On a Functional defined by means of Kullback-Leibler Measure and Its Statistical Applications

ALADDIN SHAMILOV, YELIZ MERT KANTAR, ILHAN USTA
Department of Statistics
Anadolu University
Eskisehir, 26470
TURKEY

Abstract: In the present study, we define a functional by means of Kullback-Leibler (K-L) measure for continuous random variables and consider its properties. By virtue of this functional, we suggest a new generalization of K-L optimization principle (GKOP). Due to the defined GKOP, it is possible to obtain the probability distribution of the system which is closer to the true distribution in the sense of K-L measure. Finally, the distributions of yearly wind speed data measured in 2001, 2000, 1999 are found by using GKOP and compared with the Weibull distribution which is widely used to determine distribution of wind speed. The Weibull distribution, whose parameters are estimated using maximum likelihood method, shows worse fits considered all years than the distribution derived from GKOP.

Although principles based on information theory have been used in a variety of field, the obtained distributions though GKOP are the first time to be applied to the wind energy field.

Keywords: Kullback-Leibler measure, Entropy, Wind speed data, Weibull distribution, model-selection criteria.

1 Introduction

Applications of entropy optimization principle can be found in the image process, statistical inference, pattern recognition, queuing theory, reliability analyses, the contingency table, goodness-fit-test and parameter estimation, reliability, life testing, survival analysis, engineering, wind energy studies [1-11].

The one of the important principle is Kullback-Leibler minimum cross entropy principle [12-14]. According to Kullback's minimum cross entropy principle, out of all probability distributions satisfying given moment constraints, it is choosen a probability distribution *P* that minimizes Kullback-Leibler measure. This distribution is called as MinxEnt distribution [14]. In this process, auxiliary functional constructed by Lagrange multipliers method reaches its minimum value at some distribution defined by K-L principle, which corresponds to some moment vector function. So we can define a functional that assigns to each moment vector function precisely one number- conditional minimum value of Kullback-Leibler measure. In order to discriminate mentioned functional from the Kullback-Leibler measure, we call it as Kullback-Leibler functional of moment vector functions.

Note that Kullback-Leibler measure for the continuous random variable is functional of probability density functions. At the same time, K-L functional is defined on the set of given moment vector functions and reaches its minimum value from the set of conditional minimum values of Kullback-Leibler measure. We prove that K-L functional is continuous on the set of continuous moment vector-functions.

Moment vector function which gives the least and the greatest value to K-L functional specify probability distribution. We have proved that the specified probability distribution gives to K-L functional the greatest value in the compact set of continuous moment-vector functions. However, this distribution contains the greater information than the other probability distributions. This result can be considered as a Generalized Entropy Optimization principle.

In this case, when a finite number of moment vector-functions are given, in other words, moment conditions are considered, then by using mentioned GKOP, it is possible easily to obtain the probability distribution representing the greatest information.

It should be noted that if the finite number –k moment function is given, then by virtue of possible–m (m<k) combinations of these functions we can obtain m$\times$ 1 moment vector functions which corresponds to probability distribution containing greatest information.

Moreover by testing several moment vector function GKOP allows to obtain probability distribution representing maximum information and therefore to draw as well as possible inferences about population from the given random sample.

The rest of the paper is organized as follows. Section 2 introduces Kullback-Leibler measure, Kullback-Leibler functional defined on moment functions. Properties of this functional also are given as theorems. In Section 3, the estimation of distribution of wind speed data taken from [15] by using GKOP is obtained. The Section 4 concludes the paper by summarizing the mains results and suggesting future studies.

2. Kullback-Leibler functional on moment vector functions

Let us consider the problem of minimizing Kullback-Leibler measure

$$D(f(x);q(x)) = \int_a^b f(x)\ln\frac{f(x)}{q(x)}dx, \qquad (1)$$

subject to constraint

$$\int_a^b g(x)f(x)dx = \mu\,, \qquad (2)$$

where $\boldsymbol{g(x)} = (g_1(x),...,g_m(x))'$ are moment vector functions, $\boldsymbol{\mu} = (\mu_1,...,\mu_m)$ are moment vector-values.

Auxiliary functional corresponding to problem (1)-(2) is

$$\int_a^b \underbrace{\left(f(x)\ln\frac{f(x)}{q(x)}dx - \lambda_0 f(x) - \sum_{j=1}^m \lambda_j g_j(x) f(x)\right)}_{F} dx = \mu \qquad (3)$$

According to following Euler-Lagrange equation

$$\frac{\partial F}{\partial f(x)} - \frac{d}{dx}\left(\frac{\partial F}{\partial f'(x)}\right) = 0 \qquad (4)$$

the $f(x)$ which minimize (3).

$$\ln\frac{f(x)}{q(x)} + 1 - \lambda_0 - \sum_{j=1}^m \lambda_j g_j(x) = 0\,, \qquad (5)$$

Use of the Euler-Lagrange equation of the calculus of variations gives:

$$f(x) = q(x)\exp(\lambda_0 + \sum_{j=1}^m \lambda_j g_j(x))\,. \qquad (6)$$

(6) is called as MinxEnt in [1].

If (6) is taken into account in (1), then

$$\begin{aligned} D_{\min} &= \int_a^b f(x)\ln\frac{f(x)}{q(x)}dx \\ &= \int_a^b f(x)\left(\lambda_0 + \sum_{j=1}^m \lambda_j g_j(x)\right)dx \qquad (7) \\ &= \lambda_0 + \sum_{j=1}^m \lambda_j \mu_j\,. \end{aligned}$$

It is known that (6) gives to K-L measure minimum value. The following functional

$$U(g) = D_{\min}$$

is called as the Kullback-leibler functional on the set of moment vector function $g(x)$

Let us consider functional

$$H(f) = -\int_a^b f(x)\ln f(x)dx$$

which represents entropy of random variable with $f(x)$ probability density function (p.d.f.), then this functional subject to constraints

$$\int_a^b f(x)dx = 1 \tag{8}$$

has maximum value

$$H_{\max} = \ln(b-a)\,. \tag{9}$$

Entropy of variable with MinxEnt p.d.f. is equal.

$$H(f) = -\int_a^b f(x)\ln f(x)dx$$

$$= -\int_a^b f(x)\left(\ln q(x) + \lambda_0 + \sum_{j=1}^{m} \lambda_j g_j(x)\right)dx \tag{10}$$

$$= -\int_a^b f(x)\ln q(x)dx - U(g)$$

$$\mathrm{H} = -\int_a^b f(x)\ln q(x)dx - U(g) \tag{11}$$

It is known that the change of entropy is information.

The change of entropy (9) via (11) is given as follows:

$$I = \ln(b-a) + \int_a^b f(x)\ln q(x)dx + U(g) \tag{12}$$

where $U(g)$ by defined (7) is Kullback-Leibler functional defined on the set of moment vector function $g(x)$. The functional $U(g)$ defined by (7) has the properties which are given as theorems.

<u>**Theorem 1**</u> $U(g)$ is continuous on the set of continuous vector moment functions $C[a,b]$.

<u>**Theorem 2.**</u> $U(g)$ reaches its least and greatest values in the given compact set $K \subset C[a,b]$ of continuous moment vector functions.

$$\max_{g \in K} U(g) = U\left(g^{(0)}\right)$$

The distribution corresponds to $g^{(0)}$ is called as MaxMinxEnt.

3. The Use of Wind Probability distributions derived from GKOP

This section develops a theoretical approach to the analytically determination of the wind speed distributions through the application of GKOP. On this purpose, a comparison of the two parameters- Weibull which is widely used to determine distribution of wind speed [16-20] and the distributions derived from GKOP (MaxMinxEnt) is also made to show the ability to describe the experimental mean wind power density.

The comparison between MaxMinxEnt distributions and the Weibull distributions for the wind speed data is shown in Table 1, Table 2, Table 3. Here the prior distribution is taken the distribution of previous year.

Chi-square(χ^2), root mean square error(RMSE), correlation coefficient(R^2), Kullback-Leibler measure (K-L) will be used in statistically evaluating the performance of the Weibull and the distributions obtained GKOP.

The formula of mentioned ssuitability judgment criteria are

$$\chi^2 = \frac{\sum_{i=1}^{N}(y_i - x_i)^2}{N-n} \tag{12}$$

$$RMSE = \left(\frac{\sum_{i=1}^{N}(y_i - x_i)^2}{N}\right)^2, \tag{13}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - x_i)^2}{\sum_{i=1}^{N}(y_i - z_i)^2}, \tag{14}$$

$$K-L = \sum_{i=1}^{N} y_i \log\frac{y_i}{x_i} \tag{15}$$

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, *n* is the number of parameters or the number of constrains.

The best distribution function can be determined according to the lowest values *RMSE*, χ^2, *K-L* measure and the highest value R^2 .

The wind speed data taken from [15] in time-series format was usually arranged in the frequency distribution format. It is convenient a statistical analysis. We fit the Weibull and distribution and the distributions obtained GKOP to wind speed data. The frequency distributions calculated the Weibull and the distributions obtained GKOP are given Table 1

Table 1. The estimated distribution of wind data for 1999, when prior distribution is taken as 1998

V (m/s)	1999 (h)	f_i	f_w	f_{k1}	f_{k2}
0-1	2681	0.3060	0.3272	0.2928	0.3115
1-2	3410	0.3893	0.3167	0.3863	0.3812
2-3	1558	0.1779	0.1921	0.1773	0.1675
3-4	565	0.0645	0.0961	0.0821	0.0776
4-5	247	0.0282	0.0422	0.0354	0.0343
5-6	138	0.0158	0.0167	0.0134	0.0136
6-7	63	0.0072	0.0061	0.0058	0.0061
7-8	32	0.0037	0.0020	0.0038	0.0042
8-9	23	0.0026	0.0006	0.0005	0.0006
9-10	14	0.0016	0.0002	0.0004	0.0004
10-11	14	0.0016	0.0001	0	0
11-12	9	0.0010	0.0000	0.0002	0.0003
12-13	2	0.0002	0.0000	0	0
13-14	1	0.0001	0.0000	0.0003	0.0004
14-15	2	0.0002	0.0000	0.0007	0.0010
15-16	1	0.0001	0.0000	0.0004	0.0006
16-17	0	0	0.0000	0.0005	0.0007

In the Table 1, Table 2, Table 3, f_i is probability function (p.d.f.) of observed data, $f_w(x)$ is the Weibull density function, f_{k1} (MaxMinxEnt$_1$) is the distributions obtained GKOP subject to one constraint, f_{k2}(MaxMinxEnt$_2$) is the distributions obtained GKOP subject to two constraints.

Table 2. The estimated distribution of wind data for 2000, when prior distribution is taken as 1999

V (m/s)	2000 (h)	f_i	f_w	f_{k1}	f_{k2}
0-1	650	0.3026	0.3270	0.3106	0.3041
1-2	3636	0.4152	0.3375	0.4080	0.4074
2-3	1488	0.1699	0.1970	0.1740	0.1790
3-4	521	0.0595	0.0893	0.0585	0.0608
4-5	254	0.0290	0.0338	0.0238	0.0245
5-6	114	0.0130	0.0111	0.0125	0.0126
6-7	49	0.0056	0.0032	0.0054	0.0052
7-8	16	0.0018	0.0008	0.0026	0.0024
8-9	8	0.0009	0.0002	0.0018	0.0016
9-10	7	0.0008	0.0000	0.0010	0.0009
10-11	3	0.0003	0.0000	0.0010	0.0008
11-12	8	0.0009	0.0000	0.0006	0.0005
12-13	0	0	0.0000	0.0001	0.0001
13-14	3	0.0003	0.0000	0.0001	0.0000
14-15	0	0	0.0000	0.0001	0.0001
15-16	0	0	0.0000	0.0001	0.0000
16-17	0	0	0.0000	0	0

Table 3. The estimated distribution of wind data for 2001, when prior distribution is taken as 2000

V (m/s)	2001 (h)	f_i	f_w	f_{k1}	f_{k2}
0-1	2297	0.2623	0.2903	0.2588	0.2621
1-2	3397	0.3880	0.3080	0.4004	0.3891
2-3	1638	0.1871	0.2031	0.1847	0.1878
3-4	722	0.0825	0.1102	0.0729	0.0762
4-5	332	0.0379	0.0524	0.0401	0.0422
5-6	155	0.0177	0.0224	0.0203	0.0211
6-7	89	0.0102	0.0088	0.0098	0.0100
7-8	48	0.0055	0.0032	0.0036	0.0036
8-9	38	0.0043	0.0011	0.0020	0.0019
9-10	20	0.0023	0.0003	0.0020	0.0018
10-11	10	0.0011	0.0001	0.0010	0.0008
11-12	5	0.0006	0.0000	0.0029	0.0023
12-13	2	0.0002	0.0000	0	0
13-14	3	0.0003	0.0000	0.0014	0.0010
14-15	0	0	0.0000	0	0
15-16	0	0	0.0000	0	0
16-17	0	0	0.0000	0	0

Table 4. Comparison of the actual probability distribution with the distributions obtained from GKOP and the Weibull distribution.

1999				
	RMSE	X^2	R^2	K-L
f_w	0.0204	0.00047515	0.9680	0.0376
f_{k1}	0.0057	3.7568e-005	0.9974	0.00969
f_{k2}	0.0050	3.0648e-005	0.9980	0.00801

As can be seen Table 1, Table 2, Table 3, the larger value of R^2 are obtained by using MaxMinxEnt subject to one and two constraint. The results shows that RMSE, X^2, K-L values of MaxMinxEnt in all years are lower than the values obtained by the Weibull distribution. As a results, MaxMinxEnt or the distributions obtained GKOP are suitable for the distributions of wind speed

Table 5. Comparison of the actual probability distribution with the distributions obtained from GKOP and the Weibull distribution.

2000				
	RMSE	X^2	R^2	K-L
f_w	0.0220	0.00055302	0.9586	0.0425

f_{k1}	0.0030	1.074e-005	0.9985	0.0030
f_{k2}	0.0031	1.206e-005	0.9995	0.0027

Table 6. Comparison of the actual probability distribution with the distributions obtained from GKOP and the Weibull distribution.

2001				
	RMSE	X^2	R^2	K-L
f_w	0.0222	0.00056336	0.9586	0.0405
f_{k1}	0.0041	1.931e-005	0.9986	0.0061
f_{k2}	0.0022	6.0358e-006	0.9996	0.0049

Figure 1. The probability distribution of actual data, the MaxMinxEnt f_{k1} f_{k2} distributions for 1999

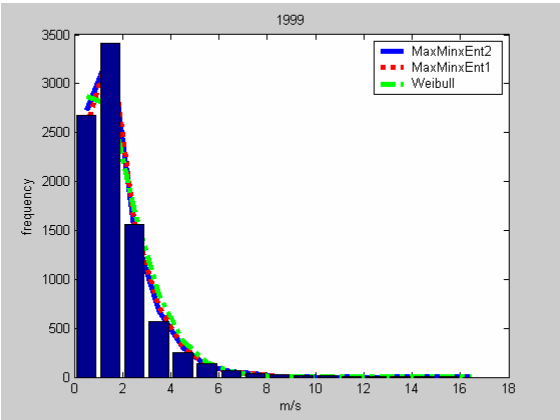
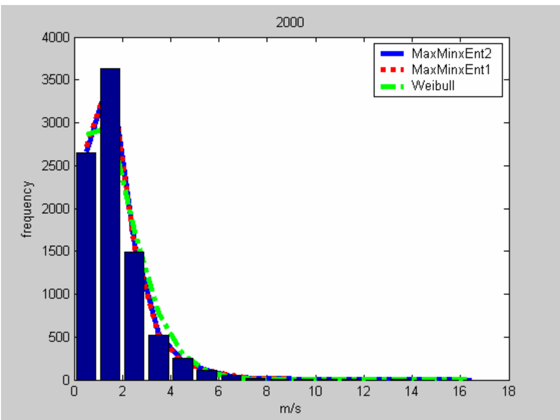
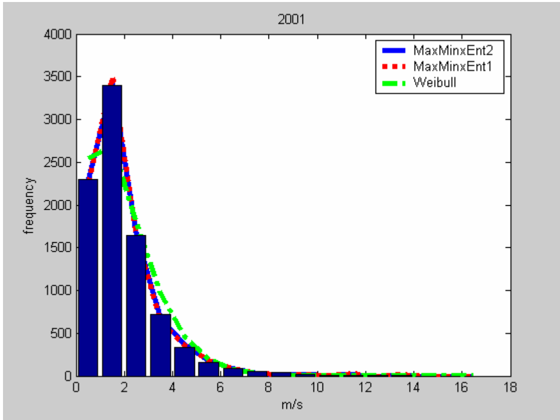


Figure 2. The probability distribution of actual data, the MaxMinxEnt f_{k1} f_{k2} distributions for 2000



The MaxMinxEnt and Weibull probability density functions of yearly wind speed data are seen inFigure 1, 2, 3

Figure 3. The probability distribution of actual data, the MaxMinxEnt f_{k1} f_{k2} distributions for 2001



4 Conclusion

The following main conclusions can be drawn from the present study:

1. The yearly distribution of wind speed was obtained through application of the GKOP
2. A comparison was made between the two parameters Weibull distribution and the distribution obtained through GKOP using various criteria. The distributions obtained GKOP are more suitable for distribution of wind speed.
3. Although principle based on information theory has been used in a variety of field, the obtained distributions GKOP is the first time to be applied to the wind energy field.
4. One of the main results is that the distributions obtained by GKOP also can represent- the wind power density much more accurately.

References:

[1]N. Ebrahimi, The maximum entropy method for lifetime distributions, *Sankhya: The Indian Journal of Statistics*, Vol.62, series A, 2000, 236-243.

[2]N. Ebrahimi, Families od distributions characterized by entropy, *IEEE Transactions on Information Theory*, Vol.47, 2001.

[3]N.C.Lind, Three information theoretical methods to estimates a random variable,

Journal of Environmental Management, Vol. 49, 1997, 43-51.
[4] N.C Lind., H.P. Hong ve V.A. Solona croos entropy method for Flood Frequency Analysis, *Stocastic Hydrology and Hydraulics*, Vol.3, 191-202.
[5] M.D. Pandey, Direct estimation of quantile functions using the maximum entropy principle, *Structural Safely*, Vol. 22, 1, 2000, 61-79.
[6] M.D.Pandey, Extreme quantile estimation using order statistics with minimum cross-entropy principle**,** *Probabilitistic Engineering mechanics*, Vol.16, 1, 2001, 31-42 .
[7]Pandey M.D. Minimum cross entropy method for extreme value estimation using peaks-over-threshold data, *Structural Safety* , Vol.23, 4 , 2001, 345-363.
[8] Li M., Li X. MEP-type distribution function: a better alternative to Weibull function for wind speed distribution, *Renewable energy*, 30, 2005, 1221-1240.
[9] M. Li, X. Li Investigation of wind characteristics and assessment of wind energy potential for Waterloo region, *Energy conversion management*, 46, 2005, 3014-3033.
[10] A Shamilov. Comparison of classical estimate methods based on maximum entropy. (*International conference, Ordered Statistical Data*). 2005
[11] W. Ximing. Calculation of maximum entropy densities with application to income distribution, *Journal of Econometrics*, Vol.115, 2003347-354.
[12] Kullback, S. and Leibler, R. A. On Information and Sufficiency, *Annals of Mathematical Statistics*, 22, 1951, 79–86.
[13] Kullback, S. *Information Theory and Statistics*. Wiley, 1959
[14] Kapur, J.N. and Kesavan, H. K. (1992), Entropy Optimization Principles with Applications*. London: Academic Press, Inc.*
[15] Genc a., Esiroglu M., Pekgor A.,Otuanc G., Hepbaslı A., Ulgen K. Estimation of Wind Power Potential Using Weibull distribution, *Energy Sources*, 27, 809-822, 2005.
[16] Celik A.N. A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey, *Renewable energy*, 29, 2003, 593-604.
[17] Akpınar E.K., Akpınar S. An assessment on seasonal analysis of wind energy characteristics and wind turbine characteristics, *Energy conversion management*, 46, 2005, 1848-1867.
[18] Weisser D. . A wind energy analysis of Grenada: an estimation using the ‘Weibull’ density function, *Renewable Energy*, 28, 1803-1812, 2003.
[19] Al-Nassar W., Alhajraf S, Al-Enizi A., Al-Awadhi L.. *Potential wind power generation in the State of Kuwait*, 30, 2005, 2149-2161.
[20] Ulgen K, Hepbasli A. Determination of Weibull parameters for wind energy analysis of Izmir, Turkey. *Int J Energy Res*, 26, 2002, 495-506