Neryškioji dichotominių testo klausimų ir socialinių rodiklių diferencijavimo savybių klasifikacija

Aleksandras KRYLOVAS, Natalja KOSAREVA, Julija KARALIŪNAITĖ

Main idea:
To construct an estimate of a test item solvability

We propose to separate all test takers into three fuzzy subsets:

W – weak students;
A – average students;
S – strong students.

For each subset we assess the solvability by a triangular fuzzy number:

$$Tr_W, Tr_A, Tr_S$$

These numbers are compared to each other, and on that basis classification of test items is performed.

1. Order relations of the fuzzy triangular numbers

If S is a crisp set, define the solvability of the test item $k(S)$ of the set S of all testees as the percent of correctly answered students' number $t(S)$:

$$k(S) = \frac{t(S)}{|S|} \cdot 100(\%), \qquad (1)$$

where $|S|$ is a number of elements of the set S.

If S is a fuzzy set, we will define the solvability of the test item by the fuzzy triangular number $Tr(L, T, R)$ $(0 \leq L \leq T \leq R \leq 100)$, which in the specific case coincides with (1): $k(S) = L = T = R$.

In general $Tr(L, T, R)$ is a fuzzy number of the set $x \in [0, 100]$ which has the triangular membership function (Zadeh 1965)

$$\mu(x) = \begin{cases} \dfrac{x - L}{T - L}, & \text{for } x \in [L, T], \\ \dfrac{R - x}{R - T}, & \text{for } x \in (T, R], \\ \quad 0, & \text{otherwise.} \end{cases}$$

It is clear that $0 \leq \mu(x) \leq 1$.

Suppose that Tr_1 (L_1, T_1, R_1) and Tr_2 (L_2, T_2, R_2) are two triangular fuzzy numbers. We denote that

$$Tr_1 \preccurlyeq Tr_2, \text{ when } L_1 \leq L_2 \ \& \ T_1 \leq T_2 \ \& \ R_1 \leq R_2. \tag{2}$$

$$Tr_1 \prec Tr_2, \text{ when } R_1 < L_2. \tag{3}$$

We suggest such classifications of the test item differentiation property:

The test item well differentiates **all** students, when:	$Tr_W \prec Tr_A \prec Tr_S$;
The test item well differentiates **strong** students, when:	$Tr_W \preccurlyeq Tr_A \prec Tr_S$;
The test item well differentiates **weak** students, when:	$Tr_W \prec Tr_A \preccurlyeq Tr_S$;
The test item **badly** differentiates students, when:	$Tr_W \preccurlyeq Tr_A \preccurlyeq Tr_S$;
The test item **inappropriate** is in all other cases.	

2. Algorithm of making the triangular fuzzy numbers

Suppose, that $A = \{a \in S: 0 \leq \mu_A (a) \leq 1\}$ is a fuzzy set (the fuzzy subset of the testees (universal) set S), $\mu_A (a)$ – its membership function.

The α – cuts (α level sets) of the set A are called (crisp) sets $A_\alpha = \{a \in S: \mu_A (a) \geq \alpha\}$.

We will take the nonempty cuts A_{α_1}, A_{α_2}, ..., A_{α_n} of the set A. For each cut A_{α_i} we calculate the values

$$p_{\alpha_i} = \frac{t_{\alpha_i}}{|A_{\alpha_i}|} \cdot 100(\%),$$

where t_{α_i} – number of the testees from the set A_{α_i} who correctly answered the test item, $|A_{\alpha_i}|$ – number of elements of the set A_{α_i} (we apply (1) formula). We get number T of the triangular fuzzy number $Tr(L, T, R)$ when $\alpha_i = 1$: $T = p_{1.0}$.
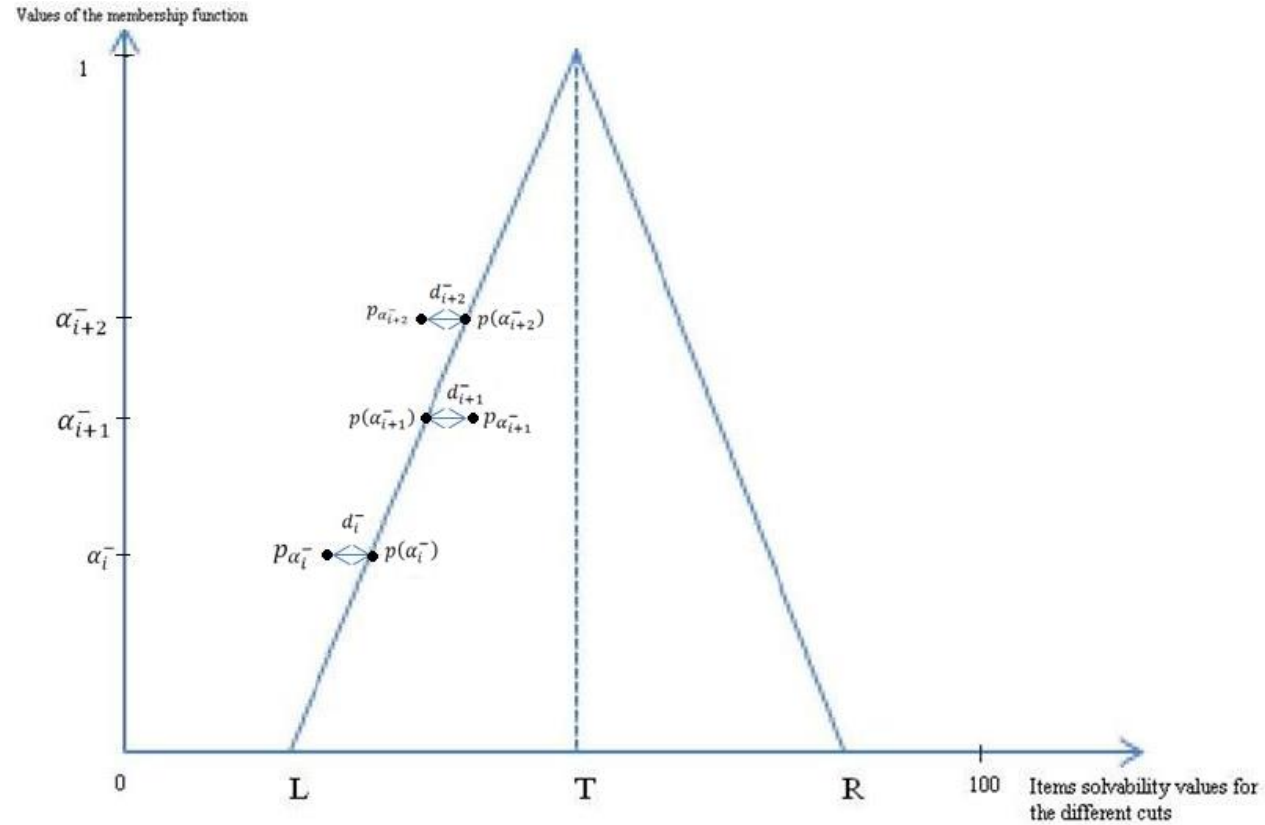
Fig. 1. The values L are obtained by the least squares method minimizing by L the sum $\sum_{i=1}^{n^-}\left(d_i^-\right)^2$, where $d_i^- = p_{\alpha_i^-} - p(\alpha_i^-)$ and $p(\alpha_i^-) = (T-L)\alpha_i^- + L$. Similarly are obtained the values R.

The values L and R are obtained by the least squares method. Denote $\alpha_1^-,\alpha_2^-,...,\alpha_{n^-}^-$ that $\alpha_i < 1.0$ values for which $p_{\alpha_i^-} < T$, and correspondingly denote $\alpha_1^+,\alpha_2^+,...,\alpha_{n^+}^+$, when $p_{\alpha_i^+} > T$; $n^- + n^+ = n - 1$. Define the objective functions for the parameters L and R:

$$f(L) = \sum_{i=1}^{n^-}\left(p_{\alpha_i^-} - (T-L)\alpha_i^- - L\right)^2, \quad g(R) = \sum_{i=1}^{n^+}\left(p_{\alpha_i^+} - R + (R-T)\alpha_i^+\right)^2.$$

The meaning of the functions $f(L)$, $g(R)$ is sum of the squares of the values $d_i^{\pm}$ shown in the Fig 1. Further we solve the equations $f'(L) = 0$ and $g'(R) = 0$:

$$L = \frac{\displaystyle\sum_{i=1}^{n^-}\left(p_{\alpha_i^-} - T\alpha_i^-\right)\cdot(1-\alpha_i^-)}{\displaystyle\sum_{i=1}^{n^-}\left(1-\alpha_i^-\right)^2}, \quad R = \frac{\displaystyle\sum_{i=1}^{n^+}\left(p_{\alpha_i^+} - T\alpha_i^+\right)\cdot(1-\alpha_i^+)}{\displaystyle\sum_{i=1}^{n^+}\left(1-\alpha_i^+\right)^2}.$$

3. Finding fuzzy subsets of the testees set

The fuzzy subsets W, A, S of the testees set we define by the trapezoidal membership functions

$$\mu_{(a,b,c,d)}(t) = \begin{cases} \dfrac{t-a}{b-a}, & \text{when } t \in [a,b], \\ \dfrac{d-t}{d-c}, & \text{when } t \in [c,d], \\ 1, & \text{when } t \in (b,c), \\ 0, & \text{otherwise,} \end{cases}$$

where $a \leq b \leq c \leq d$. Let us suppose that knowledge (or other considered property) of all tested students is valued by some scores b_1, b_2, …, b_n. Denote the least and the greatest b_i values as *min* and *max*.

Let us take four numbers

$$min < \alpha < \beta < \gamma < \delta < max$$

and let us define the membership functions of the subsets W, A, S as:

$$\mu^W_{(min,min,\alpha,t_1)}(t), \mu^A_{(t_2,\beta,\gamma,t_3)}(t), \mu^S_{t_4,\delta,max,max)}(t).$$

Parameters t_i get such values

$$\alpha \leq t_1 \leq t_2 \leq \beta, \qquad \gamma \leq t_3 \leq t_4 \leq \delta,$$

and the meaning of the parameters α, β, γ, δ is the following: when $b_i \leq \alpha$ then the student for sure is weak, when $\beta \leq b_i \leq \gamma$ – average, when $b_i \geq \delta$ – strong.
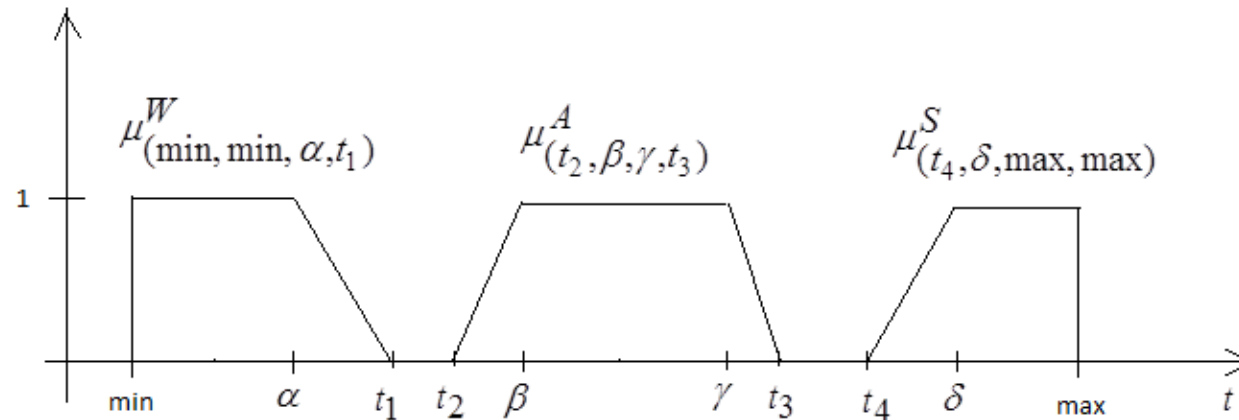

Fig. 2. Membership functions of the weak, average and strong students ($\mu^W_{(min,min,\alpha,t_1)}(t), \mu^A_{(t_2,\beta,\gamma,t_3)}(t), \mu^S_{t_4,\delta,max,max)}(t)$.).

For each subset *W*, *A*, *S* we construct the triangles $Tr(L, T, R)$ and analyze all parameters t_i, which follow restrictions. From the obtained fuzzy triangular numbers we will make optimistic triangles (we take averages of *L*, *T*, *R*, when all t_i values are reselected) and pessimistic triangles (wider intervals, when we take *minL*, *maxR* and average of *T* values, when all t_i values are reselected). Thus while classifying test items of the test we pay attention not only to the average (optimistic), but also to the pessimistic fuzzy assessments.

4. Analysis of the test items of one particular test

In this section we will show some test items from the mathematics test. The test was given to 106 students from Vilnius Gediminas Technical University faculty of Civil Engineering. Their knowledge was evaluated by the results of the test made of 20 test items, each valued by one score. We present as an example of obtained results by the Table 1.

Table 1. A part of the obtained results

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	+	+	+	−	+	−	−	−	−	−	−	−	+	+	−	−	−	−	−	−	6
2	+	+	+	−	−	+	+	+	+	−	+	−	+	+	+	−	−	−	−	+	12
3	+	+	−	+	+	+	+	+	+	−	+	+	+	+	−	−	−	−	−	−	12
4	+	+	+	+	+	+	+	+	−	+	−	+	+	−	−	−	−	+	−	+	13

In the first column there are students' numbers, in the first row – test items' numbers. The considered test items are shaded. In the last column there are total numbers of correctly answered test items of each student – scores b_i. Here we present only a part of one table.

There were 4 groups with 29, 26, 28 and 23 students correspondingly in each group.

Parameters were:

$min = 1$, $max = 20$, $\alpha = 9$, $\beta = 12$, $\gamma = 14$, $\delta = 17$.

24 - weak students $(1 \leq b_i \leq 9)$;

30 - average $(12 \leq b_i \leq 14)$;

19 - strong $(17 \leq b_i \leq 20)$.

For the rest of the students we cannot unambiguously state that they are strong, average or weak. We have $33 = 106 - (24+30+19)$ such students. We varied with assigning them to one of the three subsets, changing parameters t_i, constructing trapezoids.

Next, we present the test items from the considered test:

6	$\lim\limits_{x \to 0} x^{14} \ln x =$	1) 14;	2) 0;
		3) ∞;	4) $\ln 14$;
		5) limit does not exist;	6) 1.

The solvability of the test item is 74(%) (that is 74% of all students gave correct answer to that test item). The fuzzy triangular numbers of the groups and for all students are shown in the Figs. 4–8.

Notice that only for one (2nd) group test item 6 badly differentiates students in the pessimistic case (relations $Tr_{r_W} \lesssim Tr_{r_A} \lesssim Tr_{r_S}$). For all other groups separately and for all groups together the test item well differentiates weak students both in optimistic and pessimistic cases (relations $Tr_{r_W} \prec Tr_{r_A} \lesssim Tr_{r_S}$).
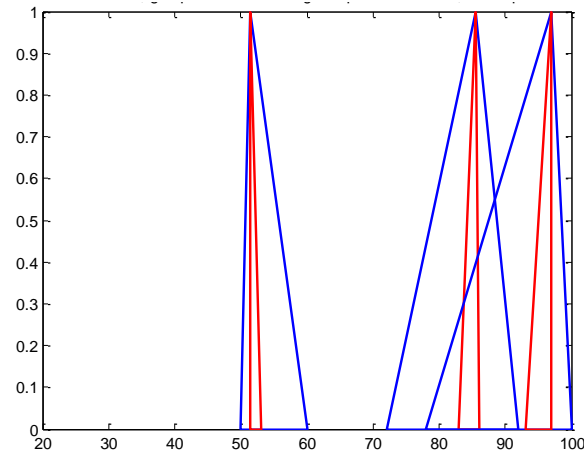

Fig. 4. Fuzzy triangles – test item 6, group 1 (blue triangles – pessimistic, red – optimistic cases).
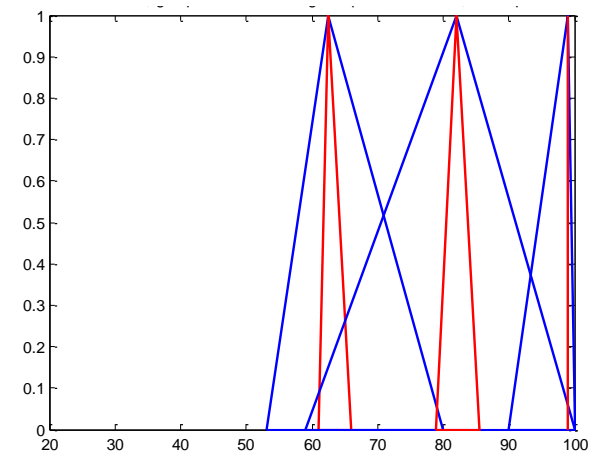
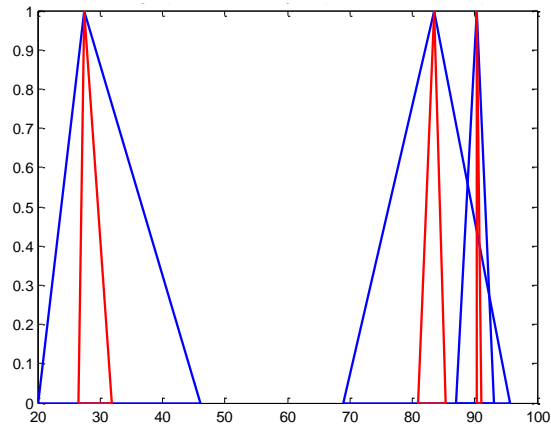Fig. 5. Fuzzy triangles – test item 6, group 2 (blue triangles – pessimistic, red – optimistic cases).

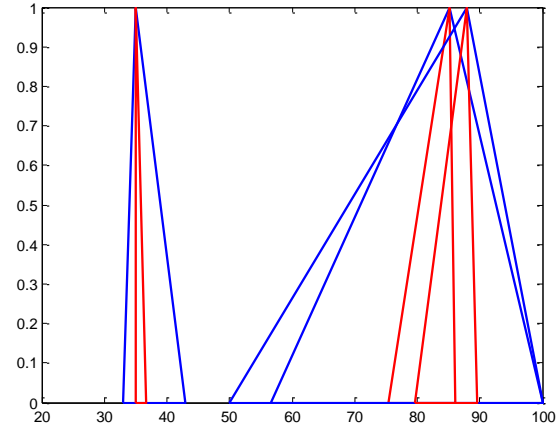Fig. 6. Fuzzy triangles – test item 6, group 3 (blue triangles – pessimistic, red – optimistic cases).

Fig. 7. Fuzzy triangles – test item 6, group 4 (blue triangles – pessimistic, red – optimistic cases).
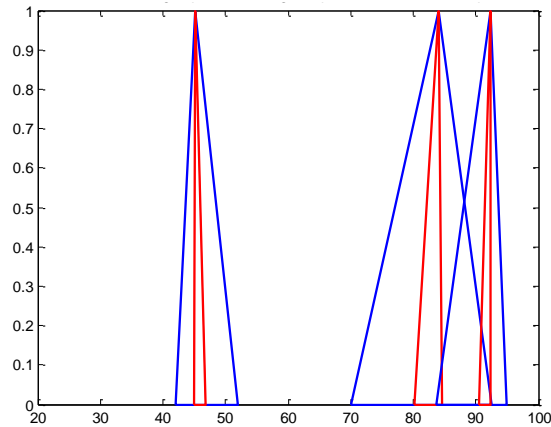

Fig. 8. Fuzzy triangles – test item 6, all groups (blue triangles – pessimistic, red – optimistic cases).

5. Conclusion and future research

- Method of establishing the test's items differentiation property was proposed.

- Considered method does not require the strict evaluation of the testees' knowledge.

- To apply it is enough to have only relative achievement scores, for example raw test scores.

- The same method can be applied both for the relatively small groups of testees and for combined groups.

- An experiment allows to expect to get a stable test items classification having not big (20–30 students) groups of testees, however to get statistically reliable conclusions it is appropriate to carry out Monte Carlo type experiments. This is the object of our future research.

- Also pay attention that in this survey all test items are considered independently from another, though the test items blocks essentially make one task and it would be right to study them together (Krylovas, Kosareva 2011).

AČIŪ UŽ DĖMESĮ!

For example, if $A = \{(1, 0.5), (2, 0.6), (3, 0.7), (4, 1.0),$
$(5, 0.8)\}$, then $A_{0.5} = S = \{1, 2, 3, 4, 5\}$, $A_{0.6} = \{2, 3, 4, 5\}$, $A_{0.7} = \{3, 4, 5\}$, $A_{0.8} = \{4, 5\}$, $A_{1.0} = \{4\}$,

Example. Suppose that the fuzzy set of testees is:
$A = \{(a,1.0),(b,0.5),(c,1.0),(d,0.3),(e,0.1),(f,0.3)\}$. Suppose that only a and b have answered the test item correctly. Then the set A has four different α-cuts. The corresponding p_α values from formula (6) are:

$$A_{1.0} = \{a,c\}, \ p_{1.0} = \frac{1}{2} \cdot 100 = 50;$$

$$A_{0.5} = \{a,b,c\}, \ p_{0.5} = \frac{2}{3} \cdot 100 = 67;$$

$$A_{0.3} = \{a,b,c,d,f\}, \ p_{0.3} = \frac{2}{5} \cdot 100 = 40;$$

$$A_{0.1} = \{a,b,c,d,e,f\}, \ p_{0.1} = \frac{2}{6} \cdot 100 = 33.$$

Thus $n^- = 2$, $\alpha_1^- = 0.1$, $p_1^- = 33$, $\alpha_2^- = 0.3$, $p_2^- = 40$, $n^+ = 1$, $\alpha_1^+ = 0.5$, $p_1^+ = 67$. We find th parameters of the fuzzy triangular number $Tr(L,T,R)$:

$T = 50$,

$$L = \frac{(33 - 50 \cdot 0.1) \cdot (1.0 - 0.1) + (40 - 50 \cdot 0.3) \cdot (1.0 - 0.3)}{(1.0 - 0.1)^2 + (1.0 - 0.3)^2} = 32.85,$$

$$R = \frac{(67 - 50 \cdot 0.5) \cdot (1.0 - 0.5)}{(1.0 - 0.5)^2} = 84,$$

Finally, we obtain triangular fuzzy number describing this test item: $(32.85, 50, 84)$.

Example. Suppose that we obtained total scores of 20 items test for 5 testees. Furthemore, we know that 3 of them responded correctly to the particular test item (+) and 2 answered incorrectly (–): $s_1(2,+), s_2(4,+), s_3(3,-), s_4(3,-), s_5(5,+)^2$. Let us construct optimistic and pessimistic fuzzy triangles describing this test item. From the test results, we see that all testees are probably weak students. The membership function of the subset of weak students W is $\mu^W_{(min,min,\alpha,t_1)}(t)$, where $min = 2$, $\alpha = 3$. By changing α and t_1 values $(3 \le \alpha \le t_1 \le 5)$ we get 6 different trapezoids (fuzzy set's membership functions): $\mu^W_1 = \mu^W_{(2,2,3,3)}(t)$, $\mu^W_2 = \mu^W_{(2,2,3,4)}(t)$, $\mu^W_3 = \mu^W_{(2,2,3,5)}(t)$, $\mu^W_4 = \mu^W_{(2,2,4,4)}(t)$, $\mu^W_5 = \mu^W_{(2,2,4,5)}(t)$, $\mu^W_1 = \mu^W_{(2,2,5,5)}(t)$. For each trapezoid, construct a fuzzy triangle describing the test item.

1. $\alpha = t_1 = 3$. $A_1 = \{(s_1, 1.0), (s_3, 1.0), (s_4, 1.0)\}$. Fuzzy set of weak students is a crisp set in this case, so $L_1 = T_1 = R_1 = \frac{1}{3} \cdot 100 = 33$.

2. $\alpha = 3, t_1 = 4$. Fuzzy set of testees is $A_2 = \{(s_1, 1.0), (s_2, 0.0), (s_3, 1.0), (s_4, 1.0)\}$ $L_2 = T_2 = \frac{1}{3} \cdot 100 = 33$. The set A_2 has 2 different α-cuts and corresponding p_α values $A_{0.0} = \{s_1, s_2, s_3, s_4\}$ and $A_{1.0} = \{s_1, s_3, s_4\}$. $\alpha_1^+ = 0.0$, $p_{0.0}^+ = \frac{2}{4} \cdot 100 = 50$.

$$R_2 = \frac{(50 - 33 \cdot 0) \cdot (1 - 0)}{(1 - 0)^2} = 50.$$

3. $\alpha = 3, t_1 = 5$. $A_3 = \{(s_1, 1.0), (s_2, 0.5), (s_3, 1.0), (s_4, 1.0), (s_5, 0.0)\}$. $L_3 = T_3 = 33$. The set A_3 has 3 α-cuts and corresponding p_α values $A_{0.0} = \{s_1, s_2, s_3, s_4, s_5\}$, $A_{0.5} = \{s_1, s_2, s_3, s_4\}$ and $A_{1.0} = \{s_1, s_3, s_4\}$. $\alpha_1^+ = 0.0$, $p_{0.0}^+ = 60$, $\alpha_2^+ = 0.5$, $p_{0.5}^+ = 50$.

$$R_3 = \frac{(50 - 33 \cdot 0.5) \cdot (1 - 0.5) + (60 - 33 \cdot 0.0) \cdot (1 - 0.0)}{(1 - 0.5)^2 + (1 - 0.0)^2} = 61.4.$$

4. $\alpha = 4, t_1 = 4$. A_4 is also a crisp set, $A_4 = \{(s_1, 1.0), (s_2, 1.0), (s_3, 1.0), (s_4, 1.0)\}$.

$L_4 = T_4 = R_4 = \dfrac{2}{4} \cdot 100 = 50$.

5. $\alpha = 4, t_1 = 5$. $A_5 = \{(s_1, 1.0), (s_2, 1.0), (s_3, 1.0), (s_4, 1.0), (s_5, 0.0)\}$. The set A_5 has 2 different α-cuts and corresponding p_α values $A_{0.0} = \{s_1, s_2, s_3, s_4, s_5\}$ and

$A_{1.0} = \{s_1, s_2, s_3, s_4\}$. $L_5 = T_5 = \dfrac{2}{4} \cdot 100 = 50$. $\alpha_1^+ = 0.0$, $p_{0.0}^+ = 60$. $R_5 = \dfrac{(60 - 33 \cdot 0)}{(1-0)^2} = 60$.

6. $\alpha = 5, t_1 = 5$. A_6 is a crisp set, $A_6 = \{(s_1, 1.0), (s_2, 1.0), (s_3, 1.0), (s_4, 1.0), (s_5, 1.0)\}$.

$L_6 = T_6 = R_6 = \dfrac{3}{5} \cdot 100 = 60$.

So, we have 6 fuzzy triangles:

$(33, 33, 33)$; $(33, 33, 50)$; $(33, 33, 61.4)$; $(50, 50, 50)$; $(50, 50, 60)$; $(60, 60, 60)$.

Construct optimistic and pessimistic triangles according to formulas (11)–(12):

$(L_{opt}, T_{opt}, R_{opt}) = (43.17, 43.17, 52.4)$, $(L_{pes}, T_{pes}, R_{pes}) = (33, 43.17, 61.4)$.

Example. Suppose that the fuzzy set of testees is: $S = \{(a, 1.0), (b, 0.5), (c, 1.0), (d, 0.3), (e, 0.1), (f, 0.3)\}$. Suppose that only a and b answered correctly to the test item. Then the set S has four different $\alpha -$ cuts and corresponding p_α values:

$$S_{1.0} = \{a, c\}, \quad p_{1.0} = \frac{1}{2} \cdot 100 = 50,$$

$$S_{0.5} = \{a, b, c\}, \quad p_{0.5} = \frac{2}{3} \cdot 100 = 67,$$

$$S_{0.3} = \{a, b, c, d, f\}, \quad p_{0.3} = \frac{2}{5} \cdot 100 = 40,$$

$$S_{0.1} = \{a, b, c, d, e, f\}, \quad p_{0.1} = \frac{2}{6} \cdot 100 = 33.$$

Thus $n^- = 2$, $\alpha_1^- = 0.1$, $p_1^- = 33$, $\alpha_2^- = 0.3$, $p_2^- = 40$; $n^+ = 1$, $\alpha_1^+ = 0.5$, $p_1^+ = 67$. We find the parameters of the fuzzy triangular number $Tr(L, T, R)$:

$T=50$,

$$L = \frac{(33 - 50 \cdot 0.1) \cdot (1.0 - 0.1) + (40 - 50 \cdot 0.3) \cdot (1.0 - 0.3)}{(1.0 - 0.1)^2 + (1.0 - 0.3)^2} = 32.85,$$

$$R = \frac{(67 - 50 \cdot 0.5) \cdot (1.0 - 0.5)}{(0.5)^2} = 84.$$