

Studentų žinių norminių ir kriterinių vertinimų derinimas

Olga Navickienė, Aleksandras Krylovas

Mykolo Romerio universitetas, Socialinės informatikos fakultetas

Ateities g. 20, LT-08303 Vilnius

E. paštas: navickiene@mruni.eu, krylovas@mruni.eu

Santrauka. Bet kuris žinių vertinimo kiekybinių rezultatų interpretavimas atliekamas laikantis vieno iš dviejų principų – kriterinio arba norminio. Kriterinis vertinimas remiasi studento žinių (gebėjimų, įgūdžių, kompetencijų) palyginimu su tam tikrais standartais, dažniausiai aprašytais dalyko programoje. Taip vertinamos beveik visų Lietuvos universitetų studentų žinios. Kitas vertinimas – norminis – numato pasiekimų palyginimą tarpusavyje. Taip, pavyzdžiui, vertinami moksleiviai valstybiniuose brandos egzaminuose. Straipsnyje aptarti norminių vertinimų teoriniai ir praktiniai aspektai, galimybės juos taikyti studentų žinioms vertinti.

Raktiniai žodžiai: žinių vertinimas, studijų kokybė.

1 Įvadas

Studentų žinių vertinimo procesą išsamiai aprašo Universitetų veiklą reglamentuojantys dokumentai, privalomi ir dėstytojams, ir studentams. Bet kuris vertinimo kiekybinių rezultatų interpretavimas atliekamas laikantis kriterinio arba norminio vertinimo principo [2, 6]. Kriterinis vertinimas, kai studento įgytos kompetencijos lyginamos su kompetencijomis, numatytomis dalyko programoje, formaliai užtikrina žinių atitikimą tam tikriems įsilavinimo standartams, tačiau turi ir nemažai trūkumų [7], visų pirma, negarantuoja pakankamo objektyvumo [8]. Kitas – norminis vertinimas numato studentų pasiekimų palyginimą tarpusavyje. Būtent taip traktuoja norminį vertinimą ir normatyviniai dokumentai [9], ir edukologų rekomendacijos [14, 10]. Mokslinėje literatūroje norminis vertinimas paprastai suprantamas siauriau ir turimas omenyje testuojamųjų rezultatų normalusis tikimybinis skirstinys. Tada tam tikra dalis studentų (pavyzdžiui, 5%) gauna aukščiausią įvertinimą, kita dalis – sekantį ir t. t. Teoriškai tai leidžia lyginti, pavyzdžiui, skirtingų metų arba universitetų vertinimus. Kadangi to pasiekti pavyksta tik esant standartizuotiems testams [1], praktikoje apsiribojama testuojamųjų procentinių rangų apskaičiavimu, taip, pavyzdžiui, vertinami valstybinių brandos egzaminų rezultatai. Nepaisant to, kad priėmimas į universitetus vyksta pagal abiturientų brandos egzaminų norminių vertinimų rezultatus, patys universitetai beveik nenaudoja studentų žinių norminių vertinimų. Juos taiko tik kai kurie universitetai [15], bet Bolonijos deklaracijoje numatytas ECTS kreditų sistemos diegimas Lietuvos universitetuose jau dabar reikalauja derinti vertinimus su ECTS pažymių skale. Pateiktas ECTS naudojimo apraše [5] pavyzdys rekomenduoja aukščiausią vertinimą (10) taikyti 5% studentų, (9) – 10%, (8) – 35% ir t. t. Problema yra ta, kad kriterinė vertinimo sistema neturi nieko bendra su jokia net ir labai orientaciniu aprioriniu pažymiu skirstiniu.

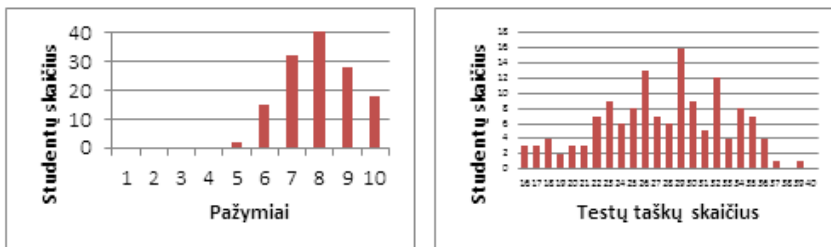
Šio darbo tikslas – parodyti kelis testo rezultatų vertinimo būdus ir juos pritaikyti konkrečiai imčiai. Taip pat juos palyginti kaip tarpusavyje, taip ir su kriterinio vertinimo rezultatu. Šiuo tikslu apskaičiuojame ankstesniame mūsų darbe [12] pasiūlytų vertinimo kokybės indikatorių reikšmes.

2 Tyrimo metodologija

Tarkime, kad turime n studentų testų ar kitų atsiskaitymų rezultatus t_1, \dots, t_n , $t_j \in 0, 1, \dots, T$. Be papildomos informacijos šie rezultatai neleidžia vertinti studentų žinių, rašyti j -ajam studentui pažymį p_j . Labiausiai paplitęs vertinimas – $p_j = [(t \cdot 10)/T]$ (čia $[]$ – sveikoji dalis, tačiau apvalinimas yra nebūtinai matematinis [6]), formaliai traktuojamas kaip kriterinis, bet tam testas turi tenkinti griežtus validumo reikalavimus [1]. Norminiu vertinimu vadiname rezultatų t_j transformaciją į dešimties balų pažymių skalę p_j : $p_j = T(j)$, kai T yra bet kuri monotonišė funkcija, t. y. $(\forall t_j < t_i) T(t_j) \leq T(t_i)$. Šiame darbe nagrinėjame penkis testo rezultatų įvertinimo formas: T_1 – *standartinius rodiklius*, kurie gaunami tiesiniais pertvarkymais, apskaičiuojant vidurkį ir standartinį nuokrypį; T_2 – *stenainus* (pagal C.M. Charlesą [3]) (angl. standard nine), pradiniai testų taškai paverčiami ne į atskirus balus, bet į balų grupes, padalijant 40 taškų seką į 9 rangus (lygius); surinktas stenainų skaičius parodo vertinimo lygmenį; T_3 – *procentilius* – procentai studentų, surinkusių tiek pat arba mažiau taškų, negu konkretus studentas; T_4 – *stenus* – testų taškai paverčiami į dešimties balų skalę; T_5 – *maksimalaus informatyvumo rodiklius*, duota imtis padalijama į 10 lygių intervalų, kiekvienam iš jų priskiriamas balas dešimties balų sistema, kad kiekvieną balą atitiktų 10% visų studentų (10% studentų surinkusių mažiausiai taškų gauna 1 balą iš 10).

Šiuos hipotetinius vertinimus mes lyginsime su realiais tų pačių studentų pažymiais, t. y. rezultatais pagal kriterinę sistemą dešimties balų skalėje. Vertinimams T_1, \dots, T_5 palyginti apskaičiuosime tris indikatorius:

- $I = \text{entr}(k_0, \dots, k_n) = -\sum_{j=0}^n \frac{k_j}{n} \ln \frac{k_j}{n}$ – testu teikiamos informacijos kiekis [13]. Šiame straipsnyje indikatorius I dalinamas iš funkcijos reikšmės, apskaičiuotos pradinėms testų taškų reikšmėms (žr. 1 pav.).
- $S = \frac{100}{1 + \frac{1}{n} \sum_{i=1}^n |p_i - r_i|}$ – įverčių sutapimas; čia p_i – i -ojo studento testo pažymį (testo taškai perskaiciuoti į 10-ties balų skalę), r_i – to paties studento pažymys iš žiniaraščio, n – studentų skaičius;



1 pav. Testų taškų ir pažymių pasiskirstymas.

- $K = \frac{n \sum_{i=1}^n p_i r_i - (\sum_{i=1}^n p_i)(\sum_{i=1}^n r_i)}{\sqrt{(n \sum_{i=1}^n p_i^2 - (\sum_{i=1}^n p_i)^2)(n \sum_{i=1}^n r_i^2 - (\sum_{i=1}^n r_i)^2)}} \cdot 100$ – Pirsono koreliacijos koeficientas [4] padaugintas iš 100.

Tyrimo dalyvavo Mykolo Romerio universiteto Politikos ir vadybos fakulteto Viešojo administravimo studijų antro kurso ir Organizacijų vadybos studijų pirmo kurso studentai, iš viso 248 studentai, kuriems matematika buvo dėstoma tik vieną semestrą. Studentai laikė du testus po 20 uždarojo tipo klausimų. Pagal iš anksto dėstytojo nustatytą reikalavimą: studentai neišlaikę pirmojo testo, negalėjo laikyti antrojo (antrojo testo metu studentai laikė pirmą). Šiame darbe išnagrinėti tik tų, studentų rezultatai ir vertinimai, kurie iš pirmo karto išlaikė pirmą testą (gavo 8 taškus ir daugiau iš 20 galimų) ir laiku laikė antrą testą. Tokių studentų – 141 (57% viso srauto studentų). Pasirinkti būtent šie duomenis, kadangi jie leidžia lyginti kriterinius ir norminius vertinimus.

3 Rezultatai

Hipotetiniai norminiai vertinimai T_1 – T_5 apskaičiuojami taip:

T_1 . Studentų testų rezultatus suskirstome į intervalus, prieš tai apskaičiuojant vidurkį $\mu = 27$ ir standartinį nuokrypį $\sigma = 5$, ir suskaičiuojame kiek studentų gavo įvertinimus, patenkančius į šiuos intervalus. Po to testų taškų intervalus transformuojame į dešimties balų sistemą: geriausi studentai (mūsų atveju vienas studentas) gauna didžiausią įvertinimą dešimties balų sistema, visi kiti – atitinkamai vienu balu mažiau (žr. 1 lentelę).

T_2 . Pirminiai studentų testų įverčiai paverčiami į stenainus, išlygiagretinant įverčius pagal jų dydį ir prirašant stenainus atitinkamai pagal normaliojo tankio reikšmes (procentais). Kad galėtume šiuos vertinimus palyginti su studentų galutiniais pažymiais, stenainus transformuojame į dešimties balų sistemą, pradedant nuo pažymio 2 (žr. 2 lentelę).

1 lentelė. Norminio vertinimo T_1 skalė.

Intervalas	Studentų skaičius	Pažymys
$[\mu - 4\sigma; \mu - 3\sigma)$	[7; 12)	0
$[\mu - 3\sigma; \mu - 2\sigma)$	[12; 17)	3
$[\mu - 2\sigma; \mu - 1\sigma)$	[17; 22)	15
$[\mu - 1\sigma; \mu)$	[22; 27)	43
$[\mu; \mu + 1\sigma)$	[27; 33)	55
$[\mu + 1\sigma; \mu + 2\sigma)$	[33; 38)	24
$[\mu + 2\sigma; \mu + 3\sigma)$	[38; 40] ¹	1

2 lentelė. Norminio vertinimo T_2 skalė.

Procentas	4%	7%	12%	17%	20%	17%	12%	7%	4%
Studentų skaičius	6	10	17	24	28	24	17	10	5
Stenainas	1	2	3	4	5	6	7	8	9
Pažymys	2	3	4	5	6	7	8	9	10

¹ Intervalas $[\mu + 2\omega; \mu + 3\omega)$ yra nuo 38 iki 43 taškų, kadangi iš testų galima surinkti daugiausiai 40 taškų, tai šį intervalą mes užrašėme kaip [38; 40].

3 lentelė. Norminio vertinimo T_3 skalė.

Taškų skaičius	Studentų skaičius	Procentilis	Pažymys	Taškų skaičius	Studentų skaičius	Procentilis	Pažymys
16	3	2	1	28	6	52	5
17	3	4	1	29	16	64	6
18	4	7	1	30	9	70	7
19	2	9	1	31	5	74	7
20	3	11	1	32	12	82	8
21	3	13	1	33	4	85	9
22	7	18	2	34	8	91	9
23	9	24	2	35	7	96	10
24	6	28	3	36	4	98	10
25	8	34	3	37	1	99	10
26	13	43	4	39	1	100	10
27	7	48	5				

4 lentelė. Norminio vertinimo T_4 skalė.

Intervalas	Dažnis	Stenas	
$[\mu - 5/2\sigma; \mu - 2\sigma)$	[14; 18)	6	1
$[\mu - 2\sigma; \mu - 3/2\sigma)$	[18; 20)	6	2
$[\mu - 3/2\sigma; \mu - 1\sigma)$	[20; 23)	13	3
$[\mu - 1\sigma; \mu - 1/2\sigma)$	[23; 25)	15	4
$[\mu - 1/2\sigma; \mu)$	[25; 28)	28	5
$[\mu; \mu + 1/2\sigma)$	[28; 31)	31	6
$[1/2\sigma; \mu + 1\sigma)$	[31; 33)	17	7
$[\mu + 1\sigma; \mu + 3/2\sigma)$	[33; 36)	19	8
$[\mu + 3/2\sigma; \mu + 2\sigma)$	[36; 38)	5	9
$[\mu + 2\sigma; \mu + 5/2\sigma)$	[38; 40)	1	10

T_3 . Kiekvienam studentui apskaičiuojama kiek studentų surinko tiek pat arba mažiau taškų, po to apskaičiuojama kiek procentų sudaro šis skaičius nuo visų studentų. Kuo žemesnis procentilis, tuo prastesnė studento pozicija standartizuotoje imtyje. Gauti procentiliai verčiami į dešimties balų sistemą, padalijant procentilių reikšmes iš 10 ir suapvalinant jas iki dešimtųjų (žr. 3 lentelę).

T_4 . Studentų testų įverčiai paverčiami į stenas, iš pradžių apskaičiuojant vidurkį ir standartinį nuokrypį, po to nuo vidurkio suskaičiuojama po penkis intervalus po $1/2\sigma$ į abi puses. Gauti intervalai sunumeruojami iš kairės į dešinę. Intervalo numeris ir yra stenas (žr. 4 lentelę).

T_5 . Studentų testų rezultatai suskirstomi į 10 intervalų, taip kad kiekvieną balą iš dešimties galimų gautų po 10% viso srauto studentų. Į kiekvieną grupę nepatenka po lygiai studentų, kadangi yra studentų, kurie gavo vienodą taškų skaičių, tuo pačiu jie turi gauti ir vienodą pažymį (žr. 5 lentelę).

Iš 6 lentelės matome, kad norminiai vertinimai gali būti pakankamai įvairūs, ir taikant šiuos skirtingus vertinimus studentai už tuos pačius testo taškus gali gauti skirtingus pažymius.

4 Išvados

Iš pateiktos vertinimo kokybę apibūdinančių indikatorių I , S , K reikšmių 7 lentelės matome, kad norminių vertinimų suderinamumas nėra prastesnis už tų pačių studen-

5 lentelė. Norminio vertinimo T_5 skalė.

Taškų skaičius	Studentų skaičius	Studentų dalis	Pažymys	Taškų skaičius	Studentų skaičius	Studentų dalis	Pažymys
16	3			27	7	9%	5
17	3			28	6		
18	4	13%	1	29	16	11%	6
19	2			30	9	10%	7
20	3			31	5		8
21	3			32	12	9%	8
22	7	11%	2	33	4	9%	9
23	9			34	8		
24	6	10%	3	35	7		
25	8			36	4	9%	10
26	13	9%	4	37	1		
				39	1		

6 lentelė. Norminių vertinimų palyginimas.

Taškų skaičius	Studentų skaičius	T_1	T_2	T_3	T_4	T_5	Taškų skaičius	Studentų skaičius	T_1	T_2	T_3	T_4	T_5
16	3	5	2	1	1	1	28	6	8	6	5	6	5
17	3	6	2	1	1	1	29	16	8	6	6	6	6
18	4	6	3	1	2	1	30	9	8	7	7	6	7
19	2	6	3	1	2	1	31	5	8	7	7	7	7
20	3	6	3	1	3	1	32	12	8	8	8	7	8
21	3	6	4	1	3	1	33	4	9	8	9	8	9
22	7	7	4	2	3	2	34	8	9	8	9	8	9
23	9	7	4	2	4	2	35	7	9	9	10	8	10
24	6	7	5	3	4	3	36	4	10	10	10	9	10
25	8	7	5	3	5	3	37	1	10	10	10	9	10
26	13	7	5	4	5	4	39	1	10	10	10	10	10
27	7	8	6	5	5	5							

7 lentelė. Trijų vertinimo kokybės indikatorių reikšmės.

	T_1	T_2	T_3	T_4	T_5
I	47	70	78	70	78
S	67	33	27	30	27
K	84	85	86	86	86

tų žinių skirtingų dėstytojų vertinimus (dėl straipsnio puslapių skaičiaus apribojimo šių indikatorių reikšmių interpretavimo neaprašome, daugiau žr. mūsų straipsniuose [12, 11]). Taigi įvairiais būdais pereinant prie norminio vertinimo, gavome, kad vertinimai pakankamai gerai suderinti su kriteriniu vertinimu. Mūsų darbo tikslas – iširti principinę galimybę konstruoti suderintus su kriteriniu norminius vertinimus. Geriausio įvairių kriterijų prasme norminio vertinimo konstravimas yra mūsų tolimesnių tyrimų objektas ir šiame darbe to nenagrinėjame. Todėl apsiribojame tik geresnių studentų duomenimis – semestro metu sėkmingai atlikusių visus atsiskaitymus, nes manome, kad tokių studentų žinių vertinimas yra patikimesnis.

Literatūra

- [1] A. Anastasi and S. Urbina. *Psychological Testing*. Upper Saddle River, NJ, Prentice Hall.
- [2] T. Bulajeva. *Žinių ir kompetencijų vertinimas: kaip sukurti studentų pasiekimų vertinimo metodiką: metodinė priemonė*. Petro ofsetas, Vilnius, 2007.
- [3] C.M. Charles. *Pedagoginio tyrimo įvadas*. Alma littera, Vilnius, 1999.
- [4] V. Čekanavičius ir G. Murauskas. *Statistika ir jos taikymai*: vadovėlis aukštųjų mokyklų studentams. TEV, Vilnius, 2006.
- [5] *ECTS naudotojo vadovas*. Europos komisijos leidinį ECTS User Guide išleista 2009 m. vertė Nacionalinė Bolonijos ekspertė Raimonda Markevičienė. [žiūrėta 2012-05-29]. Adresas internete:
http://ec.europa.eu/education/lifelong-learning-policy/doc/ects/guide_lt.pdf.
- [6] S. Girdzijauskas. *Studentų žinių kontrolė ir vertinimas*. Vilniaus universiteto leidykla, Vilnius, 1999.
- [7] N.L. Gage ir D.C. Berliner. *Pedagoginė psichologija*. Alma littera, Vilnius, 1994.
- [8] R. Kriauzienė, A. Krylovas ir N. Kosareva. Studentų matematikos žinių vertinimo subjektyvumo problema: teoriniai ir praktiniai aspektai. *Soc. techn.*, **1**(1):121–138, 2011.
- [9] Lietuvos Respublikos švietimo ir mokslo ministro 2004 m. vasario 25 d. įsakymas Nr. ISAK-256 „Dėl mokinių pažangos ir pasiekimų vertinimo sampratos“. *Valstybės žinios* 35-1150, 2004.
- [10] *Mokomoji medžiaga vertintojams: vertinimo standartas*. VDU Profesinio rengimo studijų centras, Kaunas, 2006.
- [11] O. Navickienė and A. Krylovas. Integral criterion for measuring the quality of teachers' evaluation. *Soc. techn./Social technologies*, **4**(2):1–9, 2012.
- [12] O. Navickienė ir A. Krylovas. Studentų žinių vertinimo kokybės kriterijų modeliavimas. *Stud. šiuol. visuom./Studies in modern society*, **3**(1):177–184, 2012.
- [13] V. Stakėnas. *Informacijos kodavimas*. Vilniaus universiteto leidykla, Vilnius, 1996.
- [14] *Vertinimas ugdymo procese*: knyga mokytojui: projekto „Vertinimas ugdymo procese“ (2004–2006 m.) patirtis. Leidinį sudarė Irma Neseckienė. AJA viešieji ryšiai (Spauda), Vilnius, 2006.
- [15] A. Zvicevičiūtė. Studentus išgašdino nauja vertinimo sistema. „Savas“ Lietuvos studentų laikraštis, **4**(189), 2006. [Žiūrėta 2012-05-29]. Adresas internete:
<http://www.savas.lt/archyvas-rubrikos/pirmas-puslapis/1006-studentus-igsgdino-nauja-vertinimo-sistema.html>.

SUMMARY

Adjustment of norm-referenced and criterion-referenced evaluations of students' knowledge

O. Navickienė, A. Krylovas

Any evaluation of knowledge is performed in accordance with one of the two principles of evaluation – normative or criteria. An evaluation criterion is based on student's knowledge (abilities, skills, competencies) comparison with certain standards usually described in a programme of a subject. Knowledge of students of almost all Lithuanian universities is evaluated in such principle. Another evaluation – normative – provides a comparison of outcomes with each other. For example, students are assessed in state examinations. Theoretical and practical aspects of normative evaluation and possibilities to apply them to assess students' knowledge are discussed in this article.

Keywords: evaluation of knowledge, quality of studies.