

Analiză

Simularea Evaluării Naționale și Evaluarea inițială cu testele Brio

Județul Iași



Iliescu, D. & Iancu, D. E. (2024)

Simularea Evaluării Naționale și Evaluarea inițială cu testele Brio. Studiu de caz pe Județul Iași. București: Brio

Cuprins

PREAMBUL.....	2
Întrebări de cercetare	2
Cum și ce s-a măsurat?.....	2
Participanți	3
Analiza datelor.....	3
REZULTATE	4
1. Distribuții de frecvență și comparații ale seturilor de date	4
Matematică	4
Limba și Literatura Română	4
2. Analiza corelațională și de predicție.....	5
Matematică	6
Limba și Literatura Română	7
3. Analiza diferențelor între scoruri la nivel de elevi	8
Diferență sub un punct	9
Câștiguri și pierderi mai mari de 1 punct.....	10
Praguri: Nota 4 la testul Brio în raport cu nota 5 la simularea EN.....	11
CONCLUZII	12
Rezultate principale și posibile interpretări	12
Limitări și direcții viitoare de analiză	13

Preambul

Analiza prezentată aici a investigat relația și diferențele dintre rezultatele obținute de elevii de clasa a 8-a din județul Iași la simularea Evaluării Naționale (EN) și la evaluarea inițială derulată la nivelul județului cu testele Brio. Mai specific, s-a dorit investigarea măsurii în care rezultatele obținute la simularea EN pot fi prezise pe baza scorurilor Brio obținute la Evaluarea Inițială cu 4 luni înainte, atât la Matematică, cât și la Limba și Literatura Română.

Studiul a fost realizat la solicitarea ISJ Iași care a pus la dispoziția Brio rezultatele elevilor de clasa a VIII-a la simularea EN. Aceste rezultate au fost transmise anonimizat, în baza codurilor unice ale tuturor elevilor (diferite de codurile din SIIR) generate de ISJ Iași pentru derularea evaluării inițiale din luna septembrie. Astfel, Brio nu deține date despre numele sau alte caracteristici ale elevilor care ar putea ajuta la identificarea acestora.

Întrebări de cercetare

1. În ce măsură distribuția rezultatelor obținute de elevi la simularea EN este diferită de distribuția rezultatelor obținute la evaluarea inițială derulată cu testele Brio?
2. Putem anticipa rezultatele elevilor la simularea EN pornind de la rezultatele obținute de către aceștia în urma participării la evaluarea inițială derulată cu testele Brio?
3. Care sunt și ce ne spun diferențele la nivel individual dintre notele obținute de elevi la probele simulării EN și la testele Brio?

Cum și ce s-a măsurat?

Examenul de simulare a Evaluării Naționale a fost administrat creion-hârtie în zilele de 5 și 6 februarie 2024 la Matematică, respectiv Limba și Literatura Română. Itemii din care a fost alcătuit testul administrat au fost de mai multe feluri, de la itemi de tip adevărat/fals, până la itemi care solicitau un răspuns construit. Examenul de simulare EN a vizat întreg spectrul de competențe dobândite în ciclul gimnazial la cele două discipline de studiu, iar fiecare probă a durat 3 ore.

Testele Brio au fost administrate digital, prin intermediul unui dispozitiv electronic (telefon mobil, tabletă sau computer/laptop) și au fost alcătuite din itemi cu 4 variante predefinite de răspuns dintre care doar una corectă. Derularea evaluării inițiale la nivelul județului Iași a avut loc în luna septembrie, testele de Matematică și de Limba și Literatura Română fiind aplicate în zilele de 26, respectiv 27 septembrie. Testele Brio pentru cele două discipline au vizat strict competențele dobândite în clasa a VII-a și au avut o durată de 45 de minute.

Participanți

Din numărul total de 7871 elevi de clasa a VIII-a din Județul Iași (în anul școlar 2023-2024), numărul celor care au participat atât la simularea EN cât și la evaluarea inițială cu testele Brio este de:

- 6876 elevi pentru Matematică (87.35 % din total);
- 6886 elevi pentru Limba și Literatura Română (87.48 % din total).

Analiza datelor

Notele obținute la simularea EN au fost cuprinse între 1 și 10, în timp ce scorurile Brio au fost cuprinse între 0 și 100. În scopul comparării echivalente a distribuțiilor, scorurile Brio au fost transformate astfel încât să fie încadrate în același interval (1-10) ca notele de la simulare. Această transformare liniară a avut la bază formula $1+(x/100)*9$ și nu a afectat în nici un fel restul analizelor derulate.

Pentru a răspunde la întrebările de cercetare prezentate mai sus, au fost calculate statistici descriptive (media, abaterea standard), teste t pentru eșantioane pereche și regresii liniare simple. Suplimentar, au mai fost analizate, la nivelul fiecărei discipline, diferențele dintre scorurile obținute la simulare și la evaluarea inițială de către aceiași elevi. Pentru vizualizarea datelor și a rezultatelor obținute în urma analizelor, au fost realizate grafice de tip *density plots* și *scatter plots* și grafice cu bare verticale.

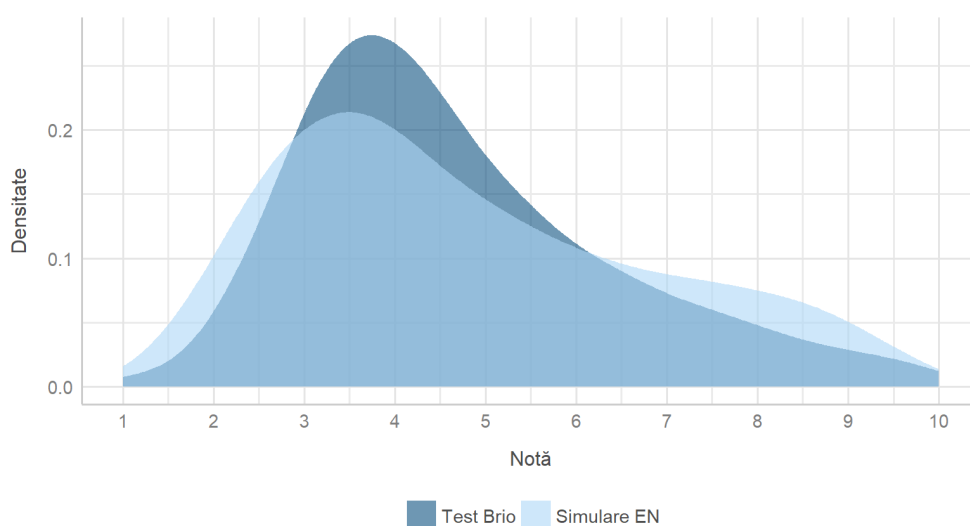
Rezultate

1. Distribuții de frecvență și comparații ale seturilor de date

Matematică

Cele două variabile legate de competențele matematice, nota la testul Brio și nota la Simularea EN, au distribuții relativ similare. În urma ambelor evaluări notele se concentrează în jurul valorii de 4. Diferența dintre valorile mediane ale celor două seturi de date este de 0.07, acestea fiind 4.33 în cazul testării Brio și 4.40 în cazul simulării EN.

Suprapunerea distribuției notelor obținute - Matematică

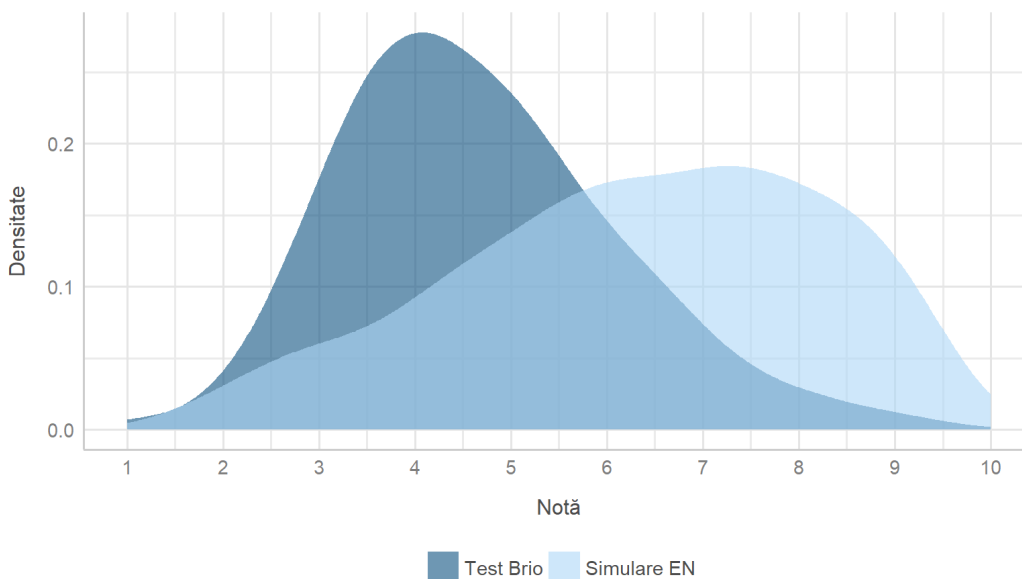


Pentru a analiza statistic diferența dintre cele două seturi de rezultate a fost derulată tehnica testelor t pentru eșantioane pereche. S-a obținut o valoare a lui t de 7.73 și o mărime a efectului măsurată prin d -ul lui Cohen de 0.09, ceea ce ne arată că diferența care există între valorile medii ale celor două seturi de date este una foarte scăzută. Rezultatele pot sugera că nivelul de performanță la matematică al elevilor a fost măsurat cu un nivel de dificultate relativ similar de către cele două teste.

Limba și Literatura Română

Cele două variabile legate de competențele lingvistice, nota la testul Brio și nota la Simularea EN, au distribuții semnificativ diferite. În urma Simulării EN notele sunt semnificativ mai mari și se concentrează în jurul valorii de 7, pe când în cazul testului Brio acestea se concentrează în jurul valorii de 4. Diferența dintre valorile mediane ale celor două seturi de date este de 1.99, acestea fiind 4.51 în cazul testării Brio și 6.50 în cazul simulării EN.

Suprapunerea distribuției notelor obținute - Limba și Literatura Română



Și în cazul acestei discipline, pentru a analiza statistic diferența dintre cele două seturi de rezultate a fost derulată tehnica testelor t pentru eșantioane pereche. De data aceasta s-a obținut o valoare a lui t de 91.55 și o mărime a efectului măsurată prin d -ul lui Cohen de 1.19, ceea ce ne arată că diferența care există între valorile medii ale celor două seturi de date este una foarte ridicată, de peste o abatere standard. Acest rezultat sugerează că proba de competențe lingvistice de la simularea EN a avut un grad de dificultate mai scăzut decât testul Brio. Vedem asta din centrarea distribuției în jurul notei de 7, ceea ce arată că s-au primit multe note mari și relativ puține note mici.

2. Analiza corelațională și de predicție

Analiza de corelație măsoară gradul și direcția unei relații între două variabile continue. Coeficientul de corelație variază între -1 și 1. O valoare apropiată de 1 indică o relație pozitivă puternică (când o variabilă crește, cealaltă de asemenea crește), o valoare apropiată de -1 indică o relație negativă puternică (când o variabilă crește, cealaltă scade), iar o valoare apropiată de 0 sugerează că nu există nicio relație liniară.

Analiza de regresie liniară este folosită pentru a prezice valoarea unei variabile, bazându-se pe valoarea altei variabile aflată în relație cu cea dintâi. Modelul de regresie liniară construiește o ecuație liniară (de forma $Y = b + a \cdot X$), unde Y este variabila dependentă (ceea ce dorim să prezicem), X este variabila independentă (predictorul), iar ' a ' (panta de regresie) și ' b ' (interceptul) sunt coeficienții modelului determinați pe baza datelor. Aceasta ecuație ne permite să prezicem valoarea lui Y pentru orice valoare dată a lui X .

Matematică

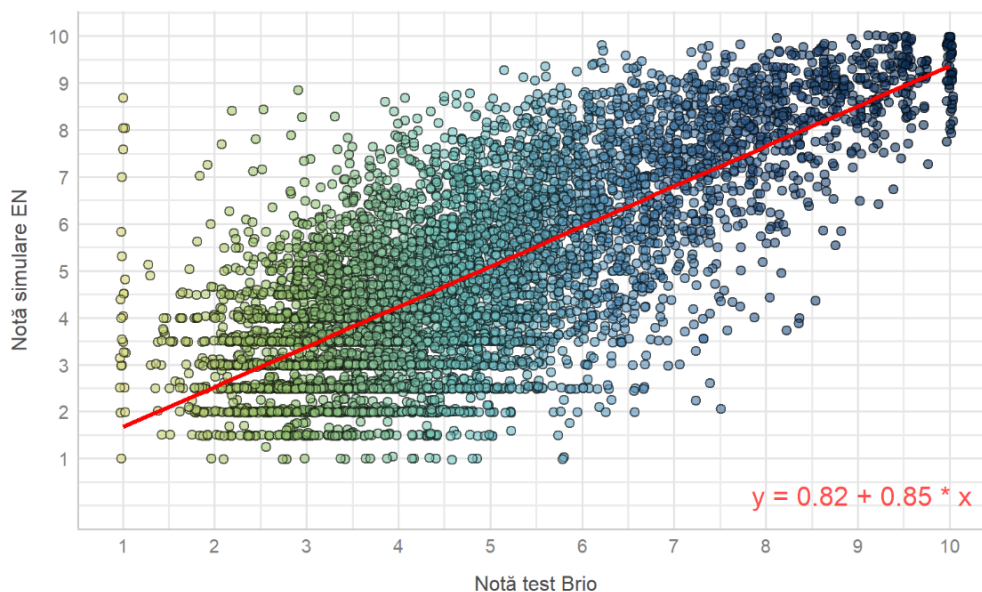
Pentru analiza relației dintre rezultatele absolvenților de clasa a VIII-a la simularea Evaluării Naționale și la testele Brio la disciplina Matematică, a fost calculat coeficientul de corelație Pearson. De asemenea, datorită distribuției ușor asimetrice a datelor a fost calculat și coeficientul de corelație Spearman (a ordinii datelor), care surprinde mai bine relația dintre două seturi de date care nu sunt distribuite perfect simetric.

Valoarea coeficientului de corelație Pearson calculat pe rezultatele a 6876 de participanți a fost de $r = 0.73$, ceea ce indică o asociere puternică între rezultatele celor două evaluări. Acest rezultat sugerează că elevii cu scoruri mai ridicate la testul Brio din septembrie 2023 au obținut în medie scoruri mai ridicate și la simularea EN din februarie 2024. Calculând și coeficientul de corelație Spearman, a fost obținută o valoare $\rho = 0.67$, care la rândul ei reprezintă o asociere puternică între cele două seturi de date.

În urma derulării analizei de regresie a fost obținută ecuația $Y = 0.82 + 0.85 * X$. Această ecuație de regresie estimează o notă medie de 0.82 la simularea EN atunci când scorul la testul Brio este 0 și un efect mediu al scorului la testul Brio de 0.85. Rezultatul sugerează că pentru fiecare creștere unitară a scorului la testarea BRIO, se așteaptă o creștere medie a notei la simularea EN de 0.85.

De exemplu, dacă un elev obține exact nota 8 la testul Brio, putem anticipa că nota pe care o va obține la simularea EN va fi în jur de 7.62. Dacă obține exact nota 4, ne vom aștepta ca nota pe care o va obține la simularea EN să fie în jurul valorii de 4.22.

Grafic corelație - Matematica



Notă: Fiecare punct de pe grafic reprezintă notele unui elev la cele două evaluări (simularea EN și testul Brio). Linia roșie reprezintă dreapta de regresie liniară utilizată pentru a descrie relația dintre cele două variabile. În partea din dreapta jos este prezentată și ecuația de regresie obținută în urma analizei datelor.

Limba și Literatura Română

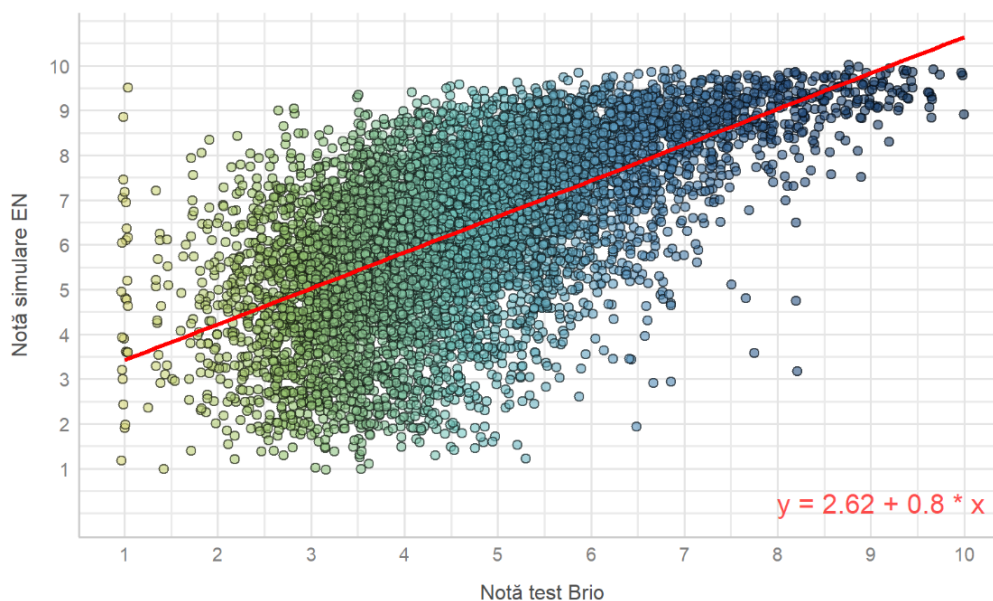
Pentru analiza relației dintre rezultatele absolvenților de clasa a VIII-a la simularea Evaluării Naționale și la testele Brio la disciplina Limba și Literatura Română, au fost calculate de asemenea coeficientul de corelație Pearson, respectiv Spearman, din aceleași considerente de distribuție ușor asimetrică a datelor.

Valoarea coeficientului de corelație Pearson calculat pe rezultatele a 6886 de participanți a fost de $r = 0.60$, ceea ce indică o asociere substanțială între rezultatele celor două evaluări. Ca în cazul matematicii, elevii cu scoruri mai ridicate la testul Brio de Lb. Română din septembrie 2023 au obținut în medie scoruri mai ridicate și la proba de Lb. Română din cadrul simulării EN din februarie 2024. Coeficientul de corelație Spearman de $\rho = 0.63$ confirmă la rândul său relația identificată între cele două seturi de date.

În urma derulării analizei de regresie liniară a fost obținută ecuația $Y = 2.62 + 0.80 * X$. Această ecuație de regresie estimează o notă medie de 2.62 la simularea EN atunci când scorul la testul Brio este 0 și un efect mediu al scorului la testul Brio de 0.80. Astfel, pentru fiecare creștere unitară a scorului la testarea BRIO, se așteaptă o creștere medie a notei la simularea EN de 0.80.

De exemplu, dacă în cazul acestei discipline un elev obține exact nota 7 la testul Brio, putem anticipa că nota pe care o va lua la simularea EN va fi în jur de 8.22. Dacă în schimb ar obține exact nota 4, ne vom aștepta ca nota de la simularea EN să fie în jurul valorii de 5.82.

Grafic corelație - Limba și Literatura Română



Notă: Fiecare punct de pe grafic reprezintă notele unui elev la cele două evaluări (simularea EN și testul Brio). Linia roșie reprezintă dreapta de regresie liniară utilizată pentru a descrie relația dintre cele două variabile. În partea din dreapta jos este prezentată și ecuația de regresie obținută în urma analizei datelor.

3. Analiza diferențelor între scoruri la nivel de elevi

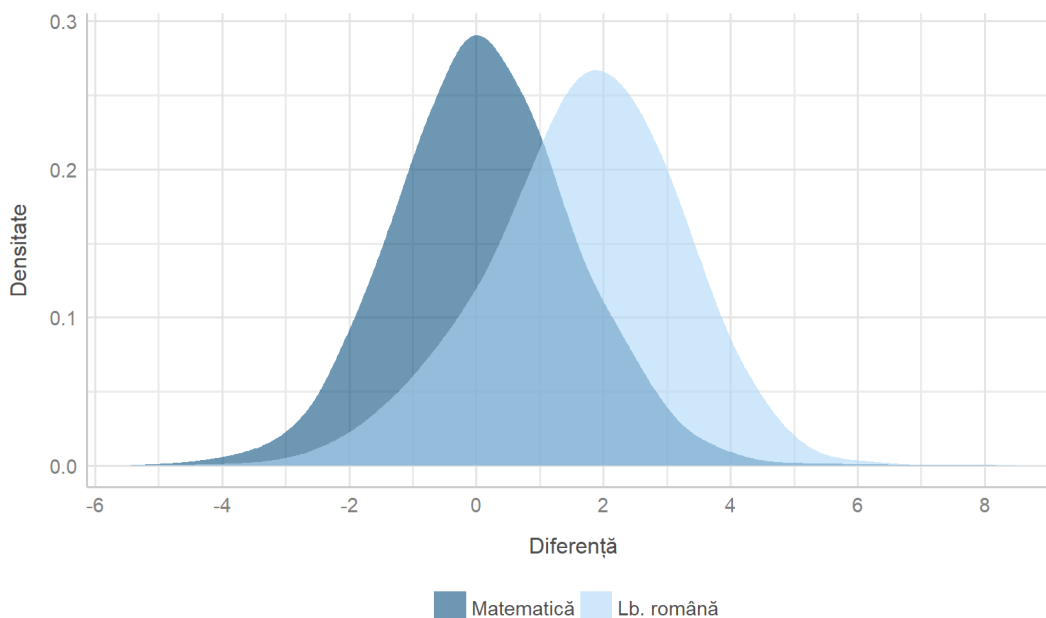
Pentru calcularea diferențelor individuale dintre rezultatele obținute la simularea EN și la testele Brio, din cele dintâi s-a scăzut valoarea celor din urmă. De exemplu, dacă un elev a obținut nota 8 la proba de la disciplina Matematică din cadrul simulării EN și nota 6 la testul Brio de Matematică, diferența dintre rezultatele obținute este de 2 (adică 8 minus 6). Dacă același elev ar fi obținut la testul Brio nota 10, diferența ar fi fost de -2 (adică 8 minus 10).

Astfel, s-a obținut o variabilă nouă, denumită *scoruri-diferență*. Distribuția valorilor acestei variabile oferă o imagine de ansamblu asupra amplitudinii diferențelor și asupra proporției de elevi care au avut note diferite la simularea EN față de evaluarea inițială. Tabelul următor prezintă statisticile descriptive ale diferenței de scor în rezultatele elevilor pentru ambele discipline la care au fost evaluați. De asemenea, graficul de sub tabel prezintă distribuția celor două seturi de diferențe.

Tabel 1. Statistici descriptive ale diferențelor dintre scoruri

<i>Disciplină</i>	<i>Medie diferențe</i>	<i>Mediană diferențe</i>	<i>Diferență minimă</i>	<i>Diferență maximă</i>
Matematică	0.13	0.10	-5.43	8.50
Lb. Română	1.69	1.78	-5.00	7.70

Suprapunerea distribuției diferențelor între notele obținute la cele două evaluări



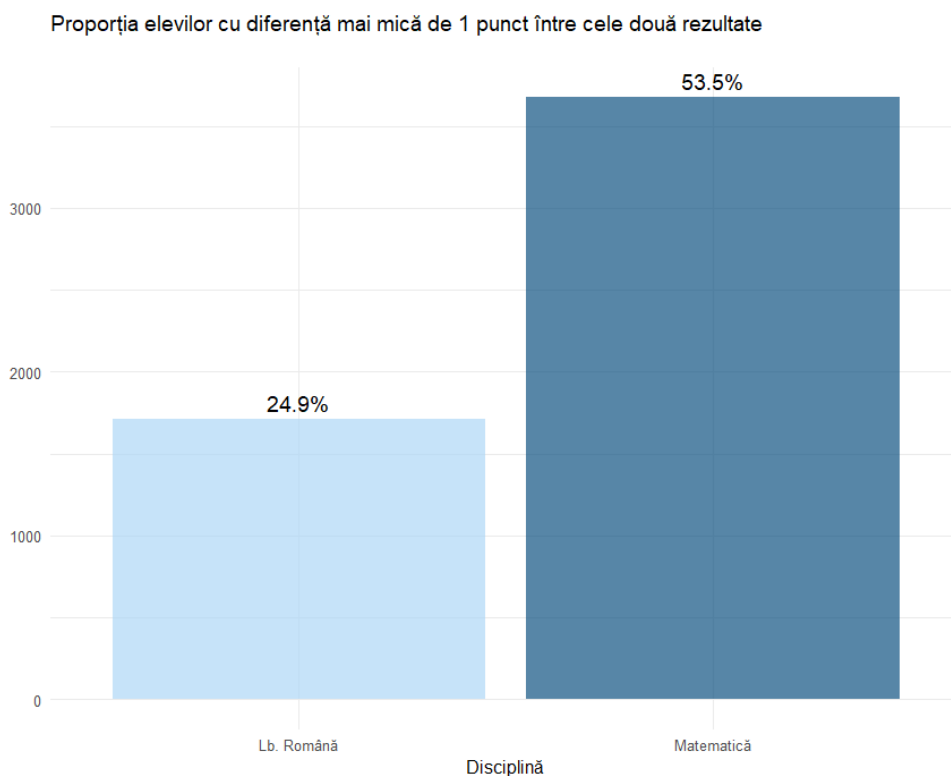
Notă: Fiecare curbă de pe grafic reprezintă distribuția diferențelor între rezultate obținute de elevi la cele două evaluări, pentru fiecare disciplină. Pe axa X ("Diferență"), este reprezentată diferența în rezultatele obținute la cele două evaluări. Aceste diferențe pot fi câștiguri sau pierderi în scor de la o evaluare la alta. O diferență de scor de zero ar însemna că rezultatele obținute în urma celor două evaluări sunt perfect egale.

Deși plaja diferențelor este una destul de largă în rândul ambelor discipline, distribuția acestora este una relativ simetrică în ambele cazuri. Totuși, în cazul disciplinei Lb. Română există mai multe diferențe pozitive, vârful distribuției și implicit majoritatea diferențelor situându-se în jurul valorii de 2. Aceste rezultate ne spun că la disciplina Matematică majoritatea elevilor au avut rezultate similare la cele două evaluări, în timp ce la disciplina Lb. Română o mare parte din elevi au luat note mai mari la această probă din cadrul simulării EN decât la evaluarea inițială.

Diferență sub un punct

A fost calculată și proporția elevilor care au avut o diferență mai mică de un punct între rezultatele obținute la cele două evaluări, pe fiecare materie în parte. După cum se poate observa în graficul de mai jos, în cazul matematicii peste jumătate din elevi, adică 3676, au avut mai puțin de un punct diferență între cele două rezultate, în timp ce la disciplina Lb. Română doar 1711, aproximativ un sfert, dintre elevi au rămas în același ecart de 1 punct al diferenței.

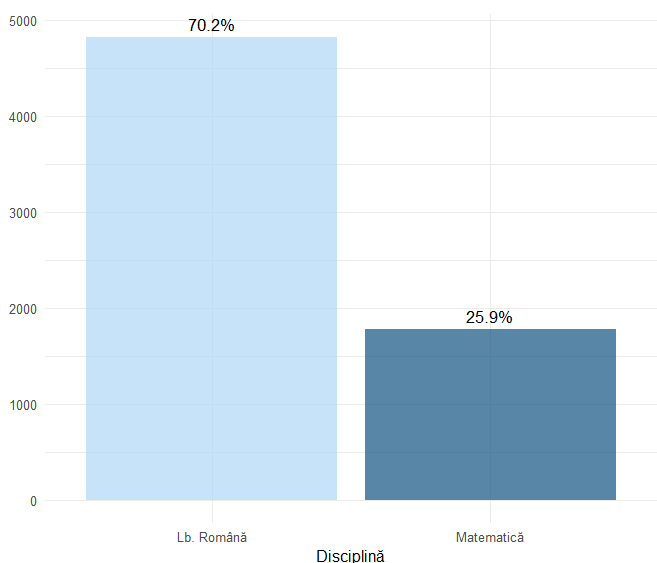
Aceste rezultate ne arată că la Matematică rezultatele la simularea EN au fost mult mai apropiate de cele obținute în urma testelor Brio, în comparație cu notele de la Limba și Literatura Română, ceea ce sugerează fie că testele măsoară cu precizie diferită, fie că nivelul de competență lingvistică al elevilor s-a îmbunătățit (în cele 4 luni dintre teste) mai mult la Lb. Română.



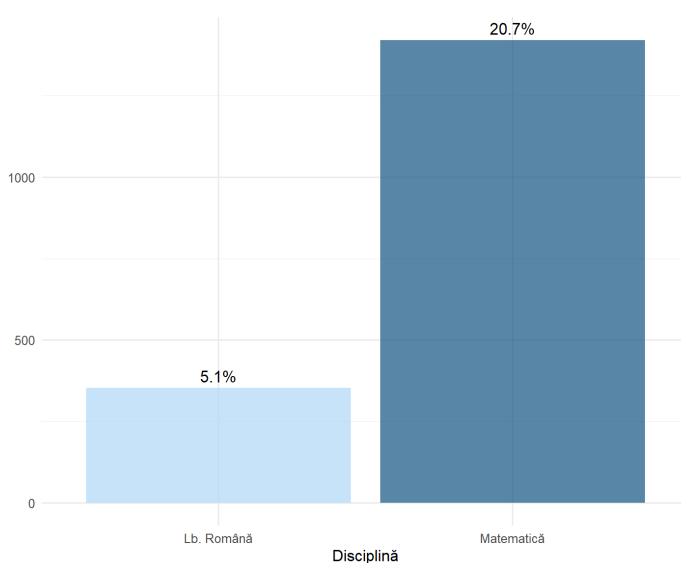
Câștiguri și pierderi mai mari de 1 punct

Numărul elevilor care au obținut o notă mai mare cu cel puțin un punct la simularea EN este de 4823 pentru Lb. Română și de 1780 pentru Matematică. O largă majoritate a elevilor, 70.2%, au luat o note mai mari cu 1 punct la proba EN de Lb. Română, în timp ce la proba de Matematică, doar aproximativ un sfert dintre elevi, 24.9%, au reușit să facă acest lucru.

Proporția elevilor cu câștiguri mai mari de 1 punct la simularea EN



Proporția elevilor cu pierderi mai mari de 1 punct la simularea EN



Pe de altă parte, numărul elevilor care au obținut o notă mai mică cu cel puțin un punct la simularea EN a fost de 352 pentru Lb. Română și de 1420 pentru Matematică. O mică parte dintre elevi, 5,1%, au avut o scădere de mai mult de un punct la rezultatele pentru Limba Română, în timp ce pentru Matematică, o porțiune semnificativă de 20,7% a obținut o notă mai mică cu mai mult de un punct.

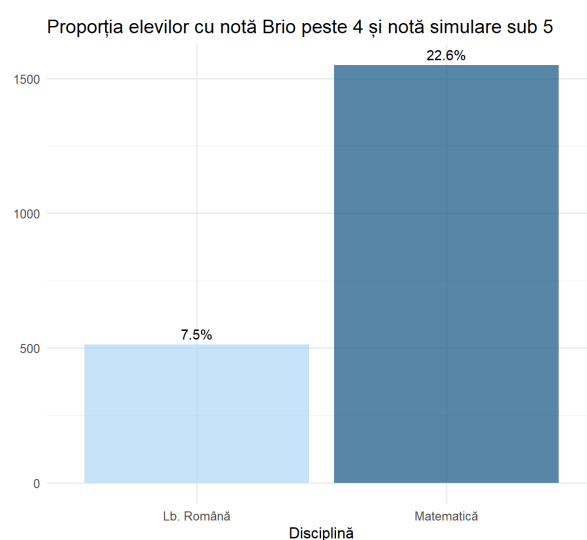
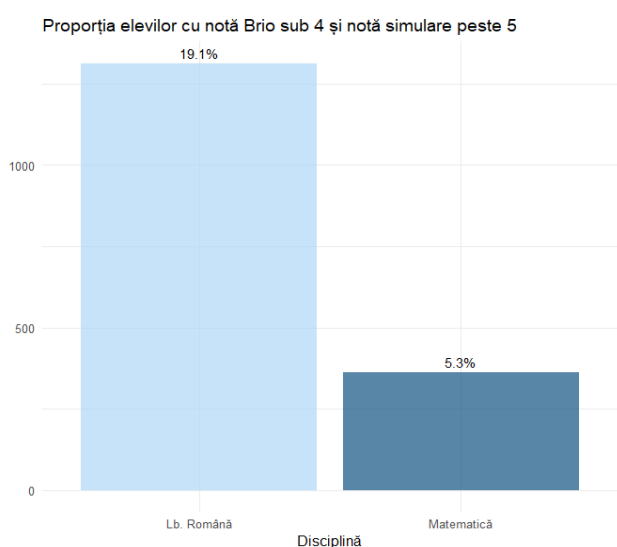
Procentul ridicat de creșteri la Lb. Română ar putea sugera că elevii au reușit să-și îmbunătățească semnificativ competențele în cele 4 luni dintre cele două evaluări. Rezultatele mai pot de asemenea sugera că intervențiile sau metodele de studiu accesate de către elevi între cele două evaluări au fost mai eficiente pentru Lb. Română decât pentru Matematică, sau că elevii au găsit în mod natural că este mai ușor să-și îmbunătățească performanța la Lb. Română în comparație cu Matematica.

O altă posibilă interpretare ar fi că testul de Lb. Română de la simularea EN a avut o dificultate mai scăzută decât testul Brio. Nu putem fi siguri care este interpretarea potrivită, deoarece nu avem date despre dificultatea testelor din cadrul simulării EN, și nici despre măsura în care elevii au depus efort sau au participat la activități remediale în vederea creșterii performanței lor la aceste discipline.

Praguri: Nota 4 la testul Brio în raport cu nota 5 la simularea EN

Tot în materie de diferențe între scoruri, a mai fost calculat procentul de elevi care au obținut note la simularea EN peste sau sub 5, în raport cu scoruri peste sau sub 4 la testul Brio.

Astfel, numărul de elevi care a luat peste nota 5 la simularea EN și sub nota 4 la testul Brio este de 1338 pentru Lb. Română și de 417 pentru Matematică. Aceste rezultate sugerează că dacă un elev obține o notă sub 4 la testul Brio, acesta are șanse foarte scăzute să obțină peste nota 5 la simularea EN, mai ales în cazul matematicii.



Pe de cealaltă parte, numărul de elevi care a luat note sub 5 la simularea EN și peste 4 la testul Brio este de 514 pentru Lb. Română și de 1550 pentru Matematică. Aceste rezultate subliniază importanța obținerii unei note de cel puțin 4 la testele Brio, sugerând că dacă un elev obține o notă peste 4 la testul Brio, acesta are șanse foarte scăzute să obțină o notă sub 5 la simularea EN, mai ales în cazul probei de Lb. Română.

Concluzii

Rezultate principale și posibile interpretări

1. În cazul disciplinei Matematică există o diferență foarte scăzută între rezultatele obținute de elevi la cele două evaluări. În același timp, în cazul disciplinei Lb. Română avem o diferență foarte ridicată între cele două seturi de rezultate, de peste o abatere standard (cca. 15%) în favoarea scorurilor de la simularea EN.
2. Rezultatele obținute de elevi la testele Brio sunt corelate cu o intensitate medie spre ridicată cu rezultatele la simularea EN. Totodată, performanța măsurată prin testele Brio poate fi utilizată pentru a anticipa/prezice notele obținute de către elevi la simularea EN.
3. La disciplina Matematică majoritatea elevilor au avut rezultate similare la cele două evaluări, în timp ce la disciplina Lb. Română o mare parte din elevi au luat note mai mari la această probă din cadrul simulării EN decât la testul Brio. De asemenea, marea majoritate a elevilor au luat note mai mari cu 1 punct (diferențe pozitive) la proba EN de Lb. Română, în timp ce doar un sfert dintre aceștia au reușit să obțină o astfel de notă mai mare și la Matematică. În privința notelor mai mici cu cel puțin 1 punct, există un număr relativ redus de astfel de cazuri, mai ales în cazul probei de Limba Română, unde o foarte mică parte din elevi au luat o astfel de notă.
4. Dacă un elev obține o notă sub 4 la testul Brio, acesta are șanse foarte scăzute să obțină peste nota 5 la simularea EN, mai ales în cazul matematicii. În același timp, dacă un elev obține o notă peste 4 la testul Brio, acesta are șanse foarte scăzute să obțină o notă sub 5 la simularea EN, mai ales în cazul probei de Lb. Română.

În urma analizei datelor s-au conturat două posibile interpretări ale rezultatelor. Pe de o parte, diferențele observate între notele elevilor la cele două testări, mai ales în cazul competențelor lingvistice, pot sugera că testele măsoară competența elevilor cu precizie (dificultate și capacitate de discriminare) diferită. Pe de altă parte, aceste rezultate pot sugera că nivelul de competență al elevilor s-a îmbunătățit (în cele 4 luni dintre teste) preponderent la Lb. Română, posibilele intervenții sau metode de studiu accesate de către elevi între cele două evaluări fiind astfel mai eficiente pentru această disciplină decât pentru Matematică.

Limitări și direcții viitoare de analiză

Este important de avut în vedere că testele Brio au fost aplicate în cadrul unui demers de evaluare inițială a tuturor elevilor de gimnaziu și liceu din Județul Iași. Acest aspect a limitat durata testelor Brio la 45 de minute, făcându-le mult mai scurte decât testele de la simularea EN care s-au derulat pe parcursul a 3 ore. Totodată, datorită aceluiasi context de evaluare inițială, testele Brio au evaluat doar nivelul de competențe dobândite în clasa a VII-a. O evaluare inițială mai complexă, acoperind o parte mai consistentă a materiei și din clasele anterioare, pe o durată mai mare a testării, va crește cu mare probabilitate puterea de predicție a testului inițial chiar mai mult.

În regresia liniară simplă, utilizată în analiza relației dintre rezultatele celor două evaluări, presupunem că toate observațiile sunt independente unele de altele, ceea ce nu este cazul aici. Avem elevi care sunt grupați în clase și școli, ceea ce înseamnă că performanțele elevilor din aceeași clasă sau școală pot fi mai asemănătoare între ele decât performanțele elevilor din clase sau școli diferite. Aceste efecte pot fi analizate, pentru a identifica școlile și/sau clasele în care diferențele dintre cele două teste sunt mai mari, identificând astfel în mod potențial zone în care educația produce efecte mai puternice și de unde pot fi sugerate exemple de bune practici.

Astfel, aceste rezultate oferă o înțelegere parțială a relației dintre nota la simularea Evaluării Naționale și scorul la testele Brio. Pentru o înțelegere mai cuprinzătoare a relației dintre rezultatele celor două evaluări, dar și a factorilor care pot influența rezultatele elevilor la examenele naționale, este necesară luarea în considerare a unor variabile suplimentare și analiza atentă a contextului educațional.

Câțiva dintre factorii care ar trebui măsurați suplimentar pentru a putea înțelege mai bine această relație într-un studiu similar viitor, de exemplu în cazul probelor de EN sau Bacalaureat, ar putea fi:

1. Dificultatea și capacitatea de discriminare a itemilor incluși în testele administrate la examenele naționale;
2. Numărul și tipul de activități care se derulează în vederea remedierii sau stimulării performanței și măsura în care elevii participă și se implică la aceste activități.
3. Alte variabile de context, cum ar fi: mediul de învățare al clasei, metodele de predare utilizate de către profesori, motivația elevilor pentru învățare etc.