

Metode de evaluare a studenților la fizică

Chistol Vitalie, Tronciu Vasile

Departamentul Fizică, Universitatea Tehnică a Moldovei, Chișinău, Republica Moldova
vitalie.chistol@fiz.utm.md

Rezumat— Aplicarea Regulamentului de notare a studenților nu întotdeauna reprezintă cea mai eficientă metodă de interpretare a rezultatelor evaluării. Uneori această metodă poate fi chiar greșită, atunci când dificultatea testului este foarte ridicată sau foarte joasă. În aceasta lucrare sunt examinate diferite metode de analiză a rezultatelor obținute de studenți în urma evaluării precum și modalități de interpretare a acestor rezultate.

Termeni cheie—Evaluare, itemi de test, prelucrare statistică, percentilă, scor Z.

I. INTRODUCERE

Predarea – învățarea – evaluarea sunt cele trei activități fundamentale ce se desfășoară în mod obligatoriu în cadrul procesului de învățământ. Aceste activități fundamentale devin eficiente numai în cazul în care ele formează un tot întreg. În această lucrare vom analiza în detalii activitatea de evaluare.

Sistemul de evaluare din învățământul universitar vizează următoarele cerințe:

- evaluarea obiectivelor curriculare și a strategiilor educaționale;
- evaluarea activității de predare-învățare, a strategiilor didactice și a metodelor de învățământ;
- evaluarea performanțelor profesionale;

- evaluarea întregului sistem de învățământ;
- informarea studenților despre rezultatele obținute și cauzele nerealizării obiectivelor curriculare propuse.

Este bine știut că evaluarea efectuată calitativ urmărește scopul nu numai de măsurare a cantității cunoștințelor dobândite de student, ci poate să ne ofere o informație bogată despre întregul proces de studii, oferindu-ne soluții pentru perfecționarea procesului didactic.

II. ANALIZA REZULTATELOR OBȚINUTE ÎN URMA EVALUĂRII

Deseori evaluarea cunoștințelor studenților la fizică se efectuează în baza itemilor de tip eseu. Dar, în ultimii ani, se utilizează foarte des și itemii de test cu alegere multiplă. Să presupunem că, în cadrul procesului de evaluare, au fost testați 10 studenți, cărora li s-a propus un test alcătuit din 20 de itemi cu alegere multiplă (de obicei atât numărul de studenți evaluați, cât și numărul de itemi este mult mai mare). Fiecare item conține câte 4 opțiuni, dintre care studentul trebuie să aleagă doar una, pe cea corectă. Pentru răspunsul corect studentul obține un punct, pentru cel greșit zero puncte. În Tabelul 1 sunt prezentate rezultatele obținute de fiecare student, în ordine descrescătoare a scorului obținut.

TABEL I. REZULTATELE TESTĂRII

Studentul	Analiza itemilor de test																				A1	A0	%
	Itemul																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20			
E	1	0	0	0	1	0	1	1	1	1	0	1	1	1	0	1	1	0	1	0	12	8	60
H	1	0	0	1	0	0	1	1	1	1	0	1	1	1	0	0	1	0	0	1	11	9	55
J	1	0	0	0	0	0	1	1	1	1	0	1	0	1	0	1	0	0	1	1	10	10	50
F	0	0	0	0	1	0	0	1	0	1	0	1	1	1	0	1	0	1	0	1	9	11	45
D	0	1	0	0	1	0	1	1	1	1	0	0	0	1	0	1	0	0	1	0	9	11	45
B	1	1	0	0	0	0	0	1	1	0	0	1	0	1	0	0	1	0	0	1	8	12	40
A	1	1	0	1	0	0	1	1	0	0	0	1	0	1	0	0	0	0	1	0	8	12	40
C	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	0	1	0	1	6	14	30
G	1	1	0	0	0	1	1	1	0	1	0	1	0	1	0	0	0	0	0	0	8	12	40
K	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	4	16	20
Corecte	6	5	0	3	3	1	6	9	5	7	0	8	3	9	0	5	3	2	5	5			
Greșite	4	5	10	7	7	9	4	1	5	3	10	2	7	1	10	5	7	8	5	5	Media	43	
Indice de dif.	0.4	0.5	1	0.7	0.7	0.9	0.4	0.1	0.5	0.3	1	0.2	0.7	0.1	1	0.5	0.7	0.8	0.5	0.5	Deviația standard	12	
Indice de discr.	0	-0.6	0	-0.2	0.6	-0.2	0.4	0.2	0.6	0.6	0	0	0.6	0.2	0	0.6	0.2	0	0.2	0.2			

În continuare analizăm datele incluse în Tabelul I. Pe verticală avem următoarele:

- Cifrul fiecărui student participant la evaluare. În cazul când rezultatele se fac publice, nu se recomandă să fie afișat

numele studentului, deoarece studenții cu un scor mai mic sunt puși într-o situație neplăcută;

- Rezultatele obținute de studenți la fiecare item, unde „1” reprezintă răspuns corect, iar „0” – răspuns greșit;
- „A1” este numărul total de răspunsuri corecte;

- „A0” este numărul total de răspunsuri greșite;
 - „%” este scorul studentului în procente, care se calculează împărțind numărul total de răspunsuri corecte ale studentului la numărul total de itemi.

Pe orizontală, după liniile ce includ rezultatele fiecărui student, sunt incluse numărul rezultatelor corecte și al celor greșite. Ulterior sunt expuși indicii de dificultate i_{dif} și de discriminare i_{discr} ai itemului. Indicele de dificultate al itemului se calculează după formula:

$$i_{dif} = N_g / N, \quad (1)$$

unde N_g este numărul studenților care au răspuns greșit la itemul dat, iar N este numărul studenților care au participat la evaluare.

Pentru a determina indicele de discriminare, inclus în ultima linie din Tabelul 1, numărul de studenți se împarte în două: prima și a doua jumătate (în caz că numărul de studenți este impar, studentul de mijloc se exclude). Notăm prin N_1 numărul răspunsurilor corecte la itemul dat din prima jumătate a studenților, prin N_2 – numărul răspunsurilor corecte din cea de a doua jumătate, iar prin N' – numărul studenților din una din jumătăți. Atunci indicele de discriminare se va determina astfel [1]:

$$i_{discr} = (N_1 - N_2) / N'. \quad (2)$$

În Tabelul I sunt indicate de asemenea și valoarea medie a rezultatelor obținute de studenți, precum și deviația standard a acestor rezultate (ambele valori pot fi modelate la calculator).

În continuare vom analiza rezultatele obținute. Se consideră că dificultatea testului este medie dacă valoarea medie a rezultatelor obținute este cuprinsă între 60 și 80% [1]. În cazul descris mai sus media este de 43%. Deci, pentru studenții examinați testul a fost foarte dificil. Cel mai bun student a acumulat un scor de doar 60%. Conform Regulamentului [2] acest student trebuie să fie apreciat cu nota 5, iar majoritatea, începând cu studentul J, trebuie să fie apreciați cu note insuficiente. Cauza reușitei scăzute poate fi motivată prin pregătirea slabă a studenților, sau dificultatea prea mare a testului. Însă, studentul nu trebuie să fie penalizat din motiv că profesorul a propus un test foarte dificil. În cele ce urmează vom prezenta și alte modalități de interpretare a rezultatelor evaluării.

Să examinăm fiecare item în parte. Vedem că itemii 3, 11 și 15 au gradul de dificultate 1, deci la acest item nu a răspuns corect nici un student. Care este cauza? Sau profesorul nu a predat bine lecția la tema respectivă, sau itemul a fost prea complicat, sau a fost formulat neclar. În toate cazurile studentul nu are nici o vină. De aceea, itemul respectiv trebuie exclus. În caz general se recomandă [1] ca itemii cu indicii de dificultate mai mare decât 0,8 să fie excluși. Astfel, analizând rezultatele obținute vedem care teme au fost însușite mai bine sau mai slab de către studenți.

Itemul 2 are indicii de discriminare -0,6. Ce ne spune acest indicator? Din formula (2) observăm că indicii de discriminare este pozitiv dacă numărul studenților mai bine pregătiți care au răspuns la itemul dat este mai mare decât cel al studenților mai slab pregătiți și negativ – dacă numărul studenților mai slab pregătiți este mai mare. În cazul itemului 2, din jumătatea studenților bine pregătiți a răspuns corect doar un singur student, iar din jumătatea celor mai slab pregătiți – 4. Care este cauza? Sau studenții mai slab pregătiți au ghicit răspunsul, sau itemul a fost formulat în așa fel, încât discriminează studenții pregătiți mai bine. Pentru a înțelege cum un item poate discrimina un student bine pregătit, vom examina următorul item. Selectați răspunsul corect:

Becul electric a fost inventat de:

- A). Jhon Whitney;
- B). Genghis Khan;
- C). Thomas Edison;
- D). Bill Clinton.

În majoritatea manualelor este scris că becul electric a fost inventat de Thomas Edison. În realitate istoricii spun că invenția becului nu îi aparține lui Thomas Edison, iar ca rezultat au creat o lista cu 22 de inventatori care au realizat ceva ce se aseamănă cu becul. Printre ei sunt doi inventatori mai puțin cunoscuți precum Sir Humphry Davy sau Sir Joseph Wilson Swan. Studentul care a citit doar manualul știe că becul electric a fost inventat de Thomas Edison. Studentul care a mai citit și literatură suplimentară știe că nu Edison a fost inventatorul becului. Din lista de mai sus Genghis Khan nu poate fi, Bill Clinton este exclus, rămâne doar Jhon Whitney (un nume inventat). Astfel, studentul mai bun va marca rezultatul greșit A), iar cel mai slab – rezultatul corect C). Astfel, se recomandă ca itemii cu indicii de discriminare mai mic ca - 0,2 să fie excluși.

Itemii 8 și 14 au indicii de dificultate 0,1, deci au fost foarte simpli. De obicei se recomandă ca itemii cu indicii de dificultate mai mic decât 0,2 să fie excluși. Totuși, pentru a nu defavoriza studenții ei ar putea fi lăsați. Decizia finală despre un astfel de item o ia profesorul. Astfel, în testul de mai sus se exclud itemii 2, 3, 6, 11 și 15, iar scorul obținut, media testului și deviația standard se calculează din nou (vezi Tabelul II).

TABEL II. ANALIZA DETALIATĂ A ITEMILOR DE TEST

Studentul	A1	%
E	12	80
H	11	73
J	10	67
F	9	60
D	8	53
B	7	47
A	7	47
C	6	40
G	6	40
K	3	20
	Media testului	53
	Deviația standard	18

TABEL III. REZULTATELE PRELUCRATE ALE TESTĂRII

Itemul de test	Opțiunea			
	A	B	C	D
1	4	0	6	0
2	2	2	1	5
3	9	1	0	0
4	3	2	4	1
5	3	2	4	1
6	2	4	1	3
7	0	0	4	6
8	9	1	0	0
9	1	5	4	5
10	0	3	7	0
11	2	0	2	8
12	1	0	0	9
13	3	3	3	1
14	0	0	1	9
15	7	0	2	1
16	0	5	5	0
17	3	2	2	3
18	3	2	1	4
19	2	0	6	2
20	2	5	0	3

Ținem să menționăm că din excluderea unor itemi studenții au numai de câștigat. Observăm că scorul maxim a ajuns la 80%. Astfel, conform Regulamentului notele studenților ar trebui să crească.

Calitatea itemilor poate fi analizată, înregistrând pentru fiecare dintre ei opțiunea selectată de student (vezi Tabelul III).

În continuare vom prezenta câteva exemple. La itemul 1 observăm că 4 studenți au selectat opțiunea A, 6 studenți au selectat opțiunea corectă C, iar opțiunile B și D nu le-au selectat nici un student. De aici rezultă că opțiunea A este un distractor bun, iar distractorii B și D n-au putut distra atenția studenților. Se propune ca distractorii B și D să fie înlocuiți cu alții.

Ce ține de itemul 3, vedem că aproape toți studenții au selectat distractorul A și nici un student nu a ales opțiunea corectă D. Concluzionăm că studenții au înțeles greșit materia la tema respectivă. În viitor trebuie să atragem o atenție mai mare la predarea temei respective.

Astfel, o analiză detaliată a rezultatelor din Tabelul III ne permite să îmbunătățim itemii pentru testele viitoare și să tragem unele concluzii ce ține de predarea temelor de către profesor.

III. INTERPRETAREA SCORULUI OBTINUT DE STUDENT

Dificultatea testului poate să varieze în dependență de cerințele profesorului care a alcătuit testul: un profesor mai exigent poate alcătui un test mai dificil, un profesor mai indulgent – unul mai simplu. În rezultat, dacă vom urma cerințele Regulamentului, două grupe de studenți aproximativ egali după capacități, în urma evaluării vor obține rezultate complet diferite.

Este avantajos ca la evaluarea studenților să ținem cont de dificultatea testului utilizând prelucrarea statistică a rezultatelor. De exemplu, la multe concursuri de fizică, participanților li se acordă diplome după metoda descrisă mai jos [3]:

- Se calculează media m a scorului acumulat de participanți și deviația standard s .

- Se acordă diplomă de gradul I participanților care au acumulat un scor mai mare decât $m+2s$ (selectează rezultatele cuprinse în regiunea N1 de 2,5% din zona rezultatelor excelente) (vezi Fig. 1) [2].

- Se acordă diplomă de gradul II și III participanților care au acumulat un scor mai mic decât $m+2s$, dar mai mare decât $m+s$ (selectează rezultatele cuprinse în regiunea N2 de 13,5% din zona rezultatelor foarte bune).

- Se acordă premii speciale (mențiuni) participanților care au acumulat un scor cuprins între $m+0,5s$ și $m+s$ (selectează rezultatele cuprinse în regiunea N3 de aproximativ 27,5% din lucrările prezentate la concurs).

În cazul nostru am putea echivala diplomele de gradul I, II și III cu notele 10, 9 și 8, iar mențiunile – cu notele de 7 și 6.

Astfel, conform rezultatelor din Tabelul II $m+2s=53+2\cdot 18=89\%$.

Scorul maximal obținut de studenți este de 80%. Deci, nota 10 nu este obținută de nici un student. Pentru notele 9 și 8 scorul trebuie să fie cuprins între 89% și $m+s=53+18=71\%$. Din Tabelul II observăm că doi studenți au scorul cuprins în acest interval: studentul E cu scorul de 80%, căruia putem să-i acordăm nota 9 și studentul H cu scorul de 73%, căruia putem să-i acordăm nota 8. Pentru notele 7 și 6 scorul trebuie să fie cuprins între 71% și $m+0,5s=53+9=62\%$. Din același Tabel II vedem că doar un singur student are scorul cuprins în acest interval: studentul J – scor 67%, căruia putem să-i punem nota 7, iar cu nota 6 nu va fi apreciat nici un student.

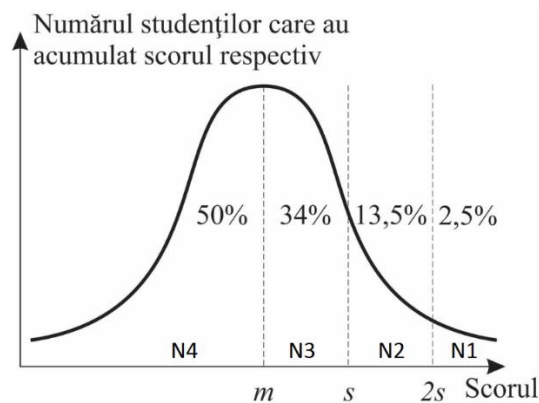


Fig. 1 Curba distribuției normale.

Pentru nota 5 putem utiliza prevederile Regulamentului [2]: cu notă trecătoare se apreciază studenții care au acumulat mai mult de 50% din totalul de puncte. În cazul descris mai sus, studenții F și D obțin nota 5, iar ceilalți din lista studenților (Tabel I) obțin note insuficiente. Astfel, la notarea studenților s-a ținut cont de dificultatea testului (scorul mediu) și, în cazul când vom avea două teste care se deosebesc foarte mult după dificultate, în rezultatul evaluării notele studenților din diferite grupe se vor deosebi puțin unele de altele.

O altă metodă de interpretare a scorului este calcularea percentilei, care se utilizează pe larg în universitățile de peste hotare. Pentru calcularea percentilei, într-un tabel se introduc următoarele date (vezi Tabelul IV):

- Scorurile studenților (în ordine descrescătoare);
- *Scorul de mai jos*, care se începe cu cifra 0 pentru ultimul student și atinge valoarea maximă pentru primul student. Este evident că valoarea maximă a scorului de mai jos va fi egală cu $N - 1$, unde N este numărul de studenți evaluați.
- *Factorul de percentilă* f , care se calculează după formula $f = 100 / (N - 1)$.
- *Percentila*, care se calculează împărțind scorul de mai jos la factorul de percentilă.

Studentul care a obținut cel mai bun scor întotdeauna va avea percentila 100, iar ultimul 0. Studentul din mijlocul listei întotdeauna va avea percentila 50. Menționăm că percentila nu ne arată scorul obținut de student, ci doar locul pe care studentul l-a deținut în grup. Pentru ce este necesară percentila? Sunt universități unde se fac studii la nivel înalt, profesorii sunt foarte exigenți și studenții absolvec universitatea cu note mici. Totodată sunt și universități în care cerințele sunt mici, dar notele studenților sunt mari. Astfel, concludem ca notele din diplomă încă nu reflectă performanța absolventului. Totuși, dacă studentul a fost printre primii în grupul său, aceasta constituie un indicator că el este o persoană ageră și poate deveni un specialist bun.

TABEL IV. CALCULAREA PERCENTILEI

Elevul	Scorul	Sc. de mai jos	f	Percent.	Scor Z
E	80	9	11.11	100	1.5
H	73	8		89	1.1
J	67	7		78	0.8
F	60	6		67	0.4
D	53	5		56	0.1
B	47	4		44	-0.3
A	47	3		33	-0.3
C	40	2		22	-0.6
G	40	1		11	-0.6
K	20	0		0	-1.7

În prezent, în suplimentul de la diplomă, în afara notelor obținute de student se mai introduce expresia „A fost pe locul 7 (de exemplu) din 15”. Această expresie este o analogie a percentilei. Totuși, ar fi mai comod când poziția persoanei în grup ar fi exprimată printr-o cifră și nu printr-o expresie.

O prelucrare statistică a rezultatelor evaluării este utilă și în cazul necesității comparării performanțelor studenților din grupe diferite, de exemplu, în cazul acordării burselor speciale. Pentru compararea performanțelor poate fi utilizat scorul Z , care se calculează astfel:

$$Z = (x - m) / s, \quad (3)$$

unde x este scorul obținut de student, m este scorul mediu al testului, iar s – deviația standard.

În continuare vom compara performanțele la fizică ale studenților M și N din grupe diferite, care, în urma evaluării, au obținut următoarele rezultate (Tabelul V).

TABEL V. COMPARAREA PERFORMANȚEI STUDENȚILOR

Studentul	x	m	s
M	80	85	10
N	60	55	10

S-ar părea că studentul A, cu scorul 80, are la fizică performanțe mai bune decât studentul B, care are un scor de 60. Însă, din Tabelul V observăm că studentul A a avut un test de dificultate mică ($m=85$), iar studentul B – un test mult mai dificil ($m=55$). Vom calcula scorul Z pentru fiecare student:

$$Z_M = (80 - 85) / 10 = -0,5,$$

$$Z_N = (60 - 55) / 10 = 0,5.$$

Din analiza scorului Z observăm că, chiar dacă studentul M are un scor mai mare decât studentul N, performanțele lui la fizică sunt mai mici. Menționăm că, estimarea scorului Z nu este necesară în cazul comparării performanțelor studenților din același grup. În acest caz performanțele lor pot fi comparate cu ajutorul scorului obținut, sau cu ajutorul percentilei. Scorul Z este important în cazul necesității de comparare a performanțelor studenților din două grupuri diferite, sau performanțele aceluiași student la două obiecte diferite.

IV. CONCLUZII

Din cele expuse în această lucrare observăm că, prelucrând statistic rezultatele evaluării, putem obține o informație bogată despre calitatea itemilor, gradul de însușire de către studenți a unor compartimente ale cursului de fizică, calitatea predării acestor compartimente de către profesor și, totodată, efectuăm o evaluare mai obiectivă și mai eficientă atât a studentului cât și a profesorului. Este adevărat că alcătuirea tabelor prezentate în această lucrare necesită mult timp. Pe de altă parte, nu toate prelucrările prezentate mai sus sunt obligatorii. Ar fi suficient ca Tabelul III să fi alcătuit periodic, pentru a judeca despre calitatea unor itemi nou alcătuiți. În cazul utilizării unor itemi de tip eseu, tabelele I și II pot fi excluse. În acest caz putem utiliza doar scorul fiecărui student. În final obținem valorile „ m ” și „ s ”, pentru calcularea notei studentului și, dacă este necesar, a percentilei sau a scorului Z .

BIBLIOGRAFIE

- [1] M. David Miller, Robert L. Linn, Norman E. Gronlund, Measurement and Assessment in Teaching, 10th ed., Pearson Education, New Jersey, 2009.
- [2] Regulamentul Universității Tehnice a Moldovei cu privire la organizarea studiilor în baza sistemului național de credite de studiu, UTM, 2016.
- [3] Regulamentul concursului de fizică creativă "Ștefan Procopiu".