Proceedings of the 9th WSEAS International Conference on Applied Mathematics, Istanbul, Turkey, May 27-29, 2006 (pp605-610)

The Distribution of Minimizing Maximum Entropy: Alternative to Weibull distribution for wind speed

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

Abstract: In the present study, we have constructed MinMaxEnt distribution alternative to the Weibull distribution for the wind speed. MinMaxEnt distribution minimizes the special maximum entropy functional in the concrete set of moment functions, therefore, maximizes the amount of information generated by the given random sample. MinMaxEnt distribution of wind speed data measured in Eskisehir is also obtained and compared with the Weibull distribution by using various statistical criterias. Consequently, it is determined that MinMaxEnt distributions is more convenient for wind speed distributions, compared to the Weibull which is widely used in wind studies.

Keywords: MaxEnt; MaxEnt distribution; MinMaxEnt distribution, Wind speed data, root mean square error, chi-square, correlation coefficient, Akaike information criteria, Bayes information criteria.

1 Introduction

Recently the methods based on information theory to obtain the distribution of random variable are very popular. The maximum entropy (MaxEnt) [1] and the minimum cross entropy (MinxEnt) are such of these methods [2], [3], [4]. There has been a lot of study about application of these method in literature. In[5], MaxEnt method was employed to approximate the size distribution of U.S. family income and MaxEnt distributions was compared with two conventional income distribution. In [6]-[7], it was used the MaxEnt principle - statistical inference method for the first time to the energy field, and proposed a theoretical approach to determine distribution of wind speed data analytically. In mentioned studies, MaxEnt method operated according to wind's mass, momentum, energy [6] is proposed that the family of distributions is developed by introducing a pre-exponential term to the theoretical distribution derived from MaxEnt principle. In [7], MaxEnt models are used to determine the distribution of diurnal, monthly, seasonal and yearly wind speed.

In this study, it is given the definition of MinMaxEnt and comparison of the MinMaxEnt distribution and Weibull distribution which was widely used in wind power studies [6-12], by using statistical criterias and concluded that MinMaxEnt distribution can be more convenient to determine distribution of wind speed data .

The rest of the paper is organized as follows. Section 2 introduces MaxEnt method. In section 3, it is defined a functional on moment functions and properties of this functional are also given as theorems. It is shown that there is distribution which minimizes the special maximum entropy functional on moment functions. This distribution is called as MinMaxEnt. In Section 4, the estimation of the distribution of wind speed data measured in Eskişehir is obtained using defined distribution. The section 5 concludes the paper with some suggestions for of further research.

2. MaxEnt Method

2.1 MaxEnt method for discrete system

The MaxEnt is a method which proposes to estimate $p(x)$ by maximizing the entropy subject to some moment constraints. The estimated $p(x)$ is said to be MaxEnt density function.

The MaxEnt density function of discrete random variable (RV) X is obtained by maximizing Shannon's entropy measure

$$H(X) = -\sum_{i=1}^{n} p_i \ln p_i \quad (1)$$

subject to the constraints:

$$\sum_{i=1}^{n} p_i = 1$$
$$E\{g\} = \sum_{i=1}^{n} g(x_i) p_i = \mu, \quad (2)$$

where $g = (g^{(1)},...,g^{(m)})^T$ is moment vector function, $\mu = (\mu_1,...,\mu_m)^T$ is moment vector value.

By using Lagrange method, it is easily indicated that the resulting functional form of MaxEnt density function is

$$p_i = \exp(-\lambda_0 - \sum_{k=1}^{m} \lambda_k g_k(x_i)), \quad (3)$$

where $\lambda_0, \lambda_1,..., \lambda_k$ are Lagrange multipliers which can be found by Newton method.

2.2 MaxEnt Method for continuous system

The MaxEnt density function of continuous random variable X is obtained by maximizing Shannon's entropy measure

$$H = -\int_a^b f(x) \ln f(x) dx \quad (4)$$

subject to the constraints:

$$\int_a^b f(x) dx = 1 \quad (5)$$

$$E\{g\} = \int_a^b g(x) f(x) dx = \mu \quad (6)$$

where $g = (g^{(1)},...,g^{(m)})^T$ is moment vector function, $\mu = (\mu_1,...,\mu_m)^T$ is moment vector value.

Obtained MaxEnt density function is

$$f(x) = \exp(-\lambda_0 - \sum_{k=1}^{m} \lambda_k g_k(x)) .$$

3 Definition of MinMaxEnt Distribution

In this section, we give the definition of MinMaxEnt Distribution for discrete and continuous random variables(RVs)

3.1 MinMaxEnt Distribution for Discrete RV

Let $x_1,...,x_n$ be random sample (RS) with the corresponding probabilities $p_1,...,p_n$. It is required to obtain amount of maximum information generated by given RS on the basis of the given moment vector function $g_k = (g_k^{(1)},...,g_k^{(k)})^T$. RS has entropy

$$H = -\sum_{i=1}^{n} p_i \ln p_i \quad (7)$$

H has maximum value $\ln n$ subject to constraint

$$\sum_{i=1}^{n} p_i = 1 \quad (8)$$

Maximum of H subject to constraint (8) and moment constraints (9) represents a functional which is denoted by $S(g)$

$$\sum_{i=1}^{n} p_i g_k(x_i) = \mu_k, \quad (9)$$

where $g_k = (g_k^{(1)},...,g_k^{(k)})^T$ is moment vector function, $\mu_k = (\mu_k^{(1)},...,\mu_k^{(k)})^T$ is moment vector value. This functional possesses some properties expressed by

Theorem 1. If $S(g)$ is the functional derived by maximizing H subject to constraints (8), (9), then $S(g)$ is continuous on the set of continuous vector moment functions C[a,b].

Theorem 2. If $S(g)$ is the functional derived by maximizing H subject to constraints (8), (9), then $S(g)$ as continuous functional reaches its least and greatest values in the given compact set $K \subset C[a,b]$.

Suppose that $g^{(0)}$ vector moment function realizes minimum value of functional $S(g)$, i.e.

$$\min_{g \in K} S(g) = S(g^{(0)})$$

The distribution $p_1,...,p_n$ corresponding to $g^{(0)}$ is called as MinMaxEnt distribution and represented in the following form

$$p_i = \exp(-\lambda_0 - \sum_{j=1}^{m} \lambda_j g_j^{(0)}(x_i)),\ i = 1,...,n.$$

Definition of MinMaxEnt distribution is based on obtaining some moment vector function $g^{(0)}$ from the given set *K*. By virtue of the mentioned moment vector function, the maximum information is taken from the given RS. In this sense, MinMaxEnt distribution realizes a possible maximum information from RS.

In general, obtaining least value of MaxEnt functional $S(g)$ in a compact set of moment vector functions can be done by applying methods of calculus of variations (see [13]).

3.2 MinMaxEnt Distribution for Continuous RV

In order to maximize H defined by (4) subject to constraints (5), (6) it is applied Lagrange multipliers method. This method requires to construct the following auxiliary functional

$$H = -\int_a^b f(x) \ln f(x) dx - \lambda_0 (\int_a^b f(x) dx - 1)$$

$$\times \sum_{j=1}^{m} \lambda_j (\int_a^b g_j(x) f(x) dx - \mu_j)$$

or

$$H = -\int_a^b f(x) \ln f(x) dx - \sum_{j=0}^{m} \lambda_j (\int_a^b g_j(x) f(x) dx - \mu_j), \quad (10)$$

where $\mu_0 = 1,\ g_0(x) = 1$.

Solution of Euler equation for functional (10) has the form

$$f(x) = \exp(-\lambda_0 - \sum_{j=1}^{m} \lambda_j g_j(x)). \quad (11)$$

Substituting $f(x)$ in (5), (6) it is possible to obtain $\lambda_0,...,\lambda_m$. Then

$$S = H_{\max} = \lambda_0 + \lambda_1 \mu_1 + ... + \lambda_m \mu_m. \quad (12)$$

Theorem 3 If $S(g)$ is the functional derived by maximizing H given by (4) subject to constraints (5), (6), then $S(g)$ is continuous on the set of continuous vector moment functions $C[a,b]$.

Theorem 4. If $S(g)$ is functional derived by maximizing H subject to constraints (5), (6), then $S(g)$ reaches its least and greatest values in the given compact set $K \subset C[a,b]$ of continuous moment vector functions.

Now, we apply the defined MinMaxEnt distribution to wind speed data measured in Eskisehir in 2005.

4 MinMaxEnt Distribution function: a better alternative to Weibull distribution function for wind speed

The wind energy is one of the most significant and rapidly developing renewable energy sources in the world and it provides a clean energy resource. For this reason, the probability distribution of wind speed is one of the most important wind characteristics for assessment of wind energy potential and for the performance of wind energy conversion systems. In this section, it is shown that the MinMaxEnt distribution not only agree better with a variety of the measured wind speed data than the conventionally used empirical Weibull distribution, but also can present the wind power density much more accurately. Furthermore, the MinMaxEnt distribution obtained by MaxEnt theory is more suitable for the assessment of the wind energy potential.

The methodology presented in this study is applied to wind speed data.

In order to calculate MinMaxEnt distribution, the following steps are realized.

1. Determine moment constrains (x, x^2, $\ln x$, $\ln(x+1)$, $(\ln x)^2$, $\ln(x^2+1)$)
2. Calculate MaxEnt distributions subject to each of moment constrains
3. Calculate entropy of MaxEnt distribution
4. Calculate information contained by MaxEnt distributions.
5. Select moment constrain which has maximum information.
6. Determine the MinMaxEnt distribution by selected moment constraint

It is repeated above steps for two moment constraints. The corresponding moment constraint to maximum information are listed Table 1, Table 2 using wind data measured on December. This process can be repeated many times. It is clear that determined MinMaxEnt subject to the more constraints shows the better fitting.

Table 1. Entropy of the calculated MaxEnts subject to one constraint for measured data on December

Constrains	Entropy
x	4.1383
x^2	**4.0909**
lnx	4.2037
ln(1+x)	4.1912
$(\ln x)^2$	4.1584
$\ln(1+x^2)$	4.1930

Table 2. Entropy of the calculated MaxEnts subject to two constraints for measured data on December

Constrains	Entropy	Constrains	Entropy
x & x^2	4.0491	x^2 & $\ln(1+x^2)$	4.0577
x & lnx	4.0938	lnx & ln(x+1)	4.1523
x & ln(x+1)	4.0807	lnx & $(\ln x)^2$	4.1346
x & $(\ln x)^2$	4.0750	lnx & $\ln(1+x^2)$	4.1710
x & $\ln(1+x^2)$	4.0765	ln(x+1) & $(\ln x)^2$	4.1294
x^2 & lnx	4.0649	ln(x+1) & $\ln(1+x^2)$	4.1710
x^2 & ln(x+1)	4.0589	$(\ln x)^2$ & $\ln(1+x^2)$	4.1222
x^2 & $\ln(x)^2$	**4.0457**		

From Table 1 and Table 2, we conclude that MinMaxEnt under one constraint is the calculated MaxEnt subject to x^2 and MinMaxEnt under two constraint is the calculated MaxEnt subject to x^2&$\ln(x)^2$

The wind speed data in time-series format is usually arranged in the frequency distribution format. It is convenient statistical analysis. We fit the Weibull distribution to wind speed data, found the MinMaxEnt distributions under the one and two moment constraints for each months of 2005. Arrangement of the measured hourly time-series data in frequency distribution format for December and the frequency distributions calculated the Weibull and the MinMaxEnt distributions are given Table 3.

In the Table 3, f_i is probability density function (PDF) of observed data, $f_w(x)$ is the Weibull density function**,** $f_{mx1}(x)$ is the MinMaxEnt PDF subject to one constraint, $f_{mx2}(x)$ is the MinMaxEnt PDF subject to two constraint.

Table 3. Distribution calculated from Weibull, MinMaxEnt distributions for December

V (m/s)	Dec. (h)	f_i	f_w	f_{mx1}	f_{mx2}
0-1	47	0.0672	0.0369	0.0848	0.0611
1-2	40	0.0572	0.0686	0.0840	0.0544
2-3	38	0.0544	0.0844	0.0823	0.0627
3-4	57	0.0815	0.0917	0.0798	0.0714
4-5	57	0.0815	0.0930	0.0767	0.0784
5-6	48	0.0687	0.0900	0.0729	0.0827
6-7	62	0.0887	0.0841	0.0686	0.0840
7-8	49	0.0701	0.0764	0.0639	0.0822
8-9	41	0.0587	0.0677	0.0590	0.0775
9-10	54	0.0773	0.0587	0.0539	0.0706
10-11	48	0.0687	0.0500	0.0487	0.0622
11-12	39	0.0558	0.0418	0.0436	0.0531
12-13	35	0.0501	0.0344	0.0386	0.0438
13-14	33	0.0472	0.0279	0.0339	0.0351
14-15	20	0.0286	0.0223	0.0294	0.0272
15-16	18	0.0258	0.0176	0.0253	0.0205
16-17	6	0.0086	0.0137	0.0215	0.0150
17-18	4	0.0057	0.0106	0.0181	0.0106
18-19	3	0.0043	0.0080	0.0151	0.0073

Now, in order to compare MinMaxEnt distributions with the Weibull distribution, we can make use of suitability criteria to identify the best distribution from amongst considered distributions

Root mean square error (*RMSE*) [8], Chi-square (X^2), Correlation coefficient (R^2) Akaike's information criterion (*AIC*) [5],[15] Bayes's information criterion (*BIC*) Kullback-Leibler measure (K-L) [4] will be used in statistically evaluating the performance of MinMaxEnt and Weibull distributions.

The formula of mentioned suitability judgment criteria are

$$RMSE = \left(\frac{\sum_{i=1}^{N} (y_i - x_i)^2}{N} \right)^{\frac{1}{2}}, \qquad (15)$$

$$\chi^2 = \frac{\sum_{i=1}^{N} (y_i - x_i)^2}{N - n}, \qquad (16)$$

$$R^2 = \left(1 - \frac{\sum_{i=1}^{N}(y_i - x_i)^2}{\sum_{i=1}^{N}(y_i - z_i)^2}\right), \qquad (17)$$

$$AIC = -2\log(L(f(x,\theta)) + 2n, \qquad (18)$$

$$BIC = -2\log(L(f(x,\theta)) + n\log N, \qquad (19)$$

$$K - L = \sum_{i=1}^{N} y_i \log\frac{y_i}{x_i} \qquad (20)$$

where y_i is the ith probability of actual data, x_i is the ith predicted probability, 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*, X^2, *K-L* measure, *AIC, BIC,* the highest values R^2.

It is obtained that comparison of MinMaxEnt and Weibull distributions for every months of 2005 but given only October, December due to the space limitation.

Table 4. Comparison of the actual probability distribution with MinMaxEnt distributions and the Weibull distribution for October

October						
	RMSE	X^2	R^2	AIC	BIC	K-L
f_w	0.0152	0.0002529	0.7036	155.21	157.39	0.10474
f_{mx1}	0.0122	0.0001644	0.8074	151.35	153.53	0.06035
f_{mx2}	0.0103	0.0001240	0.8620	156.34	159.61	0.03827

Table 5. Comparison of the actual probability distribution with MinMaxEnt distributions and the Weibull distribution for December

December						
	RMSE	X^2	R^2	AIC	BIC	K-L
f_w	0.0153	0.000262	0.6431	125.50	127.39	0.09649
f_{mx1}	0.0147	0.000243	0.6694	120.76	122.65	0.06528
f_{mx2}	0.0084	8.32e-005	0.8934	125.36	128.19	0.02009

Table 4, Table 5 shows the excellent agreement of the MinMaxEnt densities with wind data, much better than the corresponding empirical Weibull distribution in terms of all criterias. For this reason the MinMaxEnt distributions are much convenient for wind speed data than the Weibull since all criterions favour MinMaxEnt distibutions.

Figure 1. Histogram and obtained Weibull and MinMaxEnt densities for wind speed data

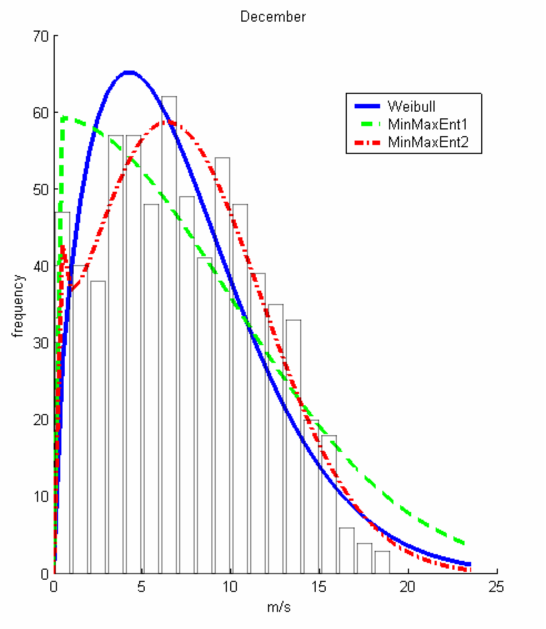

Figure 1 reports the histogram of the wind speed data and the estimated MinMaxEnt densities with one and two constraints and the Weibull density. From this Figure, obtained MinMaxEnt densities demonstrates good fitting when wind speed values<5 and at both tails.

4 Conclusion

In the present study, we have constructed MinMaxEnt distribution alternative to the Weibull distribution for the wind speed. By comparing the MinMaxEnt distributions with the Weibull distribution on wind speed data, we concluded that MinMaxEnt distributions is more convenient for wind speed distribution. Moreover, it can be shown that estimate of wind power can be studied based on MinMaxEnt distribution.

References:

[1] Jaynes, E. T. Papers on Probability, Statistics, and Statistical Physics. Netherlands: D. Reidel, Dordrecht, 1983.

[2] Kullback, S. and Leibler, R. A. On Information and Sufficiency, *Annals of Mathematical Statistics*, 22, 1951, 79–86.
[3] Kullback, S. *Information Theory and Statistics*. Wiley, 1959.
[4] Kapur, J.N. and Kesavan, H. K. (1992), *Entropy Optimization Principles with Applications.* London: Academic Press, Inc.
[5] Wu, X. Calculation of maximum entropy densities with application to income distribution, *Journal of Econometrics*, 2003, 115, 347-354.
[6]Li M., Li X. MEP-type distribution function: a better alternative to Weibull function for wind speed distribution, *Renewable energy,* 30, 2005, 1221-1240.
[7] Li M., Li X. Investigation of wind characteristics and assessment of wind energy potential for Waterloo region, *Energy conversion management*, 46, 2005, 3014-3033.
[8] 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.
[9] 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.
[10] Weisser D. A wind energy analysis of Grenada: an estimation using the 'Weibull' density function, *Renewable Energy,* 28, 1803-1812, 2003.
[11] Al-Nassar W., Alhajraf S, Al-Enizi A., Al-Awadhi L.. *Potential wind power generation in the State of Kuwait*, 30, 2005, 2149-2161.
[12] Ulgen K, Hepbaslý A. Determination of Weibull parameters for wind energy analysis of Izmir, Turkey. *Int J Energy Res*, 26, 2002, 495-506
[13] Elsgolts, L. *Differential Equations and the calculus of variations*, Mir Publisher, 1970
[14] Shamilov A., Kantar Mert Y. *On a Distribution Minimizing Maximum Entropy*, International conference, Ordered Statistical Data, 2005.
[15] Kenneth B. *Model selection and Multi-Model Inference*, Springer-Verlag, 2002.