ÁP DỤNG MÔ HÌNH RỦI RO NHÓM VÀ PHƯƠNG PHÁP BAYES TRONG ƯỚC LƯỢNG CHI PHÍ KHÁM CHỮA BỆNH Ở VIỆT NAM

Phạm Thị Hồng Thắm
Khoa Toán Kinh tế, Đại học Kinh tế Quốc dân
Email: thamtkt@neu.edu.vn

Ngày nhận: 27/8/2018
Ngày nhận bản sửa: 02/10/2018
Ngày duyệt đăng: 15/10/2018

Tóm tắt:
Bài nghiên cứu này áp dụng mô hình rủi ro nhóm và phương pháp Bayes để dự báo chi phí khám chữa bệnh do Bảo hiểm Y tế chi trả ở Việt Nam. Theo cách này, người bệnh được chia thành nhiều nhóm dựa trên các yếu tố nhân khẩu và phi nhân khẩu, và chi phí khám chữa bệnh trong năm của một người trong mỗi nhóm được giả định là có phân phối như nhau và được mô hình hóa bởi số lượt khám chữa bệnh và chi phí của mỗi lượt. Tham số của mô hình được ước lượng bằng phương pháp Bayes. Phương pháp mô phỏng Monte – Carlo được áp dụng để xác định được phân phối xác suất của tổng chi phí khám chữa bệnh do Bảo hiểm Y tế chi trả trong những năm tiếp theo. Từ kết quả mô phỏng, bài viết đưa ra được một số đặc trưng liên quan đến chi phí khám chữa bệnh bằng bảo hiểm y tế trong tương lai.
Từ khóa: Mô hình rủi ro nhóm, ước lượng Bayes, chi phí khám chữa bệnh, mô phỏng Monte – Carlo.

Collective Risk Models and Bayesian Estimation in Predicting Health Cost in Vietnam

Abstract:
This paper applies the collective risk model and the Bayesian method to predict the healthcare costs paid by the Social Health Insurance in Vietnam. By employing these approaches, the insured people are divided into various groups based on their demographic and non-demographic characteristics. In each group, the annual total health cost for each person is assumed to be independently distributed, depending on the number of visits and the cost of each visit. The parameters of the model are estimated by the Bayesian method. Using Monte-Carlo simulation technique, the author generates the future total cost along with its probability distribution and its parameters for the coming years. The simulated cost would help to draw some characteristics of the health cost in the future.
Keywords: Collective risk models, Baysian estimation, health cost, Monte – Carlo simulation.

1. Đặt vấn đề

Theo Báo cáo Các chỉ số phát triển thế giới (WDI) năm 2016 của Ngân hàng Thế giới (2017), chi tiêu y tế bình quân đầu người của Việt Nam theo giá hiện hành vào năm 2015 là 142,4 đô la Mỹ, bằng 3,9 lần so với năm 2005 và 1,8 lần so với năm 2010. Tỷ lệ chi tiêu tiền túi (tính bằng chi tiêu tiền túi so với tổng chi tiêu khám, chữa bệnh) giảm từ 68% vào năm 2005 xuống 47% vào năm 2010 và xuống 36,7% vào năm 2015. Một trong nhiều yếu tố làm giảm chi tiêu tiền túi của người dân trong khám chữa bệnh là việc bảo hiểm y tế ngày càng đóng vai trò quan trọng hơn trong việc giảm gánh nặng chi trả khám chữa bệnh (Bộ Y tế, 2011). Tính toán gần đây cho thấy tính trung bình, bảo hiểm y tế bao phủ khoảng 80-85% tổng chi phí khám chữa bệnh nội trú

và ngoại trú ở Việt Nam, đặc biệt là cho những nhóm người cao tuổi – nhóm với rủi ro sức khoẻ cao và tần suất và chi phí khám, chữa bệnh lớn (Long & cộng sự, 2016; Kelly & cộng sự, 2016).

Tuy vậy, một vấn đề phát sinh trong khám, chữa bệnh bằng bảo hiểm y tế ở Việt Nam là có sự gia tăng đột biến trong một vài năm gần đây. Ngay cả khi không tính tới các điều chỉnh chính sách có liên quan trực tiếp tới việc thay đổi giá dịch vụ (như Thông tư 37/2015 của Bộ Y tế (2015) về quy định giá viện phí) thì tổng chi phí do bảo hiểm y tế chi trả vẫn tăng tới 15-20%/năm. Một số nghiên cứu đã cho thấy rằng việc sử dụng và chi tiêu chưa có hiệu quả cho thuốc và dịch vụ kỹ thuật là nguyên nhân chính dẫn tới việc gia tăng lớn về chi phí (Kelly & cộng sự, 2016). Vì vậy, mục tiêu cải cách bảo hiểm y tế thực chất là nhằm cung cấp dịch vụ tốt hơn cho người bệnh cũng như là kiểm soát được chi phí hợp lý trong khám chữa bệnh và giảm tối đa tiền túi của người dân.

Việc ước lượng đúng chi phí khám chữa bệnh có vai trò quan trọng hàng đầu cho mục tiêu cải cách bảo hiểm y tế cũng như tăng tính giải trình trong chăm sóc sức khoẻ cho người dân. Cho tới nay, ở Việt Nam cũng có một số nghiên cứu ước lượng và dự báo chi phí khám chữa bệnh. Tuy nhiên, các nghiên cứu này chủ yếu sử dụng tính toán bằng bảng biểu thông thường. Theo hiểu biết của tác giả, tới nay, chưa có nghiên cứu nào ở Việt Nam sử dụng các mô hình Toán kinh tế trong ước lượng và dự báo chi phí khám chữa bệnh do bảo hiểm chi trả. Bài nghiên cứu này sử dụng mô hình rủi ro nhóm với các phương pháp ước lượng và mô phỏng để đưa ra các dự báo với bối cảnh của bảo hiểm y tế Việt Nam.

2. Tổng quan nghiên cứu

Theo Astolfi & cộng sự (2012) thì các mô hình dự báo chi phí y tế có thể được xây dựng cho từng cá nhân, cho một nhóm đặc trưng hoặc cho toàn bộ dân số. Bên cạnh đó, các mô hình cũng có thể chỉ chú trọng vào một loại chi phí khám chữa bệnh cụ thể như chi tiêu y tế công, chi tiêu y tế tư nhân hoặc chi phí tiền túi của hộ gia đình cho khám chữa bệnh. Bằng cách kết hợp đơn vị phân tích, mức độ chi tiết về chi phí y tế cần dự báo và các câu hỏi chính sách đặt ra thì theo Culyer & Newhouse (2000), có thể chia ra ba lớp mô hình dự báo chi phí y tế là: lớp mô hình mô phỏng vi mô (Micro-simulation models), lớp mô hình ở cấp độ vĩ mô (Macro-level models) và lớp mô hình theo thành phần (Component-based models).

Lớp mô hình mô phỏng vi mô tập trung vào từng cá nhân và coi nó như một đơn vị phân tích (Brown & cộng sự, 2009) để từ đó mô phỏng cho một tổng thể dân số. Các cá nhân được mô phỏng mang các đặc trưng và hành vi đại diện cho tổng thể. Các vấn liên quan đến sức khỏe trong toàn bộ cuộc đời mỗi con người (như mang thai và sinh nở; chịu các yếu tố rủi ro như huyết áp cao, mỡ máu hay nghiện thuốc lá; nguy cơ gia tăng các loại bệnh không lây nhiễm và mạn tính như ung thư, tiểu đường, tim mạch…) sẽ được mô phỏng lại. Các chi phí y tế liên quan cho mỗi tình huống bệnh tật cũng được mô phỏng để từ đó đưa ra được các ước lượng về chi phí y tế.

Ngược lại, lớp mô hình ở cấp độ vĩ mô lại coi tổng chi phí y tế như một đơn vị phân tích (Baltagi & Moscone, 2010; Brown & cộng sự, 2009; Oliveira Martins & De la Maisonneuve, 2006; Warshawsky, 1994). Mô hình dự báo ở cấp độ vĩ mô được thực hiện dựa vào việc sử dụng các mô hình hồi quy với số liệu chuỗi thời gian hoặc số liệu mảng. Các dự báo có thể dựa vào việc ngoại suy thuần túy các mô hình thống kê hoặc có thể dựa vào giá trị dự báo của những biến giải thích quan trọng. Ví dụ, Getzen & Poullier (1992) đã đưa ra mô hình kinh tế lượng đơn giản để ước lượng và dự báo tổng chi phí y tế dựa theo mức tăng trưởng tổng sản phẩm quốc nội (GDP) và lạm phát trong quá khứ.

Lớp mô hình theo thành phần dự báo chi phí y tế dựa vào việc phân chia các đối tượng nghiên cứu theo các thành phần như theo nhóm chi trả, theo nhà cung cấp dịch vụ, theo các dịch vụ đã sử dụng hay theo các nhóm đối tượng cá nhân. Một lớp mô hình con rất quan trọng trong lớp mô hình này là lớp mô hình dự báo theo đoàn hệ (cohort-based models), trong đó các cá nhân được nhóm lại dựa theo các đặc điểm nhân khẩu (như tuổi, giới tính) và/hoặc phi nhân khẩu (như tình trạng bệnh tật...). Mỗi người trong nhóm được coi như có chi phí đồng đều và được mô hình hóa theo một dạng nào đó. Số người trong mỗi nhóm được dự báo có tính đến sự biến động của dân số. Tổng chi phí y tế cho mỗi nhóm đối tượng sẽ được tính bằng cách nhân chi phí của một người với số người trong nhóm và từ đó suy ra dự báo cho chi phí của toàn bộ tổng thể (Baltagi & Moscone, 2010; Przywara, 2010; Ringel & cộng sự, 2010; Vos & cộng sự, 2007; Racic, 1997).

Nằm trong lớp mô hình theo thành phần, mô hình rủi ro nhóm là một trong những mô hình kinh điển

trong mô hình hóa các tổn thất trong lĩnh vực bảo hiểm, nhất là bảo hiểm phi nhân thọ. Mô hình này dựa trên giả thiết tất cả các cá nhân trong một danh mục bảo hiểm cùng đối mặt với rủi ro như nhau. Do đó giá trị của mỗi lần yêu cầu bồi thường X_i sẽ có phân phối xác suất giống nhau. Với giả thiết số lần xảy ra tổn thất/số lần yêu cầu bồi thường N cũng là biến ngẫu nhiên thì tổng thiệt hại của danh mục S sẽ có phân phối phức hợp.

Có một lượng lớn các nghiên cứu về lớp mô hình này về cả lý thuyết và thực tiễn. Các nghiên cứu lý thuyết tập trung vào xây dựng các dạng phân phối xác suất lý thuyết cho S dựa vào các phân phối xác suất của N và X_i. Phân phối Poisson là một trong những phân phối được sử dụng nhiều nhất trong mô hình hóa N (Klugman & cộng sự, 2012). Tuy nhiên, phân phối Poisson không phải khi nào cũng phù hợp với dữ liệu thực tế (ví dụ như trong trường hợp phương sai lớn hơn nhiều so với trung bình) và khi đó thì phân phối nhị thức âm lại chiếm ưu thế. Các phân phối của X_i cũng rất đa dạng như phân phối mũ, chuẩn, LogNormal, Weibull, Pareto... Các lựa chọn cũng tùy thuộc vào dạng phân phối thực tế của dữ liệu (Kozubowski & Panorska, 2005; Hernández-Bastida & cộng sự, 2011).

Ngược lại với nghiên cứu lý thuyết, các nghiên cứu thực nghiệm đưa ra các phân phối thực nghiệm của bằng các phương pháp xấp xỉ dựa vào dữ liệu thực tế. Nhiều nghiên cứu ứng dụng mô hình rủi ro nhóm trong ứng dụng vào lĩnh vực bảo hiểm phi nhân thọ đã được thực hiện như Hayne (1989), Dickson & cộng sự (1998), Hernández-Bastida & cộng sự (2011), Klugman & cộng sự (2012), Migon & Moura (2005), Fellingham & cộng sự (2015), Yu (2015), Meyers & Schenker (1983), Meyers (2009), Heckman & Meyers (1983), Meyers (2007).

Về mặt thực nghiệm, có nhiều cách tiếp cận để ước lượng phân phối xác suất của S khi biết trước phân phối xác suất của N và X_i. Các phương pháp này được chia thành ba nhóm chính là (Hayne, 1989): xấp xỉ giải tích (analytic approximation), xấp xỉ phân phối (approximate distribution) và mô phỏng Monte-Carlo. Phương pháp xấp xỉ giải tích đưa ra cách tính xấp xỉ hàm phân phối xác suất dựa vào hàm phân phối xác suất đã biết của các thành phần trong mô hình. Phương pháp xấp xỉ phân phối được thực hiện bằng cách giả sử tổng chi phí tuân theo một phân phối nào đó. Từ đó, tính xấp xỉ các tham số của phân phối xác suất của S (Dickson & cộng sự, 1998). Không giống hai phương pháp trên, phương pháp mô phỏng Monte-Carlo được thực hiện dựa trực tiếp vào phân phối xác suất của các thành phần trong mô hình (Beekman & Fuelling, 1980). Phương pháp này không cần đưa ra dạng giải tích của phân phối xác suất của tổng chi phí và dễ dàng thực hiện ngay cả khi muốn kết hợp nhiều loại hình bảo hiểm và trong khoảng thời gian dài.

Liên quan đến các phương pháp ước lượng, phương pháp Bayes được sử dụng khá rộng rãi để ước lượng các tham số trong mô hình rủi ro nhóm (Fellingham & cộng sự, 2015; Klugman, 2013; Yu, 2015; Meyers & Schenker, 1983; Mildenhall, 2006; Hernández-Bastida & cộng sự, 2009). Phương pháp này được thực hiện dựa trên phân phối xác suất hậu nghiệm của tham số cần ước lượng với điều kiện dữ liệu mẫu đã biết và được cập nhật liên tục theo thời gian. Phương pháp này được đặc biệt sử dụng trong lý thuyết về độ tin cậy (Credibility Theory) để định giá và tính toán dự phòng cho các sản phẩm bảo hiểm. Thông tin về yêu cầu bồi thường của người được bảo hiểm được cập nhật hàng năm và làm cơ sở để định giá lại sản phẩm bảo hiểm của người đó cho năm tiếp theo. So với phương pháp ước lượng theo tần suất, ước lượng Bayes cho kết quả chính xác vượt trội (Klugman & cộng sự, 2012; Makov, 2001).

Trong nghiên cứu này, tác giả áp dụng mô hình rủi ro nhóm với dữ liệu khám chữa bệnh do bảo hiểm y tế chi trả ở Việt Nam. Bằng phương pháp ước lượng Bayes, tác giả thử nghiệm một số phân phối xác suất cho số lượt khám chữa bệnh và chi phí cho mỗi lần khám chữa bệnh. Các tham số ước lượng cho năm gốc sẽ được dùng để dự báo tổng chi phí khám chữa bệnh của năm sau bằng mô phỏng Monte-Carlo với điều kiện dữ liệu thực tế được cập nhật liên tục. Tổng chi phí khám chữa bệnh sau khi mô phỏng được sẽ được sử dụng để tính toán một số tham số liên quan.

3. Dữ liệu và phương pháp nghiên cứu

3.1. Dữ liệu nghiên cứu

Dữ liệu được sử dụng trong nghiên cứu này là dữ liệu thống kê khám chữa bệnh do bảo hiểm y tế chi trả và được phân loại theo khám chữa bệnh ngoại trú và khám chữa bệnh nội trú (tương ứng với các biểu mẫu 79b và 80b được bảo hiểm xã hội Việt Nam quy định). Các trường dữ liệu cho biết thông tin về người bệnh như ngày sinh, giới tính, mã thẻ, nơi khám chữa bệnh, loại bệnh, các loại chi phí cho

cả đợt khám chữa bệnh đó và phần chi phí do bảo hiểm y tế chi trả. Mỗi bản ghi tương ứng với một lượt khám chữa bệnh của một bệnh nhân nào đó trong năm. Dựa vào bộ dữ liệu này, tác giả có thể phân chia người bệnh thành nhiều nhóm dựa vào các thuộc tính khác nhau, cụ thể là theo giới tính, nhóm tuổi, loại khám chữa bệnh (nội trú, ngoại trú) và tuyến khám chữa bệnh (trung ương, tỉnh, huyện, xã). Mã thẻ bảo hiểm y tế cho phép phân biệt từng cá nhân người bệnh nên giúp tính được số lượt khám chữa bệnh của mỗi người trong năm.

3.2. Phương pháp nghiên cứu

Trong nghiên cứu này tác giả sử dụng mô hình rủi ro nhóm để mô hình hóa tổng chi phí khám chữa bệnh trong một năm của một người trong một nhóm đặc trưng nào đó. Sử dụng dữ liệu quá khứ được mô tả ở trên, các tham số của mô hình được ước lượng dựa vào phương pháp Bayes. Áp dụng phương pháp mô phỏng Monte-Carlo, tác giả sẽ mô phỏng tổng chi phí khám chữa bệnh của mỗi người trong nhóm. Kết quả của mô phỏng sẽ cho ta ước lượng của tổng chi phí khám chữa bệnh do bảo hiểm y tế cùng phân phối xác suất thực nghiệm của nó. Sử dụng phương pháp kiểm định phân phối xác suất Kolmogorov–Smirnov, tác giả kiểm chứng sự phù hợp của mô hình với dữ liệu thực tế. Cuối cùng, với mô hình phù hợp và kết quả mô phỏng được, nghiên cứu có thể đưa ra các ước lượng về tổng chi phí khám chữa bệnh do bảo hiểm y tế chi trả trong tương lai.

Trong mục này, tác giả sẽ trình bày mô hình rủi ro nhóm, phương pháp ước lượng Bayes, phương pháp mô phỏng Monte - Carlo.

3.2.1. Mô hình rủi ro nhóm

Theo Wuthrich (2017), biến ngẫu nhiên S được gọi là tuân theo mô hình rủi ro nhóm nếu được viết dưới dạng:

$$S = \begin{cases} \sum_{j=1}^{N} X_j, & N \neq 0 \\ 0, & N = 0 \end{cases}$$

với các giả thiết sau:

N là biến ngẫu nhiên rời rạc không nhận giá trị âm;

X_1, X_2,… là các biến ngẫu nhiên với tập giá trị dương, độc lập và có cùng phân phối xác suất;

N và (X_1, X_2,…) độc lập với nhau.

Phân phối xác suất của S sẽ là phân phối dạng phức (compound distribution) do nó phụ thuộc đồng thời vào phân phối xác suất của N và các X_j . Các phân phối xác suất của N có thể kể đến là Poisson, Nhị thức (Binomial), Nhị thức âm (Negative Binomial),...Trong đó, các X_j lại có thể có các phân phối như Mũ, Log-normal, Gamma, Pareto...

Trong nghiên cứu này, S được dùng để mô hình hóa tổng chi phí khám chữa bệnh được bảo hiểm y tế chi trả của một người bệnh trong một nhóm trong một năm;N là số lượt khám chữa bệnh của người đó trong năm và X_j là chi phí khám chữa bệnh được bảo hiểm y tế chi trả ở lần khám chữa bệnh thứ *j*.

3.2.2. Ước lượng Bayes

Phần này sẽ trình bày phương pháp Bayes để ước lượng các tham số trong các phân phối của N và X_j. Phương pháp Bayes coi tham số θ cần ước lượng như một giá trị cụ thể của biến ngẫu nhiên Θ và Θ được ước lượng dựa vào phân phối xác suất hậu nghiệm của nó với điều kiện đã biết mẫu dữ liệu cụ thể $x = (x_1, x_2,\ldots, x_n)$ của mẫu ngẫu nhiên $X = (X_1,X_2,\ldots,X_n)$.

Cụ thể, phương pháp ước lượng Bayes được Tse(2009) tóm tắt như sau:

- X là biến ngẫu nhiên tổn thất (loss variable) cần nghiên cứu (số lượt khám chữa bệnh, chi phí cho mỗi lượt khám chữa bệnh), có phân phối xác suất phụ thuộc vào tham số θ, được coi như một giá trị cụ thể của biến ngẫu nhiên Θ.

- Θ có phân phối xác suất thống kê tiên nghiệm (prior) với hàm mật độ xác suất $f_\Theta(\theta)$.

- Hàm mật độ xác suất của X với điều kiện θ, kí hiệu $f_{X|\Theta}(x,\theta)$,được gọi là hàm hợp lý, được xác định bởi:

$$f_{X|\Theta}(x,\theta) = \prod_{i=1}^{n} f_{X|\Theta}(x_i,\theta)$$

Trong đó: $x = (x_1, x_2, \ldots,x_n)$ là mẫu cụ thể của mẫu ngẫu nhiên $X = (X_1, X_2,\ldots,X_n)$ có cùng phân phối xác suất với X.

- Dựa vào dữ liệu mẫu cụ thể ***x***, phân phối xác suất của θ được cập nhật. Phân phối xác suất của Θ với điều kiện ***x*** được gọi là phân phối hậu nghiệm (posterior), với hàm mật độ xác suất $f_{\Theta|X}(\theta|x)$ được xác định bởi:

$$f_{\Theta|\mathbf{X}}(\theta|x) = \frac{f_{\Theta X}(\theta,x)}{f_X(x)} = \frac{f_{X|\Theta}(x,\theta)f_\Theta(\theta)}{\int_\theta f_{X|\Theta}(x,\theta)f_\Theta(\theta)d\theta}$$

- Ước lượng tham số của X, là hàm của Θ, được

Bảng 1: Các tham số ước lượng được từ số liệu theo các năm 2012 đến 2016

Năm	$\alpha_1 = \alpha_2$	β_1	β_2
2012	4.154	24.224	11.618,33
2013	8.906	51.824	26.520,11
2014	14.935	86.112	47.777,06
2015	21.563	124.712	72.163,86
2016	25.020	143.072	83.506,74

Nguồn: Tác giả tự tính toán.

xác định dựa vào phân phối hậu nghiệm của Θ, được gọi là ước lượng Bayes, đồng thời là ước lượng của tổn thất trong tương lai.

Như vậy phân phối hậu nghiệm của Θ hoàn toàn được xác định dựa vào phân phối tiên nghiệm của nó $f_\Theta(\theta)$ và hàm hợp lý $f_{X|\Theta}(x,\theta)$. Ta xét một số trường hợp cụ thể:

Trường hợp X phân phối Poisson – Gamma: Biến ngẫu nhiên X phân phối Poisson với tham số λ, với λ là giá trị cụ thể của biến ngẫu nhiên Λ phân phối Gamma G(α,β). Hàm mật độ tiên nghiệm của Λ có dạng:

$$f_\Lambda(\lambda;\alpha,\beta) = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)}$$

và hàm hợp lý của X có dạng:

$$f_{X|\Lambda}(\boldsymbol{x},\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

Khi đó phân phối hậu nghiệm của Λ với điều kiện x cũng là phân phối Gamma G(α*,β*) với tham số được tính bởi:

$$\alpha^* = \alpha + \sum_{i=1}^{n} x_i \,;\, \beta^* = \beta + n$$

với hàm mật độ xác suất hậu nghiệm:

$$f_{\Lambda|\mathbf{X}}(\lambda|x) \propto \lambda^{\alpha+\sum_{i=1}^{n} x_i - 1} e^{-\lambda(\beta+n)}$$

Trường hợp X có phân phối Mũ – Gamma: Biến ngẫu nhiên X phân phối Mũ ε(γ), với hàm hợp lý:

$$f_{X|\Lambda}(\boldsymbol{x},\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

và biến ngẫu nhiên Γ phân phối Gamma G(a,b). Khi đó phân phối hậu nghiệm của Γ với điều kiện x cũng là phân phối Gamma G(a*,b*) với tham số được tính bởi:

$$a^* = a + n; b^* = b + \sum_{i=1}^{n} x_i$$

với hàm mật độ xác suất hậu nghiệm:

$$f_{\Gamma|\mathbf{X}}(\gamma|x) \propto \gamma^{a+n-1} e^{-\gamma(b+\sum_{i=1}^{n} x_i)}$$

3.2.3. Mô phỏng Monte – Carlo

Giá trị của tổng chi phí khám chữa bệnh do bảo hiểm y tế chi trả cho một người thuộc nhóm đối tượng đang xét được mô phỏng theo các bước sau:

- Tạo ra số λ_1 từ phân phối xác suất của biến ngẫu nhiên Λ ~ G(α,β);

- Tạo ra n_1 từ phân phối xác suất của biến ngẫu nhiên N~P(λ_1), là số lượt khám chữa bệnh;

- Tạo ra γ_1 từ phân phối xác suất của biến ngẫu nhiên Γ~ G (a,b);

- Tạo ra n_1 giá trị x_1, x_2,…, x_{n1} từ phân phối xác suất của biến ngẫu nhiên X~ ε(γ_1), là chi phí cho mỗi lượt khám chữa bệnh;

- Tổng $s_1 = x_1 + x_2 + \cdots + x_{n1}$ cho ta một giá trị cụ thể của tổng chi phí khám chữa bệnh S của một người thuộc nhóm đối tượng được xét.

Lặp lại n lần các bước từ 1 đến 5 sẽ nhận được một mẫu n giá trị ngẫu nhiên của tổng chi phí khám chữa bệnh của một người trong nhóm đối tượng được xét trong một năm $s_1, s_2, \ldots, s_n$. Kết quả này cho ta phân phối thực nghiệm của tổng chi phí khám chữa bệnh

Bảng 2: Các thống kê của giá trị mô phỏng được năm 2017

Min	1st Qu.	Median	Mean	3rd Qu.	Max
0,20021	1,04394	2,47809	3,56643	4,96188	34,89162

Nguồn: Tác giả tự tính toán.

Bảng 3: Thống kê một số phân vị của chi phí

Quantile	0,90	0,91	0,92	0,93	0,94	0,95	0,96	0,97	0,98	0,99
Chi phí	8,242	8,611	8,964	9,402	10,018	10,658	11,520	12,639	13,832	16,534

Nguồn: Tác giả tự tính toán.

của một người trong nhóm. Từ đó, ta có thể tính toán tham số đặc trưng mà ta quan tâm như chi phí trung bình, phương sai của chi phí, các giá trị rủi ro tương ứng của tổng chi phí khám chữa bệnh do bảo hiểm y tế chi trả trong tương lai.

4. Kết quả ước lượng

Nghiên cứu sử dụng mô hình rủi ro nhóm cho trường hợp số lượt khám chữa bệnh phân phối Poisson – Gamma và chi phí khám chữa bệnh cho mỗi lượt có phân phối Mũ – Gamma. Tác giả tập trung vào dữ liệu của một nhóm người có thẻ bảo hiểm y tế, cụ thể là những người thuộc nhóm tuổi từ 60 đến 65, khám chữa bệnh nội trú ở tuyến tỉnh.

Với bộ dữ liệu trong 5 năm từ 2012 đến 2016, nghiên cứu lần lượt thực hiện theo các bước sau:

- Sử dụng dữ liệu lịch sử, ước lượng các tham số α_1, β_1 trong phân phối Gamma của Λ và ước lượng các tham số α_2, β_2 trong phân phối Gamma của Γ. Do năm 2012 là năm đầu tiên có dữ liệu nên các tham số được ước lượng theo phương pháp thống kê tần suất. Từ năm 2013 trở đi các tham số được ước lượng theo phương pháp Bayes, tham số cho năm sau được ước lượng dựa vào phân phối xác suất hậu nghiệm với điều kiện đã biết dữ liệu của những năm trước.

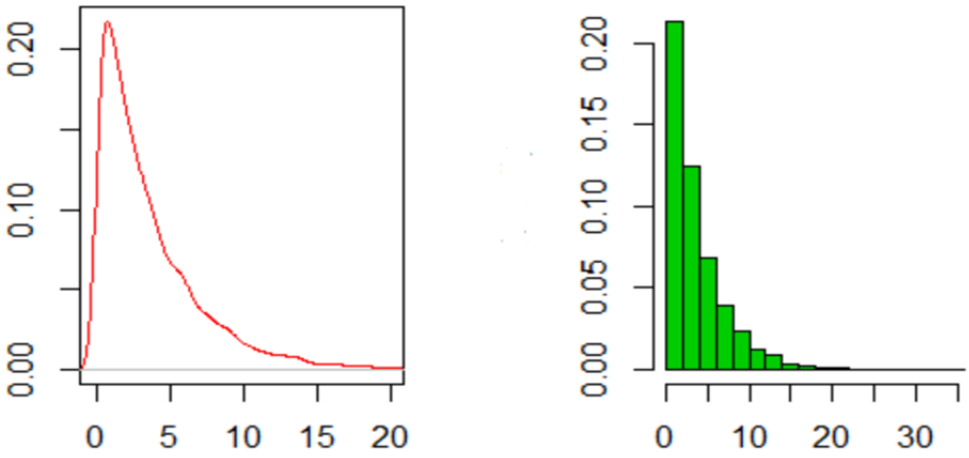
- Sử dụng tham số ước lượng được trong năm 2016 với dữ liệu được cập nhật qua từng năm theo phương pháp Bayes, ta mô phỏng tổng chi phí khám chữa bệnh của một người trong nhóm theo phương pháp Monte – Carlo đã trình bày trong mục 3.1.3 ở trên. Kết quả của mô phỏng cho ta phân phối xác suất thực nghiệm của tổng chi phí khám chữa bệnh của nhóm đó trong năm 2017.

Các tham số ước lượng được trong Bước 1 theo các năm được thống kê trong Bảng 1. Các tham số này được ước lượng lại theo phương pháp Bayes với điều kiệu dữ liệu cập nhật từng năm. Tham số ước lượng tương ứng với năm 2016 sẽ được dùng để mô phỏng trong dự báo cho năm 2017.

Thực hiện 100,000 lượt mô phỏng Monte – Carlo cho ta 100,000 giá trị tương ứng của tổng chi phí khám chữa bệnh do bảo hiểm y tế chi trả của một người trong nhóm trong năm 2017. Các giá trị này là các giá trị cụ thể của biến ngẫu nhiên , đại diện cho tổng chi phí ta đang cần nghiên cứu. Do đó, các tham số đặc trưng của biến ngẫu nhiên này có thể được ước lượng bởi các đặc trưng của mẫu vừa mô phỏng. Bảng 2 cho ta một số thống kê đơn giản của (đơn vị tính: triệu đồng).

Từ kết quả của Bảng 2, đối với một người có thẻ trong nhóm tuổi từ 60 đến 64, khám chữa bệnh nội trú ở tuyến tỉnh có cho ta chi phí bình quân vào khoảng trên 3,5 triệu đồng/năm. Tương ứng các chi

Hình 1: Đồ thị hàm mật độ xác suất và dạng cột của chi phí mô phỏng được



Nguồn: Tác giả tự tính toán.

phí lớn nhất và nhỏ nhất là hơn 30 triệu đồng và 200.000 đồng.

Sử dụng dữ liệu mô phỏng được ta cũng có thể vẽ đồ thị hàm mật độ xác suất thực nghiệm của phân phối và biểu đồ cột tương ứng như Hình 1. Hình này mô tả dạng phân phối của chi phí khám chữa bệnh. Đây là phân phối lệch phải nên khả năng mức chi phí cao là rất nhỏ so với mức chi phí thấp. Mức chi phí có khả năng xảy ra nhiều nhất, nhỏ hơn mức chi phí trung bình.

5. Kết luận

Bài nghiên cứu này là nghiên cứu đầu tiên sử dụng mô hình rủi ro nhóm cho bộ dữ liệu khám chữa bệnh bằng bảo hiểm y tế ở Việt Nam. Nghiên cứu đã trình bày phương pháp ước lượng tham số của mô hình theo cách tiếp cận Bayes, tức là xem tham số như một biến ngẫu nhiên và ước lượng nó dựa vào dữ liệu cập nhật hàng năm. Phương pháp mô phỏng Monte-Carlo cũng được sử dụng để đưa ra phân phối xác suất của chi phí khám chữa bệnh trong tương lai.

Ở những nghiên cứu sau, tác giả sẽ cố gắng đưa vào mô hình nhiều dạng phân phối khác nhau của số lượt KCB và chi phí cho mỗi lần KCB cũng như thử với nhiều nhóm hơn. Như thế sẽ có thể so sánh các mô hình để đưa ra dự báo phù hợp nhất cho mỗi nhóm người tham gia bảo hiểm.

Tài liệu tham khảo:

Astolfi Roberto, Luca Lorenzoni & Jillian Oderkirk (2012), 'A comparative analysis of health forecasting methods', *OECD Health Working Paper No. 104*, OECD.

Baltagi, B.H. & Francesco, M. (2010), 'Health care expenditure and income in the OECD reconsidered: Evidence from panel data', *Economic Modelling*, 27(4), 804-811.

Beekman, J.A. & Clinton, P.F. (1980), 'Simulation of a multi risk collective model', in *Computational Probability*, Elsevier, 287-301.

Bộ Y tế (2011), *Strengthening management capacity and reforming health financing to implement the five-year health sector plan 2011-2015*, Hà Nội.

Brown, L., Harris, A., Picton, M., Thurecht, L., Yap, M., Harding, A. & Richardson, J. (2009), 'Linking microsimulation and macro-economic models to estimate the economic impact of chronic disease prevention', in *New frontiers in microsimulation modelling*, A. Zaidi, A. Harding & P. Williamson (Eds.), Surrey, UK: Ashgate Publishing Limited, 527-555.

Culyer, A.J. & Joseph, P.N. (2000), *Handbook of health economics*, Elsevier.

Dickson David CM, Leanna M Tedesco & Ben Zehnwirth (1998), 'Predictive aggregate claims distributions', *Journal of Risk and Insurance*, 65(4), 689-709.

Fellingham, G.W., Athanasios, K. & Brian, M.H. (2015), 'Bayesian nonparametric predictive modeling of group health claims', *Insurance: Mathematics and Economics*, 60, 1-10.

Getzen, T.E. & Jean, P.P. (1992), 'International health spending forecasts: concepts and evaluation', *Social Science & Medicine*, 34(9), 1057-1068.

Giang Thanh Long , Phạm Thị Hồng Thắm & Phạm Lê Tuấn (2016), 'Bảo hiểm y tế trong chăm sóc sức khỏe cho NCT ở Việt Nam', *Tạp chí Tạp chí Kinh tế và Phát triển*, 9/2016, 38-48.

Hayne, R.M. (1989), 'Application of Collective Risk Theory to Estimate Variability in Loss Reserves', *Proceedings of the Casualty Actuarial Society*, Casualty Actuarial Society Forum, Trang 77-97.

Heckman, P.E. & Glenn, G.M. (1983), 'The calculation of aggregate loss distributions from claim severity and claim count distributions', *Proceedings of the Casualty Actuarial Society*, Casualty Actuarial Society Forum, 49-66.

Hernández-Bastida, A., Fernández-Sánchez, M.P. & Gómez-Déniz, E. (2009), 'The net Bayes premium with dependence between the risk profiles', *Insurance: Mathematics and Economics*, 45(2), 247-254.

Hernández-Bastida A., Fernández-Sánchez, M.P. & Gómez-Déniz, E. (2011), 'Collective risk model: Poisson–Lindley and exponential distributions for Bayes premium and operational risk', *Journal of Statistical Computation and*

Simulation, 81(6), 759-778.

Kelly, E., Giang, T.L. & Pham, T.H.T. (2016), *Actuarial analysis related to development of vietnam's social health insurance package*, United States Agency for International Development.

Klugman, S.A., Harry, H.P. & Gordon, E.W. (2012), *Loss models: from data to decisions*, John Wiley & Sons publisher.

Klugman Stuart A. (2013), *Bayesian statistics in actuarial science: with emphasis on credibility*, Springer Science & Business Media publisher.

Kozubowski Tomasz J. & Anna K. Panorska (2005), 'A mixed bivariate distribution with exponential and geometric marginals', *Journal of Statistical Planning and Inference*, 134(2), 501-520.

Makov Udi E. (2001), 'Principal applications of Bayesian methods in actuarial science: a perspective', *Tạp chí North American Actuarial Journal*, 5(4), 53-57.

Meyers Glenn (2009), 'Stochastic loss reserving with the collective risk model', *Variance*, 3(2), 239-269.

Meyers Glenn G. (2007), 'Estimating predictive distributions for loss reserve models', *Variance*, 1(2), 248-272.

Meyers Glenn & Nathaniel Schenker (1983), 'Parameter Uncertainty in the Collective Risk Model', *The Pacific Coast Archaeological Society 70*, 111-143.

Migon Helio S. & Fernando A.S. Moura (2005), 'Hierarchical bayesian collective risk model: an application to health insurance', *Insurance: Mathematics and Economics*, 36(2), 119-135.

Mildenhall Stephen J. (2006), 'A multivariate Bayesian claim count development model with closed form posterior and predictive distributions', *CAS Proceedings*, The Casualty Actuarial Society Forum, 451-493.

Ngân hàng Thế giới (2017), *World Development Indicators*, truy cập ngày 10 tháng 8 năm 2018, từ <https://data.worldbank.org/products/wdi>.

Oliveira Martins Joaquim & Christine De la Maisonneuve (2006), 'The drivers of public expenditure on health and long-term care: An integrated approach', *OECD Economic Studies*, 2006(2), retrieved on August 10th 2018, from <https://doi.org/10.1787/eco_studies-v2006-art11-en>.

Przywara Bartosz (2010), 'Projecting future health care expenditure at European level: drivers, methodology and main results', *European Economy Economic Papers*, 417, retrieved on August 10th 2018, from <http://ec.europa.eu/economy_finance/publications/economic_paper/2010/ecp417_en.htm>.

Racic Tatjana (1997), 'The actual uses of health service indicators and projections of health services expenditures in Croitia', *F&R Insurance Consulting*, 341-354.

Ringel Jeanne S., Christine Eibner, Federico Girosi, Amado Cordova & Elizabeth A. McGlynn (2010), 'Modeling health care policy alternatives', *Health Services Research*, 45(52), 1541-1558.

Tse Yiu-Kuen (2009), *Nonlife actuarial models: theory, methods and evaluation*, Cambridge University Press.

Vos Theo, John Goss, Stephen Begg & Nicholas Mann (2007), *Projection of health care expenditure by disease: a case study from Australia*, retrieved on August 10th 2018, from < https://www.researchgate.net/publication/43470170_Projection_of_Health_Care_Expenditure_by_Disease_A_Case_Study_from_Australia>.

Warshawsky Mark J. (1994), 'Projections of health care expenditures as a share of the GDP: actuarial and macroeconomic approaches', *Health Services Research*, 29(3), 293-313.

Wuthrich Mario V. (2017), Non-life insurance: mathematics & statistics, retrieved on August 10th 2018, from <https://ssrn.com/abstract=2319328 or http://dx.doi.org/10.2139/ssrn.2319328>.

Yu Guang Qu (2015), 'Hierarchical bayesian modeling of health insurance claims', A Project Sub of Master of Science, Simon Fraser University.