โครงการ

การเรียนการสอนเพื่อเสริมประสบการณ์

ชื่อโครงการ การศึกษาข้อมูลประกันชีวิตของประเทศไทยโดยใช้แบบจำลองอนุกรมเวลา
A study of insurance data of Thailand using time series models

ชื่อนิสิต นายศราวุฒิ ทองเชื้อ 593 35490 23

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์
สาขาวิชา คณิตศาสตร์

ปีการศึกษา 2562

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การศึกษาข้อมูลประกันชีวิตของประเทศไทยโดยใช้แบบจำลองอนุกรมเวลา

นายศราวุฒิ ทองเชื้อ

โครงงานนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตร์บัณฑิต

สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2562

A study of insurance data of Thailand using time series models

Sarawut Tongchue

A Project Submitted in Partial Fulfillment of the Requirements

for the Degree of Bachelor of Science Program in Mathematics

Department of Mathematics and Computer Science

Faculty of Science

Chulalongkorn University

Academic Year 2019

หัวข้อโครงงาน การศึกษาข้อมูลประกันชีวิตของประเทศไทยโดยใช้แบบจำลองอนุกรมเวลา

โดย นายศราวุฒิ ทองเชื้อ

สาขาวิชา คณิตศาสตร์

อาจารย์ที่ปรึกษาโครงงาน ผู้ช่วยศาสตราจารย์ ดร. จิราพรรณ สุนทรโชติ

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับโครงงานฉบับนี้เป็นส่วนหนึ่ง ของการศึกษาตามหลักสูตรปริญญาบัณฑิต ในรายวิชา 2301499 โครงงานวิทยาศาสตร์ (Senior Project)

.... หัวหน้าภาควิชาคณิตศาสตร์

(ศาสตราจารย์ ดร.กฤษณะ เนียมมณี) และวิทยาการคอมพิวเตอร์

คณะกรรมการสอบโครงงาน

.... อาจารย์ที่ปรึกษาโครงงานหลัก

(ผู้ช่วยศาสตราจารย์ ดร. จิราพรรณ สุนทรโชติ)

.... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร. พงษ์เดช มนทกานติรัตน์)

.... กรรมการ

(อาจารย์ ดร. อธิปัตย์ ธำรงธัญลักษณ์)

นายศราวุฒิ ทองเชื้อ: การศึกษาข้อมูลประกันชีวิตของประเทศไทยโดยใช้แบบจำลองอนุกรมเวลา (A study of insurance data of Thailand using time series models)

อ.ที่ปรึกษาโครงงานหลัก: ผู้ช่วยศาสตราจารย์ ดร. จิราพรรณ สุนทรโชติ, 110 หน้า.

ในโครงงานนี้ เรามีความสนใจในการศึกษาการประยุกต์ใช้อนุกรมเวลาและอนุกรมเวลาเชิงการนับในการศึกษาข้อมูลประกันชีวิตของประเทศไทยใน 2 ประเด็นคือ (1) ศึกษาแบบจำลองอนุกรมเวลาสำหรับข้อมูลจำนวนเงินเอาประกันภัยของบริษัทประกันภัยในประเทศไทย (2) ศึกษาแบบจำลองอนุกรมเวลาเชิงการนับสำหรับข้อมูลจำนวนการเอาประกันภัยของบริษัทประกันภัยในประเทศไทย ข้อมูลของประกันชีวิตที่ศึกษาในโครงงานนี้ ได้แก่ ประกันชีวิตแบบตลอดชีพ ประกันชีวิตแบบสะสมทรัพย์ และประกันชีวิตแบบชั่วระยะเวลา ข้อมูลที่ใช้ในการศึกษานี้คือข้อมูลรายเดือนประกันชีวิตของประเทศไทยในช่วงปี พ.ศ. 2552-2556 ซึ่งเก็บรวบรวมข้อมูลจากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย โดยใช้แบบจำลองและชุดคำสั่งในโปรแกรม R

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์ ลายมือชื่อนิสิต ศราวุฒิ ทองเชื้อ

สาขาวิชา คณิตศาสตร์ ลายมือชื่อ อ.ที่ปรึกษาโครงงานหลัก

ปีการศึกษา 2562

5933549023 : MAJOR MATHEMATICS

KEYWORDS : TIME SERIES, COUNT TIME SERIES, INSURANCE CLAIM COUNT

SARAWUT TONGCHUE: A STUDY OF INSURANCE DATA OF THAILAND USING TIME SERIES MODELS. ADVISOR: ASST. JIRAPHAN SUNTORNCHOST, Ph.D., 110 pp.

In this project, we are interested in studying the application of time series and count time series to study life insurance data in Thailand in two aspects which are (1) studying the application of time series models for the amount of claims for data of the insurance companies in Thailand (2) studying the count time series model for the number of claims of the insurance companies in Thailand. There are three types of life insurance studied in this study which are the whole life insurance, the term life insurance, and the endowment insurance. The data used in this study are monthly data from 2009 to 2013 produced by Office of Insurance Commission. All analyses are performed via the R software.

Department: Mathematics and Computer Science Student's Signature Sarawut Tongchue

Field of Study : Mathematics Advisor's Signature J.suntornchost

Academic Year : 2019

กิตติกรรมประกาศ

โครงงานการศึกษาจำนวนการเอาประกันภัยโดยใช้แบบจำลองอนุกรมเวลาเชิงการนับฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ด้วยความช่วยเหลือของผู้ช่วยศาสตราจารย์ ดร. จิราพรรณ สุนทรโชติ อาจารย์ที่ปรึกษาโครงงาน ซึ่งท่านได้ให้ความรู้ ให้คำปรึกษา และให้คำแนะนำต่างๆซึ่งเป็นประโยชน์ อีกทั้งยังช่วยแก้ปัญหาต่างๆที่เกิดขึ้นระหว่างการดำเนินงาน

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. พงษ์เดช มนทกานติรัตน์ และ อาจารย์ ดร. อธิปัตย์ ธำรงธัญลักษณ์ ที่ให้ความกรุณาเป็นกรรมการในการนำเสนอโครงการ ให้คำแนะนำและเสนอแนะต่างๆ เพื่อนำไปปรับปรุงและแก้ไขโครงงานให้ดียิ่งขึ้น นอกจากนี้ขอขอบพระคุณอาจารย์ทุกๆท่านและขอบคุณเพื่อนๆทุกคนที่เป็นกำลังใจ และให้ความช่วยเหลือในการทำโครงงานนี้

สุดท้ายนี้ ผู้จัดทำขอขอบพระคุณบิดา มารดา และครอบครัว ที่เปิดโอกาสให้ได้รับการศึกษาเล่าเรียน คอยให้ความช่วยเหลือทุกๆด้าน และให้กำลังใจผู้จัดทำเสมอมา

สารบัญ

หน้า

หน้า

สารบัญตาราง

หน้า

หน้า

สารบัญภาพ

หน้า

หน้า

หน้า

หน้า

บทที่ 1

บทนำ

1.1 ความเป็นมาและเหตุผลการวิจัย

ปัจจุบันการทำประกันชีวิตได้รับความนิยมเป็นอย่างมาก เนื่องจากสามารถให้ความคุ้มครองต่อบุคคลและครอบครัว ทำให้บริษัทประกันภัยนั้นจำเป็นต้องหาแนวทางในการคาดการณ์จำนวนการเอาประกันภัย ซึ่งอนุกรมเวลาและอนุกรมเวลาเชิงการนับเป็นวิธีการทางสถิติหนึ่งที่นิยมในการนำมาใช้พยากรณ์ข้อมูล

อนุกรมเวลา คือ อนุกรมของข้อมูลที่มีการเก็บรวบรวมข้อมูลของตัวแปรตามลำดับเวลา เช่น ราคาหุ้นรายวันของบริษัท และปริมาณน้ำรายเดือนในเขื่อน แบบจำลองอนุกรมเวลาที่ได้รับความนิยม คือ แบบจำลอง ARIMA (Autoregressive Integrated Moving Average) ซึ่งการวิเคราะห์อนุกรมเวลาเป็นวิธีการทางสถิติที่นิยมในการนำมาใช้พยากรณ์ข้อมูลที่เกิดขึ้นในอนาคต

อนุกรมเวลาเชิงการนับ คือ ข้อมูลจำนวนของเหตุการณ์ต่อเวลาในช่วงเวลาที่ถูกสังเกตโดยอาจเก็บข้อมูลเป็นรายวัน รายเดือน หรือรายปีขึ้นอยู่กับวัตถุประสงค์ในการนำไปใช้ อนุกรมเวลาเชิงการนับมักพบเห็นได้ในสถานการณ์ต่างๆ เช่น จำนวนการเอาประกันภัยรายเดือนของบริษัทประกันภัย และจำนวนผู้ป่วยรายสัปดาห์ที่เข้ารับการรักษาในโรงพยาบาล ซึ่งอนุกรมเวลาเชิงการนับมีประโยชน์อย่างมากในการวิเคราะห์และวางแผนตัดสินใจทางธุรกิจ หรือคาดการณ์ขั้นแผนงานให้มีความผิดพลาดน้อยที่สุด โดยใช้ข้อมูลในอดีตเป็นพื้นฐานสำหรับการพยากรณ์ข้อมูลในอนาคต

แบบจำลองอนุกรมเวลาเชิงการนับได้รับความสนใจอย่างกว้างขวางในบรรดานักสถิติในการสร้างแบบจำลองสำหรับสถานการณ์ต่างๆ เช่น Fokianos และคณะ (2009) ได้นำเสนอแบบจำลองอนุกรมเวลาเชิงการนับสำหรับตัวแบบการ์ชแบบปัวซง และศึกษาบทประยุกต์สำหรับจำนวนการซื้อขายหุ้นในตลาดหลักทรัพย์ และ Agosto และคณะ (2015) ได้นำเสนอแบบจำลองอนุกรมเวลาเชิงการนับแบบปัวซงที่มีตัวแปรร่วม และศึกษาบทประยุกต์ในการนับจำนวนล้มละลายของบริษัทในประเทศสหรัฐอเมริกาช่วงปีค.ศ. 1982-2011 ต่อมาในปี 2017 Liboschik และคณะ ได้สร้างชุดคำสั่ง tscount ในโปรแกรม R สำหรับแบบจำลองอนุกรมเวลาเชิงการนับ ซึ่งทำให้มีความสะดวกมากขึ้นในการใช้งานแบบจำลองอนุกรมเวลาเชิงการนับ

โครงงานนี้ เรามีความสนใจในการศึกษาอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมสำหรับข้อมูลประกันชีวิตของบริษัทประกันชีวิตในประเทศไทยในช่วงปีพ.ศ. 2552-2556 โดยใช้แบบจำลองและชุดคำสั่งในโปรแกรม R

1.2 วัตถุประสงค์ของการวิจัย

เราจะศึกษาหาแบบจำลองอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมที่สุดสำหรับข้อมูลประกันชีวิตของบริษัทประกันชีวิตในประเทศไทย

1.3 ขอบเขตของการวิจัย

ในโครงงานนี้ เราจะศึกษาอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมสำหรับข้อมูลประกันชีวิตของบริษัทประกันชีวิตในประเทศไทยโดยศึกษาผ่านชุดคำสั่งในโปรแกรม R

ข้อมูลที่ใช้ในการศึกษานี้คือข้อมูลจำนวนเงินเอาประกันภัย และจำนวนการเอาประกันภัย ในช่วง พ.ศ. 2552-2556 จากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย

1.4 ขั้นตอนการวิจัย

1. ศึกษาความรู้พื้นฐานเกี่ยวกับการวิเคราะห์อนุกรมเวลาและอนุกรมเวลาเชิงการนับ
2. ศึกษาการใช้โปรแกรม R
3. ประยุกต์ใช้ชุดคำสั่งในโปรแกรม R สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาและอนุกรมเวลาเชิงการนับ
4. ศึกษาหาแบบจำลองอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมสำหรับข้อมูลประกันชีวิตของบริษัทประกันชีวิตในประเทศไทย
5. สรุปผลและเขียนรายงาน

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ประโยชน์ที่คาดว่าจะได้รับต่อผู้ดำเนินงาน

1. มีความรู้เกี่ยวกับการวิเคราะห์อนุกรมเวลาและอนุกรมเวลาเชิงการนับ
2. มีความรู้เกี่ยวกับกรมธรรม์ประกันภัยของบริษัทประกันภัยในประเทศไทย
3. มีความรู้ในการใช้โปรแกรม R
4. มีความรู้ในการประยุกต์ใช้ชุดคำสั่งในโปรแกรม R

1.5.2 ประโยชน์ที่ได้จากโครงงานที่พัฒนาขึ้น

1. เพื่อใช้เป็นความรู้ต่อยอดงานวิจัย
2. สามารถคาดการณ์จำนวนเงินเอาประกันชีวิตและจำนวนการเอาประกันชีวิตของบริษัทประกันชีวิตในประเทศไทยได้

1.6 โครงสร้างของรายงาน

บทที่ 2 จะกล่าวถึงความรู้พื้นฐานใน 2 ส่วน ส่วนแรกคือ อนุกรมเวลา จะกล่าวถึง บทนิยามอนุกรมเวลา แบบจำลองอนุกรมเวลา การเลือกแบบจำลองที่เหมาะสม และการตรวจสอบความถูกต้องของแบบจำลอง ส่วนที่สองคือ อนุกรมเวลาเชิงการนับ จะกล่าวถึง บทนิยามอนุกรมเวลาเชิงการนับ แบบจำลองอนุกรมเวลาเชิงการนับ และการวิเคราะห์แบบจำลองที่เหมาะสม

บทที่ 3 จะกล่าวถึงการศึกษาอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมสำหรับข้อมูลจำนวนการเอาประกันภัยและจำนวนเงินเอาประกันภัยของบริษัทประกันชีวิตในประเทศไทย ในกลุ่มข้อมูลของประกันชีวิต 3 ประเภท ได้แก่ ประกันชีวิตแบบตลอดชีพ ประกันชีวิตแบบสะสมทรัพย์และประกันชีวิตแบบชั่วระยะเวลา ในช่วงปี พ.ศ. 2552-2556 โดยใช้แบบจำลองและชุดคำสั่งในโปรแกรม R และผลการดำเนินงาน ได้แก่ การหาอันดับของแบบจำลอง การประมาณค่าพารามิเตอร์ของแต่ละแบบจำลอง การตรวจสอบค่าสหสัมพันธ์ของเศษเหลือ การหาแบบจำลองสำหรับการจำลอง และการหาแบบจำลองสำหรับการพยากรณ์

บทที่ 2

ความรู้พื้นฐาน

ในบทนี้ จะกล่าวถึงความรู้พื้นฐานใน 2 ส่วน ส่วนแรกคือ อนุกรมเวลา จะกล่าวถึง บทนิยามอนุกรมเวลา แบบจำลองอนุกรมเวลา การเลือกแบบจำลองที่เหมาะสม และการตรวจสอบความถูกต้องของแบบจำลอง ส่วนที่สองคือ อนุกรมเวลาเชิงการนับ จะกล่าวถึง บทนิยามอนุกรมเวลาเชิงการนับ แบบจำลองอนุกรมเวลาเชิงการนับ และการวิเคราะห์แบบจำลองที่เหมาะสม

2.1 อนุกรมเวลา

บทนิยาม 1. อนุกรมเวลา (Time Series)

อนุกรมเวลา คือ อนุกรมของข้อมูลที่มีการเก็บรวบรวมข้อมูลของตัวแปรตามลำดับเวลา โดยทั่วไปอนุกรมเวลาแบบสุ่มของตัวแปรหนึ่งจะเขียนแทนด้วยสัญลักษณ์ X_t หรือ $\{X_t: t \in T\}$ เมื่อ T คือเซตของเวลา เช่น ราคาหุ้นรายวันของบริษัท และปริมาณน้ำรายเดือนในเขื่อน

บทนิยาม 2. ตัวรบกวนขาวหรือปัจจัยภายนอก (White Noise)

ถ้า $u_1, u_2, \dots, u_T$ คือ ลำดับของตัวแปรสุ่มที่เป็นอิสระต่อกันและมีการแจกแจงเหมือนกันที่มีค่าเฉลี่ยและความแปรปรวนเป็นค่าคงที่แล้ว อนุกรม u_t $(t = 1,2,\dots,T)$ ถูกเรียกว่าตัวรบกวนขาว (White Noise) และถ้าพบว่าอนุกรม u_t มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็นศูนย์ และความแปรปรวนเป็น σ^2 แล้วเราจะเรียกอนุกรมเวลา u_t ว่า ตัวรบกวนขาวแบบเกาส์เซียน (Gaussian White Noise)

บทนิยาม 3. ความนิ่งของข้อมูล (Stationary)

อนุกรมเวลาของตัวแปร X_t จะมีความนิ่งก็ต่อเมื่อมีคุณสมบัติดังต่อไปนี้

1. ค่าเฉลี่ยของตัวแปร X_t ในแต่ละช่วงเวลา t มีค่าคงที่
2. ความแปรปรวนของตัวแปร X_t ในแต่ละช่วงเวลา t มีค่าคงที่
3. ความแปรปรวนร่วมของตัวแปร X_t ณ เวลา t_1 และ t_2 ซึ่ง $t_1 \neq t_2$ มีค่าคงที่

2.1.1 แบบจำลองอนุกรมเวลา

บทนิยาม 4. แบบจำลองการถดถอยในตัว หรือแบบจำลอง Autoregressive (AR)

แบบจำลองการถดถอยในตัว หรือแบบจำลอง Autoregressive เป็นแบบจำลองที่ใช้อธิบายอนุกรมเวลาที่ขึ้นอยู่กับข้อมูลในอดีตเพื่อใช้ในการพยากรณ์ข้อมูลในอนาคต โดยมีความสัมพันธ์เป็นฟังก์ชันเชิงเส้นกับข้อมูลในอดีต ซึ่งสามารถเขียนสมการของแบบจำลอง Autoregressive ลำดับที่ p หรือใช้สัญลักษณ์ AR(p) ได้ดังนี้

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_p X_{t-p} + \varepsilon_t$$

โดย X_t คือ อนุกรมเวลา

$\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_p$ คือ ค่าพารามิเตอร์

ε_t คือ ตัวแปรสุ่มคลาดเคลื่อนที่มีคุณสมบัติเป็นตัวรบกวนขาวแบบเกาส์เซียน

จากนิยามข้างต้น เราจะยกตัวอย่างรูปแบบทั่วไปของแบบจำลอง Autoregressive ดังนี้

รูปแบบทั่วไปของแบบจำลอง AR(1) เขียนได้ดังนี้

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \varepsilon_t$$

รูปแบบทั่วไปของแบบจำลอง AR(2) เขียนได้ดังนี้

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$$

บทนิยาม 5. แบบจำลองการเฉลี่ยเคลื่อนที่ หรือแบบจำลอง Moving Average (MA)

แบบจำลองการเฉลี่ยเคลื่อนที่ หรือแบบจำลอง Moving Average เป็นแบบจำลองที่ใช้อธิบายอนุกรมเวลาที่ขึ้นอยู่กับปัจจัยภายนอก (ε) ตั้งแต่อดีตจนถึงปัจจุบันเพื่อใช้ในการพยากรณ์ข้อมูลในอนาคต โดยมีความสัมพันธ์เป็นฟังก์ชันเชิงเส้นกับปัจจัยภายนอกในอดีต ซึ่งสามารถเขียนสมการของแบบจำลอง Moving Average ลำดับที่ q หรือใช้สัญลักษณ์ MA(q) ได้ดังนี้

$$X_t = \beta_0 - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \cdots - \beta_q \varepsilon_{t-q} + \varepsilon_t$$

โดย X_t คือ อนุกรมเวลา

$\beta_0, \beta_1, \beta_2, \dots, \beta_q$ คือ ค่าพารามิเตอร์

ε_t คือ ตัวรบกวนขาวที่มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนเป็น σ^2

จากนิยามข้างต้น เราจะยกตัวอย่างรูปแบบทั่วไปของแบบจำลอง Moving Average ดังนี้

รูปแบบทั่วไปของแบบจำลอง MA(1) เขียนได้ดังนี้

$$X_t = \beta_0 - \beta_1 \varepsilon_{t-1} + \varepsilon_t$$

รูปแบบทั่วไปของแบบจำลอง MA(2) เขียนได้ดังนี้

$$X_t = \beta_0 - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} + \varepsilon_t$$

บทนิยาม 6. แบบจำลองการเฉลี่ยเคลื่อนที่ถดถอยในตัว หรือแบบจำลอง Autoregressive Moving Average (ARMA)

แบบจำลองการเฉลี่ยเคลื่อนที่ถดถอยในตัว หรือแบบจำลอง Autoregressive Moving Average เป็นแบบจำลองที่ประยุกต์จากแบบจำลอง Autoregressive และแบบจำลอง Moving Average มารวมกัน โดยเป็นแบบจำลองที่ใช้อธิบายอนุกรมเวลาที่ขึ้นอยู่กับข้อมูลในอดีตและปัจจัยภายนอกตั้งแต่อดีตถึงปัจจุบัน ซึ่งสามารถเขียนสมการของแบบจำลอง Autoregressive Moving Average ลำดับที่ (p,q) หรือใช้สัญลักษณ์ ARMA(p,q) ได้ดังนี้

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_p X_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \cdots - \beta_q \varepsilon_{t-q}$$

โดย X_t คือ อนุกรมเวลา

$\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_p$ และ $\beta_1, \beta_2, \dots, \beta_q$ คือ ค่าพารามิเตอร์

ε_t คือ ตัวรบกวนขาวที่มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนเป็น σ^2

จากนิยามข้างต้น เราจะยกตัวอย่างรูปแบบทั่วไปของแบบจำลอง Autoregressive Moving Average ดังนี้

รูปแบบทั่วไปของแบบจำลอง ARMA(1,1) เขียนได้ดังนี้

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \varepsilon_t - \beta_1 \varepsilon_{t-1}$$

รูปแบบทั่วไปของแบบจำลอง ARMA(1,2) เขียนได้ดังนี้

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2}$$

บทนิยาม 7. อนุกรมเวลา X_t อยู่ในรูปผลรวมลำดับที่ d (Integrate of order d)

เนื่องจากแบบจำลอง ARMA จะใช้ได้กับอนุกรมเวลาที่มีความนิ่งเท่านั้น หากพบว่าอนุกรมเวลาไม่มีความนิ่ง เราจะต้องแปลงอนุกรมเวลานั้นให้นิ่งก่อน และวิธีที่นิยมนำมาใช้แปลงอนุกรมเวลาคือ วิธีการหาผลต่างของอนุกรมเวลา โดยการทำผลต่างลำดับที่ d ของอนุกรมเวลา X_t เขียนได้ดังนี้

$$\Delta^d X_t = \Delta^{d-1}(X_t - X_{t-1})$$

โดยที่ Δ^d คือ ผลต่างอันดับที่ d

จากนิยามข้างต้น เราจะยกตัวอย่างผลต่างอันดับที่ d ดังนี้

ตัวอย่าง ผลต่างอันดับที่หนึ่งของอนุกรมเวลา X_t เขียนได้ดังนี้

$$\Delta X_t = X_t - X_{t-1}$$

ตัวอย่าง ผลต่างอันดับที่สองของอนุกรมเวลา X_t เขียนได้ดังนี้

$$\begin{aligned}
\Delta^2 X_t &= \Delta(\Delta X_t) \\
&= \Delta(X_t - X_{t-1}) \\
&= \Delta X_t - \Delta X_{t-1} \\
&= (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) \\
&= X_t - 2X_{t-1} + X_{t-2}
\end{aligned}$$

บทนิยาม 8. แบบจำลองรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ หรือแบบจำลอง Autoregressive Integrated Moving Average (ARIMA)

แบบจำลองรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ หรือแบบจำลอง Autoregressive Integrated Moving Average เป็นแบบจำลองของอนุกรมเวลาที่ขึ้นอยู่กับข้อมูลในอดีตและขึ้นอยู่กับปัจจัยภายนอกตั้งแต่อดีตถึงปัจจุบัน โดยผ่านวิธีการหาผลต่าง (differencing) ให้อนุกรมเวลามีความนิ่ง (Stationary) ซึ่งแบบจำลอง Autoregressive Integrated Moving Average ลำดับที่ (p,q) โดยผ่านวิธีการหาผลต่างลำดับที่ d หรือใช้สัญลักษณ์ ARIMA(p,d,q)

จากนิยามข้างต้น เราจะยกตัวอย่างรูปแบบทั่วไปของแบบจำลอง Autoregressive Integrated Moving Average ดังนี้

รูปแบบทั่วไปของแบบจำลอง ARIMA(1,1,1) เขียนได้ดังนี้

$$\Delta X_t = \alpha_0 + \alpha_1 \Delta X_{t-1} + \varepsilon_t - \beta_1 \varepsilon_{t-1}$$

แทนค่า $\Delta X_t = X_t - X_{t-1}$ จะได้ว่า

$$X_t - X_{t-1} = \alpha_0 + \alpha_1(X_{t-1} - X_{t-2}) + \varepsilon_t - \beta_1 \varepsilon_{t-1}$$

$$X_t = \alpha_0 + X_{t-1} + \alpha_1 X_{t-1} - \alpha_1 X_{t-2} + \varepsilon_t - \beta_1 \varepsilon_{t-1}$$

ดังนั้น

$$X_t = \alpha_0 + (1 + \alpha_1) X_{t-1} - \alpha_1 X_{t-2} + \varepsilon_t - \beta_1 \varepsilon_{t-1}$$

2.1.2 การเลือกแบบจำลองที่เหมาะสม

บทนิยาม 9. ค่าสหสัมพันธ์ในตัวเอง Autocorrelation Function (ACF)

กำหนดให้ $y_1, y_2, ..., y_T$ คืออนุกรมเวลาชุดหนึ่งที่มีความนิ่งจำนวน T ข้อมูล การหาค่าสหสัมพันธ์ในตัวเองอันดับที่ k นั้นจะดูความสัมพันธ์ระหว่างข้อมูลอนุกรมเวลา y_t กับอนุกรมเวลา y_{t+k} เขียนสมการได้ดังนี้

$$\rho_k = \frac{E(y_t - \mu)(y_{t+k} - \mu)}{\sqrt{Var(y_t)Var(y_{t+k})}}$$

ซึ่งสามารถประมาณได้ด้วยค่าสหสัมพันธ์จากตัวอย่าง r_k ดังนี้

$$r_k = \frac{\sum_{t=1}^{T-k}(y_t - \mu_y)(y_{t+k} - \mu_y)}{\sum_{t=1}^{T}(y_t - \mu_y)^2}$$

โดยที่ μ_y คือ ค่าเฉลี่ยของข้อมูลอนุกรมเวลา

r_k คือ ค่าสัมประสิทธิ์สหสัมพันธ์

คุณสมบัติของค่าสัมประสิทธิ์สหสัมพันธ์

1. $-1 \le r_k \le 1$
2. $r_k = 0$ คือ ตัวแปรทั้งสองไม่มีความสัมพันธ์กัน
3. $|r_k| = 1$ คือ ตัวแปรทั้งสองมีความสัมพันธ์กันอย่างมาก
4. $r_k > 0$ คือ ตัวแปรทั้งสองมีความสัมพันธ์ในทิศทางเดียวกัน
5. $r_k < 0$ คือ ตัวแปรทั้งสองมีความสัมพันธ์ในทิศทางตรงกันข้าม

บทนิยาม 10. ค่าสหสัมพันธ์ในตัวเองบางส่วน Partial Autocorrelation Function (PACF)

การหาค่าสหสัมพันธ์ในตัวเองบางส่วนจะดูความสัมพันธ์ระหว่างข้อมูลอนุกรมเวลา y_t กับอนุกรมเวลา y_{t+k} โดยไม่สนใจข้อมูลระหว่างเวลา t กับ $t+k$ เขียนสมการได้ดังนี้

$$\rho_k = \frac{Cov(y_t, y_{t+k} \mid y_{t+1}, \ldots, y_{t+k-1})}{\sqrt{Var(y_t \mid y_{t+1}, \ldots, y_{t+k-1})}\sqrt{Var(y_{t+k} \mid y_{t+1}, \ldots, y_{t+k-1})}}$$

โดยที่ ρ_k คือ ค่าสัมประสิทธิ์สหสัมพันธ์ มีค่าตั้งแต่ -1 ถึง +1

เมื่อ $\rho_k = 0$ คือ ตัวแปรทั้งสองไม่มีความสัมพันธ์กัน

$|\rho_k| = 1$ คือ ตัวแปรทั้งสองมีความสัมพันธ์กัน

$\rho_k > 0$ คือ ตัวแปรทั้งสองมีความสัมพันธ์ในทิศทางเดียวกัน

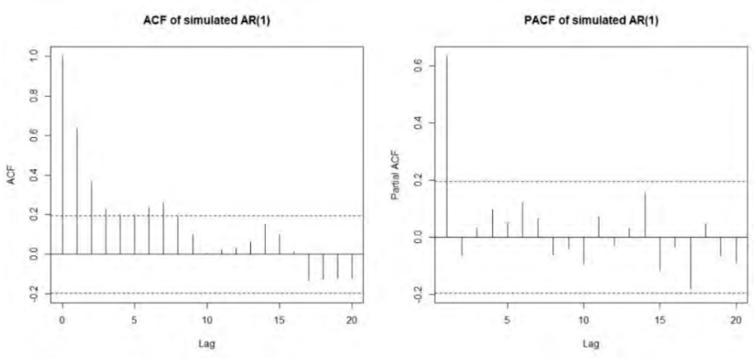
$\rho_k < 0$ คือ ตัวแปรทั้งสองมีความสัมพันธ์ในทิศทางตรงกันข้าม

การเลือกอันดับของแบบจำลอง ARMA(p,q)

การเลือกอันดับของแบบจำลอง ARMA(p,q) ดูลักษณะของกราฟค่าสหสัมพันธ์ในตัวเอง (Autocorrelation plot) และกราฟค่าสหสัมพันธ์ในตัวเองบางส่วน (Partial Autocorrelation plot) ใช้เป็นแนวทางในการเลือกแบบจำลอง

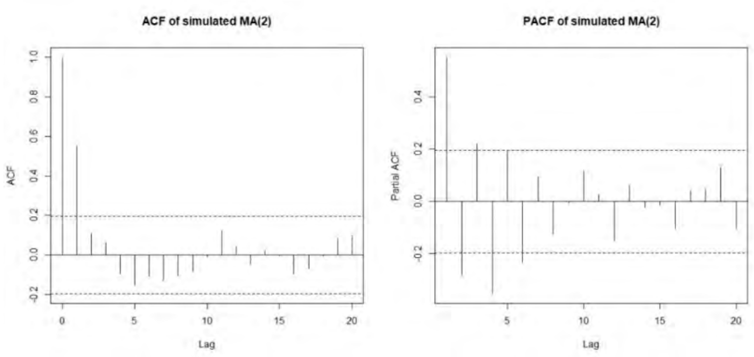
แบบจำลอง	ลักษณะของค่า ACF	ลักษณะของค่า PACF
AR(p)	ลดลงอย่างรวดเร็ว	สิ้นสุดหลัง p ช่วงเวลาที่แล้ว
MA(q)	สิ้นสุดหลัง q ช่วงเวลาที่แล้ว	ลดลงอย่างรวดเร็ว
ARMA(p,q)	ลดลงเรื่อยๆอย่างรวดเร็วหลัง q ช่วงเวลาที่แล้ว	ลดลงเรื่อยๆอย่างรวดเร็วหลัง p ช่วงเวลาที่แล้ว

ตัวอย่างแบบจำลอง AR(1)



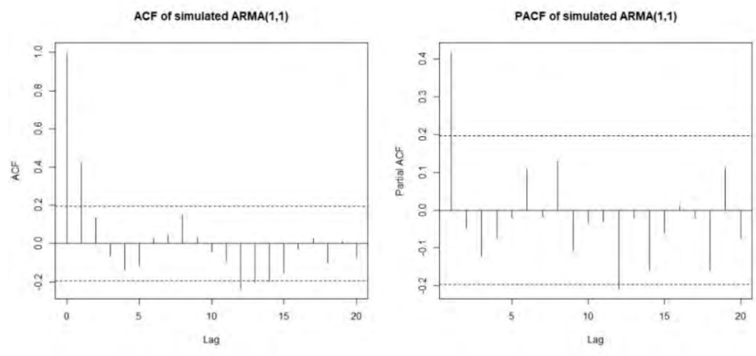
ภาพที่ ก ตัวอย่างแบบจำลอง AR(1)

ตัวอย่างแบบจำลอง MA(2)



ภาพที่ ข ตัวอย่างแบบจำลอง MA(2)

ตัวอย่างแบบจำลอง ARMA(1,1)



ภาพที่ ค ตัวอย่างแบบจำลอง ARMA(1,1)

บทนิยาม 11. เกณฑ์ข้อสนเทศอาไคเคะหรือเกณฑ์เอไอซี (Akaike Information Criterion : AIC)

เกณฑ์การคัดเลือกแบบจำลองที่เหมาะสมภายใต้แบบจำลองที่มีการใช้ข้อมูลชุดเดียวกันแต่มีจำนวนพารามิเตอร์แตกต่างกันที่ใช้อธิบายความสัมพันธ์ระหว่างตัวแปรตอบสนองกับตัวแปรอธิบายชุดหนึ่ง แบบจำลองที่เหมาะสมกว่า คือ แบบจำลองที่มีค่าเอไอซีต่ำกว่า มักใช้ในการวิเคราะห์อนุกรมเวลา เช่น ในกรณีที่มีการประมาณค่าพารามิเตอร์ของแบบจำลองด้วยภาวะน่าจะเป็นสูงสุด $AIC = -2\ln L + 2p$ และในกรณีมีการประมาณค่าพารามิเตอร์ของแบบจำลองด้วยวิธีกำลังสองน้อยสุด

$$AIC = n\ln\left[\frac{SSE}{n}\right] + 2p$$

โดยที่ n คือ จำนวนข้อมูลที่พิจารณา

p คือ จำนวนพารามิเตอร์ในแบบจำลอง

SSE คือ ผลบวกกำลังสองของความคลาดเคลื่อน

L คือ ฟังก์ชันภาวะน่าจะเป็นสูงสุด

2.1.3 การตรวจสอบความถูกต้องของแบบจำลอง (Model Validation)

ภายหลังการประมาณค่าพารามิเตอร์ต่างๆแล้วควรจะต้องมีการตรวจสอบแบบจำลองว่าสามารถพยากรณ์ข้อมูลได้ดีเพียงใดดูได้จากค่าเศษเหลือ และค่า p-value ของเศษเหลือว่ามีนัยสำคัญทางสถิติหรือไม่ และแบบจำลองที่ดีจะต้องมีจำนวนพารามิเตอร์น้อยที่สุดแต่ยังให้ค่าพยากรณ์ที่ดี

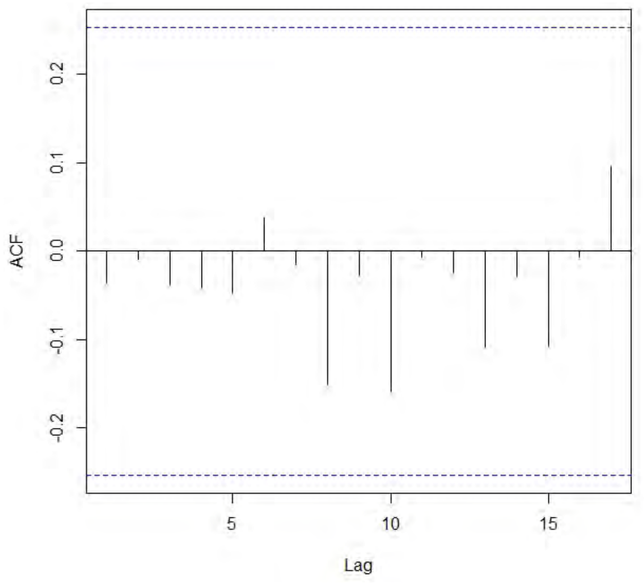
ค่าเศษเหลือ (Residuals)

ค่าเศษเหลือ คือ ค่าผลต่างระหว่างค่าที่ได้จากค่าสังเกต และ ค่าที่ได้จากการพยากรณ์ โดยแบบจำลองที่สร้างขึ้นจะต้องเป็นแบบสุ่ม (Random) ซึ่งค่าเศษเหลือควรมีการแจกแจงแบบปกติโดยทดสอบดังนี้

1. การทดสอบความสัมพันธ์ในตัวเองของค่าเศษเหลือ (Residuals Test)

พิจารณาจากกราฟค่าสหสัมพันธ์ในตัวเองของเศษเหลือ (Residuals Autocorrelation plot) แบบจำลองที่เหมาะสมควรมีค่าสหสัมพันธ์ในตัวเองของเศษเหลือน้อย เพื่อแสดงให้เห็นว่าค่าสหสัมพันธ์ในตัวเองของเศษเหลือไม่มีความสัมพันธ์กัน

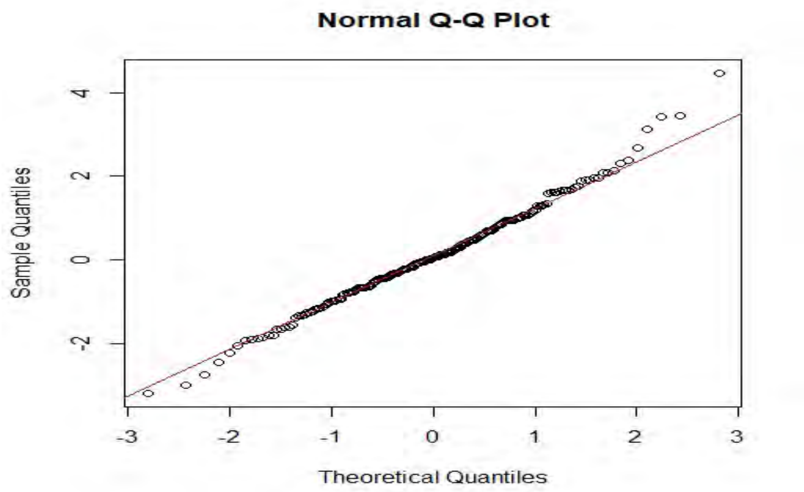
ตัวอย่าง



ภาพที่ ง แสดงกราฟค่าสหสัมพันธ์ในตัวเองของเศษเหลือ

2. การทดสอบการแจกแจงแบบปกติด้วยวิธี Quantile - Quantile Plot หรือ QQ plots

การทดสอบการแจกแจงแบบปกติด้วยวิธี Quantile - Quantile Plot ลักษณะการเกิดจุดจะต้องไม่กระจุกเป็นกลุ่มๆและอยู่ในแนวเส้นทแยงมุมจะเรียกว่าเป็นการกระจายแบบปกติ

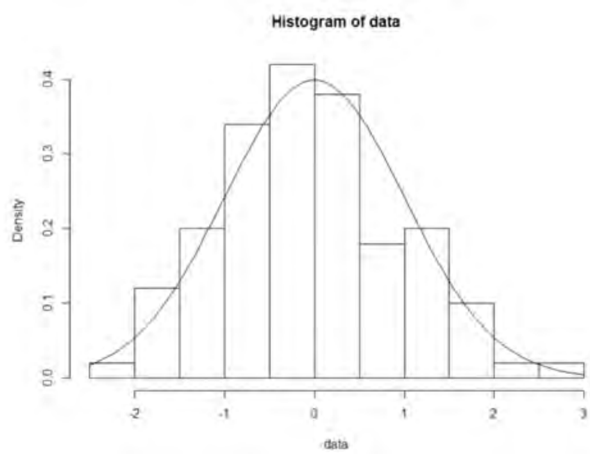


ภาพที่ จ กราฟการกระจายแบบปกติด้วยวิธี Quantile - Quantile Plot

3. การทดสอบการแจกแจงแบบปกติด้วยแผนภาพฮิสโตแกรม

พิจารณาจากรูปร่างของการกระจายของข้อมูลที่สนใจ ข้อมูลที่มีการแจกแจงแบบปกติควรมีลักษณะการกระจายของข้อมูลสมมาตรรอบค่าเฉลี่ย หรือมีรูปร่างคล้ายคลึงกับรูปทรงระฆังคว่ำ

ตัวอย่าง



ภาพที่ ฉ แสดงแผนภาพฮิตโตแกรมของการแจกแจงปกติ

4. การทดสอบการแจกแจงแบบปกติของข้อมูลด้วยการคำนวณตัวสถิติทดสอบ Kolmogorov-Smirnov

การทดสอบการแจกแจงแบบปกติของข้อมูลด้วยการคำนวณตัวสถิติทดสอบ Kolmogorov-Smirnov เป็นการทดสอบความแตกต่างระหว่างความถี่สะถมของตัวอย่างกับความถี่สะสมของการแจกแจง โดยกำหนดให้ $G(x)$ แทนฟังก์ชันของกลุ่มตัวอย่างที่ไม่ทราบการแจกแจงและ $F(x)$ แทนฟังก์ชันการแจกแจงที่ต้องการทดสอบ เราสามารถเขียนสมมติฐานของการทดสอบได้ดังนี้

สมมติฐานการทดสอบ

$$H_0: G \equiv F$$

$$H_1: G \not\equiv F$$

เราจะทดสอบสมมติฐานด้วยวิธี Kolmogorov-Smirnov คำนวณสถิติทดสอบจากความแตกต่างของค่าสัมบรูณ์ระหว่างการแจกแจงตัวอย่างกับการแจกแจกปกติ

$$K = max_x|G(x) - F(x)|$$

เกณฑ์การตัดสินใจในการทดสอบสมมติฐาน คือ ปฏิเสธสมมติฐาน H_0 เมื่อ ค่า p-value ของค่าสถิติทดสอบมีค่าน้อยกว่าค่าวิกฤต ณ ระดับนัยสำคัญ α

2.2 อนุกรมเวลาเชิงการนับ

บทนิยาม 12. อนุกรมเวลาเชิงการนับ (Count Time Series)

อนุกรมเวลาเชิงการนับ คือ ข้อมูลจำนวนของเหตุการณ์ต่อเวลาในช่วงเวลาที่ถูกสังเกตโดยอาจเก็บข้อมูลเป็นรายวัน รายเดือน หรือรายปีขึ้นอยู่กับวัตถุประสงค์ในการนำไปใช้ อนุกรมเวลาเชิงการนับมักพบเห็นได้ในสถานการณ์ต่างๆ เช่น จำนวนการเอาประกันภัยรายเดือนของบริษัทประกันภัย และจำนวนผู้ป่วยรายสัปดาห์ที่เข้ารับการรักษาในโรงพยาบาล

บทนิยาม 13. การแจกแจงปัวซง (Poisson distribution)

เราจะเรียกตัวแปรสุ่ม X ว่า ตัวแปรสุ่มปัวซง (Poisson random variable) ที่มีพารามิเตอร์ $\lambda(\lambda > 0)$ ถ้า

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad \text{เมื่อ } k = 0,1,2,\dots$$

โดยเราจะใช้สัญลักษณ์ $X \sim \text{Poisson}(\lambda)$ แทน X เป็นตัวแปรสุ่มปัวซงที่มีพารามิเตอร์ λ และเรียกการแจกแจงของตัวแปรสุ่มปัวซงว่า การแจกแจงปัวซง (Poisson distribution)

คุณสมบัติของตัวแปรสุ่มปัวซง

1. $E(X) = \lambda$
2. $Var(X) = \lambda$

บทนิยาม 14. การแจกแจงทวินามลบ (negative Binomial distribution)

เราจะเรียกตัวแปรสุ่ม X ว่า ตัวแปรสุ่มทวินามลบ (negative binomial random variable) ที่มีพารามิเตอร์ p และ r ถ้าตัวแปรสุ่ม X มีค่าเป็นจำนวนครั้งที่ล้มเหลวในการทดลองซ้ำเพื่อให้ได้ความสำเร็จเป็นครั้งที่ r ในการทดลองทวินามลบ โดยที่โอกาสที่การทดลองย่อยแต่ละครั้งจะสำเร็จมีค่าเท่ากับ p

$$P(X = x) = \frac{\Gamma(x+r)}{\Gamma(r)\Gamma(x+1)} p^r (1-p)^x \qquad \text{เมื่อ } x = 0,1,\ldots$$

โดยเราจะใช้สัญลักษณ์ $X \sim \text{NegBin}(r, p)$ แทน X เป็นตัวแปรสุ่มทวินามลบที่มีพารามิเตอร์ p และ r และเรียกการแจกแจงของตัวแปรสุ่มทวินามลบว่า การแจกแจงทวินามลบ (negative binomial distribution)

คุณสมบัติของตัวแปรสุ่มทวินามลบ

1. $E(X) = \frac{r(1-p)}{p}$
2. $Var(X) = \frac{r(1-p)}{p^2}$

บทนิยาม 15. แบบจำลองเชิงเส้นนัยทั่วไป (Generalized Linear Model)

การวิเคราะห์แบบจำลองเชิงเส้นนัยทั่วไป เป็นการหาความสัมพันธ์ระหว่างตัวแปรอิสระ $(X_1, \ldots, X_k)$ กับตัวแปรตาม (Y) โดยมีสมการดังนี้

$$g(Y) = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \varepsilon$$

โดยที่ g คือ ฟังก์ชันการเชื่อมโยง

Y คือ ตัวแปรตาม

$X_1, \dots, X_k$ คือ ตัวแปรอิสระ หรือ ตัวแปรช่วย

$\beta_0, \dots, \beta_k$ คือ สัมประสิทธิการถดถอย

ε คือ ค่าคาดเคลื่อนจากตัวแบบ

2.2.1 แบบจำลองอนุกรมเวลาเชิงการนับ

บทนิยาม 16. แบบจำลองเชิงเส้นลอการิทึม (Log-linear Model)

แบบจำลองเชิงเส้นลอการิทึม (Log-linear Model) เป็นแบบจำลองที่ใช้อธิบายอนุกรมเวลาเชิงการนับ โดยอาศัยข้อมูลในอดีตและค่าเฉลี่ยแบบมีเงื่อนไขของอนุกรมเวลาในการพยากรณ์ข้อมูลในอนาคต ซึ่งเขียนนิยามได้ดังนี้

ให้ $\{Y_t : t \in \mathbb{N}\}$ เป็นลำดับของข้อมูลอนุกรมเวลาที่สนใจ

$\{X_t : t \in \mathbb{N}\}$ เป็นอนุกรมเวลาของตัวแปรช่วย

F_t เป็นข้อมูลในอดีตของ $\{Y_t, \lambda_t, X_{t+1} : t \in \mathbb{N}\}$

$E(Y_t | F_{t-1}) = \lambda_t$ เป็นค่าเฉลี่ยแบบมีเงื่อนไขของอนุกรมเวลา

ซึ่งแบบจำลองเชิงเส้นลอการิทึม คือ

$$\log(\lambda_t) = \beta_0 + \sum_{k=1}^{p} \beta_k \log(Y_{t-i_k} + 1) + \sum_{l=1}^{q} \alpha_l \log(\lambda_{t-j_l}) + \eta^T X_t$$

เมื่อ $\eta = (\eta_1, \dots, \eta_r)^T$ เป็นพารามิเตอร์เวกเตอร์

$P = \{i_1, i_2, \dots, i_p\}$ เมื่อ $p \in \mathbb{N}_0$

$Q = \{j_1, j_2, \dots, j_q\}$ เมื่อ $q \in \mathbb{N}_0$

$\beta_0, \beta_1, \dots, \beta_p$ คือ สัมประสิทธิการถดถอยของ Y_t

$\alpha_1, \dots, \alpha_q$ คือ สัมประสิทธิการถดถอย λ_t

โดยที่

$\sum_{k=1}^{p} \beta_k \log(Y_{t-i_k} + 1)$ คือ ส่วนของข้อมูลในอดีตของตัวแปรที่สนใจ

$\sum_{l=1}^{q} \alpha_l \log(\lambda_{t-j_l})$ คือ ส่วนของค่าเฉลี่ยแบบมีเงื่อนไข

$\eta^T X_t$ คือ ส่วนของข้อมูลอนุกรมเวลาของตัวแปรช่วย

บทนิยาม 17. แบบจำลองสำหรับการแจกแจงปัวซง

แบบจำลองสำหรับการแจงแจงปัวซง เป็นการประยุกต์ใช้แบบจำลองเชิงเส้นลอการิทึมสำหรับอนุกรมเวลาเชิงการนับที่มีการแจงแจงปัวซง ซึ่งสามารถเขียนแบบจำลองสำหรับการแจกแจงปัวซงได้ดังนี้

$$\log(\lambda_t) = \beta_0 + \sum_{k=1}^{p} \beta_k \log(Y_{t-i_k} + 1) + \sum_{l=1}^{q} \alpha_l \log(\lambda_{t-j_l}) + \eta^T X_t$$

โดยที่ $Y_t|F_{t-1} \sim Poisson(\lambda_t)$

หมายความว่า $\mathrm{P}(Y_t = y|F_{t-1}) = \frac{\lambda_t^y \exp(-\lambda_t)}{y!},\ y = 0,1,\ldots$

ซึ่ง $VAR(Y_t|F_{t-1}) = E(Y_t|F_{t-1}) = \lambda_t$

บทนิยาม 18. แบบจำลองสำหรับการแจกแจงทวินามลบ

แบบจำลองสำหรับการแจงแจงทวินามลบ เป็นการประยุกต์ใช้แบบจำลองเชิงเส้นลอการิทึมสำหรับอนุกรมเวลาเชิงการนับที่มีการแจงแจงทวินามลบ ซึ่งสามารถเขียนแบบจำลองสำหรับการแจกแจงทวินามลบได้ดังนี้

$$\log(\lambda_t) = \beta_0 + \sum_{k=1}^{p} \beta_k \log(Y_{t-i_k} + 1) + \sum_{l=1}^{q} \alpha_l \log(\lambda_{t-j_l}) + \eta^T X_t$$

โดยที่ $Y_t|F_{t-1} \sim NegBin(\lambda_t, \phi), \phi \in (0, \infty)$

หมายความว่า $P(Y_t = y|F_{t-1}) = \frac{\Gamma(\phi+y)}{\Gamma(y+1)\Gamma(\phi)} (\frac{\phi}{\phi+\lambda_t})^{\phi} (\frac{\lambda_t}{\phi+\lambda_t})^{y},\ y = 0,1,\ldots$

ซึ่ง $E(Y_t|F_{t-1}) = \lambda_t$ และ $VAR(Y_t|F_{t-1}) = \lambda_t + \frac{\lambda_t^2}{\phi}$

เมื่อ $\phi \to \infty$ จะได้ว่า $\mathrm{VAR}(Y_t|F_{t-1}) = \lambda_t$

2.2.2 การวิเคราะห์แบบจำลองที่เหมาะสม

การวิเคราะห์แบบจำลองที่เหมาะสม เป็นการหาแบบจำลองที่เหมาะสมที่สุดสำหรับข้อมูลโดยการพิจารณาค่าคาดเคลื่อนกำลังสอง ซึ่งนิยามได้ดังนี้

กำหนดให้ $v_t = \sqrt{\mathrm{VAR}(Y_t|F_{t-1})}, P_t(y) = \mathrm{P}(Y_t \le y|F_{t-1})$

$p_t(y) = \mathrm{P}(Y_t = y|F_{t-1})$ และ $\|p_t\|^2 = \sum_{y=0}^{\infty} p_t^2(y)$

ซึ่งค่าคาดเคลื่อนกำลังสองเท่ากับ $\sum_{t=1}^{T}(y_t - \hat{\lambda}_t)^2$

สำหรับแบบจำลองที่เหมาะสมที่สุดควรมีค่าคาดเคลื่อนกำลังสองน้อยที่สุด

บทที่ 3

ผลการทำงาน

โครงงานนี้ เรามีความสนใจในการศึกษาอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมสำหรับข้อมูลจำนวนการเอาประกันภัยและจำนวนเงินเอาประกันภัยของบริษัทประกันภัยในประเทศไทยในกลุ่มข้อมูลของประกันชีวิต 3 ประเภท ได้แก่ ประกันชีวิตแบบตลอดชีพ ประกันชีวิตแบบสะสมทรัพย์และประกันชีวิตแบบชั่วระยะเวลา ในช่วงปี พ.ศ. 2552-2556 โดยใช้แบบจำลองและชุดคำสั่งในโปรแกรม R

การประกันชีวิตเป็นวิธีการที่บุคคลกลุ่มหนึ่งร่วมกันเฉลี่ยภัยอันเนื่องจากการตาย การสูญเสียอวัยวะ ทุพพลภาพ และการสูญเสียรายได้ในยามชรา โดยที่เมื่อบุคคลใดต้องประสบกับภัยเหล่านั้น ก็ได้รับเงินเฉลี่ยช่วยเหลือเพื่อบรรเทาความเดือดร้อนแก่ตนเองและครอบครัว โดยบริษัทประกันชีวิตจะทำหน้าที่เป็นแกนกลางในการนำเงินก้อนดังกล่าวไปจ่ายให้แก่ผู้ได้รับภัย

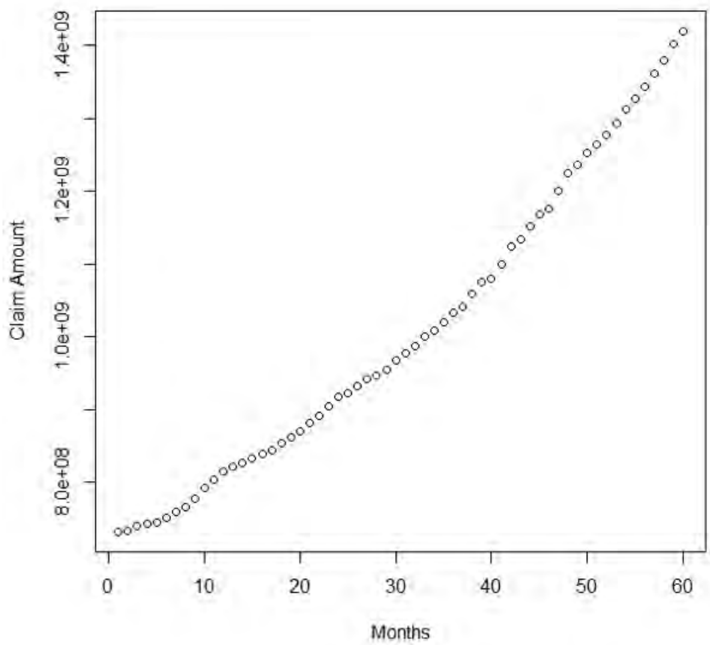
3.1 จำนวนเงินเอาประกันภัยของบริษัทประกันภัยในประเทศไทย

1. ประกันชีวิตแบบตลอดชีพ

ประกันชีวิตแบบตลอดชีพ เป็นการประกันชีวิตที่ให้ความคุ้มครองตลอดชีพ ถ้าผู้เอาประกันภัยเสียชีวิตเมื่อใดในขณะที่กรมธรรม์มีผลบังคับ บริษัทประกันชีวิตจะจ่ายเงินเอาประกันภัย ให้แก่ผู้รับประโยชน์ วัตถุประสงค์เบื้องต้นของการประกันภัยแบบนี้เพื่อจัดหาเงินทุนสำหรับจุนเจือ บุคคลที่อยู่ในความอุปการะเมื่อผู้เอาประกันภัยเสียชีวิต หรือเพื่อเป็นเงินทุนสำหรับการเจ็บป่วยครั้งสุดท้ายและค่าทำศพ ทั้งนี้เพื่อไม่ให้ตกเป็นภาระของคนอื่น โดยเราศึกษาข้อมูลจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 เนื่องจากข้อมูลสูญหาย 4 เดือน ได้แก่ เดือนมกราคม ปี พ.ศ. 2553 เดือนกุมภาพันธ์ ปี พ.ศ. 2555 เดือนพฤศจิกายน ปี พ.ศ.2555 และเดือนกรกฎาคม ปี พ.ศ. 2556 เราจึงแก้ไขข้อมูลโดยให้ข้อมูลที่สูญหายแทนด้วยค่าเฉลี่ยของข้อมูลก่อนหน้า 1 เดือนและข้อมูลหลังจากนั้น 1 เดือน

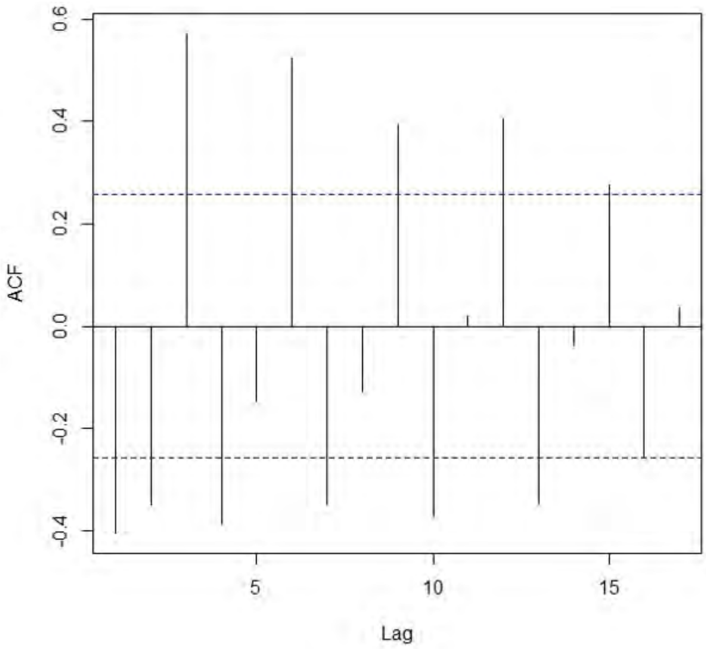
เดือน ปี พ.ศ.	จำนวนเงินเอาประกันภัย (บาท)
ม.ค. 2552	732,836,028
ก.พ. 2552	734,091,387
มี.ค. 2552	741,099,476
เม.ย. 2552	744,157,319
พ.ค. 2552	744,906,847
มิ.ย. 2552	751,603,329
⋮	⋮
ก.ค. 2556	1,328,069,060
ส.ค. 2556	1,344,178,246
ก.ย. 2556	1,361,250,898
ต.ค. 2556	1,379,325,896
พ.ย. 2556	1,401,911,576
ธ.ค. 2556	1,419,781,352

ตารางที่ 1.1 แสดงจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556

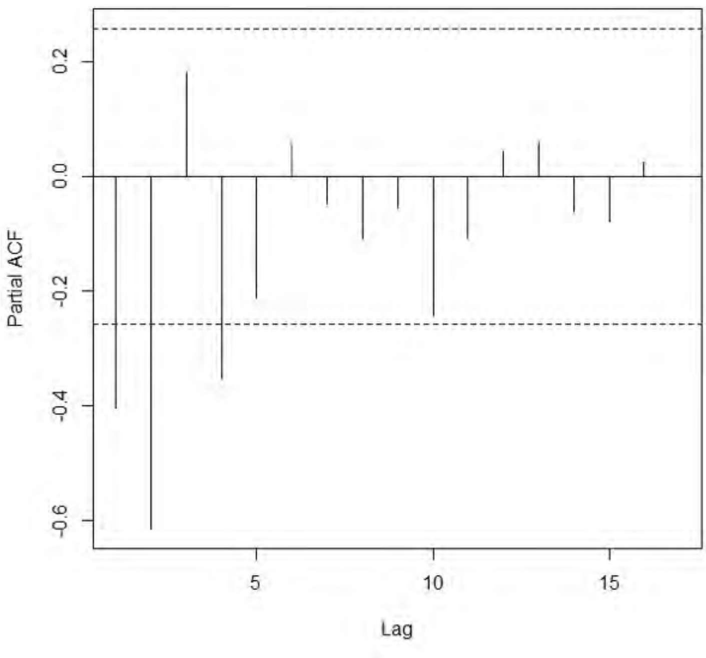


ภาพที่ 1.1 แสดงจำนวนเงินเอาประกันภัยของประกันชีวิตแบบตลอดชีพระหว่างปี พ.ศ. 2552-2556

ต่อไปเราจะนำข้อมูลจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาทดสอบความนิ่ง (Stationary) และแปลงข้อมูลให้มีความนิ่ง โดยใช้วิธีการหาผลต่าง (differencing) จะได้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งมีความนิ่ง เราจึงใช้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งในการคาดเดาลำดับของ Autoregressive (AR) และ Moving Average (MA) โดยพิจารณาค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) แสดงในภาพที่ 1.2 และภาพที่ 1.3



ภาพที่ 1.2 แสดงกราฟ ACF ของจำนวนเงินเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2



ภาพที่ 1.3 แสดงกราฟ PACF ของจำนวนเงินเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2

จากภาพที่ 1.2 และภาพที่ 1.3 พบว่าอนุกรมเวลาของผลต่างอันดับ 2 มีค่าสหสัมพันธ์ในตัวเอง (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,3,4,6,7,9,10,12 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,4 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจะนำข้อมูลจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาพิจารณา ARIMA(p,d,q) เมื่อ p∈{1,2,...,4},q=∈{1,2,...,12}, และ d=2 ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC ต่ำที่สุด ผลการศึกษาแสดงดังตารางที่ 1.2 ต่อไปนี้

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,2,1)	1945.389	ARIMA(3,2,1)	1934.419
ARIMA(1,2,2)	1940.531	ARIMA(3,2,2)	1933.567
ARIMA(1,2,3)	1940.077	ARIMA(3,2,3)	1931.381
ARIMA(1,2,4)	1940.391	ARIMA(3,2,4)	1932.406
ARIMA(1,2,5)	1940.083	ARIMA(3,2,5)	1934.407
ARIMA(1,2,6)	1934.206	ARIMA(3,2,6)	1936.406
ARIMA(1,2,7)	1936.165	ARIMA(3,2,7)	1937.433
ARIMA(1,2,8)	1935.164	ARIMA(3,2,8)	1936.045
ARIMA(1,2,9)	1936.615	ARIMA(3,2,9)	1936.360
ARIMA(1,2,10)	1936.586	ARIMA(3,2,10)	1941.870
ARIMA(1,2,11)	1939.826	ARIMA(3,2,11)	1941.732
ARIMA(1,2,12)	1938.759	ARIMA(3,2,12)	1938.519
ARIMA(2,2,1)	1932.430	**ARIMA(4,2,1)**	**1930.366**
ARIMA(2,2,2)	1934.228	ARIMA(4,2,2)	1932.346
ARIMA(2,2,3)	1932.104	ARIMA(4,2,3)	1932.466
ARIMA(2,2,4)	1930.407	ARIMA(4,2,4)	1934.449
ARIMA(2,2,5)	1932.406	ARIMA(4,2,5)	1932.810
ARIMA(2,2,6)	1936.172	ARIMA(4,2,6)	1938.391
ARIMA(2,2,7)	1938.636	ARIMA(4,2,7)	1939.303
ARIMA(2,2,8)	1939.538	ARIMA(4,2,8)	1937.737

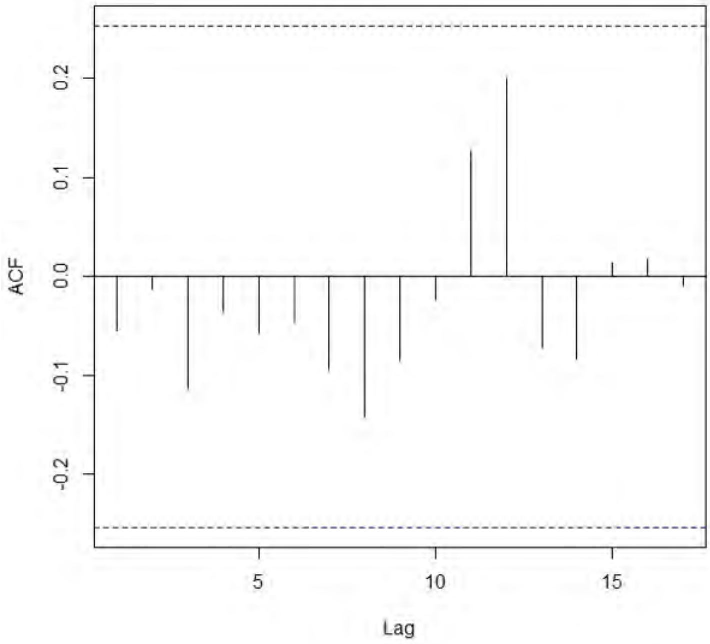
ตารางที่ 1.2 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนเงินเอาประกันภัย

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(2,2,9)	1941.548	ARIMA(4,2,9)	1939.992
ARIMA(2,2,10)	1938.513	ARIMA(4,2,10)	1941.766
ARIMA(2,2,11)	1937.713	ARIMA(4,2,11)	1940.095
ARIMA(2,2,12)	1938.187	ARIMA(4,2,12)	1940.140

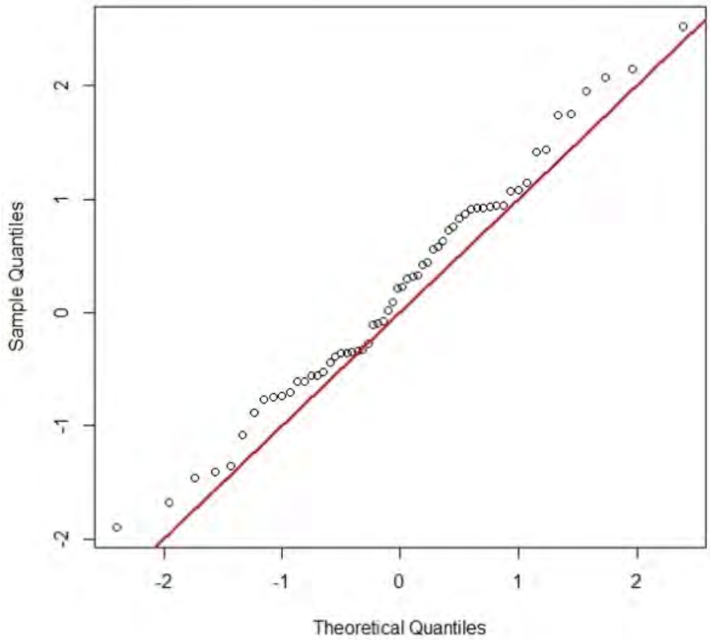
ตารางที่ 1.2 (ต่อ) แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนเงินเอาประกันภัย

จากตารางที่ 1.2 เมื่อพิจารณาจากค่า AIC เนื่องจาก ARIMA(4,2,1) ให้ค่า AIC น้อยที่สุด คือ 1930.366 ดังนั้นแบบจำลองอนุกรมเวลาของตัวแบบสำหรับการจำลองที่เหมาะสมสำหรับข้อมูลจำนวนเงินเอาประกันภัยของประกันชีวิตแบบตลอดชีพ คือ ARIMA(4,2,1)

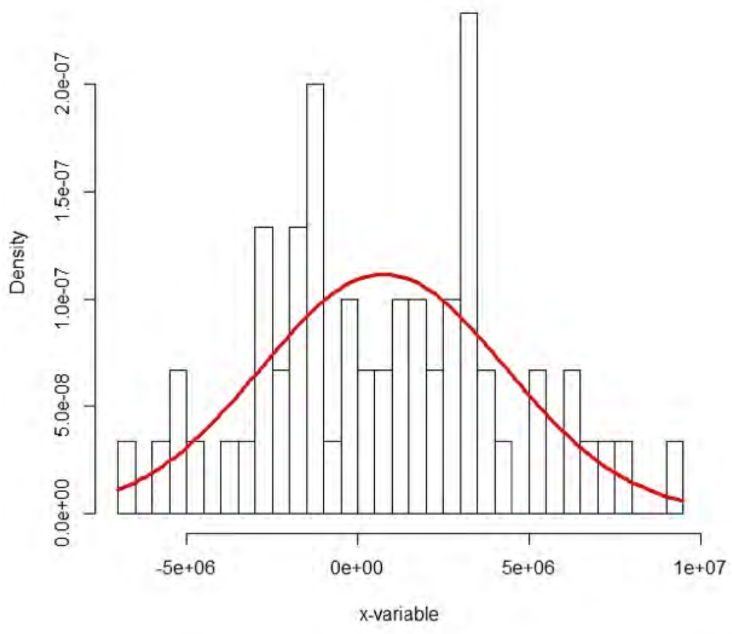
ต่อไปจะนำค่าเศษเหลือ (Residuals) ซึ่งคือค่าผลต่างระหว่างจำนวนเงินเอาประกันภัยจริงและจำนวนเงินเอาประกันภัยพยากรณ์ มาทดสอบความเป็นอิสระต่อกันของค่าเศษเหลือด้วยกราฟค่าสหสัมพันธ์ในตัวเองของเศษเหลือ (Autocorrelation residuals plot) ทดสอบการแจกแจงแบบปกติจากกราฟโดยวิธี Quantile-Quantile Plot หรือ QQ plots และทดสอบการแจกแจงปกติด้วยกราฟฮิสโตแกรม แสดงในภาพที่ 1.4 ภาพที่ 1.5 และภาพที่ 1.6 และทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov แสดงในตารางที่ 1.3



ภาพที่ 1.4 แสดงค่าสหสัมพันธ์ในตัวเองของเศษเหลือ



ภาพที่ 1.5 แสดงการทดสอบ Quantile-Quantile Plot ของค่าเศษเหลือ



ภาพที่ 1.6 แสดงทดสอบการแจกแจงปกติด้วยกราฟฮิสโตแกรม

แบบจำลอง	ค่า p-value ของ Kolmogorov-Smirnov
ARIMA(4,2,1)	0.7115

ตารางที่ 1.3 แสดงค่า p-value ของ Kolmogorov-Smirnov

จากภาพที่ 1.4 แบบจำลองแสดงให้เห็นว่าค่าสหสัมพันธ์ในตัวเองของเศษเหลือไม่มีความสัมพันธ์กัน สำหรับการทดสอบการแจกแจงปกติของเศษเหลือของแบบจำลองจะได้ว่า จากภาพที่ 1.5 แบบจำลองแสดงการกระจายของเศษเหลือที่คล้ายจะเป็นการแจกแจงปกติ จากภาพที่ 1.6 กราฟฮิสโตแกรมของแบบจำลองแสดงให้เห็นว่าเศษเหลือของแบบจำลองมีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติ แต่มีบางจุดที่ค่าความถี่มีค่าโดดจากค่าอื่น และเมื่อทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov ในตารางที่ 1.3 พบว่าค่า p-value ของ Kolmogorov-Smirnov เท่ากับ 0.7115 ซึ่งมากกว่า 0.05 แสดงว่าเศษเหลือของแบบจำลองมีการแจกแจงแบบปกติ เมื่อพิจารณาภาพที่ 1.5 – 1.6 และ ตารางที่ 1.3 จึงสามารถสรุปได้ว่าแบบจำลองเศษเหลือของ ARIMA(4,2,1) มีการแจกแจงใกล้เคียงกับการแจกแจงปกติ

ต่อไปจะแสดงการพยากรณ์จำนวนเงินเอาประกันภัยโดยใช้แบบจำลอง ARIMA(4,2,1)

พารามิเตอร์	ARIMA(4,2,1)
Ar1	-0.0864
Ar2	-0.4457
Ar3	0.2016
Ar4	-0.4132
Ma1	-0.4132

ตารางที่ 1.4 แสดงพารามิเตอร์ของจำนวนเงินเอาประกันภัย

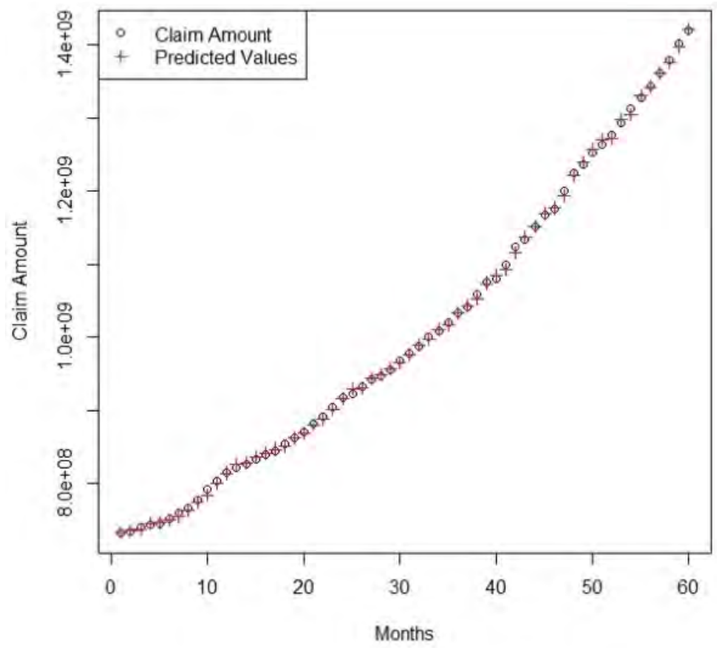
จากตารางที่ 1.4 จะได้ว่าแบบจำลอง ARIMA(4,2,1) เขียนค่าประมาณพารามิเตอร์จากการทำผลต่างอันดับ 2 ได้สมการต่อไปนี้

$$\Delta^2 y_t = -0.0864\Delta^2 y_{t-1} - 0.4457\Delta^2 y_{t-2} + 0.2016\Delta^2 y_{t-3} - 0.4132\Delta^2 y_{t-4}$$
$$+0.4132\varepsilon_{t-1} + \varepsilon_t$$

จัดรูปสมการได้ดังนี้

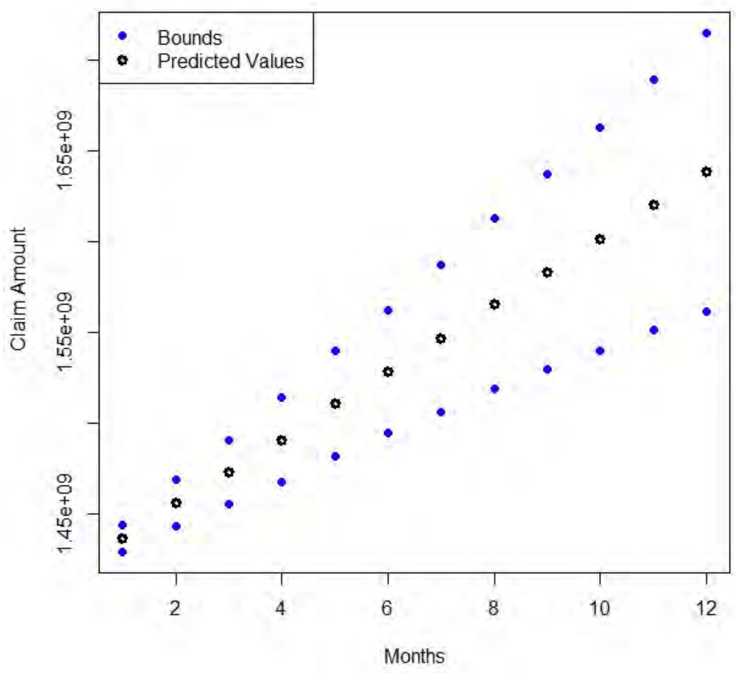
$$y_t = 1.9136 y_{t-1} - 1.2729 y_{t-2} + 1.0066 y_{t-3} - 1.2621 y_{t-4} + 1.028 y_{t-5}$$
$$-0.4132 y_{t-6} + 0.4132\varepsilon_{t-1} + \varepsilon_t$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนเงินเอาประกันภัยจริงกับจำนวนเงินเอาประกันภัยพยากรณ์ตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 แสดงจำนวนเงินเอาประกันภัยพยากรณ์อีก 12 เดือนข้างหน้า และแสดงจำนวนเงินเอาประกันภัยจริงตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ต่อด้วยจำนวนเงินเอาประกันภัยพยากรณ์อีก 12 เดือนข้างหน้า ดังภาพที่ 1.7 ภาพที่ 1.8 และภาพที่ 1.9

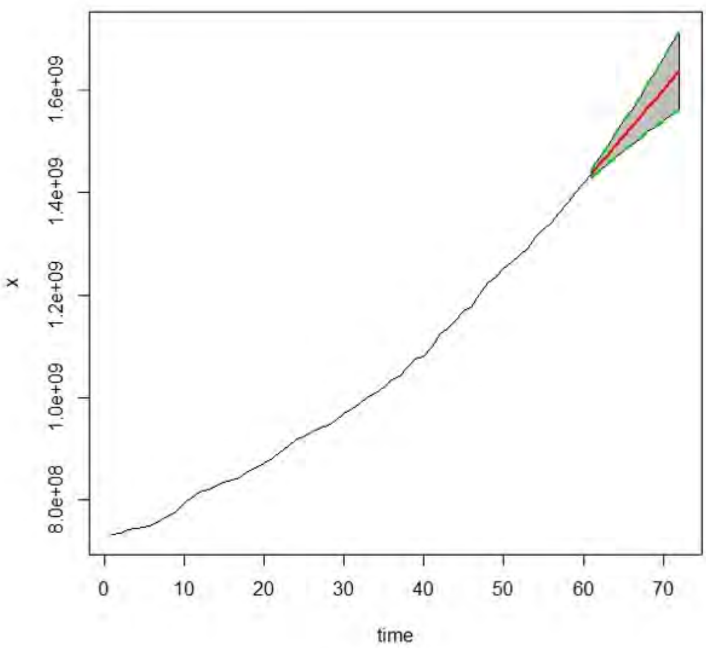


ภาพที่ 1.7 แสดงการเปรียบเทียบจำนวนเงินเอาประกันภัยจริงกับ

จำนวนเงินเอาประกันภัยจากแบบจำลอง



ภาพที่ 1.8 แสดงจำนวนเงินเอาประกันภัยพยากรณ์



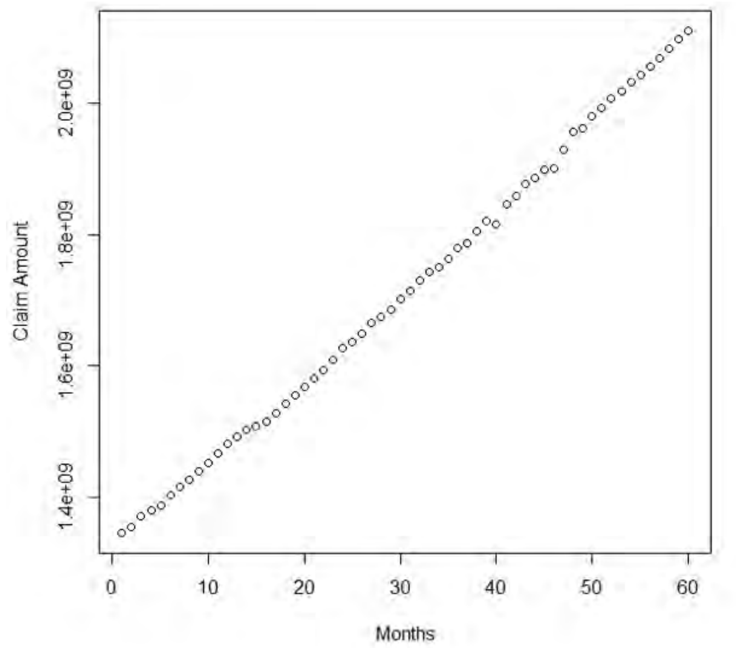
ภาพที่ 1.9 แสดงจำนวนเงินเอาประกันภัยจริงตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ต่อด้วยจำนวนเงินเอาประกันภัยพยากรณ์อีก 12 เดือนข้างหน้า

2. ประกันชีวิตแบบสะสมทรัพย์

ประกันชีวิตแบบสะสมทรัพย์ เป็นการประกันชีวิตที่บริษัทจะจ่ายเงินเอาประกันภัยให้แก่ผู้เอาประกันภัย เมื่อมีชีวิตอยู่ครบกำหนดสัญญา หรือจ่ายเงินเอาประกันภัย ให้แก่ผู้รับประโยชน์เมื่อผู้เอาประกันภัยเสียชีวิตลงภายในระยะเวลาประกันภัย การประกันชีวิตแบบสะสมทรัพย์เป็นส่วนผสมของการคุ้มครองชีวิตและการออมทรัพย์ ส่วนของการออมทรัพย์ คือส่วนที่ผู้เอาประกันภัยได้รับคืนเมื่อสัญญาครบกำหนด โดยเราศึกษาข้อมูลจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 เนื่องจากข้อมูลสูญหาย 4 เดือน ได้แก่ เดือนมกราคม ปี พ.ศ. 2553 เดือนกุมภาพันธ์ ปี พ.ศ. 2555 เดือนพฤศจิกายน ปี พ.ศ.2555 และเดือนกรกฎาคม ปี พ.ศ. 2556 เราจึงแก้ไขข้อมูลโดยให้ข้อมูลที่สูญหายแทนด้วยค่าเฉลี่ยของข้อมูลก่อนหน้า 1 เดือนและข้อมูลหลังจากนั้น 1 เดือน

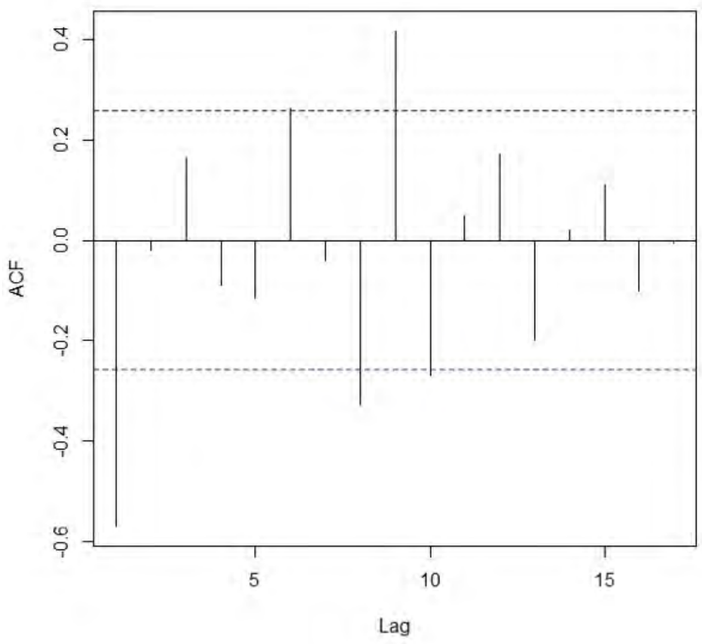
เดือน ปีพ.ศ.	จำนวนเงินเอาประกันภัย (บาท)
ม.ค. 2552	1,345,769,545
ก.พ. 2552	1,353,570,187
มี.ค. 2552	1,369,690,473
เม.ย. 2552	1,379,112,864
พ.ค. 2552	1,387,348,721
มิ.ย. 2552	1,403,224,547
⋮	⋮
ก.ค. 2556	2,044,316,113
ส.ค. 2556	2,056,029,450
ก.ย. 2556	2,069,668,440
ต.ค. 2556	2,082,783,758
พ.ย. 2556	2,098,706,646
ธ.ค. 2556	2,110,349,651

ตารางที่ 2.1 แสดงจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556

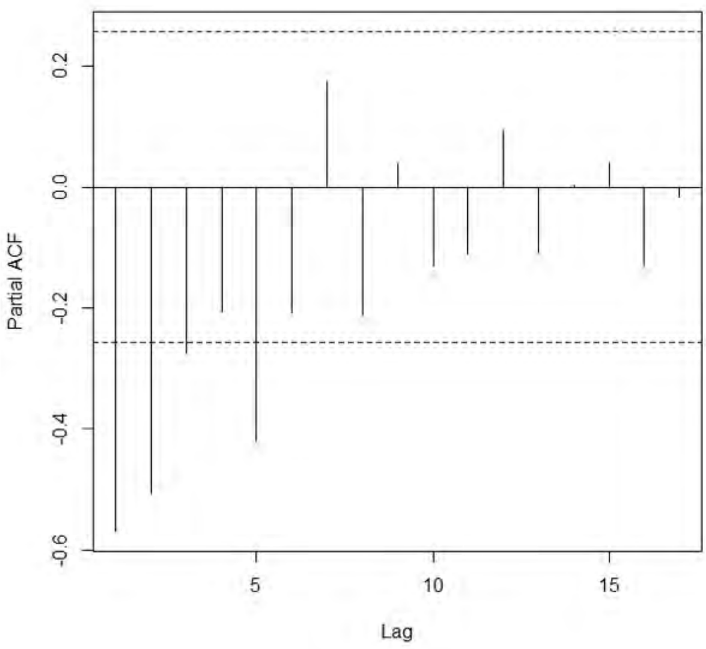


ภาพที่ 2.1 แสดงจำนวนเงินเอาประกันภัยของประกันชีวิตแบบสะสมทรัพย์ระหว่างปี พ.ศ. 2552-2556

ต่อไปเราจะนำข้อมูลจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาทดสอบความนิ่ง (Stationary) และแปลงข้อมูลให้มีความนิ่ง โดยใช้วิธีการหาผลต่าง (differencing) จะได้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งมีความนิ่ง เราจึงใช้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งในการคาดเดาลำดับของ Autoregressive (AR) และ Moving Average (MA) โดยพิจารณาค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) แสดงในภาพที่ 2.2 และภาพที่ 2.3



ภาพที่ 2.2 แสดงกราฟ ACF ของจำนวนเงินเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2



ภาพที่ 2.3 แสดงกราฟ PACF ของจำนวนเงินเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2

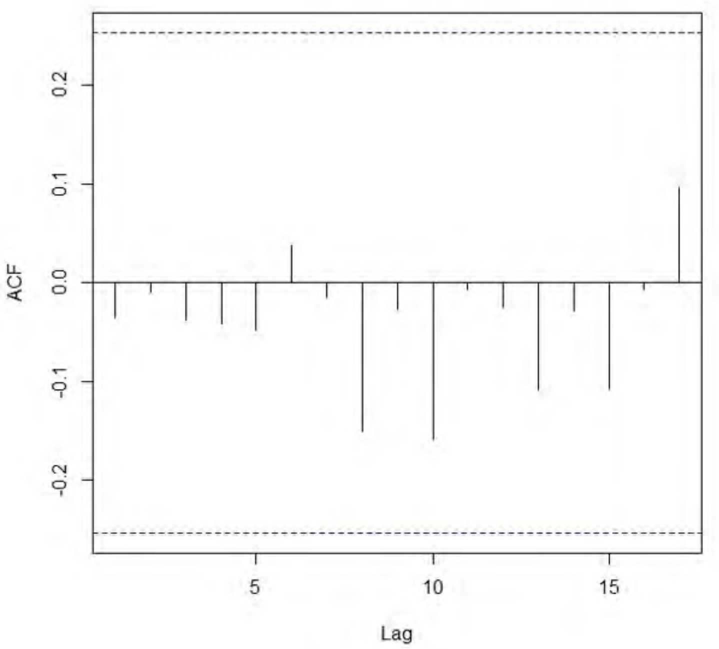
จากภาพที่ 2.2 และภาพที่ 2.3 พบว่าอนุกรมเวลาของผลต่างอันดับ 2 มีค่าสหสัมพันธ์ในตัวเอง (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,8,9 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,5 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจะนำข้อมูลจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาพิจารณา ARIMA(p,d,q) เมื่อ $p\in\{1,2,\ldots,5\}, q\in\{1,2,\ldots,9\}$ และ $d=2$ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC ต่ำที่สุด ผลการศึกษาแสดงดังตารางที่ 2.2 ต่อไปนี้

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,2,1)	1971.556	ARIMA(3,2,1)	1972.382	ARIMA(5,2,1)	1974.178
ARIMA(1,2,2)	1978.815	ARIMA(3,2,2)	1974.192	ARIMA(5,2,2)	1975.557
ARIMA(1,2,3)	1971.802	ARIMA(3,2,3)	1975.746	ARIMA(5,2,3)	1973.521
ARIMA(1,2,4)	1973.751	ARIMA(3,2,4)	1972.378	ARIMA(5,2,4)	1975.089
ARIMA(1,2,5)	1975.733	ARIMA(3,2,5)	1972.937	ARIMA(5,2,5)	1975.795
ARIMA(1,2,6)	1977.689	ARIMA(3,2,6)	1970.468	ARIMA(5,2,6)	1972.519
ARIMA(1,2,7)	1978.916	ARIMA(3,2,7)	1970.613	ARIMA(5,2,7)	1974.507
ARIMA(1,2,8)	1976.522	ARIMA(3,2,8)	1970.622	ARIMA(5,2,8)	1976.889
ARIMA(1,2,9)	1968.463	ARIMA(3,2,9)	1967.791	ARIMA(5,2,9)	1971.816
ARIMA(2,2,1)	1970.397	ARIMA(4,2,1)	1973.598		
ARIMA(2,2,2)	1972.192	ARIMA(4,2,2)	1975.311		
ARIMA(2,2,3)	1973.755	ARIMA(4,2,3)	1971.030		
ARIMA(2,2,4)	1968.532	ARIMA(4,2,4)	1971.078		
ARIMA(2,2,5)	1970.939	ARIMA(4,2,5)	1974.225		
ARIMA(2,2,6)	1968.873	ARIMA(4,2,6)	1972.460		
ARIMA(2,2,7)	1968.661	ARIMA(4,2,7)	1969.770		
ARIMA(2,2,8)	1969.650	ARIMA(4,2,8)	1971.442		
ARIMA(2,2,9)	**1965.962**	ARIMA(4,2,9)	1969.409		

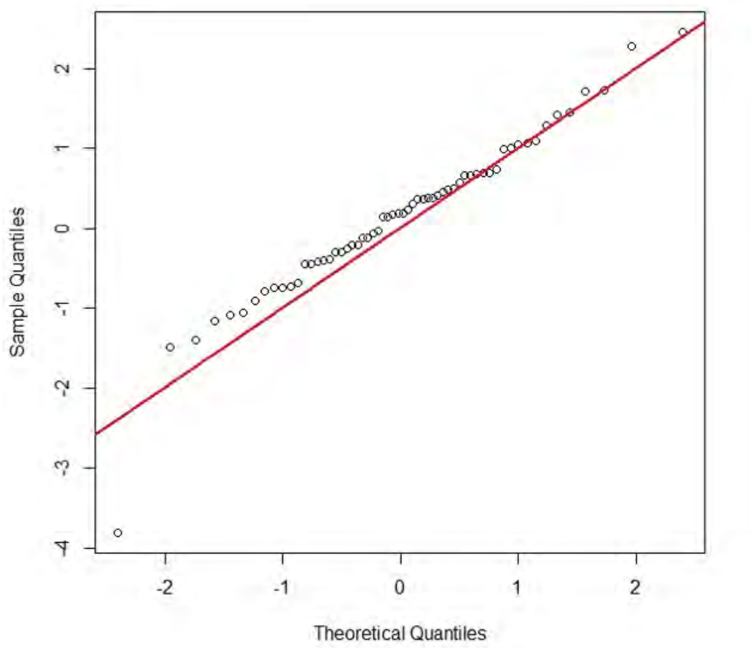
ตารางที่ 2.2 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนเงินเอาประกันภัย

จากตารางที่ 2.2 เมื่อพิจารณาจากค่า AIC เนื่องจาก ARIMA(2,2,9) ให้ค่า AIC น้อยที่สุด คือ 1965.962 ดังนั้นแบบจำลองอนุกรมเวลาของตัวแบบสำหรับการจำลองที่เหมาะสมสำหรับข้อมูลจำนวนเงินเอาประกันภัยของประกันชีวิตแบบสะสมทรัพย์ คือ ARIMA(2,2,9)

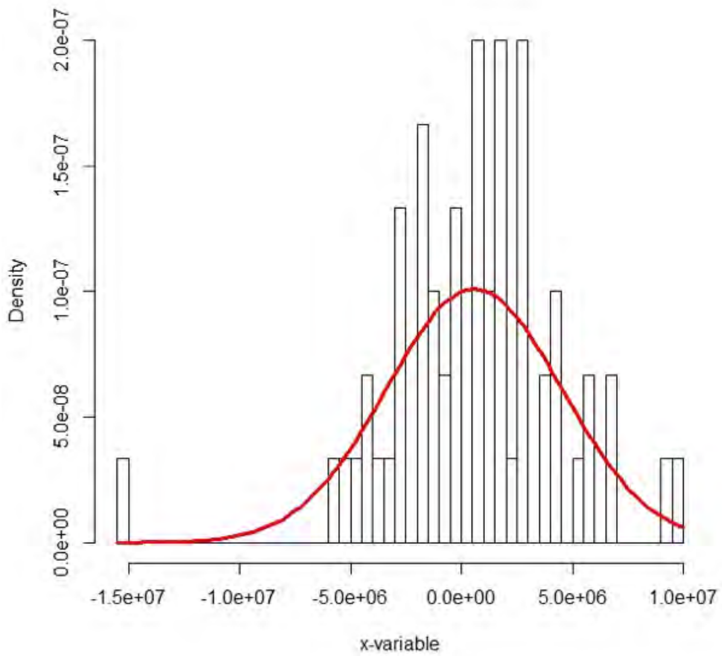
ต่อไปจะนำค่าเศษเหลือ (Residuals) ซึ่งคือค่าผลต่างระหว่างจำนวนเงินเอาประกันภัยจริงและจำนวนเงินเอาประกันภัยพยากรณ์ มาทดสอบความเป็นอิสระต่อกันของค่าเศษเหลือด้วยกราฟค่าสหสัมพันธ์ในตัวเองของเศษเหลือ (Autocorrelation residuals plot) ทดสอบการแจกแจงแบบปกติจากกราฟโดยวิธี Quantile-Quantile Plot หรือ QQ plots และทดสอบการแจกแจงปกติด้วยกราฟฮิสโตแกรม แสดงในภาพที่ 2.4 ภาพที่ 2.5 และภาพที่ 2.6 และทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov แสดงในตารางที่ 2.3



ภาพที่ 2.4 แสดงค่าสหสัมพันธ์ในตัวเองของเศษเหลือ



ภาพที่ 2.5 แสดงการทดสอบ Quantile-Quantile Plot ของค่าเศษเหลือ



ภาพที่ 2.6 แสดงทดสอบการแจกแจงปกติด้วยกราฟฮิสโตแกรม

แบบจำลอง	ค่า p-value ของ Kolmogorov-Smirnov
ARIMA(2,2,9)	0.8626

ตารางที่ 2.3 แสดงค่า p-value ของ Kolmogorov-Smirnov

จากภาพที่ 2.4 แบบจำลองแสดงให้เห็นว่าค่าสหสัมพันธ์ในตัวเองของเศษเหลือไม่มีความสัมพันธ์กัน สำหรับการทดสอบการแจกแจงปกติของเศษเหลือของแบบจำลองจะได้ว่า จากภาพที่ 2.5 แบบจำลองแสดงการกระจายของเศษเหลือที่คล้ายจะเป็นการแจกแจงปกติ จากภาพที่ 2.6 กราฟฮิสโตแกรมของแบบจำลองแสดงให้เห็นว่าเศษเหลือของแบบจำลองมีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติ แต่มีความถี่ของบางค่าสูงกว่าความถี่ของการแจกแจงแบบปกติ และมีความถี่ของบางค่าต่ำกว่าความถี่ของการแจกแจงแบบปกติ และเมื่อทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov ในตารางที่ 2.3 พบว่าค่า p-value ของ Kolmogorov-Smirnov เท่ากับ 0.8626 ซึ่งมากกว่า 0.05 แสดงว่าเศษเหลือของแบบจำลองมีการแจกแจงแบบปกติ เมื่อพิจารณาภาพที่ 2.5 – 2.6 และ ตารางที่ 2.3 จึงสามารถสรุปได้ว่าแบบจำลองเศษเหลือของ ARIMA(2,2,9) มีการแจกแจงใกล้เคียงกับการแจกแจงปกติ

ต่อไปจะแสดงการพยากรณ์จำนวนเงินเอาประกันภัยโดยใช้แบบจำลอง ARIMA(2,2,9)

พารามิเตอร์	ARIMA(2,2,9)
Ar1	-0.6723
Ar2	-0.4217
Ma1	-0.6090
Ma2	-0.2186
Ma3	-0.4862
Ma4	0.1885
Ma5	-0.0037
Ma6	0.4131
Ma7	-0.3196

ตารางที่ 2.4 แสดงพารามิเตอร์ของจำนวนเงินเอาประกันภัย

พารามิเตอร์	ARIMA(2,2,9)
Ma8	-0.3987
Ma9	0.5946

ตารางที่ 2.4 (ต่อ) แสดงพารามิเตอร์ของจำนวนเงินเอาประกันภัย

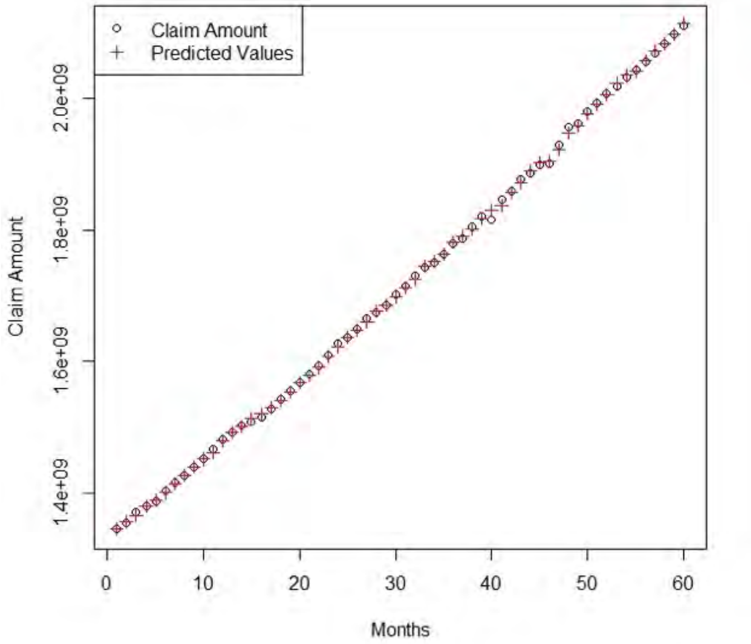
จากตารางที่ 2.4 จะได้ว่าแบบจำลอง ARIMA(2,2,9) เขียนค่าประมาณพารามิเตอร์จากการทำผลต่างอันดับ 2 ได้สมการต่อไปนี้

$$\begin{aligned}\Delta^2 y_t = &-0.6723\Delta^2 y_{t-1} - 0.4217\Delta^2 y_{t-2} + 0.6090\varepsilon_{t-1} + 0.2186\varepsilon_{t-2}\\ &+0.4862\varepsilon_{t-3} - 0.1885\varepsilon_{t-4} + 0.0037\varepsilon_{t-5} - 0.4131\varepsilon_{t-6}\\ &+0.3196\varepsilon_{t-7} + 0.3987\varepsilon_{t-8} - 0.5946\varepsilon_{t-9} + \varepsilon_t\end{aligned}$$

จัดรูปสมการได้ดังนี้

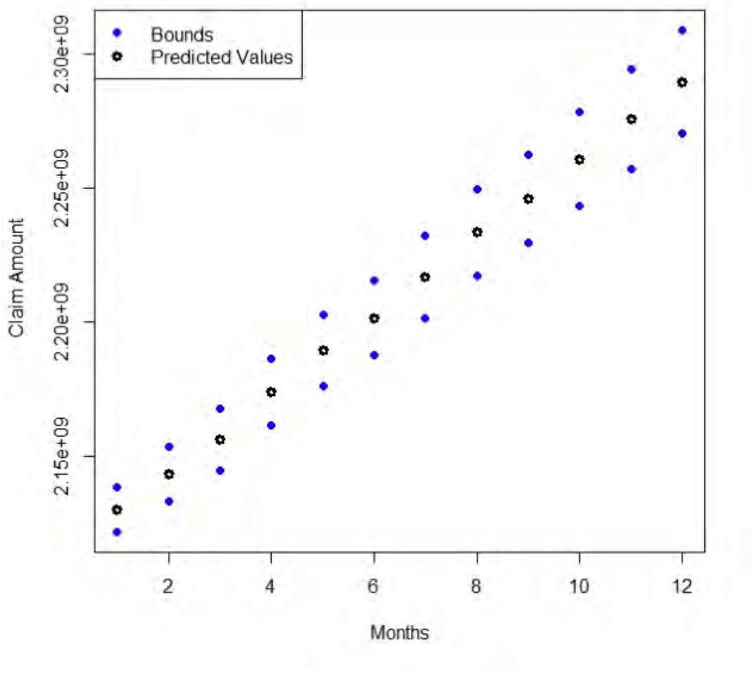
$$\begin{aligned}y_t = &\ 1.3277y_{t-1} - 0.0771y_{t-2} + 0.1711y_{t-3} - 0.4217y_{t-4} + 0.6090\varepsilon_{t-1}\\ &+0.2186\varepsilon_{t-2} + 0.4862\varepsilon_{t-3} - 0.1885\varepsilon_{t-4} + 0.0037\varepsilon_{t-5} - 0.4131\varepsilon_{t-6}\\ &+0.3196\varepsilon_{t-7} + 0.3987\varepsilon_{t-8} - 0.5946\varepsilon_{t-9} + \varepsilon_t\end{aligned}$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนเงินเอาประกันภัยจริงกับจำนวนเงินเอาประกันภัยพยากรณ์ตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 แสดงจำนวนเงินเอาประกันภัยพยากรณ์อีก 12 เดือนข้างหน้า และแสดงจำนวนเงินเอาประกันภัยจริงตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ต่อด้วยจำนวนเงินเอาประกันภัยพยากรณ์อีก 12 เดือนข้างหน้า ดังภาพที่ 2.7 ภาพที่ 2.8 และภาพที่ 2.9

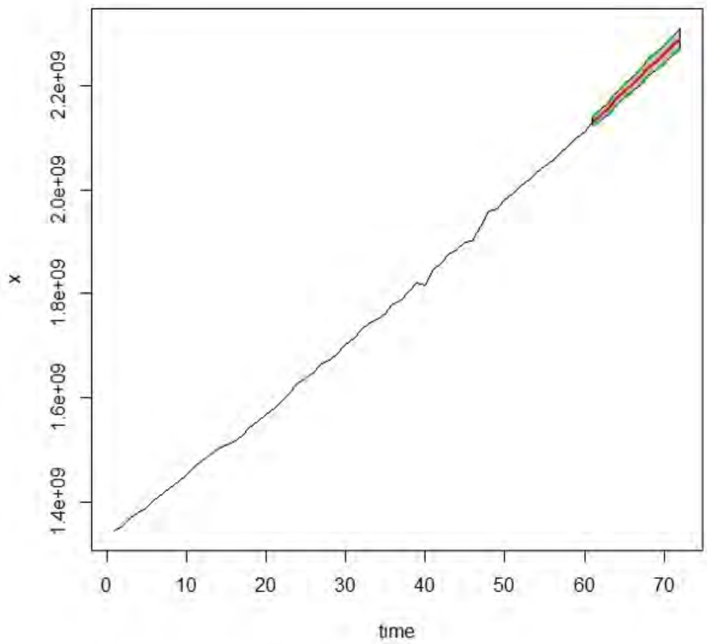


ภาพที่ 2.7 แสดงการเปรียบเทียบจำนวนเงินเอาประกันภัยจริงกับ

จำนวนเงินเอาประกันภัยจากแบบจำลอง



ภาพที่ 2.8 แสดงจำนวนเงินเอาประกันภัยพยากรณ์



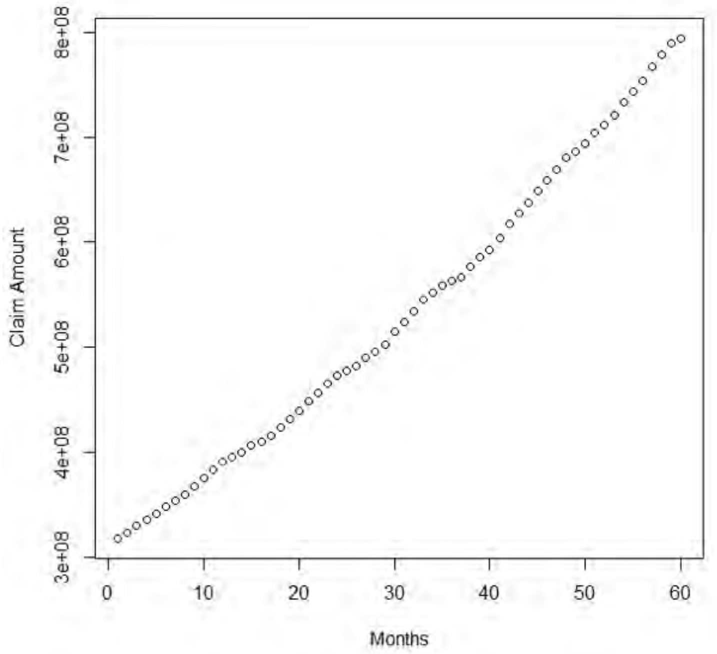
ภาพที่ 2.9 แสดงจำนวนเงินเอาประกันภัยจริงตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ต่อด้วยจำนวนเงินเอาประกันภัยพยากรณ์อีก 12 เดือนข้างหน้า

3. ประกันชีวิตแบบชั่วระยะเวลา

ประกันชีวิตแบบชั่วระยะเวลา เป็นการประกันชีวิตที่บริษัทประกันชีวิตจะจ่ายเงินให้แก่ผู้รับประโยชน์เมื่อผู้เอาประกันภัยเสียชีวิตในระยะเวลาประกันภัย วัตถุประสงค์เพื่อคุ้มครองการเสียชีวิตก่อนวัยอันสมควร การประกันชีวิตแบบนี้ไม่มีส่วนของการออมทรัพย์ เบี้ยประกันภัยจึงต่ำกว่าแบบอื่น ๆ และไม่มีเงินเหลือคืนให้หากผู้เอาประกันภัยอยู่จนครบกำหนดสัญญา โดยเราศึกษาข้อมูลจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 เนื่องจากข้อมูลสูญหาย 4 เดือน ได้แก่ เดือนมกราคม ปี พ.ศ. 2553 เดือนกุมภาพันธ์ ปี พ.ศ. 2555 เดือนพฤศจิกายน ปี พ.ศ.2555 และเดือนกรกฎาคม ปี พ.ศ. 2556 เราจึงแก้ไขข้อมูลโดยให้ข้อมูลที่สูญหายแทนด้วยค่าเฉลี่ยของข้อมูลก่อนหน้า 1 เดือนและข้อมูลหลังจากนั้น 1 เดือน

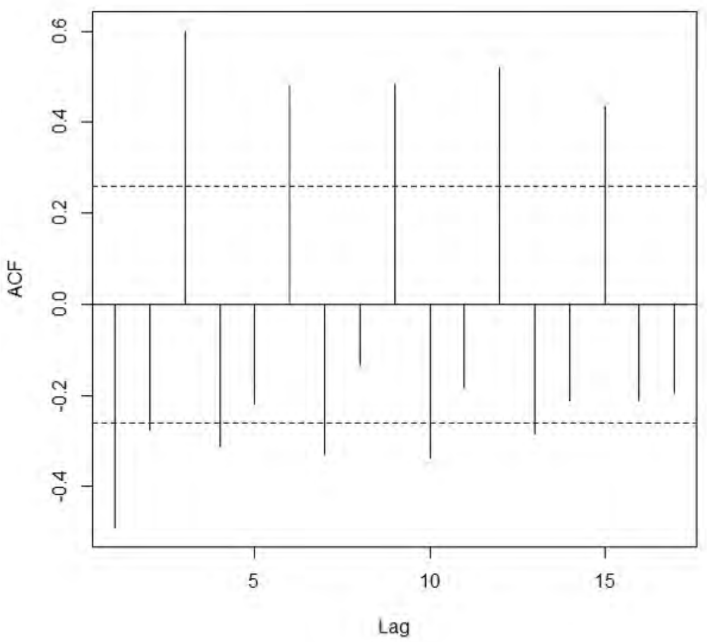
เดือน ปีพ.ศ.	จำนวนเงินเอาประกันภัย (บาท)
ม.ค. 2552	317,912,534
ก.พ. 2552	323,149,991
มี.ค. 2552	329,963,016
เม.ย. 2552	335,604,172
พ.ค. 2552	341,297,792
มิ.ย. 2552	348,528,985
⋮	⋮
ก.ค. 2556	743,451,557
ส.ค. 2556	753,455,429
ก.ย. 2556	767,815,225
ต.ค. 2556	778,299,903
พ.ย. 2556	789,896,862
ธ.ค. 2556	794,339,586

ตารางที่ 3.1 แสดงจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556

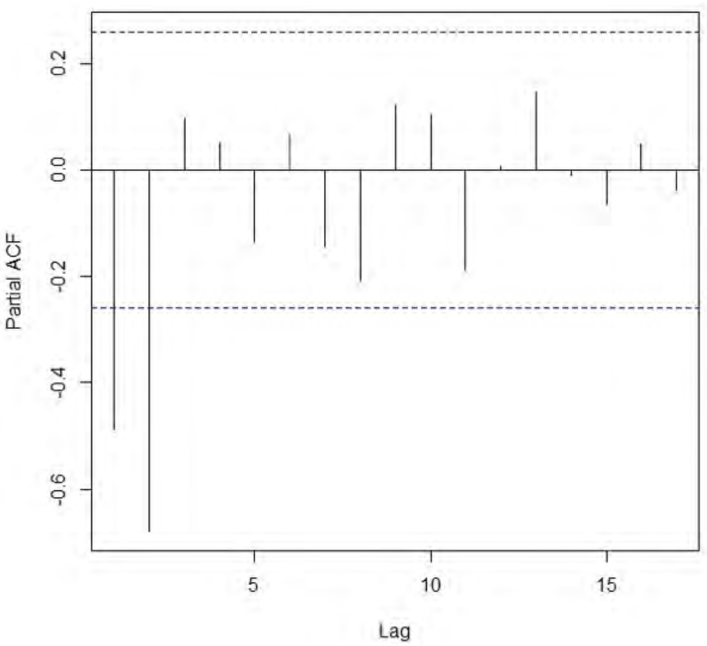


ภาพที่ 3.1 แสดงจำนวนเงินเอาประกันภัยของประกันชีวิตแบบเฉพาะกาลระหว่างปี พ.ศ. 2552-2556

ต่อไปเราจะนำข้อมูลจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปีพ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาทดสอบความนิ่ง (Stationary) และแปลงข้อมูลให้มีความนิ่ง โดยใช้วิธีการหาผลต่าง (differencing) จะได้ข้อมูลที่ผ่านการหาผลต่าง 3 ครั้งมีความนิ่ง เราจึงใช้ข้อมูลที่ผ่านการหาผลต่าง 3 ครั้งในการคาดเดาลำดับของ Autoregressive (AR) และ Moving Average (MA) โดยพิจารณาค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) แสดงในภาพที่ 2.2 และภาพที่ 2.3



ภาพที่ 3.2 แสดงกราฟ ACF ของจำนวนเงินเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 3



ภาพที่ 3.3 แสดงกราฟ PACF ของจำนวนเงินเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 3

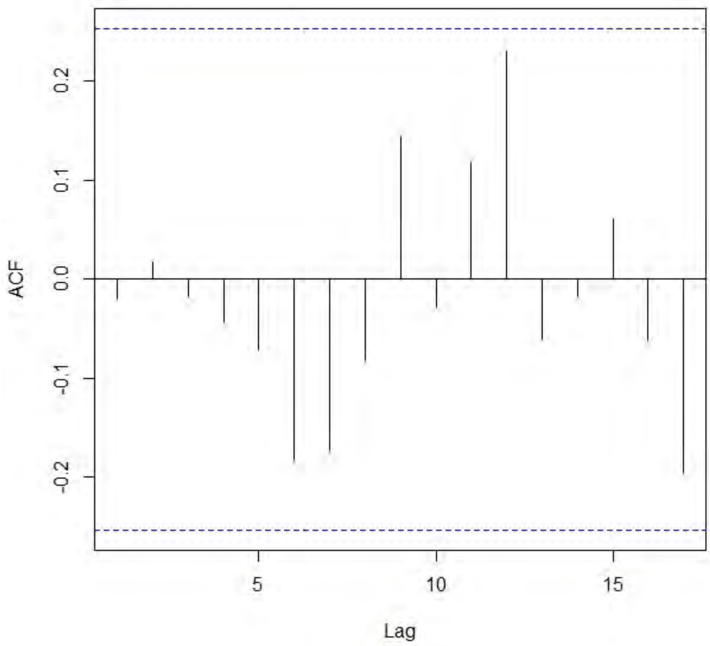
จากภาพที่ 3.2 และภาพที่ 3.3 พบว่าอนุกรมเวลาของผลต่างอันดับ 2 มีค่าสหสัมพันธ์ในตัวเอง (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,3,4,6,7,9,10,12 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจะนำข้อมูลจำนวนเงินเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาพิจารณา ARIMA(p,d,q) เมื่อ $p \in \{1,2\}, q \in \{1,2,\ldots,12\}$ และ d=3 ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC ต่ำที่สุด ผลการศึกษาแสดงดังตารางที่ 3.2 ต่อไปนี้

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,3,1)	1866.163	ARIMA(1,3,9)	1862.102	ARIMA(2,3,5)	1865.464
ARIMA(1,3,2)	1865.345	ARIMA(1,3,10)	1863.481	ARIMA(2,3,6)	1862.574
ARIMA(1,3,3)	1871.964	ARIMA(1,3,11)	1864.432	ARIMA(2,3,7)	1862.895
ARIMA(1,3,4)	1863.487	ARIMA(1,3,12)	1867.833	ARIMA(2,3,8)	1863.357
ARIMA(1,3,5)	1863.924	ARIMA(2,3,1)	1865.590	ARIMA(2,3,9)	1863.532
ARIMA(1,3,6)	1860.852	ARIMA(2,3,2)	1867.285	ARIMA(2,3,10)	1863.364
ARIMA(1,3,7)	1862.852	ARIMA(2,3,3)	1862.446	ARIMA(2,3,11)	1865.235
ARIMA(1,3,8)	**1860.516**	ARIMA(2,3,4)	1865.038	ARIMA(2,3,12)	1862.727

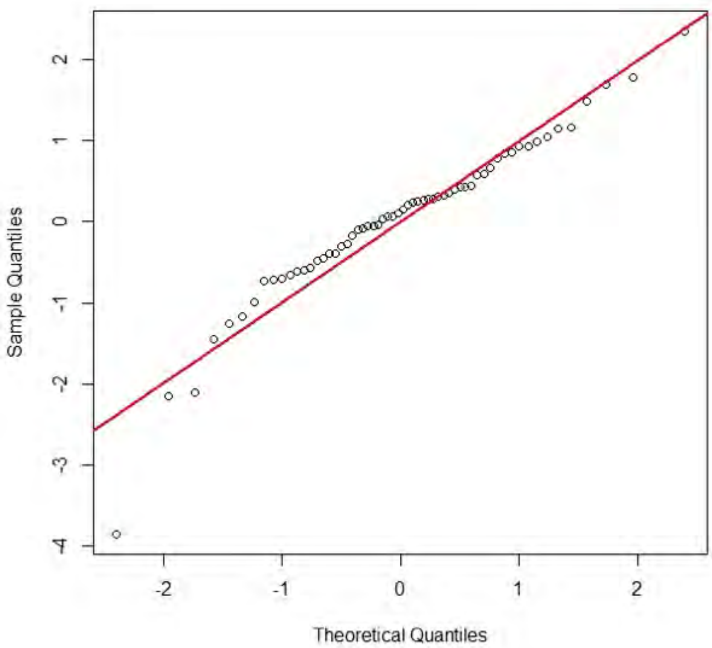
ตารางที่ 3.2 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนเงินเอาประกันภัย

จากตารางที่ 3.2 เมื่อพิจารณาจากค่า AIC เนื่องจาก ARIMA(1,3,8) ให้ค่า AIC น้อยที่สุด คือ 1860.516 ดังนั้นแบบจำลองอนุกรมเวลาของตัวแบบสำหรับการจำลองที่เหมาะสมสำหรับข้อมูลจำนวนเงินเอาประกันภัยของประกันชีวิตแบบเฉพาะกาล คือ ARIMA(1,3,8)

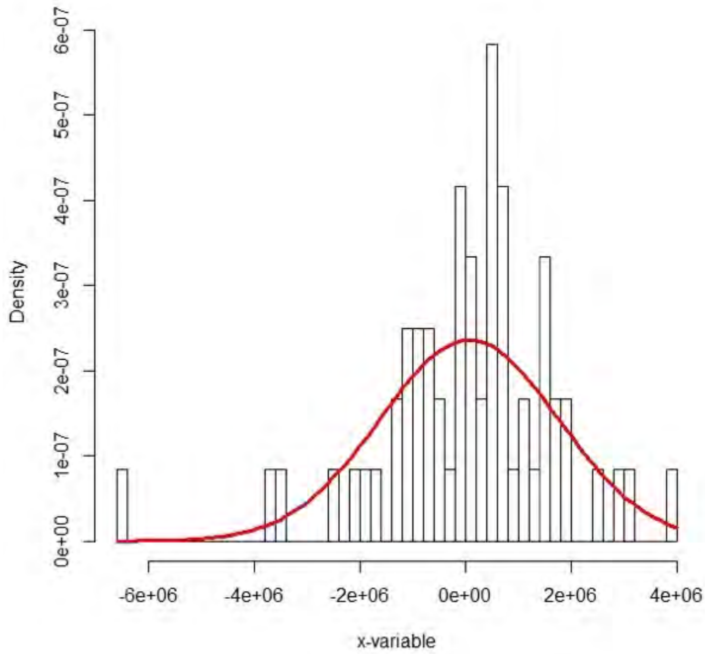
ต่อไปจะนำค่าเศษเหลือ (Residuals) ซึ่งคือค่าผลต่างระหว่างจำนวนเงินเอาประกันภัยจริงและจำนวนเงินเอาประกันภัยพยากรณ์ มาทดสอบความเป็นอิสระต่อกันของค่าเศษเหลือด้วยกราฟค่าสหสัมพันธ์ในตัวเองของเศษเหลือ (Autocorrelation residuals plot) ทดสอบการแจกแจงแบบปกติจากกราฟโดยวิธี Quantile-Quantile Plot หรือ QQ plots และทดสอบการแจกแจงปกติด้วยกราฟฮิสโตแกรม แสดงในภาพที่ 3.4 ภาพที่ 3.5 และภาพที่ 3.6 และทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov แสดงในตารางที่ 3.3



ภาพที่ 3.4 แสดงค่าสหสัมพันธ์ในตัวเองของเศษเหลือ



ภาพที่ 3.5 แสดงการทดสอบ Quantile-Quantile Plot ของค่าเศษเหลือ



ภาพที่ 3.6 แสดงทดสอบการแจกแจงปกติด้วยกราฟฮิสโตแกรม

แบบจำลอง	ค่า p-value ของ Kolmogorov-Smirnov
ARIMA(1,3,8)	0.4728

ตารางที่ 3.3 แสดงค่า p-value ของ Kolmogorov-Smirnov

จากภาพที่ 3.4 แบบจำลองแสดงให้เห็นว่าค่าสหสัมพันธ์ในตัวเองของเศษเหลือไม่มีความสัมพันธ์กัน สำหรับการทดสอบการแจกแจงปกติของเศษเหลือของแบบจำลองจะได้ว่า จากภาพที่ 3.5 แบบจำลองแสดงการกระจายของเศษเหลือที่คล้ายจะเป็นการแจกแจงปกติ จากภาพที่ 3.6 กราฟฮิสโตแกรมของแบบจำลองแสดงให้เห็นว่าเศษเหลือของแบบจำลองมีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติ แต่มีความถี่ของบางค่าสูงกว่าความถี่ของการแจกแจงแบบปกติ และมีความถี่ของบางค่าต่ำกว่าความถี่ของการแจกแจงแบบปกติ และเมื่อทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov ในตารางที่ 3.3 พบว่าค่า p-value ของ Kolmogorov-Smirnov เท่ากับ 0.4728 ซึ่งมากกว่า 0.05 แสดงว่าเศษเหลือของแบบจำลองมีการแจกแจงแบบปกติ เมื่อพิจารณาภาพที่ 3.5 – 3.6 และ ตารางที่ 3.3 จึงสามารถสรุปได้ว่าแบบจำลองเศษเหลือของ ARIMA(1,3,8) มีการแจกแจงใกล้เคียงกับการแจกแจงปกติ

ต่อไปจะแสดงการพยากรณ์จำนวนเงินเอาประกันภัยโดยใช้แบบจำลอง ARIMA(1,3,8)

พารามิเตอร์	ARIMA(1,3,8)
Ar1	-0.6919
Ma1	-1.0252
Ma2	-1.0764
Ma3	1.0013
Ma4	0.0174
Ma5	-0.4220
Ma6	1.0645
Ma7	-0.2444
Ma8	-0.3152

ตารางที่ 3.4 แสดงพารามิเตอร์ของจำนวนเงินเอาประกันภัย

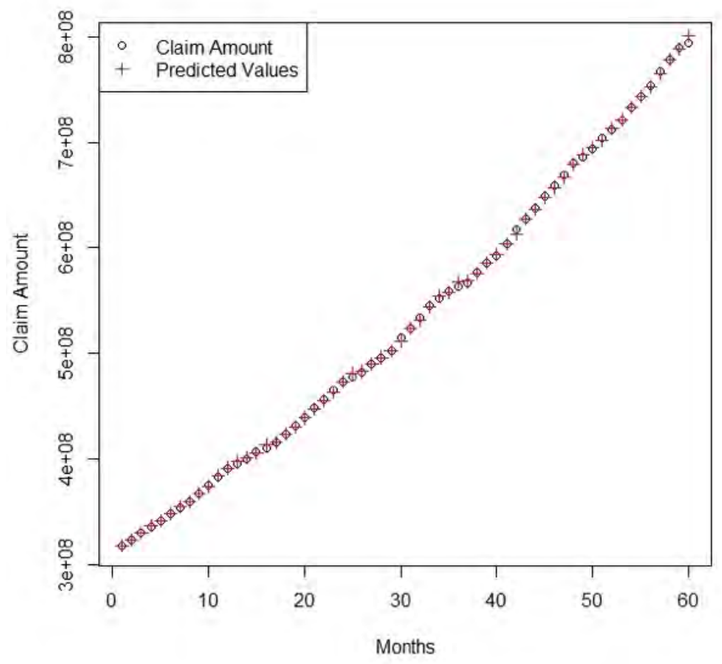
จากตารางที่ 3.4 จะได้ว่าแบบจำลอง ARIMA(1,3,8) เขียนค่าประมาณพารามิเตอร์จากการทำผลต่างอันดับ 3 ได้สมการต่อไปนี้

$$\Delta^3 y_t = -0.6919\Delta^3 y_{t-1} + 1.0252\varepsilon_{t-1} + 1.0764\varepsilon_{t-2} - 1.0013\varepsilon_{t-3} - 0.0174\varepsilon_{t-4}$$
$$+0.4220\varepsilon_{t-5} - 1.0645\varepsilon_{t-6} + 0.2444\varepsilon_{t-7} + 0.3152\varepsilon_{t-8} + \varepsilon_t$$

จัดรูปสมการได้ดังนี้

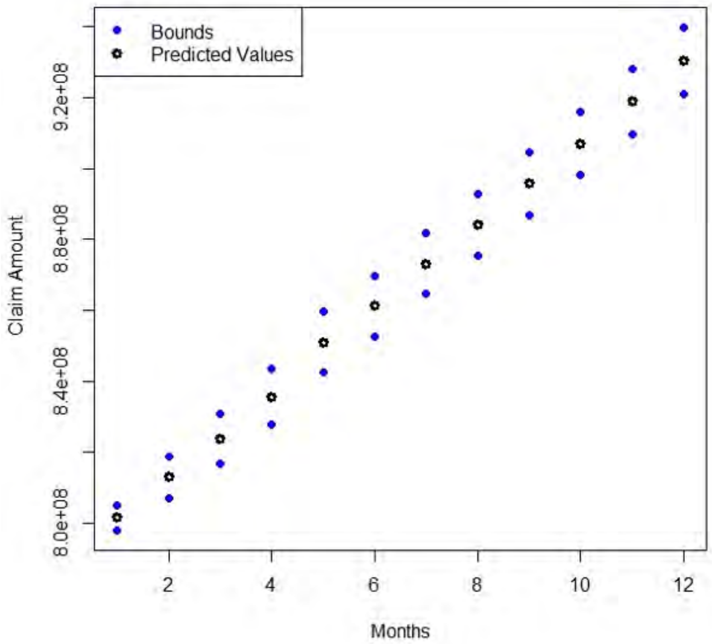
$$y_t = 2.3081y_{t-1} - 0.9241y_{t-2} - 1.0757y_{t-3} + 0.6919y_{t-4} + 1.0252\varepsilon_{t-1}$$
$$+1.0764\varepsilon_{t-2} - 1.0013\varepsilon_{t-3} - 0.0174\varepsilon_{t-4} + 0.4220\varepsilon_{t-5} - 1.0645\varepsilon_{t-6}$$
$$+0.2444\varepsilon_{t-7} + 0.3152\varepsilon_{t-8} + \varepsilon_t$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนเงินเอาประกันภัยจริงกับจำนวนเงินเอาประกันภัยพยากรณ์ตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 แสดงจำนวนเงินเอาประกันภัยพยากรณ์อีก 12 เดือนข้างหน้า และแสดงจำนวนเงินเอาประกันภัยจริงตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ต่อด้วยจำนวนเงินเอาประกันภัยพยากรณ์อีก 12 เดือนข้างหน้า ดังภาพที่ 3.7 ภาพที่ 3.8 และภาพที่ 3.9

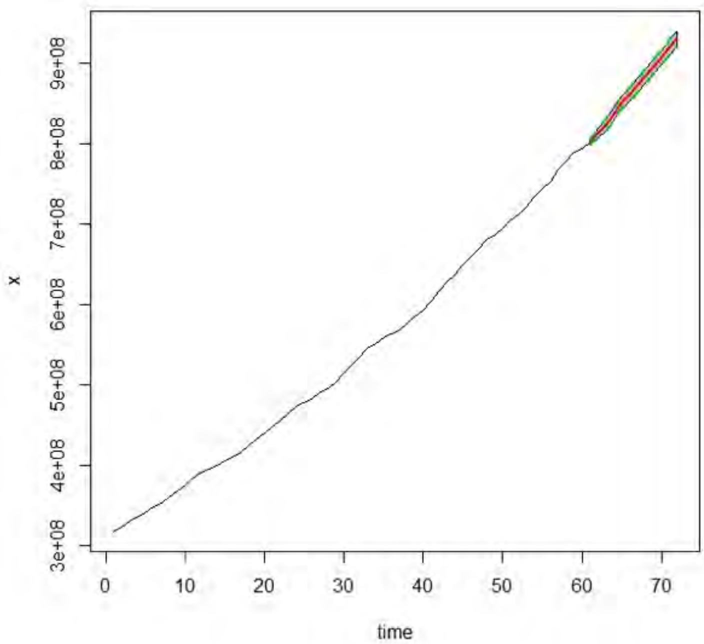


ภาพที่ 3.7 แสดงการเปรียบเทียบจำนวนเงินเอาประกันภัยจริงกับ

จำนวนเงินเอาประกันภัยจากแบบจำลอง



ภาพที่ 3.8 แสดงจำนวนเงินเอาประกันภัยพยากรณ์



ภาพที่ 3.9 แสดงจำนวนเงินเอาประกันภัยจริงตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ต่อด้วยจำนวนเงินเอาประกันภัยพยากรณ์อีก 12 เดือนข้างหน้า

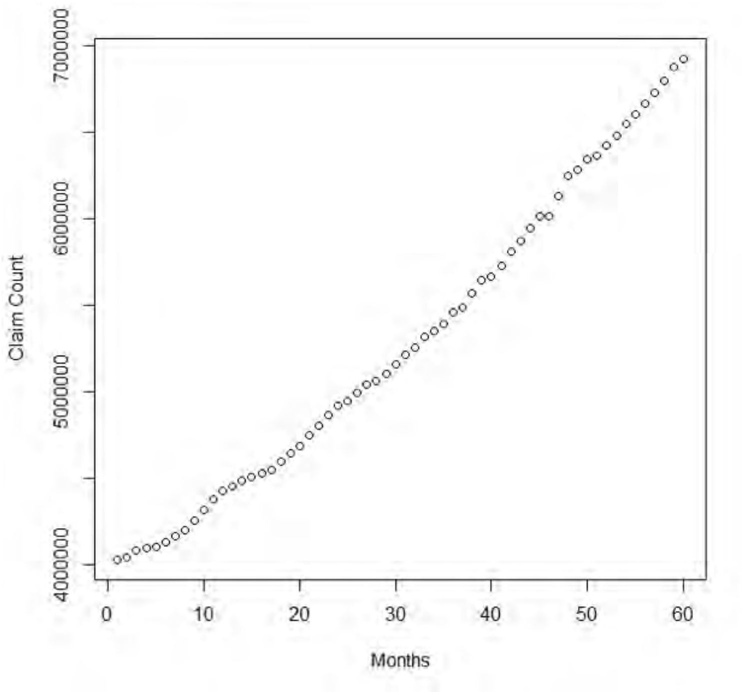
3.2 จำนวนการเอาประกันภัยของบริษัทประกันภัยในประเทศไทย

จากที่ได้ศึกษาจำนวนเงินเอาประกันภัยโดยใช้อนุกรมเวลาแบบพื้นฐาน ในหัวข้อถัดไปจะศึกษาจำนวนการเอาประกันภัยโดยใช้อนุกรมเวลาเชิงการนับโดยแบบจำลองที่ได้ศึกษา ได้แก่ แบบจำลองสำหรับการแจกแจงปัวซง และแบบจำลองสำหรับการแจกแจงทวินามลบ

1. ประกันชีวิตแบบตลอดชีพ

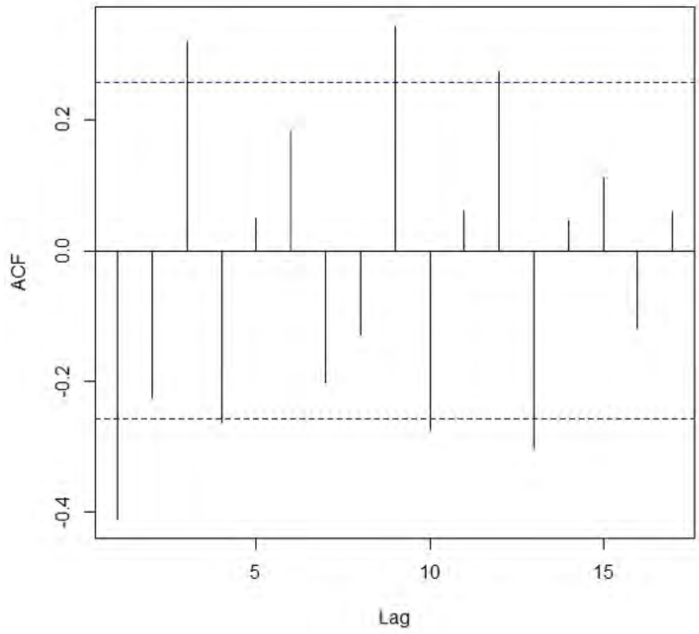
เดือน ปีพ.ศ.	จำนวนการเอาประกันภัย (ราย)
ม.ค. 2552	4,028,023
ก.พ. 2552	4,039,916
มี.ค. 2552	4,081,946
เม.ย. 2552	4,096,408
พ.ค. 2552	4,098,030
มิ.ย. 2552	4,125,634
⋮	⋮
ก.ค. 2556	6,603,031
ส.ค. 2556	6,662,321
ก.ย. 2556	6,723,709
ต.ค. 2556	6,791,727
พ.ย. 2556	6,872,659
ธ.ค. 2556	6,922,057

ตารางที่ 1.1 แสดงจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556

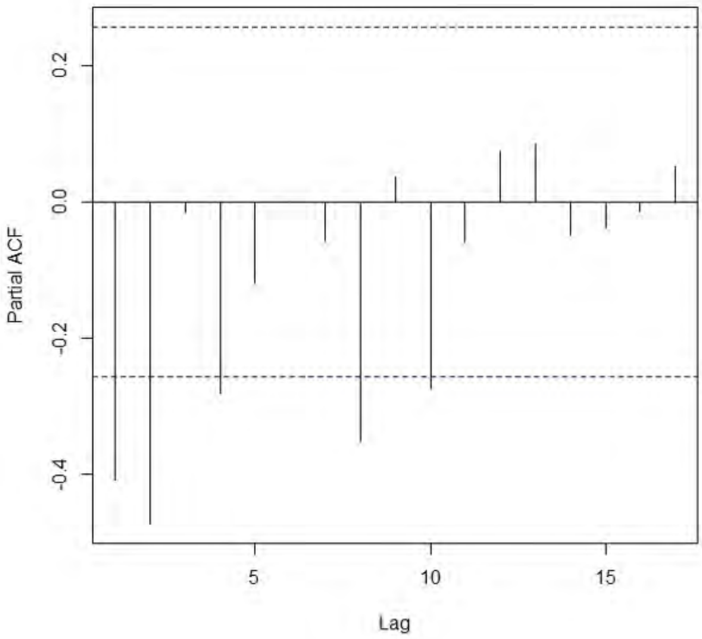


ภาพที่ 1.1 แสดงจำนวนการเอาประกันภัยของประกันชีวิตแบบตลอดชีพระหว่างปี พ.ศ. 2552-2556

ต่อไปเราจะนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาแปลงข้อมูลโดยการหารากที่สองของข้อมูล จากนั้นนำข้อมูลที่แปลงแล้วมาทดสอบความนิ่ง (Stationary) และแปลงข้อมูลให้มีความนิ่ง โดยใช้วิธีการหาผลต่าง (differencing) จะได้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งมีความนิ่ง เราจึงใช้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งในการคาดเดาลำดับของ Autoregressive (AR) และ Moving Average (MA) โดยพิจารณาค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) แสดงในภาพที่ 1.2 และภาพที่ 1.3



ภาพที่ 1.2 แสดงกราฟ ACF ของการจำนวนเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2



ภาพที่ 1.3 แสดงกราฟ PACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2

จากภาพที่ 1.2 และภาพที่ 1.3 พบว่าอนุกรมเวลาของผลต่างอันดับ 2 มีค่าสหสัมพันธ์ในตัวเอง (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,3,9 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,8 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจะนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปีพ.ศ. 2552 ถึงเดือนธันวาคม ปีพ.ศ. 2556 มาพิจารณา ARIMA(p,d,q) เมื่อ p∈{1,2,...,8},q=∈{1,2,...,9}, และ d=2 ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC ต่ำที่สุด ผลการศึกษาแสดงดังตารางที่ 1.2 ต่อไปนี้

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,2,1)	348.9918	ARIMA(4,2,1)	348.9590	ARIMA(7,2,1)	351.3789
ARIMA(1,2,2)	**343.9517**	ARIMA(4,2,2)	349.2866	ARIMA(7,2,2)	350.3828
ARIMA(1,2,3)	345.4831	ARIMA(4,2,3)	347.2932	ARIMA(7,2,3)	349.1910
ARIMA(1,2,4)	346.6970	ARIMA(4,2,4)	347.0830	ARIMA(7,2,4)	351.2474
ARIMA(1,2,5)	347.4560	ARIMA(4,2,5)	347.3897	ARIMA(7,2,5)	354.6746
ARIMA(1,2,6)	349.4725	ARIMA(4,2,6)	351.4051	ARIMA(7,2,6)	351.9834
ARIMA(1,2,7)	347.9801	ARIMA(4,2,7)	349.5827	ARIMA(7,2,7)	352.9793
ARIMA(1,2,8)	346.7761	ARIMA(4,2,8)	350.0022	ARIMA(7,2,8)	351.5484
ARIMA(1,2,9)	346.3158	ARIMA(4,2,9)	350.6477	ARIMA(7,2,9)	354.3363
ARIMA(2,2,1)	350.5741	ARIMA(5,2,1)	350.9240	ARIMA(8,2,1)	347.6621
ARIMA(2,2,2)	345.3438	ARIMA(5,2,2)	350.0326	ARIMA(8,2,2)	345.1732
ARIMA(2,2,3)	347.1449	ARIMA(5,2,3)	348.5132	ARIMA(8,2,3)	346.0942
ARIMA(2,2,4)	347.5235	ARIMA(5,2,4)	348.8967	ARIMA(8,2,4)	347.8234
ARIMA(2,2,5)	349.1593	ARIMA(5,2,5)	349.7575	ARIMA(8,2,5)	349.7876
ARIMA(2,2,6)	350.5392	ARIMA(5,2,6)	348.3581	ARIMA(8,2,6)	351.0891
ARIMA(2,2,7)	350.4285	ARIMA(5,2,7)	351.7516	ARIMA(8,2,7)	351.2092
ARIMA(2,2,8)	349.2727	ARIMA(5,2,8)	352.7814	ARIMA(8,2,8)	353.0157
ARIMA(2,2,9)	347.9666	ARIMA(5,2,9)	349.8787	ARIMA(8,2,9)	355.3995

ตารางที่ 1.2 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(3,2,1)	350.3391	ARIMA(6,2,1)	352.5694
ARIMA(3,2,2)	346.8766	ARIMA(6,2,2)	351.5068
ARIMA(3,2,3)	347.8098	ARIMA(6,2,3)	350.9673
ARIMA(3,2,4)	346.0486	ARIMA(6,2,4)	351.9562
ARIMA(3,2,5)	348.0383	ARIMA(6,2,5)	351.4518
ARIMA(3,2,6)	352.8166	ARIMA(6,2,6)	351.7515
ARIMA(3,2,7)	352.2713	ARIMA(6,2,7)	353.7395
ARIMA(3,2,8)	350.4051	ARIMA(6,2,8)	353.8109
ARIMA(3,2,9)	347.9764	ARIMA(6,2,9)	351.4822

ตารางที่ 1.2 (ต่อ) แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

จากตารางที่ 1.2 เมื่อพิจารณาจากค่า AIC เนื่องจาก ARIMA(1,2,2) ให้ค่า AIC น้อยที่สุด คือ 343.9517 ดังนั้นแบบจำลองอนุกรมเวลาของตัวแบบสำหรับการจำลองที่เหมาะสมสำหรับข้อมูลจำนวนการเอาประกันภัยของประกันชีวิตแบบตลอดชีพ คือ ARIMA(1,2,2)

ต่อไปจะแสดงการพยากรณ์จำนวนการเอาประกันภัยโดยใช้แบบจำลอง ARIMA(1,2,2)

พารามิเตอร์	ARIMA(1,2,2)
Ar1	-0.7096
Ma1	0.0846
Ma2	-0.9154

ตารางที่ 1.3 แสดงพารามิเตอร์ของจำนวนการเอาประกันภัย

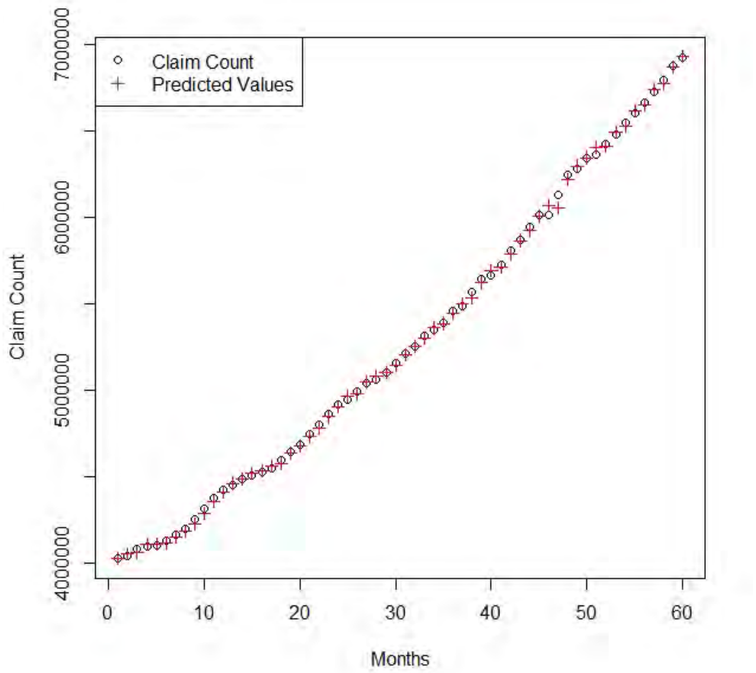
จากตารางที่ 1.3 จะได้ว่าแบบจำลอง ARIMA(1,2,2) เขียนค่าประมาณพารามิเตอร์จากการทำผลต่างอันดับ 2 ได้สมการต่อไปนี้

$$\Delta^2 y_t = -0.7096\Delta^2 y_{t-1} - 0.0846\varepsilon_{t-1} + 0.9154\varepsilon_{t-2} + \varepsilon_t$$

จัดรูปสมการได้ดังนี้

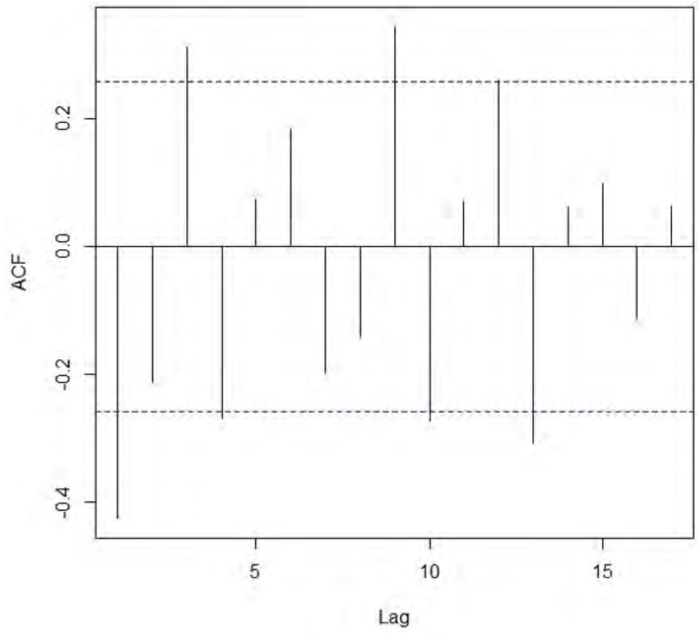
$$y_t = 1.2904y_{t-1} + 0.4192y_{t-2} - 0.7096y_{t-3} - 0.0846\varepsilon_{t-1} + 0.9154\varepsilon_{t-2} + \varepsilon_t$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ดังภาพที่ 1.4

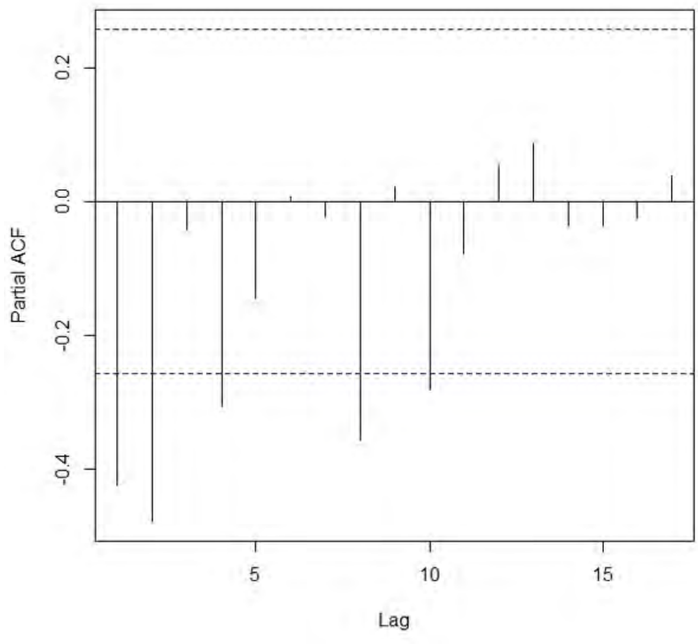


ภาพที่ 1.4 แสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับ
จำนวนการเอาประกันภัยจากแบบจำลอง

ต่อไปเราจะศึกษาหาตัวแบบสำหรับการแจกแจงปัวซงและตัวแบบสำหรับการแจกแจงทวินามลบ โดยนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาทดสอบความนิ่ง (Stationary) และแปลงข้อมูลให้มีความนิ่ง โดยใช้วิธีการหาผลต่าง (differencing) จะได้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งมีความนิ่ง เราจึงใช้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งในการคาดเดาลำดับของ Autoregressive (AR) และ Moving Average (MA) โดยพิจารณาค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) แสดงในภาพที่ 1.5 และภาพที่ 1.6



ภาพที่ 1.5 แสดงกราฟ ACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2



ภาพที่ 1.6 แสดงกราฟ PACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2

จากภาพที่ 1.5 และภาพที่ 1.6 พบว่าอนุกรมเวลาของผลต่างอันดับ 2 มีค่าสหสัมพันธ์ในตัวเอง (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,3,9 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,4,8,10 เนื่องจากข้อมูลผ่านการทำผลต่างอันดับ 2 ดังนั้นเมื่อแก้สมการแล้วจะได้ว่าต้องพิจารณาค่าสังเกตในอดีตอีกใน lag ที่ 11,12 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจะนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาพิจารณาหาแบบจำลองสำหรับการแจกแจงปัวซงและแบบจำลองสำหรับการแจกแจงทวินามลบ ตามลำดับ

พิจารณาแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซง

แบบจำลอง		ค่า AIC	แบบจำลอง		ค่า AIC
ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต		ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต	
1	1	1,038,763	11	9	5,942,483
2	1	2,784,088	12	9	6,472,901
4	1	3,717,845	1	1,3	14,890.21
8	1	5,591,418	1	1,9	947,313.1
10	1	6,400,864	**1**	**3,9**	**9,565.726**
11	1	6,763,881	2	3,9	28,314.86
12	1	6,624,292	4	3,9	1,668,224
1	3	60,847.68	8	3,9	4,859,932
2	3	2,800,805	10	3,9	5,415,916
4	3	3,756,413	11	3,9	5,824,896
8	3	5,638,186	12	3,9	6,198,775
10	3	6,451,594	1,2	3,9	618,829.9
11	3	6,812,514	1,4	3,9	153,119.6
12	3	6,925,578	1,8	3,9	113,766
1	9	1,080,115.9	1,10	3,9	410,611.5

ตารางที่ 1.4 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

แบบจำลอง		ค่า AIC	แบบจำลอง		ค่า AIC
ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต		ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต	
2	9	344,067	1,11	3,9	127,224.9
4	9	2,058,035	1,12	3,9	137,707.6
8	9	4,526,804	1	1,3,9	192,434.7
10	9	5,564,459			

ตารางที่ 1.4 (ต่อ) แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

จากตารางที่ 1.4 เมื่อพิจารณาค่า AIC จะได้ว่าแบบจำลองที่มีค่าสังเกตในอดีตเท่ากับ 1 และค่าเฉลี่ยในอดีตเท่ากับ 3 และ 9 ให้ค่า AIC น้อยที่สุดคือ 9,565.726 ดังนั้นแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซงที่เหมาะสมสำหรับจำนวนเอากรมธรรม์ประกันภัยโดยส่วนของข้อมูลในอดีตของตัวแปรที่สนใจ (ค่าสังเกตในอดีต) คือ ข้อมูลก่อนหน้า 1 หน่วยเวลา และส่วนของค่าเฉลี่ยแบบมีเงื่อนไข (ค่าเฉลี่ยในอดีต) คือ ค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 3 หน่วยเวลา และค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 9 หน่วยเวลา

ต่อไปจะแสดงการพยากรณ์จำนวนการเอาประกันภัยโดยใช้แบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซงที่มีค่าสังเกตในอดีตเท่ากับ 1 และค่าเฉลี่ยในอดีตเท่ากับ 3 และ 9

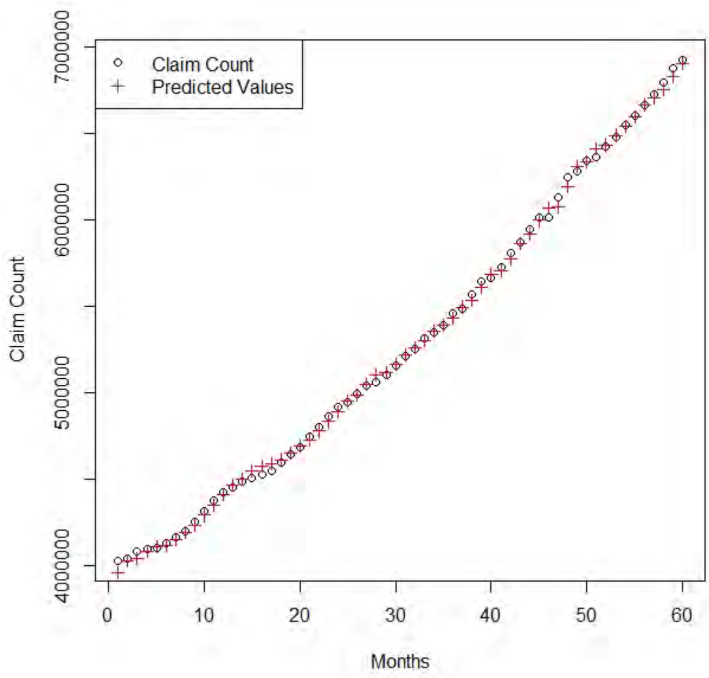
พารามิเตอร์	แบบจำลอง
Intercept	0.1519
Beta 1	0.9729
Alpha 3	0.2387
Alpha 9	-0.2216

ตารางที่ 1.5 แสดงพารามิเตอร์ของจำนวนการเอาประกันภัย

จากตารางที่ 1.5 จะได้ว่าแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซง ดังสมการต่อไปนี้

$$\log(\lambda_t) = 0.1519 + 0.9729Y_{t-1} + 0.2387\lambda_{t-3} - 0.2216\lambda_{t-9}$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ดังภาพที่ 1.7



ภาพที่ 1.7 แสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับ
จำนวนการเอาประกันภัยจากแบบจำลอง

พิจารณาแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบ

แบบจำลอง		ค่า AIC	แบบจำลอง		ค่า AIC
ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต		ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต	
1	1	1,691.965	11	9	1,796.236
2	1	1,750.939	12	9	1,801.629
4	1	1,768.055	1	1,3	1,435.196
8	1	1,792.364	1	1,9	1,688.548
10	1	1,800.47	**1**	**3,9**	**1,408.658**

ตารางที่ 1.6 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

แบบจำลอง		ค่า AIC	แบบจำลอง		ค่า AIC
ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต		ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต	
11	1	1,803.787	2	3,9	1,477.984
12	1	1,803.102	4	3,9	1,724.222
1	3	1,517.71	8	3,9	1,787.177
2	3	1,751.275	10	3,9	1,793.037
4	3	1,768.655	11	3,9	1,797.362
8	3	1,792.859	12	3,9	1,800.836
10	3	1,800.944	1,2	3,9	1,666.265
11	3	1,804.219	1,4	3,9	1,576.644
12	3	1,805.913	1,8	3,9	1,559.055
1	9	1,587.59	1,10	3,9	1,638.883
2	9	1,622.485	1,11	3,9	1,570.653
4	9	1,733.607	1,12	3,9	1,580.228
8	9	1,780.633	1	1,3,9	1,591.651
10	9	1,792.354			

ตารางที่ 1.6 (ต่อ) แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

จากตารางที่ 1.6 เมื่อพิจารณาค่า AIC จะได้ว่าแบบจำลองที่มีค่าสังเกตในอดีตเท่ากับ 1 และค่าเฉลี่ยในอดีตเท่ากับ 3 และ 9 ให้ค่า AIC น้อยที่สุดคือ 1,408.658 ดังนั้นแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบที่เหมาะสมสำหรับจำนวนการเอาประกันภัยโดยส่วนของข้อมูลในอดีตของตัวแปรที่สนใจ (ค่าสังเกตในอดีต) คือ ข้อมูลก่อนหน้า 1 หน่วยเวลา และส่วนของค่าเฉลี่ยแบบมีเงื่อนไข (ค่าเฉลี่ยในอดีต) คือ ค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 3 หน่วยเวลา และค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 9 หน่วยเวลา

ต่อไปจะแสดงการพยากรณ์จำนวนการเอาประกันภัยโดยใช้แบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบที่มีค่าสังเกตในอดีตเท่ากับ 1 และค่าเฉลี่ยในอดีตเท่ากับ 3 และ 9

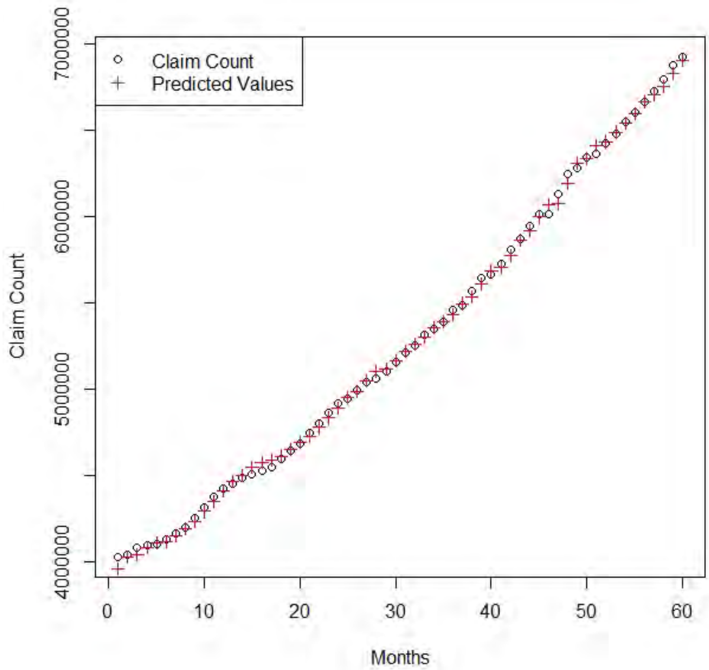
พารามิเตอร์	แบบจำลอง
Intercept	0.1519
Beta 1	0.9729
Alpha 3	0.2387
Alpha 9	-0.2216

ตารางที่ 1.7 แสดงพารามิเตอร์ของจำนวนการเอาประกันภัย

จากตารางที่ 1.7 จะได้ว่าแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบ ดังสมการต่อไปนี้

$$\log(\lambda_t) = 0.1519 + 0.9729Y_{t-1} + 0.2387\lambda_{t-3} - 0.2216\lambda_{t-9}$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ดังภาพที่ 1.8



ภาพที่ 1.8 แสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับ

จำนวนการเอาประกันภัยจากแบบจำลอง

เมื่อพิจารณาจากภาพการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยพยากรณ์ของทั้งสามแบบจำลองจะเห็นว่าแต่ละแบบจำลองมีความเหมาะสมในการพยากรณ์ เราจึงมาพิจารณาค่าคาดเคลื่อนกำลังสองเพื่อหาแบบจำลองที่เหมาะสมที่สุดเราจะเลือกแบบจำลองที่มีค่าคาดเคลื่อนกำลังสองต่ำที่สุด ผลการศึกษาแสดงดังตารางที่ 1.8 ต่อไปนี้

แบบจำลอง	ค่าคาดเคลื่อนกำลังสอง
ARIMA(1,2,2)	23,931,470,748
เชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซง ค่าสังเกตในอดีตเท่ากับ 1 ค่าเฉลี่ยในอดีตเท่ากับ 3,9	43,828,552,540
เชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบ ค่าสังเกตในอดีตเท่ากับ 1 ค่าเฉลี่ยในอดีตเท่ากับ 3,9	43,828,552,540

ตารางที่ 1.8 แสดงค่าคาดเคลื่อนกำลังสองของแบบจำลอง

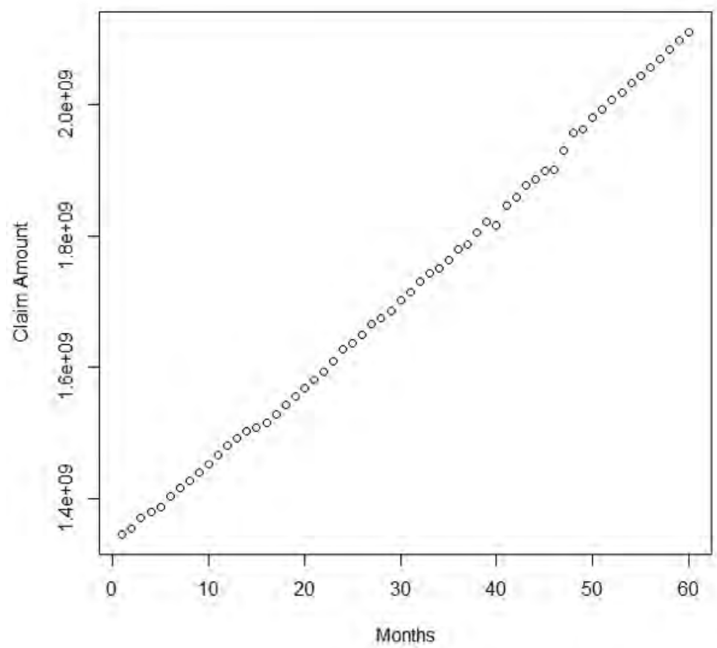
จากตารางที่ 1.8 เมื่อพิจารณาจากค่าคาดเคลื่อนกำลังสองเนื่องจาก ARIMA(1,2,2) ให้ค่าคาดเคลื่อนกำลังสองน้อยที่สุด คือ 23,931,470,748 ดังนั้นแบบจำลองอนุกรมเวลาสำหรับการจำลองที่เหมาะสมสำหรับข้อมูลจำนวนการเอาประกันภัยของประกันชีวิตแบบตลอดชีพ คือ ARIMA(1,2,2)

2. ประกันชีวิตแบบสะสมทรัพย์

เดือน ปีพ.ศ.	จำนวนการเอาประกันภัย (ราย)
ม.ค. 2552	7,043,369
ก.พ. 2552	7,089,628
มี.ค. 2552	7,165,248
เม.ย. 2552	7,198,703
พ.ค. 2552	7,231,542
มิ.ย. 2552	7,293,495
⋮	⋮
ก.ค. 2556	9,851,602
ส.ค. 2556	9,880,412
ก.ย. 2556	9,925,622
ต.ค. 2556	9,974,535
พ.ย. 2556	10,022,873
ธ.ค. 2556	10,077,049

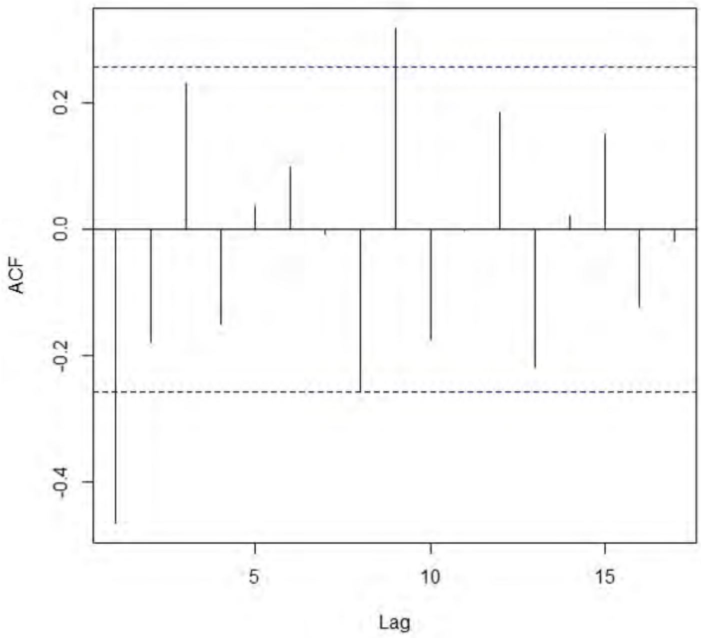
ตารางที่ 2.1 แสดงจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม

ปี พ.ศ. 2556

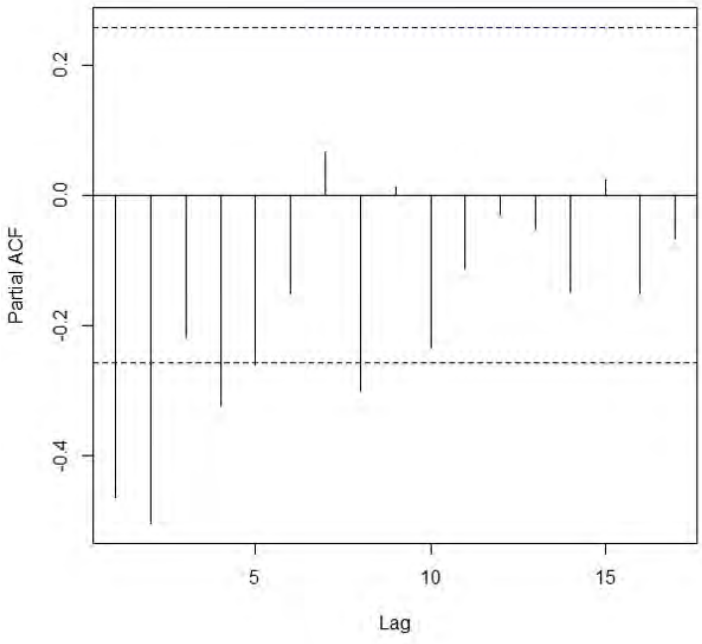


ภาพที่ 2.1 แสดงจำนวนการเอาประกันภัยของประกันชีวิตแบบสะสมทรัพย์ระหว่างปี พ.ศ. 2552-2556

ต่อไปเราจะนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาแปลงข้อมูลโดยการหารากที่สองของข้อมูล จากนั้นนำข้อมูลที่แปลงแล้วมาทดสอบความนิ่ง (Stationary) และแปลงข้อมูลให้มีความนิ่ง โดยใช้วิธีการหาผลต่าง (differencing) จะได้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งมีความนิ่ง เราจึงใช้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งในการคาดเดาลำดับของ Autoregressive (AR) และ Moving Average (MA) โดยพิจารณาค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) แสดงในภาพที่ 2.2 และภาพที่ 2.3



ภาพที่ 2.2 แสดงกราฟ ACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2



ภาพที่ 2.3 แสดงกราฟ PACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2

จากภาพที่ 2.2 และภาพที่ 2.3 พบว่าอนุกรมเวลาของผลต่างอันดับ 2 มีค่าสหสัมพันธ์ในตัวเอง (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,9 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,4,8 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจะนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปีพ.ศ. 2552 ถึงเดือนธันวาคม ปีพ.ศ. 2556 มาพิจารณา ARIMA(p,d,q) เมื่อ p∈{1,2,...,8},q=∈{1,2,...,9}, และ d=2 ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC ต่ำที่สุด ผลการศึกษาแสดงดังตารางที่ 1.2 ต่อไปนี้

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,2,1)	325.1582	ARIMA(4,2,1)	326.3313	ARIMA(7,2,1)	332.0631
ARIMA(1,2,2)	323.2902	ARIMA(4,2,2)	328.3146	ARIMA(7,2,2)	328.3097
ARIMA(1,2,3)	**320.6421**	ARIMA(4,2,3)	325.9487	ARIMA(7,2,3)	329.8152
ARIMA(1,2,4)	322.4956	ARIMA(4,2,4)	326.0985	ARIMA(7,2,4)	327.7887
ARIMA(1,2,5)	324.4219	ARIMA(4,2,5)	328.9624	ARIMA(7,2,5)	329.9061
ARIMA(1,2,6)	326.3994	ARIMA(4,2,6)	330.9479	ARIMA(7,2,6)	331.7910
ARIMA(1,2,7)	328.3993	ARIMA(4,2,7)	332.9446	ARIMA(7,2,7)	331.8700
ARIMA(1,2,8)	330.3341	ARIMA(4,2,8)	330.4291	ARIMA(7,2,8)	333.6750
ARIMA(1,2,9)	329.1373	ARIMA(4,2,9)	332.3243	ARIMA(7,2,9)	334.6551
ARIMA(2,2,1)	323.6981	ARIMA(5,2,1)	328.3001	ARIMA(8,2,1)	330.4049
ARIMA(2,2,2)	321.6734	ARIMA(5,2,2)	330.3237	ARIMA(8,2,2)	329.4311
ARIMA(2,2,3)	322.4522	ARIMA(5,2,3)	328.4440	ARIMA(8,2,3)	330.4561
ARIMA(2,2,4)	323.9581	ARIMA(5,2,4)	327.6058	ARIMA(8,2,4)	332.4126
ARIMA(2,2,5)	325.9505	ARIMA(5,2,5)	328.0693	ARIMA(8,2,5)	331.6667
ARIMA(2,2,6)	327.9336	ARIMA(5,2,6)	332.9337	ARIMA(8,2,6)	332.0244
ARIMA(2,2,7)	330.3352	ARIMA(5,2,7)	329.3659	ARIMA(8,2,7)	333.5626
ARIMA(2,2,8)	328.4683	ARIMA(5,2,8)	333.3522	ARIMA(8,2,8)	334.6872
ARIMA(2,2,9)	328.9783	ARIMA(5,2,9)	332.9101	ARIMA(8,2,9)	336.0595

ตารางที่ 2.2 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(3,2,1)	325.6276	ARIMA(6,2,1)	330.0650
ARIMA(3,2,2)	322.9650	ARIMA(6,2,2)	332.0648
ARIMA(3,2,3)	325.9597	ARIMA(6,2,3)	329.9092
ARIMA(3,2,4)	324.4755	ARIMA(6,2,4)	329.5942
ARIMA(3,2,5)	327.8812	ARIMA(6,2,5)	332.8958
ARIMA(3,2,6)	328.9641	ARIMA(6,2,6)	334.4640
ARIMA(3,2,7)	330.9381	ARIMA(6,2,7)	330.9217
ARIMA(3,2,8)	329.8211	ARIMA(6,2,8)	332.8675
ARIMA(3,2,9)	330.3316	ARIMA(6,2,9)	336.2202

ตารางที่ 2.2 (ต่อ) แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

จากตารางที่ 2.2 เมื่อพิจารณาจากค่า AIC เนื่องจาก ARIMA(1,2,3) ให้ค่า AIC น้อยที่สุด คือ 320.6421 ดังนั้นแบบจำลองอนุกรมเวลาของตัวแบบสำหรับการจำลองที่เหมาะสมสำหรับข้อมูลจำนวนการเอาประกันภัยของประกันชีวิตแบบสะสมทรัพย์ คือ ARIMA(1,2,3)

ต่อไปจะแสดงการพยากรณ์จำนวนการเอาประกันภัยโดยใช้แบบจำลอง ARIMA(1,2,3)

พารามิเตอร์	ARIMA(1,2,3)
Ar1	-0.6589
Ma1	-0.3422
Ma2	-0.9558
Ma3	0.3864

ตารางที่ 2.3 แสดงพารามิเตอร์ของจำนวนการเอาประกันภัย

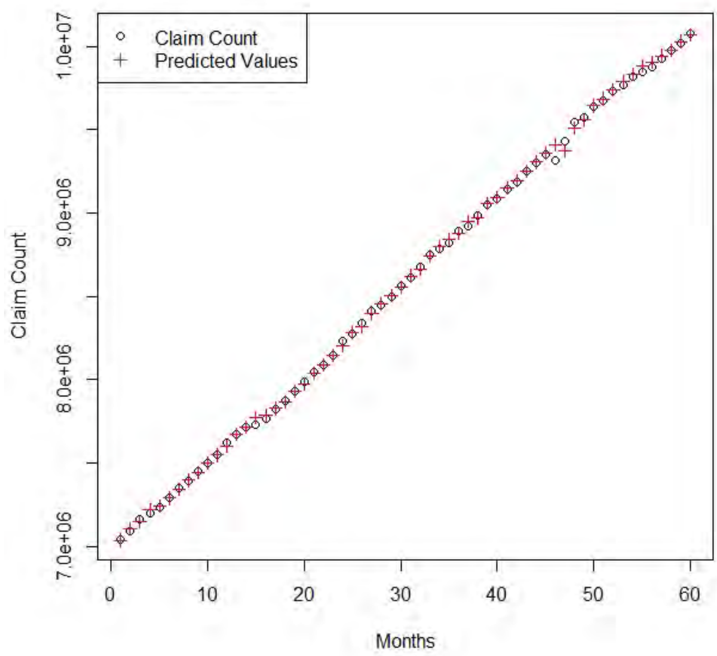
จากตารางที่ 2.3 จะได้ว่าแบบจำลอง ARIMA(1,2,3) เขียนค่าประมาณพารามิเตอร์จากการทำผลต่างอันดับ 1 ได้สมการต่อไปนี้

$$\Delta^2 y_t = -0.6589\Delta^2 y_{t-1} + 0.3422\varepsilon_{t-1} + 0.9558\varepsilon_{t-2} - 0.3864\varepsilon_{t-3} + \varepsilon_t$$

จัดรูปสมการได้ดังนี้

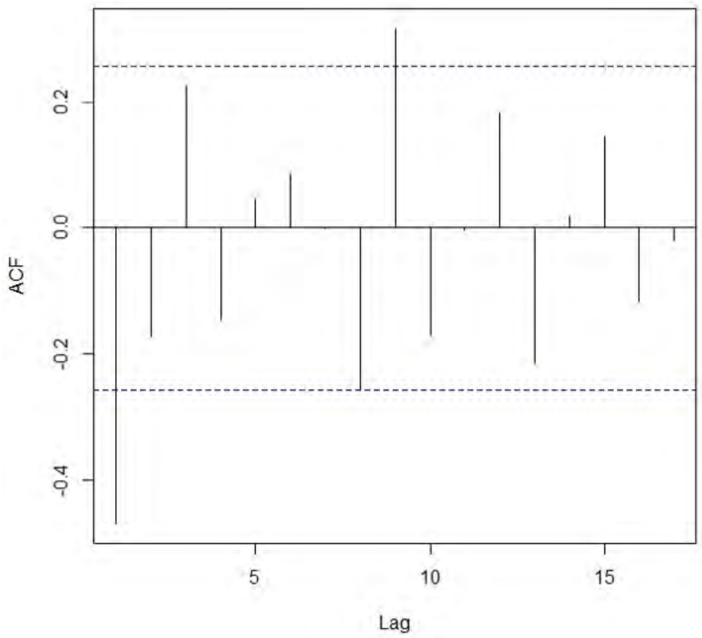
$$y_t = 1.3411y_{t-1} + 0.3178y_{t-2} - 0.6589y_{t-3} + 0.3422\varepsilon_{t-1} + 0.9558\varepsilon_{t-2}$$
$$-0.3864\varepsilon_{t-3} + \varepsilon_t$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ดังภาพที่ 2.4

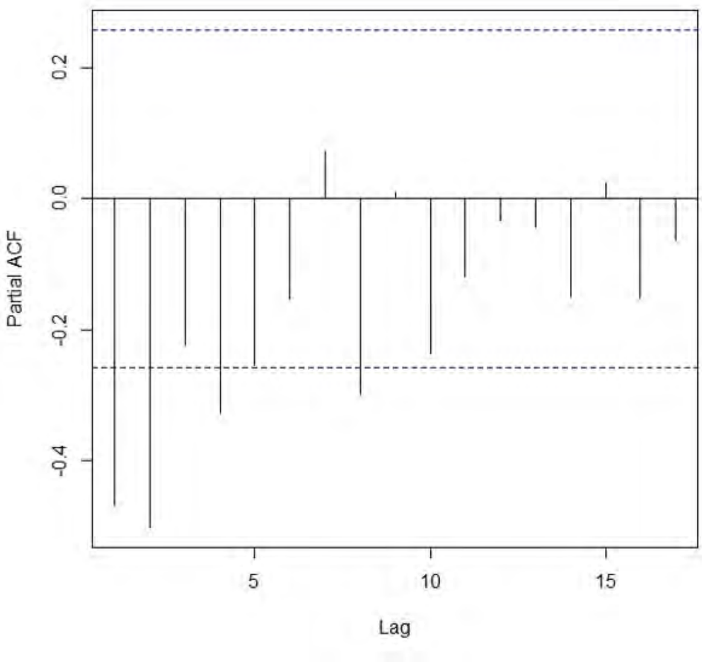


ภาพที่ 2.4 แสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับ
จำนวนการเอาประกันภัยจากแบบจำลอง

ต่อไปเราจะศึกษาหาตัวแบบสำหรับการแจกแจงปัวซงและตัวแบบสำหรับการแจกแจงทวินามลบ โดยนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาทดสอบความนิ่ง (Stationary) และแปลงข้อมูลให้มีความนิ่ง โดยใช้วิธีการหาผลต่าง (differencing) จะได้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งมีความนิ่ง เราจึงใช้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งในการคาดเดาลำดับของ Autoregressive (AR) และ Moving Average (MA) โดยพิจารณาค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) แสดงในภาพที่ 2.5 และภาพที่ 2.6



ภาพที่ 2.5 แสดงกราฟ ACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2



ภาพที่ 2.6 แสดงกราฟ PACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2

จากภาพที่ 2.5 และภาพที่ 2.6 พบว่าอนุกรมเวลาของผลต่างอันดับ 2 มีค่าสหสัมพันธ์ในตัวเอง (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,9 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,4,8 เนื่องจากข้อมูลผ่านการทำผลต่างอันดับ 2 ดังนั้นเมื่อแก้สมการแล้วจะได้ว่าต้องพิจารณาค่าสังเกตในอดีตอีกใน lag ที่ 9,10 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจะนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาพิจารณาหาแบบจำลองสำหรับการแจกแจงปัวซงและแบบจำลองสำหรับการแจกแจงทวินามลบ ตามลำดับ

พิจารณาแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซง

แบบจำลอง		ค่า AIC	แบบจำลอง		ค่า AIC
ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต		ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต	
1	1	110,744.7	9	9	4,285,882
2	1	79,421.88	10	9	3,809,727
4	1	921,861.6	1	1,9	27,777.68
8	1	3,509,396	2	1,9	304,089.3
9	1	3,499,044	1,2	9	382,314.7
10	1	3,801,666	**1,2**	**1,9**	**6,197.731**
1	9	27,514.76	1,2,4	1,9	1,671,028
2	9	374,546.3	1,2,8	1,9	2,012,117
4	9	1,320,548	1,2,9	1,9	1,906,304
8	9	4,927,445	1,2,10	1,9	2,015,269

ตารางที่ 2.4 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

จากตารางที่ 2.4 เมื่อพิจารณาค่า AIC จะได้ว่าแบบจำลองที่มีค่าสังเกตในอดีตเท่ากับ 1 และ 2 และค่าเฉลี่ยในอดีตเท่ากับ 1 และ 9 ให้ค่า AIC น้อยที่สุดคือ 6,197.731 ดังนั้นแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซงที่เหมาะสมสำหรับจำนวนการเอาประกันภัยโดยส่วนของข้อมูลในอดีตของตัวแปรที่สนใจ (ค่าสังเกตในอดีต) คือ ข้อมูลก่อนหน้า 1 หน่วยเวลา และข้อมูลก่อนหน้า 2 หน่วยเวลา ส่วนของค่าเฉลี่ยแบบมีเงื่อนไข (ค่าเฉลี่ยในอดีต) คือ ค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 1 หน่วยเวลา และค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 9 หน่วยเวลา

ต่อไปจะแสดงการพยากรณ์จำนวนการเอาประกันภัยโดยใช้แบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซงที่มีค่าสังเกตในอดีตเท่ากับ 1 และ 2 และค่าเฉลี่ยในอดีตเท่ากับ 1 และ 9

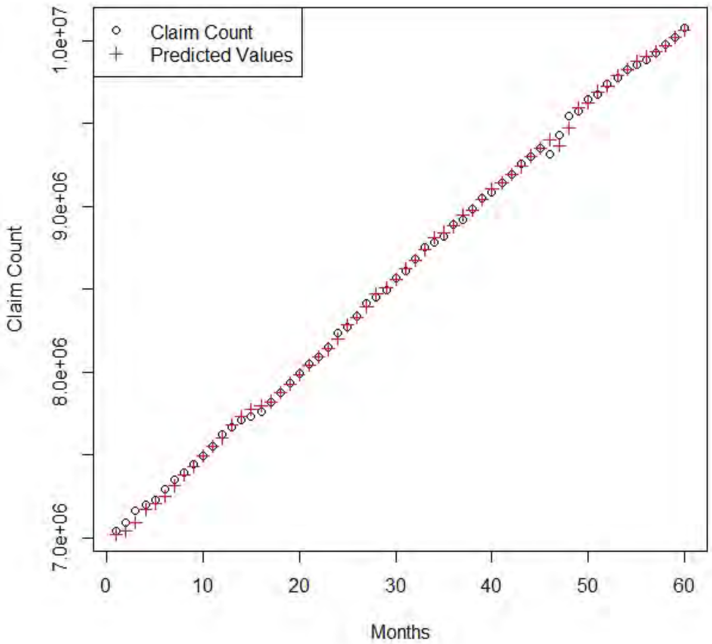
พารามิเตอร์	แบบจำลอง
Intercept	0.01579
Beta 1	0.99793
Beta 2	-0.56859
Alpha 1	0.62630
Alpha 9	-0.05664

ตารางที่ 2.5 แสดงพารามิเตอร์ของจำนวนการเอาประกันภัย

จากตารางที่ 2.5 จะได้ว่าแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซง ดังสมการต่อไปนี้

$$\log(\lambda_t) = 0.01579 + 0.99793Y_{t-1} - 0.56859Y_{t-2} + 0.62630\lambda_{t-1} - 0.05664\lambda_{t-9}$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ดังภาพที่ 2.7



ภาพที่ 2.7 แสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับ

จำนวนการเอาประกันภัยจากแบบจำลอง

พิจารณาแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบ

แบบจำลอง		ค่า AIC	แบบจำลอง		ค่า AIC
ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต		ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต	
1	1	1,587.164	9	9	1,807.561
2	1	1,567.408	10	9	1,800.492
4	1	1,717.266	1	1,9	1,501.309
8	1	1,795.138	2	1,9	1,650.184
9	1	1,794.915	1,2	9	1,665.552
10	1	1,799.724	**1,2**	**1,9**	**1,408.873**
1	9	1,501.453	1,2,4	1,9	1,756.344
2	9	1,662.11	1,2,8	1,9	1,767.373
4	9	1,737.803	1,2,9	1,9	1,764.114
8	9	1,815.933	1,2,10	1,9	1,767.465

ตารางที่ 2.6 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

จากตารางที่ 2.6 เมื่อพิจารณาค่า AIC จะได้ว่าแบบจำลองที่มีค่าสังเกตในอดีตเท่ากับ 1 และ 2 และค่าเฉลี่ยในอดีตเท่ากับ 1 และ 9 ให้ค่า AIC น้อยที่สุดคือ 1,408.873 ดังนั้นแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบที่เหมาะสมสำหรับจำนวนการเอาประกันภัยโดยส่วนของข้อมูลในอดีตของตัวแปรที่สนใจ (ค่าสังเกตในอดีต) คือ ข้อมูลก่อนหน้า 1 หน่วยเวลา และข้อมูลก่อนหน้า 2 หน่วยเวลา ส่วนของค่าเฉลี่ยแบบมีเงื่อนไข (ค่าเฉลี่ยในอดีต) คือ ค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 1 หน่วยเวลา และค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 9 หน่วยเวลา

ต่อไปจะแสดงการพยากรณ์จำนวนการเอาประกันภัยโดยใช้แบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบที่มีค่าสังเกตในอดีตเท่ากับ 1 และ 2 และค่าเฉลี่ยในอดีตเท่ากับ 1 และ 9

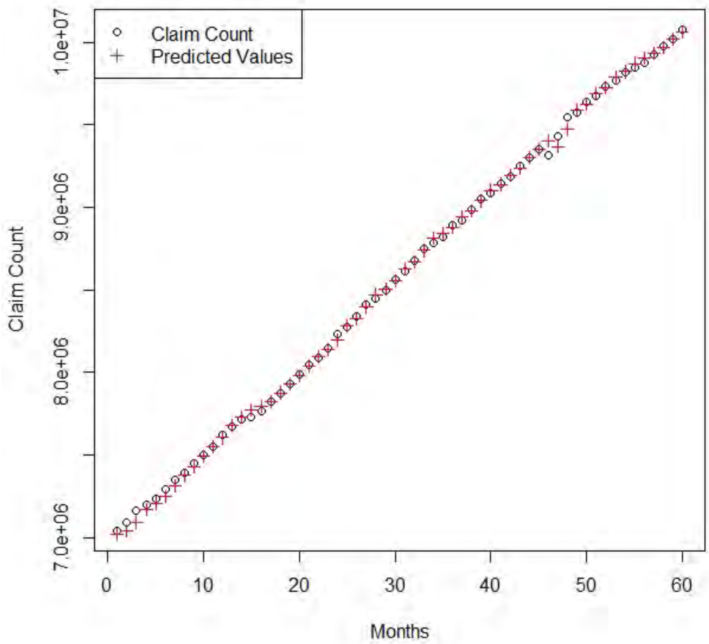
พารามิเตอร์	แบบจำลอง
Intercept	0.01579
Beta 1	0.99793
Beta 2	-0.56859
Alpha 1	0.62630
Alpha 9	-0.05664

ตารางที่ 2.7 แสดงพารามิเตอร์ของจำนวนการเอาประกันภัย

จากตารางที่ 2.7 จะได้ว่าแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบ ดังสมการต่อไปนี้

$$\log(\lambda_t) = 0.01579 + 0.99793Y_{t-1} - 0.56859Y_{t-2} + 0.62630\lambda_{t-1} - 0.05664\lambda_{t-9}$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ดังภาพที่ 2.8



ภาพที่ 2.8 แสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับ
จำนวนการเอาประกันภัยจากแบบจำลอง

เมื่อพิจารณาจากภาพการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองของทั้งสามแบบจำลองจะเห็นว่าแต่ละแบบจำลองมีความเหมาะสมในการพยากรณ์ เราจึงมาพิจารณาค่าคาดเคลื่อนกำลังสองเพื่อหาแบบจำลองที่เหมาะสมที่สุดเราจะเลือกแบบจำลองที่มีค่าคาดเคลื่อนกำลังสองต่ำที่สุด ผลการศึกษาแสดงดังตารางที่ 2.8 ต่อไปนี้

แบบจำลอง	ค่าคาดเคลื่อนกำลังสอง
ARIMA(1,2,3)	24,868,309,566
เชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซง ค่าสังเกตในอดีตเท่ากับ 1,2 ค่าเฉลี่ยในอดีตเท่ากับ 1,9	42,842,927,734
เชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบ ค่าสังเกตในอดีตเท่ากับ 1,2 ค่าเฉลี่ยในอดีตเท่ากับ 1,9	42,842,927,734

ตารางที่ 2.8 แสดงค่าคาดเคลื่อนกำลังสองของแบบจำลอง

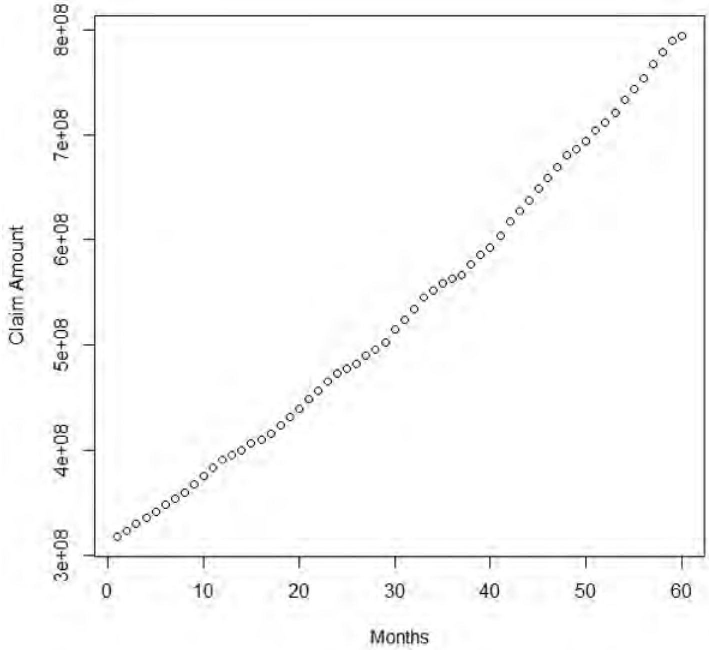
จากตารางที่ 2.8 เมื่อพิจารณาจากค่าคาดเคลื่อนกำลังสอง เนื่องจาก ARIMA(1,2,3) ให้ค่าคาดเคลื่อนกำลังสองน้อยที่สุด คือ 24,868,309,566 ดังนั้นแบบจำลองอนุกรมเวลาสำหรับการจำลองที่เหมาะสมสำหรับข้อมูลจำนวนการเอาประกันภัยของประกันชีวิตแบบสะสมทรัพย์ คือ ARIMA(1,2,3)

3. ประกันชีวิตแบบชั่วระยะเวลา

เดือน ปีพ.ศ.	จำนวนเอาประกันภัย (ราย)
ม.ค. 2552	758,491
ก.พ. 2552	769,517
มี.ค. 2552	773,990
เม.ย. 2552	781,922
พ.ค. 2552	799,178
มิ.ย. 2552	812,028
⋮	⋮
ก.ค. 2556	1,178,924
ส.ค. 2556	1,190,858
ก.ย. 2556	1,210,655
ต.ค. 2556	1,228,197
พ.ย. 2556	1,241,587
ธ.ค. 2556	1,257,550

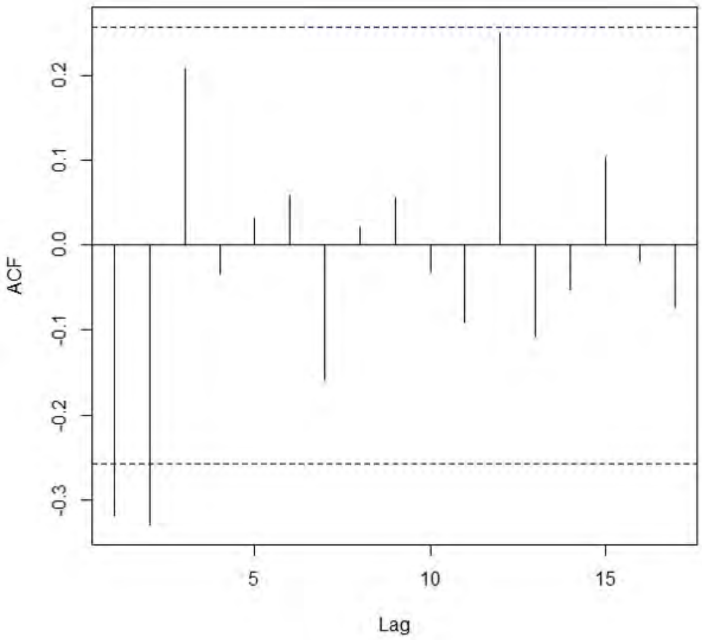
ตารางที่ 3.1 แสดงจำนวนเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม

ปี พ.ศ. 2556

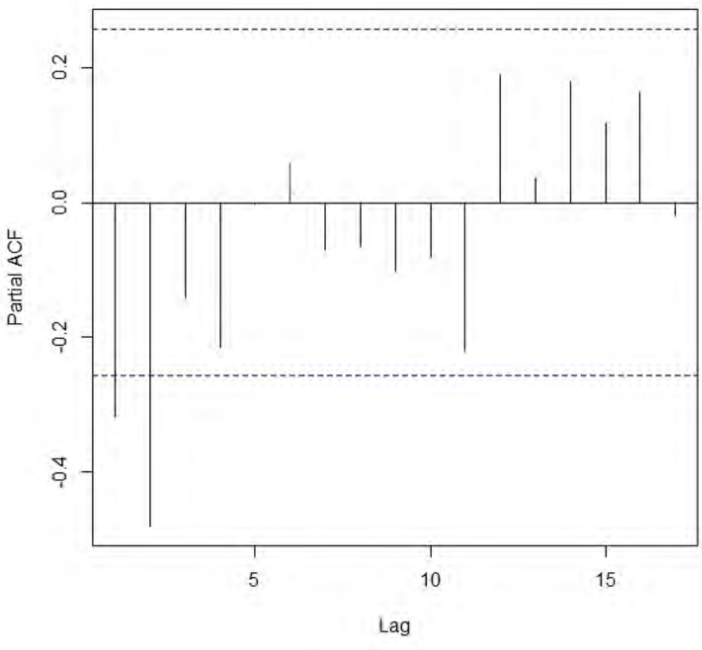


ภาพที่ 3.1 แสดงจำนวนการเอาประกันภัยของประกันชีวิตแบบเฉพาะกาลระหว่างปี พ.ศ. 2552-2556

ต่อไปเราจะนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือน ธันวาคม ปี พ.ศ. 2556 มาแปลงข้อมูลโดยการหารากที่สองของข้อมูล จากนั้นนำข้อมูลที่แปลงแล้วมาทดสอบความนิ่ง (Stationary) และแปลงข้อมูลให้มีความนิ่ง โดยใช้วิธีการหาผลต่าง (differencing) จะได้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งมีความนิ่ง เราจึงใช้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งในการคาดเดาลำดับของ Autoregressive (AR) และ Moving Average (MA) โดยพิจารณาค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) แสดงในภาพที่ 3.2 และภาพที่ 3.3



ภาพที่ 3.2 แสดงกราฟ ACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2



ภาพที่ 3.3 แสดงกราฟ PACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2

จากภาพที่ 3.2 และภาพที่ 3.3 พบว่าอนุกรมเวลาของผลต่างอันดับ 2 มีค่าสหสัมพันธ์ในตัวเอง (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจะนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือน มกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาพิจารณา ARIMA(p,d,q) เมื่อ p∈{1,2},q=∈{1,2}, และ d=2 ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC ต่ำที่สุด ผลการศึกษาแสดงดังตารางที่ 2.2 ต่อไปนี้

แบบจำลอง	ค่า AIC
ARIMA(1,2,1)	232.5448
ARIMA(1,2,2)	231.5830
ARIMA(2,2,1)	**230.5338**
ARIMA(2,2,2)	231.6532

ตารางที่ 3.2 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

จากตารางที่ 3.2 เมื่อพิจารณาจากค่า AIC เนื่องจาก ARIMA(2,2,1) ให้ค่า AIC น้อยที่สุด คือ 230.5338 ดังนั้นแบบจำลองอนุกรมเวลาของตัวแบบสำหรับการจำลองที่เหมาะสมสำหรับข้อมูลจำนวนการเอาประกันภัยของประกันชีวิตแบบสะสมชั่วระยะเวลา คือ ARIMA(2,2,1)

ต่อไปจะแสดงการพยากรณ์จำนวนการเอาประกันภัยโดยใช้แบบจำลอง ARIMA(2,2,1)

พารามิเตอร์	ARIMA(2,2,1)
Ar1	-0.1621
Ar2	-0.3902
Ma1	-0.4209

ตารางที่ 3.3 แสดงพารามิเตอร์ของจำนวนการเอาประกันภัย

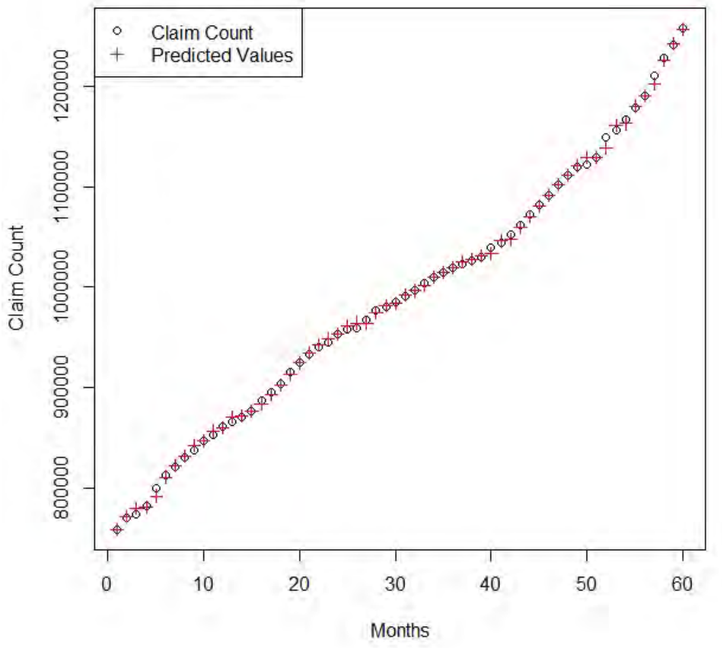
จากตารางที่ 3.3 จะได้ว่าแบบจำลอง ARIMA(2,2,1) เขียนค่าประมาณพารามิเตอร์จากการทำผลต่างอันดับ 2 ได้สมการต่อไปนี้

$$\Delta^2 y_t = -0.1621\Delta^2 y_{t-1} - 0.3902\Delta^2 y_{t-2} + 0.4209\varepsilon_{t-1} + \varepsilon_t$$

จัดรูปสมการได้ดังนี้

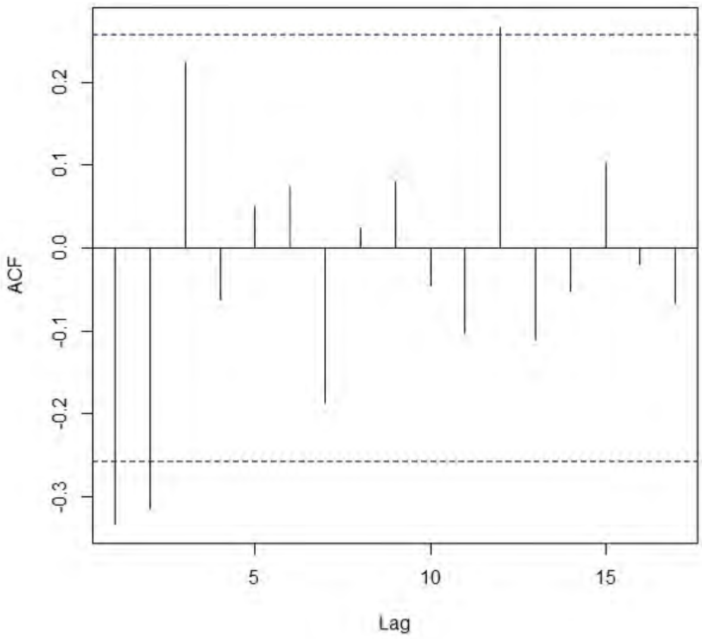
$$y_t = 1.8379y_{t-1} - 1.066y_{t-2} + 0.6183y_{t-3} - 0.3902y_{t-4} + 0.4209\varepsilon_{t-1} + \varepsilon_t$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ดังภาพที่ 3.4

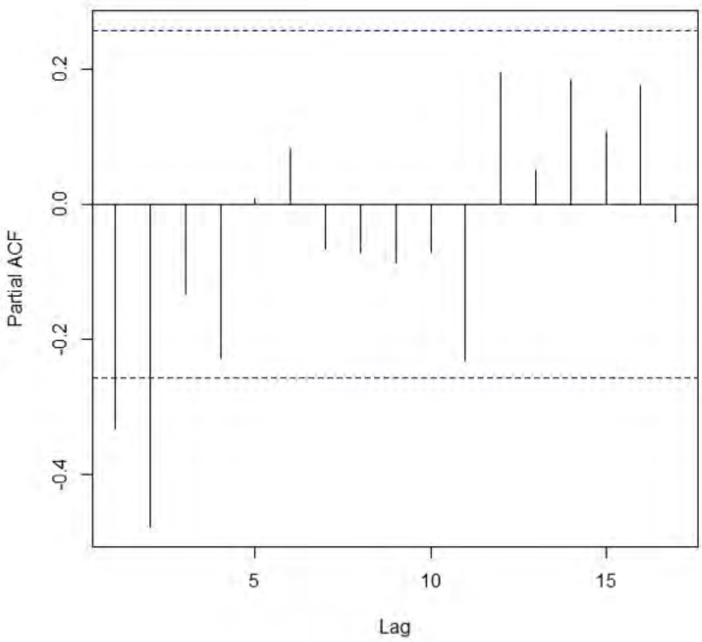


ภาพที่ 3.4 แสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับ
จำนวนการเอาประกันภัยจากแบบจำลอง

ต่อไปเราจะศึกษาหาตัวแบบสำหรับการแจกแจงปัวซงและตัวแบบสำหรับการแจกแจงทวินามลบ โดยนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาทดสอบความนิ่ง (Stationary) และแปลงข้อมูลให้มีความนิ่ง โดยใช้วิธีการหาผลต่าง (differencing) จะได้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งมีความนิ่ง เราจึงใช้ข้อมูลที่ผ่านการหาผลต่าง 2 ครั้งในการคาดเดาลำดับของ Autoregressive (AR) และ Moving Average (MA) โดยพิจารณาค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) แสดงในภาพที่ 3.5 และภาพที่ 3.6



ภาพที่ 3.5 แสดงกราฟ ACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2



ภาพที่ 3.6 แสดงกราฟ PACF ของจำนวนการเอาประกันภัย ณ lag ที่ 1,2,...,15 โดยการหาผลต่างของข้อมูลครั้งที่ 2

จากภาพที่ 3.5 และภาพที่ 3.6 พบว่าอนุกรมเวลาของผลต่างอันดับ 2 มีค่าสหสัมพันธ์ในตัวเอง (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,12 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2 เนื่องจากข้อมูลผ่านการทำผลต่างอันดับ 2 ดังนั้นเมื่อแก้สมการแล้วจะได้ว่าต้องพิจารณาค่าสังเกตในอดีตอีกใน lag ที่ 3,4 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจะนำข้อมูลจำนวนการเอาประกันภัยตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 มาพิจารณาหาแบบจำลองสำหรับการแจกแจงปัวซงและแบบจำลองสำหรับการแจกแจงทวินามลบ ตามลำดับ

พิจารณาแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซง

แบบจำลอง		ค่า AIC	แบบจำลอง		ค่า AIC
ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต		ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต	
1	1	432,569.3	2	12	69,102.89
2	1	496,752.9	3	12	227,565.1
3	1	560,166	4	12	335,109.4
4	1	619,173.7	1,2	12	20,832.73
1	2	441,217.6	1,2	1,12	177,596.7
2	2	505,705.5	1,2	2,12	5,149.327
3	2	567,566.3	**1,2**	**1,2,12**	**1,928.312**
4	2	625,702.4	1	1,2,12	2,727.491
1	12	65,927.08	2	1,2,12	13,434.04

ตารางที่ 3.4 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนการเอาประกันภัย

จากตารางที่ 3.4 เมื่อพิจารณาค่า AIC จะได้ว่าแบบจำลองที่มีค่าสังเกตในอดีตเท่ากับ 1 และ 2 และค่าเฉลี่ยในอดีตเท่ากับ 1,2 และ 12 ให้ค่า AIC น้อยที่สุดคือ 1,928.312 ดังนั้นแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซงที่เหมาะสมสำหรับจำนวนการเอาประกันภัยโดยส่วนของข้อมูลในอดีตของตัวแปรที่สนใจ (ค่าสังเกตในอดีต) คือ ข้อมูลก่อนหน้า 1 หน่วยเวลา และข้อมูลก่อนหน้า 2 หน่วยเวลา ส่วนของค่าเฉลี่ยแบบมีเงื่อนไข (ค่าเฉลี่ยในอดีต) คือ ค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 1 หน่วยเวลา ค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 2 หน่วยเวลา และค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 12 หน่วยเวลา

ต่อไปจะแสดงการพยากรณ์จำนวนการเอาประกันภัยโดยใช้แบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซงที่มีค่าสังเกตในอดีตเท่ากับ 1 และ 2 และค่าเฉลี่ยในอดีตเท่ากับ 1,2 และ 12

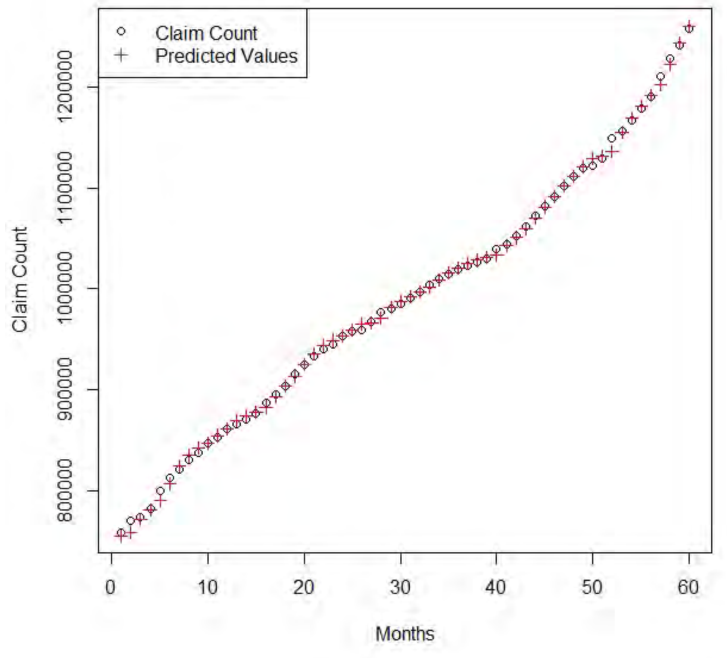
พารามิเตอร์	แบบจำลอง
Intercept	0.001482
Beta 1	0.986558
Beta 2	-0.400425
Alpha 1	0.891277
Alpha 2	-0.463767
Alpha 12	-0.013752

ตารางที่ 3.5 แสดงพารามิเตอร์ของจำนวนการเอาประกันภัย

จากตารางที่ 3.5 จะได้ว่าแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซง ดังสมการต่อไปนี้

$$\log(\lambda_t) = 0.001482 + 0.986558Y_{t-1} - 0.400425Y_{t-2} + 0.891277\lambda_{t-1} - 0.463767\lambda_{t-2} - 0.013752\lambda_{t-12}$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ดังภาพที่ 3.7



ภาพที่ 3.7 แสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับ

จำนวนการเอาประกันภัยจากแบบจำลอง

พิจารณาแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบ

แบบจำลอง		ค่า AIC	แบบจำลอง		ค่า AIC
ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต		ค่าสังเกตในอดีต	ค่าเฉลี่ยในอดีต	
1	1	1,539.233	2	12	1,430.511
2	1	1,547.451	3	12	1,502.3
3	1	1,554.607	4	12	1,525.406
4	1	1,560.585	1,2	12	1,355.501
1	2	1,540.418	1,2	1,12	1,489.395
2	2	1,548.522	1,2	2,12	1,271.179
3	2	1,555.394	**1,2**	**1,2,12**	**1,183.078**
4	2	1,561.215	1	1,2,12	1,218.679
1	12	1,428.792	2	1,2,12	1,332.837

ตารางที่ 3.6 แสดงอันดับของแบบจำลองและค่า AIC ของจำนวนเอาประกันภัย

จากตารางที่ 3.6 เมื่อพิจารณาค่า AIC จะได้ว่าแบบจำลองที่มีค่าสังเกตในอดีตเท่ากับ 1 และ 2 และค่าเฉลี่ยในอดีตเท่ากับ 1,2 และ 12 ให้ค่า AIC น้อยที่สุดคือ 1,183.078 ดังนั้นแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบที่เหมาะสมสำหรับจำนวนการเอาประกันภัยโดยส่วนของข้อมูลในอดีตของตัวแปรที่สนใจ (ค่าสังเกตในอดีต) คือ ข้อมูลก่อนหน้า 1 หน่วยเวลา และข้อมูลก่อนหน้า 2 หน่วยเวลา ส่วนของค่าเฉลี่ยแบบมีเงื่อนไข (ค่าเฉลี่ยในอดีต) คือ ค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 1 หน่วยเวลา และค่าเฉลี่ยแบบมีเงื่อนไขก่อนหน้า 9 หน่วยเวลา

ต่อไปจะแสดงการพยากรณ์จำนวนการเอาประกันภัยโดยใช้แบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามที่มีค่าสังเกตในอดีตเท่ากับ 1 และ 2 และค่าเฉลี่ยในอดีตเท่ากับ 1,2 และ 12

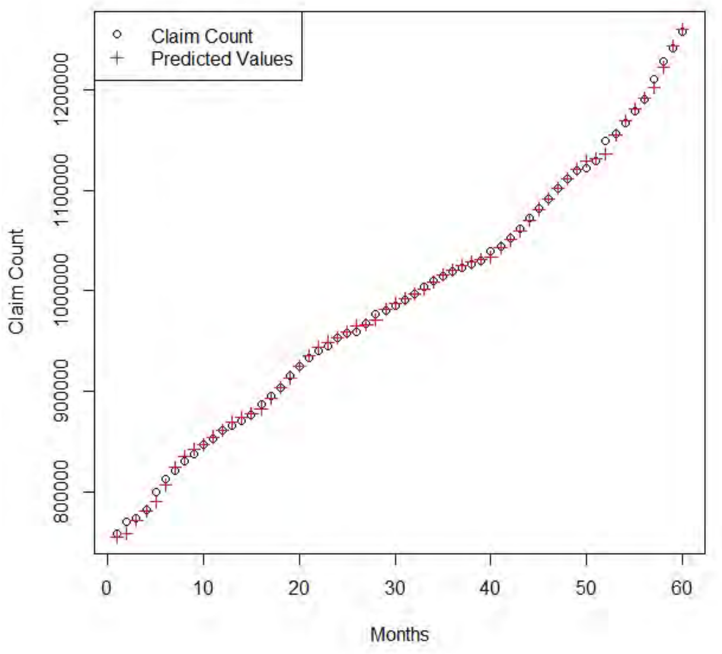
พารามิเตอร์	แบบจำลอง
Intercept	0.001482
Beta 1	0.986558
Beta 2	-0.400425
Alpha 1	0.891277
Alpha 2	-0.463767
Alpha 12	-0.013752

ตารางที่ 3.7 แสดงพารามิเตอร์ของจำนวนการเอาประกันภัย

จากตารางที่ 3.7 จะได้ว่าแบบจำลองเชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบ ดังสมการต่อไปนี้

$$\begin{aligned}\log(\lambda_t) &= 0.001482 + 0.986558Y_{t-1} - 0.400425Y_{t-2} + 0.891277\lambda_{t-1} \\ &\quad - 0.463767\lambda_{t-2} - 0.013752\lambda_{t-12}\end{aligned}$$

ต่อไปเราจะแสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยจากแบบจำลองตั้งแต่เดือนมกราคม ปี พ.ศ. 2552 ถึงเดือนธันวาคม ปี พ.ศ. 2556 ดังภาพที่ 3.8



ภาพที่ 3.8 แสดงการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับ
จำนวนการเอาประกันภัยจากแบบจำลอง

เมื่อพิจารณาจากภาพการเปรียบเทียบจำนวนการเอาประกันภัยจริงกับจำนวนการเอาประกันภัยพยากรณ์ของทั้งสามแบบจำลองจะเห็นว่าแต่ละแบบจำลองมีความเหมาะสมในการพยากรณ์ เราจึงมาพิจารณาค่าคาดเคลื่อนกำลังสองเพื่อหาแบบจำลองที่เหมาะสมที่สุดเราจะเลือกแบบจำลองที่มีค่าคาดเคลื่อนกำลังสองต่ำที่สุด ผลการศึกษาแสดงดังตารางที่ 3.8 ต่อไปนี้

แบบจำลอง	ค่าคาดเคลื่อนกำลังสอง
ARIMA(2,2,1)	651,611,370
เชิงเส้นลอการิทึมสำหรับการแจกแจงปัวซง ค่าสังเกตในอดีตเท่ากับ 1,2 ค่าเฉลี่ยในอดีตเท่ากับ 1,2,12	934,206,887
เชิงเส้นลอการิทึมสำหรับการแจกแจงทวินามลบ ค่าสังเกตในอดีตเท่ากับ 1,2 ค่าเฉลี่ยในอดีตเท่ากับ 1,2,12	934,206,887

ตารางที่ 3.8 แสดงค่าคาดเคลื่อนกำลังสองของแบบจำลอง

จากตารางที่ 3.8 เมื่อพิจารณาจากค่าคาดเคลื่อนกำลังสอง เนื่องจาก ARIMA(2,2,1) ให้ค่าคาดเคลื่อนกำลังสองน้อยที่สุด คือ 651,611,370 ดังนั้นแบบจำลองอนุกรมเวลาสำหรับการจำลองที่เหมาะสมสำหรับข้อมูลจำนวนการเอาประกันภัยของประกันชีวิตแบบชั่วระยะเวลา คือ ARIMA(2,2,1)

3.3 สรุปผลการทำงาน

จากการศึกษาหาแบบจำลองอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมที่สุดสำหรับข้อมูลจำนวนการเอาประกันภัยและจำนวนเงินเอาประกันภัยของบริษัทประกันชีวิตในประเทศไทย ในกลุ่มข้อมูลของประกันชีวิต 3 ประเภท ได้แก่ ประกันชีวิตแบบตลอดชีพ ประกันชีวิตแบบสะสมทรัพย์ และประกันชีวิตแบบชั่วระยะเวลา ในช่วงปี พ.ศ. 2552-2556 ได้ผลการศึกษาแสดงดังตารางต่อไปนี้

จำนวนเงินเอาประกันภัย	แบบจำลองที่เหมาะสม
1. ประกันชีวิตแบบตลอดชีพ	ARIMA(4,2,1)
2. ประกันชีวิตแบบสะสมทรัพย์	ARIMA(2,2,9)
3. ประกันชีวิตแบบชั่วระยะเวลา	ARIMA(1,3,8)
จำนวนการเอาประกันภัย	แบบจำลองที่เหมาะสม
1. ประกันชีวิตแบบตลอดชีพ	ARIMA(1,2,2)
2. ประกันชีวิตแบบสะสมทรัพย์	ARIMA(1,2,3)
3. ประกันชีวิตแบบชั่วระยะเวลา	ARIMA(2,2,1)

ตาราง แสดงแบบจำลองที่เหมาะสมสำหรับข้อมูลจำนวนเงินเอาประกันภัยและจำนวนการเอาประกันภัย

รายการอ้างอิง

[1] Agosto A, Cavaliere G, Kristensen D, Rahbek A (2015). "Modeling Corporate Defaults: Poisson Autoregressions with Exogenous Covariates (PARX)." CREATES Research Paper, School of Economics and Management, University of Aarhus.

[2] Fokianos, K., Rahbek, A. and Tjøstheim, D. (2009). Poisson Autoregression *Journal of the American Statistical Association*, 1430-1439.

[3] Liboschik, T., Fokianos, K. and Fried, R. (2017). tscount: An R package for analysis of count time series following generalized linear models. *Journal of Statistical Software* 82(5), 1–51.

[4] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienns, Austria. ISBN3-900051-07-0, URL http://www.R-project.org.

[5] ภูมิฐาน รังคกูลนุวัฒน์ (2556) *การวิเคราะห์อนุกรมเวลาสำหรับเศรษฐศาสตร์และธุรกิจ.* กรุงเทพมหานคร:สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ภาคผนวก

ภาคผนวก ก

แบบเสนอหัวข้อโครงงาน รายวิชา 2301399 Project Proposal

ภาคต้นปีการศึกษา 2562

ชื่อโครงงาน (ภาษาไทย)	การศึกษาข้อมูลประกันชีวิตของประเทศไทยโดยใช้แบบจำลองอนุกรมเวลา
ชื่อโครงงาน (ภาษาอังกฤษ)	A study of insurance data of Thailand using time series models
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. จิราพรรณ สุนทรโชติ
ผู้ดำเนินการ	นายศราวุฒิ ทองเชื้อ เลขประจำตัวนิสิต 5933549023
	สาขาวิชา คณิตศาสตร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
	คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

หลักการและเหตุผล

ปัจจุบันการทำประกันชีวิตได้รับความนิยมเป็นอย่างมาก เนื่องจากสามารถให้ความคุ้มครองต่อบุคคลและครอบครัว ทำให้บริษัทประกันภัยนั้นจำเป็นต้องหาแนวทางในการคาดการณ์จำนวนการเอาประกันภัย ซึ่งอนุกรมเวลาและอนุกรมเวลาเชิงการนับเป็นวิธีการทางสถิติหนึ่งที่นิยมในการนำมาใช้พยากรณ์ข้อมูล

อนุกรมเวลา คือ อนุกรมของข้อมูลที่มีการเก็บรวบรวมข้อมูลของตัวแปรตามลำดับเวลา เช่น ราคาหุ้นรายวันของบริษัท และปริมาณน้ำรายเดือนในเขื่อน แบบจำลองอนุกรมเวลาที่ได้รับความนิยม คือ แบบจำลอง ARIMA (Autoregressive Integrated Moving Average) ซึ่งการวิเคราะห์อนุกรมเวลาเป็นวิธีการทางสถิติที่นิยมในการนำมาใช้พยากรณ์ข้อมูลที่เกิดขึ้นในอนาคต

อนุกรมเวลาเชิงการนับ คือ ข้อมูลจำนวนของเหตุการณ์ต่อเวลาในช่วงเวลาที่ถูกสังเกตโดยอาจเก็บข้อมูลเป็นรายวัน รายเดือน หรือรายปีขึ้นอยู่กับวัตถุประสงค์ในการนำไปใช้ อนุกรมเวลาเชิงการนับมักพบเห็นได้ในสถานการณ์ต่างๆ เช่น จำนวนการเอาประกันภัยรายเดือนของบริษัทประกันภัย และจำนวนผู้ป่วยรายสัปดาห์ที่เข้ารับการรักษาในโรงพยาบาล ซึ่งอนุกรมเวลาเชิงการนับมีประโยชน์อย่างมากในการวิเคราะห์และวางแผนตัดสินใจทางธุรกิจ หรือคาดการณ์ขั้นแผนงานให้มีความผิดพลาดน้อยที่สุด โดยใช้ข้อมูลในอดีตเป็นพื้นฐานสำหรับการพยากรณ์ข้อมูลในอนาคต

แบบจำลองอนุกรมเวลาเชิงการนับได้รับความสนใจอย่างกว้างขวางในบรรดานักสถิติในการสร้างแบบจำลองสำหรับสถานการณ์ต่างๆ เช่น Fokianos และคณะ (2009) ได้นำเสนอแบบจำลองอนุกรมเวลาเชิงการนับสำหรับตัวแบบการ์ชแบบปัวซง และศึกษาบทประยุกต์สำหรับจำนวนการซื้อขายหุ้นในตลาดหลักทรัพย์ และ Agosto และคณะ (2015) ได้นำเสนอแบบจำลองอนุกรมเวลาเชิงการนับแบบปัวซงที่มีตัวแปรร่วม และศึกษาบทประยุกต์ในการนับจำนวนล้มละลายของบริษัทในประเทศสหรัฐอเมริกาช่วงปีค.ศ. 1982-2011 ต่อมาในปี 2017 Liboschik และคณะ ได้สร้างชุดคำสั่ง tscount ในโปรแกรม R สำหรับแบบจำลองอนุกรมเวลาเชิงการนับ ซึ่งทำให้มีความสะดวกมากขึ้นในการใช้งานแบบจำลองอนุกรมเวลาเชิงการนับ

โครงงานนี้ เรามีความสนใจในการศึกษาอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมสำหรับข้อมูลประกันชีวิตของบริษัทประกันชีวิตในประเทศไทยในช่วงปีพ.ศ. 2552-2556 โดยใช้แบบจำลองและชุดคำสั่งในโปรแกรม R

วัตถุประสงค์

เราจะศึกษาหาแบบจำลองอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมที่สุดสำหรับข้อมูลประกันชีวิตของบริษัทประกันชีวิตในประเทศไทย

ขอบเขตของโครงงาน

ในโครงงานนี้ เราจะศึกษาอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมสำหรับข้อมูลประกันชีวิตของบริษัทประกันชีวิตในประเทศไทยโดยศึกษาผ่านชุดคำสั่งในโปรแกรม R

ข้อมูลที่ใช้ในการศึกษานี้คือข้อมูลจำนวนเงินเอาประกันภัย และจำนวนการเอาประกันภัย ในช่วง พ.ศ. 2552-2556 จากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย

วิธีการดำเนินงาน

1. ศึกษาความรู้พื้นฐานเกี่ยวกับการวิเคราะห์อนุกรมเวลาและอนุกรมเวลาเชิงการนับ
2. ศึกษาการใช้โปรแกรม R
3. ประยุกต์ใช้ชุดคำสั่งในโปรแกรม R สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาและอนุกรมเวลาเชิงการนับ
4. ศึกษาหาแบบจำลองอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมสำหรับข้อมูลประกันชีวิตของบริษัทประกันชีวิตในประเทศไทย
5. สรุปผลและเขียนรายงาน

ขั้นตอนการดำเนินการ	เดือน / ปีการศึกษา 2562										
	มิ.ย	ก.ค.	ส.ค.	ก.ย	ต.ค	พ.ย	ธ.ค	ม.ค	ก.พ	มี.ค.	เม.ย
1. ศึกษาความรู้พื้นฐานเกี่ยวกับการวิเคราะห์อนุกรมเวลาและอนุกรมเวลาเชิงการนับ											
2. ศึกษาการใช้โปรแกรม R											
3. ศึกษาการประยุกต์ใช้ชุดคำสั่งในโปรแกรม R สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาและอนุกรมเวลาเชิงการนับ											
4. ศึกษาหาแบบจำลองอนุกรมเวลาและอนุกรมเวลาเชิงการนับที่เหมาะสมสำหรับข้อมูลประกันชีวิตของบริษัทประกันชีวิตในประเทศไทย											
5. สรุปผลและเขียนรายงาน											

ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ที่คาดว่าจะได้รับต่อผู้ดำเนินงาน

1. มีความรู้เกี่ยวกับการวิเคราะห์อนุกรมเวลาและอนุกรมเวลาเชิงการนับ
2. มีความรู้เกี่ยวกับกรมธรรม์ประกันภัยของบริษัทประกันภัยในประเทศไทย
3. มีความรู้ในการใช้โปรแกรม R
4. มีความรู้ในการประยุกต์ใช้ชุดคำสั่งในโปรแกรม R

ประโยชน์ที่ได้จากโครงงานที่พัฒนาขึ้น

1. เพื่อใช้เป็นความรู้ต่อยอดงานวิจัย
2. สามารถคาดการณ์จำนวนเงินเอาประกันชีวิตและจำนวนการเอาประกันชีวิตของบริษัทประกันชีวิตในประเทศไทยได้

อุปกรณ์และเครื่องมือที่ใช้

ฮาร์ดแวร์

1. เครื่องคอมพิวเตอร์
2. เครื่องพิมพ์

ซอร์ฟแวร์

1. โปรแกรมงานเอกสาร Microsoft World
2. โปรแกรมงานเอกสาร Microsoft Excel
3. โปรแกรม Adobe PDF
4. โปรแกรม R

บรรณานุกรม

[1] Agosto A, Cavaliere G, Kristensen D, Rahbek A (2015). "Modeling Corporate Defaults: Poisson Autoregressions with Exogenous Covariates (PARX)." CREATES Research Paper, School of Economics and Management, University of Aarhus.

[2] Fokianos, K., Rahbek, A. and Tjøstheim, D. (2009). Poisson Autoregression *Journal of the American Statistical Association*, 1430-1439.

[3] Liboschik, T., Fokianos, K. and Fried, R. (2017). tscount: An R package for analysis of count time series following generalized linear models. *Journal of Statistical Software* 82(5), 1–51.

[4] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienns, Austria. ISBN3-900051-07-0,
URL http://www.R-project.org.

[5] ภูมิฐาน รังคกูลนุวัฒน์ (2556) *การวิเคราะห์อนุกรมเวลาสำหรับเศรษฐศาสตร์และธุรกิจ.* กรุงเทพมหานคร:สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ประวัติผู้เขียน

นายศราวุฒิ ทองเชื้อ

รหัสนิสิต 5933549023

สาขาวิชา คณิตศาสตร์

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย