LIETUVOS MATEMATIKOS RINKINYS. LMD DARBAI ISSN 0132-2818
51 tomas, 2010, 408–413 www.mii.lt/LMR/

Studentų pasiekimų analizė naudojant ranginę logistinę regresiją

Gediminas Murauskas[1], Marijus Radavičius[1,2]

[1] *Vilniaus universitetas, Matematikos ir informatikos fakultetas*
Naugarduko g. 24, LT-03225 Vilnius
[2] *Matematikos ir informatikos institutas*
Akademijos g. 4, LT-08663 Vilnius
E. paštas: gediminas.murauskas@mif.vu.lt; mrad@ktl.mii.lt

Santrauka. Darbe nagrinėjama matematinių gebėjimų įvertinimo problema. Laikoma, kad studento matematinės analizės pažymys yra vienas iš pagrindinių matematinių gebėjimų rodiklių, leidžiantis gana patikimai prognozuoti jo tolesnių matematinės specialybės studijų eigą. Naudojant VU matematinių specialybių studentų duomenis ir ranginį logistinės regresijos su kintančiu masteliu modelį tiriama įvairių veiksnių, tarp jų ir egzaminuojančio dėstytojo vertinimo sistemos, įtaka matematinės analizės pažymiams.

Raktiniai žodžiai: matematiniai gebėjimai, pasiekimų analizė, proporcingųjų galimybių modelis, ranginė logistinė regresija, veiksnių sąveika, proporcingųjų šansų modelis.

Įvadas

Studijų kokybės ir mokslo rezultatyvumo gerinimas yra pagrindinės universitetų veiklos kryptys. Todėl faktorių, kurie lemia sėkmingas studijas nustatymas yra svarbus veiksnys tobulinant tiek priėmimo sistemą, tiek gerinant studijų kokybę. Yra nemažai darbų, kuriuose nagrinėjami pažangumo rodikliai (žr., pvz., [2]). Jie gali pateikti informacijos apie studijų eigą, tendencijas ir pan. Dažniausiai modeliuojama pažangumo rodiklių priklausomybė nuo studentų gebėjimų ir ir testų sunkumo (Rash modeliai, žr. apžvalgą [6]).

Studentų gebėjimai yra latentinis faktorius. Jiems išmatuoti sudaromi testai – vertinimo sistemų pagrindas. Kai vertinimo sistema nusistovi, ji virsta atitinkamų gebėjimų apibrėžimu, o su laiku – ir jų sinonimu. Todėl labai svarbu tinkamai parinkti išmatuojamus gebėjimų požymius ir kriterijus.

Šiame darbe nagrinėjamas vienas iš galimų gabumų matematikai indikatorių – matematinės analizės pažymys. Tyrimo tikslas yra nustatyti matematinės analizės įvertinimus įtakojančius faktorius. Pagrindinis tiriamas faktorius yra skirtingų dėstytojų vertinimo sistemos. Tai yra vienas iš rečiau tiriamų faktorių, bet jis yra tiesiogiai susijęs su matematinių gabumų vertinimo kokybe.

Darbe taikomi ranginės logistinės regresijos modeliai [4, 1]. Jų formalūs apibrėžimai pateikti kitame skyrelyje, kuriame aptariamos ir ranginio atsako kintamojo modeliavimo problemos. Trečiame ir ketvirtame – trumpai aprašytas atliktas tyrimas ir pateikti pagrindiniai rezultatai, gale – suformuluotos išvados.

Skaičiavimai atlikti su PASW Statistics 18 programa. Logistinės regresijos modelių analizei naudota PLUM procedūrą [5].

1 Ranginės regresijos modeliai

Tarkime Y_i yra i-tojo studento pasirinkto dalyko pažymys, o x_i žymi m-matį su juo susijusių aiškinančiųjų kintamųjų vektorių. Pažymys yra ranginis kintamasis. Todėl jo sąlyginio skirstinio modeliavimui įprasti regresiniai modeliai (tiesinė regresija ar net apibendrintieji tiesiniai modeliai) netinka. Ranginio kintamojo regresiniai modeliai plačiai taikomi sociologiniuose moksluose [1, 3, 6].

Klasikinis tokiu atveju taikomas modelis yra *daugialygė logistinė regresija su sukauptąja logit atsako funkcija* [1]. Taikymuose (žr., pvz., [3]) šis modelis dar vadinamas *rangine logistine regresija*. Laikoma, kad $Y_i \in \{1,\dots,K\}$, $i=1,\dots,N$, yra sąlyginai nepriklausomi atsitiktiniai dydžiai, kai žinomos visų aiškinančiųjų kintamųjų reikšmės (x_i, $i=1,\dots,N$). Pažymėkime $\pi_k(x_i) := \mathbf{P}\{Y_i = k \mid x_i\}$, $\Pi_k(x_i) := \mathbf{P}\{Y_i \leqslant k \mid x_i\}$, $k=1,\dots,K-1$, $i=1,\dots,N$. Kadangi kintamasis Y_i yra ranginis, tai jam yra prasminga palyginimo operacija. Taigi, formaliai apibrėžtos sukauptosios tikimybes $\Pi_k(x_i)$ turi ir adekvačią dalykinę interpretaciją. Didėjant k jos didėja, o tuo pačiu didėja ir įvykio $\{Y_i \leqslant k\}$ *sąlyginiai šansai* (*sąlyginė galimybė*) su sąlyga, kad $x_i = x$,

$$\frac{P(Y_i \leqslant k \mid x_i = x)}{P(Y_i > k \mid x_i = x)} = \frac{\Pi_k(x)}{1-\Pi_k(x)} := \theta_k(x). \tag{1}$$

Vadinasi,

$$\theta_k(x) \leqslant \theta_{k+1}(x), \quad \forall x \in \mathbf{R}^m, \ k=1,\dots,K-1. \tag{2}$$

Ranginės logistinės regresijos modelis [4, 1] nusako sąryšį tarp $\log\theta_k(x)$ ir tiesinio prediktoriaus $\eta_k(x) := \alpha_k + \beta' x$:

$$\operatorname{logit}(\Pi_k(x)) := \log(\theta_k(x)) = \alpha_k + \beta' x, \quad k=1,\dots,K-1. \tag{3}$$

Čia α_k, $k=1,\dots,K-1$, ir $\beta \in \mathbf{R}^m$ yra modelio parametrai. Sąlyga $\alpha_k \leqslant \alpha_{k+1}$ $\forall k=1,\dots,K-1$ garantuoja, kad išpildyta ir sąlyga (2).

Kadangi šiuo atveju $\log(\theta_{k+1}(x)/\theta_k(x)) = \alpha_{k+1} - \alpha_k$, $k=1,\dots,K-1$, tai visiems i įvykių $\{Y_i \leqslant k\}$ ir $\{Y_i \leqslant j\}$ *sąlyginių šansų santykis* $\theta_k(x)/\theta_j(x)$ ($k,j=1,\dots,K$) prie sąlygos, kad yra žinomos x reikšmės, nuo tų reikšmių nepriklauso. Todėl modelis (3) dar vadinamas *proporcingųjų šansių* modeliu. Praktikoje proporcingųjų šansų prielaida dažnai nėra išpildyta (ši prielaida nepagrįsta ir mūsų tiriamiems duomenims). Natūralus modelio (3) apibendrinimas

$$\operatorname{logit}\bigl(\Pi_k(x)\bigr) = \alpha_k + \beta_k' x, \quad k=1,\dots,K-1, \tag{4}$$

įvedant parametrų β priklausomybę nuo lygmens k, matematiškai yra nekorektiškas, nes kai kurioms $x \in \mathbf{R}^m$ reikšmėms pažeidžiama sutvarkymo sąlyga (2). Tačiau tam tikrais atvejais, pavyzdžiui, kai x komponentės yra *binariniai* kintamieji, modelis (4) tenkina sutvarkymo sąlygą (2) [6].

Nežiūrint minėtų matematinių problemų modelis (4) plačiai taikomas praktikoje. Tikimasi, kad realioms kintamojo x reikšmėms x_i sutvarkymo sąlyga galios. Tuo atveju, kai ji vis dėlto yra neišpildyta,siūloma *ignoruoti* ranginę kintamojo Y_i prigimtį ir naudoti *daugialygį logit* modelį, skirtą *kategorinių* kintamųjų skirstinio modeliavimui:

$$\log\bigl(\pi_k(x_i)/\pi_K(x_i)\bigr) = \alpha_k + \beta_k' x_i, \quad k=1,\dots,K-1. \tag{5}$$

Čia apibrėžiant apibendrintąją logit funkciją didžiausias atsako kintamojo Y_i lygis K laikomas *bazine* arba *atskaitos* (*reference*) reikšme [1, 6].

Pastaruoju metu pradėti taikyti (ir vis plačiau taikomi) *ranginiai apibendrintieji tiesiniai* (*mišrieji*) modeliai [3, 6]. Juose įvedamas papildomas mastelio parametras, kuris gali priklausyti ir nuo lygio k, ir nuo aiškinančiųjų kintamųjų bei jų sąveikų. Tai ženkliai padidina modelio lankstumą ir interpretavimo galimybes. Ranginis apibendrintasis tiesinis modelis (be atsitiktinių faktorių) bendru atveju užsirašo taip:

$$\operatorname{logit}\bigl(\Pi_k(x)\bigr) = \exp\{a_k + b_k' x\}(\alpha_k + \beta_k' x), \quad k = 1, \ldots, K-1.$$

Šis modelis dar vadinamas *nehomogeniniu pasirinkimo modeliu* [3]. Vėlgi reikia pabrėžti, kad sutvarkymo sąlyga (2) bendru atveju jam gali būti neišpildyta. Tačiau ji galioja, kai mastelio funkcijos parametrai a_k ir b_k bei β_k nuo k nepriklauso. Tuomet gauname 2-*jo tipo IRT modelio* variantą [6]

$$\operatorname{logit}\bigl(\Pi_k(x)\bigr) = \exp\{-b' x\}(\alpha_k + \beta' x). \tag{6}$$

Mastelio parametrai (6) modelyje leidžia aprašyti vertinimo skalės *nehomogeniškumą*, jos skirtingą jautrumą prie įvairių sąlygų, o tuo pačiu dalinai išspręsti proporcingųjų šansų modelio netinkamumo problemą nepažeidžiant sutvarkymo sąlygos (2). Šiame darbe (6) yra pagrindinis modelis.

2 Uždavinio formulavimas ir tiriamoji analizė

Studentų pažangumą nusako daugelio dalykų pažymių visuma. Vienas iš pažangumo rodiklių yra sesijos pažymių vidurkis. Daugelio skirtingų dalykų pažymių vidurkiai nėra tinkamas tolesnių studijų rezultatų indikatorius, ypač lyginant skirtingų specialybių studentus. Tačiau daugelyje specialybių pirmajame kurse yra *baziniai kursai*, kurių rezultatai didžiąja dalimi nulemia tolimesnius studijų rezultatus. VU Matematikos ir informatikos fakultete (MIF) tokiu baziniu kursu galima laikyti matematinę analizę. Studentas, gavęs iš matematinės analizės pažymį mažesnį nei 7, turi apie 7 kartus mažesnes galimybes (šansus) sėkmingai baigti studijas, negu studentas, gavęs pažymį 7 ir daugiau (vieno iš autorių atliktas tyrimas naudojant 2003–2005 m. įstojusiųjų duomenis). Ta prasme kitos galimos alternatyvos, pavyzdžiui, algebros ir geometrijos pažymys, yra blogesnis tolimesnių studijų rezultatų indikatorius. Todėl tikimasi, kad matematinės analizės pažymio modeliavimas padės suprasti faktorius ir priežastis, nulemiančias studijų eigą.

Duomenys. Darbe panaudoti VU informacinės sistemos duomenys apie įstojusius 2003–2006 m. į MIF bakalauro studijų programas. Tyrimo kintamieji: ATESTVID – atestato vidurkis; IVERT – dalyko pažymys; LYTIS – studento lytis; DTNR – dėstytojo indentifikatorius.

Norėdami įvertinti, kaip pirmųjų sesijų matematinės analizės rezultatai yra susiję su studento lytimi, mokykliniais pasiekimais bei egzaminuojančių dėstytojų naudojama vertinimo sistema apsiribojome tik trijų „matematinių“ specialybių duomenimis (viso 603 studentai). Priežastis – duomenys yra daug homogeniškesni matematinės analizės kurso programos ir egzaminuojančių dėstytojų atžvilgiu (taigi, galima eliminuoti programos ir pan. įtaką).

Studentų atestato vidurkis nėra vien tik mokyklinių pažymių vidurkis, tai daugiau santykinis mokyklinių pasiekimų rodiklis. Įstojusiųjų į minėtas specialybes atestato vidurkių sklaidos intervalas yra nuo 6 iki 11, tačiau tik 1% (6-ių studentų) atestato vidurkis nedidesnis už 8. Ranginės logistinės regresijos modeliai yra labai jautrūs (nestabilūs), kai didelė dalis analizuojamų kintamųjų reikšmių kombinacijų duomenyse nė karto nepasitaikė arba buvo labai retos. Todėl imtyje palikome tik stebinius, kuriuose

	IVERT							ATESTVID	
	Lytis		DTNR		ATEST_CAT			Lytis	
	V	M	1	2	8–9	9–10	>10	V	M
Mean	6,59	6,29	6,16	6,71	6,58	6,06	7,00	9,7333	9,9763
N	268	335	310	293	43	305	255	268	335
Std. Dev	1,670	1,659	1,489	1,801	1,749	1,536	1,630	0,5569	0,3827

atestato vidurkis yra didesnis už 8, ir atestato vidurkį kategorizavome: "$8-9$" $\to$ "3", "$9-10$" $\to$ "4", ">10" $\to$ "5" (kintamasis $ATEST\_CAT$). Be to, buvo transformuotas ir atsako kintamasis Y_i. Neigiami pažymiai nuo 1 iki 4 ne visada yra susiję su žinių įvertinimu, todėl neigiamų pažymių kategorijas apjungėm į vieną kategoriją 4 (kintamasis $IVERT\_TRUMP$).

Pradžioje pateiksime ir aptarsime svarbiausias aprašomąsias nagrinėjamų kintamųjų charakteristikas, „pateisinančias" logistinės regresijos modelio pasirinkimą. Aukščiau lentelėje pateikiami skirtingų grupių atestatų ir matematinės analizės pažymių vidurkiai ir standartiniai nuokrypiai.

Vaikinų pažymių vidurkis yra statistiškai reikšmingai didesnis (atestato vidurkis – priešingai, yra didesnis merginų); skiriasi ir egzaminuojančių dėstytojų vidutiniai vertinimai. Tarp atestato vidurkio ir matematinės analizės įvertinimo yra nestipri, bet statistiškai reikšminga koreliacija (Spirmeno koreliacijos koeficientas 0.385, $p < 0,01$) – studentai, turintys „geresnius" atestato vidurkius, gauna geresnius ir matematinės analizės pažymius.

3 Parinktų modelių aptarimas

VU matematinių specialybių studentų pažangumo duomenims sudarytas ranginės logistinės regresijos modelis (3) ir jo apibendrinimas su mastelio parametrais (6). „Simboliškai" su $x_i = (LYTIS_i * ATEST\_CAT_i * DTNR_i)$, jie užsirašo taip:

$$\mathrm{logit}(\Pi_k(x_i)) = \alpha_k - (\gamma_1 LYTIS_i * ATEST\_CAT_i * DTNR_i) \quad (k = 4, \dots, 9), \tag{7}$$

$$\mathrm{logit}(\Pi_k(x_i)) = \frac{\alpha_k - (\gamma_1 LYTIS_i * ATEST\_CAT_i * DTNR_i)}{\exp(\tau_1 LYTIS_i * ATEST\_CAT_i * DTNR_i)}. \tag{8}$$

Buvo bandyta pritaikyti proporcingųjų šansų logistinės regresijos modelį. Proporcingųjų šansų prielaida nėra išpildyta tiek pateiktam modeliui, tiek ir įtraukus įvairius aiškinančiųjų kintamųjų rinkinius ir sąveikas. Įvedus mastelio kintamuosius (modelis (8)) gauname adekvatų tiriamų duomenų modelį, apie ką liudija apibendrintųjų tiesinių modelių tinkamumo matai, tikėtinumo ir Pearsono deviacijos. Žemiau dviejose lentelėse pateikiama kintamajam $IVERT\_TRUMP$ parinkto daugialygio logistinės regresijos modelio (6) tinkamumo matų ir statistiškai reikšmingų faktorių standartinė PLUM procedūros išvedimo forma[1]

Parametrų įverčių lentelė rodo, kad geresnio pažymio galimybės yra didesnės, turintiems „geresnius" atestatus, tačiau yra akivaizdi dėstytojų egzaminavimo sistemos įtaka bei aiškus skirtumas tarp merginų ir vaikinų rezultatų. Pavyzdžiui, šansų santykis, kad vaikinas su atestato vidurkiu '„8–9" egzaminuojamas pirmojo dėstytojo gaus k-tąjį ar mažesnį pažymį yra maždaug 9 ($e^{2,247} = 9,459$) kartus didesnis lyginant su mergina, kurios atestato vidurkis „> 10" ir ji egzaminuojama antrojo dėstytojo.

[1] Lentelėje praleisti kelių statistiškai nereikšmingų mastelio parametrų įverčiai.

Parameter Estimates

		Estimate	Std. Error	Wald	df	Sig.
Threshold	[ivert_trump = 4.00]	-3.809	.522	53.189	1	.000
	[ivert_trump = 5.00]	-1.933	.312	38.308	1	.000
	[ivert_trump = 6.00]	-1.105	.250	19.533	1	.000
	[ivert_trump = 7.00]	-.147	.219	.447	1	.504
	[ivert_trump = 8.00]	.844	.244	12.001	1	.001
	[ivert_trump = 9.00]	1.510	.288	27.512	1	.000
Location	[ATEST_CAT=3] * [LYTIS=1] * [DTNR=1]	-2.247	.752	8.921	1	.003
	[ATEST_CAT=3] * [LYTIS=1] * [DTNR=2]	-2.027	.480	17.863	1	.000
	[ATEST_CAT=3] * [LYTIS=2] * [DTNR=1]	-3.349	1.009	11.010	1	.001
	[ATEST_CAT=4] * [LYTIS=1] * [DTNR=1]	-1.563	.339	21.242	1	.000
	[ATEST_CAT=4] * [LYTIS=1] * [DTNR=2]	-.956	.317	9.088	1	.003
	[ATEST_CAT=4] * [LYTIS=2] * [DTNR=1]	-2.240	.369	36.806	1	.000
	[ATEST_CAT=4] * [LYTIS=2] * [DTNR=2]	-1.697	.331	26.344	1	.000
	[ATEST_CAT=5] * [LYTIS=1] * [DTNR=1]	-.827	.304	7.372	1	.007
	[ATEST_CAT=5] * [LYTIS=1] * [DTNR=2]	-.026	.302	.007	1	.932
	[ATEST_CAT=5] * [LYTIS=2] * [DTNR=1]	-1.171	.290	16.260	1	.000
	[ATEST_CAT=5] * [LYTIS=2] * [DTNR=2]	0^a	.	.	0	.
Scale	[ATEST_CAT=3] * [LYTIS=1] * [DTNR=1]	.409	.288	2.017	1	.156
	[ATEST_CAT=4] * [LYTIS=2] * [DTNR=1]	-.371	.169	4.797	1	.029
	[ATEST_CAT=5] * [LYTIS=1] * [DTNR=2]	-.318	.190	2.796	1	.095
	[ATEST_CAT=5] * [LYTIS=2] * [DTNR=1]	-.212	.152	1.943	1	.163
	[ATEST_CAT=5] * [LYTIS=2] * [DTNR=2]	0^a	.	.	0	.

IVERT

LYTIS	Mean	N	Std. Deviation
V	6,59	268	1,670
M	6,29	335	1,659
DTNR			
1	6,16	310	1,489
2	6,71	293	1,801
ATEST_CAT			
8-9	5,58	43	1,749
9-10	6,06	305	1,536
>10	7,00	255	1,630

ATESTVID

LYTIS	Mean	N	Std. Deviation
V	9,7333	268	,55692
M	9,9763	335	,38268

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	33.272	40	.765
Deviance	36.209	40	.642

Link function: Logit.

Pseudo R-Square

Cox and Snell	.211
Nagelkerke	.217
McFadden	.067

Link function: Logit.

Čia pateikiame tik vieno iš modelių rezultatų lenteles. Ir klasikinis ranginis logistinės regresijos modelis (3), ir jo apibendrinimas (6) buvo taikytas tiek keičiant aiškinančiųjų kintamųjų kategorizaciją ir rinkinius tiek imant atskirų specialybių (ar jų grupių) studentų duomenis. Modelių tinkamumo matai, proporcingumo hipotezės tinkamumas keitėsi, kito modelių parametrų įverčiai, tačiau aiškinamųjų kintamųjų įtakos tendencijos liko tos pačios.

Išvados

Tyrime nustatytas statistiškai reikšmingas ryšys tarp matematinės analizės pažymio ir mokyklinių pasiekimų. Tačiau ši įtaka skirtinga merginoms ir vaikinams. Netiesiogiai tai parodo, kad atestato vidurkis nėra adekvatus matematinių gabumų indikatorius.

Studentų pasiekimų vertinime dėstytojo įtaka yra žymi. Atlikta analizė leidžia teigti, kad yra labai aktualu labiau unifikuoti studentų vertinimo taisykles, kad vertinime būtų kuo mažiau subjektyvumo.

Tiriamiems duomenims hipotezė apie galimybių proporcingumą buvo atmesta. Ranginis apibendrintasis tiesinis modelis, kuriame modelio parametrai nepriklauso nuo atsako kintamojo lygmenų, nepažeidžia ranginio kintamojo sutvarkymo sąlyga ir leidžia atspindėti skirtingą atsako kintamojo variaciją grupėse su skirtingomis aiškinančiųjų kintamųjų kombinacijomis.

Ranginės logistinės regresijos modeliai yra nestabilūs, kai aiškinančiųjų kintamųjų kombinacijų dažniai kai kuriems atsako kintamojo lygmenims yra maži. Tai susiję su tuo, kad duomenims su nuliniais kombinacijų dažniais gali būti nepakankamas statistinėse išvadose naudojamų aproksimacijų tikslumas ir didžiausio tikėtinumo įvertinys bendru atveju aplamai gali neegzistuoti. Todėl sudarant ranginės logistinės regresijos modelius patartina ieškoti balanso tarp modelio laisvės laipsnių skaičiaus ir modelio „tikslumo".

Toliau tobulinant modelius reiktų atsižvelgti ir į kitus galimus įtakojančius faktorius, pvz., studento motyvacija, jo baigta mokykla ir pan.

Literatūra

[1] A. Agresti. *Categorical Data Analysis*. Wiley-Interscience, New York, 2002.

[2] J. Čikun, J.-V. Daunoravičius, M. Radavičius. Pažangumo rodiklių tarpusavio ryšių analizė. *Lietuvos mat. rink.*, **46**(spec. nr.):138–143, 2006.

[3] A. Fielding, M. Yang and H. Goldstein. Multilevel ordinal models for examination grades. *Statistical Modeling*, **3**:127–153, 2003. Available from Internet: http://www.cmm.bristol.ac.uk/research/momeg.pdf.

[4] P. McCullagh. Regression models for ordinal data (with discussion). *Journal of the Royal Statistical Society*, **Series B, 42**:109–142, 1980.

[5] M.J. Norusis. *PASW Statistics 18 Advanced Statistical Procedures*. Prentice Hall, Upper Saddle River, New Jersey, 2010.

[6] K. Sijstma and B.G. Junker. Item response theory: past performance, present developments, and expectations. *Behaviourmetrika*, **33**(1):75–102, 2006.

SUMMARY

Analysis of student achievements using ordered logistic regression
G. Murauskas, M. Radavičius
In the paper the problem of assessment of mathematical abilities is addressed. It is assumed that student's grade in the mathematical analysis is one of the main indices of mathematical abilities which gives rather reliable guess of his further academic performance in mathematics. Relationships of the mathematical analysis grades with other factors including the examinator's effect are investigated using data of Vilnius University students in mathematics and rank logistic regression with a variable scale.

Keywords: achievement analysis, interactions, mathematical abilities, ordered logistic regression, proportional odds