

Calculation of Threshold Value by Expert Methods

Armins Laurs¹, Zigurds Markovics²

¹Latvia University of Agriculture, ²Riga Technical University

Abstract – Expert information in the form of decision or evaluation is often used to solve tasks with the help of artificial intelligence theory and methods.

The usage of information has strict requirements for the degree of expert consensus.

Unfortunately, the format of calculation of the degree of consensus does not cover in the practice existing spectrum of the task format, so it is necessary adapt a practical task to a theoretical method format. The paper describes one of the methods and research being evaluated.

Keywords – Degree of consensus, emission calculations, expert methods, herd thresholds, Pearson criterion, ranks, ranking correlation, statistical reliability.

I. INTRODUCTION

Expert methods – expert evaluation, expert voting, expert decision-making methods – have gained wide application especially in areas contacting with the theory of artificial intelligence applications. Sometimes these are the only ones capable of making the initial quantitative information or formulating partial decisions.

However, the expert methods have their limits, being formulated in the form of postulates and must be strictly followed. One of these postulates requires calculation of degree of consensus of experts and use of information provided by experts if and only if the consensus degree is greater than the threshold value.

There are a number of methods for calculating the degree of expert consensus. Unfortunately, in practice there are a lot of tasks, whose format does not comply with theoretical format of calculation. Therefore, it is necessary to adapt the task format to theoretical methods.

This paper deals with one of these adaptation methods.

II. FORMULATION OF THE PROBLEM

Let us look at the three most popular methods of determining the degree of expert consensus. They are:

- a) Spearman's rank correlation method;
- b) Concordance coefficient method;
- c) Information theory method.

All have one feature in common, i.e. they work with the ranking scales, only in a different way.

Spearman's rank correlation method works with two ranked rows – information of two experts. Concordance coefficient can be calculated at once for the entire expert team who has made the ranked rows of objects.

Information theory method also works with a team of experts as a whole and is based on the probability that i -th object occupies the k -th place in a ranked row.

Two methods intended for situations in which objects are compared in pairs may also be attributed to the popular methods. One of them works with information theory and probability that i -th object will be found to be better compared to k -th object.

The second method is substantially the same with the finding of concordance coefficient, only the input information is a comparison table of couples.

As can be seen, all the mentioned methods work if the experts have evaluated or ranked many objects or processes.

But in practice it is not only a few tasks, experts should evaluate a single object, set one limit, make a single decision, for example, as the output in the form of production rule. Then these methods cannot be applied directly, adaptation methods have to be developed.

III. THE CONTENT OF EXPERT DEGREE OF CONSENSUS

Let us assume that the set of current objects should be ranked:

$$X = x_1, x_2, \dots, x_i, \dots, x_n, \quad i = 1, n.$$

The task is managed by experts $e_j, j = 1, m, m = 2$. Then we obtain two ranked rows with elements r_{ji} .

The similarity between these rows is expressed in Spearman's rank correlation coefficient [1], [2]:

$$\rho = 1 - \frac{6}{n^3 - n} \sum_{i=1}^n (r_{1i} - r_{2i})^2 \quad (1)$$

ρ values are located in the range of +1 to -1. The value +1 is in the case, if the two rows are identical, the value -1 is in the case, if the rankings are inverse.

If the number of experts $m > 2$, we have to use concordance coefficient [3], [4]:

$$W = \frac{\sum_{i=1}^n \left\{ \sum_{j=1}^m r_{ji} - \frac{1}{2} m (n+1) \right\}^2}{\frac{1}{12} m^2 (n^3 - n)}, \quad (2)$$

or in case of corresponding ranks:

$$W = \frac{\sum_{i=1}^n \left\{ \sum_{j=1}^m r_{ji} - \frac{1}{2} m (n+1) \right\}^2}{\frac{1}{12} m^2 (n^3 - n) - m \sum_j T_j}, \quad (3)$$

$$\text{where } T_j = \frac{1}{12} \sum_{ij} (t_j^3 - t_j),$$

t_j – recurrence number of ranks in case of ranking of j -th expert.

The range of W value is $[0; 1]$.

The expert degree of consensus in case of comparison of pairs is found by the formula [3]:

$$V = \frac{4 \left(\sum \gamma_{if}^2 - m \sum \gamma_{if} + C_m^2 C_n^2 \right)}{m(m-1)n(n-1)} \quad (4)$$

where γ_{if} and γ_{fi} – numbers from input matrix boxes that show how many experts prefer a given object when x_i is compared with x_f ,

C_m^2 , C_n^2 – the number of combinations of m per two and of n per two.

The theory states that values W and V also require calculating the statistical reliability after Pearson criterion χ^2 (chi-square) with probability values 0.95; 0.99; 0.999.

In practice, threshold limits are accepted for values of W and V :

- for technical systems not less than 0.5;
- for medical and human life related systems (e.g., aviation) not less than 0.75.

In the application of information theory, the degree of consensus is [5]

$$W_{ij} = 1 - \frac{H}{H_{\max}}, \quad (5)$$

where $H_{\max} = n \log_2 n$ – maximum entropy;

$$H = - \sum_{k=1}^n \sum_{i=1}^n P_{ik} \log_2 P_{ik} \text{ – entropy;}$$

$$P_{ik} = \frac{m_{ik}}{m} \text{ – probability that the } i\text{-th object will take } k\text{-th}$$

place in ranked row;

m_{ik} – number of experts who have given the k -th place to i -th object.

In case of pair comparison changes, the content of P_{ik} may be expressed by

$$P_{if} = \frac{\gamma_{if}}{m} \text{ – the probability that object } x_i \text{ will be recognised}$$

as the best one in comparison with object x_f .

IV. THE PROBLEMS OF PRACTICAL TASK ADAPTATION

As can be seen from the previous section, all the methods for assessing the degree of expert consensus are intended for many object evaluation tasks. Problems arise when an expert team has to evaluate only one value or has to adopt one decision etc.

Let us consider the situation where the decision has to be formulated in the form of production rule – in the form of IF ..., THEN. In the condition part “IF” the characteristics of the situation are included; in the performance part “THEN” the decision has to be formulated.

Often the decision formulation and adoption become a problem when there is a need to create a knowledge base in appearance of set of production rules. It is known that a set of

production rules in case of two criteria may be presented with the finding of the so-called decision-making table (see Table I).

TABLE I
DECISION-MAKING TABLE

A load danger	Strength of wind				
	No wind (nw)	Temperate wind (tw)	Moderate wind (mw)	Strong wind (sw)	Very strong wind (vsw)
No danger (nd)	T	T	T	T	T
Little danger (ld)	g	g			
Medium danger (md)					tg
Very dangerous (vd)				tg	tg

Table stencil is drawn up for a specific example related to load transportation by crane over the area, which may be partially filled with people, animals, and kiosks.

In table vertical dimension, linguistic gradations of one criterion are located, and these are gradations of danger of load. In table horizontal dimension, linguistic gradations of the other criterion are located that present wind strength [6]. Both criteria together form the condition part “IF” of production rule.

The table boxes must be filled with performance parts “THEN” linguistic gradations which present crane route distance from the people crowd. Gradation may be adopted as one of the following:

- 1) T – straight over people’s heads;
- 2) g – directly along the border of people;
- 3) vg – average far from the border of people;
- 4) tg – far from border of people.

Set of production rules listed in the following table has the advantage that it is visible at the set of all possible situations because each box requires filling.

Part of the table boxes can be filled clearly that can be entrusted to one person. These are the boxes of the table, where in the bottom right spot criteria of maximum large value gradations are met; vd and vsw – very dangerous load and very strong wind. The decision is clear: the route should be chosen with gradation tg – far from the border of people.

Similarly, the problems are not caused by the table in the boxes in the upper left corner, where criteria for minimum values are met: if there is no wind and load danger is zero (no load, it is idle), a route can go straight over people’s heads.

In addition with gradation T all the boxes in the first line can be filled, regardless of the presence of wind.

Unfortunately, the box filling is unclear or ambiguous in the table in the lower left corner, where in upper right corner and

in the centre one criterion of maximum value complies with the second criterion of small values. Here is the place for experts' experiment to fill in the table according to expert majority opinion. There is only one problem how to meet the degree of expert consensus postulate, if the format of task does not comply with the theoretical method format – the task does not have ranked rows.

One of the task types of adaptation can be as follows.

For each possible situations represented by criteria gradation couple, for filling in the appropriate boxes an expert experiment is performed. Experts' task is to select one of the row gradations to be entered in the box.

Given row is T, g, vg, tg .

Ranked row extraction can have several variations.

- 1) If the expert chooses gradation from the given row of one or the other end, then this gradation is assigned by the first rank (1). For the next gradation there will be the second rank (2), for the subsequent – the third rank (3) etc.
- 2) If the expert chooses gradation from the middle of the given row, the ranking may include the following:
 - selected gradation obtains rank 1;
 - neighbour gradations on the left and right of the selected gradation are actually with equal rank, but theoretically they take 2 and 3 in the ranking, as a result, the reduced rank is calculated for each:
$$r_{red} = \frac{2+3}{2} = 2.5;$$
 - further gradations on the left and the right takes 4 and 5 in the ranking, as a result, the reduced rank is 3.5, etc.
- 3) If the expert chooses gradation from the middle of given row, a different version is also possible, when the closest neighbours to the selected gradation receive ranks of 2, 3, etc. only towards one side, for example, to the right from the first rank. Gradations on the left will be ranked in the same row at the far end.

With this adaptation process, we can get all the expert ranked rows for each table box. There is a possibility to determine the degree of consensus of experts with concordance or information theory method.

If the consensus degree is greater than the threshold, statistically valid with a probability of at least 0.95, any methods can be used to express the majority view of experts.

In the example with no numeric values of the linguistic variable, it may be one of the voting procedures. For numeric values, it is possible to use the calculation of the mathematical expectation of numerically weak sets – Voronin method [7].

V. STUDIES ON DISTRIBUTION OF THE NUMBER OF DAIRY CATTLE IN LATVIA

The study was carried out in the project “Agricultural Sector EGG (emissions of greenhouse gases) Emission Calculation Methodology and Data Analysis Developing the Tools of Modelling, Integrating Climate Change”, subproject No. 1.

One of the tasks of subproject No. 1 was to determine the number of Latvian dairy cattle distribution by groups according to their housing technology and find the average quantitative

indicators of those groups, which further could be used in the calculation of greenhouse gas emissions.

Practice has proved that in small farms, dairy cows are kept tied, in medium-sized farms – either tied or loose, but in large farms – only loose. In addition, small and medium-sized farms put cows out to graze in the summer. Each holding technology leaves a distinct impression on greenhouse gas emissions.

Conducting the present research, it was not possible to inspect the entire general sample – all farms engaged in the Latvian dairy farming. This required a very large workforce and financial resources. Also from farms representative and sufficient sample volume terms cannot be distinguished to be processed with mathematical statistical methods, because Latvian dairy cattle farms are located very unevenly.

If the main criteria of distribution of the cattle groups are taken by animal husbandry technologies, there are questions regarding the size of the farm, at the transition from one form to another. Similarly, the question arises on the farm threshold, to which the animals are put out to graze, as well as on the duration of grazing period.

Unfortunately, these values could not be obtained by means of calculation; comprehensive statistics was too expensive.

It was decided to use an expert evaluation method.

VI. RESEARCH METHODOLOGY

The research defined the following tasks:

- to identify the average threshold of herd at which it is kept loose;
- to identify an average value of the herd at which the dairy cows are put out to graze;
- to identify the grazing period, the average duration.

The research expert team consisted of cattle breeding consultants of the Latvian Agricultural Advisory Centre (LAAC) from 16 LAAC branches, as well as of specialists from the Latvian Association of Beef Cattle Breeders.

The acquisition method used to obtain the experts' point of view was the survey.

The task of expert team was to find the range of interval taking into account that the expert team thoughts were a fair value, and to calculate the average value of the interval mathematical expectation.

Each expert discussed 3 rows within a value range of the questions posed:

What is the average limit (herd size), which includes the transition from cow stanchion to not tethered?

Number of dairy cows in the herd				
50–60	61–70	71–80	81–90	91–100

What is the maximum size of the herd, which is further placed on grassland (if any)?

Number of dairy cows in the herd				
<50	51–80	81–100	101–120	121–150

What is the average duration of the grazing period (number of days)?

Grazing period (number of days)				
50–60	61–70	71–80	81–90	91–100

As the expert evaluation method was used, there should be the expert method postulate to determine the degree of consensus with the threshold size compliance.

Unfortunately, the task format with one specific determination of the amount does not correspond to the degree of consensus in the calculation methodology formats that works with ranked rows. This means that any of the above specified task adaptation practices must be used to obtain the necessary ranked rows.

Consequently, the experts had a task to rank specific interval rows.

The following is the calculation of the degree of expert consensus with concordance coefficient W , the statistical reliability of the probability calculation with Pearson criterion X^2 , data grading (if applicable) and cattle group value calculation by determining the mathematical expectation.

VII. RESEARCH RESULTS

Results of the research are presented in tables, which show expert ranking, including the reduced ranks, concordance coefficient calculations and statistical reliability of the calculations.

TABLE II

THE AVERAGE HERD SIZE LIMIT AT THE BEGINNING LOOSE IN HOLDING (DAIRY COWS)

Expert	Object				
	50–60	61–70	71–80	81–90	91–100
1.	3.5	2	1	3.5	5
2.	3.5	2	1	3.5	5
3.	4.5	3	1	2	4.5
4.	4.5	3	1	2	4.5
5.	4.5	3	1	2	4.5
6.	4.5	3	1	2	4.5
7.	5	4	3	1	2
8.	5	4	3	1	2
9.	5	4	3	1	2
10.	5	3.5	2	1	3.5
11.	5	3.5	2	1	3.5
12.	5	4	3	1	2
13.	5	4	3	1	2
14.	5	4	3	1	2
R_i	65.0	47.0	28.0	23.0	47.0
$rv\sum$	42.0				
Δ	23.0	5.0	–14.0	–19.0	5.0
Δ^2	529	25	196	361	25
s	1136				
smax	1960				
skor	4				
smax - skor	1956				
W	0.58				
χ^2_{apr}	32.48				
χ^2_{tab}	13.28				
Statistical reliability	>0.99				

TABLE III

THE AVERAGE SIZE OF HERD PLACED ON GRASSLAND (DAIRY COWS)

Expert	Object				
	<50	51–80	81–100	101–120	121–150
1.	5	1	2	3	4
2.	5	1	2	3	4
3.	5	1	2	3	4
4.	5	1	2	3	4
5.	5	1	2	3	4
6.	5	1.5	1	1.5	4
7.	5	1.5	1	1.5	4
8.	5	1.5	1	1.5	4
9.	5	4	2	1	3
10.	5	4	2	1	3
11.	5	4	2	1	3
12.	5	4	2	1	3
13.	5	4	2	1	3
14.	5	4	2	1	3
R_i	70.0	33.5	25.0	25.5	50.0
$rv\sum$	42				
Δ	28.0	–8.5	–17.0	–16.5	8.0
Δ^2	784.00	72.25	289.00	272.25	64.00
s	1481.5				
smax	1960				
skor	1.5				
smax - skor	1958.5				
W	0.76				
χ^2_{apr}	42.56				
χ^2_{tab}	13.28				
Statistical reliability	>0.99				

TABLE IV

THE DURATION OF THE GRAZING PERIOD (DAIRY COWS)

Expert	Object				
	145–150	151–160	161–170	171–180	181–185
1.	4	1	2	3	5
2.	4	1	2	3	5
3.	4	1	2	3	5
4.	4	1	2	3	5
5.	4	1	2	3	5
6.	4	1	2	3	5
7.	4	1	2	3	5
8.	4	1	2	3	5
9.	4	1	2	3	5
10.	4.5	1.5	1	1.5	4.5
11.	5	3	2	1	4
12.	5	3	2	1	4
13.	5	3	2	1	4
14.	5	3	2	1	4
15.	5	3	2	1	4
16.	5	3	2	1	4
R_i	70.5	28.5	31.0	34.5	73.5
$rv\sum$	48				
Δ	22.5	–19.5	–17.0	–13.5	–25.5
Δ^2	506.25	380.25	289.00	182.25	650.25
s	2008				
smax	2560				
skor	1				
smax - skor	2559				
W	0.78				
χ^2_{apr}	49.92				
χ^2_{tab}	13.28				
Statistical reliability	>0.99				

As can be seen from the tables, in all 3 experiments W values are higher than the threshold value of 0.5, namely – 0.58, 0.76, 0.78. The statistical probability of significance is 0.99. It gives the right to process the data and continue to make the necessary calculations.

To find a threshold, the iterative mathematical expectation estimation method of numerically weak cases, known as Voronin method, was applied [7]:

$$y_{ik} = \frac{\sum_{j=1}^m y_{ji} \exp \left[-\frac{(y_{ik-1} - y_{ji})^2 (m-1)}{2 \sum_{j=1}^m (y_{ik-1} - y_{ji})^2} \right]}{\sum_{j=1}^m \exp \left[-\frac{(y_{ik-1} - y_{ji})^2 (m-1)}{2 \sum_{j=1}^m (y_{ik-1} - y_{ji})^2} \right]}, \quad (6)$$

where y_{ik} – mathematical expectation of the k -th step;

y_{ik-1} – expectation in the previous step $k-1$;

y_{ik-1} – the first step filled by the simple arithmetic average;

y_{ji} – the i -th object rating by j -th expert opinion;

m – number of experts.

As in the above tables, there are not quantifiable values, but the size of the intervals, we have to adapt the task to formal method.

Values y_{ij} were obtained taking averages of these intervals evaluated by experts with the first ranking. For example, a row of numbers, using the Voronin method on behalf of the herd size, placed on the pasture, will be as follows:

60, 60, 60, 60, 60, 90, 90, 90, 110, 110, 110, 110, 110, 110, which is obtained by replacing the existing table first ranks with appropriate interval average values.

As the result each of tables is assessed by one number which is the mathematical expectation of searched threshold value. It is used as the average threshold for distinguishes a group of cattle or expresses the average number of days of grazing.

We can say that each table has obtained a new calculation interval, which does not coincide with the initial decision intervals. Its average size is finding of the mathematical expectation, interval width is found adding minimum and maximum of primary interval width.

For example, on behalf of the tethering and loose in a herd size of the new interval may be 85 ± 5 or $80-90$ animals. The average value of mathematical expectation is 85, which is also the further calculation number.

The average size of the herd to which still animals can be put out to graze is in the interval of 90 ± 9 , $81-99$ cows. Further calculations of threshold have to be adopted with 90 cows.

The average length of grazing period is in the interval of 165 ± 5 or $160-170$ days. In the following calculations, 165 days have to be taken as the threshold.

VIII. CONCLUSION

- 1) Expert evaluation tasks, whose format does not comply with theoretical methods, can be adapted so that there is an appropriate format.
- 2) Expert consensus factor or concordance can also be determined in adapted tasks, requiring from an expert to name only one number, not to rank an object row.
- 3) Concordance can also be determined from tasks with linguistic variable selection.
- 4) To find mathematical expectation, tasks can be adapted working with the value intervals, rather than the object values.

REFERENCES

- [1] N. S. Djakova and G. K. Krug, *Primenenije rangovoj korelaciji dlja obrabotki kachestvennoj informacii*, vol. 67. Trudi MEI, 1966.
- [2] Z. Markovičs, *Ekspertu novērtējumu metodes*. Riga, Latvija, 2009, 111 lpp.
- [3] M. G. Kendall, *Rank correlation methods*. Griffin, London, 1970, p. 211.
- [4] M. G. Kendall and J. D. Gobbons, *Correlation methods*, 5th ed. Amat, London, UK, 1990, pp. 114–121.
- [5] V. U. Nikolajev and V. N. Temnov, “Ob odnom metoda formirovaniya ekspertnoj ocenki,” *Tekhnicheskaja kibernetika*, vol. 5. 1973, pp. 32–36.
- [6] K. Mežale, A. Kundziņš and Z. Markovičs, “Aspects of Foundation of Knowledge Base in Decision-Making Tasks for the Needs of Intellectual Robots,” *Scientific J. of Riga Technical University, Technologies of Computer Control*, vol. 16. Riga, Latvia, 2015.
- [7] A. N. Voronin, “Metod obrabotki massila grannit ekspertnih ocenok,” in *Ergaticheskie cictemi upravlenija*. Kijev, Naukova dushka, 1974, pp. 97–103.
- [8] *Guidelines for National Greenhouse Gas Inventories: Emissions from Livestock and Manure Manageme*, IPCC: Chapter 10, 2006, pp. 87.
- [9] Autoru kolektīvs (J. Priekūla red.), “Moderna piena ražošanas ferma: tehnoloģija, tehnika, apsaimniekošana”. Jelgava, LLU, 2012, 240 lpp.
- [10] *Īpašās prasības piesārņojošo darbību veikšanai dzīvnieku novietnēs*, Ministru kabineta noteikumi Nr. 829 (no 2014.23.12), Rīga, MK.
- [11] J. Priekulis, A. Aboltins, A. Laurs and L. Melace, “Research in manure management in Latvia,” in *Proc. 14th International Scientific Conference „Engineering for Rural Development”*, vol. 14. Jelgava, Latvia University of Agriculture Faculty of Engineering, 2015, pp. 88–93.
- [12] J. Priekulis, A. Sudars and L. Berzina, “Dairy farming and greenhouse gas emissions in Latvia,” in *Proc. 14th International Scientific Conference „Engineering for Rural Development”*, vol. 14. Jelgava, Latvia University of Agriculture, Faculty of Engineering, 2015, pp. 577–583.
- [13] B. Amon and V. Krykoruchko, “Methane and ammonia emissions during storage and after application of dairy cattle slurry and influence of slurry treatment,” in *Agriculture, Ecosystem and Environment*, vol. 112. 2006, pp. 153–162. <http://dx.doi.org/10.1016/j.agee.2005.08.030>
- [14] J. R. Knapp, G. L. Laur and P. A. Vadas, “Enteric methane in dairy cattle production,” in *Journal of Dairy Science*, vol. 97. 2014, pp. 3231–3261. <http://dx.doi.org/10.3168/jds.2013-7234>
- [15] P. Gerber and T. Velling, “Productivity gains and greenhouse gas emission intensity in dairy systems,” in *Livestock Science*, vol. 139. 2011, pp. 100–109. <http://dx.doi.org/10.1016/j.livsci.2011.03.012>

Armins Laurs, Dr. sc. ing., head researcher of Latvia University of Agriculture, Faculty of Engineering, Institute of Agriculture Engineering. He has 137 scientific publications. Research interests in the live-stock breeding engineering field. He is a member of Nordic Association of Agricultural Scientist and the Latvian Association of Scientists. Address: 5 J. Čakstes bulv. Jelgava, LV-3031, Latvia. Email: armins.laurs@promedia.lv

Zigurds Markovičs, Dr. habil. sc. ing., Professor (1993) of Riga Technical University, the Faculty of Computer Science and Information Technology, the Institute of Computer Control, Automation and Computer Engineering. He has 148 scientific publications. Research interests include: computer control systems, artificial intelligence systems, robotics. He is a member of the Latvian Association of Professors and the Latvian Association of Scientists. Address: 2, Daugavgrīvas Str., Riga, LV-1007, Latvia. E-mail: Zigurds.Markovics@rtu.lv

Armīns Laurs, Zīgurds Markovičs. Slietšņa vērtības noteikšana ar ekspertu novērtējumu metodēm

Ekspertu informāciju lēmumu vai novērtējumu formā bieži lieto, lai risinātu uzdevumus ar mākslīgā intelekta teorijas un metožu palīdzību.

Šīs informācijas lietojumiem ir stingras prasības pēc ekspertu vienprātības pakāpes noteikšanas.

Diemžēl vienprātības pakāpes aprēķinu formāts nepārklāj praksē eksistējošo uzdevumu formātu spektru, tādēļ ir nepieciešama praktisko uzdevumu adaptācija teorētisko metožu formātam. Darbā aplūkota viena no šādām metodēm.

Арминс Лаурс, Зигурдс Маркович. Расчет порогового значения экспертными методами

Экспертная информация в форме парциальных решений или оценок качества объектов часто применяется в задачах искусственного интеллекта.

Применение экспертной информации разрешено лишь при соблюдении строгих правил, таких, как существование определённого уровня согласованности мнений экспертов.

К сожалению, формат определения согласованности мнений не охватывает весь спектр форматов практических задач. В этом случае появляется необходимость подстраивать их под метод теоретических расчётов. В работе рассмотрен один из таких методов.