โครงการ

การเรียนการสอนเพื่อเสริมประสบการณ์

ชื่อโครงการ การศึกษามูลค่าการเอาประกันวินาศภัยภัยโดยใช้แบบจำลองอนุกรมเวลา
A study of Non-life insurance using time series models

ชื่อนิสิต นายเจษฎากร น้อยท่าช้าง 593 35082 23

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์
สาขาวิชา คณิตศาสตร์

ปีการศึกษา 2562

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การศึกษามูลค่าการเอาประกันวินาศภัยภัยโดยใช้แบบจำลองอนุกรมเวลา

นายเจษฎากร น้อยท่าช้าง

โครงงานนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2562

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A study of Non-life insurance using time series models

Jedsadakorn Noytachang

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

Copyright of Chulalongkorn University

หัวข้อโครงงาน	การศึกษามูลค่าการเอาประกันวินาศภัยภัยโดยใช้แบบจำลองอนุกรมเวลา
โดย	นายเจษฎากร น้อยท่าช้าง
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. สำรวม บัวประดิษฐ์)

..... กรรมการ

(รองศาสตราจารย์ ดร. พันทิพา ทิพย์วิวัฒน์พจนา)

นายเจษฎากร น้อยท่าช้าง: การศึกษามูลค่าการเอาประกันวินาศภัยภัยโดยใช้แบบจำลองอนุกรมเวลา (A study of Non-life insurance using time series models)

อ.ที่ปรึกษาโครงงานหลัก: ผู้ช่วยศาสตราจารย์ ดร. จิราพรรณ สุนทรโชติ, 107 หน้า.

ในโครงงานนี้ เรามีความสนใจในการศึกษาการประยุกต์ใช้อนุกรมเวลาเพื่อใช้ในการศึกษาแนวโน้มของข้อมูลของประกันวินาศภัยของประเทศไทยโดยแบ่งออกเป็น 2 ประเด็นที่สนใจคือ (1) มูลค่าการเอาประกันวินาศภัย และ (2) จำนวนกรมธรรม์ของประกันวินาศภัย ซึ่งประกันวินาศภัยสามารถแบ่งได้ออกเป็น 3 ประเภท คือ การประกันอัคคีภัย (Fire insurance) การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance) และ การประกันภัยรถยนต์ (Automobile insurance) โดยข้อมูลที่ใช้คือ มูลค่าการเอาประกันวินาศภัยและจำนวนกรมธรรม์ของประกันวินาศภัย ในช่วงเดือนมกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 จากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) การวิเคราะห์ในโครงงานนี้ศึกษาผ่านโปรแกรม R

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์ ลายมือชื่อนิสิต เจษฎากร

สาขาวิชา คณิตศาสตร์ ลายมือชื่อ อ.ที่ปรึกษาโครงงานหลัก

ปีการศึกษา 2562

5933508223 : MAJOR MATHEMATICS
KEYWORDS : TIME SERIES, COUNT TIME SERIES, INSURANCE CLAIM AMOUNT.

JEDSADAKORN NOYTACHANG: A STUDY OF NON-LIFE INSURANCE USING TIME SERIES MODEL. ADVISOR: ASST. JIRAPHAN SUNTORNCHOST, Ph.D., 107 pp.

This project aims to study on Non-life insurance data using time series models in two aspects (1) study time series models for the claim amount of Non-life insurance (2) study time series models for the number of Non-life insurance policies. The Non-life insurance can be classified into three types which are Fire insurance, Marine and Transportation insurance and Automobile insurance. The data used in this study are monthly data from January 2012 to December 2016 obtained from the Office of Insurance Commission. All analyses are done by using R software.

Department: Mathematics and Computer Science Student's Signature Jedsadakorn

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

Academic Year : 2019

กิตติกรรมประกาศ

โครงงานเรื่อง การศึกษามูลค่าการเอาประกันวินาศภัยภัยโดยใช้แบบจำลองอนุกรมเวลา สำเร็จลุล่วงไปได้ด้วยดี เพราะได้รับการอนุเคราะห์และการช่วยเหลือจาก ผู้ช่วยศาสตราจารย์ ดร.จิราพรรณ สุนทรโชติ อาจารย์ที่ปรึกษาโครงงาน ซึ่งได้ให้ความรู้ คำปรึกษา คำแนะนำและแนวทางต่างๆ ซึ่งเป็นประโยชน์อย่างมากต่อการทำโครงงาน นอกจากนี้ยังช่วยแก้ไขปัญหาที่เกิดขึ้นระหว่างการดำเนินการทำโครงงาน ด้วยความเอาใจใส่อย่างใกล้ชิดตั้งแต่ต้นจนเสร็จสมบูรณ์ ทางผู้ดำเนินโครงงานจึงขอขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. สำรวม บัวประดิษฐ์ และรองศาสตราจารย์ ดร. พันทิพา ทิพย์วิวัฒน์พจนา ที่ให้ความกรุณาเป็นกรรมการในการนำเสนอโครงการ ให้คำแนะนำและเสนอแนะต่างๆ เพื่อนำไปปรับปรุงและแก้ไขโครงงานให้ดียิ่งขึ้น

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....ง

บทคัดย่อภาษาอังกฤษ.....จ

กิตติกรรมประกาศ.....ฉ

สารบัญ.....ช

สารบัญตาราง.....ญ

สารบัญภาพ.....ฐ

บทที่ 1 บทนำ.....1

1.1 ความเป็นมาและเหตุผลการวิจัย.....1

1.2 วัตถุประสงค์ของการวิจัย.....2

1.3 ขอบเขตการวิจัย.....2

1.4 ขั้นตอนการวิจัย.....2

1.5 ประโยชน์ที่คาดว่าจะได้รับ.....2

1.5.1 ประโยชน์ที่คาดว่าจะได้รับต่อผู้ดำเนินงาน.....2

1.5.2 ประโยชน์ที่ได้จากโครงงานที่พัฒนาขึ้น.....3

1.6 โครงสร้างของรายงาน.....3

บทที่ 2 ความรู้พื้นฐาน.....4

2.1 อนุกรมเวลา.....4

2.1.1 ลักษณะของข้อมูลอนุกรมเวลา.....4

2.2 แบบจำลอง Box-Jenkins.....4

2.2.1 แบบจำลองการถดถอยในตัว (Autoregressive).....5

2.2.2 แบบจำลองการเฉลี่ยเคลื่อนที่ (Moving Average).....6

หน้า

2.2.3 แบบจำลองการเฉลี่ยเคลื่อนที่ถดถอยในตัว (Autoregressive Moving Average) 7

2.2.4 กระบวนการ Integrated 8

2.2.5 แบบจำลองรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ (Autoregressive Integrated Moving Average) 9

2.3 การทดสอบความนิ่งของอนุกรมเวลาด้วยวิธี Augmented Dickey-Fuller.....11

2.4 การเลือกตัวแบบที่เหมาะสม 12

2.4.1 เกณฑ์ข้อสนเทศอาไคเคะ หรือเกณฑ์เอไอซี (Akaike Information Criterion, AIC).....12

2.4.2 ฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation Function ; ACF) 12

2.4.3 ฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน (Partial Auto-Correlation Function ; PACF).....13

2.4.4 การกำหนดแบบจำลองและอันดับของ p และ q 14

2.5 การวิเคราะห์การแจกแจงแบบปกติของส่วนเหลือ 15

2.5.1 แผนภาพฮิสโตแกรม 15

2.5.2 การทดสอบการแจกแจงแบบปกติโดยวิธี Quantile – Quantile Plot.....16

2.5.3 การทดสอบการแจกแจงปกติด้วยวิธี Kolmogorov-Smirnov normality test.....17

2.6 กระบวนการแปลงข้อมูล (Transformation) 17

2.7 อนุกรมเวลาเชิงการนับ (Count time series) 18

2.7.1 การแจกแจงแบบปัวซง (Poisson Distribution).....19

2.7.2 การแจกแจงแบบทวินามลบ (Negative Binomial Distribution).....19

2.7.3 แบบจำลองเชิงเส้น (Linear Models).....19

หน้า

2.7.4 แบบจำลองสำหรับอนุกรมเวลาเชิงการนับ 20

2.7.5 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง 21

2.7.6 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ 21

2.7.7 ค่าคลาดเคลื่อนกำลังสอง.....22

บทที่ 3 ผลการทำงาน 23

3.1 มูลค่าการเอาประกันวินาศภัย 23

1. การประกันอัคคีภัย (fire insurance).....24

2. การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance).....32

3. การประกันภัยรถยนต์ (Automobile insurance).....42

3.1.1 สรุปผล.....50

3.2 จำนวนกรมธรรม์ของประกันวินาศภัย.....51

1. การประกันอัคคีภัย (fire insurance).....51

2. การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance).....61

3. การประกันภัยรถยนต์ (Automobile insurance).....70

3.2.1 สรุปผล..... 82

รายการอ้างอิง 83

ภาคผนวก ก แบบเสนอหัวข้อโครงงาน

รายวิชา 2301399 Project Proposal ปีการศึกษา 2562 85

ประวัติผู้เขียน..... 89

สารบัญตาราง

หน้า

มูลค่าการเอาประกันวินาศภัย

1. การประกันอัคคีภัย (fire insurance)

ตารางที่ 3.1 มูลค่าการเอาประกันอัคคีภัย
รายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559 23

ตารางที่ 3.2 ค่า AIC ของแบบจำลอง 25

ตารางที่ 3.3 พารามิเตอร์ของแบบจำลอง 26

ตารางที่ 3.4 การทดสอบแบบจำลองด้วย Kolmogorov-Smirnov normality test 27

ตารางที่ 3.5 มูลค่าการเอาประกันจริงเทียบกับมูลค่าพยากรณ์ของการรับประกัน 27

2. การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance)

ตารางที่ 3.6 มูลค่าการเอาประกันภัยทางทะเลและขนส่ง
รายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559 31

ตารางที่ 3.7 ค่า AIC ของแบบจำลอง 33

ตารางที่ 3.8 พารามิเตอร์ของแบบจำลอง 34

ตารางที่ 3.9 การทดสอบแบบจำลองด้วย Kolmogorov-Smirnov normality test 36

ตารางที่ 3.10 การทดสอบแบบจำลองด้วย Kolmogorov-Smirnov normality test
ของเศษเหลือที่ขจัดข้อมูลผิดปกติ 39

ตารางที่ 3.11 มูลค่าการเอาประกันจริงเทียบกับมูลค่าพยากรณ์ของการเอาประกัน 40

3. การประกันภัยรถยนต์ (Automobile insurance)

ตารางที่ 3.12 มูลค่าการเอาประกันภัยรถยนต์
รายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559 42

ตารางที่ 3.13 ค่า AIC ของแบบจำลอง 44

ตารางที่ 3.14 พารามิเตอร์ของแบบจำลอง 45

ตารางที่ 3.15 การทดสอบแบบจำลองด้วย Kolmogorov-Smirnov normality test 47

หน้า

ตารางที่ 3.16 มูลค่าการเอาประกันจริงเทียบกับมูลค่าพยากรณ์ของการเอาประกัน 48
ตารางที่ 3.17 สรุปผลแบบจำลองที่เหมาะสมสำหรับการพยากรณ์ 50

จำนวนกรมธรรม์การเอาประกันวินาศภัย

1. การประกันอัคคีภัย (fire insurance)

ตารางที่ 3.18 จำนวนกรมธรรม์การประกันอัคคีภัย
รายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559 51
ตารางที่ 3.19 ค่า AIC ของแบบจำลองค่ารากที่สองของข้อมูล 54
ตารางที่ 3.20 พารามิเตอร์ของแบบจำลอง ARIMA 54
ตารางที่ 3.21 ค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง 56
ตารางที่ 3.22 ค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับ
ที่มีการแจกแจงแบบปัวซง 57
ตารางที่ 3.23 ค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ.. 58
ตารางที่ 3.24 ค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับ
ที่มีการแจกแจงแบบทวินามลบ 59
ตารางที่ 3.25 ค่าคลาดเคลื่อนกำลังสองของแบบจำลอง 59
ตารางที่ 3.26 จำนวนกรมธรรม์พยากรณ์ของแบบจำลองเทียบกับจำนวนกรมธรรม์จริง 60

2. การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance)

ตารางที่ 3.27 จำนวนกรมธรรม์การประกันภัยทางทะเลและขนส่ง
รายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559 61
ตารางที่ 3.28 ค่า AIC ของแบบจำลองค่ารากที่สองของข้อมูล 63
ตารางที่ 3.29 พารามิเตอร์ของแบบจำลอง ARIMA 64
ตารางที่ 3.30 ค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง 66
ตารางที่ 3.31 ค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับ
ที่มีการแจกแจงแบบปัวซง 66
ตารางที่ 3.32 ค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ.. 67

หน้า

ตารางที่ 3.33 ค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับ
ที่มีการแจกแจงแบบทวินามลบ 67
ตารางที่ 3.34 ค่าคลาดเคลื่อนกำลังสองของแบบจำลอง 68
ตารางที่ 3.35 จำนวนกรมธรรม์พยากรณ์ของแบบจำลองเทียบกับจำนวนกรมธรรม์จริง 69

3. การประกันภัยรถยนต์ (Automobile insurance)
ตารางที่ 3.36 จำนวนกรมธรรม์การประกันภัยรถยนต์
รายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559 70
ตารางที่ 3.37 ค่า AIC ของแบบจำลองค่ารากที่สองของข้อมูล 73
ตารางที่ 3.38 พารามิเตอร์ของแบบจำลอง ARIMA 74
ตารางที่ 3.39 ค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง 76
ตารางที่ 3.40 ค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับ
ที่มีการแจกแจงแบบปัวซง 77
ตารางที่ 3.41 ค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ.. 78
ตารางที่ 3.42 ค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับ
ที่มีการแจกแจงแบบทวินามลบ 79
ตารางที่ 3.43 ค่าคลาดเคลื่อนกำลังสองของแบบจำลอง 80
ตารางที่ 3.44 จำนวนกรมธรรม์พยากรณ์ของแบบจำลองเทียบกับจำนวนกรมธรรม์จริง 81
ตารางที่ 3.45 สรุปผลแบบจำลองที่เหมาะสมสำหรับการพยากรณ์ 82

สารบัญภาพ

หน้า

ภาพที่ 2.1 ตัวอย่าง ACF 13
ภาพที่ 2.2 ตัวอย่างแบบจำลอง PACF 14
ภาพที่ 2.3 แผนภาพฮิสโตแกรมของการแจกแจงแบบปกติ 16
ภาพที่ 2.4 แผนภาพการกระจายตัวของการแจกแจงแบบปกติด้วยวิธี QQ-plots 16

มูลค่าการรับประกันวินาศภัย

1. การประกันอัคคีภัย (fire insurance)

ภาพที่ 3.1 มูลค่าการเอาประกันระหว่างปี 2555 – 2559 25
ภาพที่ 3.2 ACF และ PACF ของมูลค่าการเอาประกัน 25
ภาพที่ 3.3 ACF และ PACF ของมูลค่าการเอาประกันโดยการหาผลต่างของข้อมูลครั้งที่ 1 26
ภาพที่ 3.4 ค่าเศษเหลือของมูลค่าการเอาประกันเทียบกับเวลาในแบบจำลอง 28
ภาพที่ 3.5 กราฟแสดงฮิสโตแกรมของเศษเหลือและ QQ-Plots ของมูลค่าการเอาประกัน 28
ภาพที่ 3.6 ค่า ACF ของเศษเหลือของมูลค่าการเอาประกันในแบบจำลอง 29
ภาพที่ 3.7 กราฟของมูลค่าการเอาประกันจริงเทียบกับมูลค่าพยากรณ์ของการเอาประกัน 31
ภาพที่ 3.8 แสดงผลพยากรณ์ของมูลค่าการเอาประกัน 31

2. การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance)

ภาพที่ 3.9 มูลค่าการเอาประกันระหว่างปี 2555 – 2559 33
ภาพที่ 3.10 ACF และ PACF ของมูลค่าการเอาประกัน 33
ภาพที่ 3.11 ACF และ PACF ของมูลค่าการเอาประกันโดยการหาผลต่างของข้อมูลครั้งที่ 1 34
ภาพที่ 3.12 ค่าเศษเหลือของมูลค่าการเอาประกันเทียบกับเวลาในแบบจำลอง 36
ภาพที่ 3.13 กราฟแสดงฮิสโตแกรมของเศษเหลือและ QQ-Plots ของมูลค่าการเอาประกัน 36
ภาพที่ 3.14 ค่า ACF ของเศษเหลือของมูลค่าการเอาประกันในแบบจำลอง 37
ภาพที่ 3.15 ค่าเศษเหลือที่ขจัดข้อมูลผิดปกติเทียบกับเวลาในแบบจำลอง 38

หน้า

ภาพที่ 3.16 กราฟแสดงฮิสโตแกรมของเศษเหลือที่ขจัดข้อมูลผิดปกติและ QQ-Plots 38
ภาพที่ 3.17 ค่า ACF ของเศษเหลือที่ขจัดข้อมูลผิดปกติ 39
ภาพที่ 3.18 กราฟของมูลค่าการเอาประกันจริงเทียบกับมูลค่าพยากรณ์ของการเอาประกัน 41
ภาพที่ 3.19 แสดงผลพยากรณ์ของมูลค่าการเอาประกัน 41

3. การประกันภัยรถยนต์ (Automobile insurance)

ภาพที่ 3.20 มูลค่าการเอาประกันระหว่างปี 2555 – 2559 43
ภาพที่ 3.21 ACF และ PACF ของมูลค่าการเอาประกัน 43
ภาพที่ 3.22 ACF และ PACF ของมูลค่าการเอาประกันโดยการหาผลต่างของข้อมูลครั้งที่ 1 44
ภาพที่ 3.23 ค่าเศษเหลือของมูลค่าการเอาประกันเทียบกับเวลาในแบบจำลอง 46
ภาพที่ 3.24 กราฟแสดงฮิสโตแกรมของเศษเหลือและ QQ-Plots ของมูลค่าการเอาประกัน 46
ภาพที่ 3.25 ค่า ACF ของเศษเหลือของมูลค่าการเอาประกันในแบบจำลอง 47
ภาพที่ 3.26 กราฟของมูลค่าการเอาประกันจริงเทียบกับมูลค่าพยากรณ์ของการเอาประกัน 49
ภาพที่ 3.27 แสดงผลพยากรณ์ของมูลค่าการเอาประกัน 49

จำนวนกรมธรรม์การเอาประกันวินาศภัย

1. การประกันอัคคีภัย (fire insurance)

ภาพที่ 3.28 อนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกัน 52
ภาพที่ 3.29 ACF และ PACF ของค่ารากที่สองของข้อมูล 52
ภาพที่ 3.30 ACF และ PACF ของค่ารากที่สองของข้อมูล
โดยการหาผลต่างของข้อมูลครั้งที่ 1 53
ภาพที่ 3.31 ACF และ PACF ของจำนวนกรมธรรม์ของการเอาประกัน 55

2. การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance)

ภาพที่ 3.32 อนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกัน 61
ภาพที่ 3.33 ACF และ PACF ของค่ารากที่สองของข้อมูล 62

หน้า

ภาพที่ 3.34 ACF และ PACF ของค่ารากที่สองของข้อมูล
โดยการหาผลต่างของข้อมูลครั้งที่ 1 63
ภาพที่ 3.35 ACF และ PACF ของจำนวนกรมธรรม์ของการเอาประกัน 65

3. การประกันภัยรถยนต์ (Automobile insurance)
ภาพที่ 3.36 อนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกัน 70
ภาพที่ 3.37 ACF และ PACF ขอค่ารากที่สองของข้อมูล 71
ภาพที่ 3.38 ACF และ PACF ของค่ารากที่สองของข้อมูล
โดยการหาผลต่างของข้อมูลครั้งที่ 1 72
ภาพที่ 3.39 ACF และ PACF ของจำนวนกรมธรรม์ของการเอาประกัน 75

บทที่ 1
บทนำ

1.1 ความเป็นมาและเหตุผลการวิจัย

ในปัจจุบันการทำประกันวินาศภัยนั้นเป็นที่ได้รับความนิยมเป็นอย่างมาก เนื่องจากสามารถให้ความคุ้มครองต่อบุคคล ครอบครัวและทรัพย์สิน แม้กระทั่งช่วยให้เกิดความมั่นคงในการประกอบธุรกิจจากความเสี่ยงที่อาจจะเกิดขึ้นได้ในอนาคต ทำให้บริษัทประกันภัยนั้นจำเป็นที่จะต้องหาแนวทางในการคาดการณ์แนวโน้มของมูลค่าการเอาประกันวินาศภัยล่วงหน้า เพื่อลดความเสี่ยงของการลงทุนและเพิ่มผลกำไรที่จะได้รับในการลงทุนของบริษัท

การวิเคราะห์อนุกรมเวลา (Time Series Analysis) เป็นวิธีการทางสถิติวิธีหนึ่งซึ่งเป็นที่นิยมในการนำมาใช้พยากรณ์ข้อมูลที่จะเกิดขึ้นในอนาคต ซึ่งจากการศึกษาเกี่ยวกับการวิเคราะห์อนุกรมเวลา (Time Series Analysis) พบว่ามีการสร้างแบบจำลองขึ้นมามากมาย แบบจำลองอนุกรมเวลาที่ได้รับความนิยม ได้แก่ แบบจำลอง ARIMA (Autoregressive Integrated Moving Average) ซึ่งพัฒนาโดย George E.P. Box และ Gwilym M. Jenkins ในปี ค.ศ. 1970 ซึ่งให้ค่าพยากรณ์ในระยะสั้นที่ดีและได้มีการนำมาประยุกต์ใช้อย่างกว้างขวาง เช่น แบบจำลองลี-คาร์เทอร์ (Lee Carter Models, 1992) ได้ใช้แบบจำลอง ARIMA ในการพยากรณ์อัตรามรณะของประชากรในประเทศสหรัฐอเมริกาและพฤทธ์สรรค์ สุทธิไชยเมธี (2553) วิเคราะห์อุปสงค์การนำเข้ายางพาราของประเทศจีนจากประเทศไทย ได้แก่ ยางแผ่นรมควัน ยางแท่งและน้ำยางข้น ในช่วงปี พ.ศ. 2538-2547 ด้วยแบบจำลอง ARIMA และพยากรณ์ล่วงหน้าช่วงไตรมาส 1-4 ในปี พ.ศ. 2548 เป็นต้น

ในโครงงานนี้ จะศึกษาหาแบบจำลองที่เหมาะสมที่ใช้ในการพยากรณ์มูลค่าการเอาประกันวินาศภัยและจำนวนกรมธรรม์รายเดือน จากข้อมูลสถิติการรับประกันวินาศภัย จำแนกตามประเภทการประกันภัย พ.ศ. 2555-2559 โดยสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย กระทรวงการคลัง

1.2 วัตถุประสงค์ของการวิจัย

ในโครงงานนี้ เรามีความสนใจในการศึกษาการประยุกต์ใช้อนุกรมเวลาเพื่อใช้ในการศึกษาแนวโน้มของข้อมูลของประกันวินาศภัยในอนาคตโดยแบ่งออกเป็น 2 ประเด็นที่สนใจ คือ (1) มูลค่าการเอาประกันวินาศภัย และ (2) จำนวนกรมธรรม์ของประกันวินาศภัย

1.3 ขอบเขตของการวิจัย

ในโครงงานนี้ เราจะศึกษาหาแบบจำลองอนุกรมเวลาที่เหมาะสมที่ใช้ในการพยากรณ์แนวโน้มมูลค่าการเอาประกันวินาศภัย และ จำนวนกรมธรรม์ของประกันวินาศภัย โดยข้อมูลมูลค่าการเอาประกันวินาศภัยและจำนวนกรมธรรม์ของประกันวินาศภัย ที่ใช้คือ ในช่วงเดือน มกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 จากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.)

1.4 ขั้นตอนการวิจัย

1. ศึกษาความรู้พื้นฐานเกี่ยวกับประกันวินาศภัย
2. ศึกษาความรู้พื้นฐานเกี่ยวกับข้อมูลการวิเคราะห์อนุกรมเวลา (Time Series Analysis)
3. ศึกษาความรู้พื้นฐานเกี่ยวกับแบบจำลองทางอนุกรมเวลา
4. ศึกษาการใช้โปรแกรม R
5. ศึกษาหาแบบจำลองอนุกรมเวลาที่เหมาะสม
6. สรุปผลและเขียนรายงาน

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ประโยชน์ที่คาดว่าจะได้รับต่อผู้ดำเนินงาน

1. มีความรู้เกี่ยวกับคุณสมบัติและการประยุกต์ใช้แบบจำลองอนุกรมเวลา
2. มีความรู้เกี่ยวกับประกันวินาศภัย
3. มีความรู้ในการใช้โปรแกรม R

1.5.2 ประโยชน์ที่ได้จากโครงงานที่พัฒนาขึ้น

1. เพื่อใช้เป็นความรู้ต่อยอดงานวิจัย
2. สามารถคาดการณ์มูลค่าการรับประกันวินาศภัยในอนาคต

1.6 โครงสร้างของรายงาน

บทที่ 2 จะกล่าวถึงความรู้พื้นฐานที่ใช้ในการทำโครงงานคือ อนุกรมเวลาและการวิเคราะห์อนุกรมเวลา แบบจำลองอนุกรมเวลาที่ใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลา การเลือกแบบจำลองที่เหมาะสมและการตรวจสอบแบบจำลองอนุกรมเวลา

บทที่ 3 จะกล่าวถึงการใช้แบบจำลองอนุกรมเวลาในการวิจัย และ ผลการดำเนินงานจะแบ่งออกเป็น การหาอันดับของแบบจำลอง การประมาณค่าพารามิเตอร์ของแต่ละแบบจำลอง การหาแบบจำลองของตัวแบบสำหรับมูลค่าการรับประกันและจำนวนกรมธรรม์ของประกันวินาศภัย การหาแบบจำลองของตัวแบบสำหรับการพยากรณ์มูลค่าการรับประกันและจำนวนกรมธรรม์ของประกันวินาศภัย และ การตรวจสอบแบบจำลองของตัวแบบสำหรับการพยากรณ์มูลค่าการรับประกันและจำนวนกรมธรรม์ของประกันวินาศภัย

บทที่ 2

ความรู้พื้นฐาน

2.1 อนุกรมเวลา (Time series)

อนุกรมเวลา (Time series) หมายถึง การเก็บรวบรวมข้อมูลของตัวแปรหนึ่งตามลำดับเวลา เช่น ปริมาณฝนเฉลี่ยรายเดือนตั้งแต่วันที่ 1 มกราคม 2554 ถึงวันที่ 30 มิถุนายน 2556 ข้อมูลอัตราแลกเปลี่ยนรายสัปดาห์ตั้งแต่สัปดาห์ที่ 1 ถึงสัปดาห์ที่ 52 ของปี 2555 ข้อมูลผลผลิตข้าวรายปีตั้งแต่ปี 2520-2554 และข้อมูล GDP รายไตรมาสตั้งแต่ไตรมาสที่ 2 ของปี 2525 ถึงไตรมาสที่ 4 ของปี 2554 เป็นต้น

การวิเคราะห์อนุกรมเวลา (Time series analysis) เป็นวิธีที่ใช้ในการวิเคราะห์ข้อมูลที่มีการเปลี่ยนแปลงในระยะเวลาใดเวลาหนึ่ง เพื่อพยากรณ์หรือคาดการณ์ข้อมูลที่จะเกิดขึ้นในอนาคตจากรูปแบบอนุกรมเวลาของข้อมูลในอดีต

2.1.1 ลักษณะของข้อมูลอนุกรมเวลา

1. ข้อมูลที่มีลักษณะนิ่ง (Stationary) คือ ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean) และค่าความแปรปรวน (Variance) คงที่ตลอดระยะเวลาที่ศึกษา
2. ข้อมูลที่มีลักษณะไม่นิ่ง (Non-Stationary) คือ ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean) และค่าความแปรปรวน (Variance) ไม่คงที่ตลอดระยะเวลาที่ศึกษา

2.2 แบบจำลอง Box-Jenkins

สำหรับการวิเคราะห์ข้อมูลที่มีลักษณะเป็นอนุกรมเวลาโดยใช้ข้อมูลในอดีตนั้น แบบจำลองหนึ่งที่ได้รับความนิยมมากที่สุดในการอธิบายอนุกรมเวลาก็คือ แบบจำลองรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ หรือแบบจำลอง ARIMA (Autoregressive Integrated Moving Average) ซึ่งถูกนำมาใช้ในสาขาเศรษฐศาสตร์โดยนักสถิติ 2 ท่านคือ จอร์จ อี.พี. บ็อกซ์ และ กวิลีม เอ็ม. เจนกินส์ (George E.P. Box & Gwilym M. Jenkins) ในช่วงทศวรรษที่ 1970 (พ.ศ. 2513) สำหรับจุดเด่นของแบบจำลอง

ARIMA ก็คือมีการรวบรวมและให้น้ำหนักความสำคัญกับข้อมูลทั้งหมดที่เกี่ยวข้องกับการพยากรณ์ ไม่ว่าจะเป็นข้อมูลจากการสังเกตในอดีตและปัจจัยภายนอกในอดีตจากการพยากรณ์ โดยจัดอยู่ในรูปแบบของสมการเชิงเส้น ซึ่งประกอบไปด้วย 3 ส่วนหลักๆ ได้แก่

1. แบบจำลองการถดถอยในตัว หรือ แบบจำลอง Auto Regressive ที่มีอันดับที่ p เขียนแทนด้วย $AR(p)$
2. กระบวนการ Integrated ที่มีอันอันดับที่ d เขียนแทนด้วย $I(d)$
3. แบบจำลองการเฉลี่ยเคลื่อนที่ หรือ แบบจำลอง Moving Average ที่มีอันดับที่ q เขียนแทนด้วย $MA(q)$

2.2.1 แบบจำลองการถดถอยในตัว หรือแบบจำลอง Autoregressive

แบบจำลองการถดถอยในตัว หรือแบบจำลอง Autoregressive เป็นแบบจำลองที่ใช้ในการอธิบายอนุกรมเวลา ณ เวลา t เขียนแทนด้วย y_t จะขึ้นกับข้อมูลในอดีต โดยมีความสัมพันธ์เป็นฟังก์ชันเชิงเส้นกับข้อมูลในอดีต ซึ่งสามารถเขียนสมการของแบบจำลอง Auto Regressive ลำดับที่ p หรือ $AR(p)$ ได้ดังนี้

$$y_t = \mu + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + \epsilon_t$$

โดยที่ y_t คือ ข้อมูลในอนุกรมเวลา ณ เวลา t

μ คือ ค่าคงที่

$\emptyset_j$ คือ สัมประสิทธิ์ของข้อมูล ณ เวลา $t-j$ โดยที่ $j = 1,2,\ldots,p$

ϵ_t คือ ปัจจัยภายนอก ณ เวลา t

จากนิยามข้างต้น เราจะยกตัวอย่างของแบบจำลอง Auto Regressive ดังนี้

ตัวอย่าง $AR(1)$ เขียนได้ดังนี้

$$y_t = \mu + \emptyset_1 y_{t-1} + \epsilon_t$$

ตัวอย่าง $AR(2)$ เขียนได้ดังนี้

$$y_t = \mu + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \epsilon_t$$

2.2.2 แบบจำลองการเฉลี่ยเคลื่อนที่ หรือแบบจำลอง Moving Average

แบบจำลองการเฉลี่ยเคลื่อนที่ หรือแบบจำลอง Moving Average เป็นแบบจำลองที่ใช้ในการอธิบายอนุกรมเวลา ณ เวลา t เขียนแทนด้วย y_t จะขึ้นกับปัจจัยภายนอกในอดีต โดยมีความสัมพันธ์เป็นฟังก์ชันเชิงเส้นกับปัจจัยภายนอกในอดีต ซึ่งสามารถเขียนสมการของแบบจำลอง Moving Average อันดับที่ q หรือ $MA(q)$ ได้ดังนี้

$$y_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q}$$

โดยที่ y_t คือ ข้อมูลในอนุกรมเวลา ณ เวลา t

μ คือ ค่าคงที่

ϵ_t คือ ปัจจัยภายนอก ณ เวลา t

θ_j คือ สัมประสิทธิ์ของข้อมูล ณ เวลา $t-j$ โดยที่ $j = 1,2,\ldots,q$

จากนิยามข้างต้น เราจะยกตัวอย่างของแบบจำลอง Moving Average ดังนี้

ตัวอย่าง $MA(1)$ เขียนได้ดังนี้

$$y_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

ตัวอย่าง $MA(2)$ เขียนได้ดังนี้

$$y_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}$$

2.2.3 แบบจำลองการเฉลี่ยเคลื่อนที่ถดถอยในตัว หรือแบบจำลอง Autoregressive Moving Average

แบบจำลองการเฉลี่ยเคลื่อนที่ถดถอยในตัว หรือแบบจำลอง Autoregressive Moving Average เป็นแบบจำลองที่ประยุกต์ใช้แบบจำลอง Autoregressive และแบบจำลอง Moving Average มาผสมผสานกัน โดยใช้ในการอธิบายอนุกรมเวลา ณ เวลา t เขียนแทนด้วย y_t จะขึ้นกับข้อมูลในอดีตและขึ้นกับปัจจัยภายนอกในอดีตซึ่งสามารถเขียนสมการของแบบจำลอง Autoregressive Moving Average อันดับที่ (p, q) หรือ $ARMA(p, q)$ ได้ดังนี้

$$y_t = \mu + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + \epsilon_t - \theta_t \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q}$$

โดยที่ μ คือ ค่าคงที่

y_t คือ ข้อมูลในอนุกรมเวลา ณ เวลา t

$\emptyset_j$ คือ ข้อ สัมประสิทธิ์ของข้อมูล ณ เวลา $t - j$ โดยที่ $j = 1,2, \dots, p$

θ_j คือ สัมประสิทธิ์ของข้อมูล ณ เวลา $t - j$ โดยที่ $j = 1,2, \dots, q$

ϵ_t คือ ปัจจัยภายนอก ณ เวลา t

จากนิยามข้างต้น เราจะยกตัวอย่างของแบบจำลอง Autoregressive Moving Average ดังนี้

ตัวอย่าง $ARMA(1,1)$ เขียนได้ดังนี้

$$y_t = \mu + \emptyset_1 y_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

ตัวอย่าง $ARMA(1,2)$ เขียนได้ดังนี้

$$y_t = \mu + \emptyset_1 y_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}$$

2.2.4 กระบวนการ Integrated

เนื่องจากแบบจำลอง $ARMA(p,q)$ จะต้องนำไปใช้กับอนุกรมเวลาที่มีความนิ่ง (Stationary) เท่านั้น ดังนั้น หากพบว่าอนุกรมเวลาไม่มีความนิ่ง (Non-Stationary) เราจะต้องแปลงอนุกรมเวลานั้นให้มีความนิ่งเสียก่อน และวิธีการหนึ่งที่มักถูกนำมาใช้แปลงอนุกรมเวลาที่ไม่มีความนิ่งให้เป็นอนุกรมเวลาที่มีความนิ่งก็คือ วิธีการหาผลต่าง

โดยรูปทั่วไปหาก อนุกรมเวลา ณ เวลา t เขียนแทนด้วย y_t ไม่มีความนิ่ง และการทำผลต่างลำดับที่ d กับอนุกรมเวลา y_t จะทำให้ได้ข้อมูลอนุกรมเวลาที่มีความนิ่ง เขียนได้ดังนี้

$$\Delta^d y_t = \Delta^{d-1}(y_t - y_{t-1})$$

โดยที่ Δ^d คือ ผลต่างอันดับที่ d

จากนิยามข้างต้น เราจะยกตัวอย่าง*ผลต่างอันดับที่* d ดังนี้

ตัวอย่าง ผลต่างอันดับที่หนึ่งของอนุกรมเวลา y_t เขียนได้ดังนี้

$$\Delta y_t = y_t - y_{t-1}$$

ตัวอย่าง ผลต่างอันดับที่สองของอนุกรมเวลา y_t เขียนได้ดังนี้

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

2.2.5 แบบจำลองรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ หรือแบบจำลอง Autoregressive Integrated Moving Average

แบบจำลองรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ หรือแบบจำลอง Autoregressive Integrated Moving Average เป็นแบบจำลองที่เกิดจากการรวมกันระหว่าง แบบจำลอง Autoregressive Moving Average อันดับ (p, q) กับ กระบวนการ Integrated อันดับที่ d ซึ่งสามารถเขียนสมการของแบบจำลอง Autoregressive Integrated Moving Average อันดับที่ (p, d, q) หรือ $ARIMA(p, d, q)$ ได้ดังนี้

$$\Delta^d y_t = \mu + \emptyset_1 \Delta^d y_{t-1} + \emptyset_2 \Delta^d y_{t-2} + \cdots + \emptyset_p \Delta^d y_{t-p} + \epsilon_t - \theta_t \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q}$$

โดยที่ y_t คือ ข้อมูลในอนุกรมเวลา ณ เวลา t

μ คือ ค่าคงที่

d คือ จำนวนครั้งที่หาผลต่างเพื่อให้อนุกรมเวลานิ่ง

Δ^d คือ ผลต่างอันดับที่ d

ϵ_t คือ ปัจจัยภายนอก ณ เวลา t โดยมีคุณสมบัติเป็นตัวรบกวนขาวแบบเกาส์เซียน คือ มีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น σ^2

จากนิยามข้างต้น เราจะยกตัวอย่างของแบบจำลอง แบบจำลอง Autoregressive Integrated Moving Average ดังนี้

ตัวอย่าง $ARIMA(1,1,1)$ เขียนได้ดังนี้

$$\Delta y_t = \mu + \emptyset_1 \Delta y_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

แทนค่า $\Delta y_t = y_t - y_{t-1}$

นั่นคือ $$y_t - y_{t-1} = \mu + \emptyset_1 (y_{t-1} - y_{t-2}) + \epsilon_t - \theta_1 \epsilon_{t-1}$$

$$y_t = \mu + y_{t-1} + \emptyset_1 y_{t-1} - \emptyset_1 y_{t-2} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

$$y_t = \mu + (1 + \emptyset_1) y_{t-1} - \emptyset_1 y_{t-2} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

2.3 การทดสอบความนิ่งของอนุกรมเวลาด้วยวิธี Augmented Dickey-Fuller

วิธี Augmented Dickey-Fuller หรือ ADF เป็นวิธีการทดสอบความนิ่งของอนุกรมเวลา โดยการตั้งสมมติฐานหลักและสมมติฐานรองดังนี้ $H_0: \gamma = 0$ และ $H_1: \gamma < 0$ ถ้าเราไม่สามารถปฏิเสธสมมติฐานหลัก นั่นหมายถึงอนุกรมเวลา y_t ไม่มีความนิ่ง (หรืออนุกรมเวลา y_t มี Unit Root) และหากปฏิเสธสมมติฐานหลัก จะหมายถึงอนุกรมเวลา y_t มีความนิ่ง โดยที่ $\gamma = \emptyset_1 + \emptyset_2 + \cdots + \emptyset_p - 1$

ถ้าค่าพารามิเตอร์ภายใต้สมมติฐานหลักคือ $\gamma = 0$ เป็นจริง แล้วตัวประมาณค่าด้วยวิธีกำลังสองน้อยที่สุด $(\hat{\gamma})$ จะไม่ใช่การแจกแจงแบบปกติ ในกรณีนี้ เราจะใช้ค่าสถิติ t ของ $\hat{\gamma}$ มาใช้เทียบค่าวิกฤติของ MacKinnon (1991,1996) ซึ่งเป็นค่าวิกฤติที่ใช้กับกรณีตัวอย่างมีขนาดเล็กด้วย ส่วนค่าสถิติ t ของค่าสัมประสิทธิ Δy_{t-i} $(i = 1,2, \dots, p-1)$ สามารถเทียบกับค่าวิกฤติจากตาราง t หรือ F ได้ โดยสมการที่ใช้ทดสอบความนิ่งของอนุกรมเวลา y_t ด้วยวิธี ADF แบ่งเป็น 3 กรณี ซึ่งแสดงได้ดังนี้

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=1}^{p-1} c_i \Delta y_{t-i} + \epsilon_t \cdots\cdots\cdots (i)$$

$$\Delta y_t = \beta_0 + \gamma y_{t-1} + \sum_{i=1}^{p-1} c_i \Delta y_{t-i} + \epsilon_t \cdots\cdots\cdots (ii)$$

$$\Delta y_t = \beta_0 + \beta_1 t + \gamma y_{t-1} + \sum_{i=1}^{p-1} c_i \Delta y_{t-i} + \epsilon_t \cdots\cdots\cdots (iii)$$

โดยที่ $\gamma = \emptyset_1 + \emptyset_2 + \cdots + \emptyset_p - 1$

โดยเมื่อเราวาดแผนภาพของอนุกรมเวลาที่ต้องการทดสอบความนิ่ง แล้วพบว่า อนุกรมเวลานั้นเคลื่อนที่ขึ้นๆ ลงๆ อยู่รอบๆศูนย์ เราควรเลือกใช้สมการ (i) และหากพบว่าอนุกรมเวลาไม่มีแนวโน้มที่เพิ่มขึ้นหรือลดลงเมื่อเวลาผ่านไป แต่เคลื่อนที่ขึ้นๆ ลงๆ รอบค่าคงที่ค่าหนึ่ง เราควรเลือกใช้สมการ (ii) และหากอนุกรมเวลานั้นมีแนวโน้มที่เพิ่มขึ้นหรือลดลงเมื่อเวลาผ่านไป เราควรเลือกใช้สมการ (iii)

2.4 การเลือกตัวแบบที่เหมาะสม

2.4.1 เกณฑ์ข้อสนเทศอาไคเคะ หรือเกณฑ์เอไอซี (Akaike Information Criterion, AIC)

ในปี ค.ศ. 1973 อาไคเคะได้เสนอเกณฑ์ในการคัดเลือกแบบจำลองที่เหมาะสมเพื่อใช้เป็นเครื่องมือในการหาแบบจำลองที่ให้ค่าพยากรณ์ที่แม่นยำที่สุด โดยเกณฑ์ข้อสนเทศอาไคเคะ เป็นเกณฑ์ที่พิจารณาจากการประมาณค่าความคลาดเคลื่อนรวมเข้ากับข้อสนเทศของค่าสังเกตและใช้แนวคิดจากการหาค่าต่ำสุดของข้อสนเทศ ด้วยหลักการคูลส์แบล็ค-ไลท์เบอร์ (Kullback-Leiber) เพื่อนำมาใช้ในการปรับค่าประมาณของการพยากรณ์ให้มีความแม่นยำมากยิ่งขึ้น

$$AIC_k = -2\log(ML_k) + 2p_k$$

โดยที่ ML_k คือ ฟังก์ชันภาวะน่าจะเป็นสูงสุด (Maximum likelihood Function) ของแบบจำลองที่ k

p_k คือ จำนวนพารามิเตอร์ของแบบจำลองที่ k ซึ่งตัวแบบที่ให้ค่า AIC ต่ำสุดจะเป็นแบบจำลองที่ให้ค่าพยากรณ์แม่นยำและถูกต้อง

2.4.2 ฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation Function ; ACF)

เป็นฟังก์ชันการวัดสหสัมพันธ์ระหว่างข้อมูล ณ เวลา ณ เวลา t (y_t) และข้อมูล ณ เวลา $t+h$ (y_{t+h}) ของช่วงเวลาห่างกัน h หน่วย เขียนสมการได้ดังนี้

$$\rho_h = \frac{Cov(y_t, y_{t+h})}{\sqrt{Var(y_t)}\sqrt{Var(y_{t+h})}}$$

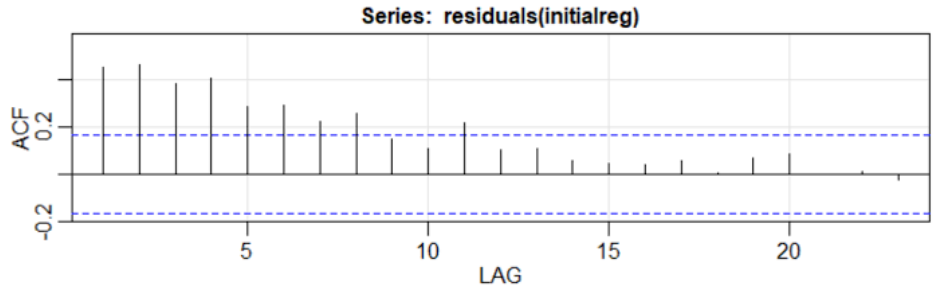
สำหรับข้อมูลที่มีลักษณะนิ่ง จะได้ว่า

$$\rho_h = \frac{Cov(y_t, y_{t+h})}{Var(y_t)}$$

ซึ่งสามารถประมาณค่าได้ด้วย

$$r_h = \frac{\sum_{t=1}^{n-h}(y_t - \bar{y})(y_{t+h} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$$

ตัวอย่าง



ภาพที่ 2.1 แสดงตัวอย่าง ACF

2.4.3 ฟังก์ชันสหสัมพันธ์ในตัวเองบางส่วน (Partial Auto-Correlation Function ; PACF)

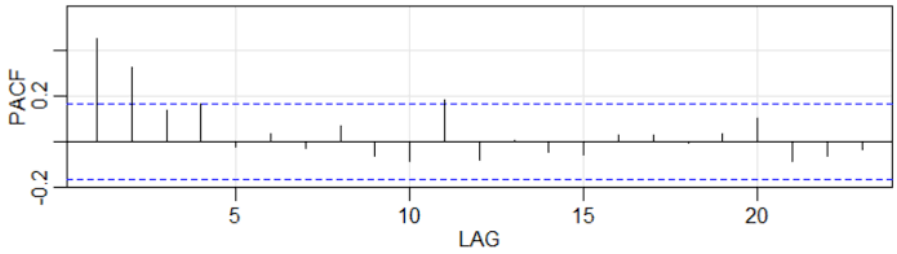
การพิจารณาสหสัมพันธ์ระหว่างข้อมูล ณ เวลา t (y_t) และข้อมูล ณ เวลา $t+h$ (y_{t+h}) โดยไม่มีอิทธิพลของข้อมูลระหว่างเวลา t (y_t) และ $t+h$ (y_{t+h}) มาเกี่ยวข้อง

$$Corr(y_t, y_{t+h} | y_{t+1}, \dots, y_{t+h-1})$$

ซึ่งจะได้ว่า PACF มีค่า

$$\emptyset_{hh} = \frac{Cov[(y_t|y_{t+1}, \dots, y_{t+h-1}), (y_{t+h}|y_{t+1}, \dots, y_{t+h-1})]}{\sqrt{Var(y_t|y_{t+1}, \dots, y_{t+h-1})Var(y_{t+h}|y_{t+1}, \dots, y_{t+h-1})}}$$

ตัวอย่าง



ภาพที่ 2.2 แสดงตัวอย่าง PACF

2.4.4 การกำหนดแบบจำลองและอันดับของ p และ q

แบบจำลอง $AR(p)$ ACF จะมีลักษณะลดลงเข้าหาศูนย์อย่างรวดเร็วและ PACF ที่ห่างกัน p ช่วงเวลา จะมีลักษณะเข้าสู่ศูนย์อย่างรวดเร็ว ส่วนแบบจำลอง $MA(q)$ ACF ที่ห่างกัน q ช่วงเวลา จะมีลักษณะเข้าสู่ศูนย์อย่างรวดเร็วและ PACF จะมีลักษณะลดลงเข้าหาศูนย์อย่างรวดเร็ว สุดท้ายแบบจำลอง $ARMA(p,q)$ ACF ที่ห่างกันเกิน q ช่วงเวลาและ PACF ที่ห่างกันเกิน p ช่วงเวลาจะลดลงเรื่อยๆ อย่างรวดเร็ว

ตัวแบบ	ACF	PACF
$AR(p)$	ลดลงเข้าหา 0 อย่างรวดเร็ว	หลัง $lag\ p$ มีค่าเท่ากับ 0
$MA(q)$	หลัง $lag\ q$ มีค่าเท่ากับ 0	ลดลงเข้าหา 0 อย่างรวดเร็ว
$ARMA(p,q)$	ลดลงเรื่อยๆ อย่างรวดเร็วหลัง q	ลดลงเรื่อยๆ อย่างรวดเร็วหลัง p

ตารางที่ 2.1 แสดงลักษณะของ ACF และ PACF ของแบบจำลอง

2.5 การวิเคราะห์การแจกแจงแบบปกติของส่วนเหลือ

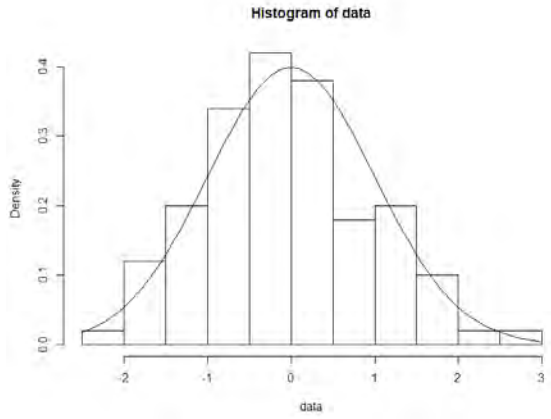
เมื่อมีการนำแบบจำลองหรือสมการมาใช้ในการพยากรณ์ ค่าที่ได้จากการพยากรณ์ก็คือค่าจากสมการประมาณ หรือค่าสารูป (Fitted Value) และสิ่งที่ไม่สามารถอธิบายได้จากแบบจำลองหรือสมการก็คือค่าความคลาดเคลื่อนจากการพยากรณ์ (Forecast Error) หรือส่วนเหลือ (Residual) ซึ่งก็คือผลต่างระหว่างค่าสังเกตของตัวแปรสุ่มหนึ่งกับค่าจากการพยากรณ์หรือค่าประมาณของตัวแปรสุ่มนั้น

โดยเฉพาะเมื่อมีการกล่าวถึงส่วนเหลือจากแบบจำลองการถดถอยเชิงเส้น ก็มักจะกล่าวว่าส่วนเหลือดังกล่าวต้องมีการแจกแจงแบบปกติด้วย สาเหตุที่เป็นเช่นนั้นก็เนื่องจากการประมาณค่าสัมประสิทธิ์ของแบบจำลองถดถอยเชิงเส้นจะใช้หลักการกำลังสองน้อยสุด ซึ่งข้อสมมติฐานหนึ่งก็คือ ค่าส่วนเหลือจะต้องมีการแจกแจงแบบปกติ ด้วยเหตุนี้หากแบบจำลองพยากรณ์นั้นสร้างขึ้นมาจากข้อมูลซึ่งมีจำนวนจำกัด ส่วนเหลือจะต้องมีการแจกแจงแบบปกติ อย่างไรก็ดี หากมีข้อมูลจำนวนมากพอ ก็อาจไม่มีความจำเป็นที่จะต้องตรวจสอบเงื่อนไขของการแจกแจงแบบปกติ ด้วยเหตุนี้หากข้อมูลมีจำนวนไม่มากพอ การตรวจสอบการแจกแจงแบบปกติของส่วนเหลือก็อาจมีความจำเป็น

2.5.1 แผนภาพฮิสโตแกรม

การใช้แผนภาพฮิสโตแกรมเพื่อใช้ตรวจสอบการแจกแจงแบบปกตินั้นจะเน้นที่การพิจารณารูปร่างของการกระจายของข้อมูลที่สนใจ ข้อมูลที่มีการแจกแจงแบบปกติควรมีลักษณะดังนี้คือ มีลักษณะสมมาตรรอบค่าเฉลี่ย หรือมีรูปร่างคล้ายคลึงกับรูปทรงระฆังคว่ำ

ตัวอย่าง

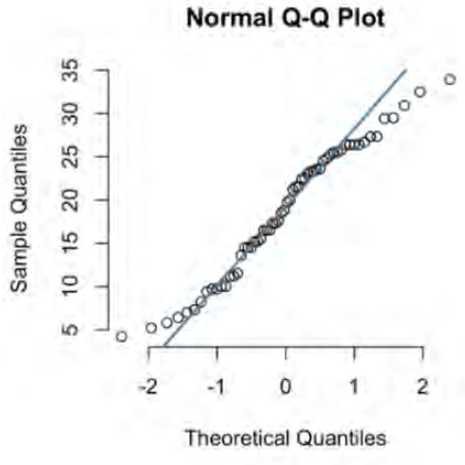


ภาพที่ 2.3 แสดงแผนภาพฮิสโตแกรมของการแจกแจงแบบปกติ

2.5.2 การทดสอบการแจกแจงแบบปกติโดยวิธี Quantile – Quantile Plot หรือ QQ – plots

การทดสอบการแจกแจงแบบปกติโดยวิธี Quantile – Quantile Plot หรือ QQ – plots คือการวาดแผนภาพเพื่อเปรียบเทียบควอนไทล์ของข้อมูลสองชนิดว่ามีการแจกแจงเดียวกันหรือไม่ โดยถ้าจุดที่ได้มีการจัดเรียงตัวในลักษณะเส้นตรงที่ 45 องศา จะได้ว่ามีการแจกแจงแบบปกติ

ตัวอย่าง



ภาพที่ 2.4 แสดงแผนภาพการกระจายตัวของการแจกแจงแบบปกติด้วยวิธี QQ-plots

2.5.3 การทดสอบการแจกแจงปกติด้วยวิธี Kolmogorov-Smirnov normality test

การทดสอบการแจกแจงปกติด้วยวิธี Kolmogorov-Smirnov normality test เป็นการทดสอบความแตกต่างระหว่างการแจงแจงความถี่สะสมของตัวอย่างกับการแจกแจงความถี่สะสมของการแจกแจงการแจกแจงใดๆ โดยกำหนดให้ $G(x)$ แทนฟังก์ชันของกลุ่มตัวอย่างที่ไม่ทราบการแจกแจงและ $F(x)$ แทนฟังก์ชันการแจกแจงใดๆที่ต้องการทดสอบ เราจะสามารถเขียนสมมติฐานของการทดสอบได้ดังนี้

$$H_0: G \equiv F$$

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

เพื่อที่จะทดสอบสมมติฐานเราจะใช้ วิธี Kolmogorov-Smirnov คำนวณสถิติทดสอบจากความแตกต่างของค่าสัมบูรณ์ระหว่างการแจกแจงของตัวอย่างกับการแจกแจงใดๆ

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

2.6 กระบวนการแปลงข้อมูล (Transformation)

กระบวนการแปลงข้อมูลนั้นจะมีขั้นตอนโดยเริ่มจากการวิเคราะห์ทางสถิติว่าข้อมูลนั้นมีลักษณะที่เป็นไปตามเงื่อนไขหรือไม่ เช่น มีการแจกแจงแบบปกติหรือมีความแปรปรวนคงที่หรือไม่ หากไม่มีลักษณะดังกล่าว สิ่งที่จะต้องทำต่อไปตามกระบวนการแปลงข้อมูลก็คือการเลือกรูปแบบหรือฟังก์ชันทางคณิตศาสตร์ที่จะนำมาใช้ในการแปลงข้อมูล หลังจากนั้นจึงทำการตรวจสอบด้วยวิธีการทางสถิติอีกครั้งหนึ่งว่าข้อมูลมีลักษณะเป็นไปตามเงื่อนไขหรือไม่ หากไม่ใช่ ผู้วิเคราะห์ก็เลือกฟังก์ชันอื่นๆ เพื่อมาทำการแปลงข้อมูล อีกครั้งจนข้อมูลจะเป็นไปตามเงื่อนไข โดยจะเป็นขั้นตอนที่ทำซ้ำไปเรื่อยๆ จนกว่าจะพบฟังก์ชันที่เหมาะสม และในขั้นตอนสุดท้ายเมื่อเลือกขั้นตอนที่เหมาะสมที่สุดแล้ว สิ่งที่ผู้วิเคราะห์จะต้องไม่ลืมก็คือ การแปลงข้อมูลกลับเป็นหน่วยเดิม

ฟังก์ชัน	ลักษณะของข้อมูล	ตัวอย่าง
การยกกำลัง $y^p (p > 1)$	• เบ้ซ้าย	ข้อมูลทางการแพทย์ของผู้ป่วย
การถอดราก	• เบ้ขวา • มีการแจกแจงแบบปัวซงหรือทวินาม	จำนวนการเคลมประกันภัย
การผกผัน	• มีการเบ้ไปทางซ้ายหรือขวาอย่างมากหรือมีลักษณะไม่เป็นเชิงเส้น	เวลาในการผลิต
ลอการิทึม	• ข้อมูลมีการเบ้ขวา • มีการแจกแจงแบบเอกโพเนนเชียล หรือแผนภาพกระจายระหว่างค่าส่วนเหลือกับค่าที่ได้จากการพยากรณ์ที่รูปร่างเป็นแบบลำโพง	ผลประกอบการลงทุน

2.7 อนุกรมเวลาเชิงการนับ (Count time series)

อนุกรมเวลาเชิงการนับ คือ ข้อมูลของจำนวนเหตุการณ์ที่เกิดขึ้นในช่วงเวลาใดเวลาหนึ่งที่เราให้ความสนใจหรือทำการสังเกต โดยสามารถพบเห็นได้ในสถานการณ์ต่างๆ อาทิเช่น จำนวนครั้งในการเคลมประกันภัยรายเดือน จำนวนธุรกรรมหรือการซื้อขายของหุ้นต่อนาที จำนวนผู้ป่วยโรคบางอย่างที่เข้ารับการรักษาที่โรงพยาบาล เป็นต้น

2.7.1 การแจกแจงแบบปัวซง (Poisson Distribution)

เป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มไม่ต่อเนื่อง โดยลักษณะของการแจกแจงนี้จะเป็นการศึกษาจำนวนครั้งของการเกิดเหตุการณ์ที่สนใจ ที่เกิดในช่วงเวลาใดเวลาหนึ่ง ตัวแปรสุ่ม X จะมีการแจกแจงแบบปัวซง ที่มีพารามิเตอร์ $\lambda > 0$ จะเขียนแทนสัญลักษณ์ด้วย $X \sim Poisson(\lambda)$ และฟังก์ชันการแจกแจงแบบปัวซงคือ

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

จากนิยามจะได้ว่า $\mu = E(X) = \lambda$ และ $\sigma^2 = Var(X) = \lambda$

2.7.2 การแจกแจงแบบทวินามลบ (Negative Binomial Distribution)

เป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มไม่ต่อเนื่อง โดยลักษณะของการแจกแจงนี้จะเป็นการศึกษาจำนวนครั้งของการเกิดเหตุการณ์ที่กระทำจนสำเร็จ รวม k ครั้ง ตัวแปรสุ่ม X จะมีการแจกแจงแบบทวินามลบที่มีพารามิเตอร์ k และ p จะเขียนแทนสัญลักษณ์ด้วย $X \sim NegBin(k,p)$ และฟังก์ชันการแจกแจงแบบทวินามลบคือ

$$P(X = x) = \binom{k+n-1}{n} q^n p^k \text{ เมื่อ } k = 0,1,2,\ldots$$

จากนิยามจะได้ว่า $\mu = E(X) = \frac{kq}{p}$ และ $\sigma^2 = Var(X) = \frac{kq}{p^2}$

2.7.3 แบบจำลองเชิงเส้น (Linear Models)

แบบจำลองเชิงเส้น คือ การหาความสัมพันธ์ระหว่างตัวแปรต้น (Explanatory Variables) กับตัวแปรตาม (Response Variable) โดยมีแนวคิดว่า ตัวแปรตาม (Y) สามารถเขียนให้อยู่ในรูปผลรวมของตัวแปรต้น (X) ซึ่งอยู่ในรูปแบบสมการเส้นตรงดังนี้

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

โดยที่ $\beta_0, \beta_1, \dots, \beta_k$ คือ สัมประสิทธิการถดถอย

ε คือ ค่าความเคลื่อนจากแบบจำลอง

2.7.4 แบบจำลองสำหรับอนุกรมเวลาเชิงการนับ

แบบจำลองสำหรับอนุกรมเวลาเชิงการนับ คือ แบบจำลองอนุกรมที่สามารถเขียนเป็นฟังก์ชันเชิงเส้นของข้อมูลสามส่วนได้แก่ ส่วนของข้อมูลในอดีต $\left(\sum_{k=1}^{p} \beta_k \tilde{g}\left(Y_{t-i_k}\right)\right)$ ส่วนของค่าเฉลี่ยแบบมีเงื่อนไข $\left(\sum_{l=1}^{q} \alpha_l g\left(\lambda_{t-j_l}\right)\right)$ และส่วนข้อมูลอนุกรมเวลาของตัวแปรอิสระ $\left(\eta^T X_t\right)$

กำหนดให้ Y_t คือ ข้อมูลอนุกรมเวลาเชิงการนับที่เราสนใจ โดยที่ $t \in \mathbb{N}$

X_t คือ ข้อมูลอนุกรมเวลาของตัวแปรอิสระ

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

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

จะได้ว่าแบบจำลองสำหรับอนุกรมเวลาเชิงการนับคือ

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

โดยที่ $\eta = (\eta_1, \eta_2, \dots, \eta_r)^T$ คือ ค่าพารามิเตอร์ของเวกเตอร์

$P = \{i_1, i_2, \dots, i_p\}$ โดยที่ $p \in \mathbb{N}$ และ $Q = \{j_1, j_2, \dots, j_q\}$ โดยที่ $q \in \mathbb{N}$

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

Y_t และ $\alpha_l, \alpha_2, \dots, \alpha_q$ ค่าสัมประสิทธิ์การถดถอยของ λ_t

2.7.5 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

กำหนดให้ $\nu_t = log(\lambda_t)$ โดยที่ $(Y_t = y|F_{t-1}) \sim Poisson(\lambda_t)$

จากแบบจำลองสำหรับอนุกรมเวลาเชิงการนับจะได้ว่า

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

โดยที่ $P(Y_t = y|F_{t-1}) = \frac{\lambda_t^y \exp(-\lambda_t)}{y!}$

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

2.7.6 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

กำหนดให้ $\nu_t = log(\lambda_t)$ โดยที่ $(Y_t = y|F_{t-1}) \sim NegBin(\lambda_t, \emptyset)$

จากแบบจำลองสำหรับอนุกรมเวลาเชิงการนับจะได้ว่า

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

โดยที่ $P(Y_t = y|F_{t-1}) = \frac{\Gamma(\emptyset+y)}{\Gamma(y+1)\Gamma(\emptyset)} \left(\frac{\emptyset}{\emptyset+\lambda_t}\right)^{\emptyset} \left(\frac{\lambda_t}{\emptyset+\lambda_t}\right)^{y}, y = 0,1, \ldots$

$$Var(Y_t|F_{t-1}) = \lambda_t + {\lambda_t}^2/\emptyset$$

2.7.7 ค่าคลาดเคลื่อนกำลังสอง

ค่าคลาดเคลื่อนกำลังสอง คือ ผลรวมกำลังสองความคลาดเคลื่อน โดยเป็นการวัดความแปรผันที่ไม่สามารถอธิบายได้โดยสมการถดถอย โดยในการเลือกแบบจำลองสำหรับอนุกรมเวลาเชิงการนับ จะใช้ค่าคลาดเคลื่อนกำลังสองมาใช้ในการพิจารณาแบบจำลองอนุกรมเวลาเชิงการนับที่เหมาะสมที่สุด โดยแบบจำลองที่ให้ค่าคลาดเคลื่อนกำลังสองน้อยที่สุด จะเป็นแบบจำลองอนุกรมเวลาเชิงการนับที่พยากรณ์ได้ดีที่สุด

$$\text{ค่าคลาดเคลื่อนกำลังสอง} = \left(y_t - \lambda_t\right)^2$$

โดยที่ คือ y_t ข้อมูลที่สังเกต ณ เวลา t

λ_t ข้อมูลพยากรณ์ ณ เวลา t

บทที่ 3

ผลการทำงาน

ในโครงงานนี้ เรามีความสนใจในการศึกษาการประยุกต์ใช้อนุกรมเวลาเพื่อใช้ในการศึกษาแนวโน้มของข้อมูลของประกันวินาศภัยในอนาคตโดยแบ่งออกเป็น 2 ประเด็นที่สนใจ คือ (1) มูลค่าการเอาประกันวินาศภัย และ (2) จำนวนกรมธรรม์ของประกันวินาศภัย ซึ่งประกันวินาศภัยสามารถแบ่งได้ออกเป็น 3 ประเภท คือ การประกันอัคคีภัย (Fire insurance) การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance) และ การประกันภัยรถยนต์ (Automobile insurance) โดยข้อมูลมูลค่าการรับประกันวินาศภัย และ จำนวนกรมธรรม์ของประกันวินาศภัย ที่ใช้คือ ในช่วงเดือน มกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 จากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.)

3.1 มูลค่าการเอาประกันวินาศภัย

จากการวิเคราะห์อนุกรมเวลาจะสามารถประมาณค่าพารามิเตอร์และพยากรณ์มูลค่าการเอาประกันวินาศภัยได้ โดยข้อมูลที่ใช้ศึกษาเป็นข้อมูลมูลค่าการเอาประกันวินาศภัย ในช่วงเดือน มกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 จากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) โดยจะพิจารณาอนุกรมเวลา y_t ด้วย ตัวแบบจำลองรวมการถดถอยในตัวกับการเฉลี่ยเคลื่อนที่ลำดับที่ (p,d,q) หรือ $ARIMA(p,d,q)$ ตามที่นิยามในหัวข้อ 2.2 โดยกำหนดให้ $t = 1,2,3,\dots,60$

y_t คือ ข้อมูลมูลค่าการเอาประกันวินาศภัย ณ เวลา t

ϵ_t คือ ปัจจัยภายนอก ณ เวลา t

โดยข้อมูลมูลค่าการเอาประกันวินาศภัยที่ใช้ในการศึกษาสามารถแบ่งออกเป็น 3 ประเภท ดังนี้

1. การประกันอัคคีภัย (Fire insurance)
2. การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance)
3. การประกันภัยรถยนต์ (Automobile insurance)

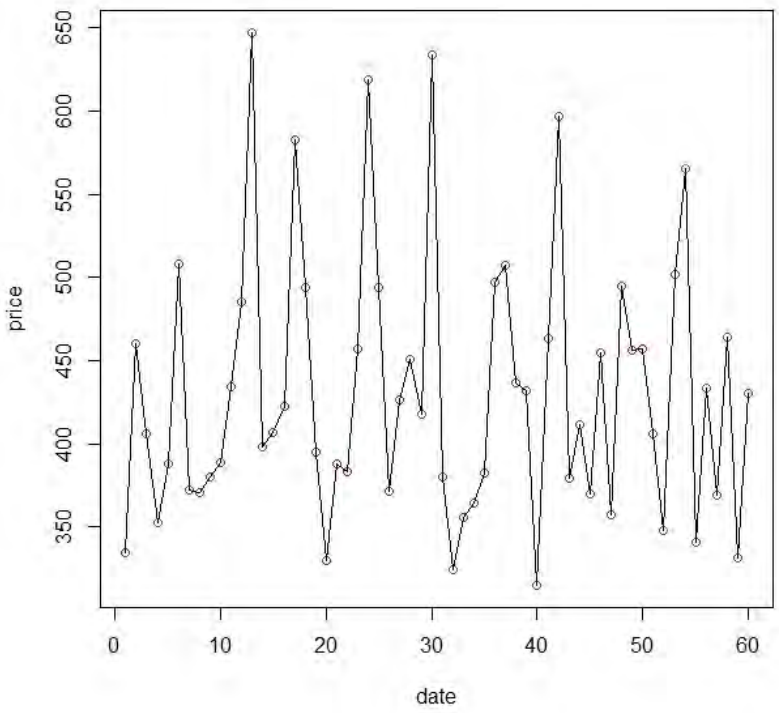
1) การประกันอัคคีภัย (fire insurance)

เป็นประกันที่คุ้มครองทรัพย์สินจากเหตุไฟไหม้ ฟ้าผ่า รวมถึงไฟฟ้าลัดวงจรที่เกิดจากฟ้าผ่า และการระเบิดของแก๊สหุงต้มหรือแก๊สที่ใช้ทำแสงสว่างเพื่อประโยชน์ในการอยู่อาศัย ภัยจากการเฉี่ยวชนของยานพาหนะ โดยสามารถเอาประกันได้ทั้งสังหาริมทรัพย์ เช่น เฟอร์นิเจอร์ เครื่องใช้ไฟฟ้า เครื่องนุ่งห่ม สต๊อกสินค้า เครื่องจักร ฯลฯ และอสังหาริมทรัพย์ประเภทสิ่งปลูกสร้าง นอกจากนี้ผู้เอาประกันยังยังสามารถซื้อความคุ้มครองอื่นๆ ได้ เช่น ภัยน้ำท่วม ภัยลมพายุ หรือภัยแผ่นดินไหว เป็นต้น

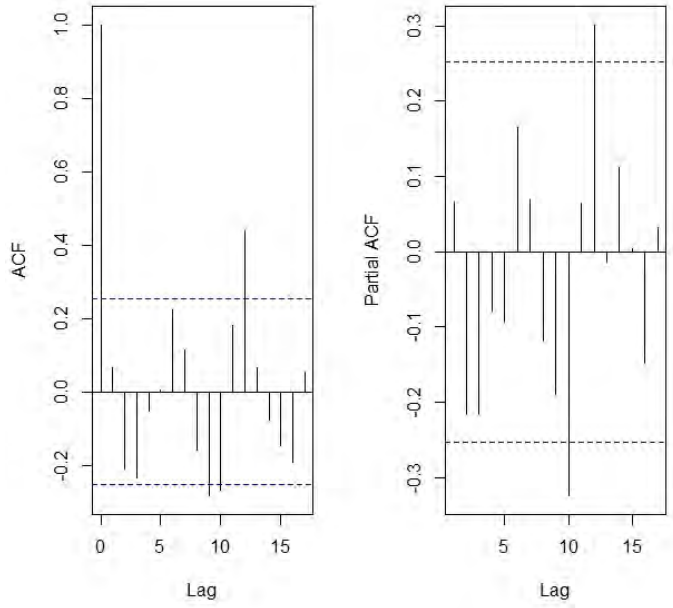
มูลค่าการเอาประกันอัคคีภัย และ อนุกรมเวลาของมูลค่าการเอาประกันอัคคีภัย ตั้งแต่ เดือนมกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 ค่าสหสัมพันธ์ในตัวเอง (ACF) ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ค่าสหสัมพันธ์ในตัวเอง (ACF) ของผลต่างอันดับ 1 และ ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ของผลต่างอันดับ 1 แสดงในตาราง 3.1 ภาพที่ 3.1 ภาพที่ 3.2 และ ภาพที่ 3.3 ตามลำดับ

วัน	ม.ค. 2555	ก.พ. 2555	มี.ค. 2555	เม.ย. 2555	พ.ค. 2555	มิ.ย. 2555
มูลค่าการเอาประกัน (พันล้านบาท)	334.2688	460.2963	405.842	352.4178	388.2161	507.9835
วัน	ก.ค. 2555	ส.ค. 2555	ก.ย. 2555	ต.ค. 2555	พ.ย. 2555	ธ.ค. 2555
มูลค่าการเอาประกัน (พันล้านบาท)	371.9625	370.4861	380.5023	388.3968	434.4513	485.628
⋮	⋮					
วัน	ก.ค. 2559	ส.ค. 2559	ก.ย. 2559	ต.ค. 2559	พ.ย. 2559	ธ.ค. 2559
มูลค่าการเอาประกัน (พันล้านบาท)	341.2226	433.8155	369.33	463.7448	331.4765	430.4088

ตาราง 3.1 แสดงมูลค่าการเอาประกันอัคคีภัยรายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559

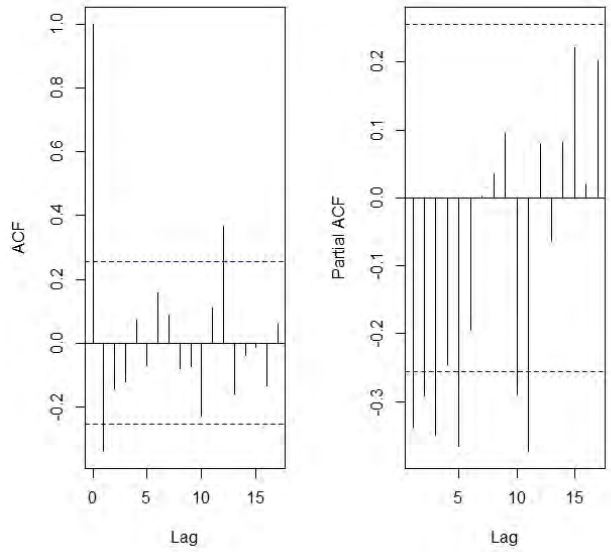


ภาพที่ 3.1 แสดงมูลค่าการเอาประกันอัคคีภัยระหว่างปี 2555 – 2559



ภาพที่ 3.2 แสดง ACF และ PACF ของมูลค่าการเอาประกันอัคคีภัย

จากภาพที่ 3.2 แสดง ACF และ PACF ของมูลค่าการเอาประกันอัคคีภัย เมื่อทำการทดสอบความนิ่งของมูลค่าการรับประกันอัคคีภัย ด้วยวิธี Augmented Dickey-Fuller Test โดยที่สมมติฐานหลักคือข้อมูลไม่คงที่ พบว่าค่า p-value มีค่ามากกว่า 0.01 จึงยอมรับสมมติฐานหลัก จึงสรุปได้ว่า ข้อมูลยังไม่นิ่ง



ภาพที่ 3.3 แสดง ACF และ PACF ของมูลค่าการเอาประกันอัคคีภัย โดยการหาผลต่างของข้อมูลครั้งที่ 1

จากภาพที่ 3.3 เมื่อทำการทดสอบความนิ่งของมูลค่าการเอาประกันอัคคีภัย ด้วยวิธี Augmented Dickey-Fuller Test พบว่าอนุกรมเวลาของผลต่างอันดับ 1 นิ่งและมีค่าสหสัมพันธ์ (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,12 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,3,5,10,11 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจึงพิจารณา $ARIMA(p,d,q)$ เมื่อ $p \in \{1,2,3,5,10,11\}$, $q \in \{1,12\}$ และ $d = 1$ ตามลำดับ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC น้อยที่สุด

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,1,1)	692.2800	ARIMA(3,1,1)	691.4265	ARIMA(10,1,1)	689.5868
ARIMA(1,1,12)	694.0888	ARIMA(3,1,12)	689.9963	ARIMA(10,1,12)	698.2324
ARIMA(2,1,1)	691.7654	ARIMA(5,1,1)	694.6677	**ARIMA(11,1,1)**	**687.6361**
ARIMA(2,1,12)	688.3474	ARIMA(5,1,12)	691.8478	ARIMA(11,1,12)	699.4741

ตารางที่ 3.2 แสดงค่า AIC ของแบบจำลองของมูลค่าการเอาประกันอัคคีภัย

จากตารางที่ 3.2 เมื่อพิจารณาค่า AIC เนื่องจากแบบจำลอง ARIMA(11,1,1) ให้ค่า AIC น้อยที่สุดคือ 687.6361 ดังนั้นแบบจำลองอนุกรมเวลาที่เหมาะสมของการเอาประกันอัคคีภัยคือ ARIMA(11,1,1) ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.3

พารามิเตอร์	ar1	ar2	ar3	ar4	ar5	ar6
ARIMA (11,1,1)	-0.6302	-0.6400	-0.5815	-0.4317	-0.5411	-0.3822
	ar7	ar8	ar9	ar10	ar11	ma1
	-0.356	-0.4466	-0.4794	-0.6374	-0.3501	-0.4725

ตารางที่ 3.3 พารามิเตอร์ของแบบจำลอง ARIMA (11,1,1)

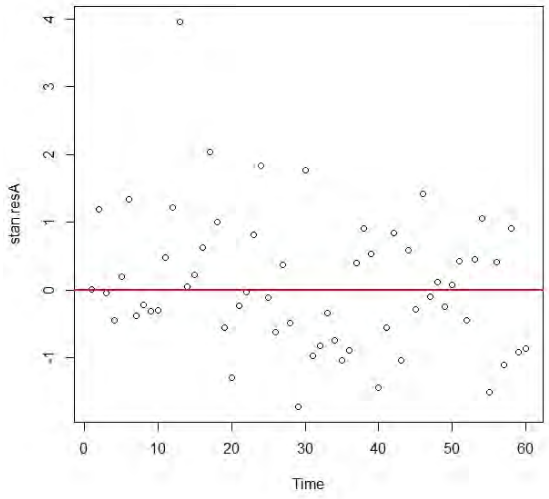
จากตารางที่ 3.3 จึงสรุปได้ว่าแบบจำลองอนุกรมเวลาที่เหมาะสมที่สุดสำหรับข้อมูลของมูลค่าการเอาประกันอัคคีภัย คือ แบบจำลอง ARIMA(11,1,1) ซึ่งสามารถเขียนประมาณพารามิเตอร์จากการทำผลต่างอันดับ 1 ได้สมการต่อไปนี้

$$\begin{aligned} y_t - y_{t-1} = & -0.6302(y_{t-1} - y_{t-2}) - 0.6400(y_{t-2} - y_{t-3}) - 0.5815(y_{t-3} - y_{t-4}) \\ & -0.4317(y_{t-4} - y_{t-5}) - 0.5411(y_{t-5} - y_{t-6}) - 0.3822(y_{t-6} - y_{t-7}) \\ & -0.3560(y_{t-7} - y_{t-8}) - 0.4466(y_{t-8} - y_{t-9}) - 0.4794(y_{t-9} - y_{t-10}) \\ & -0.6374(y_{t-10} - y_{t-11}) - 0.3501(y_{t-11} - y_{t-12}) + 0.4725\epsilon_{t-1} + \epsilon_t \end{aligned}$$

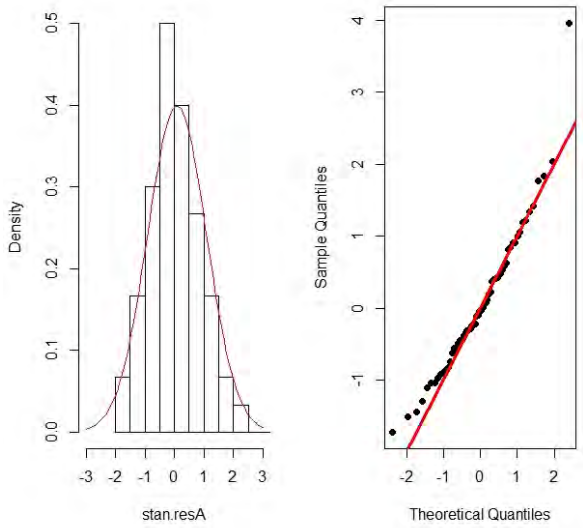
จัดรูปสมการต่อได้ดังนี้

$$\begin{aligned} y_t = & \ 0.3698y_{t-1} - 0.0098y_{t-2} + 0.0585y_{t-3} + 0.1498y_{t-4} - 0.1094y_{t-5} \\ & +0.1589y_{t-6} + 0.0262y_{t-7} - 0.0897y_{t-8} - 0.0328y_{t-9} - 0.158y_{t-10} \\ & +0.2873y_{t-11} + 0.3501y_{t-12} + 0.4725\epsilon_{t-1} + \epsilon_t \end{aligned}$$

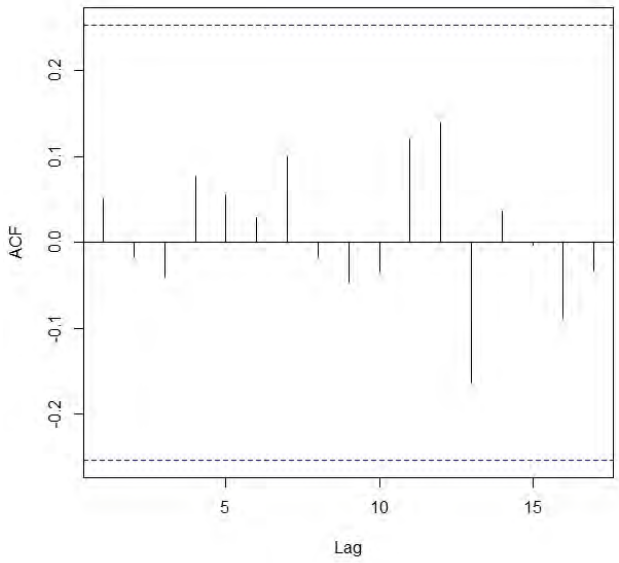
ต่อไปจะนำค่าเศษเหลือ (Residuals) ซึ่งคือค่าผลต่างของมูลค่าการเอาประกันอัคคีภัยจริงและมูลค่าการเอาประกันอัคคีภัยพยากรณ์มาทดสอบการแจกแจงแบบปกติโดยวิธีพิจารณากราฟของค่าเศษเหลือเทียบกับเวลา ดังแสดงในภาพที่ 3.4 ทดสอบการแจกแจงแบบปกติโดยวิธี Quantile – Quantile Plot หรือ QQ – plots และกราฟฮิสโตแกรม ดังแสดงในภาพที่ 3.5 ทดสอบความเป็นอิสระต่อกันของค่าเศษเหลือด้วยกราฟค่าสหสัมพันธ์ในตัวเอง ดังแสดงในภาพที่ 3.6 และทดสอบการแจกแจงปกติด้วยวิธี Kolmogorov-Smirnov normality test ดังแสดงในตารางที่ 3.4



ภาพที่ 3.4 แสดงค่าเศษเหลือของมูลค่าการประกันอัคคีภัยเทียบกับเวลาในแบบจำลอง ARIMA (11,1,1)



ภาพที่ 3.5 กราฟแสดงฮิสโตแกรมของเศษเหลือและ QQ-Plots ของมูลค่าการเอาประกันอัคคีภัยใน แบบจำลอง ARIMA (11,1,1)



ภาพที่ 3.6 แสดงค่าสหสัมพันธ์ในตัวเองของเศษเหลือของมูลค่าการเอาประกันอัคคีภัยในแบบจำลอง ARIMA (11,1,1)

แบบจำลอง	ค่า Kolmogorov-Smirnov	ค่า p-value
ARIMA(11,1,1)	0.072567	0.8874

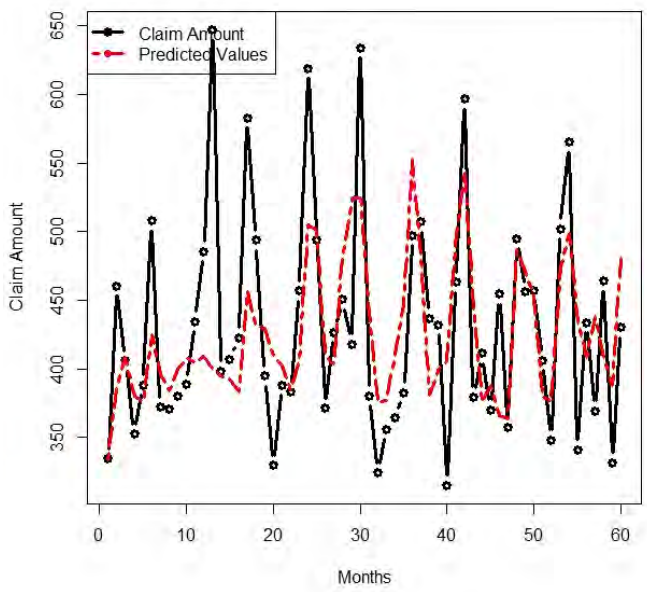
ตารางที่ 3.4 แสดงการทดสอบแบบจำลองด้วย Kolmogorov-Smirnov normality test ในแบบจำลอง ARIMA (11,1,1)

จากภาพที่ 3.4 การกระจายของเศษเหลือเทียบกับเวลาเป็นแบบสุ่ม ทำให้อนุกรมเวลาไม่มีแนวโน้ม จากภาพที่ 3.5 กราฟฮิสโตแกรมของเศษเหลือมีการแจกแจงเป็นรูประฆังคว่ำคล้ายการแจกแจงแบบปกติ และการทดสอบโดยวิธี QQ – plots พบว่ามีการกระจายของควอนไทล์เป็นเส้นตรงมีลักษณะคล้ายการแจกแจงแบบปกติ จากภาพที่ 3.6 แบบจำลองแสดงให้เห็นว่าค่าสหสัมพันธ์ในตัวเองของเศษเหลือไม่มีความสัมพันธ์กันในทุก lag ที่มากกว่า 0 และจากตารางที่ 3.4 เมื่อทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov normality test พบว่า ค่า p-value เท่ากับ 0.8874 ซึ่งมากกว่า 0.05 แสดงว่าเศษเหลือของแบบจำลองมีการแจกแจงแบบปกติ เราจึงสรุปได้ว่าเศษเหลือของแบบจำลอง ARIMA (11,1,1) มีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติ

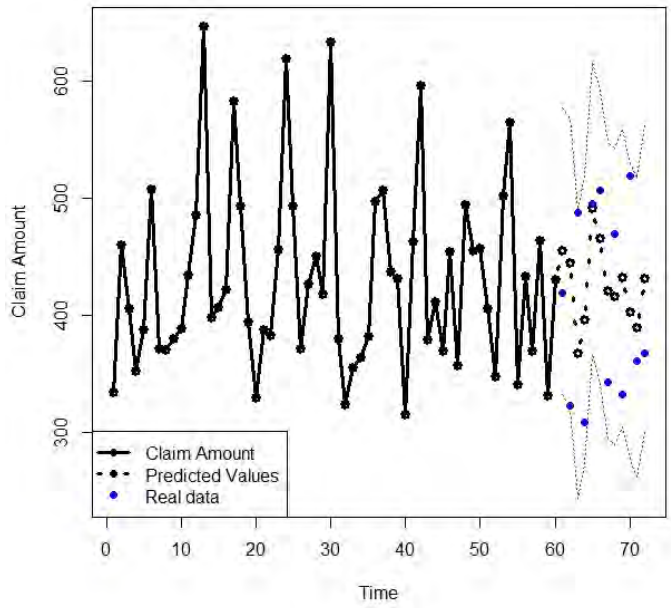
ขั้นตอนต่อไปจะเป็นการหามูลค่าพยากรณ์ของการเอาประกันอัคคีภัยของแบบจำลอง ARIMA (11,1,1) เทียบกับมูลค่าการเอาประกันอัคคีภัยจริงในอีก 12 เดือนหลังจากเดือนธันวาคม 2559 แสดงเป็นค่าในตารางที่ 3.5 กราฟของมูลค่าพยากรณ์ของการเอาประกันอัคคีภัยของแบบจำลอง ARIMA (11,1,1) เทียบกับมูลค่าการเอาประกันอัคคีภัยจริง ดังแสดงในภาพที่ 3.7 และ แสดงผลพยากรณ์ของมูลค่าการเอาประกันอัคคีภัยระหว่างเดือนมกราคม 2560 ถึงเดือนธันวาคม 2560 ดังแสดงในภาพที่ 3.8

วัน	มูลค่าการเอาประกันภัยจริง (พันล้านบาท)	ผลพยากรณ์ (พันล้านบาท)
ม.ค. 2560	418.9132	455.2769
ก.พ. 2560	322.8114	444.6380
มี.ค. 2560	488.0167	367.5620
เม.ย. 2560	308.6397	396.1330
พ.ค. 2560	495.9461	491.3500
มิ.ย. 2560	506.7226	466.3619
ก.ค. 2560	343.0739	421.3648
ส.ค. 2560	470.2843	416.5533
ก.ย. 2560	333.0318	432.3154
ต.ค. 2560	519.3943	403.4835
พ.ย. 2560	361.5313	389.7877
ธ.ค. 2560	367.5239	432.0517

ตารางที่ 3.5 แสดงมูลค่าการเอาประกันอัคคีภัยจริงเทียบกับมูลค่าพยากรณ์ของการเอาประกันอัคคีภัย



ภาพที่ 3.7 แสดงกราฟของมูลค่าการเอาประกันอัคคีภัยจริงเทียบกับมูลค่าพยากรณ์ของการเอาประกันอัคคีภัย



ภาพที่ 3.8 แสดงผลพยากรณ์ของมูลค่าการเอาประกันอัคคีภัยระหว่างเดือนมกราคม 2560 ถึงเดือนธันวาคม 2560

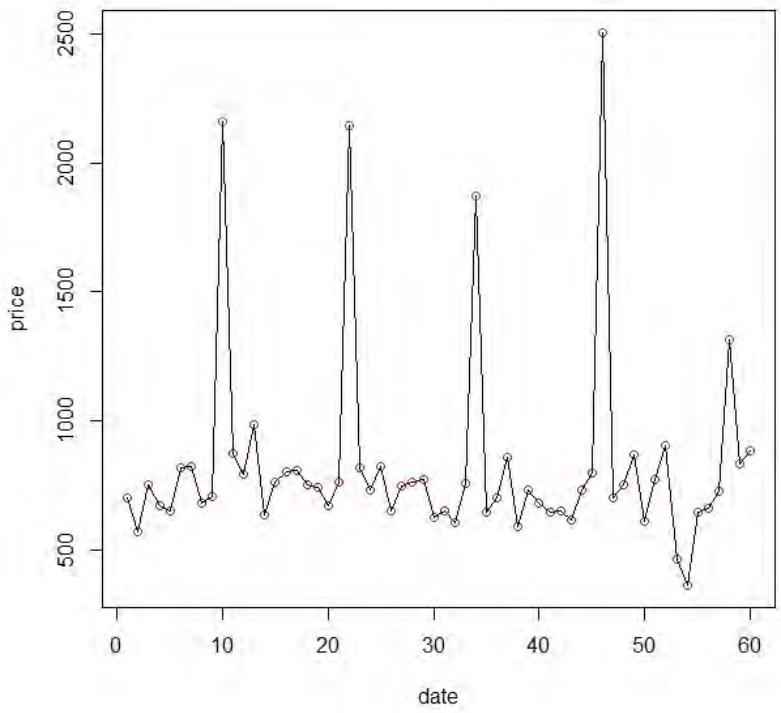
2) การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance)

เป็นประกันที่คุ้มครองความเสียหายแก่เรือและทรัพย์สินหรือสินค้าที่อยู่ในระหว่างการขนส่งทางทะเล แบ่งเป็นการประกันภัยตัวเรือและเครื่องจักร ที่อาจได้รับความเสียหายจากอุบัติเหตุต่างๆ เช่น เรือเกยตื้น เรือชนกัน เรือชนหิน หรืออุบัติเหตุอันเกิดจากภัยพิบัติทางธรรมชาติ การเกิดอัคคีภัย เป็นต้น รวมไปถึงการประกันภัยขนส่งสินค้า ในกรณีที่เรือเกิดอุบัติเหตุ การโจรกรรม หรือการทิ้งของลงทะเลในกรณีฉุกเฉิน เป็นต้น

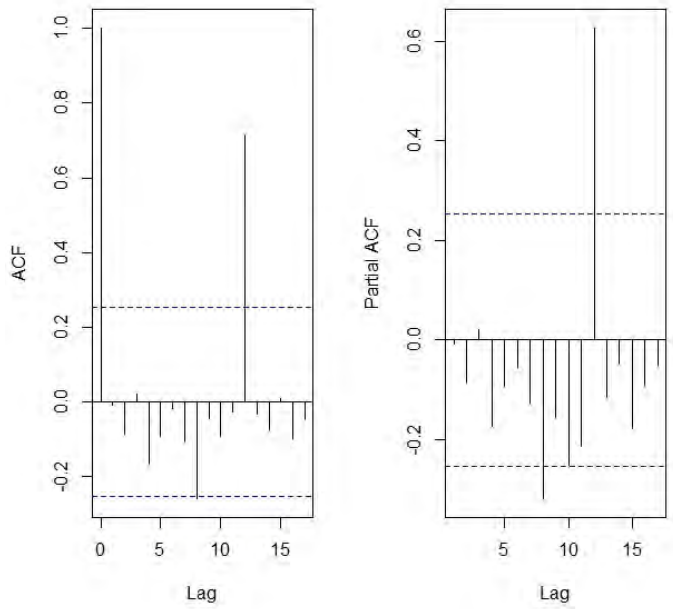
มูลค่าการเอาประกันภัยทางทะเลและขนส่งและอนุกรมเวลาของมูลค่าการเอสประกันภัยทางทะเลและขนส่ง ตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึงเดือนธันวาคม พ.ศ. 2559 ค่าสหสัมพันธ์ในตัวเอง (ACF) ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ค่าสหสัมพันธ์ในตัวเอง (ACF) และค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ของผลต่างอันดับ 1 แสดงในตาราง 3.6 ภาพที่ 3.9 ภาพที่ 3.10 และ ภาพที่ 3.11 ตามลำดับ

วัน	ม.ค. 2555	ก.พ. 2555	มี.ค. 2555	เม.ย. 2555	พ.ค. 2555	มิ.ย. 2555
มูลค่าการเอาประกัน (พันล้านบาท)	701.9927	570.1779	749.2588	671.7707	648.2885	817.497
วัน	ก.ค. 2555	ส.ค. 2555	ก.ย. 2555	ต.ค. 2555	พ.ย. 2555	ธ.ค. 2555
มูลค่าการเอาประกัน (พันล้านบาท)	819.6443	679.9884	707.7069	2160.878	874.9204	789.6109
⋮	⋮					
วัน	ก.ค. 2559	ส.ค. 2559	ก.ย. 2559	ต.ค. 2559	พ.ย. 2559	ธ.ค. 2559
มูลค่าการเอาประกัน (พันล้านบาท)	646.2631	660.9032	728.1008	1313.309	834.5041	881.3963

ตาราง 3.6 แสดงมูลค่าการเอาประกันภัยทางทะเลและขนส่งรายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ.2559

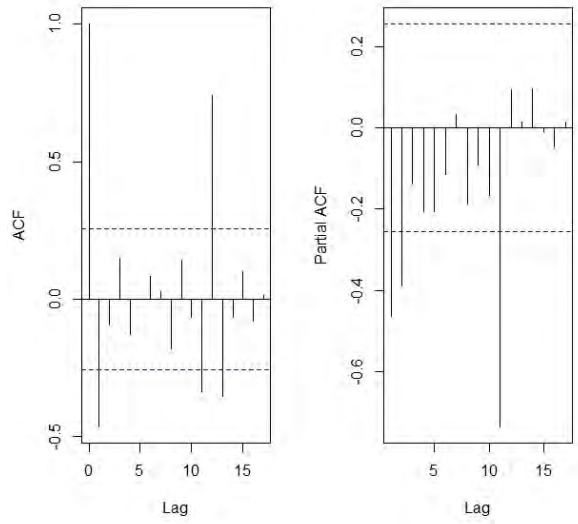


ภาพที่ 3.9 แสดงมูลค่าการเอาประกันภัยทางทะเลและขนส่ง ระหว่างปี 2555 – 2559



ภาพที่ 3.10 แสดง ACF และ PACF ของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง

จากภาพที่ 3.10 แสดง ACF และ PACF ของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง เมื่อทำการทดสอบความนิ่งของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง ด้วยวิธี Augmented Dickey-Fuller Test โดยที่สมมติฐานหลักคือข้อมูลไม่คงที่ พบว่าค่า p-value มีค่ามากกว่า 0.01 จึงยอมรับสมมติฐานหลัก จึงสรุปได้ว่า ข้อมูลยังไม่นิ่ง



ภาพที่ 3.11 แสดง ACF และ PACF ของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง โดยการหาผลต่างของข้อมูลครั้งที่ 1

จากภาพที่ 3.11 เมื่อทำการทดสอบความนิ่งของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง ด้วยวิธี Augmented Dickey-Fuller Test พบว่าอนุกรมเวลาของผลต่างอันดับ 1 นิ่งและมีค่าสหสัมพันธ์ (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,11,12,13 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,11 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจึงพิจารณา $ARIMA(p,d,q)$ เมื่อ $p \in \{1,2,11\}$, $q \in \{1,11,12,13\}$ และ $d = 1$ ตามลำดับ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC น้อยที่สุด

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,1,1)	880.9390	ARIMA(2,1,1)	882.6603	**ARIMA(11,1,1)**	**837.2113**
ARIMA(1,1,11)	877.6935	ARIMA(2,1,11)	878.1250	ARIMA(11,1,11)	839.5922
ARIMA(1,1,12)	874.8378	ARIMA(2,1,12)	871.8563	ARIMA(11,1,12)	848.9655
ARIMA(1,1,13)	865.6604	ARIMA(2,1,13)	867.4488	ARIMA(11,1,13)	845.1647

ตารางที่ 3.7 แสดงค่า AIC ของแบบจำลองของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง

จากตารางที่ 3.7 เมื่อพิจารณาค่า AIC เนื่องจากแบบจำลอง ARIMA (11,1,1) ให้ค่า AIC น้อยที่สุดคือ 839.5922 ดังนั้นแบบจำลองอนุกรมเวลาที่เหมาะสมของการเอาประกันภัยทางทะเลและขนส่งคือ ARIMA(11,1,1) ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.8

พารามิเตอร์	ar1	ar2	ar3	ar4	ar5	ar6
ARIMA (11,1,1)	-0.9782	-0.9548	-0.8848	-0.8304	-0.7960	-0.7822
	ar7	ar8	ar9	ar10	ar11	ma1
	-0.7987	-0.8675	-0.8819	-0.8789	-0.8409	0.0268

ตารางที่ 3.8 พารามิเตอร์ของแบบจำลอง ARIMA (11,1,1)

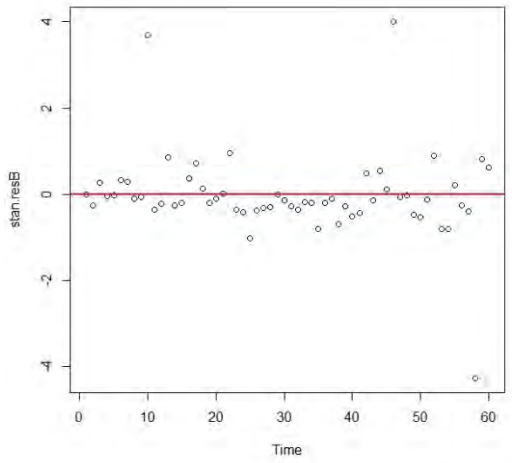
จากตารางที่ 3.8 จึงสรุปได้ว่าแบบจำลองอนุกรมเวลาที่เหมาะสมที่สุดสำหรับข้อมูลของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง คือ แบบจำลอง ARIMA (11,1,1) ซึ่งสามารถเขียนประมาณพารามิเตอร์จากการทำผลต่างอันดับ 1 ได้สมการต่อไปนี้

$$\begin{aligned} y_t - y_{t-1} = & -0.9782(y_{t-1} - y_{t-2}) - 0.9548(y_{t-2} - y_{t-3}) - 0.8848(y_{t-3} - y_{t-4}) \\ & -0.8304(y_{t-4} - y_{t-5}) - 0.7960(y_{t-5} - y_{t-6}) - 0.7822(y_{t-6} - y_{t-7}) \\ & -0.7987(y_{t-7} - y_{t-8}) - 0.8675(y_{t-8} - y_{t-9}) - 0.8819(y_{t-9} - y_{t-10}) \\ & -0.8789(y_{t-10} - y_{t-11}) - 0.8409(y_{t-11} - y_{t-12}) - 0.0268\epsilon_{t-1} + \epsilon_t \end{aligned}$$

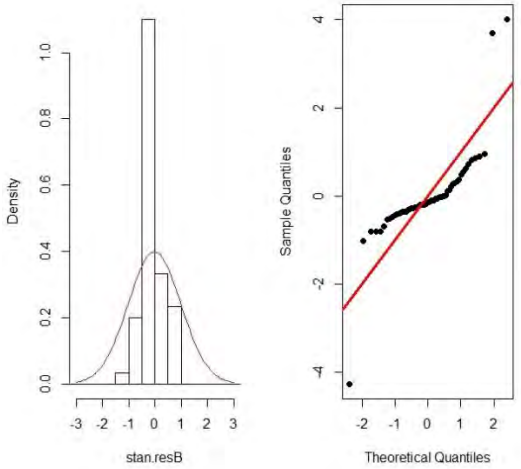
จัดรูปสมการต่อได้ดังนี้

$$\begin{aligned} y_t = & \ 0.0218y_{t-1} + 0.0234y_{t-2} + 0.07y_{t-3} + 0.0544y_{t-4} - 0.0344y_{t-5} \\ & +0.0138y_{t-6} - 0.0164y_{t-7} - 0.0689y_{t-8} - 0.0144y_{t-9} \\ & -0.003y_{t-10} - 0.038y_{t-11} + 0.8409y_{t-12} - 0.0268\epsilon_{t-1} + \epsilon_t \end{aligned}$$

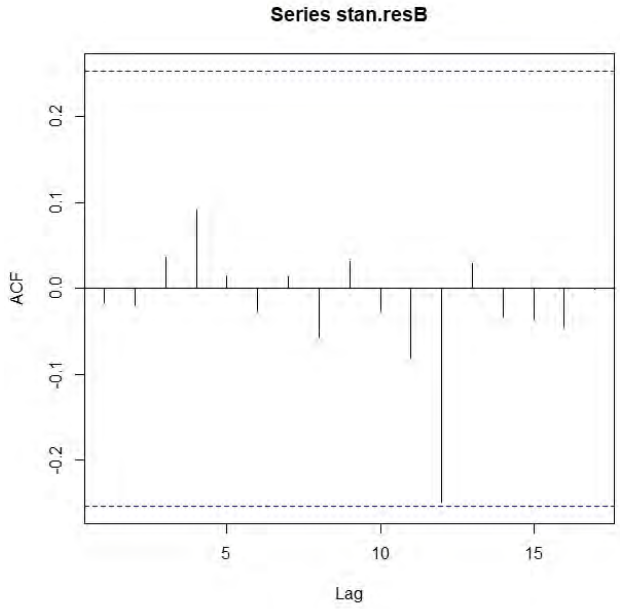
ต่อไปจะนำค่าเศษเหลือ (Residuals) ซึ่งคือค่าผลต่างของมูลค่าการเอาประกันภัยทางทะเลและขนส่งจริงและมูลค่าการเอาประกันภัยทางทะเลและขนส่งพยากรณ์มาทดสอบการแจกแจงแบบปกติโดยวิธีพิจารณากราฟของค่าเศษเหลือเทียบกับเวลา ดังแสดงในภาพที่ 3.12 ทดสอบการแจกแจงแบบปกติโดยวิธี Quantile – Quantile Plot หรือ QQ – plots และกราฟฮิสโตแกรม ดังแสดงในภาพที่ 3.13 ทดสอบความเป็นอิสระต่อกันของค่าเศษเหลือด้วยกราฟค่าสหสัมพันธ์ในตัวเองของเศษเหลือ ดังแสดงในภาพที่ 3.14 และทดสอบการแจกแจงปกติด้วยวิธี Kolmogorov-Smirnov normality test ดังแสดงในตารางที่ 3.9



ภาพที่ 3.12 แสดงค่าเศษเหลือของมูลค่าการเอาประกันภัยทางทะเลและขนส่งเทียบกับเวลาในแบบจำลอง ARIMA (11,1,1)



ภาพที่ 3.13 กราฟแสดงฮิสโตแกรมของเศษเหลือและ QQ-Plots ของมูลค่าการเอาประกันภัยทางทะเลและขนส่งในแบบจำลอง ARIMA (11,1,1)



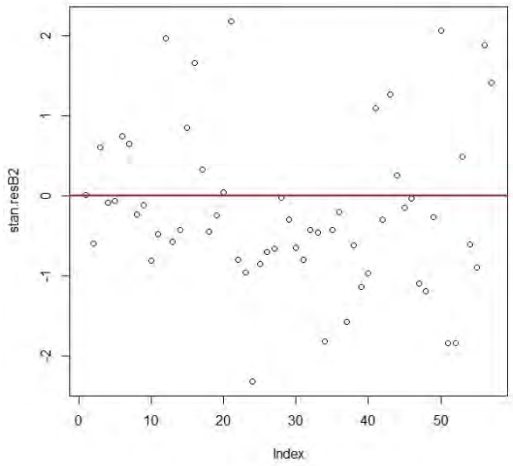
ภาพที่ 3.14 แสดงค่าสหสัมพันธ์ในตัวเองของเศษเหลือของการเอาประกันภัยทางทะเลและขนส่งในแบบจำลอง ARIMA (11,1,1)

แบบจำลอง	ค่า Kolmogorov-Smirnov	ค่า p-value
ARIMA(11,1,1)	0.20751	0.009611

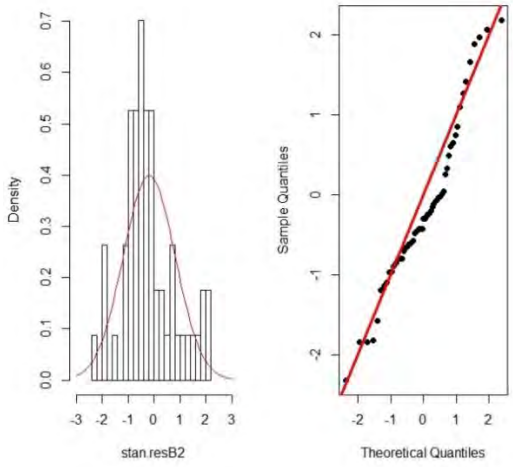
ตารางที่ 3.9 แสดงการทดสอบแบบจำลองด้วย Kolmogorov-Smirnov normality test ในแบบจำลอง ARIMA (11,1,1)

จากภาพที่ 3.12 การกระจายของเศษเหลือเทียบกับเวลาไม่เป็นแบบสุ่ม ทำให้อนุกรมเวลามีแนวโน้ม จากภาพที่ 3.13 กราฟฮิสโตแกรมของเศษเหลือมีการแจกแจงเป็นรูประฆังคว่ำคล้ายการแจกแจงแบบปกติ และการทดสอบโดยวิธี QQ – plots พบว่ามีการกระจายของควอนไทล์เป็นเส้นตรงมีลักษณะไม่คล้ายการแจกแจงแบบปกติ จากภาพที่ 3.14 แบบจำลองแสดงให้เห็นว่าค่าสหสัมพันธ์ในตัวเองของเศษเหลือไม่มีความสัมพันธ์กันในทุก lag ที่มากกว่า 0 และจากตารางที่ 3.9 เมื่อทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov normality test พบว่า ค่า p-value เท่ากับ 0.009611 ซึ่งน้อยกว่า 0.05 แสดงว่าเศษเหลือของแบบจำลองไม่มีการแจกแจงแบบปกติ เราจึงสรุปได้ว่าเศษเหลือของแบบจำลอง ARIMA (11,1,1) ไม่มีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติ

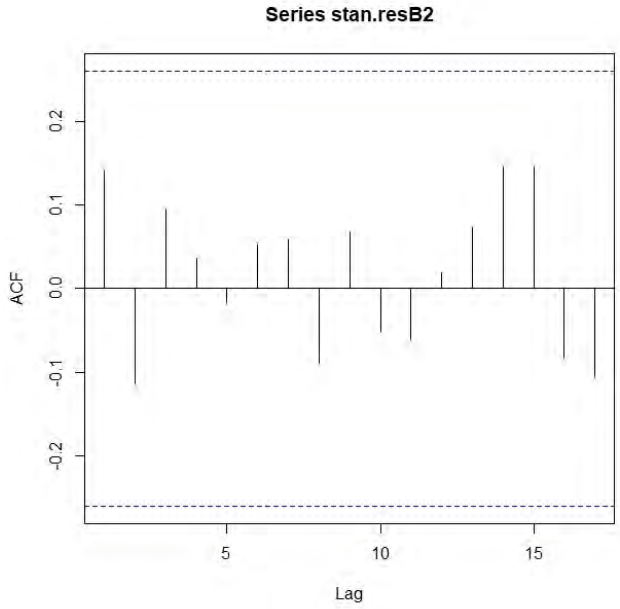
แต่เนื่องจากค่าเศษเหลือ (Residuals) บางตัวมีค่าผิดปกติ (Outliers) คือมีข้อมูลที่มีค่าแยกออกจากกลุ่มหรือผิดแผกแตกต่างไปจากข้อมูลค่าอื่นๆทำให้ผลการวิเคราะห์คลาดเคลื่อน จึงได้ขจัดข้อมูลที่ผิดปกติออกไป แล้วนำข้อมูลค่าเศษเหลือที่ได้ใหม่มาทดสอบการแจกแจงแบบปกติโดยวิธีพิจารณากราฟของค่าเศษเหลือเทียบกับเวลา ดังแสดงในภาพที่ 3.15 ทดสอบการแจกแจงแบบปกติโดยวิธี Quantile – Quantile Plot และกราฟฮิสโตแกรม ดังแสดงในภาพที่ 3.16 ทดสอบความเป็นอิสระต่อกันของค่าเศษเหลือด้วยกราฟค่าสหสัมพันธ์ในตัวเองของเศษเหลือ ดังแสดงในภาพที่ 3.17 และทดสอบการแจกแจงปกติด้วยวิธี Kolmogorov-Smirnov normality test ดังแสดงในตารางที่ 3.10



ภาพที่ 3.15 แสดงค่าเศษเหลือที่ขจัดข้อมูลผิดปกติของมูลค่าการเอาประกันภัยทางทะเลและขนส่งเทียบกับเวลาในแบบจำลอง ARIMA (11,1,1)



ภาพที่ 3.16 กราฟแสดงฮิสโตแกรมของเศษเหลือที่ขจัดข้อมูลผิดปกติและ QQ-Plots ของมูลค่าการเอาประกันภัยทางทะเลและขนส่งในแบบจำลอง ARIMA (11,1,1)



ภาพที่ 3.17 แสดงค่าสหสัมพันธ์ในตัวเองของเศษเหลือที่ขจัดข้อมูลผิดปกติของการเอาประกันภัยทางทะเลและขนส่งในแบบจำลอง ARIMA (11,1,1)

แบบจำลอง	ค่า Kolmogorov-Smirnov	ค่า p-value
ARIMA(11,1,1)	0.14616	0.1584

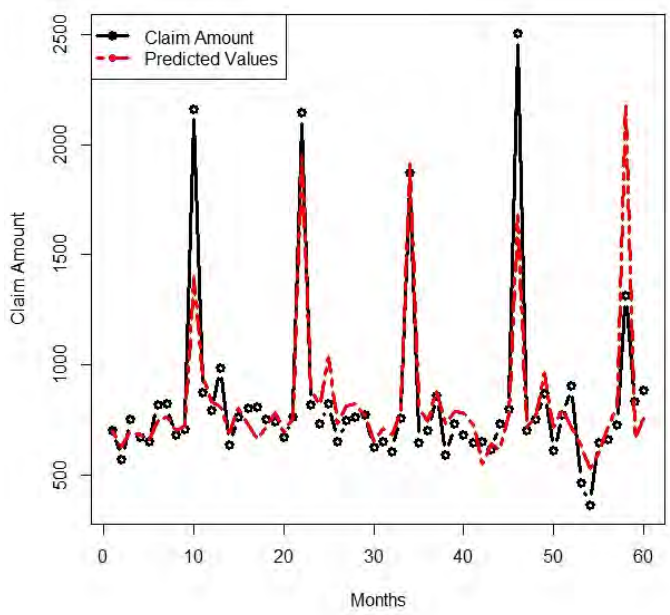
ตารางที่ 3.10 แสดงการทดสอบแบบจำลองด้วย Kolmogorov-Smirnov normality test ของเศษเหลือที่ขจัดข้อมูลผิดปกติในแบบจำลอง ARIMA (11,1,1)

จากภาพที่ 3.15 การกระจายของเศษเหลือเทียบกับเวลาเป็นแบบสุ่ม ทำให้อนุกรมเวลาไม่มีแนวโน้ม จากภาพที่ 3.16 กราฟฮิสโตแกรมของเศษเหลือมีการแจกแจงเป็นรูประฆังคว่ำคล้ายการแจกแจงแบบปกติ และการทดสอบโดยวิธี QQ – plots พบว่ามีการกระจายของควอนไทล์เป็นเส้นตรงมีลักษณะคล้ายการแจกแจงแบบปกติ จากภาพที่ 3.17 แบบจำลองแสดงให้เห็นว่าค่าสหสัมพันธ์ในตัวเองของเศษเหลือไม่มีความสัมพันธ์กันในทุก lag ที่มากกว่า 0 และจากตารางที่ 3.10 เมื่อทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov normality test พบว่า ค่า p-value เท่ากับ 0.1584 ซึ่งมากกว่า 0.05 แสดงว่าเศษเหลือของแบบจำลองมีการแจกแจงแบบปกติ เราจึงสรุปได้ว่าเศษเหลือของแบบจำลอง ARIMA (11,1,1) มีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติ

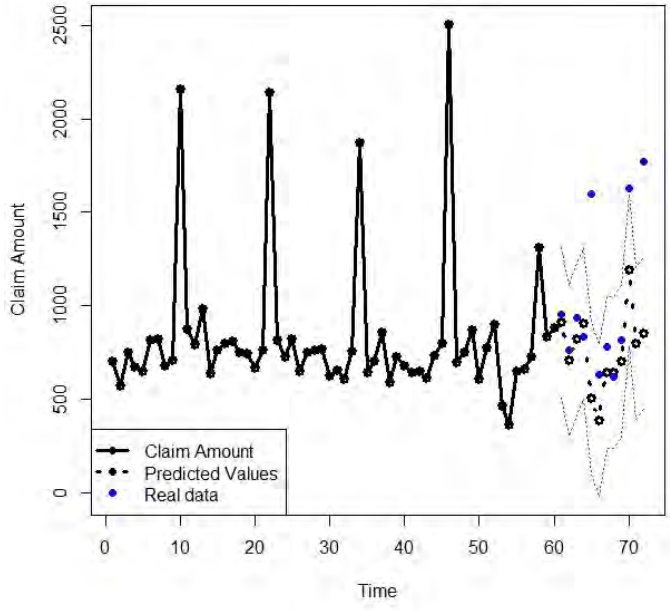
ขั้นตอนต่อไปจะเป็นการหามูลค่าพยากรณ์ของการเอาประกันภัยทางทะเลและขนส่งของแบบจำลอง ARIMA (11,1,1) เทียบกับมูลค่าการเอาประกันภัยทางทะเลและขนส่งจริงในอีก 12 เดือนหลังจากเดือนธันวาคม 2559 แสดงเป็นค่าในตารางที่ 3.11 กราฟของมูลค่าพยากรณ์ของการเอาประกันภัยทางทะเลและขนส่งของแบบจำลอง ARIMA (11,1,1) เทียบกับมูลค่าการเอาประกันภัยทางทะเลและขนส่งจริง ดังแสดงในภาพที่ 3.18 และแสดงผลพยากรณ์ของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง ระหว่างเดือนมกราคม 2560 ถึงเดือนธันวาคม 2560 ดังแสดงในภาพ 3.19

วัน	มูลค่าการเอาประกันภัยจริง (พันล้านบาท)	ผลพยากรณ์ (พันล้านบาท)
ม.ค. 2560	950.3221	909.0324
ก.พ. 2560	763.6168	706.0147
มี.ค. 2560	932.6158	822.4903
เม.ย. 2560	831.9374	903.1254
พ.ค. 2560	1,599.0215	502.8480
มิ.ย. 2560	633.2804	388.2095
ก.ค. 2560	782.0365	644.8784
ส.ค. 2560	619.0734	645.2547
ก.ย. 2560	814.3700	703.8492
ต.ค. 2560	1,626.8928	1,189.9427
พ.ย. 2560	22,767.5335	799.2709
ธ.ค. 2560	1,769.4363	852.7087

ตารางที่ 3.11 แสดงมูลค่าการเอาประกันภัยทางทะเลและขนส่งจริงเทียบกับมูลค่าพยากรณ์ของการเอาประกันภัยทางทะเลและขนส่ง



ภาพที่ 3.18 แสดงกราฟของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง
จริงเทียบกับมูลค่าพยากรณ์ของการเอาประกันภัยทางทะเลและขนส่ง



ภาพที่ 3.19 แสดงผลพยากรณ์ของมูลค่าการเอาประกันภัยทางทะเลและขนส่ง
ระหว่างเดือนมกราคม 2560 ถึงเดือนธันวาคม 2560

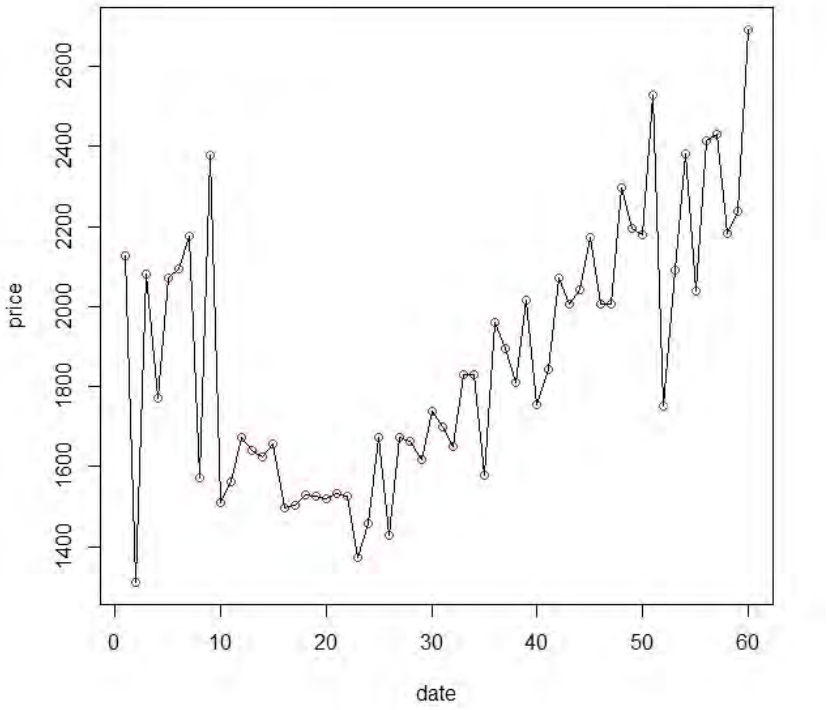
3) การประกันภัยรถยนต์ (Automobile insurance)

เป็นประกันที่ให้ความคุ้มครองความเสียหายที่เกิดจากการใช้รถยนต์ ได้แก่ ความเสียหายต่อตัวรถ ความเสียหายต่อทรัพย์สินของบุคคลภายนอก ให้ความคุ้มครองผู้ขับขี่รถยนต์ ผู้โดยสาร รวมถึงบุคคลภายนอกที่ได้รับบาดเจ็บหรือเสียชีวิต

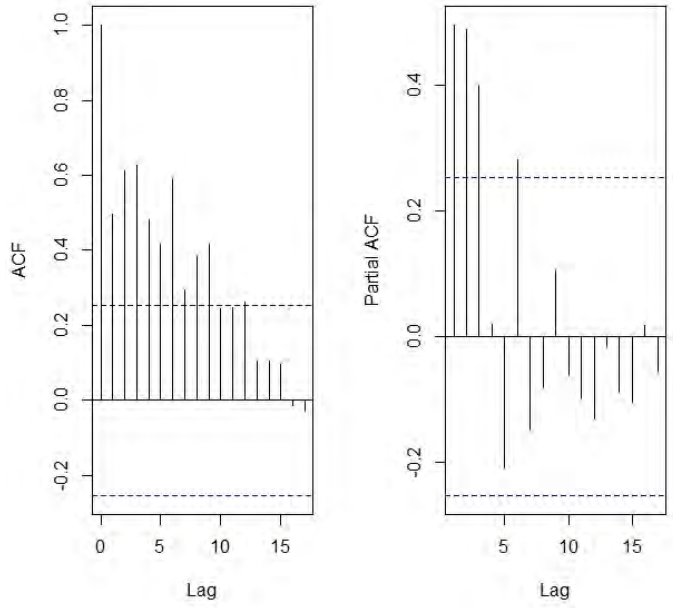
มูลค่าการเอาประกันภัยรถยนต์ และ อนุกรมเวลาของมูลค่าการเอาประกันภัยรถยนต์ ตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 ค่าสหสัมพันธ์ในตัวเอง (ACF) ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ค่าสหสัมพันธ์ในตัวเอง (ACF) ของผลต่างอันดับ 1 และ ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ของผลต่างอันดับ 1 แสดงในตาราง 3.12 ภาพที่ 3.20 ภาพที่ 3.21 และ ภาพที่ 3.22 ตามลำดับ

วัน	ม.ค. 2555	ก.พ. 2555	มี.ค. 2555	เม.ย. 2555	พ.ค. 2555	มิ.ย. 2555
มูลค่าการเอาประกัน (พันล้านบาท)	2126.3829	1312.0906	2081.0825	1770.8082	2070.4201	2095.4043
วัน	ก.ค. 2555	ส.ค. 2555	ก.ย. 2555	ต.ค. 2555	พ.ย. 2555	ธ.ค. 2555
มูลค่าการเอาประกัน (พันล้านบาท)	2177.6372	1572.4215	2377.6414	1509.757	1563.5972	1672.2114
⋮	⋮					
วัน	ก.ค. 2559	ส.ค. 2559	ก.ย. 2559	ต.ค. 2559	พ.ย. 2559	ธ.ค. 2559
มูลค่าการประกัน (พันล้านบาท)	2037.8689	2413.6491	2430.7417	2183.812	2238.7146	2691.5764

ตาราง 3.12 แสดงมูลค่าการเอาประกันภัยรถยนต์รายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559

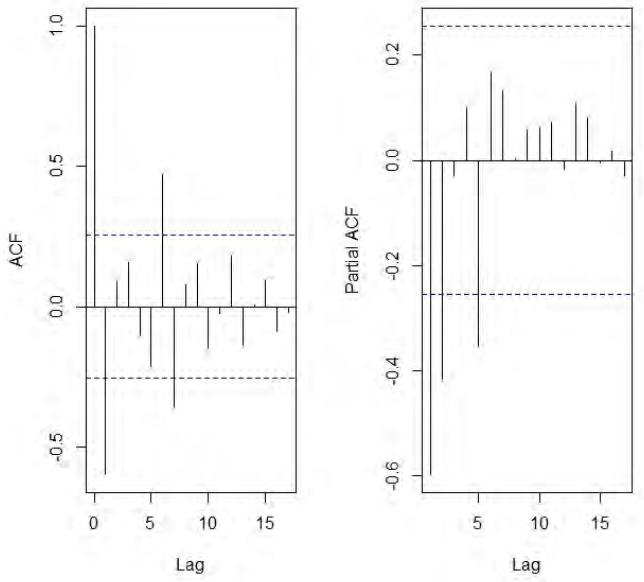


ภาพที่ 3.20 แสดงมูลค่าการเอาประกันภัยรถยนต์ระหว่างปี 2555 – 2559



ภาพที่ 3.21 แสดง ACF และ PACF ของมูลค่าการเอาประกันภัยรถยนต์

จากภาพที่ 3.21 แสดง ACF และ PACF ของมูลค่าการเอาประกันภัยรถยนต์ เมื่อทำการทดสอบความนิ่งของมูลค่าการเอาประกันภัยรถยนต์ ด้วยวิธี Augmented Dickey-Fuller Test โดยที่สมมติฐานหลักคือ ข้อมูลไม่คงที่ พบว่าค่า p-value มีค่ามากกว่า 0.01 จึงยอมรับสมมติฐานหลัก จึงสรุปได้ว่า ข้อมูลยังไม่นิ่ง



ภาพที่ 3.22 แสดง ACF และ PACF ของมูลค่าการเอาประกันภัยรถยนต์ โดยการหาผลต่างของข้อมูลครั้งที่ 1

จากภาพที่ 3.22 เมื่อทำการทดสอบความนิ่งของมูลค่าการเอาประกันภัยรถยนต์ ด้วยวิธี Augmented Dickey-Fuller Test พบว่าอนุกรมเวลาของผลต่างอันดับ 1 นิ่งและมีค่าสหสัมพันธ์ (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,6,7 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,5 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจึงพิจารณา $ARIMA(p,d,q)$ เมื่อ $p \in \{1,2,5\}$, $q \in \{1,6,7\}$ และ $d = 1$ ตามลำดับ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC น้อยที่สุด

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,1,1)	805.8530	ARIMA(2,1,1)	803.0663	**ARIMA(5,1,1)**	**794.5572**
ARIMA(1,1,6)	803.0990	ARIMA(2,1,6)	804.0728	ARIMA(5,1,6)	797.4965
ARIMA(1,1,7)	803.6498	ARIMA(2,1,7)	805.4049	ARIMA(5,1,7)	799.6322

ตารางที่ 3.13 แสดงค่า AIC ของแบบจำลองของมูลค่าการเอาประกันภัยรถยนต์

จากตารางที่ 3.13 เมื่อพิจารณาค่า AIC เนื่องจากแบบจำลอง ARIMA (5,1,1) ให้ค่า AIC น้อยที่สุดคือ 794.5572 ดังนั้นแบบจำลองอนุกรมเวลาที่เหมาะสมของการเอาประกันภัยรถยนต์คือ ARIMA(5,1,1) ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.14

พารามิเตอร์	ar1	ar2	ar3
ARIMA (5,1,1)	-1.1021	-0.5315	-0.0605
	ar4	ar5	ma1
	-0.1166	-0.3939	0.2158

ตารางที่ 3.14 พารามิเตอร์ของแบบจำลอง ARIMA (5,1,1)

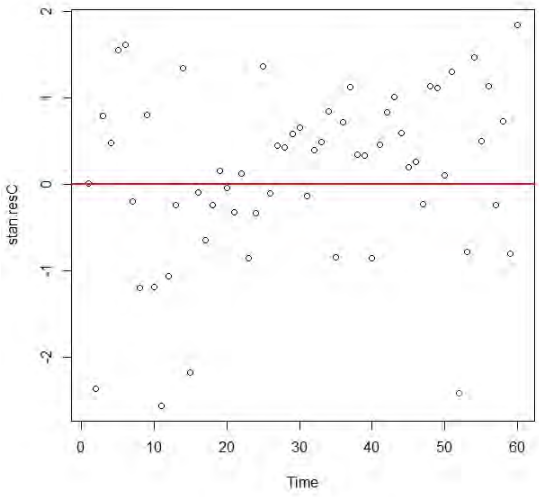
จากตารางที่ 3.14 จึงสรุปได้ว่าแบบจำลองอนุกรมเวลาที่เหมาะสมที่สุดสำหรับข้อมูลของมูลค่าการเอาประกันภัยรถยนต์ คือ แบบจำลอง ARIMA (5,1,1) ซึ่งสามารถเขียนประมาณพารามิเตอร์จากการทำผลต่างอันดับ 1 ได้สมการต่อไปนี้

$$y_t - y_{t-1} = -1.1021(y_{t-1} - y_{t-2}) - 0.5315(y_{t-2} - y_{t-3}) - 0.0605(y_{t-3} - y_{t-4}) \\ -0.1166(y_{t-4} - y_{t-5}) - 0.3939(y_{t-5} - y_{t-6}) - 0.2158\epsilon_{t-1} + \epsilon_t$$

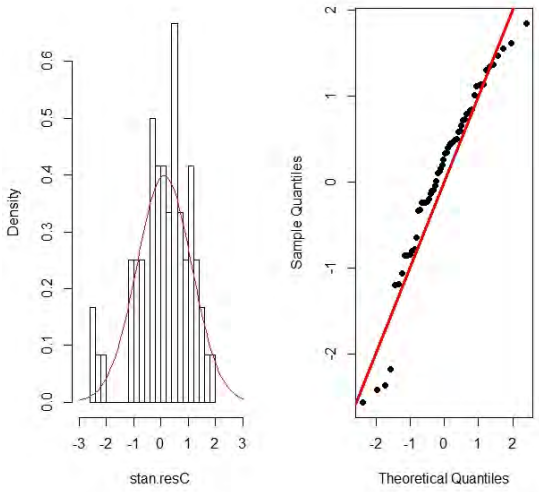
จัดรูปสมการต่อได้ดังนี้

$$y_t = -0.1021y_{t-1} + 0.5706y_{t-2} + 0.4710y_{t-3} - 0.0561y_{t-4} - 0.2773y_{t-5} \\ +0.3939y_{t-6} - 0.2158\epsilon_{t-1} + \epsilon_t$$

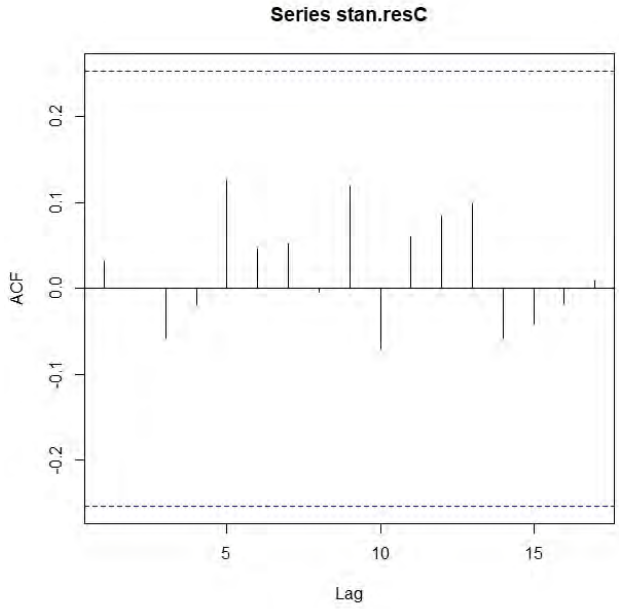
ต่อไปจะนำค่าเศษเหลือ (Residuals) ซึ่งคือค่าผลต่างของมูลค่าการประกันภัยรถยนต์จริงและมูลค่าการประกันภัยรถยนต์พยากรณ์มาทดสอบการแจกแจงแบบปกติโดยวิธีพิจารณากราฟของค่าเศษเหลือเทียบกับเวลา ดังแสดงในภาพที่ 3.23 ทดสอบการแจกแจงแบบปกติโดยวิธี Quantile – Quantile Plot หรือ QQ – plots และกราฟฮิสโตแกรม ดังแสดงในภาพที่ 3.24 ทดสอบความเป็นอิสระต่อกันของค่าเศษเหลือด้วยกราฟค่าสหสัมพันธ์ในตัวเองของเศษเหลือ ดังแสดงในภาพที่ 3.25 และทดสอบการแจกแจงปกติด้วยวิธี Kolmogorov-Smirnov normality test ดังแสดงในตารางที่ 3.15



ภาพที่ 3.23 แสดงค่าเศษเหลือของมูลค่าการเอาประกันภัยรถยนต์เทียบกับเวลาในแบบจำลอง ARIMA (5,1,1)



ภาพที่ 3.24 กราฟแสดงฮิสโตแกรมของเศษเหลือและ QQ-Plots ของมูลค่าการเอาประกันอัคคีภัยในแบบจำลอง ARIMA (5,1,1)



ภาพที่ 3.25 แสดงค่าสหสัมพันธ์ในตัวเองของเศษเหลือของมูลค่าการเอาประกันภัยรถยนต์ในแบบจำลอง ARIMA (5,1,1)

แบบจำลอง	ค่า Kolmogorov-Smirnov	ค่า p-value
ARIMA(11,1,1)	0.10891	0.4438

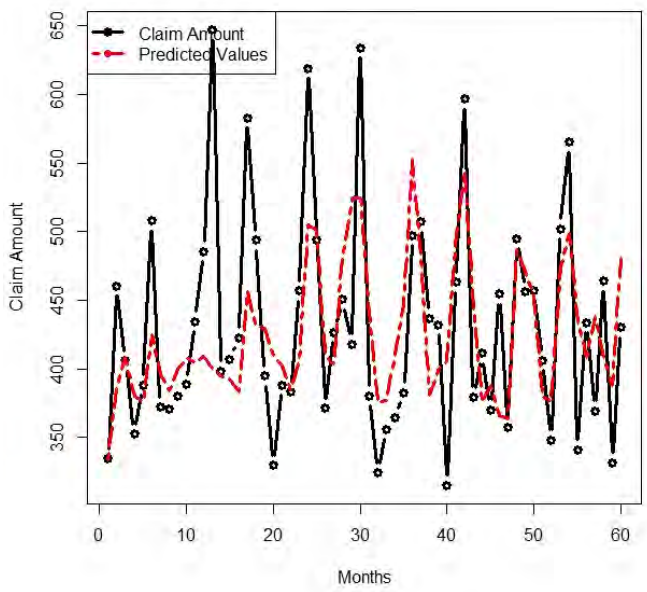
ตารางที่ 3.15 แสดงการทดสอบแบบจำลองด้วย Kolmogorov-Smirnov normality test ในแบบจำลอง ARIMA (5,1,1)

จากภาพที่ 3.23 การกระจายของเศษเหลือเทียบกับเวลาเป็นแบบสุ่ม ทำให้อนุกรมเวลาไม่มีแนวโน้ม จากภาพที่ 3.24 กราฟฮิสโตแกรมของเศษเหลือมีการแจกแจงเป็นรูประฆังคว่ำคล้ายการแจกแจงแบบปกติ และการทดสอบโดยวิธี QQ – plots พบว่ามีการกระจายของควอนไทล์เป็นเส้นตรงมีลักษณะคล้ายการแจกแจงแบบปกติมี จากภาพที่ 3.25 แบบจำลองแสดงให้เห็นว่าค่าสหสัมพันธ์ในตัวเองของเศษเหลือไม่มีความสัมพันธ์กันในทุก lag ที่มากกว่า 0 และจากตารางที่ 3.15 เมื่อทดสอบค่าสถิติด้วยวิธี Kolmogorov-Smirnov normality test พบว่า ค่า p-value เท่ากับ 0.4438 ซึ่งมากกว่า 0.05 แสดงว่าเศษเหลือของแบบจำลองมีการแจกแจงแบบปกติ เราจึงสรุปได้ว่าเศษเหลือของแบบจำลอง ARIMA (5,1,1) มีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติ

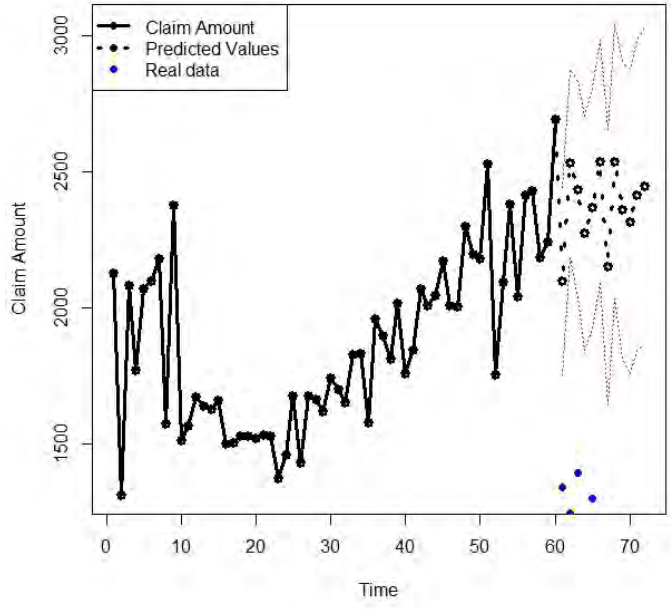
ขั้นตอนต่อไปจะเป็นการหามูลค่าพยากรณ์ของการเอาประกันภัยรถยนต์ของแบบจำลอง ARIMA (5,1,1) เทียบกับมูลค่าการเอาประกันภัยรถยนต์จริงในอีก 12 เดือนหลังจากเดือนธันวาคม 2559 แสดงเป็นค่าในตารางที่ 3.16 กราฟของมูลค่าพยากรณ์ของการเอาประกันภัยรถยนต์ของแบบจำลอง ARIMA (5,1,1) เทียบกับมูลค่าการเอาประกันภัยรถยนต์จริง ดังแสดงในภาพที่ 3.26 และ แสดงผลพยากรณ์ของมูลค่าการเอาประกันภัยรถยนต์ระหว่างเดือนมกราคม 2560 ถึงเดือนธันวาคม 2560 ดังแสดงในภาพที่ 3.27 ซึ่งคาดการณ์ว่าข้อมูลมูลค่าการเอาประกันภัยรถยนต์จริงรายเดือนปี พ.ศ. 2560 มีความผิดพลาดเพราะข้อมูลมีค่าต่างจากข้อมูลรายเดือนในปี พ.ศ. 2559 อย่างชัดเจน ดังแสดงในภาพที่ 3.27

วัน	มูลค่าการเอาประกันภัยจริง (พันล้านบาท)	ผลพยากรณ์ (พันล้านบาท)
ม.ค. 2560	1,340.0041	2,097.750
ก.พ. 2560	1,244.5607	2,530.217
มี.ค. 2560	1,391.9722	2,432.717
เม.ย. 2560	1,093.2687	2,271.793
พ.ค. 2560	1,296.5387	2,365.657
มิ.ย. 2560	1,103.5305	2,537.139
ก.ค. 2560	1,066.3351	2,149.009
ส.ค. 2560	1,016.3939	2,537.100
ก.ย. 2560	992.5706	2,357.771
ต.ค. 2560	972.3156	2,315.633
พ.ย. 2560	1,007.8010	2,411.621
ธ.ค. 2560	1,099.5879	2,446.720

ตารางที่ 3.16 แสดงมูลค่าการเอาประกันภัยรถยนต์จริงเทียบกับมูลค่าพยากรณ์ของการประกันภัยรถยนต์



ภาพที่ 3.26 แสดงกราฟของมูลค่าการเอาประกันภัยรถยนต์จริงเทียบกับมูลค่าพยากรณ์ของการเอาประกันภัยรถยนต์



ภาพที่ 3.27 แสดงผลพยากรณ์ของมูลค่าการเอาประกันภัยรถยนต์ระหว่างเดือนมกราคม 2560 ถึงเดือนธันวาคม 2560

3.1.1 สรุปผล

จากการศึกษาแบบจำลองอนุกรมเวลาของมูลค่าการเอาประกันวินาศภัย เพื่อหาแบบจำลองอนุกรมเวลาที่เหมาะสมและพยากรณ์มูลค่าการรับประกันวินาศภัย จำนวน 3 ประเภท ได้แก่ การประกันอัคคีภัย (Fire insurance) การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance) และ การประกันภัยรถยนต์ (Automobile insurance) โดยใช้ข้อมูลในช่วงเดือน มกราคม พ.ศ. 2555 ถึงเดือนธันวาคม พ.ศ. 2559 จากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) แสดงในตารางที่ 3.17

อนุกรมเวลาของมูลค่าการเอาประกันวินาศภัย	แบบจำลองที่เหมาะสม
การประกันอัคคีภัย	ARIMA (11,1,1)
การประกันภัยทางทะเลและขนส่ง	ARIMA (11,1,1)
การประกันภัยรถยนต์	ARIMA (5,1,1)

ตารางที่ 3.17 แสดงแบบจำลองที่เหมาะสมสำหรับการพยากรณ์มูลค่าการเอาประกันวินาศภัย

3.2 จำนวนกรมธรรม์ของประกันวินาศภัย

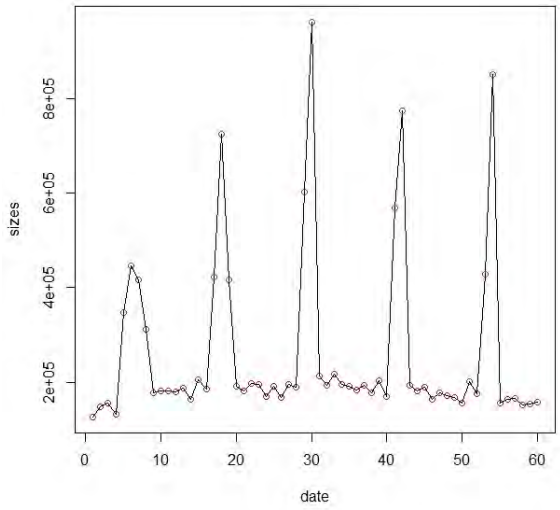
เนื่องจากข้อมูลที่ใช้ในการศึกษา คือ ข้อมูลจำนวนกรมธรรม์ของประกันวินาศภัย ซึ่งเป็นที่ชัดเจนว่าข้อมูลลักษณะดังกล่าวจะมีการแจกแจงแบบปัวซง เพราะเป็นจำนวนครั้งของการเคลมประกันวินาศภัยที่เกิดขึ้นในระยะเวลาหนึ่ง และในทำนองเดียวกันนี้ยังเหมาะกับข้อมูลที่มีการแจกแจงแบบทวินามลบด้วย นอกจากนี้ค่าความแปรปรวนที่ไม่คงที่ของข้อมูลนั้นขึ้นอยู่กับเวลาโดยเปลี่ยนไปตามช่วงเวลา ดังนั้นจึงมีการใช้การแปลงข้อมูลเข้ามาช่วยเพื่อให้ค่าความแปรปรวนคงที่ ซึ่งจากข้อมูลจำนวนกรมธรรม์ของประกันวินาศภัยมีการแจกแจงแบบปัวซง ทำให้ค่าเฉลี่ยและค่าความแปรปรวนเป็นฟังก์ชันเดียวกัน จึงใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูล ดังนั้นแบบจำลองที่เราจะศึกษาจึงประกอบไปด้วยแบบจำลองการใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซงและแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ โดยข้อมูลมูลค่าจำนวนกรมธรรม์ของประกันวินาศภัยที่ใช้ในการศึกษาสามารถแบ่งออกเป็น 3 ประเภท เหมือนดังหัวข้อที่ 3.1 ดังนี้

1) การประกันอัคคีภัย (fire insurance)

จำนวนกรมธรรม์ของการเอาประกันอัคคีภัย และ อนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย ตั้งแต่ เดือนมกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 แสดงในตาราง 3.18 และภาพที่ 3.28

วัน	ม.ค. 2555	ก.พ. 2555	มี.ค. 2555	เม.ย. 2555	พ.ค. 2555	มิ.ย. 2555
จำนวนกรมธรรม์	126891	148595	156831	131659	348351	446564
วัน	ก.ค. 2555	ส.ค. 2555	ก.ย. 2555	ต.ค. 2555	พ.ย. 2555	ธ.ค. 2555
จำนวนกรมธรรม์	415938	311213	177996	181663	182040	180353
⋮	⋮					
วัน	ก.ค. 2559	ส.ค. 2559	ก.ย. 2559	ต.ค. 2559	พ.ย. 2559	ธ.ค. 2559
จำนวนกรมธรรม์	155769	163441	165972	152067	153712	157890

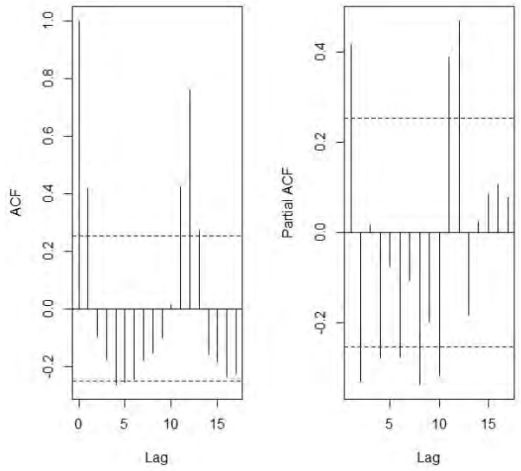
ตาราง 3.18 แสดงจำนวนกรมธรรม์การเอาประกันอัคคีภัยรายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559



ภาพที่ 3.28 แสดงอนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย

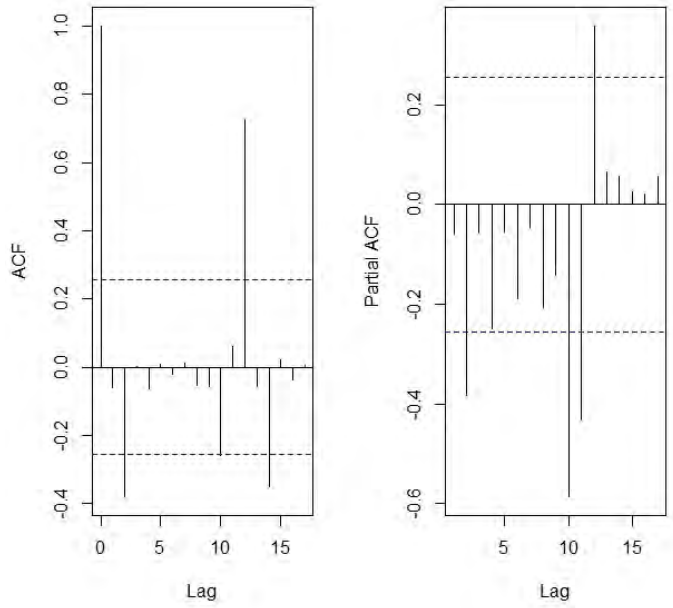
แบบจำลองที่ 1 การใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ

ต่อไปจะนำข้อมูลจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย มาเข้ากระบวนการแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูล โดยค่าสหสัมพันธ์ในตัวเอง (ACF) ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ค่าสหสัมพันธ์ในตัวเอง (ACF) ของผลต่างอันดับ 1 และ ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ของผลต่างอันดับ 1 แสดงใน ภาพที่ 3.29 และ ภาพที่ 3.30 ตามลำดับ



ภาพที่ 3.29 แสดง ACF และ PACF ของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย

จากภาพที่ 3.29 แสดง ACF และ PACF ของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย โดยเมื่อทำการทดสอบความนิ่งของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการรับประกันอัคคีภัย ด้วยวิธี Augmented Dickey-Fuller Test โดยที่สมมติฐานหลักคือข้อมูลไม่คงที่ พบว่าค่า p-value มีค่ามากกว่า 0.01 จึงยอมรับสมมติฐานหลัก จึงสรุปได้ว่าข้อมูลยังไม่นิ่ง



ภาพที่ 3.30 แสดง ACF และ PACF ค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย โดยการหาผลต่างของข้อมูลครั้งที่ 1

จากภาพที่ 3.30 เมื่อทำการทดสอบความนิ่งของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย ด้วยวิธี Augmented Dickey-Fuller Test พบว่าอนุกรมเวลาของผลต่างอันดับ 1 นิ่งและมีค่าสหสัมพันธ์ (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 2,10,12,14 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 2,10,11,12 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจึงพิจารณา $ARIMA(p,d,q)$ เมื่อ $p \in \{2,10,11,12\}$, $q \in \{2,10,12,14\}$ และ $d = 1$ ตามลำดับ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC น้อยที่สุด

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(2,1,2)	-	ARIMA(11,1,2)	-
ARIMA(2,1,10)	740.3532	ARIMA(11,1,10)	689.1419
ARIMA(2,1,12)	734.1474	ARIMA(11,1,12)	686.8865
ARIMA(2,1,14)	724.6004	ARIMA(11,1,14)	690.0746
ARIMA(10,1,2)	-	ARIMA(12,1,2)	-
ARIMA(10,1,10)	**682.7947**	ARIMA(12,1,10)	691.5227
ARIMA(10,1,12)	695.2984	ARIMA(12,1,12)	688.7717
ARIMA(10,1,14)	686.9941	ARIMA(12,1,14)	689.6483

ตารางที่ 3.19 แสดงค่า AIC ของแบบจำลองค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย

จากตารางที่ 3.19 เมื่อพิจารณาค่า AIC เนื่องจากแบบจำลอง ARIMA (10,1,10) ให้ค่า AIC น้อยที่สุดคือ 682.7947 ดังนั้นแบบจำลองอนุกรมเวลาที่เหมาะสมคือ ARIMA (10,1,10) ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.20

พารามิเตอร์	ar1	ar2	ar3	ar4	ar5
ARIMA (10,1,10)	0.0131	-1.0696	0.0377	-1.0874	0.0340
	ar6	ar7	ar8	ar9	ar10
	-1.0727	0.0218	-1.0593	-0.0047	-0.9323
	ma1	ma2	ma3	ma4	ma5
	-0.9248	1.0145	-0.8388	1.0726	-0.9978
	ma6	ma7	ma8	ma9	ma10
	1.1453	-0.9627	0.9352	-0.8533	0.5519

ตารางที่ 3.20 พารามิเตอร์ของแบบจำลอง ARIMA (10,1,10)

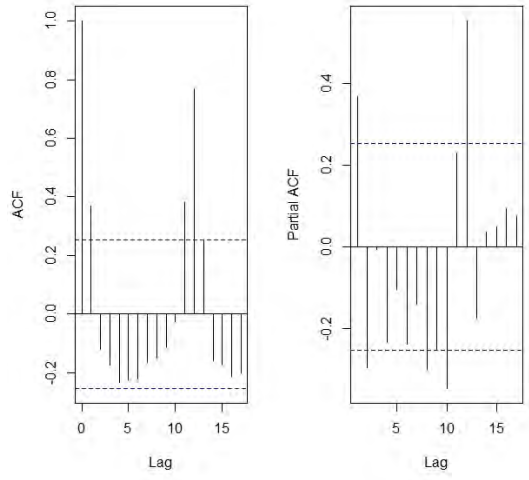
จากตารางที่ 3.20 จึงสรุปได้ว่าแบบจำลองอนุกรมเวลาที่เหมาะสมที่สุดสำหรับข้อมูลของจำนวนกรมธรรม์การเอาประกันอัคคีภัย คือ แบบจำลอง ARIMA(10,1,10) ซึ่งสามารถเขียนประมาณพารามิเตอร์จากการทำผลต่างอันดับ 1 ได้สมการต่อไปนี้

$$\begin{aligned} y_t - y_{t-1} = {} & 0.0131(y_{t-1} - y_{t-2}) - 1.0696(y_{t-2} - y_{t-3}) + 0.0377(y_{t-3} - y_{t-4}) \\ & -1.0874(y_{t-4} - y_{t-5}) + 0.0340(y_{t-5} - y_{t-6}) - 1.0727(y_{t-6} - y_{t-7}) \\ & +0.0218(y_{t-7} - y_{t-8}) - 1.0593(y_{t-8} - y_{t-9}) - 0.0047(y_{t-9} - y_{t-10}) \\ & -0.9323(y_{t-10} - y_{t-11}) + 0.9248\epsilon_{t-1} - 1.0145\epsilon_{t-2} + 0.8388\epsilon_{t-3} \\ & -1.0726\epsilon_{t-4} + 0.9978\epsilon_{t-5} - 1.1453\epsilon_{t-6} + 0.9627\epsilon_{t-7} \\ & -0.9352\epsilon_{t-8} + 0.8533\epsilon_{t-9} - 0.5519\epsilon_{t-10} + \epsilon_t \end{aligned}$$

จัดรูปสมการต่อได้ดังนี้

$$\begin{aligned} y_t = {} & 1.0131y_{t-1} - 1.0827y_{t-2} + 1.1073y_{t-3} - 1.1251y_{t-4} + 1.1214y_{t-5} \\ & -1.1067y_{t-6} + 1.0945y_{t-7} - 1.0811y_{t-8} + 1.0546y_{t-9} - 0.9276y_{t-10} \\ & +0.9323y_{t-11} + 0.9248\epsilon_{t-1} - 1.0145\epsilon_{t-2} + 0.8388\epsilon_{t-3} - 1.0726\epsilon_{t-4} \\ & +0.9978\epsilon_{t-5} - 1.1453\epsilon_{t-6} + 0.9627\epsilon_{t-7} - 0.9352\epsilon_{t-8} + 0.8533\epsilon_{t-9} \\ & -0.5519\epsilon_{t-10} + \epsilon_t \end{aligned}$$

ต่อไปจะนำข้อมูลจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย มาพิจารณาในแบบจำลองอนุกรมเวลาที่มีการแจกแจงแบบปัวซงและแบบจำลองอนุกรมเวลาที่มีการแจกแจงแบบทวินามลบ



ภาพที่ 3.31 แสดง ACF และ PACF ของจำนวนกรมธรรม์ของการเอาประกันอัคคีภัย

จากภาพที่ 3.31 เมื่อพิจารณาค่าสหสัมพันธ์ (ACF) และ ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) พบว่ามีค่าสหสัมพันธ์ (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,11 และ 12 ตามลำดับ และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,8,10 และ 12 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจึงพิจารณาส่วนของข้อมูลในอดีต (previous observations) เท่ากับ 1,2,8,10 และ 12 ตามลำดับ และส่วนของค่าเฉลี่ยแบบมีเงื่อนไข (previous conditional means) เท่ากับ 1,11 และ 12 ตามลำดับของทั้งแบบจำลองอนุกรมเวลาที่มีการแจกแจงแบบปัวซงและแบบจำลองอนุกรมเวลาที่มีการแจกแจงแบบทวินามลบ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC น้อยที่สุด

แบบจำลองที่ 2 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

แบบจำลอง	past_obs	past_mean	ค่า AIC	แบบจำลอง	past_obs	past_mean	ค่า AIC
1	1	1	3,063,419	16	12	(1,11)	1,406,998
2	1	11	2,757,643	17	12	(1,12)	1,228,700
3	1	12	4,329,759	18	12	(11,12)	1,243,755
4	2	1	3,641,248	19	(1,2)	1	3,064,447
5	2	11	5,447,105	20	(1,8)	1	3,029,645
6	2	12	5,410,784	21	(1,10)	1	3,059,897
7	8	1	5,493,531	22	(1,12)	1	2,004,890
8	8	11	5,166,619	23	(2,8)	1	3,493,278
9	8	10	5,497,840	24	(2,10)	1	3,324,134
10	10	1	5,357,301	25	(2,12)	1	1,174,870
11	10	11	5,690,221	26	(8,10)	1	4,027,297
12	10	12	5,562,068	27	(8,12)	1	1,420,328
13	12	1	1,227,187	28	(10,12)	1	1,218,034
14	12	11	1,244,359	⋮	⋮	⋮	⋮
15	12	12	1,248,951	**29**	**(2,12)**	**(1,11)**	**1,168,748**

ตารางที่ 3.21 แสดงค่า AIC แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

จากตารางที่ 3.21 นำค่า previous observations และค่า previous conditional means มาพิจารณาจับคู่ทีละตัวพบว่าที่แบบจำลองที่ 13 ให้ค่า AIC น้อยที่สุด จึงนำค่า previous observations เท่ากับ 12 และค่า previous conditional means เท่ากับ 1 ไปจับคู่กับตัวอื่น แล้วพิจารณาค่า AIC โดยกับจับคู่ไปเรื่อยๆ จึงสรุปได้ว่าว่าแบบจำลองที่ 29 ซึ่งมีค่า previous observations เท่ากับ 2 และ 12 และค่า previous conditional means เท่ากับ 1 และ 11 ให้ค่า AIC น้อยที่สุด นั่นคือ 1,168,748 ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.22

พารามิเตอร์	ค่าประมาณ
Intercept	2.5492
beta_2	-0.1427
beta_12	0.9992
alpha_1	-0.0223
alpha_11	-0.0404

ตารางที่ 3.22 แสดงค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

จากการตารางที่ 3.22 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง สามารถเขียนประมาณพารามิเตอร์ ได้สมการต่อไปนี้

$$log(\lambda_t) = 2.5492 - 0.1427Y_{t-2} + 0.9992Y_{t-12} - 0.0223\lambda_{t-1} - 0.0404\lambda_{t-11}$$

แบบจำลองที่ 3 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

แบบจำลอง	past_obs	past_mean	ค่า AIC	แบบจำลอง	past_obs	past_mean	ค่า AIC
1	1	1	1,566.968	16	12	(1,11)	1,505.020
2	1	11	1,536.986	17	12	(1,12)	1,500.524
3	1	12	1,570.921	18	12	(11,12)	1,502.284
4	2	1	1,558.406	19	(1,2)	1	1,568.965
5	2	11	1,590.571	20	(1,8)	1	1,567.138
6	2	12	1,589.762	21	(1,10)	1	1,569.718
7	8	1	1,592.001	22	(1,12)	1	1,531.158
8	8	11	1,586.227	23	(2,8)	1	1,556.993
9	8	10	1,592.094	24	(2,10)	1	1,555.584
10	10	1	1,591.710	25	(2,12)	1	1,499.331
11	10	11	1,597.249	26	(8,10)	1	1,569.303
12	10	12	1,595.156	27	(8,12)	1	1,506.084
13	**12**	**1**	**1,498.426**	28	(10,12)	1	1,499.389
14	12	11	1,500.344	⋮	⋮	⋮	⋮
15	12	12	1,500.084				

ตารางที่ 3.23 แสดงค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

จากตารางที่ 3.23 นำค่า previous observations และค่า previous conditional means มาพิจารณาจับคู่ทีละตัวพบว่าที่แบบจำลองที่ 13 ให้ค่า AIC น้อยที่สุด จึงนำค่า previous observations เท่ากับ 12 และค่า previous conditional means เท่ากับ 1 ไปจับคู่กับตัวอื่น แล้วพิจารณาค่า AIC โดยกับจับคู่ไปเรื่อยๆ จึงสรุปได้ว่าว่าแบบจำลองที่ 13 ซึ่งมีค่า previous observations เท่ากับ 12 และค่า previous conditional means เท่ากับ 1 ให้ค่า AIC น้อยที่สุด นั่นคือ 1,498.426 ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.24

พารามิเตอร์	ค่าประมาณ
Intercept	0.7742
beta_12	1.0000
alpha_1	-0.0626
sigmasq	0.0800

ตารางที่ 3.24 แสดงค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

จากการตารางที่ 3.24 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ สามารถเขียนประมาณพารามิเตอร์ ได้สมการต่อไปนี้

$$log(\lambda_t) = 0.7742 + Y_{t-12} - 0.0626\lambda_{t-1}$$

ต่อไปเราจะมาพิจารณาค่าคลาดเคลื่อนกำลังสองของทั้งสามแบบจำลอง โดยแสดงในตารางที่ 3.25

แบบจำลอง	ค่าคลาดเคลื่อนกำลังสอง
การใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ ARIMA(10,1,10)	220,825,319,055
แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง	418,817,570,147
แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ	449,628,695,323

ตารางที่ 3.25 แสดงค่าคลาดเคลื่อนกำลังสองของแบบจำลอง

จากตารางที่ 3.25 เมื่อพิจารณาค่าคลาดเคลื่อนกำลังสองเนื่องจากแบบจำลองการใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติให้ค่าคลาดเคลื่อนกำลังสองน้อยที่สุดคือ 220,825,319,055 ดังนั้นแบบจำลองอนุกรมเวลาที่เหมาะสมสำหรับข้อมูลของจำนวนกรมธรรม์การเอาประกันอัคคีภัย คือ แบบจำลองการใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ ARIMA(10,1,10)

ขั้นตอนต่อไปจะเป็นการหาจำนวนกรมธรรม์พยากรณ์ของการเอาประกันอัคคีภัยเทียบกับจำนวนกรมธรรม์พยากรณ์ของการเอาประกันอัคคีภัยจริงในอีก 12 เดือนหลังจากเดือนธันวาคม 2559 แสดงเป็นค่าในตารางที่ 3.25

วันที่	จำนวนกรมธรรม์จริง	ผลพยากรณ์
ม.ค. 2560	160,636	158,116.1
ก.พ. 2560	146,080	139,746.3
มี.ค. 2560	163,282	229,342.3
เม.ย. 2560	149,748	178,948.0
พ.ค. 2560	169,575	470,361.8
มิ.ย. 2560	1,283,628	728,352.1
ก.ค. 2560	357,438	163,702.1
ส.ค. 2560	178,944	130,473.3
ก.ย. 2560	182,325	177,724.7
ต.ค. 2560	166,459	130,339.5
พ.ย. 2560	170,655	154,603.0
ธ.ค. 2560	191,334	159,522.9

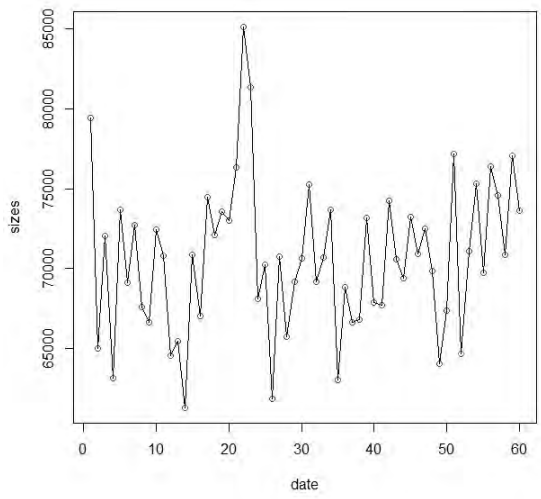
ตารางที่ 3.26 จำนวนกรมธรรม์พยากรณ์ของการเอาประกันอัคคีภัยเทียบกับจำนวนกรมธรรม์ของการเอาประกันอัคคีภัยจริง

2) การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance)

จำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง และ อนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง ตั้งแต่ เดือนมกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 แสดงในตาราง 3.27 และ ภาพที่ 3.32

วัน	ม.ค. 2555	ก.พ. 2555	มี.ค. 2555	เม.ย. 2555	พ.ค. 2555	มิ.ย. 2555
จำนวนกรมธรรม์	79435	64984	72020	63158	73667	69134
วัน	ก.ค. 2555	ส.ค. 2555	ก.ย. 2555	ต.ค. 2555	พ.ย. 2555	ธ.ค. 2555
จำนวนกรมธรรม์	72704	67617	66647	72450	70799	64537
⋮	⋮					
วัน	ก.ค. 2559	ส.ค. 2559	ก.ย. 2559	ต.ค. 2559	พ.ย. 2559	ธ.ค. 2559
จำนวนกรมธรรม์	69719	76404	74603	70884	77048	73638

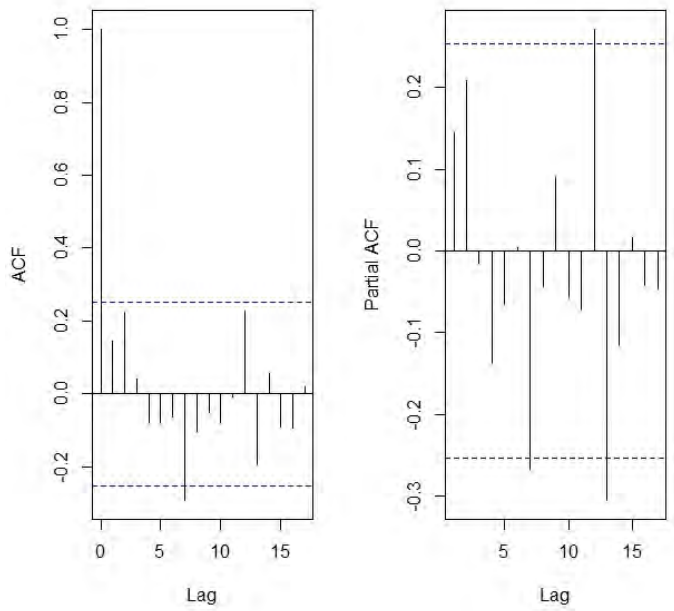
ตาราง 3.27 แสดงจำนวนกรมธรรม์การเอาประกันการประกันภัยทางทะเลและขนส่งรายเดือนตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559



ภาพที่ 3.32 แสดงอนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง

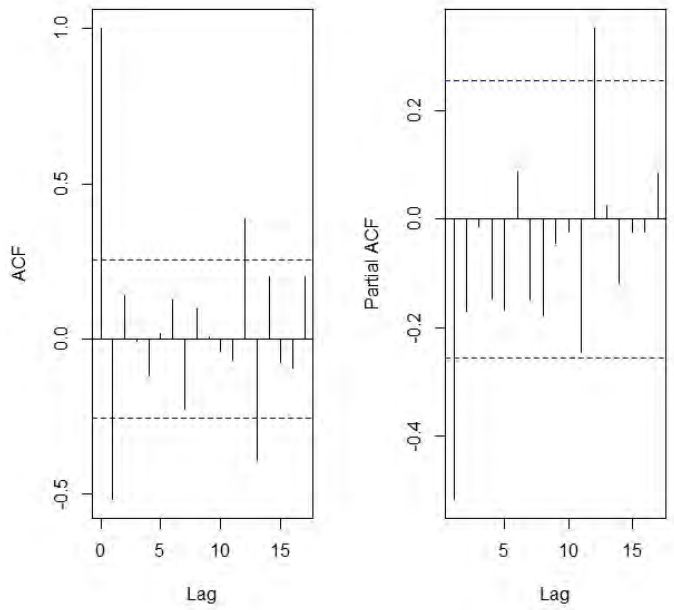
แบบจำลองที่ 1 การใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ

ต่อไปจะนำข้อมูลจำนวนกรมธรรม์ของการการเอาประกันการประกันภัยทางทะเลและขนส่ง มาเข้ากระบวนการแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูล โดยค่าสหสัมพันธ์ในตัวเอง (ACF) ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ค่าสหสัมพันธ์ในตัวเอง (ACF) ของผลต่างอันดับ 1 และ ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ของผลต่างอันดับ 1 แสดงใน ภาพที่ 3.33 และ ภาพที่ 3.34 ตามลำดับ



ภาพที่ 3.33 แสดง ACF และ PACF ของ ค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง

จากภาพที่ 3.33 แสดง ACF และ PACF ของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง โดยเมื่อทำการทดสอบความนิ่งของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง ด้วยวิธี Augmented Dickey-Fuller Test โดยที่สมมติฐานหลักคือข้อมูลไม่คงที่ พบว่าค่า p-value มีค่ามากกว่า 0.01 จึงยอมรับสมมติฐานหลัก จึงสรุปได้ว่าข้อมูลยังไม่นิ่ง



ภาพที่ 3.34 แสดง ACF และ PACF ของ ค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง โดยการหาผลต่างของข้อมูลครั้งที่ 1

จากภาพที่ 3.34 เมื่อทำการทดสอบความนิ่งของ ค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง ด้วยวิธี Augmented Dickey-Fuller Test พบว่าอนุกรมเวลาของผลต่างอันดับ 1 นิ่งและมีค่าสหสัมพันธ์ (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,12 และ 13 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1 และ 12 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจึงพิจารณา $ARIMA(p,d,q)$ เมื่อ $p \in \{1,12\}$, $q \in \{1,12,13\}$ และ $d = 1$ ตามลำดับ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC น้อยที่สุด

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,1,1)	599.3255	ARIMA(12,1,1)	601.3233
ARIMA(1,1,12)	432.1036	**ARIMA(12,1,12)**	**428.9287**
ARIMA(1,1,13)	433.1408	ARIMA(12,1,13)	433.9995

ตารางที่ 3.28 แสดงค่า AIC ของแบบจำลองของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง

จากตารางที่ 3.28 เมื่อพิจารณาค่า AIC เนื่องจากแบบจำลอง ARIMA (12,1,12) ให้ค่า AIC น้อยที่สุดคือ 428.9287 ดังนั้นแบบจำลองอนุกรมเวลาที่เหมาะสมคือ ARIMA(12,1,12) ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.29

พารามิเตอร์	ar1	ar2	ar3	ar4	ar5	ar6
ARIMA (12,1,12)	-0.7750	-0.3658	0.2306	0.7440	1.0340	0.5115
	ar7	ar8	ar9	ar10	ar11	ar12
	-0.2976	-0.4874	-0.4719	-0.4159	-0.2185	0.4261
	ma1	ma2	ma3	ma4	ma5	ma6
	0.0733	0.2861	-1.1428	-1.2955	-1.3395	0.2074
	ma7	ma8	ma9	ma10	ma11	ma12
	1.1081	1.5471	0.8120	-0.1240	-0.2557	-0.8763

ตารางที่ 3.29 พารามิเตอร์ของแบบจำลอง ARIMA (12,1,12)

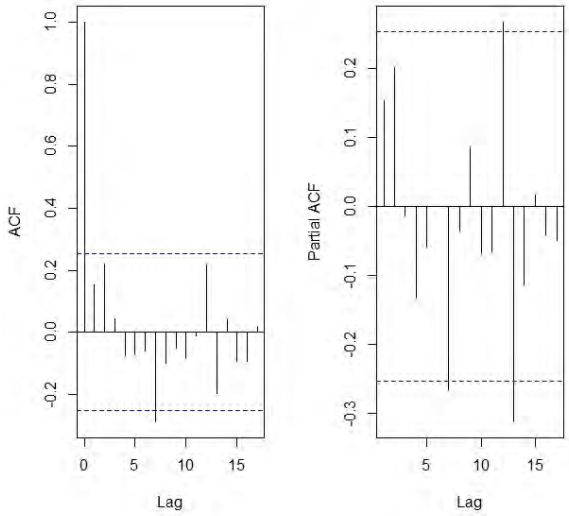
จากตารางที่ 3.29 จึงสรุปได้ว่าแบบจำลองอนุกรมเวลาที่เหมาะสมที่สุดสำหรับข้อมูลของจำนวนกรมธรรม์การเอาประกันการประกันภัยทางทะเลและขนส่ง คือ แบบจำลอง ARIMA(12,1,12) ซึ่งสามารถเขียนประมาณพารามิเตอร์จากการทำผลต่างอันดับ 1 ได้สมการต่อไปนี้

$$\begin{aligned} y_t - y_{t-1} = {} & -0.7750(y_{t-1} - y_{t-2}) - 0.3658(y_{t-2} - y_{t-3}) + 0.2306(y_{t-3} - y_{t-4}) \\ & +0.7440(y_{t-4} - y_{t-5}) + 1.0340(y_{t-5} - y_{t-6}) + 0.5115(y_{t-6} - y_{t-7}) \\ & -0.2976(y_{t-7} - y_{t-8}) - 0.4874(y_{t-8} - y_{t-9}) - 0.4719(y_{t-9} - y_{t-10}) \\ & -0.4159(y_{t-10} - y_{t-11}) - 0.2185(y_{t-11} - y_{t-12}) \\ & +0.4261(y_{t-12} - y_{t-13}) - 0.0733\epsilon_{t-1} - 0.2861\epsilon_{t-2} + 1.1428\epsilon_{t-3} \\ & +1.2955\epsilon_{t-4} + 1.3395\epsilon_{t-5} - 0.2074\epsilon_{t-6} - 1.1081\epsilon_{t-7} - 1.5471\epsilon_{t-8} \\ & -0.8120\epsilon_{t-9} + 0.1240\epsilon_{t-10} + 0.25573\epsilon_{t-11} + 0.8763\epsilon_{t-12} + \epsilon_t \end{aligned}$$

จัดรูปสมการต่อได้ดังนี้

$$\begin{aligned} y_t = {} & 0.2250y_{t-1} + 0.4092y_{t-2} - 0.6164y_{t-3} + 0.5134y_{t-4} + 0.2900y_{t-5} \\ & -0.5225y_{t-6} - 0.8091y_{t-7} - 0.1898y_{t-8} + 0.0155y_{t-9} - 0.0560y_{t-10} \\ & +0.1974y_{t-11} + 0.6446y_{t-12} + 0.4261y_{t-13} - 0.0733\epsilon_{t-1} - 0.2861\epsilon_{t-2} \\ & +1.1428\epsilon_{t-3} + 1.2955\epsilon_{t-4} + 1.3395\epsilon_{t-5} - 0.2074\epsilon_{t-6} - 1.1081\epsilon_{t-7} \\ & -1.5471\epsilon_{t-8} - 0.8120\epsilon_{t-9} + 0.1240\epsilon_{t-10} + 0.25573\epsilon_{t-11} \\ & +0.8763\epsilon_{t-12} + \epsilon_t \end{aligned}$$

ต่อไปจะนำข้อมูลจำนวนกรมธรรม์ของกาเอาประกันการประกันภัยทางทะเลและขนส่ง มาพิจารณาในแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซงและแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ



ภาพที่ 3.35 แสดง ACF และ PACF ของจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง

จากภาพที่ 3.35 เมื่อพิจารณาค่าสหสัมพันธ์ (ACF) และ ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) พบว่ามีค่าสหสัมพันธ์ (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 7 และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 7,12 และ 13 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจึงพิจารณาส่วนของข้อมูลในอดีต (previous observations) เท่ากับ 7,12 และ 13 ตามลำดับ และ ส่วนของค่าเฉลี่ยแบบมีเงื่อนไข (previous conditional means) เท่ากับ 7 ของทั้งแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซงและแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC น้อยที่สุด

แบบจำลองที่ 2 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

แบบจำลอง	past_obs	past_mean	ค่า AIC
1	7	7	17,333.50
2	12	7	17,318.38
3	13	7	18,151.89
4	(7,12)	1	16,411.28
5	(7,13)	7	16,203.00
6	(12,13)	7	15,901.18
7	**(7,12,13)**	**7**	**14,937.18**

ตารางที่ 3.30 แสดงค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

จากตารางที่ 3.30 นำค่า previous observations และค่า previous conditional means มาพิจารณาจับคู่ทีละตัว จึงสรุปได้ว่าว่าแบบจำลองที่ 7 ซึ่งมีค่า previous observations เท่ากับ 7,12 และ 13 และค่า previous conditional means เท่ากับ 7 ให้ค่า AIC น้อยที่สุด นั่นคือ 14,937.18 ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.31

พารามิเตอร์	ค่าประมาณ
Intercept	15.492
beta_7	-0.302
beta_12	0.309
beta_13	-0.317
alpha_7	-0.077

ตารางที่ 3.31 แสดงค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

จากการตารางที่ 3.31 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง สามารถเขียนประมาณพารามิเตอร์ ได้สมการต่อไปนี้

$$log(\lambda_t) = 15.492 - 0.302Y_{t-7} + 0.309Y_{t-12} - 0.317Y_{t-13} - 0.077\lambda_{t-7}$$

แบบจำลองที่ 3 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

แบบจำลอง	past_obs	past_mean	ค่า AIC
1	7	7	1,185.050
2	12	7	1,184.791
3	13	7	1,187.915
4	(7,12)	1	1,183.307
5	(7,13)	7	1,182.805
6	(12,13)	7	1,181.481
7	**(7,12,13)**	**7**	**1,179.534**

ตารางที่ 3.32 แสดงค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

จากตารางที่ 3.32 นำค่า previous observations และค่า previous conditional means มาพิจารณาจับคู่ทีละตัว จึงสรุปได้ว่าว่าแบบจำลองที่ 7 ซึ่งมีค่า previous observations เท่ากับ 7,12 และ 13 และค่า previous conditional means เท่ากับ 7 ให้ค่า AIC น้อยที่สุด นั่นคือ 1,179.534 ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.33

พารามิเตอร์	ค่าประมาณ
Intercept	15.49163
beta_7	-0.30250
beta_12	0.30905
beta_13	-0.31725
alpha_7	-0.07700

ตารางที่ 3.33 แสดงค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

จากการตารางที่ 3.33 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ สามารถเขียนประมาณพารามิเตอร์ ได้สมการต่อไปนี้

$$log(\lambda_t) = 15.49163 - 0.30250Y_{t-7} + 0.30905Y_{t-12} - 0.31725Y_{t-13} - 0.07700\lambda_{t-7}$$

ต่อไปเราจะมาพิจารณาค่าคลาดเคลื่อนกำลังสองของทั้งสามแบบจำลอง โดยแสดงในตารางที่ 3.33

แบบจำลอง	ค่าคลาดเคลื่อนกำลังสอง
การใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ ARIMA(12,1,12)	306,278,162
แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง	1,011,085,462
แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ	1,011,085,462

ตารางที่ 3.34 แสดงค่าคลาดเคลื่อนกำลังสองของแบบจำลอง

จากตารางที่ 3.34 เมื่อพิจารณาค่าคลาดเคลื่อนกำลังสอง เนื่องจากแบบจำลองการใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติให้ค่าคลาดเคลื่อนกำลังสองน้อยที่สุดคือ 306,278,162 ดังนั้นแบบจำลองอนุกรมเวลาที่เหมาะสมสำหรับข้อมูลของจำนวนกรมธรรม์การเอาประกันการประกันภัยทางทะเลและขนส่ง คือ แบบจำลองการใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ ARIMA(12,1,12)

ขั้นตอนต่อไปจะเป็นการหาจำนวนกรมธรรม์พยากรณ์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง เทียบกับจำนวนกรมธรรม์พยากรณ์ของการเอาประกันการประกันภัยทางทะเลและขนส่ง จริงในอีก 12 เดือนหลังจากเดือนธันวาคม 2559 แสดงเป็นค่าในตารางที่ 3.35

วันที่	จำนวนกรมธรรม์จริง	ผลพยากรณ์
ม.ค. 2560	67,947	69,333.07
ก.พ. 2560	66,908	77,096.63
มี.ค. 2560	79,816	70,151.81
เม.ย. 2560	64,828	61,602.60
พ.ค. 2560	76,743	69,891.59
มิ.ย. 2560	89,308	66,217.21
ก.ค. 2560	71,205	67,406.17
ส.ค. 2560	78,247	72,066.29
ก.ย. 2560	74,518	61,564.07
ต.ค. 2560	72,328	67,666.34
พ.ย. 2560	77,850	74,954.65
ธ.ค. 2560	72,074	65,476.96

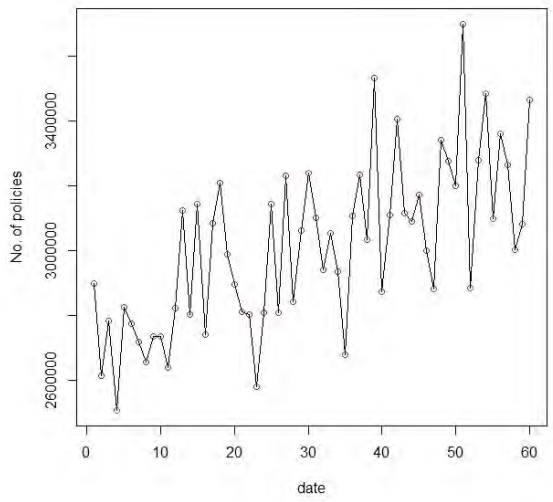
ตารางที่ 3.35 จำนวนกรมธรรม์พยากรณ์ของการเอาประกันการประกันภัยทางทะเลและขนส่งเทียบกับจำนวนกรมธรรม์ของการเอาประกันการประกันภัยทางทะเลและขนส่งจริง

3) การประกันภัยรถยนต์ (Automobile insurance)

จำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์ และ อนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์ตั้งแต่ เดือนมกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 แสดงในตาราง 3.36 และ ภาพที่ 3.36

วัน	ม.ค. 2555	ก.พ. 2555	มี.ค. 2555	เม.ย. 2555	พ.ค. 2555	มิ.ย. 2555
จำนวนกรมธรรม์	2898725	2614805	2784542	2509070	2826138	2774240
วัน	ก.ค. 2555	ส.ค. 2555	ก.ย. 2555	ต.ค. 2555	พ.ย. 2555	ธ.ค. 2555
จำนวนกรมธรรม์	2718702	2656140	2737062	2736816	2639918	2822726
⋮	⋮					
วัน	ก.ค. 2559	ส.ค. 2559	ก.ย. 2559	ต.ค. 2559	พ.ย. 2559	ธ.ค. 2559
จำนวนกรมธรรม์	3097950	3359318	3265920	3004652	3082114	3463620

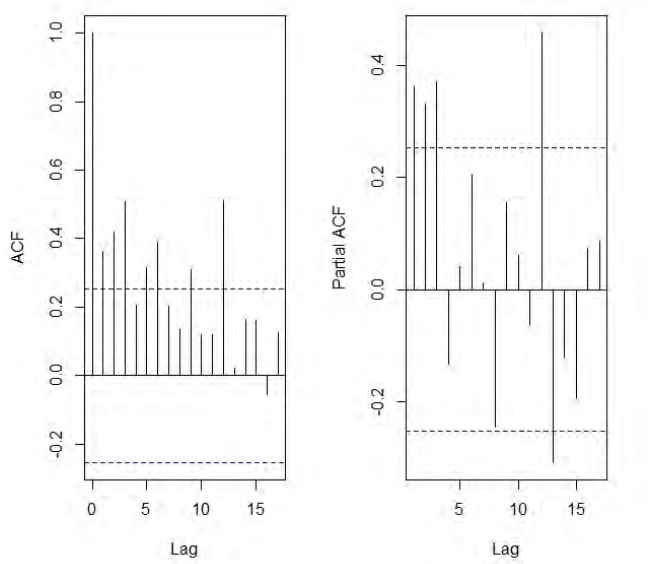
ตาราง 3.36 แสดงจำนวนกรมธรรม์การเอาประกันการประกันภัยรถยนต์รายเดือนตั้งแต่ปี พ.ศ. 2555 ถึง ปี พ.ศ. 2559



ภาพที่ 3.36 แสดงอนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์

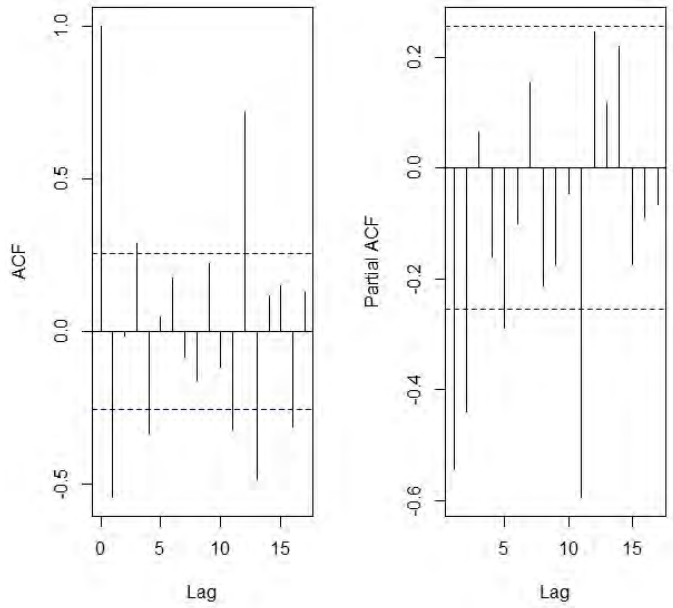
แบบจำลองที่ 1 การใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ

ต่อไปจะนำข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์มาเข้ากระบวนการแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูล โดยค่าสหสัมพันธ์ในตัวเอง (ACF) ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ค่าสหสัมพันธ์ในตัวเอง (ACF) ของผลต่างอันดับ 1 และ ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) ของผลต่างอันดับ 1 แสดงใน ภาพที่ 3.37 และ ภาพที่ 3.38 ตามลำดับ



ภาพที่ 3.37 แสดง ACF และ PACF ของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์

จากภาพที่ 3.37 แสดง ACF และ PACF ของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์การเอาประกันการประกันภัยรถยนต์ โดยเมื่อทำการทดสอบความนิ่งของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการรับประกันการประกันภัยทางทะเลและขนส่ง ด้วยวิธี Augmented Dickey-Fuller Test โดยที่สมมติฐานหลักคือข้อมูลไม่คงที่ พบว่าค่า p-value มีค่ามากกว่า 0.01 จึงยอมรับสมมติฐานหลัก จึงสรุปได้ว่าข้อมูลยังไม่นิ่ง



ภาพที่ 3.38 แสดง ACF และ PACF ของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์ โดยการหาผลต่างของข้อมูลครั้งที่ 1

จากภาพที่ 3.38 เมื่อทำการทดสอบความนิ่งค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์ ด้วยวิธี Augmented Dickey-Fuller Test พบว่าอนุกรมเวลาของผลต่างอันดับ 1 นิ่งและมีค่าสหสัมพันธ์ (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,3,4,11,12,13 และ 16 ตามลำดับ และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,5 และ 11 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจึงพิจารณา $ARIMA(p,d,q)$ เมื่อ $p \in \{1,2,5,11\}$, $q \in \{1,3,4,11,12,13,16\}$ และ $d = 1$ ตามลำดับ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC น้อยที่สุด

แบบจำลอง	ค่า AIC	แบบจำลอง	ค่า AIC
ARIMA(1,1,1)	-	ARIMA(5,1,1)	-
ARIMA(1,1,3)	-	ARIMA(5,1,3)	-
ARIMA(1,1,4)	-	ARIMA(5,1,4)	-
ARIMA(1,1,11)	638.1006	ARIMA(5,1,11)	635.7475
ARIMA(1,1,12)	633.3999	ARIMA(5,1,12)	628.4912
ARIMA(1,1,13)	633.0188	ARIMA(5,1,13)	630.0906
ARIMA(1,1,16)	638.0621	ARIMA(5,1,16)	612.2468
ARIMA(2,1,1)	-	ARIMA(11,1,1)	-
ARIMA(2,1,3)	-	ARIMA(11,1,3)	-
ARIMA(2,1,4)	-	ARIMA(11,1,4)	-
ARIMA(2,1,11)	637.5955	**ARIMA(11,1,11)**	**592.9789**
ARIMA(2,1,12)	632.1014	ARIMA(11,1,12)	598.0024
ARIMA(2,1,13)	633.7887	ARIMA(11,1,13)	600.9355
ARIMA(2,1,16)	622.6073	ARIMA(11,1,16)	596.5298

ตารางที่ 3.37 แสดงค่า AIC ของแบบจำลองของค่ารากที่สองของข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์

จากตารางที่ 3.37 เมื่อพิจารณาค่า AIC เนื่องจากแบบจำลอง ARIMA (11,1,11) ให้ค่า AIC น้อยที่สุดคือ 592.9789 ดังนั้นแบบจำลองอนุกรมเวลาที่เหมาะสมคือ ARIMA(11,1,11) ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.37

พารามิเตอร์	ar1	ar2	ar3	ar4	ar5	ar6
ARIMA (11,1,1)	-1.0613	-1.0097	-1.0783	-1.1732	-1.1027	-1.0852
	ar7	ar8	ar9	ar10	ar11	ma1
	-1.0930	-0.9793	-0.9191	-0.9706	-0.9416	0.6713
	ma2	ma3	ma4	ma5	ma6	ma7
	0.8304	1.0108	0.7419	0.6340	0.9486	0.3297
	ma8	ma9	ma10	ma11		
	0.3768	0.6682	0.1032	0.1273		

ตารางที่ 3.38 พารามิเตอร์ของแบบจำลอง ARIMA (11,1,11)

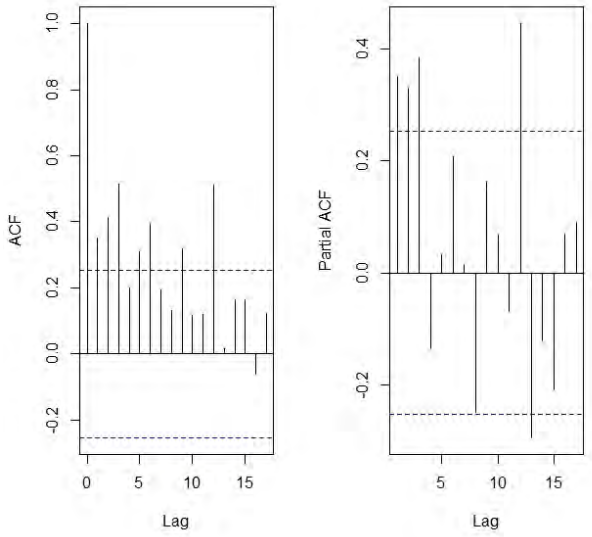
จากตารางที่ 3.38 จึงสรุปได้ว่าแบบจำลองอนุกรมเวลาที่เหมาะสมที่สุดสำหรับข้อมูลของจำนวนกรมธรรม์การเอาประกันการประกันภัยรถยนต์ คือ แบบจำลอง ARIMA(11,1,11) ซึ่งสามารถเขียนประมาณพารามิเตอร์จากการทำผลต่างอันดับ 1 ได้สมการต่อไปนี้

$$\begin{aligned} y_t - y_{t-1} = {} & -1.0613(y_{t-1} - y_{t-2}) - 1.0097(y_{t-2} - y_{t-3}) - 1.0783(y_{t-3} - y_{t-4}) \\ & -1.1732(y_{t-4} - y_{t-5}) - 1.1027(y_{t-5} - y_{t-6}) - 1.0852(y_{t-6} - y_{t-7}) \\ & -1.0930(y_{t-7} - y_{t-8}) - 0.9793(y_{t-8} - y_{t-9}) - 0.9191(y_{t-9} - y_{t-10}) \\ & -0.9706(y_{t-10} - y_{t-11}) - 0.9416(y_{t-11} - y_{t-12}) - 0.6713\epsilon_{t-1} \\ & -0.8304\epsilon_{t-2} - 1.0108\epsilon_{t-3} - 0.7419\epsilon_{t-4} - 0.6340\epsilon_{t-5} - 0.9486\epsilon_{t-6} \\ & -0.3297\epsilon_{t-7} - 0.3768\epsilon_{t-8} - 0.6682\epsilon_{t-9} - 0.1032\epsilon_{t-10} \\ & -0.1273\epsilon_{t-11} + \epsilon_t \end{aligned}$$

จัดรูปสมการต่อได้ดังนี้

$$\begin{aligned} y_t = {} & -0.0613y_{t-1} + 0.0516y_{t-2} - 0.0686y_{t-3} - 0.0949y_{t-4} + 0.0705y_{t-5} \\ & +0.0175y_{t-6} - 0.0078y_{t-7} + 0.1137y_{t-8} + 0.0602y_{t-9} - 0.0515y_{t-10} \\ & +0.0290y_{t-11} + 0.9416y_{t-12} - 0.6713\epsilon_{t-1} - 0.8304\epsilon_{t-2} - 1.0108\epsilon_{t-3} \\ & -0.7419\epsilon_{t-4} - 0.6340\epsilon_{t-5} - 0.9486\epsilon_{t-6} - 0.3297\epsilon_{t-7} - 0.3768\epsilon_{t-8} \\ & -0.6682\epsilon_{t-9} - 0.1032\epsilon_{t-10} - 0.1273\epsilon_{t-11} + \epsilon_t \end{aligned}$$

ต่อไปจะนำข้อมูลจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์ มาพิจารณาในแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซงและแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ



ภาพที่ 3.39 แสดง ACF และ PACF ของจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์

จากภาพที่ 3.39 เมื่อพิจารณาค่าสหสัมพันธ์ (ACF) และ ค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) พบว่ามีค่าสหสัมพันธ์ (ACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,3,5,6,9 และ 12 ตามลำดับ และมีค่าสหสัมพันธ์ในตัวเองบางส่วน (PACF) อย่างมีนัยสำคัญ ณ lag ที่ 1,2,3,12 และ 13 ตามลำดับ เมื่อพิจารณาข้อมูลดังกล่าว เราจึงพิจารณาส่วนของข้อมูลในอดีต (previous observations) เท่ากับ ที่ 1,2,3,12 และ 13 ตามลำดับ และ ส่วนของค่าเฉลี่ยแบบมีเงื่อนไข (previous conditional means) เท่ากับ 1,2,3,5,6,9 และ 12 ตามลำดับของทั้งแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซงและแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ ซึ่งในการพิจารณานั้นเราจะเลือกแบบจำลองที่มีค่า AIC น้อยที่สุด

แบบจำลองที่ 2 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

แบบจำลอง	past_obs	past_ mean	ค่า AIC	แบบจำลอง	past_obs	past_ mean	ค่า AIC
1	1	1	870,746.5	22	12	1	825,396.5
2	1	2	905,234.7	23	12	2	874,092.0
3	1	3	1,108,413.0	24	12	3	1,013,554.0
4	1	5	1,035,483.0	25	12	5	886,321.9
5	1	6	1,105,818.0	26	12	6	809,524.4
6	1	9	1,107,798.0	27	12	9	889,931.8
7	1	12	1,078,022.0	28	12	12	918,589.1
8	2	1	861,791.4	29	13	1	1,279,094.0
9	2	2	1,026,189.0	30	13	2	1,279,394.0
10	2	3	1,042,671.0	31	13	3	1,275,482.0
11	2	5	1,015,242.0	32	13	5	1,286,554.0
12	2	6	1,057,617.0	33	13	6	1,289,072.0
13	2	9	1,057,328.0	34	13	9	1,288,264.0
14	2	12	1,053,967.0	35	13	12	1,255,995.0
15	3	1	1,048,015.0	**36**	**12**	**(1,2)**	**535,646.2**
16	3	2	914,949.9	⋮	⋮	⋮	⋮
17	3	3	835,398.0				
18	3	5	938,809.3				
19	3	6	919,005.8				
20	3	9	917,066.7				
21	3	12	929,142.5				

ตารางที่ 3.39 แสดงค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

จากตารางที่ 3.39 นำค่า previous observations และค่า previous conditional means มาพิจารณาจับคู่ทีละตัวพบว่าที่แบบจำลองที่ 26 ให้ค่า AIC น้อยที่สุด จึงนำค่า previous observations เท่ากับ 12 และค่า previous conditional means เท่ากับ 6 ไปจับคู่กับตัวอื่น แล้วพิจารณาค่า AIC โดยกับจับคู่ไปเรื่อยๆ จึงสรุปได้ว่าว่าแบบจำลองที่ 36 ซึ่งมีค่า previous observations เท่ากับ 12 และค่า previous conditional means เท่ากับ 1 และ 2 ให้ค่า AIC น้อยที่สุด นั่นคือ 535,646.2 ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.40

พารามิเตอร์	ค่าประมาณ
Intercept	0.2971
beta_12	0.9849
alpha_1	-0.0995
alpha_2	0.0945

ตารางที่ 3.40 แสดงค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง

จากการตารางที่ 3.40 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง สามารถเขียนประมาณพารามิเตอร์ ได้สมการต่อไปนี้

$$log(\lambda_t) = 0.297 + 0.9849Y_{t-12} - 0.0995\lambda_{t-1} + 0.0945\lambda_{t-2}$$

แบบจำลองที่ 3 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

แบบจำลอง	past_obs	past_ mean	ค่า AIC	แบบจำลอง	past_obs	past_ mean	ค่า AIC
1	1	1	1,647.696	22	12	1	1,645.594
2	1	2	1,650.357	23	12	2	1,648.329
3	1	3	1,662.068	24	12	3	1,656.967
4	1	5	1,658.164	25	12	5	1,649.285
5	1	6	1,661.929	26	12	6	1,643.875
6	1	9	1,662.029	27	12	9	1,649.515
7	1	12	1,660.290	28	12	12	1,652.424
8	2	1	1,647.180	29	13	1	1,670.932
9	2	2	1,657.959	30	13	2	1,670.945
10	2	3	1,658.621	31	13	3	1,670.784
11	2	5	1,657.220	32	13	5	1,671.282
12	2	6	1,659.579	33	13	6	1,671.397
13	2	9	1,659.578	34	13	9	1,671.360
14	2	12	1,659.296	35	13	12	1,669.908
15	3	1	1,658.928	**36**	**12**	**(1,2)**	**1,623.113**
16	3	2	1,651.058	⋮	⋮	⋮	⋮
17	3	3	1,645.975				
18	3	5	1,652.555				
19	3	6	1,651.441				
20	3	9	1,651.042				
21	3	12	1,651.852				

ตารางที่ 3.41 แสดงค่า AIC ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

จากตารางที่ 3.41 นำค่า previous observations และค่า previous conditional means มาพิจารณาจับคู่ทีละตัวพบว่าที่แบบจำลองที่ 26 ให้ค่า AIC น้อยที่สุด จึงนำค่า previous observations เท่ากับ 12 และค่า previous conditional means เท่ากับ 6 ไปจับคู่กับตัวอื่น แล้วพิจารณาค่า AIC โดยกับจับคู่ไปเรื่อยๆ จึงสรุปได้ว่าว่าแบบจำลองที่ 36 ซึ่งมีค่า previous observations เท่ากับ 12 และค่า previous conditional means เท่ากับ 1 และ 2 ให้ค่า AIC น้อยที่สุด นั่นคือ 1,623.113 ซึ่งมีค่าพารามิเตอร์ดังแสดงในตารางที่ 3.42

พารามิเตอร์	ค่าประมาณ
Intercept	0.29715
beta_12	0.98491
alpha_1	-0.09950
alpha_2	0.09454
sigmasq	0.00345

ตารางที่ 3.42 แสดงค่าพารามิเตอร์ของแบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ

จากการตารางที่ 3.42 แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ สามารถเขียนประมาณพารามิเตอร์ ได้สมการต่อไปนี้

$$log(\lambda_t) = 0.29715 + 0.98491Y_{t-12} - 0.09950\lambda_{t-1} + 0.09454\lambda_{t-2}$$

ต่อไปเราจะมาพิจารณาค่าคลาดเคลื่อนกำลังสองของทั้งสามแบบจำลอง โดยแสดงในตารางที่ 3.43

แบบจำลอง	ค่าคลาดเคลื่อนกำลังสอง
การใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ ARIMA(11,1,11)	2.51×10^{11}
แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบปัวซง	1.57×10^{12}
แบบจำลองอนุกรมเวลาเชิงการนับที่มีการแจกแจงแบบทวินามลบ	1.57×10^{12}

ตารางที่ 3.43 แสดงค่าคลาดเคลื่อนกำลังสองของแบบจำลอง

จากตารางที่ 3.43 เมื่อพิจารณาค่าคลาดเคลื่อนกำลังสอง เนื่องจากแบบจำลองการใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติให้ค่าคลาดเคลื่อนกำลังสองน้อยที่สุดคือ 2.51×10^{11} ดังนั้นแบบจำลองอนุกรมเวลาที่เหมาะสมสำหรับข้อมูลของจำนวนกรมธรรม์การเอาประกันการประกันภัยทางทะเลและขนส่ง คือ แบบจำลองการใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ ARIMA(11,1,11)

ขั้นตอนต่อไปจะเป็นการหาจำนวนกรมธรรม์พยากรณ์ของการเอาประกันการประกันภัยรถยนต์ เทียบกับจำนวนกรมธรรม์พยากรณ์ของการเอาประกันการประกันภัยรถยนต์จริงในอีก 12 เดือนหลังจากเดือนธันวาคม 2559 แสดงเป็นค่าในตารางที่ 3.43

วันที่	จำนวนกรมธรรม์จริง	ผลพยากรณ์
ม.ค. 2560	3,466,472	3,292,000
ก.พ. 2560	3,320,725	3,468,182
มี.ค. 2560	3,836,351	3,743,595
เม.ย. 2560	3,014,396	2,914,315
พ.ค. 2560	3,549,549	3,436,324
มิ.ย. 2560	3,609,481	3,412,424
ก.ค. 2560	3,158,057	3,106,135
ส.ค. 2560	3,442,114	3,495,032
ก.ย. 2560	3,324,129	3,272,617
ต.ค. 2560	3,119,536	3,048,474
พ.ย. 2560	3,149,463	3,199,640
ธ.ค. 2560	3,554,476	3,388,293

ตารางที่ 3.44 จำนวนกรมธรรม์พยากรณ์ของการเอาประกันการประกันภัยรถยนต์เทียบกับจำนวนกรมธรรม์ของการเอาประกันการประกันภัยรถยนต์จริง

3.2.1 สรุปผล

จากการศึกษาแบบจำลองอนุกรมเวลาของจำนวนกรมธรรม์ของการเอาประกันวินาศภัย เพื่อหาแบบจำลองอนุกรมเวลาที่เหมาะสมและพยากรณ์จำนวนกรมธรรม์ของการเอาประกันวินาศภัย จำนวน 3 ประเภท ได้แก่ การประกันอัคคีภัย (Fire insurance) การประกันภัยทางทะเลและขนส่ง (Marine and Transportation insurance) และ การประกันภัยรถยนต์ (Automobile insurance) โดยใช้ข้อมูลในช่วงเดือน มกราคม พ.ศ. 2555 ถึง เดือนธันวาคม พ.ศ. 2559 จากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) แสดงในตารางที่ 3.45

อนุกรมเวลาของมูลค่าการเอาประกันวินาศภัย	แบบจำลองที่เหมาะสม
การประกันอัคคีภัย	การใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ ARIMA(10,1,10)
การประกันภัยทางทะเลและขนส่ง	การใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ ARIMA(12,1,12)
การประกันภัยรถยนต์	การใช้การแปลงข้อมูลโดยการหาค่ารากที่สองของข้อมูลและการแจกแจงแบบปกติ ARIMA(11,1,11)

ตารางที่ 3.44 แสดงแบบจำลองที่เหมาะสมสำหรับการพยากรณ์จำนวนกรมธรรม์ของการเอาประกันวินาศภัย

รายการอ้างอิง

[1] Anderson, O.D., “Time Series Analysis and Forecasting – The Box – Jenkins Approach,” Butterworths, London, 1975

[2] Lee, R. D., and Carter, L. R. (1992). “Modeling and Forecasting U.S. Mortality”. Journal of the American Statistical Association. 87(419): 659-671

[3] Statistical tools for high-throughput data analysis. “Normality Test in R”. แหล่งที่มา http://www.sthda.com/english/wiki/normality-test-in-r

[4] Tobias L. ,Konstantinos F. and Roland F. (2017). “An R Package for Analysis of Count Time Series following Generalized Linear Models”. Journal of Statistical Software.

[5] กรินทร์ กาญทนานนท์ (2561) *การพยากรณ์เชิงสถิติ.*กรุงเทพมหานคร:ซีเอ็ดยูเคชั่น

[6] ทรงศิริ แต้สมบัติ (2549) *การพยากรณ์เชิงปริมาณ.* กรุงเทพมหานคร:สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์

[7] พฤทธ์สรรค์สุทธิไชยเมธี. (2553) *วิเคราะห์แนวโน้มการส่งออกยางพาราของประเทศไทย ไปตลาดจีน* งานวิจัย. ทุนวิจัยวิทยาลัยเทคโนโลยีสยาม.

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

[9] สำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.). “การประกันวินาศภัย”. แหล่งที่มา http://www.oic.or.th/th/consumer/insurance/about/non-life

[10] สำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.). “รายงานการรับประกันภัยจำแนกตามช่องทางการขาย”. แหล่งที่มา https://www.oic.or.th/th/industry/statistic/data/43/2

ภาคผนวก

ภาคผนวก ก
แบบเสนอหัวข้อโครงงาน รายวิชา 2301399 Project Proposal
ภาคต้นปีการศึกษา 2562

ชื่อโครงงาน (ภาษาไทย)	การศึกษามูลค่าการรับประกันวินาศภัยภัยโดยใช้แบบจำลองอนุกรมเวลา
ชื่อโครงงาน (ภาษาอังกฤษ)	A study of Non-life insurance using time series models
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. จิราพรรณ สุนทรโชติ
ผู้ดำเนินการ	นายเจษฎากร น้อยท่าช้าง เลขประจำตัวนิสิต 5933508223 สาขาวิชา คณิตศาสตร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

หลักการและเหตุผล

ในปัจจุบันการทำประกันวินาศภัยนั้นเป็นที่ได้รับความนิยมเป็นอย่างมาก เนื่องจากสามารถให้ความคุ้มครองต่อบุคคล ครอบครัวและทรัพย์สิน แม้กระทั่งช่วยให้เกิดความมั่นคงในการประกอบธุรกิจจากความเสี่ยงที่อาจจะเกิดขึ้นได้ในอนาคต ทำให้บริษัทประกันภัยนั้นจำเป็นที่จะต้องหาแนวทางในการคาดการณ์แนวโน้มของมูลค่าการรับประกันวินาศภัยล่วงหน้า เพื่อลดความเสี่ยงของการลงทุนและเพิ่มผลกำไรที่จะได้รับในการลงทุนของบริษัท

การวิเคราะห์อนุกรมเวลา (Time Series Analysis) เป็นวิธีการทางสถิติวิธีหนึ่งซึ่งเป็นที่นิยมในการนำมาใช้พยากรณ์ข้อมูลที่จะเกิดขึ้นในอนาคต ซึ่งจากการศึกษาเกี่ยวกับการวิเคราะห์อนุกรมเวลา (Time Series Analysis) พบว่ามีการสร้างแบบจำลองขึ้นมามากมาย แบบจำลองอนุกรมเวลาที่ได้รับความนิยม ได้แก่ แบบจำลอง ARIMA (Autoregressive Integrated Moving Average) ซึ่งพัฒนาโดย George E.P. Box และ Gwilym M. Jenkins ในปี ค.ศ. 1970 ซึ่งให้ค่าพยากรณ์ในระยะสั้นที่ดีและได้มีการนำมาประยุกต์ใช้อย่างกว้างขวาง เช่น แบบจำลองลี-คาร์เทอร์ (Lee Carter Models, 1992) ได้ใช้

แบบจำลอง ARIMA ในการพยากรณ์อัตรามรณะของประชากรในประเทศสหรัฐอเมริกา, พฤทธ์สรรค์ สุทธิไชยเมธี (2553) วิเคราะห์อุปสงค์การนำเข้ายางพาราของประเทศจีนจากประเทศไทย ได้แก่ ยางแผ่นรมควัน ยางแท่งและน้ำยางข้น ในช่วงปี พ.ศ. 2538-2547 ด้วยแบบจำลอง ARIMA และพยากรณ์ล่วงหน้าช่วงไตรมาส 1-4 ในปี พ.ศ. 2548 เป็นต้น

ในโครงงานนี้ จะศึกษาหาแบบจำลองที่เหมาะสมที่ใช้ในการพยากรณ์มูลค่าการรับประกันวินาศภัยรายปี จากข้อมูลสถิติการรับประกันวินาศภัย จำแนกตามประเภทการประกันภัย พ.ศ. 2551-2560 โดยสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย กระทรวงการคลัง

วัตถุประสงค์

ในโครงงานนี้มีวัตถุประสงค์เพื่อศึกษาหาแบบจำลองอนุกรมเวลาที่เหมาะสมที่ใช้ในการพยากรณ์มูลค่าการรับประกันวินาศภัยรายปี ของข้อมูลสถิติการรับประกันวินาศภัย จำแนกตามประเภทการประกันภัย พ.ศ. 2551-2560 โดยสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย กระทรวงการคลัง

ขอบเขตของโครงงาน

แบบจำลองที่ศึกษาในโครงงานนี้คือ กลุ่มตัวแบบการถดถอยในตัวกำกับการเคลื่อนที่(ARIMA model) และข้อมูลที่ใช้ในการศึกษาคือข้อมูลสถิติการรับประกันวินาศภัย จำแนกตามประเภทการประกันภัย พ.ศ. 2551-2560 โดยสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย กระทรวงการคลัง โดยจะทำการศึกษาโครงงานนี้ผ่านโปรแกรม R

วิธีการดำเนินงาน

1. ศึกษาความรู้พื้นฐานเกี่ยวกับประกันวินาศภัย
2. ศึกษาความรู้พื้นฐานเกี่ยวกับข้อมูลการวิเคราะห์อนุกรมเวลา (Time Series Analysis)
3. ศึกษาความรู้พื้นฐานเกี่ยวกับแบบจำลองทางอนุกรมเวลา
4. ศึกษาการใช้โปรแกรม R
5. ศึกษาหาแบบจำลองอนุกรมเวลาที่เหมาะสม
6. สรุปผลและเขียนรายงาน

ตารางเวลาการดำเนินการ

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. ศึกษาความรู้พื้นฐานเกี่ยวกับประกันวินาศภัย											
2. ศึกษาความรู้พื้นฐานเกี่ยวกับข้อมูลการวิเคราะห์อนุกรมเวลา (Time Series Analysis)											
3. ศึกษาความรู้พื้นฐานเกี่ยวกับแบบจำลองทางอนุกรมเวลา											
4. ศึกษาการใช้โปรแกรม R											
5. ศึกษาหาแบบจำลองอนุกรมเวลาที่เหมาะสม											
6. สรุปผลและเขียนรายงาน											

ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ที่คาดว่าจะได้รับต่อผู้ดำเนินงาน

1.มีความรู้เกี่ยวกับคุณสมบัติและการประยุกต์ใช้แบบจำลองอนุกรมเวลา

2.มีความรู้เกี่ยวกับประกันวินาศภัย

3.มีความรู้ในการใช้โปรแกรม R

ประโยชน์ที่ได้จากโครงงานที่พัฒนาขึ้น

1.เพื่อใช้เป็นความรู้ต่อยอดงานวิจัย

2.สามารถคาดการณ์มูลค่าการรับประกันวินาศภัยในอนาคต

อุปกรณ์และเครื่องมือที่ใช้

ฮาร์ดแวร์

1. เครื่องคอมพิวเตอร์
2. เครื่องพิมพ์

ซอร์ฟแวร์

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

เอกสารอ้างอิง

[1] Anderson, O.D., “Time Series Analysis and Forecasting – The Box – Jenkins Approach,” Butterworths, London, 1975

[2] Lee, R. D., and Carter, L. R. (1992). “Modeling and Forecasting U.S. Mortality”. Journal of the American Statistical Association. 87(419): 659-671

[3] พฤทธ์สรรค์สุทธิไชยเมธี. (2553) วิเคราะห์แนวโน้มการส่งออกยางพาราของประเทศไทย ไปตลาดจีน งานวิจัย. ทุนวิจัยวิทยาลัยเทคโนโลยีสยาม.

[4] สำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.). “การประกันวินาศภัย”. แหล่งที่มา http://www.oic.or.th/th/consumer/insurance/about/non-life

ประวัติผู้เขียน

นายเจษฎากร น้อยท่าช้าง

รหัสนิสิต 5933508223

สาขาวิชา คณิตศาสตร์

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย