Ekoist: Journal of Econometrics and Statistics, 33, 1-11

DOI: 10.26650/ekoist.2020.33.810417
http://ekoist.istanbul.edu.tr

İSTANBUL ÜNİVERSİTESİ YAYINEVİ

Received: 14.10.2020
Revision Requested: 10.12.2020
Last Revision Received: 13.12.2020
Accepted: 17.12.2020

Ekoist: Journal of Econometrics and Statistics

RESEARCH ARTICLE / ARAŞTIRMA MAKALESİ

On Weibull-Pareto Distribution in Censored and Uncensored Data Structures

Kamil Alakuş*, Necati Alp Erilli**

Abstract

The Weibull distribution gives a flexible measurement that details the probability distribution associated with the lifetime characteristics of a particular part or service component. It is commonly used to assess product reliability, analyze life data, and model failure times. It is a versatile distribution that can take on the characteristics of other types of distributions, based on the value of the shape parameter. Weibull-Pareto distribution has been introduced as a new type of application of the Weibull distribution. In this article, point and interval estimations for Weibull-Pareto distribution with censored and uncensored data are investigated. With real-time data applications, it is shown that Weibull-Pareto results are better against Exponential, Gamma, and Weibull distributions' results.

Keywords

Weibull-Pareto distribution, Alternative maximum likelihood estimation, T-X family, Weibull distribution

JEL Classification

C34, C02

* Kamil Alakuş (Prof. Dr.), 19 Mayis University, Faculty of Arts and Sciences, Department of Statistics, Samsun, Turkey. E-mail: kamilal@omu.edu.tr ORCID: 0000-0002-5092-8486

** **Corresponding author:** Necati Alp Erilli (Assoc. Prof. Dr.), Sivas Cumhuriyet University, Faculty of Economics and Administrative Sciences, Department of Econometrics, Sivas, Turkey. E-mail: aerilli@cumhuriyet.edu.tr ORCID: 0000-0001-6948-0880

To cite this article: Alakus, K., & Erilli, N. A. (2020). On Weibull-Pareto distribution in censored and uncensored data structures. *EKOIST Journal of Econometrics and Statistics*, 33, 1-11. https://doi.org/10.26650/ekoist.2020.33.810417

Introduction

The Weibull distribution is a continuous probability distribution named after Swedish mathematician Waloddi Weibull. He originally proposed the distribution as a model for material breaking strength but recognized the potential of the distribution in his work (Weibull, 1951). It is a well-known distribution due to its wide use to model various types of data. With many applications of the Weibull distribution, the distribution has been widely used in survival reliability analyses and model failure times. It is a versatile distribution that can take on the characteristics of other types of distributions, based on the value of the shape parameter, β. In recent years, many researchers developed many generalizations of the Weibull distribution. These generalizations include the generalized Weibull distribution by Mudholkar and Kollia (1994), the exponentiated-Weibull distribution by Mudholkar et al. (1995), Inverse Weibull distribution by Akgul et al. (2016), Modified Weibull distribution by He et al. (2016), and the beta-Weibull distribution by Famoye et al. (2005), and so on.

In this article, point and confidence interval estimates of Weibull-Pareto distribution with censored and uncensored data structures are studied. Alzaatreh et al. (2013a, 2013b) studies can be considered pioneering studies in this regard. Nasiru and Luguterah (2015) also investigated this distribution as well. The probability density function (*pdf*) of Weibull-Pareto Distribution (WPD) can be given in Eq.1:

$$f(t)=\begin{cases}\dfrac{\alpha\lambda}{t}\left[\lambda\log\left(t/\theta\right)\right]^{\alpha-1}\exp\left\{-\left[\lambda\log\left(t/\theta\right)\right]^{\alpha}\right\}; & t\geq\theta>0;\alpha,\lambda>0\\ 0 ; & other\end{cases} \tag{1}$$

Cumulative distribution function $[F(t)]$, Survival Function $[S(t)]$, and Hazard Function $[h(t)]$ of Weibull-Pareto distribution can be given as in Eq.2, 3, and 4, respectively.

$$F(t)=1-\exp\left\{-\left[\lambda\log\left(t/\theta\right)\right]^{\alpha}\right\} \tag{2}$$

$$S(t)=\exp\left\{-\left[\lambda\log\left(t/\theta\right)\right]^{\alpha}\right\} \tag{3}$$

$$h(t)=\frac{f(t)}{s(t)}=\frac{\alpha\lambda}{t}\left[\lambda\log\left(t/\theta\right)\right]^{\alpha-1} \tag{4}$$

The definition ranges in the formulas are; $\lambda,\ \theta,\ \alpha>0$ and $t>\theta$.

Statistical Methodology

In this section, it is given point and interval estimation of censored and uncensored cases for Weibull-Pareto distribution.

Point Estimation for Uncensored Cases

To find the estimators of the parameters, an alternative maximum likelihood method (AMLE) was used. First, for θ the estimator $\hat{\theta} = Min\{t_1, t_2, ..., t_n\} = t_{(1)}$ with AMLE is defined. Thus, the number of parameters to be estimated becomes 2: parameters of α and β. When $\hat{\theta} = t_{(1)}$ the log-likelihood function (*Llik*) can be written as in Eq.5 which is taken from independent and random sample n units.

$$Llik = (n-r)\log\alpha + (n-r)\alpha\log\lambda - \sum_{t_i \neq t_{(1)}}^{n} \log(t_i) + (\alpha-1)\sum_{t_i \neq t_{(1)}}^{n} \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right] - \lambda^{\alpha}\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\alpha} ; \lambda, \theta, \alpha > 0,\ t > \theta \tag{5}$$

Here r represents the frequency of the smallest observation value. Thus, with the AMLE method, the estimators of the parameters α and λ can be given as follows:

$$\left.\frac{\partial Llik}{\partial \alpha}\right|_{\alpha=\hat{\alpha}} = 0 \Rightarrow$$

$$\left.\frac{\partial Llik}{\partial \alpha}\right|_{\alpha=\hat{\alpha}} = \frac{(n-r)}{\hat{\alpha}} + \sum_{t_i \neq t_{(1)}}^{n} \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right] - (n-r)\frac{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}} \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]}{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}} = 0$$

or

$$\hat{\alpha} : \frac{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}} \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]}{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}} - \frac{1}{\hat{\alpha}} = \frac{1}{(n-r)}\sum_{t_i \neq t_{(1)}}^{n} \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right] \tag{6}$$

and

$$\left.\frac{\partial Llik}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} = 0 \Rightarrow$$

$$\left.\frac{\partial Llik}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} = \frac{(n-r)\hat{\alpha}}{\hat{\lambda}} - \frac{\hat{\alpha}\hat{\lambda}^{\hat{\alpha}}}{\hat{\lambda}}\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}} = 0 \geq \hat{\lambda}^{\hat{\alpha}} = \frac{(n-r)}{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}}$$

As a result, $\hat{\lambda}$ can be found as;

$$\hat{\lambda}=\left\{\frac{(n-r)}{\sum_{t_i\neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}}\right\}^{1/\hat{\alpha}} \tag{7}$$

Equation (6) is calculated by the iterative process and the λ estimate is found by Eq.7.

Confidence Interval Estimation for Uncensored Cases

Let us first obtain the Fisher information matrix of the log-likelihood function given by (5). For this, the second derivatives according to α and λ should be taken. Thus, elements of the Fisher information matrix can be given in Eq.8:

$$\begin{cases} I_{\alpha\alpha}=-\dfrac{\partial^2 Llik}{\partial\alpha^2}=\dfrac{(n-r)}{\hat{\alpha}^2}+\hat{\lambda}^{\hat{\alpha}}\log^2\hat{\lambda}\hat{a}+2\lambda^{\hat{\alpha}}\log\hat{\lambda}b+\hat{\lambda}^{\hat{\alpha}}c \\ I_{\alpha\lambda}=I_{\lambda\alpha}=-\dfrac{\partial^2 Llik}{\partial\lambda\partial\alpha}=-\dfrac{(n-r)}{\hat{\lambda}}+\hat{\alpha}\hat{\lambda}^{\hat{\alpha}-1}\log\hat{\lambda}\hat{a}+\hat{\lambda}^{\hat{\alpha}-1}\hat{a}+\hat{\alpha}\lambda^{\hat{\alpha}-1}b \\ I_{\lambda\lambda}=-\dfrac{\partial^2 Llik}{\partial\lambda^2}=\dfrac{(n-r)\hat{\alpha}}{\hat{\lambda}^2}+\hat{\alpha}(\hat{\alpha}-1)\lambda^{\hat{\alpha}-2}a \end{cases} \tag{8}$$

The constants a, b, and c gave in Eq.8 are respectively defined as follows.

$$a=\sum_{t_i\neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}} \quad ; \quad b=\sum_{t_i\neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}\log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right] \text{ and}$$

$$c=\sum_{t_i\neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}\log^2\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]$$

Thus, the observed Fisher information matrix is found as, $I(\alpha,\lambda)=\begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\lambda} \\ I_{\lambda\alpha} & I_{\lambda\lambda}\end{bmatrix}$.

The inverse of this matrix is defined as the covariance matrix:

$$\sum(\alpha,\lambda)=I^{-1}(\alpha,\lambda)=\frac{1}{|I(\alpha,\lambda)|}\begin{bmatrix} I_{\lambda\lambda} & -I_{\alpha\lambda} \\ -I_{\lambda\alpha} & I_{\alpha\alpha}\end{bmatrix}=\begin{bmatrix} Var(\alpha) & Cov(\alpha,\lambda) \\ Cov(\lambda,\alpha) & Var(\lambda)\end{bmatrix}$$

So, $100(1-\alpha)\%$ confidence interval estimate of parameters α and λ can be given respectively as; $\hat{\alpha}\pm z_{\alpha/2}\sqrt{Var(\hat{\alpha})}$ and $\hat{\lambda}\pm z_{\alpha/2}\sqrt{Var(\hat{\lambda})}$. Finally, test statistics for the significance of α and λ can be given as; $t_{(\alpha)}=\dfrac{\hat{\alpha}}{\sqrt{Var(\hat{\alpha})}}$ and for β $t_{(\beta)}=\dfrac{\hat{\lambda}}{\sqrt{Var(\hat{\lambda})}}$.

Point Estimation for Censored Cases

As for the AMLE of the parameters, the first estimator of the parameters θ is determined as $\hat{\theta} = Min\{t_1, t_2, ..., t_n\} = t_{(1)}$, as in the case of the uncensored case.

Thus, the number of parameters to be estimated becomes 2: parameter of α and β. As in the uncensored case, if we take $\hat{\theta} = t_{(1)}$, the log-likelihood function can be written as in Eq.9:

$$Llik = (d-k)\log\alpha + (d-k)\alpha\log\lambda - \sum_{t_i \neq t_{(1)}}^{n} \delta_i \log(t_i) + (\alpha-1)\sum_{t_i \neq t_{(1)}}^{n} \delta_i \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right] - \lambda^{\alpha} \sum_{t_i \neq t_{(1)}}^{n} \left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\alpha} \tag{9}$$

In this equation, k represents the uncensored frequency of the smallest observation value. Thus, with the AMLE method, the estimators of the parameters α and λ can be given as follows:

$$\left.\frac{\partial Llik}{\partial \alpha}\right|_{\alpha=\hat{\alpha}} = 0 \Rightarrow$$

$$\left.\frac{\partial Llik}{\partial \alpha}\right|_{\alpha=\hat{\alpha}} = \frac{(d-k)}{\hat{\alpha}} + \sum_{t_i \neq t_{(1)}}^{n} \delta_i \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right] - (d-k)\frac{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}} \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]}{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}} = 0$$

or

$$\hat{\alpha} : \frac{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}} \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]}{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}} - \frac{1}{\hat{\alpha}} = \frac{1}{(d-k)} \sum_{t_i \neq t_{(1)}}^{n} \partial_i \log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right] \tag{10}$$

and

$$\left.\frac{\partial Llik}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} = 0 \Rightarrow$$

$$\left.\frac{\partial Llik}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} = \frac{(d-k)\hat{\alpha}}{\hat{\lambda}} - \frac{\hat{\alpha}\hat{\lambda}^{\hat{\alpha}}}{\hat{\lambda}} \sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}} = 0 \Rightarrow \hat{\lambda}^{\hat{\alpha}} = \frac{(d-k)}{\sum_{t_i \neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}}$$

As a result, $\hat{\lambda}$ can be found as;

$$\hat{\lambda}=\left\{\frac{(d-k)}{\sum_{t_i\neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}}\right\}^{1/\hat{\alpha}} \tag{11}$$

Eq.9 is calculated by the iterative process and λ the estimate is found by Eq.11. For the data used in the study, the best result was obtained with the AMLE method.

Confidence Interval Estimation for Censored Case

First, the Fisher information matrix of the log-likelihood function given by Eq.9 is obtained. For this, the second derivatives according to α and λ should be taken. Thus, Fisher information matrix elements are given with Eq.12.

$$\begin{cases} I_{\alpha\alpha}=-\dfrac{\partial^2 Llik}{\partial\alpha^2}=\dfrac{(d-k)}{\hat{\alpha}^2}+\hat{\lambda}^{\hat{\alpha}}\log^2\hat{\lambda}\hat{a}+2\hat{\lambda}^{\hat{\alpha}}\log\hat{\lambda}b+\hat{\lambda}^{\hat{\alpha}}c \\ I_{\alpha\lambda}=I_{\lambda\alpha}=-\dfrac{\partial^2 Llik}{\partial\lambda\partial\alpha}=-\dfrac{(d-k)}{\hat{\lambda}}+\hat{\alpha}\hat{\lambda}^{\hat{\alpha}-1}\log\hat{\lambda}\hat{a}+\hat{\lambda}^{\hat{\alpha}-1}a+\hat{\alpha}\hat{\lambda}^{\hat{\alpha}-1}b \\ I_{\lambda\lambda}=-\dfrac{\partial^2 Llik}{\partial\lambda^2}=\dfrac{(d-k)\hat{\alpha}}{\hat{\lambda}^2}+\hat{\alpha}(\hat{\alpha}-1)\hat{\lambda}^{\hat{\alpha}-2}a \end{cases} \tag{12}$$

The constants *a*, *b,* and *c* gave in Eq.12 are respectively defined as follows.

$$a=\sum_{t_i\neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}} \quad ; \quad b=\sum_{t_i\neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}\log\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right] \text{ and}$$

$$c=\sum_{t_i\neq t_{(1)}}^{n}\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]^{\hat{\alpha}}\log^2\left[\log\left(\frac{t_i}{t_{(1)}}\right)\right]$$

Thus, the observed Fisher information matrix is found as, $I(\alpha,\lambda)=\begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\lambda} \\ I_{\lambda\alpha} & I_{\lambda\lambda}\end{bmatrix}$. The inverse of this matrix is the covariance matrix can be defined as;

So, $100(1-\alpha)\%$ confidence interval estimate of parameters α and λ can be given respectively as; $\hat{\alpha}\pm z_{\alpha/2}\sqrt{Var(\hat{\alpha})}$ and $\hat{\lambda}\pm z_{\alpha/2}\sqrt{Var(\hat{\lambda})}$. Finally, test statistics for the significance of α and λ can be given as; $t_{(\alpha)}=\dfrac{\hat{\alpha}}{\sqrt{Var(\hat{\alpha})}}$ and for β $t_{(\beta)}=\dfrac{\hat{\lambda}}{\sqrt{Var(\hat{\lambda})}}$.

Application

In this section, 2 uncensored and 1 censored real data are tested for the power of Weibull-Pareto distribution. The results obtained from the Weibull-Pareto distribution were compared with those obtained from the Exponential, Weibull, and Gamma distributions which are frequently used in survival analysis. For comparison, 4 different model selection criteria are used. These are; Akaike (AIC), Bayesian (BIC), Log-Likelihood (LLik), and -2LLik (Akaike, 1970; Schwarz, 1978; Hurvich and Tsai, 1989).

Uncensored Data

2 uncensored real data is tested in this subsection. The first example is Head and Neck Cancer Disease Survival Times is taken from Efron (1988) and the second example is the Guinea Pigs Infected Study is taken from Bjerkedal (1960).

Head and Neck Cancer Study

The data set reported by Efron (1988), represent the survival times of a group of patients suffering from Head and Neck cancer disease and treated using radiotherapy (RT) is given in Table 1:

Table 1

Head and Neck Cancer Disease Survival Times for Radiotherapy Group

6.53	7.00	10.42	14.48	16.10	22.70	34.00	41.55	42.00	45.28
49.40	53.62	63.00	64.00	83.00	84.00	91.00	108.00	112.0	129.0
133.0	133.0	139.0	140.0	140.0	146.0	149.0	154.0	157.0	160.0
160.0	165.0	146.0	149.0	154.0	157.0	160.0	160.0	165.0	173.0
176.0	218.0	225.0	241.0	248.0	273.0	277.0	297.0	405.0	417.0
420.0	440.0	523.0	583.0	594.0	1101.0	1146.0	1417.0		

Parameter estimation results of Exponential, Weibull, Gamma, and Weibull-Pareto distributions and model selection criteria results for Head and Neck Cancer Disease data are given in Table 2.

Table 2

Parameter Estimation and Model Selection Criteria Results for Head and Neck Cancer Disease Data

Distribution	**Parameter Estimation**	**LLik**	**-2LLik**	**AIC**	**BIC**
Exponential	$\hat{\beta}$ =226.1738	-372.4356	744.8712	746.8712	748.9317
Weibull	$\hat{\alpha}$ =0.973234 $\hat{\beta}$ =223.1886	-372.3952	744.7903	748.7903	752.9112
Gamma	$\hat{\alpha}$ =1.033092 $\hat{\beta}$ =218.9289	-372.3862	744.7724	748.7724	752.8933
Weibull-Pareto	$\hat{\alpha}$ =2.951379 $\hat{\theta}$ =6.53 $\hat{\beta}$ =0.297464	**-369.7180**	**739.4361**	**743.4361**	**747.5222**

When we look at the results of the model selection criteria in Table.2, it is seen that the best model selection values (for all of 4) belong to the Weibull-Pareto distribution.

Estimated Fisher information matrix for Weibull-Pareto distribution results calculated as $\hat{I}(.) = \begin{bmatrix} 804.7569 & 64.40761 \\ 64.40761 & 5611.219 \end{bmatrix}$ an estimated covariance matrix calculated as; $\hat{\Sigma}(.) = \frac{1}{10^4} \begin{bmatrix} 12.44 & -0.14 \\ -0.14 & 1.78 \end{bmatrix}$. Thus, 95% confidence intervals can be found as (2.882256; 3.020502) for α and (0.271286; 0.323641) for λ. Table 3 shows summary results and significance values for parameter estimations of Weibull-Pareto distribution results. As for Table.3, it can be seen that both parameters are significant at 0.01 and 0.05 levels of importance.

Table 3

Parameter Results for Weibull-Pareto Distribution

Parameter	Predicted	Variance	Std. Error	t-test	p-Value
α	2.951379	0.001244	0.035267	83.687	0.000
λ	0.297464	0.000178	0.013356	22.272	0.000

Guinea Pigs Infected Study

This data set was observed and reported by Bjerkedal (1960). It represents the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli. Observations of data are given in Table 4:

Table 4

Guinea Pigs Infected Data

12	15	22	24	24	32	32	33	34	38	38	43
44	48	52	53	54	54	55	56	57	58	58	59
60	60	60	60	61	62	63	65	65	67	68	70
70	72	73	75	76	76	81	83	84	85	87	91
95	96	98	99	109	110	121	127	129	131	143	146
146	175	175	211	233	258	258	263	297	341	341	376

Parameter estimation results of Exponential, Weibull, Gamma, and Weibull-Pareto distributions and model selection criteria results for Guinea Pigs Infected Data are given in Table 5.

Table 5

Parameter Estimation and Model Selection Criteria Results for Guinea Pigs Infected Data

Distribution	**Parameter Estimation**	**LLik**	**-2LLik**	**AIC**	**BIC**
Exponential	$\hat{\beta}$ =99.81944	-403.4421	806.8843	808.8843	811.1609
Weibull	$\hat{\alpha}$ =1.393187 $\hat{\beta}$ =110.5552	-397.1477	794.2953	798.2953	802.8487
Gamma	$\hat{\alpha}$ =2.081279 $\hat{\beta}$ =47.96063	-394.2476	788.4952	792.4952	797.0485
Weibull-Pareto	$\hat{\theta}$ =12 $\hat{\alpha}$ =2.968006 $\hat{\lambda}$=0.474259	**-383.7787**	**767.5574**	**771.5574**	**776.0827**

Estimated Fisher information matrix for Weibull-Pareto distribution results calculated as $\hat{I}(.)=\begin{bmatrix}438.1815 & 377.727\\ 377.727 & 2780.716\end{bmatrix}$ an estimated covariance matrix calculated as; $\hat{\Sigma}(.)=\frac{1}{10^4}\begin{bmatrix}25.85 & -3.5\\ -3.5 & 4.07\end{bmatrix}$. Thus, 95% confidence intervals can be found as (2.868357; 3.067655) for α and (0.434702; 0.513816) for λ. Table 6 shows summary results and significance values for parameter estimations of Weibull-Pareto distribution results. As for Table.6, it can be seen that both parameters are significant at 0,01 and 0.05 levels of importance.

Table 6

Parameter Results for Weibull-Pareto Distribution

Parameter	**Predicted**	**Variance**	**Std. Error**	**t-test**	**p-Value**
α	2.968006	0.002585	0.050841	58.378	0.000
λ	0.474259	0.000407	0.020182	23.499	0.000

Censored Data

The Stanford Heart Transplant Study

In this subsection, the distribution of Weibull-Pareto is tested on censored data. The Stanford Heart Transplant Study data (Kalbfleisch and Prentice,1980) has 103 observations 28 of which are censored. The data is given in Table.7:

Table 7

The Stanford Heart Transplant Data

1	2	2	2	3	3	3	5	5	6	6	8
9	12	16	16	16	17	18	21	21	28	30	32
35	36	37	39	40	40	43	45	50	51	53	58
61	66	68	68	69	72	72	77	78	80	81	85
90	96	100	102	110	149	153	165	186	188	207	219
263	285	285	308	334	340	342	583	675	733	852	979
995	1032	1386									
11+	31+	39+	109+	131+	180+	265+	340+	370+	397+	427+	445+
482+	515+	545+	596+	630+	670+	841+	915+	941+	1141+	1321+	1400+
1407+	1571+	1586+	1799+								
+: Denotes censored time.											

Parameter estimation results for Exponential, Weibull, and Weibull-Pareto distributions and model selection criteria results for Stanford Heart Transplant Data are given in Table.8. As for Table.8, Weibull-Pareto results for Stanford Heart Transplant data have minimum values for all selection criteria.

Table 8

Parameter Estimation and Model Selection Criteria Results for The Stanford Heart Transplant Data

Distribution	Parameter Estimation	LLik	-2Llik	AIC	BIC
Exponential	$\hat{\beta}$ =425.9733	-529.0783	1058.1565	1060.1565	1062.7912
Weibull	$\hat{\alpha}$ =0.526421 $\hat{\beta}$ =362.2146	-497.6149	995.2299	999.2299	1004.4993
Weibull-Pareto	$\hat{\theta}$ =1.00 $\hat{\alpha}$ =2.250424 $\hat{\beta}$ =0.173726	**-486.4126**	**972.8252**	**976.8252**	**982.0947**

Estimated Fisher information matrix for Weibull-Pareto distribution results calculated as $\hat{I}(.)=\begin{bmatrix} 19.48399 & -17.62861 \\ -17.62861 & 12342.78 \end{bmatrix}$ an estimated covariance matrix calculated as; $\hat{\Sigma}(.)=\frac{1}{10^5}\begin{bmatrix} 5139.059 & 7.3399 \\ 7.3399 & 8.1124 \end{bmatrix}$. Thus, 95% confidence intervals can be found as (1.806102; 2.694746) for α and (0.156073; 0.191380) for λ. Table.9 shows summary results and significance values for parameter estimations of Weibull-Pareto distribution results.

Table 9

Parameter Results for Weibull-Pareto Distribution

Parameter	Predicted	Variance	Std. Error	t-test	p-Value
α	2.250424	0.051390591	0.226695	4.411	0.000
λ	0.173726	0.000081124	0.009007	98.339	0.000

As we look at Table.9 its clearly stated that, both parameters are significant at 0.01 and 0.05 level of importance.

Conclusion

Survival analysis is generally defined as a set of methods for analyzing data where the outcome variable is the time until the occurrence of an event of interest. There are many statistical distributions used in survival analysis that gives successful results: Exponential, Weibull, Gamma, Lindley, etc. Weibull-Pareto distribution is one of the generalizations of Weibull distribution. In this study, point and confidence interval estimations for censored and uncensored data of the Weibull-Pareto distribution were introduced. The effect of the distribution was compared in real-life data of censored and uncensored observations with results obtained from some other exponential family distributions (Exponential, Weibull, Gamma). The results, compared with other known distributions, revealed that the Weibull-Pareto Distribution provides a better fit for modeling real-life data.

Hakem Değerlendirmesi: Dış bağımsız.
Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.
Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Peer-review: Externally peer-reviewed.
Conflict of Interest: The authors have no conflict of interest to declare.
Grant Support: The authors declared that this study has received no financial support.

References/Kaynakça

Akaike, H. (1970). Statistical Predictor Identification, *Annals of the Institute of Statistical Mathematics*, 22, 203-217.

Akgul F.G., Şenoğlu B., & Arslan T. (2016). An alternative distribution to Weibull for modeling the wind speed data: Inverse Weibull distribution. *Energy Conversion and Management*, 114, p:234-240.

Alzaatreh, A., Famoye, F., & Lee, C. (2013a). Weibull-Pareto Distribution and Its Applications. *Communications in Statistics - Theory and Methods*, 42(9), 1673-1691.

Alzaatreh, A., Lee, C., & Famoye, F. (2013b). A new method for generating families of continuous distributions. *Metron: International Journal of Statistics*, 71, 63-79.

Bjerkedal, T. (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Am J Hyg* 72, 130-148.

Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American Statistical Association*, *83*(402), 414-425.

Famoye, F., Lee, C., & Olumolade, O. (2005). The Beta-Weibull distribution. *J. Statist. Theor. Appl, 4*(2), 121–136.

He B., Cui W., & Du X. (2016). An additive modified Weibull distribution. *Reliability Engineering and System Safety*, 145, 28-37.

Hurvich, C.M., & Tsai, C. (1989). Regression and Time Series Model Selection in Small Samples, *Biometrika*, 76, 297-307.

Kalbfleisch, J.D., & Prentice, R.L. (1980). *The Statistical Analysis of Failure Time Data*. NY: Wiley & Sons. First Edition, Appendix I, pp. 230-32.

Mudholkar, G. S., & Kollia, G. D. (1994). Generalized Weibull family: A structural analysis. *Commun. Statist. Theor. Meth. 23*(4), 1149–1171.

Mudholkar, G.S., Srivastava, D.K., & Freimer, M. (1995). The exponentiated Weibull family: a reanalysis of the bus-motor-failure data. *Technometrics*, *37*(4), 436–445.

Nasiru, S., & Luguterah, A. (2015). The New Weibull-Pareto Distribution. *Pak. J. Stat. Oper.Res*., *11*(1), 103-114.

Schwarz, G. (1978). Estimating the Dimensions of a Model, *The Annals of Statistics*, 6, 461-464.

Weibull, W. (1951). Statistical distribution function of wide applicability. *Journal of Applied Mechanics*, 18, 292-297.