Journal of Econometrics and Statistics
Vol. 1, Issue 1, 2021, pp. 61-73

URL : www.arfjournals.com/jes

GOODNESS-OF-FIT TEST AND POWER COMPARISON FOR THE NEW TWO-PARAMETER DISTRIBUTION WITH UNKNOWN PARAMETERS AND CENSORSHIP

AIDI KHAOULA[1*] and Seddik-Ameur N[1]
[1]*Laboratory of probability and statistics LaPS, University BadjiMokhtar- Annaba- Algeria*
[*]*Corr. Author Email: khaoula.aidi@yahoo.fr*

Article History
Received : 13 April 2021
Revised : 19 April 2021
Accepted : 29 April 2021
Published : 2 September 2021

To cite this paper
AIDI KHAOULA & Seddik-Ameur N. (2021). Goodness-of-Fit Test and Power Comparison for the New two-Parameter Distribution with Unknown Parameters and Censorship. *Journal of Econometrics and Statistics.* 1(1), 61-73.

JEL Classification
62F03-62G05-62G10-62N05

Abstract: With only two parameters the New Distribution (ND) introduced recently by Doostmoradi (2018) is very flexible for modeling litetime data because the failure rate function can have different shapes (increasing, decreasing and unimodal). This work is devoted to the maximum likelihood estimation of the unknown parameters and the construction of goodness-of-fit test statistics for this model when data are right censored. Also, a comparative study is provided to distinguish between this model (ND) and the competing distributions, namely the exponentiatedWeibull (EW), modified Weibull (MW), extended generalized Lindley distribution (EGL), generalized Lindley (GL), power Lindley (PL), and the inverse Lindley (IL). An important simulation study was carried out and theoretical results obtained through this study are applied to real data sets from reliability and survival analysis.

Keywords: Censored data, Fisher information matrix, maximum likelihood estimation, modified goodness-of-fit test.

1. Introduction

Until the last two decades, existing models in the statistical literature could not describe properly the collected observations obtained from experiences, so researchers tried to provide new ones. Different methods are developed to generalize classical models particularly by adding new parameters to the baseline distributions to obtain more flexibility. Except the inverse Lindley (IL) introduced by Sharma et al, (2015), the proposed models have three and more parameters such as modified Weibull (MW) (Lai *et al.* 2003), generalized Lindley (GL) (Zakerzadeh and Dolati, 2009), power Lindley (PL) (Ghitany *et al.*, 2013), generalized gamma (GG) (Stacy *et al.*,1962), extended generalized Lindley distribution (EGL) (Torabi *et al.*, 2014), Entezar distribution (ENT)

(Doostmoradi *et al.*, 2016), exponential flexible Weibull extension (EFWE) (Beih *et al.*, 2017). Recently, Doostmoradi (2018) proposed a very interesting flexible distribution so-called a new two-parameter distribution (ND). With only two parameters, this model can have an increasing, decreasing and unimodal failure rates. Using Akaike Information criterion (AIC), likelihood ratio (LR), bayesian information criterion (BIC) and rate function plots to fit reliability data to this new distribution, the author showed that this distribution can describe reliability data better than the exponentiated Weibull (EW) introduced by Mudholkar *et al.* (1995), modified Weibull MW, extended generalized Lindley EGL, Entezar ENT, generalized Lindley GL, power Lindley PL, generalized gamma GG, inverse Lindley IL, exponential flexible Weibull extension EFWE and other distributions.

The flexibility of this new model allows it to certainly model reliability and survival analysis data which are often censored. In this work, we firstly determine the maximum likelihood estimators of the unknown parameters of this model when data are right censored. Then we propose a criteria goodness-of-fit test which takes into account the unknown parameters and censorship. The construction of this statistic is based on the approach of Bagdonavicius and Nikulin (2011). Also, a comparative study is provided to distinguish between ND distribution and all its competing distributions cited above.

Theoretical results are confirmed by an important simulation study. We generated samples of different sizes and different percentage of right censoring from this model. Then, maximum likelihood estimators and mean square errors of the unknown parameters of all these samples are computed. Estimated information matrix and criteria goodness-of-fit test Y^2 are also computed and empirical proportions of rejection of the null hypothesis H_0 for 1%, 5% and 10% levels of significance are compared to the theoretical ones. A Comparative study between the ND distribution and some alternative distributions namely EW, MW, EGL, GL, PL, and IL is also evaluated.

2. New two-parameter Distribution ND

Characterized by two parameters, the distribution function of the new-two parameter distribution (ND) introduced by Doostmoradi (2018) is

$$F(t) = 1 - (1 + \alpha t^{\lambda})e^{-\alpha t^{\lambda}}$$

The density function, reliability and rate functions are respectively

$$f(t) = \alpha^2 \lambda t^{2\lambda - 1} e^{-\alpha t^{\lambda}},\ \alpha, \lambda > 0,\ t > 0$$

$$S(t) = (1 + \alpha t^{\lambda})e^{-\alpha t^{\lambda}}$$

$$h(t) = \frac{\alpha^2 \lambda t^{2\lambda - 1} e^{-\alpha t^{\lambda}}}{1 + \alpha t^{\lambda}}$$

and the cumulative rate function is

$$H(t) = -\ln S(t; \alpha, \lambda) = \alpha t^{\lambda} - \ln(1 + \alpha t^{\lambda})$$

After studying statistical properties, maximum likelihood estimators and estimated information matrix, the author used Akaike Information criterion (AIC), likelihood ratio (LR), bayesian information criterion (BIC) and rate function plots to fit reliability data to this new distribution in complete data case. Reliability data set were used to show the flexibility of this model. For more details, one can see Doostmoradi (2018).

3. Maximum Likelihood Estimation with Right Censorship

Let us consider $T = (T_1, T_2, \ldots, T_n)^T$ a sample from ND distribution with the parameter vector $\theta = (\alpha, \lambda)^T$ which can contain right censored data with fixed censoring time τ. Each T_i can be written as $T_t = (t_i, \delta_i)$ where

$$\delta_i = \begin{cases} 0 & \text{if } t_i \text{ is a censoring time} \\ 1 & \text{if } t_i \text{ is a failure time} \end{cases}$$

The right censoring is assumed to be non-informative, so the log-likelihood function can be written as:

$$L_n(\theta) = \sum_{i=1}^{n} \delta_i \ln f(t_i, \theta) + \sum_{i=1}^{n} (1-\delta_i) \ln S(t_i, \theta)$$

$$L_n(\theta) = \sum_{i=1}^{n} \delta_i [2\ln\alpha + \ln\lambda + (2\lambda - 1)\ln(t_i) - \ln(1+\alpha t_i^{\lambda})] + \sum_{i=1}^{n} \ln(1+\alpha t_i^{\lambda}) - \alpha t_i^{\lambda}$$

The maximum likelihood estimators θ and λ of the unknown parameters α and λ are derived from the nonlinear following score equations:

$$\frac{\partial L}{\partial \alpha} = \sum_{i=1}^{n} \delta_i \left[\frac{2}{\alpha} - \frac{t_1^2}{1+\alpha t_i^{\lambda}} \right] + \sum_{i=1}^{n} \frac{t_1^{\lambda}}{1+\alpha t_i^{\lambda}} = 0$$

$$\frac{\partial L}{\partial \alpha} = \sum_{i=1}^{n} \delta_i \left[\frac{1}{\lambda} - \frac{(2+\alpha t_i^{\lambda})\ln(t_i)}{1+\alpha t_i^{\lambda}} \right] + \sum_{i=1}^{n} \frac{\alpha t_1^{\lambda} \ln(t_i)}{1+\alpha t_i^{\lambda}} = 0$$

The explicit form of $\hat{\alpha}$ and $\hat{\lambda}$ cannot be obtained, so one can use numerical methods.

4. Estimated Fisher information matrix $\hat{I}$

The components of the estimated information matrix $I = (1_{ij})_{2\times 2}$ are obtained by

$$\hat{l}_{11} = \frac{1}{11} \sum\nolimits_{i=1}^{n} \delta_i \left(\frac{2}{\alpha} - \frac{t_i^2}{1+\alpha t_l^{\lambda}} \right)^2$$

$$\hat{l}_{12} = \frac{1}{n} \sum_{i=1}^{n} \delta_i \left(\frac{1}{\lambda} - \frac{(2+\alpha t_i^{\lambda})\ln(t_i)}{1+\alpha t_i^{\lambda}} \right)^2$$

$$\hat{l}_{12} = \frac{1}{n}\sum_{i=1}^{n}\delta_i\left(\frac{2}{\alpha} - \frac{t_i^{\lambda}}{1+\alpha t_i^{\lambda}}\right)\left(\frac{1}{\lambda} - \frac{(2+\alpha t_i^{\lambda})\ln(t_i)}{1+\alpha t_i^{\lambda}}\right)$$

where the parameters α and λ are replaced by their maximum likelihood estimators (MLEs) $\hat{\alpha}$ and $\hat{\lambda}$.

5. Test Statistic for Right Censored Data

Let $T_1, \ldots, T_n$ be n i.i.d. random variables grouped into k classes I_j. To assess the adequacy of a parametric model F_0

$$H_0: P(T_i \le t \mid H_0) = F_0(t;\theta),\ t \ge 0,\ \theta = (\theta_1, \ldots, \theta_s)^T \in \Theta \subset R^s$$

when data are right censored and the parameter vector β is unknown, Bagdonavicius and Nikulin (2011) proposed a statistic test Y^2 based on the vector

$$Z_j = \frac{1}{\sqrt{n}}(U_j - e_j),\ j = 1, 2, \ldots, k),\ \text{with } k = s.$$

This one represents the differences between observed and expected numbers of failures (U_j and e_j) to fall into these grouping intervals $I_j = (a_{j-1}, a_j]$ with $a_0 = 0$, $a_k = \tau$, where τ is a finite time. The authors considered a_j as random data functions such as the k intervals chosen have equal expected numbers of failures e_j.

The statistic test Y^2 is defined by

$$Y^2 = Z^T\hat{\Sigma}^{-}Z = \sum_{i=1}^{k}\frac{(U_j - e_j)^2}{U_j} + Q$$

where $Z = (Z^1, \ldots, Z_k)^T$ and $\hat{\Sigma}^{-}$ is a generalized inverse of the covariance matrix $\hat{\Sigma}$ and

$$Q = W^T\hat{G}^{-}W,\ \hat{A}_j = \frac{U_j}{n},\ U_j = \sum_{i:T_i\in I_j}\delta_i$$

$$W = (W_1, \ldots, W_s)^T,\ \hat{G} = [\hat{g}_{ll'}]_{s\times s},\ \hat{g}_{ll'} = \hat{l}_{ll'} - \sum_{j=1}^{k}\hat{C}_{lJ}\hat{G}_{l'J}\hat{A}_J^{-1}$$

$$\hat{C}_{lj} = \frac{1}{n}\sum_{i=T_i\in I_j}\delta_i\frac{\partial \ln nh(t_i,\hat{\theta})}{\partial\theta},\ \hat{l}_{ll'} = \frac{1}{n}\sum_{i=1}^{n}\delta_i\frac{\partial \ln nh(t_i,\hat{\theta})}{\partial\theta_l}\frac{\partial \ln nh(t_i,\hat{\theta})}{\partial\theta_{l'}}$$

$$\hat{W}_l = \sum_{j=1}^{k}\hat{C}_{lJ}\hat{A}_J^{-1}Z_j l,\ l' = 1, \ldots, s$$

$\hat{\theta}$ is the maximum likelihood estimator of θ on initial non-grouped data.

Under the null hypothesis H_0, the limit distribution of the statistic Y^2 is a chi-square with k = rank(Σ) degrees of freedom. The description and applications of modified chi-square tests are discussed in Voinov *et al.* (2013).

The interval limits a_j for grouping data into j classes I_j are considered as data functions and defined by

$$\hat{a}_j = H^{-1}\left(\frac{E_j - \Sigma_{l=1}^{i-1} H(t_l, \hat{\theta})}{n-i+1}, \hat{\theta}\right), \ \hat{a}_j = \max(T_{(n)}, \tau)$$

such as the expected failure times e_j to fall into these intervals are $e_j = \frac{E_k}{k}$ for any j, with $E_r = \Sigma_{l=1}^{i-1} H(t_l, \theta)$. The distribution of this statistic test Y_n^2 is chi-square (see Voinov *et al.*, 2013).

6. Criteria Test for the New Distribution

For testing the null hypothesis H_0 that data belong to the ND model, we construct a modified chi-squared type goodness-of-fit test based on the statistic Y^2. Suppose that τ is a finite time, and observed data are grouped into $k > s$ sub-intervals $I_j = (a_{j-1}, a_j]$ of $[0, \tau]$. The limit intervals a_j are considered as random variables such that the expected numbers of failures in each interval I_j are the same, so the expected numbers of failures are obtained as

Estimated Matrix $\hat{W}$ et $\hat{C}$

The components of the estimated matrix $\hat{W}$ are derived from the estimated matrix $\hat{C}$ which is given by:

$$\hat{C}_{1j} = \frac{1}{n}\sum_{i=t_i \in I_j}^{n} \delta_i \left[\frac{2}{\alpha} - \frac{t_i^{\lambda}}{1+\alpha t_i^{\lambda}}\right]$$

$$\hat{C}_{2j} = \frac{1}{n}\sum_{i=t_i \in I_j}^{n} \delta_i \left[\frac{1}{\lambda} - \frac{(2-\alpha t_i^{\lambda})\ln(t_i)}{1+\alpha t_i^{\lambda}}\right]$$

And

$$\hat{W}_l = \sum_{j=1}^{k} \hat{C}_{lJ} \hat{A}_J^{-1} Z_j l, l' = 1, 2 \ j = 1, ..., k$$

Therefore the quadratic form of the test statistic can be obtained easily:

7. Simulations

7.1 Maximum Likelihood Estimation

We generated N = 10,000 right censored samples with different sizes (n = 15, 25, 50, 130, 350, 500) from the ND model with parameters α = 2 and λ = 1.5. Using *R* statistical software and the Barzilai-

Borwein (BB) algorithm (Ravi, 2009), we calculate the maximum likelihood estimators of the unknown parameters and their mean squared errors (S.M.E). The results are given in Table 1.

Table1: Mean Simulated values of MLEs $\hat{\alpha}$ and $\hat{\lambda}$ and their corresponding square mean errors

$N = 10,000$	$n_1 = 15$	$n_2 = 25$	$n_3 = 50$	$n_4 = 130$	$n_5 = 350$	$n_6 = 500$
$\hat{\alpha}$	1.8645	1.8857	1.9345	1.9538	1.9754	2.0006
S.M.E	0.0046	0.0031	0.0029	0.0024	0.0018	0.0009
Bias($\hat{\alpha}$)	–0.0595	0.0391	–0.0274	0.0187	–0.0084	-0.0010
$\hat{\lambda}$	1.8546	1.7925	1.7285	1.6845	1.6294	1.5098
S.M.E	0.0058	0.0042	0.0035	0.0028	0.0020	0.0012
Bias($\hat{\lambda}$)	0.0533	–0.0350	0.0212	-0.0147	0.0058	–0.0015

The maximum likelihood estimated parameter values, presented in Table 1, agree closely with the true parameter values.

7.2 Test Statistic Y^2

Using right censored simulated samples with different percentage (15% and 30%) of right censoring and different sizes (n = 25, 50, 130, 350, 500) we calculate the test statistic Y^2 for each sample with respect to the ND model and we compare the obtained values with the theoretical levels of significance (ε = 0.01, 0.05, 0,1). The results are summarized in Table 2 and Table 3.

Table 2: Simulated levels of significance for Y^2 against their theoretical values (15% of censorship)

$N = 10,000$	$n_1 = 25$	$n_2 = 50$	$n_3 = 130$	$n_4 = 350$	$n_5 = 500$
$\varepsilon = 1\%$	0.0045	0.0054	0.0096	0.0101	0.0102
$\varepsilon = 5\%$	0.0488	0.0492	0.0495	0.0498	0.0504
$\varepsilon = 10\%$	0.0984	0.0991	0.0998	0.01012	0.1004

Table 3: Simulated levels of significance for Y^2 against their theoretical values (30% of censorship)

$N = 10,000$	$n_1 = 25$	$n_2 = 50$	$n_3 = 130$	$n_4 = 350$	$n_5 = 500$
$\varepsilon = 1\%$	0.0092	0.0093	0.0095	0.0098	0.0101
$\varepsilon = 5\%$	0.0478	0.0489	0.0492	0.0495	0.0499
$\varepsilon = 10\%$	0.0986	0.0992	0.0996	0.0998	0.1001

As we can see empirical proportions of rejection of the null hypothesis H_0 for ε = 1%, 5% and 10% levels of significance for all sample sizes and for different percentage of censorship (Table 2 and Table 3) are very close to the theoretical ones. Therefore, the test statistic Y^2 proposed in this work can be applied to fit data to ND.

7.3 Power Study

To evaluate the power of the test statistic Y^2 proposed whether data fits well to the ND, we have considered some competing distributions proposed by the author in his study, namely the exponentiated Weibull (EW), modified Weibull (MW), extended generalized Lindley distribution (EGL), generalized Lindley (GL), power Lindley (PL), and the inverse Lindley (IL). So, we generate N = 10,000 random samples under alternative hypotheses. The power analysis has been carried out for sample sizes (n = 25, 50, 130, 350, 500) at levels of significance (ε = 0.01, 0.05, 0.1). Using maximum likelihood estimates, the powers of Y^2 test for testing about belonging of the samples to ND distribution against that sample is from EW, MW, EGL, GL, PL, and IL distributions are given in Tables 4.a, b, c, d, e, f.

Table 4a: Power of Y^2 for ND against EW

N = 10,000	$n_1 = 25$	$n_2 = 50$	$n_3 = 130$	$n_4 = 350$	$n_5 = 500$
ε = 1%	0.4856	0.5124	0.6147	0.7487	0.7945
ε = 5%	0.6785	0.6954	0.7458	0.8125	0.8952
ε = 10%	0.7984	0.8120	0.8523	0.9525	0.9823

Table 4b: Power of Y^2 for ND against MW

N = 10,000	$n_1 = 25$	$n_2 = 50$	$n_3 = 130$	$n_4 = 350$	$n_5 = 500$
ε = 1%	0.4236	0.5349	0.6451	0.7124	0.7356
ε = 5%	0.5124	0.7452	0.7321	0.8325	0.8468
ε = 10%	0.6252	0.8021	0.8451	0.9354	0.9745

Table 4c: Power of Y^2 for ND against EGL

N = 10,000	$n_1 = 25$	$n_2 = 50$	$n_3 = 130$	$n_4 = 350$	$n_5 = 500$
ε = 1%	0.4125	0.5326	0.6954	0.7356	0.8421
ε = 5%	0.5684	0.7584	0.7954	0.8694	0.9162
ε = 10%	0.7845	0.8151	0.8456	0.9123	0.9725

Discussion

On the basis of tables above, it is observed that at level of significance 0.10, all test power values for Y^2 are higher than 80%, then we can distinguish ND from the competing distributions for all

Table 4d: Power of Y^2 for ND against GL

N = 10,000	$n_1 = 25$	$n_2 = 50$	$n_3 = 130$	$n_4 = 350$	$n_5 = 500$
ε = 1%	0.4784	0.6254	0.7854	0.7954	0.8120
ε = 5%	0.625	0.7854	0.8024	0.8562	0.9014
ε = 10%	0.7541	0.8123	0.9154	0.9512	0.9974

Table 4e: Power of Y^2 for ND against PL

N = 10,000	$n_1 = 25$	$n_2 = 50$	$n_3 = 130$	$n_4 = 350$	$n_5 = 500$
ε = 1%	0.4958	0.6147	0.7245	0.7421	0.8214
ε = 5%	0.6147	0.7689	0.7959	0.8952	0.9436
ε = 10%	0.7845	0.8214	0.9245	0.9214	0.9784

Table 4f: Power of Y^2 for ND against IL

N = 10,000	$n_1 = 25$	$n_2 = 50$	$n_3 = 130$	$n_4 = 350$	$n_5 = 500$
ε = 1%	0.5123	0.6458	0.7154	0.7536	0.8421
ε = 5%	0.6895	0.7561	0.8425	0.8954	0.9541
ε = 10%	0.8124	0.8351	0.9124	0.9541	0.9989

sample sizes. At level of significance 0.05, test power values are about 80% for sample sizes superior or equal to 50. Then we can distinguish ND from the competing distributions for big sizes sample. The power of Y^2 is least when comparing ND and modified Weibull MW (Table 4.b) which means that these distributions are quite close to each other for little sample sizes.

For all levels of significance and for all sample sizes, Power in case of testing goodness-of-fit of ND versus IL distribution is more than in other cases. So, the proposed test statistic Y^2 can detect the difference between the ND and IL distribution with high power and hence a small sample is sufficient to differentiate ND from IL distribution.

As expected, the results given in the tables above show that the new two-parameter distribution ND can be used to fit data better than several models.

8. Data Analysis

In this section, we apply the statistic Y^2 proposed in this work to fit two censored real data sets to the ND model. The first one is from survival analysis and the second one is from reliability.

Example 1:

We consider the bone marrow transplant data (Klein and Moeschberger, 2003) for patients suffering from acute lymphoblastic leukemia. This data consist of time (in days) to death or on study time after anallogenic bone marrow transplant for 38 patients. The bone marrow transplant is a standard treatment for acute leukemia.

Recovery following bone marrow transplantation is a complex process. Immediately following transplantation, patients have depressed platelet counts and have higher hazard rate for the development of infections but as the time passes the hazard decreases. Data are given in Table 5, where * denotes censored observations.

Table 5: Time to death (in days) after bone marrow transplant.

1	86	107	110	122	156	162	172	194	226*
243	262	262	269	276	350*	371	417	418	466
487	526	530*	716	781	996*	1111*	1167*	1182*	1199*
1279	1330*	1377*	1433*	1462*	1496**	1602*	2081*		

We use the statistic test provided above to verify if these data are modeled by the ND distribution. Using BB solve software, we calculate the maximum likelihood estimators of the unknown parameters

$$\hat{\theta} = (\hat{\alpha}, \hat{\lambda})^T = (0.5434, 0.1716)^T.$$

Then, we grouped the observations into k = 5 intervals I_j. The intermediate calculations are given in Table 6.

Table 6: Values of a_j, U_j, $\hat{C}_{1j}$, $\hat{C}_{2J}$, e_j

a_j	170.652	292.521	492.233	1235.215	2081
U_j	7	8	6	9	8
$\hat{C}_{1J}$	0.9512	1.2013	0.3156	1.6234	0.1578
$\hat{C}_{2J}$	–0.2845	–0.1542	-0.0145	0.1243	0.3154
e_j	4.6183	4.6183	4.6183	4.6183	4.6183

The value of the statistic test Y_n^2 is obtained as follows:

$$Y_n^2 = X^2 + Q = 3.1562 + 1.9623 = 5.1185$$

As the value of Y_n^2 = 5.1185 is less than the critical value χ_5^2 = 11.0705 (for significance evel ε = 0.05), so we can say that these data can be fitted by the ND distribution.

We also calculated the test statistics Y_n^2 to fit these data to the competing models. The results are given in table 7.

Table 7: Values of the test statistics for Leukemia Y_n^2 free survival times to fit different alternatives

Modeling distribution	Y_n^2
New two-parameter distribution ND,	5.1185
Exponentiated Weibull (EW)	9.3789
Modified Weibull (MW)	7.2347
Extended generalized Lindley distribution (EGL)	10.1967
GeneralizedLindley (GL)	7.7452
Power Lindley (PL)	8.3124
Inverse Lindley (IL)	11.0.235

From these results, we have showed that the ND distribution fit these data better than the competing distributions. It can be seen clearly in the pdfs plots (Figure 1).

Example 2:

In this example, the data set represent the number of successive failures for the air conditioning system of each member in a fleet of 13 Boeing 720 jet airplanes (1963) and reported in Doostmoradi (2018). The observations are:

194, 413, 90, 74, 55, 23, 97, 50, 359, 50, 130, 487, 57, 102, 15, 14, 10, 57, 320, 261, 51, 44, 9, 254, 493, 33, 18, 209, 41, 58, 60, 48, 56, 87, 11, 102, 12, 5, 14,14, 29, 37, 186, 29, 104, 7, 4, 72,

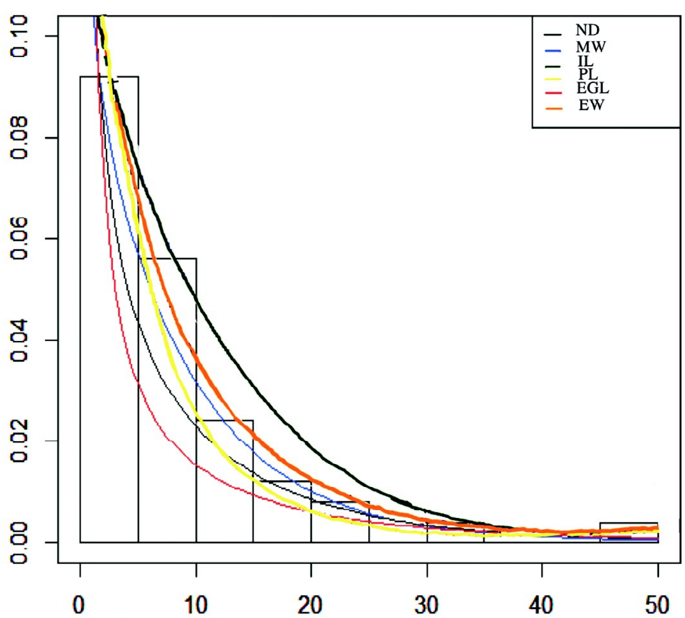

Figure 1

270, 283, 7, 61, 100, 61, 502, 220, 120, 141, 22, 603, 35, 98, 54, 100, 11, 181, 65, 49, 12, 239, 14, 18, 39, 3, 12, 5, 32, 9, 438, 43, 134, 184, 20, 386, 182, 71, 80, 188, 230, 152, 5, 36, 79, 59, 33, 246, 1, 79, 3, 27, 201, 84, 27, 156, 21, 16, 88, 130, 14, 118, 44, 15, 42, 106, 46, 230, 26, 59, 153, 104, 20, 206, 5, 66, 34, 29, 26, 35, 5, 82, 31, 118, 326, 12, 54, 36, 34, 18, 25, 120, 31, 22, 18, 216, 139, 67, 310, 3, 46, 210, 57, 76, 14, 111, 97, 62, 39, 30, 7, 44, 11, 63, 23, 22, 23, 14, 18, 13, 34, 16, 18, 130, 90, 163, 208, 1, 24, 70, 16, 101, 52, 208, 95, 62, 11, 191, 14, 71.

Using different methods, the authors showed that the ND distribution fit these data better than several alternatives. To confirm this fact, we calculated the corresponding test statistic Y^2 of this sample. Under the null hypothesis that these data are modeled by the ND distribution, the maximum likelihood estimators of the unknown parameters obtained by iterative methods are:

$$\hat{\theta} = (\hat{\alpha}, \hat{\lambda})^T = (0.1397, 0.6183)^T.$$

If data are grouped into $k = 6$ intervals I_j, the values of the components of the test statistic are calculated and summarized in Table 8.

Table 8: Values of a_j, U_j, $\hat{C}_{1j}$, $\hat{C}_{2J}$, e_j

a_j	17	40	64	115	215
U_j	39	40	30	31	27
$\hat{C}_{1J}$	3.178	1.454	5.124	3.124	2.784
$\hat{C}_{2J}$	2.365	4.157	6.214	2.154	3.715
e_j	31.361	31.361	31.361	31.361	31.361

We deduced the value of the statistic test Y_n^2:

$$Y_n^2 = X^2 + Q = 5.6258 + 4.1592 = 9.785$$

For significance level $\varepsilon = 0.05$, the critical value $\chi_6^2 = 12.3204$ is superior than the value of $Y_n^2 = 9.785$. So, the results confirm that the studied data can be described by the ND.

As several alternatives can be used to model these data, we use the proposed test statistic to distinguish between theseones.To this end, we calculated the values of the test statistics Y_n^2 for the alternatives. The results are given in Table 9.

Modeling distribution	Y_n^2
New two-parameter distribution ND,	9.785
Exponentiated Weibull (EW)	10.054
Modified Weibull (MW)	11.235
Extended generalized Lindley distribution (EGL)	10.795
Inverse Lindley (IL)	12.055
Power Lindley (PL)	10.415

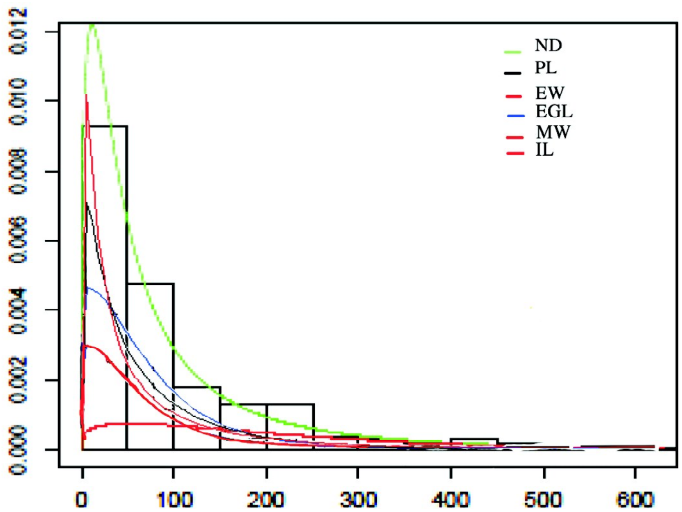

Figure 2

As expected, the new two-parameter distribution ND fit the data set better than the EW, MW, EGL,IL and PL distributions.

As in the previews example, we use the pdfs plots of the data for the different distributions which confirm our result (Fig.2).

Conclusion

Among the new models proposed in the literature, the new two-parameter distribution ND is interesting because of its different forms of the failure rate function. It can be used in several areas and particularly in reliability and medical studies where data are often censored. In this work, after calculating the maximum likelihood estimators of the model parameters, a criteria of goodness-of-fit test statistic Y^2 is provided to validate this distribution when data are right censored. Censored and complete real data sets from reliability and survival analysis are used to illustrate the usefulness of this model and the practicability of the proposed test. In addition and as shown in the power study, we can conclude that the proposed goodness-of fit test Y^2 for the new two-parameter distribution ND has a high power, so it can be used to differentiate ND from the competing distributions for small and large sample sizes.

References

Aidi, K. and Seddik-Ameur, N. (2016), Chi-square tests for generalized exponential AFT distributions with censored data, *Electronic journal of Applied Statistical Analysis,* Vol. 09, Issue 02, 371-384.

Bagdonavièius, V. and Nikulin, M. (2011), Chi-squared goodness-of-fit test for right censored data, *International Journal of Applied Mathematics and Statistics*, vol. 24, 30-50.

Bagdonavièius, V., Levuliene, R., J. and Nikulin, M., (2013), Chi-squared goodness-of-fit tests for parametric accelerated failure time models, Communications in Statistics-Theory and Methods, Volume: 42 Issue: 15, 2768-2785.

Beih SED, Abdelfattah M, Shamsan AG (2017) The Exponential Flexible Weibull Extension Distribution. *Journal of Modelling and Simulation*. 5: 83-97.

Doostmoradi, A., and Chamran, S., (2018), A New Distribution with two parameters to Lifetime Data, Biostatistics and *Biometrics Open Access Journal*. Volume 8 Issue 2.

Doostmoradi A, Zadkarami M, Akhond M (2016) A New Entezar Distribution for Lifetime modeling, *Iran J Sci Technol Trans Sci*. 42(1): 129-139.

Felipe RS, Edwin MMO, Cordeiro M (2011) The generalized inverse Weibull distribution. Stat Papers 52(3): 591-619.

Ghitany ME, Al-Mutairi DK, Balakrishnan N, Al-Enezi LJ (2013) Power Lindley Distribution and Associated Inference. Computational Statistics and Data Analysis 64: 20-33.

Ghitany ME, Al-Mutairi DK, Al-Awadhi FA, Al-Burais MM (2012) Marshall-Olkin Extended Lindley Distribution and Its Application. *International Journal of Applied Mathematics*. 25: 709-721.

Ghitany ME, Alqallaf F, Al-Mutairi DK, Husain HA (2011) A Two parameter Weighted Lindley Distribution and Its Applications to Survival Data. Mathematics and Computers in Simulation 81(6): 1190-1201.

Manal M. Nassar, Fathy H. Eissa (2003), On the Exponentiated Weibull Distribution, Communications in Statistics - Theory and Methods,Volume 32, 2003 - Issue 7

Ravi, V. and Gilbert, P. D. (2009), BB: An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function, *J. Statist. Software*. 32(4).

Silva G, Ortega E, Cordeiro G (2010) The beta modified Weibull distribution. Lifetime Data Anal 16(3): 409-430.

Sharma VK, Singha SK, Singha U, Agiwal V (2015) The inverse Lindley distribution: a stress-strength reliability model with application to head and neck cancer data, *Journal of Industrial and Production Engineering*. 32(3): 162-173.

Stacy EW (1962) A generalization of the gamma distribution. Annals of Mathematical Statistics 33: 1187-1192.

Torabi H, Falahati-Naeini M, Montazeri NH (2014) An Extended Generalized Lindley Distribution and Its Applications to Lifetime Data. *J Statist Res Iran*. 11: 203-222.

Warahena-Liyanage G, Pararai M (2014) A Generalized Power Lindley Distribution with Applications. *Asian Journal of Mathematics and Applications*. 1: 1-23.

Zakerzadeh H, Dolati A (2009) Generalized Lindley Distribution. *Journal of Mathematical Extension*. 3: 13-25.