

МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2002
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2002
Proceedings of Thirty First Spring Conference of
the Union of Bulgarian Mathematicians
Borovets, April 3–6, 2002

КАЧЕСТВЕНА ИНТЕРПРЕТАЦИЯ НА РЕЗУЛТАТИТЕ ОТ
ТЕСТОВЕ ЗА ПОСТИЖЕНИЯ*

Кирил Георгиев Банков

Разглежда се метода на базисните точки, който дава възможност да се опишат знанията и уменията на учениците в различни точки от скалата за оценяване на постиженията им при провеждане на тест. Методът предлага начин за характеристика на скалата за оценяване от качествена гледна точка. В известен смисъл той помага и за обяснение на резултатите за масовата публика.

Увод. При отчитане на резултатите от тест за постижения основните акценти са предимно от количествено естество. Например, сериозно внимание се отделя за да се сравнява представянето на отделни групи ученици; да се разпределят учениците по отношение на зададени норми или критерии и др. Този тип изследвания са сравнително ясно поставени на основата на използваните статистически методи и процедури. Значителни трудности се наблюдават, обаче, при опити за качествени анализи на скалите за постижения. Например, не е лесно да се опише какви знания и умения имат учениците, чиито постижения са в определена точка от скалата, както и какви са техните основни пропуски.

Въпросите от качествено естество са важни за по-доброто разбиране на скалата на даден тест за постижения, както и за правилната интерпретация на резултатите от гледна точка на предметната област на теста. В тази статия се разглежда един метод за качествен анализ върху скалата за постижения: *метода на базисните точки*. Той е разработен за нуждите на американската организация National Assessment of Educational Progress (NAEP) [1]. Използван е и за анализ на скалите в международното изследване на обучението по математика и природни науки TIMSS 1999 [2].

Описание на метода. Методът на базисните точки се използва за анализиране на знанията и уменията на учениците, чиито постижения са в определени точки (базисни) от скалата. Той е приложим за всякакъв вид непрекъсната скала. В зависимост от избора на базисните точки методът дава възможност да се каже какво знаят и могат учениците, представили се по различен начин (“лошо”, “средно добре”, “добре” и т.н.) на даден тест. Основните стъпки при изпълнение на метода са следните:

1. Избират се няколко точки $a_1 < a_2 < \dots < a_n$ от скалата на постиженията. Тези точки се наричат *базисни*.

2. За всяко $i = 1, 2, \dots, n$ се определя множеството A_i от ученици, чиито постижения по теста са равни на a_i .

*Тази работа е спонсорирана от фонд “Научни изследвания” на СУ “Св. Кл. Охридски”.

3. За всяко $i = 1, 2, \dots, n$ се определят задачите от теста, които “характеризират” базисните точки. Това са задачи, които са решени от “болшинството” ученици A_i . Условието е поне $p\%$ от учениците A_i да са решили съответната задача, където p е определено число. За всяко $i > 1$ се поставя и условието задачата да не е решена от “много” ученици от групата A_{i-1} , т.е. по-малко от $q\%$ ученици от A_{i-1} да са решили задачата, където q е също определено число. За определените така задачи се казва, че са *асоциирани към базисната точка a_i* .

4. Нека B_i е множеството от задачи, асоциирани към точката a_i . Екип от специалисти по предметната област на теста анализират множествата B_i , $i = 1, 2, \dots, n$ с цел да опишат знанията и уменията на всяка от групите ученици A_i .

Избиране на базисните точки. Теоретично всяка точка от скалата може да бъде базисна. Това не означава, че техният избор става хаотично. В началото се определя метод за избиране. Той може да бъде:

- *Емпиричен* – точките се определят в проценти или в интервали от стандартното отклонение;
- *Експертен* – експертна група определя точките в така наречените “нива на постижения”, като подробно описва какво учениците би трябвало да знаят и могат в съответното ниво.

Например, в началото NAEP са използвали емпиричния метод изразен чрез смислени интервали от стандартната грешка. По-късно те са преминали към експертния метод чрез определяне на три нива на постижения: основно, средно, напреднало.

Независимо от метода, по който става определянето на базисните точки, те трябва да са достатъчно много, за да дават възможност за смислено описание на скалата на постиженията, т.е. да обединят групите ученици, които имат различни постижения. От друга страна, базисните точки трябва да са на достатъчно голямо разстояние една от друга, за да разграничават учениците, чиито постижения са в две съседни точки.

Например, в международното изследване TIMSS 1999 се използват емпирично определените точки 25-ти, 50-ти, 75-ти и 90-ти перцентил.

Определяне на множеството ученици за всяка базисна точка. Обикновено броят на учениците, които влизат в някои (или всички) групи A_i (така както са дефинирани в точка 2 по-горе), е малък. За да получим достатъчно много ученици във всяка група, се постъпва по следния начин. Около всяка базисна точка a_i се определя околност $[a_i - \varepsilon, a_i + \varepsilon]$ и под A_i се разбира множеството от ученици, чиито постижения влизат в тази околност. Изборът на тези околности трябва да отговаря на следните две условия. Първо, те да са “големи”, за да може всяка от тях да осигури достатъчно много ученици. Второ, те трябва да са разположени “далеч” една от друга, за да разграничават добре учениците от две различни околности по отношение на постиженията им.

Например, по времето, когато NAEP са ползвали емпиричния метод за определяне на базисните точки, околностите са избирани плюс и минус една четвърт от стандартното отклонение около съответната базисна точка.

Скалата за ученическите постижения на изследването TIMSS е конструирана чрез Теорията за верния отговор (Item Response Theory – IRT) за средна стойност 500 и стандартно отклонение 100. Ако околностите са изберат плюс и минус 25

($\varepsilon = 25$) около всяка базисна точка, те акумулират достатъчно много ученици, но не може да се направи добра разлика в постиженията на учениците от 75-тия и 90-тия проценти. Оптимално решение се получава, като околностите се изберат плюс и минус 5 ($\varepsilon = 5$) около всяка базисна точка. Следната таблица показва разпределението на интервалите за базисните точки на данните по математика за TIMSS 1999 [2].

Базисна точка	25-ти проц.	50-ти проц.	75-ти проц.	90-ти проц.
Стойност	396	479	555	616
Интервал ($\varepsilon = 5$)	391 – 401	474 – 484	550 – 560	611 – 621
Брой ученици	3540	5690	5531	3703

Определяне на задачите, асоциирани към базисните точки. Първият въпрос, който изниква тук, е свързан с определянето на процентите p и q . Читателят лесно ще съобрази как се влияе изборът на задачите асоциирани към дадена базисна точка с увеличаване или намаляване на всеки от тях. Да разгледаме примера с TIMSS като представителен. При него е избрано $p = 65\%$ и $q = 50\%$. Това означава, че една задача се асоциира към 25-тия перцентил, ако поне 65% от учениците в групата A_1 са я решили. (За първата базисна точка процентът q не участва.) Една задача се асоциира към 50-тия перцентил, ако поне 65% от учениците в групата A_2 са я решили и по-малко от 50% от учениците в групата A_1 са я решили. По аналогичен начин се определят и задачите, асоциирани към 75-тия и 90-тия перцентил.

Пример: За някаква (хипотетична) задача в TIMSS са получени следните данни:

Базисна точка	25-ти проц.	50-ти проц.	75-ти проц.	90-ти проц.
Процент вярно решили от групата A_i	14	43	72	86

Тогава тази задача се асоциира към 75-тия перцентил.

Ако асоциираните задачи към някоя от базисните точки са малко, броят им може да се увеличи като се причислят и така наречените “почти асоциирани задачи”. Това са такива задачи, при които процентът p е “малко” занижен. Например, в TIMSS почти асоциирани са онези задачи, при които $p = 60\%$. Тогава, задачата с данни:

Базисна точка	25-ти проц.	50-ти проц.	75-ти проц.	90-ти проц.
Процент вярно решили от групата A_i	47	63	79	93

е почти асоциирана към 50-тия перцентил.

Ако се окаже, че множеството от асоциираните и почти асоциираните задачи все още е малко, може да се пренебрегне и условието за процента q . Така получаваме множество от задачи, които са асоциирани само въз основа на p критерия. Например, задачи с данни:

Базисна точка	25-ти проц.	50-ти проц.	75-ти проц.	90-ти проц.
Процент вярно решили от групата A_i	26	53	64	78

е асоциирана към 75-тия перцентил асоциирани само въз основа на p критерия.

Следващата таблица показва разпределението на броя математически задачи, асоциирани към базисните точки за TIMSS 1999 [2].

	Асоциирани	Почти асоциирани	Само по p критерия	Общо
25-ти проц.	4	2	0	6
50-ти проц.	16	7	13	36
75-ти проц.	34	14	25	73
90-ти проц.	17	4	22	43
Общо	71	27	60	158

От данните ясно се вижда, че включването на почти асоциирани задачи, както и на такива, които са асоциирани само въз основа на p критерия, значително увеличава броя на задачите, характеризиращи постиженията на учениците в базисните точки.

Описание на знанията и уменията на учениците за всяка базисна точка. Екип от 10-12 специалисти преглеждат задачите, асоциирани към всяка от базисните точки. Целта е да опишат знанията и уменията, които учениците трябва да имат, за да решат задачите от съответното множество B_i . Това описание в определен смисъл характеризира групата ученици от съответната базисна точка.

Ето какви са кратките описания на учениците в базисните точки на международно ниво по математика за учениците от 8 клас в TIMSS 1999 [3].

25-ти перцентил: Учениците могат да извършват основни действия с естествени числа.

50-ти перцентил: Учениците могат да прилагат основни математически знания в познати ситуации. Те събират и изваждат естествени числа и десетични дроби за да решат едностъпкови задачи с думи; определят нагледно представени обикновени дроби по относителната големина на представянето; намират неизвестен член на пропорция; познават основните понятия свързани с пропорционалност и процент; използват основни свойства на геометричните фигури; разчитат и интерпретират данни, представени графично, с таблици или скали; разбират смисъла на прости алгебрични изрази.

75-ти перцентил: Учениците могат да прилагат знанията и уменията си в относително по-трудни и разнородни ситуации. Те извършват пресмятания с дроби за решаване на задачи с думи; решават многостъпкови задачи с думи включително с използване на пропорционалност; решават задачи от елементарна вероятност; използват знания за геометричните фигури за решаване на задачи; извършват пресмятания с алгебрични изрази и решават (линейни) уравнения с едно неизвестно.

90-ти перцентил: Учениците могат да организират информация, да правят обобщения и да обясняват решения на нерутинни задачи в нетипични ситуации. Те прилагат числови, геометрични и алгебрични зависимости при решаване на задачи; правят еквивалентни преобразувания на алгебрични изрази.

Валидност на метода. Естествена критика на метода на базисните точки е използването на експертни оценки на различни места в прилагането му и особено в точка 4. Тези критики имат своите основания. Ясно е, че правилното определяне на математическите знания и умения, заложените във всяка задача, са основата на крайния продукт от прилагането на метода.

През 1992 г. е направено изследване на валидността на метода върху данни на NAEP, като описанието на знанията и уменията на учениците по математика във всяка от базисните точки е правено от две групи експерти независимо една от друга и описанията на двете групи са сравнени [4]. Авторите отбелязват, че описанията на двете групи си приличат и с точност до редакция описват идентично знанията и уменията на учениците в базисните точки.

Описанието на постиженията на учениците, което се получава чрез метода на базисните точки, е важно за популяризиране на резултатите от тест. На масовата публика те се представят във вид много близък до това, което групата експерти дава като характеристика на учениците в базисните точки.

ЛИТЕРАТУРА

- [1] A. E. BEATON, N. L. ALLEN. Interpreting scales through scale anchoring. *Journal of Educational Statistics*, **17** (1992), 191–204.
- [2] M. O. MARTIN et al. TIMSS 1999 Technical Report, Boston College (2000), 266–276.
- [3] M. O. MARTIN et al. TIMSS 1999 International Mathematics Report, Boston College (2000), 55–90.
- [4] I. V. S. MULLOS, E. G. JOHNSON. The NAEP scale anchoring process for the 1992 mathematics assessment. In: The NAEP 1992 technical report. (Eds E. G. Johnson, J. E. Carlson) Washington, DC, National Center for Educational Statistics (1994), 893–907.

Кирил Банков
Факултет по математика и информатика
СУ “Св. Климент Охридски”
Бул. “Джеймс Баучер” 5
1164 София
e-mail: kbankov@fmi.uni-sofia.bg

QUALITATIVE INTERPRETATION OF THE RESULTS FROM ACHIEVEMENT TESTS

Kiril Gueorguiev Bankov

The paper presents the scale anchoring method, which is used to characterize student's performance at different points on an achievement scale. The method gives a meaning of a scale score in terms of the subject matter of the test. This manner of interpretation of the scale score helps the policymakers, educators and the public to understand better results of achievement tests.