Journal of Theoretical and Applied Information Technology
28th February 2013. Vol. 48 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DECIDING WEIGHING BY ENTROPY EVALUATION METHOD IS AN ERROR

RUNCHUN HAN, JIXIAN XIAO, JIE YANG, HONGWEI DU

Hebei United University

ABSTRACT

In physics，heat energy divided by temperature，entropyis obtained, which means the conversion degree from heat to work. While in science and technology field, entropy generally refers to possible degree of certain matter system status. In social science area, entropy indicates the degree of certain society state, in the field of information theory, entropy expresses uncertain degree. In a word, the application of “entropy” is widely used. However, there’s no scientific theory support for deciding weighing by entropy evaluation method. This paper give a detailed explanation of it in three aspects to demonstrate that scientific is not decide weighing by entropy evaluation method.

Key words: *Entropy Evaluation Method; Weighing; Evaluation System*

1. INTRODUCTION

"Entropy", created by German physicist Clausius (Rudolf Clausius, 1822 - 1888) in 1850, it is used to indicate the degree of uniformity of any energy distributed in the space. The more uniform distribution of energy, the greater the entropy is. For the observed system, if the energy equably distributes, then the entropy of the system achieve its maximum. Clausius held that it is one of universal laws for the energy density tends to uniformity is. In other words, "entropy will increase with time."

2. METHOD

In recent years, some people use entropy measurement methods to select the weight of the indicator system evaluation, from the master's thesis, doctoral dissertation, to research papers and scientific research projects, such method is widely used and known as “the law of objectively valuing weights”. Just type "entropy" and "weight" to search for key words or names; you will find a lot of stuffs about the entropy evaluation method to determine the weights. According to some references [1][2], the method is presented as follows:

Suppose $X=(x_{ij})_{m\times n}$, where $x_{ij}(i=1,2,\dots,n; j=1,2,\dots,n)$ means the i^{th} observation of the j^{th} indicator.

Firstly, because the observations may have different units, it should normalize the data. There are some entropy evaluation methods to find weights, here we use

$$d_{ij}=\frac{x_{ij}-x_j^*}{\bar{x}_j-x_j^*},$$

where $\bar{x}_j=\max_{1\le i\le m} x_{ij}$ *and* $x_j^*=\min_{1\le i\le m} x_{ij}$

Secondly, calculate the probability of d_{ij}

$$p_{ij}=\frac{d_{ij}}{\sum_{i=1}^{m} d_{ij}},\ i=1,2,\cdots,m,\ j=1,2,\cdots,n$$

Then, we need entropy value as follow

$$e_j=\begin{cases}-\dfrac{1}{\ln m}\sum_{i=1}^{m} p_{ij}\ln p_{ij} & when\ \ p_{ij}\neq 0\\ 0 & when\ \ p_{ij}=0\end{cases}$$

where $j=1,2,\cdots,n$

Then we can get the weight λ_j

$$\lambda_j=\frac{1-e_j}{\sum_{j=1}^{n}(1-e_j)}\qquad j=1,2,\cdots,n$$

Journal of Theoretical and Applied Information Technology
28th February 2013. Vol. 48 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

At last, calculate the integrated evaluation for the i^{th} sample

$$s_i = \sum_{j=1}^{n} \lambda_j d_{ij}$$

3. ANALYSIS

Indubitably, to carry out any evaluation has its specific purpose, and the indicator system is closely associated with this purpose, while the weights is used to reflect the significance degree of each indicator by evaluators to achieve an evaluation purpose, without such purpose, the system is worthless. It is difficult to find out what kind of purpose and background Charies Spearman had to use entropy evaluation method to determine the weights, and how to prove such weights in line with the wishes of the evaluators, at least, however, now people use this method without sufficient demonstration or deduction. Therefore, I believe that the application of the method to determine weight lack of theoretical basis. This paper will, expound that adopting this method to determine the weights of evaluation system is unscientific from three aspects.

3.1 Data

We randomly choose 30 schools as the samples from the report about “the world university rankings 2008” published by the Times, for convenience we call the "Sample 1.", we expound that adopting this method to determine the weights of evaluation system is unscientific from three aspects. The data collected in

3.2 It is precarious to make the contribution rate of principal component as the evaluation weights when sample size is changing

We firstly apply the entropy evaluation method to determine the weight, and get the ranking of 30 schools (calculation omitted), as Rank 1 in Table 7. Then, remove the 4th school, and name it “Sample 2”, and with the same method, we calculate Rank 2 as in Table 7. Because the 4th school in the "Sample 1" ranks the 11th, according to common sense, after the removing, the top 10 schools should stay in the same places, and the others start from the 12th school should raise by 1 position. It is a feature that increase/decrease the sample size is stable for the evaluation. However, from Rank 1 and 2 in Table 7, we can find that, except the 24th school, the rankings of others are completely disordered, for example, the rankings of schools with the index 15, 20, 25, 28 in “Sample 1” have changed from the 10th, 20th, 18th, and the 12th to the 3rd, 13th, 10th, and the 5th. Using SPSS software to take the Kappa Testing and get Table 2.

Applying the entropy evaluation method, we can get the integrated position of the 15 students. Then we can remove the student in the first place, and re-rank the rest 14 students. The results are presented as following Table 2.

Table 1, The OrignInal Data Table

RANK	PEER REVIEW SCORE	EMPLOYER REVIEW SCOR	STAFF/STUDENT SCORE	CITATIONS/STASS SCORE	INTERNATIONAL STAFF SCORE	INTERNATIONAL STUDENT SCORE
1	100	100	96	100	87	81
2	100	99	98	91	78	83
3	100	99	98	94	29	89
4	100	99	90	96	28	76
5	93	98	89	70	91	85
6	96	99	82	70	91	82
7	85	84	80	57	98	86
8	98	99	52	37	99	99
9	87	59	81	58	54	81
10	59	95	74	62	92	99
11	63	90	61	61	87	83
12	65	73	49	80	71	88
13	72	89	39	61	87	95
14	72	91	67	33	95	82
15	51	72	80	72	61	100
16	89	67	10	67	64	85
17	73	75	36	53	100	88
18	71	35	47	75	63	81
19	62	32	98	34	84	80
20	41	35	90	80	42	97
21	66	17	89	40	24	90
22	44	71	78	51	91	80
23	55	66	77	23	96	92
24	54	51	66	57	30	80
25	59	58	47	46	70	98
26	37	59	63	71	77	95
27	61	66	15	68	86	82
28	65	87	18	41	88	100
29	62	50	46	39	71	85
30	52	58	77	28	46	89

Since p = 1, so Rank 1 and 2 are not consistent. It is generally believed that the Kappa Value ≥ 0.75 shows a good consistency, 0.75> Kappa Value ≥ 0.4 has a general consistent, and if the Kappa Value <0.4, the two factors have a poor consistency [3]. Because Kappa Value <0.001 <0.4, the Rank 1 and 2 do not have consistency. This result shows that

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

using entropy to determine the weight does not have stability when sample size is changing, and the unstable ranking list is worthless.

3.3 It is precarious to make the contribution rate of principal component as the evaluation weights when sample value changed.

Normally, for an ordered indicator array, changing the value of observation will move up or drop behind its position of evaluation; while the observations out of the changing range will keep their original places, namely, for a school ranks the ith, if the observation value changed, and ranking turns into the jth. When i>j, the schools ranked the kth(j≤k<i) should be the k+1th; when j>i, When i>j, the schools ranked the kth(j≤k<i) should be the k-1th. We call this characteristic the stability of changing sample value. However, if we change the index 4 school's INTERNATIONAL STUDENT SCORE value from 76 to 80 (named "Sample 3"), and use entropy evaluation method to calculate Rank 3, we can find two abnormal problem. First, the index 4 school's ranking has not increased, but fall from the 11th to 19th.

Second, even if the ranking dropped from the 11th to the 19th is recognized, according to the usual understanding, the original 12th to 19th should ramge from the 11th to 18th in turn, and other schools rank the same, however other schools' positions are totally out of order (see Table 7 Rank 4). After a Kappa Testing on Rank 3 and 4, we get Table3.

Table 2 Symmetric Measures

		Value	Asymp. Std. Errora	Approx. Tb	Approx. Sig.
Measure of Agreement	Kappa	.000	.035	.000	1.000
N of Valid Cases		29			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Table 3 Symmetric Seasures

		Value	Asymp. Std. Errora	Approx. Tb	Approx. Sig.
Measure of Agreement	Kappa	.034	.047	1.017	.309
N of Valid Cases		30			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Since p = 0.309> 0.05 and Kappa Value = 0.034 <0.4, so Rank 3 and 4 do not have consistent. The cause of such a big change is only increasing 4 points of INTERNATIONAL STUDENT SCORE. It indicates that, when sample values shift, using entropy evaluation method to determine the weight lack of stability, which also makes the ranking unauthentic.

3.4 Entropy is the inherent characteristics of the sample data, and can not reflect the significance of the indicator

In physics, entropy is obtained from heat energy divided by temperature, which represents the conversion degree from heat to work. Boltzmann said that, "When the energy reduced, the atom shows a more disordered state." Entropy is a measurement of disorder, namely, the probability of a special status -- the number of atoms in different gather ways, which can be formulized as $S = K \times \log W$, where S is the entropy, which is in direct rate with the logarithm of probability of given state W, K is the Boltzmann constant.

In the information theory, entropy is a measurement of information. Shannon, the founder of information theory, proposes an information measurement based on a statistical model in his book "Mathematical Theory in Communication". He defined information as "to eliminate the uncertainties." If there is a random trial with N possible value, and corresponding probabilities are $p_1, p_2, \dots, p_N$, then the average of information is, $H = -\sum_{i=1}^{N} p_i \ln p_i$, where N is called entropy. In the information theory, entropy can be used as a measurement for the event uncertainty, and will decline as the amount of information increases, the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

structure of system becomes more regular, or the function of system gets more comprehensive.

No matter in terms of the physical point of view or from the perspective of information theory, entropy is a measurement of disorder, but not reflects the characteristic of indicators in the evaluation system. We know that in an evaluation index system, the weight is the significance level of the evaluated index, the same index with different purpose can have different significance level, and for an evaluation system, each indicator weight should be only be determined as one true value to reflect the important degree of indicators, even though this value is too difficult to find. For the previous examples, the indicator value should be determined by the character of courses, but not by the grade distribution of each course. On the one hand, based on "Sample 1" and the ranking, the entropy of INTERNATIONAL STUDENT SCORE is larger, so the weights of the index are greater. And the weights, 0.18715, 0.11777, 0.1394, 0.18314, 0.16651, 0.20603, indicate INTERNATIONAL STUDENT SCORE is more important than other index in the evaluation course. In the "Sample 2", its weights are 0.15711, 0.10025, 0.11975, 0.15198, 0.11777, 0.35314 respectively, and the weights of "Sample 3" are 0.14938, 0.094, 0.11126, 0.14617, 0.1329 and 0.36629. The data above show that to increase/decrease in the sample size, or change the observations is significant for the weights, and the ranking is irresponsible. On the other hand, supposing there is a "Sample 4" with 30 schools and as the same scores as "Sample 1", namely, the INTERNATIONAL STUDENT SCORE of "Sample 4" is equal to the INTERNATIONAL STUFF SCORE of "Sample 1", and the INTERNATIONAL STUDENT SCORE of "Sample 1" is the same as the INTERNATIONAL STUFF SCORE of "Sample 4". So the weights in "Sample 4" are 0.18715, 0.11777, 0.1394, 0.18314, 0.16651 and 0.20603, which means the index, INTERNATIONAL STAFF SCORE, is the most important, and it is clearly not in line with our objective of the original evaluation. It is because that the entropy is the inherent characteristics of the sample data, but not reflects the significance level of indicators characteristics of the data, which does not reflect the importance of indicators, therefore the weights calculated from entropy do not indicate the authenticity of the evaluation.

From the university rankings published by the Times, the weight coefficients are 0.4, 0.1, 0.2, 0.2, 0.05, 0.05 according to importance, and based on such weights, we obtained Rank 5 in Table 7.

Compare Rank 1 with Rank 5, only 3 schools have the same orders, and there are 17 schools move more than 4 rankings, even more, the 28th school raise 16 rankings. We make a Kappa testing on Rank 1 and 5 by SPSS, and get table 4.

Since $p = 0.042 < 0.05$, we can not exclude Rank 1 and 5 for they have a certain consistency, but also because Kappa Value=0.069 <0.4, indicates the consistency is very weak. This shows that the entropy evaluation method that used to determine the weights and the ranks is inconsistent, so Rank 1 is worthless.

We find that the ranking of "Sample 3" calculated with the Times' weights is the same as the ranking of "Sample 1", which means the changing of INTERNATIONAL STUDENT SCORE (4 scores) of the 4th school does not change the ranking. However, compare Rank 1 with 3, the whole ranking of the 30 schools have changed. And based on Kappa testing, we get Table 5.

Table 5 Symmetric Measures

		Value	Asymp. Std. Errora	Approx. Tb	Approx. Sig.
Measure of Agreement	Kappa	-.034	.000	-1.017	.309
N of Valid Cases		30			

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.

Table 6 Symmetric Measures

		Value	Asymp. Std. Errora	Approx. Tb	Approx. Sig.
Measure of Agreement	Kappa	.000	.034	.000	1.000
N of Valid Cases		30			

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Since p = 0.309> 0.05, so Rank 1 and 3 do not have consistency, and Kappa Value =- 0.034 <0.4, also shows that. The huge different ranking between the Times' weights and entropy evaluation method show that, when sample values shift, using entropy evaluation method to determine the weight is lack of stability, which also makes the ranking unauthentic. And make the Kappa Testing on Rank 3 and 5 from Table 7

Table 7 Overall Ranking Table

Index	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
1	1	8	8	1	1
2	2	7	7	2	2
3	5	4	9	5	3
4	11	removed	19	19	4
5	4	9	18	4	5
6	7	14	12	7	6
7	8	12	1	8	7
8	3	1	5	3	8
9	21	22	23	21	9
10	6	2	14	6	10
11	15	19	3	14	11
12	14	17	21	13	12
13	9	6	22	9	13
14	17	21	16	16	14
15	10	3	2	10	15
16	22	20	27	22	16
17	13	18	6	12	17
18	25	27	10	25	18
19	26	29	13	26	19
20	20	13	11	20	20
21	29	23	26	29	21
22	23	28	15	23	22
23	19	16	29	18	23
24	30	30	24	30	24
25	18	10	28	17	25
26	16	15	20	15	26
27	24	25	30	24	27
28	12	5	17	11	28
29	27	26	25	27	29
30	28	24	4	28	30

Since p = 1 and Kappa Value <0.001 <0.4, the Rank 3 and 5 do not have consistency. It can be seen that use the entropy evaluation method to determine the weight, the ranking results lack credibility. To sum up, it is a misapplication to use the entropy evaluation method to determine the weight of revaluation, such method is lack of scientific basis, which can not indicate the significance of the evaluation indicators, and the evaluation result is just nonsensical.

4. CONCLUSION

As a result, we draw the following conclusions:1. Weight maked by Entropy is instability for evaluating the change of entity, thus the instability makes the really rank is invalid; 2. Eight maked by Entropy is instability for evaluating the change of observation, thus the instability makes the rank of weight is invalid. Based on those, we think Entroy isn't a scentific method, and we will study on the reasonable scope of the weight of entropy, and study on the new way of making weight.

ACKNOWLEDGEMENT

Hebei Society Science Funds：HB11GL061，Strategy of Hebei Ports' Services Industry

REFERENCES

[1] Jinhui Li, Xuhong Li, and Jie He, *Transport and Computer[J]* , April 2008, pp. 193–196.

[2] Huiyuan Jiang, and Hao Wang. Water Carriour Engineer[J], July 2008, pp.1-3.

[3] Wentong Zhang, and Jie Yan, *SPSS Statistical Analysis Basic Testbook[M]*, Beijing Higher Education Press, 2004.9, pp.318

[4] Xiaohai Han, Yaohui Zhang, Fujun Sun, Shaohua Wang, *Journal of sichuan ordnance*, October 2012, pp.124-126

[5] Xiuying Guo, *Statistics and Decision*, 2012[17], pp.32-34

[6] Li Zhao, Yongming Zhu, Meichen Fu, Pengtao Zhang, Yingui Cao, *Transactions of the Chinese Society of Agricultural Engineering*, April 2012, pp.235-242