วารสารคณิตศาสตร์ MJ-MATh 62(692) May–Aug, 2017
โดย สมาคมคณิตศาสตร์แห่งประเทศไทย ในพระบรมราชูปถัมภ์
http://MathThai.Org MathThaiOrg@gmail.com

การใช้ตัวแบบเชิงเส้นวางนัยทั่วไปในการประมาณค่าเบี้ยประกันภัยรถยนต์

Using Generalized Linear Models to Estimate Pure Premium for Motor Insurance

ชนกนาต สุขประยูร ตวงพร พัฒนบัณฑิต ภัทรพงศ์ สกุลวโรภาส
และ อุไรวรรณ เจริญกีรติกุล

Chanoknat Sukprayoon[1] Tuangporn Pattanabandit[2] Puttharapong Sakulwaropas[3]
and Uraiwan Jaroengeratikun[4]

[1-4]Department of Applied Statistics, Faculty of Applied Science,
King Mongkut's University of Technology North Bangkok,
Bangsue, Bangkok 10800, Thailand

Email: [3]tojob8445@gmail.com [4]uraiwan.j@sci.kmutnb.ac.th

บทคัดย่อ

ในธุรกิจประกันวินาศภัย การกำหนดราคาเบี้ยประกันภัยที่เหมาะสมและเพียงพอต่อความเสี่ยงที่บริษัทสามารถรับไว้ได้ถือเป็นหัวใจสำคัญในการบริหารจัดการความเสี่ยง วัตถุประสงค์ของงานวิจัยนี้เพื่อประยุกต์ใช้ตัวแบบเชิงเส้นวางนัยทั่วไปในการประมาณค่าเบี้ยประกันภัยแท้จริงในการประกันภัยรถยนต์ โดยใช้ข้อมูลกรมธรรม์รถยนต์ประเภทชั้น 1 ของบริษัทประกันแห่งหนึ่งในประเทศไทย ซึ่งพบว่ามีค่าส่วน ลด-เพิ่ม จากประวัติการเรียกร้องสินไหมทดแทน (NCB) และลักษณะการใช้งาน (Car Code) ที่มีผลกระทบต่อการประมาณเบี้ยประกันภัยแท้จริง ที่ระดับนัยสำคัญ 0.05

คำสำคัญ: เบี้ยประกันภัยแท้จริง ตัวแบบเชิงเส้นวางนัยทั่วไป ปัจจัยเสี่ยง ตัวแบบถดถอย

ABSTRACT

In casualty business, the optimal premium pricing and adequate cost for an insurance company is important in risk management. The aim of this research was to study in the application of generalized linear models to estimate pure premium for motor insurance. In this study, the data set was the claim of comprehensive motor insurance, which was provided by one of the insurance company in Thailand. The results of this study show that the risk factors significantly related to pure premium at the 0.05 level consisted of no claim bonus (NCB) and use of motor vehicle (Car Code).

Keywords: Pure Premium, Generalized Linear Models, Risk Factor, Regression Model

1. บทนำ

ในธุรกิจประกันวินาศภัย บริษัทประกันภัยทุกบริษัทต้องมีการกำหนดเบี้ยประกันภัย (Insurance Premium) หรือพิกัดที่เหมาะสมและเพียงพอต่อความเสี่ยงที่บริษัทสามารถรับไว้ได้ซึ่งผู้เอาประกันภัย (Insured) ชำระค่าเบี้ยประกันภัยให้กับบริษัทประกันภัย (Insurer) เพื่อได้รับความคุ้มครองความเสี่ยงภัย การกำหนดเบี้ยประกันภัยจำเป็นต้องใช้ข้อมูลที่เคยเกิดขึ้นจริงในอดีตมาทำการวิเคราะห์และใช้วิธีการทางสถิติเพื่อคาดการณ์เบี้ยประกันภัยที่เหมาะสมในอนาคต ซึ่งเป็นรายรับหลักของบริษัทที่เก็บจากผู้เอาประกันภัย จากสถิติที่เกิดภัยของผู้เอาประกันภัยแต่ละรายหรือเรียกว่าหน่วยเสี่ยงภัย (Exposure Unit) มีความแตกต่างกัน ดังนั้นกำหนดค่าเบี้ยประกันภัยให้มีความเพียงพอกับความเสี่ยงภัยที่อาจจะเกิดขึ้นในอนาคตจึงต้องคำนึงถึงประสบการณ์การเกิดภัยของหน่วยเสี่ยงภัยหรือตัวแปรปัจจัยเสี่ยง (Risk Factor) และถ้าบริษัทสามารถทราบถึงปัจจัยเสี่ยงที่มีผลกระทบกับค่าเบี้ยประกันภัยและตัวแบบการคำนวณค่าประมาณเบี้ยประกันภัยแท้จริง (Pure Premium) ที่เหมาะสมและมีความพอเพียงกับความเสี่ยงภัยที่บริษัทรับความเสี่ยงซึ่งจะเป็นสารสนเทศสำคัญให้กับบริษัทประกันภัยใช้กำหนดค่าเบี้ยประกันภัยในอนาคตได้ โดยปกติเบี้ยประกันภัยแท้จริงประกอบด้วย จำนวนครั้ง

(Claim Frequency) และมูลค่าความเสียหาย (Claim Cost or Claim Amount) ที่เกิดการเรียกร้องสินไหมทดแทนซึ่งในทางประกันภัยจะมีลักษณะเป็นการแจกแจงที่ไม่เป็นการแจกแจงปรกติ (Normal Distribution)

จากปัญหาและความสำคัญข้างต้น ในด้านคณิตศาสตร์ประกันภัยจึงได้มีผู้ที่ศึกษาและพัฒนาตัวแบบตัวแปรปัจจัยเสี่ยงและตัวแบบเชิงเส้นวางนัยทั่วไป (Generalized Linear Models: GLMs) เป็นวิธีการทางสถิติที่นำไปประยุกต์ใช้ในงานด้านการประกันภัย อาทิเช่น [1] ได้ศึกษา GLMs ในการประมาณความน่าจะเป็นของการเกิดความเสี่ยงภัยในการประกันภัยรถยนต์เพื่อประมาณเบี้ยประกันภัย ซึ่งต่อมา [2-4] ได้ใช้หลักวิธีการของ GLMs ไปประยุกต์ใช้ในการสร้างตัวแบบความเสี่ยงโดยเฉพาะในการประกันวินาศภัย แล้วในปี ค.ศ. 2015 [5] ได้พัฒนาวิธีการคำนวณค่าเบี้ยประกันภัยในข้อมูลการประกันภัยรถยนต์ของประเทศฝรั่งเศสที่เริ่มสัญญาประกันภัยในปี 2009 โดยใช้ GLMs ในการวิเคราะห์ข้อมูล ซึ่งทำให้ได้ทราบถึงปัจจัยเสี่ยงที่มีผลกระทบต่อเบี้ยประกันภัยแท้จริง

ในงานวิจัยนี้ใช้การวิเคราะห์ GLMs เพื่อคำนวณหาค่าเบี้ยประกันภัยแท้จริงที่เหมาะสมและเพียงพอกับความเสี่ยงภัยในอนาคตที่บริษัทประกันภัยควรจะเรียกจากผู้เอาประกันภัยแต่ละรายอย่างเป็นธรรมและข้อมูลที่นำมาศึกษานี้ เป็นข้อมูลกรมธรรม์การประกันภัยรถยนต์ประเภทชั้น 1 ของบริษัทประกันภัยแห่งหนึ่งในประเทศไทย ซึ่งข้อมูลกรมธรรม์ประเภทรถยนต์นั่ง ที่เริ่มสัญญาในปี พ.ศ.2550 ระยะเวลาสัญญา 1 ปีจำนวน 26,959 กรมธรรม์

2. วิธีการวิจัย

การวิจัยนี้ใช้ตัวแปรปัจจัยเสี่ยงที่จะศึกษาเป็นตามปัจจัยเสี่ยงภายใต้ข้อกำหนดเงื่อนไขพระราชบัญญัติประกันวินาศภัย พ.ศ.2535 ที่ได้กำหนดตัวแปรที่ใช้ศึกษาและมีลักษณะข้อมูลเบื้องต้นจากตารางข้อมูลพื้นฐานของหน่วยเสี่ยงภัยดังนี้

ตารางที่ 1 ตารางลักษณะเบื้องต้นของข้อมูลที่ใช้ในงานวิจัย

ส่วนของข้อมูลที่เป็นแบบกลุ่ม			
ปัจจัย	**ตัวแปร (X)**	**คำอธิบาย**	**ร้อยละของหน่วยเสี่ยงภัย**
1.ลักษณะการใช้งาน	-Car Code110 (base) -Car Code120	การใช้ส่วนบุคคล (รหัส110) การใช้เพื่อการพาณิชย์ (รหัส 120)	94.63% 5.37%
2.กลุ่มรถยนต์จำแนกตามยี่ห้อและรุ่น	-Car Group1 -Car Group2 -Car Group3 (base) -Car Group4 -Car Group5	กลุ่มที่ 1 กลุ่มที่ 2 กลุ่มที่ 3 กลุ่มที่ 4 กลุ่มที่ 5	0.11% 2.33% 67.19% 10.87% 19.50%
ส่วนของข้อมูลที่เป็นค่าตัวเลข			
ปัจจัย	**ตัวแปร (X)**	**คำอธิบาย**	**ค่าเฉลี่ย (ส่วนเบี่ยงเบนมาตรฐาน)**
1.ขนาดเครื่องยนต์	CC		2,243.44 (581)
2.จำนวนที่นั่ง	Seat		6.74 (0.68)
3.ทุนเอาประกันภัย	Sum Insured		654,762.90 (382,865)
4.ค่าส่วนลด-เพิ่ม จากประวัติการเรียกร้องสินไหมทดแทน	NCB (No Claim Bonus)	หน่วยเป็นร้อยละของเบี้ยประกันภัย มีค่าส่วนลดเพิ่มทีละ 5% ของเบี้ยประกันภัย -กรณีส่วนเพิ่มเบี้ยประกันภัยจะมีค่าเป็นจำนวนลบ -กรณีส่วนลดเบี้ยประกันภัยจะมีค่าเป็นจำนวนบวก	11.95 (17)
5.เบี้ยประกันภัยสุทธิ	Net Premium	หน่วยเป็นบาท ตั้งแต่ 1000 บาทขึ้นไป	17,707.81 (7,470)

การวิเคราะห์ GLMs ในการศึกษาตัวแปรปัจจัยเสี่ยงที่มีผลกระทบกับเบี้ยประกันภัยและตัวแบบการประมาณค่าเบี้ยประกันภัยนั้นจะพิจารณาทั้งตัวแบบการประมาณจำนวนครั้งการเรียกร้องสินไหมทดแทนและตัวแบบการประมาณมูลค่าสินไหมทดแทนด้วยวิธีการวิเคราะห์ตัวแบบถดถอยทวินามลบ (Negative Binomial Regression Model) [6] และตัวแบบถดถอยแกมมา (Gamma Regression Model) ตามลำดับ โดยจะเริ่มคัดเลือกปัจจัยเสี่ยงด้วยการวิเคราะห์ตัวแปรปัจจัยเดี่ยว (Univariate Analysis) ที่ระดับนัยสำคัญ (Significance Level) 0.25 [7] และนำตัวแปรปัจจัยเสี่ยงที่คัดเลือกได้ไปวิเคราะห์โดยใช้การวิเคราะห์หลายตัวแปร (Multivariate Analysis) ด้วยวิธีลดตัวแปร (Backward Elimination) ที่ระดับนัยสำคัญ 0.05 ซึ่งในการศึกษานี้จะใช้โปรแกรม R ในการวิเคราะห์และประมวลผล

2.1 ตัวแบบเชิงเส้นวางนัยทั่วไป (GLMs)

[8] ได้คิดค้นและนำ GLMs ไปใช้ในตัวแบบคณิตศาสตร์ประกันภัย ซึ่ง GLMs จะเป็นการศึกษาตัวแปรตาม (Dependent Variable: Y) ที่มีความสัมพันธ์กับตัวแปรปัจจัยเสี่ยง (X) โดยที่ Y มีการแจกแจงอยู่ในกลุ่มวงศ์เลขชี้กำลัง (Exponential Family) ซึ่งมีฟังก์ชันเชื่อมโยง (Link Function) แทนด้วย η ในรูปความสัมพันธ์เชิงเส้นกับ X ดังนี้

$$\eta = \mathrm{XB} \tag{1}$$

โดยที่ X คือเมตริกซ์ของปัจจัยเสี่ยงและ B เวกเตอร์ของสัมประสิทธิ์การถดถอยของปัจจัยเสี่ยง ที่ประมาณค่าด้วยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method)

การคำนวณเบี้ยประกันภัยแท้จริงจะพิจารณาจำนวนครั้งการเรียกร้องสินไหมทดแทน (N) และมูลค่าสินไหมทดแทน (C) ซึ่งให้เป็นอิสระต่อกัน จึงได้สมการคำนวณค่าเบี้ยประกันภัยแท้จริง ดังนี้

$$E(\textstyle\sum_{k=1}^{N_i} C_{ik}) = E(N_\mathrm{i})E(C_\mathrm{i}) \quad , \mathrm{i} = 1,2,3,\ldots,26959 \tag{2}$$

ในงานวิจัยนี้ใช้ GLMs ในการวิเคราะห์ โดยผลการวิเคราะห์จะเป็นสารสนเทศเพื่อใช้ประกอบการตัดสินใจให้กับผู้ปฏิบัติงานที่เกี่ยวข้องในสายงาน

ธุรกิจประกันวินาศภัย ซึ่งในทางปฏิบัติไม่ต้องใช้กระบวนการที่ซับซ้อนที่อาจทำให้เกิดต้นทุนที่สูงเกินความจำเป็นของบริษัท

ดังนั้น การศึกษาตัวแบบ N และ C ของงานวิจัยนี้จึงเลือกใช้การแจกแจงของทั้งสองตัวแบบให้เป็นรูปแบบทั่วไปที่ใช้ในสายงานธุรกิจประกันวินาศภัย นั่นคือ ให้ E(N) อยู่ในตัวแบบถดถอยแบบทวินามลบ และ E(C) อยู่ในตัวแบบถดถอยแบบแกมมา ซึ่งทั้งสองตัวแบบมีฟังก์ชันการเชื่อมโยงเป็น Log-Link function ดังนี้

$$ln(\mu) = \beta^t x$$

$$\mu = exp(\beta^t x) \qquad \mu \text{ คือ } E(N) \text{ หรือ } E(C) \quad (3)$$

โดยที่ β คือ ค่าสัมประสิทธิ์การถดถอยของปัจจัยเสี่ยงที่มีผลกระทบต่อ μ และ x คือ เวกเตอร์ของปัจจัยเสี่ยง

จากสมการที่ (2) ใช้ GLMs ในการประมาณเบี้ยประกันภัยแท้จริงได้ ดังนี้

$$E(\sum_{k=1}^{N_i} C_{ik}) = E(N_i)E(C_i) = exp(\beta_{freq}{}^t x_i)exp(\beta_{cost}{}^t x_i)$$

$$= exp((\beta_{freq}+\beta_{cost})^t x_i) \quad (4)$$

โดยที่ β_{freq} คือ ค่าสัมประสิทธิ์การถดถอยของปัจจัยเสี่ยงที่มีผลกระทบต่อ N และ β_{cost} คือ ค่าสัมประสิทธิ์การถดถอยของปัจจัยเสี่ยงที่มีผลกระทบต่อ C

3. ผลการวิจัย

จากผลการวิจัย พบว่ามีปัจจัยเสี่ยง NCB และ Car Code ที่มีผลกระทบต่อเบี้ยประกันภัยแท้จริง โดยที่ Car Code เป็นปัจจัยเสี่ยงเดียวเท่านั้นที่มีผลกระทบต่อมูลค่าสินไหมทดแทน ที่ระดับนัยสำคัญ 0.05 และเมื่อให้ค่าตัวแปรปัจจัยเสี่ยงอื่นๆ คงที่ในตัวแบบประมาณค่าเบี้ยประกันภัยแท้จริง ค่าปัจจัยเสี่ยง NCB มีค่าเปลี่ยนแปลง 1 หน่วย (ได้รับส่วนเพิ่ม 5% ของเบี้ยประกันภัย) จะส่งผลกระทบให้ค่าประมาณเบี้ยประกันภัยแท้จริงมีค่าเป็น e–0.0052(1) = 0.995 ซึ่งหมายความว่าจะมีค่าเบี้ยประกันภัยแท้จริงลดลง 0.5% และในส่วนปัจจัยเสี่ยง Car Code120 เมื่อให้ค่าตัวแปรปัจจัยเสี่ยงอื่นๆคงที่ในตัวแบบค่าประมาณเบี้ยประกันภัยแท้จริงจะส่งผลกระทบให้ค่าประมาณเบี้ยประกันภัยแท้จริงมีค่าเป็น e0.0638(1) = 1.066

เปรียบเทียบกับกลุ่ม Car Code110 นั่นคือ กรมธรรม์ในกลุ่ม Car Code120 จะต้องจ่ายเบี้ยประกันภัยแท้จริงมากกว่ากรมธรรม์ในกลุ่ม Car Code110 เท่ากับ 6.6% และค่าจากตารางที่ 2 จะได้ตัวแบบประมาณเบี้ยประกันภัยแท้จริง ดังนี้

$$E(\sum_{k=1}^{N_i} C_{ik}) = e^{9.4969 - 0.0052x_{NCB_i} + 0.0638x_{Car\ Code120_i}}$$

ตารางที่ 2 ผลการวิเคราะห์ GLMs กับการประมาณค่าเบี้ยประกันภัยแท้จริง

ปัจจัยเสี่ยง		**Intercept**	**NCB**	**Car Code120**
ตัวแบบถดถอยทวินามลบ	$\widehat{\boldsymbol{\beta}}_{freq}$	0.0491	–0.0009	0.0638
	SE($\widehat{\boldsymbol{\beta}}_{freq}$)	0.0076	0.0004	0.0222
	z-value	6.48	–2.60	2.87
	p-value	<.0500	<.0500	<.0500
ตัวแบบถดถอยแกมมา	$\widehat{\boldsymbol{\beta}}_{cost}$	9.4478	–0.0043	–
	SE($\widehat{\boldsymbol{\beta}}_{cost}$)	0.0260	0.0012	–
	t-value	362.91	–3.44	–
	p-value	<.0500	<.0500	–
ตัวแบบเบี้ยประกันภัยแท้จริง	$\widehat{\boldsymbol{\beta}}_{freq} + \widehat{\boldsymbol{\beta}}_{cost}$	9.4969	–0.0052	0.0638

จากตัวแบบในข้างต้น สามารถประมาณเบี้ยประกันภัยแท้จริง ได้ดังตัวอย่างต่อไปนี้ จากข้อมูลในรายกรมธรรม์ ผู้เอาประกันภัย มีค่า NCB เท่ากับ 35% และใช้รถยนต์เพื่อการพาณิชย์ (Car Code120) จะได้ค่าประมาณเบี้ยประกันภัยแท้จริงรายปีที่ต้องจ่ายให้บริษัทประกันภัย คือ

$$E(\sum_{k=1}^{N_i} C_{5k}) = e^{9.4969 - 0.0052(7) + 0.0638(1)} = 12{,}924.50 \text{ บาท}$$

4. สรุปผล

การใช้ GLMs ในการประมาณค่าเบี้ยประกันภัยรถยนต์ ซึ่งจากการวิเคราะห์ GLMs กับข้อมูลการประกันภัยรถยนต์ประเภทชั้น 1 ของบริษัทประกันภัยแห่งหนึ่งในประเทศไทย จะได้สมการดังนี้

$$E(\sum_{k=1}^{N_i} C_{ik}) = e^{9.4969 - 0.0052x_{NCB_i} + 0.0638x_{Car\ Code120_i}}$$

ซึ่งมีเพียงปัจจัยเสี่ยง NCB และ Car Code ที่มีผลกระทบต่อค่าประมาณเบี้ยประกันภัยแท้จริง ที่ระดับนัยสำคัญ 0.05 ดังนั้นการกำหนดค่าเบี้ยควรคำนึงถึงปัจจัยดังกล่าว ซึ่งผลการศึกษานี้จะเป็นแนวทางให้กับบริษัทประกันภัยนำไปประยุกต์ใช้กับข้อมูลที่รับเสี่ยงเพื่อสร้างตัวแบบการประมาณค่าเบี้ยประกันภัยแท้จริง ซึ่งผลการวิเคราะห์ข้อมูลของแต่ละบริษัทอาจจะให้ค่าปัจจัยเสี่ยงที่มีผลกระทบต่อค่าประมาณเบี้ยประกันภัยแท้จริงที่แตกต่างไปจากผลการศึกษานี้ เนื่องจากลักษณะข้อมูลการรับประกันความเสี่ยงภัยมีความแตกต่างกัน

เอกสารอ้างอิง

[1] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*, United States: John Wiley & Sons Inc., 2000.

[2] E. Ohlsson and B. Johansson, *Non-Life Insurance Pricing with Generalized Linear Models*, Springer-Verlag Berlin Heidelberg, 2010.

[3] J. Lemaire, *Automobile Insurance: Actuarial Models*, Dordrecht, Netherlands: Kluwer Academic, 1985.

[4] Joseph M. Hilbe, *Negative Binomial Regression*, 2nd ed. Cambridge, Cambridge University Press, 2011.

[5] M. David, "Auto Insurance Premium Calculation Using Generalized Linear Models," *Procedia Economics and Finance*, vol. 20, pp. 147-156, 2015

[6] P. Jong and G. Z. Heller, *Generalized Linear Models for Insurance Data*, Cambridge, Cambridge University Press, 2008.

[7] P. McCullagh and J. A. Nelder, *Generalized Linear Models*, 2nd ed. London, United States: John CRC Press, 1989.

[8] R. Kaas, M. Goovaerts, J. Dhaene and M. Denuit, *Modern Actuarial Risk Theory: Using R*, 2nd ed. Springer, Berlin, 2009