

SO SÁNH ĐIỂM HỌC PHẦN TOÁN VÀ VẬT LÝ CỦA SINH VIÊN TRƯỜNG ĐẠI HỌC XÂY DỰNG MIỀN TÂY

COMPARISON OF MATHS AND PHYSICS MODULE GRADES OF MIEN TAY CONSTRUCTION UNIVERSITY STUDENTS

ThS. Hoàng Công Thiện

Khoa Khoa học cơ bản – Trường ĐHXD Miền Tây

Email: hoangcongthien@mtu.edu.vn

Điện thoại: 0907 916 049

ThS. Trần Quang Đăng

Khoa Khoa học cơ bản – Trường ĐHXD Miền Tây

Email: tranquangdang@mtu.edu.vn

Điện thoại: 0909 531 321

Ngày nhận bài: 27/11/2023

Ngày gửi phản biện: 07/12/2023

Ngày chấp nhận đăng: 19/12/2023

Tóm tắt:

Trong bài báo này, tôi nghiên cứu so sánh điểm học phần toán và điểm học phần vật lý của sinh viên Trường Đại học Xây dựng Miền Tây theo quan điểm thống kê toán học. Kiểm định giả thuyết thường được sử dụng trong trường hợp này là kiểm định theo cặp. Kiểm định theo cặp là một kỹ thuật thống kê được sử dụng để so sánh hai trung bình tổng thể trong trường hợp hai mẫu có tương quan. Dữ liệu được xét theo cặp vì các quan sát trong tập dữ liệu thứ nhất có quan hệ tương ứng với các quan sát trong tập dữ liệu còn lại. Áp dụng nghiên cứu với khóa đào tạo 2022 bậc đại học của Trường Đại học Xây dựng Miền Tây. Có hay không sự khác biệt về điểm học phần toán và điểm học phần vật lý.

Từ khóa: kiểm định theo cặp, kiểm định giả thuyết, thống kê toán học, Đại học Xây dựng Miền Tây.

Abstract:

In this paper, comparison of maths and physics module grades of Mien Tay Construction University students is studied from the point of view of mathematical statistics. Hypothesis testing often are used in this case is paired samples t-test. It is a statistical technique that is used to compare two population means in the case of two samples that are correlated. The data are considered paired since the observations in one dataset are matched to observations in the other dataset. Applied research with undergraduate level 2022 course of Mien Tay Construction University. There is or not a difference in maths and physics module grades.

Keywords: paired samples t-test, hypothesis testing, mathematical statistics, Mien Tay Construction University.

1. Giới thiệu

Đánh giá việc học tập của sinh viên Trường Đại học Xây dựng Miền Tây có ý nghĩa thống kê nếu phân tích số liệu theo thống kê toán học. Đào tạo bậc đại học theo hệ thống tín chỉ, điểm học phần của sinh viên tại một trường đại học thường được xác định thông qua một biểu thức đại số nào đó. Chẳng hạn, đối với Trường Đại học Xây dựng Miền Tây [9], điểm học phần - kí hiệu G_{mod} được xác định bởi công thức

$$G_{mod} = G_{cla} * W_{cla} + G_{exa} * W_{exa},$$

trong đó G_{cla} là điểm quá trình với trọng số W_{cla} và G_{exa} là điểm thi với trọng số W_{exa} . Lưu ý rằng các điểm $G_{cla}, G_{exa}, G_{mod}$ là các số thực thuộc đoạn $[0,10]$ và các trọng số W_{cla}, W_{exa} là các phần trăm tương ứng lần lượt là 40%, 60%. Điểm học phần toán và điểm học phần vật lý được giảng dạy trong năm thứ nhất tại Trường Đại học Xây dựng Miền Tây, hai điểm học phần này có sự khác biệt hay không theo quan điểm thống kê toán học. Biết rằng trong nghiên cứu khoa học ứng dụng, vai trò của thống kê toán học được thể hiện qua thống kê mô tả, so sánh và liên hệ dữ liệu. Do đó, với mức ý nghĩa α cho trước, phương pháp t-test có thể được nghĩ đến (hoặc paired sample t-test hoặc

independent sample t-test) trong nghiên cứu kiểm định sự khác biệt về giá trị trung bình giữa hai tổng thể (hoặc phụ thuộc hoặc độc lập). Hơn nữa, nhận thấy rằng phương pháp paired sample t-test (kiểm định theo cặp) là phù hợp và được chọn cho mục tiêu nghiên cứu.

2. Kiểm định theo cặp

2.1. Bài toán

Cho XX, YY tương ứng lần lượt là hai biến ngẫu nhiên của hai tổng thể mà trong đó hai biến XX, YY là phụ thuộc và tuân theo dạng phân phối chuẩn với phương sai chưa biết [3, 5, 7]. Chọn cặp mẫu ngẫu nhiên (X, Y) có kích thước m là $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$.

Xét một biến ngẫu nhiên DD được thành lập từ hai biến ngẫu nhiên XX, YY với viết dưới dạng kí hiệu hình thức $D = X - Y$ và

$$D_k = X_k - Y_k, k = \overline{1, n}.$$

Vì vậy, với cặp mẫu cụ thể gồm m quan sát

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

của cặp mẫu ngẫu nhiên (X, Y) .

Khi đó một mẫu cụ thể gồm m quan sát $d_1, d_2, \dots, d_n$ của biến ngẫu nhiên DD được tính theo công thức

$$d_k = x_k - y_k, k = \overline{1, n}.$$

Với một mức ý nghĩa α cho trước, chúng ta cần kiểm định giả thuyết

$$H_0: \mu_{dif} = 0,$$

và có đối thuyết

$$H_1: \mu_{dif} \neq 0,$$

trong đó μ_{dif} là trung bình của một tổng thể mà có biến ngẫu nhiên D . Khi đó trung bình mẫu cụ thể và phương sai mẫu cụ thể hiệu chỉnh của biến ngẫu nhiên D tương ứng lần lượt được kí hiệu và xác định bởi các công thức

$$\bar{d} = \frac{1}{n} \sum_{k=1}^n d_k,$$

$$s_{dif}^2 = \frac{1}{n-1} \sum_{k=1}^n (d_k - \bar{d})^2.$$

2.2. Các bước thực hiện kiểm định

Biết rằng bài toán kiểm định theo cặp cũng được thực hiện theo các bước chung của bài toán kiểm định tổng quát. Do đó, cách thức tiến hành kiểm định giả thuyết về giá trị trung bình giữa hai tổng thể phụ thuộc hoặc phối hợp từng cặp được thực hiện sau [4, 7]:

Bước 1. Đặt giả thuyết của kiểm định

$$\begin{cases} H_0: \mu_{dif} = 0, \\ H_1: \mu_{dif} \neq 0. \end{cases}$$

Bước 2. Chọn thống kê

$$T = \frac{\bar{D}}{S_{dif}} \sqrt{n},$$

trong đó trung bình mẫu ngẫu nhiên và phương sai mẫu ngẫu nhiên hiệu chỉnh của biến ngẫu nhiên D tương ứng lần lượt được kí hiệu và xác định bởi các công thức

$$\bar{D} = \frac{1}{n} \sum_{k=1}^n D_k,$$

$$S_{dif}^2 = \frac{1}{n-1} \sum_{k=1}^n (D_k - \bar{D})^2.$$

Bước 3. Miền bác bỏ

$$W_\alpha = \left\{ t: |t| > t_{\frac{\alpha}{2}}(n-1) \right\},$$

trong đó $t_{\frac{\alpha}{2}}(n-1)$ là giá trị tới hạn hai đuôi với bậc tự do là $(n-1)$.

Bước 4. Tính giá trị quan sát

$$t = \frac{\bar{d}}{s_{dif}} \sqrt{n}.$$

Bước 5. Kết luận

● $t \in W_\alpha$: bác bỏ giả thuyết H_0 , chấp nhận đối thuyết H_1 .

● $t \notin W_\alpha$: chưa đủ bằng chứng để bác bỏ giả thuyết H_0 , nên hiện tại tạm thời phải chấp nhận giả thuyết H_0 .

2.3. Điều kiện áp dụng kiểm định

● Kích cỡ hai mẫu so sánh phải bằng nhau và ở đây giả sử bằng n .

● n quan sát trong mẫu thứ nhất phối

hợp từng cặp với m quan sát trong mẫu thứ hai.

● Tính toán chênh lệch giữa m giá trị của hai mẫu sẽ thu được một mẫu có đúng m giá trị, chẳng hạn mẫu thu được kí hiệu là

$$d_1, d_2, \dots, d_n$$

và mẫu này bắt buộc phải có dạng phân phối chuẩn hoặc cỡ mẫu phải đủ lớn để có thể xem như xấp xỉ một phân phối chuẩn.

3. Dữ liệu quan sát và phương pháp thống kê dữ liệu

3.1. Dữ liệu quan sát

Giả sử chọn ngẫu nhiên hai giảng viên tham gia giảng dạy hai học phần trong Trường Đại học Xây dựng Miền Tây, một học phần toán (Toán cao cấp 1) và một học phần vật lý (Vật lý 1). Khi đó, tập dữ liệu được cung cấp bởi Phòng Đào tạo với 78 cặp quan sát từ học phần toán và học phần vật lý (78 sinh viên có trong hai mã lớp học khác nhau) mà do chính hai giảng viên được chọn ở trên trực tiếp tham gia giảng dạy trong năm học 2022 – 2023 [8]. Cụ thể:

● Có 40 sinh viên trong mã lớp XD22D03.

● Có 38 sinh viên trong mã lớp XD22D05.

Để việc phân tích dữ liệu được chính xác nhất có thể, trước hết ta phải có một tập dữ liệu tốt. Vì vậy, từ tập dữ liệu được cung cấp với 78 cặp quan sát, ta loại bỏ 32 cặp quan sát không hợp lý (một cặp quan sát được xem là không hợp lý nếu điểm

học phần hoặc không có điểm quá trình hoặc không có điểm thi hoặc không có cả điểm quá trình và điểm thi hoặc tồn tại trị tuyệt đối của điểm chênh lệch giữa điểm quá trình và điểm thi lớn hơn 2), ta sẽ thu được tập dữ liệu tốt với chỉ có 46 cặp quan sát (tất cả 46 sinh viên đều học tập cả học phần toán và học phần vật lý). Cụ thể:

Toán	6.5	8.1	8.7	9.3	6.9	8.0	9.9
Lý	6.6	8.4	7.0	8.1	7.8	7.8	6.9
Toán	9.5	8.5	9.9	9.2	7.8	10.0	9.6
Lý	8.3	7.5	8.6	9.2	7.9	8.9	8.6
Toán	8.9	7.7	7.3	8.7	9.5	7.1	9.1
Lý	8.5	6.0	6.0	6.4	6.0	6.0	8.5
Toán	8.1	6.5	10.0	7.1	6.8	4.4	10.0
Lý	8.2	6.6	8.6	7.0	6.5	3.8	8.6
Toán	7.6	10.0	5.5	8.2	10.0	10.0	8.6
Lý	5.3	8.7	6.3	6.5	10.0	8.5	6.4
Toán	8.4	9.0	9.0	6.2	8.4	8.2	9.3
Lý	8.8	6.2	5.2	5.4	5.5	5.7	8.0
Toán	9.1	9.7	9.3	7.2			
Lý	6.9	8.7	7.9	6.6			

3.2. Phương pháp kiểm tra dữ liệu

Ta nhận thấy rằng mẫu dữ liệu điểm toán (MATHS) và mẫu dữ liệu điểm vật lý (PHYSICS) có cùng kích thước mẫu và bằng 46. Hơn nữa, mỗi quan sát trong mẫu dữ liệu điểm toán có liên hệ tương ứng theo cặp với một quan sát trong mẫu dữ liệu điểm vật lý. Cuối cùng, phương pháp kiểm định theo cặp (paired samples t-test) sẽ được sử dụng trong phân tích dữ liệu nếu 46 quan sát của mẫu chênh lệch điểm (DISTANCE) giữa mẫu điểm toán và mẫu điểm vật lý có dạng phân phối chuẩn

[1, 2, 6]. Thật vậy (phân tích thống kê với phần mềm hỗ trợ SPSS),

Trong thống kê mô tả ở Bảng 1, ta có một số kết quả như: trị số trung bình bằng 1.1283, trung vị bằng 1.1500 và độ xiên bằng 0.382. Ta thấy rằng trị số trung bình và trung vị gần bằng nhau, độ xiên dao động trong khoảng từ -0.5 – 0.5 đến $+0.5$ – 0.5 nên DISTANCE được xem như có dạng phân phối chuẩn.

Hơn nữa, nhận thấy rằng biểu đồ tần suất của DISTANCE ở Hình 1 với một đường cong chuẩn có dạng hình chuông, có trị số trung bình bằng 1.1283 và đồ thị phân tán khá đồng đều hai bên nhưng bên phải nhỉnh hơn không đáng kể.

Hoàn toàn có thể kết hợp với một biểu đồ xác suất chuẩn Hình 2 để cung cấp thêm bằng chứng rằng DISTANCE có dạng phân phối chuẩn bởi vì các trị số quan sát và trị số mong đợi đều có xu hướng theo một đường thẳng.

Ta có thể kiểm tra lại dạng chuẩn của DISTANCE lần cuối cùng bằng phương pháp kiểm định của Kolmogorov-Smirnov và Shapiro-Wilk cho mẫu có kích cỡ trung bình từ 30 đến 300 (nếu cỡ mẫu từ 30 đến 50 thì sử dụng Shapiro-Wilk và nếu cỡ mẫu từ 50 đến 300 thì sử dụng Kolmogorov-

Smirnov). Do đó, quan sát giá trị Sig. (đọc kết quả từ Shapiro-Wilk vì cỡ mẫu bằng 46) ở cuối Bảng 2 ta thấy Sig. bằng 0.399 > 0.05, suy ra DISTANCE có dạng phân phối chuẩn.

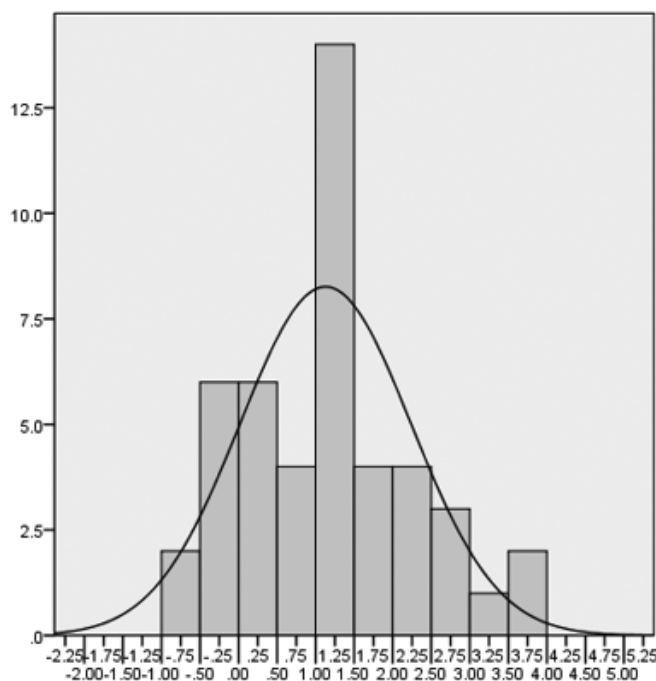
Bảng 1. Descriptive Statistics of DISTANCE

N	Valid	46
	Missing	0
Mean		1.1283
Median		1.1500
Mode		-.10 ^a
Std. Deviation		1.11069
Variance		1.234
Skewness		.382
Std. Error of Skewness		.350
Kurtosis		-.234
Std. Error of Kurtosis		.688
Minimum		-.90
Maximum		3.80
Sum		51.90

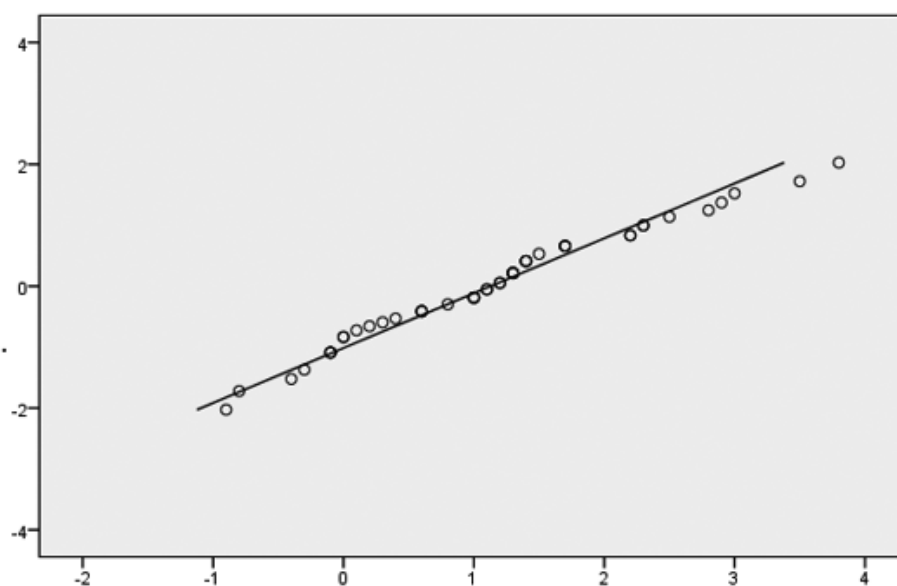
Với một số kết quả kiểm tra như thống kê mô tả ở Bảng 1, kiểm định Kolmogorov-Smirnov và Shapiro-Wilk ở Bảng 2, biểu đồ tần suất ở Hình 1, và biểu đồ xác suất chuẩn ở Hình 2 nên có thể khẳng định chắc chắn rằng DISTANCE là dạng phân phối chuẩn. Chính vì vậy, phương pháp kiểm định theo cặp (paired samples t-test) sẽ được thực hiện trong việc phân tích dữ liệu quan sát và đọc kết quả.

Bảng 2. Test of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
DISTANCE	.099	46	.200 [*]	.974	46	.399



Hình 1. Histogram of DISTANCE with normal curve



Hình 2. Normal Q-Q Plot of DISTANCE

4. Phân tích dữ liệu và đọc kết quả

Có sự khác biệt giữa điểm toán và điểm vật lý [1, 2, 6] của sinh viên Trường Đại học Xây dựng Miền Tây hay không? Bằng

cách đặt

- Giả thuyết H_0 : không có sự khác biệt giữa điểm toán và điểm vật lý,

● Đối thuyết H_1H_1 : có sự khác biệt giữa điểm toán và điểm vật lý.

Trước khi thực hiện kiểm định, biết rằng điều kiện áp dụng của kiểm định theo cặp được thỏa mãn vì đã kiểm ra chi tiết trong mục 3.2. Tiến hành thực hiện việc kiểm định (phân tích thống kê với phần mềm hỗ trợ SPSS) sẽ thu được 3 kết quả như ở Bảng 3, Bảng 4 và Bảng 5.

Thống kê mô tả đơn biến được cung cấp bởi Bảng 3, cột mean thể hiện trung bình mức điểm đánh giá của hai biến MATHS và PHYSICS. Có thể thấy rằng điểm toán bằng 8.4087 cho kết quả trung bình trội hơn điểm vật lý bằng 7.2804 là đáng kể. Từ đây có thể kết luận rằng có sự khác biệt giữa điểm toán và điểm vật lý.

Bảng 3. Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	MATHS	8.4087	46	1.33230	.19644
	PHYSICS	7.2804	46	1.33344	.19660

Bảng 4. Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	MATHS & PHYSICS	46	.653	.000

Bảng 5. Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	MATHS - PHYSICS	1.12826	1.11069	.16376	.79843	1.45809	6.890	45	.000

5. Kết luận

Điểm toán và điểm vật lý của sinh viên Trường Đại học Xây dựng Miền Tây trong trường hợp sinh viên khóa đào tạo 2022

Hơn nữa, hệ số tương quan Pearson được đưa ra trong Bảng 4 cho một cặp biến MATHS-PHYSICS với cột kết quả Sig. bằng $0.000 < 0.05$, tức là điểm toán và điểm vật lý có tương quan với nhau mà hệ số tương quan bằng 0.653 dao động trong khoảng từ $|\pm 0.5|$ đến $|\pm 1|$ nên tương quan này được xem là mạnh.

Cuối cùng, kết quả kiểm định giả thuyết được đưa ra trong Bảng 5. Quan sát giá trị Sig. ở cuối bảng ta thấy Sig. bằng $0.000 < 0.05$, suy ra bác bỏ H_0H_0 nên chấp nhận H_1H_1 . Điều này có nghĩa là có sự khác biệt giữa điểm toán và điểm vật lý có ý nghĩa thống kê.

học hai học phần toán và vật lý tương ứng lần lượt với hai giảng viên phụ trách giảng dạy được chọn trong nghiên cứu (một giảng viên dạy một học phần) này là có sự

khác biệt có ý nghĩa thống kê.

Mặc dù có sự khác biệt nhưng điểm toán và điểm vật lý đều cùng một điểm chữ đó là điểm B trong hệ thống tín chỉ bởi vì điểm toán là cận trên của điểm B và điểm lý là cận dưới của điểm B.

Vì vậy, trong cách đánh giá điểm học phần, giảng viên giảng dạy toán tăng độ khó một chút và giảng viên giảng dạy vật lý giảm độ khó một chút thì có thể điểm toán và điểm vật lý sẽ thuộc trong lân cận của chính giữa điểm B. Khi đó chắc chắn rằng điểm toán và điểm vật lý không có sự khác biệt có ý nghĩa thống kê

Điểm toán và điểm vật lý của sinh viên Trường Đại học Xây dựng Miền Tây ứng với từng trường hợp được mở rộng mà ta có thể đưa ra kết luận có hay không sự khác biệt giữa điểm toán và điểm vật lý có ý nghĩa thống kê, chẳng hạn một số trường hợp được mở rộng như sau:

TÀI LIỆU THAM KHẢO

- [1] Hoàng Trọng và Chu Nguyễn Mộng Ngọc, *Phân tích dữ liệu nghiên cứu với SPSS tập 1 và tập 2*. HCM: Nxb Hồng Đức, 2008.
- [2] Ngô Văn Thứ và Nguyễn Mạnh Thế, *Giáo trình thống kê thực hành*. Hà Nội: Nxb ĐH Kinh tế Quốc dân, 2019.
- [3] Nguyễn Văn Hữu, Đào Hữu Hổ và Hoàng Hữu Như, *Thống kê toán học*. Hà Nội: Nxb ĐHQG Hà Nội, 2004.
- [4] Nguyễn Đức Khiêm, *Toán 3*. Vĩnh Long: Trường ĐHXD Miền Tây, 2018.
- [5] Phạm Xuân Kiều, *Giáo trình xác suất và thống kê*. Hà Nội: Nxb Giáo dục, 2005.
- [6] Võ Văn Tài, *Đề tài nghiên cứu khoa học: Phân tích thống kê điểm học tập và rèn luyện của sinh viên Trường Đại học Cần Thơ*. Cần Thơ: Trường ĐH Cần Thơ, 2011.
- [7] Võ Văn Tài, *Giáo trình xử lý số liệu thống kê*. Cần Thơ: Trường ĐH Cần Thơ, 2012.
- [8] Phòng Quản lý Đào tạo, *Tập dữ liệu*, Đại học Xây dựng Miền Tây.
- [9] <http://mtu.edu.vn/>.

● Sinh viên khóa đào tạo 2022 học hai học phần toán và vật lý tương ứng với nhiều giảng viên giảng dạy (nhiều giảng viên dạy một học phần).

● Sinh viên khóa đào tạo khác học hai học phần toán và vật lý tương ứng với hai giảng viên giảng dạy (một giảng viên dạy một học phần).

● Sinh viên khóa đào tạo khác học hai học phần toán và vật lý tương ứng với nhiều giảng viên giảng dạy (nhiều giảng viên dạy một học phần).

Hơn nữa, bài báo này cũng giúp ta có những kết quả nghiên cứu tương tự để khảo sát về cách đánh giá điểm học phần thứ nhất và điểm học phần thứ hai tương ứng lần lượt với bất kỳ một giảng viên giảng dạy học phần thứ nhất và một giảng viên giảng dạy học phần thứ hai trong Trường Đại học Xây dựng Miền Tây.