

SỬ DỤNG MÔ HÌNH COX ĐỂ TÍNH XÁC SUẤT NỢ QUÁ HẠN CỦA CÁC KHÁCH HÀNG DOANH NGHIỆP VỪA VÀ NHỎ

Đoàn Trọng Tuyền

Khoa Toán kinh tế, Trường Đại học Kinh tế quốc dân

Email: doantrongtuyen@gmail.com

Nguyễn Thị Minh

Khoa Toán kinh tế, Trường Đại học Kinh tế quốc dân

Email: minhkthn@gmail.com

Nguyễn Thị Lương

Phòng Phân tích công cụ mô hình rủi ro, Ngân hàng TMCP Hàng hải Việt Nam

Email: nguyenluong.neu@gmail.com

Ngày nhận: 01/8/2016

Ngày nhận bản sửa: 5/9/2016

Ngày duyệt đăng: 15/9/2016

Tóm tắt:

Một yêu cầu quan trọng trong nghiệp vụ của ngân hàng là bắt buộc phải dành một khoản dự phòng cho các rủi ro có thể xảy ra. Theo các hiệp ước Basel thì ngân hàng bắt buộc phải giữ lại lượng vốn bằng ít nhất 8% tổng tài sản, được tính theo nhiều phương pháp khác nhau và phụ thuộc vào độ rủi ro của chúng, do vậy việc đánh giá rủi ro của khách hàng rất quan trọng. Bài viết này sẽ sử dụng phương pháp phân tích sống sót để đánh giá xác suất một khách hàng bị quá hạn theo thời gian, từ đó có thể ước tính mức dự phòng cần thiết. Khác với các phương pháp ước lượng xác suất vỡ nợ thông thường chỉ quan tâm tới việc sự kiện xảy ra hay không xảy ra trong một đơn vị thời gian (thường là một năm), trong nghiên cứu này, chúng tôi còn quan tâm đến thời điểm nó xảy ra, do đó các kết quả tính toán dự phòng rủi ro sẽ có giá trị thực tiễn hơn, giúp cho công tác dự phòng đạt hiệu quả tốt hơn.

Từ khóa: Hiệp ước Basel; nợ quá hạn; phân tích sống sót; rủi ro tín dụng.

Applying Cox model to estimate probability of overdue debt of small and medium enterprises

Abstract:

Banking business is required to save a provision for risks that may occur. According to the Basel Capital Accord, banks must hold capital equal to at least 8% of the basket of assets production calculated by different methods depending on their risk level. Thus, risk assessment of customers is very important. This paper uses the survival analysis methods to assess the probability that a customer is past due over time to estimate the level of redundancy required in a reasonable manner. Different from the method of calculating the probability of default in which the only common interest in events happening or not in a time (usually a year), in this study, the time the events happened are considered. By this way, the results calculated for risk provisions would have more practical value.

Keywords: Basel Capital Accord; credit risk; overdue debt; survival analysis.

1. Giới thiệu

Đánh giá rủi ro tín dụng của khách hàng là một khâu quan trọng trong hoạt động của các ngân hàng thương mại cũng như các tổ chức tín dụng khác. Việc đánh giá được mức độ rủi ro của một khách hàng sẽ cho phép các ngân hàng đưa ra các tính toán hợp lý trong việc quyết định cho vay hay không cho vay, và nếu cho vay thì với mức lãi suất nào là phù hợp để nhằm tối đa hóa lợi nhuận của ngân hàng và đồng thời đảm bảo an toàn vốn của chủ sở hữu. Ngoài ra, dựa vào tính chất đặc biệt của hoạt động tín dụng và vai trò quan trọng của sự ổn định trong hệ thống tín dụng đối với sự phát triển của nền kinh tế, từ đó ngân hàng nhà nước đưa ra các quy định riêng dành cho các hoạt động tín dụng. Một trong những quy định quan trọng nhất là quy định về sự bảo đảm an toàn vốn, trong đó yêu cầu về việc tính toán và trích lập dự phòng đối với các khoản vay có rủi ro. Như vậy có thể nói việc ước lượng – dự báo mức độ rủi ro của các khoản vay với độ tin cậy cao không chỉ là yêu cầu nội tại của mỗi ngân hàng mà còn là đòi hỏi từ phía quản lý nhà nước nhằm đảm bảo hoạt động ổn định của hệ thống ngân hàng và an toàn vốn cho người chủ vốn. Tại Việt Nam, Ngân hàng Nhà nước đã đưa ra lộ trình tiến tới thực hiện trích lập dự phòng và bảo đảm an toàn vốn theo chuẩn Basel II, trong đó đặt mục tiêu đến năm 2018, sẽ có 10 ngân hàng, bao gồm: BIDV, VietinBank, Vietcombank, Techcombank, ACB, VPBank, MB, Maritime Bank, Sacombank và VIB hoàn thành việc thí điểm thực hiện phương pháp quản trị vốn và rủi ro theo tiêu chuẩn Basel II. Trong quá trình thực hiện lộ trình này, việc xây dựng được một công cụ nhằm đánh giá – dự báo tốt nhất có thể về rủi ro tín dụng là một yếu tố quan trọng, nhất là trong bối cảnh Việt Nam hội nhập ngày càng sâu rộng với kinh tế quốc tế, làm cho các hoạt động tín dụng ngày càng có quy mô lớn và độ phức tạp cao.

Có nhiều phương pháp đã được sử dụng để ước lượng độ rủi ro vỡ nợ, chẳng hạn sử dụng các mô hình logit - probit hoặc phương pháp tách biệt. Các phương pháp này cho phép ước lượng xác suất để một khoản vay bị quá hạn trong một khoảng thời gian nhất định (thường là một năm). Tuy nhiên trong bối cảnh vòng quay của các hoạt động tín dụng ngày càng gia tăng, cũng như sức ép về việc sử dụng một cách có hiệu quả và an toàn đồng vốn thì việc tính toán xác suất trong một khoảng thời gian cố định

cho thấy những hạn chế nhất định. Chẳng hạn, đối với khách hàng cá nhân, thì mỗi ngân hàng có đến hàng triệu tài khoản cho vay, khi đó việc tính toán khả năng các khoản vay bị quá hạn tại mỗi thời điểm lại có thể giúp cho ngân hàng đưa ra được một cách chính xác hơn lượng trích lập cần thiết cũng như có các biện pháp bảo đảm an toàn vốn phù hợp và hiệu quả. Khi đó có thể sử dụng phương pháp phân tích sống sót (survival analysis) cho phép tính toán khả năng bị quá hạn của các khoản vay tại từng thời điểm, phương pháp này trở nên hữu hiệu trong bài toán tính xác suất nợ quá hạn. Phương pháp phân tích sống sót đã được sử dụng trong nhiều nghiên cứu trên thế giới, tuy nhiên các nghiên cứu ở Việt Nam vẫn còn rất hạn chế và hầu như chưa được thực hiện để ước lượng rủi ro tín dụng trong hoạt động của các ngân hàng. Vì vậy chúng tôi hy vọng bài viết sẽ có những đóng góp nhất định trong các nghiên cứu thực nghiệm, đặc biệt là nghiên cứu thực nghiệm tại Việt Nam về đo lường rủi ro tín dụng.

2. Tổng quan nghiên cứu

Trên thế giới đã có nhiều nghiên cứu nhằm xây dựng mô hình và ước lượng đánh giá rủi ro tín dụng. Có thể nói, mô hình tính toán rủi ro tín dụng được giới thiệu lần đầu bởi Altman (1968) với mô hình Z-score. Trong nghiên cứu này, tác giả đã sử dụng phương pháp phân tích tách biệt (discriminant analysis) để tính giá trị của hàm phân biệt Z của từng công ty dựa trên các đặc điểm tài chính của các công ty này và sử dụng chúng để dự báo xác suất vỡ nợ của các công ty trong vòng hai năm tới. Phương pháp phân tích tách biệt hay Z-score hiện nay vẫn đang được sử dụng khá phổ biến trong việc tính toán và dự báo xác suất vỡ nợ cũng như xếp hạng tín dụng.

Một phương pháp khác cũng thường được sử dụng trong đánh giá xác suất vỡ nợ của khách hàng là dựa trên các mô hình hồi quy dạng logit - probit. Chẳng hạn Martina (2012) đã sử dụng mô hình logit dạng đa thức (và một số phương pháp khác) để xếp hạng các công ty tài chính châu Âu. Kết quả nghiên cứu cho thấy kết quả dự báo từ mô hình logit dạng đa thức khá tốt với xác suất đúng đến trên 70%.

Gần đây, với sự phong phú của số liệu và sự cải thiện trong tốc độ tính toán của máy tính, các nhà nghiên cứu đã sử dụng một số phương pháp khác, như phương pháp mạng nơ-ron nhân tạo (artificial

neural networks - ANN), hay máy vector hỗ trợ (support vector machines - SVM) cũng như các mô hình kết hợp giữa chúng. Các nghiên cứu điển hình theo hướng này là của Altman & Saunders (1998), Saunders (1999), Crouhy & cộng sự (2000), Hand (2001), Hamerle & cộng sự (2003), Chen & cộng sự (2006), Hanson & Schuermann (2004), Wang & cộng sự (2005).

Các nghiên cứu trên đều sử dụng các phương pháp truyền thống để ước lượng rủi ro tín dụng trong đó quan tâm đến xác suất vỡ nợ của các công ty trong một khoảng thời gian, thường là một năm. Tuy nhiên trong tính toán rủi ro, sẽ rất hữu ích nếu có thể tính toán được xác suất vỡ nợ của các công ty tại các thời điểm, do đó gần đây các nhà nghiên cứu bắt đầu quan tâm đến việc tính toán các xác suất này, và một trong những phương pháp mới được sử dụng phương pháp phân tích sống sót, chẳng hạn các nghiên cứu của Narain (1992), Stepanova & Thomas (2002), Baba & Goko (2006), Malix & Thomas (2006), Beran & Djaidja (2007). Điểm đặc trưng của các nghiên cứu này là tác giả sử dụng các mô hình hồi quy tham số và bán tham số để mô hình hóa thời gian tới khi phá sản (thời gian sống sót), bao gồm các mô hình mũ, mô hình Weibull và mô hình tỷ lệ nguy cơ hằng số Cox.

Tại Việt Nam, việc tính toán và đánh giá rủi ro nói chung và rủi ro tín dụng nói riêng đã được các ngân hàng quan tâm, đặc biệt là trong mấy năm gần đây, khi mà Ngân hàng Nhà nước đưa ra quy định về việc thực hiện Hiệp ước Basel II. Do đó đã có một số nghiên cứu ban đầu về đánh giá rủi ro tín dụng tại các ngân hàng thương mại, ví dụ luận án thạc sĩ của Nguyễn Văn Đức (2012), luận án thạc sĩ của Trương Thị Thảo My (2014) và luận án tiến sĩ của Nguyễn Trọng Hòa (2010), tuy nhiên các nghiên cứu này chỉ tập trung tính toán xác suất vỡ nợ của khách hàng trong một khoảng thời gian, cụ thể là trong một năm và các phương pháp truyền thống được các tác giả sử dụng là các mô hình Z-score của Altman, mô hình logit, probit hoặc phân tích tách biệt. Với quy mô và độ phức tạp ngày càng tăng của các hoạt động tín dụng tại Việt Nam, trong đó số khách hàng có thể lên đến con số hàng triệu, và vòng quay của dòng tiền ngày càng nhanh trên thị trường tín dụng, việc có thể đưa ra các ước lượng về độ rủi ro theo thời điểm là rất cần thiết. Tuy nhiên các nghiên cứu về vấn đề này tại Việt Nam còn rất hạn chế. Vì vậy

trong nghiên cứu này chúng tôi sẽ sử dụng phương pháp phân tích sống sót để ước lượng xác suất vỡ nợ quá hạn cho các khách hàng của ngân hàng thương mại.

3. Phương pháp phân tích sống sót

Phân tích sống sót quan tâm đến thời gian sống sót của đối tượng được nghiên cứu, ví dụ như lượng thời gian sống (đến khi chết) của bệnh nhân điều trị một loại bệnh, thời gian đến lúc tái phát một căn bệnh mãn tính, thời gian đến khi khách hàng vay vốn bị nợ quá hạn...

Tính từ thời điểm bắt đầu nghiên cứu, một đối tượng có thể sống sót đến khi kết thúc nghiên cứu, hoặc xảy ra sự kiện (ví dụ như tái phát bệnh, tử vong hay nợ quá hạn...) trong khoảng thời gian nghiên cứu hoặc ta có thể mất thông tin về đối tượng (bệnh nhân không tiếp tục điều trị,...), ta gọi những trường hợp như thế là những sự kiện chưa xảy ra (censored).

Gọi thời gian sống sót của đối tượng là T , ở đây $T \geq 0$ là biến ngẫu nhiên có thể rời rạc hoặc liên tục, t là một giá trị cụ thể của T . Ta có một số định nghĩa sau:

Hàm sống sót (survivor function): $S(t) = P(T \geq t)$;

Hàm nguy cơ (hazard function):

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

Từ định nghĩa trên ta suy ra mối liên hệ giữa hàm sống sót và hàm nguy cơ:

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} [\log S(t)]; \quad S(t) = e^{-\int_0^t h(u) du}$$

Trong đó $f(t)$ là hàm mật độ:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t)}{\Delta t}$$

Có nhiều mô hình phân tích sống sót đã được xây dựng như mô hình phi tham số (ước lượng Kaplan-Meier), mô hình bán tham số (mô hình Cox) và các mô hình tham số.

Phương pháp ước lượng Kaplan-Meier đã đưa ra công thức ước lượng hàm sống sót $S(t) = P(T \geq t)$ như sau:

$$\hat{S}(t) = \prod_{k: \tau_k < t} \frac{r_k - d_k}{r_k} = \prod_{k: \tau_k < t} \left(1 - \frac{d_k}{r_k} \right)$$

Trong đó:

$\tau_1, \tau_2, \dots, \tau_p$ là tập hợp p thời gian phân biệt xảy ra sự kiện trong mẫu;

Bảng 1: Hàm sống sót $S(t)$ và hàm nguy cơ $h(t)$ trong mô hình có tham số

Phân phối	$S(t)$	$h(t)$
Mũ	$e^{-\lambda t}$	λ
Weibull	$e^{-\lambda t^p}$	$\lambda p t^{p-1}$
Log-logistic	$\frac{1}{1 + \lambda t^p}$	$\frac{\lambda p t^{p-1}}{1 + \lambda t^p}$

d_k là số lượng đối tượng xảy ra sự kiện ở τ_k ;

r_k là số đối tượng có nguy cơ ở bên phải của thời gian xảy ra sự kiện thứ k (bao gồm các đối tượng xảy ra sự kiện hoặc chưa xảy ra sự kiện ở thời gian thứ k trở đi).

Phương pháp này có ưu điểm là công thức tính toán khá đơn giản. Tuy nhiên một nhược điểm của nó là chỉ dựa vào các con số thống kê về sự xảy ra hay không xảy ra của sự kiện, mà không tính đến các yếu tố có tác động đến sự kiện này. Trong thực tế, các yếu tố này thường xuyên thay đổi, do đó tác động của nó tới xác suất xảy ra sự kiện cũng sẽ thay đổi vì vậy để có thể đưa ra các tính toán cũng như dự báo chính xác hơn thì mô hình nghiên cứu cần tính đến vai trò của các yếu tố này. David R. Cox (1972) đã đề xuất mô hình phân tích sống sót trong đó có xem xét đến các yếu tố tác động trên.

Mô hình Cox mô tả hàm nguy cơ tại thời điểm t với tập hợp các biến giải thích X dưới dạng sau:

$$h(t, X) = h_0(t) \cdot e^{\sum_{k=1}^p \beta_k \cdot x_k}$$

với $X = (x_1, x_2, \dots, x_p)$ là vectơ các biến giải thích, $h_0(t)$ là hàm nguy cơ cơ sở (baseline hazard function), ở đây $h_0(t)$ chỉ phụ thuộc vào t chứ không phụ thuộc vào các biến giải thích. Các hệ số β_k , $k = 1, 2, \dots, p$ là hệ số thể hiện mối liên hệ giữa x_k và hàm sống sót.

Mô hình sống sót có tham số (parametric survival model) được sử dụng với giả thiết thời gian sống sót tuân theo một phân phối đã biết nào đó. Các phân phối thường được sử dụng trong mô hình sống sót có tham số là phân phối Weibull, phân phối mũ và phân phối Log-logistic. Bảng 1 là một số giá trị hàm sống sót $S(t)$, hàm nguy cơ $h(t)$ tương ứng với các phân phối của thời gian sống sót hay được sử dụng.

Số tham số trong một mô hình có thể là 1 (Mô hình với phân phối mũ) hay 2 (Weibull hay Log-logistic). Vấn đề còn lại của bài toán ước lượng mô hình là xác định giá trị các tham số một cách phù hợp từ dữ liệu đang nghiên cứu.

4. Sử dụng mô hình Cox ước lượng xác suất vỡ nợ của khách hàng

4.1. Số liệu

- Khái niệm

Trong nghiên cứu này, nếu một khách hàng đến ngày quy định không trả được nợ gốc hoặc trả lãi thì khách hàng này được xếp vào nợ quá hạn. Thời gian kể từ khi bắt đầu mở hạn mức cho khách hàng tại ngân hàng đến thời điểm lần đầu tiên xảy ra hiện tượng quá hạn được định nghĩa là thời gian sống sót của đối tượng (survival time) và đồng thời đối tượng này cũng được xếp vào trường hợp xảy ra sự kiện. Thời gian sống sót của đối tượng được đo theo tháng.

Số liệu nghiên cứu được trích dẫn nguồn từ Ngân hàng TMCP Hàng hải Việt Nam từ 31/01/2012 đến 30/06/2015, gồm 249 quan sát với đối tượng nghiên cứu là các khách hàng doanh nghiệp nhỏ và vừa. Trong 249 khách hàng thì có 151 khách hàng để xảy ra nợ quá hạn.

- Các biến số

Status: biến phụ thuộc, nhận giá trị 1 khi có xảy ra hiện tượng nợ quá hạn và nhận giá trị 0 nếu ngược lại và khi đó khách hàng này chỉ nằm trong nhóm các khách hàng có nguy cơ xảy ra nợ quá hạn. Các biến giải thích trong mô hình bao gồm:

donbay (đo bằng Tổng nợ/Tổng tài sản): nếu nợ quá lớn so với tổng tài sản thì khả năng trả được nợ đúng hạn sẽ giảm sút.

loinhuan (đo bằng Lợi nhuận sau thuế/Vốn chủ sở hữu): lợi nhuận sau thuế càng lớn thì khả năng trả được nợ đúng hạn càng gia tăng.

Bảng 2: Kết quả mô hình Cox

t	Coef.	Std. Err.	z	$P > z $	[95% Conf. Interval]	
$x_1 = donbay$	0,7906522	0,3689548	2,14	0,032	0,0675141	1,51379
$x_2 = loinhuan$	-1,181008	0,6236522	-1,89	0,058	-2,403344	0,0413281
$x_3 = nganhnghe$	0,3201661	0,1649506	1,94	0,052	-0,003131	0,6434633
$x_4 = kinhnghiem$	-0,8220808	0,1766531	-4,65	0,000	-1,168315	-0,4758471

Nguồn: Tác giả tính toán từ dữ liệu nghiên cứu

nganhnghe (ngành nghề hoạt động của doanh nghiệp): nhận giá trị 0 với các doanh nghiệp hoạt động trong lĩnh vực dịch vụ, nhận giá trị 1 nếu doanh nghiệp hoạt động trong lĩnh vực sản xuất, chế tác hoặc xây dựng. Ngành nghề kinh doanh cũng có thể có tác động khác nhau đến khả năng trả nợ, chẳng hạn những ngành có độ rủi ro cao thì khả năng không trả nợ đúng hạn cũng gia tăng.

kinhngghiem (kinh nghiệm hoạt động của doanh nghiệp): thể hiện thời gian hoạt động của doanh nghiệp, các doanh nghiệp mới hoạt động thì tính ổn định về tài chính, về thị trường có thể bị hạn chế hơn các doanh nghiệp đã tham gia thị trường lâu hơn. Ở đây chúng tôi tạm chia thành 2 nhóm: biến này nhận giá trị 0 nếu doanh nghiệp có thời gian hoạt động chưa đến 3 năm và 1 nếu doanh nghiệp đã hoạt động từ 3 năm trở lên.

4.2. Mô hình và kết quả ước lượng

Mô hình Cox được xác định như sau:

$$\hat{h}(t, \mathbf{X}) = \hat{h}_0(t) \cdot e^{\hat{\beta}_1 * x_1 + \hat{\beta}_2 * x_2 + \hat{\beta}_3 * x_3 + \hat{\beta}_4 * x_4}$$

Sau khi thực hiện các kiểm định (Phụ lục), kết quả kiểm định cho thấy mô hình Cox với bộ biến số: $X = (x_1 = donbay, x_2 = loinhuan, x_3 = nganhnghe, x_4 = kinhnghiem)$ là thỏa mãn các giả thiết cần thiết, và hệ số ước lượng được cho trong Bảng 2.

Khi đó hàm nguy cơ được viết dưới dạng

$$\hat{h}(t, \mathbf{X}) = \hat{h}_0(t) \cdot e^{0,7906522 * x_1 - 1,181008 * x_2 + 0,3201661 * x_3 - 0,8220808 * x_4}$$

Từ mối liên hệ giữa hàm nguy cơ $h(t, \mathbf{X})$ với hàm sống sót $S(t, \mathbf{X})$ ta được

$$\hat{S}(t, \mathbf{X}) = \left[\hat{S}_0(t) \right]^{e^{0,7906522 * x_1 - 1,181008 * x_2 + 0,3201661 * x_3 - 0,8220808 * x_4}}$$

Kết quả ước lượng cho thấy các biến số đều có ý nghĩa thống kê và có dấu phù hợp với kỳ vọng ban đầu, cụ thể như sau:

Hệ số $\hat{\beta}_1 = 0,7906522 > 0$, ngụ ý rằng *khi đòn bẩy càng bé thì* $\sum_{i=1}^4 \hat{\beta}_i * x_i$ càng bé, và do đó giá trị của hàm sống sót này càng lớn (do $0 \leq S(t) \leq 1$), điều này

là phù hợp vì khi nợ càng ít so với tổng tài sản của doanh nghiệp thì doanh nghiệp càng có khả năng thanh toán nợ tốt hơn.

Lập luận tương tự, hệ số của biến $x_2 = loinhuan$ là $\hat{\beta}_2 = -1,181008$, khi $x_2 = loinhuan$ càng lớn thì $\sum_{i=1}^4 \hat{\beta}_i * x_i$ càng bé, vì vậy $\hat{S}(t, \mathbf{X})$ càng lớn, khả năng thanh toán nợ của doanh nghiệp tăng theo $x_2 = loinhuan$, nghĩa là khi lợi nhuận của doanh nghiệp tính trên vốn chủ sở hữu càng cao thì doanh nghiệp càng có khả năng thanh toán nợ tốt hơn.

Theo số liệu tính toán thì $\hat{\beta}_3 = 0,3201661 > 0$, có nghĩa là các doanh nghiệp hoạt động trong lĩnh vực dịch vụ có khả năng thanh toán nợ với ngân hàng tốt hơn so với các doanh nghiệp hoạt động trong lĩnh vực phi sản xuất. Tương tự như trên, biến $x_4 = kinhnghiem$ có hệ số $\hat{\beta}_4 = -0,8220808 < 0$ chứng tỏ các doanh nghiệp có kinh nghiệm hoạt động trên 3 năm (*kinhngghiem* = 1) sẽ thanh toán nợ tốt hơn các doanh nghiệp mới hoạt động dưới 3 năm (*kinhngghiem* = 0).

- *Tính toán hàm sống sót – minh họa với giá trị cụ thể:*

Giả sử muốn tính toán hàm sống sót cho hai doanh nghiệp với giá trị biến số như sau: $X_1 = (0, 7115; 0,0934; 1; 0)$, $X_2 = (0,7115; 0,0934; 1; 1)$, nghĩa là hai doanh nghiệp trong ngành sản xuất – chế tác, có cùng hệ số đòn bẩy 71,15% và tỷ lệ lợi nhuận sau thuế là 9,34%, một doanh nghiệp mới hoạt động và một doanh nghiệp đã có nhiều năm hoạt động. Khi đó ta có:

$$\hat{S}(t, \mathbf{X}_1) = \left[\hat{S}_0(t) \right]^{2,1650}, \hat{S}(t, \mathbf{X}_2) = \left[\hat{S}_0(t) \right]^{0,95154}$$

Do $0 \leq \hat{S}_0(t) \leq 1$ nên đường sống sót $\hat{S}(t, \mathbf{X}_1)$ luôn nằm phía dưới $\hat{S}(t, \mathbf{X}_2)$, nghĩa là doanh nghiệp có kinh nghiệm hoạt động lâu năm hơn thì xác suất nợ quá hạn tại mỗi thời điểm sẽ luôn nhỏ hơn.

Để tính toán được hàm sống sót $\hat{S}(t, \mathbf{X})$ tại mỗi giá trị cụ thể của X thì còn cần phải tìm giá trị của hàm sống sót cơ sở $\hat{S}_0(t)$. Với mô hình trên, chẳng

Bảng 3: Trích xuất giá trị của hàm sống sót cơ sở $\widehat{S}_0(t)$

Time	Survivor Function
5	0,7548 , khi đó $\widehat{S}_0(5) = 0,7548$
10	0,5242 , khi đó $\widehat{S}_0(10) = 0,5242$
25	0,3361 , khi đó $\widehat{S}_0(25) = 0,3361$
30	0,2940 , khi đó $\widehat{S}_0(35) = 0,2940$

Nguồn: Tác giả tính toán từ phần mềm STATA với dữ liệu nghiên cứu

hạn khi tính tại các mốc thời gian $t_1=5; t_2=10; t_3=25; t_4=35$ thì thu được kết quả là:

Các đường sống sót sơ sở và đường sống sót tương ứng với các vector $X_1 = (0,7115;0,0934;1;0)$, $X_2 = (0,7115;0,0934;1;1)$ được tính toán và biểu diễn trong Hình 1.

- So sánh với mô hình logit

Cũng với bộ số liệu đã thực hiện trên mô hình Cox trên đây, ta tính xác suất để một đối tượng xảy ra hiện tượng nợ quá hạn bằng mô hình logit. Kết quả được tóm tắt trong Bảng 4.

Khi đó, mô hình logit là:

$$\text{logit } P(\text{status} = 1|X) = -0,1897811 + 1,158026*x_1 - 3,220027*x_2 + 0,6067207*x_3 - 0,5197397*x_4,$$

Để tiện so sánh hai mô hình này, ta xét một trường hợp cụ thể của vector biến giải thích X là $X_1 = (0,7115;0,0934;1;0)$, khi đó ta có $\text{logit } P(\text{status}$

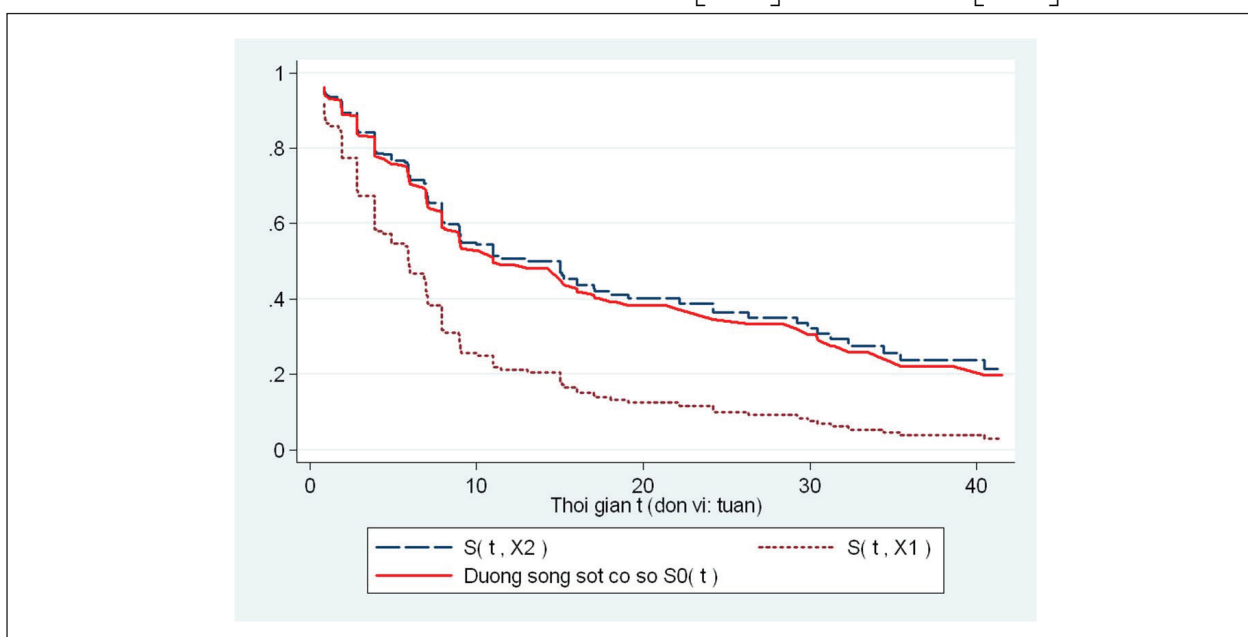
$= 1|X_1) = 0,94012$, từ đó suy ra $P(\text{status} = 1|X_1) = 0,71912$.

Rõ ràng ta thấy rằng, kết quả của mô hình logit trên đây cho thấy khi $X_1 = (0,7115;0,0934;1;0)$, xác suất để khách hàng nợ quá hạn là 0,71912; trong khi đó cũng với vector X_1 này thì mô hình Cox cho kết quả khác, $\hat{S}(t, X_1) = [\widehat{S}_0(t)]^{2,1650}$, ta thấy xác suất để xảy ra nợ quá hạn còn phụ thuộc vào thời gian nữa, ở đây khoảng tháng thứ 4 thì mới xảy ra nợ quá hạn với xác suất 0,71912, còn đến tháng thứ 30 thì xác suất này chỉ còn lại gần 0,070627.

5. Thảo luận và khuyến nghị

Tại hội thảo chuyên đề về quản trị rủi ro ngân hàng trong khuôn khổ Banking Vietnam 2013, rủi ro tín dụng tiếp tục được tái khẳng định là rủi ro lớn nhất trong hoạt động ngân hàng hiện nay. Bởi vì tín dụng là nghiệp vụ quan trọng, cơ bản của ngân hàng,

Hình 1: Đường ước lượng của $\widehat{S}_0(t)$, $\hat{S}(t, X_1) = [\widehat{S}_0(t)]^{2,1650}$, $\hat{S}(t, X_2) = [\widehat{S}_0(t)]^{0,95154}$



Bảng 4: Kết quả mô hình logit

Status	Coef.	Std. Err.	z	P > z	[95% Conf. Interval]	
$x_1 = donbay$	1,158026	0,5501504	2,10	0,035	0,0797506	2,236301
$x_2 = loinhuan$	-3,220027	1,168675	-2,67	0,006	-5,510588	-0,9294657
$x_3 = nganhnghe$	0,6067207	0,2761069	2,20	0,028	0,0655612	1,14788
$x_4 = kinhnghiem$	-0,5197397	0,2744079	-1,89	0,058	-1,057569	0,01809
_cons	-0,1897811	0,4429909	-0,43	0,668	-1,058027	0,6784652

Nguồn: Tác giả tính toán từ dữ liệu nghiên cứu

chiếm tỷ lệ lớn trong các hoạt động kinh doanh của ngân hàng cả về doanh thu và lợi nhuận. Vì vậy, rủi ro trong lĩnh vực này cũng tỷ lệ thuận với tầm quan trọng của nó và chiếm phần lớn trong tổng mức rủi ro của hoạt động ngân hàng.

Trong những năm gần đây, nợ xấu và rủi ro tín dụng tiếp tục là vấn đề lớn cản trở sự phát triển toàn diện của hệ thống ngân hàng. Vì vậy, việc lượng hóa rủi ro tín dụng trở thành một vấn đề cấp thiết tại các ngân hàng thương mại. Việc lượng hóa rủi ro tín dụng một cách chính xác không chỉ giúp cho các ngân hàng chọn lọc khách hàng, ra quyết định cho vay đúng đắn, có chiến lược định lãi suất cho vay hiệu quả mà còn có thể giúp cho các ngân hàng thiết lập dự phòng rủi ro tín dụng và mức vốn kinh tế cần thiết để chống đỡ rủi ro.

Mô hình Cox có thể triển khai thành những mô hình phức tạp và chuyên sâu hơn cho các nghiên cứu khác nhau với nhiều biến và tương tác giữa các yếu tố rủi ro. Trong bài viết, những kết quả đạt được từ mô hình Cox sẽ cung cấp cho các tổ chức tín dụng cách nhìn thấu đáo hơn trong việc đo lường rủi ro vỡ nợ của khách hàng. Các ngân hàng và các tổ chức tín dụng khác nên cân nhắc lựa chọn mô hình này như là một công cụ phân tích, nhận diện và đo lường rủi ro tín dụng từ đó có những giải pháp thích hợp nhằm kiểm soát rủi ro.

Phương pháp này giúp định lượng xác suất vỡ nợ của khách hàng khác với phương pháp truyền thống vẫn dùng trong các ngân hàng và trong các nghiên cứu hiện tại ở Việt Nam.

PHỤ LỤC: KIỂM TRA GIẢ THIẾT PH VỚI CÁC BIẾN SỐ

Có nhiều cách để kiểm tra giả thiết PH với các biến số, ở đây ta dùng phương pháp kiểm định Goodness of Fit (GOF), phương pháp này cho giá trị p – value đối với mỗi biến để từ đó có kết luận về sự thỏa mãn giả thiết PH đối với các biến dự báo. Kiểm định được sử dụng ở đây là của Harrel & Lee (1986), đây là kiểm định đã được sửa đổi so với kiểm định ban đầu của Schoenfeld (1982). Với mỗi biến dự báo trong mô hình, phần dư Schoenfeld được xác định với mọi quan sát có xảy ra sự kiện, trong bài này ta có 4 biến dự báo là *donbay*, *loinhuan*, *nganhnghe* và *kinhngghiem* khi đó có 4 phần dư Schoenfeld được xác định với mỗi quan sát có xảy ra quá hạn nợ. Giả sử rằng quan sát k xảy ra quá hạn nợ tại thời điểm t_k , khi đó phần dư Schoenfeld của *donbay* là giá trị của *donbay* của quan sát này trừ đi trung bình của *donbay* của các quan sát khác vẫn còn nằm trong tập nguy cơ tại thời điểm t_k , tương tự với 3 biến dự báo còn lại. Ý tưởng của kiểm định thống kê là nếu giả thiết PH thỏa mãn đối với các biến cụ thể thì các phần dư Schoenfeld với các biến đó sẽ không tương quan với thời gian sống sót (hoặc thứ hạng của thời gian sống sót).

Kiểm tra giả thiết PH đối với các biến *donbay*, *loinhuan*, *nganhnghe* và *kinhngghiem*, kết quả được tóm tắt trong Bảng 5.

Bảng 5: Kiểm tra giả thiết PH đối với các biến dự báo

Test of proportional-hazards assumption				
Time: Rank(t)				
	rho	chi2	df	Prob>chi2
$x_1 = donbay$	0,01139	0,03	1	0,8618
$x_2 = loinhuan$	-0,01390	0,03	1	0,8658
$x_3 = nganhnghhe$	-0,05213	0,40	1	0,5294
$x_4 = kinhnghiem$	0,02904	0,13	1	0,7178
global test		0,56	4	0,9677

Nguồn: Tác giả tính toán từ dữ liệu nghiên cứu

Trong các nghiên cứu thực nghiệm, chỉ cần p-value > 0,10 thì ta đã có thể kết luận rằng giả thiết PH đối với biến đang nghiên cứu được thỏa mãn. Như kết quả ở trên, các giá trị p-value đối với các biến khá lớn, vì vậy ta có thể kết luận rằng giả thiết PH thỏa mãn với tất cả các biến, p-value tổng thể của mô hình trên là 0,9677 chứng tỏ mô hình Cox với 4 biến dự báo là phù hợp.

Tài liệu tham khảo

- Altman Edward I., Saunders A. (1998), 'Credit risk measurement: Developments over the last 20 years'. *Journal of Banking and Finance*, vol. 21, 1721-1742.
- Altman, Edward I. (1968), 'Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy', *Journal of Finance*, 23, 4.
- Baba, N. & Goko, H. (2006), *Survival analysis of hedge funds*, Bank of Japan, Working Papers Series No.06-E-05.
- Beran, J. & Djaidja, A.K. (2007), 'Credit risk modeling based on survival analysis with inmunes', *Statistical Methodology*, 4, 251- 276.
- Chen, S., Härdle, W.K. & Moro, R.A. (2006), 'Estimation of default probabilities with support vector machines', *Center of Applied Statistics and Economics (CASE)*, Humboldt Universität zu Berlin, Germany, SFB 649 discussion paper No. 2006-077.
- Cox, D.R. (1972), 'Regression models and life-tables (with discussion)', *J. Royal Statist. Society, Series B*, 74 187-220.
- Crouhy, M., Galai, D. & Mark, R. (2000), 'A comparative analysis of current credit risk models', *Journal of Banking and Finance*, 24, 59-117.
- Hamerle, A., Liebig, T. & Rösch, D. (2003), 'Credit risk factor modeling and the Basel II IRB Approach', *Deutsche Bundesbank Discussion Paper Series 2, Banking and Financial Supervision*, document No. 02/2003.
- Hand, D.J. (2001), 'Modelling consumer credit risk', *IMA Journal of Management Mathematics*, 12, 139-155.
- Hanson, S. & Schuermann, T. (2004), 'Estimating probabilities of default', *Federal Reserve Bank of New York*,
- Harrel, F. & Lee, K. (1986), *Proceedings of the Eleventh Annual SAS User's Group International*, 823-828.
- International Scientific Conference Managing and Modelling of Financial Risks Ostrava VSB-TU Ostrava*, Faculty of Economics, Finance Department.
- Malik, M. & Thomas, L. (2006), *Modeling Credit Risk of Portfolio of Consumer Loans*, University of Southampton, School of Management, Working Paper, Series No. CORMSIS-07-12.
- Martina Novotná (2012), 'The use of different approaches for credit rating prediction and their comparison', 6th
- Narain, B. (1992), 'Survival analysis and the credit granting decision', in *Credit Scoring and Credit Control*, Thomas L., Crook, J.N. & Edelman, D.B. (eds.), OUP: Oxford, 109-121.
- Nguyễn Trọng Hòa (2010), 'Xây dựng mô hình xếp hạng tín dụng đối với các doanh nghiệp Việt Nam trong nền kinh tế chuyên đổi', luận án tiến sĩ kinh tế.
- Nguyễn Văn Đức (2012), 'Phân tích danh mục tín dụng - xác suất không trả được nợ PD', luận văn thạc sĩ.
- Saunders, A. (1999), *Credit Risk Measurement: New Approaches to Value at Risk and Other Paradigms*, John Wiley & Sons, New York.
- Schoenfeld, D. (1982), 'Partial residuals for the proportional hazards model', *Biometrika*, 69: 51-55.
- Staff Report, no. 190.
- Stepanova, M. & Thomas, L. (2002), 'Survival analysis methods for personal loan data', *Operations Research*, 50, 277-289.
- Trương Thị Thảo My (2014), 'Ước lượng xác suất không trả được nợ của khách hàng doanh nghiệp tại Ngân hàng TMCP Công Thương Việt Nam – Chi nhánh Bến Tre', luận văn thạc sĩ kinh tế.
- Wang, Y., Wang, S. & Lai, K.K. (2005), 'A fuzzy support vector machine to evaluate credit risk', *IEEE Transactions on Fuzzy Systems*, 13, 820-831.