

МОДЕЛЬ РЕТЕСТОВОЇ НАДІЙНОСТІ, АЛЬТЕРНАТИВНА МОДЕЛІ ВНУТРІКЛАСОВОЇ КОРЕЛЯЦІЇ

Заневський І.П., Заневська Л.Г.

Львівський державний університет фізичної культури

Постановка проблеми. Тест, зазвичай, вважається надійним, якщо при його застосуванні отримують ті ж самі результати для кожного обстежуваного при різних тестуваннях в однакових умовах. Для практики найбільшої ваги мають три форми прояву надійності: ретестова надійність, надійність паралельних форм тесту та надійність частин тесту (Л.К. Джекобс, 1991).

Ретестова надійність – це спосіб визначення надійності тесту, при якому показники точності та усталеності результатів визначаються шляхом повторного дослідження. Надійність у цьому разі встановлюється як міра збігу результатів першого й повторних обстежень або як ступінь збереження рангових місць обстежених у вибірці при ретесті (К. Черрі, 2003). Іншими словами ретестова надійність – це стабільність результатів повторних тестувань.

За кількісну міру надійності прийнято коефіцієнт надійності, величина якого дорівнює відношенню дисперсії істинних результатів до дисперсії отриманих результатів тесту (А. Маховські, 1993). Арифметичне значення кореня квадратного з величини коефіцієнта надійності називають індексом надійності, який є теоретичним коефіцієнтом кореляції між істинними й зареєстрованими результатами тесту. За визначенням коефіцієнт надійності та індекс надійності можуть приймати значення від нуля до одиниці.

Оскільки в практичній роботі з тестами істинні значення результатів тесту невідомі, для визначення величини коефіцієнта надійності застосовують опосередковані методи. Найбільш поширеним з них є метод двофакторного дисперсійного аналізу з визначенням коефіцієнта внутрікласової кореляції як оцінки величини коефіцієнта надійності (Дж. Бартко, 1966). Величина цього коефіцієнта дорів-

нює відношенню різниці дисперсії результатів об'єктів тестування і сумарної дисперсії повторних вимірювань та дисперсії взаємодії з об'єктами тестування до дисперсії результатів об'єктів тестування. Верхньою границею величини коефіцієнта внутрікласової кореляції є одиниця, а нижня границя може бути й від'ємним числом, що входить у протиріччя з визначенням коефіцієнта надійності.

За різними оцінками (Дж. Морроу, 2006; В.М. Заціорський, 1982; С.М. Морозов, 1994) вважаються прийнятними для практичного застосування тести, коефіцієнт надійності яких не менший від числа, яке знаходиться в межах 0,5 ... 0,8. Використовують шкалу оцінок надійності тесту із розділенням загального діапазону прийнятних значень коефіцієнта надійності тесту (тобто, принаймні, додатних) на інтервали: відмінна, добра, задовільна, сумнівна й неприйнятна надійність. Оскільки нульове значення коефіцієнта внутрікласової кореляції не є його мінімальним значенням, відповідна шкала оцінок не повністю відповідає визначенню коефіцієнта, прийнятому у класичній теорії надійності тестів. Тобто, від'ємні значення нижньої границі діапазону величин коефіцієнта внутрікласової кореляції суперечать моделі коефіцієнта надійності тесту, в якій ця границя є нульовою.

Отож у теорії спортивних тестів існує проблема неповної відповідності моделі внутрікласової кореляції теоретичній моделі надійності тесту.

Аналіз попередніх досліджень і публікацій. Основи класичної теорії тестів розроблено в середині минулого століття Г. Гулліксеном (1950), а пізніше розвинуто Ф.М. Лордом і М.Р. Новіком (1986), а також Л. Крокером і Дж. Алгино (1986) стосовно теорії психометрії. Початки дисперсійного аналізу та внутрікласової кореляції, які застосовують для оцінки величини коефіцієнта надійності тесту, були закладені наприкінці дев'ятнадцятого століття Р.А. Фішером (1925). Обґрунтуванню різних варіантів методу внутрікласової кореляції для оцінки величини коефіцієнта надійності тесту присвячено роботи Дж. Бартко (1966), Л.Дж. Кронбаха, Г.К. Глесера, Г. Нанда й Н. Раяратнам (1972), а також – П.Е. Шраута із Дж.Л. Фляйссем (1979).

Специфіці застосування теорії тестів у спортивній практиці присвячено роботи Х. Бубе зі співавторами (1968), а також В.Л Уткіна (1978), В.М. Заціорського (1979) й У.Г. Гопкінса (2000). Найбільш докладно стан проблеми надійності моторних тестів розглянуто П. О'Донохью (2010).

Починаючи з останньої декади минулого століття математична теорія надійності тестів розвивалася у напрямі створення моделей, альтернативних моделі внутрікласової кореляції. Дж.М. Бланд і Д.Г. Альтман (1986) запропонували метод оцінки узгодженості двох тестів з використанням стандартного відхилення повторних спроб. А.М. Невілл і Г. Аткінсон (1997) розробили метод визначення коефіцієнта надійності для вимірів, проведених у шкалі відношень, на основі аналізу кількості вимірів, які потрапили до 95-процентного інтервалу довіри.

На початку нового століття А. Брутон, Дж.Х. Конвей і С.Т. Голгейт (2000) підвели ризик під піввіковим періодом інтенсивного розвитку математичних методів наближеної оцінки надійності тестів, представивши порівняльний аналіз запропонованих моделей і визначивши їхні переваги та недоліки. У.Г. Гопкінс (2000) доводить, що найточнішим параметром для кількісної оцінки надійності тесту є коефіцієнт варіації результатів повторних спроб.

На основі аналізу публікацій останньої декади може скластися враження, що математична теорія надійності тестів досягла завершеності, оскільки головні її здобутки були пов'язані зі створенням відповідного програмного забезпечення загального користування, наприклад, в комп'ютерних пакетах Statistica, SPSS, Mathematica, Mathcad та ін.

У відкритих джерелах інформації останніх трьох декад нами не виявлено публікацій стосовно проблеми, що розглядається.

Метою роботи було створення моделі коефіцієнта надійності тесту, застосування якої дало б можливість у практичній роботі з тестами отримувати значення коефіцієнта надійності в теоретичних границях від нуля до одиниці, як це прийнято у класичній теорії надійності тестів. Основна ідея розв'язання цієї проблеми полягає у прямому визначенні дисперсії істинних результатів тесту, оцінка величини яких приймається рівною середньому арифметичному результатів повторних вимірювань.

Завдання дослідження: показати неповну відповідність моделі внутрікласової кореляції класичній теорії ретестової надійності; розробити модель коефіцієнта надійності, який би загалом відповідав цій теорії; провести кількісний аналіз подібності моделі внутрікласової кореляції та альтернативної моделі.

Методика дослідження ґрунтувалася на положеннях класичної теорії спортивних тестів, на математичній теорії надійності мото-

рних тестів, на моделі внутрікласової кореляції, на дисперсійному аналізі й на методі Шапіра – Уїлка. Для визначення рівних величин коефіцієнта внутрікласової кореляції й альтернативного коефіцієнта надійності тесту було складено алгебричне квадратне рівняння, розв'язки якого були отримані в радикалах.

Обчислення проводилися з використанням комп'ютерної програми Excel (функції CORREL, VAR, AVERAGE, а також підпрограми з пакету “Tools >> Data Analysis >> Anova: Two Factor without Replication”) й програми Mathematica (функція Graphs >> Histograms >> Advanced >> Fit type: Normal >> Shapiro-Wilk test).

Основні результати. Від'ємне значення внутрікласового коефіцієнта кореляції, обчислене за результатами двофакторного дисперсійного аналізу, суперечить основам класичної теорії надійності тестів, згідно з якою коефіцієнт надійності не може бути від'ємним числом. Цього протиріччя можна уникнути, приймаючи модель коефіцієнта надійності тесту, в якій за істинні значення взяти середні арифметичні індивідуальних результатів повторних спроб, оскільки середнє арифметичне — це найкраще наближення для оцінки істинної величини у генеральній сукупності.

Висновки. Модель внутрікласової кореляції неповною мірою відповідає положенням класичної теорії надійності: величина відповідного коефіцієнта ретестової надійності, при її погіршенні, прямує до від'ємної нескінченності. Пропонована модель ретестової надійності позбавлена цього недоліку, а відповідний коефіцієнт надійності існує в межах, визначених положеннями класичної теорії надійності ($0 \div 1$). Найкраще наближення до істинного результату – це середнє арифметичне послідовних тестувань. В діапазоні практично важливих величин пропонована модель та модель внутрікласової кореляції призводять до різних оцінок надійності тесту не тільки кількісно ($\approx 20\%$), але й якісно, тобто із суттєвим перевищенням ширини діапазонів оцінки надійності тесту за величиною відповідного коефіцієнта.