Lietuvos matematikos rinkinys ISSN 0132-2818
Lietuvos matematikų draugijos darbai, ser. B www.mii.lt/LMR/
58 t., 2017, 39–44

Žinių tikrinimo testų klausimų skiriamoji geba

Aleksandras Krylovas, Natalja Kosareva, Julija Karaliūnaitė

Vilniaus Gedimino technikos universitetas, Fundamentinių mokslų fakultetas
Saulėtekio al. 11, LT-10223 Vilnius
E. paštas: aleksandras.krylovas@vgtu.lt, natalja.kosareva@vgtu.lt
E. paštas: julija.karaliunaite@vgtu.lt

Santrauka. Pasiūlytas dichotominio testo klausimų neryškiosios klasifikacijos algoritmas. Priklausomai nuo testu matuojamo žinių lygio, klausimas gali gerai diferencijuoti visus testuojamosios arba gerai diferencijuoti turinčius stiprias žinias arba silpnas žinias; taip pat klausimas gali prastai diferencijuoti visus testuojamuosius ir būti netinkamas. Tyrimo metodika naudoja neryškiųjų aibių matematinį aparatą.

Raktiniai žodžiai: matematinis modeliavimas, neryškiosios aibės, dichotominiai testai.

Įvadas

Straipsnyje [1] buvo pasiūlyta apibūdinti dichotominių testų klausimus nemažėjančiomis funkcijomis $k(p) : [0,1] \to [0,1]$, kai p yra testuojamųjų žinių (arba kitos testu matuojamos savybės) lygis. Šią prielaidą gana sunku pagrįsti metodologiškai, nes žinios (ir panašios savybės) yra sudėtinga, turinti daugybę aspektų sąvoka. Dažnai ji traktuojama kaip tam tikras įsivaizduojamas konstruktas ir apsiribojama jo matavimo procedūromis. Matuojant žinias daug paprasčiau gauti santykinius testuojamųjų įverčius, kuriuos teikia, pavyzdžiui, neapdoroti testo balai.

Turėdami tokius santykinius žinių vertinimus, konstruojame klausimų išsprendžiamumo (teisingai atsakiusių į testo klausimą procentinė dalis) įverčius. Šiame straipsnyje siūloma išskirti visų testuojamųjų aibės tris neryškiuosius poaibius W, M, S (silpnų, vidutinių ir stiprių studentų) ir kiekvienam iš jų išsprendžiamumą vertinti neryškiuoju skaičiumi (žr., pvz., [3]): I^W, I^M, I^S. Šiuos skaičius (šiame straipsnyje trikampinius) lyginame tarpusavyje ir tokiu pagrindu atliekame klausimų klasifikaciją.

Šis straipsnis organizuotas taip. Pirmajame skyriuje apibrėžti trikampinių neryškiųjų skaičių palyginimo sąryšiai ir atitinkama dichotominių testų klausimų skiriamosios gebos klasifikacija. Antrajame skyriuje aprašytas trikampinių neryškiųjų skaičių konstravimo algoritmas. Trečiajame – trapecinio pavidalo priklausomumo funkcijomis apibrėžiami poaibiai W, M, S. Ketvirtajame – atlikta vieno konkretaus aukštosios matematikos testo klausimų analizė.

1 Neryškiųjų trikampinių skaičių tvarkos sąryšiai

Klausimo išspendžiamumą $k(S')$ testuojamųjų aibei S' apibrėžiame kaip teisingai atsakiusiųjų $t(S')$ į šį klausimą procentinę dalį: $k(S') = \frac{t(S')}{|S'|} \cdot 100(\%)$, čia $|S'|$ aibės

S' elementų skaičius. Jei S' yra neryškioji aibė, klausimo sunkumą apibūdinsime neryškiu trikampiniu skaičiumi $Tr(L,T,R)$ ($0 \leqslant L \leqslant T \leqslant R \leqslant 100$), kuris atskiruoju atveju sutampa su $k(S') = L = T = R$. Bendruoju atveju $Tr(L,T,R)$ yra aibės $x \in [0,100]$ neryškusis poaibis, turintis priklausomumo funkciją

$$\mu(x) = \begin{cases} \frac{x-L}{T-L}, & \text{kai } x \in [L,T], \\ \frac{R-x}{R-T}, & \text{kai } x \in (T,R], \\ 0, & \text{kitais atvejais.} \end{cases}$$

Šio straipsnio 2 skyriuje bus aprašytas neryškių trikampinių skaičių konstravimo algoritmas, t. y. skaičių L, T, R skaičiavimo metodika, kai yra žinoma, kurie iš neryškiosios testuojamųjų aibės teisingai atsakė į testo klausimą.

Tarkime, kad $Tr_1(L_1,T_1,R_1)$ ir $Tr_2(L_2,T_2,R_2)$ yra du trikampiniai skaičiai. Sakome, kad

$$Tr_1 \preccurlyeq Tr_2, \quad \text{kai } L_1 \leqslant L_2 \ \& \ T_1 \leqslant T_2 \ \& \ R_1 \leqslant R_2. \tag{1}$$

Pastebėkime, kad sąryšis $\preccurlyeq$ yra refleksyvusis ($\forall Tr: \ Tr \preccurlyeq Tr$), antisimetrinis (jei $Tr_1 \preccurlyeq Tr_2$ ir $Tr_2 \preccurlyeq Tr_1$, tai $Tr_1 = Tr_2$) ir tranzityvusis (jei $Tr_1 \preccurlyeq Tr_2$ ir $Tr_2 \preccurlyeq Tr_3$, tai $Tr_1 \preccurlyeq Tr_3$). Tai reiškia, kad sąryšis $\preccurlyeq$ yra negriežtosios tvarkos sąrysis (žr.: [2]). Jis nėra pilnasis, nes yra trikampinių skaičių kuriems negalioja nei $Tr_1 \preccurlyeq Tr_2$, nei $Tr_2 \preccurlyeq Tr_1$.

Sakome, kad

$$Tr_1 \prec Tr_2, \quad \text{kai } R_1 < L_2 \text{ arba } R_1 = L_2 \text{ ir } T_1 < T_2. \tag{2}$$

Sąryšis $\prec$ irgi yra tranzityvusis ir antisimetrinis, tačiau yra antirefleksyvusis: $(\forall Tr)\ Tr \not\prec Tr$. Todėl jis yra griežtosios tvarkos sąryšis.

Testuojamųjų aibėje nagrinėsime neryškiuosius poaibius: silpnų studentų W, vidutinių – M ir stiprių – S. Tarkime, kad kiekvienam poaibiui apibrėžtas neryškusis trikampinis klausimo sunkumo skaičius Tr_W, Tr_M, Tr_S. Siūloma tokia klausimo skiriamosios gebos klasifikacija: klausimas gerai diferencijuoja **visus** studentus, kai galioja $Tr_W \prec Tr_M \prec Tr_S$; klausimas gerai diferencijuoja **stiprius** studentus, kai galioja $Tr_W \preccurlyeq Tr_M \prec Tr_S$; klausimas gerai diferencijuoja **silpnus** studentus, kai $Tr_W \prec Tr_M \preccurlyeq Tr_S$; klausimas **prastai** diferencijuoja visus studentus, kai $Tr_W \preccurlyeq Tr_M \preccurlyeq Tr_S$; klausimas yra **netinkamas** visais kitais atvejais.

2 Neryškiųjų trikampinių skaičių konstravimo algoritmas

Tarkime, kad $A = \{a \in S' : \ 0 \leqslant \mu_A(a) \leqslant 1\}$ yra neryškioji aibė (testuojamųjų (universaliosios) aibės S' neryškusis poaibis), $\mu_A(a)$ – jos elementų priklausomumo funkcija. Pareikalaukime, kad A būtų normalioji, t. y. $\sup_{a\in S'} \mu_A(a) = 1$. Aibės A α-pjūviais (α lygmens aibėmis) vadinamos (ryškios) aibės $A_\alpha = \{a \in S' : \mu_A(a) \geqslant \alpha\}$. Žinomas neryškiųjų aibių teorijos faktas (žr., pvz. [3]) yra tokia formulė $A = \bigcup_{\alpha\geqslant 0} \alpha A_\alpha$. Čia $\alpha A_\alpha = \{a \in S' : 0 \leqslant \alpha\mu_A(a) \leqslant 1\}$. Pavyzdžiui, jei $A = \{(1,0.5),(2,0.6),(3,1.0)\}$, tai $A_{0.5} = S = \{1,2,3\}$, $A_{0.6} = \{2,3\}$, $A_{1.0} = \{3\}$, $\bigcup_{\alpha\geqslant 0} \alpha A_\alpha = \{(1,0.5),(2,0.5),(3,0.5)\} \cup \{(2,0.6),(3,0.6)\} \cup \{(3,1.0)\} = \{(1,0.5),(2,\max\{0.5,0.6\}),(3,\max\{0.5,0.6,1.0\})\} = A$.

Imsime aibės A netuščius pjūvius $A_{\alpha_1}, A_{\alpha_2}, \dots A_{\alpha_n}$, kurių yra baigtinis skaičius, nes testuojamųjų aibė S' yra baigtinė. Kiekvienam pjūviui A_{α_i} apskaičiuojame reikšmes $p_{\alpha_i} = \frac{t_{\alpha_i}}{|A_{\alpha_i}|} \cdot 100(\%)$, čia t_{α_i} – teisingai atsakiusių į testo klausimą testuojamųjų iš aibės A_{α_i} skaičius, $|A_{\alpha_i}|$ – aibės A_{α_i} elementų skaičius. Neryškaus trikampio $Tr(L, T, R)$ skaičių T gauname, kai $\alpha_i = 1$: $T = p_{1.0}$. Pastebėkime, kad aibė A yra normalioji, todėl $A_{1.0} \neq \emptyset$.

Reikšmes L ir R konstruojame mažiausių kvadratų metodu. Pažymėkime $\alpha_1^-, \alpha_2^-, \dots, \alpha_{n^-}^-$ tas $\alpha_i < 1.0$ reikšmes, esant kurioms $p_{\alpha_i^-} \leqslant T$ ir atitinkamai pažymėkime $\alpha_1^+, \alpha_2^+, \dots, \alpha_{n^+}^+$, kai $p_{\alpha_i^+} \geqslant T$; $n^- + n^+ = n - 1$. Apibrėžkime tikslo funkcijas parametrams L ir R ieškoti:

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

Sprendžiame lygtis $f'(L) = 0$ ir $g'(R) = 0$:

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

1 pavyzdys. Tarkime, kad testuojamųjų studentų neryškioji aibė yra ši: $A = \{(s_1^+, 1), (s_2^+, 1), (s_3^+, 0,7), (s_4^-, 0.3), (s_5^-, 0.1)\}$; viršutinis indeksas reiškia, kad į testo klausimą buvo atsakyta teisingai (+) arba klaidingai (−). Aibė A turi keturis skirtingus α-pjūvius ir atitinkamas p_α reikšmes:

$$\begin{aligned} A_{1.0} &= \{s_1^+, s_2^+\}, & p_{1.0} &= \frac{2}{2} \cdot 100 = 100,\\ A_{0.7} &= \{s_1^+, s_2^+, s_3^+\}, & p_{0.7} &= \frac{3}{3} \cdot 100 = 100,\\ A_{0.3} &= \{s_1^+, s_2^+, s_3^+, s_4^-\}, & p_{0.3} &= \frac{3}{4} \cdot 100 = 75,\\ A_{0.1} &= \{s_1^+, s_2^+, s_3^+, s_4^-, s_5^-\}, & p_{0.1} &= \frac{3}{5} \cdot 100 = 60. \end{aligned}$$

Taigi $n^- = 3$, $\alpha_1^- = 0.1$, $p_1^- = 60$, $\alpha_2^- = 0.3$, $p_2^- = 75$, $\alpha_3^- = 0.7$, $p_3^- = 100$; $n^+ = 0$. Surandame neryškaus trikampinio skaičiaus $Tr(L, T, R)$ parametrus:

$$\begin{aligned} T &= 100,\\ L &= \frac{(60 - 100 \cdot 0.1) \cdot (1 - 0.1) + (75 - 100 \cdot 0.3) \cdot (1 - 0.3) + (100 - 100 \cdot 0.7) \cdot (1 - 0.7)}{(1 - 0.1)^2 + (1 - 0.3)^2 + (1 - 0.7)^2}\\ &= \frac{50 \cdot 0.9 + 45 \cdot 0.7 + 30 \cdot 0.3}{0.9^2 + 0.7^2 + 0.3^2} = \frac{45 + 31.5 + 9}{0.81 + 0.49 + 0.09} = 61.51,\\ R &= T = 100. \end{aligned}$$

Taigi gauname neryškųjų trikampį skaičių, apibūdinantį šį klausimą: $(61.51, 100, 100)$.

3 Testuojamųjų aibės neryškiųjų poaibių konstravimas

Testuojamųjų aibės neryškiuosius poaibius W, M, S (silpnų, vidutinių ir stiprių studentų) apibrėžiame trapecinėmis priklausomumo funkcijomis

$$\mu_{(a,b,c,d)}(t) = \begin{cases} \frac{t-a}{b-a}, & \text{kai } t \in [a,b], \\ \frac{d-t}{d-c}, & \text{kai } t \in [c,d], \\ 1, & \text{kai } t \in (b,c), \\ 0, & \text{kitais atvejais,} \end{cases}$$

čia $a \leqslant b \leqslant c \leqslant d$. Tarkime, kad visų testuojamųji žinios (arba kita vertinama savybė) įvertintos tam tikrais balais $b_1, b_2, \dots, b_n$. Pažymėkime mažiausią ir didžiausią b_i reikšmes min ir max. Pasirinksime keturis skaičius

$$\min < \alpha < \beta < \gamma < \delta < \max \tag{3}$$

ir poaibių W, M, S trapecines funkcijas apibrėžkime taip:

$$\mu^W_{(\min,\min,\alpha,t_1)}(t), \qquad \mu^M_{(t_2,\beta,\gamma,t_3)}(t), \qquad \mu^S_{(t_4,\delta,\max,\max)}(t). \tag{4}$$

Parametrai t_i įgyja reikšmes

$$\alpha \leqslant t_1 \leqslant t_2 \leqslant \beta \leqslant \gamma \leqslant t_3 \leqslant t_4 \leqslant \delta, \tag{5}$$

o parametrų α, β, γ, δ prasmė yra ši: kai $b_i \leqslant \alpha$ studentas tikrai yra silpnas, kai $\beta \leqslant b_i \leqslant \gamma$ – vidutinis, o kai $b_i \geqslant \delta$ – stiprus.

Kiekvienam poaibiui W, M, S konstruojame trikampius $Tr(L,T,R)$ ir nagrinėjame visus parametrus t_i (jų yra baigtinis skaičius), tenkinančius (5) apribojimus. Iš konstruojamų trikampinių neryškių skaičių $Tr(L_i,T_i,R_i)$ sudarysime optimistinius (imame dydžių L, T, R vidurkius, kai perrenkamos visos t_i reikšmės: $L_{opt} = \frac{1}{N}\sum_{i=1}^N L_i$, $T_{opt} = \frac{1}{N}\sum_{i=1}^N T_i$, $R_{opt} = \frac{1}{N}\sum_{i=1}^N R_i$) ir pesimistinius (platesnius, kai skirtumai $R-L$ gali būti dideli) trikampius, kai imame $L_{pes} = \min_i L_i$, $R_{pes} = \max_i R_i$ ir reikšmių T vidurkį $T_{opt} = \frac{1}{N}\sum_{i=1}^N T_i$.

Taigi klasifikuodami testo klausimus atsižvelgiame ne tik į vidutinius (optimistinius), bet ir į pesimistinius neryškiuosius vertinimus.

4 Testo klausimo skiriamosios gebos nustatymo procedūra

Siūlome tokią vieno dichotominio testo klausimo skiriamosios gebos nustatymo procedūrą.

1. Pradiniai duomenys: Įvedami kiekvieno studento surinkti testo taškai, b) į kiekvieną klausimą studentas atsako teisingai (+) arba klaidingai (−).
2. Nagrinėjamam testo klausimui parenkame parametrų α, β, γ, δ reikšmes, kurios turi tenkinti (3) sąlygas.
3. Parenkame parametro t_1 reikšmę, tenkinančią sąlygą (5).
4. Skaičiuojame silpnų studentų poaibio W trapecines funkcijas $\mu^W_{(\min,]min,\alpha,t_1)}(t)$; skaičiuojame $A_{1.0}$ ir $T = p_{1.0}$.

5. Konstruojame visus skirtingus α − pjūvius atitinkančius p_α reikšmes.
6. Skaičiuojame L ir R reikšmes pagal pateiktas 2-ame skyriuje formules. Fiksuojame neryškųjį trikampį (L, T, R).
7. Renkamės kitą parametro t_1 reikšmę, tenkinančią sąlygą (5). Jeigu tokia reikšmė egzistuoja, grįžtame į žingsnį 4, kitu atveju pereiname į žingsnį 8.
8. Iš visų gautų (L, T, R) reikšmių klausimui ir silpnų studentų poaibiui W skaičiuojame optimistinius ir pesimistinius neryškiuosius trikampius.
9. Analogiškai vidutinių ir stiprių studentų poaibiams A ir S kartojame žingsnius 3–8 ir gauname optimistinius ir pesimistinius neryškiuosius trikampius šiems poaibiams.
10. Rezultatas: brėžiame klausimo optimistinius ir pesimistinius neryškiuosius trikampius poaibiams W, A ir S. Nustatome klausimo skiriamąją gebą.

5 Vieno konkretaus testo klausimų analizė

Parodysime kaip buvo analizuojami kai kurie aukštosios matematikos testo klausimai. Testą laikė 106 VGTU statybos fakulteto studentai, jų žinios buvo vertinamos testo balais skaičiuojant teisingų atsakymų į 20 klausimų skaičius.

Studentų testų balai pateikti lentelėje:

Bbalai	1	2	3	4	5	6	7	8	9	10
Stud. sk.	1	0	2	2	3	4	1	7	4	10
Balai	11	12	13	14	15	16	17	18	19	20
Stud. sk.	7	5	14	11	13	3	6	7	4	2

Parametrai (3) buvo tokie: min = 1, max = 20, $\alpha = 9$, $\beta = 12$, $\gamma = 14$, $\delta = 17$. Taigi studentų dėl kurių neabejojama, kad jie yra silpni ($1 \leqslant b_i \leqslant 9$), yra 24; vidutinių ($12 \leqslant b_i \leqslant 14$) − 30; stiprių ($17 \leqslant b_i \leqslant 20$) − 19. Kiti $33 = 106 - (24 + 30 + 19)$ studentai buvo priskiriami prie vieno iš šių trijų poaibių, keičiant (5) parametrus t_i, konstruojant trapecijas (4), kurių yra atitinkamai 10, 100 ir 10.

Dėl apribojimų straipsnio apimčiai praleisime atliktų skaičiavimų detales ir pateisime tik du analizuojamus testo klausimus:

A	$\lim_{x\to 0} x^{14} \ln x =$	① 14;	② 0;
		③ ∞;	④ $\ln 14$;
		⑤ riba neegzistuoja;	⑥ 1.

Klausimo išsprendžiamumas (sunkumas) yra 74(%), neryškieji trikampiniai skaičiai pesimistiniu ir optimistiniu atvejais yra šie:

$$Tr_W^{pes}(42, 45.1, 52), \quad Tr_M^{pes}(70, 84.11, 92.86), \quad Tr_S^{pes}(83.57, 92.4, 95),$$
$$Tr_W^{opt}(44.9, 45.1, 46.84), \quad Tr_M^{opt}(80.24, 84.11, 84.87), \quad Tr_S^{opt}(90.46, 92.4, 92.4).$$

Matome, kad pesimistiniu atveju klausimas gerai diferencijuoja silpnus studentus: $Tr_W^{pes} \prec Tr_M^{pes} \preccurlyeq Tr_S^{pes}$). Optimistiniu atveju turime $Tr_W^{opt} \prec Tr_M^{opt} \prec Tr_S^{opt}$, ir tai reiškia, kad klausimas gerai diferencijuoja visus studentus.

Dar vienas testo klausimas

$$\boxed{\text{B}} \quad \lim_{z\to 2} \frac{z^2+5z-14}{\mathrm{tg}(17z-34)} =$$

① $\frac{9}{17}$; ② $\frac{25}{17}$; ③ $-\frac{9}{17}$; ④ $-\frac{25}{17}$; ⑤ $-\frac{90}{17}$; ⑥ 0; ⑦ ∞.

turėjo išsprendžiamumą 32(%) ir gerai diferencijuoja visus studentus ir optimistiniu ir pesimistiniu atvejais:

$$Tr_W^{pes}(17, 18.2, 22), \quad Tr_M^{pes}(22, 22.85, 33), \quad Tr_S^{pes}(44, 64.2, 74),$$
$$Tr_W^{opt}(18.2, 18.2, 19.2), \quad Tr_M^{opt}(25.18, 25.85, 27.25), \quad Tr_S^{opt}(57.34, 64.2, 64.2).$$

6 Išvados ir būsimieji tyrimai

Pasiūlyta testo klausimų skiriamosios gebos analizės metodika, kuri nereikalauja tikslaus testuojamų žinių vertinimo ir apsiriboja tik santykiniais pasiekimų įverčiais, pvz., neapdorotais testų balais. Atliktas eksperimentas leidžia tikėtis gauti stabilią klausimų klasifikaciją, tačiau statistiškai patikimoms išvadoms daryti tikslinga atlikti Monte Carlo tipo eksperimentus. Tai yra būsimųjų mūsų tyrimų objektas.

Literatūra

[1] A. Krylovas, N. Kosareva. Mathematical modeling of forecasting the results of knowledge testing. *Techn. Econ. Devel. Econ.* **14**(3):388–401, 2008.

[2] A. Krylovas. *Diskrecioji matematika: vadovelis.* Technika, Vilnius, 2009.

[3] H.-J. Zimmermann. *Fuzzy Set Theory–And Its Applications.* Kluwer Academic Publishers, 4th edition, 2001.

SUMMARY

Discrimination power of knowledge testing items

A. Krylovas, N. Kosareva, J. Karaliūnaitė

In the article one fuzzy classification algorithm of dichotomous test questions is proposed. Depending on the level of knowledge measured by the test, the test item for example may well differentiate whether all testees or only testees with strong or only with weak level of knowledge; as well as the test item may badly differentiate testees and therefore be inappropriate. In the research fuzzy sets mathematical theory is used.

Keywords: mathematical modeling, fuzzy sets, dicotomous tests.