

Umumta'lim fanlaridan milliy sertifikat imtihonlari tahlili

Ma'lumki, Bilim va malakalarni baholash agentligi tomonidan umumta'lim fanlaridan milliy sertifikat uchun test sinovlari o'tkazib kelinmoqda. Mazkur test sinovlari Vazirlar Mahkamasining "Umumta'lim fanlarini bilish darajasini baholashning milliy test tizimini joriy etish to'g'risida" 2020-yil 12-oktabrdagi 646-son qaroriga muvofiq tashkil etilmoqda.

Milliy sertifikat imtihonlarida ishtirok etgan ayrim talabgorlar tomonidan test topshiriqlari qiyin, ular uchun ajratilgan vaqt esa kam degan mazmunda murojaatlar bo'lmoqda. Mazkur murojaatlar qanchalik o'rinli? Haqiqatan ham test topshiriqlari qiyinmi? Ushbu savollarga javob topish uchun test natijalarining statistik tahlillariga e'tibor qaratish lozim.

Xalqaro imtihon tizimlarida test sinovlaridan oldin har bir test topshirig'i mutaxassislar tekshiruvdan va aprobatsiyadan o'tkaziladi. Test sinovlaridan so'ng esa test sinovi natijalari tahlil qilinadi. Xalqaro tashkilotlar (*masalan, ETS, Cambridge Assessment kabi tashkilotlar*) tomonidan bunday tahlillar uchun klassik va zamonaviy test nazariyalaridan foydalaniladi. Bundan maqsad test sinovlari natijalarining validlik va ishonchlilik darajasini, testning oldindan belgilangan maqsadlarga qanchalik erishilganini va har bitta sinoluvchi (*test topshiruvchi*) qanchalik o'zining qobiliyat darajasiga mos ball olganligini baholashdan iborat.

Klassik test nazariyasi yordamida test topshiriqlariga qo'yilgan asosiy talablar ko'rsatkichlari aniqlanadi. Test topshiriqlariga qo'yiladigan asosiy talablar – topshiriqning qiyinligi, test ballarining dispersiyasi (*o'zgaruvchanligi, farqlanishi*), topshiriqning boshqa topshiriqlar bilan, shuningdek, umumiy ballar yig'indisi bilan korrelyatsiyasi (*mutanosibli*)dan iborat. Topshiriqning qiyinlik darajasini aniqlash usullaridan biri topshiriqni empirik sinovdan o'tkazib, to'g'ri javoblar salmog'ini aniqlashdan iborat.

Zamonaviy test nazariyalari klassik test nazariyasi doirasida hal qilib bo'lmaydidan muammolarni yechish uchun qilingan ilmiy tadqiqotlar natijasida paydo bo'lgan. Bu nazariyalar bilan standartlashtirilgan testlar uchun sifatli test bazalari yaratiladi. Hozirgi kunda tadqiqotchilar tomonidan juda ko'p zamonaviy test nazariyalari rivojlantirilgan. Zamonaviy test nazariyalaridan kompyuter adaptiv testlarining bazasi uchun ham foydalanish mumkin. Hozirgi kunda ham bu nazariya bo'yicha xalqaro tashkilotlar tomonidan ko'plab tadqiqotlar o'tkazilmoqda va turli xil modellar rivojlantirilmoqda.

Klassik test nazariyasi va zamonaviy test nazariyalari bir-birini inkor qilmaydi, balki bir-birini to'ldiradi. Shu sababli Bilim va malakalarni baholash agentligida test sinovlari natijalari ham klassik, ham zamonaviy test nazariyalari asosida tahlil qilinadi. Bunda xalqaro tajribada zamonaviy test nazariyalari ichida eng ko'p Rash modelidan foydalanilgani va bu model bo'yicha olingan natijalarni boshqa modellarga nisbatan

talqin qilish qulay bo‘lgani uchun agentlikda test natijalarini tahlil qilish uchun Rash modelidan foydalaniladi.

Umumta’lim fanlaridan milliy sertifikat imtihonlari natijalarining klassik va zamonaviy test nazariyalari asosidagi statistik tahlili agentlik rasmiy veb-sayti (uzbmb.uz)ning “Milliy sertifikatlash tizimi” bo‘limiga joylashtirilgan bo‘lib, qiziqqanlar mazkur tahlillar bilan tanishishi mumkin.

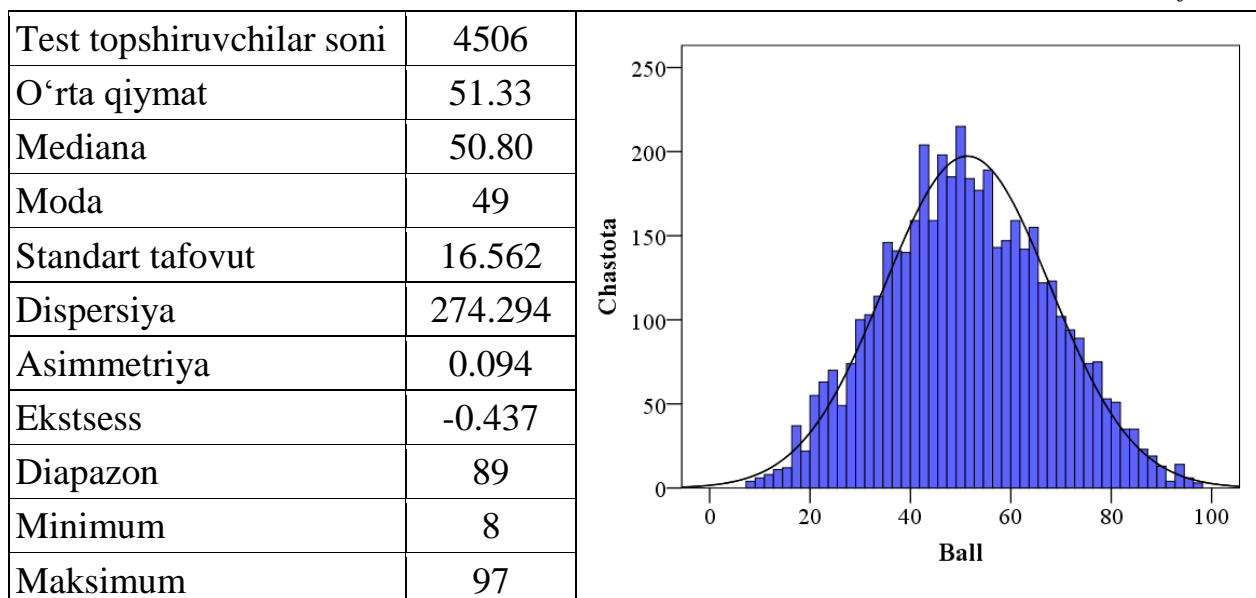
Talabgorlar tomonidan eng ko‘p murojaat **matematika** fani bo‘yicha bo‘layotgani uchun ushbu fan bo‘yicha qilingan tahlilni ko‘rib chiqsak.

2022-yil 4-dekabr kuni o‘tkazilgan matematika fanidan milliy sertifikat imithonida jami **4 506 nafar** talabgor ishtirok etgan. Unda talabgorlarga **35 ta yopiq** turdagi va **10 ta ochiq** turdagi (*har biri A va B qismlarga ajratilgan*) test topshirig‘i berilgan. Ochiq turdagi test topshiriqlari 2 qismga ajratilgani uchun ularni 20 ta teb oladigan bo‘lsak, talabgorlarga jami **55 ta test topshirig‘i** berilgan deb qarash mumkin. Mazkur test topshiriqlarini ishlash uchun talabgorlarga **150 daqiqa** (*2 soat-u 30 daqiqa*) **vaqt** berilgan.

I. Matematika fanidan milliy sertifikat imtihoni natijalarining klassik test nazariyasi asosidagi statistik tahlili

Klassik test nazariyasida test sinovi natijalari asosida aniqlangan test ballari taqsimoti gistogrammasi quriladi va uning normal taqsimotga yaqinligi baholanadi. Gistogrammaning normal taqsimotga yaqinligi testning sifatini va test sinovlarining obyektiv o‘tkazilganini bildiradi. Test ballarining eng ko‘p takrorlanadigan qiymati statistikada moda, o‘shish tartibida joylashtirilgan test ballari qatorining o‘rtasida joylashgan qiymati esa mediana deyiladi. O‘rta arifmetik qiymat, moda va mediana qiymatlari o‘zaro teng bo‘lganda test ballari taqsimoti simmetrik bo‘ladi. Ushbu statistik ko‘rsatkichlar biri-biridan qanchalik ko‘p farq qilsa, ballar taqsimoti normal taqsimotdan shunchalik uzoq bo‘ladi.

1- jadval



Test natijalarining statistik tahliliga ko‘ra, matematika fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalari bo‘yicha test variantining **ishonchlilik koeffitsiyenti** (*Kronbax alfa koeffitsiyenti*) $0,891 \approx 0,9$ ga teng bo‘ldi. Kronbax alfa koeffitsiyentining bu qiymati ushbu test variantining ichki muvofiqligi “a’lo” ekanligini ko‘rsatadi. Bu test variantiga kiritilgan test topshiriqlari barcha darajadagi talabgorlarning qobiliyatini juda yaxshi (*a’lo darajada*) baholayotganidan darak beradi.

2- jadval

Kronbax alfa	Ichki muvofiqlik
$0,9 \leq \alpha$	a’lo
$0,8 \leq \alpha < 0,9$	yaxshi
$0,7 \leq \alpha < 0,8$	qoniqarli
$0,6 \leq \alpha < 0,7$	munozarali
$0,5 \leq \alpha < 0,6$	yomon
$\alpha < 0,5$	qoniqarsiz

Test topshiriqlarining ichki muvofiqligi har bir test topshirig‘iga berilgan to‘g‘ri javoblarning umumiy ball bilan korrelyatsiyasiga, talabgorlar olgan umumiy ballarning standart og‘ishiga, har bir test topshirig‘iga berilgan javoblarning standart og‘ishlari yig‘indisiga hamda test topshiriqlari va test topshiruvchilar soniga bog‘liq bo‘ladi. Bundan tashqari test topshiriqlarining ichki muvofiqligi nafaqat test topshiriqlarining sifatiga, balki talabgorlarning tayyorgarlik darajasining past yoki yuqoriligiga ham bog‘liqdir.

Test topshiriqlarining qiyinlik darajasi

3-jadvalda test natijalari bo‘yicha aniqlangan test topshiriqlarining qiyinlik darajalari keltirilgan. Bunda:

- ID – test topshiriqlarining ID raqami;
- N – ishtirokchilar soni;
- X_{max} – ishtirokchilarning barchasi test topshirig‘iga to‘g‘ri javob berganida hosil bo‘ladigan ballarning yig‘indisi;
- M – test topshirig‘iga to‘g‘ri javob berganlar soni;
- X_i – talabgorlarning ushbu test topshirig‘idan olgan ballari yig‘indisi;
- Ans – to‘g‘ri javob berganlar foizi;
- V – test natijalari bo‘yicha aniqlangan test topshiriqlarining qiyinlik darajalari.

To‘g‘ri javob berganlar foizi (Ans) quyidagi formula yordamida aniqlanadi:

$$Ans = \frac{X_i}{X_{max}} 100\% \quad \text{yoki} \quad Ans = \frac{M}{N} 100\%;$$

№	ID	N	X_{max}	M	X_i	Ans (foizda)	V
1	T18	4 506	5 857.8	4 499	5 848.7	99.84	1
2	T2	4 506	9 913.2	4 280	9 416.0	94.98	1
3	T7	4 506	9 913.2	4 220	9 284.0	93.65	1
4	T16	4 506	5 857.8	4 180	5 434.0	92.77	1
5	T24	4 506	5 857.8	4 106	5 337.8	91.12	1
6	T25	4 506	5 857.8	4 065	5 284.5	90.21	1
7	T6	4 506	5 857.8	4 044	5 257.2	89.75	1
8	T10	4 506	5 857.8	4 000	5 200.0	88.77	1
9	T11	4 506	9 913.2	3 912	8 606.4	86.82	1
10	T3	4 506	9 913.2	3 813	8 388.6	84.62	1
11	T20	4 506	5 857.8	3 811	4 954.3	84.58	1
12	T12	4 506	5 857.8	3 705	4 816.5	82.22	1
13	T1	4 506	5 857.8	3 644	4 737.2	80.87	1
14	T5	4 506	5 857.8	3 561	4 629.3	79.03	1
15	T26	4 506	9 913.2	3 469	7 631.8	76.99	1
16	T28	4 506	9 913.2	3 239	7 125.8	71.88	2
17	T31	4 506	9 913.2	3 141	6 910.2	69.71	2
18	T17	4 506	9 913.2	2 984	6 564.8	66.22	2
19	T19	4 506	9 913.2	2 941	6 470.2	65.27	2
20	T8	4 506	9 913.2	2 937	6 461.4	65.18	2
21	T13	4 506	9 913.2	2 912	6 406.4	64.62	2
22	O38B	4 506	7 660.2	2 908	4 943.6	64.54	2
23	T9	4 506	9 913.2	2 685	5 907.0	59.59	2
24	T14	4 506	9 913.2	2 609	5 739.8	57.90	2
25	T33	4 506	9 913.2	2 494	5 486.8	55.35	2
26	T32	4 506	9 913.2	2 356	5 183.2	52.29	2
27	T23	4 506	9 913.2	2 308	5 077.6	51.22	2
28	O38A	4 506	6 759.0	2 160	3 240.0	47.94	2
29	O37A	4 506	6 759.0	2 146	3 219.0	47.63	2
30	T21	4 506	9 913.2	2 120	4 664.0	47.05	2
31	T34	4 506	9 913.2	2 053	4 516.6	45.56	2
32	T29	4 506	9 913.2	2 007	4 415.4	44.54	2
33	T4	4 506	9 913.2	1 995	4 389.0	44.27	2
34	T15	4 506	9 913.2	1 944	4 276.8	43.14	2
35	O37B	4 506	7 660.2	1 928	3 277.6	42.79	2
36	O40A	4 506	6 759.0	1 852	2 778.0	41.10	2
37	T30	4 506	9 913.2	1 795	3 949.0	39.84	2
38	T27	4 506	9 913.2	1 743	3 834.6	38.68	2
39	O39A	4 506	6 759.0	1 648	2 472.0	36.57	2
40	T22	4 506	9 913.2	1 599	3 517.8	35.49	2
41	O41A	4 506	6 759.0	1 279	1 918.5	28.38	2
42	O42A	4 506	6 759.0	1 059	1 588.5	23.50	3
43	T35	4 506	9 913.2	1 056	2 323.2	23.44	3

44	O45B	4 506	7 660.2	851	1 446.7	18.89	3
45	O43B	4 506	7 660.2	824	1 400.8	18.29	3
46	O42B	4 506	7 660.2	736	1 251.2	16.33	3
47	O44A	4 506	6 759.0	711	1 066.5	15.78	3
48	O36A	4 506	6 759.0	589	883.5	13.07	3
49	O44B	4 506	7 660.2	578	982.6	12.83	3
50	O40B	4 506	7 660.2	462	785.4	10.25	3
51	O39B	4 506	7 660.2	415	705.5	9.21	3
52	O43A	4 506	6 759.0	297	445.5	6.59	3
53	O36B	4 506	7 660.2	247	419.9	5.48	3
54	O41B	4 506	7 660.2	192	326.4	4.26	3
55	O45A	4 506	6 759.0	58	87.0	1.29	3

Klassik test nazariyasiga ko‘ra:

- **75 %** va undan ko‘p to‘g‘ri javob berilgan test topshiriqlarini **1-darajali** (*oson*);
- **25 %** dan **75 %** gacha to‘g‘ri javob berilgan test topshiriqlarini **2-darajali** (*o‘rta*);
- **25 %** gacha to‘g‘ri javob berilgan test topshiriqlarini **3-darajali** (*qiyin*) deb hisbolash mumkin.

Demak, matematika fanidan o‘tkazilgan milliy sertifikat imtihonida talabgorlarga berilgan test topshiriqlarining **15 tasi** (27,27 %) oson, **26 tasi** (47,27 %) o‘rta, **14 tasi** (25,46 %) qiyin darajadagi test topshiriqlaridir (3-jadval).

II. Matematika fanidan milliy sertifikat imtihoni natijalarining zamonaviy test nazariyasi (Rash modeli) asosidagi statistik tahlili

Rash modeli ingliz tilida “item response theory” (IRT) deb nomlanadi. Bu “elementlarga javob nazariyasi” yoki “elementlarga reaksiya” degan ma’noni anglatadi. IRT faqat savollarga javoblarnigina baholamasdan, balki turli xildagi so‘rovnomalarni baholashda ham ishlatiladi. Shuning uchun ham bu nazariyada “**test topshirig‘i**” o‘rnida “**element (item)**” so‘zini ishlatish maqsadga muvofiq hisoblanadi. Shu sababli tahlilda “**test topshirig‘i**” tushunchasi va “**element**” tushunchasi bir xil ma’noda ishlatiladi.

Rash modeliga ko‘ra, dixotomik elementlarga individual javoblar ehtimoli shaxsning qobiliyat va element qiyinlik darajalari bilan aniqlanadi. Bu quyidagi matematik formula orqali ifodalanadi:

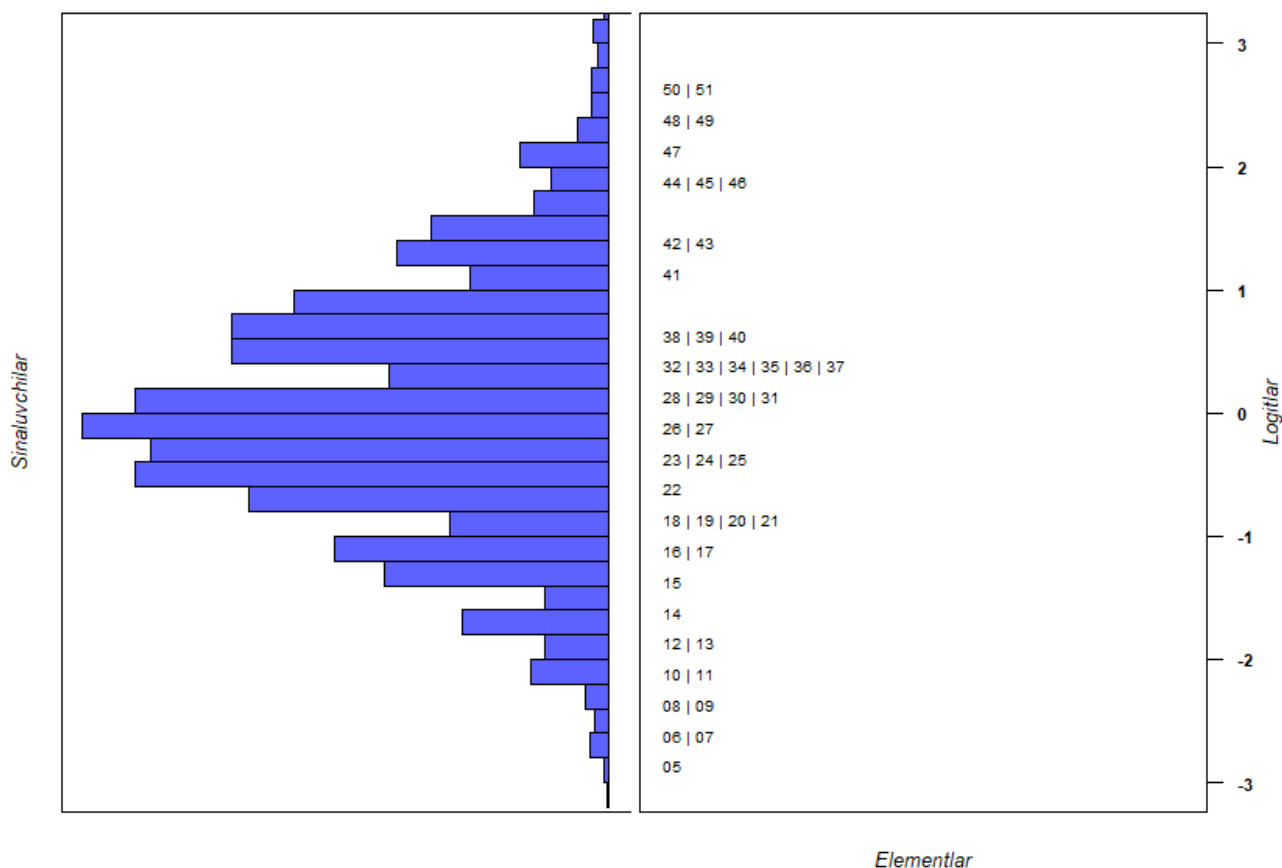
$$P(X_{is} = 1 | \theta_s, b_i) = \frac{e^{\theta_s - b_i}}{1 + e^{\theta_s - b_i}}$$

Bunda:

- $X_{is} = 1 - s$ o‘quvchining i elementga to‘g‘ri javob berish ehtimolligi;
- θ_s – qobiliyat o‘zgaruvchisi;
- b_i – test topshirig‘ining qiyinlik darajasi;
- e – natural logarifm asosi ($e = 2,7182818 \dots$).

IRT parametrlari tajriba va modellarni moslashtirish orqali topiladi. Bizning tahlilimizda tajriba bu matematika fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalaridir.

Rash modeli asosida aniqlangan qiyinlik darajalarining sinaluvchilar qobiliyatlariga qanchalik mosligini Rayt xaritasi yordamida tahlil qilish mumkin. Rayt xaritasi – test topshiriqlarining qiyinlik darajalari va sinaluvchilarning qobiliyat darajalarini o‘zaro mos kelishini aniqlovchi diagrammadir. Matematika fanidan o‘tkazilgan test sinovlari natijasi asosida chizilgan Rayt xaritasi 1-rasmda keltirilgan.



1-rasm. Qobiliyat va qiyinlik darajalarining mosligi (Rayt xaritasi)

Pedagogik o‘lchovlar nazariyasiga ko‘ra sinaluvchilar qobiliyatining 99,7 foizi (-3:3) logit birligi oralig‘ida, asosiy sinaluvchilar qobiliyatlari esa taqsimotning o‘rtasida (0 logit atrofida) bo‘ladi. Shu sababli test topshiriqlarining qiyinlik darajalari (-3:3) oralig‘ida, eng ko‘p qismi esa taqsimot markazida (0 atrofida) bo‘lishi lozim. Agar test topshiriqlarining ko‘proq qismi (-3:0) oralig‘ida bo‘lsa, test varianti osonroq, (0:3) oralig‘ida bo‘lsa, test varianti qiyinroq hisoblanadi. 1-rasmdan ko‘rish mumkinki, matematika fanidan test topshiriqlarining qiyinlik darajalari variantda yaxshi taqsimlangan.

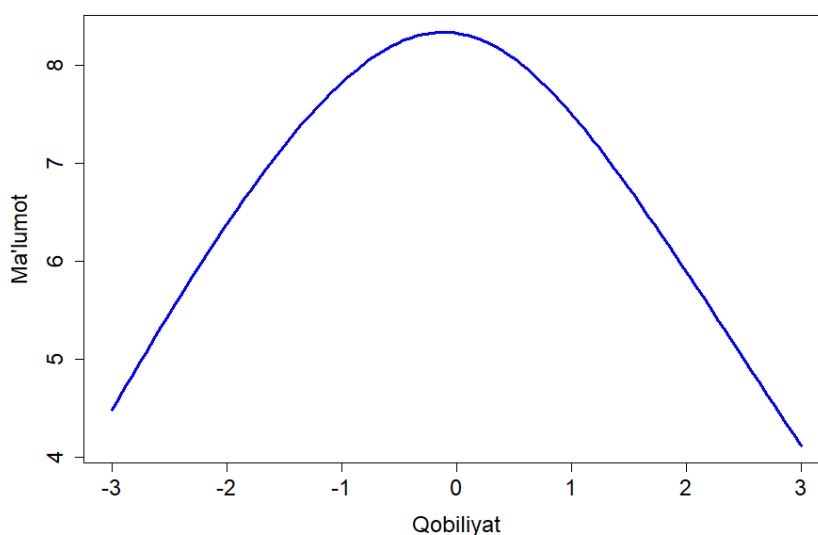
4-jadvalda matematika fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalari asosida bir parametrlil Rash modeli bo‘yicha aniqlangan test topshiriqlarining qiyinlik darajalari keltirilgan:

4- jadval

№	ID	b
1	T18	-7,026
2	T2	-3,423
3	T7	-3,155
4	T16	-3,002
5	T24	-2,758
6	T25	-2,639
7	T6	-2,582
8	T10	-2,468
9	T11	-2,262
10	T3	-2,057
11	T20	-2,053
12	T12	-1,857
13	T1	-1,753
14	T5	-1,618
15	T26	-1,477
16	T28	-1,158
17	T31	-1,031
18	T17	-0,837
19	T19	-0,786
20	T8	-0,781
21	T13	-0,751
22	O38B	-0,747
23	T9	-0,489
24	T14	-0,404
25	T33	-0,276
26	T32	-0,124
27	T23	-0,071
28	O38A	0,091

№	ID	b
29	O37A	0,106
30	T21	0,135
31	T34	0,209
32	T29	0,26
33	T4	0,274
34	T15	0,33
35	O37B	0,348
36	O40A	0,434
37	T30	0,499
38	T27	0,558
39	O39A	0,669
40	T22	0,727
41	O41A	1,128
42	O42A	1,436
43	T35	1,44
44	O45B	1,767
45	O43B	1,814
46	O42B	1,975
47	O44A	2,023
48	O36A	2,279
49	O44B	2,303
50	O40B	2,593
51	O39B	2,728
52	O43A	3,133
53	O36B	3,348
54	O41B	3,636
55	O45A	4,927

4-jadvaldan Rash modeli bo'yicha "O45A" ID raqamli test topshirig'i eng qiyin, "T18" ID raqamli test topshirig'i esa eng oson ekanligini ko'rish mumkin.



2-rasm. Test ma'lumoti chizig'i

2-rasmda test ma'lumoti chizig'i (TMCh) keltirilgan. Testning umumiy ma'lumot miqdori 55,00 ga teng bo'lib, shundan (-3:3) oralig'idagi qobiliyatga ega bo'lganlar uchun ma'lumot miqdori 40.53 (73.8 foiz) ga teng. (-3:0) va (0:3) oraliqlardagi qobiliyatga ega bo'lganlar miqdori mos ravishda 20.78 (37.84 %) va 19.75 (35.96 %)ga to'g'ri keladi. Pedagogik o'lchovlar nazariyasiga ko'ra (-3:0) va (0:3) oraliqlardagi ma'lumot miqdorlari bir-biriga qanchalik yaqin bo'lsa, test variantining statistik ko'rsatkichlari shunchalik yaxshi bo'ladi. Bizda (-3:0) oraliqdagi ma'lumot miqdori (0:3) oraliqdagi ma'lumot miqdoridan biroz (1,88 %) ko'proq bo'lib, bu test varianati sinaluvchilar uchun me'yorga nisbatan (juda kichik miqdorda – variantning o'rtacha qiyinlik darajasi -0,12 ga teng) osonroq ekanini anglatadi. Umuman olganda esa test variantidagi test topshiriqlarining qiyinlik darajalari bo'yicha taqsimoti ideal holatga juda yaqin.

III. Xalqaro imtihonlarda vaqt me'yori

SAT (*Scholastic Aptitude Test*) matematika imtihonida **58 ta** test toshirig'i uchun **80 daqiqa** (*1 soat-u 20 daqiqa*) vaqt beriladi (*1 ta test topshirig'iga taxminan 1 daqiqa-yu 23 soniya*). Bunda ularning 45 tasi yopiq turdagi, 13 tasi ochiq turdagi test topshiriqlaridir.

GMAT (*Graduate Management Admission Test*) matematika imtihonida jami **31 ta** test topshirig'i uchun **62 daqiqa** (*1 soat-u 2 daqiqa*) vaqt ajratiladi (*1 ta test topshirig'iga taxminan 2 daqiqa*).

Milliy sertifikat imtihonida esa **35 ta** yopiq va **10 ta** ochiq turdagi (*har biri A va B qismlarga ajratilgan*) test topshirig'i uchun **150 daqiqa** (*2 soat-u 30 daqiqa*) vaqt beriladi. Ochiq turdagi test topshiriqlari 2 qismga ajratilgani uchun jami test topshiriqlari sonini **55 ta** deb qaraydigan bo'lsak, har bir test topshirig'iga taxminan **2 daqiqa-yu 44 soniya** to'g'ri keladi.

Tahlillardan ko'rish mumkinki, matematika fanidan milliy sertifikat imtihoniga ajratilgan vaqt va test topshiriqlarining qiyinlik darajalari xalqaro standartlarga va sinaluvchilarning qobiliyatiga to'la mos keladi.