

A Size-Biased Poisson-Gamma Lindley Distribution with Application

Fatma Zohra Seghier¹, Halim Zeghdoudi^{1,*}, Abbas Benchaabane²

¹LaPS Laboratory, Badji-Mokhtar University, Annaba, Algeria

fatmas81@yahoo.fr, halimzeghdoudi77@gmail.com

²Laboratory of Analysis and Control of Differential Equations, Univ. 8 May
1945 Guelma, Algeria

benchaabane.abbes@univ-guelma.dz

*Correspondence: halimzeghdoudi77@gmail.com

ABSTRACT. In this paper, a size-biased Poisson-Gamma Lindley distribution (SBPGLD) has been obtained by size-biasing the Poisson-Gamma Lindley distribution (PGLD) introduced recently by Nedjar and Zeghdoudi (2020). Some of its statistical properties have been discussed. The method of maximum likelihood and the method of moments for the estimation of its parameters have been discussed. Also, an application on the real data of survival times of (56) Indian state of Kerala individus infected with Nipah virus is given.

1. Introduction

Nedjar and Zeghdoudi (2020) introduced a new distribution, named Poisson-gamma Lindley (PGLD) having probability mass function (p.m.f.)

$$P_0(x; \theta, \beta) = \frac{\theta^2 (\beta + \beta\theta - \theta)x + \beta(1 + \theta) + 1}{\beta (1 + \theta)^{x+3}}, \quad x = 0, 1, \dots; \quad \theta > 0; \quad \beta > \frac{\theta}{1 + \theta} \quad (1.1)$$

The first four moments about origin and the variance of PGLD obtained by Nedjar and Zeghdoudi (2020) are as follows:

$$\begin{aligned} \mu_1 &= \frac{-\theta + 2\beta(1 + \theta)}{\beta\theta(1 + \theta)}, \quad \mu_2 = \frac{-4\theta + 8\beta\theta + 2\beta\theta^2 + 6\beta - \theta^2}{\beta\theta^2(1 + \theta)}, \\ \mu_3 &= \frac{2\theta^3\beta^2 - \theta^3\beta + 6\theta^2\beta^2 - \theta^2\beta - \theta^2 + 6\theta\beta^2 + 2\beta^2}{\beta^2\theta^2(1 + \theta)^2}. \end{aligned}$$

Received: 7 Oct 2021.

Key words and phrases. Lindley distribution; estimation; gamma Lindley distribution.

The PSD arises from the Poisson distribution when its parameter λ follows a lifetime distribution named Gamma Lindley distribution introduced by Zeghdoudi, Nedjar (2016a) having probability density function (p.d.f.).

$$f_0(x; \theta, \beta) = \frac{\theta^2 ((\beta + \beta\theta - \theta)x + 1)e^{-\theta x}}{\beta(1+\theta)}, \quad x > 0; \theta > 0; \beta > \frac{\theta}{1+\theta} \quad (1.2)$$

Size-biased distributions are introduced firstly by Fisher (1937) to model ascertainment biases and in a unifying theory. Further, Van Deusen (1986) and Lappi and Bailey (1987) have studied and applied size-biased distributions to analyze HPS diameter increment data. Also, Patil and Rao (1978) used this distribution to modeling ecology problem and human population. There are many works study the size-biased distributions and their applications ,

including Ghitany, M.E., Al-Mutairi, D.K. (2008), Alavi and Chinipardaz (2009), Ducey and Gove (2015), Tanka and Srivastava (2014), Shanker (2016,2017), Shanker and Shukla (2017), Shanker and Mishra (2017), Shanker and Shukla (2018), Ayesha (2017), Beghriche and Zeghdoudi (2019), are some among others.

Let a random variable X has the original probability distribution, then a $P_0(x; \theta, \beta)$ simple size-biased distribution is given by its probability function:

$$P_1(x; \theta, \beta) = \frac{xP_0(x; \theta, \beta)}{\mu_1'} \quad (1.3)$$

Where $\mu_1' = E(X) = \sum_{x=0}^{\infty} xP_0(x; \theta, \beta)$ is the mean of the original probability distribution.

In the present paper, a size-biased Poisson-Gamma Lindley distribution (SBPGLD) has been proposed. It's raw and central moments based properties including coefficient of variation, skewness, kurtosis , index of dispersion and other statistical properties have been obtained and discussed. The method of moment and the method of maximum likelihood estimation have been discussed for estimating the parameter of SBPGLD. Is a first time in literature a work used individus infected with Nipah virus in distribution theory. These contributions are motivated by the following:

- the size-biased Poisson-Gamma Lindley distribution (SBPGLD) use may be restricted to the tail of a distribution, but it is easy to apply;

- The formulas of the mean, variance, and coefficient of variation, skewness, kurtosis and index of dispersion are simple in form and may be used as quick approximations in many cases. But in general it is applicable to try out simpler distributions than more complicated ones;
- This new distribution has advantages including a number of parameters (two) which we can modeled actuarial phenomena (collective model);
- The SBPGLD distribution can be used quite effectively in analyzing many real lifetime data sets: application to Ebola and Nipah virus and application to bladder cancer patients;
- The SBPGLD distribution gives adequate fits to many data sets.

The paper is organized as follows. In Section 2, we introduce the SBPGL distribution, and give immediate properties. In Section 3 we introduce the statistical properties of SBPGLD. In Section 4, we are interested in parameters estimation using both the maximum likelihood and the moment method. In this last section, simulation studies are reported and as well, are provided data driven applications allowing comparisons between others distributions. We finish the paper with a conclusive section.

2. Size-Biased Poisson-Gamma Lindley Distribution

Using (1.1) and (1.3), the p.m.f. of the size-biased Poisson-Gamma Lindley distribution (SBPGLD) with parameters θ , β can be obtained as

$$\begin{aligned}
 P_1(x; \theta, \beta) &= \frac{xP_0(x; \theta, \beta)}{\mu_1} \\
 &= \frac{\theta^2}{\beta} \frac{\beta\theta(1+\theta)}{-\theta+2\beta(1+\theta)} \frac{x((\beta+\beta\theta-\theta)x+\beta(1+\theta)+1)}{(1+\theta)^{x+3}} \\
 &= \frac{\theta^3}{-\theta+2\beta(1+\theta)} \frac{x((\beta+\beta\theta-\theta)x+\beta(1+\theta)+1)}{(1+\theta)^{x+2}}
 \end{aligned}$$

Where $\mu_1 = \frac{-\theta+2\beta(1+\theta)}{\beta\theta(1+\theta)}$ and $x=1,2,3,\dots$; $\theta > 0$; $\beta > \frac{\theta}{1+\theta}$

$$P_1(x; \theta, \beta) = \frac{\theta^3}{-\theta + 2\beta(1 + \theta)} \frac{x((\beta + \beta\theta - \theta)x + \beta(1 + \theta) + 1)}{(1 + \theta)^{x+2}} \tag{2.1}$$

To study the nature and behavior of SBPGLD for varying values of parameters θ and β , a number of graphs of the pmf of SBPGLD have been drawn and presented in the figure 1.

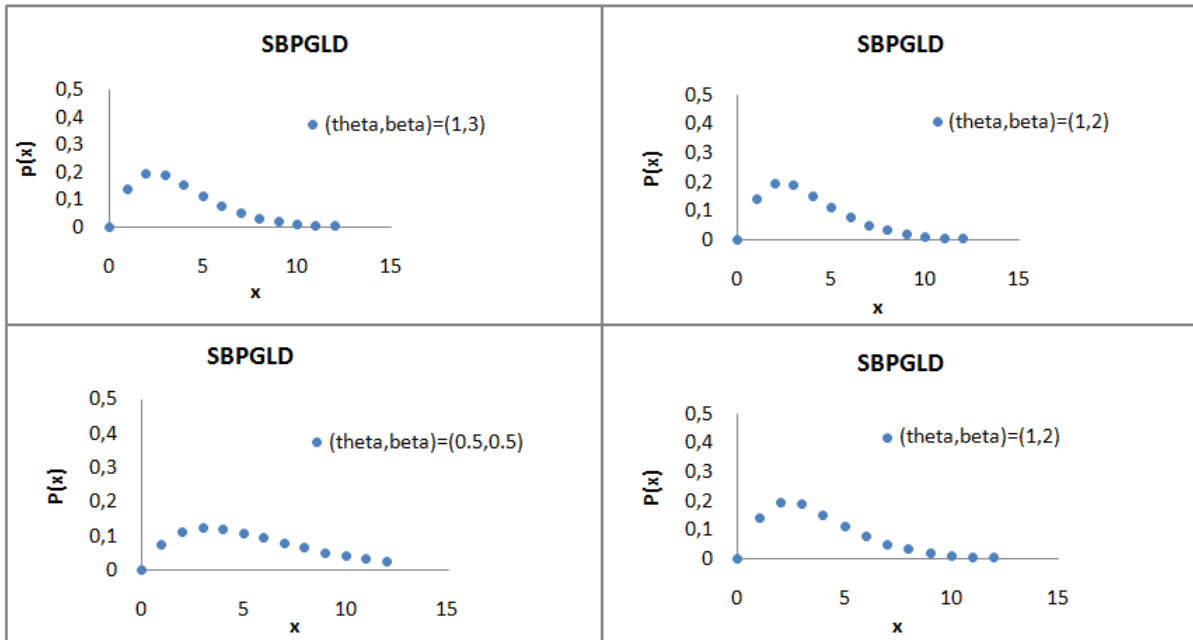


Figure.1. Graph of the probability mass function of SBPGLD for varying values of parameters θ and β .

The SBPGLD distribution also arises from the size-biased Poisson (SBPD) distribution with p.m.f.

$$g(x|\lambda) = e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!}; \quad x = 1, 2, \dots, \quad \theta > 0 \tag{2.2}$$

when its parameter λ follows size-biased Gamma Lindley distribution (SBGLD) with p.d.f.

$$h(\lambda; \theta, \beta) = \frac{\theta^3 \lambda ((\beta + \beta\theta - \theta)\lambda + 1) e^{-\theta\lambda}}{2\beta(1 + \theta) - \theta}, \quad \lambda > 0; \quad \theta > 0; \quad \beta > \frac{\theta}{1 + \theta} \tag{2.3}$$

Thus, the pmf of SBPGLD can be obtained as

$$\begin{aligned}
P_1(x; \theta, \beta) &= \int_0^{\infty} g(x|\lambda)h(\lambda; \theta, \beta)d\lambda \\
&= \frac{\theta^3}{(2\beta(1+\theta)-\theta)(x-1)!} \int_0^{\infty} \lambda^{x-1} e^{-\lambda} \lambda((\beta + \beta\theta - \theta)\lambda + 1) e^{-\theta\lambda} d\lambda \\
&= \frac{\theta^3}{(2\beta(1+\theta)-\theta)(x-1)!} \left[\int_0^{\infty} \lambda^x e^{-(1+\theta)\lambda} d\lambda + (\beta + \beta\theta - \theta) \int_0^{\infty} \lambda^{x+1} e^{-(1+\theta)\lambda} d\lambda \right] \\
&= \frac{\theta^3}{(2\beta(1+\theta)-\theta)(x-1)!} \left[\frac{x!}{(1+\theta)^{x+1}} + \frac{(\beta + \beta\theta - \theta)(x+1)x!}{(1+\theta)^{x+2}} \right] \\
&= \frac{\theta^3}{(2\beta(1+\theta)-\theta)} \left[\frac{x(1+\theta)}{(1+\theta)^{x+2}} + \frac{(\beta + \beta\theta - \theta)(x+1)x}{(1+\theta)^{x+2}} \right] \\
&= \frac{\theta^3}{(2\beta(1+\theta)-\theta)} \left[\frac{x((\beta + \beta\theta - \theta)x + \beta(1+\theta) + 1)}{(1+\theta)^{x+2}} \right]
\end{aligned}$$

Where $x=1,2,3,\dots$; $\theta > 0$; $\beta > \frac{\theta}{1+\theta}$.

3. Moments and Moments Based Measures

The r^{th} factorial moment about origin of the SBPGLD (2.1) can be obtained as

$$\begin{aligned}
\mu'_{(r)} &= E(E(X^r|\lambda)) = \int_0^{\infty} \left[\sum_{x=1}^{\infty} x^r e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} \right] \cdot h(\lambda; \theta, \beta) d\lambda \\
&= \int_0^{\infty} \left[\sum_{x=1}^{\infty} x^r e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} \right] \cdot \frac{\theta^3 \lambda((\beta + \beta\theta - \theta)\lambda + 1) e^{-\theta\lambda}}{2\beta(1+\theta) - \theta} d\lambda \\
&= \frac{\theta^3}{2\beta(1+\theta) - \theta} \int_0^{\infty} \lambda^{r-1} \left[\sum_{x=r}^{\infty} x e^{-\lambda} \frac{\lambda^{x-r}}{(x-r)!} \right] \cdot \lambda((\beta + \beta\theta - \theta)\lambda + 1) e^{-\theta\lambda} d\lambda \\
&= \frac{\theta^3}{2\beta(1+\theta) - \theta} \int_0^{\infty} \lambda^r \left[\sum_{x=r}^{\infty} x e^{-\lambda} \frac{\lambda^{x-r}}{(x-r)!} \right] \cdot ((\beta + \beta\theta - \theta)\lambda + 1) e^{-\theta\lambda} d\lambda
\end{aligned}$$

Taking $y = x - r$, we get

$$\begin{aligned}
 \mu_{(r)}' &= \frac{\theta^3}{2\beta(1+\theta)-\theta} \int_0^\infty \lambda^r \left[\sum_{y=0}^\infty (y+r)e^{-\lambda} \frac{\lambda^y}{y!} \right] \cdot ((\beta + \beta\theta - \theta)\lambda + 1)e^{-\theta\lambda} d\lambda \\
 &= \frac{\theta^3}{2\beta(1+\theta)-\theta} \int_0^\infty \lambda^r (\lambda + r) \cdot ((\beta + \beta\theta - \theta)\lambda + 1)e^{-\theta\lambda} d\lambda \\
 &= \frac{\theta^3}{2\beta(1+\theta)-\theta} \int_0^\infty \lambda^r ((\beta + \beta\theta - \theta)\lambda^2 + (r(\beta + \beta\theta - \theta) + 1)\lambda + r)e^{-\theta\lambda} d\lambda \\
 &= \frac{\theta^3}{2\beta(1+\theta)-\theta} \left[(\beta + \beta\theta - \theta) \int_0^\infty \lambda^{r+2} e^{-\theta\lambda} d\lambda + (r(\beta + \beta\theta - \theta) + 1) \int_0^\infty \lambda^{r+1} e^{-\theta\lambda} d\lambda + r \int_0^\infty \lambda e^{-\theta\lambda} d\lambda \right] \\
 &= \frac{\theta^3}{2\beta(1+\theta)-\theta} \left[(\beta + \beta\theta - \theta) \frac{r!(r+1)(r+2)}{\theta^{r+3}} + (r(\beta + \beta\theta - \theta) + 1) \frac{r!(r+1)}{\theta^{r+2}} + r \frac{r!}{\theta^{r+1}} \right] \\
 &= \frac{r! [(\beta + \beta\theta - \theta)(r+1)(r+2) + (r(\beta + \beta\theta - \theta) + 1)\theta(r+1) + r\theta^2]}{(2\beta(1+\theta) - \theta)\theta^r}
 \end{aligned}$$

The first four moments about origin of SBPGD can be obtained as

$$\begin{aligned}
 \mu_{(1)}' &= \frac{2\theta^2\beta - \theta^2 + 8\theta\beta - 4\theta + 6\beta}{(2\beta(1+\theta) - \theta)\theta} = \frac{\theta^2(2\beta - 1) + 4\theta(2\beta - 1) + 6\beta}{(2\beta(1+\theta) - \theta)\theta} \\
 \mu_{(2)}' &= \frac{12\theta^2\beta - 8\theta^2 + 36\theta\beta - 18\theta + 24\beta}{(2\beta(1+\theta) - \theta)\theta^2} = \frac{4\theta^2(3\beta - 2) + 18\theta(2\beta - 1) + 24\beta}{(2\beta(1+\theta) - \theta)\theta^2} \\
 \mu_{(3)}' &= \frac{72\theta^2\beta - 54\theta^2 + 192\theta\beta - 96\theta + 120\beta}{(2\beta(1+\theta) - \theta)\theta^3} = \frac{18\theta^2(4\beta - 3) + 96\theta(2\beta - 1) + 120\beta}{(2\beta(1+\theta) - \theta)\theta^3} \\
 \mu_{(4)}' &= \frac{480\theta^2\beta - 384\theta^2 + 1200\theta\beta - 600\theta + 720\beta}{(2\beta(1+\theta) - \theta)\theta^4} = \frac{96\theta^2(5\beta - 4) + 600\theta(2\beta - 1) + 720\beta}{(2\beta(1+\theta) - \theta)\theta^4}
 \end{aligned}$$

Now using the relationship between factorial moments about origin and the moments about origin, the first four moments about origin of SBPGLD (2.1) can be obtained as

$$\begin{aligned}
 \mu_1' &= \frac{\theta^2(2\beta - 1) + 4\theta(2\beta - 1) + 6\beta}{\theta(2\beta(1+\theta) - \theta)} \\
 \mu_2' &= \frac{\theta^3(2\beta - 1) + 4\theta^2(5\beta - 3) + 6\theta(7\beta - 3) + 24\beta}{\theta^2(2\beta(1+\theta) - \theta)} \\
 \mu_3' &= \frac{\theta^4(2\beta - 1) + 4\theta^3(11\beta - 7) + 6\theta^2(31\beta - 18) + 24\theta(11\beta - 4) + 120\beta}{\theta^3(2\beta(1+\theta) - \theta)} \\
 \mu_4' &= \frac{\theta^5(2\beta - 1) + 4\theta^4(23\beta - 15) + 30\theta^3(23\beta - 15) + 120\theta^2(15\beta - 8) + 120\theta(16\beta - 5) + 720\beta}{\theta^4(2\beta(1+\theta) - \theta)}
 \end{aligned}$$

Now, using the relationship between central moments and the moments about origin, the central moments of the SBPGLD (2.1) are thus obtained as

$$\mu_2 = \frac{2\theta^3(6\beta^2 - 7\beta + 2) + 2\theta^2(18\beta^2 - 13\beta + 1) + 12\theta(3\beta^2 - \beta) + 12\beta^2}{\theta^2(2\beta(1+\theta) - \theta)^2}$$

$$\mu_3 = \frac{\left[2\theta^5(12\beta^3 - 20\beta^2 + 11\beta - 2) + 2\theta^4(72\beta^3 - 94\beta^2 + 35\beta - 3) + 4\theta^3(84\beta^3 - 82\beta^2 + 21\beta - 2) \right. \\ \left. + 4\theta^2(96\beta^3 - 63\beta^2 + 9\beta) + 72\theta(3\beta^3 - \beta^2) + 48\beta^3 \right]}{\theta^3(2\beta(1+\theta) - \theta)^3}$$

$$\mu_4 = \frac{\left[2\theta^7(24\beta^4 - 52\beta^3 + 41\beta^2 + 15\beta + 2) + 2\theta^6(9898\beta^4 - 996\beta^3 + 760\beta^2 - 253\beta + 31) \right. \\ \left. + 4\theta^5(1200\beta^4 - 2094\beta^3 + 1273\beta^2 - 311\beta + 24) + 8\theta^4(1410\beta^4 - 1957\beta^3 + 871\beta^2 - 132\beta + 3) \right. \\ \left. + 16\theta^3(905\beta^4 - 932\beta^3 + 270\beta^2 - 18\beta) + 16\theta^2(678\beta^4 - 449\beta^3 + 63\beta^2) \right. \\ \left. + 32\theta(130\beta^4 - 45\beta^3) + 720\beta^4 \right]}{\theta^4(2\beta(1+\theta) - \theta)^4}$$

The coefficient of variation ($C.V$), coefficient of Skewness ($\sqrt{\beta_1}$), coefficient of Kurtosis ($\sqrt{\beta_2}$) and index of dispersion (γ) of the SBPGLD (2.1) are thus obtained as

$$C.V = \frac{\sqrt{2\theta^3(6\beta^2 - 7\beta + 2) + 2\theta^2(18\beta^2 - 13\beta + 1) + 12\theta(3\beta^2 - \beta) + 12\beta^2}}{\theta^2(2\beta - 1) + 4\theta(2\beta - 1) + 6\beta}$$

$$\sqrt{\beta_1} = \frac{\left[2\theta^5(12\beta^3 - 20\beta^2 + 11\beta - 2) + 2\theta^4(72\beta^3 - 94\beta^2 + 35\beta - 3) \right. \\ \left. + 4\theta^3(84\beta^3 - 82\beta^2 + 21\beta - 2) + 4\theta^2(96\beta^3 - 63\beta^2 + 9\beta) + 72\theta(3\beta^3 - \beta^2) + 48\beta^3 \right]}{\left(2\theta^3(6\beta^2 - 7\beta + 2) + 2\theta^2(18\beta^2 - 13\beta + 1) + 12\theta(3\beta^2 - \beta) + 12\beta^2 \right)^{\frac{3}{2}}}$$

$$\sqrt{\beta_2} = \frac{\left[2\theta^7(24\beta^4 - 52\beta^3 + 41\beta^2 + 15\beta + 2) + 2\theta^6(9898\beta^4 - 996\beta^3 + 760\beta^2 - 253\beta + 31) \right. \\ \left. + 4\theta^5(1200\beta^4 - 2094\beta^3 + 1273\beta^2 - 311\beta + 24) + 8\theta^4(1410\beta^4 - 1957\beta^3 + 871\beta^2 - 132\beta + 3) \right. \\ \left. + 16\theta^3(905\beta^4 - 932\beta^3 + 270\beta^2 - 18\beta) + 16\theta^2(678\beta^4 - 449\beta^3 + 63\beta^2) \right. \\ \left. + 32\theta(130\beta^4 - 45\beta^3) + 720\beta^4 \right]}{\left(2\theta^3(6\beta^2 - 7\beta + 2) + 2\theta^2(18\beta^2 - 13\beta + 1) + 12\theta(3\beta^2 - \beta) + 12\beta^2 \right)^2}$$

$$\gamma = \frac{2\theta^3(6\beta^2 - 7\beta + 2) + 2\theta^2(18\beta^2 - 13\beta + 1) + 12\theta(3\beta^2 - \beta) + 12\beta^2}{(\theta(2\beta(1+\theta) - \theta))(\theta^2(2\beta - 1) + 4\theta(2\beta - 1) + 6\beta)}$$

4. Statistical Properties of SBPSD

4.1. Reliability Properties

We now give some basic properties of the SBPGLD model

(i) Since

$$\mu - \sigma^2 = \frac{\theta^4(4\beta^2 - 4\beta + 1) + 4\theta^3(2\beta^2 - \beta) + 2\theta^2(-4\beta^2 + 6\beta - 1) + 12\theta(-2\beta^2 + \beta) + 12\beta^2}{\theta^2(2\beta(1+\theta) - \theta)^2}$$

(ii) Since

$$\frac{P_1(x+1; \theta, \beta)}{P_1(x; \theta, \beta)} = \left(\frac{1}{1+\theta}\right) \left(1 + \frac{1}{x}\right) \left[1 + \frac{(\beta + \beta\theta - \theta)}{(\beta + \beta\theta - \theta)x + \beta(1+\theta) + 1}\right]$$

is a decreasing function of x , $P_1(x; \theta, \beta)$ is log-concave. Therefore, SBPGLD is unimodal, has an increasing failure rate (IFR), and hence increasing failure rate average (IFRA). It is new better than used in expectation (NBUE) and has decreasing mean residual life (DMRL).

4.2. Generating Function

Probability Generating Function: The probability generating function of the SBPGLD

(2.1) can be obtained as

$$\begin{aligned} P_X(t) &= E(t^X) = \sum_{x=1}^{\infty} t^x P_1(x; \theta, \beta) \\ &= \frac{\theta^3}{(2\beta(1+\theta) - \theta)(1+\theta)^2} \left[(\beta + \beta\theta - \theta) \sum_{x=1}^{\infty} x^2 \left(\frac{t}{1+\theta}\right)^x + (\beta(1+\theta) + 1) \sum_{x=1}^{\infty} x \left(\frac{t}{1+\theta}\right)^x \right] \\ &= \frac{\theta^3}{(2\beta(1+\theta) - \theta)(1+\theta)^2} \left[\frac{(\beta + \beta\theta - \theta)t(\theta + 1 + t)(1+\theta)}{(\theta + 1 - t)^3} + \frac{(\beta(1+\theta) + 1)t(1+\theta)}{(\theta + 1 - t)^2} \right] \\ &= \frac{\theta^3 t}{(2\beta(1+\theta) - \theta)(1+\theta)} \left[\frac{(\beta + \beta\theta - \theta)(\theta + 1 + t)}{(\theta + 1 - t)^3} + \frac{(\beta(1+\theta) + 1)}{(\theta + 1 - t)^2} \right] \end{aligned}$$

So

$$E(t^x) = \frac{\theta^3 t}{(2\beta(1+\theta) - \theta)(1+\theta)} \left[\frac{(\beta + \beta\theta - \theta)(\theta + 1 + t)}{(\theta + 1 - t)^3} + \frac{(\beta(1+\theta) + 1)}{(\theta + 1 - t)^2} \right] \quad (4.1)$$

Moment Generating Function: The moment generating function of the SBPGLD (2.1) is given by

$$M_x(t) = E(e^{tx}) = \frac{\theta^3 e^t}{(2\beta(1+\theta) - \theta)(1+\theta)} \left[\frac{(\beta + \beta\theta - \theta)(\theta + 1 + e^t)}{(\theta + 1 - e^t)^3} + \frac{(\beta(1+\theta) + 1)}{(\theta + 1 - e^t)^2} \right] \quad (4.2)$$

4.3. Maximum Likelihood Estimate (MLE)

Given a random sample x_1, x_2, \dots, x_n of size n from the SBPGLD distribution with p.m.f. (2.1) is,

$$P(x; \theta, \beta) = \frac{\theta^3}{-\theta + 2\beta(1+\theta)} \frac{x((\beta + \beta\theta - \theta)x + \beta(1+\theta) + 1)}{(1+\theta)^{x+2}}$$

Where $x = 1, 2, 3, \dots$; $\theta > 0$; $\beta > \frac{\theta}{1+\theta}$

The likelihood function will be

$$L(x_i; \theta, \beta) = \left(\frac{\theta^3}{-\theta + 2\beta(1+\theta)} \cdot \frac{1}{(1+\theta)^2} \right)^n \prod_{i=1}^n \frac{1}{(1+\theta)^{x_i}} \cdot x_i ((\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1)$$

Taking log both sides

$$\begin{aligned} \log L &= 3n \log(\theta) - n \log(2\beta(1+\theta) - \theta) - 2n \log(1+\theta) \\ &\quad + \sum_{i=1}^n \log x_i + \sum_{i=1}^n \log((\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1) - \sum_{i=1}^n x_i \log(1+\theta) \end{aligned}$$

$$\frac{\partial \log L}{\partial \beta} = -\frac{2n(1+\theta)}{2\beta(1+\theta) - \theta} + \sum_{i=1}^n \frac{(1+\theta)x_i + (1+\theta)}{(\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1} = 0 \quad (4.3)$$

$$\begin{aligned}
\frac{\partial \log L}{\partial \theta} &= \frac{3n}{\theta} + \frac{n(2\beta-1)}{2\beta(1+\theta)-\theta} - \frac{2n}{(1+\theta)} + \sum_{i=1}^n \frac{(\beta-1)x_i + \beta}{(\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1} - \sum_{i=1}^n x_i \frac{1}{(1+\theta)} \\
&= \frac{3n}{\theta} + \frac{n(2\beta-1)}{2\beta(1+\theta)-\theta} - \frac{2n}{(1+\theta)} - \frac{n\bar{X}}{(1+\theta)} + \sum_{i=1}^n \frac{(\beta-1)x_i + \beta}{(\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1} \\
&= \frac{3n}{\theta} + \frac{n(2\beta-1)}{2\beta(1+\theta)-\theta} - \frac{n(2+\bar{X})}{(1+\theta)} + \sum_{i=1}^n \frac{(\beta-1)x_i + \beta}{(\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1}
\end{aligned}$$

Or

$$\frac{\partial \log L}{\partial \theta} = \frac{3n}{\theta} + \frac{n(2\beta-1)}{2\beta(1+\theta)-\theta} - \frac{n(2+\bar{X})}{(1+\theta)} + \sum_{i=1}^n \frac{(\beta-1)x_i + \beta}{(\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1} \quad (4.4)$$

where \bar{X} is the sample mean.

The maximum likelihood estimate (MLE), $\hat{\theta}$ of θ is the solution of the equation $\frac{\partial \log L}{\partial \theta}$

and is given by the solution of the non-linear equation.

The two equations (4.3) and (4.4) may be solved by appropriate numerical methods.

However, the Fisher Scoring Method can be applied to solve these equations. We have

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{3n}{\theta^2} - \frac{n(2\beta-1)^2}{(2\beta(1+\theta)-\theta)^2} + \frac{n(2+\bar{X})}{(1+\theta)^2} - \sum_{i=1}^n \frac{((\beta-1)x_i + \beta)^2}{((\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1)^2} \quad (4.5)$$

$$\frac{\partial^2 \log L}{\partial \beta^2} = \frac{4n(1+\theta)^2}{(2\beta(1+\theta)-\theta)^2} - \sum_{i=1}^n \frac{((1+\theta)x_i + (1+\theta))^2}{((\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1)^2} \quad (4.6)$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \beta} = \frac{-2n}{(2\beta(1+\theta)-\theta)^2} + \sum_{i=1}^n \frac{(x_i + 1)^2}{((\beta + \beta\theta - \theta)x_i + \beta + \beta\theta + 1)^2} \quad (4.7)$$

The following equations for $\hat{\theta}$ and $\hat{\beta}$ can be solved

$$\begin{bmatrix} \frac{\partial^2 \log L}{\partial \theta^2} & \frac{\partial^2 \log L}{\partial \theta \partial \beta} \\ \frac{\partial^2 \log L}{\partial \beta \partial \theta} & \frac{\partial^2 \log L}{\partial \beta^2} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\beta}=\beta_0}} \begin{bmatrix} \hat{\theta} - \theta_0 \\ \hat{\beta} - \beta_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \log L}{\partial \theta} \\ \frac{\partial \log L}{\partial \beta} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\beta}=\beta_0}}$$

Where θ_0 and β_0 are the initial values of θ and β respectively. These equations are solved iteratively till sufficiently closed estimates of $\hat{\theta}$ and $\hat{\beta}$ are obtained.

4.4. Estimates from Moments

Using the first moment μ_1' and second moment μ_2' about SBPGLD, we have

$$\begin{cases} E(X) = \mu_1' = \frac{\theta^2(2\beta-1) + 4\theta(2\beta-1) + 6\beta}{\theta(2\beta(1+\theta) - \theta)} \\ E(X^2) = \mu_2' = \frac{\theta^3(2\beta-1) + 4\theta^2(5\beta-3) + 6\theta(7\beta-3) + 24\beta}{\theta^2(2\beta(1+\theta) - \theta)} \end{cases} \quad (4.8)$$

where $\mu_2' = s^2 + \mu_1'^2$ and s^2 is the variance. By solving this non linear system, we find the couple $(\hat{\theta}, \hat{\beta})$ where $(\hat{\theta}, \hat{\beta}) > 0$ for all $\mu_1' > 0$, $s > 0$.

$$\begin{aligned} \mu_1' &= \frac{\theta^2(2\beta-1) + 4\theta(2\beta-1) + 6\beta}{\theta(2\beta(1+\theta) - \theta)} \\ \theta^2(2\beta-1) + 4\theta(2\beta-1) + 6\beta &= 2\beta(1+\theta)\theta\mu_1' - \theta\mu_1' \\ 2\beta\theta(1+\theta)\mu_1' - 2\beta(\theta^2 + 4\theta + 3) &= \theta\mu_1' - \theta^2 + 4\theta \\ \beta(2\theta(1+\theta)\mu_1' - 2\theta^2 - 8\theta - 6) &= \theta\mu_1' - \theta^2 + 4\theta \\ \beta &= \frac{\theta^2\mu_1' - \theta^2 + 4\theta}{2\mu_1'(\theta^2 + \theta) - 2\theta^2 - 8\theta - 6} \end{aligned}$$

So

$$\hat{\beta} = \frac{\theta^2\mu_1' - \theta^2 + 4\theta}{2\mu_1'(\theta^2 + \theta) - 2\theta^2 - 8\theta - 6}$$

Note that $\mu_1' = \bar{X} = 1$ if and only if $x_i = 1$ for all $i = 1, 2, \dots, n$. A data set where all observations are ones is not worth analyzing. This situation, of course, will not lead to any estimate of θ . However, such situation may arise in a simulation experiment when n is small. For this reason, we will assume throughout this paper that $\bar{X} > 1$.

5. Simulation

This section present a simulation study is carried out for each triplet (n, θ, β) , to examine the bias and mean square error of the maximum likelihood estimators for a finite sample size n .

In this simulation, we can choose $\theta = 0.1, 0.5, 5, \beta = 0.2, 0.5, 2$, and $n = 20, 50, 100$.

The plan of this simulation as follows:

- choose the initial values of θ_0, β_0 for the corresponding elements of the parameter vector $\Theta = (\theta, \beta)$ to specify SBPGaL distribution;
- choose sample size n ;
- generate N independent samples of size n from SBPGaL(θ, β);
- compute the ML estimate Θ_n of Θ_0 for each of the N samples;
- compute the mean of the obtained estimators over all N samples,

$$\text{Average bias}(\theta) = \frac{1}{N} \sum_{i=1}^N (\Theta_i - \Theta_0)^2 \text{ and the average square error } \text{MSE}(\theta) = \frac{1}{N} \sum_{i=1}^N (\Theta_i - \Theta_0)^2.$$

Table 1 Average bias of the simulated estimates

	$\theta = 0.1, \beta = 0.2$		$\theta = 0.1, \beta = 0.5$		$\theta = 0.1, \beta = 2$	
	bias(θ)	bias(β)	bias(θ)	bias(β)	bias(θ)	bias(β)
$n=20$	0.002	0.085	0.086	0.0086	0.025	-0.014
$n=50$	0.0006	0.042	2.5×10^{-2}	0.001	1.6×10^{-2}	-0.001
$n=100$	0.0001	0.0071	4×10^{-3}	0.0006	0.01	-0.0008
	$\theta = 0.1, \beta = 2$		$\theta = 0.5, \beta = 2$		$\theta = 5, \beta = 2$	
	bias(θ)	bias(β)	bias(θ)	bias(β)	bias(θ)	bias(β)
$n=20$	2×10^{-4}	-0.005	10^{-2}	-0.03	0.092	-0.056
$n=50$	0.6×10^{-4}	-0.01	2×10^{-3}	-0.013	0.022	-0.01
$n=100$	5×10^{-5}	-0.006	10^{-4}	-0.0064	0.007	-0.008
	$\theta = 5, \beta = 0.2$		$\theta = 0.5, \beta = 0.5$		$\theta = 0.1, \beta = 0.5$	
	bias(θ)	bias(β)	bias(θ)	bias(β)	bias(θ)	bias(β)
$n=20$	0.45	-0.426	0.015	-0.0049	0.0005	-0.024
$n=50$	0.12	-0.122	3.3×10^{-3}	-0.0029	1.2×10^{-4}	-0.073
$n=100$	0.05	-0.067	0.0001	-0.0004	0.00002	-0.0005

Table 2 Average MSE of the simulated estimates

	$\theta = 0.1, \beta = 0.2$		$\theta = 0.1, \beta = 0.5$		$\theta = 0.1, \beta = 2$	
	MSE(θ)	MSE(β)	MSE(θ)	MSE(β)	MSE(θ)	MSE(β)
$n=20$	0.8×10^{-4}	1.76	0.14	0.0011	0.017	0.00252
$n=50$	2.1×10^{-5}	0.23	0.03	2.7×10^{-4}	2.1×10^{-3}	0.0013
$n=100$	4×10^{-6}	0.03	0.002	0.00001	0.0003	0.0006
	$\theta = 0.1, \beta = 2$		$\theta = 0.5, \beta = 2$		$\theta = 5, \beta = 2$	
	MSE(θ)	MSE(β)	MSE(θ)	MSE(β)	MSE(θ)	MSE(β)
$n=20$	1.2×10^{-6}	0.046	0.0005	0.024	0.142	0.035
$n=50$	2.8×10^{-7}	0.012	1.1×10^{-4}	0.0075	0.051	0.015
$n=100$	10^{-8}	0.006	0.00002	0.0003	0.008	0.005
	$\theta = 5, \beta = 0.2$		$\theta = 0.5, \beta = 0.5$		$\theta = 0.1, \beta = 0.5$	
	MSE(θ)	MSE(β)	MSE(θ)	MSE(β)	MSE(θ)	MSE(β)
$n=20$	3.0	3.6	1.1×10^{-3}	0.001	1×10^{-5}	0.08
$n=50$	0.5	0.7	0.4×10^{-3}	0.0003	3×10^{-6}	0.002
$n=100$	0.06	0.08	0.00006	0.00004	2×10^{-6}	0.0004

Table 1 shows how the two biases vary with respect to n . Table 2 shows how the mean squared errors vary with respect to n . The mean squared errors for each parameter decrease to zero as $n \rightarrow \infty$. Also, the result of table 1 confirm that θ is positively biased and the bias $(\theta) \rightarrow 0, \theta \rightarrow 0$, and β is negatively biased if $\beta > 0.5$ and the bias $(\beta) \rightarrow 0, \beta \rightarrow 0$ and $\theta \rightarrow 0$. Moreover, table 2 shows that $MSE(\theta)$ and $MSE(\beta) \rightarrow 0$ where $\theta \rightarrow 0$ and $n \rightarrow \infty$.

6. Goodness of fit: Application to real data set

In this section, we illustrate the applicability of SBPGL distribution by considering a real data set on **Nipah virus infection** (*world health organization source*).

Table 3 represents the data of survival times (in months) of (56 Indian state of Kerala individus infected with Nipah virus in 2017, which we compare Poisson distribution, Poisson Quasi Lindley distribution (PQLD) and SBPGLD.

According to Table 3, we can observe that size-biased gamma-Lindley distribution provide smallest χ^2 values as compared others distributions and hence best fits the data among all the models considered.

Table 3 Comparison between distributions

<i>Number of attached particles</i>	Obs freq.	Poisson	PQL	SBPGLD
0	33	26.45	31.85	32.56
1	12	19.84	13.83	13.522
2	6	7.44	5.94	5.84
3	3	1.86	2.53	2.51
4	1	0.35	1.07	1.08
5	1	0.05	0.45	0.47
Total	56	56	56	56
		$\theta = 0.75$	$\theta = 1.53$ $\beta = 5.93$	$\theta = 1.53,$ $\beta = 0.75$
χ^2		24.1	1.057	0.756

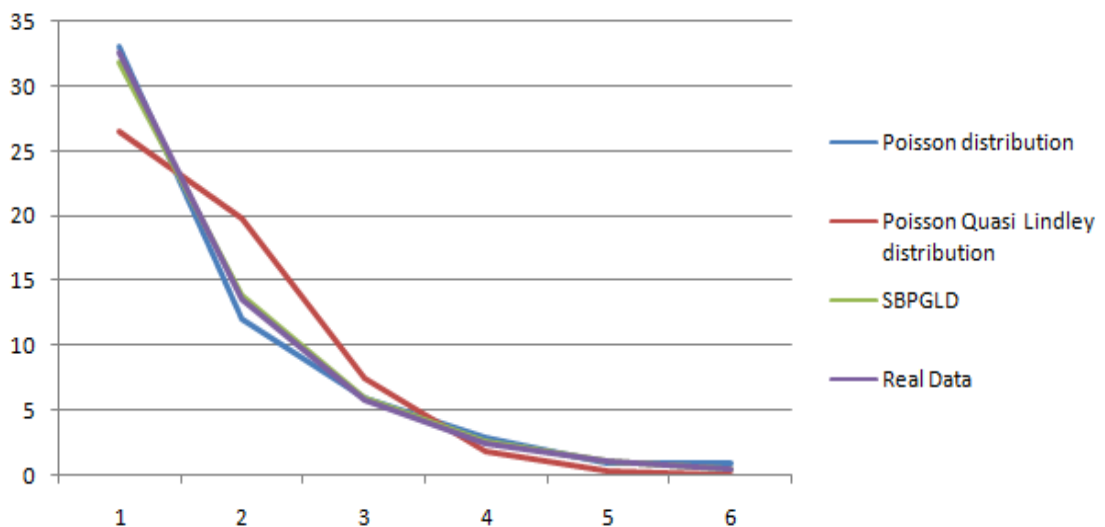


Figure 2 Density plots of Poisson, PQL, SBPGL distributions and real data

7. Conclusion

The size-biased Poisson-Gamma Lindley distribution (SBPGLD), of which the size-biased Poisson Lindley distribution (SBPLD) is a particular case, has been introduced. The estimation of its parameters has been discussed using the method of maximum likelihood and the method of moments. A simulation study is carried out to examine the bias and mean square error of the maximum likelihood estimators of the parameters. An application of the model to a real data set are presented (data of survival times (in months) of (56) Indian state of Kerala individuals infected with Nipah virus). Also, the comparison fit study with other well-known one and two parameters. The adequacy of fits was assessed in term χ^2 value and density plots. We can show that the size-biased Poisson gamma-Lindley distribution can be used quite effectively in analyzing real lifetime data and actuarial science.

For future studies, we can discuss the collective risk model by considering the proposed distribution (SBPGLD) as primary distribution and exponential and Erlang distributions as secondary ones.

Acknowledgment

The authors are grateful for the comments and suggestions by the referee and the Editor. Their comments and suggestions greatly improved the article.

REFERENCES

- [1] S.M.R. Alavi , R. Chinipardaz, Form-invariance under weighted sampling, *Statistics* 43 (1) (2009) 81-90. <https://doi.org/10.1080/02331880802172826>.
- [2] A. Ayesha, Size biased Lindley distribution and its properties a special case of weighted distribution. *Appl. Math.* 8(6) (2017) 808-819. <https://doi.org/10.4236/am.2017.86063>.
- [3] A. Beghriche, H. Zeghdoudi, A Size Biased Gamma Lindley Distribution. *Thailand Statistician*; 17(2), (2019) 179-189. <https://ph02.tci-thaijo.org/index.php/thaistat/article/view/202288/141165>
- [4] M.J. Ducey, JH.Gove, Size-biased distributions in the generalized beta distribution family with applications to forestry. *Int. J. Forest Res.* 88 (2015) 143-151. <https://doi.org/10.1093/forestry/cpu038>
- [5] R.A. Fisher, The effects of methods of ascertainment upon the estimation of frequency. *Ann. Eugenics.* 6(1) (1937) 13-25. <https://doi.org/10.1111/j.1469-1809.1934.tb02105.x>

- [6] M.E. Ghitany, D.K. Al-Mutairi, Size-biased Poisson-Lindley distribution and Its applications, *Metron – Int. J. Stat.* LXVI (3), (2008) 299-311. <https://EconPapers.repec.org/RePEc:mtn:ancoec:080303>
- [7] R. Grine, H. Zeghdoudi, On Poisson Quasi-Lindley Distribution and its Applications. *J. Mod. Appl. Stat. Meth.* 16(2) (2017) 403-417. <https://doi.org/10.22237/jmasm/1509495660>.
- [8] J. Lappi, R.L. Bailey, Estimation of the Diameter Increment Function or Other Tree Relations Using Angle-Count Samples. *Forest Sci.* 33(3) (1987) 725-739. <https://doi.org/10.1093/forestscience/33.3.725>
- [9] G.P. Patil, C.R. Rao, Weighted distributions and size-biased sampling with applications to wild life populations and human families. *Biometrics* 34(2) (1978) 179-189. <https://doi.org/10.2307/2530008>
- [10] R. Shanker, A Size-Biased Poisson-Amarendra Distribution and Its Applications, *Int. J. Stat. Appl.* 6(6) (2016) 376-385. <https://doi.org/10.5923/j.statistics.20160606.06>
- [11] R. Shanker, A Size-Biased Poisson-Shanker Distribution and Its Applications, *Int. J. Probab. Stat.* 6(3) (2017) 33-44. <https://doi.org/10.5923/j.ijps.20170603.01>
- [12] R. Shanker, K.K. Shukla, Size-biased poisson-garima distribution with applications. *Biom. Biostat. Int. J.* 6(3) (2017) 335-340. <https://doi.org/10.15406/bbij.2017.06.00167>
- [13] R. Shanker, A. Mishra, On Size- Biased Two Parameter Poisson-Lindley Distribution and Its Applications, *Amer. J. Math. Stat.* 7(3) (2017) 99-107. <https://doi.org/10.5923/j.ajms.20170703.02>
- [14] R. Shanker, K.K. Shukla, A generalized size-biased poisson-lindley distribution and its applications to model size distribution of freely-forming small group, *J. Sci. Res.* 10(2) (2018) 145-157. <http://doi.org/10.3329/jsr.v10i2.34905>
- [15] T.R. Adhikari, R.S. Srivastava, Size-biased discrete two parameter Poisson-Lindley Distribution for modeling and waiting survival times data, *IOSR J. Math.* 10 (2014) 39-45. <https://www.iosrjournals.org/iosr-jm/papers/Vol10-issue1/Version-3/F010133945.pdf>
- [16] J. Van Deusen, Occupational therapy theory: values held by undergraduate students. *Occupational Therapy J. Res.* 6 (1986) 115-122. <https://doi.org/10.1177/153944928600600205>
- [17] H. Zeghdoudi, S. Nedjar, Gamma Lindley distribution and its application. *J. Appl. Probab. Stat.* 11 (2016) 129-138. <https://doi.org/10.1504/IJCSM.2018.090722>
- [18] H. Zeghdoudi, S. Nedjar, On Gamma Lindley distribution: properties and simulation. *J. Comput. Appl. Math.* 298 (2016) 167-174. <https://doi.org/10.1016/j.cam.2015.11.047>
- [18] S. Nedjar, H. Zeghdoudi, New compound Poisson distribution: properties, inflated distribution and applications. *J. Agric. Stat. Sci.* 16 (2020) 519-526. <https://connectjournals.com/03899.2020.16.519>