TIẾN TRÌNH PHÂN TÍCH & ĐÁNH GIÁ ĐỀ THI TRẮC NGHIỆM KHÁCH QUAN

ThS. NGUYỄN CHÍNH THẮNG *

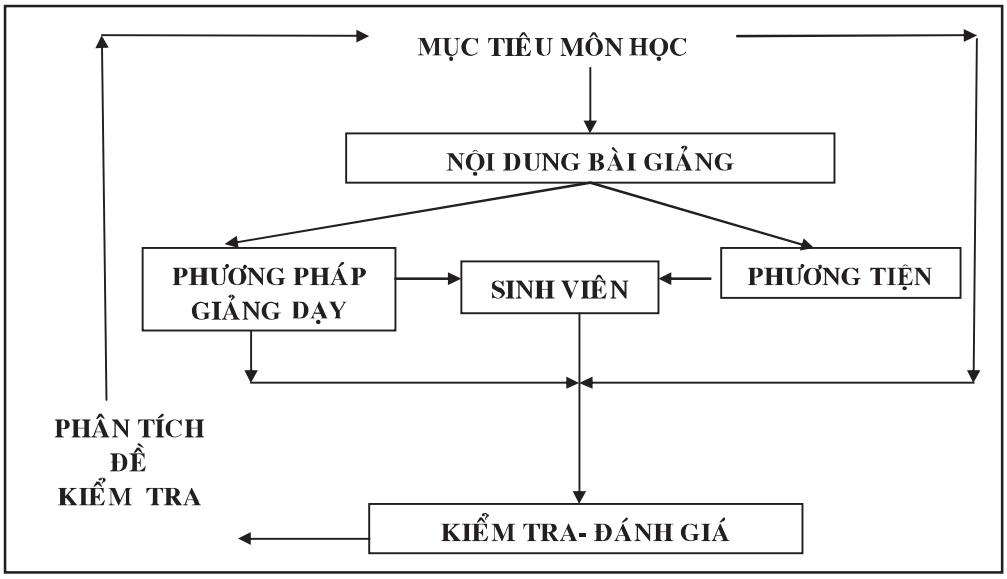
Mở đầu :

Bài viết này tôi xin kính tặng Cố GS-TS Dương Thiệu Tống. Nội dung của bài viết nhằm minh họa dưới dạng thực hành, từng bước phân tích đánh giá một đề thi tốt nghiệp thực tế. Các kiến thức sử dụng trong bài viết này hầu hết tham khảo trong các tài liệu của Cố GS-TS Dương Thiệu Tống, phần còn lại từ các tác giả khác. Phần kiểm định, so sánh giá trị trung bình 2 nhóm thi cùng một đề thi gốc chỉ khác nhau về thứ tự được hóan vị bằng chương trình máy tính, tác giả sử dụng tiêu chuẩn Manwithney. Hỗ trợ tính toán là các đoạn chương trình thực dụng do tác giả viết.

A. PHÂN TÍCH CÂU HỎI

I. ĐẶT VẤN ĐỀ

Việc giảng viên ra một đề thi, chấm thi và công bố điểm cho sinh viên được biết thành quả của



họ chưa phải là giảng viên đó đã hoàn tất một chu trình giảng dạy môn học của mình. Để đánh giá được mức độ thành công trong việc giảng dạy của giảng viên và sự tiếp thu môn học của sinh viên, việc phân tích được đề thi, các câu hỏi đã cho trong kỳ kiểm tra hết môn đó là rất cần thiết. Nhằm 2 mục tiêu:

- Hỗ trợ giảng viên cải tiến phương pháp giảng dạy, điều chỉnh mục tiêu và nội dung bài giảng cho chu kỳ giảng dạy - kiểm tra kế tiếp.
- Tìm hiểu xem sinh viên đã trả lời các câu hỏi thi như thế nào?
- Qua đó thông báo, phân tích cho sinh viên biết những kiến thức cần bổ sung, phát triển.

* *Phó Giám đốc Trung tâm khảo thí - Trường Đại học Mở TP.HCM*

Bài viết này sẽ hỗ trợ giai đoạn phân tích đề thi, câu hỏi thi và độ tin cậy của bài thi kiểm tra dưới hình thức trắc nghiệm khách quan (TN) nhiều lựa chọn. Số liệu sử dụng để minh họa trong bài viết này là số liệu thực, được lấy từ một đề thi tốt nghiệp gồm 50 câu hỏi (4 lựa chọn) với số sinh viên dự thi 488 người. Khảo sát với giả thiết 50 câu hỏi đều có tỷ trọng điểm bằng nhau (mỗi câu tương ứng 0.20 theo thang điểm 10), trong bài viết mỗi câu làm đúng được tính 1, làm sai được tính 0. Số điểm đạt được của mỗi sinh viên dự thi là số câu đúng trong 50 câu.Với mỗi mục minh họa đều có trình bày các công thức, bảng tính, qui tắc tương ứng.

II. PHÂN TÍCH CÂU HỎI MỘT ĐỀ THI TRẮC NGHIỆM KHÁCH QUAN (ĐTN).

Phương pháp thông dụng là chúng ta so sánh câu trả lời của sinh viên dự thi (SV) tại mỗi câu hỏi với điểm số chung của toàn bộ ĐTN. Chúng ta kỳ vọng: *tại lựa chọn có giá trị đúng của mỗi câu hỏi số lượng SV thuộc nhóm có điểm cao chọn nhiều hơn số lượng sinh viên thuộc nhóm có điểm thấp*; và tại các lựa chọn có giá trị sai (câu trả lời nhiễu), *số lượng SV thuộc nhóm cao sa bẫy ít hơn số SV thuộc nhóm thấp sa bẫy*.

2.1 Tính toán - sắp xếp số liệu :

Nhằm khảo sát sự tương quan giữa trả lời mọi câu hỏi với điểm số tổng quát, chúng ta so sánh điểm số giữa 2 nhóm SV: nhóm SV có điểm làm bài cao (số câu đúng nhiều) và nhóm SV có điểm thấp (số câu đúng ít. Theo lý thuyết thống kê, chúng ta có thể dùng phân bố chuẩn khi kích thước mẫu khá lớn (trên 50 SV).

Bước 1 (xử lý - tinh lọc dữ liệu): Từ danh sách SV dự thi chúng ta loại bỏ những SV vi phạm thi và cách làm bài như tô sai mã đề thi, tô sai trả lời (chọn 2 lựa chọn trở lên,…)

Bước 2 (chia nhóm): Chia mẫu SV dự thi đã có điểm thành 2 nhóm này như sau: từ danh sách đã chấm chúng ta sắp xếp theo thứ tự giảm dần theo số lượng câu trả lời đúng: nhóm cao được lấy 27% đến 30% SV có điểm cao nhất (từ trên danh sách đã sắp lấy xuống) và lấy từ 27% đến 30% SV có điểm thấp nhất (từ dưới danh sách lấy lên). Số sinh viên còn lại gọi là nhóm trung bình. Dù lấy với tỷ lệ phần trăm nào trong khỏang 27% đến 30%, ta cần chọn sao cho số lượng SV nhóm cao và nhóm thấp bằng nhau.

Bước 3 (tính toán sơ khởi): Để phân tích cần đếm được số lượng chọn trả lời cho mỗi câu hỏi của ĐTN. Cụ thể ta cần biết trong số các lựa chọn (A, B, C, D) cho sẵn ở mỗi câu hỏi :

- Có bao nhiêu SV chọn câu lựa chọn đúng?
- Có bao nhiêu SV chọn câu lựa chọn sai (câu nhiễu) ? và
- Có bao nhiêu SV không trả lời câu hỏi đó (bỏ trống không làm)?

Bước 4: Sau khi thống kê số lượng từng câu trả lời của ĐTN như trên ở các nhóm có điểm cao , điểm thấp và điểm trung bình, chúng ta có thể suy ra :

- Mức độ khó của câu hỏi
- Mức độ phân biệt giữa 2 nhóm có điểm cao và nhóm thấp của từng câu hỏi
- Mức độ hấp dẫn sinh viên của các câu lựa chọn có nhiệm vụ gây nhiễu cho sẵn để chọn trong mỗi câu hỏi thi.
- Phương sai của mỗi câu lựa chọn.

Kết quả chúng ta nhận được bảng thống kê có cấu trúc sau:

Cột 1	Cột 2	Cột 3	Cột 4	Cột 5	Cột 6	Cột 7
Câu hỏi thứ	Câu trả lời để chọn	SỐ NGƯỜI CHỌN			Tổng số người chọn	Nhóm Cao trừ Nhóm Thấp
		Nhóm cao	Nhóm Trung Bình	Nhóm Thấp		
1	A B C D E				C3+C4+C5	Cột 3-Cột 5
	Bỏ trống không làm				Tổng cột 6 = Số SVDT	
	Tổng : SL nhóm cao SL nhóm TB SL nhóm thấp Tổng cột 6 = Số SVDT bằng 0 Ghi chú : Tổng SL nhóm cao = Tổng SL nhóm thấp . Tổng cột 7 = 0					

Ghi chú :

+ Cột 6 = Cột 3 + Cột 4 + Cột 5 tại mỗi dòng.

+ Tổng của cột 6 cho số sinh viên dự thi.

+ Giá trị trong cột 7 bằng giá trị trong cột 3 trừ giá trị trong cột 5 gọi là hệ số phân biệt. Giá trị cột 7 có thể âm khi giá trị trong cột 3 nhỏ hơn cột. Tính tổng các con số ở cột 7 cho câu hỏi tương ứng, tổng đại số này phải bằng không; nếu không bằng 0 chúng ta cần xem lại các kết quả tính các giá trị trên cho câu này.

Ví dụ : Khảo sát đề thi trắc nghiệm đã được tổ chức tháng 04/2009 nói trong I.

Tổng SV dự thi: 488 sinh viên. Số sinh viên làm bài TN không đúng qui định : 13

Danh sách dùng làm mẫu xử lý – phân tích minh họa: 475 sinh viên

Nhóm cao có: 128 SV ; Nhóm thấp có: 128 SV và Nhóm trung bình có: 219 SV;

Ví dụ : Xét câu hỏi số 2 của đề thi gốc, trình bày theo cấu trúc bảng trên ta có:

PHÂN TÍCH MỘT ĐỀ THI TRẮC NGHIỆM KHÁCH QUAN 4 LỰA CHỌN

Câu	Đ-ÁN	LỰA CHỌN	CAO	TBINH	THẤP	TỔNG	HS-PC
2	D	A	6	14	16	36	-10
2	D	B	2	4	4	10	-2
2	D	C	33	75	52	160	-19
2	D	D	87	123	54	264	33
2		K		3	2	5	-2
			128	219	128	475	0

2.2 Tính các giá trị đặc trưng

2.2.1. Độ khó của một câu hỏi .

Độ khó p_i của một câu hỏi thứ i bằng tỷ số giữa số SV dự thi trả lời đúng câu hỏi i đó với tổng số sinh viên dự thi: $$p_i = \frac{so\ SV\ tra\ loi\ dung\ cau\ thu\ i}{tong\ so\ \ sinh\ \ vien\ du\ thi}$$

2.2.2. Độ phân biệt của một câu hỏi.

Độ phân biệt hay hệ số phân cách của một câu hỏi bằng hiệu số của số lượng sinh viên thuộc nhóm cao trả lời đúng với số lượng sinh viên thuộc nhóm thấp đã trả lời đúng. Tại các câu nhiễu chúng ta cũng tính hiệu số này.

Để có độ phân biệt không phụ thuộc vào số người dự thi chúng ta tìm hệ phân cách lớn nhất trong tất cả các phân cách

Đặt H_i = số SV nhóm cao trả lời đúng câu hỏi i;

L_i =số SV nhóm thấp trả lời đúng câu hỏi i

$PC(i) = H_i - L_i$; $Mpc = \max\{PC(i) ; i = 1, 2, \ldots, \text{số câu hỏi ĐTN}\}$

Trong ĐTN khảo sát mẫu này ta có Mpc = 50. Một số tài liệu chọn đơn giản Mpc bằng số SV thuộc nhóm cao (cũng bằng số sinh viên thuộc nhóm thấp).

Tỷ số $h_i = \frac{PC(i)}{Mpc}$ cũng được gọi là độ phân biệt của câu hỏi thứ i

Kỳ vọng :

- Tại các câu lựa chọn có giá trị đúng : PC(i) > 0
- Tại các câu lựa chọn có giá trị sai : PC(i) < 0

Bảng tiêu chuẩn định tính một câu hỏi với d là độ phân cách cho lựa chọn đúng:

d ≥ 40%	Phân cách rất tốt
0.30 ≤ d < 0.39	Phân cách khá tốt
0.20 ≤ d < 0.29	Phân cách trung bình
d < 0.19	Phân cách kém, cần xem lại

2.2.3. Tính độ khó vừa phải của câu trong bài thi trắc nghiệm

- Đặt SOCAU = số câu hỏi của ĐTN.
- SPC = số lựa chọn (số lựa chọn của mỗi câu trắc nghiệm)
- TBBTN = Tổng điểm thi / SOCAU gọi là trung bình bài trắc nghiệm
- TBDMR = Trung bình điểm may rủi = $\frac{1}{SLC}$
- ĐỘ KHÓ VỪA PHẢI CÂU TN = $DKVP = \frac{1 + TBDMR}{2}$
- Kết Quả : so sánh độ khó của câu p(i) và độ khó vừa phải DKVP.

Nếu có 4 lựa chọn A, B, C, D tại mỗi câu

$$DKVP = \frac{1 + \frac{1}{4}}{2} = \frac{5}{8} \approx 62\ \%$$

Nếu p(i) xấp xỉ DKVP thì kết luận câu hỏi thứ i có độ khó vừa phải.
Nếu p(i) khá lớn so với DKVP thì kết luận câu hỏi thứ i là câu dễ.
Nếu p(i) khá nhỏ so với DKVP thì kết luận câu hỏi thứ i là câu hỏi khó.

Ví dụ (trích phân tích các câu hỏi đề TN đã dẫn)

PHÂN TÍCH CÂU HỎI MỘT ĐỀ THI TRẮC NGHIỆM KHÁCH QUAN (4 LỰA CHỌN)

CÂU	ĐÁP ÁN	CHỌN	CAO	TBINH	THẤP	TỔNG	HIỆU SỐ PHÂN CÁCH	ĐỘ PHÂN CÁCH	ĐỘ KHÓ	PHƯƠNG SAI	NHẬN XÉT ĐỘ PHÂN CÁCH CỦA CÂU	NHẬN XÉT ĐỘ KHÓ
2	D	A	6	14	16	36	-10	-0.2	0.08			
	D	B	2	4	4	10	-2		0.02			
	D	C	33	75	52	160	-19	-0.3	0.34			
	D	D	87	123	54	264	33	0.58	0.56	0.3	Phân cách rất tốt	Câu khó vừa phải
		K		3	2	5	-2					
			128	219	128	475	0					
13	D	A	1	1	1	3			0.01		Câu nhiễu kém điều chỉnh lại	
	D	B	12	47	35	94	-23	-0.4	0.2			
	D	C	96	131	73	300	23	0.4	0.63		Câu nhiễu kém điều chỉnh lại	
	D	D	18	39	19	76	-1		0.16	0.1	Phân cách kém	Câu khó
		K	1	1		2	1					
			128	219	128	475	0					
36	C	A	23	71	49	143	-26	-0.5	0.3			
	C	B	5	6	19	30	-14	-0.3	0.06			
	C	C	100	138	59	297	41	0.72	0.63	0.2	Phân cách rất tốt	Câu hơi dễ
	C	D		3	1	4	-1	-0	0.01			
		K		1		1						
			128	219	128	475						
44	C	A	12	12	12	36			0.08		Câu nhiễu kém điều chỉnh lại	
	C	B	23	41	35	99	-12	-0.2	0.21			
	C	C	67	120	42	229	25	0.44	0.48	0.3	Phân cách rất tốt	Câu khó vừa phải
	C	D	23	44	34	101	-11	-0.2	0.21			
		K	3	2	5	10	-2					
			128	219	128	475	0					
45	C	A	19	44	18	81	1	0.02	0.17		Câu nhiễu kém điều chỉnh lại	
	C	B	64	121	68	253	-4	-0.1	0.53			
	C	C	33	33	21	87	12	0.21	0.18	0.2	Phân cách trung bình	Câu khó
	C	D	11	20	17	48	-6	-0.1	0.1			
		K	1	1	4	6	-3					
			128	219	128	475	0					

Diễn giải kết quả:

Tổng lớn nhất của hiệu phân cách trong 50 câu là Mpc = 57

+ Câu 2 (câu đúng D):

- Số lượng nhóm cao chọn D : 87 sinh viên
- Số lượng nhóm thấp chọn D : 54 sinh viên

=> Hiệu phân cách cho lựa chọn D là : 87 – 54 = 33

Độ phân cách cho lựa chọn D là :(87–54)/57=33/57=58% => phân cách rất tốt

Độ khó của câu (số làm đúng/số dự thi)=264/475=56% => câu có độ khó vừa phải.

Các lựa chọn nhiễu đều có độ phân cách âm=> dùng được.

Chi tiết hơn :

- Lựa chọn A có độ phân cách 20% (tốt)
- Lựa chọn B có độ phân cách xấp xỉ 0 (kém)
- Lựa chọn C có độ phân cách xấp xỉ -30% (tốt). Đặc biệt có khá nhiều sinh viên nhóm cao và thấp đều sa bẫy: câu nhiễu rất tốt và cũng cần xem lại đáp án vì có 160 trong 475 sinh viên chọn lựa chọn này.

+ Câu 13 (câu đúng D):

- Số lượng nhóm cao chọn D : 18 sinh viên
- Số lượng nhóm thấp chọn D : 19 sinh viên

=> Hiệu phân cách cho lựa chọn D là : 18-19 = -1

Độ phân cách cho lựa chọn D là :(18-17)/57=-1/57 xấp xỉ 0% => phân cách kém

Độ khó của câu (số làm đúng/số dự thi)=16% => câu khó.

Các lựa chọn nhiễu đề có độ phân cách âm=> dùng được.

Chi tiết hơn :

- Lựa chọn A có độ phân cách bằng 0 và rất ít SV chọn bẫy này. Hãy làm lại lựa chọn này
- Lựa chọn B có độ phân cách tốt và nhiều sinh viên bị bẫy này hấp dẫn
- Lựa chọn (nhiễu) C có độ phân cách dương khá lớn, xấp xỉ +40% .Nếu là câu nhiễu thì độ phân cách này có tương quan nghịch => là câu nhiễu kém.Nhưng chú ý rằng tại lựa chọn C này số nhóm giỏi chọn khá lớn 96 SV so với 18 SV nhóm thấp và SV thuộc số trung bình chọn C cũng khá lớn. Với tình huống bất thường này giảng viên cần xem lại câu nhiễu C. Vì nếu lựa chọn đúng là C câu đúng thì câu này có độ phân cách khá tốt và là câu có độ khó vừa phải.

Những câu còn lại có thể khảo sát và diễn giải tương tự.

2.3 Chọn - lưu trữ câu hỏi tốt - chỉnh sửa câu kém - xem lại bài học đã giảng

2.3.1 Sau khi phân tích chi tiết từng câu, chúng ta có thể giữ và lưu trữ để sử dụng sau này những câu hỏi có các tính chất sau:

(1). Độ khó vào khoảng từ 40 đến 70 phần trăm.

(2). Độ phân biệt ≥ 30% .

(3). Các câu lựa chọn nhiễu (lựa chọn có giá trị sai) cần thỏa các điều kiện :

+ Hấp dẫn được SV thuộc nhóm thấp và trung bình.

+ Có độ phân cách âm.

Qui trình phân tích các câu hỏi giúp chúng ta biết được khuyết điểm trong câu hỏi hoặc trong công việc giảng dạy. Tuy nhiên, từ các con số thống kê chúng ta không nhất thiết phải loại các câu hỏi không thỏa các điều kiện lưu giữ trên. Những đặc trưng thống kê qua các con số trên chỉ cho

ta thấy những tình huống bất thường cần xem xét lại một cách tỉnh táo. Bất thường có thể do nội dung bài trắc nghiệm không thuần nhất, hoặc cách diễn đạt các lựa chọn gây nhầm lẫn dẫn đến câu hỏi có tương quan nghịch, hoặc tương quan không đáng kể với toàn bài trắc nghiệm. Trong bài trắc nghiệm cũng cần có câu dễ, câu khó theo tỷ lệ qua một tiêu chí nào đó chứ không nhất thiết phải là một ĐTN phải luôn luôn thuần nhất, đồng đều tại mọi câu. Những bài ĐTN quá thuần nhất có khi làm giảm phạm vi–chất lượng của mục đích kiểm tra-đánh giá môn học.

B. ĐÁNH GIÁ ĐỀ THI TRẮC NGHIỆM (ĐTN) - ĐỘ TIN CẬY CỦA (ĐTN)

I. Khảo sát ban đầu :

Trường hợp dùng một đề thi cho nhiều nhóm SV có trình độ như nhau về mặt lý thuyết. Chẳng hạn một đề thi tốt nghiệp dùng cho một khóa đào tạo. Một đề thi gốc đã được một nhóm giảng viên soạn, *các câu hỏi độc lập nhau* và theo đánh giá "chủ quan" của các giảng viên ra đề các câu hỏi này tương đương về độ khó. Từ đề gốc tạo các đề, mỗi đề là một hoán vị các câu hỏi của đề gốc và được ghi mã đề để phân biệt. Danh sách sinh viên được sắp ngẫu nhiên, khi phát đề cán bộ coi thi cũng phát ngẫu nhiên. Số đề phát ra của các đề hoán vị cũng không bằng nhau.

Vấn đề đầu tiên thường được quan tâm là giữa một đề thi có các câu hỏi được sắp xếp theo trình tự như trong quá trình học, ôn tập và một đề thi các câu hỏi được sắp xếp một cách ngẫu nhiên theo một chương trình máy tính có ảnh hưởng đến điểm trung bình của sinh viên dự thi không?

Giải quyết vấn đề này chúng ta có thể làm bằng một trong 2 cách sau:

+ Phương pháp kiểm định thông thường:

Với mức ý nghĩa thích hợp α là một số dương từ 1% ,..., 10% khá bé, chúng ta cần tìm hiểu xem trung bình điểm số của hai nhóm sinh viên thi trên 2 mã đề khác nhau có bằng nhau không? Đây là bài toán kiểm định trong thống kê với giả thiết H_0: "Trung bình của 2 nhóm sinh viên thi theo mã đề bằng nhau".

Trong trường hợp số sinh viên dự thi trên 30 sinh viên chúng ta có thể ước lượng theo phân phối chuẩn. Nếu giả thiết H_0 cần kiểm định được chấp nhận thì ta có thể nói : với mức ý nghĩa α , việc hoán vị câu hỏi từ một đề gốc không ảnh hưởng đến điểm trung bình chung của sinh viên dự thi

Ví dụ : Khảo sát đề thi TN đã nêu. Đề thi này dùng làm một tiêu chuẩn để xét tốt nghiệp (đề thi tốt nghiệp). Đề gốc được chia thành hai đề có mã đề 001 và 002.

Nhóm theo mã đề	Đặc trưng	Giá trị
	Số đề phát ra	N_1 =240
001	Tổng số câu làm đúng	5411
	Trung bình số câu đúng	22.55
	Phương sai mẫu	$S_1^2 = 19.87$
	Số đề phát ra	N_2 = 244
002	Tổng số câu làm đúng	5365
	Trung bình số câu đúng	21.98
	Phương sai mẫu	$S_2^2 = 17.77$

Trong Lý Thuyết Thống kê ta có : $T = \frac{(N_1 + N_2)\sqrt{N_1 * N_2 * (N_1 + N_2 - 2)}}{\sqrt{(N_1 + N_2).(N_1.S_1^2 + N_2.S_2^2)}}$

có phân phối F với (N_1 + N_2 -2) độ tự do

Áp dụng cho ĐTN trên ta có T = 1.41

Cụ thể : với mức ý nghĩa α = 5% ta có t_α = 1.96

Do | T | = 1.41 < t_α = 1.96 nên ta chấp nhận được giả thiết "trung bình điểm số của 2 nhóm (làm theo 2 mã đề của cùng đề gốc) bằng nhau".

II . Tiêu chuẩn Manwithney

Giả thiết H_0 : Trung bình 2 nhóm bằng nhau

Giả thiết đối của H_0 : Trung bình 2 nhóm khác nhau

Gọi N_1 : số lượng sinh viên thi đề có mã đề 001 và làm bài đúng qui định

Gọi N_2 : số lượng sinh viên thi đề có mã đề 002 và làm bài đúng qui định

N_1 + N_2 = tổng sinh viên dự thi và làm bài đúng qui định

Sắp xếp N_1 + N_2 sinh viên dự thi theo thứ tự tăng dần theo số câu đúng)

Tính hạng của mỗi sinh viên (những sinh viên cùng số câu đúng được lấy trung bình cộng của hạng là hạng của mỗi người)

Gọi R1, R2 bằng tổng hạng của nhóm mã đề 001, 002 theo thứ tự. Tính các giá trị:

$$U_1 = N_1 + N_2 + \frac{N_1 + N_2}{2} - R_1$$

$$U_2 = N_1 + N_2 + \frac{N_1 + N_2}{2} - R_2$$

$$U = \min\{U_1, U_2\}$$

$$EU = \frac{N_1 . N_2}{2}; DU = \frac{N_1 . N_2 . (N_1 + N_2 + 1)}{12}$$

$$U_\alpha = \frac{U - EU}{\sqrt{DU}}$$

Với mức ý nghĩa đã cho, tra bảng chúng ta tính được Zalpha.

Nếu | U_α | < Z_α thì ta chấp nhận giả thiết H_0

Ngược lại, bác bỏ H_0.

Áp dụng với số liệu là bài thi trên, ta có:

Mẫu dữ liệu sau khi tính hạng có dạng:

BẢNG TÍNH HẠNG CỦA SINH VIÊN THEO SỐ CÂU LÀM ĐÚNG

STT	TTDSTHI	SO BD	MA DE	SOCAUDUNG	RANK	GHI CHU
1	044	022044	001	10	1.00	
2	027	008038	002	11	2.00	
3	008	005008	002	12	4.00	
4	028	005028	002	12	4.00	
5	_05	007059	001	12	4.00	
6	009	002010	002	13	8.50	
7	057	002058	002	13	8.50	
8	022	006022	002	13	8.50	
9	015	009017	001	13	8.50	
10	009	008010	002	13	8.50	
11	041	008054	002	13	8.50	
..						
..						
..						
..						
482	007	022007	002	33	481.50	
483	_01	007019	001	33	481.50	
484	010	002011	001	34	484.00	

Tính toán số liệu (Nguyễn Chính Thắng)

Ta tính được:

$N_1 = 240$; $N_2 = 244$

$R_1 = 60014$; $R_2 = 57356$

$U = 27466$; $U\alpha = -1.7192$ và $Z\alpha = 1.96$ với mức ý nghĩa 5%.

Do $|U\alpha| < Z\alpha$: nên chấp nhận giả thiết H_0, nghĩa là 2 trung bình điểm số của 2 nhóm bằng nhau

Qua việc tính tóan trên chúng ta rút ra kết luận: có thể yên tâm trộn một đề thi gốc một cách ngẫu nhiên thành nhiều đề cho cùng một nhóm thí sinh cùng trình độ, chỉ với yêu cầu các câu hỏi của đề thi độc lập nhau.

III . Tính độ tin cậy của ĐTN

Có nhiều công thức thống kê thông dụng dùng việc tính độ tin cậy. Ở đây chúng ta dùng phương pháp phân đôi đề TN theo câu chẵn - câu lẻ

Trong phương pháp này từ dữ liệu chấm ta có bảng như sau:

SD = số lượng câu của SV làm đúng ; X = số lượng câu lẻ của SV làm đúng

Y = số lượng câu chẵn của SV làm đúng; SD = X + Y

3.1. Theo công thức Spearman - Brown

Dòng cuối cùng bằng tổng các số cùng cột

TT_SV	SD	X	Y	X.Y	X.X	Y.Y
1	10	5	5	25	25	25
2	14	7	7	49	49	49
3	12	8	4	32	64	16
4	5	3	2	6	9	4
.	.	.	.	.	.	.
.	.	.	.	.	.	.
482	12	6	6	36	36	36
483	4	2	2	4	4	4
484	11	7	4	28	49	16
484	4684	2750	1934	11129	17098	9458

Ta có : $N = 484$

$$\sum X = 2750; \sum Y = 1934; \sum X.Y = 11129$$

$$\sum X^2 = 17098; \sum Y^2 = 9458$$

Hệ số tương quan được tính theo công thức

$$r_s = \frac{\overline{X.Y} - \overline{X}.\overline{Y}}{S_X.S_X}$$ Ta có $$r_s = \frac{22,994 - 5,682x3,996}{\sqrt{3,043x3,574}} = 0,088$$

Hệ số tin cậy của ĐTN trên theo Spearman - Brown : $r_e = \frac{2.r_s}{r_s + 1} =$ 16,16%

3.2. Theo công thức Bulon

Để dùng công thức tính này, từ dữ liệu chấm bài ta lập bảng có dạng như sau:

SD = số lượng câu của SV làm đúng X = số lượng câu lẻ của SV làm đúng

Y = số lượng câu chẵn của SV làm đúng SD = T = X + Y

H = bằng hiệu của số trả lời đúng câu chẵn – số trả lời đúng câu lẻ

TT_SV	X	Y	T = X+Y	H = X-Y	T_2	H_2
1	5	5	10	0	100	0
2	7	7	14	0	196	0
3	8	4	12	4	144	16
.	.	.	.	.	.	.
.	.	.	.	.	.	.

483	2	2	4	0	4	16
484	7	4	11	3	16	256
484	**2750**	**1934**	**4684**	**816**	**48814**	**4298**

$$\sigma_T^2 = \frac{\sum T^2}{N} - \left(\frac{\sum T}{N}\right)^2 = \frac{48814}{484} - \left(\frac{4684}{484}\right)^2 = 7,198$$

$$\sigma_H^2 = \frac{\sum H^2}{N} - \left(\frac{\sum H}{N}\right)^2 = \frac{4298}{484} - \left(\frac{816}{484}\right)^2 = 6,038$$

$$=> r_e = 1 - \frac{\sigma_H^2}{\sigma_T^2} = 1 - \frac{6,038}{7,198} = 16,12\%$$

Hai cách tính cho chúng ta cùng độ tin cậy của đề thi trên: r_e xấp xỉ 16% .

C. KẾT LUẬN

Qua khảo sát ĐTN, chúng ta rút ra được nhiều kết luận khác nhau. Quan trọng nhất là cần xem lại các câu hỏi, câu lựa chọn đúng và nhiễu của tòan bộ đề thi. Từ đó giảng viên góp phần từng bước nâng cao chất lượng bài giảng. Những luận giải của bài viết mong muốn cung cấp một góc nhìn nhằm thúc đẩy tiến trình ra đề thi trắc nghiệm khách quan đối với các môn học cần thiết tại trường.

TÀI LIỆU THAM KHẢO

1. Dương Thiệu Tống (1995), *Trắc nghiệm và đo lường thành quả học tập (phương pháp thực hành) TSKH - GD,* NXB Giáo dục

2. Dương Thiệu Tống (2005), *Thống kê ứng dụng trong nghiên cứu khoa học giáo dục. TSKH - GD,* NXB Thống kê

3. Nguyễn Thị Hồng Minh & Đào Hữu Hồ (1996), *Xử lý số liệu thống kê toán học.*

4. Võ Ngọc Lan - Nguyễn Phụng Hoàng (1998), *Phương pháp trắc nghiệm trong kiểm tra và đánh giá thành quả học tập,* NXB Giáo dục

5. Trần Thị Tuyết Oanh (2006), *Đánh giá và đo lường kết quả học tập,* NXB Giáo dục

6. Nguyễn Ngọc Tiềng (1996), *Thống kê học trong nghiên cứu khoa học,* NXB Giáo dục

7. Nguồn dữ liệu từ kỳ thi TN tháng 04/2009 (ĐHM TPHCM)

8. Nguyễn Chính Thắng. *Bài giảng tóm lược xác suất thống kê toán.*

9. Một số tài liệu toán thống kê khác.