การเปรียบเทียบตัวแบบอัตรามรณะเพื่อพยากรณ์อัตรามรณะไทย

A Comparison of Mortality Models for Forecasting Thai Mortality Rate

Received : April 1, 2021
Revised : May 28, 2021
Accepted : June 2, 2021

ณัฐสุรางค์ ยะสูงเนิน, วท.ด. (Natthasurang Yasungnoen, D.Sc.)[1*]

บทคัดย่อ

วัตถุประสงค์: งานวิจัยนี้มีวัตถุประสงค์ เพื่อทำการศึกษาตัวแบบอัตรามรณะ และเพื่อคัดเลือกตัวแบบอัตรามรณะ ที่เหมาะสมในการพยากรณ์อัตรามรณะของประชากรไทย ในช่วงปี 2563-2572

วิธีการวิจัย: ตัวแบบที่ทำการศึกษาคือ ตัวแบบลี-คาร์เตอร์ (Lee-Carter Model) และตัวแบบเอช-โคฮอท (Age-Cohort Model) ภายใต้สมมติฐาน จำนวนการตายมีการแจกแจงปัวซง และจำนวนการตายมีการแจกแจงทวินามลบ ข้อมูลที่ใช้ประกอบด้วย จำนวนประชากรกลางปี และจำนวนการตายของประชากรจำแนกตามเพศ และช่วงอายุ มาจาก กระทรวงมหาดไทย และจากสำนักนโยบายและยุทธศาสตร์ กระทรวงสาธารณสุข ตามลำดับ การวิเคราะห์ข้อมูลโดยทำการประมาณค่าอัตรามรณะ เพื่อหาตัวแบบที่เหมาะสม เพื่อที่ใช้ในการพยากรณ์อัตรามรณะ นอกจากนี้การวิเคราะห์ได้พิจารณาหาวิธีการประมาณค่าที่เหมาะสมกับข้อมูลโดยพิจารณาการเกิดปัญหา overdispersion ของข้อมูลจำนวนการตาย อีกด้วย

ผลการวิจัย: พบว่า สำหรับเพศชาย ตัวแบบที่เหมาะสมในการนำไปพยากรณ์ค่าอัตรามรณะ คือ ตัวแบบลี-คาร์เตอร์ ภายใต้สมมติฐานจำนวนการตายมีการแจกแจงทวินามลบ กรณีพารามิเตอร์การกระจายขึ้นอยู่กับอายุ และสำหรับเพศหญิง ได้ตัวแบบที่เหมาะสมคือ ตัวแบบเอช-โคฮอท ภายใต้สมมติฐานจำนวนการตายมีการแจกแจงทวินามลบ กรณีพารามิเตอร์การกระจายไม่ขึ้นอยู่กับอายุ ผลการพยากรณ์อัตรามรณะพบว่าอัตรามรณะมีแนวโน้มลดลงทั้งเพศชายและเพศหญิง และอายุขัยโดยเฉลี่ยมีแนวโน้มเพิ่มขึ้นทั้งเพศชายและหญิง นอกจากนี้พบว่าอายุขัยเฉลี่ยของเพศหญิง สูงกว่าอายุขัยเฉลี่ยของเพศชาย

คำสำคัญ: ตัวแบบลี-คาร์เตอร์, ตัวแบบเอช-โคฮอท, อัตรามรณะไทย

[1]อาจารย์ประจำสาขาวิทยาการคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยวงษ์ชวลิตกุล จังหวัดนครราชสีมา
(Lecturer, Computer Science Branch, Faculty of Engineering, Vongchavalitkul University, Nakhon Ratchasima)
[*]ผู้เขียนหลัก (Corresponding author)
E-mail: natthasurang_yas@vu.ac.th

Abstract

Objective: The objectives of this research were to study morality model and comparison mortality model that appropriate to forecast Thai mortality rate from 2563-2572.

Methods: The mortality model was composed of the Lee-Carter model and the Age-Cohort model. The estimation of the parameter was considered under the hypothesis of a distribution of the death count as a Poisson distribution and a Negative Binomial distribution. The main source of data on the number of age-specific people in the population came from

the Thai Ministry of the Interior. The death counts were obtained from the Bureau of Policy and Strategy, Ministry of Public Health. Data analysis was performed by using mortality rate estimation for the suitable model to forecast mortality rate. Furthermore, we also considered the method for finding suitable model with the problem of overdispersion of the death counts.

Results: The result showed the suitable model for male was the Lee-Carter model under assumption of the death count as negative binomial distribution with age-dependent dispersion parameter. For female, the suitable model was the Age-Cohort model under assumption of the death count as negative binomial distribution with age-independent dispersion parameter. Furthermore, the forecasted mortality rate both genders trended to decline, span from 2563-2572. The life expectancy trended to increase both male and female. Moreover, the life expectancy of female was almost greater than the male life expectancy.

Keywords: Lee-Carter Model, Age-Cohort Model, Thai Mortality Rate

1. ความเป็นมาและความสำคัญของปัญหา

การจัดการความเสี่ยง (Risk Management) ในธุรกิจประกันชีวิต เป็นการจัดการความเสี่ยง เกี่ยวกับการใช้ความคุ้มครองต่อการเสียชีวิตของผู้เอาประกันภัย โดยกลุ่มผู้เอาประกันภัยจะมีความเสี่ยง หรือ อัตราการเสียชีวิตที่มีความแตกต่างกันในแต่ละช่วงอายุ ซึ่งในการจัดการความเสี่ยงในธุรกิจประกันชีวิต ตารางชีพหรือตารางมรณะ ได้ถูกนำมาใช้เป็นเครื่องมือในการประมาณค่าความเสี่ยง การเสียชีวิตของผู้เอาประกันภัยในแต่ละช่วงอายุที่แตกต่างกัน โดยการประมาณค่าหรือคำนวณค่าอัตรามรณะที่ใช้ในตารางมรณะ มีความสำคัญเป็นอย่างยิ่งในการดำเนินธุรกิจประกันชีวิต ทั้งนี้เนื่องจากการเลือกใช้ตารางมรณะที่เหมาะสม จะส่งผลให้บริษัทสามารถจัดการความเสี่ยงภัยได้อย่างมีประสิทธิภาพ

ในปัจจุบัน ได้มีประกาศใช้ตารางมรณะ ปี 2560 เพื่อเป็นข้อมูลที่ใช้ในการดำเนินธุรกิจประกันชีวิต โดยตารางมรณะไทย (Thai Mortality Ordinary) จัดทำโดยสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) โดยทาง คปภ. จะมีประกาศตารางมรณะในทุก ๆ 10 ปี จากที่กล่าวข้างต้น ตารางมรณะมีไว้เพื่อจัดการความเสี่ยงในธุรกิจประกันชีวิต กล่าวคือ เพื่อใช้ในการคำนวณอัตราเบี้ยประกันชีวิต นอกจากนี้ตารางมรณะสามารถ นำมาประยุกต์ใช้ในการคำนวณผลประโยชน์พนักงาน เพื่อคำนวณหาโอกาสที่พนักงานจะทำงานกับบริษัทจนเกษียณอายุเป็นเท่าไรได้อีกด้วย สำหรับการประยุกต์ใช้ตารางมรณะกับธุรกิจประกันชีวิต สามารถนำตารางมรณะมาประยุกต์ใช้ ในการกำหนดอัตราเบี้ยประกันภัย โดยการกำหนดอัตราเบี้ยประกันของแบบประกันจะคำนวณได้ โดยอาศัยปัจจัย 3 ปัจจัย คือ อัตรามรณะ อัตราดอกเบี้ย และค่าใช้จ่ายดำเนินงาน นอกจากนี้ ในการวิเคราะห์เงินส่วนเกิน (Surplus Analysis) เพื่อประเมินผลการทำงานของบริษัท ซึ่งเป็นการวิเคราะห์ถึงปัจจัยในการกำหนดอัตราเบี้ยประกันภัยของการประกันชีวิตในแต่ละรูปแบบ

ตัวแบบที่ใช้ในการคำนวณอัตรามรณะมีตัวแบบที่หลากหลาย การเลือกใช้ตัวแบบที่เหมาะสมจะส่งผลให้การประมาณค่า และการพยากรณ์มีความแม่นยำมากขึ้น โดยการศึกษาตัวแบบอัตรามรณะ ตัวแบบที่รู้จักกันดีคือ ตัวแบบลี-คาร์เตอร์ (Lee-Carter Model) (Lee และ Carter, 1992) ซึ่งถูกพัฒนาโดย Lee และ Carter ในปี ค.ศ. 1992 ซึ่งต่อมา ได้มีผู้พัฒนาตัวแบบเพื่อใช้ในการคำนวณอัตรามรณะอีกหลายตัวแบบ ไม่ว่าจะเป็นตัวแบบ เอช-โคฮอท (Age-Cohort) (Ranshaw และ Haberman, 2006) ซึ่งเป็นตัวแบที่ถูกขยายมาจาก

ตัวแบบลี-คาร์เตอร์ โดยมีปัจจัยด้านอายุ และดัชนีการเปลี่ยนแปลงเวลาของรุ่น ในแต่ละช่วงอายุในตัวแบบที่เพิ่มลงไปจากตัวแบบลี-คาร์เตอร์ ตัวแบบเอช-พี เรียต-โคฮอท (Age-Period-Cohort) (Currie, 2006) ซึ่งเป็นตัวแบบขยายจากตัวแบบลี-คาร์เตอร์ ซึ่งมีปัจจัยทั้งด้านอายุ เวลา และดัชนีการเปลี่ยนแปลงเวลาของรุ่น ในแต่ละช่วงอายุในตัวแบบ เป็นต้น ในงานวิจัยนี้ ได้ทำการศึกษาตัวแบบลี-คาร์เตอร์ และตัวแบบ เอช-โคฮอท ภายใต้สมมติฐานจำนวนการตาย แบ่งเป็นสองกรณีคือ จำนวนการตายมีการแจกแจงปัวซง และ มีการแจกแจงทวินามลบ ซึ่งตัวแบบที่เหมาะสมจะถูกนำมาใช้ในการพยากรณ์อัตรามรณะไทย และคำนวณอายุขัยโดยเฉลี่ย ของประชากรไทย จำแนกตามช่วงอายุและเพศ

2. วัตถุประสงค์การวิจัย

2.1 เพื่อคัดเลือกหาตัวแบบอัตรามรณะที่เหมาะสมในการประมาณค่าอัตรามรณะ ระหว่างตัวแบบ Lee-Carter และตัวแบบ Age-Cohort Model โดยพิจารณาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มจำนวนการตายของประชากร 2 กรณี ดังนี้

2.1.1 มีการแจกแจงแบบปัวซง

2.1.2 มีการแจกแจงแบบทวินามลบ

2.2 เพื่อประมาณค่าและพยากรณ์อัตรามรณะไทย

2.3 เพื่อพยากรณ์อายุขัยโดยเฉลี่ยของประชากรไทย

3. วิธีดำเนินการวิจัย

3.1 ข้อมูล

ข้อมูลที่ใช้ในงานวิจัยประกอบด้วยข้อมูลการตายของประชากรไทย จากสำนักนโยบายและยุทธศาสตร์ กระทรวงสาธารณสุข และข้อมูลประชากรกลางปี จากกระทรวงมหาดไทย โดยจำแนกตามเพศ และช่วงอายุ 0, 1-4, 5-9, 10-14, 10-15, 16-19, 20-24, 25-29,.., 80-84 และ 85 ปี ขึ้นไป ในปี พ.ศ. 2552-2562

โดยค่าประมาณอัตรามรณะกลางปี $(\hat{m}_{x,t})$ ของประชากรอายุ x ปี ในปีที่ t มีค่าดังนี้

$$\hat{m}_{x,t} = \frac{D_{x,t}}{ETR_{x,t}}$$

เมื่อ $D_{x,t}$ คือ จำนวนการตายกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t และ

$ETR_{x,t}$ คือ จำนวนประชากรกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t

x คือ ช่วงอายุ 0, 1-4, 5-9, 10-14, 10-15, 16-19, 20-24, 25-29, .., 80-84 และ 85 ปี ขึ้นไป

t แทน ปี คือ พ.ศ. 2552- 2562

3.2 ตัวแบบอัตรามรณะ (Mortality Model)

3.2.1 ตัวแบบลี-คาร์เตอร์ (Lee-Carter Model) (Lee และ Carter, 1992) ตัวแบบลี-คาร์เตอร์ กำหนดโดย

$$m_{x,t} = exp(\alpha_x + \beta_x k_t + \varepsilon_{xt}),$$

เมื่อ $m_{x,t}$ แทนอัตรามรณะกลางปี,

α_x แทนค่าเฉลี่ยลอการิทึมของอัตรามรณะรายอายุ x,

β_x แทนอัตราการลดลงของอัตรามรณะในรายอายุต่างๆ,

k_t แทนดัชนีการเปลี่ยนแปลงเวลาของอัตรามรณะปีที่ t และ

$\epsilon_{x,t}$ แทนความคลาดเคลื่อนของตัวแบบ

3.2.2 ตัวแบบเอช-โคฮอท (Age-Cohort Model) (Renshaw และ Haberman, 2006)

$$m_{x,t} = exp(\alpha_x + \beta_x^{(0)} k_t + \beta_x^{(1)} \gamma_{t-x} + \epsilon_{x,t})$$

เป็นตัวแบบลีคาร์เตอร์ที่เพิ่มปัจจัยอิทธิพลกระทบตามรุ่น (cohort effect) เข้าไปในตัวแบบ

เมื่อ γ_{t-x} แทนดัชนีการเปลี่ยนแปลงเวลาของรุ่น (Cohort Effect) ช่วงอายุ $t-x$

3.3 การประมาณค่าพารามิเตอร์

ในงานวิจัยนี้เราทำการประมาณค่าพารามิเตอร์ของตัวแบบ โดยพิจารณา 2 สมมติฐานเกี่ยวกับจำนวนการตายของประชากร คือ

3.3.1 ภายใต้สมมติฐานจำนวนการตายมีการแจกแจงปัวซง

การประมาณค่าพารามิเตอร์โดยกำหนดจำนวนการตายกลางปีของประชากรมีการแจกแจงแบบปัวซง

$$D_{x,t} \sim Poisson(E_{x,t} m_{x,t})$$

โดย $E(D_{x,t}) = E_{x,t} m_{x,t}$

ใช้วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation, MLE) ในการประมาณค่าเมื่อกำหนดให้จำนวนการตายของประชากรกลางปีมีการแจกแจงปัวซง การประมาณค่าพารามิเตอร์วิธีด้วยภาวะน่าจะเป็นสูงสุดภายใต้ตัวแปรสุ่มจำนวนการตายกลางปีมีการแจกแจงปัวซง โดยพิจารณาจาก ค่า Deviance ที่สอดคล้อง (Renshaw and Harberman, 2003) ดังนี้

$$D(d_{x,t}, \hat{d}_{x,t}) = \sum_{x,t} dev(d_{x,t}, \hat{d}_{x,t}) = \sum_{x,t} 2\omega_{x,t} \left\{ d_{x,t} ln\left(\frac{d_{x,t}}{\hat{d}_{x,t}}\right) - (d_{x,t} - \hat{d}_{x,t}) \right\}$$

โดย $\omega_{x,t} = \begin{cases} 1, E_{x,t} > 0 \\ 0, E_{x,t} = 0 \end{cases}$, $\hat{d}_{x,t} = E_{x,t} \hat{m}_{x,t}$.

3.3.2 ภายใต้สมมติฐานจำนวนการตายมีการแจกแจงทวินามลบ

สำหรับการพิจารณาจำนวนการตายมีการแจกแจงทวินามลบ ในการประมาณค่าพารามิเตอร์ จะถูกแบ่งเป็น 2 กรณี คือ

1) พารามิเตอร์การกระจายขึ้นอยู่กับอายุ (Age-dependent Dispersion Parameter)

ภายใต้การพิจารณาจำนวนการตาย $D_{x,t}$ มีการแจกแจง ทวินามลบ กรณี พารามิเตอร์การกระจาย $(\bar{\gamma}_x)$ ขึ้นอยู่กับอายุ (Yasungnoen, 2015)

$$D_{x,t} \sim NBin(\bar{\gamma}_x^{-1}, \frac{\left(E_{x,t} m_{x,t} \bar{\gamma}_x\right)^{-1}}{1 + \left(E_{x,t} m_{x,t} \bar{\gamma}_x\right)^{-1}}) = NBin\left(\bar{\gamma}_x^{-1}, \frac{1}{E_{x,t} m_{x,t} \bar{\gamma}_x + 1}\right)$$

ค่าเฉลี่ย และความแปรปรวน คือ

$$E[D_{x,t}] = E_{x,t} m_{x,t},$$
$$Var[D_{x,t}] = E_{x,t} m_{x,t} + \bar{\gamma}_x (E_{x,t} m_{x,t})^2$$

2) พารามิเตอร์การกระจายไม่ขึ้นอยู่กับอายุ (Age-independent Dispersion Parameter) (Yasungnoen, 2015)

$$D_{x,t} \sim NBin(\bar{\gamma}_x^{-1}, \frac{\left(E_{x,t} m_{x,t} \bar{\gamma}\right)^{-1}}{1 + \left(E_{x,t} m_{x,t} \bar{\gamma}\right)^{-1}}) = NBin\left(\bar{\gamma}_x^{-1}, \frac{1}{E_{x,t} m_{x,t} \bar{\gamma} + 1}\right).$$

การประมาณค่าพารามิเตอร์วิธีด้วยภาวะน่าจะเป็นสูงสุดภายใต้ตัวแปรสุ่มจำนวนการตายมีการแจกแจงทวินามลบ พารามิเตอร์การกระจาย $(\bar{\gamma})$ ไม่ขึ้นอยู่กับอายุ โดยพิจารณาจากค่า Deviance ที่สอดคล้อง กับตัวแบบทวินามลบ ดังนี้

$$D(d_{x,t}, \hat{d}_{x,t}) = \sum_{x,t} dev(d_{x,t}, \hat{d}_{x,t}) = \sum_{x,t} 2\omega_{x,t} \left\{ d_{x,t} ln\left(\frac{d_{x,t}}{\hat{d}_{x,t}}\right) - (d_{x,t} + \frac{1}{\bar{\gamma}}) \ln(\frac{1+\bar{\gamma} d_{x,t}}{1+\bar{\gamma} \hat{d}_{x,t}}) \right\}$$

โดย $\omega_{x,t} = \begin{cases} 1, E_{x,t} > 0 \\ 0, E_{x,t} = 0 \end{cases}$, $\hat{d}_{x,t} = E_{x,t} \hat{m}_{x,t}$.

3.4 การประมาณค่าพารามิเตอร์

ในขั้นตอนการประมาณค่าพารามิเตอร์ θ เราใช้วิธีการปรับปรุงค่าพารามิเตอร์ θ ด้วยวิธีการทำซ้ำ ที่เรียกว่า Goodman Algorithm (Goodman, 1979; Brounhns et al., 2002) ดังนี้

$$update(\hat{\theta}^{k+1}) = \hat{\theta}^{k} - \frac{\frac{\partial D}{\partial \theta}}{\frac{\partial^2 D}{\partial \theta^2}}$$

เมื่อ k=0,1,2,..... โดยการกำหนดค่าเริ่มต้น(initial value) เป็นค่าสุ่ม (random number)

3.5 การประมาณค่าตัวแบบที่เหมาะสม

3.5.1 ตัวแบบลี-คาร์เตอร์ (Lee-Carter Model) ตัวแบบลี-คาร์เตอร์ :

$$m_{x,t} = exp(\alpha_x + \beta_x k_t + \varepsilon_{xt})$$

จะได้ตัวแบบเพื่อประมาณค่า $\hat{m}_{x,t}$ คือ

$$\hat{m}_{x,t} = exp(\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t).$$

3.5.2 ตัวแบบเอช-โคฮอท (Age-Cohort Model) ตัวแบบเอช-โคฮอท :

$$m_{x,t} = exp(\alpha_x + \beta_x^{(0)} k_t + \beta_x^{(1)} \gamma_{t-x} + \varepsilon_{x,t})$$

จะได้ตัวแบบเพื่อประมาณค่า $\hat{m}_{x,t}$ คือ

$$\hat{m}_{x,t} = exp(\hat{\alpha}_x + \hat{\beta}_x^{(0)} \hat{k}_t + \hat{\beta}_x^{(1)} \hat{\gamma}_{t-x})$$

3.6 Overdispersion test

การทดสอบ overdispersion ด้วยวิธีของ Dean และ Lawless (1989) และ Dean (1992) ใช้ทดสอบการเกิด overdispersion ซึ่งจะเป็นการตรวจสอบข้อมูลว่ามีการแจกแจงทวินามลบหรือไม่ โดยการพิจารณาฟังก์ชันความแปรปรวนของการแจกแจงทวินามลบ

$$Var(d_{x,t}) = E[d_{x,t}] + \tau[E[d_{x,t}]]^2$$

กำหนด สมมติฐานการทดสอบ $H_0: \tau = 0$ และ $H_1: \tau > 0$

สถิติที่ใช้ในการทดสอบ Score Statistic Q :

$$Q = \frac{\sum_{x,t}((d_{x,t} - \hat{d}_{x,t})^2 - d_{x,t})}{\sqrt{2\sum_{x,t} \hat{d}_{x,t}^2}} \sim N(0,1)$$

3.7 การเปรียบเทียบตัวแบบด้วยการวิเคราะห์หาความคลาดเคลื่อนจากการประมาณค่า

สถิติความคลาดเลื่อนที่ใช้ในการวิเคราะห์ ประกอบด้วย

- ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error, MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

- ค่าเฉลี่ยของร้อยละความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error, MAPE)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \times 100$$

- ค่าเฉลี่ยความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Deviation, MAD)

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|$$

เมื่อ Y_i คือ ค่าอัตรามรณะที่คำนวณจากข้อมูลจริง และ $\hat{Y}_i$ คือ ค่าอัตรามรณะที่ประมาณจากตัวแบบที่ทำการศึกษา

4. ผลการวิจัย

4.1 ผลการประมาณค่าพารามิเตอร์ของตัวแบบภายใต้สมมติฐานการแจกแจงปัวซง

4.1.1 ความคลาดเคลื่อนจากการประมาณค่าอัตรามรณะ

การประมาณค่าพารามิเตอร์ เพื่อประมาณค่าอัตรามรณะ $(\hat{m}_{x,t})$ ภายใต้วิธีการประมาณค่าพารามิเตอร์ โดยสมการการแจกแจงปัวซง จากค่าความคลาดเคลื่อน แสดงดังตารางที่ 1

ตารางที่ 1 แสดงค่าความคลาดเคลื่อนจากการประมาณค่าอัตรามรณะภายใต้สมมติฐานการแจกแจงปัวซง

ค่าความคลาดเคลื่อน	เพศชาย		เพศหญิง	
	Lee-Carter Model	Age-Cohort Model	Lee-Carter Model	Age-Cohort Model
MSE	**0.00000134**	0.00000659	0.000000648	**0.000000241**
MAD	**0.000314**	0.00556	0.000244	**0.000174**
MAPE	1.968758	**1.749657**	1.825886	**1.797509**

จากตารางแสดงค่าความคลาดเคลื่อน โดยใช้ค่าความคลาดเคลื่อน MSE, MAD และ MAPE พบว่า สำหรับเพศชาย ตัวแบบ Lee-Carter ให้ค่า MSE และ MAD ต่ำที่สุด และตัวแบบ Age-Cohort ให้ค่า ความคลาดเลื่อนต่ำที่สุด และสำหรับเพศหญิง พบว่า ตัวแบบ Age-Cohort ให้ค่าความคลาดเคลื่อนต่ำที่สุดทั้งสามค่า

4.1.2 ผลการประมาณค่าพารามิเตอร์และตัวแบบ

จากผลการประมาณค่าและค่าคลาดเคลื่อนจากการประมาณค่าของตัวแบบอัตรามรณะ ในกรณีการประมาณค่าภายใต้สมมติฐานจำนวนการตายมีการแจกแจงปัวซง พบว่า สำหรับเพศหญิงตัวแบบที่เหมาะสม คือ ตัวแบบ Lee-Cater และสำหรับเพศชาย ตัวแบบที่เหมาะสมคือ ตัวแบบ Age-Cohort

4.2 ผลการทดสอบ Overdispersion test

ตารางที่ 2 แสดงค่า Score Statistics ทุกค่า ให้ค่า p-value น้อยกว่า 0.01 ดังนั้น เราสามารถปฏิเสธ H_0 นั่นคือ ผลลัพธ์ แสดงว่าข้อมูลที่ใช้ประมาณค่าอัตรามรณะ เกิด Overdispersion ดังนั้น ผลลัพธ์ จะนำไปสู่วิธีการประมาณค่าพารามิเตอร์ ภายใต้สมมติฐานจำนวนการตายมีการแจกแจงแบบทวินามลบ

ตารางที่ 2 แสดงค่า Score Statistics

เพศ	ตัวแบบ Lee-Carter	ตัวแบบ Age-Cohort
ชาย	46.39999	32.2963
หญิง	2.311341	167.2319

4.3 ผลการประมาณค่าพารามิเตอร์ของตัวแบบภายใต้สมมติฐานจำนวนการตายมีการแจกแจงทวินามลบ

การประมาณค่าตัวแบบอัตรามรณะ ภายใต้สมมติฐานจำนวนการตายมีการแจกแจงทวินามลบ จะพิจารณาออกเป็น 2 กรณี คือ พารามิเตอร์การกระจายขึ้นอยู่กับอายุ (Age-dependent dispersion parameter) และ พารามิเตอร์การกระจายไม่ขึ้นอยู่กับอายุ (Age-independent dispersion parameter)

4.3.1 ความคลาดเคลื่อนจากการประมาณค่าตัวแบบ

จากการประมาณค่าพารามิเตอร์ และประมาณค่าอัตรามรณะ โดยตัวแบบ Lee-Carter และ ตัวแบบ Age-Cohort คำนวณค่าคลาดเคลื่อนของตัวแบบ ดังตารางที่ 3

ตารางที่ 3 แสดงค่าคลาดเคลื่อนจากการประมาณค่าอัตรามรณะ สำหรับเพศชาย

ความคลาดเคลื่อน	ตัวแบบ Lee-Carter		ตัวแบบ Age-Cohort	
	Age-dependent	Age-independent	Age-dependent	Age-independent
MSE	**0.000000122**	0.000000124	0.00000123	0.0000016
MAD	**0.000171**	0.000172	0.000302	0.00037
MAPE	1.894202	1.892517	1.642232	**1.59431**

ตารางที่ 4 แสดงค่าคลาดเคลื่อนจากการประมาณค่าอัตรามรณะ สำหรับเพศหญิง

ความคลาดเคลื่อน	ตัวแบบ Lee-Carter		ตัวแบบ Age-Cohort	
	Age-dependent	Age-independent	Age-dependent	Age-independent
MSE	0.000000346	0.000000242	0.00000522	**0.00000209**
MAD	**0.000174**	0.00019	0.000427	0.000311
MAPE	1.797415	1.780661	1.778041	**1.551575**

4.3.2 ผลการคัดเลือกตัวแบบ

จากการประมาณค่าอัตรามรณะ โดยตัวแบบ Lee-Carter และตัวแบบ Age-Cohort ภายใต้สมมติฐานการแจกแจงของจำนวนการตายมีการแจกแจงทวินามลบ ซึ่งแบ่งออกเป็น 2 กรณีคือ กรณีพารามิเตอร์การกระจายขึ้นอยู่กับอายุ และกรณีพารามิเตอร์การกระจายไม่ขึ้นอยู่กับอายุ เมื่อพิจารณาค่าความคลาดเคลื่อนจากการประมาณค่าตัวแบบ จำแนกตามเพศ พบว่า

1) สำหรับเพศชาย ตัวแบบ Lee-Cater กรณีพารามิเตอร์การกระจายขึ้นอยู่กับอายุ ให้ค่า MSE คือ 0.00000012 และค่า MAD คือ 0.000171 ซึ่งเป็นค่าน้อยที่สุดเมื่อเปรียบเทียบกับ ตัวแบบ Age-Cohort และสำหรับค่า MAPE ให้ค่าต่ำสุดจากการคำนวณจากตัวแบบ Age-Cohort กรณีพารามิเตอร์การกระจายไม่ขึ้นกับอายุ และ

2) สำหรับเพศหญิง ตัวแบบ Age-Cohort กรณีพารามิเตอร์การกระจายไม่ขึ้นกับอายุ ให้ค่า MSE และ MAPE ต่ำสุด

ดังนั้น ในการพยากรณ์ค่าอัตรามรณะ สำหรับเพศชาย เราจะเลือกใช้ตัวแบบ Lee-Carter ในการพยากรณ์อัตรามรณะ ซึ่งในการประมาณค่าพารามิเตอร์ จะใช้กรณีพารามิเตอร์การกระจายขึ้นอยู่กับอายุ และสำหรับเพศหญิง เราจะเลือกใช้ตัวแบบ Age-Cohort ในการพยากรณ์อัตรามรณะ โดยการประมาณค่าพารามิเตอร์ของตัวแบบ จะใช้กรณีพารามิเตอร์การกระจายไม่ขึ้นกับอายุ

4.3.3 การพยากรณ์อัตรามรณะและอายุขัยเฉลี่ย

จากผลการประมาณค่าพารามิเตอร์ และค่าอัตรามรณะไทย สำหรับเพศชาย เราเลือกใช้ตัวแบบ Lee-Carter เป็นตัวแบบเพื่อพยากรณ์อัตรามรณะ โดยใช้ การประมาณค่าพารามิเตอร์ กรณีจำนวนการตายมีการแจกแจงทวินามลบ โดยพารามิเตอร์การกระจายขึ้นอยู่กับอายุ และสำหรับเพศหญิง ใช้ตัวแบบ Age-Cohort เป็นตัวแบบเพื่อพยากรณ์อัตรามรณะ โดยวิธีการประมาณค่าอัตรามรณะ โดยใช้การประมาณค่าพารามิเตอร์ กรณีจำนวนการตายมีการแจกแจงทวินาม โดยพารามิเตอร์การกระจายไม่ขึ้นกับอายุ

ในการพยากรณ์อัตรามรณะ พารามิเตอร์ที่ขึ้นอยู่กับเวลา จะเป็นพารามิเตอร์เพื่อใช้ในการพยากรณ์เพื่อให้ได้อัตรามรณะในอนาคตข้างหน้า ด้วยวิธีการสร้างแบบจำลอง Box-Jenkins สำหรับการพยากรณ์พารามิเตอร์ที่ขึ้นอยู่กับเวลา คือ k_t และ γ_{t-x} ด้วยวิธีการสร้างแบบจำลอง Autoregression Integrated Moving Average Model (ARIMA)

สำหรับตัวแบบ Lee-Carter ค่า $\dot{k}_{t_n+s}, s = 1,2,3,\ldots,10$ และ $n = 2019$ ค่าพยากรณ์อัตรามรณะ สามารถคำนวณได้ดังนี้

$$\dot{m}_{x,t_n+s} = exp(\hat{\alpha}_x + \hat{\beta}_x \hat{k}_{t_n+s})$$

สำหรับตัวแบบ Age-Cohort ค่า $\dot{k}_{t_n+s}, s = 1,2,3,\ldots,10$ และ $n = 2019$ ค่าประมาณ ของพารามิเตอร์ cohort effect $\hat{\gamma}_z$ เมื่อ $z\epsilon[t_1 - x_k, t_n - x_1]$ คำนวณได้ดังนี้ (Renshaw and Harberman, 2006)

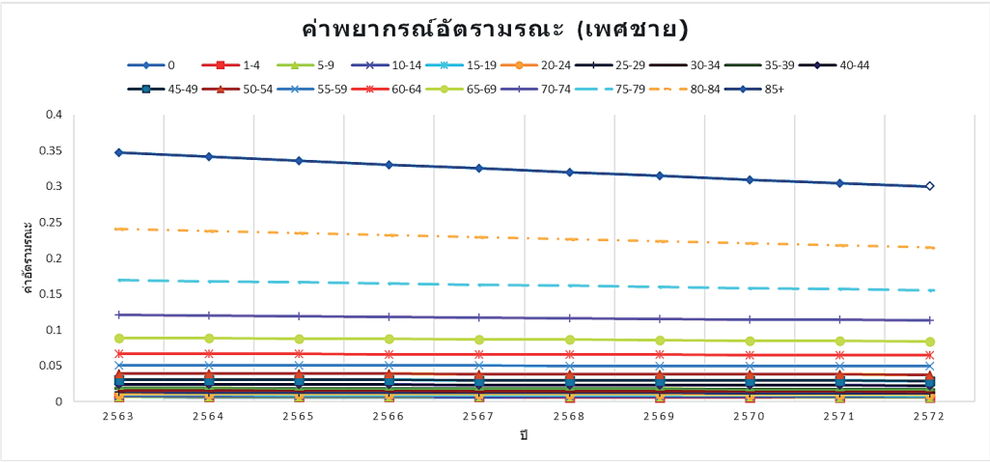
$$\tilde{\gamma}_{t_n-x+s} = \begin{cases} \hat{\gamma}_{t_n-x+s}, & 0<s\leq x-x_1 \\ \dot{\gamma}_{t_n-x+s}, & s>x-x_1 \end{cases}.$$

โดยค่าพยากรณ์อัตรามรณะ คำนวณได้ดังนี้

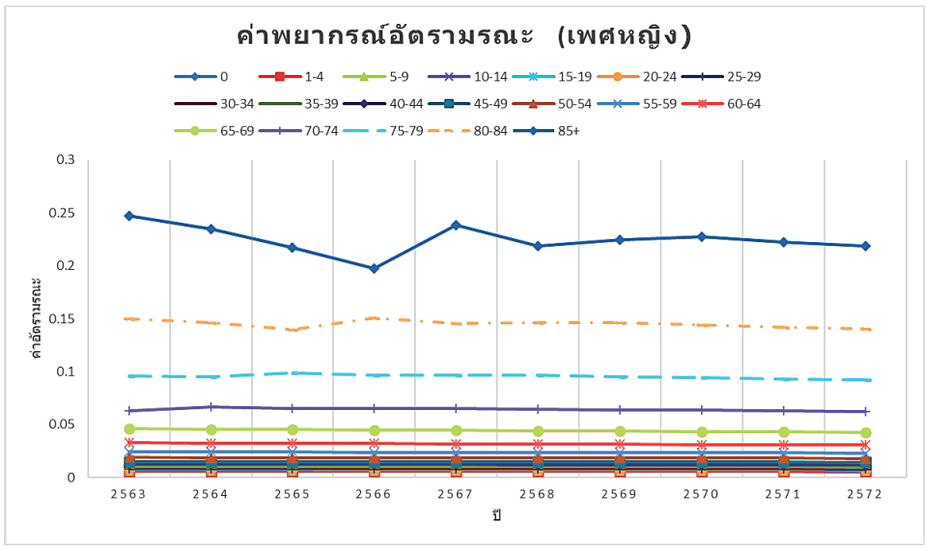
$$\dot{m}_{x,t_n+s} = \hat{m}_{x,t_n} exp\left\{\hat{\beta}_x^{(0)}\left(\dot{k}_{t_n+s} - \hat{k}_{t_n}\right) + \hat{\beta}_x^{(1)}\left(\tilde{\gamma}_{t_n-x+s} - \hat{\gamma}_{t_n-x}\right)\right\}$$

1) ผลการพยากรณ์อัตรามรณะ

จากการคัดเลือกตัวแบบ ตัวแบบที่เหมาะสม สำหรับเพศชาย คือตัวแบบ Lee-Carter และสำหรับเพศหญิง คือ ตัวแบบ Age-Cohort ดังนั้น เรานำตัวแบบดังกล่าว มาพยากรณ์อัตรามรณะไทย ในปี พ.ศ. 2563-2572 ผลการพยากรณ์อัตรามรณะ แสดงดังภาพที่ 1 และ 2 ของเพศชาย และเพศหญิงตามลำดับ



ภาพที่ 1 แสดงค่าพยากรณ์อัตรามรณะ เพศชาย ปี พ.ศ. 2563-2572



ภาพที่ 2 แสดงค่าพยากรณ์อัตรามรณะ เพศหญิง ปี พ.ศ. 2563-2572

2) ผลการพยากรณ์อายุขัยเฉลี่ย

จากการพยากรณ์อัตรามรณะสามารถนำมาคำนวณอายุขัยเฉลี่ยในแต่ละช่วงอายุได้ (e_x) โดยสำหรับเพศชาย จากการประมาณค่าอัตรามรณะด้วยตัวแบบ Lee-Carter โดยวิธีการประมาณค่าพารามิเตอร์ ภายใต้สมมติฐานจำนวนการตายมีการแจกแจงทวินามลบ กรณีพารามิเตอร์การกระจายขึ้นอยู่กับอายุ จะได้อายุขัยเฉลี่ยดังตาราง

ตารางที่ 5 แสดงอายุขัยโดยเฉลี่ย (e_x) ของแต่ละช่วงอายุ สำหรับเพศชาย

ช่วงอายุ	ปี									
	2563	2564	2565	2566	2567	2568	2569	2570	2071	2572
0	74.87	74.93	74.98	75.04	75.09	75.15	75.20	75.25	75.30	75.36
1-4	74.35	74.40	74.45	74.49	74.54	74.59	74.63	74.68	74.72	74.77
5-9	70.52	70.56	70.61	70.65	70.69	70.74	70.78	70.82	70.86	70.91
10-14	65.63	65.67	65.71	65.75	65.79	65.83	65.87	65.91	65.95	65.99
15-19	60.80	60.84	60.88	60.92	60.96	61.00	61.04	61.08	61.12	61.16
20-24	56.24	56.28	56.32	56.36	56.40	56.44	56.48	56.52	56.56	56.59
25-29	51.69	51.73	51.76	51.80	51.84	51.87	51.91	51.94	51.98	52.02
30-34	47.16	47.19	47.22	47.26	47.29	47.32	47.35	47.39	47.42	47.45
35-39	42.69	42.72	42.75	42.77	42.80	42.82	42.85	42.88	42.90	42.93
40-44	38.37	38.39	38.41	38.43	38.45	38.48	38.50	38.52	38.54	38.57
45-49	34.19	34.22	34.24	34.27	34.29	34.32	34.34	34.37	34.39	34.42
50-54	30.10	30.13	30.16	30.19	30.21	30.24	30.27	30.30	30.32	30.35
55-59	26.12	26.15	26.18	26.20	26.23	26.26	26.29	26.32	26.34	26.37
60-64	22.23	22.26	22.29	22.32	22.35	22.38	22.41	22.44	22.47	22.50
65-69	18.43	18.46	18.48	18.51	18.54	18.56	18.59	18.62	18.64	18.67
70-74	14.64	14.66	14.68	14.70	14.72	14.74	14.76	14.77	14.79	14.81
75-79	10.83	10.84	10.85	10.87	10.88	10.89	10.90	10.91	10.92	10.93
80-84	6.86	6.86	6.87	6.87	6.88	6.88	6.89	6.89	6.90	6.90
85+	2.50	2.50	2.50	2.50	2.50	2.50	2.50	2.50	2.50	2.50

ตารางที่ 6 แสดงอายุขัยโดยเฉลี่ย (e_x) ของแต่ละช่วงอายุ สำหรับเพศหญิง

ช่วงอายุ	ปี									
	2563	2564	2565	2566	2567	2568	2569	2570	2071	2572
0	79.94	79.93	79.98	79.97	80.01	80.05	80.09	80.14	80.19	80.24
1-4	79.40	79.40	79.45	79.44	79.48	79.52	79.56	79.61	79.66	79.71
5-9	75.54	75.53	75.58	75.57	75.61	75.64	75.68	75.73	75.78	75.83
10-14	70.63	70.62	70.66	70.65	70.69	70.72	70.76	70.81	70.86	70.90
15-19	65.73	65.72	65.77	65.76	65.79	65.83	65.86	65.91	65.96	66.00
20-24	60.89	60.88	60.92	60.91	60.95	60.98	61.02	61.06	61.11	61.15
25-29	56.05	56.03	56.08	56.07	56.10	56.13	56.17	56.21	56.26	56.30
30-34	51.22	51.20	51.24	51.23	51.26	51.29	51.33	51.37	51.42	51.46
35-39	46.43	46.41	46.45	46.43	46.46	46.49	46.52	46.56	46.60	46.64
40-44	41.70	41.68	41.71	41.70	41.72	41.75	41.78	41.81	41.86	41.89
45-49	37.04	37.02	37.05	37.03	37.06	37.08	37.11	37.14	37.18	37.22
50-54	32.47	32.44	32.48	32.45	32.48	32.50	32.53	32.56	32.60	32.63
55-59	27.99	27.97	28.00	27.97	28.00	28.01	28.04	28.07	28.11	28.14
60-64	23.62	23.58	23.61	23.58	23.60	23.61	23.64	23.66	23.70	23.73
65-69	19.38	19.34	19.36	19.32	19.34	19.34	19.36	19.38	19.41	19.43
70-74	15.26	15.20	15.22	15.18	15.19	15.18	15.19	15.21	15.23	15.25
75-79	11.07	11.15	11.12	11.10	11.12	11.11	11.11	11.12	11.13	11.15
80-84	6.92	6.94	6.99	6.92	6.95	6.94	6.94	6.94	6.95	6.95
85+	2.50	2.50	2.50	2.50	2.50	2.50	2.50	2.50	2.50	2.50

5. อภิปรายผล

จากการศึกษาตัวแบบเพื่อพยากรณ์อัตรามรณะไทย โดยตัวแบบ Lee-Carter และตัวแบบ Age-Cohort ภายใต้สมมติฐาน จำนวนการตายมีการแจกแจงปัวซง และมีการแจกแจงทวินามลบ กรณีแรกภายใต้สมมติฐาน จำนวนการตายมีการแจกแจงปัวซง ผลลัพธ์พบว่า สำหรับเพศชาย ตัวแบบที่เหมาะสมคือตัวแบบ Lee-carter และสำหรับ เพศหญิง คือตัวแบบ Age-Cohort ซึ่งข้อมูลจำนวนการตายที่มี อาจจะไม่มีการแจกแจงปัวซง จึงได้ทดสอบ Overdispersion ผลการทดสอบสมมติฐานพบว่า ข้อมูลจำนวนการตายไม่สามารถปฏิเสธได้ว่ามีการแจกแจงทวินามลบที่ระดับนัยสำคัญ 0.01 ดังนั้น จึงนำการประมาณค่า พารามิเตอร์ภายใต้สมมติฐานจำนวนการตายมีการแจกแจงทวินามลบมาประยุกต์ใช้ในการประมาณค่าอัตรามรณะ ภายใต้เงื่อนไขนี้ แบ่งออกเป็น 2 กรณี คือ พารามิเตอร์การกระจายไม่ขึ้นอยู่กับอายุ และพารามิเตอร์การกระจายขึ้นกับอายุ ซึ่งผลการประมาณค่าพบว่าตัวแบบที่เหมาะสมสำหรับเพศชาย ตัวแบบที่เหมาะสม คือ ตัวแบบ Lee-Cater กรณีพารามิเตอร์การกระจายขึ้นอยู่กับอายุ และ สำหรับเพศหญิง พบว่าตัวแบบที่เหมาะสมคือ ตัวแบบ Age-Cohort กรณีพารามิเตอร์การกระจายไม่ขึ้นอยู่กับอายุ

6. สรุปและข้อเสนอแนะ

6.1 สรุป

จากการประมาณค่าอัตรามรณะสำหรับข้อมูลประเทศไทย ซึ่งแบ่งตามช่วงอายุ โดยพิจารณาข้อมูลตั้งแต่ปี 2552-2562 เพื่อทำการพยากรณ์อัตรามรณะ ในปี พ.ศ. 2563-2572 โดยตัวแบบ Lee-Carter และ ตัวแบบ Age-Cohort จากผลลัพธ์การทดสอบ Overdispersion test และการประมาณค่าอัตรามรณะภายใต้สมมติฐานจำนวนการตายของประชากรไทยมีการแจกแจงทวินามลบ พบว่าตัวแบบที่เหมาะสมสำหรับเพศชายคือ ตัวแบบ Lee-Carter กรณีพารามิเตอร์การกระจายขึ้นอยู่กับอายุ และสำหรับเพศหญิง พบว่าตัวแบบที่เหมาะสมคือตัวแบบ Age-Cohort กรณีพารามิเตอร์การกระจายไม่ขึ้นอยู่กับอายุ ดังนั้นจึงพยากรณ์อัตรามรณะ ในปี พ.ศ. 2563-2572 ด้วยตัวแบบดังกล่าว ผลการพยากรณ์พบว่า แนวโน้มอัตรามรณะของทั้งเพศชายและเพศหญิง มีแนวโน้มลดลง นอกจากนี้จากผลการคำนวณเพื่อหาอายุขัยเฉลี่ยมีแนวโน้มเพิ่มขึ้น

6.2 ข้อเสนอแนะ

ผลจากการศึกษาตัวแบบอัตรามรณะเพื่อการพยากรณ์ค่าอัตรามรณะในอนาคต พบว่าตัวแบบที่เหมาะสมจะสามารถทำให้ทราบถึงแนวโน้มของอัตรามรณะของประชากรในอนาคตได้เป็นอย่างดี การเลือกใช้ตัวแบบที่เหมาะสมกับข้อมูลแต่ละชุดมีความจำเป็นอย่างยิ่ง เมื่อชุดข้อมูลต่างออกไป การเลือกใช้ตัวแบบในการพยากรณ์จะต่างออกไปด้วย ซึ่งจะทำให้ได้ตัวแบบที่เหมาะสมมากที่สุด ทำให้การพยากรณ์แม่นยำมากขึ้น ซึ่งในปัจจุบันตัวแบบเพื่อใช้ในการพยากรณ์มีหลายตัวแบบ รวมถึงการเลือกใช้วิธีการประมาณค่าพารามิเตอร์ที่เหมาะสมกับชุดข้อมูลและตัวแบบ มีความสำคัญมาก ผู้วิจัยควรคำนึงถึงและควรเลือกใช้ให้เหมาะสม เพื่อให้ได้ผลลัพธ์ หรือผลการพยากรณ์ที่แม่นยำหรือผลการพยากรณ์ที่ใกล้เคียงค่าจริงมากที่สุด

7. เอกสารอ้างอิง

1. Brouhns, N., Denuit, M., & Vermunt, J. K. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetable. Insurance: *Mathematics and Economics*, 373-393.
2. Currie, I. D. (2006). Smooting and forecasting mortality rates with P-splins. *Talk given at the Institute of actuaries.* http://www.macs.hw.ac.uk/~iain/research/talks/Mortality.pdf
3. Dean, C. B. (1992). Testing for overdispersion in Poisson and binomial regression. *Journal of the American Statistical Association*, *87,* 451-457.
4. Dean, C., & Lawless, J. F. (1989). Tests for detecting overdispersion in Poisson. *Journal of the American Statistical Association*, *84,* 467-472.
5. Goodman, L. A. (1979). Simple Models for the Analysis of Associate in Cross-Classification having Ordered categories. *Journal of the American Statistical Association*, *74*(367), 537-552.
6. Lee, R. D., & Carter, L. R. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association, 87*(419), 659-671.

7. Renshaw, A., & Haberman, S. (2003). Lee-Carter mortality forecasting with age-specific enhancement. Insurance: *Mathematics and Economics*, *33,* 255-272.
8. Renshaw, A., & Haberman, S. (2006). A Cohort-based extension to the Lee-Carter model for mortality reduction facstors. *Insurance: Mathematics and Economics*, *38,* 556-570.
9. Yasungnoen, N. (2015). *Forecasting Thai mortality by using a modified Lee-Carter model.* (Doctoral dissertation). Suranaree University of Technology, Nakhonratchasima, Thailand.