

MATEMATIKA FANIDAN O'TKAZILGAN DIAGNOSTIK TEST SINOV NATIJALARINING STATISTIK TAHLILLARI

Q.A. Amonov, A.A. Baratov

*O'zbekiston Respublikasi Oliy ta'lim, fan va innovatsiyalar vazirligi huzuridagi
Bilim va malakalarni baholash agentligi huzuridagi Ilmiy – o'quv amaliy markazi,
100084, Toshkentsh, Bog'ishamol k., 12*

Qisqacha mazmuni. Ushbu maqolada matematika fanidan o'tkazilgan diagnostika test sinovlari natijalarining klassik va zamonaviy test nazariyalari asosidagi statistik tahlili bayon etilgan. Statistik tadqiqot natijalariga ko'ra, test variantining ishonchlilik ko'effitsiyenti – Kronbax alfa ko'effitsiyenti 0,94 ga teng ekanligi aniqlandi. Zamonaviy test nazariyasi bilan aniqlangan test topshiriqlarining qiyinlik darajasi hisoblanganda o'ta qiyin va o'ta oson test topshiriqlari mavjudligi aniqlandai. Sinaluvchilarning test topshiriqlariga bergan javoblari tahlil qilinganda matematika fanidagi fizik jarayonlarning matematik yechimlari, murakkab tenglamalar va tengsizliklar, murakkab funksiyalar, matematik analiz asoslari va geometriya fanining planimetriya bo'limlaridan o'zlashtirish past ekanligi aniqlandi.

Kalit so'zlar: Test topshiriqlari, Moda, Mediana, Standart tafovut, Dispersiya, Kronbax alfa ko'effitsiyenti, validlik, qiyinlik darajasi, Rash modeli, Rayt xaritasi, qobiliyat darajalari.

I. Kirish

Maktab va oliy ta'lim sohasida amalga oshirilayotgan islohotlar – yetuk salohiyatli yoshlarni tarbiyalash va bilim berish, raqobatbardosh kadrlar tayyorlash va yangi pedagogik texnologiyalarni amalga joriy qilishni talab qiladi. Ta'lim sifatini oshirishning muhim omillaridan yana biri bu baholashni takomillashtirish, o'quvchi va talabalarning bilimlarini o'lchashning ratsional, obyektiv va shaffof shakllarini yo'lga qo'yishdir. O'rta va oliy ta'lim muassasalarida baholashning nomutanosibliigi – ta'lim

tizimida baholashga turlicha yondashuv mavjud ekanligini anglatadi. Ta'lim jarayonida o'quvchilarni bilimlarini to'g'ri o'lchash nafaqat ularning shaxs sifatida o'z qobiliyati va intilishlari bo'yicha kasb va hunar tanlashida, balki oliy o'quv yurtlariga kirish uchun ham muhim omil hisoblanadi.

O'quvchilarning haqiqiy bilim va ko'nikmalarini aniqlash uchun o'tkaziladigan pedagogik o'lchashlarda standart testlar va bir o'lchovli

shkalalardan foydalanish talab etiladi [1-3].

Standartlashtirilgan test sinovlarini o'tkazish, natijalarini ilmiy tahlil qilish bilan turli xil guruhlardagi sinaluvchilarning qobiliyat darajalarini solishtirish, o'quvchi-larning o'quv dasturi va fan mavzularining o'zlashtirish darajasini aniqlash, pedagoglar va ta'lim muassasalari faoliyati samaradorligini baholash, shuningdek, ularning yutuq va kamchilliklari haqida ham ma'lumotlar olish mumkin [4, 5]. Standartlik nafaqat test variantiga nisbatan ishlatiladi, balki butun jarayonga tegishli bo'lib test topshiriqlarining sifati, test o'tkazish shart-sharoitlarining barcha sinaluvchilar uchun bir xilligi va baholash jarayoni asosli ravishda mezonlashtirilganligi va sinaluvchilarga beriladigan ballarni oldindan aniqlangan tartib asosida talqin qilishni anglatadi [6, 7].

Bunda, nafaqat test varianti yopiq test topshiriqlaridan, balki barcha turdagi test topshiriqlaridan iborat test varianti ham standartlik shartlari doirasida bo'lishi maqsadga muvofiqdir.

Pedagogik o'lchovlarda sinaluvchilarning turli xil qobiliyat darajalariga mos bo'lgan turli xil qiyinlik darajadagi test topshiriqlari mavjud bo'lishi ham maqsadga muvofiqdir [8]. Bu o'z navbatida sinaluvchilarning bilimlarini samarali baholash hamda quyi va yuqori

darajadagi o'zlashtiruvchi sinaluvchilarni yaxshi ajratish imkonini beradi. Odatda testlar turli maqsadda ishlab chiqiladi va qo'llaniladi. Bulardan eng muhimlari maqsadga ko'ra me'yorga asoslangan, mezonga mo'ljallangan va bashorat qilish uchun foydalaniladigan testlardir [9].

Me'yorlarga asoslangan test topshiriqlari yordamida sinaluvchilarning natijalari shu yoshdagi, o'rganishning shu bosqichidagi yoki qobiliyat darajalari bir biriga yaqin bo'lgan test topshiruvchilarning yirik guruhi natijalari bilan solishtiriladi va ularning reyting o'рни aniqlanadi. Mezonga asoslangan test topshiriqlari yordamida esa sinaluvchilarning natijalari oldindan belgilangan biror bir mezon talablariga javob berish imkoniyati baholanadi. Jumladan sinaluvchilarning natijalarini tahlil qilish orqali ta'lim dasturini qay darajada o'zlashtirganligini yoki fan mavzularini qay darajada o'zlashtirmaganligini ham aniqlash mumkin.

Ushbu maqolada matematika fani bo'yicha o'tkazilgan diagnostik test sinovi natijalari va test sinovlarida foydalanilgan test topshiriqlari sifati tahlil qilindi. Diagnostik test sinovlarida jami 79 nafar talabgor ishtirok etdilar. Test sinovlarida jami 45 ta test topshiriqlaridan foydalanildi. Jumladan test variantida 35 tasi yopiq turdagi test topshiriqlari va 36-45- test topshiriqlarining A va B qismlarga

ajratilishi hisobiga 20 ta ochiq turdagi test topshiriqlaridan foydalanildi.

II. Test sinovlari natijalarining klassik test nazariyasi bo'yicha tahlili

Pedagogik o'lchovlar nazariy asoslariga ko'ra, tuzilgan test topshiriqlarining mazmuni shu fan mutaxassis-ekspertlari tomonidan ko'rib chiqilgandan so'ng aprobatsiya test sinovlariga qo'yiladi va uning natijalari asosida test topshiriqlarining sifatini tashxislash uchun statistik tavsiflari aniqlanadi. Statistik tavsiflar orqali test topshiriqlariga qo'yilgan asosiy ko'rsatkichlar aniqlanadi, bu ko'rsatkichlar – test topshirig'ining qiyinlik darajasi, test ballarining dispersiyasi (test topshirig'ining boshqa test topshiriqlari bilan farqlanishi, o'zgaruvchanligi), shuningdek, umumiy ballar yig'indisi bilan korrelyatsiyasidan iborat bo'ladi. Test topshirig'ining qiyinlik darajasini aniqlash usullaridan biri test topshirig'ini empirik sinovdan o'tkazib, to'g'ri javoblar salmog'ini aniqlashdan iboratdir. Test ballari (yoki to'g'ri javoblar)ning dispersiyasi test topshiruvchilarning tayyorgarlik darajasini aniqlashga, biladiganlarni bilmaydiganlardan ajratishga imkon beradi. Pedagogik o'lchovlarda klassik test nazariyasining asosiy statistik tavsiflari qatoriga o'rta qiymat, gistogrammani qurish, moda va mediana kabi ko'rsatkichlarni hisoblash hamda test ballarining

umumiy dispersiyasi (standart tafovut) ko'rsatkichi ham kiradi [10-15].

Test ballari (yoki to'g'ri javoblar)ning o'rta arifmetik qiymati fanlar, ta'lim muassasalari va boshqa muhim belgilar kesimida aniqlanadi. Bu ko'rsatkich test ballari o'rtasidagi tafovutni umumlashtiradi, ularga xos bo'lgan qonuniyatni ochib beradi. Test sinovi natijalari asosida aniqlangan test ballari taqsimotining gistogrammasi quriladi va u normal taqsimotga yaqin yoki uzoqligi baholanadi. Test ballarining eng ko'p takrorlanadigan qiymati statistikada moda, o'sish 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 bir-biridan qanchalik ko'p farq qilsa, ballar taqsimoti normal taqsimotdan shuncha uzoqda bo'ladi.

1-jadvalda matematika fanidan test sinovi ballarining o'rta qiymati, medianasi, modasi, standart xatoligi, dispersiyasi, diapazoni, maksimum, minimum qiymatlari keltirilgan.

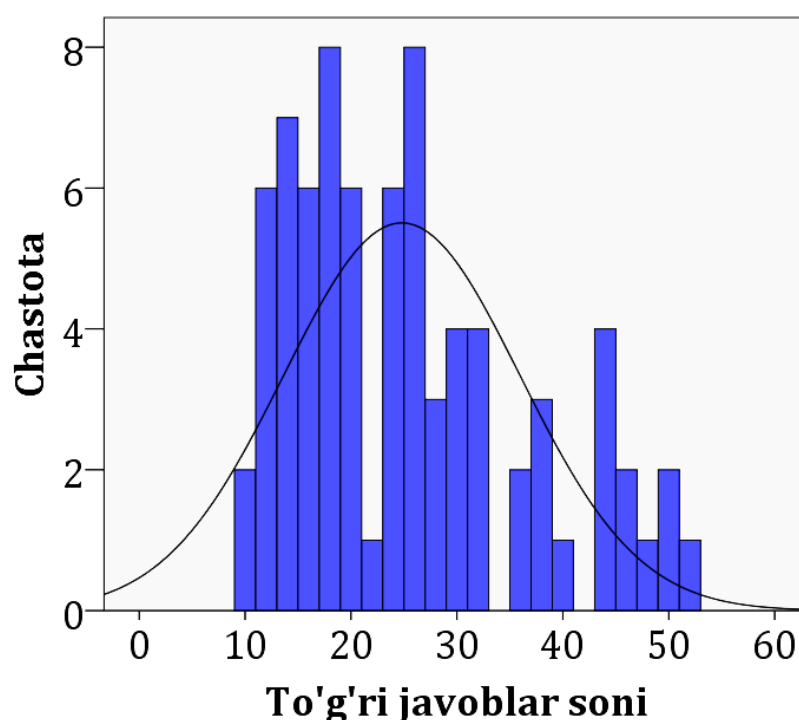
Matematika fanidan o'tkazilgan diagnostik test sinovi natijalarining tavsif statistikasi ma'lumotlari

1	Test topshiriqlari soni	55
2	O'rta qiymat	24,79
3	Moda	14
4	Mediana	23
5	Dispersiya	124,53
6	Standart tafovut	11,16
7	O'lchashning standart xatoligi	1,3
8	Kronbax alfa koeffitsiyenti	0,94
9	Diapazon	42
10	Minimum	10
11	Maksimum	52

Statistik tadqiqot natijalariga ko'ra, matematika fanidan test variantining ishonchlilik koeffitsiyenti, ya'ni Kronbax alfa koeffitsiyenti 0,94 ga tengligi aniqlandi. Kronbax alfa koeffitsiyentining 0,9 va undan kattaligi ushbu test sinovlari uchun tanlab olingan test variantlarining ishonchliligi a'lo darajada ekanligini ko'rsatmoqda [16].

1-rasmda matematika fanidan o'tkazilgan diagnostik test sinovlari natijalarining gistogrammasi

keltirilgan. Rasmdan ko'rinib turibdiki, test sinovi natijalari bo'yicha test topshiriqlarining individual ballari taqsimoti bir-biriga yaqin, ammo normal taqsimotga nisbatan chap tomonga siljigan. Gistogrammaning o'ng tomoni normal taqsimotga nisbatan keskin farqlanishi va test topshiriqlarining individual ballari taqsimoti bilan farqlanishi test sinovlarida sinaluvchilarning sonini kamligi bilan izohlash mumkin.



1-rasm. Matematika fanidan o'tkazilgan diagnostik test sinovi natijalarining gistogrammalari

Test topshiriqlarining qiyinlik darajalari tahlil qilinganda test sinovi natijalari bo'yicha 55 ta test topshiriqlaridan 11 tasi (20 foiz) 1-qiyinlik darajasidagi test topshiriqlaridan, 29 tasi (52,75 foiz) 2-qiyinlik darajasidagi test topshiriqlaridan va 15 tasi (27,27 foiz) 3-qiyinlik darajasidagi test topshiriqlaridan iborat ekanligi

aniqlandi (2-jadval). Ushbu test variantidagi test topshiriqlarining qiyinlik darajalari (1-, 2- va 3-daraja bo'yicha) normada taqsimlanganligi aniqlandi, biroq 1-o'rindagi (ID raqami-5000002) test topshirig'ining qiyinlik darajalari juda past va 55-o'rindagi (ID raqami-5000054) test topshirig'ining qiyinlik darajasi esa juda yuqori ekanligi aniqlandi.

2-Jadval

Matematika fanidan o'tkazilgan diagnostik test sinovi natijalari bo'yicha test topshiriqlarining aniqlangan qiyinlik darajalari

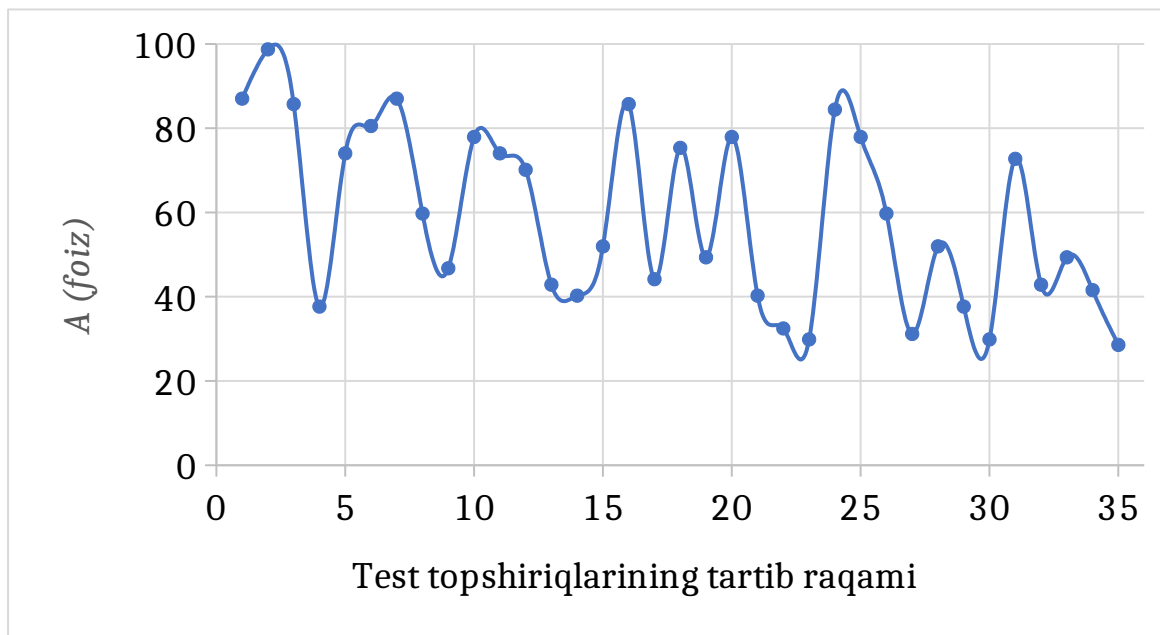
Nº	ID	X_i	A(foiz)	V
1	5000002	76	98,7	1
2	5000001	67	87,01	1
3	5000007	67	87,01	1
4	5000003	66	85,71	1
5	5000016	66	85,71	1

6	5000024	65	84,42	1
7	5000006	62	80,52	1
8	5000010	60	77,92	1
9	5000020	60	77,92	1
10	5000025	60	77,92	1
11	5000018	58	75,32	1
12	5000005	57	74,03	2
13	5000011	57	74,03	2
14	5000031	56	72,73	2
15	5000012	54	70,13	2
16	5000008	46	59,74	2
17	5000026	46	59,74	2
18	5000041	46	59,74	2
19	5000015	40	51,95	2
20	5000028	40	51,95	2
21	5000019	38	49,35	2
22	5000033	38	49,35	2
23	5000009	36	46,75	2
24	5000017	34	44,16	2
25	5000013	33	42,86	2
26	5000032	33	42,86	2
27	5000034	32	41,56	2
28	5000039	32	41,56	2
29	5000040	32	41,56	2
30	5000014	31	40,26	2
31	5000021	31	40,26	2
32	5000004	29	37,66	2
33	5000029	29	37,66	2
34	5000038	27	35,06	2
35	5000022	25	32,47	2
36	5000027	24	31,17	2
37	5000042	24	31,17	2
38	5000023	23	29,87	2
39	5000030	23	29,87	2
40	5000035	22	28,57	2
41	5000044	18	23,38	3
42	5000046	18	23,38	3

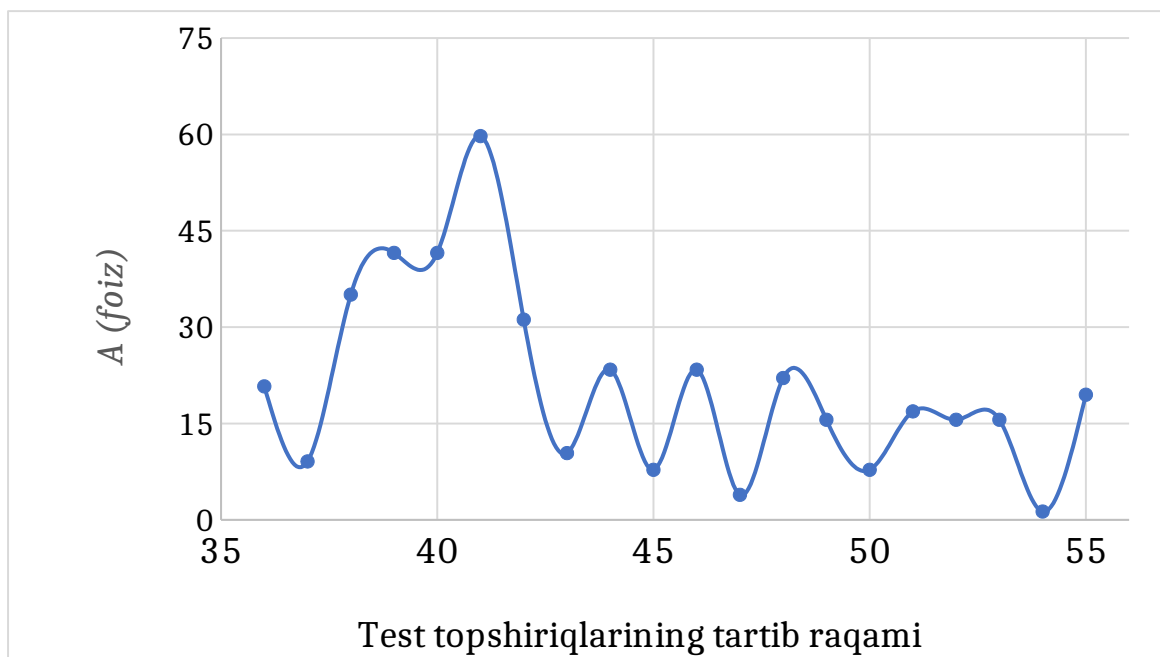
43	5000048	17	22,08	3
44	5000036	16	20,78	3
45	5000055	15	19,48	3
46	5000051	13	16,88	3
47	5000049	12	15,58	3
48	5000052	12	15,58	3
49	5000053	12	15,58	3
50	5000043	8	10,39	3
51	5000037	7	9,09	3
52	5000045	6	7,79	3
53	5000050	6	7,79	3
54	5000047	3	3,9	3
55	5000054	1	1,3	3

Matematika fanidan o'tkazilgan diagnostik test sinovida foydalanilgan ochiq (2-rasm) va yopiq (3-rasm) test topshiriqlarining sinaluvchilarga nisbatan bajarilganlik foizlari tahlil qilinganda bir qancha minimum nuqtalar aniqlandi. Ushbu nuqtalarga mos keluvchi test topshiriqlarining test tafsilotlarida keltirilgan fan mavzulari bilan taqqoslanganda sinaluvchilar tomonidan matematika fanidagi fizik jarayonlarning matematik yechimlari, murakkab tenglamalar va tengsizliklar, murakkab funksiyalar, matematik analiz asoslari va geometriya fanining planimetriya bo'limlaridan o'zlashtirish pastligi aniqlandi. Sinaluvchilar tomonidan ochiq test topshiriqlarini bajarish ko'nikmasi ham past ekanligi va ularning to'g'ri yechilgan test topshiriqlarining o'rtacha qiymati 14 ga tengligi aniqlandi. Matematika

fanidan o'tkazilgan diagnostik test sinovida test topshiriqlarining ichki muvofiqligi har bitta test topshirig'iga berilgan to'g'ri javoblarning umumiy ball bilan korrelyatsiyasiga, sinaluvchilar olgan umumiy ballarning standart og'ishiga, har bitta 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 sinaluvchilarning tayyorgarlik darajasining past yoki yuqoriligiga ham bog'liqdir. Har bir test topshirig'iga berilgan javoblarning umumiy test bali bilan korrelyatsiyasi test topshiriqlarining diskriminat-siyasi (ajratish darajasi)ni bildiradi.



2- rasm. Yopiq test topshiriqlarining bajarilganlik foizining test topshiriqlarining tartib raqamiga bog'liqligi.



3- rasm. Ochiq test topshiriqlarining bajarilganlik foizining test topshiriqlarining tartib raqamiga bog'liqligi.

Umumiy ball bilan korrelyatsiya koeffitsiyenti (UBBKK) qiymati 2-qiyinlik darajasidagi test topshiriqlari uchun 0,5 va undan katta bo'lsa, 1- va 3-qiyinlik darajasidagi test topshiriqlari uchun esa 0,25 va undan

katta bo'lsa, valid hisoblanadi. Umumiy ball bilan korrelyatsiya koeffitsiyenti qiymati manfiy bo'lgan test topshiriqlari esa variantdan chiqariladi. Aks holda bilim darajalari past bo'lgan sinalluvchilar g'olib bo'lib, bilim

darajalari yuqori bo'lgan sinaluvchilar
test topshiriqlarini yechishda noto'g'ri

javobni tanlaydilar yoki ularni o'tkazib
yuboradilar.

3-Jadval

Nº	ID	Har bitta test topshirig'i bilan UBBKK
1	5000002	0,132
2	5000001	0,108
3	5000007	0,341
4	5000003	0,324
5	5000016	0,083
6	5000024	0,367
7	5000006	0,497
8	5000010	0,419
9	5000020	0,439
10	5000025	0,411
11	5000018	0,405
12	5000005	0,464
13	5000011	0,189
14	5000031	0,349
15	5000012	0,415
16	5000008	0,262
17	5000026	0,577
18	5000041	0,479
19	5000015	0,282
20	5000028	0,601
21	5000019	0,403
22	5000033	0,677
23	5000009	0,431
24	5000017	0,529
25	5000013	0,546
26	5000032	0,634
27	5000034	0,529
28	5000039	0,536
29	5000040	0,565
30	5000014	0,577
31	5000021	0,749

32	5000004	0,389
33	5000029	0,619
34	5000038	0,723
35	5000022	0,468
36	5000027	0,422
37	5000042	0,794
38	5000023	0,138
39	5000030	0,424
40	5000035	0,559
41	5000044	0,810
42	5000046	0,650
43	5000048	0,628
44	5000036	0,587
45	5000055	0,459
46	5000051	0,518
47	5000049	0,696
48	5000052	0,680
49	5000053	0,631
50	5000043	0,394
51	5000037	0,564
52	5000045	0,512
53	5000050	0,425
54	5000047	0,319
55	5000054	0,282

3-jadvalda test sinovi natijalari tahlili asosida olingan test topshiriqlarining umumiy ball bilan korrelyatsiya koeffitsiyenti qiymatlari qiyinlik darajalari ortib borishi tartibida keltirildi.

Har bitta test topshirig'i bilan umumiy ball korrelyatsiya koeffitsiyenti qiymati 0,25 dan kichik (3-jadvalda ajratib ko'rsatilgan ID raqamlari-5000002, 5000001,

5000016, 50000011 va 5000023) bo'lgan test topshiriqlarini o'rganib chiqib kerakli o'zgarishlar qilish, lozim bo'lsa test topshiriqlar bazaga qo'shmaslik tavsiya etiladi.

Har bitta test topshirig'i bilan umumiy ball korrelyatsiya koeffitsiyenti qiymatlari 0,5 dan kichik bo'lgan test topshiriqlari esa, o'rganib chiqilib qiyinlik darajalarini hisobga olgan holda kerakli o'zgarishlar qilish

tavsiya qilinadi. Har bitta test topshirig'i bilan umumiy ball korrelyatsiya koeffitsiyenti qiymatlari

0,5 dan katta bo'lgan test topshiriqlari testologiya qoidalariga ko'ra valid hisoblanadi.

III. Test sinovlari natijalarining Rash modeli asosida tahlili

Jahonning ko'plab rivojlangan davlatlarida pedagogik o'lchov vositalarining sifatini aniqlashda Rash modeli asosida matematik-statistik tadqiqotlar olib borilmoqda [17-19]. Rash modeli asosidagi matematik-statistik tahlillar o'lchanayotgan xususiyatlarga obyektiv va xolis yondashuvni ta'minlab beradi. Test topshiriqlarining natijalarini Rash modeli asosida tahlil qilish uchun R dasturida ishlatiladigan dexter to'plamidan foydalanildi [20]. Rash modelining muhim xususiyati o'lchovning nimaligini, ta'lim tizimida o'lchovlarni qanday sifatli amalga oshirish imkoniyatini beradi [18, 19]. Test topshiruvchilarning yashirin qobiliyati va test topshiriqlarning qiyinlik darajasi kabi parametrlarini

Rash modeli bilan ochib berish mumkin. Bu ikki kattalikdan birinchisi o'zgaruvchi sifatida, ikkinchisi esa parametr sifatida kiritiladi. Chunki test topshiruvchilarning qobiliyati (bilimi) bu modelda elementlarga ya'ni topshiriqlarga berilgan javoblarga qarab belgilanadi, shuning uchun topshiriqlarning qiyinlik darajasini parametr sifatida qarash qulay. Rash modeliga ko'ra, dixotomik elementlarga individual javoblar shaxsning qobiliyat darajasi va element qiyinligi bilan aniqlanadi. Ma'lum bir qobiliyatga ega bo'lgan shaxsning ma'lum bir qiyinlikdagi elementga to'g'ri javob berish ehtimolligini aniqlaydi. Bu quyidagi matematik formula orqali ifodalanadi [17-19]:

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

u yerda $X_{is} = 1$, s -o'quvchining i -elementga to'g'ri javob berish ehtimolligi, θ_s -qobiliyat o'zgaruvchisi, b_i -topshiriq qiyinlik darajasi. Ushbu (1) ifodaga asosan to'g'ri javoblar ehtimolligi qobiliyat va qiyinlik kabi o'zgaruvchilarining farqiga bog'liq, bu esa qobiliyat va qiyinlik o'zgaruvchilariga ixtiyoriy o'zgarmas

son qo'shilganda ehtimollik o'zgarmasligini bildiradi. [21-23].

Ushbu modelga asosan sinaluvchilarning test natijalari asosida test topshiriqlarining qiyinlik darajalari aniqlandi (3- jadval). 3- jadvaldan test sinovlarida qiyinlik darajalari bo'yicha test topshiriqlari (-3: 3) logit birligi oralig'iga tushmagan oson (ID raqamlari- **5000018**,

5000002 va **5000007** bo'lgan test topshiriqlari) va qiyin (ID raqamlari **5000050**, **5000037**, **5000047** va

5000054 bo'lgan test topshiriqlari) bo'lgan test topshiriqlari mavjudligi kuzatildi.

4-Jadval

Rash modeli bilan aniqlangan qiyinlik darajalari

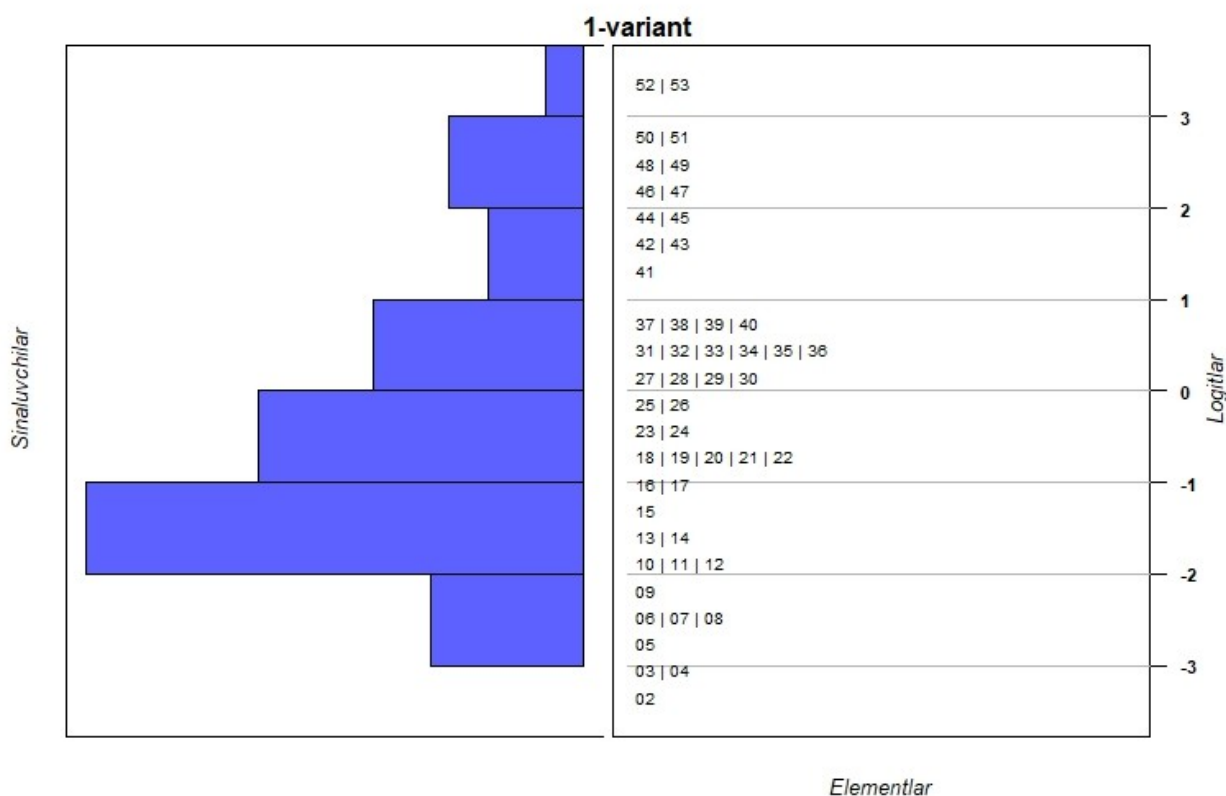
Nº	ID	b	Nº	ID	b
1	5000018	-5,65	29	5000038	0,20
2	5000002	-3,37	30	5000021	0,22
3	5000007	-3,07	31	5000034	0,29
4	5000016	-2,92	32	5000029	0,35
5	5000024	-2,68	33	5000004	0,36
6	5000025	-2,55	34	5000015	0,41
7	5000006	-2,50	35	5000039	0,43
8	5000010	-2,38	36	5000044	0,53
9	5000011	-2,18	37	5000030	0,59
10	5000003	-1,99	38	5000027	0,65
11	5000020	-1,98	39	5000042	0,76
12	5000012	-1,77	40	5000022	0,82
13	5000001	-1,69	41	5000046	1,23
14	5000005	-1,54	42	5000035	1,53
15	5000026	-1,39	43	5000048	1,54
16	5000028	-1,07	44	5000055	1,87
17	5000031	-0,96	45	5000051	1,92
18	5000017	-0,74	46	5000049	2,09
19	5000008	-0,70	47	5000052	2,13
20	5000019	-0,70	48	5000036	2,38
21	5000041	-0,67	49	5000053	2,42
22	5000013	-0,66	50	5000045	2,72
23	5000009	-0,40	51	5000043	2,85
24	5000014	-0,31	52	5000050	3,27
25	5000033	-0,19	53	5000037	3,48
26	5000032	-0,04	54	5000047	3,79
27	5000023	0,03	55	5000054	5,10
28	5000040	0,18			

Rash modeli asosida aniqlangan test topshiriqlarining qiyinlik darajalari va sinaluvchilarning qobiliyat darajalarining o'zaro mos kelishini Rayt xaritasi yordamida tahlil qilish mumkin[24-26].

4-rasmlarda matematika fanidan o'tkazilgan diagnostik test sinovi natijalari asosida olingan Rayt xaritasi keltirilgan. 4-rasmdan qobiliyat darajalari ham test topshiriqlari qiyinlik darajalari ham (-3:3) logit birligidan katta oraliqda tashqarida taqsimlanganligi aniqlandi. Qiyinlik darajasi juda past va juda yuqori test topshiriqlaridan juda kam miqdordagi ma'lumot olinadi, shuning uchun

bunday test topshiriqlari o'rniga mos ravishda (-3:3) oraliqdagi test topshiriqlarini kiritish maqsadga muvofiq bo'ladi.

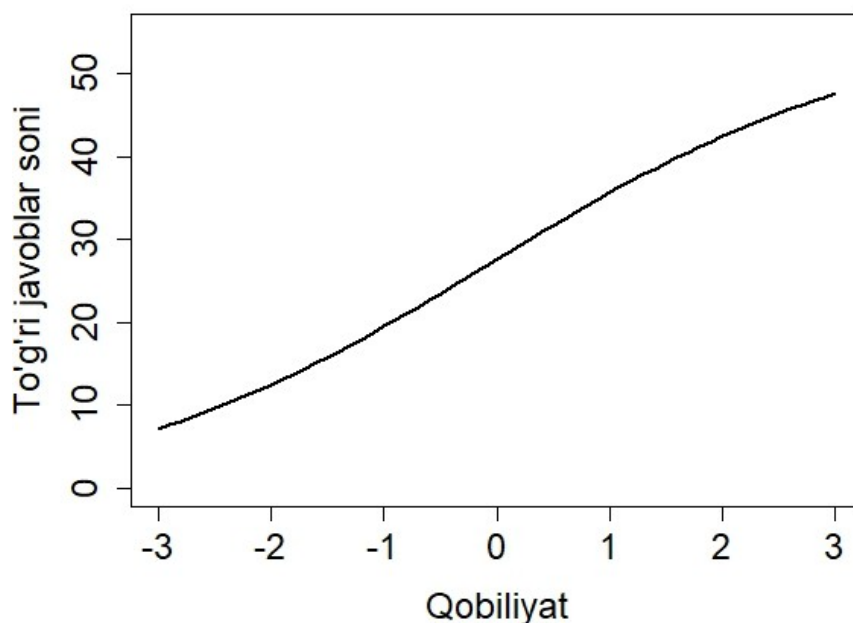
Rayt xaritasi asosida aniqlangan taqsimotdagi bo'sh joylarga va bir xil qiyinlikdagi test topshiriqlari o'rniga bo'sh joylarga mos keladigan qiyinlik darajasidagi test topshiriqlaridan qo'yish taqsimotni yanada yaxshilash imkonini beradi. 4-rasmdan ham talabgorlarning qobiliyat darajalari (-2,46: 4,38) logit birligi oralig'ida, test topshiriqlarining qiyinlik darajalari esa (-5,65: 5,10) oralig'ida ekanligi ko'rinadi.



4-rasm. Test sinovi natijalari asosida aniqlangan qobiliyat va qiyinlik darajalarining mosligi (Rayt xaritasi)

5-rasmda matematikada fanidan o'tkazilgan diagnostik test sinovlarida foydalanilgan test variantining test xarakteristikasi chizig'i (TXCh) ko'rsatilgan. Test variantining test xarakteristikasi chizig'i qanchalik tek bo'lsa test topshiriqlari shunchalik osonligini, aksincha bo'lsa shunchalik qiyinligini anglatadi. Umuman olganda ushbu holatda test variantlarining sinaluvchilarni qobiliyatlariga nisbatan qiyinligini yoki osonligini aniqlab bo'lmaydi. Zamonaviy test nazariyasi amaliyotda va xalqaro tajribalarda [27]

kalibrovkalangan test topshiriqlaridan iborat bazalardan foydalanish orqali baholashlardan foydalanish maqsadga muvofiq bo'ladi. Bunday holatda test topshiriqlarining qiyinlik darajalari va qobiliyat darajalari bir xil shkalada bo'ladi va ular test natijalarini hisoblash jarayonidayoq aniqlanadi. Shuningdek, bunday baholashda statistik tahlillar izchilligi ta'minlanadi. Bu esa test natijalarining ishonchligi va validligi haqida to'g'ri xulosalar chiqarish imkonini beradi.



5-rasm. Test sinovida ishlatilgan variantlardan aniqlangan test xarakteristikasi chizig'i

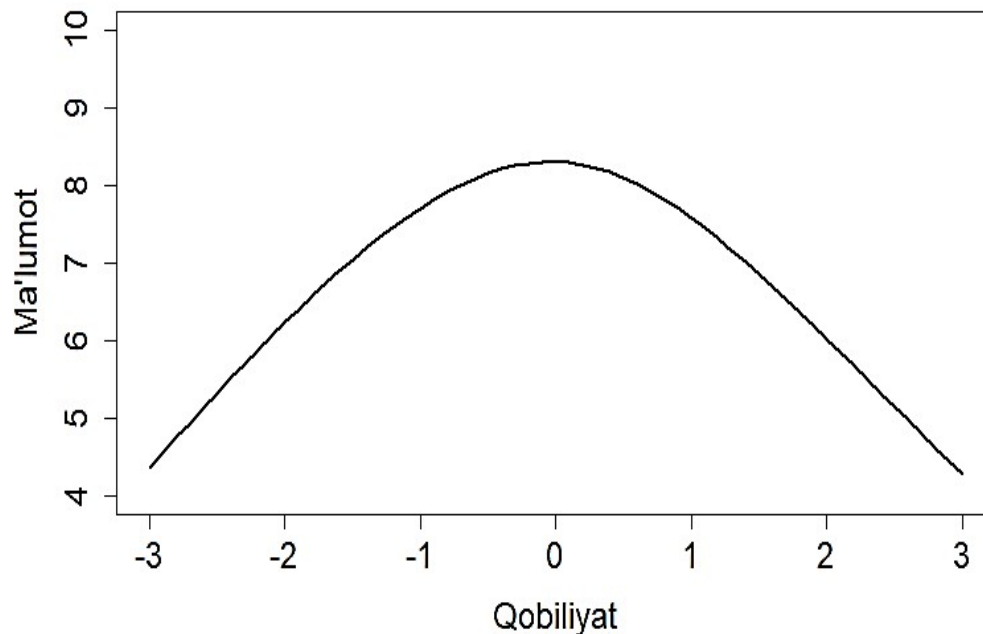
6-rasm test sinovida ishlatilgan variantlardan aniqlangan test ma'lumoti chizig'i ko'rsatilgan. Qiyinlik darajasi juda past (3- jadval, ID raqamlari- 5000018 va 5000002 bo'lgan test topshiriqlari) va juda yuqori (3-jadval, ID raqamlari 5000050, 5000037, 5000047 va

5000054 bo'lgan test topshiriqlari) test topshiriqlaridan juda kam miqdordagi ma'lumotlar olinadi, shuning uchun bunday test topshiriqlari o'rniga (-3: 3) oraliqdagi test topshiriqlarini kiritish maqsadga muvofiq bo'ladi.

Ma'lumot chizig'i cho'qqisi nolga nisbatan deyarli simmetrik joylashgan

bo'lib, bu esa o'z navbatida ushbu test varianti yordamida qobiliyat darajasi bir xil bo'lgan sinaluvchilar to'g'risida ko'proq ma'lumot berishini anglatadi. Agar Ma'lumot chizig'i cho'qqisining nolga nisbatan o'ng tomonga surilishi,

qobiliyat darajasi yuqoriroq bo'lgan, yoki nolga nisbatan chap tomonga surilishi qobiliyat darajasi past bo'lgan talabgorlar to'g'risida ma'lumotlar beradi.



6-rasm. Test sinovida ishlatilgan variantlardan aniqlangan test ma'lumoti chizig'i

Qiyinlik darajasi juda past bo'lgan test topshiriqlarini o'rniga -3 logit birligi atrofida test topshiriqlarini va qiyinlik darajasi juda yuqori bo'lgan test topshiriqlarining o'rniga 3 logit birligi atrofida test topshiriqlarini kiritib, yuqori va past qobiliyat darajalaridan olinadigan ma'lumot miqdori orasidagi tafovutni yanada kamaytirish mumkin [24-26].

Klassik test nazariyasi bilan har bitta sinaluvchi bitta test topshirig'iga bergan javobini ularning to'plagan umumiy ballari bilan korrelyatsiyasini va shu test topshirig'i natijasi chiqarilganda, qolgan test topshiriqlaridan hosil bo'lgan umumiy

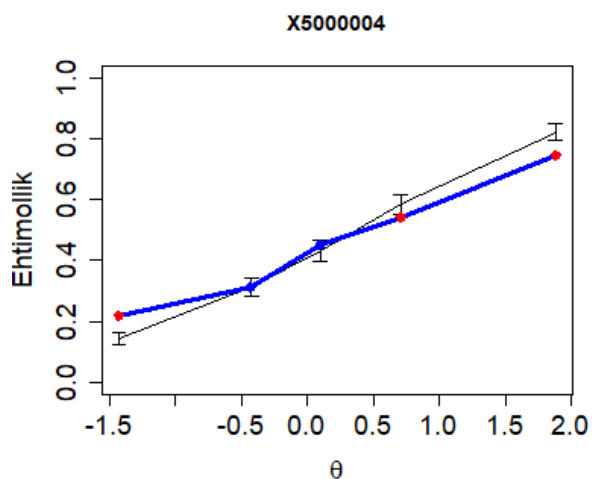
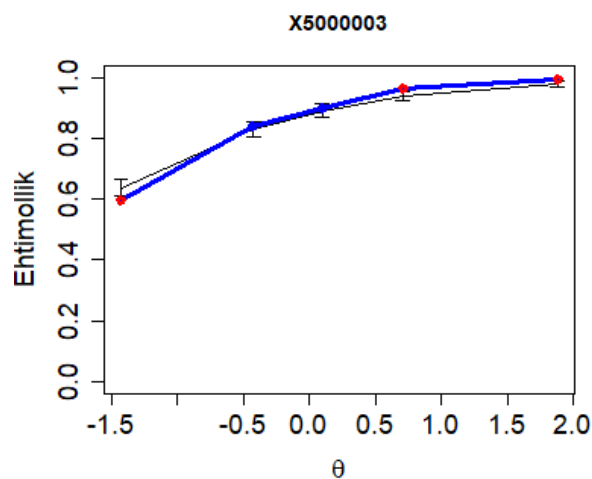
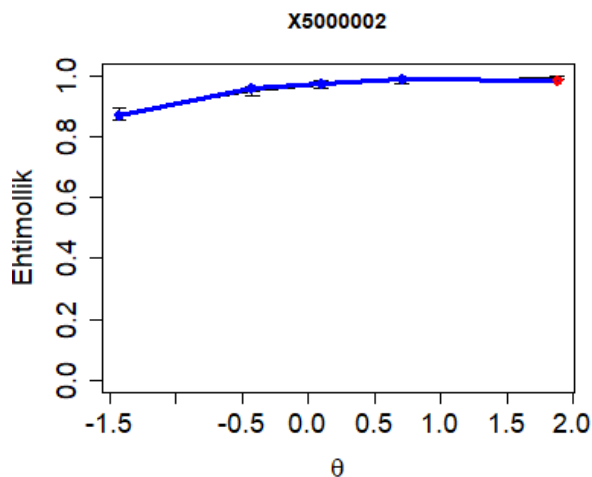
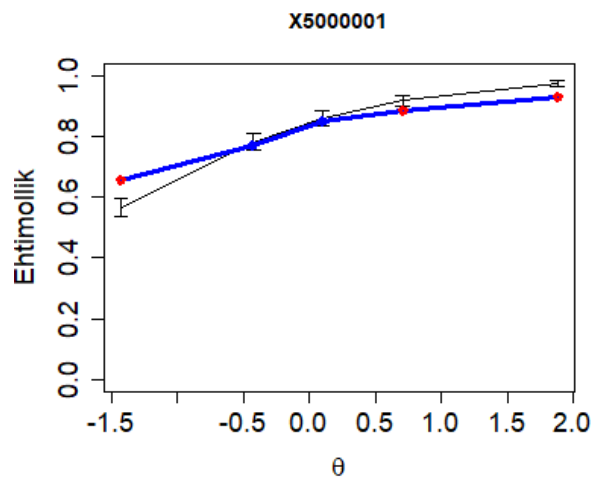
ball bilan uning korrelyatsiyasini ko'rish va ma'lum bir ball to'plagan sinaluvchilar guruhining bitta test topshirig'iga to'g'ri va noto'g'ri javob bergan javoblari o'rtasidagi bog'liqlikni ko'rishimiz mumkin [20]. Bu bog'liqlik sifati yuqori va past bo'lgan test topshiriqlari uchun bunday bog'lanish qanday bo'lishi to'g'risida tasavvur beradi. Test topshiriqlari bazasi uchun esa Rash modeli bilan moslik ham ahamiyatga ega. Yuqori sifatli test topshiriqlarini ham Rash modeli bilan mosligi yaxshi bo'lmasligi mumkin.

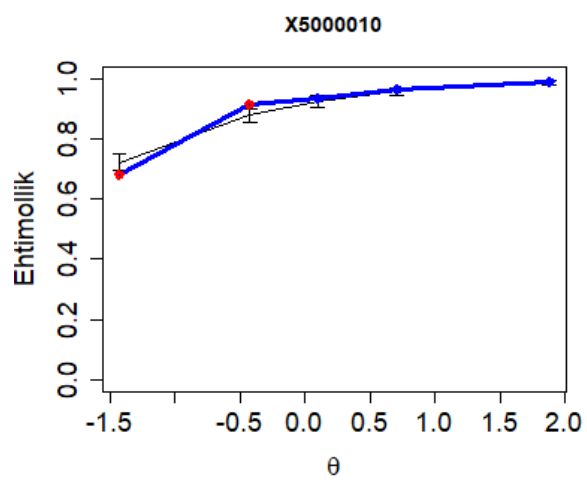
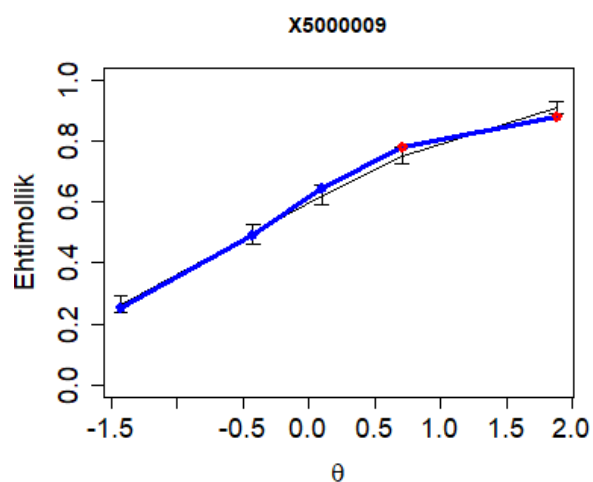
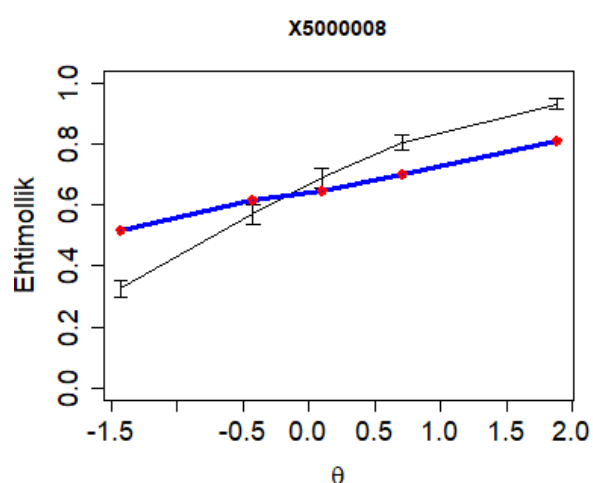
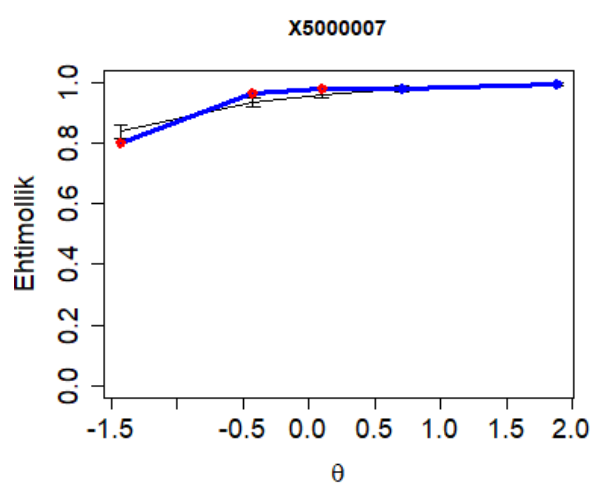
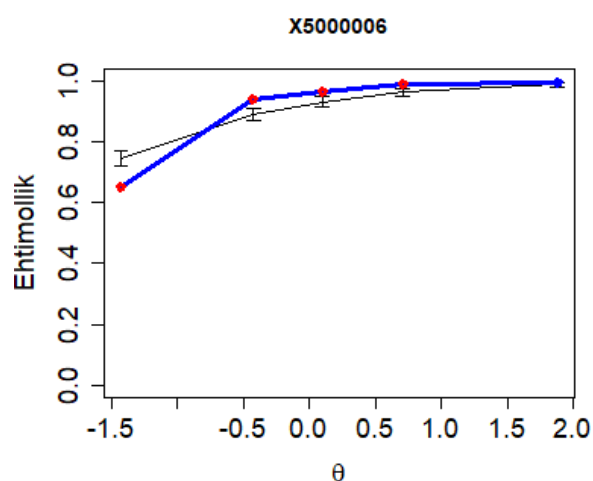
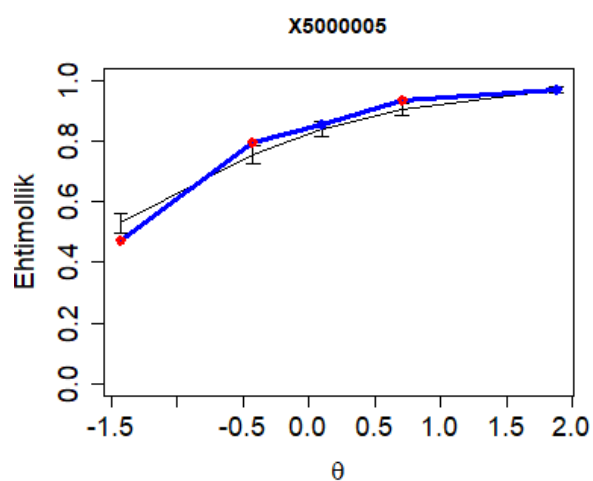
7-rasmda har bir test topshirig'ining Rash modeli bilan

qanchalik mos kelishini ko'rsatuvchi grafiklar ko'rsatilgan.

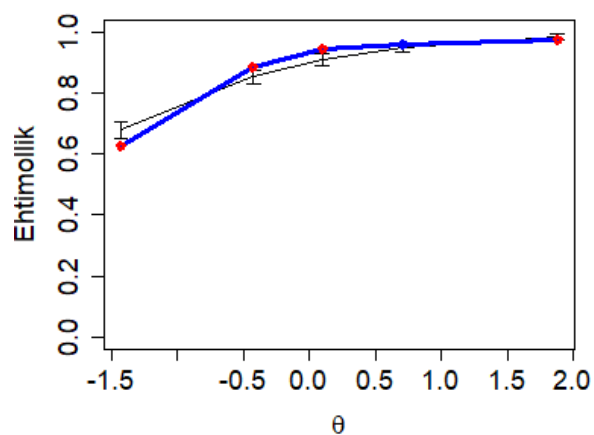
Sinaluvchilarning qobiliyati qalin ko'k chiziqlar bilan test sinovlaridan olingan natijalar, ingichka qora chiziq

bilan kutiladigan qiymatlar esa vertikal standart xatolik chiziqlari bilan birga ko'rsatilgan. Standart xatolik chegarasidan chiqib ketgan nuqtalar qizil doiralar bilan ko'rsatilgan.

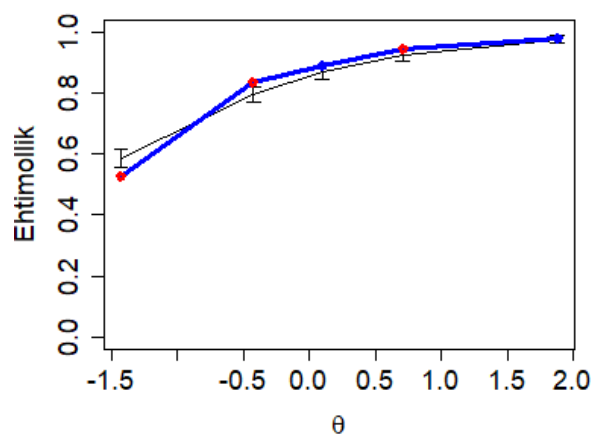




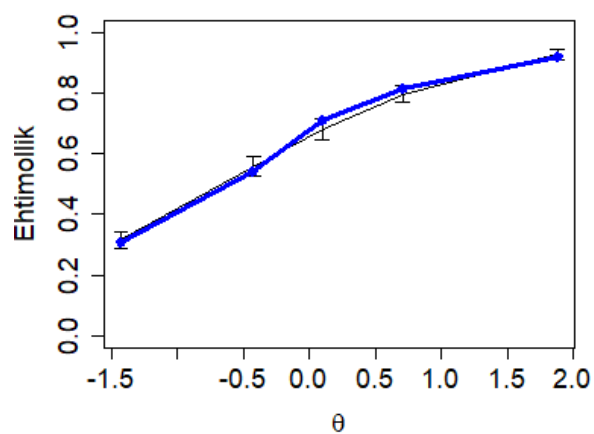
X5000011



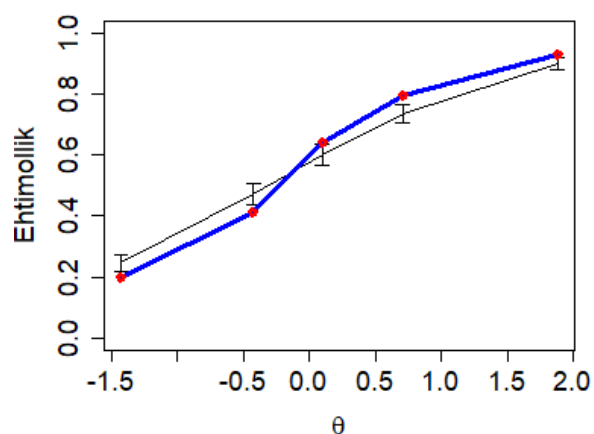
X5000012



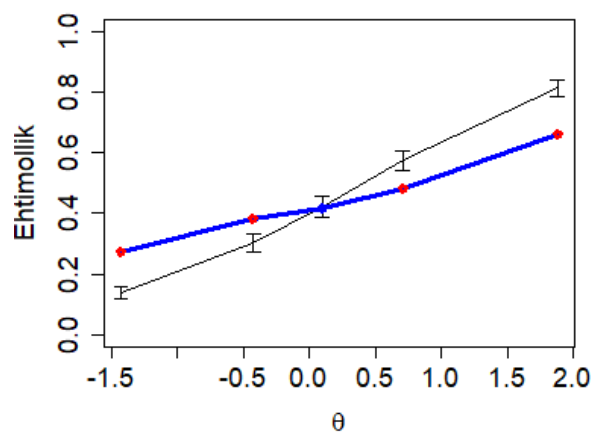
X5000013



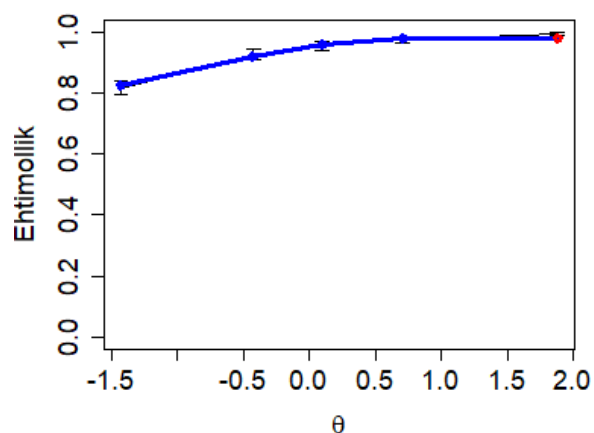
X5000014



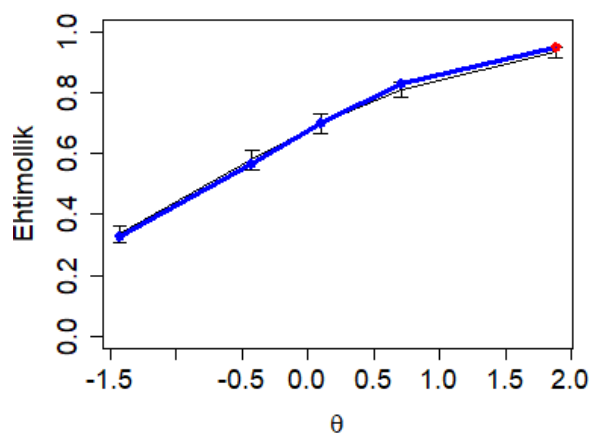
X5000015



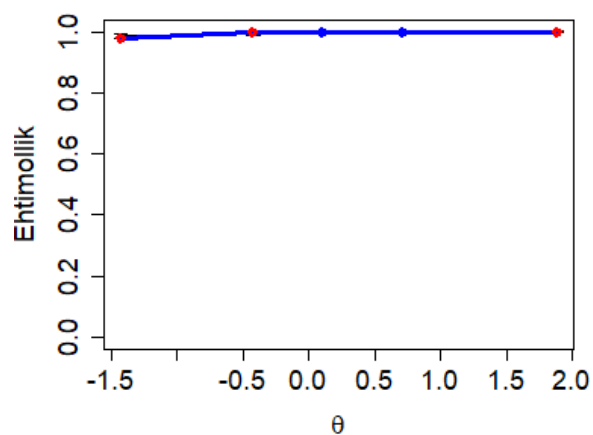
X5000016



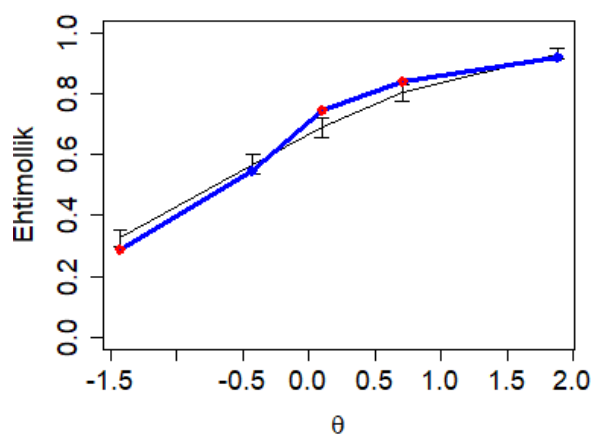
X5000017



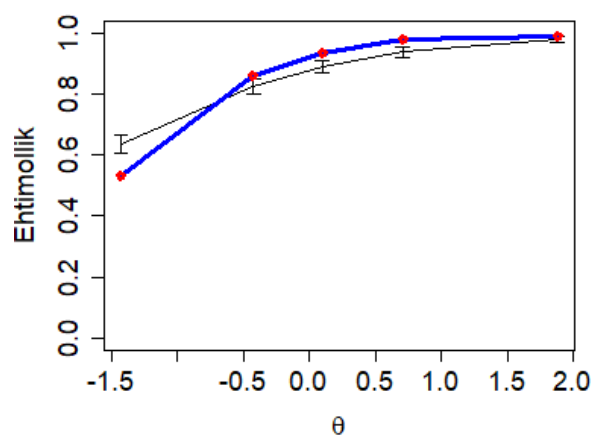
X5000018



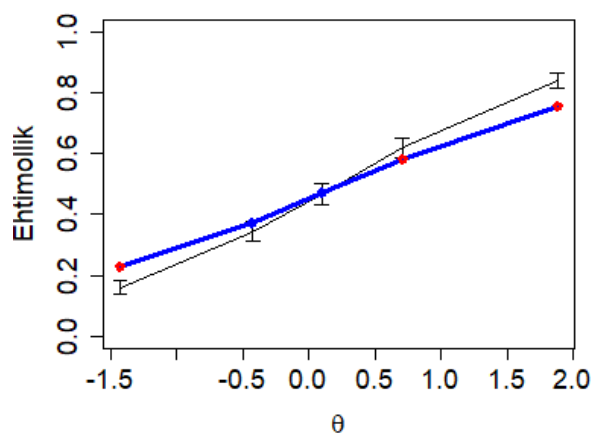
X5000019



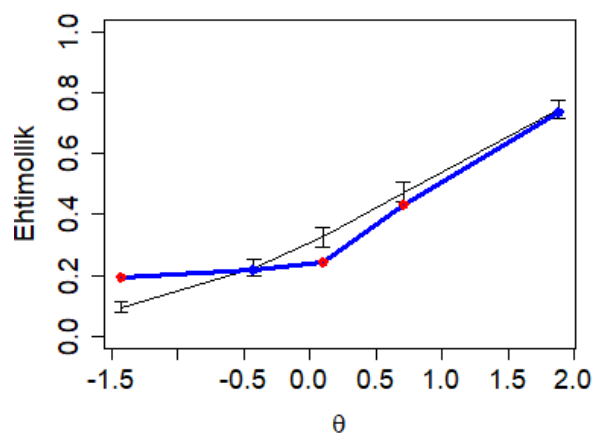
X5000020



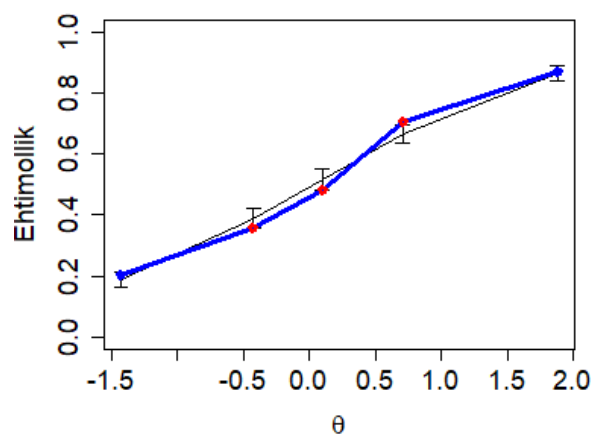
X5000021



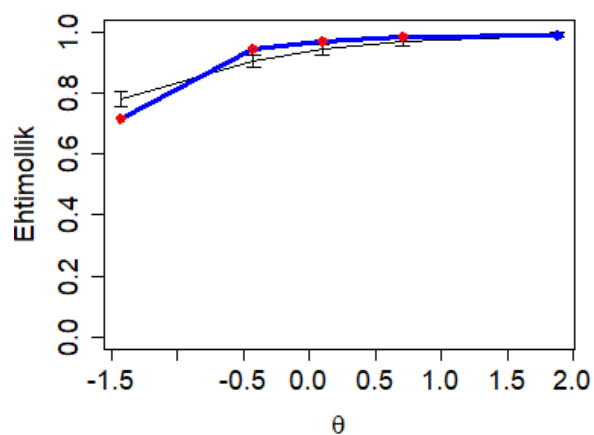
X5000022



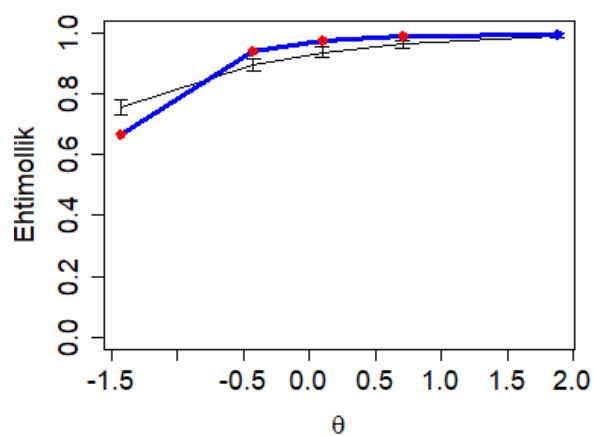
X5000023



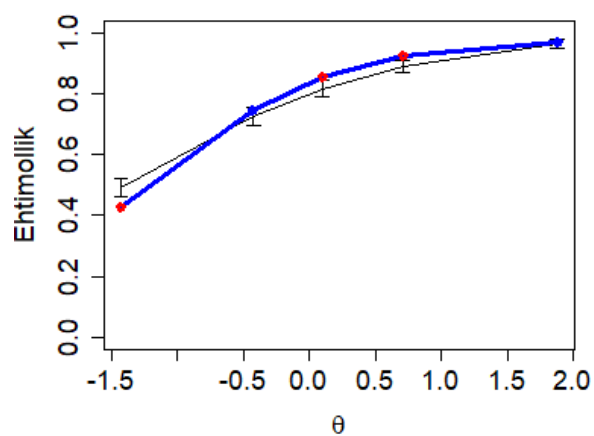
X5000024



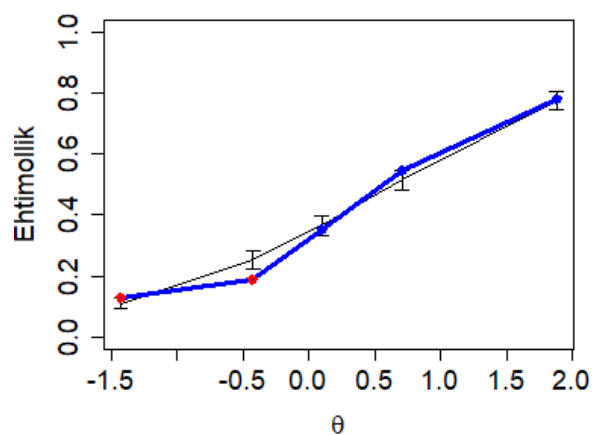
X5000025



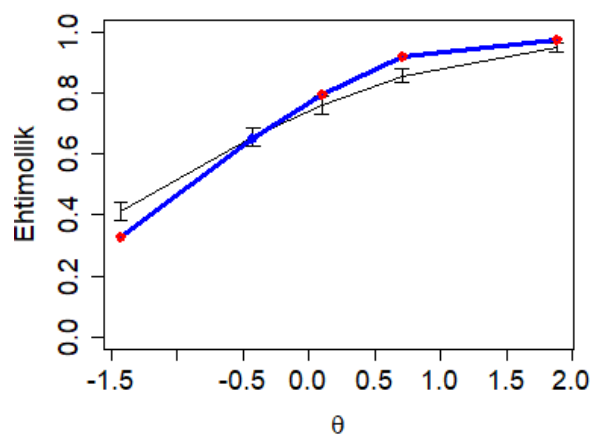
X5000026



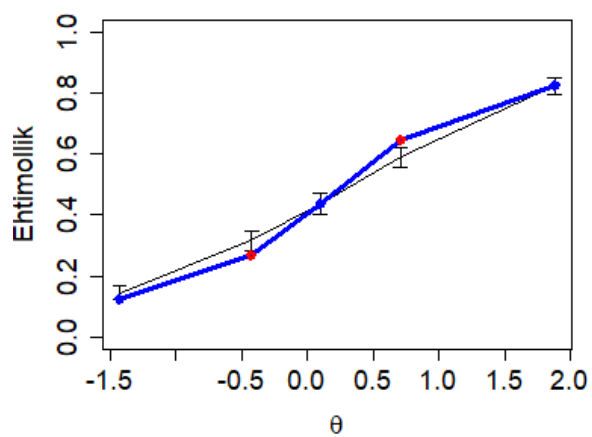
X5000027



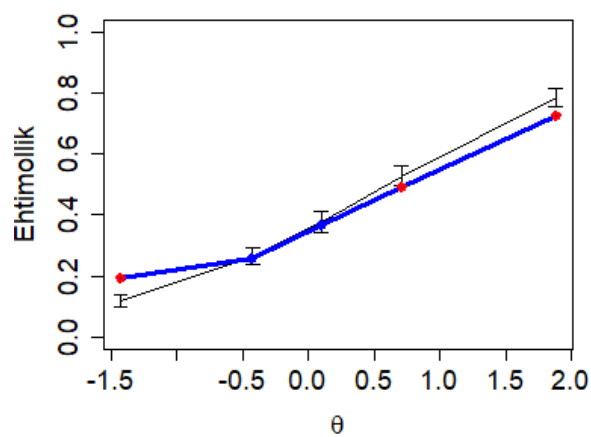
X5000028



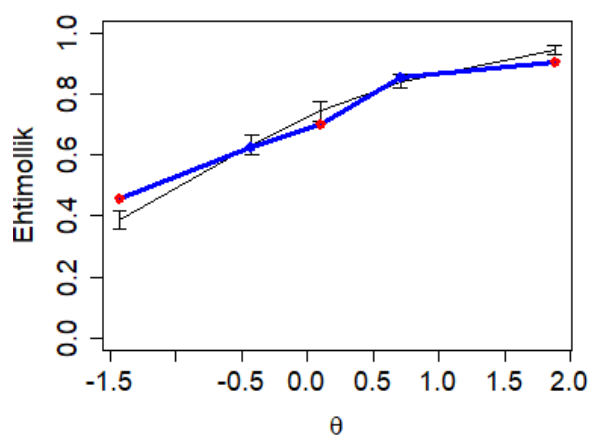
X5000029



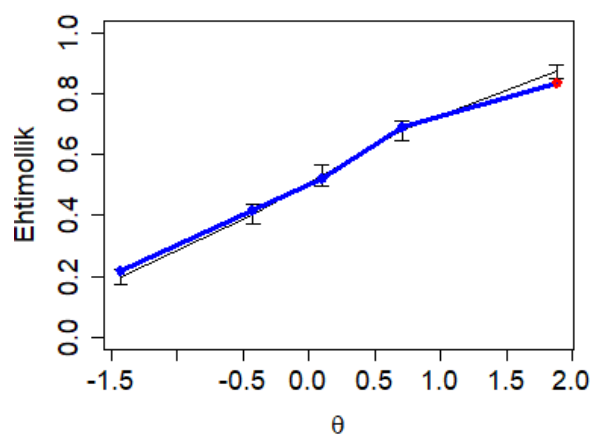
X5000030



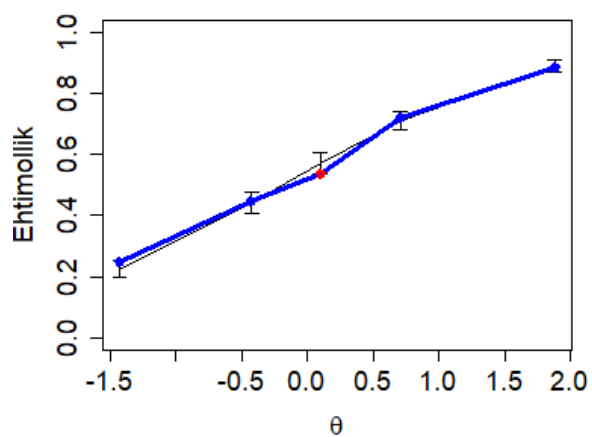
X5000031



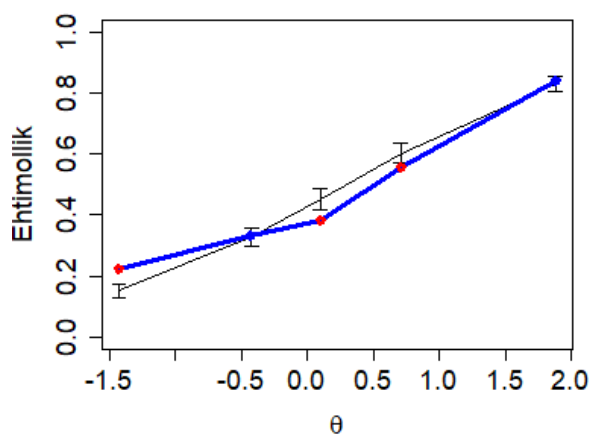
X5000032



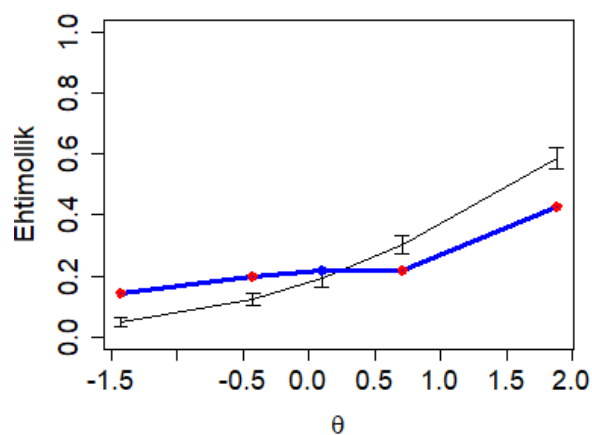
X5000033



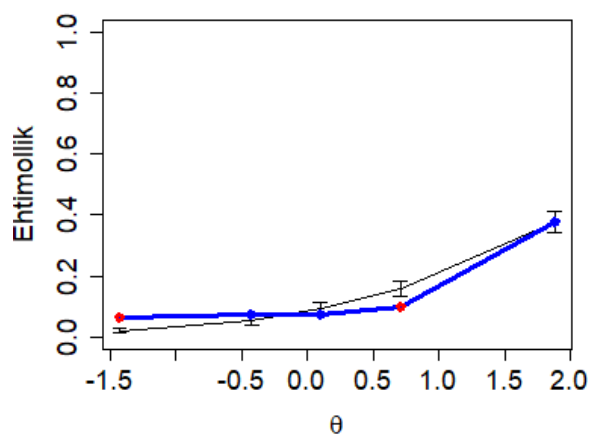
X5000034



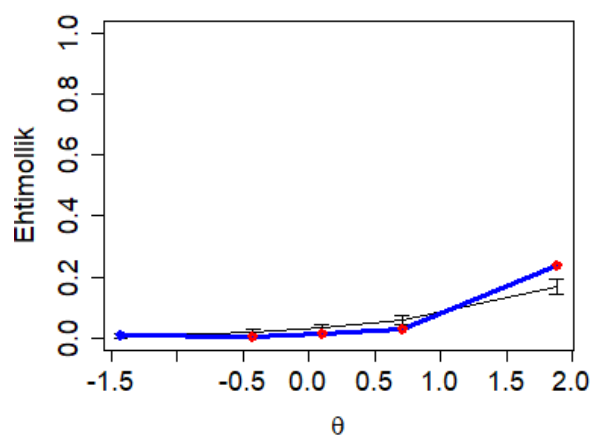
X5000035



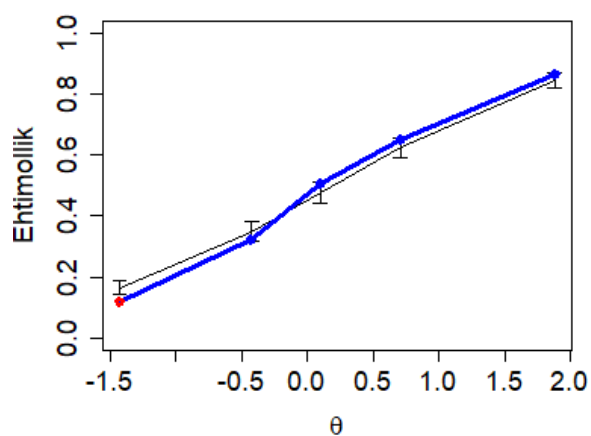
X5000036



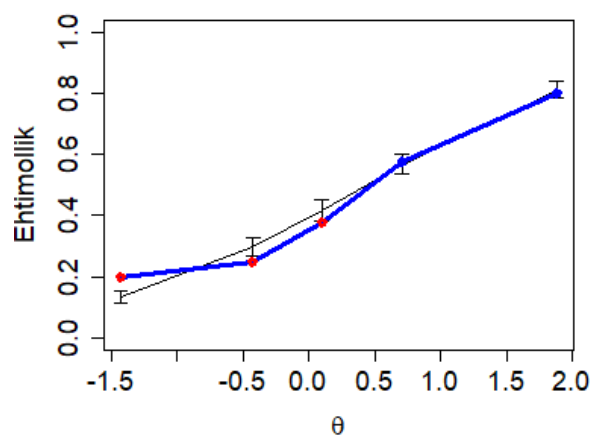
X5000037



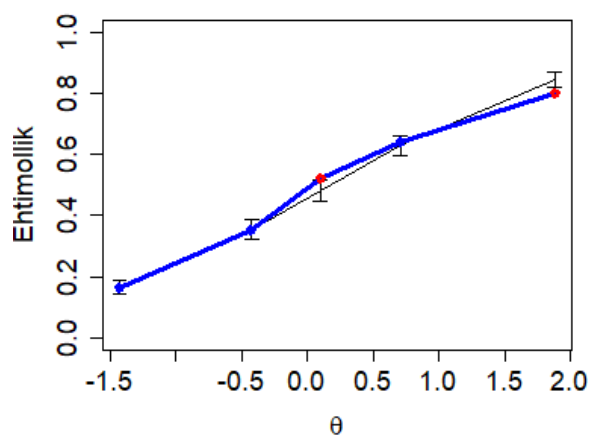
X5000038



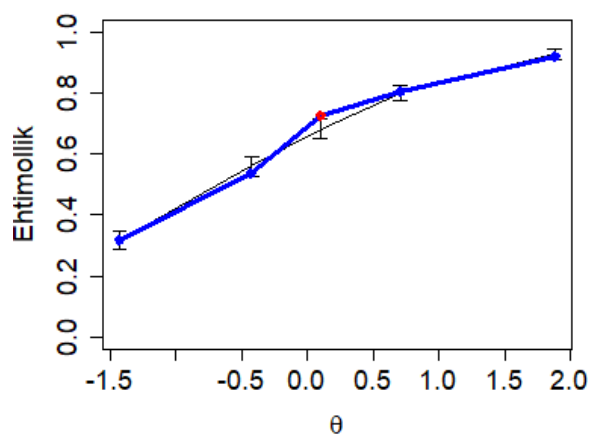
X5000039



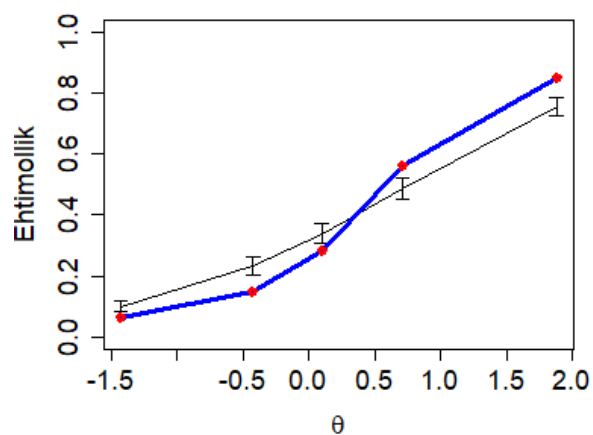
X5000040



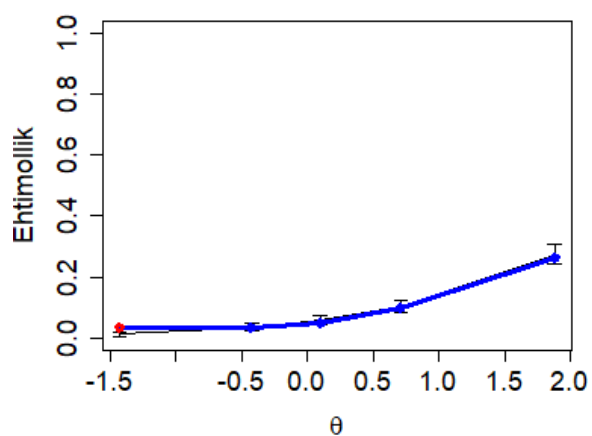
X5000041



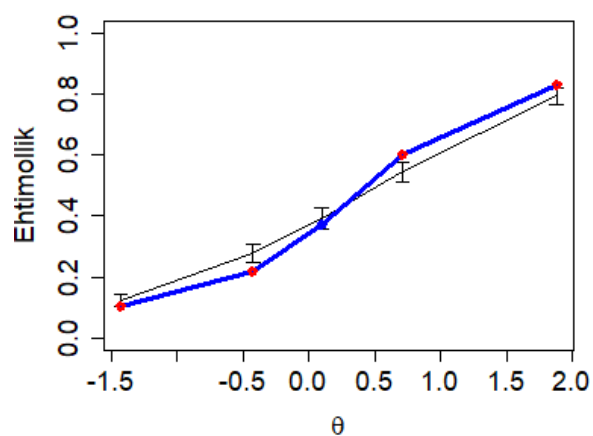
X5000042



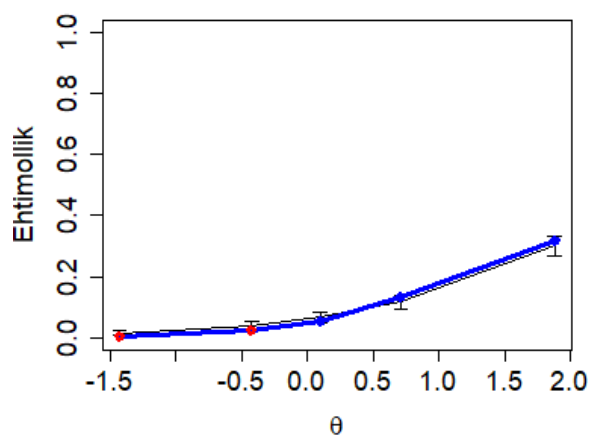
X5000043



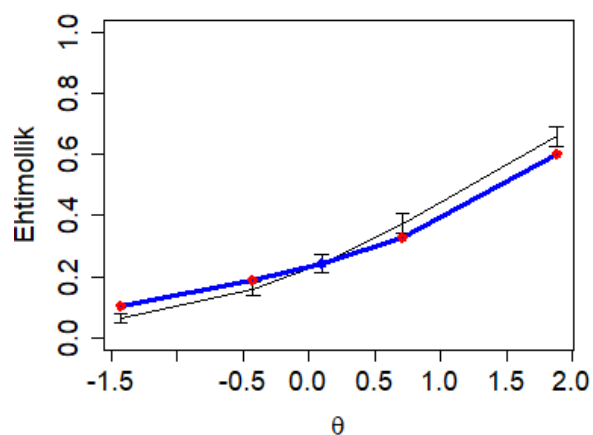
X5000044



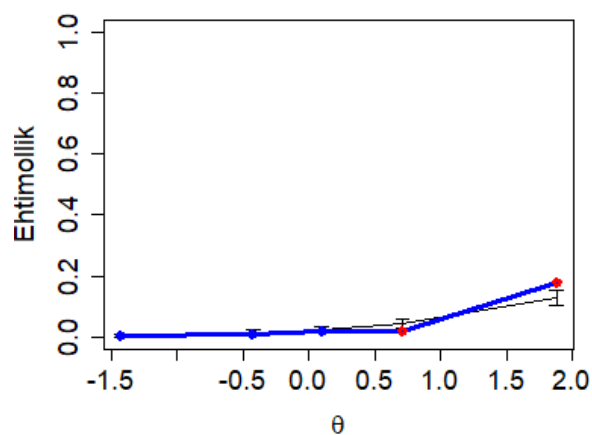
X5000045



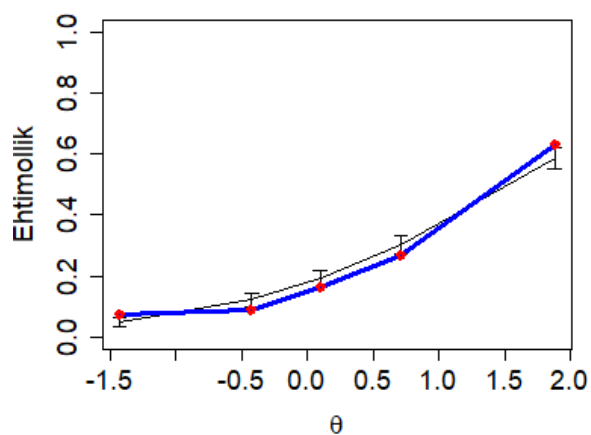
X5000046



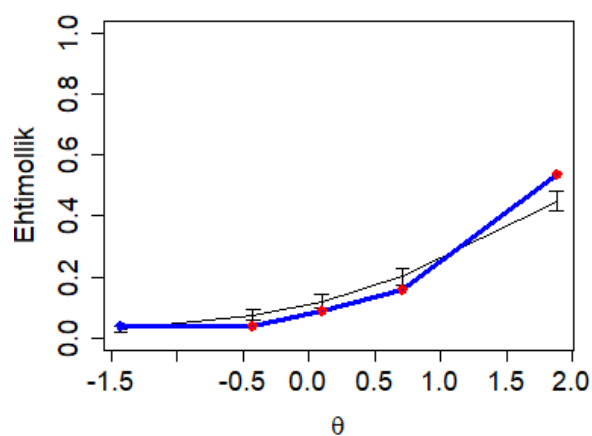
X5000047



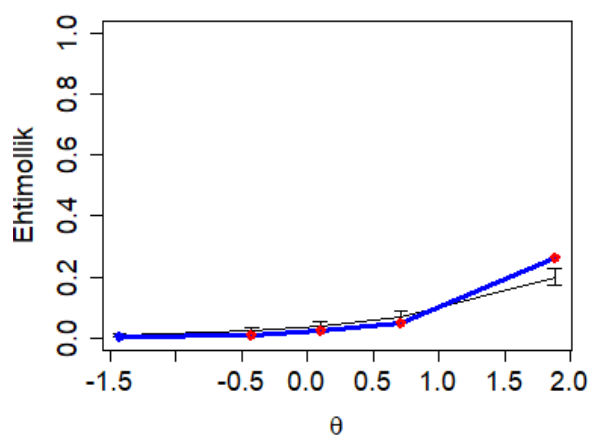
X5000048



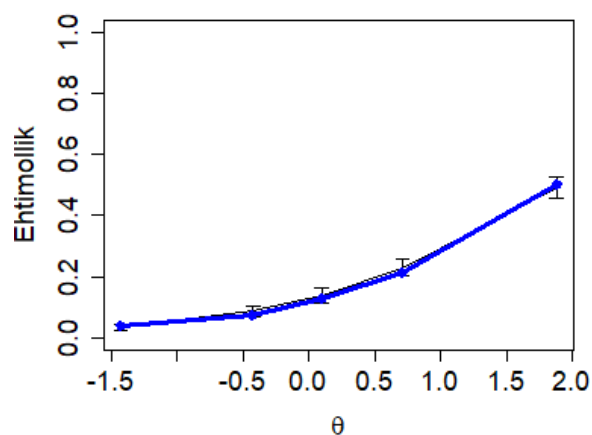
X5000049



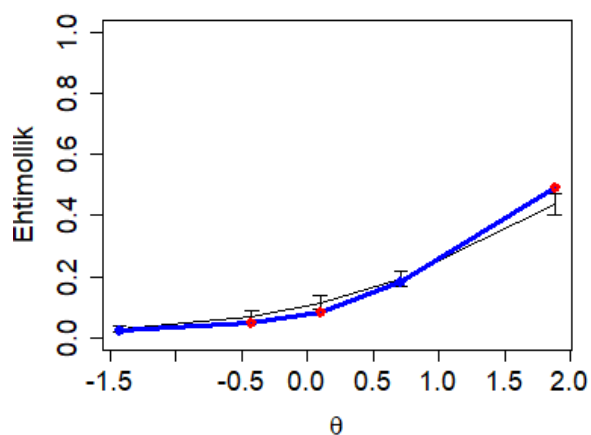
X5000050

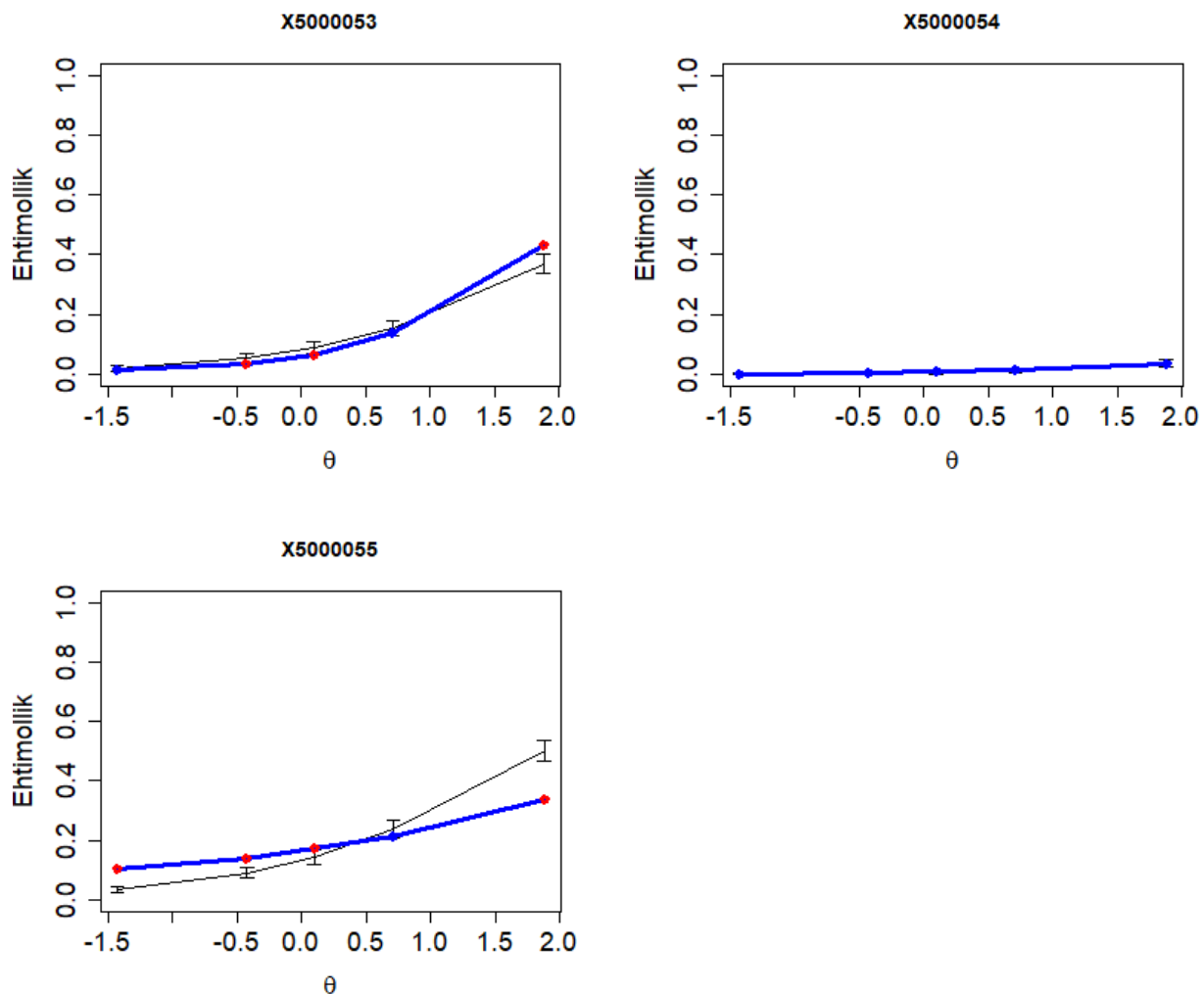


X5000051



X5000052





7-rasm. Test topshiriqlarining Rasch modeli bilan mosligi

ID raqamlari 5000004, 5000008, 5000014, 5000015, 5000021, 5000022, 5000035, 5000042, 5000044 va 5000055 bo'lgan test topshiriqlarining Rasch modeli bilan mosligi yaxshi emasligini ya'ni ajratilgan qobiliyat guruhlarining barchasi bilan mos tushmaganligini bildiradi aniqlandi.

Umuman olganda esa test topshiriqlarining korrelyatsiyalari juda

kichik bo'lgan test topshiriqlaridan tashqari barcha test topshiriqlarining Rasch modeli bilan mosligini qoniqarli deyish mumkin. Ushbu test variantining Kronbax alfasi 0,94 ga, test topshiriqlarining umumiy ball bilan korrelyatsiya koeffitsienti o'rtacha 0,43 bo'lib, bu variantning statistik ko'rsatkichlari me'yorda ekanligini ko'rsatadi.

Xulosa

Ushbu nuqtalarga mos keluvchi test topshiriqlari test tafsilotlarida keltirilgan fan mavzulari bilan taqqoslanganda sinaluvchilar tomonidan matematika fanidagi fizik jarayonlarning matematik yechimlari, murakkab tenglamalar va tengsizliklar, murakkab funksiyalar, matematik analiz asoslari va geometriya fanining planimetriya bo'limlaridan o'zlashtirish pastligi aniqlandi.

Sinaluvchilar tomonidan ochiq test topshiriqlarini bajarish ko'nikmasi ham past ekanligi va ularning o'rtacha bali 14 ga tengligi aniqlandi.

Rash modeli bilan baholash xom ball bilan baholashga nisbatan standartlik, validlik va ishonchlilikni aniqroq talqin qilish imkonini beradi. Rash modeli bilan hisoblangan test topshiriqlarining qiyinlik darajalari va qobiliyat darajalarining o'zaro mosligini Rayt xaritasi bilan tahlil qilish va mo'ljallangan guruh uchun test topshiriqlarini tanlash mumkin. Ilmiy tadqiqot uchun turli xil moslash usullaridan foydalanish mumkin, lekin test topshiriqlari bazasini yaratishda bitta usulni tanlash maqsadga muvofiqdir.

ADABIYOTLAR

1. H. Tyrone Black, David L. Duhon. Evaluating and Improving Student Achievement in Business Programs: The Effective Use of Standardized Assessment Test. Journal of Education for Business, 2014, 90-98.
2. Bond, Linda A. Norm- and Criterion-Referenced Testing, Practical Assessment, Research, and Evaluation, 1996, Vol. 5, 1-3. 3.
3. Verger A., Parcerisa L., Fontdevila C. The growth and spread of national assessments and test based accountabilities: A political sociology of global education reforms. Educational Review, 2019, 71, 5-30.
4. Verger A., Parcerisa L., Fontdevila C. The growth and spread of national assessments and test based accountabilities: A political sociology of global education reforms. Educational Review, 2019, 71, 5-30.
5. A.B. Normurodov, M.Dj. Ermamatov, T.X. Jumayev. Qiyinlik va qobiliyat darajalari invariantligi: fizika fanidan test natijalari tahlili. Axborotnoma. №1. 2024. 4-17
6. Popham, W.J. Why standardized tests don't measure educational quality. Educational Leadership, 1999, 56 (6), 8-15.,6.
7. Bond, Linda A. Norm- and Criterion-Referenced Testing, Practical Assessment, Research, and Evaluation, 1996, Vol. 5, 1-3.
8. Maja Planinic, William J. Boone, Ana Susac, and Lana Ivanjek. Rasch analysis in physics education research: Why measurement matters. PHYSICAL REVIEW PHYSICS EDUCATION RESEARCH 15, 020111 (2019)
9. Dylan Wiliam. Standardized Testing and School Accountability. Educational Psychologist, 2010, 107-122.
10. Spearman, C. (1904a). "General intelligence," objectively determined and measured. American Journal of Psychology, 15, 201-293.
11. Spearman, C. (1904b). The proof and measurement of association between two things. American Journal of Psychology, 15, 72-101.
12. Spearman, C. 1907. Demonstration of formulae for true measurement of correlation. Am. J. of Psychology. 18, 160-169.
13. Spearman, C. 1910. Correlation from faulty data, British J. of Psychology. 3, 271-295.
14. Аванесов В.С. Теория и методика педагогических измерений. МКО УГТУ-УПУ, 2005.
15. Челышкова М.Б. Теория и практика конструирования педагогических тестов: учеб. Пособие М.: Логос, 2002, с. 432.
16. [https://ru.wikipedia.org/wiki/Alfa_Kronbaха#:~: text = Коеффициент алфа Кронбаха, и для проверки их надёжности.](https://ru.wikipedia.org/wiki/Alfa_Kronbaха#:~:text=Коеффициент%20алфа%20Кронбаха,и%20для%20проверки%20их%20надёжности.)

17. Rasch G., Probabilistic models for some intelligence and attainment tests, Copenhagen, Danish Institute for Educational research. 1960.
18. Rasch G. "An item analysis which takes individual differences into account." British journal of mathematical and statistical psychology 19.1 1966, 49-57.
19. Hattie J. Methodology review: assessing unidimensionality of tests and items, Applied psychological measurement, 1985, T, 9, №2, 139-164.
20. Gunter Maris, Timo Bechger, Jesse Koops and Ivailo Parchev, Data Management and Analysis of Tests, 2022, p. 1-49.
21. M.D. Ermamatov, M.D. Alimov, A.A. Sulaymonov, A.R. Sattiyev. Kalibrovkalangan test topshiriqlari: Sharq tillaridan o'tkazilgan test sinovi natijalarining statistik tahlili, Axborotnoma №. 3-4, 16-83 b., 2022.
22. M.D. Ermamatov, A. Abbosov, A.A. Baratov, Test topshiriqlarini kalibrovkalash va qobiliyatlarini tenglashtirish, Axborotnoma №. 3-4, 4-16 b., 2022.
23. M.Dj. Ermamatov, A.R. Sattiyev, A.B. Normurodov, Z.O. Olimbekov, A.A. Baratov. Fizika fanidan o'tkazilgan test sinovi natijalari: Rayt xaritasi, ichki va tashqi moslik statistikalari, Rash modeli bilan moslik, Axborotnoma №1, 2023, 4-62
24. B.D. Wright and M.H. Stone, Best Test Design (MESA Press, Chicago, 1979).
25. Maja Planinic, William J. Boone, Ana Susac, and Lana Ivanjek. Rasch analysis in physics education research: Why measurement matters. Physical review physics education research 15, 020111 (2019).
26. A.R. Sattiyev, M.Dj. Ermamatov. Matematika fanidan milliy sertifikat uchun o'tkazilgan test sinovlari natijalari tahlili. Axborotnoma №2. 2023. 35-55 betlar.
27. Wobbe Zijlstra, ALTE-accreditation Uzbekistan Psychometrics-Online Seminar, Tashkent, 2022-June.

STATISTICAL ANALYSES OF THE RESULTS OF DIAGNOSTIC TEST TRIALS IN MATHEMATICS

K.A. Amonov, A.A. Baratov

Scientific-study Practical Center under the Agency for Assessment of Knowledge and Competences under the ministry of higher education, science and innovation of the republic of Uzbekistan, Tashkent 100084, Bogishamol st. 12

Abstract. This article presents a statistical analysis of the results of diagnostic test trials in mathematics based on classical and modern test theories. According to the statistical research findings, the reliability coefficient of the test version – Cronbach's alpha coefficient was determined to be 0.94. However, when the difficulty levels of the test items were assessed using modern test theory, the presence of extremely difficult and extremely easy test items was observed. An analysis of the responses provided by test-takers revealed low proficiency in areas such as mathematical solutions to physical processes, complex equations and inequalities, advanced functions, the foundations of mathematical analysis, and the planimetry section of geometry.

Keywords: Test items, Mode, Median, Standard deviation, Variance, Cronbach's alpha coefficient, validity, level of difficulty, Rasch model, Wright-map, ability levels