

# SỬ DỤNG PHẦN MỀM IATA ĐỂ PHÂN TÍCH, ĐÁNH GIÁ VÀ NÂNG CAO CHẤT LƯỢNG CÂU HỎI TRẮC NGHIỆM KHÁCH QUAN TRONG CHƯƠNG HÀM SỐ LŨY THỪA, HÀM SỐ MŨ, HÀM SỐ LÔGARIT

Bùi Anh Kiệt<sup>1</sup> và Bùi Nguyên Phương<sup>2\*</sup>

<sup>1</sup>Khoa Sư phạm, Trường Đại học Cần Thơ

<sup>2</sup>Học viên cao học ngành Lý luận và Phương pháp dạy học Bộ môn Toán khóa 23, Trường Đại học Cần Thơ

\*Người chịu trách nhiệm về bài viết: Bùi Nguyên Phương (email: [buinguyenphuong1991@gmail.com](mailto:buinguyenphuong1991@gmail.com))

## Thông tin chung:

Ngày nhận bài: 12/03/2018

Ngày nhận bài sửa: 02/06/2018

Ngày duyệt đăng: 27/12/2018

## Title:

Using IATA to analyze, evaluate and improve the quality of the multiple-choice questions in chapter power functions, exponential functions and logarithmic functions

## Từ khóa:

Câu hỏi trắc nghiệm khách quan, IATA, IRT, lý thuyết ứng đáp câu hỏi

## Keywords:

IATA, IRT, Item Response Theory, multiple-choice questions

## ABSTRACT

This article presents an overview about Item Response Theory (IRT) as well as introduce how to install and use IATA – a software used to analyze and evaluate multiple-choice questions (MCQ) based on the IRT. In addition, the article propose a process for compiling and evaluating a MCQ test with the support of IATA. Finally, the article cover analyzing and evaluating multiple choice questions that were compiled in chapter power functions, exponential functions and logarithmic functions according to the specified process, in order to improve the quality of multiple-choice questions.

## TÓM TẮT

Bài viết này trình bày tổng quan về Lý thuyết ứng đáp câu hỏi (IRT) cũng như giới thiệu sơ lược về cách cài đặt và sử dụng phần mềm IATA – một phần mềm dùng để phân tích, đánh giá câu hỏi trắc nghiệm khách quan (TNKQ) dựa trên nền tảng IRT. Đồng thời, quy trình biên soạn và thẩm định một đề kiểm tra TNKQ dưới sự hỗ trợ của phần mềm IATA cũng được đề xuất trong bài viết. Cuối cùng, cách phân tích, đánh giá một số câu hỏi được biên soạn trong chương hàm số mũ, hàm số lũy thừa và hàm số lôgarit theo quy trình đã nêu cũng được trình bày, nhằm mục đích nâng cao chất lượng cho các câu hỏi trắc nghiệm khách quan.

Trích dẫn: Bùi Anh Kiệt và Bùi Nguyên Phương, 2018. Sử dụng phần mềm IATA để phân tích, đánh giá và nâng cao chất lượng câu hỏi trắc nghiệm khách quan trong chương hàm số lũy thừa, hàm số mũ, hàm số lôgarit. Tạp chí Khoa học Trường Đại học Cần Thơ. 54(9C): 81-93.

## 1 ĐẶT VẤN ĐỀ

Trắc nghiệm khách quan (TNKQ) là một hình thức kiểm tra đánh giá kết quả học tập được sử dụng ngày càng phổ biến những năm gần đây ở nước ta. Hơn thế nữa, kể từ năm 2017, trong kì thi Trung học phổ thông quốc gia (THPTQG), môn Toán đã thi theo hình thức TNKQ. Dù hình thức thi TNKQ có nhiều ưu điểm nhưng hạn chế lớn nhất của nó là không thể đánh giá được một cách chính xác khả năng lập luận, phán xét, tư duy của học sinh (HS). Do đó, vấn đề được đặt ra là làm sao chúng ta có thể

biên soạn và thẩm định được một câu hỏi, một đề kiểm tra (ĐKT) TNKQ nào đó có tốt hay không, có đảm bảo chất lượng và phù hợp với mục tiêu kiểm tra đánh giá hay không? Bài viết trình bày cách thẩm định và nâng cao chất lượng câu hỏi TNKQ dựa trên cơ sở của IRT dưới sự hỗ trợ của phần mềm IATA.

## 2 SƠ LƯỢC VỀ IRT

Trong nửa cuối thế kỉ XX nhiều nhà nghiên cứu về đo lường trong tâm lý và giáo dục đã cố gắng tìm giải pháp nhằm nâng cao độ chính xác và tính khách quan của các phép đo dùng TNKQ. Trong một bài

kiểm tra hoặc một bài thi, việc HS trả lời câu hỏi đúng hay sai phụ thuộc vào năng lực của HS đó trong phạm vi được kiểm tra và độ khó của từng câu hỏi cụ thể. Lý thuyết ứng đáp câu hỏi (Item Response Theory – IRT) là một phương pháp được dùng để tính xác suất một HS trả lời đúng một câu hỏi thông qua mối tương quan giữa năng lực HS và các tham số đặc trưng của câu hỏi. Ở nước ngoài, có nhiều học giả đã nghiên cứu chuyên sâu về IRT như Linda and James (2008) đã trình bày về các lý thuyết đo lường, độ tin cậy, độ giá trị, IRT,... Thomas (2004) nghiên cứu về TNKQ và ứng dụng IRT vào TNKQ, Ayala (2009) cũng trình bày một cách tương đối chi tiết về IRT (có sự so sánh với lý thuyết khảo thí cổ điển) và ứng dụng vào phân tích các câu hỏi TNKQ... Ở Việt Nam, Lâm Quang Thiệp (2008), (2012), Dương Thiệu Tống (2005), ... cũng có nhiều công trình nghiên cứu về ứng dụng IRT trong kiểm tra, đánh giá giáo dục.

Có hai mô hình IRT thường được sử dụng là mô hình một tham số và mô hình hai tham số.

### 2.1 Mô hình irt một tham số

Theo Leigh (2010), mô hình IRT đơn giản nhất là mô hình một tham số hay còn gọi là mô hình Rasch. Nhà toán học Đan Mạch, Georg Rasch, đã đưa ra một mô hình “ứng đáp CH” để mô tả mối tương tác nguyên tố giữa một HS với một câu hỏi của đề trắc nghiệm (ĐTN), và dùng mô hình đó để phân tích các dữ liệu thu được từ ĐTN. Theo mô hình Rasch, xác suất để HS  $i$  trả lời đúng câu hỏi  $j$  là

$$P_j(\theta_i) = \frac{e^{\theta_i - b_j}}{1 + e^{\theta_i - b_j}}, \text{ trong đó } \theta_i \text{ là năng lực của HS}$$

$i$  và  $b_j$  là độ khó của câu hỏi  $j$ . Giá trị của  $\theta$  tỉ lệ thuận với năng lực của HS và giá trị của  $b$  càng lớn khi và chỉ khi câu hỏi càng khó. Một HS muốn có xác suất trả lời đúng một câu hỏi là 0,5 thì HS đó phải có năng lực bằng với độ khó của câu hỏi đó, cụ thể là  $\theta_i = b_j$ . Năng lực của HS thường được giả sử là một phân phối chuẩn với kì vọng là 0 và phương sai là 1. Tập giá trị thường gặp của năng lực  $\theta$  là từ -3 đến 3 hoặc từ -4 đến 4. Các giá trị bên ngoài khoảng này cho thấy câu hỏi có vấn đề.

### 2.2 Mô hình irt hai tham số

Bên cạnh độ khó, mô hình có thể đưa thêm một tham số thứ hai liên quan đến độ phân biệt của câu hỏi là  $a_j$ , và ta thường gọi nó là mô hình hai tham số. Theo mô hình đó, xác suất để HS  $i$  trả lời đúng

$$\text{câu hỏi } j \text{ là } P_j(\theta_i) = \frac{e^{a_j(\theta_i - b_j)}}{1 + e^{a_j(\theta_i - b_j)}}, \text{ trong đó } \theta_i,$$

$b_j$  có ý nghĩa như trong mô hình một tham số và  $a_j$  là độ phân biệt của câu hỏi thứ  $j$  (Leigh, 2010).

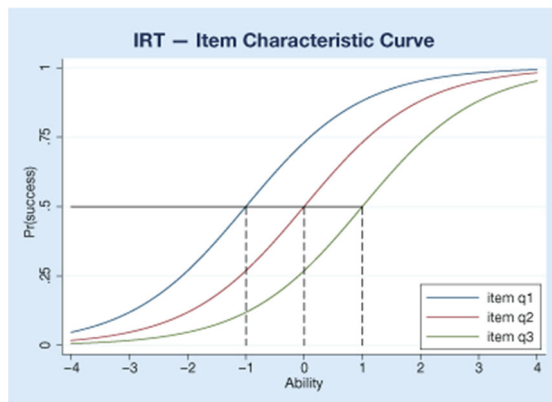
Giá trị của  $a_j$  thường nằm trong khoảng từ 0 đến 2.

Khi độ phân biệt có giá trị càng lớn thì một sự khác biệt rất nhỏ về năng lực cũng sẽ gây ra một độ chênh lệch lớn về xác suất trả lời đúng. Dễ dàng nhận thấy rằng nếu  $a_j = 1$  với mọi  $j$  thì ta thu được mô hình

Rasch. Tương tự như độ khó, một câu hỏi có độ phân biệt bằng các giá trị biên hoặc vượt khỏi khoảng từ 0 đến 2 cho thấy câu hỏi có thể có vấn đề về nội dung hoặc cách diễn đạt. Đặc biệt, độ phân biệt âm ám chỉ rằng khả năng trả lời đúng của những HS có năng lực cao lại nhỏ hơn so với những HS có năng lực thấp. Những câu hỏi này cần được xem xét lại để đảm bảo rằng chúng phù hợp với nội dung mà ĐKT đang kiểm tra.

### 2.3 Đường cong đặc trưng câu hỏi

Nếu cho giá trị của  $\theta$  biến đổi liên tục trên trục số (giả sử từ -4 đến 4) thì chúng ta thu được một đồ thị được gọi là **đường cong đặc trưng câu hỏi (item characteristic curve – ICC)** có dạng như sau:



**Hình 1: Đường cong đặc trưng câu hỏi theo mô hình Rasch**

Các tính chất của một câu hỏi có thể tổng kết được bằng cách nhìn vào ICC. ICC thể hiện xác suất trả lời đúng một câu hỏi ứng với các giá trị khác nhau của  $\theta$  và nó là một đường cong logistic có biên từ 0 đến 1. Độ khó một câu hỏi chính là tọa độ điểm uốn của ICC theo thang năng lực  $\theta$ . Theo Lâm Quang Thiệp (2008), nếu kẻ một đường thẳng song song với trục hoành và cắt trục tung tại vị trí  $P=0,5$  thì đường thẳng này sẽ cắt các ICC tại các điểm có hoành độ bằng độ khó của câu hỏi tương ứng, vì điểm của đường cong ứng với tung độ  $P=0,5$  sẽ có hoành độ  $\theta=b$ . Các đường cong càng dịch chuyển về phía bên trái (phải) biểu thị cho câu hỏi càng dễ (khó). Ngoài ra, độ phân biệt mô tả độ dốc của ICC tại điểm có hoành độ  $\theta=b$  và tung độ

$P(\theta)=0,5$ . Một ICC có độ dốc (hệ số góc) càng lớn cho thấy câu hỏi đó có khả năng phân biệt năng lực HS càng tốt. Một ICC có hệ số góc âm khi và chỉ khi độ phân biệt âm. Các ICC càng gần với đồ thị hàm bậc thang biểu thị cho câu hỏi có độ phân biệt đặc biệt lớn.

### 3 PHẦN MỀM IATA

IATA (Item and Test Analysis) là một phần mềm (được cung cấp miễn phí bởi Fernando Cartwright) dùng để phân tích dữ liệu trong đánh giá giáo dục và tâm lý học. Nó thực hiện phân tích các câu hỏi TNKQ cũng như toàn bộ bài kiểm tra dựa trên IRT, từ đó giúp xây dựng được các câu hỏi TNKQ có chất lượng phù hợp với năng lực người học và mục đích kiểm tra đánh giá.

#### 3.1 Hướng dẫn cài đặt và sử dụng phần mềm iata

Phần mềm IATA có thể cài đặt và sử dụng theo các bước sau:

**Bước 1:** Truy cập vào địa chỉ:

<https://polymetrika.com/Downloads/Downloader?content=IATAsetup.exe> để tải file cài đặt phần mềm (IATAsetup.exe – 13,97Mb).

**Bước 2:** Tiến hành cài đặt ứng dụng bình thường (Next → I accept the agreement → Next → Chọn đường dẫn thư mục → Next → Next → Next → Install → Finish)



**Hình 2:** Giao diện khi khởi động của phần mềm IATA

**Bảng 1:** Dữ liệu câu trả lời của học sinh

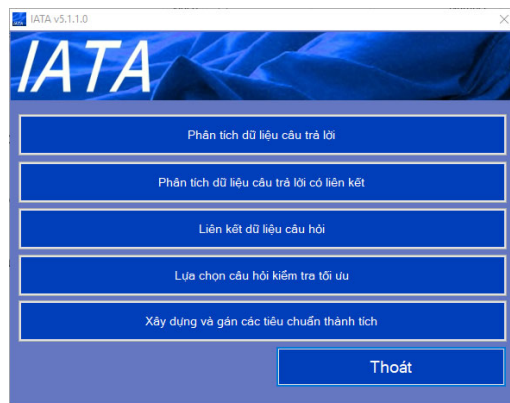
| Ho              | Ten | SBD   | Diem | Cau 1 | Cau 2 | Cau 3 |
|-----------------|-----|-------|------|-------|-------|-------|
| Nghiem Hoai     | An  | 10112 | 2.5  | A     | B     | D     |
| Nguyen Thi Thuy | An  | 10001 | 7.75 | A     | B     | B     |

Nhấn vào nút “Mở tập tin” và chọn đường dẫn thư mục chứa tập tin *Dữ liệu câu trả lời của học sinh* (có thể đặt tên khác cho tập tin này). Sau đó nhấn OK → “Tiếp tục”.

**Bước 3:** Phần mềm IATA được tự động khởi động sau khi cài đặt hoặc người dùng có thể tự khởi động phần mềm. Giao diện khi khởi động của phần mềm như Hình 2.

Để sử dụng phần mềm, người dùng không bắt buộc phải đăng ký tài khoản. Họ có thể thay đổi ngôn ngữ sang Tiếng Việt tại bảng chọn “Select language for IATA”. Để bắt đầu phân tích dữ liệu, người dùng nhấn vào Main Menu (Menu chính).

**Bước 4:** Menu chính của phần mềm gồm các lựa chọn như sau:



**Hình 3:** Menu chính của phần mềm IATA

Sau khi cài đặt phần mềm, trên màn hình Desktop sẽ tự động xuất hiện một thư mục có tên là IATA, trong đó có chứa các tập tin mẫu về Dữ liệu trả lời của HS, dữ liệu câu hỏi... để chạy thử phần mềm. Người dùng có thể tải về hướng dẫn chi tiết (bằng tiếng Anh) cách sử dụng phần mềm IATA tại địa chỉ <https://polymetrika.com/Downloads/Downloader?content=IATAMannual.pdf>

#### 3.2 Hướng dẫn phân tích dữ liệu câu trả lời

**Bước 1:** Tại Menu chính của phần mềm, ta chọn mục Phân tích dữ liệu câu trả lời.

**Bước 2:** Chúng ta nhập tập tin *Dữ liệu câu trả lời của học sinh*, tập tin này chứa các câu trả lời của từng HS cho từng câu hỏi. Tập tin này có thể được nhập bằng Excel hoặc chỉnh sửa theo tập tin mẫu trong thư mục IATA, ví dụ như sau:

**Bước 3:** Người dùng nhập tập tin *Dữ liệu câu hỏi*, tập tin này chứa đáp án của từng câu hỏi, mức độ nhận thức, nội dung kiểm tra. Tập tin này có thể

được nhập bằng Excel hoặc chỉnh sửa theo tập tin mẫu trong thư mục IATA, ví dụ như sau:

**Bảng 2: Dữ liệu câu hỏi**

| Name  | Key | Level | Content |
|-------|-----|-------|---------|
| Cau 1 | A   | 1     | CD1     |
| Cau 2 | B   | 2     | CD2     |
| Cau 3 | D   | 3     | CD2     |

Nhấn vào nút “Mở tập tin” và chọn đường dẫn thư mục chứa tập tin *Dữ liệu câu hỏi* (có thể đặt tên khác cho tập tin này). Sau đó nhấn OK → “Tiếp tục”.

**Bước 4:** Mục này nhập các thông số dùng để phân tích cũng như cách chấm điểm cho từng câu hỏi, cách xử lý với những câu hỏi lỗi (không có đáp án, chọn hai đáp án, chọn đáp án khác...). Ta có thể nhấn “Tiếp tục” để bỏ qua bước này nếu điểm số cho từng câu hỏi là bằng nhau và tất cả các đáp án không đúng quy định đều được chấm là sai. Các trang còn lại là kết quả phân tích câu hỏi và bài kiểm tra, tùy vào mục đích nghiên cứu mà ta sẽ xem xét kỹ những trang tương ứng.

#### 4 NHỮNG ƯU ĐIỂM CỦA PHẦN MỀM IATA SO VỚI MỘT SỐ PHẦN MỀM KHÁC

Gần đây, trên thế giới cũng như ở Việt Nam, các nhà nghiên cứu sử dụng nhiều phần mềm thống kê khác nhau để phân tích các câu hỏi TNKQ; hầu hết các phần mềm này đều dựa trên IRT. Chẳng hạn, Sudol và Studer (2010) sử dụng phần mềm R để phân tích câu hỏi TNKQ, Ayala (2009) dùng phần mềm Multiblog, Trần Thị Anh Đào (2008) dùng phần mềm Quest, Đặng Thị Hương (2012) và Lâm Quang Thiệp (2008) dùng phần mềm Vitesta, Lê Ngọc và Đào Thị Trang (2017) dùng phần mềm IATA... Phần mềm IATA có một số ưu điểm vượt trội hơn so với các phần mềm phân tích ĐKT TNKQ khác như sau:

- Phần mềm IATA được cung cấp hoàn toàn miễn phí; cho nên rất tiện lợi trong việc phổ biến cho các giáo viên (GV) và không lo ngại vấn đề bản quyền của phần mềm. Phần mềm R cũng là phần mềm miễn phí, tuy nhiên sử dụng phần mềm R đòi hỏi thông qua một số câu lệnh; điều này sẽ gây khó khăn cho nhiều giáo viên phổ thông.

- Ngôn ngữ hiển thị là Tiếng Việt (phần mềm Quest sử dụng ngôn ngữ là Tiếng Anh) nên dễ sử dụng.

- Các file dữ liệu đầu vào có thể được nhập bằng file Excel, gửi, dễ sử dụng đối với GV (phần mềm Vitesta yêu cầu GV nhập thủ công file dữ liệu từ file text).

- Có giao diện trực quan, sử dụng bằng bảng chọn và chuột nên dễ sử dụng (phần mềm Quest yêu cầu người dùng phải viết các câu lệnh).

- Phần mềm IATA có đầy đủ các tính năng cần thiết của một phần mềm thống kê trong việc phân tích ĐKT TNKQ; chẳng hạn như độ khó, độ phân biệt của mỗi câu hỏi... và đặc biệt là phần mềm IATA cho phép ước lượng được năng lực thực sự của mỗi học sinh; từ đó có thể đánh giá học sinh một cách chính xác hơn dựa vào năng lực thực sự của họ.

Từ những ưu điểm trên, phần mềm IATA được lựa chọn để giới thiệu cho GV sử dụng trong việc phân tích các câu hỏi TNKQ của một ĐKT nhằm nâng cao chất lượng của các câu hỏi; đồng thời cũng giúp GV có công cụ để đánh giá một cách chính xác năng lực thực sự của học sinh.

#### 5 QUY TRÌNH BIÊN SOẠN VÀ THẨM ĐỊNH MỘT ĐỀ KIỂM TRA TNKQ

Qua nghiên cứu các tài liệu tập huấn Kỹ thuật viết câu hỏi TNKQ, tài liệu tập huấn Dạy học và kiểm tra, đánh giá kết quả học tập của HS theo định hướng phát triển năng lực của HS..., bài viết trình bày việc biên soạn, thẩm định một đề kiểm tra TNKQ chương Hàm số mũ, hàm số lũy thừa và hàm số lôgarit lớp 12 theo một quy trình như sau:

##### **Bước 1: Xác định mục đích và hình thức của đề kiểm tra (ĐKT)**

ĐKT được xây dựng nhằm mục đích kiểm tra hai năng lực chính của HS là năng lực giải quyết vấn đề và năng lực mô hình hóa theo hình thức trắc nghiệm.

##### **Bước 2: Thiết lập ma trận ĐKT**

Dựa vào chuẩn kiến thức và kỹ năng, phân phối chương trình, thời gian làm bài kiểm tra, số lượng câu hỏi, mục đích kiểm tra đánh giá... mà GV thiết lập một ma trận đề cho phù hợp.

##### **Bước 3: Biên soạn câu hỏi theo ma trận ĐKT**

Dựa vào ma trận ĐKT, GV biên soạn lời dẫn cho các câu hỏi, đáp án đúng và các đáp án nhiễu. Giải thích lý do xây dựng đáp án nhiễu. Các câu hỏi cần đảm bảo nội dung, hình thức và mục đích kiểm tra đánh giá.

##### **Bước 4: Xây dựng bảng đáp án và thang điểm**

Dựa vào ĐKT đã biên soạn, GV xây dựng bảng đáp án và thang điểm cho từng câu hỏi (nếu cần thiết).

##### **Bước 5: Kiểm tra, chỉnh sửa ĐKT trước khi thẩm định**

Rà soát, kiểm tra, thẩm định nội dung, từ ngữ, ký hiệu, hình thức của từng câu hỏi và bảng đáp án trước khi cho HS làm kiểm tra.



**Bước 6: Thẩm định, phân tích, đánh giá, chỉnh sửa ĐKT lần 1**

Dựa vào kết quả phân tích từng câu hỏi và ĐKT từ phần mềm IATA, tiến hành nghiên cứu, đánh giá từng câu hỏi, thực hiện chỉnh sửa hoặc thay thế những câu hỏi có vấn đề hoặc chưa phù hợp.

**Bước 7: Thẩm định, phân tích, đánh giá, chỉnh sửa ĐKT lần 2**

Dựa vào kết quả thẩm định lần 2, tiến hành chỉnh sửa lần 2 (nếu có). Sau đó rà soát, lựa chọn các câu hỏi phù hợp để sử dụng hoặc làm ngân hàng đề thi.

**6 DÙNG PHẦN MỀM IATA ĐỂ PHÂN TÍCH, ĐÁNH GIÁ VÀ NÂNG CAO CHẤT LƯỢNG CÂU HỎI TNKQ CHƯƠNG HÀM SỐ LŨY THỪA, HÀM SỐ MŨ, HÀM SỐ LÔGARIT**

Bài viết chỉ giới thiệu cách phân tích, đánh giá một số câu hỏi TNKQ được biên soạn và thẩm định từ một ĐKT cuối chương Hàm số lũy thừa, hàm số mũ, hàm số lôgarit. ĐKT này gồm 40 câu hỏi TNKQ và thời gian làm bài là 75 phút. Theo Bộ Giáo dục và Đào tạo (2010), ma trận chi tiết mô tả chuẩn kiến thức, kĩ năng, mục tiêu kiểm tra của hai đề TNKQ được trình bày như sau:

**Bảng 3: Ma trận chuẩn kiến thức, kĩ năng của 2 đề kiểm tra TNKQ**

| Chủ đề  | Nhận thức<br>Nhận biết   | Thông hiểu  | Vận dụng thấp   | Vận dụng cao  | Cộng           |
|---|--|---|---|---|----------------|
| <b>Chủ đề 1<br/>(4 tiết)</b><br>Lũy thừa, hàm số thừa, hàm số lũy thừa, hàm số lũy thừa, lôgarit.<br>Tỉ lệ: 20% | – Biết các khái niệm và tính chất của lũy thừa, hàm số thừa, hàm số lũy thừa và lôgarit.<br>– Biết các khái niệm biểu thức, so sánh những lôgarit thập phân và biểu thức có chứa lôgarit tự nhiên.<br>– Biết công thức tính đạo hàm của hàm số lũy thừa. | – Biết dạng đồ thị của các hàm số lũy thừa, hàm số lôgarit.<br>– Biết dùng tính chất của lũy thừa để đơn giản thức chứa lôgarit phức tạp.<br>– Biết dùng định nghĩa có chứa lũy thừa phức tạp.<br>– Biết tính đạo hàm của hàm số chứa lôgarit đơn giản. | – Biết vận dụng các tính chất của lôgarit vào các bài tập của lôgarit để đơn giản biểu thức, so sánh những biểu thức.<br>– Biết vận dụng các tính chất của lũy thừa để đơn giản biểu thức, so sánh những biểu thức.<br>– Biết tính đạo hàm của hàm số lũy thừa dưới dạng hàm hợp.   |   | 8 câu<br>20%   |
| <b>Chủ đề 2<br/>(3 tiết)</b><br>Hàm số mũ, hàm số lôgarit.<br>Tỉ lệ: 15%  | – Biết khái niệm và tính chất của hàm số mũ, hàm số lôgarit.<br>– Biết công thức tính đạo hàm của các hàm số mũ, hàm số lôgarit.<br>– Biết tập xác định của một hàm số lôgarit.  | – Biết dạng đồ thị của các hàm số mũ, hàm số lôgarit.<br>– Biết dùng tính chất của các hàm số mũ, hàm số lôgarit vào việc so sánh hai số, hai biểu thức chứa mũ và lôgarit.<br>– Tính được đạo hàm các hàm số mũ và hàm số lôgarit.                     | – Biết vận dụng khái niệm, tính chất của hàm số mũ, hàm số lôgarit để giải quyết một số bài toán thực tiễn (đã được giới thiệu trong SGK)<br>– Biết vận dụng tính chất của các hàm lôgarit để tìm tập xác định của một số hàm số lôgarit phức tạp.<br>– Tính được đạo hàm của các hàm số mũ, hàm số lôgarit dưới dạng hàm hợp phức tạp. |   | 6 câu<br>15%   |
| <b>Chủ đề 3<br/>(7 tiết)</b><br>Phương trình mũ, phương trình lôgarit<br>Tỉ lệ: 35%                             | – Nhận biết được phương trình mũ và phương trình lôgarit.<br>– Biết các cách giải các phương trình mũ, phương trình lôgarit cơ bản.  | – Giải được các phương trình mũ, phương trình lôgarit đơn giản bằng các phương pháp: đưa về cùng cơ số, đặt ẩn phụ, lôgarit hóa và mũ hóa.  | – Biết vận dụng tổng hợp các phương pháp để giải những phương trình mũ, phương trình lôgarit phức tạp.  | – Biết vận dụng phương trình mũ, phương trình lôgarit để giải quyết các vấn đề trong thực tiễn.         | 14 câu<br>35%  |
| <b>Chủ đề 4<br/>(6 tiết)</b><br>Bất phương trình mũ, bất phương trình lôgarit<br>Tỉ lệ: 30%                     | – Nhận biết được bất phương trình mũ và bất phương trình lôgarit.<br>– Biết các cách giải các bất phương trình mũ, bất phương trình lôgarit cơ bản.  | – Giải được các bất phương trình mũ, bất phương trình lôgarit đơn giản bằng các phương pháp: đưa về cùng cơ số, đặt ẩn phụ, lôgarit hóa và mũ hóa.  | – Biết vận dụng tổng hợp các phương pháp để giải những bất phương trình mũ, bất phương trình lôgarit phức tạp.  | – Biết vận dụng bất phương trình mũ, bất phương trình lôgarit để giải quyết các vấn đề trong thực tiễn. | 12 câu<br>30%  |
| Tổng<br>Tỉ lệ   | 8 câu<br>20%   | 12 câu<br>30%   | 16 câu<br>40%   | 4 câu<br>10%  | 40 câu<br>100% |

| Lần 1                               |   |        |       |      | Lần 2                               |   |        |       |      |
|-------------------------------------|---|--------|-------|------|-------------------------------------|---|--------|-------|------|
| Use                                 | O | Name   | Discr | PVal | Use                                 | O | Name   | Discr | PVal |
| <input checked="" type="checkbox"/> |   | Cau 1  | 0.05  | 0.97 | <input checked="" type="checkbox"/> |   | Cau 1  | 0.14  | 0.92 |
| <input checked="" type="checkbox"/> |   | Cau 2  | 0.02  | 0.95 | <input checked="" type="checkbox"/> |   | Cau 2  | 0.51  | 0.73 |
| <input checked="" type="checkbox"/> |   | Cau 3  | 0.10  | 0.94 | <input checked="" type="checkbox"/> |   | Cau 3  | 0.40  | 0.67 |
| <input checked="" type="checkbox"/> |   | Cau 4  | 0.62  | 0.55 | <input checked="" type="checkbox"/> |   | Cau 4  | 0.46  | 0.71 |
| <input checked="" type="checkbox"/> |   | Cau 5  | 0.71  | 0.45 | <input checked="" type="checkbox"/> |   | Cau 5  | 0.57  | 0.69 |
| <input checked="" type="checkbox"/> |   | Cau 6  | 0.22  | 0.88 | <input checked="" type="checkbox"/> |   | Cau 6  | 0.59  | 0.56 |
| <input checked="" type="checkbox"/> |   | Cau 7  | 0.58  | 0.35 | <input checked="" type="checkbox"/> |   | Cau 7  | 0.67  | 0.35 |
| <input checked="" type="checkbox"/> |   | Cau 8  | 0.39  | 0.68 | <input checked="" type="checkbox"/> |   | Cau 8  | 0.43  | 0.70 |
| <input checked="" type="checkbox"/> |   | Cau 9  | 0.74  | 0.43 | <input checked="" type="checkbox"/> |   | Cau 9  | 0.65  | 0.55 |
| <input checked="" type="checkbox"/> |   | Cau 10 | 0.41  | 0.72 | <input checked="" type="checkbox"/> |   | Cau 10 | 0.27  | 0.69 |
| <input checked="" type="checkbox"/> |   | Cau 11 | 0.38  | 0.77 | <input checked="" type="checkbox"/> |   | Cau 11 | 0.48  | 0.67 |
| <input checked="" type="checkbox"/> |   | Cau 12 | 0.51  | 0.68 | <input checked="" type="checkbox"/> |   | Cau 12 | 0.62  | 0.70 |
| <input checked="" type="checkbox"/> |   | Cau 13 | 0.56  | 0.55 | <input checked="" type="checkbox"/> |   | Cau 13 | 0.51  | 0.66 |
| <input checked="" type="checkbox"/> |   | Cau 14 | 0.67  | 0.50 | <input checked="" type="checkbox"/> |   | Cau 14 | 0.51  | 0.60 |
| <input checked="" type="checkbox"/> |   | Cau 15 | 0.72  | 0.55 | <input checked="" type="checkbox"/> |   | Cau 15 | 0.59  | 0.57 |
| <input checked="" type="checkbox"/> |   | Cau 16 | 0.79  | 0.35 | <input checked="" type="checkbox"/> |   | Cau 16 | 0.56  | 0.56 |
| <input checked="" type="checkbox"/> |   | Cau 17 | 0.22  | 0.12 | <input checked="" type="checkbox"/> |   | Cau 17 | 0.34  | 0.27 |
| <input checked="" type="checkbox"/> |   | Cau 18 | 0.56  | 0.63 | <input checked="" type="checkbox"/> |   | Cau 18 | 0.56  | 0.59 |
| <input checked="" type="checkbox"/> |   | Cau 19 | 0.44  | 0.40 | <input checked="" type="checkbox"/> |   | Cau 19 | 0.64  | 0.37 |
| <input checked="" type="checkbox"/> |   | Cau 20 | 0.39  | 0.31 | <input checked="" type="checkbox"/> |   | Cau 20 | 0.42  | 0.37 |
| <input checked="" type="checkbox"/> |   | Cau 21 | 0.02  | 0.86 | <input checked="" type="checkbox"/> |   | Cau 21 | 0.54  | 0.61 |
| <input checked="" type="checkbox"/> |   | Cau 22 | 0.12  | 0.90 | <input checked="" type="checkbox"/> |   | Cau 22 | 0.75  | 0.47 |
| <input checked="" type="checkbox"/> |   | Cau 23 | 0.34  | 0.74 | <input checked="" type="checkbox"/> |   | Cau 23 | 0.54  | 0.57 |
| <input checked="" type="checkbox"/> |   | Cau 24 | 0.38  | 0.22 | <input checked="" type="checkbox"/> |   | Cau 24 | 0.59  | 0.58 |
| <input checked="" type="checkbox"/> |   | Cau 25 | 0.37  | 0.47 | <input checked="" type="checkbox"/> |   | Cau 25 | 0.81  | 0.60 |
| <input checked="" type="checkbox"/> |   | Cau 26 | 0.43  | 0.81 | <input checked="" type="checkbox"/> |   | Cau 26 | 0.54  | 0.44 |
| <input checked="" type="checkbox"/> |   | Cau 27 | 0.42  | 0.71 | <input checked="" type="checkbox"/> |   | Cau 27 | 0.65  | 0.49 |
| <input checked="" type="checkbox"/> |   | Cau 28 | 0.48  | 0.76 | <input checked="" type="checkbox"/> |   | Cau 28 | 0.64  | 0.45 |
| <input checked="" type="checkbox"/> |   | Cau 29 | 0.60  | 0.37 | <input checked="" type="checkbox"/> |   | Cau 29 | 0.78  | 0.40 |
| <input checked="" type="checkbox"/> |   | Cau 30 | 0.77  | 0.53 | <input checked="" type="checkbox"/> |   | Cau 30 | 0.57  | 0.53 |
| <input checked="" type="checkbox"/> |   | Cau 31 | 0.72  | 0.47 | <input checked="" type="checkbox"/> |   | Cau 31 | 0.64  | 0.36 |
| <input checked="" type="checkbox"/> |   | Cau 32 | -0.17 | 0.17 | <input checked="" type="checkbox"/> |   | Cau 32 | 0.43  | 0.41 |
| <input checked="" type="checkbox"/> |   | Cau 33 | 0.64  | 0.60 | <input checked="" type="checkbox"/> |   | Cau 33 | 0.75  | 0.47 |
| <input checked="" type="checkbox"/> |   | Cau 34 | 0.72  | 0.54 | <input checked="" type="checkbox"/> |   | Cau 34 | 0.70  | 0.46 |
| <input checked="" type="checkbox"/> |   | Cau 35 | 0.45  | 0.36 | <input checked="" type="checkbox"/> |   | Cau 35 | 0.51  | 0.40 |
| <input checked="" type="checkbox"/> |   | Cau 36 | 0.57  | 0.54 | <input checked="" type="checkbox"/> |   | Cau 36 | 0.48  | 0.44 |
| <input checked="" type="checkbox"/> |   | Cau 37 | -0.18 | 0.17 | <input checked="" type="checkbox"/> |   | Cau 37 | 0.45  | 0.40 |
| <input checked="" type="checkbox"/> |   | Cau 38 | 0.46  | 0.27 | <input checked="" type="checkbox"/> |   | Cau 38 | 0.45  | 0.36 |
| <input checked="" type="checkbox"/> |   | Cau 39 | 0.41  | 0.26 | <input checked="" type="checkbox"/> |   | Cau 39 | 0.50  | 0.34 |
| <input checked="" type="checkbox"/> |   | Cau 40 | 0.41  | 0.23 | <input checked="" type="checkbox"/> |   | Cau 40 | 0.45  | 0.34 |

Hình 4: Kết quả phân tích câu hỏi sau 2 lần thẩm định

Theo Linacre (1994), để việc ước lượng các tham số trong IRT chính xác với độ tin cậy 95% thì cỡ mẫu tối thiểu là 100 và tối thiểu là 150 nếu độ tin cậy là 99%. Nghiên cứu sử dụng độ tin cậy 95%; nên cỡ mẫu tối thiểu cần thiết là 100. Ayala (2009) đề xuất cỡ mẫu tối thiểu cần thiết khi phân tích các câu hỏi TNKQ phải gấp 3 lần số câu hỏi. Vì nghiên cứu sử dụng để kiểm tra có 40 câu hỏi nên cỡ mẫu ít nhất phải là 120.

Lần thẩm định đầu tiên được thực hiện trên 139 HS đang học lớp 12 thuộc các trường THPT chuyên Nguyễn Đình Chiểu – Đồng Tháp, THPT Phú Quốc – Kiên Giang, THPT Bình Thủy – Cần Thơ. Các HS thực hiện bài kiểm tra trong thời gian 75 phút, sau khi học xong chương Hàm số lũy thừa, hàm số mũ và hàm số lôgarit. Phiếu trả lời trắc nghiệm của HS được thu về xử lý dữ liệu bài làm.

Lần thẩm định thứ hai được thực hiện trên 131 HS đang học lớp 12 thuộc các trường THPT chuyên Huỳnh Mẫn Đạt – Kiên Giang, THCS và THPT Lạc Hồng – TP Hồ Chí Minh. Lần thẩm định thứ hai sau lần thẩm định đầu tiên 3 ngày. Các HS thực hiện bài

kiểm tra trong thời gian 75 phút, sau khi học xong chương Hàm số lũy thừa, hàm số mũ và hàm số lôgarit. Phiếu trả lời trắc nghiệm của HS được thu về xử lý dữ liệu bài làm. Các tập tin *Dữ liệu câu trả lời của học sinh* và *Dữ liệu câu hỏi* được xử lý thông qua phần mềm Excel theo đúng cấu trúc mà phần mềm IATA yêu cầu.

Kết quả tổng quát phân tích các câu hỏi trong ĐKT ở 2 lần như Hình 4.

Sau khi thẩm định đề lần 1, có 7 câu có dấu báo hiệu màu đỏ (hình tam giác), phân tích câu 2:

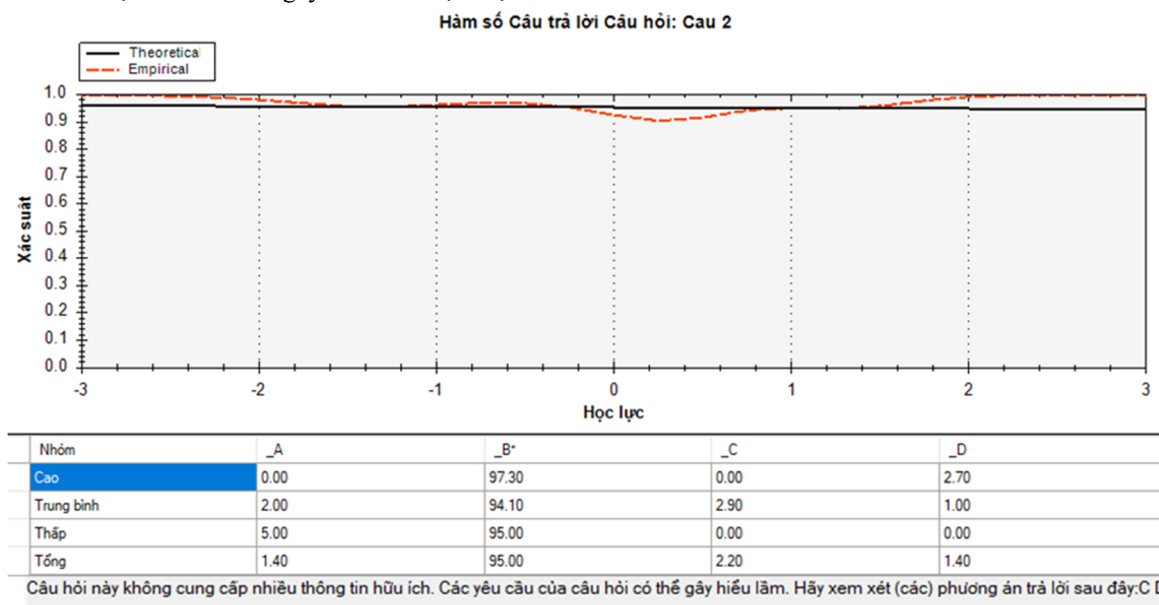
**Câu 2.** Cho hai số dương  $a$  và  $b$ ,  $a \neq 1$ ,  $\alpha$  là số thực bất kì.

Tính chất nào dưới đây **sai**?

A.  $\log_a a = 1$ . B.  $\log_a 1 = a$ .

C.  $\log_a (a^\alpha) = \alpha$ . D.  $a^{\log_a b} = b$ .

Kết quả phân tích chi tiết câu 2 như sau:



**Hình 5: Kết quả phân tích câu hỏi Câu 2 sau lần thẩm định đầu tiên**

Dấu \* được đánh dấu phía sau chữ cái đại diện cho đáp án đúng. Từ số liệu thống kê ta thấy Câu 2 có độ khó (Pval) là 0,95 và độ phân biệt (Discr) là 0,02. Theo IRT thì câu hỏi này quá dễ và độ phân biệt kém. Các đáp án nhiễu C và D cần xem xét lại vì nhóm HS điểm thấp không có ai lựa chọn nhưng nhóm HS điểm cao và trung bình lại lựa chọn. Do đó, câu 2 đã được thay đổi như sau:

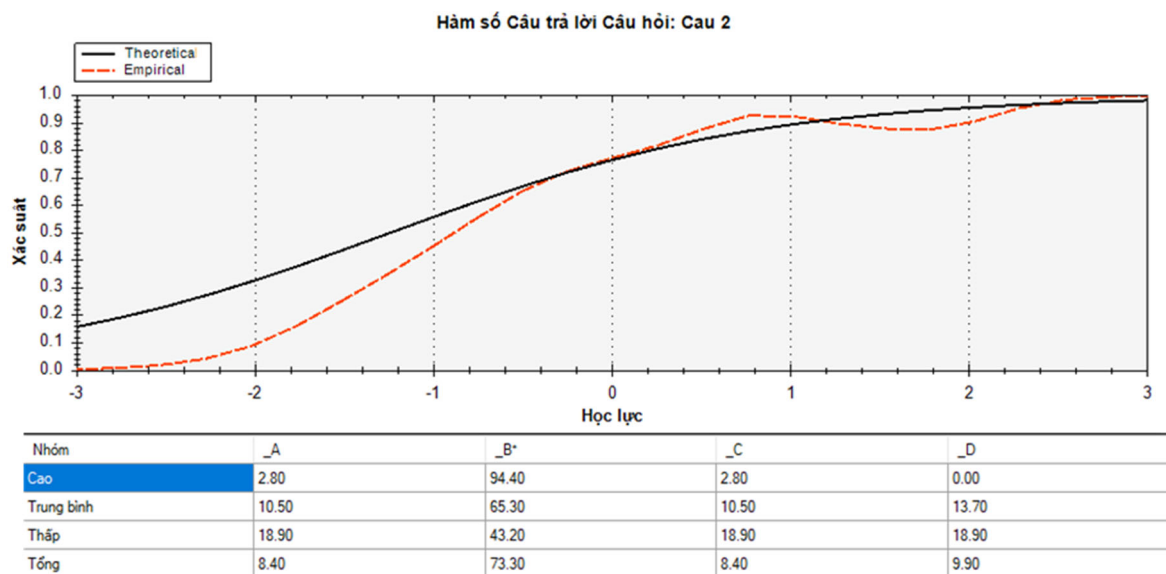
**Câu 2.** Cho ba số dương  $a$ ,  $b$ ,  $c$  với  $a \neq 1$ ,  $b \neq 1$  và  $c \neq 1$ .

Mệnh đề nào dưới đây đúng?

A.  $\log_a b = \frac{\log_b c}{\log_a c}$ . B.  $\log_a b = \frac{\log_c b}{\log_c a}$ .

C.  $\frac{\log_a c}{\log_b c} = \log \frac{a}{b} c$ . D.  $\frac{\log_c a}{\log_c b} = \log_c (a-b)$ .

Kết quả phân tích chi tiết câu 2 lần thứ 2 như sau:



**Hình 6: Kết quả phân tích câu hỏi Câu 2 sau lần thẩm định thứ hai**

Từ số liệu thu được, sau khi thay đổi câu hỏi thì độ khó của câu hỏi mới là 0,73 và độ phân biệt là 0,51. Theo IRT câu hỏi này ở mức độ dễ và độ phân biệt khá tốt. Các đáp án nhiều đều có HS lựa chọn và tỉ lệ HS theo năng lực khi chọn đáp án nhiều hợp lí.

Chuẩn kiến thức muốn kiểm tra HS ở câu hỏi này đều là khả năng ghi nhớ, nhận biết các công thức cơ bản của lôgarit được trình bày trong sách giáo khoa. Tuy nhiên từ kết quả phân tích qua hai lần thẩm định, ta thấy công thức được kiểm tra ở lần thẩm định đầu tiên HS nắm vững, ít nhầm lẫn hơn công thức được kiểm tra ở lần thẩm định thứ hai. Như vậy, GV có thể cân nhắc tùy đối tượng HS mà lựa chọn câu hỏi cho phù hợp. Nếu đối tượng là HS có học lực trung bình – yếu, GV có thể lựa chọn câu hỏi ở lần 1, nếu đối tượng là HS có học lực trên trung bình, GV có thể lựa chọn câu hỏi ở lần 2.

Câu 17 cũng là một câu có cảnh báo màu đỏ do độ khó cao (0,12) và độ phân biệt cũng khá thấp (0,22). Nội dung câu 17 như sau:

**Câu 17.** Một học sinh đã giải phương trình  $2\log_3(x-2)+\log_3(x-4)^2=0$  như sau:

**Bước 1:** Điều kiện:  $\begin{cases} x-2>0 \\ (x-4)^2>0 \end{cases} \Leftrightarrow \begin{cases} x>2 \\ x\neq 4 \end{cases}$

**Bước 2:** Ta có:

$$2\log_3(x-2)+\log_3(x-4)^2=0$$

$$\Leftrightarrow 2\log_3(x-2)+2\log_3(x-4)=0$$

**Bước 3:**

$$\Leftrightarrow \log_3(x-2)(x-4)=0 \Leftrightarrow (x-2)(x-4)=1 \Leftrightarrow x^2-6x+7=0$$

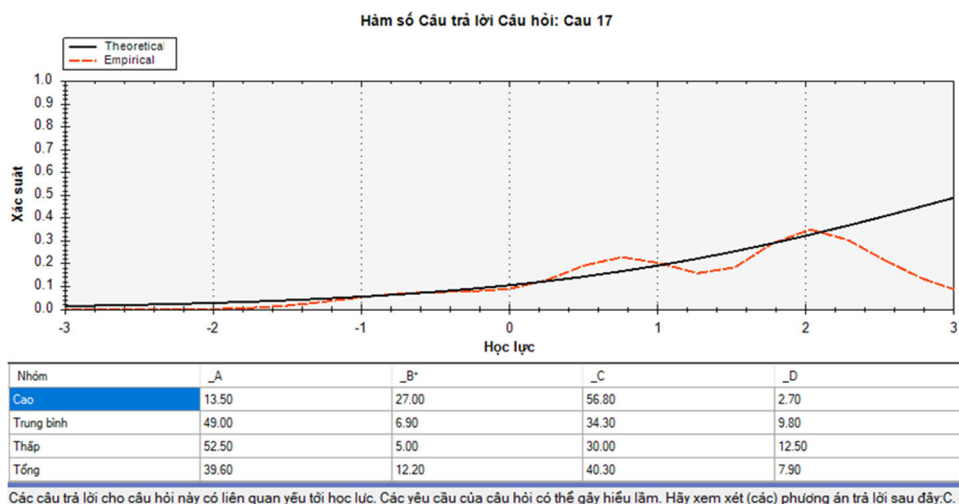
$$\Leftrightarrow x=3\pm\sqrt{2}. \text{ Vậy phương trình đã cho có nghiệm là } x=3\pm\sqrt{2}.$$

Mệnh đề nào dưới đây đúng?

- A. Bài giải trên sai từ bước 1.
- B. Bài giải trên sai từ bước 2.
- C. Bài giải trên sai từ bước 3.
- D. Bài giải trên đúng hết.

Kết quả phân tích câu 17 từ IATA như sau:





Hình 6: Kết quả phân tích câu hỏi 17 sau lần thẩm định thứ nhất

Mục đích của chúng tôi là kiểm tra kiến thức về phương trình lôgarit và kĩ năng biến đổi công thức của HS. Cụ thể là HS thường sai lầm khi cho rằng  $\log_a b^2 = 2 \log_a b$  với  $0 < a \neq 1$ . Tuy nhiên, để giải phương trình ban đầu còn đòi hỏi thêm ở HS một số cách biến đổi hàm số lôgarit khác. Do đó, câu hỏi này có độ khó cao và ít HS trả lời đúng. Ta thấy các đáp án nhiều A và C có nhiều HS lựa chọn (80%), trong đó đáp án nhiều C có số HS điểm cao chọn nhiều hơn số HS điểm trung bình, số HS điểm thấp chọn đáp án C là ít nhất trong ba nhóm. Từ đó cho thấy khi giải dạng toán này HS mắc sai lầm này khá nhiều. Từ ý tưởng ban đầu của bài toán, câu 17 được chỉnh sửa lại như sau:

**Câu 17.** Một học sinh đã giải phương trình  $\log_3(x-2)^4 = 8$  như sau:

**Bước 1:** Điều kiện:  $(x-2)^4 > 0 \Leftrightarrow x \neq 2$ .

**Bước 2:** Ta có:

$$\log_3(x-2)^4 = 8 \Leftrightarrow 4 \log_3(x-2) = 8 \Leftrightarrow \log_3(x-2) = 2$$

**Bước 3:**  $\Leftrightarrow x-2=3^2 \Leftrightarrow x-2=6 \Leftrightarrow x=8$  (thỏa mãn điều kiện).

Vậy phương trình đã cho có nghiệm là  $x = 8$ .

Mệnh đề nào dưới đây đúng?

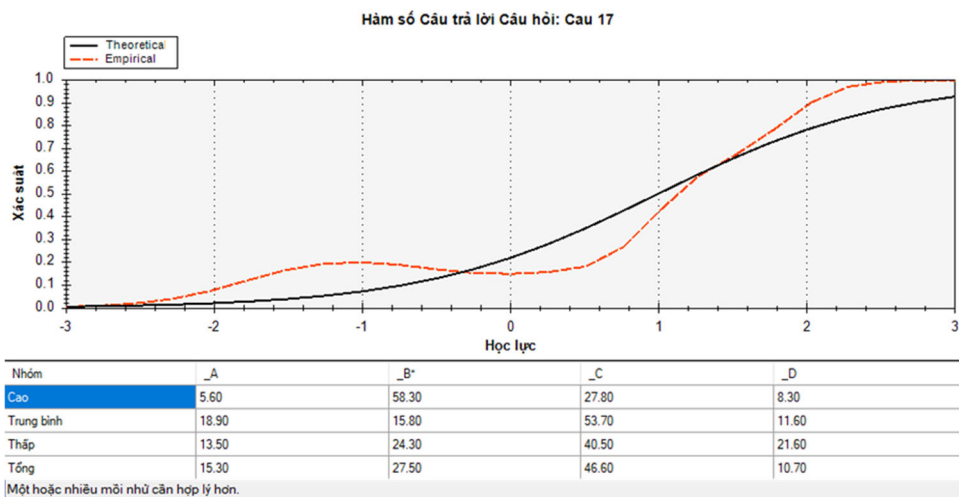
A. Bài giải trên sai từ bước 1.

B. Bài giải trên sai từ bước 2, bước 1 đúng.

C. Bài giải trên sai từ bước 3, bước 1 và bước 2 đúng.

D. Bài giải trên đúng hết.

Kết quả phân tích từ IATA sau khi thẩm định lần 2 cho ta thấy câu hỏi sau khi chỉnh sửa đã dễ hơn (độ khó tăng lên được 0,27) và độ phân biệt đã tăng lên 0,34 (độ phân biệt chấp nhận được). Kết quả phân tích cụ thể như sau:



Hình 7: Kết quả phân tích câu hỏi 17 sau lần thẩm định thứ nhất

Mục đích kiểm tra ban đầu vẫn không thay đổi, tuy nhiên phương trình được cho trong đề bài đơn giản hơn và HS sẽ dễ dàng bộc lộ hơn lỗi mà ta dự đoán ban đầu. Sau khi chỉnh sửa cho câu hỏi dễ hơn nhưng vẫn giữ nguyên ý tưởng ban đầu của bài toán, số HS chọn đúng đã nhiều hơn, số HS chọn đáp án nhiều A đã giảm. Tuy nhiên đáp án nhiều C vẫn còn nhiều HS lựa chọn, đặc biệt là số HS có điểm số trung bình lựa chọn đáp án nhiều C nhiều hơn số HS có điểm thấp. Những điều này khiến cho câu 17 vẫn còn cảnh báo màu vàng (hình 8thoại).

Đây là một lỗi sai lớn mà nhiều HS mắc phải, thậm chí là HS khá giỏi. Qua đó GV cần giúp HS lưu ý để tránh lỗi sai này và có thể sử dụng câu hỏi này ở mức nhận thức là Vận dụng thấp (do câu hỏi khó).

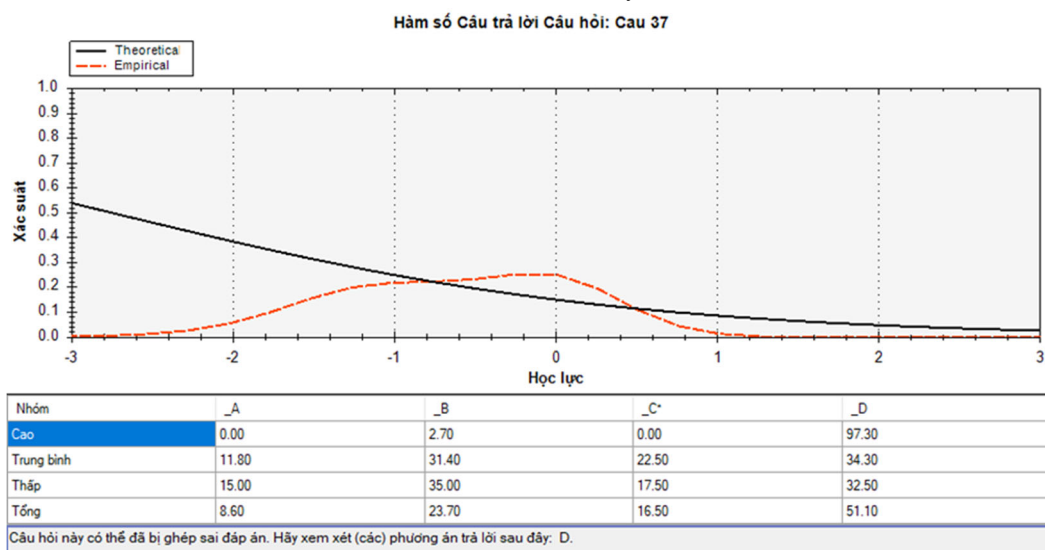
Ta thấy câu 37 cũng có báo hiệu màu đỏ. Câu này có độ khó phân tích được là 0,17 và độ phân biệt là -0,18. Nội dung câu 37 như sau:

**Câu 37.** Số lượng tin nhắn rác trên điện thoại di động trong năm 2009 là mười triệu. Vào năm 2013, tổng số tin nhắn rác tăng theo hàm mũ đến 500 triệu. Biết số tin nhắn rác được tính theo hàm số mũ

$y=ab^x$ , trong đó  $y$  là số tin nhắn rác của năm  $k$ ,  $a$  là số tin nhắn rác trong năm 2009,  $x=k-2009$  là số năm tính từ năm 2009 đến năm  $k$  và  $b$  là hằng số. Hỏi vào năm nào số lượng thư rác trên điện thoại di động đạt đúng 160 tỉ ?

- A. Năm 2016. B. Năm 2017.  
C. Năm 2018. D. Năm 2019.

Phần mềm IATA cho số liệu phân tích cụ thể câu hỏi này như sau:



**Hình 8:** Kết quả phân tích câu hỏi 37 sau lần thẩm định thứ nhất

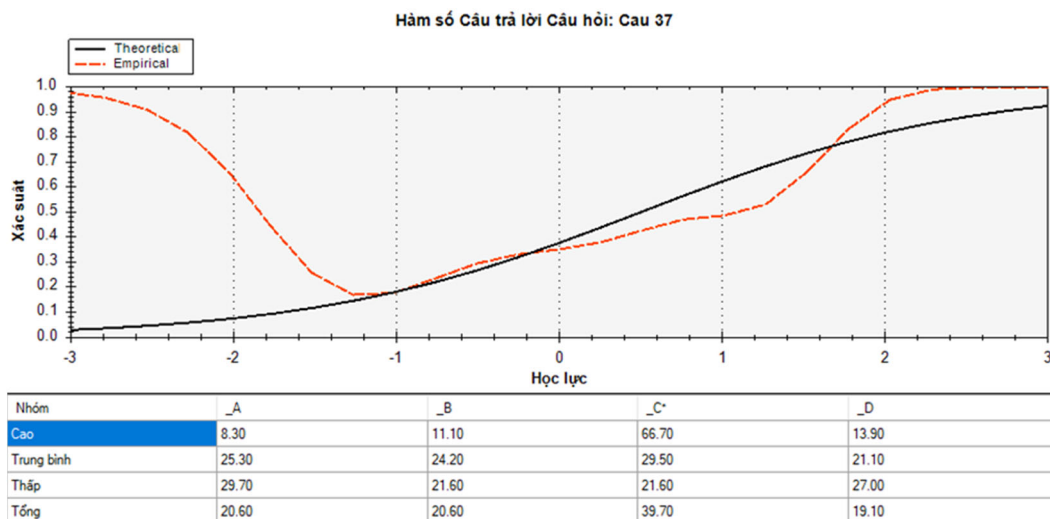
Ta thấy số HS có điểm thấp lại chọn đáp án đúng nhiều hơn số HS có điểm cao, trong khi đó có tới 97,3% HS có điểm cao chọn đáp án nhiều D. Điều này cho thấy câu hỏi có thể đã ghép sai đáp án hoặc có sự nhầm lẫn, chưa chính xác trong cách sử dụng từ ngữ hoặc là HS đã hiểu lầm yêu cầu của bài toán. Câu 37 được chỉnh sửa lại như sau:

**Câu 37.** Số lượng tin nhắn rác trên điện thoại di động trong cả năm 2009 là mười triệu. Trong cả năm 2013, tổng số tin nhắn rác tăng đến 500 triệu. Biết số tin nhắn rác được tính theo hàm số mũ  $y = ab^x$ , trong đó  $y$  là số tin nhắn rác của năm  $k$ ,  $a$  là số tin nhắn rác trong năm 2009,  $x=k-2009$  là số năm tính

từ cuối năm 2009 đến cuối năm  $k$  và  $b$  là hằng số. Hỏi trong năm nào số lượng thư rác trên điện thoại di động đạt **đúng** 160 tỉ (giả sử tất cả các năm đều chỉ có 365 ngày)?

- A. Năm 2016. B. Năm 2017.  
C. Năm 2018. D. Năm 2019.

Ở đây một số từ ngữ đã được chỉnh sửa và bổ sung thông tin khiến bài toán rõ ràng và chặt chẽ hơn. Đồng thời in đậm từ **“đúng”** để gây chú ý cho HS hơn, khiến cho các em hiểu rõ hơn yêu cầu bài toán. Do đó, độ khó câu 37 đã tăng lên được 0,4 và độ phân biệt đã tăng lên được 0,45. Cụ thể như sau:



Hình 9: Kết quả phân tích câu hỏi Câu 37 sau lần thẩm định thứ hai

Ta thấy các đáp án nhiễu đã được lựa chọn khá đồng đều, số lượng HS điểm cao, điểm trung bình, điểm thấp chọn đáp án đúng C giảm dần. Ta thấy câu 37 sau khi chỉnh sửa đã thu được kết quả kiểm định tốt hơn lần đầu nhiều.

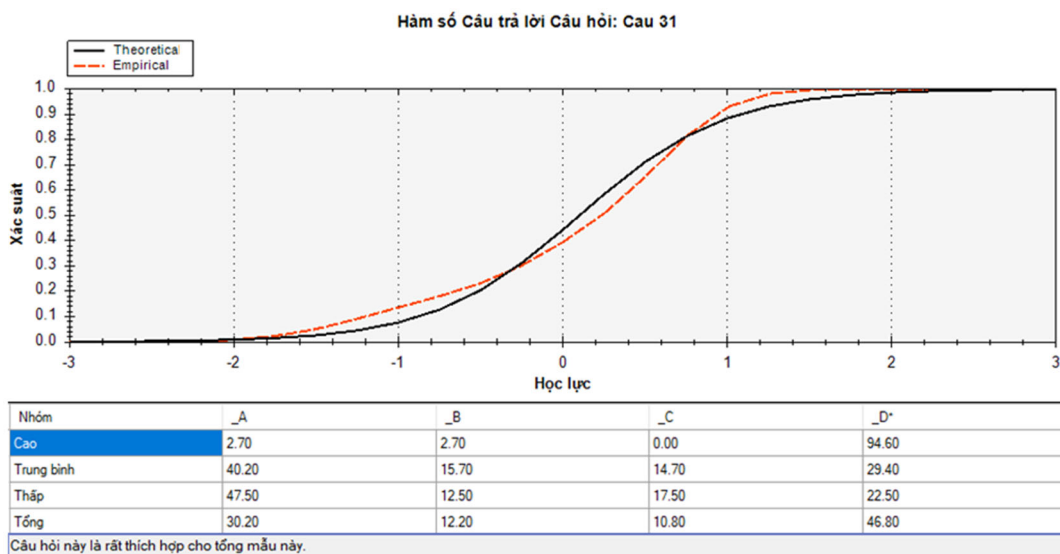
Phần mềm cũng cho chúng ta biết được mức độ phù hợp và hiệu quả của các câu hỏi được biên soạn tốt. Ví dụ như câu 31 có màu xanh (hình tròn).

**Câu 31.** Nhà địa chất học Charles Francis Richter đã định nghĩa cường độ động đất là  $n = \log \frac{I}{I_0}$ , trong đó  $I$  là biên độ chấn động của trận động đất và  $I_0$  là biên độ chuẩn. Trận động đất

Loma Prieta năm 1989 rung chuyển San Francisco (Mỹ) có cường độ 7,1 độ Richter. Năm 2004, trận động đất Sumatra – Andaman kèm sóng thần ở Indonesia có cường độ mạnh 9,2 độ Richter. Hỏi biên độ chấn động của trận động đất Sumatra – Andaman mạnh gấp bao nhiêu lần của trận động đất Loma Prieta?

- A. 1,3 lần.    B. 8,2 lần.  
C. 19,8 lần.    D. 125,9 lần.

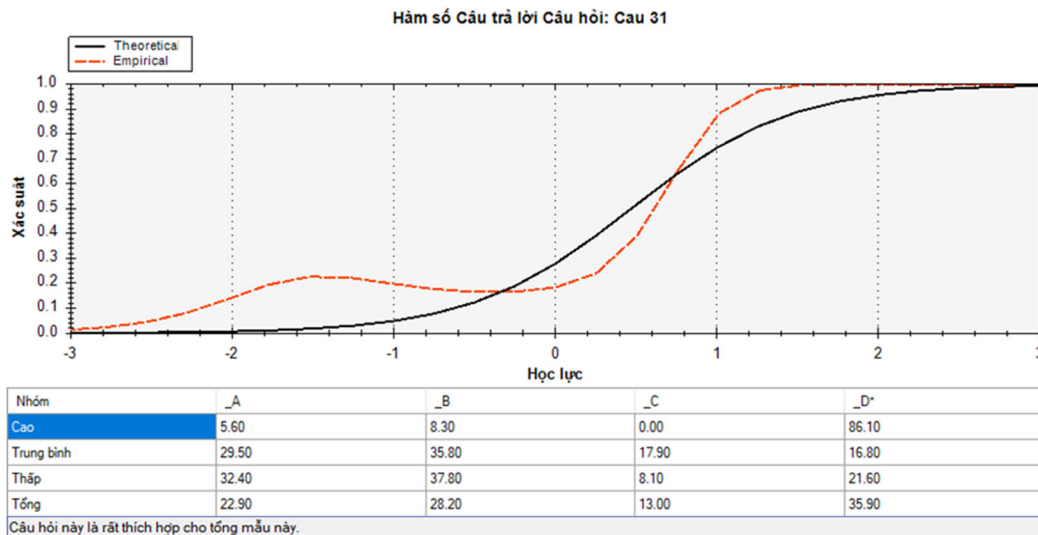
Ở lần thẩm định đầu tiên câu 31 có độ khó là 0,47 và độ phân biệt là 0,72. Câu hỏi này có độ khó vừa phải và độ phân biệt cao. Kết quả phân tích lần 1 như sau:



Hình 10: Kết quả phân tích câu hỏi 31 sau lần thẩm định thứ nhất

Với mẫu HS đầu tiên, câu hỏi này rất thích hợp, đáp án nhiều A là một mối như hấp dẫn. Bên cạnh đó, các đáp án nhiều B và C cũng có một số HS lựa chọn.

Ở lần phân tích thứ 2, câu 31 có độ khó là 0,36 và độ phân biệt là 0,64. Như vậy với mẫu HS thứ 2, có ít HS trả lời đúng câu này hơn so với mẫu HS thứ nhất. Cụ thể như sau:



Hình 11: Kết quả phân tích câu hỏi 31 sau lần thẩm định thứ hai

Với mẫu HS thứ 2, các em lựa chọn đáp án nhiều B nhiều hơn. Dù ICC của câu 31 ở lần kiểm định thứ 2 có biến động hơn so với ICC ở lần kiểm định thứ 1 nhưng xét về tổng thể, câu hỏi này cũng rất thích hợp với mẫu HS thứ 2.

## 7 KẾT LUẬN VÀ ĐỀ XUẤT

Để biên soạn một ĐKT tốt với các đáp án nhiều hấp dẫn cần có một quá trình nghiên cứu và biên soạn thực sự nghiêm túc. Tuy nhiên phải thông qua nhiều lần kiểm định với những số liệu cụ thể dựa trên một cơ sở lý thuyết về TNKQ vững chắc mới đảm bảo được một ĐTN chất lượng, đáng tin cậy và thỏa mãn các yêu cầu thiết yếu của các phần mềm thống kê thẩm định câu hỏi. Trong khuôn khổ một bài báo, bài viết này chỉ mới mang đến những yếu tố cơ bản nhất khi phân tích, đánh giá một câu hỏi TNKQ nói riêng và một ĐTN nói chung. Để khai thác tối đa và hiệu quả hết các tính năng của phần mềm IATA cũng như nắm được nhiều thông tin nâng cao hơn về ĐKT, cần có những nghiên cứu chuyên sâu, chi tiết và quy mô hơn.

Bên cạnh đó, có một tiềm năng to lớn mang ý nghĩa quan trọng trong dạy học khi kết hợp giữa IATA với phân tích lỗi sai của từng nhóm đối tượng HS theo năng lực (với điều kiện là các đáp án nhiều phải dự đoán được sai lầm có thể gặp phải của HS khi giải một dạng bài toán cụ thể nào đó). Từ đó, GV có thể rút ra được những kết luận khách quan về sai lầm của HS khi giải toán và có những phương pháp giảng dạy HS, thậm chí là từng nhóm đối tượng với

những cấp độ năng lực khác nhau, một cách hợp lý để mang lại hiệu quả giảng dạy tốt nhất.

## TÀI LIỆU THAM KHẢO

- Bộ Giáo dục và Đào tạo, 2010. Công văn số 8773/BGDĐT-GDTrH, ngày 30/12/2010 về việc “Hướng dẫn biên soạn đề kiểm tra”, ngày truy cập 19/12/2018. Địa chỉ: <https://thuvienphapluat.vn/cong-van/Giao-duc/Cong-van-8773-BGDDT-GDTrH-huong-dan-bien-soan-de-kiem-tra-148324.asp>.
- Dương Thiệu Tống, 2005. Trắc nghiệm và đo lường thành quả học tập (phương pháp thực hành). Nxb Khoa học xã hội. Tp Hồ Chí Minh, 545 trang.
- Đăng Thị Hương, 2012. Xây dựng hệ thống câu hỏi trắc nghiệm khách quan chương động học chất điểm với sự hỗ trợ của phần mềm Vitesta. (Tuyển tập Báo cáo Hội nghị Sinh viên Nghiên cứu Khoa học lần thứ 8 Đại học Đà Nẵng), Hội nghị Sinh viên Nghiên cứu Khoa học lần thứ 8, ngày 22/05/2012, Đại học Đà Nẵng, Đà Nẵng, 21 – 32.
- Lâm Quang Thiệp, 2008. Trắc nghiệm và ứng dụng. Nxb Khoa học và Kỹ thuật. Hà Nội, 214 trang.
- Lâm Quang Thiệp, 2012. Đo lường và đánh giá hoạt động học tập trong nhà trường. Nxb Đại học Sư phạm Hà Nội, Hà Nội, 207 trang.
- Lê Ngọc và Đào Thị Trang, 2017. Áp dụng phần mềm IATA để phân tích, đánh giá đề thi trắc nghiệm khách quan môn Toán cao cấp A1. Kỷ yếu hội thảo Dạy học Toán và Đào tạo giáo viên đáp ứng yêu cầu đổi mới giáo dục hiện nay. ĐHSP Huế, 3-2017, 33-39.

- Trần Thị Anh Đào, 2008. Sử dụng phần mềm Quest để phân tích câu hỏi trắc nghiệm dùng trong kiểm tra đánh giá kết quả học tập ở học sinh lớp 10. Tuyển tập Báo cáo Hội nghị Sinh viên Nghiên cứu Khoa học lần thứ 6 Đại học Đà Nẵng, ngày 28/5/2008, Đà Nẵng, 237-250.
- Ayala, R., 2009. The Theory and Practice of Item Response Theory. The Guilford Press. New York, USA, 448 pages.
- Sudol, L.A. and Studer, C., 2010, March. Analyzing test items: using item response theory to validate assessments. In Proceedings of the 41st ACM technical symposium on Computer science education (SIGCSE 1010), March 10 – 13, 2010, Milwaukee, Wisconsin. USA. 410-416.
- Thomas, M.H., 2004. Developing and Validating Multiple – Choice Test Items. Lawrence Erlbaum Associates. London, 306 pages.
- Calibration Stability. Rasch Measurement Transactions, 328 pages.
- Linda, C. and James, A., 2008. Introduction to Classical and Modern Test Theory. Cengage Learning. USA, 527 pages.
- Sudol, L.A. and Studer, C., 2010, March. Analyzing test items: using item response theory to validate assessments. In Proceedings of the 41st ACM technical symposium on Computer science education (SIGCSE 1010), March 10 – 13, 2010, Milwaukee, Wisconsin. USA. 410-416.
- Thomas, M.H., 2004. Developing and Validating Multiple – Choice Test Items. Lawrence Erlbaum Associates. London, 306 pages.