หัวข้อพิเศษทางการประกันภัย

SPECIAL TOPICS IN INSURANCE

สาขาวิชาการประกันภัย

ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี

จุฬาลงกรณ์มหาวิทยาลัย

เรื่อง

การเปรียบเทียบการประมาณค่าพารามิเตอร์ของตัวแบบพยากรณ์อัตรามรณะของประชากรไทยโดยวิธีของลี-คาร์เตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์

Comparing Parameter Estimation of Thai Mortality Rate Modeling and Forecasting with Lee-Carter between Maximum Likelihood Method and Bayesian Method

รหัสโครงงาน 18

โดย

6042019626 นายจิรัฐิติ ผ่องเกษม

6042062526 นายปัณณวิชญ์ ขอพลอยกลาง

ปีการศึกษา 2563

หัวข้อพิเศษทางการประกันภัย

SPECIAL TOPICS IN INSURANCE

เรื่อง

การเปรียบเทียบการประมาณค่าพารามิเตอร์ของตัวแบบพยากรณ์อัตรามรณะของประชากรไทยโดยวิธีของลี-คาร์เตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์

Comparing Parameter Estimation of Thai Mortality Rate Modeling and Forecasting with Lee-Carter between Maximum Likelihood Method and Bayesian Method

รหัสโครงงาน 18

โดย

6042019626 นายจิรัฐิติ ผ่องเกษม

6042062526 นายปัณณวิชญ์ ขอพลอยกลาง

ปีการศึกษา 2563

ข้าพเจ้าขอรับรองว่า การจัดทำโครงงานฉบับนี้ ข้าพเจ้าไม่ได้คัดลอกเอกสารใด ๆ ไม่ว่าจะเป็นบางส่วนหรือทั้งหมด หากตรวจพบข้อความใด ๆ ที่สามารถสืบค้นหรือได้รับแจ้งว่าโครงงานฉบับนี้คัดลอกทั้งหมดหรือบางส่วน ข้าพเจ้ายอมรับในความผิดทั้งทางแพ่งและอาญาแต่เพียงผู้เดียว

ลงชื่อนิสิตคนที่ 1

(นายจิรัฐิติ ผ่องเกษม)

วันที่

ลงชื่อนิสิตคนที่ 2

(นายปัณณวิชญ์ ขอพลอยกลาง)

วันที่

คำนำ

โครงงานเรื่อง การเปรียบเทียบการประมาณค่าพารามิเตอร์ของตัวแบบพยากรณ์อัตรามรณะของ ประชากรไทยโดยวิธีของลี-คาร์เตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์ เป็นส่วนหนึ่งของวิชา หัวข้อพิเศษทางประกันภัย (Special topics in insurance) รหัสวิชา 2603494 ปีการศึกษา 2563 โดยมีจุดประสงค์เพื่อปรียบเทียบการประมาณค่าพารามิเตอร์ของตัวแบบพยากรณ์อัตรามรณะโดยวิธีของลี-คาร์เตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์ โดยใช้ข้อมูลจำนวนประชากรและข้อมูลจำนวนการตายของประเทศไทยตั้งปี พ.ศ. 2541-2563 ซึ่งเนื้อหาภายในโครงงานประกอบด้วย วิธีการประมาณโดยใช้วิธีภาวะน่าจะเป็นสูงสุด วิธีการประมาณโดยใช้วิธีแบบเบย์ซึ่งประกอบด้วยการทำ การสุ่มแบบมอนติคาร์โล ลูกโซ่มาร์คอฟ และการพยากรณ์เพื่อวัดผลความแม่นยำ

คณะผู้จัดทำ หวังว่าการทำโครงงานครั้งนี้จะแสดงให้เห็นถึงประสิทธิภาพของการประมาณโดยวิธีต่าง ๆ รวมถึงแสดงให้เห็นแนวคิดการประมาณโดยใช้วิธีแบบเบย์กับข้อมูลในไทย คณะผู้จัดทำหวังว่าโครงงานนี้จะเป็นประโยชน์สำหรับผู้ที่สนใจศึกษาเพื่อนความรู้ไปต่อยอดได้ หากมีข้อเสนอแนะประการใด ผู้เขียนขอรับไว้ด้วยคำขอบพระคุณ

จิรัฐิติ ผ่องเกษม

ปัณณวิชญ์ ขอพลอยกลาง

กิตติกรรมประกาศ

โครงงานเรื่อง การเปรียบเทียบการประมาณค่าพารามิเตอร์ของตัวแบบพยากรณ์อัตรามรณะของ ประชากรไทยโดยวิธีของลี-คาร์เตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์ สามารถผ่านพ้นไปได้ด้วยดีเพราะได้รับการช่วยเหลือจากหลาย ๆ ท่าน

ผู้ศึกษาขอกราบขอบพระคุณ รศ.ดร. สุวาณี สุรเสียงสังข์ ที่ให้ข้อมูลจำนวนประชากรและจำนวนการตายของประเทศไทยตั้งแต่ปีพ.ศ.2540-2559 รวมถึงให้คำแนะนำไขข้อสงสัยใน งานวิจัยหลาย ๆ ฉบับที่ได้ทำการทบทวนวรรณกรรม

ขอกราบขอบพระคุณ ผศ.ดร. อนุภาพ สมบูรณ์สวัสดี ในการให้คำแนะนำในเรื่องวิธีการเขียนโปรแกรมอาร์เพื่อคำนวนใช้ในการประมาณค่าพารามิเตอร์ด้วยวิธีแบบเบย์ รวมถึงคำแนะนำการเลือกและปรับใช้การแจกแจงแบบต่าง ๆ ในการประมาณค่าพารามิเตอร์ด้วยวิธีแบบเบย์ และคำแนะนำในการแก้ปัญหาการเขียนโปรแกรมที่ผิดพลาด

ขอกราบขอบพระคุณ อ.ดร. ปุณฑวิกา นาคา ที่ให้คำแนะนำในเรื่องการเลือกใช้ ไลบรารี่สำเร็จรูปในการประมาณค่าพารามิเตอร์โดยใช้วิธีภาวะน่าจะเป็นสูงสุด เพื่อความสะดวกค่อการคำนวณ

สุดท้ายขอบขอบพระคุณเพื่อน ๆ พี่น้อง ครอบครัว ทุกท่านที่คอยอยู่เคียงข้างช่วยเหลือให้คำแนะนำให้กำลังใจกันและกันโดยตลอดเสมอมา

สารบัญ

สารบัญตาราง

สารบัญรูปภาพ

บทคัดย่อ

ชื่อโครงการ : การเปรียบเทียบการประมาณค่าพารามิเตอร์ของตัวแบบพยากรณ์อัตรามรณะของประชากรไทยโดยวิธีของลี-คาร์เตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์

รหัสโครงงาน : 18

ชื่อนิสิต : 1. นาย จิรัฐิติ ผ่องเกษม รหัสนิสิต 6042019626

2. นาย ปัณณวิชญ์ ขอพลอยกลาง รหัสนิสิต 6042062526

จำนวนหน้า : 42 หน้า

วัตถุประสงค์ของโครงงานชิ้นนี้มีเพื่อเปรียบเทียบประสิทธิภาพการประมาณค่าพารามิเตอร์ของตัวแบบอัตรามรณะลีคาร์-คาร์เตอร์ ระหว่างวิธีแบบภาวะน่าจะเป็นสูงสุดกับวิธีแบบเบย์ โดยการนำค่าที่ได้จากการประมาณการทั้งสองแบบ มาทำการพยการณ์อัตรามรณะ และใช้ค่า MAPE ในการเปรียบเทียบระหว่างค่าที่ได้จากการพยากรณ์และค่าที่แท้จริง

ผลการศึกษาพบว่าในการประมาณค่าพารามิเตอร์ทั้งวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์นั้นได้ตัวค่าพารามิเตอร์ไปในทิศทางเดียวกัน ค่าคงที่ของประชากรกลางปี (α) มีลักษณะสูงช่วงอายุ 0 ถึง 5 ปีและลดลงจนถึงอายุ 20 ปี หลังจากนั้นจะมีค่าเพิ่มขึ้นเรื่อย ๆ ในขณะที่ค่าอัตราเสื่อมของเวลา (β) มีลักษณะที่แกว่งไม่แน่นอน สุดท้ายค่า ดัชนีเวลา (κ) มีค่าลดลงเรื่อย ๆ เมื่อเวลาผ่านไป เมื่อนำค่าพารามิเตอร์ที่ได้มาพยากรณ์แล้วเปรียบเทียบโดยใช้ค่า MAPE ในรายอายุพบว่า ค่า MAPE จากวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์มีค่าใกล้เคียงกัน โดย ค่า MAPE ในเพศหญิงนั้น วิธีแบบภาวะน่าจะเป็นสูงสุดจะได้ต่ำกว่าวิธีแบบเบย์ แต่ในทางตรงกันข้าม เพศชายวิธีแบบเบย์จะได้ค่า MAPE ที่ต่ำกว่า วิธีภาวะน่าจะเป็นสูงสุด

คำสำคัญ : ตัวแบบลี-คาร์เตอร์, วิธีแบบเบย์, วิธีภาวะน่าจะเป็นสูงสุด, ประชากรไทย, การสุ่มแบบมอนติคาร์โล ลูกโซ่มาร์คอฟ

ภาควิชา	สถิติ	ลายมือชื่อนิสิตคนที่ 1...........................
สาขาวิชา	การประกันภัย	ลายมือชื่อนิสิตคนที่ 2...........................
ปีการศึกษา	2563	

ABSTRACT

Project Name : Comparing Parameter Estimation of Thai Mortality Rate Modeling and Forecasting with Lee-Carter between Maximum Likelihood Method and Bayesian Method

Project ID : 18

Student name : 1. Mr. Jiratdhiti Phongkasem student ID 6042019626

2. Mr. Punnawit Khoployklang student ID 6042062526

Number of pages : 42 pages

The goal of this project is to compare parameter estimation of Thai mortality rate modeling and to forecast with Lee-Carter between the Maximum Likelihood Method(MLE) and the Bayesian method. To achieve the objective, we use estimated parameters from those two methodologies. Then, we forecast and use the MAPE value to compare efficiency between MLE and Bayesian.

The result shows, that parameters from both the MLE method and Bayesian method are not different. The average age profile of mortality value is increased highly between 0 to 5 years and declined from 6 to 20 years. After that, it has increased following age. Meanwhile, the factor of time index has fluctuated. Finally, the index of time value has declined following the calendar year. We forecasted the mortality rate by using estimated parameters and found that the MAPE value of the MLE approach similar to the MAPE value of Bayesian. However, the female MAPE value of the MLE approach is less than the MAPE value of Bayesian. Conversely, the male MAPE value of the MLE approach is greater than the MAPE value of the Bayesian approach.

Keywords: Lee-Carter Model, Bayesian Method, Maximum Likelihood Method, Thai Population, Markov Chain Monte Carlo

---

Department	Statistics	Student's Signature 1 ....
Field of study	Insurance	Student's Signature 2...........................
Academic Year	2020	

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญ

ในอุตสาหกรรมประกันภัย การสร้างตารางมรณะมีความสำคัญเป็นอย่างยิ่ง เนื่องจากข้อมูลตารางมรณะจะถูกนำไปใช้อย่างมากเช่น คิดเบี้ยประกันภัย การคิดเงินสำรองแบบต่าง ๆ เป็นต้น โดยบริษัทประกันภัยจะใช้ข้อมูลตารางมรณะทั้งจากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) และ ประสบการณ์ในอดีตของบริษัทประกันภัย โดยปกติแล้ว บริษัทประกันภัยจะมีเวลาคิดเบี้ยประกันจะใช้ตารางมรณะ ซึ่งค่าตารางมรณะนี้เองไม่มีความยืดหยุ่นตามการเวลา ดังนั้นบริษัทประกันภัยจึงจะกำหนดค่าขอบเขตความปลอดภัย (Safety of margin) มาเพื่อจัดการปัญหาความไม่ยืดหยุ่นของตารางมรณะ การกำหนดค่าขอบเขตความปลอดภัย (Safety of margin) นั้นอาจจะอิงจากเกณฑ์ของคปภ. หรือ การพยากรณ์อัตรามรณะในอนาคต ซึ่งหากการสร้างตัวแบบนั้นหากตัวแบบมีประสิทธิภาพในการพยากรณ์หรือความแม่นยำที่ไม่สูงมากอาจจะก่อให้เกิดความผิดพลาดได้ซึ่งจะส่งผลกระทบให้เกิดความสูญเสียทางธุรกิจประกันภัย

ผู้ศึกษาจึงเล็งเห็นถึงความสำคัญของการพยากรณ์อัตรามรณะ จึงใคร่ครวญต้องการทำโครงเกี่ยวกับอัตรามรณะ จากนั้นผู้ศึกษาจึงเริ่มทำการศึกษาการสร้างตัวแบบและพยากรณ์อัตรามรณะในไทยนั้นและได้พบว่าโดยทั่วไปจะใช้ ตัวแบบลี-คาร์เตอร์ (Lee and Carter, 1992) ผู้ศึกษาจึงเล็งเห็นถึงความสำคัญของการใช้ตัวแบบลี-คาร์เตอร์ในการพยากรณ์ และจากการทบทวนวรรณกรรม พบว่ามีวิธีในการประมาณในการประมาณค่าพารามิเตอร์ของตัวแบบ ลี-คาร์เตอร์ ได้หลายวิธีเช่น วิธีแบบ การแยกค่าเอกฐาน (SVD : Singular-Value-Decomposition), วิธีประมาณภาวะน่าจะเป็นสูงสุด (MLE : Maximum Likelihood Estimator), วิธีแบบเบย์โดยใช้มอนติคาร์โลลูกโซ่มาร์คอฟ (Bayesian estimator by using Markov Chain Monte Carlo) เมื่อผู้ศึกษาลองศึกษางานวิจัยในประเทศไทยพบว่าในงานวิจัยส่วนใหญ่โดยเฉพาะการพยากรณ์อัตรามรณะในประเทศไทยมีการนำตัวแบบลี-คาร์เตอร์นำตัวแบบลี-คาร์เตอร์ไปพัฒนาต่อยอดอย่างหลากหลาย เช่น Lee-Cater Model and Extensions to Forecast Thai mortality Rate (Natthasurang Yasungnoen, 2018), แต่ยังไม่มีการนำหลักการอนุมานแบบเบย์มาประยุกต์ใช้ในการประมาณค่าพารามิเตอร์ของตัวแบบลี-คาร์เตอร์ ผู้ศึกษาจึงยกบทความของ Claudia, Antone และ Michel ซึ่งอธิบายรูปแบบการประมาณค่าพารามิเตอร์ของตัวแบบลี-คาร์เตอร์ด้วยวิธีแบบเบย์กับข้อมูลประชากรและการตายของประเทศฝรั่งเศสมาเป็นแนวทางและปรับใช้กับประเทศไทยเพื่อทดสอบความเหมาะสมของการประมาณค่าพารามิเตอร์ของตัวแบบลี-คาร์เตอร์ด้วยวิธีแบบเบย์เพื่อพิจารณาถึงความสามารถในการนำบทความวิจัยนั้นมาประยุกต์ใช้ให้เข้ากับข้อมูลในประเทศไทยได้หรือไม่

1.2 วัตถุประสงค์ของการศึกษา

1.2.1 เพื่อประมาณค่าพารามิเตอร์ของตัวแบบลี-คาร์เตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์

1.2.2 เพื่อเปรียบเทียบประสิทธิภาพในการประมาณค่าพารามิเตอร์ทั้งสองแบบของลี-คาร์เตอร์

1.3.3 เพื่อวัดแนวโน้มอัตรามรณะของประชากรไทยในอนาคต

1.3 ขอบเขตการศึกษา

โครงงานนี้มุ่งศึกษาถึงการเปรียบเทียบประสิทธิภาพของวิธีการทางสถิติที่ใช้ในการพยากรณ์อัตรามรณะของประเทศไทย โดยเป็นการเน้นการเปรียบเทียบประสิทธิภาพระหว่างการใช้ตัวแบบลี-คาร์เตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดกับตัวแบบลี-คาร์เตอร์ด้วยวิธีแบบเบย์ รวมถึงพยากรณ์แนวโน้มอัตรามรณะของประชากรไทยในอนาคต

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 เพื่อเป็นแนวทางในการสร้างตัวแบบพยากรณ์อัตรามรณะโดยใช้วิธีอนุมานแบบเบย์

1.4.2 เพื่อศึกษาแนวโน้มอัตรามรณะในไทย

บทที่ 2

ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

ในส่วนของทฤษฎีและวรรณกรรมที่เกี่ยวข้องนั้นทางผู้ศึกษาของแบ่งเป็นสองส่วนด้วยกันคือ ทฤษฎีที่เกี่ยวข้องในส่วนแรก และวรรณกรรมที่เกี่ยวข้องในส่วนหลัง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ตัวแบบลี-คาร์เตอร์

$$m_{x,t} = e^{\alpha_x + \beta_x k_t + \varepsilon_{x,t}}$$

เมื่อ $m_{x,t}$ คืออัตรามรณะกลางปีของประชากรอายุ x ปี ในปีที่ t

α_x คือค่าคงที่กลางปีของประชากรอายุ x ปี

β_x คืออัตราเสื่อมของดัชนีเวลาที่อายุ x ปี

k_t คือดัชนีการเปลี่ยนแปลงของเวลาในปีที่ t

$\varepsilon_{x,t}$ คือความคาดเคลื่อนของอัตรามรณะกลางปีของประชากรอายุ x ปี ในปีที่ t ซึ่งมีค่าเฉลี่ยและความแปรปรวนเป็น 0 และ ${\sigma_\varepsilon}^2$ ตามลำดับ

โดย t คือปีที่ศึกษาซึ่ง $t = t_1, t_2, \dots, T$

x คืออายุที่ศึกษาซึ่ง $x = x_1, x_2, \dots, X$

ซึ่งทุก ๆ พารามิเตอร์นั้นเป็นค่าคงที่ที่ไม่ทราบค่า ดังนั้นจึงต้องหาค่าของอัตรามรณะกลางปีดังนี้

$$m_{x,t} = \frac{D_{x,t}}{L_{x,t}}$$

เมื่อ $D_{x,t}$ คือจำนวนการตายของประชากรอายุ x ปี ในปีที่ t

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

และในการประมาณค่าพารามิเตอร์นั้นเพื่อให้ได้คำตอบที่แน่นอนและมีเพียงคำตอบเดียวจึงจำเป็นที่จะต้องกำ หนดเงื่อนไข 2 เงื่อนไขคือ

$$\sum_{t=t_1}^{T} k_t = 0 \text{ และ } \sum_{x=x_1}^{X} \beta_x = 1$$

2.1.2 การกำหนดการแจกแจงก่อนการสังเกต (prior distribution) ของตัวแบบลี-คาร์เตอร์

ในการประมาณแบบเบย์สิ่งที่สำคัญคือการกำหนดกการแจกแจงก่อนการสังเกต (prior distribution) จะมีค่าพารามิเตอร์ที่สำคัญหลัก ๆ ของตัวแบบบลี-คาร์เตอร์สามตัวแปรคือ α, β, κ ซึ่งในงานวิจัยนี้จะอ้างอิงการกำหนดการแจกแจงก่อนการสังเกตตามงานวิจัย Bayesian Poisson log-bilinear mortality projections ของ Claudia, Antone และ Michel เพื่อสะดวกต่อการศึกษา โดยที่กำหนดให้มีข้อกำหนดเบื้องต้นดังนี้

$T = t_{\max} - t_{min} + 1$ คือช่วงระยะเวลาปีปฏิทิน

$M = x_{\max} - x_{\min} + 1$ คือช่วงระยะอายุ

2.1.4.1 การกำหนดการแจกแจงก่อนการสังเกตของค่า $\boldsymbol{\kappa}$

เนื่องจากค่า $\boldsymbol{\kappa}$ เป็นค่าดัชนีคาดเคลื่อนตามเวลาดังนั้น ค่าการแจกแจงก่อนการสังเกตจะขึ้นกับลักษณะของอนุกรมเวลาโดยมีรูปแบบดังนี้

$$(\kappa - \boldsymbol{X}\gamma) = \boldsymbol{P}(\kappa - \boldsymbol{X}\gamma) + \epsilon$$

$$P = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \rho & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \rho & 0 \end{pmatrix} \quad \boldsymbol{X} = \begin{pmatrix} 1 & t_{min} \\ \vdots & \vdots \\ 1 & t_{max} \end{pmatrix} \quad \gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \quad \epsilon \sim Normal(0, \sigma_\kappa^2)$$

ซึ่งจากสมการนี้สามารถสรุปเป็นการแจกแจงก่อนการสังเกตได้ดังนี้ $\kappa \sim Normal_T(\boldsymbol{X}\gamma, \sigma_\kappa^2 \boldsymbol{Q}^{-1})$

โดยที่ $$\boldsymbol{Q} = \begin{pmatrix} 1+\rho^2 & -\rho & 0 & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & \ddots & \vdots \\ 0 & -\rho & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1+\rho^2 & -\rho \\ 0 & \cdots & \cdots & -\rho & 1 \end{pmatrix}$$

จะสังเกตว่าในค่าในการแจกแจงก่อนการสังเกตมีตัวแปรที่ต้องหาค่าอีกทีเช่น $\rho, \gamma, \sigma_\kappa^2$ ซึ่งตัวแปรเหล่านี้จะเรียกว่าค่า ไฮเปอร์พารามิเตอร์ (hyperparameter) และในไฮเปอร์พารามิเตอร์จะมีการแจกแจงก่อนการสังเกตดังนี้

$$\gamma \sim Normal_2(\gamma_0, \Sigma_0)$$

$$\rho \sim Normal\left(0, \sigma_\rho^2\right)$$ ซึ่งถูกตัดที่ช่วง [0,1]

$$\sigma_{\kappa}^{-2} \sim Gamma(a_{\kappa}, b_{\kappa})$$

โดยค่า $\gamma_0, \Sigma_0, \sigma_{\rho}^2 a_{\kappa}, b_{\kappa}$ คือค่าคงที่ซึ่งจะถูกกล่าวในบทถัดไป

2.1.4.2 การกำหนดการแจกแจงก่อนการสังเกตของ β

$$\beta \sim Normal_M(0, \sigma_{\beta}^2 \boldsymbol{I_M})$$

โดยที่ $\boldsymbol{I_M}$ คือเมทริกซ์เอกลักษณ์มิติขนาด M X M และเพื่อง่ายต่อการคำนวนจะมีไฮเพอร์พารามิเตอร์ คือ

$$\sigma_{\beta}^{-2} \sim Gamma(a_{\beta}, b_{\beta})$$

2.1.4.3 การกำหนดการแจกแจงก่อนการสังเกตของ α

กำหนดให้ $\boldsymbol{e} = \exp(a)$ โดยที่ $e \sim Gamma(a_x, b_x)$

2.1.3 การพยากรณ์ค่าดัชนีการเปลี่ยนของเวลาโดยใช้ตัวแบบ ARIMA

จากตัวแบบลีคาร์เตอร์ หากเราต้องการพยากรณ์ค่าอัตรามรณะนั้นจะต้องพยากรณ์ค่าดัชนีการเปลี่ยนของเวลา โดยวิธีการพยากรณ์ค่านั้นเราจะใช้ตัวแบบ ARIMA(p,d,q) ในการพยากรณ์ดังนี้

$$\kappa_t = \phi_0 + \phi_1\kappa_{t-1} + \phi_2\kappa_{t-2} + \ldots + \phi_p\kappa_{t-p} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \ldots + \theta_q a_{t-q}$$

เมื่อ κ_t คือผลต่างระดับที่ d ของ k_t

a_t คือตัวรบกวนขาว (White noise) ซึ่ง $a_t \sim N(0, {\sigma_a}^2)$

ϕ_0 คือแนวโน้ม (Drift term)

2.1.4 มอนติคาร์โลลูกโซ่มาร์คอฟ (MCMC : Markov Chain Monte Carlo)

มอนติคาร์โลลูกโซ่มาร์คอฟคือกระบวนการสุ่มจากการแจกแจงความน่าจะเป็นโดยการสร้างลูกโซ่มาร์คอฟเพื่อให้ได้การลู่เข้าสู่การแจกแจงที่ต้องการเนื่องจากลูกโซ่มาร์คอฟมีคุณสมบัติในการลู่เข้าสู่การแจกแจงที่สมดุล (equilibrium distribution) ซึ่งลักษณะอัลกอริทึมการสุ่มสามารถแบ่งได้ออกเป็น 2 ประเภทคือ การสุ่มแบบ เมโทโปลิสแฮซติ้ง และ การสุ่มแบบกิบ

2.1.4.1 การสุ่มแบบเมโทโปลิศแฮซติ้ง (Metropolis-Hastings)

Metropolis-Hastings algorithms คือ กระบวนการสุ่มกลุ่มตัวอย่าง เพื่อประมาณการแจกแจง ร่วมของตัวแปรหลายตัวโดยใช้การแจกแจงแบบมีเงื่อนไข การสุ่มแบบเมโทโปลิศแฮซติ้ง นั้นจะมีกระบวนการในการยอมรับหรือปฏิเสธ ตัวอย่างที่สุ่มมาด้วย ก่อนอื่นจะต้องกำหนดการแจกแจง π ที่เป็น การแจกแจงเป้าหมาย (target distribution) ที่ต้องการประมาณการแจกแจงซึ่ง การสุ่มแบบเมโทโปลิศแฮซติ้ง จะทำการจำลองลูกโซ่มาร์คอฟ ที่มี π เป็น การแจกแจงที่สมดุล (equilibrium distribution) โดย ลูกโซ่มาร์คอฟ นี้จะกำหนดให้เป็น Q ซึ่งเป็น one-step transition probability หรือเรียกอีกอย่างว่า proposal distribution

สำหรับกรณีทั่วไปในการสุ่มตัวอย่างแบบ Metropolis-Hastings algorithms จะมีขั้นตอนดังนี้

1. กำหนดตัวอย่างเริ่มต้น $X^{(1)} = (x_1^{(1)}, x_2^{(1)}, \dots , x_k^{(1)})$

2. สำหรับ i ≥ 2 ตัวอย่าง $X^{(i)}$ สามารถหาได้โดย

2.1 สุ่ม X^* จากการแจกแจง $Q(X^*|X^{(i-1)})$

2.2 คำนวณ Hastings ratio $h = \frac{Q(X^{(i-1)}|X^*)}{Q(X^*|X^{(i-1)})}$

2.3 คำนวณ acceptance probability $\alpha = \min\,(1, \frac{\pi(X^*)}{\pi(X^{(i-1)})} \times h)$

2.4 สุ่ม u จากการแจกแจง U(0,1)

2.5 ถ้า u ≤ α จะกำหนดให้ $X^{(i)} = X^*$และถ้านอกเหนือจากนั้นจะกำหนดให้ $X^{(i)} = X^{(i-1)}$

2.1.4.2 การสุ่มแบบกิบ (Gibbs sampling)

การสุ่มแบบกิบ คือ กระบวนการสุ่มกลุ่มตัวอย่างจากการแจกแจงความน่าจะเป็นหลายตัวแปร (multivariate probability distribution) เพื่อประมาณการแจกแจงร่วมของตัวแปรเหล่านั้น โดยใช้การแจกแจงแบบมีเงื่อนไข ซึ่งขั้นตอนมีดังนี้

หากต้องการทราบการแจกแจงของ $p(x_1, x_2, \dots , x_k)$ โดยรู้การแจกแจงของ $p(x_j|x_1, x_2, \dots , x_{j-1}, x_{j+1} \dots , x_k)$ มีขั้นตอนดังนี้

1. กำหนดตัวอย่างเริ่มต้น $\left(x_1^{(1)}, x_2^{(1)}, \dots\ x_k^{(1)}\right)$
2. สำหรับ $i \geq 2$ ตัวอย่าง $\left(x_1^{(i)}, x_2^{(i)}, \dots\ x_k^{(i)}\right)$ สามารถหาได้โดย

   2.1 สุ่ม $x_1^{(i)}$ จากการแจกแจง $p(x_1|x_2^{(i-1)}, \dots , x_k^{(i-1)})$

   2.2 สุ่ม $x_2^{(i)}$ จากการแจกแจง $p(x_2|x_1^{(i)}, x_3^{(i-1)}, \dots , x_k^{(i-1)})$

   2.3 สุ่ม $x_3^{(i)}$ จากการแจกแจง $p(x_3|x_1^{(i)}, x_2^{(i)}, x_4^{(i-1)}, \dots , x_k^{(i-1)})$

   $\vdots$

   2.4 สุ่ม $x_k^{(i)}$ จากการแจกแจง $p(x_k|x_1^{(i)}, x_2^{(i)}, \dots , , x_{k-2}^{(i)}, x_{k-1}^{(i-1)})$

2.1.5 การวัดความแม่นยำของตัวแบบ

การวัดความแม่นยำของตัวแบบ เนื่องจากเป็นการวัดความแม่นยำของตัวแบบของการพยากรณ์จึงจะใช้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE : Mean Absolute Percentage Error) โดยตัวแบบที่ให้ค่า MAPE ต่ำที่สุดจะถือว่าเป็นตัวแบบที่มีประสิทธิภาพมากที่สุด โดยมีสูตรดังนี้

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

เมื่อ A_t คือ ค่าที่แท้จริง

เมื่อ F_t คือ จากการพยากรณ์

โดยค่า MAPE จะมีค่าออกมาอยู่ในรูปค่าความผิดพลายที่เป็นร้อยละโดยเฉลี่ย

2.2 วรรณกรรมที่เกี่ยวข้อง

ในปี 1992 Lee-Cater ได้ทำการตีพิมพ์วิธีการพยากรณ์อัตรามรณะในระยะยาวของอายุ ซึ่งใช้การผสมผสานวิธีการทางอนุกรมเวลา ซึ่งตัวแบบของลี-คาร์เตอร์ อธิบายลอการิทึมของอัตรามรณะ ด้วยค่าผลรวมระหว่างค่าคงที่ของอายุซึ่งอิสระจากเวลากับค่าที่ขึ้นกับการเวลาที่เกิดจากผลคูณระหว่างค่าเสื่อมของอัตรามรณะซึ่งแสดงถึงความไวของการลดลงของ อัตรามรณะ และค่าดัชนีเของวลา ค่าเหล่านี้จะเป็นค่าพารามิเตอร์ ซึ่งวิธีหาค่าพารามิเตอร์เหล่านี้จะใช้วิธีการแยกค่าเอกฐาน (SVD : Singular Value Decomposition) โดยใช้ข้อมูลการตายของประชากรประเทศสหรัฐอเมริกาตั้งแต่ปี 1933 ถึงปี 1987 ซึ่งพบว่า การพยากรณ์ในช่วงปี 1989 ถึง 1997 มีค่าใกล้เคียงกับค่าจริงของการตายและการพยากรณ์อัตรามรณะ ตั้งแต่ปี 1990 ถึงปี 2065 ซึ่งพบว่าอัตรามรณะที่ลดลง และตัวแบบ Lee-Cater นี้เองถูกนำมาใช้เป็นตัวแบบพื้นฐานในการพยากรณ์อัตรามรณะและถูกนำมาใช้อย่างยิ่งโดยเฉพาะอย่างยิ่งในกลุ่มประเทศกลุ่ม G7 และพบว่า ค่าดัชนีของเวลามีลักษณะเป็น Simple Random Walk (ARIMA (0,1,0))

ถัดมาในปี 2002 Bronhus และคณะได้ตีพิมพ์วิจัยชื่อ A poisson log-bilinear regression approach to construction of projected lifetables โดยมีการพัฒนานำแนวคิด ตัวแบบถดถอยแบบปัวซงค์ (Poisson regression model) มาใช้ในตัวแบบลี-คาร์เตอร์ โดยให้จำนวนคนตาย ณ ปีที่ t มีลักษณะการแจกแจงแบบปัวซงค์ ซึ่งจำนวนคนตายเท่ากับจำนวนประชากรคูณอัตรามรณะ ซึ่งจะได้ว่าอัตรามรณะมีลักษณะเป็น log-bilinear หลังจากนั้นมีการพัฒนาตัวแบบ Lee-Cater โดยประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด (MLE : Maximum likelihood estimator) โดยประมาณค่า จากจำนวนการตายที่แจกแจงแบบปัวซงค์โดยใช้วิธีการทางคณิตศาสตร์เพื่อหาค่าพารามิเตอร์ มาแทนการใช้ วิธี SVD ซึ่งพบว่าการใช้ MLE เหมาะสมกับการประยุกต์ใช้ในการคำนวนทางเบี้ยประกัน

Claudia, Antone และ Michel ได้มีต่อยอดงานวิจัยจาก Bronhus โดยนำแนวคิดการมองค่าการตายเป็นการแจกแจงแบบปัวซงค์ โดยใช้วิธีการอนุมานทางเบย์มาประมาณหาค่าพารามิเตอร์ในตัวแบบลี-คาร์เตอร์ โดยนำวิธีสุ่มมอนติคาร์โลลูกโซ่มาร์คอฟ (MCMC: Markov Chain Monte Carlo) มาคำนวณและเปรียบเทียบกับวิธี SVD กับ MLE ซึ่งผลที่ได้พบว่า ผลลัพธ์ค่าคงที่ของประชากรกลางปีจากการใช้ทั้งสามวิธีนั้นมีค่าใกล้เคียงกันแต่ค่าอัตราเสื่อมของดัชนีกับค่าดัชนีตามเวลาที่หาจากวิธี SVD มีค่าที่ไม่ตกอยู่ในช่วงความน่าเชื่อถือของวิธีแบบเบย์ และค่าอัตรามรณะของวิธี MLE อยู่ในช่วงค่าความน่าเชื่อถือของวิธีแบบเบย์

ในปี 2005 Alho & Spencer ได้เขียนในหนังสือ Statistical Demography & Forecasting (หน้า 242) แนะนำว่า ความยาวของข้อมูลที่ใช้ในการพยากรณ์ต้องมีความยาวมากกว่าปีที่พยากรณ์ โดยแนะนำให้ยาวกว่า 2 ถึง 3 เท่า

ปี 2008 Booth และ Tickle ได้ทำการทบทวนวิธีสร้างตัวแบบและพยากรณ์แบบต่าง ๆ พบว่า ตัวแบบลี-คาร์เตอร์ มีประสิทธิภาพและมีความแม่นยำสูง ในขณะที่ตัวแบบเชิงเส้นมาตรฐาน (GLM: Generalized Linear Model) มีความประสบความสำเร็จน้อยกว่าเนื่องจากภาวะไม่เป็นเส้นตรงของเวลา ในการประยุกต์ใช้ตัวแบบ ลี-คาร์เตอร์ กับข้อมูลต่าง ๆ ค่าดัชนีของเวลาเกือบทุกงานจะมีลักษณะเป็น Simple Random Walk (ARIMA (0,1,0)) หรือเป็น ARIMA (1,1,0) โดยในปัจจุบันตัวแบบที่มีความนิยมคือ ตัวแบบ ช่วงอายุ ลี-คาร์เตอร์ ซึ่งปัจจุบันโปรแกรมที่ใช้ในการทำตัวแบบมีการพัฒนาเป็นอย่างมาก

ถัดมาปี 2015 Arkadiusz Wisniowski & Peter W. F. Smith & Jakub Bijak & James Raymer & Jonathan J. Forster ได้ตีพิมพ์บทความ Bayesian Population Forecasting: Extending the Lee-Carter Method ซึ่งพยากรณ์จำนวนประชากรของประเทศอังกฤษโดยอาศัยปัจจัย 4 อย่างคือ การตาย การเกิด การอพยพเข้า การอพยพออก ซึ่งคณะผู้วิจัยแสดงให้เห็นถึงความยืดหยุ่นและข้อดีของการเอาวิธีแบบเบย์มาใช้

ปี 2016 จันท์ธิดา บุญมหาสิทธิ์ และ สำรวม จงเจริญ ได้ทำการศึกษาการเปรียบเทียบการใช้ตัวแบบเพื่อพยากรณ์อัตรามรณะโดยใช้ข้อมูลจำนวนประชากรและข้อมูลการตายจำแนกตามอายุและเพศของประชากรไทยปี ระหว่าง ปี 2539 ถึง 2555 เปรียบเทียบระหว่าง วิธีลี-คาร์เตอร์ที่ประมาณค่าพารามิเตอร์ด้วยวิธีค่าเฉพาะเจาะจง ตัวแบบลี-คาร์เตอร์ที่ประมาณพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและตัวแบบลี-มิลเลอร์ ซึ่งพบว่าตัวแบบที่ใช้แล้วเหมาะสมที่สุดในการพยากรณ์อัตรามรณะของประชากรไทยในอีก 10 ปีข้างหน้าคือตัวแบบลี-คาร์เตอร์ที่ประมาณพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด นอกจากนี้ยังพบว่าค่าพยากรณ์ของทั้ง 3 วิธีมีแนวโน้มลดลงในอีก 10 ปีข้างหน้า

ปี 2018 ณัฐสุรางค์ ยาสูงเนิน ได้ทำการพยากรณ์อัตรามรณะไทยโดยใช้ 2 ตัวแบบเพิ่มเติมจากลี-คาร์เตอร์ คือ ตัวแบบช่วงอายุประชากรตามรุ่น (Age period cohort model) และตัวแบบช่วงอายุประชากรตามรุ่นแบบกรณีเฉพาะเจาะจง (Specific case age cohort period model) โดยใช้วิธีประมาณค่าพารามิเตอร์แบบภาวะน่าจะเป็นสูงสุดโดยให้การตายแจกแจงปัวซงค์และการแจกแจงทวินามลบ โดยพบว่าอัตรามรณะในไทยในอนาคตมีการลดลงอย่างช้า ๆ ทั้งเพศชายและหญิง เนื่องจากอายุขัยเฉลี่ยที่เพิ่มมากขึ้น

บทที่ 3
ระเบียบวิธีศึกษา

3.1 ข้อมูลที่ใช้ศึกษา

3.1.1. จำนวนประชากรที่ตายของประชากรไทยทั้งเพศชายและหญิงตั้งแต่อายุ 0-99 ปี ตั้งแต่ปี 2540-2562

3.1.2. จำนวนประชากรกลางปีของประชากรไทยทั้งชายและหญิงตั้งแต่อายุ 0-99 ปี ตั้งแต่ปี 2540-2562

3.2 ขั้นตอนศึกษา

ผู้ศึกษาได้มีวิธีการเลือกเครื่องมือที่ใช้ทำวิจัย และการคัดเลือกข้อมูลสำหรับการวิจัยดังนี้

3.2.1 เครื่องมือที่ใช้ในการวิจัย

โปรแกรมทางสถิติ อาร์ (R Studio)

3.2.2 การแบ่งข้อมูล

ผู้ศึกษาแบ่งข้อมูลของการจำนวนประชากรไทยชายและหญิงเป็นสองชุดข้อมูล คือ

1. ข้อมูลฝึกฝน (Training Set) เพื่อใช้ในการสร้างตัวแบบโดยใช้จำนวนคนตายของประชากรไทยชายและหญิงตั้งแต่ปี พ.ศ. 2540-2559

2. ข้อมูลทดสอบ (Test Set) เพื่อใช้ในการทดสอบประสิทธิภาพของตัวแบบใช้จำนวนคนตายและจำนวนประชากรกลางปีของคนไทยทั้งชายและหญิงตั้งแต่ปีพ.ศ. 2560-2562

หลังจากที่ได้ทำการแบ่งข้อมูลเรียบร้อยแล้วจะทำการประมาณค่าพารามิเตอร์โดยทั้งสองแบบด้วยข้อมูลฝึกฝนและประเมินความแม่นยำของการพยากรณ์ด้วยการนำค่าที่พยากรณ์ได้มาเทียบกับข้อมูลทดสอบ

3.3 วิธีการประมาณค่าพารามิเตอร์

3.3.1 การประมาณค่าพารามิเตอร์ตัวแบบลี-คาร์เตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด

โดยวิธีการประมาณค่าพารามิเตอร์ในตัวแบบลี-คาร์เตอร์จะประมาณโดยสมมติให้จำนวนการตายแจกแจงแบบปัวซงค์ (Poisson) ดังนี้

$$D_{x,t} \sim Poi(L_{x,t}m_{x,t})$$

จากนั้นจึงทำการหาลอการิทึมของฟังก์ชันภาวะน่าจะเป็น (Log-likelihood function) ดังนี้

$$L(\alpha_x, \beta_x, \kappa_t | D_{x,t}) = \prod_{x,t} \frac{e^{-L_{x,t}m_{x,t}}(L_{x,t}m_{x,t})^{D_{x,t}}}{D_{x,t}!} = \prod_{x,t} \frac{e^{-L_{x,t}e^{\alpha_x+\beta_x\kappa_t}}(L_{x,t}e^{\alpha_x+\beta_x\kappa_t})^{D_{x,t}}}{D_{x,t}!}$$

$$log\, L\,(a_x, \beta_x, k_t | D_{x,t}) = \sum_{x,t} [\, D_{x,t}(\alpha_x + \beta_x\kappa_t) - L_{x,t}e^{\alpha_x+\beta_x\kappa_t} + D_{x,t}\, ln\, L_{x,t} - ln(\, D_{x,t}!)]$$

ซึ่งหลังจากนี้จะสามารถหาค่าประมาณพารามิเตอร์ได้โดยการหาอนุพันธ์ (Differentiate) เทียบกับแต่ละพารามิเตอร์แล้วให้ผลลัพธ์มีเท่ากับ 0 และใช้วิธีการ นิวตัน-ราฟสัน (Newton-Raphson method) เพื่อให้ได้ค่าประมาณพารามิเตอร์ $\{\alpha_x\},\{\beta_x\}$ และ $\{k_t\}$ โดยมีวิธีการดำเนินการสำหรับค่าประมาณพารามิเตอร์ในขั้นที่ v+1 ดังนี้

$$\hat{\theta}^{(v+1)} = \hat{\theta}^{(v)} - \frac{\frac{\partial l^{(v)}}{\partial \theta}}{\frac{\partial^2 l^{(v)}}{\partial \theta^2}}$$

เมื่อ

$$l^{(v)} = log\, L\,(\hat{\theta}^{(v)}|D_{x,t})$$

ซึ่งกำหนดค่าเริ่มต้นโดยให้ $\hat{\alpha}_x^{(0)} = 0$, $\hat{\beta}_x^{(0)} = 1$ และ $\hat{\kappa}_t^{(0)} = 0$ จากนั้นจึงแทนค่าประมาณพารามิเตอร์ลงในสมการดังนี้

$$\hat{\alpha}_x^{(v+1)} = \hat{\alpha}_x^{(v)} - \frac{\sum_t (D_{x,t} - \hat{D}_{x,t}^{(v)})}{-\sum_t \hat{D}_{x,t}^{(v)}}, \hat{\beta}_x^{(v+1)} = \hat{\beta}_x^{(v)}, \hat{\kappa}_t^{(v+1)} = \hat{\kappa}_t^{(v)}$$

$$\hat{\kappa}_t^{(v+2)} = \hat{\kappa}_t^{(v+1)} - \frac{\sum_x (D_{x,t} - \hat{D}_{x,t}^{(v+1)})\, \hat{b}_x^{(v+1)}}{-\sum_x \hat{D}_{x,t}^{(v+1)} (\hat{b}_x^{(v+1)})^2}, \hat{\alpha}_x^{(v+2)} = \hat{\alpha}_x^{(v+1)}, \hat{\beta}_x^{(v+2)} = \hat{\beta}_x^{(v+1)}$$

$$\hat{\beta}_x^{(v+3)} = \hat{\beta}_x^{(v+2)} - \frac{\sum_t (D_{x,t} - \hat{D}_{x,t}^{(v+2)}) \hat{k}_t^{(v+2)}}{-\sum_t \hat{D}_{x,t}^{(v+2)} (\hat{k}_t^{(v+2)})^2}, \hat{\alpha}_x^{(v+3)} = \hat{\alpha}_x^{(v+2)}, \hat{\kappa}_t^{(v+3)} = \hat{\kappa}_t^{(v+2)}$$

โดยที่

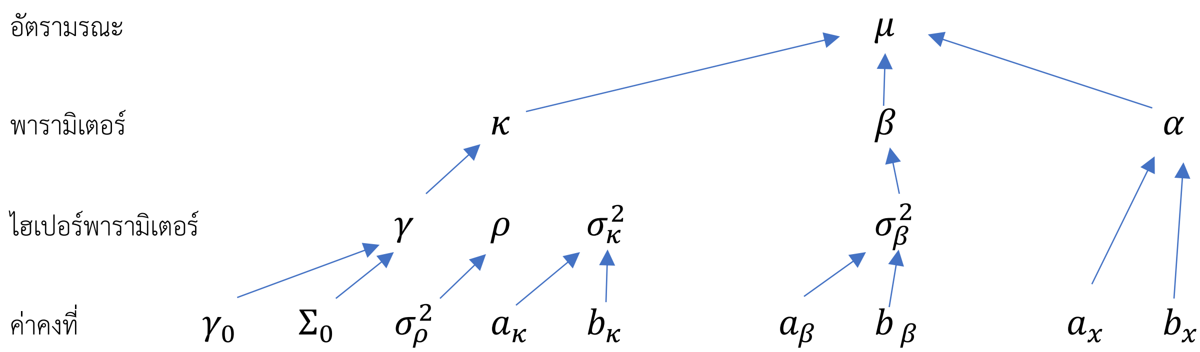
$$\hat{D}_{x,t}^{(v)} = L_{x,t} e^{\hat{\alpha}_x^{(v)} + \hat{\beta}_x^{(v)} \hat{\kappa}_t^{(v)}}$$

และจะดำเนินการวนซ้ำไปเรื่อย ๆ จนกว่าค่าประมาณพารามิเตอร์มีค่าที่เปลี่ยนแปลงน้อยมากหรือมีค่าลู่เข้าสู่ค่าเดียวกัน

โดยการทำการประมาณค่าโดยใช้วิธีภาวะน่าจะเป็นสูงสุดจะใช้โปรแกรม R ใช้ไลบรารี่ชื่อว่า STMOMO ในการประมาณ

3.3.2 การประมาณค่าโดยใช้วิธีแบบเบย์

จากบทที่แล้วจะเห็นว่าการแจกแจงก่อนการสังเกตจะมีค่าพารามิเตอร์หลากหลายค่าและในค่าพารามิเตอร์เองก็มีตัวแปรสุ่มในนั้นด้วย ซึ่งสามารถเขียนได้เป็นสรุปดังนี้



จะเห็นว่าในเริ่มต้นจะต้องกำหนดค่าคงที่เริ่มต้นซึ่งจะกำหนดดังนี้

ในค่าประมาณเริ่มต้นเราจะประมาณ γ_0จากการประมาณค่าพารามิเตอร์ของการถดถอยเชิงเส้นระหว่าง ค่า κ กับปีปฏิทิน ซึ่งค่า κ จะได้มาจากการประมาณโดยวิธีภาวะน่าจะเป็นสูงสุด นอกจากนี้เราจะได้ค่า Σ_0 จากเมทริกซ์แปรปรวน (Variance Co-variance matrix) หลังจากนั้นเราจะนำ $\kappa - X\gamma_0$ มาสร้างตัวแบบ AR(1) เพื่อหาค่า ρ และ σ_ρ^2

ในส่วนของค่า a_β, b_β จะเริ่มจากการหา σ_β^2 ซึ่งสามารถหาได้จากการหาการแปรปรวนจากค่า β ที่ถูกประมาณโดยวิธีภาวะน่าจะเป็นสูงสุด เรารู้ว่า $\sigma_\beta^{-2} \sim Gamma(a_\beta, b_\beta)$ จะได้ว่า $E\left(\sigma_\beta^2\right) = \frac{b_\beta}{a_\beta - 1}$ จะได้ว่า $b_\beta = \left(a_\beta - 1\right)\sigma_\beta^2$ ซึ่งเราจะกำหนดค่า $a_\beta = 2.1$ เพื่อควบคุมความแปรปรวน โดยค่าของ a_κ, b_κ ก็หาได้ในลักษณะเดียวกัน ถัดมาการหาค่า a_x, b_x เราจะกำหนดให้ $b_x = 0.001$ และ $a_x = b_X e^{\alpha_x}$ โดยค่า α_x หามาจากวิธีภาวะน่าจะเป็นสูงสุด

หลังจากกำหนดค่าเริ่มต้นขั้นตอนถัดไปคือการทำการสุ่มโดยจะทำการสุ่ม 20,000 รอบ โดย 10,000 รอบแรกเป็นการเบิร์นอิน (burn-in) เนื่องจากใน 10,000 รอบแรกถือว่าเป็นช่วงแรกการสุ่มจะทำให้การแจกแจงมีลักษณะไม่คงที่และมีรูปลักษณะที่คาดเดายากดังนั้นใน 10,000 รอบแรกจะไม่นำมารวมกับการแจกแจงค่าเฉลี่ยของการแจกแจงหลังการสังเกต และอีก 10,000 รอบหลังจากนั้นจะเป็นการสุ่มเพื่อหาการแจกแจงหลังการสังเกต ในกระบวนการสุ่มจะแยกเป็นสองแบบคือการสุ่มแบบเมโทรโปลิศแฮชติ้งและการสุ่มแบบกิบโดยวิธีการสุ่มจะขึ้นกับพารามิเตอร์ดังนี้

3.3.2.1 การสุ่มแบบเมโทรโปลิศแฮชติ้งสำหรับ κ

กำหนดให้ $\kappa_{-t} = (\kappa_1, \kappa_2, \dots, \kappa_{t-1}, \kappa_{t-1}, \dots \kappa_{t_{max}})'$

จะได้ค่าการแจกแจงหลังการสังเกตดังนี้

เมื่อ $t = t_{min}$

$$f(\kappa|\kappa_{-t}, \alpha, \beta, D, \sigma_\kappa^2, \sigma_\beta^2, \rho, \gamma)$$

$$\propto \prod_x \exp(-E_{xt}\exp(\alpha_x + \beta_x\kappa_t)) \prod_x \exp(\beta_x \kappa_t D_{xt})$$

$$\exp\left(-\frac{1}{2\sigma_\kappa^2}(\kappa_t-\eta_t)^2\right)\exp\left(-\frac{1}{2\sigma_\kappa^2}(\kappa_{t+1}-\eta_{t+1}-\rho(\kappa_t-\eta_t))^2\right)$$

เมื่อ $t = t_{max}$

$$f(\kappa|\kappa_{-t},\alpha,\beta,D,\sigma_\kappa^2,\sigma_\beta^2,\rho,\gamma)$$
$$\propto \prod_x \exp(-E_{xt}\exp(\alpha_x+\beta_x\kappa_t))\prod_x \exp(\beta_x\kappa_t D_{xt})$$
$$\exp\left(-\frac{1}{2\sigma_\kappa^2}(\kappa_t-\eta_t-\rho(\kappa_{t-1}-\eta_{t-1}))^2\right)$$

เมื่อ $t_{min} < t < t_{max}$

$$f(\kappa|\kappa_{-t},\alpha,\beta,D,\sigma_\kappa^2,\sigma_\beta^2,\rho,\gamma)$$
$$\propto \prod_x \exp(-E_{xt}\exp(\alpha_x+\beta_x\kappa_t))\prod_x \exp(\beta_x\kappa_t D_{xt})$$
$$\exp\left(-\frac{1}{2\sigma_\kappa^2}(\kappa_t-\eta_t-\rho(\kappa_{t-1}-\eta_{t-1}))^2\right)$$
$$\exp\left(-\frac{1}{2\sigma_\kappa^2}(\kappa_{t+1}-\eta_{t+1}-\rho(\kappa_t-\eta_t))^2\right.$$

ทุก ๆ รอบที่ $i+1$ ของการสุ่มเมโทลิศแฮชติ้งจากการปรับปรุงค่า κ_t จะมีขั้นตอนต่าง ๆ ดังนี้ สมมติให้เรามีค่า κ_t^i ที่ได้จากการที่เมโทโปลิศแฮชติ้งรอบที่ i โดยมีขั้นตอนดังนี้

1. สุ่มค่า κ_t^*จาก $Normal(\kappa_t^i,\sigma_t^2)$ ซึ่งรู้ค่าความแปรปรวนของ σ_t^2
2. คำนวณค่าความน่าจะเป็นที่ยอมรับได้ (acceptance probability)

$$\Psi\left(\kappa_t^i,\kappa_t^*\right)=\min\left(1,\frac{f(\kappa_i^*|\kappa_{-t}^i,D,\alpha,\beta,\sigma_\kappa^2,\sigma_\beta^2,\rho,\gamma)}{f(\kappa_i^i|\kappa_{-t}^i,D,\alpha,\beta,\sigma_\kappa^2,\sigma_\beta^2,\rho,\gamma)}\right)$$

โดยที่

$$\kappa_{-t}=(\kappa_1,\kappa_2,\ldots,\kappa_{t-1},\kappa_{t-1},\ldots\kappa_{t_{max}})'$$

3. หลังจากนั้นสร้างค่า u จากการแจกแจง $Uniform(0,1)$ ในกรณีที่ $u \leq \Psi\left(\kappa_t^i, \kappa_t^*\right)$ จะได้ว่า $\kappa_t^{i+1} = \kappa_i^*$ ในทางกลับกันจะได้ว่าถ้า $u > \Psi\left(\kappa_t^i, \kappa_t^*\right)$

   ค่าที่ถูกสุ่มมาจะปฏิเสธและลูกโซ่มาร์คอฟจะไม่ถูกขยับ ซึ่งจะได้ว่า $\kappa_t^{i+1} = \kappa_i^i$

4. โดยสุดท้ายแล้วจะได้ว่า

$$\kappa_t^{i+1} = (\kappa_1^{i+1}, \kappa_2^{i+1}, \dots, \kappa_t^{i+1}, \kappa_{t-1}^{i+1}, \dots \kappa_{t_{max}}^{i+1})'$$

และสำหรับ α^i เพื่อที่จะให้สอดคล้องกับเงื่อนไขที่

$$\sum_{t=t_1}^{T} k_t = 0 \text{ และ } \sum_{x=x_1}^{X} \beta_x = 1$$

เราจะใช้สูตร
$$\kappa^{i+1} \leftarrow \kappa^{i+1} - \bar{\kappa}$$
$$\alpha^i \leftarrow \alpha^i + \beta^i \bar{\kappa}$$

โดยที่

$$\bar{\kappa} = \frac{1}{T}\left(\sum_{s \leq t} \kappa_s^{i+1} + \sum_{s \leq t} \kappa_s^i\right)$$

ค่าพารามิเตอร์ σ_t^2 เป็นค่าที่ควบคุมความน่าจะเป็นในการยอมรับ โดยหากกำหนดค่าความแปรปรวนมากจะเพิ่มค่าความน่าจะเป็นในการยอมรับค่า κ_i^* โดยในปกติแล้วเราต้องการค่าการยอมรับจากการสุ่มอยู่ในช่วง [20%, 50%] โดยหากค่าการยอมรับมีค่าน้อยให้เพิ่มค่าความแปรปรวนแต่หากการยอมรับมีค่ามากให้ลดความแปรปรวน

3.3.2.2 การสุ่มแบบเมโทรโปลิสแฮชติ้งสำหรับ β

กำหนดให้
$$\beta_{-x} = (\beta_{x_{min}}, \beta_2, \dots \beta_{x-1}, \beta_{x+1}, \dots \beta_{x_{max}})'$$

จะได้การแจกแจงหลังการสังเกตดังนี้

$$f(\beta | \beta_{-x}, \alpha, D, \sigma_\kappa^2, \sigma_\beta^2, \rho, \gamma)$$
$$\propto \prod_x \exp(-E_{xt} \exp(\alpha_x + \beta_x \kappa_t)) \prod_x \exp(\beta_x \kappa_t D_{xt}) \exp\left(-\frac{1}{2\sigma_\beta^2} \beta_x^2\right)$$

ทุก ๆ รอบที่ $i+1$ของการสุ่มเมโทลิศแฮชติ้งจากการปรับปรุงค่า β_x จะมีขั้นตอนต่าง ๆ ดังนี้ สมมติให้เรามีค่า β_x^i ที่ได้จากการที่เมโทโปลิศแฮชติ้งรอบที่ i โดยมีขั้นตอนดังนี้

1. สุ่มค่า β_x^* จาก $Normal(\beta_x^i, \sigma_x^2)$ ซึ่งรู้ค่าความแปรปรวนของ σ_t^2
2. คำนวณค่าความน่าจะเป็นที่ยอมรับได้ (acceptance probability)

$$\Psi\left(\beta_x^i, \beta_x^*\right) = \min\left(1, \frac{f(\beta_x^* | \beta_{-x}^i, D, \alpha, \kappa, \sigma_\kappa^2, \sigma_\beta^2, \rho, \gamma)}{f(\beta_i^i | \beta_{-t}^i, D, \alpha, \kappa, \sigma_\kappa^2, \sigma_\beta^2, \rho, \gamma)}\right)$$

   โดยที่

$$\beta_{-x}^i = (\beta_{x_{\min}}^{i+1}, \dots, \beta_{x-1}^i, \beta_{x+1}^i, \dots \beta_{x_{max}}^i)'$$

3. หลังจากนั้นสร้างค่า u จากการแจกแจง $Uniform(0,1)$ ในกรณีที่ $u \leq \Psi\left(\beta_x^i, \beta_x^*\right)$ จะได้ว่า $\beta_x^{i+1} = \beta_x^*$ ในทางกลับกันจะได้ว่าถ้า $u > \Psi\left(\beta_x^i, \beta_x^*\right)$

   ค่าที่ถูกสุ่มมาจะปฏิเสธและลูกโซ่มาร์คอฟจะไม่ถูกขยับ ซึ่งจะได้ว่า $\beta_x^{i+1} = \beta_x^i$
4. โดยสุดท้ายแล้วจะได้ว่า

$$\beta_x^i = \left(\beta_{x_{\min}}^{i+1}, \dots, \beta_x^i, \beta_{x+1}^i, \dots \beta_{x_{max}}^i\right)'$$

และสำหรับ α^i เพื่อที่จะให้สอดคล้องกับเงื่อนไขที่

$$\sum_{t=t_1}^{T} k_t = 0 \text{ และ } \sum_{x=x_1}^{X} \beta_x = 1$$

เราจะใช้สูตร

$$\beta^{i+1} \leftarrow \frac{\beta^{i+1}}{\beta_.}$$

$$\kappa^{i+1} \leftarrow \kappa^{i+1}\beta_.$$

โดยที่

$$\beta_. = (\sum_{y \leq x} \beta_y^{i+1} + \sum_{y > x} \beta_y^i)$$

ค่าพารามิเตอร์ σ_x^2 เป็นค่าที่ควบคุมความน่าจะเป็นในการยอมรับ โดยหากกำหนดค่าความแปรปรวนมากจะเพิ่มค่าความน่าจะเป็นในการยอมรับค่า β_x^* โดยในปกติแล้วเราต้องการค่าการยอมรับจากการสุ่มอยู่ในช่วง [20%, 50%] โดยหากค่าการยอมรับมีค่าน้อยให้เพิ่มค่าความแปรปรวนแต่หากการยอมรับมีค่ามากให้ลดความแปรปรวน

3.3.2.3 การสุ่มแบบกิบสำหรับ α

กำหนดให้

$$c_x = \sum_x E_{xt} \exp\left(\beta_x \kappa_t\right) \text{ และ } D_{x\cdot} = \sum_t D_{xt}$$

จะได้การแจกแจงหลังการสังเกตของ α เป็นดังนี้

ให้ $e = \exp(\alpha)$ จะได้ว่า

$$e|\beta, \kappa, D, \sigma_\kappa^2, \sigma_\beta^2, \rho, \gamma \sim Gamma(a_X + D_{x\cdot}, b_x + c_x)$$

3.3.2.4 การสุ่มแบบกิบสำหรับ ρ

เมื่อทำการสุ่มแบบกิบจะได้การแจกแจงหลังการสังเกตดังนี้

$$\rho|\alpha, \beta, \kappa, D, \sigma_\kappa^2, \sigma_\beta^2, \gamma \sim Normal\left(\mu_\rho^*, \sigma_\rho^{2^*}\right) \text{ ตัดช่วงที่ } (-1,1)$$

โดยที่

$$\mu_\rho^* = \frac{b_\rho}{a_\rho + \frac{\sigma_\kappa^2}{\sigma_\rho^2}} \quad and\ \sigma_\rho^{2^*} = \frac{\sigma_\kappa^2}{a_\rho + \frac{\sigma_\kappa^2}{\sigma_\rho^2}}$$

เมื่อ

$$a_\rho = \sum_t (\kappa_{t-1} - \eta_{t-1})^2 \text{ และ } b_\rho = \sum_t (\kappa_t - \eta_t)(\kappa_{t-1} - \eta_{t-1})$$

3.3.2.5 การสุ่มแบบกิบสำหรับ σ_κ^2

เนื่องจากการแจกแจงก่อนการสังเกตมีลักษณะเป็น $\sigma_\kappa^{-2} \sim Gamma(a_\kappa, b_\kappa)$ ดังนั้นลักษณะการแจกแจงหลังจากการควรจะเป็นตามรูปที่สอดคล้องกับการแจกแจงก่อนการสังเกตดังนี้

$$\sigma_\kappa^{-2}|\alpha, \beta, \kappa, D, \rho, \sigma_\beta^2, \gamma \sim Gamma\left(a_\kappa + \frac{T}{2}, b_\kappa + \frac{1}{2}\sum_t \left(\kappa_t - \eta_t - \rho(\kappa_{t-1} - \eta_{t-1})\right)^2\right)$$

<u>3.3.2.6 การสุ่มแบบกิบสำหรับ σ_β^2</u>

การแจกแจงหลังการสังเกตของ σ_β^2 เป็นดังนี้

$$\sigma_\beta^{-2}|\alpha,\beta,\kappa,D,\rho,\sigma_\kappa^2,\gamma \sim Gamma\left(a_\beta + \frac{M}{2}, b_\beta + \frac{1}{2}\beta'\beta\right)$$

<u>3.3.2.7 การสุ่มแบบกิบสำหรับ γ</u>

การแจกแจงหลังการสังเกตของ γ เป็นดังนี้

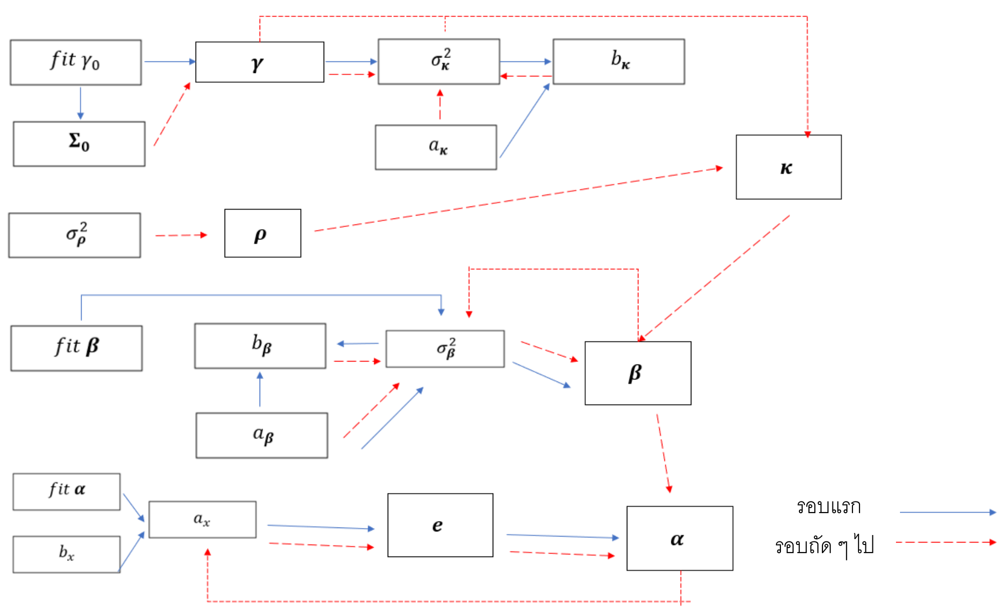
$$\gamma|\kappa,\sigma_\beta^2,\sigma_\kappa^2,\rho \sim Normal_2(\gamma^*,\sigma_\kappa^2 \textstyle\sum^*)$$

โดยที่

$$\Sigma^* = (X'QX + \sigma_\kappa^2\Sigma_0^{-1})^{-1}, \qquad \gamma^* = \Sigma^*(X'Q\kappa + \sigma_\kappa^2\Sigma_0^{-1}\gamma_0)$$

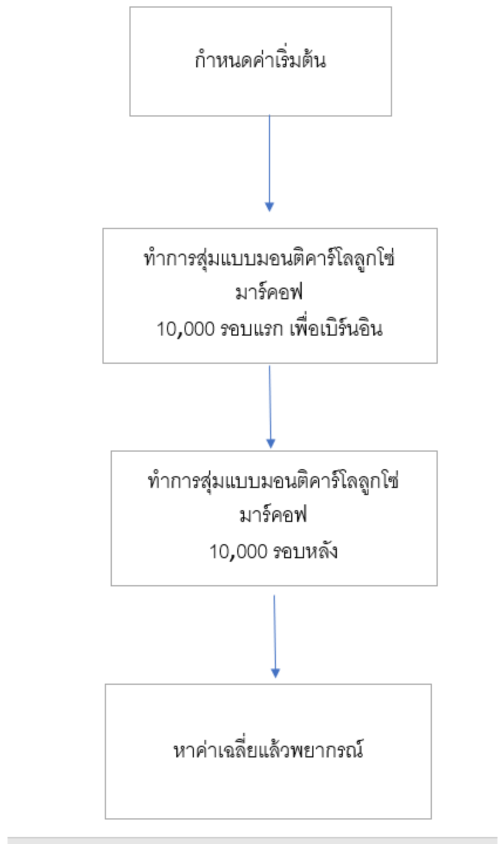
ซึ่งเราจะสุ่มตัวอย่างจากการแจกแจงหลังการสังเกตโดยมีขั้นตอนการสุ่มตามพารามิเตอร์เป็นลำดับตามรูปภาพดังนี้

รูปที่ 3.1 แสดงภาพกระบวนการทำประมาณค่าพารามิเตอร์



และเมื่อมองเป็นภาพรวมจะได้กระบวนการดังนี้

รูปที่ 3.2 แสดงกระบวนการคัดเลือกพารามิเตอร์



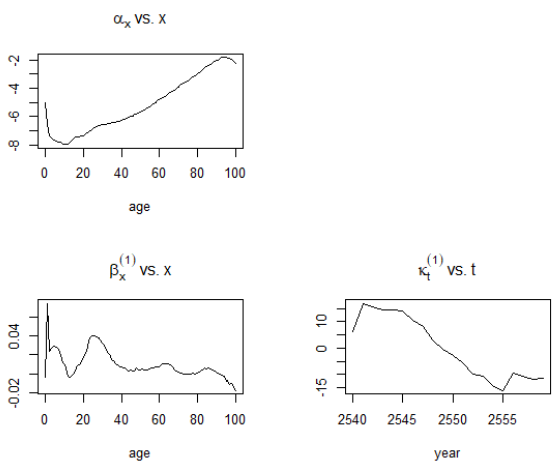
บทที่ 4

ผลการศึกษา

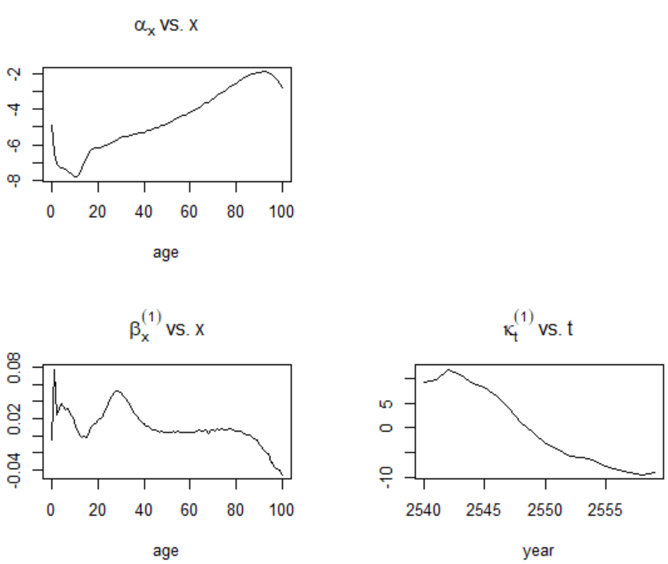
4.1 ผลการประมาณค่าโดยใช้วิธีภาวะน่าจะเป็นสูงสุด

จากการประมาณค่าพารามิเตอร์ผ่านโปรแกรม R ได้ผลลัพธ์เป็นค่าพารามิเตอร์ดังรูป

รูปที่ 4.1 แสดงถึงค่าพารามิเตอร์ของเพศหญิงที่ได้จากการประมาณค่าโดยใช้วิธีภาวะน่าจะเป็นสูงสุด

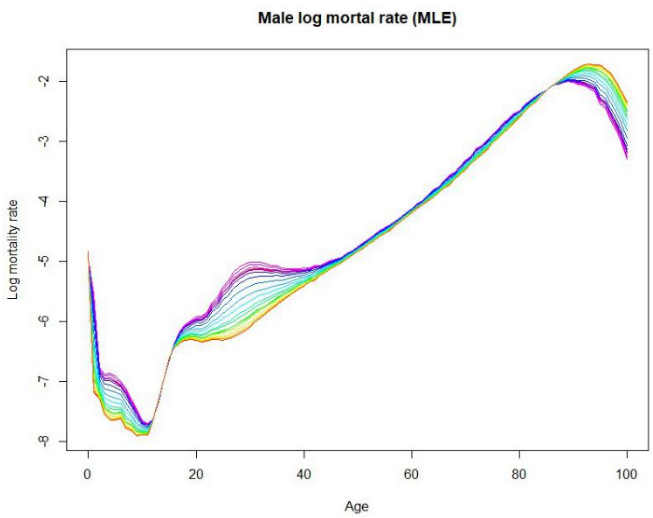


รูปที่ 4.2 แสดงถึงค่าพารามิเตอร์ของเพศชายที่ได้จากการประมาณค่าโดยใช้วิธีภาวะน่าจะเป็นสูงสุด

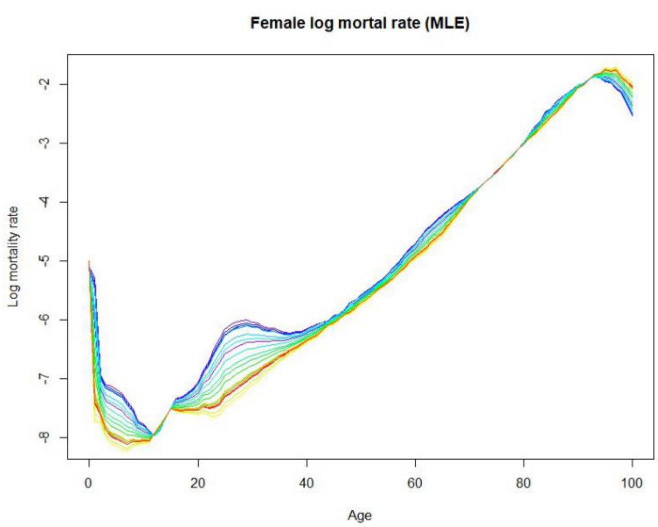


จากรูปที่ 4.1 และ 4.2 จะเห็นว่าค่าพารามิเตอร์ของแอลฟาซึ่งแสดงถึงค่าอายุเฉลี่ยมีแนวโน้มเพิ่มขึ้นเมื่ออายุเพิ่มขึ้น ในทางกลับกันค่าเบต้าที่แสดงถึงค่าเสื่อมของดัชนีเวลามีค่าลดลงเรื่อย ๆ เมื่ออายุเพิ่มขึ้นเรื่อย ๆ ซึ่งจะมีเพิ่มขึ้นบ้างในช่วงอายุแรกเกิด 1-6 ปี ค่าแคปปาซึ่งแสดงถึงดัชนีของเวลามีค่าลดลงเรื่อย ๆ และเมื่อนำมาสร้างเป็นอัตรามรณะรายปีตั้งแต่ พ.ศ. 2540-2559 สามารถสร้างเป็นกราฟได้ดังรูป

รูปที่ 4.3 อัตรามรณะแบบลอการิทึมของรายอายุในแต่ละปีในเพศชาย โดยไล่สีจากม่วงไปแดงตามปี พ.ศ.2540-2559



รูปที่ 4.4 อัตรามรณะแบบลอการิทึมของรายอายุในแต่ละปีในเพศหญิง โดยไล่สีจากม่วงไปแดงตามปี พ.ศ.2540-2559



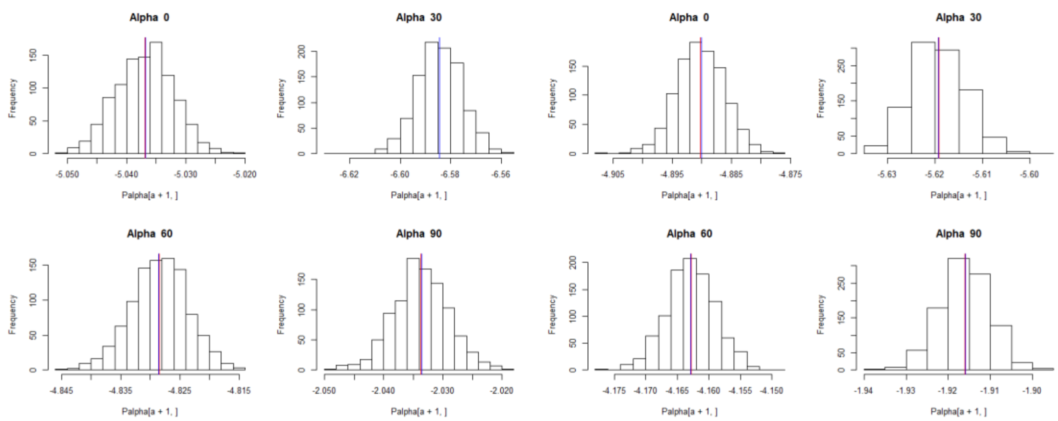
จะเห็นชัดเจนว่าในช่วง ประมาณช่วง 0-15 ปี, 20-40 ปี และ 80-100 ปี อัตรามรณะจะมีความผันผวนเป็นพิเศษ ซึ่งอาจจะเกิดมาจากมาจากในช่วงปี 2540 ทางประเทศไทยได้มีการบันทึกการเสียชีวิตของคนไทยเข้าสู่ระบบคอมพิวเตอร์ ซึ่ง ในช่วงแรกจะมีการปรับค่าโดยเอาของปีห่อนหน้ามาใส่เป็นจำนวนมากทำให้ค่าในช่วงปี 2540 อาจเกิดสูงเกินจริง

4.2 ผลการประมาณค่าโดยใช้วิธีแบบเบย์

เนื่องจากเป็นการสุ่มโดยวิธีการสุ่มมอนติคาร์โลลูกโซ่มาร์คอฟดังนั้นสามารถนำค่าพารามิเตอร์มาแสดงเป็นรูปแบบฮิสโตรแกรมได้ดังรูป

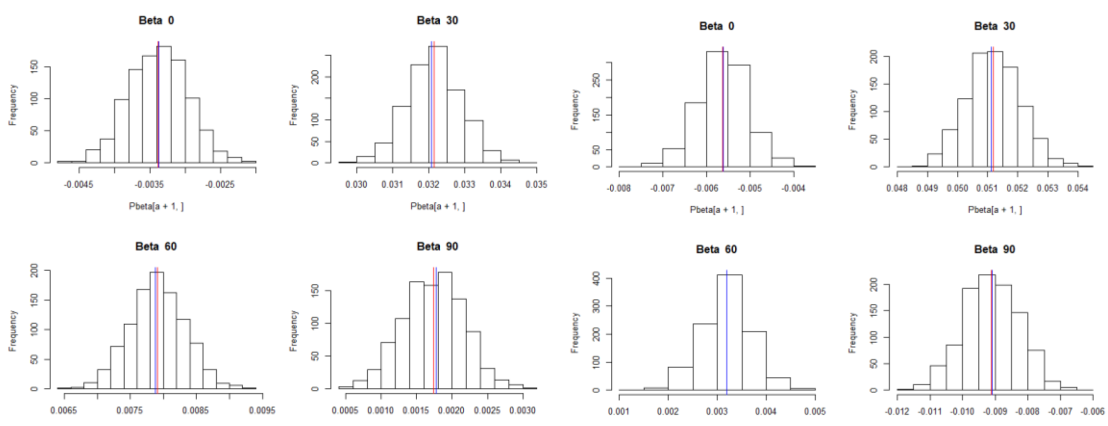
รูปที่ 4.5 แสดงฮิสโตรแกรมของค่าแอลฟาหลังจากการทำการสุ่มมอนติคาร์โลลูกโซ่มาร์คอฟ

เพศหญิง เพศชาย



รูปที่ 4.6 แสดงฮิสโตรแกรมของค่าเบต้าหลังจากการทำการสุ่มมอนติคาร์โลลูกโซ่มาร์คอฟ

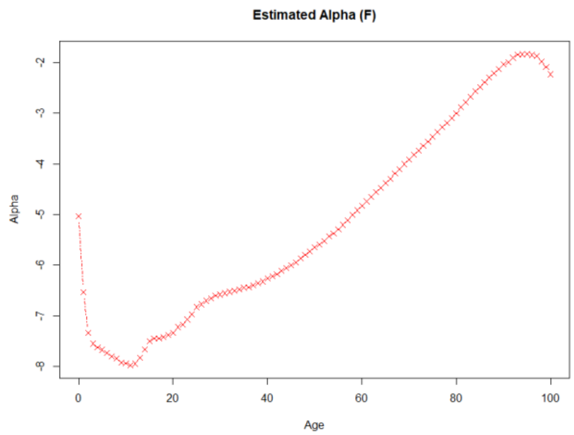
เพศหญิง เพศชาย



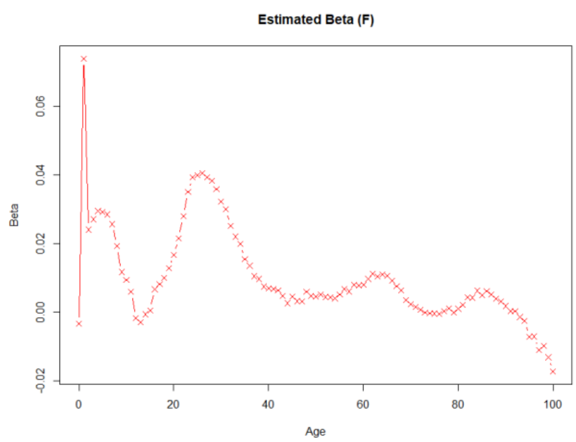
โดยเส้นสีแดงคือค่าเฉลี่ยของการแจกแจงที่สุ่มได้มาส่วนเส้นสีน้ำเงินคือ ค่าพารามิเตอร์ที่ได้จากการประมาณโดยใช้วิธีแบบภาวะน่าจะเป็นสูงสุด จะเห็นได้ว่าส่วนใหญ่ค่ามีลักษณะใกล้เคียงกันระหว่างเบย์กับวิธีภาวะน่าจะเป็นสูงสุด

เมื่อนำค่าผลลัพธ์ของค่าพารามิเตอร์มาหาค่าเฉลี่ยจะสามารถสร้างกราฟได้ดังรูป

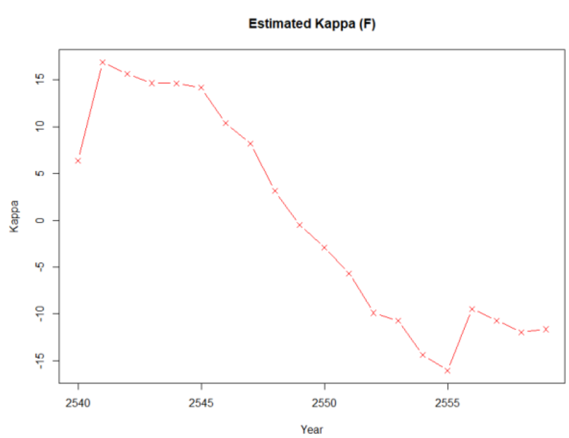
รูปที่ 4.7 แสดงค่าเฉลี่ยของค่าอายุเฉลี่ยหรือค่าแอลฟาของเพศหญิงซึ่งประมาณโดยวิธีแบบเบย์



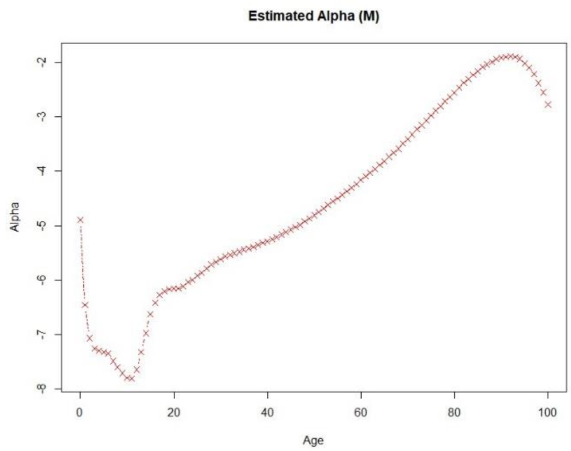
รูปที่ 4.8 แสดงค่าเสื่อมของดัชนีเวลาหรือค่าเบต้าของเพศหญิงซึ่งประมาณโดยวิธีแบบเบย์



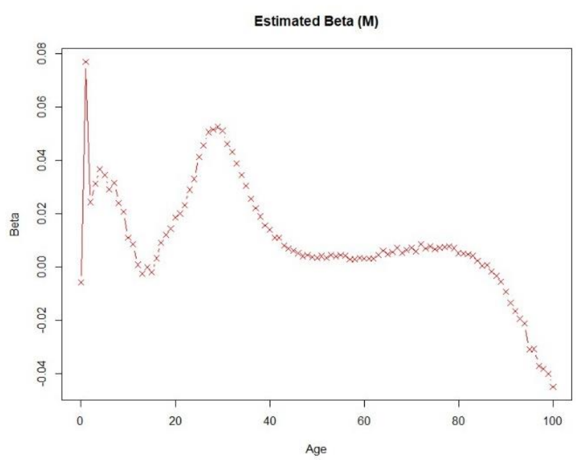
รูปที่ 4.9 แสดงค่าดัชนีเวลาของเพศหญิงหรือค่าแคปปาจากการประมาณโดยวิธีแบบเบย์



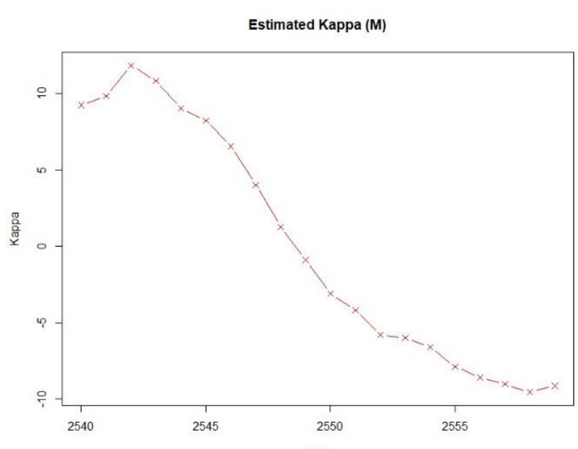
รูปที่ 4.10 แสดงค่าเฉลี่ยของค่าอายุเฉลี่ยหรือค่าแอลฟาของเพศชายซึ่งประมาณโดยวิธีแบบเบย์



รูปที่ 4.11 แสดงค่าเสื่อมของดัชนีเวลาหรือค่าเบต้าของเพศชายซึ่งประมาณโดยวิธีแบบเบย์

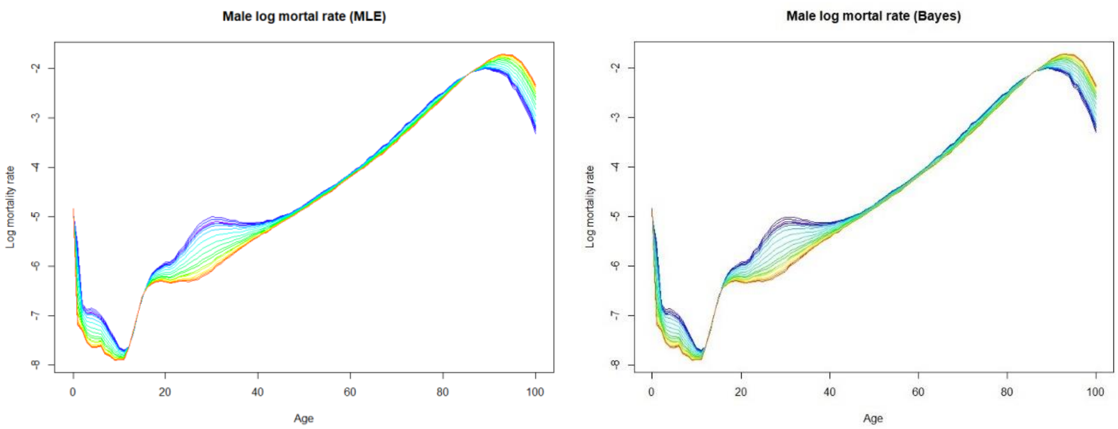


รูปที่ 4.12 แสดงค่าดัชนีเวลาของเพศชายหรือค่าแคปปาจากการประมาณโดยวิธีแบบเบย์

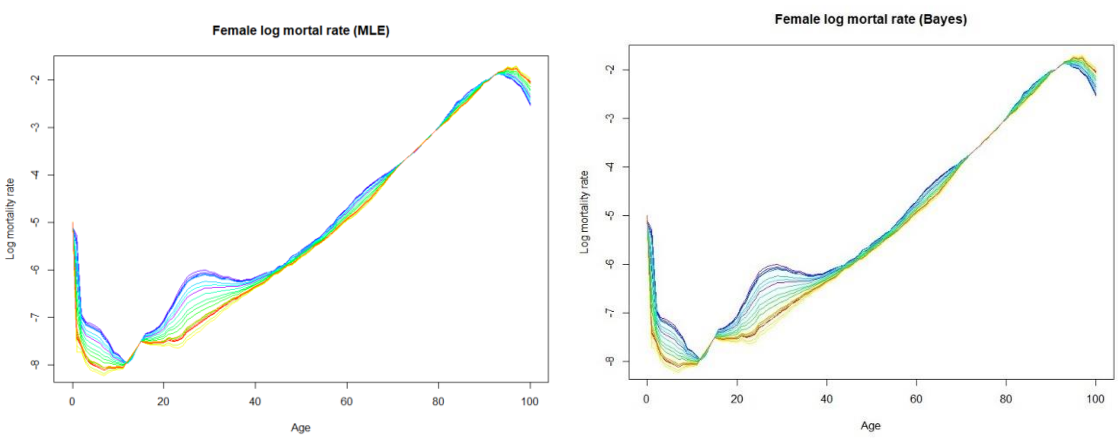


เมื่อนำค่าพารามิเตอร์ต่าง ๆ สามารถนำมาสร้างเป็นกราฟได้ดังรูป

รูปที่ 4.13 อัตรามรณะแบบลอการิทึมของรายอายุในแต่ละปีในเพศชายโดยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์ โดยไล่สีจากม่วงไปแดงตามปี พ.ศ.2540-2559



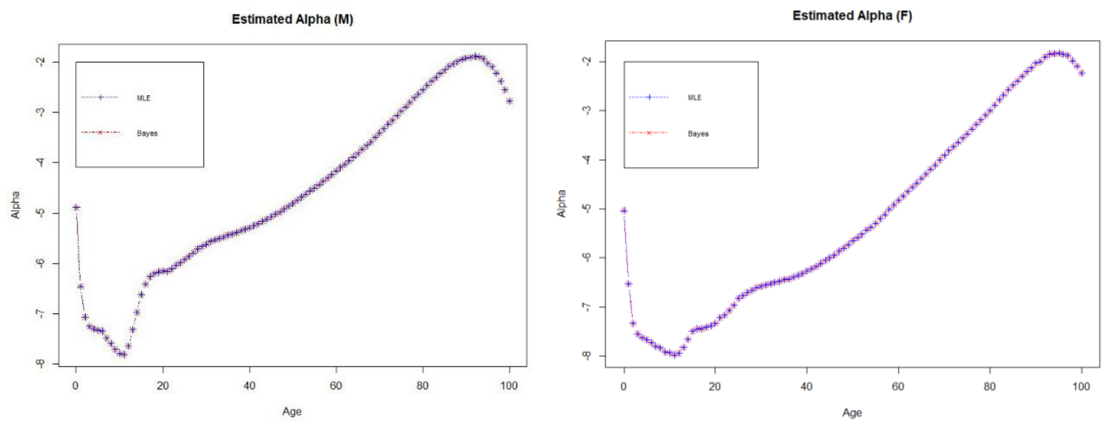
รูปที่ 4.14 อัตรามรณะแบบลอการิทึมของรายอายุในแต่ละปีในเพศหญิงโดยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์ โดยไล่สีจากม่วงไปแดงตามปี พ.ศ.2540-2559



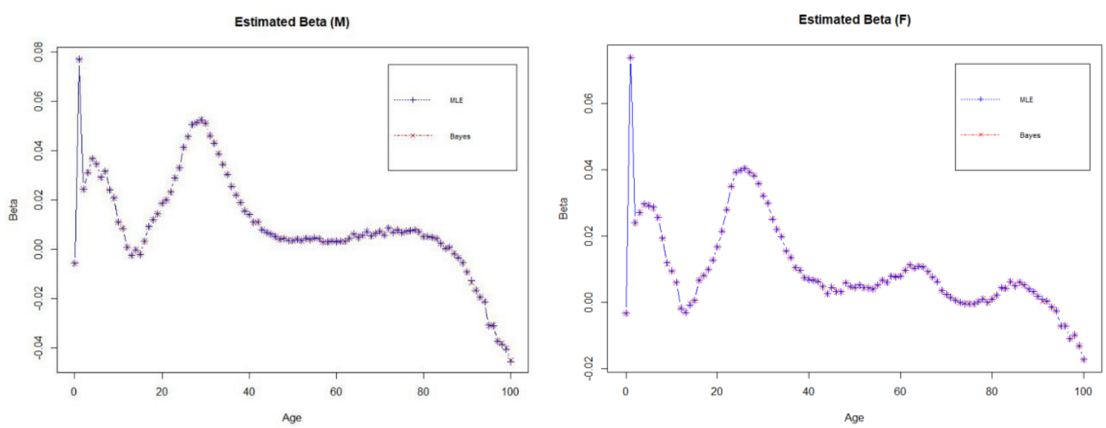
จะเห็นชัดเจนว่าในช่วงค่าอัตรามรณะแบบลอการิทึมจะมีลักษณะกราฟจะมีลักษณะใกล้เคียงกับอัตรามรณะแบบลอการิทึมของวิธีแบบภาวะน่าจะเป็นสูงสุด

โดยเมื่อนำมาเทียบค่าพารามิเตอร์ต่าง ๆ มาเทียบกันระหว่างการประมาณค่าโดยใช้วิธีภาวะน่าจะเป็นสูงสุดกับวิธีแบบเบย์สามารถเทียบได้จากกราฟดังรูป

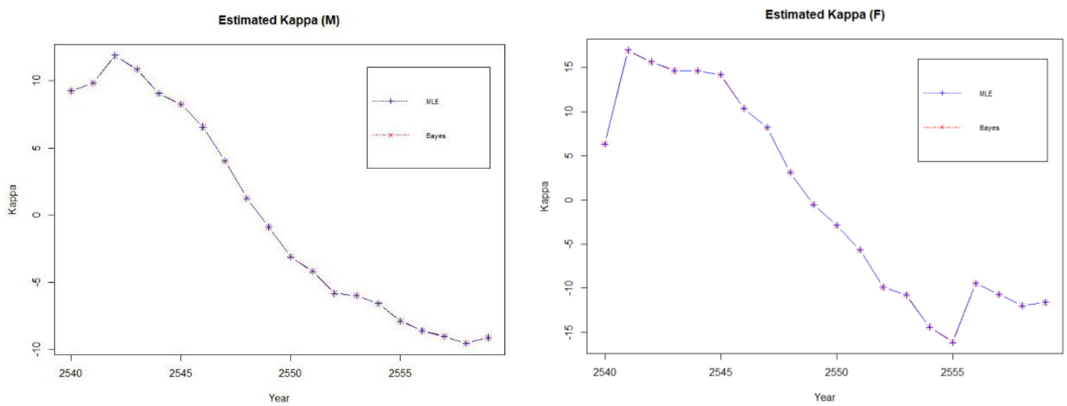
รูปที่ 4.15 แสดงค่าคงที่กลางปีของประชากรจากวิธีแบบเบย์และแบบภาวะน่าจะเป็นสูงสุด



รูปที่ 4.16 แสดงค่าเสื่อมเวลาของประชากรจากวิธีแบบเบย์และแบบภาวะน่าจะเป็นสูงสุด



รูปที่ 4.17 แสดงดัชนีเวลาของประชากรจากวิธีแบบเบย์และแบบภาวะน่าจะเป็นสูงสุด



จากกราฟทั้งสามที่แสดงการเปรียบเทียบค่าพารามิเตอร์ที่ได้ข้างต้นจะเห็นได้ชัดว่าการประมาณค่าพารามิเตอร์โดยใช้วิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์ทำให้ผลลัพธ์ที่ได้ออกมามีลักษณะใกล้เคียงกัน ซึ่งสอดคล้องกับค่าอัตรามรณะแบบลอการิทึมที่มีลักษณะใกล้เคียงกัน

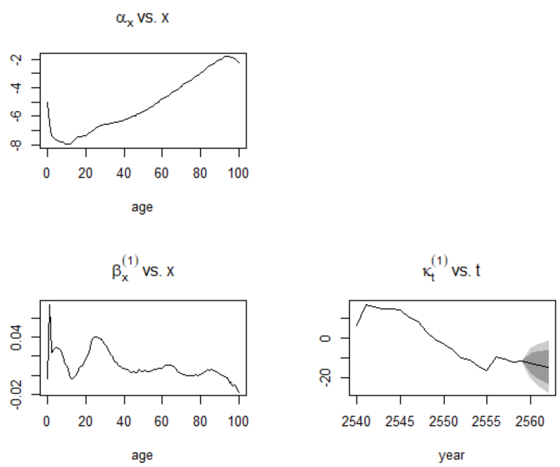
4.3 การพยากรณ์

ผู้ศึกษาได้ทำการพยากรณ์ด้วยหลักวิธีทางอนุกรมเวลาโดยใช้วิธีอาริมาโดยกำหนดให้มีการสร้างตัวแบบให้เหมาะสมกับข้อมูล โดยค่าที่ใช้สร้างตัวแบบอนุกรมเวลาคือค่าดัชนีเวลาหรือแคปปา

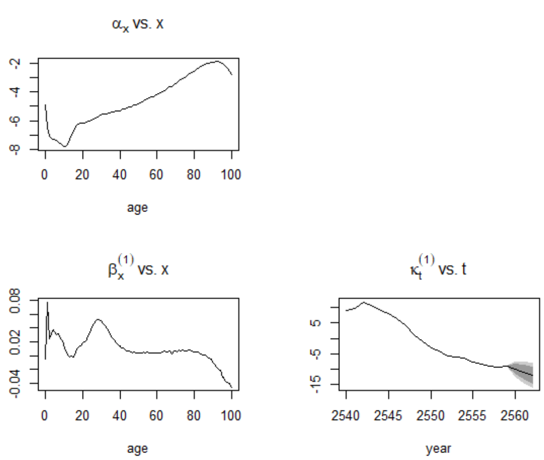
4.3.1 การพยากรณ์ของตัวแบบภาวะน่าจะเป็นสูงสุด

เมื่อนำค่าดัชนีเวลาที่ได้ออกมาจากวิธีการประมาณค่าความน่าจะเป็นสูงสุดมาสร้างตัวแบบอนุกรมเวลา มีลักษณะเป็นรูปแบบอาริมา (1,1,0) โดยจะใช้ตัวแบบอาริมากับทั้งการพยากรณ์กับวิธีการประมาณค่าพารามิเตอร์ทั้งสองแบบ เมื่อทำการพยากรณ์ค่าแคปปาสามารถพยากรณ์ได้ออกมาดังนี้

รูปที่ 4.18 พยากรณ์ค่าแคปปาของอัตรามรณะหญิง



รูปที่ 4.19 พยากรณ์ค่าแคปปาของอัตรามรณะชาย



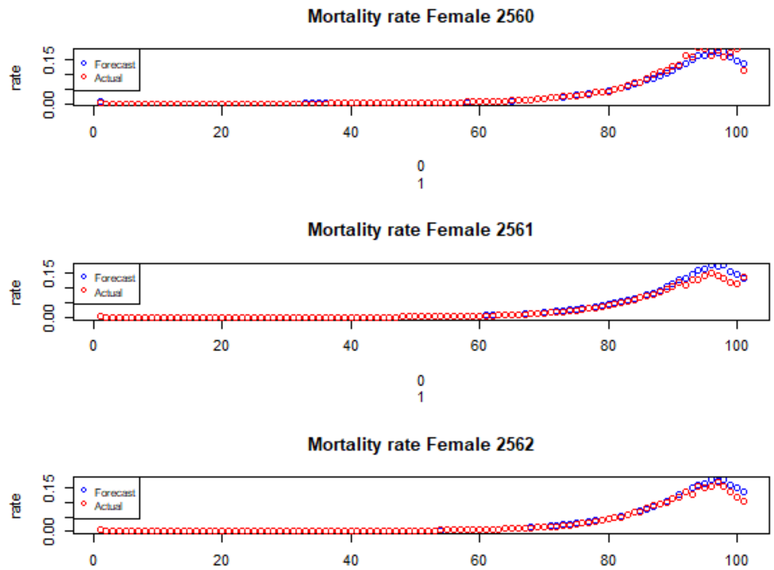
เมื่อนำค่าแคปปาที่ได้จากการพยากรณ์มาทำการคำนวณเป็นอัตรามรณะสามารถแสดงได้เป็นตารางด้านล่างดังนี้

ตารางที่ 4.1 แสดงอัตรามรณะของเพศชายและหญิงของทุก ๆ อายุ 5 ปี โดยวิธีภาวะน่าจะเป็นสูงสุด

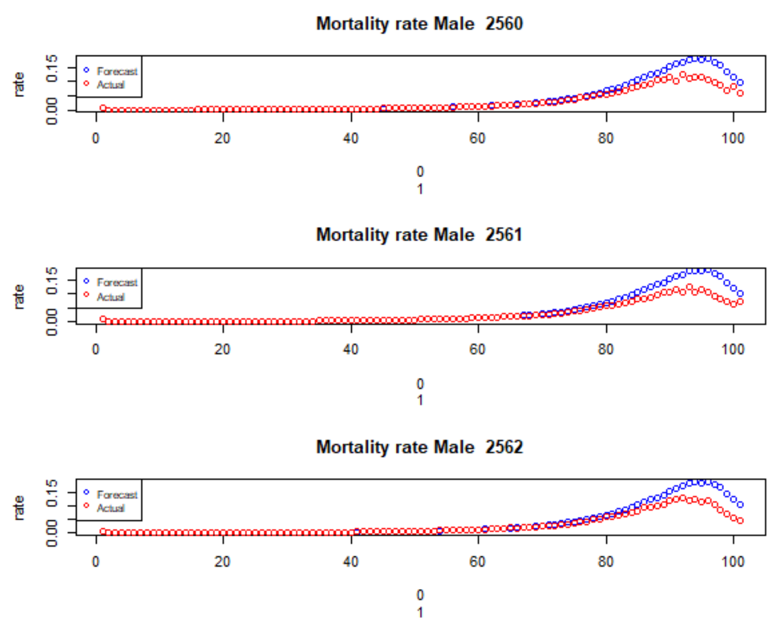
	ชาย			หญิง		
	2560	2561	2562	2560	2561	2562
0	0.007964819	0.008004627	0.008044635	0.007965	0.008005	0.008045
5	0.000429397	0.000413696	0.00039857	0.000429	0.000414	0.000399
10	0.000357445	0.000352951	0.000348513	0.000357	0.000353	0.000349
15	0.001353339	0.001355562	0.001357789	0.001353	0.001356	0.001358
20	0.001708532	0.001675476	0.001643061	0.001709	0.001675	0.001643
25	0.001628665	0.001558923	0.001492168	0.001629	0.001559	0.001492
30	0.001942755	0.001839874	0.00174244	0.001943	0.00184	0.001742
35	0.00300732	0.002912106	0.002819905	0.003007	0.002912	0.00282
40	0.004279278	0.004217957	0.004157514	0.004279	0.004218	0.004158
45	0.005840054	0.00580391	0.005767991	0.00584	0.005804	0.005768
50	0.007816905	0.007786418	0.007756051	0.007817	0.007786	0.007756
55	0.01051113	0.010460126	0.01040937	0.010511	0.01046	0.010409
60	0.014949661	0.014896809	0.014844144	0.01495	0.014897	0.014844
65	0.020771976	0.020657926	0.020544502	0.020772	0.020658	0.020545
70	0.030177917	0.029933705	0.029691469	0.030178	0.029934	0.029691
75	0.046915801	0.046564485	0.0462158	0.046916	0.046564	0.046216
80	0.072472356	0.072008515	0.071547642	0.072472	0.072009	0.071548
85	0.113934154	0.113753727	0.113573586	0.113934	0.113754	0.113574
90	0.161555723	0.162845807	0.164146193	0.161556	0.162846	0.164146
95	0.185962417	0.191497487	0.197197306	0.185962	0.191497	0.197197
100	0.101005755	0.105343463	0.109867454	0.101006	0.105343	0.109867

โดยเมื่อนำมาสร้างกราฟเทียบอัตรามรณะที่พยากรณ์และค่าจริงเป็นได้ออกมาดังรูป

รูปที่ 4.20 อัตรามรณะของเพศหญิงโดยเทียบระหว่างค่าที่พยากรณ์และค่าที่แท้จริงในช่วงปีพ.ศ.2560-2563



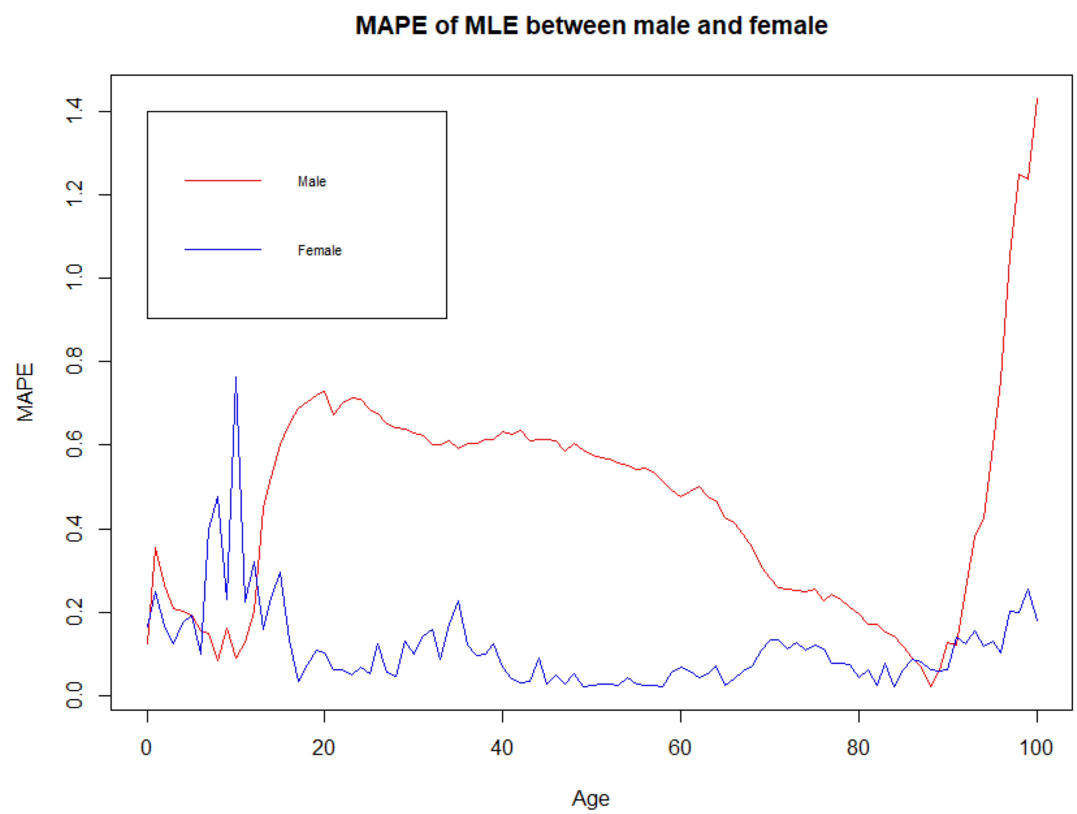
รูปที่ 4.21 อัตรามรณะของเพศชายโดยเทียบระหว่างค่าที่พยากรณ์และค่าที่แท้จริงในช่วงปีพ.ศ.2560-2563



จากรูปที่ 4.20 และ รูปที่ 4.21 จะเห็นได้ว่าค่าการพยากรณ์ทั้งปี 2560, 2561 และ 2562 จะมีค่าสูงกว่าความเป็นจริงในช่วงอายุประมาณ 85 ขึ้นไปค่าพยากรณ์ที่ได้จะมีค่าสูงเกินจริงอย่างมากแต่ในทางกลับกันในช่วงอายุ 0-40 ค่าที่ได้จากการพยากรณ์มีค่าใกล้เคียงกับค่าจริงมาก

ซึ่งเมื่อวัดค่า MAPE ออกมาจะได้ค่าดังตารางข้างต้นดังนี้

รูปที่ 4.22 แสดงค่า MAPE ของเพศชายและหญิงโดยวิธีภาวะน่าจะเป็นสูงสุด



โดยค่า MAPE แสดงถึงประสิทธิภาพของการพยากรณ์ซึ่งค่า MAPE ที่น้อยแสดงถึงความสามารถในการพยากรณ์ที่ดี จะเห็นได้ว่าค่า MAPE เมื่อเทียบกันระหว่างหญิงและชายในเกือบทุกช่วงอายุ ค่าของเพศหญิงมีค่า MAPE ที่ต่ำกว่า

4.3.1 การพยากรณ์ของตัวแบบเบย์

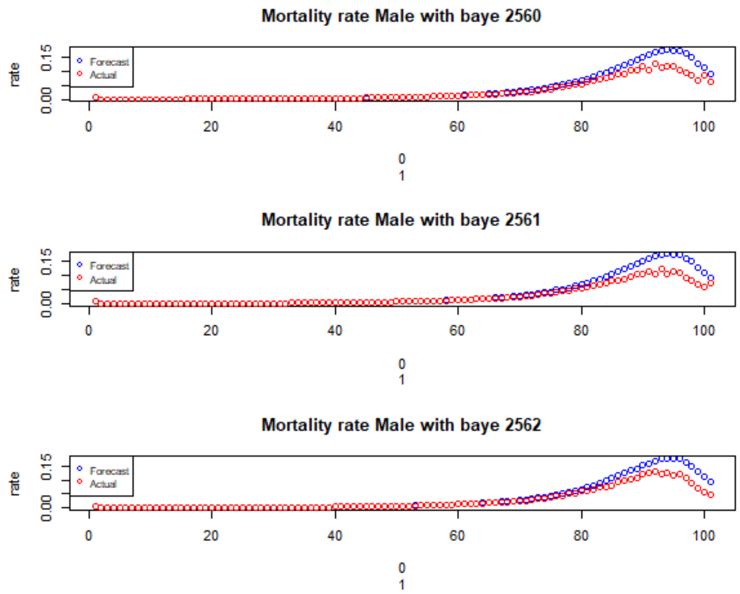
เมื่อนำค่าดัชนีเวลาที่เกิดจากการประมาณค่าแบบเบย์มาสร้างตัวแบบอนุกรมเวลา จะได้ตัวแบบอนุกรมเวลาที่มีลักษณะเป็น อาริมา (1,1,0) หรือก็คือตัวแบบออโต้รีเกรซซีฟที่ 1 (AR(1)) เมื่อนำไปพยากรณ์ค่าอัตรามรณะจะได้เป็นผลดังนี้

ตารางที่ 4.2 แสดงอัตรามรณะเพศชายและหญิงของทุก ๆ อายุ 5 ปี โดยวิธีแบบเบย์

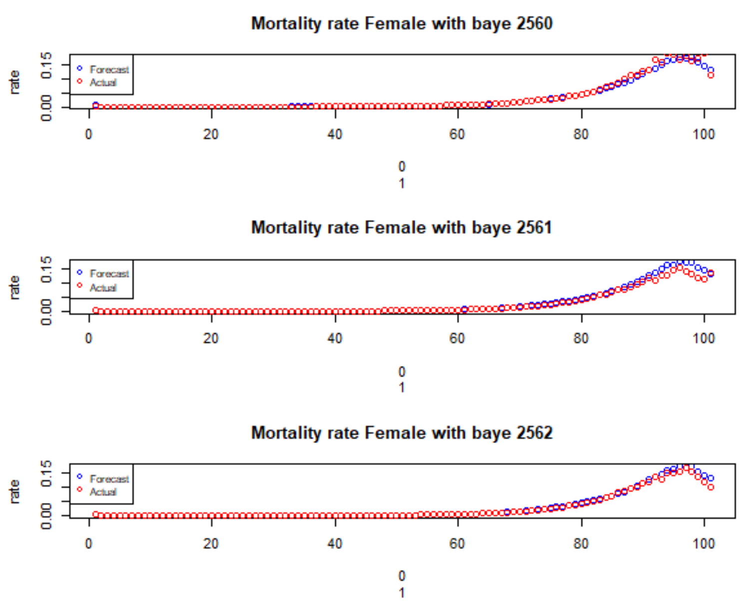
	ชาย			หญิง		
	2560	2561	2562	2560	2561	2562
0	0.007903	0.007886	0.007893	0.006691	0.006682	0.006686
5	0.000492	0.000489	0.000485	0.000363	0.000361	0.000359
10	0.000376	0.000375	0.000376	0.000329	0.000328	0.000329
15	0.001351	0.00135	0.00135	0.000551	0.000551	0.000551
20	0.001823	0.001818	0.00181	0.000565	0.000563	0.000561
25	0.001881	0.001864	0.001893	0.000767	0.000761	0.000772
30	0.002308	0.002353	0.002334	0.00104	0.001053	0.001048
35	0.00336	0.003344	0.003321	0.00139	0.001386	0.001382
40	0.004473	0.004459	0.004483	0.001788	0.001786	0.00179
45	0.005938	0.005952	0.005946	0.002375	0.002379	0.002378
50	0.007929	0.007924	0.007918	0.003371	0.003368	0.003365
55	0.010673	0.010662	0.010681	0.004775	0.00477	0.004779
60	0.01513	0.015149	0.015141	0.007459	0.007481	0.007472
65	0.021222	0.021206	0.021183	0.011389	0.01137	0.011344
70	0.031079	0.031029	0.031116	0.019557	0.019547	0.019564
75	0.048119	0.048243	0.048191	0.031167	0.031162	0.031164
80	0.074466	0.074404	0.074319	0.049407	0.049399	0.049389
85	0.115309	0.115298	0.115316	0.080225	0.080139	0.080287
90	0.159544	0.158994	0.159226	0.128861	0.128946	0.12891
95	0.17209	0.172938	0.174105	0.170819	0.171017	0.171289
100	0.091953	0.09286	0.091295	0.12429	0.124763	0.123945

โดยเมื่อนำมาสร้างกราฟเทียบกันระหว่างค่าที่พยากรณ์กับค่าที่เกิดขึ้นจริงสามารถเป็นกราฟได้ดังรูป

รูปที่ 4.23 อัตรามรณะของเพศชายโดยเทียบระหว่างค่าที่พยากรณ์ของวิธีแบบเบย์และค่าที่แท้จริงในช่วงปีพ.ศ.2560-2563



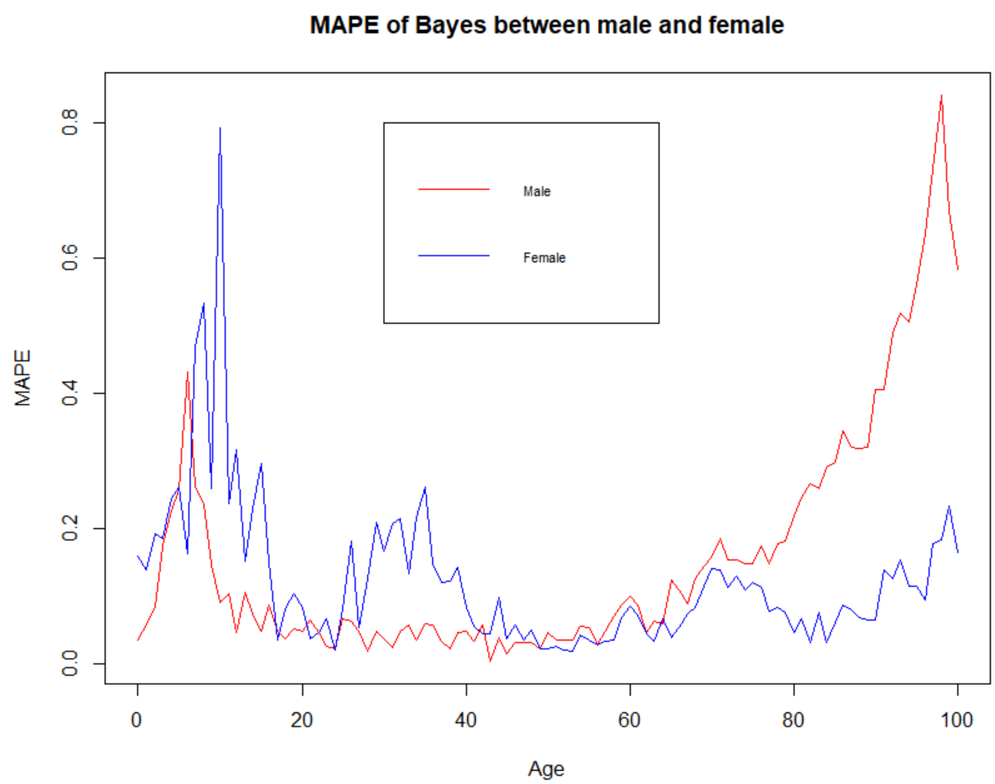
รูปที่ 4.24 อัตรามรณะของเพศหญิงโดยเทียบระหว่างค่าที่พยากรณ์ของวิธีแบบเบย์และค่าที่แท้จริงในช่วงปีพ.ศ.2560-2563



จากกราฟที่ 4.23 และ 4.24 จะเห็นได้ว่ามีค่าการพยาพรณ์กับค่าที่แท้จริงจะมีลักษณะใกล้เคียงกันในช่วงอายุ 0-85 แต่ว่าหลังจาก 85 ปีเป็นต้นไปจะเห็นได้ชัดว่าค่าที่พยากรณ์ได้โดยเฉพาะเพศชายมีลักษณะสูงเกินจริงมากกว่าเพศหญิง

เมื่อนำค่าพยากรณ์ที่ได้มาคำนวณค่า MAPE สามารถแสดงได้เป็นตารางดังตารางด้านล่าง

รูปที่ 4.25 แสดงค่า MAPE ของเพศชายและหญิงโดยวิธีการแบบเบย์

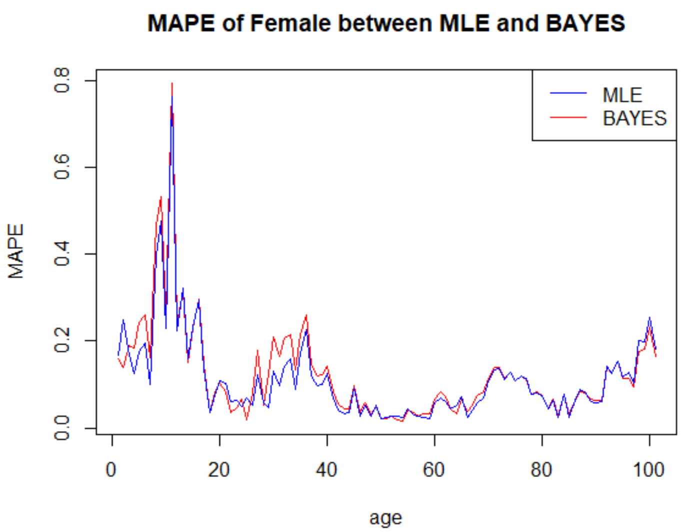


จากตารางจะเห็นชัดเจนว่าค่า MAPE ของเพศหญิงนั้นจะต่ำกว่าของเพศชายในช่วงอายุประมาณ 80-100 ปี ซึ่งมีลักษณะคล้ายกับค่า MAPE ในการประมาณแบบภาวะน่าจะเป็นสูงสุด แต่ในทางกลับกันช่วงอายุเริ่มต้น ค่า MAPE ของเพศชายกลับมีลักษณะต่ำกว่าเพศหญิง

4.4 การเปรียบเทียบความสามารถในการพยากรณ์

เมื่อนำค่า MAPE มาสร้างกราฟสามารถสร้างเป็นเส้นได้ดังรูป

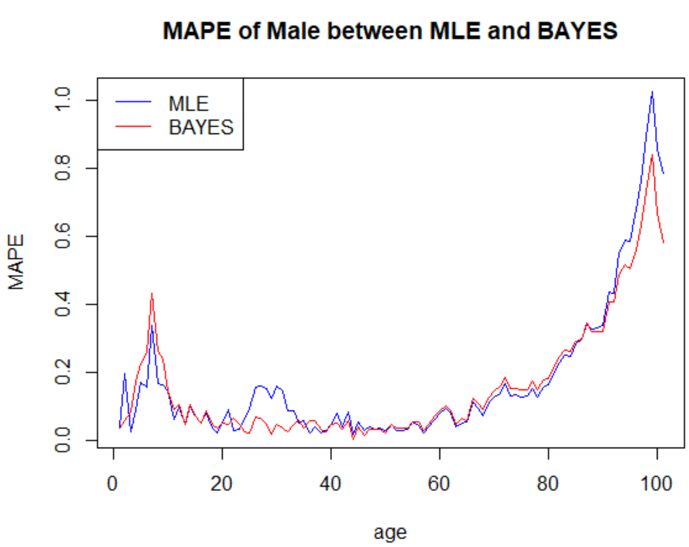
รูปที่ 4.26 รูปแสดงค่า MAPE ของเพศหญิงเทียบกันระหว่างวิธีแบบเบย์และวิธีภาวะน่าจะเป็นสูงสุด



จะเห็นได้ว่าค่า MAPE มีลักษณะใกล้เคียงกันแต่ในช่วงค่าอายุระหว่าง 20 ถึง 40 ปี ค่า MAPE ของวิธีแบบเบย์มีค่าสูงกว่า

โดยเมื่อนำค่าเฉลี่ย MAPE ที่ได้ของวิธีภาวะความน่าจะเป็นสูงสุดคือ 0.11 และวิธีแบบเบย์คือ 0.12

รูปที่ 4.27 รูปแสดงค่า MAPE ของเพศชายเทียบกันระหว่างวิธีแบบเบย์และวิธีภาวะความน่าจะเป็นสุงสุด



จะเห็นได้ว่าค่า MAPE มีลักษณะใกล้เคียงกันแต่ในช่วงค่าอายุระหว่าง 20 ถึง 40 ปี ค่า MAPE ของวิธีภาวะน่าจะเป็นสูงสุดนั้นสูงกว่าวิธีแบบเบย์ และในช่วงอายุประมาณ 85 ปีขึ้นไป ค่า MAPE ของวิธีภาวะน่าจะเป็นสูงสุดสูงกว่าวิธีแบบเบย์ โดยเมื่อนำค่าเฉลี่ย MAPE ที่ได้ของวิธีภาวะความน่าจะเป็นสูงสุดคือ 0.17 และวิธีแบบเบย์คือ 0.15

บทที่ 5

สรุปและข้อเสนอแนะ

5.1 สรุปผลการศึกษา

ในงานศึกษาชิ้นนี้ผู้ศึกษาต้องการเปรียบเทียบประสิทธิภาพของการประมาณค่าพารามิเตอร์ของแต่ละวิธีโดยใช้ข้อมูลจำนวนประชากรกลางปีและการตายของประชากรประเทศไทยตั้งแต่ปี พ.ศ. 2540-2562 โดยในวิธีแบบเบย์ ผู้ศึกษาทำการกำหนดการแจกแจงก่อนการสังเกตอ้างอิงตามงานวิจัยของ Claudia, Antone และ Michel เพื่อความสะดวกในการคำนวน ในส่วนของวิธีแบบภาวะน่าจะเป็นสูงสุดได้มีการกำหนดการตายแจกแจงแบบปัวซงค์ เมื่อนำค่าพารามิเตอร์ที่ถูกประมาณโดยสองวิธีมาเปรียบเทียบกัน ค่าพารามิเตอร์มีลักษณะใกล้เคียงกัน โดยค่าคงที่กลางปีของประชากรหรือแอลฟานั้นมีลักษณะมากในช่วงอายุ 0-5 ปีและมีลักษณะลดลงในช่วง 5-20 ปี และมีลักษณะเพิ่มขึ้นเรื่อย ๆ ในขณะที่ค่าอัตราเสื่อมของดัชนีเวลาหรือค่าเบต้านั้นมีลักษณะค่อนข้างแกว่งในช่วงอายุ 0 ถึง 40 ปี หลังจากนั้นจะมีค่าลดลงเรื่อย ๆ ในส่วนของค่าดัชนีเวลาหรือแคปปา มีลักษณะลดลงเรื่อย ๆ ตามปีที่เพิ่มขึ้น เมื่อนำค่าพารามิเตอร์มาคำนวณอัตรามรณะในแต่ละปี พบว่าในช่วง 0 ถึง 15 ปี และก็ 20-40 ปี อัตรามรณะมีลักษณะที่ค่อนข้างแกว่ง จากการประมาณค่าพารามิเตอร์จะเห็นได้ว่าทั้งสองวิธีนั้นค่าพารามิเตอร์ที่ได้มีลักษณะใกล้เคียงกันค่อนข้างมาก

เมื่อนำค่าพารมิเตอร์มาทำการพยากรณ์ ด้วยตัวแบบ ARIMA พบว่าค่าแคปปาของทั้งสองแบบมีลักษณะเป็น ARIMA(1,1,0) เมื่อนำค่าไปพยากรณ์พบว่า อัตรามรณะที่ได้จากการพยากรณ์จากวิธีแบบเบย์และวิธีภาวะน่าจะเป็นสูงสุดในเพศชายนั้นมีลักษณะไปในทิศทางเดียวกัน และเมื่อพิจารณาค่า MAPE ของเพศชาย พบว่าในช่วงอายุ 85 ปี มีค่า MAPE ที่สูงขึ้นอย่างเห็นได้ชัด ในขณะเดียวกันเมื่อพิจารณาเพศหญิงพบว่าค่าอัตรามรณะที่พยากรณ์ได้และค่า MAPE ที่เกิดขึ้นในช่วงอายุ 85 ปีขึ้นไปมีลักษณะต่ำกว่าเพศชาย แต่ในช่วงอายุแรกเกิดจนถึงประมาณ 20 ปี ค่า MAPE ที่ได้นั้นมีลักษณะค่อนข้างสูงเมื่อเทียบกับอายุช่วงอื่น และเป็นที่สังเกตว่าค่า MAPE ที่ได้จากการประมาณแบบเบย์ในเพศหญิงนั้นมีลักษณะที่ค่าสูงกว่าการใช้วิธีแบบภาวะน่าจะเป็นสูงสุด

5.2 อภิปรายผล

จากผลการศึกษาจะเห็นได้ชัดว่าวิธีการประมาณค่าแบบเบย์และแบบภาวะน่าจะเป็นสูงสุดนั้นให้ลักษณะค่าพารามิเตอร์ที่ไม่ต่างกันมากอย่างชัดเจน ส่งผลให้อัตรามรณะที่เกิดขึ้นมีลักษณะไม่ต่างกันมากและทั้งคู่มีความผันผวนในช่วงอายุ 0-15 ปี 20-40 ปี และ 85 ปีขึ้นไป ซึ่งมีสาเหตุมาจากการปรับค่าอัตรามรณะในช่วงปี 2540-2545 เพราะว่าในขณะนั้น ได้เริ่มมีการบันทึกการเสียชีวิตลงระบบคอมพิวเตอร์ทำให้อาจจะให้เกิดค่าที่สูงเกินจริงหรืออาจจะเกิดจากการแพร่ระบาดของโรคหรือเหตุการณ์ด้านภัยพิบัติทางธรรมชาติ ในส่วนของการพยากรณ์จะเห็นได้ว่าการพยากรณ์นั่นค่า MAPE เฉลี่ยมีลักษณะใกล้เคียงกันมากโดยที่ค่า MAPE ของ ผู้หญิงนั้นจะมีค่าต่ำกว่าของผู้ชาย แต่เมื่อพิจารณาลงแยกแล้วในเพศหญิงวิธีประมาณโดยใช้วิธีภาวะน่าจะเป็นสูงสุดจะได้ค่า MAPE ต่ำกว่าวิธีแบบเบย์ ในทางกลับกันในเพศชาย วิธีแบบเบย์จะให้ค่า MAPE เฉลี่ยต่ำกว่าวิธีประมาณแบบภาวะน่าจะเป็นสูงสุด ซึ่งจากผลลัพธ์แสดงให้เห็นว่าวิธีแบบภาวะความน่าจะเป็นสุงสุดและวิธีแบบเบย์นั้นไม่ได้มีลักษณะที่แตกต่างกัน แต่วิธีแบบเบย์นั้นเนื่องจากทางผู้ศึกษาได้นำรูปแบบการแจกแจงที่ใช้กับประชากรประเทศฝรั่งเศส ดังนั้นถ้าหากเป็นไปได้เรายังสามารถเลือกปรับ

ค่าคงที่อยู่ในตัวพารามิเตอร์ที่ทำการประมาณได้ อาทิเช่น a_x, b_x เป็นต้นเพื่อปรับให้เหมาะสมกับลักษณะประชากรประเทศไทย หรือเปลี่ยนรูปแบบกลักษณะค่าพารามิเตอร์หรือการแจกแจงได้

ข้อเสนอแนะ

1. ในการพิจารณาเรื่องการเพิ่มลดอัตราการมรณะของประเทศไทยอาจจะนำเหตุผลด้านการแพทย์หรือเศรษฐกิจมาใช้ช่วยในการอธิบาย

2. ในการเลือกปีที่ใช้ในการคำนวณอัตรามรณะนั้น เนื่องจากข้อมูลในช่วงปี พ.ศ. 2540-2545 อาจจะมีปัญหาจึงแนะนำให้ใช้ชุดข้อมูลที่ใหม่และล่าสุดเพื่อป้องกันการประมาณที่ผิดพลาด

3. ในการใช้โปรแกรม R ในการประมาณแบบเบย์ ผู้ศึกษาคาดว่าตัวอัลกอริทึมสามารถนำไปพัฒนาต่อเพื่อให้มีความรวดเร็วในการประมวลผลที่รวดเร็วมากขึ้นได้

4. ศึกษาเพิ่มเติมเรื่องการปรับพารามิเตอร์ให้เข้ากับลักษณะประชากรไทย

5. ศึกษาเพิ่มเติมเกี่ยวกับความเอนเอียง (Bias) ของพารามิเตอร์ต่าง ๆ เพื่อสังเกตผลกระทบและความเอนเอียงที่เกิดขึ้นกับตัวแบบลี-คาร์เตอร์

6. เนื่องจากอัตรามรณะแบบลอการิทึมของรายอายุตั้งแต่ 85 ถึง 90 ปีเป็นต้นไปมีแนวโน้มลดลง ซึ่งขัดแย้งกับข้อเท็จจริงที่ว่าอัตรามรณะของผู้สูงอายุจะมีแนวโน้มสูงขึ้นเรื่อย ๆ ดังนั้นข้อมูลที่ใช้ในการสร้างตัวแบบลี-คาร์เตอร์จึงควรศึกษาช่วงอายุระหว่าง 0 ถึง 85 ปีเท่านั้น และศึกษาเพิ่มเติมเกี่ยวกับตัวแบบอื่นที่ใช้ในการประมาณและพยากรณ์อัตรามรณะสำหรับช่วงวัยสูงอายุดังกล่าวโดยเฉพาะ

บรรณานุกรม

*Ronald Demon Lee (*1992). *THE LEE-CARTER METHOD FOR FORECASTING MORTALITY, WITH VARIOUS EXTENSIONS AND APPLICATIONS. Departments of Demography and Economics at the University of California*

*Natacha Brouhns, Michel Denuit, Jeroen K. Vermunt (*2002). *A Poisson log-bilinear regression approach to the construction of projected life tables. Institut de Statistique, Université Catholique de Louvain, Voie du Roman Pays,* 20, *B-*1348 *Louvain-la-Neuve, Belgium*

*Czado Claudia*1, *Delwarde Antoine*2, *Denuit,Michel*3 (2004) : *Bayesian Poisson log-bilinear mortality projections,*1 *Technische Universit¨at M¨unchen D-*85748 *Garching bei Munich, Germany,*2 *Institut des Sciences Actuarielles Universit´ e Catholique de Louvain B-*1348 *Louvain-la-Neuve, Belgium,*3 *Institut de Statistique Universit´ e Catholique de Louvain B-*1348 *Louvain-la-Neuve, Belgium Alho,Spencer (*2005). *Statistical Demography and Forecasting. University of Helsinki.*

*Heather Booth(*2008)*, MORTALITY MODELLING AND FORECASTING: A REVIEW OF METHODS Australian Demographic and Social Research Institute, Coombs Building* 9*, Australian National University, ACT* 0200*, Australia.*

*Arkadiusz Wiśniowski*1 *& Peter W. F. Smith & Jakub Bijak*1 *& James Raymer*2 *& Jonathan J. Forster*1 *MORTALITY MODELLING AND FORECASTING: A REVIEW OF METHODS.* 1 *EconomicandSocialResearchCouncilCentreforPopulationChange, University of Southampton, Highfield, SO*17 1*BJ Southampton, UK* 2 *Australian Demographic & Social Research Institute, The Australian National University, Acton ACT* 2601*, Australia*

จันท์ธิดา บุญมหาสิทธิ์ และสำรวม จงเจริญ (2016). การเปรียบเทียบการใช้ตัวแบบอัตรามรณะ เพื่อการพยากรณ์อัตรามรณะโดยข้อมูลจากประเทศไทย . คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ ถนนเสรีไทย แขวงคลองจั่น เขตบางกะปิ กรุงเทพมหานคร 10240

ณัฐสุรางค์ ยาสูงเนิน*, Lee-Carter Model and Extension to Forecast Thai Mortality Rate,* ภาควิชา วิทยาการคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยวงษ์ชวลิตกุล บ้านเกาะเมือง นครราชสีมา 3000

ภาคผนวก

ภาคผนวก ก

ค่าประมาณพารามิเตอร์ด้วยวิธีแบบเบย์

ตาราง ก1 ค่าประมาณแอลฟาด้วยวิธีแบบเบย์สำหรับเพศชาย

อายุ x (ปี)	$\hat{\alpha}_x$	อายุ x (ปี)	$\hat{\alpha}_x$	อายุ x (ปี)	$\hat{\alpha}_x$	อายุ x (ปี)	$\hat{\alpha}_x$
0	-4.890226	26	-5.860528	52	-4.682576	78	-2.710797
1	-6.460265	27	-5.789821	53	-4.618858	79	-2.629630
2	-7.066161	28	-5.715116	54	-4.551721	80	-2.553341
3	-7.253387	29	-5.666249	55	-4.499984	81	-2.462058
4	-7.298221	30	-5.619234	56	-4.437734	82	-2.374960
5	-7.325282	31	-5.562089	57	-4.367542	83	-2.303915
6	-7.348444	32	-5.533263	58	-4.302032	84	-2.224915
7	-7.488050	33	-5.504055	59	-4.232268	85	-2.156486
8	-7.594450	34	-5.477885	60	-4.162864	86	-2.078047
9	-7.711259	35	-5.438138	61	-4.091876	87	-2.036855
10	-7.791238	36	-5.417802	62	-4.030235	88	-1.990654
11	-7.809197	37	-5.387435	63	-3.956390	89	-1.940293
12	-7.644685	38	-5.348962	64	-3.882240	90	-1.915987
13	-7.320006	39	-5.315634	65	-3.811908	91	-1.900883
14	-6.975672	40	-5.288510	66	-3.730740	92	-1.882751
15	-6.624264	41	-5.247775	67	-3.657902	93	-1.897218
16	-6.411834	42	-5.211556	68	-3.586272	94	-1.930846
17	-6.269996	43	-5.163684	69	-3.492748	95	-2.019645
18	-6.203079	44	-5.118914	70	-3.407974	96	-2.091611
19	-6.167676	45	-5.071796	71	-3.323418	97	-2.217534
20	-6.149010	46	-5.022985	72	-3.230041	98	-2.377620
21	-6.159004	47	-4.982504	73	-3.152126	99	-2.548010
22	-6.113273	48	-4.919700	74	-3.064663	100	-2.773269
23	-6.036809	49	-4.865543	75	-2.973942		
24	-5.993011	50	-4.807026	76	-2.885384		
25	-5.919262	51	-4.743784	77	-2.800438		

ตาราง ก2 ค่าประมาณเบต้าด้วยวิธีแบบเบย์สำหรับเพศชาย

อายุ x (ปี)	$\hat{\beta}_x$	อายุ x (ปี)	$\hat{\beta}_x$	อายุ x (ปี)	$\hat{\beta}_x$	อายุ x (ปี)	$\hat{\beta}_x$
0	-0.005627	26	0.045708	52	0.003498	78	0.007860
1	0.077060	27	0.050635	53	0.004485	79	0.007123
2	0.024412	28	0.051507	54	0.004040	80	0.005214
3	0.031215	29	0.052514	55	0.004647	81	0.005197
4	0.036701	30	0.051194	56	0.004366	82	0.004845
5	0.034544	31	0.046231	57	0.003034	83	0.004395
6	0.029135	32	0.043217	58	0.003002	84	0.002394
7	0.031589	33	0.038863	59	0.003387	85	0.000425
8	0.023969	34	0.034498	60	0.003192	86	0.000819
9	0.020840	35	0.030481	61	0.003250	87	-0.001742
10	0.011128	36	0.025623	62	0.003270	88	-0.003336
11	0.008514	37	0.022062	63	0.004459	89	-0.005453
12	0.000927	38	0.019029	64	0.006256	90	-0.009120
13	-0.002367	39	0.015637	65	0.004826	91	-0.013417
14	-0.000093	40	0.014081	66	0.005708	92	-0.016560
15	-0.001968	41	0.011037	67	0.007254	93	-0.019344
16	0.003356	42	0.011098	68	0.005440	94	-0.021218
17	0.009249	43	0.008004	69	0.006452	95	-0.030745
18	0.012032	44	0.006931	70	0.007342	96	-0.030696
19	0.014405	45	0.006181	71	0.005909	97	-0.037080
20	0.018721	46	0.005229	72	0.008682	98	-0.038226
21	0.020115	47	0.004153	73	0.006948	99	-0.039979
22	0.023237	48	0.004524	74	0.007819	100	-0.044905
23	0.029010	49	0.003681	75	0.006809		
24	0.033102	50	0.003580	76	0.007365		
25	0.041429	51	0.004148	77	0.007632		

ตาราง ก3 ค่าประมาณแคปปาด้วยวิธีแบบเบย์สำหรับเพศชาย

ปี	$\hat{\kappa}_t$
2540	9.244805
2541	9.824172
2542	11.833440
2543	10.814437
2544	9.036372
2545	8.225570
2546	6.541674
2547	4.021701
2548	1.249533
2549	-0.893956
2550	-3.101402
2551	-4.196515
2552	-5.814657
2553	-6.002381
2554	-6.597534
2555	-7.887261
2556	-8.592860
2557	-9.034445
2558	-9.539569
2559	-9.131123

ตาราง ก4 ค่าประมาณแอลฟาด้วยวิธีแบบเบย์สำหรับเพศหญิง

อายุ x (ปี)	$\hat{\alpha}_x$	อายุ x (ปี)	$\hat{\alpha}_x$	อายุ x (ปี)	$\hat{\alpha}_x$	อายุ x (ปี)	$\hat{\alpha}_x$
0	-5.036940	26	-6.773037	52	-5.526135	78	-3.190200
1	-6.532385	27	-6.706400	53	-5.436551	79	-3.096274
2	-7.338234	28	-6.660111	54	-5.381406	80	-2.999764
3	-7.551091	29	-6.608098	55	-5.300124	81	-2.886937
4	-7.629736	30	-6.584457	56	-5.204868	82	-2.783613
5	-7.673876	31	-6.554679	57	-5.117825	83	-2.681621
6	-7.732601	32	-6.529338	58	-5.010757	84	-2.570695
7	-7.806702	33	-6.506102	59	-4.919660	85	-2.480912
8	-7.840103	34	-6.479546	60	-4.828482	86	-2.396020
9	-7.922890	35	-6.447675	61	-4.743429	87	-2.300233
10	-7.939600	36	-6.437338	62	-4.649172	88	-2.209218
11	-7.981169	37	-6.394637	63	-4.561313	89	-2.130544
12	-7.947081	38	-6.358289	64	-4.477501	90	-2.033708
13	-7.825959	39	-6.325562	65	-4.385542	91	-1.992228
14	-7.664369	40	-6.266776	66	-4.295366	92	-1.905404
15	-7.500178	41	-6.225823	67	-4.192408	93	-1.845442
16	-7.449049	42	-6.175947	68	-4.108667	94	-1.838355
17	-7.448691	43	-6.112837	69	-4.006215	95	-1.828599
18	-7.416429	44	-6.056583	70	-3.914525	96	-1.851402
19	-7.386237	45	-6.003328	71	-3.821515	97	-1.877153
20	-7.338393	46	-5.944654	72	-3.736943	98	-1.982095
21	-7.223410	47	-5.866940	73	-3.645355	99	-2.092242
22	-7.172706	48	-5.800448	74	-3.562611	100	-2.234728
23	-7.074222	49	-5.732093	75	-3.472659		
24	-6.969482	50	-5.654518	76	-3.376004		
25	-6.829853	51	-5.590547	77	-3.282179		

ตาราง ก5 ค่าประมาณเบต้าด้วยวิธีแบบเบย์สำหรับเพศหญิง

อายุ x (ปี)	$\hat{\beta}_x$	อายุ x (ปี)	$\hat{\beta}_x$	อายุ x (ปี)	$\hat{\beta}_x$	อายุ x (ปี)	$\hat{\beta}_x$
0	-0.003383	26	0.040512	52	0.004384	78	0.000987
1	0.073878	27	0.039321	53	0.004285	79	-0.000113
2	0.023998	28	0.038222	54	0.003991	80	0.000935
3	0.027030	29	0.035772	55	0.005138	81	0.002123
4	0.029468	30	0.032145	56	0.006687	82	0.004322
5	0.029209	31	0.029959	57	0.005988	83	0.004165
6	0.028454	32	0.025141	58	0.007870	84	0.006212
7	0.025594	33	0.021973	59	0.007658	85	0.004877
8	0.019213	34	0.019805	60	0.007908	86	0.006054
9	0.011660	35	0.015476	61	0.009623	87	0.005158
10	0.009408	36	0.013467	62	0.011212	88	0.003914
11	0.005947	37	0.010450	63	0.010289	89	0.003097
12	-0.001752	38	0.009637	64	0.010952	90	0.001734
13	-0.003033	39	0.007423	65	0.010591	91	0.000236
14	-0.000681	40	0.006932	66	0.009182	92	0.000272
15	0.000474	41	0.006651	67	0.007588	93	-0.001505
16	0.006617	42	0.006235	68	0.006281	94	-0.002676
17	0.008076	43	0.004696	69	0.003539	95	-0.007269
18	0.009910	44	0.002566	70	0.002311	96	-0.007133
19	0.012706	45	0.004446	71	0.001420	97	-0.011037
20	0.016677	46	0.003160	72	0.000582	98	-0.009918
21	0.021476	47	0.003080	73	-0.000124	99	-0.013218
22	0.027862	48	0.005892	74	-0.000450	100	-0.017367
23	0.035007	49	0.004710	75	-0.000484		
24	0.039212	50	0.004509	76	-0.000503		
25	0.039786	51	0.005090	77	0.000151		

ตาราง ก6 ค่าประมาณแคปปาด้วยวิธีแบบเบย์สำหรับเพศหญิง

ปี	$\hat{\kappa}_t$
2540	9.244805
2541	9.824172
2542	11.833440
2543	10.814437
2544	9.036372
2545	8.225570
2546	6.541674
2547	4.021701
2548	1.249533
2549	-0.893956
2550	-3.101402
2551	-4.196515
2552	-5.814657
2553	-6.002381
2554	-6.597534
2555	-7.887261
2556	-8.592860
2557	-9.034445
2558	-9.539569
2559	-9.131123

ภาคผนวก ข

ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE)

ตาราง ข1 ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) สำหรับเพศชาย

อายุ x (ปี)	MAPE	อายุ x (ปี)	MAPE	อายุ x (ปี)	MAPE	อายุ x (ปี)	MAPE
0	0.033959	26	0.062282	52	0.033843	78	0.177037
1	0.056836	27	0.046824	53	0.034129	79	0.180683
2	0.082447	28	0.016840	54	0.054291	80	0.217698
3	0.172662	29	0.046566	55	0.052709	81	0.245662
4	0.225291	30	0.036011	56	0.029229	82	0.266178
5	0.257692	31	0.022594	57	0.047133	83	0.259163
6	0.431747	32	0.046797	58	0.069406	84	0.289618
7	0.262754	33	0.056819	59	0.085538	85	0.298714
8	0.237425	34	0.034908	60	0.099867	86	0.345007
9	0.144984	35	0.058351	61	0.086867	87	0.320280
10	0.089198	36	0.056671	62	0.044956	88	0.318053
11	0.103085	37	0.032861	63	0.062243	89	0.320488
12	0.045865	38	0.021964	64	0.057613	90	0.405449
13	0.105069	39	0.044994	65	0.123034	91	0.405920
14	0.072705	40	0.047896	66	0.109148	92	0.488144
15	0.047645	41	0.031730	67	0.088559	93	0.518027
16	0.085560	42	0.057202	68	0.126191	94	0.505270
17	0.047240	43	0.003699	69	0.144087	95	0.560545
18	0.036493	44	0.038378	70	0.157628	96	0.634687
19	0.051068	45	0.014103	71	0.183947	97	0.732136
20	0.047115	46	0.031324	72	0.152657	98	0.840869
21	0.063666	47	0.030792	73	0.152735	99	0.670686
22	0.045104	48	0.030914	74	0.146720	100	0.582670
23	0.024837	49	0.021575	75	0.147257		
24	0.020653	50	0.046303	76	0.173662		
25	0.065864	51	0.035065	77	0.147226		

ตาราง ข2 ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) สำหรับเพศหญิง

อายุ x (ปี)	MAPE	อายุ x (ปี)	MAPE	อายุ x (ปี)	MAPE	อายุ x (ปี)	MAPE
0	0.158847	26	0.180058	52	0.019249	78	0.081958
1	0.138359	27	0.052463	53	0.016950	79	0.074946
2	0.191285	28	0.125047	54	0.042095	80	0.045253
3	0.184091	29	0.209513	55	0.035136	81	0.065933
4	0.244563	30	0.165504	56	0.027074	82	0.030023
5	0.260166	31	0.206127	57	0.033218	83	0.075824
6	0.162845	32	0.214694	58	0.033483	84	0.031499
7	0.469304	33	0.133581	59	0.067494	85	0.060857
8	0.533322	34	0.216496	60	0.084505	86	0.085901
9	0.259384	35	0.261435	61	0.072336	87	0.079697
10	0.793995	36	0.148126	62	0.044192	88	0.067342
11	0.236699	37	0.119170	63	0.033302	89	0.063163
12	0.316696	38	0.121282	64	0.065169	90	0.063652
13	0.151729	39	0.141315	65	0.038099	91	0.138066
14	0.228719	40	0.081978	66	0.053224	92	0.124693
15	0.296437	41	0.054849	67	0.073362	93	0.152779
16	0.148517	42	0.044483	68	0.081734	94	0.114346
17	0.034732	43	0.043411	69	0.113865	95	0.113910
18	0.081234	44	0.096872	70	0.139464	96	0.093817
19	0.103351	45	0.036358	71	0.138660	97	0.177339
20	0.083101	46	0.057433	72	0.112786	98	0.182598
21	0.036598	47	0.033542	73	0.128469	99	0.233260
22	0.045234	48	0.049683	74	0.109225	100	0.164121
23	0.065770	49	0.020968	75	0.119687		
24	0.018789	50	0.021914	76	0.112102		
25	0.080622	51	0.025783	77	0.077887		

ภาคผนวก ค

ชุดคำสั่งของโปรแกรมอาร์ในการประมาณและพยากรณ์อัตรามรณะ

ชุดคำสั่ง ค1 โปรแกรมประมาณอัตรามรณะโดยวิธีแบบเบย์สำหรับเพศชาย

```
library(MASS)
library(forecast)
library(TruncatedNormal)
library(ggplot2)

# import and transform data
setwd("C:/Users/ASUS/Desktop/Senior Project")
ABmle <- read.csv("C:/Users/ASUS/Desktop/Senior Project/output_male.csv")
Ktmle <- read.csv("C:/Users/ASUS/Desktop/Senior Project/output_kt_male.csv")
Dxt <- as.matrix(read.csv("C:/Users/ASUS/Desktop/Senior Project/Mortality
Male.csv"))[,-1:-2]
Ext <- as.matrix(read.csv("C:/Users/ASUS/Desktop/Senior Project/Population
Male.csv"))[,-1:-2]
Amle <- as.numeric(ABmle[,2])
Bmle <- as.numeric(ABmle[,3])
Kmle <- data.frame(cbind(1:20, t(Ktmle)))

# set loops
set.seed(2500)
N = 20000
M = 101
t = 20

### initial values ###

# fit kt in SLR to find g0 and S0
SLR <- lm(Kmle[,2]~Kmle[,1], data = Kmle)
g0 = t(t(SLR$coefficients))
S0 = vcov(SLR)

# fit transformed kt in AR(1) to find roh and s2k
```

```
X <- matrix(1, nrow = t, ncol = 2)
X[, 2] = 1:t
Eta <- X %*% g0
plot.ts(Kmle[2] - Eta)
AR <- arima(Kmle[2] - Eta, order = c(1,0,0))
roh.ar = AR$coef[1]
s2k.ar = AR$sigma2

# other initial values
s2b.mle = var(Bmle)
alpha.mle = Amle
beta.mle = Bmle
kappa.mle = Kmle[,2]
s2p = 1
Ab = 2.1
Bb = (Ab - 1) * s2b.mle
Ak = 2.1
Bk = (Ak - 1) * s2k.ar
Ba = 0.001
Aa = Ba * exp(alpha.mle)

P <- matrix(0, nrow = t, ncol = t)
for(i in 1:t - 1){
  P[i + 1, i] = roh.ar
}

Q <- matrix(0, nrow = t, ncol = t)
for(i in 1:t - 1){
  Q[i, i + 1] = -roh.ar
  Q[i, i] = 1 + roh.ar^2
  Q[i + 1, i] = -roh.ar
}
```

```
Q[t, t] = 1

gamma <- matrix(0, nrow = 2, ncol = N)
roh <- rep(0,N)
s2k <- rep(0,N)
s2b <- rep(0,N)
alpha <- matrix(0, nrow = M, ncol = N)
beta <- matrix(0, nrow = M, ncol = N)
kappa <- matrix(0, nrow = t, ncol = N)

gamma[,1] <- g0
roh[1] <- roh.ar
s2k[1] <- s2k.ar
s2b[1] <- s2b.mle
alpha[,1] <- alpha.mle
beta[,1] <- beta.mle
kappa[,1] <- kappa.mle

Pgamma <- c()
Proh <- c()
Ps2k <- c()
Ps2b <- c()
Palpha <- c()
Pbeta <- c()
Pkappa <- c()

### MCMC ###

countk = 0
countb = 0
sigmax = sqrt(s2b.mle / 32)
sigmat = 1/2
```

```
for(i in 2:N){
  # gamma
  S. <- solve(t(X) %*% Q %*% X + s2k[i - 1] * solve(S0))
  g. <- S. %*% (t(X) %*% Q %*% kappa[,i - 1] + s2k[i - 1] * solve(S0) %*% g0)
  gamma[,i] <- t(mvrnorm(1, g., s2k[i - 1] * S.))

  # s2b
  s2b[i] <- 1 / rgamma(1, Ab + M / 2, Bb + (t(beta[,i - 1]) %*% beta[,i - 1]
/ 2))

  # s2k
  eta <- X %*% gamma[,i]
  S1 = (kappa[1,i - 1] - eta[1]) ^ 2
  for(j in 2:t){
    s1 <- (kappa[j,i - 1] - eta[j] - roh[i - 1] * (kappa[j - 1,i - 1] - eta[j
- 1])) ^ 2
    S1 = S1 + s1
  }
  s2k[i] <- 1 / rgamma(1, Ak + t / 2, Bk + S1 / 2)

  # roh
  Ap = 0
  Bp = 0
  for(j in 2:t){
    ap <- (kappa[j - 1,i - 1] - eta[j - 1]) ^ 2
    Ap = Ap + ap
    bp <- (kappa[j,i - 1] - eta[j]) * (kappa[j - 1,i - 1] - eta[j - 1])
    Bp = Bp + bp
  }
  mup. <- Bp / (Ap + s2k[i] / s2p)
  s2p. <- s2k[i - 1] / (Ap + s2k[i] / s2p)
  roh[i] <- rtnorm(1, mup., s2p., -1, 1)
```

```
  # kappa
  # x = kappa[j,i - 1]
  # y = 1
  # k = kappa[,i - 1]
  # a = alpha[,i - 1]
  # b = beta[,i - 1]
  llh.kt <- function(x, y, k, a, b){
    f.Dt = 0
    fkt <- -(x - eta[1]) ^ 2 / (2 * s2k[i])
    for(m in 1:M){
      fDt <- as.numeric(-Ext[m,y] * exp(a[m] + b[m] * x) + b[m] * x *
Dxt[m,y])
      f.Dt = f.Dt + fDt
    }
    if(y < t){
      fkt1.kt <- -(k[y + 1] - eta[y + 1] - roh[i] * (x - eta[y])) ^ 2 / (2 *
s2k[i])
    }
    if(y > 1){
      fkt.kt1 <- -(x - eta[y] - roh[i] * (k[y - 1] - eta[y - 1])) ^ 2 / (2 *
s2k[i])
    }
    if(y == 1){
      llh <- f.Dt + fkt + fkt1.kt
    }
    else if(y == t){
      llh <- f.Dt + fkt.kt1
    }
    else{
      llh <- f.Dt + fkt.kt1 + fkt1.kt
    }
    return(llh)
```

```
  }
  kap <- kappa[,i - 1]
  alp <- alpha[,i - 1]
  bet <- beta[,i - 1]
  j = 1
  for(j in 1:t){
    pp.kt <- rnorm(1, kappa[j,i - 1], sigmat)
    psi1 <- min(1, exp(llh.kt(pp.kt, j, kap, alp, bet) - llh.kt(kappa[j,i -
1], j, kap, alp, bet)))
    u <- runif(1)
    if(u > psi1 | is.nan(psi1) == T){
      kap[j] <- kappa[j,i - 1]
    }
    else{
      kap[j] <- pp.kt
      countk = countk + 1
    }
    kbar <- mean(kap)
    kap <- kap - kbar
    alp <- alp + bet * kbar
  }
  kappa[,i] <- kap
  alpha[,i] <- alp


  # beta
  # x = pp.bx
  # y = 68
  # k = kap
  # a = alp
  # b = bet
  llh.bx <- function(x, y, k, a, b){
    f.Dt = 0
```

```
    fbx <- -x ^ 2 / (2 * s2b[i])
    for(n in 1:t){
      fDt <- as.numeric(-Ext[y,n] * exp(a[y] + x * k[n]) + x * k[n] *
Dxt[y,n])
      f.Dt = f.Dt + fDt
    }
    llh <- f.Dt + fbx
    return(llh)
  }
  kap <- kappa[,i]
  alp <- alpha[,i]
  bet <- beta[,i - 1]
  for(j in 1:M){
    pp.bx <- rnorm(1, beta[j,i - 1], sigmax)
    psi2 <- min(1, exp(llh.bx(pp.bx, j, kap, alp, bet) - llh.bx(beta[j,i -
1], j, kap, alp, bet)))
    v <- runif(1)
    if(v > psi2 | is.nan(psi2) == T){
      bet[j] <- beta[j,i - 1]
    }
    else{
      bet[j] <- pp.bx
      countb = countb + 1
    }
    bsum <- sum(bet)
    bet <- bet / bsum
    kap <- kap * bsum
  }
  kappa[,i] <- kap
  beta[,i] <- bet


  # alpha
  c <- rep(0,M)
```

```
  D. <- rep(0,M)
  for(j in 1:M){
    c[j] <- Ext[j,] %*% exp(beta[j,i] * kappa[,i])
    D.[j] <- rowSums(Dxt)[j]
    alpha[j,i] <- log(rgamma(1, Aa[j] + D.[j], Ba + c[j]))
  }

  # posterior distribution (burn-in)
  if(i > N / 2 & i %% 10 == 1){
    Pgamma <- cbind(Pgamma, gamma[,i])
    Proh <- cbind(Proh, roh[i])
    Ps2k <- cbind(Ps2k, s2k[i])
    Ps2b <- cbind(Ps2b, s2b[i])
    Palpha <- cbind(Palpha, alpha[,i])
    Pbeta <- cbind(Pbeta, beta[,i])
    Pkappa <- cbind(Pkappa, kappa[,i])
  }
}

### Checking ###

rateK <- countk / (t * N)
rateB <- countb / (M * N)
rateK
rateB
hist(Pgamma)
plot(Pgamma[1,], type = "b")
plot(Pgamma[2,], type = "b")
hist(Proh)
plot(Proh[1,], type = "b")
hist(Ps2k)
plot(Ps2k[1,], type = "b")
```

```
hist(Ps2b)
plot(Ps2b[1,], type = "b")

# Convergence of Parameters
par(mfrow = c(2,2))
for(i in 1:4){
  a <- 30 * (i - 1)
  ma <- paste("Alpha ", a)
  plot(1:200, alpha[a + 1,1:200], main = ma, type = "l")
  abline(h = alpha.mle[i], col = "blue")
}
for(i in 1:4){
  a <- 30 * (i - 1)
  mb <- paste("Beta ", a)
  plot(1:200, beta[a + 1,1:200], main = mb, type = "l")
  abline(h = beta.mle[i], col = "blue")
}
for(i in 1:t){
  mk <- paste("Kappa ", i + 1997)
  plot(1:200, kappa[i,1:200], main = mk, type = "l")
  abline(h = kappa.mle[i], col = "blue")
}
plot(1:100, kappa[1,1:100], type = "b")

# Posterior Distribution
par(mfrow = c(1,1))
par(mfrow = c(2,2))
for(i in 1:4){
  a <- 30 * (i - 1)
  m1 <- paste("Alpha ", a)
  hist(Palpha[a + 1,], main = m1)
  abline(v = mean(Palpha[a + 1,]), col = "red")
```

```
  abline(v = alpha.mle[a + 1], col = "blue")
}
for(i in 1:4){
  a <- 30 * (i - 1)
  m1 <- paste("Beta ", a)
  hist(Pbeta[a + 1,], main = m1)
  abline(v = mean(Pbeta[a + 1,]), col = "red")
  abline(v = beta.mle[a + 1], col = "blue")
}
for(i in 1:t){
  m2 <- paste("Kappa ", (i + 1997))
  hist(Pkappa[i,], main = m2)
  abline(v = mean(Pkappa[i,]), col = "red")
  abline(v = kappa.mle[i], col = "blue")
}
for(i in 1:4){
  a <- 30 * (i - 1)
  m1 <- paste("Alpha ", a)
  plot(Palpha[a + 1,], main = m1, type = "b")
  abline(h = mean(Palpha[a + 1,]), col = "red")
  abline(h = alpha.mle[a + 1], col = "blue")
}
for(i in 1:4){
  a <- 30 * (i - 1)
  m1 <- paste("Beta ", a)
  plot(Pbeta[a + 1,], main = m1, type = "b")
  abline(h = mean(Pbeta[a + 1,]), col = "red")
  abline(h = beta.mle[a + 1], col = "blue")
}
for(i in 1:t){
  m2 <- paste("Kappa ", (i + 1997))
  plot(Pkappa[i,], main = m2, type = "b")
```

```
  abline(h = mean(Pkappa[i,]), col = "red")
  abline(h = kappa.mle[i], col = "blue")
}

### Bayesian Estimation ###

PostmeanA <- c()
PostmeanB <- c()
PostmeanK <- c()
for(i in 1:M){
  PostmeanA <- cbind(PostmeanA, mean(Palpha[i,]))
  PostmeanB <- cbind(PostmeanB, mean(Pbeta[i,]))
  if(i <= t){
    PostmeanK <- cbind(PostmeanK, mean(Pkappa[i,]))
  }
}

par(mfrow = c(1,1))
plot(x = 0:(M - 1), y = PostmeanA, main = "Estimated Alpha (M)",
     xlab = "Age", ylab = "Alpha", type = "b", col = "red", pch = 4, lty = 4)
lines(x = 0:(M - 1), y = alpha.mle, type = "b", col = "blue", pch = 3, lty =
3)
legend(x = 0, y = -2, legend = c("MLE","Bayes"), col = c("blue","red"),
       lty = c(3,4), pch = c(3,4), cex = 0.7)
plot(x = 0:(M - 1), y = PostmeanB, main = "Estimated Beta (M)",
     xlab = "Age", ylab = "Beta", type = "b", col = "red", pch = 4)
lines(x = 0:(M - 1), y = beta.mle, type = "b", col = "blue", pch = 3)
legend(x = 72, y = 0.075, legend = c("MLE","Bayes"), col = c("blue","red"),
       lty = c(3,4), pch = c(3,4), cex = 0.7)
plot(x = 2540:2559, y = PostmeanK, main = "Estimated Kappa (M)",
     xlab = "Year", ylab = "Kappa", type = "b", col = "red", pch = 4)
lines(x = 2540:2559, y =  kappa.mle, type = "b", col = "blue", pch = 3)
```

```
legend(x = 2553.5, y = 11, legend = c("MLE","Bayes"), col = c("blue","red"),
       lty = c(3,4), pch = c(3,4), cex = 0.7)
PostmeanA - alpha.mle
PostmeanB - beta.mle
PostmeanK - kappa.mle

### Log mortailty rate ###

logm.mle <- function(i){
  return(alpha.mle + beta.mle * kappa.mle[i])
}
plot.logm.mle <- function(a, b){
  plot(x = 0:(M - 1), y = logm.mle(b - 2539), main = "Male log mortal rate
(MLE)",
       xlab = "Age", ylab = "Log mortality rate", type = "l", col = "red")
  for(i in (b - a):1){
    lines(x = 0:(M - 1), y = logm.mle(b - 2539 - i), type = "l",
          col = rainbow(1, start = (i + 0.226 * (b - a - i) - 4.3) / (b -
a)))
  }
}
plot.logm.mle(2540,2559)

logm.bayes <- function(i){
  return(PostmeanA + PostmeanB * PostmeanK[i])
}
plot.logm.bayes <- function(a, b){
  plot(x = 0:(M - 1), y = logm.bayes(b - 2539), main = "Male log mortal rate
(Bayes)",
       xlab = "Age", ylab = "Log mortality rate", type = "l", col = "red")
  for(i in (b - a):1){
    lines(x = 0:(M - 1), y = logm.bayes(b - 2539 - i), type = "l",
```

```
          col = rainbow(1, start = (i + 0.226 * (b - a - i) - 4.3) / (b -
a)))
  }
}
plot.logm.bayes(2540,2559)


### Print value ###


write.csv(PostmeanA, "output_axbay_male.csv", row.names = F)
write.csv(PostmeanB, "output_bxbay_male.csv", row.names = F)
write.csv(PostmeanK, "output_ktbay_male.csv", row.names = F)
```

ชุดคำสั่ง ค2 โปรแกรมประมาณอัตรามรณะโดยวิธีแบบเบย์สำหรับเพศหญิง

```
library(MASS)

library(forecast)

library(TruncatedNormal)

library(ggplot2)


# import and transform data

setwd("C:/Users/ASUS/Desktop/Senior Project")

ABmle <- read.csv("C:/Users/ASUS/Desktop/Senior Project/output_female.csv")

Ktmle <- read.csv("C:/Users/ASUS/Desktop/Senior
Project/output_kt_female.csv")

Dxt <- as.matrix(read.csv("C:/Users/ASUS/Desktop/Senior Project/Mortality
Female.csv"))[,-1]

Ext <- as.matrix(read.csv("C:/Users/ASUS/Desktop/Senior Project/Population
Female.csv"))[,-1]

Amle <- as.numeric(ABmle[,2])

Bmle <- as.numeric(ABmle[,3])

Kmle <- data.frame(cbind(1:20, t(Ktmle)))


# set loops

set.seed(2500)

N = 20000

M = 101

t = 20


### initial values ###


# fit kt in SLR to find g0 and S0

SLR <- lm(Kmle[,2]~Kmle[,1], data = Kmle)

g0 = t(t(SLR$coefficients))

S0 = vcov(SLR)
```

```
# fit transformed kt in AR(1) to find roh and s2k
X <- matrix(1, nrow = t, ncol = 2)
X[, 2] = 1:t
Eta <- X %*% g0
plot.ts(Kmle[2] - Eta)
AR <- arima(Kmle[2] - Eta, order = c(1,0,0))
roh.ar = AR$coef[1]
s2k.ar = AR$sigma2

# other initial values
s2b.mle = var(Bmle)
alpha.mle = Amle
beta.mle = Bmle
kappa.mle = Kmle[,2]
s2p = 1
Ab = 2.1
Bb = (Ab - 1) * s2b.mle
Ak = 2.1
Bk = (Ak - 1) * s2k.ar
Ba = 0.001
Aa = Ba * exp(alpha.mle)

P <- matrix(0, nrow = t, ncol = t)
for(i in 1:t - 1){
  P[i + 1, i] = roh.ar
}

Q <- matrix(0, nrow = t, ncol = t)
for(i in 1:t - 1){
  Q[i, i + 1] = -roh.ar
  Q[i, i] = 1 + roh.ar^2
  Q[i + 1, i] = -roh.ar
```

```
}
Q[t, t] = 1

gamma <- matrix(0, nrow = 2, ncol = N)
roh <- rep(0,N)
s2k <- rep(0,N)
s2b <- rep(0,N)
alpha <- matrix(0, nrow = M, ncol = N)
beta <- matrix(0, nrow = M, ncol = N)
kappa <- matrix(0, nrow = t, ncol = N)

gamma[,1] <- g0
roh[1] <- roh.ar
s2k[1] <- s2k.ar
s2b[1] <- s2b.mle
alpha[,1] <- alpha.mle
beta[,1] <- beta.mle
kappa[,1] <- kappa.mle

Pgamma <- c()
Proh <- c()
Ps2k <- c()
Ps2b <- c()
Palpha <- c()
Pbeta <- c()
Pkappa <- c()

### MCMC ###

countk = 0
countb = 0
sigmax = sqrt(s2b.mle / 32)
```

```
sigmat = 1
for(i in 2:N){
  # gamma
  S. <- solve(t(X) %*% Q %*% X + s2k[i - 1] * solve(S0))
  g. <- S. %*% (t(X) %*% Q %*% kappa[,i - 1] + s2k[i - 1] * solve(S0) %*% g0)
  gamma[,i] <- t(mvrnorm(1, g., s2k[i - 1] * S.))


  # s2b
  s2b[i] <- 1 / rgamma(1, Ab + M / 2, Bb + (t(beta[,i - 1]) %*% beta[,i - 1]
/ 2))


  # s2k
  eta <- X %*% gamma[,i]
  S1 = (kappa[1,i - 1] - eta[1]) ^ 2
  for(j in 2:t){
    s1 <- (kappa[j,i - 1] - eta[j] - roh[i - 1] * (kappa[j - 1,i - 1] - eta[j
- 1])) ^ 2
    S1 = S1 + s1
  }
  s2k[i] <- 1 / rgamma(1, Ak + t / 2, Bk + S1 / 2)


  # roh
  Ap = 0
  Bp = 0
  for(j in 2:t){
    ap <- (kappa[j - 1,i - 1] - eta[j - 1]) ^ 2
    Ap = Ap + ap
    bp <- (kappa[j,i - 1] - eta[j]) * (kappa[j - 1,i - 1] - eta[j - 1])
    Bp = Bp + bp
  }
  mup. <- Bp / (Ap + s2k[i] / s2p)
  s2p. <- s2k[i - 1] / (Ap + s2k[i] / s2p)
  roh[i] <- rtnorm(1, mup., s2p., -1, 1)
```

```
  # kappa
  # x = kappa[j,i - 1]
  # y = 1
  # k = kappa[,i - 1]
  # a = alpha[,i - 1]
  # b = beta[,i - 1]
  llh.kt <- function(x, y, k, a, b){
    f.Dt = 0
    fkt <- -(x - eta[1]) ^ 2 / (2 * s2k[i])
    for(m in 1:M){
      fDt <- as.numeric(-Ext[m,y] * exp(a[m] + b[m] * x) + b[m] * x *
Dxt[m,y])
      f.Dt = f.Dt + fDt
    }
    if(y < t){
      fkt1.kt <- -(k[y + 1] - eta[y + 1] - roh[i] * (x - eta[y])) ^ 2 / (2 *
s2k[i])
    }
    if(y > 1){
      fkt.kt1 <- -(x - eta[y] - roh[i] * (k[y - 1] - eta[y - 1])) ^ 2 / (2 *
s2k[i])
    }
    if(y == 1){
      llh <- f.Dt + fkt + fkt1.kt
    }
    else if(y == t){
      llh <- f.Dt + fkt.kt1
    }
    else{
      llh <- f.Dt + fkt.kt1 + fkt1.kt
    }
    return(llh)
```

```
  }
  kap <- kappa[,i - 1]
  alp <- alpha[,i - 1]
  bet <- beta[,i - 1]
  j = 1
  for(j in 1:t){
    pp.kt <- rnorm(1, kappa[j,i - 1], sigmat)
    psi1 <- min(1, exp(llh.kt(pp.kt, j, kap, alp, bet) - llh.kt(kappa[j,i -
1], j, kap, alp, bet)))
    u <- runif(1)
    if(u > psi1 | is.nan(psi1) == T){
      kap[j] <- kappa[j,i - 1]
    }
    else{
      kap[j] <- pp.kt
      countk = countk + 1
    }
    kbar <- mean(kap)
    kap <- kap - kbar
    alp <- alp + bet * kbar
  }
  kappa[,i] <- kap
  alpha[,i] <- alp

  # beta
  # x = pp.bx
  # y = 68
  # k = kap
  # a = alp
  # b = bet
  llh.bx <- function(x, y, k, a, b){
    f.Dt = 0
```

```
    fbx <- -x ^ 2 / (2 * s2b[i])
    for(n in 1:t){
      fDt <- as.numeric(-Ext[y,n] * exp(a[y] + x * k[n]) + x * k[n] *
Dxt[y,n])
      f.Dt = f.Dt + fDt
    }
    llh <- f.Dt + fbx
    return(llh)
  }
  kap <- kappa[,i]
  alp <- alpha[,i]
  bet <- beta[,i - 1]
  for(j in 1:M){
    pp.bx <- rnorm(1, beta[j,i - 1], sigmax)
    psi2 <- min(1, exp(llh.bx(pp.bx, j, kap, alp, bet) - llh.bx(beta[j,i -
1], j, kap, alp, bet)))
    v <- runif(1)
    if(v > psi2 | is.nan(psi2) == T){
      bet[j] <- beta[j,i - 1]
    }
    else{
      bet[j] <- pp.bx
      countb = countb + 1
    }
    bsum <- sum(bet)
    bet <- bet / bsum
    kap <- kap * bsum
  }
  kappa[,i] <- kap
  beta[,i] <- bet


  # alpha
  c <- rep(0,M)
```

```
  D. <- rep(0,M)
  for(j in 1:M){
    c[j] <- Ext[j,] %*% exp(beta[j,i] * kappa[,i])
    D.[j] <- rowSums(Dxt)[j]
    alpha[j,i] <- log(rgamma(1, Aa[j] + D.[j], Ba + c[j]))
  }

  # posterior distribution (burn-in)
  if(i > N / 2 & i %% 10 == 1){
    Pgamma <- cbind(Pgamma, gamma[,i])
    Proh <- cbind(Proh, roh[i])
    Ps2k <- cbind(Ps2k, s2k[i])
    Ps2b <- cbind(Ps2b, s2b[i])
    Palpha <- cbind(Palpha, alpha[,i])
    Pbeta <- cbind(Pbeta, beta[,i])
    Pkappa <- cbind(Pkappa, kappa[,i])
  }
}

### Checking ###

rateK <- countk / (t * N)
rateB <- countb / (M * N)
rateK
rateB
hist(Pgamma)
plot(Pgamma[1,], type = "b")
plot(Pgamma[2,], type = "b")
hist(Proh)
plot(Proh[1,], type = "b")
hist(Ps2k)
plot(Ps2k[1,], type = "b")
```

```
hist(Ps2b)
plot(Ps2b[1,], type = "b")

# Convergence of Parameters
par(mfrow = c(2,2))
for(i in 1:4){
  a <- 30 * (i - 1)
  ma <- paste("Alpha ", a)
  plot(1:200, alpha[a + 1,1:200], main = ma, type = "l")
  abline(h = alpha.mle[i], col = "blue")
}
for(i in 1:4){
  a <- 30 * (i - 1)
  mb <- paste("Beta ", a)
  plot(1:200, beta[a + 1,1:200], main = mb, type = "l")
  abline(h = beta.mle[i], col = "blue")
}
for(i in 1:t){
  mk <- paste("Kappa ", i + 1997)
  plot(1:200, kappa[i,1:200], main = mk, type = "l")
  abline(h = kappa.mle[i], col = "blue")
}
plot(1:100, kappa[1,1:100], type = "b")

# Posterior Distribution
par(mfrow = c(1,1))
par(mfrow = c(2,2))
for(i in 1:4){
  a <- 30 * (i - 1)
  m1 <- paste("Alpha ", a)
  hist(Palpha[a + 1,], main = m1)
  abline(v = mean(Palpha[a + 1,]), col = "red")
```

```
  abline(v = alpha.mle[a + 1], col = "blue")
}
for(i in 1:4){
  a <- 30 * (i - 1)
  m1 <- paste("Beta ", a)
  hist(Pbeta[a + 1,], main = m1)
  abline(v = mean(Pbeta[a + 1,]), col = "red")
  abline(v = beta.mle[a + 1], col = "blue")
}
for(i in 1:t){
  m2 <- paste("Kappa ", (i + 1997))
  hist(Pkappa[i,], main = m2)
  abline(v = mean(Pkappa[i,]), col = "red")
  abline(v = kappa.mle[i], col = "blue")
}
for(i in 1:4){
  a <- 30 * (i - 1)
  m1 <- paste("Alpha ", a)
  plot(Palpha[a + 1,], main = m1, type = "b")
  abline(h = mean(Palpha[a + 1,]), col = "red")
  abline(h = alpha.mle[a + 1], col = "blue")
}
for(i in 1:4){
  a <- 30 * (i - 1)
  m1 <- paste("Beta ", a)
  plot(Pbeta[a + 1,], main = m1, type = "b")
  abline(h = mean(Pbeta[a + 1,]), col = "red")
  abline(h = beta.mle[a + 1], col = "blue")
}
for(i in 1:t){
  m2 <- paste("Kappa ", (i + 1997))
  plot(Pkappa[i,], main = m2, type = "b")
```

```
  abline(h = mean(Pkappa[i,]), col = "red")
  abline(h = kappa.mle[i], col = "blue")
}

### Bayesian Estimation ###

PostmeanA <- c()
PostmeanB <- c()
PostmeanK <- c()
for(i in 1:M){
  PostmeanA <- cbind(PostmeanA, mean(Palpha[i,]))
  PostmeanB <- cbind(PostmeanB, mean(Pbeta[i,]))
  if(i <= t){
    PostmeanK <- cbind(PostmeanK, mean(Pkappa[i,]))
  }
}

par(mfrow = c(1,1))
plot(x = 0:(M - 1), y = PostmeanA, main = "Estimated Alpha (F)",
     xlab = "Age", ylab = "Alpha", type = "b", col = "red", pch = 4, lty = 4)
lines(x = 0:(M - 1), y = alpha.mle, type = "b", col = "blue", pch = 3, lty =
3)
legend(x = 0, y = -2, legend = c("MLE","Bayes"), col = c("blue","red"),
       lty = c(3,4), pch = c(3,4), cex = 0.7)
plot(x = 0:(M - 1), y = PostmeanB, main = "Estimated Beta (F)",
     xlab = "Age", ylab = "Beta", type = "b", col = "red", pch = 4)
lines(x = 0:(M - 1), y = beta.mle, type = "b", col = "blue", pch = 3)
legend(x = 72, y = 0.072, legend = c("MLE","Bayes"), col = c("blue","red"),
       lty = c(3,4), pch = c(3,4), cex = 0.7)
plot(x = 2540:2559, y = PostmeanK, main = "Estimated Kappa (F)",
     xlab = "Year", ylab = "Kappa", type = "b", col = "red", pch = 4)
lines(x = 2540:2559, y =  kappa.mle, type = "b", col = "blue", pch = 3)
```

```
legend(x = 2553.5, y = 16, legend = c("MLE","Bayes"), col = c("blue","red"),
       lty = c(3,4), pch = c(3,4), cex = 0.7)
PostmeanA - alpha.mle
PostmeanB - beta.mle
PostmeanK - kappa.mle

### Log mortailty rate ###

logm.mle <- function(i){
  return(alpha.mle + beta.mle * kappa.mle[i])
}
plot.logm.mle <- function(a, b){
  plot(x = 0:(M - 1), y = logm.mle(b - 2539), main = "Female log mortal rate
(MLE)",
       xlab = "Age", ylab = "Log mortality rate", type = "l", col = "red")
  for(i in (b - a):1){
    lines(x = 0:(M - 1), y = logm.mle(b - 2539 - i), type = "l",
          col = rainbow(1, start = (i + 0.226 * (b - a - i) - 4.3) / (b -
a)))
  }
}
plot.logm.mle(2540,2559)

logm.bayes <- function(i){
  return(PostmeanA + PostmeanB * PostmeanK[i])
}
plot.logm.bayes <- function(a, b){
  plot(x = 0:(M - 1), y = logm.bayes(b - 2539), main = "Female log mortal
rate (Bayes)",
       xlab = "Age", ylab = "Log mortality rate", type = "l", col = "red")
  for(i in (b - a):1){
    lines(x = 0:(M - 1), y = logm.bayes(b - 2539 - i), type = "l",
```

```
          col = rainbow(1, start = (i + 0.226 * (b - a - i) - 4.3) / (b -
a)))
  }
}
plot.logm.bayes(2540,2559)

### Print value ###

write.csv(PostmeanA, "output_axbay_female.csv", row.names = F)
write.csv(PostmeanB, "output_bxbay_female.csv", row.names = F)
write.csv(PostmeanK, "output_ktbay_female.csv", row.names = F)
```

ชุดคำสั่ง ค3 โปรแกรมแสดงกราฟเปรียบเทียบประสิทธิภาพในการพยากรณ์อัตรามรณะสำหรับเพศชายและเพศหญิง

```
MAPE <- read.csv("C:/Users/ASUS/Desktop/Senior Project/MAPE Use.csv").
MAPE.mle.m <- c()
MAPE.mle.f <- c()
MAPE.bay.m <- c()
MAPE.bay.f <- c()
for(i in 1:M){
  MAPE.mle.m[i] <- (MAPE[,3])[i]
}
for(i in 1:M){
  MAPE.mle.f[i] <- (MAPE[,2])[i]
}
for(i in 1:M){
  MAPE.bay.m[i] <- (MAPE[,5])[i]
}
for(i in 1:M){
  MAPE.bay.f[i] <- (MAPE[,4])[i]
}

plot(x = 0:(M - 1), y = MAPE.mle.m, main = "MAPE of MLE between male and
female",
     xlab = "Age", ylab = "MAPE", type = "l", col = "red")
lines(x = 0:(M - 1), y = MAPE.mle.f, type = "l", col = "blue")
legend(x = 0, y = 1.4, legend = c("Male","Female"), col = c("red","blue"),
lty = c(1,1), cex = 0.7)
plot(x = 0:(M - 1), y = MAPE.bay.m, main = "MAPE of Bayes between male and
female",
     xlab = "Age", ylab = "MAPE", type = "l", col = "red")
lines(x = 0:(M - 1), y = MAPE.bay.f, type = "l", col = "blue")
legend(x = 30, y = 0.8, legend = c("Male","Female"), col = c("red","blue"),
lty = c(1,1), cex = 0.7)
```

ชุดคำสั่ง ค4 โปรแกรมประมาณอัตรามรณะโดยวิธีภาวะน่าจะเป็นสูงสุด พยากรณ์อัตรามรณะและเปรียบเทียบประสิทธิภาพในการพยากรณ์อัตรามรณะโดยวิธีภาวะน่าจะเป็นสูงสุดและวิธีแบบเบย์

```
library(StMoMo)

library(demography)

library(forecast)

# prepare data for calculating mortality #


# import mortality data and change its format to martix form #


df.molf<-read.csv("D:/Senior project/Prepared data/Mortality
Female.csv",header = TRUE,sep=",")

df.molf<-df.molf[,2:21]

colnames(df.molf)<-c(2540:2559)

rownames(df.molf)<-c(0:100)

mat.molf<-data.matrix(df.molf)


# import population data and change its format to martix form #


df.popf<-read.csv("D:/Senior project/Prepared data/Population
Female.csv",header = TRUE,sep=",")

df.popf<-df.popf[,2:21]

colnames(df.popf)<-c(2540:2559)

rownames(df.popf)<-c(0:100)

mat.popf<-data.matrix(df.popf)


# create matrix mortality rate#


mat.rate<-mat.molf/mat.popf

mat.rate

# prepare to STMOMO format#


year<-c(2540:2559)
```

```
age<-c(0:100)

# prepare demogdata format #

tf<-demogdata(data = mat.rate,pop = mat.popf,ages = age,years = year,type =
"mortality",label = "Thai",name = "female" )



tfstmomo<-StMoMoData(tf,"female")
class(tfstmomo)
tfstmomo

# fit moidel#
Lcfit<-fit(lc(),data=tfstmomo, ages.fit = 0:100)
plot(Lcfit)
# forecast k #
Lcfor<-forecast(Lcfit,h=3)
plot(Lcfor)
# extract k value#
c<-Lcfor$kt.f
eps<-c$mean
# forcast actual balue$
ax<-data.matrix(Lcfit$ax)
# duplicate matrix#
axf<-cbind(ax,ax,ax)
bx<--data.matrix(Lcfit$bx)
bxf<-cbind(bx,bx,bx)
for (i in 1:3){
  bxf[,i]<-bxf[,i]*eps[i]
}
# forecast mortality rate#
mx<-exp(axf+bxf)
```

```
df<-data.frame(age=c(0:100),Lcfit$ax,Lcfit$bx)

write.csv(df,"D:/Senior project/output_female.csv",row.names=F)

test.mf<-read.csv("D:/Senior project/Prepared data/test set/Mortality
Female.Test.csv",header = TRUE,sep=",")

test.mf<-subset(test.mf,select = -c(ï..))

colnames(test.mf)<-c(2560,2561,2562)

rownames(test.mf)<-c(0:100)

mattest.mf<-data.matrix(test.mf)

test.pf<-read.csv("D:/Senior project/Prepared data/test set/Population
Female.Test.csv",header = TRUE,sep=",")

test.pf<-subset(test.pf,select = -c(ï..))

colnames(test.pf)<-c(2560,2561,2562)

rownames(test.pf)<-c(0:100)

mattest.pf<-data.matrix(test.pf)

mattest.rate<-mattest.mf/mattest.pf

f<-Lcfor$rates

V<-abs((f-mattest.rate)/mattest.rate)

MAPE<-matrix(nrow = 101,ncol = 1)

for(i in 1:101){

  MAPE[i,]<-sum(V[i,])/3

}

MAPE

par(mfrow=c(3,1))

plot(mx[,2],col="blue")

points(mattest.rate[,2],col="red")

for (i in 1:3) {

  plot(Lcfor$rates[,i],col="blue",main = paste("Mortality rate
Female",yeartest[i]),xlab = age , ylab = "rate")

  points(mattest.mrate[,i],col="red")

  legend("topleft",legend =
c("Forecast","Actual"),col=c("blue","red"),cex=0.8,pch = c(1,1),text.font =
36 )

}
```

```
#----------------------- End for female mortality modeling ------------------
----#

# prepare data for calculating mortality #

# import mortality data and change its format to martix form #

df.molm<-read.csv("D:/Senior project/Prepared data/Mortality Male.csv",header
= TRUE,sep=",")

df.molm<-df.molm[,3:22]

colnames(df.molm)<-c(2540:2559)

rownames(df.molm)<-c(0:100)

mat.molm<-data.matrix(df.molm)

# import population data and change its format to martix form #

df.popm<-read.csv("D:/Senior project/Prepared data/Population
Male.csv",header = TRUE,sep=",")

df.popm<-df.popm[,3:22]

colnames(df.popm)<-c(2540:2559)

rownames(df.popm)<-c(0:100)

mat.popm<-data.matrix(df.popm)

head(df.popm)

# create matrix mortality rate#

mat.mrate<-mat.molm/mat.popm

mat.mrate

# prepare to STMOMO format#

year<-c(2540:2559)

age<-c(0:100)

# prepare demogdata format #

tm<-demogdata(data = mat.mrate,pop = mat.popm,ages = age,years = year,type =
"mortality",label = "Thai",name = "male" )

tmstmomo<-StMoMoData(tm,"male")

class(tfstmomo)

tfstmomo$Dxt

# fir moidel#
```

```
Lcfitm<-fit(lc(),data=tmstmomo, ages.fit = 0:100)


plot(Lcfitm)

# forecast k #

Lcform<-forecast(Lcfitm,h=3)

Lcform$rates

plot(Lcform)


dfm<-data.frame(age=c(0:100),Lcfitm$ax,Lcfitm$bx)

#write.csv(dfm,"D:/Senior project/output_male.csv",row.names=F)

#write.csv(Lcfitm$kt,"D:/Senior project/output_kt_male.csv",row.names=F)


test.mm<-read.csv("D:/Senior project/Prepared data/test set/Mortality
Male.Test.csv",header = TRUE,sep=",")

test.mm<-test.mm[,2:4]

colnames(test.mm)<-c(2560,2561,2562)

rownames(test.mm)<-c(0:100)

mattest.mm<-data.matrix(test.mm)


test.pm<-read.csv("D:/Senior project/Prepared data/test set/Population
Male.Test.csv",header = TRUE,sep=",")

test.pm<-test.pm[,2:4]

colnames(test.pm)<-c(2560,2561,2562)

rownames(test.pm)<-c(0:100)

mattest.pm<-data.matrix(test.pm)


mattest.mrate<-mattest.mm/mattest.pm


dim(mattest.mrate)

a<-Lcform$rates


Vm<-abs((a-mattest.mrate)/mattest.mrate)
```

```
MAPEM<-matrix(nrow = 101,ncol = 1)
for(i in 1:101){

  MAPEM[i,]<-sum(Vm[i,])/3
}
MAPEM
par(mfrow=c(3,1))
yeartest<-c(2560,2561,2562)

for (i in 1:3) {

  plot(Lcform$rates[,i],col="blue",main = paste("Mortality rate Male
",yeartest[i]),xlab = age , ylab = "rate")
  points(mattest.mrate[,i],col="red")
  legend("topleft",legend =
c("Forecast","Actual"),col=c("blue","red"),cex=0.8,pch = c(1,1),text.font =
36 )
}

#----------------------------------------- End ------------------------------
-----------#

par(mfrow=c(1,1))
plot(MAPE,col="red" , xlab="age")
points(MAPEM,col="blue", xlab="age")
legend("topleft",legend =
c("Forecast","Actual"),col=c("blue","red"),cex=0.8,pch = c(1,1),text.font =
36 )

#------------- for forecast baye --------------#
```

```
kt_bay_male<-read.csv("D:/Senior project/output_ktbay_male.csv",header =T
,sep=",")

kt_bay_male<-t(data.matrix(kt_bay_male))

auto.arima(kt_bay_male)

k_baye_fit<-auto.arima(kt_bay_male)
k_for_baye_<-forecast(k_baye_fit)

kap_baye_for<-k_for_baye_$mean
kap_baye_for<-kap_baye_for[1:3]

a_baye<-read.csv("D:/Senior project/output_axbay_male.csv")
a_baye<-t(as.matrix(a_baye))
a_baye_for<-cbind(a_baye,a_baye,a_baye)


b<-read.csv("D:/Senior project/output_bxbay_male.csv")
b<-t(as.matrix(b))
b_baye_for<-cbind(b,b,b)

baye_rate_male<-exp(a_baye_for+b_baye_for*kap_baye_for)

rownames(baye_rate)<-c(0:100)
colnames(baye_rate)<-c(2560:2562)

baye_rate_male
Vmb<-abs((baye_rate_male-mattest.mrate)/mattest.mrate)
```

```
MAPEMB<-matrix(nrow = 101,ncol = 1)
for(i in 1:101){

  MAPEMB[i,]<-sum(Vmb[i,])/3
}
par(mfrow=c(3,1))
for (i in 1:3) {

  plot(baye_rate_male[,i],col="blue",main = paste("Mortality rate Male with
baye",yeartest[i]),xlab = age , ylab = "rate")
  points(mattest.mrate[,i],col="red")
  legend("topleft",legend =
c("Forecast","Actual"),col=c("blue","red"),cex=0.8,pch = c(1,1),text.font =
36 )
}
#---------- female ---------- #

kt_bay_female<-read.csv("D:/Senior project/output_ktbay_female.csv",header =T
,sep=",")

kt_bay_female<-t(data.matrix(kt_bay_female))

k_baye_fit_f<-auto.arima(kt_bay_female)
k_for_baye_f<-forecast(k_baye_fit_f)

kap_baye_for_f<-k_for_baye_f$mean
kap_baye_for_f<-kap_baye_for_f[1:3]

a_baye_f<-read.csv("D:/Senior project/output_axbay_female.csv")
a_baye_f<-t(as.matrix(a_baye_f))
```

```
a_baye_for_f<-cbind(a_baye_f,a_baye_f,a_baye_f)


b_f<-read.csv("D:/Senior project/output_bxbay_female.csv")
b_f<-t(as.matrix(b_f))
b_baye_for_f<-cbind(b_f,b_f,b_f)



baye_rate_female<-exp(a_baye_for_f+b_baye_for_f*kap_baye_for_f)

rownames(baye_rate_female)<-c(0:100)
colnames(baye_rate_female)<-c(2560:2562)


Vfb<-abs((baye_rate_female-mattest.rate)/mattest.rate)

MAPEFB<-matrix(nrow = 101,ncol = 1)
for(i in 1:101){

  MAPEFB[i,]<-sum(Vfb[i,])/3
}
par(mfrow=c(3,1))
for (i in 1:3) {


  plot(baye_rate_female[,i],col="blue",main = paste("Mortality rate Female with baye",yeartest[i]),xlab = age , ylab = "rate")
  points(mattest.rate[,i],col="red")
  legend("topleft",legend = c("Forecast","Actual"),col=c("blue","red"),cex=0.8,pch = c(1,1),text.font = 36 )
}
```