

Sử dụng entropy xấp xỉ để so sánh tính ngẫu nhiên của các chuỗi dữ liệu trên thị trường chứng khoán các nước Asean

Trần Thị Tuấn Anh

Tóm tắt—Bài viết sử dụng giá đóng cửa hàng ngày thu thập từ Datastream trong giai đoạn từ tháng 01/2000 cho đến tháng 12/2016 của các chỉ số chứng khoán trên thị trường các quốc gia ASEAN để tính toán entropy xấp xỉ theo thuật toán do Pincus (2008) đề xuất. Entropy xấp xỉ trong bài viết này được dùng để đo lường tính ngẫu nhiên trong biến động chuỗi thời gian chứng khoán ở các quốc gia ASEAN. Kết quả tính toán trên toàn bộ dữ liệu cho thấy rằng chuỗi tỷ suất sinh lợi biến động mang tính ngẫu nhiên cao hơn rất nhiều so với chuỗi chỉ số chứng khoán và Singapore là quốc gia được xem là có tính ngẫu nhiên trong biến động các chuỗi thời gian trên thị trường chứng khoán cao nhất. Indonesia là quốc gia có tính ngẫu nhiên trong biến động chỉ số chứng khoán là thấp nhất. Trong giai đoạn sau khủng hoảng, sự cải thiện trong tính ngẫu nhiên của thị trường Việt Nam được thể hiện rõ rệt. Philippines trở thành quốc gia có tiềm năng cho các nhà đầu tư dự đoán biến động chứng khoán và tìm kiếm cơ hội kinh doanh chênh lệch giá thu lợi nhuận bất thường.

Từ khóa—Tính ngẫu nhiên của chuỗi thời gian, entropy xấp xỉ, tính hiệu quả của thị trường chứng khoán, dự đoán chỉ số chứng khoán, tính hình mẫu.

1 GIỚI THIỆU

Bachelier [1] là người đầu tiên đề xuất ý tưởng giá chứng khoán biến động tuân theo chuyển động Brown. Ý tưởng này hàm việc giá chứng khoán phản ánh hết các thông tin khả dụng trên thị trường. Ý tưởng này được thể hiện rõ ràng hơn trong giả thuyết thị trường hiệu quả đề xuất bởi [2], theo đó, giá hiện tại của chứng khoán chính là dự báo tốt nhất cho giá chứng khoán trong tương lai. Khi đó, sự thay đổi giá chứng khoán có thể được mô tả bằng một chuỗi nhiễu trắng, cũng đồng nghĩa với việc chuỗi giá chứng khoán tuân

theo biến động của một bước ngẫu nhiên (random walk).

Tuy nhiên, đã có rất nhiều các nghiên cứu cho thấy bằng chứng chống lại giả thuyết thị trường hiệu quả của [2]. [3] nhận định rằng các chuỗi tỷ suất sinh lợi có “trí nhớ lâu dài” (long memory) và có thể được mô hình hóa bằng một chuyển động Brown phân dạng (fractal Brownian motion). [4] cung cấp bằng chứng cho thấy khả năng phân dạng của các chuỗi thời gian tài chính bằng cách dùng đại lượng Hurst (Hurst exponent) để đo lường tính bền theo thời gian của dữ liệu. Nhiều mô hình cho phép sự biến động của phương sai theo thời gian cũng được sử dụng để mô tả tính không ngẫu nhiên của chuỗi thời gian tài chính như mô hình ARCH [5], mô hình chuyển trạng thái Markov [6], kiểm định tỷ số phương sai [7].

Bên cạnh những công cụ kiểm định truyền thống, sự phát triển của Kinh tế học vật lý (Econophysics) – lĩnh vực ứng dụng các khái niệm và cách tiếp cận trong vật lý vào phân tích các mô hình động phức tạp trong tài chính - đã giúp cho các nhà nghiên cứu có thêm nhiều công cụ để kiểm định tính ngẫu nhiên trong các chuỗi thời gian tài chính. Trong số đó, entropy và các mở rộng của entropy được xem là một hướng ứng dụng nhiều tiềm năng nhất. Entropy vốn là khái niệm dùng để mô tả sự biến động hỗn độn trong nhiệt động lực học. Một hệ vật lý chuyển động càng hỗn độn thì entropy của hệ càng lớn và ngược lại. Nếu biến động của các chuỗi dữ liệu trên thị trường chứng khoán hoàn toàn ngẫu nhiên thì cũng có thể được xem như có tính tương đồng với biến động hỗn độn của các hệ vật lý. Do vậy, ngày càng nhiều các nhà nghiên cứu vận dụng entropy trong việc kiểm định tính ngẫu nhiên của

Ngày nhận bản thảo: 01-9-2018; Ngày chấp nhận đăng: 7-11-2018; Ngày đăng: 31-12-2018

Tác giả Trần Thị Tuấn Anh, công tác tại Trường Đại học Kinh tế TP.HCM (Email: anhttt@ueh.edu.vn).

các chuỗi dữ liệu thị trường. Khái niệm entropy ngày càng được mở rộng, từ những khái niệm cơ bản như Shannon entropy, Tsallis entropy, Renyi entropy... đến những khái niệm phức tạp hơn như entropy hoán vị (permutation entropy), entropy xấp xỉ (approximate entropy), entropy mẫu (sample entropy), entropy đa hướng (multiscaled entropy).v.v... Mỗi đại lượng entropy đều có những điểm mạnh riêng khi khai thác và kiểm định sự ngẫu nhiên của chuỗi thời gian thông qua kiểm tra tính lặp lại của các hình mẫu.

Bài viết này lựa chọn giới thiệu khái niệm entropy xấp xỉ và ứng dụng entropy xấp xỉ để so sánh tính ngẫu nhiên của các chuỗi giá chứng khoán và tỷ suất sinh lợi trên thị trường chứng khoán Việt Nam cũng như thị trường chứng khoán của các quốc gia Đông Nam khác như Philippines, Malaysia, Indonesia, Thái Lan và Singapore để có cơ sở nhận định về mức độ ngẫu nhiên giữa các thị trường chứng khoán của các quốc gia. Với mục tiêu như vậy, bài viết này được cấu trúc như sau: Mục 2 của bài viết thực hiện tổng quan một số nghiên cứu có liên quan đến việc ứng dụng entropy xấp xỉ đối với chuỗi thời gian tài chính; Mục 3 trình bày khái niệm entropy xấp xỉ và phương pháp tính toán entropy xấp xỉ và sử dụng entropy xấp xỉ để so sánh sự ngẫu nhiên của chuỗi thời gian; Mục 4 trình bày kết quả nghiên cứu và các thảo luận; Mục 5 kết luận và đề xuất một số hàm ý từ kết quả nghiên cứu.

2 TỔNG QUAN LÝ THUYẾT

Cùng với sự phát triển của kinh tế học vật lý và những kết quả khả quan khi ứng dụng entropy vào nghiên cứu tài chính, nhiều nhà nghiên cứu đã đề xuất áp dụng khái niệm entropy xấp xỉ (Approximate entropy - ApEn) để khảo sát tính ngẫu nhiên của các dữ liệu chứng khoán trên thị trường.

[8] sử dụng entropy hoán vị để đặc trưng hóa mức độ ngẫu nhiên cũng như khám phá sự bất quy tắc trong các hệ sinh lý. [9] sử dụng trong nghiên cứu tính ngẫu nhiên của nhịp tim hoặc [10] ứng dụng trong nghiên cứu động lực EEG. [11] nhận xét rằng việc tính toán entropy hoán vị không phụ thuộc vào các mô hình thống kê nhưng vẫn có thể được sử dụng kết hợp trong các nghiên cứu dựa

trên mô hình.

[12] sử dụng complexity-entropy như một công cụ hữu hiệu để kiểm tra tính ngẫu nhiên của các chuỗi dữ liệu thị trường và phân chia thị trường thành các giai đoạn phát triển khác nhau. Các tác giả ứng dụng dữ liệu chứng khoán của 32 quốc gia trên thế giới và nhận định rằng cách tiếp cận thông qua complexity-entropy giúp dễ dàng phân biệt các giai đoạn phát triển của thị trường chứng khoán. Sự khác nhau giữa các thị trường chứng khoán mới nổi cũng như các thị trường phát triển có thể dễ dàng được nhận thấy với công cụ hữu hiệu này.

[13] đề xuất sử dụng entropy khuếch tán để phân tích tính ổn định của thị trường chứng khoán và áp dụng thực nghiệm với chỉ số chứng khoán công nghiệp Dow Jones (Mỹ). Kết quả cho thấy sự hiệu quả vượt trội của phương pháp entropy khuếch tán so với các phương pháp khác khi phản ánh được mức độ biến động và các trường hợp cực trị của thị trường.

Như vậy, có thể có nhiều khái niệm entropy khác nhau được sử dụng để đo lường mức độ ngẫu nhiên trong chuỗi thời gian. Bài viết này dựa trên một phần cách tiếp cận bằng entropy hoán vị của [14] để đo lường và so sánh mức độ ngẫu nhiên trong chuỗi thời gian chứng khoán của thị trường các quốc gia ASEAN. Thứ nhất, do các khái niệm entropy vận dụng vào trong nghiên cứu kinh tế và tài chính ở Việt Nam còn khá mới mẻ nên bài viết này hướng đến thử nghiệm các khái niệm này trong điều kiện Việt Nam. Thứ hai, việc so sánh tính ngẫu nhiên trong chuỗi thời gian tài chính của các thị trường cũng hàm ý mức độ về tính hiệu quả của thông tin. Khi thị trường đạt trạng thái hiệu quả thông tin, các biến động về giá chứng khoán hoặc tỷ suất sinh lợi trên thị trường hoàn toàn mang tính ngẫu nhiên và không thể dự đoán được bằng một hình mẫu nào để có thể mang lại cơ hội kinh doanh chênh lệch giá thu lợi nhuận bất thường. Thứ ba, các nghiên cứu hiện tại hầu hết dùng cách tiếp cận định lượng thông qua các mô hình thống kê truyền thống. Hướng tiếp cận mới mẻ thông qua entropy có thể được sử dụng như một nguồn thông tin bổ sung giúp đối chiếu và so sánh các kết quả đạt được để nhà đầu tư có những quyết định kinh doanh hợp lý hơn.

Bảng I. Các chỉ số chứng khoán của các quốc gia Đông Nam Á

Quốc gia	Chỉ số chứng khoán	Diễn giải
Việt Nam	VN-Index	Vietnam Stock index
Philippines	FTWIPHLL	FTSE Philippines Index FTSE Bursa
Malaysia	FBMKLCI	Malaysia KLCI Index Jakarta Stock
Indonesia	JCT	Exchange Composite Index
Thái Lan	SET	Stock Exchange of Thailand SET Index
Singapore	STI	FTSE Straits Times Index

3 DỮ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

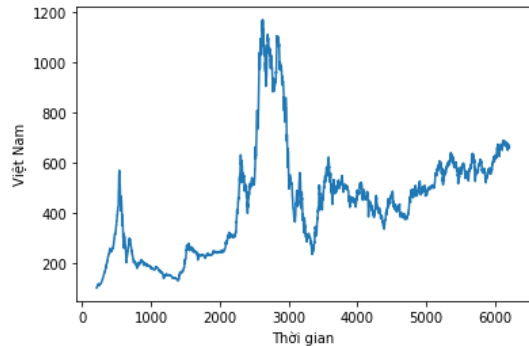
3.1 Dữ liệu

Với cách tiếp cận vận dụng entropy xấp xỉ để đo lường tính ngẫu nhiên của chuỗi thời gian tài chính, bài viết sử dụng dữ liệu về chuỗi chỉ số chứng khoán và tỷ suất sinh lợi chứng khoán hàng ngày trên thị trường chứng khoán Việt Nam và thị trường chứng khoán Đông Nam Á. Các chuỗi chỉ số chứng khoán sử dụng được liệt kê trong bảng I.

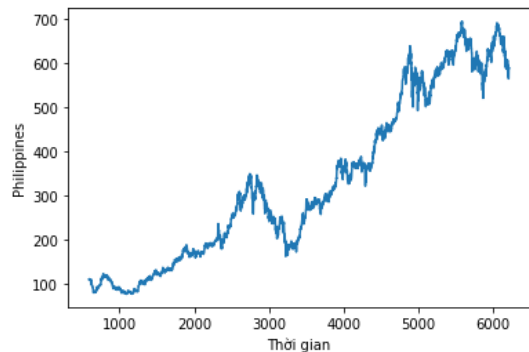
Chỉ số chứng khoán hàng ngày của 6 quốc gia Đông Nam Á được thu thập từ nguồn Datastream trong thời gian từ tháng 01 năm 2000 đến tháng 12 năm 2016. Các tính toán entropy xấp xỉ và vẽ đồ thị trên dữ liệu được thực hiện với sự hỗ trợ của phần mềm Python.

Biến động chỉ số chứng khoán của các quốc gia Đông Nam Á được thể hiện trên đồ thị từ hình 1 đến hình 6. Trong giai đoạn dữ liệu được thu thập, có thể thấy rằng chỉ số chứng khoán các quốc gia đều có xu hướng tăng dần trong dài hạn nhưng biến động rất khó dự đoán trong ngắn hạn. Đặc biệt, trong giai đoạn khủng hoảng kinh tế 2008 – 2009, chỉ số chứng khoán các quốc gia đều có sự sụt giảm đáng kể và có xu hướng gia tăng trở lại sau khủng hoảng. Tuy nhiên, đồ thị không thể hiện được chính xác tính ngẫu nhiên của chuỗi chỉ số chứng khoán, do đó không thể dựa vào đồ thị để kết luận chuỗi giá chứng khoán ở thị trường chứng khoán quốc gia nào biến động ngẫu nhiên

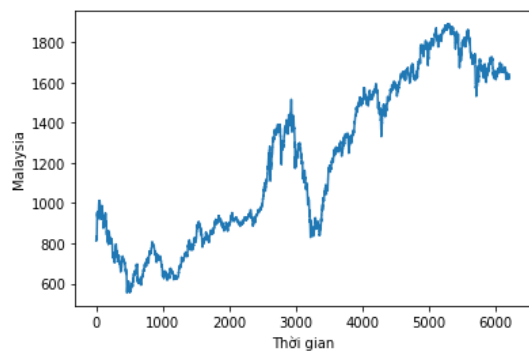
hơn. Việc so sánh độ ngẫu nhiên trong các chuỗi tài chính của mỗi quốc gia cũng hàm ý về tính hiệu quả thông tin của thị trường. Thị trường chứng khoán càng hiệu quả, thì chuỗi thời gian tài chính càng trở nên ngẫu nhiên vì không có tính hình mẫu và khó, hoặc không thể, dự báo được biến động để có cơ hội kinh doanh chênh lệch giá thu lợi nhuận phi rủi ro.



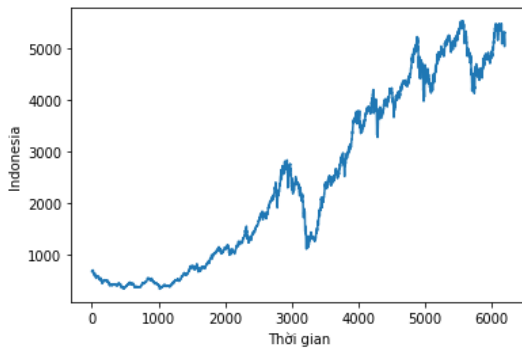
Hình 1. Biểu đồ chỉ số VN-Index của thị trường chứng khoán Việt Nam



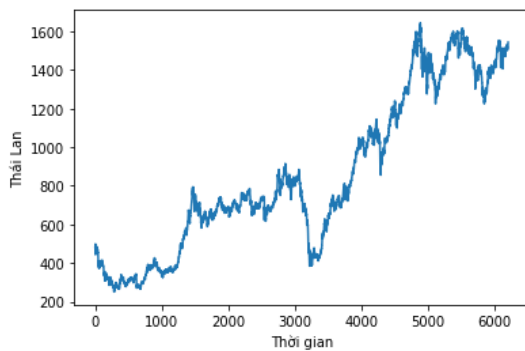
Hình 2. Biểu đồ chỉ số FTWIPHLL của thị trường chứng khoán Philippines



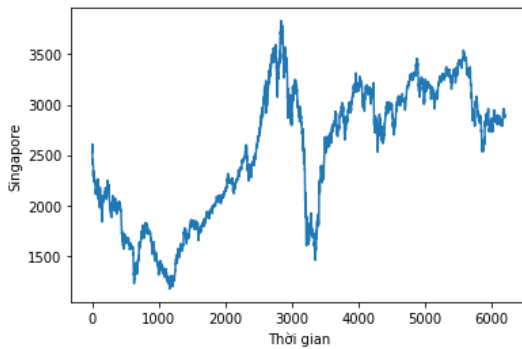
Hình 3. Biểu đồ chỉ số FBMKLCI của thị trường chứng khoán Malaysia



Hình 4. Biểu đồ chỉ số JCT của thị trường chứng khoán Indonesia



Hình 5. Biểu đồ chỉ số SET của thị trường chứng khoán Thái Lan



Hình 6. Biểu đồ chỉ số STI của thị trường chứng khoán Singapore

3.2 Phương pháp nghiên cứu

Entropy xấp xỉ là đại lượng ước lượng khả năng mà các hình mẫu tương tự nhau của một nhóm quan sát nhưng lại được nối tiếp theo bởi các hình mẫu không còn tương tự nhau nữa. Vì vậy, những chuỗi thời gian có chứa những hình mẫu lặp lại nhiều thường sẽ có entropy xấp xỉ khá nhỏ so với một chuỗi ngẫu nhiên hoàn toàn.

Entropy xấp xỉ được giới thiệu lần đầu tiên bởi [15] và cách tính toán được Pincus làm rõ trong các nghiên cứu tiếp theo sau đó. [14] đề xuất các

bước tính entropy xấp xỉ một cách rõ ràng như một thuật toán tin học để có thể cài đặt dễ dàng. Xét một dãy số liệu y_1, y_2, \dots, y_N cùng với số nguyên dương m , được gọi là độ dài, và số thực dương r , được gọi là mức lọc. Thuật toán để tính entropy xấp xỉ được thực hiện theo các bước như sau:

1. Lập một vector Y_i gồm m phần tử $Y_i = (y_i, y_{i+1}, \dots, y_{m+i-1})$ trong đó $i = 1, 2, \dots, N - m + 1$. Khoảng cách giữa hai vector Y_i và Y_j với $i, j = 1, 2, \dots, N - m + 1$ được định nghĩa như sau:

$$d(Y_i, Y_j) = \max_s |y_{is} - y_{js}|$$

trong đó y_{is} là phần tử thứ s của vector Y_i và y_{js} là phần tử thứ s của vector Y_j .

2. Xét vector Y_i với $i = 1, 2, \dots, N - m + 1$. Gọi B_i là tập hợp các vector $Y_j, j = 1, 2, \dots, N - m + 1$, sao cho khoảng cách từ Y_i đến Y_j không quá r . Nghĩa là

$$B_i = \{Y_j, j = 1, 2, \dots, N - m + 1 \mid d(Y_i, Y_j) \leq r\}$$

Gọi

$$p_{B_i} = \frac{n_{B_i}}{N - m + 1}$$

Trong đó n_{B_i} là số phần tử thuộc B_i và $N - m + 1$ là tổng số vector có thể có.

Khi đó p_{B_i} chính là tỷ lệ số vector được xem là gần với Y_i . Với cách tính khoảng cách như trên, việc vector Y_i gần với vector Y_j cũng có nghĩa là các thành phần của Y_i gần với các phần tử của Y_j . Từ đó suy ra hình mẫu biến động của Y_i và Y_j là tương tự nhau với mức độ tương tự cho phép bởi biên độ r .

Với N đủ lớn, p_{B_i} có thể xem là xác suất để một vector Y_j có hình mẫu biến động gần với Y_i .

3. Gọi

$$\Phi_m(r) = \frac{\sum_{i=1}^{N-m+1} \ln p_{B_i}}{N - m + 1}$$

Và entropy xấp xỉ, ký hiệu là ApEn, được tính bằng cách:

$$ApEn = \Phi_m(r) - \Phi_{m+1}(r)$$

Trong khi $\Phi_m(r)$ đo lường mức độ tương tự trong các hình mẫu có độ dài m thì $\Phi_{m+1}(r)$ sẽ đo lường mức độ tương tự các hình mẫu có độ dài $m+1$. Nếu mọi vector Y_i có độ dài m có tỷ lệ số hình mẫu tương tự là n_{B_i} vẫn tiếp tục giữ được số lượng các vector có hình mẫu tương tự với độ dài $m + 1$, thì khi đó $ApEn$ bằng 0. Nghĩa là, nếu $ApEn$ càng lớn, thì có nghĩa là sự thay đổi về mức độ tương tự hình mẫu ở độ dài $m + 1$ càng nhiều khi so với độ dài m . Nói tóm lại, chuỗi càng có tính quy luật thì entropy xấp xỉ tính toán được càng nhỏ; chuỗi càng ngẫu nhiên, càng ít có tính quy luật thì entropy xấp xỉ tính được càng tăng.

Thuật toán trên được xây dựng giúp việc hiểu biết và tính toán entropy xấp xỉ được trực quan hơn, được phát triển trong [16]. Trong tính toán entropy xấp xỉ, mặc dù về mặt lý thuyết, có thể thực hiện với một độ dài m bất kỳ. Tuy nhiên [17] chỉ ra rằng khi kiểm định tính ngẫu nhiên của chuỗi trong thực nghiệm, không cần thiết phải xét những trường hợp $m > 4$.

Theo [14], $ApEn$ gần như không bị ảnh hưởng bởi các nhiễu có độ lớn dưới mức lọc r . Trong thực nghiệm, mức lọc cũng thường được xác định trong khoảng từ 0,1 đến 0,25 lần độ lệch chuẩn (theo [14]). Bài viết này dùng mức lọc là $0,2 * sd$ với sd là độ lệch chuẩn của chuỗi thời gian. $ApEn$ ổn định và không nhạy cảm với các quan sát bất thường nếu chúng xảy ra không thường xuyên. [14] cũng thảo luận về những điểm mạnh của entropy xấp xỉ và tiềm năng ứng dụng của đại lượng này trong kinh tế lượng. Cách tính toán entropy xấp xỉ không phụ thuộc vào việc lập mô hình mà được xác định bởi các phân phối tần số (chung). Entropy xấp xỉ được áp dụng cho từng chuỗi đơn biến, không cần thiết lập mô hình. Ngoài ra, $ApEn$ hữu ích để đánh giá liệu dữ liệu chuỗi thời gian có thỏa mãn đặc điểm của một quá trình cụ thể nào đó hay không (ví dụ, đặc điểm ngẫu nhiên "ngẫu nhiên"). Các chuỗi có giá trị entropy xấp xỉ không dưới 80% giá trị tối đa của một chuỗi hoàn toàn ngẫu nhiên thì có thể xem là ngẫu nhiên. Entropy xấp xỉ còn có thể được áp

dụng để đánh giá sự ổn định hệ thống; sự gia tăng đáng kể các giá trị entropy xấp xỉ có thể báo trước các thay đổi trạng thái rõ rệt.

4 KẾT QUẢ VÀ THẢO LUẬN

4.1 Thống kê mô tả dữ liệu

Bảng II cho thấy kết quả mô tả giá trị trung bình trong cả giai đoạn 2000 – 2016 của tỷ suất sinh lợi hàng ngày của thị trường chứng khoán các quốc gia ASEAN. Kết quả mô tả cho thấy Singapore là quốc gia có tỷ suất sinh lợi trung bình thấp nhất trong số sáu quốc gia ASEAN. Điều này có thể thấy qua đồ thị ở Hình 6 khi mà chỉ số chứng khoán trong năm 2016 của Singapore không cao hơn nhiều so với những năm đầu thế kỷ 21. Việt Nam, Philippines và Indonesia là các quốc gia có tỷ suất sinh lợi trung bình hàng ngày cao nhất.

Bảng II. Tỷ suất sinh lợi chứng khoán trung bình trong giai đoạn 2000-2016

Quốc gia	Tỷ suất sinh lợi			Độ lệch chuẩn
	Trung bình	Nhỏ nhất	Lớn nhất	
Việt Nam	0,030	-13,276	10,303	1,14
Philippines	0,031	-7,722	6,669	1,27
Malaysia	0,011	-9,979	4,503	0,68
Indonesia	0,033	-10,954	7,623	1,14
Thái Lan	0,019	-16,063	10,577	1,12
Singapore	0,002	-8,696	7,531	0,97

Nguồn: tác giả tính toán từ số liệu thu thập được

Đề có thể thấy được diễn biến thị trường theo thời gian, Bảng III tiến hành mô tả tỷ suất sinh lợi trung bình hàng ngày của chỉ số chứng khoán các quốc gia theo từng năm. Đáng chú ý nhất trong kết quả mô tả ở bảng III là việc tỷ suất sinh lợi trung bình của cả 6 quốc gia trong năm 2008 đều mang dấu âm, cho thấy sự ảnh hưởng tiêu cực của khủng hoảng kinh tế đến thị trường chứng khoán của các quốc gia. Cũng có thể thấy rằng năm 2009 là năm phục hồi của thị trường khi mà tất cả các quốc gia đều có tỷ suất sinh lợi cao hơn hẳn những năm trước và sau đó thị trường các quốc gia trở lại mức bình thường. Vì năm 2008 có sự khác biệt với các năm khác, đánh dấu năm tác động rõ rệt của cuộc đại khủng hoảng, nên bài viết

này, ngoài việc tính toán entropy xấp xỉ trên toàn bộ mẫu dữ liệu của các quốc gia để so sánh tính ngẫu nhiên của chuỗi chỉ số chứng khoán, còn tính toán entropy xấp xỉ cho từng giai đoạn trước

và khủng hoảng. Những năm trước 2008 được tính vào giai đoạn trước khủng hoảng (pre - crisis) và những năm sau năm 2008 được xem là giai đoạn sau khủng hoảng (post – crisis).

Bảng III. Tỷ suất sinh lợi trung bình theo năm ở các quốc gia

Năm	Quốc gia					
	Việt Nam	Philippines	Malaysia	Indonesia	Thái Lan	Singapore
2000	0,4626		-0,0489	-0,1332	-0,1596	-0,0673
2001	0,0348	-0,1047	0,0066	-0,0165	0,0332	-0,0521
2002	-0,0685	-0,0540	-0,0203	0,0221	0,0438	-0,0621
2003	-0,0251	0,1226	0,0564	0,1336	0,2118	0,0777
2004	0,0979	0,0724	0,0365	0,1007	-0,0395	0,0396
2005	0,0695	0,0444	-0,0023	0,0412	0,0181	0,0360
2006	0,2446	0,1016	0,0541	0,1206	-0,0133	0,0676
2007	0,0573	0,0368	0,075	0,1149	0,0638	0,0471
2008	-0,2940	-0,1410	-0,136	-0,1929	-0,1764	-0,1849
2009	0,1230	0,1194	0,102	0,1715	0,1343	0,1364
2010	-0,0056	0,0757	0,048	0,1039	0,0934	0,0263
2011	-0,0878	-0,0015	0,002	0,0086	-0,0020	-0,0512
2012	0,0443	0,0829	0,026	0,0333	0,0835	0,0491
2013	0,0544	0,0016	0,027	-0,0027	-0,0190	0,0000
2014	0,0214	0,0521	-0,016	0,0551	0,0391	0,0166
2015	0,0161	-0,0099	-0,010	-0,0354	-0,0413	-0,0424
2016	0,0378	-0,0037	-0,008	0,0389	0,0493	-0,0002

Nguồn: tác giả tính toán từ số liệu thu thập được

4.2 Kết quả nghiên cứu

Với dữ liệu về chỉ số chứng khoán hàng ngày trên thị trường các quốc gia ASEAN thu thập trong giai đoạn 2000 - 2016, bài viết tính tỷ suất sinh lợi hàng ngày và áp dụng phương pháp tính entropy xấp xỉ như đã mô tả ở mục 3 để đạt được kết quả nghiên cứu. Trước hết, bài viết tiến hành tính entropy xấp xỉ với độ dài m lần lượt bằng 2, 3 và 4 trên toàn bộ mẫu dữ liệu đối với chuỗi chỉ số chứng khoán thị trường và chuỗi tỷ suất sinh lợi hàng ngày của thị trường, kết quả được thể hiện ở bảng IV. Sau đó, bài viết chia dữ liệu thành hai mẫu con: giai đoạn trước khủng hoảng và giai đoạn sau khủng hoảng; kết quả tính toán tương ứng được thể hiện ở bảng V và bảng VI.

Bảng IV thể hiện kết quả tính toán entropy hoán vị với chuỗi chỉ số chứng khoán và chuỗi tỷ suất sinh lợi ứng với độ dài m lần lượt nhận giá trị 2, 3 và 4. Tương tự như các nghiên cứu thông

thường khác, mức lọc mặc định được sử dụng là $0,2*sd$ với sd là độ lệch chuẩn của chuỗi thời gian. Với nhận định khi entropy xấp xỉ càng lớn, độ ngẫu nhiên của chuỗi thời gian càng cao, có thể thấy rằng chuỗi tỷ suất sinh lợi có độ ngẫu nhiên cao hơn rất nhiều so với chuỗi giá chứng khoán. Điều này hàm ý rằng việc dự đoán chuỗi giá có thể dễ dàng hơn rất nhiều so với chuỗi tỷ suất sinh lợi. Khi so sánh giữa các thị trường, có thể thấy Singapore là quốc gia có chuỗi thời gian chứng khoán trên thị trường với mức ngẫu nhiên cao nhất với entropy xấp xỉ tính được cao hơn các quốc gia khác trong hầu hết các trường hợp. Trong khi đó, nếu xét về tính ngẫu nhiên của chỉ số chứng khoán, thị trường Indonesia cho thấy biến động ít có tính ngẫu nhiên nhất; còn nếu xét về tính ngẫu nhiên của tỷ suất sinh lợi thì ở vị trí tính ngẫu nhiên trong biến động thấp nhất lại là Việt Nam. Tuy nhiên, tính ngẫu nhiên trong biến động chỉ số chứng khoán tính bằng entropy xấp

xi, Việt Nam chỉ đứng sau Singapore.

Bảng IV. Entropy xấp xỉ của các nước Asean trên toàn bộ mẫu

	Chỉ số thị trường (P_i)	Tỷ suất sinh lợi (R_i)	
m = 2	Việt Nam	0,0604	1,5849
	Philippines	0,0332	1,6480
	Malaysia	0,0348	1,6597
	Indonesia	0,0267	1,6838
	Thái Lan	0,0376	1,6725
	Singapore	0,0728	1,6879
m = 3	Việt Nam	0,0576	1,2318
	Philippines	0,0298	1,2825
	Malaysia	0,0310	1,3128
	Indonesia	0,0218	1,3159
	Thái Lan	0,0336	1,3194
	Singapore	0,0678	1,3248
m = 4	Việt Nam	0,0571	0,8856
	Philippines	0,0275	0,9290
	Malaysia	0,0289	0,9569
	Indonesia	0,0218	0,9510
	Thái Lan	0,0308	0,9595
	Singapore	0,0642	0,9357

Nguồn: tác giả tính toán từ số liệu thu thập được

Kết quả nghiên cứu này hàm ý rằng, thị trường Singapore có tính hiệu quả thông tin nhất trong số các quốc gia ASEAN. Điều này cũng khá hợp lý nếu đánh giá dựa trên mức độ phát triển kinh tế của quốc gia. Để dự đoán biến động trên thị trường, nếu dựa trên tính ngẫu nhiên của chuỗi thời gian chứng khoán, nhà đầu tư nên lựa chọn dự báo chuỗi chỉ số chứng khoán hơn là chuỗi tỷ suất sinh lợi vì tính hình mẫu tiềm ẩn trong chuỗi chỉ số chứng khoán cao hơn, sẽ có những hình mẫu có xu hướng lặp lại trong chuỗi và dựa vào những hình mẫu biến động lặp lại đó, nhà đầu tư có thể đưa ra dự đoán về biến động trong tương lai. Thị trường Việt Nam và thị trường Singapore là hai quốc gia có chỉ số chứng khoán biến động ngẫu nhiên nhất trong số các quốc gia ASEAN; ngược lại Indonesia là quốc gia có chỉ số chứng khoán biến động ít ngẫu nhiên nhất. Sau khi xem xét với dữ liệu toàn bộ mẫu, bài viết xem xét tính ngẫu nhiên đo lường bằng entropy xấp xỉ của giai đoạn trước và sau khủng hoảng để tìm hiểu sự thay đổi về tính hiệu quả trên thị trường các quốc

gia với biến cố khủng hoảng kinh tế.

Bảng V. Entropy xấp xỉ của các nước Asean giai đoạn trước khủng hoảng

	Chỉ số thị trường (P_i)	Tỷ suất sinh lợi (R_i)	
m = 2	Việt Nam	0,0406	1,4675
	Philippines	0,0536	1,3725
	Malaysia	0,0646	1,5857
	Indonesia	0,0269	1,6194
	Thái Lan	0,0656	1,5878
	Singapore	0,0696	1,6200
m = 3	Việt Nam	0,0379	1,0993
	Philippines	0,0494	1,0482
	Malaysia	0,0588	1,1738
	Indonesia	0,0253	1,1623
	Thái Lan	0,0577	1,1736
	Singapore	0,0622	1,1771
m = 4	Việt Nam	0,0379	0,7590
	Philippines	0,0480	0,7238
	Malaysia	0,0568	0,7874
	Indonesia	0,0253	0,7415
	Thái Lan	0,0577	0,7740
	Singapore	0,0589	0,7475

Nguồn: tác giả tính toán từ số liệu thu thập được

Trong bài viết này, dựa trên kết quả mô tả tỷ suất sinh lợi ở bảng III, bài viết chia thời gian được xét thành hai giai đoạn: giai đoạn trước khủng hoảng tính từ năm 2000 đến trước năm 2008 và giai đoạn sau khủng hoảng tính từ năm 2009 đến hết năm 2016. Bảng V thể hiện kết quả tính toán entropy xấp xỉ giai đoạn trước khủng hoảng. Kết quả của giai đoạn trước khủng hoảng khá tương đồng với kết quả tính toán trên toàn mẫu dữ liệu. Kết quả phân tích vẫn cho thấy rằng chuỗi tỷ suất sinh lợi chứng khoán biến động ngẫu nhiên hơn rất nhiều so với chuỗi giá chứng khoán và Singapore vẫn là quốc gia có các chuỗi thời gian trên thị trường chứng khoán có tính ngẫu nhiên cao nhất; và Indonesia vẫn là quốc gia có biến động chỉ số chứng khoán ít ngẫu nhiên và có tính hình mẫu lặp lại cao nhất. Tuy nhiên, trong giai đoạn trước khủng hoảng, Việt Nam cũng có chỉ số chứng khoán và tỷ suất sinh lợi có tính ngẫu nhiên thấp, chỉ sau Indonesia.

Bảng VI. Entropy xấp xỉ của các nước Asean giai đoạn sau khủng hoảng

		Chỉ số thị trường (P_t)	Tỷ suất sinh lợi (R_t)
m = 2	Việt Nam	0,1460	1,5659
	Philippines	0,0586	1,5928
	Malaysia	0,0639	1,5940
	Indonesia	0,0714	1,5880
	Thái Lan	0,0598	1,5731
	Singapore	0,1403	1,5850
m = 3	Việt Nam	0,1453	1,1379
	Philippines	0,0544	1,1599
	Malaysia	0,0591	1,1620
	Indonesia	0,0657	1,1704
	Thái Lan	0,0552	1,1672
	Singapore	0,1364	1,1611
m = 4	Việt Nam	0,1455	0,7392
	Philippines	0,0514	0,7611
	Malaysia	0,0563	0,7498
	Indonesia	0,0628	0,7573
	Thái Lan	0,0522	0,7678
	Singapore	0,1374	0,7491

Nguồn: tác giả tính toán từ số liệu thu thập được

Kết quả tính entropy xấp xỉ giai đoạn sau khủng hoảng thể hiện trong bảng VI cho thấy rằng tính ngẫu nhiên trong chuỗi giá chứng khoán của các quốc gia thay đổi rõ rệt. Cả Việt Nam và Indonesia đều cho thấy sự biến động chỉ số chứng khoán trở nên ngẫu nhiên hơn rất nhiều so với giai đoạn trước khủng hoảng. Trong khi đó, Singapore không còn là quốc gia có tính hiệu quả thông tin trên thị trường cao nhất nữa mặc dù vẫn ở mức cao so với các quốc gia khác. Việt Nam chính là quốc gia đạt được vị trí này khi xét theo entropy xấp xỉ của giai đoạn sau khủng hoảng, và Philippines đã thay thế vị trí của Indonesia, trở thành quốc gia có biến động giá chứng khoán ít ngẫu nhiên khác. Tuy nhiên, một kết quả khá trái ngược có thể nhận ra, đó là Việt Nam lại là quốc gia ít có tính hiệu quả thông tin nhất khi xét theo chuỗi tỷ suất sinh lợi mặc dù lại là quốc gia đứng đầu về tính hiệu quả thông tin trong chỉ số chứng khoán trong giai đoạn khủng hoảng. Philippines trở thành quốc gia có tiềm năng cho các nhà đầu tư dự đoán biến động chứng khoán và tìm kiếm cơ hội kinh doanh chênh lệch giá thu lợi nhuận bất thường.

5 KẾT LUẬN

Bài viết sử dụng tính toán entropy xấp xỉ theo thuật toán do Pincus (2008) đề xuất để đo lường tính ngẫu nhiên trong biến động chuỗi thời gian chứng khoán ở các quốc gia ASEAN với dữ liệu hàng ngày thu thập từ Datastream trong giai đoạn từ tháng 01/2000 cho đến tháng 12/2016. Kết quả tính toán trên toàn bộ dữ liệu cho thấy rằng chuỗi tỷ suất sinh lợi biến động mang tính ngẫu nhiên cao hơn rất nhiều so với chuỗi chỉ số chứng khoán và Singapore là quốc gia được xem là có tính ngẫu nhiên trong biến động các chuỗi thời gian trên thị trường chứng khoán cao nhất. Indonesia là quốc gia có tính ngẫu nhiên trong biến động chỉ số chứng khoán là thấp nhất. Trong giai đoạn sau khủng hoảng, sự cải thiện trong tính ngẫu nhiên của thị trường Việt Nam được thể hiện rõ rệt.

Kết quả nghiên cứu của bài viết khi sử dụng entropy hoán vị hàm ý rằng nếu cần dự toán các chuỗi thời gian trên thị trường chứng khoán các quốc gia Đông Nam Á, nhà đầu tư nên dự đoán trên chuỗi chỉ số chứng khoán thay vì chuỗi tỷ suất sinh lợi, vì tính ngẫu nhiên trong các chuỗi này thấp hơn, và tính hình mẫu lặp lại cao hơn. Thị trường càng kém hiệu quả thông tin, nhà đầu tư càng có cơ hội tìm kiếm cơ hội kinh doanh chênh lệch giá thu lợi nhuận bất thường. Vì vậy, trong giai đoạn sau khủng hoảng, thì Philippines đang là quốc gia hấp dẫn nhất ở khía cạnh này.

Kết quả nghiên cứu của bài viết cũng chỉ mới là bước đầu của việc ứng dụng khái niệm entropy xấp xỉ nói riêng và các đại lượng liên quan đến entropy nói chung vào trong nghiên cứu tính ngẫu nhiên của chuỗi thời gian tài chính, đặc biệt là ở Việt Nam. Các khái niệm entropy ứng dụng trong kinh tế đang ngày càng được phát triển mạnh mẽ. Vì vậy, những nghiên cứu mở rộng bài viết này có thể được thực hiện theo hướng thực hiện kiểm định tính ngẫu nhiên của chuỗi thời gian tài chính theo các đại lượng entropy khác để so sánh kết quả, hoặc phát triển sâu hơn nữa kiểm định tính ngẫu nhiên bằng entropy hoán vị với các kỹ thuật kiểm định nâng cao.

TÀI LIỆU THAM KHẢO

- [1] L. Bachelier, *Théorie de la spéculation*, Ph.D. Thesis, Sorbonne, Paris, 1900.
- [2] E.F. Fama, "Efficient capital markets: A review of theory and empirical work", vol. 25, pp. 383–417, 1965.
- [3] B. B. Mandelbrot, J. W. Van Ness, "Fractional Brownian motion, fractional noises and applications", *SIAM Rev.*, vol. 10, pp. 422–437, 1968.
- [4] E. Peters, *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*, John Wiley & Sons: New York, 1994.
- [5] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation", *Econometrica*, pp. 987–1007, 1982.
- [6] J. D. Hamilton, *Time series analysis*, Princeton University Press, Princeton, NJ, 1994.
- [7] A. Lo, A. C. MacKinlay, "Stock prices do not follow random walks: Evidence from a simple specification test", *Rev. Financial Stud.*, vol. 1, pp. 41–66, 1988.
- [8] J. A. Posener, C. DeBattista, J. D. Veldhuis, M. A. Province, G. H. Williams, A. F. Schatzberg, "Process irregularity of cortisol and adrenocorticotropin secretion in men with major depressive disorder", vol. 29, no. 9, pp. 1129 – 1137, 2004.
- [9] S. A. C. Schuckers, "Use of approximate entropy measurements to clarify ventricular tachycardia and fibrillation", *J. Electrocardiol.*, vol. 31, pp.101–105, 1998.
- [10] J. Bruhn, H. Ropcke, B. Rehberg, T. Bouillon, A. Hoefl, "Electroencephalogram approximate entropy correctly classifies the occurrence of burst suppression pattern as increasing anesthetic drug effect", *Anesthesiology*, vol. 93, pp. 981–985, 2000.
- [11] J. D. Veldhuis, M. L. Johnson, O. L. Veldhuiss, M. Straume, S. M. Pincus, "Impact of pulsatility attributes on the ensemble orderliness (approximate entropy) of neurohormone secretion", *Am. J. Physiol.*, vol. 281, pp. R1975–R1985, 2001.
- [12] L. Zunino, M. Zanin, B. M. Tabake, D.G. Pérez, O.A. Rosso, "Complexity-entropy causality plane: A useful approach to quantify the stock market inefficiency", *Physica A*, vol. 389, pp. 1891–1901, 2010.
- [13] S. Li, Y. Zhuang, J. He, "Stock market stability: Diffusion entropy analysis," *Physica A: Statistical Mechanics and its Applications*, Elsevier, vol. 450, no. C, pp. 462-465, 2016.
- [14] S. M. Pincus, "Approximate entropy as a measure of system complexity", *Proc. Natl. Acad. Sci. USA*, vol. 88, pp. 2297–2301, 1991.
- [15] S. M. Pincus, "Approximate Entropy as an Irregularity Measure for Financial Data", *Econometric Reviews, Taylor & Francis Journals*, vol. 27, no. 4-6, pp. 329-362, 2008.
- [16] S. M. Pincus, A. L. Goldberger, "Physiological time-series analysis: What does regularity quantify?," *Am. J. Physiol.*, vol. 266, pp. H1643–H1656, 1994.
- [17] S. Chatterjee, M. R. Yilmaz, M. Habibullah, M> Laudato, "An approximate entropy test for randomness", *Commun. Statist. – Theory Meth.*, vol. 29, pp. 655–675, 2000.

Applying approximate entropy to compare the randomness of data series in Aseans' stock markets

Tran Thi Tuan Anh

University of Economics Ho Chi Minh City
Corresponding author: anhttt@ueh.edu.vn

Received: Sept 1st, 2018; Accepted: Nov 7th, 2018; Published: Dec 31st, 2018

Abstract—The paper calculates the approximate entropy using the algorithm proposed by Pincus (2008) on the daily closing price of ASEAN countries' stock indices collected from the Datastream from January 2000 to December 2016. The approximate entropy is employed to measure the randomness of financial time series in ASEAN countries' stock markets. The results on the whole data show that the fluctuation rate of return is much

higher than the stock index and Singapore has the most stochastic time series, including stock index and its return. Indonesia's stock index exhibits the lowest randomness as suggested by approximate entropy. After crisis, the randomness of time series in the Vietnam's market is sharply enhanced and the Philippines has become a potential country for investors to seek arbitrage opportunities.

Index Terms—The randomness of time series, approximate entropy, the information efficiency of stock market, forecast stock index, pattern.