PARAMETER ESTIMATION USING THE MEASURE OF SYMMETRIC CROSS ENTROPY DIVERGENCE MEASURE

ALADDIN SHAMILOV*
YELIZ MERT KANTAR*
ILHAN USTA*
*Anadolu University
Science and Art Faculty-Department of Statistics
16470-Eskisehir
TURKEY

Abstact: - In this paper, a method for parameter estimation is presented by using the measure of symmetric cross entropy divergence (J divergence). The presented method is used to estimate two-parameter Weibull distribution which is widely used in reliability, life testing, and survival analysis, engineering, wind energy studies. A comparison of the proposed method and usual methods, such as Maximum likelihood methods, moment method, least square method, is also given. In application, it is used monthly wind speed data measured in Karabayir, Turkey.

Key words:-Parameter estimation methods, *J* divergence measure, Weibull distribution, Wind speed

1 Introduction

Many information-theoretical divergence measures between two probability distributions have been introduced and widely studied by many author in literature [1-7]. The applications of these information-divergences can be found in the contingency table, statistical inference, pattern recognition, queuing theory, goodness-fit-test and parameter estimation. [1], [11] can be given for excellent reference about these measures.

Parameter estimation plays an important role in many natural sciences, engineering and other disciplines. A lot of studies were done to find more precise, accurate methods for parameter estimation. The most widespread methods may be maximum likelihood method, least square method, moment method in literature. But there are a lot of modifications of these methods, Bayes estimation method [19], the minimum-distance estimation method [8], maximum spacing method [9-10]

In this paper, the *J* divergence measure to find parameters of the probability distributions is used and illustrated by estimating parameters of Weibull distribution which is widely used in wind power studies [11-14]. Parameters are found by minimizing *J* measure between the expectation of the order statistics of the sample and the probability mass function (p.m.f.) obtained the functional form of candidate probability density function (p.d.f). Since this approach needs to be minimization, it is said to be optimization method.

The rest of the paper is organized as follows. Section 2 introduces the measure of symmetric cross entropy divergence (J divergence) and optimization method for parameter estimation. A brief of review of entropy, Kullback-Leibler divergence measure and the other information-theoretic divergence measures are also included in this section. In section 3, we find the parameter of the Weibull distribution by using a measure of symmetric cross entropy divergence and recall the other usual methods. It is compared all methods according to different statistical criteria in this section.. The section 4 concludes the paper with some suggestions of future research.

2 Parameter Estimation Using *J* Divergence

Let $\boldsymbol{p} = (p_1, p_2, ..., p_n)$ and $\boldsymbol{q} = (q_1, q_2, ..., q_n)$ be two probability mass functions. Then, Kullback-Leibler measure is

$$\boldsymbol{I}(\boldsymbol{p};\boldsymbol{q}) = \sum_{j=1}^{n} p_j \ln \frac{p_j}{q_j},$$

where logarithmic base is 2. It is well known that *I* is nonnegative, additive, but not symmetric [1]. To obtain a symmetric measure, *J* divergence was defined in the following form

$$\boldsymbol{J}(\boldsymbol{p};\boldsymbol{q}) = \boldsymbol{I}(\boldsymbol{p};\boldsymbol{q}) + \boldsymbol{I}(\boldsymbol{q};\boldsymbol{p})$$

It is clear that *I* and *J* share a lot of properties but *J* has symmetric measure property.

Now, *J* measure is used to estimate parameters of the Weibull distribution. The parameters of the Weibull distribution can be obtained by minimizing *J* measure between the expectation of the order statistics and probability mass function obtained the functional form of candidate p.d.f (see Figure 1). The procedures can be given as follows:

Suppose that $x_1,\ldots,x_n$ is random sample of size *n* from a distribution F (.) and $-\infty = x_{(0)} < x_{(1)} < \ldots < x_{(n)} < x_{(n+1)} = \infty$ denotes the order statistics of the observed sample. It is well known that

$$G(x_{(j)}) = \frac{j}{n+1},\ G(x_{(0)}) = 0,\ G(x_{(n+1)}) = 1. \quad (1)$$

From (1), probabilities of ordered sample are $p'_j = G(x_{(j)}) - G(x_{(j-1)}),\ \ j = 0,\ldots,n+1.$

Let $F(x,\theta)$ be distribution function and $f(x,\theta)$ be p.d.f. with unknown parameter θ and probability of every subinterval $(x_{(j)}, x_{(j-1)})$ be p_j. It is seen Figure 1. In this treatment, the p.d.f. for continuous random variable is converted to p.m.f. for discrete random variable

That is,

$$p_j = \int_{x_{(j-1)}}^{x_{(j)}} f(x,\theta)dx = F(x_{(j)},\theta) - F(x_{(j-1)},\theta) \quad (2)$$

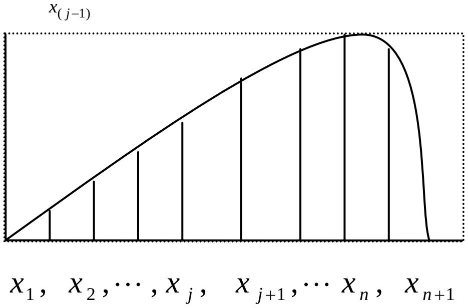


Figure 1 Unknown $f(x,\theta)$

Since $P = (p_0,.p_1,\ldots,p_{n+1})$ and $P' = (p'_0,.p'_1,\ldots,p'_{n+1})$ are two discrete distribution function, *J* measure between these distributions is given following from.

$$\boldsymbol{J}(\boldsymbol{P};\boldsymbol{P}') = \sum_{j=1}^{n+1} p'_j \ln \frac{p'_j}{p_j} + \sum_{j=1}^{n+1} p_j \ln \frac{p_j}{p'_j} \quad (3)$$

$$= \sum_{j=1}^{n+1} p'_j \ln p'_j - \sum_{j=1}^{n+1} p'_j \ln p_j + \sum_{j=1}^{n+1} p_j \ln p_j - \sum_{j=1}^{n+1} p_j \ln p'_j$$

$$= \sum_{j=1}^{n+1} \frac{1}{n+1} \ln \frac{1}{n+1} - \sum_{j=1}^{n+1} \frac{1}{n+1}\left(\ln \int_{x_{(j-1)}}^{x_{(j)}} f(x,\theta)dx \right) + \sum_{j=1}^{n+1}\left(\int_{x_{(j-1)}}^{x_{(j)}} f(x,\theta)dx \right)\left(\ln \int_{x_{(j-1)}}^{x_{(j)}} f(x,\theta)dx \right) + \sum_{j=1}^{n+1}\left(\int_{x_{(j-1)}}^{x_{(j)}} f(x,\theta)dx \right)\left(\frac{1}{n+1} \right) \quad (4)$$

In last expression, he first term of (4) does not depend on θ, the last term of (4) is equal to 1. For this reason, unknown parameters can be obtained by minimizing following equation with respect to θ

$$-\sum_{j=1}^{n+1} \frac{1}{n+1}\left(\ln \int_{x_{(j-1)}}^{x_{(j)}} f(x,\theta)dx \right) + \sum_{j=1}^{n+1}\left(\int_{x_{(j-1)}}^{x_{(j)}} f(x,\theta)dx \right)\left(\ln \int_{x_{(j-1)}}^{x_{(j)}} f(x,\theta)dx \right) \quad (5)$$

From (5), it can be concluded that

$$= -\sum_{j=1}^{n+1} \frac{1}{n+1} \ln((F_{(j)},\theta) - F(x_{(j-1)},\theta)) + \sum_{j=1}^{n+1} ((F_{(j)},\theta) - F(x_{(j-1)},\theta)) \quad (6)$$

$$\ln((F_{(j)},\theta) - F(x_{(j-1)},\theta))$$

$$= -\sum_{j=1}^{n+1} \frac{1}{n+1} \ln(p_j) + \sum_{j=1}^{n+1} p_j \ln(p_j) \quad (7)$$

It is noted that if (6) is multiplied -1, the firs term of (6) is equal to maximum spacing method in [9] and the last term of (7) is equal to the entropy of *P* distribution [1].

Estimation of parameters of the Weibull distribution can be obtained by minimizing (6) with respect to θ. This approach can be considered as minimization of distance between the expectation of the order statistics and the functional form of candidate p.d.f.

3 Parameter estimation for the Weibull distribution

Optimization method used *J* divergence and the usual methods such as the maximum likelihood method, least square method, the moment's method are used to find the parameters of Weibull distribution of wind speed data.
The wind speed data used in this study were taken from [18]. It was measured and recorded hourly at Karabayir, in Kutahya in Turkey at 10 *m* above ground level, between February and June in 2002.

Estimated parameters are listed in Table 2. Figure 2 shows a comparison of the observed cumulative distribution and cumulative curves obtained by mentioned methods.

The estimated parameters by the mentioned methods are listed in Table 2. From Table 2, the proposed method can be seen as competitor of the known methods. For January, optimization method performs well than moment and maximum likelihood methods in terms of RMSE and also performs well than least square method in terms of AIC. For March and April, optimization method performs well than moment and maximum likelihood methods in terms of both criteria. For this reason, it can be said that an optimization method can used to estimate parameter of the Weibull distribution

Root mean square error (*RMSE*) [12], Akaike's information criterion (*AIC*) [17], will be used in statistically evaluating the performance of mentioned methods in estimating parameters of the Weibull distribution. Therefore, the best parameter estimation can be selected according to the lowest values RMSE and AIC.

$$RMSE = \left[\frac{1}{N} \sum_{i=1}^{N} (y_i - x_i)^2 \right]^2 \quad (8)$$

$$AIC = -2\log(L(f(v_i, \theta)) + 2K \,, \quad (9)$$

where y_i is the *i*th actual data, x_i is the *i*th predicted data, *N* is number of all observed wind speed data, *L* is log likelihood function, *K* is the number of parameters of distribution.

Table 1 Estimated parameter by different methods

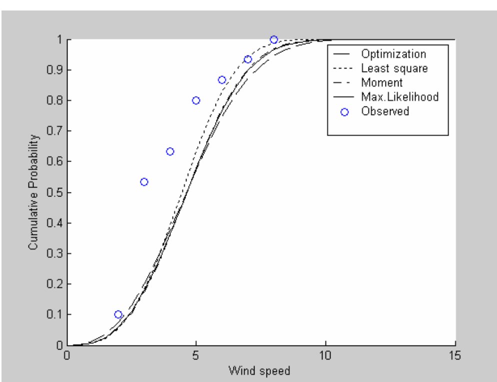
Methods	$\hat{c}$	$\hat{k}$
February		
Optimization	4.4585	2.9314
Least-square	4.3027	3.5879
Moment	4.4192	3.4134
Maximum likelihood	4.4197	3.2626
March		
Optimization	5.3170	2.6245
Least-square	4.9958	3.1294
Moment	5.2553	2.9575
Maximum likelihood	5.2623	2.8932
April		
Optimization	4.6943	3.3610
Least-square	4.5243	3.8134
Moment	4.6729	3.9275
Maximum likelihood	4.6791	3.7080
May		
Optimization	4.6826	3.2658
Least-square	4.4702	3.8525
Moment	3.8602	4.6339
Maximum likelihood	4.6394	3.5578

The estimated parameters by the mentioned methods are listed in Table 2. From Table 2, the proposed method can be seen as competitor of the known methods. For January, optimization method performs well than moment and maximum likelihood methods in terms of RMSE and also performs well than least square method in terms of AIC. For March and April, optimization method performs well than moment and maximum likelihood methods in terms of both criteria. For this reason, it can be said that an optimization method can used to estimate parameter of the Weibull distribution

Table 2 A comparison of different method using various criteria for distribution of wind speed data

February		
Methods	RMSE	AIC
Optimization	0.0005	106.56
Least-square	0.0003	107.02
Moment	0.0048	232.16
Maximum likelihood	0.0004	105.35
March		
Methods	RMSE	AIC
Optimization	0.0010	124.47
Least-square	0.0010	125.58
Moment	0.0012	123.84
Maximum likelihood	0.0011	123.8
April		
Methods	RMSE	AIC
Optimization	0.0011	101.09
Least-square	0.0010	101.12
Moment	0.0015	100.67
Maximum likelihood	0.0013	100.44
June		
Methods	c	k
Optimization	0.0007	105.02
Least-square	0.0006	106.03
Moment	0.0011	159.5
Maximum likelihood	0.0007	105.04

Figure 2 A comparisons of the wind speed data distribution and theoretical cumulative distributions obtained by different methods



The probability distribution of wind speed is very important piece of information needed in assessment of wind energy potential. There are several density functions that can be used to describe the wind speed data. Patel claims [20] that Weibull distribution is the best one to describe the variation in wind speed

4 Conclusions

In this paper, a method for parameter estimation was presented using a measure of symmetric cross entropy divergence (J divergence). The method was used to estimate two-parameter Weibull distribution which was widely accepted in wind power studies. A comparison of the proposed method and usual methods, such as Maximum likelihood methods, moment method, least square method, was also given. It was shown that optimization method performs well than the other methods in terms of some criteria in some cases.

References:

[1]. J.N Kapur, H.K. Kesevan, *Entropy Optimization Principle with Applications*, Academic Press, 1992

[2] S. Kullback, *Information Theory and Statistics*, Wiley, 1959

[3]. U. Kumar, V. Kumar, J.N Kapur Some normalized measures of directed divergence, *Int.J.Gen.Syst*, Vol.13, 1986, pp.5-16.

[4]. A Rengi, On measures of entropy and information, *Proc. 4. Syms.Math. Stat.Prob*., Berkeley, Vol1, 1961, pp.547-561,.

[5]. H. Havrda, F. Chavrat. Quantifications methods of classification Procesess: Concepts of Structural α entropy. *Kybernatica*, Vol.3, 1967, pp.30-35.

[6] S.M. Ali, S.D. Silvey. A general class of coefficient of divergence of one distribution from another, *J.Roy.Statist. Soc., Ser, B*, Vol. 28, 1966, pp.131-142.

[7] J.N. Kapur. Generalized entropy of order α and type β, *The mathematical seminar*, Vol.4, 1967, 78-94.

[8] J.R. Hobbs, A.H. Moore, Miller R.M., A Bayes estimation of the parameters and reliability function of the three parameter Weibull distribution. *IEEE Trans.Reliability*, Vol.37, 1988, pp.364-369.

[9] R.C.H Cheng, N.A.K Amin. Estimating parameters in continuous univariate distributions with a shifted origin, *J. Roy. Statist. Soc.Ser.B*, 1983, Vol.45, pp.394-403.

[10] B.Ranneby. The maximum spacing method: an estimated method related to the maximum likelihood method. *Scand.J.Statists* Vol.11, 1984, pp.93-112.

[11] T. M. Cover, Elements of information theory. (John Wiley, 1991)

[12] E.K Akinar, S. Akinar, An assessment on seasonal anaysis of wind energy characteristics and wind turbine characteristics, *Energy conversion management*, Vol.46, 2005, pp.1848-1867.

[13]. D. Weisser A wind energy analysis of Grenada: an estimation using the 'Weibull' density function, *Renewable Energy*, Vol.28, 2003, pp.1803-1812.

[14]. W. Al-Nassar, S. Alhajraf, A. Al-Enizi, L. Al-Awadhi, Potential wind power generation in the State of Kuwait, Vol.30, 2005, pp.2149-2161.

[15]. K. Ulgen, A.Hepbasli Determination of Weibull parameters for wind energy analysis of Izmir, Turkey. *Int J Energy Res*, 2002, Vol.26, pp.495-506

[16]. K. Ulgen, A. Genc, Hepbasli A, Oturanc G. (2004) . An assessment of wind characteristics for energy generation, *Energy Sources*, Vol.26, pp.1227-1237.

[17]. X. Wu. Calculation of Maximum entropy densities with application to income distribution, *Journal of Econometrics* Vol.115, 2003, pp.347-354.

[18]. M.Arif ÖZGÜRT(thesis), The Investigation of the electric energy production potential with wind results measured in a place in Kütahya, 2002

[19]. J. Berger *Statistical Decision Theory*, Spinger Verlag, 1980.

[20] M.R. Patel *Wind and Solar power systems*, CRC Press, 1999