บทความวิจัย

การพยากรณ์อัตรามรณะของประชากรไทย

นฎกร หมัดเลียด[1*] และ สำรวม จงเจริญ[2]

บทคัดย่อ

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อทำการพยากรณ์ค่าอัตรามรณะของประชากรไทยในอีก 10 ปีข้างหน้า โดยใช้วิธีการพยากรณ์อัตรามรณะจาก 2 ตัวแบบ คือ การพยากรณ์อัตรามรณะด้วยตัวแบบของลี-คาร์เตอร์ (The Lee-Carter Model) และการพยากรณ์อัตรามรณะด้วยตัวแบบฮินแมน-อุลลาห์ (The Hyndman-Ullah Model) โดยอาศัยข้อมูลจำนวนประชากรจากกรมการปกครอง กระทรวงมหาดไทย และจำนวนประชากรตายจากสำนักงานปลัดกระทรวงกระทรวงสาธารณสุขในปี พ.ศ. 2546-พ.ศ. 2555 จำแนกตามเพศและอายุ พร้อมทั้งทำการเปรียบเทียบค่าพยากรณ์อัตรามรณะของ 2 ตัวแบบ เพื่อหาตัวแบบที่มีความเหมาะสมที่สุด ด้วยการเปรียบเทียบค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ย (Mean Absolute Percent Error) ของแต่ละตัวแบบ ผลการวิจัยพบว่า ค่าพยากรณ์อัตรามรณะในอีก 10 ปีข้างหน้าของ 2 ตัวแบบทั้งเพศชายและเพศหญิงมีลักษณะทิศทางเดียวกัน คือมีแนวโน้มลดลงเมื่อระยะเวลาผ่านไป และค่าพยากรณ์อัตรามรณะของเพศชายมีค่าสูงกว่าเพศหญิงเล็กน้อย เมื่อเปรียบเทียบค่า MAPE พบว่าการพยากรณ์อัตรามรณะด้วยตัวแบบฮินแมน-อุลลาห์ มีค่า MAPE น้อยกว่าของตัวแบบลี-คาร์เตอร์ทั้งเพศชายและเพศหญิง โดยจำนวนปีที่คาดว่าจะมีชีวิตอยู่ของประชากรทั้งสองเพศมีแนวโน้มเพิ่มขึ้น

คำสำคัญ: อัตรามรณะ ตัวแบบลี-คาร์เตอร์ ตัวแบบฮินแมน-อุลลาห์

[1] คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

[2] ศาสตราจารย์ประจำ คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

[*] ผู้นิพนธ์ประสานงาน, e-mail: ferinyfern@hotmail.com

Methods for Forecasting Thai Mortality Rate

Natakorn Madliad[1*] and Samruam Chongcharoen[2]

ABSTRACT

This research aims forecast the mortality rate of Thai population in the next 10 years by using 2 estimating methods: Lee-Carter Model and Hyndman-Ullah Model. Data used in the study are the number of population and the number of death by age and sex of year 2003-2012 from the Ministry of Interior and the Ministry of Public Health respectively. In addition to comparing the estimated mortality rates from all methods to find the model that is most appropriate using mean absolute percentage error (MAPE). The results show that both models have the predicted mortality rates for the next 10 years in the same direction and trend to decrease over time for both males and females. The predicted mortality rates of males are slightly higher than those of females. The comparison of the estimated mortality rates showed that the Hyndman-Ullah Model provides the most minimum MAPE for both males and females. The life expectancy has an increasing trend over time.

Keywords: Mortality Rate Lee-Carter Model Hyndman-Ullah Model

[1]Graduate School of Applied Statistics, National Institute of Development Administration

[2]Professor of Graduate School of Applied Statistics, National Institute of Development Administration

[*]Corresponding author, e-mail: ferinyfern@hotmail.com

บทนำ

เนื่องจากในปัจจุบันมีความเจริญก้าวหน้าในวิทยาการแพทย์ค่อนข้างมากทำให้แนวโน้มการเสียชีวิตของประชากรต่ำลง อัตราการตายหรืออัตรามรณะเป็นตัวหนึ่งที่ใช้ในตัวกำหนดโครงสร้างประชากร ที่มักจะถูกใช้เป็นดัชนีชี้วัดสุขภาพของประชากรอีกทั้งในธุรกิจประกันชีวิตและประกันสุขภาพ อัตรามรณะเป็นปัจจัยที่ใช้ในการคำนวณเบี้ยประกันภัยอีกด้วย เพื่อที่จะให้การคำนวณค่าเบี้ยประกันภัยมีความเหมาะสมจึงมีความจำเป็นที่จะต้องทราบแนวโน้มของอัตรามรณะ

จากข้อมูลอายุคาดเฉลี่ยเมื่อแรกเกิด* [1] ของสถาบันวิจัยประชากรและสังคม มหาวิทยาลัยมหิดล พบว่าอายุคาดเฉลี่ยเมื่อแรกเกิดของประชากรไทยทั้งเพศชายและเพศหญิงมีแนวโน้มเพิ่มขึ้นเรื่อยๆ ดังตารางต่อไปนี้

ตารางที่ 1 อายุคาดเฉลี่ยเมื่อแรกเกิดของประชากรไทย พ.ศ. 2540-พ.ศ. 2557

	อายุคาดเฉลี่ยเมื่อแรกเกิด																	
เพศ/ปี	2540	2541	2542	2543	2544	2545	2546	2547	2548	2549	2550	2551	2552	2553	2554	2555	2556	2557
ชาย	67.74	68.18	66.09	66.32	66.82	66.78	67.11	67.26	67.78	68.36	68.85	69.12	69.71	69.55	69.50	69.60	71.10	71.30
หญิง	76.15	75.52	74.17	74.17	74.22	74.30	74.74	74.68	75.13	75.64	75.89	76.19	76.85	76.90	76.30	76.90	78.10	78.20

ที่มา : สถาบันวิจัยประชากรและสังคม มหาวิทยาลัยมหิดล

จากตารางที่ 1 แสดงให้เห็นว่าประชากรไทยทั้งเพศชายและเพศหญิงมีแนวโน้มที่จะมีอายุยืนมากขึ้นกว่าในอดีต การพยากรณ์อัตรามรณะจึงถือเป็นสิ่งที่สำคัญและจำเป็นอย่างยิ่งสำหรับผู้ที่ดำเนินธุรกิจประกันชีวิต เนื่องจากธุรกิจประกันชีวิตเป็นธุรกิจที่มีความเสี่ยงต่อการให้ความคุ้มครองต่อการเสียชีวิตของกลุ่มผู้เอาประกันภัยก่อนหรือหลังวัยอันควร ซึ่งโดยทั่วไปแล้วคนทั่วไปจะมีอัตราการเสียชีวิตไม่แน่นอนในแต่ละช่วงอายุ นักคณิตศาสตร์ประกันภัยต้องใช้ตารางมรณะในการกำหนดเบี้ยประกันชีวิตและเงินสำรองของบริษัทให้เหมาะสม เพื่อให้บริษัทประกันชีวิตสามารถจัดการความเสี่ยงได้อย่างมีประสิทธิภาพและส่งผลดีต่อการดำเนินงานโดยรวมของบริษัท นอกจากนี้อัตรามรณะยังสามารถนำไปใช้ประโยชน์ในสายงานด้านอื่นๆ เช่น การแพทย์ สาธารณสุข ประชากรศาสตร์ เศรษฐศาสตร์ เป็นต้น

การคาดการณ์อัตรามรณะเป็นสิ่งที่ค่อนข้างยุ่งยาก เนื่องจากปัญหาในหลายๆ ด้าน อาทิเช่น ภัยพิบัติ ภัยธรรมชาติ โรคติดต่อใหม่ หรือผลกระทบจากปัจจัยอื่นๆ เช่น ความก้าวหน้าทางการแพทย์ เป็นต้น โดยการศึกษาครั้งนี้จะไม่นำปัจจัยเหล่านี้เข้ามาพิจารณา เพียงแค่พยากรณ์อัตรามรณะในอนาคตโดยอาศัยการคาดคะเนแนวโน้มสถิติของการตายของประชากรในอดีตเท่านั้น

*อายุคาดเฉลี่ยเมื่อแรกเกิดหมายถึงจำนวนปีที่คาดว่าคนจะมีชีวิตอยู่ต่อไปนับตั้งแต่เกิดจนกระทั่งตาย

ที่ผ่านมาในหลายภูมิภาคได้มีการสร้างและนำตัวแบบคณิตศาสตร์หลายตัวแบบไปใช้ในการพยากรณ์อัตรามรณะ ซึ่งในปี ค.ศ. 1992 ลีและคาร์เตอร์ [2] ได้เสนอวิธีการพยากรณ์อัตรามรณะโดยนำพื้นฐานด้านอนุกรมเวลา (Time Series Analysis) และวิธีการอย่างง่าย (Simple Method) มาประยุกต์ใช้ โดยพบว่าอัตรามรณะมีค่าลดลงอย่างต่อเนื่อง หลังจากนั้นในปี ค.ศ. 2007 ฮินแมนด์และอุลลาห์ [3] ได้สมมติให้การตายเป็นฟังก์ชั่นปรับเรียบมาประยุกต์กับวิธีของลีและคาร์เตอร์ พบว่าผลไปในทิศทางเดียวกับวิธีของลีและคาร์เตอร์ โดยค่าอัตรามรณะที่ได้จะมีความราบเรียบมากกว่า

ซึ่งหลักในการศึกษาครั้งนี้คือความเหมาะสมของตัวแบบที่ใช้พยากรณ์อัตรามรณะของประชากรไทย โดยจะศึกษาและเปรียบเทียบตัวแบบในการพยากรณ์อัตรามรณะจาก 2 วิธี ได้แก่

1. ตัวแบบลี-คาร์เตอร์ (The Lee Carter Model)
2. ตัวแบบฮินแมน-อุลลาห์ (The Hyndman-Ullah Model)

โดยพิจารณาว่าตัวแบบใดพยากรณ์อัตรามรณะไทยได้ใกล้เคียงกับค่าจริงมากกว่ากัน ซึ่งตัวแบบนั้นจะเป็นตัวแบบที่เหมาะสมที่สุด

การดำเนินการวิจัย

1. ข้อมูล

ข้อมูลจำนวนประชากรและข้อมูลการตายที่ใช้ในการศึกษาครั้งนี้เป็นข้อมูลทุติยภูมิทั้งหมด จำแนกตามอายุและเพศของประชากรไทยปี พ.ศ. 2546-พ.ศ. 2555 จากสำนักงานปลัดกระทรวง กระทรวงสาธารณสุข และกรมการปกครอง กระทรวงมหาดไทย โดยจะแบ่งข้อมูลออกเป็นข้อมูลเพศชายและข้อมูลเพศหญิงที่อายุน้อยกว่า 1 ปี จนถึงอายุมากกว่า 100 ปี ซึ่งอัตรามรณะกลางปีของประชากรที่อายุปี ในปีที่จะสามารถประมาณได้จาก

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

เมื่อ $\hat{m}_{x,t}$ แทนอัตรามรณะกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t

$D_{x,t}$ แทนจำนวนการตายของประชากรกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t

$L_{x,t}$ แทนจำนวนประชากรกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t

2. ตัวแบบในการพยากรณ์อัตรามรณะ

ในการศึกษาครั้งนี้จะทำการพยากรณ์อัตรามรณะของประชากรไทยในอีก 10 ปีข้างหน้าโดยใช้ตัวแบบลี-คาร์เตอร์ (The Lee-Carter Model) และตัวแบบฮินแมน-อุลลาห์ (The Hyndman-Ullah Model)

2.1 ตัวแบบลี-คาร์เตอร์ (The Lee-Carter Model)

ตัวแบบอัตรามรณะกลางปีรายอายุของลี-คาร์เตอร์ สามารถเขียนอยู่ในรูปของลอการิทึมของอัตรามรณะกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t ได้ดังนี้

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$$

โดยที่ $m_{x,t}$ แทนอัตรามรณะกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t

a_x แทนค่าคงที่กลางปีประชากรที่อายุ x ปี

b_x แทนอัตราเสื่อมของดัชนีเวลาที่อายุ x ปี

k_t แทนดัชนีการเปลี่ยนแปลงของเวลาในปีที่ t

$\varepsilon_{x,t}$ แทนความคลาดเคลื่อนของค่าลอกการึทึ่มอัตรามรณะกลางปีกลุ่มอายุ x ปี ในปีที่ t

จากตัวแบบลี-คาร์เตอร์ข้างต้น พบว่า a_x, b_x และ k_t เป็นค่าคงที่ที่ไม่ทราบค่าหรือค่าพารามิเตอร์ดังนั้นในการพยากรณ์อัตรามรณะกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t หรือ $m_{x,t}$ จากตัวแบบจำเป็นต้องประมาณค่าพารามิเตอร์เหล่านี้เสียก่อน

การประมาณค่าพารามิเตอร์

จากตัวแบบลี-คาร์เตอร์ ในการประมาณค่าพารามิเตอร์เพื่อให้ได้คำตอบที่แน่นอนและมีเพียงคำตอบเดียว ลี-คาร์เตอร์ จึงได้กำหนดเงื่อนไขเพิ่มเติม 2 เงื่อนไข คือ ผลรวมของสัมประสิทธิ์ของ k_t เท่ากับ 0 และผลรวมของสัมประสิทธิ์ของ b_x เท่ากับ 1 นั่นคือ

$$\sum_{t=t_1}^{T} k_t = 0 \text{ เมื่อ } t = t_1, t_2, t_3, \ldots, \mathrm{T} \quad \text{และ} \quad \sum_{x=x_1}^{X} b_x = 1 \text{ เมื่อ } x = x_1, x_2, x_3, \ldots, \mathrm{X}$$

ซึ่ง $\hat{a}_x$ คือ ค่าประมาณค่าเฉลี่ยลอการิทึมของอัตรามรณะกลางปีของประชากรในแต่ละอายุ x ปี ซึ่งประมาณได้จาก

$$\hat{a}_x = \frac{1}{T}\sum_{t=1}^{T} ln(\hat{m}_{x,t})$$

ดังนั้นจึงสามารถเขียนตัวแบบได้ใหม่เป็น

$$Z_{x,t} = ln(\hat{m}_{x,t}) - \hat{a}_x = b_x k_t$$

ในการประมาณค่า b_x และ k_t จะประมาณโดยวิธีการแยกด้วยค่าเจาะจง (The Singular Value Decomposition : SVD) ซึ่งเป็นการแก้ระบบสมการเชิงเส้นแบบหนึ่งที่นิยมใช้ในการแก้ปัญหาในการแยกเมตริกซ์ Z ซึ่งเป็นเมตริกซ์ของสมการเชิงเส้น โดยจะแยกเมตริกซ์ Z ออกเป็นสามเมตริกซ์คือ U, W และ V ซึ่งเมตริกซ์ W จะเป็นเมตริกซ์ทแยงค่าใดๆ ส่วนเมตริกซ์ U และ V จะเป็นเมตริกซ์แบบเชิงตั้งฉากปกติ (orthonormal) ดังนั้น

$$Z_{X\times T} = U_{X\times X} W_{X\times T} V'_{T\times T} \qquad \text{เมื่อ} \qquad W_{X\times T} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}_{X\times T}$$

โดยที่ Z คือ เมตริกซ์ขนาด $X \times T$

Z' คือ เมตริกซ์ขนาด $T \times X$

U คือ เมตริกซ์ตั้งฉากปกติขนาด $X \times X$ (orthonormal matrix) ของ ZZ'

V คือ เมตริกซ์ตั้งฉากปกติขนาด $T \times T$ (orthonormal matrix) ของ $Z'Z$

และ D คือ เมตริกซ์ทแยง (diagonal matrix) ที่ D_j เมื่อ $j = 1,2,\ldots,r$ คือรากที่สองของค่าเจาะจงของ Z

$$Z=\begin{bmatrix}\ln(\hat{m}_{x_1,t_1})-\hat{a}_{x_1} & \ln(\hat{m}_{x_1,t_2})-\hat{a}_{x_1} & \dots & \ln(\hat{m}_{x_1,\mathrm{T}})-\hat{a}_{x_1}\\ \ln(\hat{m}_{x_2,t_1})-\hat{a}_{x_2} & \ln(\hat{m}_{x_2,t_2})-\hat{a}_{x_2} & \dots & \ln(\hat{m}_{x_2,\mathrm{T}})-\hat{a}_{x_2}\\ \vdots & \vdots & \ddots & \vdots\\ \ln(\hat{m}_{\mathrm{X}t_1})-\hat{a}_{X} & \ln(\hat{m}_{\mathrm{X},t_2})-\hat{a}_{x_m} & \dots & \ln(\hat{m}_{\mathrm{X},T})-\hat{a}_{X}\end{bmatrix}$$

$$Z_{XxT}=\begin{bmatrix}u_{1,1} & u_{1,2} & . & . & . & u_{1,X-1} & u_{1,X}\\ u_{2,1} & u_{2,2} & . & . & . & u_{1,X-1} & u_{1,x}\\ . & & . & & & . & .\\ . & & & . & & . & .\\ . & & & & . & . & .\\ u_{X-1,1} & u_{X-1,1} & . & . & . & u_{X-1,1} & u_{X-1,X}\\ u_{X,1} & u_{X,1} & . & . & . & u_{X,1} & u_{X-1,X}\end{bmatrix}_{XxX}\begin{bmatrix}w_{1,1} & w_{1,2} & \cdots & w_{1,T-1} & w_{1,T}\\ w_{2,1} & w_{2,2} & \cdots & w_{2,T-1} & w_{2,T}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ w_{X-1,1} & w_{X-1,2} & \cdots & w_{X-1,T-1} & w_{X-1,T}\\ w_{X,1} & w_{X,1} & \cdots & w_{X,T-1} & w_{X,T}\end{bmatrix}_{XxT}\begin{bmatrix}v_{1,1} & v_{1,2} & \cdots & v_{1,T-1} & v_{1,T}\\ v_{2,1} & v_{2,2} & \cdots & v_{2,T-1} & v_{2,T}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ v_{T-1,1} & v_{T-1,2} & \cdots & v_{T-1,T-1} & v_{T-1,T}\\ v_{T,1} & v_{T,1} & \cdots & v_{T,T-1} & v_{T,T}\end{bmatrix}_{TxT}^{T}$$

ตัวประมาณ $\hat{b}_x$ คือคอลัมน์แรกของเมตริกซ์ U สำหรับทุกค่า x (อายุ) และตัวประมาณ $\hat{k}_t$ คือ $W_{1,1}V'_{t,1}$ สำหรับทุกๆ ค่า t (ปี) ดังนั้นพารามิเตอร์ที่ประมาณได้จะมี $X + T$ ตัว ($\hat{b}_x$ มี X ตัว และ $\hat{k}_t$ มี T ตัว)

เนื่องจากที่กำหนดให้อายุและเวลาเป็นอิสระจากกันจึงทำให้แถวและคอลัมน์ของเมตริกซ์ Z เป็นอิสระจากกัน ดังนั้นจึงส่งผลให้ไม่เกิดสหสัมพันธ์ (correlation) ระหว่างอายุและเวลาดังนั้น b_X จึงถูกกำหนดให้คงที่ในทุกปีสำหรับทุกอายุ x ปี และ k_t จะถูกกำหนดคงที่ทุกกลุ่มอายุในทุกปี t

การพยากรณ์ค่าพารามิเตอร์ k_t

หลังจากที่หาค่าประมาณพารามิเตอร์ a, b, และ k_t ได้แล้ว เนื่องจากเราต้องการศึกษาอัตรามรณะในอีก 10 ปีข้างหน้า ทำให้เราต้องทำการพยากรณ์ค่าประมาณของพารามิเตอร์ k_t ซึ่งมีอยู่ทั้งหมด T ค่า ในการพยากรณ์ค่าประมาณ k_t จะใช้ตัวแบบอนุกรมเวลา *ARIMA(p,d,q)* ซึ่งเป็นตัวแบบที่เหมาะสมในการสร้างแบบจำลองและพยากรณ์ข้อมูลในอนาคตเขียนอยู่ในรูปของ

$$X_t=\varphi_1 X_{t-1}+\varphi_2 X_{t-2}+\dots+\varphi_p X_{t-p}+\delta+u_t-\theta_1 u_{t-1}-\theta_2 u_{t-2}-\dots-\theta_q u_{t-q}$$

จากการศึกษาข้อมูลประเทศอื่นๆ พบว่า มีการใช้ตัวแบบ ARIMA(0,1,0) ค่อนข้างมากในระยะหลัง อีกทั้งผู้ที่เคยศึกษาข้อมูลประชากรไทยเสนอว่า ARIMA(1,0,0) มีความเหมาะสมกับข้อมูลประชากรไทยมากกว่า [4] แต่ในการศึกษาครั้งนี้มีข้อมูลเพียง 10 ปี ทำให้การใช้ตัวแบบ AR(1) ไม่เหมาะสม ดังนั้นจึงเลือกใช้การถดถอยเชิงเส้นปกติในการพยากรณ์ค่า ดังนี้

$$\hat{k}_t=\beta_0+\beta_1\hat{k}_{t-1}$$

จะได้ค่า $\hat{k}_{t+s};\ s=1,2,\dots,10$ เพื่อพยากรณ์อัตรามรณะกลางปีในอีก 10 ปีข้างหน้าต่อไป

การพยากรณ์ค่าอัตรามรณะกลางปีในอีก 10 ปี ข้างหน้า

ดังนั้นเมื่อได้ค่าพยากรณ์ $\hat{k}_{t+s}; s = 1,2,...,10$ แล้วนำค่าพยากรณ์ $\hat{k}_{t+s}; s = 1,2,...,10$ มาคำนวณหาค่าพยากรณ์อัตรามรณะกลางปีในอีก 10 ปีข้างหน้าด้วยตัวแบบลี-คาร์เตอร์ ซึ่งคำนวณหาค่าพยากรณ์อัตรามรณะกลางปีดังนี้

$$\tilde{m}_{x,t+s} = e^{\hat{a}_x + \hat{b}_x \hat{k}_{t+s}}; s = 1,2,...,10$$

2.2 ตัวแบบฮินแมนด์-อุลลาห์ (The Hyndman-Ullah Model)

ฮินแมน-อุลลาห์ (Hyndman & Ullah, 2007) ได้เสนอตัวแบบโดยการใช้เทคนิคการวิเคราะห์ใหม่ซึ่งได้ต่อยอดจากตัวแบบลี-คาร์เตอร์ โดยลอกการิทึมของอัตรามรณะกลางปีจะใช้วิธีการปรับเรียบ (Smoothing Methods) ปรับเรียบข้อมูลลอกการิทึมของอัตรามรณะ [5] โดยสังเกตค่าคลาดเคลื่อนของแต่ละอายุ ซึ่งในการพิจารณานี้จะให้อายุเป็นตัวแปรต่อเนื่องเพื่อที่จะให้ความสำคัญของความคลาดเคลื่อนจากในกรณีที่อายุไม่ต่อเนื่อง จะได้สมการค่าอัตรามรณะกลางปีปรับเรียบดังนี้

$$\mathrm{y}_t(\mathrm{x}) = f_t(x) + \sigma_t(x)\varepsilon_t$$

โดยที่ $y_t(x)$ แทนค่าลอกการึทึ่มอัตรามรณะกลางปีของประชากรอายุ x ปี ในปีที่ t

$f_t(x)$ แทนค่าลอกการึทึ่มอัตรามรณะกลางปีของประชากรอายุ x ปี ในปีที่ t ที่ปรับเรียบแล้ว

$\sigma_t(x)$ แทนค่ารบกวน (noise) ของการปรับเรียบข้อมูลลอกการิทึ่มของอัตรามรณะ

ε_t แทนความคลาดเคลื่อนของการปรับเรียบข้อมูลลอกการิทึ่มของอัตรามรณะ

วิธีการปรับเรียบข้อมูล จะใช้การปรับเรียบนอนพาราเมตริกซ์ตามงานวิจัยของฮินแมนด์และอุลลาห์ ซึ่งจะไม่มีการสร้างข้อสมมติของข้อมูล แต่เป็นการสร้างฟังก์ชั่นของตัวแปรตามกับตัวแปรอิสระเพื่อปรับค่าของตัวแปรตามให้เรียบไปตามค่าของตัวแปรอิสระ ฮินแมนด์-อุลลาห์ได้เสนอให้ใช้เทคนิคการปรับเรียบ weighted penalized spline โดยกำหนดค่าน้ำหนักในการปรับเรียบอัตรามรณะกลางปีให้เท่ากับอินเวอร์สของความแปรปรวนของการปรับเรียบข้อมูล ดังนี้

กำหนดความแปรปรวนของการปรับเรียบข้อมูลลอกการิทึ่มของอัตรามรณะ

$$\hat{\sigma}_t^2(x) \approx \frac{1 - m_{x,t}}{L_{x,t}\ m_{x,t}}$$

$$\frac{1}{\hat{\sigma}_t^2(x)} \approx \frac{L_{x,t}\ m_{x,t}}{1 - m_{x,t}}$$

จะได้ค่าน้ำหนักในการปรับเรียบ

$$w_t(x) = \frac{L_{x,t} m_{x,t}}{1 - m_{x,t}}$$

ซึ่งจะสามารถหาค่า $f_t(x_i)$ ได้จากการหาเวคเตอร์ β ซึ่งคำนวณจากการทำให้สมการต่อไปนี้มีค่าน้อยที่สุด

$$\|w(y - X\beta)\|^2 + \lambda^2 \beta^T D\beta$$

โดยที่ y แทนค่าเวคเตอร์ลอกการิทึมอัตรามรณะกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t

X แทนเมตริกซ์แสดงค่าฐานของ linear spline

λ แทนค่าพารามิเตอร์ปรับเรียบ [6]

D แทนเมตริกซ์ทแยงมุม $diag\ (0,0,1,1,...,1)_{xxx}$

w แทนน้ำหนักในการปรับเรียบหรืออินเวอร์สของความแปรปรวนของการปรับเรียบข้อมูลลอกการิทึมของอัตรามรณะ

เนื่องจากการคำนวณเวคเตอร์ β นั้นมีขั้นตอนที่ยุ่งยาก ผู้วิจัยจึงใช้ชุดคำสั่ง smooth.demogdata ในโปรแกรมคอมพิวเตอร์ R version 3.2.3 ในการคำนวณค่าเวคเตอร์ β หรือค่า $f_t(x)$ เพื่อใช้ในการพยากรณ์ตัวแบบฮินแมน-อุลลาห์ต่อไป

ซึ่งทำให้ได้สมการอัตรามรณะกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t ด้วยตัวแบบฮินแมน-อุลลาห์ดังนี้

$$f_t(x) = a(x) + \sum_{j=1}^{J} b_j(x) k_{t,j} + e_t(x)$$

โดยที่ $f_t(x)$ แทนค่าลอกการิทึมอัตรามรณะกลางปีของประชากรอายุ x ปี ในปีที่ t ที่ปรับเรียบแล้ว

$a(x)$ แทนค่าคงที่กลางปีประชากรกลุ่มอายุ x ปี

$b_j(x)$ แทนค่าพารามิเตอร์องค์ประกอบที่ j กลุ่มอายุ x ปี

$k_{t,j}$ แทนค่าพารามิเตอร์องค์ประกอบที่ j ในปีที่ t

$e_t(x)$ แทนความคลาดเคลื่อนของตัวแบบของประชากรกลุ่มอายุ x ปี ในปีที่ t

จากตัวแบบฮินแมน-อุลลาห์ข้างต้น พบว่า $a(x)$, $b_j(x)$ และ $k_{t,j}$ เป็นค่าคงที่ที่ไม่ทราบค่าหรือค่าพารามิเตอร์ ดังนั้นในการพยากรณ์อัตรามรณะกลางปีของประชากรกลุ่มอายุ x ปี ในปีที่ t หรือ $m_{x,t}$ จากตัวแบบจำเป็นต้องประมาณค่าพารามิเตอร์เหล่านี้เสียก่อน

การประมาณค่าพารามิเตอร์

การประมาณค่าพารามิเตอร์ $a(x)$ หรือค่าคงที่กลางปีประชากรกลุ่มอายุ x ปี คำนวณได้จากค่าประมาณค่าเฉลี่ยลอการิทึมของอัตรามรณะกลางปีของประชากรในแต่ละอายุ x ปี เช่นเดียวกับตัวแบบลี-คาร์เตอร์ ดังนี้

$$\hat{a}(x) = \frac{1}{T} \sum_{t=1}^{T} f_t(x_i)$$

สามารถเขียนตัวแบบใหม่ดังนี้

$$\left[f_t(x) - \hat{a}(x) \right] = b_j(x) k_{t,j}$$

ค่าพารามิเตอร์ $b_j(x)$ และ $k_{t,j}$ ที่นำมาใช้ในการพยากรณ์จะนำมาพิจารณามากกว่า 1 ชุด ซึ่งพัฒนาจากตัวแบบลี-คาร์เตอร์ที่ประมาณค่าโดยวิธีการแยกด้วยค่าเจาะจงที่พิจารณาเฉพาะองค์ประกอบแรกองค์ประกอบเดียว แต่การประมาณค่าพารามิเตอร์ของตัวแบบฮินแมนด์-อุลลาห์ครั้งนี้ จะใช้วิธีการวิเคราะห์องค์ประกอบหลักตามฟังก์ชั่น (Functional Principle Component Analysis : FPCA) ในการแยกพารามิเตอร์ $b_j(x)$ และ $k_{t,j}$

การวิเคราะห์องค์ประกอบหลักตามฟังก์ชั่น เป็นเทคนิคการลดตัวแปรโดยการสร้างเซตของตัวแปรหรือองค์ประกอบใหม่ ซึ่งตัวแปรใหม่นี้จะต้องสกัดหรือดึงรายละเอียดหรือค่าความแปรปรวนจากตัวแปรเดิมมาไว้ในตัวแปรใหม่ให้มากที่สุดดังนั้นตัวแปรใหม่ที่ 1 จึ่งถูกสร้างให้มีความแปรปรวนมากที่สุด ตัวแปรใหม่ที่ 2 มีความแปรปรวนรองลงมา ตามลำดับ และตัวแปรใหม่ทั้งหมดจะต้องไม่มีความสัมพันธ์กัน

ในการสร้างตัวแปรใหม่ $\phi_k(x)$ นั้น จะกำหนดให้ตัวแปรใหม่ที่ 1 หรือ $\phi_1(x)$ มีค่าความแปรปรวนมากที่สุด ดังนี้

$$z_{1,k} = \int \phi_k(x)\hat{f}_k^*(x)\,dx$$

กำหนดให้ $z_{1,k}$ แทน ค่าคะแนนของตัวแปรใหม่

$\phi_k(x)$ แทน ตัวแปรใหม่ ที่ k

$\hat{s}_t^*(x)$ แทน เซตของค่าสัมประสิทธิ์บนเส้นโค้ง

ซึ่งทำให้ $\int \phi_k^2(x) = 1$ และ $\int \phi_k(x)\phi_{k-1}(x)dx = 0$ เมื่อ $k \geq 2$

โดยที่ $k = 1,\ldots, \mathrm{K} < \mathrm{n}$ เมื่อ n แทนจำนวนตัวแปรเดิม

และ $\hat{s}_k^*(x) = \hat{s}_{k-1}^*(x) - \mathrm{z}_{1,k}\,\phi_k(x)$

ทำการทำซ้ำเพื่อหาตัวแปรใหม่ทั้งหมดที่ต้องการ ซึ่งพารามิเตอร์ $b_j(x)$ ก็คือค่า $\phi_k(x)$ ส่วนพารามิเตอร์ $k_{\mathrm{t},j}$ คือค่า $z_{1,k}$ นั่นเอง การวิเคราะห์องค์ประกอบหลักตามฟังก์ชั่นทำให้ได้ชุดพารามิเตอร์ของ $\{b_1(x), b_2(x),\ldots, b_j(x)\}$ และ $\{k_{t,1}, k_{t,2},\ldots, k_{t,j}\}$

ซึ่งการพิจารณาจำนวนตัวแปรใหม่ที่เหมาะสมนั้นจะใช้ตัวแปรเพียงไม่กี่ตัว ที่มีความแปรปรวนสูงโดยเฉพาะอย่างยิ่งตัวแปรใหม่ตัวแรกๆ ส่วนตัวแปรใหม่ท้ายๆ นั้นจะมีความแปรปรวนต่ำ อีกทั้งวัตถุประสงค์ของวิธีนี้มีเพื่อลดจำนวนตัวแปรลง ดังนั้นตัวแปรใหม่จึงควรจะมีจำนวนน้อยกว่าตัวแปรเดิม

การพยากรณ์ค่าพารามิเตอร์

หลังจากที่หาค่าประมาณพารามิเตอร์ $a(x)$, $b_j(x)$, และ $k_{\mathrm{t},j}$ ได้แล้ว เนื่องจากเราต้องการศึกษาอัตรามรณะในอีก 10 ปีข้างหน้า ทำให้เราต้องทำการพยากรณ์ค่าประมาณของพารามิเตอร์ $k_{\mathrm{t},j}$ ซึ่งมีอยู่ทั้งหมด T ค่า ในการพยากรณ์ค่าประมาณ $k_{\mathrm{t},j}$ จะใช้ตัวแบบอนุกรมเวลา (Time Series Model) ในที่นี้จะพยากรณ์ค่า $k_{\mathrm{t},j}$ โดยใช้วิธีเอกซ์โพเนนเซียลสมูทติ้ง (Exponential smoothing) [7] ตามงานวิจัยของฮินแมนด์และอุลลาห์ ดังนี้

$$\hat{k}_{t+1} = l_t$$

$$l_t = \alpha\hat{k}_t + (1-\alpha)l_{t-1}$$

เมื่อ α คือค่าสัมประสิทธิ์การปรับเรียบหรือตัวแปรถ่วงน้ำหนักมีค่าได้ตั้งแต่ 0 ถึง 1 กำหนดโดยวิธีลองผิดลองถูก ซึ่งการพิจารณาว่าควรใช้ α เท่าไหร่นั้นจะพิจารณาจากความคลาดเคลื่อนการพยากรณ์ที่มีค่าน้อยที่สุดจะได้ค่า $\hat{k}_{t+s}\,; s$ = 1,2...,10 เพื่อพยากรณ์อัตรามรณะกลางปีในอีก 10 ปีข้างหน้าต่อไป

การพยากรณ์ค่าอัตรามรณะกลางปีในอีก 10 ปี ข้างหน้า

ดังนั้นเมื่อได้ค่าพยากรณ์ $\hat{k}_{t,j+s}; s = 1,2...,10$ แล้วนำค่าพยากรณ์ $\hat{k}_{t+s}; s = 1,2...,10$ มาคำนวณหาค่าพยากรณ์อัตรามรณะกลางปีในอีก 10 ปีข้างหน้าด้วยตัวแบบฮินแมน-อุลลาห์ ซึ่งคำนวณหาค่าพยากรณ์อัตรามรณะกลางปีดังนี้

$$\tilde{m}_{x,t+s} = e^{\hat{a}(x)+\sum_{j=1}^{J}\hat{k}_{t+s,j}\hat{b}_j(x)} \quad ; s = 1,2,...,10$$

3. การพิจารณาความแม่นยำของวิธีการพยากรณ์ค่าอัตรามรณะแต่ละตัวแบบ

พิจารณาการเปรียบเทียบค่าอัตรามรณะว่าตัวแบบใดเหมาะสมที่สุดสำหรับข้อมูลชุดนี้ ในรูปของวิธีการหาค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ยหรือ MAPE เป็นค่าที่ใช้วัดความคลาดเคลื่อนของการประมาณตัวแบบ โดยวัดความคลาดเคลื่อนจากค่าจริงและค่าที่พยากรณ์ได้ ซึ่งค่าพยากรณ์ที่มีความใกล้เคียงค่าจริงมากที่สุดหรือทำให้เกิดค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ยน้อยที่สุด จะเป็นค่าที่เหมาะสม โดยจะหาค่าได้ดังนี้

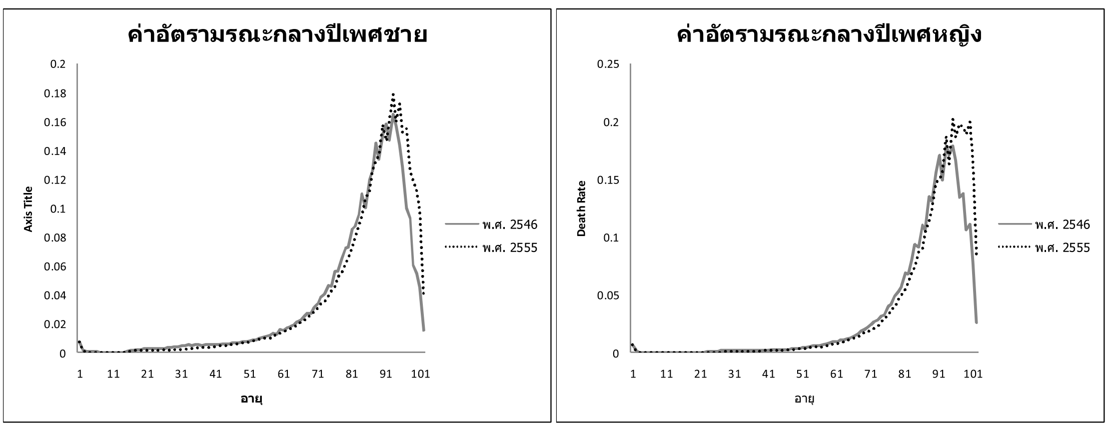
$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{A_t - F_t}{A_t}\right| \times 100$$

โดย A_t แทน ค่าจากข้อมูลจริง

F_t แทน ค่าที่ได้จากการพยากรณ์

ผลการวิจัย

จากข้อมูลจำนวนประชากรและข้อมูลการตายจำแนกตามอายุและเพศของประชากรไทยปี พ.ศ. 2546-พ.ศ. 2555 จะเห็นได้ว่าค่าอัตรามรณะกลางปีสำหรับข้อมูลเพศชายและเพศหญิงมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่องในลักษณะของเส้นตรงตั้งแต่อายุแรกเกิดถึงช่วงอายุ 70 ปี หลังจากนั้นค่าอัตรามรณะกลางปีมีแนวโน้มเพิ่มขึ้นอย่างรวดเร็วส่วนในช่วงอายุตั้งแต่ 90 ปีขึ้นไป ซึ่งจะมีค่าอัตรามรณะกลางปีมีแนวโน้มลดลงอย่างรวดเร็วดังรูปที่ 1



รูปที่ 1 ค่าอัตรามรณะกลางปีของประชากรเพศชายและเพศหญิง พ.ศ. 2546 และ พ.ศ. 2555

จากรูปที่ 1 ผู้วิจัยได้เลือกเปรียบเทียบระหว่างค่าอัตรามรณะกลางปีระหว่าง พ.ศ. 2546 และ พ.ศ. 2555 จะเห็นได้ว่าทั้งเพศชายและเพศหญิงอัตรามรณะกลางปี พ.ศ. 2555 มีค่าต่ำกว่าอัตรามรณะกลางปี พ.ศ. 2546 ตั้งแต่อายุแรกเกิดจนถึงอายุประมาณ 90 ปี จึงค่อยๆ เพิ่มสูงขึ้น จึงสรุปได้ว่าเมื่อเวลาผ่านไป อัตรามรณะในอนาคตจะต่ำกว่าในอดีต ซึ่งตรงกับสมมติฐานที่ว่าประชากรไทยทั้งเพศชายและเพศหญิงมีอายุยืนขึ้นกว่าในอดีต

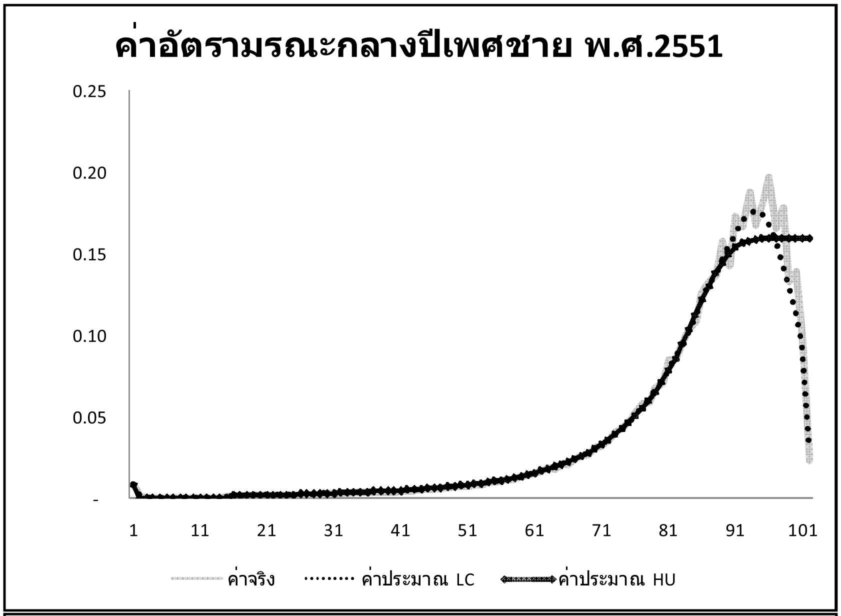
1. ผลการประมาณค่าอัตรามรณะกลางปีจากทั้ง 2 ตัวแบบ

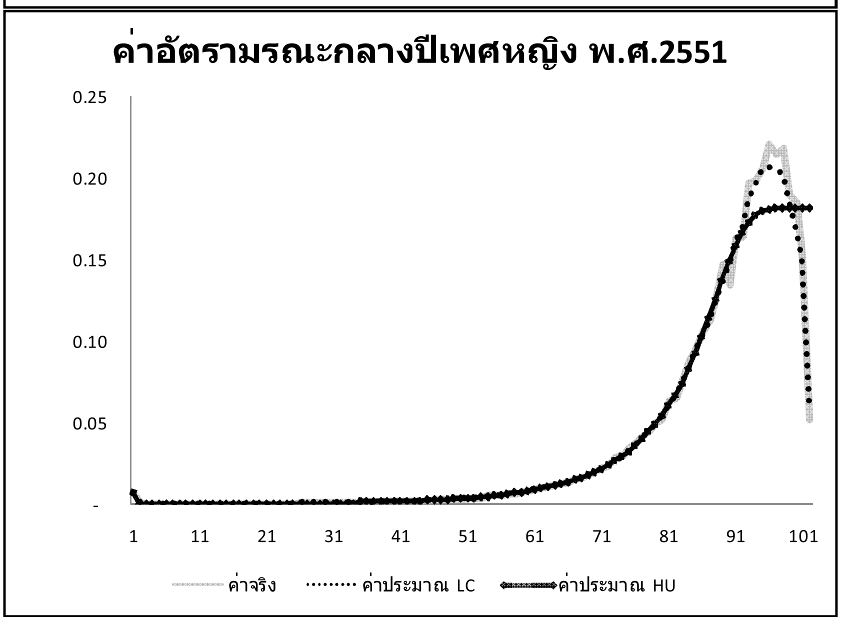
จากการศึกษาจะได้ผลการประมาณค่าอัตรามรณะกลางปีของตัวแบบ-ลีคาร์เตอร์และตัวแบบฮินแมนด์-อุลลาห์ของทุกปี ตั้งแต่ พ.ศ. 2546 ถึง พ.ศ. 2555 ในงานวิจัยนี้ขอยกตัวอย่างปี พ.ศ. 2551 เพื่อแสดงการเปรียบเทียบค่าอัตรามรณะกลางปีจริงกับค่าประมาณอัตรามรณะกลางปีนำเสนอในตารางที่ 2 ดังนี้

ตารางที่ 2 ค่าอัตรามรณะกลางปีจริงและค่าประมาณด้วยตัวแบบลี-คาร์เตอร์และตัวแบบฮินแมนด์-อุลลาห์ของประชากรเพศชายและเพศหญิง พ.ศ. 2551

อายุ (ปี)	เพศชาย			เพศหญิง		
	ค่าจริง $\hat{m}_{x,t}$	ค่าประมาณ $\tilde{m}_{x,t}$		ค่าจริง $m_{x,t}$	ค่าประมาณ $\tilde{m}_{x,t}$	
		LC	HU		LC	HU
0	0.008215	0.008134	0.008086	0.007127	0.006814	0.006820
1	0.001074	0.001092	0.001065	0.000798	0.000933	0.000935
2	0.000793	0.000804	0.000779	0.000511	0.000548	0.000557
3	0.000615	0.000639	0.000628	0.000448	0.000446	0.000451
4	0.000600	0.000590	0.000587	0.000357	0.000405	0.000407
5	0.000664	0.000593	0.000581	0.000408	0.000395	0.000388
6	0.000592	0.000605	0.000563	0.000351	0.000363	0.000370
7	0.000477	0.000515	0.000523	0.000348	0.000347	0.000353
8	0.000516	0.000486	0.000470	0.000310	0.000348	0.000343
9	0.000465	0.000444	0.000426	0.000333	0.000340	0.000337
10	0.000418	0.000416	0.000408	0.000350	0.000335	0.000333
.	.	.	.	.	.	.
.	.	.	.	.	.	.
50	0.007833	0.008042	0.007989	0.003698	0.003771	0.003743
.	.	.	.	.	.	.
.	.	.	.	.	.	.
90	0.172848	0.160085	0.153349	0.163448	0.159708	0.158037
91	0.166688	0.170240	0.155958	0.163889	0.169361	0.166080
92	0.187933	0.172880	0.157567	0.196905	0.185699	0.172221
93	0.167106	0.176485	0.158441	0.198991	0.195240	0.176501
94	0.180982	0.174220	0.158836	0.207002	0.205675	0.179174
95	0.196458	0.168298	0.158966	0.220894	0.206316	0.180628
96	0.165358	0.158810	0.158986	0.214608	0.204689	0.181278
97	0.177606	0.142674	0.158994	0.218315	0.202293	0.181495
98	0.132497	0.130524	0.159039	0.190114	0.186227	0.181547
99	0.138715	0.113346	0.159132	0.185516	0.168381	0.181586
100	0.099517	0.095636	0.159262	0.152595	0.150181	0.181668
มากกว่า 100	0.023604	0.033148	0.159410	0.051838	0.062700	0.181781

จากตารางที่ 2 จะเห็นได้ว่าค่าอัตรามรณะกลางปีจากการประมาณด้วยตัวแบบลี-คาร์เตอร์และตัวแบบฮินแมนด์-อุลลาห์เป็นไปในทิศทางเดียวกับค่าอัตรามรณะกลางปีจริง แต่ค่าอัตรามรณะจะมีความแตกต่างกันอย่างเห็นได้ชัดตั้งแต่อายุ 90 ปีขึ้นไปในทั้งประชากรเพศชายและเพศหญิง





รูปที่ 2 เปรียบเทียบค่าอัตรามรณะกลางปีจริงและค่าประมาณด้วยตัวแบบลี-คาร์เตอร์และตัวแบบฮินแมนด์-อุลลาห์ของประชากรเพศชายและเพศหญิง พ.ศ. 2551

เมื่อนำข้อมูลจากตารางที่ 2 มาแสดงผลด้วยกราฟเส้นดังรูปที่ 2 จะเห็นได้ชัดเจนว่าค่าอัตรามรณะกลางปีจากการประมาณด้วยตัวแบบลี-คาร์เตอร์และตัวแบบฮินแมนด์-อุลลาห์ เป็นไปในทิศทางเดียวกับค่าอัตรามรณะกลางปีจริงตั้งแต่อายุที่ 0 ปีถึงอายุ 89 ปี ส่วนตั้งแต่ที่อายุ 90 ปีขึ้นไป ตัวแบบฮินแมนด์-อุลลาห์จะให้ค่าอัตรามรณะกลางปีแตกต่างจากค่าจริงอย่างเห็นได้ชัด ส่วนตัวแบบลี-คาร์เตอร์นั้นให้ค่าอัตรามรณะกลางปีแตกต่างจากค่าจริงเล็กน้อย

2. ผลการพิจารณาความแม่นยำของตัวแบบ

ผลการประมาณค่าอัตรามรณะกลางปีข้างต้นนั้นไม่สามารถสรุปได้ว่าตัวแบบใดเหมาะสมกับการพยากรณ์อัตรามรณะของประชากรไทยมากที่สุด จึงต้องพิจารณาความแม่นยำของตัวแบบโดยเปรียบเทียบค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ยหรือ MAPE

ตารางที่ 3 เปรียบเทียบค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ยทุกอายุของประชากรเพศชายและเพศหญิงด้วยตัวแบบลี-คาร์เตอร์และตัวแบบฮินแมน-อุลลาห์

Average MAPE	**Male**		**Female**	
	LC	HU	LC	HU
	3.9065	2.3986	4.2039	2.7984

จากตารางที่ 3 ซึ่งเปรียบเทียบค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ยทุกอายุของประชากรเพศชายและเพศหญิงด้วยตัวแบบลี-คาร์เตอร์และตัวแบบฮินแมน-อุลลาห์ ค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ยของค่าอัตรามรณะเพศชายและเพศหญิงของการพยากรณ์ด้วยตัวแบบฮินแมน-อุลลาห์มีค่าต่ำกว่าการประมาณด้วยตัวแบบลี-คาร์เตอร์

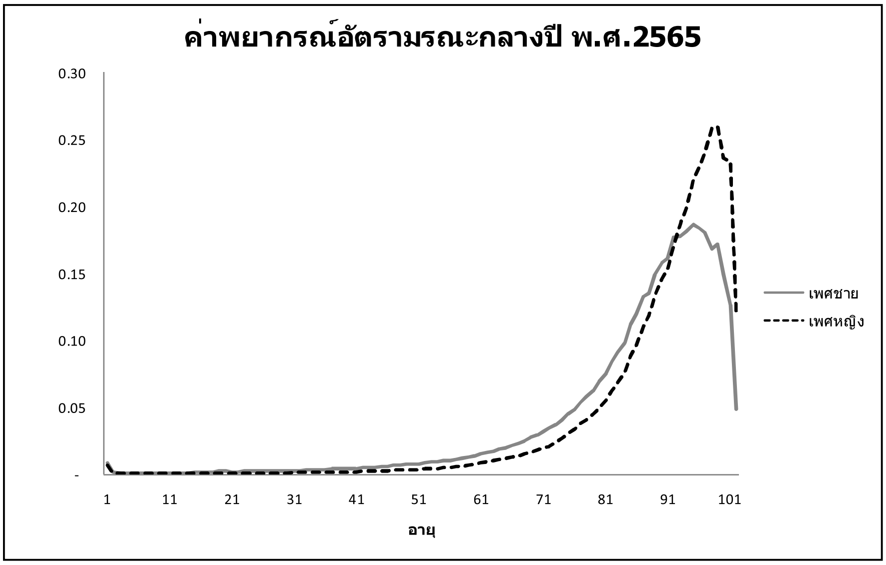
แต่จากการเปรียบเทียบค่าอัตรามรณะกลางปีจริงและค่าประมาณด้วยตัวแบบลี-คาร์เตอร์และตัวแบบฮินแมนด์-อุลลาห์ ของประชากรเพศชายและเพศหญิงของทุกปี ตั้งแต่ พ.ศ. 2546 ถึง พ.ศ. 2555 กราฟของค่าประมาณด้วยตัวแบบลี-คาร์เตอร์ มีแนวโน้มไปในทิศทางเดียวกันกับค่าอัตรามรณะกลางปีจริงมากกว่าค่าประมาณด้วยตัวแบบฮินแมน-อุลลาห์ ซึ่งเห็นได้จากตัวอย่างการเปรียบเทียบค่าอัตรามรณะ พ.ศ. 2551 ในรูปที่ 2 ตั้งแต่อายุ 90 ปีขึ้นไปนั้นตัวแบบฮินแมน-อุลลาห์ มีแนวโน้มแตกต่างจากค่าอัตรามรณะกลางปีจริงอย่างเห็นได้ชัดซึ่งแม้ว่าค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ยของการพยากรณ์ตัวแบบฮินแมน-อุลลาห์จะมีค่าต่ำกว่า แต่เมื่อพิจารณาแนวโน้มของกราฟแล้วผู้วิจัยจึงเลือกตัวแบบลี-คาร์เตอร์ในการพยากรณ์ค่าอัตรามรณะไทยในอีก 10 ปีข้างหน้า

3. ผลการพยากรณ์อัตรามรณะกลางปีของประชากรไทยในอีก 10 ปีข้างหน้า

จากผลการพิจารณาความแม่นยำของตัวแบบสรุปได้ว่า ในการศึกษาครั้งนี้จะพยากรณ์อัตรามรณะกลางปีของประชากรไทยในอีก 10 ปีข้างหน้าด้วยตัวแบบลี-คาร์เตอร์ ซึ่งสามารถพยากรณ์อัตรามรณะกลางปีได้ตั้งแต่ พ.ศ. 2555 ถึง พ.ศ. 2565 งานวิจัยนี้จะขอยกตัวอย่างผลการพยากรณ์อัตรามรณะกลางปีของประชากรไทยในปี พ.ศ. 2565 ซึ่งได้ผลการพยากรณ์ดังตารางต่อไปนี้

ตารางที่ 4 ค่าพยากรณ์อัตรามรณะกลางปีประชากรเพศชายและเพศหญิงในปี พ.ศ. 2565

อายุ (ปี)	เพศชาย	เพศหญิง	อายุ (ปี)	เพศชาย	เพศหญิง	อายุ (ปี)	เพศชาย	เพศหญิง
0	0.007876	0.006388	34	0.003019	0.001029	68	0.027108	0.015762
1	0.000866	0.000612	35	0.003257	0.001129	69	0.029265	0.017505
2	0.000696	0.000414	36	0.003472	0.001171	70	0.031591	0.019415
3	0.000556	0.000336	37	0.003601	0.001257	71	0.034191	0.020555
4	0.000486	0.000321	38	0.003754	0.001358	72	0.036728	0.023734
5	0.000496	0.000307	39	0.004154	0.001459	73	0.040384	0.026280
6	0.000533	0.000270	40	0.004292	0.001538	74	0.044474	0.030225
7	0.000431	0.000244	41	0.004563	0.001656	75	0.048166	0.032933
8	0.000420	0.000258	42	0.004718	0.001760	76	0.053351	0.037409
9	0.000398	0.000274	43	0.005048	0.001956	77	0.057985	0.040492
10	0.000390	0.000270	44	0.005316	0.002222	78	0.062402	0.044903
11	0.000393	0.000259	45	0.005610	0.002252	79	0.069255	0.049099
12	0.000471	0.000324	46	0.006096	0.002459	80	0.075129	0.054713
13	0.000661	0.000366	47	0.006565	0.002656	81	0.083924	0.061795
14	0.000891	0.000424	48	0.006992	0.002782	82	0.091083	0.067814
15	0.001286	0.000484	49	0.007256	0.003041	83	0.098103	0.076412
16	0.001462	0.000475	50	0.007781	0.003253	84	0.111388	0.088457
17	0.001637	0.000485	51	0.008196	0.003518	85	0.119545	0.095645
18	0.001714	0.000504	52	0.008764	0.003734	86	0.132883	0.110116
19	0.001699	0.000496	53	0.009084	0.004113	87	0.135675	0.118494
20	0.001615	0.000508	54	0.009790	0.004425	88	0.149218	0.133562
21	0.001650	0.000517	55	0.010264	0.004794	89	0.158130	0.145338
22	0.001800	0.000492	56	0.010941	0.005128	90	0.161469	0.153417
23	0.001750	0.000476	57	0.011794	0.005840	91	0.177241	0.170324
24	0.001830	0.000500	58	0.012819	0.006090	92	0.178053	0.187122
25	0.001872	0.000499	59	0.013647	0.006885	93	0.181549	0.199168
26	0.001889	0.000535	60	0.014827	0.007800	94	0.186577	0.219746
27	0.001965	0.000598	61	0.016212	0.008559	95	0.184035	0.229536
28	0.002112	0.000586	62	0.017254	0.009599	96	0.180216	0.240809
29	0.002207	0.000611	63	0.018253	0.010690	97	0.168451	0.259261
30	0.002270	0.000695	64	0.019761	0.011745	98	0.171888	0.259751
31	0.002528	0.000768	65	0.021462	0.012550	99	0.149307	0.236652
32	0.002638	0.000870	66	0.023087	0.013697	100	0.126498	0.233618
33	0.002833	0.000962	67	0.024651	0.014823	> 100	0.048398	0.118987



รูปที่ 3 เปรียบเทียบค่าพยากรณ์อัตรามรณะกลางปีประชากรเพศชายและเพศหญิงในปี พ.ศ. 2565

ผลการพยากรณ์อัตรามรณะกลางปีของประชากรไทยในอีก 10 ปีข้างหน้าด้วยตัวแบบลี-คาร์เตอร์ ผู้วิจัยยกตัวอย่างผลการพยากรณ์อัตรามรณะกลางปีของประชากรไทยในปี พ.ศ. 2565 มาแสดงผลดังตารางที่ 4 และรูปที่ 3 นั้น พบว่าค่าอัตรามรณะกลางปีเพศชายที่อายุ 0 ปีหรือทารกแรกเกิดนั้นมีค่าต่ำจนถึงอายุ 40 ปี และค่อยๆ สูงขึ้นเรื่อยจนถึงช่วงอายุ 70 ปี หลังจากนั้นสูงขึ้นอย่างรวดเร็วจนกระทั่งอายุ 90 ปี ส่วนของเพศหญิงก็มีแนวโน้มสูงขึ้นเช่นกันค่าอัตรามรณะกลางปีเพศหญิงที่อายุ 0 ปีหรือทารกแรกเกิดนั้นมีค่าต่ำจนถึงอายุ 50 ปีและค่อยๆ สูงขึ้นเรื่อยจนถึงช่วงอายุ 96 ปีจากนั้นค่าอัตรามรณะกลางปีจะลดลงอย่างรวดเร็วอย่างต่อเนื่องจนถึงช่วงอายุสุดท้าย

ซึ่งค่าอัตรามรณะกลางปีของเพศหญิงจะมีค่าต่ำกว่าเพศชายตั้งแต่อายุแรกเกิดจนถึงอายุ 91 ปี แต่เมื่อถึงอายุ 92 ปี ค่าอัตรามรณะกลางปีของเพศหญิงจะสูงกว่าเพศชายหมายความว่าตั้งแต่อายุแรกเกิดจนถึง 91 ปีนั้น ประชากรเพศชายมีแนวโน้มที่จะเสียชีวิตมากกว่าประชากรเพศหญิง แต่เมื่อถึงอายุ 92 ปี ประชากรเพศหญิงจะมีแนวโน้มการเสียชีวิตมากกว่าประชากรเพศชาย

จากผลของการพยากรณ์ค่าอัตรามรณะกลางปีที่ได้จากการพยากรณ์อัตรามรณะด้วยตัวแบบลี-คาร์-เตอร์ จะสามารถนำไปสร้างตารางมรณะไทยเพื่อใช้ในการคำนวณเบี้ยประกันภัยและมูลค่าต่างๆ ของบริษัทประกันชีวิตต่อไป

สรุปและวิจารณ์ผลการวิจัย

การศึกษาตัวแบบพยากรณ์อัตรามรณะของไทยในครั้งนี้ ได้ทำการศึกษาตัวแบบของลี-คาร์เตอร์ (Lee-Carter Model) และตัวแบบฮินแมน-อุลลาห์ (Hyndman-Ullah Model) โดยอาศัยข้อมูลจำนวนประชากรจากกรมการปกครอง กระทรวงมหาดไทยและจำนวนประชากรตายจากสำนักงานปลัดกระทรวงกระทรวงสาธารณสุขในปี พ.ศ. 2546-พ.ศ. 2555 จำแนกตามเพศและอายุ ปรากฏว่าตัวแบบฮินแมน-อุลลาห์ ให้ค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ยน้อยที่สุด แต่กราฟของค่าประมาณด้วยตัว

แบบลี-คาร์เตอร์ มีแนวโน้มไปในทิศทางเดียวกันกับค่าอัตรามรณะกลางปีจริงมากกว่าผู้วิจัยจึงเลือกตัวแบบลี-คาร์เตอร์ในการพยากรณ์อัตรามรณะของไทยในอีก 10 ปีข้างหน้า ซึ่งแตกต่างกับงานวิจัยของฟางและฮาร์เดิล [8] ที่พยากรณ์อัตรามรณะของประชากรจีนและญี่ปุ่น ได้สรุปผลการศึกษาโดยเลือกตัวแบบฮินแมน-อุลลาห์จากการพิจารณาค่าคลาดเคลื่อนของแต่ละตัวแบบ

การศึกษาในครั้งนี้ได้สรุปผลโดยพิจารณาค่าร้อยละค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ยคู่กับแนวโน้มของกราฟทั้งสองตัวแบบ ในการศึกษาครั้งถัดไปอาจเสนอการพิจารณาความแม่นยำของวิธีการพยากรณ์อื่นๆ และการสรุปผลในครั้งนี้ได้ใช้ตัวแบบเพียงตัวแบบเดียวพยากรณ์อัตรามรณะในทุกอายุ การศึกษาครั้งต่อไปควรแบ่งผลการทดลองเป็นกลุ่มอายุเด็ก วัยรุ่ย วัยกลางคน และผู้สูงอายุ เป็นต้น อีกทั้งข้อมูลที่ใช้ในการศึกษามาจากระบบการจดทะเบียนและมักจะมีปัญหาเรื่องการตกจดทะเบียนโดยเฉพาะอย่างยิ่งในกลุ่มผู้สูงอายุ ซึ่งในการศึกษาครั้งนี้ไม่ได้ปรับข้อมูลที่ตกจดทะเบียน การศึกษาในครั้งต่อไปจึงควรพิจารณาในประเด็นนี้ด้วย

เอกสารอ้างอิง

1. สถาบันวิจัยประชากรและสังคม มหาวิทยาลัยมหิดล. 2557. *สารประชากร มหาวิทยาลัยมหิดล.*
2. Lee, R. D., and Carter, L. 1992. Modeling and Forecasting US Mortality. *Journal of the American Statistical Association, 87*(419): 659-671.
3. Hyndman, R. J., and Ullah, M. S. 2007. Robust Forecasting of Mortality and Fertility Rate: A Functional Data Approach, *Computational Statistics & Data Analysis, 51*(10): 4942-4956.
4. ณิชา ราชฤทธิ์ และ สุวาณี สุรเสียงสังข์. 2549. ตัวแบบพยากรณ์อัตรามรณะของประชากรไทยโดยวิธีของลีและคาร์เตอร์. *วารสารประชากรศาสตร์, 22*(2): 25-42.
5. Wood, S. N. 2003. Thin plate regression splines. *Journal of the Royal Statistical Society Series B, 65*(1): 95-114.
6. Ruppert, D., Wand, M. P., and Carroll, R. J. 2003. *Semiparametric regression.* Cambridge university press.
7. Hyndman, R. J., and Khandakar Y. 2008. Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software,* 27(3).
8. Fang, L., and Härdle, W. 2015. *Stochastic Population Analysis: A Functional Data Approach.* Humboldt University, Collaborative Research Center 649.

ได้รับบทความวันที่ 2 มีนาคม 2559
ยอมรับตีพิมพ์วันที่ 20 เมษายน 2559