

Exponential-Discrete Lindley Distribution: Properties and Applications

Sibel Açık Kemaloğlu* and Mehmet Yılmaz

Abstract

In this paper, we introduce a new compounding distribution named Exponential Discrete Lindley distribution which compounds the discrete Lindley distribution and exponential distribution. We obtain several properties of the new distribution such as its probability density function, survival function, hazard rate function, mean residual life function. Moments and expression for the Rényi entropy of the proposed distribution are also given. Moreover, the maximum likelihood method using the EM algorithm is developed for parameter estimation. Two real data sets are used to illustrate the usefulness of the proposed distribution.

Keywords: Discrete Lindley distribution; Compound distribution; Exponential distribution; Maximum likelihood estimation; EM-algorithm; Rényi entropy.

AMS Subject Classification (2020): Primary: 62E10 ; Secondary: 62F10; 62P99.

*Corresponding author

1. Introduction

In recent years, many distributions have been proposed to model lifetime data by several authors. [1], introduced a two-parameter exponential-geometric (E-G) distribution by compounding an exponential distribution with a geometric distribution. Similarly, [2] introduced an exponential-Poisson (E-P) distribution by compounding an exponential distribution with a Poisson distribution. [3] studied the exponential-logarithmic (EL) distribution. [4] introduced the family of exponential-power series distributions (EPS). [5] proposed the Weibull-geometric (WG) distribution. [6] studied a distribution family by mixing the Weibull and power-series distributions. [7] have considered the Lindley distribution which is composed of the mixture of the exponential and the gamma distributions. Moreover, [8] proposed a new lifetime distribution by compounding exponential and modified-zero-truncated discrete Lindley distributions.

The purpose of this paper is to introduce a new lifetime distribution with decreasing hazard rate by compounding exponential and discrete Lindley distributions. In this context, we propose and improve the statistical properties of Exponential Discrete Lindley (EDL) distribution and show that it is suitable to use this distribution for reliability analysis. Some statistical properties such as the shapes of the density function and hazard rate function, moments, moment generating function (MGF), and parameter estimation by using the maximum likelihood method and EM algorithm are given in the study. Finally, applications on real data sets are presented to show the feasibility and usefulness of the distribution.

This paper is organized as follows: In Section 2, the model and some of its statistical properties including moments, survival and hazard rate functions are presented. In Section 3, the estimation of parameters is studied via maximum likelihood and EM algorithm methods. Applications on real data sets are given in Section 4 and finally conclusions are given in Section 5.

2. Exponential Discrete Lindley Distribution (EDL) and Its Properties

In this section, we introduce pdf of EDL distribution, and investigate raw moments, survival, hazard rate and mean residual life functions, an algorithm to generate random number.

2.1 Distribution and moments

Let Y_1, Y_2, \dots, Y_M be random sample from an exponential distribution with probability density function

$$f_Y(y; \beta) = \beta e^{-\beta y}, \quad y > 0, \quad \beta > 0.$$

M is a random variable from zero truncated discrete Lindley distribution with probability function as follows:

$$P(M = m) = \frac{e^{-\theta m - 1}}{1 + 2\theta} [\theta (1 - 2e^{-\theta}) + (1 - e^{-\theta})(1 + \theta m)], \quad \theta > 0, \quad m = 1, 2, 3, \dots$$

Suppose that Y and M are independent and $X = \min \{Y_1, Y_2, \dots, Y_M\}$. Then the conditional density function of $X|M$ is

$$f(x|m; \beta) = \beta m e^{-\beta m x},$$

and the marginal probability density function of X is

$$f(x; \theta, \beta) = \beta e^{-\beta x} (1 - e^{-\theta}) \left[A_1(\theta) \frac{1}{(1 - e^{-\beta x - \theta})^2} + A_2(\theta) \frac{(1 - e^{-\theta})(1 + e^{-\beta x - \theta})}{(1 - e^{-\beta x - \theta})^3} \right],$$

where $A_1(\theta) = \frac{1 - e^{-\theta} + \theta(1 - 2e^{-\theta})}{(1 - e^{-\theta})(2\theta + 1)}$ and $A_2(\theta) = \frac{\theta}{(1 - e^{-\theta})(2\theta + 1)}$ with $A_1(\theta) + A_2(\theta) = 1$. This distribution defines as Exponential Discrete Lindley (EDL) Distribution.

Moment generation function of X can be obtained as follows:

$$M_X(t) = \beta (1 - e^{-\theta}) e^\theta \left[A_1(\theta) \sum_{j=1}^{\infty} \frac{j e^{-\theta j}}{(\beta j - t)} + A_2(\theta) (1 - e^{-\theta}) \sum_{j=1}^{\infty} \frac{j^2 e^{-\theta j}}{(\beta j - t)} \right].$$

Furthermore, first two raw moments and k . th raw moment is calculated as follows:

$$\begin{aligned} E(X) &= (1 - e^{-\theta}) \beta^{-1} [A_2(\theta) - A_1(\theta) e^\theta \ln(1 - e^{-\theta})], \\ E(X^2) &= 2(1 - e^{-\theta}) e^\theta \beta^{-2} [A_1(\theta) Li_2(e^{-\theta}) - A_2(\theta) (1 - e^{-\theta}) \ln(1 - e^{-\theta})], \\ E(X^k) &= k! (1 - e^{-\theta}) e^\theta \beta^{-k} [A_1(\theta) Li_k(e^{-\theta}) + A_2(\theta) (1 - e^{-\theta}) Li_{k-1}(e^{-\theta})]. \end{aligned}$$

where $Li_v(z)$ denotes *Polylogarithm* function defined as $\sum_{j=1}^{\infty} \frac{z^j}{j^v}$ with $|z| < 1$.

2.2 The survival, hazard rate and mean residual life functions

The survival function of the EDL distribution is given by

$$S(x) = A_1(\theta) (1 - e^{-\theta}) \frac{e^{-\beta x}}{(1 - e^{-\beta x - \theta})} + A_2(\theta) (1 - e^{-\theta})^2 \frac{e^{-\beta x}}{(1 - e^{-\beta x - \theta})^2}.$$

The hazard rate function of EDL distribution is given as

$$h(x) = \frac{\beta}{1 - e^{-\beta x - \theta}} \left[1 + \frac{A_2(\theta)(1 - e^{-\theta})(e^{-\beta x - \theta})}{A_1(\theta)(1 - e^{-\beta x - \theta}) + A_2(\theta)(1 - e^{-\theta})} \right].$$

Let $r = e^{-\beta x - \theta}$, then we can rewrite the hazard rate function as

$$h(r) = \frac{\beta}{1 - r} \left[1 + \frac{r}{\frac{A_1(\theta)}{A_2(\theta)(1 - e^{-\theta})} (1 - r) + 1} \right].$$

This expression is an increasing function of r . Then we can say that hazard rate function is decreasing function of x , and also the following expressions can be written:

$$\lim_{x \rightarrow 0} h(x) = \beta \frac{1+A_2(\theta)e^{-\theta}}{(1-e^{-\theta})} \geq \beta \text{ and } \lim_{x \rightarrow \infty} h(x) = \beta.$$

The mean residual life function of EDL distribution is given as follows:

$$m(x) = \frac{(1 - e^{-\beta x - \theta}) [A_2(\theta) (1 - e^{-\theta}) (e^{-\beta x - \theta}) - A_1(\theta) (1 - e^{-\beta x - \theta}) \ln(1 - e^{-\beta x - \theta})]}{\beta (e^{-\beta x - \theta}) [A_1(\theta) (1 - e^{-\beta x - \theta}) + A_2(\theta) (1 - e^{-\theta})]}.$$

2.3 Random number generation

An algorithm is given to generate random numbers from EDL distribution.

- Step0. Generate $u \sim U(0, 1)$.
- Step1. Input the values of parameters θ and β .
- Step2. If $u < \frac{1-e^{-\theta} + \theta(1-2e^{-\theta})}{(1-e^{-\theta})(2\theta+1)}$ then go to Step 3 else go to Step 4.
- Step3. $x = \frac{-\ln u + \ln(u+e^\theta-1) - \theta}{\beta}$.
- Step4. $x = \frac{1}{\beta} \left[\ln \left(1 + \frac{1}{2u} (1 - e^{-\theta})^2 u - \frac{(1-e^{-\theta})e^{\theta/2}}{u} \sqrt{u + \frac{1}{4}(1 - e^{-\theta})^2 e^\theta} \right) + \theta \right]$.

2.4 Relationship Pareto Distribution

Under the transformation $Y = \frac{e^{\beta x} - e^{-\theta}}{1 - e^{-\theta}}$, the survival function can be written as

$$S(y) = \left(1 - A_2(\theta)e^{-\theta}\right) \frac{1}{y} + A_2(\theta)e^{-\theta} \frac{1}{y^2}, \quad y \geq 1$$

The life function represents a mixture of Pareto Type I distributions with $\alpha = 1$ and 2 , respectively, with ratios of $1 - A_2(\theta)e^{-\theta}$ and $A_2(\theta)e^{-\theta}$.

2.5 Rényi Entropy

The Rényi entropy is defined by

$$I_R(\gamma) = \frac{1}{1-\gamma} \log \int_0^\infty (f(x; \theta, \beta))^\gamma dx, \quad \gamma > 0 \text{ and } \gamma \neq 1.$$

By taking $\left[A_1(\theta) + A_2(\theta)(1 - e^{-\theta}) \frac{(1+e^{-\beta x - \theta})}{(1-e^{-\beta x - \theta})} \right] = z$, above integral can be written as

$$\begin{aligned} \int_0^\infty (f(x; \theta, \beta))^\gamma dx &= \int_{1-A_2(\theta)e^{-\theta}}^{1+A_2(\theta)e^{-\theta}} \left(\frac{e^\theta}{2A_2(\theta)} \right)^\gamma z^\gamma dz \\ &= e^{\theta\gamma} (2A_2(\theta))^{-\gamma} \left[\frac{(1 + A_2(\theta)e^{-\theta})^{\gamma+1} - (1 - A_2(\theta)e^{-\theta})^{\gamma+1}}{\gamma + 1} \right] \\ &= \frac{e^{\theta\gamma} (2A_2(\theta))^{-\gamma}}{\gamma + 1} (1 + A_2(\theta)e^{-\theta})^{\gamma+1} \left[1 - \left(\frac{1 - A_2(\theta)e^{-\theta}}{1 + A_2(\theta)e^{-\theta}} \right)^{\gamma+1} \right]. \end{aligned}$$

Then the entropy is obtained as follows:

$$I_R(\gamma) = \frac{1}{1-\gamma} \left[\frac{\theta\gamma - \log(\gamma + 1) - \gamma \log(2) - \gamma \log(A_2(\theta)) + (\gamma + 1) \log(1 + A_2(\theta)e^{-\theta}) + (\gamma + 1) \log \left(1 - \left(\frac{1 - A_2(\theta)e^{-\theta}}{1 + A_2(\theta)e^{-\theta}} \right)^{\gamma+1} \right)}{(\gamma + 1) \log \left(1 - \left(\frac{1 - A_2(\theta)e^{-\theta}}{1 + A_2(\theta)e^{-\theta}} \right)^{\gamma+1} \right)} \right].$$

3. Estimation of the Parameters

To estimate parameters of EDL distribution, we apply two techniques that are Maximum Likelihood and Expectation- Maximization (EM) algorithm.

3.1 Maximum Likelihood Estimation

Let X_1, X_2, \dots, X_n be a random sample from the EDL distribution with observed values $\underline{x} = x_1, x_2, \dots, x_n$ and (θ, β) be parameter vector. The log-likelihood function can be expressed

$$\begin{aligned} \log L(\theta, \beta | \underline{x}) = & 2n \log \beta - 2\beta \sum_{i=1}^n x_i + 2 \log(1 - e^{-\theta}) + n \log A_1(\theta) + n \log A_2(\theta) \\ & + n \log(1 - e^{-\theta}) + \sum_{i=1}^n \log(1 + e^{-\beta x_i - \theta}) - 5 \sum_{i=1}^n \log(1 - e^{-\beta x_i - \theta}). \end{aligned}$$

To obtain the maximum likelihood estimators, the derivative of the above statement based on parameters is taken to zero as below:

$$\begin{aligned} \frac{\partial \log L}{\partial \beta} &= \frac{2n}{\beta} - 2 \sum_{i=1}^n x_i - \sum_{i=1}^n \frac{x_i e^{-\beta x_i - \theta}}{1 + e^{-\beta x_i - \theta}} - 5 \sum_{i=1}^n \frac{x_i e^{-\beta x_i - \theta}}{1 - e^{-\beta x_i - \theta}} = 0, \\ \frac{\partial \log L}{\partial \theta} &= \frac{2e^{-\theta}}{1 - e^{-\theta}} + n \frac{\frac{dA_1(\theta)}{d\theta}}{A_1(\theta)} + n \frac{\frac{dA_2(\theta)}{d\theta}}{A_2(\theta)} + n \frac{e^{-\theta}}{1 - e^{-\theta}} - \sum_{i=1}^n \frac{e^{-\beta x_i - \theta}}{1 + e^{-\beta x_i - \theta}} - 5 \sum_{i=1}^n \frac{e^{-\beta x_i - \theta}}{1 - e^{-\beta x_i - \theta}} = 0. \end{aligned}$$

Since it has no closed form solution, we can use numerical methods such as the quasi-Newton algorithm to numerically optimize the log-likelihood function.

3.2 EM Algorithm

To start the EM algorithm with an E-step, the pdf of the conditional distribution is obtained as follows:

$$f(m|x) = \frac{f(x|m)P(M=m)}{f(x)} = \frac{r^{m-1} [A_1(\theta)m + A_2(\theta)(1 - e^{-\theta})m^2] (1-r)^2}{A_1(\theta) + A_2(\theta)(1 - e^{-\theta}) \frac{1+r}{1-r}}, \quad m = 1, 2, \dots$$

For M step of EM algorithm, we take a sample with n sample size, then

$$\sum_{i=1}^n \sum_{m=1}^{\infty} \frac{\partial \log(f(x_i, m))}{\partial \theta} f(m|x_i; \theta^0, \beta^0) = 0, \quad (3.1)$$

$$\sum_{i=1}^n \sum_{m=1}^{\infty} \frac{\partial \log(f(x_i, m))}{\partial \beta} f(m|x_i; \theta^0, \beta^0) = 0. \quad (3.2)$$

After the equations are solved simultaneously, we get the following expressions:

$$\begin{aligned} \frac{\partial \log f(x, m)}{\partial \theta} &= \frac{\partial \log(f(x|m)P(M=m))}{\partial \theta} = -m + \frac{1}{1 - e^{-\theta}} \\ &+ \frac{\frac{dA_1}{d\theta} + \left((1 - e^{-\theta}) \frac{dA_2(\theta)}{d\theta} + A_2(\theta) e^{-\theta} \right) m}{A_1(\theta) + A_2(\theta)(1 - e^{-\theta})m}. \end{aligned}$$

For simplicity, let $a_1^{(1)} = \frac{A_1(\theta)}{A_2(\theta)(1 - e^{-\theta})}$, $a_2^{(1)} = \frac{(A_2(\theta)e^{-\theta} + (1 - e^{-\theta}) \frac{dA_2(\theta)}{d\theta})}{A_2(\theta)(1 - e^{-\theta})}$ and $a_3^{(1)} = \frac{\frac{dA_1(\theta)}{d\theta}}{(A_2(\theta)e^{-\theta} + (1 - e^{-\theta}) \frac{dA_2(\theta)}{d\theta})}$, and taking derivative of $\log f(x, m)$ with respect to β , we get

$$\frac{\partial \log f(x, m)}{\partial \beta} = \frac{\partial \log(f(x|m)P(M=m))}{\partial \beta} = \frac{1}{\beta} - xm.$$

When we substitute the above expressions in equations (3.1) and (3.2), we have

$$\sum_{i=1}^n \sum_{m=1}^{\infty} \left[-m + \frac{1}{1 - e^{-\theta^{(1)}}} + a_2^{(1)} \frac{a_3^{(1)} + m}{a_1^{(1)} + m} \right] f(m | x_i; \theta^{(0)}, \beta^{(0)}) = 0,$$

and

$$\sum_{i=1}^n \sum_{m=1}^{\infty} \left[\frac{1}{\beta^{(1)}} - x_i m \right] f(m | x_i; \theta^{(0)}, \beta^{(0)}) = 0.$$

For simplicity the coefficients which include parameter in $f(m | x_i; \theta^{(0)}, \beta^{(0)})$, is reorganized as

$$f(m | x_i; \theta^{(0)}, \beta^{(0)}) = \frac{(r_i^{(0)})^{m-1} (1 - r_i^{(0)})^2 m [a_1^{(0)} + m]}{a_1^{(0)} + \frac{1+r_i^{(0)}}{1-r_i^{(0)}}}.$$

After simplifying the equations and performing some algebraic operations, the following expression is obtained:

$$-\sum_{i=1}^n E[M | x_i; \theta^{(0)}, \beta^{(0)}] + \frac{n}{1 - e^{-\theta^{(1)}}} + a_2^{(1)} \sum_{i=1}^n E\left[\frac{a_3^{(1)} + M}{a_1^{(1)} + M} | x_i; \theta^{(0)}, \beta^{(0)}\right] = 0, \tag{3.3}$$

$$\frac{n}{\beta^{(1)}} - \sum_{i=1}^n x_i E[M | x_i; \theta^{(0)}, \beta^{(0)}] = 0. \tag{3.4}$$

Here, first conditional expectations in equation (3.3) are calculated respectively as follows:

$$\begin{aligned} E[M | x_i; \theta^{(0)}, \beta^{(0)}] &= \frac{(1 - r_i^{(0)})^2}{a_1^{(0)} + \frac{1+r_i^{(0)}}{1-r_i^{(0)}}} \sum_{m=1}^{\infty} m^2 [a_1^{(0)} + m] (r_i^{(0)})^{m-1} \\ &= \frac{(1 - r_i^{(0)})^1}{a_1^{(0)} + \frac{1+r_i^{(0)}}{1-r_i^{(0)}}} \left[a_1^{(0)} \sum_{m=1}^{\infty} m^2 (r_i^{(0)})^{m-1} (1 - r_i^{(0)}) + \sum_{m=1}^{\infty} m^3 (r_i^{(0)})^{m-1} (1 - r_i^{(0)}) \right] \end{aligned}$$

The above sums represent the second and third raw moment, respectively, of the geometric distribution with probability of success $(1 - r_i^{(0)})$. Using this information, $E[M | x_i; \theta^{(0)}, \beta^{(0)}]$ can be written as follows:

$$\begin{aligned} E[M | x_i; \theta^{(0)}, \beta^{(0)}] &= \frac{(1 - r_i^{(0)})^1}{a_1^{(0)} + \frac{1+r_i^{(0)}}{1-r_i^{(0)}}} \left[a_1^{(0)} \frac{1 + r_i^{(0)}}{(1 - r_i^{(0)})^2} + \frac{1 + 4r_i^{(0)} + (r_i^{(0)})^2}{(1 - r_i^{(0)})^3} \right] \\ &= \frac{a_1^{(0)} (1 - (r_i^{(0)})^2) + 1 + 4r_i^{(0)} + (r_i^{(0)})^2}{a_1^{(0)} (1 - r_i^{(0)})^2 + (1 - (r_i^{(0)})^2)}. \end{aligned}$$

The second conditional expectation in equation (3.3) is calculated from the following expression

$$E\left[\frac{a_3^{(1)} + M}{a_1^{(1)} + M} | x_i; \theta^{(0)}, \beta^{(0)}\right] = \frac{(1 - r_i^{(0)})^2}{a_1^{(0)} + \frac{1+r_i^{(0)}}{1-r_i^{(0)}}} \sum_{m=1}^{\infty} \frac{m [a_3^{(1)} + m] [a_1^{(0)} + m] (r_i^{(0)})^{m-1}}{a_1^{(1)} + m}. \tag{3.5}$$

The sum at the right hand side of the above equation can be written as

$$\begin{aligned} \sum_{m=1}^{\infty} \frac{m \left[a_3^{(1)} + m \right] \left[a_1^{(0)} + m \right] \left(r_i^{(0)} \right)^{m-1}}{a_1^{(1)} + m} &= \sum_{m=1}^{\infty} m \left[a_3^{(1)} + m \right] \left[a_1^{(0)} + m \right] \int_0^{r_i^{(0)}} z^{a_1^{(1)} + m - 1} dz \left(r_i^{(0)} \right)^{-a_1^{(1)} - 1} \\ &= \left(r_i^{(0)} \right)^{-a_1^{(1)} - 1} \int_0^{r_i^{(0)}} z^{a_1^{(1)}} \sum_{m=1}^{\infty} z^{m-1} \left[m^3 + m^2 \left(a_1^{(0)} + a_3^{(1)} \right) + a_1^{(0)} a_3^{(1)} m \right] dz, \end{aligned}$$

and the sum in the integral can be written as

$$\begin{aligned} \sum_{m=1}^{\infty} z^{m-1} \left[m^3 + m^2 \left(a_1^{(0)} + a_3^{(1)} \right) + a_1^{(0)} a_3^{(1)} m \right] &= \sum_{m=1}^{\infty} m^3 z^{m-1} + \left(a_1^{(0)} + a_3^{(1)} \right) \sum_{m=1}^{\infty} m^2 z^{m-1} + a_1^{(0)} a_3^{(1)} \sum_{m=1}^{\infty} m z^{m-1} \\ &= \frac{1 + 4z + z^2}{(1-z)^4} + \left(a_1^{(0)} + a_3^{(1)} \right) \frac{1+z}{(1-z)^3} + a_1^{(0)} a_3^{(1)} \frac{1}{(1-z)^2}. \end{aligned}$$

After these results are substituted in (3.5), the sum in the expression can be written as follows:

$$\begin{aligned} \sum_{m=1}^{\infty} \frac{\left[m^3 + m^2 \left(a_1^{(0)} + a_3^{(1)} \right) + a_1^{(0)} a_3^{(1)} m \right] \left(r_i^{(0)} \right)^{m-1}}{a_1^{(1)} + m} &= \left(r_i^{(0)} \right)^{-a_1^{(1)} - 1} \int_0^{r_i^{(0)}} z^{a_1^{(1)}} \left(\frac{1 + 4z + z^2}{(1-z)^4} + \left(a_1^{(0)} + a_3^{(1)} \right) \frac{1+z}{(1-z)^3} \right. \\ &\quad \left. + a_1^{(0)} a_3^{(1)} \frac{1}{(1-z)^2} \right) dz. \end{aligned}$$

Taylor series expansions of $\frac{1+4z+z^2}{(1-z)^4}$, $\frac{1+z}{(1-z)^3}$ and $\frac{1}{(1-z)^2}$ expressions will be used to obtain a more simple expression of the integral.

$$\begin{aligned} &\int_0^{r_i^{(0)}} z^{a_1^{(1)}} \left(\frac{1 + 4z + z^2}{(1-z)^4} + \left(a_1^{(0)} + a_3^{(1)} \right) \frac{1+z}{(1-z)^3} + a_1^{(0)} a_3^{(1)} \frac{1}{(1-z)^2} \right) dz \\ &= \sum_{j=0}^{\infty} \int_0^{r_i^{(0)}} z^{a_1^{(1)} + j} dz \left((1+j)^3 + \left(a_1^{(0)} + a_3^{(1)} \right) (1+j)^2 + \left(a_1^{(0)} a_3^{(1)} \right) (1+j) \right) \\ &= \sum_{j=0}^{\infty} \frac{\left(r_i^{(0)} \right)^{a_1^{(1)} + j + 1}}{\left(a_1^{(1)} + j + 1 \right)} \left((1+j)^3 + \left(a_1^{(0)} + a_3^{(1)} \right) (1+j)^2 + \left(a_1^{(0)} a_3^{(1)} \right) (1+j) \right). \end{aligned}$$

Then, we finally obtain $E \left[\frac{a_3^{(1)} + M}{a_1^{(1)} + M} \mid x_i; \theta^{(0)}, \beta^{(0)} \right]$ as follows:

$$\begin{aligned} E \left[\frac{a_3^{(1)} + M}{a_1^{(1)} + M} \mid x_i; \theta^{(0)}, \beta^{(0)} \right] &= \frac{\left(1 - r_i^{(0)} \right)^2}{a_1^{(0)} + \frac{1+r_i^{(0)}}{1-r_i^{(0)}}} \times \\ &\quad \sum_{j=0}^{\infty} \frac{\left(r_i^{(0)} \right)^j}{\left(a_1^{(1)} + j + 1 \right)} \left((1+j)^3 + \left(a_1^{(0)} + a_3^{(1)} \right) (1+j)^2 + \left(a_1^{(0)} a_3^{(1)} \right) (1+j) \right) \end{aligned}$$

After the expected values are submitted in equations (3.3) and (3.4), the following expressions are obtained:

$$\begin{aligned} \frac{n}{1 - e^{-\theta^{(1)}}} + a_2^{(1)} \sum_{i=1}^n \left[\frac{\left(1 - r_i^{(0)} \right)^2}{a_1^{(0)} + \frac{1+r_i^{(0)}}{1-r_i^{(0)}}} \sum_{j=0}^{\infty} \frac{\left(r_i^{(0)} \right)^j}{\left(a_1^{(1)} + j + 1 \right)} \left((1+j)^3 + \left(a_1^{(0)} + a_3^{(1)} \right) (1+j)^2 + \left(a_1^{(0)} a_3^{(1)} \right) (1+j) \right) \right] \\ = \sum_{i=1}^n \frac{a_1^{(0)} \left(1 - \left(r_i^{(0)} \right)^2 \right) + 1 + 4r_i^{(0)} + \left(r_i^{(0)} \right)^2}{a_1^{(0)} \left(1 - r_i^{(0)} \right)^2 + \left(1 - \left(r_i^{(0)} \right)^2 \right)}, \end{aligned}$$

and

$$\beta^{(1)} = \frac{n}{\sum_{i=1}^n x_i \frac{a_1^{(0)} \left(1 - (r_i^{(0)})^2\right) + 1 + 4r_i^{(0)} + (r_i^{(0)})^2}{a_1^{(0)} \left(1 - r_i^{(0)}\right)^2 + \left(1 - (r_i^{(0)})^2\right)}}.$$

So that, the parameter values are obtained iteratively. Here

$$a_1^{(s)} = \frac{1}{\theta^{(s)}} + \frac{\left(1 - 2e^{-\theta^{(s)}}\right)}{\left(1 - e^{-\theta^{(s)}}\right)}, \text{ with } s = 0, 1,$$

$$a_2^{(1)} = \frac{\left(A_2(\theta)e^{-\theta} + (1 - e^{-\theta}) \frac{dA_2}{d\theta}\right)}{A_2(\theta) \left(1 - e^{-\theta}\right)} = \frac{1}{\theta^{(1)} \left(2\theta^{(1)} + 1\right)},$$

$$a_3^{(1)} = \frac{\frac{dA_1(\theta)}{d\theta}}{\left(A_2(\theta)e^{-\theta} + (1 - e^{-\theta}) \frac{dA_2(\theta)}{d\theta}\right)} = \frac{e^{-\theta^{(1)}} \left(1 + \theta^{(1)} + 2\theta^{(1)2}\right) - 1}{\left(1 - e^{-\theta^{(1)}}\right)^2},$$

and

$$r_i^{(0)} = e^{-\beta^{(0)} x_i - \theta^{(0)}}.$$

If we want to obtain an equation for $E \left[\frac{a_3^{(1)} + M}{a_1^{(1)} + M} \mid x_i; \theta^{(0)}, \beta^{(0)} \right]$ using the generalized hypergeometric function, we can write the following expression:

$$\begin{aligned} & \sum_{m=1}^{\infty} \frac{m^3 + m^2 \left(a_1^{(0)} + a_3^{(1)}\right) + a_1^{(0)} a_3^{(1)} m \left(r_i^{(0)}\right)^{m-1}}{a_1^{(1)} + m} \\ &= \sum_{m=1}^{\infty} \frac{m^3 \left(r_i^{(0)}\right)^{m-1}}{a_1^{(1)} + m} + \left(a_1^{(0)} + a_3^{(1)}\right) \sum_{m=1}^{\infty} \frac{m^2 \left(r_i^{(0)}\right)^{m-1}}{a_1^{(1)} + m} + \left(a_1^{(0)} a_3^{(1)}\right) \sum_{m=1}^{\infty} \frac{m \left(r_i^{(0)}\right)^{m-1}}{a_1^{(1)} + m} \\ &= \frac{1}{a_1^{(1)} + 1} \left\{ {}_4F_3 \left[2, 2, 2, a_1^{(1)} + 1; 1, 1, a_1^{(1)} + 2; r_i^{(0)} \right] + \left(a_1^{(0)} + a_3^{(1)}\right) {}_3F_2 \left[2, 2, a_1^{(1)} + 1; 1, a_1^{(1)} + 2; r_i^{(0)} \right] \right. \\ & \quad \left. + a_1^{(0)} a_3^{(1)} {}_2F_1 \left[2, a_1^{(1)} + 1; a_1^{(1)} + 2; r_i^{(0)} \right] \right\}. \end{aligned}$$

Thus, the iterative solution for θ is obtained as follows:

$$\begin{aligned} & \frac{n}{1 - e^{-\theta^{(1)}}} + \frac{a_2^{(1)}}{a_1^{(1)} + 1} \sum_{i=1}^n \left[\frac{{}_4F_3 \left[2, 2, 2, a_1^{(1)} + 1; 1, 1, a_1^{(1)} + 2; r_i^{(0)} \right]}{\left(a_1^{(0)} + a_3^{(1)}\right) {}_3F_2 \left[2, 2, a_1^{(1)} + 1; 1, a_1^{(1)} + 2; r_i^{(0)} \right]} \right. \\ & \quad \left. + a_1^{(0)} a_3^{(1)} {}_2F_1 \left[2, a_1^{(1)} + 1; a_1^{(1)} + 2; r_i^{(0)} \right] \right] \\ &= \sum_{i=1}^n \frac{a_1^{(0)} \left(1 - (r_i^{(0)})^2\right) + 1 + 4r_i^{(0)} + (r_i^{(0)})^2}{a_1^{(0)} \left(1 - r_i^{(0)}\right)^2 + \left(1 - (r_i^{(0)})^2\right)}. \end{aligned}$$

For simplicity, let $e^{-\theta} \cong 1 - \theta$, $\theta \in (0, 1)$, except for r_i , then the following expressions are obtained:

$$a_1^{(1)} = a_1^{(0)} = 2, \quad a_2^{(1)} = \frac{1}{\theta^{(1)} \left(2\theta^{(1)} + 1\right)} \quad \text{and} \quad a_3^{(1)} = 1 - 2\theta^{(1)},$$

$$f \left(m \mid x_i; \theta^{(0)}, \beta^{(0)} \right) = \frac{\left(r_i^{(0)}\right)^{m-1} \left(1 - r_i^{(0)}\right)^3 m [2 + m]}{3 - r_i^{(0)}},$$

$$E \left[M \mid x_i; \theta^{(0)}, \beta^{(0)} \right] = \frac{3 + 4r_i^{(0)} - \left(r_i^{(0)}\right)^2}{\left(3 - r_i^{(0)}\right) \left(1 - r_i^{(0)}\right)} = \frac{6}{\left(3 - r_i^{(0)}\right) \left(1 - r_i^{(0)}\right)} - 1 = \frac{3}{\left(1 - r_i^{(0)}\right)} + \frac{-3}{\left(3 - r_i^{(0)}\right)} - 1,$$

$$\begin{aligned} E \left[\frac{1 - 2\theta^{(1)} + M}{2 + M} \mid x_i; \theta^{(0)}, \beta^{(0)} \right] &= \frac{\left(1 - r_i^{(0)}\right)^3}{3 - r_i^{(0)}} \sum_{m=1}^{\infty} \frac{m \left(1 - 2\theta^{(1)} + m\right) [2 + m] \left(r_i^{(0)}\right)^{m-1}}{2 + m} \\ &= \frac{\left(1 - r_i^{(0)}\right)^3}{3 - r_i^{(0)}} \left[\left(1 - 2\theta^{(1)}\right) \sum_{m=1}^{\infty} m \left(r_i^{(0)}\right)^{m-1} + \sum_{m=1}^{\infty} m^2 \left(r_i^{(0)}\right)^{m-1} \right] \\ &= \left(1 - 2\theta^{(1)}\right) \frac{1 - r_i^{(0)}}{3 - r_i^{(0)}} + \frac{1 + r_i^{(0)}}{3 - r_i^{(0)}} \\ &= - \left(1 + 2\theta^{(1)}\right) \frac{1 - r_i^{(0)}}{3 - r_i^{(0)}} + 1. \end{aligned}$$

When the above expressions are substituted in (3.3) and (3.4), the following equations are obtained:

$$\begin{aligned} \frac{2}{\theta^{(1)}} \sum_{i=1}^n \frac{1}{3 - r_i^{(0)}} + \frac{n}{\theta^{(1)} (1 + 2\theta^{(1)})} &= 3 \sum_{i=1}^n \frac{1}{1 - r_i^{(0)}} - 3 \sum_{i=1}^n \frac{1}{3 - r_i^{(0)}} - n, \\ \beta^{(1)} &= \frac{n}{6 \sum_{i=1}^n \frac{x_i}{\left(3 - r_i^{(0)}\right) \left(1 - r_i^{(0)}\right)} - \sum_{i=1}^n x_i}. \end{aligned}$$

Taking $m_1 = \frac{1}{n} \sum_{i=1}^n \frac{1}{1 - r_i^{(0)}}$ and $m_2 = \frac{1}{n} \sum_{i=1}^n \frac{1}{3 - r_i^{(0)}}$, the derivative expression with respect to θ can be rewritten as follows:

$$\frac{2}{\theta^{(1)}} m_2 + \frac{1}{\theta^{(1)} (1 + 2\theta^{(1)})} = 3m_1 - 3m_2 - 1$$

After the necessary adjustments, a second-order polynomial for θ is obtained as;

$$2(3m_1 - 3m_2 - 1)\theta^{(1)2} - \theta^{(1)}(1 - 3m_1 + 7m_2) - 1 - 2m_2 = 0.$$

Then, the appropriate solution for θ is founded as

$$\theta^{(1)} = \frac{(1 - 3m_1 + 7m_2) + \sqrt{(1 - 3m_1 + 7m_2)^2 + 8(3m_1 - 3m_2 - 1)(1 + 2m_2)}}{4(3m_1 - 3m_2 - 1)}.$$

Let $e^{-\theta} \cong \frac{1}{1+\theta}$, $\theta \geq 1$, except for r_i , then the following expressions are obtained:

$$a_1^{(1)} = a_1^{(0)} = 1, \quad a_2^{(1)} = \frac{1}{\theta(1+2\theta)} \quad \text{and} \quad a_3^{(1)} = 2(1+\theta),$$

$$f \left(m \mid x_i; \theta^{(0)}, \beta^{(0)} \right) = \frac{\left(r_i^{(0)}\right)^{m-1} \left(1 - r_i^{(0)}\right)^3 m [1 + m]}{2},$$

$$E \left[M \mid x_i; \theta^{(0)}, \beta^{(0)} \right] = \frac{1 + 2r_i^{(0)}}{1 - r_i^{(0)}},$$

$$E \left[\frac{2(1 + \theta^{(1)}) + M}{1 + M} \mid x_i; \theta^{(0)}, \beta^{(0)} \right] = \frac{(1 + 2\theta^{(1)})}{2} \left(1 - r_i^{(0)}\right) + 1.$$

When the above expressions are substituted in (3.3) and (3.4), the following equations are obtained:

$$n \frac{(1 + \theta^{(1)})}{\theta^{(1)}} + \frac{n}{\theta^{(1)} (1 + 2\theta^{(1)})} + \frac{1}{2\theta^{(1)}} \sum_{i=1}^n (1 - r_i^{(0)}) = \sum_{i=1}^n \frac{1 + 2r_i^{(0)}}{1 - r_i^{(0)}},$$

$$\beta^{(1)} = \frac{n}{3 \sum_{i=1}^n \frac{x_i}{(1 - r_i^{(0)})} - 2 \sum_{i=1}^n x_i}.$$

Taking $m_3 = \frac{1}{n} \sum_{i=1}^n \frac{1+2r_i^{(0)}}{1-r_i^{(0)}}$ and $m_4 = \frac{1}{n} \sum_{i=1}^n \frac{1-r_i^{(0)}}{2}$, the derivative expression with respect to θ can be rewritten as follows:

$$2(m_3 - 1)\theta^{(1)^2} - \theta^{(1)}(3 - m_3 + 2m_4) - (2 + m_4) = 0$$

Then, the appropriate solution for θ is founded as

$$\theta^{(1)} = \frac{(3 - m_3 + 2m_4) + \sqrt{(3 - m_3 + 2m_4)^2 + 8(m_3 - 1)(2 + m_4)}}{4(m_3 - 1)}.$$

4. Application

To show the usefulness of EDL distribution, we take into consideration two different data sets. The parameters are estimated with the ML method and to decide the best fit distribution, we use Kolmogorov-Smirnov (K-S) statistics.

Data 1.

The first data set used by [9] is vinyl chloride data obtained from clean upgradient monitoring wells in mg/l. The data are 5.1, 1.2, 1.3, 0.6, 0.5, 2.4, 0.5, 1.1, 8, 0.8, 0.4, 0.6, 0.9, 0.4, 2, 0.5, 5.3, 3.2, 2.7, 2.9, 2.5, 2.3, 1, 0.2, 0.1, 0.1, 1.8, 0.9, 2, 4, 6.8, 1.2, 0.4, 0.2.

Lindley, Weibull and exponential distributions are considered as the competitor because these are commonly used in literature for fitting lifetime data. ML estimates for data set 1 of the parameters of the EDL distribution respectively are calculated as $\hat{\theta} = 2.0363$, $\hat{\beta} = 0.4804$.

Table 1. Kolmogorov-Smirnov and p values of the models based on data set 1

Model	Kolmogorov-Smirnov	p value
EDL	0.0878	0.9353
Lindley	0.1326	0.5880
Weibull	0.0919	0.9364
Exp	0.0889	0.9508

Table1 shows that the EDL distribution gives the better fit than the other three distributions.

Data 2.

The second data set corresponding to intervals in days between 109 successive coal-mining disasters in Great Britain, for the period 1875-1951, published by [10]. The sorted data are given as follows: 1, 4, 4, 7, 11, 13, 15, 15, 17, 18, 19, 19, 20, 20, 22, 23, 28, 29, 31, 32, 36, 37, 47, 48, 49, 50, 54, 54, 55, 58, 59, 59, 61, 61, 66, 72, 72, 75, 78, 78, 81, 93, 96, 99, 108, 113, 114, 120, 120, 123, 124, 129, 131, 137, 145, 151, 156, 171, 176, 182, 188, 189, 195, 203, 208, 215, 217, 217, 217, 224, 228, 233, 255, 271, 275, 275, 275, 286, 291, 312, 312, 315, 326, 326, 329, 330, 336, 338, 345, 348, 354, 361, 364, 369, 378, 390, 457, 467, 498, 517, 566, 644, 745, 871, 1205, 1312, 1357, 1613, 1630.

[1] and [2] and some different authors are proposed many models on this data set.[1] suggested to use Exponential-Geometric (E-G) and [2] suggested to use Exponential-Poisson (E-P) distributions. On the other hand, [11] and [8] have proposed two-component mixed exponential distribution (2MED) and exponential modified discrete Lindley distribution (EMDL) for modeling this data set. According to reported result of [12], Exponentiated exponential (EE) distribution is also used for modeling this data set. ML estimates for data set 2 of the parameters of the EDL distribution respectively are calculated as $\hat{\theta} = 0.7603$, $\hat{\beta} = 0.0025$.

According to Kolmogorov-Smirnov values from Table 2, 2MED, E-P and EDL distributions give better fit than the other three distributions. Note that, EDL has the third place within three competitors.

Table 2. Kolmogorov-Smirnov and p values for E-P, E-G, 2MED, EMDL, EDL and EE.

Model	Kolmogorov-Smirnov	p value
E-P	0.0625	0.7876
E-G	0.0761	0.5524
2MED	0.0578	0.8386
EMDL	0.0752	0.5436
EDL	0.0748	0.5508
EE	0.0830	0.4402

Conclusions

In this article, we have proposed and discussed a new model named as the EDL distribution which compounds the discrete Lindley distribution and exponential distribution. Some statistical properties of the proposed distribution including the survival, hazard rate and mean residual life functions, moments, moment generating function, generating random number and R enyi entropy are presented. Moreover, we have derived the ML estimates and EM estimates of the parameters. The applicability of the proposed model is illustrated by using the real data sets.

References

- [1] Adamidis K., Loukas, S.: *A lifetime distribution with decreasing failure rate*. Stat Probab Lett. **39** (1), 35-42 (1998). [https://doi.org/10.1016/S0167-7152\(98\)00012-1](https://doi.org/10.1016/S0167-7152(98)00012-1)
- [2] Kus, C.: *A new lifetime distribution*. Comput Stat Data Anal. **51** (9), 4497-4509 (2007). <https://doi.org/10.1016/j.csda.2006.07.017>.
- [3] Tahmasbi, R., Rezaei, S.: *A two-parameter lifetime distribution with decreasing failure rate*. Comput Stat Data Anal. **52** (8), 3889-3901 (2008). <https://doi.org/10.1016/j.csda.2007.12.002>
- [4] Chahkandi, M., Ganjali, M.: *On some lifetime distributions with decreasing failure rate*. Comput Stat Data Anal. **53** (12), 4433-4440 (2009). <https://doi.org/10.1016/j.csda.2009.06.016>
- [5] Barreto-Souza, W., Bakouch, H.S.: *A new lifetime model with decreasing failure rate*. Statistics. **47** (2), 465-476 (2013) <https://doi.org/10.1080/02331888.2011.595489>
- [6] Lu, W., Chiang, J. Y.: *On some life distributions with flexible failure rate*. Quality Technology & Quantitative Management. **15** (3), 413-433 (2018). <https://doi.org/10.1080/16843703.2016.1226596>
- [7] Gui, W., Zhang, S., Lu, X.: *The Lindley-Poisson distribution in lifetime analysis and its properties*. Hacet J Math Stat. **43** (6), 1063-1077 (2014).
- [8] Yilmaz, M., Hameldarbandi, M., Kemaloglu, S.A.: *Exponential-modified discrete Lindley distribution*. SpringerPlus. **5** (1660), (2016b). <https://doi.org/10.1186/s40064-016-3302-2>
- [9] Bhaumik, D.K., Kapur, K., Gibbons, R.D.: *Testing Parameters of a Gamma Distribution for Small Samples*. Technometrics. **51**, 326-334 (2009). <https://doi.org/10.1198/tech.2009.07038>
- [10] Maguire, B.A., Pearson, E.S., Wynn, A.H.A.: *The time intervals between industrial accidents*. Biometrika. **39** (1/2), 168-180 (1952). <https://doi.org/10.2307/2332475>
- [11] Yilmaz, M., Potas, N., Buyum, B.: *A Classical Approach to Modeling of Coal Mine Data*. In: Er etin  . (eds) Chaos, Complexity and Leadership 2014. Springer Proceedings in Complexity. Springer, Cham. **Chapter 7**, 65-72 (2016). https://doi.org/10.1007/978-3-319-18693-1_7
- [12] Mahdavi, A. and Kundu, D.: *A New Method for Generating Distributions with an Application to Exponential Distribution*. Communications in Statistics-Theory and Methods. **46** (13), 6543-6557 (2017). <https://doi.org/10.1080/03610926.2015.1130839>

Affiliations

SIBEL ACIK KEMALOGU

ADDRESS: Ankara University, Dept. of Statistics, Ankara-Turkey.

E-MAIL: acik@ankara.edu.tr

ORCID ID: orcid.org/0000-0003-0449-6966

MEHMET YILMAZ

ADDRESS: Ankara University, Dept. of Statistics, Ankara-Turkey.

E-MAIL: mehmetyilmaz@ankara.edu.tr

ORCID ID: orcid.org/0000-0002-9762-6688