

Studentų matematikos žinių kompleksinis vertinimas semestro metu

Aleksandras KRYLOVAS, Juozas RAULYNAITIS (VGTU)

el. paštas: akr@fm.vtu.lt

1. VGTU studentų žinių vertinimas reglamentuotas keliomis instrukcijomis, kurios numato, kad egzamino pažymys atitinka tam tikrą dalyko modulyje aprašytą žinių įsisavinimo lygį: pažymys 5 (silpnai) rašomas už 50–55 % žinių, pažymys 10 (puikiai) – už 96–100 %. Kitas šių instrukcijų reikalavimas yra tas, kad tam tikrą egzamino pažymio dalį studentas gauna už atliktus semestro metu laboratorinius, kontrolinius, namų darbus, išlaikytus kolokviumus ir pan. Visų šių darbų pažymiai bei sesijos egzamino pažymys sumuojami su iš anksto nustatytais ir praneštais studentams svertiniais koeficientais. Šiame darbe nagrinėjamas straipsnio autorių atliktas eksperimentas, kai studentas galutinio vertinimo pažymį galėjo gauti vien už darbą semestro metu, o sesijos egzaminą laikė nesurinkus pakankamo taškų skaičiaus arba norint pagerinti turimą įvertį. Egzamino įvertis p_e buvo skaičiuojamas taip:

$$p_e = (0.167t_{15}/15 + 0.333t_{30}/30 + 0.500p_{45}/45) \cdot 10. \quad (1)$$

Čia indeksais pažymėtas maksimalus skaičius taškų, gaunamų laikant du testus ir sprendžiant uždavinius bei atliekant įvairius darbus pratybų metu.

2. Šiam eksperimentui atlikti buvo pasirinktas vienas VGTU Statybos fakulteto antrojo kurso srautas (4 grupės), kai vienas šio straipsnio autorius skaitė modulio Matematika 3 paskaitas, o kitas vedė pratybas visoms srauto grupėms. Šio modulio turinį sudaro tikimybių teorijos ir matematinės statistikos kursas, kuriam skiriama 3 savaitinės valandos paskaitų ir 2 valandos pratybų. Pratybų metu vieną akademinę valandą dažniausiai buvo paaiškinama, kaip sprendžiami atitinkamos temos uždaviniai, o kitą valandą studentai savarankiškai sprendė analogiškus uždavinius. Jų įvertinimas per kiekvieną užsiėmimą vidutiniškai buvo 3 taškai, o per visą semestrą pratybų didžiausias vertinimas galėjo būti 45 taškai. Matematinės statistikos kurse dar buvo ir namų darbai, kurių vertinimas analogiškas.

Paskaitas skaitęs dėstytojas tikrino žinių įsisavinimą, pateikdamas du testus per semestrą. Pirmasis testas (t_{15}) buvo skirtas apibrėžimams, žymėjimams bei pagrindinėms sąvokoms (formulėms ir teoremų formuluotėms) tikrinti. Atsakyti į pirmojo testo klausimus reikėjo per 20 minučių. Jis buvo laikomas nesinaudojant literatūra ir nepateikus studentams klausimų pavyzdžių. Skirtingi testo variantai buvo išlygiagretinti: turėjo įvairias formules bei įvairius kintamųjų žymėjimus (žr. [2]). Testai buvo uždarojo tipo: kiekvienas klausimas turėjo nuo 3 iki 10 atsakymo variantų.

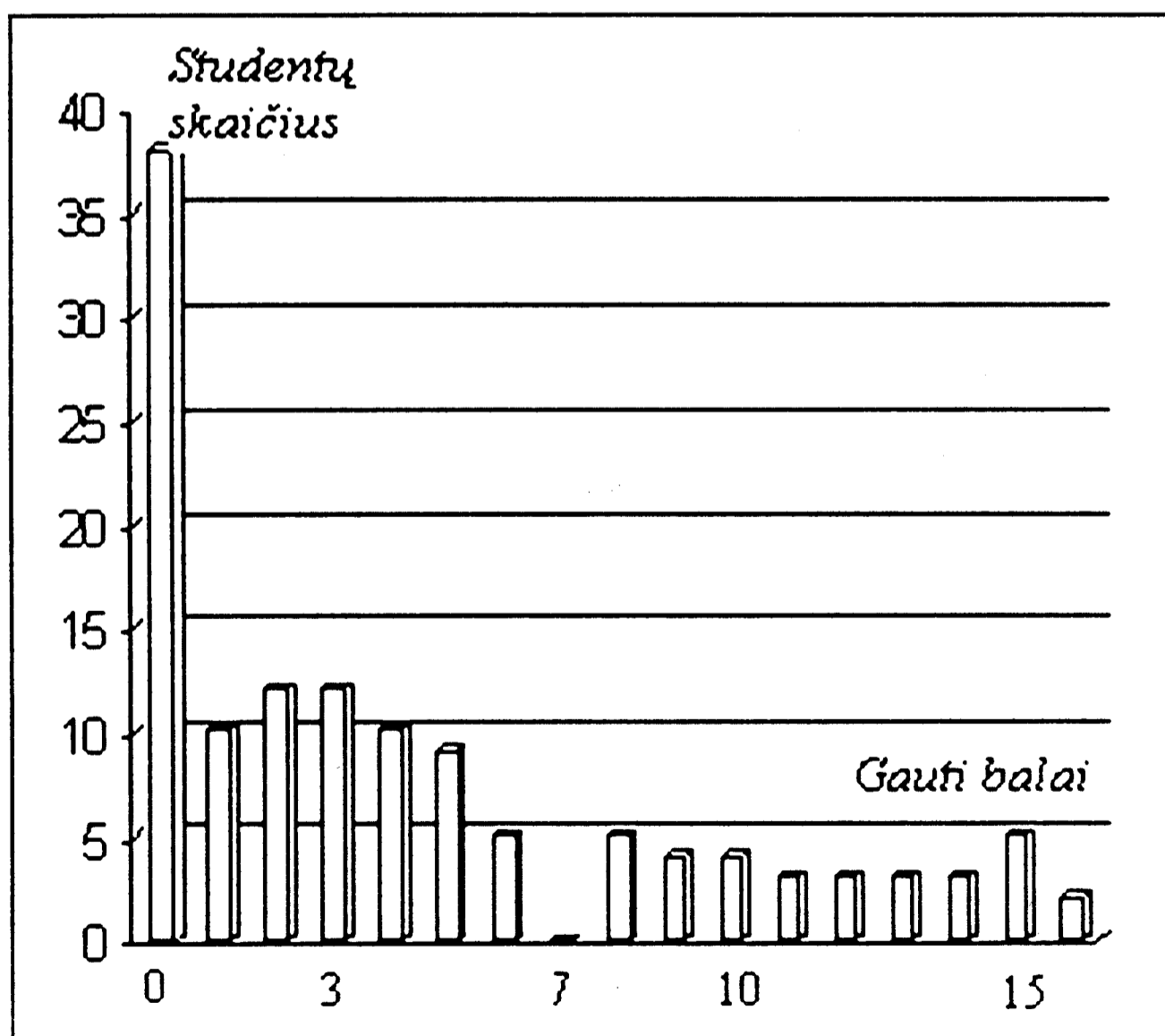
3. Vertinant žinias uždarojo tipo testais visada sprendžiama rezultato skaičiavimo problema. Vien teisingų atsakymų įskaitymas niekaip neatsižvelgiant į klaidingus skatina studentus spėlioti ir ne tik toleruoja tam tikrą nesąžiningumą, bet ir iškreipia įvertį. Esant pakankamai dideliame klausimų skaičiui (mūsų atveju 15 ir 30), gausime teigiamą įverčio postūmį, sąlygotą atspėtais atsakymais. Tai nėra labai blogai, taikant norminius žinių vertinimus (kai vertinami santykiniai testuojamųjų pasiekimai, pavyzdžiui, taip vertinamos abiturientų žinios Valstybiniuose Brandos egzaminuose). Tačiau toks įverčio iškreipimas blogai suderinamas su kriteriniais vertinimais. Būtent kriterinis vertinimas, kai bandoma išmatuoti studento įsisavintų žinių dalį, taikomas VGTU.

Mes siūlome tokį šios problemos sprendimą. Testo rezultatas r skaičiuojamas taip:

$$r = t - b(k). \quad (2)$$

Čia t – teisingų atsakymų skaičius, k – klaidingų atsakymų skaičius, $b(k)$ – baudos funkcija. Pirmojo testo $b(k) = k$, antrojo – $b(k) = [(k - 1)/2]$. Studentai iš anksto informuojami apie šį atėmimą ir jiems rekomenduojama ne tik nespėlioti, bet ir praleisti klausimus, dėl kurių yra abejojama. Patirtis rodo, kad nemažiau kaip pusė studentų vadovaujasi šiomis rekomendacijomis tik gavus gerokai prastesnį rezultatą negu tikėjosi. Pakartotiniams testų laikymams būdingas daug mažesnis spėliojančiųjų skaičius.

Pirmąją testą pirmą kartą laikė 118 studentų. Teigiamų atsakymų skaičių vidurkis buvo 6,68, o pirmojo testo perskaičiuotų rezultatų, atimant baudos taškus, vidurkis lygus 2,59. Pastebėkime, kad pirmojo testo vertinimų, taikant ir netaikant baudos taškus, koreliacijos koeficientas buvo 0,92. Diagramoje (1 pav.) parodyti galutiniai pirmojo



1 pav. Pirmojo testo rezultatai.

testo pirmojo laikymo rezultatai. Pirmas stulpelis atitinka gavusius nulį ir mažiau balų, atėmus baudos taškus. Išlaikiusiems (gavusiems 7 ir daugiau balų) pridėtas vienas balas todėl atitinkamo stulpelio aukštis lygus nuliui, o maksimalus rezultatas yra 16 balų. Visi stulpeliai į dešinę nuo jo atitinka išlaikiusiųjų gautus balus.

Pirmojo testo rezultatai buvo palyginti su pirmųjų 5 pratybų įvertinimais, už kurias buvo galima gauti irgi 15 taškų. Pratybų įvertinimų vidurkis buvo 8,85, koreliacijos koeficientai 0,47 (testas be baudos) ir 0,45 (su bauda). Tai eilinią kartą rodo gerai žinomą didaktikoje kriterinių vertinimų nestabilumą (žr. [1]) ir nereikėtų tikėtis, kad bet kuri (1) tipo formulė iš tikrųjų matuoja įgytų žinių lygį.

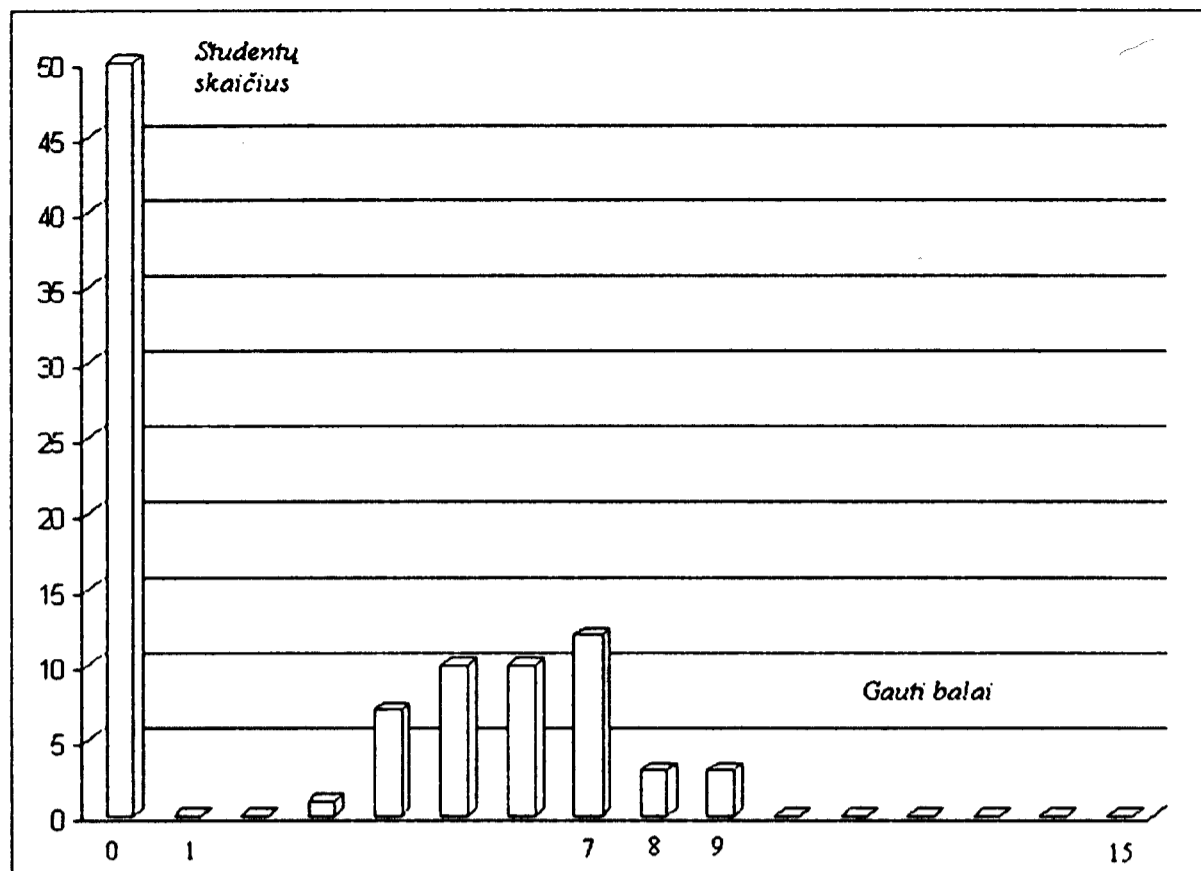
4. Antrasis testas skirtas žinių supratimui bei taikymo gebėjimams tikrinti. Jis buvo sudarytas iš 30 uždarojo tipo teorinio pobūdžio klausimų ir uždavinių. Kai kurie klausimai buvo sujungti į blokus. Vienas uždavinys galėjo būti taip suskaidytas į kelis klausimus, kad atsakymas į antrąjį panaudoja pirmąjį, į trečiąjį–antrąjį ir pan. Taigi už vieną kiek sudėtingesnę teisingai išspręstą uždavinį studentas galėjo gauti 4–6 taškus. Studentams buvo pateikti antrojo testo išlygiagretintų variantų pavyzdžiai. Jį laikant buvo leidžiama naudotis literatūra bei išspręstais pavyzdžiais.

Antrąjį testą galėjo laikyti tik išlaikiusieji pirmąjį. Pirmojo testo išlaikymo kriterijus (8 taškai iš 15) buvo iš anksto žinomas studentams. Po testo laikymo jis buvo sumažintas iki 7 balų ir apie šią galimybę studentai irgi žinojo prieš testui rašant. Po pirmo laikymo išlaikiusių pirmąjį testą buvo tik 27 studentai (23%).

Kiti (96 studentai) pakartotinai laikė pirmąjį testą. Šį kartą testo klausimai (irgi 15) buvo parinkti iš dviejų temų: atsitiktiniai įvykiai ir atsitiktiniai dydžiai (pirmą kartą buvo tik viena tema). Šio laikymo rezultatai (su bauda) parodyti 2 pav. Pirmasis stulpelis rodo, kiek studentų gavo nulį ir mažiau balų. Kiti nenuliniai stulpeliai atitinka gavusių 3, 4, 5, 6, 7, 8 ir 9 balus studentų skaičius. Gavusių 10 arba daugiau taškų (iš 15) nebuvo. Teigiamų atsakymų vidurkis 5,76, o vidurkis su baudos taškais buvo 3,05. Šie rezultatai irgi buvo palyginti su pirmųjų 5 pratybų rezultatais. Pakartotinai laikusiųjų pirmąjį testą pratybų įvertinimo vidurkis 7,74, pratybų koreliacijos koeficientai su pirmuoju testu 0,32 (be baudos), 0,34 (su bauda).

Per antrą pirmojo testo laikymą išlaikiusiųjų buvo 42 (44%), kai išlaikymo kriterijus buvo dar sumažintas iki 6 taškų. Jiems semestro pabaigoje buvo leista laikyti antrąjį testą. Taigi iš viso antrąjį testą semestro metu laikė 69 studentai. Antrojo testo, apskaičiuotų pagal (2) formulę, įverčių vidurkis lygus 20,96. Šie rezultatai palyginti su 10 pratybų, už kurias irgi buvo galima gauti 30 taškų, rezultatais. Laikiusiųjų antrąjį testą pratybų įvertinimų vidurkis 23,29, koreliacijos koeficientas su antruoju testu 0,36 (be baudos taškų). Įdomu, kad nebuvo pastebėta statistiškai reikšmingos koreliacijos tarp pirmojo ir antrojo testo rezultatų. Šis faktas, be abejo, vertas tolesnių tyrimų, tačiau paprastas jo paaiškinimas yra pernelyg žemas pirmojo testo įskaitymo kriterijus.

5. Tikrinant bet kuriuos raštu atliktus darbus sunku išvengti subjektyvių faktorių įtakos vertinimui. Vienas šių faktorių yra tvarkingumas, kuris, be abejo, turi įtaką (1) formulės dėmeniui p_{45} . Turint objektyvaus žinių vertinimo įrankį (testą), galima įvertinti tvarkingumo įtaką pažangumui. Buvo atrinkti $n_x = 23$ namų darbai, kurių atlikėjus galima pavadinti tvarkingais ir $n_y = 15$ darbų, kurie atlikti netvarkingai, skubotai, neįsigilinus į užduotį. Tvarkingų atlikėjų pirmojo testo pirmo laikymo teisingų atsakymų skaičiaus vidurkis $\bar{x} = 7,26$, pataisytoji dispersija $s_{1x}^2 = 7,29$. Netvarkingų



2 pav. Pirmojo testo antro laikymo rezultatai.

atlikėjų: $\bar{y} = 5,60$, $s_{1y}^2 = 3,97$. Taikydami Fišerio kriterijų $F = \frac{s_{1x}^2}{s_{1y}^2}$, kurio apskaičiuotoji reikšmė yra 1,84, priimame nulinę hipotezę, kad teoriniai vidutiniai kvadratiniai nuokrypiai yra lygūs (alternatyvioji hipotezė $\sigma_x > \sigma_y$; reikšmingumo lygmuo $\alpha = 0.05$; Fišerio skirstinio kritinė reikšmė $f_{\alpha; n_x-1; n_y-1} = 2,3667$). Po to taikydami Stjudento reikšmingumo kriterijų $T = \frac{\bar{x} - \bar{y}}{\sqrt{s_{1x}^2/n_x + s_{1y}^2/n_y}}$, kurio apskaičiuotoji reikšmė yra 2,687, tikriname nulinę hipotezę, kad teoriniai vidurkiai lygūs ($a_x = a_y$). Kai reikšmingumo lygmuo 0,05 ir alternatyvioji hipotezė $H_a : a_x > a_y$, Stjudento skirstinio kritinė reikšmė $t_{0,05; n_x+n_y-2} = 1,688$. Taigi nulinę hipotezę atmetame ir galime teigti, kad tvarkingumas turi įtakos pažangumui.

6. Taigi straipsnyje aprašytas studentų darbo vertinimo semestro metu priemonių kompleksas leidžia padaryti kai kurias išvadas.

Tik nedidelė studentų dalis, net esant tokiems rimtiems skatinimo stimulams kaip nelaikyti sesijos egzamino, intensyviai dirba semestro metu. Nemaža studentų dalis nesugeba patenkinamai išsivinti bazinę dalyko dalį net po kelių laikymų. Svarstytinas klausimas, ar reikia bandyti vertinti aukštesniųjų lygių gebėjimus bei igūdžius, kai tik minimaliai išsivintos dalyko žinios (mūsų atveju pagrindinės tikimybių teorijos formulės ir teoremos).

Autoriai mano, kad (1) tipo formulės žinioms vertinti gali būti taikomos sesijos egzamino pažymiui skaičiuoti su sąlyga, kad studentas laiku ir pakankamai gerai atlieka visus ar beveik visus privalomuosius darbus. Nepažangiems studentams turėtų būti suformuluoti kiti vertinimo kriterijai.

Vienas iš pasiūlymų galėtų būti toks. Studentų, kelis kartus neįveikusių tam tikrų reikalavimų (neišlaikiusių kolokviumų, neparašiusių kontrolinių darbų, neatlikusių namų darbų ir pan.), žinioms tikrinti nebetaikyti (1) tipo formulę, o vertinti tik dalyko pagrindo žinias. Tokiu atveju egzamino pažymys galėtų būti tik 5 arba 6, o įvertinimas gali turėti norminio vertinimo elementus, kai suskaičiuojami visų laikusiųjų rezultatai

ir išlaikymo kriterijus nustatomas, atsižvelgiant ne tik į minimalių reikalingų žinių lygį, bet ir į neišlaikiusių skaičių. Esant masiniam aukštajam išsilavinimui ši skaičių turi nustatyti ne dėstytojas, o aukštosios mokyklos administracija.

Literatūra

1. A. Krylovas, J. Raulynaitis, J. Jaurienė, Apie matematikos žinių įvairių vertinimų suderinamumą, *Liet. matem. rink.*, **42** (spec. nr.), 397–401 (2002).
2. A. Krylovas, J. Raulynaitis, Vieno tikimybių teorijos uždavinio išlygiagretinimo patirtis, *Liet. matem. rink.*, **43** (spec. nr.), 357–360 (2003).

SUMMARY

A. Krylovas, J. Raulynaitis. Integrated estimation of mathematical knowledge of students at semester time

The experiment of integrated estimation of students mathematical knowledge at semester time in Vilnius Gediminas technical university is described.

Keywords: didactic, training requirements, tests.