O‘ZBEKISTON RESPUBLIKASI OLIY TA’LIM, FAN VA INNOVATSIYALAR VAZIRLIGI HUZURIDAGI BILIM VA MALAKALARNI BAHOLASH AGENTLIGI

AXBOROTNOMA

BULLETIN

1

2025

Toshkent

“AXBOROTNOMA”
O‘zbekiston Respublikasi Oliy ta’lim, fan va innovatsiyalar vazirligi huzuridagi Bilim va malakalarni baholash agentligi ilmiy-uslubiy jurnali
Bir yilda 4 marta chiqadi.
O‘zbekiston Matbuot va axborot agentligida 2007-yil 19-aprelda qayta ro‘yxatdan o‘tkazilgan.
Guvohnoma № 0247

Ta’sischi:
Bilim va malakalarni baholash agentligi

Bosishga ruxsat etildi: 30.04.2025
Shartli bosma tabog‘i: 5,4
Nashriyot hisob tabog‘i: 5,3
Adadi 100. Buyurtma № 1 - A - 25

Bilim va malakalarni baholash agentligi bosmaxonasida chop etildi.

Nashrga tayyorlovchilar:
A. Normurodov
Muharrir:
N.Abduraxmanova
Kompyuterda sahifalovchi:
Y. To‘rayev
Tahririyat manzili:
100084, Toshkent shahri,
Bog‘ishamol ko‘chasi, 12-uy.

MUNDARIJA

CONTENTS

KIRISH

“Axborotnoma” ilmiy-uslubiy jurnalining ushbu sonida pedagogik o‘lchovlar bo‘yicha ilmiy-uslubiy izlanishlar olib boruvchi mutaxassislar tomonidan olib borilgan tadqiqotlar natijalari haqida to‘rtta maqola berilgan.

Birinchi maqolada oliy ta’lim muassasalari bakalavriat bosqichi kirish test sinovlarida ingliz tili fanidan abituriyentlarning o‘qib tushunish ko‘nikmasini baholovchi yopiq test topshiriqlarining autentlik xususiyatlarini tadqiq qilish orqali ushbu muammoni qanday qilib bartaraf etish mumkinligi muhokama qilinadi.

Ikkinchi maqolada maqolada milliy sertifikat uchun kimyo va biologiya fanlaridan o‘tkazilgan imtihonda qatnashgan talabgorlarning test va yozma ish ballari o‘rtasidagi korrelyatsiya koeffitsiyentlarini hisoblash, test va yozma ish uchun berilgan ballar o‘rtasidagi korrelyatsiyani aniqlash usullari bayon qilingan va hisoblash natijalari tahlil qilingan.

Uchinchi maqolada shkala va uning turlari, sinaluvchilarning xom ballarini z- va T-ballga aylantirish hamda ona tili va adabiyot fanidan o‘tkazilgan test natijalarining xom ballarini z- va T-ballarga o‘tkazish, ya’ni bir xil shkalaga keltirish usullari muhokama qilingan.

To‘rtinchi maqolada 2024-yilda geografiya fani bo‘yicha milliy sertifikat uchun o‘tkazilgan test sinovlari natijalari klassik test nazariyasi va Rash modeli asosida tahlil qilingan. Test sinovi natijalarining tavsif statistikasi va har bitta test topshirig‘iga berilgan javoblarning umumiy ball bilan korrelyatsiyasi muhokama qilingan. Klassik test nazariyasi va Rash modeli bilan o‘tkazilgan uchta test sinovida ishlatilgan variantlardagi test topshiriqlarining qiyinlik darajasi tahlil qilingan. Rash modeli asosida aniqlangan qobiliyat va qiyinlik darajalaridan foydalanib, har bir variant uchun Rayt xaritalari olingan. Shuningdek, test ma’lumoti va xarakteristikasi chiziqlari hamda test variantlaridagi har bir test topshiriqlarining Rash modeli bilan moslik statistikalari o‘rganilgan.

ENHANCING READING COMPREHENSION TEST DESIGN: INTEGRATING AUTHENTIC VISUAL MATERIALS TO IMPROVE ASSESSMENT

K.O. Fayziyeva and Sh.Sh. Nishonboyev

Agency for Assessment of Knowledge and Competences under the Ministry of Higher Education, Science and Innovations of the Republic of Uzbekistan, 100084, Tashkent., Bogishamol str., 12

Abstract. Entrance exams play crucial role in determining students' academic futures. However, there has been inappropriacy in the test format and content which fails to reflect real-world language use when traditional testing methods are used. This article attempts to explore how features of authenticity can bridge this gap when applied in reading comprehension tests. In response to this concern, this article attempts to clarify how the authenticity and its application can address this issue in tailoring test items designed to assess reading skills in the English language in the entrance exams. In particular, it examines the notion of authenticity, its characterization and application in designing test items which more effectively assess reading comprehension. It also mentions that, the term authenticity has long been studied thoroughly in countries abroad but requires more meticulous investigation and practical application in the language testing context in Uzbekistan.

Keywords: reading comprehension, authenticity, closed-ended test items, test details, infographics.

Authenticity is a key quality of language test items as it is a distinctive feature that should be present in the relationship between the passage and the reader. In this sense, Widdowson defined as a 'genuine input' in the use of texts that are not simplified and tasks that encourage test takers to perform in a 'real-world' context. [1].

Bachman and Palmer [2], experts in the field of language assessment, raised concerns about the content with which test takers would perform in the test, claiming that it would define to what extent they would engage in the language use outside the test environment. Building on this idea, they further introduced the term target language use (TLU) to describe the context in which language is used in the real world, wherein test takers will use the language.

Expanding on this perspective, Bachman [3] claimed that authenticity is a characteristic which

demonstrates to what extent test tasks are similar or look like to target language use (TLU) tasks. In this regard, an authentic assessment entails activities and tasks which students are likely to encounter in real life situations.

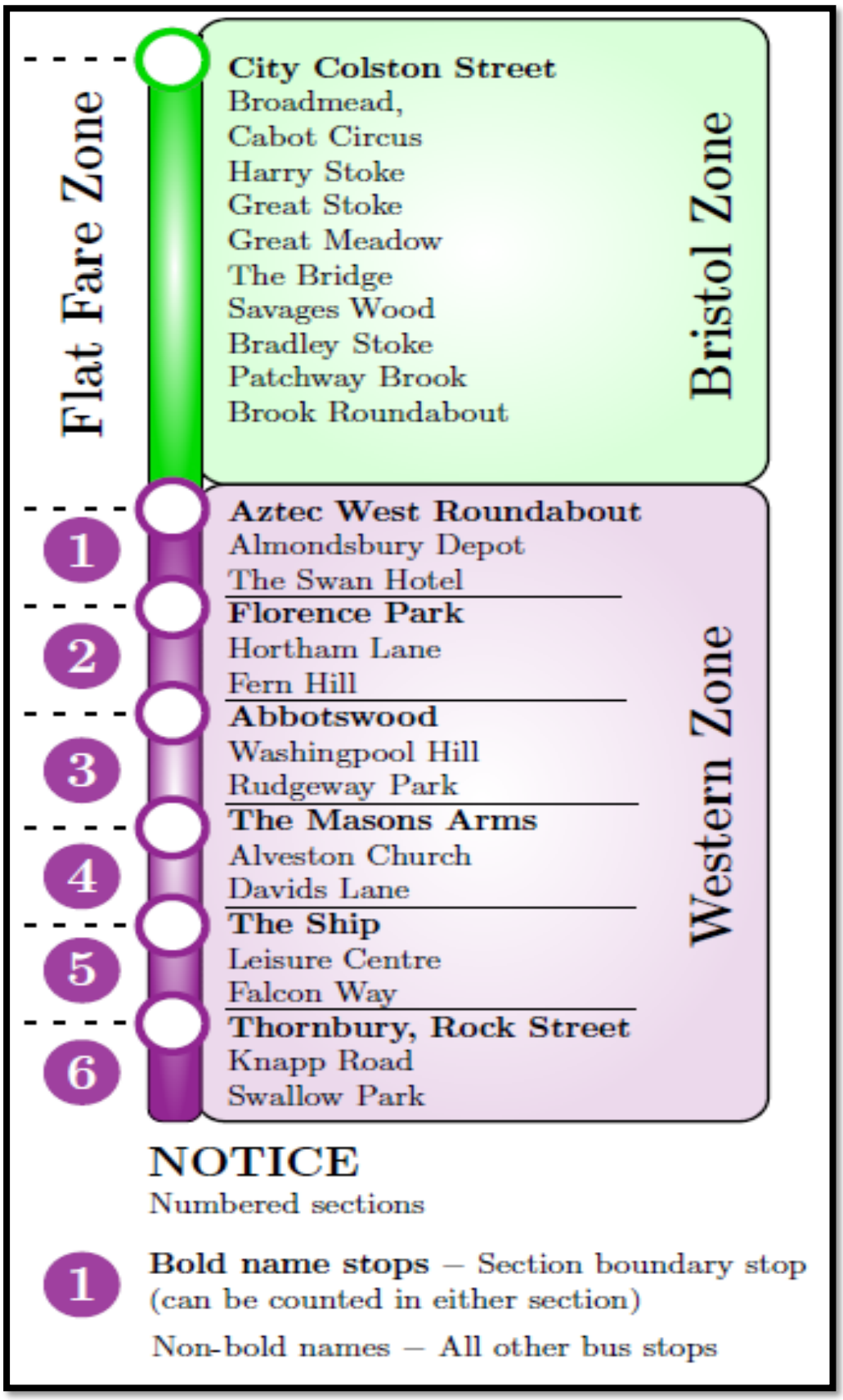
In terms of designing test items for reading comprehension it is important to note that the fact that texts that are produced or available in the real world are quite different, especially when considering the complexity, target audience, and amount of information [4]. This underscores the need to carefully select reading materials that maintain a balance between authenticity and other relevant factors.

There are a number of contributing factors involved in maintaining the authenticity of the task, as follows [5]:

- the language of the test should be kept natural;
- provide a context for the items rather than making it isolated;
- choose meaningful – relevant and interesting topics for the learner;
- present items in a storyline or in episodes so as to keep thematic organization;
- tasks should resemble real world tasks.

In recent years, the shift towards authenticity in Uzbekistan's language testing environment has brought about a number of noticeable changes. Several decades ago, foreign language testing items designed to assess applicants consisted of boring and artificial grammar-based multiple-choice items accompanied by short reading passages that lacked spontaneity and naturalness. Since then, in line with developments in language assessment research, numerous changes and approaches have been implemented to improve the quality of test materials.

One of the recent changes was to integrate authentic graphics, charts, diagrams, and infographics together with reading passages to prompt the test taker to analyze and synthesize the given content simultaneously. O'Malley and Valdez Pierce [6] make a related argument proposing various approaches that can enhance the authenticity of the tasks: reports, drawings and diagrams, interviews, inventories of work place items, logs, simulations and models with descriptions. As a result of this and other studies, several modifications were made to the reading test items designed to assess the applicants for university entrance exams.



Sample 1. Read the guide about the fare system in the West of England and answer the questions according to it.

A flat fare structure is in place in the Bristol Zone. This means that wherever you travel in that zone, the price of a single journey will be the same – £2.00. Single fares in the Western Zone are based on the distance you are travelling.

Distances are calculated using 'fare stage' sections rather than individual bus stops, with each section being one-mile long. If you travel up to 3 miles it'll be £2.30, 4-6 miles will be £3.30. Here is an example route and some of the fares along it:

From Aztec West Roundabout to the Rudgeway Park would cost £2.30 as you are travelling in three sections (numbers 1, 2 and 3 - the Rudgeway Park is classed as section 3 in this example as it's where you are getting off the bus). From the Masons Arms to Thornbury, Rock Street would also be £2.30 as you are again travelling in three sections (4, 5 and 6 - The Masons Arms is counted as section 4 in this example as it's your boarding point). In the other direction, from Swallow Park to Rudgeway Park would cost £3.30 as you are travelling in four sections (6, 5, 4 and 3).

1. A passenger wants to travel from Almondsbury Depot to Washingpool Hill. How much will he have to pay for his trip?

A) £2.30
B) £2.00
C) £3.00
D) £3.30

2. As stated in the guide, within the Western Zone … .

A) the distance is calculated for every section you pass
B) you should pay the same fare as in the Bristol Zone
C) the fare system requires passengers to pay £2.00
D) every bus stop is counted as one mile long

3. How much will the fare be for the route starting from Knapp Road to Abbotswood?

A) £2.00
B) £2.30
C) £3.00
D) £3.30

4. As stated in the guide, in the Bristol Zone … .
   A) the same fare is on regardless of the bus stop
   B) passengers pay for every bus they pass
   C) single journey will cost less than £2.00
   D) three bus stops are equal to one section

The above presented infographic and the text (Sample 1) are taken from an authentic – real life guidebook for the passengers travelling in the territories of West of England and Bristol city [7]. Certain differentiations and adaptations have been implemented to this material such as replacing a few difficult words in the text with B1 level alternatives and changing complex sentences into simple versions. Overall, the names of bus sections, the length and the content of the text and graphic remain almost unchanged. The multiple choice items created based on this material are intended to test scanning skills and simulate the applicant to choose the correct option from the given stem presented as a problem-solving scenario. This approach aligns with "constructivist learning principles" proposed by Jonassen [8] who suggests that integrating assessment tasks into situations involving problem-solving scenarios improves engagement and fosters deeper cognitive processing.

Open day sessions	Activities				Location (for each of the activity)
Time	9.am-10.am	10.50am-11.50am	12.30pm-1.10pm	2.50pm-3.50pm	
Durham Difference	✓	✓	✓		Ground Floor, Room 4B
Careers & Enterprise	✓		✓	✓	Lovalls Lecture Hall
Student Guide	✓	✓		✓	Holmes Lecture Hall
Support Service		✓		✓	Dunelm House

= Presentations, = Tours, = Exhibitions

= Drop-in sessions (for those who have a private individual question and have not made prior booking to attend)

✓ = Available activities at given time and place

Sample 2. Look at the guide about open day sessions at Durham university and answer the questions according to it.

Durham Difference

This session will introduce the university and its values, vision and goals. It will introduce you to our academic staff. Our current students will share their experiences about what it means to become a member of university community.

Careers & Enterprise

This session is about exploring your job opportunities whilst studying a subject you enjoy. It's also about gaining a qualification and developing skills and experience that can lead to a fulfilling and rewarding career. Find out how we support students, how we work with employers and the types of opportunities available to students during and after their studies.

Student Guide

Find out about fees, and government loans. This session will provide basic factual student finance information, explaining region-specific privileges and the application process.

Support Service

Here you will find a specialist who supports at any time during your studies if you are struggling with academic or personal challenges. It operates as an addition to the support offered in colleges and academic departments. Our teams of advisors help students to manage any difficulties impacting study or wider university life.

1. A student would like to know more about the expenses of studying at this university and whether there are any financial benefits.

Which of these sessions is the most suitable for him?

A) Durham Difference
B) Careers & Enterprise
C) Student Guide
D) Support Service

2. Students who have problems with their studies or personal life need advice. Which session should they attend?
A) Durham Difference- Exhibition
B) Careers & Enterprise - Presentation
C) Student Guide - Drop-in session
D) Support Service - Tour

3. A parent wants to take his child to a session to learn about the university goals and student life. What time and where should they go?
A) Holmes Lecture Hall at 9:00 am
B) Lovalls Lecture Hall at 12:30 pm
C) Dunelm House at 10:50 am
D) Ground Floor, Room 4B at 9:00 am

The chosen reading passage (Sample 2) is taken from an open day guidebook [9] and the items correspond to the criteria of providing authenticity in the task [5]:

- It maintains authenticity through the use of relevant topic area 'university open day sessions' that applicants are most likely to encounter in the future;

The items are composed based on the passage together with appropriate visual and real-life context, not being confined in the text format;

- multiple choice questions mainly pose students to analyze the problematic case and make informed decision by scanning both text and table (visual);

Improving the authenticity of language tests is crucial for both building learner's real world language proficiency and ensuring the validity of test results. This article attempts to address this case by presenting a strong evidence that using real-world visual aids may enhance reading comprehension. It analyses the approach of using real world examples with theoretical foundation. However, key points for improvement include avoiding the dependence on multiple choice forms, the absence of empirical validation and the limited contextual discussions. Future research should concentrate on measuring the efficiency of authenticity-based evaluation. Alternative question formats should be

explored to better capture the complexity of real-world language use. Comparative studies between different educational systems would also be valuable in understanding how authenticity can be adapted to diverse testing environments.

REFERENCES

1. Widdowson, H. 1978: Teaching language as communication. Oxford: Oxford University Press.
2. Bachman, L and Palmer, A. 1996: Language testing in practice. Oxford: Oxford University Press.
3. Bachman, L. 1991: What does language testing have to offer? TESOL Quarterly 25, 671–704.
4. Seliger, H. W., & Shohamy, E. (1989). Second Language Research Methods. Oxford University Press.
5. Brown, D. H. (2004). Language assessment: Principles and classroom practices. White Plains, New York: Pearson Education.
6. O'Malley, J. M. & Valdez Pierce, L. (1996). Authentic assessment for English language learners: Practical approaches for teachers. Addison-Wesley Publishing.
7. http://www.amazon.com/Authentic-Assessment-English-Language-Learners/dp/0201591510.
8. https://www.firstbus.co.uk/uploads/node_images/J13039%20Ticket%20Guide%20AW2.pdf.
9. Jonassen, D. H. 1997. Instructional design models for well-structured and ill-structured problem-solving learning outcomes. Educational Technology Research and Development, 45(1), 65–94.
10. https://www.durham.ac.uk/media/durham-university/visit-us/documents/open-days/open-day-presentations/Preparing-for-your-Open-Day-Visit-June-2023.pdf.

INGLIZ TILI FANIDAN O‘QIB TUSHUNISH BO‘YICHA YOPIQ TEST TOPSHIRIQLARINI TAKOMILLASHTIRISH: AUTENTIK VIZUAL MATERIALLAR ORQALI SAMARALI BAHOLASH

K.O‘.Fayziyeva va Sh.Sh.Nishonboyev

Oliy ta’lim, fan va innovatsiyalar vazirligi huzuridagi
Bilim va malakalarni baholash agentligi, 10084, Toshkent sh., Bog‘ishamol k.,12

Qisqacha mazmuni. Talabalarning ta’limdagi kelajagini belgilashda kirish imtihonlari muhim ahamiyatga ega. Ammo ko‘pincha an’anaviy test usullari qo‘llanilganda, test formati va mazmunida kamchiliklar yuzaga keladi va real hayotda tilning qo‘llanilishini to‘laqonli aks ettira olmaydi. Ushbu maqolada oliy ta’lim muassasalari bakalavriat bosqichi kirish test sinovlarida ingliz tili fanidan abituriyentlarning o‘qib tushunish ko‘nikmasini baholovchi yopiq test topshiriqlarining autentlik xususiyatlarini tadqiq qilish orqali ushbu muammoni qanday qilib bartaraf etish mumkinligi muhokama qilinadi.

Kalit so‘zlar: o‘qib tushunish ko‘nikmasi, autentlik, yopiq test topshiriqlari, test tafsiloti, infografiklar.

KORRELYATSIYA TAHLILI: KIMYO VA BIOLOGIYA FANLARIDAN TEST NATIJALARI

A.B. Normurodov, M.Dj. Ermamatov

Bilim va malakalarni baholash agentligi huzuridagi Ilmiy-o‘quv amaliy markazi, 100084, Toshkent sh., Bog‘ishamol k., 12, anormurodov@gmail.com

Qisqacha mazmuni. Ushbu maqolada milliy sertifikat uchun kimyo va biologiya fanlaridan o‘tkazilgan imtihonda qatnashgan talabgorlarning test va yozma ish ballari o‘rtasidagi korrelyatsiya koeffitsiyentlari hisoblangan. Test va yozma ish uchun berilgan ballar o‘rtasidagi korrelyatsiyani aniqlash usullari bayon qilingan va hisoblash natijalari tahlil qilingan.

Kalit so‘zlar: Korrelyatsiya koeffitsiyenti, Pirson korrelyatsiyasi, Spirman korrelyatsiyasi, quti chizmalari.

1. Kirish

Ta‘limda bilimni baholash uchun turli pedagogik o‘lchov vositalaridan foydalaniladi. Masalan, test topshiriqlari orqali qisqa va aniq javoblar olinsa, yozma ishlar orqali bilimlarni chuqur va kengroq aniqlash mumkin bo‘ladi. Test topshiriqlari talabgorlarning bilim darajasini qisqa vaqt ichida baholaydi. Javoblar odatda to‘g‘ri (kalit) yoki noto‘g‘ri (distraktorlar) ko‘rinishda bo‘ladi. Yozma ishda esa talabgorlarning mavzuni chuqur tushunish, fikrlash va mantiqiy tahlil qilish kabi qobiliyatlari aniqlanadi. Odatda javoblar qisqa va kengaytirilgan shaklda bo‘ladi. Ushbu ikkita baholash usuli orasidagi bog‘lanishni (korrelyatsiyani) aniqlash, baholash jarayoni sifatini oshirish va baholash tizimini takomillashtirishda muhim hisoblanadi.

Korrelyatsiya – bu ikki o‘zgaruvchi orasidagi bog‘lanish darajasini aniqlaydigan statistik ko‘rsatkichdir. Korrelyatsiya qiymatiga ko‘ra musbat korrelyatsiya, manfiy korrelyatsiya va nol qiymatli korrelyatsiyalarga bo‘linadi. Musbat korreliyatsiyada birinchi o‘zgaruvchi oshganda, ikkinchisi ham oshadi. Manfiy korrelyatsiyada esa birinchi o‘zgaruvchi oshganda, ikkinchisi kamayadi. Nol qiymatli korrelyatsiyada o‘zgaruvchilar orasida sezilarli bog‘lanish mavjud bo‘lmaydi. Ko‘plab tadqiqotlar korrelyatsiya tahlilidan foydalanib, o‘rganilayotgan o‘zgaruvchilar orasidagi bog‘liqlik darajasini tahlil qiladi. Ayniqsa, ijtimoiy fanlar tadqiqotlarida chiziqli

korrelyatsiya tahlili bir o‘zgaruvchining boshqasiga qanchalik yaqinligini ifodalash uchun keng qo‘llaniladigan vositadir [1]. Chiziqli korrelyatsiya koeffitsiyenti (r) ikki o‘zgaruvchi orasidagi yaqin bog‘liqlik darajasi haqida ma‘lumot beruvchi kattalikdir. Korrelyatsiya tahlilini o‘tkazishdan maqsad deyarli barcha tadqiqotlarda bir xil bo‘lib, ko‘pincha mustaqil va o‘zaro bog‘liq o‘zgaruvchilar orasidagi bog‘liqlikni o‘rganish hisoblanadi.

Statistik tahlillarda o‘zgaruvchilar orasidagi bog‘lanishni aniqlash va baholash muhim ahamiyatga ega. Bunday bog‘lanishni o‘lchash usullaridan biri – Pirson korrelyatsiyasidir. Ushbu usul ikki o‘zgaruvchi orasidagi chiziqli bog‘lanish darajasini aniqlash uchun ishlatiladi. Pirson korrelyatsiyasi - bu ikki o‘zgaruvchi orasidagi chiziqli bog‘lanishni aniqlashda samarali va keng qo‘llaniladigan usul. Biroq uning cheklovlarini hisobga olgan holda, kerak bo‘lganda Spirman kabi alternativ usullarni qo‘llash zarur. Spirman korrelyatsiyasi - ikki o‘zgaruvchi orasidagi nochiziqli (monoton) bog‘lanishni baholashda qo‘llaniladigan usuldir. Bu usul ma‘lumotlar chiziqli bo‘lmaganda yoki normal taqsimlanmagan hollarda samarali hisoblanadi. Spirman korrelyatsiyasi ikki o‘zgaruvchi orasidagi monoton bog‘lanishni samarali baholaydi va o‘zgaruvchilar qiymatlari normal taqsimlanmagan hollarda barqarordir. Agar ma‘lumotlar chiziqli bo‘lmasa yoki normal taqsimlanmagan bo‘lsa, Spirman usulini qo‘llash tavsiya etiladi. Statistik tahlilni to‘g‘ri amalga oshirish uchun Pirson va Spirman usullarini to‘g‘ri tanlash zarurdir.

Ushbu maqolada milliy sertifikat uchun kimyo va biologiya fanlaridan o‘tkazilgan imtihonda qatnashgan talabgorlarning test va yozma ish savollari bo‘yicha ballari o‘rtasidagi korrelyatsiyani aniqlash usullari va hisoblash natijalarini tahlil qilamiz.

2. Natijalar va muhokamalar

Korrelyatsiyalar muhokamasidan oldin aprobatsiya va test sinovlarining quti chizmalarini muhokama qilamiz. Quti chizmasi - (Ing. boxplots) ma‘lumotlar to‘plamining taqsimoti va umumiy statistikasini vizual taqdim qilish uchun kuchli vositadir. Ular ma‘lumotlarning asosiy jihatlarini tezda baholash va xulosa chiqarish imkonini beradi [2].

1- va 2-jadvallarda kimyo va biologiya fanlaridan milliy sertifikat uchun o‘tkazilgan aprobatsiya test sinovlari va haqiqiy test sinovi natijalarining tavsif statistikalari solishtirilgan.

1-jadval

Kimyo fanidan o‘tkazilgan test sinovlarining tavsif statistikasi ma‘lumotlari

Tavsif statistikasi	**Kimyo**		
	Aprobatsiya (22.09.2025)	**Test (25.01.2025)**	
	1	1	2
Test topshiriqlari soni	40	40	40
Test topshiruvchilar soni	482	6399	2611
Kronbax alfa koeffitsiyenti	0,89	0,88	0,86
O‘rta qiymat	18,27	18,61	17,56
Moda	12	12	11
Mediana	16,5	17	16
Dispersiya	66,58	59,97	53,29
Standart tafovut	8,13	7,74	7,3
Asimmetriya	0,640	0,465	0,408
Ekstsess	-0,381	-0,457	-0,579
O‘lchashning standart xatoligi	2,70	2,68	2,73

2-jadval

Biologiya fanidan o‘tkazilgan test sinovlarining tavsif statistikasi ma‘lumotlari

Tavsif statistikasi	**Biologiya**			
	Aprobatsiya (22.09.2025)	**Test (25.01.2025)**		
	1	1	2	3
Test topshiriqlari soni	40	40	40	40
Test topshiruvchilar soni	483	5185	6343	4530
Kronbax alfa koeffitsiyenti	0,91	0,89	0,88	0,87
O‘rta qiymat	25,06	21,99	21,96	21,51
Moda	30	29	22	19
Mediana	26	22	22	21
Dispersiya	73,86	65,07	59,08	56,52
Standart tafovut	8,59	8,07	7,69	7,52
Assymmetriya	-0,327	-0,002	0,010	0,066
Ekstess	-0,927	-0,909	-0,866	-0,792
O‘lchashning standart xatoligi	2,58	2,68	2,66	2,71

1- va 2-jadvallarda kimyo va biologiya fanlaridan aprobatsiya va test sinovlarining tavsif statistikasi ma‘lumotlari biri biriga juda yaqin ekanligini ko‘rish mumkun, bu kimyo va biologiya fanlaridan aprobatsiyada guruhlarni reprezentativligi yaxshi ta‘minlanganligini anglatadi.

Quti chizmasida mediana qutining o‘rtasiga yaqinligi taqsimotning nisbatan simmetrikligini bildiradi. Quti o‘rtasiga nisbatan tepaga yoki pastga siljiganlik taqsimotning mos yo‘nalishda surilganligini bildiradi.

Qutining o‘zi inter-kvartil diapazonni (IKD) bildiradi. IKD ma‘lumotlarning birinchi (Q1) va uchinchi (Q3) kvartillar orasidagi 50 foizni bilidiradi: IKD = Q3-Q1.

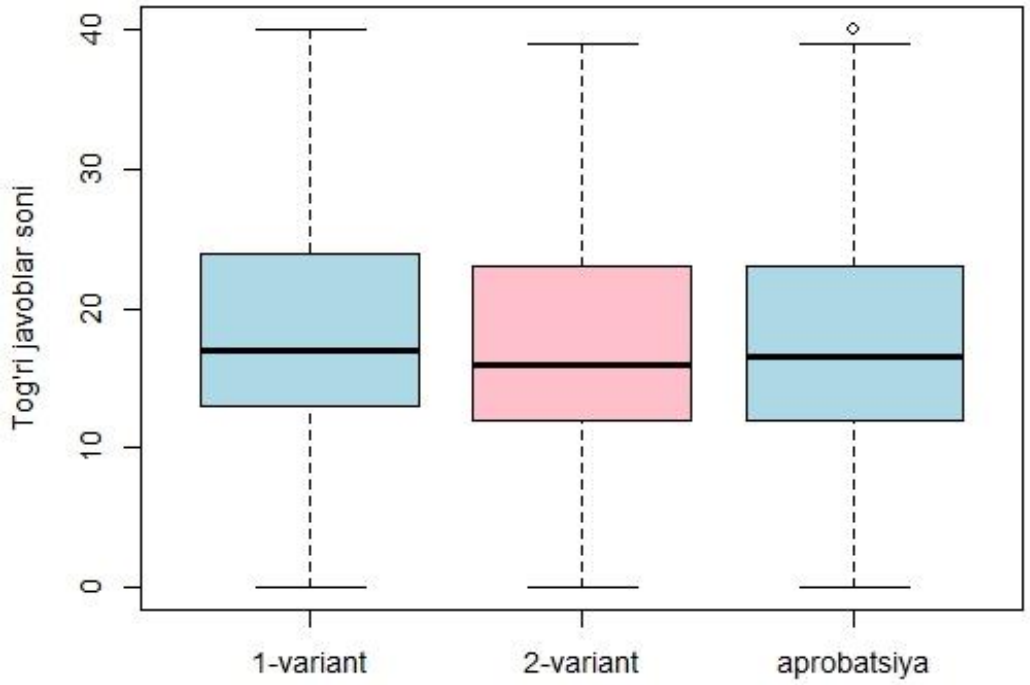
Kengroq quti ma‘lumot ko‘proq tarqalganligini, tor quti esa kamroq tarqalganligini (ma‘lumot nuqtalari konsentratsiyasi kattaligini) bildiradi.

Mediana va quti chiziqlarining qutiga nisbatan joylashuviga qarab taqsimot cho‘zilganligini baholash mumkin. Medianani pastki kvartil (Q1) ga yaqinligi va quti tepasidagi chiziqning uzunligi taqsimotning chapga surilganligini (musbat assimmetriya) bildiradi. Medianani yuqori kvartil (Q3) ga yaqinligi va qutining pastidagi chiziqning uzunligi esa taqsimotning o‘ngga surilganligini (manfiy assimmetriya) anglatadi. Agar mediana qutini o‘rtasida bo‘lsa va qutining yuqori va pastki chiziqlari uzunligi teng bo‘lsa, taqsimot simmetrik bo‘ladi.

Quti chiziqlari chekkasida joylashgan nuqtalar (Ing. outliers) chekka nuqtalar deb ataladi. Chekka nuqtalarning quyi va yuqori chegarasi mos ravishda Q1-1.5*IKD hamda Q3+1.5*IKD formulalar orqali aniqlanadi. Undan tashqari taqsimotda ekstrim chekka nuqtalar ham bo‘lishi mumkin, ular esa Q1-3*IKD hamda Q3+3*IKD formulalar bilan hisoblanishi mumkin. Chekka nuqtalar o‘rganilishi lozim bo‘lgan “maxsus hol”, “o‘lchash xatoligi”, “ekstrim kuzatishlar” kabi holatlarni ko‘rsatishi mumkin. Shuningdek, chekka nuqtalar ma‘lumot yig‘ishdagi potensial muammolar yoki ma‘lomotlarning o‘ta o‘zgaruvchanligi tufayli paydo bo‘lishi ham mumkin.

Kimyo va biologiya fanlaridan o‘tkazilgan aprobatsiya va milliy sertifikat uchun o‘tkazilgan test sinovlarida ishlatilgan test variantlarining tavsif statistikasi ma‘lumotlari mos ravishda 1- va 2-jadvallarda berilgan.

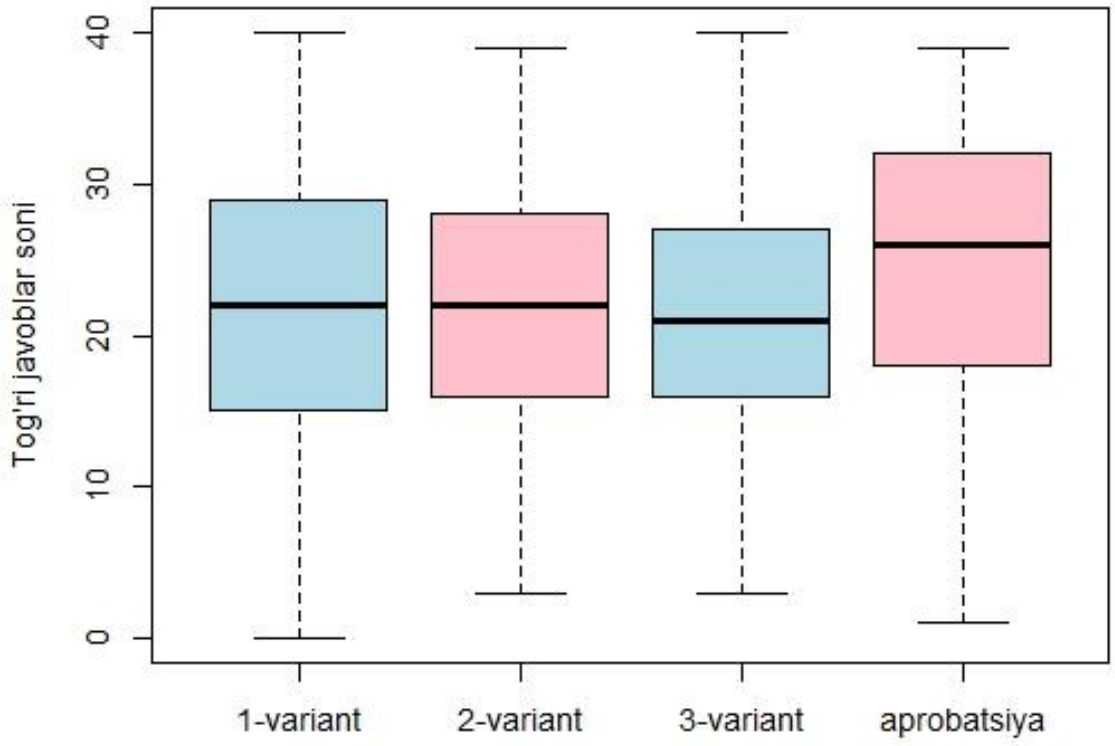
1-rasmda kimyo fanidan o‘tkazilgan test sinovlarida ishlatilgan test variantlarining quti chizmasi keltirilgan.



1-rasm. Kimyo fanidan test va aprobatsiya variantlaridan olingan natijalarining quti chizmalari

Quti chizmasining yuqoridagi talqinidan foydalanib, 1-rasmdan vizual xulosalar chiqarishimiz mumkin. Avvalo 1-rasmdan medianalar 1- va 2-variant hamda aprobatsiyada ishlatilgan variantlarning medianalari bir-biriga juda yaqin, bu variantlarning deyarli parallelligini bildiradi. Quti chizmasida barcha variantlarda mediana qutining o‘rtasidan biroz pastda ekanligi va quti tepasidagi chiziqning uzunligi asimmetriya musbatligini ko‘rsatadi va buni 1-jadvaldagi asimmetriya qiymatlari ham tasdiqlaydi.

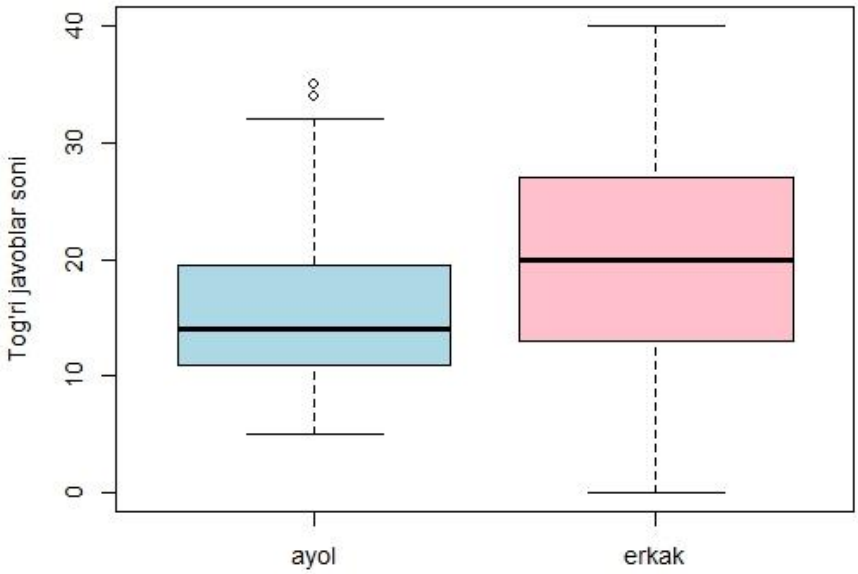
Qutilarning qalinligi bir-biriga juda yaqin ekanligi ularning IKD lari qiymatlari bir-biriga yaqin ekanligini anglatadi. Qutilarning qalinligi qiymatining kattaligi esa ma’lumotlar yaxshi tarqalganligini bildiradi. 1–jadvaldagi dispersiyalar qiymat-larini solishtirish ham shu natijani beradi. Shuningdek, 1-rasmdan apro-batsiya test sinovida 1 ta chekka nuqta mavjudligi ko‘rinadi.



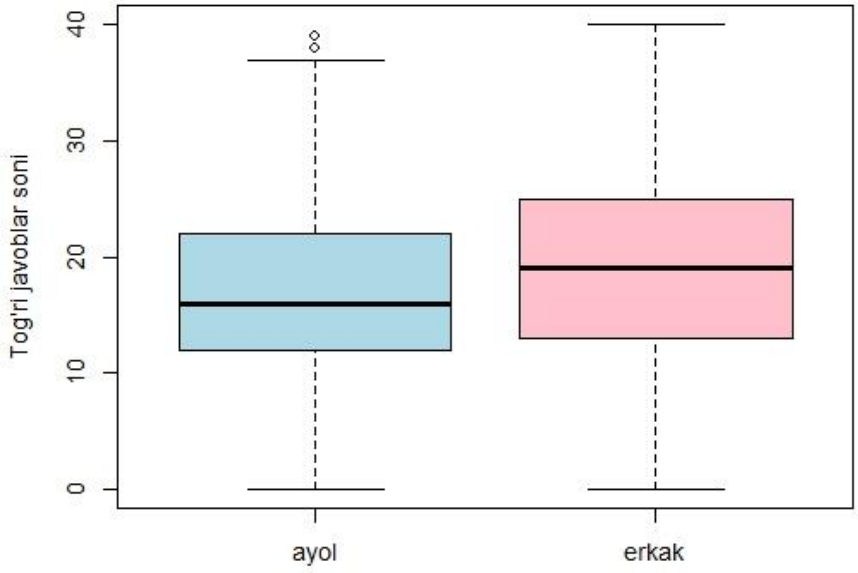
2-rasm. Biologiya fanidan test va aprobatsiya variantlaridan olingan natijalarining quti chizmalari

2-rasmda biologiya fanidan o‘tkazilgan test sinovlarida ishlatilgan test variantlarining quti chizmasi keltirilgan. Kimyo fanidagi natijalardan farqli ravishda, biologiya fanidan aprobatsiya test sinovida mediana quti o‘rtasidan balandroqda joylashgan va quyining pastki qismidagi chiziq uzunroq bo‘lib, bu manfiy asimmetriya kattaligini bilidiradi. Qolgan 3 ta variantning medianalari qutilarning o‘rtasiga juda yaqinligi, qutining tepasi va pastidagi chiziqlar uzunligining bir-biriga yaqinligi taqsimot deyarli simmetrikligini bildiradi. Haqiqatan, 2-jadvaldagi asimmetriya qiymatlari ham shuni ko‘rsatadi.

Shunga o‘xshash, aprobatsiyada va testda qatnashgan ayol va erkak talabgorlarning quti chizmasini ham tahlil qilish mumkin. 3-rasmda kimyo fanidan aprobatsiya (a) va test sinovlarida (b) qatnashgan ayol va erkak talabgorlar to‘g‘ri javoblarining quti chizmalari solishtirilgan.



a)



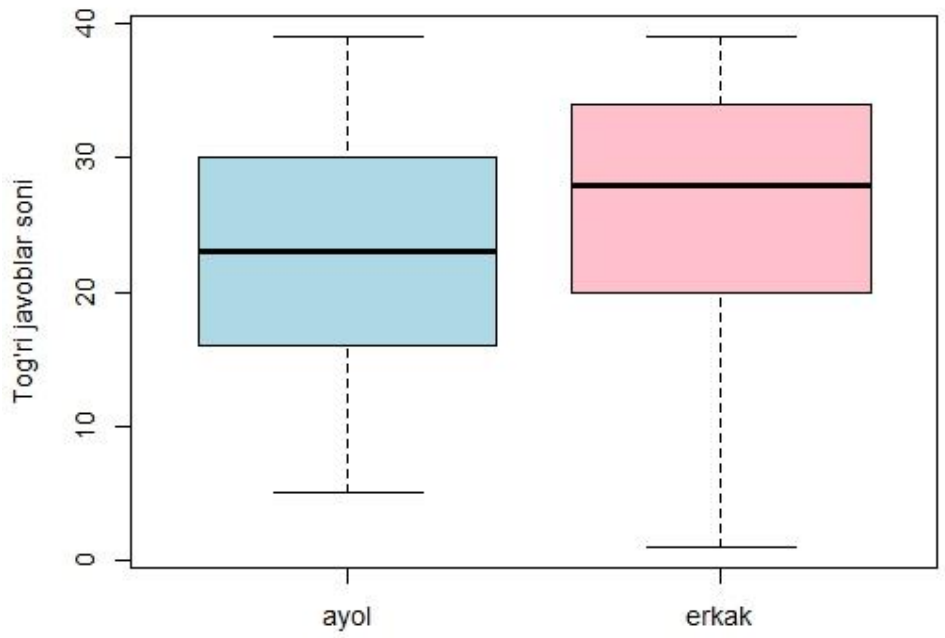
b)

3-rasm. Kimyo fanidan aprobatsiya (a) va test (b) sinovida qatnashgan talabgorlar to‘g‘ri javoblarining quti chizmalari

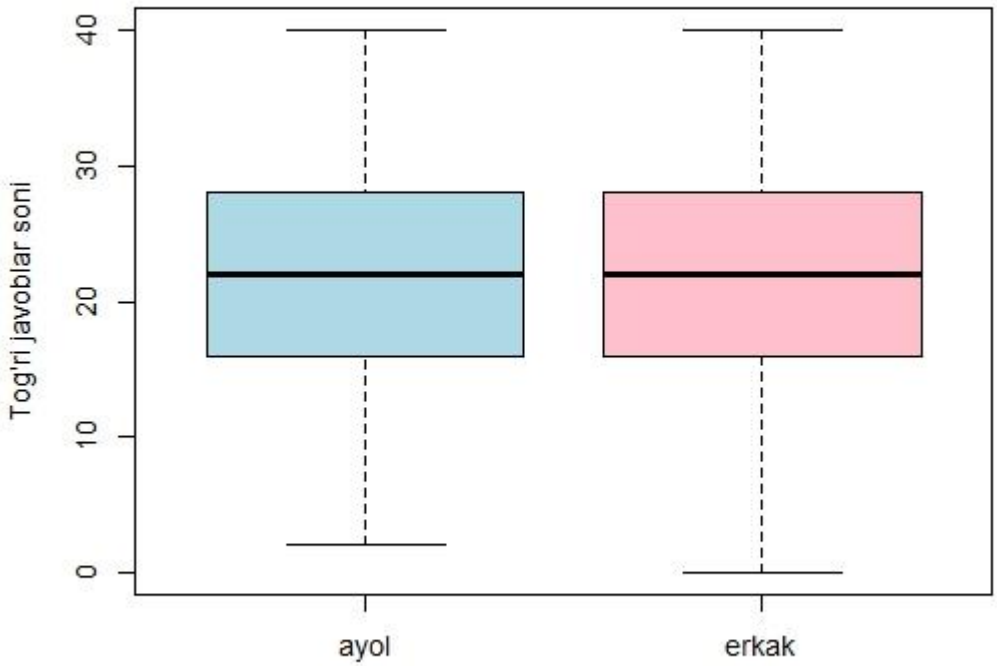
3-rasmdan aprobatsiya test sinovlarida medianalarning qiymatlari bir-biriga yaqin emasligini, test sinovlarida esa ular bir-biriga

yaqinroq kelganini, lekin aprobatsiyada ayollar uchun qutilar qalinligi erkaklar uchun quti qalinligiga nisbatan kichikroqligini ko‘rish mumkin. Shuningdek, aprobratsiya test sinovlarida ham medianalar ayollar uchun qutining o‘rtasidan biroz pastda joylashgan. Erkaklar uchun esa medianalar quti o‘rtasiga yaqin bo‘lib, bu o‘rtadagi taqsimot simmetrik ekanligini bildiradi. Ayollar uchun quti tepasidagi chiziq aprobatsiyada ham testda ham pastdagiga nisbatan uzun bo‘lib, bu taqsimot cho‘zilganini anglatadi. Erkaklar uchun hatto taqsimot chekkalarida ham simmetriya borligi - rasmdagi quti chiziqlari uzunliklari bir-biriga yaqinligidan ko‘rinib turibdi.

Test sinovidagi holat aprobatsiyadagidan ancha yaxshi: aprobatsiyda 482 nafar talabgor (219 nafar ayol va 263 nafar erkak) ishtirok etganligi hisobga olinsa, aprobatsiya natijalari test natijalarini yaxshi bashorat qilgan deyish mumkin (qiyoslash uchun test sinovlarida 4821 nafar ayol va 4189 nafar erkak talabgorlar, jami 9010 nafar talabgorlar ishtirok etganlar). 3-rasmdan aprobatsiyada ham, test sinovlarida ham ayollar uchun chekka nuqtalar mavjudligini ko‘rish mumkin.



a)



b)

4-rasm. Biologiya fanidan aprobatsiya (a) va test (b) sinovida qatnashgan talabgorlar to‘g‘ri javoblarining quti chizmalari

Biologiya fani bo‘yicha aprobatsiyada erkaklar uchun mediana quti o‘rtasidan biroz balandroqda joylashgan. Erkaklar uchun qutilarning qalinligi bir-biriga yaqin bo‘lsada, erkaklar uchun quti ayollarga nisbatan biroz balandda joylashgan, quti tepasidagi chiziquzunliklari esa bir-biridan ancha farq qiladi va bu erkaklar uchun taqsimot cho‘zilganligini bilidiradi. Test sinovlarida esa medianalar qutining o‘rtasiga yaqin, qutilarning qalinligi deyarli bir xil va ularning yuqori va pastidagi chiziqlarining uzunliklari bir-biriga juda yaqin. Bu biologiyadan test sinovlarida test topshiriqlari ayollar uchun erkaklarga nisbatan (yoki aksincha) ustunlik bermasligini anglatadi.

Quti chizmalarini elementning differentsial ishlash ko‘rsatkichi (*Ing.* Differential Item Functioning - DIF) kontekstida jinslar bo‘yicha taqqoslash bilan biz test sinovlarida qatnashgan guruhlar o‘rtasida sistematik farqlar mavjud yoki mavjud emasligini baholashimiz mumkin. Bu test variantidagi savollar biror guruhga boshqasiga nisbatan(yuqoridagi holda, ayol va erkaklar guruhi) imtiyoz yoki to‘siq berish bermasligiga (*ing.* bias)

ishonch hosil qilish uchun muhim, chunki bu baholashning eng muhim adolat mezoni bilan bog‘liq.

Quti chizmalarida DIF media-nalarining, IKDlarining, quti qalin-liklarining bir-biridan keskin farq qilishi, chekka nuqtalarning biror guruhda mavjud bo‘lishi - test topshirig‘i alohida olingan guruhlarga imtiyoz yoki to‘siq tufayli bo‘lishi mumkin.

Kimyo fani bo‘yicha o‘tkazilgan aprobatsiyada va test sinovida 2 tadan chekka nuqta mavjud. Biologiya fani bo‘yicha DIF muammosi quti chizmasi usulida ko‘zga tashlanmaydi.

Statistik tahlilda to‘g‘ri usulni tanlash orqali tadqiqot natijalarining ishonchliligini oshirish mumkin. Pirson korrelyatsiyasi - bu ikki o‘zgaruvchi orasidagi chiziqli bog‘lanish darajasini o‘lchaydigan statistik ko‘rsatkich bo‘lib, u birinchi marta 1895-yilda ingliz statisti Karl Pirson tomonidan taklif qilingan [3,4]. Pirson korrelyatsiyasi odatda Pirsonning r - koeffitsiyenti ham deb ataladi va u −1 dan +1 gacha qiymatlar oralig‘ida bo‘ladi. Buni quyidagicha tavsiflash mumkin. r=(+1) to‘liq musbat chiziqli bog‘lanish bo‘lib, birinchi o‘zgaruvchi oshganda, ikkinchisi ham proporsional ravishda oshadi. r = (−1) esa to‘liq manfiy chiziqli bog‘lanish bo‘lib, birinchi o‘zgaruvchi oshganda ikkinchisi proporsional ravishda kamayadi. r = 0 da hech qanday chiziqli bog‘lanish mavjud emas.

Pirsonning r koeffitsiyenti quyidagi formula orqali hisoblanadi [3]:

$$r = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum(X_i - \overline{X})^2 \sum(Y_i - \overline{Y})^2}} \quad (1)$$

bu erda:

X_i va Y_i – i-chi talabgorning qiymatlari.

$\overline{X}$ va $\overline{Y}$ – x va y larning o‘rtacha qiymatlari.

Umuman olganda Pirson korrelyatsiyasidan chiziqli bog‘-lanishni aniqlash uchun foydalaniladi. Bundan tashqari o‘zgaruvchilar orasidagi to‘g‘ridan-to‘g‘ri (chiziqli) munosabat mavjudligini tekshirish, bog‘lanish yo‘nalishini musbat yoki manfiy ekanligini, bog‘lanishning kuchli yoki kuchsiz ekanini aniqlash uchun ham ishlatiladi [4]. Demak Pirson korrelyatsiyasi ma‘lumotlar chiziqli bog‘langan bo‘lsa, o‘zgaruv-chilar interval yoki nisbat (*ing*. ratio) ko‘rinishida bo‘lsa, normal taqsimotga yaqin bo‘lsa, keskin o‘zgaruvchi qiymatlar kam bo‘lsagina foydalaniladi. Agar bog‘lanishlar monoton (qandaydir tartibda o‘sishi yoki kamayishini) bo‘lsa bunda Spirman korrelyatsiyasidan foydalaniladi.

Spirman korrelyatsiyasi (*ing* . spirman‘s rank correlation coefficient) – ikki o‘zgaruvchi orasidagi monoton (lekin chiziqli bo‘lishi shart emas) bog‘lanish darajasini o‘lchaydigan statistik usuldir [5]. U Spirmanning reyting raqamlari (asl qiymatlar orasidagi korrelyatsiya emas, balki

ularning reyting raqamlari orasidagi korrelyatsiya hisoblanadi) korrelyatsiya koeffitsiyenti yoki Spirmanning ro (ρ) – koeffitsiyenti deb ham ataladi. Bu usul Karl Pirson korrelyatsiyasiga alternativ bo‘lib, ayniqsa, chiziqli bog‘lanish talabi bajarilmagan hollarda qo‘llaniladi. Odatda monoton bog‘lanish ikki turga yani musbat va manfiy monoton bog‘lanishlarga bo‘linadi. Musbat monoton bog‘lanishda bir o‘zgaruvchi oshganda, ikkinchisi ham uzluksiz oshadi yoki kamayadi, manfiy monoton bog‘lanishda esa bir o‘zgaruvchi oshganda, ikkinchisi uzluksiz kamayadi. Spirman korrelyatsiyasi ham −1 dan +1 gacha qiymatlar oralig‘ida bo‘ladi. Agar Spirman korrelyatsiyasi koeffitsiyenti ρ = (+1) bo‘lsa to‘liq musbat monoton bog‘lanish, ρ = (−1) bo‘lsa to‘liq manfiy monoton bog‘lanish, ρ ≈ 0 bo‘lsa hech qanday monoton bog‘lanish bo‘lmaydi.

Spirman korrelyatsiyasi quyidagi 2-formula orqali hisoblanadi [4]:

$$\rho = 1 - \frac{6\sum d_i^2}{N(N^2 - 1)} \qquad (2)$$

bu yerda:

d_i –ikki o‘zgaruvchining i-chi reyting raqamlari orasidagi farq.

N – o‘zgaruvchilar soni.

Spirman korrelyatsiyasidan chiziqlilik sharti bajarilmaganda, o‘zgaruvchi qiymatlar mavjud bo‘lganda, ma‘lumotlar tartibli (*ing*. ordinal) bo‘lganda va ma‘lumotlar normal taqsimlanmagan holatlarda foydalaniladi. Umumiy holda qiymatlar o‘rtasidagi bog‘lanish darajalariga mos Pirson va Spirman korrelyatsiya koeffitsiyentlarining qiymatlari 3-jadvalda keltirilgan [5].

3-jadval

Korrelyatsiyasi koeffitsiyenti qiymatlarining bog‘lanish bo‘yicha darajalari

Korrelyatsiya koeffitsiyenti (r yoki ρ)	**Korrelyatsiya**
0,9 dan 1	Juda kuchli
0,7 dan 0,9	Kuchli
0,5 dan 0,7	O‘rtacha
0,3 dan 0,5	Kuchsiz
0 dan 0,3	Juda kuchsiz

Statistik tahlillarda ikkita ma‘lumot to‘plami orasida bog‘liqlik mavjudligi to‘g‘risidagi savolga javob korrelyatsiya yordamida olinadi. Bizda bu milliy sertifikat uchun kimyo va biologiya fanlaridan o‘tkazilgan imtihonda qatnashgan talabgorlarning test topshiriqlari va yozma ishlari

uchun olgan ballaridir. Test topshiriqlari ballarini *X* bilan, yozma ish ballarini esa *Y* bilan belgilab olamiz. *X* va *Y* ma‘lumotlar to‘plami o‘rtasidagi moslik darajasini ifodalash uchun kovariatsiya deb ataluvchi maxsus o‘lchov qo‘llaniladi. “Kovariatsiya” tushunchasining ma‘nosini talabgorlar tomonidan bajarilgan ikki – *X* va *Y* ya‘ni test va yozma savollar natijalari asosidagi ballari misolida tushuntirish mumkin. Birinchi *X* test topshiriqlari bo‘yicha natijalar X_i *(i = l, 2, ..., N)*, ikkinchi *Y* yozma savollar natijalari bo‘yicha Y_i *(i = l, 2, ..., N)* bo‘lsin [7]. U holda ushbu test va yozma ish bo‘yicha talabgorlar natijalari orasidagi bog‘liqlik o‘lchovini aniqlash uchun har bir talabgorning *X*-test va *Y*-yozma ish natijalari bo‘yicha ballarining o‘rta qiymatiga nisbatan taqqoslash kerak. Test topshiriqlari (*X*) va yozma ish (*Y*) bo‘yicha *i*-talabgor natijalarining moslik darajasi tafovutlar ko‘paytmasi $(X_i - \overline{X})(Y_i - \overline{Y})$ orqali aniqlanadi. Bu yerda X_i, Y_i – mos ravishda (*i = 1, 2, ..., N*) test va yozma ish bo‘yicha *i*-talabgorning natijalari, $\overline{X}$, $\overline{Y}$ – *X* va *Y* test va yozma ish natijalari bo‘yicha ballarning o‘rta qiymatlari, *N* – talabgorlar soni.

X – test va *Y* – yozma ish bo‘yicha X_i va Y_i *(i = 1, 2, ..., N)* qiymatlar to‘g‘ri bog‘liqlikda bo‘lganda barcha ko‘paytmalar yig‘indisi katta va musbat qiymatga ega bo‘ladi, ya‘ni,

$$\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y}) \quad (3)$$

Test va yozma ish natijalari qiymatlari teskari bog‘liq bo‘lganda ushbu summa noldan kichik (manfiy) bo‘lib, absolyut qiymati katta bo‘ladi.

Agar *X* – test va *Y* – yozma savollar natijalari bo‘yicha ballar orasida tizimli bog‘liqlik kuzatilmasa, tafovutlar ko‘paytmasi $(X_i - \overline{X})(Y_i - \overline{Y})$ ishorasi o‘zgarib turadi. Bunday hollarda, talabgorlarning yetarlicha katta tanlanma guruhida quyidagi yig‘indi nolga yaqin bo‘ladi:

$$\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y}) \quad (4)$$

Shunday qilib, tafovutlar ko‘paytmasining $(X_i - \overline{X})(Y_i - \overline{Y})$ ishorasi va absolyut qiymati *X* – test va *Y* – yozma ish natijalari bo‘yicha ma‘lumotlar to‘plami orasidagi bog‘liqlik xarakterini aks ettiradi. Tafovutlar ko‘paytmasi yig‘indisini tanlanmadagi talabgorlar soniga bo‘lib, tanlanma o‘lchamiga bog‘liq bo‘lmagan, kovariatsiya deb ataluvchi o‘rtacha kattalik olinadi. Uni turli hajmdagi tanlanmalar bo‘yicha test natijalari o‘lchovlari orasidagi bog‘liqlikni taqqoslashda qo‘llash mumkin bo‘ladi. Shunday qilib, kovariatsiya quyidagi formula bo‘yicha hisoblanadi:

$$S_{xy} = \frac{\sum_{i=1}^{N}(X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{N - 1} \quad (5)$$

Turli dispersiyali tanlanmalar bo‘yicha bog‘liqlik ko‘rsatkichlari qiymatlari taqqoslanishini oshirish uchun kovariatsiya standart tafovutga bo‘linadi. Shunday qilib, S_{xy} kovariatsiya S_x va S_y ga bo‘linadi, bunda S_x va S_y mos ravishda X va Y to‘plamlar bo‘yicha standart tafovutlardir.

Natijada, o‘zgartirishlardan keyin, (1) – formuladan foydalanib Pirson korrelyatsiya koeffitsiyenti r_{xy} – kattalik olinadi (6-formula).

$$\boldsymbol{r_{xy} = \frac{S_x}{S_x S_y} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2 \sum_{i=1}^{N}(Y_i - \bar{Y})^2}}}, \quad (6)$$

bu erda:

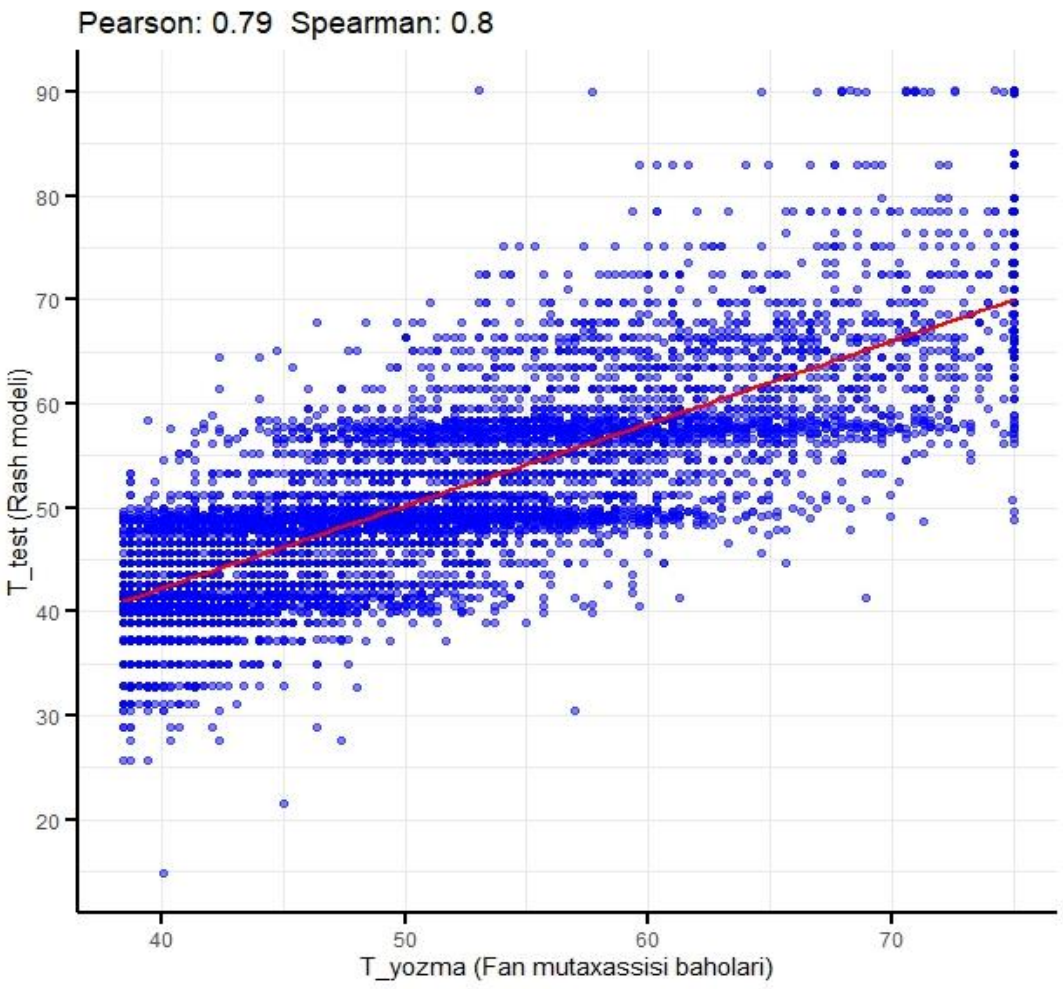
$\bar{X}$ – test ballari bo‘yicha olingan o‘rtacha qiymat.

$\bar{Y}$ – yozma ish ballari bo‘yicha olingan o‘rtacha qiymat.

Bu formula orqali milliy sertifikat uchun kimyo fanidan o‘tkazilgan test sinovida qatnashgan talabgorlarning test va yozma ish ballari o‘rtasidagi korreliyatsiyasi hisoblandi. Quyidagi 5-rasmda talabgorlarning test va yozma ish bo‘yicha T- ballar korreliyatsiyasi berilgan. 0 ball olgan talabgorlar korreliyatsiya ko‘rsatkichlariga ta‘sir qilganligi uchun tahlildan chiqarilib hisoblandi. Bunda Pirson korrelyatsiyasi koeffitsiyenti $r_{xy} = 0,79$, Spirman korrelyatsiyasi koeffitsiyenti $\rho = 0,80$ qiymatiga teng bo‘ldi.

Pirson korrelyatsiyasi koeffitsiyentining qiymati 0,79 ga tengligi test sinovida qatnashgan talabgorlarning test va yozma ish ballari orasida kuchli bog‘liqlik (1-jadval) borligini ko‘rsatadi.

Spirman korrelyatsiyasi koeffitsiyenti qiymatining 0,80 ga tengligi esa talabgorlarning test va yozma ish ballari o‘rtasida kuchli bog‘liqlik borligini bildiradi.



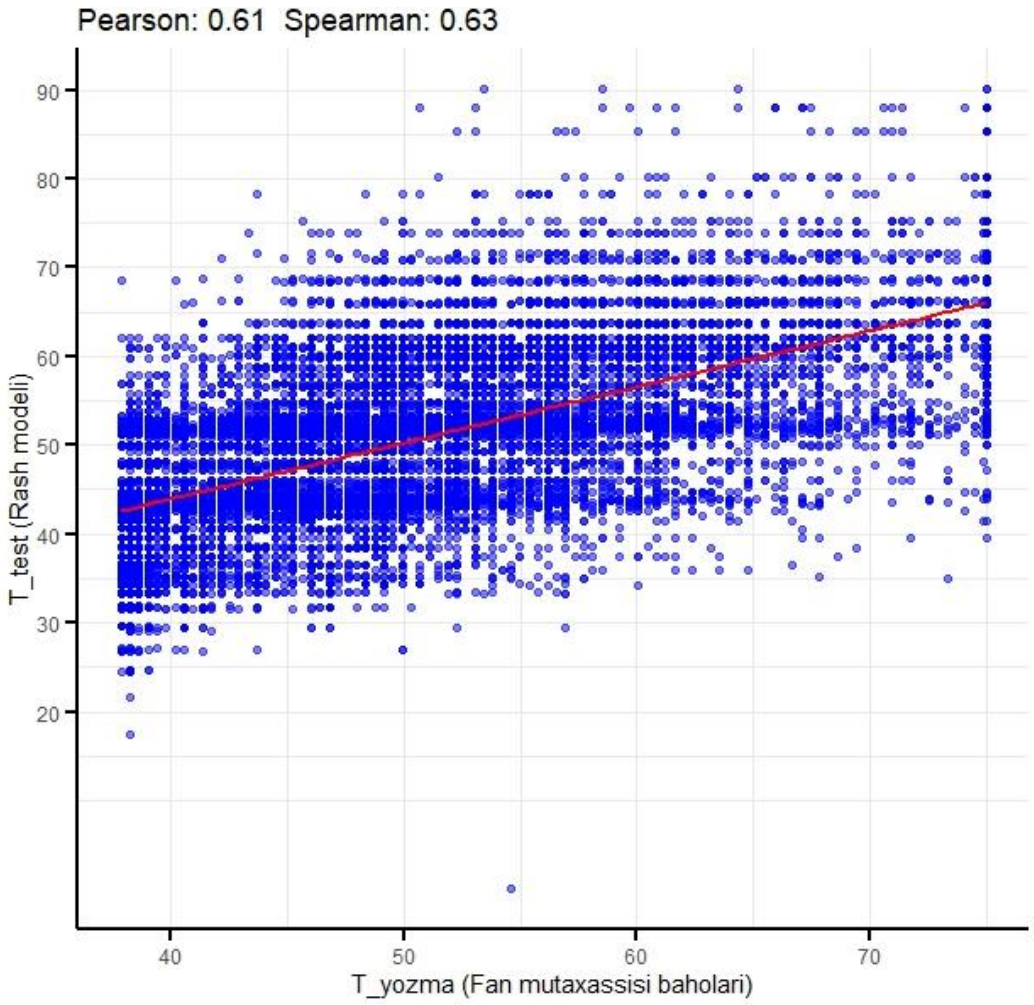
5-rasm. Kimyo fanidan o‘tkazilgan test sinovlarida qatnashgan talabgorlarning test (T_test) va yozma ish (T_yozma) ballari o‘rtasidagi korreliyatsiya

5-rasmdagi grafikda talabgorlarning fan mutaxassislari tomonidan qo‘yilgan yozma ish ballari (T_yozma) va Rash modeli bilan baholangan test ballari (T_test) orasidagi bog‘liqlik ko‘rsatilgan. Umuman olgana, rasmdan ballar o‘zaro bog‘ligi, ya‘ni Rash modeli bilan baholangan qobiliyatlar yuqori (past) bo‘lsa, yozma ishga fan mutaxassislari ham yuqori (past) ball qo‘yganliklarini ko‘rish mumkin. Biroq past yoki yuqori qobiliyatlar sohasida farqlar borligi ko‘rinadi.

Milliy sertifikat uchun biologiya fanidan o‘tkazilgan test sinovidagi talabgorlarning test va yozma ish ballari o‘rtasidagi korreliyatsiya hisoblandi (6-rasm). 0 ball olgan talabgorlar korreliyatsiya ko‘rsatkichlariga ta‘sir qilganligi uchun tahlildan chiqarilib hisoblandi. 6–rasmdagi grafikda talabgorlarning Rash modeli bilan baholangan test topshiriqlari asosidagi (T_test) va fan mutaxassislari tomonidan qo‘yilgan yozma savollar asosidagi haqiqiy ballari (T_yozma) orasidagi bog‘liqlik ko‘rsatilgan. Bunda talabgorlarning test topshiriqlari va yozma savollar natijari asosidagi T-

haqiqiy ballar korreliyatsiyalari berilgan. Pirson korrelyatsiyasi koeffitsiyenti r_{xy} = 0,61, Spirman korrelyatsiyasi koeffitsiyenti ρ = 0,63 qiymatiga teng.

Pirson korrelyatsiyasi koeffitsiyentining qiymati 0,61 ga tengligi test sinovidagi talabgorlarning test va yozma savollar natijasi asosidagi ballari orasida o‘rtacha bog‘liqlik (3-jadval) borligini ko‘rsatadi.



6-rasm. Biologiya fanidan o‘tkazilgan test sinovlarida qatnashgan talabgorlarning test (T_test) va yozma ish (T_yozma) ballari o‘rtasidagi korreliyatsiya

Spirman korrelyatsiyasi koeffitsiyenti qiymatining ρ = 0,63 ga tengligi esa talabgorlarning reytingi bo‘yicha test va yozma ish ballari orasida nisbatan kuchliroq bog‘liqlik borligini bildiradi. Spirman korreliyatsiya koeffitsiyenti ko‘rsatkichi Pirson korrelyatsiyasi koeffitsiyentiga nisbatan yuqoriroq bo‘lgani uchun, tartib jihatidan o‘xshashlik bor, lekin individual qiymatlarda farqlar mavjud. Grafikdagi vertikal chiziqlar, ba‘zi talabgorlarning yozma ish ballari 0 ga

yaqin yoki juda past bo‘lganda, test ballarining katta ekanligini ko‘rsatadi. Keskin farq qiluvchi qiymatlar mavjudligi intuitiv baholashni mukammal amalga oshirish mushkul ekanligini anglatadi.

Bunday qiymatlar ma‘lumot to‘plamidagi boshqa qiymatlardan keskin farq qiladigan va umumiy tendensiyaga mos kelmaydigan qiymatlar hisoblanadi. Ular odatda juda katta yoki juda kichik bo‘lishi mumkin. Pirson korrelyatsiyasi keskin farq qiluvchi qiymatlarga juda sezgir bo‘ladi, Spirman korrelyatsiyasi esa bunday qiymatlarga unchalik sezgir emas. Agar ma‘lumotda bitta yoki bir nechta keskin farq qiluvchi qiymatlar bo‘lsa, bu korrelyatsiya koeffitsiyentini past bo‘lishiga olib keladi.

Xulosa

Kimyo fanidan o‘tkazilgan test sinovlarida qatnashgan talabgorlarning test (T_test) va yozma ish ballari (T_yozma) o‘rtasida kuchli bog‘liqlik bor ekanligi aniqlandi.

Spirman korreliyatsiya koeffitsiyenti ko‘rsatkichi yuqori ekanligi, talabgorlarning qobiliyatlari Rash modeli va fan mutaxassislari tomonidan deyarli bir xil aniqlanganlini ko‘rsatadi. Shunga qaramay, talabgorlarning test va yozma ish bo‘yicha ballari o ‘rtasida tafovut har doim ma‘lum miqdorda bo ‘lishi tabiiy hol, chunki 0 xatolikka erishish imkoniyat amaliyotda mavjud emas.

Biologiya fanidan o‘tkazilgan test sinovlarida qatnashgan talabgorlarning test (T_test) va yozma ish ballari (T_yozma) o‘rtasida sezilarli bog‘liqlik bor.

Spirman korreliyatsiya koeffitsiyenti ko‘rsatkichi yuqori bo‘lgani uchun, tartib jihatidan o‘xshashlik bor, lekin individual qiymatlarda farqlar katta.

Grafikdagi vertikal chiziqlar, ba‘zi talabgorlarning yozma ish ballari 0 ga yaqin yoki juda past bo‘lganda, test ballarining katta ekanligini ko‘rsatadi.

ADABIYOTLAR

1. S. Senthilnathan. Usefulness of Correlation Analysis. Available at SSRN, 2019, pp.9. doi: http://dx.doi.org/10.2139/ssrn.3416918
2. Mcgill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of Box Plots. The American Statistician, 32(1), 12–16. https://doi.org/10.1080/00031305.1978.10479236
3. K. Pearson. Notes on the history of correlation. Biometrika, 1920, 13: 25–45.
4. J. Hauke and T. Kossowski. Comparison of values of Pearson‘s and Spearman‘s correlation coefficients on the same sets of data. Questiones Geographicae, 2011, 30(2), pp. 87-93, doi: http://dx.doi.org/10.2478/v10117-011-0021-1 .
5. C. Spearman. The Proof and Measurement of Association between Two Things The American Journal of Psychology , 1904, Vol. 15, No. 1, pp. 72-101
6. D.E. Hinkle, W. Wiersma, S.G. Jurs. Applied Statistics for the Behavioral Sciences. 5th ed. Boston: Houghton Mifflin, 756 pages, 2003.
7. M.Sh. Zakirova. Test topshiriqlari sifatining statistik tahlili. Uslubiy qo‘llanma, Toshkent, 2019, 57 b.

CORRELATION ANALYSIS: TESTS ON CHEMISTRY AND BIOLOGY

A.B. Normurodov, M.Dj. Ermamatov

Scientific and Educational Practical Center Under the Agency for Assessment of Knowledge and Competences, anormurodov@gmail,com

Abstract. In this paper, the correlation coefficients between the test scores and the written work scores of candidates who participated in the chemistry and biology exams for the national certificate,are calculated. The methods for determining the correlation between the scores given for the test and the written work are described, and the calculation results are analyzed.

Keywords: Correlation coefficient, Pearson correlation, Spearman correlation, boxplots

BILIMLARNI BAHOLASHDA SHKALALASHNING AHAMIYATI

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. Maqolada shkala va uning turlari, sinaluvchilarning xom ballarini z- va T-ballga aylantirish hamda ona tili va adabiyot fanidan o‘tkazilgan test natijalarining xom ballarini z- va T-ballarga o‘tkazish ya’ni bir xil shkalaga keltirish usullari muhokama qilingan.

Kalit so‘zlar: test topshiriqlari, standart og‘ish, test topshiriqlarining qiyinlik darajasi, o‘rtacha ball, shkalalar turlari, z-shkala, T-shkala, persentil shkala

I. Kirish

Ma’lumki o‘quv jarayonida ta’lim sifatini baholashning ishonchli va ob’ektiv vositalarini ishlab chiqish, hamda sinaluvchilar kognitiv ko‘nikmalarini o‘lchashda adolatlilik va shaffoflik tamoyillarini ta’minlash muhim ahamiyatga ega. Jamiyatda ro‘y berayotgan o‘zgarishlar, ta’lim sifatini oshirish uchun maxsus o‘lchov vositalari, ilmiy metodlar va zamonaviy pedagogik texnologiyalardan foydalanish muhim ekanini ko‘rsatmoqda. Shu sababli, an’anaviy nazorat va baholash tizimi bilan bir qatorda, ta’lim jarayoniga yangi pedagogik o‘lchash tizimlari joriy etilmoqda.

Sinaluvchilarning bilim, malaka va ko‘nikmalarini o‘lchashdan maqsad esa o‘lchanayotgan o‘zgaruvchi bilan bog‘liq miqdoriy – raqamli qiymatlarni olishdan iborat. Bu o‘zgaruvchi – sinaluvchilarning o‘quv dasturini o‘zlashtirganlik ko‘rsatkichi bo‘lib, u ma’lum darajada sinaluvchining yashirin xususiyati-tayyorgarlik darajasini aks ettiradi.

Raqamli qiymatlarni olish sifat xususiyatlarini o‘zgartirish orqali amalga oshiriladi va buning uchun turli shkalalash usullari qo‘llaniladi. Bu shuni anglatadiki, sifat, ya'ni tavsifiy yoki sifatiy ko‘rsatkichlar sonli - miqdoriy qiymatlarga aylantiriladi. Buni amalga oshirish uchun maxsus shkalalash usullari ishlatiladi. Sinaluvchilar bilim va fan dasturini o‘zlashtirish darajasiga ko‘ra “a’lo”, “yaxshi”, “o‘rtacha” yoki “qoniqarli”, “qoniqarsiz” deb baholanishi mumkin. Bu sifat tavsiflarini aniqroq va taqqoslanadigan miqdoriy qiymatlarga

aylantirish uchun ularga raqamli qiymatlar beriladi.

Shkalalash – ma'lum ob'yektlarning xususiyatini sonli tizimlar yordamida modellashtirish usulidir. Ya'ni jarayonlarni raqamlar orqali ifodalash shaklidir.

Pedagogika va psixologiyada shkalalash o'rganilayotgan hodisani matematik tahlil qilishning eng muhim vositalaridan biri bo'lib, kuzatish, so'rovnomalar, tajribalar va test sinovlari orqali olingan empirik ma'lumotlarni tartibga solish usuli hisoblanadi. Shkalalashning umumiy jarayoni esa ma'lum qoidalarga asoslangan holda shkalani yaratishdan iborat [1].

II. Shkalalar va ularning turlari

Bugungi kunda ijtimoiy va psixologik sohalarda shkalalashning turli shakllaridan foydalaniladi va har bir shkala o'z vazifasiga ko'ra turli maqsadlarda foydalaniladi. Bu shkalalar quyidagilardan iborat bo'lib: 1) nominal, 2) tartib, 3) interval va 4) nisbat shkalalardir. Ushbu shkalalarga tegishli kattaliklarning xususiy qiymatlari sonlar to'plamidagi xususiyatlarini qanchalik saqlab qolish darajasi bilan farqlanadi.

Nominal shkala – o'lchovning eng oddiy turi bo'lib, obyektlarni yoki hodisalarni kategoriya va guruhlarga ajratish uchun ishlatiladi. Bu shkala faqat identifikatsiya va tasniflash maqsadida qo'llaniladi. Undagi raqamlar son sifatida emas, balki belgilar sifatida tushuniladi. Nominal shkalaning o'ziga xos xususiyatlari: tartib yo'q – guruhlar, turlar, turkumlar yoki sinf orasida ustunlik yoki tartib mavjud emas.

1-Jadval

Fanlarning tasniflanishi (klassifikatsiyasi)

№	Fanlar		
	Aniq fanlar	Tabiiy fanlar	Ijtimoiy fanlar
1	Fizika	Kimyo	Tarix
2	Matematika	Biologiya	Falsafa
3	Informatika	Geografiya	Huquqshunoslik
4	Astronomiya	Ekologiya	Iqtisodiyot
5			Sotsiologiya

Matematik amallar bajarilmaydi – o'rtacha qiymat yoki yig'indini hisoblab bo'lmaydi. Faqat nomlash va ajratish uchun foydalaniladi [2,3]. Misol uchun: 1-jadvaldagi fanlarning tasniflari - aniq, tabiiy va ijtimoiy [2,4]. Bu jadvalda keltirilgan yorliqlar

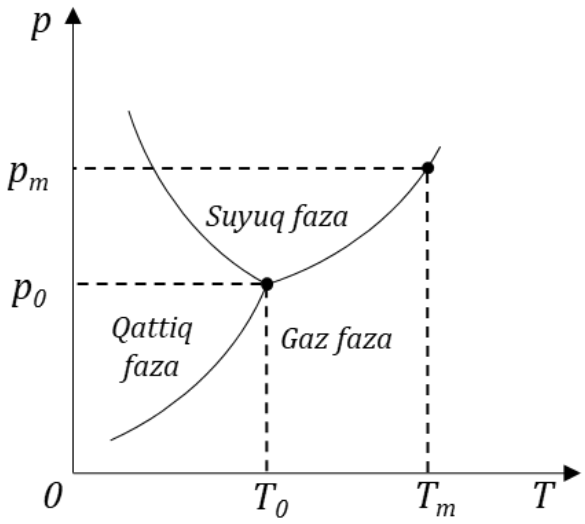
(nomlar) faqat ma’lumot sifatida qabul qilinadi.

Tartib shkalasi – bu ma’lumotlarni tartibga solish va qiyoslash imkonini beruvchi o‘lchov shkalasi bo‘lib, unda toifalar yoki darajalar mavjud. Lekin bu shkalada oraliqlari aniq emas, ya’ni ikki daraja orasidagi farq qanchalik katta ekanligini aniq hisoblab bo‘lmasligi mumkin. Tartib shkalasida toifalar yoki darajalar o‘zaro joylashuv tartibiga ega, bu shkalada toifalar yoki darajalarni qo‘shish va ayirish mumkin emas, lekin tartib va taqqoslash mumkin [4,5]. Masalan, o‘quvchi-larning bilimini baholash tizimida “a’lo”, “yaxshi”, “o‘rta”, “qoniqarli” va “qoniqarsiz” kabi darajalarga yoki sport musobaqalarida g‘oliblarni 1-o‘rin (Oltin medal), 2-o‘rin (Kumush medal), 3-o‘rin (Bronza medal) kabi darajalarga ajratish mumkin.

Interval shkalasi – o‘lchov tizimlaridan biri bo‘lib, unda ob’ektlar yoki hodisalar orasidagi farqlar izchil va teng masofalarda o‘lchanadi. Ushbu shkalada “nol” nuqtasi mavjud bo‘lib, bu nuqta mutlaq emas, ya’ni shartli belgilangan nuqta hisoblanadi. Interval shkalasi tartib shkalasining mukammal shakli bo‘lib, unda nafaqat obyekt-larning ketma-ketligi, balki ular orasidagi farqlar ham aniq o‘lchanadi. Ushbu shkalada kattaliklarning qiymatlarini qo‘shish va ayirish mumkin, ammo ko‘paytirish va bo‘lish amallarini qo‘llash mumkin emas, chunki “nol” mutlaq nuqta emasligi tufayli proporsional hisob-kitoblar ma’no anglatmaydi [2]. Misol tariqasida jism haroratni o‘lchashda foydalanadigan Selsiy shkalasini olaylik. Bu shkalaning eng quyi nuqtasi t = −273,15 °C ni tashkil etadi. Bu tabiatda eng past harorat chegarasi bo‘lib, bu nuqtada moddadagi barcha atom va molekulalar harakati to‘xtaydi**,** ya’ni ularning kinetik energiyasi nolga teng bo‘ladi [5**,**6]**.** Ammo bu shkalaning “nol” nuqtasi ham mavjud bo‘lib bu nuqtada tabiiy jarayonlarning va hodisalarning mavjud emasligini anglatmaydi. Selsiy shkalasining “nol” nuqtasi sifatida shartli ravishda suvning uch nuqtasi tanlangan. Bu haroratda suv bir vaqtning o‘zida va muvozanatda uchta fazada - qattiq, suyuq va gaz holatlarida bo‘lishi mumkin (1-rasm).

Suvning qaynash nuqtasidagi harorat esa shartli ravishda 100 °C deb belgilangan va shu harorat oralig‘i 100 qismga bo‘lingan, yuqori haroratli jarayonlarda ularning haroratini aniq-lash uchun aynan suvning uch nuqtasiga nisbatan belgilab olinadi. Selsiy shkalasida bir jismning harorati 10 °C, ikkinchi jismning harorati 20 °C bo‘lsa, u holda bu ikki jismning harorat nisbatlari 2 barobarni tashkil qilmaydi, ya’ni jismlardagi mavjud issiqlik miqdorlari nisbati 2 ga teng bo‘lmaydi. Shuning uchun bu shkalada nisbiy “nol” nuqtaning mavjudligi matematik hisob kitoblarda noqulaylik keltirib chiqa-

radi. Shu sababli fizik jarayonlarning matematik hisob-kitoblarida, ilmiy tadqiqot ishlarida Selsiy shkalasidan Kelvin shkalasiga $T(K) = t\,(°C) + 273{,}15$ ifoda orqali o‘tiladi. Bu esa o‘z navbatida hisob - kitoblarni qulaylashtiradi.



1-rasm. Suvning fazaviy diagrammasi, T_0– Suvning uch fazali nuqtasi, T_m–suvning qaynash harorati

Nisbat shkalasi – o‘lchov tizimidagi eng aniq va mukammal shkala bo‘lib, u nol nuqtasi mutlaq bo‘lgan, tartib, interval va nisbat xususiyatlariga ega bo‘lgan shkala hisoblanadi. Ushbu shkalaning o‘ziga xos xususiyati mutloq nol qiymatda o‘lchanayotgan kattalikning yo‘qligini bildiradi, qiymatlar tartiblanadi, qiymatlar orasidagi masofa (interval) aniq va o‘zgarmas qilib tanlanadi**,** boshqa shkalalardan farqli o‘laroq, bu shkalada nisbatlarni taqqoslash imkoniyati mavjud, bu shkalada barcha matematik amallar bajarilishi mumkin (2-jadval). Bu shkala fizika, matematika, iqtisodiyot, ijtimoiy fanlar va statistika kabi ko‘plab sohalarda keng qo‘llaniladi**.** Misol tariqasida uzunlik, og‘irlik, vaqt yoki daromad, pul miqdori kabi ko‘plab o‘lchov tizimlarini keltirish mumkin [7].

Yuqorida ko‘rsatib o‘tilgan shkalalash jarayonlarini pedagogika va psixologiya sohalariga ham tadbiq etish, o‘rganilayotgan ob’yektlarning xususiyatlarini tizimlashtirish, ular haqida to‘liq xulosalar chiqarishga asos bo‘ladi. Jumladan pedagogik o‘lchovlar sohasida sinaluvchilarning bilim, malaka va ko‘nikmalarini o‘lchashda ularning xom ballarini haqiqiy ballga aylantirish pedagogik test natijalarining adekvat talqini va taqqoslashini ta’minlaydi.

Pedagogik o‘lchovlar sohasida sinaluvchilar natijalariga beriladigan dastlabki ballar ularning xom ballari deyiladi.

Xom (boshlang‘ich) ball - bu test yoki baholash natijasida test topshiruvchi olgan dastlabki (o‘zgartirilmagan) balldir. Xom ball o‘quvchining nechta savolga to‘g‘ri javob berganligini yoki umumiy to‘plagan ballni ifodalaydi [8-10]. Xom ballarni turli usullar bilan hisoblash mumkin - dixotomik baholash, politomik baholash va qiyinlik

darajalari orqali [11-13]. Ammo bu ballar nisbiy o‘lchov bo‘lib, u turli testlarda yoki turli sinov sharoitlarida bir xil ahamiyatga ega bo‘lmasligi mumkin. Xom ball o‘z ichiga subyektiv va obyektiv ta’sirlarni olishi mumkin (masalan, testning qiyinligi, vaqt cheklovi, test topshiruvchining jismoniy va ruhiy holati).

2-jadval

Shkalalarning bir-biridan farqlari

Xususiyatlar	Nominal shkala	Tartib shkalasi	Interval shkala	Nisbat shkala
Ma’lumotni tasniflash	✓	✓	✓	✓
Tartibni aniqlash	✗	✓	✓	✓
Qiymatlar orasidagi farqlarni hisoblash	✗	✗	✓	✓
Mutlaq nol mavjudligi	✗	✗	✗	✓
Matematik amallar	✗	✗	✗	✓

Demak, xom ball – test natijalari bo‘yicha test topshiruvchining dastlabki bahosi bo‘lib, u me’yorlashtirish va qayta ishlash jarayonlari orqali yanada ma’noli shaklga keltiriladi. Test natijalarini to‘g‘ri tahlil qilish va taqqoslash uchun xom balldan keyingi ishlov berish jarayonlari sinaluvchilar haqida chiqariladigan ilmiy xulosalar uchun muhim ahamiyatga ega [14-16].

Pedagogik o‘lchovlarda sinaluvchilarning test natijalarini baholashda eng keng tarqalgan va samarali usullardan biri - **z-shkala**sidir. Bu o‘rtacha qiymat va dispersiyadan foydalangan holda test natijalarini normallashtirish uchun ishlatiladigan me’yorlashtirilgan o‘lchov tizimidir. Ushbu shkala yordamida turli test variantlaridan olingan natijalarni taqqoslash, sinaluvchilarning natijalarini umumiy taqsimotga nisbatan joylashishini aniqlash va baholash imkoniyati yaratiladi [17,18]. Har qanday test natijalarini z-shkalasiga o‘tkazish uchun quyidagi ifodadan foydalaniladi:

$$z = \frac{X - \overline{X}}{s_x} \qquad (1)$$

bu yerda, X – sinaluvchining test natijasi asosida hisoblangan xom balli, $\overline{\mathbf{X}}$ – test natijalari yig‘indisining o‘rta arfmetik qiymati:

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \qquad (2)$$

1-formuladagi s_x – test natijalarining standart og‘ishi bo‘lib o‘rtacha qiymatdan individual natijalarning qanchalik farq qilishini ifodalaydi :

$$\boldsymbol{s_x = \sqrt{\frac{S}{N}}} \qquad (3)$$

bu yerda, $\boldsymbol{S = \sum_{i=1}^{N} X^2 - \frac{\left(\sum_{i=1}^{N} X_i\right)^2}{N}}$ ga teng bo‘lib, test topshiriqlar dispersiyasi deyiladi. *z*-shkalasi bo‘yicha hisoblanganda sinaluvchilar ballarining o‘rtacha qiymati “0” ga teng bo‘ladi. Agar natijalar *z*-shkalasida musbat qiymatlarini qabul qilsa, u holda sinaluvchining natijasi guruhning o‘rtacha natijasidan yuqoriroq ekanligini ko‘rsatadi. Aksincha, manfiy qiymatlarini qabul qilsa, guruhning o‘rtacha natijasidan pastroq ekanligini bildiradi.

Ammo z-ball shkalasida manfiy qiymatlar navjudligi mutaxassis bo‘lmaganlar uchun murakkab shkaladir, sinaluvchilar orasida o‘rtacha ball olgan ishtirokchilarning ballari “0”ga teng bo‘ladi, bu esa o‘z navbatida sinaluvchida tushunmov-chilikni keltirib chiqaradi. Shuni hisobga olib xom ballni standart ball shkalada aks ettirishning boshqa bir yo‘li T- balli shkalaga o‘tishdir.

***T*-shkala** – bu test natijalarini standartlashtirish va ularni tushu-narliroq formatga o‘tkazish uchun ishlatiladigan shkaladir. U z-shkala asosida quyidagi ifoda orqali hisoblanadi:

$$\boldsymbol{T = 50 + 10Z} \qquad (4)$$

Bu yerda z – z-shkala bo‘yicha hisoblangan natija, 50 – *T*-shkaladagi o‘rtacha qiymat, 10 – *T*-shkaladagi standart og‘ish. Ushbu shkalaning manfiy qiymatlar mavjud emas, natijalarni 0-100 oralig‘ida baholanadi, har xil variantdagi test natijalarini bir xil tizimda ifodalashga yordam beradi [18,19].

Sinaluvchilarning individual natijalarini guruh ichidagi joylashuviga qarab baholash tizimi ham mavjud bo‘lib bu persentil shkalasi deb nomlanadi. Bu usul test yoki o‘lchov natijalarining nisbiy taqsimotini ko‘rsatish uchun ishlatiladi [20].

P=$\boldsymbol{\frac{R}{N} \cdot 100\%}$, bu yerda R-sinaluvchining nechanchi o‘rin (Reyting)da turishini ko‘rsatuvchi kattalik, persentil shkalasi mutlaq natijani emas, balki shaxsning boshqalar bilan solishtirgandagi o‘rnini ifodalaydi. Bu sinaluvchining test natijalarini taqqoslashda yoki guruh ichida qanday taqsimlanganini aniqlashda foydalaniladi. Ba’zan noto‘g‘ri talqin qilinishi ham mumkin (masalan, natijalarning normal taqsimotga ega bo‘lmasligi).

III. Natijalar va ularning muhokamasi

Yuqorida keltirilgan shkalalar asosida sinaluvchilarning xom ballari qanday standart ballga o‘zgarishini ko‘rib chiqaylik.

2024-yilning 21-22-dekabr kunlari Bilim va malakalarni baholash agentligi tomonidan ona tili va adabiyot fanidan milliy sertifikat uchun test sinovlari o‘tkazildi. Test sinovlarida ishtirokchilarga 44 ta (40-44-ochiq test topshiriqlarining A va B qismlarga ajratilishi hisobiga 49 ta) va 1 ta yozma ish (esse) dan iborat test topshiriqlar to‘plami berildi. 3-jadvalda test sinovlarida ishtirok etgan sinaluvchilardan ixtiyoriy tanlab olingan 39 ta ishtirokchining natijalari keltirilgan. Bu jadvalda sinaluvchilarning yopiq test topshiriqlarining natijalari asosida hisoblangan xom ballarining (X) monoton kamayish tartibida tanlab olingan. Jadvalning keyingi ustunida sinaluvchilarning yozma ish (esse) natijalariga fan ekspertlari tomonidan qo‘yilgan dastlabki xom ballari-(X_{in}) keltirilgan, T_{in}-yozma ish (esse) ballarining T-ball shkalasidagi qiymatlari.

3-Jadval

Yopiq test va yozma ish natijalarining xom ballari va bu ballarning z va T shkaladagi qiymatlari

№	X	X_{in}	T_{in}	z-ball	T-ball
1	46	18,25	63,5	3,425	84,26
2	45	19,75	66,5	3,072	80,72
3	44	19,5	66	2,747	77,47
4	43	15,75	58,5	2,452	74,52
5	42	19,25	65,5	2,247	72,48
6	41	20,25	67,5	2,121	71,22
7	40	14,75	56,5	2,008	70,08
8	39	20,75	68,5	1,839	68,39
9	38	17	61	1,592	65,92
10	37	17,5	62	1,341	63,42
11	36	17,75	62,5	1,174	61,74
12	35	19,5	66	1,089	60,9
13	34	20,75	68,5	1,048	60,49
14	33	17,75	62,5	1,017	60,17
15	32	17,75	62,5	0,966	59,66
16	31	16,75	60,5	0,858	58,58
17	30	15,5	58	0,659	56,6

18	29	16,5	60	0,4007	54,01
19	28	17,25	61,5	0,1849	51,85
20	27	15	57	0,0621	50,62
21	26	18,75	64,5	0,00615	50,06
22	25	17,75	62,5	-0,019	49,81
23	24	14,75	56,5	-0,036	49,64
24	23	12,25	51,5	-0,061	49,38
25	22	10,75	48,5	-0,119	48,81
26	21	10,5	48	-0,245	47,55
27	20	15,25	57,5	-0,463	45,36
28	19	9,75	46,5	-0,721	42,78
29	18	11	49	-0,917	40,82
30	17	12	51	-1,022	39,77
31	16	8,5	44	-1,072	39,28
32	15	14	55	-1,102	38,98
33	14	20,5	68	-1,141	38,59
34	13	9,5	46	-1,195	38,04
35	12	12	51	-1,383	36,16
36	11	11,75	50,5	-1,631	33,69
37	10	10,25	47,5	-1,881	31,19
38	9	11	49	-2,054	29,45
39	8	8,25	43,5	-2,169	28,31

Yozma ishga fan ekspertlari tomonidan qo‘yilgan xom ballar asosan 10 va 20 ball oralig‘ida tebranayotganligini ko‘rish mumkin, ammo yopiq test natijalari esa uzluksiz kamayib bormoqda. 33-o‘rindagi sinaluvchining yopiq test natijasida olgan xom balli 14 ga teng bo‘lsa-da, yozma ishga qo‘yilgan xom ball 20,5 ga teng yoki 1-o‘rindagi sinaluvchining yopiq test natijasida olgan xom balli 46 ga, yozma ishga quyilgan xom balli esa 18,25 ga teng, shu kabi holatlar boshqa sinaluvchilarda ham uchraydi. Yozma ishlarni baholashda asosiy muammolardan biri – baholovchining subyektivligi. Bu esa o‘z navbatida shaxsiy qarashlarga, yozma ish mazmunini turlicha talqin qilishiga va emotsional holatga bog‘liq bo‘lishi mumkin. Ba’zi mavzularni tushunish va ularga munosabat bildirish osonroq bo‘lsa, ba’zilarida sinaluvchilar qiynalishi mumkin. Ayrim fan ekspertlari qat’iy baholasa, boshqalar yumshoqroq baholaydi, bu esa umumiy natijalarga ta’sir qiladi. Shuning uchun yozma ishga qo‘yilgan xom ballar

to‘g‘ridan-to‘g‘ri natija sifatida qabul qilinishi unchalik to‘g‘ri bo‘lmaydi. Sinaluvchilarning yopiq test natijalariga qo‘yilgan xom ballari bilan yozma ishga qo‘yilgan xom ballarini arifmetik qo‘shish orqali umumiy ballni chiqarish adolatlilk tamoyillarining buzilishiga olib keladi. Sinaluvchilarning standart ballarini hisoblash uchun esa o‘zaro muvofiqlashtirish talab etiladi. 3-jadvalda sinaluvchilarning yopiq test natijalari va yozma ish(esse) natijalarining *z* va *T* shkaladagi ballari keltirilgan, bu ballar asosida sinaluvchilarning umumiy balli hisoblangan.

Xulosa

Test va esse ballarini o‘zaro muvofiqlashtirib standart ballar (z-ball, T-ball, persentil ballar va boshqalar)ga aylantirish bir qator muammolarni hal qiladi va baholash tizimining adolatli va aniq bo‘lishini ta’minlaydi. Standartlashtirish barcha ballarni bitta umumiy tizimga keltirib, ularni solishtirishni osonlashtiradi va subyektivlikdan xalos bo‘lib baholashning ob’ektivligini oshiradi (3-jadval *z*-ball va *T*-ball ustunlari). Bir nechta test natijalarining bir xil shkalaga o‘tkazish ularning o‘rtacha qiymatlarini, test savollarining qiyinlik darajalari kabi kattaliklarini va normal taqsimotga mosligini to‘g‘ri taqqoslash imkonini beradi. Test va yozma ish natijalarini standartlashtirish orqali o‘rtacha ball, mediana, ballarning taqsimot darajasi va standart og‘ish kabi statistik ko‘rsatkichlarni olish mumkin. Bu esa ta’lim sifati bo‘yicha yaxshiroq tahlil olib borish va natijalar asosida qarorlar qabul qilishga yordam beradi.

Q.A. Amonov va A.A. Baratov Bilim va malakalarni baholash agentligi huzuridagi Ilmiy – o‘quv amaliy markazi, Pedagogik o‘lchovlarni rivojlantirish ilmiy tadqiqot bo‘limi boshlig‘i M.Dj. Ermamatovga ilmiy maslahatlari va ko‘rsatmalari uchun minnatdorlik bildiradi.

ADABIYOTLAR

1. Стивенс С.С. Экспериментальная психология. Т. 1. //М.: Изд-во иностр. литературы, 1960. С. 19–89
2. Ильясов Ф. Н. Шкалы и специфика социологического измерения //Мониторинг общественного мнения: экономические и социальные перемены. – 2014. – №. 1 (119). – С. 3-16.
3. Сафонцева Н. Ю., Марочкина А. С. Номинальная шкала дидактического обеспеченияучебного процесса //Вестник непрерывного образования. – 2009. – №. 5. – С. 59-63.
4. https://vawilon.ru/statistika-grupp-krovi/
5. Atkins P. W., De Paula J., Keeler J. Atkins' physical chemistry. – Oxford university press, 2023.
6. Callen H. B. Thermodynamics and an Introduction to Thermostatistics. – John wiley & sons, 1991.
7. Зайцев С. А. и др. Контрольно-измерительные приборы и инструменты //М.: Академия. – 2002., Спиридонов О. П. Фундаментальные физические постоянные. – Высшая школа, 1991.
8. T.G. Bond and C.M. Fox, Applying the Rasch Model: Fundamental Measurement in the Human Sciences, 2nd ed. (Lawrence Erlbaum Associates, Mahwah, NJ, 2007).
9. B. D. Wright and M. H. Stone, Best Test Design (MESA Press, Chicago, 1979)
10. 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” ilmiy-uslubiy jurnali. – 2023. №1,– C. 4-62.
11. Rasch G. (1960), Probabilistic models for some intelligence and attainment tests, Copenhagen, Danish Institute for Educational research.
12. Wright, B. D., & Stone, M. H. (1979). Best test design. Chicago, IL: Mesa Press.
13. Birnbaum, A. Some latent trait models and their use in inferring an examinee’s ability. In F.M. Lord M.R. Novick, statisticaltheories of mental test scores (chapter 17-20), Reading, M. A.: Addison-Wisley.
14. Ermamatov M.D., Boyxonov I.A. Umumiy elementli test variantlari: biologiya fanidan olingan test natijalari, // Axborotnoma, ilmiy-uslubiy jurnali. – 2023. – C. 32-50.
15. M.D. Ermamatov, M.D. Alimov, A.A. Sulaymonov, A.R. Sattiyev, Kalibrovkalangan test topshiriqlair: Sharq tillaridan o‘tkazilgan test sinovi natijalarining statistik tahlili, Axborotnoma 3-4-son, 16-83 b., 2022. 11.

16. M.D. Ermamatov, A. Abbosov, A.A. Baratov, Test topshiriqlarini kalibrovkalash va qobiliyatlarini tenglashtirish, Axborotnoma 3-4-son, 4- 16 b., 2022.
17. Henderi H., Wahyuningsih T., Rahwanto E. Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer // International Journal of Informatics and Information Systems. – 2021. – Т. 4. – №. 1. – С. 13-20.
18. Аванесов В. Шкалирование тестовых результатов //Педагогические измерения. – 2013. – №. 4. – С. 65-78.
19. K. Jalilov. Baholash nazariyasi asoslari (testologiya, pedagogik o‘lchovlar)// T.: Akademnashr.-2020.

IMPORTANCE OF SCALING IN ASSESSMENT

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. The article discusses the scale and its types, the conversion of raw scores of the examinees into Z-scores and T-scores, as well as the methods for converting raw scores of the tests conducted in native language and literature into z- and T-scores, i.e., bringing them to a common scale.

Keywords: Test items, standard deviation, difficulty level of test items, average score, types of scales, z-scale, T-scale, percentile scale.

TEST XARAKTERISTIKALARI: GEOGRAFIYA FANIDAN O‘TKAZILGAN TEST SINOVLARI

A.R. Sattiyev, M.Dj. Ermamatov

Bilim va malakalarni baholash agentligi huzuridagi
Ilmiy-o‘quv amaliy markazi, 100084, Toshkent sh., Bog‘ishamol k., 12

Qisqacha mazmuni. Mazkur maqolada 2024-yilda geografiya fani bo‘yicha milliy sertifikat uchun o‘tkazilgan test sinovlari natijalari klassik test nazariyasi va Rash modeli asosida tahlil qilingan. Test sinovi natijalarining tavsif statistikasi va har bitta test topshiriqlariga berilgan javoblarning umumiy ball bilan korrelyatsiyasi muhokama qilingan. Klassik test nazariyasi va Rash modeli bilan o‘tkazilgan uchta test sinovida ishlatilgan variantlardagi test topshiriqlarining qiyinlik darajasi tahlil qilingan. Rash modeli asosida aniqlangan qobiliyat va qiyinlik darajalaridan foydalanib, har bir variant uchun Rayt xaritalari olingan. Shuningdek, test ma’lumoti va xarakteristikasi chiziqlari hamda test variantlaridagi har bir test topshirig‘ining Rash modeli bilan moslik statistikalari o‘rganilgan.

Kalit so‘zlar: Test topshiriqlari, Kronbax alfa koeffitsiyenti, qiyinlik darajasi, korrelyatsiya koeffitsiyenti, Rash modeli, Rayt xaritasi, qobiliyat darajalari, Rash modeli bilan moslik

I. Kirish

“Umumta’lim fanlarini bilish darajasini baholashning milliy test tizimini joriy etish to‘g‘risida”gi qaroriga asosan Bilim va malakalarni baholash agentligi umumta’lim fanlari bo‘yicha milliy sertifikat uchun test sinovlarini tashkil etib kelmoqda.

Milliy sertifikat uchun o‘t-kaziladigan test sinovlari natijalari klassik test nazariyasi [1-2] hamda zamonaviy nazariyalaridan biri bo‘lgan Rash modeli asosida [3-5] tahlil qilinadi va test topshiriqlarining xususiyatlarini optimal holga keltirish uchun fan mutaxassislariga taklif va tavsiyalar berib boriladi [6-8].

Avvalgi maqolalarimizda [8-14] ham umumta’lim fanlaridan milliy sertifikat uchun o‘tkazilgan test sinovlari natijalari bo‘yicha testning klassik nazariyasi hamda Rash modeli bilan matematik-statistik tadqiqotlar olib borilgan va tegishli tavsiyalar, xulosalar taqdim qilingan.

Mazkur maqola ham umumta’lim fanlaridan biri bo‘lgan geografiya fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalari haqida

bo‘ladi. Ilmiy tadqiqot obyekti sifatida milliy sertifikat uchun geografiya fanidan 2024-yil may (1-test sinovi), oktabr (2-test sinovi) va dekabr (3-test sinovi) oylarida o‘tkazilgan test sinovi natijalaridan foydalanildi. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 1-test sinovida 1358 nafar, 2-test sinovida 245 nafar va 3-test sinovida 701 nafar talabgorlar ishtirok etdilar. Test sinovlarida foydalanilgan har bir variant 45 ta (36-45-ochiq test topshiriqlarining A va B qismlarga ajratilishi hisobiga 55 ta) test topshirig‘idan iborat bo‘lib, ajratilgan vaqt javoblar varaqasini bo‘yash bilan birgalikda 150 daqiqani tashkil qilgan.

Maqolada geografiya fanidan milliy sertifikat uchun o‘tkazilgan 3 ta test sinovi natijasi asosida klassik test nazariyasi hamda Rash modeli bilan aniqlangan test xarakteristikalari o‘rganib chiqilgan.

II. Test sinovlari natijalarining klassik test nazariyasi asosida tahlili

Pedagogik o‘lchovlarning nazariy asoslariga ko‘ra test topshiriqlari mazmuni ekspert tekshiruvidan o‘tkazilgandan keyingi bosqichda test sinovlari o‘tkaziladi va uning natijalari asosida testlarning xususiyatlari sifatini tashxislash uchun ularning statistik xarakteristikalari aniqlanadi. Testlar statistikada tanlanma to‘plam hisoblanadi va u real testlar to‘g‘risida xulosalar chiqarish imkonini beradi.

Test topshiriqlariga qo‘yilgan asosiy talablarning ko‘rsatkichlari statistik tavsiflar yordamida aniq-lanadi. Unga qo‘yiladigan asosiy talablar topshiriqning qiyinligi, test ballarining dispersiyasi (o‘zgaruvchan-ligi, farqlanishi), topshiriqning boshqa topshiriqlar bilan, shuningdek, umumiy ballar yig‘indisi bilan korrelyatsiyasidan iborat. Berilgan topshiriqning qiyinlik darajasini aniqlash usullaridan biri uni empirik sinovdan o‘tkazib, to‘g‘ri javoblar salmog‘ini aniqlashdan iborat. Test ballarining (yoki to‘g‘ri javoblarning) dispersiyasi test topshiruvchilarning tayyorgarlik darajasini aniqlashga, bilim darajalari bo‘yicha ajratishga imkon beradi.

Test variantlari va test topshiriqlarining asosiy statistik tavsiflari qatoriga o‘rta qiymat, gistogrammani qurish, moda va mediana kabi ko‘rsatkichlarni hisoblash hamda test ballarining umumiy dispersiyasi (standart og‘ish) ham kiradi [1-2,15-18]. Test ballarining (yoki to‘g‘ri javoblarning) o‘rta arifmetik qiymati fanlar, oliy ta’lim muassasalari va boshqa muhim belgilar kesimida aniqlanadi. Bu ko‘rsatkich test ballari o‘rtasidagi tafovutlarni umumlashtiradi, ularga xos bo‘lgan yo‘nalishni, qonuniyatni ochib beradi. Test sinovlari natijalari asosida aniqlangan test ballari taqsimoti gistogrammasi quriladi va

uning normal taqsimotga yaqinligi baholanadi. Gistogrammaning normal taqsimotga yaqinligi testning sifatini, test sinovlarining obyektiv o‘tkazilganligini bildiradi. 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 biri-biridan qanchalik ko‘p farq qilsa, ballar taqsimoti normal taqsimotdan shunchalik uzoqda bo‘ladi.

1-jadvalda o‘tkazilgan test sinovlari natijalarining tavsif statistikasi ma’lumotlari keltirilgan.

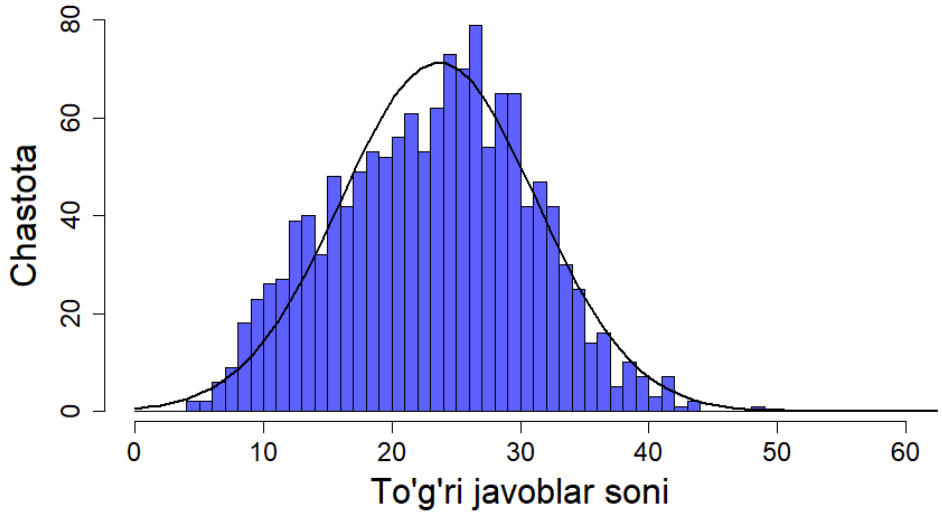
1-jadval

Natijalarining tavsif statistikasi

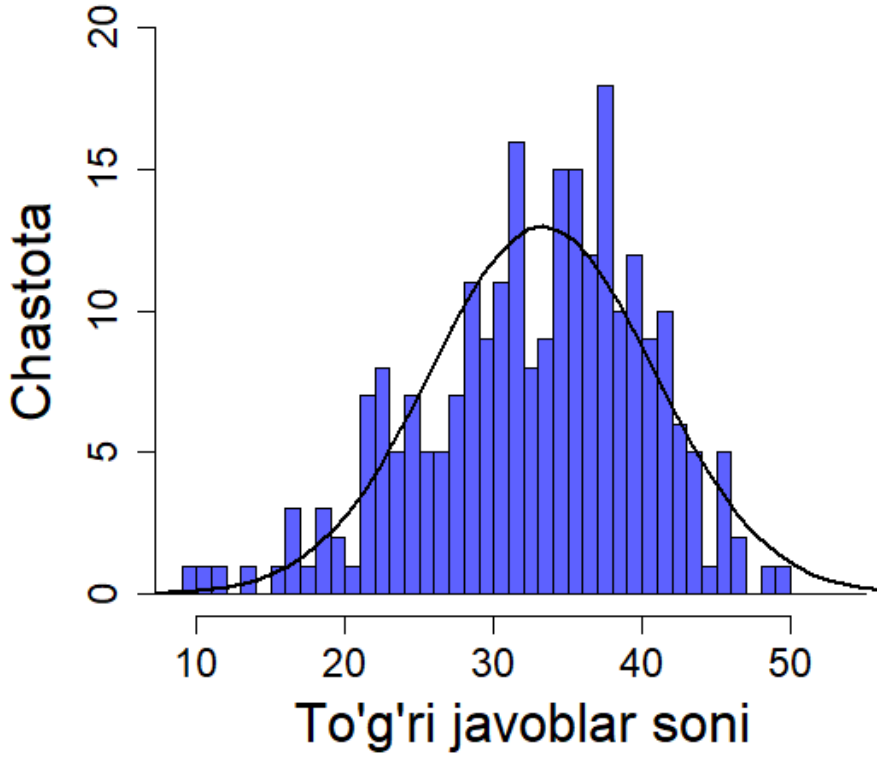
№	Test sinovlari	O‘rta qiymat	Mediana	Moda	Dispersiya	Kronbax alfa koeffitsiyenti	O‘lchashning standart xatoligi
1.	1-test sinovi	23,63	24	27	54,64	0,84	3,04
2.	2-test sinovi	33,28	34	38	56,63	0,84	3,01
3.	3-test sinovi	33,17	35	38	76,39	0,87	3,15

1-jadvalda geografiya fanidan test sinovlari natijalarining tavsif statistikalari, ya’ni test sinovi ballarining o‘rta qiymati, medianasi, modasi va dispersiyasi qiymatlari keltirilgan. 1*a*-, 1*b*- va 1*c*-rasmlarda esa ularning gistogrammalari keltirilgan. Gistogrammalardan ko‘rinib turibdiki, o‘tkazilgan 3 ta test sinovi natijalari bo‘yicha test topshiriqlarining individual ballari taqsimoti bir-biriga juda yaqin va normal taqsimotdan juda kam farq qiladi. Statistik tadqiqot natijalariga ko‘ra, geografiya fanidan test variantining ishonchlilik koeffitsiyenti, ya’ni Kronbax alfa koeffitsiyenti 1-, 2- va 3-test sinovi natijalari bo‘yicha mos ravishda 0,84; 0,84 va 0,87 ga teng ekanligi aniqlandi. Kronbax alfa koeffitsiyentining 0,8 va undan kattaligi ushbu test sinovlari uchun tanlab olingan test variantlarining ishonchliligi yaxshi darajada ekanligini ko‘rsatmoqda [19]. Test topshiriqlarining ichki muvofiqligi har bitta test topshirig‘iga berilgan to‘g‘ri javoblarning umumiy ball bilan korrelyatsiyasiga, talabgorlar olgan umumiy ballarning standart og‘ishiga, har bitta test topshirig‘iga berilgan javoblarining standart og‘ishlari yig‘indisiga hamda test topshiriqlari va

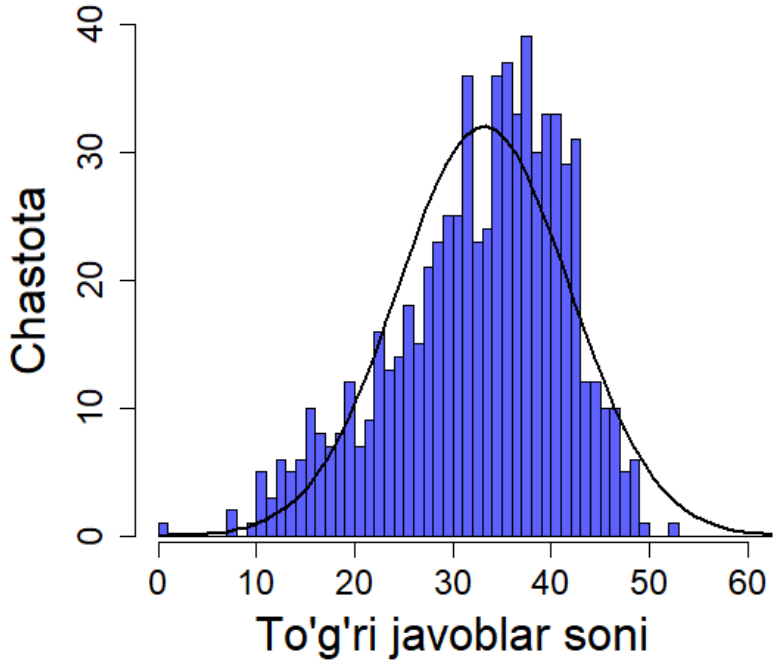
test topshiruvchilar soniga bog‘liq bo‘ladi. Bundan tashqari test topshiriqlarining ichki muvofiqligi nafaqat test topshiriqlarining sifatiga, balki talabgorlarning tayyorgarlik darajasining past yoki yuqoriligiga ham bog‘liqdir.



1*a*-rasm. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 1-test sinovi natijalarining gistogrammasi



1*b*-rasm. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 2-test sinovi natijalarining gistogrammasi



1*c*-rasm. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 3-test sinovi natijalarining gistogrammasi

Klassik test nazariyasi asosida geografiya fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalari bo‘yicha test topshiriqlarining qiyinlik darajalari aniqlandi. Quyida test topshiriqlari natijalari asosida, test topshiriqlarining - ID raqami, to‘g‘ri javob berganlar foizi – Ans (foizda) va test natijalari bo‘yicha aniqlangan qiyinlik darajalari V (1-, 2- va 3-qiyinlik darajalari) 2-jadvalda ko‘rsatilgan.

2-jadval

Geografiya fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalari bo‘yicha test topshiriqlarining aniqlangan qiyinlik darajalari

№	1-test sinovi			2-test sinovi			3-test sinovi		
	ID	**Ans *(foizda)***	**V**	***ID***	**Ans *(foizda)***	**V**	***ID***	**Ans *(foizda)***	**V**
1.	115548	96,24	1	114169	95,10	1	1073812	92,15	1
2.	115765	85,35	1	107211	91,84	1	107318	91,44	1
3.	115555	84,98	1	107304	91,84	1	113866	87,02	1
4.	115781	80,41	1	113873	89,80	1	107286	86,88	1
5.	115755	78,72	1	107322	88,98	1	112568	86,16	1

6.	115770	75,48	1	112546	87,76	1	107331	85,88	1
7.	115763	69,73	2	112612	85,71	1	114196	83,88	1
8.	115762	69,51	2	107344	84,49	1	112607	82,17	1
9.	115767	69,15	2	112539	84,08	1	114179	80,88	1
10.	115771	68,04	2	113869	83,27	1	112612	79,89	1
11.	115782	67,30	2	107235	80,41	1	107333	77,75	1
12.	115776	66,94	2	113832	79,59	1	107381	77,75	1
13.	115768	66,64	2	107406	79,18	1	107235	77,46	1
14.	1155552	61,27	2	104179	77,14	1	107212	77,32	1
15.	115775	60,97	2	11447433	76,73	1	114115	76,46	1
16.	115783	60,82	2	114099	76,73	1	114193	73,32	2
17.	115772	60,24	2	114126	76,33	1	114191	72,61	2
18.	115754	57,51	2	107391	75,10	1	113848	72,04	2
19.	115759	57,29	2	107331	74,69	2	112616	71,61	2
20.	115780	57,00	2	104183	73,88	2	112592	71,18	2
21.	115756	54,86	2	112562	72,65	2	107190	69,90	2
22.	115758	54,71	2	113847	72,24	2	107230	68,90	2
23.	115777	51,91	2	114474	72,24	2	113851	66,33	2
24.	115550	48,45	2	113861	71,43	2	114125	66,33	2
25.	115786	46,91	2	114116	70,61	2	107339	66,19	2
26.	11575435	46,91	2	112580	68,98	2	107206	63,91	2
27.	115773	44,92	2	104190	66,94	2	107246	60,63	2
28.	115764	44,18	2	112592	66,94	2	112547	59,91	2
29.	115761	41,46	2	107292	66,12	2	107270	57,77	2
30.	115757	39,03	2	107204	64,90	2	1141252	57,63	2
31.	115785	37,19	2	107230	64,49	2	1141712	57,49	2
32.	1155502	35,27	2	114091	63,67	2	1141152	56,06	2
33.	115784	35,05	2	1073912	61,63	2	114132	55,21	2
34.	115769	34,54	2	112600	61,22	2	107376	54,35	2
35.	115779	32,47	2	1141162	59,18	2	107239	53,07	2
36.	115760	31,37	2	114190	55,92	2	1140832	52,64	2
37.	115766	28,94	2	113952	54,69	2	113949	51,36	2
38.	115546	27,39	2	114110	54,29	2	10740935	50,64	2

39.	115551	26,22	2	107341	50,61	2	114171	50,36	2
40.	11575434	24,45	3	114122	49,80	2	113957	48,07	2
41.	115547	21,72	3	1141222	48,16	2	1073762	47,93	2
42.	115554	20,47	3	1141102	46,94	2	107314	46,79	2
43.	115774	20,32	3	11447434	44,49	2	107402	46,79	2
44.	1155482	19,07	3	1141262	40,41	2	107250	45,36	2
45.	1155522	16,13	3	107311	36,73	2	1073672	43,51	2
46.	115778	14,65	3	113955	35,10	2	112560	41,80	2
47.	115552	12,37	3	107316	33,06	2	111033	40,37	2
48.	1155492	11,41	3	112586	31,84	2	107397	37,52	2
49.	1155542	11,27	3	114097	29,39	2	114083	34,95	2
50.	115549	9,13	3	1140992	24,49	3	107367	34,81	2
51.	1155512	7,88	3	1140972	17,14	3	1074022	34,38	2
52.	1155462	6,77	3	1074062	10,20	3	1141322	33,95	2
53.	1155472	4,79	3	1140912	4,08	3	107409	26,39	2
54.	1155532	4,12	3	1141072	2,86	3	10740934	24,68	3
55.	115553	2,80	3	114107	2,04	3	1073972	6,70	3

1-test sinovi natijalari bo‘yicha (2-jadval), 55 ta test topshirig‘idan 6 tasi (10,91 foiz) 1-qiyinlik darajasidagi test topshirig‘idan, 33 tasi (60,00 foiz) 2-qiyinlik darajasidagi test topshirig‘idan va 16 tasi (29,09 foiz) 3-qiyinlik darajasidagi test topshirig‘idan iborat ekanligi aniqlandi.

1-test sinovida foydalanilgan variantda 1-qiyinlik darajasidagi test topshiriqlari normadan biroz kam, 2-qiyinlik darajasidagi test topshiriqlari normada taqsimlanganligi, 3-qiyinlik darajasidagi test topshiriqlari esa normadan ko‘p taqsimlanganligi kuzatildi. ID raqami - 115548 bo‘lgan test topshirig‘ining qiyinlik darajasi juda past va ID raqamlari - 1155532 va 115553 bo‘lgan test topshiriqlarining qiyinlik darajalari juda yuqori ekanligi ko‘rish mumkin.

2-test sinovi natijalari bo‘yicha esa (2-jadval), 55 ta test topshirig‘idan 18 tasi (32,73 foiz) 1-qiyinlik darajasidagi test topshirig‘idan, 31 tasi (56,36 foiz) 2-qiyinlik darajasidagi test topshirig‘idan va 6 tasi (10,91 foiz) 3-qiyinlik darajasidagi test topshirig‘idan iborat ekanligi aniqlandi.

2-test sinovida foydalanilgan variantda 1-qiyinlik darajasidagi test topshiriqlari normadan ko‘p taqsimlanganligi, 2-qiyinlik darajasidagi test topshiriqlari esa normada

taqsimlanganligi va 3-qiyinlik darajasidagi test topshiriqlari esa normadan biroz kam taqsimlanganligi kuzatildi. ID raqamilari - 114169, 107211 va 107304 bo‘lgan test topshiriqlarining qiyinlik darajasi juda past va ID raqamlari - 1140912, 1141072 va 114107 bo‘lgan test topshiriqlarining qiyinlik darajalari esa juda yuqori ekanligi aniqlandi.

3-test sinovi natijalari bo‘yicha esa (2-jadval), 55 ta test topshiriqlaridan 15 tasi (27,27 foiz) 1-qiyinlik darajasidagi test topshiriqlaridan, 38 tasi (69,10 foiz) 2-qiyinlik darajasidagi test topshiriqlaridan va 2 tasi (3,63 foiz) 3-qiyinlik darajasidagi test topshiriqlaridan iborat ekanligi aniqlandi.

3-test sinovida foydalanilgan variantda 1- va 2-qiyinlik darajasidagi test topshiriqlari normadan biroz ko‘p taqsimlanganligi va 3-qiyinlik darajasidagi test topshiriqlari normadan juda kam taqsimlanganligi kuzatildi. ID raqamlari- 1073812 va 107318 bo‘lgan test topshiriqlarining qiyinlik darajasi juda past va ID raqami - 1073972 bo‘lgan test topshirig‘ining qiyinlik darajasi esa juda yuqori ekanligi aniqlandi.

Klassik test nazariyasi bo‘yicha test topshiriqlarining normal taqsimotini ta’minlash uchun, 1- va 3-darajali test topshiriqlari sonini test variantidagi test topshiriqlari sonining 16-25 (9-14 ta) foizi, 2-darajali test topshiriqlari sonini esa 50-68 (27-37 ta) foiz qilib olish maqsadga muvofiq bo‘ladi.

Geografiya fanidan milliy sertifikat uchun o‘tkazilgan test sinovida test topshiriqlarining ichki muvofiqligi har bitta test topshirig‘iga berilgan to‘g‘ri javoblarning umumiy ball bilan korrelyatsiyasiga, talabgorlar olgan umumiy ballarning standart og‘ishiga, har bitta test topshirig‘iga berilgan javoblarining 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.

Har bitta test topshirig‘iga berilgan javoblarning umumiy test bali bilan korrelyatsiyasi shu test topshirig‘ining qobiliyatlarni qanchalik yaxshi ajaratishini bildiradi.

Umuman olganda, umumiy ball bilan korrelyatsiya koeffitsiyenti 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 talabgorlar g‘olib bo‘lib, bilim darajalari yuqori bo‘lgan talabgorlar test topshiriqlarini yechishda noto‘g‘ri

javobni tanlaydilar yoki ularni o‘tkazib yuboradilar. 3-jadvalda 1-, 2- va 3-test sinovi natijalari tahlili asosida olingan test topshiriqlarining umumiy ball bilan korrelyatsiya koeffitsiyenti (UBBKK) qiymatlari qiyinlik darajalari ortib borishi tartibida keltirilgan. Statistik tadqiqotlar natijalariga ko‘ra 1-test sinovida foydalanilgan test topshiriqlarining 14 tasining (3-jadvalning 1-test sinovi qismida ajratib ko‘rsatilgan ID raqamlari – 115548, 115776, 115758, 115784, 115769, 115779, 115760, 11575434, 115774, 1155522, 115778, 1155462, 1155472 va 115553) va 2-test sinovida foydalanilgan test topshiriqlarining 15 tasining (3-jadvalning 2-test sinovi qismida ajratib ko‘rsatilgan ID raqamlari – 114169, 107304, 113873, 113869, 114126, 114190, 107341, 11447434, 107316, 112586, 1140992, 1074062, 1140912, 1141072 va 114107) hamda 3-test sinovida foydalanilgan test topshiriqlarining 7 tasining (3-jadvalning 3-test sinovi qismida ajratib ko‘rsatilgan ID raqamlari – 107331, 107314, 107250, 112560, 107409, 10740934 va 1073972) umumiy ball bilan korrelyatsiya koeffitsiyenti qiymati 0,25 dan kichikligi aniqlandi. 3-jadvaldan aksariyat test topshiriqlarining qiyinlik darajasi yuqori yoki past ekanligini ko‘rish mumkin.

1-test sinovidagi ID raqamlari-115784, 115769, 115779, 115760 bo‘lgan test topshiriqlarinig qiyinlik darajasi taqsimotida o‘rtaga yaqinligidan uning umumiy ball bilan korrelyatsiyasi ancha kichikligini hisobga olish lozim.

3-jadval

Individual test topshiriqlariga berilgan javoblarining umumiy ball bilan korrelyatsiyalari

№	1-test sivoni		2-test sivoni		3-test sivoni	
	ID	UBBKK	ID	UBBKK	ID	UBBKK
1.	**115548**	**0,087**	**114169**	**0,225**	1073812	0,311
2.	115765	0,457	107211	0,352	107318	0,295
3.	115555	0,331	**107304**	**0,235**	113866	0,334
4.	115781	0,456	**113873**	**0,242**	107286	0,385
5.	115755	0,444	107322	0,432	112568	0,313
6.	115770	0,285	112546	0,356	**107331**	**0,233**
7.	115763	0,407	112612	0,358	114196	0,447
8.	115762	0,436	107344	0,290	112607	0,349

9.	115767	0,500	112539	0,359	114179	0,363
10.	115771	0,352	**113869**	**0,190**	112612	0,496
11.	115782	0,369	107235	0,314	107333	0,426
12.	**115776**	**0,246**	113832	0,265	107381	0,516
13.	115768	0,433	107406	0,508	107235	0,498
14.	1155552	0,408	104179	0,458	107212	0,444
15.	115775	0,384	11447433	0,413	114115	0,416
16.	115783	0,318	114099	0,322	114193	0,335
17.	115772	0,353	**114126**	**0,210**	114191	0,447
18.	115754	0,431	107391	0,398	113848	0,440
19.	115759	0,390	107331	0,309	112616	0,279
20.	115780	0,317	104183	0,412	112592	0,402
21.	115756	0,298	112562	0,314	107190	0,439
22.	**115758**	**0,144**	113847	0,265	107230	0,395
23.	115777	0,422	114474	0,457	113851	0,343
24.	115550	0,404	113861	0,471	114125	0,390
25.	115786	0,373	114116	0,322	107339	0,413
26.	11575435	0,433	112580	0,265	107206	0,299
27.	115773	0,341	104190	0,287	107246	0,250
28.	115764	0,363	112592	0,434	112547	0,284
29.	115761	0,429	107292	0,528	107270	0,335
30.	115757	0,293	107204	0,323	1141252	0,295
31.	115785	0,343	107230	0,336	1141712	0,452
32.	1155502	0,522	114091	0,423	1141152	0,407
33.	**115784**	**0,223**	1073912	0,434	114132	0,427
34.	**115769**	**0,203**	112600	0,291	107376	0,475
35.	**115779**	**0,148**	1141162	0,433	107239	0,259
36.	**115760**	**0,165**	**114190**	**0,192**	1140832	0,334
37.	115766	0,313	113952	0,444	113949	0,334
38.	115546	0,278	114110	0,414	10740935	0,289
39.	115551	0,351	**107341**	**0,194**	114171	0,404
40.	**11575434**	**0,181**	114122	0,391	113957	0,320
41.	115547	0,351	1141222	0,429	1073762	0,440

42.	115554	0,398	1141102	0,295	**107314**	**0,230**
43.	**115774**	**0,100**	**11447434**	**0,060**	107402	0,541
44.	1155482	0,325	1141262	0,344	**107250**	**0,237**
45.	**1155522**	**0,181**	107311	0,330	1073672	0,351
46.	**115778**	**0,096**	113955	0,319	**112560**	**0,073**
47.	115552	0,327	**107316**	**0,204**	111033	0,376
48.	1155492	0,317	**112586**	**0,229**	107397	0,444
49.	1155542	0,356	114097	0,387	114083	0,357
50.	115549	0,287	**1140992**	**0,182**	107367	0,427
51.	1155512	0,293	1140972	0,338	1074022	0,442
52.	**1155462**	**0,221**	**1074062**	**0,220**	1141322	0,251
53.	**1155472**	**0,230**	**1140912**	**0,192**	**107409**	**0,157**
54.	1155532	0,284	**1141072**	**0,143**	**10740934**	**0,118**
55.	**115553**	**0,225**	**114107**	**0,156**	**1073972**	**0,185**

III. Test sinovlari natijalarining Rash modeli asosida tahlili

Klassik test nazariyasining ham, zamonaviy test nazariyasining ham asosiy maqsadi test va so‘rovnomalar yordamida olingan natijalarni tahlil qilib, shaxslarning yashirin xususiyati (latent trait) bo‘yicha ma’lum bir shkaladagi o‘rnini aniqlashdir [20]. Mazkur maqolada ham zamonaviy test nazariyasining modellaridan biri bo‘lgan Rash modeli asosida tahlil natijalarini keltiramiz.

Test topshiriqlarini sifatini Rash modeli asosida matematik-statistik tadqiqoti hozirda keng tarqalgan usullardan biri bo‘lib, u dunyoning qator yetakchi mamlakatlari ta’lim tizimida samarali foydalanib kelinmoqda.

Rash modeli turli xildagi so‘rovnoma va testlar yordamida obyektiv o‘lchashlarni amalga oshirish tomon qilingan harakatlar tufayli Daniyalik olim Jorj Rash tomonidan yaratilgan. Bir o‘lchovlilikni ta’minlash mushkul bo‘lishiga qaramasdan, uni ta’minlash uchun oldindan tayyorgarlik ishlarinini amalga oshirish va bu ishlar qanchalik amalga oshirilganini empirik usullar bilan tekshirish imkoni mavjud. Chiziqli mavhum shkalaga esa Rash modelida logit birliklari orqali o‘tiladi. Rash modelining muhim xususiyati u shunchaki ma’lumotlarni tahlil qilish uchun statistik usul emas, balki u o‘lchovni nimaligini, ta’lim tizimida o‘lchovlarni qanday sifatili amalga oshirish imkoniyatini beradi [8].

Rash modelida [3-4] yashirin qobiliyat va elementlar qiyinligi kabi

parametrlarini aniqlash muhim o‘rin tutadi. Bu ikkita kattalikdan birinchisi o‘zgaruvchi sifatida, ikkinchisi esa parametr sifatida qaralishi mumkin. Test natijalarini tahlil qilishda elementlar qiyinlik darajasini parametr sifatida qarash qulay, chunki qobiliyat (bilim) bu modelda elementlarga berilgan javoblarga qarab belgilanadi.

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:

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

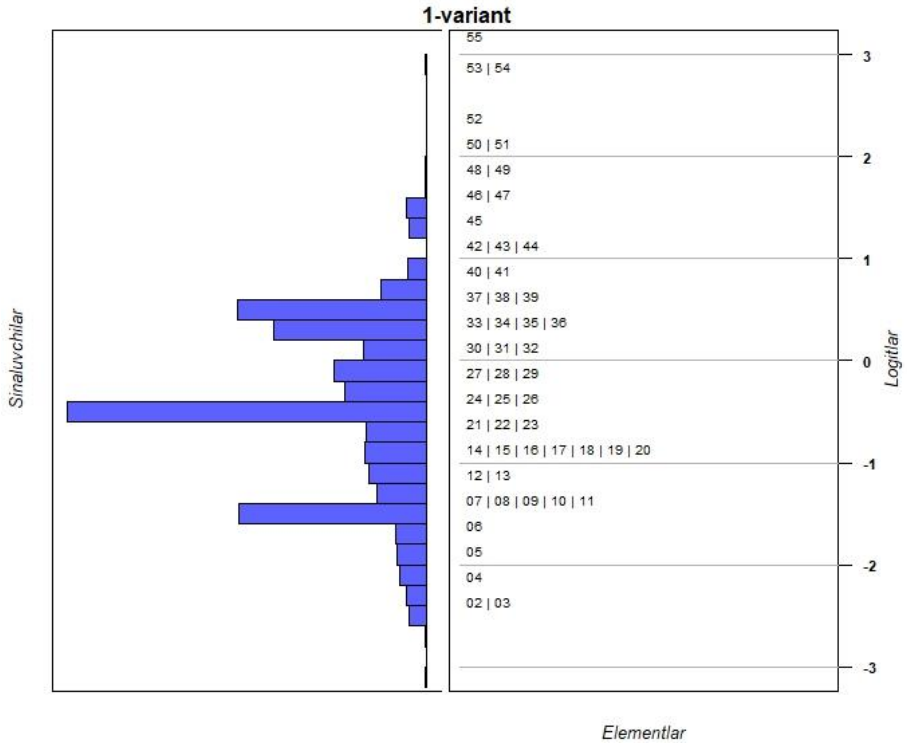
Bu yerda, $\boldsymbol{X_{is}}$ = s- talabgorning $\boldsymbol{i}$ – elementga to‘g‘ri javob berish ehtimolligi, $\boldsymbol{\theta_s}$– qobiliyat o‘zgaruvchisi, $\boldsymbol{b_i}$– topshiriq qiyinlik darajasi, $\boldsymbol{e}$– natural logarifm asosi (e=2,7182818...).

Geografiya fanidan test sinovlari natijalarining Rash modeli bo‘yicha tahlilini maxsus dastur asosida amalga oshirish uchun ishlab chiqilgan turli xil dasturiy paketlardan foydalanamiz. Masalan, qiyinlik darajasi $\boldsymbol{b}$ ni aniqlashda biz dexter dasturiy paketidan foydalanamiz [5], chunki Rash modeli uchun bu dasturiy paket yordamida tajribaning (test natijalari) modelga qanchalik mosligini hisoblash mumkin bo‘ladi.

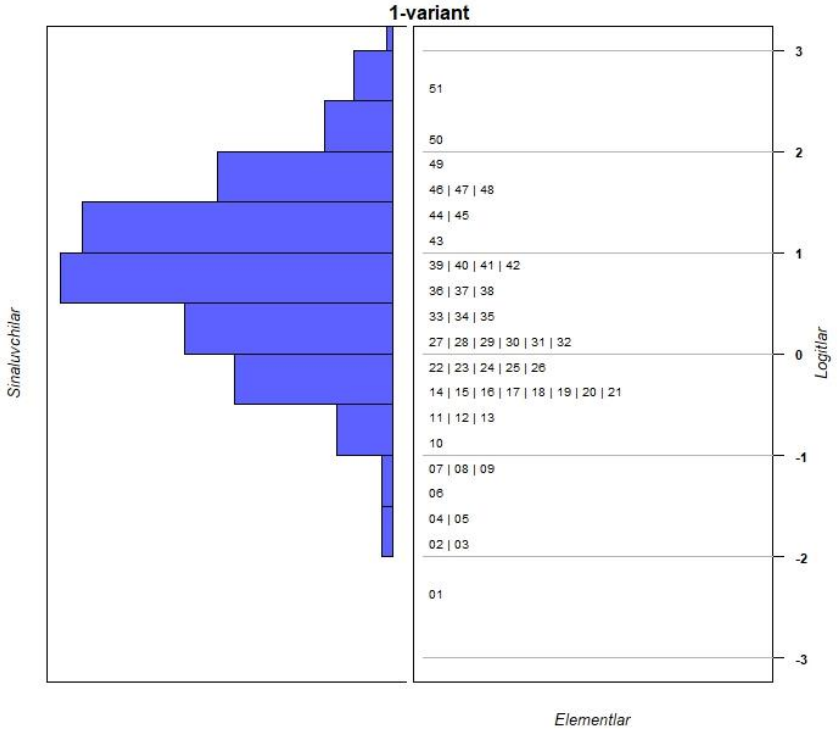
Rash modeli asosida aniqlangan qiyinlik darajalarini talabgorlar qobiliyatlariga qanchalik mosligini Rayt xaritasi yordamida tahlil qilish mumkin.

Rayt xaritasi – test topshiriqlarining qiyinlik darajalari va talabgorlarning qobiliyat darajalari o‘zaro mos kelishini aniqlovchi diagrammadir [21]. Rayt xaritasining chizmasini olishda [22] dasturiy paketidan foydalanildi. 2-, 3- va 4-rasmlarda mos ravishda geografiya fanidan milliy sertifikat uchun o‘tkazilgan 1-, 2- va 3-test sinovi natijalari asosida chizilgan Rayt xaritalari keltirilgan. 2-rasmdan qobiliyat darajalari $\boldsymbol{-3,05}$ va $\boldsymbol{2,83}$ logit birligi orasida, test topshiriqlari qiyinlik darajalari esa $(\boldsymbol{-3,95:3,36})$ oraliqda va 3-rasmdan qobiliyat darajalari $\boldsymbol{-1,62}$ va $\boldsymbol{3,58}$ logit birligi orasida, test topshiriqlari qiyinlik darajalari esa $(\boldsymbol{-2,39:4,97})$ oraliqda hamda 4-rasmdan qobiliyat darajalari $\boldsymbol{-4,34}$ va $\boldsymbol{4,16}$ logit birligi orasida, test topshiriqlari qiyinlik darajalari esa $(\boldsymbol{-1,93:3,63})$ oraliqda taqsimlanganligi kuzatildi.

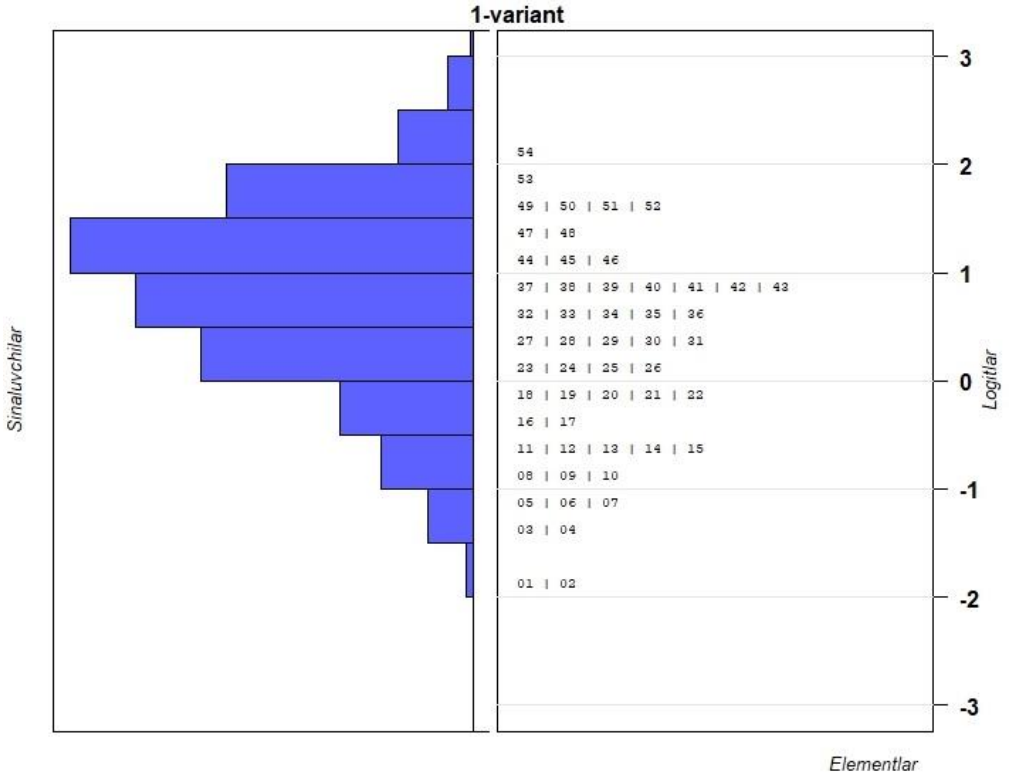
Qiyinlik darajasi bo‘yicha 2-rasmdagi 55-o‘rinda turgan, 3-rasmdagi 52-55-o‘rinlarda turgan va 4-rasmdagi 55-o‘rinda turgan test topshirilari - test topshiriqlarining qiyinlik darajalari bo‘yicha 3 logit birligidan tashqarida ekanligi ko‘rinib turibdi.



2-rasm. 1-test sinovi natijalari asosida aniqlangan qobiliyat va qiyinlik darajalarining mosligi



3-rasm. 2-test sinovi natijalari asosida aniqlangan qobiliyat va qiyinlik darajalarining mosligi



4-rasm. 3-test sinovi natijalari asosida aniqlangan qobiliyat va qiyinlik darajalarining mosligi

Demak, -3 logit birligidan kichik va 3 logit birligidan katta qiyinlik darajasiga ega bo‘lgan test topshiriqlari, mos ravishda, juda oson yoki juda qiyin test topshiriqlaridir.

Qiyinlik darajasi bo‘yicha $(-\mathbf{3}:\mathbf{3})$ logit birligi oralig‘idan tashqarida joylashgan (yuqorida keltirilgan) test topshiriqlaridan kam miqdorda ma’lumot olinishi sababli bunday test topshiriqlarining o‘rniga $(-\mathbf{3}:\mathbf{3})$ logit birligi oralig‘iga to‘g‘ri keladigan qiyinlikdagi test topshiriqlaridan qo‘yish maqsadga muvofiq bo‘ladi. Biroq bunday test topshiriqlari agarda talabgorlar orasida shu test topshiriqlariga mos qobiliyatli talabgorlar bo‘lsa, shunday test topshiriqlarining mavjudligi ular to‘g‘risida ma’lumot beradi. Aks holda qiyinlik darajasi juda past va juda yuqori bo‘lgan test topshiriqlarining o‘rniga $(-\mathbf{3}:\mathbf{3})$ logit birligi atrofidagi test topshiriqlaridan kiritish, yuqori va past qobiliyat darajalaridan olinadigan ma‘lumot miqdori orasidagi tafovutni yanada kamaytirish imkonini beradi.

2-rasmdan 1-test sinovi uchun tanlab olingan variantda test topshriqlari talabgorlarning qobiliyatlariga mos ekanligi ko‘rinadi.

3-rasmdan qiyinlik darajalari bo‘yicha 52-55-o‘rinlarda turgan test topshiriqlari 3 logit birligidan tashqarida bo‘lsa-da, ushbu test topshiriqlarini yecha oladigan

qobiliyatli talabgorlar ham mavjud ekanligi ko‘rinadi.

4-rasmda esa bunday test topshiriqlari 1 dona bo‘lib, bu qiyinlik darajasi bo‘yicha 55-o‘rinda turgan test topshirig‘idir. Bunda ham ushbu topshiriqni yecha oladigan yuqori qobiliyatli talabgorlarning mavjudligini ko‘rish mumkin.

Quyida test topshiriqlarining Rash modeli bilan aniqlangan qiyinlik darajalarini ko‘rib chiqamiz. 4-jadvalda geografiya fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalarining Rash modeli bilan aniqlangan qiyinlik darajalari keltirilgan.

4-jadvaldagi 1-test sinovi natijalari bo‘yicha aniqlangan qiyinlik darajalariga e’tibor bersak, 115553 ID raqamli test topshirig‘i eng qiyin, 115548 ID raqamli test topshirig‘i esa eng oson ekanligi, 2-test sinovi natijalari bo‘yicha aniqlangan qiyinlik darajalarini ko‘radigan bo‘lsak, 114107 ID raqamli test topshirig‘i eng qiyin, 114169 ID raqamli test topshirig‘i esa eng oson ekanligi va 3-test sinovi natijalari bo‘yicha aniqlangan qiyinlik darajalarini ko‘radigan bo‘lsak, 1073972 ID raqamli test topshirig‘i eng qiyin, 1073812 ID raqamli test topshirig‘i esa eng oson ekanligi ko‘rinadi. Bu esa 2-, 3- va 4-rasmlardagi Rayt xaritasida ham o‘z aksini topgan.

4-jadval

Rash modeli bilan aniqlangan qiyinlik darajalari

№	1- test sinovi		2- test sinovi		3- test sinovi	
	ID	***b***	**ID**	***b***	**ID**	***b***
1.	115548	-3,96	114169	-2,39	1073812	-1,93
2.	115765	-2,41	107211	-1,82	107318	-1,83
3.	115555	-2,37	107304	-1,82	113866	-1,32
4.	115781	-2,03	107322	-1,74	107286	-1,31
5.	115755	-1,91	113873	-1,56	112568	-1,24
6.	115770	-1,71	112546	-1,34	114196	-1,04
7.	115763	-1,38	107344	-1,19	107331	-1,00
8.	115762	-1,37	112612	-1,15	112607	-0,91
9.	115767	-1,35	112539	-1,01	112612	-0,88
10.	115771	-1,30	113869	-0,94	114179	-0,81

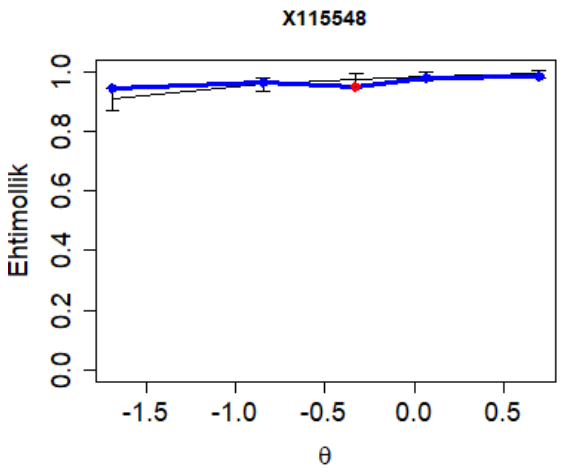
11.	115782	-1,26	107235	-0,73	107235	-0,66
12.	115776	-1,24	113832	-0,68	107381	-0,59
13.	115768	-1,22	107406	-0,65	107333	-0,59
14.	1155552	-0,96	11447433	-0,49	107212	-0,57
15.	115775	-0,95	114099	-0,49	114115	-0,51
16.	115783	-0,94	114126	-0,47	114193	-0,32
17.	115772	-0,91	104179	-0,44	114191	-0,28
18.	115754	-0,78	107391	-0,39	113848	-0,25
19.	115759	-0,77	107331	-0,37	112616	-0,22
20.	115780	-0,76	104183	-0,30	112592	-0,18
21.	115756	-0,66	112562	-0,25	107190	-0,13
22.	115758	-0,66	114474	-0,23	107230	-0,05
23.	115777	-0,53	113847	-0,23	113851	0,06
24.	115550	-0,37	113861	-0,18	114125	0,06
25.	115786	-0,30	114116	-0,14	107339	0,07
26.	11575435	-0,30	112580	-0,05	107206	0,18
27.	115773	-0,21	104190	0,05	107246	0,34
28.	115764	-0,18	112592	0,05	112547	0,37
29.	115761	-0,05	107204	0,16	107270	0,48
30.	115757	0,06	107230	0,18	1141252	0,48
31.	115785	0,15	114091	0,22	1141712	0,49
32.	1155502	0,24	107292	0,22	1141152	0,56
33.	115784	0,25	1073912	0,31	114132	0,59
34.	115769	0,28	112600	0,33	107376	0,63
35.	115779	0,38	1141162	0,43	107239	0,69
36.	115760	0,44	114190	0,58	1140832	0,71
37.	115766	0,56	113952	0,64	113949	0,77
38.	115546	0,65	114110	0,65	10740935	0,80
39.	115551	0,71	107341	0,82	114171	0,82
40.	11575434	0,82	114122	0,86	113957	0,92
41.	115547	0,99	1141222	0,93	1073762	0,93

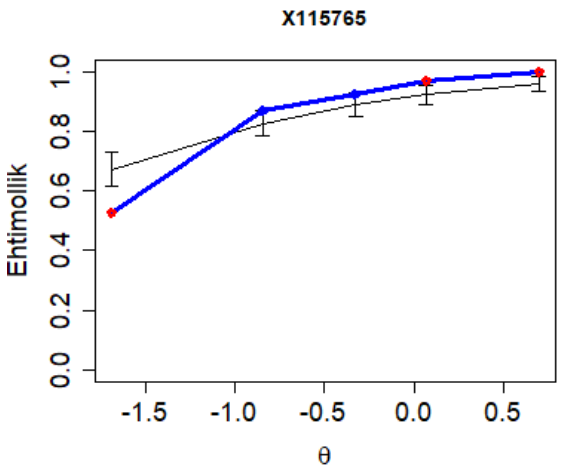
42.	115554	1,07	1141102	0,99	107402	0,98
43.	115774	1,08	11447434	1,10	107314	0,98
44.	1155482	1,16	1141262	1,28	107250	1,04
45.	1155522	1,38	107311	1,46	1073672	1,13
46.	115778	1,50	113955	1,54	112560	1,21
47.	115552	1,71	107316	1,64	111033	1,28
48.	1155492	1,81	112586	1,70	107397	1,41
49.	1155542	1,82	114097	1,83	114083	1,53
50.	115549	2,07	1140992	2,10	107367	1,54
51.	1155512	2,24	1140972	2,59	1074022	1,56
52.	1155462	2,41	1074062	3,23	1141322	1,58
53.	1155472	2,79	1140912	4,25	107409	1,98
54.	1155532	2,95	1141072	4,62	10740934	2,08
55.	115553	3,36	114107	4,97	1073972	3,69

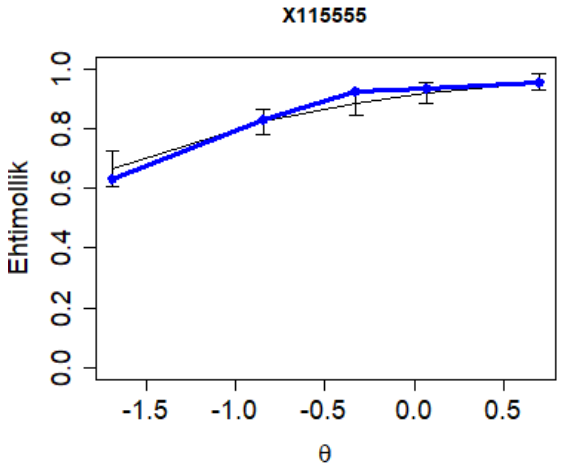
Test topshiriqlarining Rash modeli bilan mosligi ham muhim ahamiyatga kasb etadi. Har xil qobiliyat darajalarini aniqlash uchun kalibrovkalangan test topshiriqlari bazasini yaratishda test topshiriqlarining Rash modeli bilan moslik statistikalarini o‘rganib chiqish kerak.

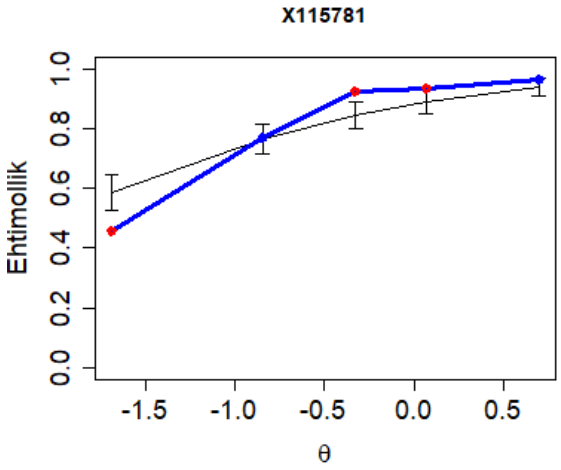
5*a*-, 5*b*- va 5*c*-rasmlarda mos ravishda geografiya fanidan milliy sertifikat uchun o‘tkazilgan 1-, 2- va 3-test sinovi natijalarining Rash modeli bilan mosligi – talabgorlar 5 ta qobiliyat guruhiga bo‘lingan va qalin ko‘k chiziqlar bilan test sinovlaridan olingan natijalar, ingichka qora chiziq bilan kutiladigan qiymatlar esa vertikal standart xatolik chiziqlari bilan birga ko‘rsatilgan. Standart xatolik chegarasidan chiqib ketgan nuqtalar qizil doirachalar bilan belgilangan. Rasmlarning yuqorisidagi raqamlar test topshiriqlarining ID raqamini bildiradi.

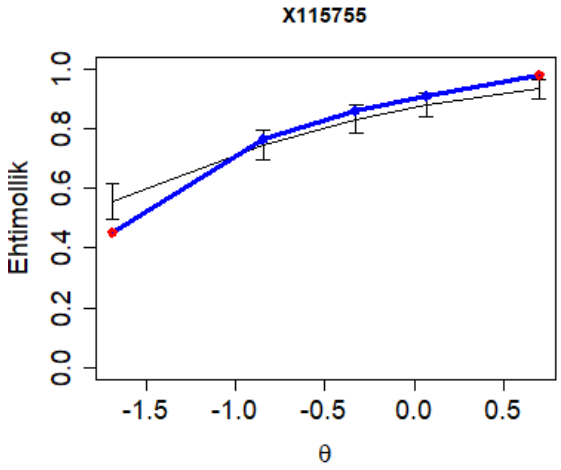
Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 1-test sinovi tahlili natijalariga ko‘ra ID raqamlari – 101155, 101166, 101412 va 101590 bo‘lgan test topshiriqlarining Rash modeli bilan moslik darajasi yaxshi emasligini ya’ni ajratilgan qobiliyat guruhlarining barchasi bilan mos tushmaganligini bildiradi (5*a*-rasm).

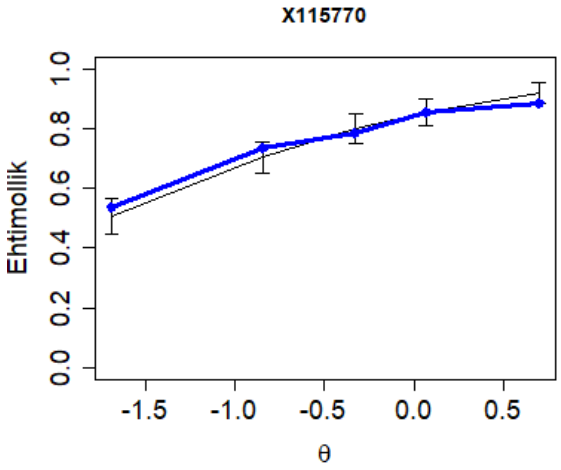

X115548
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

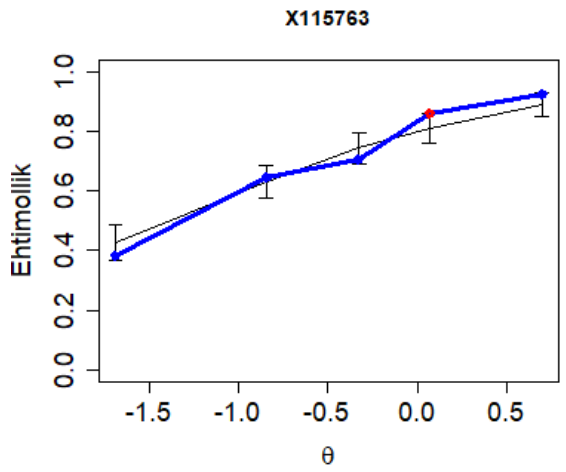

X115765
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

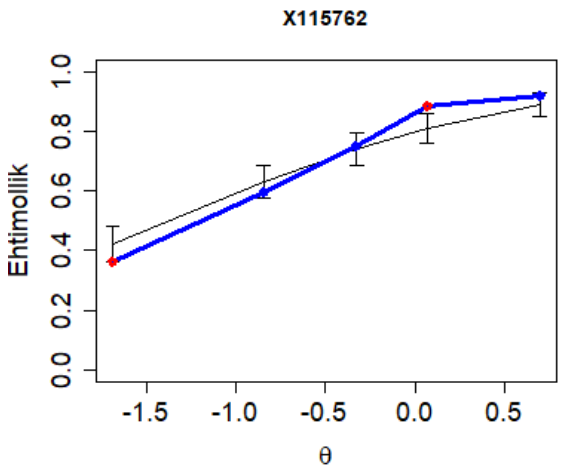

X115555
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

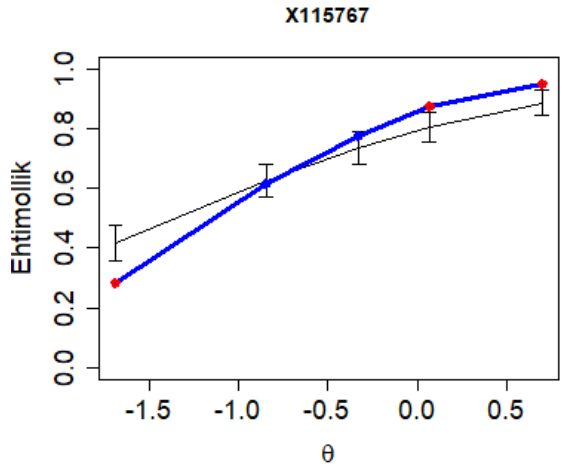

X115781
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

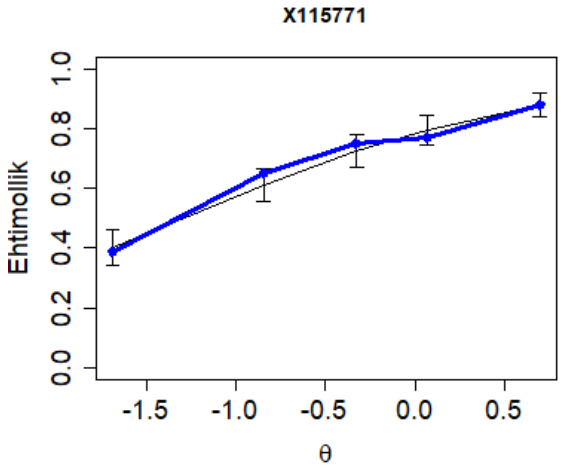

X115755
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

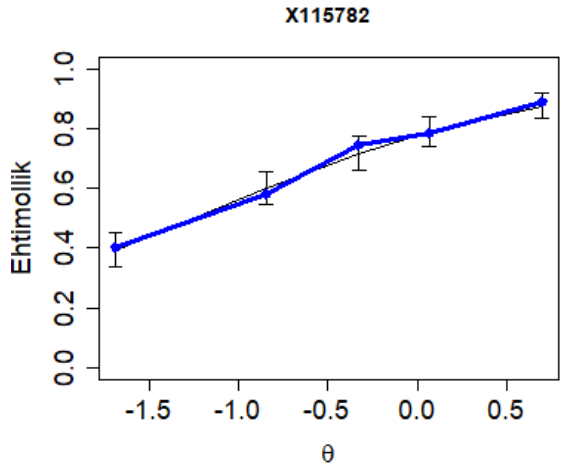

X115770
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

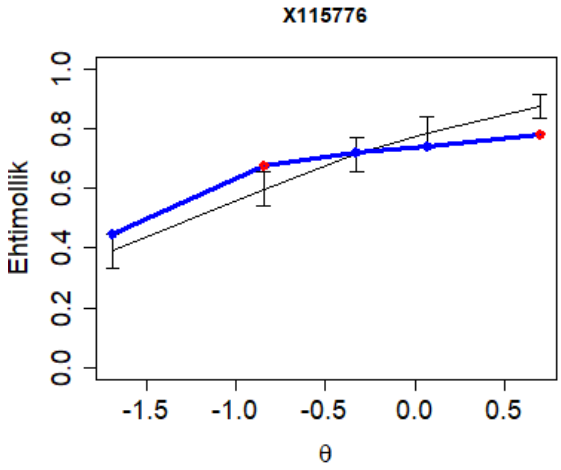

X115763
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

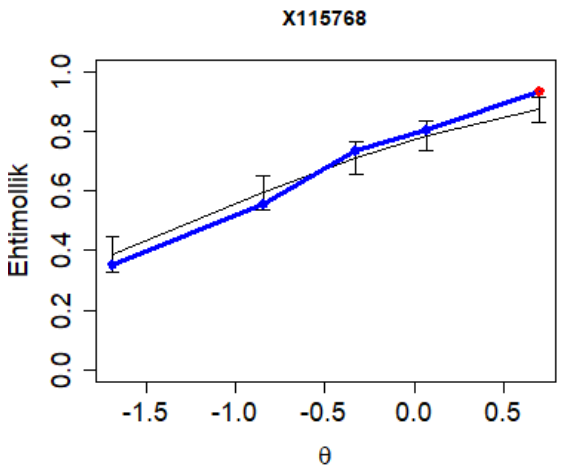

X115762
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

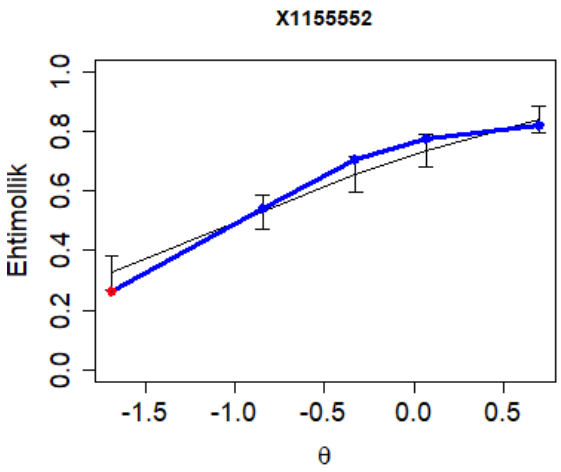

X115767
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

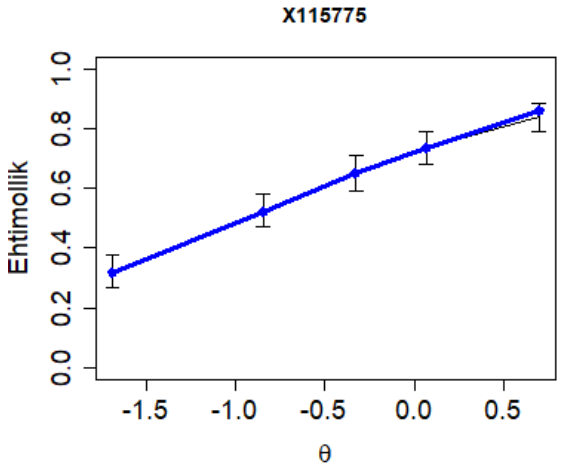

X115771
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

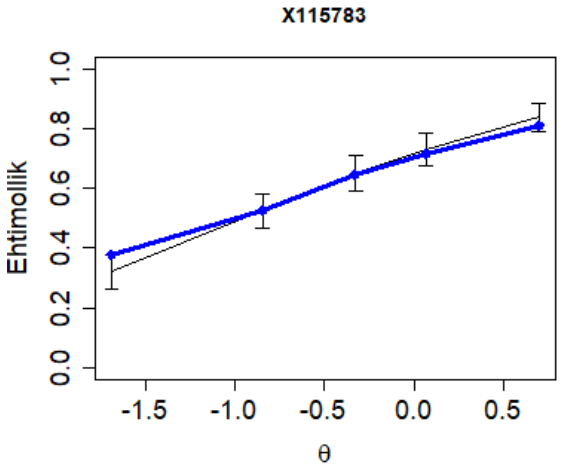

X115782
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

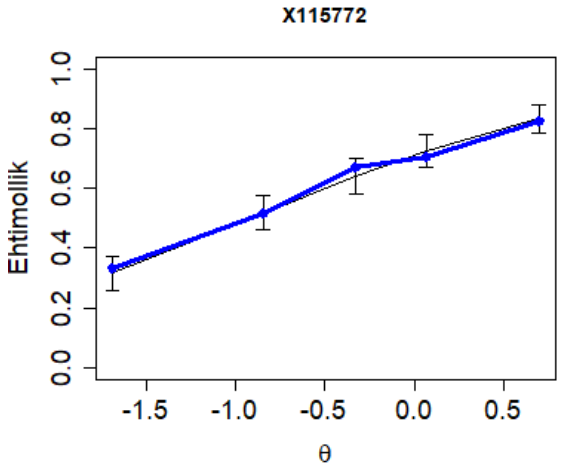

X115776
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

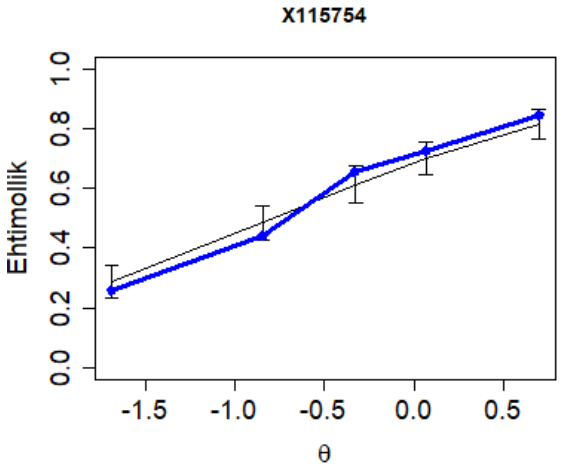

X115768
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

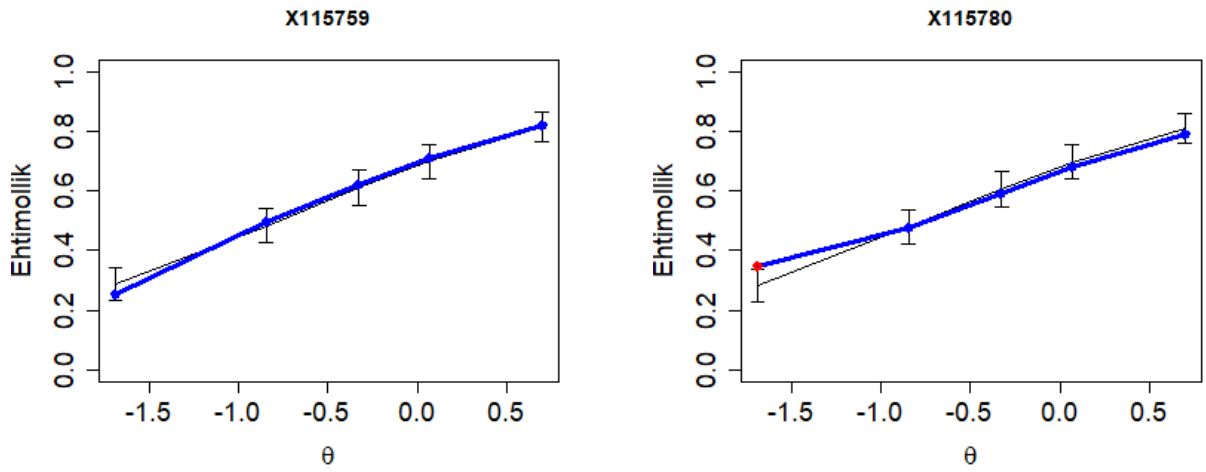

X1155552
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ


X115775
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

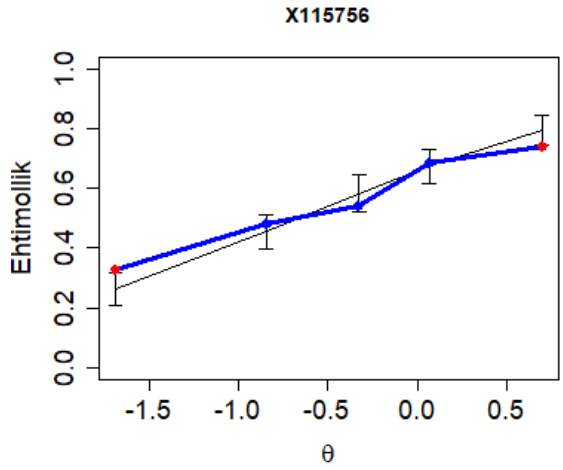

X115783
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

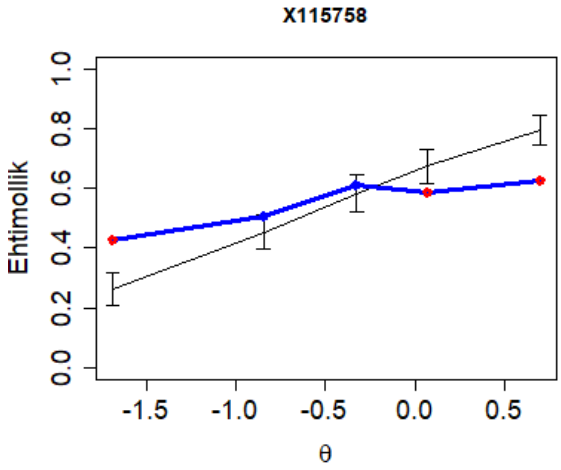

X115772
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

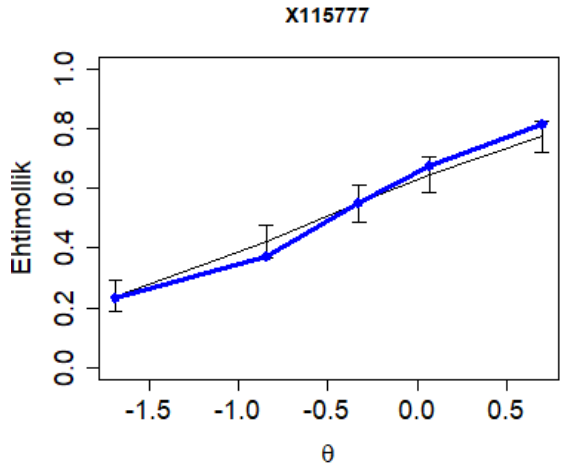

X115754
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

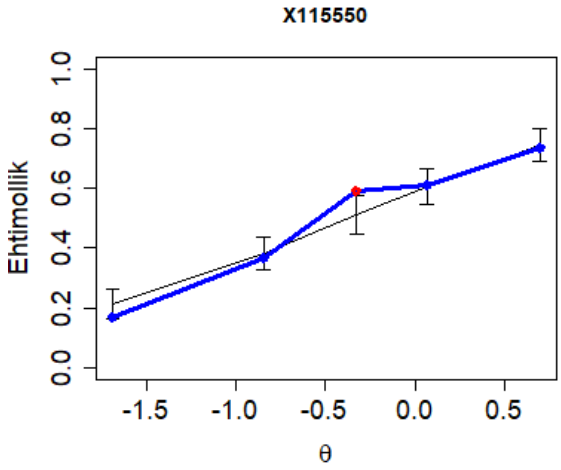

X115759
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

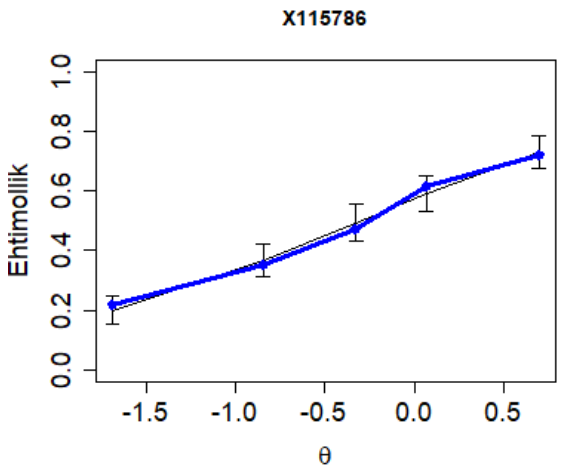
X115780
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

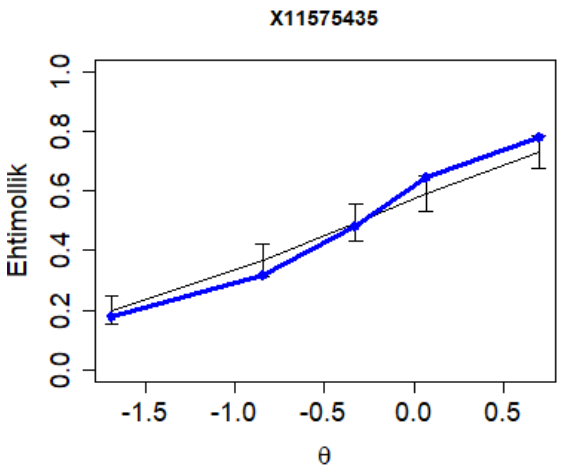

X115756
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

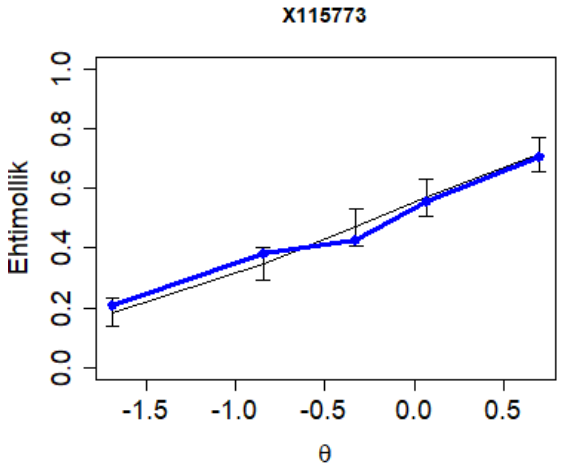

X115758
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

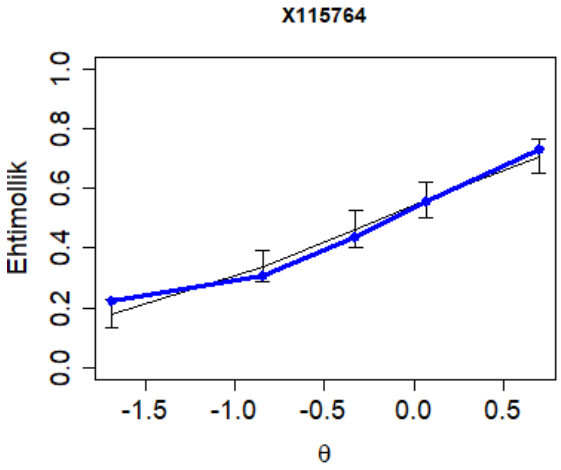

X115777
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

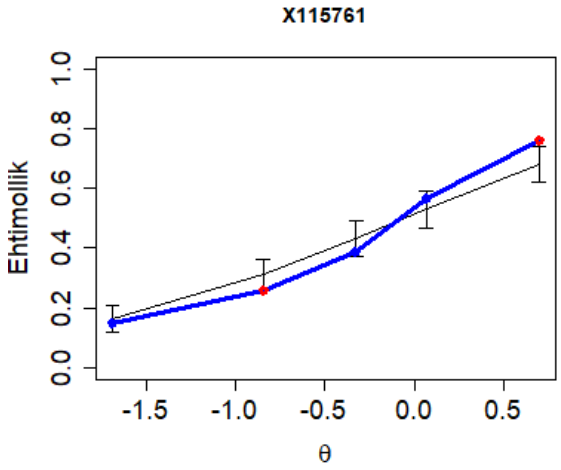

X115550
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

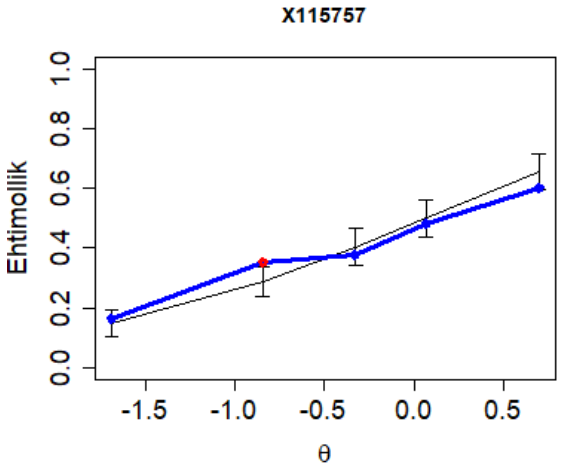

X115786
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

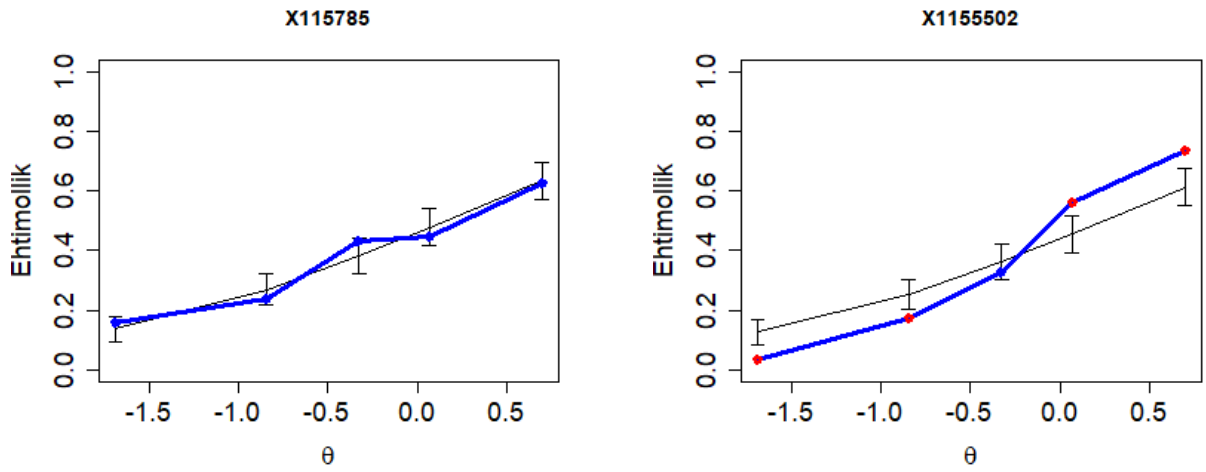

X11575435
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

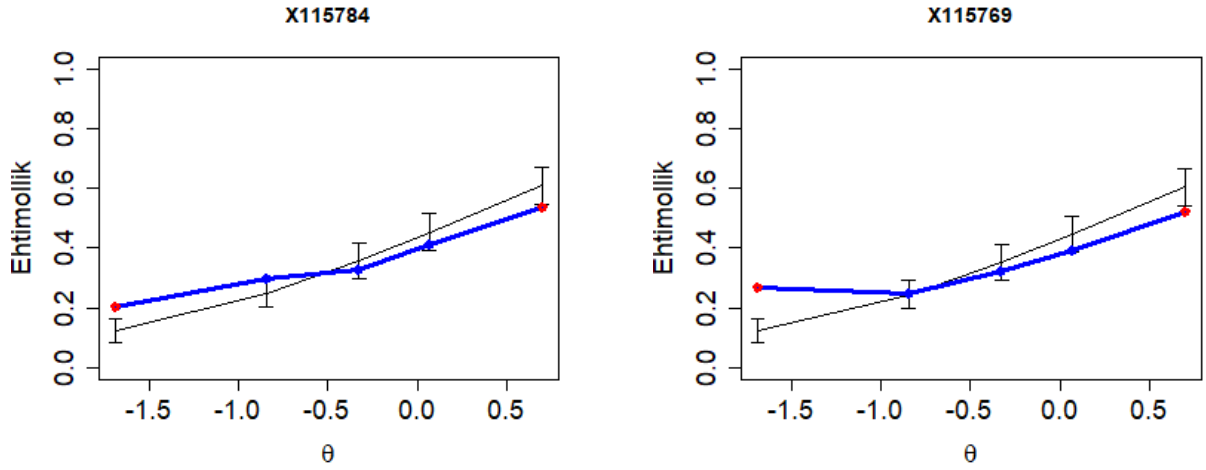

X115773
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

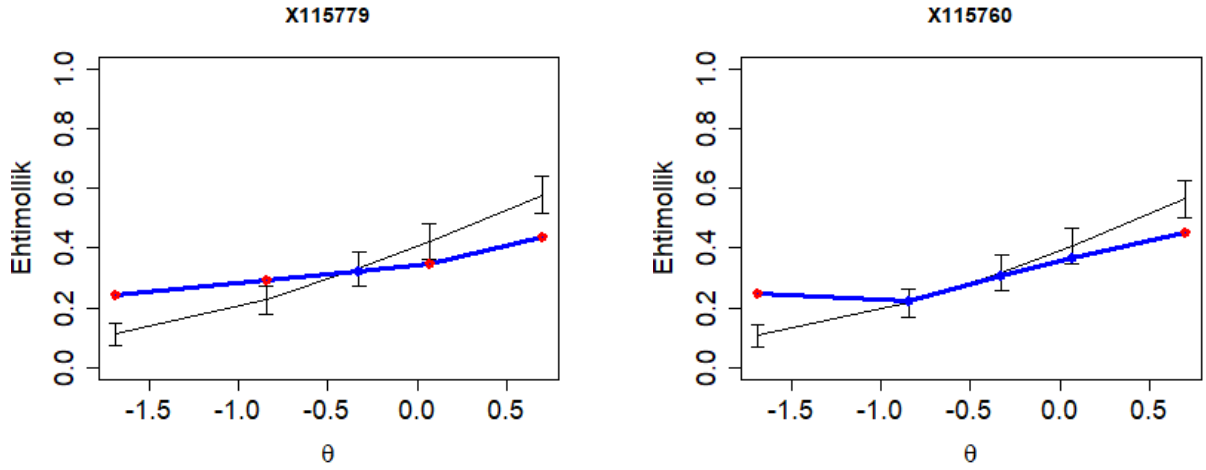

X115764
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

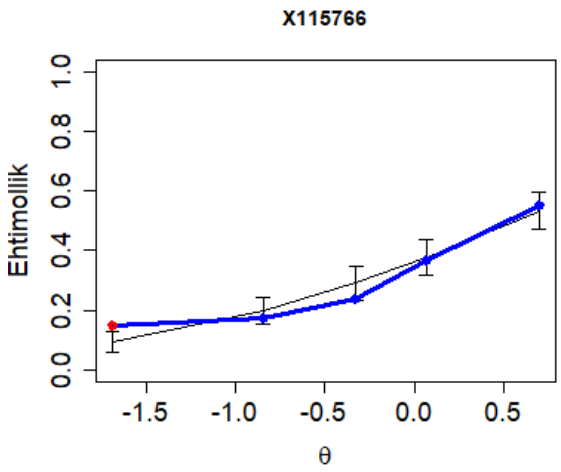

X115761
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

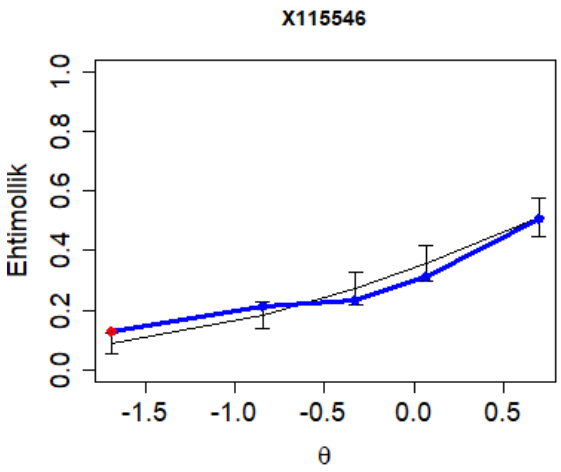

X115757
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

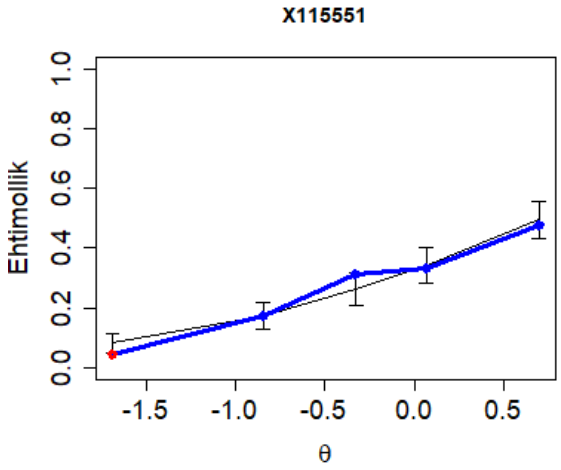

X115785
X1155502
Ehtimollik
θ
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5

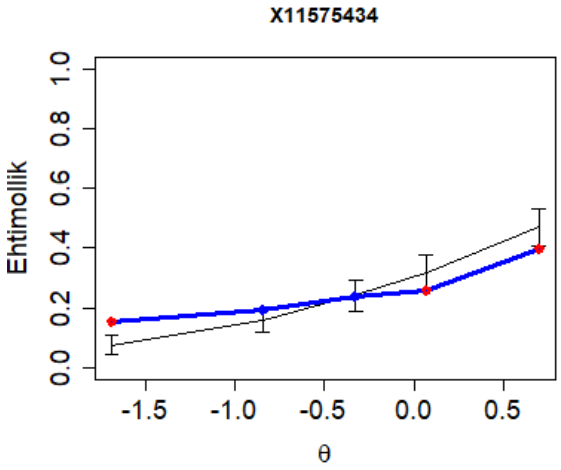

X115784
X115769
Ehtimollik
θ
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5

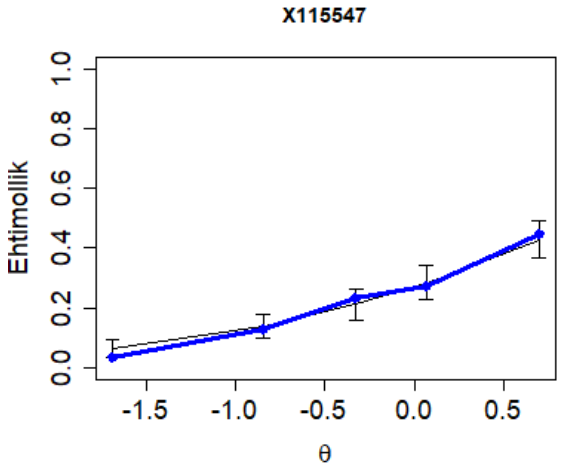

X115779
X115760
Ehtimollik
θ
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5

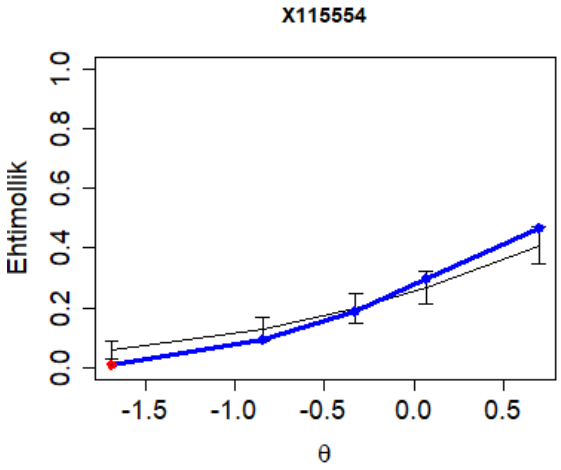

X115766
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

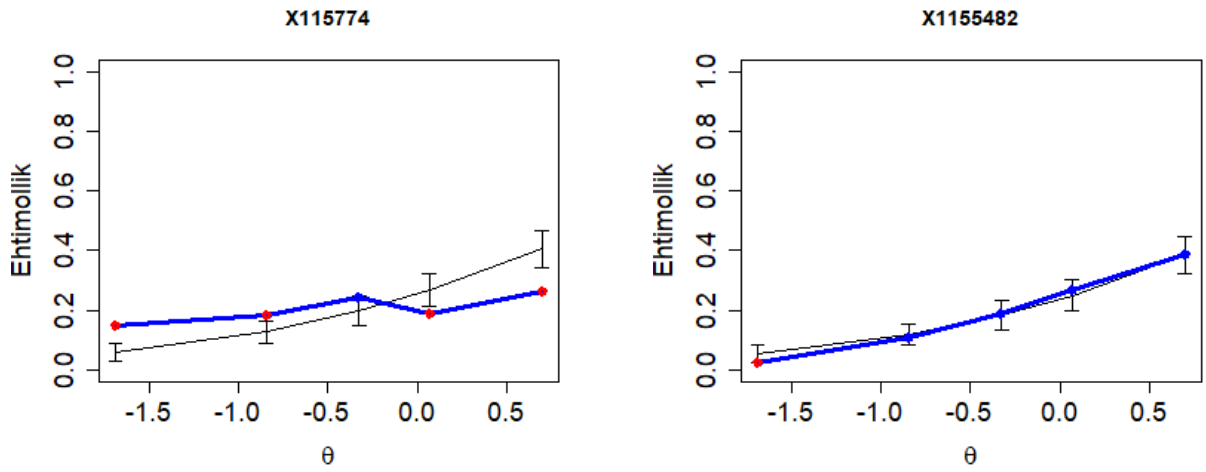

X115546
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ


X115551
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

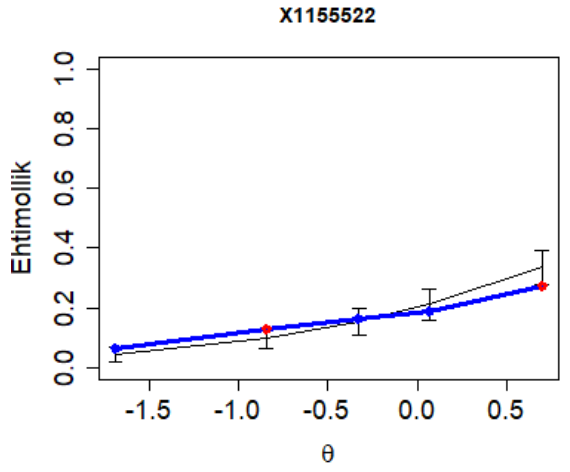

X11575434
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

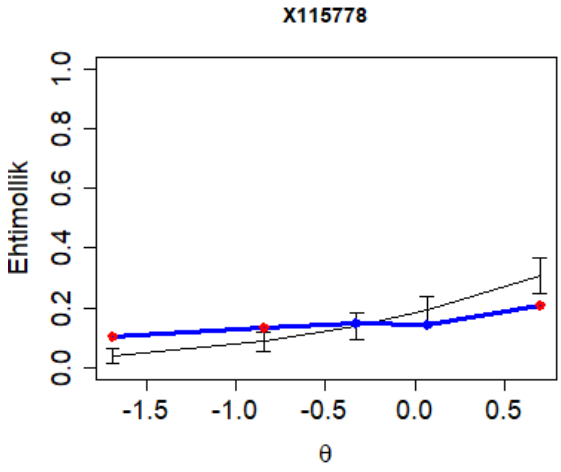

X115547
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

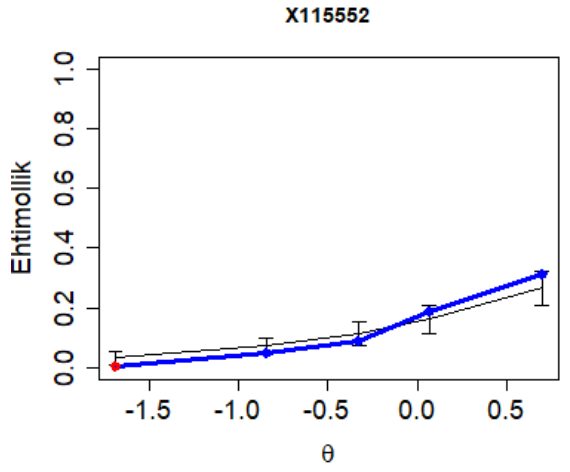

X115554
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

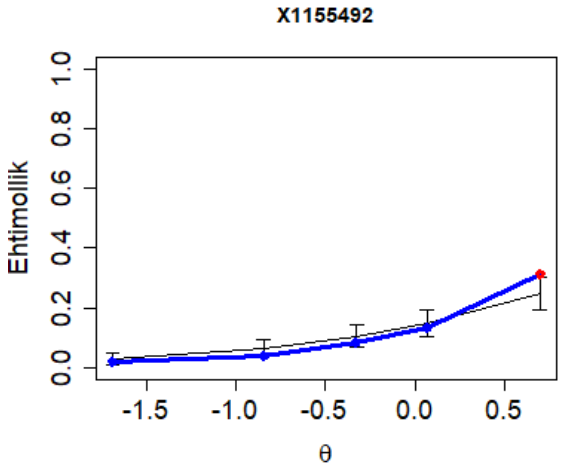

X115774
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

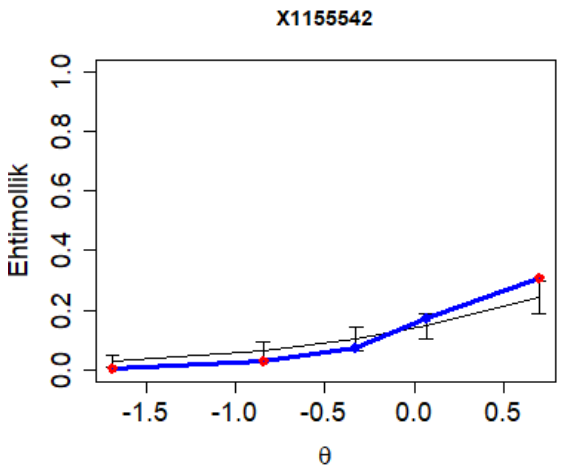
X1155482
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

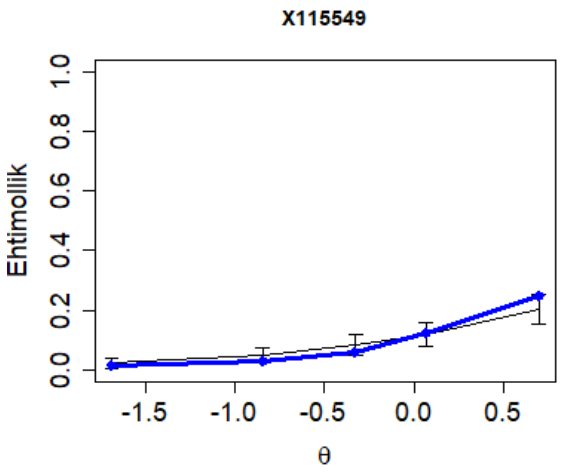

X1155522
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

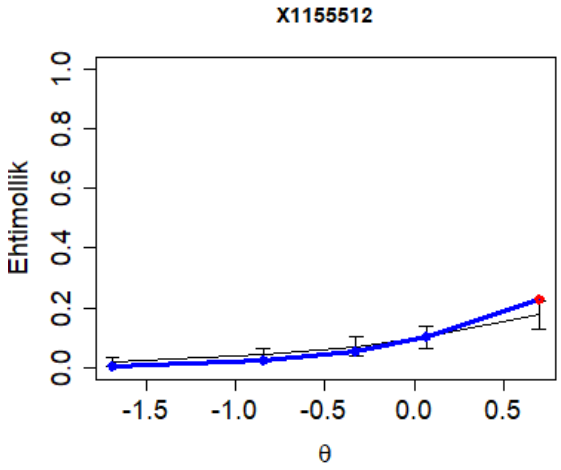

X115778
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

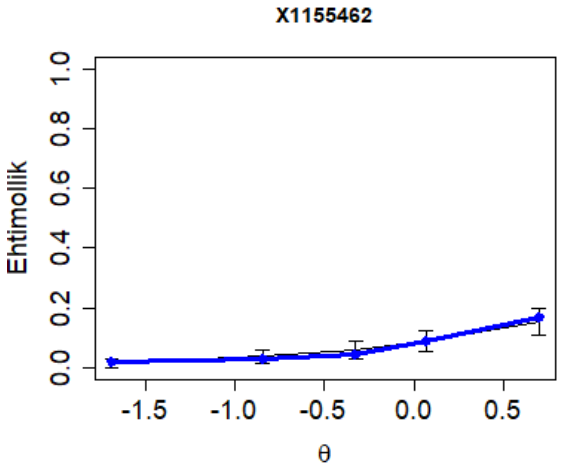

X115552
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

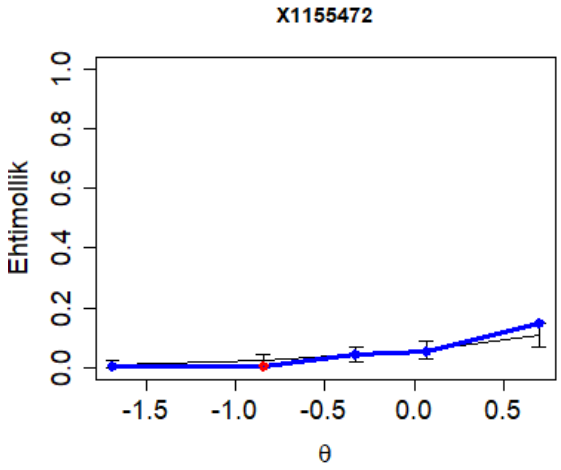

X1155492
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-1.5
-1.0
-0.5
0.0
0.5
θ

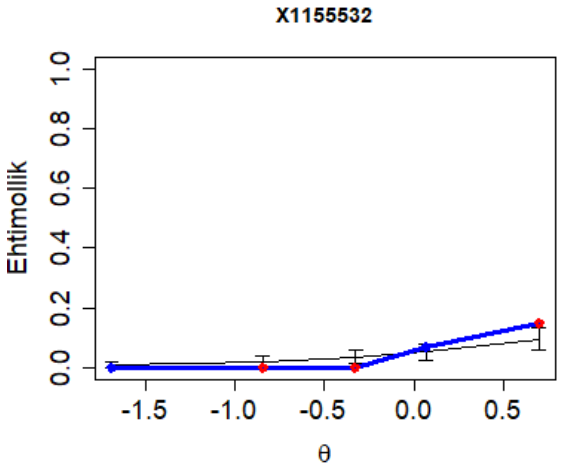

X1155542
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

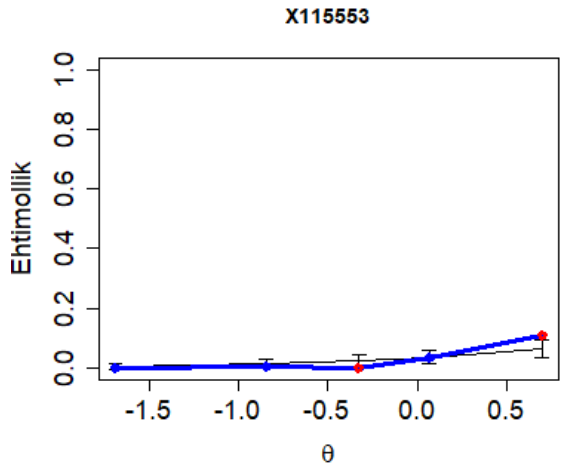

X115549
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ


X1155512
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ


X1155462
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

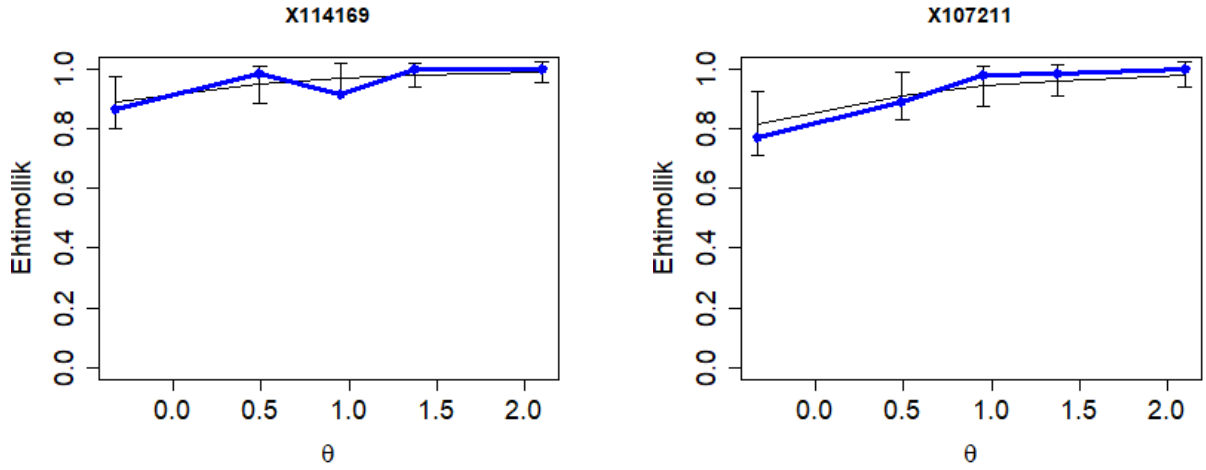

X1155472
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

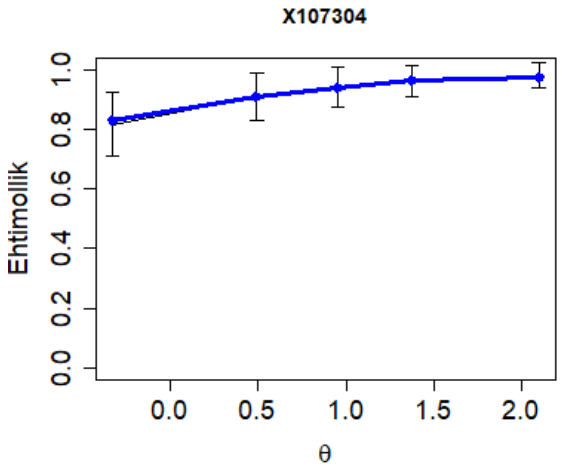

X1155532
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.5 -1.0 -0.5 0.0 0.5
θ

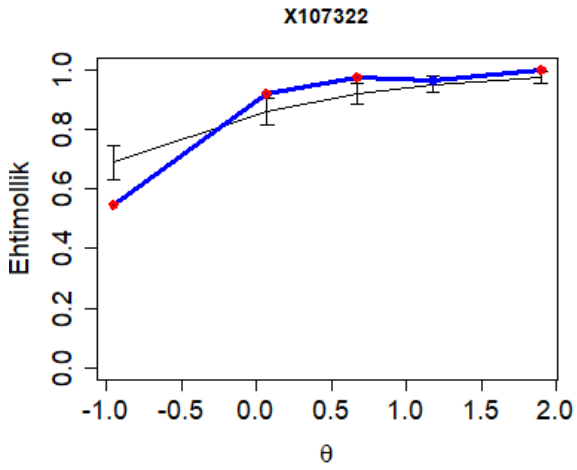


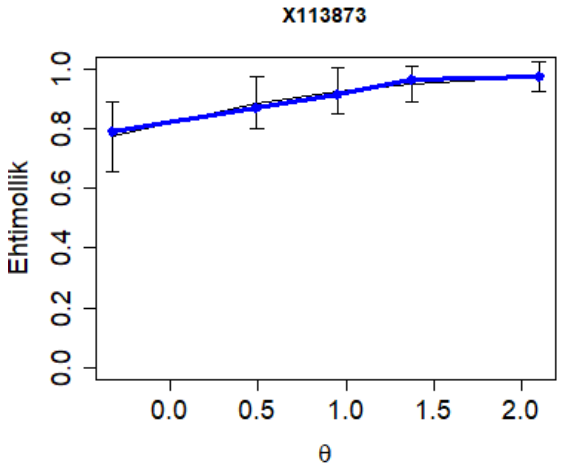
5*a*-rasm. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 1-test sinovi natijalarining Rash modeli bilan mosligi

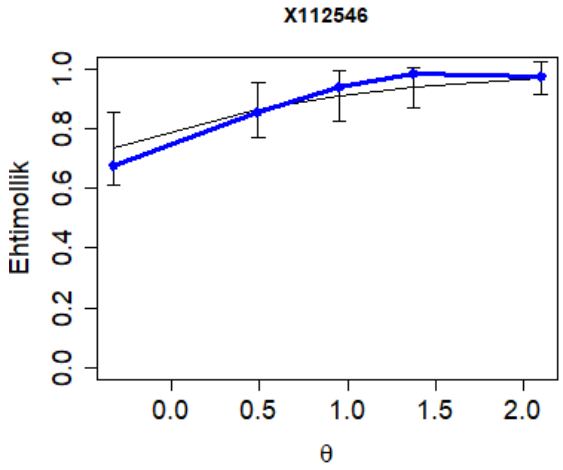
2-test sinovi tahlil natijalariga ko‘ra esa, ID raqamlari - 107322 va 11447434 bo‘lgan test topshiriq-larining Rash modeli bilan moslik darajasi yaxshi emasligini ya’ni bunda ham ajratilgan qobiliyat guruhlarining barchasi bilan mos tushmaganligini bildiradi (5*b*-rasm).

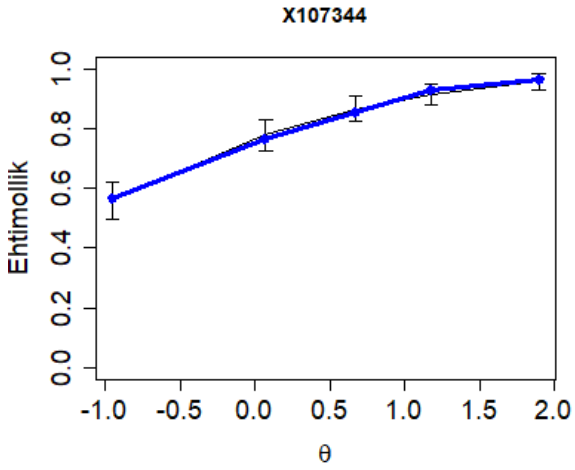


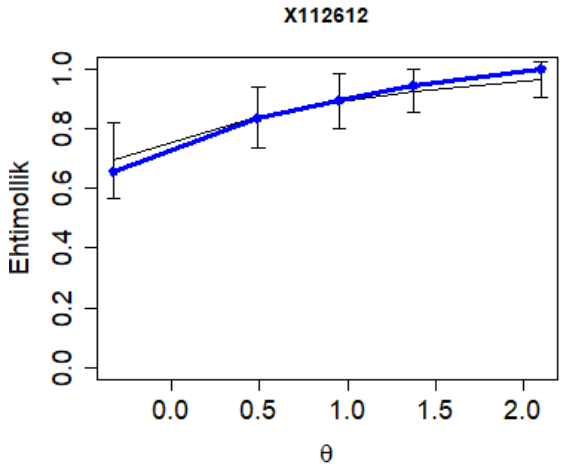

X107304
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

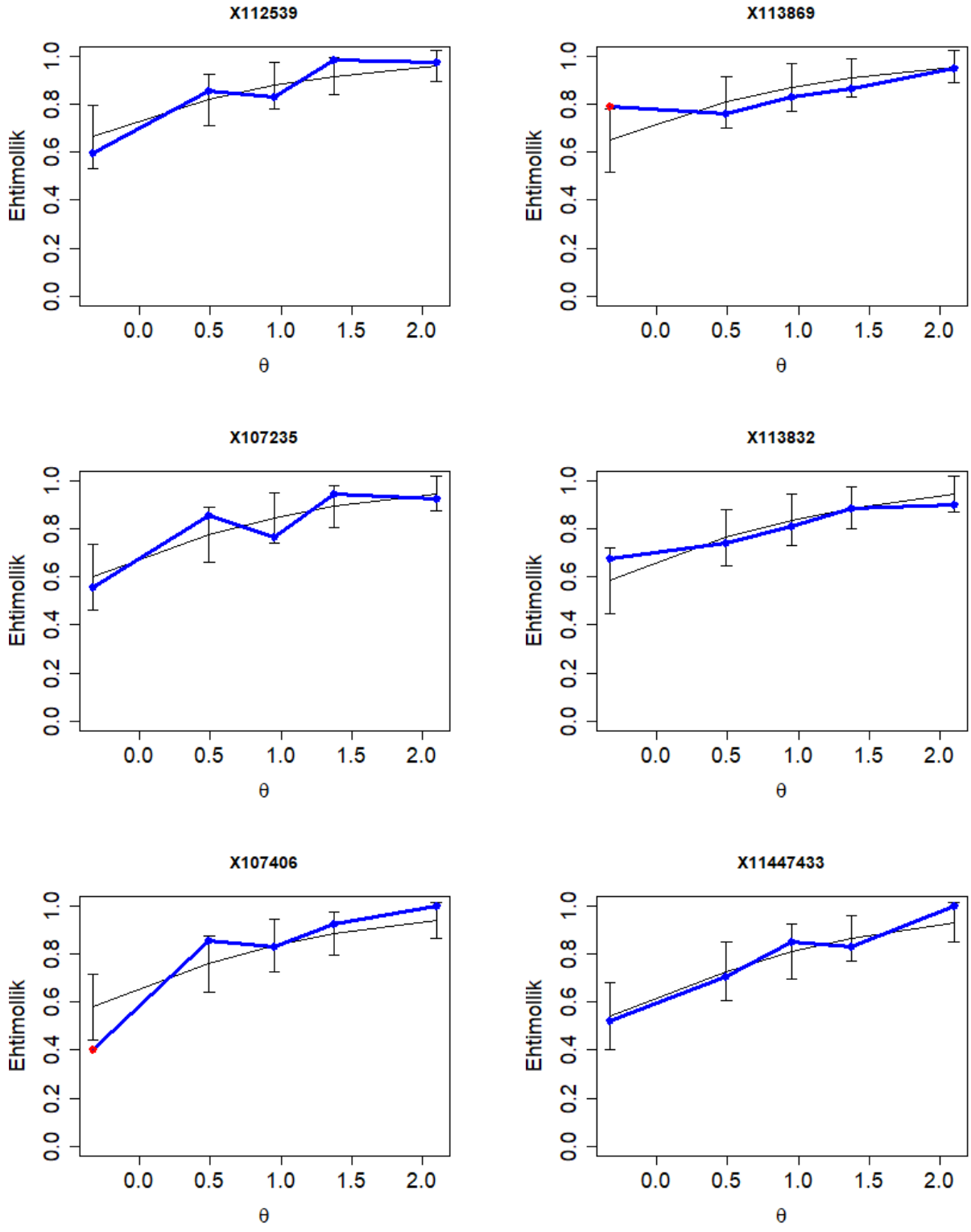

X107322
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.0 -0.5 0.0 0.5 1.0 1.5 2.0
θ

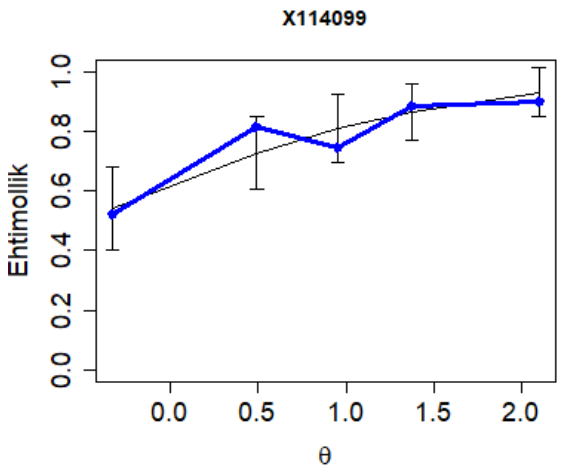

X113873
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

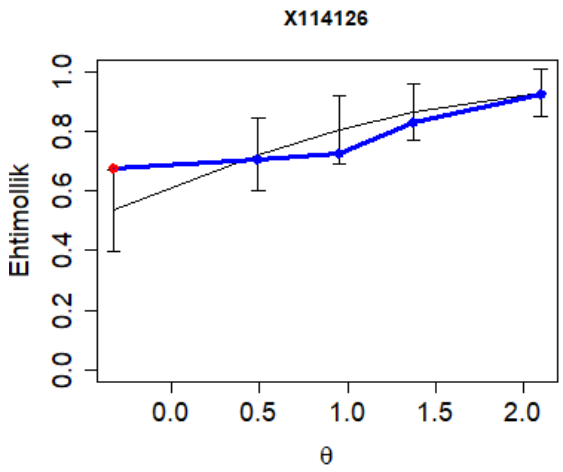

X112546
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

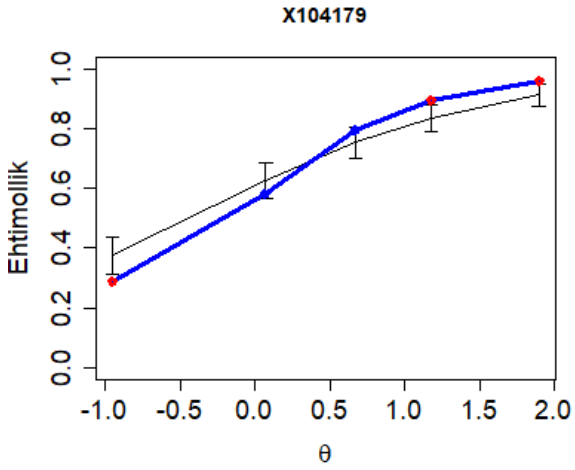

X107344
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.0 -0.5 0.0 0.5 1.0 1.5 2.0
θ

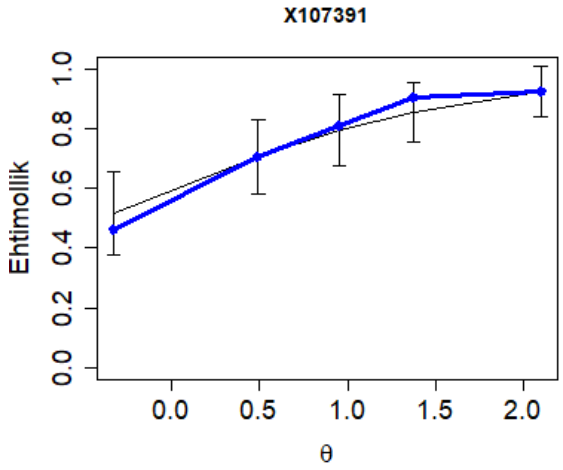

X112612
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

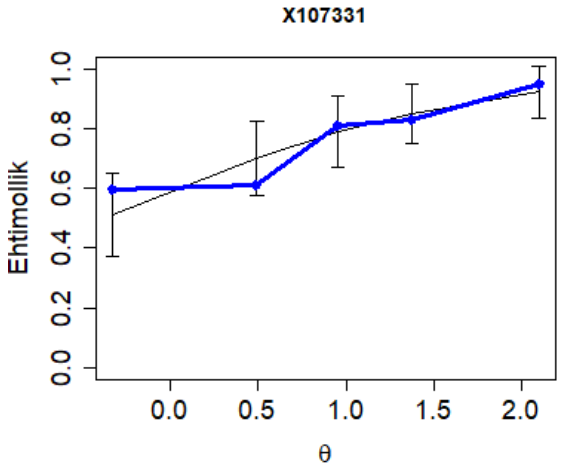

X112539
X113869
X107235
X113832
X107406
X11447433
Ehtimollik
θ
0.0
0.2
0.4
0.6
0.8
1.0
0.5
1.5
2.0

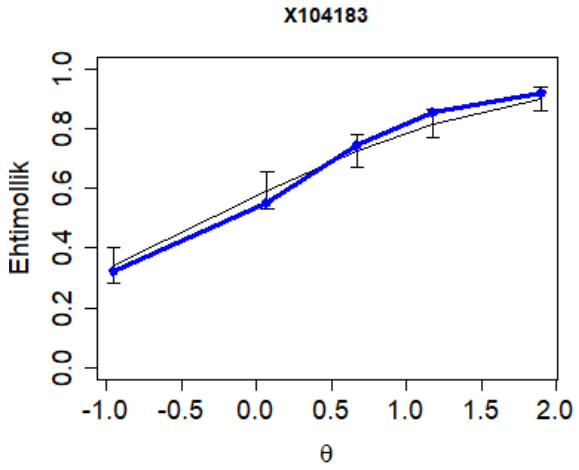

X114099
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

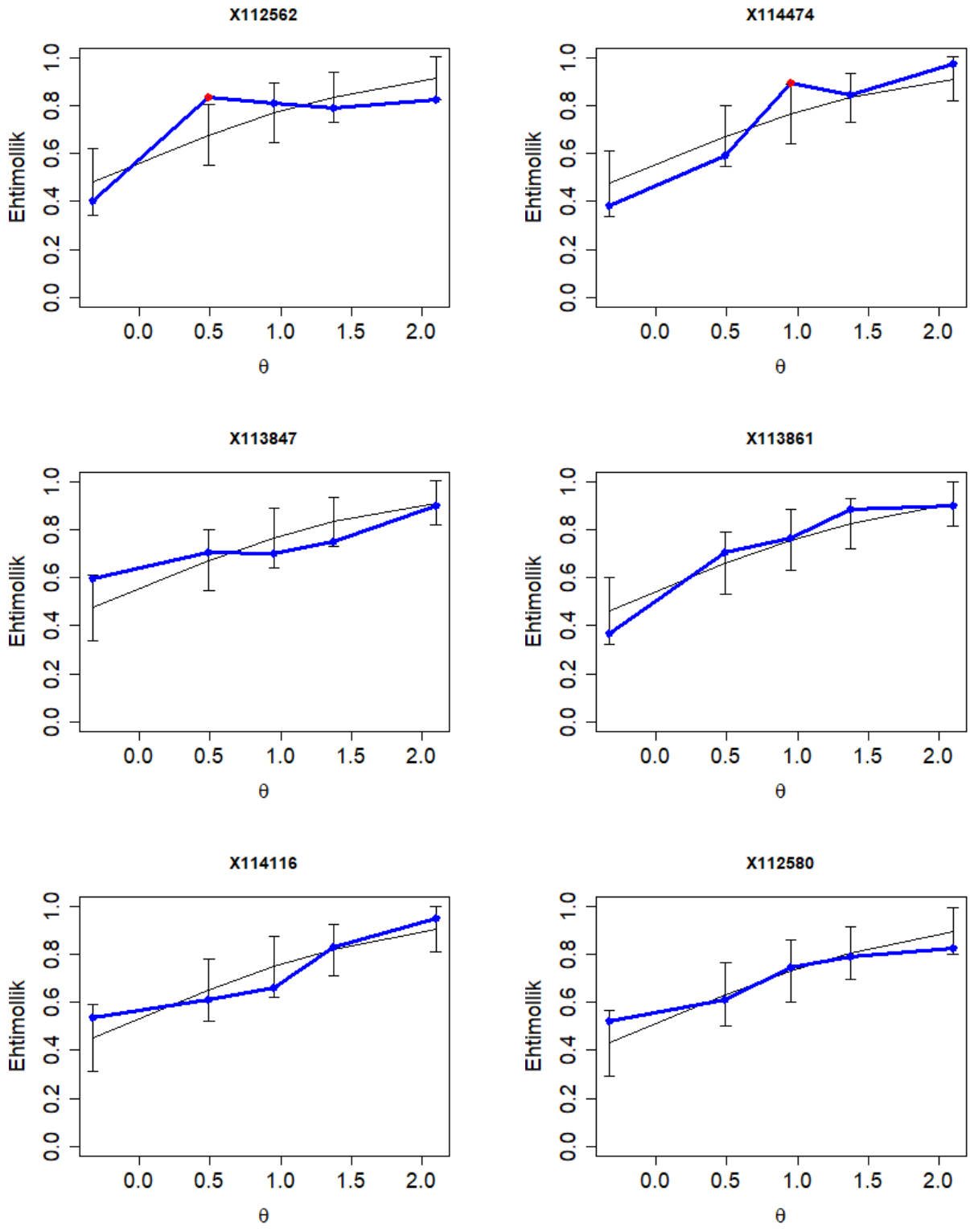

X114126
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

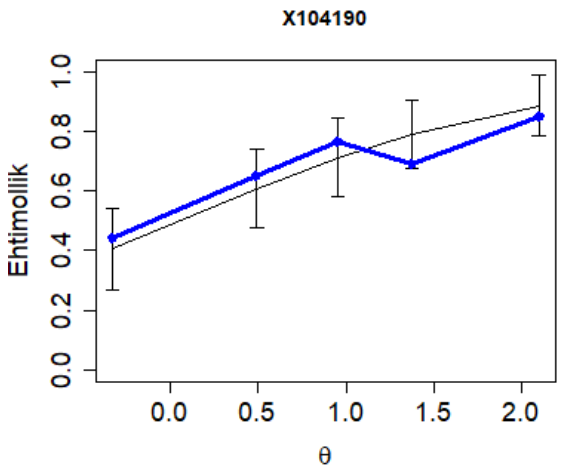

X104179
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.0 -0.5 0.0 0.5 1.0 1.5 2.0
θ

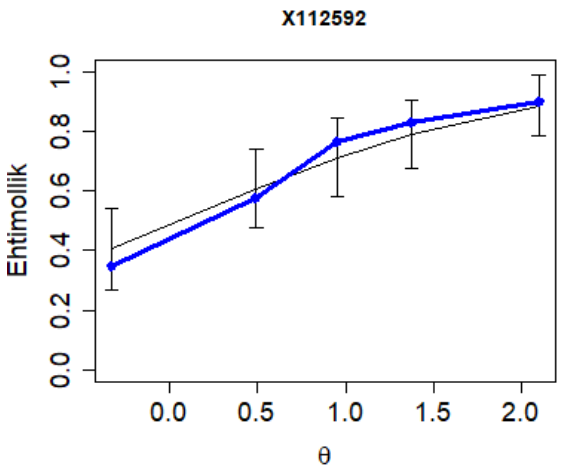

X107391
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

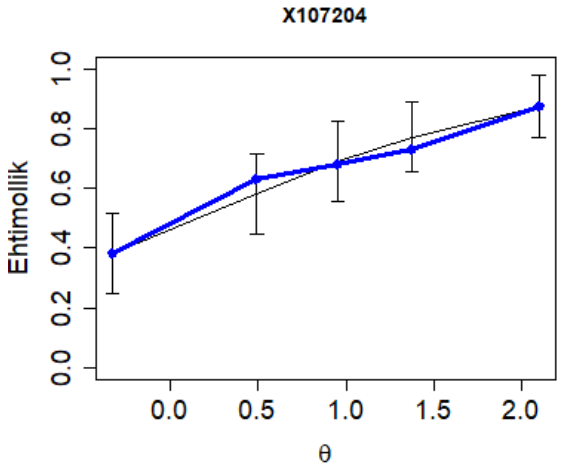

X107331
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

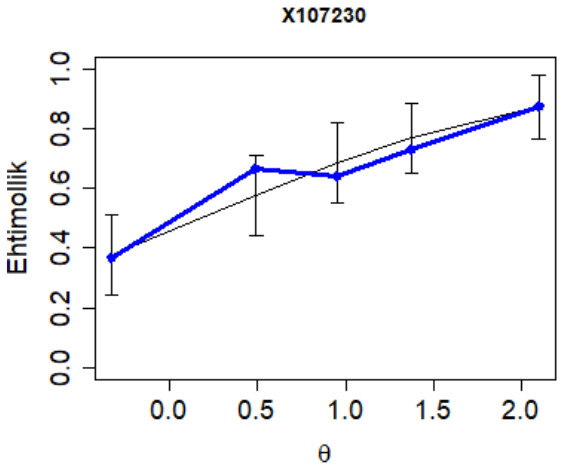

X104183
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.0 -0.5 0.0 0.5 1.0 1.5 2.0
θ

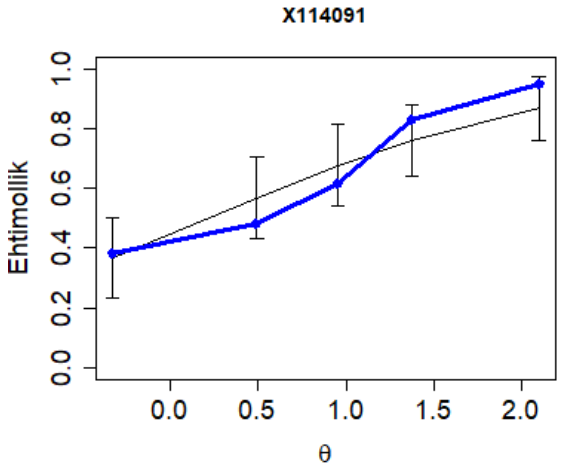

X112562
X114474
X113847
X113861
X114116
X112580
Ehtimollik
θ
0.0
0.2
0.4
0.6
0.8
1.0
0.5
1.5
2.0

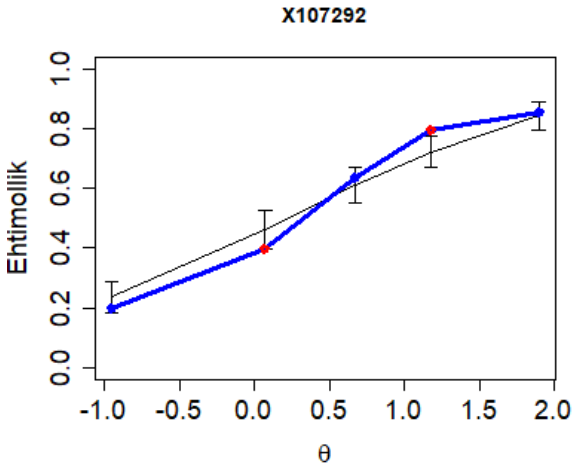

X104190
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

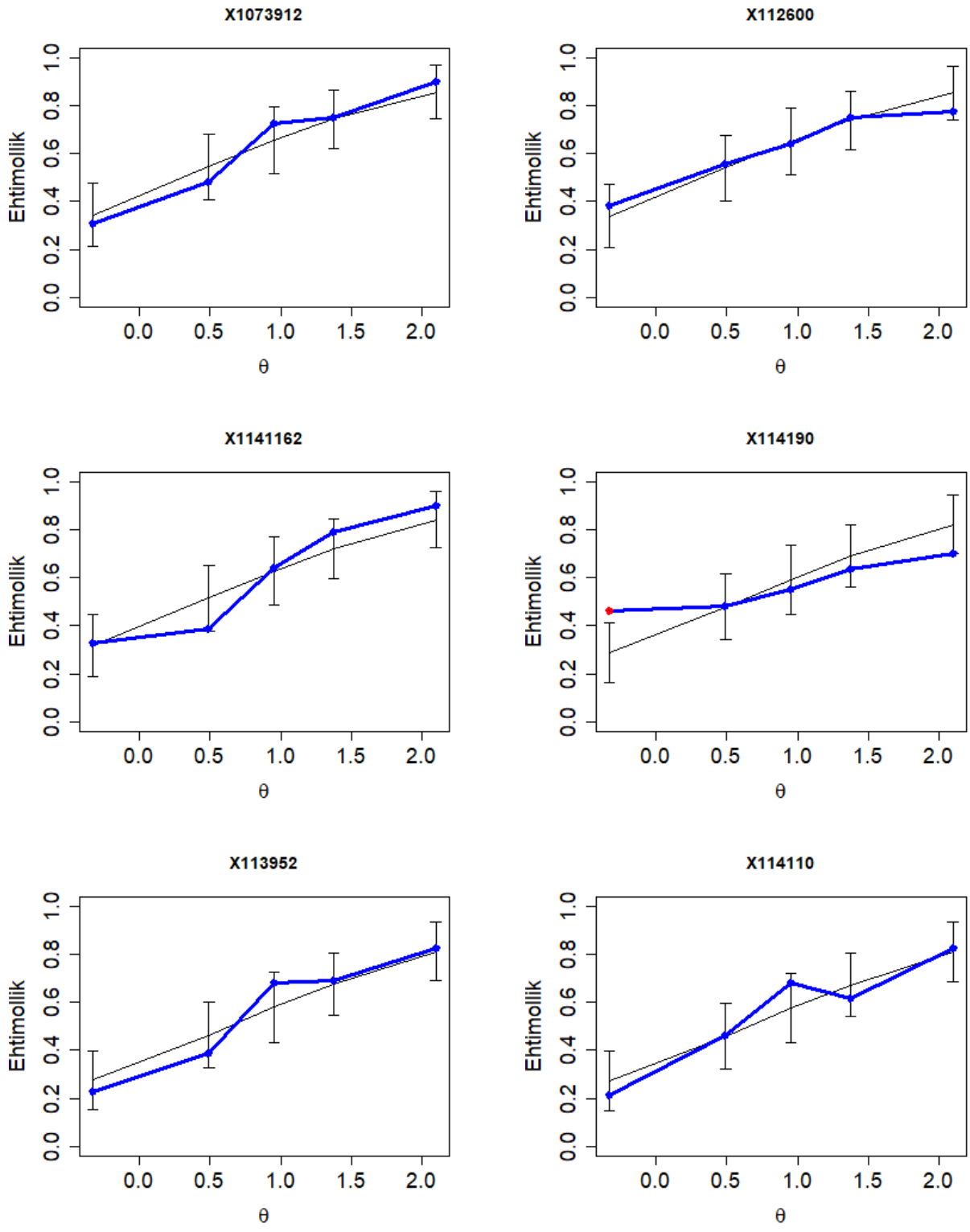

X112592
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

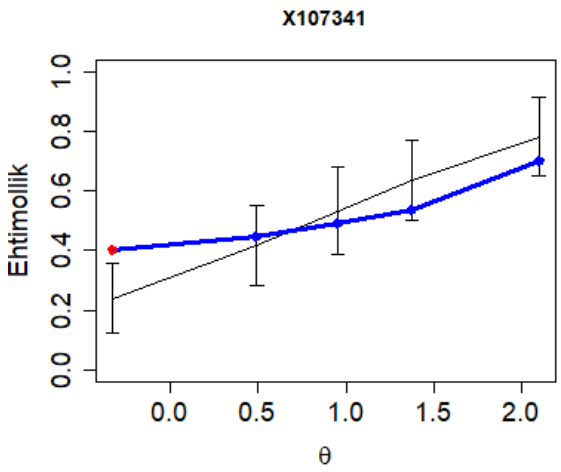

X107204
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

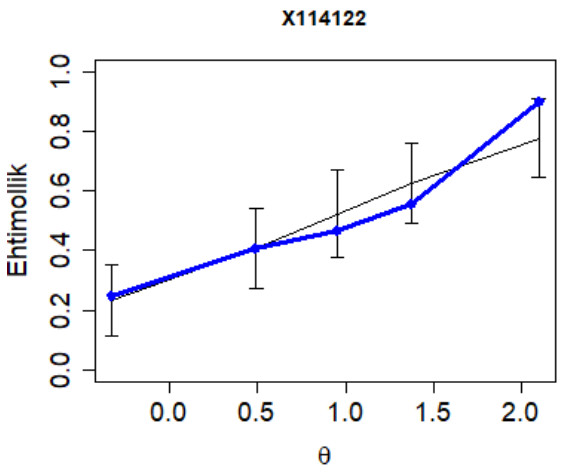

X107230
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

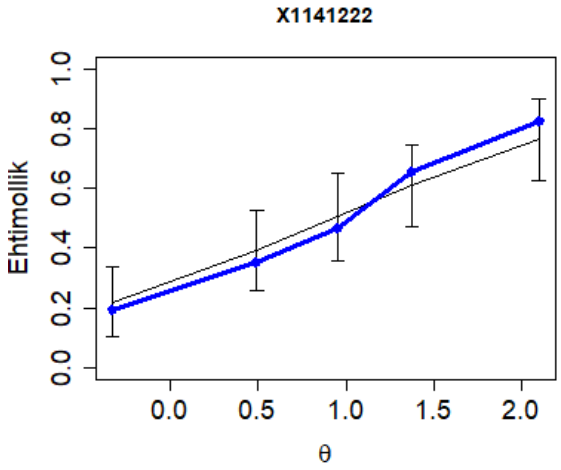

X114091
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

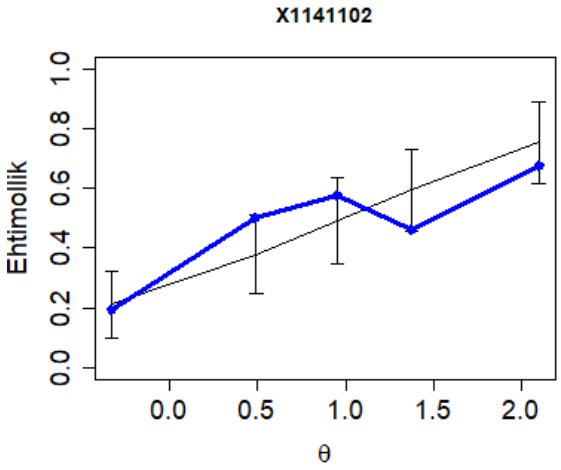

X107292
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-1.0 -0.5 0.0 0.5 1.0 1.5 2.0
θ

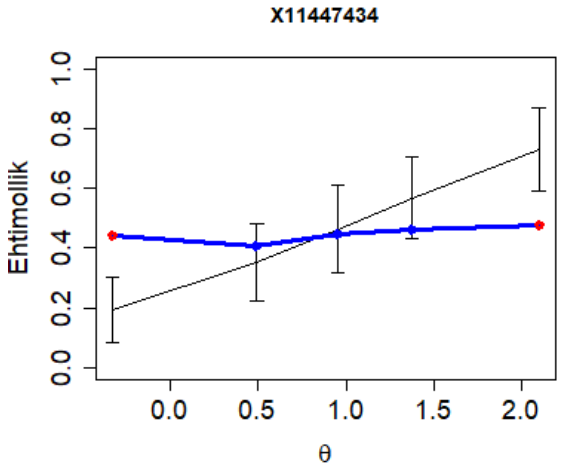

X1073912
X112600
X1141162
X114190
X113952
X114110
Ehtimollik
θ
0.0
0.2
0.4
0.6
0.8
1.0
0.5
1.5
2.0

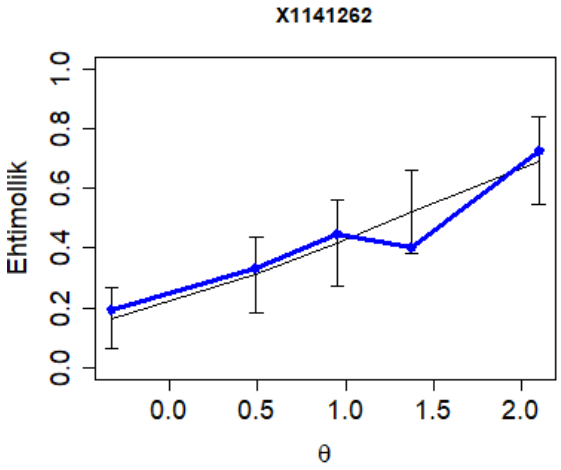

X107341
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

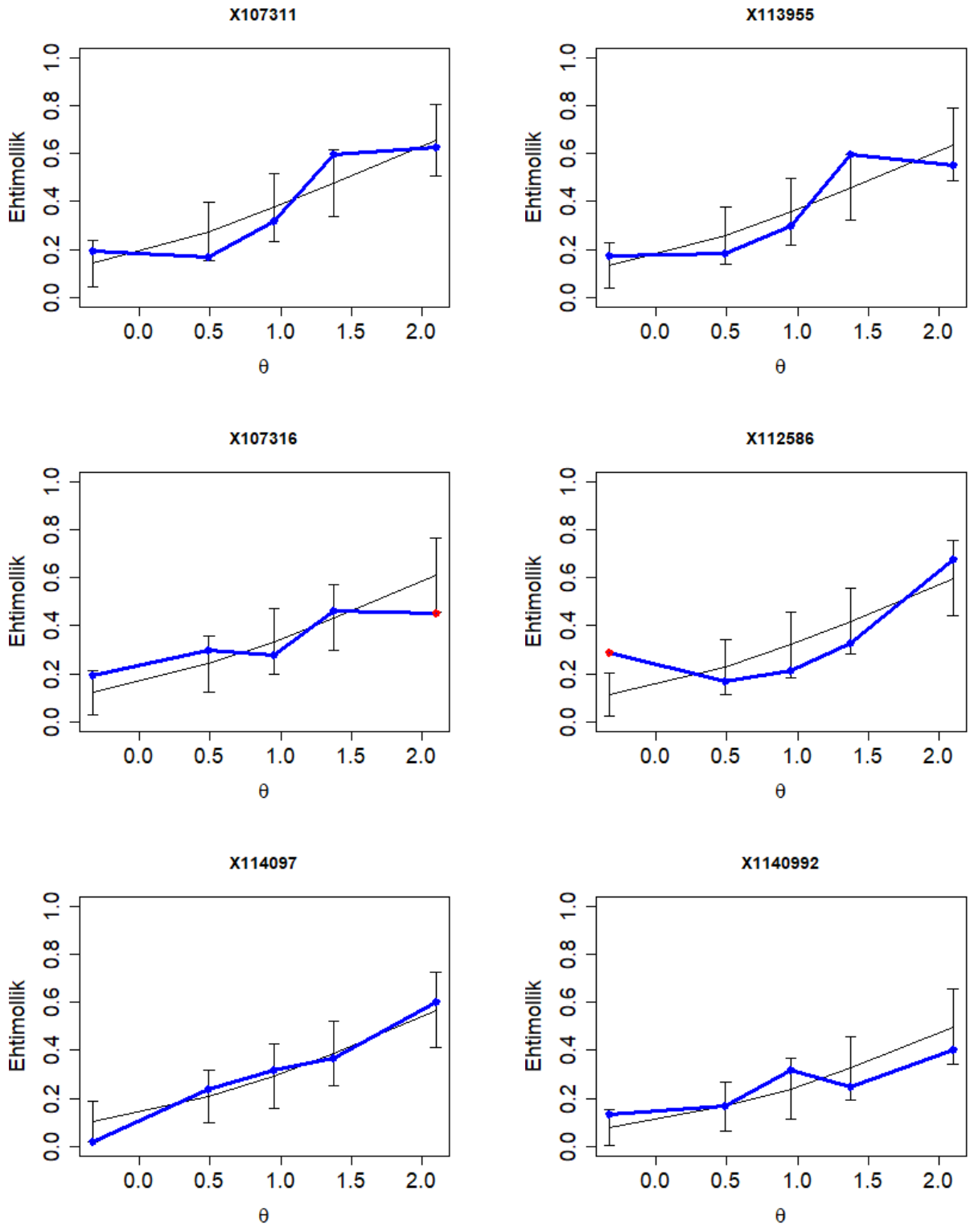

X114122
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ

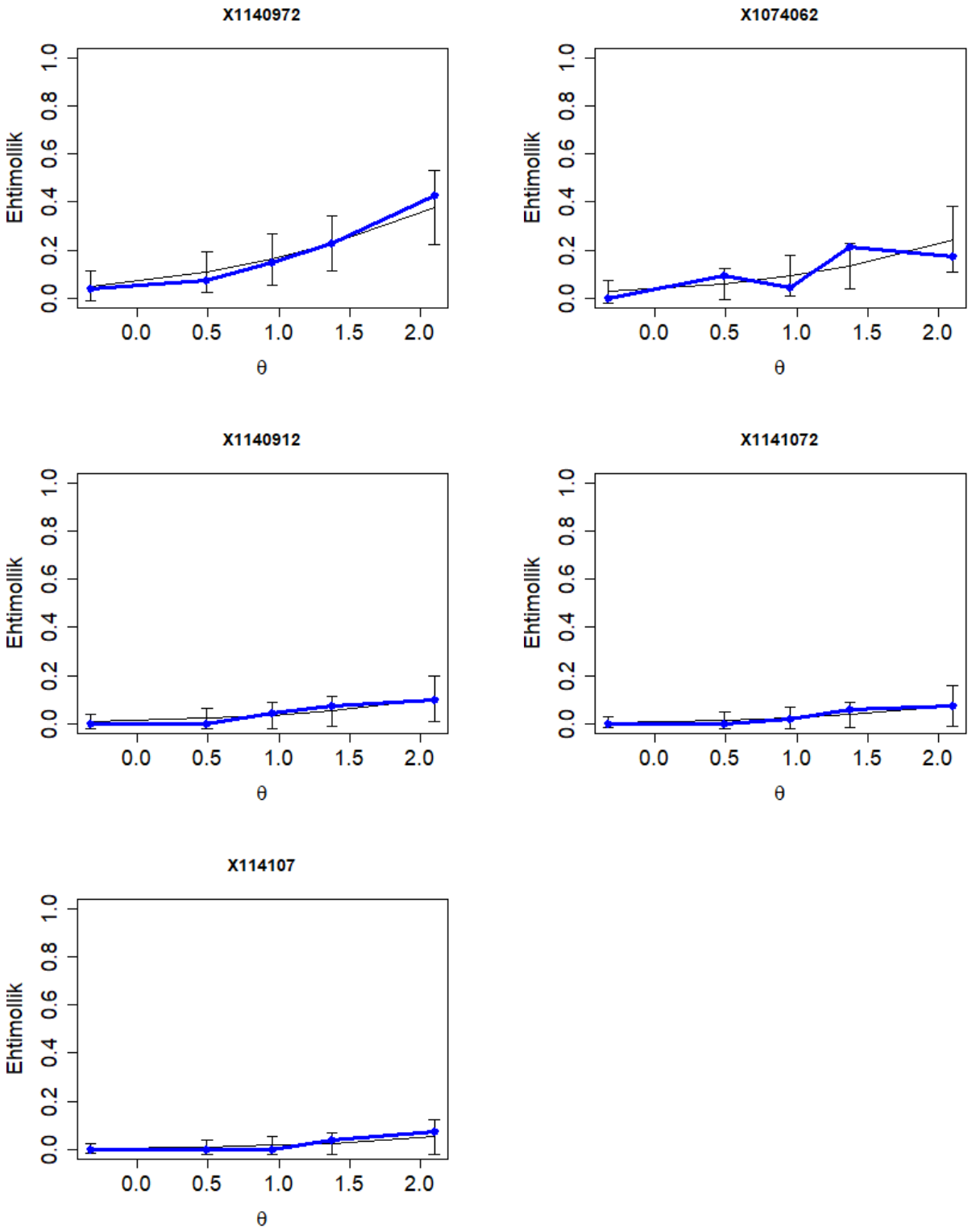

X1141222
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ


X1141102
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ


X11447434
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ


X1141262
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
θ


X107311
X113955
X107316
X112586
X114097
X1140992
Ehtimollik
θ
0.0
0.2
0.4
0.6
0.8
1.0
0.5
1.5
2.0



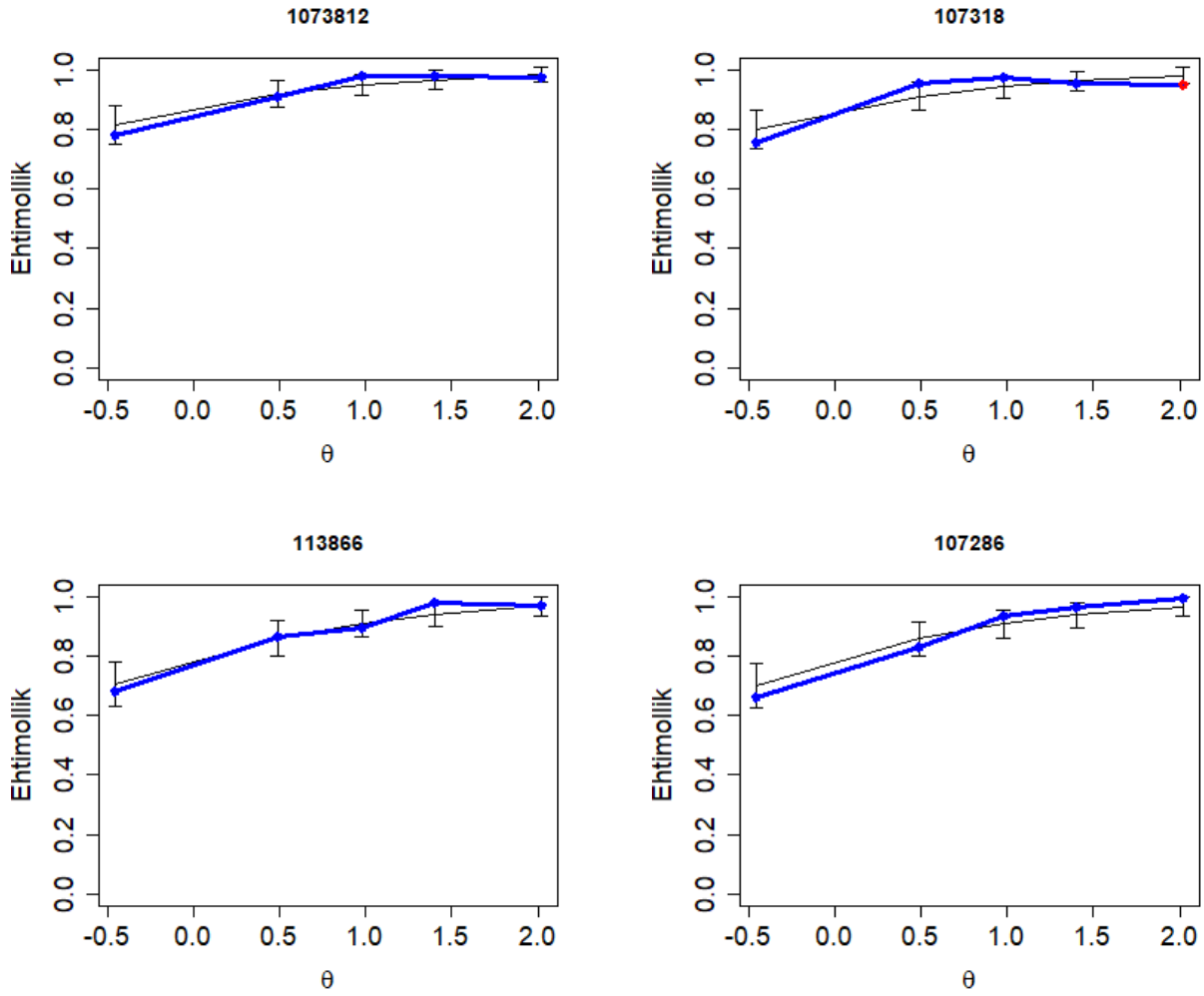
5*b*-rasm. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 2-test sinovi natijalarining Rash modeli bilan mosligi

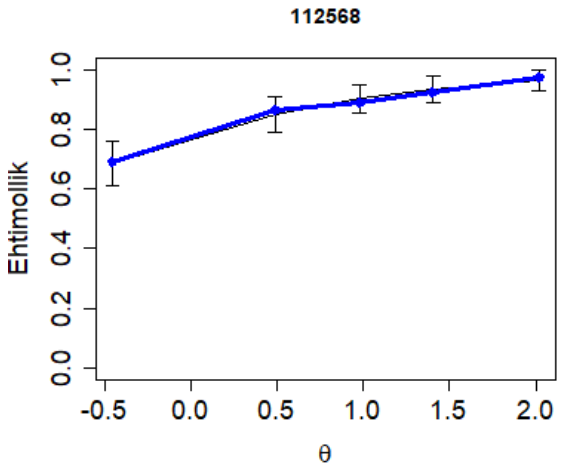
3-test sinovi tahlil natijalariga ko‘ra, ID raqamlari - 107402 va 112560 bo‘lgan test topshiriqlarining Rash modeli bilan moslik darajasi yaxshi emasligini ya’ni bunda ham ajratilgan qobiliyat guruhlarining barchasi bilan

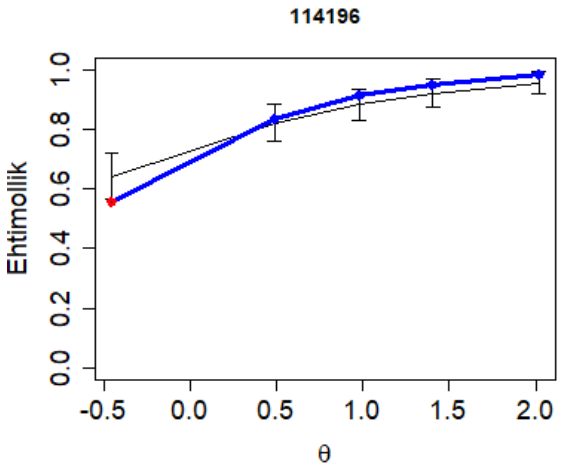
mos tushmaganligini bildiradi (5*c*-rasm).

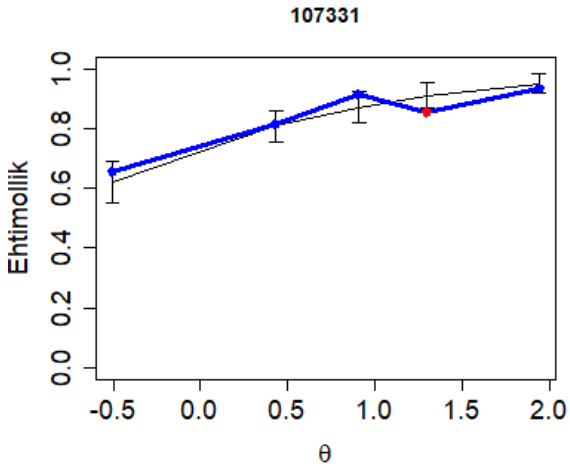
Geografiya fanidan yuqori sifatli kalibrovkalangan test topshiriqlari bazasini yaratishda yuqorida keltirilgan Rash modeli bilan mos kelmagan test topshiriqlarini kalibirovkalangan test bazasiga qo‘shishdan oldin qayta ko‘rib chiqish lozim.

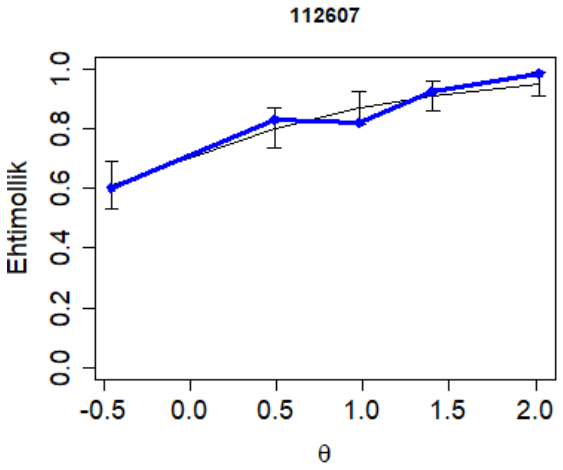
Umumiy holda test topshiriqlari xususiyatlarini Rash modelga moslashtirish muammoli elementlarni va o‘ziga xos xususiyatlarga ega bo‘lgan qobiliyatlarni ajratish imkonini beradi.

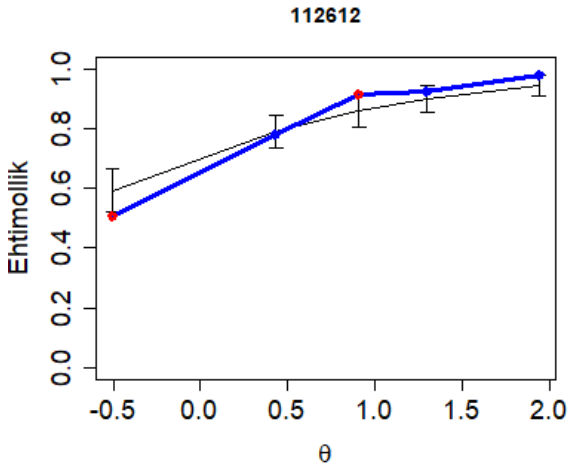


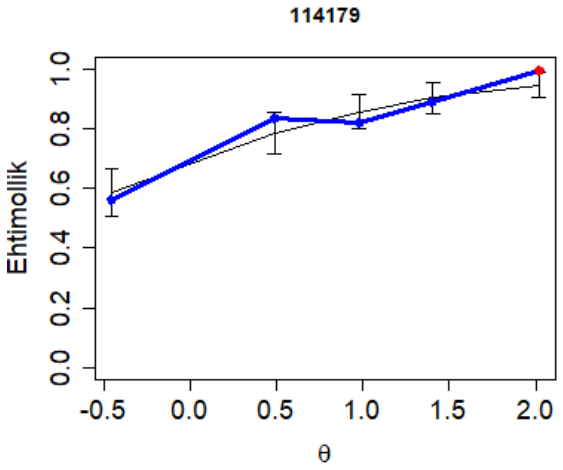

112568
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

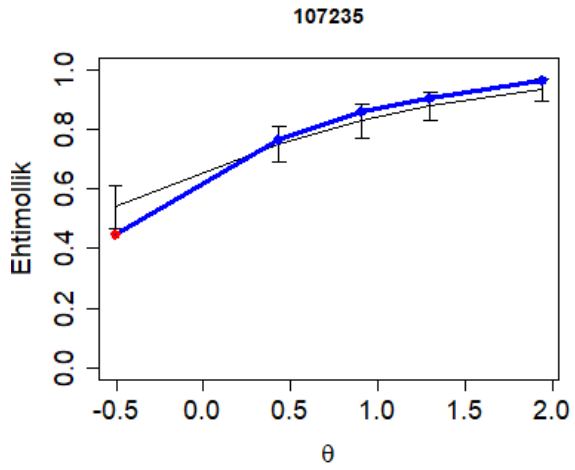

114196
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

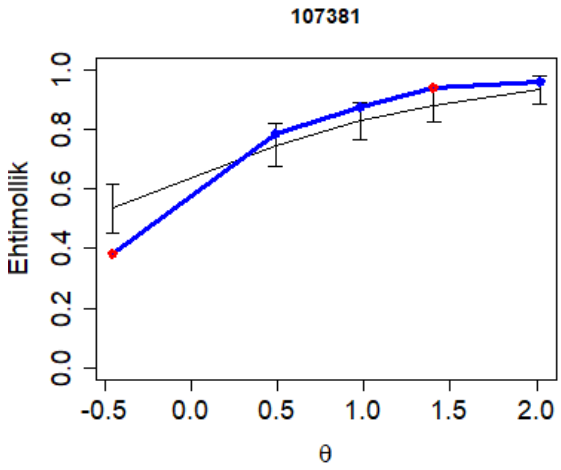

107331
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

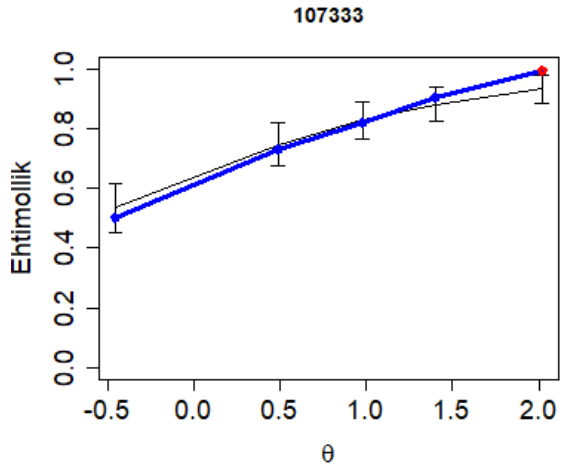

112607
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

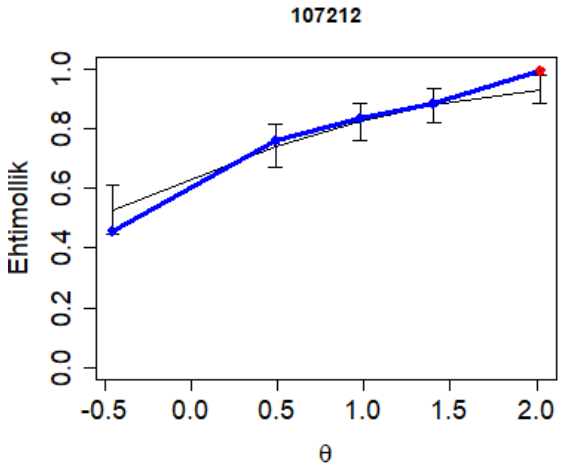

112612
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

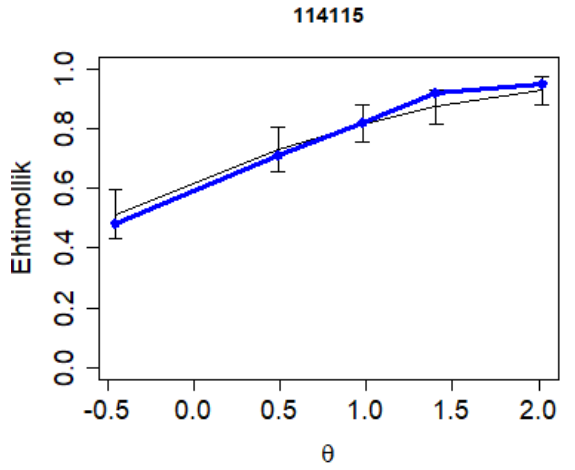

114179
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

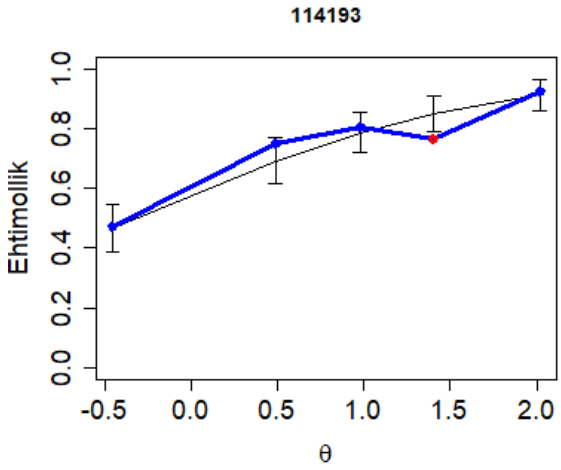

107235
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

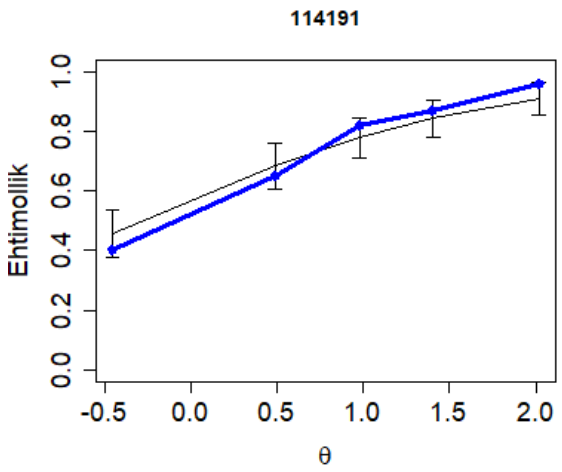

107381
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

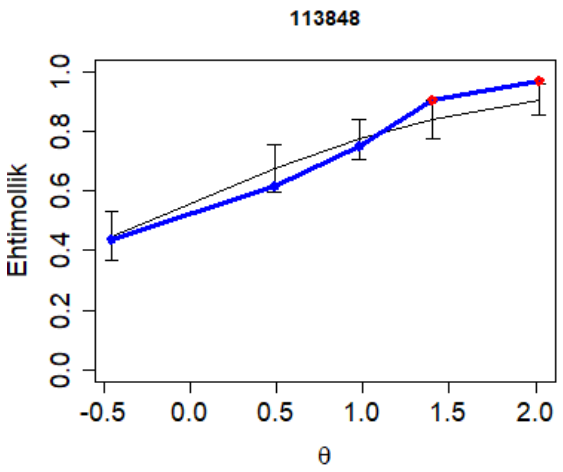

107333
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

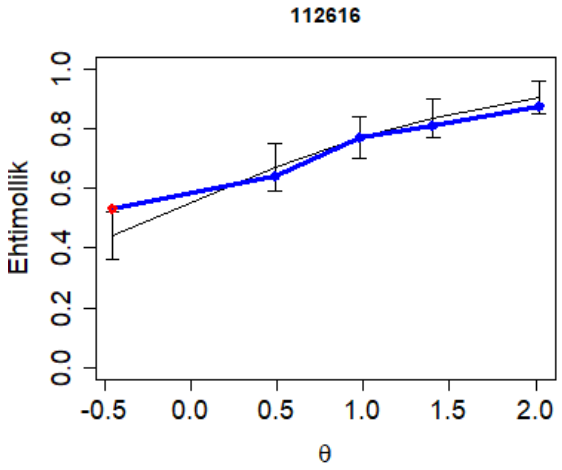

107212
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

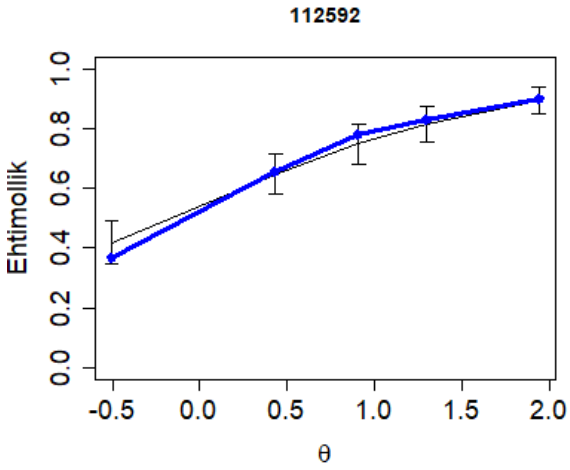

114115
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

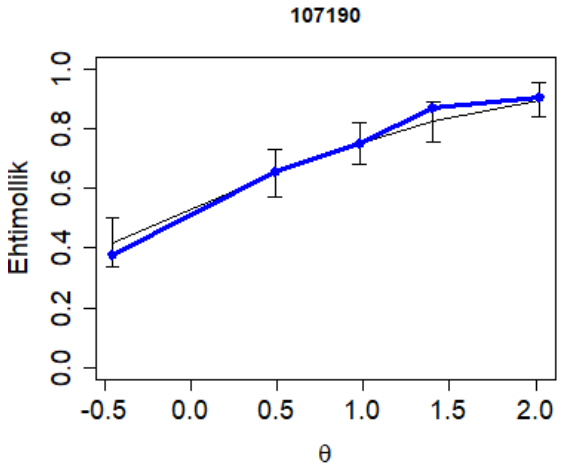

114193
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

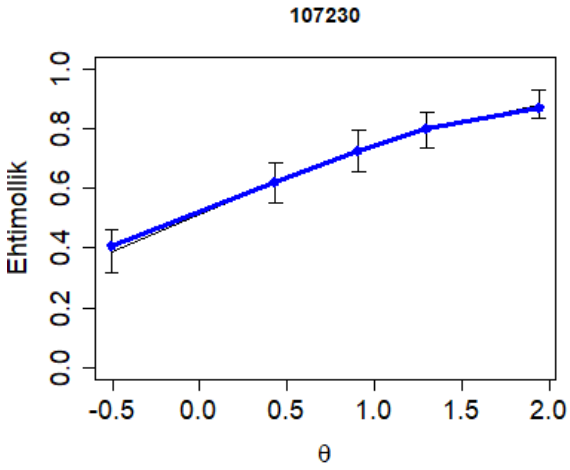

114191
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

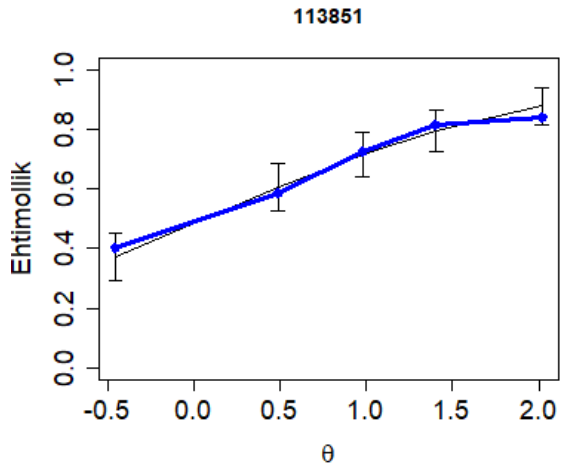

113848
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

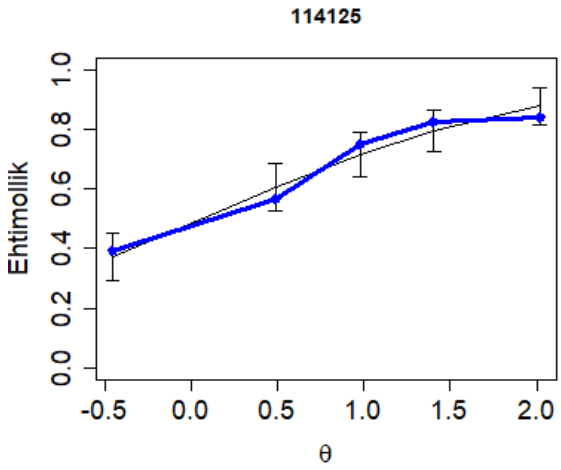

112616
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

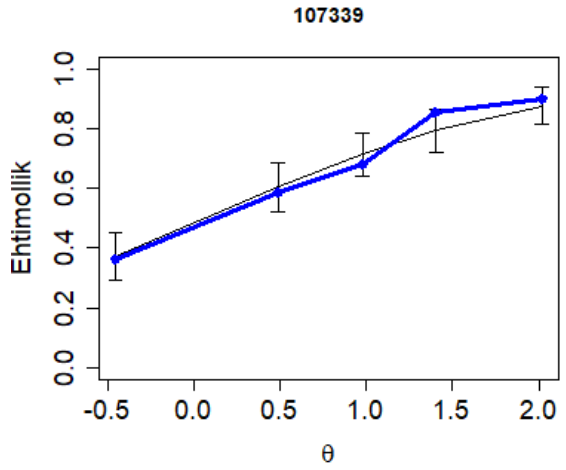

112592
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

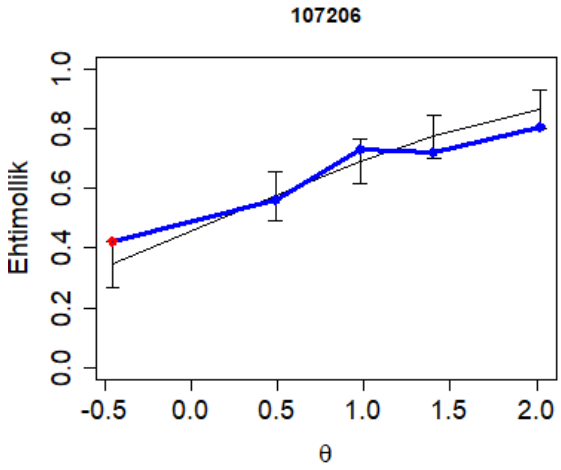

107190
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

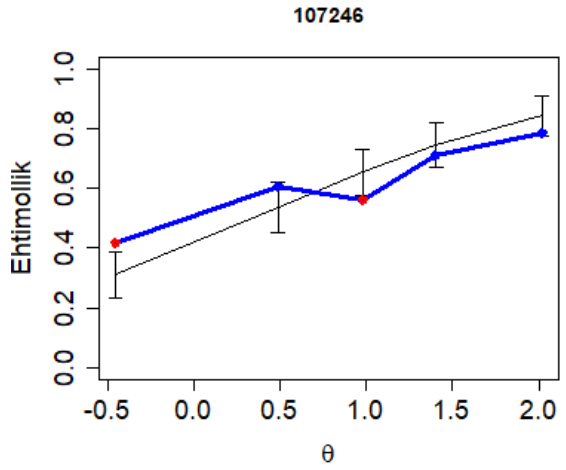

107230
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

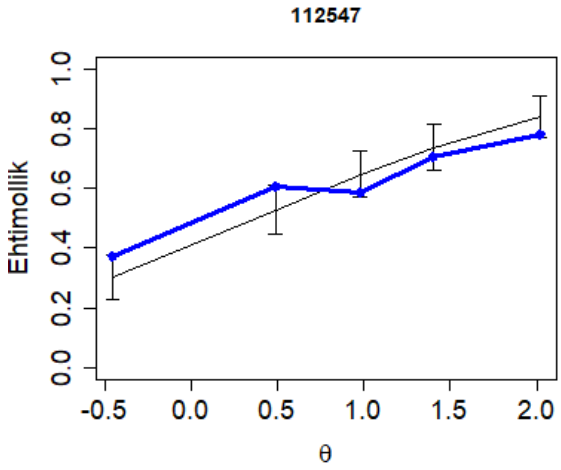

113851
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

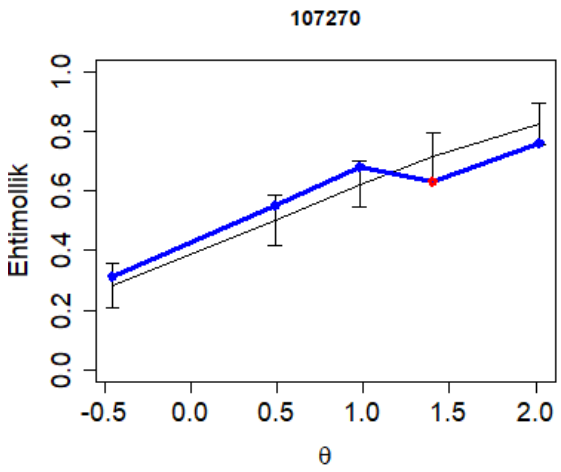

114125
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

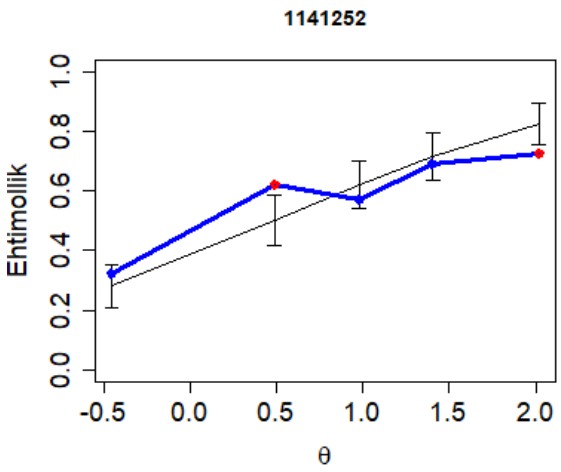

107339
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

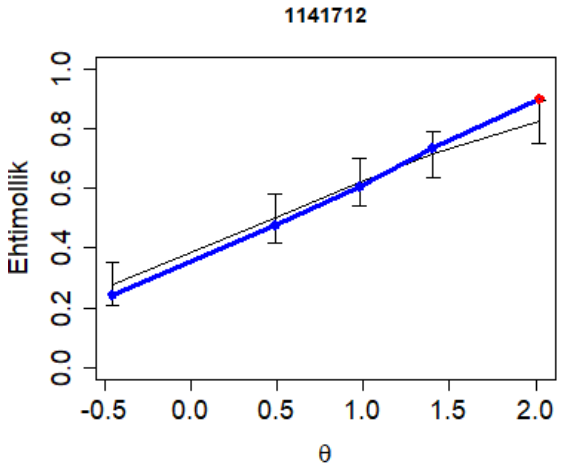

107206
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

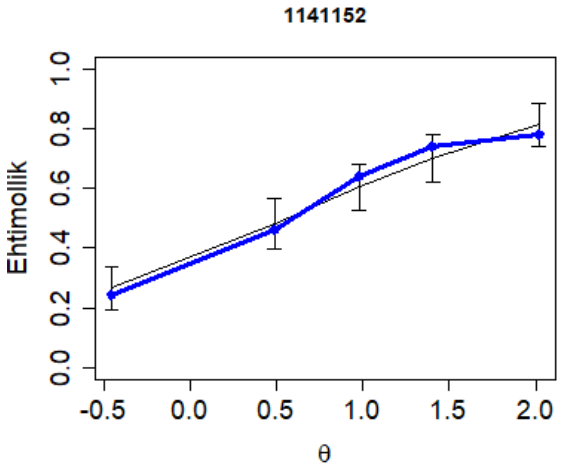

107246
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

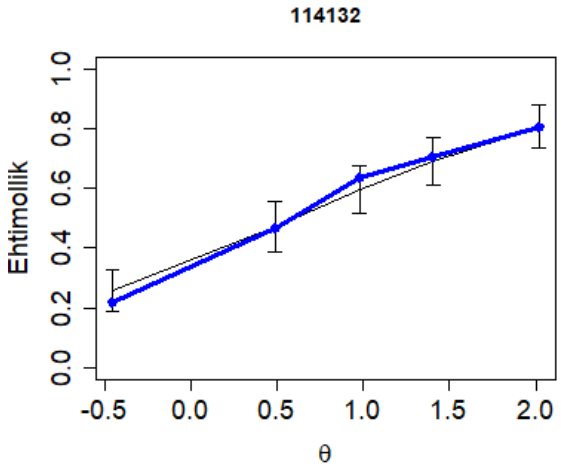

112547
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

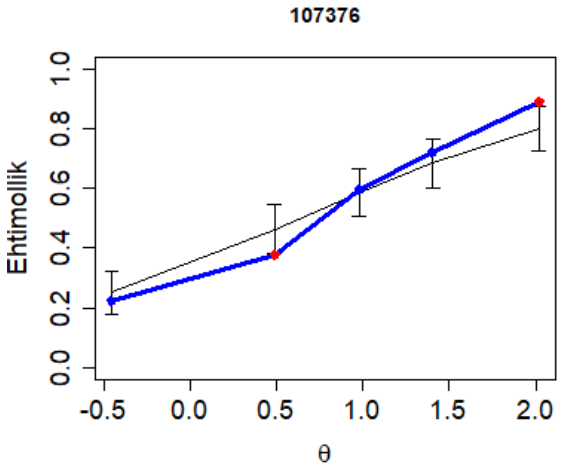

107270
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

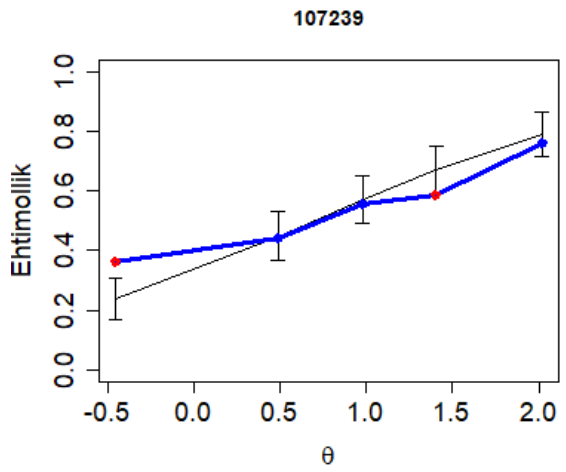

1141252
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

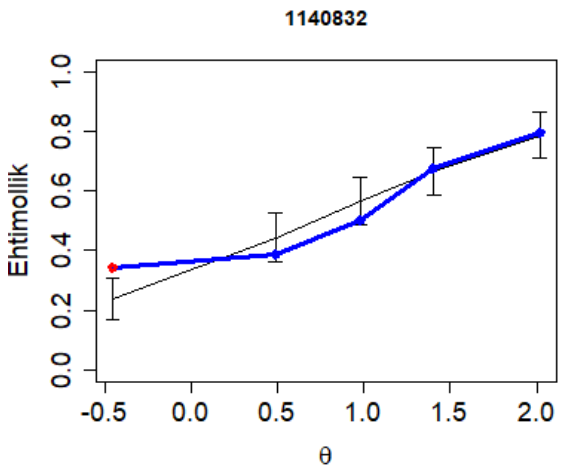

1141712
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

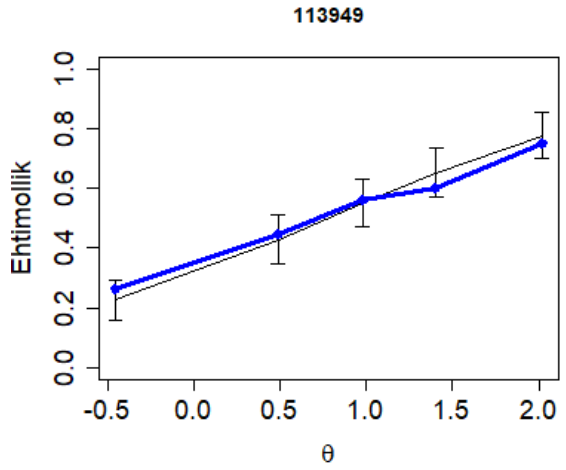

1141152
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

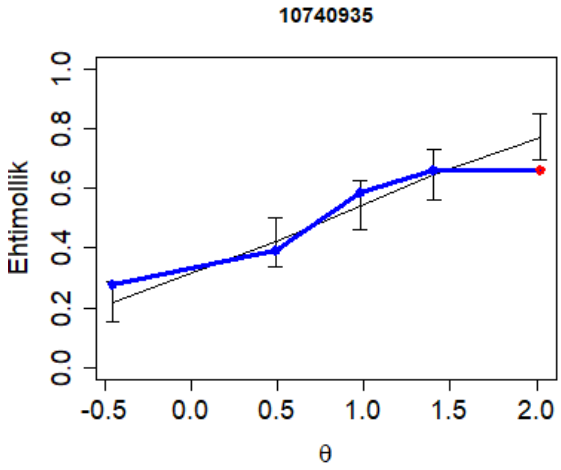

114132
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

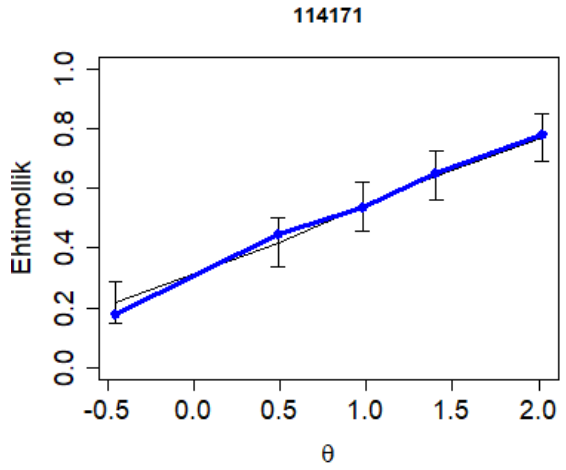

107376
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

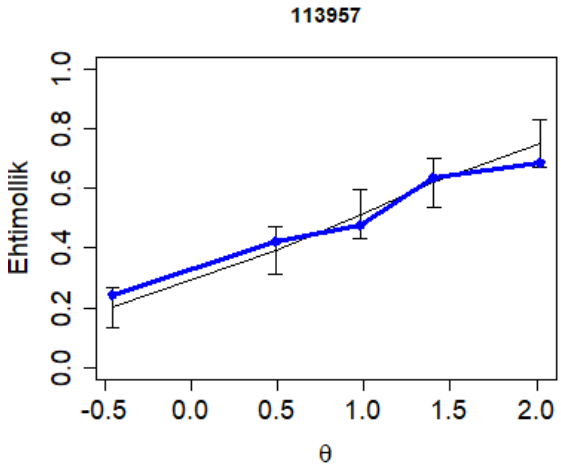

107239
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

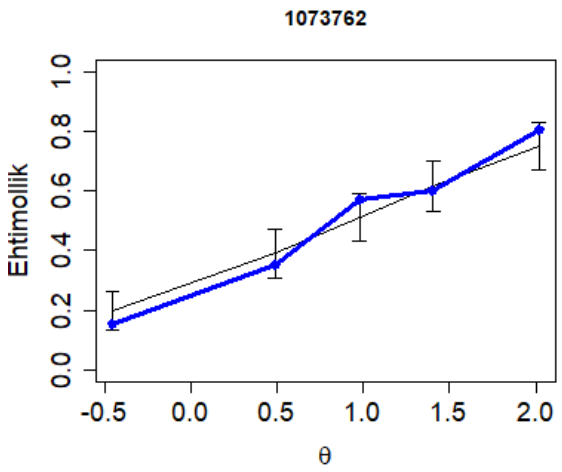

1140832
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

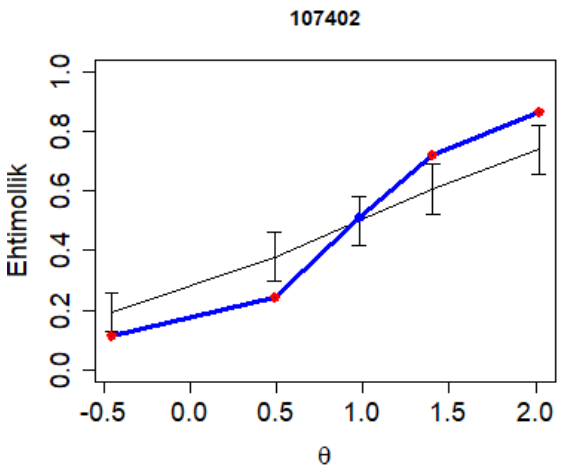

113949
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

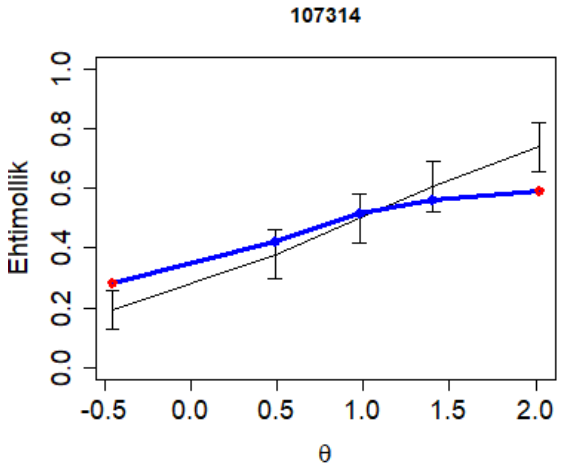

10740935
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

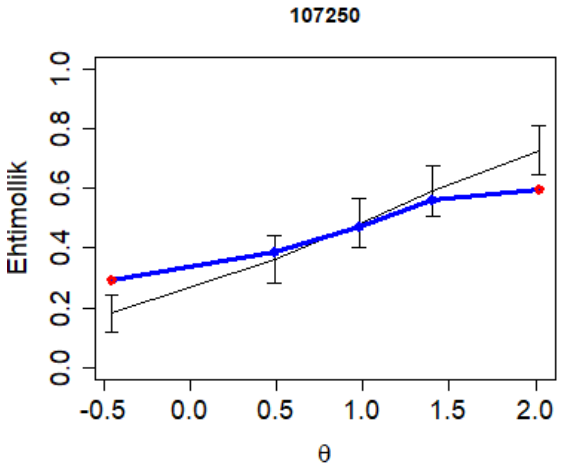

114171
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

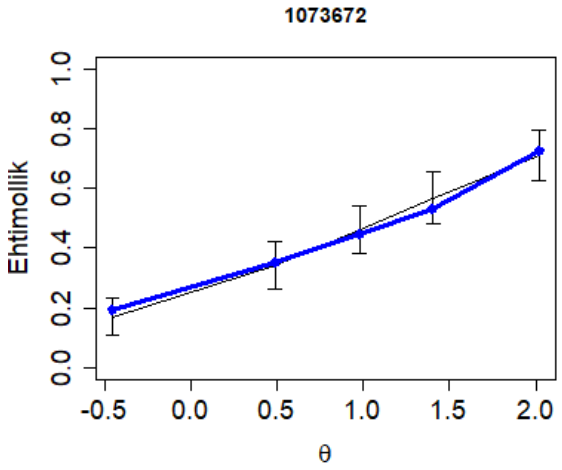

113957
Ehtimollik
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0
θ

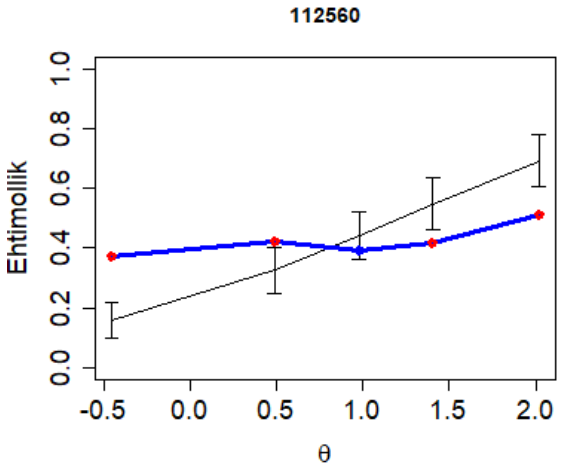

1073762
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

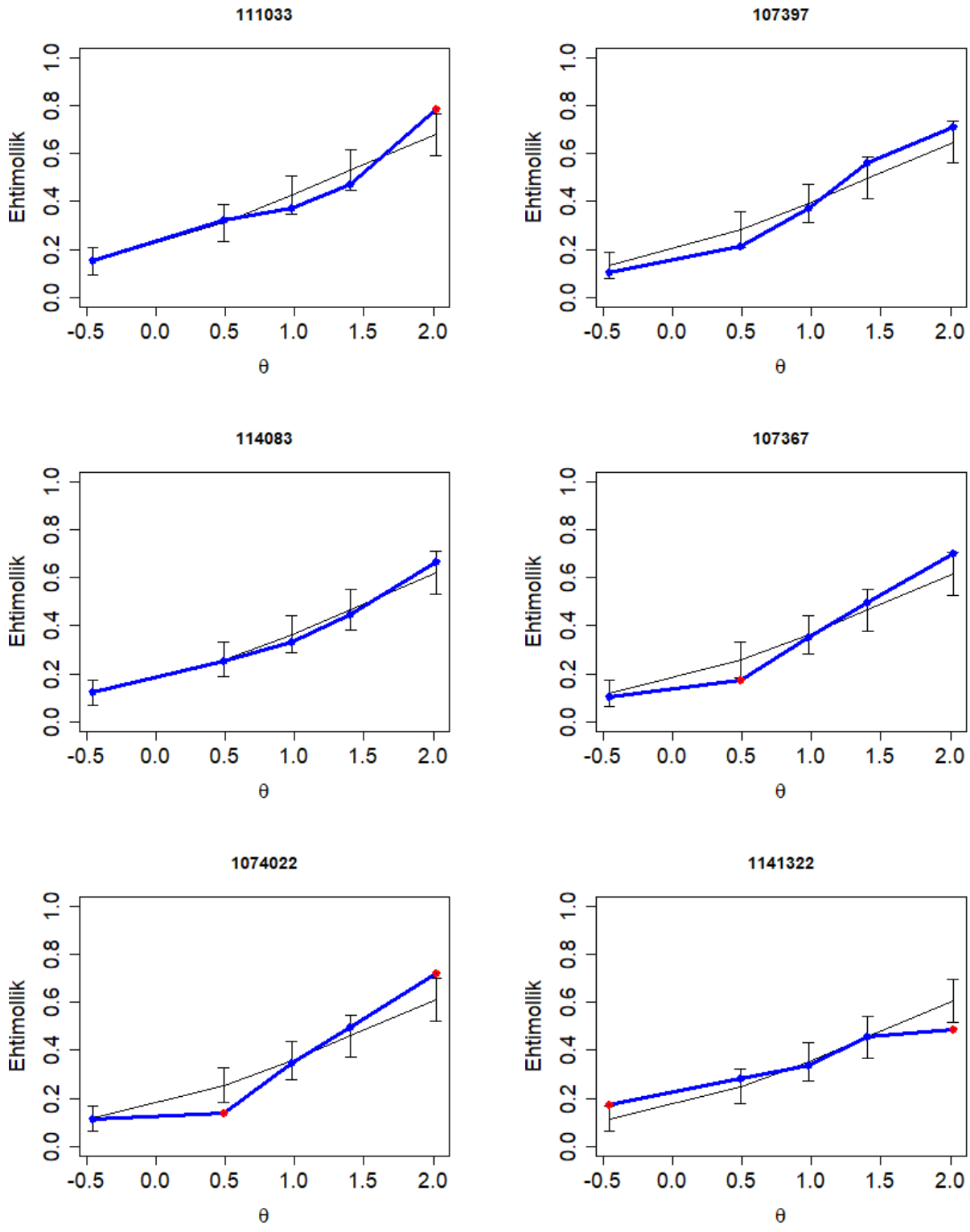

107402
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ

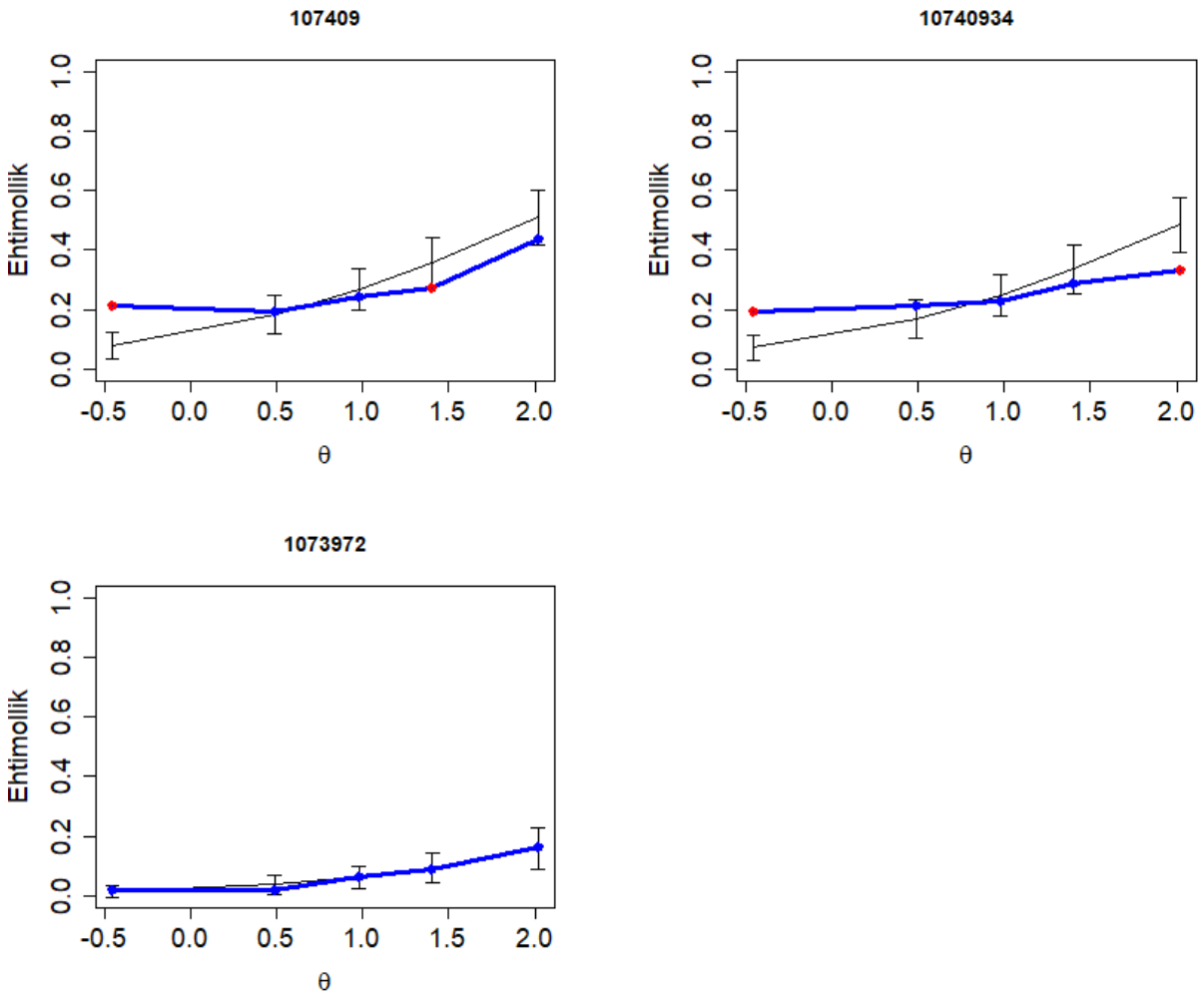

107314
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ


107250
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ


1073672
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ


112560
Ehtimollik
0.0 0.2 0.4 0.6 0.8 1.0
-0.5 0.0 0.5 1.0 1.5 2.0
θ


111033
107397
114083
107367
1074022
1141322
Ehtimollik
θ
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
2.0



5*c*-rasm. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 3-test sinovi natijalarining Rash modeli bilan mosligi

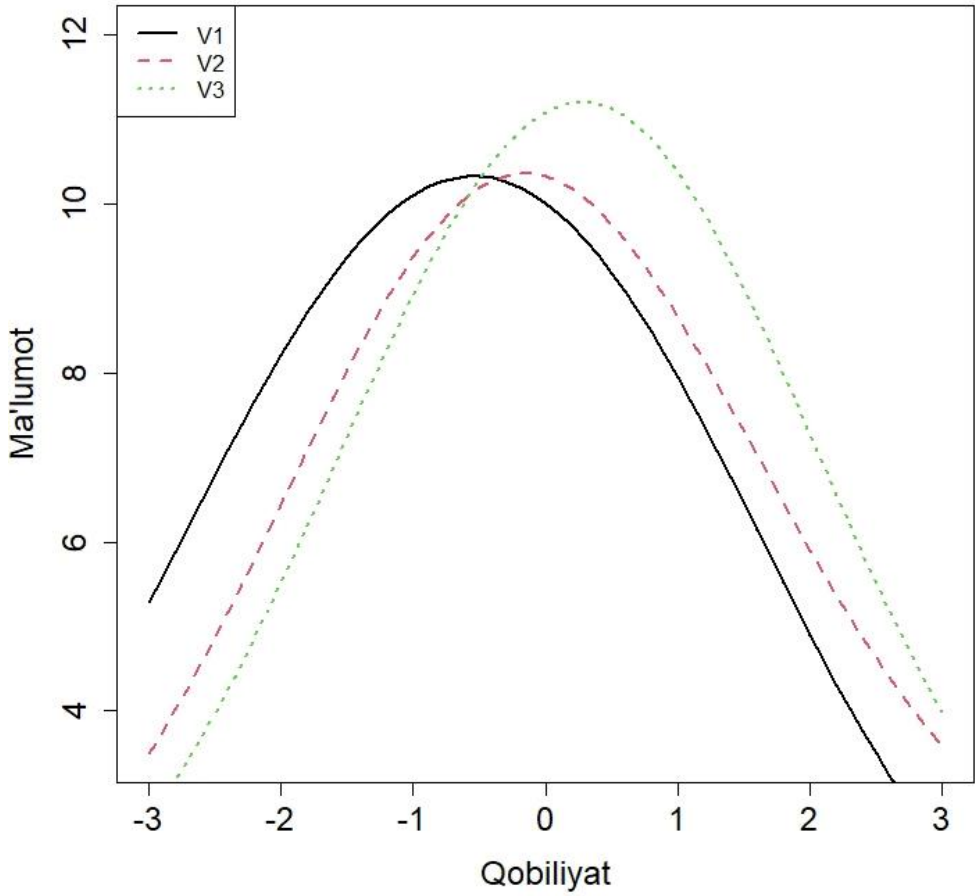
6-rasmda geografiya fanidan o‘tkazilgan 1-, 2- va 3-test sinovi natijalari bo‘yicha test ma’lumoti chiziqlari (TMCh) keltirilgan.

Test ma’lumoti chizig‘i cho‘q-qisining nolga nisbatan chap tomonga biroz surilganligi ushbu test varianti qobiliyat darajasi past bo‘lgan talab-gorlar to‘g‘risida va aksincha, ma’lumot chizig‘i cho‘qqisining nolga nisbatan o‘ng tomonga surilganligi ushbu test varianti qobiliyat darajasi yuqori bo‘lgan talabgorlar to‘g‘risida ko‘proq ma’lumot beradi. Test ma’lumoti chizig‘i cho‘qqisi nol qiymatga to‘g‘ri kelishi yoki juda yaqinligi esa mazkur test varianti qobiliyat darajasi past va yuqori bo‘lgan talabgorlar to‘g‘risida teng ma’lumot beradi.

5-rasmda ko‘rsatilgan test ma’lumoti chiziqlariga ko‘ra, geo-grafiya fanidan o‘tkazilgan 1- va 2-test sinovi natijalari ma’lumot chiziqlari cho‘qqilarining nolga nisbatan chap tomonga (1-test sinovi ko‘proq, 2-test sinovi bo‘yicha esa

biroz) surilganligi ushbu test variantlari orqali qobiliyat darajasi past bo‘lgan talabgorlar haqida ko‘proq ma‘lumot olinganligini bildiradi.

3-test sinovi ma’lumot chiziqlari cho‘qqisining nolga nisbatan o‘ng tomonga surilganligi esa, ushbu test varianti orqali qobiliyat darajasi yuqori bo‘lgan talabgorlar haqida ko‘proq ma‘lumot olinganligini bildiradi.



6-rasm. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 1-(V1), 2-(V2) va 3-(V3) test sinovi natijalarining test ma’lumoti chiziqlari

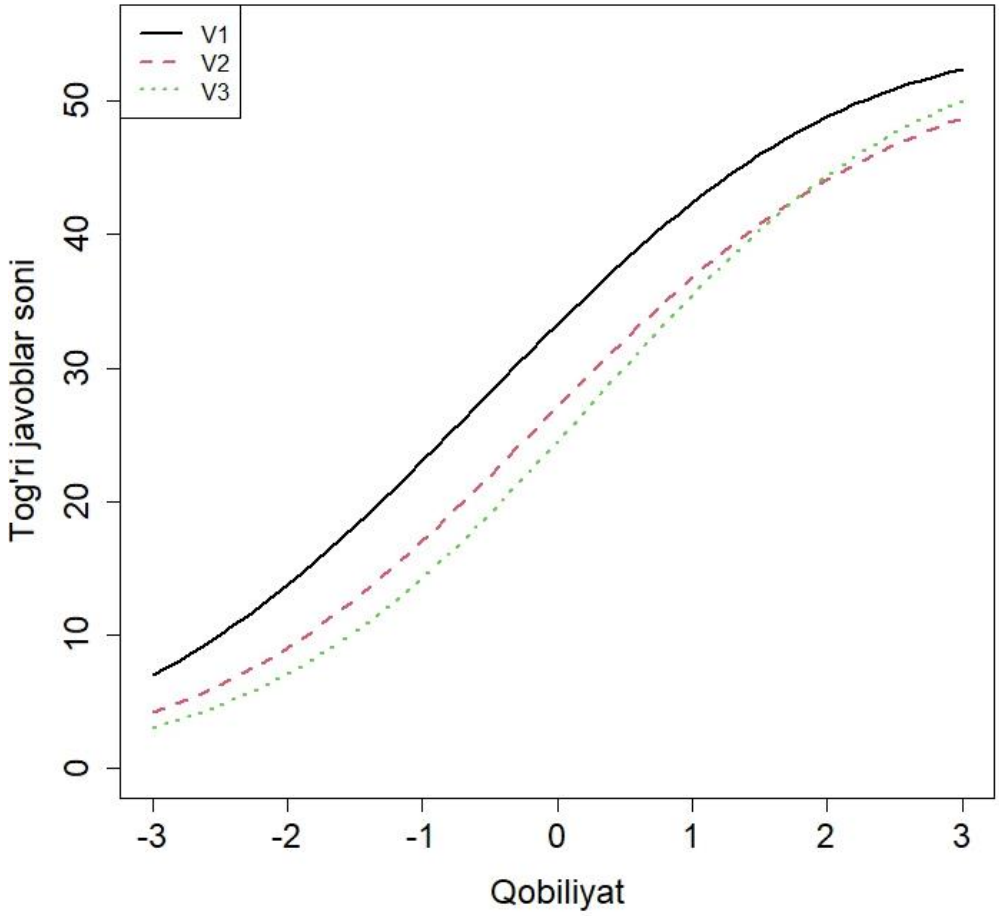
7-rasmda geografiya fanidan o‘tkazilgan 1-, 2- va 3-test sinovlari natijalari bo‘yicha test xarakteristikasi chiziqlari keltirilgan.

7-rasmdagi test xarakteristikasi chizig‘idan to‘g‘ri javoblar sonining qobiliyat darajalari oshishi bilan ortib borishi ko‘rinadi.

Test xarakteristikasi chizig‘idan 1-test sinovida ishlatilgan umumiy elementli test varianti (V1) va 2-test sinovida foydalanilgan umumiy elementli test varianti (V2) hamda 3-test sinovida ishlatilgan umumiy elementli test varianti (V3) ning qiyinlik darajalarini bir-biri bilan

solishtirish mumkin. Bunga ko‘ra, 1-test sinovida ishlatilgan umumiy elementli test varianti (V1) 2- va 3-test sinovi ishlatilgan umumiy elementli test varianti (V2, V3) ga nisbatan osonroq, 3-test sinovida ishlatilgan umumiy elementli test varianti (V3) esa 1-test va 2-test sinovida foydalanilgan umumiy elementli test varianti (V1, V2) ga nisbatan qiyinroq ekanligi ko‘rinadi.

7-rasmdan har bir variantning qiyinlik darajalari hisobga olingan holda qobiliyatlar aniqlayotganligini ko‘rish mumkin. Bu esa variantlarni qiyinligini hisobga olgan holda talabgorlarning qobiliyatlarni aniq-lashda Rash modelini klassik test nazariyasiga nisbatan afzalligini bil-diradi. Biroq, klassik test nazariyasi tahlil xulosalarini yanada to‘ldirishini yodda tutish lozim. Shuning uchun ham odatda, test topshiriqlarining klassik test nazariyasi bilan aniqlangan xususiyatlari Rash modeli natijalari bilan birgalikda beriladi.



7-rasm. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 1-(V1), 2-(V2) va 3-(V3) test sinovi natijalarining test xarakteristikasi chiziqlari

Xulosa

Geografiya fanidan milliy sertifikat uchun o‘tkazilgan 1-, 2- va 3-test sinovlarida foydalanilgan test topshiriqlarining ishonchlilik koeffitsiyenti (Kronbax alfa koeffitsiyenti) mos ravishda 0,84, 0,84 va 0,87 ga teng ekanligi, ishonchlilik koeffitsiyentiga qo‘yilgan mezon talabi bo‘yicha test topshiriqlarining o‘zaro ichki muvofiqligi har uchala test varianti bo‘yicha “yaxshi” darajada ekanligi ko‘rsatib berildi.

Har bitta test topshirig‘i bilan umumiy ball korrelyatsiya koeffitsiyenti qiymatlari 1- va 3-qiyinlik darajasidagi test topshiriqlari uchun 0,25 dan kichik bo‘lgan test topshiriqlarini o‘rganib chiqib kerakli o‘zgarishlar qilish, lozim bo‘lsa, variantdan chiqarib tashlash maqsadga muvofiq bo‘ladi. Har bitta test topshirig‘i bilan umumiy ball korrelyatsiya koeffitsiyenti qiymatlari 2-qiyinlik darajasidagi test topshiriqlari uchun 0,5 dan kichik bo‘lgan test topshiriqlari esa o‘rganib chiqib qiyinlik darajalarini hisobga olgan holda kerakli o‘zgarishlar qilish test topshiriqlarining sifatini yaxshilash uchun xizmat qiladi.

Qiyinlik darajasi bo‘yicha $(-\mathbf{3}:\mathbf{3})$ logit birligi oralig‘idan tashqarida joylashgan test topshiriqlaridan kam miqdorda ma’lumot olinishi sababli bunday test topshiriqlarining o‘rniga $(-\mathbf{3}:\mathbf{3})$ logit birligi oralig‘iga to‘g‘ri keladigan qiyinlikdagi test topshiriqlaridan qo‘yish maqsadga muvofiq. Biroq bunday test topshiriqlari agarda talabgorlar orasida shu test topshiriqlariga mos qobiliyatli talabgorlar bo‘lsa, shunday test topshiriqlarining mavjudligi shu qobiliyatli talabgorlar to‘g‘risida ma’lumot beradi. Aks holda qiyinlik darajasi juda past va juda yuqori bo‘lgan test topshiriqlarining o‘rniga $(-\mathbf{3}:\mathbf{3})$ logit birligi atrofidagi test topshiriqlaridan kiritish, yuqori va past qobiliyat darajalaridan olinadigan ma‘lumot miqdori orasidagi tafovutni yanada kamaytirish imkonini beradi.

Test topshiriqlari xususiyatlarini Rash modeliga moslashtirish muammoli elementlarni va o‘ziga xos xususiyatlarga ega bo‘lgan qobiliyatlarni ajratish imkonini beradi. Geografiya fanidan milliy sertifikat uchun o‘tkazilgan test sinovlarida foydalanilgan test variantlarida 1-test sinovi bo‘yicha (ID raqamlari – 101155, 101166, 101412 va 101590) bo‘lgan, 2-test sinovi bo‘yicha (ID raqamlari - 107322 va 11447434) bo‘lgan va 3-test sinovi bo‘yicha (ID raqamlari - 107402 va 112560) bo‘lgan test topshiriqlarining Rash modeli bilan moslik darajasi mos tushmaganligi, ya’ni ajratilgan qobiliyat guruhlarining barchasi bilan mos tushmaganligi aniqlandi. Bunday test topshiriqlarini

ham kalibirovkalangan test bazasiga qo‘shishdan avval mutaxassislar tomonidan qayta ko‘rib chiqilishi lozim bo‘ladi.

ADABIYOTLAR

1. Crocker, L. and Algina, J. (1986) Introduction to Classical and Modern Test Theory. Harcourt, New York, 527.
2. Allen, M. J., & Yen, W. M. (1979). Introduction to Measurement Theory. Monterey, CA: Brooks/Cole Publishing Company.
3. Hambleton, R. K., Swaminathan, H., and Rogers, H.J. (1991), Fundamentals of item response theory. Newbury Park, CA: Sage
4. Ivailo Partchev (2004), A visual guide to item response theory, Friedrich-Schiller-Universitat¨ Jena
5. Gunter Maris, Timo Bechger, Jesse Koops and Ivailo Parchev, Data Management and Analysis of Tests, p. 1-49, 2022.
6. D. M. Alimov, M. Dj. Ermamatov, A.R.Sattiyev, Z.Z.Mirvaliyev, A.A. Sulaymonov “Matematika fanidan test sinovi natijalarini klassik test nazariyasi asosida statistik tahlili”, Ajiniyoz nomidagi Nukus Davlat pedagogika institutining “Fan va jamiyat” ilmiy-uslubiy jurnali, Nukus, 2021-yil, №1, 75-77 betlar.
7. D.M.Alimov, M.Dj.Ermamatov, A.R.Sattiyev, Z.Z.Mirvaliyev, A.A.Sulaymonov “Matematika fanidan test sinovi natijalarining zamonaviy test nazariyasi asosida statistik tahlili”, Ilm sarchashmalari jurnali, Urganch, 2021-yil, №6, 116-119 betlar.
8. 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” ilmiy-uslubiy jurnali, 2023-yil, 1-son, 4-62-betlar.
9. A.R. Sattiyev, M. Dj. Ermamatov “Matematika fanidan milliy sertifikat uchun o‘tkazilgan test sinovlari natijalari tahlili”, “Axborotnoma” ilmiy-uslubiy jurnali, 2023-yil, 2-son, 35-55-betlar.
10. A.R. Sattiyev, M. Dj. Ermamatov “Tarix fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalarining klassik test nazariyasi va Rash modeli asosida tahlili”, “Axborotnoma” ilmiy-uslubiy jurnali, 2023-yil, 3-son, 51-72-betlar.
11. A.R. Sattiyev, M. Dj. Ermamatov “Test xarakteristikalarini o‘rganish: milliy sertifikat bo‘yicha test sinovi natijalari”, “Axborotnoma” ilmiy-uslubiy jurnali, 2023-yil, 4-son, 67-96-betlar.
12. A.R. Sattiyev, M. Dj. Ermamatov “Rayt xaritasi, ichki va tashqi moslik statistikalari: kimyo fanidan milliy sertifikat uchun o‘tkazilgan test sinovi natijalari”, “Axborotnoma” ilmiy-uslubiy jurnali, 2024-yil, 1-son, 18-96-betlar.

13. A.R. Sattiyev, M. Dj. Ermamatov, D.M. Alimov “Ona tili va adabiyot fanidan milliy sertifikat uchun o‘tkazilgan test sinovida foydalanilgan test topshiriqlarining distraktorlari va korrelyatsiyalari”, “Axborotnoma” ilmiy-uslubiy jurnali, 2024-yil, 2-son, 4-96-betlar.
14. A.R. Sattiyev, M. Dj. Ermamatov “Milliy sertifikat uchun geografiya fanidan o‘tkazilgan test sinovi natijalarining test xarakteristikalari”, “Axborotnoma” ilmiy-uslubiy jurnali, 2024-yil, 4-son, 14-52-betlar.
15. Spearman, C. 1907. Demonstration of formulae for true measurement of correlation. Am. J. of Psychology. 18, 160-169.
16. Spearman, C. 1910. Correlation from faulty data, British J. of Psychology. 3, 271-295.
17. Аванесов В.С. Теория и методика педагогических измерений. ЦТ и МКО УГТУ-УПУ, 2005
18. Челышкова М.Б. Теория и практика конструирования педагогических тестов: учеб. Пособие М.: Логос, 2002, с. 432.
19. https://ru.wikipedia.org/wiki/%D0%90%D0%BB%D1%8C%D1%84%D0%B0_%D0%9A%D1%80%D0%BE%D0%BD%D0%B1%D0%B0%D1%85%D0%B0.
20. М.Дж. Эрмаматов Замонавий тест назарияси асослари: Услубий қўлланма., Тошкент, 2020 йил, 68 б.
21. B.D. Wright and M.H. Stone, Best Test Design (MESA Press, Chicago, 1979).
22. David Torres and Rebecca Freund, Package WrightMap, 2024.

TEST CHARACTERISTICS: TESTS ON GEOGRAPHY

A.R. Sattiev, M.Dj. Ermamatov

Scientific-study Practical Center under the Agency for Assessment of Knowledge and Competences, Tashkent 100084, Bogishamol st. 12.

Abstract. This article analyzes the results on the geography tests conducted in 2024 for the national certification based on Classical Test Theory (CTT) and the Rasch model. The descriptive statistics of the test results and the correlation between individual test items and the total score are discussed. The difficulty levels of the test items used in three test administrations conducted using CTT and the Rasch model have been analyzed. Wright maps were generated for each test form based on the ability and difficulty levels identified using the Rasch model. Additionally, test data and characteristic curves, as well as the fit statistics of each test item within the test forms according to the Rasch model, were examined.

Keywords: Test items, Cronbach's alpha, difficulty, correlation coefficient, Rasch model, Wright-map, ability, fit to the Rasch model.