Research article
Open Access
CrossMark

Statistical and reliability properties of poisson-suja distribution

Abstract

The Poisson-Suja distribution, the Poisson compound of Suja distribution, was introduced to model count data of over-dispersed nature from biological sciences. The main purpose of this paper is to derive some statistical and reliability properties of Poisson-Suja distribution which have not been studied earlier. It has been observed that the proposed distribution is log-concave and is a two-component mixture of negative binomial distributions. The cumulative distribution function, survival function and hazard function of the Poisson-Suja distribution have been derived and their natures for varying values of parameter have been presented. The expressions for other reliability parameter including reverse hazard rate function, the second rate of failure, the cumulative hazard function and the Mills ratio of the Poisson-Suja distribution have also been derived. The simulation study has been presented to show the consistency of maximum likelihood estimator of the parameter of Poisson-Suja distribution. The Poisson-Suja distribution has been applied to two real datasets and found to provide quite satisfactory fit over other competing one parameter over-dispersed count distributions namely Poisson-Lindley distribution, Poisson-Komal distribution, Poisson-Shanker distribution, and Poisson-Garima distribution.

Keywords: suja distribution, poisson-suja distribution, statistical properties, reliability properties, estimation of parameter, simulation, applications

Volume 14 Issue 2 - 2025

Rama Shanker, Riki Tabassum, Jyotirmoyee Baishya, Mousumi Ray
Department of Statistics, Assam University, Silchar, India

Correspondence: Mousumi Ray, Department of Statistics, Assam University, Silchar, India

Received: June15, 2025 | **Published:** November 10, 2025

Introduction

The Poisson distribution which is a limiting distribution of binomial distribution is the first discrete distribution for count data of equi-dispersed (mean equal to variance) nature. Count data arise in almost every field of knowledge including insurance, agriculture, environmental science, biological sciences, clinical trials, engineering etc. It has been found that the count data are in general either over-dispersed (variance greater than mean) or under-dispersed (variance less than mean) and hence Poisson distribution has limited applications due to its limitation of mean equal to variance. During recent decades several over-dispersed count distributions have been introduced in statistical literature including negative binomial distribution which is the Poisson compound of gamma distribution, Poisson-Lindley distribution (PLD) of Sankaran[1] which is the Poisson compound of Lindley distribution of Lindley,[2] Poisson-Shanker distribution (PShD) of Shanker[3] which is the Poisson compound of Shanker distribution of Shanker,[4] Poisson-Garima distribution (PGD) of Shanker[5] which is the Poisson compound of Garima distribution of Shanker,[6] Poisson-Komal distribution (PKD) of Shanker et al.[7] which is the Poisson compound of Komal distribution of Shanker.[8] It should be noted that Shanker et al.[9] have detailed study on statistical and reliability properties along with some new applications of Poisson-Garima distribution.

The Suja distribution introduced by Shanker (2017)[10] is defined by its probability density function (PDF)

$$f(\lambda;\theta)=\frac{\theta^5}{\theta^4+24}\left(1+\lambda^4\right)e^{-\theta\lambda}\quad;\lambda>0,\ \theta>0$$

Length- biased Suja distribution, power length-biased Suja distribution and weighted Suja distribution have been proposed and studied by Al-Omari and Alsmairan,[11] Al-Omari et al.[12] and Alsmairan and Al-Omari,[13] respectively. Todorka et al.[14] have studied on the CDF of various modifications of Suja distributions and discussed their applications in the field of the analysis of computer- viruses' propagation and debugging theory. Shanker et al.[15] introduced quasi Suja distribution and discussed its statistical and reliability properties along with applications. A bimodal extension of Suja distribution using quadratic rank transformation map been proposed by Enogwe et al.[16] An extended Suja distribution has been suggested by Shanker et al.,[17] Enogwe et al.,[18] proposed generalized Suja distribution using Kumaraswamy generator and discussed its properties and applications. Abu-Zinadah and Alsumairi[19] discussed several estimating methods for estimating parameter of Suja distribution.

Recently, Shanker et al.[20] derived Poisson-Suja distribution (PSD) by compounding Poisson distribution with Suja distribution. The genesis of PSD is that a random variable X is said to be PSD if it follows the following stochastic representation $X\mid\lambda\sim\text{Poisson}(\lambda)$ distribution and $\lambda\mid\theta\sim\text{Suja}(\theta)$ distribution for $\lambda>0,\theta>0$ and the unconditional distribution of this stochastic representation is the PSD(θ). That is, the probability mass function (PMF) of PSD is obtained as

$$P(x;\theta)=\int_0^\infty P(X=x\mid\lambda)f(\lambda;\theta)\,d\lambda$$

$$=\int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}\frac{\theta^5}{\theta^4+24}\left(1+\lambda^4\right)e^{-\theta\lambda}\,d\lambda$$

$$=\frac{\theta^5}{\theta^4+24}\frac{x^4+10x^3+35x^2+50x+24+(\theta+1)^4}{(\theta+1)^{x+5}}\ ;x=0,1,2,...,\ \theta>0.$$

The natures of PMF of PSD for varying values of parameter are shown in Figure 1 and it is clear that it is highly positively skewed for increasing values of parameter.

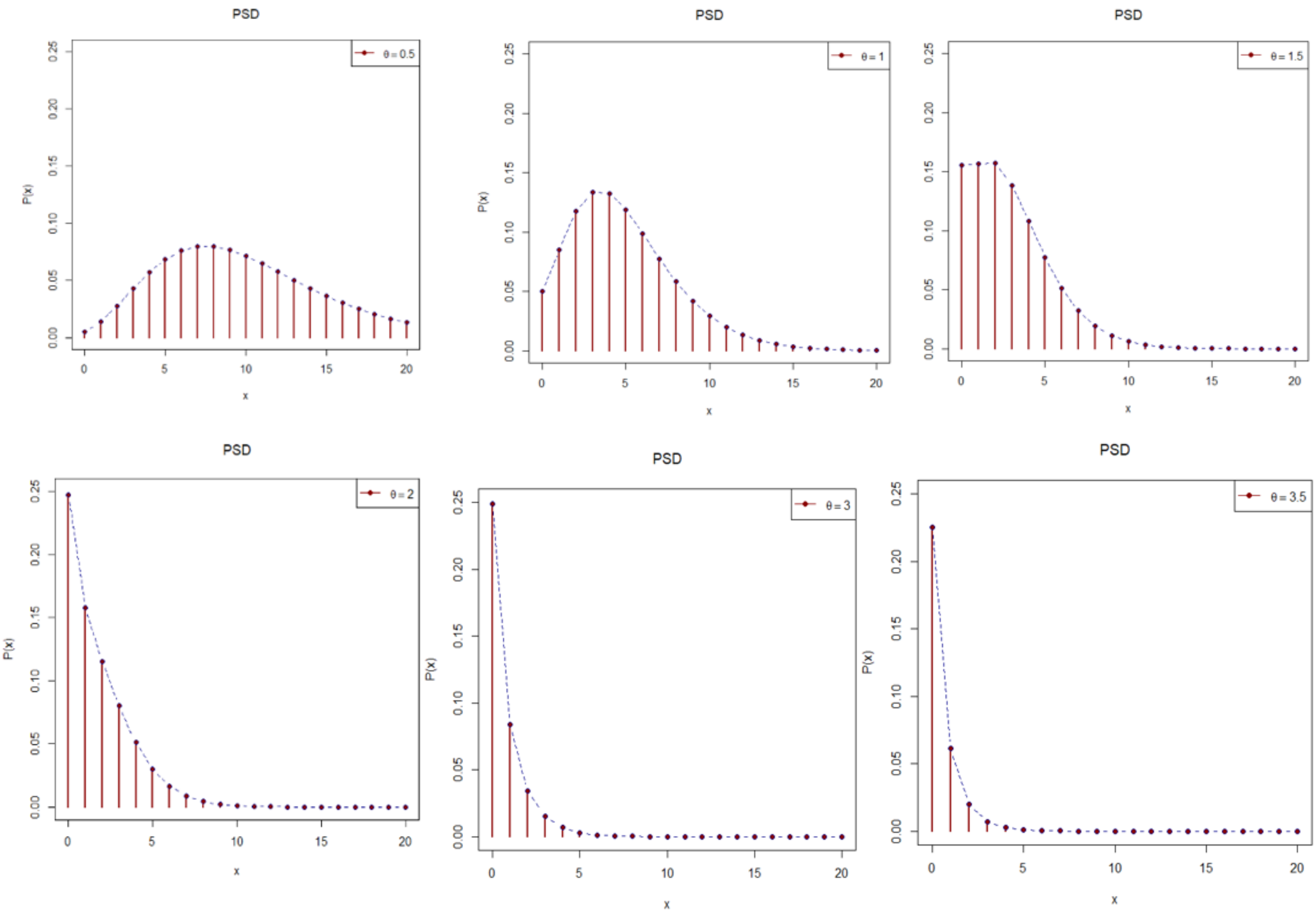


Figure 1 PMF of PSD for varying values of parameter.

The first four moments about origin and the variance of PSD are given by

$$\mu_1' = \mu_{(1)}' = \frac{\theta^4 + 120}{\theta\left(\theta^4 + 24\right)}$$

$$\mu_2' = \frac{\theta^5 + 2\theta^4 + 120\theta + 720}{\theta^2\left(\theta^4 + 24\right)}$$

$$\mu_3' = \frac{\theta^6 + 6\theta^5 + 6\theta^4 + 120\theta^2 + 2160\theta + 5040}{\theta^3\left(\theta^4 + 24\right)}$$

$$\mu_4' = \frac{\theta^7 + 14\theta^6 + 36\theta^5 + 24\theta^4 + 120\theta^3 + 5040\theta^2 + 30240\theta + 40320}{\theta^4\left(\theta^4 + 24\right)}$$

$$\mu_2 = \sigma^2 = \frac{\theta^9 + \theta^8 + 144\theta^5 + 528\theta^4 + 2880\theta + 2880}{\theta^2\left(\theta^4 + 24\right)^2}$$

The detailed discussion about its properties, estimation of parameter, and applications has been discussed by Shanker et al.[7] It seems that some interesting and useful statistical and reliability properties of PSD distribution have not been discussed by Shanker et al.[20]

The main purpose of this paper is to derive some interesting statistical and reliability properties of PSD, which has not been studied earlier by Shanker et al.[20] We have shown that the proposed distribution is log-concave and is a two-component mixture of negative binomial distributions. The natures of cumulative distribution function, survival function and hazard function of the distribution has been presented graphically. For showing the consistency of the maximum likelihood estimator of the distribution, simulation study has been presented. The goodness of fit of the PSD has been established with two datasets and observed that it provides best fit over other competing one parameter over-dispersed distributions.

Statistical properties of poisson-suja distribution

In this section an attempt has been made to prove some important statistical results of PSD. The PSD is positively skewed, unimodal and decreasing which is supported by theorems 1 and 2. In theorem 3, it has been shown that PSD is also a two-component mixture of negative binomial distributions with different parameter (number of successes) and for the same probability of success.

The PSD has increasing hazard rate (IHR) and is unimodal. Since

$$Q(x,\theta)=\frac{P(x+1,\theta)}{P(x,\theta)}=\frac{1}{\theta+1}\left[1+\frac{4x^3+36x^2+104x+96}{\left(x^4+10x^3+35x^2+50x+24\right)+(\theta+1)^4}\right]$$

is a decreasing function of x for a given θ, $P(x,\theta)$ is log-concave. This implies that PSD has an increasing hazard rate and is unimodal.

Grandell[21] has detailed discussion about relationship between log-concavity, IHR and Unimodality of discrete distributions.

Theorem 1: The $Q(x,\theta)$ is decreasing function of x for given θ.

Proof: We have

$$Q(x,\theta)=\frac{P(x+1,\theta)}{P(x,\theta)}=\frac{1}{\theta+1}\left[1+\frac{4x^3+36x^2+104x+96}{\left(x^4+10x^3+35x^2+50x+24\right)+(\theta+1)^4}\right]$$

Differentiating partially with respect to x, we get

$$Q'(x,\theta)=-\frac{\begin{bmatrix}\left(4x^6+72x^5+532x^4+2064x^3+8072x^2+4992x-1336\right)\\ -\left(12x^2+72x+104\right)(\theta+1)^4\end{bmatrix}}{(\theta+1)\left[\left(x^4+10x^3+35x^2+50x+24\right)+(\theta+1)^4\right]^2}.$$

Since $Q'(x,\theta)<0$, $Q(x,\theta)$ is decreasing function of x for given θ.

Theorem 2: The pmf $P(x,\theta)$ of PSD is log-concave

Proof: We have

$$P(x;\theta)=\frac{\theta^5}{\theta^4+24}\,\frac{x^4+10x^3+35x^2+50x+24+(\theta+1)^4}{(\theta+1)^{x+5}}$$

$$\log P(x,\theta)=5\log\theta+\log\left\{x^4+10x^3+35x^2+50x+24+(\theta+1)^4\right\}-\log\left(\theta^4+24\right)-(x+5)\log(\theta+1)$$

Assuming $g(x,\theta)=\log P(x,\theta)$, and differentiating partially twice with respect to x, we get

$$g'(x,\theta)=\frac{4x^3+30x^2+70x+50}{x^4+10x^3+35x^2+50x+24+(\theta+1)}-\log(\theta+1)$$

and

$$g''(x,\theta)=-\frac{\begin{bmatrix}\left(4x^6+60x^5+370x^4+1200x^3+2162x^2+3500x+820\right)\\ -\left(12x^2+60x+70\right)(\theta+1)^4\end{bmatrix}}{\left[x^4+10x^3+35x^2+50x+24+(\theta+1)^4\right]^2}<0$$

.

This means that the PMF of PSD is log-concave.

Theorem 3: The PSD is a two-component mixture of negative binomial distributions and can be expressed as

$$P(x;\theta)=p_1\,P_1(x;\theta)+p_2P_2(x;\theta),$$

Where $P_i(x;\theta)$ is the PMF of the negative binomial distribution (NBD) with parameters, the number of successes i and $p_1=\frac{\theta^4}{\theta^4+24}$ with $P_1(x;\theta)=\frac{\theta}{(\theta+1)^{x+1}}$ as $NBD\left(1,\frac{\theta}{\theta+1}\right)$, $p_2=\frac{24}{\theta^4+24}$ with $P_2(x;\theta)=\frac{(x+1)(x+2)(x+3)(x+4)\theta^5}{24(\theta+1)^{x+5}}$ as the $NBD\left(5,\frac{\theta}{\theta+1}\right)$ .respectively.

Proof: We have

$$P(x;\theta)=\int_0^\infty\frac{e^{-\lambda}\lambda^x}{x!}\frac{\theta^5}{\theta^4+24}\left(1+\lambda^4\right)e^{-\theta\lambda}d\lambda$$

$$=\int_0^\infty\frac{e^{-\lambda}\lambda^x}{x!}\left\{\frac{\theta^4}{\theta^4+24}\left(\theta e^{-\theta\lambda}\right)+\frac{24}{\theta^4+24}\left(\frac{\theta^5}{\Gamma(5)}e^{-\theta\lambda}\lambda^4\right)\right\}d\lambda$$

$$=\frac{\theta^4}{\theta^4+24}\left[\frac{\theta}{x!}\int_0^\infty e^{-(\theta+1)\lambda}\lambda^{x+1-1}d\lambda\right]+\frac{24}{\theta^4+24}\left[\frac{\theta^5}{x!\Gamma(5)}\int_0^\infty e^{-(\theta+1)\lambda}\lambda^{x+5-1}d\lambda\right]$$

$$=\frac{\theta^4}{\theta^4+24}\left[\frac{\theta}{x!}\frac{\Gamma(x+1)}{(\theta+1)^{x+1}}\right]+\frac{24}{\theta^4+24}\left[\frac{\theta^5}{x!\Gamma(5)}\frac{\Gamma(x+5)}{(\theta+1)^{x+5}}\right]$$

$$=\frac{\theta^4}{\theta^4+24}\left[\frac{\theta}{(\theta+1)^{x+1}}\right]+\frac{24}{\theta^4+24}\left[\frac{(x+1)(x+2)(x+3)(x+4)\theta^5}{24(\theta+1)^{x+5}}\right]$$

$$=\frac{\theta^4}{\theta^4+24}\left[\binom{x+1-1}{x}\left(\frac{\theta}{\theta+1}\right)^1\left(\frac{1}{\theta+1}\right)^x\right]+\frac{24}{\theta^4+24}\left[\binom{x+5-1}{x}\left(\frac{\theta}{\theta+1}\right)^5\left(\frac{1}{\theta+1}\right)^x\right]$$

$$=\frac{\theta^4}{\theta^4+24}\left[NBD\left(1,\frac{\theta}{\theta+1}\right)\right]+\frac{24}{\theta^4+24}\left[NBD\left(5,\frac{\theta}{\theta+1}\right)\right]$$

.

Now, even though the PSD is a two-component mixture of NBD's, the presence of two modes is not perceptible in any of the plots of PMF of PSD in figure 1 for the selected values of θ. This suggests that the two modes coming from the two sub-populations are very close to each other. Tajuddin et al.[22] observed that if the modes of the sub-populations are very close to each other, then the population will be unimodal. This means that the distribution in which the existence of the modes of sub-populations are very close, the distribution would be suitable for over-dispersed count data.

Reliability properties of poisson-suja distribution

Various interesting and useful reliability properties including reverse hazard rate function, second rate of failure, cumulative hazard function and Mills ratio of a distribution depends on cumulative distribution function, survival function and hazard function of the distribution. The following theorem 4 deals with the cumulative distribution function (CDF), survival function and the hazard function of PSD. The expression for reverse hazard rate function, second rate of failure, cumulative hazard function and Mills ratio of PSD have also been obtained.

Theorem 4: The cumulative distribution function (CDF), survival function and the hazard function of PSD are given by

$$F(x)=F(x;\theta)=1-\frac{\left\{\begin{array}{l}\theta^4x^4+\left(14\theta^4+4\theta^3\right)x^3+\left(71\theta^4+48\theta^3+12\theta^2\right)x^2\\+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)x\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4\left(\theta+1\right)^4\end{array}\right\}}{\left(\theta^4+24\right)\left(\theta+1\right)^{x+5}}$$

$$S(x)=S(x;\theta)=\frac{\left\{\begin{array}{l}\theta^4x^4+\left(14\theta^4+4\theta^3\right)x^3+\left(71\theta^4+48\theta^3+12\theta^2\right)x^2\\+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)x\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4\left(\theta+1\right)^4\end{array}\right\}}{\left(\theta^4+24\right)\left(\theta+1\right)^{x+5}}\text{, and}$$

Proof: We have

$$F(x)=F(x,\theta)=P(X\le x)=1-P(X\ge x+1)$$

$$=1-\sum_{t=x+1}^{\infty}\int_0^{\infty}P(X=t\,|\,\lambda)f(\lambda\,|\,\theta)\,d\lambda$$

$$=1-\sum_{t=x+1}^{\infty}\int_0^{\infty}\frac{e^{-\lambda}\lambda^t}{t!}\frac{\theta^5}{\theta^4+24}\left(1+\lambda^4\right)e^{-\theta\lambda}d\lambda$$

$$=1-\sum_{t=x+1}^{\infty}\frac{\theta^5}{\left(\theta^4+24\right)t!}\int_0^{\infty}e^{-(\theta+1)\lambda}\lambda^t\left(1+\lambda^4\right)d\lambda$$

$$=1-\sum_{t=x+1}^{\infty}\frac{\theta^5}{\left(\theta^4+24\right)t!}\left[\frac{\Gamma(t+1)}{(\theta+1)^{t+1}}+\frac{\Gamma(t+5)}{(\theta+1)^{t+5}}\right]$$

$$=1-\sum_{t=x+1}^{\infty}\frac{\theta^5}{\theta^4+24}\left[\frac{\left(t^4+10t^3+35t^2+50t+24\right)+(\theta+1)^4}{(\theta+1)^{t+5}}\right]$$

$$=1-\frac{\theta^5}{\left(\theta^4+24\right)(\theta+1)^5}\sum_{t=x+1}^{\infty}\frac{\left(t^4+10t^3+35t^2+50t+24\right)+(\theta+1)^4}{(\theta+1)^t}$$

$$=1-\frac{\left\{\begin{array}{l}\theta^4x^4+\left(14\theta^4+4\theta^3\right)x^3+\left(71\theta^4+48\theta^3+12\theta^2\right)x^2\\+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)x\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4\left(\theta+1\right)^4\end{array}\right\}}{\left(\theta^4+24\right)\left(\theta+1\right)^{x+5}}$$

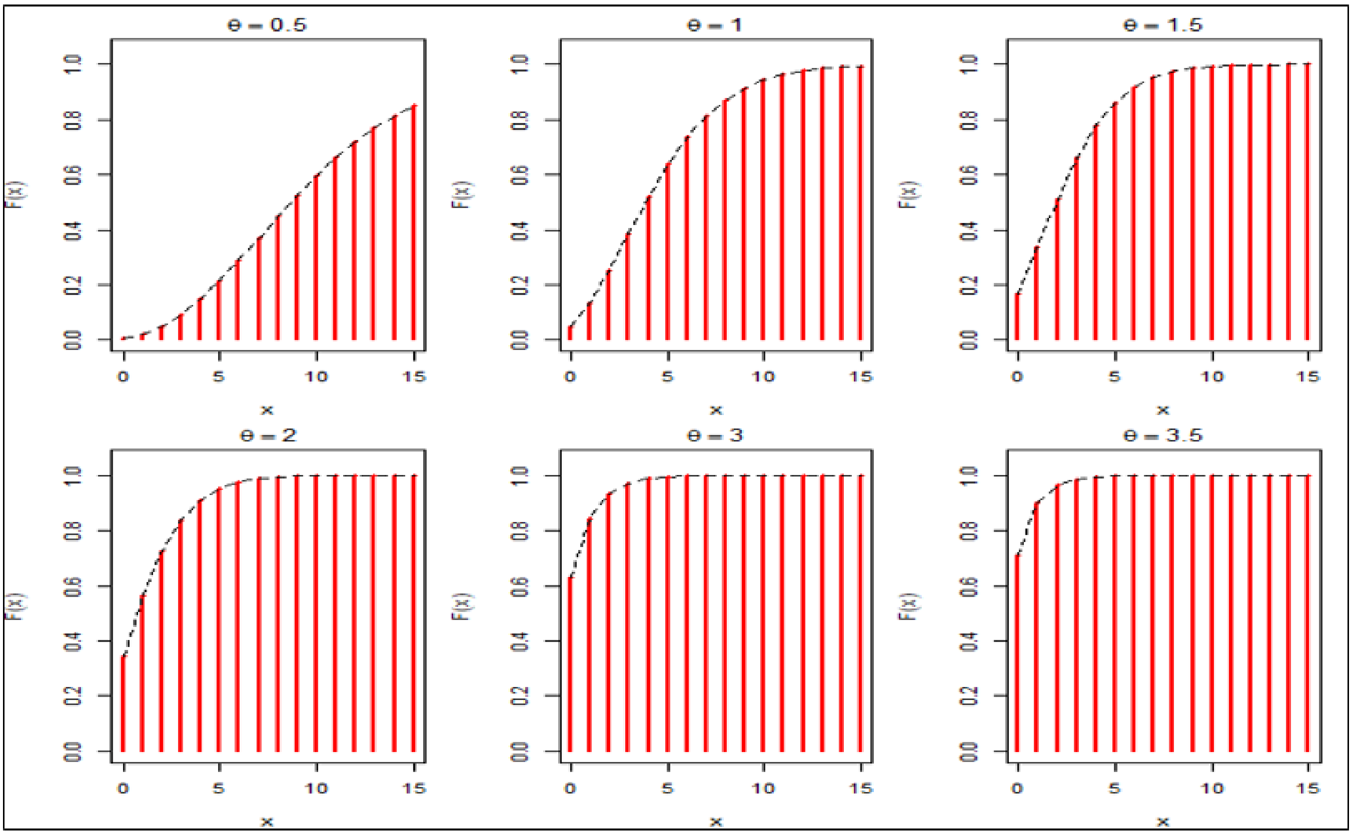
The survival function of PSD can thus be given by

$$S(x)=S(x,\theta)=1-F(x,\theta)=\frac{\left\{\begin{array}{l}\theta^4x^4+\left(14\theta^4+4\theta^3\right)x^3+\left(71\theta^4+48\theta^3+12\theta^2\right)x^2\\+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)x\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4\left(\theta+1\right)^4\end{array}\right\}}{\left(\theta^4+24\right)\left(\theta+1\right)^{x+5}}$$

The hazard function of PSD can be expressed as

$$h(x)=h(x,\theta)=\frac{P(x,\theta)}{S(x,\theta)}=\frac{\theta^5\left\{x^4+10x^3+35x^2+50x+24+\left(\theta+1\right)^4\right\}}{\left\{\begin{array}{l}\theta^4x^4+\left(14\theta^4+4\theta^3\right)x^3+\left(71\theta^4+48\theta^3+12\theta^2\right)x^2\\+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)x\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4\left(\theta+1\right)^4\end{array}\right\}}$$

The natures of CDF, survival function and hazard function of PSD for varying values of parameter are shown in the following Figure 2 and it is obvious from the figure that the PSD has a valid CDF since $F(x)\to 1$ as $x\to\infty$. Further, the hazard rate function shows an increasing pattern with a limiting value of θ, which means that $\lim_{x\to\infty} h(x)=\theta$.



Citation: Shanker R, Tabassum R, Baishya J, et al. Statistical and reliability properties of poisson-suja distribution. *Biom Biostat Int J.* 2025;14(2):85‒94.
DOI: 10.15406/bbij.2025.14.00438

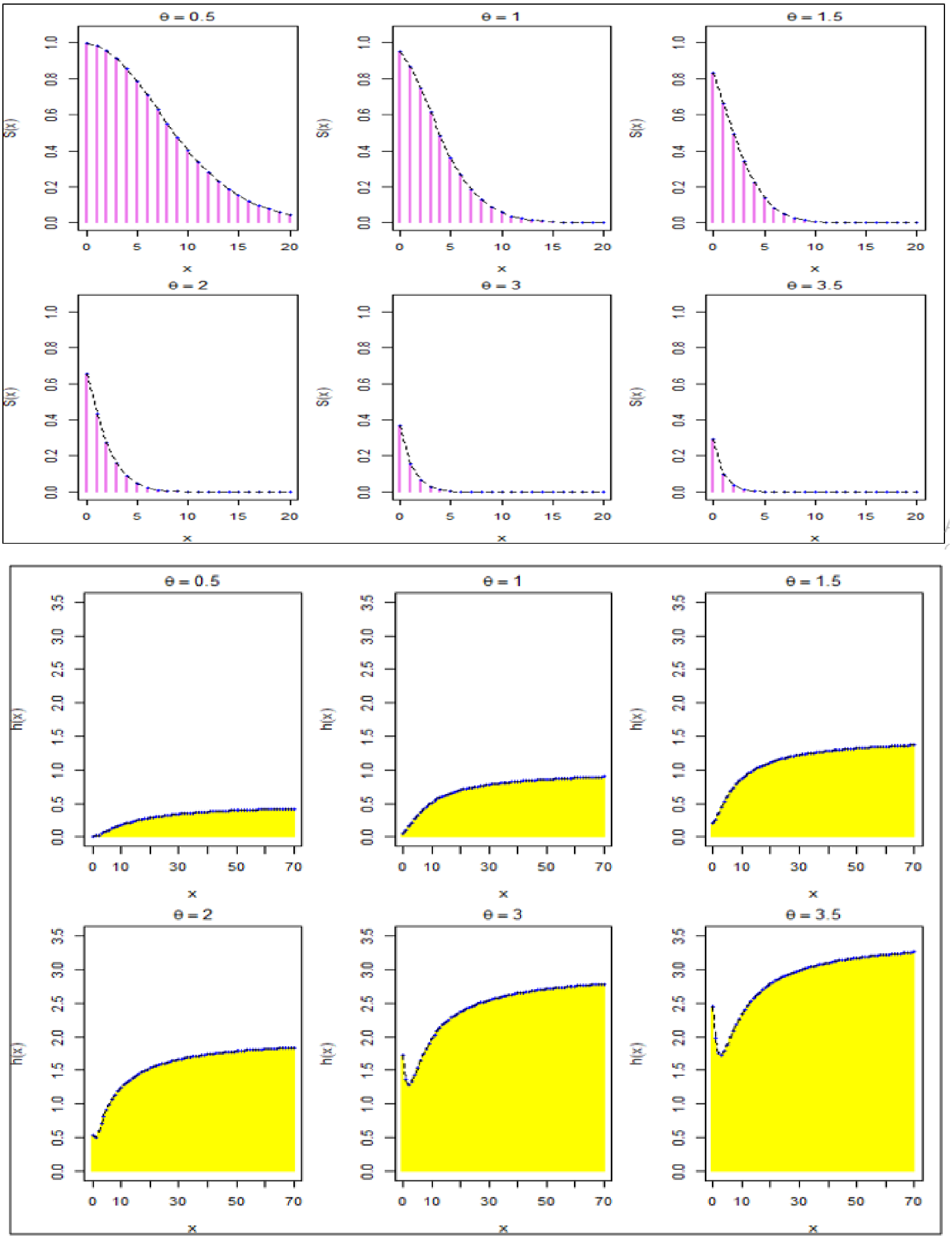


Figure 2 CDF, survival function and hazard function of PSD for varying values of parameter.

Citation: Shanker R, Tabassum R, Baishya J, et al. Statistical and reliability properties of poisson-suja distribution. *Biom Biostat Int J.* 2025;14(2):85‒94.
DOI: 10.15406/bbij.2025.14.00438

The reverse hazard rate function $Rh(x;\theta)$ and the second rate of failure $SRF(x;\theta)$ of the PSD can be obtained as

$$Rh(x;\theta)=\frac{P(x;\theta)}{F(x;\theta)}=\frac{\theta^5\left\{x^4+10x^3+35x^2+50x+24+(\theta+1)^4\right\}}{\left[(\theta^4+24)(\theta+1)^{x+5}-\left\{\begin{array}{l}\theta^4x^4+\left(14\theta^4+4\theta^3\right)x^3+\left(71\theta^4+48\theta^3+12\theta^2\right)x^2\\+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)x\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4(\theta+1)^4\end{array}\right\}\right]}$$

, and

$$SRF(x;\theta)=\ln\left[\frac{S(x;\theta)}{S(x+1;\theta)}\right]=\ln\left[\frac{(\theta+1)\left\{\begin{array}{l}\theta^4x^4+\left(14\theta^4+4\theta^3\right)x^3+\left(71\theta^4+48\theta^3+12\theta^2\right)x^2\\+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)x\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4(\theta+1)^4\end{array}\right\}}{\left\{\begin{array}{l}\theta^4(x+1)^4+\left(14\theta^4+4\theta^3\right)(x+1)^3+\left(71\theta^4+48\theta^3+12\theta^2\right)(x+1)^2\\+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)(x+1)\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4(\theta+1)^4\end{array}\right\}}\right]$$

The cumulative hazard function $H(x;\theta)$ and Mills ratio $M(x;\theta)$ of PSD are given by

$$H(x;\theta)=-\ln S(x;\theta)$$

$$-\ln\left(\frac{\begin{array}{l}\theta^4x^4+\left(14\theta^4+4\theta^3\right)x^3+\left(71\theta^4+48\theta^3+12\theta^2\right)x^2+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)x\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4(\theta+1)^4\end{array}}{(\theta^4+24)(\theta+1)^{x+5}}\right)$$

and

$$M(x;\theta)=\frac{S(x;\theta)}{P(x;\theta)}=\frac{\left\{\begin{array}{l}\theta^4x^4+\left(14\theta^4+4\theta^3\right)x^3+\left(71\theta^4+48\theta^3+12\theta^2\right)x^2\\+\left(154\theta^4+188\theta^3+108\theta^2+24\theta\right)x\\+\left(120\theta^4+240\theta^3+240\theta^2+120\theta+24\right)+\theta^4(\theta+1)^4\end{array}\right\}}{\theta^5\left\{x^4+10x^3+35x^2+50x+24+(\theta+1)^4\right\}}.$$

Estimation of parameter

The likelihood function L of the PSD is given by

$$L=\left(\frac{\theta^5}{\theta^4+24}\right)^n\frac{1}{(\theta+1)^{\sum_{x=1}^{k}(x+5)f_x}}\prod_{x=1}^{k}\left[x^4+10x^3+35x^2+50x+\left(\theta^4+4\theta^3+6\theta^2+4\theta+25\right)\right]^{f_x},$$

Where 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 the largest is observed value having non-zero frequency. The log likelihood function is thus obtained as

Citation: Shanker R, Tabassum R, Baishya J, et al. Statistical and reliability properties of poisson-suja distribution. *Biom Biostat Int J.* 2025;14(2):85‒94. DOI: 10.15406/bbij.2025.14.00438

$$\log L = n\log\left(\frac{\theta^5}{\theta^4+24}\right) - \sum_{x=1}^{k}(x+5)f_x\log(\theta+1) + \sum_{x=1}^{k} f_x\log\begin{bmatrix} x^4+10x^3+35x^2+50x \\ +\left(\theta^4+4\theta^3+6\theta^2+4\theta+25\right)\end{bmatrix}$$

The first derivative of the log likelihood function is given by

$$\frac{d\log L}{d\theta} = \frac{5n}{\theta} - \frac{4n\theta^3}{\theta^4+24} - \sum_{x=1}^{k}\frac{(x+5)f}{\theta+1} + \sum_{x=1}^{k}\frac{\left(4\theta^3+12\theta^2+12\theta+4\right)f_x}{\left[x^4+10x^3+35x^2+50x+\left(\theta^4+4\theta^3+6\theta^2+4\theta+25\right)\right]}$$

Where $\overline{x}$ is the sample mean?

The maximum likelihood estimate (MLE), $\hat{\theta}$ of θ PSD 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

$$\frac{5n}{\theta} - \frac{4n\theta^3}{\theta^4+24} - \sum_{x=1}^{k}\frac{(x+5)f}{\theta+1} + \sum_{x=1}^{k}\frac{\left(4\theta^3+12\theta^2+12\theta+4\right)f_x}{\left[x^4+10x^3+35x^2+50x+\left(\theta^4+4\theta^3+6\theta^2+4\theta+25\right)\right]} = 0$$

This non-linear equation can be solved by any numerical iteration methods such as Newton- Raphson, Bisection method, Regula –Falsi method etc.

Simulation study

To evaluate the performance of the maximum likelihood estimator (MLE) for the parameter θ of the PSD, a simulation study was carried out. Random samples were generated using the inverse transformation method, wherein the cumulative distribution function (CDF) was constructed numerically over a sufficiently large support, and each observation was drawn by matching a uniform random number with the CDF. The simulations were conducted for fixed parameter value θ=2 and θ=2.5, considering various sample sizes: n=20, 80, 100, 200, 300. For each scenario, 10,000 independent samples were generated, and the corresponding MLEs were obtained by numerically minimizing the negative log-likelihood using the L-BFGS-B optimization method.

For every sample size, the mean and variance of the estimated values were computed, along with the bias and mean squared error (MSE) with respect to the true parameter. The simulation results confirm that the MLE performs well; the bias is minimal and decreases as the sample size increases, while both the variance and MSE consistently decrease, and indicating improved estimation precision. Moreover, the MSE values are nearly identical to the variances, confirming that the estimator is nearly unbiased. These findings support the consistency and efficiency of the MLE under the PSD (Table1&2).

Table 1 Simulation Study for the parameter value θ =2

n	Mean	MSE	BIAS	Variance
20	2.0241	0.0581	0.0241	0.0575
80	2.0074	0.0132	0.0074	0.0132
100	2.0046	0.0106	0.0046	0.0105
200	2.0014	0.0053	0.0014	0.0052
300	2.0017	0.0034	0.0016	0.0034

Table 2 Simulation Study for the parameter value θ =2.5

n	Mean	MSE	BIAS	Variance
20	2.5702	0.1506	0.0702	0.1456
80	2.5166	0.0261	0.0166	0.0257
100	2.5118	0.0208	0.0117	0.2062
200	2.5053	0.01	0.0053	0.01
300	2.5037	0.0065	0.0037	0.0064

Applications

In this section two examples of observed count datasets has been taken to test the goodness of fit of Poisson-Suja distribution and its goodness of fit has been compared with the goodness of fit of Poisson distribution (PD), Poisson-Lindley distribution (PLD), Poisson-Komal distribution (PKD), Poisson-Shanker distribution (PShD) and Poisson-Garima distribution (PGD). The first dataset presented in Table 3 is related to the number of days that experienced the number of thunderstorms events at Cape Kennedy, Florida for the month of August, 11-year period of record January 1957 to December 1967 available in Falls et al.[23] The second dataset in Table 4 is related to the number of European red mites on apple leaves available in Bliss.[24] Based on the values of chi-square, PSD provides best fit for both the datasets among all considered one parameter over-dispersed count distributions.

Table 3 Number of thunderstorm events in the month of August

Number of observations	Observed Frequency	Expected Frequency					
		PD	PLD	PKD	PShD	PGD	PSD
0	185	151.8	184.8	187	186	186.2	194.1
1	89	122.9	87.2	85.3	86.1	86	75.9
2	30	49.7	39.3	38.4	38.8	38.7	36.4
3	24	13.4	17.1	17	17.1	17.1	18.4
4	10	2.7	7.3	7.5	7.4	7.4	9
5	3	0.5	5.3	5.8	5.6	5.6	7.2
Total	**341**	**341**	**341**	**341**	**341**	**341**	**341**
$\hat{\theta}(SE)$		0.8094	1.6934 (0.1132)	1.4944	1.5844	1.5800 (0.1141)	2.7087 (0.0886)
		-0.0487		-0.0977	-0.0951		
$-2\log L$ χ^2		888.617	847.6051	847.994	847.779	847.7928	849.5542
		49.49	7.0191	7.0868	7.0038	6.9731	6.1494
df		2	4	4	4	4	4
p-value		0	0.135	0.131	0.135	0.137	0.1883

Table 4 Number of European red mites available in bliss (1953)

Number of European red mites	Observed frequency	Expected Frequency					
		PD	PLD	PKD	PSHD	PGD	PSD
0	70	47.7	67.3	68.6	67.7	68.4	71.3
1	38	54.6	38.9	38	38.6	38.1	34.3
2	17	31.3	21.2	20.6	21	20.7	19.9
3	10	12	11.2	11	11.1	11	11.7
4	9	3.4	5.7	5.8	5.7	5.8	6.5
5	3	0.8	2.9	3	2.9	3	3.3
6	2	0.2	1.4	1.5	1.5	1.5	1.6
7	1	0	0.7	0.8	0.7	0.8	0.7
8	0	0	0.7	0.7	0.8	0.7	0.7
Total	**150**	**150**	**150**	**150**	**150**	**150**	**150**
$\hat{\theta}(SE)$		1.1467 (0.0874)	1.2601 (0.1139)	1.1289 (0.0983)	1.2247 (0.0970)	1.1491 (0.1142)	2.3741 (0.1041)
$-2\log L$ χ^2		485.6199	445.0218	444.7778	444.9169	444.7943	443.985
		23.295	3.0161	2.5141	2.8706	2.5555	2.0683
df		3	4	4	4	4	4
p-value		0.0001	0.6975	0.7744	0.7199	0.7681	0.8396

Conclusion

The Poisson-Suja distribution (PSD) is log-concave and is a two-component mixture of negative binomial distributions. The cumulative distribution function, survival function and hazard function of the PSD have been derived and their behaviors' for different values of parameter have been discussed. The expressions reverse hazard rate function, the second rate of failure; the cumulative hazard function and the Mills ratio of the PSD have also been derived. The consistency of maximum likelihood estimator of the parameter of the PSD has been shown through simulation study. The PSD has been applied to two real datasets and found to provide quite satisfactory fit over other competing one parameter over-dispersed count distributions namely Poisson-Lindley distribution, Poisson-Komal distribution, Poisson-Shanker distribution, and Poisson-Garima distribution. Therefore, for count data, PSD can be considered as an important discrete distribution.

Acknowledgments

None.

Conflicts of interest

None.

References

1. Sankaran M. The discrete Poisson-Lindley distribution. *Biometrics*. 1970;26(1):145–149.
2. Lindley DV. Fiducial distributions and Bayes theorem. *J R Stat Soc.* 1958;20(1):102–107.
3. Shanker R. The discrete Poisson-Shanker distribution. *Jacobs J Biostatistics*. 2016;1(1):1–7.
4. Shanker R. Shanker distribution and its applications. *Int J Stat Appl.* 2015;5(6):338–348.
5. Shanker R. The discrete Poisson-Garima distribution. *Biom Biostat Int J*. 2017;5(2):48–53.
6. Shanker R. Garima distribution and its application to model behavioral science data. *Biom Biostat Int J*. 2016;4(7):275–281.
7. Shanker R, Tabassum R, Sarma I, et al. The Poisson compound of Komal distribution with statistical properties and applications. *Biom Biostat Int J*. 2025;14(1):35–43.
8. Shanker R. Komal distribution with properties and application in survival analysis. *Biom Biostat Int J.* 2023;12(2):40–44.
9. Shanker R, Baishya J, Tabassum R. Statistical properties and applications of Poisson-Garima distribution. *Biom Biostat Int J*. 2025;14(2):45–51.
10. Shanker R. Suja distribution and its application. *Int J Probab Stat.* 2017;6(2):11–19.
11. Al-Omari AI, Alsmairam IK. Length-biased Suja distribution and its application. *J Appl Probab Stat*. 2019;14(3):95–116.
12. Al-Omari AI, Alhyasat IK, Abu BMA. Power length-biased Suja distribution: properties and application. *Electron J Appl Stat Anal*. 2019;12(2):429–452.
13. Alsmairan IK, Al-Omari AI. Weighted Suja distribution with application to ball bearings data. *Life Cycle Reliab Saf Eng*. 2020;9:195–211.
14. Todorka T, Anton I, Asen R, et al. Comments on some modification of Suja cumulative functions with applications to the theory of computer viruses propagation. *Int J Differ Equ Appl*. 2020;19(1):83–95.
15. Shanker R, Upadhaya R, Shukla KK. A quasi Suja distribution. *Reliab Theory Appl.* 2022;17(3):162–178.
16. Enonwe SU, Okereke EW, Ibeh GC. A bimodal extension of Suja distribution with applications. *Stat Appl.* 2023;21(2):155–173.
17. Shanker R, Das R, Shukla KK. An extended Suja distribution with properties and applications. *Biom Biostat Int J*. 2024;13(1):16–21.
18. Enogwe SU, Christiana UA, Chinonye EO, et al. A generalized Suja distribution with applications to lifetime data. *Asian J Probab Stat*. 2024;26(10):127–140.
19. Abu-Zinadah H, Alsimairi T. The estimations of parameter of Suja distribution with application. *Alexandria Eng J*. 2024;86:327–334.
20. Shanker R, Saharia J, Shukla KK. The Poisson-Suja distribution with properties and applications in biological count data sets. *Reliab Theory Appl*. 2025;20(1):1009–1019.
21. Grandell J. Mixed Poisson Processes. Chapman & Hall; 1997.
22. Tajuddin RRM, Ismail N, Ibrahim K. Several two-component mixture distributions for count data. *Commun Stat Simul Comput*. 2022;51:3760–3771.
23. Falls LW, Wilford WO, Carter MC. Probability distribution for thunderstorm activity at Cape Kennedy, Florida. *J Appl Meteorol*. 1971;10(1):97–104.
24. Bliss CI. Fitting negative binomial distribution to biological data. *Biometrics*. 1953;9:177–200.