Tạp chí Khoa học Trường Đại học Cần Thơ

Số chuyên đề: Khoa học tự nhiên

website: sj.ctu.edu.vn

DOI:10.22144/ctu.jsi.2020.093

PHÂN PHỐI TỔNG CÓ TRỌNG SỐ CỦA HAI BIẾN NGẪU NHIÊN PHỤ THUỘC VÀ ỨNG DỤNG TRONG LỰA CHỌN DANH MỤC ĐẦU TƯ

Võ Văn Tài[1*], Danh Ngọc Thắm[2] và Nguyễn Ái Quỳnh[3]

[1]*Trường Đại học Cần Thơ*

[2]*Khoa Sư phạm và Xã hội nhân văn, Trường Đại học Kiên Giang*

[3]*Học viên cao học K25, Trường Đại học Cần Thơ*

**Người chịu trách nhiệm về bài viết: Võ Văn Tài (email: vvtai@ctu.edu.vn)*

Thông tin chung:
Ngày nhận bài: 04/03/2020
Ngày nhận bài sửa: 17/04/2020
Ngày duyệt đăng: 29/06/2020

Title:
Weighted distribution for sum of two dependent random variables and application in selecting the portfolio

Từ khóa:
Danh mục đầu tư, phân phối xác suất, phụ thuộc, tổng hai biến

Keywords:
Dependent, probability distribution, portfolio, sum of two variables

ABSTRACT

This article, based on the Copula theory, is to establishe the weighted density and distribution functions for sum of two dependent random variables. The research also surveyed some measures to evaluate the risk for investing in a certain stock. These measures and the established functions has been used in evaluating for a plan when investing on two stocks at the same time. The article also proposes the estimations for the parameters to apply the theory to real data. Performing for the real data in Vietnam, the research they not only illustrate the complex calculations of the present theories, but also shows the potential in the application

TÓM TẮT

Dựa vào lý thuyết Copula, bài viết này thiết lập hàm mật độ và hàm phân phối có trọng số của tổng hai biến ngẫu nhiên phụ thuộc. Nghiên cứu cũng khảo sát một số độ đo để đánh giá sự rủi ro khi đầu tư vào một cổ phiếu nào đó. Những độ rủi ro này và các hàm số đã thiết lập được sử dụng để đánh giá phương án khi đầu tư cùng lúc hai cổ phiếu. Bài viết cũng đề xuất những ước lượng cho các tham số khi áp dụng lý thuyết vào số liệu thực. Thực hiện từ số liệu thực ở Việt Nam, bài viết không chỉ minh họa những tính toán phức tạp của lý thuyết đã trình bày mà còn cho thấy tiềm năng trong ứng dụng vào tài chánh của nghiên cứu này.

Trích dẫn: Võ Văn Tài, Danh Ngọc Thắm và Nguyễn Ái Quỳnh, 2020. Phân phối tổng có trọng số của hai biến ngẫu nhiên phụ thuộc và ứng dụng trong lựa chọn danh mục đầu tư. Tạp chí Khoa học Trường Đại học Cần Thơ. 56(Số chuyên đề: Khoa học tự nhiên)(1): 54-62.

1 GIỚI THIỆU

Trong thống kê, việc xác định được phân phối xác suất của một biến ngẫu nhiên có vai trò rất quan trọng, bởi vì đó là cách tốt nhất để hiểu rõ bản chất của biến ngẫu nhiên đó. Khi có nhiều biến ngẫu nhiên, việc tìm hàm mật độ xác suất cho tổng, hiệu, tích và thương của chúng là vấn đề có nhiều ứng dụng cho nhiều lĩnh vực khác nhau (Galambos and Simonelli, 2004). Trong nghiên cứu nhiều biến, việc nghiên cứu hai biến luôn được xem là nền tảng, bởi vì từ kết quả hai biến, trường hợp nhiều biến có thể được mở rộng.

Phân phối tổng của hai biến ngẫu nhiên luôn được các nhà thống kê quan tâm vì nó được xem là nền tảng để nghiên cứu những vấn đề quan trọng trong tài chính như lựa chọn phương án đầu tư, tính độ rủi ro, hay bài toán cân bằng cung cầu của thị trường (Ekaterina *et al*., 2010). Phân phối tổng cũng liên quan đến bài toán phân loại và phân tích dữ liệu (Tai *et al*., 2016). Mặc dù được quan tâm nhiều, nhưng các nghiên cứu phần lớn chỉ tập trung cho trường hợp độc lập (Dettmann and Georgiou, 2009; Yang and Wang, 2013; Garg *et al*., 2016). Khi giả thiết hai biến độc lập, việc xây dựng các biểu thức liên quan đến tổng của hai biến ngẫu nhiên như hàm mật độ, hàm phân phối, các tham số đặc trưng đã được thiết lập khá đơn giản. Tuy nhiên việc giả sử này đã làm mất đi bản chất của nhiều vấn đề trong áp dụng thực tế (Sel *et al*., 2019). Trong nhiều trường hợp, các biến không độc lập mà có sự liên hệ mật thiết với nhau. Chẳng hạn, tổng các tỉ lệ bằng 1, do đó nếu giảm hoặc tăng một tỉ lệ thành phần thì sẽ ảnh hưởng đến các tỉ lệ khác, hay tổng số tiền không đổi được đầu tư vào các cổ phiếu, tỉ trọng đầu tư cho cổ phiếu này thay đổi sẽ ảnh hưởng đến cổ phiếu kia. Do đó có thể khẳng định rằng việc nghiên cứu cấu trúc phụ thuộc của hai cũng như nhiều biến ngẫu nhiên là đòi hỏi của thực tế.

Nghiên cứu về tổng độc lập của hai biến ngẫu nhiên đã được nhiều nhà thống kê thực hiện. Pham-Gia *et al*. (1993) đã tìm hàm mật độ xác suất cho tổng và hiệu của hai tỉ lệ để từ đó áp dụng trong các ước lượng Bayes. Garg *et al*. (2016) đã nghiên cứu nhiều vấn đề liên quan đến hai biến ngẫu nhiên, trong đó có phân phối tổng để từ đó áp dụng trong lĩnh vực tài chánh. Tai *et al*. (2018) đã xây dựng phân phối tổng của 2 tiên nghiệm để áp dụng trong bài toán phân loại hai tổng thể bằng phương pháp Bayes. Ngoài các nghiên cứu kể trên, có rất nhiều nghiên cứu liên quan đến lý thuyết và ứng dụng của tổng hai biến ngẫu nhiên, tuy nhiên các ứng dụng này chỉ xét trong điều kiện độc lập.

Sklar (1959) đã đề xuất một lý thuyết mới để đo cấu trúc phụ thuộc của các đại lượng được gọi là Copula. Copula đã nhanh chóng thu hút được sự quan tâm rất lớn của cộng đồng thống kê thế giới để trở thành một mảng nghiên cứu quan trọng của thống kê nhiều chiều. Nó cũng trở thành một công cụ mạnh, không thể thiếu trong các nghiên cứu liên quan đến kinh tế và tài chánh (Cherubini *et al*., 2004, 2011; Tang, 2014; Hien *et al*, 2015, 2017). Ứng dụng Copula để mở rộng từ cấu trúc phụ thuộc sang cấu trúc độc lập của tổng hai biến ngẫu nhiên đã được nghiên cứu bởi Sel (2019). Nghiên cứu này cũng áp dụng lý thuyết vào việc lựa chọn danh mục đầu tư. Tuy nhiên vì Copula có rất nhiều họ, nên nghiên cứu này cũng xem xét một số họ cụ thể. Hơn nữa nghiên cứu chỉ xem xét ứng dụng cho một tập dữ liệu cụ thể mà không phải là tổng quát.

Trong bài viết này, sau khi tổng kết lại lý thuyết liên quan đến Coupla, hàm mật độ và hàm phân phối cho tổng có trọng số của hai biến ngẫu nhiên phụ thuộc được xây dựng. Áp dụng kết quả này, các bước xây dựng và vấn đề tính toán trong việc lựa chọn phương án đầu tư tối ưu trong hai danh mục được chọn lựa. Một áp dụng cụ thể với số liệu thực được thực hiện để minh họa cho các lý thuyết phức tạp được trình bày. Ứng dụng này cũng cho thấy tiềm năng ứng dụng thực tế trong tài chánh của nghiên cứu khi có thể áp dụng tương tự cho nhiều vấn đề khác.

2 PHÂN PHỐI TỔNG CÓ TRỌNG SỐ CỦA HAI BIẾN NGẪU NHIÊN PHỤ THUỘC

2.1 Copula

Định nghĩa 1. Đặt $\mathrm{I}=\left[0,1\right]$ và $\mathrm{I}^2=\left[0,1\right]^2$, khi đó Copula hai biến là một hàm $C(.,.): \mathrm{I}^2 \to \mathrm{I}$ thỏa mãn ba điều kiện sau:

i) $C(0,v)=C(u,0)=0,\ \forall u,v\in \mathrm{I}$,

ii) $C(1,v)=v, C(u,1)=u,\ \forall u,v\in \mathrm{I}$,

iii) Với mọi $u_1,u_2,v_1,v_2\in \mathrm{I}$ sao cho $u_1<u_2$ và $v_1<v_2$, ta có

$$C(u_2,v_2)-C(u_2,v_1)-C(u_1,v_2)+C(u_1,v_1)\ge 0.$$

Từ định nghĩa này, các nhà thống kê đã đưa ra rất nhiều họ Copula cụ thể khác nhau. Trong áp dụng, các họ Copula sau thường được đề cập:

Họ Copula Gauss:

$$C_r(u,v)=\frac{1}{2\pi\sqrt{1-r^2}}\int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)}\exp\left(-\frac{s^2-2rst+t^2}{2(1-r^2)}\right)dsdt,$$

Họ Copula Student:

$$C_r \; u,v \; = \frac{1}{2\pi\sqrt{1-r^2}} \int_{-\infty}^{t_v^{-1}\; u} \int_{-\infty}^{t_v^{-1}\; v} \left(1 + \frac{s^2 - 2rst + t^2}{v \; 1 - r^2}\right)^{v+2\;/2} dsdt,$$

trong đó trong đó r là hệ số tương quan Pearson, $\left|r\right| < 1$ và $v > 2$ là bậc tự do.

Họ Copula Gumbel:

$$C_\theta \; u,v \; = \exp\left(-\left[\; -\ln u \;^{\theta} + \; -\ln v \;^{\theta}\right]^{\frac{1}{\theta}}\right) dsdt,$$

$\theta \geq 1.$

Họ Copula Clayton:

$$C_\theta \; u,v \; = \max \; u^{-\theta} + v^{-\theta} - 1, 0 \;^{-\frac{1}{\theta}},$$

$\theta \in \left[-1, +\infty \right.$.

Họ Copula Frank:

$$C_\theta \; u,v \; = -\frac{1}{\theta}\ln\left(1 + \frac{e^{-\theta u} - 1 \;\; e^{-\theta v} - 1}{e^{-\theta} - 1}\right),$$

Định lý 1. (Sklar, 1959) Giả sử $H\left(x,y\right)$ là một hàm phân phối xác suất hai chiều có phân phối lề là $F\left(x\right)$ và $G\left(y\right)$. Khi đó, tồn tại một Copula $C\left(.,.\right)$ sao cho với mọi $\theta \in \mathbb{R} \setminus \; 0$. $x, y \in \mathbb{R}$ thỏa

$$H \; x,y \; = C \; F \; x \; ,G \; y \;\; .$$

Nếu các và $G\left(y\right)$. là các hàm phân phối liên tục thì Copula $C\left(.,.\right)$ là duy nhất. Ngược lại, nếu C là $F \; x$ một Copula và $G\left(y\right)$. là các hàm phân phối một chiều, khi đó hàm số xác định bởi $H\left(x,y\right) = C\left(F_1\left(x\right), G\left(y\right)\right)$ là một hàm phân phối xác suất hai chiều với các lề F(x) và G(y).

Chứng minh định lý này được trình bày chi tiết trong Sklar (1959).

2.2 Hàm mật độ và hàm phân phối xác suất

Định lý 2. Giả sử $\left(x_1, x_2\right)$ là một vectơ $F \; x$, biến ngẫu nhiên liên tục có hàm phân phối lề tương ứng là và $F_2(x)$. Gọi C là Copula mô tả cấu trúc phụ thuộc của X_1 và X_2. Đặt

$$Y = w_1 X_1 + w_2 X_2, \text{ với } w_1, w_2 \in \mathbb{R} \setminus \left\{0\right\}. \quad (1)$$

Khi đó hàm mật độ $f_Y\left(y\right)$ và hàm phân phối xác suất $F_Y\left(y\right)$ của Y được xác định như sau:

$$f_Y\left(y\right) = \frac{1}{\left|w_2\right|}\int_0^1 c\left(u, F_2\left(\frac{y - w_1 F_1^{-1}\left(u\right)}{w_2}\right)\right) f_2\left(\frac{y - w_1 F_1^{-1}\left(u\right)}{w_2}\right) du. \qquad (2)$$

$$F_Y\left(y\right) = 1_{\left\{w_2 < 0\right\}} + \operatorname{sgn}\left(w_2\right)\int_0^1 \frac{\partial}{\partial u} C\left(u, F_2\left(\frac{y - w_1 F_1^{-1}\left(u\right)}{w_2}\right)\right) du. \qquad (3)$$

trong đó

1_A là hàm $F_1(x)$ chỉ số của tập

c là hàm mật độ của Copula C,

$F_1^{-1}\left(u\right)$ là hàm tựa ngược của $F_1\left(u\right)$,

$\operatorname{sgn}\left(w_2\right)$ là một hàm dấu của w_2,

$$\operatorname{sgn}\left(w_2\right) = \begin{cases} 1, & \text{khi } w_2 > 0. \\ -1, & \text{khi } w_2 < 0. \end{cases}$$

Chứng minh.

Gọi F và f lần lượt là hàm phân phối và hàm mật độ đồng thời của X_1, X_2.

$$\text{Đặt} \begin{cases} Y_1 = w_1 X_1 + w_2 X_2. \\ Y_2 = X_1. \end{cases}$$

Theo định lý Sklar, khi đó sẽ tồn tại d A, uy nhất một Copula C sao cho hàm phân phối và hàm mật độ như sau:

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)).$$

$$f(x_1, x_2) = c(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2).$$

Với c *là hàm mật độ của Copula* C:

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}.$$

Ta có

$$\begin{cases} X_1 = Y_2. \\ X_2 = \dfrac{Y_1 - w_1 Y_2}{w_2}. \end{cases}$$

Jacobian của phép biến đổi này được xác định bởi

$$J = \begin{vmatrix} \dfrac{\partial X_1}{\partial Y_1} & \dfrac{\partial X_1}{\partial Y_2} \\ \dfrac{\partial X_2}{\partial Y_1} & \dfrac{\partial X_2}{\partial Y_2} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \dfrac{1}{w_2} & -\dfrac{w_1}{w_2} \end{vmatrix} = -\frac{1}{w_2} \neq 0.$$

Khi đó mật độ đồng thời của (Y_1, Y_2) được xác định như sau:

$$h(y_1, y_2) = f\left(y_2, \frac{y_1 - w_1 y_2}{w_2}\right)|J|$$
$$= \frac{1}{|w_2|} c\left(F_1(y_2), F_2\left(\frac{y_1 - w_1 y_2}{w_2}\right)\right) f_1(y_2) f_2\left(\frac{y_1 - w_1 y_2}{w_2}\right).$$

Do đó, ta có

Hàm mật độ xác suất của Y_1:

$$f_{Y_1}(y_1) = \int_{-\infty}^{+\infty} h(y_1, y_2)\, dy_2$$

$$= \frac{1}{|w_2|} \int_{-\infty}^{+\infty} c\left(F_1(y_2), F_2\left(\frac{y_1 - w_1 y_2}{w_2}\right)\right) f_1(y_2) f_2\left(\frac{y_1 - w_1 y_2}{w_2}\right) dy_2$$

$$= \frac{1}{|w_2|} \int_0^1 c\left(u, F_2\left(\frac{y_1 - w_1 F_1^{-1}(u)}{w_2}\right)\right) f_2\left(\frac{y_1 - w_1 F_1^{-1}(u)}{w_2}\right) du. \quad (4)$$

Hàm phân phối xác suất Y_1:

$$F_{Y_1}(t) = \int_{-\infty}^{t} f_{Y_1}(y_1)\, dy_1$$

$$= \int_{-\infty}^{t} \frac{1}{|w_2|} \int_0^1 c\left(u, F_2\left(\frac{y_1 - w_1 F_1^{-1}(u)}{w_2}\right)\right) f_2\left(\frac{y_1 - w_1 F_1^{-1}(u)}{w_2}\right) du\, dy_1$$

$$= \frac{1}{|w_2|} \int_0^1 \int_{-\infty}^{t} c\left(u, F_2\left(\frac{y_1 - w_1 F_1^{-1}(u)}{w_2}\right)\right) f_2\left(\frac{y_1 - w_1 F_1^{-1}(u)}{w_2}\right) dy_1\, du. \quad (5)$$

Đổi biến $v = F_2\left(\dfrac{y_1 - w_1 F_1^{-1}(u)}{w_2}\right)$, ta xét hai trường hợp sau:

Trường hợp $w_2 > 0$. Từ (5) ta được

$$F_{Y_1}(t) = \frac{w_2}{|w_2|} \int_0^1 \int_0^{F_2\left(\frac{t - w_1 F_1^{-1}(u)}{w_2}\right)} c(u, v)\, dv\, du$$

$$= \int_0^1 \int_0^{F_2\left(\frac{t - w_1 F_1^{-1}(u)}{w_2}\right)} \frac{\partial^2 C}{\partial u \partial v}(u, v)\, dv\, du$$

$$= \int_0^1 \frac{\partial}{\partial u} C\left(u, F_2\left(\frac{t - w_1 F_1^{-1}(u)}{w_2}\right)\right) du.$$

Trường hợp $w_2 < 0$. Từ (5) ta được

$$F_{Y_1}(t) = \frac{w_2}{|w_2|}\int_0^1\int_0^{F_2\left(\frac{t-w_1F_1^{-1}(u)}{w_2}\right)} c(u,v)\,dv\,du$$

$$= -\int_0^1\int_0^{F_2\left(\frac{t-w_1F_1^{-1}(u)}{w_2}\right)} \frac{\partial^2 C}{\partial u\partial v}(u,v)\,dv\,du$$

$$= -\int_0^1\left[\frac{\partial}{\partial u}C\left(u, F_2\left(\frac{t-w_1F_1^{-1}(u)}{w_2}\right)\right) - \frac{\partial}{\partial u}C(u,1)\right]du$$

$$= 1 - \int_0^1 \frac{\partial}{\partial u}C\left(u, F_2\left(\frac{t-w_1F_1^{-1}(u)}{w_2}\right)\right)du.$$

3 ỨNG DỤNG TRONG LỰA CHỌN DANH MỤC ĐẦU TƯ

3.1 Độ đo rủi ro trong danh mục đầu tư

Giả sử một phương án đầu tư P gồm hai danh mục kết hợp. Khi ta đầu tư vào cổ phiếu A với tỉ trọng w_1 và đầu tư vào cổ phiếu B với tỉ trọng w_2. Tỉ suất sinh lợi tại thời điểm hiện tại cho biết giá cổ phiếu đã lời hay lỗ bao nhiêu (phần trăm) so với ngày hôm qua.

Gọi lần lượt là giá của phương án đầu tư, giá cổ phiếu A và giá cổ phiếu B tại thời điểm t. P_t, P_{1t}, P_{2t} R_t, R_{1t}, R_{2t} lần lượt là tỉ suất đầu tư cả phương án, cổ phiếu A và cổ phiếu B tại thời điểm Khi đó

$$P_t = w_1P_{1t} + w_2P_{2t}. \quad (6)$$

$$R_t = w_1R_{1t} + w_2R_{2t}, \quad (7)$$

với t. $R_{1t} = \dfrac{P_{1t} - P_{1t-1}}{P_{1t-1}}, R_{2t} = \dfrac{P_{2t} - P_{2t-1}}{P_{2t-1}}$. (8)

Khi $R_{it} > 0,\ i = 1,2$ ta gọi tỉ suất đầu tư này là tỉ suất lời và ngược lại gọi là tỉ suất lỗ. Nếu $R_t > 0$ thì giá trị đầu tư sẽ tăng, còn nếu $R_t < 0$ thì giá đầu tư sẽ giảm. Với một xác suất $\alpha \in (0,1)$ cho trước, ta cần xác định độ đo rủi ro khi đầu tư.

Trong áp dụng, dựa vào số liệu quá khứ, chúng ta tìm w_1 và w_2 sao cho phương sai của R_t là nhỏ nhất với điều kiện $w_1 + w_2 = 1$, tức là giải quyết bài toán tối ưu sau:

$$\arg\min_{w1+w2=1} \left(w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2 Cov(R_{1t}, R_{2t})\right),$$

trong đó

σ_1^2, σ_2^2 là phương sai của R_{1t}, R_{2t}.

$Cov(R_{1t}, R_{2t})$ là hiệp phương sai của R_{1t}, R_{2t}.

Khi đó, lời giải bài toán trên là $w = \dfrac{aS^{-1}}{aS^{-1}a^T}$,

với $w = (w_1, w_2)$, $a = (1,1)$ và $S = Cov(R_{1t}, R_{2t})$ là ma trận hiệp phương sai của R_{1t}, R_{2t}.

Trong tài chính, người ta thường sử dụng VaR và ES là hai đại lượng để đo lường rủi ro. Chúng được xác định như sau:

$$VaR_\alpha(R_t) = \sup\{x \in \mathbb{R} : F_{R_t}(x) \le \alpha\} = \inf\{x \in \mathbb{R} : F_{R_t}(x) \ge \alpha\}, \quad (9)$$

$$ES_\alpha(R_t) = E(R_t | R_t \le VaR_\alpha(R_t)), \quad (10)$$

trong đó F_{R_t} là phân phối xác suất của phương án đầu tư với một xác suất $\alpha \in (0,1)$ cho trước và điều kiện thị trường bình thường.

Như vậy VaR của một danh mục đầu tư ở mức xác suất α tại thời điểm t được hiểu một cách đơn giản chính là phân vị xác suất đuôi dưới với mức α trong khi ES là kỳ vọng các R_t thỏa điều kiện dưới ngưỡng VaR

Ngoài hai độ đo rủi ro trên, chúng ta còn có một lớp độ đo rủi ro khá tổng quát gọi là độ đo rủi ro biến dạng. Giả sử Y là biến ngẫu nhiên tổn thất, độ đo rủi ro biến dạng của Y với mức ý nghĩa $\beta \in [0,1]$ được nghĩa như sau:

$$PH_\beta[Y] = \int_0^{+\infty} (1 - F_Y(y))^{1/\beta}\,dy. \quad (11)$$

3.2 Một ứng dụng cụ thể

Trong phần này nghiên cứu thực hiện trên số liệu thực về cổ phiếu của hai công ty lớn của Việt Nam, được thu thập trên website *finance.vietstock.vn* từ ngày 01/6/2018 đến 31/05/2019. Tuy nhiên, vì đây là vấn đề nhạy cảm của giai đoạn nghiên cứu nên bài viết chỉ kí hiệu hai công ty này là VIC và VNM mà không nêu cụ thể tên các công ty.

Phương án đầu tư lúc này là $R_t = w_1 R_{1t} + w_2 R_{2t}$,

trong đó

R_{1t}, R_{2t} lần lượt là tỉ suất đầu tư của cổ phiếu VIC và cổ phiếu VNM,

w_1, w_2 lần lượt là trọng số đầu tư vào cổ phiếu VIC và cổ phiếu VNM.

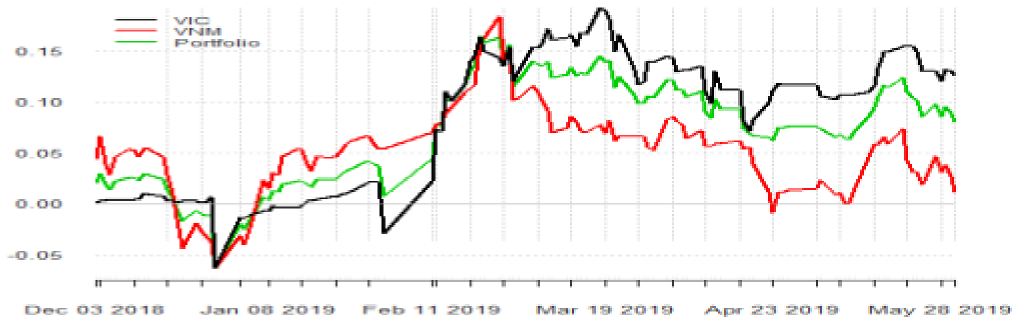
Trong áp dụng này, bài viết sử dụng số liệu từ 01/6/2018 đến ngày 03/12/2018 để tìm trọng số tối ưu w_1 và w_2. Ta có

$$S = \begin{bmatrix} 0.0004289374 & 0.0000764533 \\ 0.0000764533 & 0.0005493200 \end{bmatrix},$$

$$w = \ 0.57, 0.43 \ .$$

Nghĩa là trong tổng số vốn đầu tư thì ta sẽ đầu tư 57% vào VIC và 43% vào VNM. Như vậy phương án đầu tư cụ thể là $R_t = 0.57 R_{1t} + 0.43 R_{2t}$.

Tiếp theo ta tính hiệu suất lời tích luỹ của VIC, VNM và phương án đầu tư (Portfolio). Kết quả thực hiện được cho bởi Hình 1.



Hình 1: Hiệu suất sinh lời tích luỹ của VIC, VNM và Portfolio

Hình 1 cho thấy hiệu suất lời của VIC ở giai đoạn đầu thấp hơn VNM, tuy nhiên nó có xu hướng tăng mạnh hơn so với VNM ở giai đoạn sau. Hiệu suất sinh lời của Portfolio gần như là trung bình cộng của hai cỗ phiếu này.

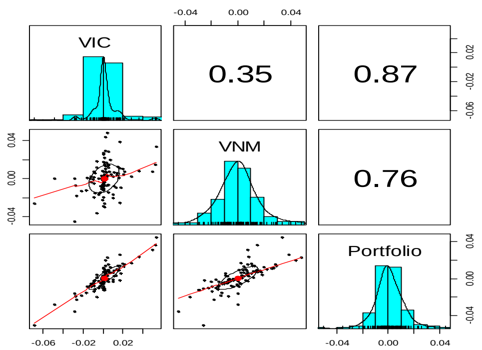
Tính thêm các tham số thống kê và ước lượng với độ tin cậy 95% cho các hiệu suất đầu tư, ta có Bảng 1.

Bảng 1: Các tham số cho tỉ suất đầu tư của VIC, VNM và Portfolio

Tham số	VIC	VNM	Portfolio
Cở mẫu	119.0000	119.0000	119.0000
Giá trị nhỏ nhất	-0.0693	-0.0447	-0.0507
Trung bình	0.0011	0.0002	0.0007
Giá trị lớn nhất	0.0535	0.0480	0.0448
Cận dưới của ước lượng	-0.0016	-0.0025	-0.0015
Cận trên của ước lượng	0.0039	0.0029	0.0030
Phương sai	0.0002	0.0002	0.0002
Độ lệch chuẩn	0.0151	0.0151	0.0125

Bảng 1 cho thấy hiệu suất đầu tư của VIC dao động từ -0.16% đến 0.39% và độ lệch chuẩn là 1.51%, trong khi của VNM dao động từ -0.25% đến 0.29% và độ lệch chuẩn là 1.51%. Hiệu suất đầu tư của Porfolio lần lượt −0.015% đến 0.3% và độ lệch chuẩn bằng 1.25%. Các tham số của Bảng 1 một lần nữa cho thấy khi đầu tư vào VIC tốt hơn VNM.

Tính hệ số tương quan giữa hiệu xuất đầu tư của các đại lượng ta có Hình 2.



Hình 2: Tương quan giữa hiệu suất sinh lời của VIC, VNM và Portfolio

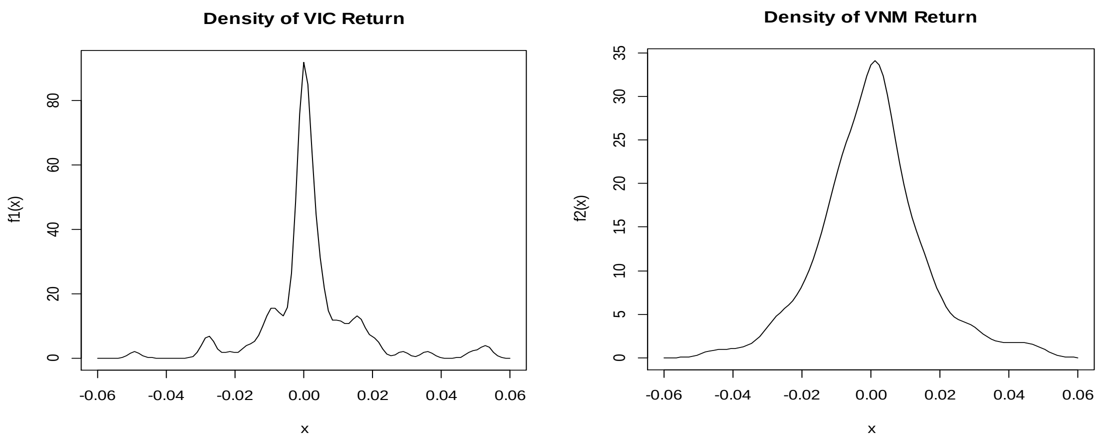
Biểu đồ tương quan giữa hiệu suất đầu tư của VIC, VNM và Portfolio cho biết VIC và VNM có tương quan với nhau nhưng khá yếu (hệ số tương quan là 0.35). Bên cạnh đó ta thấy rằng tỉ trọng đầu tư vào cổ phiếu VIC và VNM đều có tác động khá lớn đến đầu ra của Portfolio với hệ số tương quan lần lượt là 0.87 và 0.76.

Chúng ta sẽ thực hiện thêm kiểm định Shapiro-Wilk để kiểm tra phân phối chuẩn hiệu suất sinh lời của VIC và VNM. Kết quả kiểm định được cho bởi Bảng 2.

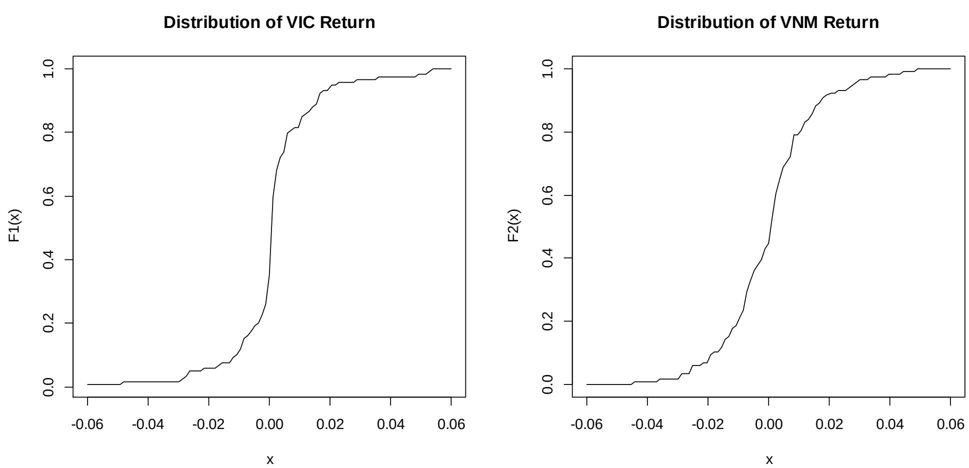
Bảng 2: P- giá trị của kiểm định Shapiro-Wilk

VIC	VNM	Portfolio
1.716×10^{-10}	0.03811	4.587×10^{-5}

Kết quả này cho thấy các đại lượng xem xét không có phân phối chuẩn. Do đó để ước lượng hàm mật độ xác suất của chúng, nghiên cứu này sử dụng phương pháp hàm hạt nhân (Scott,1979; Xu *et al.*, 2015). Kết quả được cho bởi Hình 3 và Hình 4.

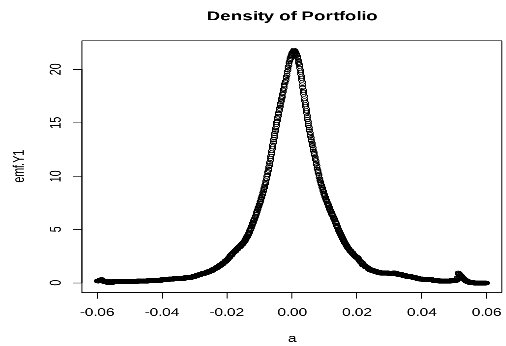


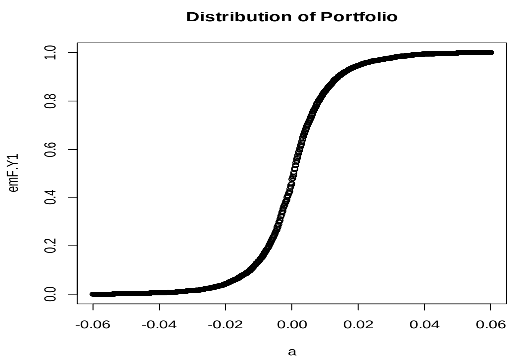
Hình 3: Hàm mật độ hiệu suất đầu tư của VIC và VNM



Hình 4: Hàm phân phối hiệu suất đầu tư của VIC và VNM

Lựa chọn cấu trúc phụ thuộc giữa VIC và VNM theo tiêu chuẩn AIC (Aho *et al.,* 2015), ta nhận được Copula Gauss là phù hợp nhất. Ước lượng hàm mật độ xác suất cho phương án đầu tư với cấu trúc Copula Gauss, ta có Hình 5.





Hình 6: Hàm mật độ và hàm phân phối của Portfolio

Tính các độ đo rủi ro cho Portfolio với mức ý nghĩa 1% và 5%, ta có Bảng 3.

Bảng 3: Các độ đo rủi ro

Độ rủi ro	VIC	VNM	Portfolio
VaR	-2.28%	-2.29%	-1.94%
ES	-4.13%	-2.97%	-3.10%
PH	-3.23%	-2.79%	-2.38%

Từ Bảng 3 ta có những nhận xét sau:

* Theo độ đo *VaR* thì VIC, VNM và Portfolio có nguy cơ giảm lần lượt là 2.28%, 2.29% và 1.94% giá trị.

* Theo độ đo *ES* thì VIC, VNM và Portfolio có nguy cơ giảm lần lượt là 5.88% ,4.73% và 3.94% giá trị.

* Theo độ đo *PH* thì VIC, VNM và Portfolio có nguy cơ giảm lần lượt là 3.23%, 2.79% và 2.38% giá trị.

Mặc dù VIC có hiệu suất sinh lời cao hơn, nhưng từ kết quả Bảng 3 ta thấy VIC lại có độ rủi ro cao hơn VNM và khi kết hợp lại hai cổ phiếu này thì rủi ro sẽ giảm đi.

4 KẾT LUẬN

Bài viết đã thiết lập được hàm mật độ và hàm phân phối xác suất cho tổng có trọng số của hai biến ngẫu nhiên phụ thuộc. Hàm mật độ và hàm phân phối này được sử dụng để tính các độ rủi ro khi đầu tư vào hai cổ phiếu. Nghiên cứu cũng đề xuất những ước lượng quan trọng cho các tham số khi áp dụng số liệu thực vào các mô hình lý thuyết. Quy trình thực hiện trong áp dụng của nghiên cứu này được minh họa qua số liệu thực tế. Áp dụng này cũng cho thấy tiềm năng của nghiên cứu cho những vấn đề khác nhau trong tài chính. Trong tương lai chúng tôi sẽ mở rộng nghiên cứu cho tổng có trọng số của nhiều hơn hai biến phụ thuộc và áp dụng nó vào những vấn đề phức tạp hơn của tài chính.

TÀI LIỆU THAM KHẢO

Aho, K., Derryberry, D. and Peterson, T., 2014. Model selection for ecologists: the world views of AIC and BIC. Ecology, 95(3): 631–636.

Cherubini, U., Elisa, L. and Walter, V., 2004. Copula methods in finance. John Wiley & Sons. New York. 512 pages.

Cherubini, U., Mulinacci, S. and Romagnoli, S., 2011. A Copula-based model of speculative pricedynamics in discrete time. Journal of Multivariate Analysis, 102(6): 1047-1063.

Dettmann, C. P. and Orestis, G., 2009. Product of *n* independent uniform random variables. Statistics & Probability Letters, 79: 2501–3.

Sereda, E. N., Bronshtein, E. M., Rachev, S. T., Fabozzi, F. J., Sun, W. and Stoyanov, S. V., 2010. Distortion risk measures in portfolio optimization. In Handbook of portfolio construction (pp. 649-673). Springer, Boston, MA.

Galambos, J. and Italo, S., 2004. Products of Random Variables: Applications to Problems of Physics and to Arithmetical Functions. Chappan & Hall. New York, 344 pages.

Garg, M., Sharma, A. and Manohar, P., 2016. The distribution of the product of two independent generalized trapezoidal random variables. Communications in Statistics-Theory and Methods, 45(21): 6369–6384.

Hien, T. D., Uyen, P. H., Sel, L. and Trung, V. D., 2015. A new measure of monotone dependence by using Sobolev norms for Copula. In

International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making, 4: 126–37.

Pham-Gia, T., Turkkan, N. and Eng, P., 1993. Bayesian analysis of the difference of two proportions. Communications in Statistics-Theory and Methods, 22(6): 1755–1771.

Sel, L., Hung, P. K., Sal, L. and Wing, K. W., 2019. Distribution of quotient of dependent and independent random variables using Copulas. Journal of Risk and Financial Management, 12: 1-27.

Sklar, A.,1959. N-dimensional distribution functions and their margins. Publ. Inst. Statist,8: 229–231.

Scott, D., 1979. On optimal and data-based histograms. Biometrika, 66(3): 605–610.

Tai, V.V., Ha, C.N. and Thao, N.T., 2016. The prior probability in classifying two populations by Bayesian method. Applied mathematics in Engineering and Reliability. CRC Press, 1: 35-40.

Tang, J., Songsak, S., Vicente, R. and Wing, K.W., 2014. Modelling dependence between tourism demand and exchange rate using Copula-based GARCH model. Current Issues in Method and Practice, 19: 1–19.

Yang, Y. and Yuebao, W., 2013. Tail behavior of the product of two dependent random variables with applications to risk theory. Extremes, 16: 55–74.

Xu, X., Yan, Z. and Xu, S., 2015. Estimating wind speed probability distribution by diffusion-based kernel density method. Electric Power Systems Research, 121: 28–37.