



Different Parameter Estimation Methods for Exponential Geometric Distribution and Its Applications in Lifetime Data Analysis



Feyza Günay^{1*} and Mehmet Yilmaz²

¹Graduate School of Natural and Applied Sciences, Ankara University, Turkey

²Department of Statistics, Faculty of Science, Ankara University, Turkey

Submission: July 19, 2018; **Published:** September 25, 2018

***Corresponding author:** Feyza Günay, Graduate School of Natural and Applied Sciences, Ankara University, Ankara, Turkey;
Email: fgunay@ankara.edu.tr

Abstract

The new compound distributions which are started to be used with the study of Adamidis, et al. [1] still take place in recent studies. Exponential Geometric distribution, introduced by them, is a flexible distribution for modeling the lifetime data sets. They have used maximum likelihood method with expectation-maximization algorithm to estimate unknown parameters. In this paper, we use maximum likelihood and also least squares, weighted least squares, maximum product of spacings and l-moments methods to estimate the unknown parameters of exponential geometric distribution family. Then we compare the efficiency of these estimators via a simulation study for different sample sizes and parameter settings. At the end of the study, two lifetime data sets such as coal mine data and medical data about occupational safety and duration hospitalization studies are illustrated for application.

Keywords: Exponential geometric distribution; Maximum likelihood; Least squares; Weighted least squares; Maximum product of spacings; L-moments; Lifetime data analysis

Abbreviations: EG: Exponential Geometric; EP: Exponential Poisson; DFR: Decreasing Failure Rate; PE: Poisson-Exponential; CEG: Complementary Exponential Geometric; ML: Maximum Likelihood; EM: Expectation-Maximization; LS: Least Squares; WLS: Weighted Least Squares; MPS: Maximum Product of Spacings; LM: L-moments; AIC: Akaike Information Criterion

Introduction

Survival analysis includes modeling the time until the occurrence of an event of interest. It is generally considered as time until death or failure. In recent years, various compound distributions are proposed by using well known lifetime distributions. Firstly, the Exponential Geometric (EG) distribution proposed by [1] became the focus point of various studies [2-9]. In the same way with EG, [10] also proposed Exponential Poisson (EP) distribution which has decreasing failure rate (DFR). Barreto, et al. [4] generalized EP with the help of a power parameter and investigated its properties. [5] studied Poisson-Exponential (PE) distribution and its properties. [11] examined Complementary Exponential Geometric (CEG) distribution and its characteristics [12]. Similarly, [13-18] conducted various studies on compound distributions.

In this study, our aim is to estimate the unknown parameters of EG distribution family by using Maximum Likelihood (ML) estimation with Expectation-Maximization (EM) algorithm and Least Squares (LS), Weighted Least Squares (WLS), Maximum Product of Spacings (MPS) and L-moments (LM) estimation methods. We then compare the efficiencies of these estimators via a simulation study for different sample sizes and parameter settings. It should

be noted that we have used different estimation methods from earlier studies. Additionally, we have illustrated the performance of fitting two data sets into the EG distribution.

To the best of our knowledge, this is the first study to obtain LS, WLS, MPS, LM estimators in the context of parameter estimation for EG distribution family. [1] have used ML estimators using the EM algorithm to estimate unknown parameters of EG distribution. So we compare the efficiency of estimators that we use with ML estimators used by them. The remaining sections of this paper are organized as follows. In Section 1, EG distribution is introduced briefly. Parameter estimation methods are presented in detail in Section 2. Results for an extensive simulation study to compare the performances of parameter estimation methods are presented in Section 3. Section 4 illustrates real data applications. Finally, main results are summarized in Section 5.

Exponential geometric distribution

Let N be a random variable has a geometric distribution with parameter p (probability for failure) and Y_1, Y_2, \dots, Y_N be random variables from exponential distribution with independent param-

eter β . For $i \geq 1$, Y_i and N are independent from each other; Let X be a random variable defined as $X = \{Y_1, Y_2, \dots, Y_N\}$ then it follows a $EG(\beta, p)$ distribution. The cumulative distribution function (cdf) is given as follows.

$$F(x | p; \beta) = P(X \leq x) = \frac{1 - e^{-x\beta}}{1 - pe^{-x\beta}}, \quad x > 0, \beta \in \mathcal{R}^+, p \in (0,1) \quad (1)$$

The survival and hazard functions of $EG(\beta, p)$ distribution are

$$L(\beta, p; \bar{x}) = \prod_{i=1}^m f(x_i; \beta, p) = \prod_{i=1}^m \frac{(1-p)\beta e^{-x_i\beta}}{(1-pe^{-x_i\beta})^2}, \quad (2)$$

$$\log(\mathcal{L}(\beta, p; \bar{x})) = m \log(1-p) + m \log(\beta) - \beta \sum_{i=1}^m x_i - \sum_{i=1}^m (-pe^{-x_i\beta})$$

$$\beta^{(t+1)} = m \sum_{i=1}^m \left\{ \frac{x_i (1 + p^{(t)} e^{-\beta^{(t)} x_i})}{1 - p^{(t)} e^{-\beta^{(t)} x_i}} \right\}^{-1}, \quad (3)$$

respectively. The probability density function (pdf) of X is given by

$$\frac{\partial S}{\partial \beta} = \frac{1}{m+1} \left(\frac{(1-p)x_i e^{-x_i\beta}}{(1-pe^{-x_i\beta})(1-e^{-x_i\beta})} - \frac{x_m}{1-pe^{-x_m\beta}} \right) + \frac{1}{m+1} \sum_{i=1}^m \left\{ \frac{x_{(i)} e^{-x_{(i-1)}\beta} - x_{(i-1)} e^{-x_{(i)}\beta}}{e^{x_{(i)}\beta} - e^{x_{(i-1)}\beta}} \right\}$$

$$- \frac{1}{m+1} \sum_{i=1}^m \left(\frac{px_{(i)} e^{x_{(i)}\beta} (1-pe^{x_{(i-1)}\beta}) + px_{(i-1)} e^{x_{(i-1)}\beta} (1-pe^{x_{(i)}\beta})}{(1-pe^{x_{(i)}\beta})(1-pe^{x_{(i-1)}\beta})} \right) = 0 \quad (4)$$

The probability density and hazard functions plotted for different parameter values for EG distribution are shown in Figure 1 and 2.

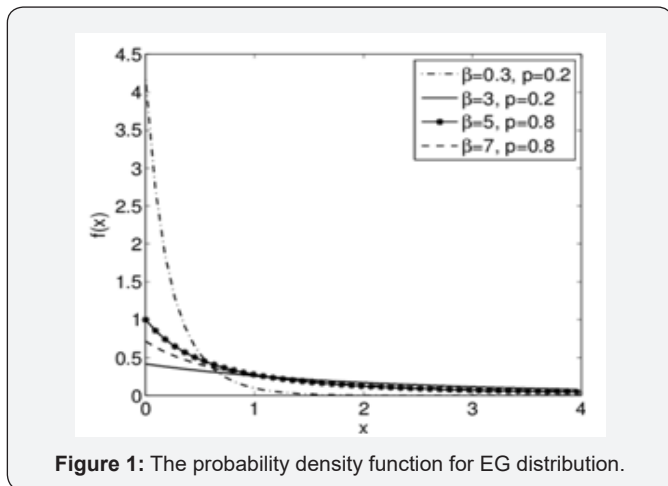


Figure 1: The probability density function for EG distribution.

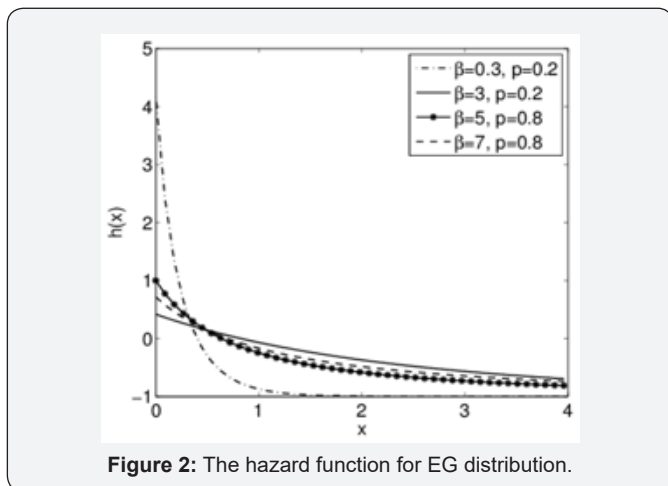


Figure 2: The hazard function for EG distribution.

Parameter estimation

In this study ML, MPS, LS, WLS and LM methods are used to estimate the parameters of the EG distribution. In this section, we give details of these methods for estimation of parameters of the EG distribution.

ML estimation method

Let X_1, X_2, \dots, X_m be a random sample from $EG(\beta, p)$ distribution. In order to estimate the parameters of the EG distribution, the log-likelihood function is maximized with respect to the parameters. The likelihood and log-likelihood function can be written as

$$L(\beta, p; \bar{x}) = \prod_{i=1}^m f(x_i; \beta, p) = \prod_{i=1}^m \frac{(1-p)\beta e^{-x_i\beta}}{(1-pe^{-x_i\beta})^2},$$

$$\log(\mathcal{L}(\beta, p; \bar{x})) = m \log(1-p) + m \log(\beta) - \beta \sum_{i=1}^m x_i - \sum_{i=1}^m (-pe^{-x_i\beta}) \quad (6)$$

From these equations, the ML estimators of parameters cannot be obtained in an analytical form. So this nonlinear equations are solved by a numerical methods such as EM. [7] used a method called EM to obtain ML estimates in case of missing data. The EM algorithm gives remarkable results by computing the ML estimates at each renewal and repeating steps E and M until convergence occurs. Thus, an EM iteration, taking $(\beta^{(t)}, p^{(t)})$ into, $(\beta^{(t+1)}, p^{(t+1)})$

$$\beta^{(t+1)} = m \sum_{i=1}^m \left\{ \frac{x_i (1 + p^{(t)} e^{-\beta^{(t)} x_i})}{1 - p^{(t)} e^{-\beta^{(t)} x_i}} \right\}^{-1}, \quad (7)$$

$$p^{(t+1)} = 1 - m \left\{ \sum_{i=1}^m \frac{1 + p^{(t)} e^{-\beta^{(t)} x_i}}{1 - p^{(t)} e^{-\beta^{(t)} x_i}} \right\}^{-1}. \quad (8)$$

In case of convergence, ML estimators are obtained for the parameters of EG distribution.

MPS method

This method which was proposed as an alternative to the maximum likelihood method, was first introduced by [6]. For the MPS and ML estimators emphasized that they are asymptotically equal and have asymptotic sufficiency, consistency and efficiency properties. The definition of uniform spacings for the MPS of the EG distribution as follows.

$$D_1 = F(x_1) = \frac{1 - e^{-x_1\beta}}{1 - pe^{-x_1\beta}}, \quad (9)$$

$$D_{(m+1)} = 1 - F(x_m) = 1 - \frac{1 - e^{-x_m\beta}}{1 - pe^{-x_m\beta}} = \frac{(1-p)e^{-x_m\beta}}{1 - pe^{-x_m\beta}}.$$

And the general term for spacings is given as

$$D_i = F(x_{(i)}) - F(x_{(i-1)}) = \left[\frac{1 - e^{-x_{(i)}\beta}}{1 - pe^{-x_{(i)}\beta}} \right] - \left[\frac{1 - e^{-x_{(i-1)}\beta}}{1 - pe^{-x_{(i-1)}\beta}} \right]$$

Where $\sum D_i = 1$. The MPS method maximizes the geometric mean of intervals for finding the estimates of parameters. Let geometric mean G and its logarithm S are given by

$$G = \left(\prod_{i=1}^{m+1} D_i \right)^{\frac{1}{m+1}}, \quad S = \frac{1}{m+1} \sum_{i=1}^{m+1} \log D_i,$$

respectively. S is regulated and the equation is written as

$$S = \frac{1}{m+1} \left\{ \log D_1 + \sum_{i=2}^m \log D_i + \log D_{m+1} \right\}$$

$$= \frac{1}{m+1} \left\{ \log \left[\frac{1 - e^{-x_m \beta}}{1 - p e^{-x_m \beta}} \right] + \sum_{i=2}^m \log \left[\frac{(p-1)(e^{-x_{(i)} \beta} - e^{-x_{(i-1)} \beta})}{(1 - p e^{-x_{(i)} \beta})(1 - p e^{-x_{(i-1)} \beta})} \right] \right\} + \frac{1}{m+1} \log \left[\frac{(1-p)e^{-x_m \beta}}{1 - p e^{-x_m \beta}} \right]$$

$\hat{\beta}_{MPS}$ and \hat{p}_{MPS} are obtained by equating the nonlinear derivative equations to zero.

$$\frac{\partial S}{\partial \beta} = \frac{1}{m+1} \left(\frac{(1-p)x_i e^{-x_i \beta}}{(1 - p e^{-x_i \beta})(1 - e^{-x_i \beta})} - \frac{x_m}{1 - p e^{-x_m \beta}} \right) + \frac{1}{m+1} \sum_{i=1}^m \left(\frac{x_{(i)} e^{-x_{(i)} \beta} - x_{(i-1)} e^{-x_{(i-1)} \beta}}{e^{x_{(i)} \beta} - e^{x_{(i-1)} \beta}} \right)$$

$$- \frac{1}{m+1} \sum_{i=1}^m \left(\frac{p x_{(i)} e^{x_{(i)} \beta} (1 - p e^{x_{(i)} \beta}) + p x_{(i-1)} e^{x_{(i-1)} \beta} (1 - p e^{x_{(i-1)} \beta})}{(1 - p e^{x_{(i)} \beta})(1 - p e^{x_{(i-1)} \beta})} \right) = 0 \quad (10)$$

$$\frac{\partial S}{\partial p} = \frac{1}{m+1} \left(\frac{1}{e^{-x_m \beta} - p} + \frac{1 - e^{-x_m \beta}}{(1-p)(e^{-x_m \beta} - p)} \right) + \frac{1}{m+1} \sum_{i=1}^m \left(\frac{1}{p-1} - \frac{e^{x_{(i)} \beta} - e^{x_{(i-1)} \beta}}{(p - e^{x_{(i)} \beta})(p - e^{x_{(i-1)} \beta})} \right) = 0 \quad (11)$$

The analytical solution of these equations cannot be obtained, but the estimates can be reached by iterative methods.

LS estimation method

Let $x_{(1)} < x_{(2)} < \dots < x_{(m)}$ be the m order statistics of the random sample from EG distribution. The objective of LS estimation of the parameters is based on minimizing of the sum of difference between CDF F and empirical distribution \hat{F} . The LS estimators for parameters β and p can be found by minimizing of the following function

$$H = \sum_{i=1}^m \left(F(x_{(i)}) - \hat{F}(x_{(i)}) \right)^2$$

It is well known that $E[F(x_{(i)})] = \hat{F}(x_{(i)}) = \frac{i}{n+1} = u_{(i)}$. Putting the cdf in (12) we get the following equation.

$$H(\beta, p) = \sum_{i=1}^m \left(\frac{1 - e^{-x_{(i)} \beta}}{1 - p e^{-x_{(i)} \beta}} - u_{(i)} \right)^2 \quad (13)$$

The arrangement of the mathematical equation in parentheses and its logarithm are

$$(1 - u_{(i)}) = e^{-x_{(i)} \beta} (1 - u_{(i)} p), \log(1 - u_{(i)}) = -x_{(i)} \beta + \log(1 - u_{(i)} p), \quad (14)$$

respectively. H can be written in logarithmic form as follows.

$$H^* = \sum_{i=1}^m \left[x_{(i)} \beta - \log(1 - u_{(i)} p) + \log(1 - u_{(i)}) \right]^2 \quad (15)$$

The LS estimator for parameter β is as given below.

$$\frac{\partial H^*}{\partial \beta} = 2 \sum_{i=1}^m \left[x_{(i)} \beta + \log(1 - u_{(i)} p) + \log(1 - u_{(i)}) \right] x_{(i)}$$

$$\Rightarrow \sum_{i=1}^m x_{(i)}^2 \beta - x_{(i)} \log(1 - u_{(i)} p) + x_{(i)} \log(1 - u_{(i)}) = 0$$

$$\hat{\beta}_{LSE} = \frac{\sum_{i=1}^m x_{(i)} \log(1 - u_{(i)} p) - \sum_{i=1}^m x_{(i)} \log(1 - u_{(i)})}{\sum_{i=1}^m x_{(i)}^2} \quad (16)$$

With the help of (14), the LS estimator of p is obtained as follows.

$$\frac{\partial H}{\partial p} = 2 \sum_{i=1}^m \left[1 - e^{-x_{(i)} \beta} - u_{(i)} + u_{(i)} p e^{-x_{(i)} \beta} \right] u_{(i)} e^{-x_{(i)} \beta}$$

$$\Rightarrow \sum_{i=1}^m \left[1 - e^{-x_{(i)} \beta} - u_{(i)} + u_{(i)} p e^{-x_{(i)} \beta} \right] u_{(i)} e^{-x_{(i)} \beta} = 0$$

$$\hat{p}_{LSE} = \frac{\sum_{i=1}^m \check{w}_{(i)} e^{-2x_{(i)} \beta} - u_{(i)} e^{-x_{(i)} \beta} + u_{(i)}^2 e^{-x_{(i)} \beta}}{\sum_{i=1}^m \check{w}_{(i)}^2 e^{-2x_{(i)} \beta}} \quad (17)$$

Here $\hat{\beta}_{LSE}$ and \hat{p}_{LSE} are obtained as iterative equations.

WLS estimation method

For the EG distribution, the WLS estimates $\hat{\beta}_{WLS}$ and \hat{p}_{WLS} are obtained by minimizing the function given in the following equation.

$$\sum_{i=1}^m \check{w}_{(i)} \left(F(x_{(i)}) - u_{(i)} \right)^2$$

It is well known that: $w_{(i)} = \frac{1}{V(F(x_{(i)}))} = \frac{(n+1)^2 (n+2)}{i(n-i+1)}$. The cdf of the EG Distribution is written in the function as follows.

$$\sum_{i=1}^m \check{w}_{(i)} \left(\frac{1 - e^{-x_{(i)} \beta}}{1 - p e^{-x_{(i)} \beta}} - u_{(i)} \right)^2 \quad (18)$$

The WLS estimators $\hat{\beta}_{WLS}$ and \hat{p}_{WLS} of β and p are obtained as iterative equations by solving (18).

$$\hat{\beta}_{WLS} = \frac{\sum_{i=1}^m w_{(i)} x_{(i)} \log(1 - \log(p)) - \sum_{i=1}^m w_{(i)} x_{(i)} (-u_{(i)})}{\sum_{i=1}^m w_{(i)} x_{(i)}^2} \quad (19)$$

$$\hat{p}_{WLS} = \frac{\sum_{i=1}^m \check{w}_{(i)} u_{(i)} e^{-2x_{(i)} \beta} - w_{(i)} u_{(i)} e^{-x_{(i)} \beta} + w_{(i)} u_{(i)}^2 e^{-x_{(i)} \beta}}{\sum_{i=1}^m \check{w}_{(i)} u_{(i)}^2 e^{-2x_{(i)} \beta}} \quad (20)$$

L-moments estimation method

The L-moments method was firstly proposed by [9] is similar to method of moment estimators. When the parameters are unknown, we need to equate sample L-moments with the population L-moments. The L-moment estimates for the EG distribution can be obtained by equating first two sample L-moments to first two population L-moments. For the EG distribution, the quantile function is given by

$$Q(u | p, \beta) = \frac{1}{\beta} \log \left(\frac{1 - up}{1 - u} \right). \quad (21)$$

The first two sample L-moments are

$$l_1(p, \beta) = \frac{1}{m} \sum_{i=1}^m x_{(i)}, \quad l_2(p, \beta) = \frac{2}{m(m-1)} \sum_{i=1}^m (i-1) x_{(i)} - l_1$$

The first two population L-moments are

$$\mu_1(p, \beta) = \int_0^1 Q(u | p, \beta) du = E(X | p, \beta) = \frac{(p-1) \log(1-p)}{\beta p}$$

$$\mu_2(p, \beta) = \int_0^1 Q(u | p, \beta) (2u-1) du = \frac{(p-1)(p + \log(1-p))}{\beta p^2}$$

By equating these two sample and population L-moments, it is possible to obtain the \hat{p}_{LME} estimator with the help of nonlinear equation given in the following.

$$\frac{\bar{x}}{p} + \frac{\bar{x}}{\log(1-p)} - l_2 = 0$$

In the same way, $\hat{\beta}_{LME}$ is calculated as follows.

$$\hat{\beta}_{LME} = \frac{(\hat{p}_{LME} - 1) \log(1 - \hat{p}_{LME})}{\bar{x}(\hat{p}_{LME})}$$

Simulation Study

In this section, an extensive simulation study is conducted to compare the performances of parameter estimation methods under different scenarios. Performances are examined by using estimations, biases and mean square errors (MSE). Bias is absolute difference between the true value of the parameter and the expected value of the estimator of parameter. MSE is a well-known and widely-used criterion in the literature for comparing biased (3;0.3), (3;0.5), (3;0.9), (5;0.3), (5;0.5), (5;0.9), (9;0.3), (9;0.5), (9;0.9).

estimators with respect to their efficiencies. It is evaluated via sum of the variance of the estimator of parameter and the square value of bias.

Each experiment is repeated 1000 times in this study. Settings for sample sizes and parameter values are given in the following:

$$n = 50;100;1000 \text{ and } (\beta; p) = (0.5;0.3), (0.5;0.5), (0.5;0.9),$$

In Table 1, bias and MSE values of ML, MPS, LS, WLS and LM estimation with regards to parameters of EG distribution is presented for $n = 50$ and sample size. In Table 2 and 3, it is presented for $n = 100$ and $n = 1000$, respectively.

Table 1: The Bias and MSE(Parenthesis) for $n = 50$.

Parameter β					
$(\beta; p)$	ML	MPS	LS	WLS	LM
(0.5;0.3)	-0.0051(0.0180)	-0.0764(0.0336)	-0.0931(0.0362)	-0.0276(0.0358)	-0.0248(0.0196)
(0.5;0.5)	0.0543(0.0326)	-0.0393(0.0399)	-0.0588(0.0413)	0.0241(0.0506)	0.0305(0.0311)
(0.5;0.7)	0.0791(0.0607)	-0.0660(0.0493)	-0.0886(0.0496)	0.0062(0.0907)	0.0398(0.0540)
(3;0.3)	0.2200(0.5389)	-0.2105(0.8753)	-0.3648(1.0143)	0.2775(1.5466)	0.1097(0.5166)
(3;0.5)	0.2450(1.2195)	-0.3276(1.4115)	-0.3994(1.4972)	-0.2263(1.5477)	0.0857(1.1945)
(3;0.7)	0.7495(2.8894)	-0.1453(2.1989)	-0.3693(2.0037)	0.4058(4.9698)	0.5128(2.5018)
(5;0.3)	0.0990(1.720)	-0.5505(3.4522)	-0.7949(3.4846)	0.1933(6.9193)	-0.0969(1.793)
(5;0.5)	0.4914(2.812)	-0.4508(3.7212)	-0.6585(4.019)	0.1207(4.9554)	0.2129(2.5952)
(5;0.7)	0.9627(7.4471)	-0.4193(5.9371)	-0.6009(5.9504)	0.1559(9.5742)	0.5604(6.4064)
(9;0.3)	0.4267(4.9681)	-0.5202(9.4591)	-0.8507(9.7064)	0.3211(14.0827)	0.0444(4.9115)
(9;0.5)	0.8712(10.029)	-0.7313(12.868)	-1.1503(12.522)	0.2537(20.662)	0.7135(8.955)
(9;0.7)	1.7784(22.3519)	-0.8672(22.9207)	-1.9195(16.8329)	1.5467(67.0252)	1.0810(19.1324)
Parameter p					
$(\beta; p)$	ML	MPS	LS	WLS	LM
(0.5;0.3)	0.0132(0.0587)	0.0991(0.1582)	0.1325(0.1602)	0.0030(0.1862)	0.0583(0.0642)
(0.5;0.5)	-0.1081(0.0774)	-0.0249(0.1977)	0.0057(0.1740)	-0.1202(0.2391)	-0.0669(0.071)
(0.5;0.7)	-0.0843(0.0499)	0.0230(0.0348)	0.0379(0.0332)	-0.0351(0.0649)	-0.0461(0.0401)
(3;0.3)	-0.0685(0.0472)	0.0269(0.1391)	0.0790(0.1423)	-0.1375(0.2434)	-0.0213(0.0451)
(3;0.5)	-0.0616(0.0660)	0.0430(0.0826)	0.0595(0.0781)	0.0061(0.1234)	-0.0211(0.0627)
(3;0.7)	-0.11289(0.0602)	-0.0011(0.0394)	0.0272(0.0312)	-0.0942(0.1646)	-0.0728(0.0471)
(5;0.3)	0.0139(0.0572)	0.0800(0.1603)	0.1266(0.1525)	-0.0999(0.5501)	0.0581(0.0616)
(5;0.5)	-0.0971(0.0706)	0.0059(0.1052)	0.0371(0.1028)	-0.0841(0.1535)	-0.0478(0.0591)
(5;0.7)	-0.1128(0.0652)	-0.0073(0.0500)	0.0056(0.0484)	-0.0590(0.0813)	-0.0710(0.0506)
(9;0.3)	-0.0771(0.0639)	-0.0257(0.1542)	0.0169(0.1397)	-0.1559(0.2997)	-0.0231(0.0557)
(9;0.5)	-0.1059(0.0765)	-0.0087(0.1064)	0.0270(0.0933)	-0.1029(0.2072)	-0.1110(0.0600)
(9;0.7)	-0.0877(0.0535)	0.0123(0.0577)	0.0629(0.0277)	-0.1108(0.2153)	-0.0465(0.0392)

Table 2: The Bias and MSE(Parenthesis) for $n = 100$.

Parameter β					
$(\beta; p)$	ML	MPS	LS	WLS	LM
(0.5;0.3)	0.0054(0.0101)	-0.0406(0.0147)	-0.0574(0.0163)	-0.0043(0.0183)	-0.0043(0.0104)
(0.5;0.5)	0.0328(0.0223)	-0.0296(0.0226)	-0.0435(0.0245)	0.0170(0.0275)	0.0197(0.0215)
(0.5;0.7)	0.0378(0.0301)	-0.0507(0.0287)	-0.0704(0.0296)	0.0290(0.0525)	0.0213(0.0291)

(3;0.3)	0.0014(0.3670)	-0.2907(0.5283)	-0.3662(0.6027)	-0.0681(0.5744)	-0.0672(0.3728)
(3;0.5)	0.2657(0.8309)	-0.0903(0.8510)	-0.2075(0.8508)	0.2360(1.3835)	0.1795(0.7920)
(3;0.7)	0.2416(1.0371)	-0.2950(0.9457)	-0.3910(1.0515)	0.0665(1.6049)	0.1339(0.9840)
(5;0.3)	0.2414(0.9986)	-0.2183(1.1648)	-0.3925(1.3427)	0.2250(1.7647)	0.1274(0.9590)
(5;0.5)	0.2104(1.3685)	-0.4412(1.5442)	-0.6155(1.7667)	0.1150(2.0980)	0.0727(1.2722)
(5;0.7)	0.5745(3.1742)	-0.3316(2.5814)	-0.5563(2.7443)	0.4390(5.082)	0.3883(2.9017)
(9;0.3)	0.1456(3.1205)	-0.7007(4.1331)	-0.9248(4.4103)	-0.1189(5.5615)	-0.1404(2.7674)
(9;0.5)	0.5902(5.7176)	-0.5188(5.5191)	-0.7471(5.8137)	-0.1209(7.2110)	0.2305(4.7879)
(9;0.7)	0.6339(10.4726)	-1.0320(10.7229)	-1.3823(11.1817)	0.5308(15.8358)	0.3110(8.9796)
Parameter					
($\beta; p$)	ML	MPS	LS	WLS	LM
(0.5;0.3)	-0.0226(0.0397)	0.0514(0.0628)	0.0864(0.0657)	-0.0198(0.0841)	0.0022(0.0397)
(0.5;0.5)	-0.0600(0.0507)	0.0186(0.0460)	0.0371(0.0488)	-0.0470(0.0673)	-0.0370(0.0474)
(0.5;0.7)	-0.0526(0.0260)	0.0149(0.0184)	0.0312(0.0182)	-0.0528(0.0431)	-0.0354(0.0233)
(3;0.3)	0.0075(0.0401)	0.0852(0.0630)	0.1093(0.0667)	0.0120(0.0783)	0.0350(0.0406)
(3;0.5)	-0.0878(0.0529)	-0.0202(0.0770)	0.0084(0.0776)	-0.1041(0.1170)	-0.0613(0.0469)
(3;0.7)	-0.0390(0.0201)	0.0268(0.0149)	0.0372(0.0160)	-0.0232(0.0285)	-0.0225(0.0182)
(5;0.3)	-0.0418(0.0430)	0.0317(0.0554)	0.0661(0.0617)	-0.0582(0.0863)	-0.0125(0.0405)
(5;0.5)	-0.0501(0.0358)	0.0336(0.0336)	0.0594(0.0352)	-0.0468(0.0613)	-0.0245(0.0309)
(5;0.7)	-0.0487(0.0203)	0.0179(0.0141)	0.0366(0.0143)	-0.0460(0.0320)	-0.0300(0.0171)
(9;0.3)	-0.0412(0.0408)	0.0409(0.0518)	0.0672(0.0525)	-0.0280(0.0798)	0.0024(0.0313)
(9;0.5)	-0.0710(0.0395)	0.0118(0.0311)	0.0318(0.0310)	-0.0238(0.0495)	-0.0337(0.0279)
(9;0.7)	-0.0371(0.0260)	0.0303(0.0212)	0.04561(0.0208)	-0.0379(0.0383)	-0.0172(0.0201)

Table 3: The Bias and MSE(Parenthesis) for $n=1000$.

Parameter					
($\beta; p$)	ML	MPS	LS	WLS	LM
(0.5;0.3)	0.0030(0.0013)	-0.0055(0.0013)	-0.0109(0.0015)	0.0041(0.0017)	0.0018(0.0013)
(0.5;0.5)	0.0051(0.0015)	-0.0051(0.0014)	-0.0108(0.0016)	0.0047(0.0019)	0.0036(0.0014)
(0.5;0.7)	0.0035(0.0031)	-0.0108(0.0032)	-0.0165(0.0035)	-0.0013(0.0046)	0.0015(0.0031)
(3;0.3)	0.0194(0.0586)	-0.0306(0.0585)	-0.0506(0.0638)	-0.0054(0.0706)	0.0094(0.0564)
(3;0.5)	0.0053(0.0601)	-0.0558(0.0629)	-0.0921(0.0738)	0.0106(0.0799)	-0.0029(0.0588)
(3;0.7)	-0.0008(0.0862)	-0.0859(0.0922)	-0.1257(0.1080)	-0.0148(0.1523)	-0.0120(0.0877)
(5;0.3)	-0.0044(0.1425)	-0.0906(0.1505)	-0.1389(0.1917)	-0.0303(0.1721)	-0.0197(0.1357)
(5;0.5)	0.0604(0.1814)	-0.0442(0.1796)	-0.0940(0.2129)	0.0275(0.2187)	0.0461(0.1735)
(5;0.7)	-0.0078(0.2519)	-0.1521(0.2755)	-0.2306(0.3403)	0.0010(0.4262)	-0.0222(0.2419)
(9;0.3)	-0.0159(0.4249)	-0.1704(0.4528)	-0.2747(0.5151)	-0.0317(0.6417)	-0.0809(0.3544)
(9;0.5)	0.0551(0.4066)	-0.1318(0.4164)	-0.2537(0.4590)	0.0628(0.7222)	0.0200(0.3725)
(9;0.7)	0.0338(0.5766)	-0.2238(0.6310)	-0.3601(0.8473)	0.0039(0.9020)	0.0033(0.4901)
Parameter					
($\beta; p$)	ML	MPS	LS	WLS	LM
(0.5;0.3)	-0.0054(0.0071)	0.0111(0.0069)	0.0221(0.0074)	-0.008(0.0092)	-0.0023(0.0071)
(0.5;0.5)	-0.0086(0.0030)	0.0047(0.0028)	0.0130(0.0032)	-0.0084(0.0037)	-0.0059(0.0029)
(0.5;0.7)	-0.0048(0.0021)	0.0055(0.0021)	0.0096(0.0023)	-0.0015(0.0028)	-0.0032(0.0021)
(3;0.3)	-0.0096(0.0078)	0.0066(0.0075)	0.0128(0.0082)	-0.0016(0.0084)	-0.0054(0.0073)
(3;0.5)	-0.0053(0.0032)	0.0080(0.0032)	0.0167(0.0036)	-0.0071(0.0045)	-0.0026(0.0031)
(3;0.7)	-0.0003(0.0016)	0.0098(0.0016)	0.0147(0.0019)	0.0007(0.0026)	0.0012(0.0016)

(5;0.3)	-0.0055(0.0065)	0.0112(0.0064)	0.0201(0.0076)	-0.00075(0.0029)	-0.0005(0.0023)	0.015
(5;0.5)	-0.0094(0.0036)	0.0042(0.0034)	0.0107(0.0071)	-0.0005(0.0023)	-0.0004(0.0017)	0.010
(5;0.7)	-0.0002(0.0019)	0.0101(0.0019)	0.0161(0.0023)	-0.0014(0.0030)	0.0012(0.0018)	
(9;0.3)	-0.0021(0.0064)	0.0145(0.0064)	0.0260(0.0068)	-0.0044(0.0095)	0.0077(0.0050)	

In view of MSE, it can be seen that MSE values of all estimators increases while β increases. Additionally, when β is greater than

5, MSE values of estimators of ρ decreases while ρ increases for $n = 100$. And while β increases, MSE values of estimators of ρ decreases. For MSE values of estimators of ρ decreases while ρ increases. When $n = 1000$, estimation methods show ML, MPS and LM, show better performance than LS and WLS estimation methods, according to MSE criterions. In situations that are mentioned above, decreasing of MSE values indicates that estimators show better performance.

Table 4: The time intervals (in days) between coal mine accidents ($n = 109$).

378	96	59	108	54	275	498	228	217	19	156	36	124	61	188
217	78	49	271	120	329	47	15	50	1	233	113	17	131	208
275	330	129	31	120	13	28	32	1205	182	517	20	312	1630	215
203	189	22	23	644	255	1613	66	171	29	11	176	345	61	151
467	195	54	291	145	217	137	55	20	78	361	871	224	326	4
75	7	4	93	81	99	312	48	56	181	36	36	59	69	69
286	326	354	123	390	348	338	37	19	37	315	27	38	15	15
72	745	336	19											

Applications
 In this section, we illustrate the applicability of EG distribution by considering two different data sets. First of them is a data set which is first obtained by [12]. This data set obtains the time intervals (in days) between coal mine accidents concluded death

of 10 or more men, and the data set consists of 109 observations. It is named as coal mine data in literature. The coal mine data is given in Table 4. The second data set shows the duration of hospitalization (on daily basis) in a neurosurgical service of a research hospital in Turkey for 68 male patients. The data set is given in Table 5.

Table 5: Duration of hospitalization (on daily basis) in a neurosurgical service of a research hospital in Turkey for male patients ($n = 68$).

1	1	1	1	1	1	2	3	4	4	4	5	6	6	6	7
7	7	7	8	8	8	9	10	11	11	11	12	12	13	14	17
18	19	19	19	20	21	23	24	24	25	26	31	32	32	33	39
39	41	41	42	47	51	54	58	61	67	74	77	80	81	82	84
95	120	125	150												

Akaike information criterion (AIC) and Kolmogorov-Smirnov (KS) statistics which are good indicators to examine the fitting performance of distributions. Mathematical definitions of AIC is $AIC = -2LL + 2k$ where k is the number of parameters and LL is the log-likelihood function value of considered model. KS statistics is $KS = \sup_x |F_n(x) - F(x)|$ where $F_n(x)$ denotes the empirical distribution based on ordered data and F is the distribution function of considered distribution. It is known that, the distribution having smaller AIC and KS statistics values shows a better modeling capacity.

Parameter estimates of EG distribution are $\hat{\rho} = 0.5602$, $\hat{\beta} = 0.0026$ for 109 coal mine data set. Calculated K-S statistic and AIC criteria for EG, Weibull, gamma, exponential distributions can be seen in Table 6. From the Table 6, it is seen that EG distribution gives the best fit to coal mine data in five models. In Figure 3, we give the graph of histogram with the fitted pdf of EG distribution for coal mine data.

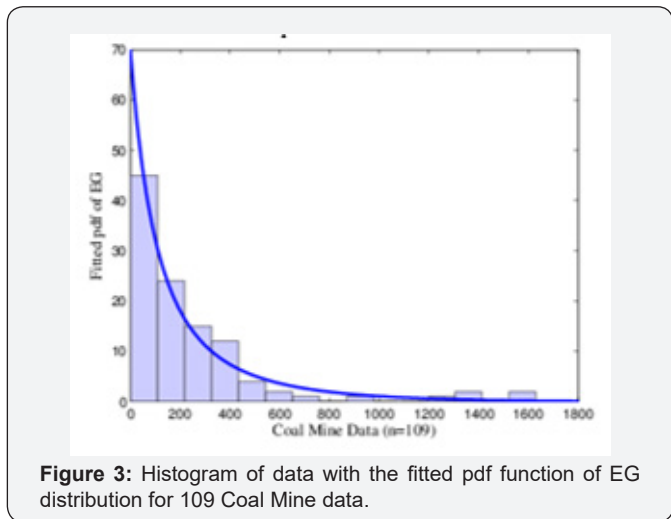


Figure 3: Histogram of data with the fitted pdf function of EG distribution for 109 Coal Mine data.

Table 6: Kolmogorov-Smirnov statistics, -2log-likelihood and AIC values obtained from the fit of each of the five distributions for coal mine data set ($n = 109$).

Distribution	KS	-2LL	AIC
EG	0.0765	1407.832	1411.832

In this study, these data sets are modeled by EG, Weibull, lognormal, gamma and exponential distributions. The parameters of EG distribution are estimated using the ML method. Some criteria are used for comparison among distributions. These are

Weibull	0.0777	1409.274	1413.274
Log-normal	0.0848	1414.696	1418.696
Gamma	0.0830	1410.900	1414.900
Exponential	0.0906	1413.636	1415.636

We fit 68 male patients in neurosurgical service data to EG distribution and calculated parameter estimates as $\hat{\rho} = 0.4173$, $\hat{\beta} = 0.0246$. Calculated KS statistic and AIC criteria for EG, Weibull, gamma, exponential distributions can be seen in Table 7 for neurosurgical service data. According to this results, EG distribution gives the best fit to neurosurgical service data in five models. In Figure 4, we give the graph of histogram with the fitted pdf of EG distribution for neurosurgical service data. This graph supports the conclusion.

Table 7: Kolmogorov-Smirnov statistics, -2log-likelihood and AIC values obtained from the fit of each of the five distributions for male patients in neurosurgical service data set ($n = 68$).

Distribution	KS	-2LL	AIC
EG	0.0547	600.291	604.291
Weibull	0.0679	600.896	604.896
Log-normal	0.0716	604.756	608.756
Gamma	0.0752	601.138	605.138
Exponential	0.1034	602.986	604.986

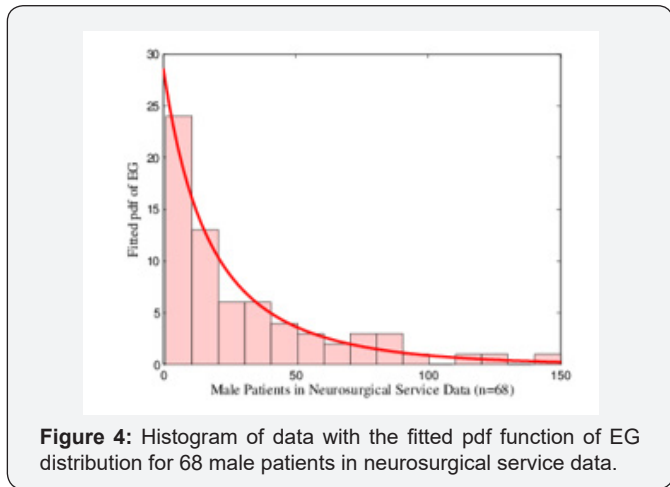


Figure 4: Histogram of data with the fitted pdf function of EG distribution for 68 male patients in neurosurgical service data.

In this section we show that the EG distribution given by Equation (1) shows better performance with the lowest AIC and KS statistics for fitting the data sets than the other distributions.

Conclusion and Discussion

In this paper, some properties of EG distribution and used parameter estimation methods have been introduced. Parameter estimates of EG distribution have been obtained by the estimation methods which are not used before. We have compared the efficiencies of the estimators of EG distribution parameters in

terms of Bias and MSE criteria via simulation study. The availability of EG distribution has been illustrated for two lifetime data sets. These data sets are coal mine data and neurosurgical service data which are two different study fields. It can be said that the EG distribution can be modeled better than the gamma, lognormal and Weibull distributions in terms of AIC and KS criteria for the data sets. EG distribution can be used also for different survival datasets and parameters of EG distribution can be estimated via estimators which is given in the study. In addition, Parameters of EG distributions can be estimated with also different estimator methods. And can be compared with the estimator methods which is mentioned in this paper.

References

- Adamidis K, and Loukas S (1998) A lifetime distribution with decreasing failure rate. *Statist Probab Lett* 39(1): 35-42.
- Adamidis K, Dimitrakopoulou T, Loukas S (2005) On an extension of the exponential geometric distribution. *Statist Probab Lett* 73(3): 259-269.
- Al-Zahrani B, Sagor H (2014) The Poisson-Lomax distribution. *Revista Colombiana de Estadística* 37(3): 223-243.
- Barreto-Souza W, Cribari-Neto F (2009) A generalization of the exponential-Poisson distribution. *Statist. Probab Lett* 79(24): 2493-2500.
- Cancho VG, Louzada-Neto F, Barriga GDC (2011) The Poisson-exponential lifetime distribution. *Comput. Statist Data Anal* 55(1): 677-686.
- Cheng RCH, Amin NAK (1983) Estimating parameters in continuous univariate distributions with a shifted origin. *JR Statist Soc B* 45(1): 394-403.
- Dempster N Laird, Rubin D (1977) Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc B* 39(1): 1-38.
- Gupta RD, and Kundu D (2000) Generalized exponential distribution: different method of estimations. *J Statist Comput Simul* 69(4) 315-337.
- Hosking JRM (1990) L-moment: analysis and estimation of distributions using linear combinations of order statistic. *J R Statist Soc B* 52(1): 105-124.
- Kus C (2007) A new lifetime distribution. *Comput Statist Data Anal* 51(9): 4497-4509.
- Louzada F, Roman M, Cancho VG (2011) The complementary exponential geometric distribution: Model, properties, and comparison with its counter part. *Comput Statist Data Anal* 55(8): 2516-2524.
- Maguir BA, Pearson ES, Wynn AHA (1952) The time intervals between industrial accidents. *Biometrika* 39(1): 168-180.
- Mahmoudi E, Sepahdar A (2013) Exponentiated Weibull-Poisson distribution: Model, properties and applications. *Math Computers Simul* 92: 76-97.
- Rezaeia S, Nadarajah S, Tahghighniac N (2013) A new three-parameter lifetime distribution. *J Theor App Statist* 47(4): 835-860.
- Ristić MM, Nadarajah S (2014) A new lifetime distribution. *J Statist Comput Simul* 84(1): 135-150.
- Singh U, Kumar M (2014) Estimation for the parameter of Poisson-exponential distribution under Bayesian paradigm. *J Data Sci* 12(2014): 157-173.



This work is licensed under Creative Commons Attribution 4.0 License
DOI: [10.19080/BBOAJ.2018.08.555735](https://doi.org/10.19080/BBOAJ.2018.08.555735)

**Your next submission with Juniper Publishers
will reach you the below assets**

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats
(Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission
<https://juniperpublishers.com/online-submission.php>