

CÁC LOẠI ĐIỂM TRẮC NGHIỆM THƯỜNG ĐƯỢC SỬ DỤNG TRONG ĐÁNH GIÁ NĂNG LỰC NGƯỜI HỌC

TRẦN THỊ CẨM THO*

Ngày nhận bài: 29/10/2016; ngày sửa chữa: 11/11/2016; ngày duyệt đăng: 21/11/2016.

Abstract: To measure the capacity of students in learning subjects, ten-band scale has been applied commonly at schools. In this article, author introduces types of scores of tests used to assess learner's competence in Vietnam in current period. The article can be seen as a reference for assessing students' competence in the planned National Multiple-choice tests in the 2016-2017 academic year.

Keywords: Test scores, learner's competence.

Đánh giá năng lực người học yêu cầu ứng dụng phương pháp để phân tích quá trình và kết quả giáo dục, khuynh hướng là cần lượng hóa chính xác. Thành tích của người học về một bài trắc nghiệm (TN) cần đo lường theo một thang bậc để hiểu và giải thích, so sánh được một cách đồng nhất. Do vậy, trong quá trình lượng hóa, cần một thang điểm, loại điểm TN phù hợp.

1. Điểm thô

Tính điểm thô dựa vào phương pháp điểm số tích lũy, nghĩa là người đánh giá dựa trên các tiêu chuẩn để tiến hành cho điểm, sau đó cộng các điểm số cho từng hạng mục lại sẽ thu được tổng số điểm của người học. Giáo viên dùng tổng số điểm này để đưa ra phán đoán đối với người học.

Năng lực của một người học về một môn học có thể được đánh giá thông qua đề TN. Một đề TN thường gồm nhiều câu hỏi (CH), mỗi CH được gán với một điểm số, chẳng hạn: CH nhị phân là điểm 1 nếu làm đúng, điểm 0 nếu làm sai. Sau khi chấm bài TN và cộng các điểm số của từng thí sinh, ta thu được điểm số của các thí sinh, được gọi là điểm thô (raw score). Đối với đề tự luận, người ta có thể gán một điểm nào đó cho từng ý, từng nội dung mà thí sinh trả lời được; sau khi chấm xong cộng điểm của mỗi thí sinh lại, ta cũng có một điểm thô.

Hiện nay, ở bậc phổ thông, để đánh giá kết quả học tập của học sinh (HS), giáo viên dùng điểm thô với thang điểm 10 là chủ yếu.

2. Điểm phần trăm đúng

Điểm phần trăm đúng được xác định dựa trên việc so sánh điểm thô của một thí sinh với điểm tối đa có thể đạt được từ đề thi TN, sau đó tính theo tỉ lệ phần trăm, do đó sẽ không bị ảnh hưởng bởi điểm số của

những người khác trong nhóm. Điểm số loại này có thể được tính theo tỉ lệ phần trăm số CH làm đúng trên tổng số CH của đề TN, theo công thức dưới đây:

— ; trong đó: X là điểm tính theo tỉ lệ phần trăm; D là số câu TN làm đúng; T là tổng số câu TN.

Cách tính loại điểm này đơn giản và thường quy định nếu đạt từ 50% trở lên là đạt yêu cầu, còn dưới 50% là chưa đạt. Việc quy định như vậy không mang tính khách quan và không đo lường chính xác được khả năng của HS (vì ta có thể viết các câu TN dễ hơn nếu muốn HS có điểm cao hoặc ngược lại).

3. Điểm chữ

Điểm chữ cũng căn cứ trên tỉ lệ phần trăm số câu trả lời đúng, nhưng được biểu thị bằng các chữ cái như: A, B, C, D, E, F,... Chẳng hạn: người ta ấn định như sau: A cho các điểm phần trăm đúng từ 90-100, B cho các điểm từ 80-89,...

Theo Văn bản hợp nhất số 17/VBHN-BGDĐT, ngày 15/5/2014 của Bộ trưởng Bộ GD-ĐT quyết định về việc ban hành Quy chế đào tạo đại học, cao đẳng hệ chính quy theo hệ thống tín chỉ thì cách tính điểm theo thang điểm chữ chuyển đổi điểm học phần từ điểm hệ 10 sang điểm chữ và từ điểm chữ sang điểm hệ 4, cụ thể như sau: - Từ 8,5-10: điểm chữ A loại giỏi; - Từ 7,0-8,4: điểm chữ B loại khá; - Từ 5,5-6,9: điểm chữ C loại trung bình; - Từ 4,0-5,4: điểm chữ D loại trung bình yếu; - Dưới 4,0: điểm chữ F loại kém và không đạt.

Cũng như điểm phần trăm đúng, loại điểm chữ là hoàn toàn ngẫu nhiên, không khách quan trong việc ấn định các mức điểm.

* Trưởng Cao đẳng Sư phạm Nghệ An

4. Điểm tiêu chuẩn (standard scores)

Một cách biến đổi các điểm thô trong bài TN ra thành một loại điểm khác gọi là điểm tiêu chuẩn, chúng được đặt cơ sở trên độ lệch tiêu chuẩn của phân bố điểm số. Điểm tiêu chuẩn có thể được quan niệm như là một cách phát biểu khoảng cách của điểm ấy tính từ giá trị trung bình (mean), hay có thể là những điểm số đã được gán cho một giá trị trung bình và độ lệch tiêu chuẩn nào đó. Theo [1; tr 188], điểm tiêu chuẩn được sử dụng nhiều hơn trong TN và nghiên cứu vì có những đặc tính sau: - Mỗi loại điểm tiêu chuẩn có giá trị trung bình và độ lệch tiêu chuẩn chung cho các bài TN và tất cả mọi người; - Điểm tiêu chuẩn vẫn giữ nguyên hình thái phân bố các điểm thô và chỉ thay đổi các số định cỡ (calibration); - Cho phép ta thực hiện việc so sánh TN này với TN kia, nhóm này với nhóm kia. Đó là điều không thực hiện được với đa số các loại điểm khác; - Chúng có thể được xử lý bằng các phương pháp toán học trong khi một số các loại điểm khác không thực hiện được.

Có 2 loại điểm tiêu chuẩn thông dụng đã được sử dụng ở nước ta từ trước đến nay:

4.1. Điểm Z (Z score). Điểm Z là loại điểm tiêu chuẩn căn bản, tất cả các loại điểm tiêu chuẩn khác đều có thể được thiết lập từ đó. Nó cho ta biết hiệu số giữa một trị số điểm thô nào đó với điểm trung bình của nhóm làm bài TN với công thức:

$$Z = \frac{X - \bar{X}}{s} \quad (1); \text{ trong đó: } X \text{ là một điểm thô nào đó;}$$

\bar{X} là điểm thô trung bình của nhóm làm TN; s là độ lệch tiêu chuẩn của nhóm ấy.

Mặc dù điểm tiêu chuẩn Z có nhiều thuận lợi cho quá trình nghiên cứu, nhưng có sự bất tiện là khoảng một nửa số điểm Z là số âm và tất cả các điểm Z đều phải chứa một hay nhiều số lẻ. Vì vậy, cần tìm ra các loại điểm tiêu chuẩn khác để loại đi các trị số âm này.

4.2. Điểm tiêu chuẩn V. Điểm tiêu chuẩn V được đưa ra dựa trên điểm tiêu chuẩn Z, nhưng loại trừ được các điểm âm, đồng thời phù hợp với hệ thống điểm từ 1-20; quy về phân bố bình thường, với trung bình là 10 và độ lệch tiêu chuẩn là 4. Như vậy, để biến đổi các điểm thô ra điểm tiêu chuẩn V, trước hết cần biến đổi điểm thô ra điểm tiêu chuẩn Z, sau đó áp dụng công thức đơn giản sau:

$$\text{Điểm chuẩn } V = 4Z + 10 \quad (2)$$

Ví dụ: Một thí sinh có điểm thô là 30 ở một bài TN có 32 câu, biết rằng điểm trung bình của nhóm thí sinh ấy là 19 và độ lệch tiêu chuẩn là 7. Trước hết, cần

biến đổi điểm thô đó thành điểm tiêu chuẩn Z thông qua công thức (1): $Z = (30 - 19)/7 = 2$.

Sau đó, áp dụng công thức (2) để tính điểm tiêu chuẩn V cho điểm số 30 của thí sinh: $V = (4 \times 2) + 10 = 18$. Như vậy, thí sinh có điểm tiêu chuẩn V là 18.

Để phù hợp với hệ thống điểm từ 1-10, ta chọn điểm trung bình là 5 và độ lệch tiêu chuẩn là 2, biến đổi điểm tiêu chuẩn Z thành điểm tiêu chuẩn V bằng công thức: $V = 2Z + 5$ (3).

Điểm tiêu chuẩn có nhiều ưu điểm so với các loại điểm TN khác, có thể được sử dụng để thực hiện các phép tính. Tuy nhiên, vì điểm tiêu chuẩn phụ thuộc vào độ lệch tiêu chuẩn nên khó giải thích với phụ huynh HS về ý nghĩa của các điểm số TN.

5. Điểm thực

5.1. Điểm thực trong Lí thuyết cổ điển. Theo [2; tr 117], giả sử t là điểm trung bình của một thí sinh qua các phép đo đề TN. Điểm quan sát X của đề TN qua các phép đo được coi là một biến ngẫu nhiên, với một phân bố tần suất nào đó chưa biết. Giá trị trung bình (kì vọng toán học) của phân bố đó được gọi là điểm thực τ của thí sinh. Gọi là sai số của phép đo, có thể biểu diễn mối quan hệ giữa điểm thực τ , các điểm quan sát X và sai số ε như sau: $\varepsilon = X - \tau$.

Điểm thực định nghĩa như trên theo Lí thuyết cổ điển là khá trừu tượng, không có quy trình nào để xác định. Do đó, sai số của phép đo ε là một đại lượng có tính chất trung bình của phép đo nói chung với toàn bộ dải năng lực của thí sinh.

5.2. Điểm thực trong Lí thuyết ứng đáp câu hỏi. Theo [2; tr 118], giả sử câu hỏi TN được xét là CH nhị phân: trả lời đúng được 1 điểm, trả lời sai được 0 điểm. Điểm thô của một thí sinh sẽ thu được bằng cách cộng các điểm của tất cả CH trong đề TN. Như vậy, điểm thô của thí sinh đối với đề TN thường là một số nguyên nằm giữa 0 và n , trong đó n là số CH của TN. Giả sử một thí sinh làm lại đề TN (khi làm lại người đó không nhớ những gì đã làm lần trước), người đó sẽ được một điểm thô khác. Giả thiết là thí sinh làm đề TN nhiều lần và nhận được nhiều điểm thô khác nhau, các điểm này phân bố quanh một giá trị trung bình nào đó. Theo lí thuyết về đo lường, giá trị trung bình đó gắn với một giá trị được gọi là điểm thực, điểm thực phụ thuộc vào một lí thuyết đo lường xác định.

Giả sử X là điểm thô, thì $X = \sum_{j=1}^n U_j$, trong đó U là vectơ ứng đáp, được biểu diễn: $U = (U_1, U_2, \dots, U_j, \dots, U_n)$; U_j bằng 1 nếu trả lời đúng và bằng 0 nếu trả lời sai CH

thứ j . Nếu biểu diễn điểm thực là τ , ta có thể tính điểm thực theo biểu thức kì vọng toán học của X như sau:

$$\tau = E(X) = E\left(\sum_{j=1}^n U_j\right) = \sum_{j=1}^n E(U_j),$$
 trong đó E là kì vọng toán học. Vì U_j có giá trị bằng 1 với xác suất $P_j(\theta)$ và có giá trị bằng 0 với xác suất $Q_j(\theta) = [1 - P_j(\theta)]$, nên: $E(U_j) = 1 \cdot P_j(\theta) + 0 \cdot Q_j(\theta) = P_j(\theta)$

Do đó:
$$\tau = \sum_{j=1}^n P_j(\theta)$$

Điểm thực của một thí sinh có năng lực θ là tổng của các xác suất trả lời đúng các CH của đề TN tại giá trị θ

6. Các thang điểm được sử dụng ở Việt Nam

Thang điểm đang sử dụng phổ biến ở nước ta hiện nay là thang điểm 10, đó là điểm tuyệt đối dựa vào sự ấn định ngầm điểm tối đa là 10, điểm tối thiểu là 0 và điểm trung bình nằm ở khoảng 5. Để hỗ trợ cho quá trình tuyển chọn trong các kì thi với nhiều môn học, người ta ấn định hệ số dựa vào mức độ quan trọng của các môn học đó đối với việc học tập trong tương lai. Trong các kì thi tuyển sinh đại học, người ta cộng điểm thô của các môn lại sẽ thu được tổng điểm. Trên thực tế, sự phân bố điểm của các môn thường rất khác nhau, do đó giá trị trung bình và độ lệch tiêu chuẩn cũng hoàn toàn khác nhau.

Chúng ta có thể thiết lập thang điểm tiêu chuẩn Z với điểm trung bình bằng 5 và độ lệch tiêu chuẩn bằng 2, khi ấy khoảng $[-2,5\sigma; +2,5\sigma]$ sẽ ứng với khoảng điểm $[0; 10]$, gần với thang điểm 10 đang sử dụng phổ biến ở nước ta hiện nay.

Nếu thiết lập thang điểm tiêu chuẩn với điểm trung bình bằng 10 và độ lệch tiêu chuẩn là 4, thì khoảng $[-2,5\sigma; +2,5\sigma]$ gần với thang điểm 20.

Ví dụ: Đề kiểm tra học phần môn “*Phương pháp dạy học đại cương môn Toán*” ở Trường Cao đẳng Sư phạm Nghệ An, gồm 39 CH TN nhiều lựa chọn và 4 phương án trả lời, được thiết kế dựa trên các tiêu chí kĩ thuật và ma trận đề kiểm tra đã được xây dựng. Qua thử nghiệm đề kiểm tra tại Trường vào năm 2014 trên 30 sinh viên, kết quả được phân tích thông qua phần mềm Vitesta cho biết năng lực, số CH trả lời đúng trên tổng số CH (điểm thô), điểm thực của từng sinh viên, điểm Z tính theo công thức (1) và điểm V tính theo công thức (3). Dưới đây là kết quả của 10 sinh viên theo thứ tự từ 1-10 như *bảng 1* (kết quả điểm được tính theo Lí thuyết ứng

đáp câu hỏi), trong đó điểm thô trung bình của nhóm là 19, độ lệch tiêu chuẩn là 8.

Bảng. Kết quả của 10 sinh viên theo thứ tự từ 1-10

TT	Năng lực	Điểm thô	Điểm thực	Điểm Z	Điểm V (thang điểm 10)
1	1,05633	29/39	605,631	1,25	7,5
2	0,35955	24/39	535,95	0,625	6,25
3	-0,53504	17/39	446,50	-0,25	4,5
4	0,25710	24/39	525,71	0,625	6,25
5	2,82630	35/39	782,63	2	9
6	1,16520	28/39	616,52	1,125	7,25
7	1,30682	28/39	630,68	1,125	7,25
8	1,23957	30/39	623,96	1,375	7,75
9	0,34225	22/39	534,23	0,375	5,75
10	-0,12034	22/39	487,97	0,375	5,75

Kết quả điểm thực, điểm Z và điểm V như *bảng* đã phản ánh đúng năng lực của sinh viên. Như vậy, có thể vận dụng các công thức nêu trên để tính điểm và đánh giá kết quả học tập của HS.

* * *

Theo phương án mà Bộ GD-ĐT vừa công bố, kì thi Trung học phổ thông Quốc gia năm 2017 sẽ tổ chức gồm 5 bài thi ở các môn: *Toán, Ngữ văn, Ngoại ngữ* (3 môn bắt buộc), 2 bài thi tổ hợp Khoa học Tự nhiên (*Lí, Hóa, Sinh*) và Khoa học Xã hội. Mỗi thí sinh phải làm 4 bài thi, gồm: 3 bài bắt buộc *Toán, Ngữ văn, Ngoại ngữ* và một môn tự chọn là Khoa học tự nhiên hoặc Khoa học xã hội. Trong đó, môn *Ngữ văn* thi theo hình thức tự luận còn tất cả các môn khác thi theo hình thức TN khách quan. Do vậy, cần sử dụng một thang đo năng lực và loại điểm TN phù hợp để có thể định lượng được chính xác năng lực của người học trên thang đo ấy. □

Tài liệu tham khảo

[1] Dương Thiệu Tống (1995). *Trắc nghiệm và đo lường thành quả học tập*. NXB Khoa học xã hội.
 [2] Lâm Quang Thiệp (2011). *Đo lường trong giáo dục lí thuyết và ứng dụng*. NXB Đại học Quốc gia Hà Nội.
 [3] Harry Hsu. *Ngân hàng câu hỏi dựa trên thuyết ứng đáp câu hỏi*. Dự án Phát triển giáo dục trung học cơ sở.
 [4] Trần Bá Hoàn (1995). *Đánh giá trong giáo dục*. NXB Giáo dục.
 [5] Nguyễn Như Ý (2011). *Đại từ điển tiếng Việt*. NXB Đại học Quốc gia TP. Hồ Chí Minh.