Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 45 (2012) 101 – 105

Procedia Engineering

www.elsevier.com/locate/procedia

2012 International Symposium on Safety Science and Technology

Kappa analysis of industrial park risk evaluation index system

HUANG Wenhong*, WANG Hailong, LI Xuesheng, BAO Qifu

Zhejiang Province Key Laboratory of Safety Engineering and Science Research, Zhejiang Province Institute of Safety Science and Technology, Hangzhou 310012, Zhejiang, China

Abstract

In the industrial park regional risk evaluation process, expert evaluation method is usually used to budget the constructed index system, so that the overall park risk value can be quantified. During setting the estimation system and expert assignment, in order to comply with the objective factor, Kappa statistics analysis method is adopted to evaluate the consistency analysis of the system setting of regional risk estimation and the expert's assignment situation. The Kappa statistic is an statistical index which via comparing two or more observers on the same things, or the twice or more observation results of one observer on the same thing ,and the difference between the consistency and the consistency of actual observed value ,which is caused by opportunity or by actual measurement, is used as statistical index. Kappa statistic and weighted Kappa statistic not only can be used to inspect the consistency and the reproducibility of ordered and unordered variable data, but also can give a magnitude which reflects the consistency. During the course of assessment, experts coming from different departments give respective grades of safety conditions in industrial park. Then indexes setup and grades by experts are discussed. The results indicate that grades by experts show good consistency.

Keywords: industrial park; regional risk evaluation; kappa statistic; consistency

1. Introduction

At present, the research on risk evaluation and safety management for the industrial park have become a hot point in the field of public safety. Many domestic regions and units are carrying out industrial park safety plan and risk evaluation. However, there is a core problem in safety plan, risk evaluation and safety management: how to judge the risk level? What can be used to measure the security level? What level can industrial park reach? What primary insurance can get for safety production?

Risk evaluation index system is a quantifiable system reflecting the relationship and the important degree of each part, also a scale to judge the risk level and the criterion to safety case [1-2]. The author established risk evaluation index system framework for industrial parks, by taking Zhejiang Province major science and technology projects and a series of regional risk evaluation projects. And the index system has been applied in some industrial parks.

In the regional risk evaluation process of industrial park, using the evaluation index system for risk evaluation, the statistics significance is usually receive attention and query, because it is related to whether the final evaluation results can accurately reflect the actual risks of the park. The author combines the regional risk evaluation in an industrial park, the Kappa analysis and statistics are used for a further verification and analysis of the park risk evaluation index system.

Corresponding author. Tel.: 86-571-88083488; fax: 86-571-88083488.
E-mail address: huangwenhonghz@163.com

1877-7058 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.proeng.2012.08.128

2. Kappa analyses

2.1. Concept

During the research, the problem that two kinds of inspection or determination results are consistent and whether two responses can reproduce is frequently arise. The analysis methods which were often used, such as t test, x^2 test and simple correlation, however, there are some limitations. Aforementioned statistical method can not definitely reflect the consistency of index factors. For example: the coincidence rate includes a considerable part of opportunities; x^2 test can only distinguish whether two difference was statistically significant, but cannot reflect the consistency; Related factors can only express related, but cannot express its indeed consistent. Even in some cases, analyses a same data with different analysis methods may come to completely opposite conclusion.

The Kappa statistic [3-6] is an statistical index which via comparing two or more observers on the same things, or the twice or more observation results of one observer on the same thing, and the difference between the consistency and the consistency of actual observed value, which is caused by opportunity or by actual measurement, is used as statistical index. Kappa statistic and weighted Kappa statistic not only can be used to inspect the consistency and the reproducibility of ordered and unordered variable data, but also can give a magnitude which reflects the consistency. When a response as the unordered variables or two variables use Kappa statistic; when a Response as the ordered variable, use weighted Kappa statistic.

2.2. Kappa statistic of two evaluators

Two evaluators according to the result of n indicators of park risk evaluation indexes to judge which a class in C categories, Judgment results can be written as in table $C \times C$ form. We can use Kappa statistic consistency test to determine whether the two evaluators judged results are consistent.

$$\text{Kappa} = \frac{F_0 - F_e}{n - F_e} = \frac{F_0 - P_e}{n - P_e} \tag{1}$$

In this formula:

$F_0 = \sum O(i,j)/n$ Means two evaluators judged consistent with the observed count value

$O(i,j)$ Means the assessed A_1 be judged to Class I; The assessed A_2 be judged to Class J

$F_e = \sum O_1(i)O_2(i)/n$ Means two evaluators judgment consistency by the opportunity to gauge theory of value, the value come from Evaluator classification completely independent.

$P_0 = F_0/n, P_e = F_e/n$ Means the observed frequency and theory frequency Correspond to F_0 and F_e.

For the H_o (Kappa = 0) and H_1 (Kappa> 0) hypothesis test, Landis and Koch Recommendations for the use of the table below the acceptable range of Kappa coefficients.

Table 1. Kappa coefficient assignment range

Kappa	Consistency degree	Kappa	Consistency degree
＜0.00	Very poor	0.41-0.60	moderate
0.00-0.20	Litter poor	0.61-0.80	better
0.21-0.40	Poor	0.81-1.00	best

The Table 1 can be referenced, but lack of reliability assurance, then Schouten Pointed out that the available Z test

$$Z = \frac{\text{Kappa} - 0}{\sqrt{Var(\text{Kappa})}} \tag{2}$$

In this formula, Var(Kappa) means Kappa statistic's variance under the hypothesis which the Evaluator classification completely independent. The Z approximation to the standard normal distribution. So we can use $Z_{0.05} = 1.645$, $Z_{0.01} = 2.326$ to make statistical inference. When the judging categories appear in order, like as: Considering the $O(1, 2)$ and the $O(1, 3)$ has a different consistency. The former expresses two evaluators judged results is only have a difference of one level, while the latter two levels. So, introducing weight $\omega(i, j)$ which convicted for class I by first Evaluator, for class J by second Evaluator.

For Linear weight $w(i,j) = 1 - \frac{|i-j|}{c-1}$, for Square weight $w(i,j) = 1 - \frac{(i-j)^2}{(c-1)^2}$.

The weighted P_0, P_e value:

$$P_{0(w)} = \sum_i \sum_i O(i,j)w(i,j) / n \tag{3}$$

$$P_{e(w)} = \sum_i \sum_i O_1(i)O_2(j)w(i,j) / n^2 \tag{4}$$

The weighted Kappa statistic:

$$\text{Kappa}(w) = \frac{P_{0(w)} - P_{e(w)}}{1 - P_{e(w)}} \tag{5}$$

2.3. Kappa statistic of multiple evaluators and multiple classes

Assume there are K evaluators (A_1, A_2, ….., A_k) according to the result of n indicators of park risk evaluation indexes to judge which a class in C categories it belong to in this case.

$$P_0 = \frac{1}{nk(k-1)} \sum_l \sum_{m^2 l} \sum_i O_{lm}(i,i) \tag{6}$$

In this formula: $O_{lm}(i,i)$ means the evaluators A_1 ...A_m be judged to Class I;

$\sum_i O_{lm}^{(i,i)}$ Means the evaluators A_1 ...A_m judged consistent with the observed count value;

$\sum_l \sum_{m\square} \sum O_{lm}(i,i)$ Means the observed count value which be judged consistent by any of the two evaluators.

Assume there are N indicators are judged a class I is $O_l(i)$ by evaluator A_l, Then $O_l(i) = \sum_j O_{lm}(i,j)$, so we get the formula:

$$P_e = \frac{1}{n^2 k(k-1)} \sum_l \sum_{m^2 l} \sum_i O_l(i)O_m(i) \tag{7}$$

Same with two evaluators, for ordinal categorical data with linear or square weighting $\omega(i,j)$ Kappa(w) 's Approximate formula is:

$$Var[\text{Kappa}(w)] = \frac{2\sum_l \sum_{m^2 l} u_{lm}}{n^5 k^2 (k-1)^2 [1 - P_e(w)]^2} \tag{8}$$

And:

$$u_{lm} = \sum_i \sum_j O_l(i)O_m(j)[nw(i,j) - \sum_{i_1} O_l(i_1)w(i_1,j)] - \sum_{j_1} O_m(j_1)w(i,j_1)]^2 - [\sum_i \sum_j O_l(i)O_m(j)w(i,j)]^2 \tag{9}$$

3. An industrial park risk evaluation index Kappa analysis

3.1. An industrial park risk evaluation procedure

June 2011, we worked in an industrial park of regional risk evaluation. According to the established regional risk evaluation index system (As shown in Table 2).

Using the Table 1 index system, ten experts from safety management, enterprise, professional research institute were chosen to build a risk assessment panel. Then we scored these indexes according to the park's actual situation. The analysis curve about experts score's average value and standard deviation shown in Fig 1. As seen from the graph. Standard deviation curve has a certain fluctuation and change; it means there are some differences in expert's evaluating. In-depth analysis can know: B_1 (*Park economic carrying capacity*), C_5 (*Fire explosion and poisoning scope of influence's standard deviation*) is the highest. It means experts have relatively large differences in these two indexes evaluation process. However, they had basic agreement to A_4 (*Business impact on the surrounding environment*), A_1 (*Planning and compliance*), C_1 (*Hazardous material property*), C_9 (*The perfect degree of safety management*), D_2 (*safety management emergency resource*), and the other views are slightly different.

3.2. Kappa analysis about industrial park risk assessment index system data

In this industrial park risk evaluation process, ten experts to judge these indexes which category belongs to 0.1.2.3. It is a multiple categories and more raters' consistency checking problem (c=3, k=9, n=405), On the Kappa statistics were

calculated using Matlab 6. 5 available, we can get the following result:

1) Without using a weighted Kappa statistic, Kappa =0. 51, z = 84.13, p = 1.

2) The weight is 1, Kappa (ω) = 0. 59, z = 82. 85, p = 1.

Table 2. Industrial park risk evaluation index system

The first level index	No.	The second level index
A The park overall planning	1	A_1 Planning and compliance
	2	A_2 The surrounding population density
	3	A_3 The surrounding traffic
	4	A_4 Business impact on the surrounding environment
B Socio economic status	5	B_1 Park economic carrying capacity
	6	B_2 Public service investment proportion
C Enterprise safety production situation	7	C_1 Hazardous material property
	8	C_2 The amount of hazardous materials
	9	C_3 The production process of danger
	10	C_4 Potential occupation disease harm
	11	C_5 Fire explosion and poisoning scope of influence
	12	C_6 Waste disposal
	13	C_7 Building fire proof grade
	14	C_8 Safety monitoring of protective measures for perfection
	15	C_9 The perfect degree of safety management
D The emergency rescue capability	16	D_1 Fire emergency resource
	17	D_2 Safety management emergency resource
	18	D_3 Medical resources
	19	D_4 Emergency organization supporting condition
E The mutual influence between the factors of enterprise	20	E_1 Between enterprise and the rationality of layout
	21	E_2 The complexity of the factors affecting each other between the enterprise
	22	E_3 Cascading accident severity
	23	E_4 Cascading accident probability

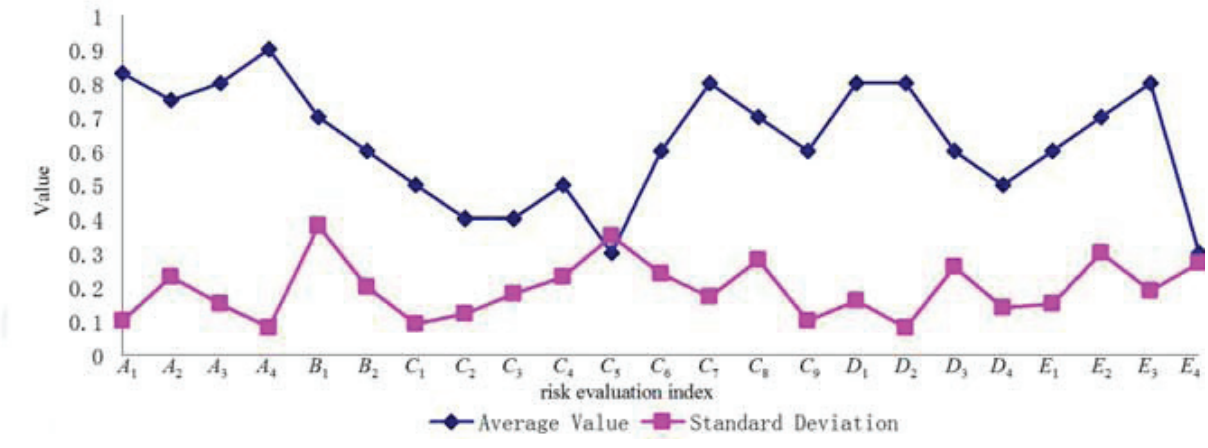

Fig. 1. An industrial park risk evaluation index system data curve.

3) The weight is 2, Kappa (ω) = 0. 68, z = 72.15, p = 1.

In these calculation. When significance level α= 0. 05, ten experts judgment has good consistency. We can get the same

result from Fig 1, the first two cases of Kappa value between 0.41-0.60, the consistency in a Moderate degree, but when the weight is 2, Consistency degree obviously.

4. Conclusions and prospects

The author using industrial park regional risk evaluation index system, selecting a chemical industrial park regional risk assessment to a project, using kappa for assessment of data obtained from a statistical analysis.

1) Kappa statistic and weighted Kappa statistic not only can be used to inspect the consistency and the reproducibility of ordered and unordered variable data, but also can give a magnitude which reflects the consistency.

2) Analysis of a chemical industry park regional risk evaluation index standard deviation curve, the results showed there are some differences in evaluating the same category.

3) Using Kappa to analyze expert assignment, the results showed when the weight reaches 2, the expert assessment of data consistency degree from different branches is improved significantly.

Refining and screening of the industrial park risk evaluation index system is a gradual process, Should keep renewing with the improvement of regional risk bearing capacity . The whole park security positioning and the grade of these indexes has not published the relevant standards in our country. So we have not o be standardized, standardization and systematization to the grading range, expert assignment of the industrial park risk evaluation index system. And kappa analysis can provide effective and reasonable methods to this work. It is worth to researching and expanding.

Acknowledgements

The support from the Laboratory construction program of the Science & Technology Agency of Zhejiang Province (2011F10011) is acknowledged.

References

[1] Jia Wei, Zhu Jianxin, Gao Zenglian, Bao Qifu. 2009. The Method of Quantitative Area Risk Assessment and Its Application in Chemical Industrial Park. China Safety Science Journal 19(5), p. 140.
[2] Chen Xiaodong, Duo Yingquan. 2009. Researches on Safety Capacity of Chemical Industry Park. Journal of Safety Science and Technology 5(2), p. 10.
[3] Zheng Shuangzhong, Deng Yunfeng,Jiang, Tian Han. 2006. Kappa Analysis on Assessment of Urban Emergency Capability. China Safety Science Journal, 16(2), p. 69.
[4] Ruan Jianan, Wang Jixian.1995. Multi class multi rater kappa analysis. Chinese Health Statistics, 12(6), p. 20.
[5] Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, p. 37.
[6] Wang Jiezhen, Hang Rui. 1996. Kappa statistics in consistency and reproducibility of the test application. Shandong Medical University Journal, 34(3), p. 209.