresponding sample mean, the MOME $\tilde{\theta}$ of θ of PSD (1.1) is the solution of the following cubic equation

$$\bar{x}\theta^3 - \theta^2 + \bar{x}\theta - 2 = 0$$

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

Goodness of fit and applications

Since the condition for the applications for Poisson distribution is the independence of events and equality of mean and variance, this condition is rarely satisfied completely in biological and medical science due to the fact that the occurrences of successive events are dependent. Further, the negative binomial distribution is a possible alternative to the Poisson distribution when successive events are possibly dependent, (see Johnson et al.,[9]) but for fitting negative binomial distribution (NBD) to the count data, mean should be less than the variance (over-dispersion). In biological and medical sciences, these conditions are not fully satisfied. Generally, the count data in biological science and medical science are either over-dispersed or under-dispersed. The main reason for selecting PLD and PSD to fit data from biological science and thunderstorms are that these two distributions are always over-dispersed and PSD has some flexibility over PLD.

Count data from ecology and biological sciences

In this section we fit Poisson distribution (PD), Poisson -Lindley distribution (PLD) and Poisson-Shanker distribution (PSD) to many count data from ecology and biological sciences using maximum likelihood estimate. The data were on haemocytometer yeast cell counts per square, on European red mites on apple leaves and European corn borers per plant. Recall that Shanker & Hagos[7] have fitted Poisson-Lindley distribution(PLD) to the same data sets.

It is obvious from above tables that in Table 1, PD gives better fit than PLD and PSD; in Table 2, PSD gives better fit than PD and PLD while in Table 3, PLD gives better fit than PD and PSD.

Count data from genetics

In this section we fit PSD, PLD and PD using maximum likelihood estimate to count data relating to genetics. Recall that Shanker & Hagos[8] have fitted Poisson-Lindley distribution to the same data sets. The data set in Table 4 is available in Loeschke & Kohler,[13] and Janardan & Schaeffer.[14] The data sets in Tables 5-7 are available in Catcheside et al.,[15,16]

It is obvious from the fitting of PSD, PLD, and PD that both PSD and PLD gives much satisfactory fit than PD. Further, PSD gives much closer fit than both PLD and PD in almost all data sets.

Count data from thunderstorms

In this section, we fit PSD, PLD and PD to count data from thunderstorms available in Falls et al.,[17]

It is obvious from the fitting of PSD, PLD and PD to thunderstorms data that PLD gives better fit than both PSD and PD in Table 8, 9 and 11 while PSD gives better fit than both PLD and PD in Table 10.

Table 1 Observed and expected number of Haemocytometer yeast cell counts per square observed by Gosset.[10]

Number of yeast cells per square	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	213	202.1	234.0	233.2
1	128	138.0	99.4	99.6
2	37	47.1	40.5	41.0
3	18	10.7	16.0	16.3
4	3	1.8	6.2	6.7
5	1	0.2	2.4	2.3
6	0	0.1	1.5	0.9
Total	400	400.0	400.0	400.0
ML estimate		$\hat{\theta}$=0.6825	$\hat{\theta}$=1.950236	$\hat{\theta}$=1.795126
χ^2		10.08	11.04	12.25
d.f.		2	2	2
p-value		0.0065	0.0040	0.0023

Table 2 Observed and expected number of red mites on Apple leaves, available in Fisher et al.,[11]

Number mites per leaf	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	38	25.3	35.8	36.0
1	17	29.1	20.7	20.6
2	10	16.7	11.4	11.2
3	9	6.4	6.0	6.0
4	3	1.8	3.1	3.1
5	2	0.4	1.6	1.6
6	1	0.2	0.8	0.8
7+	0	0.1	0.6	0.7
Total	80	80.0	80.0	80.0
ML estimate		$\hat{\theta}$=1.15	$\hat{\theta}$=1.255891	$\hat{\theta}$=1.219731
χ^2		18.27	2.47	2.37
d.f.		2	3	3
p-value		0.0001	0.4807	0.4992

Citation: Shanker R, Fesshaye H, Shanker R, et al. On discrete poisson-shanker distribution and its applications. *Biom Biostat Int J.* 2017;5(1):6–14. DOI: 10.15406/bbij.2017.05.00121

Table 3 Observed and expected number of European corn- borer of Mc Guire et al.,[12]

Number of bores per plant	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	188	169.4	194.0	195.0
1	83	109.8	79.5	78.4
2	36	35.6	31.3	31.0
3	14	7.8	12.0	12.1
4	2	1.2	4.5	4.6
5	1	0.2	2.7	2.9
Total	324	324.0	324.0	324.0
ML estimate		$\hat{\theta}$=0.648148	$\hat{\theta}$=2.043252	$\hat{\theta}$=1.879553
χ^2		15.19	1.29	1.67
d.f.		2	2	2
p-value		0.0005	0.5247	0.4338

Table 4 Distribution of number of Chromatid aberrations (0.2 g chinon 1, 24 hours)

Number of aberrations	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	268	231.3	257.0	258.3
1	87	126.7	93.4	92.1
2	26	34.7	32.8	32.4
3	9	6.3	11.2	11.3
4	4	0.8	3.8	3.9
5	2	0.1	1.2	1.3
6	1	0.1	0.4	0.5
7+	3	0.1	0.2	1.5
Total	400	400.0	400.0	400.0
ML estimate		$\hat{\theta}$=0.5475	$\hat{\theta}$=2.380442	$\hat{\theta}$=2.162674
χ^2		38.21	6.21	3.45
d.f.		2	3	3
p-value		0.0000	0.1018	0.3273

Table 5 Mammalian cytogenetic dosimetry lesions in rabbit lymphoblast induced by streptonigrin (NSC-45383), Exposure-60 $\mu g|kg$

Class/exposure ($\mu g|kg$)	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	413	374.0	405.7	407.1
1	124	177.4	133.6	131.9
2	42	42.1	42.6	42.3
3	15	6.6	13.3	13.5
4	5	0.8	4.3	4.3
5	0	0.1	1.3	1.3
6	2	0.0	0.6	0.6
Total	601	601.0	601.0	601.0
ML Estimate		$\hat{\theta}$=0.47421	$\hat{\theta}$=2.685373	
χ^2		48.17	1.34	0.82
d.f.		2	3	3
p-value		0.0000	0.7196	0.8446

Table 6 Mammalian cytogenetic dosimetry lesions in rabbit lymphoblast induced by streptonigrin (NSC-45383), Exposure-70 $\mu g|kg$

Class/exposure ($\mu g|kg$)	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	200	172.5	191.8	192.7
1	57	95.4	70.3	69.4
2	30	26.4	24.9	24.6
3	7	4.9	8.6	8.7
4	4	0.7	2.9	3.0
5	0	0.1	1.0	1.0
6	2	0.0	0.5	0.6
Total	300	300.0	300.0	300.0
ML Estimate		$\hat{\theta}$=0.55333	$\hat{\theta}$=2.353339	$\hat{\theta}$=2.138048
χ^2		29.68	3.91	3.66
d.f.		2	2	2
p-value		0.0000	0.1415	0.1604

Citation: Shanker R, Fesshaye H, Shanker R, et al. On discrete poisson-shanker distribution and its applications. *Biom Biostat Int J.* 2017;5(1):6–14. DOI: 10.15406/bbij.2017.05.00121

Table 7 Mammalian cytogenetic dosimetry lesions in rabbit lymphoblast induced by streptonigrin (NSC-45383), Exposure -90 $\mu g|kg$

Class/exposure ($\mu g|kg$)	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	155	127.8	158.3	159.3
1	83	109.0	77.2	76.3
2	33	46.5	35.9	35.4
3	14	13.2	16.1	16.1
4	11	2.8	7.1	7.2
5	3	0.5	3.1	3.2
6	1	0.2	2.3	2.5
Total	300	300.0	300.0	300.0
ML Estimate		$\hat{\theta}$=0.853333		$\hat{\theta}$=1.520805
χ^2		24.97	1.51	1.48
d.f.		2	3	3
p-value		0.0000	0.6799	0.6868

Table 8 Observed and expected number of days that experienced X thunderstorms events at Cape Kennedy, Florida for the 11-year period of record for the month of June, January 1957 to December 1967, Falls et al.,[17]

No. of thunderstorms	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	187	155.6	185.3	186.4
1	77	117.0	83.5	82.3
2	40	43.9	35.9	35.5
3	17	11.0	15.0	15.0
4	6	2.1	6.1	6.3
5	2	0.3	2.5	2.6
6	1	0.1	1.7	1.9
Total	330	330.0	330.0	330.0
ML estimate		$\hat{\theta}$=0.751515		$\hat{\theta}$=1.679053
χ^2		31.93	1.43	1.48
d.f.		2	3	3
p-value		0.0000	0.6985	0.6869

Table 9 Observed and expected number of days that experienced X thunderstorms events at Cape Kennedy, Florida for the 11-year period of record for the month of July, January 1957 to December 1967, Falls et al.,[17]

No. of thunderstorms	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	177	142.3	177.7	178.7
1	80	124.4	88.0	86.9
2	47	54.3	41.5	41.0
3	26	15.8	18.9	18.9
4	9	3.5	8.4	8.6
5	2	0.7	6.5	6.9
Total	341	341.0	341.0	341.0
ML estimate		$\hat{\theta} = 0.873900$	$\hat{\theta}$=1.583536	$\hat{\theta}$=1.497274
χ^2		39.74	5.15	5.41
d.f.		2	3	3
p-value		0.0000	0.1611	0.1441

Table 10 Observed and expected number of days that experienced X thunderstorms events at Cape Kennedy, Florida for the 11-year period of record for the month of August, January 1957 to December 1967, Falls et al.,[17]

No. of thunderstorms	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	185	151.8	184.8	186.0
1	89	122.9	87.2	86.1
2	30	49.7	39.3	38.8
3	24	13.4	17.1	17.1
4	10	2.7	7.3	7.4
5	3	0.5	5.3	5.6
Total	341	341.0	341.0	341.0
ML estimate		$\hat{\theta}$=0.809384	$\hat{\theta}$=1.693425	$\hat{\theta}$=1.586731
χ^2		49.49	5.03	4.87
d.f.		2	3	3
p-value		0.0000	0.1696	0.1816

Citation: Shanker R, Fesshaye H, Shanker R, et al. On discrete poisson-shanker distribution and its applications. *Biom Biostat Int J.* 2017;5(1):6–14. DOI: 10.15406/bbij.2017.05.00121

Table 11 Observed and expected number of days that experienced X thunderstorms events at Cape Kennedy, Florida for the 11-year period of record for the summer, January 1957 to December 1967, Falls et al.,[17]

No. of thunderstorms	Observed frequency	Expected frequency		
		PD	PLD	PSD
0	549	547.5	547.5	550.8
1	246	364.8	259.0	255.7
2	117	148.2	116.9	115.5
3	67	40.1	51.2	51.1
4	25	8.1	21.9	22.3
5	7	1.3	9.2	9.6
6	1	0.3	6.3	7.0
Total	1012	1012.0	1012.0	1012.0
ML estimate		$\hat{\theta}$=0.812253	$\hat{\theta}$=1.688990	$\hat{\theta}$=1.582475
χ^2		141.42	9.60	10.09
d.f.		3	4	4
p-value		0.0000	0.0477	0.0389

Concluding remarks

In the present paper, a simple and interesting method for finding moments of Poisson-Shanker distribution (PSD) has been suggested and thus the first four moments about origin and the variance have been obtained. The goodness of fit of PSD has been discussed with several data from ecology, genetics and thunderstorms and the fit has been compared with Poisson distribution (PD) and Poisson-Lindley distribution (PLD).

Acknowledgement

None.

Conflict of interest

None.

References

1. Shanker R. The discrete Poisson–Shanker distribution. *Jacobs Journal of Biostatistics*. 2016;1(1):1–7.
2. Sankaran M. The discrete Poisson–Lindley distribution. *Biometrics*. 1970;26(1):145–149.
3. Shanker R. Shanker distribution and its applications. *International Journal of Statistics and Applications*. 2015;5(6):338–348.
4. Shanker R, Hagos F. On modeling of lifetime data using Akash, Shanker, Lindley and exponential distributions. Biometrics & *Biostatistics International Journal*. 2016;3(2):1–10.
5. Shanker R. Akash distribution and its applications. *International Journal of Probability and Statistics*. 2015;4(3):65–75.
6. Lindley DV. Fiducial distributions and Bayes theorem. *Journal of Royal Statistical Society*. 1958;20(1):102–107.
7. Shanker R, Hagos F, Sujatha S. On modeling of lifetime data using exponential and Lindley distributions. *Biometrics & Biostatistics International Journal*. 2015;2(5):1–9.
8. Shanker R, Hagos F. On Poisson–Lindley distribution and Its applications to Biological Sciences. *Biometrics & Biostatistics International Journal*. 2015;2(4):1–5.
9. Johnson NL, Kotz S, Kemp AW. Univariate Discrete Distributions. *John Wiley & sons Inc*. 2nd ed. USA: 1992.
10. Loeschke V, Kohler W. Deterministic and Stochastic models of the negative binomial distribution and the analysis of chromosomal aberrations in human leukocytes. *Biometrische Zeitschrift*. 1976;18(6):427–451.
11. Janardan KG, Schaeffer DJ. Models for the analysis of chromosomal aberrations in human leukocytes. *Biometrical Journal*. 1977;19(8):599–612.
12. Catcheside DG, Lea DE, Thoday JM. Types of chromosome structural change induced by the irradiation on Tradescantia microspores. *J Genet*. 1946;47:113–136.
13. Catcheside DG, Lea DE, Thoday JM. The production of chromosome structural changes in Tradescantia microspores in relation to dosage, intensity and temperature. *J Genet*. 1946;47: 137–149.
14. Falls LW, Williford WO, Carter MC. Probability distributions for thunderstorm activity at Cape Kennedy, Florida. Journal of *Applied Meteorology*. 1971;10(1):97–104.

15. Gosset WS. The probable error of a mean. *Biometrika.* 1908;6(1):1–25.
16. Fisher RA, Corpet AS, Williams CB. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology*. 1943;12(1):42–58.
17. Mc Guire JU, Brindley TA, Bancroft TA. The distribution of European corn–borer larvae pyrausta in field corn. *Biometrics*. 1957;13(1):65–78.

Citation: Shanker R, Fesshaye H, Shanker R, et al. On discrete poisson-shanker distribution and its applications. *Biom Biostat Int J.* 2017;5(1):6–14. DOI: 10.15406/bbij.2017.05.00121