МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ

Кафедра загальної математики

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

на тему: **«ЯКІСНИЙ АНАЛІЗ ПІДГОТОВКИ АБІТУРІЄНТІВ ВНЗ ЗАСОБАМИ ITEM RESPONSE THEORY»**

Виконала: студентка 2 курсу, групи 8.1118

спеціальності 111 математика

(шифр і назва спеціальності)

освітньої програми математика

Н.С. Дзундза

(ініціали та прізвище)

Керівник завідувач кафедри загальної математики, доцент, к.ф.-м.н. Зіновєєв І.В

(посада, вчене звання, науковий ступінь, прізвище та ініціали)

Рецензент доцент кафедри комп'ютерних наук, доцент, к.т.н. Матвіїшина Н.В

(посада, вчене звання, науковий ступінь, прізвище та ініціали)

Запоріжжя – 2020

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет математичний

Кафедра загальної математики

Рівень вищої освіти магістр

Спеціальність 111 математика

(шифр і назва)

Освітня програма математика

ЗАТВЕРДЖУЮ

Завідувач кафедри загальної математики, к.ф.-м.н., доцент

_____________ Зіновєєв І.В.

(підпис)

« 30 » 05 2019 р.

З А В Д А Н Н Я

НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТЦІ

Дзундзі Наталії Сергіївні

(прізвище, ім'я та по батькові)

1. Тема роботи Якісний аналіз підготовки абітурієнтів ВНЗ засобами Item Response Theory

керівник роботи Зіновєєв Ігор Валерійович, к.ф.-м.н, доцент

(прізвище, ім'я та по батькові, науковий ступінь, вчене звання)

затверджені наказом ЗНУ від « 29 » 05 2019 року № 811-с

2. Строк подання студентом роботи 26.12.2019

3. Вихідні дані до роботи 1. Постановка задачі.
2. Перелік літератури.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)
1. Огляд літературних джерел з питання теорії тестів
2. Основні засади Item Response Theory
3. Аналіз підготовки абітурієнтів засобами IRT

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)
ілюстрації до тексту роботи, графіки розрахунків, презентація

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання 30.05.2019

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1.	Розробка плану роботи.	09.09.19	
2.	Збір вихідних даних.	10.09.19	
3.	Обробка методичних та теоретичних	15.09.19	
	джерел.		
4.	Розробка першого розділу.	20.09.19	
5.	Розробка другого розділу.	10.10.19	
6.	Розробка третього розділу.	05.11.19	
7.	Оформлення та нормо контроль	25.11.19	
	кваліфікаційної роботи.		
8.	Попередній захист кваліфікаційної роботи.	05.12.19	

Студент ______________ Н. С. Дзундза
(підпис) (ініціали та прізвище)

Керівник роботи ______________ І. В. Зіновєєв
(підпис) (ініціали та прізвище)

Нормоконтроль пройдено

Нормоконтролер ______________ О. Г. Спиця
(підпис) (ініціали та прізвище)

РЕФЕРАТ

Кваліфікаційна робота магістра «Якісний аналіз підготовки абітурієнтів ВНЗ засобами Item Response Theory»: 84 с., 31 рис., 76 джерел, 5 додатки.

ДІАГНОСТИКА, ЗОВНІШНЄ НЕЗАЛЕЖНЕ ОЦІНЮВАННЯ, ЛАТЕНТНІ ПАРАМЕТРИ, ОЦІНЮВАННЯ, ПЕДАГОГІЧНІ ТЕСТИ, ТЕСТИ, ТЕСТУВАННЯ, IRT МОДЕЛІ, ITEM RESPONSE THEORY.

Об'єкт дослідження – педагогічні тести, IRT теорія обробки результатів тестування.

Мета роботи: аналіз підготовки абітурієнтів ВНЗ на основі педагогічних тестів засобами Item Response Theory.

Методи дослідження – аналітичний, синтез-метод, аналітико-синтетичний, практичний, порівняльний.

У кваліфікаційній роботі розглянуто історичні аспекти розвитку тестування у світі і в Україні, історію розвитку основних підходів до обробки результатів тестування – Classic Response Theory і Item Response Theory. Розглянуто основні теоретичні засади IRT теорії та моделі IRT: однопараметрична модель Г. Раша, двопараметрична та трипараметрична модель А. Бірнбаума, політомічна модель Раша-Мастерса, політомічна модель Е. Андерсенена та модель Тіссена-Стейнберга для завдань множинного вибору. А також інформаційні функції для оцінювання ефективності тесту. Наведено приклад застосування Item Response Theory для якісного аналізу підготовки абітурієнтів ЗНУ на основі результатів тренінгу із зовнішнього незалежного оцінювання 2019 року.

SUMMARY

Master's Qualification Thesis «Qualitative Analysis of the Education of the University-Entering Students with the Help of Item Response Theory»: 84 pages, 31 figures, 76 references, 5 supplements.

DIAGNOSTICS, EXTERNAL INDEPENDENT TESTING, LATENT PARAMETERS, EVALUATION, PEDAGOGICAL TESTS, TESTS, TESTING, IRT MODELS, ITEM RESPONSE THEORY.

The object of the study is the pedagogical tests, IRT theory of the processing of the test results.

The aim of the study is to analysis of the preparation of the university entrants on the basis of the pedagogical tests by means of Item Response Theory.

The methods of the research are analytical, synthesis, analytical-synthetic, practical and comparative.

In the qualification paper the historical aspects of test development in the world and in Ukraine, and the history of the development of basic the approaches to the processing of the test results – Classic Response Theory and Item Response Theory. The basic theoretical foundations of IRT theory and IRT models: one-parameter G. Rasch model, two-parameter and three-parameter A. Birnbaum model, Rasch-Masters polytomous model, E. Andersenen polytomous model and Thissen-Steinberg model the for multiple choice items have been studied. As well as information features to evaluate test performance. An example of the application of Item Response Theory for the qualitative analysis of ZNU entrants preparation based on the results of the 2019 External Independent testing training has been given.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

ВНЗ – вищий навчальний заклад

ДПА – державна підсумкова атестація

ЄС – Європейський Союз

ЗНО – зовнішнє незалежне оцінювання

ЗНУ – Запорізький національний університет

ПЗ – програмне забезпечення

ACE – American Council on Education

ASC – Assessment Systems Corporation

CRT – Classical Response Theory

CTT – Classical Test Theory

ETS – Educational Testing Service

GRE – Graduate Record Examinations

IRF – Item Response Functions

IRT – Item Response Theory

NAEP – National Assessment of Educational Progress

NRM – Nominal Response Model

PCM – Partial Credit Model.

RUMM – Rasch Unidimensional Measurement Model

SAT – Scholastic Aptitude Test

Tempus –Trans-European Mobility Programme for University Studies

ВСТУП

Актуальність теми. Впровадження зовнішнього незалежного тестування для випускників шкіл, ліцеїв та коледжів, а у майбутньому заміна ДПА в 9 класі на ЗНО, дає можливість закріпити використання тестування як інструменту вимірювання знань. У зв'язку з цим аналіз результатів тестування, як можливість диференціювання знань учнів чи студентів, набуває вагомого значення.

Мета та завдання дослідження: аналіз підготовки абітурієнтів ВНЗ на основі педагогічних тестів засобами Item Response Theory.

Об'єкт дослідження – педагогічні тести, IRT теорія обробки результатів тестування.

Методи дослідження. У роботі використовуються такі методи, як: аналітичний, синтез, аналітико-синтетичний, практичний, порівняльний.

Практичне значення одержаних результатів. Результати роботи можуть бути використанні вчителями загальноосвітніх шкіл та викладачами закладів вищої освіти для аналізу та підвищення якості педагогічних тестів.

Структура й обсяг кваліфікаційної роботи. Робота складається з трьох розділів. У першому розділі наведено історичні аспекти розвитку тестування у світі і в Україні, а також історію розвитку основних підходів до обробки результатів тестування – Classic Response Theory і Item Response Theory. У другому розділі наведено теоретичні засади теорії IRT, розглянуто її моделі: однопараметрична модель Г. Раша, двопараметрична та трипараметрична моделі А. Бірнбаума, політомічна модель Раша-Мастерса, політомічна модель Е. Андерсенена та модель Тіссена-Стейнберга для завдань множинного вибору. А також інформаційні функції для оцінювання ефективності тесту. У третьому розділі наведено приклад застосування Item Response Theory для якісного аналізу підготовки абітурієнтів ЗНУ на основі результатів тренінгу із зовнішнього незалежного оцінювання 2019 року.

1 ОГЛЯД ЛІТЕРАТУРНИХ ДЖЕРЕЛ З ПИТАННЯ ТЕОРІЇ ТЕСТІВ

1.1 Історичні аспекти розвитку тестування та підходів до обробки результатів тестування

Активне використання тестів у якості інструменту вимірювання знань почалось в кінці XIX на початку XX ст. Одним із перших вчених, який вивчав теорії тестового контролю був Ф. Гальтон (F. Galton). Він працював над створенням об’єктивних методів оцінки здібностей і властивостей особистості. Ф. Гальтон зазначав, що методично впорядковане тестування вимагає певних умов експерименту і висунув ідею статистичної обробки результатів експерименту. У 1883 році у його роботі «Inquiries Into Human Faculty and Its Development» [1] були викладені теоретичні основи тестування. А у 1894 р. в школах вперше з'явилися тести успішності (для перевірки знань, умінь і навичок учнів) з правопису.

Важливим внеском в тестологію було розділення тестів на педагогічні (Educational Test) і психологічні (Intelligence Test), обґрунтування мети використання педагогічних тестів – об'єднання в групи учнів, що засвоюють рівний за обсягом матеріал з однаковою швидкістю, а також введення поняття педагогічного та розумового віку в роботах В. А. Макколла (W. A. McCall) «How to Measure in Education» [2], «How to Experiment in Education» [3].

У напрямку психологічних тестів працював американський психолог Дж. Кеттелл (J. Cattell). Він є основоположником методів тестування, автором ряду психологічних тестів, один з перших фахівців із психометрії та психодіагностики, ввів термін «розумові тести» в 1890 році. А також сформулював фундаментальні принципи стандартизації процедури проведення тестування [4]:

– рівність умов для всіх випробуваних;
– обмеженість у часі;

– відсутність сторонніх у лабораторії;

– надійність обладнання;

– забезпечення однакового інструментарію і чіткості умов завдань;

– оброблення результатів тестування статистичними методами.

Ще один психолог, який зробив внесок в тестологію, був А. Біне (A. Binet). З його ім'ям пов'язано початок розвитку сучасної тестології масового застосування тестів. Він розподілив тести на чотири класи: тести інтелекту, тести здібностей, тести досягнень, тести особи [5].

У співавторстві із Т. Симон (T. Simon) працював над якістю тестових завдань. У результаті чого в 1905 році вони разом склали метричну шкалу розумового розвитку, відому під назвою «тест Біне-Симона» (в 1916 р. була перероблена Л. Терменом в «шкалу інтелекту Стенфорд-Біне»), чим дали поштовх до створення нових психологічних, психолого-педагогічних та педагогічних тестів.

Варто зазначити, важливість внеску математика К. Пірсона (K. Pearson), який працював над теорію аналізу тестування з 1896 р. Він вдосконалив методи кореляційного, регресійного і факторного аналізу тестів, визначив коефіцієнт кореляції. Розроблена ним теорія почала використовуватися для визначення надійності та валідності тестів. У своїй статті [6] 1901 року К. Пірсон писав про ідею «методу головних осей», яка є важливою внеском в факторний аналіз у статистичному плані.

Ще одним внеском К. Пірсона у розвиток теорії тестів було визначення трьох основних принципів проведення дослідження за допомогою тестування, зокрема [5]:

– використання серії однакових тестів для великої кількості людей;

– статистичне оброблення результатів;

– виділення еталону оцінки.

Е. Торндайк (E. Thorndike) один із перших, хто працював над питанням методичних аспектів розроблення тестів у педагогіці. В 1903 р. видав книгу «Education Psychology» [7] в якій було описано види тестів, за допомогою яких можливо ефективно визначити успіхи у навчанні. Крім того розробив систему тестів для кількісного визначення рівня психологічного розвитку.

Подальший розвиток математичних методів у тестології пов'язано з ім'ям Ч. Спірмена (C. Spearman), який у своїй роботі «The Proof and Measurement of Association between Two Things» [8] 1904 року, зробив значний крок у використанні методу кореляції для тестування. Він, запропонувавши загальні основи використання кореляційного аналізу для стандартизації тестів, показав, що кореляція між двома змінними свідчить про наявність загального фактора (причини, що визначає величини цих змінних) і специфічних чинників, властивих кожній змінній.

Об'єктивністю одержаних результатів в 1905 р. займалися Дж. Орлеанс (J. Orleans) та Г. Сілі (G. Seeley), які у своїй роботі «Об'єктивні тести» визначили головні цілі тестування: забезпечення викладачів інформацією як про рівень засвоєння вивченого матеріалу, так і про обсяг матеріалу, призначеного для подальшого вивчення та надання допомоги викладачам у виборі методів навчання. Вони ж і запровадили класифікацію за формою діагностичних тестів: індивідуальні і групові [9].

У 1913 році розробкою групових тестів, поштовхом до створення яких були шкали А. Біне, які давали можливість вводити одночасно велику кількість суб'єктів, займався В. Пайл (V. Pyle). Продовжував роботу у цьому напрямку Р. Пінтер (R. Pintner), який у 1919 р. в статті «Anon-language group intelligence test» [10] писав про набір з шести невербальних тестів інтелекту які можна було б використати глухими та не англомовними людьми.

Після активного розвитку тестів у педагогіці тестування почало розвиватися у двох основних напрямках: створення та використання тестів інтелектуального розвитку і створення та використання педагогічних тестів, призначених для оцінки академічних знань та здібностей учнів [9].

Упродовж 1925-1926 рр. С. Прессі (S. Pressey) працював над теорією програмованого навчання, яка базувалася на засадах біхевіоризму і полягала у дотриманні схеми: стимул – реакція, оскільки учень переходив до наступного завдання тільки після того, як дасть правильну відповідь на попереднє питання [11].

Крім того він сконструював і запатентував кілька машин для перевірки знань дітей за допомогою методу тестування. В роботах «Machine for automatic teaching of drill material» [12] і «Simple apparatus that gives tests and scores and teaches» [13] описав свою ідею використання тестових запитань із множинним вибором яка полягала в тому, щоб зафіксувати питання в машині так, щоб воно не рухалася далі, поки студент не обрав правильну відповідь. Це була перша демонстрація того, що машина може навчити, а також того, що знання результатів є причиною навчання.

Історичний аналіз розвитку тестування в США провів Ф. Фрімен (F. Freeman) у своїй роботі 1926 р. «Mental Tests. Their History, Principles and Applications» [14]. Він намітив основні напрямки його подальшого розвитку: створення педагогічних тестів, метою яких є вимірювання рівня засвоєння знань, умінь і навичок на різних етапах навчання, а також створення тестів для визначення рівня загального розумового розвитку і спеціальних здібностей для подальшого навчання.

У 1926 році Ч. Грін (Ch. Greene) у монографії «New Type Tests» [15] проаналізував переваги і недоліки тестів, розробив класифікацію і рекомендації щодо створення тестів: визначення точного обсягу матеріалу для тестування, тривалості проведення тестів, розміщення питань у порядку зростання складності і уникнення регулярності чергування вірних та невірних відповідей.

З 1937 року ідеї факторного аналізу отримала свій розвиток у роботах Л. Терстоун (L. Thurstone) «Multiple factor analysis» [16], «Multiple-factor Analysis: A Development and Expansion of The Vectors of the Mind» [17]. Він вважав, що інтелект складається з декількох первинних здібностей (вербальної, перцептивної, здатності до сприйняття простору, пам'яті, швидкого мовлення), які є рівноправні. За допомогою факторного аналізу не просто встановлюється зв'язок зміни однієї змінної з іншою змінною, а визначається міра зв'язку з ними і виявляються основні причини, що лежать в основі зазначених змін.

Внесок у тестологію зробив і Д. Векслер (D. Wechaler), який в 1939 р. розробив тест розумового розвитку для дорослих (шкалу Wechsler-Bellevue, в яку входило одинадцять субтестів). В результаті модифікацій в 1949 р. Д. Векслер розробив шкалу розумового розвитку для дітей. В 1955 р. з'явилася нова шкала Векслера для вимірювання інтелекту дорослих WAIS (Wechsler Adult Intelligence Scale) і для дітей WISC (Wechsler Intelligence Scale for Children), яка мала аналогічну структуру з шкалою Векслер-Бельвью і спираючись на роботи Спірмена про структуру інтелекту включала факторний аналіз шкали і виділяла такі основні фактори: загальний інтелект, вербальний фактор, невербальний фактор і фактор пам'яті.

У 1955 р. Д. Векслер написав книгу «Manual for the Wechsler Adult Intelligence Scale» [18], яка базується на роботі з вибіркою в 1700 дорослих віком від 16 до 64 років і містить вказівки щодо IQ таблиць та таблиць масштабованих балів. Крім того він перше використав вербальні тести і тести на виконання, які є одними з найбільш поширені і використовуються у світі.

З розвитком тестології розвивалися методи розробки і перевірки тестів, якими займалися спеціальні державні служби. В 1947 року почала працювати приватна некомерційна служба Educational Testing Service – національна служба з тестування в освіті засновники якої Американська Рада з питань освіти (ACE), Фонд Карнегі з підвищення рівня освіти та викладання і Рада коледжів.

ETS працює над розробкою методики складання тестів і програм тестування, до яких належать як загальнонаціональні програми визначення рівня знань школярів, студентів коледжів та університетів, так і програм визначення професійної придатності – MMPI (Minnesota Multiphasic Personality Inventory) [19].

Вже у другій половині XX ст. у практичній педагогіці починають використовувати критерійні орієнтовані тести. Їх особливістю є орієнтація не на середню норму виконання тесту учнями, а на особово-орієнтовані досягнення в конкретній освітній області [20].

У 1954 році Б. Скіннер (B. Skinner) розглядає ідею підкріплення (тобто рефлексології) стосовно навчання дитини у своїй роботі «The science of learning and art of teaching» [21]. Він вважав, що підкріплення у процесі навчання є важливим аспектом, адже віддалені перспективи набуття знань, – недостатній стимул для учнів і її потрібно підкріплювати (заохочувати) оцінками. Б. Скіннер запропонував поділити навчальний матеріал на невеликі порції, що сприятиме його кращому засвоєнню, тобто успіху учнів у навчанні. Такий підхід забезпечує максимальну частоту заохочень або «підкріплень», а значна або оптимальна кількість підкріплень сприяє наближенню поведінки до такої, яку потрібно сформувати. У цьому випадку рефлексологічна формула «стимул – реакція» виглядає так: «успіх – прагнення до навчання» [11].

У 60-х роках XX ст. тестування почалося активно використовуватися у системі освіти США. Була розроблена програма National Assessment of Educational Progress, мета якої стимулювати внутрішню зацікавленість людей для вдосконалення системи шкільної освіти.

У 1983 році NAEP починає займатися ETS, яка проводила систематизацію робіт з тестування, стандартизацію тестів, займалася встановленням єдиних правил процедури тестування, а також розробкою критеріїв для визначення якості освіти [22].

У 1959 р. Міжнародною асоціацією з оцінювання шкільної успішності було уперше проведено міжнародне порівняльне моніторингове дослідження успішності з математики учнів початкової і середньої школи. У 1969 році проведено перше Національне вимірювання поступу освіти (при фінансуванні урядом США), а в 1978 році ухвалено закон про обов’язкове періодичне загальнодержавне вимірювання освітнього поступу [23].

Для вступу до закладів вищої освітив США абітурієнти повинні пройти тестування для визначення їхньої академічної успішності. Тести, які використовуються, є стандартизованими, що дає можливість проводити порівняльний аналіз в масштабах країни.

Найвідомішими стандартизованими тестами у світі є:

– Scholastic Aptitude Test (SAT – тест для визначення академічної успішності, який складається з двох розділів: вербального SAT-Verbal та математичного SAT-Math);

– Achievement Test (ACT – тест для на визначення глибини знань з окремих дисциплін);

– Overseas Registration Exam (ORE – іспит, який повинні скласти кваліфіковані стоматологи, щоб зареєструватися в General Dental Council);

– Graduate Management Admission Test (GMAT – тест для визначення здатності успішно навчатися в бізнес-школах);

– Law School Admission Test (LSAT – вступний тест для юридичних вузів);

– Medical College Admission Test (MCAT – комп'ютерний тест для потенційних студентів-медиків);

– Test of English as a Foreign Language (TOEFL – тест для вимірювання знання англійської не-носіями мови, які бажають вступити до закладів вищої освіти США або Канади);

– International English Language Testing System (IELTS – міжнародна система тестування англійської мови).

В 90-і рр. ХХ ст. головним в освіті стала «підзвітність», тобто увага звертається на конкурентну здатність дітей, що в свою чергу призвело до більшого акцентування уваги на використанні стандартизованих тестів [24].

В 1983 р. у США до публікації виходить доповідь «A Nation at Risk: The Imperative for Educational Reform» [25], над якою працювала комісія з 18 членів протягом двох років. Причиною роботи була необхідність у зміні освітньої системи США, яка не задовольняла національних потреб у конкурентній робочій силі. У звіті зазначається, що середній бал за SAT знизився «на понад 50 балів» в усному розділі та «майже 40 балів» у розділі математики за період 1963-1980 рр. Відповідно до цих та подібних проблем комісія дала 38 рекомендацій, розділених на 5 основних категорій: зміст, стандарти та очікування, час, навчання, лідерство та фіскальна підтримка.

В 1990 році була заснована програма Trans-European Mobility Programme for University Studies, для успішного співробітництва та вдосконалення систем вищої освіти в країнах-партнерах ЄС. Tempus фінансує між університетську співпрацю в таких сферах: розробка навчальних планів; управління університетами; взаємодії науковців та громадянського суспільства та структурні реформи в галузі вищої освіти. Програма в продовж всього часу існування поділялась на декілька етапів: Tempus I – 1990-1994 рр.; Tempus II – 1994-1998 рр.; Tempus II ibis – 1998-2000 рр.; Tempus III – 2000-2006 рр.; Tempus IV – 2007-2013 рр. [26].

У 1990 році була започаткована програма імені Жана Моне, яка має за мету підвищити рівень знань та поінформованості суспільства з питань європейської інтеграції, через стимулювання викладання, дослідницької діяльності з європейської інтеграції, зокрема стосунків ЄС з іншими країнами, та міжлюдського та міжкультурного діалогу.

У 2004 році була започаткована програма Erasmus Mundus, яка є механізмом студентських та академічних обмінів між державами Євросоюзу та іншими країнами. Перший етап програми працював з 2004-2008 рр. і реалізувався за трьома напрямками:

– спільні магістерські та докторські програми, що передбачають стипендіальну підтримку;
– партнерство університетів;
– підвищення привабливості європейської вищої освіти.

Другий етап удосконаленої програми Erasmus Mundus проходив в період 2009-2013 рр.

У 2013 р. починає працювати програма Союзу для освіти, підготовки, молоді та спорту «Erasmus+» відповідно до регламенту (ЄС) № 1288/2013 Європейського Парламенту та Ради [27]. Вона сприяє створенню та і розвитку європейських мереж, забезпечуючи можливості для співпраці між зацікавленими сторонами та обміну і передачі знань та інновацій в різних сферах.

Активна робота над теорією тестування сприяла розвитку підходів до конструювання та обробки результатів педагогічних тестів. Серед них виділяють Classical Response Theory і Item Response Theory (в рамках теорії латентно-структурного аналізу)

В основу Classical Response Theory покладено класичний статистичний апарат для дослідження результатів вимірювань. Засновником класичної теорії тестів є Ч. Спірмен (C. Spearman). Однак всебічно основи класичної теорії тестів вперше були викладені в 1950 році у роботі Г. Гулліксена (H. Gulliksen) «Theory of Mental Tests» [28]. Значний вклад в розвиток цієї теорії вніс Л. Гуттман (L. Guttman) у роботах [29, 30]. Більш детальніше про його вклад описано в роботі «Louis Guttman's Contributions to Classical Test Theory» [31]. З розвитком математичного апарату CTT тільки вдосконалювалась і була викладена в роботах Ф. М. Лорд (F. M. Lord) і М. Р.Новік (M. R. Novick) [32], а також Л. Крокером (L. Crocker) і Дж. Алгино (J. Aligna) [33]. Початки дисперсійного аналізу та внутрікласової кореляції, які застосовують для оцінки величини коефіцієнта надійності тесту, були закладені наприкінці дев’ятнадцятого століття Р. А. Фішером (R. A. Fisher) [34].

Першими поштовхами до створення IRT були дослідницькі роботи А. Біне (A. Binet) і Т. Симон (T. Simon), в яких вони графічним шляхом прагнули виявити якість завдань. Вони першими помітили закономірність, яка нагадувала криву (пізніше буде названа характеристичною).

У 1936 році М. Річардсон при обробці тесту студентів з 12 груп, по сто чоловік у кожній, звернула увагу на різницю крутизни кривих тестових завдань. Вона усвідомила плідність використання усереднених точок для графічної презентації формальних характеристик завдань проектованих тестів і запропонувала розглядати міру крутизни як приблизну оцінку, що диференціює здатності завдання.

Великий вклад в теорію IRT зробив датський математик Г. Раш (G. Rasch), який запропонував математичну модель, яка більш відома як однопараметрична модель Раша і ввів дві міри: «логіт рівня знань» і «логіт рівня складності завдання».

Важливим внеском в теорію були роботи А. Бірнбаума, його моделі (двохпараментична і трипараметирична), моделі Раша-Мастерса і Е. Андерсенена для політомічних завдань та Тіссена-Стейнберга для завдань множинного вибору [35]. Сучасні дослідження проводилися в роботах В. Аванесова, М. Б.Челишкової [36], В. С. Ким [37], Ю. М. Нейман та В. А. Хлебникова [38].

1.2 Розвиток тестування в Україні і сучасні підходи в теорії тестів

В Україні перші роботи з тестології почали з’являтися на початку XX ст., а самі тести використовувалися як допоміжний інструмент перевірки знань. У 1925 році при педологічному відділі Інституту методів шкільної роботи створено спеціальну тестову комісію, яка в 1926 році підготувала стандартизовані тести з природознавства, суспільствознавства, арифметики, географії та правопису. Однак подальший розвиток теорії вітчизняного тестування в 30 – 50 році XX ст. призупинився, причиною стало прийняття Постанови ЦК ВКП(б) від 4 липня 1936 р. «Про педологічні перекручення в системі Наркомосів». Надалі використовувалися лише зарубіжні зразки тестів в обмеженій кількості і переважно у сфері психологічних досліджень [39].

З 60 років XX ст. розвивався інтерес до теорії програмованого навчання, ідея якого – управління навчанням, навчально-пізнавальними діями учнів. Вже в 1963 р. відбулася перша республіканська конференція з програмованого навчання. А 1964 р. у Науково-дослідному інституті психології була створена лабораторія програмованого навчання. Її фахівці займалися дослідженням психологічних основ програмованого навчання, розробкою програмованих матеріалів з низки тем з різних предметів, перевіркою нових навчаючих і контролюючих машин, вивченням зарубіжної теорії та практики програмування навчання [40]. Проблемами програмного навчання займалися Г. Костюк і В. Глушков в роботі «Наукові проблеми програмованого навчання та шляхи їх розробки» [41], психологією програмованого навчання займався Г. Балл у статті [42]. Важливим внеском у розвиток програмованого навчання були роботи Є. Морокішко «Експериментальна перевірка програмованих посібників» [43] і М. Розенберг «Експериментальне дослідження ефективності програмованого навчання» [44].

Упродовж 60-х–90-х рр. XX ст. знову почали використовувати тестування однак переважно у військових училищах та інших спеціальних навчальних закладах. Спроби відродити тестування були невдалими, так само як невдалою була організація тестування успішності випускників загальноосвітніх навчальних закладів, здійснена Міністерством освіти 1993-1994 рр.

У 90 р. XX ст. почалося впровадження тестування на випускних іспитах у школах, результати яких зараховувалися як складова частина вступних іспитів закладів вищої освіти України.

У 1993 році Міністерство охорони здоров'я України за підтримки міжнародних фондів і безпосередньої участі західних спеціалістів запровадило систему перевірки й оцінювання професійних знань і майстерності студентів та випускників медичних навчальних закладів – «Кроки». У рамках цього масштабного проекту створено банк завдань, у який збирають емпіричні дані, проводять їх психометричний аналіз [39].

29 квітня 1993 р. Україна приєдналася до міжнародної програми Tempus під час другого етапу її діяльності, яка була спрямована на удосконалення управління закладами освіти, оновлення або розробку нових навчальних програм і підвищення кваліфікації викладачів. Таким чином, Україна отримала можливість брати участь у двох конкурсах попередньої підготовки проектів (Pre-JEPs) 1993 та 1994 рр.

У 1997 році створюється Центр моніторингу освіти при Інституті змісту і методів навчання Міністерства освіти України. З 1999 року його реформовано у відділ моніторингу якості загальної середньої освіти Науково-методичного центру середньої освіти, який започатковує всеукраїнські моніторингові дослідження [45]:

– якості засвоєння курсу фізики (1999);

– якості основної навчальної літератури для загальноосвітніх навчальних закладів (2001);

– якості математичної освіти випускників початкової, основної та старшої школи (2002);

– стану фізичного, морального і психічного розвитку учнів та інфраструктури навчальних закладів, яка забезпечує збереження здоров’я школярів (2005).

У 1999 році Міжнародний фонд «Відродження» ініціював дослідження з незалежного тестування для нових вступних іспитів, як антикорупційний захід.

У 2002 році центр тестових технологій за підтримки Міністерства освіти й науки України та Міжнародного фонду «Відродження» почав готувати нормативну база щодо впровадження зовнішнього незалежного оцінювання. Вони провели тестування на 200 студентах перших курсів для апробації тестових завдань і розроблення технології адміністрування. В 2003 році тестування проводилося на 3121 випускниках 670 школах України з історії та математики і 4 заклади вищої освіти прийняли рішення щодо зарахування результатів тестування як вступних випробувань. В 2004 р. тестування проходили 4485 випускники з історії, математики, української мови та економіки і 31 заклад вищої освіти зарахували результати як вступні випробування. В 2005 р. вийшов указ про здійснення переходу щодо проведення вступних випробувань у форматі ЗНО. І цього ж року тестування пройшло 10030 учнів 1567 шкіл України [46].

В 2006 р. у Державному бюджеті України уперше передбачаються кошти на запровадження ЗНО. Створюються 8 регіональних центрів оцінювання якості освіти і залучаються 6300 інструкторів для проведення тестування майже 42 тис. випускників. В 2007 р. завершується створення регіональних центрів якості освіти, і всі випускники, які проходили ЗНО, а це понад 116 тис., отримали сертифікати з результатами. З 2008 р. ЗНО є обов'язковою умовою вступу до ЗВО [47].

Упродовж 2000-2006 рр. проходив III етап програми Tempus, в якій все більше уваги приділялось на нові предметні області: аграрні науки, інформаційні й комп'ютерні технології, екологію. Шляхом міжнародної співпраці у закладах вищої освіти були створені офіси міжнародних відносин, які зараз продовжують свою діяльність.

З 2001 по 2004 рр. працює проект програми Tempus «Вивчення міжнародних фінансів в українських університетах». Протягом цього часу кожного року 4 студенти, 6 аспірантів, 8 викладачів стажувалися у зарубіжних закладах вищої освіти. Було розроблено 22 навчальних курси, підготовлено 2 підручники, один з них «Міжнародні фінанси» з грифом Міністерства освіти і науки України [48], крім того було проведено курси лекцій 8 іноземних професорів і створена магістерська програма з міжнародних фінансів.

У 2006 році публікується начальний посібник для викладачів, учителів і методистів І. Є. Булах «Створюємо якісний тест» [9] за підтримки програми TEMPUS/TACIS Європейської Комісії в межах проекту «Справедливе оцінювання».

В 2010 виходить навчально-методичний посібник І. В. Лупана та О. В. Авраменко, у якому наведено завдання, приклади та методичні рекомендації в роботі комп'ютерних статистичних пакетів. Показано у порівнянні інструментарій пакетів MSExcel, SPSS, Statistica [38]. А також посібник «Зовнішнє незалежне оцінювання в освіті України. Курс лекцій» у якому узагальнено вітчизняний та закордонний досвід впроваджування стандартизованого тестування, зародження, розвиток, організації та проведення ЗНО, опрацювання та представлення результатів тестування, структур, змісту та прикладів тестових завдань [49].

Подальший розвиток освіти зосереджено на проведені моніторингу якості освіти згідно до Указу Президента України від 30 вересня 2010 р. №926

(926/2010) «Про заходи щодо забезпечення пріоритетного розвитку освіти в Україні» [50].

З 2012 р. у рамках IV етапу проекту Tempus починається активна робота над реформуванням навчальних планів і програм для дисциплін спеціалізації «Освітні вимірювання». Насамперед – це роботи О. В. Авраменко, Д. С. Сільвестрова і О. Д. Борисенко по підготовці фахівців з освітніх вимірювань в Україні [51, 52].

Підручник О. В. Авраменко «Вимірювання в освіті» [53], в якому наведено теоретичні основи для конструювання тестів та тестових завдань, теорія з освітніх вимірювань, класичних та сучасних моделей тестування, комп’ютерних засобів тестування, моніторингу якості освіти в Україні та світі, а також методики навчання освітніх вимірювань.

Основам теорії педагогічного оцінювання, огляд традиційних та інноваційних підходів до оцінювання присвячена роботі Т. М. Канівець [54]. А основи теорії освітніх вимірювань для закладів вищої освіти розглянуто в роботах Ю. О Ковальчука [55] та Л. І. Лутченко [56].

Важливими є роботи О. В. Авраменко і Г. Ю. Павличенко щодо СТТ «Статистичні методи в освітніх вимірюваннях. Частина I. Класична теорія тестування» [57] і Т. Лісової щодо IRT в роботі «Моделі та методи сучасної теорії тестів» [58].

Крім того П. І. Андронатій, В. В. Котяк розробили посібник «Комп’ютерні технології в освітніх вимірюваннях» [59], який був призначений для систематизації знань з конструювання тестів та тестових завдань, комп’ютерних засобів тестування, а також методики навчання освітніх вимірювань.

Теоретико-методологічні основи і особливості методико-практичного застосування інтерсуб’єктних методів психологічного оцінювання особистості, почуттів і емоцій студентів наведені в роботі Г. В. Дьяконов [60].

1.3 Висновки за 1 розділом

Впровадження в освітній процес тестова система оцінки знань і вмінь учнів та студентів потребує компетентного аналізу якості створених тестів. Тому вміння створювати тести та обробляти їх результати залишається актуальним.

Розроблені підходи до обробки результатів тестування розвиваються у двох напрямках Classical Response Theory і Item Response Theory. Широко використовувана IRT теорія в NAEP, SAT, GRE все більше привертає увагу вітчизняних фахівців в галузі педагогічних та психологічний вимірювань. Удосконалюються методи розрахунку латентних параметрів і розвиваються теоретико-методичні засади створення та використання комп’ютерно-орієнтованих систем і засобів навчання.

2 ОСНОВНІ ЗАСАДИ ITEM RESPONSE THEORY

2.1 Теоретичні засади IRT теорії

Основою Item Response Theory є припущення про існування взаємозв'язку між результатами тестування і латентними параметрами, що показують здібності особистості, які недоступні для прямого спостереження. Зазвичай ці латентні параметри трактують як рівень підготовленості випробуваних і рівень складності завдання [58].

За допомогою застосування математико-статистичних моделей вимірювання оцінюються латентні параметри випробовуваних і параметрів завдань тесту. IRT як теорія тестів дозволяє вимірювати рівень досягнень випробовуваного в спеціальних одиницях виміру – логітах. А оскільки Г. Раш запропонував математичну модель зв'язку між латентними параметрами то параметри оцінюються в одній шкалі, що дає змогу співвіднести рівень знань будь-якого випробовуваного з мірою складності кожного завдання тесту.

Рівень підготовки випробовуваних позначають θ_i $(i = \overline{1, N})$, де i – номер того, хто проходить тест, N – кількість випробовуваних. Вважається, що кожному випробовуваному відповідає одне значення латентного параметра, який може змінюватися від $(-\infty, \infty)$. І чим кращий рівень підготовки, тим більше ймовірність правильної відповіді на питання тесту.

Геометрична інтерпретація зв'язку між отриманими оцінками і оцінками латентної характеристики θ (рівень підготовки) має вигляд монотонної та нелінійної S-подібної кривої, яка називається характеристичною кривою завдання [33].

На рисунку 2.1 зображено характеристичну криву деякого завдання по горизонталі якої рівень підготовки випробовуваних θ, а по вертикалі ймовірність правильної відповіді відповідно до латентної характеристики.

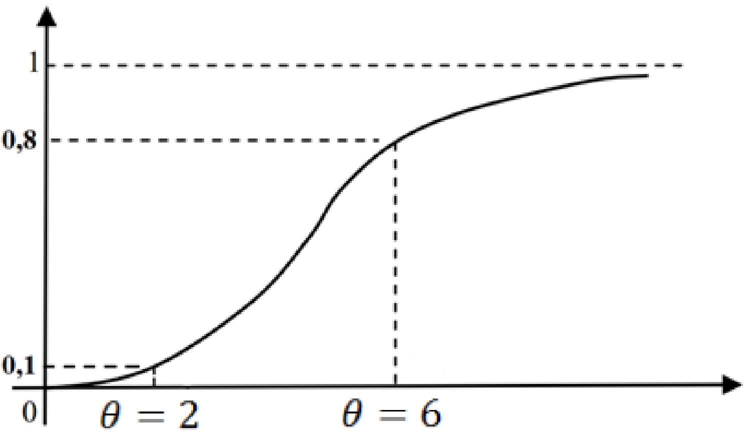

Рисунок 2.1 – Характеристична крива завдання

Аналізуючи рис. 2.1 можна сказати, що випробовуваний зі значення латентної характеристики $\theta = 2$, правильно відповість на питання із ймовірністю 0,1, а випробуваний із значенням $\theta = 6$ із ймовірністю 0,8. Тобто більшому значенню θ, відповідає більше ймовірність виконання.

Рівень складності j-го завдання позначають β_j $(j = \overline{1,n})$, де n – кількість завдань у тесті. Геометрична інтерпретація ймовірності відповіді на питання в залежності від рівня складності має вигляд монотонно-спадної та нелінійної кривої.

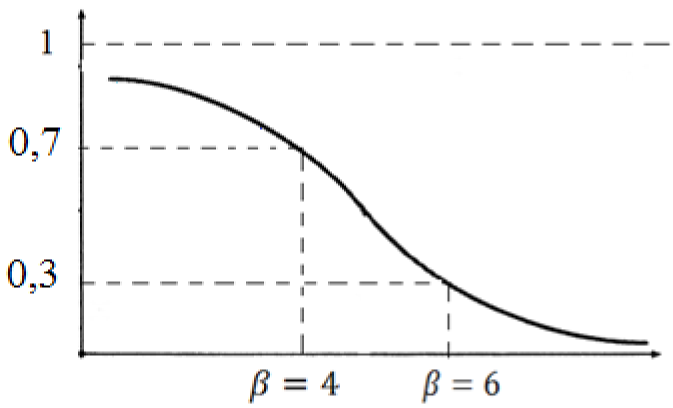

Рисунок 2.2 – Індивідуальна крива випробовуваного

Аналізуючи рис. 2.2 можна сказати, що при $\beta = 4$ ймовірність виконання завдання 0,7, а при $\beta = 6$ ймовірність рівна 0,3 тобто при більшому значенні складності завдання отримуємо меншу ймовірність виконання завдання тесту.

Якщо розглядати значення θ_i – як положення i-го випробуваного на шкалі логітів, а значення β_j – як положення j-го завдання на цій же шкалі, тоді $|\theta_i - \beta_j|$ – відстань на якій знаходиться випробовуваний з рівнем підготовки θ_i від завдання з складністю β_j. Якщо ця різниця додатна і велика, то завдання не підходить для контролю рівня знань випробовуваного, оскільки завдання такої складності є для випробовуваного занадто легким. Якщо ж ця різниця від'ємна і велика за модулем, то завдання не підходить для вимірювання рівня навчальних досягнень випробовуваного, оскільки ймовірність досягнення успіху i-го випробуваного в j-му завданні буде дуже низькою [36].

При застосуванні IRT розглядають умовну ймовірність P_i вірного виконання i-тим випробовуваним з рівнем підготовки θ_i, різних за складність завдань тесту. При цьому вважають θ_i параметром, β – незалежною змінною.

$$P_i\left\{x_{ij} = 1 \middle| \theta_i\right\} = f(\theta_i - \beta), i = \overline{1, N}, j = \overline{1, n}, \tag{2.1}$$

де $x_{ij} = \begin{cases} 1, & \text{якщо відповідь } i \text{ випробовуваного на } j \text{ завадання вірна;} \\ 0, & \text{якщо відповідь } i \text{ випробовуваного на } j \text{ завадання невірна.} \end{cases}$

Аналогічно вводиться P_j для позначення ймовірності правильного виконання j-го завдання складності β_j різними випробуваними, де θ – незалежна змінна, а β_j – параметр, що визначає складність j-го завдання тесту.

$$P_j\left\{x_{ij} = 1 \middle| \beta_j\right\} = \varphi\left(\theta - \beta_j\right), i = \overline{1, N}, j = \overline{1, n}. \tag{2.2}$$

У теорії IRT функції $f(\beta)$ і $\varphi(\theta)$ називаються Item Response Functions, а їх графіки відповідно індивідуальною кривою i-го випробуваного і характеристичною кривою j-го завдання.

2.2 Базові моделі IRT теорії: моделі Г. Раша та А. Бірнбаума

2.2.1 Однопараметрична модель Г. Раша

Найпростіша аналітична однопараметрична модель запропонована Г. Рашем [61] , яку часто називають логістичною моделлю, має вигляд:

$$P_j(\theta) = \frac{e^{1,7(\theta-\beta_j)}}{1+e^{1,7(\theta-\beta_j)}}, \tag{2.3}$$

$$P_i(\beta) = \frac{e^{1,7(\theta_i-\beta)}}{1+e^{1,7(\theta_i-\beta)}}, \tag{2.4}$$

де θ, β – незалежні змінні для відповідних функції, $i=\overline{1,N}$, $j=\overline{1,n}$, N – кількість випробовуваних, n – кількість завдань у тесті.

При використанні однопараметричної моделі Раша крутизна кривих завдань вважається однаковою для всіх кривих. Цей недолік простежується, коли потрібно віддати перевагу одному з завдань рівної складності. При проведенні аналізу можна зробити похибки в інтерпретації результатів, що знизить надійність і валідність тесту, позбувшись від завдань з більш крутими характеристичними кривими, залишивши з пологіші криві [62].

2.2.2 Двопараметрична модель А. Бірнбаума

Для вирішення цієї проблеми використовують двопараметричну модель А. Бірнбаума, яка включає в себе параметр диференційованої спроможності завдання тесту, що дозволяє розрізняти учасників тестування з різним рівнем навчальних досягнень.

Двопараметрична модель А.Бірнбаума для умовної ймовірності вірного виконання завдання тесту випробовуваними знаходиться за формулами:

$$P_j(\theta) = \frac{e^{1,7a_j(\theta-\beta_j)}}{1+e^{1,7a_j(\theta-\beta_j)}}, \tag{2.5}$$

$$P_i(\beta) = \frac{e^{1,7a_i(\theta_i-\beta\,)}}{1+e^{1,7a_i(\theta_i-\beta\,)}}, \tag{2.6}$$

де a_j – параметр характеристики диференційованої спроможності завдання при зміні різних значень θ, a_i – параметр, що вказує на міру структурованості знань випробовуваного, що проходить тест.

При геометричній інтерпретації формули (2.5) параметр β_j розглядається як характеристику положення кривої j-го завдання відносно осі θ.

Другий параметр a_j пов'язаний з крутизною кривої завдання в точці її перегину, він прямо пропорційне тангенсу кута нахилу дотичної до характеристичної кривої завдання тесту в точці $\theta = \beta_j$. А це означає, що більш круті криві відповідають більшим значенням a_j, відповідно для пологих кривих $a_j \to 0$ [36].

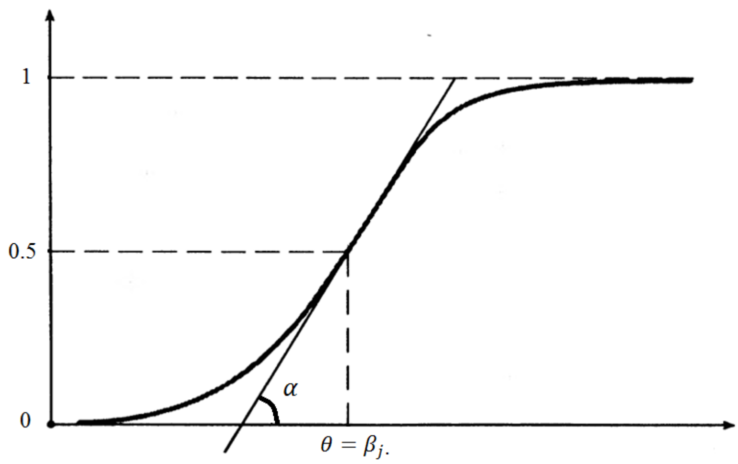

Рисунок 2.3 – Характеристична крива j-го завдання

Для порівняльної характеристики розглянемо характеристичні криві трьох завдань однокової складності з різною крутизною (рис.2.4).

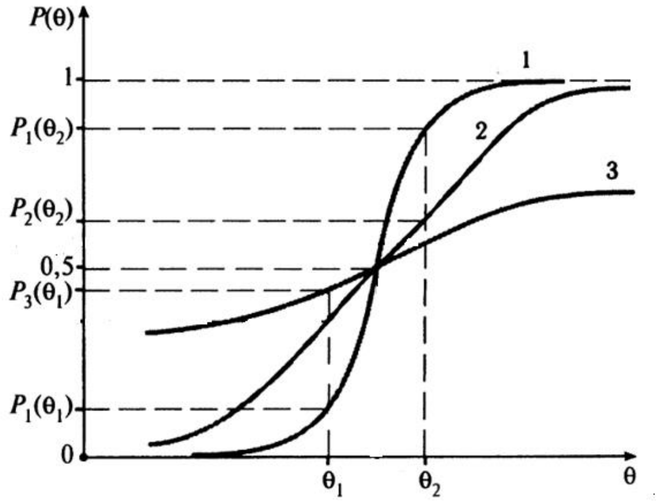

Рисунок 2.4 – Характеристичні криві трьох завдань рівної складності

Крива 1 найбільш крута і відповідає більшому значенню a_1, а крива 3 найбільш пологa, що означає $a_3 \rightarrow 0$. Для випробовуваних з рівнем підготовки θ_1 і θ_2 розташованих на осі θ по різні сторони від точки перегину кривих ймовірність вірного виконання 1 завдання тесту відрізняються $P_1(\theta_2) > P_1(\theta_1)$.

Таким чином, значення a_j, які близькі до нуля, відповідають випадку, коли випробовувані з різним рівнем підготовки вірно відповідають на j-е завдання з приблизно рівною ймовірністю, що відповідно суперечить очікуваним прогнозам розробника тесту. Ці завдання виявляються непотрібними при диференціації випробовуваних групи за оцінюваним параметром, тому що не несуть інформації про індивідуальні відмінності особистості.

Також є непродуктивними завдання з від’ємними значеннями a_i, на них відповідають вірно з більшою ймовірністю випробовувані з низьким рівнем підготовки, а для тих, хто має кращу підготовку, значення θ ймовірності вірної відповіді буде прагнути до нуля. Тому число завдань в тесті має зменшуватися в першу чергу через видалення таких завдань навіть в тому випадку, коли інші їх характеристики підходять розробнику тесту. Як правило таке скорочення приводить до підвищення надійності і валідності тесту.

Проведений аналіз виявляє роль a_j при диференціації випробовуваних. Відповідно цьому параметр a_j отримав назву диференційованої спроможності j-го завдання тесту, який можна отримати за наступною формулою [33]:

$$a_j = \frac{(r_{bis})_j}{\sqrt{1-(r_{bis})_j^2}}, \tag{2.7}$$

де $(r_{bis})_j$ – бісеріальний коефіцієнт кореляції, який знаходиться за формулою:

$$(r_{bis})_j = \frac{(\bar{X}_1)_j \ - (\bar{X}_0)_j}{S_x} \cdot \frac{(N_1)_j - (N_0)_j}{uN\sqrt{N^2 - N}}, \tag{2.8}$$

де $(\bar{X}_1)_j$ – середнє значення індивідуальних балів випробовуваних, які виконали вірно j-те завдання тесту, $(\bar{X}_0)_j$ – середнє значення індивідуальних балів випробовуваних, які виконали невірно j-те завдання тесту, S_x – стандартне відхилення за множиною значень індивідуальних балів, $(N_1)_j$ – число випробовуваних, які виконали вірно j-те завдання, $(N_0)_j$ – число випробовуваних, які виконали невірно j-те завдання, N – загальна кількість випробовуваних $(N = N_1 + N_0)$, u – ордината нормованого нормального розподілу в точці, за якої лежить $100 \cdot \frac{N_1}{N}\%$ площі під нормальною кривою.

Теоретично значення параметра a_j можуть змінюватися в інтервалі $(-\infty, +\infty)$, але на практиці як правило залишають завдання зі значеннями a_j, які належать інтервалу $(0{,}5;\ 0{,}25)$ [36].

2.2.3 Трипараметрична модель А. Бірнбаума

Дана модель враховує ймовірність вірної відповіді в тому випадку, якщо відповідь була вгадана, а не основана на знаннях випробовуваного.

Трипараметрична логістична модель А. Бірнбаума ймовірності вірної відповіді випробовуваним на j-е завдання тесту знаходиться за формулою:

$$P_j\left\{x_{ij}=1 \middle| \beta_j\right\}=c_j+\left(1-c_j\right)\frac{e^{1,7a_j(\theta-\beta_j)}}{1+e^{1,7a_j(\theta-\beta_j)}}, \qquad (2.9)$$

де c_j – параметр вгадування.

Параметр c_j визначається кількістю відповідей до закритих завдань тесту. Наприклад, для завдання з п'ятьма відповідями за класичною теорією ймовірності $c_j = 0{,}2$, при чотирьох запропонованих відповідях $c_j = 0{,}25$.

Характеристична крива j-го завдання тесту у випадку трипараметричної моделі зображена на рисунку 2.5. При наявності параметра c_j характеристична крива стає більш пологою, ніж у випадку з характеристичною кривою, яка має таку ж точку перегину, але в якої нижня асимптота є вісь θ ($c_j = 0$). Таким чином, ефект вгадування знижує диференційовані здатність завдань тесту.

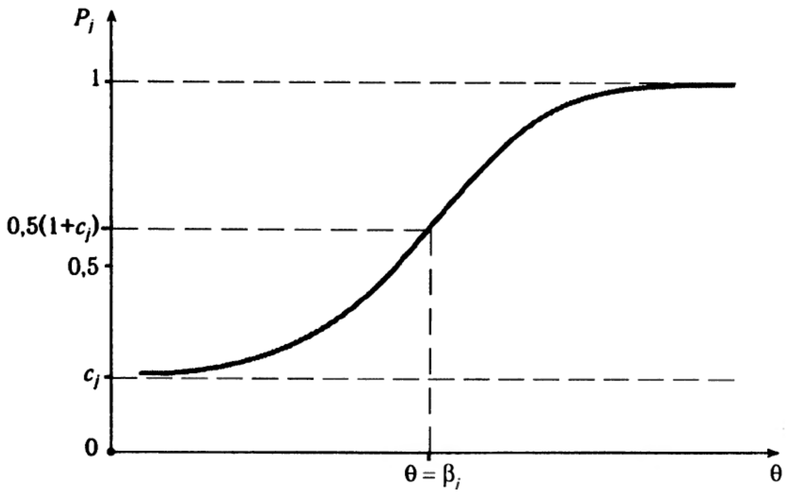

Рисунок 2.5 – Характеристична крива j-го завдання тесту (трипараметрична модель)

Застосування трипараметричної моделі значно ускладнює аналіз і обробку статистичних даних в процесі конструювання тесту.

Введення параметра c_j не тільки істотно знижує оцінки параметрів параметра θ і β, але і погіршує збіжність ітераційних методів, які застосовуються для підвищення точності оцінки латентних змінних θ і β.

2.3 Політомічні моделі: моделі Раша-Мастерса, Е. Андерсена та модель Тіссена-Стейнберга

2.3.1 Політомічна модель Раша-Мастерса

Разом з вище викладеними моделями існує політомічна модель Раша-Мастерса або Partial Credit Model. PCM використовується для аналізу тестових робіт, що містять політомічні завдання і заснована на основі базової моделі Г. Раша, яку детально вивчав Дж. Мастерс (G. Masters) у своїй роботі [63].

Розглянемо тест, що містить L завдань і виконується N випробовуваними. Кожне i-те завдання має m_i підрівнів. Тобто за i-е завдання випробовуваний може отримати $j = 0, 1, \dots, m_i$ балів. Будемо казати, що в такому випадку i-е завдання матиме $m_i + 1$ рівнів, кожному з яких відповідає число j. Якщо $m_i > 1$, то таке завдання називається політомічним [64].

У PCM дві множини латентних параметрів приймаються у іншому вигляді: θ_n – параметри підготовки n-го випробовуваного, де $n = \overline{1, N}$, N – кількість випробовуваних. β_{ij} – параметри складності j-го рівня i-го завдання, де $i = \overline{1, L}$, $j = \overline{0, m_i}$, m_i – максимальний рівень i-го завдання.

Параметр складності β_{ij} визначає складність переходу із $(j-1)$-го рівня i-го завдання на j-й і чим більша ця складність, тим більше значення β_{ij} [65].

У політомічній моделі Раша-Мастерса ймовірність отримання n-м випробовуваним j балів за i-те завдання означується таким чином:

$$P_{ij}(\theta_n) = \frac{e^{\sum_{l=0}^{j}(\theta_n - \beta_{il})}}{\sum_{k=0}^{m_i} e^{\sum_{l=0}^{k}(\theta_n - \beta_{il})}}. \tag{2.10}$$

Оцінювання відповідних латентних параметрів можна отримати шляхом розв’язання нелінійної системи рівнянь [65]:

$$\begin{cases} r_n - \sum_{i=1}^{L}\sum_{k=1}^{m_i} kP_{ik}(\theta_n) = 0, \\ -S_{ij} + \sum_{n=1}^{N}\sum_{k=j}^{m_i} P_{ik}(\theta_n) = 0, \end{cases} \tag{2.11}$$

де S_{ij} – кількість випробовуваних, які отримали за i-те завдання не менш як j балів, а $r_n = \sum_{i=1}^{L} x_{ni}$, $n = \overline{1, N}, i = \overline{1, L}, j = \overline{1, m_\iota}$.

Один із методів розв’язання системи (2.11) є метод Ньютона-Рафсона (детальніше про цей метод описано в книзі J. Raphson [66] і Б.Демидовича [67]), який дає такі ітераційні формули:

$$\theta_n^{(t+1)} = \theta_n^{(t)} - \frac{r_n - \sum_{i=1}^{L}\sum_{k=1}^{m_i} kP_{ik}(\theta_n)}{-\sum_{i=1}^{L}\left[\sum_{k=1}^{m_i} k^2 P_{ik}(\theta_n) - \left(\sum_{k=1}^{m_i} kP_{ik}(\theta_n)\right)^2\right]}, \tag{2.12}$$

$$\beta_{il}^{(t+1)} = \beta_{il}^{(t)} - \frac{-S_{ij} + \sum_{n=1}^{N}\sum_{k=j}^{m_i} P_{ik}(\theta_n)}{-\sum_{n=1}^{N}\left[\sum_{k=j}^{m_i} P_{ik}(\theta_n) - \left(\sum_{k=j}^{m_i} P_{ik}(\theta_n)\right)^2\right]}, \tag{2.13}$$

де $n = \overline{1, N}, i = \overline{1, L}, j = \overline{1, m_\iota}$, за такої умови завершення ітераційного процесу:

$$\sqrt{\sum_{n=1}^{N}\left(\theta_n^{(t)} - \theta_n^{(t-1)}\right)^2 + \sum_{i=1}^{L}\sum_{j=1}^{m_i}\left(\beta_{ij}^{(t)} - \beta_{ij}^{(t-1)}\right)^2} < \varepsilon, \tag{2.14}$$

де ε – заздалегідь задана точність обчислення.

Залежності ймовірностей від рівня підготовленості випробовуваного зображують за допомогою характеристичних кривих рівнів завдання.

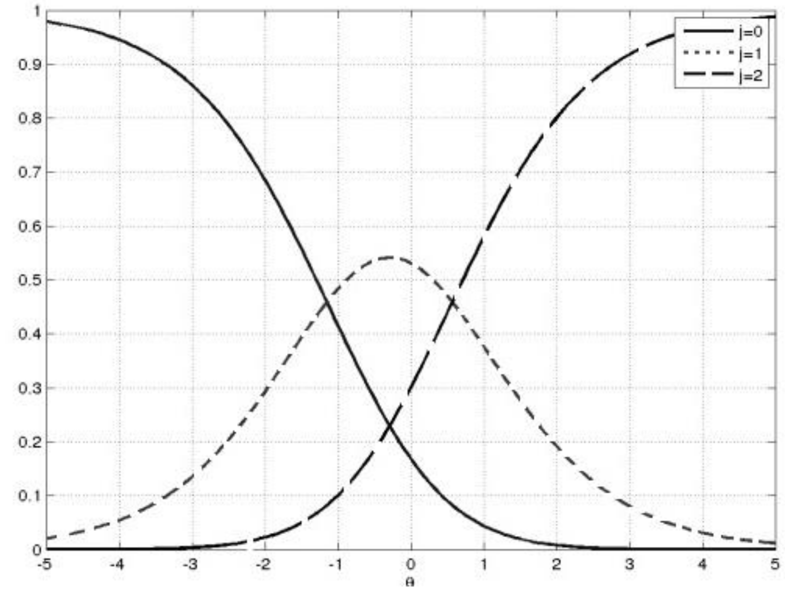

Рисунок 2.6 – Характеристичні криві двох рівнів політомічного завдання за моделлю Раша-Мастерса

2.3.2 Політомічна модель Е. Андерсена

Розглянута вище модель Раша-Мастерса є частинним випадком політомічної моделі Е. Андерсена (E. Andersen). Для означення політомічної моделі Андерсена введемо такі множини параметрів: θ_n, $n = \overline{1, N}$ – параметр підготовленості n-го випробовуваного, параметр складності j-го рівня i-го завдання позначатимемо η_{ij}, $i = \overline{1, L}$, $j = \overline{1, m_\iota}$, $\eta_{i0} \equiv 0$ та бали, що даються випробовуваним за досягнення j-го рівня i-го завдання позначатимемо a_{ij}, $i = \overline{1, L}, j = \overline{1, m_\iota}$. Причому параметри a_{ij} мають утворювати арифметичну прогресію $d = a_{ij} - a_{i(j-1)}$, $i = \overline{1, L}, j = \overline{2, m_\iota}$ [68].

Тоді ймовірність досягнення n-м випробовуваним j-го рівня за i-е завдання визначається як

$$P_{ij}(\theta_n) = \frac{e^{(a_{ix}\theta_n - \eta_{ix})}}{\sum_{k=0}^{m_i} e^{(a_{ik}\theta_n - \eta_{ik})}}. \tag{2.15}$$

У випадку коли $a_{ij} = j,\ i = \overline{1, L}, j = \overline{0, m_\iota}$ та $\eta_{ij} = \sum_{k=1}^{j} \beta_{ik},\ i = \overline{1, L},\ j = \overline{1, m_\iota}$ модель Андерсена збігається з моделлю Раша-Мастерса.

Для тривіального рівня $e^{-\eta_{ij}} \equiv 0, a_{ij} \equiv 0$. Окрім того $a_{ib_i} \equiv 0, \eta_{ib_i} \equiv 0$, де b_i – порядковий номер найнижчого нетривіального рівня i-го завдання. Модель Андерсена охоплює випадок появи тривіальних стовпців у моделі Раша-Мастерса. Значення параметрів знаходять використовуючи систему [69]:

$$\begin{cases} \sum_{i=1}^{L} a_{ix_{ni}} - \sum_{i=1}^{L} \sum_{k=0}^{m_i} a_{ik} P_{ik}(\theta_n) = 0, \\ -R_{ij} + \sum_{n=1}^{N} P_{ij}(\theta_n) = 0\,, \end{cases} \tag{2.16}$$

де $n = \overline{1, N},\ i = \overline{1, L},\ j = \overline{1, m_\iota},\ R_{ij}$ – кількість випробовуваних, які в i-му завданні здобули результат j, x_{ni} – кількість n-го випробовуваного за i-те завдання.

Відповідні ітераційні формули оцінювання параметрів набувають вигляду:

$$\theta_n^{(t+1)} = \theta_n^{(t)} - \frac{\sum_{i=1}^{L} a_{ix_{ni}} - \sum_{i=1}^{L} \sum_{k=0}^{m_i} a_{ik} P_{ik}(\theta_n)}{-\sum_{i=1}^{L} \left[\sum_{k=0}^{m_i} a^2{}_{ik} P_{ik}(\theta_n) - \left(\sum_{k=0}^{m_i} a_{ik} P_{ik}(\theta_n) \right)^2 \right]}, \tag{2.17}$$

$$\eta_{ij}^{(t+1)} = \eta_{ij}^{(t)} - \frac{-R_{ij} + \sum_{n=1}^{N} P_{ij}(\theta_n)}{-\sum_{n=1}^{N} \left(P_{ij}(\theta_n) - \left(P_{ij}(\theta_n) \right)^2 \right)}, \tag{2.18}$$

де $n = \overline{1, N},\ i = \overline{1, L},\ j = \overline{1, m_\iota}$.

За такої умови збіжності ітераційного процесу:

$$\sqrt{\sum_{n=1}^{N}\left(\theta_n^{(t)}-\theta_n^{(t-1)}\right)^2+\sum_{i=1}^{L}\sum_{j=1}^{m_i}\left(\eta_{ij}^{(t)}-\eta_{ij}^{(t-1)}\right)^2} < \varepsilon, \qquad (2.19)$$

де ε – заздалегідь задана точність обчислення.

2.3.3 Модель Тіссена-Стейнберга

В 1984 році Д. Тіссен (D. Thissen) і Л. Стейнберг (L. Steinberg) у своїй роботі «Response model for multiple choice items» [70] представили модель для аналізу якості завдань з множинним вибором.

Для означення моделі Тіссена-Стейнберга введемо такі множини параметрів: a_{jk} – диференціююча спроможність k-го варіанту відповіді j-го завдання, c_{jk} – складність k-го варіанту відповіді j-го завдання, d_{jk} – відносна вага k-ї відповіді на j-те завдання або ймовірність вибору k-го варіанту відповіді j-го завдання, якщо відповідь була вгадана, а не основана на знаннях випробовуваного, θ_i – рівень підготовленості i-го випробовуваного.

Відповідно до модель Тіссена-Стейнберга j-те завдання тесту має h_j варіантів відповіді j-го завдання, а ймовірність вибору i-м випробовуваним k-го варіанту j-го завдання знаходиться за формулою:

$$P(k,\theta_i)=\frac{e^{(a_{jk}\theta_i+c_{jk})}+d_{jk}e^{(a_{j0}\theta_i+c_{j0})}}{\sum_{s=0}^{h_j}e^{(a_{js}\theta_i+c_{js})}}, \qquad (2.20)$$

де $i=\overline{1,N}$, $j=\overline{1,K}$, N – кількість випробовуваних, K – загальна кількість завдань у тесті.

Геометрична інтерпретація залежності ймовірностей від рівня підготовленості випробовуваного зображують за допомогою характеристичних кривих варіантів відповіді (рис. 2.7) [71].

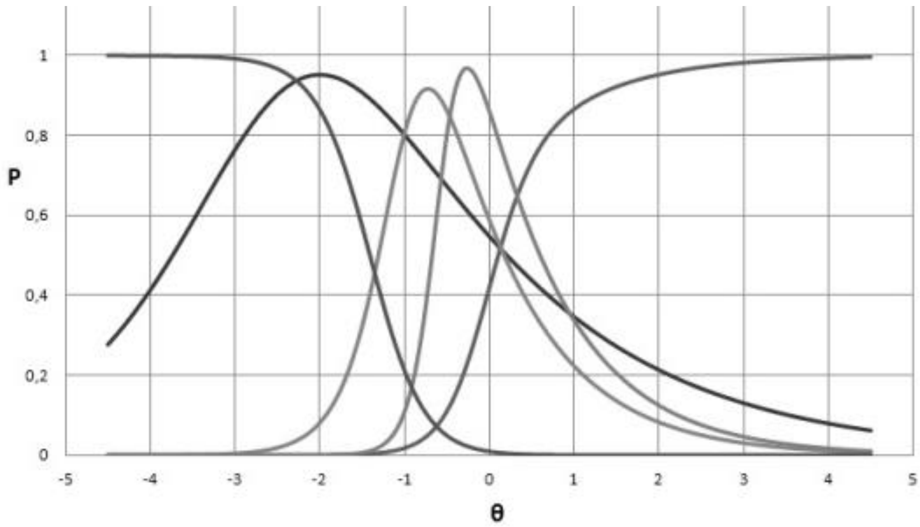

Рисунок 2.7 – Характеристичні криві варіантів відповіді завдання множинного вибору за моделлю Тіссена-Стейнберга

Будь-яке політомічне завдання може бути розглянуто як багатокрокове. Щоб досягти будь-якої категорії k, випробовуваний повинен послідовно подолання k кроків, $k = \overline{0,m}$, m – загальна кількість кроків виконання завдання. За правильне виконання кожного кроку випробуваний отримує 1 бал. Отже, загальний бал за виконання завдання дорівнює кількості правильно виконаних кроків [72].

В якості математичної моделі тестування виберемо модель Раша з довільними проміжними категоріями виконання завдань [73]. Відповідно до цієї моделі ймовірність P_{nk} того, що випробовуваний n з рівнем підготовки θ_n отримає k балів за виконання завдань знаходиться за формулою:

$$P_{nk} = \frac{e^{(k\theta_n - \sum_{j=0}^{k} \beta_j)}}{\sum_{l=0}^{m} e^{(l\theta_n - \sum_{j=0}^{l} \beta_j)}}, \tag{2.21}$$

де β_j – складність виконання j-го кроку завдання, $j = \overline{1,m}$, $\beta_0 = 0$.

Однак складність досягнення наступної категорії завдання вираховується за умови, що попередня категорія вже досягнута.

Подібна модель дозволяє кожному завданню тесту мати свою шкалу оцінювання, за виконання якого випробовуваний може отримати від 0 до m балів, де m може змінюватися в різних завданнях тесту. Тоді загальна складність завдань $\bar{\beta} = \frac{1}{m}\sum_{j=1}^{m}\beta_j$ і може бути використано для порівняння різних завдань між собою.

Розглянемо випадок коли складності кроків завдання β_j, де $j = \overline{1,m}$ відомі і фіксовані. Тоді ймовірність є функцією параметра θ (рівня підготовленості випробуваного) і визначається формулою:

$$P_k = \frac{e^{(k\theta - \sum_{j=0}^{k}\beta_j)}}{\sum_{l=0}^{m} e^{(l\theta - \sum_{j=0}^{l}\beta_j)}}, \qquad (2.22)$$

де $k = \overline{0,m}$.

Нехай a_n – бал отриманий n випробовуваним за виконання завдання. Якщо розглядати цю величину як дискретну випадкову величину з можливими значеннями $0,1,..m$ і ймовірностями $P_0,\ P_1, \dots, P_m$ то математичне очікування і дисперсія знаходяться за відповідними формулами [72]:

$$M(a_n) = \sum_{k=1}^{m} kP_k, \qquad (2.23)$$

$$D(a_n) = \sum_{k=0}^{m}\left(k - \sum_{l=1}^{m} l \cdot P_l\right)^2 \cdot P_k = \sum_{k=0}^{m} k^2 \cdot P_k - \left(\sum_{k=1}^{m} k \cdot P_k\right)^2. \qquad (2.24)$$

Характеристичною функцією політомічного завдання назвемо функцію яка може бути інтерпретована як функція математичного очікуванням бала за виконання завдання випробуваними з різним рівнем підготовленості. Вона обчислюється за формулою:

$$f(\theta) = M(a_n) = \sum_{k=1}^{m} k \cdot P_k. \qquad (2.25)$$

Характеристичною функцією шага політомічного завдання назвемо функцією ймовірності виконання випробовуваними з різним рівнем підготовки k-го кроку в завданні і знаходиться за формулою [73]:

$$f_k(\theta) = \frac{e^{(\theta-\beta_k)}}{1+e^{(\theta-\beta_k)}}, \qquad (2.26)$$

де β_k – складність виконання k-го кроку завдання, $k = \overline{1,m}$.

При розгляданні найпростішого політомічного двокрокового завдання ймовірність отримати 0,1,2 бали за виконання завдань позначається відповідно $P_0,\ P_1,\ P_2$. В такому випадку формула (2.22) при $m = 2$ має вигляд:

$$P_0 = \frac{1}{1 + e^{(\theta-\beta_1)} + e^{(2\theta-\beta_1-\beta_2)}}, \qquad (2.27)$$

$$P_1 = \frac{e^{(\theta-\beta_1)}}{1 + e^{(\theta-\beta_1)} + e^{(2\theta-\beta_1-\beta_2)}}, \qquad (2.28)$$

$$P_2 = \frac{e^{(2\theta-\beta_1-\beta_2)}}{1 + e^{(\theta-\beta_1)} + e^{(2\theta-\beta_1-\beta_2)}}, \qquad (2.29)$$

де β_1, β_2 – складність кроків розглянутого завдання.

В ЗНО використовують політомічні чотирикрокові завдання, тому звернемо увагу і на них. Ймовірність отримати $0, 1, 2, 3, 4$ бали за виконання завдань позначимо відповідно $P_0,\ P_1,\ P_2,\ P_3,\ P_4$ і знаходиться за формулами:

$$P_0 = \frac{1}{1 + e^{(\theta-\beta_1)} + e^{(2\theta-\beta_1-\beta_2)} + e^{(3\theta-\beta_1-\beta_2-\beta_3)} + e^{(4\theta-\beta_1-\beta_2-\beta_3-\beta_4)}}, \qquad (2.30)$$

$$P_1 = \frac{e^{(\theta-\beta_1)}}{1 + e^{(\theta-\beta_1)} + e^{(2\theta-\beta_1-\beta_2)} + e^{(3\theta-\beta_1-\beta_2-\beta_3)} + e^{(4\theta-\beta_1-\beta_2-\beta_3-\beta_4)}}, \qquad (2.31)$$

$$P_2 = \frac{e^{(2\theta-\beta_1-\beta_2)}}{1 + e^{(\theta-\beta_1)} + e^{(2\theta-\beta_1-\beta_2)} + e^{(3\theta-\beta_1-\beta_2-\beta_3)} + e^{(4\theta-\beta_1-\beta_2-\beta_3-\beta_4)}}, \qquad (2.32)$$

$$P_3 = \frac{e^{(3\theta-\beta_1-\beta_2-\beta_3)}}{1+e^{(\theta-\beta_1)}+e^{(2\theta-\beta_1-\beta_2)}+e^{(3\theta-\beta_1-\beta_2-\beta_3)}+e^{(4\theta-\beta_1-\beta_2-\beta_3-\beta_4)}}, (2.33)$$

$$P_4 = \frac{e^{(4\theta-\beta_1-\beta_2-\beta_3-\beta_4)}}{1+e^{(\theta-\beta_1)}+e^{(2\theta-\beta_1-\beta_2)}+e^{(3\theta-\beta_1-\beta_2-\beta_3)}+e^{(4\theta-\beta_1-\beta_2-\beta_3-\beta_4)}}, (2.34)$$

де β_1, β_2, β_3, β_4 – складність кроків розглянутого завдання.

2.4 Інформаційні функції тесту

Одним з компонентів аналізу якості тесту у IRT теорії є дослідження інформаційної функції для оцінювання його ефективності, яке дає підставу для постановки питання про створення тесту, який би найбільш оптимально відповідав заданому рівню підготовки групи випробовуваних [64].

Поняття інформаційної функції і загальної формули інформаційної функції для дихотомічних завдань було введене у 1968 році у роботі А. Бірнбаума «Some Latent Trait Models and Their Use in Inferring an Examinee's Ability» [74]. У 1969 році у роботі Ф. Самеджима (F. Samejima) «Estimation of Latent Ability Using a Response Pattern of Graded Scores» [75] була введена формула інформаційної функції для політомічних завдань. А в 1993 році в роботі Е. Муракі (E. Muraki) «Information Functions of the Generalized Partial Credit Model» [76] досліджувалась інформаційна функція для узагальненої Partial Credit моделі.

Інформаційна функція базується на понятті кількості інформації, яка забезпечується оцінюванням параметра рівня підготовки θ за j-те завдання і знаходиться за формулою [64]:

$$I_j(\theta) = \frac{\left(E_j'(\theta)\right)^2}{\sigma_j^2(\theta)}, \qquad (2.35)$$

де $j = \overline{1,K}, K$ − кількість завдань в тесті, $E_j'(\theta)$ − похідна математичного сподівання оцінки параметра θ, $\sigma_j(\theta)$− стандартне відхилення параметра θ.

Інформаційна функція дихотомічного завдання можна записати формулою:

$$I_j(\theta) = \frac{\left(P_j'(\theta)\right)^2}{P_j(\theta) \cdot Q_j(\theta)} = \frac{\left(P_j'(\theta)\right)^2}{P_j(\theta) \cdot (1 - P_j(\theta))}, \qquad (2.36)$$

де $P_j(\theta)$ – ймовірність вірної відповіді на j-те завдання, $Q_j(\theta)$ – ймовірність неправильної відповіді на j-те завдання тесту, θ – параметр підготовки випробовуваного, $j = \overline{1,K},\ K$ − кількість завдань в тесті [58].

Ймовірність вірної відповіді на j-те завдання $P_j(\theta)$ залежить від обраної моделі, якщо розглядати однопараметичну модель Г. Раша то $P_j' = P_j(1 - P_j)$ тому формула (2.35) приймає вигляд :

$$I_j(\theta) = P_j(\theta) \cdot \left(1 - P_j(\theta)\right) = P_j(\theta) \cdot Q_j(\theta), \qquad (2.37)$$

де $j = \overline{1,K}$.

Для двопараметричної моделі А. Бірнбаума інформаційна функція набуває вигляду:

$$I_j(\theta) = a_j^2 \cdot P_j(\theta) \cdot Q_j(\theta), \qquad (2.38)$$

де a_j – диференціююча спроможність j-го завдання тесту, $P_j(\theta) = \frac{e^{\left(\theta - \beta_j\right)}}{1 + e^{\left(\theta - \beta_j\right)}}$ – ймовірність вірної відповіді на j-те завдання, $j = \overline{1,K}$.

Розглянемо на прикладі інформаційну функцію двох тестів (криві 4) з трьох завдань зображених на рисунку 2.8 [58].

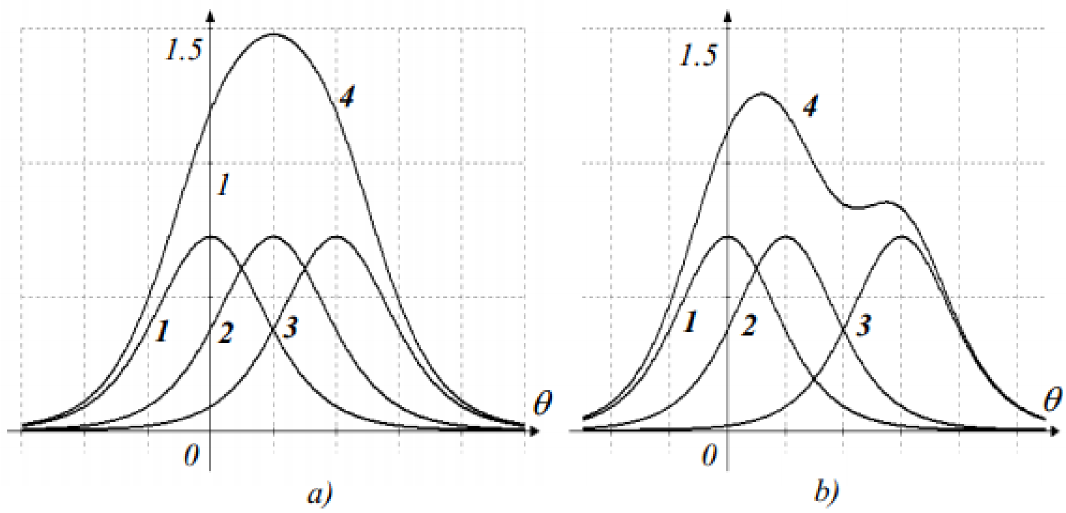

Рисунок 2.8 – Інформаційні функції двох тестів з трьома завданнями

На рисунку (a) завдання тесту мають рівномірно розподілену складність вздовж осі θ і інформаційна функція має один виражений екстремум. У другому тесті рисунок (b) інформаційна функція має «провал». Причиною цього є третє завдання, яке має значно більшу складність, ніж інші два. Інформаційна функція гарного збалансованого тесту повинна мати один чітко виражений екстремум, щоб це зробити в тесті (b) потрібно додати завдання проміжної складності, або змінити складність існуючих завдань [58].

У випадку, коли графік інформаційної функції має пологий, не чітко виражений екстремум, говорять про зниження ефективності усього тесту. Якщо існує кілька локальних екстремумів тоді тест потребує вдосконалення.

Для означення інформаційної функції політомічних завдань означимо такі множини параметрів: m_j – кількість рівнів у j-му завданні, θ – параметр підготовки випробовуваного, $P_{lj}(\theta)$ – ймовірність досягнення l-го рівня випробовуваним з рівнем підготовки θ у j-му завданні тесту, $l = \overline{1, m_j}, j = \overline{1, K}$.

Відповідно до роботи Ф. Самеджими над політомічними завданнями інформаційну функцію j-го завдання знаходять за формулою [75]:

$$I_j(\theta) = \sum_{l=0}^{m_j} \frac{\left(P'_{lj}(\theta)\right)^2}{P_{lj}(\theta)}. \qquad (2.39)$$

Інформаційна функція для моделі Раша-Мастерса знаходиться за формулою [76] :

$$I_j(\theta) = \sum_{l=0}^{m_j} \left[l - \sum_{k=0}^{m_j} k P_{kj}(\theta) \right]^2 \cdot P_{lj}(\theta), \qquad (2.40)$$

де

$$P_{jg}(\theta) = \frac{e^{\sum_{l=0}^{g}(\theta-\beta_{jl})}}{\sum_{k=0}^{m_j} e^{\sum_{l=0}^{k}(\theta-\beta_{jl})}}, \qquad (2.41)$$

де θ – параметр підготовки випробовуваного, β_{jg} – складність переходу з $(g-1)$-го рівня j-го завдання на наступний, m_j – кількість рівнів у j-му завданні, $j = \overline{1, K}, K$ – кількість завдань в тесті.

Формула (2.38) для моделі Тіссена-Стейнберга набуває вигляду [64]:

$$I_j(\theta) = \frac{\left(e_{jh}(a_{jh}\sum_{k=0}^{m_j} e_{jk} - \sum_{k=0}^{m_j} a_{jk}e_{jk}) + d_{jh}e_{j0}(a_{j0}\sum_{k=0}^{m_j} e_{jk} - \sum_{k=0}^{m_j} a_{jk}e_{jk})\ \right)^2}{\left(e_{jh} + d_{jh}e_{j0}\right)\left(\sum_{k=0}^{m_j} e_{jk}\right)^3},$$

де $e_k = a_{jk}\theta + c_{jk}\theta,\ k = \overline{0, m_j},\ m_j$ – кількість категорій у j-му завданні, $j = \overline{1, K}$ $h = \overline{1, m_j},\ a_{jk}$ – диференціюоча спроможність k-го варіанту відповіді j-го завдання, c_{jk}– складність k-го варіанту відповіді j-го завдання, d_{jk} – відносна вага k-ї відповіді на j-те завдання або ймовірність вибору k-го варіанту відповіді j-го завдання, якщо відповідь була вгадана, а не основана на знаннях випробовуваного, θ – рівень підготовленості випробовуваного.

2.5 Висновки за 2 розділом

Дослідження основних теоретичних засад Item Response Theory дає змогу дійти до висновку про доцільність використання цієї теорії не тільки для підвищення точності створення збалансованих і якісних тестів, а й для аналізу підготовки випробовуваних. Розглянуті базові та політомічні моделі IRT теорії такі, як однопараметрична модель Г. Раша, двопараметрична та трипараметрична моделі А. Бірнбаума, політомічна модель Раша-Мастерса, політомічна модель Е. Андерсенена та модель Тіссена-Стейнберга для завдань множинного вибору, кожна з яких має свої особливості, дають змогу побудувати ансамбль характеристичних кривих за допомогою відповідних формул ймовірності для подальшої інтерпретація результатів оброблених даних.

3 АНАЛІЗ ПІДГОТОВКИ АБІУРІЕНТІВ ЗНУ ЗАСОБАМИ IRT

3.1 Обробка і аналіз даних засобами IRT теорії

Першим кроком для аналізу підготовки абітурієнтів за допомогою тестування є перевірка якості тестових завдань і тесту в цілому. Оцінка якості тесту відбувається на основі аналізу розподілу вибірки результатів, знайдених значень коефіцієнтів надійності, критеріальної валідності тесту та графіку інформаційної функції. На рис. 3.1 зображена схема за допомогою якою проводиться аналіз тесту, тут: НД – надійність, ВЛД – валідність, ЕФЕКТ – ефективність; Н.РЗП.ІБ – нормальний розподіл індивідуальних балів і висновки щодо аналізу якості тесту [64].

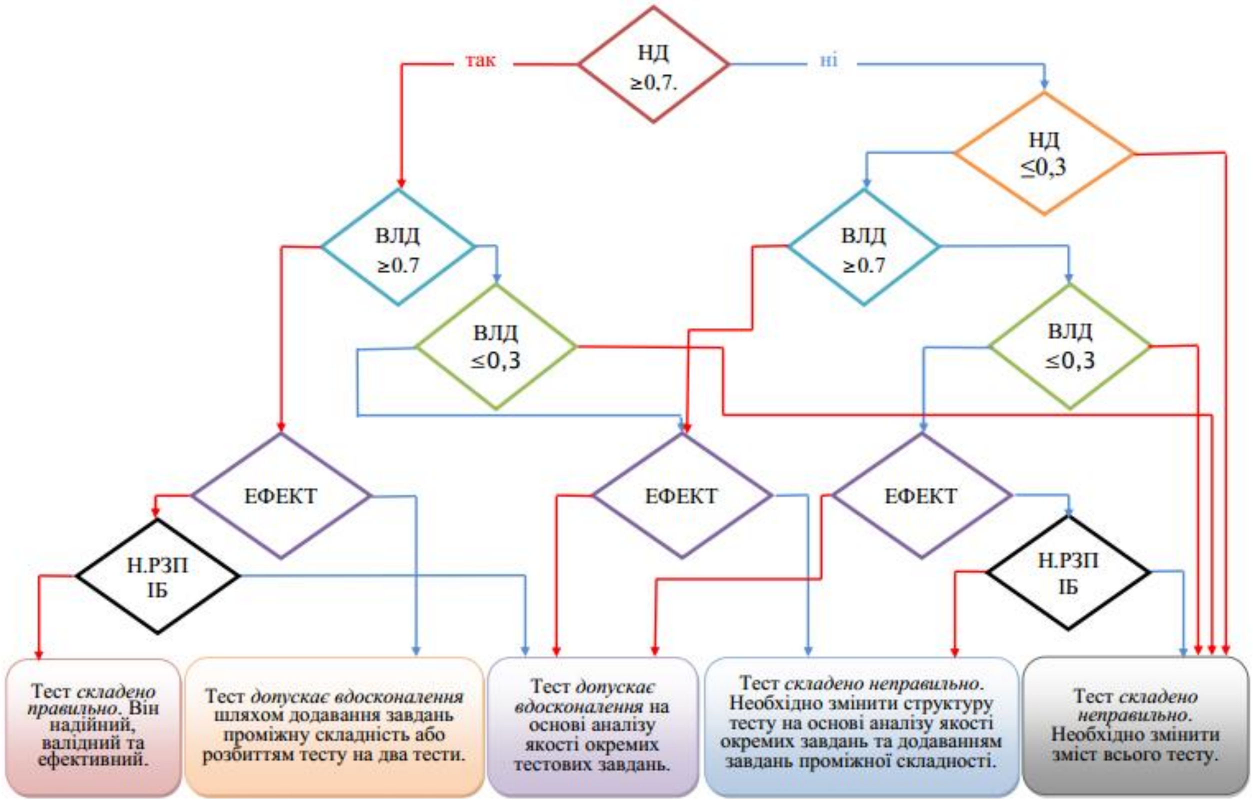

Рисунок 3.1 – Оцінювання якості тесту в цілому

На рис. 3.2 зображена схема для оцінювання якості окремих тестових завдань, яка здійснюється на основі аналізу кореляційної матриці завдань, ансамблю характеристичних кривих завдань, інформаційної функції та характеристичних кривих рівнів політомічних завдань, дистракторів завдань множинного вибору та значень коефіцієнтів валідності завдань.

Тут ЛЕГК/СКЛАД – завдання є занадто складним або занадто легким, $R \notin (0; 0{,}3]$ – коефіцієнти кореляції завдання з більшістю інших завдань не належать проміжку (0;0,3]; ПОЛІТ – завдання є політомічним, МНЖН ВИБІР – завдання множинного вибору; ДИСТР НЕ СПРЦ – виявлено дистрактор завдання, що не спрацьовує; ПОМИЛК В ТЕКСТІ – виявлено помилку в тексті завдання [64].

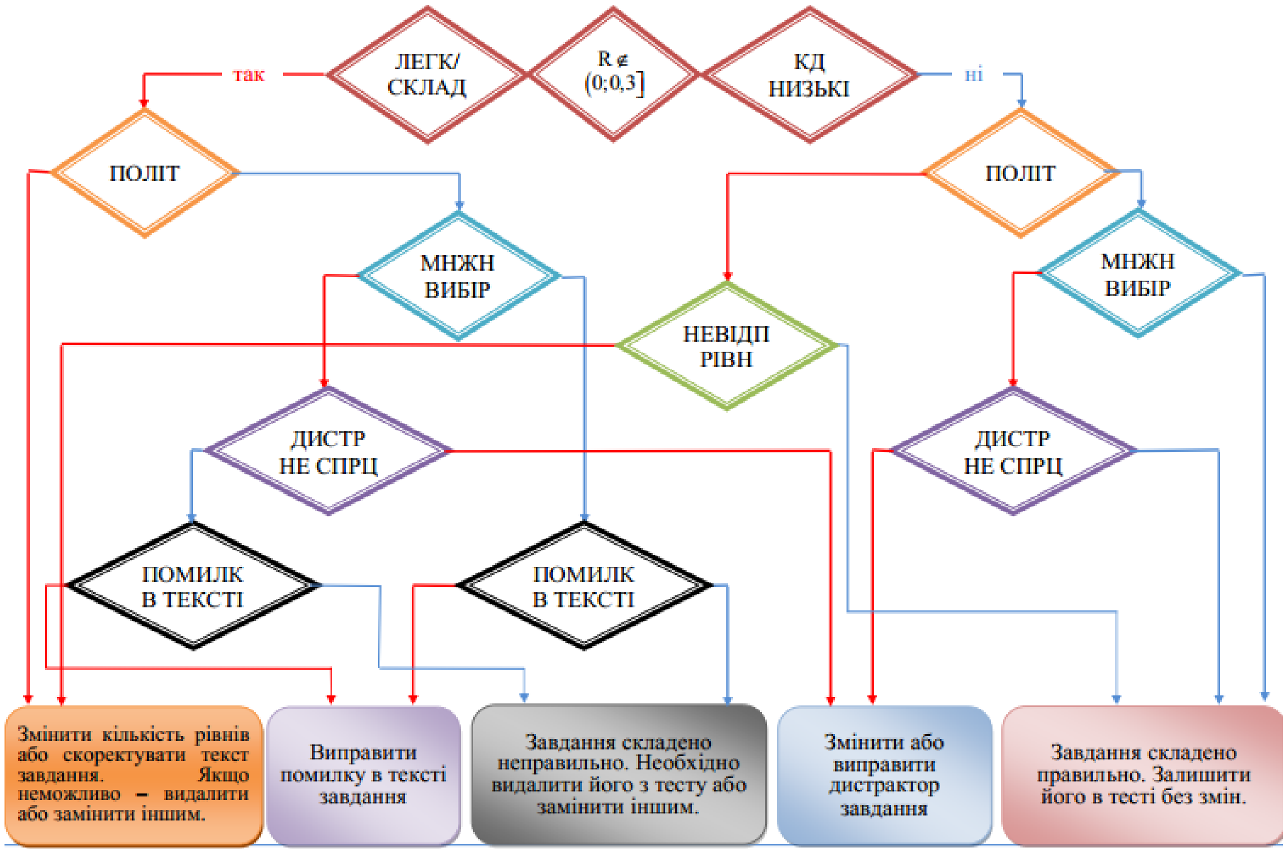

Рисунок 3.2 – Оцінювання якості тестових завдань

Дослідження проводилось на основі обробки результатів навмання взятих 50 робіт тренінгу із зовнішнього незалежного оцінювання (надалі пробного ЗНО) з математики 2019 року. Інструкція до проходження тренінгу із ЗНО 2019 наведено в додатку А, тренінг із ЗНО – в додатку Б, бланк відповідей – в додатку В. Розглядалися тільки тестові частини цих робіт, а саме: завдання з вибором правильної відповіді 1–20 і завдання на встановлення відповідності 21–24 закритої форми з п’ятьма дистракторами.

Алгоритм розрахунку оцінок параметрів складності завдань тесту і підготовки випробовуваних відбувається згідно роботи [36].

Спершу за результатами тестування складається матриця з елементами x_{ij} ($x_{ij} = 1,$ якщо відповідь i-го випробовуваного на j-е завдання вірна, $x_{ij} = 0$ – якщо відповідь невірна). В завданнях 21–24 на встановлення відповідності для зручності користування таблицею використовуємо позначення 21-1 для першого питання 21 тесту, 21-2 для другого питання 21 тесту і т.д.

Далі обчислюються індивідуальні бали випробовуваних X_i (тобто кількість правильних відповідей на тест) і кількість правильних відповідей випробовуваних на кожне завдання тесту Y_j із подальшим упорядкуванням рядків за спаданням X_i, а потім стовпців за зростанням Y_j (див. рис. 3.3).

Номер	Номер завдання j										
учасника	1	2	7	22-1	21-1	22-2	24-1		21-3	21-4	8	
1	0	1	0	1	0	0	0	…	1	0	0	5
2	0	0	0	0	0	0	0	…	0	0	0	8
3	1	1	1	0	0	0	0	…	0	1	0	8
4	1	1	0	0	1	1	0	…	0	0	0	9
5	1	1	0	0	1	0	0	…	0	0	0	10
6	1	1	0	0	0	0	0	…	1	0	0	10
7	1	1	1	0	0	0	1	…	0	0	0	12
…	…	…	…	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…	…	…	…
48	1	1	1	1	1	1	1	…	0	0	1	32
49	1	1	1	1	1	1	1	…	1	1	1	34
50	1	1	1	1	1	1	1	…	1	1	1	35
Rj	45	41	40	40	39	39	39	…	17	15	12	

Рисунок 3.3 – Фрагмент таблиці упорядкованої матриці результатів тестування пробного ЗНО з математики

Проводиться підрахунок долей вірних відповідей i-го випробовуваного $p_i = \frac{X_i}{n}$ і долей невірних відповідей $q_i = 1 - p_i$, де $i = \overline{1, N}$, N – кількість студентів, n – число завдань в тесті.

Після цього проводиться початкова оцінка значень параметра, що характеризує рівень підготовки студентів в логітах (див. рис. 3.4). Логіт рівня підготовки i-го студента знаходять за формулою:

$$\theta_i^0 = ln\frac{p_i}{g_i}, \tag{3.1}$$

де p_i, q_i – долі відповідних вірних і невірних відповідей i-го студента на завдання тесту.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Xi	5	8	8	9	10	10	12	12	13	15	15	15	16	16	17	17	17	18	18	18	19	19	19	20	21
pi	0,14	0,22	0,22	0,25	0,28	0,28	0,33	0,33	0,36	0,42	0,42	0,42	0,44	0,44	0,47	0,47	0,47	0,5	0,5	0,5	0,53	0,53	0,53	0,56	0,58
qi	0,86	0,78	0,78	0,75	0,72	0,72	0,67	0,67	0,64	0,58	0,58	0,58	0,56	0,56	0,53	0,53	0,53	0,5	0,5	0,5	0,47	0,47	0,47	0,44	0,42
θi	-1,82	-1,25	-1,25	-1,1	-0,96	-0,96	-0,69	-0,69	-0,57	-0,34	-0,34	-0,34	-0,22	-0,22	-0,11	-0,11	-0,11	0	0	0	0,11	0,11	0,11	0,22	0,34

	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
Xi	21	21	21	22	22	23	24	24	24	24	25	25	25	25	26	26	28	28	28	28	28	30	32	34	35
pi	0,58	0,58	0,58	0,61	0,61	0,64	0,67	0,67	0,67	0,67	0,69	0,69	0,69	0,69	0,72	0,72	0,78	0,78	0,78	0,78	0,78	0,83	0,89	0,94	0,97
qi	0,42	0,42	0,42	0,39	0,39	0,36	0,33	0,33	0,33	0,33	0,31	0,31	0,31	0,31	0,28	0,28	0,22	0,22	0,22	0,22	0,22	0,17	0,11	0,06	0,03
θi	0,34	0,34	0,34	0,45	0,45	0,57	0,69	0,69	0,69	0,69	0,82	0,82	0,82	0,82	0,96	0,96	1,25	1,25	1,25	1,25	1,25	1,61	2,08	2,83	3,56

Рисунок 3.4 – Початкові значення логітів підготовки 50 випробовуваних

Після чого підраховується долі вірних p_j і невірних q_j відповідей на кожне завдання тесту $p_j = \frac{Y_j}{N}$, $q_j = 1 - p_j$, де N – кількість випробовуваних.

Потім проводиться початкова оцінка значень параметра β, в логітах, що характеризує складність завдань тесту, знаходиться за формулою:

$$\beta_j^0 = ln\frac{q_j}{p_j}, \tag{3.2}$$

де p_j, q_j – долі вірних і невірних відповідей на j-е завдання тесту (рис. 3.5).

	1	2	7	22-1	21-1	22-2	24-1	24-2	5	24-3	6	23-3	17	9	21-2	24-4	12	18
Rj	45	41	40	40	39	39	39	38	35	35	32	32	31	30	30	30	28	28
pj	0,9	0,82	0,8	0,8	0,78	0,78	0,78	0,76	0,7	0,7	0,64	0,64	0,62	0,6	0,6	0,6	0,56	0,56
qj	0,1	0,18	0,2	0,2	0,22	0,22	0,22	0,24	0,3	0,3	0,36	0,36	0,38	0,4	0,4	0,4	0,44	0,44
βj	-2,2	-1,52	-1,39	-1,39	-1,27	-1,27	-1,27	-1,15	-0,85	-0,85	-0,58	-0,58	-0,49	-0,41	-0,41	-0,41	-0,24	-0,24

	23-1	4	22-4	11	3	15	19	20	23-2	10	14	23-4	22-3	13	16	21-3	21-4	8
Rj	26	25	25	24	23	23	23	23	23	22	22	22	21	19	19	17	15	12
pj	0,52	0,5	0,5	0,48	0,46	0,46	0,46	0,46	0,46	0,44	0,44	0,44	0,42	0,38	0,38	0,34	0,3	0,24
qj	0,48	0,5	0,5	0,52	0,54	0,54	0,54	0,54	0,54	0,56	0,56	0,56	0,58	0,62	0,62	0,66	0,7	0,76
βj	-0,08	0	0	0,08	0,16	0,16	0,16	0,16	0,16	0,24	0,24	0,24	0,32	0,49	0,49	0,66	0,85	1,15

Рисунок 3.5 – Початкові значень логітів складності завдань тесту

Наступним кроком підраховується середнє значення логітів рівня підготовки і логітів складності завдань тесту. Середнє значення $\bar{\theta}$ для множини θ_i^0 знаходиться за формулою:

$$\bar{\theta} = \frac{\sum_{i=1}^{N} \theta_i^0}{N}, \tag{3.3}$$

де θ_i^0 – початкові значення логітів підготовки i-го абітурієнта, N – число випробовуваних абітурієнтів.

Середнє значення $\bar{\beta}$ для множини початкових значень логіту складності β_j^0 $(j = \overline{1,n})$, де n – кількість завдань у тесті, знаходиться за формулою:

$$\bar{\beta} = \frac{\sum_{j=1}^{n} \beta_j^0}{n}. \tag{3.4}$$

Для розглянутого тесту пробного ЗНО середнє значення $\bar{\theta} = 0{,}332$ логітів рівня підготовки та середнє значення $\bar{\beta} = -0{,}3$ логітів складності.

Для оцінки параметрів рівня підготовки і складності завдань знаходять дисперсії $V = \frac{\sum_{i=1}^{N}(\theta_i^0)^2 - N(\bar{\theta})^2}{N-1}$ для множини $\theta_i^0 (i = \overline{1,N})$ та $U = \frac{\sum_{i=1}^{n}(\beta_j^0)^2 - n(\bar{\beta})^2}{n-1}$ для множини $\beta_j^0 (j = \overline{1,n})$. А також поправочні коефіцієнти $X = \sqrt{\frac{1+U/2.89}{1-UV/8.35}}$, $Y = \sqrt{\frac{1+V/2.89}{1-UV/8.35}}$. Для розглянутого тесту $V \approx 0{,}998$; $U \approx 0{,}574$; $X \approx 1{,}1$; $Y \approx 1{,}2$.

Оцінки параметрів рівня підготовки θ і складності завдань β в єдиній інтервальній шкалі знаходяться за формулами:

$$\theta_i = \bar{\beta} + X \cdot \theta_i^0, \; \beta_j = \bar{\theta} + Y \cdot \beta_j^0, \tag{3.5}$$

де всі позначення залишаються колишніми, а параметри θ і β мають оцінки θ_i і β_j в стандартній інтервальній шкалі.

Для розглянутого тесту пробного ЗНО з математики 2019 року рівень підготовки i-тих абітурієнтів $\theta_i = -0{,}2 + 1{,}1 \cdot \theta_i^0$ логіт $(i = \overline{1, N})$, а складність j-го завдання $\beta_j = 0{,}3 + 1{,}2 \cdot \beta_j^0$ логіт $(j = \overline{1, n})$.

Потім оцінюється стандартна похибка вимірювання $S(\theta_i)$ за формулою (3.6), яка знаходиться для кожного значення θ_i $(i = \overline{1, N})$.

$$S(\theta_i) = \frac{X}{\sqrt{p_i(n - X_i)}} = \frac{X}{\sqrt{np_i(1 - p_i)}} = \frac{X}{\sqrt{np_i q_i}}. \tag{3.6}$$

Результати підрахунку оцінок параметрів рівня підготовки θ_i і похибок параметру θ_i для розглянутого прикладу наведені на рис. 3.6.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
θi	-1,82	-1,25	-1,25	-1,1	-0,96	-0,96	-0,69	-0,69	-0,57	-0,34	-0,34	-0,34	-0,22	-0,22	-0,11	-0,11	-0,11	0	0	0	0,11	0,11	0,11	0,22	0,34
S(θi)	0,55	0,45	0,45	0,44	0,42	0,42	0,4	0,4	0,39	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38	0,38

	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
θi	0,34	0,34	0,34	0,45	0,45	0,57	0,69	0,69	0,69	0,69	0,82	0,82	0,82	0,82	0,96	0,96	1,25	1,25	1,25	1,25	1,25	1,61	2,08	2,83	3,56
S(θi)	0,38	0,38	0,38	0,39	0,39	0,39	0,4	0,4	0,4	0,4	0,41	0,41	0,41	0,41	0,42	0,42	0,45	0,45	0,45	0,45	0,45	0,51	0,6	0,83	1,15

Рисунок 3.6 – Оцінка параметру підготовки і стандартної похибки вимірювання параметра θ_i

А також підраховується стандартна похибка вимірювання $S(\beta_j)$, яка знаходиться для кожного значення складність j-го завдання за формулою (3.7). Результати підрахунку похибок $S(\beta_j)$ наведені на рис. 3.7.

$$S(\beta_j) = \frac{Y}{\sqrt{p_j(N - Y_j)}} = \frac{Y}{\sqrt{Np_j(1 - p_j)}} = \frac{Y}{\sqrt{Np_j q_j}}. \tag{3.7}$$

	1	2	7	22-1	21-1	22-2	24-1	24-2	5	24-3	6	23-3	17	9	21-2	24-4	12	18
βj	-2,2	-1,52	-1,39	-1,39	-1,27	-1,27	-1,27	-1,15	-0,85	-0,85	-0,58	-0,58	-0,49	-0,41	-0,41	-0,41	-0,24	-0,24
S(βj)	0,57	0,44	0,42	0,42	0,41	0,41	0,41	0,4	0,37	0,37	0,35	0,35	0,35	0,35	0,35	0,35	0,34	0,34

	23-1	4	22-4	11	3	15	19	20	23-2	10	14	23-4	22-3	13	16	21-3	21-4	8
βj	-0,08	0	0	0,08	0,16	0,16	0,16	0,16	0,16	0,24	0,24	0,24	0,32	0,49	0,49	0,66	0,85	1,15
S(βj)	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,35	0,35	0,36	0,37	0,4

Рисунок 3.7 – Оцінка параметру β_j і стандартної похибка оцінок складності завдань

Після підрахунку значень параметрів θ та β на шкалі логітів будують характеристичні криві завдань тесту, аналіз їх взаємного розташування дозволить намітити шлях подальшого удосконалення тесту (у разі потреби) і дозволить зробити аналіз підготовки абітурієнтів.

Згідно до алгоритму на основі статистичних даних будуються логістичні функції відповідно до обраної моделі IRT теорії. Розглянемо двопараметричну модель А. Бірнбаума, яка обчислюється формулами (2.5) і (2.6).

Для використання формули (2.5) потрібно знайти параметр диференційованої спроможності (дискримінативності) j-го завдання тесту за формулою: $a_j = \frac{(r_{bis})_j}{\sqrt{1-(r_{bis})_j^2}}$.

Для спрощення розрахунку коефіцієнта a_j використовуються точково-бісеріальний коефіцієнт який знаходиться за формулою [36]:

$$\left(r_{pbis}\right)_j = \frac{(\bar{X}_1)_j \ - (\bar{X}_0)_j}{S_x} \cdot \sqrt{\frac{(N_1)_j \cdot (N_0)_j}{N(N-1)}}, \tag{3.8}$$

де всі позначення ті ж що і для формули 2.8.

Перевагою використання точково-бісеріального коефіцієнту замість бісеріального коефіцієнту є простота в підрахунках і відсутність обов’язкових гіпотез, які висуваються в силу необхідності нормального характеру розподілу дихотомічних даних при знаходженні міри зв’язку.

Знайдемо коефіцієнти $(\bar{X}_1)_j$, $(\bar{X}_0)_j$ і точково-бісеріальний коефіцієнт для першого завдання:

$(\bar{X}_0)_1 = \frac{5+8+16+19+21}{5} = 13{,}8,\ (\bar{X}_1)_1 = (8 + 9 + 10 * 2 + 12 * 2 + 13 + 15 * 3 + 16 + 17 * 3 + 18 * 3 + 19 * 2 + 20 + 21 * 3 + 22 * 2 + 23 + 24 * 4 + 25 * 4 + 26 * 2 + 28 * 5 + 30 + 32 + 34 + 35) : 45 = 21.$

$$(N_1)_1 = 45, (N_0)_1 = 5, S_x \approx 6{,}994, \ \left(r_{pbis}\right)_j = \frac{21-13{,}8}{6{,}994}\sqrt{\frac{45*5}{50*49}} \approx 0{,}31.$$

Аналогічно для інших завдань тесту, результати наведені на рис. 3.8.

i	1	2	7	22-1	21-1	22-2	24-1	24-2	5	24-3	6	23-3	17	9	21-2	24-4	12	18
rp_bis	0,31	0,12	0,56	0,58	0,51	0,65	0,62	0,4	0,61	0,48	0,4	0,23	0,19	0,37	0,39	0,55	0,56	0,44

i	23-1	4	22-4	11	3	15	19	20	23-2	10	14	23-4	22-3	13	16	21-3	21-4	8
rp_bis	0,4	0,44	0,59	0,33	0,28	0,29	0,37	0,28	0,37	0,33	0,38	0,4	0,42	0,53	0,41	0,28	0,35	0,39

Рисунок 3.8 – Значення точково-бісеріального коефіцієнта

Завдання вважається валідним, коли значення коефіцієнтів $\left(r_{pbis}\right)_j \approx 0{,}5$ [57]. А отже при аналізі даних можна сказати, що потрібно звернути увагу на завдання в яких точково-бісеріальний коефіцієнт суттєво відрізняється від норми (№2, 23-3).

Наступним знаходимо диференційовану спроможність кожного завдань тесту a_j (див. рис. 3.9), наприклад для першого $a_1 = \frac{0{,}31}{\sqrt{1-0{,}31^2}} \approx 0{,}33$.

i	1	2	7	22-1	21-1	22-2	24-1	24-2	5	24-3	6	23-3	17	9	21-2	24-4	12	18
aj	0,33	0,12	0,68	0,71	0,6	0,86	0,8	0,43	0,77	0,55	0,43	0,24	0,19	0,4	0,42	0,66	0,68	0,49

i	23-1	4	22-4	11	3	15	19	20	23-2	10	14	23-4	22-3	13	16	21-3	21-4	8
aj	0,43	0,49	0,73	0,35	0,29	0,3	0,4	0,29	0,4	0,34	0,42	0,44	0,46	0,62	0,45	0,3	0,38	0,42

Рисунок 3.9 – Значення диференційованої спроможності a_j завдань тесту

Після знаходження всіх параметрів в формулі (2.5) будуємо характеристичні криві завдань тесту. Для побудови характеристичної кривої завдання його складність β вважають параметром, а θ – незалежною змінною, значення якої вибирається довільно. Розглянемо параметр $\theta \in [-5{,}5]$, що відповідає шкалі оцінювання від 0 = «-5» до 10 = «5».

Для побудови графіка функції $P_1(\theta)$ – імовірності виконання 1-го завдання випробовуваним із рівнем підготовки $\theta \in (-5{,}5)$ та складністю $\beta_1 = -2{,}3$ логіта $P_1(\theta) = \frac{e^{1{,}7\cdot 0{,}33\cdot(\theta-(-2{,}3))}}{1+e^{1{,}7\cdot 0{,}33\cdot(\theta-(-2{,}3))}} = \frac{e^{0{,}561\cdot(\theta+2{,}3)}}{1+e^{0{,}561\cdot(\theta+2{,}3)}}$.

Аналогічно виконуємо для інших завдань, а потім нанесемо значення функції на координатну площину і отримаємо характеристичні криві 36 завдань тесту пробного ЗНО з математики.

Аналізуючи поведінку цих функцій можемо всі завдання умовно поділити на три типи:

– Завдання високого рівня виконання – відповідні логістичні функції яких відповідають завданням з ймовірністю виконання ($P_j(\theta) \geq 0{,}95$, для $\theta \geq 2$ – рівень знань вище 7 балів);

– Завдання середнього рівня виконання – відповідні логістичні функції яких відповідають завданням з ймовірністю ($0{,}8 \leq P_j\ (\theta) < 0{,}95$, для $\theta \geq 2$);

– Завдання низького рівня виконання ($0{,}5 \leq P_j(\theta) \leq 0{,}8$, для $\theta \geq 0$).

Відповідні цим типам графіки зображено на рис.3.10.

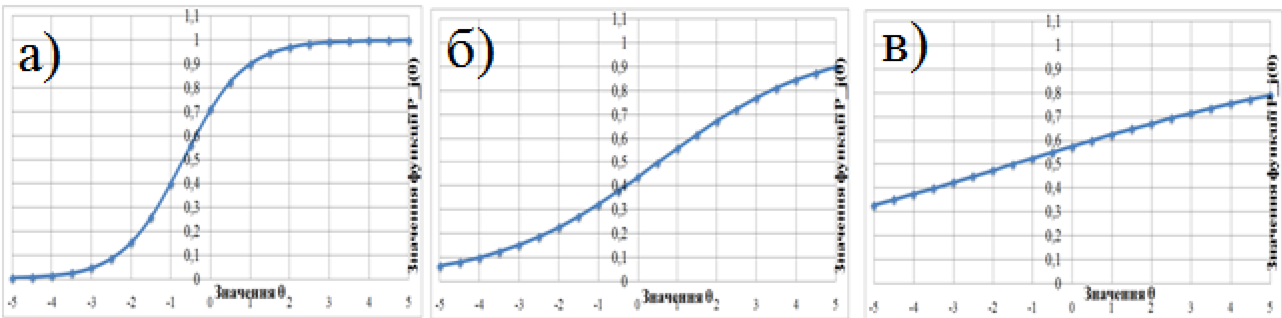

Рисунок 3.10 – Види характеристичних кривих ймовірностей виконання завдань в залежності від рівня підготовки: а) криві високого рівня виконання; б) криві середнього рівня виконання; в) криві низького рівня виконання

Розглянемо детальніше функції 1 типу: завдання високого рівня виконання, тобто з ймовірністю виконання $P_j(\theta) \geq 0{,}95$, для рівня знань випробовуваних вище 7 балів ($\theta \geq 2$) рис. 3.11.

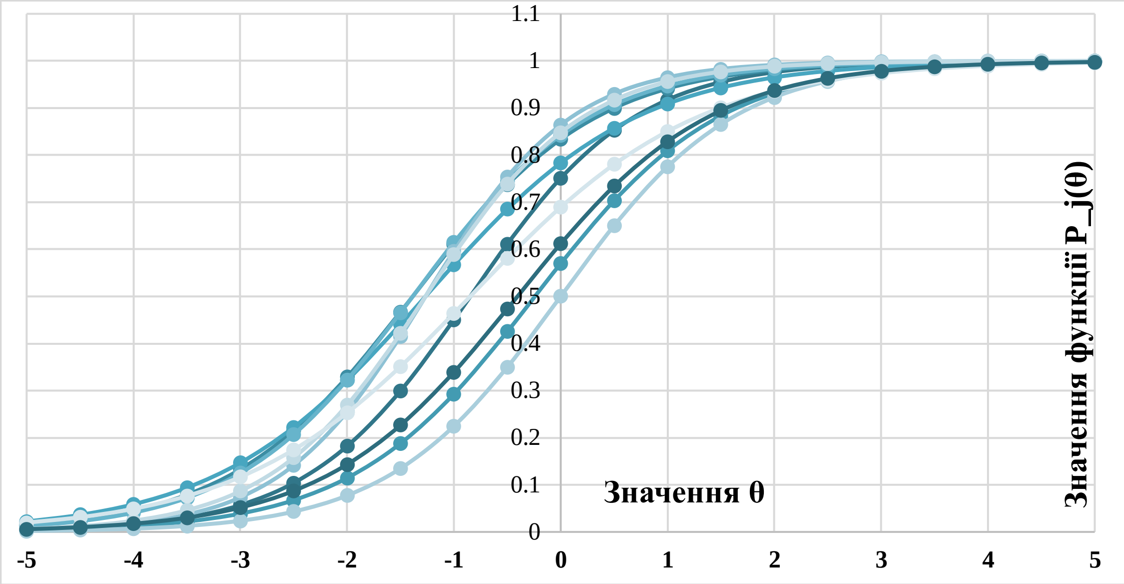

Рисунок 3.11 – Характеристичні криві завдань, для $P_j(\theta) \geq 0{,}95$

Аналізуючи отримані дані і рисунок 3.11 на якому зображені логістичні криві $P_j(\theta)$ ймовірності виконання завдань абітурієнтами з рівнем підготовки $\theta \in (-5{,}5)$. Завдання з номерами 5, 7, 12, 21-1, 22-1, 22-2, 22-4, 24-1, 24-3, 24-4 мають високий рівень ймовірності виконання. Тематика цих завдань: прогресії, нерівності, логарифми, знаходження об'ємів тіл.

Завдання 1, 4, 6, 9, 13, 14, 16,18, 19, 21-2, 22-3, 23-1, 23-2, 23-4, 24-2 мають менше значення $P_j(\theta)$, в них відсутній ефект насищення на абітурієнтах підготовлених вище середнього рівня, тобто ці завдання середнього рівня виконання з відповідними логістичними функції які відповідають значенню ймовірності $0{,}8 \leq P_j\ (\theta) < 0{,}95$, для $\theta \geq 2$.

Розглянемо детальніше завдання 23: установіть відповідність між функціями (1-4) та їх похідними у точці x_0 (А-Д):

1)	$y = 2\sqrt{x},\ x_0 = 4;$	А) 1
2)	$y = sin2x, x_0 = \frac{\pi}{6};$	Б) 48
3)	$y = 4x^3 + 1, x_0 = 2;$	В) 20
4)	$y = \frac{2}{x} + 2x,\ x_0 = 1.$	Г) 0
		Д) 0,5

Побудовані характеристичні криві 23 завдання двопараметричною моделлю А. Бірнбаума (див. рис. 3.12) характеризують його як те, що включає в себе завдання середнього та низького рівня складності. Третє завдання №23 (тобто 23-3) для абітурієнтів виявилось складнішим за інші три (у порівнянні з іншими трьома найгірші результати), хоча є найлегшим. Такий результат можна пояснити помилкою – використанням формул інтегрування замість формул диференціювання.

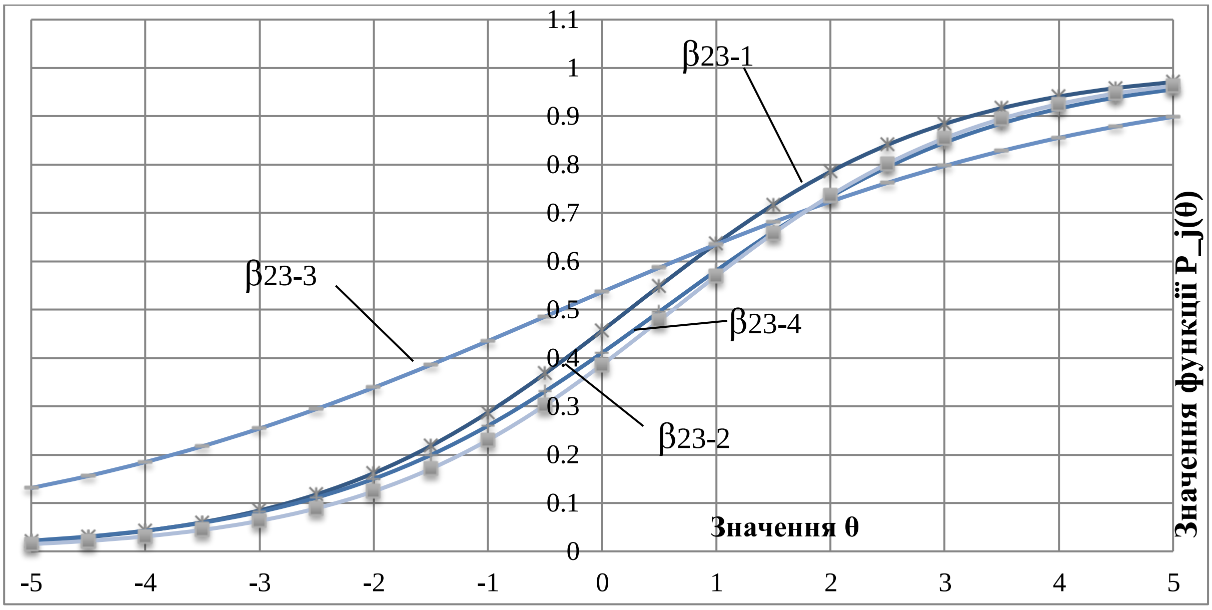

Рисунок 3.12 – Характеристичні криві завдання № 23 тесту

Далі будуються логістичні функції відповідно до трипараметричної модель А. Бірнбаума, яка обчислюється формулою (2.9), де третій параметр характеризує ймовірність вірної відповіді на завдання j в тому випадку, якщо відповідь була вгадана. Для використання цієї формули потрібно знайти параметр c_j, який визначається кількістю відповідей з закритими завданнями тесту, тобто в нашому випадку з п'ятьма дистракторами $c_j = 0{,}2$.

Для побудови графіка функції $P_1(\theta)$ – імовірності виконання 1-го завдання випробовуваним із рівнем підготовки $\theta \in (-5{,}5)$ та складністю $\beta_1 = -2{,}3$, $P_1(\theta) = 0{,}2 + (1 - 0{,}2) \cdot \frac{e^{1{,}7 \cdot 0{,}33 \cdot (\theta - (-2{,}3))}}{1 + e^{1{,}7 \cdot 0{,}33 \cdot (\theta - (-2{,}3))}} = 0{,}2 + 0{,}8 \cdot \frac{e^{0{,}561 \cdot (\theta + 2{,}3)}}{1 + e^{0{,}561 \cdot (\theta + 2{,}3)}}$.

Аналогічно будуються ймовірності для виконання всіх завдань в тесті.

Розглянемо детальніше 5 завдання тесту за моделями А. Бірнбаума з двома $\beta_5(2)$ та трьома $\beta_5(3)$ параметрами (див. рис.3.13).

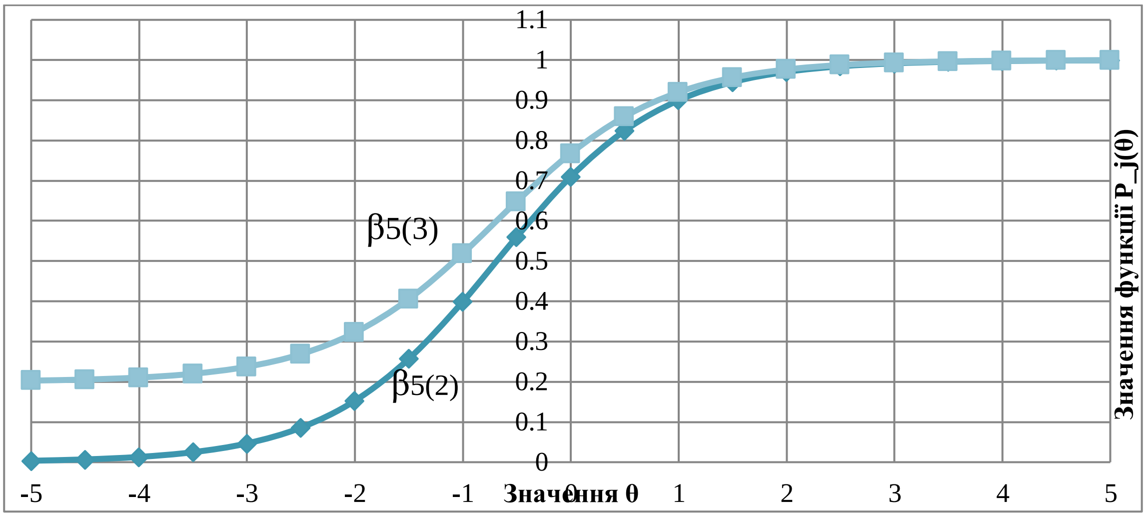

Рисунок 3.13 – Характеристичні криві, за моделями А. Бірнбаума ($\beta_5(2)$ – два параметри, $\beta_5(3)$ – три параметри)

Аналізуючи дані можна сказати, що вгадування дає суттєву похибку в оцінюванні реальних знань абітурієнтів з рівнем підготовки нижче 5, та є несуттєвим для рівня вище 8.

Аналіз завдань 2, 3, 8, 10, 11, 15, 17, 20, 21-3, 21-4, 23-3 за трипараметричною моделлю А. Бірнбаума, криві якої зображені на рис. 3.14 показують, що отримані результати були нижчі ніж очікувані у відповідності до стандартів освіти. Навіть з урахуванням ефекту вгадування, $P_j(\theta)$ – ймовірність вірної відповіді абітурієнтами не досягає 1.

Дані питання в основному були орієнтовані на теоретичні та практичні завдання з геометрії (розташування прямих в ромбі, знаходження периметру паралелограми, знаходження сторони паралелограми за даними координатами вершин та обчислення об’єму призми), а також на стандартні задачі тригонометрії (розв’язати нерівність $tgx \geq \sqrt{3}$, де $-\frac{\pi}{2} < x < \frac{\pi}{2}$; знайти період функції $y = sin4x$) та на знаходження області визначення функції. Такий рівень знань може бути оснований на недостатньо розвиненому просторовому мислені, недосконалому знанні формул та невірного їх використання.

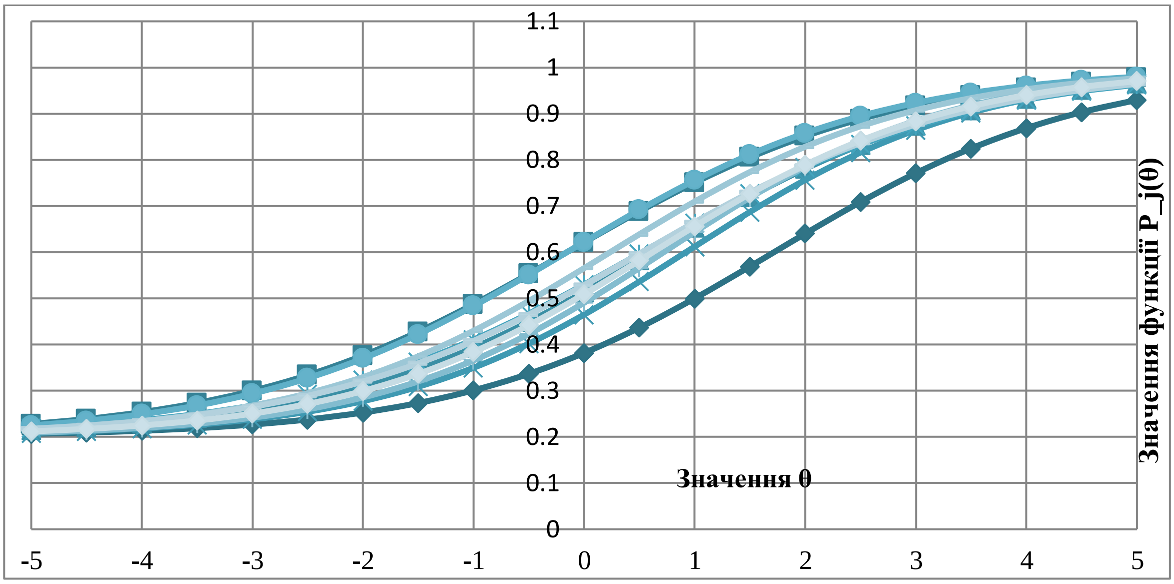

Рисунок 3.14 – Характеристичні криві завдань тесту трипараметричною моделлю А.Бірнбаума

Наступним кроком проведемо аналіз завдання на встановлення відповідності 21–24 закритої форми з п'ятьма дистракторами політомічними моделями розглянутими в розділі 2.

Розглянемо детальніше завдання 21: у паралелограмі $ABCD$ бісектриса AE кута BAD ділить протилежну сторону у відношенні $BE : EC = 5 : 3$, $\angle BAD = 60°$, $AD = 8$ см. Установіть відповідність між вказаними відрізками (1-4) ті їх величинами (А-Д):

1) EC	А) 7 см
2) AB	Б) $\sqrt{129}$ см
3) BD	В) 3 см
4) AC	Г) $\sqrt{137}$ см
	Д) 5 см

При розгляданні політомічного чотирьох крокового завдання ймовірність отримати $0, 1, 2, 3, 4$ бали за виконання завдань відповідно позначаються P_0, P_1, P_2, P_3, P_4 (формули (2.30)–(2.34)).

В нашому прикладі для відповідних складностей завдань $\beta_1 = -1{,}189; \beta_2 = -0{,}155; \beta_3 = 1{,}129; \ \beta_4 = 1{,}35, P_i$ рівні:

$P_0 = \frac{1}{1+e^{(\theta+1,189)}+e^{(2\theta+2,045)}+e^{(3\theta+0,916)}+e^{(4\theta-0,434)}}$. При $\theta = -5, P_0 \approx 0,978$.

$P_1 = \frac{e^{(\theta+1,19)}}{1+e^{(\theta+1,189)}+e^{(2\theta+2,045)}+e^{(3\theta+0,916)}+e^{(4\theta-0,434)}}$. При $\theta = -5, P_1 \approx 0,0216$.

$P_2 = \frac{e^{(2\theta+1,35)}}{1+e^{(\theta+1,189)}+e^{(2\theta+2,045)}+e^{(3\theta+0,916)}+e^{(4\theta-0,434)}}$. При $\theta = -5, P_2 \approx 0,00017$.

$P_3 = \frac{e^{(3\theta+0,22)}}{1+e^{(\theta+1,189)}+e^{(2\theta+2,045)}+e^{(3\theta+0,916)}+e^{(4\theta-0,434)}}$. При $\theta = -5, P_3 \approx 3,71092\text{E}-07$.

$P_4 = \frac{e^{(4\theta-1,13)}}{1+e^{(\theta+1,189)}+e^{(2\theta+2,045)}+e^{(3\theta+0,916)}+e^{(4\theta-0,434)}}$. При $\theta = -5, P_4 \approx 6,48004\text{E}-10$.

Аналізуючи отримані графіки (див. рис. 3.15) ймовірність отримати 0 балів при збільшенні значення θ прямує до 0, тобто при збільшенні рівня підготовки шанси отримати 0 балів зменшуються. Ймовірність отримати 4 бали при $\theta \leq -1$ наближається 0 і дорівнює 0,5 лише при $\theta \approx 1,8$. Тобто ймовірність отримати 4 балів збільшується пропорційно до збільшення рівня підготовки. Найбільша ймовірність отримати 1 бал відповідає значенню $\theta = -1$, найбільша ймовірність отримати 2 бал відповідає значенню $\theta = 0,5$ і найбільша ймовірність отримати 3 бал відповідає значенню $\theta = 1,5$.

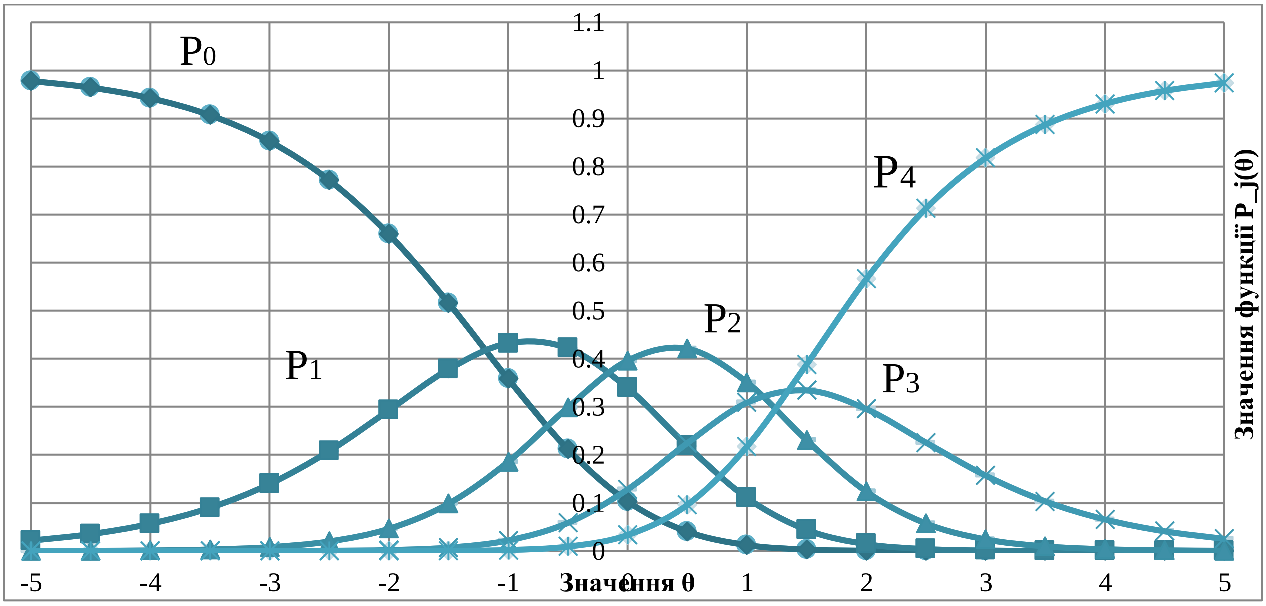

Рисунок 3.15 – Характеристичні криві з політомічного чотирьох крокового завдання

Таким чином наведена методика дозволяє провести оцінку латентних параметрів випробовуваних і параметрів завдань тесту для побудувати ансамблю характеристичних кривих за допомогою відповідних формул ймовірності, інтерпретація яких дає змогу провести аналіз рівнів підготовки абітурієнтів.

Отримані результати дають підстави зробити висновок, що завдання з номерами 5, 7, 12, 21-1, 22-1, 22-2, 22-4, 24-1, 24-3, 24-4 мають високий рівень ймовірності виконання. Тематика цих завдань: прогресії, нерівності, логарифми, знаходження об'ємів тіл.

Завдання 1, 4, 6, 9, 13, 14, 16,18, 19, 21-2, 22-3, 23-1, 23-2, 23-4, 24-2 – завдання середнього рівня виконання з відповідними логістичними функції які відповідають значенню ймовірності $0{,}8 \leq P_j\ (\theta) < 0{,}95$, для $\theta \geq 2$.

Завдання 2, 3, 8, 10, 11, 15, 17, 20, 21-3, 21-4, 23-3 показують, що отримані результати були нижчі ніж очікувані у відповідності до стандартів освіти. Навіть з урахуванням ефекту вгадування, $P_j(\theta)$ – ймовірність вірної відповіді абітурієнтами не досягає 1.

Дані питання в основному були орієнтовані на теоретичні та практичні завдання з геометрії (розташування прямих в ромбі, знаходження периметру паралелограми, знаходження сторони паралелограми за даними координатами вершин та обчислення об'єму призми), а також на стандартні задачі тригонометрії (розв'язати нерівність $tgx \geq \sqrt{3}$, де $-\frac{\pi}{2} < x < \frac{\pi}{2}$; знайти період функції $y = sin4x$) та на знаходження області визначення функції. Такий рівень знань може бути оснований на недостатньо розвиненому просторовому мислені, недосконалому знанні формул та невірного їх використання.

В цілому спостерігається недостатній рівень знань майбутніх абітурієнтів в питаннях орієнтованих на теоретичні та практичні завдання з геометрії та стандартні задачі тригонометрії.

За описаною схемою були проведені розрахунки у створеній програмі «Дослідження тестів» див. додаток Д. Розроблена програма призначена для перевірки якості тестування на основі моделей Item Response Theory для дихотомічних завдань.

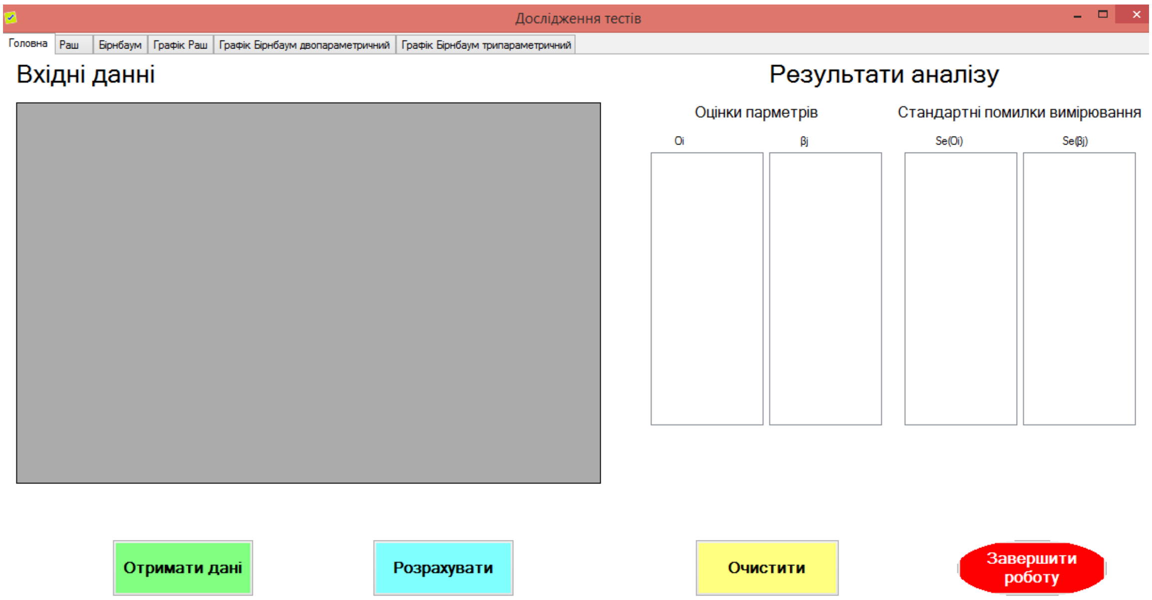

Рисунок 3.16 – Головна вкладка розробленого програмного забезпечення

Інструкція користувача програмного забезпечення: програма може використовуватися при аналізі будь-яких результатів тестування за наявності комп'ютеру з операційною системою Windows 7 та вище.

Дані повинні відповідати строгим критеріям:

– занесені до Excel;

– лист Excel з потрібними даними, має бути 1 листом в документі;

– правильні відповіді позначаються «1», а неправильні «0»;

– аналіз не виконується за умови всіх правильних, або всіх неправильних відповідей випробовуваного чи по завданню тесту.

Відкриті дані будуть показані на головній вкладці програми. При отримані некоректних даних буде показано повідомлення «Помилкові дані !!!», «Виконайте вимоги до БАЗИ ДАНИХ», після чого користуючись кнопкою «Очистити» можна знову почати отримання даних, які відповідають заданим критеріям.

3.2 Висновки за 3 розділом

Розглянута методика оцінки латентних параметрів Item Response Theory за результатами тренінгу із зовнішнього незалежного оцінювання з математики 2019 року дозволяє провести аналіз підготовки абітурієнтів та виявити тематику завдань і відповідний рівень виконання.

Тематика завдань високого рівня виконання: числа і вирази, прогресії, нерівності, логарифми, знаходження об’ємів тіл.

Питання середнього рівня виконання ґрунтувалися в основному на вирази з дробами та модулем, рівняннях та похідних.

Проведений аналіз результатів дослідження за допомогою трипараметричної моделі А. Бірнбаума дозволяє виявити тематику завдань низького рівня виконання: теоретичні та практичні завдання з геометрії (розташування прямих в ромбі, знаходження периметру паралелограма, знаходження сторони паралелограма за даними координатами вершин та обчислення об’єму призми), стандартні задачі тригонометрії на розв’язок нерівності, а також завдання на знаходження періоду та області визначення функцій.

ВИСНОВКИ

Історичні аспекти розвитку тестування у світі і в Україні, історія розвитку основних підходів до обробки результатів тестування – Classic Response Theory і Item Response Theory показують важливість використання тестового методу оцінки знань в освіті.

Розглянуті основні теоретичні засади Item Response Theory та її моделі (однопараметрична модель Г. Раша, двопараметрична та трипараметрична моделі А. Бірнбаума, політомічна модель Раша-Мастерса, політомічна модель Е. Андерсенена та модель Тіссена-Стейнберга для завдань множинного вибору) дають змогу дійти до висновку про доцільність використання цієї теорії не тільки для підвищення точності створення збалансованих і якісних тестів, а й для аналізу підготовки випробовуваних, який був зроблений у дослідженні. Була виявлена тематика завдань і відповідний рівень їх виконання.

Викладений матеріал дозволяє досягти поставленої мети аналізу підготовки абітурієнтів ВНЗ на основі педагогічних тестів.

ПЕРЕЛІК ПОСИЛАНЬ

1. Francis G. Inquiries Into Human Faculty and Its Development. London : Macmillan and co., 1883. 387 p.

2. McCALL W. A. How to Measure in Education. New York : The Macmillan Co., 1922. 416 p.

3. McCALL W. A. How to Experiment in Education. New York : The Macmillan Co., 1923. 281 p.

4. Аванесов В. С. Тесты в социологическом исследовании. Москва : Наука, 1982. 200 с.

5. Фігурська Л. В. Становлення та розвиток тестування як методу педагогічної діагностики [Електронний ресурс] / Народна освіта. № 1. 2009 : https://www.narodnaosvita.kiev.ua/Narodna_osvita/vupysku/7/statti/6figyrska.htm (дата звернення : 19.12.2019).

6. Pearson K. On lines and planes of closest fit to systems of points in space // *Philosophical Magazine*. 1901. Vol 2, № 11. pp. 559–572.

7. Thorndike E. L., Education Psychology. New York : Lemcke and Buechner, 1903. 188 p.

8. Spearman C. The Proof and Measurement of Association between Two Things // *The American Journal of Psychology*. 1904. Vol. 15, № 1. pp. 72–101.

9. Булах І. Є., Мруга М. Р. Створюємо якісний тест : Навч. посіб. Київ : Майстер клас, 2006. 160 с.

10. Pintner R. A non-language group intelligence test // *Journal of Applied Psychology*. 1919. Vol 3. № 3. pp. 199–214.

11. Янченко Т. В. Програмоване навчання як результат еволюції ідей педології та біхевіоризму // *Молодий вчений*. 2016. № 12 (39). С. 550–554.

12. Pressey S. L. A machine for automatic teaching of drill material // *School and Society*. 1927. № 25. pp. 549–552.

13. Pressey S. L. A simple apparatus that gives tests and scores and teaches // *School and Society*. 1926. № 23. pp. 373–376.

14. Freemen F. N. Mental Tests. Their History, Principles and Applications. Boston : Houghton, Mifflin and Company, 1926. 503 p.

15. Greene Ch. E. New Type Tests: [Monograph]. Number Three : Denver, 1926. 35 p.

16. Thurstone L. L. Multiple factor analysis // *Psychological Review*. 1931. Vol 38 № 5. pp. 406–427.

17. Thurstone L. L. Multiple-factor Analysis: A Development and Expansion of The Vectors of the Mind. Chicago : University of Chicago Press, 1947. 535 p.

18. Wechsler D. Manual for the Wechsler Adult Intelligence Scale. New York : The Psychological Corporation, 1955. 230 p.

19. Крутько О. Часопис «Радянська школа» про програмоване навчання в Україні в 60-ті роки XX століття // *Історія педагогічної думки*. 2010. №1. С.15–19.

20. Lord F. M. Application of Item Response Theory to Practical Testing Problems. Hillsdale. New Jersey : Lawrence Erlbaum Ass. 1980. 266 p.

21. Skinner B. F. The science of learning and art of teaching // *Harward Education Review.* 1954. № 24. pp. 86–97.

22. Булах І. Є. Комп’ютерна діагностика навчальної успішності. Київ : ЦМК МОЗ України, УДМУ. 1995. 221 с.

23. Гаврилюк Н. М. Моніторинг якості освіти: зарубіжний досвід // *Наукові записки Вінницького державного педагогічного університету імені Михайла Коцюбинського*. 2012. № 38. С.350–354.

24. Мишко С.А. Становлення системи тестування у сфері освіти США в XX столітті // *Науковий вісник УжНУ*. 2017. №1(36). С.81–85.

25. A Nation at Risk: the imperative For Educational Reform. A Report to the Nation and the Secretary of Education United States Department of Education by The National Commission on Excellence in Education. Washington D. C. : The Commission, 1983. 65 p.

26. Попович Л. В. Правове регулювання освітніх програм Європейського Союзу // *Науковий вісник Міжнародного гуманітарного університету*. 2015. №14. С. 89–92.

27. Регламент (ЄС) № 1288/2013 Європейського Парламенту та Ради від 11 грудня 2013 року про встановлення «Erasmus+»: програма Союзу для освіти, підготовки, молоді та спорту.

28. Gulliksen H. Theory of Mental Tests. New York : Wiley, 1950. 486 p.

29. Guttman L. A special review of Harold Gulliksen, Theory of Mental Tests // *Psychometrika*. 1953. №18. pp. 133–130.

30. Guttman L. A basis for analyzing test-retest reliability // *Psychometrika*. 1945. №10. pp. 255–282.

31. Zimmerman D. W., Williams R. H., Zumbo B. D. Louis Guttman's Contributions to Classical Test Theory // *International Journal of Testing*. 2005. №5. pp. 81–95.

32. Lord F. M., Novick R. M. Statistical theories of mental test scores. Massachuserrs : Addison-Wesley, 1986. 192 p.

33. Crocker L., Algina J. Introduction to classical and modern test theory. New York : Holt, Rinehart, and Winston, 1986. 218 p.

34. Fisher R. A. Statistical Methods for Research Workers. London : Oliver and Boyd, 1925. 276 p.

35. Диховичний О. О., Дудко А. Ф. Аналіз поняттєво-категоріального апарату дослідження якості тестів з вищої математики // *Science Rise: Pedagogical Education.* 2017. №5 (13). С.17–21.

36. Челышкова М. Б. Теория и практика конструирования педагогических тестов : Учебное пособие. Москва : Логос, 2002. 432 с.

37. Ким В. С. Тестирование учебных достижений. Монография. Уссурийск : Издательство УГПИ, 2007. 214 с.

38. Нейман Ю. М., Хлебников В. А. Введение в теорию моделирования и параметризации педагогических тестов. Москва : Прометей, 2000. 168 с.

39. Кухар Л. О., Сергієнко В. П. Конструювання тестів. Курс лекцій : навч. посіб. Луцьк, 2010. 182 с.

40. Янченко Т. В. Програмоване навчання як результат еволюції ідей педології та біхевіоризму // *Молодий вчений*. 2016. №12 (39) . С. 550–554.

41. Глушков В. М., Костюк Г. С. та ін. Наукові проблеми програмованого навчання та шляхи їх розробки // *Радянська школа.* 1966. № 6. С. 17–23. (Передруковано : Психологія і суспільство. 2007. № 4 С.151–167).

42. Психологія програмованого навчання / за ред. Г.С. Костюка. Київ : Радянська школа, 1973. 125 с.

43. Морокішко Є. М., Дутко Л. П. Експериментальна перевірка програмованих посібників // *Радянська школа*. 1966. №8. С. 16–18.

44. Розенберг М. Й. Експериментальне дослідження ефективності програмованого навчання // *Радянська школа*. 1965. №8. С. 25–30.

45. Гаврилюк Н. М. Моніторинг якості освіти: зарубіжний досвід // *Наукові записки Вінницького державного педагогічного університету імені Михайла Коцюбинського*. 2012. №38. С. 350–353.

46. Кашина Г. С., Сергієнко В. П. Зовнішнє незалежне оцінювання в освіті України. Курс лекцій : навч. посіб. Луцьк : видано за підтримки програми Tempus IV, 2010. 115 с.

47. Зовнішнє незалежне оцінювання навчальних досягнень випускників загальноосвітніх навчальних закладів. 2008 р. : Інформаційні матеріали / Український центр оцінювання якості освіти : Уклад. : І. Л. Лікарчук (наук. ред.) та ін. Київ, 2007. 288 с.

48. Рогача О. І. Міжнародні фінанси : підручник. Київ : Либідь, 2003. 784 с.

49. Лупан І. В., Авраменко О. В. Комп’ютерні статистичні пакети : навчально-методичний посібник. Кіровоград, 2010. 218 с.

50. Указ Президента України «Про заходи щодо забезпечення пріоритетного розвитку освіти в Україні» від 30 вересня 2010 р. № 926/2010 // *Інформація і право*. 2011. № 1(1). С. 96–98.

51. Сільвестров Д. С., Борисенко О. Д., Авраменко О. В. та ін. Підготовка фахівців з освітніх вимірювань в Україні : [навчально-методичний комплекс]. Частина 1. Ніжин : Видавець ПП Лисенко М. М., 2012. 362 с.

52. Авраменко О. В., Ковальчук Ю. О., Сергієнко В. П. та ін. Підготовка фахівців з освітніх вимірювань в Україні : [навчально-методичний комплекс] Частина 2. Ніжин : Видавець ПП Лисенко М. М., 2012. 398 с.

53. Авраменко О. В. Вимірювання в освіті : Підручник. Кіровоград : Лисенко В. Ф., 2011. 360 с.

54. Канівець Т. М. Основи педагогічного оцінювання : [навчально-методичний посібник]. Ніжин : Видавець ПП Лисенко М. М., 2012. 102 с.

55. Ковальчук Ю. О. Теорія освітніх вимірювань. Ніжин : Видавець ПП Лисенко М. М., 2012. 200 с.

56. Лутченко Л. І., Пасічник Н. О. Основи педагогічного оцінювання : навчально-методичний посібник. Кіровоград : Лисенко В. Ф, 2012. 72 с.

57. Авраменко О. В., Павличенко Г. Ю., Паращук С. Д. Статистичні методи в освітніх вимірюваннях. Частина І. Класична теорія тестування : навчально-методичний посібник. Кіровоград : Лисенко В. Ф., 2012. 120 с.

58. Лісова Т. В. Моделі та методи сучасної теорії тестів : [навчально-методичний посібник]. Ніжин : Видавець ПП Лисенко М. М., 2012. 112 с.

59. Андронатій П. І., Котяк В. В. Комп'ютерні технології в освітніх вимірюваннях: навчально-методичний посібник. Кіровоград : Лисенко В. Ф., 2011. 144 с.

60. Дьяконов Г. В. Інтерсуб'єктні методи оцінювання психології особистості. Навчальний посібник для вищої школи. Кіровоград : Лисенко В. Ф., 2012. 36 с.

61. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago : Univ. of Chicago Press, 1980. 199 p.

62. Аванесов В. С. Метрическая система Георга Раша – Rasch Measurement (RM) // *Педагогические измерения*. 2010. № 2. С.3–30.

63. Masters G. N. A Rasch model for partial credit scoring // *Psychometrika*. 1982. Vol. 47 (2). pp. 149–174.

64. Дудко А. Ф. Комп’ютерно орієнтована методика оцінювання якості тестів з вищої математики викладачами закладів вищої освіти : дис. на здобуття наук. ступеня канд. педаг. наук : 13.00.10 / Дудко Анна Федорівна. Київ, 2019. 293 с.

65. Алєксєєва І. В., Гайдей В. О., Диховичний О. О., Коновалова Н. Р., Федорова Л. Б. Статистичний аналіз тестових завдань із застосуванням сучасних математичних моделей // *Вісник Національного університету «Львівська політехніка».* 2012. №731. С. 10–13.

66. Raphson J. Analysis Æequationum Universalis seu ad Æequationes Algebraicas Resolvendas Methodus Generalis, & Expedita, Ex nova Infinitarum Serierum Methodo, Deducta ac Demonstrata (in Latin) London : Typis TB. for A. & I. Churchill et al. 1697. 95 p.

67. Демидович Б. П., Марон И. А., Шувалова Э. З. Численные методы анализа. Москва : Наука, 1967. 368 с.

68. Andersen E. B. A general latent structure model for contingency table data / Wainer H., Messick S. Principles of modern psychological measurement. New Jersey : Lawrence Erlbaum Associates, 1983. 377 p.

69. Диховичний О. О., Шепель М. О., Удовенко А. Ф. Політомічні моделі Мастерса та Андерсена в аналізі якості тестових завдань // *Теорія та методика електронного навчання: збірник наукових праць*. 2012. №3. С.83–87.

70. Thissen D., Steinberg L. Response model for multiple choice items // *Psychometrika*. 1984. Vol 49. pp. 501–519.

71. Диховичний О. О., Дудко А. Ф. Автоматизована система аналізу результатів комп’ютерного тестування з вищої математики // *Наукові праці ДонНТУ*. 2013. №2 (14). С. 103–110.

72. Карданова Е. Ю., Карданов Р. С. О некоторых свойствах характеристической и информационной функций политомического тестового

задания // *Вестник Новгородского государственного университета им. Ярослава Мудрого*. 2010. № 55. С. 19–24.

73. Wright B. D., Masters G. N. Rating Scale Analysis. Rasch Measurement. Chicago: Mesa Press. 1979. 206 p.

74. Birnbaum A. Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In Lord F. M. and Novick M. Statistical Theories of Test scores. Reading Mass.: Addison–Wesley. 1968. pp. 397–479.

75. Samejima F. Estimation of Latent Ability Using a Response Pattern of Graded Scores // *Psychometrika Monograph*. 1969. № 17. pp. 1–100.

76. Muraki E. Information Functions of the Generalized Partial Credit Model // *Applied Psychological Measurement*. 1993. Vol 17(4). pp. 351–363.

ДОДАТОК А

Інструкція до проходження тренінгу із ЗНО 2019

ЗОШИТ 1

ТРЕНІНГ ІЗ ЗОВНІШНЬОГО НЕЗАЛЕЖНОГО ОЦІНЮВАННЯ – 2019

СЕРТИФІКАЦІЙНА РОБОТА З МАТЕМАТИКИ

Час виконання – 180 хвилин

Сертифікаційна робота складається з 33 завдань різних форм. Відповіді на завдання 1-24 та 25-30 Ви маєте позначити в бланку **А**, повне розв'язання та відповіді до задач 31-33 ви повинні навести у бланку **Б**.

Інструкція щодо роботи у тестовому зошиті

1. Правила виконання завдань зазначені перед кожною новою формою завдань.
2. Відповідайте тільки після того, як Ви уважно прочитали та зрозуміли завдання.
3. У разі необхідності використовуйте чернетку.
4. Намагайтесь виконати всі завдання.

Інструкція щодо заповнення бланків відповідей А та Б

1. До бланка **А** записуйте лише правильні, на Вашу думку відповіді.
2. Відповіді вписуйте чітко, згідно з інструкціями до кожної форми завдань.
3. Неправильно позначені, закреслені та підчищені відповіді у бланку **А** вважатимуться помилкою.
4. Якщо ви позначили відповідь до котрогось із завдань 1-24 неправильно, то можете виправити її, замалювавши попередню позначку та поставивши нову, як показано на зразку:

	А	**Б**	**В**	**Г**	**Д**
1	☒		■		

5. Якщо Ви записали відповідь до котрогось з завдань 25-30 неправильно, то зможете виправити її, записавши новий варіант відповіді у спеціально відведеному місці бланка **А**.
6. У бланку **Б** потрібно навести повний розв'язок задач 31 – 33 з необхідними рисунками , а також записати у спеціально відведених місцях відповіді до цих задач.

Позначте номер Вашого зошита у відповідному місці бланків **А** та **Б** так:

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
☒														

БАЖАЄМО ВАМ УСПІХУ!

ДОДАТОК Б

Тренінг із зовнішнього незалежного оцінювання 2019

Завдання 1 – 20 мають по п’ять варіантів відповідей, серед яких лише одна відповідь є правильною. Виберіть правильну, на вашу думку, відповідь і позначте її у бланку А. Не робіть інших позначок – комп’ютерна програма реєструватиме їх як помилки.

1. Обчисліть значення виразу: $\dfrac{\left(3\frac{1}{2}-0,25\right)}{2^{-3}}\cdot\dfrac{16}{13}$.

А	**Б**	**В**	**Г**	**Д**
4	0,25	32	16	1

2. Знайдіть область визначення функції $f(x)=\sqrt{5x-x^2-6}$.

А	**Б**	**В**	**Г**	**Д**
$(0;+\infty)$	$[0;+\infty)$	$[2;3]$	$[1;3]$	$[-2;1]$

3. Основою прямої призми з бічним ребром довжиною 8 см є паралелограм з діагоналями 5 см та 7 см, кут між якими дорівнює 60°. Знайдіть об’єм призми.

А	**Б**	**В**	**Г**	**Д**
70см^3	$70\sqrt{3}\,\text{см}^3$	$70\sqrt{2}\,\text{см}^3$	140см^3	$140\sqrt{3}\,\text{см}^3$

4. Спростити вираз $|1-a|+|a-3|-1$ при $1<a<3$.

А	**Б**	**В**	**Г**	**Д**
$2a$	-3	a	$a-1$	1

5. Ділене, частка та остача відповідно дорівнюють 1869, 24 і 69. Знайти дільник.

А	**Б**	**В**	**Г**	**Д**
75	65	58	85	71

6. Для яких значень x функція $y=-\frac{5}{x-1}$ набуває від'ємних значень?

А	Б	В	Г	Д
$x\in(1;5)$	$x\in(-\infty;1]$	$x\in[1;+\infty)$	$x\in(-5;1)$	$x\in(1;+\infty)$

7. Знайдіть перший член арифметичної прогресії $\{a_n\}$, якщо її третій член $a_3=15$, а різниця прогресії $d=7$.

А	Б	В	Г	Д
1	2	3	0	–1

8. Значення виразу $\dfrac{\sin\left(-\frac{\pi}{6}\right)-1}{\cos\left(-\frac{\pi}{6}\right)}+3\mathrm{tg}\left(-\frac{\pi}{6}\right)$ дорівнює

А	Б	В	Г	Д
$-3\sqrt{3}$	0	$\sqrt{3}$	$2\sqrt{3}$	$-2\sqrt{3}$

9. Група складається з 6 чоловіків та 4 жінок. Скількома способами з неї можна вибрати делегацію з 4 осіб, до складу якої входять 2 жінки?

А	Б	В	Г	Д
25	70	90	95	8

10. Сторони паралелограма відносяться як 2 : 1, а діагоналі дорівнюють 2 см та 6 см. Знайдіть периметр паралелограма.

А	Б	В	Г	Д
24 см	15 см	8 см	12 см	18 см

11. Діагоналі ромба $ABCD$ перетинаються у точці O, MA – перпендикуляр до площини ABC. З'ясуйте взаємне розташування прямих BD та MO.

А	Б	В	Г	Д
Перетинаються під гострим кутом	Мимобіжні	Перпендикулярні	Паралельні	Збігаються

12. Знайдіть $x+y$, якщо $x+y<9$ і $\begin{cases}\log_2 x+7^{\log_7 y}=4,\\ x^y=8.\end{cases}$

А	Б	В	Г	Д
–1	0	1	3	5

13. Знайдіть суму найменшого та найбільшого значень функції $y = 1 + 3\sin(x+1)$

А	Б	В	Г	Д
–6	–2	0	2	4

14. Кутовий коефіцієнт дотичної до графіка функції $y = x - x^3$, проведеної у точці з абсцисою $x = 1$, дорівнює

А	Б	В	Г	Д
3	2	–8	–2	5

15. Знайдіть довжину меншої діагоналі паралелограма $ABCD$, якщо $A(4;5)$, $B(3;-3)$, $C(-1;-7)$.

А	Б	В	Г	Д
13	5	7	4	8

16. Знайти повну поверхню правильної чотирикутної піраміди, сторона основи якої дорівнює 6 см, а бічне ребро – 5 см.

А	Б	В	Г	Д
84 см^2	60 см^2	12 см^2	96 см^2	48 см^2

17. Період функції $y = \sin 4x$ дорівнює

А	Б	В	Г	Д
$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{\pi}{4}$	$\frac{\pi}{6}$	4

18. З одного пункту вийшли два туристи. Перший пішов на схід, другий – на південь. Через який час відстань між ними по прямій дорівнюватиме 15 км, якщо швидкість першого – 4 км/год, другого – 3 км/год?

А	Б	В	Г	Д
1 год.	3 год.	2год.	4год.	5год.

19. Знайти $\frac{5x+6y}{2z}$, якщо $2x + y = 3z$, $3x + 5y = 7z$.

А	Б	В	Г	Д
5	10	12	4	1

20. Розв’язком нерівності $\operatorname{tg} x \geq -\sqrt{3}$, що належить проміжку $\left(-\frac{\pi}{2}; \frac{\pi}{2}\right)$, є проміжок:

А	Б	В	Г	Д
$\left(-\frac{\pi}{2}; \frac{\pi}{3}\right)$	$\left[0; \frac{\pi}{6}\right]$	$\left[-\frac{\pi}{3}; \frac{\pi}{2}\right)$	$\left[-\frac{\pi}{2}; \frac{\pi}{4}\right)$	$\left(-\frac{\pi}{6}; \frac{\pi}{4}\right)$

Завдання 21 – 24 передбачають установлення відповідності. До кожного рядка, позначеного ЦИФРОЮ, доберіть один відповідник, позначений ЛІТЕРОЮ, і поставте позначки в *бланку А* на перетині відповідних рядків (цифри) і колонок (літери). Усі інші види Вашого запису в *бланку А* комп’ютерна програма реєструватиме як помилки!

21. У паралелограмі $ABCD$ бісектриса AE кута BAD ділить протилежну сторону у відношенні $BE:EC = 5:3$, $\angle BAD = 60^\circ$, $AD = 8$ см. Установіть відповідність між вказаними відрізками (1 – 4) та їх величинами (А-Д).

1 EC	**А 7** см
2 AB	**Б** $\sqrt{129}$ см
3 BD	**В 3** см
4 AC	**Г** $\sqrt{137}$ см
	Д 5 см

	А	Б	В	Г	Д
1					
2					
3					
4					

22. Установіть відповідність між нерівностями (1-4) та множинами їх розв’язків (А-Д).

1 $\frac{x+1}{x-2} \geq 0$	**А** $(-\infty; 3) \cup (3; +\infty)$
2 $x^2 - 6x + 9 > 0$	**Б** $(-3; +\infty)$
3 $\log_2(1-x) \geq -\log_2 \frac{1}{x}$	**В** $(-\infty; -1] \cup (2; +\infty)$
4 $2^{x+1} < 4^{x+2}$	**Г** $[0,5; 2]$
	Д $(0; 0,5]$

	А	Б	В	Г	Д
1					
2					
3					
4					

23. Установіть відповідність між функціями (1-4) та їх похідними у точці x_0 (А-Д).

1 $y = 2\sqrt{x}$, $x_0 = 4$	**А** 1
2 $y = \sin 2x$, $x_0 = \pi/6$	**Б** 48
3 $y = 4x^3 + 1$, $x_0 = 2$	**В** 20
4 $y = \frac{2}{x} + 2x$, $x_0 = 1$	**Г** 0
	Д 0,5

	А	**Б**	**В**	**Г**	**Д**
1					
2					
3					
4					

24. Установіть відповідність між заданими многогранниками (1-4) та їх об’ємами (А-Д).

1 Куб із ребром 4	**А** 9
2 Прямокутний паралелепіпед з вимірами 3, 4, 5	**Б** 10
3 Піраміда з висотою 10, основою якої є прямокутний трикутник з катетами 3 і 4	**В** 20
4 Правильний тетраедр з ребром $3\sqrt{2}$	**Г** 60
	Д 64

	А	**Б**	**В**	**Г**	**Д**
1					
2					
3					
4					

Розв’яжіть завдання 25 – 30. Одержані відповіді запишіть у зошиті та бланку *А*.
Пам’ятайте, що відповіді у бланку *А* необхідно записувати лише десятковими дробами

25. Середня лінія трикутника ABC на 6 см менша за основу, якій вона паралельна. Знайти у сантиметрах: 1) основу трикутника ABC; 2) середню лінію трапеції, основами якої є середня лінія трикутника ABC і його основа.

1) Відповідь: ____

2) Відповідь: ____

26. Для нерівності $\dfrac{x^2+6x}{\sqrt{x^2-6x+9}}<0$ знайти: 1) середину інтервалу, на якому вона виконується; 2) суму цілих розв'язків цієї нерівності.

1) Відповідь: ____

2) Відповідь: ____

27. Знайти n, якщо $C_{n+4}^{n+1}-C_{n+3}^{n}=15(n+2)$.

Відповідь: ____

28. Знайдіть довжину проміжку, на якому виконується нерівність $2\log_2(x-1)>\log_2(5-x)+1$.

Відповідь: ____

29. Знайдіть у сантиметрах гіпотенузу прямокутного трикутника, радіус вписаного кола якого дорівнює 3 см, а один з катетів – 8 см.

Відповідь: ____

30. У двох урнах лежать по 10 кульок, причому у першій з них 7 червоних, а у другій – 4 червоних. З кожної урни навмання виймають по одній кульці. Знайдіть ймовірність того, що обидві вийняті кульки будуть червоними.

Відповідь: ____

Розв'яжіть завдання 31 – 33, запишіть повний варіант вашого розв'язання до бланку Б, за необхідності наведіть рисунок. Відповіді до завдань наведіть у відведених для цього місцях бланку Б. Розв'язання до завдань 31 – 33 повинно містити повне пояснення, записане у вигляді послідовних логічних дій із посиланням на математичні факти, з яких випливає те або інше твердження. Якщо це необхідно, виконайте креслення або побудуйте графік.

31. Обчисліть площу криволінійної трапеції, обмеженої графіком функції $y = x\sqrt{x}$, віссю Ox і прямими $x = 1$ та $x = 4$.

Відповідь: ____

32. Знайти бічну поверхню правильної чотирикутної піраміди, якщо її об’єм дорівнює $\frac{32}{\sqrt{3}}$, а двогранний кут при основі дорівнює 60^0.

Відповідь: ____

33. Нехай x_1 та x_2 – відповідно точка максимуму та точка мінімуму функції $f(x) = 2x^3 - 9ax^2 + 12a^2x + 1$. При якому значенні параметра a $x_1^2 = x_2$?

Відповідь: ____

ЧЕРНЕТКА

Таблиця значень тригонометричних функцій деяких кутів

α	0°	30°	45°	60°	90°
$\sin\alpha$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\cos\alpha$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\mathrm{tg}\alpha$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	не існує
$\mathrm{ctg}\alpha$	не існує	$\sqrt{3}$	1	$\frac{\sqrt{3}}{3}$	0

Кінець тестового зошита

ДОДАТОК В

Бланк відповідей

1. В
2. В
3. Б
4. Д
5. А
6. Д
7. А
8. Д
9. В
10. Г
11. В
12. Д
13. Г
14. Г
15. Б
16. А
17. Б
18. Б
19. А
20. В

21. 1 – В, 2 – Д, 3 – А, 4 – Б
22. 1 – В, 2 – А, 3 – Д, 4 – Б
23. 1 – Д, 2 – А, 3 – Б, 4 – Г
24. 1 – Д, 2 – Г, 3 – В, 4 – А
25. 1) 12. 2) 9.
26. 1)–3. 2)–15.
27. 27.
28. 2.
29. 17.
30. 0.28
31. 12.4
32. 32.
33. 2.

ДОДАТОК Г

Таблиця результатів пробного тренінгу ЗНО

Номер учасника i	Номер завдання j																
	1	2	7	22-1	21-1	22-2	24-1	24-2	5	24-3	6	23-3	17	9	21-2	24-4	12	18
1	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	0	0	0
3	1	1	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
4	1	1	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	1
5	1	1	0	0	1	0	0	0	0	1	0	1	0	0	0	0	1	0
6	1	1	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0
7	1	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0
8	1	1	1	0	0	0	0	0	1	0	0	0	1	0	0	1	0	1
9	1	0	0	1	1	0	1	0	0	0	1	1	1	1	0	1	1	1
10	1	1	1	1	1	0	1	1	0	1	1	0	1	0	0	0	1	1
11	1	1	1	1	0	1	1	1	0	0	1	1	0	1	0	0	0	0
12	1	0	1	0	1	1	1	1	1	1	1	0	0	0	1	0	0	0
13	0	1	1	1	1	1	1	1	0	0	1	1	0	1	1	0	1	1
14	1	1	0	1	0	0	1	1	0	1	0	0	0	0	0	1	1	0
15	1	0	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
16	1	1	0	1	0	1	0	0	1	0	1	0	1	1	1	1	0	0
17	1	1	1	1	1	1	1	1	0	1	0	1	1	0	1	1	0	0
18	1	1	1	1	1	1	0	0	1	1	1	0	1	1	0	0	0	1
19	1	0	1	0	1	1	1	0	0	0	0	1	1	0	1	1	0	1
20	1	1	0	1	1	1	0	1	1	1	0	0	1	1	0	0	0	0
21	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1
22	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	0	0	0
23	0	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	0
24	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	0	1
25	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	0	1
26	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1
27	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	0
28	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0
29	1	0	1	1	1	1	1	1	0	1	0	1	0	0	0	1	1	0
30	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1
31	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1
32	1	1	1	1	1	1	1	0	1	0	1	1	1	1	0	0	1	1
33	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0
34	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1
35	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1
36	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
37	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0
38	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1
39	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0
40	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1
41	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
42	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
43	1	1	1	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1
44	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1
45	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
46	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0
47	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
48	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
49	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
50	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Рисунок Г.1 – Фрагмент таблиці результатів пробного тренінгу ЗНО 2019 з математики

Номер учасника i	Номер завдання j																	
	23-1	4	22-4	11	3	15	19	20	23-2	10	14	23-4	22-3	13	16	21-3	21-4	8
1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0
2	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0
3	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0
4	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
5	1	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0
6	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0
7	1	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0	0
8	0	0	0	1	1	0	1	0	0	1	0	0	0	0	0	1	0	0
9	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
10	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0
11	1	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0
12	0	0	0	1	0	0	1	1	0	0	1	0	1	0	0	0	0	0
13	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0
14	1	0	0	0	0	1	0	1	1	1	0	1	0	0	1	0	0	1
15	0	0	1	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0
16	0	1	1	1	0	1	0	0	0	1	0	0	1	0	0	0	0	1
17	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0
18	0	1	0	0	0	1	1	0	1	0	1	0	1	0	0	0	0	0
19	0	1	1	0	1	1	0	1	1	0	1	0	0	0	1	0	0	0
20	1	1	1	1	0	1	0	0	0	1	0	0	1	0	0	0	0	1
21	0	1	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1
22	0	1	1	1	0	1	0	0	1	1	0	0	1	0	0	0	0	1
23	1	1	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0
24	0	0	0	1	0	1	1	0	0	0	1	1	0	0	0	1	0	0
25	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	0	0
26	0	0	1	0	0	0	1	0	1	0	0	1	0	1	1	0	0	0
27	1	1	0	1	0	1	0	1	0	1	1	1	0	0	0	0	0	0
28	0	1	1	0	0	0	1	1	0	1	0	0	0	1	0	1	1	0
29	1	1	1	0	1	1	0	1	1	1	0	1	1	0	0	0	1	0
30	0	0	1	1	1	0	0	0	0	0	0	0	1	0	1	1	1	0
31	1	1	1	0	1	0	0	0	1	1	0	1	0	0	0	0	1	0
32	1	0	1	1	1	1	1	0	1	0	0	1	0	1	0	1	0	0
33	1	1	0	0	1	1	0	1	1	0	0	1	1	0	0	0	0	0
34	0	0	1	1	0	1	0	0	0	0	0	0	1	1	1	1	1	0
35	0	0	1	1	0	0	1	1	0	0	1	0	1	0	1	1	0	0
36	1	1	1	0	0	0	1	1	1	0	0	1	0	0	1	0	0	1
37	1	1	0	1	0	1	0	1	1	1	1	1	0	0	0	0	0	0
38	1	1	0	1	1	0	1	0	1	1	0	1	0	1	1	0	0	0
39	0	0	1	1	1	1	1	1	0	0	1	1	1	0	0	0	0	0
40	1	1	1	1	1	0	1	0	0	0	1	1	1	1	1	0	0	0
41	0	0	0	1	1	1	0	0	1	0	0	0	1	1	1	1	1	0
42	1	1	0	1	1	0	0	0	1	1	1	1	0	1	1	0	0	0
43	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	0
44	1	1	1	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1
45	1	1	1	0	1	0	1	1	0	1	1	1	0	1	0	1	1	0
46	1	0	1	1	0	1	0	0	0	1	1	0	1	1	1	1	1	1
47	0	1	1	0	1	0	1	1	0	0	1	1	1	1	1	1	1	1
48	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1
49	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1
50	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Рисунок Г.2 – Фрагмент таблиці результатів пробного тренінгу ЗНО 2019 з математики

ДОДАТОК Д

Результати обробки даних тренінгу ЗНО за створеним ПЗ

Для підрахунку усіх коефіцієнтів за розглянутою схемою в 3 розділі необхідно натиснути кнопку «Розрахувати». Початкова таблиця даних і отримані розрахунки оцінок параметрів рівня підготовки θ і складності завдань β в єдиній інтервальній шкалі, а також стандартна похибка вимірювання $S(\theta_i), S\left(\beta_j\right)$ показані у вікні на головній вкладці (рис. Д.1).

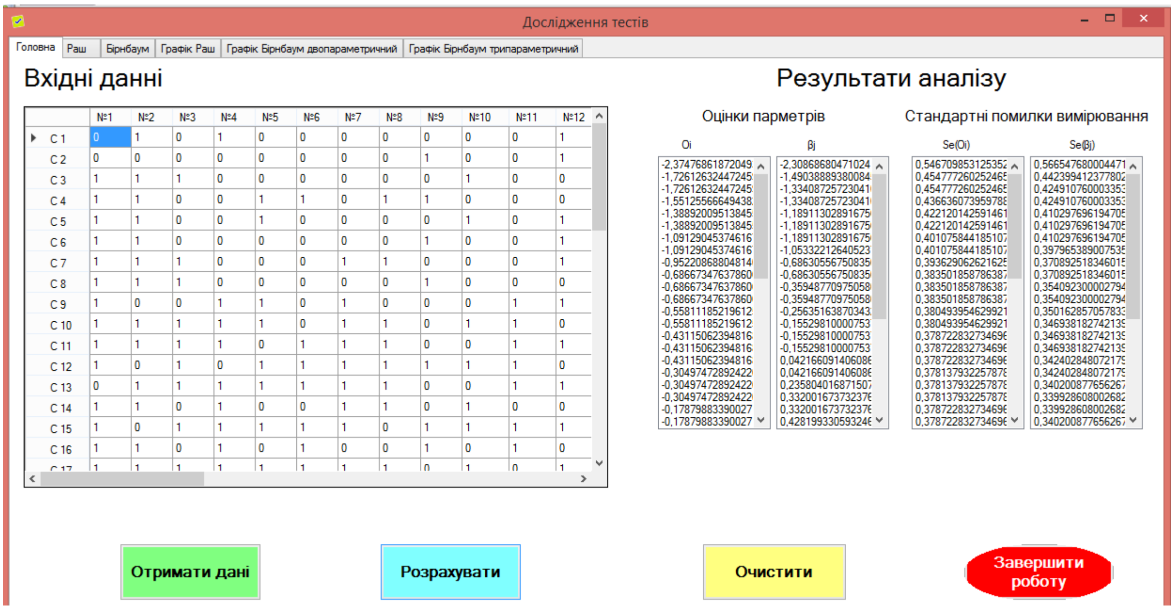

Рисунок Д.1 – Головна вкладка розробленого програмного забезпечення з відповідними розрахунками

Всі допоміжні розрахунки долей вірних і долей невірних відповідей i-го студента, долей вірних і долей невірних відповідей на кожне завдання тесту; початкова оцінка значень параметра, що характеризує рівень підготовки студентів і складності завдань тесту; середнє значення логітів рівня підготовки і логітів складності завдань тесту, а також дисперсії для множин θ_i^0 і β_j^0 можна переглянути на вкладці «Раш». Результати наведено на рисунку Д.2.

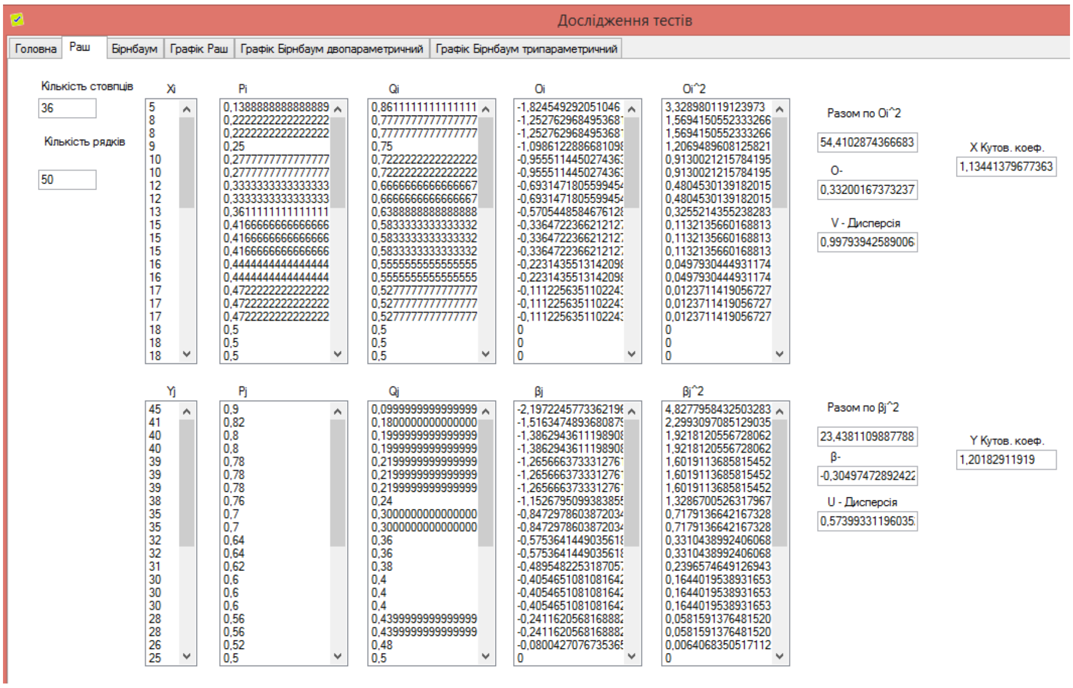

Рисунок Д.2 – Сторінка для додаткових розрахунків по однопараметричній моделі Г. Раша

Ансамбль характеристичних кривих завдань пробного тренінгу з ЗНО 2019 року з математики, знайдений за допомогою формули однопараметричної моделі Г. Раша, наведено на рисунку Д.3.

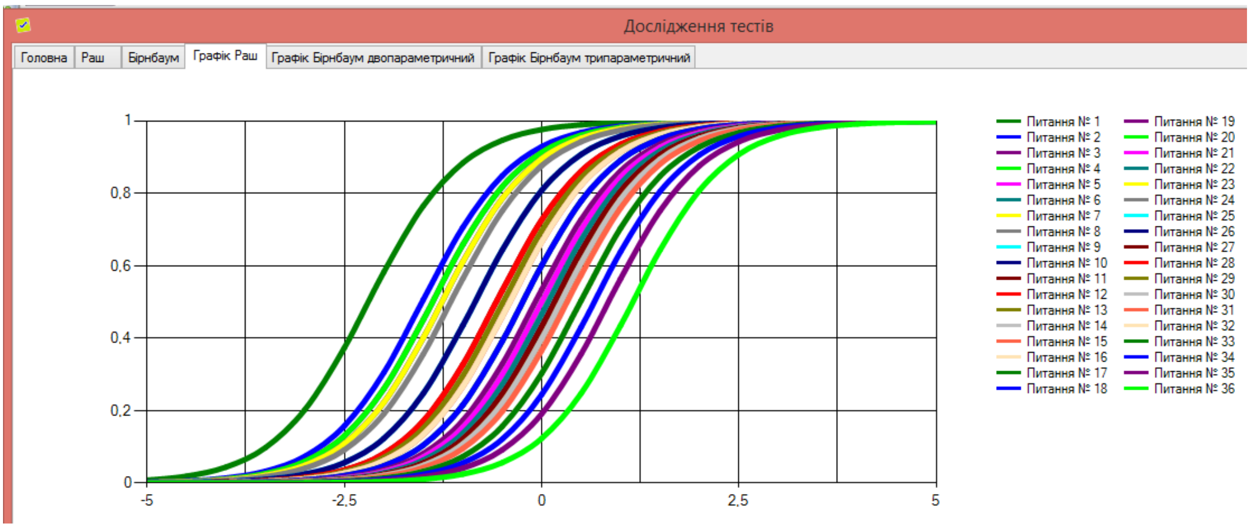

Рисунок Д.3 – Характеристичні криві 36 завдань тесту за однопараметричною моделлю Г. Раша

Отримані розрахунки усіх коефіцієнтів за двопараметричною моделлю А. Бірнбаума показані у вкладці «Бірнбаум» (рис. Д.4).

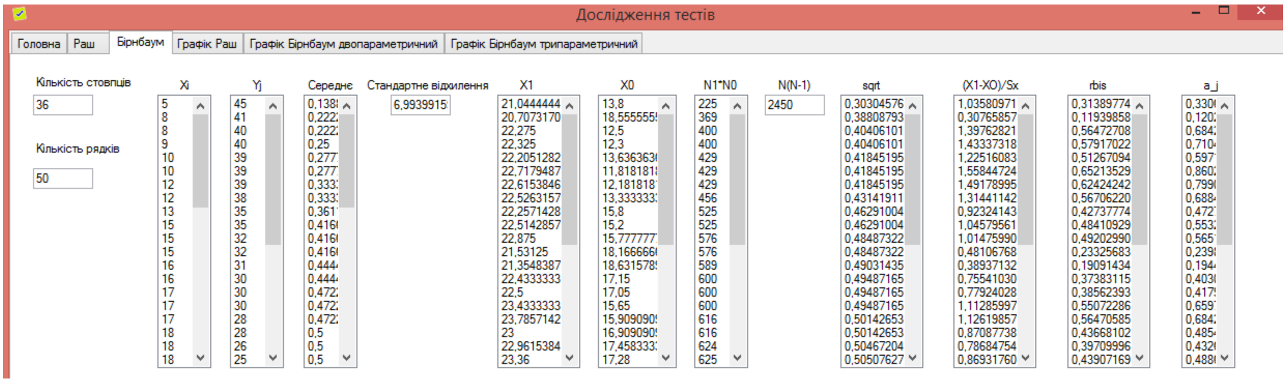

Рисунок Д.4 – Сторінка для додаткових розрахунків за моделлю А. Бірнбаума з двома параметрами

Ансамбль характеристичних кривих завдань пробного тренінгу з ЗНО 2019 з математики, знайдений за допомогою формули двопараметричної моделі А. Бірнбаума, наведено на рисунку Д.5.

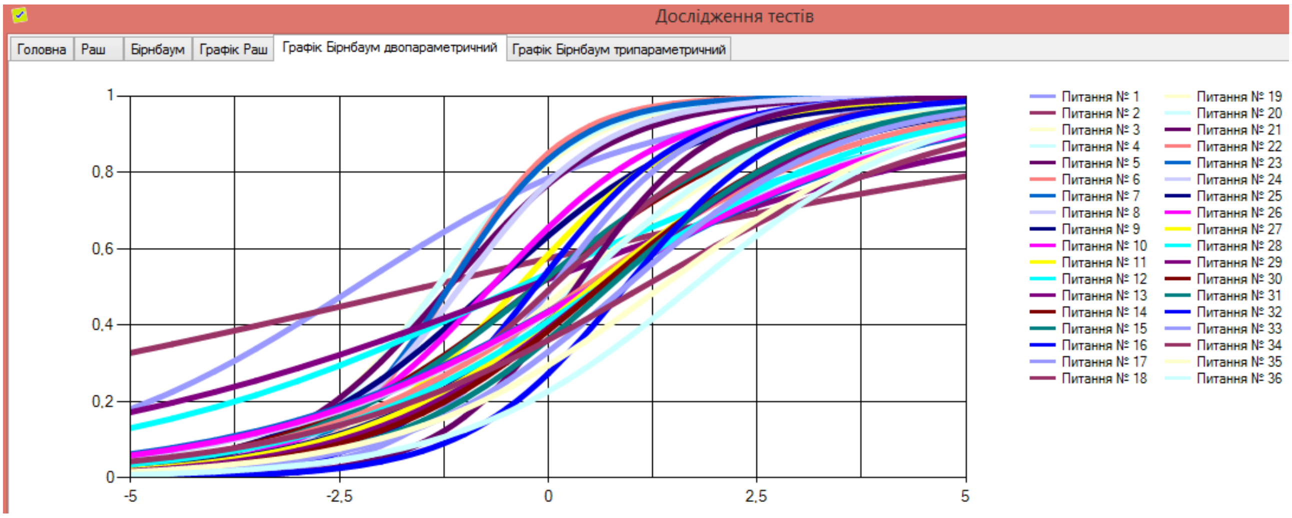

Рисунок Д.5 – Характеристичні криві 36 завдань тесту за двопараметричною моделлю А. Бірнбаума