

QIYINLIK VA QOBILIYAT DARAJALARI INVARIANTLIGI: FIZIKA FANIDAN TEST NATIJALARI TAHLILI

A.B. Normurodov, M.Dj. Ermamatov, T.X. Jumayev

*Bilim va malakalarni baholash agentligi huzuridagi Ilmiy-o'quv amaliy
markazi,*

tursunboyjumayev@gmail.com

Qisqacha mazmuni. Ushbu maqolada fizika fanidan sinaluvchilarning bilimlarini baholashda standart testlardan foydalanish orqali turli xil test topshiruvchilar guruhlarining natijalarini solishtirish, ularning fizika fanidan o'quv dasturini qay darajada o'zlashtirganligini baholash o'rganilgan. Turli xil guruhlarining qobiliyat darajalari va test variantidagi test topshiriqlarining qiyinlik darajalarini bir xil shkalada bo'lishini ta'minlash muhimligi test sinovi natijalarini Rash modeli bilan tahlil qilish orqali o'rganilgan. Rash modeli bilan baholash ya'ni turli xil guruhlarining qobiliyat darajalari va turli xil test variantlaridagi test topshiriqlarining qiyinlik darajalarini bir xil shkalada bo'lishi orqali reyting o'rinlarni aniqlash va bu haqida xulosalar chiqarish mumkinligi ko'rsatildi.

Kalit so'zlar: Standart test, Rash modeli, qobiliyat darajasi, qiyinlik darajasi, Rayt xaritasi.

1. Kirish

Bugungi kunga kelib ko'plab mamlakatlarda bilimlarni va ta'lim muassasalari hamda pedagoglar faoliyati samaradorligini baholashda standartlashtirilgan testlardan keng ko'lamda foydalanilmoqda [1]. Standartlashtirilgan testlar bu muvofiq-
lashtirilgan yoki "standart" tarzda o'tkaziladigan va baholanadigan testlardir [2]. Standartlashtirilgan test sinovlari natijalari bilan turli xil guruhlardagi sinaluvchilar bilimlarini solishtirish, ularning o'quv dasturini qay darajada o'zlashtirganligini aniqlash, pedagoglar va ta'lim muassasalari

faoliyati samaradorligini baholash, shuningdek, ularning yutuq va kamchilliklari haqida ma'lumotlar olish mumkin bo'ladi. Standart testlar o'quvchilar, ota-onalar, pedagoglar va ta'lim tizimining boshqa ishtirokchilariga ham o'ziga xos ta'sir ko'rsatadi. Xususan, standartlashtirilgan testlar o'quvchilarning natijalari asosida ularning o'zlashtirishlarini nazorat qilishga va shaxs sifatida o'z qobiliyatlariga ko'ra shakllanishiga hissa qo'shadi [3]. Lekin standartlashtirilgan testlardan to'g'ri foydalanish uchun pedagoglar ham

baholash bo'yicha bilimlarga ega bo'lishi talab etiladi [4]. Shuning bilan birga talabgorlarning haqiqiy bilimni baholash uchun standart test variantlarini yaratishda test topshiriqlari sifatini aniqlash, aprotatsiya test sinovini o'tkazish, test sinovlari natijalarini statistik tahlil asosida ilmiy asoslash va test topshiriqlari tavsiflarini yaxshilash juda muhim hisoblanadi [5]. Fizika fanidan bilimni baholash uchun foydalaniladigan standart testlar bo'yicha ko'plab fikrlar keltirilgan bo'lsa-da, bugungi kunda mamlakatimizda standartlashtirilgan testlarning ahamiyati kam o'rganilgan. Standartlashtirilgan test bu - test topshiriqlari va o'tkazish sharoitlari barcha sinaluvchilar uchun bir xil va baholash jarayoni asosli ravishda mezonlashtirilganligi sababli test topshiriqlarining sifatini va sinaluvchilarga beriladigan ballarni oldindan aniqlangan tartib asosida talqin qilish imkonini beradigan jarayondir [2,6]. Demak, standartlik faqat test variantiga nisbatan ishlatilmasdan, balki butun jarayonga tegishli bo'ladi. Bunda, yuqoridagi shartlar bajarilganda faqat muqobil javobli test topshiriqlaridan iborat test

varianti emas, balki barcha turdagi test topshiriqlaridan iborat test varianti ham standartlik shartlari doirasida bo'lishi mumkin.

Odatda testlar turli maqsadda ishlab chiqiladi va qo'llaniladi. Bulardan eng muhimlari maqsadga ko'ra me'yorga asoslangan, mezonga mo'ljallangan va bashorat qilish uchun foydalaniladigan testlardir [7]. Shu kabi fizika fanidan milliy sertifikat uchun qo'llaniladigan standart testlar nafaqat sinaluvchilarning bilimlarini baholashni, balki mamlakatimizda ta'lim standarti asosidagi dasturlarni qay darajada o'zlashtirilganligini yoki pedagoglar va ta'lim muassasasi faoliyatining samadorligini ham ko'rsatib beradi.

Ushbu maqolada talabgorlarning haqiqiy bilim darajasini aniqlash, pedagoglar faoliyatining samaradorligi va ta'lim sifatini qay darajada ekanligi haqida to'g'ri xulosa chiqarish uchun standartlashtirilgan testlardan foydalanishning o'ziga xos xususiyatlari haqida bayon etilgan. Bunda ilmiy tadqiqot obyekti sifatida fizika fanidan milliy sertifikat uchun o'tkazilgan test sinovi natijalari olingan.

2. Asosiy qism

Ilmiy tadqiqot fizika fanidan haqiqiy bilimlarini baholash uchun standart test variantini yaratish va foydalanish maqsadida o'tkazildi. Ilmiy tadqiqot obyekti sifatida fizika fanidan

milliy sertifikat uchun o'tkazilgan test natijalari olindi. Fizika fanidan milliy sertifikat uchun o'tkazilgan test sinovlarida foydalanilgan variant 45 ta (36–45-ochiq test topshiriqlarining a

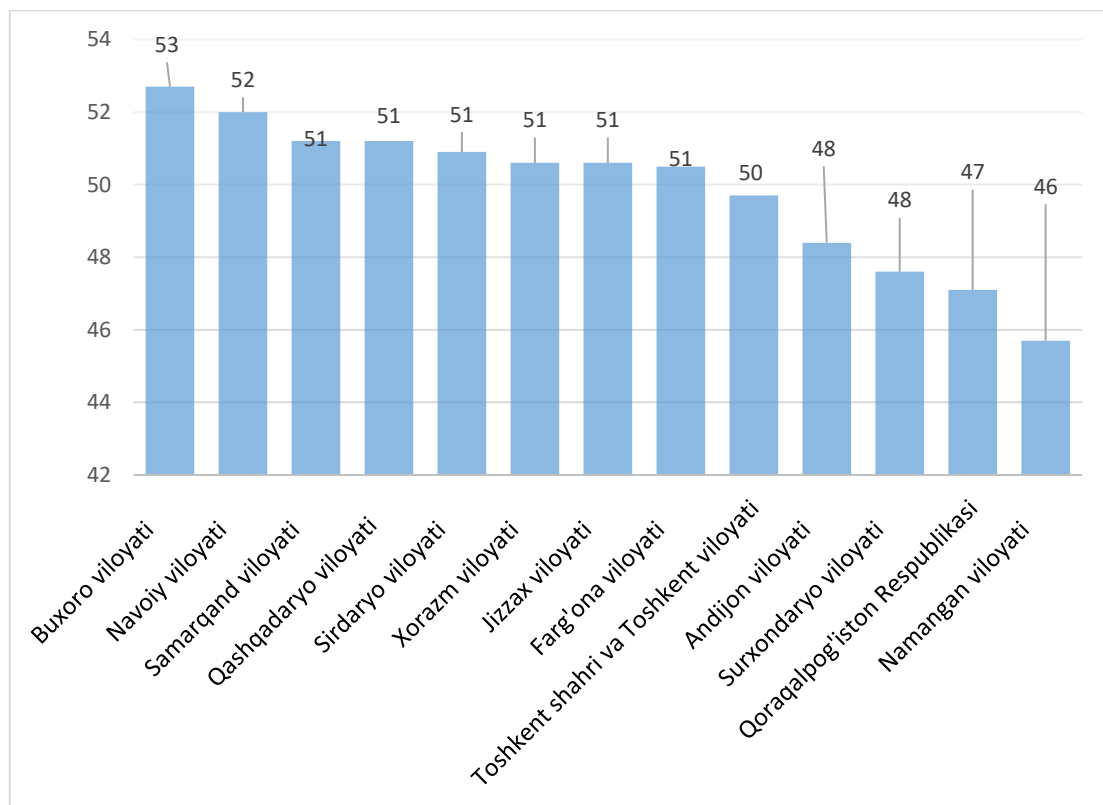
va b qismlarga ajratilishi hisobiga 55 ta) test topshirig'idan iborat bo'lib, ajratilgan vaqt javoblar varaqasini to'ldirish bilan birgalikda, 150 daqiqani tashkil etadi. Test sinovida jami 1141 nafar talabgor qatnashishdi va test sinovi natijalari Rash modeli [8-10] bilan tahlil qilindi. Rash modeliga ko'ra, dixotomik elementlarga individual javoblar ehtimoli shaxsning qobiliyat darajasi va element (test topshirig'i) qiyinligi bilan aniqlandi [11].

1-rasmda talabgorlarning Rash modeli bilan aniqlangan o'rtacha T-ballari hududlar kesimida keltirilgan. Rasmdan barcha hududlar uchun Rash modeliga asosan hisoblangan o'rtacha ballari turli xil qiymatlarga ega: Buxoro viloyatida fizika fanidan talabgorlarning o'rtacha ballari eng yuqori, Namangan viloyatidagi talabgorlarning o'rtacha ballari esa eng past ko'rsatkichga ega ekanligini ko'rish mumkin. Bunda turli xil hududlardagi talabgorlarning fizika fanidan test sinovi natijalarini o'zaro solishtirish orqali ularning reytingi haqida xulosa chiqarish mumkin bo'ladi. Lekin turli xil qobiliyat egalariga (guruhlariga) bir xil test topshiriqlari varianti berilgan bo'lsa-da, hududlar uchun hisoblangan o'rtacha ballar yoki alohida guruhlar uchun ajratib hisoblangan talabgorlar qobiliyat darajalari va test topshiriqlarining qiyinlik darajalarini

solishtirish maqsadga muvofiq emas. Chunki qobiliyat va qiyinlik darajalari bu yerda har bir guruh uchun turli xil shkalada bo'ladi. Ular bir xil shkalada bo'lishi uchun mo'ljallangan guruh (bizning tadqiqotda hududlar) uchun qobiliyat va qiyinlik darajalari birgalikda hisoblanishi lozim. Hozirgi kunda bitta guruhga har xil test varianti berilganda ham qobiliyat va qiyinlik darajalarini bir xil shkalada bo'lishini ta'minlaydigan usullar ishlab chiqilgan. Masalan, shunday usullardan biri [12] maqolada yoritilgan. Keyingi tadqiqotlar maqsadi fizika fanidan test sinovi natijalarining shu usul bilan kengaytirilgan tahlilidir.

Quyida biz talabgorlarning qobiliyat darajalari va test topshiriqlarining qiyinlik darajalarini bir xil shkalada bo'lishini ta'minlash uchun turli xil hududlardagi sinaluvchilarning test sinovi natijalarini Rash modeli bilan birgalikda hisoblab, amalga oshirilgan tahlillarini keltirib o'tamiz.

Qobiliyat va qiyinlik darajalari bir xil shkalada bo'lishi ta'minlangan turli xil hududlardagi talabgorlarning fizika fanidan test sinovi natijalarini solishtirish orqali ular ta'lim olayotgan ta'lim muassasalarining samaradorligini, unda faoliyat yuritayotgan pedagoglar va talabgorlarning bilimlari haqida xulosalar chiqarish mumkin (1-rasm).



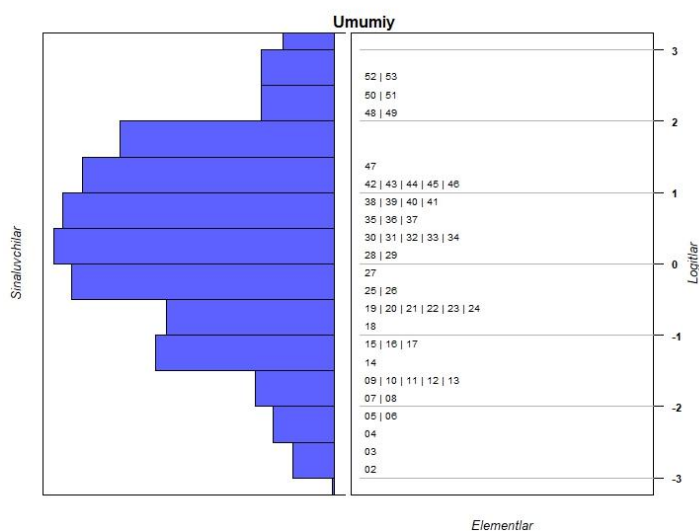
1-rasm. Hududlar kesimida talabgorlarning Rash modeli bilan aniqlangan o'rtacha T-ballari

Rash modeli asosida aniqlangan qiyinlik darajalarini sinaluvchilar qobiliyatlariga qanchalik mosligini Rayt xaritasi yordamida tahlil qilish mumkin [13]. Bunda ham test topshiruvchilarning qobiliyat darajalari va test topshiriqlari qiyinlik darajalarini bir xil shkalada bo'lishini alohida ta'kidlab o'tish lozim. Rayt xaritasi – test topshiriqlarining qiyinlik darajalari va sinaluvchilarning qobiliyat darajalari o'zaro mos kelishini aniqlovchi diagrammadir. O'tkazilgan test sinovlari natijasi asosida umumiy holda chizilgan Rayt xaritasi va Rash modeli bilan aniqlangan haqiqiy ball (T-ball) taqsimoti 2-rasmda keltirilgan. Rasmda test topshiriqlarining

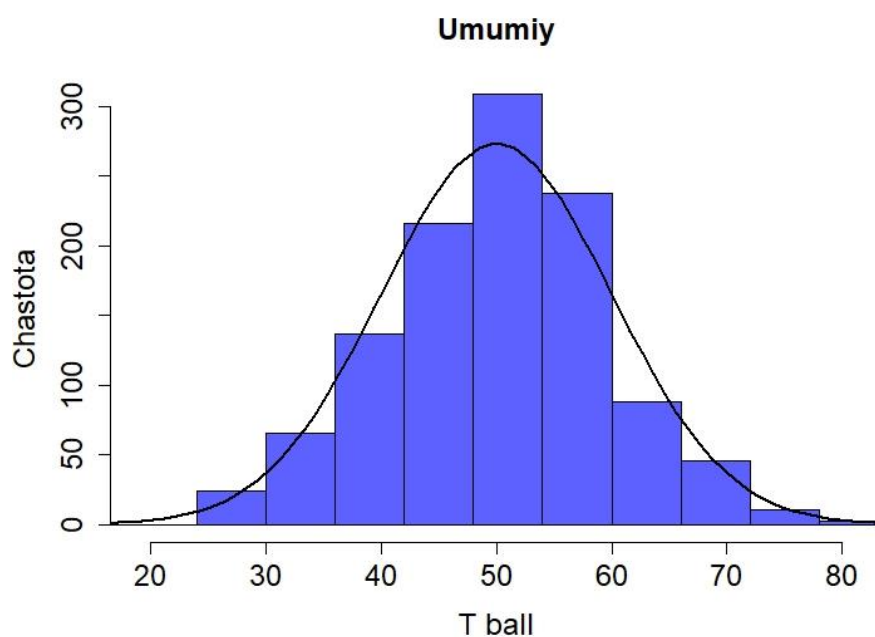
qiyinlik darajasi o'sib borish tartibida raqamlangan (1-test topshirig'i eng oson, ..., 55-test topshirig'i eng qiyin). Hisoblashlarga asosan talabgorlarning qobiliyat darajalari ($-2,93:3,22$) logit birligi oralig'ida, test topshiriqlarining qiyinlik darajalari esa ($-3,60:4,18$) oralig'ida ekanligi aniqlandi. 2-rasmdan umuman olganda test topshiriqlari qiyinlik darajalari savollar kitobida variantda yaxshi taqsimlanganligini ko'rish mumkin. Variantdagi test topshiriqlarining qiyinlik darajalari bo'yicha taqsimotini yanada yaxshilash mumkin. Buning uchun taqsimotdagi bo'sh joylarga va bir xil qiyinlikdagi test topshiriqlariga e'tibor berish

tavsiya qilinadi. Chunki bir xil qiyinlikdagi test topshiriqlari o'rniga bo'sh joylarga qiyinlik darajasi mos keladigan boshqa test topshiriqlarini qo'yish taqsimotni yanada yaxshilaydi.

Rayt xaritasidan, shuningdek, 1- va 54-55-o'rinlarda turgan test topshiriqlarining qiyinlik darajalari ($-3:3$) logit birligi oralig'idan tashqarida ekanligi ko'rinadi.



a)



b)

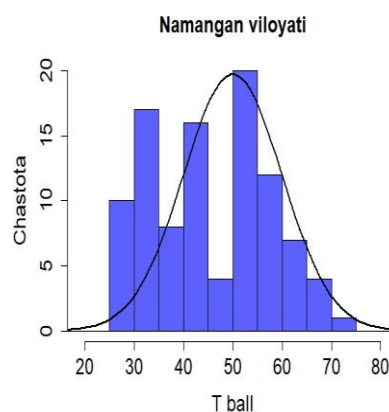
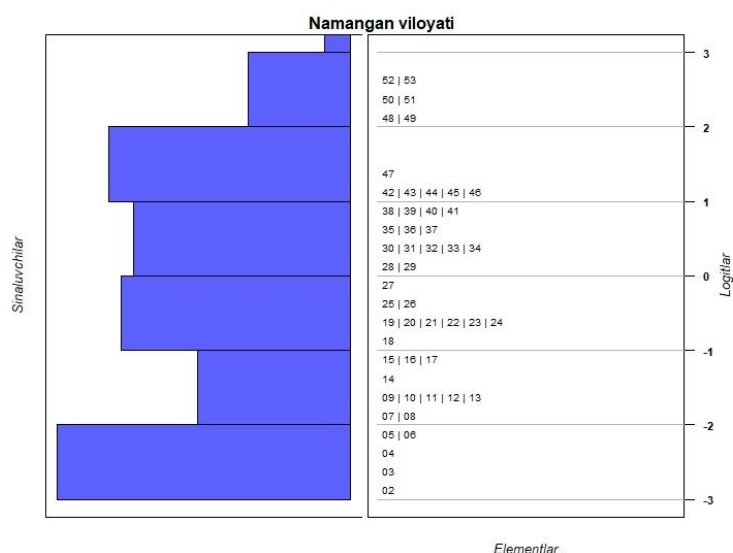
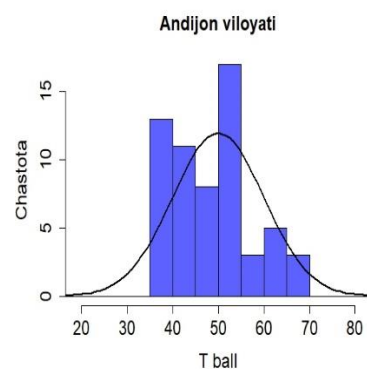
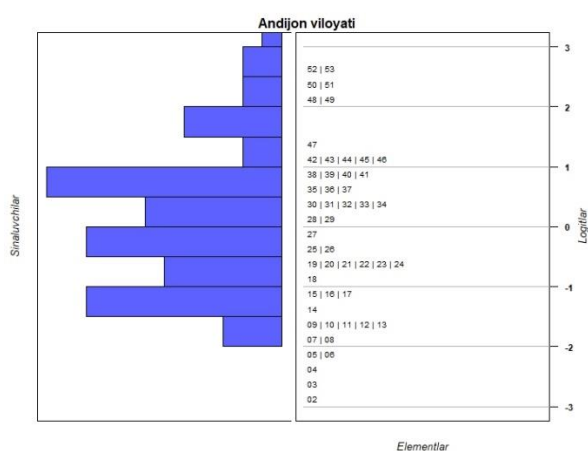
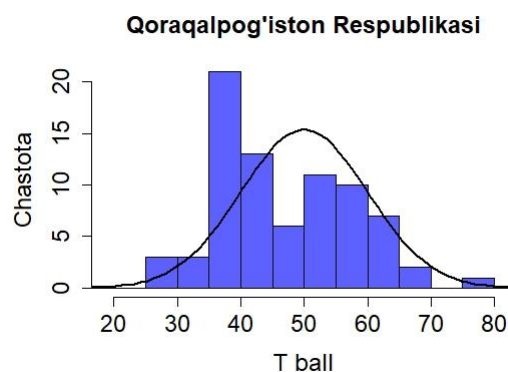
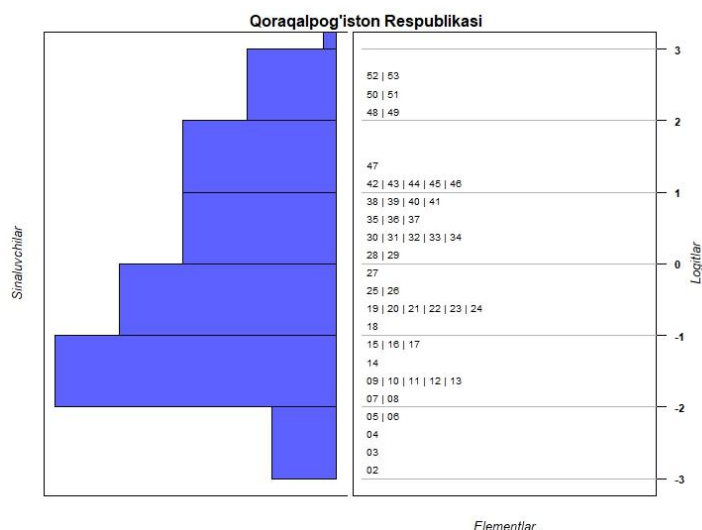
2-rasm. Talabgorlarning umumiy guruhi (1140 nafar) uchun (a) qobiliyat va qiyinlik darajalarining mosligi (Rayt xaritasi), (b) Rash modeli bilan aniqlangan haqiqiy ball (T-ball) taqsimoti

Qiyinlik darajasi bo'yicha 1- va 54–55-o'rinlarda turgan test topshiriqlarining qiyinlik darajasiga ishonch hosil qilish, agar mutaxassislar nuqtayi nazaridan ham ularning qiyinlik darajalari juda past yoki yuqori bo'lsa, uning o'rniga ($-3:3$) logit birligi oralig'iga to'g'ri keladigan qiyinlikdagi test topshiriqlaridan qo'yish tavsiya etiladi. Agar mutaxassislar nuqtayi nazarida ularning qiyinlik darajasi me'yorda deb topilsa, ushbu test topshiriqlari qiyinlik darajalarining ($(-3:3)$ logit birligi oralig'idan tashqarida ekanligini) sababini aniqlash lozimdir.

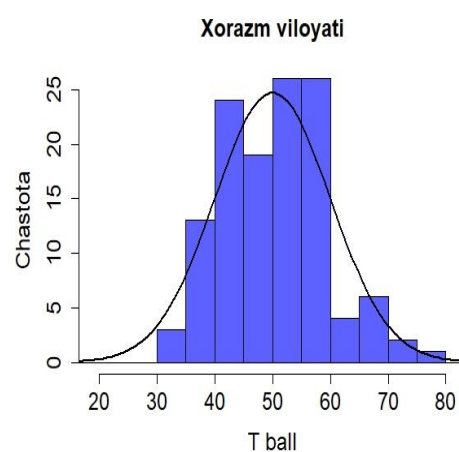
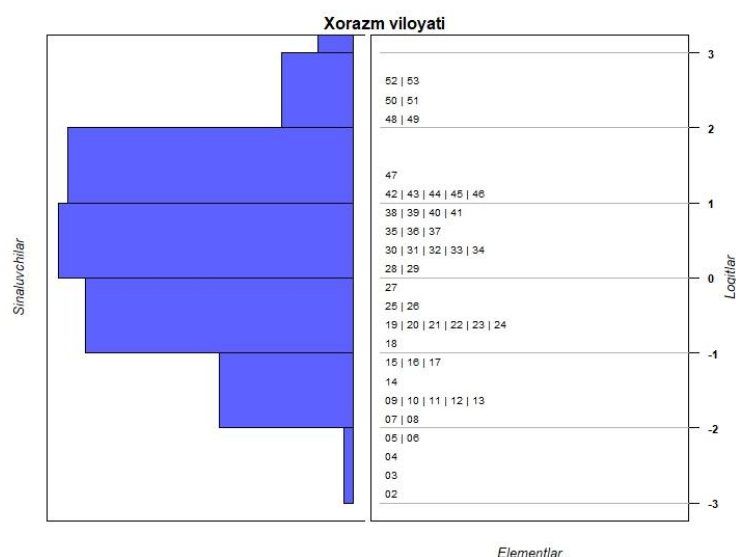
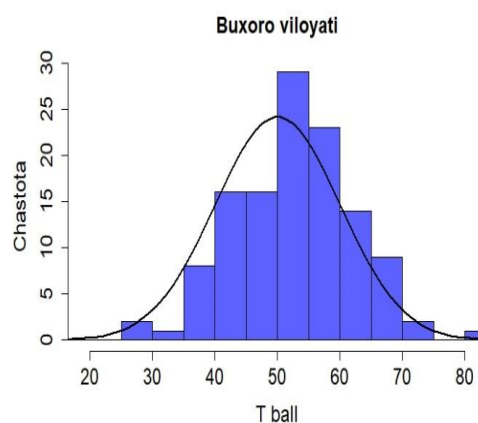
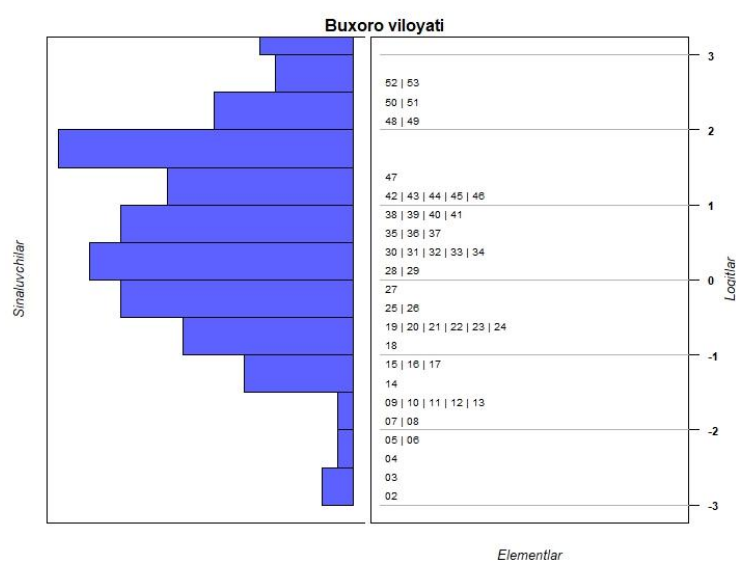
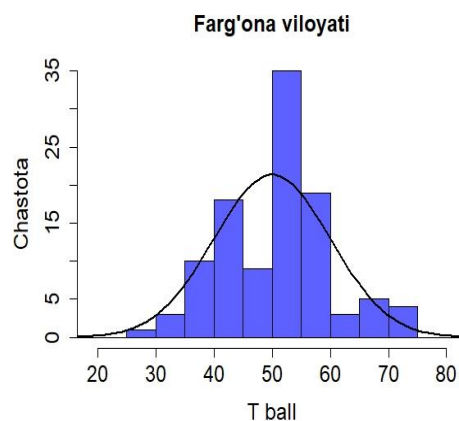
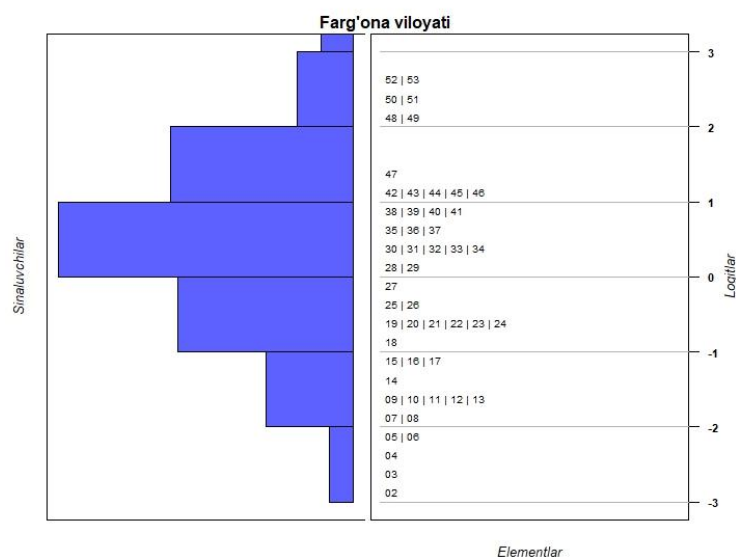
Rash modeli bilan aniqlangan haqiqiy ballar (T-ball) taqsimoti gistogrammasi normal taqsimotga yaqin (2-rasm). Gistogrammaning normal taqsimotga yaqinligi testning sifatini, test sinovlarining obyektiv o'tkazilganligini bildiradi. Bundan tashqari taqsimotga ko'ra o'rtacha qiymat taxminan 50 ball ekanligidan ham taqsimotning normalligi haqida xulosa qilish mumkin bo'ladi.

O'tkazilgan test sinovlari natijasi asosida hududlar kesimida chizilgan Rayt xaritasi va Rash modeli bilan aniqlangan haqiqiy ball (T-ball) taqsimoti 3–6-rasmlarda keltirilgan.

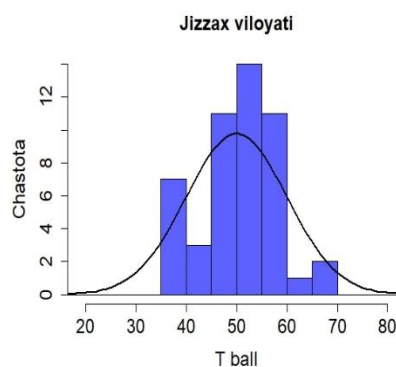
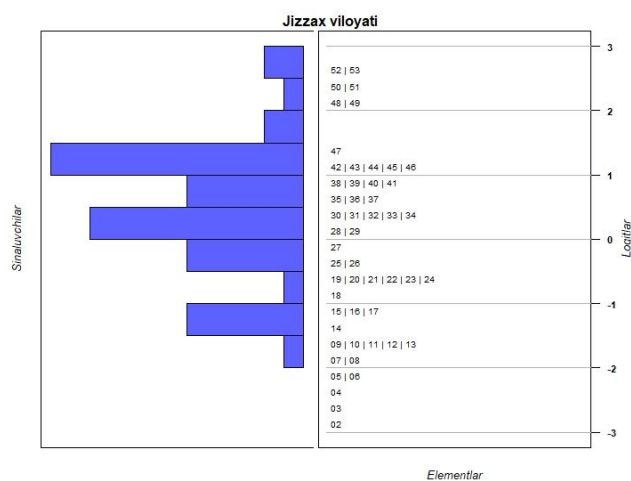
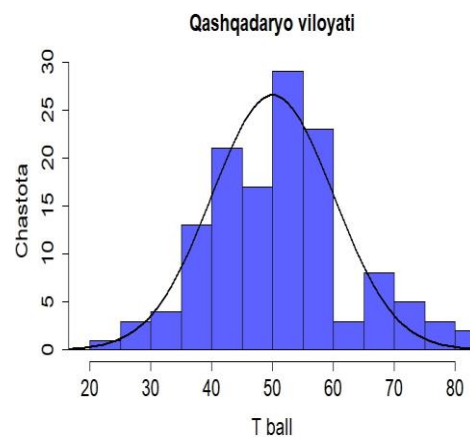
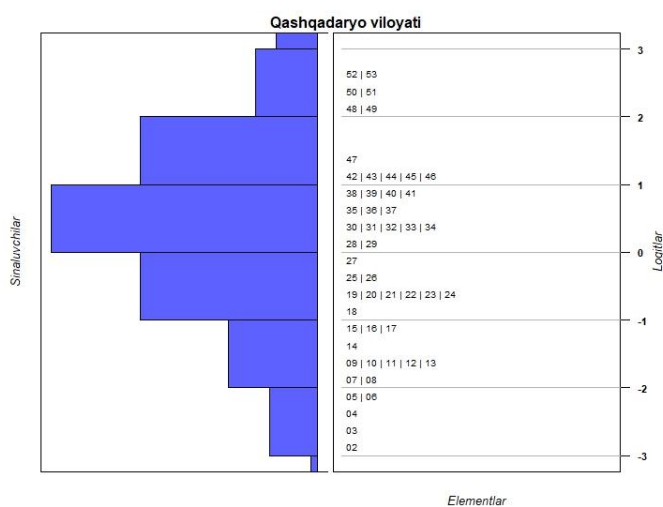
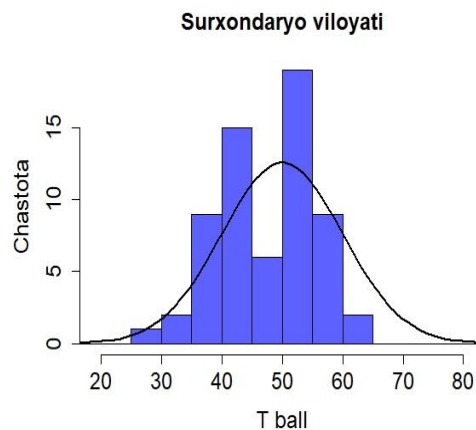
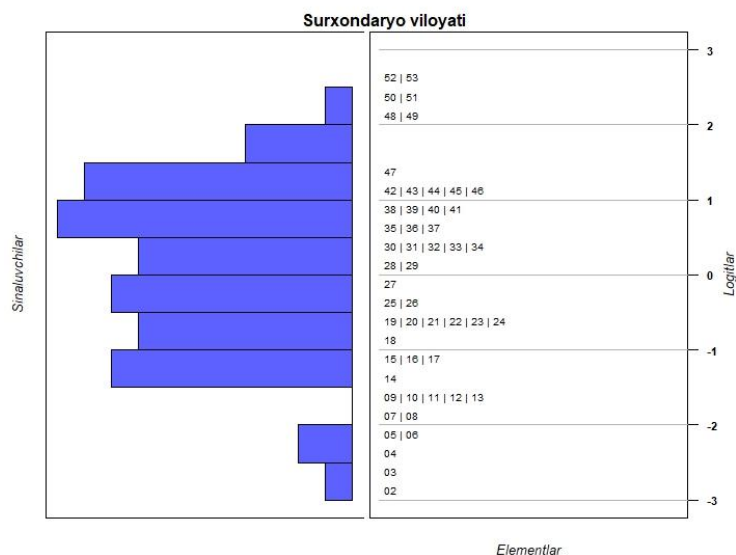
Qoraqalpog'iston Respublikasi, viloyatlar va Toshkent shahrida o'tkazilgan test sinovi natijalari asosida alohida chizilgan Rayt xaritalaridan ham umumiy holdagi kabi test topshiriqlari qiyinlik darajasining o'sib borish tartibida raqamlangan va 1-test topshirig'i eng oson, ..., 55-test topshirig'i eng qiyin ekanligini ko'rish mumkin. Hududlardagi test sinovi natijalarini hisoblashlarga asosan turli xil hududlarda ham talabgorlarning qobiliyat darajalari ($-2,93:3,22$) logit birligi oralig'ida, test topshiriqlarining qiyinlik darajalari esa ($-3,60:4,18$) oralig'ida ekanligi aniqlandi. Bu hududlarda o'zlashtirish darajalari o'rtachadan past va yuqori bo'lgan sinaluvchilar soni deyarli bir-biriga yaqin ekanligini bildiradi. Buxoro va Navoiy viloyatlarida esa o'zlashtirish darajalari o'rtachadan yuqori bo'lgan sinaluvchilar soni ko'proq ekanligi aniqlandi (4-rasm). Rasmlardan umuman olganda test topshiriqlari qiyinlik darajalari savollar kitobida variantda yaxshi taqsimlanganligini ko'rish mumkin. Bu hududlar bo'yicha Rash modeli bilan aniqlangan haqiqiy ball (T-ball) taqsimoti grafiklarda ham aks etgan.



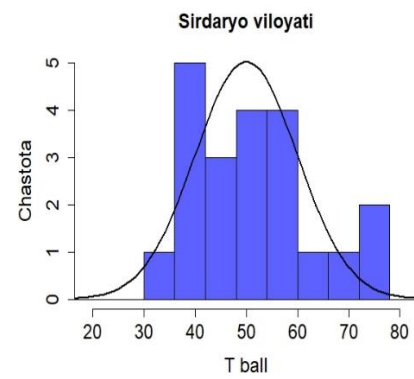
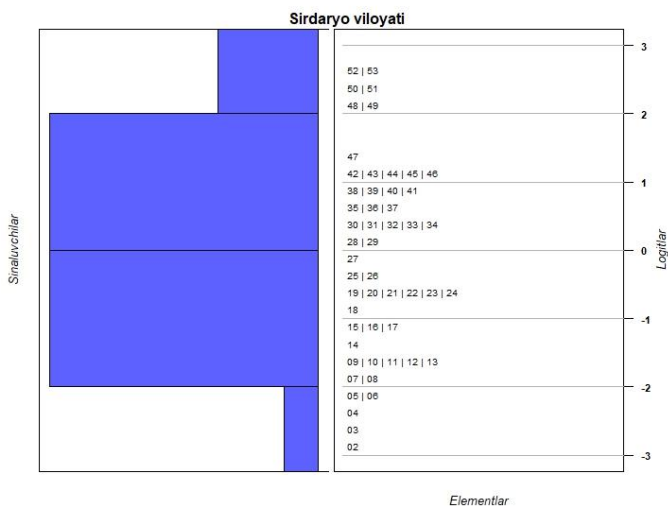
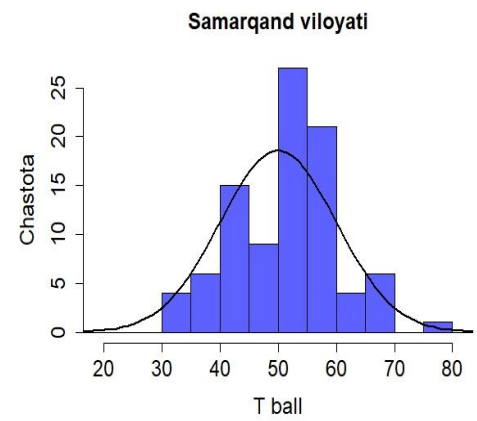
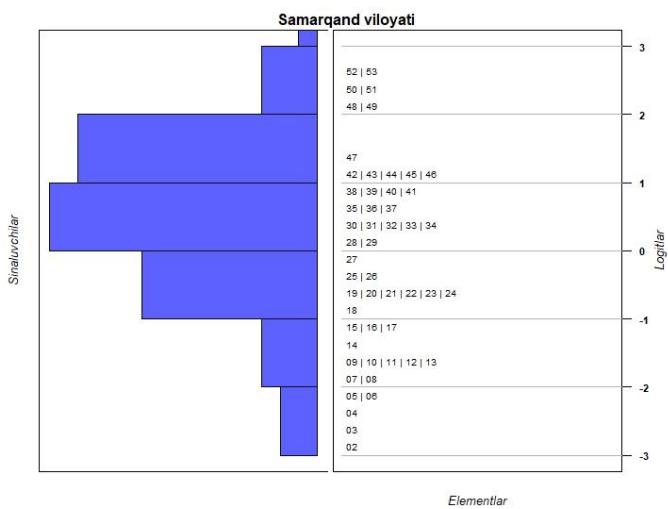
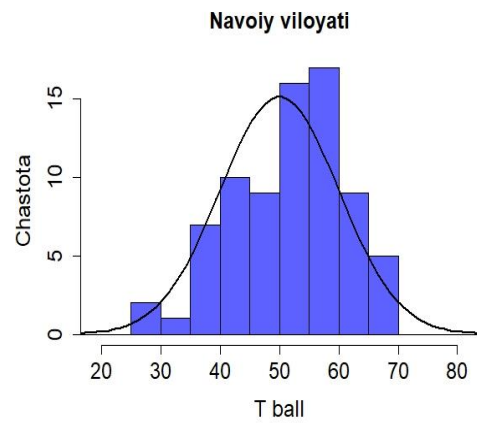
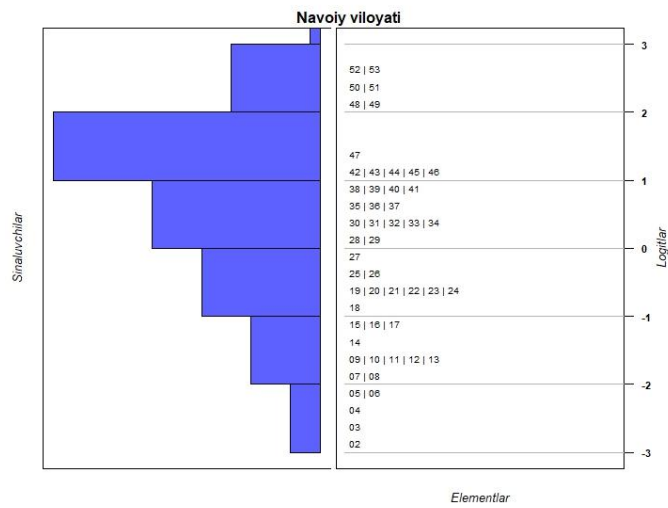
3-rasm. Qoraqalpog'iston Respublikasi (77 nafar), Andijon viloyati (60 nafar) Namangan viloyati (99 nafar) uchun qobiliyat va qiyinlik darajalarining mosligi (Rayt xaritasi), Rash modeli bilan aniqlangan haqiqiy ball (T-ball) taqsimoti.

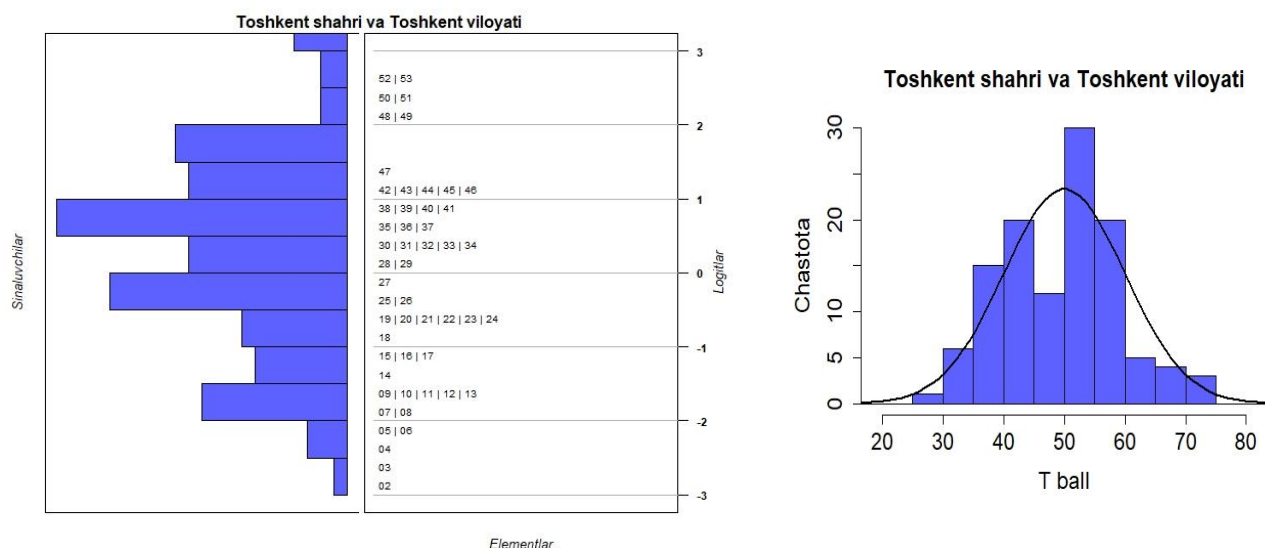


4-rasm. Farg'ona viloyati (107 nafar), Buxoro viloyati (121 nafar), Xorazm viloyati (124 nafar) uchun qobiliyat va qiyinlik darajalarining mosligi (Rayt xaritasi), Rash modeli bilan aniqlangan haqiqiy ball (T-ball) taqsimoti.



5-rasm. Surxondaryo viloyati (63 nafar), Qashqadaryo viloyati (133 nafar), Jizzax viloyati (49 nafar) uchun qobiliyat va qiyinlik darajalarining mosligi (Rayt xaritasi), Rash modeli bilan aniqlangan haqiqiy ball (T-ball) taqsimoti.





6-rasm. Navoiy (76 nafar), Samarqand (93 nafar), Sirdaryo (21 nafar), Toshkent viloyati va Toshkent shahri uchun qobiliyat va qiyinlik darajalarining mosligi (Rayt xaritasi), Rash modeli bilan aniqlangan haqiqiy ball (T-ball) taqsimoti.

Rash modeli bilan baholash, ya'ni turli xil guruhlarining qobiliyat darajalari va turli xil test variantlaridagi test topshiriqlarining qiyinlik darajalarini bir xil shkalada bo'lishi orqali ta'lim muassasalarining reytingini aniqlash va bu haqida xulosalar chiqarish mumkin bo'ladi.

Maqolada keltirilgan usullar bilan o'quvchilar, pedagoglar, ta'lim muassasalarining reytingini adolatli aniqlash mumkin. Bu esa xalqaro tizimda ham raqobatlasha oladigan reyting tizimini yaratish imkonini berish bilan birgalikda, standart, valid va ishonchli baholashni ta'minlaydi.

Xulosa

Fizika fani bo'yicha sinaluvchilardan standart test asosida olingan natijalarga ko'ra hududlarda o'zlashtirish darajalari o'rtachadan past va yuqori bo'lgan sinaluvchilar soni deyarli bir-biriga yaqin ekanligi aniqlandi.

Qobiliyat va qiyinlik darajalari bir xil shkalada bo'lishi ta'minlangan turli

xil hududlardagi talabgorlarning fizika fanidan test sinovi natijalarini solishtirish orqali ular ta'lim olayotgan ta'lim muassasalarining samaradorligini, unda faoliyat yuritayotgan pedagoglar va talabgorlarning bilimlari haqida xulosalar chiqarish mumkin.

Fizika fanidan bilimlarni baholashda standart testlardan foydalanish orqali turli xil test topshiruvchilar guruhlarini solishtirish, o'quvchilarning o'quv

dasturini qay darajada o'zlashtirishini, shuningdek, ularning yutuq va kamchilliklari haqida ma'lumotlar olish mumkin bo'ladi.

ADABIYOTLAR

1. Verger A., Parcerisa L., Fontdevila C. The growth and spread of national assessments and test based accountabilities: A political sociology of global education reforms. *Educational Review*, 2019, 71, 5–30.
2. Popham, W.J. Why standardized tests don't measure educational quality. *Educational Leadership*, 1999, 56 (6), 8–15.
3. Cameron Graham and Dean Neu. Standardized testing and the construction of governable persons. *Journal curriculum studies*, 2004, Vol. 36, 295–319.
4. W. James Popham. Uses and Misuses of Standardized Tests. *NASSP Bulletin*, 2001, Vol. 85, No. 622, 24-31.
5. H. Tyrone Black, David L. Duhon. Evaluating and Improving Student Achievement in Business Programs: The Effective Use of Standardized Assessment Test. *Journal of Education for Business*, 2014, 90-98.
6. Bond, Linda A. Norm- and Criterion-Referenced Testing, Practical Assessment, Research, and Evaluation, 1996, Vol. 5, 1-3.
7. Dylan Wiliam. Standardized Testing and School Accountability. *Educational Psychologist*, 2010, 107-122.
8. Rasch G., Probabilistic models for some intelligence and attainment tests, Copenhagen, Danish Institute for Educational research. 1960.
9. Hambleton, R., K., Swaminathan, H.,& Rogers, H.J., Fundamentals of item response theory. Newbury Park, CA: Sage. 1991.
10. Ivailo Partchev. A visual guide to item response theory, Friedrich-Schiller-Universitat Jena. 2004.
11. M.Dj. Ermamatov, A. Abbosov, A.A. Baratov, “Test topshiriqlarini kalibrovkalash va qobiliyatlarni tenglashtirish”. DTM “Axborotnoma” ilmiy-uslubiy jurnali, 2022 yil, 3-4-son, 4-15 betlar.
12. M.D. Ermamatov, M.D. Alimov, A.A. Sulaymonov, A.R. Sattiyev, Kalibrovkalangan test topshiriqlari: Sharq tillaridan o'tkazilgan test sinovi natijalarining statistik tahlili, “Axborotnoma” ilmiy-uslubiy jurnali, 2022 yil, 3-4-son, 16-83 betlar.
13. B. D. Wright and M. H. Stone. Best Test Design. MESA Press, Chicago, 1979.

INVARIANCE OF DIFFICULTY AND ABILITY ESTIMATION USING RASCH MODEL: ANALYZING PHYSICS TEST RESULTS

A.B. Normurodov, M. J Ermamatov, T.X. Jumayev

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

Abstract. In this paper, the evaluation of the knowledge of physics students through the use of standardized tests by various test administering groups is compared, and the extent to which their physics curriculum has been mastered is assessed. It has been found through the analysis with the Rash model that ensuring the equivalence of the difficulty levels of test tasks in various groups and ability levels is important for interpreting test results. Rating positions can be determined and conclusions can be drawn about this through the evaluation with the Rash model, which ensures equating of the ability levels of various groups and the difficulty levels of test tasks in various test variants.

Keywords: Standard test, Rasch model, ability, difficulty, Wright map.