

2 Славинская, О. В. Опыт использования платформы Online Test Pad в методике преподавания психолого-педагогических дисциплин / Славинская О. В. // Дистанционное обучение – образовательная среда XXI века: материалы XII Международной научно-методической конференции, Минск, 26 мая 2022 г. / редкол.: Е. Н. Шнейдеров [и др.]. – Минск : БГУИР, 2022. – С. 142–143.

УДК 004.9

НАДЕЖНОСТЬ И СТАНДАРТНАЯ ОШИБКА ИЗМЕРЕНИЯ ПЕДАГОГИЧЕСКОГО ТЕСТА

О.Б. Ганак, И.П. Овсянникова

Белорусский государственный университет пищевых и химических технологий,
г. Могилев, Республика Беларусь

В настоящее время в высшей школе широко применяется метод компьютерного педагогического тестирования как один из методов контроля усвоения студентами учебных компетенций по дисциплине. Инструментом для измерения по шкале достижений студента является правильно сконструированный тест, который соответствует не только предмету обучения, но и его задачам и служит развитию системного подхода к изучению учебной дисциплины.

Компьютерное тестирование имеет ряд преимуществ:

- быстрое получение результатов испытания и освобождение преподавателя от трудоемкой работы по обработке результатов тестирования;
- объективность в оценке;
- тестирование на компьютере более интересно по сравнению с традиционными формами опроса, что создает положительную мотивацию у студентов.

Чтобы тест адекватно оценивал знания, умения и навыки, он должен обладать определенными характеристиками. Одной из таких характеристик и важным количественным показателем качества теста является надежность. Она отражает объективность и точность педагогического измерения, степень постоянства, стабильности результатов тестирования.

Авторами была произведена оценка надежности тестов при проведении итогового контроля знаний студентов 1 курса по дисциплине «Информатика». Использовался метод расщепления, при котором оценка надежности строится на подсчете корреляции между двумя наборами результатов выполнения одного и того же теста или двух его параллельных форм. Чем выше корреляция, тем надежнее тест. Хорошим коэффициентом надежности теста считается тот, когда показатель колеблется в пределах $0,8 < r < 1$. Были получены следующие результаты: бинарная матрица результатов тестирования итогового контроля знаний студентов; коэффициент надежности теста по формуле Кьюдера-Ричардсона, равный 0,8353 и коэффициент надежности по формуле Спирмена-Брауна, равный 0,8148 [1, с.164].

Один из аспектов применения коэффициента надежности связан с определением стандартной ошибки измерения. Ошибка измерения определяется как статистическая величина, отражающая степень отклонения наблюдаемого балла от истинного балла испытуемого.

Стандартную ошибку измерения надежности теста индивидуального балла испытуемого можно вычислить по формуле[2]:

$$S_E = S_X \cdot \sqrt{1 - r_t} , \quad (1)$$

где S_X — стандартное отклонение по распределению индивидуальных баллов;
 r_t — коэффициент надежности теста;

S_E – стандартная ошибка измерения.

Для бинарной матрицы было вычислено стандартное отклонение $S_x = 1,02$ по формуле:

$$S_x = \sqrt{\frac{1}{N(n-1)} \left(N \sum_{i=1}^N X_i^2 - \left(\sum_{i=1}^N X_i \right)^2 \right)} \quad (2)$$

Коэффициент надежности для этой же матрицы, рассчитанный по формуле Спирмена-Брауна, равен $r_r = 0,8148$. Тогда для стандартной ошибки измерения получим

$$S_E = 1,02 \cdot \sqrt{1 - 0,8148} = 0,4389$$

Допустим, что i -й испытуемый выполнял много раз один и тот же тест. Если предположить, что эффект запоминания отсутствует, то результаты тестирования образуют нормальное распределение вокруг истинного балла T_i со стандартным отклонением S_E . На практике S_E рассматривается как статистическая величина, отражающая степень точности отдельных измерений, поэтому величину S_E используют для определения границ доверительного интервала, внутри которого должен находиться истинный балл оцениваемого студента группы.

Принято, что доверительный интервал выстраивается как две симметричные окрестности (левая и правая) вокруг наблюдаемого показателя студента. Этот факт вынуждено игнорируется в прикладных исследованиях в силу отсутствия истинного балла, и доверительный интервал при заданном риске допустить ошибку $t=0,05$, принимается равным $(X_i - 1,96 S_E; X_i + 1,96 S_E)$, где i —наблюдаемый балл i -го испытуемого; 1,96 – константа, табличное число, используемое при $t=0,05$.

Тогда доверительный интервал для истинного балла студента со значением $X_i=7$ будет $(7 - 0,86; 7 + 0,86)$ или $(6,14; 7,86)$. Следовательно, истинный балл испытуемого студента находится в промежутке от 6 до 8 баллов.

При оценке надежности нельзя полагаться лишь на один показатель, поскольку каждый из них имеет свои ограничения, смещающие оценки надежности теста в сторону завышения или занижения. Для достоверной проверки качества теста следует учитывать несколько показателей надежности. Таким образом, вопросы определения надежности теста, его стандартной ошибки, области локализации истинного тестового балла очень важны для создания качественного педагогического теста.

Список литературы

1. Овсянникова, И. П. Анализ надежности теста для оценки учебных компетенций студентов / И. П. Овсянникова, О. Б. Ганак // Качество подготовки специалистов в техническом университете: проблемы, перспективы, инновационные подходы: материалы V Международной научно-методической конференции 19-20 ноября 2020 года / Учреждение образования "Могилевский государственный университет продовольствия"; редкол.: А. С. Носиков (отв. ред.) [и др.]. - Могилев: МГУП, 2020. - с. 164 – 166.

2. Чельшкова М.Б. Теория и практика конструирования педагогических тестов: учебное пособие. М.: Логос, 2002- 432с.