

Бахрушин Володимир Євгенович  
професор, доктор фізико-математичних наук  
Класичний приватний університет, м. Запоріжжя, Україна  
Vladimir.Bakhrushin@zhu.edu.ua

## **Статистичний аналіз ЗНО 2009 – 2011**

### **Анотація**

*Проаналізовано показники якості тестів ЗНО останніх років за допомогою статистичних методів. Показано, що для значної кількості тестових завдань коефіцієнти їх кореляції й дискримінації, а також показники складності не відповідають вимогам до освітніх тестів. Отримані результати свідчать, що в існуючому вигляді ЗНО не є якісним інструментом оцінювання якості освіти й потребує суттєвого доопрацювання. Зокрема пропонується накопичення бази даних якісних тестових завдань, а також запровадження різномісних тестів, призначених для задоволення різних вимог університетів і спеціальностей до абітурієнтів та рівня їх підготовки з дисциплін тестування.*

**Ключові слова:** ЗНО, тестове завдання, якість, статистичний аналіз

Останнім часом тестові технології все ширше використовуються на різних рівнях освіти. З 2008 р. уведено Зовнішнє незалежне оцінювання (ЗНО) абітурієнтів вищих навчальних закладів як основний інструмент відбору студентів на навчання. Це висуває підвищені вимоги до якості відповідних тестів, але на сьогодні практично немає незалежних оцінок їх відповідності загальним вимогам теорії освітнього тестування. Втім є багато публікацій, що свідчать про наявність окремих недоліків і проблем. З іншого боку особливістю тестів ЗНО є великі обсяги випробуваних, що дає змогу обґрунтовано використовувати статистичні методи для аналізу їх результатів й опрацювати відповідні методики досліджень.

При аналізі освітніх тестів необхідно враховувати як якісні показники тесту у цілому, які насамперед показують відповідність його результатів меті тестування, так і показники якості окремих тестових завдань.

Метою роботи є статистичний аналіз показників якості завдань тестів ЗНО 2009 – 2011 р. Попередні результати цього дослідження наведено у працях: Bakhrushin and Gorban (2011), Бахрушин и Горбань (2011), Бахрушин, Горбань и Игнахина (2011).

Основи теорії й результати досліджень окремих аспектів освітнього тестування докладно викладено у працях: Авансов (2002), Ащепкова (2003), Кім (2007),

Челишкова (2002), а також у багатьох публікаціях в журналах Вісник "Тестування і моніторинг в освіті", "Інформаційні технології в освіті", "Інформаційні технології і засоби навчання", "Педагогические измерения", Journal of Educational Measurement та інших.

В літературі використовують різні визначення педагогічних тестів. Зокрема Авансов (2005, с. 19) дає таке визначення: "Педагогічний тест – це система паралельних завдань зростаючої складності, специфічної форми, що дає змогу якісно й ефективно оцінювати рівень та структуру підготовленості учнів". У роботі Кім (2007, с. 19) наведено дещо інше визначення: "Педагогічний тест це система тестових завдань різної складності, яка дає змогу якісно й ефективно вимірювати рівень та структуру підготовленості випробуваних".

Аналіз цих та інших визначень дає змогу розглядати освітні тести як певні моделі знань, поєднані із засобами вимірювання відповідності випробуваних цим моделям. Втім відомо (Бахрушин, 2009, с. 9-10), що будь-яка модель характеризується такими властивостями як: цільове призначення, наближеність, неточність, неповнота тощо. Стосовно тестування це означає, що форма та зміст освітнього тесту повинні залежати від мети тестування. При цьому його результати навіть для якісного тесту є правильними для групи в цілому, але для кожного конкретного випробуваного можуть містити істотну статистичну похибку, що визначається не знаннями й вміннями особи, а випадковими статистичними чинниками. Для мінімізації похибок тестування тест має розроблятися згідно з певними правилами та процедурами, які зокрема передбачають перевірку якості як самого тесту, так і його окремих завдань.

Якщо тестування використовується для відбору студентів на навчання, то модель тесту має відповідати моделі успішного студента, яка у свою чергу визначається моделлю успішного випускника. Сьогодні для різних університетів та різних спеціальностей використовують три основні моделі успішного студента (Bakhrushin and Gorban, 2011; Бахрушин, Горбань и Игнахина, 2011).

Згідно з першою, успішним є той студент, що може запам'ятовувати надані йому знання й використовувати їх для вирішення типових професійних завдань. Звідси впливає модель успішного абітурієнта, що покладена до основи сучасного ЗНО: успішним є той абітурієнт, що за обмежений час може надати правильні відповіді на велику кількість простих питань або (для математики й фізики)

правильно розв'язати велику кількість простих задач. При цьому вимога великої кількості завдань, а відповідно й їх простоти впливає з вимоги якомога повнішого охоплення змісту шкільної програми. Такий підхід в цілому задовольняє потреби технічних і деяких гуманітарних спеціальностей у більшості ВНЗ. Але є принаймні два виключення.

По-перше, університетська підготовка з технічних чи, наприклад, філологічних спеціальностей базується на 11-річній шкільній математичній та мовній підготовці. Але існує велика кількість напрямів підготовки (правознавство, психологія, соціальна робота, менеджмент, туризм, готельне господарство тощо), де успішне навчання не використовує такою мірою шкільні знання. Тому успішний абітурієнт для відповідних спеціальностей – це не той, хто добре опанував шкільну програму, а той, хто має здатності для подальшого навчання. Саме тому останнім часом керівництво освіти все частіше згадує про необхідність розробки й запровадження відповідних тестів.

По-друге, існує обмежена кількість університетів (КНУ, НТУУ "КПІ" й деякі інші ВНЗ, що готують переважно науковців, педагогів вищої школи, фахівців нових наукоємних виробництв) для яких успішність студента має полягати не в умінні швидко відповідати на прості питання, а в його здатності вирішувати складні завдання. А це висуває зовсім інші вимоги до якостей абітурієнтів. Такі тести також можуть бути розроблені на загальнодержавному рівні, але раціональнішим є їх запровадження на рівні зацікавлених університетів, або надання їм права на проведення власних вступних іспитів.

Далі розглянемо основні результати статистичного аналізу показників якості завдань тестів ЗНО 2009 – 2011 р. Вихідні дані взято із офіційних звітів (УЦОЯО, 2009; УЦОЯО, 2010; УЦОЯО, 2011), опублікованих на сайті Українського центру оцінювання якості освіти (<http://www.testportal.gov.ua>). Особливостями цих тестів є: наявність офіційної статистики; великі обсяги вибірок, що підвищує достовірність статистичних показників; а також наявність спеціалізованого центру тестування, що мало б забезпечити дотримання науково обґрунтованої процедури розробки тестів.

Загальна характеристика якості тестів ЗНО 2009 – 2011 наведена в публікаціях Bakhrushin and Gorban (2011), Бахрушин и Горбань (2011), Бахрушин, Горбань и Игнахина (2011). Відповідно до наведених у цих працях результатів, за офіційними

даними важно охарактеризувати якість тестів у цілому. Це пов'язано з тим, що валідність тестів не визначається з міркувань забезпечення таємності тестових завдань. З тієї самої причини надійність тестів визначається за допомогою коефіцієнта  $\alpha$ -Кронбаха, який не передбачає необхідності повторення тестових випробувань. Його значення знаходяться в межах 0,80 – 0,97, що у цілому вважається задовільним показником. Але згідно з результатами проведеного нами моделювання, значення  $\alpha$ -Кронбаха може бути досить високим (0,95 – 0,98) навіть при високому рівні випадковості відповідей. Тому наявні дані не є достатніми для висновку про високу надійність аналізованих тестів без врахування інших показників.

За відсутності повторних випробувань необхідну інформацію може надати аналіз показників якості окремих тестових завдань. Для тестів ЗНО в офіційних звітах наводяться дані про коефіцієнти кореляції й дискримінації завдань, а також про показники їх складності. Розглянемо ці дані більш докладно.

– рівень складності ( $p_i$ ), який визначають як середнє арифметичне балів, отриманих всіма учнями за відповідне завдання), і розраховують за формулою:

$$p_i = \frac{\sum_{j=1}^k Y_{ij} k_{ij}}{k Y_{i\max}}, \quad (1)$$

де  $k$  – кількість випробуваних,  $Y_{ij}$  – кількість балів за  $j$ -ого випробуваного на  $i$ -те завдання,  $k_{ij}$  – кількість випробуваних, що вибрали  $j$ -ту відповідь на  $i$ -те завдання (у найпоширенішому випадку, коли оцінки за відповіді можуть набувати лише значень 0 та 1, він дорівнює середньо арифметичному всіх балів за відповідне завдання).

На рис. 1 наведено графіки емпіричних функцій розподілу показників складності завдань для тестів ЗНО-2011. З наведених даних зокрема бачимо, що для тестів з математики частка занадто складних завдань (показник складності менше 20) дорівнює приблизно 0,23 (тобто 23 %), частка оптимальних за складністю (значення показника складності від 40 до 60) – близько 20 %.

Частка дуже легких завдань в тестах ЗНО 2010 – 2011 коливається в межах від 0 до 10 %, а частка дуже складних – від 0 до 33 %. При цьому, дуже легких завдань немає у обох тестах з фізики, географії, англійської та німецької мов, а найбільшою їх

частка є в тестах 2010 р. з української мови та літератури (1 сесія) та іспанської мови, а також тесті 2011 р. з російської мови. Дуже складних завдань немає в тесті 2010 р. з іспанської мови, а також в тестах 2011 р. з української мови та літератури (2 сесія) й іспанської мови. Найбільшою їх частка є у всіх тестах з фізики та математики.

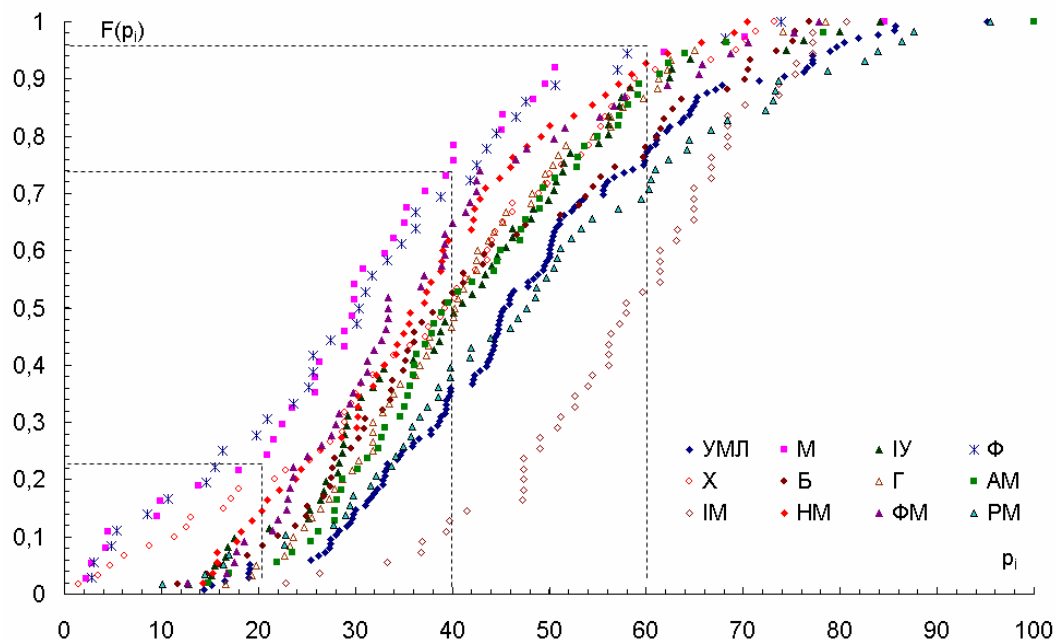


Рис. 1. Емпіричні функції розподілу показників складності завдань тестів ЗНО-2011.

Згідно з вимогами методики розробки тестів занадто складні й занадто легкі завдання повинні вилучатися з тестів, але порогове значення рівня складності залежить від мети тестування. Зокрема для відбіркових тестів використовують завдання з високим рівнем складності, а для відсіювальних – з низькою (Казиев, 2008).

Особливо актуальною вимога відбраковування занадто складних завдань є для завдань з вибором відповіді. Але для ЗНО цього не роблять, незважаючи на те, що навіть за офіційними оцінками частка неякісних за цим показником завдань для окремих дисциплін перевищує чверть всіх завдань. Частка оптимальних за складністю завдань не перевищує 40 % (хімія, географія). Найнижчою вона є для тесту з математики.

Порівняння даних для повних масивів значень показників складності для кожного року показує, що загалом відмінність між відповідними функціями розподілу є несуттєвою. Середні значення знаходяться у межах 42 – 44, а стандартні відхилення

коливаються від 17,7 до 19,6. Частки занадто складних і занадто легких завдань дорівнюють приблизно 10 й 1 %, а частки оптимальних за складністю завдань дорівнюють 30 – 35 %.

Втім наведені в офіційних звітах дані щодо показників складності занижують істинну складність тестових завдань. Це пов'язано з тим, що при їх розрахунку УЦОЯО не враховує імовірність вгадування правильних відповідей. Аналіз розподілу обрання варіантів відповідей (Bakhrushin and Gorban, 2011; Бахрушин и Горбань, 2011; Бахрушин, Горбань и Игнахина, 2011) вказує, що в окремих випадках це завищує значення показника складності на 15 – 25 %. Тобто насправді частка оптимальних за складністю завдань є значно меншою, ніж визначено в звітах УЦОЯО.

Як приклад наведемо розподіл відповідей для завдання 43 тесту ЗНО-2011 з української мови та літератури – "Польовою царівною" названо: А) Лесю Череванівну (Чорна рада); Б) Мотрю Довбиш (Кайдашів сім'я); В) Галю Гудзь (Хиба ревуть воли, як ясна повні); Г) Олену Ляуфлер (Людина); Д) Марічку Гутенюк (Тіні забутих предків). Частоти вибору варіантів відповідей учасниками тестування дорівнюють, відповідно: 16,50; 11,66; 29,10; 12,76 та 29,22. Правильною є відповідь В, тобто показник складності дорівнює 29,10. Але насправді бачимо, що приблизно 13 – 14 % правильних відповідей є результатом випадкового вибору між усіма варіантами, а інші – вибором між варіантами С й Д. Тобто істинне значення показника складності є близьким до нуля.

Іншим прикладом є завдання 36 тесту з історії України ЗНО 2011 р: Прочитайте уривок з історичного джерела та виконайте завдання. "Експропріація непродуктивних класів, анулювання боргів, зосередження доходів від промисловості, держторгівлі (внутрішньої та зовнішньої) і всієї кредитної системи в руках держави і т. ін. – самі по себе дають можливість такого нагромадження всередині країни, яке забезпечує необхідний для соціалістичного будівництва темп розвитку індустрії". В уривку вказано на: А) причини індустріалізації; Б) джерела індустріалізації (правильна відповідь); В) наслідки індустріалізації; Г) етапи індустріалізації. Розподіл відповідей для цього завдання є таким: 22,30; 27,02; 28,88; 21,63. Тобто в межах випадкового розсіювання імовірність вибору відповіді коливається поблизу 1/4, тобто відповідає випадковому вибору з чотирьох рівноімовірних варіантів. Це свідчить про те, що

істинний показник складності цього завдання є близьким до нуля, а не 27,02, як визначено в офіційному звіті УЦОЯО. Можна також зазначити, що це питання перевіряє скоріше не знання історії України, а розуміння того, що таке причини, наслідки, джерела й етапи, тобто загальний рівень культури випробуваних. Нажаль відповіді на нього свідчать про те, що цей рівень є досить низьким.

Ще однією проблемою тестів ЗНО є те, що рівень складності завдань не впливає на бали, що надаються за їх правильне виконання. На нашу думку існує необхідність доопрацювання методики обробки результатів тестування, й зокрема передбачити корегування одержуваних балів за правильні відповіді з огляду на отримані після його проведення оцінки складності завдань. Деякі підходи до вирішення цієї проблеми розглянуто в працях: Оганесян, Дещинский и Бирюлев, 2010; Бахрушин, Журавель и Игнахина, 2010.

Коефіцієнт кореляції ( $R_i$ ) тестового завдання показує, наскільки корелюють між собою успішність його виконання й підсумкові бали випробуваних за весь тест. У загальному випадку цей коефіцієнт визначають за формулою (Крисілов, Онищенко и Русинова, 2004):

$$R_i = \frac{\sum_{j=1}^n Y_{ij}x_j - \bar{Y}_i \bar{x}}{S_x S_Y} \frac{k}{k-1}; \quad (2)$$

де  $Y_{ij}$  – кількість балів  $j$ -го випробуваного за  $i$ -те завдання,  $x_j$  – загальна кількість балів  $j$ -го випробуваного,  $\bar{Y}_i$  – середня кількість балів за  $i$ -те завдання,  $\bar{x}$  – середня кількість балів за весь тест,  $S_x$ ,  $S_Y$  – стандартні відхилення відповідних показників.

Завдання з низькими коефіцієнтами кореляції рекомендується вилучати з тестів. Різні автори пропонують брати порогове значення у межах 0,15 – 0,5 (Аванесов, 2002; Аванесов, 2005; Ащепкова, 2003; Ким, 2007; Чельшкова, 2002). На нашу думку мінімальне порогове значення для коефіцієнта кореляції 0,3 є найбільш обґрунтованим, оскільки відповідає загальноприйнятому у прикладній статистиці порогу значущості коефіцієнта кореляції для типових значень обсягів вибірок (Бахрушин, 2009). Нижчі порогові значення зумовлені не стільки можливою значущістю відповідних завдань, скільки складністю підготовки якісних завдань.

Іноді пропонують не вилучати лише завдання з від'ємними коефіцієнтами кореляції, а інші завдання з низькими значеннями коефіцієнта кореляції корегувати. Але і у цьому випадку такі завдання не потрапляють до підсумкового варіанта тесту.

На рис. 2 наведено емпіричні функції розподілу коефіцієнтів кореляції завдань ЗНО-2011. З наведених даних бачимо, що частка завдань, для яких коефіцієнт кореляції є меншим ніж 0,5, для всіх тестів є дуже великою. Найменшою вона є для тесту з іспанської мови й дорівнює приблизно 1/3 усього масиву завдань. Для інших тестів вона перевищує 1/2, а для деяких (фізика, історія України, географія, біологія) наближається до 100 %.

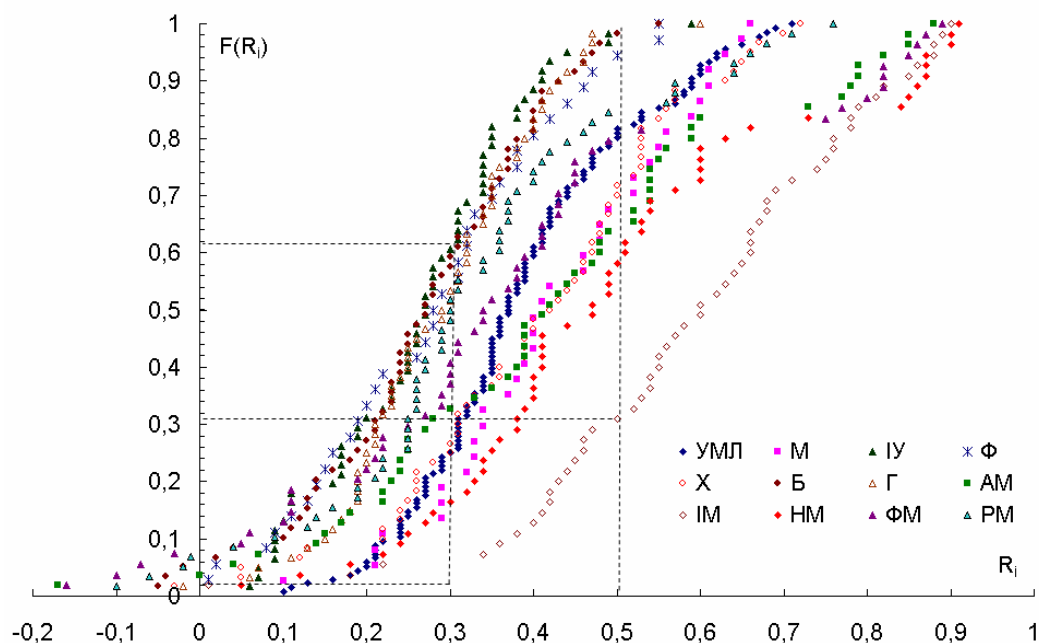


Рис. 2. Емпіричні функції розподілу коефіцієнтів кореляції тестових завдань

Навіть, якщо за граничне значення взяти 0,3 (тобто врахувати завдання, що потребують корегування), частка незадовільних за цим показником завдань знаходиться в межах від 5 % (іспанська мова) до 62 % (біологія й історія України). Аналогічними є результати для ЗНО 2009, 2010 р.

Порівняння даних для повних масивів значень кореляції для кожного року показує, що загалом відмінність між відповідними функціями розподілу є більш помітною, ніж у випадку показників складності. Але вона залишається не досить суттєвою. Середні значення знаходяться у межах 0,34 – 0,39, а стандартні відхилення коливаються від 0,17 до 0,36. Частки завдань з коефіцієнтом кореляції менше 0,5



варіюються у межах 76 – 83 %, а частки завдань з коефіцієнтом кореляції менше 0,3 – у межах 34 – 40 %, тобто є дуже високими.

Наведені дані дають вагомі підстави для сумнівів в адекватності результатів ЗНО 2009 – 2011 р. з багатьох дисциплін, бо важко зрозуміти, що саме характеризує підсумковий бал, який для деяких тестів не корелює з результатами виконання переважної більшості його окремих завдань. Наприклад, для згаданих вище завдання 43 тесту з української мови та літератури коефіцієнт кореляції дорівнює 0,24, а для завдання 36 з історії України – 0,23. Тобто в межах статистичної похибки ці величини не відрізняються від нуля. Але чого ще можна очікувати, якщо частки тих, хто впорався з цими завданнями завдяки своїм знанням, а не вгадуванню, насправді є близькими до нуля.

Основними причинами низьких значень коефіцієнтів кореляції, на нашу думку є: низький загальний рівень знань з відповідної дисципліни; наявність завдань, що потребують відповіді на другорядні питання, які не є принциповими з погляду оцінювання рівня підготовки абітурієнта; а також наявність окремих некоректних завдань, які не мають правильних відповідей або мають декілька правильних відповідей.

Зокрема для тестів з географії основною є перша причина. Рівень знань абітурієнтів можуть охарактеризувати такі дані: понад 42 % випробуваних вважають, що земна вісь проходить через екватор; близько 45 % вважають що навколосвітню подорож можна здійснити дрейфуючи Гольфстрімом; майже 50 % відповіли, що розвиток кольорової металургії у Запоріжжі зумовлений наявністю місцевої сировини. При цьому, на відміну від української мови та літератури, математики, історії України, англійської мови, тест з географії складали переважно абітурієнти, що обрали спеціальності, пов'язані з поглибленим знанням географії.

Коефіцієнт (індекс) дискримінації тестового завдання вказує на спроможність за результатами його виконання вирізнити групи учнів, що успішно й невдало виконали тест у цілому (найчастіше для його розрахунку беруть 1/3 кращих та 1/3 гірших за підсумковими результатами учнів). Очевидно, що за своєю сутністю цей показник є спрощеною формою коефіцієнта кореляції.

Для перевірки цього припущення нами було перевірено наявність кореляції між значеннями цих показників для повного масиву тестових завдань ЗНО 2009 – 2011.

Згідно з отриманими даними (рис. 3), коефіцієнт детермінації для лінійної моделі зв'язку дорівнює приблизно 0,73, що відповідає наявності сильної кореляції між цими показниками. Тому, на нашу думку, одночасний розрахунок коефіцієнтів кореляції й дискримінації для тестових завдань є зайвим. Це підтверджується результатами аналізу емпіричних функцій розподілу для значень коефіцієнтів дискримінації тестових завдань, які загалом є подібними для аналогічних функцій значень коефіцієнтів кореляції (рис. 4).

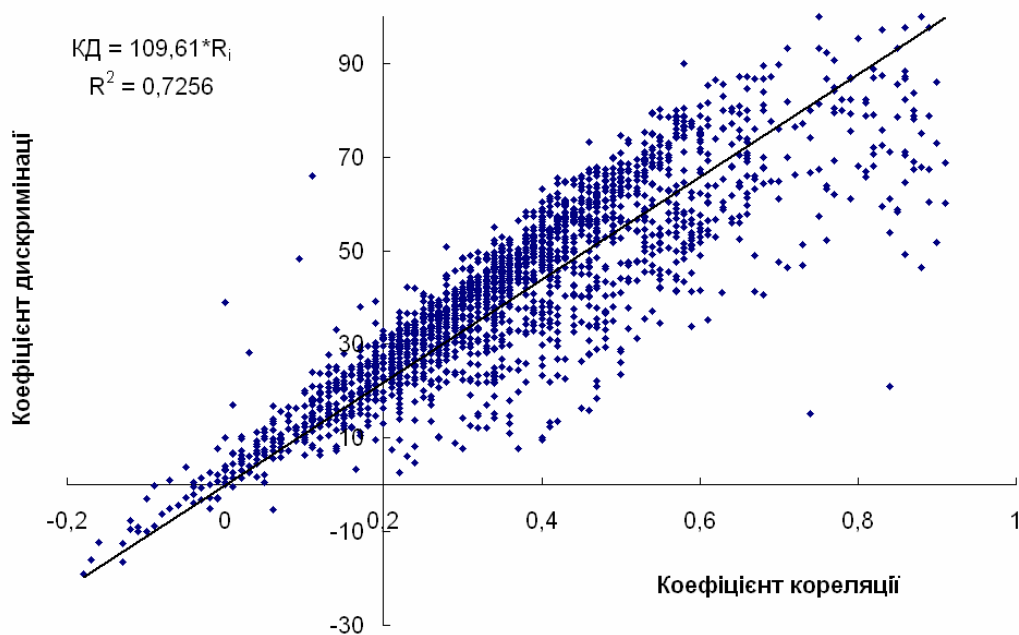


Рис. 3. Зв'язок між коефіцієнтами кореляції й дискримінації тестових завдань ЗНО 2009 – 2011.

Цікаво зазначити, що між показниками складності завдань і коефіцієнтами їх кореляції й дискримінації зв'язок відсутній. Зокрема для завдань, що мають високі (понад 0,7) значення коефіцієнта кореляції, показники складності варіюються від 4 до 75, а для завдань з низькими (менше 0,2) коефіцієнтами кореляції – від 3 до 100. Для „оптимальних” за складністю завдань зі значенням показника складності 40 – 60 значення коефіцієнта кореляції варіюються в межах від  $-0,09$  до  $+0,90$ .

Із загальних міркувань можна було б очікувати, що правильні відповіді на складні питання свідчать про більш високий рівень підготовки абітурієнта, тобто має бути певна від'ємна кореляція (не обов'язково лінійна) між показником складності й коефіцієнтом кореляції (дискримінації). Відсутність зв'язку, що спостерігається для

результатів ЗНО, може певною мірою пояснюватися зазначеними вище методичними помилками при розрахунку показника складності.

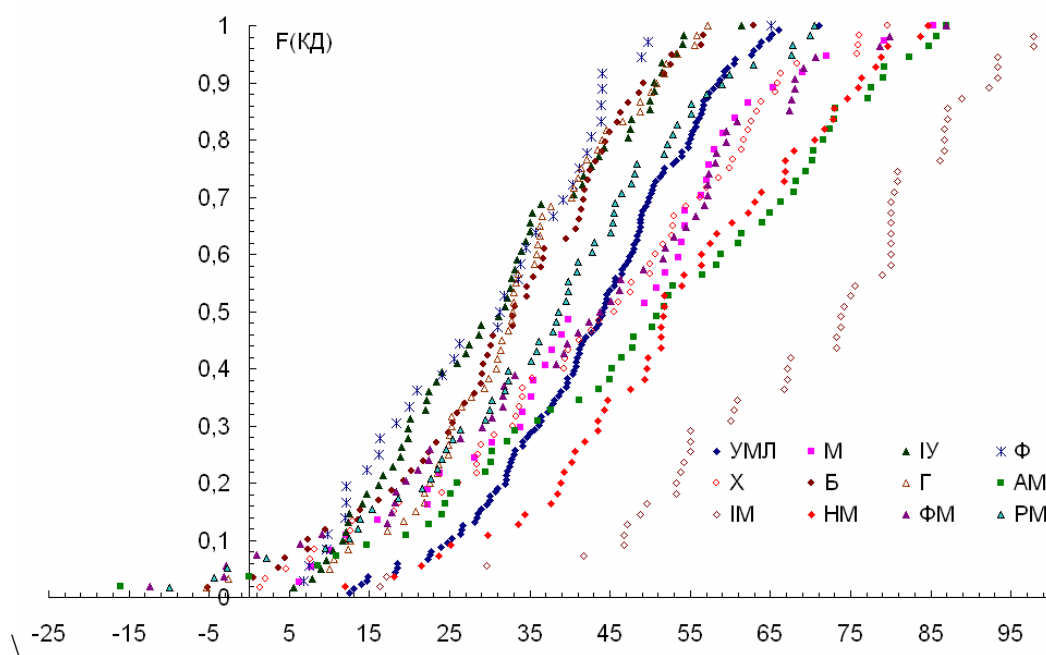


Рис. 4. Емпіричні функції розподілу коефіцієнтів дискримінації тестових завдань ЗНО-2011

Розподіл первинних балів ЗНО у більшості випадків істотно відрізняється від нормального, а часто є багатомодальним, що свідчить про неоднорідність груп абітурієнтів, що складають відповідні тести. Зокрема істотно неоднорідні розподіли для ЗНО-2011 отримано для результатів з української мови та літератури (друга сесія); хімії й іноземних мов. При цьому для тесту з англійської мови є помітний поділ на дві групи з істотно різним рівнем підготовки. Це може бути пов'язане з тим, що частина абітурієнтів складала тест з англійської мови як профільний для подальшого навчання на філологічних спеціальностях або спеціальностях потребують поглибленої мовної підготовки (міжнародна економіка, міжнародне право тощо). Для інших абітурієнтів він був непрофільним. Наприклад, такий тест могли скласти абітурієнти, що вступали на такі напрями, як прикладна математика, менеджмент організацій та багато інших.

Для деяких дисциплін (математика, фізика) існує висока імовірність вгадування необхідної для отримання позитивного результату (124 бали) кількості правильних відповідей (відповідно: 38 і 15 % у 2011 р; 22 і 44 % у 2010 р; 59 і 62 % у 2009 р.). При

цьому мінімальний прохідний первинний бал (у % від максимально можливого) не перевищує 30, а для тестів з математики, фізики й іноземних мов він знаходиться в межах 7 – 15 %.

Наведені дані свідчать про те, що в існуючому вигляді ЗНО не є якісним інструментом оцінювання якості освіти. Серед заходів, що можуть покращити ситуацію, на нашу думку слід у першу чергу використовувати таке: 1) накопичення бази даних тестових завдань з відсіюванням або корегуванням неякісних завдань; 2) запровадження різнотипних тестів (рівень знань, спроможність навчатися, спроможність до вирішення складних завдань) оскільки різні університети й спеціальності висувають різні вимоги до шкільної підготовки з дисциплін тестування; 3) вдосконалення методики визначення результатів тестування, зокрема врахування рівня складності завдань при призначенні балів за правильні відповіді на них; 4) надання провідним університетам права на проведення власних вступних випробувань.

#### **Список використаних джерел**

1. Bakhrushin, V.E. & Gorban, A.N. 2011 'Test technologies in education: The problem of test quality', *Ukr. J. Phys. Opt.*, V.12, *Suppl. 2 Sc. Horiz.*, pp. 1-10.
2. Аванесов, В.С. 2002, *Композиция тестовых заданий*. – М.: Центр тестирования, 240 с.
3. Аванесов, В.С. 2005 *Форма тестовых заданий*, М.: Центр тестирования, 156 с.
4. Ащепкова, Л.Я. 2003 *Конструирование тестовых заданий и обработка результатов тестирования*. – Владивосток: ДГУ, 50 с.
5. Бахрушин, В.Є. 2009 *Математичні основи моделювання систем*, Запоріжжя: КПУ, 224 с.
6. Бахрушин, В.Є. та Горбань, О.М. 2011, 'Тестові технології в освіті: Проблеми якості тестів', *Наукові записки Академії наук вищої школи України*, 6, с. 24- 34.
7. Бахрушин, В.Є.; Горбань, О.М. та Игнахина, М.А. 2011 'Статистичний аналіз тестів ЗНО 2009 – 2011', *Вища освіта України. Тематичний випуск "Вища освіта України в контексті інтеграції до Європейського освітнього простору"*, Додаток 2 до № 3, Т. 3 (28), с. 29-35.

8. Бахрушин, В.Е., Журавель, С.В. и Игнахина, М.А. 2010 'Автоматизация определения результатов тестирования' // *Управляющие системы и машины*, № 2, с. 10-12.
9. Казиев, В.М. 2008 *Введение в практическое тестирование*, М.: ИНТУИТ, <<http://www.intuit.ru/department/informatics/practest>>
10. Ким, В.С. 2007 *Тестирование учебных достижений*. – Уссурийск: Издательство УГПИ, 214 с.
11. Крисилов, В.А., Онищенко, Т.В. и Русинова, Н.В. 2004 'Методика анализа педагогических тестов по результатам тестирования', *Труды Одесского политехнического университета*, Вып. 2, с. 1-6.
12. Оганесян, А.Г., Дещинский, Ю.Л. и Бирюлев, К.Ю. 2010 'Тестирование или экзамен на компьютере?', *Образовательные технологии и общество*, Т. 13, № 1, с. 1-17.
13. Чельшкова, М.Б. 2002 *Теория и практика конструирования педагогических тестов: Учебное пособие*, М.: Логос, 432 с.
14. Український центр оцінювання якості освіти 2009, *Офіційний звіт про проведення зовнішнього незалежного оцінювання знань випускників загальноосвітніх навчальних закладів України в 2009 р.*, Київ, 321 с.
15. Український центр оцінювання якості освіти 2010, *Офіційний звіт про проведення зовнішнього незалежного оцінювання знань випускників загальноосвітніх навчальних закладів України в 2010 р.*, Київ, 352 с.
16. Український центр оцінювання якості освіти 2011, *Офіційний звіт про проведення зовнішнього незалежного оцінювання навчальних досягнень випускників загальноосвітніх навчальних закладів України в 2011 р.*, Київ, 359 с.