

การประยุกต์ใช้แบบจำลองปรับเรียบล็อกลิเนียร์เพื่อสร้างการแจกแจงความสูญเสีย ของสวัสดิการรักษายาบาลข้าราชการไทย

Applications of Log-linear Smoothing Models to fit Loss Distribution of Thai Government's officer Healthcare Benefit

รัติกานต์ สงขาว (Ratikarn Songkhaw)¹* ดร.อาณนท ศักดิ์วีระวิชญ์ (Dr.Armond Sakworawich)**

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองปรับเรียบล็อกลิเนียร์เพื่อพยากรณ์ความถี่และความรุนแรงของความสูญเสียของสวัสดิการรักษายาบาลข้าราชการไทย ประเภทผู้ป่วยนอกปี พ.ศ. 2558 โดยได้นำแบบจำลองปรับเรียบหนึ่งตัวแปรมาประยุกต์ใช้กับแบบจำลองความสูญเสีย วิธีในการนำค่าของฟังก์ชันของคะแนนของความถี่และฟังก์ชันของคะแนนของความรุนแรงของความสูญเสียของแต่ละโมเมนต์มาเข้าแบบจำลอง จะมีการปรับให้เป็นคะแนนมาตรฐาน (Z-score) ผลการวิจัยพบว่า จากการคัดเลือกแบบจำลองจากค่า G^2 และ AIC โดยเลือกค่าที่ต่ำที่สุด ปรากฏว่าทั้งสองเกณฑ์ได้เลือกแบบจำลองเดียวกัน (1) ผลการคัดเลือกแบบจำลองปรับเรียบล็อกลิเนียร์หนึ่งตัวแปรของความถี่ของความสูญเสีย คือ แบบจำลองที่ 2 ซึ่งประกอบด้วย โมเมนต์ที่ 1 - 5 (2) ผลการคัดเลือกแบบจำลองปรับเรียบล็อกลิเนียร์หนึ่งตัวแปรของความรุนแรงของความสูญเสีย คือ แบบจำลองที่ 1 ซึ่งประกอบด้วย โมเมนต์ที่ 1 - 4

ABSTRACT

The purpose of this research was to fit log-linear smoothing models to forecast frequency and severity of loss of Thai government's officer healthcare benefit in case of outpatients who admitted to hospital in 2015. This study attempts to apply the univariate smoothing model to fit loss modeling. The methods take the value of score function of frequency and score function of severity of loss each moments are generalized as standard scores (Z-score). The results of the study revealed, from the selected results for one model from G^2 and AIC based on the lowest value showed that both criteria have chosen the same. (1) Univariate smoothing model selects results for one model of frequency of loss from the second model, which consists of 1st - 5th moment. (2) Univariate smoothing model selects results for one model of severity of loss from the first model, which consists of 1st - 4th moment.

คำสำคัญ: แบบจำลองปรับเรียบล็อกลิเนียร์ แบบจำลองความสูญเสีย

Keywords: Log-linear smoothing models, Loss modeling

¹ Correspondent author: ratikarnsongkhaw@gmail.com

* นักศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์ คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

** ผู้ช่วยศาสตราจารย์ หลักสูตรการวิเคราะห์ธุรกิจและวิทยาการข้อมูล คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์, ผู้อำนวยการศูนย์คลังปัญญาสารสนเทศ



บทนำ

การประกันภัยเป็นการบริหารความเสี่ยงรูปแบบหนึ่ง เมื่อมีความสูญเสียเกิดขึ้นจากเหตุการณ์ที่ไม่แน่นอนจึงเป็นหน้าที่ของนักคณิตศาสตร์ประกันภัยที่จะประมาณการความสูญเสียกับเหตุการณ์ที่จะเกิดในอนาคต ดังนั้นจึงเป็นสิ่งจำเป็นที่นักคณิตศาสตร์ประกันภัยต้องใช้ข้อมูลที่เป็นความถี่ (Frequency) และความรุนแรง (Severity) ของความสูญเสียที่เกิดขึ้นในอดีต เพื่อสร้างแบบจำลองการแจกแจงความสูญเสีย (Loss distribution Model) ว่าเป็นแบบจำลองในลักษณะใด หากทราบลักษณะของแบบจำลองแล้วจะสามารถสร้างแบบจำลองของข้อมูลด้วยวิธีหาตัวประมาณทางสถิติ [1]

โดยวิธีที่จะใช้ในการหาตัวประมาณของแบบจำลองนั้นมีหลายวิธี ซึ่งข้อมูลที่น่ามาสร้างแบบจำลองความสูญเสียมี 2 ส่วน คือ ความถี่ของความสูญเสียจะเป็นตัวแปรสุ่มที่ไม่ต่อเนื่อง ส่วนความรุนแรงของความสูญเสียจะเป็นตัวแปรสุ่มต่อเนื่อง โดยวิธีการสร้างแบบจำลองความสูญเสียนั้นมีสองวิธีการหลัก คือ 1) การสร้างแบบจำลองแบบพารามเมตริก 2) การสร้างแบบจำลองนอนพารามเมตริก

วิธีการที่หนึ่งเป็นการสร้างแบบจำลองความสูญเสียด้วยแบบจำลองแบบพารามเมตริก (Parametric Model) โดยทั่วไปแล้วแบบจำลองลักษณะนี้นักขึ้นอยู่กับค่าพารามิเตอร์ ได้แก่ การแจกแจงที่อยู่ในวงศ์เอกซ์โปเนนเชียล (Exponential family) ส่วนใหญ่จะหาตัวประมาณของแบบจำลองความสูญเสียโดยใช้วิธีภาวะน่าจะเป็นสูงสุด (Maximum likelihood estimation method) [2] แบบจำลองความสูญเสียที่เป็นแบบพารามิเตอร์มีข้อเสียคือต้องใช้ความชำนาญและต้องตั้งข้อสมมุติว่าข้อมูลน่าจะแจกแจงตามการแจกแจงตามทฤษฎี (Theoretical distribution) ซึ่งเป็นเรื่องยากที่จะเป็นจริงได้ และในหลายครั้งก็ไม่ได้สอดคล้องกับการแจกแจงตามทฤษฎี (Theoretical distribution) ใดๆ เลย อย่างไรก็ตามการสร้างแบบจำลองการแจกแจงความสูญเสียด้วยวิธีการพารามิเตอร์ก็ยังมิข้อดีบางประการคือใช้พารามิเตอร์น้อย มีพารามิเตอร์ที่เข้าใจได้ง่ายและนำไปใช้งานต่อในคณิตศาสตร์ประกันภัยได้ง่าย

วิธีการที่สองเป็นการสร้างแบบจำลองนอนพารามเมตริก (Non-parametric model) จะใช้การปรับเรียบเส้นโค้ง (Smoothing splines) สำหรับวิธีการปรับเรียบเส้นโค้งที่นิยมใช้มากที่สุดคือการปรับเรียบแบบคิวบิกสไปไลน์ (Cubic smoothing spline) [3] เพราะการปรับเรียบแบบคิวบิกสไปไลน์เป็นการประมาณค่าในช่วงเส้นโค้งที่ไม่มีข้อสมมุติเบื้องต้นของข้อมูลแต่จะเป็นการสร้างฟังก์ชันของตัวแปรตามกับตัวแปรอิสระ ซึ่งแบบจำลองการแจกแจงที่ได้อาจจะผ่านหรือไม่ผ่านจุดทุกจุดของข้อมูลก็ได้ แต่ต้องเป็นแบบจำลองการแจกแจงที่ต่อเนื่องและใกล้เคียงกับจุดทุกจุดมากที่สุด แต่ถ้าข้อมูลมีการกระจายแบบไม่เป็นระบบแล้วค่าที่ประมาณในช่วงก็จะมีค่าผันแปรตามข้อมูลที่กระจายซึ่งค่านี้อาจถูกให้รองรับกับแบบจำลอง ซึ่งเทคนิคที่ใช้ในการเลือกค่าพารามิเตอร์ปรับให้เรียบคือ การประมาณค่าในช่วงแบบคิวบิกสไปไลน์ (Cubic spline interpolation) วิธีการนี้มีข้อดีคือไม่ต้องสมมุติว่าข้อมูลแจกแจงแบบใดตามทฤษฎี และมีความยืดหยุ่นที่จะสร้างการแจกแจงให้สอดคล้องกับข้อมูลเชิงประจักษ์แบบใดก็ได้เสมอ แต่วิธีการนี้มีข้อเสียคือขาดพารามิเตอร์ทำให้นำไปใช้งานต่อได้ยากกว่า

ในงานชิ้นนี้ได้นำเสนอวิธีการสร้างแบบจำลองความสูญเสียด้วยแบบจำลองที่เป็นแบบจำลองกึ่งพารามเมตริก (Semiparametric Model) ทั้งนี้แบบจำลองปรับเรียบลอกลินียร์ (Log-linear smoothing models) เป็นวิธีการที่ยืดหยุ่นและไม่จำเป็นต้องสมมุติว่าข้อมูลเชิงประจักษ์นั้นมีการแจกแจงเชิงทฤษฎีเป็นเช่นไรอันเป็นจุดเด่นของแบบจำลองนอนพารามเมตริก ในขณะเดียวกันก็มีข้อดีคือยังคงมีค่าพารามิเตอร์ในแบบจำลองไม่ว่าจะเป็นโมเมนต์ศูนย์กลาง (Central moment) และโมเมนต์ข้าม (Cross-moment) ได้อย่างยืดหยุ่นอีกด้วย ทำให้การคำนวณเบี้ยในทางคณิตศาสตร์ประกันภัยสามารถนำไปประยุกต์ต่อได้ง่าย

แบบจำลองปรับเรียบลือกลิเนียร์เป็นแบบจำลองเชิงเส้นนัยทั่วไป (Generalized linear models) ที่ใช้กับข้อมูลแจกแจง ซึ่งใช้กันโดยทั่วไปสำหรับตารางแจกแจงความถี่หลายทาง (Multi-way frequency tables) และยังเป็นการแจกแจงที่อยู่ในวงศ์เอกซ์โปเนนเชียลอีกด้วย ซึ่งในการเลือกแบบจำลองปรับเรียบลือกลิเนียร์นั้นมิใช่หมายความว่าต้องนำโมเมนต์สู่ศูนย์กลางใดเข้ามาในแบบจำลอง และโมเมนต์ข้ามใดที่ควรนำมาใช้สำหรับแบบจำลองที่เหมาะสมที่สุดอีกทั้งยังมีการรักษาโมเมนต์แบบดั้งเดิมไว้ด้วย ดังนั้นการจับคู่โมเมนต์ (Moment matching) จะมีความสำคัญกับแบบจำลองปรับเรียบลือกลิเนียร์มาก [4] เพราะจะจับคู่หาโมเมนต์ไปเพื่ออธิบายลักษณะของแบบจำลองการแจกแจงและจะทำให้ได้แบบจำลองที่เหมาะสมที่สุด ในการสร้างแบบจำลองของความถี่จะเป็นการแจกแจงของตัวแปรสุ่มที่ไม่ต่อเนื่อง ส่วนความรุนแรงจะเป็นการแจกแจงของตัวแปรสุ่มที่ต่อเนื่อง แต่สำหรับแบบจำลองปรับเรียบลือกลิเนียร์ที่ใช้ในการวิจัยจะต้องทำให้เป็นค่าที่ไม่ต่อเนื่องเสียก่อน เพราะถ้ามีช่วงที่ถี่พอจะสามารถประมาณค่าการแจกแจงของตัวแปรสุ่มที่ต่อเนื่องได้

ปัจจุบันโรคภัยไข้เจ็บเป็นสิ่งที่หลีกเลี่ยงไม่ได้ เมื่อไม่สบายหรือเจ็บป่วยสิ่งที่ตามมาก็คือ ค่าใช้จ่ายสำหรับค่ารักษาพยาบาล ซึ่งค่ารักษาพยาบาลนั้นจะสูงมาก ดังนั้นเราจึงต้องทำประกันสุขภาพเพราะประกันสุขภาพจะให้ความคุ้มครองสำหรับค่าใช้จ่ายในการรักษาพยาบาลเนื่องจากการบาดเจ็บจากอุบัติเหตุหรือการเจ็บป่วยจากโรคภัย ระบบสวัสดิการรักษายาบาลข้าราชการจะเป็นระบบที่จ่ายตามค่าใช้จ่ายจริงที่มีการการเรียกเก็บ และเป็นระบบปลายเปิดที่ไม่มีการกำหนดวงเงินสูงสุดจึงเป็นปัญหาในการควบคุมค่าใช้จ่าย [5] ซึ่งในการใช้แบบจำลองการแจกแจงความสูญเสีย นั้นจะมีประโยชน์ คือ สามารถนำไปใช้คำนวณเบี้ยทางคณิตศาสตร์ประกันภัยและตั้งสำรองได้ดังนี้

โดยงานวิจัยที่ศึกษาเกี่ยวกับสวัสดิการรักษายาบาลข้าราชการไทย พบว่า มีงานวิจัยที่ได้จัดทำไว้มากมาย เช่น การศึกษาเรื่องค่าใช้จ่ายที่พยากรณ์จากการใช้ยาราคาแพงของสวัสดิการรักษายาบาลข้าราชการเปรียบเทียบกับหลักประกันสุขภาพถ้วนหน้า [6] โดยใช้วิธีการวิเคราะห์ห้อนุกรมเวลา พบว่า งานวิจัยมีข้อจำกัดที่สำคัญคือแบบจำลองวิเคราะห์ทางสถิติที่ใช้ในการพยากรณ์ค่าใช้จ่ายยังไม่ได้ควบคุมความผันแปรระหว่างช่วงเวลาในปัจจัยด้านอื่นๆที่มีผลต่อโอกาสการใช้ยาเป้าหมาย

การศึกษาเรื่องการวิเคราะห์ปัจจัยที่มีผลต่อค่าใช้จ่ายผู้ป่วยนอกของระบบสวัสดิการรักษายาบาลของข้าราชการ กรณีศึกษาของโรงพยาบาลสระบุรี [7] พบว่า ผลกระทบของโครงการจ่ายตรงโรคเรื้อรัง 4 โรคที่ได้จากวิธีการเปรียบเทียบอย่างง่ายและวิธี Differences-in-Differences แสดงให้เห็นว่า การเปรียบเทียบอย่างง่ายไม่มีความแม่นยำมากเพียงพอที่จะนำมาใช้ประเมินผลกระทบของโครงการจ่ายตรงได้ และการนำผลที่ได้จากการเปรียบเทียบอย่างง่ายไปใช้ควรทำด้วยความระมัดระวังอย่างมาก

จากตัวอย่างงานวิจัยข้างต้นจะเห็นว่ายังไม่มียานวิจัยที่นำแบบจำลองปรับเรียบลือกลิเนียร์มาใช้กับระบบสวัสดิการรักษายาบาลข้าราชการเลย ดังนั้นวัตถุประสงค์การวิจัยในครั้งนี้คือนำแบบจำลองปรับเรียบลือกลิเนียร์มาประยุกต์ใช้กับแบบจำลองการแจกแจงความสูญเสียโดยใช้ข้อมูลการจ่ายค่ารักษาพยาบาลด้วยระบบสวัสดิการรักษายาบาลข้าราชการ ประเภทผู้ป่วยนอก จากกรมบัญชีกลาง ซึ่งการใช้แบบจำลองปรับเรียบลือกลิเนียร์นั้นจะมีประโยชน์มากในการนำมาวิเคราะห์เนื่องจากมีความยืดหยุ่นมากพอที่จะทำให้มีความเหมาะสมกับแบบจำลองปรับเรียบหนึ่งตัวแปร (Univariate Smoothing Model)



วัตถุประสงค์การวิจัย

สร้างแบบจำลองปรับเรียบลือกลิเนียร์เพื่อพยากรณ์ความถี่และความรุนแรงของความสูญเสียของสวัสดิการค่ารักษาพยาบาลข้าราชการไทยและสมาชิกในครอบครัว ประเภทผู้ป่วยนอก ในหมวดค่ารักษาพยาบาลรวม ปี พ.ศ. 2558

วิธีการวิจัย

กระบวนการดำเนินงานวิจัยครั้งนี้ได้นำแบบจำลองปรับเรียบลือกลิเนียร์ที่เป็นแบบจำลองปรับเรียบหนึ่งตัวแปรมาประยุกต์ใช้กับแบบจำลองความสูญเสีย

1. **แบบจำลองความสูญเสีย** เป็นแบบจำลองทางประกันภัยส่วนใหญ่จะเป็นแบบจำลองความสูญเสียที่มีการแจกแจงที่อยู่ในลักษณะเบ้ขวา ในการสร้างแบบจำลองความสูญเสียจะเป็นการผสมของแบบจำลองระหว่างการแบบจำลองของความถี่ (Frequency) และแบบจำลองของความรุนแรง (Severity) จึงเขียนสมการได้ดังนี้

$$\text{Loss} = \text{Frequency} \times \text{Severity}$$

ซึ่งแบบจำลองของความถี่จะเป็นการแจกแจงของตัวแปรสุ่มที่ไม่ต่อเนื่อง ซึ่งส่วนใหญ่จะเป็นการแจกแจงทวินามลบ (Negative binomial distribution) ส่วนความรุนแรงจะเป็นการแจกแจงของตัวแปรสุ่มที่ต่อเนื่องแต่บางครั้งก็เป็นการแจกแจงของตัวแปรสุ่มที่ไม่ต่อเนื่องได้เช่นกัน ซึ่งส่วนใหญ่จะเป็นการแจกแจงล็อกนอร์มอล (Lognormal distribution) [1]

2. **แบบจำลองปรับเรียบลือกลิเนียร์** เป็นแบบจำลองเชิงเส้นน้อยทั่วไปที่ใช้กับข้อมูลจำนวนนับซึ่งใช้กันโดยทั่วไปสำหรับตารางแจกแจงความถี่หลายทาง ซึ่งสามารถอธิบายได้ในรูปของโมเมนต์ตัวอย่าง (Sample moment) [8] ได้ใช้แบบจำลองการปรับเรียบลือกลิเนียร์กับการแจกแจงหนึ่งตัวแปร ซึ่งในการเลือกแต่ละแบบจำลองของแบบจำลองการปรับเรียบลือกลิเนียร์นั้นจะมีเป้าหมายว่าโมเมนต์ใดที่ควรนำมาใช้สำหรับแบบจำลองที่มีความเหมาะสมที่สุด และอีกทั้งในแต่ละแบบจำลองจะต้องมีโมเมนต์ที่ 1 - 4 (ซึ่งประกอบไปด้วย ค่าเฉลี่ย, ความแปรปรวน, ความเบ้ และความโค้ง)

แบบจำลองปรับเรียบหนึ่งตัวแปร (Univariate Smoothing Model)

ในการสร้างแบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปร สามารถเขียนแบบจำลองได้ 2 แบบจำลอง คือ แบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรของความรุนแรงของความสูญเสีย และแบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรของความถี่ของความสูญเสีย

แบบจำลองที่ 1 แบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรของความรุนแรงของความสูญเสีย เขียนได้ดังนี้

ให้ $i = 1, \dots, I$ คือ อนุกรมของโมเมนต์ของคะแนนของความรุนแรงของความสูญเสีย
 $j = 1, \dots, J$ คือ หมายเลขของหมวดหมู่สำหรับคะแนนของความรุนแรงของความสูญเสีย

$$\log(p_j) = \alpha + \sum_{i=1}^I \beta_i (x_j)^i \quad (1)$$

โดยที่ p_j คือ ค่าสัดส่วนที่พยากรณ์ได้จากแบบจำลองของความรุนแรงของความสูญเสียของหมวดหมู่ที่ j
 α คือ ค่าคงที่
 β_i คือ สัมประสิทธิ์ของโมเมนต์อันดับอนุกรม i
 $(x_j)^i$ คือ ฟังก์ชันของคะแนนของความรุนแรงของความสูญเสียของหมวดหมู่ที่ j โดยที่ $j = 1, \dots, J$

แบบจำลองที่ 2 แบบจำลองปรับเรียบลอกลินีร์หนึ่งตัวแปรของความถี่ของความสูญเสีย เขียนได้ดังนี้

ให้ $k = 1, \dots, K$ คือ องศาพหุนามของโมเมนต์ของคะแนนของความถี่ของความสูญเสีย

$l = 1, \dots, L$ คือ หมายเลขของหมวดหมู่สำหรับคะแนนของความถี่ของความสูญเสีย

$$\log(p_l) = \alpha + \sum_{k=1}^K \beta_k (y_l)^k \quad (2)$$

โดยที่ p_l คือ ค่าสัดส่วนที่พยากรณ์ได้จากแบบจำลองของความถี่ของความสูญเสียของหมวดหมู่ที่ l

α คือ ค่าคงที่

β_k คือ สัมประสิทธิ์ของโมเมนต์อันดับพหุนามที่ k

$(y_l)^k$ คือ ฟังก์ชันของคะแนนของความถี่ของความสูญเสียของหมวดหมู่ที่ l โดยที่ $k = 1, \dots, K$

3. วิธีการวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ได้นำแบบจำลองปรับเรียบลอกลินีร์ที่เป็นแบบจำลองปรับเรียบหนึ่งตัวแปรมาประยุกต์ใช้กับแบบจำลองความสูญเสีย ในการประมวลผลข้อมูลโดยเลือกใช้โปรแกรม Statistical Analysis System (SAS) เนื่องจากข้อมูลในการจ่ายค่ารักษาพยาบาลด้วยระบบสวัสดิการรักษายาบาลข้าราชการและสมาชิกในครอบครัวประเภทผู้ป่วยนอกของปี 2558 ที่ได้รับมาจากกรมบัญชีกลางเป็นข้อมูลขนาดใหญ่ [9] ได้ใช้ SAS PROC GENMOD ซึ่งมีสมมติฐานดำเนินการเหมือนกับการประมาณโดยวิธีภาวะน่าจะเป็นสูงสุด (Maximum likelihood estimation method)

ในการนำค่าของฟังก์ชันของคะแนนของความถี่และฟังก์ชันของคะแนนของความรุนแรงของความสูญเสียของแต่ละโมเมนต์มาเข้าแบบจำลอง จะมีการปรับให้เป็นคะแนนมาตรฐาน (Z-score) เนื่องจากค่าของแต่ละโมเมนต์มีค่าสูงมากและจะทำให้การเข้าสู่ของการวิเคราะห์เชิงตัวเลข (Numerical analysis) ซึ่งใช้วิธีการนิวตัน-ราฟสัน (Newton-Raphson) เกิดปัญหา จึงต้องแก้ปัญหาดังกล่าวด้วยการปรับให้เป็นคะแนนมาตรฐานเสียก่อน โดยมีตัวแปรที่ใช้ในงานวิจัย ได้แก่ ความถี่และความรุนแรง ซึ่งจากข้อมูลที่ได้รับมานั้นเป็นข้อมูลทุกมิติที่มีขนาดใหญ่ จึงเห็นสมควรที่จะแบ่งข้อมูลเป็นช่วงเพื่อง่ายต่อการวิเคราะห์ข้อมูล ทั้งนี้เนื่องจากความรุนแรงของความเสียหายเป็นตัวแปรต่อเนื่องจึงต้องทำให้เป็นค่าที่ไม่ต่อเนื่องก่อน และใช้ค่ากลางของแต่ละช่วงเป็นคะแนนของความรุนแรงในแบบจำลองปรับเรียบลอกลินีร์

3.1 การระบุแบบจำลองปรับเรียบลอกลินีร์ (Log-linear Smoothing Models Specification)

การระบุแบบจำลองปรับเรียบลอกลินีร์หนึ่งตัวแปร แสดงในตาราง 1 ทั้งสิ้น 10 แบบจำลอง โดยให้แบบจำลองที่ง่ายที่สุดยังคงจับคู่ (Matching moments) ไปจนถึงจำนวนโมเมนต์ที่ 4 และแบบจำลองที่ซับซ้อนที่สุดจะไปถึงจำนวนโมเมนต์ที่ 13 และรักษาลำดับชั้นของโมเมนต์ โดยจะมีโมเมนต์ที่สูงกว่าได้ต้องมีโมเมนต์ที่ต่ำกว่าในแบบจำลองแล้วเท่านั้น ดังตารางที่ 1

3.2 การคัดเลือกแบบจำลองปรับเรียบลอกลินีร์ มี 2 เกณฑ์ คือ ค่าสถิติดีไวเียนซ์ (G^2) และเกณฑ์

สารสนเทศของ Akaike (AIC)

ค่าสถิติดีไวเียนซ์ (Deviance Statistic หรือ G^2) [9] เป็นสถิติทดสอบอัตราส่วนภาวะน่าจะเป็น (Likelihood ratio test) หรือเรียกอีกอย่างหนึ่งว่า สถิติทดสอบไคสแควร์อัตราส่วนภาวะน่าจะเป็น (Likelihood ratio chi-squared statistic) ซึ่งเป็นสถิติที่ใช้ในการบ่งบอกถึงความเหมาะสมของแบบจำลอง

$$G^2 = 2 \sum_{j=1}^J \left[n_j \log_e \left(\frac{n_j}{\hat{p}_j N} \right) - (n_j - \hat{p}_j N) \right]$$

โดยที่ N คือ จำนวนทั้งหมด



n_j คือ ขนาดตัวอย่างของหมวดหมู่ที่ j โดยที่ $j = 1, \dots, J$

p_j คือ ค่าสัดส่วนที่พยากรณ์ได้จากแบบจำลองของหมวดหมู่ที่ j โดยที่ $j = 1, \dots, J$

ในการคัดเลือกแบบจำลองที่ดีจะพิจารณาจากแบบจำลองที่มีค่า G^2 น้อยที่สุด

เกณฑ์สารสนเทศของ Akaike (Akaike's Information Criterion : **AIC**) Akaike [10] ได้คิดค้นและนำเสนอเกณฑ์สารสนเทศนี้ โดยมีสมการดังนี้

$$AIC = n_j \log_e \left(\frac{n_j}{p_j} \right) + 2k$$

โดยที่ k คือ ผลรวมของจำนวนพารามิเตอร์

n_j คือ ขนาดตัวอย่างของหมวดหมู่ที่ j โดยที่ $j = 1, \dots, J$

p_j คือ ค่าสัดส่วนที่พยากรณ์ได้จากแบบจำลองของหมวดหมู่ที่ j โดยที่ $j = 1, \dots, J$

ในการคัดเลือกแบบจำลองที่ดีจะพิจารณาจากแบบจำลองที่มีค่า **AIC** น้อยที่สุด

สำหรับเกณฑ์ในการคัดเลือกแบบจำลองปรับเรียบลือกลิเนียนั้น เมื่อใช้เกณฑ์ในการคัดเลือกที่ไม่เหมือนกันแบบจำลองที่ได้ก็จะไม่เหมือนกัน [11]

ผลการวิจัย

ผลการสร้างแบบจำลองปรับเรียบลือกลิเนียน์หนึ่งตัวแปรของความถี่และความรุนแรงของความสูญเสียประเภทผู้ป่วยนอก ในหมวดคำรักษาพยาบาลรวมของปี พ.ศ. 2558 มีดังนี้

1. ผลการสร้างแบบจำลองปรับเรียบลือกลิเนียน์หนึ่งตัวแปรของความถี่ของความสูญเสีย

ผลการคัดเลือกแบบจำลองปรับเรียบลือกลิเนียน์หนึ่งตัวแปรจากค่า G^2 และ **AIC** โดยเลือกค่าที่ต่ำที่สุด จากตารางที่ 2 พบว่า หมวดคำรักษาพยาบาลรวมของปีงบประมาณ พ.ศ. 2558 ได้เลือกแบบจำลองที่ไม่มีความซับซ้อนมาก คือ แบบจำลองที่ 2 ดังนั้นแบบจำลองที่ใช้พยากรณ์ความถี่ของความสูญเสีย เขียนได้ดังนี้

$$\log(p_l) = 1.86 - 25.41(y_l)^1 + 214.90(y_l)^2 - 1,149.43(y_l)^3 + 2,167.36(y_l)^4 - 1,212.08(y_l)^5$$

นำผลการปรับเรียบจากแบบจำลองที่เลือกมาข้างต้นมาเปรียบเทียบกับข้อมูลที่ยังไม่ได้ปรับเรียบของปีงบประมาณ พ.ศ. 2558 โดยผลการเปรียบเทียบแสดงดังภาพที่ 1 - 2

เมื่อเปรียบเทียบผลการปรับเรียบจากแบบจำลองที่เลือกกับข้อมูลที่ยังไม่ได้ปรับเรียบดังภาพที่ 1 - 2 พบว่า ข้อมูลปรับเรียบมีความใกล้เคียงกับข้อมูลจริง แต่มีการแจกแจงความถี่ที่ชัดเจนกว่าข้อมูลเดิม จึงสามารถนำข้อมูลที่ปรับเรียบแล้วไปใช้ในการวิเคราะห์ในลำดับถัดไปได้

2. ผลการสร้างแบบจำลองปรับเรียบลือกลิเนียน์หนึ่งตัวแปรของความรุนแรงของความสูญเสีย

ผลการคัดเลือกแบบจำลองปรับเรียบลือกลิเนียน์หนึ่งตัวแปรจากค่า G^2 และ **AIC** โดยเลือกค่าที่ต่ำที่สุด จากตารางที่ 3 พบว่า หมวดคำรักษาพยาบาล รวมของปีงบประมาณพ.ศ. 2558 ได้เลือกแบบจำลองที่ไม่มีความซับซ้อนมาก คือ แบบจำลองที่ 1 ดังนั้นแบบจำลองที่ใช้พยากรณ์ความรุนแรงของความสูญเสีย เขียนได้ดังนี้

$$\log(p_j) = 22.62 - 39.78(x_j)^1 + 117.77(x_j)^2$$

นำผลการปรับเรียบจากแบบจำลองที่เลือกมาข้างต้นมาเปรียบเทียบกับข้อมูลที่ยังไม่ได้ปรับเรียบของปีงบประมาณพ.ศ. 2558 โดยผลการเปรียบเทียบแสดงดังภาพที่ 3 - 4

เมื่อเปรียบเทียบผลการปรับเรียบจากแบบจำลองที่เลือกกับข้อมูลที่ยังไม่ได้ปรับเรียบดังภาพที่ 3 - 4 พบว่าข้อมูลปรับเรียบมีความใกล้เคียงกับข้อมูลจริง แต่มีการแจกแจงความถี่ที่ชัดเจนกว่าข้อมูลเดิม จึงสามารถนำข้อมูลที่ปรับเรียบแล้วไปใช้ในการวิเคราะห์ในลำดับถัดไปได้

จะเห็นได้ว่า จากกราฟของความถี่และความรุนแรงของความสูญเสียที่ได้จากค่าปรับเรียบ สังเกตได้ว่ามีลักษณะเบ้ขวา

สรุปและอภิปรายผลการวิจัย

ในหัวข้อนี้จะแสดงผลสรุป และอภิปรายผล จากผลการวิเคราะห์ดังกล่าวในหัวข้อที่ผ่านมา แยกเป็นหัวข้อ ดังนี้

1. สรุปผลการสร้างแบบจำลองการปรับเรียบลือกลิเนียร์หนึ่งตัวแปรของความถี่และความรุนแรงของความสูญเสีย ประเภทผู้ป่วยนอก

1.1 ในการสร้างแบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรของความถี่ของความสูญเสีย พบว่า ค่า G^2 และ AIC มีแนวโน้มที่จะให้ผลไปในทิศทางเดียวกันที่จะเลือกแบบจำลองที่ไม่มีความซับซ้อนมาก คือ แบบจำลองที่ 2 ซึ่งประกอบด้วย โมเมนต์ที่ 1 - 5 และเห็นได้ว่าการสร้างแบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรมีความยืดหยุ่นที่มากพอกับข้อมูลของความถี่ของความสูญเสีย

1.2 ในการสร้างแบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรของความรุนแรงของความสูญเสีย ประเภทผู้ป่วยนอก หมวดคำรักษาพยาบาลรวมของ ปี พ.ศ. 2558 พบว่า ค่า G^2 และ AIC มีแนวโน้มที่จะให้ผลไปในทิศทางเดียวกันที่จะเลือกแบบจำลองที่ไม่มีความซับซ้อนมาก คือ แบบจำลองที่ 1 ซึ่งประกอบด้วย โมเมนต์ที่ 1 - 4 และเห็นได้ว่าการสร้างแบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรมีความยืดหยุ่นที่มากพอกับข้อมูลของความรุนแรงของความสูญเสีย

1.3 จากภาพที่ 1 - 4 จะเห็นว่า กราฟมีลักษณะเบ้ขวาแต่มีลักษณะที่เบ้มากปลายหางค่อนข้างยาว เพราะจำนวนครั้งของความถี่และความรุนแรงในการเบิกจ่ายยิ่งมากผู้มีส่วนเกี่ยวข้องในการเบิกจ่ายจะยิ่งน้อย เมื่อพิจารณาการเปรียบเทียบการแจกแจงความถี่และความรุนแรงของค่าจริงและค่าปรับเรียบของการเลือกแบบจำลองด้วยค่า G^2 และ AIC จะเห็นได้ว่ากราฟมีลักษณะไปในทิศทางเดียวกันและสอดคล้องกัน แต่ค่าปรับเรียบจะมีความยืดหยุ่นและราบเรียบกว่าค่าจริง ดังนั้นจึงเป็นไปตามทฤษฎีที่ได้กล่าวไว้ข้างต้น

2. อภิปรายผลการสร้างแบบจำลองการปรับเรียบลือกลิเนียร์หนึ่งตัวแปรของความถี่และความรุนแรงของความสูญเสีย ประเภทผู้ป่วยนอก

ในการสร้างแบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรของความถี่และความรุนแรงของความสูญเสีย หมวดคำรักษาพยาบาลรวม พบว่า แบบจำลองความถี่และความรุนแรงของความสูญเสียมีแนวโน้มจะไปอยู่ที่แบบจำลองที่ไม่มีความซับซ้อนมาก ซึ่งแบบจำลองไม่เกิดกับโมเมนต์สูงๆ จากภาพที่ 1 - 4 จะเห็นว่ากราฟจะมีความเบ้มากและโค้งมาก แสดงว่าการแจกแจงค่อนข้างผิดปกติ แต่ก็เพียงพอต่อการสร้างแบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรของความถี่และความรุนแรง ซึ่งแบบจำลองมีโมเมนต์อันดับสูงที่สุดเพียงโมเมนต์ที่ 5 และโมเมนต์ที่ 4 ตามลำดับ สำหรับความรุนแรงของความสูญเสีย สามารถอธิบายได้ว่าหากไม่มีความรุนแรงมากนักผู้ป่วยนอกก็จะไม่มีความต้องการเข้ารับการรักษาที่โรงพยาบาล และจากการสร้างแบบจำลองปรับเรียบลือกลิเนียร์ในการพยากรณ์ความถี่และความรุนแรงของความสูญเสียสามารถนำไปพยากรณ์สวัสดิการคำรักษาพยาบาลข้าราชการไทย ในหมวดคำรักษาพยาบาลรวมในอนาคตได้ จากทฤษฎีข้างต้นในการนำแบบจำลองปรับเรียบลือกลิเนียร์หนึ่งตัวแปรมาประยุกต์ใช้กับแบบจำลองความสูญเสียสามารถนำมาใช้ได้ดีทั้งกับข้อมูลที่เป็นความถี่และความรุนแรง และสามารถนำแบบจำลองปรับเรียบลือกลิเนียร์มา



คำนวณเบี่ยงทางคณิตศาสตร์ประกันภัยและเอาไปตั้งสำรองได้ แต่งานวิจัยนี้มีข้อจำกัดที่สำคัญคือ การแบ่งค่าต่อเนื่องออกเป็นช่วงย่อยๆ ในกรณีของข้อมูลของความรุนแรงอาจจะสูญเสียความต่อเนื่องของข้อมูลได้

ถ้าเป็นแบบจำลองปรับเรียบลือกลิเนียร์สองตัวแปรและแบบจำลองปรับเรียบลือกลิเนียร์ตัวแปรพหุ อาจจะมีรูปแบบที่คล้ายเดิมหรือต่างออกไป ทั้งเรื่องของความถี่และความรุนแรงของความสูญเสีย แต่ถ้าหากพิจารณารวมกัน อาจจะได้เห็นความสัมพันธ์ระหว่างความถี่และความรุนแรงของความสูญเสียได้

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับข้อมูลของสวัสดิการรักษายาบาลข้าราชการไทยมาจาก “กรมบัญชีกลาง”

เอกสารอ้างอิง

1. Klugman SA, Panjer HH, Willmot GE. Loss Models from Data to Decisions. 4th ed. The United State of America: John Wiley & Sons; 2012.
2. Holland PW, Thayer DT. Notes on the use of log-linear models for fitting discrete probability distributions. ETS. 1987.
3. Moses T, Davier AAv, Casabianca J. Loglinear Smoothing : An Alternative Numerical Approach Using SAS. ETS. 2004.
4. Nikomborirak D, Ruenthip K. Study on organizational reform guidelines to improve the efficiency of the management of the welfare system of government officials. 2011. Thai.
5. Moses T, Holland PW. Selection strategies for univariate loglinear smoothing models and their effect on equating function accuracy. Journal of Educational Measurement. 2009; 46(2): p. 159-176.
6. Limwattananon S, Limwattananon C, Cheawchanwattana A, Silkavute P, Tangcharoensatien V. Forecasted expenditure due to the use of expensive drugs in Civil Servant Medical Benefit Scheme. Journal of Health Systems Research. 2011; 5(2). Thai.
7. Panpiemras J, Wasi N, Manachotphon W. Decomposing health expenditure : The case of CSMBS outpatients in Saraburi hospital. Thailand Development Research Institute; 2013. Thai.
8. Suwatthi P. Statistical inference theory Bangkok, Thailand: Project development and technical documentation. NIDA; 2010. Thai.
9. Holland PW, Thayer DT. Univariate and bivariate loglinear models for discrete test score distribution. ETS. 2000.
10. Akaike H. Likelihood of a model and information criteria. Journal of Econometrics. 1981; 16(1): p. 3-14.
11. Pollock DSG. Smoothing with Cubic Splines. Queen Mary and Westfield College. 1999.

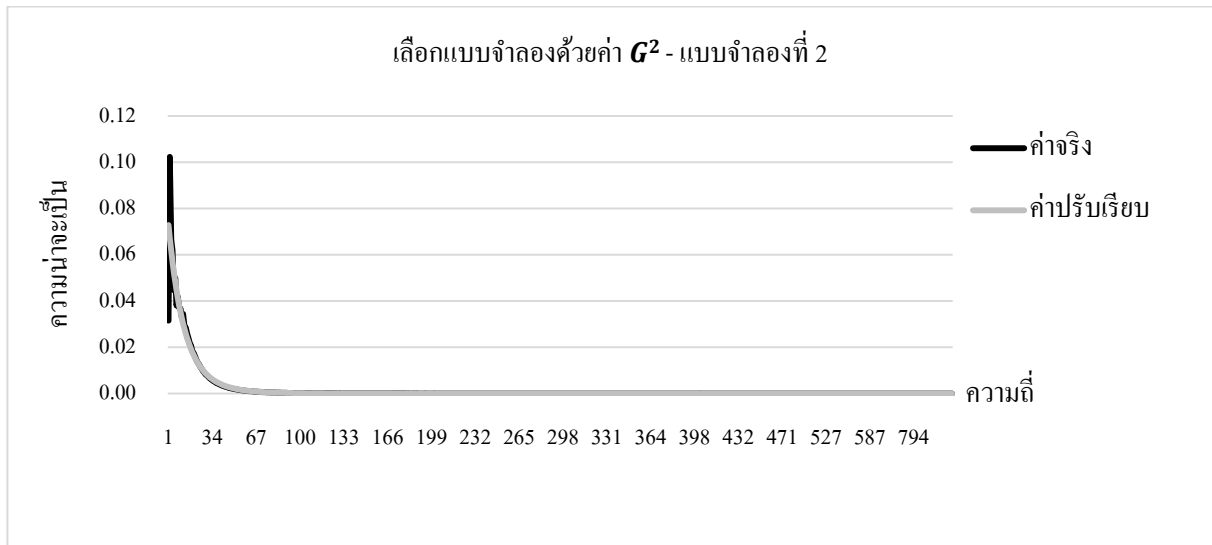
ตารางที่ 1 การระบุแบบจำลองปรับเรียบล้อยกเนียร์หนึ่งตัวแปร

แบบจำลอง	โมเมนต์ลำดับที่												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1	✓	✓	✓	✓									
2	✓	✓	✓	✓	✓								
3	✓	✓	✓	✓	✓	✓							
4	✓	✓	✓	✓	✓	✓	✓						
5	✓	✓	✓	✓	✓	✓	✓	✓					
6	✓	✓	✓	✓	✓	✓	✓	✓	✓				
7	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
8	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
9	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
10	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

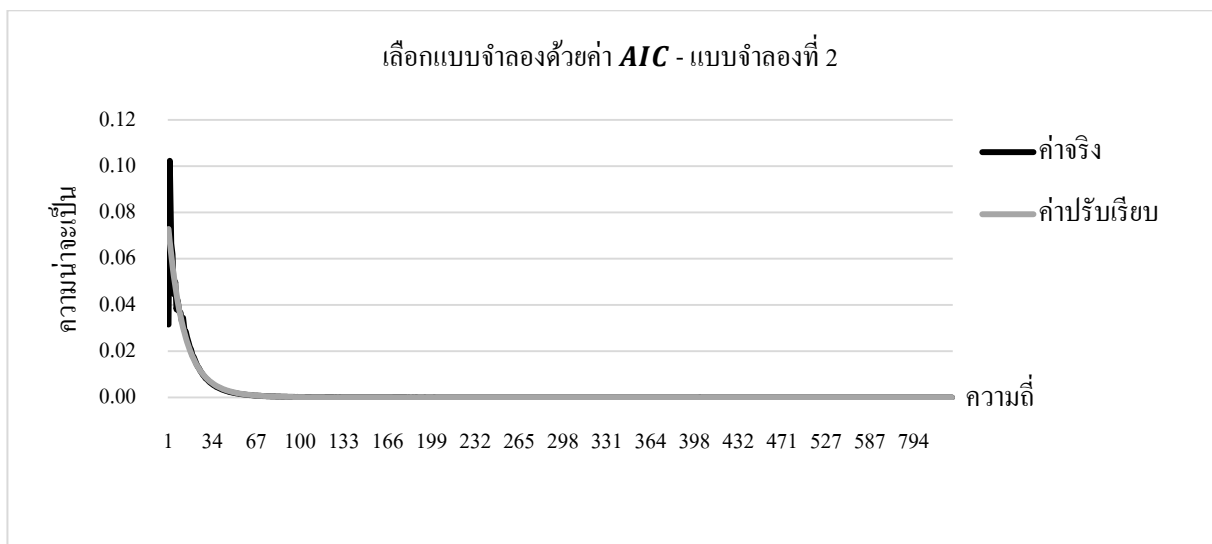
หมายเหตุ : ✓ คือ สัญลักษณ์ของการระบุโมเมนต์ในแต่ละแบบจำลอง

ตารางที่ 2 ผลการคัดเลือกแบบจำลองปรับเรียบล้อยกเนียร์หนึ่งตัวแปรจากค่า G^2 และ AIC ของความถี่ของความสูญเสีย ในหมวดค่ารักษาพยาบาลรวม ปี พ.ศ. 2558

แบบจำลอง	โมเมนต์													เกณฑ์การคัดเลือกแบบจำลอง	
	y^1	y^2	y^3	y^4	y^5	y^6	y^7	y^8	y^9	y^{10}	y^{11}	y^{12}	y^{13}	G^2	AIC
1	✓	✓	✓	✓										6.2×10^7	6.2×10^7
2	✓	✓	✓	✓	✓									1.8×10^5	1.8×10^5
3	✓	✓	✓	✓	✓	✓								2×10^5	2×10^5
4	✓	✓	✓	✓	✓	✓	✓							2×10^5	2×10^5
5	✓	✓	✓	✓	✓	✓	✓	✓						2×10^5	2×10^5
6	✓	✓	✓	✓	✓	✓	✓	✓	✓					2×10^5	2×10^5
7	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				2×10^5	2×10^5
8	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			2×10^5	2×10^5
9	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		2×10^5	2×10^5
10	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	2×10^5	2×10^5



ภาพที่ 1 เปรียบเทียบการแจกแจงความถี่ของค่าจริงและค่าปรับเรียบของการเลือกแบบจำลองด้วยค่า G^2 สำหรับความถี่ของความสูญเสียประเภทผู้ป่วยนอก ปี พ.ศ. 2558 ค่ารักษาพยาบาลรวม



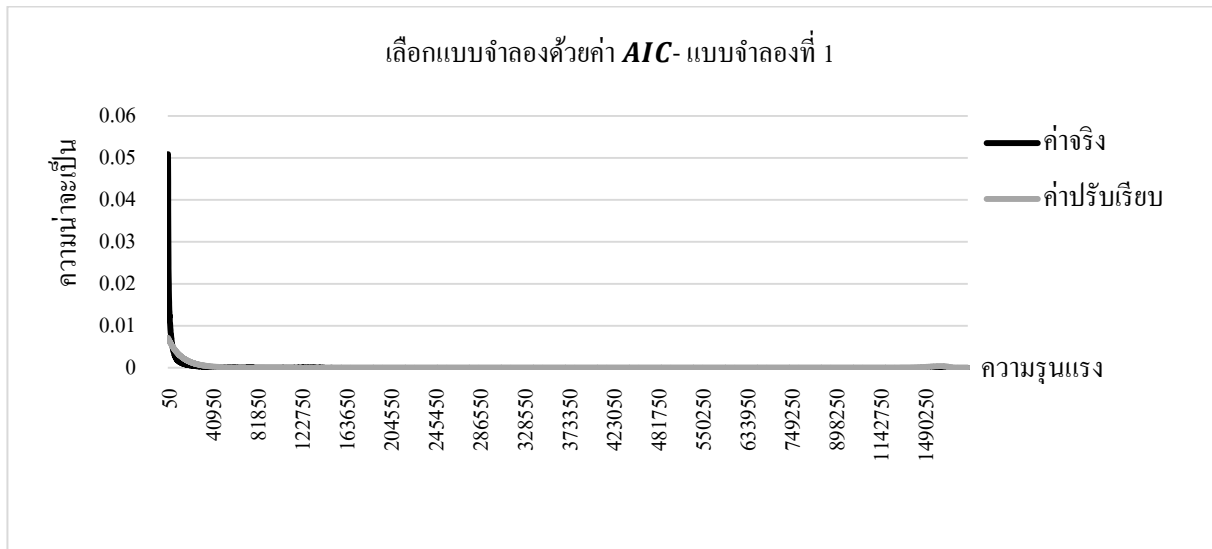
ภาพที่ 2 เปรียบเทียบการแจกแจงความถี่ของค่าจริงและค่าปรับเรียบของการเลือกแบบจำลองด้วยค่า AIC สำหรับความถี่ของความสูญเสียประเภทผู้ป่วยนอก ปี พ.ศ. 2558 ค่ารักษาพยาบาลรวม

ตารางที่ 3 ผลการคัดเลือกแบบจำลองปรับเรียบสีกลิเนียร์หนึ่งตัวแปรจากค่า G^2 และ AIC ของความรุนแรงของความเสี่ยง ในหมวดค่ารักษาพยาบาลรวม ปี พ.ศ. 2558

แบบจำลอง	โมเมนต์													เกณฑ์การคัดเลือกแบบจำลอง	
	x^1	x^2	x^3	x^4	x^5	x^6	x^7	x^8	x^9	x^{10}	x^{11}	x^{12}	x^{13}	G^2	AIC
1	✓	✓	✓	✓										5×10^{22}	5×10^{22}
2	✓	✓	✓	✓	✓									2.9×10^{52}	2.9×10^{52}
3	✓	✓	✓	✓	✓	✓								2.7×10^{49}	2.7×10^{49}
4	✓	✓	✓	✓	✓	✓	✓							2.7×10^{49}	2.7×10^{49}
5	✓	✓	✓	✓	✓	✓	✓	✓						3.4×10^{37}	3.4×10^{37}
6	✓	✓	✓	✓	✓	✓	✓	✓	✓					3.4×10^{37}	3.4×10^{37}
7	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				3.4×10^{37}	3.4×10^{37}
8	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			3.4×10^{37}	3.4×10^{37}
9	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		3.4×10^{37}	3.4×10^{37}
10	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	3.4×10^{37}	3.4×10^{37}



ภาพที่ 3 เปรียบเทียบการแจกแจงความรุนแรงของค่าจริงและค่าปรับเรียบของการเลือกแบบจำลองด้วยค่า G^2 สำหรับความรุนแรงของความเสี่ยงประเภทผู้ป่วยนอก ปี พ.ศ. 2558 ค่ารักษาพยาบาลรวม



ภาพที่ 4 เปรียบเทียบการแจกแจงความรุนแรงของค่าจริงและค่าปรับเรียบของการเลือกแบบจำลองด้วยค่า AIC

สำหรับความรุนแรงของความสูญเสียประเภทผู้ป่วยนอก ปี พ.ศ. 2558 ค่ารักษาพยาบาลรวม