

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
ІМЕНІ ІГОРЯ СІКОРСЬКОГО»**

**НАЦІОНАЛЬНА АКАДЕМІЯ ПЕДАГОГІЧНИХ НАУК УКРАЇНИ
ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ І ЗАСОБІВ НАВЧАННЯ**

Кваліфікаційна наукова
праця на правах рукопису

ДУДКО АННА ФЕДОРІВНА

УДК 37.091.26:004.9

**КОМП'ЮТЕРНО ОРІЄНТОВАНА МЕТОДИКА
ОЦІНЮВАННЯ ЯКОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ
ВИКЛАДАЧАМИ ЗАКЛАДІВ ВИЩОЇ ОСВІТИ**

13.00.10 – інформаційно-комунікаційні технології в освіті

01 «Освіта / Педагогіка»

Подається на здобуття наукового ступеня кандидата педагогічних наук
Дисертація містить результати власних досліджень. Використання ідей,
результатів і текстів інших авторів мають посилання на відповідне джерело

Дудко А.Ф.

Науковий керівник: **Диховичний Олександр Олександрович**, кандидат
фізико-математичних наук, доцент

Київ – 2019

АНОТАЦІЯ

Дудко А.Ф. Комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики викладачами закладів вищої освіти. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня кандидата педагогічних наук (доктора філософії) за спеціальністю 13.00.10 «Інформаційно-комунікаційні технології в освіті» (01 «Освіта / Педагогіка»). – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Інститут інформаційних технологій і засобів навчання НАПН України, Київ, 2019.

Зміст анотації

Дисертаційне дослідження присвячено проблемі оцінювання якості тестів з вищої математики викладачами закладів вищої освіти (ЗВО). Проаналізовано зарубіжний та вітчизняний досвід з теми дослідження та уточнено поняття «тест», «тестове завдання», «якість тесту», «оцінювання якості тесту», «комп'ютерно орієнтована методика оцінювання якості тестів».

Обґрунтовано необхідність поєднання методів класичної теорії тестів (КТТ) та Item Response Theory (IRT) для оцінювання якості тестів з вищої математики. Визначено основні методи оцінювання якості тестів з вищої математики в рамках КТТ, які передбачають як аналіз окремих тестових завдань, так і тесту у цілому. Аналіз тесту у цілому полягає у такому: аналізі розподілу вибірки результатів; оцінюванні надійності тесту; оцінюванні критеріальної валідності тесту. Аналіз тестових завдань передбачає: аналіз кореляційної матриці завдань; оцінювання валідності окремих завдань; аналіз дистракторів завдань множинного вибору.

Визначено моделі IRT для оцінювання якості тестів з вищої математики: Раша та Бірнбаума (для дихотомічних завдань), Андерсенена для політомічних завдань та Тіссена-Стейнберга для завдань множинного вибору. IRT-аналіз тестових завдань і тесту у цілому проводиться із

використанням визначених моделей і передбачає: оцінювання значень латентних параметрів використовуваних IRT-моделей; аналіз ансамблю характеристичних кривих завдань тесту; аналіз характеристичних кривих рівнів політомічних тестових завдань; аналіз характеристичних кривих дистракторів завдань множинного вибору.

Досліджено питання оцінювання латентних параметрів IRT-моделей. Для оцінювання латентних параметрів обрано метод максимальної вірогідності. Розглянуто його застосування на прикладі моделі Раша-Мастерса. Вирішено такі проблеми практичного оцінювання параметрів IRT-моделей: вибір початкових значень; зсув оцінок параметрів; наявність тривіальних рівнів політомічних тестових завдань.

Розроблено авторську систему автоматизованого аналізу якості тестових завдань (СААЯТЗ). Основні функціональні можливості СААЯТЗ: гнучкий вибір алгоритму статистичного опрацювання результатів тестування; наочність графічного представлення результатів; формування та обслуговування бази каліброваних завдань. В основу статистичного аналізу системи покладено методи КТТ та IRT.

Розроблено комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики викладачами ЗВО. *Метою оцінювання* якості тестів є покращення контролю знань студентів з вищої математики. Основним *засобом* оцінювання якості тестів є СААЯТЗ. *Методами* оцінювання якості тестів з вищої математики є дібрані методи теорії ймовірностей та математичної статистики, а саме методи КТТ та IRT.

Змістовий компонент оцінювання відображає загальна схема, що включає в себе такі етапи: формування таблиці результатів; аналіз окремих тестових завдань і тесту в цілому; приймання остаточних рішень. Формування таблиці результатів тестування передбачає підготовку даних до подальшої роботи з ними в СААЯТЗ. Аналіз тесту у цілому полягає у такому: аналізі розподілу вибірки результатів; оцінюванні надійності тесту; оцінюванні критеріальної валідності тесту; оцінюванні ефективності тесту.

Аналіз тестових завдань передбачає: аналіз кореляційної матриці завдань; оцінювання валідності окремих завдань; аналіз ансамблю характеристичних кривих завдань; аналіз політомічних завдань; аналіз дистракторів завдань множинного вибору. Приймання остаточних рішень здійснюється на підставі сумісного аналізу взаємодоповнюючих характеристик IRT та КТТ. При цьому робляться висновки щодо якості як окремих тестових завдань, так і тесту в цілому.

Впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики доцільно здійснювати поетапно: розвиток компетентності викладачів щодо оцінювання якості тестів та безпосереднє застосування викладачами комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики у своїй професійній діяльності.

Застосування викладачами комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики доцільно здійснювати відповідно до загальної схеми роботи викладача з тестами з вищої математики, яка передбачає: вирішення питання визначення підходу до розроблення тестів; створення тесту; оцінювання змістової валідності; завантаження тестів у систему комп'ютерного тестування (Moodle); проведення пробного тестування та вирішення питань, пов'язаних з його організацією та проведенням; імпортування результатів тестування та оцінювання якості тестів за розробленою методикою; прийняття рішення щодо якості тесту та подальшої роботи з ним; збереження тесту та завдання тесту разом з їх статистичними характеристиками у базі каліброваних завдань. Робота з тестами за описаною схемою вимагає від викладача високого рівня компетентності щодо оцінювання якості тестів з вищої математики.

Визначено *компетентність викладача щодо оцінювання якості тестів з вищої математики* як готовність та здатність викладача із застосуванням ІКТ обчислювати характеристики тесту, тестових завдань та на основі цих характеристик оцінювати якість окремих завдань і об'єктивно складати висновки щодо якості тесту в цілому, його покращення та доцільності використання в

освітньому процесі для контролю знань студентів з вищої математики. Визначено та описано основні критерії компетентності викладачів щодо оцінювання якості тестів з вищої математики: ціннісно-мотиваційний, когнітивний та операційно-діяльнісний. Представлено характеристики визначених критеріїв для трьох рівнів: базового, поглибленого та професійного.

Розроблено модель розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики, яка складається з таких компонентів: мотиваційно-цільового, змістового, операційно-технологічного, діагностичного та результативного. На підґрунті даної моделі побудовано технологію розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики.

Загальні результати педагогічного експерименту показали, що впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики сприятиме розвитку компетентності викладача щодо оцінювання якості тестів з вищої математики та покращенню контролю знань студентів.

Результати впровадження розробленої комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики викладачами ЗВО підтверджують її ефективність.

Наукова новизна та теоретичне значення полягає в тому, що *вперше* обґрунтовано комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики викладачами ЗВО на основі спільного застосування методів КТТ та IRT; визначено критерії та рівні розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики; розроблено модель розвитку компетентності викладача щодо оцінювання якості тестів з вищої математики; *уточнено* поняття тесту, тестового завдання, якості тесту, оцінювання якості тесту, комп'ютерно орієнтованої методики оцінювання якості тесту та компетентності викладача щодо оцінювання якості тестів з вищої математики; *удосконалено* методи розрахунку латентних параметрів тестових завдань; методи покращення якості тестів; *дістали подальшого*

розвитку теоретико-методичні засади створення та використання комп'ютерно-орієнтованих систем і засобів навчання, теоретичні та методичні проблеми розроблення та використання інформаційно-комунікаційних технологій для вимірювання та оцінювання рівня навчальних досягнень і моніторингу освітньої діяльності, теорія та методика підвищення кваліфікації кадрів інформатизації освіти.

Практичне значення одержаних результатів полягає в тому, що розроблено: основні компоненти комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики викладачами ЗВО; автоматизовану систему аналізу якості тестових завдань з вищої математики; технологію розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики; навчальну програму «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики» для підвищення кваліфікації наукових і науково-педагогічних працівників, розміщену на сайті Навчально-методичного комплексу «Інститут післядипломної освіти» КПІ ім. Ігоря Сікорського (<http://ipo.kpi.ua/ua/pidvishchennja-kvalifikaciji.html>); дистанційний курс за навчальною програмою «Формування компетентності щодо оцінювання якості тестів з вищої математики» в системі Moodle, розміщений на сайті Навчально-методичного комплексу «Інститут післядипломної освіти» КПІ ім. Ігоря Сікорського (<https://do.ipo.kpi.ua/course/view.php?id=16>); базу каліброваних завдань для тестів з вищої математики (1647 завдань); пакет тестів з вищої математики для студентів технічних спеціальностей ЗВО, що використовується викладачами фізико-математичного факультету КПІ ім. Ігоря Сікорського.

Теоретичні та практичні результати дисертаційної роботи можуть бути використані: у ЗВО викладачами для аналізу та підвищення якості тестів з вищої математики, призначених для контролю знань студентів; у системі підвищення кваліфікації наукових та науково-педагогічних працівників для розвитку компетентності щодо оцінювання якості тестів з вищої математики,

а також у процесі розроблення навчальних і методичних посібників з оцінювання якості тестів з вищої математики.

Проведене дослідження не вичерпує всієї повноти проблеми і може бути продовжено за такими напрямками наукового пошуку: розроблення та впровадження методики комп'ютерно орієнтованого оцінювання якості тестів інших дисциплін; дослідження інноваційних методів IRT оцінювання якості тестів та вдосконалення засобів розробленої методики.

Ключові слова: комп'ютерно орієнтована методика оцінювання якості тестів, тести з вищої математики, компетентність викладачів щодо оцінювання якості тестів, система автоматизованого аналізу якості тестових завдань, класична теорія тестів, Item Response Theory.

ABSTRACT

Dudko A.F. Computer-oriented methodology of assessing the quality of tests in higher mathematics by teachers of higher education establishments. - Qualification academic paper, manuscript.

The thesis for a Candidate Degree in Pedagogical Sciences (Philosophy Doctor) specialty 13.00.10. "Information and communication technology in education" (011 – Educational, pedagogical science). - National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Institute of Information Technologies and Learning Tools NAES of Ukraine, Kyiv, 2019.

Abstract content

The paper studies the problem of assessing the quality of tests in higher mathematics by teachers of higher education establishments. The international and domestic experience on the research issue is analyzed and the concepts of "test", "test item", "quality of the test", "test quality assessment", "computer-oriented methodology of assessing the quality of tests" are specified.

The necessity of combining methods of Classical Test Theory and Item Response Theory for assessing the quality of tests in higher mathematics is substantiated. The basic methods for assessing the quality of tests in higher mathematics in the context of CTT are determined. They include analysis of individual test items and the test in general. The test analysis consists in analysis of distribution of the results' sample, estimation of the test reliability, validity and efficiency. The analysis of test items consists in analysis of the correlation matrix of items, analysis of the ensemble of characteristic curves of items, the analysis of polytomous items, the analysis of the distractors of the multiple choice items.

The IRT-models for assessing the quality of tests in higher mathematics are determined: Rash and Birnbaum for dichotomous items, Andersen for polytomous items and Thyssen-Steinberg for multiple choice items. The IRT analysis of test items and the test in general is realized with the use of determined models and includes: evaluation of the latent parameters of IRT-models; analysis of the ensemble of characteristic curves of the test items; analysis of characteristic curves

of levels of political test items; analysis of characteristic curves of distractors of multiple choice items.

The issue of estimating of latent parameters of IRT-models is studied. The maximum likelihood is chosen for estimating of latent parameters. Its application is demonstrated by the example of the partial credit model. The following problems of practical parameter estimation of IRT-models are solved: the choice of initial values; the shift of estimated parameter; the presence of trivial levels of polytomous test items.

The author's computer-based system of quality analysis of test items is developed. The fundamental functionality of this system: flexible choice of algorithm of statistical analysis of test results; visualization of graphical presentation of results; generation and support of the base of calibrated items. Statistical analysis of the system is based on the methods of CTT and IRT.

The computer-oriented methodology of assessing the quality of the tests in higher mathematics by teachers of higher education establishments is developed. The purpose of assessing the quality of the tests is to improve the control of students' knowledge of higher mathematics. The main means of assessing the quality of tests is computer-based system of quality analysis of test items. Methods of assessing the quality of tests in higher mathematics are selected methods of probability theory and mathematical statistics, namely, methods of CTT and IRT.

The general scheme reflects the content component of assessing. It includes the following steps: generation of test matrix; analysis of test and individual test items; final decision making. The generation of test matrix is the data preparation for follow-up analysis by computer-based system of quality analysis of test items. The analysis of the test in general includes: analysis of distribution of the results' sample, estimation of the test reliability, validity and efficiency. Analysis of test items includes: analysis of the correlation matrix of items, analysis of the ensemble of characteristic curves of items, the analysis of polytomous items, and the analysis of the distractors of the multiple choice items. The final decisions are made on the basis of a consistent analysis of the complementary characteristics of IRT and

CTT. At the same time, conclusions on the quality of both individual test items and the test in general are drawn.

It is expedient to realize the introduction of a computer-based methodology of assessing the quality of tests in higher mathematics step-by-step: the development of teachers' competence in assessing the quality of tests and the application of a computer-oriented methodology for assessing the quality of tests in higher mathematics by teachers in their professional activities.

It is expedient to realize the use of computer-oriented methodology for assessing the quality of tests on higher mathematics by teachers in accordance with the general scheme of the teacher's work with tests in higher mathematics. This scheme provides: solution of the problem of determining the approach to creation of tests; creation the test; estimation of content validity; test loading into the computer test system (Moodle); trial test administration; importation of test results and quality assessment of tests based on the developed methodology; decision making about the test quality and further actions with it; retention of test and test items together with their statistical characteristics in the base of calibrated items. The use of tests requires high level teachers' competence concerning the quality of tests in higher mathematics.

The teacher's competence concerning the quality of tests in higher mathematics is defined as teacher's readiness and ability to calculate the characteristics of the test, test items with the use of ICT and on the basis of these characteristics to evaluate the quality of single item and objectively draw conclusions about the quality of the test, its improvement and the expediency of using in the learning process for knowledge control of higher mathematics. The basic criteria of teachers' competence concerning the quality of tests in higher mathematics are defined and described: value-motivational, cognitive, and operational-activity. The characteristics of the defined criteria for basic, advanced and professional levels are presented.

The development model of teachers' competence concerning the quality of tests in higher mathematics is proposed. It consists of the following components:

motivational, content, technological, diagnostic and resultative. The development methodology of teachers' competence concerning the quality of tests in higher mathematics is based on this model.

The general results of the pedagogical experiment gives evidences to confirm that the introduction of the computer-oriented methodology for assessing the quality of tests in higher mathematics will promote the development of teachers' competence concerning the quality of tests in higher mathematics and improvment of knowledge control.

The result of introduction of the developed computer-oriented methodology of assessing the quality of tests in higher mathematics by teachers of higher education establishments confirms its effectiveness.

Scientific novelty and theoretical meaning of the research: *for the first time* the computer-oriented methodology of assessing the quality of tests in higher mathematics by teachers of higher education establishments has been substantiated based on interaction of the methods of KTT and IRT; the criteria and levels of teachers' competence concerning the quality of tests in higher mathematics has been defined; the development model of teachers' competence concerning the quality of tests in higher mathematics has been developed; the concepts of test, test item, quality of the test, test quality assessment and computer-oriented methodology of assessing the quality of tests have been *specified*; methods of test quality improvement and methods of estimation of latent parameters of test items have been improved; *future developments* of theoretical and methodical foundations of the creating and using of computer-based systems and tutorials, theoretical and methodical problems of development and use of information and communication technologies for measurement and evaluation of academic performance and monitoring of educational activities, the theory and methods of training, retraining and advanced training of informatization of education have been suggested.

Practical meaning of the received results: the main components of the computer-oriented methodology of assessing the quality of tests in higher

mathematics by teachers of higher education establishments has been developed; the computer-based system of quality analysis of test items has been developed; the development technology of teachers' competence concerning the quality of tests in higher mathematics has been developed; the curriculum "Development of the teachers' competence concerning the quality of tests in higher mathematics" for the advanced training course of scientific, scientific and pedagogical staff has been developed and placed on the site of the Educational methodical complex "Institute of Postgraduate Education" of Igor Sikorsky Kyiv Polytechnic Institute (<http://ipo.kpi.ua/ua/pidvishchennja-kvalifikaciji.html>); the distance course "Development of the teachers' competence concerning the quality of tests in higher mathematics" has been developed and placed on the site of the Educational methodical complex "Institute of Postgraduate Education" of Igor Sikorsky Kyiv Polytechnic Institute (<http://ipo.kpi.ua/ua/pidvishchennja-kvalifikaciji.html>); the base of calibrated items in higher mathematics (1647 items); the test package in higher mathematics for students of technical specialties higher education establishments, used by teachers of Faculty of Physics and Mathematics of Igor Sikorsky Kyiv Polytechnic Institute, has been developed.

The materials and the results of the research can be used by teachers of mathematics in higher education establishments for analysis and improvement of the quality of tests in higher mathematics, in the system of advanced training of scientific, pedagogical and scientific staff for development of teachers' competence concerning the quality of tests in higher mathematics and in the process of developing educational and methodological manuals for the assessment of the quality of tests in higher mathematics.

The research does not exhaust the completeness of the problem and can be proceeded in the following areas of scientific research: development and introduction of the computer-oriented methodology of assessing the quality of tests of other branches of science; research of innovative methods of IRT for assessing the quality of tests and improving the developed methodology

Keywords: computer-oriented methodology of assessing the quality of

tests, tests in higher mathematics, teachers' competence concerning the quality of tests in higher mathematics, computer-based system of quality analysis of test items, classical test theory, Item Response Theory.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковано основні наукові результати

1. Дудко А. Ф., Диховичний О. О. Автоматизована система аналізу результатів комп'ютерного тестування з вищої математики. *Наукові праці ДонНТУ. Серія: «Педагогіка, психологія і соціологія»*. 2013. №2 (14). С. 103-110.
2. Дудко А. Ф., Диховичний О. О. Застосування інформаційної функції для аналізу та підвищення ефективності тестів з вищої математики [Електронний ресурс]. *Інформаційні технології і засоби навчання*. 2014. Том 41. №3. С. 55-69. URL: <http://journal.iitta.gov.ua/index.php/itlt/article/view/1040/790#.VFK7rzSsXm4> (включений до міжнар. наукометрич. баз).
3. Дудко А. Ф., Диховичний О. О. Практичний аналіз якості тестів з вищої математики. *Науковий вісник Ужгородського національного університету*. Серія «Педагогіка. Соціальна робота». 2014. № 32. С. 73-78.
4. Дудко А. Ф. Концепція розробки комп'ютерних тестів з вищої математики. *Науковий часопис Національного педагогічного університету імені М. П. Драгоманова*. Серія 5. Педагогічні науки: реалії та перспективи: зб. наук. пр. Київ: Вид-во НПУ імені М. П. Драгоманова. 2014. Випуск 50. С. 44-48.
5. Дудко А. Ф., Диховичний О. О. Комплексна методика аналізу якості тестів з вищої математики. *Науковий часопис НПУ імені М.П. Драгоманова*. Серія №2. Комп'ютерно-орієнтовані системи навчання: Зб. наук. праць. К.: НПУ імені М. П. Драгоманова. 2015. № 15 (22). С. 140-144.
6. Дудко А. Ф., Диховичний О. О. Оцінювання змістовної валідності тестів з вищої математики з використанням Google Docs додатку [Електронний ресурс]. *Інформаційні технології і засоби навчання*. 2016. Том 52 №2. С. 52-61 <http://journal.iitta.gov.ua/index.php/itlt/article/view/1383/1032> (включений до міжнар. наукометрич. баз).
7. Дудко А. Ф., Диховичний О. О. Аналіз поняттєво-категоріального апарату дослідження якості тестів з вищої математики. *Journal «ScienceRise: Pedagogical Education»*. 2017. № 5(13). С. 17-21. (включений до міжнар. наукометрич. баз)

8. Удовенко А.Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Аналіз якості тестових завдань для комплекту дистанційної освіти «Вища математика». *Теорія та методика електронного навчання*: зб. наук. пр. Кривий Ріг: Видавничий відділ НМетАУ. 2010. Вип. І. С. 3-9.

9. Удовенко А.Ф., Диховичний О. О., Шепель М. О. Політомічні моделі Мастерса та Андерсена в аналізі якості тестових завдань. *Теорія та методика електронного навчання*: збірник наукових праць. Кривий Ріг: Видавничий відділ НМетАУ. 2012. Випуск III. С. 83-87.

Наукові праці, які засвідчують апробацію матеріалів дисертації

10. Удовенко А.Ф. Сучасні математичні моделі педагогічного тестування. *«Історія розвитку науки, техніки та освіти»*: Зб. пр. VIII міжнародної молодіжної науково-практ. конф. за темою «Взаємодія академічної та університетської науки», 18 травня 2010 р. Київ. 2010. С. 202-206.

11. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Статистичний аналіз тестових завдань із застосуванням сучасних математичних моделей IRT. *Вісник Національного університету «Львівська політехніка» (серія «Інформатизація вищого навчального закладу»)*. 2012. №731. С. 10-13.

12. Дудко А.Ф., Диховичний О. О., Шепель М. О. Про особливості застосування політомічних математичних моделей тестових завдань. *Чотирнадцята міжнар. Конф. імені академіка М. Кравчука*, 19-21 квітня, 2012 р. Київ: Матеріали конф. Т.4. Історія та методика викладання математики. К.: НТУУ «КПІ». 2012. С. 91.

13. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про досвід застосування тестових контрольних робіт з вищої математики. *Міжнар. наук. конф. «Математика в сучасному технічному університеті»*, 19-20 квітня 2013 р., м. Київ. 2013. С. 442-445.

14. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про електронне навчання вищої математики в НТУУ «КПІ». *Міжнар. наук. конф. «Математика в сучасному технічному університеті»*, 19-20 квітня 2013 р., м. Київ. 2013. С. 438-441.

15. Дудко А. Ф., Диховичний О. О., Тарабара І. С. Статистичний аналіз якості тестових завдань для моделі з множинним вибором. *Третя міжнар. наук. конф. молодих вчених з математики та фізики*, 25-27 квітня, 2013 р., м. Київ. 2013. С. 169-170.
16. Дудко А. Ф., Диховичний О. О., Філіппова Г. В. Система автоматизованого аналізу якості тестових завдань. *Третя міжнар. наук. конф. молодих вчених з математики та фізики*, 25-27 квітня, 2013 р., м. Київ. 2013. С. 170-172.
17. Dudko A. F., Dykhovichnyi O. O. On the application of IRT-models for the statistical analysis of the quality of educational tests in higher mathematics. *Ukrainian-German Workshop on Empirical Complete Convergence and other Limit Theorems of Probability Theory*, September 23-27, 2013, Koktebel, Ukraine. 2013.
18. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про оцінку ефективності тестів з вищої математики. *Друга міжнародна наук. Конф. «Математика в сучасному технічному університеті»*, 20-21 грудня, 2013 р., м. Київ. 2013. С. 271-274.
19. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про електронне тестування з вищої математики в НТУУ «КПІ». *Вища освіта: проблеми і шляхи забезпечення якості. X Всеукр. науково-метод. конф.*, 28-29 листопада, 2013 р., м. Київ, НТУУ «КПІ». 2013. С. 7.
20. Дудко А. Ф., Диховичний О. О. Методика аналізу та підвищення ефективності тестів з вищої математики. *П'ятнадцята міжнар. конф. імені академіка М. Кравчука*, 15-17 травня, 2014 р. Київ: Матеріали конф. Т.4. Історія та методика математики. К.: НТУУ «КПІ». 2014. С. 96-97.
21. Дудко А. Ф., Диховичний О. О. Застосування IRT-моделей у дослідженні якості педагогічних тестів. *Всеукр. наукова конф. «Стохастичні та детерміновані математичні моделі у наукових дослідженнях»*, 27-29 червня 2014 р., м. Київ. 2014.
22. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Вибір підходу до розроблення комп'ютерних тестів з вищої математики. *Математика в сучасному технічному університеті*:

матеріали III Міжнар. наук.-практ. конф. (Київ, 25–26 груд. 2014 р.). К.: НТУУ «КПІ». 2015. С. 317-318.

23. Дудко А. Ф., Диховичний О. О. Методика аналізу якості тестів та її застосування. *Четверта міжнар. наук. конф. молодих вчених з математики та фізики*, 23-25 квітня, 2015 р. – С. 78.

24. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Аналіз якості комп'ютерних тестів методами класичної теорії тестів. *Шістнадцята міжнар. конф. імені академіка М. Кравчука*, 14-15 травня, 2015 р. Київ: Матеріали конф. Т. 3. Теорія ймовірностей та математична статистика. Історія та методика математики. К.: НТУУ «КПІ». 2015. С. 94-97.

25. Дудко А. Ф., Диховичний О. О. Комп'ютерно-орієнтована методика статистичного аналізу тестів з вищої математики. *П'ята міжнар. наук. конф. молодих вчених з математики та фізики*, 25-26 квітня, 2016 р., м. Київ. 2016. – С. 71.

26. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Комп'ютерно-орієнтована методика аналізу якості педагогічних тестів як засіб ІК-підтримки викладачів вищої математики НТУУ «КПІ». *Математика в сучасному технічному університеті: матеріали IV Міжнар. наук.-практ. конф. (Київ, 24–25 груд. 2015 р.)*. К.: НТУУ «КПІ». 2016. С. 277-278.

27. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Оцінювання якості змісту тестів з вищої математики з використанням Google Docs додатку. *Сімнадцята міжнар. конф. імені академіка М. Кравчука*, 19-20 травня, 2016 р. Київ: Матеріали конф. Т. 3. К.: НТУУ «КПІ». 2015. С. 176-179.

28. Dudko A., Dykhovychnyi A. Statistical analysis of tests in higher mathematics. *Limit theorem in probability theory, number theory and mathematical statistics. International workshop in honour of Prof. V.V. Buldygin (October 10-12, 2016)*. NTUU «KPI» Kyiv, Ukraine. 2016. – P. 18.

29. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Схема роботи викладача з тестами для контролю знань студентів з вищої математики. *Математика в сучасному технічному*

університеті: матеріали V Міжнар. наук.-практ. конф. (Київ, 29-30 грудня 2016 р.). Київ : НТУУ «КПІ». 2017. С. 229-232.

30. Дудко А. Ф., Диховичний О. О., Прохоренко Н. В. Застосування мови R у статистичному аналізі якості тестів з вищої математики. *Матеріали Вісімнадцятої міжнар. наук. конф. імені академіка Михайла Кравчука, 7-10 жовтня 2017 року*, Київ: Т. 2. Київ: НТУУ «КПІ». 2017. С. 222-224.

31. Дудко А. Ф., Диховичний О. О., Прохоренко Н. В. Моделювання даних результатів тестування з вищої математики за допомогою мови R. *Матеріали Вісімнадцятої міжнар. наук. конф. імені академіка Михайла Кравчука, 7-10 жовтня 2017 року*, Київ: Т. 2. Київ: НТУУ «КПІ». 2017. С. 225-226.

32. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про необхідність розвитку компетентності щодо оцінювання якості тесту викладача вищої математики. *Матеріали Вісімнадцятої міжнар. наук. конф. імені академіка Михайла Кравчука, 7-10 жовтня 2017 року*, Київ: Т. 2. Київ: НТУУ «КПІ». 2017. С. 166-167.

33. Дудко А. Ф., Диховичний О. О., Круглова Н. В., Ординська З. П. Застосування моделі Бірнбаума до визначення ставлення викладачів вищої математики до тестового контролю знань. *Сьома всеукр. наук. конф. студентів, аспірантів та молодих вчених з математики, 18-20 квітня, 2018 р., м. Київ. 2018. С. 73.*

ЗМІСТ

ВСТУП.....	22
РОЗДІЛ I. ОСНОВНІ ЗАСАДИ ТЕСТОВОГО КОНТРОЛЮ ЗНАНЬ СТУДЕНТІВ ЗАКЛАДІВ ВИЩОЇ ОСВІТИ З ВИЩОЇ МАТЕМАТИКИ	31
1.1. Основні поняття щодо оцінювання якості тестів.....	31
1.2. Тестування як засіб контролю знань студентів	47
1.3. Сучасний стан дослідження оцінювання якості тестів.....	59
<i>Висновки до розділу 1.....</i>	<i>70</i>
РОЗДІЛ II. ПРОЕКТУВАННЯ КОМП'ЮТЕРНО ОРІЄНТОВАНОЇ МЕТОДИКИ ОЦІНЮВАННЯ ЯКОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ	74
2.1. Загальна методика дослідження	74
2.2. Аналіз математичних методів оцінювання якості тестів.....	77
2.2.1. Аналіз методів класичної теорії тестів	78
2.2.2. Аналіз методів Item Response Theory.....	86
2.2.3. Оцінювання ефективності тестів	99
2.2.4. Оцінювання латентних параметрів.....	106
2.3. Проектування системи автоматизованого аналізу якості тестових завдань	114
<i>Висновки до розділу 2.....</i>	<i>120</i>
РОЗДІЛ III. ОБҐРУНТУВАННЯ КОМП'ЮТЕРНО ОРІЄНТОВАНОЇ МЕТОДИКИ ОЦІНЮВАННЯ ЯКОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ	123
3.1. Розроблення тестів з вищої математики.....	123
3.2. Комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики	139

3.3. Практичне застосування методики та особливості роботи викладача з тестами з вищої математики.....	148
3.4. Критерії та рівні розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики.....	166
3.5. Модель та технологія розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики.....	171
<i>Висновки до розділу 3.....</i>	188
РОЗДІЛ IV. ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА КОМП'ЮТЕРНО ОРІЄНТОВАНОЇ МЕТОДИКИ ОЦІНЮВАННЯ ЯКОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ.....	193
4.1. Проведення та результати педагогічного експерименту	193
4.2. Експертне оцінювання якості контролю знань студентів з вищої математики.....	211
<i>Висновки до розділу 4.....</i>	217
ВИСНОВКИ	220
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	225
ДОДАТКИ	248

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

ЕГ	– експериментальна група
ЄДЕ	– єдиний державний екзамен
ЗВО	– заклад (-и) вищої освіти
ЗНО	– зовнішнє незалежне оцінювання
ІЗ	– індивідуальні завдання
ІКТ	– інформаційно-комунікаційні технології
ІК-підтримка	– інформаційно-комунікаційна підтримка
КГ	– контрольна група
КОТ	– критеріально орієнтований тест
КТТ	– класична теорія тестів
НОТ	– нормативно орієнтований тест
ПК	– персональний комп'ютер
РСО	– рейтингова система оцінювання знань
СААЯТЗ	– система автоматизованого аналізу якості тестових завдань
GRE	– Graduate Record Examinations
IRF	– Item Response Functions
IRT	– Item Response Theory
Moodle	– Modular Object-Oriented Dynamic Learning Environment
SAT	– Scholastic Aptitude Tests

ВСТУП

Актуальність досліджуваної теми. У часи модернізації системи освіти підвищення якості вищої технічної освіти набуває особливої актуальності. Дана проблема породжує ряд задач, однією з яких є задача покращення контролю знань студентів. На сьогоднішній день центральне місце серед методів контролю знань посідає тестування. У зв'язку із всебічним впровадженням інформаційно-комунікаційних технологій, що відповідають запитам сучасного інформаційного суспільства, особливого значення набуває комп'ютерне тестування.

Питанню комп'ютерного тестування присвячено роботи В. Ю. Бикова [18, 19, 20], Ю. М. Богачкова [20], І. Є. Булах [26], Ю. О. Жука [18, 20], В. С. Кіма [91], Л. О. Кухар [104], О. Г. Колгатіна [94, 95, 96], М. Р. Мруги [25], С. А. Ракова [18], В. П. Сергієнка [104], М. Б. Челишкової [160, 162]. Основними перевагами комп'ютерного оцінювання навчальних досягнень студентів є:

- автоматизація та оперативність оброблення результатів тестування;
- прискорення зворотного зв'язку за результатами тестування, що надає студенту можливість самостійно виявляти прогалини в структурі своїх знань і вживати заходів до їх ліквідації;
- забезпечення об'єктивності оцінювання;
- можливість регулярно поповнювати та модифікувати систему тестових завдань;
- звільнення викладача від виконання рутинних робіт.

Незважаючи на переваги тестування, застосування тестів з вищої математики для студентів технічних спеціальностей на практиці досі не стало масштабним з таких причин: прихильність викладачів до традиційних методів контролю та недовіра до результатів тестування; недостатня інтеграція тестових технологій контролю в процеси навчання і атестації студентів; тестовий контроль вимагає забезпечення комп'ютерної бази для

його проведення; для охоплення тестуванням великої кількості студентів необхідно створити потужну базу тестових завдань; процес розроблення та аналізу педагогічного тесту досить тривалий і трудомісткий, потребує високого рівня компетентності викладачів. Проте необхідно зауважити, що тестування стає більш розповсюдженим і тести все ширше застосовуються у вищій та середній школах при підсумковому контролі. У зв'язку з цим виникає необхідність пошуку шляхів підвищення якості педагогічних тестів з метою забезпечення достовірного оцінювання навчальних досягнень студентів.

Процес оцінювання якості тестів пов'язаний з опрацюванням великих обсягів даних та із застосуванням складних математико-статистичних методів, тому вимагає відповідної ІК-підтримки викладача і передбачає розроблення комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики. Реалізація такої методики сприятиме підвищенню ефективності роботи викладача вищої математики, рівня контролю знань студентів, що, в свою чергу, надасть можливість покращити якість освіти в цілому.

Основу аналізу якості тестів традиційно складають методи теорії ймовірностей та математичної статистики, а саме методи КТТ та сучасної теорії тестів IRT. Методи КТТ опрацювання результатів тестування розроблені та досліджені в наукових працях Дж. Алгіни (J. Algina) [99], А. Анастасі (A. Anastasi) [10], Г. Гулліксена (H. Gulliksen) [179], Л. Гуттмана (L. Guttman) [180], В. І. Звоннікова [80], В. С. Кіма [91], Л. Крокер (L. Croker) [99], Л. О. Кухар, В. П. Сергієнко [104], Р. К. Хемблтона (R. K. Hambleton) [187], Ч. Е. Спірмена (C. E. Spearman) [207, 208], Н. Б. Челишкової [162], методи IRT – в працях В. С. Аванесова [8], Ф. Бейкера (F. Baker) [168], А. Бірнбаума (A. Birnbaum) [169], Г. Вайнера (H. Wainer) [211], Д. Ендріча (D. Andrich) [167], Н. Ф. Єфремової [70], В. І. Звоннікова [80], Р. Лінн (R. Linn) [188], Т. В. Лісової [107], Ф. М. Лорда (F. M. Lord) [189], Г. Раша (G. Rasch) [198], Р. К. Хемблтон (R. K. Hambleton) [182], Н. Б. Челишкової.

Вітчизняними науковцями, які на основі сучасних математичних теорій впроваджують комп'ютерно орієнтовані технології, розглянуто теоретичне розроблення і практичне застосування методів тестології для застосування в навчальному процесі (В. Ю. Биков, Ю. М. Богачков, Ю. О. Жук, С. А. Раков).

Актуальність дослідження зумовлена його спрямованістю на розв'язання *суперечностей* між:

- доцільністю впровадження комп'ютерного тестування для контролю знань з вищої математики та недостатнім рівнем його поширення;
- необхідністю оцінювання викладачами якості тестів з вищої математики та недостатнім рівнем розробленості технологій розвитку відповідної компетентності викладачів;
- необхідністю забезпечення високих показників валідності, надійності та ефективності тестів з вищої математики та недостатньою розробленістю комп'ютерно орієнтованих засобів оцінювання якості тестів з вищої математики.

Проблемою дослідження є наукове обґрунтування та впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики викладачами ЗВО. Актуальність дослідження, недостатня розробленість означеної проблеми та необхідність розв'язання окреслених суперечностей зумовили вибір теми дисертаційної роботи *«Комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики викладачами закладів вищої освіти»*.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконана в Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського» у межах НДР «Статистичний аналіз якості тестів з вищої математики» (ДР № 0116U008958, 2016-2020 рр.), що здійснюється на кафедрі математичного аналізу та теорії ймовірностей, одним із виконавців якої був дисертант.

Тему дисертаційної роботи затверджено на засіданні Вченої ради Національного технічного університету України «Київський політехнічний

інститут імені Ігоря Сікорського» (протокол № 2 від 4.02.2013), перезатверджено на засіданні Вченої ради фізико-математичного факультету (протокол № 1 від 22.01.2016 р.) та узгоджено Міжвідомчою радою з координації наукових досліджень з педагогічних та психологічних наук в Україні при НАПН України (протокол № 3 від 29.03.2016).

Мета дослідження – розробити комп’ютерно орієнтовану методику оцінювання якості тестів з вищої математики викладачами закладів вищої освіти.

Досягнення поставленої мети передбачає розв’язання таких **завдань**:

1. Визначити теоретичні засади тестового контролю знань студентів з вищої математики та основні поняття дослідження.
2. Визначити та обґрунтувати вибір статистичних методів КТТ та IRT для аналізу якості тестів з вищої математики та їхньої програмної реалізації.
3. Спроекувати систему автоматизованого аналізу якості тестових завдань з вищої математики.
4. Визначити критерії та рівні розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики, розробити модель та технологію розвитку цієї компетентності.
5. Розробити та описати основні компоненти комп’ютерно орієнтованої методики оцінювання якості тестів з вищої математики та перевірити її ефективність експериментальним шляхом.

Об’єкт дослідження – розвиток компетентності викладачів закладів вищої освіти щодо оцінювання якості тестів з вищої математики.

Предмет дослідження – комп’ютерно орієнтована методика оцінювання якості тестів з вищої математики викладачами закладів вищої освіти.

Для досягнення мети і реалізації завдань дослідження застосовувався комплекс **методів**: *теоретичні* – аналіз, порівняння, узагальнення психолого-педагогічної, тестологічної, методичної літератури для обґрунтування теоретичних положень та уточнення основних понять дисертаційного

дослідження; аналіз та систематизація алгоритмів і методів опрацювання результатів тестування для розроблення комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики; *емпіричні* – математико-статистичні методи для аналізу якості тестів з вищої математики; педагогічний експеримент для перевірки ефективності комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики викладачами ЗВО; анкетування, тестування, спостереження з метою визначення рівня розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики; експертне опитування з метою оцінювання змістової валідності тестів з вищої математики та для проведення експертами педагогічного оцінювання стану контролю знань студентів з вищої математики; моделювання педагогічних процесів для розроблення моделі розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики; математично-статистичні методи для кількісного та якісного аналізу результатів педагогічного експерименту і перевірки гіпотези дослідження.

Наукова новизна та теоретичне значення полягає в тому, що

- *вперше* обґрунтовано комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики викладачами ЗВО на основі спільного застосування методів КТТ та IRT; визначено критерії та рівні розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики; розроблено модель розвитку компетентності викладача щодо оцінювання якості тестів з вищої математики;
- *уточнено* поняття тесту, тестового завдання, якості тесту, оцінювання якості тесту, комп'ютерно орієнтованої методики оцінювання якості тесту та компетентності викладача щодо оцінювання якості тестів з вищої математики;
- *удосконалено* методи розрахунку латентних параметрів тестових завдань; методи покращення якості тестів;
- *дістали подальшого розвитку* теоретико-методичні засади

створення та використання комп'ютерно-орієнтованих систем і засобів навчання, теоретичні та методичні проблеми розроблення та використання інформаційно-комунікаційних технологій для вимірювання та оцінювання рівня навчальних досягнень і моніторингу освітньої діяльності, теорія та методика підвищення кваліфікації кадрів інформатизації освіти.

Практичне значення одержаних результатів дослідження полягає в тому що розроблено:

- основні компоненти комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики викладачами ЗВО;
- автоматизовану систему аналізу якості тестових завдань з вищої математики;
- технологію розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики;
- навчальну програму «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики» для підвищення кваліфікації наукових і науково-педагогічних працівників, розміщену на сайті Навчально-методичного комплексу «Інститут післядипломної освіти» КПІ ім. Ігоря Сікорського (<http://ipo.kpi.ua/ua/pidvishchennja-kvalifikaciji.html>);
- дистанційний курс за навчальною програмою «Формування компетентності щодо оцінювання якості тестів з вищої математики» в системі Moodle, розміщений на сайті Навчально-методичного комплексу «Інститут післядипломної освіти» КПІ ім. Ігоря Сікорського (<https://do.ipo.kpi.ua/course/view.php?id=16>);
- базу каліброваних завдань для тестів з вищої математики (1647 завдань);
- пакет тестів з вищої математики для студентів технічних спеціальностей ЗВО, що використовується викладачами фізико-математичного факультету КПІ ім. Ігоря Сікорського.

Теоретичні та практичні результати дисертаційної роботи можуть

бути використані: у ЗВО викладачами для аналізу та підвищення якості тестів з вищої математики, призначених для контролю знань студентів; у системі підвищення кваліфікації наукових та науково-педагогічних працівників для розвитку компетентності щодо оцінювання якості тестів з вищої математики, а також у процесі розроблення навчальних і методичних посібників з оцінювання якості тестів з вищої математики.

Особистий внесок здобувача. У працях, опублікованих у співавторстві, автору належать такі результати: проаналізовано поняттєво-категоріальний апарат дослідження якості тестів з вищої математики [41]; досліджено питання, пов'язані із застосуванням тестів для контролю знань студентів з вищої математики [59, 60, 61]; досліджено методи КТТ для оцінювання якості тестів [42, 153, 175]; проаналізовано основні сучасні математичні моделі педагогічного тестування [47, 174, 175]; досліджено моделі IRT для аналізу дихотомічних завдань [67, 153]; досліджено моделі IRT для аналізу політомічних завдань [154, 63]; досліджено моделі IRT з множинним вибором [66]; досліджено методи IRT для оцінювання ефективності тестів [64]; розроблено методику аналізу та підвищення ефективності тестів з вищої математики із застосуванням інформаційної функції [44, 52]; обґрунтовано вибір концепції розробки комп'ютерних тестів з вищої математики [43]; досліджено питання аналізу якості педагогічних тестів на підставі застосування взаємодоповнюючих методів КТТ та IRT, розроблено комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики [48, 49, 50, 53]; розглянуто основні функціональні можливості СААЯТЗ [40, 65]; розглянуто практичний аналіз якості тестів на основі комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики [58]; розглянуто особливості роботи викладача з тестами з вищої математики [155]; проаналізовано основні програмні засоби опрацювання результатів тестування [46]; досліджено питання аналізу змістової валідності комп'ютерних тестів з вищої математики [56, 57]; досліджено застосування моделювання даних результатів тестування з вищої математики в процесі

розвитку компетентності викладачів щодо оцінювання якості тестів [55]; обґрунтовано необхідність розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики [62], розроблено тести для визначення ставлення викладачів вищої математики до тестового контролю знань [45].

Апробація. Основні теоретичні та практичні результати проведеного дослідження, а також концептуальні положення й загальні висновки були представлені у вигляді доповідей, зокрема на: *міжнародних конференціях*: VIII молодіжна науково-практична конференція «Історія розвитку науки, техніки та освіти» за темою «Взаємодія академічної та університетської науки» (2010 р., Київ); «Конференція імені академіка М. Кравчука» (2012, 2014-2016, 2017 рр., Київ), «Наукова конференція молодих вчених з математики та фізики» (2013, 2015-2016, 2018 рр., Київ), «Математика в сучасному технічному університеті» (2013-2015 рр., Київ), «Вища освіта: проблеми і шляхи забезпечення якості» (2013 р., Київ); *всеукраїнській конференції* «Стохастичні та детерміновані математичні моделі в наукових дослідженнях» (2014 р., Київ).

Результати дослідження апробовано на семінарах для наукових і науково-педагогічних працівників: на семінарах кафедри математичного аналізу та теорії ймовірностей КПІ ім. Ігоря Сікорського (Київ, 2012-2017 рр.), на Всеукраїнському методологічному семінарі для молодих учених «Інформаційно-комунікаційні технології в освіті та наукових дослідженнях» в Інституті інформаційних технологій і засобів навчання НАПН України (м. Київ, 2015, 2017-2018 рр.), на Всеукраїнському науково-методичному семінарі «Системи навчання і освіти в комп'ютерно орієнтованому середовищі» (Київ, 2018 р.).

Результати дослідження впроваджено в навчальний процес Національного транспортного університету (довідка від 25.06.2018 р. № 1524/01), Національного педагогічного університету імені М. П. Драгоманова (довідка від 26.06.2018 р. № 07-10/1035), та в Навчально-методичному комплексі «Інституту післядипломної освіти» Національного

технічного університету України «Київського політехнічного інституту імені Ігоря Сікорського» (довідка від 10.07.2018 р. № 90/07-18).

Публікації. За матеріалами дослідження опубліковано 33 наукові праці, з них 7 статей у фахових виданнях (3 статті у виданнях, внесених до міжнародних наукометричних баз даних), 3 статті в інших наукових виданнях, 23 тези доповідей у матеріалах конференцій.

Структура дисертації. Робота складається з анотацій, переліку умовних позначень і скорочень, вступу, чотирьох розділів, висновків до кожного розділу, висновків, списку використаних джерел (212 найменування, з них – 46 іноземними мовами) та додатків. Загальний обсяг дисертації – 293 сторінки, з них 224 сторінки основного тексту, в якому міститься 11 таблиць та 55 рисунків. Додатки розміщено на 46 сторінках.

РОЗДІЛ І. ОСНОВНІ ЗАСАДИ ТЕСТОВОГО КОНТРОЛЮ ЗНАНЬ СТУДЕНТІВ ЗАКЛАДІВ ВИЩОЇ ОСВІТИ З ВИЩОЇ МАТЕМАТИКИ

1.1. Основні поняття щодо оцінювання якості тестів

Для подальшого дослідження необхідно з'ясувати та уточнити поняття «тест», «тестове завдання», «якість тесту», «оцінювання якості тесту», «комп'ютерно орієнтована методика оцінювання якості тесту». Слово «тест» походить від англійського слова «test» і означає випробування або дослідження. В педагогічній літературі можна зустріти різні визначення поняття «тесту», які суттєво відрізняються одне від одного.

За В. П. Беспальком [15] тестами можна назвати завдання на виконання діяльності певного рівня складності в поєднанні з системою оцінювання. Порівняно з традиційними засобами контролю тестування є більш ефективним методом оцінювання знань. Правильно складені тести дають змогу за той самий час перевірити знання з більш широкої області, ніж інші види контролю. Крім того, тестування (зокрема, комп'ютерне тестування) викликає додаткову зацікавленість студентів. Для дистанційних навчальних курсів тест є основним засобом для швидкого моніторингу процесу навчання студентів, який також сприяє організації самостійної роботи в навчанні через самоконтроль [80].

М.Б. Челишкова поняття «тесту» тлумачить як сукупність завдань, які дозволяють дати об'єктивну, таку, що можна порівняти, і навіть кількісну оцінку якості підготовки іспитника в заданій освітній галузі [160]. Тестом можна вважати єдину зв'язану систему завдань, яка відповідає вимогам валідності й надійності [3]. За визначенням А. О. Кузнецова, тест – це стандартизовані, короткі, обмежені в часі випробування, призначені для встановлення кількісних і якісних індивідуальних розбіжностей [102].

У теорії педагогічних вимірювань існує багато визначень поняття "*педагогічний тест*". Деякі дослідники, розглядаючи це поняття, значно розширюють його зміст. Так за О. М. Майоровим тест є інструментом, що складається з кваліметрично вивіреної системи тестових завдань,

стандартизованої процедури проведення і наперед спроектованої технології опрацювання та аналізу результатів [109]. Наприклад, Т. Балихіна наводить близько двох десятків визначень і видів тестів, різних за призначенням [13].

Іноді поняття «педагогічний тест» розглядається у двох значеннях: як метод педагогічного вимірювання і як результат використання тесту, що складається з сукупності завдань [36]. Останнім часом з'явилися терміни, які враховують обидва значення тесту: і як методу, і як результату.

У дослідженні під поняттям *тест* будемо розуміти систему завдань специфічної форми, певного змісту для об'єктивного оцінювання рівня навчальних досягнень студентів з наперед заданою методикою аналізу результатів.

Основним напрямом досліджень в педагогічних вимірюваннях є розроблення якісних тестів [3]. *Якість педагогічного тесту* задається ще на етапі його розроблення через дотримання рекомендованих вимог. Процедура розробки однакова для всіх видів педагогічних тестів. Спеціалісти з тестології такі як А. Анастасі, В.С. Аванесов, А.Н. Майоров, В.Ю. Переверзєв та ін. описують різну кількість етапів даної процедури. Найбільш повний опис процедури розроблення тестів представлено у роботі М.Б. Челишкової. Вона виділяє і описує такі етапи [162]:

- визначення мети тестування, вибір виду тесту і підходу до його створення;
- аналіз змісту навчальної дисципліни;
- визначення структури тесту і стратегії розташування завдань;
- розроблення специфікації тесту, вибір довжини тесту і часу його виконання;
- створення передтестових завдань;
- відбір завдань в тест і їх ранжування відповідно до обраної стратегії пред'явлення на підставі авторських оцінок складності завдань;
- експертиза змісту передтестових завдань і тесту;
- експертиза форми передтестових завдань;

- зміна змісту і форми завдань за результатами експертизи;
- розроблення методики пробного тестування;
- розроблення інструкцій для студентів і для викладачів, які проводять апробацію тесту;
- проведення пробного тестування;
- збирання емпіричних результатів;
- статистичне опрацювання результатів тестування;
- інтерпретація результатів опрацювання з метою поліпшення якості тесту. Перевірка відповідності характеристик тесту науково обґрунтованим критеріям якості;
- корекція змісту і форми завдань на підставі даних попереднього етапу. Видалення невдалих та додавання нових завдань для оптимізації діапазону значень параметра складності. Оптимізація довжини тесту і часу його виконання. Оптимізація порядку розташування завдань у тесті;
- повторення етапу апробації для виконання чергових кроків щодо підвищення якості тесту;
- інтерпретація опрацьованих даних, встановлення норм тесту і створення шкали для оцінювання результатів іспитників.

Л. О. Кухар та В. П. Сергієнко характеризують якість тесту як окрему категорію, що базується на певних теоріях і закономірностях [104]. Вони зазначають, що визначення якості тесту базується на певних теоретико-методологічних засадах та методах, що використовуються під час аналізу результатів тестування. Методи, що використовуються під час аналізу, поділяються на аналітичні (експертні) та емпіричні (математико-статистичні). Є. В. Драган визначає якість навчального тесту як величину, що описує сукупність взаємопов'язаних параметрів навчального тесту, які характеризують його властивості: валідність, надійність, об'єктивність, ефективність [38].

Багато авторів не дають чіткого визначення поняття якості тесту, а розглядають лише критерії цієї якості. Так В. С. Аванесов зазначає, що серед

критеріїв якості можна виділити: відповідність мети розроблення тесту тій чи іншій освітній політиці, мети розроблення тесту – контингенту іспитників, відбору змісту кожного завдання – цілям тесту в цілому, рівня складності завдань – рівню підготовленості кожного іспитника. Головними критеріями якості тестів є точність результатів вимірювання і адекватність інтерпретації результатів відповідно до цілей тестування [3].

Аналіз наукової педагогічної і психологічної літератури дає змогу виділити дві основні теорії, які визначають критерії якості тесту: Item Response Theory та класичну теорію тестів. Необхідно наголосити на високій зацікавленості науковців в досягненнях обох теорій та методах розроблення високоякісних педагогічних тестів.

В основу класичної теорії тестів покладено розвинутий класичний статистичний апарат для дослідження результатів вимірювань. Засновником класичної теорії тестів є відомий британський психолог, автор факторного аналізу Чальз Едвард Спірмен (Charles Edward Spearman) (1863-1945 pp). Послідовниками Спірмена були А. Анастасі (A. Anastasi), Дж. П. Гілфорд (J.P. Guilford), П. Вернон (P. Vernon), С. Берт (C. Burt), А. Дженсен (A. Jensen), Л. Гуттман (L. Guttman) та ін.

Сучасна теорія тестів Item Response Theory призначена для оцінювання латентних параметрів іспитників та параметрів завдань тесту за допомогою застосування математико-статистичних моделей вимірювання [161, 191]. Засновником IRT є датський математик Георг Раш [198].

Назва Item Response Theory перекладається по-різному. У вітчизняній та російськомовній літературі поширена назва «сучасна теорія тестів» [80]. Ю. М. Нейман і В. А. Хлебніков пропонують називати її «Теорія моделювання та параметризації педагогічних тестів» [124]. В. С. Аванесов – «Математико-статистична теорія оцінювання латентних параметрів завдань тесту і рівня підготовленості іспитників» [5]. Оскільки у вітчизняній літературі загальноприйнятої назви немає, будемо називати її без перекладу – IRT.

Методи IRT зводяться до отримання оцінок параметрів складності завдання рівня підготовленості іспитників. До однієї з найзначніших переваг IRT можна віднести використання однієї шкали при вимірюваннях значень параметрів студентів і завдань тестів. Саме це дає змогу зіставляти рівні навчальних досягнень студентів з рівнями складності тестових завдань. На цій властивості сумісних оцінок параметрів студентів і завдань базується математичний апарат, що використовується більшістю тестологів [156].

Методи КТТ та IRT дають змогу оцінити не лише якість тесту у цілому, а також якість окремих тестових завдань. А оскільки *тестове завдання* є елементарною одиницею тесту, то, очевидно, що якість тесту залежить від якості кожного відібраного для нього завдання.

Створення тестового завдання – це складний процес, що включає створення деякої заготовки тестового завдання (предтестове завдання або завдання в тестовій формі) і власне тестового завдання. Наведемо визначення цих типів завдань.

М. Б. Челишкова визначає предтестове завдання як одиницю контрольного матеріалу, зміст, логічна структура і форма подання якого задовольняють ряду вимог і забезпечують однозначність оцінок результатів іспитників [162]. Вимоги, що пред'являються до передтестових завдань, можна умовно розділити на спеціальні, що відображають специфіку форми завдань, і загальні, інваріантні щодо обраної форми.

В. С. Аванесов завдання в тестовій формі визначає як педагогічний засіб, що відповідає таким загальним вимогам [4, 7]:

1. *Мета*. Розробник повинен чітко усвідомлювати, яка мета пред'явлення тестового завдання іспитнику. Мета створення тестового завдання залежить, зокрема, від типу тесту (критеріальною-орієнтований або нормативно-орієнтований), від рівня навчальних досягнень, від виду тестування (підсумкове, поточний, для самоконтролю) тощо.

Одним із важливих чинників, що визначає мету завдання, є отримання відомостей про рівень засвоєння іспитником тієї чи іншої одиниці

навчального матеріалу. Формулюванню мети завдань сприяє створення технологічної матриці, що представляє собою перелік навчальних тем з відносним розподілом тестових завдань по цих темах [109].

Також важливим є визначення типу вміння, що визначає тестове завдання. Так за таксономією Б. С. Блума виділяються такі рівні засвоєння навчального матеріалу: запам'ятовування, розуміння, застосування, аналіз, синтез, оцінювання [170]. В. П. Беспалько [16], П. В. Симонов [142], М. В. Кларін [93] пропонують інші таксономії. Завдання з однієї і тієї ж навчальної одиниці, що орієнтовані на різні рівні засвоєння матеріалу, мають різні цілі і тому матимуть різний зміст.

2. *Стислість*. Формулювання завдання повинно бути максимально лаконічним, але зрозумілим. Необхідно виключити повторювання одних і тих самих слів чи фраз. Лаконічне завдання краще сприймається іспитником.

3. *Технологічність*. Ця вимога є важливою при комп'ютерному тестуванні. Технологічність передбачає використання при створенні тестових завдань малюнків, формул, креслень. Для тестів з вищої математики вимога технологічності є особливо важливою.

4. *Логічна форма висловлювання*. За В. С. Аванесовим, це засіб упорядкування і ефективної організації змісту тесту. Ця форма зазвичай замінює питання. Перевага завдання в тестовій формі полягає в можливості після відповіді іспитника перетворення твердження в форму істинного або хибного висловлювання [7]. Такої ж думки дотримується, наприклад, і М. Б. Челишкова. А. А. Н. Майоров вважає, що правильно сформульоване завдання в питальній формі не поступається завданню у формі твердження. За В. Ю. Переверзевим [129] основа тестового завдання також може формулюватися в питальній формі. Це дає змогу завданням відповідати більш високим пізнавальним рівням таксономії.

5. *Визначеність місця для відповіді*. Важливою вимогою до завдання, є «економність» при його оформленні [7]. На бланку або на моніторі комп'ютера не повинно бути нічого зайвого. Форма завдання повинна

максимально сприяти розумінню його змісту. Визначеність місця для відповіді є одним з прийомів такої мінімізації. Якщо місце для відповіді змінюватиметься від завдання до завдання, це ускладнить роботу іспитника, що може призвести до зниження результату тестування з причин, не пов'язаних з рівнем його навчальних досягнень.

6. *Однаковість правил оцінювання відповідей.* Однаковість правил оцінювання відповідей полягає в тому, що всі іспитники повинні знаходитись в рівних умовах при оцінюванні правильності відповіді. Це необхідна умова використання тесту в якості засобу оцінювання знань.

7. *Правильність розташування елементів завдання.* Ця вимога полягає в строгому дотриманні розміщення елементів завдання згідно його структури.

8. *Однакові інструкції для усіх студентів.* Це очевидне вимога, в іншому випадку з'явиться велика систематична похибка вимірювання.

9. *Адекватність інструкції формі і змісту завдання.* Дана вимога означає взаємну відповідність компонентів. Невідповідність форми змісту і, навпаки, змісту формі, призводить до неправильного розуміння завдання [7].

Тест складається не з завдань у тестовій формі, не з питань і завдань, а лише з тестових завдань [7]. Передтестове завдання може називатись тестовим лише тільки після статистичної перевірки, якщо кількісні оцінки його характеристик задовольняють певним критеріям якості змісту та форми. Змістова відповідність завдань тесту визначається експертно. Головною умовою при цьому є його відповідність специфікації та змісту дисципліни. На експертному рівні завдання в тестовій формі перевіряються з точки зору правильності їх змісту, а також правильності використаних тестових форм. На практиці експертиза дуже часто пропускається, і в результаті для тестування застосовуються неякісні тестові матеріали.

М. Б. Челишкова дає таке визначення тестового завдання: завдання називається тестовим, якщо апостеріорні кількісні оцінки його характеристик задовольняють певним критеріям, націленим на перевірку якості змісту,

форми і на виявлення системоутворюючих властивостей завдань тесту [162]. Відзначимо також, що тестові завдання повинні задовольняти умові локальної незалежності [182]. В. Ю. Переверзев зазначає, що відповідь іспитника на кожне тестове завдання є статистично незалежною від відповіді на будь-яке інше тестове завдання [129]. Локальна незалежність передбачає, що іспитник, відповідаючи на завдання, не може використовувати знання, отримане з відповіді на будь-яке інше тестове завдання. З локальної незалежності випливає, що при комп'ютерному тестуванні не варто інформувати іспитника про успішність виконання чергового завдання, оскільки це може допомогти йому відповісти на інше завдання, тобто умова локальної незалежності завдань буде порушено.

Підсумовуючи вищезазначене, у дослідженні під **тестовим завданням** будемо розуміти складову одиницю тесту, яка відповідає таким вимогам: зміст відповідає меті тестування; сформульовано максимально лаконічно, але зрозуміло; відповідає умові локальної незалежності; однакова для всіх іспитників інструкція і структура завдання; чітко визначене місце для фіксації відповідей; однаковість принципів оцінювання та інтерпретації результатів; кількісні характеристики задовольняють певним критеріям якості.

Якість тесту у дослідженні будемо розуміти як відповідність вимогам, що висуваються до характеристик тесту у цілому та до окремих тестових завдань. Базовими характеристиками якості тесту є надійність, валідність, ефективність тесту, набір латентних параметрів та структура ансамблів характеристичних кривих. Наведемо визначення основних понять.

Челишкова М.Б. надійністю називає характеристику тесту, що відображає точність тестових вимірювань, а також стійкість тестових результатів до дії випадкових факторів [162]. Вона зазначає, що термін «надійність» має два значення. По-перше, тест вважається надійним, якщо він забезпечує високу точність вимірювань. І по-друге, тест вважається

надійним, якщо він дає при повторному тестуванні близькі результати за умови, що підготовка іспитника не змінилася до повторного виконання тесту.

За П. Клайном термін «надійність» у психометрії також має два значення [92]. Тест називається надійним, якщо він є внутрішньо узгодженим. Тест також називається надійним, якщо він дає одні й ті ж показники для кожного іспитника при повторному тестуванні. Надійність при повторному тестуванні через певний час називають також ретестовою надійністю.

У дослідженні під *надійністю тесту* будемо розуміти характеристику тесту, що відображає стійкість тестових результатів до дії випадкових факторів. Тест можна вважати надійним, якщо він відтворює результати при повторному тестуванні тієї ж вибірки іспитників.

Термін «валідність» походить від англійського «valid» і означає обґрунтований, дійсний, придатний, той, що має силу. А. Анастасі означає валідність як характеристику тесту, яка визначає, що вимірює тест і наскільки якісно це здійснюється [10].

Поняття валідності використовується для загальної характеристики тесту в аспекті відповідності одержаних результатів меті та вимогам оцінювання. Воно відповідає на питання про те, що саме виявляє тест, наскільки він придатний для виявлення того, для чого він призначений [34]. Невалідний тест – такий, у якому одержані результати не вимірюють того, що було визначено метою тестування [104].

Аванесов В. С. зазначає, що валідність – це методологічна характеристика здатності тесту вимірювати те, для чого його було створено [6]. Вона залежить від якості та кількості завдань; від міри повноти і глибини охоплення змісту навчальної дисципліни; від балансу і розподілу завдань за складністю; від методу відбору завдань із тестового банку; від інтерпретації результатів тестування; від організації збирання даних, від властивостей вибіркової сукупності іспитників.

У дослідженні під **валідністю тесту** будемо розуміти характеристику тесту, що відображає відповідність результатів тестування тій меті, заради якої тест було створено.

Поняття надійності та валідності педагогічного тесту надзвичайно важливі, оскільки саме вони характеризують тест як вимірювальний засіб. Тест з невідомими надійністю і валідністю непридатний для оцінювання знань студентів.

За В. С Аванесовим ефективність – це порівняльна характеристика, яка дає змогу назвати тест, який краще, ніж інші тести, вимірює знання студентів потрібного рівня підготовки з меншою кількістю завдань, якісніше, швидше, дешевше, і все це, якщо можливо, одночасно [2].

У дослідженні під **ефективністю тесту** будемо розуміти порівняльну характеристику, що дає змогу назвати тест, який точніше та з меншими витратами, ніж інші тести, вимірює знання студентів певного рівня підготовки.

Складність тестових завдань може бути визначена по-різному в межах різних теорій тестування. Вона відображає рівень можливості правильного виконання завдання в даній статистичній вибірці. Так в межах КТТ складність тестового завдання – це частка правильних відповідей на завдання групою іспитників. У рамках IRT складність є латентним параметром завдання. Одне завдання вважається складнішим за інше, якщо ймовірність правильної відповіді на перше завдання менша, ніж на друге, незалежно від того, хто їх виконує. Слід зазначити, що для різних IRT-моделей завдання характеризують різні набори латентних параметрів. Детально дане питання буде розглянуто в п. 2.2.2.

В. Ю. Биков, Ю. М. Богачков, Ю. О. Жук в роботі [20] зазначають, що для того, щоб визначити поняття складності тесту, треба встановити деяку шкалу та точку відліку, відносно якої будемо оцінювати цей параметр. Згідно з теорією тестів, як правило, складність тесту має відповідати групі

тестування, тобто, щоб середній студент давав правильні відповіді приблизно на половину запитань. Але ця умова залежить від мети тестування.

У дослідженні під *складністю тестового завдання* будемо розуміти латентний параметр завдання в рамках IRT. Латентними називають «приховані» параметри, що визначаються за результатами тестування. *Набір латентних параметрів* включає параметри тестового завдання, зв'язок між якими задається ймовірністю правильної відповіді на завдання іспитника з певним рівнем підготовленості. Різним IRT-моделям відповідають різні набори латентних параметрів, що детально будуть розглянуті в пункті 2.2.2. Графіком залежності ймовірності правильної відповіді іспитника на завдання від параметра підготовленості при заданому наборі латентних параметрів завдання називають *характеристичною кривою тестового завдання*.

Поняття якості тесту є нерозривним з поняттям *оцінювання якості тесту*. Під оцінюванням розуміють складання уявлення, висновку про що-небудь, визначення суті, характеру, значення, роль і т. ін. чогось [143].

У дослідженні під *оцінюванням якості тесту* будемо розуміти процедуру встановлення ступеня відповідності характеристик окремих тестових завдань та тесту у цілому критеріям якості та складання висновку про якість тесту.

Розробник тестів також повинен чітко розуміти, що на результат оцінювання якості впливає ряд факторів, пов'язаних безпосередньо із *індивідуальними особливостями іспитника*. До таких факторів можна віднести: психологічний стан іспитника під час тестування; мотивацію іспитника до вдалого складання тесту; його підготовку до тестування; можливість вгадування іспитником відповіді; час, необхідний конкретному іспитнику для вдалого складання тесту; умови проведення тестування тощо.

Викладач повинен вжити всіх засобів для мінімізації впливів цих факторів як під час розроблення тесту, так під час підготовки студентів до тестування та безпосередньо під час тестування. Оптимальний час виконання тесту іспитниками визначається при оцінюванні змістової валідності (п. 3.3).

Визначення оптимального часу тестування є важливим з декількох причин. По-перше, збільшення часу тестування може призвести до стомлення іспитників, що в свою чергу вплине на результати тестування. Тому час від початку тестування до моменту стомлення іспитників можна вважати оптимальним часом для виконання тесту. Стан втоми виникає швидше під час інтенсивної або монотонної роботи. Тому тест повинен містити завдання, різноманітні за формою та за видами діяльності. По друге, неправильно визначений час виконання тесту призводить до двох негативних наслідків: при збільшенні часу до кінця тестування знижується дисципліна, тому що іспитники, що впоралися з тестом, мають можливість "допомогти" своїм друзям; зменшення часу тестування провокує іспитників на вгадування правильних відповідей. Оптимальний час виконання тесту може бути скоректовано після аналізу поведінки студентів на пробному тестуванні. Питання вибору оптимального часу тестування досліджено в роботі [90].

До початку тестування бажано дати студентам можливість пройти тренувальне тестування, особливо якщо вони ніколи раніше не виконували тести. Психологами встановлено, що попередня підготовка, ознайомлення з приблизним змістом тесту, знижує хвилювання іспитників під час тестування і мотивує їх до виконання тесту [80]. Тренування перед тестуванням найбільш корисними є для студентів з недостатнім запасом знань, оскільки змушує їх повірити в свої сили і мобілізує до виконання тесту. Фахівці з педагогічних вимірювань також вважають за необхідне проводити тренування перед тестуванням. Тренування мінімізує втрати часу при роботі іспитників з різними типами інтерфейсу тестових завдань у тестовій системі без звертання до екзаменатора.

Важливою є і правильна підготовка іспитників до тестування. На основі кривої забування, або кривої Еббінгауза, було розроблено методику ефективного вивчення [28]. Крива Еббінгауза – це графік, який зображує процеси забування людиною певного роду інформації. Відповідно до розробленої методики, вивчення матеріалу не повинно відбуватися за один

раз, необхідне повторювання матеріалу. Будь-який матеріал можна вивчити за два дні, якщо перше повторення зробити відразу ж після першого засвоєння матеріалу, друге – через 20 хвилин, третє – через вісім годин, а четверте – через день після останнього повторення. Така підготовка до тестування є дуже ефективною, її застосування варто радити студентам.

Вгадування іспитником відповідей завдань, крім обмеження часом, можна уникнути різноманітним форм тестових завдань. Зокрема, тест повинен включати завдання відкритої форми. Завдання на відповідність повинні складатися таким чином, щоб елементів другої множини відповіді було більше, ніж першої. Це дасть змогу уникнути автоматичного визначення відповідності останньої пари. Також рекомендуємо нараховувати штрафні бали за встановлення неправильної пари, що мінімізує встановлення відповідності іспитниками намання. Нарахування штрафних балів також є ефективним у завданнях множинного вибору з декількома правильними відповідями. Також ефективним психологічним прийомом проти вгадування є відповідний пункт в інструкції до тестування із зазначенням того, що обирати відповідь намання в разі сумніву або незнання не можна.

Забезпечення зручних умов для іспитників вимагає від викладача завчасного підбору місця проведення тестування. В аудиторії повинно бути все необхідне технічне забезпечення, з гарним освітленням та зручними робочими місцями. Розміщення робочих місць повинно виключати можливість списування іспитників один у одного.

Також велике значення має розроблення правильних інструкцій для іспитників щодо виконання тесту. Інструкція до тесту це керівна вказівка, у якій описано, як виконувати завдання тесту [104, ст. 175]. Вона має бути простою та стислою, але водночас містити такі відомості:

- 1) мета тесту;
- 2) час, відведений для виконання тесту;
- 3) пораду не вдаватися до вгадування, якщо є сумніви щодо правильної відповіді;

- 4) порядок виконання завдань тесту;
- 5) оцінювання виконання завдань тесту.

Перед початком тестування викладач повинен:

- пояснити, навіщо потрібен тест, повідомити, які результати очікуються;
- повільно, чітко, природним голосом прочитати інструкцію до тесту, що була роздана іспитникам;
- повідомити про правила виправлення допущених помилок, розповісти про те, чого не рекомендується робити при виконанні завдань і до кого звертатися в разі виникнення питань;
- відповісти на наявні питання до початку виконання тесту.

Під час тестування викладачу необхідно уважно слідкувати за іспитниками, щоб уникнути випадки шахрайства. За 15 хвилин до закінчення часу тестування необхідно зробити відповідне повідомлення.

Перейдемо до понять «метод» та «методика». Термін «метод» походить від грецького *metodos*, що означає «шлях дослідження, теорія, вчення». Метод – це засіб досягнення будь-якої мети, вирішення конкретного завдання; сукупність прийомів або операцій практичного чи теоретичного засвоєння (пізнання) дійсності [166]. Гончаренко С. У. визначає метод як спосіб організації практичного й теоретичного освоєння дійсності, зумовлений закономірностями розглядуваного об'єкта [34].

Терміном «методика» зазвичай позначають конкретні принципи, форми та засоби використання методів, за допомогою яких здійснюється більш глибоке пізнання різноманітних педагогічних проблем та їх розв'язання. [103].

У тлумачних словниках можна зустріти такі визначення методики: система правил, виклад методів виконання якоїсь роботи [149]; сукупність методів практичного виконання чого-небудь [148].

В. І. Воронов, В. П. Сідоров визначають методику як сукупність прийомів, методів дослідження, порядок їх застосування та інтерпретацію

отриманих з її допомогою результатів. Вона залежить від характеру об'єкта вивчення; мети дослідження; розроблених методів; загального рівня кваліфікації дослідника [31]

Методика оцінювання якості тестів передбачає застосування складних математичних методів і вимагає використання ІКТ. Використання ІКТ дає змогу суттєво збільшити обсяг даних, які мають бути накопичені та опрацьовані для оцінювання та покращення якості тестів.

Поняття інформаційно-комунікаційних технологій визначають з різних боків. Так М. І. Жалдак визначає ІКТ як сукупність методів, засобів та прийомів, що використовуються для збирання, систематизації, зберігання, опрацювання, передавання, подання різних повідомлень і даних [72]. Він акцентує увагу на необхідності збереження, вдосконалення та посилення кращих досягнень педагогічної науки минулого, в т. ч. за рахунок використання останніх досягнень комп'ютерної техніки.

Н. В. Морзе уточнює дане визначення [121]: «ІКТ – це сукупність методів, засобів і прийомів, що використовується для реалізації конкретного складного процесу, шляхом поділу його на систему послідовних взаємопов'язаних процедур і операцій, що виконуються більш або менш однозначно і мають на меті досягнення результативності щодо пошуку, накопичення, опрацювання, зберігання, подання, передавання відомостей та даних за допомогою засобів обчислювальної техніки та зв'язку, а також засобів їх раціонального поєднання з процесами опрацювання даних без використання машин».

О. М. Спірін наголошує, що ІКТ в освіті є не тільки сукупністю методів, засобів і прийомів, що використовуються для розроблення інформатичних систем та побудови освітніх комунікаційних мереж, а також технологіями формалізації та розв'язування задач у певних предметних галузях з використанням таких систем і мереж [145].

Ю. О. Жук зазначає, що впровадження засобів ІКТ в освітній процес, потребує розроблення спеціальних засобів, що надають можливість

використання засобів ІКТ у навчально-виховному процесі відповідно до змісту педагогічної ситуації, які отримали назву «комп'ютерно орієнтовані засоби навчальної діяльності». Такі засоби надають користувачу певний набір послуг, використання яких розширює спектр навчальної діяльності та змінює структуру навчального середовища [73].

Визначення ІКТ Ю. І. Триуса включає комп'ютер як засіб управління навчально-пізнавальною діяльністю, сукупність комп'ютерно орієнтованих методів, засобів та організаційних форм навчання [151].

І. В. Роберт визначає ІКТ як програмні, програмно-апаратні та технічні засоби і пристрої, що функціонують на базі засобів мікропроцесорної обчислювальної техніки, а також сучасних засобів і систем трансляції інформації, інформаційного обміну, що забезпечують операції щодо збирання, накопичення, опрацювання, зберігання, продукування, передачі, використання інформації, можливість доступу до інформаційних ресурсів комп'ютерних мереж, в тому числі і глобальних [136].

В. А. Трайнев [150] вважає, що до складу ІКТ потрібно включати сукупність методів та програмно-технічних засобів, які об'єднуються в технологічний ланцюг, що забезпечує збирання, опрацювання, збереження та відображення відомостей та даних з метою зниження трудомісткості їх використання, а також для підвищення їх надійності й оперативності.

Проаналізувавши різні тлумачення поняття ІКТ, їх можна об'єднати в такі групи:

- засоби, що забезпечують роботу з джерелами відомостей та даних, наприклад, персональний комп'ютер, смартфон, мультимедійна дошка та ін., програмні засоби; Web технології та ін.;

- джерела відомостей та даних, під якими розуміються матеріали у цифровому форматі (тексти, зображення та ін.), наприклад, електронні бібліотеки, підручники, посібники, енциклопедії та ін.; бази даних; ресурси дистанційних курсів та ін.;

– сукупність методів і прийомів, що використовуються для збирання, систематизації, зберігання, опрацювання, передавання даних у різних форматах.

У дослідженні будемо використовувати поняття ІКТ, тлумачення якого відповідає першій групі.

У дослідженні під **комп'ютерно орієнтованою методикою оцінювання якості тесту** розумітимемо теоретично обґрунтовану сукупність змісту, методів та засобів оцінювання якості тесту з використанням ІКТ.

Не зважаючи на велику кількість робіт, присвячених питанню аналізу якості тестів, актуальним залишається питання спільного застосування методів КТТ та ІРТ при опрацюванні результатів тестування, поєднання їх в єдину комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики, а також впровадження такої методики.

1.2. Тестування як засіб контролю знань студентів

Актуальною педагогічною проблемою є проблема покращення контролю знань студентів. Контроль знань є одним з найважливіших елементів навчального процесу. Його основне призначення – встановити «зворотній зв'язок» для оцінювання динаміки засвоєння навчального матеріалу, аналізу рівня володіння системою знань, умінь і навичок та внесення відповідних коректив в організацію навчального процесу. Питанню вдосконалення контролю знань присвячені праці В.С. Аванесова [2], С.І. Архангельського [12], В.П. Беспалька [16], В. Ю. Бикова, Ю. М. Богачкова, Ю. О. Жука [20], Н. Ф. Тализіної [147] та інших дослідників. На сьогоднішній день до числа найбільш визнаних методів контролю знань у світовій практиці відноситься тестування. Починаючи з 2008 року, завдяки впровадженню в Україні ЗНО, використання тестів опинилося у фокусі освітянського інтересу нашої країни. Тестування все ширше застосовується у вищій та середній школах при підсумковому контролі, тестування стає більш розповсюдженим. Деякі освітні установи видають сертифікати на основі

даних комп'ютерного тестування. Також зростає роль дистанційного навчання, в якому в основному тестування застосовується як форма контролю знань студентів.

Необхідно зазначити, що сучасний етап розвитку суспільства характеризується вираженою інформатизацією усіх сфер соціального життя. Побудова інформаційного суспільства є одним із стратегічних напрямів політики нашої держави, визначеним Законами України «Про національну програму інформатизації» [75], «Про пріоритетні напрями інноваційної діяльності в Україні» [79] та ін. Рівень інформатизації суспільства безпосередньо пов'язаний із рівнем інформатизації освіти. На сьогоднішній день світова практика розвитку та використання ІКТ вказує на необхідність змін традиційних форм організації освітнього процесу. Напрями перебудови та вдосконалення процесу навчання із впровадженням інформаційно-комунікаційних технологій пропонують у своїх дослідженнях В. Ю. Биков [18, 19], Ю. І. Жук [73], А. О. Білощицький [22].

Зокрема спостерігаються зміни в процесі контролю знань студентів. І у зв'язку із реформуванням вітчизняної освітньої галузі і впровадженням ІКТ особливого значення набуває комп'ютерне тестування. Застосування комп'ютерного тестування як засобу контролю знань вимагає ґрунтовної підготовки з боку викладачів для складання надійних педагогічних тестів. Насамперед необхідні глибоке оволодіння строгими правилами й математичним апаратом в області педагогічних вимірювань, формування компетентності щодо оцінювання якості тестів та забезпечення інформаційно-комунікаційною підтримкою. Неправильно складені тести можуть давати помилкові результати.

Вітчизняні тестологи І. Є. Булах [26], О.Г. Виноградова [30], М. М. Олійник [125] стверджують, що використання тестів в українській освіті дає позитивні результати. Тестування є швидким та ефективним засобом контролю та оцінювання навчальних досягнень, спрямоване на підвищення якості освіти, що дає змогу регулярно проводити моніторинг

рівня навчальних досягнень студентів і своєчасно забезпечувати необхідну взаємодію між студентами та викладачами.

Основними перевагами комп'ютерного оцінювання навчальних досягнень студентів є:

- автоматизація та оперативність опрацювання результатів тестування;
- прискорення зворотного зв'язку за результатами тестування, що дає змогу студенту самостійно виявляти прогалини в структурі своїх знаннях і вживати заходів до їх ліквідації;
- забезпечення об'єктивності оцінювання;
- можливість регулярно поповнювати та модифікувати систему тестових завдань;
- звільнення викладача від виконання рутинних робіт.

Можна виділити ряд істотних недоліків тестування. Наприклад, жорсткі часові обмеження тестування виключають можливості визначити рівень підготовленості тих іспитників, які в силу своїх психофізіологічних особливостей думають і виконують завдання повільно, але правильно. Аналіз способів розв'язання завдань і обчислювальних операцій іспитників у більшості випадках є складним або взагалі неможливим. Крім того, при тестуванні не задіяний мовний апарат, що унеможливорює дослідження логіки міркування іспитника. Іншим недоліком є ситуації вибору відповідей навмання або методом виключення. Однак, не зважаючи на вищезазначені обмеження діагностування рівня сформованості знань, виїнь та навичок студентів, тестування є одним з найнадійніших та найоб'єктивніших засобів контролю. Визначальною умовою при цьому є використання лише правильно складених завдань та тесту у цілому.

Вищезазначені чинники вказують на те, що тестування не слід розглядати як ідеальний і єдиний метод об'єктивного контролю знань. Тестування не заміняє й не скасовує традиційних форм педагогічного контролю, заснованих на безпосередньому спілкуванні викладача зі студентом. Для підвищення ефективності навчального процесу, на нашу

думку, тестування повинно не підміняти, а доповнювати традиційні, сформовані форми контролю знань. Традиційний контроль, окрім надання відомостей про рівень знань студентів та прогалини в їхній підготовці, допомагає дослідити емоційно-психологічний стан студентської групи. З іншого боку, традиційні форми педагогічного контролю носять багато в чому суб'єктивний характер і не дають змогу одержати порівняльні дані, настільки необхідні для керування процесом освіти в масштабах навчального закладу, району, області, країни. А завдання такого плану все більше набувають своєї актуальності.

Питання використання тестів як засобу педагогічного контролю досліджували в своїх роботах В. С. Аванесов [2], В. П. Беспалько [14], О. М. Майоров [109], Л. І. Петрова [130], та ін.

Обмежене застосування тестування з вищої математики у ЗВО має місце у зв'язку з тим, що викладачі сприймають тести лише як певні психодіагностичні методики, більше придатні до гуманітарних наук. Це зумовлено недостатнім упровадженням математичної теорії, яка може забезпечити розроблення надійних та валідних тестів [54].

Також спостерігається помилкове стереотипне уявлення викладачів про нібито елементарність процесу оцінювання знань студентів. Прихильність традиційних методів контролю, заперечення інновацій та недовіра до тестування та його об'єктивності призводять до відмови викладачів від тестування, як засобу контролю знань.

Т. О. Лукіна зазначає, що «Тестування, як практика застосування тестів, може розвиватися, спираючись на теорію створення тесту, а може існувати відокремлено від неї, лише імітуючи наявність наукового підґрунтя. На жаль, в Україні в цілому характерна друга тенденція, коли на ринку педагогічних матеріалів та в практичній діяльності з'явилася величезна кількість псевдотестів та прикладів псевдотестування навчальних досягнень» [108, с. 34]. Це пов'язано з низьким рівнем компетентності викладачів щодо розроблення та оцінювання якості тестів. Псевдотести наносять значну

шкоду як іспитнику, так і викладачу. У першому випадку, іспитник отримує неправильні результати оцінювання його навчальних досягнень, в другому – некомпетентний у створенні тестів викладач зазнає не тільки морального, але і професійного краху. Звідси, у суспільстві формується хибне уявлення про тестування як методу оцінювання знань та зневіра в його можливостях [108]. Для застосування тестів як засобу контролю знань необхідною умовою є використання лише тих тестів, надійність, валідність та ефективність яких була підтверджена. При цьому викладачі, які розроблятимуть тести, повинні пройти відповідне навчання з метою розвитку компетентності у галузі тестології.

Тестологія – наука про тести, їх створення й використання (В. С. Аванесов, І. Є. Булах, Т. Н. Тягунова). У галузі педагогічних вимірювань – це теоретико-методологічне й методичне обґрунтування процесів розроблення та застосування педагогічних тестів. Завданням тестології є розроблення тестів для об'єктивного контролю рівня навчальних досягнень студентів. Педагогічна тестологія, як зазначають В. С. Аванесов, М. М. Олійник та Ю. А. Романенко, – це прикладна методична теорія наукової педагогіки. Ключовими поняттями тестології є вимірювання, тест, зміст і форма завдань, надійність і валідність результатів вимірювання [3]. Також в тестології використовують такі поняття статистики як вибірка й генеральна сукупність, розподіл вибірки, середні показники, дисперсія, кореляція та ін. [185]. У тестології є як педагогічна складова, так і математична й психологічна, а у зв'язку з комп'ютерною реалізацією – ще й технологічна.

Центральним поняттям тестології є поняття *тесту*, яке детально розглянуто в пункті 1.1. Структурним компонентом педагогічного тесту є: тестове завдання. Поняття завдання є базовим поняттям теорії тестів [172]. В тестології завдання можна визначити як засіб інтелектуального розвитку, освіти, що сприяє активізації навчання, підвищенню рівня навчальних досягнень студентів, а також ефективності педагогічної праці.

Завдання в тесті розглядається як одиниця контрольного матеріалу, яка оцінюється за параметрами складності, диференціюючої спроможності, ефективності, відповідності змісту дисципліни, часу виконання тощо. Завдання можуть бути використані як для самоперевірки, так і для визначення рейтингу студентів, і для вимірювання рівня сформованості основних компетентностей студентів.

Існують різні способи класифікації **форм тестових завдань**. Автори одностайні в тому, що виділяють дві основні групи завдань: відкритої та закритої форми. Найчастіше та, зокрема, в роботах [75] та [96] ці групи об'єднують шість форм завдань. Вважаємо доцільним виділити в цих групах сім форм завдань, представлених схематично на Рис. 1.1., відокремивши завдання з пропусками та завдання з короткою відповіддю із завдань на доповнення.



Рис. 1.1. Основні форми тестових завдань

Тестові завдання **відкритої форми** не мають варіантів відповіді. Ці завдання розділимо на:

- завдання з пропусками;
- завдання вільного викладу;
- завдання з короткою відповіддю.

Задання з пропусками. В такому завданні іспитник повинен його зміст доповнити одним або кількома словами (числами, буквами, символами, словосполученнями). В результаті завдання повинно перетворитися в істинне твердження.

Приклад.

Доповніть твердження.

Оберненою називається матриця, при множенні на яку вихідна матриця перетворюється на _____ матрицю. (Відповідь: одиничну)

Задання з пропусками у тестах з вищої математики найбільш придатні для оцінювання знань студентів теоретичного матеріалу. Зазвичай вони широко використовуються в лінгвістичних тестах.

Завдання вільного викладу. В такому завданні іспитник повинен самостійно сформулювати розгорнуту відповідь. Ніякі обмеження на нього в завданні не накладається.

Приклад.

1) Дайте відповідь на запитання.

У чому полягає дія транспонування матриць?

(Відповідь: Дія транспонування матриці полягає у переведенні рядка у стовпець, і навпаки.)

Завдання з короткою відповіддю. У завданні цієї форми іспитник повинен з клавіатури відповісти на поставлене питання, ввівши необхідне слово, словосполучення або число. Завдання може мати декілька правильних відповідей з різним значенням оцінок.

Приклад.

1) Дайте відповідь на запитання.

Які матриці можна множити?

(Відповідь: узгоджені)

2) Знайти похідну y'_x в точці $(1; 0)$, якщо $e^y + xy = x^2$.

(Відповідь: 1)

Тестові завдання **закритої форми** передбачають наявність варіантів

відповіді. Вони поділяються на:

- завдання з множинним вибором (з однією правильною відповіддю, з декількома правильними відповідями);
- завдання на відповідність;
- завдання альтернативних відповідей;
- завдання на встановлення правильної послідовності.

Завдання з множинним вибором. Ця форма завдання вимагає від іспитника обрати одну або декілька відповідей із запропонованого набору варіантів. Кожен варіант відповіді може мати свою вагу оцінювання. Загальна оцінка за завдання складається з оцінок обраних варіантів.

Наведемо приклад завдання множинного вибору з однією правильною відповіддю.

Приклад.

Виберіть одну відповідь.

Якщо функція $y = f(x)$ має похідну в точці x_1 , то обернена функція $x = \varphi(y)$ має похідну у відповідній точці y_1 , яку знаходять за формулою:

a) $\varphi'(y) = \frac{1}{f'(x)}$;

b) $\varphi'(y) = f'(x)$;

c) $\varphi'(y) = -\frac{1}{f(x)}$;

d) $\varphi'(y) = -f'(x)$.

(Відповідь: а)

Наведемо приклад завдання множинного вибору з двома правильними відповідями.

Приклад.

Виберіть принаймні одну відповідь.

Знайдіть асимптоти графіка функції $f(x) = \frac{5x^2 - 1}{x + 2}$.

- a) $y = 10x + 5$;
- b) $x = -2$;
- c) не існує похилої асимптоти;
- d) $y = 5x - 10$.

(Відповідь: b, d)

Для завдань множинного вибору радимо для дистракторів встановлювати від'ємну вагу оцінювання.

Завдання на відповідність. У таких завданнях необхідно встановити відповідність між елементами двох множин. Елементами першої множини, наприклад, можуть бути питання на задану тему, другої – відповіді на ці питання. При тестуванні для кожного питання треба обрати із переліку відповідну правильну відповідь. Кількість елементів другої множини повинна бути більшою, ніж першої. Наведемо приклад тестового завдання на відповідність.

Приклад.

Побудуйте частинний розв'язок диференціального рівняння

$$y''' - 2y'' + y' = f(x), \text{ якщо}$$

- 1. $f(x) = xe^{3x}$;
- 2. $f(x) = 2e^x$;
- 3. $f(x) = 5x$;
- 4. $f(x) = e^{-x} + 1$.

Варіанти:

- a) $y_{\text{ч.н.}} = (ax + b)e^{3x}$;
- б) $y_{\text{ч.н.}} = ax^2e^x$;
- в) $y_{\text{ч.н.}} = ax^2 + bx$;
- г) $y_{\text{ч.н.}} = ae^{-x} + bx$;
- д) $y_{\text{ч.н.}} = ax^2e^{-x} + bx$.

(Відповідь: 1.а, 2.б, 3.в, 4.г.)

Завдання альтернативних відповідей. У таких завданнях необхідно дати однозначну відповідь (правильно – не правильно) на поставлене питання. Наведемо приклад такого тестового завдання.

Приклад.

Чи правильно, що $(\sin x)' = 1$?

- a) Правильно;
- b) Не правильно.

(Відповідь: b.)

Завдання альтернативних відповідей не радимо використовувати для тестів з вищої математики через високу ймовірність вгадування правильної відповіді..

Завдання на встановлення правильної послідовності. У таких завданнях необхідно розмістити елементи списку у правильній послідовності. Наведемо приклад такого тестового завдання.

Приклад.

Встановіть правильну послідовність кроків перетворення матриці до зведеного східчастого вигляду метод Гауса — Йордана.

1. Звести матрицю до східчастого вигляду.
2. Відкинути нульові рядки.
3. Поділити останній рядок на його лідера, отримуючи 1.
4. Додати до решти рядків новий останній рядок, помножений на відповідні коефіцієнти, отримуючи нулі над 1.
5. Повторити попередні кроки для решти рядків.

Відповідь: 1,2,3,4,5.

Тестові завдання закритої форми є найбільш поширеними. До переваги завдань закритої форми можна віднести:

- оцінювання завдань повністю об'єктивне: між оцінками різних викладачів не може бути відмінностей;
- швидке опрацювання відповідей, яка не потребує великих зусиль від викладача;

- неважливо, чи вміють іспитники добре формулювати відповідь;
- простота заповнення бланку відповідей знижує ймовірність випадкових помилок та описок;
- ці завдання дають змогу охопити великі обсяги навчального матеріалу, що для тестів досягнень особливо важливо;
- можливе отримання точного оцінювання змістовності тесту, що особливо важливо для визначення відповідності тесту цілям дослідження.

Незважаючи на різноманітність видів тестових завдань, а також на переваги тестування у цілому, застосування тестів на практиці досі не стало масштабним з кількох причин: прихильність викладачів до традиційних методів контролю та недовіра до тестування; недостатня інтеграція тестових технологій контролю в процеси навчання і атестації студентів; тестовий контроль вимагає комп'ютерної бази для його проведення; процес розроблення та аналізу педагогічного тесту досить тривалий і трудомісткий, вимагає високого рівня компетентності викладачів. У зв'язку з цим «виникає необхідність пошуку шляхів підвищення якості педагогічних тестів з метою забезпечення достовірної оцінки навчальних досягнень студентів» [110, с. 28].

Важливим і актуальним є розгляд проблеми застосування тестів з вищої математики у ЗВО. Це пов'язано з рядом факторів. Вища математика – це одна з базових дисциплін студентів технічних спеціальностей, на яку спираються інші дисципліни. Тому актуальною є проблема створення нових, адекватних цілям і завданням навчання засобів контролю знань студентів з вищої математики, необхідних для моделювання та розв'язування задач за обраним фахом [54].

Студенти в короткі строки опановують великі обсяги формалізованих знань. При цьому вони мають опунувати як теоретичний матеріал, так і навчитись розв'язувати практичні завдання. За умов значного скорочення аудиторного навантаження необхідним є впровадження нових методів

управління навчально-пізнавальною діяльністю студентів. Найбільш придатним типом управління процесом навчання є управління зі зворотним зв'язком. Реалізація такого управління потребує ефективних форм контролю, що мінімізують витрати навчального часу на виконання робіт студентами та часу їх перевірки викладачами.

Таким чином, тестовий контроль з вищої математики може стати для студентів не лише засобом визначення рівня навчальних досягнень, але й сучасною технологією навчання та розвитку. Оперативність опрацювання результатів і повідомлення їх студенту надає можливість проводити тестування багаторазово з високою періодичністю, що сприяє реалізації функцій зворотного зв'язку та контролю. Оскільки студент є самостійною особистістю, здатною до самостійної роботи та самоконтролю, то інформація, яку він отримує про структуру власних навчальних досягнень, часто є достатньою для спрямування власних навчальних дій. Таким чином, зменшується потреба в прямому керуванні або спеціальному заохоченні з боку викладача. Необхідно зазначити, що самостійна робота студента відіграє вирішальну роль у навчанні вищої математики.

За допомогою комп'ютера можна посилити контроль не тільки основних теоретичних знань, але і навичок вирішення типових завдань. Застосування комп'ютерного тестування з вищої математик надає значні переваги при створенні малюнків, графіків, об'ємних зображень та ілюстрацій до запитань [138].

Використання ІКТ для контролю знань сприймається студентами позитивно, оскільки тестування на комп'ютері цікавіше в порівнянні з традиційними формами опитування. Це можна пояснити застосуванням комп'ютерних технологій в багатьох сферах їхнього життя, зокрема в більшості навчальних курсах. Також комп'ютерне тестування допомагає вирішувати студенту психолого-педагогічні проблеми, такі як неприйняття особистості викладача, страх перед екзаменаційною лотереєю та необ'єктивністю викладача.

Оскільки тестування з вищої математики унеможлиблює контроль навичок усного мовлення та простеження логіки міркувань іспитника; діагностику оригінальності мислення у вирішенні навчальних проблем і завдань; аналіз всіх етапів та способів розв'язання задачі, то спільно з використанням тестів необхідно практикувати і традиційні форми контролю. В роботі [140] зазначено, що оптимального поєднання форм контролю можна досягти так: за допомогою комп'ютерних тестів проводиться попередній контроль (залік) знань визначень і відомостей; після залікових випробувань проводиться зустріч студентів із викладачем, де вони демонструватимуть вже не формальні знання, а навички комунікації гостроту та оригінальність мислення, здатність до навчання.

Доцільність використання комп'ютерних тестів саме викладачами вищої математики можна пояснити їх здатністю зрозуміти математико-статистичний апарат, що використовується при оцінюванні якості тестів. До того ж, тестування дає змогу викладачеві зменшити рутинну роботу з контролю знань. Оскільки навантаженні викладача передбачає перевірку розрахунково-графічних робіт, типових розрахунків, самостійних, контрольних, модульних, залікових та екзаменаційних робіт, то даний фактор є досить вагомим при виборі викладачем тестування як засобу контролю. Використання комп'ютерного тестування як засобу контролю знань студентів з вищої математики дає змогу викладачам розробляти ефективні моделі контролю навчальних досягнень студентів з максимальним урахуванням сучасних підходів, специфіки, цілей і завдань вивчення математичних дисциплін.

1.3. Сучасний стан дослідження оцінювання якості тестів

У відповідності до цілей дослідження доцільно проаналізувати результати досліджень вітчизняних та закордонних вчених щодо проблеми оцінювання якості тестів та відповідної ІК-підтримки. Огляд наукової педагогічної і психологічної літератури засвідчив надзвичайно високу зацікавленість освітньої спільноти в методах розроблення високоякісних

педагогічних тестів. Розроблення та застосування педагогічних тестів в системі освіти є предметом досліджень багатьох авторів. Традиційно аналіз якості тестів проводиться у рамках двох теорій: КТТ та IRT.

На сьогоднішній день особливої актуальності набуває підхід до створення тестів на основі саме сучасної теорії IRT, що спрямована на оцінювання латентних якостей іспитника та параметрів завдань тесту на основі математико-статистичних моделей вимірювання. У цілому IRT розрахована на створення індустрії тестів і вимагає спеціальної текстологічної підготовки фахівців. Досвід західних країн засвідчує, що вони значно випереджають педагогічну науку країн СНД за масштабами застосування тестів, числом публікацій, підготовкою наукових кадрів, обчислювальною технікою і сучасними комунікаційними технологіями. В багатьох країнах практично вже створено індустрію тестування з власною інфраструктурою. Тому і основні теоретичні здобутки IRT належать саме закордонним вченим.

Одним із перших внесків у теорію і практику вимірювання в освіті стала монографія Г. Раша (G. Rasch) [198]. Вона стала поштовхом для наукових досліджень у галузі математичної статистики та педагогічних вимірювань. Теорія освітніх вимірювань у рамках моделі Раша та основи її застосування викладені у роботах Е. В. Сміт (E.V. Smith) та М. С. Сміт (M.S. Smith) [204], О. Д. Дункан (O.D. Duncan) [176] Р. М. Сміт (R.M. Smith) [205] та ін.

Розвитком дихотомічної моделі Раша є двопараметрична модель Бірнбаума, яка передбачає впровадження параметра диференціюючої спроможності [169]. У рамках IRT побудовано та досліджено моделі для аналізу політомічних завдань [196]. У роботах Дж. Н. Мастерса (G. N. Masters) [192, 193] досліджено політомічну модель Раша-Мастерса (або Partial credit model), розглянуто процес оцінювання параметрів методом максимальної вірогідності. У роботі М. Уілсона (M. Wilson) та Дж. Н.

Мастерса (G. N. Masters) [194] досліджено проблему наявності тривіальних рівней у моделі Раша-Мастерса.

Ряд робіт присвячено моделям IRT для аналізу завдань з множинним вибором. Так в праці Р. Д. Бока (R. D. Bock) [171] запропоновано та досліджено модель Nominal Response Model. В роботі Ф. Самеджими (F. Samejima) [202] запропоновано та досліджено модифікацію моделі Бока, модель Самеджими. Д. Тіссен (D. Thissen) та Л. Стейнберг (L. Steinberg) в роботі [209], в свою чергу, вдосконаливши модель Самеджими, розробили свою модель для завдань з множинним вибором та дослідили питання оцінювання параметрів моделі.

Одним із основних питань IRT є питання оцінювання латентних параметрів. Методи аналізу складності тестових завдань та рівня підготовленості іспитників за моделлю Раша демонструються у праці Р. М. Сміт (R.M. Smith) [206]. В роботі [182] розглянуто методи оцінювання латентних параметрів, а саме методи максимальної правдоподібності, граничну оцінку максимальної правдоподібності, Байєсовські оцінки, тощо.

Іншим важливим питанням IRT є дослідження інформаційної функції завдань для різних моделей. Поняття інформаційної функції та її загальна формула для дихотомічних завдань було введено у роботі [169] А. Бірнбаумом (A. Birnbaum). Дослідженню інформаційної функції для дихотомічних завдань присвячено роботи багатьох авторів, зокрема, роботи Ф. Бейкера (F. Baker) [168]. Формула інформаційної функції для політомічних завдань була отримана Ф. Семеджимою (F. Samejima) [201]. В роботі [195] Е. Муракі (E. Muraki) досліджувалась інформаційна функція для узагальненої Partial Credit моделі.

Деякі аспекти практичного застосування IRT наведено в роботах Ф. Б. Бейкера (F. Baker) [168], Р. К. Хемблтона (R. K. Hambleton) [181], Ф. М. Лорда (F. M. Lord) [189], М. Вілсона (M. Wilson) [212] та ін. Зокрема, Ф. Б. Бейкер (F. Baker) [168] наводить рекомендації щодо інтерпретації результатів тестування, проаналізованих в рамках IRT, Р. К. Хемблтон (R. K.

Hambleton) [181] розглядає можливості застосування IRT щодо оцінювання параметрів іспитників у різних галузях, таких як освіта, психологія та ін.

Фундаментальною роботою, в якій досліджено сучасний стан моделей і методів IRT є збірник Handbook of modern Item Response Theory [187]. Він надає широкі можливості щодо вивчення сучасної теорії тестування та її застосування в освіті.

У вітчизняних та російськомовних роботах в основному представлено огляд існуючих моделей IRT та досліджено питання оцінювання латентних параметрів різних моделей. Так найпростіші ідеї та методи IRT з основними поняттями теорії ймовірностей викладено в роботі Ю. М. Неймана та В. А. Хлебникова [124]. Вона розрахована на широке коло викладачів, зацікавлених у використанні тестування як засобу оцінювання знань.

Моделі IRT для дихотомічних завдань та поняття інформаційної функції були розглянуті в роботах В. С. Аванессова [5, 8]. Проблему оцінювання латентних параметрів висвітлено в роботах А. А. Маслака [113] Т. С. Анісімової [11] С. А. Позднякова [132] та ін. Зокрема С. А. Поздняков досліджує питання точності вимірювання латентних параметрів в освіті, надає практичні рекомендації з управління якістю тестових завдань. Україномовною роботою, в якій описано основні моделі IRT, методи оцінювання латентних параметрів та рекомендації до інтерпретації якості тестів є робота Лісової [107].

Засновником *класичної теорії тестів* є відомий британський психолог, автор факторного аналізу, Чальз Едвард Спирмен (C. E. Spearman) [207]. На Ч. Е. Спирмена сильний вплив спричинили роботи щодо тестування інтелекту Ф. Гальтона (F. Galton), який був засновником ідеї статистичного опрацювання результатів експерименту. Учнями Ч. Е. Спирмена були Р. Б. Кеттел (R. B. Cattell) і Д. Векслер (D. Wechsler). У числі його послідовників можна назвати А. Анастасі, Дж.П. Гілфорд (J. P. Guilford), П. Вернон (P. Vernon) та С. Берт (C. Burt).

Значний внесок у розвиток класичної теорії тестів вніс Л. Гуттман (L. Guttman) [180]. Всебічно і досить повно класична теорія тестів вперше викладена в фундаментальній праці Г. Гулліксена (H. Gulliksen) [179]. З того часу теорія кілька видозмінювалася, зокрема удосконалювався математичний апарат. З російських дослідників вперше опис цієї теорії у 1989 р. дав В. Аванесов [3].

Класичну теорію тестів в сучасному вигляді висвітлено в книзі Л. Крокер (L. Croker) та Дж. Алгіни (J. Algina) [99]. В роботі подано введення в теорію вимірювань, представлено статистичні концепції в теорії тестування, описано процес конструювання тестів, розтлумачено поняття надійності, висвітлено проблему валідності тестів та шляхи її вирішення. Деякі аспекти опрацювання результатів тестування із використанням КТТ висвітлено у роботі вітчизняних дослідників Л. О. Кухар та В. П. Сергієнка [104].

Існує ряд робіт, в яких *одночасно* розглянуто *КТТ і IRT*. Наприклад у роботі Д. Тіссена (D. Thissen) та Г. Вайнера (H. Wainer) [210] викладено основи цих двох теорій, зазначено шляхи їхньої інтеграції, а також визначено перспективи розвитку освітніх вимірювань методами психометрії.

Частково процес опрацювання результатів тестування на основі класичної теорії тестів та із застосуванням методів IRT розглянуто в роботах В. І. Звоннікова [80], В. С. Кіма [91], та М. Б. Чельшкової [161, 162]. В роботі вітчизняних вчених В. Ю. Бикова, Ю. М. Богачкова та Ю. О. Жука описанні деякі аспекти опрацювання результатів тестування в рамках КТТ та IRT [20].

Було досліджено питання оцінювання якості відомих тестів, таких як ЗНО, ЄДЕ, SAT, NAEP, GRE та ін. Інформація про методики оцінювання якості цих тестів є закритою. Можна знайти лише деякі аспекти такого оцінювання.

Так відповідно до офіційних звітів Українського центру оцінювання якості освіти про проведення зовнішнього незалежного оцінювання результатів навчання оцінювання якості тестів ЗНО відбувається із

застосування КТТ [127]. У звіті обчислюються основні характеристики вибірки балів учасників тестування: середнє, мода, медіана, стандартне відхилення, асиметрія та ексцес. Оцінюються такі характеристики тесту як складність тесту, розподільна здатність тесту, надійність тесту за коефіцієнтом Кронбаха та стандартна похибка вимірювань. Також будується розподіл учасників тестування за кількістю набраних балів. Оцінюються такі характеристики окремих тестових завдань як складність, дискримінативність (D-index), кореляція між результатами виконання певного завдання тесту учасниками тестування та загальним балом. Надаються рекомендації щодо інтерпретації різних значень цих характеристик. Будуються діаграми розподілів тестових завдань за складністю та розподільною здатністю, точкові діаграми розсіювання тестових завдань за складністю та розподільною здатністю, діаграми розсіювання тестових завдань за коефіцієнтом кореляції.

Щодо оцінювання якості ЄДЕ мало що відомо. Зі звітів Федерального інституту педагогічних вимірювань можна зробити висновок, що обчислюється складність завдань в рамках КТТ та будуються розподіл балів учасників, гістограми виконання різних груп завдань [84]. Є також відомості, що одна з моделей IRT, а саме модель Раша-Мастерса, використовується для шкалювання результатів тестування [120].

Для оцінювання якості SAT також використовують методи КТТ, оцінюються такі характеристики: складність завдань, валідність окремих завдань (коефіцієнт точкової бісеріальної кореляції дихотомічних завдань, коефіцієнт кореляції між балами за завдання та загальними балами учасників), надійність (за методом Кьюдера-Річардсона та Кронбаха α), змістова та прогностична валідності. Відомо також, що оцінювання якості SAT здійснюється із використанням IRT, зокрема однопараметричної моделі Раша [203]. Також IRT використовується для створення паралельних форм тестів SAT.

Аналіз вищезазначених робіт виявив відсутність загальноприйнятої методики оцінювання якості тесту, зокрема тестів з вищої математики. Розробнику тесту необхідно докласти чималих зусиль для аналізу наявних методів оцінюванні, вибору з них найбільш придатних та застосуванню їх на практиці. Існуючі рекомендації, в основному, описують процес математичного опрацювання даних, не приділяючи належної уваги їхньої реалізації та інтерпретації. Проблема підготовки викладачів вищої математики до застосування тестів також не знайшла достатнього відображення у дослідженнях як самостійна наукова проблема, що має велике теоретичне і практичне значення. Тому актуальною є задача систематизації існуючих методів оцінювання якості тесту, розроблення методики оцінювання якості тестів з вищої математики, її впровадження, чому і присвячено наше дослідження.

Та слід зауважити, що опрацювання великих обсягів результатів тестування, обчислення великої кількості параметрів, зокрема латентних параметрів IRT-моделей, вимагає виконання громіздких математико-статистичних обрахунків, що, в свою чергу, призводить до необхідності використання ІКТ. Розглянутий досвід використання тестів у вищій освіті вказує на те, що оцінювання рівня навчальних досягнень за допомогою комп'ютерних тестів набуває все ширшого застосування. Широкі можливості використання ІКТ у зберіганні та обробці великих обсягів статистичних даних дають змогу впроваджувати новітні математичні методи тестології. Дослідження західних вчених призвели до побудови не лише численних теорій, але і комп'ютерних систем на їхній основі, які в останні десятиліття впроваджуються і в Україні. Ймовірнісні моделі і теорії показали себе ефективним і потужним засобом для оцінювання якості тестів, тому використовуються в зарубіжних і деяких вітчизняних ЗВО. Проте комп'ютерної реалізація цієї теорії для тестів з вищої математики потребує подальшої розробки.

Створення розвиненого математичного апарату КТТ та IRT

забезпечило розроблення *програмних засобів опрацювання результатів тестування*. Існують як безкоштовні, так і пропрієтарні програмні засоби. Їх розрізняють за такими критеріями як об'єм вхідних даних, розмір результатних таблиць, використання конкретних статистичних методів тощо. Програми дають змогу опрацьовувати великий об'єм даних та обчислювати необхідні статистичні показники. Користувач в результаті отримує інформацію про характеристики тестових завдань та тесту у цілому, про рівень підготовленості іспитників, про ступінь відповідності рівня навчальних досягнень іспитників рівню складності завдань тощо. Розглянемо деякі з них.

Програма LOGIST використовувалась у свій час для обчислень за методом максимальної вірогідності на основі трипараметричної моделі Раша. Вона давала змогу одночасно оцінювати параметри складності завдань та параметри підготовленості іспитників за сумісною процедурою максимальної вірогідності. Недоліком даної процедури є вимога великої вибірки іспитників (більше 1000) для забезпечення більш точних оцінок.

Програма BILOG дає змогу отримати слушні оцінки параметрів тестових завдань за одно-, дво- і трипараметричної логістичних моделей з використанням процедури маргінальної максимальної вірогідності, раніше відомої.

Програма BICAL також оцінює параметри моделі Раша за сумісною процедурою максимальної вірогідності оскільки її результати дуже прості для інтерпретації [99]. BICAL дає змогу отримати майже слушні оцінки параметрів.

Популярною серед користувачів є програма WINSTEPS або її Windows версія BIGSTEPS, призначена для оцінювання та калібрування параметрів дихотомічних завдання за моделлю Раша та політомічних завдань за моделями Partial Credit і Rating Scale. Для оцінювання параметрів у програмі використовується метод максимальної вірогідності. Принципи роботи у WINSTEPS наведено в роботі [185]. Дана програма набула популярності

частково завдяки своїй безкоштовній академічній версії MINISTEP, що дає змогу аналізувати тест, складений не більше, ніж з 25 завдань, який виконувало не більше 75 іспитників.

Програми RUMM та WINMIRA використовують умовну процедуру максимальної вірогідності CML для моделі Раша. Приклад застосування програми RUMM можна знайти у монографії В.С. Кіма [91]. RUMM2020 дає можливість аналізувати параметри тестових завдань відповідно до ймовірнісної моделі Раша за спрощеними алгоритмами обчислень параметрів PROX та NORM. Алгоритм опрацювання результатів тестування за її допомогою можна знайти в роботі [144].

Пакет QUEST дає змогу оцінювати тестові завдання та питання соціологічних анкет. Аналіз даних можна проводити із використанням як КТТ, так і IRT.

До переліку популярних програм можна також віднести MULTILog, PARSCALE, продукти компанії ASC (Assessment System Corporation) та багато інших. ASC Item and Test Analysis Package – пакет програм ASC для аналізу якості тестів і шкалювання результатів їх виконання. Програми пакету дають змогу здійснити найпростіший аналіз якості завдань за класичною теорією тестів, в тому числі і поглиблений дистракторний аналіз (ITEMAN), отримати оцінки параметрів підготовленості іспитників і складності завдань за однопараметричною моделлю IRT та побудувати графіки характеристичної та інформаційної функцій тесту (RASCAL), оцінити латентні параметри з використанням маргінального методу максимальної вірогідності (MML), проаналізувати результати тесту за дво- та три- параметричною моделями IRT, провести калібрування завдань (XCALIBRE), оцінити надійність тесту (TESTINFO), оцінити валідність тесту (TESTVAL), порівняти параметри рівня підготовленості іспитників, оцінені із використанням різних моделей IRT (SCOREALL) і т.д. Програма дає змогу проводити аналіз коротких тестів з кількістю іспитників менше 1000. До недоліків пакету ITAP можна віднести незручну графічну інтерпретацію

даних та відсутність функції побудови характеристичних та інформаційних функцій кожного завдання.

ASC є приватною компанією, створеною у 1979 р. американськими вченими Д. Вайсом (D. Weiss). Крім розроблення програмних засобів для оцінювання тестів, вона також спеціалізується на розробці та оцінюванні педагогічних та професійних тестів, зокрема на комп'ютерному адитивному тестуванні. Оцінювання якості тестів здійснюється на основі методів КТТ та IRT (моделі Раша та Бірнбаума). Також компанія допомагає при сертифікації організацій в процесі акредитації Американським національним інститутом стандартів (American National Standards Institute) або Національною комісією із сертифікуючим агенством (National Commission for Certifying Agencies).

Програмний продукт BILOG-MG 3 надає змогу отримати оцінки параметрів тестових завдань на основі для одно-, дво- та трипараметричної моделі IRT, а також оцінити надійність і валідність тесту. CITAS (Classical Item and Test Analysis Spreadsheet) - дозволяє здійснити найпростішу первинну обробку даних тестування на основі класичної теорії тестів. Вхідним файлом є книга MS Excel, що містить результати тестування. Програма має обмеження на опрацювання аних (кількість іспитників та завдань не більше 100). Вихідним файлом є та ж книга з обчисленими характеристиками тесту та тестових завдань.

Програмне середовище **R** з відкритим вихідним кодом також містить ряд пакетів для обчислення характеристик КТТ, оцінювання латентних параметрів завдань та іспитників для різних моделей IRT та моделювання репрезентативних вибірок даних тестування із заданими параметрами. Опис пакетів R для моделей IRT представлено у роботі [199]. До переваг програмного середовища можна віднести його безкоштовне розповсюдження. До недоліків – відносну складність у використанні для користувача, незнайомого з мовами програмування, незручний графічний інтерфейс.

Спроби створення програмного забезпечення на базі теорії латентних змінних в поєднанні з класичною теорією тестування вже були в Росії з середини 90-х років. Першою такою програмою, створеною на базі MS DOS, була програма R-Latent, розроблена колективом авторів за участю В.С. Аванесова і М.Б. Челишкової. Реалізація в ній можливостей теорії латентних змінних в розумному поєднанні з класичною теорією тестування давали змогу підвищити точність оцінювання якості тестів. Однак функціональні можливості R-Latent були досить обмежені. Перш за все, це виражалось в обмеженні розміру матриці результатів тестування. Так число завдань не повинно перевищувати 95, іспитників – кількох сотень.

Зі створених пізніше програмних засобів, що реалізують положення теорії латентних змінних можна програму «КаТеЗа» Федерального центру тестування Росії та програму «Камертон» Центру державної акредитації Міносвіти Росії [114]. До недоліків даних програм можна віднести відсутність функції побудови інформаційних кривих тестових завдань і тестів у цілому.

Також необхідно окремо виділити розроблену у 2012 р. українськими дослідниками в «Криворізькому державному педагогічному університеті» інформаційну систему «Logit», призначену для створення дихотомічних тестів перевірки рівня навчальних досягнень, рецензування тестів, проведення випробувань характеристик тестів на контрольних групах, побудови профілів запитань та респондентів в автоматичному режимі, проведення тестування індивідуально та у групах, уточнення характеристик тестів в процесі використання, реалізації різних стратегій тестового контролю. Особливістю ІС «Logit» є застосування оригінальної сплайн-моделі для моделювання профілів запитань та респондентів у рамках IRT. Це дало змогу реалізувати розрахунки профілів в автоматичному пакетному режимі без участі фахівця що має фахові знання із статистики та IRT.

Слід зазначити, що обчислення за класичною теорією більш прості, їх можна проводити з використанням різних статистичних пакетів. В рамках

класичної теорії тестування забезпечує опрацювання вхідних даних та отримання необхідної інформації відомий пакет програм SPSS Statistics.

Метою існуючих програм з обчислення характеристик IRT є, здебільшого, доведення теоретичних положень і вони майже не пристосовані для використання емпіричних даних у зв'язку із складністю інтерфейсів та введеними обмеженнями на кількість вхідних даних.

Більша частина розроблених у світі програм, зокрема вітчизняні програми, для конструювання тестів носить закритий характер. Матеріали, розміщені в Інтернеті, зазвичай включають лише описи технологій, перелік можливостей програм, демонстраційні версії і т.д. У вільному доступі відсутні програмні засоби для аналізу якості тестів з україномовним інтерфейсом.

Отже, виникає необхідність у розробці власного програмного забезпечення, яке слугуватиме засобом розробленої комп'ютерно орієнтованої методики оцінювання якості тестів та матиме такі переваги:

- зрозумілий україномовний інтерфейс;
- взаємодія із результатами тестування в системі Moodle;
- обчислення взаємодоповнюючих характеристик якості тестів КТТ та IRT;
- графічне представлення результатів опрацювання даних;
- формування та обслуговування бази каліброваних завдань.

Аналіз досліджень щодо проблеми оцінювання якості тестів та відповідних наявних програмних засобів дає змогу зробити висновок, що питання розроблення комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики викладачами ЗВО залишається відкритим.

Висновки до розділу 1

Аналіз проблеми тестового контролю знань студентів з вищої математики та пов'язаних з нею понять, вивчення вітчизняного та

зарубіжного досвіду й узагальнення результатів наукового пошуку дають змогу дійти таких висновків:

1. На основі результатів аналізу науково-педагогічних праць В.С. Аванесова, В.Ю. Бикова Т.М. Балихіної, В.П. Беспалька, М.І. Жалдака, А.О. Кузнєцова, Н.В. Морзе, О.М. Майорова, С.Л. Рубінштейн, О.М. Спіріна, Ю.В. Триуса, М.Б. Челишкової та ін. уточнено поняття «тест», «тестове завдання», «якість тесту», «оцінювання якості тесту», «комп'ютерно орієнтована методика оцінювання якості тесту». У дослідженні під **тестом** будемо розуміти систему завдань специфічної форми, певного змісту для об'єктивного оцінювання рівня навчальних досягнень з наперед заданою методикою аналізу результатів; під **тестовим завданням** – складову одиницю тесту, яка відповідає таким вимогам: зміст відповідає меті тестування; сформульовано максимально лаконічно, але зрозуміло; відповідає умові локальної незалежності; однакова для всіх іспитників інструкція і структура завдання; чітко визначене місце для фіксації відповідей; однаковість принципів оцінювання та інтерпретації результатів; кількісні характеристики задовольняють певним критеріям якості; під **якістю тесту** – відповідність вимогам, які висуваються до характеристик тесту у цілому та до окремих тестових завдань (базовими характеристиками якості тесту є надійність, валідність та ефективність, тестових завдань – складність, диференціююча здатність та характеристична крива); під **оцінюванням якості тесту** – процедуру встановлення ступені відповідності характеристик окремих тестових завдань та тесту у цілому критеріям якості та складання висновку про якість тесту; під **комп'ютерно орієнтованою методикою оцінювання якості тесту** – теоретично обґрунтовану сукупність змісту, методів та засобів оцінювання якості тесту з використанням ІКТ.

Визначення цих понять дало змогу означити напрямки аналізу результатів тестування, покладені в основу авторської комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики

2. Відмічено, що у зв'язку із реформуванням вітчизняної освітньої галузі та впровадженням ІКТ особливого значення для контролю знань набуває комп'ютерне тестування. Визначено переваги контролю знань студентів з вищої математики: автоматизація та оперативність опрацювання результатів тестування; прискорення зворотного зв'язку за результатами тестування, що дає змогу студенту самостійно виявляти прогалини в структурі своїх знань і вживати заходів до їх ліквідації; забезпечення об'єктивності оцінювання; можливість регулярно поповнювати та модифікувати систему тестових завдань; можливість використання малюнків, графіків, об'ємних зображень та ілюстрацій до завдань; звільнення викладача від виконання рутинних робіт, вирішення для студента психолого-педагогічних проблем, таких як неприйняття особистості викладача, страх перед екзаменаційною лотереєю та необ'єктивністю викладача. Наголошено, що тестування не замінює та не скасовує традиційних форм педагогічного контролю, заснованих на безпосередньому спілкуванні викладача зі студентом. Для зростання ефективності навчального процесу, на нашу думку, тестування повинно не підміняти, а доповнювати перевірені традиційні форми контролю знань.

3. Визначено основні причини обмеженого використання тестів з вищої математики для контролю знань студентів ЗВО: викладачі сприймають тести лише як певні психодіагностичні методики, більше придатні до гуманітарних наук; прихильність до традиційних методів контролю, заперечення інновацій; недовіра до тестування та його об'єктивності; величезна кількість псевдотестів та прикладів псевдотестування навчальних досягнень в практичній діяльності; процес розроблення та аналізу педагогічного тесту досить тривалий і трудомісткий, вимагає високого рівня компетентності викладачів. Відмічено, що використання комп'ютерних тестів саме викладачами вищої математики можна пояснити їх здатністю зрозуміти математико-статистичний апарат, що використовується при

оцінювання якості тестів. Обґрунтовано доцільність використання тестів з вищої математики для контролю знань студентів.

4. Проаналізовано дослідження щодо проблеми оцінювання якості тестів і виявлено відсутність загальноприйнятої методики оцінювання якості тестів з вищої математики. Існуючі рекомендації, в основному, описують процес математичного опрацювання даних, не приділяючи належної уваги їхньої реалізації та інтерпретації. Проблема підготовки викладачів вищої математики до застосування тестів також не знайшла достатнього відображення у дослідженнях як самостійна наукова проблема, що має велике теоретичне і практичне значення.

5. Розглянуто основні програмні засоби опрацювання результатів тестування: BICAL, BILOG, LOGIST, MULTILOG, RUMM, WINSTEPS та ін. Виявлено, що більша частина розроблених у світі програм, зокрема вітчизняні програми, для дослідження тестів носить закритий характер. Матеріали, розміщені в Інтернеті, зазвичай включають лише описи технологій, перелік можливостей програм, демонстраційні версії і т.д. У вільному доступі відсутні програмні засоби для аналізу якості тестів з україномовним інтерфейсом. Підкреслено необхідність у розробці власного програмного забезпечення, яке слугуватиме засобом розробленої комп'ютерно орієнтованої методики оцінювання якості тестів та матиме такі переваги: зрозумілий україномовний інтерфейс; взаємодія із результатами тестування в системі Moodle; обчислення взаємодоповнюючих характеристик якості тестів KTT та IRT; графічне представлення результатів опрацювання даних; формування та обслуговування бази каліброваних завдань.

Результати першого розділу дисертаційного дослідження висвітлено у науковій роботі автора [41].

РОЗДІЛ II. ПРОЕКТУВАННЯ КОМП'ЮТЕРНО ОРІЄНТОВАНОЇ МЕТОДИКИ ОЦІНЮВАННЯ ЯКОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ

2.1. Загальна методика дослідження

Актуальність, теоретичне й практичне значення проблеми оцінювання якості тестів з вищої математики, відсутність комплексної комп'ютерно орієнтованої методики оцінювання та підвищення якості тестів з вищої математики, відповідного програмного забезпечення та недостатня розробленість проблеми в науковій літературі зумовили вибір теми дослідження *«Комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики викладачами ЗВО»*.

Провідною ідеєю дослідження є положення про те, що для використання тесту як засобу оцінювання знань студентів з вищої математики необхідною є умова аналізу якості розроблених тестів та відповідно розвиток компетентності викладачів щодо оцінювання якості тестів. Провідні ідеї дослідження відображені у **гіпотезі дослідження**, яка полягає у тому, що педагогічно виважене впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики сприятиме розвитку у викладачів ЗВО відповідної компетентності за умови реалізації розробленої технології розвитку компетентності та покращенню контролю знань студентів.

Нормативно-правову базу дослідження становлять:

Закон України «Про вищу освіту» [76] ;

– Закон України «Про Концепцію Національної програми інформатизації [77];

– Закон України «Про Національну програму інформатизації» [78];

– Закон України «Про інформацію [77];

– Указ Президента України «Про Національну доктрину розвитку освіти» [78];

- Указ Президента України «Про Національну стратегію розвитку освіти в Україні на період до 2021 року» [123] ;
- проект Концепції розвитку професійно-технічної освіти в Україні на 2010-2020 роки (проект) [98].

Теоретично-джерельну базу дослідження становлять положення наукових праць щодо: розроблення технологій комп'ютерного тестування (В. Ю. Биков, Ю. М. Богачков [20] І. Є. Булах [25], Ю. О. Жук[20], В. С. Кім [91], Л. О. Кухар [104], О. Г. Колгатін [96], М. Р. Мруга [25], С. А. Раков [18], В. П. Сергієнко [140], М. Б. Челишкова [162]); педагогічної тестології (В. С. Аванесов [3], Т. М. Балихіна [13], Н. А. Гулюкіна [37], В. Ю. Переверзєв [128], М. Б. Челишкова [161] та ін.), застосування методів КТТ для опрацювання результатів тестування Дж. Алгіна (J. Algina) [99], А. Анастасі (A. Anastasi) [10], Л. Гуттман (L. Guttman) [180], В. І. Звонніков [80], В. С. Кім [91], Т. Дж. Б. Клайн (T. J. B. Kline) [92], Л. Крокер (L. Croker) [99], Л. О. Кухар, В. П. Сергієнко [104], Р. К. Хемблтона (R. K. Hambleton) [182], Ч. Е. Спірмен (C. E. Spearman) [207], Н. Б. Челишкова [162]), методів ІРТ аналізу якості тестів (В. С. Аванесов [5, 8], Ф. Бейкер (F. Baker) [168], А. Бірнбаум (A. Birnbaum) [169], Г. Вайнер (H. Wainer) [211], Дж. Гласс (G. Glass) [33], Д. Ендріч (D. Andrich) [167], Н. Ф. Єфремова [70], В. І. Звонніков [80], К. Лінн (K. Linn) [186], Т. В. Лісова [107], Ф. М. Лорд (F. M. Lord) [189], Г. Раш (G. Rasch) [198], Н. Б. Челишкова [161].

З метою розв'язання поставлених задач використовували такі **методи** дослідження:

- *теоретичні* – аналіз, порівняння, узагальнення психолого-педагогічної, тестологічної, методичної літератури для обґрунтування теоретичних положень та уточнення основних понять дисертаційного дослідження; аналіз та систематизація алгоритмів і методів опрацювання результатів тестування для розроблення комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики;

– *емпіричні* – математико-статистичні методи для аналізу якості тестів з вищої математики; педагогічний експеримент для перевірки ефективності комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики викладачами ЗВО; анкетування, тестування, спостереження з метою визначення рівня розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики; експертне опитування з метою оцінювання змістової валідності тестів з вищої математики та для проведення експертами педагогічного оцінювання стану контролю знань студентів з вищої математики; моделювання педагогічних процесів для розроблення моделі розвитку компетентності викладача щодо оцінювання якості тестів з вищої математики; математично-статистичні методи для кількісного та якісного аналізу результатів педагогічного експерименту і перевірки гіпотези дослідження.

Організація дослідження. Робота над дослідженням здійснювалася протягом 6 років і охопила такі етапи науково-педагогічного пошуку.

На першому етапі (2012-2013 рр.) здійснено теоретичний аналіз наукової літератури з досліджуваної проблеми, обґрунтовано теоретичні та методичні засади дослідження, проаналізовано існуючі програмні засоби для аналізу якості тестів з вищої математики; проаналізовано законодавчі та нормативні документи з питань ІК-підтримки викладачів; визначено науково-термінологічний апарат, об'єкт, предмет, мету, завдання дослідження; визначено основні методи аналізу якості тестів у дослідженні, а саме методи КТТ та IRT; здійснено програмну реалізацію моделей Раша, Бірнбаума та Андерсена.

На другому етапі (2013-2015 рр.) визначено гіпотезу дослідження, досліджено питання аналізу ефективності тестових завдань методами IRT; розроблено комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики, розширено функціональні можливості СААЯТЗ.

На третьому етапі (2015-2018 рр.) розроблено модель та технологію розвитку компетентності викладачів щодо оцінювання якості тестів з вищої

математики; визначено основні критерії та рівні розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики; розроблено програму експериментального дослідження; визначено учасників експерименту; проведено формувальний етап експерименту; розроблено навчальну програму «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики» та відповідний їй дистанційний курс; здійснено аналіз та опрацювання результатів експериментальної роботи, їх узагальнення.

Експериментальна база дослідження. Дослідно-експериментальну роботу було виконано в Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського». Для експериментальної перевірки ефективності комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики проведено педагогічний експеримент. У ході педагогічного експерименту було здійснено інтегральне опрацювання даних та аналіз результатів дослідження; опис проведення, опрацювання результатів дослідження на основі методів статистичного опрацювання даних; оформлення результатів.

Вірогідність результатів дослідження забезпечувалась теоретичним обґрунтуванням вихідних положень, застосуванням комплексу методів педагогічного дослідження, адекватних меті й завданням дослідження, результатами проведеного педагогічного експерименту, опрацьованих статистичними методами.

2.2. Аналіз математичних методів оцінювання якості тестів

На сьогоднішній день завдяки потужному математичному апарату та повсюдному впровадженню комп'ютерної техніки найпопулярнішими методами аналізу якості тестів є статистичні методи. Існують дві базові теорії тестування: класична та сучасна. Розглянемо їх детально у даному пункті.

2.2.1. Аналіз методів класичної теорії тестів

Базовим у класичній теорії тестів є припущення про існування істинного балу. Істинний бал вважають параметром іспитника який ототожнюють з об'єктивною оцінкою іспитника, яка не залежить від впливу будь-яких похибок вимірювання. З поняттям істинного балу пов'язані такі п'ять основних положень КТТ [162]:

1. Емпірично отриманий результат вимірювання є сумою істинного балу та похибки вимірювання:

$$X_i = T_i + E_i$$

де X_i – спостережуваний бал i -ого іспитника за тест, T_i – істинний бал i -ого іспитника за тест f , E_i – похибка вимірювання при оцінюванні T_i . Величини T_i і E_i зазвичай невідомі.

2. Другий постулат пов'язаний з поняттям «паралельні форми тесту». Форми тесту вважаються паралельними, якщо вони розроблені на основі однієї специфікації, мають однакову кількість завдань з попарно рівними характеристиками, породжують на одній і тій же вибірці однакові розподіли спостережуваних балів (з рівними вибірковими характеристиками).

3. Третій постулат побудовано на припущенні про існування результатів нескінченної кількості тестувань одного і того ж іспитника з використанням паралельних форм тесту. При цьому, якщо кількість тестувань прямує до нескінченності, середнє значення похибок вимірювання істинного бала іспитника прямує до нуля.

4. Четвертий постулат побудовано на припущенні тестування нескінченної кількості іспитників з використанням набору паралельних форм тесту. Відповідно до цього постулату, якщо кількість іспитників прямує до нескінченності, середнє значення похибок вимірювання балів іспитників прямує до нуля.

5. Істинний бал іспитника не змінюється при використанні різних паралельних форм тесту.

Використання зазначених постулатів призводить до фундаментального співвідношення КТТ, згідно якого виконується рівність:

$$\sigma^2(X) = \sigma^2(T) + \sigma^2(E) \quad (2.1)$$

де $\sigma^2(X)$ – дисперсія спостережуваних балів генеральної сукупності, $\sigma^2(T)$ – дисперсія істинних балів генеральної сукупності, $\sigma^2(E)$ – дисперсія похибки вимірювання. Отже, дисперсія спостережуваних балів дорівнює сумі дисперсій істинних і помилкових складових. Для побудови КТТ систему постулатів доцільно доповнити припущенням про нормальність закону розподілу спостережуваних балів, істинних балів і похибок вимірювання.

Рівність (2.1) перепишемо у вигляді

$$\frac{\sigma^2(X)}{\sigma^2(X)} = \frac{\sigma^2(T)}{\sigma^2(X)} + \frac{\sigma^2(E)}{\sigma^2(X)}$$

або

$$\frac{\sigma^2(T)}{\sigma^2(X)} = 1 - \frac{\sigma^2(E)}{\sigma^2(X)} \quad (2.2)$$

У виразі (2.3) $\sigma^2(E)$ є середнім арифметичним дисперсій похибок для оцінок різних іспитників з генеральної сукупності. Необхідність введення середнього значення пояснюється тим, що похибка при оцінюванні істинного балу є різною у групах іспитників.

Чим ближчі значення $\sigma^2(X)$ і $\sigma^2(T)$, тим вища кореляція між множиною спостережуваних балів та множиною істинних балів. Відношення

$\frac{\sigma^2(T)}{\sigma^2(X)}$ трактується як концептуальне визначення коефіцієнта надійності

тесту, який є одним з найважливіших показників якості тесту. Фундаментальний аналіз якості тесту передбачає оцінювання надійності, валідності та аналіз розподілу вибірки індивідуальних балів.

Надійність тесту. Поняття надійності було наведено в п. 1.1. Позначивши коефіцієнт надійності r , перепишемо формулу (2.2) у вигляді

$$r = 1 - \frac{\sigma^2(E)}{\sigma^2(X)},$$

або

$$\hat{r} = 1 - \frac{s^2(E)}{s^2(X)}, \quad (2.4)$$

де $s^2(X)$ – вибіркова дисперсія спостережуваних балів іспитників, $s^2(E)$ – вибіркова дисперсія похибки вимірювання. Хоча формула (2.4) є базовою, вона не є зручною для використання, оскільки за емпіричними результатами виконання тесту не можна визначити $s^2(E)$.

Очевидно, що коефіцієнт надійності не може приймати значення більше одиниці. Максимальне значення $r = 1$ має місце, коли $\frac{s^2(E)}{s^2(X)} = 0$.

Такий випадок на практиці неможливий.

Оцінювання надійності тестів проводиться різними методами, які за способом реалізації можна поділити на дві групи. Перша група методів базується на двократному тестуванні, проведеному за допомогою одного і того ж тесту або за допомогою двох паралельних форм тесту. Друга група передбачає одноразове тестування при оцінюванні надійності тесту. Звичайно, на практиці віддають перевагу другій групі методів, оскільки і організація повторного тестування і створення паралельних форм пов'язані з рядом ускладнень для розробників тестів.

Незалежно від методу, оцінювання надійності найчастіше будується на обчисленні кореляції між двома наборами результатів виконання одного і того ж тесту або двох його паралельних форм. При цьому логіка міркувань така: чим вище кореляція, тим вище надійність тесту.

Найпоширенішими методами оцінювання надійності тестів є:

- метод паралельних форм;
- ретестовий метод;
- метод розщеплення;
- метод Кьюдера-Річардсона;

— метод Кронбаха α .

Метод паралельних форм вважають найбільш підходящим для оцінювання надійності в рамках класичної теорії тестів, оскільки ідея існування паралельних форм закладена в ній на рівні концепцій та постулатів [162]. Проте з практичної точки зору – це найменш ефективний метод. Це пояснюється тим, що дуже складно скласти тести паралельні як за змістом, так і за результатами. Метод паралельних форм передбачає обчислення коефіцієнту кореляції Пірсона між індивідуальними балами іспитників тестів.

Ретестовий метод з технічної точки зору є простішим, але також досить незручним, оскільки передбачає дворазове тестування студентів. Перше тестування змінює рівень підготовленості іспитників. Це відбувається з різних причин, зокрема, через можливість запам'ятовування завдань тесту. Тому повторне тестування необхідно проводити через значний проміжок часу після першого тестування.

Також при повторному тестуванні необхідно враховувати те, що змінюються зовнішні умови: соціальне середовище, взаємодія з членами іспитної групи, погодні умови та ін. Крім того, змінюються і самі іспитники, змінюється їх рівень знань. В результаті, повторне тестування проводиться в інакших умовах та в інакшій групі іспитників. У зв'язку з цим бажано, щоб інтервал між тестуваннями був коротшим. Таким чином, отримуємо несумісні вимоги до інтервалу між тестуваннями. Тому ретестовий метод зазвичай використовують досить рідко і віддають перевагу методам, що передбачає одноразове тестування студентів.

На підставі одноразового тестування здійснюється метод розщеплення. Для оцінювання коефіцієнта надійності за цим методом впорядковані результати тестування ділять на дві частини: в одну включають дані студентів по парних, а в іншу – по непарних завданнях. Формула для обчислення коефіцієнта надійності r має такий вигляд:

$$r = \frac{2\tilde{r}}{1 + \tilde{r}}, \quad (2.5)$$

де $\tilde{r}$ – коефіцієнт надійності для половини завдань тесту, що є коефіцієнтом кореляції між тестовими балами іспитників по парних та непарних завданнях тесту.

Формула Кьюдера-Річардсона дає змогу обчислити коефіцієнт надійності за варіацією тестового завдання:

$$r_{KR-20} = \frac{K}{K-1} \left(1 - \frac{\sum_{j=1}^K p_j q_j}{S^2(X)} \right), \quad j = \overline{1, K}$$

де p_j – частка правильних відповідей на j -те завдання тесту, q_j – частка неправильних відповідей на j -те завдання тесту, $q_j = 1 - p_j$, $S^2(X)$ – вибіркова дисперсія по розподілу індивідуальних балів, K – загальна кількість завдань у тесті. Недоліком даної формули є те, що вона дає змогу оцінити надійність тесту, який складається лише з дихотомічних завдань.

Коефіцієнт надійності Кронбаха α знаходять за формулою:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{j=1}^L S_j^2(X)}{S^2(X)} \right), \quad j = \overline{1, K}, \quad (2.6)$$

де $S_j^2(X)$ – вибіркова дисперсія по розподілу балів за j -е завдання; $S^2(X)$ – вибіркова дисперсія по розподілу індивідуальних балів, K – загальна кількість завдань у тесті.

Враховуючи перелічені вище недоліки методів паралельних форм, Кьюдера-Річардсона та ретестового методу, для оцінювання надійності тесту у роботі використовується метод розщеплення та метод Кронбаха α .

Валідність тесту. Іншим фундаментальним критерієм якості тесту є валідність. Поняття валідності наведено у п. 1.1.

Е. Стоунс зазначає, що «валідність – ахіллесова п'ята тестування» [146]. Розрізняють валідності: змістову, критеріальну та прогностичну та ін. [117, 124].

Валідність змісту (змістова валідність) – характеристика репрезентативності змісту тесту по відношенню до запланованих для перевірки знань і вмінь [162]. Вона залежить від глибини попередньо проведеного розробниками тесту аналізу навчального плану та програми, змісту підручників, а також від точності формулювання мети навчання.

Критеріальна валідність передбачає наявність зовнішнього критерію, кореляція з яким визначає валідність тесту [91]. У якості критерію беруться оцінки експертів, виставлені при традиційній перевірці знань студентів без використання тестів. Отже, для оцінювання критеріальної валідності обчислюється коефіцієнт кореляції між тестовими індивідуальними балами студентів і їх балами за традиційну контрольну роботу.

Валідність прогнозу (прогностична валідність) – це характеристика тесту, що відображає ступінь упевненості, що отримані за тест оцінки добре прогнозують майбутні досягнення іспитника (наприклад, у США результати за складання тесту SAT передбачають певний рівень успішності студента в ЗВО) [165]. Для визначення прогностичної валідності результати даного тесту порівнюють із результатами тесту, який проводять через деякий проміжок часу (наприклад, через рік навчання). Індивідуальні бали іспитника можуть бути використані для передбачення його успіху у майбутньому навчанні, професійній діяльності тощо.

Прогностична валідність не є обов'язковою характеристикою тестів для контролю знань студентів з вищої математики. Тому у роботі оцінюються змістова та критеріальна валідності тестів.

Аналіз розподілу вибірки результатів. Також у межах КТТ дослідження якості тесту у цілому передбачає аналіз розподілу вибірки результатів. Аналіз розподілу вибірки результатів полягає в графічній інтерпретації емпіричних даних у вигляді гістограми, аналізі розподілу

вибірки індивідуальних балів та обчисленні вибірових характеристик: математичного сподівання, дисперсії, моди, медіани, вибірового середнього, асиметрії.

Аналіз якості тестових завдань. Аналіз якості тесту неможливий без детального аналізу тестових завдань, який полягає в:

- аналізі кореляційної матриці завдань;
- оцінюванні валідності окремих завдань;
- аналізі дистракторів завдань множинного вибору.

Аналіз кореляційної матриці завдань передбачає обчислення кореляційної матриці завдань.

Оцінювання валідності окремих завдань передбачає обчислення значення індексу розрізняльної здатності дихотомічних завдань D_j , коефіцієнту точкової бісеріальної кореляції та бісеріального коефіцієнту кореляції між дихотомічним завданням та індивідуальними балами іспитників і коефіцієнту кореляції Пірсона між політомічним завданням та індивідуальними балами іспитників ρ_j .

Значення індексу розрізняльної здатності дихотомічних завдань знаходять за формулою:

$$D_j = p_{uj} - p_{lj}, j = \overline{1, K} \quad (2.7)$$

де p_{uj} – частка правильних відповідей на j -те завдання для 27% іспитників з найвищими індивідуальними балами; p_{lj} – частка правильних відповідей на j -те завдання для 27% іспитників з найнижчими індивідуальними балами, K – загальна кількість завдань у тесті.

Коефіцієнт кореляції Пірсона обчислюється за формулою

$$\rho_j = \frac{\sum_{i=1}^N (X_i - \bar{X})(X_{ij} - \bar{X}_j)}{\sqrt{\left(\sum_{i=1}^N (X_i - \bar{X})^2\right) \left(\sum_{i=1}^N (X_{ij} - \bar{X}_j)^2\right)}}, i = \overline{1, N}, j = \overline{1, K},$$

де $X_i - \bar{X}$ – відхилення тестового бала i -го студента від середнього балу по тесту; $X_{ij} - \bar{X}_j$ – відхилення тестового бала i -го студента за j -те завдання від середнього балу по j -ому завданню; N – кількість іспитників, K – загальна кількість завдань у тесті.

Формула для обчислення коефіцієнта точкової бісеріальної кореляції має вигляд

$$\rho_j = \frac{\bar{X}_{1j} - \bar{X}_{0j}}{S_X} \sqrt{\frac{N_{1j} \cdot N_{0j}}{N(N-1)}}, \quad j = \overline{1, K},$$

де $\bar{X}_{1j}$ – середнє значення індивідуальних балів іспитників, які правильно виконали j -те завдання тесту; $\bar{X}_{0j}$ – середнє значення індивідуальних балів іспитників, які неправильно виконали j -те завдання тесту; S_X – стандартне відхилення вибірки індивідуальних балів; N_{1j} – кількість іспитників, які правильно виконали j -те завдання тесту; N_{0j} – кількість іспитників, які неправильно виконали j -те завдання тесту; $N = N_{1j} + N_{0j}$, N – кількість іспитників, K – загальна кількість завдань у тесті.

Обчислюється бісеріальний коефіцієнт кореляції [99; с. 424] між дихотомічним завданням та індивідуальними балами іспитників ρ_{bisj} за формулою

$$\rho_{bisj} = \frac{\bar{X}_{1j} - \bar{X}}{S_X} \sqrt{\frac{p_j}{Y_j}}, \quad j = \overline{1, K},$$

де $\bar{X}_{1j}$ – середнє значення індивідуальних балів іспитників, які правильно виконали j -те завдання тесту; $\bar{X}$ – середнє значення індивідуальних балів іспитників; S_X – стандартне відхилення вибірки індивідуальних балів; p_j – частка іспитників, які відповіли на j -те завдання правильно; Y_j – ордината стандартної нормальної кривої в точці з абсцисою, яка дорівнює z -оцінці, що відповідає величині p_j (визначається за відповідною таблицею); K – загальна кількість завдань у тесті.

Аналіз дистракторів завдань множинного вибору полягає у тому, що дистрактори, які обирають менше 5% іспитників, мають бути видалені або замінені.

Методи КТТ є базовими в аналізі якості тестів. Але вони не дають можливості у повному обсязі здобувати інформацію з емпіричних результатів тестування. Для більш детального аналізу, як тесту у цілому, так і тестових завдань необхідне використання сучасної теорії тестів IRT.

2.2.2. Аналіз методів Item Response Theory

Математичну основу аналізу якості окремих тестових завдань (Item Analysis) складає Item Response Theory. IRT все більше привертає увагу вітчизняних фахівців з галузей педагогічних і психологічних вимірюванням. Одна з можливих причин популярності цієї теорії – це її застосування в таких широко відомих тестових службах: Національна Рада розвитку освіти (National Assessment of Educational Progress, NAEP), служба з розробки тестів для оцінювання здатності до опунування освітніх програмам різного рівня складності SAT, а також американська атестаційна служба GRE. Крім цих служб, IRT використовується в таких масштабних міжнародних порівняльно-оціночних дослідженнях, як третє міжнародне дослідження рівня підготовленості з математики і природничих наук (Third International Math and Science Survey), а також у міжнародній програмі оцінювання якості підготовленості студентів (the Programme of International Student Assessment) [67]. У Росії методи IRT використовують для аналізу результатів ЄДЕ. Однією з основних причин підвищеної уваги до IRT є можливість її застосування в освітніх технологіях.

IRT призначена для оцінювання латентних параметрів іспитників і параметрів завдань тесту за допомогою застосування математико-статистичних моделей вимірювання. До найбільш значимих переваг IRT зазвичай відносять такі:

– неповнота даних, а саме пропуск деяких комбінацій «іспитник – тестове завдання», не є критичною. Це дійсно вагома перевага IRT, оскільки аналіз результатів тестування з розширеним банком завдань з використанням методик КТТ неможливий;

– стійкість і об'єктивність оцінок параметра, що характеризує рівень підготовки іспитників. Стійкість можна вважати найбільш важливою перевагою IRT. Джерелом її є відносна інваріантність оцінок параметра іспитників від складностей завдань тесту;

– стійкість і об'єктивність оцінок параметра складності завдань, їх незалежність від властивостей вибірки іспитників, які виконують тест.

– можливість вимірювання значення параметрів іспитників та завдань тесту в одній шкалі.

Остання перевага є дуже важливою, оскільки перетворення вихідних величин різного походження в одну стандартну шкалу дає змогу співвіднести рівень знань будь-якого іспитника з мірою складності кожного завдання тесту.

За допомогою IRT можна передбачити ймовірність правильного виконання завдань тесту будь-яким іспитником у вибірці до пред'явлення тесту групі студентів, виявити ефективність різних за складністю завдань, що використовуються для оцінювання знань.

На відміну від класичної теорії, де індивідуальний бал іспитника розглядається як постійне число, у IRT параметр іспитника є деякою змінною. Це дає можливість послідовного наближення до об'єктивних оцінок параметра із використанням ітераційних методів. Початкове значення параметра при цьому обчислюється на основі емпіричних даних тестування.

Математичні моделі сучасної теорії тестів. Основним припущенням сучасної теорії тестів є наявність зв'язку між латентними параметрами іспитників і завдань та результатами тестування. При цьому припускають, що результати тестування є результатом взаємодії двох множин значень латентних параметрів.

Елементами першої множини є значення латентного параметра підготовленості іспитників θ_i , $i=1, \dots, N$. Другу множину утворюють значення латентного параметра складності завдань тесту β_j , $j=1, \dots, K$. Зв'язок між елементами цих двох множин зображено на Рис. 2.1.

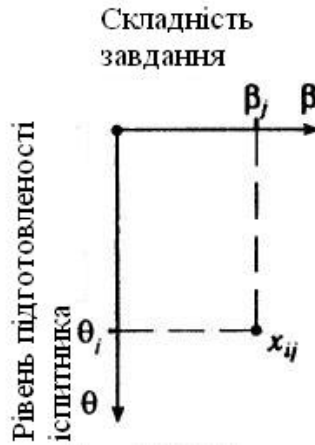


Рис. 2.1. Зв'язок між латентними параметрами іспитників та завдань

На практиці зазвичай ставиться обернена задача: за результатами тестування оцінити значення латентних параметрів. Для її розв'язання необхідно задати співвідношення між латентними параметрами θ і β . Ідея встановлення зв'язку між латентними параметрами належить датському математикові Г. Рашу, який запропонував ввести його у вигляді різниці $\theta - \beta$, припускаючи, що параметри θ і β оцінюються в одній шкалі [198].

Значення параметра θ_i можна розглядати як координату i -го іспитника, а значення β_j — як координату j -ого завдання на одній осі змінних θ і β . Тоді абсолютна величина різниці $|\theta_i - \beta_j|$ є відстанню, на якій знаходиться іспитник з рівнем підготовленості θ_i від завдання зі складністю β_j . Якщо ця різниця велика за модулем і від'ємна, то завдання не підходить для вимірювання рівня навчальних досягнень i -ого іспитника. Іспитник, напевно, виконає його неправильно. Велике додатнє значення цієї різниці також не підходять для контролю знань i -ого іспитника. Завдання такої складності є для нього занадто легким, і він, напевно, виконає його правильно.

У випадку, коли θ ненабагато більше β , іспитник хоча і може помилитися в завданні, але, скоріше за все, виконає його правильно. При від'ємних значеннях різниці $\theta - \beta$ іспитник, найімовірніше, виконає завдання

неправильно за умови, що не спробує вгадати правильну відповідь. Отже, для оцінювання іспитника із заданим параметром підготовленості з мінімальною помилкою вимірювання необхідно обирати завдання із рівнем складності, близьким до θ .

Розглянемо умовну ймовірність правильного виконання i -м іспитником з рівнем підготовленості θ_i різних за складністю завдань тесту, вважаючи θ_i параметром i -ого іспитника, а β — незалежною змінною. В цьому випадку умовна ймовірність буде функцією латентної змінної β :

$$P_i \{x_{ij} = 1 / \theta_i\} = f(\theta_i - \beta), \quad i = \overline{1, N}, j = \overline{1, K}.$$

$$X = \{x_{ij}\} = \begin{cases} 1, & \text{якщо відповідь } i\text{-ого студента на } j\text{-е завдання правильна,} \\ 0, & \text{якщо відповідь } i\text{-ого студента на } j\text{-е завдання неправильна,} \end{cases}$$

де N —кількість іспитників; K — кількість завдань в тесті.

Аналогічно записується умовна ймовірності правильного виконання j - ого завдання складністю β_j іспитниками з різним рівнем підготовленості. Тут незалежною змінною є θ , а β_j — параметр, що визначає складність j -ого завдання тесту:

$$P_j \{x_{ij} = 1 / \beta_j\} = \varphi(\theta - \beta_j), \quad i = \overline{1, N}, j = \overline{1, K},$$

де K — загальна кількість завдань у тесті.

Функції $f(\beta)$ і $\varphi(\theta)$ називають IRF. Графік функції P_j є характеристичною кривою j -ого завдання, а графік функції P_i — індивідуальною кривою i -ого іспитника.

IRF розподіляють на класи за кількістю параметрів, що входять в аналітичне завдання функцій. Серед них розрізняють:

- однопараметричну модель Г. Раша [198]

$$P_j(\theta) = \frac{\exp(1,7(\theta - \beta_j))}{1 + \exp(1,7(\theta - \beta_j))}, \quad j = \overline{1, K}, \quad (2.8)$$

$$P_i(\beta) = \frac{\exp(1,7(\theta_i - \beta))}{1 + \exp(1,7(\theta_i - \beta))}, \quad i = \overline{1, N}, \quad (2.9)$$

де θ і β незалежні змінні для першої та другої функцій відповідно; θ_i — параметр підготовленості i -ого завдання тесту; β_j — параметр складності j -ого завдання тесту; N — кількість іспитників, K — загальна кількість завдань у тесті;

- двопараметричну модель А. Бірнбаума [169]:

$$P_j(\theta) = \frac{\exp(1,7\alpha_j(\theta - \beta_j))}{1 + \exp(1,7\alpha_j(\theta - \beta_j))}, j = \overline{1, K},$$

$$P_i(\beta) = \frac{\exp(1,7\gamma_i(\theta_i - \beta))}{1 + \exp(1,7\gamma_i(\theta_i - \beta))}, i = \overline{1, N},$$

де θ_i — параметр підготовленості i -ого завдання тесту; β_j — параметр складності j -ого завдання тесту; α_j — параметр диференціюючої спроможності завдань тесту, введений А. Бірнбаумом (А. Birnbaum) [169]; γ_i — параметр, що характеризує структурованість знань студента, N — кількість іспитників, K — загальна кількість завдань у тесті;

- трипараметричну модель А. Бірнбаума [169]

$$P_j(\theta) = c_j + (1 - c_j) \frac{\exp(1,7\alpha_j(\theta - \beta_j))}{1 + \exp(1,7\alpha_j(\theta - \beta_j))}, j = \overline{1, K}.$$

де θ_i — параметр підготовленості i -ого завдання тесту; β_j — параметр складності j -ого завдання тесту; α_j — параметр диференціюючої спроможності завдань тесту; c_j — параметр, що характеризує ймовірність вгадування іспитником правильної відповіді на завдання j , K — загальна кількість завдань у тесті.

У кожній вищенаведеній моделі параметри θ і β виражаються в одній шкалі логітів. Завдяки цьому можна порівняти результати студентів, отримані за допомогою різних тестів, а також підібрати оптимальні значення β , що дають змогу оцінити шукане θ з мінімальною помилкою вимірювання. Ці переваги, доповнюючи КТТ, підвищують ефективність тестових вимірювань.

Однопараметрична модель Г. Раша. Однопараметрична модель є однією із сімейства логістичних кривих, описаних Г.Рашем у роботі [198].

Аналітичне задання однопараметричної моделі представлено формулами (2.8) і (2.9). Імовірність правильного виконання j -ого завдання тесту є зростаючою функцією від змінної θ . З практичної точки зору це можна пояснити тим, що чим більше рівень підготовки іспитника, тим більша ймовірність правильного виконання їм j -ого завдання тесту.

На Рис. 2.2 зображено характеристичну криву j -ого завдання тесту, що відображає взаємозв'язок між значеннями незалежної змінної θ і величиною $P_j(\theta)$. У точці перегину характеристичної кривої $\theta = \beta_j$, а $P_j(\theta)$ у цій точці дорівнює $\frac{1}{2}$. Отже, ймовірність відповіді іспитника з рівнем підготовленості,

що дорівнює складності j -ого завдання тесту, дорівнює $\frac{1}{2}$. Для іспитника з рівнем підготовленості набагато більшими за β_j , ймовірність правильної відповіді прямує до одиниці. Якщо ж рівень підготовленості іспитника θ набагато менший від β_j , то ймовірність правильного виконання j -ого завдання тесту прямує до нуля.

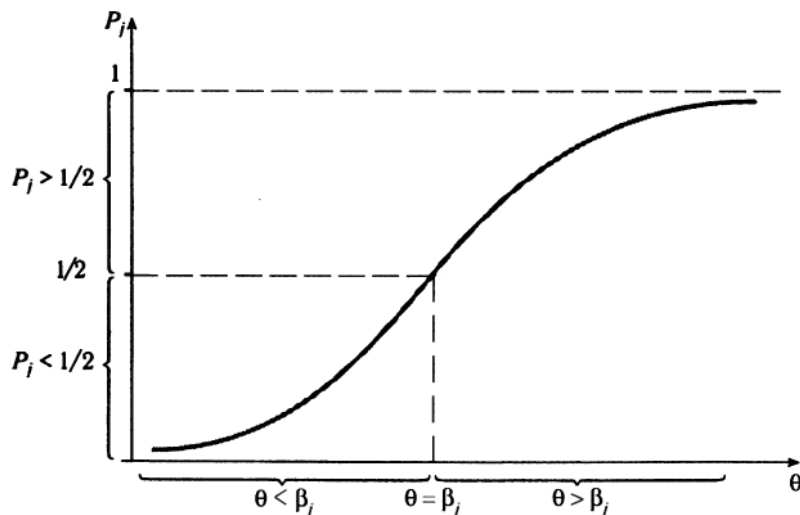


Рис. 2.2. Характеристична крива j -ого завдання тесту

Імовірність правильного виконання i -им іспитником різних за складністю завдань навпаки є спадною функцією змінної β . Це означає, що чим більше складність завдань, тим менше значення ймовірності $P_i(\beta)$. Графік функції $P_i(\beta)$ представлено на Рис. 2.3.

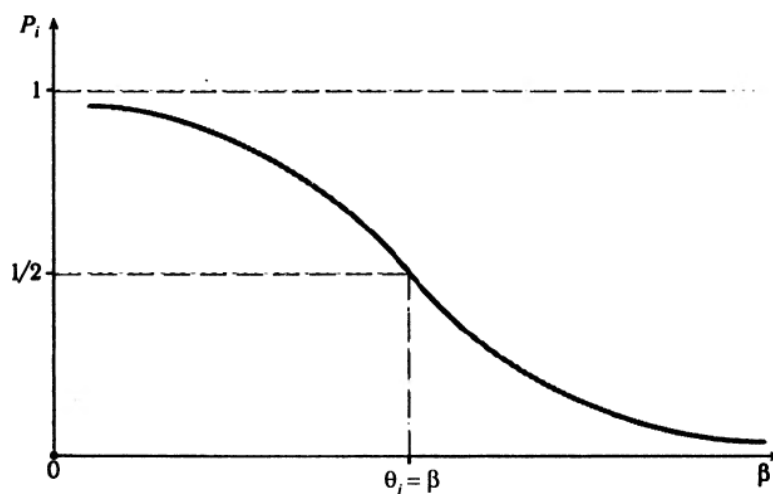


Рис. 2.3. Індивідуальна крива i -ого іспитника

Модель Раша є найбільш застосованою для дихотомічних завдань та найбільш простою з точки зору оцінювання латентних параметрів. У своїй роботі для оцінювання латентних параметрів за моделлю Раша обрали метод максимальної вірогідності. Оцінки, отримані на основі даного метода мають ряд переваг, зокрема є слушними та стійкими [182]. Також оцінки латентних параметрів паралельно можуть бути перевірені за допомогою методу PROX, описаного у роботі [161].

Побудова окремої характеристичної кривої завдання тесту за моделлю Раша не дає достатніх відомостей про якість завдання. Тому при оцінюванні якості тестів використовуємо модель Раша для побудови ансамблю характеристичних кривих тесту. Це дає змогу виявити занадто легкі та занадто складні завдання тесту, а також перевірити рівномірність наповнення тесту завданнями різної складності.

Двопараметрична модель Бірнбаума. Розвитком моделі Раша є модель Бірнбаума, яка передбачає введення параметра диференціюючої спроможності α , який показує наскільки добре завдання може розрізняти іспитників з різним рівнем підготовленості. За цією моделлю ймовірність правильної відповіді i -ого іспитника на j -те завдання тесту можна записати таким чином:

$$P_j(\theta_i) = \frac{1}{1 + \exp(-\alpha_j(\theta_i - \beta_j))}, \quad i = \overline{1, N}, j = \overline{1, K},$$

де θ_i — параметр підготовленості i -ого завдання тесту; β_j — параметр складності j -ого завдання тесту; α_j — параметр диференціюючої спроможності завдань тесту; N — кількість іспитників, K — загальна кількість завдань у тесті.

Параметр диференціюючої спроможності визначається за формулою:

$$\alpha_j = \frac{r_{bis}(j)}{\sqrt{1 - r_{bis}^2(j)}}, \quad j = \overline{1, K}$$

де $r_{bis}(j)$ — коефіцієнт кореляції між балами, отриманими за кожне тестове завдання і балами за тест в цілому, K — загальна кількість завдань у тесті.

У рамках двопараметричної моделі кожному тестовому завданню певної складності може відповідати кілька кривих з різними кутами нахилу дотичної (Рис. 2.4), які перетинаються в єдиній точці перегину. У роботі [162] надаються умови значення диференціюючої спроможності правильно складеного завдання. Кращим вважається те завдання, параметр спроможності якого більший. При малих α характеристична крива завдання є похилою. На практиці це означає, що іспитники з різними рівнями підготовленості виконують дане завдання з приблизно рівним успіхом. Такі завдання не дають об'єктивної інформації про рівень знань іспитників. Навпаки, якщо α велике, то шанси успішного виконання даного завдання іспитників з різними рівнями підготовленості істотно відрізняються. Параметр α може змінюватися в інтервалі $(-\infty; \infty)$, але завдання з від'ємними значеннями α не повинні включатися у тест, оскільки для них іспитники з вищим рівнем підготовленості мають менші шанси на правильну відповідь. На практиці, як правило, залишають завдання зі значенням α , що лежить в межах інтервалу від 0,5 до 2,5 [162].

У роботі використовуємо модель Бірнбаума для дихотомічних завдань, оскільки вона також є достатньо простою і дає змогу оцінити диференціюючу спроможність, на відміну від моделі Раша. Побудова характеристичних

кривих за моделлю Бірнбаума дає змогу наочно виявити дихотомічні завдання з низькою та від'ємною диференціюючою спроможністю.

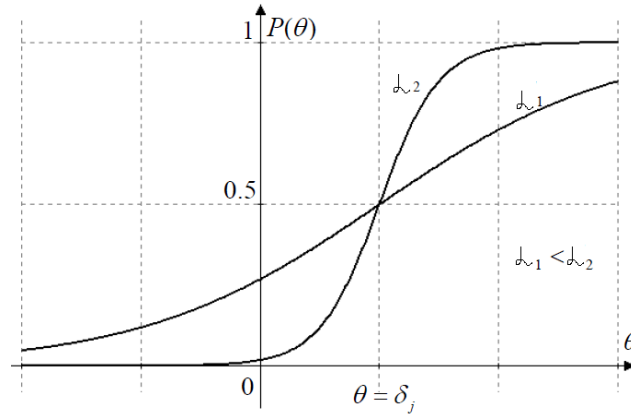


Рис. 2.4. Характеристичні криві завдань за моделлю Бірнбаума з різними значеннями диференціюючої спроможності

Політомічна модель Раша-Мастерса. Для аналізу тестових робіт, що містять політомічні завдання, застосовується політомічна модель Раша-Мастерса, або Partial Credit Model.

Розглянемо тестову роботу, що містить L завдань і яка виконується N іспитниками. За i -е завдання такої роботи іспитник може отримати j балів: $j = 0, 1, \dots, m_i$. 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.

В даній моделі разом з параметром підготовленості іспитників вводять параметри складності рівнів завдань, які позначають $\beta_{ij}, i = 1, 2, \dots, L, j = 1, 2, \dots, m_i$. Параметр складності β_{ij} визначає складність переходу із $(j - 1)$ -го рівня i -го завдання на j -й; що вищою є ця складність, то більшим є значення β_{ij} .

У політомічній моделі Раша-Мастерса ймовірність досягнення n -м іспитником j -го рівня i -е завдання означається таким чином:

$$P_{ij}(\theta_n) = \frac{\exp \sum_{l=0}^j (\theta_n - \beta_{il})}{\sum_{k=0}^{m_i} \exp \sum_{l=0}^k (\theta_n - \beta_{il})}, \quad n = \overline{1, N}, \quad i = \overline{1, L}, \quad j = \overline{1, m_i} \quad (2.10)$$

де θ_n — параметр підготовленості n -ого завдання тесту; β_{ij} — параметр складності переходу із $(j-1)$ -го рівня i -го завдання на j -й; L — кількість завдань тесту; N — кількість іспитників. Для визначеності приймаємо

$$\sum_{l=0}^0 (\theta_n - \beta_{il}) \equiv 0.$$

Залежності ймовірностей від рівня підготовленості іспитника зображують за допомогою характеристичних кривих рівнів завдання [107]. На Рис. 2.5 представлені криві для завдань з двома рівнями.

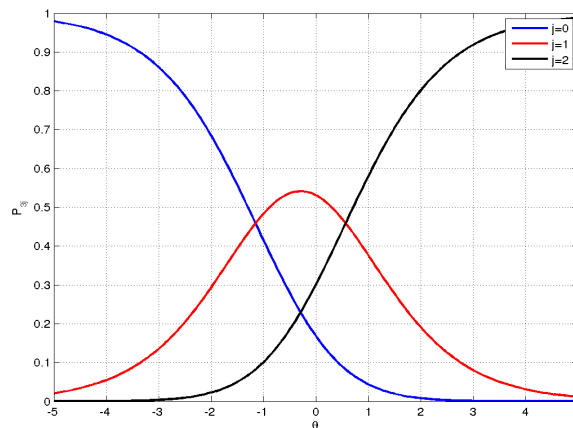


Рис. 2.5. Характеристичних кривих рівнів політоміного завдання за моделлю Раша-Мастерса

Політомічна модель Андерсена. Модель Раша-Мастерса є частинним випадком політомічної моделі Андерсена. Для означення політомічної моделі Андерсена введемо такі множини параметрів: параметри підготовленості іспитників, параметри складності рівнів завдань та бали, що даються за досягнення відповідних їм рівнів [187].

Параметр складності j -го рівня i -го завдання позначатимемо $\eta_{ij}, i=1,2,\dots,L, j=0,1,\dots,m_i$. Бал, який іспитник отримує за досягнення j -го рівня i -го завдання позначатимемо $a_{ij}, i=1,2,\dots,L, j=0,1,\dots,m_i$. Вимагається, щоб параметри a_{ij} утворювали арифметичну прогресію: $d = a_{ij} - a_{i(j-1)}, i=1,2,\dots,L, j=1,2,\dots,m_i$.

У політомічній моделі Андерсена ймовірність досягнення n -м іспитником x -го рівня i -ого завдання означається таким чином [194]:

$$P_{ix}(\theta_n) = \begin{cases} 0, & x\text{-й рівень – тривіальний,} \\ 1, & \\ \frac{\sum_{k=0}^{m_i} \exp(a_{ik}\theta_n - \eta_{ik})}{\sum_{k=0}^{m_i} \exp(a_{ik}\theta_n - \eta_{ik})}, & x = b_i, \\ \frac{\exp(a_{ix}\theta_n - \eta_{ix})}{\sum_{k=0}^{m_i} \exp(a_{ik}\theta_n - \eta_{ik})}, & x > b_i. \end{cases}$$

де η_{ix} – параметр складності x -го рівня i -го завдання; θ_n – рівень підготовленості n -ого іспитника; a_{ix} – бал, який іспитник отримує за досягнення x -го рівня i -го завдання (вимагається, щоб параметри a_{ix} утворювали арифметичну прогресію); b_i – порядковий номер найнижчого нетривіального рівня i -го завдання.

У роботі для оцінювання політомічних завдань використовується саме модель Андерсена. Її використання дає змогу оминати проблему тривіальних рівнів, яку детально буде розглянуто у п. 2.2.4.

Моделі для завдань з множинним вибором. Особливої уваги при аналізі якості потребують дихотомічні завдання з множинним вибором, тобто завдання, які мають декілька варіантів відповіді, серед яких лише один варіант є правильним [7]. Використовуючи моделей Раша та Бірнбаума, не можна зрозуміти наскільки добре спрацювують дистрактори. Тому виникає необхідність розглянути моделі IRT для завдань множинного вибору.

Р.Бок (R. Bock) у 1972 р. запропонував модель Nominal Response Model (NRM), яка припускає, що ймовірність того, що особа з рівнем підготовленості θ_i обере категорію (варіант відповіді) l у j -ому завданні ($l=1,2,\dots,m$) має вигляд [171]:

$$P_{lij} = \frac{\exp(a_{jl}\theta_i + c_{jl})}{\sum_{k=1}^m \exp(a_{jk}\theta_i + c_{jk})},$$

де a_{jl} – параметри диференціації, які змінюються для різних характеристичних кривих варіантів відповіді навіть у межах одного завдання, c_{jl} – інтерсепти кривих, які відображають популярність категорії. Величину

$z_{jl} = a_{jl}\theta_i + c_{jl}$ називають багатовимірним логітом, а сама модель будується за принципом багатовимірної моделі Раша. Додатково вводяться обмеження на параметри [197]:

$$\sum_{k=0}^{m_j} a_{jk} = \sum_{k=0}^{m_j} c_{jk} = 0.$$

У рамках моделі Бока один із варіантів відповідей повинен мати найбільше додатне значення a_k ; характеристична крива для цього варіанту монотонно зростає. Теоретично такий варіант ймовірно є правильною відповіддю. Інший варіант повинен мати найбільше за модулем від'ємне значення a_k ($\sum a_k = 0$ є однією з лінійних обмежень), і характеристична крива цього варіанту повинна бути монотонно спадною. Тоді, оскільки параметр здібності зменшується, ймовірність вибору однієї конкретної неправильної відповіді наближається до одиниці, а всі інші до нуля. Це не відповідає дійсності.

У роботі [202] запропоновано вирішення цієї проблеми в концептуальній модифікації моделі Бока, в яку додається латентна категорія відповіді, названа «нульовою». Іноді цю категорію називають як «не знаю» (НЗ). В роботі [190] вводиться аналогічний параметр, який вказує на відсоток студентів, які «абсолютно не визначилися з відповіддю». В моделі Самеджими [202], НЗ – це змінна d_h , яка чисельно дорівнює $1/m_j$ і характеризує ймовірність вибору варіанта відповіді h в тому випадку, коли студент обирає відповідь навмання. Таким чином, модель Бока набуває вигляду:

$$P(k, \theta_i) = \frac{e^{(a_{jk}\theta_i + c_{jk})} + d_{jk}e^{(a_{j0}\theta_i + c_{j0})}}{\sum_{s=0}^{h_j} e^{(a_{js}\theta_i + c_{js})}}, i=1, N, j=1, K, \quad (2.11)$$

де k – номер варіанту відповіді; θ_i – параметр підготовленості i -ого іспитника; a_{jk} – диференціююча спроможність k -ого варіанту відповіді j -ого завдання; c_{jk} – складність k -ого варіанту відповіді j -ого завдання; d_{jk} – ймовірність вибору варіанта відповіді k в тому випадку, коли студент обирає

відповідь навмання; h_j – кількість варіантів відповідей j -ого завдання; N – кількість іспитників; K – загальна кількість завдань у тесті.

Але при цьому параметр d_{jk} є константою, а це є припущенням того, що іспитники обирають різні варіанти відповіді навмання з рівними ймовірностями для кожного з варіантів відповіді. Це не є правдоподібним. А, отже, виникає необхідність розглянути модель Тіссена-Стейнберга [209], яка припускає, що ймовірність вибору i -м іспитником k -ого варіанту j -ого завдання задається формулою (2.11). Параметр d_{jk} в цій моделі є не константною, а латентним параметром, який необхідно оцінити.

Відповідні залежності ймовірностей від рівня підготовленості іспитника зображують за допомогою характеристичних кривих варіантів відповіді (Рис. 2.6).

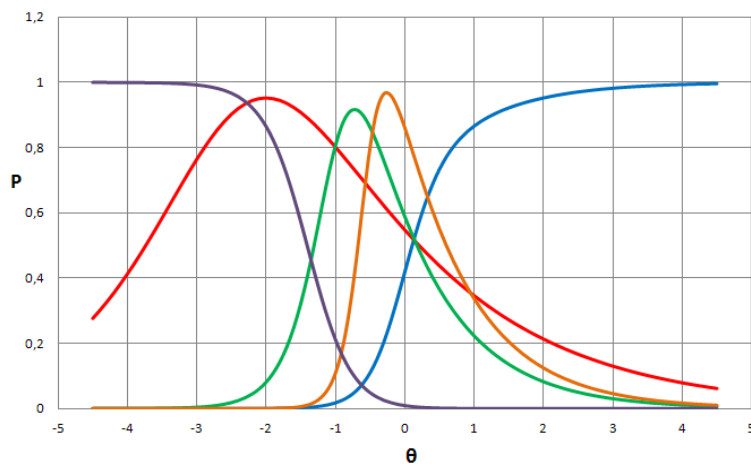


Рис. 2.6. Характеристичні криві варіантів відповіді завдання множинного вибору за моделлю Тіссена-Стейнберга

Враховуючи те, що моделі Раша та Бірнбаума не дають змогу оцінити правильність складання варіантів відповідей завдань, необхідним є використання IRT-моделей для завдань множинного вибору. Недолік моделі Бока, пов'язаний із тим, що при зменшенні рівня підготовленості іспитника, ймовірність вибору однієї конкретної неправильної відповіді наближається до одиниці, а всі інші до нуля. Недоліком моделі Самеджими є припущення того, що іспитники обирають різні варіанти відповіді навмання з рівними ймовірностями для кожного з варіантів відповіді. Враховуючи вищесказане, у роботі для оцінювання завдань множинного вибору використовується

модель Тіссена-Стейнберга. Хоча вона є відносно складною з точки зору оцінювання латентних параметрів, але найбільш точно описує завдання з множинним вибором.

2.2.3. Оцінювання ефективності тестів

Однією з компонент аналізу якості тесту у рамках IRT є дослідження інформаційної функції для оцінювання його ефективності. Це дослідження дає підставу для постановки питання про створення тесту, який би найбільш оптимально відповідав заданому рівню підготовленості групи студентів. На відміну від класичної теорії, яка не дає змогу підвищити ефективність тестових вимірювань, в IRT можна ставити питання про диференційовану оцінку ефективності j -го завдання тесту для оцінювання кожного значення θ .

Поняття кількості інформації розглядалось у роботах Р. А. Фішера, А. Бірнбаума, Ф. М. Лорда, Г. Кендала та А. Стюарта, Р. К. Хемблтона. У 1968 році у роботі [169] А. Бірнбаумом (A. Birnbaum) було введено поняття інформаційної функції та виведено загальну формулу інформаційної функції для дихотомічних завдань. Дослідженню інформаційної функції для дихотомічних завдань присвячено роботи багатьох авторів, зокрема, роботи Ф. Бейкера [168]. Формула інформаційної функції для політомічних завдань була отримана Ф. Самеджимою [201] у 1969 році. В роботі [195] Е. Муракі досліджувалась інформаційна функція для узагальненої Partial Credit моделі.

Вітчизняні дослідження в цій галузі були практично відсутні. Лише частково інформаційна функція розглядалась Т. В. Лісовою [107], М. Б. Челишковою [162], В. С. Аванессовим [8], Н. Ф. Єфремовою [70], В. С. Кімом [91].

Для оцінювання ефективності окремих тестових завдань і тесту в цілому в рамках IRT впроваджується інформаційна функція. Вона показує, наскільки ефективно дане завдання тесту оцінює іспитника з певним значенням параметра підготовленості.

Інформаційна функція базується на понятті кількості інформації, яка забезпечується оцінюванням параметра θ за j -те завдання, має такий вигляд [107]:

$$I_j(\theta) = \frac{(E'_j(\theta))^2}{\sigma_j^2(\theta)}, j = \overline{1, K}, \quad (2.12)$$

де K – кількість завдань в тесті; $E_j(\theta)$ – математичне сподівання оцінки параметра підготовленості θ ; $E'_j(\theta)$ – похідна математичного сподівання оцінки параметра θ ; $\sigma_j(\theta)$ – стандартне відхилення оцінки параметра θ .

Якщо завдання дихотомічне, то на підставі (2.12) його інформаційну функцію можна записати таким чином:

$$I_j(\theta) = \frac{(P'_j(\theta))^2}{P_j(\theta)(1 - P_j(\theta))}, j = \overline{1, K} \quad (2.13)$$

де $P_j(\theta)$ – ймовірність правильної відповіді на j -те завдання, яка залежить від обраної моделі; θ – параметр підготовленості іспитника; K – кількість завдань в тесті.

У чисельнику функції (2.13) маємо похідну характеристичної функції завдання, яка досягає максимуму у точці перегину. Оскільки точці перегину відповідає значення $\theta = \beta_j$ (Рис. 2.7), то можемо зробити висновок, що для оцінювання конкретного значення рівня підготовленості i -ого іспитника θ_i найбільш інформативними будуть завдання з околу точки θ_i . Такі завдання вважаються найбільш інформативними і вони мають найменшу стандартну похибку.

Для однопараметричної моделі Раша, враховуючи формулу (2.8),

$$P'_j = P_j Q_j, j = \overline{1, K},$$

тому

$$I_j(\theta) = P_j(\theta) Q_j(\theta), j = \overline{1, K},$$

де $Q_j = 1 - P_j, j = \overline{1, K}$ є ймовірністю неправильної відповіді студентів на j -е завдання тесту; θ – параметр підготовленості іспитника; K – кількість завдань в тесті.

Максимум інформаційної функції завдання у моделі Раша дорівнює 0,25 (Рис. 2.7) у точці $\theta = \beta_j$, де характеристична функція завдання має перегин, а значення ймовірності правильної відповіді на j -те завдання: $P_j(\theta) = 0,5$.

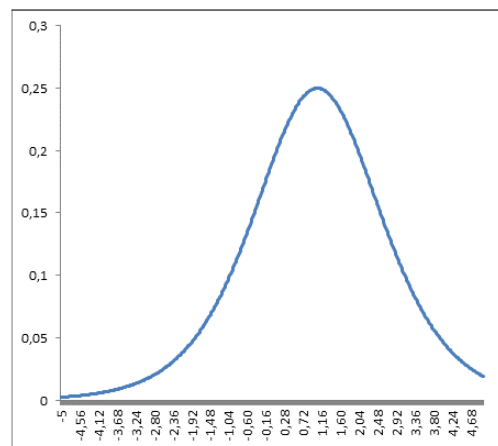


Рис. 2.7. Інформаційна функція завдання за однопараметричною моделлю Раша

Для двопараметричної моделі Бірнбаума [169]

$$I_j(\theta) = \alpha_j^2 P_j(\theta) Q_j(\theta), j = \overline{1, K},$$

де

$$P_j(\theta) = \frac{1}{1 + \exp(-\alpha_j(\theta - \beta_j))}, j = \overline{1, K};$$

$$Q_j(\theta) = 1 - P_j, j = \overline{1, K};$$

θ – параметр підготовленості іспитника; α_j – диференціююча спроможність j -ого завдання тесту; β_j – параметр складності j -ого завдання тесту K – кількість завдань в тесті.

Для моделі Бірнбаума максимум інформаційної функції також досягається у точці $\theta = \beta_j$. Можна припустити, що найбільш інформативні завдання з більш крутими характеристичними кривими. Чим крутіша крива,

тим більший внесок завдання у вимірювання даного значення θ . Однак, прагнучи включити до тесту завдання з найбільш крутими характеристичними кривими, можна зробити помилку і прийти до зниження ефективності вимірювання в окремих точках осі θ за рахунок не виправданого її збільшення в інших точках цієї ж осі.

Справа в тому, що зростання крутизни характеристичної кривої крім позитивних ефектів, супроводжується і негативними. Останні пов'язані із зменшенням довжини інтервалу осі θ , що забезпечує диференціацію індивідуальних відмінностей іспитників. Цей негативний ефект характерний для значень θ , що лежать по один бік від точки перегину кривої. Чим крутіша крива завдання, тим менше розрізняються значення ймовірностей правильного виконання завдання в таких точках. Тому при відборі завдань для тесту потрібен додатковий аналіз, що враховує характер розподілу значень латентної змінної θ і внесок всіх завдань у сумарну інформацію для кожної точки осі θ .

Для політомічних завдань інформаційну функцію j -ого завдання знаходять за формулою [201]:

$$I_j(\theta) = \sum_{l=0}^{m_j} \frac{[P'_{lj}(\theta)]^2}{P_{lj}(\theta)}, j = \overline{1, K}, \quad (2.14)$$

де $P_{lj}(\theta)$ – ймовірність того, що особа з рівнем підготовленості θ у j -му завданні досягне l -ого рівня, $l = \overline{1, m_j}$; m_j – кількість рівнів у j -му завданні; θ – параметр підготовленості іспитника; K – кількість завдань в тесті.

Формула (2.14) для моделі Тіссена-Стейнберга [107] приймає вигляд

$$I_j(\theta) = \frac{\left(e_{jh} \left(a_{jh} \sum_{k=0}^{m_j} e_{jk} - \sum_{k=0}^{m_j} a_{jk} e_{jk} \right) + d_{jh} e_{j0} \left(a_{j0} \sum_{k=0}^{m_j} e_{jk} - \sum_{k=0}^{m_j} a_{jk} e_{jk} \right) \right)^2}{\left(e_{jh} + d_{jh} e_{j0} \right) \left(\sum_{k=0}^{m_j} e_{jk} \right)^3}, j = \overline{1, K}, h = \overline{1, m_j}.$$

де $e_k = a_{jk}\theta + c_{jk}$, $k = \overline{0, m_j}$ – номер категорії відповіді, m_j – кількість категорій у j -му завданні; θ – параметр підготовленості іспитника, a_{jk} –

диференціююча спроможність k -ої категорії відповіді j -ого завдання; c_{jk} – складність k -ої категорії відповіді j -ого завдання; d_{jk} – параметр вгадування k -ої категорії відповіді j -ого завдання.

Для моделі Раша-Мастерса або Partial Credit [195] формула (2.14) набуває вигляду

$$I_j(\theta) = \sum_{l=0}^{m_j} \left[l - \sum_{k=0}^{m_j} k P_{kj}(\theta) \right]^2 \cdot P_{lj}(\theta), \quad j = \overline{1, K},$$

де

$$P_{jg}(\theta) = \frac{e^{\sum_{l=0}^g (\theta - \beta_{jl})}}{\sum_{k=0}^{m_j} e^{\sum_{l=0}^k (\theta - \beta_{jl})}}, \quad j = \overline{1, K}, \quad g = \overline{1, m_j},$$

де θ – параметр підготовленості іспитника, β_{jg} – складність переходу з $(g-1)$ -го рівня j -го завдання на g -й, m_j – кількість рівнів у j -му завданні; K – кількість завдань в тесті.

На прикладі трирівневих завдань, які досліджуються на підставі моделі Раша-Мастерса, розглянемо можливі варіанти вигляду інформаційних кривих завдань та відповідних характеристичних кривих рівнів завдань.

Інформаційна функція правильно складеного завдання повинна мати дзвоноподібну, але не занадто загострену форму (Рис. 2.8а). При цьому значення різниці параметрів складності третього та другого рівнів має бути невеликим і додатнім ($\Delta\beta_j = \beta_{j3} - \beta_{j2}$, $\Delta\beta_j \in (1; 2,8)$), що відобразатиметься впорядкованим розташуванням характеристичних кривих рівнів завдання (Рис. 2.8б). Таке завдання є найбільш інформативним, якщо його використовувати для оцінювання знань студентів з рівнями підготовленості, розкиданими по всій шкалі θ .

На Рис. 2.9а зображено інформаційна функція завдання, графік якої має провал в середині шкали θ . Тобто для відповідних значень параметра підготовленості дане завдання неінформативне. При цьому значення різниці параметрів складності третього та другого рівнів є додатнім і досить великим

($\Delta\beta_j > 2,8$), що відображається на вигляді характеристичних кривих рівнів завдання (Рис. 2.9б). Таке завдання потребує вдосконалення.

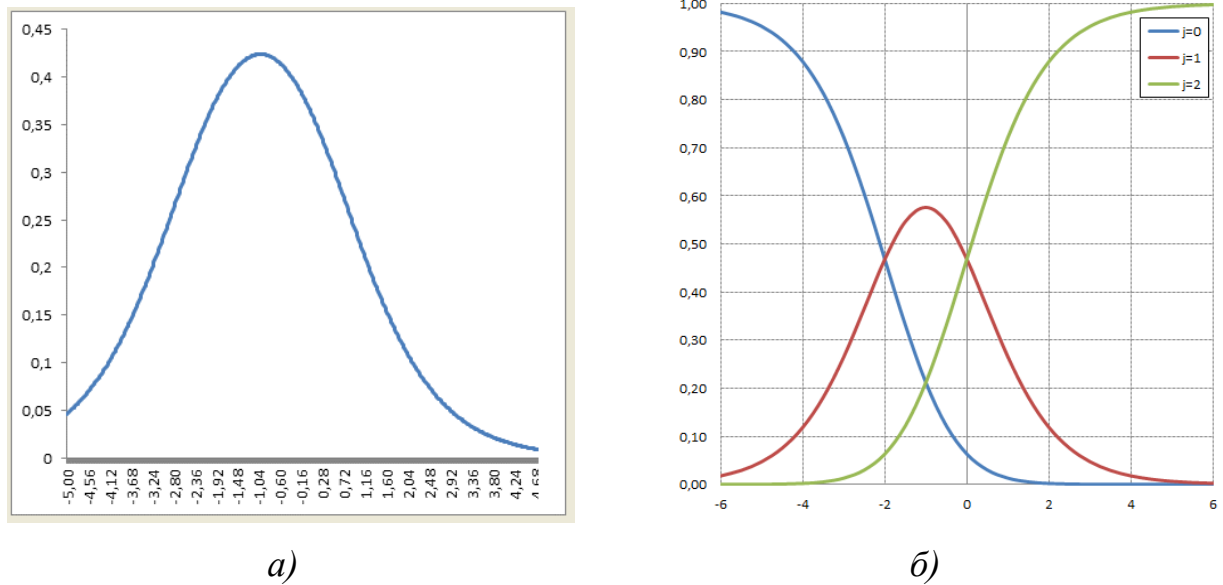


Рис. 2.8 Інформаційна крива завдання (а) та характеристичні функції рівнів (б), $\bar{\beta}_j = (0, -2, 0)$

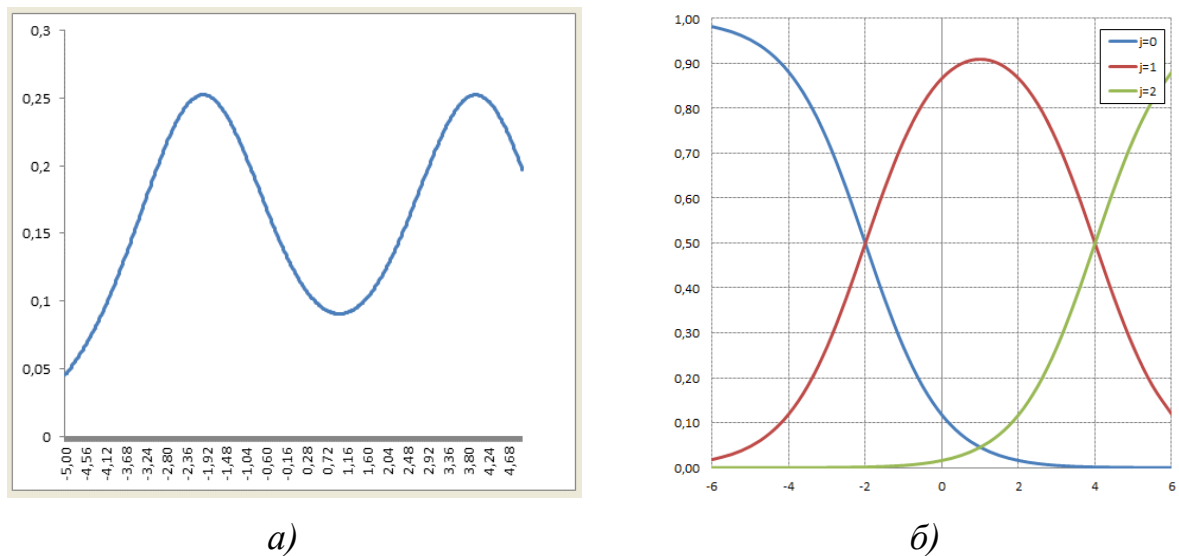


Рис. 2.9. Інформаційна крива завдання (а) та характеристичні функції рівнів (б), $\bar{\beta}_j = (0, -2, 4)$

Коли різниця між значеннями параметрів складності третього та другого рівнів стає від'ємною ($\Delta\beta_j < 0$) і зростає за модулем, завдання є інформативним лише на невеликому діапазоні рівня підготовленості θ (Рис. 2.10а). Аналізуючи вигляд характеристичних функцій рівнів завдання (Рис.

2.10б), можна зробити висновок, що другий рівень завдання не спрацьовує. Таке завдання є недосконалим, його доречно вилучити з тесту.

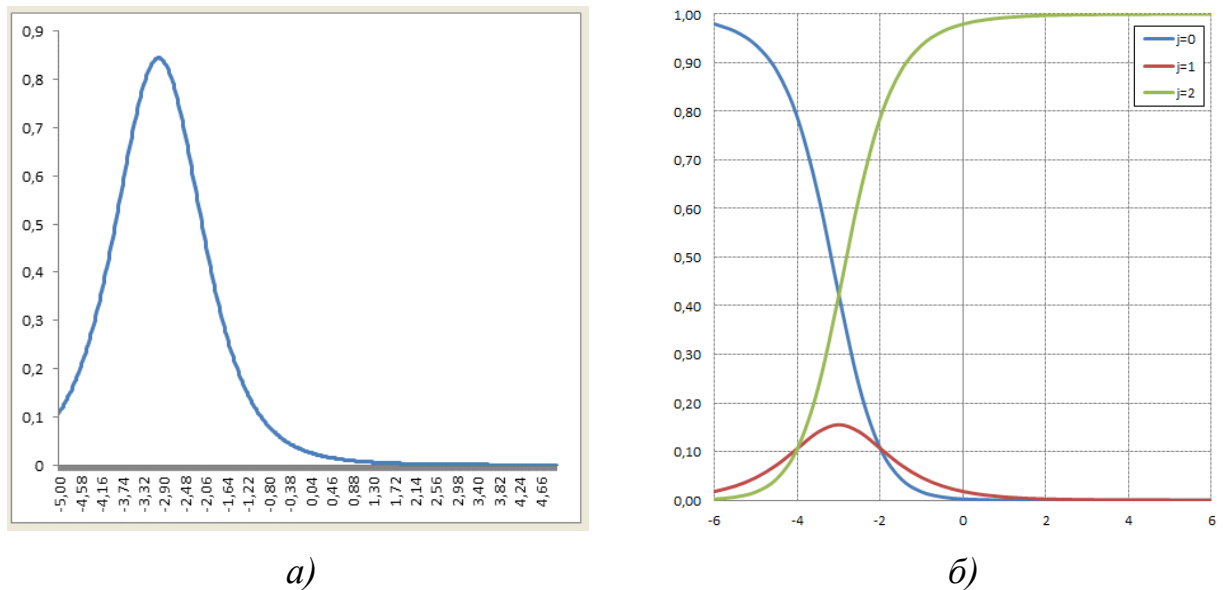


Рис. 2.10. Інформаційна крива завдання (а) та характеристичні функції рівнів (б), $\bar{\beta}_j = (0, -2, -4)$.

У випадку, коли має місце «провал» інформаційної кривої, можна збільшити кількості категорій у завданні, що призведе до підвищення точності оцінювання. Ілюструє це Рис. 2.11. Перше завдання має усього два кроки, складності яких дуже сильно відрізняються, у другому завданні додано проміжний крок, а третє завдання має п'ять кроків.

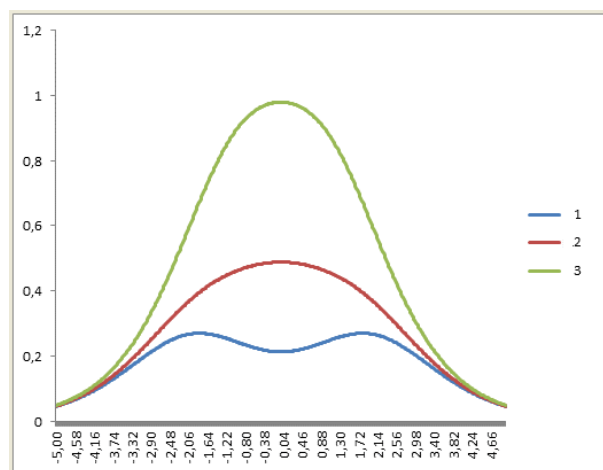


Рис. 2.11. Інформаційні криві трьох завдань:

$$\bar{\beta}_1 = (0; -2; 2), \bar{\beta}_2 = (0; -2; 0; 2), \bar{\beta}_3 = (0; -2; -1; 0; 1; 2)$$

Завдяки властивості адитивності інформація, отримана при вимірюванні даного θ за допомогою всього тесту, є сумою окремих складових $I_j(\theta)$ для кожного завдання:

$$I(\theta) = \sum_{j=1}^K I_j(\theta), \quad j = \overline{1, K},$$

де $I(\theta)$ - інформаційна функція тесту.

Після обчислення значення інформаційних функцій, можна порівняти ефективність різних новостворених тестів. Припустимо, що два тести X і Y оцінюють одну й ту ж саму латентну змінну θ . Порівняльну ефективність тесту X по відношенню до тесту Y можна охарактеризувати спеціальною функцією $E(X, Y)$, значення якої дорівнюють відношенню значень інформаційних функцій тестів Y і X у відповідних точках осі θ [162]:

$$E(Y, X) = \frac{I_Y(\theta)}{I_X(\theta)},$$

де $I_Y(\theta), I_X(\theta)$ – відповідно інформаційні функції тестів Y та X ; θ – параметр підготовленості іспитника. Функція $E(Y, X)$ отримала назву функції порівняльної ефективності. Вона дає змогу оцінити ефект після корекції тесту.

При розробці тесту зазвичай обирають область, що забезпечує високу інформативність оцінок θ на тому інтервалі, де розташована вибірка іспитників.

У своїй роботі для оцінювання ефективності тесту у цілому виконуємо побудову інформаційної функції за моделлю Раша-Мастерса оскільки, що вона дає змогу охопити як дихотомічні, так і політомічні завдання тесту.

2.2.4. Оцінювання латентних параметрів

Для оцінювання латентних параметрів IRT-моделей використовують метод максимальної вірогідності. Розглянемо його застосування на прикладі моделі Раша-Мастерса.

Розглянемо матрицю результатів виконання тестової роботи $\mathbf{X} = (x_{ni}), n = 1, 2, \dots, N, i = 1, 2, \dots, L, x_{ni} = 0, 1, \dots, m_i$. Значення елементу на перетині n -го рядка та i -го стовпця матриці $\mathbf{X}$ дорівнює кількості балів, яку n -й іспитник отримав за i -е завдання тестової роботи.

Згідно (2.10) функція вірогідності Λ для моделі Раша-Мастерса має вигляд

$$\Lambda = \prod_{n=1}^N \prod_{i=1}^L P_{ix_{ni}}(\theta_n) = \frac{\exp \sum_{n=1}^N \sum_{i=1}^L \sum_{j=0}^{x_{ni}} (\theta_n - \beta_{ij})}{\prod_{n=1}^N \prod_{i=1}^L \sum_{k=0}^{m_i} \exp \sum_{j=0}^k (\theta_n - \beta_{ij})}.$$

Для зручності шукатимемо точку максимуму функції

$$\lambda = \ln \Lambda = \sum_{n=1}^N \sum_{i=1}^L x_{ni} \theta_n - \sum_{n=1}^N \sum_{i=1}^L \sum_{j=1}^{x_{ni}} \beta_{ij} - \sum_{n=1}^N \sum_{i=1}^L \ln \sum_{k=0}^{m_i} \exp \sum_{j=0}^k (\theta_n - \beta_{ij}). \quad (2.15)$$

Вираз (2.15) доцільно спростити. Розглянемо величини $r_n = \sum_{i=1}^L x_{ni}, n = 1, 2, \dots, N$, значення якої дорівнює загальній кількості балів, отриманій n -м іспитником за розглядувану тестову роботу, та S_{ij} , значення якої дорівнює кількості іспитників, які досягли j -го рівня i -го завдання, або, інакше кажучи, отримали за i -е завдання не менш як j балів. Тоді матиме місце рівність $\sum_{n=1}^N \sum_{j=1}^{x_{ni}} \beta_{ij} = \sum_{j=1}^{m_i} S_{ij} \beta_{ij}, i = 1, 2, \dots, L$, завдяки чому вираз (2.15) набуває вигляду

$$\lambda = \sum_{n=1}^N r_n \theta_n - \sum_{i=1}^L \sum_{j=1}^{m_i} S_{ij} \beta_{ij} - \sum_{n=1}^N \sum_{i=1}^L \ln \sum_{k=0}^{m_i} \exp \sum_{j=0}^k (\theta_n - \beta_{ij}). \quad (2.16)$$

Продиференціювавши вираз (2.16) за параметрами $\boldsymbol{\theta}$ та $\boldsymbol{\beta}$, одержимо

$$\begin{aligned} \frac{\partial \lambda}{\partial \theta_n} &= r_n - \sum_{i=1}^L \sum_{k=1}^{m_i} k P_{ik}(\theta_n), & n &= 1, 2, \dots, N, \\ \frac{\partial \lambda}{\partial \beta_{ij}} &= -S_{ij} + \sum_{n=1}^N \sum_{k=j}^{m_i} P_{ik}(\theta_n), & i &= 1, 2, \dots, L, j = 1, 2, \dots, m_i. \end{aligned} \quad (2.17)$$

Для знаходження точки максимуму функції λ треба знайти такі значення параметрів θ та β , при яких похідні у (2.17) дорівнюють нулеві.

Інакше кажучи, отримуємо систему $N + \sum_{i=1}^L m_i$ рівнянь відносно $N + \sum_{i=1}^L m_i$

змінних

$$\begin{cases} r_n - \sum_{i=1}^L \sum_{k=1}^{m_i} k P_{ik}(\theta_n) = 0, & n = 1, 2, \dots, N, \\ -S_{ij} + \sum_{n=1}^N \sum_{k=j}^{m_i} P_{ik}(\theta_n) = 0, & i = 1, 2, \dots, L, j = 1, 2, \dots, m_i, \end{cases} \quad (2.18)$$

розв'язком якої є шукана точка максимуму.

Розв'язуватимемо систему (2.18) за допомогою методу Ньютона-Рафсона [195]. З огляду на те, що ліві частини рівнянь системи (2.18) є першими похідними функції λ , для застосування методу Ньютона-Рафсона знайдемо її другі похідні:

$$\begin{aligned} \frac{\partial^2 \lambda}{\partial \theta_n^2} &= - \sum_{i=1}^L \left[\sum_{k=1}^{m_i} k^2 P_{ik}(\theta_n) - \left(\sum_{k=1}^{m_i} k P_{ik}(\theta_n) \right)^2 \right], & n = 1, 2, \dots, N, \\ \frac{\partial^2 \lambda}{\partial \beta_{ij}^2} &= - \sum_{n=1}^N \left[\sum_{k=j}^{m_i} P_{ik}(\theta_n) - \left(\sum_{k=j}^{m_i} P_{ik}(\theta_n) \right)^2 \right], & i = 1, 2, \dots, L, j = 1, 2, \dots \end{aligned} \quad (2.19)$$

Застосовуючи (2.17) та (2.19), запишемо ітераційні формули для оцінювання значень параметрів θ та β :

$$\begin{aligned} \theta_n^{(t+1)} &= \theta_n^{(t)} - \frac{\frac{\partial \lambda}{\partial \theta_n}}{\frac{\partial^2 \lambda}{\partial \theta_n^2}} = \theta_n^{(t)} - \frac{r_n - \sum_{i=1}^L \sum_{k=1}^{m_i} k P_{ik}(\theta_n)}{- \sum_{i=1}^L \left[\sum_{k=1}^{m_i} k^2 P_{ik}(\theta_n) - \left(\sum_{k=1}^{m_i} k P_{ik}(\theta_n) \right)^2 \right]}, & n = 1, 2, \dots, N \\ \beta_{ij}^{(t+1)} &= \beta_{ij}^{(t)} - \frac{\frac{\partial \lambda}{\partial \beta_{ij}}}{\frac{\partial^2 \lambda}{\partial \beta_{ij}^2}} = \beta_{ij}^{(t)} - \frac{-S_{ij} + \sum_{n=1}^N \sum_{k=j}^{m_i} P_{ik}(\theta_n)}{- \sum_{n=1}^N \left[\sum_{k=j}^{m_i} P_{ik}(\theta_n) - \left(\sum_{k=j}^{m_i} P_{ik}(\theta_n) \right)^2 \right]}, & \begin{aligned} i &= 1, 2, \dots, L, \\ j &= 1, 2, \dots, m_i \end{aligned} \end{aligned} \quad (2.20)$$

де $\theta_n^{(t)}$ та $\beta_{ij}^{(t)}$ — значення параметрів θ_n та β_{ij} , отримані на t -му кроці ітераційного процесу.

Запропонований у роботі [201] алгоритм оцінювання параметрів полягає в такому. Спочатку здійснюється 10 кроків зі знаходження θ_n , потім 10 кроків зі знаходження β_{ij} , і так далі. В якості умови завершення ітераційного процесу після t ітерацій визначимо виконання нерівності

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

де ε — заздалегідь задана точність обчислення.

Основні проблеми практичного застосування моделі та шляхи їхнього розв’язання. Практичне застосування розробленого методу оцінювання параметрів моделі Раша-Мастерса виявило такі проблеми:

- вибір початкових значень;
- зсув оцінок параметрів;
- наявність тривіальних рівнів.

Вибір початкових значень. З огляду на те, що система рівнянь (2.18) розв’язується чисельно, постає питання вибору початкових значень параметрів, що оцінюються. Доступні джерела з теорії моделювання та параметризації педагогічних тестів не містять жодних рекомендацій щодо вибору початкового наближення значень параметрів політомічної моделі Раша-Мастерса, окрім наведених у роботі [192], проте вони містять очевидні помилки. Для вибору початкових значень скористаємось рекомендаціями щодо визначення початкових значень параметрів дихотомічної моделі Раша.

Якщо i -е завдання є дихотомічним, то $m_i = 1$. Відповідно, для i -го дихотомічного завдання визначено лише один параметр складності β_i , що відповідає першому і єдиному його рівню. Якщо іспитник здобуває за таке

завдання 0 балів, то кажуть, що він відповів на нього неправильно, якщо 1 бал – правильно. Розглянемо вищезгадані рекомендації [162]:

$$\theta_n^{(0)} = \ln \frac{p_n}{q_n}, n=1,2,\dots,N, \beta_i^{(0)} = \ln \frac{Q_i}{P_i}, i=1,2,\dots,L, \quad (2.21)$$

де p_n і q_n – долі відповідно правильних і неправильних відповідей n -го іспитника, а P_i та Q_i – долі відповідно правильних і неправильних відповідей на i -е завдання:

$$p_n = \frac{r_n}{L}, q_n = 1 - p_n, P_i = \frac{R_i}{N}, Q_i = 1 - P_i$$

(R_i – кількість правильних відповідей на i -е завдання). Перепишемо (2.21) у вигляді

$$\theta_n^{(0)} = \ln \frac{r_n}{L \left(1 - \frac{r_n}{L}\right)} = \ln \frac{r_n}{L - r_n}, \beta_i^{(0)} = \ln \frac{1 - \frac{R_i}{N}}{\frac{R_i}{N}} = \ln \frac{N - R_i}{R_i}, \quad (2.22)$$

$n=1,2,\dots,N, i=1,2,\dots,L$

і узагальнимо (2.22) на випадок політомічних завдань:

$$\theta_n^{(0)} = \ln \frac{r_n}{M - r_n}, n=1,2,\dots,N, \beta_{ij}^{(0)} = \ln \frac{N - S_{ij}}{S_{ij}}, i=1,2,\dots,L, j=1,2,\dots,m_i,$$

де M – максимальний бал за розглядувану тестову роботу (очевидно, якщо робота містить лише дихотомічні завдання, то $M = L$, а $S_{i1} = R_i, i=1,2,\dots,L$).

Впроваджений таким чином метод визначення початкових значень параметрів моделі Раша-Мастерса у більшості випадків призвів до збіжності ітераційного процесу знаходження їхніх оцінок. Незважаючи на це, питання збіжності обраного ітераційного методу потребує більш детального дослідження.

Зсув оцінок параметрів. Система (2.18) має нескінченну кількість розв'язків в тому сенсі, що вони визначаються з точністю до додавання константи. Це твердження справджується при практичній реалізації процедури оцінювання параметрів моделі: при різних початкових значеннях отримуємо розв'язки, віддалені один від одного на сталу величину.

Переконаємося, що це дійсно так. Нехай $\theta = (\theta_1, \theta_2, \dots, \theta_N)$ і $\beta = (\beta_{11}, \dots, \beta_{1m_1}, \dots, \beta_{L1}, \dots, \beta_{Lm_L})$ – розв’язок системи рівнянь (2.18). Розглянемо тепер параметри $\theta^* = (\theta_1 + C, \theta_2 + C, \dots, \theta_N + C)$ і $\beta^* = (\beta_{11} + C, \dots, \beta_{1m_1} + C, \dots, \beta_{L1} + C, \dots, \beta_{Lm_L} + C)$, $C = \text{const}$. Підставивши θ^* і β^* у (2.18) замість θ і β , одержимо такий самий результат. Це явище має місце в силу того, що система (2.18) містить параметри θ і β лише в якості параметрів функцій імовірності, що визначаються за формулою (2.10). У виразі (2.10) параметри θ і β згруповані у вигляді різниць $(\theta_n - \beta_{ij}), n = 1, 2, \dots, N, i = 1, 2, \dots, L, j = 1, 2, \dots, m_i$, завдяки чому константи C у (2.10) взаємно знищуються. Таким чином, розв’язок системи рівнянь (2.18) дійсно визначається з точністю до додавання сталої величини.

У зв’язку з цим, розв’язавши систему (2.18), можемо дізнатися лише взаємне розташування параметрів моделі на числовій осі, але не їхні істинні значення. В роботах [192] та [159] наведено рекомендації до знаходження істинних значень параметрів. Сутність рекомендацій полягає в тому, що середнє арифметичне значень параметрів підготовленості повинно дорівнювати нулеві. Таким чином, по завершенні процедури оцінювання параметрів обчислюється середнє арифметичне $\tilde{\beta}$ оцінок параметрів підготованості, після чого здійснюється зсув усіх значень θ та β ліворуч на величину $\tilde{\beta}$ [192].

Відмінність підходу до розв’язання цієї проблеми, описаного у роботі [159], полягає в тому, що пропонується обчислювати $\tilde{\beta}$ після кожного кроку ітерації. В цьому немає сенсу: зсув усіх значень параметрів на однакову сталу величину до завершення ітераційного процесу не вплине на нього, оскільки дані сталі величини взаємно знищуються при обчисленні ймовірностей (2.10). З огляду на це, більш раціональним є одноразове обчислення величини зсуву, аніж багаторазове виконання цієї операції до завершення ітераційного процесу.

Тривіальні рівні тестових завдань. Розглянемо матрицю $\mathbf{R} = (R_{ij}), i=1,2,\dots,L, j=0,1,\dots,\max_i m_i$. Елемент R_{ij} цієї матриці дорівнює кількості іспитників, які в i -му завданні розглядуваної тестової роботи здобули результат j . Очевидно, що в такому випадку маємо:

$$R_{ij} = S_{ij} - S_{i(j+1)}, j=0,1,\dots,m_i-1, R_{im_i} = S_{im_i}, i=1,2,\dots,L. \quad (2.23)$$

Для визначеності покладемо $R_{kj} = 0$ для $k=1,2,\dots,L, m_k < j \leq \max_i m_i, i=1,2,\dots,L$. Ці елементи не впливають на подальший аналіз результатів тестової роботи.

Припустимо, що існують такі числа i та j , $i=1,2,\dots,L, j=1,2,\dots,m_i$, що $R_{ij} = 0$, тобто, в розглядуваній тестовій роботі жоден іспитник не отримав j балів за i -е завдання. В такому випадку будемо казати, що j -й рівень i -го завдання є тривіальним.

У практичній реалізації оцінювання параметрів моделі Раша-Мастерса, наявність тривіальних рівнів унеможлиблює завершення цієї процедури. Розглянемо причини цього явища.

Припустимо, що кожен іспитник отримав за i -е завдання тестової роботи щонайменше 1 бал, тобто $R_{i0} = 0$ та $S_{i1} = N$. Розглянемо рівняння системи (2.18), яке в такому випадку відповідає першому рівневі i -го завдання:

$$-N + \sum_{n=1}^N \sum_{k=1}^{m_i} P_{ik}(\theta_n) = 0. \quad (2.24)$$

Оскільки за побудовою та за властивостями ймовірності $\sum_{k=0}^{m_i} P_{ik}(\theta_n) = 1$,

то $\sum_{k=j}^{m_i} P_{ik}(\theta_n) \leq 1, j=1,2,\dots,L, j=1,2,\dots,m_i, n=1,2,\dots,N$. З огляду на це, для

виконання рівності (2.24) повинна виконуватися умова

$\sum_{k=1}^{m_i} P_{ik}(\theta_n) = 1, n=1,2,\dots,N$. Але в такому разі у формулі (2.19) для β_{i1} при

досягненні в процесі ітерацій доброго наближення до розв'язку системи (2.18) дорівнюватимуть нулеві як чисельник, так і знаменник, що унеможливилює подальший процес обчислення.

Тепер припустимо, що жоден іспитник у i -му завданні тестової роботи не досягнув рівня m_i , тобто $R_{im_i} = S_{im_i} = 0$. Розглянемо рівняння системи (2.18), яке в такому випадку відповідає максимальному рівневі i -го завдання:

$$\sum_{n=1}^N P_{im_i}(\theta_n) = 0.$$

Звідси за властивостями ймовірності $P_{im_i}(\theta_n) = 0, n = 1, 2, \dots, N$. В такому разі при досягненні в процесі ітерацій наближення із заданою точністю до розв'язку системи (2.18) нулеві будуть дорівнювати чисельник і знаменник ітераційної формули (2.20) для β_{im_i} , що, як і в попередньому випадку унеможливилює подальший процес обчислення.

Розглянемо тепер загальний випадок: $R_{ij} = 0, i = 1, 2, \dots, L, j = 1, 2, \dots, m_i - 1$. Тоді із рівності (2.23) випливає, що $S_{ij} = S_{i(j+1)}$. Це означає, що ліві частини рівнянь системи (2.18), що відповідають j -му та $(j+1)$ -му рівням i -го завдання, співпадають, тобто має місце рівність

$$\sum_{n=1}^N \sum_{k=j}^{m_i} P_{ik}(\theta_n) = \sum_{n=1}^N \sum_{k=j+1}^{m_i} P_{ik}(\theta_n),$$

яка після спрощення набуває вигляду

$$\sum_{n=1}^N P_{ij}(\theta_n) = 0.$$

За властивостями ймовірності робимо висновок, що $P_{ij}(\theta_n) = 0, n = 1, 2, \dots, N$. За визначенням ймовірності у моделі Раша-Мастерса

(2.10) це означає, що $\exp \sum_{t=0}^j (\theta_n - \beta_{it}) = 0, n = 1, 2, \dots, N$. Звідси випливає, що для

задоволення умов системи повинна виконуватись рівність

$\sum_{t=0}^j (\theta_n - \beta_{it}) = -\infty, n = 1, 2, \dots, N$. Внаслідок цього процедура знаходження розв'язку системи (2.18) за допомогою ітерацій (2.19) призводить до нескінченного віддалення значень параметрів θ_n та β_{it} , тобто спостерігається розбіжність ітераційного методу.

Найпростішим методом уникнення подібних ситуацій є об'єднання тривіального рівня із наступним. Ця процедура здійснюється перед оцінюванням значень параметрів моделі і фактично змінює структуру тестової роботи.

Припустимо, що j -й рівень i -го завдання є тривіальним, причому $j < m_i$. Об'єднуючи його з $(j + 1)$ -м рівнем i -го завдання, необхідно потім змінити нумерацію подальших рівнів, а саме, зменшити відповідні їм номери на одиницю. Таким чином, зменшиться на одиницю і максимальна кількість балів за i -е завдання – m_i . Якщо ж тривіальним є m_i -й рівень i -го завдання, то здійснюється його об'єднання з $(m_i - 1)$ -м рівнем того ж завдання, відповідно, максимальний бал за це завдання зменшується на одиницю, а нумерація рівнів не змінюється.

2.3. Проектування системи автоматизованого аналізу якості тестових завдань

Описані та дібрані вище методи КТТ та IRT було закладено та реалізовано в СААЯТЗ. СААЯТЗ є авторською програмою, яка слугує засобом комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики.

Основними задачами СААЯТЗ є:

- 1) статистичний аналіз результатів тестування методами КТТ та IRT;
- 2) формування та обслуговування бази каліброваних завдань.

Функції системи реалізовано у СААЯТЗ за допомогою окремих модулів, які схематично представлено на Рис. 2.12. Опишемо їх детальніше. *Головний управляючий модуль.* Забезпечує формування інтерфейсу та

викликає інші модулі. Статистичний аналіз тестів забезпечують групи модулів КТТ- та IRT-методів. До групи модулів КТТ-аналізу входять:

- *Модуль первинного аналізу вибірки результатів тестування.* Дає змогу обчислити математичне сподівання μ , дисперсію σ^2 , критерій узгодженості χ^2 , середнє вибіркове Sx , моду Mo , медіану Me , асиметрію As , перевірити гіпотезу про нормальний розподіл вибірки, а також вивести на екран гістограму та криву теоретичного нормального розподілу.

- *Модуль оцінювання валідності тесту.* Дає змогу оцінити, як результати тестування співвідносяться з іншими незалежними експертними оцінками. Порівнюючи оцінки, модуль обчислює коефіцієнт критеріальної валідності та робить висновок про співвідношення оцінок.

- *Модуль оцінювання надійності тесту.* Дає змогу обчислити коефіцієнт надійності тесту r , знайдений на основі методу розщеплення, і коефіцієнт надійності Кронбаха α та отримати висновки про надійність тесту на основі обчислених коефіцієнтів

- *Модуль аналізу валідності окремих тестових завдань.* Дає змогу обчислити показники валідності окремих завдань (значення індексу розрізняльної здатності дихотомічних завдань D_j , коефіцієнт точкової бісеріальної кореляції та бісеріального коефіцієнту кореляції між дихотомічним завданням та індивідуальними балами іспитників і коефіцієнт кореляції Пірсона між політомічним завданням та індивідуальними балами іспитників $\rho_{j\cdot}$) та зробити висновки щодо якості завдання та подальших дій із ним.

- *Модуль аналізу кореляційної матриці завдань.* Дає змогу обчислити кореляційну матрицю завдань та отримати інформацію про можливі некоректність та дублювання завдань

Група модулів IRT-аналізу включає:

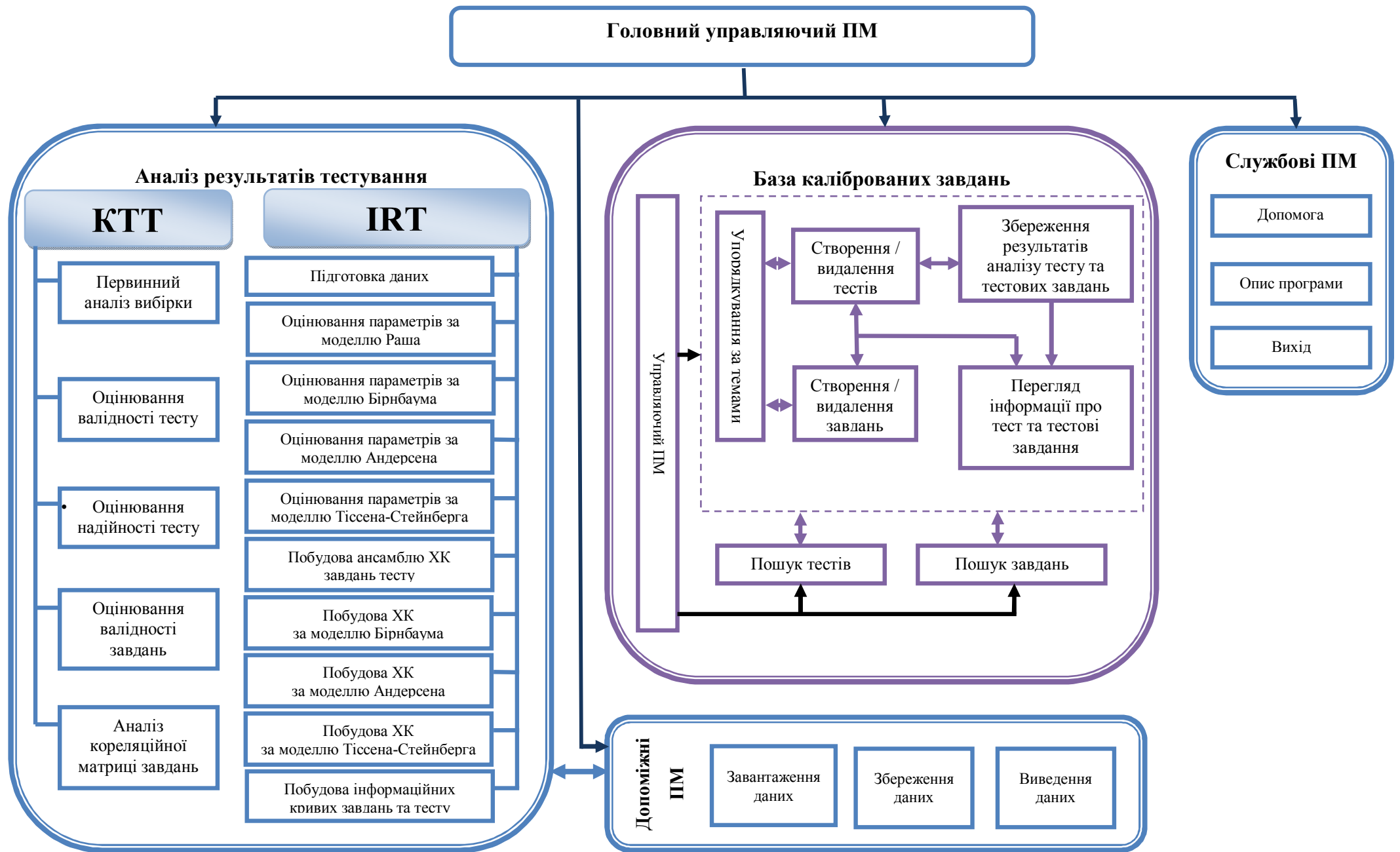


Рис. 2.12. Схема програмних модулів СААЯТЗ

- *Модуль підготовки та завантаження даних для IRT-аналізу.* Забезпечує попередню підготовку даних до IRT-аналізу та завантаження їх у систему для подальшої роботи. Даний модуль є дуже важливим, оскільки для IRT-аналізу політомічних завдань недостатньо результатів тестування. Для нього необхідною також є інформація про кількість рівнів у тестовому завданні та кількість балів за досягнення j –го рівня. Даний модуль було розроблено для зручного та швидкого внесення цих даних до результатів тестування

- *Модуль оцінювання латентних параметрів за моделлю Раша.* Дає змогу оцінити параметри складності дихотомічних завдань тесту.

- *Модуль оцінювання латентних параметрів за моделлю Бірнбаума.* Дає змогу оцінити параметри складності та параметри диференціюючої спроможності дихотомічних завдань тесту.

- *Модуль оцінювання латентних параметрів за моделлю Андерсена.* Дає змогу оцінити параметри складності рівнів завдань тесту та параметри підготовленості іспитників.

- *Модуль оцінювання латентних параметрів за моделлю Тіссена-Стейнберга.* Дає змогу оцінити параметри складності варіантів відповіді завдань множинного вигляду, їх диференціюючу спроможність, параметри ймовірності вибору варіанта відповіді навмання та параметри іспитників.

- *Модуль побудови ансамблю характеристичних кривих тесту.* Дає змогу побудувати ансамбль характеристичних кривих тесту.

- *Модуль побудови характеристичних кривих завдань тесту за моделлю Бірнбаума.* Дає змогу побудувати окремі характеристичні криві дихотомічних завдань тесту за моделлю Бірнбаума.

- *Модуль побудови характеристичних кривих рівнів завдань тесту за моделлю Андерсена.* Дає змогу побудувати характеристичні криві рівнів політомічних завдань тесту за моделлю Андерсена.

- *Модуль побудови характеристичних кривих варіантів відповіді завдань тесту за моделлю Тіссена-Стейнберга.* Дає змогу побудувати

характеристичні криві рівнів завдань множинного вигляду за моделлю Тіссена-Стейнберга.

- *Модуль побудови інформаційних кривих завдань та тесту у цілому.* Дає змогу побудувати окремі інформаційні функції завдань тесту за моделями Бірнбаума, Андерсена та Тіссена-Стейнберга та інформаційну функцію тесту за моделлю Андерсена.

Завантаження, виведення та збереження даних при статистичному аналізі забезпечено відповідними **допоміжними модулями**.

Проектування СААЯТЗ на основі вищезазначених модулів дає змогу викладачеві обирати гнучкий алгоритм оцінювання якості тестів з вищої математики.

Невід’ємною складовою СААЯТЗ є **база каліброваних завдань**, необхідна для роботи викладача з тестами з вищої математики, а саме для їх формування та редагування. Формування та обслуговування бази каліброваних завдань у СААЯТЗ забезпечено окремою підпрограмою (Рис. 2.12), реалізованою також за допомогою окремих модулів.

Головний управляючий модуль бази забезпечує формування інтерфейсу та викликає інші модулі, а саме:

- *Модуль створення та видалення тем.* Дає змогу створювати теми тестування, за якими упорядковуються тести та окремі тестові завдання.

- *Модулі створення та видалення тестів.* Дають змогу створити та видалити тест. Модуль створення тесту передбачає віднесення тесту до однієї зі створених тем, привласнення йому ідентифікаційного номеру та наповнення тесту раніше створеними тестовими завданнями.

- *Модулі створення та видалення тестових завдань.* Дають змогу відповідно створити та видалити тестове завдання. Модуль створення тестового завдання передбачає віднесення завдання до однієї зі створених тем, зазначення форми завдання, привласнення ідентифікаційного номеру завдання та написання тексту самого завдання.

- *Модуль перегляду інформації про тест та окремі тестові завдання.* Модуль перегляду інформації про тест дає змогу переглянути наповнення тексту, зміст його тестових завдань. Модуль перегляду інформації про тестове завдання дає змогу переглянути його зміст та інформацію, введену при його створенні.

- *Модуль збереження результатів аналізу тесту та тестових завдань.* Дає змогу обчислити та зберегти числові та графічні характеристики тесту та тестових завдань, якщо це не було виконано раніше.

- *Модуль пошуку тестів за заданими параметрами.* Дає змогу виконати пошук за ідентифікаційним номером тесту та назвою теми, до якої він відноситься.

- *Модуль пошуку тестових завдань за заданими параметрами.* Дає змогу виконати пошук за ідентифікаційним номером тесту та назвою теми, до якої він відноситься за такими критеріями: темою завдання, показниками валідності завдання, складністю (для дихотомічного завдання) та формою завдання. Даний модуль є особливо важливим, оскільки призначений для спрощення процедури пошуку завдань з конкретними характеристиками при створенні тесту.

Для зручної роботи з програмою створено **службові модулі СААЯТЗ**: *модуль допомоги*, який пояснює спосіб правильного налаштування програми та з окремими її роботи з окремими її складовими; *модуль опису програми*, який відображає інформацію про СААЯТЗ у цілому та *модуль виходу із системи*

Програмна реалізація СААЯТЗ. Для побудови СААЯТЗ було використано Windows Form додаток. Додаток являє собою екранний об'єкт, зазвичай прямокутник, який можна застосовувати для надання інформації викладачеві і для опрацювання введеної інформації викладачем. Форми можуть мати вигляд стандартного діалогового вікна, багатодокументного інтерфейсу (MDI) або поверхні для відображення графічної інформації.

Для реалізації програмного продукту СААЯТЗ обрано сучасну мову програмування C#. Мова C# як засіб програмування має ряд безперечних переваг. Вона добре організована, більшість її конструкцій логічні і зручні. Розвинені засоби діагностики і редагування коду роблять процес програмування ефективним. Потужна бібліотека класів .NET бере на себе масу рутинних операцій, що дає можливість вирішувати завдання, використовуючи готові «будівельні блоки». Оцінювання латентних параметрів реалізовано у середовищі MATLAB і за допомогою набору динамічних бібліотек MATLAB Compiler Runtime інтегровано в середовище «Visual Studio C++».

СААЯТЗ є портативною програмою, для коректної роботи якої викладач має встановити її на свій ПК за допомогою програми інсталятора. Установка включає в себе розміщення всіх необхідних програмі файлів у відповідних місцях файлової системи, а також зміну та створення конфігураційних файлів.

Для побудови Бази каліброваних завдань використана мова програмування PowerBuilder (SAP) – інструмент, який можна назвати класикою систем, призначених для роботи з SQL СУБД. PowerBuilder є успішним засобом розроблення клієнт-серверних додатків. Основою його успіху стала вдала технологія роботи з даними через механізм DataWindow, а також відкритість і масштабованість (дає змогу працювати з таким широким спектром джерел даних – від DBF-файлів до корпоративних БД масштабу Sybase SQL Server або Oracle).

Висновки до розділу 2

Проведений аналіз сучасних досліджень щодо методів КТТ та моделей IRT для оцінювання якості тестів з вищої математики, питання аналізу та підвищення ефективності тестів методами IRT, проблеми оцінювання латентних параметрів моделей IRT дає змоги дійти таких висновків:

1. Гіпотеза дослідження ґрунтується на припущенні, що педагогічно виважене впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики сприятиме розвитку у викладачів ЗВО відповідної компетентності за умови реалізації розробленої технології розвитку компетентності та покращенню контролю знань студентів.

2. Обґрунтовано необхідність поєднання методів КТТ та IRT для оцінювання якості тестів з вищої математики. Визначено основні методи аналізу якості тестів з вищої математики в рамках КТТ, які передбачають як аналіз окремих тестових завдань, так і тесту у цілому. Аналіз тесту у цілому полягає у такому: аналізі розподілу вибірки результатів; оцінюванні надійності тесту; оцінюванні критеріальної валідності тесту. Аналіз тестових завдань передбачає: аналіз кореляційної матриці завдань; оцінювання валідності окремих завдань; аналіз дистракторів завдань множинного вибору.

3. Визначено моделі IRT для оцінювання якості тестів з вищої математики: Раша та Бірнбаума (для дихотомічних завдань), Андерсенена для політомічних завдань та Тіссена-Стейнберга для завдань множинного вигляду. IRT-аналіз тестових завдань і тесту у цілому проводиться із використанням визначених моделей і передбачає: оцінювання значень латентних параметрів використовуваних IRT-моделей; аналіз ансамблю характеристичних кривих завдань тесту; аналіз характеристичних кривих рівнів політомічних тестових завдань; аналіз характеристичних кривих дистракторів завдань множинного вибору.

4. Досліджено питання оцінювання та підвищення ефективності тестових завдань в рамках IRT із використанням інформаційної функції. Інформаційна функція показує, наскільки ефективно дане завдання тесту оцінює іспитника з певним значенням параметра підготовленості. Наведено алгоритми обчислення та аналізу інформаційної функції в залежності від форми тестового завдання і від обраної IRT-моделі.

5. Досліджено питання оцінювання латентних параметрів IRT-моделей. Для оцінювання латентних параметрів обрано метод максимальної

вірогідності. Розглянуто його застосування на прикладі моделі Раша-Мастерса. Вирішено такі проблеми практичного оцінювання параметрів IRT-моделей: вибір початкових значень; зсув оцінок параметрів; наявність тривіальних рівнів політомічних тестових завдань.

6. Спроековано авторську систему автоматизованого аналізу якості тестових завдань. Функції системи реалізовано у СААЯТЗ за допомогою окремих модулів. Основні задачі СААЯТЗ: статистичний аналіз результатів тестування методами КТТ та IRT; формування та обслуговування бази каліброваних завдань. Формування бази каліброваних завдань є важливим результатом роботи системи. В ній зберігаються окремі тестові завдання, систематизовані за темами, призначенням (контрольна, іспит) та формою тестового завдання. Для кожного завдання існує відповідний паспорт, який містить основні характеристики тестового завдання, оцінені системою. База каліброваних завдань є невід'ємною складовою, необхідною для роботи викладача з тестами з вищої математики, а саме для формування та редагування тестів. Робота системи з обслуговування бази каліброваних тестових завдань передбачає: редагування та оновлення бази; підбір та перегляд завдань за заданими параметрами; створення нових тестів; розподіл тестів за темами; пошук та перегляд тестів.

Результати другого розділу дисертаційного дослідження висвітлено у наукових роботах автора [44, 154, 58].

РОЗДІЛ III. ОБҐРУНТУВАННЯ КОМП'ЮТЕРНО ОРІЄНТОВАНОЇ МЕТОДИКИ ОЦІНЮВАННЯ ЯКОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ

У розділі описано загальну схему комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики викладачами ЗВО та функціональні можливості авторської системи автоматизованого аналізу якості тестових завдань, яка є засобом розробленої методики. Розглянуто практичне застосування методики та наведено особливості роботи викладача з тестами для оцінювання знань студентів. Визначено компетентність викладача щодо оцінювання якості тестів з вищої математики, критерії та рівні розвитку даної компетентності. Розроблено модель та технологію розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики.

3.1. Розроблення тестів з вищої математики

Якість педагогічного тесту з вищої математики задається ще на етапі його розроблення та вимагає вирішення таких питань:

- визначення підходу до розроблення тестів;
- оцінювання змістової валідності.

Одним із ключових питань при створенні тесту є питання **визначення підходу до розроблення тестів**. Традиційно виділяють критеріально-орієнтований та нормативно-орієнтований підходи до розроблення тестів. Перед розробниками комп'ютерних тестів з вищої математики, які використовують для оцінювання знань студентів, постає задача визначення місця створених тестів серед критеріально-орієнтованих (КОТ) та нормативно-орієнтованих тестів (НОТ). Це питання є важливим з точки зору визначення мети створення, рівня деталізації області змісту тесту, методів відбору тестових завдань, методів аналізу тестових завдань і тесту у цілому, принципу інтерпретації результатів тестування.

Основні характеристики критеріально- та нормативно-орієнтованого підходів до розроблення комп'ютерних тестів розглянуті в роботах В. С. Ким [91], Л. О. Кухар [104], Л. Крокер та Дж. Алгіни [99], М. Б. Чельшкової [162] та ін. Проаналізувавши ці роботи, можна зробити висновок, що, з одного боку, не можна використовувати один і той же тест в якості нормативно-орієнтованого і в якості критеріально-орієнтованого, з іншого – інтерпретація результатів тесту по відношенню до рівня навчальних досягнень не знімає можливості диференціації індивідуальних відмінностей студентів, що спонукає поєднати нормативно-орієнтований та критеріально-орієнтований підходи. Таке поєднання в силу існуючих значних відмінностей між підходами значно ускладнює процес розроблення тесту. Проблема об'єднання двох підходів до розроблення комп'ютерних тестів залишається дискусійною і тому є актуальною.

Основні характеристики критеріально- та нормативно-орієнтованих тестів. В залежності від цілей створення та інтерпретації результатів до розроблення підсумкових комп'ютерних тестів застосовують або критеріально-, або нормативно-орієнтований підходи. В рамках нормативно-орієнтованого підходу розробляються тести для порівняння іспитників за рівнем навчальних досягнень. При критеріально-орієнтованому підході результати іспитників інтерпретуються по відношенню до степені досягнення заданого рівня знань, умінь та навичок. При цьому важливішим є те, що може виконати студент і що він знає, а не те, як він виглядає на фоні інших.

Порівняємо КОТ та НОТ за такими характеристиками.

Мета застосування. Мета застосування КОТ – визначити степінь засвоєння іспитником певного розділу в заданій предметній області. Мета застосування НОТ – ранжування іспитників за рівнем навчальних досягнень.

Рівень деталізації області змісту. Для НОТ рівень деталізації області змісту – несуттєвий. В НОТ включають завдання на перевірку найбільш значущих елементів змісту навчального матеріалу. Для цих тестів набагато важливіше отримати варіативні тестові завдання. Рівень деталізації області

змісту КОТ докладний. Автори тесту розробляють специфікацію (план) тесту, що включає всі елементи змісту. Потім по цій специфікації розробляються завдання. Критеріально-орієнтований підхід оцінювання знань студентів вимагає стовідсоткового охоплення змісту дисципліни. Це призводить до того, що підсумкові тести містять занадто багато завдань (50-150 завдань). Виконання студентами таких тестів при одноразовому пред'явленні просто неможливо.

Відсоток студентів, які правильно виконали майже всі завдання тесту, варіативність балів студентів. Для КОТ 80-90 % іспитників виконують правильно майже всі завдання тесту. При цьому має місце дуже низька варіативність балів студентів. Для НОТ 5-10 % іспитників виконують правильно майже всі завдання тесту, має місце дуже висока варіативність балів студентів.

Відбір тестових завдань. В КОТ завдання можуть мати різні значення складності, але обов'язково відповідати специфікації та елементам змісту, що перевіряється. В НОТ основна частина завдань має складність з високою диференціюючою спроможністю. В НОТ не допускається змістова варіативність – всі іспитники повинні виконувати одні й ті ж завдання.

Розподіл індивідуальних балів іспитників. Розподіл індивідуальних балів іспитників має різний характер для обох видів тестів. Для НОТ крива індивідуальних балів іспитників симетрична і має розподіл, близький до нормального. У разі КОТ ця крива несиметрична і зазвичай зсунута в область високих індивідуальних балів.

Аналіз тесту. Надійність КОТ оцінюється стабільністю отримання «заліку» або «незаліку» при дворазовому тестуванні. Надійність НОТ оцінюється або шляхом знаходження кореляції між результатами двох тестувань, або методом розщеплення тесту на дві половини при одноразовому тестуванні.

При аналізі КОТ Особлива увага приділяється змістовій та критеріальній валідностям. Для НОТ наряду зі змістовою валідністю особлива увага приділяється високим показникам прогностичної валідності.

Аналіз тестових завдань. Аналіз тестових завдань КОТ і НОТ здійснюється на підставі методів КТТ та ІРТ.

Сфера застосування. КОТ, в основному, застосовують для оцінювання підсумкового рівня навчальних досягнень, оцінювання ефективності програми навчання, НОТ – для конкурсного відбору (вступні іспити до вузів).

Інтерпретація результатів тестування. Інтерпретація результатів НОТ дає можливість визначити місце даного студента щодо групи іспитників. Відомості ж про те, як засвоєні ті чи інші розділи, елементи змісту дисципліни, НОТ дає мало. Результати КОТ інтерпретуються з точки зору повноти засвоєння змісту навчального матеріалу.

Очевидно, що КОТ і НОТ суттєво відрізняються один від одного. Тому при створенні тестів з вищої математики необхідно вирішити питання вибору підходу до розроблення тестів.

Основні характеристики тестів з вищої математики. Комп'ютерне тестування проводиться у вигляді тестових модульних та залікових контрольних робіт. Модульні і залікові контрольні тести можна віднести відповідно до проміжних та підсумкових тестів. Проміжне та підсумкове тестування проводиться по закінченню певного етапу навчання і охоплює достатньо суттєвий обсяг навчального матеріалу. Модульні і залікові контрольні тести включають завдання на перевірку знань студентів щодо найважливіших елементів змісту, сформованості у них необхідних навичок.

Проведення проміжних та підсумкових тестів з вищої математики переслідує дві мети: визначення ступеня засвоєння студентом навчального матеріалу та диференціювання іспитників за рівнем навчальних досягнень. Відповідно до цілей застосування до цих тестів можна використовувати як критеріально-, так і нормативно-орієнтований підходи.

Мета створення НОТ відповідає меті рейтингової системи оцінювання у ЗВО, яка є одним із засобів забезпечення змагальності та здорової конкуренції в навчанні, підвищенні мотивації студентів до активного, свідомого навчання, систематичної самостійної роботи протягом семестру та відповідальності за результати навчальної діяльності.

Результати модульних та залікових контрольних робіт є формуючим фактором семестрової атестації. За результатами семестрової атестації деканати складають академічні рейтинги студентів. Академічний рейтинг використовується для [131]:

- ✓ забезпечення об'єктивних критеріїв та умов конкурсного відбору студентів навчання за магістерською програмою;
- ✓ ранжування студентів університету, факультету, базової бакалаврської програми (професійного спрямування), окремої навчальної групи за рівнем навчальних досягнень та обґрунтоване надання різних пільг (направлення на навчання або на практику за кордон тощо);
- ✓ призначення іменних стипендій: державних, університету, факультету (інституту);
- ✓ конкурсного відбору студентів, які бажають проходити військову підготовки за програмою підготовки офіцерів запасу;
- ✓ пріоритетного надання путівок для оздоровлення в канікулярний період;
- ✓ надання інших пільг.

При цьому КОТ дають змогу порівнювати знання, вміння та навички студентів щодо відповідності вимогам, визначених у державних освітніх стандартах, відбирати студентів, які досягли планованого рівня навчальних досягнень, що також є дуже важливим при оцінюванні знань іспитників.

Для тестів з вищої математики важливою є охоплення всього змісту навчального матеріалу. Тому обов'язковою є розроблення специфікації тесту. Кількість завдань, які студент може виконати при одноразовому пред'явленні не повинна перевищувати 30. У зв'язку з цим на основі багаторічного досвіду

викладачів виділяються найсуттєвіші та найінформативніші завдання, зміст яких близький до змісту завдань відповідних традиційних контрольних робіт.

Для тестів з вищої математики приблизно 10% іспитників мають правильно виконувати майже всі завдання тесту. При цьому висока варіативність індивідуальних балів має місце при великій кількості іспитників. В підсумкових тестах допустимою є варіативність тестових завдань.

При інтерпретації результатів тестів з вищої математики важливими є як можливість визначити місце даного студента щодо групи іспитників, так і можливість оцінити повноту засвоєння змісту навчальної дисципліни.

Тому згідно з тенденціями сучасного розвитку освіти та запитам ЗВО, разом з вищезазначеним, вказують на необхідність об'єднання критеріально-орієнтованого та нормативно-орієнтованого підходів при оцінюванні навчальних досягнень студентів. Саме таке об'єднання рекомендується при створенні тестів з вищої математики. Саме за такої реалізації інтерпретація результатів тесту по відношенню до певного змісту навчальної дисципліни не знімає можливості диференціації індивідуальних відмінностей студентів [162].

Іншим важливим питанням при розробці тестів з вищої математики є **питання оцінювання змістової валідності**, яке детально було розглянуто нами у роботі [56]. Змістова валідність визначається як характеристика репрезентативності змісту тесту по відношенню до запланованих для перевірки знань та умінь. Найбільше значення цей вид валідності має для тестів досягнень (тестів успішності, екзаменаційних тестів).

Оцінювання змістової валідності, у свою чергу, породжує задачу організації експертизи та вибору інструментарію для її проведення. Використання ІКТ для проведення експертизи дає можливість підвищити ефективність роботи експертів. Тому набувають значущості задачі вибору засобу ІКТ у якості інструменту проведення експертизи та розроблення технології оцінювання змістової валідності тесту з його використанням [56].

Аналізу валідності педагогічного тесту присвячені роботи з проблем педагогічних вимірювань А. Анастасі [10], В.С. Аванесова [1], К. Інгенкампа [83], А. С. Казарінова [86], І. О. Караєвої [89], Є. А. Михайлевича [118], А. І. Самиловського [137] та інших дослідників. Пошук адекватних засобів забезпечення змістовою валідністю тестів у тестології проводився В. А. Аванесовим [2], В. Ю. Биковим, Ю. М. Богачковим, Ю. О. Жуком [20], Грабар М.І. та Краснянською К.О. [35], А. Н. Майоровим [109], С. А. Сафонцевим [139], М. Б. Челишковою [162], Ю.М. Нейманом та В.А. Хлебніковим [124]. Автори одностайні в тому, що оцінювання змістової валідності тесту зазвичай проводиться незалежними експертами з великим стажем роботи, які не брали участь в розробці тесту. Експертиза якості змісту тесту обов'язково має включати аналіз якості змісту окремих тестових завдань. Аналізуючи роботи вищезазначених фахівців, можна зробити висновок, що методи забезпечення змістової валідності базуються на аналізі змісту завдань у термінах вимог до рівня навчальних досягнень іспитників або до цілей тестування; оцінюванні рівня значущості змісту завдань тесту; оцінюванні відповідності змісту тесту до його специфікації; виявленні невдалих завдань тесту; обчисленні відсотку правильно виконаних завдань гіпотетичною групою іспитників.

Серед методів проведення експертизи, описаних зокрема в роботі [126], можна виділити такі:

- очне опитування;
 - вільне інтерв'ю експертів;
 - анкетне опитування експертів;
 - «Мозковий штурм»;
- заочне опитування;
 - поштове анкетне опитування експертів;
 - метод дельфі.

Обрана нами дельфійська техніка полягає у багаторазовому поштовому анкетуванні однієї і тієї ж групи експертів. Після першого опитування

експертів і опрацювання його результатів, підсумки повідомляються учасникам експертної групи. Вони повинні або підтвердити свою точку зору, висловлену на попередньому етапі, і якщо вона значно відрізняється від думки більшості, розгорнуто її мотивувати, або змінити свою оцінку відповідно до думок більшості учасників. Потім знову обробляються висновки експертів, результати знову розсилаються експертам до тих пір, поки не будуть отримані узгоджені оцінки експертів [56].

Виділимо *задачі*, які необхідно вирішити для забезпечення високої *змістової валідності тесту*:

- розроблення специфікації тесту;
- відбір групи експертів, компетентних у змістовій області;
- оцінювання змістової валідності створеного тесту.

Розроблення специфікації тесту. Специфікація тесту в розгорнутій формі включає такі пункти:

1. Вступ, в якому вказано тему тесту, для кого і для чого він призначений, яким стандартам та програмам підготовки він відповідає, які навчальні матеріали будуть покладені в основу тесту.
2. ПІБ автора специфікації.
3. Дата створення специфікації.
4. Предмет (галузь), до якої відноситься тест.
5. Мета створення тесту.
6. Перелік можливих рішень (суджень), що можуть бути прийняті за результатами тестування
7. Мова тесту.
8. Загальна кількість завдань тесту. Кількість завдань тесту різних форм із зазначенням кількості варіантів відповідей завдань закритої форми.
9. Час на виконання тесту.
10. Форми проведення тестування (бланкова, комп'ютерна, змішана тощо)
11. Організаційні умови проведення тестування.

12. Момент проведення тестування (наприкінці вивчення теми, проміжний, підсумковий тощо)
13. Спосіб аналізу результатів тестування.
14. Форма та структура представлення результатів.
15. Зміст (перелік тем), що виноситься на тестування.
16. Детальний перелік та опис підтем тесту.
17. Матриця тесту, в якій наводиться відсоток тестових завдань, що мають бути включені до конкретної підтеми тесту.
18. Список вимог до рівня навчальних досягнень іспитників.
19. Відсоток вимог до рівня навчальних досягнень іспитників, охоплених специфікацією.

Відбір групи експертів, компетентних у змістовій області.

Оцінювання якості змісту тесту проводиться незалежними експертами, які не брали участь у розробці тесту. Кількість експертів повинна бути не менше трьох з кожного тесту. До експертизи розробники залучають найбільш досвідчених викладачів, які мають великий стаж роботи з тими студентами, для яких призначено тест. Основними процедурами процесу вибору експертів є складання списку можливих експертів та вибір з них експертної комісії у відповідності до компетентності кандидатів. Детально ці процедури описані в роботі [126]. Для кількісного оцінювання рівня компетентності використовується коефіцієнт компетентності, з урахуванням якого зважуються думки експертів. Цей коефіцієнт визначається по апіорним і апостеріорним даним. При використанні апіорних даних оцінювання коефіцієнта компетентності здійснюється до проведення експертизи на основі самооцінки експерта і взаємної оцінки інших експертів. При використанні апостеріорних даних оцінювання коефіцієнта компетентності здійснюється на основі опрацювання результатів експертизи. Використовуємо апіорні дані для вибору експертів на основі взаємної оцінки інших експертів.

Оцінювання змістової валідності створеного тесту. Технологія оцінювання змістової валідності створеного тесту передбачає [56]:

– організацію та проведення експертизи зі встановлення відповідності між тестовими завданнями і змістовою областю виконання тесту.

– приймання узагальнюючих висновків розробником щодо поліпшення змісту тесту.

Організація експертизи. Правильна організації експертизи є однією з найвагоміших задач розробників тесту, оскільки від неї залежить ефективність роботи експертів. У якості інструменту проведення експертизи тестів з вищої математики було обрано Google Docs додаток, використання якого має ряд переваг [56]:

- ✓ швидкі та ефективні збирання і оброблення даних;
- ✓ мінімальні витрати;
- ✓ можливість проведення експертизи незалежно від часу та відстані;
- ✓ зручність спільної роботи з документами;
- ✓ можливість обговорення та коментування в он-лайн режимі.

За 2-3 дні до початку роботи кожен експерт повинен ознайомитися зі специфікацією рецензованого тесту, що містить пояснення щодо його структури та запланованого до перевірки змісту. Для цього створюється файл у Google Docs, в якому розміщується текст специфікації (список вимог до рівня навчальних досягнень іспитників додається як додаток на окремій сторінці в кінці документа) та текст тесту. Доступ до документу відкривається всім експертам.

Проведення експертизи. Технологія експертизи якості змісту тесту включає три напрямки роботи експертів [56]:

- оцінювання завдань тесту;
- оцінювання тесту у цілому;
- оформлення узагальнюючих висновків і рекомендацій щодо поліпшення змісту тесту.

Оцінювання завдань тесту. Робота експерта за першим напрямом полягає в аналізі змісту окремих завдань тесту. Для цього експерт повинен заповнити Таблиця 3.1. Таблиця розміщується розробником тесту в файлі Google Docs. Для кожного експерта створюється окремий файл експертизи, до якого йому відкривається доступ. Не рекомендуємо створювати єдиний файл з дублюванням таблиці для всіх експертів, оскільки можливість бачити коментарії інших експертів впливає на вираження об'єктивної думки [56].

Таблиця 3.1.

Результати експертизи змісту завдань

1.	№ завдання	1	2	3	...	<i>L</i>	Σ
2.	№ правильної відповіді в закритих завданнях						
3.	Правильна відповідь завдань з відкритою відповіддю						
4.	№ вимоги до рівня навчальних досягнень						
5.	Рівень базовості						
6.	Значущість змісту завдання						
7.	Очікуваний відсоток виконання тесту іспитниками з задовільним рівнем навчальних досягнень						
8.	Очікуваний відсоток виконання тесту групою, де рівномірно представлені іспитники з різними рівнями навчальних досягнень						
9.	Очікуваний час виконання завдання, хв.						
10.	Завдання, зміст яких не відповідає специфікації в цілому						
11.	Завдання, зміст яких не відповідає підтемі специфікації						
12.	Завдання, умова яких є нечіткою та неоднозначною						
13.	Завдання, в умові яких міститься не вся інформація для надання правильної відповіді						
14.	Завдання, в умові яких міститься інформація, яка дає відповідь на інші завдання тесту						
15.	Завдання, в яких зустрічаються неправильно складені дистрактори						
16.	Інші невдалі завдання						
17.	Коментарії про якість завдання						

У таблиці перший рядок містить номери завдань тесту.

У другому рядку для закритих завдань тестів наводяться номери правильних відповідей, обраних експертом.

Відповіді до відкритих завдань наводяться експертом у третьому рядку таблиці. Тут необхідно звернути увагу розробників тесту на можливі випадки неоднозначності, коли виникають додаткові, частково правильні відповіді.

У четвертому рядку таблиці наводяться результати аналізу змісту завдань в термінах вимог до рівня навчальних досягнень іспитників. Для проведення аналізу експерту необхідно зіставити зміст кожного завдання тесту і перелік пронумерованих вимог. За результатами зіставлення експерт визначає номер вимоги, на яку переважно орієнтовано зміст даного завдання, і виставляє цей номер у четвертому рядку для кожного завдання тесту. У процесі аналізу можливі ситуації, коли в окремих завданнях контролюється не одне, а два або більше вимог до рівня навчальних досягнень іспитників. У цьому випадку експерт наводить тільки номер тієї вимоги, яка в основному перевіряється за допомогою даного завдання тесту.

У п'ятому рядку таблиці проставляється рівень завдання на основі зіставлення змісту завдань з вимогами до рівня навчальних досягнень іспитників. Завдання розділяють на три групи: Б (базовий), Д (достатній), В (високий). У тому випадку, коли завдання перевіряє ступінь досягнення вимог на мінімальному рівні, достатньому для виставлення оцінки «задовільно», воно вважається відповідним групі Б. До групи Д відносяться завдання, правильне виконання яких дає змогу виставити студентів оцінки «добре» і «відмінно». Таким чином, завдання групи Д за складністю перевищують мінімально достатній рівень, але знаходяться в межах мінімуму змісту освіти з даного предмету. До групи В включаються завдання, що виходять за межі мінімуму змісту та вимог до рівня навчальних досягнень іспитників.

У шостому рядку таблиці експерт оцінює рівень значущості змісту кожного завдання тесту. Оцінювання значущості пов'язана з необхідністю включення до тесту лише тих елементів змісту, які є найбільш важливими, ключовими для освоєння навчального курсу в межах вимог до рівня навчальних досягнень.

Оцінку значущості змісту завдань експерти виставляють за чотирибальною шкалою. Бал 0 виставляється в тому випадку, коли експерт вважає невиправданим внесення даного завдання в тест. Бал 1 виставляється

в тому разі, якщо зміст завдання представляється експерту незначним, 2 – значущим, а 3 – найбільш важливим опорним елементом. Отримані оцінки завдань тесту необхідно усереднити, обчисливши сумарну оцінку завдань тесту:

$$\mu = \frac{\sum_{j=1}^L Z_j}{L}$$

де Z_j – бал j -ого завдання, L – кількість завдань в тесті. Сумарна оцінка наводиться в комірці, яка знаходиться на перетині останнього стовпчика та п'ятого рядка.

Для заповнення сьомого рядку експерту необхідно представити гіпотетичний випадок, коли завдання з вибором відповіді виконувала вибірка студентів з задовільним рівнем навчальних досягнень. При цьому необхідно уявляти порівняно невелику вибірку, наприклад, з 10 осіб, і потім підрахувати гіпотетичний відсоток студентів, які правильно виконали завдання тесту.

У восьмому рядку експерт також наводить передбачуваний відсоток іспитників, які правильно виконали завдання тесту, орієнтуючись в своїх гіпотетичних оцінках на вибірку, де рівномірно представлені іспитники з різними навчальних досягнень.

Дев'ятий рядок містить очікуваний час виконання студентом кожного завдання тесту. При виборі часу експерту слід орієнтуватися на студента з середньою підготовкою. Перше уявлення про час виконання завдань експерт може скласти в процесі своєї роботи над тестом. Тому, приступаючи до виконання завдань, експерту потрібно підготувати все необхідне: годинник, бланк з номерами завдань, і строго фіксувати початок і кінець роботи над кожним завданням тесту. Отриманий час необхідно скорегувати з урахуванням того, що завдання виконував викладач, а не студент. Для корекції зафіксований час виконання кожного завдання збільшують в 4-5 разів. Отриманий результат (у хвилинах) слід записати в графу для очікуваного часу виконання студентом кожного завдання тесту.

У десятому та одинадцятому рядках експерт відповідно проставляє зірочки в комірках тих завдань, зміст яких не відповідає специфікації в цілому або обраній підтемі.

У дванадцятому рядку зірочками відмічаються ті завдання, умова яких є нечіткою та неоднозначною, так що всі іспитники працюють над різним завданням, у тринадцятому – ті, в умові яких міститься не всі відомості для надання правильної відповіді.

У чотирнадцятому рядку експерт відмічає завдання, в умові яких містяться відомості, які дають відповідь на інші завдання тесту (у дужках зазначає, на які саме).

У п'ятнадцятому рядку експерт проставляє зірочки напроти тих завдань, в яких зустрічаються неправильно складені дистрактори. До таких завдань можуть належати такі:

- в яких немає логічної та граматичної відповідності між умовою та кожним дистрактором;
- в яких один або більше дистракторів є неправдоподібними;
- в яких є дистрактори, які можуть бути прийняті іспитниками за правильну відповідь;
- в яких іспитник може віднайти правильну відповідь логічним шляхом, маючи досвід тестування;
- в яких довжина запропонованих відповідей різної довжини.

Позначки експертів, виставлені в 10-15 рядках таблиці, дають змогу виявити невдало складені завдання.

У шістнадцятому рядку експерт відмічає інші невдалі завдання тесту. До таких слід віднести, наприклад, відкриті завдання з короткою відповіддю, в яких можливі випадки появи неоднозначності, коли на місці планованого єдиної відповіді можна привести синоніми, різні числа або більш-менш правильні відповіді.

Всі пояснення щодо виявлення невдалих завдань та додаткові зауваження щодо якості змісту експерт повинен привести в сімнадцятому

рядку. Якщо експерт вважає за потрібне дати розгорнуті рекомендації щодо поліпшення змісту завдань, то їх слід привести на окремих сторінках файлу експертизи з заголовком «Додаток 1. Зауваження щодо якості змісту завдань». У цьому ж рядку експерт характеризує якість формулювань завдань тесту, якість представлення графічної інформації, якщо вона є в завданнях тесту. При оцінюванні якості формулювань експерту слід пам'ятати про те, що тестові завдання повинні бути сформульовані чітко, точно, лаконічно і коректно з точки зору загальноприйнятої в навчальному предметі термінології і символіки предмета.

Оцінювання тесту у цілому. Другий напрям роботи експерта пов'язаний з аналізом якості змісту всього тесту.

При роботі за другим напрямом слід мати на увазі, що зміст тесту визначається як оптимальне відображення вимог до рівня навчальних досягнень іспитників у системі завдань тесту. Експерт, використовуючи дані четвертого рядка Таблиця 3.1, повинен обчислити відсоток вимог, охоплених тестуванням:

$$v = \frac{s}{q} \cdot 100\%,$$

де s – кількість вимог, охоплених тестом, q – загальна кількість вимог. Отриманий експертом відсоток охоплення порівнюється з наведеними в специфікації тесту. Потім обчислюється міра відхилення у вигляді різниці відсотків, яка заноситься до Таблиця 3.2, яка розміщується на окремій сторінці файлу експертизи з заголовком «Додаток 2. Відхилення відсотку охоплення вимог до рівня навчальних досягнень іспитників». Тут же для зручності попередньо записують відсоток вимог, охоплених всім тестом, заявлений в специфікації, та відсоток вимог v , отриманий експертом.

Експерт має оцінити правильність пропорцій змісту тексту. Необхідно перевірити, чи всі важливі аспекти предметної області і в правильній пропорції охоплює тест. Найчастіше при розробці тесту можливе зміщення пропорцій, оскільки тест легко перенаситити тими розділами змісту, за якими

легше складати завдання. Для оцінювання правильності пропорцій тесту експерт підраховує відсотки завдань у тесті, орієнтованих на матеріал кожної підтеми, які разом із своїм баченням оптимального співвідношення розділів наводить на окремій сторінці файлу експертизи з заголовком «Додаток 3. Оцінювання правильності пропорцій тесту та перевірка відповідності змісту специфікації». Тут же записують різницю, що характеризує відхилення думки розробників від експертних оцінок

Таблиця 3.2.

**Відхилення відсотку охоплення вимог до рівня
навчальних досягнень іспитників**

Відсоток охоплених варіантами заявлений в специфікації	вимог, всіма тесту,	Відсоток охоплених варіантами отриманий експертом	вимог, всіма тесту,	Відхилення охоплення вимог навчальних іспитників	відсотку до рівня досягнень

Експерт перевіряє відповідність змісту тесту його специфікації. Для цього експерт обчислює загальну кількість завдань, які не відповідають специфікації тесту, використовуючи дані 10 та 11 рядків Таблиця 3.1 та записує це число у Додатку 3.

Використовуючи дані 9 рядка Таблиця 3.1 експерт обчислює очікуваний час виконання всього тесту і записує його в комірці, яка знаходиться на перетині 9 рядка та останнього стовпця.

Оформлення узагальнюючих висновків і рекомендацій щодо поліпшення змісту тесту. Третій напрям роботи експерта полягає в оформленні узагальнюючих висновків і рекомендацій щодо покращення змісту тесту. На окремій сторінці файлу експертизи з заголовком «Висновки та рекомендації щодо поліпшення змісту тесту» експерт наводить свій загальний висновок про зміст тесту та рекомендації щодо його покращення. При цьому обов'язково необхідно:

- зробити висновки щодо правильності пропорцій тесту та відповідності змісту специфікації;
- оцінити значущість змісту завдань та тесту в цілому;

- оцінити, на який рівень підготовки іспитників орієнтований тест;
- навести рекомендації щодо коректування часу виконання тесту;
- навести рекомендації щодо виявлених невдалих завдань тесту;
- виставити оцінку змістової валідності від 0 до 1, враховуючи, такі інтерпретації її значень: 0,7-1 – тест має високу змістову валідність; 0,4-0,6 – необхідна деяка корекція тесту та специфікації; 0-0,3 – необхідно докорінно змінити структуру тесту.

Приймання узагальнюючих висновків розробником щодо поліпшення змісту тесту. На основі результатів експертизи розробник перевіряє узгодженість думок експертів та робить узагальнюючі висновки щодо якості змісту тесту. Узгодженість висновків експертів перевіряється на основі обчислення коефіцієнта рангової конкордації Кендалла-Сміта [126]. У випадку, коли висновки експертів не є узгодженими, то відповідно до методу Дельфі експертиза повторюється. Коли ж висновки експертів узгоджені, то усереднюючи думки всіх експертів, наприклад, за методом Джона Кемені [126], розробник приймає рішення щодо подальшої роботи з тестом:

- залишити тест без змін;
- необхідна деяка корекція тестових завдань або специфікації тесту;
- необхідно докорінно змінити структуру тесту, переглянути специфікацію та провести повторну експертизу якості тесту.

3.2. Комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики

Комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики включає мету, зміст, методи та засоби оцінювання. *Метою* оцінювання якості тестів є покращення контролю знань студентів з вищої математики. *Засобом* оцінювання є авторська система автоматизованого аналізу якості тестових завдань, основні функції якої розглянуто в п. 2.3. Опис інтерфейсу СААЯТЗ наведено в Додатку А. *Методами* методики є

дібрані методи теорії ймовірностей та математичної статистики, а саме взаємодоповнюючі методи КТТ та IRT, описані в пункті 2.2.

Змістовий компонент оцінювання відображає загальна схема на Рис. 3. 1, що включає в себе такі етапи [48]:

- формування таблиці результатів;
- аналіз окремих тестових завдань і тесту в цілому;
- приймання остаточних рішень.

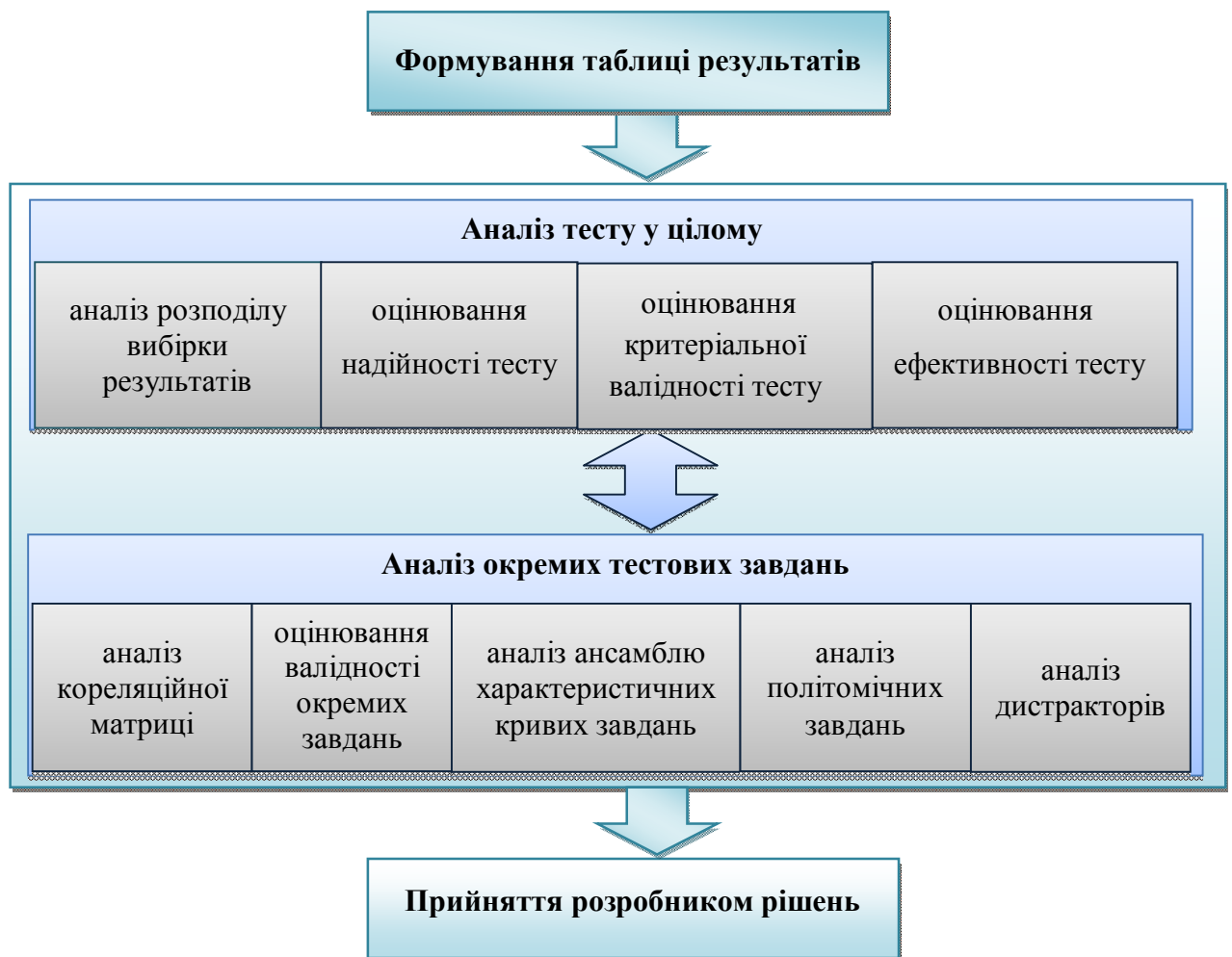


Рис. 3. 1. Загальна схема методики комп'ютерно орієнтованого оцінювання якості тестів

На **першому етапі** здійснюється формування таблиці результатів

На **першому етапі** здійснюється формування таблиці результатів тестування (Рис. 3.2). Обчислюються індивідуальні бали іспитників X_i , $i = \overline{1, N}$, де N - кількість іспитників, і загальні бали за завдання R_j , $j = \overline{1, K}$,

де K - кількість завдань тесту. Далі всюди у формулах використовуються ці позначення. Здійснюється впорядкування матриці результатів тестування спочатку в порядку спадання значень X_i , а потім – значень R_j .

На **другому етапі** проводиться аналіз окремих тестових завдань і тесту в цілому із використанням СААЯТЗ.

Аналіз тесту у цілому полягає у такому [48]:

- аналізі розподілу вибірки результатів;
- оцінюванні надійності тесту;
- оцінюванні критеріальної валідності тесту;
- оцінюванні ефективності тесту.

Аналіз розподілу вибірки результатів. Здійснюються графічна інтерпретація емпіричних даних у вигляді гістограми, аналіз розподілу вибірки індивідуальних балів та обчислення вибірових характеристик.

Оцінювання надійності тесту. Обчислюються коефіцієнт надійності тесту r (2.4), знайдений на основі методу розщеплення, та коефіцієнт надійності Кронбаха α (2.5). В якості нижньої границі допустимих значень коефіцієнта надійності обрано 0,7 [162].

ІПБ іспитника i	Номер завдання j															X_i
	№1	№2	№3	№4	№5	№6	№7	№8	№9	№10	№11	№12	№13	№14	№15	
Артем Юрійович Лашко	0,27	0,4	0,53	0,53	0	0,53	0	0,53	0,53	0	0,53	0,53	0,53	0,53	0,53	5,97
Рустам Олегович Срохін	0,27	0	0,53	0	0	0	0,53	0	0,53	0	0	0	0,53	0,53	0	2,92
Олександр Євгенійович Хорунжий	0,27	0,53	0,53	0,53	0	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	7,16
Микита Миколайович Кузімович	0,4	0,53	0,53	0	0	0	0	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	5,7
Дар'я Вікторівна Кравець	0,4	0,53	0,53	0	0	0	0	0,53	0,53	0	0,53	0,53	0,53	0,53	0	4,11
Андрій Геннадійович Захарчук	0,27	0,53	0,53	0	0	0,53	0	0	0	0,53	0,53	0,53	0,53	0,53	0,53	5,04
Марія Олександрівна Ігнатівська	0,27	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0	0,53	7,16
Антон Олександрович Чічков	0,4	0,53	0,53	0,53	0	0,53	0,53	0,53	0,53	0,53	0,53	0	0,53	0	0,53	6,23
...	...															...
R_j	13,65	16,58	16,96	11,13	5,06	11,13	9,54	15,9	19,61	15,37	20,14	14,31	18,55	13,25	14,31	

Рис. 3.2. Формування таблиці результатів тестування

Оцінювання критеріальної валідності тесту. Для оцінювання критеріальної валідності обчислюється коефіцієнт кореляції Пірсона між результатами тестування і результатами зовнішнього критерію. Як критерій зазвичай беруться оцінки експертів, виставлені при традиційній перевірці знань студентів без використання тестів. При значенні $\rho \geq 0,7$ вважаємо, що тест має високу валідністю.

Оцінювання ефективності тесту. Будується інформаційна функція тесту, оцінюється його ефективність. Визначається інтервал значень параметра підготовленості іспитника, для яких даний тест є найінформативнішим.

Графік інформаційної функції правильно складеного тесту повинен мати дзвоноподібну, але не занадто загострену форму. Якщо інформаційна функція має декілька локальних екстремумів, то тест потребує вдосконалення. При цьому, якщо кількість завдань у тесті невелика, то потрібно додавати завдання, які мають проміжну складність, щоб ліквідувати «провали» між сусідніми екстремумами. Якщо кількість завдань у тесті досить велика, то його доречно розбити на два тести.

Найкраще працює тест на тому інтервалі значень параметра θ на якому інформаційна функція приймає найбільші значення.

Аналіз тестових завдань передбачає [48]:

- аналіз кореляційної матриці завдань;
- оцінювання валідності окремих завдань;
- аналіз ансамблю характеристичних кривих завдань;
- аналіз політомічних завдань;
- аналіз дистракторів завдань множинного вибору.

Аналіз кореляційної матриці завдань. Обчислюється кореляційна матриця завдань. Для того, щоб тестові завдання не дублювали одне одного, значення кореляції між ними має бути не занадто високим ($r_{ij} \leq 0,3$). Від’ємна кореляція завдання з більшістю інших завдань говорить про потребу перегляду цього завдання, виявлення в ньому помилок, або навіть видалення з тесту.

Оцінювання валідності окремих завдань. Для оцінювання валідності j -го завдання обчислюються значення індексу розрізняльної здатності дихотомічних завдань D_j , коефіцієнту точкової бісеріальної кореляції між дихотомічним завданням та індивідуальними балами іспитників і коефіцієнту

кореляції Пірсона між політомічним завданням та індивідуальними балами іспитників ρ_j .

На основі практичного досвіду обрано такі інтерпретації значень D_j (див. (2.6)): $D_j \geq 0,4$ – завдання добре функціонує; $0,3 \leq D_j \leq 0,39$ – можливо необхідна деяка корекція завдання; $0,2 \leq D_j \leq 0,29$ – завдання, яке потребує перегляду; $D_j \leq 0,19$ – завдання, яке необхідно вилучити з тесту.

Для правильно складених завдань мають виконуватись такі умови [99, ст. 433]:

$$\rho_j \geq \frac{2}{\sqrt{N-1}}; \rho_{bisj} \geq \frac{\sqrt{pq / (N-1)}}{Y}.$$

Аналіз ансамблю характеристичних кривих завдань. Перевіряється рівномірність покриття характеристичними кривими завдань інтервалу у межах від -5 до 5 логітів. Якщо рівномірність покриття порушено характеристичними кривими, розміщеними зліва (справа) на графіку, то відповідні завдання мають мінімальне (максимальне) значення складності. Їх необхідно вилучити з тесту, якщо й інші показники цих тестових завдань вказують на неправильність їх складання. Якщо ж прогалина спостерігається в середині між характеристичними кривими, то для її заповнення до тесту необхідно додати завдання, які мають проміжну складність [48].

Аналіз політомічних завдань. Проводиться сумісний аналіз характеристичних кривих рівнів та інформаційних кривих політомічних завдань.

Аналіз дистракторів завдань множинного вибору. Дистрактори, які вибирають менше 5% іспитників, повинні бути видалені або замінені. Також для аналізу дистракторів завдань множинного вибору можна застосувати характеристичні криві IRT-моделей множинного вибору.

На **третьому етапі** аналізу на підставі сумісного аналізу взаємодоповнюючих характеристик IRT та КТТ робляться висновки щодо якості як окремих тестових завдань, так і тесту в цілому, приймаються остаточні рішення.

Якість тесту в цілому оцінюється відповідно до схеми, зображеної на (Рис. 3. 3), на основі аналізу розподілу вибірки результатів, знайдених значень коефіцієнтів надійності та критеріальної валідності тесту, графіку інформаційної функції.

Нижче наведемо умовні позначення на Рис. 3. 3.

-  - коефіцієнт надійності більше або дорівнює 0,7;
-  - коефіцієнт надійності менше або дорівнює 0,3;
-  - коефіцієнт валідності більше або дорівнює 0,7;
-  - коефіцієнт валідності менше або дорівнює 0,3;
-  - інформаційна функція правильної форми;
-  - розподіл індивідуальних балів наближений до нормального;
-  - тест складено неправильно. Необхідно змінити зміст всього тесту.
-  - тест складено неправильно. Необхідно змінити структуру тесту на

основі аналізу якості окремих завдань та додаванням завдань проміжної складності.

 - тест допускає вдосконалення на основі аналізу якості окремих тестових завдань.

 - тест допускає вдосконалення шляхом додавання завдань, які мають проміжну складність або розбиттям тесту на два тести.

 - тест складено правильно. Він є надійним, валідним та ефективним.

НД – надійність;

ВЛД – валідність;

ЕФЕКТ – ефективність;

Н. РЗП. ІБ – нормальний розподіл індивідуальних балів.

Оцінювання якості окремих тестових завдань здійснюється відповідно до схеми, зображеної на Рис. 3. 4, 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.

Нижче наведемо умовні позначення на Рис. 3. 4.

 - виконується хоча б одна з умов: завдання занадто легке/складне; переважно від'ємна або занадто висока (більша за 0,3) кореляція з іншими завданнями; низькі значення коефіцієнтів дискримінативності;

 - завдання політомічне;

 - завдання множинного вибору;

 - виявлено невідповідність характеристичних кривих рівнів завдання;

 - виявлено дистрактор, що не спрацьовує;

 - виявлено помилку в тексті.

 - завдання складено неправильно. Необхідно видалити його з тесту або замінити іншим.

 - змінити або виправити дистрактор тестового завдання.

 - виправити помилку в тексті завдання;

 - змінити кількість рівнів (умови нарахування балів) або скоректувати текст завдання на підставі сумісного аналізу інформаційних функцій та характеристичних кривих рівнів завдань. Якщо це неможливо – видалити завдання з тесту або замінити іншим.

 - завдання складено правильно. Залишити його в тесті без змін.

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.

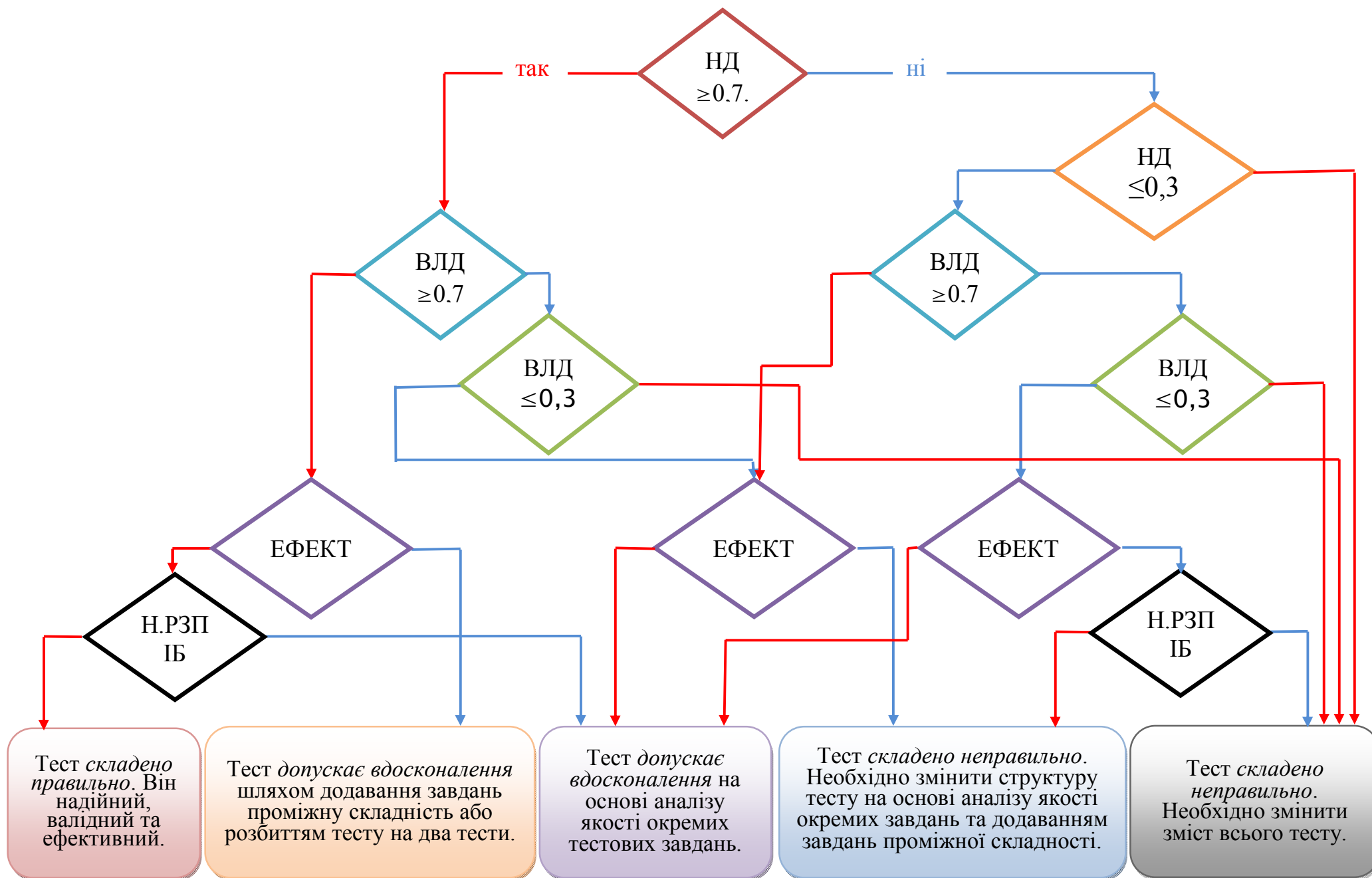


Рис. 3. 3. Оцінювання якості тесту в цілому

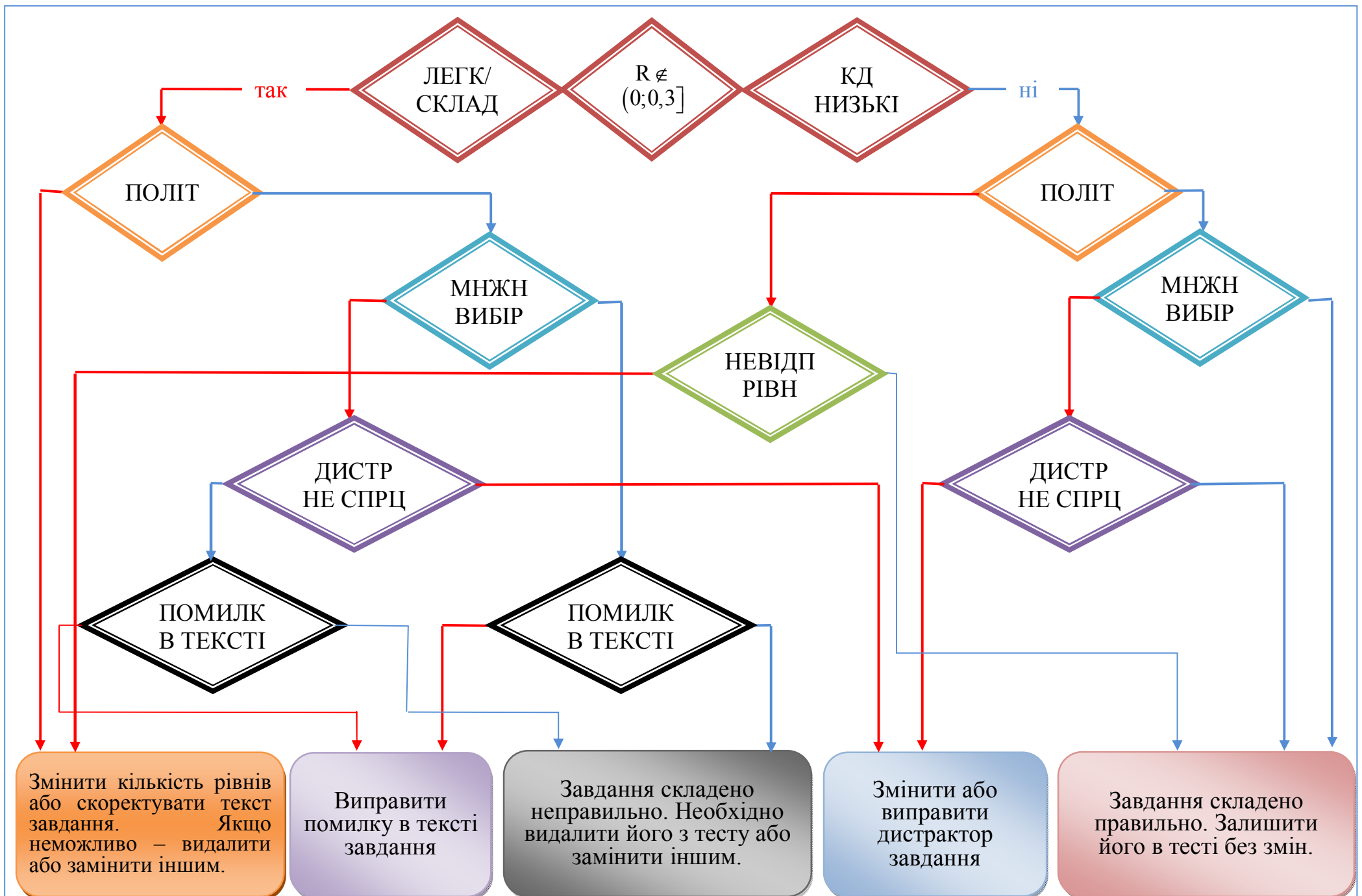


Рис. 3. 4. Оцінювання якості тестових завдань

3.3. Практичне застосування методики та особливості роботи викладача з тестами з вищої математики

Питання впровадження комп'ютерно орієнтованого оцінювання якості тестів з вищої математики потребує визначення її місця в процесі роботи викладача з тестами з вищої математики для оцінювання знань студентів.

Загальну схему роботи викладача з тестами з вищої математики зображено на Рис. 3. 5. Загальна схема роботи викладача з тестами. Розберемо її детальніше. На першому етапі викладач створює тест. Одне із ключових питань при цьому є ***визначення підходу до розроблення тестів***. Воно детально розглянуто в пункті 2.3.1. Підкреслимо, що для тестів з вищої математики є необхідним об'єднання критеріально-орієнтованого та нормативно-орієнтованого підходів при оцінюванні навчальних досягнень студентів. Саме таке об'єднання рекомендується при створенні тестів з вищої математики. Саме за такої реалізації інтерпретація результатів тесту по відношенню до певного змісту навчальної дисципліни не знімає можливості диференціації індивідуальних відмінностей студентів.

Іншим важливим питанням при розробці тестів з вищої математики є ***задача оцінювання змістової валідності***. Процес оцінювання змістової валідності тестів з використанням Google Docs додатку детально описаний в пункті 3.1.

Коли розробник переконується, що тест має високу змістову валідність, він переходить до такого етапу, ***завантаження тестів у систему комп'ютерного тестування***. Пропонується для розміщення тестів використовувати систему Moodle, яка має потужний засіб для створення тестів та має ряд переваг [24]:

- Moodle є відкритою системою з ліцензією, яка передбачає безкоштовне використання і має відкритий програмний код;
- простий, ефективний, сумісний з більшістю браузерів web-інтерфейс;
- широкий вибір типів тестових завдань;

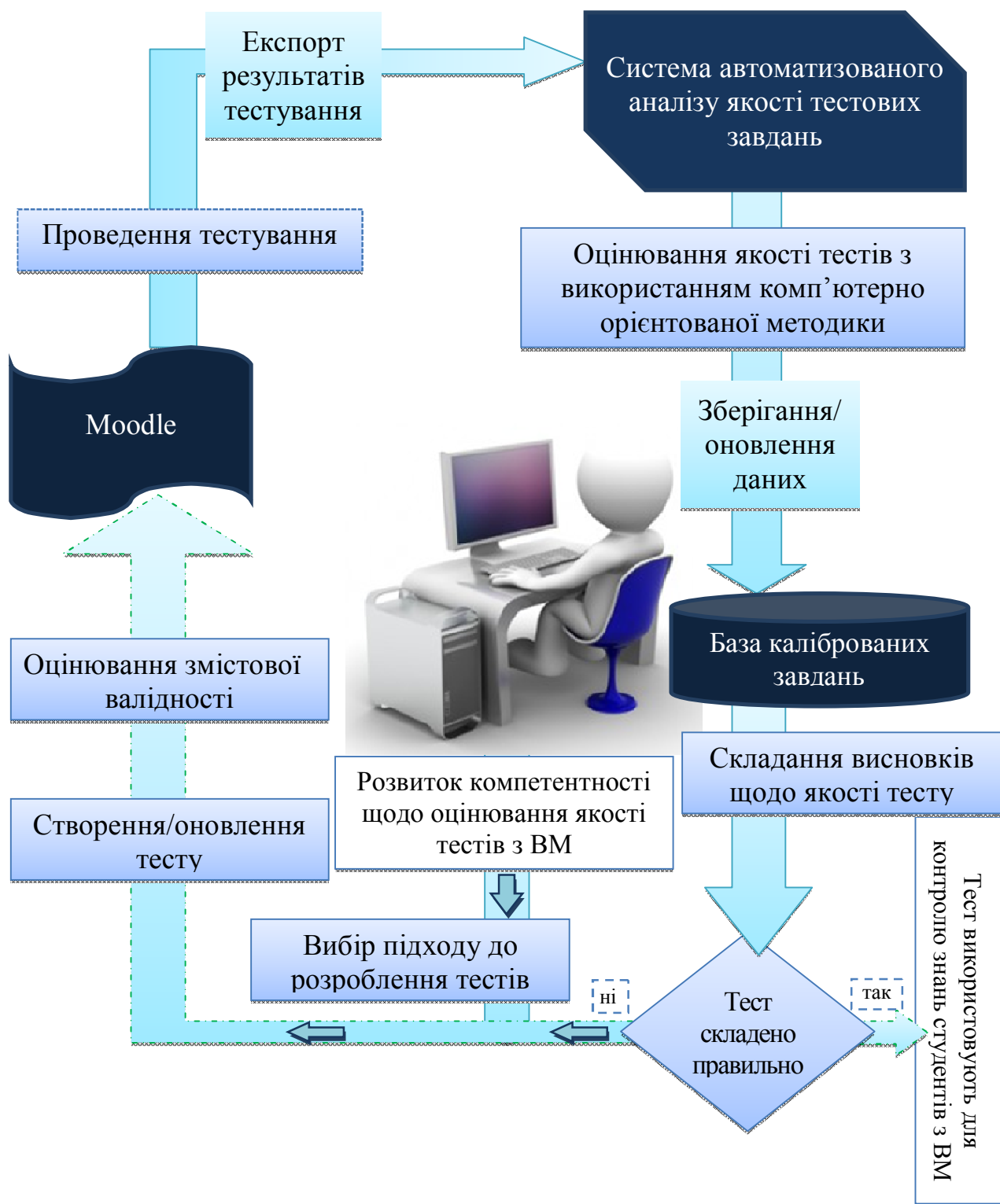


Рис. 3. 5. Загальна схема роботи викладача з тестами

- порядок подання питань в тесті і варіантів відповідей в питаннях може бути як заданим, так і випадковим;
- завдання можуть містити HTML та графічні елементи;

- тест може бути обмежений за часом і за кількістю спроб його проходження;
- кожна спроба може залежати від результатів попередньої;
- день і час доступності тесту можуть бути жорстко задані, або тест може бути відкритий завжди;
- можливість встановлення пароллю для отримання дозволу на проходження тестування;
- звіт про результати тесту для іспитника може бути гнучко настроєний;
- підсумкова оцінка за тест складається з балів, отриманих за кожне питання тесту, з урахуванням ваги питання;
- можливість нарахування штрафних балів за неправильну відповідь у завданнях множинного вибору з декількома правильними варіантами відповіді та у завданнях на відповідність (політомічних завданнях);
- питання можна вводити вручну через інтерфейс у Moodle або імпортуватися у різноманітних формах;
- обчислення деяких статистичних характеристик якості тесту;
- зручний експорт результатів тестування та ін.

Після розміщення тестів у системі комп'ютерного тестування **викладач проводить пробне тестування**. При цьому виникає ряд питань, пов'язаних з організацією та проведенням тестування. Викладачу необхідно вжити спеціальних заходів, що забезпечують стандартизацію процедури виконання тесту. Стандартизація цієї процедури торкається питань формування репрезентативної вибірки іспитників, підготовку студентів до тестування, правильного інструктування, вибору оптимального часу тестування, вибору місця проведення. Дані питання було розглянуто у п. 1.1.

Після проведення тестування викладач експортує результати тестування у форматі .xls. Далі **із використанням СААЯТЗ проводиться оцінювання якості тестів** за методикою, описаною в пункті 3.1, на основі чого **приймаються рішення щодо якості тесту та подальшої роботи з ним**.

Тест та завдання тесту разом з їх статистичними характеристиками *зберігаються у базі каліброваних завдань*.

Якщо зроблено висновок, що тест складений правильно і в ньому не виявлено невдалих завдань, його у подальшому використовують для оцінювання навчальних досягнень з вищої математики.

Якщо ж виявлено, що тест потребує деякої корекції, то шляхом видалення з нього неправильно складених завдань або додання нових завдань (з бази каліброваних завдань) тест оновлюється. Далі цикл повторюється: знову оцінюється змістова валідність, завантажуються тести у систему комп'ютерного тестування, проводиться ще одне пробне тестування, оцінюється якість тесту із використанням СААЯТЗ, зберігаються дані в базі каліброваних завдань, приймаються рішення щодо якості тесту та подальшої роботи з ним. Цикл завершується, коли прийнято висновок, що тест складений правильно.

Необхідно зазначити, що для роботи з тестами за описаною схемою, викладач повинен мати професійний рівень компетентності щодо оцінювання якості тестів з вищої математики. Розвиток даної компетентності у дослідженні пропонується проводити відповідно до технології, описаної в п. 3.5.

Розроблену комп'ютерно орієнтовану методику оцінювання якості тестів було застосовано до пакету тестів дистанційного курсу «Вища математика» [134] для студентів технічних спеціальностей КПІ ім. Ігоря Сікорського.

Розглянемо приклад застосування методики до тестової контрольної роботи на тему «Диференціальне числення функцій однієї змінної».

При розробці даного тесту проблему вибору підходу до розроблення тестів було вирішено відповідно п.3.1. Після розроблення тесту було оцінено його змістову валідність відповідно до технології, описаної в п. 3.2. В результаті було зроблено висновок, що тест має високу валідність.

Після цього тест було розміщено у системі Moodle та проведено пробне тестування 151 студента першого курсу РТФ та ІТС у 2013 році. Результати тестування було експортовано з системи Moodle у форматі .xls. Далі із використанням СААЯТЗ було оцінено якість тесту за методикою, описаною в пункті 3.2.

На першому етапі сформовано таблицю результатів тесту, фрагмент якої представлено на Рис. 3.6

ІПБ	Номер завдання j															Xi
	№15	№3	№6	№10	№1	№8	№4	№11	№5	№14	№2	№12	№9	№13	№7	
Альона Олександрівна Бикова	0,53	0,53	0,53	0,53	0,4	0,53	0,53	0,53	0,53	0,53	0,27	0,53	0,53	0,53	0,53	7,6
Наталія Андріївна Гвоздецька	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0	0,53	0,53	0,53	7,47
Юрій Вікторович Яценко	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0	0,53	0,53	0,53	0,53	0,53	0,53	7,47
Максим Райз	0,53	0	0	0	0	0,53	0	0	0,27	0	0,27	0	0	0,53	0	2,13
Джанет Обенева Дансо	0	0	0,53	0	0,13	0,53	0	0	0,27	0	0	0,53	0	0	0	2
<i>Rj</i>	25,97	21,73	21,2	20,67	18,63	18,55	18,02	15,9	15,43	15,37	14,43	13,78	13,25	11,66	4,77	

Рис. 3.6. Фрагмент таблиці результатів тесту «Диференціальне числення функцій однієї змінної»

На наступному етапі проведено аналіз тестових завдань і тесту в цілому.

Дані представлено у вигляді частотного розподілу індивідуальних балів на Рис. 3.7.

Бал	[0;1]	(1;2]	(2;3]	(3;4]	(4;5]	(5;6]	(6;7]	(7;8]
Частота	0	1	7	6	11	11	11	4

Рис. 3.7. Частотний розподіл індивідуальних балів

Гіпотеза про нормальність розподілу індивідуальних балів тесту підтверджено за критерієм Пірсона. Гістограму частотного розподілу індивідуальних балів представлено на Рис. 3.8.

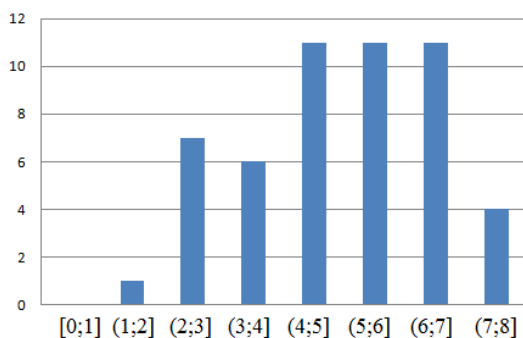


Рис. 3.8. Гістограма частотного розподілу індивідуальних балів

Мода частотного розподілу індивідуальних балів дорівнює 6, середнє вибіркове – 4,88. Правильно складений тест має бути унімодальним, і значення моди має бути близьким до значення вибіркового середнього [107]. Отже, даний тест потребує деякої корекції.

Значення асиметрії вибірки дорівнює -0,23, що говорить про дещо завищену складність тесту.

Коефіцієнт надійності тесту, знайдений на основі методу розщеплення, складає $r_t=0,78$; коефіцієнт надійності Кронбаха $\alpha=0,72$. В якості нижньої границі припустимих значень коефіцієнта надійності обрано 0,7. Отже, можна говорити про високу надійність тесту.

Оцінено критеріальну валідність тесту. Критерієм обрано оцінювання експертів, виставлені при традиційній перевірці знань студентів. Коефіцієнт валідності $\rho=0,82$. В якості нижньої границі припустимих значень коефіцієнта валідності обрано 0,7. Отже, даний тест має високу валідність.

Побудовано інформаційну функція тесту, яка має не занадто загострену дзвоноподібну форму (Рис. 3.9). Можна зробити висновок, що, в цілому, вигляд інформаційної функції відповідає правильно складеному тесту. Найбільш інформативним тест є для іспитників з рівнем підготовленості від -2,5 до 2,5 логітів.

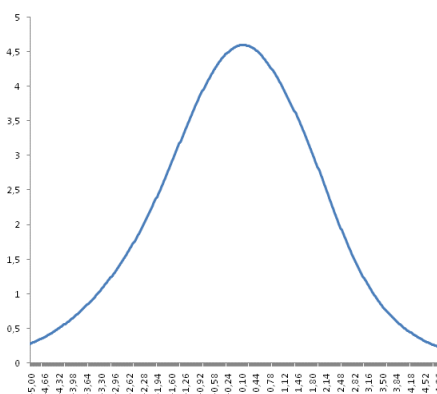


Рис. 3.9. Інформаційна функція тесту «Диференціальне числення функцій однієї змінної»

Обчислено кореляційну матрицю завдань. (Рис. 3.10). Переважно від’ємна або занадто висока (більша за 0,3) кореляція з іншими завданнями спостерігається у завдань № 5 та №12.

№ завдання	№15	№3	№6	№10	№1	№8	№4	№11	№5	№14	№2	№12	№9	№13	№7
№15	1,00	0,15	0,14	0,36	0,05	0,08	0,29	0,24	-0,09	0,03	0,21	-0,20	0,20	0,18	0,09
№3		1,00	0,22	0,19	0,34	0,20	0,17	0,09	0,36	0,27	0,10	0,01	0,19	0,23	0,23
№6			1,00	0,50	0,27	0,36	0,03	-0,05	-0,08	0,02	0,05	0,06	0,23	0,07	0,24
№10				1,00	0,29	0,22	0,20	0,10	0,09	0,26	0,19	0,10	0,36	0,30	0,26
№1					1,00	0,19	0,33	0,13	-0,05	0,19	0,32	0,09	0,16	0,09	0,11
№8						1,00	0,15	0,04	0,21	-0,08	0,03	-0,07	0,24	0,08	0,09
№4							1,00	0,17	-0,03	0,22	0,27	0,06	0,36	0,28	0,11
№11								1,00	-0,14	0,32	0,35	0,22	0,18	0,17	0,07
№5									1,00	0,19	-0,14	-0,03	0,03	0,18	0,16
№14										1,00	0,31	0,10	0,22	0,44	0,20
№2											1,00	-0,11	0,29	0,38	0,05
№12												1	0,098462	-0,01708	0,042366
№9													1	0,254669	0,163411
№13														1	0,323773
№7															1

Рис. 3.10. Кореляційна матриця завдань тесту «Диференціальне числення функцій однієї змінної»

Обчислено показники валідності окремих завдань тесту (Рис. 3.11). Знайдено значення індексу розрізняльної здатності D_j . На основі практичного досвіду обрано такі інтерпретації значень D_j : $D_j \geq 0,4$ – завдання добре функціонує; $0,3 \leq D_j \leq 0,39$ – можливо необхідна деяка корекція завдання; $0,2 \leq D_j \leq 0,29$ – завдання, яке потребує перегляду; $D_j \leq 0,19$ – завдання, яке є складеним неправильно. Завдання №15 з індексом розрізняльної здатності $D_1=0,14$ є неправильно складеним. Привертають до себе увагу завдання №6, 7, 12 з $D_6 = D_7 = D_{12}=0,36$. Текст завдань № 15, 6, 7, 12 відповідно на Рис. 3.12-Рис. 3.15. Далі розглянемо їх детальніше.

	#15	#3	#6	#10	#1	#8	#4	#11	#5	#14	#2	#12	#9	#13	#7
D_j	0,14	0,50	0,36	0,50	0,36	0,50	0,64	0,71	0,36	0,64	0,43	0,36	0,71	0,71	0,36
ρ_j	0,32	0,53	0,43	0,62	0,46	0,40	0,52	0,43	0,27	0,56	0,43	0,25	0,59	0,59	0,45
ρ_{bisj}	0,22	0,42	0,39	0,54	0,51	0,39	0,49	0,44	0,33	0,54	0,56	0,28	0,66	0,67	0,78
$t: t \leq \rho_{bisj}$	0,32	0,40	0,40	0,38	0,57	0,37	0,37	0,36	0,40	0,36	0,41	0,35	0,35	0,36	0,43

Рис. 3.11. Показники валідності завдань тесту «Диференціальне числення функцій однієї змінної»

Обчислено коефіцієнти точкової бісеріальної кореляції між дихотомічними завданнями та індивідуальними балами іспитників і коефіцієнти кореляції Пірсона між політомічними завданням та індивідуальними балами іспитників ρ_j . В якості нижньої границі припустимих значень коефіцієнта точкової бісеріальної кореляції 0,28. Завдання №5 і №12 потребують детального вивчення: $\rho_5=0,27$ $\rho_{12}=0,25$.

15 Баллов: 1

Похідну параметрично заданої функції $\begin{cases} x = x(t), \\ y = y(t) \end{cases}$ знаходять за формулою

Выберите один ответ.

a. $\begin{cases} x = x(t), \\ y'_x = \frac{y'_t}{x'_t} \end{cases}$ ✓

b. $\begin{cases} x = x(t), \\ y'_x = \frac{x'_t}{y'_t} \end{cases}$ ✗

c. $\begin{cases} x = x(t), \\ y'_x = -\frac{x'_t}{y'_t} \end{cases}$ ✗

d. $\begin{cases} x = x(t), \\ y'_x = -\frac{y'_t}{x'_t} \end{cases}$ ✗

Рис. 3.12. Текст завдання №15 «Диференціальне числення функцій однієї змінної»

6 Баллов: 1

Закінчить теорему Коші: Якщо функції $f(x)$ і $g(x)$ неперервні на $[a, b]$, диференційовні в (a, b) , причому $g'(x) \neq 0$, коли $x \in (a, b)$, то існує точка $c \in (a, b)$ така, що правдива рівність

Выберите один ответ.

a. $\frac{f(b)f(a)}{g(b)g(a)} = \frac{f'(c)}{g'(c)}$ ✗

b. $\frac{f(b) - g(b)}{f(a) - g(a)} = \frac{f'(c)}{g'(c)}$ ✗

c. $\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}$ ✓

d. $\frac{f(b) + f(a)}{g(b) + g(a)} = \frac{f'(c)}{g'(c)}$ ✗

Рис. 3.13. Текст завдання №6 «Диференціальне числення функцій однієї змінної»

7 Баллов: 1

Знайдіть $\lim_{x \rightarrow +\infty} (x + 3^x)^{1/x}$.

Выберите один ответ.

a. 3 ✓

b. ∞ ✗

c. $\ln 3$ ✗

d. e^3 ✗

Рис. 3.14. Текст завдання №7 «Диференціальне числення функцій однієї змінної»

12 Баллов: 1

Знайдіть Тейлорів многочлен 3-го порядку в точці $x = 0$ для функції $f(x) = \sin 2x$.

Выберите один ответ.

a. $2x + \frac{4x^3}{3}$ ✗

b. $x - \frac{x^3}{6}$ ✗

c. $2x - 8x^3$ ✗

d. $2x - \frac{4x^3}{3}$ ✓

Рис. 3.15. Текст завдання №12 «Диференціальне числення функцій однієї змінної»

Обчислено бісеріальний коефіцієнт кореляції між дихотомічним завданням та індивідуальними балами іспитників ρ_{bisj} . Для кожного завдання встановлено нижню границю допустимих значень бісеріальних коефіцієнтів кореляції. Підтверджено неправильність складання завдань №12: $\rho_{bis12}=0,28$ (при встановленому значенні нижньої границі 0,35) та №15: $\rho_{bis15}=0,22$ (при встановленому значенні нижньої границі 0,32). Дещо виділяється завдання №6 : $\rho_{bis6}=0,39$ (при встановленому значенні нижньої границі 0,4).

Побудовано ансамбль характеристичних кривих завдань тесту (Рис. 3.16а). Можна побачити, що характеристичні криві досить рівномірно покривають інтервал від -5 до 5 логітів. Рівномірність покриття порушує лише завдання під номером 15 (Рис. 3.12), яке дійсно виявилось занадто легким. Це тільки підтверджує неправильність його складання, виявлену раніше

Побудовано інформаційні криві всіх завдань тесту (Рис. 3.16б).

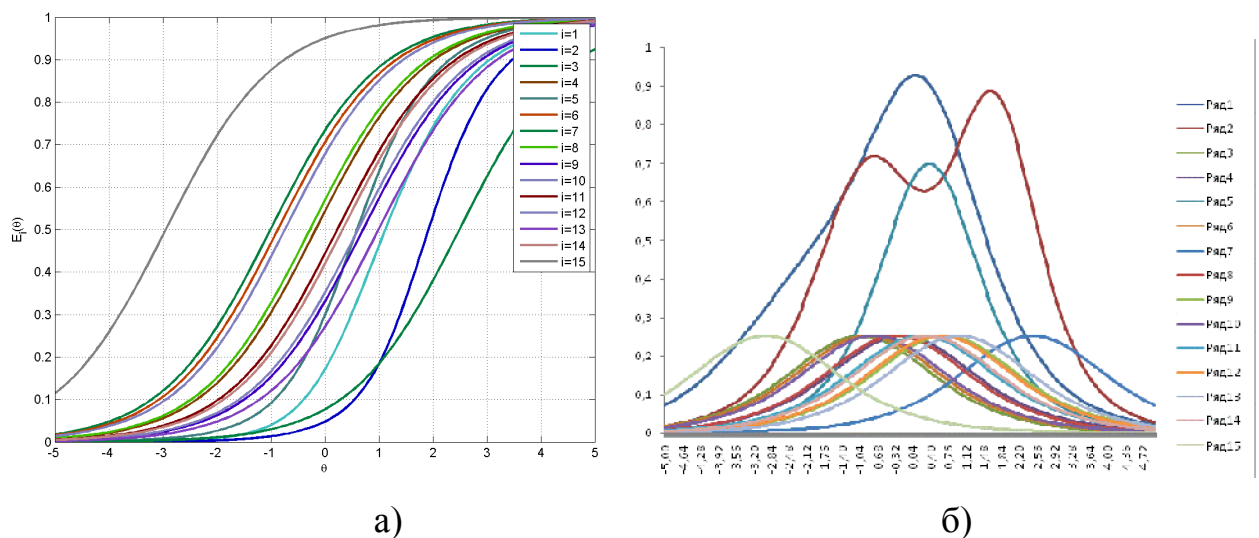


Рис. 3.16. Ансамблі характеристичних та інформаційних кривих завдань тесту «Диференціальне числення функцій однієї змінної»

Завдання №1 є інформативним (Рис. 3.16б) для іспитників з рівнем підготовленості від -3,5 до 2,5 логітів. Аналіз характеристичних кривих рівнів завдання (Рис. 3.17а) з параметрами складності

$\overline{\beta}_1 = (-2,47; -0,58; 0,27; 1,02)$ підтверджує, що останнє є правильно складеним. Текст завдання №1 наведений Рис. 3.17б.

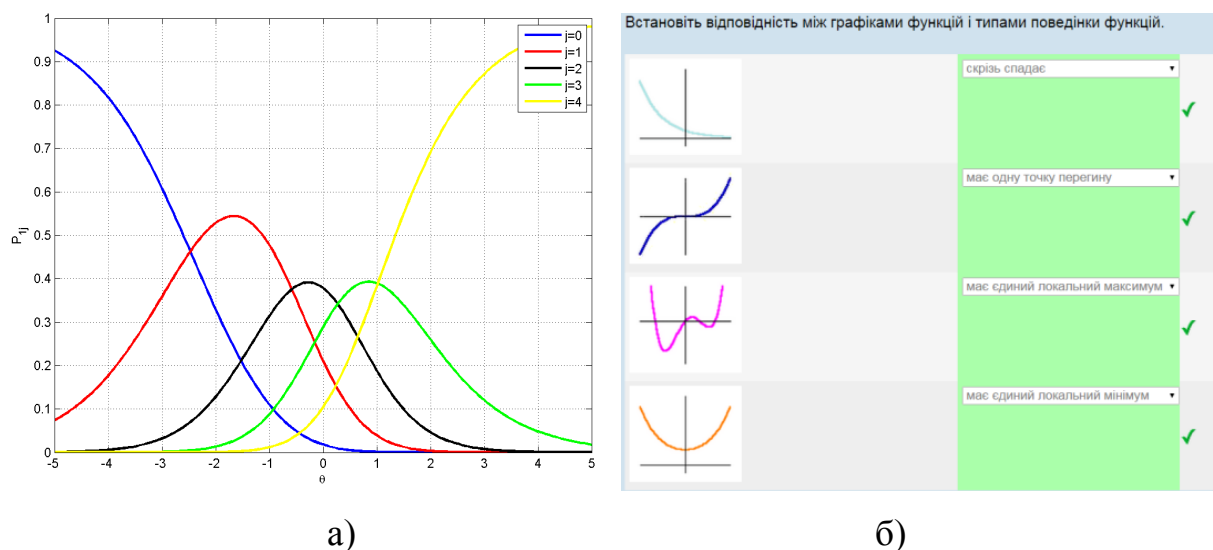


Рис. 3.17. Характеристичні криві рівнів та текст завдання №1 тесту «Диференціальне числення функцій однієї змінної»

Інформаційна крива завдання №2 (Рис. 3.16б) має провал, що означає, що рівень підготовленості студентів з різним значенням θ дане завдання оцінює з різною точністю. «Недосконалість» завдання підтверджено і виглядом характеристичних кривих рівнів, зображених на Рис. 3.17а. Завдання №2 має 5 рівнів з параметрами складності $\overline{\beta}_2 = (0; -0,95; -0,84; 2,3; 1,18)$. Як бачимо 3 рівень цього завдання не спрацьовує (Рис. 3.18а). Отже, це завдання не може бути ефективним для оцінювання рівня підготовленості іспитників з широкого інтервалу θ . Отже, у подальшому доречно зменшити кількість рівнів цього завдання або змінити варіанти першої множини відповідей. Текст завдання №2 наведено на Рис. 3.18б.

Інформаційна крива завдання №5 (Рис. 3.16б) має «загострену» дзвоноподібну форму, що говорить про різну точність оцінювання знань студентів з різними значеннями параметра підготовленості в інтервалі від -1,5 до 2,5 логітів. Розглянемо вигляд характеристичних кривих завдання №5 (Рис. 3.19а). Дане завдання має 3 рівні з параметрами складності

$\bar{\beta}_5 = (0; 0,53; 0,23)$. Перший рівень цього завдання не спрацював. Отже, враховуючи «невдалість» завдання №5, виявлену на 4-ому і 5-ому етапах, його у подальшому доречно замінити іншим. Текст завдання №5 наведено на Рис. 3.19б.

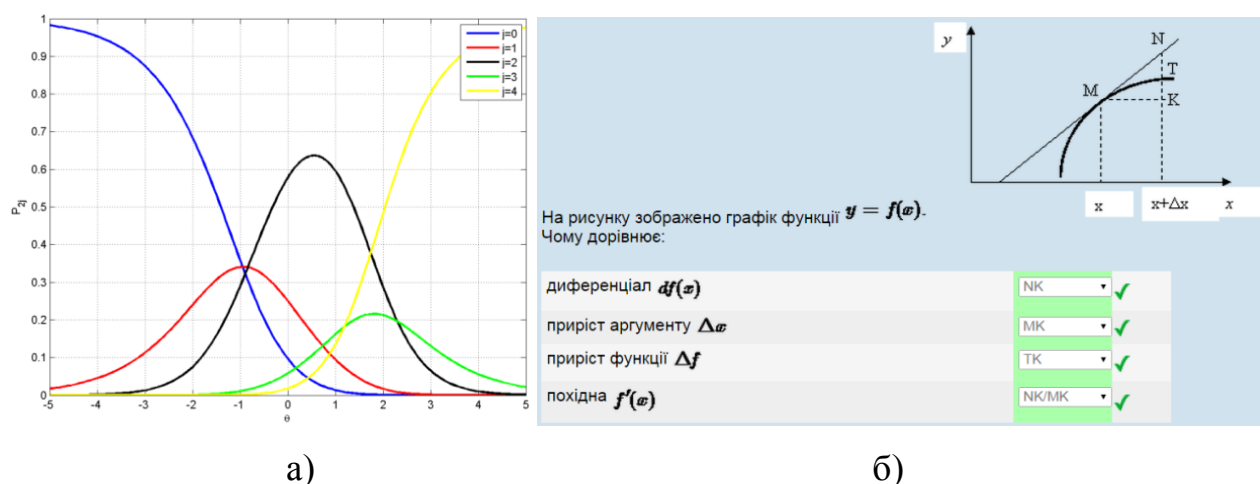


Рис. 3.18. Характеристичні криві рівнів та текст завдання №2 тесту «Диференціальне числення функцій однієї змінної»

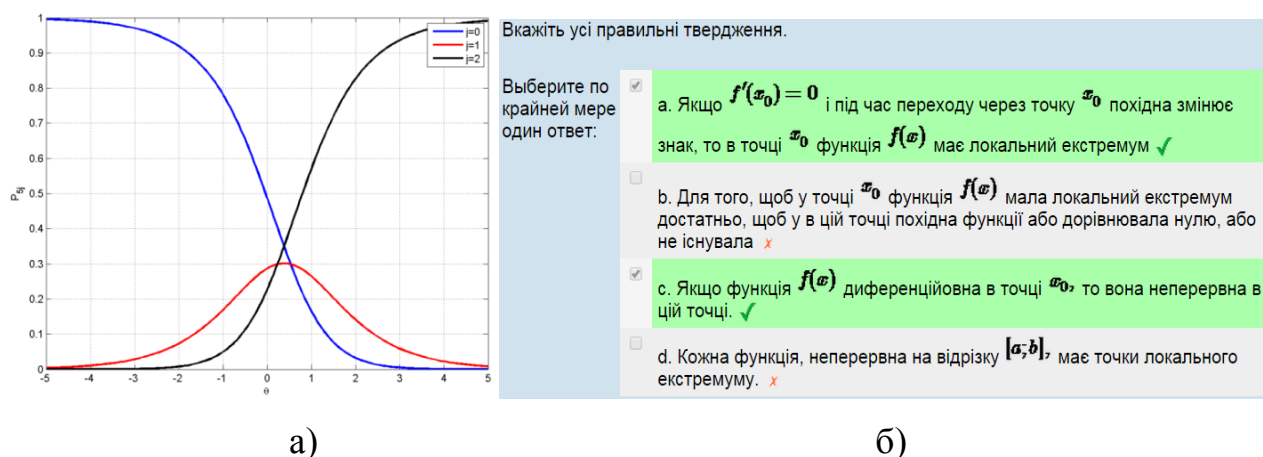


Рис. 3.19. Характеристичні криві рівнів та текст завдання №5 тесту «Диференціальне числення функцій однієї змінної»

Проведено аналіз дистракторів завдань з множинним вибором. Дистрактори, які вибирають менше 5% іспитників, повинні бути видалені або замінені.

У завданні №15 (Рис. 3.12) кожний дистрактор обрали лише 2% іспитників. Отже, це завдання складене неправильно.

У завданні №13 (Рис. 3.20) дистрактор а) обрали лише 4% іспитників, отже, його необхідно замінити іншим.

У завданні №6 (Рис. 3.13) дистрактор а) відповідно обрали 4%. Отже, його необхідно замінити іншим.

13  Баллов: 1

Знайдіть усі асимптоти графіка функції $y = \frac{x^2 - 2x}{x - 1}$.

Выберите один ответ.

- ☐ a. $y = 1, y = x$ ✗
- ☐ b. $x = 1, y = x$ ✗
- ☒ c. $x = 1, y = x - 1$ ✓
- ☐ d. $x = 1$ ✗

Рис. 3.20. Текст завдання №13 тесту «Диференціальне числення функцій однієї змінної»

В завданні №8 (Рис. 3.21) дистрактор b), який обрали лише 4% іспитників, необхідно також замінити іншим, а саме -1.

8  Баллов: 1

Знайдіть y'_x у точці (1;1), якщо $x^4 + x^2 y^2 + y - 3 = 0$.

Выберите один ответ.

- ☐ a. 1 ✗
- ☐ b. 0 ✗
- ☒ c. -2 ✓
- ☐ d. 2 ✗

Рис. 3.21. Текст завдання №8 тесту «Диференціальне числення функцій однієї змінної»

В завданні №3 (Рис. 3.22) дистрактор с) обрали лише 4% іспитників, отже, його необхідно замінити іншим.

3  Баллов: 1

Визначте точки локального екстремуму функції $f(x) = 1 - 3x^2 - x^3$.

Выберите один ответ.

- ☐ a. $x_{\max} = -1, x_{\min} = 0$ ✗
- ☒ b. $x_{\max} = 0, x_{\min} = -2$ ✓
- ☐ c. $x_{\max} = 0, x_{\min} = 1$ ✗
- ☐ d. $x_{\max} = 1, x_{\min} = -2$ ✗

Рис. 3.22. Текст завдання №3 тесту «Диференціальне числення функцій однієї змінної»

Підведемо підсумки щодо якості тесту відповідно до схеми, зображеної на Рис. 3. 3. Коефіцієнти надійності та критеріальної валідності більші за 0,7 графік інформаційної функції має дзвоноподібну не занадто загострену форму, розподіл вибірки індивідуальних балів близький до нормального. Отже, в цілому, тест складений правильно. Найінформативнішим тест є для іспитників з рівнем підготовленості від -2 до 2,5 логітів.

Висновки щодо якості окремих тестових завдань було зроблено відповідно до схеми, зображеної на Рис. 3. 4.

На неправильність завдання множинного вибору №15, що виявилось занадто легким на етапі аналізу ансамблю характеристичних кривих, вказують значення індексу розрізняльної здатності та коефіцієнту бісеріальної кореляції, а також аналіз дистракторів. Тому завдання №15 складено неправильно. Його було замінено на завдання з бази каліброваних завдань з проміжним рівнем складності $\beta = 1,2$ ($\beta \in (-2; -1) \cup (1; 2)$), текст якого зображено на Рис. 3.23.

Знайти похідну y' функції $y = \cos^3(\ln x)$:

- 1) $-\frac{\sin^3(\ln x)}{x}$;
- 2) $-\frac{3\sin^2(\ln x)}{x}$;
- 3) $-\frac{3\cos^2 x \sin x}{x}$;
- 4+) $-\frac{3\cos^2(\ln x)\sin(\ln x)}{x}$.

Рис. 3.23. Текст завдання, на яке було замінено завдання №15 тесту «Диференціальне числення функцій однієї змінної»

На неправильність завдання множинного вибору №12 вказують значення індексу розрізняльної здатності, коефіцієнтів бісеріальної кореляції та точкової бісеріальної кореляції, а також аналіз кореляційної матриці. Не виявлено невдалий дистрактор або помилку в тексті завдання. Отже, завдання №12 складено неправильно. Його було замінено на завдання з бази

каліброваних завдань з проміжним рівнем складності $\beta = 1,5$ ($\beta \in (-2; -1) \cup (1; 2)$), текст якого зображено на. Рис. 3.24.

Знайдіть Тейлорів многочлен $\tilde{P}_n(x)$ 3-го порядку в точці $x = 0$ для функції $f(x) = \ln(1 + x)$.

$$\tilde{P}_n(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

1) $1 - x + x^2 - x^3$

2) $x - x^2 + x^3$

3) $1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3$

+4) $x - \frac{1}{2}x^2 + \frac{1}{3}x^3$

Рис. 3.24. Текст завдання, на яке було замінено завдання №12 тесту «Диференціальне числення функцій однієї змінної»

На неправильність політомічного завдання №5 вказують значення коефіцієнта Пірсона, аналіз кореляційної матриці та сумісний аналіз характеристичних кривих рівнів та інформаційної кривої завдання. Його було замінено на завдання, текст якого зображено на Рис. 3.25.

Вкажіть усі правильні твердження.

+1) Якщо $f''(a)$ додатна, то графік f є опуклим донизу в точці $x = a$.

+2) Якщо графік функції f має точкою перегину $x = a$ і існує скінченна $f''(a)$, то $f''(a) = 0$

3) Якщо $f'(x_0) = 0$, то x_0 — точка екстремуму функції $f(x)$.

4) Кожна функція, неперервна на відрізку $[a; b]$, має точки локального екстремуму.

Рис. 3.25. Текст завдання, на яке було замінено завдання №5 тесту «Диференціальне числення функцій однієї змінної»

На неправильність політомічного завдання №2 вказує аналіз кореляційної матриці. На основі сумісного аналізу характеристичних кривих рівнів та інформаційної кривої завдання прийнято рішення замінити текст завдання з першої множини варіант «диференціал функції $df(x)$ » на «кутовий коефіцієнт MN».

Завданню множинного вибору №6 відповідають низькі значення індексу розрізняльної здатності та коефіцієнту бісеріальної кореляції. Аналіз дистракторів виявив невдалий дистрактор а). Розробити вдалу заміну

дистрактору а) не вдалось, тому було прийнято видалити завдання №6 з тесту. Його було замінено на завдання з бази каліброваних завдань, текст якого зображено на Рис. 3.26.

Нехай відношення $\frac{f(x)}{g(x)}$ є невизначеністю $\frac{0}{0}$ чи $\frac{\infty}{\infty}$, коли $x \rightarrow x_0$. Яке з наступних тверджень є правильним?

+1) якщо існує $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$, то існує $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$ і $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$

2) якщо існує $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$, то існує $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$ і $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$

3) якщо існує $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$, то існує $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$ і $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \left(\frac{f(x)}{g(x)} \right)'$

4) якщо існує $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$, то існує $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$ і $\lim_{x \rightarrow x_0} \left(\frac{f(x)}{g(x)} \right)' = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$

Рис. 3.26. Текст завдання, на яке було замінено завдання №6 тесту «Диференціальне числення функцій однієї змінної»

На основі аналізу дистракторів завдань множинного вибору було прийнято рішення: в завданні № 13 замінити дистрактор а) на дистрактор $x=1; y=x-3$; в завданні №3 замінити дистрактор с) на $x_{\max}=0; x_{\min}=2$; в завданні №8 замінити дистрактор b) на -1 .

Після заміни завдань тесту цикл було повторено: було підтверджено змістову валідність оновленого тесту, тест завантажено у систему MOODLE, проведено тестування та оцінено якість тесту із використанням СААЯТЗ за роробленою методикою. У повторному тестуванні взяло участь 183 студента першого курсу РТФ, ФІОТ та ІТС.

Гістограму частотного розподілу індивідуальних балів представлено на Рис. 3.27. Гіпотеза про нормальність розподілу індивідуальних балів тесту підтверджено за критерієм Пірсона. Мода частотного розподілу індивідуальних балів дорівнює 6,4, середнє вибіркве – 5,9

Обчислено коефіцієнти точкової бісеріальної кореляції між дихотомічними завданнями та індивідуальними балами іспитників і коефіцієнти кореляції Пірсона між політомічними завданням та індивідуальними балами іспитників ρ_j . В якості нижньої границі

припустимих значень коефіцієнта точкової бісеріальної кореляції 0,28. Завдання №5 і №12 потребують детального вивчення: $\rho_5=0,27$ $\rho_{12}=0,25$.

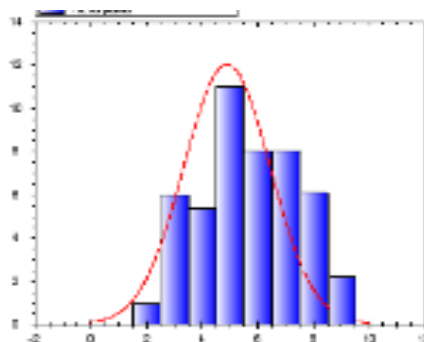


Рис. 3.27. Гістограма частотного розподілу індивідуальних балів оновленого тесту

Коефіцієнт надійності тесту, знайдений на основі методу розщеплення, складає $r_t=0,83$; коефіцієнт надійності Кронбаха $\alpha=0,81$. Отже, можна говорити про високу надійність тесту.

Оцінено критеріальну валідність тесту. Коефіцієнт валідності $\rho=0,85$. Отже, даний тест має високу валідність. Інформаційна функція оновленого тесту, має правильну дзвоноподібну не занадто загострену форму.

Обчислено кореляційну матрицю завдань. (Рис. 3.28). Завдань з переважно від'ємною або занадто високою кореляцією з іншими завданнями не виявлено.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1														
2	0,15	1													
3	0,14	0,22	1												
4	0,36	0,19	0,5	1											
5	0,05	0,34	0,27	0,29	1										
6	0,08	0,2	0,36	0,22	0,19	1									
7	0,29	0,17	0,03	0,2	0,33	0,15	1								
8	0,24	0,09	0,05	0,1	0,13	0,04	0,17	1							
9	0,07	0,36	0,09	0,09	0,05	0,21	0,05	-0,14	1						
10	0,03	0,27	0,02	0,26	0,19	-0,08	0,22	0,32	0,19	1					
11	0,21	0,1	0,05	0,19	0,32	0,03	0,27	0,35	-0,14	0,31	1				
12	-0,2	0,01	0,06	0,1	0,09	-0,07	0,06	0,22	0,05	0,1	-0,11	1			
13	0,2	0,19	0,23	0,36	0,16	0,24	0,36	0,18	0,03	0,22	0,29	0,1	1		
14	0,18	0,23	0,07	0,3	0,09	0,08	0,28	0,17	0,18	0,26	0,38	-0,02	0,25	1	
15	0,09	0,23	0,24	0,26	0,11	0,09	0,11	0,07	0,16	0,2	0,05	0,04	0,16	0,32	1

Рис. 3.28. Кореляційна матриця завдань оновленого тесту

Обчислено показники валідності окремих завдань тесту (Рис. 3.29). Всі завдання тесту мають високу валідність.

Показники валідності окремих завдань тесту															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Dj	0,15	0,54	0,38	0,54	0,54	0,54	0,77	0,69	0,62	0,38	0,77	0,69	0,31		
pj	0,32	0,53	0,43	0,62	0,46	0,4	0,52	0,43	0,32	0,56	0,43	0,25	0,59	0,59	0,45
pbisj	0,38	0,44	0,37	0,54	0,38	0,49	0,43	0,58	0,28	0,66	0,7	0,81			
t<=pbisj	0,32	0,2	0,2	0,19	0,19	0,18	0,18	0,18	0,18	0,18	0,18	0,18	0,18	0,21	
pj	0,961	0,804	0,784	0,765	0,686	0,667	0,588	0,569	0,51	0,49	0,431	0,37			

Умовні позначки:

- червоний в заголовці - політомічне
- зелений в заголовці - дихотомічне
- зелений - завдання добре функціонує
- жовтий - потребує деяку корекцію
- сірий - потребує перегляду
- червоний при Dj - завдання критичне
- червоний - не відповідає умовам

Рис. 3.29. Показники валідності завдань оновленого тесту

Побудовано ансамбль характеристичних кривих завдань тесту (Рис. 3.30). Можна побачити, що характеристичні криві досить рівномірно покривають інтервал від -5 до 5 логітів.

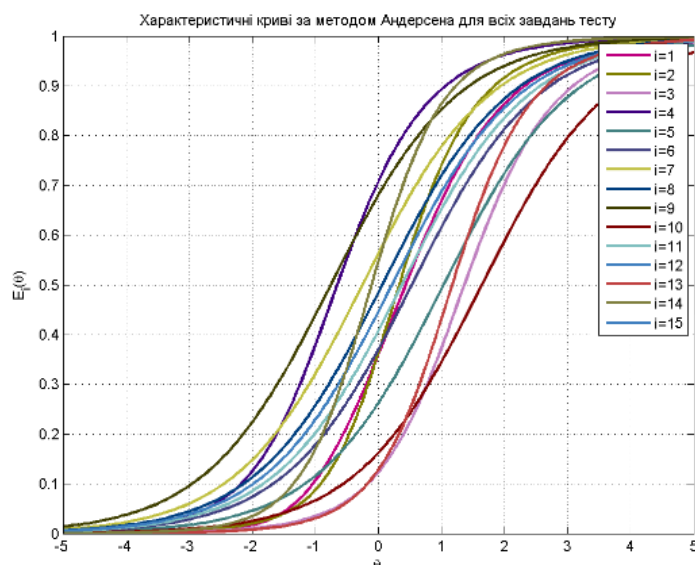


Рис. 3.30. Ансамбль характеристичних кривих завдань оновленого тесту

Аналіз характеристичних кривих рівнів політомічних завдань підтвердив, що завдання складені правильно, всі рівні завдань спрацьовують. Характеристичні криві змінених політомічних завдань № 2 та № 5 зображено відповідно на Рис. 3.31 та Рис. 3.32

Отже, можна зробити висновок, що застосування розробленої комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики до тесту з теми «Диференціальне числення функцій однієї змінної» дало змогу підвищити надійність, валідність та ефективність тесту у цілому, а також позбутись неправильно складених тестових завдань. Оновлений тест складено правильно – його можна застосовувати для оцінювання знань студентів з вищої математики.



Рис. 3.31. Характеристичні криві рівнів завдання №2 оновленого тесту

Аналіз дистракторів завдань множинного вибору підтвердив відсутність невдалих дистракторів.

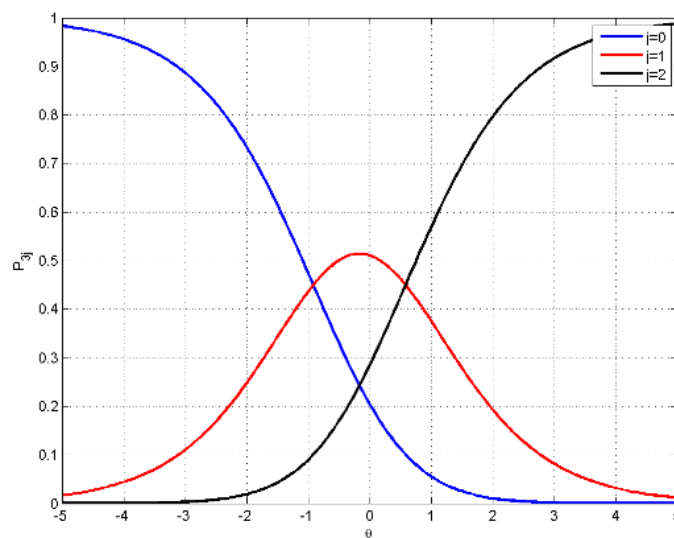


Рис. 3.32. Характеристичні криві рівнів завдання №5 оновленого тесту

Застосування комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики до пакету тестів з вищої математики для студентів технічних спеціальностей КПІ ім. Ігоря Сікорського дало змогу оцінити та покращити їх якість.

3.4. Критерії та рівні розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики

Процес реформування системи освіти, пошук шляхів покращення контролю знань та впровадження ІКТ для визначення рівня навчальних досягнень студентів призводять до необхідності виділення нових складових професійних компетентностей викладачів, зокрема компетентності щодо оцінювання якості тестів з вищої математики [54].

Дослідники визначають професійну компетентність по-різному: як оволодіння людиною здатністю й уміннями виконувати визначені професійні функції (А. К. Маркова) [111]; як наявність спеціальної освіти, глибокої загальної й спеціальної ерудиції, постійне підвищення власної науково-професійної підготовки (В. Г. Зазикін та А. П. Чернишова) [74]; як потенційну готовність розв'язувати задачі зі знанням справи (П. В. Симонов) [141]. В. Адольф стверджує, що "професійна компетентність – це складне утворення, що вміщує комплекс знань, умінь, властивостей і якостей особистості, що забезпечують варіативність, оптимальність та ефективність побудови навчально-виховного процесу" [9, с. 118].

Л. М. Мітіна пов'язує педагогічну компетентність із гармонійним поєднанням знання предмета, методики й дидактики викладання, умінь і навичок (культури) педагогічного спілкування, а також прийомів і засобів саморозвитку, самовдосконалення, самореалізації [116]. І. Колесникова вважає, що педагогічна компетентність – це інтегральна професійно-особистісна характеристика, що зумовлює готовність і здатність виконувати педагогічні функції згідно прийнятих у соціумі в конкретно-історичний момент норм, стандартів та вимог [97].

Аналізуючи ці поняття, можна зробити висновок, що в умовах використання тестів з вищої математики для контролю знань студентів компетентність викладача щодо оцінювання якості розроблених тестів є невід'ємною складовою його професійних компетентностей. Під компетентністю викладача щодо оцінювання якості тестів з вищої

математики розумітимемо готовність та здатність викладача із застосуванням ІКТ обчислювати характеристики тесту, тестових завдань та на основі цих характеристик оцінювати якість окремих завдань і об'єктивно скласти висновки щодо якості тесту в цілому, його покращення та доцільності використання в освітньому процесі для контролю знань студентів з вищої математики.

Застосування тестів з вищої математики для контролю знань студентів вимагає такого рівня компетентності викладача щодо оцінювання цих тестів, що забезпечить високі показники надійності, валідності та ефективності розроблених тестів. Тому необхідним є з'ясування критеріїв, завдяки яким визначаються рівні компетентності щодо оцінювання якості тестів з вищої математики. Зауважимо, що при визначенні характеристик та критеріїв компетентності викладачів щодо оцінювання якості тестів виходили з підходів, що використані при розробці Міжнародної стандартної класифікації освіти (МСКО) [184] та Європейської рамки кваліфікацій (ЄРК) [177].

Під час вибору критеріїв компетентності щодо оцінювання якості тестів з вищої математики прийнято до уваги зміст компонентів структури компетентності: ціннісно-мотиваційний, когнітивний та операційно-діяльнісний. Тобто критеріями у визначенні рівнів розвитку компетентності викладачів щодо оцінювання якості тестів слугують критерії рівнів розвитку її компонентів [85].

Для аналізу рівня розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики розглянемо зміст кожного з її компонентів.

Ціннісно-мотиваційний компонент компетентності викладачів щодо оцінювання якості тестів з вищої математики включає мотиви, цілі, мету, потреби щодо застосування комп'ютерних тестів для аналізу знань студентів, саморозвиток та ін. Компонент передбачає наявність інтересу у викладача до своєї професійної діяльності та в опануванні ефективних засобів її організації.

Для визначення характеристик рівнів цього компонента розглянемо поняття “мотивація”. Є. П. Ільїн [81], А. К. Маркова [112], Г. І. Щукіна [164] та ін вивчали чинники формування навчальної мотивації, її рівні, характер та етапи становлення.

Л. С. Виготський [29], С. С. Канюк [88], О. М. Леонтьєв [106], В. А. Болотов і В. В. Серіков [23] розглядають мотивацію як обов’язкову умову розвитку особистості. А. Маслоу [115] відокремлює одну з основних причин мотивації – потребу в самовираженні особистості. Під мотивом будемо розуміти спонукання людини до активності, пов’язане з прагненням задовольнити конкретні потреби.

Зазначимо, що усвідомлення викладачами значущості використання тестів та аналізі їхньої якості проявляється у ціннісному ставленні та зацікавленості до підвищення ефективності своєї роботи та покращення рівня навчальних досягнень студентів. Ціннісно-мотиваційний компонент передбачає наявність у викладача мети покращення контролю знань студентів

Відповідно до характеристик ціннісно-мотиваційного компоненту у Таблиця 3.3 представлено його характеристики для трьох рівнів: базового, поглибленого та професійного.

Таблиця 3.3

**Характеристика рівнів ціннісно-мотиваційного компонента
компетентності викладачів щодо оцінювання якості тестів з вищої
математики**

<i>Рівні</i>	<i>Характеристика</i>
Базовий	Усвідомлення викладачем потреби у необхідності використання тестів з вищої математики та оцінюванні їх якості з метою покращення контролю знань студентів, прагнення і готовність до отримання знань, умінь і навичок у галузі аналізу якості тестів.

Поглиблений	Вмотивованість до створення валідних надійних тестів з вищої математики, до використання комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики та до покращення створених тестів; мотивація досягнення успіху в професійній діяльності на основі використання тестів з вищої математики та методики оцінювання їхньої якості, прагнення отримати визнання у своїх колег; націленість на досягнення високого рівня компетентності щодо оцінювання якості тестів з вищої математики.
Професійний	Відповідальність та готовність до активного застосування комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики для розроблення пакету валідних та надійних тестів з вищої математики, для підвищення професійного світогляду відповідно до вимог інформаційного суспільства; прагнення досягти покращення контролю знань, а, отже, і підвищення рівня навчальних досягнень студентів.

Когнітивний компонент спирається на знання в галузі оцінювання якості тестів з вищої математики з використанням статистичних методів.

До базових когнітивних умінь Евсеєв А.І. відносить уміння спостерігати, мислити логічно, аналізувати, зіставляти і протиставляти факти та явища, систематизувати та класифікувати, структурувати, оцінювати навчальний матеріал, робити узагальнення і висновки, встановлювати причинно-наслідкові зв'язки, проводити уявний експеримент тощо [69].

Відповідно до зазначених характеристик когнітивного компонента, надамо його характеристики для трьох рівнів (базового, поглибленого та професійного), що представлено у Таблиця 3.4.

Операційно-діяльнісний компонент є процесуальною сутністю використання статистичних методів для розроблення, оцінювання та покращення якості комп'ютерних тестів з вищої математики. Він охоплює вміння та навички щодо застосування набутих знань при розробленні пакету тестів з вищої математики для контролю знань студентів.

**Характеристика рівнів когнітивного компонента компетентності
викладачів щодо оцінювання якості тестів з вищої математики**

<i>Рівні</i>	<i>Характеристика</i>
Базовий	Теоретичні знання про основні поняття та методи оцінювання якості тестів з вищої математики; знання, що дають змогу викладачеві з використанням окремих елементів КТТ на низькому рівні оцінювати якість тестів.
Поглиблений	Знання, що дають змогу викладачеві оцінювати якість тестів з вищої математики на середньому рівні з використанням окремих методів КТТ та IRT; вміння робити висновки на основі використаних методів.
Професійний	Знання, що дають змогу викладачеві, використовуючи взаємодоповнюючі методи КТТ та IRT, оцінювати якість тестів з високим ступенем професійності, вміння систематизувати отримані данні та робити на їх основі висновки щодо подальшої роботи з тестом.

В. А. Крутецький розглядає діяльність як активність людини, спрямовану на досягнення свідомо поставлених цілей, пов'язаних із задоволенням її потреб та інтересів, на виконання вимог до нього з боку суспільства і держави [100]. Діяльність людини формує і змінює її саму, її здібності. В процесі діяльності знання можуть бути осмислені та впорядковані. Тільки діяльність може бути індикатором набутих знань і стимулом до оволодіння новими знаннями [27]. Н. Н. Михайлова, В. Н. Орлова і Г. М. Шеламова відзначають, що знання і вміння є об'єктивними характеристиками діяльності вчителя, а ціннісні установки і особистісні властивості – суб'єктивними характеристиками викладача, які необхідні для його відповідності вимогам професії [119].

Відповідно до зазначених характеристик операційно-діяльнісного компонента, визначено його характеристики для трьох рівнів (базового, поглибленого та професійного), які представлено у Таблиця 3.5.

**Характеристика рівнів операційно-діяльнісного компонента
компетентності викладачів щодо оцінювання якості тестів з вищої
математики**

<i>Рівні</i>	<i>Характеристика</i>
Базовий	Усвідомлення й розуміння викладачем відповідності рівня своїх можливостей і вмінь оцінювання якості тестів з вищої математики. Вміння використовувати окремі методи КТТ для оцінювання якості тестів.
Поглиблений	Вміння оцінювати якість тестів з вищої математики з використанням окремих методів КТТ та IRT; вміння робити висновки на основі використаних методів щодо подальшої роботи з тестом;
Професійний	Здатність самостійно розробляти, оцінювати та покращувати тести з вищої математики, забезпечуючи їхню високу валідність та надійність.

При впровадженні методики комп'ютерно орієнтованого оцінювання якості тестів з вищої математики особливого значення набуває орієнтація на професійний рівень компетентності викладачів щодо оцінювання якості тестів, важливий для професійного зростання та покращення контролю знань студентів.

3.5. Модель та технологія розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики

Розвиток сучасної системи освіти, впровадження тестування в контроль знань як засобу визначення рівня навчальних досягнень студентів призводять до необхідності забезпечення якості тестів, що використовуються. Це, в свою чергу, вимагає підвищення рівня компетентності викладачів щодо оцінювання якості тестів з вищої математики. Про важливість підвищення кваліфікації педагогічних працівників зазначено у Законі України «Про освіту» у ст. 47 [133]. У «Національній доктрині розвитку освіти України» [122] визначено, що важливою умовою модернізації освіти є підготовка,

перепідготовка та підвищення кваліфікації педагогічних працівників шляхом опанування педагогічними кадрами сучасних інформаційних технологій.

Процес підвищення кваліфікації з метою розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики можна забезпечити такими шляхами:

1) додати до змісту курсів Інститутів післядипломної педагогічної освіти спецкурс, спрямований на розвиток компетентності викладачів щодо оцінювання якості тестів з вищої математики;

2) організовувати на базі ЗВО семінари, майстер-класи або тренінги для викладачів, спрямовані на розвиток компетентності щодо оцінювання якості тестів з вищої математики.

Підвищення кваліфікації викладачів щодо оцінювання якості тестів з вищої математики вимагає розроблення відповідної моделі розвитку цієї компетентності. Поняття “модель” (від лат. *modulus* – міра, аналог, зразок) у словнику іншомовних слів [<http://www.jnsn.com.ua/cgi-bin/u/book/sis.pl?Qry=%CC%E4%E5%EB%FC>] визначається як зразок, примірник чого-небудь; взірець; схема для пояснення якогось явища або процесу

У процесі розроблення моделі розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики спирались на визначення поняття «моделі» В. Ю. Биковим, як деяке подання системи, яка проектується та відображає особливості й властивості цієї системи, що забезпечують досягнення цілей побудови та використання моделі [21].

Згідно із загальною теорією структури діяльності запропонована модель розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики, яка складається з таких компонентів [82, 87, 101]: мотиваційно-цільового; змістового; операційно-технологічного; діагностичного; результативного. Модель розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики представлено на Рис. 3.33.

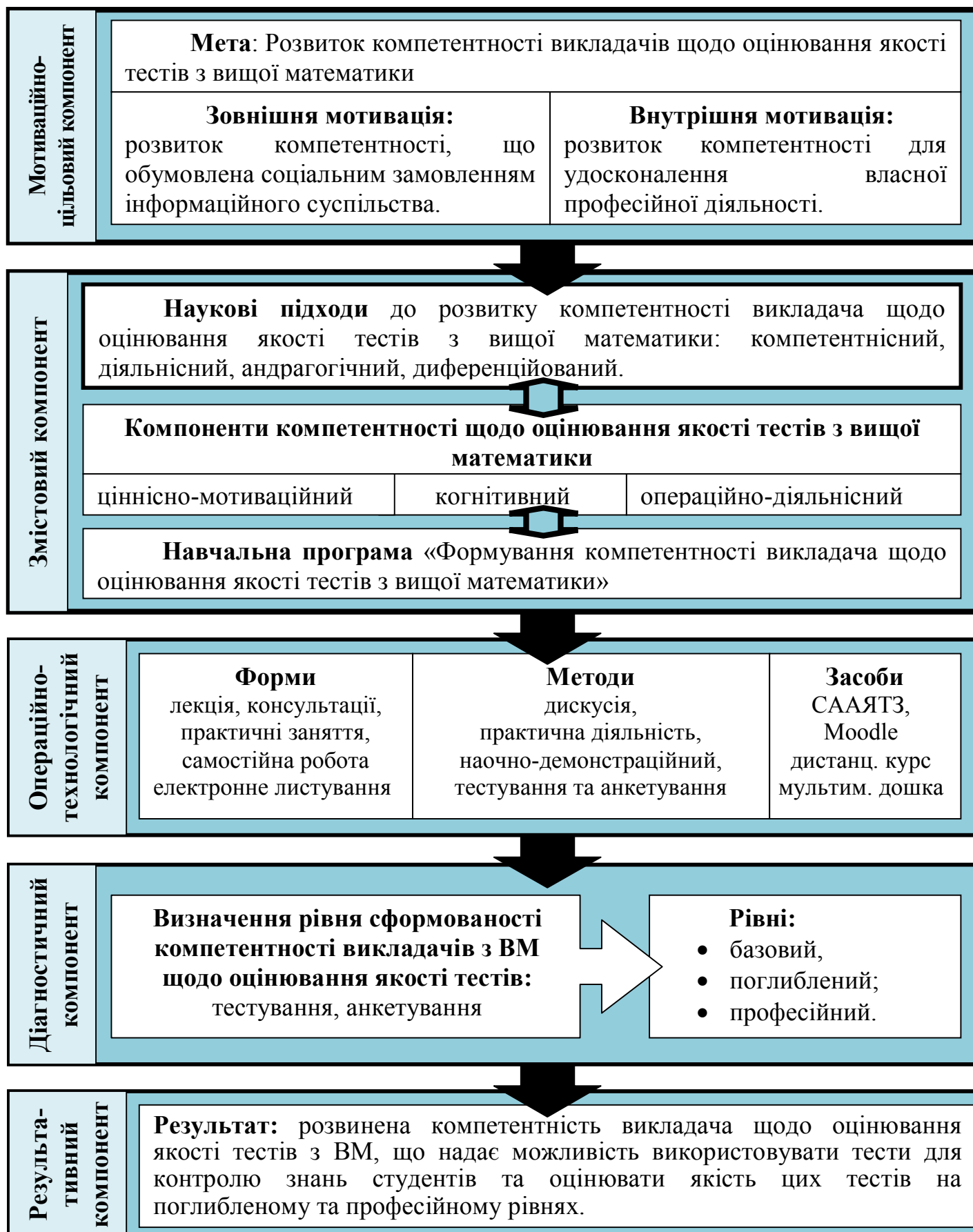


Рис. 3.33. Модель розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики

Мотиваційно-цільовий компонент включає мету: розвиток компетентності викладачів щодо оцінювання якості тестів з вищої математики, що базується на вимогах інформатизації суспільства. Відповідно до розробленої нами моделі мотиваційно-цільовий компонент складається з зовнішньої та внутрішньої мотивації. Зовнішньою мотивацією є мотивація розвитку власної компетентності щодо оцінювання якості тестів з вищої математики, що обумовлена соціальним замовленням інформаційного суспільства. Внутрішньої – є мотивація розвитку компетентності викладача щодо оцінювання якості тестів з вищої математики для удосконалення власної професійної діяльності та для покращення якості контролю.

Змістовий компонент включає наукові підходи до розвитку компетентності викладача щодо оцінювання якості тестів з вищої математики, а саме: компетентнісний, діяльнісний, андрагогічний, диференційований.

Компетентнісний підхід передбачає розвиток компетентностей як головних якостей особистості сучасного суспільства та включає результативність навчання, поєднання інтелектуальної та наукової складових освіти, реалізацію отриманих знань, вмінь, досвіду і способів поведінки в умовах конкретної діяльності. Компетентнісний підхід акцентує увагу на результатах освіти, причому як результат освіти розглядається не сума засвоєної інформації, а здатність людини діяти у різних проблемних ситуаціях [158].

Діяльнісний підхід передбачає розвиток компетентності викладача щодо оцінювання якості тестів з вищої математики та застосування здобутих ним знань, умінь та навичок безпосередньо у професійній діяльності. Такий підхід розглядає діяльність викладача як сукупність взаємозв'язаних елементів: мотивація, мета, завдання, зміст, методи, форми, результат, рефлексія. Ці елементи складають повний цикл діяльності.

Реалізація *андрагогічного підходу* є невід'ємною складовою розвитку компетентності викладача щодо оцінювання якості тестів з вищої

математики. В українському педагогічному словнику поняття «андрагогіка» визначається як педагогіка дорослих, що досліджує проблеми освіти, самоосвіти та виховання дорослих [34].

В процесі реалізації моделі розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики нами було враховано такі андрагогічні підходи:

- пріоритетність самостійного навчання, яке забезпечує ознайомлення з навчальним матеріалом у вільному режимі та дає змогу осмислити та опанувати як процеси, так і технології їх виконання;
- значимість результатів навчання для практичної діяльності викладача, що забезпечується попереднім визначенням його потреб;
- самонавчання та саморозвиток у цілому, що базуються на мотивації особистості до навчання та професійного зростання;
- співпраця, що базується на взаємоповазі та комунікаційних здібностях учасників навчання;
- поєднання індивідуальних та групових форм навчальної роботи, що базується на колективному характері навчального процесу та враховує індивідуально-психологічні особливості слухачів;
- системність навчання, в основі якого лежить відповідність мети і змісту навчання його формам, методам, засобам та оцінюванню результатів. Реалізується шляхом поступового поглиблення теоретичного та практичного матеріалу в програмі навчання. З іншого боку, системність може розглядати як регулярність навчання;
- принцип наочності в навчанні сприяє підвищенню ефективності навчання, виявленню зв'язків між знаннями та їх практичним застосуванням. Він покращує процес засвоєння знань та мотивує до пізнання;
- актуалізація результатів навчання, тобто їх вчасне використання на практиці, вимагає надання не лише знань певного обсягу навчального матеріалу, але і створення умов для опанування вміннь та навичок використання отриманих знань на практиці.

При навчанні дорослих також необхідно враховувати такі вікові особливості, які можуть ускладнити навчання:

- складність зміни статусу викладача на статус та роль того, кого навчають;
- бар'єри упереджень щодо нововведень та змін;
- складність відмови від застарілої системи професійних знань та умінь, отриманих раніше;
- внутрішня невпевненість при заниженій професійній самооцінці та при необхідності зміни своєї діяльності з урахуванням нових вимог.

Диференційований підхід орієнтований на професійні потреби кожного з учасників навчання. Він передбачає створення комфортних режимів пізнавальної діяльності дорослих із різним рівнем знань з метою забезпечення ефективного засвоєння навчального матеріалу в процесі фронтальних, групових та індивідуальних форм роботи на кожному етапі заняття. Даний підхід побудовано на принципах врахування стажу роботи, врахування індивідуальних потреб та врахування специфіки галузі науки фахівця. В роботі Н. Г. Протасової [135, с. 90-91] зазначено, що «мета післядипломної освіти полягає не в усередненні рівня розвитку слухачів, а в допомозі кожному досягти найвищого для себе рівня. Тому диференціація змісту і методики навчання – це єдиний можливий шлях досягнення його індивідуалізації».

Зокрема, при навчанні викладачів було враховано їх розуміння математико-статистичного апарату та вміння роботи з ПК, що дало змогу зосередити увагу саме на практичному оцінюванні якості тестів без втрати часу на детальне роз'яснення математичних аспектів теорії. Зазначимо, що перелічені вище підходи є взаємодоповнюючими та їх потрібно враховувати при навчанні дорослих.

Змістовий компонент моделі забезпечується навчальною програмою курсу «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики» (Додаток В), розробленою та апробованою нами

в Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського».

Операційно-технологічний компонент моделі базується на формах навчання(лекція, практичні заняття, індивідуальні та групові консультації, , електронне листування та ін.), методах навчання(візуалізована лекція, дискусія, самостійна діяльність, практична діяльність, наочно-демонстраційний метод, оцінювання знань, умінь та навичок за допомогою тестування та анкетування) і засобах навчання (ПК, Moodle, комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики, зокрема СААЯТЗ, методичні матеріали: навчальну програму (Додаток В), дистанційний курс за даною програмою, мультимедійна дошка та ін), що відповідають віковій категорії учасників навчання.

Розроблена модель розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики, побудована на діяльнісній основі, що включає такі елементи: мета – мотив – об'єкт – проблема – зразок дії – результат – самоконтроль – оцінка – корекція.

В моделі розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики використовуємо модульний підхід до навчання, засновником якого є Дж. Д. Рассел [200]. Він визначає поняття «модуль» як один навчальний пакет, що охоплює концептуальну одиницю навчального матеріалу.

Форми методи та засоби розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики розробляються відповідно до Положення про підвищення кваліфікації та стажування педагогічних і науково-педагогічних працівників ЗВО (наказ №48 МОН України від 24.01.2013). Під час підвищення кваліфікації використовується дві форми навчання: очна та дистанційна та відповідні методи.

Одним із основних засобів навчання є дистанційний науковий курс, створений в Moodle відповідно до програми «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики»,

структурований за темами програми навчання і складається з: ментальної карти: конспектів лекцій, практичних робіт, самостійних робіт, тестів для самоперевірки: системи контролю та оцінювання знань тощо. Moodle – система підтримки дистанційного навчання, яка надає викладачам та студентам розвинутий набір інструментів для комп'ютеризованого навчання. Це безкоштовна відкрита система управління навчанням, яка отримала надзвичайно широке розповсюдження в світовому освітньому просторі [152].

Діагностичний компонент моделі окреслює компоненти та рівні розвитку компетентності викладача щодо оцінювання якості тестів з вищої математики, визначені та описані у п. 3.4; включає методи оцінювання компетентності викладача щодо оцінювання якості тестів з вищої математики на базовому, поглибленому та професійному рівнях: тестування, анкетування, індивідуальні завдання. За умови отримання недостатнього рівня розвитку компетентності викладачів щодо оцінювання якості тестів модель передбачає внесення змін та доповнень в організацію навчального процесу.

Результативний компонент моделі. Результатом моделі розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики є сформована компетентність викладачів щодо оцінювання якості тестів з вищої математики, що дає змогу використовувати тести для контролю знань студентів та оцінювати якість цих тестів на поглибленому та професійному рівнях, що свідчить про досягнення поставленої мети.

Процес розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики є досить складним, тому вимагає розроблення відповідної технології його здійснення. Розроблена модель розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики слугує підґрунтям для розроблення такої технології.

Технологія розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики включає мету, зміст, форми, методи та засоби навчання, орієнтовані на досягнення запланованого результату.

Метою навчання викладачів є розвиток їх компетентності щодо оцінювання якості тестів з вищої математики.

Завданнями навчання є надання слухачам теоретичних знань з питань оцінювання якості тестів при використанні тестів для контролю знань з вищої математики та розуміння ролі такого оцінювання; розвиток вмінь та навичок оцінювання якості тестів з вищої математики на основі комп'ютерно орієнтованої методики оцінювання якості тестів.

Зміст і організація процесу розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики були розроблені на модульній основі з урахуванням професійних та особистісних вимог слухачів відповідно до навчальної програми «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики». Дана навчальна програма складається з п'яти модулів.

Модуль 1. «Проблеми розроблення тестів з вищої математики» передбачає опанування тем: «Використання тестів для контролю знань студентів», «Форми тестових завдань», «Проблема вибору підходу до розроблення тестів з вищої математики», «Питання оцінювання змістової валідності».

Очікуваний результат від опанування змісту модуля 1:

- ✓ слухач усвідомить переваги використання тестів для контролю знань студентів з вищої математики;
- ✓ слухач з'ясує основні проблеми, пов'язані із процесом розроблення тестів з вищої математики, та дізнається про методи їх розв'язання;
- ✓ слухач усвідомить роль оцінювання якості тестів з вищої математики при їхньому використанні для контролю знань студентів;
- ✓ слухач отримає знання основних принципів розроблення завдань різних форм з вищої математики;
- ✓ слухач навчиться створювати завдання з вищої математики різних форм та типів інтерфейсу у системі Moodle.

Модуль 2. «Компетентність викладача щодо оцінювання якості тестів з вищої математики» передбачає опанування тем: «Поняття компетентності», «Компетентнісний підхід у навчанні», «Компоненти та рівні розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики».

Очікуваний результат від опанування змісту модуля 2:

- ✓ слухач ознайомиться з компетентнісним підходом у навчанні;
- ✓ слухач зрозуміє поняття компетентності;
- ✓ слухач отримає знання компонент та рівней розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики.

Модуль 3. «Статистичні методи оцінювання якості педагогічних тестів» передбачає опанування тем: «Оцінювання якості тестів на основі методів КТТ», «Оцінювання якості тестів з використанням моделей IRT», «Оцінювання ефективності тестів методами IRT».

Очікуваний результат від опанування змісту модуля 3:

- ✓ слухач засвоїть базові теоретичні відомості про оцінювання якості тестів з вищої математики методами КТТ щодо: аналізу розподілу вибірки результатів; оцінювання надійності тесту; оцінювання критеріальної валідності тесту, аналізу кореляційної матриці завдань; аналізу дистракторів завдань множинного вибору;
- ✓ слухач засвоїть базові теоретичні відомості про оцінювання якості тестів з вищої математики методами IRT щодо: застосування моделей Раша та Бірнбаума (для дихотомічних завдань), Андерсенена для політомічних завдань та Тіссена-Стейнберга для завдань множинного вибору; оцінювання значень латентних параметрів використовуваних IRT-моделей; аналізу ансамблю характеристичних кривих завдань тесту; аналізу характеристичних кривих рівнів політомічних тестових завдань; аналізу характеристичних кривих дистракторів завдань множинного вибору;

- ✓ слухач засвоїть базові теоретичні відомості про оцінювання та підвищення ефективності тестових завдань в рамках IRT із використанням інформаційної функції;

- ✓ слухач навчиться робити висновки про якість тестів і тестових завдань на основі взаємодоповнюючих методів IRT та КТТ.

Модуль 4. «Система автоматизованого аналізу якості тестових завдань» передбачає опанування тем: «Функціональні можливості СААЯТЗ», «База каліброваних завдань для комп'ютерних тестів з вищої математики»

Очікуваний результат від опанування змісту модуля:

- ✓ слухач із використанням СААЯТЗ навчиться оцінювати якість тестів з вищої математики методами КТТ, а саме: аналізувати розподіл вибірки результатів; оцінювати надійність тесту; оцінювати критеріальну валідність тесту; обчислювати та аналізувати кореляційну матрицю завдань; оцінювати валідність окремих завдань, обчислюючи значення індексу розрізняльної здатності дихотомічних завдань D_j , коефіцієнт точкової бісеріальної кореляції та бісеріального коефіцієнту кореляції між дихотомічним завданням та індивідуальними балами іспитників і коефіцієнт кореляції Пірсона між політомічним завданням та індивідуальними балами іспитників $\rho_{j.}$;

- ✓ слухач із використанням СААЯТЗ навчиться оцінювати якість тестів з вищої математики методами IRT, а саме: оцінювати значення латентних параметрів моделей Раша та Бірнбаума (для дихотомічних завдань), Андерсенена для політомічних завдань та Тіссена-Стейнберга для завдань множинного вигляду; будувати та аналізувати ансамбль характеристичних кривих завдань тесту; будувати та аналізувати характеристичні криві рівнів політомічних тестових завдань за моделлю Андерсена; будувати та аналізувати характеристичні криві дистракторів завдань множинного вибору за моделлю Тіссена-Стейнберга;

- ✓ слухач опанує навички роботи з базою каліброваних завдань, а саме: створювати теми тестування; створювати та видаляти тести та тестові

завдання, упорядковуючи їх за темами; зберігати результати аналізу тестів у базу каліброваних завдань; переглядати інформації про тест та тестові завдання; проводити пошук та тестових завдань за заданими характеристиками для зручного складання нових тестів; проводити пошук тестів за темами та призначенням.

Модуль 5. «Комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики» передбачає опанування тем: «Базові положення методики», «Місце методики в роботі викладача з тестами з вищої математики», «Практичне застосування методики».

Очікуваний результат від опанування змісту модуля 5:

- ✓ слухач отримає знання базових положень методики оцінювання якості тестів з вищої математики;
- ✓ слухач з'ясує місце запропонованої методики в роботі викладача з тестами з вищої математики;
- ✓ слухач із використанням запропонованої методики навчиться оцінювати якість тестів у цілому методами КТТ та IRT на основі: аналізу розподілу вибірки результатів; оцінювання надійності тесту; оцінювання критеріальної валідності тесту; оцінювання ефективності тесту.
- ✓ слухач із використанням запропонованої методики навчиться оцінювати якість окремих тестових завдань методами КТТ та IRT на основі: аналізу кореляційної матриці завдань; оцінювання валідності окремих завдань; аналізу ансамблю характеристичних кривих завдань; аналізу політомічних завдань; аналізу дистракторів завдань множинного вибору;
- ✓ слухач навчиться у цілому застосовувати комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики до результатів тестування.

Узагальнюючи вищезазначене, можна зробити висновок, що вивчення модуля 1 спрямовано на розуміння питань, пов'язаних безпосередньо із розробленням тестів; модуля 2 – на розуміння ролі оцінюванн якості тестів та підвищення мотивації викладачів; модуля 3 – на засвоєння базових

теоретичних відомостей щодо оцінювання якості тестів методами КТТ та IRT; модуля 4 – на формування вміння роботи із СААЯТЗ; модуля 5 – на комплексне застосування комп'ютерно орієнтованої методики до оцінювання якості тестів на основі знань та вмінь, отриманих у попередніх модулях.

Згідно програми курсу навчання викладачів розпочинається вхідним моніторингом та визначенням рівня розвитку їх компетентності щодо оцінювання якості тестів з вищої математики. Після вхідного моніторингу викладачі розподіляються на групи відповідно до рівня розвитку компетентності.

Загальні вимоги до знань та вмінь, які повинен мати викладач вищої математики до проходження навчання:

1. Наявність знань про загальні принципи розроблення та використання комп'ютерних тестів з вищої математики.
2. Наявність знань про специфіку контролю знань студентів із використанням комп'ютерних тестів. Розуміння ролі і місця комп'ютерних тестів для контролю знань студентів.
3. Наявність уявлень про можливості інформатизації педагогічної діяльності, зокрема використання засобів ІКТ для контролю знань студентів.
4. Наявність базових знань та умінь роботи в системі Moodle.
5. Наявність загальних уявлень про системи комп'ютерного тестування.
6. Наявність базової ІК-компетентності, що включає, загальні уявлення щодо призначення та функціонування персональних комп'ютерів та технологій Інтернет.

Технологія відповідно до розробленої моделі ґрунтується на андрагогічному, диференційованому, компетентнісному та діяльнісному *підходах*, що застосовуються при навчанні дорослих.

При навчанні використовуються такі *засоби*: ПК, комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики, зокрема СААЯТЗ, мультимедійна дошка, система Moodle, дистанційний курс «Формування компетентності викладачів щодо оцінювання якості тестів з

вищої математики» та ін. Перед проходженням навчання СААЯТЗ необхідно встановити на ПК кожного слухача.

Навчання проводиться за допомогою групових *форм навчання* (лекції, практичні заняття, групові консультації) та індивідуальних (індивідуальні заняття, самостійна робота, індивідуальні консультації, електронне листування). Схематично організація форм навчання у межах однієї теми модуля представлено на Рис. 3.34

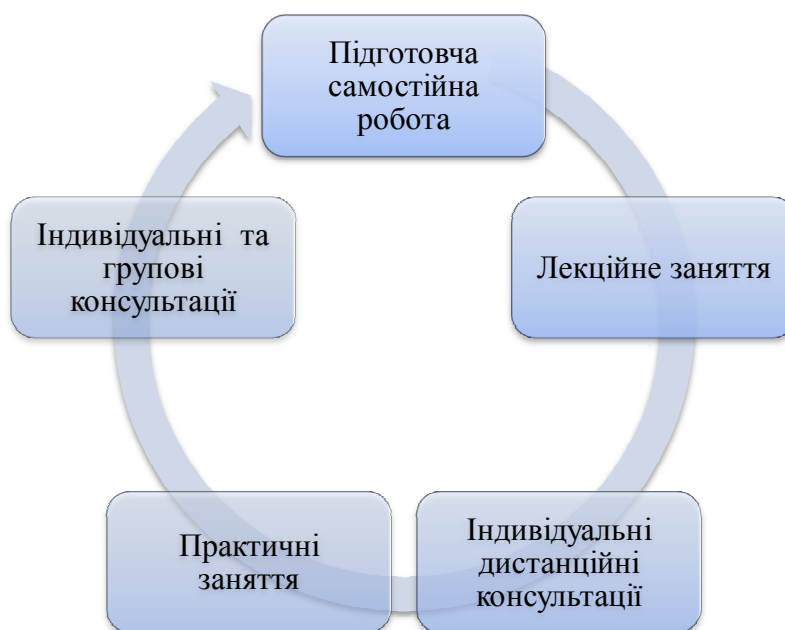


Рис. 3.34. Організація форм навчання у межах однієї теми модуля

Розпочинається цикл занять із підготовчої самостійної позааудиторної роботи, яка полягає у перегляді теоретичного матеріалу теми. Теоретичний матеріал розміщено у дистанційному курсі «Формування компетентності викладача щодо оцінювання якості тестів з вищої математики» у системі Moodle.

Поглиблене опрацювання теоретичного матеріалу проходить на лекційних заняттях. Лекції проводяться в аудиторіях, що оснащено необхідними засобами навчання: ПК з підключенням до Інтернет-мережі, мультимедійна дошка тощо.

Під час лекцій відбуваються:

- пояснення теоретичного матеріалу, що викликав труднощі при самостійному опрацюванні;
- обговорення теоретичного матеріалу в межах зазначеної теми;
- демонстрація за допомогою мультимедійної дошки прикладів використання набутих теоретичних знань на практиці;
- поточний контроль опрацювання теоретичного матеріалу.

Важливо на лекціях використовувати інтерактивні методи, які передбачають діалог, дискусію тощо, під час яких відбувається взаємодія між учасниками навчального процесу з метою взаєморозуміння, спільного вирішення навчальних завдань та розвитку особистісних умінь та навичок слухачів. Дискусія передбачає спільне обговорення та аналіз конкретної проблеми, питання або завдання. Дискусію можна зкерувати за допомогою поставлених запитань та тем для обговорення, або вона може бути неструктурованою, тобто її хід буде залежити від учасників. Один із кращих методів отримання ідей вирішення питання є поєднання групової дискусії з методом мозкового штурму. Мозковий штурм складається з двох етапів. На першому етапі генерації ідей усі учасники навчання мають право висувати свої пропозиції щодо вирішення поставленого завдання. Критику ідей при цьому заборонено. На другому етапі аналізуються всі ідеї та обґрунтовується правильне рішення.

Лекції варто починати з опитування слухачів для перевірки розуміння основних теоретичних положень матеріалу, поданого на самостійне опрацювання. При виявленні прогалин необхідно детальніше зупинитися на поясненні матеріалу, що викликав труднощі. Після пояснень доцільно надати можливість слухачам задати додаткові питання. Далі необхідно переходити до демонстрації прикладів практичного застосування опрацьованого теоретичного матеріалу. Для цього можна запропонувати практичні задачі, які буде вирішено за допомогою дискусії або мозкового штурму. В кінці лекції проводиться контроль опрацювання теоретичного матеріалу. Контроль саме в кінці лекції сприяє підвищенню мотивації та

активізації учасників навчання упродовж лекції. По завершенню кожного модуля відбувається комплексне оцінювання знань слухачів із застосуванням тестів.

Після лекції проходять індивідуальні навчальні консультації за певним графіком в очній та дистанційній формі за допомогою Інтернет сервісів, зокрема, електронною поштою та Skype. Їх, в основному, призначено для викладачів, які не змогли з поважних причин бути присутніми на лекціях і у яких є прогалини у розумінні данного теоретичного матеріалу. Специфіка консультацій полягає у їх добровільному відвідуванні учасниками навчання.

Наступним етапом циклу є практичне заняття, яке зорієнтовано на поглиблення та розширення здобутих на лекціях знань. Практичні заняття проводяться в аудиторіях, що оснащено необхідними засобами навчання: ПК з підключенням до Інтернет-мережі, мультимедійна дошка тощо. На практичних заняттях слухачі можуть користуватись і власними ПК. Для деяких тем практичне заняття не передбачено. В межах практичних занять проводиться:

- роз'яснення змісту індивідуальних завдань;
- контроль виконання індивідуальних завдань;
- контроль опрацювання теоретичного матеріалу;
- розгляд окремих ситуацій оцінювання якості тестових завдань та тестів, що викликали труднощі при виконанні практичних завдань.

На практичних заняттях слухачі отримують індивідуальні завдання, виконання яких спрямоване на опанування вмінь та навичок використання на практиці отриманих теоретичних знань. На початку практичного заняття надаються необхідні вказівки до їх виконання. Протягом проведення практичного заняття викладач має приділяти час індивідуально кожному слухачу: оцінити виконану роботу або ж допомогти. На основі виконаних індивідуальних завдань відбувається контроль знань теоретичного матеріалу, а також набуття вмінь та навичок розв'язання практичних задач.

По завершенню вивчення всіх модулів, окрім першого, слухачі отримують підсумкові індивідуальні роботи.

Після практичних занять проходять групові навчальні консультації за певним графіком в очній формі та індивідуальні навчальні консультації у дистанційній формі за допомоги електронної пошти та Skype. Їх призначено для викладачів, які потребують додаткового пояснення матеріалу.

Методами навчання є дискусія, практична діяльність, самостійна діяльність, наочно-демонстраційний метод, мозковий штурм, візуалізована лекція, оцінювання знань, умінь та навичок за допомогою тестування та анкетування.

Діагностика набутих слухачами знань, умінь та навичок проводиться за допомогою тестування та анкетування на комплексній основі, яка складається з таких компонентів:

- оцінювання результатів виконання поточних завдань до кожного з модулів курсу;
- оцінювання результатів підсумкового тесту, анкети та індивідуального завдання, що дають змогу оцінити рівень сформованості компетентності викладача щодо оцінювання якості тестів з вищої математики на кінець навчання;
- оцінювання випускної роботи, яка полягає у розробці власного тесту з вищої математики та оцінюванні його якості із використанням комп'ютерно орієнтованої методики.

Очікуваними результатами реалізації навчальної програми є:

- усвідомлення слухачами переваг використання тестів для контролю знань студентів з вищої математики;
- знання слухачів основних проблем, пов'язаних із процесом розроблення тестів з вищої математики, та вміння їх вирішувати;
- з'ясування та усвідомлення слухачами ролі оцінювання якості тестів з вищої математики при їхньому використанні для контролю знань студентів.

- теоретичні знання слухачів питань оцінювання якості тестів методами КТТ та IRT;
- знання слухачів базового функціоналу СААЯТЗ;
- вміння та навички слухачів роботи з СААЯТЗ;
- вміння та навички слухачів використання комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики у цілому.

Описану вище методику впроваджено на курсах підвищення кваліфікації науково-педагогічних працівників за програмою «Формування компетентності щодо оцінювання якості тестів з вищої математики» (<http://uiite.kpi.ua/ua/regulatory-framework/documents/math-test.html>) в Навчально-методичному комплексі «Інститут післядипломної освіти» КПІ ім. Ігоря Сікорського. Запропонована технологія сприятиме розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики, що підтверджується експериментальним дослідженням у розділі 4.1.

Висновки до розділу 3

1. Досліджено проблему вибору підходу до розроблення тестів, яка полягає у визначенні місця тестів з вищої математики серед критеріально-орієнтованих та нормативно-орієнтованих тестів. Враховуючи специфіку тестів з вищої математики, тенденції сучасного періоду розвитку освіти та запити ЗВО, зроблено висновки про необхідність об'єднання критеріально-орієнтованих та нормативно-орієнтованого підходів при оцінюванні навчальних досягнень студентів. Саме за такої реалізації інтерпретація результатів тесту по відношенню до певного змісту навчальної дисципліни не знімає можливості диференціації індивідуальних відмінностей студентів.

2. Досліджено питання оцінювання змістової валідності тестів з вищої математики. Зроблено висновок, що для забезпечення високої змістової валідності розробник тесту повинен вирішити такі задачі:

розробити специфікації тесту; відібрати групи експертів, компетентних у змістовій області; оцінити змістову валідність створеного тесту. Оцінювання змістової валідності тесту проводиться за методом Делфі та передбачає: організацію та проведення експертизи зі встановлення відповідності між тестовими завданнями і змістовою областю виконання тесту; приймання узагальнюючих висновків розробником щодо поліпшення змісту тесту. Технологія експертизи якості змісту тесту включає три напрямки роботи експертів: оцінювання завдань тесту; оцінювання тесту у цілому; оформлення узагальнюючих висновків і рекомендацій щодо поліпшення змісту тесту. У якості засобу проведення експертизи тестів з вищої математики нами був обраний Google Docs додаток.

3. Розроблено комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики. *Метою* комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики є оцінювання якості тесту та визначення подальших дій щодо його покращення. *Засобом* комп'ютерно орієнтованої методики оцінювання якості тестів є СААЯТЗ. *Методами* методики є ретельно дібрані методи теорії ймовірностей та математичної статистики, а саме методи КТТ та ІРТ

Загальна схема методики включає в себе такі етапи: формування таблиці результатів; аналіз окремих тестових завдань і тесту в цілому; приймання остаточних рішень. Формування таблиці результатів тестування передбачає підготовку даних для подальшої роботи з ними в СААЯТЗ. Аналіз тесту у цілому полягає у такому: аналізі розподілу вибірки результатів; оцінюванні надійності тесту; оцінюванні критеріальної валідності тесту; оцінюванні ефективності тесту. Аналіз тестових завдань передбачає: аналіз кореляційної матриці завдань; оцінювання валідності окремих завдань; аналіз ансамблю характеристичних кривих завдань; аналіз політомічних завдань; аналіз дистракторів завдань множинного вибору. Приймання остаточних рішень здійснюється на підставі сумісного аналізу

взаємодоповнюючих характеристик IRT та КТТ. При цьому робляться висновки щодо якості як окремих тестових завдань, так і тесту в цілому.

4. Визначено особливості роботи викладачів ЗВО з тестами з вищої математики для оцінювання знань студентів. Розроблено загальну схему роботи викладача з тестами з вищої математики, яка передбачає: вирішення питаннями визначення підходу до розроблення тестів; створення тесту; оцінювання змістової валідності; завантаження тестів у систему комп'ютерного тестування (Moodle); проведення пробного тестування та вирішення питань, пов'язаних з його організацією та проведенням; імпортування результатів тестування та оцінювання якості тестів за розробленою методикою; прийняття рішення щодо якості тесту та подальшої роботи з ним; збереження тесту та завдання тесту разом з їх статистичними характеристиками у базі каліброваних завдань. Робота з тестами за описаною схемою вимагає від викладача високого рівня компетентності щодо оцінювання якості тестів з вищої математики.

5. Описано практичне застосування розробленої методики оцінювання якості тестів з вищої математики. У якості контрольного прикладу використано модульну контрольну роботу на тему «Диференціальне числення функцій однієї змінної», яку було проведено у формі комп'ютерного тесту для студентів першого курсу РТФ, ФІОТ та ІТС.

6. Застосування комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики до пакету тестів з вищої математики для студентів технічних спеціальностей КПІ ім. Ігоря Сікорського дало змогу оцінити та покращити їх якість.

7. Визначено компетентність викладача щодо оцінювання якості тестів з вищої математики як готовність та здатність викладача із застосуванням ІКТ обчислювати характеристики тесту, тестових завдань та на основі цих характеристик оцінювати якість окремих завдань і об'єктивно складати висновки щодо якості тесту в цілому, його покращення та доцільності використання в освітньому процесі для контролю знань студентів

з вищої математики. Визначено та описано основні критерії компетентності викладачів щодо оцінювання якості тестів з вищої математики: ціннісно-мотиваційний, когнітивний та операційно-діяльнісний. Представлено характеристики визначених критеріїв для трьох рівнів: базового, поглибленого та професійного.

8. Розроблено модель розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики, яка складається з таких компонентів: мотиваційно-цільового, змістового, операційно-технологічного, діагностичного та результативного.

Мотиваційно-цільовий компонент включає мету: розвиток компетентності викладачів щодо оцінювання якості тестів з вищої математики, що базується на вимогах інформатизації суспільства. Змістовий компонент включає наукові підходи до розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики, а саме: компетентнісний, діяльнісний, андрагогічний, диференційований. Змістовий компонент моделі забезпечується навчальною програмою «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики». Операційно-технологічний компонент моделі базується на формах (лекція, практичні заняття, індивідуальні та групові консультації, електронне листування та ін.), методах (візуалізована лекція, дискусія, самостійна діяльність, практична діяльність, наочно-демонстраційний метод, оцінювання знань, умінь та навичок за допомогою тестування та анкетування) і засобах навчання (Moodle, ІКТ, комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики, зокрема СААЯТЗ, методичні матеріали: навчальну програму, дистанційний курс та ін.). Діагностичний компонент моделі включає методи оцінювання компетентності викладачів щодо оцінювання якості тестів з вищої математики: тестування, анкетування. Результатом моделі розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики є сформована компетентність викладачів щодо оцінювання якості тестів з вищої математики, що дає змогу використовувати

тести для контролю знань студентів та оцінювати якість цих тестів на поглибленому та професійному рівнях, що свідчить про досягнення поставленої мети.

9. Розроблено технологію розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики на основі відповідної моделі. *Метою* навчання викладачів є розвиток їх компетентності щодо оцінювання якості тестів з вищої математики. *Завданнями* навчання є надання слухачам теоретичних знань з питань оцінювання якості тестів при використанні тестів для контролю знань з вищої математики та розуміння ролі такого оцінювання; розвиток вмінь та навичок оцінювання якості тестів з вищої математики.

Зміст і організація процесу розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики розроблено на модульній основі з урахуванням професійних та особистісних вимог слухачів відповідно до навчальної програми «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики». Навчання проводиться за допомогою групових *форм навчання* (лекції, практичні заняття, групові консультації) та індивідуальних (індивідуальні консультації, електронне листування, самостійні завдання). При цьому використовуються такі *засоби навчання*: ПК, комп'ютерно орієнтована методика оцінювання якості тестів з вищої математики, зокрема СААЯТЗ, мультимедійна дошка, система Moodle, дистанційний курс «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики» та ін. *Методами* навчання є візуалізована лекція, дискусія, самостійна діяльність, практична діяльність, наочно-демонстраційний метод, мозковий штурм, оцінювання знань, умінь та навичок за допомогою тестування та анкетування.

Результати третього розділу дисертаційного дослідження висвітлено у наукових роботах автора [40, 48, 51, 56]

РОЗДІЛ IV. ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА КОМП'ЮТЕРНО ОРІЄНТОВАНОЇ МЕТОДИКИ ОЦІНЮВАННЯ ЯКОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ

4.1. Проведення та результати педагогічного експерименту

Експериментальне дослідження полягало у проведенні педагогічного експерименту, метою якого була перевірка ефективності впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики. Під час роботи було враховано основні вимоги до проведення педагогічного експерименту [17].

Етапи експериментального дослідження. Основну роботу експериментального дослідження було спрямовано на перевірку гіпотези та основних теоретичних положень дослідження. Розроблена програма експериментального дослідження передбачала констатувальний (2015-2016 рр.) та формувальний (2016-2017 рр.) етапи (Таблиця 4.1):

1) констатувальний етап передбачав з'ясування стану застосування та оцінювання якості комп'ютерних тестів з вищої математики для контролю знань студентів;

2) формувальний етап включав такі підетапи:

а) проведення констатувальних зрізів у контрольних (КГ) та експериментальних (ЕГ) групах щодо визначення рівнів розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики;

б) проведення навчання відповідно до методики розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики;

в) проведення контрольних зрізів у контрольних та експериментальних групах щодо визначення рівнів розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики;

г) проведення експертного оцінювання контролю знань студентів з вищої математики;

д) здійснення аналізу одержаних результатів експериментальної роботи та їх узагальнення.

Таблиця 4.1

Етапи експериментального дослідження

Назва етапу	Група/ к-ть респондентів		Основні експериментальні зрізи	Методи
Констатувальний, 2015-2016 р.	140		З’ясування стану застосування та оцінювання якості комп’ютерних тестів з вищої математики для контролю знань студентів	Анкетування, статистичний аналіз
Формувальний 2016-2017 р.	КГ, 68	ЕГ, 68	Визначення рівнів розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики	Анкетування, тестові завдання, статистичний аналіз, порівняльний аналіз динаміки змін
		Модель та технологія розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики		
	КГ, 68	ЕГ, 68	Визначення рівнів розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики	
	Експертне оцінювання контролю знань студентів з вищої математики			Анкетування, статистичний аналіз
	Опрацювання та аналіз результатів експериментального дослідження			Статистичний аналіз, систематизація, узагальнення, інтерпретація

У дослідженні було висунуто *гіпотезу*, що виважене впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики сприятиме розвитку у викладачів ЗВО відповідної компетентності за умови реалізації розробленої технології розвитку компетентності та покращенню контролю знань студентів.

Завдання експериментального дослідження. Проведення експерименту включило такі завдання:

1. З'ясувати стану застосування та оцінювання якості комп'ютерних тестів з вищої математики для контролю знань студентів;

2. Визначити рівень сформованості компетентності викладачів щодо оцінювання якості тестів з вищої математики у КГ та ЕГ на початок і кінець педагогічного експерименту за трьома критеріальними компонентами: ціннісно-мотиваційним, когнітивним, операційно-діяльнісним та трьома рівнями: базовим, поглибленим, професійним;

3. З'ясувати шляхом експертного оцінювання стану контролю знань студентів з вищої математики;

4. Провести аналіз отриманих результатів педагогічного експерименту шляхом використання математично-статистичних методів обчислення.

Констатувальний етап. Констатувальний етап експериментального дослідження проводився протягом 2015-2016 років. Мета констатувального етапу експерименту полягала у з'ясуванні стану застосування та оцінювання якості комп'ютерних тестів з вищої математики для контролю знань студентів.

Експериментальну базу дослідження на констатувальному етапі експериментального дослідження склали 140 викладачів КПІ ім. Ігоря Сікорського, Київського національного університету імені Тараса Шевченка та Національного педагогічного університету імені М.П. Драгоманова.

Завдання констатувального етапу експериментального дослідження:

- розроблення анкети для виявлення стану застосування та оцінювання якості комп'ютерних тестів з вищої математики для контролю знань студентів (Додаток А);

- проведення у 2016 році. анкетування викладачів КПІ ім. Ігоря Сікорського, Київського національного університету імені Тараса Шевченка та Національного педагогічного університету імені М.П. Драгоманова;

- аналіз результатів анкетування.

Результати констатувального етапу експериментального дослідження. Враховуючи той факт, що не всі викладачі використовують тести для контролю знань студентів, анкету було розроблено з двох блоків питань. Якщо викладач використовує тести для контролю знань студентів з вищої математики – він відповідав тільки на питання блоку А, якщо не використовує – тільки блоку Б.

Аналіз результати анкетування на констатувальному етапі показав, що тести для оцінювання знань студентів з вищої математики при проміжному або підсумковому контролі використовую лише 21% (або 29 респондентів), всі з яких використовують саме комп'ютерне тестування у системі Moodle. 96% респондентів блоку Б, що не використовують тести для контролю знань студентів, також надають перевагу комп'ютерному тестуванню над бланковим. 56% респондентів блоку Б має досвід роботи із системою комп'ютерного тестування, а саме із системою Moodle.

Використання тестів для оцінювання знань студентів з вищої математики при проміжному та підсумковому контролі лише 17% респондентів блоку Б вважають доцільним, 45% – вважають доцільним використання тестів лише як допоміжного засобу контролю знань, залишивши основним традиційні контрольні роботи, 36% – вважають недоцільним використання тестів, 2% – не змогли дати відповідь на задане запитання (Рис. 4.1).

На запитання «Чому Ви не застосовуєте тестування для контролю знань студентів з вищої математики?» 67% опитуваних блоку Б дали

відповідь «не довіряю тестуванню, як засобу контролю», 100% - «не вмію скласти якісні тести», 34% - «не хочу щось змінювати в контролі знань, оскільки традиційні контрольні роботи мене цілком влаштовують».

Всі респонденти блоку А вважають, що застосування комп'ютерного тестування як засобу контролю знань з вищої математики спрощує процедуру перевірки робіт студентів. 57% вважає, що при цьому підвищується об'єктивність контролю знань студентів з вищої математики.

При цьому лише 19% респондентів блоку Б вважають, що використання тестування сприяє підвищенню об'єктивності контролю знань студентів з вищої математики. Це можна пояснити їх недовірою до тестування, як до засобу контролю. При цьому 86% викладачів цього блоку хотіли б спростити процедуру перевірки контрольних робіт студентів з вищої математики. Враховуючи відповіді викладачів блоку А, це можливо із впровадженням комп'ютерного тестування для контролю знань в освітній процес.

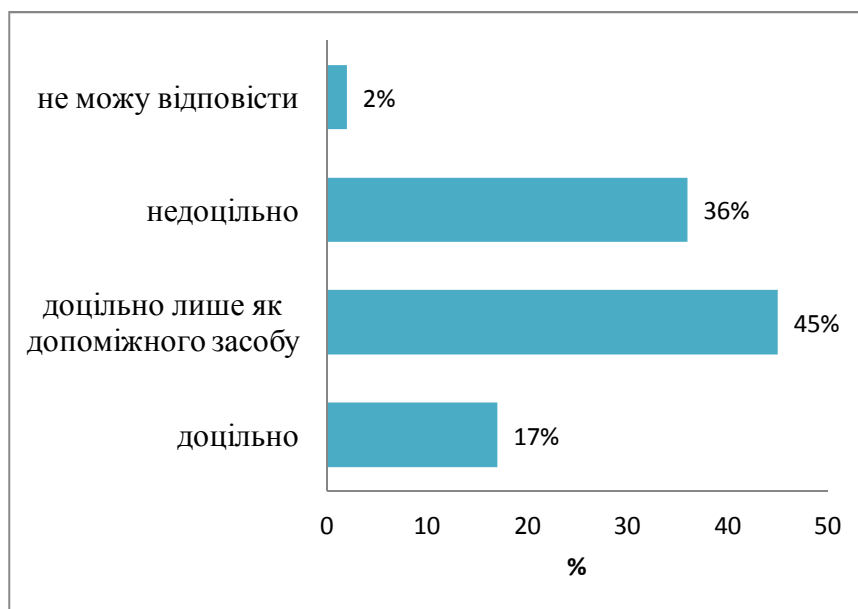


Рис. 4.1 Результати відповідей викладачів на запитання «Чи вважаєте Ви за доцільне використання комп'ютерних тестів для оцінювання знань студентів з вищої математики при проміжному та підсумковому контролі?»

Водночас на запитання «Чи необхідне, на Вашу думку, оцінювання якості тестів з вищої математики перед впровадженням їх в освітній процес?»

більшість викладачів (81%) дали позитивну відповідь, 8% вважають такий аналіз непотрібним, і 11% опитуваних не змогли дати відповідь на запитання (Рис. 4.2).

При цьому тільки п'ятеро респондентів блоку А (17%) для оцінювання якості розроблених тестів використовують класичні статистичні методи; один викладач (3%) використовує метод експертних оцінок, інші 80% не використовують жодних методів. Це можна пояснити, враховуючи відповіді на запитання №7, №9, №13 блоку А та №9 та №12 блоку Б низьким рівнем компетентності викладачів щодо оцінювання якості тестів з вищої математики.

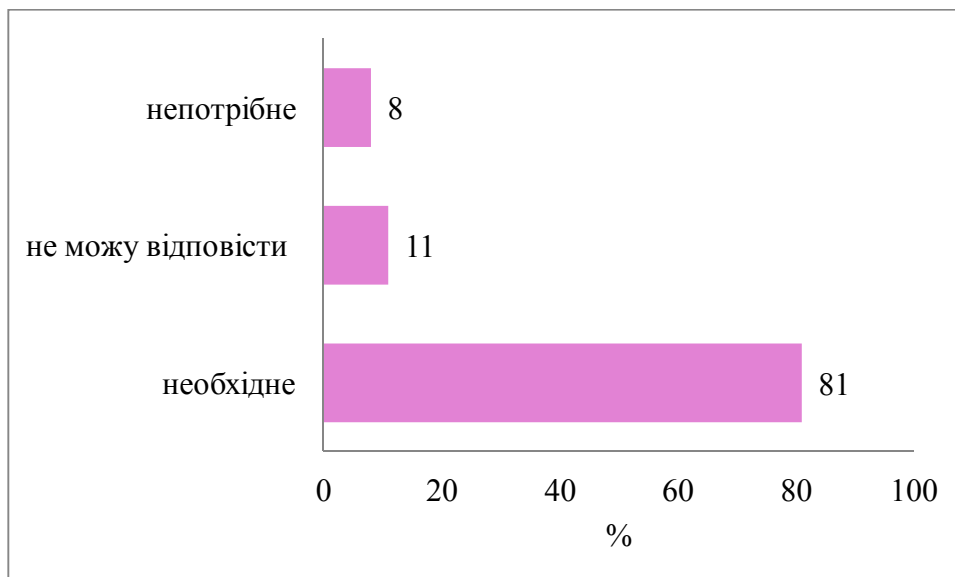


Рис. 4.2. Результати відповідей викладачів на запитання «Чи необхідне, на Вашу думку, оцінювання якості тестів з вищої математики перед впровадженням їх в освітній процес?»

На запитання «Чи вважаєте Ви за потрібне використовувати ІКТ для оцінювання якості тестів з вищої математики?» 92% опитуваних дали позитивну відповідь, 1% - негативну, і 7% не змогли дати відповідь на запитання.

При цьому лише 21 % викладачів на питання «Чи достатньо у Вашому розпорядженні програмних засобів для оцінювання якості тестів з вищої математики?» дали позитивну відповідь. Це ще раз підтверджує необхідність розроблення СААЯТЗ.

На запитання «Чи сприяла б використанню Вами тестування з вищої математики для контролю знань студентів достатня кількість знань щодо оцінювання якості тестів?» дали позитивну відповідь 52% респондентів блоку А. Це свідчить про те, що викладачі обирають традиційні засоби контролю через недовіру до псевдо тестів, а не до тестів у цілому. 71% всіх викладачів вважають за потрібне підвищити рівень компетентності викладачів щодо оцінювання якості тестів з вищої математики.

На запитання блоку А «Чи хотіли б Ви опанувати методику оцінювання якості тестів з вищої математики з використанням ІКТ?» позитивно відповіли 95% викладачів, коли на запитання блоку Б «Чи хотіли б Ви підвищити свій рівень компетентності щодо оцінювання якості тестів з вищої математики??» позитивну відповідь дали лише 28% викладачів. Це можна пояснити недостатньою вмотивованістю викладачів щодо використання тестів для контролю знань в цілому.

Підсумовуючи результати проведеного анкетування, можна зробити такі висновки:

- незважаючи на те, що тестування для контролю знань студентів на практиці використовує незначна частка опитуваних, викладачі не заперечують можливості та доцільності його використання;
- використання тестів є привабливим для викладачів, оскільки дає можливість позбавитись від трудомісткої для нього перевірки традиційних контрольних та самостійних робіт;
- більшість викладачів вважає доцільним поєднання традиційних та тестових форм контролю знань студентів з вищої математики;
- значна більшість викладачів розуміють необхідність оцінювання якості тестів з вищої математики із застосуванням ІКТ перед їхнім впровадженням в освітній процес;
- значна більшість викладачів обирають традиційні засоби контролю через недовіру до псевдо тестів, а не до тестів у цілому;

– значна більшість викладачів вважають за потрібне підвищити загальний рівень компетентності викладачів щодо оцінювання якості тестів з вищої математики;

– має місце недостатня вмотивованість викладачів щодо використання тестів для контролю знань в цілому та щодо навчання оцінювання їх якості.

Формувальний етап. Формувальний етап експериментального дослідження проводився протягом 2016-2017 років. Мета формувального етапу полягала у перевірці гіпотези дослідження.

Експериментальну базу дослідження на формувальному етапі склали 136 викладачів КПІ ім. Ігоря Сікорського, Київського національного університету імені Тараса Шевченка та Національного педагогічного університету імені М.П. Драгоманова. Було сформовано експериментальну та контрольну групи, до яких увійшло по 68 викладачів.

На початку та в кінці формувального етапу у контрольній та експериментальній групах було оцінено рівні розвитку компонентів компетентності викладачів щодо оцінювання якості тестів з вищої математики за базовим, поглибленим та професійним рівнями. З цією метою застосовувались методи анкетування та тестування. Експериментальна група проходила навчання, яке здійснювалось відповідно до технології розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики. Викладачі з контрольної групи навчались самостійно, використовуючи дистанційний курс «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики», розміщений у системі MOODLE, та СААЯТЗ, встановлену на їх персональні комп'ютери.

Метою навчання за програмою «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики» є розвиток компетентності викладачів щодо оцінювання якості тестів з вищої математики.

Завданнями навчання за програмою є надання слухачам теоретичних знань з питань розуміння ролі оцінювання якості тестів при використанні тестів для контролю знань з вищої математики; розвинути вміння та навички оцінювання якості тестів з вищої математики.

При цьому викладачі, які пройшли навчання в ЕГ, були підготовлені як тренери. Такий підхід застосовується спільною програмою МОН України та Корпорацією Intel «Навчання для майбутнього» [183] та описаний у роботах Л. Аройо (L.Aroyo), Г. Вагнера (G.Wagner), Р. Гізарді (R. Guizzardi) та Л. Дулея (L. Dooley), Е. Мартінеза (A.Martinez), Т. Меткалфа (T. Metcalf) [173, 178]. Він передбачає підготовку викладачів за тренінговою формою навчання, що у подальшому може бути трансльована колегам.

На формувальному етапі експерименту проводилось опрацювання даних, співставлення та аналіз результатів дослідження; опис ходу та результатів дослідження на основі статистичних методів опрацювання даних; оформлення результатів. Основним напрямом роботи є порівняння та узагальнення результатів, отриманих на початок та кінець формувального етапу експериментального дослідження в КГ і ЕГ.

Результати формувального етапу експериментального дослідження. Для оцінювання когнітивного компонента компетентності викладачів щодо оцінювання якості тестів з вищої математики було розроблено тест з 20 запитань (Додаток В). Правильна відповідь на одне завдання оцінювалась в 1 бал. Результат тестування і визначав рівень когнітивного критерію компетентності. За допомогою тестових завдань перевірялися знання щодо основних термінів і понять тестології, пов'язані з оцінюванням якості тестів з вищої математики, знання, що дають змогу викладачеві проводити аналіз якості тестів з вищої математики з використанням методів КТТ та IRT.

Опрацювання отриманих результатів на початок та кінець формувального етапу педагогічного експерименту у КГ та ЕГ здійснювалося шляхом переведенням абсолютних значень індивідуальних балів викладачів у відносні за формулою (4.1):

$$\tilde{X}_i = \frac{X_i}{X_{\max}} \cdot 100\%, \quad (4.1)$$

де X_i - індивідуальний бал i -ого викладача за тест;

$\tilde{X}_i$ - відносне значення індивідуального балу i -ого викладача за тест;

$X_{\max}$ - максимальне значення індивідуального балу за тест. Для даного тесту $X_{\max} = 20$.

Значення на початок та кінець етапу педагогічного експерименту у КГ та ЕГ отримано як середні арифметичні відносних значень індивідуальних балів викладачів відповідно у КГ та ЕГ. Динаміка змін когнітивного компонента на початок та кінець педагогічного експерименту між КГ та ЕГ представлено на Рис. 4.3.

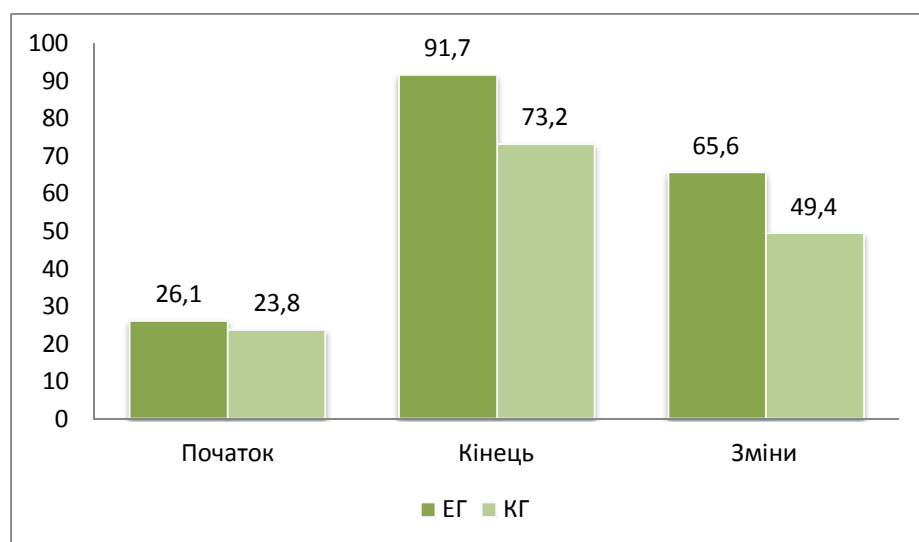


Рис. 4.3. Динаміка змін когнітивного компонента на початок та кінець педагогічного експерименту між КГ та ЕГ (середні арифметичні відносних значень індивідуальних балів)

Для визначення рівнів розвитку когнітивного компонента компетентності викладачів щодо оцінювання якості тестів з вищої математики на формувальному етапі експерименту здійснено порівняння отриманих даних з загальними оцінками рівнів розвитку компонентів компетентності, що наведено у

Таблиця 4 2.

**Загальні оцінки рівнів розвитку компонентів компетентності викладачів
щодо оцінювання якості тестів з вищої математики**

<i>Рівень</i>	<i>Показник</i>
Базовий	30-50%
Поглиблений	51-75%
Професійний	76-100%

Аналіз результатів педагогічного експерименту дозволив дійти такого висновку: на початок педагогічного експерименту рівень розвитку когнітивної складової компетентності викладачів щодо оцінювання якості тестів з вищої математики у КГ становив 24%, у ЕГ 26% і знаходився нижче базового рівня (<30%), а на кінець педагогічного експерименту у КГ групи рівень знань компетентності щодо оцінювання якості тестів з вищої математики зріс до 73% поглибленого рівня (51-75%), а у ЕГ – до 92% (76-100%) професійного рівня.

Динаміка змін між рівнями розвитку когнітивного компонента на 16% вище у ЕГ, ніж у КГ. Підвищення рівня розвитку когнітивного компонента в експериментальній групі обумовлюється тим, що викладачі пройшли навчання за відповідною технологією розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики.

Опрацювання інших компонентів компетентності викладачів щодо оцінювання якості тестів з вищої математики проводилось аналогічно опрацюванню когнітивного компонента.

Для виявлення рівня розвитку ціннісно-мотиваційного компонента компетентності щодо оцінювання якості тестів з вищої математики було проведено анкетування (Додаток Г).

Аналіз результатів педагогічного експерименту дозволив дійти висновку: на початок педагогічного експерименту рівень розвитку ціннісно-

мотиваційного компоненту компетентності щодо оцінювання якості тестів з вищої математики у КГ становив 25 %, у ЕГ 29 % і знаходився нижче базового рівня (<30%), а на кінець педагогічного експерименту у КГ групи рівень розвитку ціннісно-мотиваційного компоненту компетентності щодо оцінювання якості тестів з вищої математики зріс до 45% базового рівня (30-50%), а у ЕГ – до 72 % (51-75%) поглибленого рівня. Динаміка змін між рівнями розвитку ціннісно-мотиваційного компонента на 24 % вище у ЕГ, ніж у КГ.

Динаміку змін рівнів розвитку ціннісно-мотиваційного компонента на початок та кінець педагогічного експерименту між КГ та ЕГ представлено на Рис. 4.4.

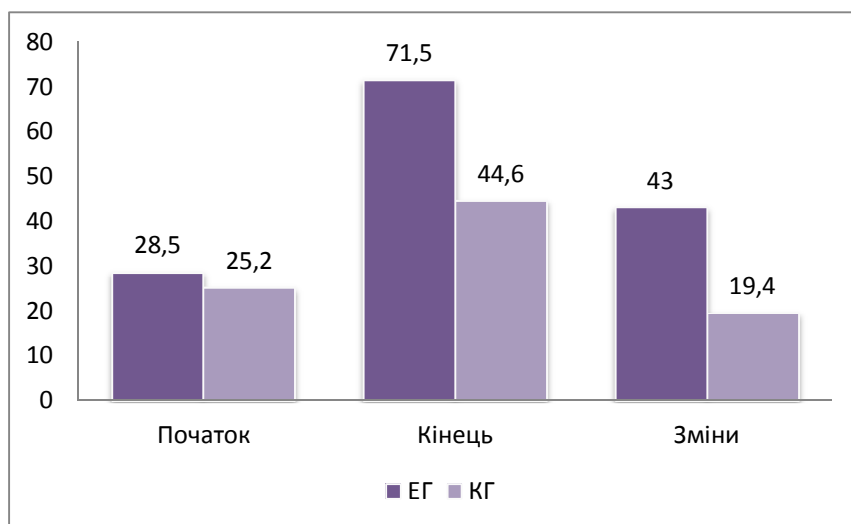


Рис. 4.4. Динаміка змін рівнів розвитку ціннісно-мотиваційного компонента на початок та кінець педагогічного експерименту між КГ та ЕГ (середні арифметичні відносних значень індивідуальних балів)

Здійснення педагогічних впливів в процесі цілеспрямованого навчання на розвиток ціннісно-мотиваційного компонента компетентності викладачів щодо оцінювання якості тестів з вищої математики стимулює зростання зацікавленості щодо використання тестів для оцінювання знань студентів, створює емоційну готовність до вирішення задачі забезпечення високих валідності та надійності розроблених тестів, формує позитивне ставлення викладачів вищої математики до процесу оцінювання якості тестів із використанням ІКТ.

Визначення рівня розвитку операційно-діяльнісного компонента компетентності викладачів щодо оцінювання якості тестів з вищої математики відбувалось на основі результатів індивідуальних завдань (ІЗ) викладачів. ІЗ полягали в оцінюванні якості тесту за змодельованими результатами тестування. В результаті ІЗ перевірялося набуття викладачами вмінь та навичок щодо оцінювання якості тестів з вищої математики. На початку експерименту ІЗ включали перелік підзавдань, які повинен виконати викладач (Додаток Е). Підзавдання було сформовано відповідно до розробленої методики оцінювання якості тестів з вищої математики. У кінці експерименту ІЗ полягало в оцінюванні якості тестів за змодельованими результатами тестування відповідно до розробленої методики.

Аналіз результатів педагогічного експерименту дозволив зробити висновок: на початок педагогічного експерименту рівень розвитку операційно-діяльнісного компонента компетентності щодо оцінювання якості тестів з вищої математики у КГ становив 8 %, у ЕГ 10 %, що нижче базового рівня (<30%), а на кінець педагогічного експерименту в КГ рівень розвитку операційно-діяльнісного компонента компетентності щодо оцінювання якості тестів з вищої математики зріс до 45% і досяг базового рівня (31-50%), а у ЕГ – до 78% (76-100 %) професійного рівня. Динаміка змін між рівнями розвитку ціннісно-мотиваційного компонента на 30% вище у ЕГ, ніж у КГ.

Динаміку змін рівнів розвитку операційно-діяльнісного компонента на початок та кінець педагогічного експерименту в КГ та ЕГ представлено на Рис. 4.5.

Операційно-діяльнісний компонент компетентності викладачів щодо оцінювання якості тестів з вищої математики формується безпосередньо під час цілеспрямованого навчання викладачів і практичної роботи щодо

використання СААЯТЗ для оцінювання якості тестів з вищої математики. Розвиток умінь та навичок складають важливу частину навчання викладачів.

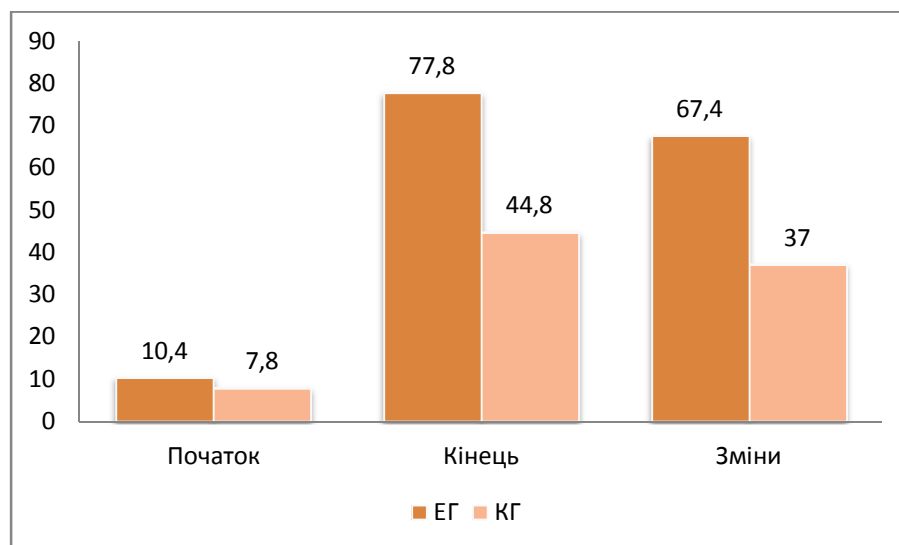


Рис. 4.5. Динаміка змін рівнів розвитку операційно-діяльнісного компонента на початок та кінець педагогічного експерименту в КГ та ЕГ (середні арифметичні відносних значень індивідуальних балів)

З метою підвищення рівня розвитку операційно-діяльнісного компонента компетентності викладачів щодо оцінювання якості тестів з вищої математики в кінці навчання було передбачено виконання викладачами випускної роботи. Дана робота полягала в розробці власного валідного та надійного тесту з вищої математики, проведенні його та оцінюванні якості результатів тестування відповідно до запропонованої методики.

Після опрацювання складових компетентності викладачів щодо оцінювання якості тестів з вищої математики (когнітивного, ціннісно-мотиваційного та операційно-діяльнісного) значення компетентності у цілому визначалося як середнє арифметичне від відповідних значень її складових.

Аналіз результатів педагогічного експерименту щодо всіх складових компетентності викладачів щодо оцінювання якості тестів з вищої математики дозволив зробити висновок: на початок педагогічного експерименту рівень розвитку компетентності викладачів щодо оцінювання

якості тестів з вищої математики в КГ становив 22%, а у ЕГ 19 %, що нижче базового рівня (<30%), а на кінець педагогічного експерименту в КГ групи рівень компетентності щодо оцінювання якості тестів з вищої математики збільшився до 54%, що відповідає поглибленому рівню (51-75 %), а у ЕГ – до 80 %, що відповідає професійному рівню (75-100 %). Динаміка змін між рівнями розвитку компетентності щодо оцінювання якості тестів з вищої математики на 23 % вище у ЕГ, ніж у КГ.

Динаміку змін рівнів розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики на початок та кінець педагогічного експерименту між КГ та ЕГ представлено на Рис. 4.6.

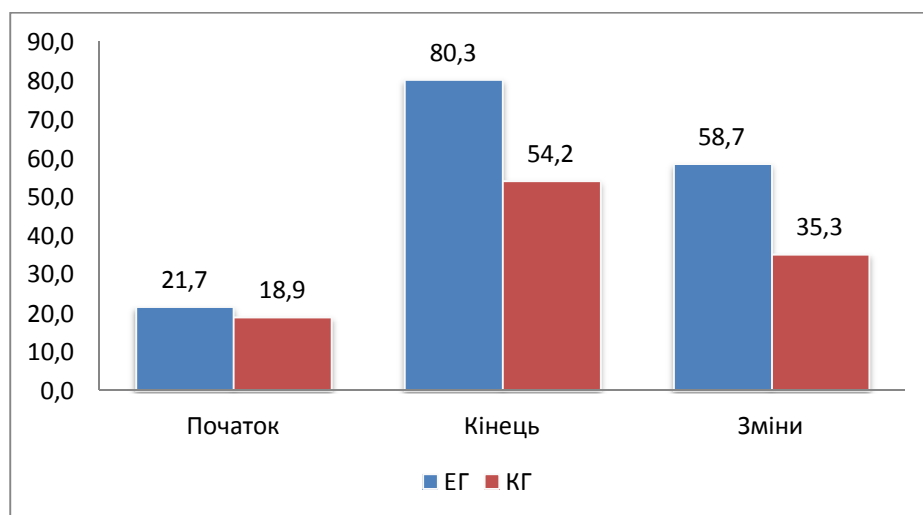


Рис. 4.6. Динаміка змін рівнів розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики на початок та кінець педагогічного експерименту між КГ та ЕГ (середні арифметичні відносних значень індивідуальних балів)

Аналіз результатів педагогічного експерименту дав змогу виявити, що найвищі показники рівня розвитку компетентності отримала операційно-діяльнісна складова компетентності викладачів щодо оцінювання якості тестів з вищої математики.

Для оцінювання достовірності відмінності двох вибірок в дослідженні було обрано метод Фішера (або ϕ^* - критерій) [157].

За кутовим перетворенням Фішера необхідно обчислити величини центрального кута, що відповідають заданим відносним значенням, за формулою

$$\varphi = 2 \arcsin \sqrt{P} \quad (4.2)$$

де P — відносне значення (частка).

Емпіричне значення φ^* визначається за формулою:

$$\varphi^* = (\varphi_1 - \varphi_2) \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

де φ_1 — кут, що відповідає відносним значенням ЕГ,

φ_2 — кут, що відповідає відносним значенням КГ,

n_1 — кількість учасників ЕГ,

n_2 — кількість учасників КГ.

Спочатку за критерієм Фішера було перевірено достовірність гіпотези про відсутність відмінностей між рівнями розвитку компетентності щодо оцінювання якості тестів з вищої математики експериментальних і контрольних груп до проходження навчання. Було визначено такі гіпотези:

H_0 : До проходження навчання викладачі ЕГ мають рівень розвитку компетентності щодо оцінювання якості тестів з вищої математики не більший, ніж викладачі КГ, що навчались самостійно.

H_1 : До проходження навчання викладачі ЕГ мають рівень розвитку компетентності щодо оцінювання якості тестів з вищої математики вищий, ніж викладачі КГ, що навчались самостійно.

За результатами анкетування та тестування було визначено відносні значення за кожним із зазначених рівнів розвитку компонентів компетентності щодо оцінювання якості тестів з вищої математики. Значення кутів на початок експерименту у контрольній та експериментальній групах представлено у Таблиця 4.3.

Аналіз значень кутів у Таблиця 4.3 показав, що емпіричні значення менше критичних, тобто не входять до зони значимості. Таким чином, можна зробити висновок, що значних розбіжностей між показниками у контрольній

та експериментальній групі на початок експерименту немає, тому гіпотеза H_0 підтверджується.

Таблиця 4.3

**Значення кутів для компетентності викладачів
щодо оцінювання якості тестів з вищої математики на початок
експерименту у контрольній та експериментальній групах**

Когнітивний компонент	φ_1 (26,1)	1,07			
	φ_2 (23,8)	1,02			
	φ^*	0,31	$\varphi^* < \varphi_{кр} = 1,64; (p = 0,05)$		
Ціннісно-мотиваційний компонент	φ_1 (28,5)	1,13			
	φ_2 (25,2)	1,05			
	φ^*	0,43	$\varphi^* < \varphi_{кр} = 1,64; (p = 0,05)$		
Операційно-діяльнісний компонент	φ_1 (10,4)	0,66			
	φ_2 (7,8)	0,57			
	φ^*	0,53	$\varphi^* < \varphi_{кр} = 1,64; (p = 0,05)$		
Компетентність викладачів щодо оцінювання якості тестів з ВМ	φ_1 (21,7)	0,97			
	φ_2 (18,9)	0,90			
	φ^*	0,40	$\varphi^* < \varphi_{кр} = 1,64; (p = 0,05)$		

Також за критерієм Фішера було перевірено достовірність гіпотези про наявність відмінностей між рівнями компетентності щодо оцінювання якості тестів з вищої математики експериментальних і контрольних груп за результатами кінцевого зрізу. Було визначено такі гіпотези:

H_0 : Після проходження навчання викладачі ЕГ мають рівень розвитку компетентності щодо оцінювання якості тестів з вищої математики не більший, ніж викладачі КГ, що навчались самостійно.

H_1 : Після проходження навчання викладачі ЕГ мають рівень розвитку компетентності щодо оцінювання якості тестів з вищої математики вищий, ніж викладачі КГ, що навчались самостійно.

За результатами анкетування та тестування було визначено відносні значення за кожним із зазначених рівнів компонентів компетентності щодо оцінювання якості тестів з вищої математики. Значення кутів на кінець експерименту у контрольній та експериментальній групах представлено у Таблиця 4.4. Аналіз значень кутів у Таблиця 4.4 показав, що емпіричні

значення більше критичних, тобто входять до зони значимості. Таким чином, можна зробити висновок, що між показниками у контрольній та експериментальній групі на кінець педагогічного експерименту є значні розбіжності, тому гіпотеза H_1 підтверджується.

Після аналізу робимо висновок, що на початку педагогічного експерименту компетентність викладачів щодо оцінювання якості тестів з вищої математики в КГ та ЕГ була нижче базового рівня, а на кінець педагогічного експерименту у КГ рівень компетентності зріс до поглибленого, а у ЕГ – до професійного.

Таблиця 4.4

**Значення кутів для компетентності викладачів
щодо оцінювання якості тестів з вищої математики на кінець
експерименту у контрольній та експериментальній групах**

Когнітивний компонент	φ_1 (91,7)	2,56			
	φ_2 (73,2)	2,05			
	φ^*	2,94	$\varphi^* > \varphi_{кр} = 2,31; (p = 0,01)$		
Ціннісно-мотиваційний компонент	φ_1 (71,5)	2,02			
	φ_2 (44,6)	1,46			
	φ^*	3,22	$\varphi^* > \varphi_{кр} = 2,31; (p = 0,01)$		
Операційно-діяльнісний компонент	φ_1 (77,8)	2,16			
	φ_2 (44,8)	1,47			
	φ^*	4,05	$\varphi^* > \varphi_{кр} = 2,31; (p = 0,01)$		
Компетентність викладачів щодо оцінювання якості тестів з ВМ	φ_1 (80,3)	2,22			
	φ_2 (54,2)	1,65			
	φ^*	3,31	$\varphi^* > \varphi_{кр} = 2,31; (p = 0,01)$		

Аналіз змін компонентів компетентності щодо оцінювання якості тестів з вищої математики показав таке:

– рівень розвитку когнітивного компоненту на початок педагогічного експерименту у КГ і ЕГ був нижче базового рівня, а на кінець експерименту у КГ досяг поглибленого рівня, а у ЕГ – професійного рівня.

– рівень розвитку ціннісно-мотиваційного компоненту на початок проведення педагогічного експерименту у КГ і ЕГ знаходився нижче базового рівня, а на кінець експерименту у ЕГ збільшився до поглибленого, а у КГ підвищився до базового рівня;

– рівень розвитку операційно-діяльнісного компоненту в КГ і ЕГ на початок педагогічного експерименту знаходилися нижче базового рівня; а на кінець експерименту у КГ досяг базового рівня, а у ЕГ – професійного рівня.

Таким чином, можна зробити висновок, що використання запропонованої методики позитивно впливає на динаміку рівнів розвитку компетентності щодо оцінювання якості тестів з вищої математики.

Проведений педагогічний експеримент підтвердив гіпотезу дослідження.

4.2. Експертне оцінювання якості контролю знань студентів з вищої математики.

Відповідно до *гіпотези* дослідження педагогічно виважене впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики сприятиме розвитку у викладачів ЗВО відповідної компетентності за умови реалізації розробленої методики покращенню контролю знань студентів. Перевірка даного твердження відбувалось із застосуванням методу експертного оцінювання.

Метод експертного оцінювання — це науковий метод, який на основі сукупності індивідуальних висновків експертів дає змогу отримати об'єктивну оцінку. Основними характеристиками методу експертних оцінок є: науково обґрунтована організація проведення всіх етапів експертизи; застосування кількісних методів при організації експертизи та при обробці суджень експертів [105].

Експертне оцінювання якості контролю знань студентів при впровадженні комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики було проведено за методом Делфі. Дельфійська техніка полягає у багаторазовому поштовому анкетуванні однієї і тієї ж групи експертів [56]. Після першого опитування експертів і опрацювання його результатів, підсумки повідомляються учасникам експертної групи. Вони повинні або підтвердити свою точку зору, висловлену на попередньому етапі,

і якщо вона значно відрізняється від думки більшості, розгорнуто її мотивувати, або змінити свою оцінку відповідно до думок більшості учасників. Потім знову обробляються висновки експертів, результати знову розсилаються експертам поки не будуть отримані узгоджені оцінки експертів.

Експертне оцінювання включало такі етапи:

1. підготовка документа експертного оцінювання;
2. формування групи експертів;
3. проведення експертизи;
4. опрацювання результатів експертного оцінювання.

Підготовка документа експертного оцінювання. На першому етапі було складено документ експертного оцінювання, в якому визначалися мета та завдання експерименту та термін виконання робіт. Також було розроблено анкети для визначення компетентності експертів та для безпосереднього проведення експертизи.

Формування групи експертів. На другому етапі для формування групи експертів було використано методи, описані в роботі [126]. Спочатку було складено список з двадцяти можливих експертів. До цього списку увійшли найбільш досвідчені співробітники кафедри математичного аналізу та теорії ймовірностей КПІ ім. Ігоря Сікорського, які викладають вищу математику, мають стаж роботи не менше п'яти років та вчене звання, мають досвід використання тестів для оцінювання знань студентів.

Далі було сформовано групу експертів із застосуванням методу колективного оцінювання. Даний метод є доцільним для формування групи експертів, оскільки викладачі тривалий час працюють в одному колективі та знають один одного як спеціалісти.

Метод колективного оцінювання полягав у такому. Кожному з двадцяти можливих експертів було роздано список з їхніми прізвищами, в якому вони мали зазначити 10 на їхню думку найбільш компетентних колег.

На підставі відповідей було побудовано таблицю рейтингів викладачів, фрагмент якої представлено на Рис. 4.7.

Хто обирав	Кого обирали									
	1	2	3	4	5	6	7	8	...	20
1	-	0	1	0	1	1	0	1	...	1
2	1	-	1	1	0	1	0	1		0
3	0	1	-	1	0	0	1	0		1
4	1	0	0	-	1	1	1	0		0
5	0	1	0	0	-	0	0	1		0
6	0	1	0	1	1	-	0	1		0
7	1	0	1	1	0	1	-	0		1
8	0	0	0	0	1	1	1	-		0
...	...									
20	1	1	0	1	1	0	0	1	...	-
Рейтинг експерта	12	11	9	18	14	15	7	13	...	8

Рис. 4.7. Таблиця рейтингів викладачів за методом колективного оцінювання

Номерами від 1 до 10 в першому стовпці позначені ті, хто обирав прізвища експертів, а в першому рядку – ті, кого обирали. В таблиці число 1 означає вибір експерта, 0 – відсутність вибору, тире вказує на те, що ніхто сам себе не може обрати. В останньому рядку таблиці наводиться сума «голосів», яку отримав кожен експерт. Ці числа і визначають рейтинг відповідних експертів. За показниками рейтингу було відібрано групу з 10 компетентних експертів.

Для визначення компетентності відібраних експертів та підтвердження репрезентативності експертної групи було використано метод самооцінки, який передбачав заповнення експертами анкети (Додаток Ж).

Рейтинг кожного експерта підраховувався у балах на основі методики Б. С. Гершунського [32]:

1) наявність наукового ступеня, вченого звання, рівня кваліфікації: доктор наук, професор – 1; кандидат наук, доцент – 0,8; кандидат наук, старший викладач – 0,6; кандидат наук, асистент – 0,4; асистент без вченого ступеня – 0,2;

2) стаж роботи: до 3 років – 0,2; від 3 до 5 років – 0,4; від 5 до 10 років – 0,6; від 10 років до 15 років – 0,8; понад 15 років – 1;

3) обґрунтування думки за проблемою (відповідно до проведених досліджень та педагогічного досвіду): 5 відповідей «так» на запитання 5-9 анкети – 1; 4 відповіді «так» на запитання 5-9 анкети – 0,8; 3 відповіді «так» на запитання 5-9 анкети – 0,6; 2 відповіді «так» на запитання 5-9 анкети – 0,4; 1 відповідь «так» на запитання 5-9 анкети – 0,2.

Зведені показники експертів, отримані на основі даних про стаж їхньої наукової діяльності, наукове звання, вчений ступінь та обґрунтування думки за досліджуваною проблемою, наведено у Таблиця 4.5.

Достовірність експертного оцінювання, головним чином, залежить від обізнаності експертів щодо проблеми дослідження, їх наукового та педагогічного стажу, уміння обґрунтовано та об'єктивно висловити власну думку. Загальну компетентність i -ого експерта за усіма показниками було обчислено за результатами анкетування за формулою:

$$K_i = \frac{\sum_{j=1}^3 K_{ij}}{K_{\max}}$$

де K_{ij} – рейтинг i -ого експерта за j -им показником;

$K_{\max}$ – максимальний загальний рейтинг експерта. В даному випадку $K_{\max} = 3$.

Група експертів вважається репрезентативною, якщо середнє значення загальних рейтингів експертів знаходиться в проміжку від 0,67 до 1. Середнє значення компетентності даної групи експертів дорівнювало 0,84. При цьому значення компетентності всіх експертів не виходили за проміжок $[0,67;1]$. Отже, експерти, які брали участь в оцінюванні, мають право залишитися у сформованій групі, а результати експертного оцінювання можна вважати репрезентативними.

Показники фахової компетентності експертів

		Абсолют ні значення	Відносн і значенн я (%)	Рейтинг експерта за Б.Гершунськи м
Стаж наукової діяльності	До 3 років	0	0	0,2
	3-5 років	0	0	0,4
	5-10 років	2	20	0,6
	10-15 років	4	40	0,8
	Більше 15 років	4	40	1
	Всього експертів	10	100	
Наукове звання, рівень кваліфікації та вчений ступінь	Доктор наук, професор	1	10	1
	Кандидат наук, доцент	7	70	0,8
	Кандидат наук, старший викладач	2	20	0,6
	Кандидат наук, асистент	0	0	0,4
	Асистент без вченого ступеня	0	0	0,2
	Всього експертів	10	100	
Обґрунтуван ня думки за досліджуван ою проблемою	Проведені дослідження та педагогічний досвід високого рівня	6	60	1
	Проведені дослідження та педагогічний досвід середнього рівня	3	30	0,8
	Педагогічний досвід високого рівня	1	10	0,6
	Педагогічний досвід середнього рівня	0	0	0,4
	Інтуїтивні уявлення	0	0	0,2
	Всього експертів	10	100	

Проведення експертизи. Для опитування експертів було розроблено анкету, що містила 16 питань (Додаток 3). Запитання було укладено за структурно-ієрархічною схемою: від широких питань до більш вузьких.

При проведенні анкетування експертів було забезпечено однозначність розуміння окремих запитань. Експертам було запропоновано для аналізу обрати шкалу оцінювання. В результаті було прийнято рішення оцінювати якість контролю знань за такою шкалою: «повністю не погоджуюсь», «скоріш не погоджуюсь, ніж погоджуюсь», «скоріш погоджуюсь, ніж не погоджуюсь», «погоджуюсь», «повністю погоджуюсь». Така шкала є найзрозумілішою та дає змогу експерту обрати відповідь відповідно до вагомості та свого ставлення до поставленого запитання.

Опрацювання результатів експертного оцінювання. Для узагальнення висновків експертів було використано статистичні методи. Коефіцієнти вагомості оцінок експертів обчислювались як відношення отриманої сумарної кількості балів за кожною оцінкою до загально можливої кількості балів. Значення коефіцієнтів вагомості оцінок експертів у відсотках представлено в Таблиця 4.6.

Таблиця 4.6

Коефіцієнти вагомості оцінок експертів

	Середнє значення у групі експертів (%)
<i>Повністю погоджуюсь</i>	40
<i>Погоджуюсь</i>	36
<i>Скоріш погоджуюсь, ніж не погоджуюсь</i>	22
<i>Скоріш не погоджуюсь, ніж погоджуюсь</i>	2
<i>Повністю не погоджуюсь</i>	0

На Рис. 4.8 представлено діаграму коефіцієнтів вагомості оцінок експертів.

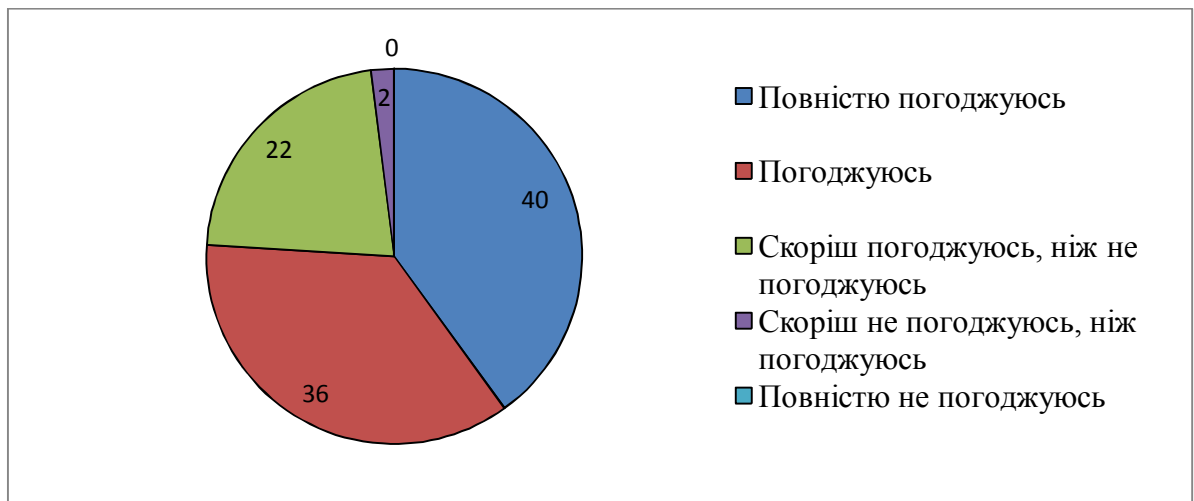


Рис. 4.8. Діаграма коефіцієнтів вагомості оцінок експертів у відсотках.

Експертна оцінка характеризує узагальнену думку й ступінь узгодженості індивідуальних оцінок експертів. Відповідно до рис. 4.8. покращення контролю знань в результаті впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики відмітили 76 % експертів (40 % повністю погоджуються, 36 % погоджуються) та 22 % скоріше погоджуються, ніж не погоджуються. Таким чином, можна зробити висновок, що впровадження комп'ютерно орієнтованої методики оцінювання якості тестів сприятиме покращенню контролю знань з вищої математики.

Висновки до розділу 4

1. Основну роботу експериментального дослідження було спрямовано на перевірку гіпотези та основних теоретичних положень дослідження. Розроблена програму експериментального дослідження передбачала констатувальний (2015-2016 рр.) та формувальний (2016-2017 рр.) етапи

2. В результаті констатувального етапу педагогічного дослідження з'ясовано, що незважаючи на те, що тестування для контролю знань студентів на практиці використовує незначна частка опитуваних, викладачі не заперечують можливості та доцільності його використання. Використання тестів є привабливим для викладачів (за умови поєднання традиційних та

тестових форм контролю знань студентів), оскільки дає можливість позбавитись від трудомісткої для них перевірки традиційних контрольних та самостійних робіт. Значна більшість викладачів готова використовувати тести з вищої математики, валідність і надійність яких було підтверджено. Викладачі розуміють необхідність оцінювання якості тестів з вищої математики із застосуванням ІКТ перед їхнім впровадженням в освітній процес, але при цьому усвідомлюють недостатній рівень компетентності щодо розроблення та оцінювання якості тестів. Пройти навчання, спрямоване на розвиток цієї компетентності, готові лише 28% викладачів.

3. Мета проведення формувального етапу педагогічного дослідження – розробити модель та технологію розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики та перевірити ефективність комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики. У формувальному етапі експериментального дослідження взяло участь 136 науково-педагогічних працівників КПІ ім. Ігоря Сікорського, з них 68 увійшло в експериментальну групу (ЕГ) та 68 – у контрольну групу (КГ). Навчання викладачів експериментальної групи було проведено відповідно до моделі через реалізацію навчальної програми «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики».

4. Результати початкового і підсумкового зрізів із використанням методів статистичного опрацювання (кутового перетворення φ^* Фішера) та порівняльного аналізу підтвердили позитивну динаміку розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики при впровадженні комп'ютерно орієнтованої методики оцінювання якості тестів, що дозволяє зробити висновок про позитивний вплив методики. Динаміка змін показників ЕГ і КГ засвідчила позитивні тенденції розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики за кількістю респондентів щодо рівнів компетентності.

5. Покращення контролю знань студентів з вищої математики внаслідок впровадження комп'ютерно орієнтованої методики оцінювання якості тестів перевірялось методом експертного оцінювання. Експертне оцінювання проводилося за методом Дельфі та охоплювало такі етапи: підготовка документа експертного оцінювання; формування групи експертів; проведення експертизи та опрацювання результатів експертного оцінювання.

6. Проведений педагогічний експеримент підтвердив гіпотезу, що педагогічно виважене впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики сприятиме розвитку у викладачів ЗВО відповідної компетентності за умови реалізації розробленої технології розвитку компетентності та покращенню контролю знань студентів.

Результати четвертого розділу дисертаційного дослідження висвітлено у науковій роботі автора [54].

ВИСНОВКИ

Відповідно до мети та поставлених завдань у процесі дослідження отримано такі результати: уточнено основні поняття з проблеми дослідження; здійснено аналіз сучасного стану проблеми оцінювання якості тестів з вищої математики; досліджено методи КТТ та ІРТ оцінювання якості тестів з вищої математики; розроблено комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики викладачами ЗВО; розроблено СААЯТЗ як засіб комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики; сформовано базу каліброваних завдань з вищої математики; оцінено та покращено якість пакету тестів з вищої математики для студентів технічних спеціальностей КПІ ім. Ігоря Сікорського; визначено критерії та рівні розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики; розроблено модель та технологію розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики.

На підставі отриманих результатів дослідження зроблено такі висновки:

1. На сьогоднішній день комп'ютерне тестування є ефективним засобом контролю знань студентів, спрямованим на підвищення якості освіти. Доведено, що для зростання ефективності навчального процесу тестування повинно не підміняти, а доповнювати перевірені традиційні форми контролю знань.

Виявлено, що, незважаючи на усвідомлення вітчизняними та зарубіжними вченими необхідності оцінювання якості тестів з вищої математики та розроблення відповідної комп'ютерно орієнтованої методики, зазначена проблема наразі залишається недостатньо вирішеною. Це підтверджується відсутністю єдиного визначення основних понять дослідження, зокрема, поняття якості тестів. Нами уточнено поняття тесту,

якості тесту, оцінювання якості тестів та комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики.

Наголошено, що для застосування тестів з вищої математики для контролю знань студентів у викладача повинна бути розвинена компетентність щодо оцінювання якості тестів. При цьому проблема підготовки викладачів вищої математики до застосування та оцінювання якості тестів також не знайшла достатнього відображення в наукових дослідженнях.

Аналіз основних програмних засобів опрацювання результатів тестування виявив необхідність розроблення системи автоматизованого аналізу якості тестових завдань з вищої математики як засобу комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики.

2. Дібрано методи КТТ та IRT для оцінювання якості тестів з вищої математики та їхньої програмної реалізації. Методи оцінювання якості тестів з вищої математики в рамках КТТ передбачають як аналіз окремих тестових завдань, так і тесту в цілому. Аналіз тесту в цілому полягає в: аналізі розподілу вибірки результатів; оцінюванні надійності тесту; оцінюванні критеріальної валідності тесту. Аналіз тестових завдань передбачає: аналіз кореляційної матриці завдань; оцінювання валідності окремих завдань; аналіз дистракторів завдань множинного вибору. Визначено, що для оцінювання якості тестів з вищої математики доцільно застосовувати наступні моделі IRT: Раша та Бірнбаума (для дихотомічних завдань), Андерсена для політомічних завдань та Тіссена-Стейнберга для завдань множинного вигляду. IRT-аналіз тестових завдань і тесту в цілому проводиться із використанням визначених моделей і передбачає: оцінювання значень латентних параметрів використовуваних IRT-моделей; аналіз ансамблю характеристичних кривих завдань тесту; аналіз характеристичних кривих підрівнів політомічних тестових завдань; аналіз характеристичних кривих дистракторів завдань множинного вибору; аналіз

інформаційних кривих тестових завдань та тесту в цілому. Встановлено, що всебічний досконалий аналіз якості тесту можливий лише при поєднанні взаємодоповнюючих методів КТТ та IRT. Дібрані методи підтвердили свою ефективність та були покладені в основу комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики.

3. Спроектовано систему автоматизованого аналізу якості тестових завдань з вищої математики як засіб розробленої методики. Функції системи реалізовано в СААЯТЗ за допомогою окремих модулів. Основними задачами СААЯТЗ є статистичний аналіз результатів тестування методами КТТ та IRT, формування та обслуговування бази каліброваних завдань. У базі каліброваних завдань зберігаються тести та окремі тестові завдання. Тести систематизовано за темами та призначенням (контрольна, іспит). Для кожного завдання існує відповідний паспорт, який містить тему тестування, форму завдання та основні характеристики тестового завдання, оцінені системою. База каліброваних завдань є невід'ємною складовою СААЯТЗ, необхідною для роботи викладача з тестами з вищої математики, а саме для формування та редагування тестів. Застосування СААЯТЗ для оцінювання якості тестів з вищої математики підтвердило свою ефективність.

4. Розвиток компетентності викладачів щодо оцінювання якості тестів з вищої математики доцільно здійснювати на основі технології, розробленої на підґрунті відповідної моделі з урахуванням визначених критеріїв (ціннісно-мотиваційного, когнітивного та операційно-діяльнісного) та рівнів (базового, поглибленого та професійного). Запропонована модель складається з наступних компонентів: мотиваційно-цільового (передбачає цілі та завдання освітнього процесу), змістового (визначає зміст навчальної діяльності за рівнями розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики), операційно-технологічного (конкретизує організаційні форми та методи

розвитку цієї компетентності), діагностичного та результативного (визначають рівні компетентності викладачів щодо оцінювання якості тестів з вищої математики, експертизу й аналіз результативності даного процесу).

5. Розроблено комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики. Дана методика надає можливість проводити швидкий, зручний та наглядний аналіз якості завдань та тестів. Загальна схема методики включає в себе три етапи: формування таблиці результатів; аналіз окремих тестових завдань і тесту в цілому; приймання остаточних рішень. Застосування методики надає можливість не лише оцінити, а й у подальшому здійснити покращення якості як окремих тестових завдань, так і тесту в цілому. Розроблено систему автоматизованого аналізу якості тестових завдань з вищої математики як засіб розробленої методики. На основі узагальнення результатів, теоретичного аналізу й експериментальної роботи розроблено дистанційний курс «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики».

Впровадження комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики доцільно здійснювати поетапно: навчання викладачів на основі запропонованої технології розвитку компетентності щодо оцінювання якості тестів з вищої математики та безпосереднє застосування викладачами комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики в свої професійній діяльності для покращення контролю знань. Експериментальна перевірка комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики та отримані результати підтвердили викладену в дисертації гіпотезу.

Разом із тим проведене дослідження не вичерпує всієї повноти проблеми і може бути продовжено за такими напрямками наукового

пошуку: розроблення та впровадження методики комп'ютерно орієнтованого оцінювання якості тестів інших дисциплін; дослідження інноваційних методів IRT оцінювання якості тестів та вдосконалення засобів розробленої методики.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Аванесов В. С. Композиция тестовых заданий. Учебная книга для преподавателей вузов, учителей школ, аспирантов и студентов педвузов. 2 изд., испр. и доп. / В. С. Аванесов – М.: Адепт 1998. – 217с.
2. Аванесов В. С. Научные проблемы тестового контроля знаний / В. С. Аванесов. – М. : Исслед. центр. пробл. кач. подгот. спец., 1994. – 135 с.
3. Аванесов В.С. Основы научной организации педагогического контроля в высшей школе. / В. С. Аванесов. – М., 1989. – 167 с.
4. Аванесов В. С. Основы педагогической теории измерений / В. С. Аванесов // Педагогические измерения. – 2004. – №1. – С. 3-19.
5. Аванесов В. С. Применение тестовых форм в Raschmeasurement / В. С. Аванесов // Педагогические Измерения. – 2005. – № 4. – С. 3-20.
6. Аванесов В. С. Проблема объективности педагогических измерений / Вадим Сергеевич Аванесов // Педагогические измерения. – 2008. – № 3. – С. 3–39.
7. Аванесов В. С. Форма тестовых заданий / В. С. Аванесов. – М. : Центр тестирования, 2005. – 156 с.
8. Аванесов В.С. Item Response Theory: Основные понятия и положения / В. С. Аванесов // Педагогические измерения. – 2007. – №2. – С.3-28
9. Адольф В.А. Профессиональная компетентность современного учителя: монография / Красноярский Гос. университет / В.А. Адольф. – Красноярск: КрГУ, 1998. – 286 с.
10. Анастаси А. Психологическое тестирование / А. Анастаси, С. Урбина – Спб.: Питер, 2006. – 688 с.
11. Анисимова Т.С. Измерение латентных переменных в образовании / Т.С. Анисимова – М.: Исследовательский центр проблем качества подготовки специалистов, 2004. – 148 с.

12. Архангельский С. И. Учебный процесс в высшей школе, его закономерные основы и методы / С. И. Архангельский. – М.: Высшая школа, 1980. – 368 с.
13. Балыхина Т. М. Словарь терминов и понятий тестологии / Т. М. Балыхина. – М. : Изд-во РУДН, 2000. – 164 с.
14. Беспалько В. П. Инструменты диагностики качества знаний учащихся / В.П. Беспалько // Школьные технологии. – 2006. – №2. – С. 138–150.
15. Беспалько В. П. О критериях качества подготовки специалистов / В. П. Беспалько // Вестник высшей школы. – 1988. – №1. – С.3–9.
16. Беспалько В. П. Слагаемые педагогической технологии. / В. П. Беспалько. – Москва, 1989 – 192 с.
17. Бех І. Наукові засади проведення експерименту / І. Бех, О. Кононко // Рідна школа. – 2001. – №10. – С. 36–40.
18. Биков В.Ю. До питання інформатизації вищих педагогічних навчальних закладів / В.Ю. Биков, І.Ф. Прокопенко, С.А. Раков // Комп'ютер у школі та сім'ї. – 2002. – №4. – С. 8-13.
19. Биков В. Ю. Інформаційне забезпечення навчально-виховного процесу: інноваційні засоби та технології: Колективна монографія / В. Ю. Биков, Гриценчук О. О., Жук Ю. І. та ін.// – К.: Атіка, 2005. – 252 с.
20. Биков В.Ю. Моніторинг рівня навчальних досягнень з використанням Інтернет-технологій: монографія / В. Ю. Биков, Ю. М. Богачков, Ю. О. Жук; за ред. В. Ю. Бикова, чл.-кор. АПН України, д. тех. наук, проф.; Ю. О. Жука, канд. пед. наук, доц. – К.: Педагогічна думка, 2008. – 128с.
21. Биков В.Ю. Наукове забезпечення дистанційної професійної освіти: проблеми та напрями досліджень / В. Ю. Биков // Професійна освіта: педагогіка і психологія. За ред.: І. Зазюна, Н. Нічкало, Т. Левовицького, І. Вільш. Україно-польський журнал. Видання II.. Видавництво: ЗАТ «ВІПОЛ», Київ-Ченстохова. – 2000 р. С. 93–116.

22. Білощицький А. О. Методи та моделі комплексного інформаційно-освітнього середовища в умовах розвитку вищого навчального закладу : автореф. дис. на здобуття наук. ступеня канд. техн. наук : спец. 05.13.06 / А. О. Білощицький; Київ. нац. ун-т буд-ва і архіт. – К., 2007. – 20 с.
23. Болотов В. А. Компетентностная модель: от идеи к образовательной программе / В. А. Болотов, В. В. Сериков // Педагогика. – 2003. – № 10. – С. 8-11.
24. Болюбаш Н. М. Створення тестів для інформаційно-освітньої системи на базі електронної платформи MOODLE: навч. посібник / Н. М. Болюбаш. – Миколаїв : МДГУ ім. Петра Могили, 2008. – 68 с.
25. Булах І. Є. Створюємо якісний тест : навчальний посібник / І. Є. Булах, М. Р. Мруга. – К. : Майстер-клас, 2006. – 160 с.
26. Булах І. Є. Теорія і методика комп'ютерного тестування успішності навчання (на матеріалах медичних навчальних закладів): дис. ... доктора пед. наук. : 13.00.01 / Булах Ірина Євгенівна. – К., 1995. – 430 с.
27. Бурцева Ю.О. Структура компетентності як педагогічного явища в контексті сучасного навчально-виховного процесу // Вища освіта України (Педагогіка вищої школи: методологія, теорія, технології). — 2011. — № 3. — Т. 1. — С. 576—582.
28. Васильева Е. Е. Суперпамять для всех. / Е. Е. Васильева, В. Ю. Васильев. — М.: Аст, 2006. — 71 с.
29. Выготский Л. С. Психология развития человека / Л. С. Выготский. — М. : Изд-во “Смысл”; Изд-во “Эксмо”, 2005. — 1136 с.
30. Виноградов О. Г. Тести як соціальна інновація в Україні / О. Г. Виноградов // Управління освітою. — 2003. — №1. — С.3.
31. Воронов В. И. Основы научных исследований. Учебное пособие / В. И. Воронов, В. П. Сидоров, редактор: М. А. Касаткина. – Владивосток: ВГУСИ, 2014. [Электронный ресурс] – Режим доступа: https://abc.vvsu.ru/books/osnovy_nauchn_issled/page0007.asp

32. Гершунский Б. С. Прогнозирование содержания обучения в техникумах: Учебно- методическое пособие / Гершунский Б. С.– М.: Высшая школа, 1980 – 144 с.
33. Глас Дж., Стэнли Дж. Статистические методы в педагогике и психологии. / Дж. Глас, Дж. Стэнли. – М.: Прогресс, 1976. – 495 с.
34. Гончаренко С. У. Український педагогічний словник. / С. У. Гончаренко. – К. : Либідь, 1997. – 376 с.
35. Грабарь М.И. Применение математической статистики в педагогических исследованиях. Непараметрические методы. / М. И. Грабарь, К. А. Краснянская. – М.: Педагогика, 1997. – 135 с.
36. Гронлунд К., Оцінювання студентської успішності: Практ. посіб. / К. Гронлунд, Е. Норман. – К. : Навчально-методичний центр "Консорціум із удосконалення менеджмент-освіти в Україні", 2005. – 312 с.
37. Гулюкина Н. А. Педагогический тест: этапы и особенности конструирования и использования. Пособие для преподавателей / Н. А. Гулюкина, С. В. Клишина. – Новосибирск : Изд-во НГТУ, 2001. – 147 с.
38. Драган Є. В. Аналіз завдань в тестовій формі для моніторингу якості знань з фізики / Є. В. Драган // Вища освіта України. – 2007. – Т. 7 – С. 224–230.
39. Драган Є. В. Комп'ютерно орієнтована технологія оцінювання навчальних досягнень студентів фізичних спеціальностей на основі ймовірнісних теорій тестування [Текст]: дис. ... канд. пед. наук. : 13.00.10 / Є. В. Драган Євген Вікторович. – К. 2012. – 218 с.
40. Дудко А. Ф. Автоматизована система аналізу результатів комп'ютерног тестування з вищої математики / О. О. Диховичний, А. Ф. Дудко // Наукові праці ДонНТУ. Серія: «Педагогіка, психологія і соціологія». – №2 (14). – Донецьк , 2013. – С .103-110.
41. Дудко А. Ф. Аналіз поняттєво-категоріального апарату дослідження якості тестів з вищої математики / О.О. Диховичний,

А.Ф. Дудко // Journal «ScienceRise: Pedagogical Education». – № 5(13). – 2017. – С. 17-21.

42. Дудко А. Ф. Аналіз якості комп'ютерних тестів методами класичної теорії тестів / І. В. Алексєєва, В. О. Гайдей, О. О. Диховичний, А. Ф. Дудко, Н. Р. Коновалова, Л. Б. Федорова // Шістнадцята міжнародна конференція імені академіка М. Кравчука, 14-15 травня, 2015 р. Київ: Матеріали конф. Т. 3. Теорія ймовірностей та математична статистика. Історія та методика математики. – К.: НТУУ «КПІ», 2015. – С. 94-97.

43. Дудко А. Ф. Вибір підходу до розробки комп'ютерних тестів з вищої математики. / І. В. Алексєєва, В. О. Гайдей, О. О. Диховичний, А. Ф. Дудко, Н. Р. Коновалова, Л. Б. Федорова // Математика в сучасному технічному університеті: матеріали III Міжнар. наук.-практ. конф. (Київ, 25–26 груд. 2014 р.). — К.: НТУУ «КПІ», 2015. — С. 317-318.

44. Дудко А. Ф. Застосування інформаційної функції для аналізу та підвищення ефективності тестів з вищої математики [Електронний ресурс] / О. О. Диховичний, А. Ф. Дудко // Інформаційні технології і засоби навчання. – Том 41 №3 (2014). – 2014. – Режим доступу: <http://journal.iitta.gov.ua/index.php/itlt/article/view/1040/790#.VFK7rzSsXm4>

45. Дудко А. Ф. Застосування моделі Бірнбаума до визначення ставлення викладачів вищої математики до тестового контролю знань / О. О. Диховичний, А. Ф. Дудко, Н. В. Круглова, З. П. Ординська // Сьома всеукраїнська наукова конференція студентів, аспірантів та молодих вчених з математики, 18-20 квітня, 2018 р. – С. 73.

46. Дудко А. Ф. Застосування мови R у статистичному аналізі якості тестів з вищої математики / О. О. Диховичний, А. Ф. Дудко, Н. В. Прохоренко // Матеріали Вісімнадцятої міжнародної наукової конференції імені академіка Михайла Кравчука, 7—10 жовтня 2017 року, Київ: Т. 2. — Київ: НТУУ «КПІ», 2017. – С. 222-224.

47. Дудко А. Ф. Застосування IRT-моделей у дослідженні якості педагогічних тестів / О. О. Диховичний, А. Ф. Дудко // Всеукраїнська

наукова конференція «Стохастичні та детерміновані математичні моделі у наукових дослідженнях», 27-29 червня 2014 р., м. Київ.

48. Дудко А. Ф. Комплексна методика аналізу якості тестів з вищої математики / О.О.Диховичний, А.Ф.Дудко // Науковий часопис НПУ імені М.П. Драгоманова. Серія №2. Комп'ютерно-орієнтовані системи навчання: Зб. наук. праць / Редрада. – К.: НПУ імені М.П. Драгоманова, 2015. – № 15 (22). – С.140-144.

49. Дудко А. Ф. Комп'ютерно-орієнтована методика статистичного аналізу тестів з вищої математики / О. О. Диховичний, А. Ф. Дудко, Тарасова Є. О. // П'ята міжнародна наукова конференція молодих вчених з математики та фізики, 25-26 квітня, 2016 р. – С. 71

50. Дудко А. Ф. Комп'ютерно-орієнтована методика аналізу якості педагогічних тестів як засіб ІК-підтримки викладачів вищої математики НТУУ «КПІ» / І.В. Алексєєва, В.О. Гайдей, О.О. Диховичний, А.Ф. Дудко, Н.Р. Коновалова, Л.Б. Федорова // Математика в сучасному технічному університеті: матеріали IV Міжнар. наук.-практ. конф. (Київ, 24–25 груд. 2015 р.). — К.: НТУУ «КПІ», 2016. — С. 277-278.

51. Дудко А. Ф. Концепція розробки комп'ютерних тестів з вищої математики / А.Ф.Дудко // Науковий часопис Національного педагогічного університету імені М. П. Драгоманова. Серія 5. Педагогічні науки: реалії та перспективи. – Випуск 50: збірник наукових праць. – Київ: Вид-во НПУ імені М. П. Драгоманова, 2014. – С. 44-48.

52. Дудко А. Ф. Методика аналізу та підвищення ефективності тестів з вищої математики / О.О. Диховичний, А.Ф. Дудко // П'ятнадцята міжнародна конференція імені академіка М. Кравчука, 15-17 травня, 2014 р. Київ: Матеріали конф. Т.4. Історія та методика математики. –К.: НТУУ «КПІ», 2014. – С. 96-97.

53. Дудко А. Ф. Методика аналізу якості тестів та її застосування / О. О. Диховичний, А. Ф. Дудко // Четверта міжнародна наукова конференція молодих вчених з математики та фізики, 23-25 квітня, 2015 р. – С. 78.

54. Дудко А. Ф. Модель розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики [Електронний ресурс] / О. О. Диховичний, А. Ф. Дудко // Інформаційні технології і засоби навчання. [Прийнято до друку]

55. Дудко А. Ф. Моделювання даних результатів тестування з вищої математики за допомогою мови R / О. О. Диховичний, А. Ф. Дудко, Н. В. Прохоренко // Матеріали Вісімнадцятої міжнародної наукової конференції імені академіка Михайла Кравчука, 7—10 жовтня 2017 року, Київ: Т. 2. — Київ: НТУУ «КПІ», 2017. — С. 225-226.

56. Дудко А. Ф. Оцінювання змістовної валідності тестів з вищої математики з використанням Google Docs додатку [Електронний ресурс] / О. О. Диховичний, А. Ф. Дудко // Інформаційні технології і засоби навчання. — Том 52 №2 (2016). — 2016. — Режим доступу: <http://journal.iitta.gov.ua/index.php/itlt/article/view/1383/1032>

57. Дудко А. Ф. Оцінювання якості змісту тестів з вищої математики з використанням Google Docs додатку / І. В.Алексєєва, В. О. Гайдей, О. О. Диховичний, А. Ф. Дудко, Н. Р. Коновалова, Л. Б. Федорова // Сімнадцята міжнародна конференція імені академіка М. Кравчука, 19-20 травня, 2016 р. Київ: Матеріали конф. Т. 3. Теорія ймовірностей та математична статистика. Історія та методика математики. — К.: НТУУ «КПІ», 2015. — С. 176-179.

58. Дудко А. Ф. Практичний аналіз якості тестів з вищої математики / О.О.Диховичний, А.Ф.Дудко // Науковий вісник Ужгородського національного університету: Серія «Педагогіка. Соціальна робота». — № 32. — 2014. — С. 73-78.

59. Дудко А. Ф. Про досвід застосування тестових контрольних робіт з вищої математики. / І. В.Алексєєва, В. О. Гайдей, О. О. Диховичний, Н. Р. Коновалова, Л. Б. Федорова, А. Ф. Дудко // Міжнародна наукова конференція «Математика в сучасному технічному університеті», 19-20 квітня 2013 р., м. Київ — С. 442-445.

60. Дудко А. Ф. Про електронне навчання вищої математики в НТУУ «КПІ». / І. В.Алексєєва, В. О. Гайдей, О. О. Диховичний, Н. Р. Коновалова, Л. Б. Федорова, А. Ф. Дудко // Міжнародна наукова конференція «Математика в сучасному технічному університеті», 19-20 квітня 2013 р., м. Київ – С. 438-441.

61. Дудко А. Ф. Про електронне тестування з вищої математики в НТУУ «КПІ» / І. В.Алексєєва, В. О. Гайдей, О. О. Диховичний, Н. Р. Коновалова, Л. Б. Федорова, А. Ф. Дудко // Вища освіта: проблеми і шляхи забезпечення якості. X Всеукраїнська науково-методична конференція, 28-29 листопада, 2013 р. – м. Київ, НТУУ «КПІ». – С. 7.

62. Дудко А. Ф. Про необхідність розвитку компетентності щодо оцінювання якості тесту викладача вищої математики / І. В.Алексєєва, В. О. Гайдей, О. О. Диховичний, А. Ф. Дудко, Н. Р. Коновалова, Л. Б. Федорова // Матеріали Вісімнадцятої міжнародної наукової конференції імені академіка Михайла Кравчука, 7—10 жовтня 2017 року, Київ: Т. 2. — Київ: НТУУ «КПІ», 2017. – С. 166-167.

63. Дудко А.Ф. Про особливості застосування політоміних математичних моделей тестових завдань / О.О. Диховичний, М.О. Шепель, А.Ф. Дудко // Чотирнадцята міжнародна конференція імені академіка М. Кравчука, 19-21 квітня, 2012 р. Київ: Матеріали конф. Т.4. Історія та методика викладання математики. –К.: НТУУ «КПІ», 2012. – С. 91.

64. Дудко А. Ф. Про оцінку ефективності тестів з вищої математики. / І. В.Алексєєва, В. О. Гайдей, О. О. Диховичний, Н. Р. Коновалова, Л. Б. Федорова, А. Ф. Дудко // Друга міжнародна наукова конференція «Математика в сучасному технічному університеті», 20-21 грудня, 2013 р., м. Київ – С. 271-274.

65. Дудко А. Ф. Система автоматизованого аналізу якості тестових завдань / О. О. Диховичний, А. Ф. Дудко, Г. В. Філіппова // Третя міжнародна наукова конференція молодих вчених з математики та фізики, 25-27 квітня, 2013 р. – С. 170-172.

66. Дудко А. Ф. Статистичний аналіз якості тестових завдань для моделі з множинним вибором / О. О. Диховичний, А. Ф. Дудко, І. С. Тарабара // Третя міжнародна наукова конференція молодих вчених з математики та фізики, 25-27 квітня, 2013 р. – С. 169-170.

67. Дудко А. Ф. Статистичний аналіз тестових завдань із застосуванням сучасних математичних моделей IRT / І. В. Алексєєва, В. О. Гайдей, О. О. Диховичний, А. Ф. Дудко, Н. Р. Коновалова, Л. Б. Федорова // Вісник Національного університету «Львівська політехніка» (серія «Інформатизація вищого навчального закладу»). – №731. – 2012. – С. 10-13.

68. Дудко А. Ф. Схема роботи викладача з тестами для контролю знань студентів з вищої математики / І. В. Алексєєва, В. О. Гайдей, О. О. Диховичний, А. Ф. Дудко, Н. Р. Коновалова, Л. Б. Федорова // Математика в сучасному технічному університеті : матеріали V Міжнар. наук.-практ. конф. (Київ, 29—30 грудня 2016 р.). — Київ : НТУУ «КПІ», 2017. — С. 229-232.

69. Евсеев, А.И. Разработка компьютерных средств обучения / А.И. Евсеев, Е.А. Ахромюшкин. – Режим доступа: http://cnit.mpei.ac.ru/textbook/00_00_00_02.htm

70. Ефремова Н.Ф. Тестовый контроль в образовании: Учеб. пособие / Н. Ф. Ефремова. – М.: Университетская книга, Логос, 2005. – 368 с.

71. Єфіменко, В. І. Створення критеріально-орієнтованих тестів досягнень особистості та інтерпретація результатів тестування на основі положень Item Response Theory. / В. І. Єфіменко, О. П. Касярум, В. М. Власенко //Вісник Черкаського університету. Випуск 26. Серія: педагогічні науки.: Збірник. – Черкаси: ЧДПУ, 2001. – С. 40-46.

72. Жалдак М. И. Система подготовки учителя к использованию информационной технологии в учебном процессе: дис. ... в форме научн. доклада ... доктора педагогических наук: 13.00.02 /М. И. Жалдак; АПН СССР; НИИ содержания и методов обучения. – М., 1989. – 48 с.

73. Жук Ю. О. Комп'ютерно орієнтовані засоби навчальної діяльності: проблеми створення та впровадження / Ю. О. Жук // Науковий вісник Ізмаїльського державного гуманітарного університету. – 2004. – Вип. 16. – С. 11–15.
74. Зазыкин В. Г.. Акмеологические проблемы профессионализма/ В. Г. Зазыкин, А. П. Чернышев // М. НИИ ВО, 1993. – 48 с.
75. Закон України “Про Національну програму інформатизації [Електронний ресурс] – Режим доступу: <http://zakon3.rada.gov.ua/laws/show/74/98-%D0%B2%D1%80>
76. Закон України «Про вищу освіту». Відомості Верховної Ради України, 2014, №37-38, – с. 2004.
77. Закон України «Про інформацію». Відомості Верховної Ради України, 1992, № 48, с. 650.
78. Закон України «Про Національну програму інформатизації». Відомості Верховної Ради України, 1998, № 27-28, –с.181.
79. Закон України «Про пріоритетні напрями інноваційної діяльності в Україні» [Електронний ресурс] – Режим доступу: <http://zakon3.rada.gov.ua/laws/show/3715-17>
80. Звонников В. И. Современные средства оценивания результатов обучения: учеб. пособие для студ. высш. учеб. заведений / В. И. Звонников, М.Б. Челышкова. — М. : Издательский центр «Академия», 2007. - 224 с.
81. Ильин Е. П. Мотивация и мотивы / Е. П. Ильин. – СПб: Питер 2002. – 512 с.
82. Ильина Т. А. Системно-структурный подход к организации обучения. Вып. 1 / Т. А. Ильина. – М.: Знание, 1972. – 72 с.
83. Ингенкамп К. Педагогическая диагностика / К. Ингенкамп. – М.: Педагогика, 1991. – 240 с.
84. Итоговый аналитический отчет о результатах ЕГЭ 2012 г. – Федеральный институт педагогических измерений. – М.: Москва, 2012. –

31 с. [Електронний ресурс] – Режим доступу:
http://www.fipi.ru/sites/default/files/document/1408709880/1_0.pdf

85. Іванова С. М. Використання системи EPrints як засобу інформаційно-комунікаційної підтримки наукової діяльності в галузі педагогічних наук [Текст]: дис. ... канд. пед. наук. : 13.00.10 / С. М. Іванова. – К. 2014. – 317 с.

86. Казаринов А.С. Технология педагогического эксперимента / А. С. Казаринов. – Гла-зов: Глазов, 1999. – 192 с.

87. Калашникова С.А. Навчання дорослих на основі компетентнісно-орієнтованого підходу: навч.-метод. матеріали / Світлана Андріївна Калашнікова. – К.: Проект Світового банку «Рівний доступ до якісної освіти», 2008. – 57 с.

88. Канюк С.С. Психологія мотивації: Навчальний посібник / С. С. Канюк – К.: Либідь, 2002. – С. 238-248.

89. Караваева И.А. Валидность педагогических тестов: Монография. / И. А. Караваева. – Ижевск: Издатель-ский дом «Удмуртский университет», 2004. – 150 с.

90. Ким В. С. Выбор оптимального времени тестирования / В. С. Ким // Педагогические измерения. – 2004. – № 1. – С. 43–50.

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

92. Клайн П. Введение в психометрическое проектирование. Справочное руководство по конструированию тестов. / П. Клайн –Киев: ПАН Лтд, 1994. -184 с.

93. Кларин М. В. Инновационные модели обучения в зарубежных педагогических поисках / М. В. Кларин. – М. Арена, 1994. – 223 с.

94. Колгатін О. Г. Автоматизована педагогічна діагностика і точність вимірювання / Олександр Колгатін // Вісник «Тестування і моніторинг в освіті». – 2006. – № 10-11. – С. 29-33.

95. Колгатін О. Г. Автоматизовані системи тестування // Вісник. «Тестування і моніторинг в освіті» – 2007. – № 10/11. – С. 38-41.
96. Колгатін О. Г. Теоретико-методичні засади проектування комп'ютерно орієнтованої системи педагогічної діагностики майбутніх учителів природничо-математичних спеціальностей : дис. ... доктора педагогічних наук : спец. 13.00.10 / Колгатін Олександр Геннадійович. – Харків, 2011. – 486 с.
97. Колесникова И. А. Коммуникативная деятельность педагога: учеб. пособие для студ. высш. пед. заведений / И. А. Колесникова; под ред. В. А. Сластенина. – М.: Издательский центр «Академия», 2007. – 336 с.
98. Концепція розвитку професійно-технічної освіти в Україні на 2010-2020 роки (проект) // Інформаційний збірник Міністерства освіти і науки України. – Київ: Педагогічна преса, 2010, № 17/18. – С.13-19.
99. Крокер Л. Введение в классическую и современную теорию тестов: учебник / Л. Крокер, Дж. Алгина; пер. с англ. Н. Н. Найденовой, В. Н. Симкина, М. Б. Чельшковой; под общ. ред. В. И. Звонникова, М. Б. Чельшковой. – М.: Логос, 2010. – 668 с.
100. Крутецкий В.А. Психология.: Учебник для пед.училищ / В. А. Крутецкий. – М.: Просвещение, 1980. – 352с.
101. Кудряшова Т.Г. Деятельностные способности как основа для формирования компетенций / Т.Г. Кудряшова // Компетентность. – № 6. – 2010. – с.5.
102. Кузнецов А. А. Оценка достижения требований образовательных стандартов / А. А. Кузнецов. – М. : Нац. центр стандартов и мониторинга образования, 1998. – 26 с.
103. Кузьмінський А. І. Педагогіка: підручник / А. І. Кузьмінський, В. Л. Омеляненко / [3-тє вид., перероб. і доп.]. – К.: Знання-Прес, 2008. – 447 с.
104. Кухар Л.О. Конструювання тестів. Курс лекцій: навч. посіб. / Л.О. Кухар, В.П.Сергієнко. – Луцьк, 2010. – 182 с.

105. Кыверялг А. А. Методы исследований в профессиональной педагогике / А. А. Кыверялг – Таллин: Валгус, 1980. – 334 с.
106. Леонтьев А. Н. Деятельность. Сознание. Личность / Алексей Николаевич Леонтьев. – М.: Политиздат, 1977. – 307 с.
107. Лісова Т. В. Моделі та методи сучасної теорії тестів / Т. В. Лісова. – Ніжин: Видавець ПП Лисенко М. М., 2012. – 112 с.
108. Лукіна Т. О. Педагогічна діагностика: завдання, методи, інструменти. Модуль 10. Навчально-методичні матеріали. – К.: «Проект «Рівний доступ до якісної освіти», 2007. – 60 с.
109. Майоров А. Н. Теория и практика создания тестов для системы образования. (Как выбирать, содавать и использовать тесты для целей образования) / А. Н. Майоров. – М. : Интеллект-центр, 2002. – 296 с.
110. Максимова О. А. Модель и технология комплексной экспертизы качества тестовых материалов для системы профессионального образования : дис. ... канд. пед. наук : 13.00.08. / О. А. Макстмова – М., 2010. – 191 с.
111. Маркова А.К. Психологический анализ профессиональной компетентности учителя // Советская педагогика. – 1990. - № 8. – С. 82-88.
112. Маркова А. К. Психология профессионализма / А. К. Маркова. – М.: Международный гуманитарный фонд “Знание”, 1996. – 312 с.
113. Маслак А.А. Измерение латентных переменных в образовании и других социально-экономических системах: теорія и практика. / А.А. Маслак. – Славянск-на-Кубани: Изд. центр СГПИ, 2007. – 424 с.
114. Масленников, А.С. О технологии анализа результатов аттестационных измерений «Камертон» / А.С. Масленников, Д.И. Петров, М.В. Петропавловский //Развитие системы тестирования в России: тезисы докладов всероссийской конференции. - М., 1999. Ч. 3. - С. 52-54.
115. Маслоу А. Г. Мотивация и личность / Маслоу А. Г. / Перевод с англ. Татлыбаевой А. М. – СПб.: Евразия, 1999. – 478 с.

116. Митина Л.М. Личностное и профессиональное развитие человека в новых социально-экономических условиях // Вопросы психологии. – 1997. – № 4. – С. 28-38.

117. Михайлов В. В. База даних тестових завдань як стандартний засіб діагностики рівня якості віщої освіти / В. В. Михайлов, В. М. Черінько // Вища школа : науково-практичне видання. – 2009. – № 8. – С. 42-47.

118. Михайлычев Е. А. Дидактическая тестология / Е. А. Михайлычев. – М.: Народное образование, 2001. – 432 с.

119. Михайлова Н. Н. Повышение профессиональной компетентности педагога в условиях освоения образовательной технологии / Н. Н. Михайлова, В. Н. Орлова, Г. М. Шеламова // Профессиональное образование. Столица. — 2008. — № 12. — С. 27—37.

120. Молчанова О. Г. Анализ систем тестирования в компьютерных обучающих комплексах / О. Г. Молчанова, О. И. Морозова, И. В. Чумаченко // Системи обробки інформації. - 2012. - Вип. 7. - С. 292-296. - Режим доступу: http://nbuv.gov.ua/UJRN/soi_2012_7_63

121. Морзе Н. В. Основы інформаційно-комунікаційних технологій / Н. В. Морзе – К.: Видавнича група BVH, 2006. – 298 с.

122. Національна доктрина розвитку освіти. Освіта. 2002. № 26. С. 2-4.

123. Національна стратегія розвитку освіти в Україні на період до 2021 року, схвалена Указом Президента України від 25 червня 2013 року №344/2013. [Електронний ресурс]– Режим доступу: <http://zakon0.rada.gov.ua/laws/show/344/2013>.

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

125. Олійник М. М. Тест як інструмент кількісної діагностики рівня знань в сучасних технологіях навчання / М. М. Олійник, Ю. А. Романенко. – Донецьк: ДонНУ, 2001. – 84 с.

126. Орлов А.И. Экспертные оценки. Учебное пособие. / А. И. Орлов – М.: 2002. - 31 с.

127. Офіційний звіт про проведення в 2018 році зовнішнього незалежного оцінювання результатів навчання, здобутих на основі повної загальної середньої освіти. Том 2. — Український центр оцінювання якості освіти. — Київ, 2018. — 350 с. [Електронний ресурс]– Режим доступу: https://dneprtest.dp.ua/docs/2017/pdf/zvit_zno_2018_tom_2.pdf

128. Переверзев В.Ю. Критериально-ориентированные педагогические тесты для итоговой аттестации студентов / В. Ю. Переверзев. – М.: Из-во НМЦ СПО Минобразования РФ, 1998. – 152 с.

129. Переверзев В.Ю. Технология разработки тестовых заданий: справочное руководство / В. Ю. Переверзев. – М.: Е-Медиа, 2005. – 265 с.

130. Петрова Л.И. Организационно-педагогические условия использования информационных технологий в обучении / Л. И. Петрова // Приложение к СПО. – 2005. – №8. – С. 113

131. Положення про кредитно-модульну організацію навчального процесу в НТУУ «КПІ» / Уклад. В. П. Головенкін – К.: ІВЦ “Видавництво «Політехніка»”, 2006. – 55 с. [Електронний ресурс]– Режим доступу: <https://kpi.ua/files/bologna/Pologennya-KMSONP.doc>

132. Поздняков С.А. Исследование точности измерения латентных переменных в образовании. / С.А. Поздняков. – Славянск-на-Кубани: Издательский центр СГПИ, 2007. – 118 с.

133. Про освіту [Електронний ресурс] : закон України від 05.09.2017 р. – Режим доступу : <http://zakon2.rada.gov.ua/laws/show/1060-12/page3>.

134. Про розвиток та досвід експлуатації комплексу дистанційної освіти «Вища математика» [Текст] / [І. В. Алексєєва, В. О. Гайдей, О. О. Диховичний та ін.] // Дидактика математики: проблеми і дослідження : міжнар. зб. наук. робіт. – Вип. 31. – Донецьк: Вид-во ДонНТУ, 2009. – С.49-56.

135. Протасова Н. Г. Післядипломна освіта педагогів: зміст, структура, тенденції розвитку / Н. Г. Протасова. – Державна академія керівних кадрів освіти. – К., 1998. – 176 с.

136. Роберт И. В. Теория и методика информатизации образования (психолого-педагогический и технологический аспекты) / И. В. Роберт. – М.: ИИО РАО, 2007. – 256 с.

137. Самыловский А.И. Тест как объективный измерительный инструмент в образовании / А. И. Самыловский // Вопросы тестирования в образовании. – 2001. – № 1. – С. 10—39.

138. Сафонова Е.И. Рекомендации по использованию инновационных образовательных технологий в учебном процессе / Е. И. Сафонова. – М.: Российский государственный гуманитарный университет, 2011. – 71 с.

139. Сафонцев С. А. Критериально-ориентированное тестирование / С. А. Сафонцев // Педагогическая диагностика. – 2002. – № 1. – С. 77—80.

140. Сергієнко В. П. Комп'ютерні технології в тестуванні: навч. посіб. для студ. вищ. пед. навч. закл. / В. П. Сергієнко, М. П. Малежик, Т. В. Сіткар. – Луцьк: Волиньполіграф, 2012. – 295 с.

141. Симонов П.В. Междисциплинарная концепция человека: потребностно-информационный подход // Человек в системе наук. - М.: Наука, 1989.-С. 58-73.

142. Симонов П. В. Педагогический менеджмент. Учебное пособие / П. В. Симонов. – М. РПА, 1997. – 160 с.

143. Словник української мови: в 11 томах. — Том 5, 1974. — С. 828. [Електронний ресурс] – Режим доступу: <http://sum.in.ua/s/ocinjuvaty>

144. Смирнова, Г. И. Алгоритм обработки матриц результатов тестирования с оценкой 0-1-2 и более с помощью программы RUMM-2010 / Г. И. Смирнова // Педагогические измерения. – 2007. – №4. – С. 86-90.

145. Спірін О. М. Інформаційно-комунікаційні та інформатичні компетентності як компоненти системи професійно-спеціалізованих компетентностей вчителя інформатики [Електронний ресурс] / О. М. Спірін //

Інформаційні технології і засоби навчання. – 2009. – №5. – Режим доступу:
<http://www.ime.edu-ua.net/em13/content/09somtio.htm>.

146. Стоунс Э. Психопедагогика. Психологическая теория и практика обучения / Э. Стоунс. – М. : Педагогика, 1984. – 470 с.

147. Талызина Н. Ф. Теоретические основы контроля в учебном процессе / Н. Ф. Талызина. – М.: Знание, 1983. – 96 с.

148. Толковый словарь Ожегова. [Электронный ресурс] – Режим доступа:

<http://dic.academic.ru/dic.nsf/ogegova/106518/%D0%9C%D0%95%D0%A2%D0%9E%D0%94%D0%98%D0%9A%D0%90>

149. Толковый словарь Ушакова. [Электронный ресурс] – Режим доступа:

<http://dic.academic.ru/dic.nsf/ushakov/859474/%D0%9C%D0%95%D0%A2%D0%9E%D0%94%D0%98%D0%9A%D0%90>

150. Трайнев В. А. Новые информационные коммуникационные технологии в образовании / В. А. Трайнев, В. Ю. Теплышев, И. В. Трайнев. – М.: Дашков и К, 2008. – 320 с.

151. Триус Ю. В. Комп'ютерно-орієнтовані методичні системи навчання математичних дисциплін у вищих навчальних закладах: автореф. дис. на здобуття учен. ступеня доктора пед. наук: спец. 13.00.02 “Теорія і методика професійної освіти” / Ю. В. Триус. – К., 2005. – 51 с.

152. Триус Ю. В. Система електронного навчання ВНЗ на базі MOODLE: Методичний посібник / Ю. В. Триус, І. В. Герасименко, В. М. Франчук // За ред. Ю. В. Триуса. — Черкаси. — 220 с.

153. Удовенко А.Ф. Аналіз якості тестових завдань для комплекту дистанційної освіти «Вища математика» / І.В. Алексєєва, В.О. Гайдей, О.О. Диховичний, Н.Р. Коновалова, Л.Б. Федорова, А.Ф. Удовенко // Теорія та методика електронного навчання: збірник наукових праць. Випуск І. – Кривий Ріг: Видавничий відділ НМетАУ, 2010. – С. 3-9.

154. Дудко А.Ф. Політомічні моделі Мастерса та Андерсена в аналізі якості тестових завдань / О.О. Диховичний, М.О. Шепель, А.Ф. Дудко // Теорія та методика електронного навчання: збірник наукових праць. Випуск III. – Кривий Ріг: Видавничий відділ НМетАУ, 2012. – С. 83-87.

155. Удовенко А.Ф. Сучасні математичні моделі педагогічного тестування / Удовенко А.Ф. // Збірник праць VIII міжнародної молодіжної науково-практичної конференції «Історія розвитку науки, техніки та освіти» за темою «Взаємодія академічної та університетської науки», 18 травня 2010 р. / Укладач Л.П. Пономаренко. – Київ, 2010. – С. 202-206.

156. Ушаков А. А. Диагностика качества физико-математической подготовки студентов в техническом вузе на основе тестовых технологий: дис. ... канд. пед. наук : 13.00.04 / А. А.Ушаков. – Казань, 2010. – 188 с.

157. Фишер Р. А. Статистические методы для исследователей / Р. А. Фишер. – М.: Госстатиздат, 1958. – 267с.

158. Хуторской, А. В. Ключевые компетенции : технология конструирования /А. В. Хуторской // Народное образование. – 2003. – № 5. – С. 55–61.

159. Цыганов, Ш. И. Математические методы педагогических измерений / Ш. И. Цыганов // Вестник Башкирского университета. – 2009. – Т. 14, №3. – С. 1-8.

160. Чельшкова М. Б. Методические рекомендации по разработке педагогических тестов для комплексной оценки подготовленности студентов в вузе / М. Б. Чельшкова, Б. А. Савельев. – М. : ИЦПКПС, 1995. – 77 с.

161. Чельшкова М. Б. Разработка педагогических тестов на основе современных математических моделей / М. Б. Чельшкова. – М.: МИСИС, 1995. – 320 с.

162. Чельшкова М. Б. Теория и практика конструирования педагогических тестов / М. Б. Чельшкова. – М.: Логос, 2002. – 431 с.

163. Шабанова Ю.О. Системний підхід у вищій школі: підруч. для студ. магістратури. Д.: НГУ, 2014. 120 с.

164. Щукина Г. И. Педагогические проблемы формирования познавательных интересов учащихся / Г. И. Щукина – М.: Педагогика, 1988. – 208 с.
165. Энциклопедия психологических тестов. Личность. Мотивация. Потребность / Под ред. А. Карелина. – М., 1997. – 312 с.
166. Ягупов В.В. Педагогіка: Навч. посібник. / В. В. Ягупов. – К.: Либідь, 2002. – 560 с.
167. Andrich D. Rasch Models for Measurement (Quantitative Applications in the Social Sciences). – Sage Publications, Inc, 1988. – 96 p.
168. Baker F. The Basics of Item Response Theory / F. Baker – ERIC Clearinghouse on Assessment and Evaluation, 2001. – 172 p.
169. 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 / A. Birnbaum. — Reading Mass.: Addison–Wesley, 1968. – pp. 397-479.
170. Bloom B. S. Handbook On Formative and Summative Evaluation of Student Learning / B. S. Bloom, J. T. Hastings, G. F. Madaus. — New-York: McGraw-Hill, 1971. – pp. 923.
171. Bock R.D. Estimating item parameters and latent ability when the responses are scored in two or more nominal categories. // Psychometrika, 37. – 1972. – P. 29-51.
172. Brown G. Assessing Student Learning in Higher Education / G. Brown, J. Bull, M. Pendlebury. – London : Routledge, 1997 – 321 p.
173. Dooley L. M. A study of the adoption of computer technology by teachers / L. Dooley, T. Metcalf, A. Martinez // Educational Technology & Society. – 1999. – № 2(4). – P.107-115.
174. Dudko A. F. On the application of IRT-models for the statistical analysis of the quality of educational tests in higher mathematics / O.O. Dykhovichnyi, A.F. Dudko // Ukrainian-German Workshop on Empirical Complete Convergence and other Limit Theorems of Probability Theory, September 23-27, 2013, Koktebel, Ukraine.

175. Dudko A. Statistical analysis of tests in higher mathematics / A. Dykhovychnyi, A. Dudko // Limit theorem in probability theory, number theory and mathematical statistics. International workshop in honour of Prof. V.V. Buldygin (October 10-12, 2016) – NTUU «KPI» Kyiv, Ukraine, 2016. – P. 18.
176. Duncan O.D. Rasch measurement: Further examples and discussion. / O.D. Duncan. In C. F. Turner & E. Martin (Eds.) // Surveying subjective phenomena. – New York, New York: Russell Sage Foundation, 1984. – Vol. 2 – P. 367-403.
177. European Commission. The European Qualifications Framework for Lifelong Learning (EQF). – Luxembourg: Office for official publications of the European Communities. – 2008. – 15 p.
178. Guizzardi R. Agent-oriented knowledge management in learning environments: a peer-to-peer helpdesk case study / R. Guizzardi, L.Aroyo, G.Wagner // American Association for Artificial Intelligence. – 2002. – 57-72 pp.
179. Gulliksen H. Theory of mental tests. – N-Y. Willey. – 1950. – 486 pp.
180. Guttman L. A basis for analyzing test-retest reliability. / L. A Guttman // Psychometrika. – 1945. – №10. – P.255-282.
181. Hambleton R.K. Application of Item Response Theory. / R.K.Hambleton. – Vancouver. BC: Educ Res Inst B.C., 1983. – 167 p.
182. Hambleton R.K. Fundamentals of Item Response Theory / R.K. Hambleton, H. Swaminathan, H.J. Rogers – Sage publications, 1991.
183. Intel. Навчання для майбутнього. – К.: Видавництво “Нора-прінт”, 2006. – 1032 с.
184. International Standard Classification of Education. ISCED 1997/ UNESCO [Electronic resource]. 1997. – Mode of access: www.uis.unesco.org/en/pub/pub. – date of access 2009.
185. Linacre J.M. Auser’s guide and manual to WINSTEPS: Raschmodel computer programs. / J.M. Linacre. – Chicago: Linacre, 2004. – 213 pp.

186. Linn R. Educational measurement (3 Ed-n) / R. Linn. – N.Y. : Macmillan, 1989. – 610 p.
187. Linden W. Handbook of Modern Item Response Theory / W. Linden, R. Hambleton. – NY: Springer-Verlag, 1997. – 510 p.
188. Linn R. Educational measurement (3 Ed-n) / R. Linn. – N.Y. : Macmillan, 1989. – 610 p.
189. Lord F. M. Application of Item Response Theory to Practical Testing Problems. Hillsdale N-J. Lawrence Erlbaum Ass., Publ., 1980, -266 p
190. Lord F.M. Maximum likelihood estimation of item response parameters when some responses are omitted / F. M. Lord // Psychometrika, 48. – P. 477-482.
191. Lord F.M. Statistical Theories of Mental Test Scores / F. M. Lord, Novick M. – Mass. Addison-Wesley Publ. Co. Reading, 1968.
192. Masters. G. N. A Rasch model for partial credit scoring // Psychometrika. – 1982. – Vol. 47, №2, June 1982. – P. 149-174.
193. Masters, G. N. The analysis of partial credit scoring / G. N. Masters // Applied measurement in education. – 1988. – Issue 7 (4). – P. 279-297.
194. Masters, G. N. The partial credit model and null categories / G. N. Masters, M. Wilson // Psychometrika. – 1993. – Vol. 58, №1, March 1993. – P. 87-99.
195. Muraki E. Information Functions of the Generalized Partial Credit Model / E. Muraki // Applied Psychological Measurement. – Vol 17(4). – 1993. – pp. 351-363.
196. Ostini, R., Nering, M. L. Polytomous item response theory models. – Thousand Oaks: Sage, 2006. – 107 p.
197. Penfield R.D. A New Response Model for Multiple-Choice Items. / R.D. Penfield. – University of Miami & Jimmy de la Torre Rutgers, The State University of New Jersey. – 2008. – 39 p.
198. Rasch G. Probabilistic models for some intelligence and attainment tests / G. Rasch. – Chicago: The University of Chicago Press, 1981. – 199 p.

199. Rusch T. Psychometrics with R: A review of CRAN packages for Item Response Theory / T. Rusch, P. Mair, R. Hatzinger. – Discussion Paper Series of the Center for Empirical Research, 2013. – 31 p.
200. Russell J.D. Modular Instruction / J.D. Russell / Minneapolis, Minn., Burgess Publishing Co., 1974. – 118 p.
201. Samejima F. Estimation of Latent Ability Using a Response Pattern of Graded Scores / F. Samejima // Psychometrika Monograph. – № 17. – 1969. – pp. 1-100.
202. Samejima F. A new family of models for the multiple choice item. – Research Report No. 79-4. – University of Tennessee, Department of Psychology. – 1979.
203. SAT Suite of Assessments Technical Manual. – 2017. – 156 p. – <https://collegereadiness.collegeboard.org/pdf/sat-suite-assessments-technical-manual.pdf>
204. Smith E.V. Introduction to Rasch Measurement. Theory, Models and Applications / E.V. Smith, M.S. Smith. – Maple Grove, Minnesota: JAM Press, 2004. – 689 pp.
205. Smith R.M. Applications of Rasch measurement / R.M. Smith. – Chicago, Illinois: MESA Press, 1992. – 268 pp.
206. Smith R.M. IPARM: Item and person analysis with the Rasch model / R.M. Smith. – Chicago, Illinois: MESA Press, 1991. – 197 pp.
207. Spearman C. Correlation calculated from faulty data //British Journal of Psychology, 1910, Vol.3, N2. -P.271-295.
208. Spearman. C. “General intelligence,” objectively determined and measured / C. Spearman // American Journal of Psychology, 15, 1904. – P. 201-293.
209. Thissen D. Response model for multiple choice items / D. Thissen , L. Steinberg // Psychometrika, 49. – 1984. – P. 501-519.
210. Thissen D. Test scoring. / D. Thissen, H. Wainer. – Lawrence Erlbaum Associates, Mahwah NJ, 2001. – 440 pp.

211. Wainer H., Messik S. Principals of Modern Psychological Measurement. – New Jersey, Lawrence Erlbaum Associates, 1983. – 377 p.

212. Wilson, M. Constructing measures: an item response modeling approach. – Mahwah: Lawrence Erlbaum Associates, 2005. – 228 p.

ДОДАТКИ

Додаток А

Опис інтерфейса СААЯТЗ

Керування системою та вибір інструментів системи здійснюється за допомогою п'яти пунктів меню (рис. А1):

- Аналіз даних,
- База даних,
- Допомога,
- Про програму,
- Вихід.

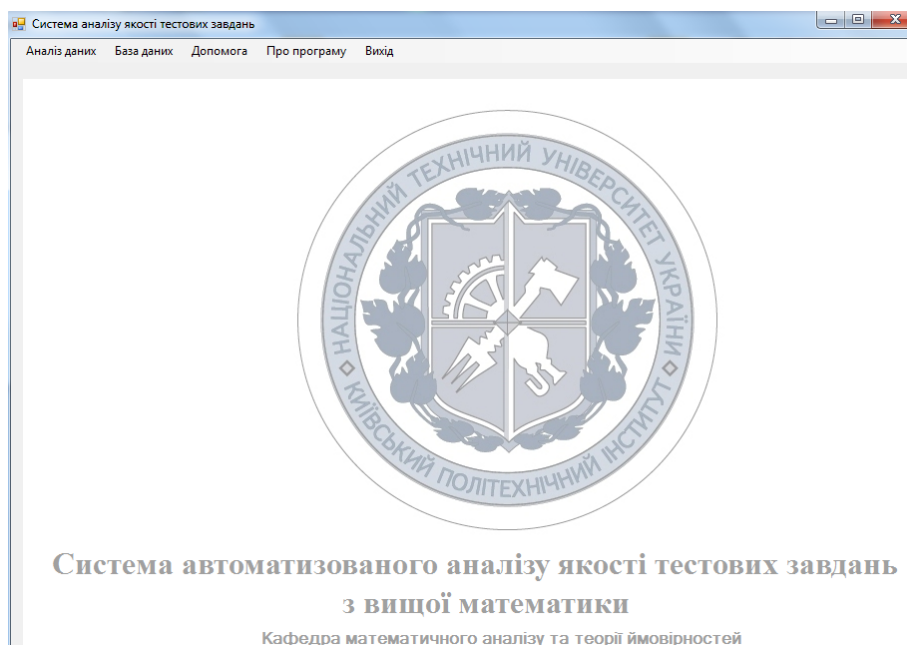


Рис. А1. Стартова сторінка СААЯТЗ

Опишемо кожний з перелічених пунктів

Пункт «**Аналіз даних**» містить у собі необхідні інструменти для аналізу якості окремих тестових завдань і тесту в цілому та складається з підпунктів:

1. КТТ аналіз:

- Характеристики тесту;
- Дисперсійний аналіз,
- Кореляційний аналіз,
- Перевірка рівності середніх (t-критерій Стьюдента),
- Характеристики завдань:

2. IRT аналіз:
 - Розрахунок,
 - Результати:
3. Продивитись текст завдань.
4. Довідка.

Аналіз даних → **КТТ аналіз** → **Характеристики тесту** → **Первинний аналіз**. Пункт меню дає змогу завантажити таблицю результатів тестування у форматі .xls, експортовану із системи Moodle, продивитись результати тестування та провести первинний аналіз завантажених (рис. А2).

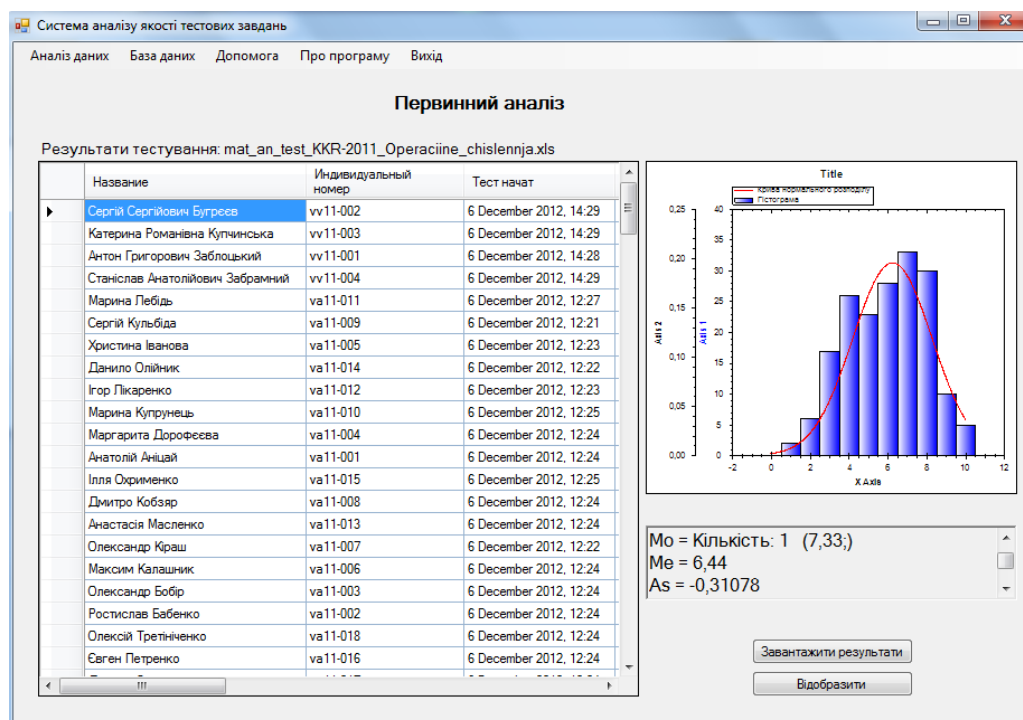


Рис. А2. Первинний аналіз даних в СААЯТЗ

При цьому розраховуються математичне сподівання μ , дисперсія σ^2 , критерій узгодженості χ^2 , середнє вибіркве Sx , мода Mo , медіана Me , асиметрія As , перевірка гіпотези про нормальний розподіл вибірки. На екран виводиться гістограма та крива теоретичного нормального розподілу.

Аналіз даних → **КТТ аналіз** → **Характеристики тесту** → **Коефіцієнт надійності**. Меню дає змогу завантажити таблицю результатів тестування у форматі .xls, експортовану із системи Moodle, обчислити коефіцієнт надійності тесту r , знайдений на основі методу розщеплення, і коефіцієнт

надійності Кронбаха α та отримати висновки про надійність тесту на основі обчислених коефіцієнтів (рис. А3).

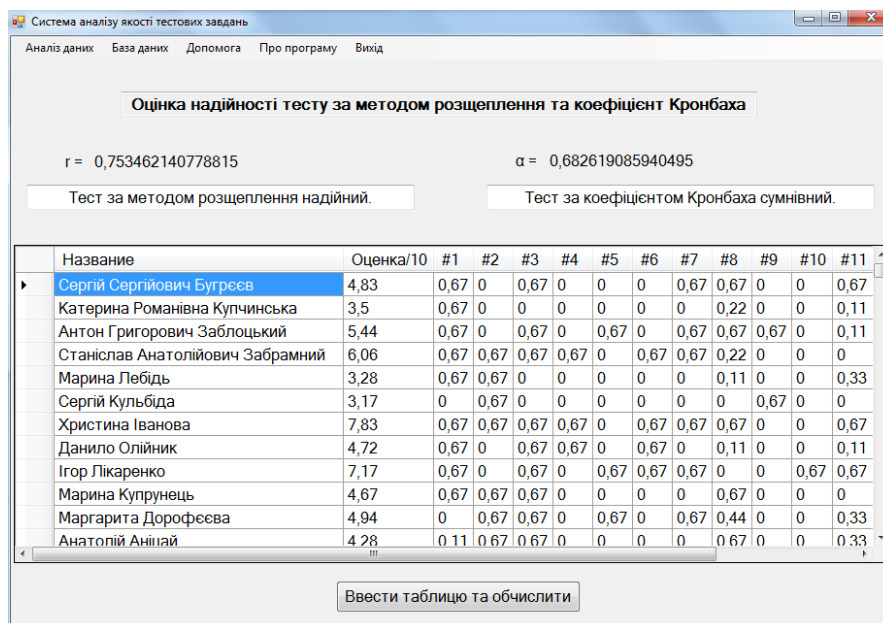


Рис. А3. Обчислення коефіцієнтів надійності тесту в СААЯТЗ

Аналіз даних → **КТТ аналіз** → **Характеристики тесту** → **Коефіцієнт валідності**. Меню дає змогу оцінити, як результати тесту співвідносяться з іншими незалежними експертними оцінками. Для цього необхідно завантажити дві таблиці: таблицю тестових оцінок у форматі .xls, експортовану із системи Moodle та таблицю експертних оцінок у форматі .xls, яку заповнює експерт. Порівнюючи оцінки, програма обчислює коефіцієнт критеріальної валідності та робить висновок про співвідношення оцінок (рис. А4).

Аналіз даних → **КТТ аналіз** → **Характеристики тесту** → **Дисперсійний аналіз**. За допомогою однофакторного дисперсійного аналізу (ANOVA), викладач вивчає вплив тих чи інших факторів на мінливість середніх. У програмі за фактор, що впливає на оцінювання знань студентів, приймається тема тестування. Пункт меню дає змогу одночасно для декількох таблиць результатів тестування обчислити середнє, дисперсію та об'єм вибірки, а також провести дисперсійний аналіз – факторний (між вибірками), залишковий (у вибірці), загальний (обчислюються сума квадратів

відхилень, кількість ступенів вільності, середній квадрат (оцінка дисперсії) та F -критерій) (рис. A5).

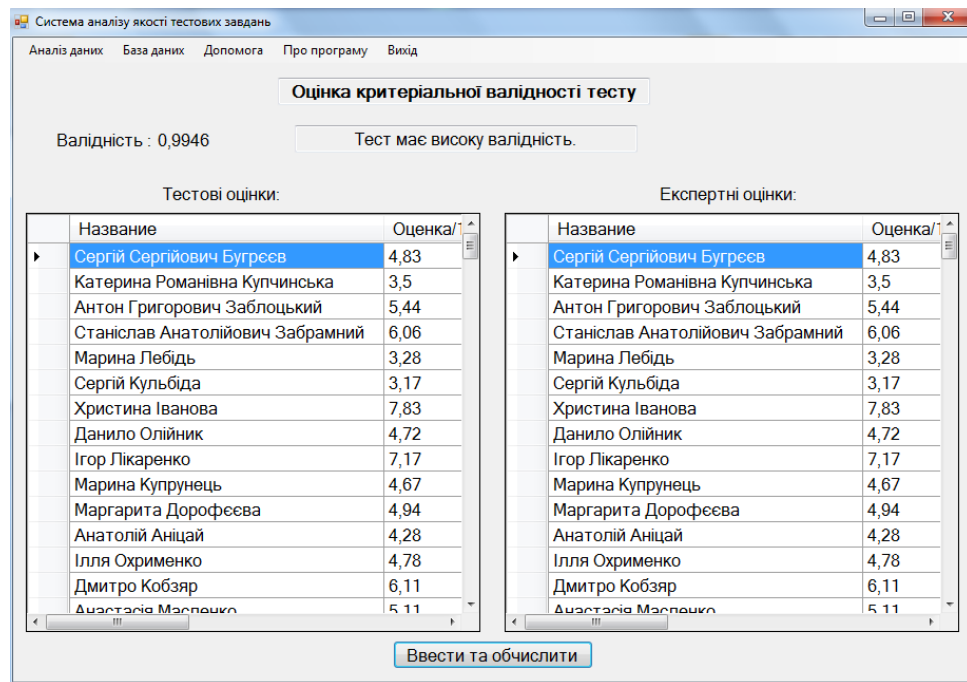


Рис. A4. Оцінювання критеріальної валідності в СААЯТЗ

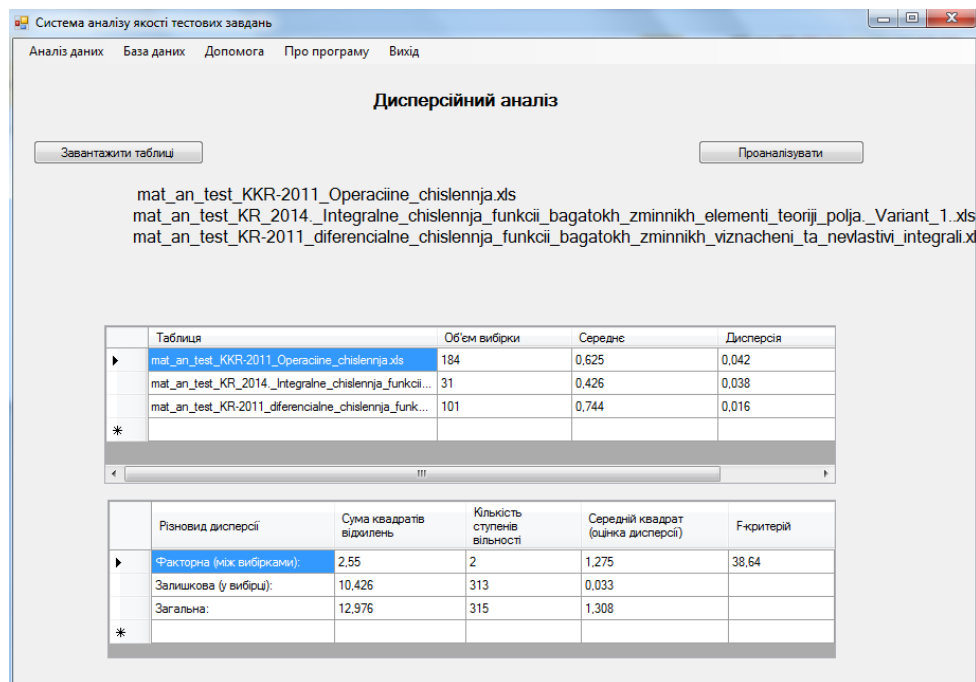


Рис. A5. Дисперсійний аналіз тесту в СААЯТЗ

Аналіз даних → **КТТ аналіз** → **Характеристики тесту** → **Кореляційний аналіз**. Пункт меню дає змогу переглядати статистичне дослідження залежності між випадковими величинами, яке у програмі виражено у кореляційному аналізі – розрахунку коефіцієнта Пірсона та

коефіцієнта рангової кореляції Спірмена індивідуальних балів іспитників різних тестів (рис. А6).

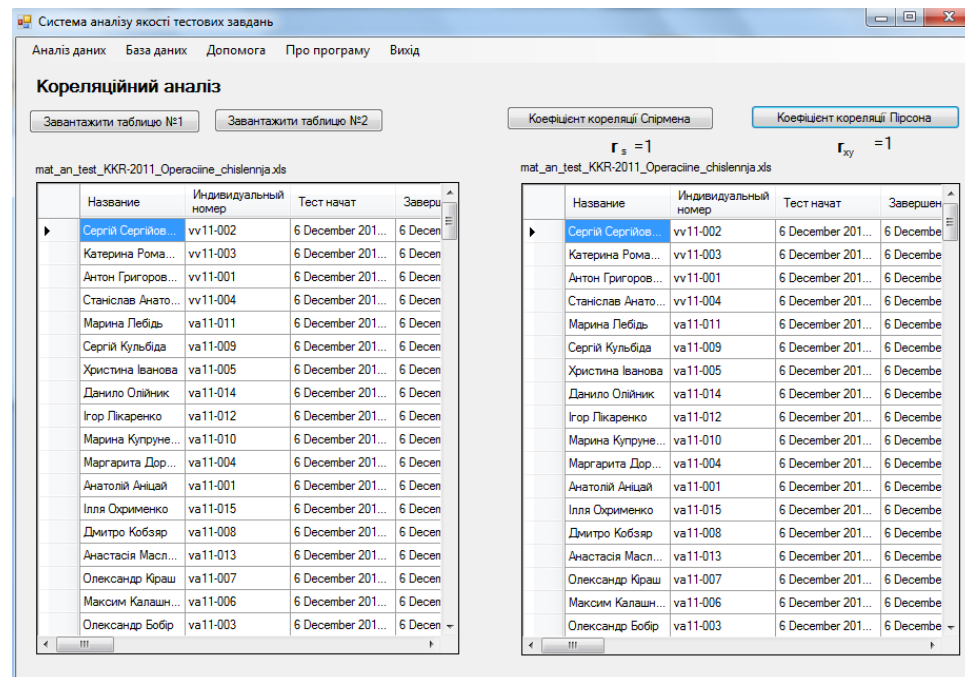


Рис. А6. Кореляційний аналіз тесту в СААЯТЗ

Аналіз даних → КТТ аналіз → Характеристики тесту → Перевірка рівності середніх (t-критерій Стюдента). Пункт меню дає змогу перевірити рівності середніх значень у двох вибірках результатів тестування (рис. А7).

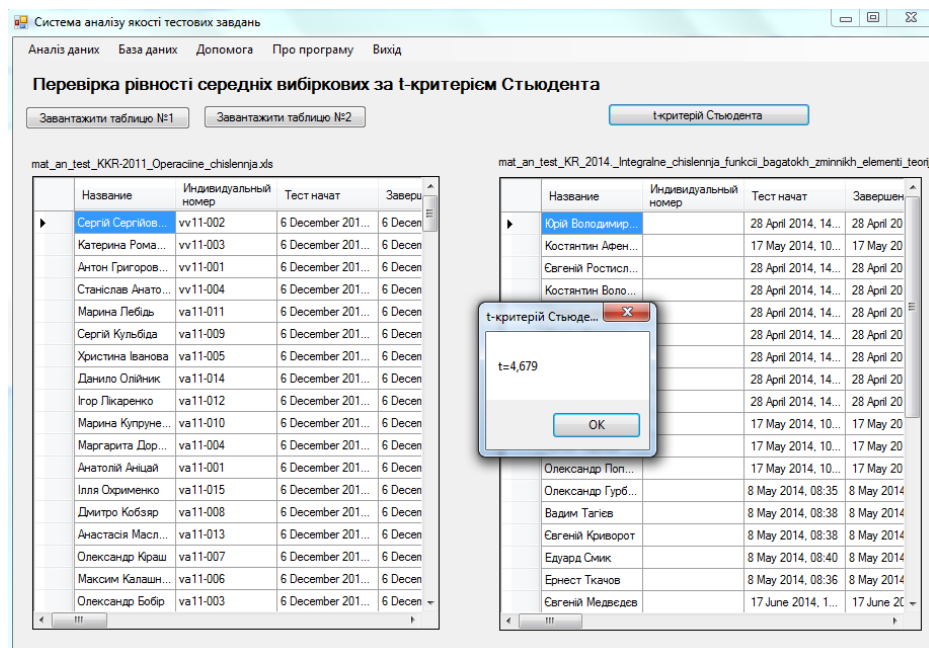


Рис. А7. Перевірка рівності середніх значень у двох вибірках результатів тестування в СААЯТЗ

Аналіз даних → **КТТ аналіз** → **Характеристики тесту** → **Кореляційна матриця**. Пункт меню дає змогу обчислити кореляційну матрицю завдань та отримати інформацію про можливі некоректність та дублювання завдань (рис. А8).

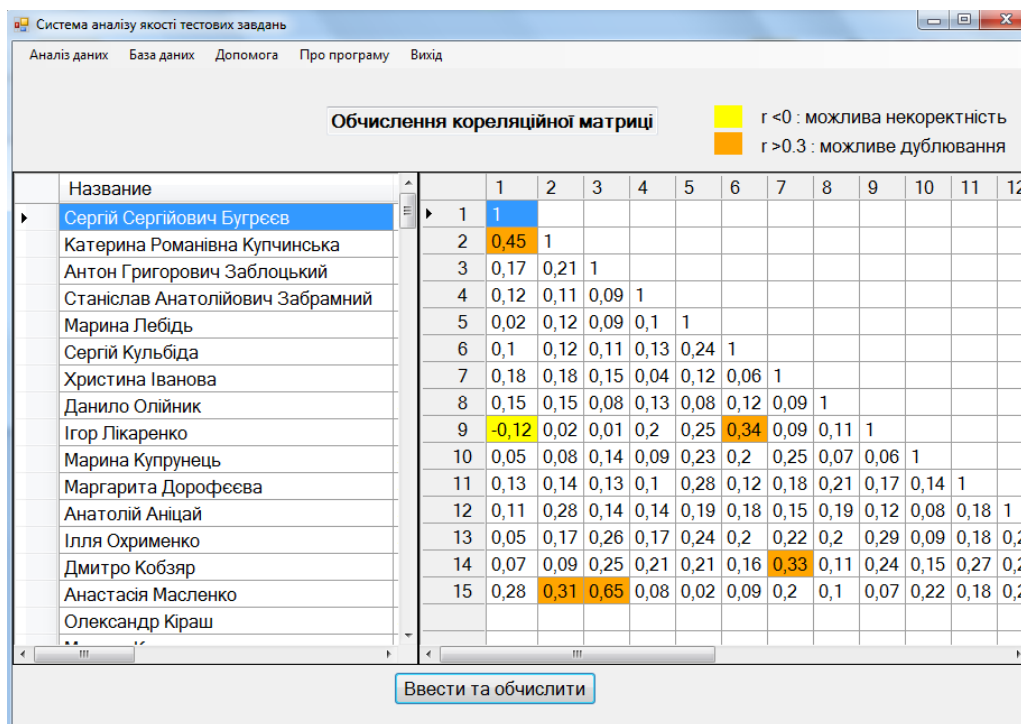


Рис. А8. Обчислення кореляційної матриці завдань в СААЯТЗ

Аналіз даних → **КТТ аналіз** → **Характеристики тесту** → **Оцінювання валідності завдань**. Меню дає змогу обчислити показники валідності окремих завдань в рамках КТТ (значення індексу розрізняльної здатності дихотомічних завдань D_j , коефіцієнту точкової бісеріальної кореляції та бісеріального коефіцієнту кореляції між дихотомічним завданням та індивідуальними балами іспитників і коефіцієнту кореляції Пірсона між політомічним завданням та індивідуальними балами іспитників $r_{j\cdot}$) та зробити висновки щодо якості завдання та подальших дій із ним. (рис. А9).

Аналіз даних → **IRT аналіз** → **Розрахунок**. Пункт меню дає змогу обчислити, зберегти на комп'ютері та переглянути латентні параметри тестових завдань за моделями Раша, Бірнбаума та Андерсена.

Аналіз даних → **IRT аналіз** → **Результати** → **Результати IRT аналізу**. Меню надає можливість переглянути збережені на комп'ютері

результати проведеного IRT-аналізу. Викладач має можливість обрати номер завдання, продивитися відповідні параметри складності дихотомічних завдань, параметри складності рівнів політомічних завдань та графіки характеристичних та інформаційних кривих (рис. А10).

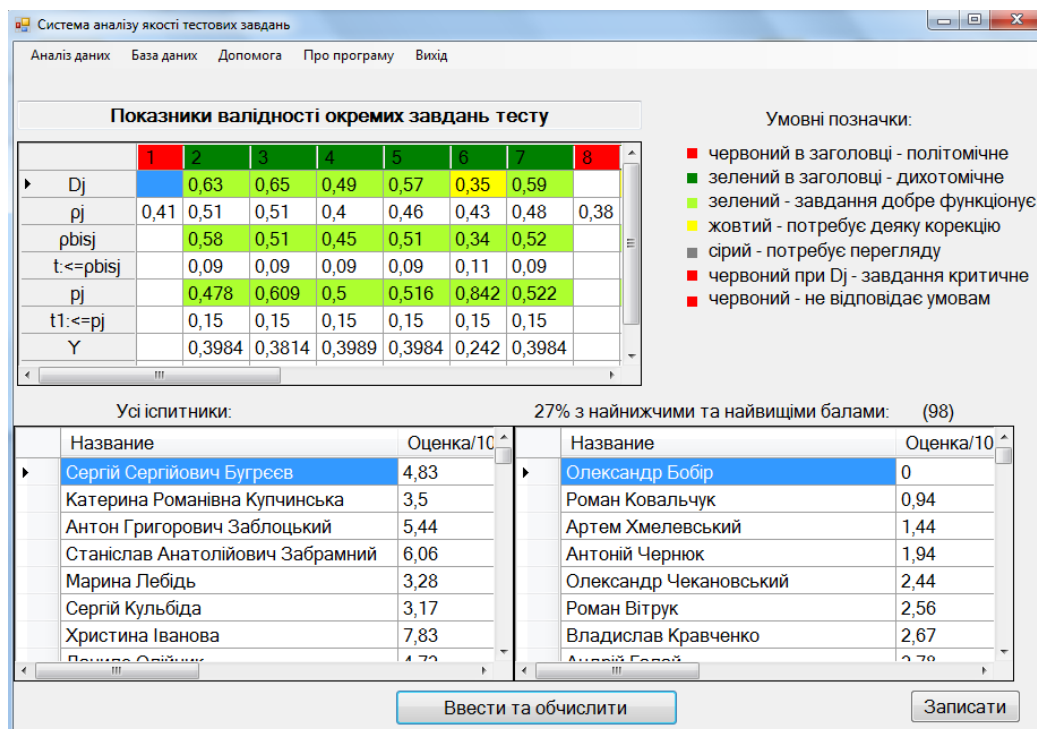


Рис. А9. Обчислення показників валідності тестових завдань в СААЯТЗ

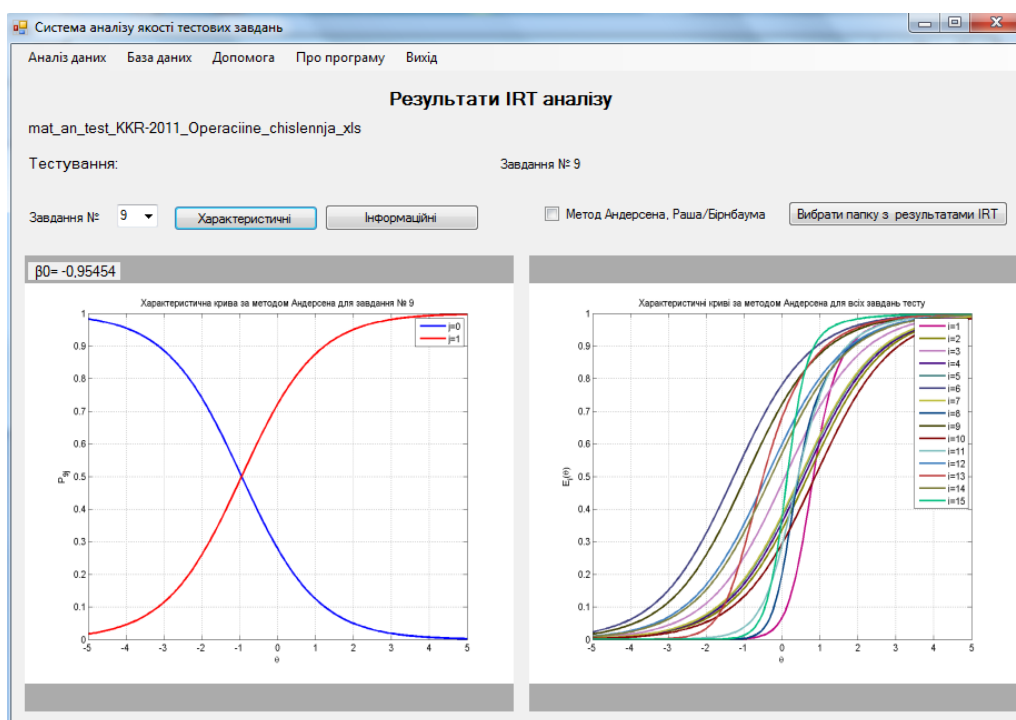


Рис. А10. Переглядд результатів IRT-аналізу в СААЯТЗ

Аналіз даних → **IRT аналіз** → **Результати** → **Продивитися параметри студентів**. Пункт меню дає змогу продивитись оцінені та збережені на комп'ютері параметри підготовленості студентів. (рис. A11).

	Название	Индивидуальный номер	Тест начат	Завершено	Затраченное время	Оценка/10	Подготовленность
1	Сергей Сергеевич	vv11-002	6 December 2012, 1...	6 December 2012, ...	1 ч 15 мин	4,83	0,463065
2	Катерина Романів...	vv11-003	6 December 2012, 1...	6 December 2012, ...	1 ч 14 мин	3,5	0,155585
3	Антон Григорович ...	vv11-001	6 December 2012, 1...	6 December 2012, ...	1 ч 13 мин	5,44	0,648858
4	Станіслав Анатолій...	vv11-004	6 December 2012, 1...	6 December 2012, ...	1 ч 6 мин	6,06	0,380931
5	Марина Лебідь	va11-011	6 December 2012, 1...	6 December 2012, ...	1 ч 6 мин	3,28	0,007299
6	Сергій Кульбіда	va11-009	6 December 2012, 1...	6 December 2012, ...	1 ч 7 мин	3,17	-0,35508
7	Христина Іванова	va11-005	6 December 2012, 1...	6 December 2012, ...	1 ч	7,83	1,226492
8	Данило Олійник	va11-014	6 December 2012, 1...	6 December 2012, ...	1 ч 12 мин	4,72	0,228764
9	Ігор Пикаренко	va11-012	6 December 2012, 1...	6 December 2012, ...	1 ч 12 мин	7,17	0,551615
10	Марина Купрунець	va11-010	6 December 2012, 1...	6 December 2012, ...	1 ч 9 мин	4,67	0,228764
11	Маргарита Дорощ...	va11-004	6 December 2012, 1...	6 December 2012, ...	1 ч 9 мин	4,94	0,155585
12	Анатолій Аніцай	va11-001	6 December 2012, 1...	6 December 2012, ...	1 ч 10 мин	4,28	0,303372
13	Ілля Охрименко	va11-015	6 December 2012, 1...	6 December 2012, ...	1 ч 10 мин	4,78	0,463065
14	Дмитро Кобзар	va11-008	6 December 2012, 1...	6 December 2012, ...	1 ч 12 мин	6,11	0,648858
15	Анастасія Масленко	va11-013	6 December 2012, 1...	6 December 2012, ...	1 ч 9 мин	5,11	0,380931
16	Олександр Кіраш	va11-007	6 December 2012, 1...	6 December 2012, ...	1 ч 13 мин	4,67	0,380931
17	Максим Калашник	va11-006	6 December 2012, 1...	6 December 2012, ...	1 ч 10 мин	3,72	0,007299
18	Олександр Бобір	va11-003	6 December 2012, 1...	6 December 2012, ...	1 ч 12 мин	0	Нет
19	Ростислав Бабенко	va11-002	6 December 2012, 1...	6 December 2012, ...	1 ч 21 мин	7,17	0,757922
20	Олексій Тютніченко	va11-018	6 December 2012, 1...	6 December 2012, ...	1 ч 9 мин	10	Нет

Рис. A11. Перегляд параметрів підготовленості студентів в СААЯТЗ

Аналіз даних → **Продивитись текст завдання**. Пункт меню дає можливість у будь-який момент продивитись текст завдання тесту (рис. A12).

Виберіть правильні твердження.

Выберите по крайней мере один ответ:

☐ a. Градієнтом функції $z = f(x, y)$ називають скаляр $\text{grad} z = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}$.

☒ b. Градієнтом функції $z = f(x, y)$ називають вектор $\text{grad} z = \frac{\partial z}{\partial x} \bar{i} + \frac{\partial z}{\partial y} \bar{j}$.

☒ c. Якщо функція $z = f(x, y)$ диференційовна в точці (x_0, y_0) , то вона неперервна в цій точці.

☐ d. Якщо функція $z = f(x, y)$ диференційовна в точці (x_0, y_0) , то вона дорівнює нулеві в цій точці.

Рис. A12. Перегляд тексту завдання в СААЯТЗ

Аналіз даних → **Довідка**. Пункт меню надає інформацію про загальні характеристики тесту: кількість студентів, кількість завдань, тему тестування та шлях xls-файлу (рис. А13).

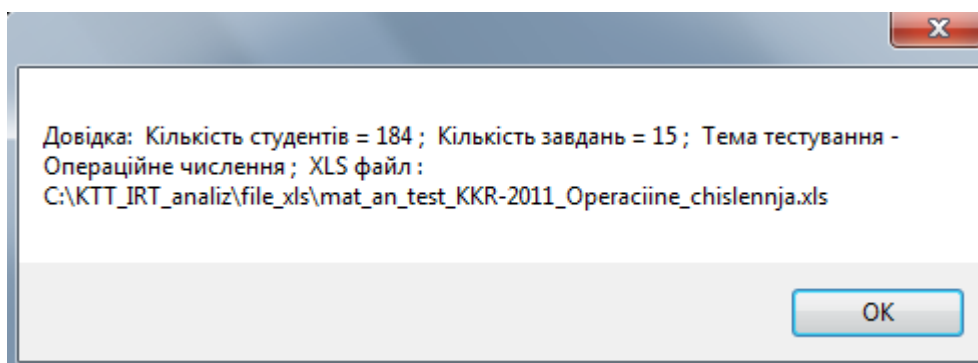


Рис. А13. Довідка про загальну характеристику тесту в СААЯТЗ

Пункт меню **База даних**. Основним результатом роботи системи є формування бази каліброваних завдань, в якій зберігаються окремі тестові завдання, систематизовані за темами, призначенням (контрольна, іспит тощо) та формою тестового завдання. Для кожного завдання існує відповідний паспорт, який містить основні характеристики тестового завдання, оцінені системою. База каліброваних завдань є невід’ємною складовою, необхідною для роботи викладача з тестами з вищої математики, а саме для формування та редагування тестів.

Робота системи з обслуговування бази каліброваних тестових завдань передбачає:

- редагування та оновлення бази;
- підбір та перегляд завдань за заданими параметрами;
- створення нових тестів;
- розподіл тестів за темами;
- перегляд тестів;

Пункт меню **Допомога** призначено для зручності роботи з програмою в цілому, з окремими її складовими та пояснення способу правильного налаштування програми (рис. А14).

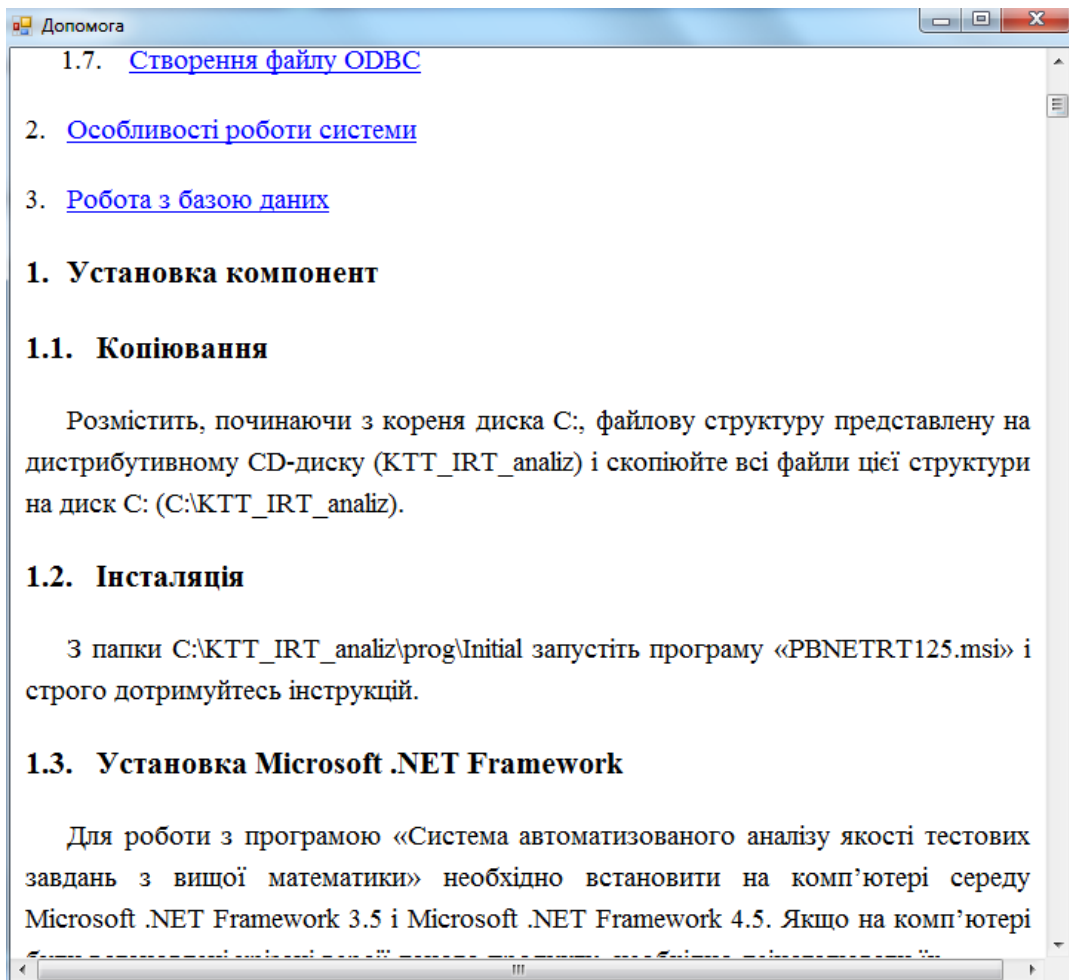


Рис. А14. Пункт меню «Допомога» в СААЯТЗ

Пункт меню ***Про програму*** відкриває вікно з відомостями про програму СААЯТЗ. Пункт меню ***Вихід*** дає змогу закрити вікно програми.

Анкета для викладачів вищої математики

Шановні колеги!

Просимо Вас відповісти на запитання анкети, основною метою якої є з'ясування сучасного стану використання та оцінювання якості комп'ютерних тестів з вищої математики для контролю знань студентів.

Інструкція: Анкета складається з двох блоків питань. Якщо Ви використовуєте тести для контролю знань студентів з вищої математики – відповідайте тільки на питання блоку А, якщо не використовуєте – тільки блоку Б.

Для заповнення анкети Вам необхідно обрати варіант відповіді, що відповідає Вашій точці зору та поставити позначку (✓) у відповідній графі.

Оскільки дані анкети є анонімними, просимо Вас надати відверті відповіді.

Дякуємо за увагу!

Блок А

1. Який тип тестування Ви використовуєте для контролю знань студентів з вищої математики?
 - ☐ бланковие;
 - ☐ комп'ютерне;
 - ☐ комп'ютерне адаптоване.
2. Які форми тестових завдань Ви використовуєте для контролю знань студентів з вищої математики?
 - ☐ завдання з пропусками;
 - ☐ завдання вільного викладу;
 - ☐ завдання з короткою відповіддю;
 - ☐ завдання з множинним вибором;
 - ☐ завдання на відповідність;
 - ☐ завдання альтернативних відповідей;
 - ☐ завдання на встановлення правильної послідовності.
 - ☐ інші (_____).
3. Яку систему комп'ютерного тестування Ви використовуєте?
 - ☐ Moodle
 - ☐ Інше (_____)
4. Чи спрощує процедуру перевірки робіт студентів застосування комп'ютерного тестування як засобу контролю знань з вищої математики?
 - ☐ так;
 - ☐ ні;
 - ☐ не можу відповісти.
5. Чи підвищує об'єктивність контролю знань студентів з вищої математики використання тестів?
 - ☐ так;
 - ☐ ні;
 - ☐ не можу відповісти.

6. Чи необхідне, на Вашу думку, оцінювання якості тестів з вищої математики перед впровадженням їх в освітній процес?
- ☐ так;
 - ☐ ні;
 - ☐ не можу відповісти.
7. Чи можете Ви самостійно оцінювати якість тестів з вищої математики?
- ☐ так;
 - ☐ ні.
8. Чи оцінюєте Ви якість тестів перед впровадженням їх в освітній процес?
- ☐ так;
 - ☐ ні.
9. Які методи оцінювання якості тестів Ви знаєте та вмієте застосовувати?
- ☐ методи класичної теорії тестів;
 - ☐ методи Item Response Theory;
 - ☐ інші (_____);
 - ☐ жодні.
10. Які методи оцінювання якості тестів Ви застосовуєте до розроблених тестів?
- ☐ класичні статистичні методи;
 - ☐ методи Item Response Theory;
 - ☐ метод експертного оцінювання;
 - ☐
 - ☐ інші (_____);
 - ☐ жодні.
11. Чи вважаєте Ви за потрібне використовувати ІКТ для оцінювання якості тестів з вищої математики?
- ☐ так;
 - ☐ ні;
 - ☐ не можу відповісти.

12. Чи достатньо у Вашому розпорядженні програмних засобів для оцінювання якості тестів з вищої математики?
- ☐ так;
- ☐ ні;
- ☐ не можу відповісти.
13. Чи достатній Ваш рівень знань для оцінювання якості тестів з вищої математики?
- ☐ так;
- ☐ ні;
- ☐ не можу відповісти.
14. Чи хотіли б Ви опанувати методику оцінювання якості тестів з вищої математики з використанням ІКТ?
- ☐ так;
- ☐ ні.
15. Чи вважаєте Ви за потрібне підвищити рівень компетентності викладачів щодо оцінювання якості тестів з вищої математики?
- ☐ так;
- ☐ ні;
- ☐ не можу відповісти.

Стаж викладацької діяльності: до 5 років ☐; від 5 до 10 років ☐; від 10 до 15 років ☐; більше 15 років ☐

Науковий ступінь: доктор наук ☐; кандидат наук ☐; без ступеня ☐

Наукове звання: професор ☐; доцент ☐; ін. ☐ (уточніть)

Посада: професор ☐; доцент ☐; старший викладач ☐; асистент ☐.

Блок Б

1. Чи вважаєте Ви за доцільне використання комп'ютерних тестів для оцінювання знань студентів з вищої математики при проміжному та підсумковому контролі?
 - ☐ вважаю доцільним;
 - ☐ вважаю доцільним використання тестів лише як допоміжного засобу контролю знань (основний – традиційні контрольні роботи);
 - ☐ вважаю недоцільним;
 - ☐ не можу відповісти.
2. Чи є у Вас досвід розроблення тестів для контролю знань студентів з вищої математики?
 - ☐ так;
 - ☐ ні.
3. Якому типу тестування для контролю знань студентів з вищої математики Ви б надали перевагу?
 - ☐ бланковому;
 - ☐ комп'ютерному;
4. З якою системою комп'ютерного тестування Ви маєте досвід роботи?
 - ☐ з Moodle;
 - ☐ Інше (_____);
 - ☐ з жодною.
5. Чому Ви не застосовуєте тестування для контролю знань студентів з вищої математики?
 - ☐ не довіряю тестуванню, як засобу контролю;
 - ☐ не вмію складати якісні тести;
 - ☐ не хочу щось змінювати в контролі знань, оскільки традиційні контрольні роботи мене цілком влаштовують;
 - ☐ інше (_____).

6. Чи сприяє, на Вашу думку, використання тестування підвищенню об'єктивності контролю знань студентів з вищої математики?
- ☐ так;
 - ☐ ні;
 - ☐ не можу відповісти.
7. Чи хотіли б Ви спростити процедуру перевірки контрольних робіт студентів з вищої математики?
- ☐ так;
 - ☐ ні;
 - ☐ не можу відповісти.
8. Чи необхідне, на Вашу думку, оцінювання якості тестів з вищої математики перед впровадженням їх в освітній процес?
- ☐ так;
 - ☐ ні;
 - ☐ не можу відповісти.
9. Які методи оцінювання якості тестів Ви знаєте?
- ☐ методи класичної теорії тестів;
 - ☐ методи Response Theory;
 - ☐ інші (_____);
 - ☐ жодні.
10. Чи вважаєте Ви за потрібне використовувати ІКТ для оцінювання якості тестів з вищої математики?
- ☐ так;
 - ☐ ні;
 - ☐ не можу відповісти.
11. Використовували б Ви для контролю знань студентів комп'ютерні тести з вищої математики, які пройшли апробацію та якість яких було підтверджено статистично?
- ☐ так;

- ☐ ні;
- ☐ не можу відповісти.

12. Чи достатній Ваш рівень знань для оцінювання якості тестів з вищої математики?

- ☐ так;
- ☐ ні;
- ☐ не можу відповісти.

13. Чи сприяла б використанню Вами тестування з вищої математики для контролю знань студентів достатня кількість знань щодо оцінювання якості тестів?

- ☐ так;
- ☐ ні;
- ☐ не можу відповісти.

14. Чи вважаєте Ви за потрібне підвищити рівень компетентності викладачів щодо оцінювання якості тестів з вищої математики?

- ☐ так;
- ☐ ні;
- ☐ не можу відповісти.

15. Чи хотіли б Ви підвищити свій рівень компетентності щодо оцінювання якості тестів з вищої математики?

- ☐ так;
- ☐ ні;
- ☐ не можу відповісти.

Стаж викладацької діяльності: до 5 років ☐; від 5 до 10 років ☐; від 10 до 15 років ☐; більше 15 років ☐

Науковий ступінь: доктор наук ☐; кандидат наук ☐; без ступеня ☐

Наукове звання: професор ☐; доцент ☐; ін. ☐ (уточніть)

Посада: професор ☐; доцент ☐; старший викладач ☐; асистент ☐.

Програма курсу «Формування компетентності викладачів щодо оцінювання якості тестів з вищої математики»

Мета курсу: отримати знання про методику оцінювання якості тестів з вищої математики на принципах компетентнісного підходу та опанувати цю методику.

Завдання курсу:

- Засвоїти базові теоретичні відомості про оцінювання якості тестів з вищої математики.
- Здобути навички розроблення тестів з вищої математики та оцінювання їхньої якості.
- Вивчити базовий функціонал СААЯТЗ.
- Опанувати комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики.
- Розробити в системі Moodle тест з вищої математики, оцінити якість даного тесту з використанням комп'ютерно орієнтованої методики.

Очікувані результати навчання:

- Ознайомлення з компетентнісним підходом у навчанні.
- З'ясування основних проблем, пов'язаних із процесом розроблення тестів з вищої математики, та вміння їх вирішувати.
- З'ясування та усвідомлення переваг використання тестів для контролю знань студентів з вищої математики.
- З'ясування та усвідомлення ролі оцінювання якості тестів з вищої математики при їхньому використанні для контролю знань студентів.
- Знання методів КТТ та IRT оцінювання якості тестів.
- Знання базового функціоналу СААЯТЗ.

- Навички роботи з СААЯТЗ.
- Опанування комп'ютерно орієнтованої методики оцінювання якості тестів з вищої математики.
- Розроблення в системі Moodle тесту з вищої математики.
- Оцінювання якості тесту з вищої математики з використанням комп'ютерно орієнтованої методики.

Загальні вимоги до початкового рівня викладачів вищої математики:

1. Наявність знань про загальні принципи розроблення та використання комп'ютерних тестів з вищої математики

2. Наявність знань про специфіку контролю знань студентів із використанням комп'ютерних тестів. Розуміння ролі і місця комп'ютерних тестів в контролі знань студентів.

3. Наявність уявлень про можливості інформатизації педагогічної діяльності, зокрема використання засобів ІКТ для контролю знань студентів.

4. Наявність базових знань та умінь роботи в системі Moodle

5. Наявність загальних уявлень про системи комп'ютерного тестування.

6. Наявність базової ІК-компетентності, що включає, загальні уявлення щодо призначення і функціонування персональних комп'ютерів, периферійних пристроїв, комп'ютерних мереж та технологій Інтернет.

Залежно від рівня ІК-компетентності слухача є можливість змінити кількість тематичних модулів Навчальної програми. Визначається рівень компетентності викладача щодо оцінювання знань студентів завдяки тестам та анкетам.

Навчальна програма складається з п'яти модулів:

МОДУЛЬ 1. ПРОБЛЕМИ РОЗРОБЛЕННЯ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ

Тематичний модуль 1

Назви тем	Кількість навчальних годин		
	Семінари, практичні заняття, тренінги	Лекції	Всього
1. Використання тестів для контролю знань студентів		1	1
2. Форми тестових завдань	1	1	2
3. Проблема вибору підходу до розроблення тестів з вищої математики		1	1
4. Питання оцінювання змістової валідності		1	1
Всього годин	1	4	5

Результат 1 модулю: слухач усвідомить переваги використання тестів для контролю знань студентів з вищої математики; з'ясує основні проблеми, пов'язані із процесом розроблення тестів з вищої математики, та дізнається про методи їх розв'язання; усвідомить роль оцінювання якості тестів з вищої математики при їхньому використанні для контролю знань студентів; отримає знання основних принципів розроблення завдань різних форм з вищої математики; навчиться створювати завдання з вищої математики різних типів у системі Moodle.

ЛІТЕРАТУРА

1. Орлов А.И. Экспертные оценки. Учебное пособие. / А. И. Орлов – М.: 2002. - 31 с.
2. Максимова О. А. Модель и технология комплексной экспертизы качества тестовых материалов для системы профессионального образования : дис. ... канд. пед. наук : 13.00.08. / О. А. Макстмова – М., 2010. – 191 с.
3. Сергієнко В. П. Комп'ютерні технології в тестуванні: навч. посіб. для студ. вищ. пед. навч. закл. / В. П. Сергієнко, М. П. Малезик, Т. В. Сіткарь.

– Луцьк: Волиньполіграф, 2012. – 295 с.

4. Чельшкова М. Б. Теория и практика конструирования педагогических тестов / М. Б. Чельшкова. – М.: Логос, 2002. – 431 с.

5. Дудко А. Ф. Оцінювання змістовної валідності тестів з вищої математики з використанням Google Docs додатку [Електронний ресурс] / О. О. Диховичний, А. Ф. Дудко // Інформаційні технології і засоби навчання. – Том 52 №2 (2016). – 2016.
<http://journal.iitta.gov.ua/index.php/itlt/article/view/1383/1032>

6. Дудко А. Ф. Концепція розробки комп'ютерних тестів з вищої математики / А.Ф.Дудко // Науковий часопис Національного педагогічного університету імені М. П. Драгоманова. Серія 5. Педагогічні науки: реалії та перспективи. – Випуск 50: збірник наукових праць. – Київ: Вид-во НПУ імені М. П. Драгоманова, 2014. – С. 44-48.

МОДУЛЬ 2. КОМПЕТЕНТНІСТЬ ВИКЛАДАЧА ЩОДО ОЦІНЮВАННЯ ЯКОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ

Тематичний модуль 2

Назви тем	Кількість навчальних годин		
	Семінари , практичні заняття, тренінги	Лекції	Всього
1. Поняття компетентності		1	1
2. Компетентнісний підхід у навчанні		1	1
3. Компоненти та рівні розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики		1	1
Всього годин		3	3

Результат 2 модулю: слухач ознайомиться з компетентнісним підходом у навчанні; зрозуміє поняття компетентності; отримає знання компонент та рівней розвитку компетентності викладачів щодо оцінювання якості тестів з вищої математики

ЛІТЕРАТУРА

1. Симонов П.В. Междисциплинарная концепция человека: лотребностно -информационный подход // Человек в системе наук. - М.: Наука, 1989.-С. 58-73.
2. Кудряшова Т.Г. Деятельностные способности как основа для формирования компетенций / Т.Г. Кудряшова // Компетентность. – № 6. – 2010. – с.5.
3. International Standard Classification of Education. ISCED 1997/ UNESCO [Electronic resource]. 1997. – Mode of access: www.uis.unesco.org/en/pub/pub. – date of access 2009.
4. European Commission. The European Qualifications Framework for Lifelong Learning (EQF). – Luxembourg: Office for official publications of the European Communities. – 2008. – 15 p.

МОДУЛЬ 3. СТАТИСТИЧНІ МЕТОДИ ОЦІНЮВАННЯ ЯКОСТІ ПЕДАГОГІЧНИХ ТЕСТІВ

Тематичний модуль 3

Назви тем	Кількість навчальних годин		
	Семінари, практичні заняття, тренінги	Лекції	Всього
1.Оцінювання якості тестів на основі методів класичної теорії тестів (КТТ)	1	1	2
2.Оцінювання якості тестів з використанням моделей Item Response Theory (IRT)	1	1	2
3.Оцінювання ефективності тестів методами IRT	1	1	2
Всього годин	3	3	6

Результат 3 модулю: слухач засвоїть базові теоретичні відомості про оцінювання якості тестів з вищої математики методами IRT та КТТ; навчиться робити висновки про якість тестів і тестових завдань, знайдених відповідно до методів IRT та КТТ.

ЛІТЕРАТУРА

1. Чельшкова М. Б. Теория и практика конструирования педагогических тестов / М. Б. Чельшкова. – М.: Логос, 2002. – 431 с.
2. Стоунс Э. Психопедагогика. Психологическая теория и практика обучения / Э. Стоунс. – М. : Педагогика, 1984. – 470 с.
3. Энциклопедия психологических тестов. Личность. Мотивация. Потребность / Под ред. А. Карелина. – М., 1997. – 312 с.
4. Rasch G. Probabilistic models for some intelligence and attainment tests / G. Rasch. – Chicago: The University of Chicago Press, 1981. – 199 p.
5. Linden W. Handbook of Modern Item Response Theory / W. Linden, R. Hambleton. – NY: Springer-Verlag, 1997. – 510 p.
6. Bock R.D. Estimating item parameters and latent ability when the responses are scored in two or more nominal categories. // Psychometrika, 37. – 1972. – P. 29-51.
7. Penfield R.D. A New Response Model for Multiple-Choice Items. / R.D. Penfield. – University of Miami & Jimmy de la Torre Rutgers, The State University of New Jersey. – 2008. – 39 p.
8. Samejima F. A new family of models for the multiple choice item. – Research Report No. 79-4. – University of Tennessee, Department of Psychology. – 1979.
9. Lord F.M. Maximum likelihood estimation of item response parameters when some responses are omitted / F. M. Lord // Psychometrika, 48. – P. 477-482.
10. Лісова Т. В. Моделі та методи сучасної теорії тестів / Т. В. Лісова. – Ніжин: Видавець ПП Лисенко М. М., 2012. – 112 с.
11. Дудко А. Ф. Застосування інформаційної функції для аналізу та підвищення ефективності тестів з вищої математики [Електронний ресурс] / О. О. Диховичний, А. Ф. Дудко // Інформаційні технології і засоби навчання.

МОДУЛЬ 4. СИСТЕМА АВТОМАТИЗОВАНОГО АНАЛІЗУ ЯКОСТІ ТЕСТОВИХ ЗАВДАНЬ (СААЯТЗ)

Тематичний модуль 4

Назви тем	Кількість навчальних годин		
	Семінари, практичні заняття, тренінги	Лекції	Всього
1. Функціональні можливості СААЯТЗ використання методів КТТ	1	1	2
2. Функціональні можливості СААЯТЗ використання методів IRT	1	1	2
3. База каліброваних завдань для комп'ютерних тестів з вищої математики	1	1	2
Всього годин	3	3	6

Результат 4 модулю: слухач навчиться оцінювати якість тестів та тестових завдань методами КТТ та IRT із використанням СААЯТЗ.

ЛІТЕРАТУРА

1. Дудко А. Ф. Автоматизована система аналізу результатів комп'ютерного тестування з вищої математики / О. О. Диховичний, А. Ф. Дудко // Наукові праці ДонНТУ. Серія: «Педагогіка, психологія і соціологія». – №2 (14). – Донецьк, 2013. – С. 103-110.

МОДУЛЬ 5. КОМП'ЮТЕРНО ОРІЄНТОВАНА МЕТОДИКА ОЦІНЮВАННЯ ЯКОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ

Тематичний модуль 5

Назви тем	Кількість навчальних годин		
	Семінари, практичні заняття, тренінги	Лекції	Всього

1. Базові положення методики		1	1
2. Місце методики в роботі викладача з тестами з вищої математики		1	1
3. Практичне застосування методики	1	1	1
Всього годин	1	3	3

Результат 5 модулю: слухач отримає знання базових положень методики; з'ясує місце методики в роботі викладача з тестами з вищої математики; навчиться застосовувати комп'ютерно орієнтовану методику оцінювання якості тестів з вищої математики до результатів тестування.

ЛІТЕРАТУРА

1. Дудко А. Ф. Комплексна методика аналізу якості тестів з вищої математики / О.О.Диховичний, А.Ф.Дудко // Науковий часопис НПУ імені М.П. Драгоманова. Серія №2. Комп'ютерно-орієнтовані системи навчання: Зб. наук. праць / Редрада. – К.: НПУ імені М.П. Драгоманова, 2015. – № 15 (22). – С.140-144.

2. Дудко А. Ф. Практичний аналіз якості тестів з вищої математики / О.О.Диховичний, А.Ф.Дудко // Науковий вісник Ужгородського національного університету: Серія «Педагогіка. Соціальна робота». – № 32. – 2014. – С. 73-78.

3. Дудко А. Ф. Схема роботи викладача з тестами для контролю знань студентів з вищої математики / І. В.Алексєєва, В. О. Гайдей, О. О. Диховичний, А. Ф. Дудко, Н. Р. Коновалова, Л. Б. Федорова // Математика в сучасному технічному університеті : матеріали V Міжнар. наук.-практ. конф. (Київ, 29—30 грудня 2016 р.). — Київ : НТУУ «КПІ», 2017. — С. 229-232.

Модулі включають відповідні теми занять. Оцінювання знань слухачів навчальної програми проводиться на комплексній основі, яка складається з таких компонентів:

- результатів виконання поточних завдань до кожного з модулів курсу;
- оцінки випускної роботи.

**Анкета для виявлення рівня ціннісно-мотиваційного компонента
компетентності викладачів щодо оцінювання якості тестів з вищої
математики**

Шановні колеги!

Просимо Вас відповісти на запитання анкети, основною метою якої є перевірка ціннісно-мотиваційного компоненту компетентності викладачів щодо оцінювання якості тестів з вищої математики

Інструкція: Анкета складається з 6-ти питань, що стосуються ціннісно-мотиваційного компоненту компетентності викладачів щодо оцінювання якості тестів з вищої математики. Висловіть Вашу позицію щодо поданих питань: позначте ті відповіді, з якими Ви погоджуєтесь (✓) Оскільки дані анкети є анонімними, просимо Вас надати відверті відповіді.

Дякуємо за увагу!

1. Я зацікавлений у застосуванні комп'ютерного тестування для оцінювання знань студентів з вищої математики.
☐ так;
☐ ні;
2. Я маю бажання підвищити об'єктивність контролю знань студентів з вищої математики.
☐ так;
☐ ні;
3. Для мене важливо, щоб тести, які я застосовую для оцінювання якості тестів з вищої математики, проходили апробацію та щоб їх якість було підтверджено статистично.
☐ так;
☐ ні;
4. Я маю потребу в ознайомленні з методами оцінювання якості тестів з вищої математики.
☐ так;
☐ ні;
5. Я маю потребу в ознайомленні з програмними засобами для оцінювання якості тестів з вищої математики.

☐ так;

☐ ні;

6. Я зацікавлений у підвищенні рівня своєї компетентності щодо оцінювання якості тестів з вищої математики

☐ так;

☐ ні;

**Тест для діагностування рівня когнітивного компонента
компетентності викладачів щодо оцінювання якості тестів з вищої
математики**

1. Дихотомічне завдання – це:

- a. завдання, яке має єдину правильну відповідь;
- b. завдання, яке має 2 рівня правильності відповіді: правильне та неправильне;
- c. завдання, яке має два правильних варіанта відповіді;
- d. завдання, яке має більше двох рівнів правильності відповіді.

2. Дихотомічними є завдання:

- a. множинного вибору з одним правильним варіантом відповіді;
- b. множинного вибору з декількома правильними варіантами відповіді;
- c. на відповідність;
- d. з короткою відповіддю.

3. Політомічними є завдання:

- a. множинного вибору з одним правильним варіантом відповіді;
- b. множинного вибору з декількома правильними варіантами відповіді;
- c. на відповідність;
- d. з короткою відповіддю.

4. Надійність тесту – це:

- a. характеристика тесту, що відображає відповідність результатів тестування тій меті, заради якої тест було створено;
- b. характеристика тесту, що дає змогу назвати тест, який точніше та з меншими витратами, ніж інші тести, вимірює навчальні досягнення студентів певного рівня підготовленості;
- c. характеристика тесту, що відображає стійкість тестових результатів до дії випадкових факторів;
- d. характеристика тесту, що дає змогу розділити студентів з різним рівнем підготовленості;

- e. характеристика тесту, що відображає ймовірність отримання відмінної оцінки іспитником з даним рівнем підготовленості;

5. Який з методів не дає змогу обчислити надійність тесту, до якого входять політомічні завдання?

- a. метод Кьюдера-Річардсона;
- b. метод паралельних форм;
- c. ретестовий метод;
- d. метод Кронбаха α .

6. Які з методів обчислення надійності тесту здійснюються на підставі одноразового тестування?

- a. ретестовий метод;
- b. метод розщеплення;
- c. метод Кьюдера-Річардсона;
- d. метод Кронбаха α .

7. Валідність тесту – це:

- a. характеристика тесту, що відображає відповідність результатів тестування тій меті, заради якої тест було створено;
- b. характеристика тесту, що дає змогу назвати тест, який точніше та з меншими витратами, ніж інші тести, вимірює знання студентів певного рівня підготовленості;
- c. характеристика тесту, що відображає стійкість тестових результатів до дії випадкових факторів;
- d. характеристика тесту, що дає змогу розділити студентів з різним рівнем підготовленості;
- e. характеристика тесту, що відображає ймовірність отримання відмінної оцінки іспитником з даним рівнем підготовленості.

8. Для правильно складеного тесту коефіцієнти надійності та валідності повинні:

- a. бути невід'ємними;
- b. лежати в інтервалі $(0; 0,3]$;

- c. бути не меншими, ніж 0,7;
- d. бути не меншими, ніж 0,3.

9. Елементами кореляційної матриці тесту є:

- a. кореляція між завданнями тесту та індивідуальними балами іспитників;
- b. кореляція між завданнями тесту;
- c. кореляція між балами іспитника та загальними балами по завданню;
- d. кореляція між балами іспитників.

10. Для правильно складеного тесту елементи кореляційної матриці повинні:

- a. бути невід'ємними;
- b. лежати в інтервалі $(0; 0,3]$;
- c. бути не меншими, ніж 0,7;
- d. бути не меншими, ніж 0,3.

11. Для оцінювання валідності лише дихотомічних завдань обчислюють:

- a. індекс розрізняльної здатності дихотомічних завдань D_j ;
- b. коефіцієнт точкової бісеріальної кореляції між балами по завданню завданням та індивідуальними балами іспитників;
- c. коефіцієнт кореляції Пірсона між балами по завданню та індивідуальними балами іспитників ρ_j ;
- d. бісеріальний коефіцієнт кореляції $\rho_{bis j}$.

12. Завдання добре функціонує, якщо значення індексу розрізняльної здатності дихотомічних завдань D_j :

- a. невід'ємне;
- b. лежати в інтервалі $(0; 0,3]$;
- c. не менше, ніж 0,4;
- d. не менше, ніж 0,7.

13. Які з методів оцінювання якості тестів відносять до методів IRT:

- a. оцінювання критеріальної валідності тестів;
- b. оцінювання надійності тестів;
- c. оцінювання латентних параметрів завдання;
- d. оцінювання ефективності тестів із використанням інформаційної функції.

14.Складність завдання в рамках IRT:

- a. вимірюється у логітах;
- b. вимірюється у відсотках;
- c. вимірюється в одній шкалі із параметрами іспитників;
- d. набувають значення з проміжку $[-1;1]$.

15.Яка IRT-модель відповідає політомічним завданням:

- a. Раша;
- b. Раша-Мастерса;
- c. Бірнбаума;
- d. Тіссена-Стейнберга.

16.Яка IRT-модель призначена для оцінювання дистракторів завдання з множинним вибором:

- a. Раша;
- b. Раша-Мастерса;
- c. Бірнбаума;
- d. Тіссена-Стейнберга.

17.Двопараметрична модель Бірнбаума є розвитком моделі Раша, у яку додатково введено параметр:

- a. вгадування;
- b. диференціюючої спроможності;
- c. підготовленості іспитника;
- d. складності дистракторів завдання.

18.В рамках IRT завдання залишають у тесті, якщо його диференціююча спроможність набуває:

- a. додатних значень;

- b. значень, що лежать в межах інтервалу $(0, 4; 1]$
- c. значень, що лежать в межах інтервалу $(0; 0,3]$
- d. значень, що лежать в межах інтервалу $[0,5; 2,5]$.

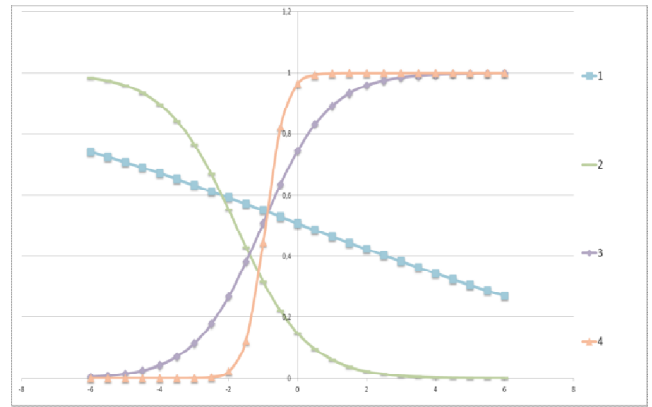


Рис. Д1.

19. Якій характеристичній кривій на рис. Д1 відповідає завдання з більшою диференціюючою спроможністю?

- a. 1;
- b. 2;
- c. 3;
- d. 4;

20. Характеристична крива завдання в рамках моделі Раша— це:

- a. графік залежності ймовірності правильної відповіді на дане завдання від параметра підготовленості іспитника;
- b. графік залежності ймовірності правильної відповіді даного іспитника від параметра складності завдання;
- c. графік залежності ймовірності правильної відповіді на дане завдання від параметра вгадування;
- d. графік залежності ймовірності правильної відповіді на дане завдання від параметра структурованості знань іспитника.

21. Інформаційна функція тесту дозволяє оцінити:

- a. надійність тесту;
- b. валідність тесту;
- c. ефективність тесту;
- d. диференціюючу спроможність тесту.

22. Який ансамбль характеристичних кривих рівнів завдань на рис. Д2 відповідає правильно складеному завданню відповідно до моделі Андерсена:

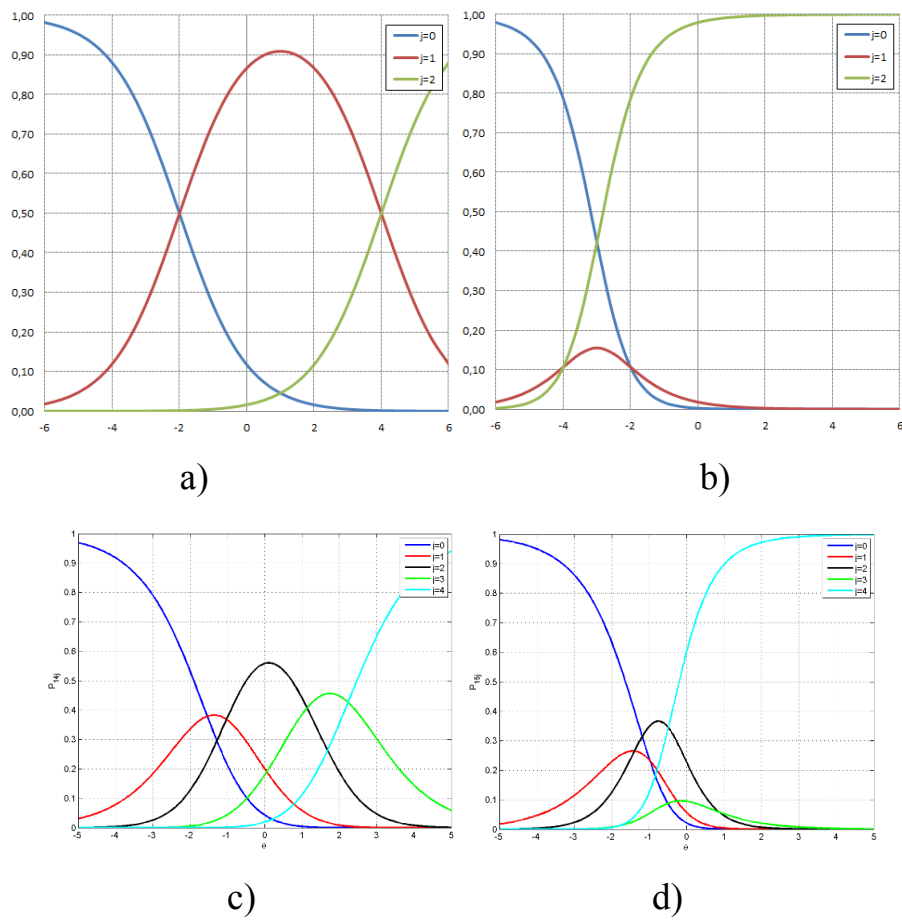


Рис. Д2

- a.
- b.
- c.
- d.

23. Графік інформаційної функції правильно складеного тесту

- a. має дзвоноподібну форму;
- b. є зростаючою функцією;
- c. має єдиний екстремум;
- d. не має екстремумів.

Індивідуальне завдання для діагностування рівня операційно-діяльнісного компонента компетентності викладачів щодо оцінювання якості тестів з вищої математики

I. За результатами тестування із застосуванням ПК...

1. ... обчисліть вибіркові характеристики індивідуальних балів іспитників.
2. ... здійсніть графічну інтерпретацію емпіричних даних у вигляді гістограм.
3. ... проаналізуйте розподіл вибірки індивідуальних балів.
4. ... обчисліть коефіцієнт надійності тесту на основі методу розщеплення.
5. ... обчисліть коефіцієнт надійності тесту Кронбаха α .
6. ... обчисліть коефіцієнт критеріальної валідності тесту.
7. ... обчисліть кореляційну матрицю завдань тесту.
8. ... обчисліть індекс розрізняльної здатності дихотомічних завдань D_j .
9. ... обчисліть коефіцієнт точкової бісеріальної кореляції між дихотомічним завданням та індивідуальними балами іспитників.
- 10.... обчисліть бісеріальний коефіцієнт кореляції між дихотомічним завданням та індивідуальними балами іспитників.
- 11.... обчисліть коефіцієнт кореляції Пірсона між політомічним завданням та індивідуальними балами іспитників ρ_j .
- 12.... оцініть латентні параметри тесту за моделлю Раша.
- 13.... оцініть латентні параметри тесту за моделлю Бірнбаума.
- 14.... оцініть латентні параметри тесту за моделлю Раша-Мастерса.
- 15.... оцініть латентні параметри тесту за моделлю Тіссена-Стейнберга.
- 16.... побудуйте ансамбль характеристичних кривих завдань за моделлю Раша.
- 17.... побудуйте характеристичні криві завдань за моделлю Бірнбаума.
- 18.... побудуйте ансамбль характеристичних кривих рівнів завдань за моделлю Раша-Мастерса.

- 19.... побудуйте ансамбль характеристичних кривих варіантів відповідей за моделлю Тіссена-стейнберга.
- 20.... побудуйте інформаційну функцію тесту.

II. Використовуючи результати I частини IЗ...

1. ... на основі обчислених характеристик зробіть висновок щодо валідності та надійності тесту.
2. ... на основі аналізу кореляційної матриці завдань та значень коефіцієнтів валідності завдань виявіть неправильно складені завдання.
3. ... проведіть аналіз ансамблю характеристичних кривих завдань за моделлю Раша та зробіть висновок щодо якості тесту та подальшої роботи з ним.
4. ... проведіть сумісний аналіз ансамблю характеристичних кривих рівнів завдань та інформаційної функції за моделлю Раша-Мастерса та зробіть висновок щодо якості завдань та подальшої роботи з ним.
5. ... проведіть аналіз ансамблю характеристичних кривих варіантів відповідей завдань за моделлю Тіссена-Стейнберга та зробіть висновок щодо якості завдань та подальшої роботи з ним.
6. ... на основі аналізу інформаційної функції тесту оцініть ефективність тесту та визначіть подальші дії.
7. ... на основі аналізу розподілу вибірки результатів, знайдених значень коефіцієнтів надійності та критеріальної валідності тесту, графіку інформаційної функції тесту оцініть якість тесту в цілому.
8. ... 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 виявіть неправильно складені завдання та визначіть подальші кроки.
9. ...зробіть висновки на основі використаних методів КТТ та IRT щодо подальшої роботи з тестом.

Анкета для визначення компетентності групи експертів

Інструкція: Шановні колеги, для заповнення анкети Вам необхідно обрати варіант відповіді та поставити позначку (✓) у відповідній графі.

1. Стаж викладацької діяльності:

- ☐ до 5 років;
- ☐ від 5 до 10 років;
- ☐ від 10 до 15 років;
- ☐ більше 15 років;

2. Науковий ступінь:

- ☐ доктор наук;
- ☐ кандидат наук;
- ☐ без ступеня.

3. Наукове звання:

- ☐ професор;
- ☐ доцент;
- ☐ ін. _____ (уточніть)

4. Посада:

- ☐ професор;
- ☐ доцент;
- ☐ старший викладач;
- ☐ асистент.

5. Який досвід проведення контролю знань з вищої математики Ви маєте?

- ☐ поточний контроль;
- ☐ підсумковий контроль (заліки та іспити);
- ☐ державна атестація;
- ☐ факультетський / ректорський контроль.

6. Чи маєте Ви досвід проведення наукових досліджень з питань контролю знань студентів з вищої математики?

- ☐ так;
- ☐ ні.

7. Чи маєте Ви публікації з питань контролю знань студентів з вищої математики?

- ☐ так;
- ☐ ні.

8. Чи маєте Ви досвід використання тестів для контролю знань студентів з вищої математики?

- ☐ так;
- ☐ ні.

9. Чи самостійно Ви розробляєте тести, які використовуєте для контролю знань студентів з вищої математики?

- ☐ так;
- ☐ ні.

10. Чи маєте Ви досвід проведення наукових досліджень з питань використання тестів для контролю знань студентів?

- ☐ так;
- ☐ ні.

11. Чи маєте Ви публікації з питань використання тестів для контролю знань студентів?

- ☐ так;
- ☐ ні.

Анкета для оцінювання якості контролю знань групою експертів

Інструкція: Для заповнення анкети Вам необхідно обрати варіант відповіді, що відповідає Вашій точці зору та поставити позначку (✓) у відповідній графі.

Оскільки дані анкети є анонімними, просимо Вас надати відверті відповіді.

Дякуємо за увагу!

№	Питання	Повністю погоджуюсь	Скоріш погоджуюсь, ніж не погоджуюсь	Скоріш не погоджуюсь, ніж погоджуюсь	Повністю не погоджуюсь
1.	Використання тестування дає змогу підвищити об'єктивність контролю знань студентів з вищої математики.				
2.	Використання комп'ютерного тестування спрощує процедуру розроблення робіт для контролю знань студентів.				
3.	Використання комп'ютерного тестування як засобу контролю знань з вищої математики спрощує процедуру перевірки робіт студентів.				
4.	Використання комп'ютерного тестування прискорює процес зворотнього зв'язку та покращує самоконтроль студентів.				
5.	Тестування дає змогу зменшити в цілому обсяг рутинних робіт викладача, тим самим, полегшуючи процес контролю знань студентів.				
6.	Тестування сприймається студентами краще, ніж традиційні засоби контролю знань.				
7.	Тестування для оцінювання навчальних досягнень доцільно використовувати у поєднанні з традиційними засобами контролю.				

8.	Використання тестів з вищої математики для оцінювання навчальних досягнень студентів вимагає оцінювання їх якості.				
9.	Для оцінювання якості з вищої математики доцільно використовувати засоби ІКТ.				
10.	Для оцінювання якості з вищої математики доцільно використовувати комп'ютерно орієнтовану методику оцінювання якості тестів.				
11.	Використання комп'ютерно орієнтованої методики оцінювання якості тестів дає змогу покращити якість розроблених тестів.				
12.	Використання комп'ютерно орієнтованої методики оцінювання якості тестів дає змогу забезпечити високі показники валідності, надійності та ефективності розроблених тестів.				
13.	СААЯТЗ є зручним засобом для оцінювання якості розроблених тестів.				
14.	Для контролю знань студентів з вищої математики доцільно використовувати тести, правильність складання яких було підтверджено з використанням комп'ютерно орієнтованої методики оцінювання якості тестів.				
15.	Результати тестів з вищої математики є узгодженими із результатами традиційних засобів контролю знань.				
16.	Впровадження комп'ютерно орієнтованої методики оцінювання якості тестів покращує контроль знань студентів з вищої математики				

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковано основні наукові результати

1. Дудко А. Ф., Диховичний О. О. Автоматизована система аналізу результатів комп'ютерного тестування з вищої математики. *Наукові праці ДонНТУ. Серія: «Педагогіка, психологія і соціологія»*. 2013. №2 (14). С. 103-110.
2. Дудко А. Ф., Диховичний О. О. Застосування інформаційної функції для аналізу та підвищення ефективності тестів з вищої математики [Електронний ресурс]. *Інформаційні технології і засоби навчання*. 2014. Том 41. №3. С. 55-69. URL: <http://journal.iitta.gov.ua/index.php/itlt/article/view/1040/790#.VFK7rzSsXm4> (включений до міжнар. наукометрич. баз).
3. Дудко А. Ф., Диховичний О. О. Практичний аналіз якості тестів з вищої математики. *Науковий вісник Ужгородського національного університету*: Серія «Педагогіка. Соціальна робота». 2014. № 32. С. 73-78.
4. Дудко А. Ф. Концепція розробки комп'ютерних тестів з вищої математики. *Науковий часопис Національного педагогічного університету імені М. П. Драгоманова*. Серія 5. Педагогічні науки: реалії та перспективи: зб. наук. пр. Київ: Вид-во НПУ імені М. П. Драгоманова. 2014. Випуск 50. С. 44-48.
5. Дудко А. Ф., Диховичний О. О. Комплексна методика аналізу якості тестів з вищої математики. *Науковий часопис НПУ імені М.П. Драгоманова*. Серія №2. Комп'ютерно-орієнтовані системи навчання: Зб. наук. праць. К.: НПУ імені М. П. Драгоманова. 2015. № 15 (22). С. 140-144.
6. Дудко А. Ф., Диховичний О. О. Оцінювання змістовної валідності тестів з вищої математики з використанням Google Docs додатку [Електронний ресурс]. *Інформаційні технології і засоби навчання*. 2016. Том 52 №2. С. 52-61 <http://journal.iitta.gov.ua/index.php/itlt/article/view/1383/1032> (включений до міжнар. наукометрич. баз).
7. Дудко А. Ф., Диховичний О. О. Аналіз поняттєво-категоріального апарату дослідження якості тестів з вищої математики. *Journal «ScienceRise: Pedagogical Education»*. 2017. № 5(13). С. 17-21. (включений до міжнар. наукометрич. баз)
8. Удовенко А.Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Аналіз якості тестових завдань для комплексу дистанційної освіти «Вища математика». *Теорія та методика електронного навчання*: зб. наук. пр. Кривий Ріг: Видавничий відділ НМетАУ. 2010. Вип. I. С. 3-9.
9. Удовенко А.Ф., Диховичний О. О., Шепель М. О. Політомічні моделі Мастерса та Андерсена в аналізі якості тестових завдань. *Теорія та методика електронного навчання*: збірник наукових праць. Кривий Ріг: Видавничий відділ НМетАУ. 2012. Випуск III. С. 83-87.

Наукові праці, які засвідчують апробацію матеріалів дисертації

10. Удовенко А.Ф. Сучасні математичні моделі педагогічного тестування. *«Історія розвитку науки, техніки та освіти»*: Зб. пр. VIII міжнародної молодіжної науково-практ. конф. за темою «Взаємодія академічної та університетської науки», 18 травня 2010 р. Київ. 2010. С. 202-206.

11. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Статистичний аналіз тестових завдань із застосуванням сучасних математичних моделей IRT. *Вісник Національного університету «Львівська політехніка» (серія «Інформатизація вищого навчального закладу»)*. 2012. №731. С. 10-13.
12. Дудко А.Ф., Диховичний О. О., Шепель М. О. Про особливості застосування політоміних математичних моделей тестових завдань. *Чотирнадцята міжнар. Конф. імені академіка М. Кравчука*, 19-21 квітня, 2012 р. Київ: Матеріали конф. Т.4. Історія та методика викладання математики. К.: НТУУ «КПІ». 2012. С. 91.
13. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про досвід застосування тестових контрольних робіт з вищої математики. *Міжнар. наук. конф. «Математика в сучасному технічному університеті»*, 19-20 квітня 2013 р., м. Київ. 2013. С. 442-445.
14. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про електронне навчання вищої математики в НТУУ «КПІ». *Міжнар. наук. конф. «Математика в сучасному технічному університеті»*, 19-20 квітня 2013 р., м. Київ. 2013. С. 438-441.
15. Дудко А. Ф., Диховичний О. О., Тарабара І. С. Статистичний аналіз якості тестових завдань для моделі з множинним вибором. *Третя міжнар. наук. конф. молодих вчених з математики та фізики*, 25-27 квітня, 2013 р., м. Київ. 2013. С. 169-170.
16. Дудко А. Ф., Диховичний О. О., Філіппова Г. В. Система автоматизованого аналізу якості тестових завдань. *Третя міжнар. наук. конф. молодих вчених з математики та фізики*, 25-27 квітня, 2013 р., м. Київ. 2013. С. 170-172.
17. Dudko A. F., Dykhovichnyi O. O. On the application of IRT-models for the statistical analysis of the quality of educational tests in higher mathematics. *Ukrainian-German Workshop on Empirical Complete Convergence and other Limit Theorems of Probability Theory*, September 23-27, 2013, Koktebel, Ukraine. 2013.
18. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про оцінку ефективності тестів з вищої математики. *Друга міжнародна наук. Конф. «Математика в сучасному технічному університеті»*, 20-21 грудня, 2013 р., м. Київ. 2013. С. 271-274.
19. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про електронне тестування з вищої математики в НТУУ «КПІ». *Вища освіта: проблеми і шляхи забезпечення якості. X Всеукр. науково-метод. конф.*, 28-29 листопада, 2013 р., м. Київ, НТУУ «КПІ». 2013. С. 7.
20. Дудко А. Ф., Диховичний О. О. Методика аналізу та підвищення ефективності тестів з вищої математики. *П'ятнадцята міжнар. конф. імені академіка М. Кравчука*, 15-17 травня, 2014 р. Київ: Матеріали конф. Т.4. Історія та методика математики. К.: НТУУ «КПІ». 2014. С. 96-97.
21. Дудко А. Ф., Диховичний О. О. Застосування IRT-моделей у дослідженні якості педагогічних тестів. *Всеукр. наукова конф. «Стохастичні та детерміновані математичні моделі у наукових дослідженнях»*, 27-29 червня 2014 р., м. Київ. 2014.

22. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Вибір підходу до розроблення комп'ютерних тестів з вищої математики. *Математика в сучасному технічному університеті: матеріали III Міжнар. наук.-практ. конф.* (Київ, 25–26 груд. 2014 р.). К.: НТУУ «КПІ». 2015. С. 317-318.

23. Дудко А. Ф., Диховичний О. О. Методика аналізу якості тестів та її застосування. *Четверта міжнар. наук. конф. молодих вчених з математики та фізики*, 23-25 квітня, 2015 р. – С. 78.

24. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Аналіз якості комп'ютерних тестів методами класичної теорії тестів. *Шістнадцята міжнар. конф. імені академіка М. Кравчука*, 14-15 травня, 2015 р. Київ: Матеріали конф. Т. 3. Теорія ймовірностей та математична статистика. Історія та методика математики. К.: НТУУ «КПІ». 2015. С. 94-97.

25. Дудко А. Ф., Диховичний О. О. Комп'ютерно-орієнтована методика статистичного аналізу тестів з вищої математики. *П'ята міжнар. наук. конф. молодих вчених з математики та фізики*, 25-26 квітня, 2016 р., м. Київ. 2016. – С. 71.

26. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Комп'ютерно-орієнтована методика аналізу якості педагогічних тестів як засіб ІК-підтримки викладачів вищої математики НТУУ «КПІ». *Математика в сучасному технічному університеті: матеріали IV Міжнар. наук.-практ. конф.* (Київ, 24–25 груд. 2015 р.). К.: НТУУ «КПІ». 2016. С. 277-278.

27. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Оцінювання якості змісту тестів з вищої математики з використанням Google Docs додатку. *Сімнадцята міжнар. конф. імені академіка М. Кравчука*, 19-20 травня, 2016 р. Київ: Матеріали конф. Т. 3. К.: НТУУ «КПІ». 2015. С. 176-179.

28. Dudko A., Dykhovychnyi A. Statistical analysis of tests in higher mathematics. *Limit theorem in probability theory, number theory and mathematical statistics. International workshop in honour of Prof. V.V. Buldygin* (October 10-12, 2016). NTUU «KPI» Kyiv, Ukraine. 2016. – P. 18.

29. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Схема роботи викладача з тестами для контролю знань студентів з вищої математики. *Математика в сучасному технічному університеті: матеріали V Міжнар. наук.-практ. конф.* (Київ, 29-30 грудня 2016 р.). Київ : НТУУ «КПІ». 2017. С. 229-232.

30. Дудко А. Ф., Диховичний О. О., Прохоренко Н. В. Застосування мови R у статистичному аналізі якості тестів з вищої математики. *Матеріали Вісімнадцятої міжнар. наук. конф. імені академіка Михайла Кравчука*, 7-10 жовтня 2017 року, Київ: Т. 2. Київ: НТУУ «КПІ». 2017. С. 222-224.

31. Дудко А. Ф., Диховичний О. О., Прохоренко Н. В. Моделювання даних результатів тестування з вищої математики за допомогою мови R. *Матеріали Вісімнадцятої міжнар. наук. конф. імені академіка Михайла Кравчука*, 7-10 жовтня 2017 року, Київ: Т. 2. Київ: НТУУ «КПІ». 2017. С. 225-226.

32. Дудко А. Ф., Диховичний О. О., Алексєєва І. В., Гайдей В. О., Коновалова Н. Р., Федорова Л. Б. Про необхідність розвитку компетентності щодо оцінювання якості тесту викладача вищої математики. *Матеріали Вісімнадцятої міжнар. наук. конф. імені академіка Михайла Кравчука*, 7-10 жовтня 2017 року, Київ: Т. 2. Київ: НТУУ «КПІ». 2017. С. 166-167.

33. Дудко А. Ф., Диховичний О. О., Круглова Н. В., Ординська З. П. Застосування моделі Бірнбаума до визначення ставлення викладачів вищої математики до тестового контролю знань. *Сьома всеукр. наук. конф. студентів, аспірантів та молодих вчених з математики*, 18-20 квітня, 2018 р., м. Київ. 2018. С. 73.

ВІДОМОСТІ ПРО АПРОБАЦІЮ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЇ

Міжнародні конференції – 14, а саме:

– VIII Міжнародна молодіжна науково-практична конференція «Історія розвитку науки, техніки та освіти» за темою «Взаємодія академічної та університетської науки» (Київ, 2010);

– Чотирнадцята міжнародна конференція імені академіка М. Кравчука (Київ, 2012);

– Міжнародна наукова конференція «Математика в сучасному технічному університеті» (Київ, 2013);

– Друга міжнародна наукова конференція «Математика в сучасному технічному університеті» (Київ, 2013);

– Третя міжнародна наукова конференція молодих вчених з математики та фізики (Київ, 2013)

– П'ятнадцята міжнародна конференція імені академіка М. Кравчука (Київ, 2014)

– Математика в сучасному технічному університеті: матеріали III Міжнар. наук.-практ. конф. (Київ, 2014)

– Четверта міжнародна наукова конференція молодих вчених з математики та фізики (Київ, 2015)

– Шістнадцята міжнародна конференція імені академіка М. Кравчука (Київ, 2015)

– П'ята міжнародна наукова конференція молодих вчених з математики та фізики (Київ, 2016)

– Математика в сучасному технічному університеті: IV Міжнар. наук.-практ. конф. (Київ, 2016)

– Сімнадцята міжнародна конференція імені академіка М. Кравчука (Київ, 2016)

– Математика в сучасному технічному університеті : V Міжнар. наук.-практ. конф. (Київ, 2017)

– Вісімнадцята міжнародна наукова конференція імені академіка Михайла Кравчука (Київ, 2017)

Всеукраїнські конференції – 3, а саме:

- Вища освіта: проблеми і шляхи забезпечення якості. X Всеукраїнська науково-методична конференція (Київ, 2013)
- Всеукраїнська наукова конференція «Стохастичні та детерміновані математичні моделі у наукових дослідженнях» (Київ, 2014)
- Сьома всеукраїнська наукова конференція студентів, аспірантів та молодих вчених з математики (Київ, 2018)

Воркшопи – 2, а саме:

- Ukrainian-German Workshop on Empirical Complete Convergence and other Limit Theorems of Probability Theory (Koktebel, 2013);
- Limit theorem in probability theory, number theory and mathematical statistics. International workshop in honour of Prof. V.V. Buldygin (Kyiv, 2016).

ДОВІДКИ ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЇ (Скановані копії)



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ПЕДАГОГІЧНИЙ УНІВЕРСИТЕТ

імені М.П. ДРАГОМАНОВА
01601, м.Київ-30, вул. Пирогова, 9
Телефон 234-11-08

E-mail: shef@ukr.net, код ЄДРПОУ 02125295

26.06.18 № 07-10/1035

ДОВІДКА

Про впровадження результатів дослідження
Дудко Анни Федорівни
у навчальний процес Національного педагогічного університету
імені М.П. Драгоманова

Результати дисертаційного дослідження А.Ф. Дудко було частково впроваджено протягом 2017-2018 рр. у навчальний процес Національного педагогічного університету імені М.П. Драгоманова. Впровадження було здійснено викладачами вищої математики студентів 1 та 2 курсів спеціальностей «111 Математика», «014 Середня освіта (Математика)».

Застосування запропонованої А.Ф. Дудко комп'ютерно орієнтованої методики оцінювання якості тестів дозволило розробити комплект ефективних, надійних та валідних тестів, які було застосовано для контролю знань студентів з вищої математики.

В результаті використання розробленої А.Ф. Дудко системи автоматизованого аналізу якості тестових завдань (СААЯТЗ) викладачами кафедри вищої математики було сформовано базу каліброваних завдань з вищої математики, яку у подальшому буде використано для модифікації комплексу розроблених тестів.

Результати експериментального впровадження розробленої А.Ф. Дудко комп'ютерно орієнтованої методики оцінювання якості тестів дозволяють констатувати підвищення ефективності роботи викладачів щодо розробки та оцінювання якості тестів з вищої математики та покращення контролю знань студентів з вищої математики у цілому.

Декан фізико-математичного факультету
НПУ імені М. П. Драгоманова
д. ф.-м. н., професор

Авба

Працьовитий М.В.

Проректор з наукової роботи
НПУ імені М. П. Драгоманова
д. ф.-м. н., професор



[Signature]

Торбін Г.М.



УКРАЇНА

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

03056, м. Київ, пр-т Перемоги, 37; тел. (+38 044) 204-82-82 тел./факс (+38 044) 204-97-88
<http://www.kpi.ua> e-mail: mail@kpi.ua ЄДРПОУ 02070921

10.07.2018 № 90/07-18
на № _____ від _____

ДОВІДКА

Про впровадження результатів дослідження *Дудко Анни Федорівни*
у Навчально-методичному комплексі «Інститут післядипломної освіти»
Національного технічного університету України
«Київського політехнічного інституту імені Ігоря Сікорського»

Починаючи з 2017/2018 н.р. результати дисертаційного дослідження А.Ф. Дудко впроваджуються у Навчально-методичному комплексі «Інститут післядипломної освіти» Національного технічного університету України «Київського політехнічного інституту імені Ігоря Сікорського» у процесі підвищення кваліфікації науково-педагогічних працівників вищих навчальних закладів за допомогою програми підвищення кваліфікації та відповідного до неї дистанційного курсу «Формування компетентності щодо оцінювання якості тестів з вищої математики», який включає електронний підручник, індивідуальні практичні завдання та тести.

В результаті експериментального навчання викладачів за розробленою А. Ф. Дудко методикою формування компетентності щодо оцінювання якості тестів з вищої математики із використанням вищезазначених програми та дистанційного курсу викладачі на професійному рівні здатні самостійно розробляти, оцінювати та вдосконалювати тести з вищої математики, забезпечуючи їхню високу валідність та надійність. Застосування методики формування компетентності щодо оцінювання якості тестів з вищої математики сприяє розвитку рівнів сформованості відповідної компетентності.

Результати впровадження розробленої А.Ф. Дудко методики формування компетентності щодо оцінювання якості тестів з вищої

математики дозволяють констатувати її ефективність та перспективність у процесі підвищення кваліфікації науково-педагогічних працівників вищих навчальних закладів.

Директор НМК ІПО
КПІ ім. Ігоря Сікорського



І. Г. Малюкова

004940



УКРАЇНА
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТРАНСПОРТНИЙ УНІВЕРСИТЕТ

вул.М.Омеляновича-Павленка, 1, м. Київ, 01010 т.ф. +38 (044) 280 82 03, т. +38 (044) 280 87 65
e-mail: general@ntu.edu.ua, код ЄДРПОУ 02070915

25.06.2018 № 1524/01

на № _____

ДОВІДКА

про впровадження комп'ютерно орієнтованої
методики оцінювання якості тестів,
розробленої Дудко Анною Федорівною,
у навчальний процес Національного транспортного університету

Національний транспортний університет протягом 2017-2018 рр. здійснював упровадження розробленої А.Ф. Дудко комп'ютерно орієнтованої методики оцінювання якості тестів під час викладання дисциплін «Дискретна математика» та «Теорія алгоритмів» для студентів денної форми навчання за спеціальністю «Комп'ютерні науки».

Застосування запропонованої комп'ютерно орієнтованої методики оцінювання якості тестів дозволило розробити та застосувати нові ефективні засоби контролю знань з вищенаведених дисциплін.

Використання комп'ютерно орієнтованої методики оцінювання якості тестів сприяло підвищенню об'єктивності проведення модульного контролю, заліків та іспитів.

Результати впровадження розробленої А.Ф. Дудко комп'ютерно орієнтованої методики оцінювання якості тестів дозволяють констатувати її ефективність щодо покращення контролю знань студентів.

Проректор з навчальної роботи
Національного транспортного
університету, професор



О.К. Гришук