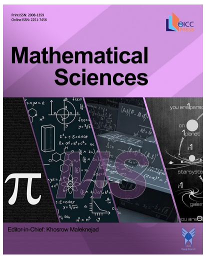

Math. Sci. 2025; 19(1) : 1-14

Mathematical Sciences (MS)

https://doi.org/10.57647/mathsci.2025.1901.02

Sushila-Poisson Distribution: A Flexible Tool for Survival and Reliability Modeling

Sepideh Daghagh[1], Anis Iranmanesh[1,*], Najmeh Nakhaei Rad[2], Ehsan Ormoz[1]

[1]*Department of Mathematics and Computer Sciences, Ma.C., Islamic Azad University, Mashhad, Iran*
[2]*Department of Statistics, University of Pretoria, Pretoria, South Africa*

*Corresponding author: anis.iranmanesh@iau.ac.ir

Original Research

Received:
06 January 2025
Revised:
14 February 2025
Accepted:
28 February 2025
Published in Issue:
31 March 2025

Abstract:

This paper presents a new type of Sushila distribution that provides greater flexibility for modelling lifetime data. This model, called the Sushila-Poisson (SP) distribution, is created by combining the Sushila and Poisson distributions. The three-parameter SP distribution represents various shapes of hazard rate functions, including upside-down bathtub, bathtub-shaped, increasing, and decreasing hazard rates, which are commonly encountered in fields such as medicine, engineering, economics, and the natural sciences. Therefore, the proposed model offers great potential for applications in these areas. The new model includes several known distributions, such as the Lindley, Lindley-Poisson, and Sushila distributions, as special cases. Several statistical properties of the SP distribution have been derived in this study. Simulation studies were conducted to examine the performance of the maximum likelihood estimators, which were developed using the Expectation-Maximization (EM) algorithm. The flexibility of the new model was further demonstrated through its application to three real data sets.

Keywords: EM algorithm; Maximum likelihood estimation; Poisson distribution; Sushila distribution

Cite this article: Daghagh S., Iranmanesh A., Nakhaei Rad N., Ormoz E. Sushila-Poisson Distribution: A Flexible Tool for Survival and Reliability Modeling. Math. Sci 2025; *19*(1) : 1-14 https://doi.org/10.57647/mathsci.2025.1901.02

1. Introduction

In reliability and survival analysis, understanding how the risk of failure or event occurrence changes over time is often more important than simply estimating average lifetimes. This is typically done through the hazard function, which describes the instantaneous risk at any given time. While standard distributions such as the exponential and Weibull have been widely used for this purpose, they impose restrictive assumptions, for example, a constant or monotonic hazard. Yet, in many practical situations, real data suggest more complicated hazard shapes: for instance, a bathtub curve in mechanical systems, or a hazard that first increases and then declines in biological contexts. Relying solely on classical models in such cases can lead to poor fit and unreliable conclusions. As a result, there is a growing need for new distributions that allow for greater flexibility in hazard shapes. Having such models makes it possible to better represent the underlying dynamics of time-to-event data, especially when dealing with heterogeneous populations or complex failure mechanisms. This, in turn, leads to more accurate inference and better-informed decision-making in fields where lifetime modeling plays a critical role.

To address the need for more flexible models capable of capturing diverse hazard function shapes, this paper introduces a new variant of the Sushila distribution. The Sushila distribution was introduced by [1] with the density function given by

$$f(y,\theta,\alpha)=\frac{\theta^2}{\alpha(\theta+1)}\left(1+\frac{y}{\alpha}\right)e^{-\frac{\theta}{\alpha}y};\, y>0,\theta,\alpha>0 \quad (1)$$

and the distribution function

$$F(y,\theta,\alpha)=1-\left(1+\frac{\theta y}{\alpha(\theta+1)}\right)e^{-\frac{\theta}{\alpha}y}. \qquad y>0,\ \alpha>0,\ \theta>0 \quad (2)$$

Several researchers have proposed modified versions of the Sushila distribution with the primary aim of embedding existing distributions into more flexible and adaptable frameworks. Borah and Saikia [2] introduced the discrete Sushila distribution. This new proposed model seemed to be suitable for modeling and thus provided a better alternative to the discrete Poisson-Lindley for modeling different types of count data. After that, the zero-truncated discrete Shanker distribution was studied more by the same authors [3]. Yamrubboon et al. [4] introduced a negative binomial-Sushila distribution, which was a new mixed negative binomial distribution. It was shown that the new distribution provided a better fit of the data than the Poisson and negative binomial distributions. Elgarhy and Shawki [5] proposed a new three-parameter exponentiated Sushila (ES) distribution and studied its different properties. It was observed that the proposed ES distribution had several desirable properties and included some distributions as special cases. Shawki and Elgarhy [6] introduced a new four-parameter Kumaraswamy Sushila (KwS) distribution and studied its different properties. It was observed that the proposed KwS distribution provided several desirable properties. Rather and Subramanian [7] discussed the statistical properties and applications of length-biased Sushila distribution, which was a special case of weighted distributions. Recently, Borah and Hazarika [8] introduced Poisson-Sushila distribution and its applications. Rather and Subramanian [9] introduced a new generalization of the Sushila distribution, namely the weighted-Sushila distribution with three shape parameters. The subject distribution was generated by using the weighted technique. Some mathematical properties, along with reliability measures, were discussed. Finally, the results were compared over Sushila distribution, and it was found that the weighted Sushila distribution provides a better fit than the Sushila distribution. Pudprommarat [10] introduced the hurdle Poisson-Sushila distribution as an extension to the Poisson-Sushila distribution and indicated this distribution provided a better fit than Poisson. Richardo de Oliveria et al.[11] used a discrete sushila distribution in the analysis of right-censored lifetime data. This new model, in the same way as the continuous Sushila distribution, has the discrete Lindley distribution as a special case. The new proposed model performed at least as well as its special case and some other traditional discrete models, such as the Poisson and geometric distributions. Pudprommarat [12] presented a zero-one inflated Poisson-Sushila distribution, which was a discrete distribution. An application of the distribution to a real data set was presented and compared with the fit attained by some other well-known distributions for count data. Adetunji [13] introduced a three-parameter generalization of the Sushila distribution using the Quadratic transmuted technique by [14] and compared the performance of the new distribution with the Sushila distribution. Boonthiem et al. [15] presented a two-parameter continuous distribution called the new Sushila distribution, and some properties of the distribution were discussed. The new distribution contains the Sushila distribution as a particular case $p = \frac{1}{2}(\theta = 1)$. Furthermore, the new Sushila distribution was compared with the Lindley and Sushila distributions for the waiting time data. The results showed that the new Sushila provided a better fit than both the Lindley and Suhila distributions. Ricardo Puziol de Oliveira et al. [16] introduced a new bivariate distribution based on the Sushila distribution to model bivariate lifetime data in the presence of a cure fraction, right-censored data, and covariates. The new bivariate probability distribution was obtained by a methodology used in the reliability theory based on fatal shocks. For more information see Aryuyuen et al. [17], Yamrubboon et al. [18], Bodhisuwan et al. [19] and Atikankul et al. [20]. Daghagh et al. [21] introduced a three-parameter Sushila-Geometric distribution by compounding Sushila and Geometric distributions, which offered a flexible model for lifetime data.

This paper aims to introduce an extension of the Sushila distribution, which offers a more flexible distribution for modeling lifetime data, namely in reliability and survival analysis, in terms of its failure rate or hazard rate shapes. The new distribution can accommodate both decreasing and increasing failure rates as well as unimodal and bathtub-shaped failure rates.

The rest of this paper is organized as follows: in Section 2 we introduce the new Sushila-Poisson distribution and investigate its basic properties, including the shape properties of its density function and the hazard rate function, stochastic orderings, quantile function, moments, order statistics, residual lifetime, reversed residual lifetime, Bonferroni and Lorenz curves, entropies, and mean deviations. The maximum likelihood inference using the EM algorithm, asymptotic properties of the estimates, and Fisher information matrix are discussed in Section 3. Simulation studies are also conducted in Section 4. Section 5 gives the application (for three real datasets) and reports the results. Our work is concluded in Section 6.

2. Sushila-Poisson distribution

To construct the new distribution, follow the steps below. Assume that the failure of a device is caused by the presence of Z (an unknown number) of initial defects. Let $Y_1, \cdots, Y_Z$ represent the failure times of these initial defects. Then, the failure time of the device is given by, $X = Min(Y_1, \cdots, Y_Z)$. Suppose the failure times of the initial defects $Y_1, \cdots, Y_Z$ follow a Sushila distribution (1) and Z has a zero-truncated Poisson distribution with probability mass function as follows:

$$P(Z = z) = \frac{\lambda^z e^{-\lambda}}{z!(1 - e^{-\lambda})}, \quad z = 1, 2, 3, \cdots . \tag{3}$$

By assuming that the random variables Y_i and Z are independent, then the conditional CDF $X|Z = z$ is given by

$$F_{X|Z}(x|z) = 1 - \left[(1 + \frac{\theta x}{\alpha(\theta + 1)})e^{-\frac{\theta}{\alpha}x}\right]^z .$$

The Sushila-Poisson (SP) distribution with three-parameter, denoted by $SP(\theta, \alpha, \lambda)$, is defined by the marginal CDF of X;

$$F(x) = \frac{e^{\lambda} - e^{\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}}}{e^{\lambda}-1}, \quad x>0. \tag{4}$$

The PDF of SP distribution for $\theta, \lambda, \alpha > 0$ is given by

$$f(x)=\frac{\lambda\theta^2(1+\frac{x}{\alpha})}{\alpha(\theta+1)(e^{\lambda}-1)}e^{\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}-\frac{\theta}{\alpha}x}, x>0. \tag{5}$$

The SP distribution includes several submodels. If $\alpha = 1$, it becomes the Lindley-Poisson (LP) distribution introduced [22]. When λ goes to zero it changes to the Sushila distribution due to Shanker et al. [1].

Theorem 2.1 *Considering the SP distribution with the probability density function in (5), we have the following property:*
If $\theta^2(\lambda+1) \geq 1$, f(x) is decreasing in x. If $\theta^2(\lambda+1) < 1$, f(x) is a unimodal function at x_0, where x_0 is the solution of the equation $\alpha(\theta+1)e^{\frac{\theta}{\alpha}x}(\theta x+\theta\alpha-\alpha)+\lambda\theta^2(\alpha+x)^2 = 0$.

Proof. $f(0) = \frac{\lambda\theta^2 e^{\lambda}}{\alpha(\theta+1)(e^{\lambda}-1)}$ and $f(\infty) = 0$. The first derivative of $\log f(x)$ is

$$\frac{d\log f(x)}{dx} = -\frac{[\alpha(\theta+1)e^{\frac{\theta}{\alpha}x}(\theta x+\theta\alpha-\alpha)+\lambda\theta^2(\alpha+x)^2]}{\alpha^2(\theta+1)(\alpha+x)}e^{-\frac{\theta}{\alpha}x}. \tag{6}$$

let

$$s(x) = \alpha(\theta+1)e^{\frac{\theta}{\alpha}x}(\theta x+\theta\alpha-\alpha)+\lambda\theta^2(\alpha+x)^2, \tag{7}$$

then $s(0) = \alpha^2\theta^2(\lambda+1) - \alpha^2$ and $s(\infty) = \infty$. $s'(x) = \theta^2(\alpha+x)[(\theta+1)e^{\frac{\theta}{\alpha}x} + 2\lambda] > 0$,
If $\alpha^2\theta^2(\lambda+1) \geq \alpha^2$ or $\theta^2(\lambda+1) \geq 1$ then $s(x) \geq 0$, $\frac{dlogf(x)}{dx} \leq 0$, i.e., $f(x)$ is decreasing in x.
If $\alpha^2\theta^2(\lambda+1) < \alpha^2$ or $\theta^2(\lambda+1) < 1$, $f(x)$ is a unimodal function at x_0, where x_0 is the solution of the equation $s(x) = 0$. Figure Figure 1 shows different shapes of (5) for selected values of parameters.

2.1 Survival and failure rate function

The survival function for (2) is provided as follows

$$\bar{F}(x) = \frac{e^{\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}}-1}{e^{\lambda}-1}. \tag{8}$$

From (1) and (8) the failure(or hazard)rate function of the SP distribution,

$$h(x) = \frac{\lambda\theta^2(1+\frac{x}{\alpha})e^{\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}-\frac{\theta}{\alpha}x}}{\alpha(\theta+1)\left[e^{\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}}-1\right]}, \tag{9}$$

where $x > 0, \theta, \lambda, \alpha > 0$. Figure 2 illustrates some shapes of (9) with some different values of α, λ, θ.

2.2 Stochastic orderings

For positive continuous random variables, stochastic ordering serves as an important tool for comparing their behavior. A random variable X is said to be stochastically smaller than a random variable Y in the

- stochastic order ($X \leq_{st} Y$) if $F_X(x) \geq F_Y(x)$ for all x.
- hazard rate order ($X \leq_{hr} Y$) if $h_X(x) \geq h_Y(x)$ for all x.
- mean residual life order ($X \leq_{mrl} Y$) if $m_X(x) \leq m_Y(x)$ for all x.
- likelihood ratio order ($X \leq_{lr} Y$) if $\frac{f_X(x)}{f_Y(x)}$ decreases in x.

From Shaked and Shanthikumar [23] we have the following implications:

$$\begin{array}{c} X \leq_{\text{lr}} Y \Rightarrow X \leq_{\text{hr}} Y \Rightarrow X \leq_{\text{mrl}} Y \\ \Downarrow \\ X \leq_{\text{st}} Y \end{array} \tag{10}$$

The following theorem demonstrates that the SP distribution follows a specific ordering with respect to the strongest “likelihood ratio” ordering.

Theorem 2.2 *Let $X \sim SP(\theta,\alpha,\lambda_1)$ and $Y \sim SP(\theta,\alpha,\lambda_2)$. If $\lambda_1 > \lambda_2$ then $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.*

Proof: First note that

$$\frac{f_X(x)}{f_Y(x)} = \frac{\lambda_1(e^{\lambda_2}-1)\exp\left(\frac{\lambda_1 e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)-\lambda_2 e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}\right)}{\lambda_2(e^{\lambda_1}-1)}.$$

Now

$$\log\left(\frac{f_X(x)}{f_Y(x)}\right) = \log(e^{\lambda_2}-1) + \log(\lambda_1) + \left(\frac{\lambda_1 e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)-\lambda_2 e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}\right).$$

Then

$$\frac{d}{dx}\log\left(\frac{f_X(x)}{f_Y(x)}\right) = \frac{\theta^2(\lambda_2-\lambda_1)}{\alpha(\theta+1)}\left(e^{-\frac{\theta}{\alpha}x}(1+\frac{x}{\alpha})\right),$$

where $\frac{f_X(x)}{f_Y(x)}$ is decreasing in x. That is $X \leq_{lr} Y$. The remaining statements follow from the implications in (10).

2.3 Quantile Function

Let X be a SP random variable with the CDF in (4). By inverting $F(x) = u$ for $0 < u < 1$, we obtain

$$e^{\lambda} - u(e^{\lambda}-1) = e^{\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}}$$

$$\frac{\alpha((\theta+1)\log(e^{\lambda}-ue^{\lambda}+u)}{\lambda} = e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x))$$

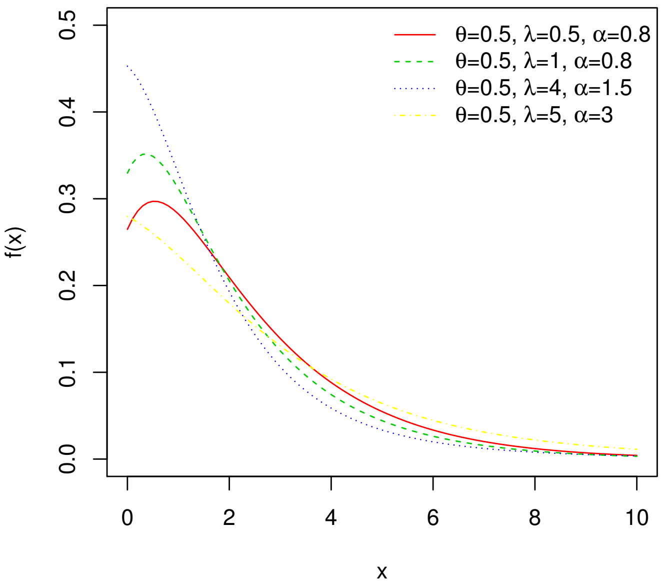

Figure 1. Plots of PDF of SP for selected values of parameters.

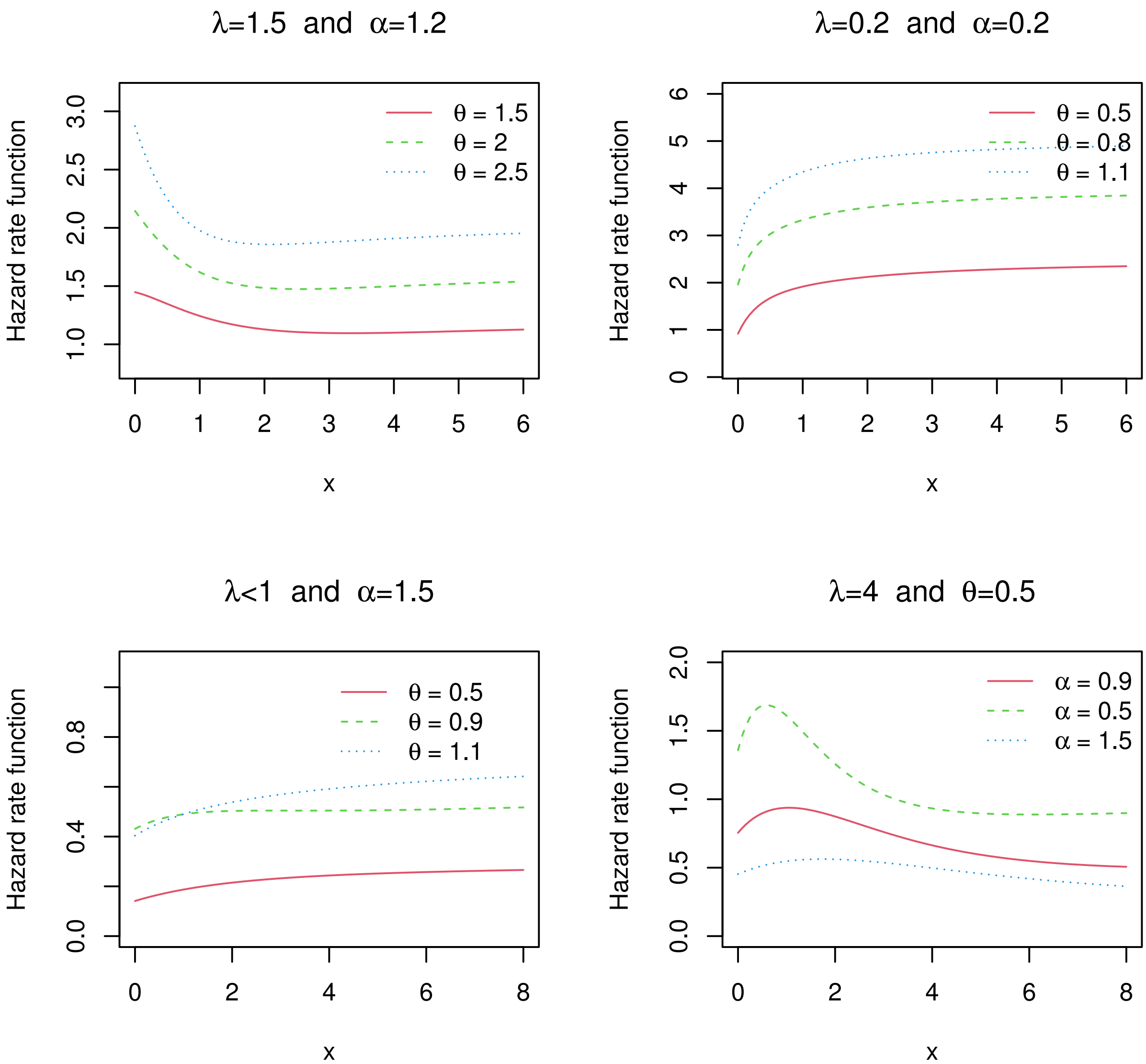

Figure 2. Plots of the hazard rate function for selected values of parameters

https://doi.org/10.57647/mathsci.2025.1901.02

Multiplying by $-\exp(-\theta-1)$ both sides of above Equation, we have

$$\frac{-(\theta+1)\log(e^{\lambda}-ue^{\lambda}+u)}{\lambda\exp(\theta+1)}=-e^{-(\frac{\theta}{\alpha}x+\theta+1)}(\theta+1+\frac{\theta}{\alpha}x)$$

From $W(z)\exp(W(z))=z$, (see Adler [24] in detail), we notice that $-(\theta+1+\frac{\theta}{\alpha}x)$ is the Lambert W function of the real argument $-\frac{1}{\lambda}(\theta+1)\log(e^{\lambda}-ue^{\lambda}+u)\exp(-\theta-1)$. Then, we have

$$W\left(\frac{-(\theta+1)\log(e^{\lambda}-ue^{\lambda}+u)}{\lambda\exp(\theta+1)}\right) = -\theta-1-\frac{\theta}{\alpha}x$$

Then the quantile function of the SP distribution is given by

$$F^{-1}(u)=\frac{-\theta-W_{-1}\left(\frac{-(\theta+1)\log(e^{\lambda}-ue^{\lambda}+u)}{\lambda\exp(\theta+1)}\right)-1}{\frac{\theta}{\alpha}}. \quad (11)$$

Note that $-\frac{1}{e}<\left(\frac{-(\theta+1)\log(e^{\lambda}-ue^{\lambda}+u)}{\lambda\exp(\theta+1)}\right)<0$ so the W_{-1} is unique, which implies that $F^{-1}(u)$ is also unique. Thus, one can use (11) for generating random data from the SP distribution. In addition, the qth quantile x_q of $SP(\theta,\alpha,\lambda)$ is given by

$$x_q = -\alpha-\frac{\alpha}{\theta}W_{-1}\left(\frac{-(\theta+1)\log(e^{\lambda}-qe^{\lambda}+q)}{\lambda\exp(\theta+1)}\right)-\frac{\alpha}{\theta}; \quad 0<q<1. \quad (12)$$

In particular, we obtain the median by putting $q=0.5$ in (12).

2.4 Moments

Many important characteristics and features of a distribution can be effectively examined through its moments. Moments are fundamental statistical measures that capture key quantitative aspects of a distribution. Let X be a random variable following the SP distribution with parameters $\Theta=(\theta,\lambda,\alpha)'$. Expressions for mathematical expectation and the kth moment on the origin of X can be obtained using the well-known formula.

$$\mu_k=E(X^k)=\int_0^{\infty}kx^{k-1}\bar{F}(x)dx.$$

Hence, it follows that

$$\begin{aligned}
\mu_k &= \int_0^{\infty}kx^{k-1}\left[\frac{e^{\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}}-1}{e^{\lambda}-1}\right]dx \\
&= \frac{1}{e^{\lambda}-1}\int_0^{\infty}kx^{k-1}\sum_{j=1}^{\infty}\left(\frac{(\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)})^j}{j!}\right)dx \\
&= \frac{1}{e^{\lambda}-1}\sum_{j=1}^{\infty}\frac{\lambda^j}{j!(\alpha(\theta+1))^j}\int_0^{\infty}kx^{k-1}e^{-\frac{\theta j}{\alpha}x}(\alpha(\theta+1)+\theta x)^j dx \\
&= \frac{1}{e^{\lambda}-1}\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\lambda^j\theta^i}{j!(\alpha(\theta+1))^i}\int_0^{\infty}x^{i+k-1}e^{-\frac{\theta j}{\alpha}x}dx \\
&= \frac{1}{e^{\lambda}-1}\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\lambda^j\theta^i}{j!(\alpha(\theta+1))^i}\times\frac{\Gamma(i+k)}{(\frac{\theta j}{\alpha})^{i+k}}.
\end{aligned}$$

In particular

$$E(X)=\frac{1}{e^{\lambda}-1}\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\alpha\lambda^j\Gamma(i+1)}{j!\theta(\theta+1)^i(j)^{i+1}}, \quad (13)$$

and

$$E(X^2)=\frac{1}{e^{\lambda}-1}\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\alpha^2\lambda^j\Gamma(i+2)}{j!\theta^2(\theta+1)^i(j)^{i+2}}.$$

Moreover, when λ goes to zero, we get $\mu=\frac{\alpha(\theta+2)}{\theta(\theta+1)}$, which is the mean of the original Sushila distribution. Also, one can calculate the $Var(X)$.

2.5 Order statistics

Order statistics investigate the properties and behavior of ordered random variables and their functions. These concepts are particularly relevant in the analysis of natural phenomena such as floods, lifespan, breaking strength, and atmospheric conditions like temperature, pressure, and wind. Understanding and predicting the recurrence of extreme events is critically important, which motivates a focused study on extreme observations. In this section, we get the probability density function and the cumulative distribution function of the kth order statistic of the SP distribution. Suppose $X_1, X_2, \cdots, X_n$ is a random sample from (5). Let $X_{1:n}<X_{2:n}<\cdots<X_{n:n}$ denote the corresponding order statistics. The PDF and CDF of the kth order statistic, say $Y=X_{k:n}$, are given by

$$f_Y(y)=\frac{\lambda\theta^2(1+\frac{y}{\alpha})}{\alpha(\theta+1)(e^{\lambda}-1)}e^{\frac{\lambda e^{-\frac{\theta}{\alpha}y}(\alpha(\theta+1)+\theta y)}{\alpha(\theta+1)}} -\frac{\theta}{\alpha}y\sum_{l=0}^{n-k}\binom{n-k}{l}(-1)^l\left[\frac{e^{\lambda}-e^{\frac{\lambda e^{-\frac{\theta}{\alpha}y}(\alpha(\theta+1)+\theta y)}{\alpha(\theta+1)}}}{e^{\lambda}-1}\right]^{k+l-1}$$

and

$$\begin{aligned}
F_Y(y) &= \sum_{j=k}^{n}\binom{n}{j}F^j(y)\,[1-F(y)]^{n-j} \\
&= \sum_{j=k}^{n}\sum_{l=0}^{n-j}\binom{n}{j}\binom{n-j}{l}(-1)^l\left[\frac{e^{\lambda}-e^{\frac{\lambda e^{-\frac{\theta}{\alpha}y}(\alpha(\theta+1)+\theta y)}{\alpha(\theta+1)}}}{e^{\lambda}-1}\right]^{j+l}
\end{aligned}$$

If $X_1,\cdots,X_n$ is a random sample from (5) and $\bar{X}=\dfrac{X_1+X_2+\cdots+X_n}{n}$ denotes the sample mean, by the central limit theorem as $n\to\infty$ then $\dfrac{\sqrt{n}(\bar{X}-E(X))}{\sqrt{Var(X)}}$ approaches the standard normal distribution. Sometimes one would be interested in the asymptotics of the extreme values $M_n=\max(X_1,\cdots,X_n)$ and $m_n=\min(X_1,\cdots,X_n)$.

For the CDF in (4), by using L'Hospital's rule, we have

$$\lim_{t\to\infty}\frac{1-F(t+x)}{1-F(t)}=\exp(-\frac{\theta}{\alpha}x)$$

In addition, by using L'Hospital's rule, it can be easily shown that

$$\lim_{t\to 0}\frac{F(tx)}{F(t)}=\lim_{t\to 0}\frac{xf(tx)}{f(t)}=x$$

By following Theorem 1.6.2 in Leadbetter et al. [25], we observe that there must be some normalizing constants $a_n>0, b_n$, $c_n>0$ and d_n, such that

$$\begin{aligned}Pr[a_n(M_n-b_n)\le x] &\rightarrow \exp(-\exp(-\frac{\theta}{\alpha}x))\\ Pr[c_n(m_n-d_n)\le x] &\rightarrow 1-\exp(-x)\end{aligned}$$

as $n\to\infty$. The form of the normalizing constants can be determined by using Corollary 1.6.3 in Leadbetter et al. [25]. As an illustration, one can see that $a_n=1$ and $b_n=F^{-1}(1-\frac{1}{n})$ where $F^{-1}(\cdot)$ denotes the inverse function of $F(\cdot)$.

2.6 Residual life time and reversed residual life time

Given that a system component has survived up to a time $t>0$, the residual life refers to the remaining time until failure, and is represented by the conditional random variable $X-t|X>t$. The mean residual life (MRL) is a key concept in survival analysis and reliability theory, as it helps characterize the lifetime of a system and can uniquely determine the corresponding lifetime distribution. The rth moment of the residual life for the *SP* distribution can be derived using a general formula.

$$\begin{aligned}\mu_r(t) &= E[(X-t)^r|X>t]\\ &= \frac{1}{\bar{F}(t)}\int_t^\infty r(x-t)^{r-1}\bar{F}(x)dx,\end{aligned}$$

Using the binominal expansion to $(x-t)^r$ and substituting $\bar{F}(x)$ given by (8) into the formula above imply

$$\mu_r(t)=\left[\frac{e^\lambda-1}{e^{\frac{\lambda e^{-\frac{\theta}{\alpha}t}(\alpha(\theta+1)+\theta t)}{\alpha(\theta+1)}}-1}\right]\times \sum_{j=1}^{\infty}\sum_{i=0}^{j}\sum_{k=0}^{r-1}\frac{\binom{j}{i}\binom{r-1}{k}\lambda^j\theta^i(-1)^k t^k r}{j!(\alpha(\theta+1))^i}(\frac{\alpha}{\theta j})^{i+r-k}\Gamma(r+i-k,\frac{\theta jt}{\alpha}).$$

Thus, we obtain the mean residual life (MRL) of the SP distribution as

$$\mu_1(t)=\mu(t)=\left[\frac{e^\lambda-1}{e^{\frac{\lambda e^{-\frac{\theta}{\alpha}t}(\alpha(\theta+1)+\theta t)}{\alpha(\theta+1)}}-1}\right]\times \sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\alpha\lambda^j\theta^i}{j!(\theta+1)^i(\theta j)^{i+1}}\Gamma(1+i;\frac{\theta jt}{\alpha}).$$

In particular, we obtain

$$\mu(0)=E(X)=\frac{1}{e^\lambda-1}\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\alpha\lambda^j\Gamma(i+1)}{j!\theta(\theta+1)^i(j)^{i+1}}.$$

Also, if $\lambda=0$, then

$$\mu(t)=\frac{\alpha}{\theta}\left[\frac{\alpha(\theta+1)+\theta t+\alpha}{\alpha(\theta+1)+\theta t}\right],$$

which is the MRL function of the original Sushila distribution. The variance of the residual life of the SP distribution can be obtained easily by using $\mu_2(t)$ and $\mu(t)$, and consequently its coefficient of variation.

The rth-order moment of the reversed residual life can be obtained by the well-known formula

$$\begin{aligned}m_r(t) &= E[(t-X)^r|X<t]\\ &= \frac{1}{F(t)}\int_0^t r(t-x)^{r-1}F(x)dx,\end{aligned}$$

$$m_r(t)=\frac{r}{e^\lambda-e^{\frac{\lambda e^{-\frac{\theta}{\alpha}t}(\alpha(\theta+1)+\theta t)}{\alpha(\theta+1)}}}\sum_{k=0}^{r-1}\binom{r-1}{k}t^{r-k-1}(-1)^k \times\left[\frac{t^{k+1}}{k+1}+\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\lambda^j\alpha^{k+1}}{j!(\theta+1)^i\theta^{k+1}j^{i+k+1}}\gamma(i+k+1,\frac{\theta jt}{\alpha})\right].$$

Thus, the mean of the reversed residual life of the SP distribution is given by

$$m_1(t)=\frac{1}{e^\lambda-e^{\frac{\lambda e^{-\frac{\theta}{\alpha}t}(\alpha(\theta+1)+\theta t)}{\alpha(\theta+1)}}}\times \left[t+\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\lambda^j\alpha}{j!\theta(\theta+1)^i(j)^{i+1}}\gamma(i+1,\frac{\theta jt}{\alpha})\right].$$

$$m_2(t)=\frac{2t}{e^\lambda-e^{\frac{\lambda e^{-\frac{\theta}{\alpha}t}(\alpha(\theta+1)+\theta t)}{\alpha(\theta+1)}}}\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\lambda^j\alpha}{j!\theta(\theta+1)^i(j)^{i+1}} \times\left[\gamma(i+1,\frac{\theta jt}{\alpha})-\frac{\alpha}{j\theta^2}\gamma(i+2,\frac{\theta jt}{\alpha})\right].$$

2.7 Bonferroni and Lorenz Curves

Bonferroni and Lorenz curves have a wide range of practical applications extending beyond economics and poverty analysis. They are also utilized in fields such as reliability engineering, lifetime data analysis, insurance, and medical research. For a random variable X with CDF (4), the Bonferroni curve is defined by

$$B_F[F(x)] = \frac{1}{\mu F(x)}\int_0^x uf(u)du,$$

Or equivalently, given by

$$B_F(p) = = \frac{1}{\mu p}\int_0^p F^{-1}(t)dt,$$

where $p=F(x)$ and $F^{-1}(t)=inf[x;F(x)>t]$. From the relationship between the Bonferroni curve and the mean residual lifetime given by Theorem 2.1 [26], the Bonferroni curve of the distribution function F of SP distribution is given by

$$B_F[F(x)]=\frac{e^\lambda-1}{e^\lambda-e^{\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}}} \times\left[1-\frac{1}{\mu(e^\lambda-1)}\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\alpha\lambda^j\Gamma(i+1;\frac{\theta jx}{\alpha})}{j!\theta(\theta+1)^i j^{i+1}} -\frac{x}{\mu}\left(\frac{e^{\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}}-1}{e^\lambda-1}\right)\right], \quad (14)$$

where μ is given in (13). Additionally, the Lorenz curve of F that follows the SP distribution can be obtained via the expression$L_F[F(x)] = B_F[F(x)]F(x)$. The scaled total time and cumulative total time on test transform of a distribution function F ([26]) are defined by

$$S_F[F(t)] = \frac{1}{\mu}\int_0^t \bar{F}(u)du$$
$$C_F = \int_0^1 S_F[F(t)]f(t)dt,$$

where $S(\cdot)$ and $F(\cdot)$ denotes the survival function and CDF of X. Then, for SG distribution we get

$$S_F[F(t)] = \frac{1}{\mu(e^\lambda-1)}\sum_{j=1}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\alpha\lambda^j}{j!\theta(\theta+1)^i j^{i+1}} \times\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\frac{(\frac{\theta jt}{\alpha})^{i+1+k}}{i+1+k},$$

and

$$C_F = \frac{1}{\mu(e^\lambda-1)^2}\sum_{j=1}^{\infty}\sum_{i=0}^{j}\sum_{k=0}^{\infty}\sum_{\gamma=0}^{\infty}\sum_{b=0}^{\gamma} \frac{\binom{j}{i}\binom{\gamma}{b}(-1)^k\lambda^{j+\gamma+1}\theta^{i+k+b+1}j^k}{j!\gamma!\alpha^b(\theta+1)^{b+i}(i+k+1)} \times\left[\gamma(i+k+b+2,)+\frac{1}{\alpha}\gamma(i+k+b+3,)\right].$$

The Gini index can be derived from the $G = 1 - C_F$.

2.8 Entropies

The entropy quantifies the level of uncertainty associated with random variables. The Rényi entropy is defined as

$$\mathcal{J}_R(\gamma) = \frac{1}{1-\gamma}\log\left[\int f^\gamma(x)dx\right], \quad \gamma > 0, \gamma \neq 1.$$

Suppose $X \sim SP(\theta,\lambda,\alpha)$, Then, one can calculate

$$\int f^\gamma(x)dx = \left(\frac{\lambda\theta^2}{\alpha(\theta+1)(e^\lambda-1)}\right)^\gamma \times\int_0^\infty\left(1+\frac{x}{\alpha}\right)^\gamma e^{\gamma(\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}-\frac{\theta}{\alpha}x)}dx$$
$$= (\frac{\lambda\theta^2}{\alpha(\theta+1)(e^\lambda-1)})^\gamma \times\sum_{j=0}^{\gamma}\frac{\binom{\gamma}{j}}{\alpha^j}\int_0^\infty x^j e^{-\frac{\gamma\theta x}{\alpha}}\sum_{k=0}^{\infty}\frac{\gamma^k(\frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)})^k}{k!}dx$$
$$= \left(\frac{\lambda\theta^2}{\alpha(\theta+1)(e^\lambda-1)}\right)^\gamma\sum_{j=0}^{\gamma}\sum_{k=0}^{\infty}\frac{\binom{\gamma}{j}}{\alpha^j}\frac{\gamma^k}{k!}\frac{\lambda^k}{(\alpha(\theta+1))^k} \times\int_0^\infty x^j e^{-\frac{\theta(\gamma+k)x}{\alpha}}(\alpha(\theta+1)+\theta x)^k dx$$
$$= \left(\frac{\lambda\theta^2}{\alpha(\theta+1)(e^\lambda-1)}\right)^\gamma\times \sum_{j=0}^{\gamma}\sum_{k=0}^{\infty}\sum_{i=0}^{k}\frac{\binom{k}{i}\binom{\gamma}{j}(\gamma\lambda\theta)^k}{k!(\alpha)^{i+j}(\theta+1)^i}\int_0^\infty x^{k+j}e^{-\frac{\theta(\gamma+k)x}{\alpha}}dx.$$

Thus, the Rényi entropy of the SP distribution can be expressed as

$$\mathcal{J}_R(\gamma) = \frac{\gamma}{1-\gamma}\log\left(\frac{\lambda\theta^2}{\alpha(\theta+1)(e^\lambda-1)}\right)+\frac{1}{1-\gamma}\times \log\left[\sum_{j=0}^{\gamma}\sum_{k=0}^{\infty}\sum_{i=0}^{k}\frac{\binom{k}{i}\binom{\gamma}{j}(\gamma\lambda)^k(\alpha)^{k-i+1}\Gamma(k+j+1)}{k!(\theta)^{j+1}(\theta+1)^i(k+\gamma)^{k+j+1}}\right]. \quad (15)$$

Shannon entropy defined by $-E[\log f(X)]$ is the particular case of (15) for $\gamma \uparrow 1$. Limiting $\gamma \uparrow 1$ in (15) and using L'Hospital's rule, after considerable algebraic manipulation, we get

$$E[-\log f(X)] = -\log\left(\frac{\lambda\theta^2}{\alpha(\theta+1)(e^\lambda-1)}\right) +\frac{\theta}{\alpha}E(X)+\sum_{i=1}^{\infty}\frac{(-1)^i}{i\alpha^i}E(X^i) -\sum_{j=0}^{\infty}\sum_{k=0}^{j}\frac{\binom{j}{k}\lambda^{j+2}\Gamma(k+1)}{j!(e^\lambda-1)(\theta+1)^{k+2}(j+1)^{k+1}} \times\left[\theta(\theta+1)+\frac{(k+1)(2\theta+1)}{(j+1)}+\frac{(k+1)(k+2)}{(j+1)^2}\right]. \quad (16)$$

2.9 Mean deviations

Population dispersion can be quantified by evaluating the deviations from central values like the mean and median. For a random variable X, with $\mu = E(X)$ and $M = Median(X)$, the mean deviation about the mean and the mean deviation about the median, are defined respectively by

$$\delta_1 = \int_0^\infty |x-\mu|f(x)dx = 2\mu F(\mu)-2+2\int_\mu^\infty xf(x)dx,$$
$$\delta_2 = \int_0^\infty |x-M|f(x)dx = 2MF(M)-M-\mu+2\int_M^\infty xf(x)dx,$$

where $L(b) = \int_b^\infty xf(x)dx$. If X is SP random variable specified by (5)

$$\delta_1 = 2\mu F(\mu)-2\mu+2L(\mu),$$
$$\delta_2 = 2MF(M)-M-\mu+2L(M),$$

$$L(b) = \frac{\lambda}{e^{\lambda-1}}\sum_{j=0}^{\infty}\sum_{i=0}^{j}\frac{\binom{j}{i}\alpha^{j+1}}{(j+1)^{i+2}}\times \left[\Gamma\left(i+2,\frac{\theta\mu(j+1)}{\alpha}\right) +\frac{1}{\theta(j+1)}\Gamma\left(i+3,\frac{\theta\mu(j+1)}{\alpha}\right)\right], \quad (17)$$

where μ and M are defined in (13) and (12), respectively.

https://doi.org/10.57647/mathsci.2025.1901.02

3. Maximum Likelihood Estimation and Fisher information matrix

MLE is a commonly used method for estimating unknown parameters of a distribution, valued for its appealing statistical properties such as consistency and asymptotic normality. In this section, the MLEs of the parameters θ, α and λ are derived. Let$X_1, \cdots, X_n$ be a random sample from the SP distribution with unknown vector of parameters $\Theta = (\theta, \alpha, \lambda)'$. Then the log-likelihood function is given by

$$\begin{aligned}\log f(x;\Theta) &= \lambda\sum_{i=1}^{n} e^{-\frac{\theta}{\alpha}x_i} + \frac{\theta\lambda}{\alpha(\theta+1)}\\ &\quad \sum_{i=1}^{n} x_i e^{-\frac{\theta}{\alpha}x_i} - \frac{\theta}{\alpha}\sum_{i=1}^{n} x_i\\ &\quad + \sum_{i=1}^{n}\log(1+\frac{x_i}{\alpha}) + 2n\log\theta\\ &\quad -n\log\alpha - n\log(\theta+1)\\ &\quad -n\log(e^{\lambda}-1) + n\log\lambda. \qquad (18)\end{aligned}$$

The MLEs of the unknown parameters can be obtained by taking the first partial derivatives of (18) with respect to $\Theta = (\theta, \lambda, \alpha)'$. We get the following likelihood equations

$$\begin{aligned}\frac{\partial \log f(x;\boldsymbol{\Theta})}{\partial\lambda} &= \sum_{i=1}^{n} e^{-\frac{\theta}{\alpha}x_i} + \frac{\theta}{\alpha(\theta+1)}\\ &\quad \sum_{i=1}^{n} x_i e^{-\frac{\theta}{\alpha}x_i} - \frac{ne^{\lambda}}{e^{\lambda}-1} + \frac{n}{\lambda}.\\ \frac{\partial \log f(x;\boldsymbol{\Theta})}{\partial\theta} &= -\frac{1}{\alpha}\sum_{i=1}^{n} x_i + \frac{2n}{\theta}\\ &\quad -\frac{n}{(\theta+1)} - \frac{\theta(\theta+2)\lambda}{\alpha(\theta+1)^2}\sum_{i=1}^{n} x_i e^{-\frac{\theta}{\alpha}x_i}\\ &\quad -\frac{\theta\lambda}{\alpha^2(\theta+1)}\sum_{i=1}^{n} x_i^2 e^{-\frac{\theta}{\alpha}x_i}.\\ \frac{\partial \log f(x;\boldsymbol{\Theta})}{\partial\alpha} &= -\frac{1}{\alpha}\sum_{i=1}^{n}\left(\frac{x_i}{\alpha+x_i}\right) - \frac{n}{\alpha}\\ &\quad +\frac{\lambda\theta^2}{\alpha^3(\theta+1)}\sum_{i=1}^{n} x_i^2 e^{-\frac{\theta}{\alpha}x_i}\\ &\quad +\frac{\lambda\theta^2}{\alpha^2(\theta+1)}\sum_{i=1}^{n} x_i e^{-\frac{\theta}{\alpha}x_i} + \frac{\theta}{\alpha^2}\sum_{i=1}^{n} x_i.\end{aligned}$$

The solutions of these nonlinear equations do not have a closed form, so numerical methods can be employed to get the MLEs. The EM algorithm has emerged as a widely used method for statistical estimation in the presence of incomplete data, and in analogous contexts such as mixture model estimation (see, [27] and [28]). This iterative approach systematically substitutes missing data with estimates and updates the parameter values.

The complete-data distribution is defined by the density function

$$\begin{aligned}f(x,z;\Theta) &= \frac{z\theta^2 e^{-\frac{\theta}{\alpha}xz}(1+\frac{x}{\alpha})\left[\alpha(\theta+1)+\theta x\right]^{z-1}}{(\alpha(\theta+1))^z}\\ &\quad \times\frac{\lambda^z e^{-\lambda}}{z!(1-e^{-\lambda})},\end{aligned}$$

where $\Theta = (\theta, \alpha, \lambda)'$, $\quad x > 0,\ z \in N$.
Under the formulation, the E-step of an EM cycle requires the expectation of $(Z|X;\Theta^r)$ where $\Theta^{(r)} = (\theta^{(r)}, \alpha^{(r)}, \lambda^{(r)})'$, is the current estimate (in the rth iteration) of Θ.
The pdf of Z given X is given by

$$\begin{aligned}f(z|x) &= \Big((\alpha(\theta+1))^{1-z}\lambda^{z-1}(\alpha(\theta+1)+\theta x)^{z-1}\\ &\quad\times\exp\left[\frac{-\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)} + \frac{\theta}{\alpha}x - \frac{\theta}{\alpha}xz\right]\Big)\\ &\quad /(z-1)!.\end{aligned}$$

Therefore, its expected value is given by

$$E(Z|X;\Theta) = 1 + \frac{\lambda e^{-\frac{\theta}{\alpha}x}(\alpha(\theta+1)+\theta x)}{\alpha(\theta+1)}.$$

The EM cycle is completed with the M-step by using the maximum likelihood estimation over Θ, with the missing Zs replaced by their conditional expectations given above. Thus, an EM iteration is given by

$$\begin{aligned}\hat{\lambda}^{(t+1)} &= n^{-1}\left(1-e^{-\lambda(t)}\right)\sum_{i=1}^{n} w_i^{(t)}\\ \hat{\theta}^{(t+1)} &= 2n\left(\sum_{i=1}^{n}\frac{x_i+\alpha^{(t)}}{\alpha^{(t)}(\theta^{(t)}+1)+\theta^{(t)}x_i}\right.\\ &\quad -\sum_{i=1}^{n}\frac{(x_i+\alpha^{(t)})w_i^{(t)}}{\alpha^{(t)}(\theta^{(t)}+1)+\theta^{(t)}x_i}\\ &\quad \left.+\frac{1}{\alpha^{(t)}}\sum_{i=1}^{n} w_i^{(t)}x_i + \frac{\sum_{i=1}^{n} w_i^{(t)}}{\theta^{(t)}+1}\right)^{-1}\\ \hat{\alpha}^{(t+1)} &= \left(\frac{\theta^{(t)}}{\alpha^{(t)}}\sum_{i=1}^{n} x_i w_i^{(t)} - \sum_{i=1}^{n}(\frac{x_i}{\alpha^{(t)}+x_i}) - \sum_{i=1}^{n} w_i^{(t)}\right)\\ &\quad\times\left((\theta^{(t)}+1)\sum_{i=1}^{n}\frac{(w_i^{(t)}-1)}{\alpha^{(t)}(\theta^{(t)}+1)+\theta^{(t)}x_i}\right)^{-1},\end{aligned}$$

where $\hat{\theta}^{(r+1)}$ and $\hat{\alpha}^{(r+1)}$ are found numerically. Hence for $i = 1, \ldots, n$ we have

$$w_i^{(t)} = 1 + \frac{\lambda^{(t)} e^{-\frac{\theta^{(t)}}{\alpha^{(t)}}x_i}\left(\alpha^{(t)}(\theta^{(t)}+1)+\theta^{(t)}x_i\right)}{\alpha^{(t)}(\theta^{(t)}+1)}.$$

The relations described above outline an iterative computational method for obtaining maximum likelihood estimators, which we have implemented on a real dataset in Section 5.

3.1 Asymptotic variance and covariance of MLEs

It is known that under some regular conditions, as the sample size increases, the distribution of the MLE tends

https://doi.org/10.57647/mathsci.2025.1901.02

to the trivariate normal distribution with mean $(\theta, \lambda, \alpha)$ and covariance matrix equal to the inverse of the Fisher information matrix. The trivariate normal distribution can be used to construct approximate confidence intervals for the parameters $\Theta = (\theta, \lambda, \alpha)'$. Let $I(\Theta) = I(\theta, \lambda, \alpha; x_{obs})$ be the observed matrix with elements I_{ij} with i,j = 1, 2, 3. The elements of the observed information matrix are given by:

$$
\begin{aligned}
I_{11} &= \frac{2n}{\theta^2} - \frac{n}{(\theta+1)^2} + \frac{2\lambda \sum_{i=1}^n x_i e^{-\frac{\theta}{\alpha}x_i}}{\alpha(\theta+1)^3} \\
&\quad - \frac{\theta\lambda \sum_{i=1}^n x_i^3 e^{-\frac{\theta}{\alpha}x_i}}{\alpha^3(\theta+1)} \\
&\quad - \frac{\lambda((\theta+1)^2-2) \sum_{i=1}^n x_i^2 e^{-\frac{\theta}{\alpha}x_i}}{\alpha^2(\theta+1)^2}. \\
I_{12} &= I_{21} = \frac{\theta(\theta+2) \sum_{i=1}^n x_i e^{-\frac{\theta}{\alpha}x_i}}{\alpha(\theta+1)^2} \\
&\quad + \frac{\theta \sum_{i=1}^n x_i^2 e^{-\frac{\theta}{\alpha}x_i}}{\alpha^2(\theta+1)}. \\
I_{22} &= -\frac{e^{\lambda} n}{(e^{\lambda}-1)^2} + \frac{n}{\lambda^2}. \\
I_{13} &= I_{31} = -\frac{\lambda\theta^2 \sum_{i=1}^n x_i^3 e^{-\frac{\theta}{\alpha}x_i}}{\alpha^4(\theta+1)} \\
&\quad + \frac{\lambda\theta(\theta-1) \sum_{i=1}^n x_i^2 e^{-\frac{\theta}{\alpha}x_i}}{\alpha^3(\theta+1)} \\
&\quad + \frac{\theta\lambda(\theta-2) \sum_{i=1}^n x_i e^{-\frac{\theta}{\alpha}x_i}}{\alpha^2(\theta+1)^2}. \\
I_{23} &= I_{32} = \frac{\theta^2 \sum_{i=1}^n x_i^2 e^{-\frac{\theta}{\alpha}x_i}}{\alpha^3(\theta+1)} + \frac{\theta^2 \sum_{i=1}^n x_i e^{-\frac{\theta}{\alpha}x_i}}{\alpha^2(\theta+1)}. \\
I_{33} &= -\frac{\theta^3\lambda \sum_{i=1}^n x_i^3 e^{-\frac{\theta}{\alpha}x_i}}{\alpha^5(\theta+1)} - \frac{\theta^2\lambda(\theta-3) \sum_{i=1}^n x_i^2 e^{-\frac{\theta}{\alpha}x_i}}{\alpha^4(\theta+1)} \\
&\quad + \frac{2\theta^2\lambda \sum_{i=1}^n x_i e^{-\frac{\theta}{\alpha}x_i}}{\alpha^3(\theta+1)} \\
&+ \frac{2\theta \sum_{i=1}^n x_i}{\alpha^3} - \frac{1}{\alpha^2}\sum_{i=1}^n \frac{x_i}{x_i+\alpha} \\
&\quad - \frac{1}{\alpha}\sum_{i=1}^n \frac{x_i}{(x_i+\alpha)^2} - \frac{n}{\alpha^2}.
\end{aligned}
$$

The elements of the expected information matrix, $J(\Theta)$, are calculated by taking the expectations of I_{ij}, i,j=1,2,3 with respect to the distribution of X. When the expectations of I_{ij}, i,j=1,2,3 is obtained, we would have the matrix $J(\Theta)$, the inverse of $J(\Theta)$, evaluated at $\hat{\Theta}$ provides the asymptotic variance-covariance matrix of MLEs. Alternative estimates can be obtained from the inverse of the observed information matrix since it is a consistent estimator of $J^{-1}(\Theta)$.

4. Simulation study

In this section, we demonstrate the methodology from Section 3 using simulated data. A numerical study is conducted to apply the EM algorithm for parameter estimation, and the accuracy of the estimates is evaluated. The simulation study is carried out using R software, with a relative tolerance of 10^{-4} set as the convergence criterion for the algorithms. The accuracy of the algorithms is examined for a proposed model (11) with different values for $(\theta, \alpha, \lambda)$ and $n = 100, 200, 500$. The simulation was again performed 1000 times and the means and standard deviations reported in Table Table 1, which shows that the EM method has a good performance-complexity trade-off as the number of observations increases.

5. Application to Real-Data

In this section, we assess the practical performance of the Sushila-Poisson distribution by comparing it with several existing distributions. To facilitate this comparison, three real-world datasets are analyzed. In the first dataset, we demonstrate that the SP distribution is a strong competitor to two important distributions, namely the gamma and Weibull distributions, which are two-parameter distributions. In the second dataset, we show that the SP distribution performs well when compared to a three-parameter distribution introduced by Daghagh et al. [21]. Since the SP and the distribution of Daghagh et al. [21] have approximately the same structures. For the third dataset, we compare the Sushila-Poisson distribution with the modified Weibull (MW) distribution introduced by [29].

In order to identify the shape of the hazard rate function of the data, we consider a graphical method based on the Total Time on Test (TTT) plot. As we know, the empirical TTT plot is given by

$$G\left(\frac{r}{n}\right) = \frac{(\sum_{i=1}^r X_{i:n} + (n-r)X_{r:n})}{\sum_{i=1}^n X_{i:n}},$$

where $X_{i:n}$ denotes the ith order statistic of the sample. If the empirical TTT transform is convex, concave, convex, then concave and concave, then convex, the shape of the corresponding hazard rate function is, respectively, decreasing, increasing, bathtub-shaped, and upside-down bathtub, see [30]. In order to compare distributions, we consider the values of the log-likelihood $(-\log L)$, Akaike Information Criterion (AIC), and the AICC (a version of Akaike information criterion, that has a correction for small sample sizes). The better distribution corresponds to smaller $-\log L$, AIC, and AICc. In addition, we apply the Kolmogorov-Smirnov statistic (and associated p-value) to verify which distribution fits the data better.

The first data set consists of the number of successive failures for the air conditioning system of each member in feet of 13 Boeing 720 jet airplanes. The pooled data with 213 observations are considered here. First data set is given by as follows:
194, 15, 41, 29, 33, 181, 413, 14, 58, 37, 100, 65, 9, 180, 447, 184, 36, 201, 118, 34, 31, 18, 18, 67, 57, 62, 7, 22, 34, 90, 10, 60, 186, 61, 49, 14, 24, 56, 20, 79, 84, 44, 59, 29,156, 118, 25, 310, 76, 26, 44, 23, 62, 130, 208, 70 , 101, 208, 74, 57, 48, 29, 502, 12, 70, 21, 29, 386, 59, 27, 153, 26, 326,55, 320, 56, 104, 220, 239, 47, 246, 176, 182, 33, 15, 104, 35, 23, 261, 87, 7, 120, 14, 62,

Table 1. Means and standard deviations (in parentheses) of the estimated parameters in 1000 simulated realizations from SP model

$(\alpha, \theta, \lambda)$	n	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$
(0.5, 0.5, 0.2)	100	0.5998(0.4610)	0.5662(0.3442)	0.2250(0.2398)
	200	0.5583(0.2802)	0.5424(0.2221)	0.2055(0.0685)
	500	0.5172(0.1203)	0.5138(0.1012)	0.2013(0.0129)
(0.5, 1.25, 1)	100	0.6837(1.1262)	1.4891(1.7831)	1.2498(0.9593)
	200	0.5555(0.3312)	1.2925(0.6151)	1.2165(0.8501)
	500	0.50825(0.1718)	1.2358(0.3293)	1.096(0.5367)
(1.25, 1, 0.5)	100	1.5332(2.8815)	1.0912(1.5087)	0.7701(0.8164)
	200	1.3365(0.7898)	1.0140(0.4611)	0.6535(0.6110)
	500	1.2884(0.3713)	1.0144(0.2363)	0.5359(0.2687)
(1.25, 0.5, 0.5)	100	1.4146(1.2501)	0.5187(0.3788)	0.7611(0.8110)
	200	1.3853(0.9504)	0.5223(0.3288)	0.7350(0.7009)
	500	1.2538(0.4426)	0.4902(0.1687)	0.6427(0.5390)

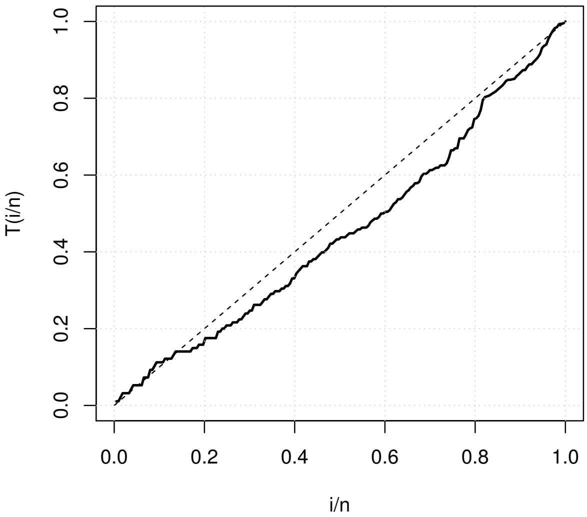

Figure 3. The empirical TTT plot of the data set 1

Table 2. MLEs of the parameters and corresponding criteria for data set 1

Model	*parameters*	$-\log L$	*AIC*	*AICC*
Gamma	$\hat{\alpha} = 0.9134$	1176.902	2360.8210	2360.7880
	$\hat{\beta} = 0.0098$			
Weibull	$\hat{\alpha} = 0.9244$	1178.4105	2359.3950	2359.4520
	$\hat{\beta} = 0.0112$			
SP	$\hat{\theta} = 1.9355$	1175.902	2357.804	2357.919
	$\hat{\alpha} = 225.1443$			
	$\hat{\lambda} = 2.0048$			

Table 3. K-S test for data set 1

Model	$K-S$	p−value
Gamma	0.0625	0.3763
Weibull	0.0524	0.6027
SP	0.04367	0.8112

47, 225, 71, 246, 21, 42, 20,5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95, 97, 51, 11, 4, 141, 18, 142, 68, 77, 80, 1, 16, 106, 206, 82, 54,31, 216, 46, 111, 39, 63, 18, 191, 18, 163, 24, 50, 44, 102, 72, 22, 39, 3, 15, 197, 188, 79, 88, 46, 5, 5, 36, 22, 139, 210,97, 30, 23, 13, 14, 359, 9, 12, 270, 603, 3, 104, 2, 438, 50, 254, 5, 283, 35, 12, 130, 493, 487, 18, 100, 7, 98, 5, 85, 91,43, 230, 3, 130, 102, , 209, 14, 57, 54, 32, 67, 59, 134, 152, 27, 14, 230, 66, 61, 34.
From Figure 3, we can notice that the hazard rate of the data set is decreasing. We observe from Table 2 and Table 3 that the SP distribution provides an improved fit over other distributions that fit lifetime data. The fitted density, the empirical CDF plot, and the p-p plot of the *SP* distribution model are presented in Figure 6. The Figure indicates a desirable fit of the SP distribution.

The second data set reported in [31] corresponds to the survival times in years of a group of forty-five patients given chemotherapy treatment. The data set is 0.047, 0.115, 0.121, 0.132, 0.164, 0.197, 0.203, 0.260, 0.282, 0.296, 0.334, 0.395,0.458, 0.466, 0.501, 0.507, 0.529, 0.537, 0.540, 0.641, 0.644, 0.696, 0.841, 0.863,1.099, 1.219, 1.271, 1.326, 1.447, 1.485, 1.553, 1.581, 1.589, 2.178, 2.343, 2.416, 2.444, 2.825, 2.830, 3.578, 3.658, 3.743, 3.978, 4.003, 4.033.
We intend to illustrate the applicability of the new distribution, hence we fit *SP* distribution for the data. We compare the Sushila-Poisson distribution with three-parameter distribution with the same structure *SP* called Sushila-Geometric with pdf given by

$$f(x) = \frac{\theta^2}{\alpha(\theta+1)}\left(1+\frac{x}{\alpha}\right)(1-p)e^{-\frac{\theta}{\alpha}x}$$
$$\left(1-p\left(1+\frac{\theta x}{\alpha(\theta+1)}\right)e^{-\frac{\theta}{\alpha}x}\right)^{-2},$$
$$\alpha, \theta > 0, 0 < p < 1.$$

From Figure 4, we can notice that the hazard rate of the data set is almost constant. We observe from Table 4 and Table 5 that the *SP* distribution provides an improved fit over a distribution that is commonly used for fitting lifetime data.

The third data set is presented Table 1 [32] and consist of 61 observed recidivism failure times (in days) of in-

https://doi.org/10.57647/mathsci.2025.1901.02

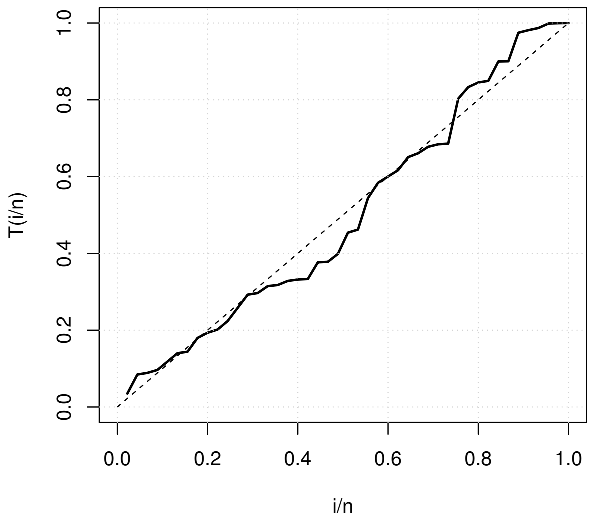

Figure 4. The empirical TTT plot of the data set 2

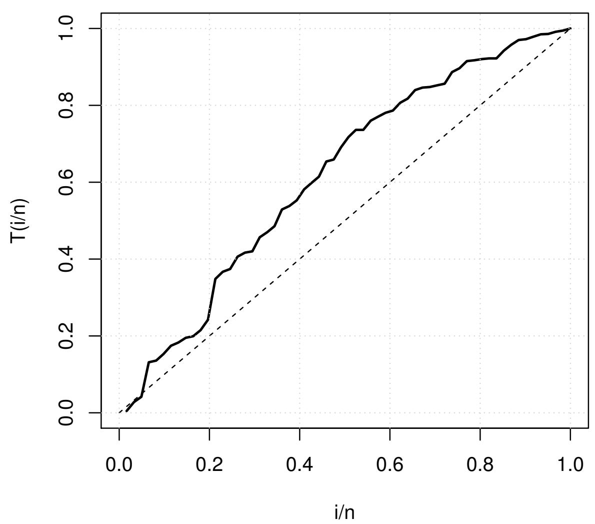

Figure 5. The empirical TTT plot of the data set 3

Table 4. MLEs of the parameters and corresponding criteria for data set 2

$Model$	$parameters$	$-\log L$	AIC	$AICC$
SG	$\hat{\theta} = 0.06207$	59.3628	124.7257	125.311
	$\hat{\alpha} = 0.07522$			
	$\hat{p} = 0.77726$			
SP	$\hat{\theta} = 2.3246$	58.14909	122.2982	122.8835
	$\hat{\alpha} = 2.5758$			
	$\hat{\lambda} = 0.3016$			

Table 5. K-S test for data set 2

$Model$	$K-S$	$p-$value
SG	0.11449	0.5581
SP	0.09445	0.7821

Table 6. MLEs of the parameters and corresponding criteria for data set 3

$Model$	$parameters$	$-\log L$	AIC	$AICC$
MW	$\hat{\alpha} = 0.1352$	393.9610	793.9221	794.3432
	$\hat{\beta} = 0.0300$			
	$\hat{\lambda} = 0.0071$			
SP	$\hat{\theta} = 0.00357$	387.3743	780.7486	781.1697
	$\hat{\alpha} = 0.380746$			
	$\hat{\lambda} = 0.019892$			

Table 7. K-S test for data set 3

$Model$	$K-S$	$p-$value
MW	0.1109	0.4411
SP	0.1004	0.5707

dividuals released directly from correctional institutions to parole in the District of Columbia, USA is
1, 6, 9, 29, 30, 34, 39, 41, 45, 49, 56, 84, 89, 91, 100, 103, 104, 115, 119, 124, 138, 141, 146, 156, 162, 168, 183, 185, 198, 209, 217, 228, 238, 233, 241, 252, 258, 271, 275, 276, 279, 282, 305, 313, 329, 331, 334, 336, 362, 384, 404, 408, 422, 438, 441, 465, 486, 556.
We fit the SP distribution to the real dataset and compare its fitting with the Modified Weibull (MW) distribution [29]with pdf given by
$f(x) = \alpha x^{\beta-1}(\beta+\lambda x)e^{\lambda x}e^{-\alpha x^{\beta}\exp(\lambda x)}, \alpha, \beta \geq 0, \lambda > 0,$
that the MW distribution is a distribution commonly used in literature for fitting lifetime data. From Figure 5, we can notice that the hazard rate of the data set is increasing. We observe from Table 6 and Table 7 that the SP distribution provides an improved fit over a distribution that is commonly used for fitting lifetime data. The fitted density, the empirical CDF plot, and the p-p plot of the SP distribution model are presented in Figure 8. The Figure indicates a desirable fit of the SP distribution.

From TTT plots, we noticed that the hazard rate functions for three data sets are decreasing, constant and increasing respectively. Figure 9 shows the estimated hazard rate functions for all three data sets. We notice that what we predicted about the hazard rate functions from the TTT plots is correct.

6. Conclusion

This paper proposes a new three-parameter distribution, referred to as the Sushila-Poisson (SP) distribution. The proposed model generalizes seversl well-known distributions, including the Lindley, Lindley-Poisson and Sushila distributions, as special cases. The hazard rate function of the SP distribution is highly flexible and can exhibit various shapes such as increasing, decreasing, bathtub-shaped and upside-down bathtub forms making it suitable for a wide range of applications. Several key mathematical properties of the distribution are derived including its density expansion, hazard rate function, quantile function, moments, mean deviation, order statistics and both Shannon and Renyi entropies, all

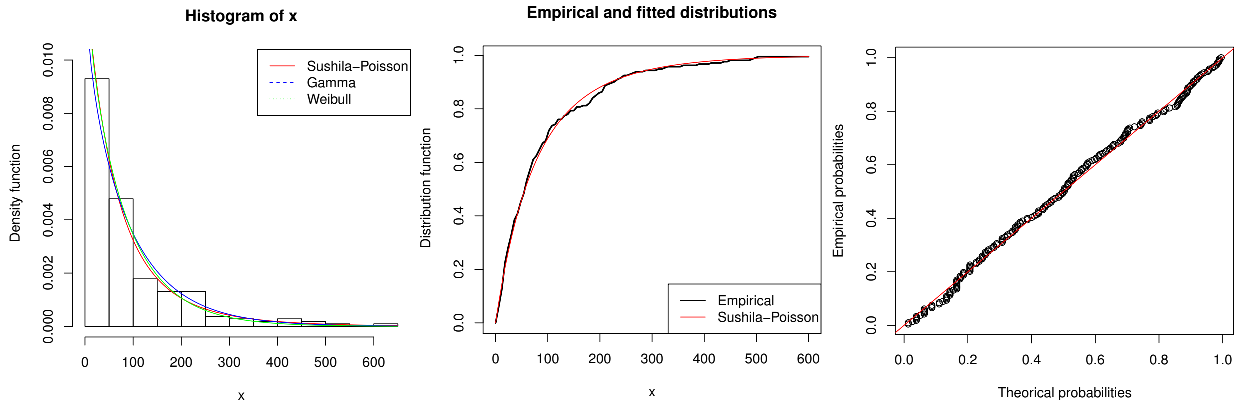

Figure 6. Plots of the estimated pdf and cdf and p-p plot of the SP model for data set1

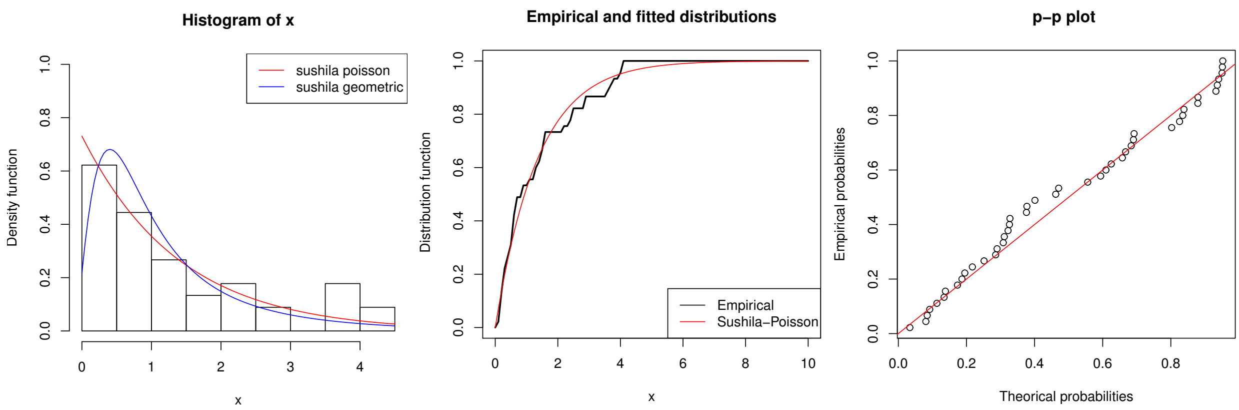

Figure 7. Plots of the estimated pdf and cdf and p-p plot of the SP model for data set2

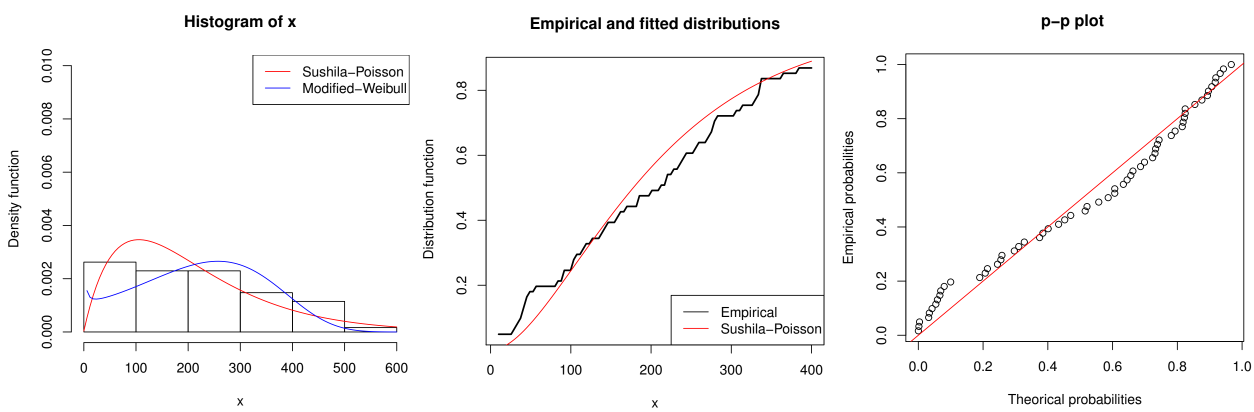

Figure 8. Plots of the estimated pdf and cdf and p-p plot of the SP model for data set3

https://doi.org/10.57647/mathsci.2025.1901.02

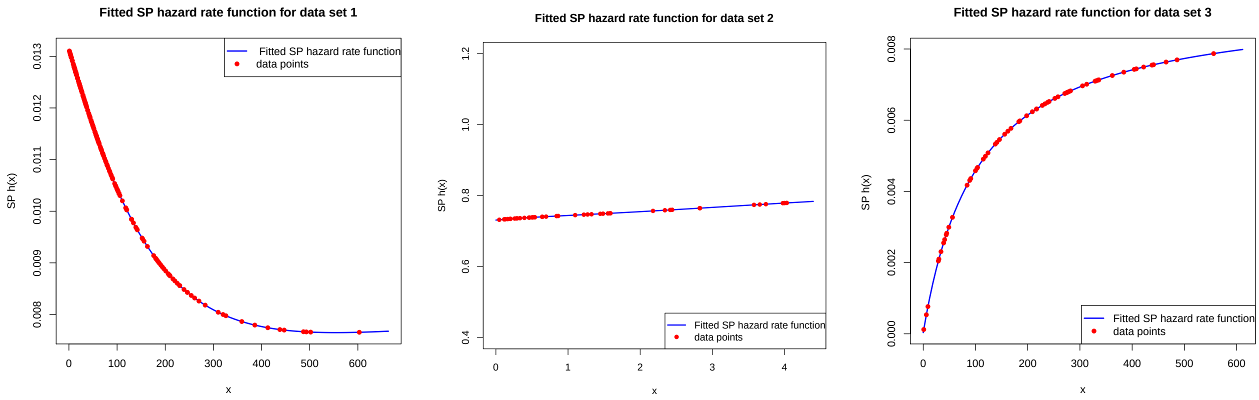

Figure 9. Plots of the estimated SP hazard rate functions for 3 data sets

expressed in closed form using familiar mathematical functions. Parameter estimation is carried out using the maximum likelihood method, with the corresponding observed information matrix also obtained. Additionally, the Expectation-Maximization(EM) algorithm is employed as a general purpose approach for computing the maximum likelihood estimates (MLEs) of the model parameters.

To demonstrate the practical utility of the proposed distribution, the SP model is fitted to three real-world datasets.

Authors contributions

All the authors have participated sufficiently in the intellectual content, conception and design of this work or the analysis and interpretation of the data (when applicable), as well as the writing of the manuscript.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Conflict of interests

The author declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open access

References

1. Shanker R, Sharma S, Shanker U, and Shanker R. Sushila distribution and its application to waiting times data. International Journal of Business Management 2013; 3:1–11
2. Borah M and Saikia KR. Certain properties of discrete Sushila. Statistics 2016; 5
3. Borah M and Saikia K. Zero-Truncated Discrete Shanker Distribution and Its Applications. Biometrics Biostatistics International Journal 2017; 5:00152
4. Yamrubboon D, Bodhisuwan W, Pudprommarat C, and Saothayanun L. The negative binomial-Sushila distribution with application in count data analysis. Thailand Statistician 2017; 15:69–77
5. Elgarhy M and Shawki A. Exponentiated SUSHILA distribution. International Journal of Scientific Engineering and Science 2017; 1:9–12
6. Shawki A and Elgarhy M. Kumaraswamy Sushila distribution. Int J Sci Eng Sci 2017; 1:29–32
7. Rather A and Subramanian C. Length biased Sushila distribution. Universal Review 2018; 7:1010–23
8. Borah M and Hazarika J. Poisson-Sushila distribution and its applications. International Journal of Statistics & Economics 2018; 19:37–45
9. Rather A and Subramanian C. On weighted sushila distribution with properties and its applications. Int. J. Sci. Res. in Mathematical and Statistical Sciences Vol 2019; 6
10. Pudprommarat C. HURDLE POISSON–SUSHILA DISTRIBUTION AND ITS APPLICATION. *INTERNATIONAL ACADEMIC MULTIDISCIPLINARY RESEARCH CONFERENCE IN AMSTERDAM 2019*. 2019 :98–103
11. Oliveira RP de, Oliveira Peres MV de, Martinez EZ, and Achcar JA. Use of a discrete Sushila distribution in the analysis of right-censored lifetime data. Model Assisted Statistics and Applications 2019; 14:255–68
12. Pudprommarat C. ZERO-ONE INFLATED NEGATIVE BINOMIAL-SUSHILA DISTRIBUTION AND ITS APPLICATION. *International Academic Multidisciplinary Research Conference in Rome 2020*. 2020 :20–8

https://doi.org/10.57647/mathsci.2025.1901.02

13. Adetunji AA. Transmuted Sushila Distribution and its application to lifetime data. Journal of Mathematical Analysis and Modeling 2021; 2:1–14
14. Shaw WT and Buckley IR. The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. arXiv preprint arXiv:0901.0434 2009
15. Boonthiem S, Moumeesri A, Klongdee W, and Ieosanurak W. A new Sushila distribution: properties and applications. European Journal of Pure and Applied Mathematics 2022; 15:1280–300
16. Oliveira RP de, Oliveira Peres MV de, Achcar JA, and Martinez EZ. A new class of bivariate Sushila distributions in presence of right-censored and cure fraction. Brazilian Journal of Probability and Statistics 2023; 37:55–72
17. Aryuyuen S, Panphut W, and Pudprommarat C. Bivariate Sushila Distribution Based on Copulas: Properties, Simulations, and Applications. Lobachevskii Journal of Mathematics 2023; 44:4592–609
18. Yamrubboon D, Thongteeraparp A, Bodhisuwan W, Jampachaisri K, and Volodin A. Bayesian inference for the negative binomial-Sushila linear model. Lobachevskii Journal of Mathematics 2019; 40:42–54
19. Bodhisuwan R, Denthet S, and Acoose T. Zero-Truncated Negative Binomial Weighted-Lindley Distribution and Its Application. Lobachevskii Journal of Mathematics 2021; 42:3105–11
20. Atikankul Y, Wattanavisut A, and Liu S. The Negative Binomial-Generalized Lindley Distribution for Overdispersed Data. Lobachevskii Journal of Mathematics 2022; 43:2378–86
21. Daghagh Iranmanesh O. Sushila-Geometric distribution, properties, and applications. Statistics, Optimization and Information Computing, accepted
22. Gui W, Zhang S, and Lu X. The Lindley-Poisson distribution in lifetime analysis and its properties. Hacettepe journal of mathematics and statistics 2014; 43:1063–77
23. Shaked M and Shanthikumar JG. Supermodular stochastic orders and positive dependence of random vectors. Journal of Multivariate Analysis 1997; 61:86–101
24. Peng B, Xu Z, and Wang M. The exponentiated lindley geometric distribution with applications. Entropy 2019; 21:510
25. Arratia R, Gordon L, and Waterman M. An extreme value theory for sequence matching. The annals of statistics 1986 :971–93
26. Pundir S, Arora S, and Jain K. Bonferroni curve and the related statistical inference. Statistics & probability letters 2005; 75:140–50
27. Dempster AP, Laird NM, and Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the royal statistical society: series B (methodological) 1977; 39:1–22
28. McLachlan GJ and Krishnan T. The EM algorithm and extensions. John Wiley & Sons, 2007
29. Lai C, Xie M, and Murthy D. A modified Weibull distribution. IEEE Transactions on reliability 2003; 52:33–7
30. Aarset MV. How to identify a bathtub hazard rate. IEEE transactions on reliability 1987; 36:106–8
31. Bekker A, Roux J, and Mosteit P. A generalization of the compound Rayleigh distribution: using a Bayesian method on cancer survival times. Communications in Statistics-Theory and Methods 2000; 29:1419–33
32. Stollmack S and Harris CM. Failure-rate analysis applied to recidivism data. Operations Research 1974; 22:1192–205

https://doi.org/10.57647/mathsci.2025.1901.02