

## Тема 10.2. “Аналіз об’єктивності результатів”

### КРИТЕРІЇ АНАЛІЗУ ОБ’ЄКТИВНОСТІ РЕЗУЛЬТАТІВ

// Система управління якістю медичної освіти в Україні: Монографія / Кол. авт. – Д., “АРТ-ПРЕС”, 2003. – С.162 -164.

Аналіз валідності результатів базується на спеціально розроблених методиках співставлення з критеріями результатів, в основі яких лежить оцінка міри відхилення від нормального закону результатів тестування та процедура їхньої нормалізації. Міру відхилення від нормального закону визначають за такими параметрами як асиметрія  $As$  і ексцес  $Ex$ , та встановлюванням їхньої статистичної достовірності за  $t$ -критерієм Стюдента. Залежно від значення та знаку асиметрії або ексцесу результати можуть бути поділені на п’ять загальних груп, які дають змогу встановити критерії прийняття рішення щодо валідності процесу вимірювання та наявності або відсутності статистично достовірної дії чинників, що впливають на результати вимірювання.

#### Критерій I

Якщо дані аналізу засвідчують, що  $As = 0$ ,  $Ex = 0$ , то можна зробити такі висновки: *результати вимірювання є об’єктивними, а отже і процес вимірювання є об’єктивним, що є необхідною умовою його валідності.*

#### Критерій II

Якщо дані аналізу засвідчують наявність від’ємної асиметрії та додатного ексцесу  $As = “-”$ ,  $Ex = “+”$ , то можна зробити такі висновки: *результати вимірювання є завищеними, тобто процес вимірювання є невалідним.*

Причиною цього може бути дія таких чинників як невалідність тесту, моделі або умов тестування. Така ситуація, окрім названих випадків, найчастіше має місце у разі втручання екзаменаторів у процес тестування, або за наявності у осіб, яких тестують, правильних відповідей або інших допоміжних матеріалів.

#### Критерій III

Якщо дані аналізу засвідчують наявність від’ємних асиметрії та ексцесу  $As = “-”$ ,  $Ex = “-”$ , то можна зробити такі висновки: *результати є завищеними за рахунок невалідності тесту.*

#### Критерій IV

Якщо дані аналізу засвідчують наявність додатної асиметрії та від’ємного ексцесу:  $As = “+”$ ,  $Ex = “-”$ , то можна зробити такі висновки: *результати є заниженими за рахунок невалідності тесту.*

#### Критерій V

Якщо дані аналізу засвідчують, що  $As = “+”$ ,  $Ex = “+”$ , то можна зробити такі висновки: *результати вимірювання є заниженими за рахунок невалідності тесту, форми діалогу, умов тестування.*

#### Критерії аналізу якості тесту

Представлені вище критерії об’єктивності результатів дають змогу на практиці провести визначання валідності інструментарію та процедури

вимірювання, процедури оцінювання; візуалізувати наявність порушення валідності, якщо це має місце; встановити чинники, що спричинили порушення. Кількісний аналіз параметрів інструментарію може бути проведений лише після експерименту, тобто на апостеріорному рівні з використанням основних положень і висновків теорії тестування. Зупинимося лише на аналізуванні характеристик тесту.

Аналізування характеристик тесту, так само як і аналізування тестових завдань, проводять за єдиною методикою, яка ґрунтується на системному аналізі функції розподілу результатів тестування та характеристичної кривої тесту.

### *Аналіз функції розподілу результатів тестування*

Залежно від значення та знаків асиметрії та ексцесу закону розподілу результати можна згрупувати у дев'ять варіантів:

#### **Критерій I**

Тест є валідним, якщо асиметрія та ексцес закону розподілу дорівнюють нулю ( $As = 0$ ,  $Ex = 0$ ).

#### **Критерій II**

Якщо асиметрія є від'ємною, а ексцес додатнім ( $As = "-"$ ,  $Ex = "+"$ ), то тест є невалідним (легкий та з малою розподільчою здатністю). Це – результат дії декількох чинників:

- тест сконструйовано із невалідних тестових завдань (він мусить валідизуватися за допомогою їх ускладнення та збільшення їхньої розподільчої здатності);
- процедура тестування є невалідною.

#### **Критерій III**

Якщо асиметрія та ексцес є від'ємними ( $As = "-"$ ,  $Ex = "-"$ ), то тест є легким, але з великою розподільчою здатністю. Його валідизація полягає в ускладненні тестових завдань.

#### **Критерій IV**

Якщо асиметрія є від'ємною, а ексцес відсутній ( $As = "-"$ ,  $Ex = 0$ ), то тест є легким. Його валідизація полягає в ускладненні тестових завдань.

#### **Критерій V**

Якщо асиметрія є додатною, а ексцес від'ємним ( $As = "+"$ ,  $Ex = "-"$ ), то тест є складним з великою розподільчою здатністю. Його валідизація полягає у валідизації змісту за допомогою вилучення тестових завдань, які не відповідають цій вимозі.

#### **Критерій VI**

У разі додатних асиметрії та ексцесу ( $As = "+"$ ,  $Ex = "+"$ ) тест є складним з малою розподільчою здатністю. Це – наслідок дії двох чинників, що впливають один на одний: порушення валідності змісту, невалідність тестових завдань. Валідизація тесту полягає у валідизації тестових завдань.

#### **Критерій VII**

У разі додатної асиметрії та відсутності ексцесу ( $As = "+"$ ,  $Ex = 0$ ), тест є складним. Це є наслідком дії двох чинників, що впливають один на одного – порушення валідності змісту, невалідність тестових завдань. Валідизація тесту

полягає у валідазації тестових завдань. Особливої уваги потребують два останніх критерії – VIII та IX, які дають можливість визначити конструктивну валідність тесту, а саме перевірити тест на однорідність. Нагадаємо, що тест вважається однорідним, якщо його сконструйовано із тестових завдань з близькими характеристиками. Неоднорідним будемо вважати такий тест, який складається із окремих субтестів, кожний з яких можна розглядати як окремий тест з певними характеристиками.

### **Критерій VIII**

Якщо крива розподілу симетрична, асиметрія дорівнює нулю, а ексцес додатний ( $As = 0$ ,  $Ex = "+"$ ), то тест є неоднорідним і сконструйованим із двох субтестів: складного та легкого. Валідація тесту полягає у його розщепленні на два окремих тести.

### **Критерій IX**

Якщо крива розподілу симетрична, асиметрія дорівнює нулю, а ексцес від'ємний ( $As = 0$ ,  $Ex = "-"$ ), то тест є неоднорідним і сконструйованим із трьох субтестів: легкого, середнього та складного. Валідація тесту полягає у його розщепленні на окремі тести.

*Узагальнення.* Визначаючи основні категорії вимірювання рівня знань, можна зробити висновок про те, що валідність методу вимірювання залежить від валідності кожної із означених категорій. Таким чином, поруч з валідністю інструментарію важливе місце посідає валідність процедури вимірювання, через те, що невідповідність характеристик цієї категорії критеріям валідності робить невалідним результат вимірювання, навіть якщо інструментарій тестування атестовано як валідний. Наведені загальні критерії об'єктивності результатів тестування та критерії кількісного аналізування тестів дають можливість ідентифікувати порушення валідності процедури вимірювання.

Найчастіше наслідком порушення валідності процедури вимірювання є випадки, які класифіковано як II та V критерії загальних критеріїв об'єктивності результатів. У цих випадках порушення нормального закону розподілу результатів тестування пов'язані з наявністю додатного ексцесу, що засвідчує зменшення розподільчої здатності тесту.