

ՀԱՄԱԿԱՐԳԶԱՅԻՆ ԹԵՍՏԱՅԻՆ ԱՌԱՋԱԴՐԱՆՔՆԵՐԻ ՈՐԱԿԻ ԳՆԱՀԱՏՄԱՆ ՄԵԹՈԴԻԿԱ

Յ. Ա. ՄԵԼԻՔՅԱՆ

Տեխնիկական գիտությունների թեկնածու, դոցենտ,
ԻՂՏ ամբիոնի վարիչ

Համակարգչային թեստավորման թեստային առաջադրանքների որակի գնահատման ամենակարևոր ցուցանիշներից են [1].

- առաջադրանքի դժվարությունը,
- առաջադրանքի տարբերիչ ունակությունը:

Թեստային առաջադրանքի՝ որպես խնդրի դժվարությունը բնութագիր է, որն արտացոլում է նրա լուծելիության ստատիկ մակարդակը: Սովորաբար այն սահմանվում է որպես առաջադրանքը ճիշտ կատարող սովորողների թվի և սովորողների ընդհանուր թվի հարաբերություն: Այդ ցուցանիշը փոփոխվում է $[0, 1]$ հատվածում և ինչքան մեծ է նրա արժեքը, այնքան հեշտ է առաջադրանքը:

Առաջադրանքի դժվարության հիմնական ցուցանիշը դժվարության ինդեքսն է, որը հաշվվում է հետևյալ բանաձևով.

$$U_i = (1 - \frac{N_p}{N}) \cdot 100\% ,$$

որտեղ՝ U_i - i -րդ առաջադրանքի դժվարության ինդեքսն է տոկոսներով,

N_p - առաջադրանքը ճիշտ կատարող սովորողների քանակն է,

N - փորձարկվողների ընդհանուր քանակն է:

Համակարգչային թեստավորման պրակտիկան ցույց է տալիս [1], որ սովորողների գիտելիքների և ունակությունների անհատական մակարդակի մասին առավելագույն ինֆորմացիա ապահովում են միջին դժվարության ($= 50\%$) առաջադրանքները:

Մանկավարժական թեստերի մշակման գործընթացում հատուկ ուշադրություն է դարձվում առաջադրանքների հատկապես կարևոր մյուս բնութագրի՝ "առաջադրանքների տարբերիչ ունակության" վրա, որից զգալիորեն կախված է թեստի համապատասխանությունը (валидность) չափվող գիտելիքներին: Այս բնութագիրը ցույց է տալիս, թե թեստային առաջադրանքը ինչքանով է տարբերում ուսումնական նյութին տիրապետող և չտիրապետող ուսանողներին:

Թեստային առաջադրանքի տարբերիչ ունակությունը (D) հաշվելու համար այդ առաջադրանքի կատարումը հանձնարարվում է սովորող-փորձարկողների ըստ ունակության երկու հակադիր ("բարձր" – "ցածր") խմբերի: Ստանալով փորձարկման արդյունքները, D -ն հաշվվում է հետևյալ բանաձևով.

$$D = \frac{n_1}{N_1} - \frac{n_2}{N_2} ,$$

որտեղ՝ D - առաջադրանքի տարբերիչ ունակությունն է,

N_1, N_2 - "բարձր" - "ցածր" հակադիր խմբերի սովորող-փորձարկողների քանակներն են համապատասխանաբար,

n_1, n_2 - այդ խմբերի այն սովորող-փորձարկողների քանակներն են, որոնք ճիշտ են կատարել առաջադրանքը:

D ցուցանիշը արժեքներ է ընդունում $[-1,+1]$ հատվածում: Եթե $D = 1$, ապա առաջադրանքը ունի առավելագույն տարբերիչ ունակություն: $D = 0$ նշանակում է, որ առաջադրանքը բոլորովին չի տարբերում ուսումնական նյութին տիրապետող և չտիրապետող սովորողներին: $D = -1$ տարբերակը հանդիպում է շատ հազվադեպ և նշանակում է, որ առաջադրանքը տարբերում է փորձարկողներին, բայց ինվերսված, այսինքն՝ ճիշտ պատասխանում են ուսումնական նյութին չտիրապետողները, իսկ նյութին տիրապետողները պատասխանում են սխալ:

Թեստի որակի գնահատման լավագույն արդյունք են ապահովում մաթեմատիկական վիճակագրության հիմնական բնութագրերի հաշվարկները սովորողների ընտրված խմբի (քաղվածքի) համար [2]: Այդ հաշվարկները թույլ են տալիս մեծ ճշտությամբ պատասխանել այն հարցին, թե սովորողի գիտելիքների գնահատականը ինչքանով է համապատասխանում նրա պատրաստվածությանը: Հաշվարկները կատարելու համար որպես ելման ինֆորմացիա վերցվում է սովորողների ընտրված խմբի թեստավորման արդյունքների վիճակագրական մատրիցը:

Կոնկրետ հաշվարկներ կատարելու համար օգտվենք աղյուսակ 1-ում ներկայացված վիճակագրական մատրիցից: Հաշվարկների պարզեցման նպատակով սովորողների քանակը (քաղվածքի ծավալը) աղյուսակ 1-ում վերցված է 14:

Աղյուսակ 1

Փորձարկողի համարը (i)	Թեստի առաջադրանքների համարները										Փորձարկողի հավաքած գումարային քալը (x _i)
	1	2	3	4	5	6	7	8	9	10	
1	1	1	1	1	1	1	0	0	0	0	6
2	1	1	0	0	0	0	0	0	0	0	2
3	0	0	0	0	0	0	0	1	0	0	1
4	1	1	0	1	1	1	1	1	1	1	9
5	1	0	1	0	1	1	0	0	0	0	4
6	1	1	1	0	0	0	0	1	0	0	4
7	1	1	1	1	0	1	0	0	0	0	5
8	1	1	1	1	0	0	0	0	0	0	4
9	1	1	1	1	1	1	1	1	1	0	9
10	0	0	1	1	1	1	0	1	0	0	5
11	1	0	1	0	0	0	0	0	0	0	2
12	1	1	0	1	1	0	1	1	1	0	7
13	0	0	0	0	0	0	0	0	0	0	0 (հեռացնել)
14	1	1	1	1	1	1	1	1	1	1	10 (հեռացնել)
Ճիշտ պատասխանների քանակը 10 առաջադրանքների համար	10	8	8	7	6	6	3	6	3	1	

Ակնհայտ է, որ ինչքան մեծ է քաղվածքի ծավալը, այնքան արժանահավատ են ստացված արդյունքները: Հետագա գնահատման ժամանակ հաշվի չեն առն-

վում 13-րդ և 14-րդ փորձարկողների արդյունքները, քանի որ դրանք չեն կրում անհրաժեշտ ինֆորմացիա (13-րդի համար այդ թեստը խիստ բարդ է, իսկ 14-րդի համար՝ խիստ թեթև):

Աղյուսակ 1-ի հիման վրա ձևավորվում են ստացված բալերի վարիացիոն շարքը (աճման կարգով ռանժիրավորված շարքը- աղյուսակ 2) և հաճախությունների բաշխման շարքը (աղյուսակ 3):

Աղյուսակ 2

Բալը	1	2	2	4	4	4	5	5	6	7	9	9
Ռանգը	1	2	2	3	3	3	4	4	5	6	7	7
Աղյուսակ 1-ի տողի համարը	3	2	11	5	6	8	7	10	1	12	4	9

Աղյուսակ 3

Բալը	1	2	4	5	6	7	9
Հաճախությունը	1	2	3	2	1	1	2

Ծանոթագրություն - Բոլոր հաճախությունների գումարը պետք է հավասար լինի փորձարկողների քանակին (մեր օրինակում՝ 12):

Այնուհետև գնահատվում է թեստավորման արդյունքների (բալերի) մոդան՝ ամենահաճախ հանդիպող թիվը: Քննարկվող օրինակում մոդան հավասար է 4-ի՝ հանդես է գալիս 3 անգամ: Եթե բալերի հաճախությունների բաշխման շարքը բազմամոդալ է, այսինքն կան ամենամեծ հաճախությամբ մի քանի թվեր, ապա դա խոսում է թեստի անհաջող կառուցվածքի մասին: Բալերի հաճախությունների շարքը պետք է ունենա նորմալ բաշխում, այսինքն՝ լինի ունիմոդալ (ունենա մոդայի մեկ արժեք):

Նորմատիվակողմնորոշված լավ թեստը ապահովում է անհատական բալերի նորմալ բաշխում, եթե բալերի միջին արժեքը գտնվում է շարքի կենտրոնում, իսկ մնացած արժեքները կենտրոնացված են նրա շուրջը: Միջին արժեքի տեղաշարժը դեպի ձախ խոսում է թեստի բարդ առաջադրանքների հավաքածուի և հակառակը, դեպի աջ տեղաշարժը՝ հեշտ առաջադրանքների հավաքածուի մասին:

Միջին բալի արժեքը հաշվվում է որպես բալերի միջին թվաբանական:

Քննարկվող օրինակում.

$$\bar{X} = \frac{1+2+2+4+4+4+5+5+6+7+9+9}{12} \approx 4,83 :$$

Ունենալով շարքի միջին արժեքը, կարելի է հաշվել նրա դիսպերսիան՝ որպես թեստավորման արդյունքների անհամասեռության չափանիշ: Դիսպերսիան հաշվվում է հետևյալ ֆորմուլայով.

$$S^2 = \frac{1}{n} * \sum_{i=1}^n (X_i - \bar{X})^2 :$$

Դիտարկվող օրինակում ($n = 12$) ստացվում է $S^2 \approx 6,14$: Դիսպերսիայի ցածր

արժեքը վկայում է, որ կատարված է փորձարկողների՝ ըստ պատրաստվածության մակարդակների, ոչ օպտիմալ բաժանում, իսկ շատ մեծ արժեքը խոսում է բաշխման պատկերի աղջատման մասին:

Թեստային առաջադրանքների մշակման պրակտիկայում, բացի վերը նշվածներից, օգտագործում են նաև մի շարք այլ վիճակագրական բնութագրեր [2], ինչպես, օրինակ, արդյունքների կոռելյացիայի գործակիցը, մեդիանան (միջնաթիվը), միջին քառակուսային շեղումը, հարաբերական հաճախությունների հիստոգրամման, քաղվածքների համեմատման Ստուդենտի t – հայտանիշը և այլն:

ԳՐԱԿԱՆՈՒԹՅԱՆ ՑԱՆԿ

1. **В. С. Аванесов** – Научные проблемы тестового контроля знаний. М., 2004г.
2. **М. С. Ермаков, А. Ф. Сизова, Т. М. Товстик** – Элементы математической статистики. Учебное пособие, Санкт-Петербург, 2001г.

МЕТОДИКА ОЦЕНКИ КАЧЕСТВА КОМПЬЮТЕРНЫХ ТЕСТОВЫХ ЗАДАНИЙ

Г. А. Меликян

В работе как показатели оценки качества тестовых заданий указаны: «сложность задания» и «различающая способность задания», численные значения которых оцениваются методами математической статистики. Приведен конкретный вычислительный пример.

METHODS OF ASSESSMENT THE QUALITY OF COMPUTER TEST QUESTIONS

H. A. Meliqyan

In this work the indicators for evaluation of the quality of test items are: “the complexity of the task” and “discriminating ability of the task,” the numerical values of which are estimated by methods of mathematical statistics. There is a specific computational example.

ԹԵՍՏԱՅԻՆ ԱՌԱՋԱԴՐԱՆՔՆԵՐԻ ՄՇԱԿՄԱՆ ՍԿԶԲՈՒՆՔՆԵՐԻ ԵՎ ՓՈՒԼԵՐԻ ՄԱՍԻՆ

Յ. Ա. ՄԵԼԻԿՅԱՆ

*Տեխնիկական գիտությունների թեկնածու, դոցենտ,
ԻՖՏ ամբիոնի վարիչ*

Թեստային առաջադրանքների բազան մշակելիս յուրաքանչյուր դասախոս ձգտում է ընտրել առաջադրանքների իր համակարգը, որի հետևանքով միևնույն ուսումնական առարկայի համար կարող են ստեղծվել մի քանի տարբեր բազա-