



ใบรับรองวิทยานิพนธ์  
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตร์มหาบัณฑิต (สถิติ)

ปริญญา

สถิติ

สถิติ

สาขา

ภาควิชา

เรื่อง การเปรียบเทียบวิธีการพยากรณ์ระหว่างวิธีเครือข่ายเบย์เซียนและวิธีการวิเคราะห์อนุกรม  
เวลาสำหรับดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย

A Comparison of Forecasting Methods between Bayesian Network and Time Series  
Analysis for the Stock Price Index of Property and Construction Groups in Thailand

นามผู้วิจัย นางสาวเมธินี กวินภาส

ได้พิจารณาเห็นชอบโดย

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

( รองศาสตราจารย์ประสิทธิ์ พย์คณพงษ์, M.S. )

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

( ผู้ช่วยศาสตราจารย์บุญอ้อม โคมที, Ph.D. )

หัวหน้าภาควิชา

( รองศาสตราจารย์ประสิทธิ์ พย์คณพงษ์, M.S. )

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

( รองศาสตราจารย์กัญญา วีระกุล, D.Agr. )

คณบดีบัณฑิตวิทยาลัย

วันที่ ..... เดือน ..... พ.ศ. ....

วิทยานิพนธ์

เรื่อง

การเปรียบเทียบวิธีการพยากรณ์ระหว่างวิธีเครือข่ายเบย์เซียนและวิธีการวิเคราะห์อนุกรมเวลา  
สำหรับดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย

A Comparison of Forecasting Methods between Bayesian Network and Time Series Analysis for  
the Stock Price Index of Property and Construction Groups in Thailand

โดย

นางสาวเมธินี กวินภาส

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์  
เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ)

พ.ศ. 2558

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เมธีณี กวินภาส 2558: การเปรียบเทียบวิธีการพยากรณ์ระหว่างวิธีเครือข่ายเบย์เซียนและวิธีการวิเคราะห์อนุกรมเวลาสำหรับดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย ปรินญาวิทยาศาสตรมหาบัณฑิต (สถิติ) สาขาสถิติ ภาควิชาสถิติ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก: รองศาสตราจารย์ประสิทธิ์ พยัคฆพงษ์, M.S. 108 หน้า

ในงานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาเปรียบเทียบเทคนิคการวิเคราะห์อนุกรมเวลาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย โดยทำการศึกษาเปรียบเทียบวิธีการพยากรณ์ 3 วิธี คือ วิธีเครือข่ายเบย์เซียน วิธีบ็อกซ์-เจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ซึ่งวิธีการพยากรณ์ที่เหมาะสมจะพิจารณาจากค่าวัดความคลาดเคลื่อนทั้ง 3 ค่า คือ ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ที่ต่ำที่สุด

อนุกรมเวลาที่ทำการศึกษาเป็นข้อมูลทุติยภูมิประกอบด้วย ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง ดัชนีเซท 50 ดัชนีราคาวัสดุก่อสร้าง ดัชนีราคาที่อยู่อาศัย ดัชนีความเชื่อมั่นทางธุรกิจ อัตราดอกเบี้ย และดัชนีราคาที่ดิน ซึ่งเป็นอนุกรมเวลารายเดือนตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556 รวมระยะเวลาทั้งสิ้น 60 เดือน

ผลการวิจัยพบว่า อนุกรมเวลาของดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างมีลักษณะการเคลื่อนไหวของแนวโน้มเส้นตรงแต่ไม่มีอิทธิพลของฤดูกาล และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโสมต์มีสมการพยากรณ์ คือ  $\hat{Y}_{t+p} = \hat{Y}_t + pS_{t-1}$  เป็นตัวแบบที่เหมาะสมสำหรับการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทยมากที่สุด เพราะให้ค่าวัดความคลาดเคลื่อน MAD MSE และ MAPE ต่ำที่สุด รองลงมา คือ วิธีบ็อกซ์-เจนกินส์ ส่วนวิธีเครือข่ายเบย์เซียนจะให้ค่าวัดความคลาดเคลื่อนสูงที่สุด จากนั้นนำรูปแบบที่เหมาะสมที่สุดมาหาช่วงการพยากรณ์ล่วงหน้า ผลการศึกษาพบว่า การพยากรณ์ล่วงหน้าในระยะเวลาสั้น ๆ 1 - 4 เดือน เหมาะสำหรับอนุกรมเวลาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย

Maythinee Kawinpas 2015: A Comparison of Forecasting Methods between Bayesian Network and Time Series Analysis for the Stock Price Index of Property and Construction Groups in Thailand. Master of Science (Statistics), Major Field: Statistics, Department of Statistics. Thesis Advisor: Associate Professor Prasit Payakkapong, M.S. 108 pages.

The purpose of this research is to study and compare time series analysis for the stock price of property and construction groups in Thailand by using the three forecasting methods : Bayesian Network, Box-Jenkins and Exponential Smoothing method. The suitable forecasting methods are chosen by considering the smallest value of mean absolute deviation (MAD), mean square error (MSE) and mean absolute percent error (MAPE).

The time series data in this study consists of stock price of property and construction, set 50 index, construction materials price index, housing price index, business sentiment index, interest rate and land price index from January 2009 to December 2013, totally 60 months.

The results of the study are follows: the data of stock price of property and construction have trend but no seasonal variation. Holt's exponential smoothing method with forecasting model  $\hat{Y}_{t+p} = \hat{Y}_t + p\hat{S}_{t-1}$  is the best method for forecasting stock price of property and construction in Thailand because it displays the smallest value of MAD, MSE and MAPE. The next one is Box-Jenkins and the last one is Bayesian network which displays the largest value. Then, using the best method for finding the best suitable forecasting period and the result of study show that the 1 - 4 month periods are suitable for time series of stock price of property and construction in Thailand.

---

Student's signature

---

Thesis Advisor's signature

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สามารถเสร็จสมบูรณ์ได้ด้วยความกรุณาและความช่วยเหลือเอาใจใส่อย่างยิ่งจากอาจารย์ที่ปรึกษาวิทยานิพนธ์ ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ประสิทธิ์ พยัคฆพงษ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ผู้ช่วยศาสตราจารย์บุญอ้อม โฉมทิ อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ที่ได้ให้คำปรึกษา คำแนะนำ ตลอดจนการตรวจแก้ไขข้อบกพร่องต่าง ๆ จนกระทั่งวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี และขอกราบขอบพระคุณ ผศ. ดร.กศยา ปลั่งพงษ์พันธ์ ผู้ทรงคุณวุฒิภายนอกสถาบันที่ให้คำแนะนำและข้อคิดเห็นเพิ่มเติม

ขอกราบขอบพระคุณอาจารย์ภาควิชาสถิติทุกท่านที่ได้อบรมสั่งสอนและมอบความรู้อันเป็นประโยชน์อย่างยิ่งในการนำไปใช้ประโยชน์ต่อไป และขอขอบคุณเจ้าหน้าที่ภาควิชาสถิติที่ช่วยให้คำแนะนำและช่วยอำนวยความสะดวกในด้านต่างๆ ด้วยดี

ขอกราบขอบพระคุณบิดา มารดาและครอบครัวที่คอยอบรมสั่งสอน สนับสนุน ให้กำลังใจ และให้โอกาสทางการศึกษาตลอดจนกำลังทรัพย์เป็นอย่างดียิ่งเสมอมา ขอขอบคุณพี่ๆ น้องๆ และเพื่อนๆ ทุกคนที่ให้กำลังใจและความช่วยเหลือในการทำวิทยานิพนธ์ฉบับนี้มาโดยตลอด

สำหรับวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชาพระคุณแก่บิดา มารดา อาจารย์ และผู้ที่มีส่วนเกี่ยวข้องทุกท่านทั้งที่ได้เอ่ยนาม และมิได้เอ่ยนามมา ณ ที่นี้ หากมีข้อผิดพลาดหรือข้อบกพร่องประการใด ผู้วิจัยขอน้อมรับไว้แต่เพียงผู้เดียว

เมธินี กวินภาส  
ธันวาคม 2557

## สารบัญ

หน้า

สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(4)
คำนำ	1
วัตถุประสงค์	4
การตรวจเอกสาร	6
อุปกรณ์และวิธีการ	53
อุปกรณ์	53
วิธีการ	53
ผลและวิจารณ์	57
ผล	57
วิจารณ์	85
สรุปและข้อเสนอแนะ	87
สรุป	87
ข้อเสนอแนะ	88
เอกสารและสิ่งอ้างอิง	89
ภาคผนวก	92
ภาคผนวก ก ข้อมูลปัจจัยทางเศรษฐกิจต่างๆ ที่ทำการศึกษตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	93
ภาคผนวก ข ตัวอย่างการวิเคราะห์อนุกรมเวลา	101
ประวัติการศึกษาและการทำงาน	108

## สารบัญตาราง

ตารางที่		หน้า
1	แสดงค่า $r_k$ ที่สัมพันธ์กับอนุกรมเวลาลักษณะต่างๆ	31
2	ค่า $\sim_Y, \dagger_Y^2, \dots_k$ และ $\dots_{kk}$ ของตัวแบบ ARMA(p,q)	34
3	แสดงน้ำหนักที่ให้กับค่าสังเกตด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว	40
4	ค่าวัดความคลาดเคลื่อนและค่าคลาดเคลื่อนมาตรฐาน	46
5	tracking signal และช่วงควบคุมเมื่อค่าความคลาดเคลื่อนเป็นอิสระกัน	47
6	ค่าความน่าจะเป็น $P(X_7 = x)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)	68
7	ค่าความน่าจะเป็น $P(X_6 = x)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)	68
8	ค่าความน่าจะเป็น $P(X_5   X_6)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)	68
9	ค่าความน่าจะเป็น $P(X_4   X_5, X_6)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)	69
10	ค่าความน่าจะเป็น $P(X_3   X_7)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)	70
11	ค่าความน่าจะเป็น $P(X_2   X_3, X_4)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)	70
12	ค่าความน่าจะเป็น $P(X_1   X_2)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)	71
13	ค่าความน่าจะเป็น $P(X_4   X_5)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ข)	72
14	ค่าความน่าจะเป็น $P(X_3 = x)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ค)	73
15	ค่าความน่าจะเป็น $P(X_2   X_3)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ง)	74
16	ค่าความน่าจะเป็น $P(X_2   X_4)$ ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (จ)	75
17	แสดงค่าความคลาดเคลื่อนจากการพยากรณ์ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างด้วยวิธีเครือข่ายเบย์เซียน	75
18	แสดงการประมาณค่าพารามิเตอร์จากรูปแบบ White Noise ของดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง	77
19	แสดงค่าสถิติของ Box-Pierce จากตัวแบบ White Noise ของดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง	80
20	แสดงค่าความคลาดเคลื่อนจากการพยากรณ์ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างด้วยวิธีบ็อกซ์-เจนกินส์	80

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
21	แสดงค่าความคลาดเคลื่อนจากการพยากรณ์ข้อมูลดัชนีราคาหุ่นกลุ่มอสังหาริมทรัพย์และก่อสร้างด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮสต์	81
22	ค่าวัดความคลาดเคลื่อนของการพยากรณ์ดัชนีราคาหุ่นกลุ่มอสังหาริมทรัพย์และก่อสร้างช่วงเดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	82
23	แสดงการเปรียบเทียบค่าพยากรณ์ ค่าจริง และค่าคลาดเคลื่อนด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮสต์กับค่าจริงตั้งแต่เดือนมกราคม 2557 ถึงเดือนมิถุนายน 2557	83
24	แสดงการคำนวณค่า $TS_1$ , $TS_2$ , $TS_3$ และ $TS_4$	84
25	แสดงการเปรียบเทียบวิธีการพยากรณ์ทั้ง 3 วิธี	87
ตารางผนวกที่		
ก1	ข้อมูลดัชนีราคาหุ่นกลุ่มอสังหาริมทรัพย์และก่อสร้าง ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	94
ก2	ข้อมูลดัชนีเซท 50 ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	95
ก3	ข้อมูลดัชนีราคาวัสดุก่อสร้าง ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	96
ก4	ข้อมูลดัชนีราคาที่อยู่อาศัย ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	97
ก5	ข้อมูลดัชนีความเชื่อมั่นทางธุรกิจ ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	98
ก6	ข้อมูลอัตราดอกเบี้ย ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	99
ก7	ข้อมูลดัชนีราคาที่ดิน ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	100
ข1	การวิเคราะห์ความแปรปรวนสำหรับอนุกรมเวลาดัชนีราคาหุ่นกลุ่มอสังหาริมทรัพย์และก่อสร้าง	107
ข2	การวิเคราะห์รูปแบบด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮสต์	107

## สารบัญภาพ

ภาพที่		หน้า
1	ข้อมูลมีลักษณะเป็นเส้นตรง	7
2	ข้อมูลมีลักษณะเป็นเอ็กซ์โปเนนเชียล	7
3	ข้อมูลมีลักษณะเป็นตัว S	8
4	ข้อมูลมีอิทธิพลของฤดูกาล	8
5	ข้อมูลที่มีอิทธิพลของวัฏจักร	9
6	กราฟ DAG ของเครือข่ายเบย์เซียน	13
7	แสดงการเชื่อมต่อแบบอนุกรม	15
8	แสดงการเชื่อมต่อแบบลู่ออก	15
9	แสดงการเชื่อมต่อแบบลู่ออก	16
10	แสดงแผนผังการคำนวณค่า K2 Algorithm	25
11	แผนภาพสรุปขั้นตอนการดำเนินการวิจัย	56
12	แสดงการเคลื่อนไหวของอนุกรมเวลาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และ ก่อสร้างของประเทศไทย ระหว่างเดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	57
13	แสดงการเคลื่อนไหวของอนุกรมเวลาดัชนีเซท 50 ของประเทศไทย ระหว่างเดือน มกราคม 2552 ถึงเดือนธันวาคม 2556	58
14	แสดงการเคลื่อนไหวของอนุกรมเวลาดัชนีราคาวัสดุก่อสร้างของประเทศไทย ระหว่างเดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	59
15	แสดงการเคลื่อนไหวของอนุกรมเวลาดัชนีราคาที่อยู่อาศัยของประเทศไทย ระหว่างเดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	60
16	แสดงการเคลื่อนไหวของอนุกรมเวลาดัชนีความเชื่อมั่นทางธุรกิจของประเทศไทย ระหว่างเดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	62
17	แสดงการเคลื่อนไหวของอนุกรมเวลาอัตราดอกเบี้ยของประเทศไทย ระหว่าง เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	63
18	แสดงการเคลื่อนไหวของอนุกรมเวลาดัชนีราคาที่ดินของประเทศไทย ระหว่าง เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556	64
19	เครือข่ายผลลัพท์ทั้ง 5 เครือข่ายที่ให้คะแนนสูงที่สุดโดยวิธีเครือข่ายเบย์เซียน	67

## สารบัญภาพ (ต่อ)

ภาพที่		หน้า
20	คอลเรลโลแกรมแสดง ACF และ PACF ของข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง หลังจากที่ได้มีการแปลงอนุกรมเวลาให้เป็นอนุกรมเวลาที่สแตชันนารีแล้ว	76
21	แสดงความสัมพันธ์ระหว่างค่าความคลาดเคลื่อนกับเวลาที่พยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์	78
22	แสดงกราฟ Normal Probability Plot ของค่าความคลาดเคลื่อนที่พยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์	79
23	แสดงความสัมพันธ์ระหว่างค่าคลาดเคลื่อนกับค่าจริงด้วยวิธีบ็อกซ์-เจนกินส์	79
<b>ภาพผนวกที่</b>		
ข1	แสดงหน้าต่างโปรแกรม Weka GUI Chooser	104
ข2	แสดงหน้าต่างโปรแกรม Weka Explorer	105

การเปรียบเทียบวิธีการพยากรณ์ระหว่างวิธีเครือข่ายเบย์เซียนและวิธีการวิเคราะห์อนุกรมเวลาสำหรับดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย

## A Comparison of Forecasting Methods between Bayesian Network and Time Series Analysis for the Stock Price Index of Property and Construction Groups in Thailand

### คำนำ

ธุรกิจอสังหาริมทรัพย์เป็นธุรกิจที่ต้องการเงินทุนระยะยาวในการลงทุนหรือเป็นธุรกิจที่ต้องใช้เงินจำนวนมากในการดำเนินงานและให้ผลตอบแทนที่สูงในระยะยาว ทำให้ผู้มีเงินทุนถึงแม้จะไม่มี ความชำนาญให้ความสนใจที่จะประกอบธุรกิจด้านนี้เป็นจำนวนมาก และธุรกิจอสังหาริมทรัพย์มีความจำเป็นต่อระบบเศรษฐกิจของประเทศไทยเป็นอย่างมาก โดยเป็นปัจจัยพื้นฐานอย่างหนึ่งที่มีความจำเป็นต่อการดำรงชีวิตของมนุษย์ในส่วนของที่อยู่อาศัย จึงเป็นเหตุให้ภาคอุตสาหกรรมอสังหาริมทรัพย์มีทั้งผู้ซื้อ และผู้ขายที่มีความต้องการที่อยู่อาศัยเป็นของตนเองที่แท้จริง หรือผู้ที่ต้องการประกอบเป็นธุรกิจเพื่อผลประโยชน์หรือการเก็งกำไร ดังนั้น เมื่อเวลาผ่านไป จำนวนประชากรเพิ่มมากขึ้น แต่ที่ดินที่อยู่อาศัยลดลง จึงส่งผลให้ราคาที่ดินเติบโตแบบคูณสองในเวลาเพียงไม่กี่ปี

แต่ปลายปี 2554 ที่ผ่านมา ประเทศไทยได้ประสบกับปัญหาอุทกภัยครั้งยิ่งใหญ่ซึ่งส่งผลต่อการเปลี่ยนแปลงอย่างมีนัยสำคัญต่อธุรกิจอสังหาริมทรัพย์ของประเทศไทยเป็นอย่างมาก โดยตลอดทั้งปี 2555 ได้เกิดปรากฏการณ์ใหม่ในภาคธุรกิจอสังหาริมทรัพย์อย่างชัดเจน ผู้ประกอบการได้มีการปรับเปลี่ยนกลยุทธ์ แนวทางในการทำธุรกิจที่ผ่านมา โดยส่วนใหญ่หันไปเปิดโครงการคอนโดมิเนียมเพิ่มมากขึ้น เพื่อทดแทนโครงการบ้านจัดสรรของผู้ที่ได้รับผลกระทบจากปัญหาใหญ่ในครั้ง นี้ ซึ่งนับตั้งแต่นั้นมา ธุรกิจอสังหาริมทรัพย์ได้ฟื้นตัวขึ้น ถือเป็นช่วงแห่งการเริ่มต้นธุรกิจอีกครั้ง หลังจากผ่านวิกฤตนี้ไปแล้ว แต่เป็นการฟื้นตัวของธุรกิจในส่วนของคนโดมเนียมเป็นหลัก ส่วนในครึ่งปีหลังของปี 2555 รัฐบาลได้มีแผนงานที่ชัดเจนในการป้องกันและแก้ไขปัญหาอุทกภัยทั้งในระยะสั้นและระยะยาว ทำให้ความเชื่อมั่นของผู้บริโภคและผู้ประกอบการต่อที่อยู่อาศัยในแนวราบ หวนกลับคืนมา

ในปัจจุบันธุรกิจอสังหาริมทรัพย์มีการพัฒนาเป็นไปในทิศทางที่ดีขึ้น มีการปรับปรุง แก้ไข แนวทางในการดำเนินการธุรกิจที่เน้นความมั่นคงให้มีประสิทธิภาพมากยิ่งขึ้น เช่น มีเกณฑ์การรับรู้ รายได้ให้กับลูกค้า หรือกฎหมายการควบคุมการก่อสร้างอาคารบ้านเรือน แต่ทั้งนี้ปัจจัยทาง เศรษฐกิจมีการเปลี่ยนแปลงอยู่ตลอดเวลา ทำให้ธุรกิจอสังหาริมทรัพย์ต้องประสบกับความผันผวน และยากต่อการควบคุม ทั้งผู้ประกอบการ นักลงทุน หรือทุกฝ่ายที่เกี่ยวข้องต่างก็จำเป็นที่จะต้องหา สาเหตุและปัจจัยที่มีผลต่อธุรกิจ โดยเฉพาะเหตุการณ์ที่เกิดขึ้นกับประเทศสหรัฐอเมริกา ซึ่งมีสาเหตุ มาจากสภาวะฟองสบู่ในภาคธุรกิจอสังหาริมทรัพย์ ทำให้เกิดผลกระทบในด้านอื่น ๆ ต่อเศรษฐกิจ ในหลายประเทศทั่วโลกตามมา เพราะฉะนั้นก่อนที่นักลงทุนจะเข้ามาลงทุนในกลุ่มอสังหาริมทรัพย์ นั้น จึงมีความจำเป็นที่จะต้องศึกษาหาความรู้ มีความเข้าใจเกี่ยวกับปัจจัยทางเศรษฐกิจต่างๆ ที่ส่งผล กระทบต่อดัชนีราคาหุ้นในกลุ่มวัสดุก่อสร้างนี้ เพื่อนำข้อมูลที่ได้จากการศึกษามาประกอบการ ตัดสินใจในการลงทุนได้อย่างเหมาะสม อีกทั้งยังป้องกันความเสี่ยงที่จะเกิดขึ้นและเพื่อให้ได้ ผลตอบแทนจากการลงทุนในระดับที่พึงพอใจ

หลายปีที่ผ่านมา เทคนิคการพยากรณ์ถูกพัฒนาอย่างต่อเนื่องเพื่อใช้ในการตัดสินใจ ดำเนินการอย่างถูกต้อง มีประสิทธิภาพ และน่าเชื่อถือ รูปแบบการพยากรณ์ที่เหมาะสมจะพิจารณา จากค่าความคลาดเคลื่อนที่น้อยที่สุด การเลือกวิธีพยากรณ์ที่เหมาะสมควรพิจารณาจากหลายๆ ปัจจัย ได้แก่ ช่วงเวลาของการพยากรณ์ เวลาที่ใช้ในการพยากรณ์ ลักษณะของข้อมูลอนุกรมเวลา และขนาดอนุกรมเวลา ซึ่งการพยากรณ์อนุกรมเวลาระยะสั้นที่นิยมใช้กัน คือ วิธีการปรับให้เรียบ แบบเอ็กซ์โปเนนเชียล และวิธีบ็อกซ์-เจนกินส์ แต่เนื่องจากการเก็บรวบรวมข้อมูลอนุกรมเวลาเป็น ข้อมูลดิบที่เกิดขึ้นในอดีต ข้อมูลอาจจะมีความคลุมเครือ ไม่ครบถ้วนสมบูรณ์ หรือข้อมูลไม่ แน่นนอน จึงเสนอการพยากรณ์โดยวิธีเครือข่ายเบย์เซียน ซึ่งประโยชน์ของวิธีเครือข่ายเบย์เซียน คือ เป็นเครื่องมือสำหรับหาเหตุผลภายใต้ความไม่แน่นอน และมีคุณสมบัติเป็น directed acyclic graph (DAG) ที่มีโหนดเป็นตัวแปรสุ่มและเส้นเชื่อมระหว่างโหนดต้องเป็นอิสระจากกัน ดังนั้น ใน การศึกษาครั้งนี้ผู้วิจัยสนใจศึกษาเทคนิคการพยากรณ์อนุกรมเวลาด้วยวิธีเครือข่ายเบย์เซียน วิธีบ็อกซ์-เจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล โดยนำวิธีดังกล่าวทั้ง 3 วิธี มา เปรียบเทียบกันเพื่อหาวิธีที่เหมาะสมที่สุด โดยการนำค่าพยากรณ์ที่ได้มาวัดความคลาดเคลื่อน ซึ่งค่า วัดความคลาดเคลื่อนของการพยากรณ์ที่ใช้ ได้แก่ ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย ค่าความคลาดเคลื่อน กำลังสองเฉลี่ย และเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย ซึ่งการวิจัยนี้มุ่งวิเคราะห์ถึงปัจจัย ทางเศรษฐกิจที่สำคัญที่ส่งผลกระทบต่อดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างอันได้แก่ ดัชนีเซท 50 ดัชนีราคาวัสดุก่อสร้าง ดัชนีราคาที่อยู่อาศัย อัตราดอกเบี้ย ดัชนีความเชื่อมั่นทางธุรกิจ และดัชนีราคาที่ดิน โดยทำการศึกษาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในตลาด

หลักทรัพย์แห่งประเทศไทย เพื่อเป็นแนวทางสำหรับผู้ที่มีความสนใจที่จะลงทุนในธุรกิจ  
อสังหาริมทรัพย์และก่อสร้าง และเพื่อเป็นข้อมูลให้กับนักลงทุน ผู้ประกอบการที่จะนำไปใช้ในการ  
ตัดสินใจในการเริ่มโครงการอสังหาริมทรัพย์ใหม่ๆ ที่จะเกิดขึ้น



## วัตถุประสงค์

1. เพื่อศึกษาความสัมพันธ์ระหว่างปัจจัยทางเศรษฐกิจที่สำคัญ ที่ส่งผลกระทบต่อ การเปลี่ยนแปลงดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง ได้แก่ ดัชนีเซท 50 ดัชนีราคาวัสดุก่อสร้าง ดัชนีราคาที่อยู่อาศัย ดัชนีความเชื่อมั่นทางธุรกิจ อัตราดอกเบี้ย และดัชนีราคาที่ดิน
2. เพื่อศึกษา เปรียบเทียบเทคนิคการวิเคราะห์อนุกรมเวลาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทยโดยวิธีเครือข่ายเบย์เซียน วิธีบ็อกซ์-เจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล
3. เพื่อสร้างสมการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทยโดยใช้ปัจจัยทางเศรษฐกิจที่กำหนด
4. เพื่อวิเคราะห์แนวโน้มดัชนีราคาหุ้นอสังหาริมทรัพย์และก่อสร้างในอนาคต

## ขอบเขตของการวิจัย

ในการศึกษาครั้งนี้ ใช้ข้อมูลทุติยภูมิประกอบด้วย ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย ข้อมูลดัชนีเซท 50 ดัชนีราคาวัสดุก่อสร้าง ดัชนีราคาที่อยู่อาศัย ดัชนีความเชื่อมั่นทางธุรกิจ อัตราดอกเบี้ย และดัชนีราคาที่ดิน ซึ่งข้อมูลเหล่านี้เป็นข้อมูลรายเดือนตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556 โดยที่ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง ดัชนีเซท 50 ได้ข้อมูลจากตลาดหลักทรัพย์แห่งประเทศไทย (ตลาดหลักทรัพย์แห่งประเทศไทย, 2557) ข้อมูลดัชนีราคาวัสดุก่อสร้าง ได้ข้อมูลจากสำนักดัชนีเศรษฐกิจการค้า (สำนักดัชนีเศรษฐกิจการค้า, 2557) ข้อมูลดัชนีราคาที่อยู่อาศัย อัตราดอกเบี้ย ดัชนีความเชื่อมั่นทางธุรกิจและดัชนีราคาที่ดิน ได้ข้อมูลจากธนาคารแห่งประเทศไทย (ธนาคารแห่งประเทศไทย, 2557)

เทคนิคการวิเคราะห์หอนุกรมเวลาที่ใช้ในการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์ และก่อสร้างในประเทศไทยที่ศึกษามี 3 วิธี ดังต่อไปนี้

1. วิธีเครือข่ายเบย์เซียน (Bayesian Network)
2. วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins)
3. วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing)

เกณฑ์ที่ใช้ในการเปรียบเทียบวิธีพยากรณ์ทั้ง 3 วิธี พิจารณาจากค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (Mean Absolute Deviation : MAD) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) และเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE)

#### ประโยชน์ที่คาดว่าจะได้รับ

1. ทราบถึงปัจจัยทางเศรษฐกิจสำคัญที่ส่งผลกระทบต่อดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์ และก่อสร้างในประเทศไทย
2. ทราบวิธีการพยากรณ์ที่เหมาะสมสำหรับการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างเมื่อข้อมูลเป็นรายเดือน
3. สามารถนำสมการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างไปใช้ในการวางแผนสำหรับนักลงทุนให้สามารถวิเคราะห์ พิจารณา ตัดสินใจลงทุนได้อย่างมีประสิทธิภาพสูงสุด
4. เป็นแนวทางสำหรับผู้สนใจที่ต้องการศึกษาวิธีการพยากรณ์ สามารถนำรูปแบบพยากรณ์ไปประยุกต์ใช้กับข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวคล้ายคลึงกับลักษณะการเคลื่อนไหวของอนุกรมเวลาของดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง

## การตรวจเอกสาร

การตรวจเอกสารแบ่งออกเป็น 3 ส่วน ได้แก่ ส่วนที่ 1 กล่าวถึงหัวข้อและทฤษฎีที่เกี่ยวข้องในการวิจัย ส่วนที่ 2 กล่าวถึงเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการพยากรณ์ และส่วนที่ 3 กล่าวถึงผลงานวิจัยที่เกี่ยวข้อง ดังรายละเอียดต่อไปนี้

### หัวข้อและทฤษฎีที่เกี่ยวข้องในการวิจัย

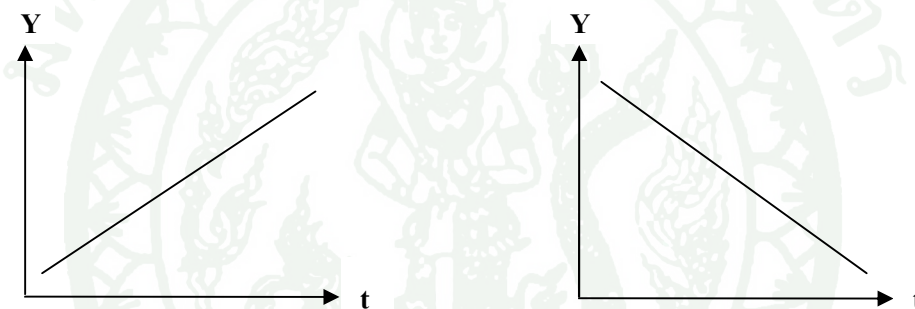
การพยากรณ์ หมายถึง การคาดคะเนหรือการทำนายลักษณะการเกิดของเหตุการณ์ต่าง ๆ หรือสภาพการณ์ในอนาคต โดยศึกษาจากแนวโน้มและรูปแบบการเกิดของเหตุการณ์หรือสภาพการณ์จากข้อมูลในอดีตที่มีการเก็บรวบรวมอย่างมีระบบ และ/หรือจากความรู้ ความสามารถ ประสบการณ์ และวิจารณญาณของผู้พยากรณ์ ภายใต้ข้อสมมติที่ว่า ลักษณะการเกิดของเหตุการณ์ต่าง ๆ ในอนาคต ไม่แตกต่างจากการเกิดของเหตุการณ์ในอดีต การพยากรณ์มีความสำคัญอย่างยิ่งต่อการวางแผนและการตัดสินใจเกี่ยวกับการดำเนินงานของบุคคลในทุกสาขาอาชีพ และองค์กรในสาขาต่าง ๆ ไม่ว่าจะองค์กรนั้นจะเป็นองค์กรขนาดเล็กหรือขนาดใหญ่ องค์กรของรัฐหรือเอกชน ในการพยากรณ์แต่ละครั้งย่อมต้องการค่าพยากรณ์ที่มีประสิทธิภาพและความแม่นยำสูง เพื่อช่วยในการตัดสินใจวางแผนงานต่าง ๆ เพราะเมื่อผู้บริหารองค์กรทราบว่าจะเกิดเหตุการณ์ใดหรือสภาพการณ์ใดในอนาคตด้วยความเชื่อมั่นระดับหนึ่งจะทำให้การวางแผนหรือการตัดสินใจในการดำเนินงานเป็นไปได้อย่างถูกต้อง ดังนั้น การพยากรณ์ที่ให้ความถูกต้องสูงจึงมีความจำเป็นอย่างยิ่ง

หลายปีที่ผ่านมา เทคนิคการพยากรณ์ได้รับการพัฒนาอย่างรวดเร็วและก้าวหน้าไปไกลมาก มีการใช้หลักการทางสถิติพัฒนามาอย่างต่อเนื่อง เพื่อใช้ในการตัดสินใจดำเนินการอย่างถูกต้อง มีประสิทธิภาพ และน่าเชื่อถือ ซึ่งการคำนวณวิธีการพยากรณ์แต่ละวิธีที่มีประสิทธิภาพพิจารณาจากค่าความคลาดเคลื่อนน้อยที่สุดของการพยากรณ์ นอกจากนี้การตัดสินใจเลือกเทคนิคการพยากรณ์วิธีใดวิธีหนึ่งนั้น จะต้องพิจารณาเลือกใช้ให้เหมาะสม โดยพิจารณาจากประเด็นต่างๆ หลายประการ ประกอบกัน เช่น ขนาดของอนุกรมเวลา ช่วงเวลาในการเก็บข้อมูล และระยะเวลาในการพยากรณ์ เป็นต้น เพราะการศึกษาปัจจัยต่าง ๆ ที่เกี่ยวข้องกับการพยากรณ์นี้ จะเป็นประโยชน์ต่อการนำค่าพยากรณ์ไปใช้ได้อย่างถูกต้องและเหมาะสมกับลักษณะงานต่าง ๆ

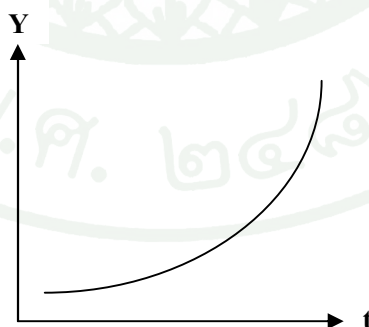
อนุกรมเวลา หมายถึง กลุ่มของค่าสังเกตที่เก็บรวบรวมตามเวลาอย่างต่อเนื่อง ซึ่งแต่ละค่าสังเกตจะเก็บรวบรวมในช่วงเวลาที่เท่ากัน ได้แก่ รายปี รายไตรมาส รายเดือน รายสัปดาห์ รายวัน รายชั่วโมง เป็นต้น

การเคลื่อนไหวของอนุกรมเวลาขึ้นกับส่วนประกอบหลัก 4 ส่วนประกอบ ได้แก่ แนวโน้ม (trend : T) อิทธิพลของฤดูกาล (seasonal effect : S) อิทธิพลของวัฏจักร (cyclical effect : C) และ เหตุการณ์ที่ผิดปกติ (irregular effect : I) การเคลื่อนไหวของอนุกรมเวลาอาจเกิดจากส่วนประกอบหลักมากกว่าหนึ่งส่วนประกอบ ซึ่งความหมายของแต่ละส่วนประกอบหลักของอนุกรมเวลามีดังนี้

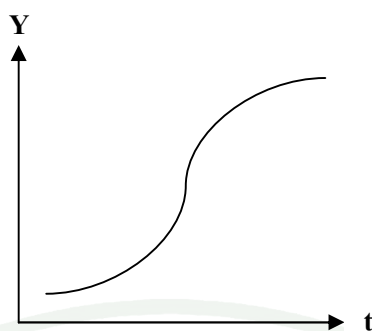
1. แนวโน้ม หมายถึง การเคลื่อนไหวของอนุกรมเวลาในระยะยาวเป็นไปได้ทั้งแนวโน้มขึ้นหรือแนวโน้มลง (upward trend or downward trend) ลักษณะแนวโน้มมีได้หลายลักษณะ เช่น แนวโน้มเส้นตรง (linear trend) แนวโน้มกำลังสอง (quadratic trend) แนวโน้มเอ็กซ์โปเนนเชียล (exponential trend) แนวโน้มกำลัง (power trend) แนวโน้มเอ็กซ์โปเนนเชียลดัดแปลง (modified exponential trend) และแนวโน้มแบบตัว S (S-shaped trend) เป็นต้น



ภาพที่ 1 ข้อมูลมีลักษณะเป็นเส้นตรง

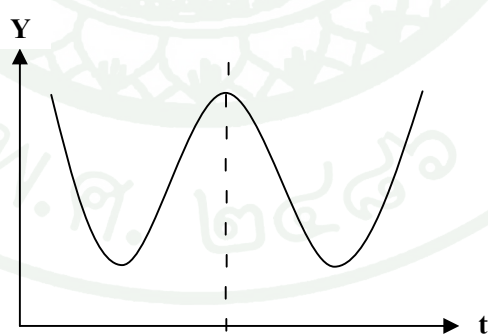


ภาพที่ 2 ข้อมูลมีลักษณะเป็นเอ็กซ์โปเนนเชียล



ภาพที่ 3 ข้อมูลมีลักษณะเป็นตัว S

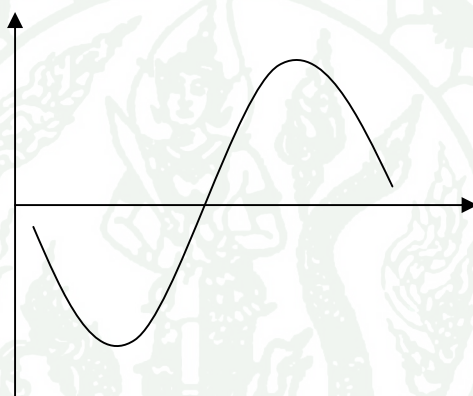
2. อิทธิพลของฤดูกาล หมายถึง การเคลื่อนไหวของอนุกรมเวลาที่เนื่องมาจากอิทธิพลของฤดูกาลเป็นการเคลื่อนไหวที่เกิดขึ้นซ้ำ ๆ ในแต่ละช่วง โดยลักษณะการเคลื่อนไหวในแต่ละช่วงไม่แตกต่างกัน ปัจจัยที่มีต่ออิทธิพลของฤดูกาลมีหลายปัจจัย เช่น สภาพอากาศ อุณหภูมิ สภาพทางสังคม วัฒนธรรม งบประมาณของหน่วยงานของรัฐ กำหนดการตามปฏิทินที่หน่วยงานกำหนด ช่วงของฤดูกาลอาจจะมีแบบที่ต่างกัน เช่น ในหนึ่งปีมี 52 สัปดาห์ ในหนึ่งปีมี 12 เดือน ในหนึ่งปีมี 4 ไตรมาส ในหนึ่งภาคการศึกษา มี 2 ภาคเรียน ในหนึ่งสัปดาห์มี 7 วัน เป็นต้น ในทางปฏิบัติอนุกรมเวลาที่นำมาศึกษาอิทธิพลของฤดูกาลมักเป็นอนุกรมเวลารายเดือน รายไตรมาส หรือรายงวดที่มีแบบแผนคงที่ (constant pattern) หรือไม่คงที่ ซึ่งการวิเคราะห์อนุกรมเวลาส่วนใหญ่จะสมมติให้มีแบบแผนคงที่



ภาพที่ 4 ข้อมูลมีอิทธิพลของฤดูกาล

3. อิทธิพลของวัฏจักร หมายถึง การเคลื่อนไหวของอนุกรมเวลาที่เก็บรวบรวมเป็นระยะเวลาต่อเนื่องกันหลายปีอาจจะแสดงอิทธิพลของวัฏจักร โดยหนึ่งวัฏจักรจะคลุมระยะเวลาหลายปี

แบบแผนวัฏจักรในแต่ละช่วงของวัฏจักรมักจะต่างกัน และช่วงของวัฏจักรจะสั้นยาวต่างกัน การเคลื่อนไหวของวัฏจักรที่พบบ่อย ๆ เช่น วัฏจักรธุรกิจ ซึ่งแบ่งเป็นช่วง ๆ ได้หลายแบบ ได้แก่ ช่วงรุ่งเรือง (growth) ช่วงคงที่ (maturity) และช่วงตกต่ำ (decline) หรือแบบแผนที่แบ่งเป็น 6 ช่วง ได้แก่ ช่วงเติบโต (growth) ช่วงรุ่งเรือง (prosperity) ช่วงเตือน (warning) ช่วงซบเซา (recession) ช่วงตกต่ำ (depression) และช่วงฟื้นตัว (recovery) การพยากรณ์การเคลื่อนไหวของวัฏจักรทำได้ยากกว่าการพยากรณ์การเคลื่อนไหวของฤดูกาล เพราะในแต่ละช่วงวัฏจักรมักมีแบบแผนการเคลื่อนไหวที่ต่างกันและช่วงวัฏจักรคลุมช่วงเวลาที่ไม่เท่ากัน ในแต่ละช่วงวัฏจักรค่าต่ำที่สุดและค่าสูงที่สุดมักมีการเปลี่ยนตำแหน่งบ่อยครั้ง สำหรับอนุกรมเวลาที่มีขนาดไม่ใหญ่มากอาจจะปรากฏวัฏจักรรวมอยู่ด้วยกับแนวโน้มเนื่องจากยังไม่แสดงอิทธิพลของวัฏจักรชัดเจน



ภาพที่ 5 ข้อมูลที่มีอิทธิพลของวัฏจักร

4. เหตุการณ์ที่ผิดปกติ เป็นเหตุการณ์ที่เกิดจากปัจจัยอื่น ๆ ที่นอกเหนือจากแนวโน้ม อิทธิพลของฤดูกาล และอิทธิพลของวัฏจักร นั่นคือ เป็นเหตุการณ์ที่ไม่ได้คาดมาก่อนว่าจะเกิดขึ้น หรือเป็นเหตุการณ์ที่ไม่เกิดขึ้นบ่อยครั้งนัก เช่น น้ำท่วม พายุ อุบัติเหตุ ปฏิวัติ สงคราม การเมือง แผ่นดินไหว เป็นต้น เหตุการณ์ดังกล่าวมีผลทำให้การเคลื่อนไหวของอนุกรมเวลาไม่มีแบบแผนที่แน่นอน

วิธีการทางสถิติที่ใช้ในการวิจัยนี้มี 3 วิธี ได้แก่ วิธีเครือข่ายเบย์เซียน วิธีบ็อกซ์-เจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ซึ่งมีรายละเอียดดังต่อไปนี้

## 1. วิธีเครือข่ายเบย์เซียน

การทำนายความสัมพันธ์ระหว่างตัวแปรของข้อมูลโดยวิธีเบย์เซียน เป็นวิธีทางคณิตศาสตร์ที่ใช้ทฤษฎีของเบย์ (Baye's theorem) ซึ่งแบ่งได้เป็น 2 กลุ่ม ได้แก่ กลุ่มแรกเป็นการทำนายแบบง่าย (Naïve Bayesian prediction) และอีกกลุ่มหนึ่งจะมีการคำนวณที่ซับซ้อนกว่าเรียกว่า เครือข่ายเบย์เซียน ซึ่งเป็นวิธีที่ผู้วิจัยจะนำมาใช้ในการวิจัย สามารถอธิบายรายละเอียดได้ดังนี้

### 1.1 ทฤษฎีของเบย์

กำหนดให้  $X$  เป็นตัวอย่างข้อมูลที่สนใจศึกษา  $H$  เป็นสมมติฐานเกี่ยวกับตัวอย่างข้อมูล  $X$  แล้ว  $P(H|X)$  หมายถึง ความน่าจะเป็นที่สมมติฐาน  $H$  จะเป็นจริงเมื่อกำหนดตัวอย่างข้อมูล  $X$  มาให้ โดยทั่วไป  $P(H)$ ,  $P(X)$  และ  $P(X|H)$  จะถูกประมาณได้จากข้อมูลที่มีอยู่ ส่วน  $P(H|X)$  สามารถคำนวณได้โดยใช้ทฤษฎีของเบย์มาช่วยในการคำนวณจาก  $P(H)$ ,  $P(X)$  และ  $P(X|H)$  ซึ่งเขียนเป็นสมการได้ดังนี้

$$P(H|X) = \frac{P(X|H) P(H)}{P(X)} \quad (1)$$

$P(H|X)$  เป็นการแจกแจงภายหลัง (posterior distribution) ของ  $H$

$P(X|H)$  เป็นความน่าจะเป็นอย่างมีเงื่อนไข (conditional probability) ของ  $X$

$P(X)$  เป็นการแจกแจงความน่าจะเป็น (probabilistic distribution) ของ  $X$

$P(H)$  เป็นการแจกแจงก่อนหน้า (prior distribution) ของ  $H$

กำหนดให้แต่ละตัวอย่างข้อมูล  $X=(x_1, \dots, x_n)$  ถูกแทนด้วยเวกเตอร์ขนาด  $n$  โดยที่  $n$  เป็นจำนวนของคุณลักษณะต่าง ๆ ของตัวอย่างข้อมูล  $X$  หรือ คือ จำนวนแอททริบิวต์ (attribute) ของข้อมูล  $(A_1, \dots, A_n)$  นั้นเอง จากนั้นแต่ละค่าความน่าจะเป็น  $P(X|H)$ ,  $P(H)$  และ  $P(X)$  สามารถคำนวณได้ดังต่อไปนี้

ค่าของ  $P(H)$  มีวิธีการคำนวณ 2 แบบ ขึ้นอยู่กับชนิดของข้อมูล ถ้าเป็นข้อมูลชนิดไม่ต่อเนื่อง (discrete data) จะได้ว่า

$$P(H) = \frac{\text{จำนวนข้อมูลที่เป็นไปตามสมมติฐาน}}{\text{จำนวนข้อมูลทั้งหมดที่พิจารณา}} \quad (2)$$

ถ้าเป็นข้อมูลชนิดต่อเนื่อง (continuous data) จะได้ว่า

$$P(H) = \frac{1}{\sqrt{2f} t_H} e^{-\frac{(H-\bar{H})^2}{2t_H^2}} \quad (3)$$

โดยที่  $\bar{H}$  คือ ค่าเฉลี่ยของข้อมูลที่เป็นไปตามสมมติฐาน H

$t_H$  คือ ส่วนเบี่ยงเบนมาตรฐานของข้อมูลที่เป็นไปตามสมมติฐาน H

ค่าของ P(X) มีวิธีการคำนวณ 2 แบบ ขึ้นอยู่กับชนิดของข้อมูล ถ้าเป็นข้อมูลชนิดไม่ต่อเนื่องจะได้ว่า

$$P(X) = \frac{\text{จำนวนข้อมูลที่มีค่าของข้อมูลตรงกับ X}}{\text{จำนวนข้อมูลทั้งหมดที่พิจารณา}} \quad (4)$$

ถ้าเป็นข้อมูลชนิดต่อเนื่อง จะได้ว่า

$$P(X) = \frac{1}{\sqrt{2f} t_x} e^{-\frac{(X-\bar{x})^2}{2t_x^2}} \quad (5)$$

โดยที่  $\bar{x}$  คือ ค่าเฉลี่ยของข้อมูลที่มีค่าของข้อมูลตรงตาม X

$t_x$  คือ ส่วนเบี่ยงเบนมาตรฐานของข้อมูลที่มีค่าของข้อมูลตรงตาม X

ค่าของ P(X|H) มีการคำนวณได้ 2 แบบ ขึ้นอยู่กับความสัมพันธ์กันของคุณลักษณะหรือแอททริบิวต์ของข้อมูลที่สนใจศึกษา ในกรณีที่แต่ละแอททริบิวต์เป็นอิสระจากกัน จะได้ว่า

$$P(X|H) = \prod_{i=1}^n P(x_i|H) \quad (6)$$

กรณีที่บางแอททริบิวต์ของข้อมูลที่มีความสัมพันธ์กัน จะได้ว่า

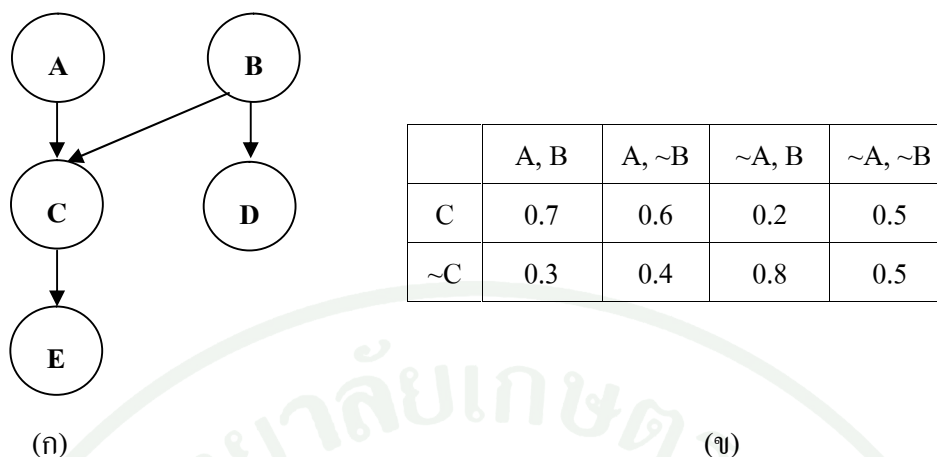
$$P(X|H) = \prod_{i=1}^n P(x_i | H, \text{Parent}(x_i)) \quad (7)$$

โดยที่  $x_i$  คือ ข้อมูลในแอททริบิวต์ที่  $i$

$\text{Parent}(x_i)$  คือ ข้อมูลในแอททริบิวต์อื่นที่มีผลต่อค่าของข้อมูลในแอททริบิวต์ที่  $i$

## 1.2 เครือข่ายเบย์เซียน

Ben-Gal (2007) กล่าวว่า เครือข่ายเบย์เซียนเป็นที่รู้จักกันในนามของเครือข่ายความเชื่อ (Bayes nets for short) ในการสร้างเครือข่ายเบย์เซียนเพื่อใช้งาน อาจทำได้โดยการเรียนรู้โครงสร้างเครือข่ายเบย์เซียนจากข้อมูลดิบซึ่งบอกถึงเหตุการณ์ที่เกิดขึ้น ปัญหาที่สำคัญในการเรียนรู้โครงสร้างเครือข่ายเบย์เซียน คือ การที่ต้องใช้เวลานานมากในการเรียนรู้โครงสร้างดังกล่าว ซึ่งเวลาที่ใช้ในการเรียนรู้จะขึ้นอยู่กับจำนวนตัวแปรและขนาดของข้อมูล ในกรณีที่ในข้อมูลมีจำนวนตัวแปรมากจะทำให้มีรูปแบบโครงสร้างที่จะต้องพิจารณามีจำนวนมากด้วย และในกรณีที่ขนาดของฐานข้อมูลใหญ่จะทำให้การเรียนรู้โครงสร้างใช้เวลามากขึ้นไปด้วย ซึ่งเครือข่ายเบย์เซียนนี้ประกอบด้วย 2 ส่วน ส่วนแรก คือ กราฟแบบมีทิศทางและไม่วนเป็นวงกลม (directed acyclic graph : DAG) ที่แต่ละโหนดแทนตัวแปรของข้อมูลหรือแอททริบิวต์ และเส้นเชื่อม (edge) แต่ละเส้นเชื่อมระหว่างโหนดแทนความน่าจะเป็นที่โหนดเหล่านั้นสัมพันธ์กัน ถ้าเส้นเชื่อมเส้นหนึ่งถูกลากจากโหนด A ไปยังโหนด B จะเรียก A ว่าเป็นโหนดพ่อแม่ (parent node) ของโหนด B และเรียก B ว่าเป็นโหนดที่สืบต่อ (descendent node) มาจาก A หรือเรียกอีกอย่างหนึ่งว่า เป็นโหนดลูก (children node) โดยทั่วไปแล้วตัวแปรจะมีได้หลายสถานะ เช่น สีของรถยนต์ (จะมีสถานะเป็น สีฟ้า, สีเขียว, สีแดง, สีดำ) จำนวนเด็กในครอบครัว (จะมีสถานะเป็น 0, 1, 2, 3, >3) หรือชนิดของเชื้อโรค (จะมีสถานะเป็น ไวรัส แบคทีเรีย) เป็นต้น โดยที่ความสัมพันธ์กันหรือขึ้นต่อกันในกราฟนี้สามารถประมาณได้โดยวิธีทางสถิติและคอมพิวเตอร์ ดังนั้น เครือข่ายเบย์เซียนจึงเป็นการรวมกันระหว่างทฤษฎีกราฟ ทฤษฎีความน่าจะเป็น วิทยาศาสตร์คอมพิวเตอร์ และสถิติ จากภาพที่ 6 เป็นตัวอย่างกราฟ DAG ซึ่งจะเห็นได้ว่า C จะทำงานได้ก็ต่อเมื่อถูกควบคุมโดย A และ B แต่ไม่ขึ้นอยู่กับ D และ E



ภาพที่ 6 กราฟ DAG ของเครือข่ายเบย์เซียน

ส่วนที่ 2 เป็นตารางความน่าจะเป็นอย่างมีเงื่อนไข (conditional probability table) หรือเรียกอีกอย่างหนึ่งว่า ตารางซีพีที (CPT) สำหรับแต่ละตัวแปร โดยที่ตารางซีพีทีของตัวแปร B เป็นการกระจายของความน่าจะเป็นอย่างมีเงื่อนไข ภาพที่ 6 (ข) แสดงตารางซีพีทีสำหรับตัวแปร C แต่ละค่าในตารางเป็นการรวมกันของค่าความเป็นไปได้ของค่าของโหนดพ่อแม่ของโหนด C เช่น

$$P(C = \text{“ทำงาน”} \mid A = \text{“ทำงาน”}, B = \text{“ทำงาน”}) = 0.7$$

$$P(C = \text{“ไม่ทำงาน”} \mid A = \text{“ไม่ทำงาน”}, B = \text{“ไม่ทำงาน”}) = 0.5$$

ความน่าจะเป็นของเหตุการณ์ A หรือ  $P(A)$  จะมีค่าอยู่ในช่วง  $[0,1]$  โดยที่

$$P(A) = 1 \text{ ก็ต่อเมื่อ เหตุการณ์ A เกิดขึ้นแน่นอน}$$

$$P(A) = 0 \text{ ก็ต่อเมื่อ เหตุการณ์ A ไม่เกิดขึ้นแน่นอน}$$

ถ้า A และ B เป็นเหตุการณ์ที่ไม่เกิดขึ้นร่วมกัน (mutually exclusive) แล้ว

$$P(A \cup B) = P(A) + P(B) \quad (8)$$

แนวคิดขั้นพื้นฐานของเบย์เซียนในเครือข่ายสาเหตุ คือ ความน่าจะเป็นอย่างมีเงื่อนไข นั่นคือ ความน่าจะเป็นของเหตุการณ์ A เมื่อเหตุการณ์ B เกิดขึ้นแล้วจะมีค่าเท่ากับ x แทนสัญลักษณ์

$P(A|B)=x$  ซึ่งไม่ได้หมายความว่า เมื่อเหตุการณ์ B เป็นจริงแล้วความน่าจะเป็นของ A เท่ากับ x แต่หมายความว่า ถ้า B เป็นจริงและทุก ๆ เหตุการณ์ที่ไม่เกี่ยวข้องกับ A แล้ว ความน่าจะเป็นของเหตุการณ์ A มีค่าเท่ากับ x หรือ  $P(A)=x$

กฎพื้นฐานสำหรับความน่าจะเป็น คือ

$$P(A|B) P(B) = P(A \cap B) \quad (9)$$

เมื่อ  $P(A \cap B)$  แทน ความน่าจะเป็นของเหตุการณ์ A และ B

แต่ถ้ามี 3 เหตุการณ์ จะได้ความน่าจะเป็น ดังนี้

$$P(A|B \cap C) P(B|C) = P(A \cap B|C) \quad (10)$$

จาก (1) จะได้ว่า  $P(A|B) P(B) = P(B|A) P(A)$  ดังนั้น กฎของเบย์ คือ

$$P(B|A) = \frac{P(A|B) P(B)}{P(A)} \quad (11)$$

ในกรณีที่มีเหตุการณ์ C อีกเหตุการณ์หนึ่ง จะได้ว่า

$$P(B|A \cap C) = \frac{P(A|B \cap C) P(B|C)}{P(A|C)} \quad (12)$$

ประโยชน์ของเครือข่ายเบย์เซียน คือ เป็นเครื่องมือสำหรับการหาเหตุผลภายใต้ความไม่แน่นอน เครือข่ายเบย์เซียนมีคุณสมบัติเป็น DAG นั่นคือ มีโหนดเป็นตัวแปรสุ่ม และเส้นเชื่อมระหว่างโหนดต้องเป็นอิสระจากกัน การจะหาค่าการแจกแจงความน่าจะเป็นอย่างมีเงื่อนไขของแต่ละโหนดในเครือข่ายเบย์เซียนต้องรู้ค่าความน่าจะเป็นจากเครือข่ายเบย์เซียนที่กำหนดให้ และต้องทราบค่าความน่าจะเป็นร่วม (joint probability) ของตัวแปรทั้งหมด เมื่อกำหนดค่าของแต่ละโหนด การหาค่าความน่าจะเป็นร่วมของตัวแปรทั้งหมดในเครือข่ายเบย์เซียน ความน่าจะเป็นร่วมกันของ

แถวใด ๆ  $(b_1, \dots, b_n)$  ที่สอดคล้องกับตัวแปรหรือแอททริบิวต์  $B_1, \dots, B_n$  สามารถคำนวณได้จากกฎลูกโซ่ (chain rule) ต่อไปนี้

$$P(b_1, \dots, b_n) = \prod_{i=1}^n P(b_i | \text{Parent}(B_i)) \quad (13)$$

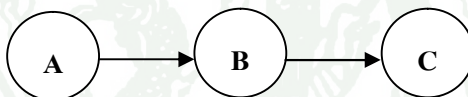
โดยที่ค่าสำหรับ  $P(b_i | \text{Parent}(B_i))$  สอดคล้องกับค่าในตารางซีพีทีสำหรับ  $B_i$

บางครั้งจะเรียก  $P(A|B)$  ว่า likelihood ของ B เมื่อเหตุการณ์ A เกิดขึ้นแล้ว และแทนด้วยสัญลักษณ์  $L(B|A)$  สมมติให้  $B_1, \dots, B_n$  เป็นเหตุการณ์ที่มีผลกระทบกับเหตุการณ์ A แล้ว

$$P(B_i | A) = \frac{P(A | B_i) P(B_i)}{P(A)} = k P(A | B_i) \text{ เมื่อ } k \text{ เป็นอิสระจาก } i$$

นิยามของคำศัพท์ที่เกี่ยวข้องกับเครือข่ายเบย์เซียน

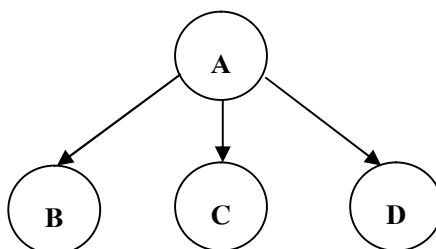
#### 1. การเชื่อมต่อแบบอนุกรม (serial connections)



ภาพที่ 7 แสดงการเชื่อมต่อแบบอนุกรม

จากภาพที่ 7 จะเห็นว่า A มีอิทธิพลต่อ B และ B มีอิทธิพลต่อ C นั่นคือ A จะส่งผลต่อ B อย่างแน่นอน และส่งผลต่อ C ด้วย แต่ถ้ารู้สถานะของ B แล้ว A และ C เป็นอิสระจากกัน จะเรียก A และ C ว่าเป็น d-separated เมื่อกำหนด B มาให้

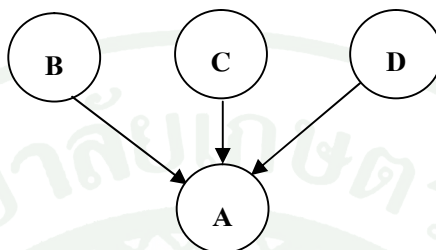
#### 2. การเชื่อมต่อแบบลู่ออก (diverging connections)



ภาพที่ 8 แสดงการเชื่อมต่อแบบลู่ออก

เมื่อรู้สถานะของ A อิทธิพลแต่ละตัวจะส่งผ่านไปยัง children ของ A ทุกตัว และจะเรียก B, C, D ว่าเป็น d-separated เมื่อกำหนด A มาให้

### 3. การเชื่อมต่อแบบลู่เข้า (converging connections)



ภาพที่ 9 แสดงการเชื่อมต่อแบบลู่เข้า

ถ้าไม่รู้สถานะเกี่ยวกับ A ยกเว้นอาจจะคาดการณ์ได้จาก parent B, C, D แล้ว parent นั้นจะเป็นอิสระจากกัน

จากทั้ง 3 กรณีข้างต้นนี้ จะครอบคลุมทุกเงื่อนไขที่มีการส่งผ่านระหว่างตัวแปร ทำให้สามารถเลือกได้ว่า จะเลือกตัวแปรคู่ใดเข้ามาในเครือข่ายสาเหตุไม่ว่าจะเป็นการขึ้นต่อกันของตัวแปรที่อยู่ในเครือข่ายก็ตาม

นิยาม d-separated ตัวแปร A และ B ในเครือข่ายสาเหตุจะเป็น d-separated ถ้าทุก ๆ เส้นเชื่อมระหว่าง A และ B มีความสัมพันธ์กันแบบอนุกรม หรือแบบลู่ออก แต่ถ้า A และ B ไม่เป็น d-separated จะเรียกว่า A และ B ว่า d-connected

#### 1.3 คะแนนเครือข่าย

เนื่องจากรงานวิจัยนี้ต้องการหาเครือข่ายที่เหมาะสมที่สุดที่จะเป็นตัวแทนของชุดข้อมูลที่สนใจศึกษา ดังนั้นต้องพิจารณาเครือข่ายทุกเครือข่ายที่เป็นไปได้ โดยการคิดคะแนนให้กับเครือข่ายทุกเครือข่ายที่เป็นไปได้แล้วเลือกเครือข่ายที่ทำให้ได้คะแนนสูงสุด กำหนดให้ D คือเครือข่าย d คือ ข้อมูล จากนั้นคำนวณหาค่าความน่าจะเป็นที่จะเกิด D เมื่อรู้ข้อมูล d

$$P(D|d) = \frac{P(D,d)}{P(d)} \quad (14)$$

Gregory and Herskovits (1993) ได้กำหนดให้ทุกเครือข่ายมีความน่าจะเป็นเริ่มต้นเท่ากัน คือ  $P(D)$  จะเห็นได้ว่า Posterior Distribution ของเครือข่ายเมื่อกำหนดข้อมูลมาให้จะขึ้นอยู่กับค่าของ  $P(D,d)$  ซึ่งคำนวณได้จาก

$$P(D|d) = \int_P P(d|D,P) f(P|D) P(D) dP \quad (15)$$

เมื่อ  $P$  คือ เซตของความน่าจะเป็นอย่างมีเงื่อนไข และ  $P(D)$  เป็นค่าคงที่ จึงสามารถเขียนสมการ (15) ใหม่ได้ดังนี้

$$P(D,d) = P(D) \int_P P(d|D,P) f(P|D) dP \quad (16)$$

กำหนดให้ การเก็บข้อมูลแต่ละครั้งเป็นอิสระจากกัน

$$P(D,d) = P(D) \int_P \left[ \prod_{h=1}^m P(C_h | D,P) \right] f(P|D) dP \quad (17)$$

เมื่อ  $m$  คือ จำนวนครั้งในการเก็บข้อมูลใน  $d$

$C_h$  คือ การเก็บข้อมูลครั้งที่  $h$  ใน  $d$

ในการเก็บข้อมูลแต่ละครั้ง ผู้ที่เก็บข้อมูลจะได้ข้อมูลของทุกตัวแปรที่สนใจ เรียกว่า ข้อมูลสมบูรณ์ (data complete) ดังนั้น จะเขียนสมการ (17) ให้เป็นตัวแปรแต่ละตัวแต่ละครั้งของการสังเกตดังนี้

$$P(D,d) = P(D) \int_P \left[ \prod_{h=1}^m \prod_{i=1}^n P(x_i = d_{ih} | W_{i \uparrow (i,h)}, P) \right] f(P|D) dP \quad (18)$$

เมื่อ  $n$  คือ จำนวนของตัวแปรทั้งหมด

$w_i$  คือ เซตของค่าที่เป็นไปได้ทั้งหมดของโหนดพ่อแม่ของตัวแปรที่  $i$

$w_{ij}$  คือ สมาชิกตัวที่  $j$  ในเซต  $w_i$

$T(i,h)$  แสดงถึงตำแหน่งของค่าของโหนดพ่อแม่ของตัวแปรที่  $i$  ในการเก็บข้อมูลครั้งที่  $h$  ว่าอยู่ในตำแหน่งที่เท่าไรของ  $w_i$

กำหนดให้  $q_i = |w_i|$

$V_i$  คือ เซตของค่าที่เป็นไปได้ทั้งหมดของตัวแปร  $x_i$  โดยที่  $r_i = |V_i|$

$r_{ijk}$  คือ จำนวนครั้งของการเก็บข้อมูลใน  $d$  ที่  $x_i = v_{ik}$  และโหนดพ่อแม่ของตัวแปร  $i$  มีค่าตรงกับสมาชิกลำดับที่  $j$  ของเซต  $w_i$  เขียนแทนด้วย  $w_{ij}$

$$P(D,d) = P(D) \int_p \left[ \prod_{h=1}^m \prod_{i=1}^n \prod_{k=1}^{r_i} P(x_i = v_{ik} | w_{ij}, P)^{r_{ijk}} \right] f(P|D) dP \quad (19)$$

กำหนดให้  $n_{ijk} = P(x_i = v_{ik} | w_{ij}, P)$  คือ ค่าความน่าจะเป็นอย่างมีเงื่อนไขที่ตัวแปรที่  $i$  จะมีค่าที่  $k$  เมื่อรู้ว่าโหนดพ่อแม่ของ  $i$  มีค่าที่  $j$

$f(n_{ij1}, \dots, n_{ijr_i})$  คือ ฟังก์ชันความหนาแน่นของความน่าจะเป็น (probability density function)

$$f(P|D) = \prod_{i=1}^n \prod_{j=1}^{q_i} f(n_{ij1}, \dots, n_{ijr_i}) \quad (20)$$

แทนค่าสมการ (20) ในสมการ (19)

$$P(D,d) = P(D) \int_{n_{ijk}} \left[ \prod_{i=1}^n \prod_{j=1}^{q_i} \prod_{k=1}^{r_i} n_{ijk}^{r_{ijk}} \right] \left[ \prod_{i=1}^n \prod_{j=1}^{q_i} f(n_{ij1}, \dots, n_{ijr_i}) \right] d_{n_{111}} \dots d_{n_{ijk}} \dots d_{n_{nq_i r_i}} \quad (21)$$

$$P(D,d) = P(D) * \prod_{i=1}^n \prod_{j=1}^{q_i} \int_{n_{ijk}} \left[ \prod_{k=1}^{r_i} n_{ijk}^{r_{ijk}} \right] f(n_{ij1}, \dots, n_{ijr_i}) d_{n_{ij1}} \dots d_{n_{ijr_i}} \quad (22)$$

เนื่องจากการกระจาย  $f(n_{ij1}, \dots, n_{ijr_i})$  มีรูปแบบเหมือนกันสำหรับทุก  $i$  ที่  $1 \leq i \leq n$  และทุก  $j$  ที่  $1 \leq j \leq n$  จะได้ว่า  $f(n_{ij1}, \dots, n_{ijr_i}) = C_{ij}$  เป็นค่าคงที่

$$\int \dots \int_{d_{ijk}} C_{ij} d_{ij1}, \dots, d_{ijr_i} = 1 \quad (23)$$

$$\text{โดยที่ } 0 \leq d_{ijk} \leq 1 \text{ และ } \sum_k d_{ijk} = 1$$

แทนค่าสมการ (23) ลงในสมการ (22)

$$P(D, d) = P(D) * \prod_{i=1}^n \prod_{j=1}^{q_i} C_{ij} \int \dots \int \left[ \prod_{k=1}^{r_i} d_{ijk}^{r_{ijk}} \right] d_{ij1}, \dots, d_{ijr_i} \quad (24)$$

การอินทิเกรตข้างต้นเรียกว่า Dirichlet's integral ซึ่งสามารถเขียนใหม่ได้ดังนี้

$$\int \dots \int_{d_{ijk}} \prod_{k=1}^{r_i} d_{ijk}^{r_{ijk}} d_{ij1}, \dots, d_{ijr_i} = \frac{\prod_{k=1}^{r_i} r_{ijk}!}{(N_{ij} + r_i - 1)!} \quad (25)$$

$$\text{โดยที่ } N_{ij} = \sum_{k=1}^{r_i} r_{ijk}$$

จากสมการ (25) ถ้าให้  $r_{ijk} = 0$  แล้ว  $N_{ij} = 0$  จะเห็นว่า  $(r_i - 1)! = C_{ij}$  นั่นเอง จากนั้นแทนสมการ (25) ลงในสมการ (24)

$$P(D, d) = P(D) \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} r_{ijk}! \quad (26)$$

ซึ่งสมการข้างต้นจะเป็นอัลกอริทึมที่ใช้ในการเรียนรู้โครงสร้างของเครือข่ายเบย์เซียนจากข้อมูล เรียกอีกอย่างหนึ่งว่า “K2 Algorithm”

$$f(i, f_i) = \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} r_{ijk}! \quad (27)$$

และจะใช้สมการ (27) นี้ในการคำนวณคะแนนเครือข่ายต่อไป

ตัวอย่าง การหาค่าคะแนน  $f(i, f_i)$  จากฐานข้อมูล (Ruiz, 2005) ต่อไปนี้

case	$x_1$	$x_2$	$x_3$
1	1	0	0
2	1	1	1
3	0	0	1
4	1	1	1
5	0	0	0
6	0	1	1
7	1	1	1
8	0	0	0
9	1	1	1
10	0	0	0

สมมติให้  $x_1$  เป็นตัวแปรเป้าหมาย

พิจารณา  $x_1(i=1)$  สามารถหาค่า  $f(1, w)$  ได้ดังนี้

$$f(1, w) = \frac{(r_1 - 1)!}{(N_{1-} + r_1 - 1)!} \prod_{k=1}^{r_1} r_{1,k}!$$

$r_{1,1} = 5$ : จำนวนกรณีที่  $x_1 = 0$  (กรณีที่ 3, 5, 6, 8, 10)

$r_{1,2} = 5$ : จำนวนกรณีที่  $x_1 = 1$  (กรณีที่ 1, 2, 4, 7, 9)

$$N_{1-} = r_{1,1} + r_{1,2} = 10$$

$$\text{จะได้ว่า } f(1, w) = \frac{(2-1)!}{(10+2-1)!} \cdot 5! \cdot 5! = \frac{1}{2772}$$

พิจารณา  $x_2(i=2)$  เมื่อ parent ของ  $x_2$  คือ  $\{x_1\}$  สามารถหาค่า  $f(2, \{x_1\})$  ได้ดังนี้

$$f(2, \{x_1\}) = \prod_{j=1}^{q_2} \frac{(r_2 - 1)!}{(N_{2j} + r_2 - 1)!} \prod_{k=1}^{r_2} r_{2jk} !$$

เนื่องจาก  $x_2$  มีค่าที่เป็นไปได้ 2 ค่า คือ  $\{0,1\}$  ดังนั้น  $r_2 = 2$

$$w_2 = ((x_1 = 0), (x_1 = 1))$$

$$q_2 = |w_2| = 2$$

$$r_{211} = 4: \text{จำนวนกรณีที่ } x_1 = 0 \text{ และ } x_2 = 0 \text{ (กรณีที่ 3, 5, 8, 10)}$$

$$r_{212} = 1: \text{จำนวนกรณีที่ } x_1 = 0 \text{ และ } x_2 = 1 \text{ (กรณีที่ 6)}$$

$$r_{221} = 1: \text{จำนวนกรณีที่ } x_1 = 1 \text{ และ } x_2 = 0 \text{ (กรณีที่ 1)}$$

$$r_{222} = 4: \text{จำนวนกรณีที่ } x_1 = 1 \text{ และ } x_2 = 1 \text{ (กรณีที่ 2, 4, 7, 9)}$$

$$N_{21} = r_{211} + r_{212} = 5$$

$$N_{22} = r_{221} + r_{222} = 5$$

$$\begin{aligned} \text{จะได้ว่า } f(2, \{x_1\}) &= \left[ \frac{(r_2 - 1)!}{(N_{21} + r_2 - 1)!} \cdot r_{211} ! \cdot r_{212} ! \right] \left[ \frac{(r_2 - 1)!}{(N_{22} + r_2 - 1)!} \cdot r_{221} ! \cdot r_{222} ! \right] \\ &= \left[ \frac{(2-1)!}{(5+2-1)!} \cdot 4! \cdot 1! \right] \left[ \frac{(2-1)!}{(5+2-1)!} \cdot 1! \cdot 4! \right] = \frac{1}{900} \end{aligned}$$

$$\text{เนื่องจาก } \frac{1}{900} > \frac{1}{2772} \text{ ดังนั้น } f_2 = \{x_1\}$$

พิจารณา  $x_3 (i=3)$  เมื่อ parent ของ  $x_3$  คือ  $\{x_1\}$  สามารถหาค่า  $f(3, \{x_1\})$  ได้ดังนี้

$$f(3, \{x_1\}) = \prod_{j=1}^{q_3} \frac{(r_3 - 1)!}{(N_{3j} + r_3 - 1)!} \prod_{k=1}^{r_3} r_{3jk} !$$

เนื่องจาก  $x_3$  มีค่าที่เป็นไปได้ 2 ค่า คือ  $\{0,1\}$  ดังนั้น  $r_3 = 2$

$$w_3 = ((x_1 = 0), (x_1 = 1))$$

$$q_3 = |w_3| = 2$$

$r_{311} = 3$ : จำนวนกรณีที่  $x_1 = 0$  และ  $x_3 = 0$  (กรณีที่ 5, 8, 10)

$r_{312} = 2$ : จำนวนกรณีที่  $x_1 = 0$  และ  $x_3 = 1$  (กรณีที่ 3, 6)

$r_{321} = 1$ : จำนวนกรณีที่  $x_1 = 1$  และ  $x_3 = 0$  (กรณีที่ 1)

$r_{322} = 4$ : จำนวนกรณีที่  $x_1 = 1$  และ  $x_3 = 1$  (กรณีที่ 2, 4, 7, 9)

$$N_{31} = r_{311} + r_{312} = 5$$

$$N_{32} = r_{321} + r_{322} = 5$$

$$\begin{aligned} \text{จะได้ว่า } f(3, \{x_1\}) &= \left[ \frac{(r_3 - 1)!}{(N_{31} + r_3 - 1)!} \cdot r_{311}! \cdot r_{312}! \right] \left[ \frac{(r_3 - 1)!}{(N_{32} + r_3 - 1)!} \cdot r_{321}! \cdot r_{322}! \right] \\ &= \left[ \frac{(2-1)!}{(5+2-1)!} \cdot 3! \cdot 2! \right] \left[ \frac{(2-1)!}{(5+2-1)!} \cdot 1! \cdot 4! \right] = \frac{1}{1800} \end{aligned}$$

$$f(3, \{x_2\}) = \prod_{j=1}^{q_3} \frac{(r_3 - 1)!}{(N_{3j} + r_3 - 1)!} \prod_{k=1}^{r_3} r_{3jk}!$$

$$w_3 = ((x_2 = 0), (x_2 = 1))$$

$$q_3 = |w_3| = 2$$

$r_{311} = 4$ : จำนวนกรณีที่  $x_2 = 0$  และ  $x_3 = 0$  (กรณีที่ 1, 5, 8)

$r_{312} = 1$ : จำนวนกรณีที่  $x_2 = 0$  และ  $x_3 = 1$  (กรณีที่ 3)

$r_{321} = 0$ : จำนวนกรณีที่  $x_2 = 1$  และ  $x_3 = 0$

$r_{322} = 5$ : จำนวนกรณีที่  $x_2 = 1$  และ  $x_3 = 1$  (กรณีที่ 2, 4, 6, 7, 9)

$$N_{31} = r_{311} + r_{312} = 5$$

$$N_{32} = r_{321} + r_{322} = 5$$

$$\begin{aligned} \text{จะได้ว่า } f(3, \{x_2\}) &= \left[ \frac{(r_3 - 1)!}{(N_{31} + r_3 - 1)!} \cdot r_{311}! \cdot r_{312}! \right] \left[ \frac{(r_3 - 1)!}{(N_{32} + r_3 - 1)!} \cdot r_{321}! \cdot r_{322}! \right] \\ &= \left[ \frac{(2-1)!}{(5+2-1)!} \cdot 4! \cdot 1! \right] \left[ \frac{(2-1)!}{(5+2-1)!} \cdot 0! \cdot 5! \right] = \frac{1}{180} \end{aligned}$$

เนื่องจาก  $\frac{1}{180} > \frac{1}{1800}$  ดังนั้น  $f_3 = \{x_2\}$

ผลจากการคำนวณ สรุปได้ว่า

โหนด  $x_1$  parent ของ  $x_1$  คือ  $f_1 = W$

โหนด  $x_2$  parent ของ  $x_2$  คือ  $f_2 = \{x_1\}$

โหนด  $x_3$  parent ของ  $x_3$  คือ  $f_3 = \{x_2\}$

ในการศึกษาครั้งนี้ จะพยากรณ์ดัชนีราคาหุ้นตามลำดับขั้นตอนดังต่อไปนี้

1. คำนวณค่าอัตราส่วนราคาต่อกำไร (price-earnings ratio หรือ P/E ratio) แทนสัญลักษณ์  $r_t$  ได้จากสูตรดังนี้

$$r_t = (\ln P_t - \ln P_{t-1}) \times 100$$

เมื่อ  $P_t$  แทนราคาปิดของหุ้น ณ เวลา  $t$

2. คำนวณเครือข่ายเบย์เซียนโดยใช้ K2 algorithm จากดังนี้

$$K2 = \prod_{i=1}^N \prod_{j=1}^M \frac{(L-1)!}{(N_{ij} + L - 1)!} \prod_{k=1}^L N_{ijk}!$$

โดยที่  $N_{ij} = \sum_{k=1}^L N_{ijk}$  แทน จำนวนตัวอย่างของ  $x_i$

เมื่อ  $N$  แทน จำนวน โหนด

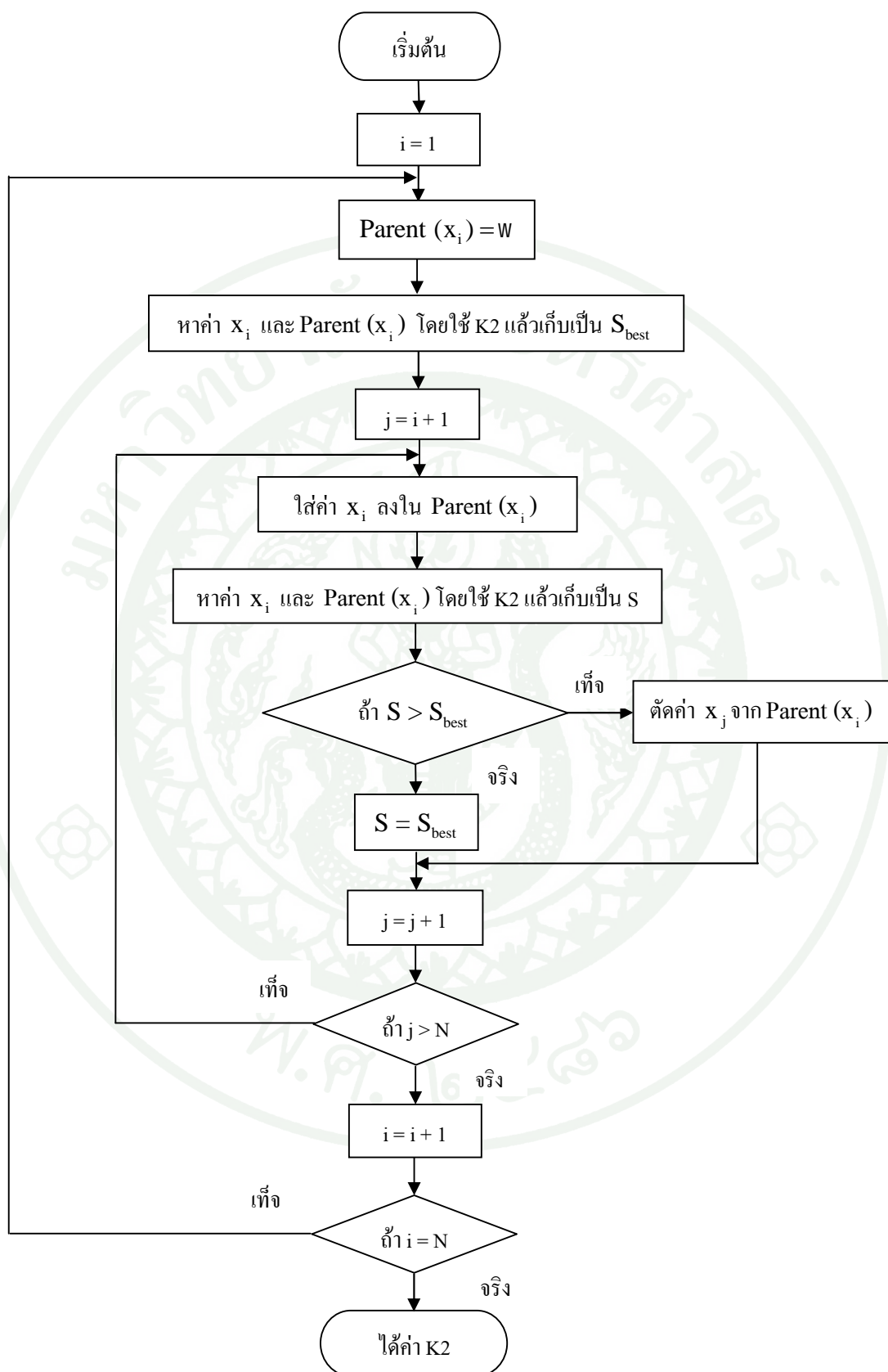
$L$  แทน จำนวนสถานะทั้งหมดของ  $x_i$

$M$  แทน จำนวนสถานะทั้งหมดของ Parent( $x_i$ )

การคำนวณขั้นตอนของ K2 ทำได้ดังนี้

1.  $i = 1$
2. ให้เซตของ parent สำหรับโหนด  $x_i$  เป็นเซตว่าง
3. ประมาณค่า K2 และ  $S_{best}$  ของเครือข่ายประกอบด้วย  $x_i$  และ Parent ( $x_i$ )
4. สำหรับ  $i = 1, \dots, N$
5. ใส่ค่า  $x_i$  ลงใน Parent ( $x_i$ )
6. ประมาณค่า K2 และ S ของเครือข่ายประกอบด้วย  $x_i$  และ Parent ( $x_i$ )
7. ตัดค่า  $x_i$  จาก Parent ( $x_i$ ) ถ้า  $S < S_{best}$
8.  $i = i + 1$
9. ย้อนกลับไปขั้นตอนที่ 2 ถ้า  $i \leq N$

จากขั้นตอนการคำนวณ K2 สรุปเป็นแผนภาพได้ดังนี้



ภาพที่ 10 แสดงแผนผังการคำนวณค่า K2 Algorithm

3. สำหรับการคำนวณเครือข่ายเบย์เซียน มีสมการพยากรณ์จากสูตรดังนี้

$$Y_t = \arg \left[ \max_{r_t} (P(r_t | B)) \right]$$

โดยที่

$r_t$  แทน เซตของอัตราส่วนราคาต่อกำไร เมื่อ  $t = 1, 2, \dots, L$

$L$  แทน จำนวนทั้งหมด

$B$  แทน เครือข่ายเบย์เซียน

$\arg[\max f(x)] = \{x | \forall y : (y \neq x \rightarrow f(y) \leq f(x))\}$  ค่าของ  $x$  ที่ทำให้  $f(x)$  มีค่าสูงสุด

สามารถประมาณค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) ได้จาก

$$r = \frac{\sum_{t=1}^n (r_t - \bar{r})(r'_t - \bar{r}')}{\sqrt{\sum_{t=1}^n (r_t - \bar{r})^2} \cdot \sqrt{\sum_{t=1}^n (r'_t - \bar{r}')^2}}$$

โดยที่  $\bar{r} = \frac{1}{n} \sum_{t=1}^n r_t$  เมื่อ  $r_t$  แทน ราคาหุ้นตามจริง ณ เวลา  $t$

$\bar{r}' = \frac{1}{n} \sum_{t=1}^n r'_t$  เมื่อ  $r'_t$  แทน ราคาหุ้นที่ได้จากการพยากรณ์ ณ เวลา  $t$

## 2. วิธีบ็อกซ์-เจนกินส์

ทรงศิริ (2549) อนุกรมเวลาบ็อกซ์-เจนกินส์เป็นวิธีการพยากรณ์ค่าในอนาคตที่พัฒนาโดยนักสถิติที่มีชื่อเสียงสองท่าน คือ George E.P.Box และ Gwilym M. Jenkins ในการพยากรณ์จะนำอนุกรมเวลาในอดีตมาพยากรณ์อนุกรมเวลาในอนาคต โดยจะมีการวิเคราะห์ถึงความเหมาะสมของตัวแบบก่อน ซึ่งในการวิเคราะห์เพื่อหาตัวแบบจะใช้ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (coefficient autocorrelation) และค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (partial coefficient autocorrelation) เมื่อได้ตัวแบบที่เหมาะสมแล้วจึงนำมาสร้างเป็นตัวแบบพยากรณ์ การวิเคราะห์

ข้อมูลอนุกรมเวลาตามวิธีบ็อกซ์-เจนกินส์จะใช้กับข้อมูลอนุกรมเวลาที่เป็นสเตชันนารี (stationary) ถ้าข้อมูลอนุกรมเวลาที่ไม่เป็นสเตชันนารี (non-stationary) เนื่องจากมีแนวโน้มหรือความผันแปรเนื่องจากฤดูกาล ต้องแปลงอนุกรมเวลาที่ไม่เป็นสเตชันนารีให้เป็นอนุกรมเวลาใหม่ที่เป็นสเตชันนารีแล้วจึงกำหนดตัวแบบให้กับอนุกรมเวลาใหม่นั้น ดังนั้นการวิเคราะห์ข้อมูลด้วยวิธีบ็อกซ์-เจนกินส์ แบ่งออกได้เป็น 2 ลักษณะตามลักษณะการเคลื่อนไหวของอนุกรมเวลาดังนี้

## 2.1 อนุกรมเวลาที่เป็นสเตชันนารี

อนุกรมเวลา  $\{y_t\}$  ที่มีค่าเฉลี่ยและค่าความแปรปรวนของ  $Y_t$  คงที่ นอกจากนั้นค่าสัมประสิทธิ์สหสัมพันธ์แบบออโตที่ช่วงเวลาห่าง  $k$  ขึ้นกับค่า  $k$  อย่างเดียว นั่นคือ ไม่ขึ้นกับเวลา  $t$  อนุกรมเวลาที่จะกำหนดรูปแบบ ARMA(p,q) ต้องเป็นอนุกรมเวลาที่เป็นสเตชันนารี ตัวอย่างของอนุกรมเวลาที่เป็นสเตชันนารี เช่น อนุกรมเวลาที่มีลักษณะการเกิดคงที่ ได้แก่ จำนวนผู้ป่วยนอกรายเดือน จำนวนผู้เข้ามาปรึกษาคลินิกกับสภานายความ อนุกรมเวลาที่มีการปรับการเจริญเติบโตแล้ว อนุกรมเวลาของค่าความคลาดเคลื่อน อนุกรมเวลาที่ได้จากการแปลงค่าสังเกต เป็นต้น

## 2.2 อนุกรมเวลาที่ไม่เป็นสเตชันนารี

อนุกรมเวลาที่ไม่มีความสมบัติสเตชันนารี เป็นอนุกรมเวลาที่ค่าเฉลี่ยหรือความแปรปรวนของ  $Y_t$  ไม่คงที่ที่เวลา  $t$  และค่าสัมประสิทธิ์สหสัมพันธ์แบบออโตที่ช่วงเวลาห่าง  $k$  ไม่ขึ้นอยู่กับเฉพาะค่า  $k$  แต่ขึ้นอยู่กับเวลา  $t$  ด้วย สำหรับอนุกรมเวลาที่มีแนวโน้มหรือฤดูกาลเป็นลักษณะของอนุกรมเวลาที่มีค่าเฉลี่ยหรือค่าคาดหวังของ  $Y_t$  หรือ  $\sim Y_t$  ไม่คงที่ ส่วนอนุกรมเวลาที่ความผันแปรของ  $Y_t$  ที่เวลา  $t$  ไม่คงที่จะเป็นลักษณะของอนุกรมเวลาที่ค่าความแปรปรวนของ  $Y_t$  หรือ  $t^2$  ที่เวลา  $t$  ไม่คงที่ อนุกรมเวลาที่มีค่าเฉลี่ยหรือค่าความแปรปรวนไม่คงที่ดังกล่าวจะไม่มีคุณสมบัติเป็นสเตชันนารี การแปลงอนุกรมเวลาเดิมที่ไม่เป็นสเตชันนารีให้เป็นอนุกรมเวลาใหม่ที่เป็นสเตชันนารีทำได้โดยการกำจัดแนวโน้ม กำจัดฤดูกาล หรือการแปลงอนุกรมเวลาให้มีค่าความแปรปรวนคงที่ที่เวลา  $t$  ด้วยการหาผลต่าง ผลต่างฤดูกาล หรือการแปลงค่าสังเกตในอนุกรมเวลาเป็นค่าสังเกตใหม่ด้วยบางฟังก์ชันคณิตศาสตร์ ตามลำดับ ซึ่งมีรายละเอียดดังนี้

### 2.2.1 อนุกรมเวลาผลต่าง (regular differencing)

ถ้าอนุกรมเวลาเดิม  $\{y_t\}$  มีการเคลื่อนไหวเนื่องจากแนวโน้ม จะแปลงอนุกรมเวลาเดิม  $\{y_t\}$  ให้เป็นอนุกรมเวลาใหม่  $\{z_t\}$  ที่ไม่มีแนวโน้ม โดยการหาผลต่างดังนี้

$$z_t = \nabla^d Y_t$$

โดยที่  $z_t$  เป็นค่าสังเกตใหม่ ณ เวลา  $t$  หลังจากกำจัดอิทธิพลของแนวโน้ม  
 $\nabla$  เป็นผลต่าง  
 $d$  เป็นจำนวนครั้งของการหาผลต่าง

จำนวนครั้งที่หาผลต่างขึ้นกับว่าเมื่อหาผลต่างแล้วอนุกรมเวลาใหม่เป็นสเตชันนารีแล้วหรือไม่ ถ้ายังไม่เป็นสเตชันนารีต้องหาผลต่างต่อไป โดยทั่วไปถ้าอนุกรมเวลามีแนวโน้มเส้นตรงจะหาผลต่างหนึ่งครั้ง และถ้าอนุกรมเวลามีแนวโน้มกำลังสองจะหาผลต่างสองครั้ง

### 2.2.2 อนุกรมเวลาผลต่างฤดูกาล (seasonal differencing)

ถ้าอนุกรมเวลาเดิม  $\{y_t\}$  มีการเคลื่อนไหวเนื่องจากฤดูกาล จะแปลงอนุกรมเวลาเดิม  $\{y_t\}$  ให้เป็นอนุกรมเวลาใหม่  $\{z_t\}$  ที่ไม่มีฤดูกาล โดยการหาผลต่างดังนี้

$$z_t = \nabla_L^D Y_t$$

โดยที่  $z_t$  เป็นค่าสังเกตใหม่ ณ เวลา  $t$  หลังจากกำจัดอิทธิพลของฤดูกาล  
 $\nabla$  เป็นผลต่าง  
 $D$  เป็นจำนวนครั้งของการหาผลต่างฤดูกาล  
 $L$  เป็นจำนวนฤดูกาลต่อปี

จำนวนครั้งที่หาผลต่างฤดูกาลขึ้นกับว่าเมื่อหาผลต่างฤดูกาลแล้วอนุกรมเวลาใหม่เป็นสเตชันนารีแล้วหรือไม่ ถ้ายังไม่เป็นสเตชันนารีต้องหาผลต่างฤดูกาลต่อไป

### 2.2.3 อนุกรมเวลาผลต่างและผลต่างฤดูกาล

ถ้าอนุกรมเวลาเดิม  $\{y_t\}$  มีการเคลื่อนไหวเนื่องจากแนวโน้มและฤดูกาล จะแปลงอนุกรมเวลาเดิม  $\{y_t\}$  ให้เป็นอนุกรมเวลาใหม่  $\{z_t\}$  ที่ไม่มีแนวโน้มและฤดูกาล โดยการหาผลต่างดังนี้

$$z_t = \nabla^d \nabla_L^D Y_t$$

จำนวนครั้งที่หาผลต่างและผลต่างฤดูกาลที่กำหนดด้วย  $d$  และ  $D$  ตามลำดับ ขึ้นกับว่าอนุกรมเวลาใหม่ที่สร้างขึ้นเป็นสเตชันนารีแล้วหรือไม่

### 2.2.4 อนุกรมเวลาที่แปลงแล้ว

เมื่อความผันแปรของอนุกรมเวลาไม่คงที่ นั่นคือ ค่าความแปรปรวนของ  $Y_t$  หรือ  $t_Y^2$  ที่เวลา  $t$  ไม่คงที่ ต้องทำการแปลงอนุกรมเวลาเดิม  $\{y_t\}$  ให้เป็นอนุกรมเวลาใหม่  $\{z_t\}$  เช่น  $Z_t = \ln(Y_t)$   $Z_t = \sqrt{Y_t}$  เป็นต้น

อนุกรมเวลาที่แปลงแล้วจะมีค่าความแปรปรวนคงที่ นอกจากนั้นผลการแปลงอนุกรมเวลาที่รูปแบบแนวโน้มฤดูกาลแบบคูณจะได้อนุกรมเวลาใหม่ที่มีรูปแบบแนวโน้มฤดูกาลแบบบวก นอกจากการแปลงด้วยฟังก์ชันของตัวแปรเดิมแล้ว บางครั้งการแปลงค่าสังเกตจากจำนวนครั้งที่เกิดเหตุการณ์ใดเหตุการณ์หนึ่งเป็นอัตราการเกิดเหตุการณ์ใดเหตุการณ์หนึ่ง จะได้อนุกรมเวลาที่เป็นสเตชันนารี การตรวจสอบว่าอนุกรมเวลาเป็นสเตชันนารีหรือไม่ พิจารณาเบื้องต้นได้จากการพล็อตอนุกรมเวลาหรือพิจารณาจากคอลเรลโลแกรม (correlogram) ของ  $r_k$

## 2.3 สัมประสิทธิ์สหสัมพันธ์ในตัวเองและสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน

### 2.3.1 สัมประสิทธิ์สหสัมพันธ์ในตัวเอง (...)

เป็นค่าวัดสหสัมพันธ์ระหว่างค่าสังเกตในอนุกรมเวลาที่อยู่ห่างกัน  $k$  ช่วงเวลา หรือเป็นค่าวัดสหสัมพันธ์ระหว่าง  $Y_t$  กับ  $Y_{t-k}$  โดยอธิบายได้ว่ามีสหสัมพันธ์ในทิศทางตามกัน

หรือตรงกันข้ามกันและสามารถบอกถึงขนาดของสหสัมพันธ์ ซึ่ง  $r_k$  เป็นค่าประมาณของ  $\dots_k$  โดยมีสูตรการคำนวณดังนี้

$$r_k = \frac{\sum_{t=k+1}^{n-k} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

โดยที่  $Y_t$  คือ ค่าสังเกต ณ เวลา  $t$

$\bar{Y}$  คือ ค่าเฉลี่ยของข้อมูลทั้งหมด ;  $\bar{Y} = \frac{\sum_{t=1}^n Y_t}{n}$

$k$  คือ จำนวนช่วงเวลาที่ข้อมูลอยู่ห่างกัน ;  $k = 1, 2, 3, \dots$

$n$  คือ จำนวนค่าสังเกตของอนุกรมเวลา

และ

$$V(r_k) = \begin{cases} \frac{1}{n} & ; k=1 \\ \frac{1}{n} \left( 1 + 2 \sum_{j=1}^{k-1} r_j^2 \right) & ; k=2, 3, \dots \end{cases}$$

คุณสมบัติของ  $r_k$  มีดังนี้

1.  $r_k$  มีค่าอยู่ระหว่าง -1 ถึง 1 เมื่อขนาดของ  $r_k$  ที่วัดด้วย  $|r_k|$  มีค่าเข้าใกล้ศูนย์ แสดงว่ามีความสัมพันธ์กันต่ำ ถ้า  $r_k$  มีค่าน้อยกว่าศูนย์ แสดงว่ามีความสัมพันธ์ในทิศทางที่ตรงกันข้าม แต่ถ้า  $r_k$  มีค่ามากกว่าศูนย์ แสดงว่ามีความสัมพันธ์กันในทิศทางเดียวกัน

2.  $r_k = r_{-k}$  นั่นคือ การวัดความสัมพันธ์ระหว่าง  $Y_t$  กับ  $Y_{t+k}$  หรือระหว่าง  $Y_t$  กับ  $Y_{t-k}$  ที่อยู่ห่างกัน  $k$  ค่าไม่แตกต่างกัน

ตารางที่ 1 แสดงค่า  $r_k$  ที่สัมพันธ์กับอนุกรมเวลาลักษณะต่างๆ

ลักษณะอนุกรมเวลา	ค่า $r_k$
1. ไม่มีแนวโน้มและความคลาดเคลื่อนเป็นอิสระ	$r_k$ มีค่าเท่ากับศูนย์ สำหรับทุกค่าของ $k$
2. ไม่มีแนวโน้มและความคลาดเคลื่อนไม่เป็นอิสระ	$r_k$ มีค่าสูง เมื่อ $k$ มีค่าน้อย และมีค่าเท่ากับศูนย์ เมื่อ $k$ มีค่ามาก
3. มีแนวโน้ม	$r_k$ มีค่ามากและเป็นบวก เมื่อ $k$ มีค่าน้อย และมีค่าลดลงอย่างต่อเนื่องซ้ำๆ เมื่อ $k$ มีค่าเพิ่มขึ้น
4. มีฤดูกาล	$r_k$ มีค่ามาก สำหรับ $k$ ที่มีค่าเป็นผลคูณของ $L, 2L, \dots$ เมื่อ $L$ เป็นจำนวนฤดูกาลต่อปี นั่นคือ สำหรับอนุกรมเวลารายเดือน $r_k$ มีค่ามาก เมื่อ $k = 6, 12, 24, \dots$ และสำหรับอนุกรมรายไตรมาส $r_k$ มีค่ามาก เมื่อ $k = 2, 4, 6, \dots$
5. มีแนวโน้มและฤดูกาล	$r_k$ มีค่าทำนองเดียวกับกรณีมีแนวโน้มและกรณีมีฤดูกาลรวมกัน

ที่มา: ทรงศิริ (2549)

### 2.3.2 สัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน ( $\dots_{kk}$ )

เป็นค่าวัดสหสัมพันธ์ระหว่างค่าสังเกตในอนุกรมเวลาที่อยู่ห่างกัน  $k$  ช่วงเวลา หรือเป็นค่าวัดสหสัมพันธ์ระหว่าง  $Y_t$  กับ  $Y_{t-k}$  เมื่อกำหนดให้ค่าสังเกต ณ เวลาอื่นคงที่ ซึ่ง  $r_{kk}$  เป็นค่าประมาณของ  $\dots_{kk}$  โดยมีสูตรการคำนวณดังนี้

$$r_{kk} = \begin{cases} r_1 & ; k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & ; k = 2, 3, 4, \dots \end{cases}$$

เมื่อ  $r_{k-j}$  คือ ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนที่ช่วงเวลาห่างกัน  $k$  และ  $j$

$$\text{โดยที่ } r_{k,j} = (r_{k-1,j} - r_{kk})r_{k-1,k-j} \quad ; j = 1, 2, \dots, k-1$$

$$k = 2, 3, \dots$$

$$\text{และ } V(r_{kk}) = \begin{cases} \frac{1}{n} & ; k = 1 \\ \frac{1}{n} \left( 1 + 2 \sum_{k=1}^{k-1} r_{kk}^2 \right) & ; k = 2, 3, \dots \end{cases}$$

## 2.4 ขั้นตอนการพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์

### 2.4.1 การกำหนดรูปแบบ (identification)

การกำหนดรูปแบบ ARMA(p,q) ให้กับอนุกรมเวลา โดยการเปรียบเทียบค่า  $r_k$  และ  $r_{kk}$  ของอนุกรมเวลาที่ศึกษากับค่า  $\dots_k$  และ  $\dots_{kk}$  ของรูปแบบ ARMA(p,q) ที่มีค่า  $p$  และ  $q$  ต่างกัน เนื่องจากต้องพิจารณาค่า  $r_k, r_{kk}, \dots_k$  และ  $\dots_{kk}$  พร้อมกันสำหรับหลายค่าของ  $k$  ดังนั้นแทนการเปรียบเทียบค่า จึงพิจารณาแผนภาพที่เรียกว่า คอเคลดโลแกรม ของ  $r_k, r_{kk}, \dots_k$  และ  $\dots_{kk}$  ที่ได้จากพล็อต  $r_k, r_{kk}, \dots_k$  และ  $\dots_{kk}$  กับค่า  $k$  แทนในทางปฏิบัติการกำหนดรูปแบบ

#### ก. รูปแบบ moving average หรือ รูปแบบ MA (q)

$$Y_t = u - \alpha_1 V_{t-1} - \dots - \alpha_q V_{t-q} + V_t$$

โดยที่  $Y_t$  เป็นค่าสังเกต ณ เวลา  $t$

$u$  เป็นค่าคงที่

$\alpha_i$  เป็นค่าพารามิเตอร์ MA อันดับที่  $i$

$V_t$  เป็นค่าความคลาดเคลื่อน

$q$  เป็นอันดับที่ของค่าเฉลี่ยเคลื่อนที่

ข. รูปแบบ autoregressive หรือ รูปแบบ AR (p)

$$Y_t = u + w_1 Z_{t-1} + \dots + w_p Z_{t-p} + v_t$$

โดยที่  $Y_t$  เป็นค่าสังเกต ณ เวลา  $t$

$u$  เป็นค่าคงที่

$w_i$  เป็นค่าพารามิเตอร์ AR อันดับที่  $i$

$v_t$  เป็นค่าความคลาดเคลื่อน

$p$  เป็นอันดับที่ของออโตรีเกรสซีฟ

ค. รูปแบบ mixed autoregressive – moving average หรือ รูปแบบ ARMA (p,q)

$$Y_t = u + w_1 Y_{t-1} + \dots + w_p Y_{t-p} - v_{t-1} - \dots - v_{t-q} + v_t$$

ง. รูปแบบ autoregressive integrated moving average หรือ รูปแบบ ARIMA(p,d,q)

$$Z_t = u + w_1 Z_{t-1} + \dots + w_p Z_{t-p} + v_t - v_{t-1} - \dots - v_{t-q}$$

โดยที่  $Z_t$  เป็นค่าสังเกต ณ เวลา  $t$  เมื่อ  $Z_t = \nabla^d Y_t$

$u$  เป็นค่าคงที่

$w_i$  เป็นค่าพารามิเตอร์ AR อันดับที่  $i$

$v_i$  เป็นค่าพารามิเตอร์ MA อันดับที่  $i$

$v_t$  เป็นค่าความคลาดเคลื่อน

$p$  เป็นอันดับที่ของออโตรีเกรสซีฟ

$q$  เป็นอันดับที่ของค่าเฉลี่ยเคลื่อนที่

จ. รูปแบบ seasonal integrated autoregressive and moving average order หรือ รูปแบบ SARIMA(P,D,Q)<sub>L</sub>

$$Z_t = u + w_1 Z_{t-s} + \dots + w_p Z_{t-ps} + v_t - v_{t-s} - \dots - v_{t-Qs}$$

โดยที่  $Z_t$  เป็นค่าสังเกต ณ เวลา  $t$   
 $u$  เป็นค่าเฉลี่ยคงที่  
 $w_i$  เป็นค่าพารามิเตอร์ SAR อันดับที่  $i$   
 $\alpha_i$  เป็นค่าพารามิเตอร์ SMA อันดับที่  $i$   
 $v_t$  เป็นค่าความคลาดเคลื่อน

ค่าลักษณะที่สำคัญของตัวแบบที่มีลักษณะสเตชันนารี ประกอบด้วยค่าเฉลี่ย ค่าความแปรปรวน ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองที่ช่วงเวลา  $k$  และค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนที่ช่วงเวลา  $k$  ซึ่งแต่ละตัวแบบมีค่าลักษณะที่ต่างกัน ดังตาราง

ตารางที่ 2 ค่า  $\gamma_Y$ ,  $\tau_Y^2$ ,  $\dots_k$  และ  $\dots_{kk}$  ของตัวแบบ ARMA(p,q)

ตัวแบบ	$\gamma_Y$	$\tau_Y^2$	$\dots_k$	$\dots_{kk}$
White noise	$u$	$\tau^2$	$\dots_k = 0, k \geq 1$	$\dots_{kk} = 0, k \geq 1$
AR(1)	$\frac{u}{1-w_1}$	$\frac{\tau^2}{1-w_1^2}$	$\dots_k = w_1^k, k=1$	$\dots_{11} = w_1$
AR(2)	$\frac{u}{1-w_1-w_2}$	$\frac{(1-w_2)\tau^2}{(1+w_2)((1-w_2)^2-w_1^2)}$	$\dots_0 = 1$	$\dots_{11} = \frac{w_1}{1-w_2}$
			$\dots_1 = \frac{w_1}{1-w_2}$	$\dots_{22} = w_2$
			$\dots_k = w_1 \dots_{k-1} + w_2 \dots_{k-2}, k > 1$	$\dots_{kk} = 0, k > 2$
MA(1)	$u$	$(1+\alpha_1^2)\tau^2$	$\dots_1 = \frac{\alpha_1}{1+\alpha_1^2}$	$\dots_{kk} = \frac{-\alpha_1^k(1-\alpha_1^2)}{1-\alpha_1^{2k+2}}$
			$\dots_k = 0, k > 1$	
MA(2)	$u$	$(1+\alpha_1^2+\alpha_2^2)\tau^2$	$\dots_1 = \frac{-\alpha_1(1-\alpha_2)}{1+\alpha_1^2+\alpha_2^2}$	ค่าลดลงเข้าสู่ 0
			$\dots_2 = \frac{-\alpha_2}{1+\alpha_1^2+\alpha_2^2}$	
			$\dots_k = 0, k > 2$	

## ตารางที่ 2 (ต่อ)

ตัวแบบ	$\sim_Y$	$\dagger_Y^2$	$\dots_k$
ARMA(1,1)	$\frac{u}{1-w_1}$	$\frac{(1-w_1^2 - 2w_1w_1) \dagger^2}{1-w_1^2}$	$\dots_1 = \frac{(1-w_1w_1)(w_1 - w_1)}{1-w_1^2 - 2w_1w_1}$ ค่าลดลงเข้าสู่ 0 $\dots_k = w_1^{k-1} \dots_1, k > 1$

ที่มา: ทรงศิริ (2549)

### 2.4.2 การประมาณค่าพารามิเตอร์ในรูปแบบ (estimation)

เป็นการหาค่าประมาณของพารามิเตอร์ในรูปแบบด้วยวิธีการประมาณแบบต่างๆ ได้แก่ วิธีการประมาณแบบง่าย วิธีการกำลังสองน้อยที่สุด คือการทำให้ค่าผลรวมกำลังสองของความคลาดเคลื่อนยกกำลังสองมีค่าน้อยที่สุด โดยวิธีที่ใช้คือ วิธีการกำลังสองน้อยที่สุดแบบไม่เป็นเชิงเส้นตรง (nonlinear least square) และวิธีภาวะน่าจะเป็นสูงสุด เป็นการประมาณค่าพารามิเตอร์ได้จากวิธีการประมาณค่าทางเชิงตัวเลข (numerical) รายละเอียดของการประมาณค่าพารามิเตอร์ด้วยวิธีต่างๆ มีดังนี้

1) วิธีการประมาณแบบง่าย เริ่มจากการสร้างสมการความสัมพันธ์ระหว่าง  $\dots_k$  และพารามิเตอร์ในรูปแบบ จำนวนสมการที่สร้างขึ้นเท่ากับจำนวนพารามิเตอร์ในรูปแบบที่ไม่รวมค่าคงที่  $w_0$  หรือเท่ากับ  $p+q$  เมื่อแทน  $\dots_k$  ด้วย  $r_k$  และแทนพารามิเตอร์  $w$  ด้วยตัวประมาณ  $\hat{w}$  ค่าตอบของสมการจะเป็นค่าประมาณของพารามิเตอร์ ส่วนค่าประมาณของ  $w_0$  ได้จากสมการความสัมพันธ์ระหว่าง  $\sim$  กับพารามิเตอร์ในรูปแบบ เมื่อแทน  $\sim$  ด้วยค่าเฉลี่ย  $\bar{Y}$  และแทนพารามิเตอร์  $w$  ด้วยตัวประมาณ  $\hat{w}$

2) วิธีการกำลังสองน้อยที่สุด ค่าประมาณของพารามิเตอร์ในรูปแบบที่ได้จากวิธีการกำลังสองน้อยที่สุดเป็นค่าประมาณที่ให้ผลรวมกำลังสองของความคลาดเคลื่อนของการพยากรณ์มีค่าน้อยที่สุด ค่าประมาณหาได้ไม่ง่ายจากการแก้สมการ เพราะมีการคำนวณที่ยุ่งยากและใช้เวลานาน จึงมีการใช้หลักการวิเคราะห์ตัวเลขด้วยคอมพิวเตอร์ช่วยในการคำนวณ การประมาณจะเริ่มจากการ

กำหนดค่าเริ่มต้นของพารามิเตอร์ซึ่งอาจเป็นค่าประมาณจากวิธีการประมาณแบบง่าย การคำนวณจะทำหลายรอบจนกว่าจะได้ค่าประมาณที่คงที่และให้ค่า SSE ต่ำที่สุด

3) วิธีภาวน่าจะเป็นสูงสุด มีหลักการทำนองเดียวกับวิธีกำลังสองน้อยที่สุด แต่วิธีภาวน่าจะเป็นสูงสุดต้องการข้อสมมติเกี่ยวกับลักษณะการแจกแจงแบบปกติของ  $v_t$  ค่าประมาณจากวิธีภาวน่าจะเป็นสูงสุดเป็นค่าประมาณที่ทำให้ฟังก์ชันภาวน่าจะเป็นมีค่าสูงที่สุด ขณะที่วิธีกำลังสองน้อยที่สุดให้ค่าประมาณที่ทำให้ SSE มีค่าต่ำที่สุด

#### 2.4.3 การตรวจสอบรูปแบบ (diagnostic checking)

หลังการกำหนดรูปแบบและประมาณค่าพารามิเตอร์ในรูปแบบแล้ว การวิเคราะห์ตรวจสอบความเหมาะสมของรูปแบบที่กำหนด โดยการพิจารณาลักษณะการเคลื่อนไหวของอนุกรมเวลาของค่าความคลาดเคลื่อน เพื่อตรวจสอบว่า ค่าคลาดเคลื่อนที่เกิดขึ้นมีการแจกแจงปกติ มีความแปรปรวนคงที่ และเป็นอิสระจากกัน ซึ่งการตรวจสอบจะพิจารณาจากค่าคลาดเคลื่อนตามวิธีการดังนี้

1) ตรวจสอบความแปรปรวนคงที่หรือไม่ โดยการพิจารณาจากกราฟของความคลาดเคลื่อนเทียบกับเวลา ถ้ากราฟมีลักษณะการกระจายอย่างสุ่มรอบเส้นที่ลากผ่าน  $v_t = 0$  ในแนวนอนสามารถยอมรับได้ว่าความแปรปรวนของค่าคลาดเคลื่อนคงที่ และขณะเดียวกันก็สามารถยอมรับได้ว่าค่าเฉลี่ยของค่าคลาดเคลื่อนเท่ากับศูนย์

2) ตรวจสอบอนุกรมเวลาของค่าคลาดเคลื่อนที่แต่ละเวลา  $t$  ว่ามีลักษณะการเคลื่อนไหวที่เป็นอิสระจากกัน โดยการทดสอบสมมติฐาน

$$H_0 : \dots_k (v_t) = 0$$

$$H_1 : \dots_k (v_t) \neq 0 ; k = 1, 2, \dots$$

ค่าสถิติทดสอบ

$$r_k(e_t) = \frac{\sum_{t=1}^{n-k} (e_t - \bar{e})(e_{t+k} - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2}$$

$$\text{โดยที่ } \bar{e} = \frac{\sum_{t=1}^n e_t}{n}$$

จะยอมรับ  $H_0$  ก็ต่อเมื่อ  $r_k(e_t)$  ตกอยู่ในช่วง  $\pm 2$  เท่าของส่วนเบี่ยงเบนมาตรฐานของค่าคลาดเคลื่อน นั่นคือ ตัวแบบที่กำหนดให้กับอนุกรมเวลามีความเหมาะสม

3) ตรวจสอบว่าพารามิเตอร์ในรูปแบบมีค่าเท่ากับศูนย์หรือไม่ โดยการทดสอบสมมติฐาน

$$H_0 : \mu = 0$$

$$H_1 : \mu \neq 0$$

ค่าสถิติทดสอบ

$$t = \frac{\hat{\mu}}{S_{\hat{\mu}}}$$

โดยที่  $\hat{\mu}$  เป็นค่าประมาณของพารามิเตอร์  $\mu$

$S_{\hat{\mu}}$  เป็นค่าประมาณของค่าคลาดเคลื่อนมาตรฐานของตัวประมาณ  $\hat{\mu}$

จะยอมรับ  $H_0$  ก็ต่อเมื่อ  $-t_{\frac{r}{2}, n-1} \leq t \leq t_{\frac{r}{2}, n-1}$  นั่นคือ พารามิเตอร์ใด ๆ ไม่เหมาะสมที่จะอยู่ในตัวแบบที่ระดับนัยสำคัญ  $r$

4) ตรวจสอบสัมประสิทธิ์สหสัมพันธ์ในตัวเองของค่าคลาดเคลื่อน ( $\dots_k(v_t)$ ) นั่นคือ เป็นการพิจารณาว่าค่าคลาดเคลื่อนของการพยากรณ์ที่อยู่ห่างกัน 1, 2, 3, ..., m ช่วงเวลาเป็นอิสระจากกันหรือไม่ โดยการทดสอบสมมติฐาน

$$H_0 : \dots_1(v_t) = \dots_2(v_t) = \dots = \dots_m(v_t) = 0$$

$$H_1 : \dots_k(v_t) \text{ อย่างน้อยหนึ่งค่าไม่เท่ากับศูนย์}$$

### ค่าสถิติทดสอบ

$$\text{Box-Pierce} \quad Q = n \sum_{k=1}^m r_k^2(e_t)$$

$$\text{Box-Ljung} \quad Q_m = n(n+2) \sum_{k=1}^m \frac{r_k^2(e_t)}{n-k}$$

โดยที่  $n$  คือ ขนาดของอนุกรมเวลาที่เป็นสเตชันนารี

$m$  คือ ช่วงเวลาห่างสูงสุดที่พิจารณาสหสัมพันธ์แบบอัตโนมัติ

$r_k(e_t)$  คือ ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองของตัวอย่างของค่าคลาดเคลื่อนที่อยู่ห่างกัน  $k$  ช่วงเวลา

จะปฏิเสธ  $H_0$  ที่ระดับนัยสำคัญ  $\alpha$  เมื่อตัวสถิติ  $Q$  หรือ  $Q_m$  มากกว่าหรือเท่ากับ  $t_{\alpha, m}^2$  แสดงว่าค่าคลาดเคลื่อนไม่เป็นอิสระต่อกัน นั่นคือ ตัวแบบที่ได้ยังไม่เหมาะสมต้องย้อนกลับไปกำหนดตัวแบบใหม่ในขั้นที่ 1 เพื่อหาตัวแบบอื่น ๆ ที่มีความเหมาะสมต่อไป

#### 2.4.4 การพยากรณ์ (forecasting)

จากรูปแบบที่กำหนดและผ่านการตรวจสอบในขั้นตอนที่ 2.4.3 แล้ว จะนำมาสร้างสมการพยากรณ์และหาค่าพยากรณ์ การพยากรณ์นี้จะเป็นทั้งการพยากรณ์แบบจุด (point forecast) และการพยากรณ์แบบช่วง (interval forecast)

### 3. วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล

วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล เป็นวิธีการสร้างสมการพยากรณ์จากค่าสังเกตในอนุกรมเวลาบางส่วนหรือทั้งหมด โดยน้ำหนักที่ให้กับค่าสังเกตแต่ละค่าแตกต่างกัน วิธีปรับให้เรียบเหมาะกับกรณีที่ลักษณะการเคลื่อนไหวของอนุกรมเวลามีแผนแบบที่เปลี่ยนแปลงอยู่ตลอดเวลา โดยทุกครั้งที่มีการสังเกตใหม่เข้ามาจะนำค่าสังเกตใหม่ดังกล่าวไปปรับสมการพยากรณ์ ซึ่งวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมีอยู่หลายวิธีการ โดยแต่ละวิธีการจะนำไปใช้กับลักษณะของข้อมูลอนุกรมเวลาที่แตกต่างกัน สรุปรายละเอียดดังนี้

### 3.1 วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (simple exponential smoothing)

เป็นวิธีการพยากรณ์สำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเคลื่อนไหวของแนวโน้มและความผันแปรของฤดูกาล โดยมีการสร้างสมการพยากรณ์จากค่าสังเกตทั้งหมดของอนุกรมเวลา หลักการปรับให้เรียบของวิธีนี้จะมีการให้น้ำหนักค่าสังเกตแต่ละค่าไม่เท่ากัน นั่นคือ จะให้น้ำหนักกับค่าสังเกตล่าสุด ( $Y_t$ ) มากที่สุดและลดหลั่นกันไปในลักษณะแบบเอ็กซ์โปเนนเชียลสำหรับค่าสังเกตที่อยู่ห่างออกไป ( $Y_{t-1}, Y_{t-2}, \dots$ ) ตามลำดับ น้ำหนักที่ให้นี้จะขึ้นอยู่กับค่าถ่วงน้ำหนัก ( $r$ ) ซึ่งมีผลรวมทั้งหมดเท่ากับ 1

สมการพยากรณ์หนึ่งช่วงเวลาล่วงหน้า ณ เวลา  $t$  เมื่อใช้ค่าสังเกตทั้งหมดจนถึงเวลา  $t$

$$\hat{Y}_{t+1} = rY_t + r(1-r)Y_{t-1} + r(1-r)^2Y_{t-2} + \dots \quad ; 0 < r < 1$$

วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว มีสมการพยากรณ์ดังนี้

$$\hat{Y}_{t+1} = rY_t + (1-r)\hat{Y}_t \quad ; 0 < r < 1$$

โดยที่  $\hat{Y}_{t+1}$  คือ ค่าพยากรณ์หนึ่งช่วงเวลาล่วงหน้า ณ เวลา  $t$

$Y_t$  คือ ค่าสังเกต ณ เวลา  $t$

$\hat{Y}_t$  คือ ค่าพยากรณ์ ณ เวลา  $t$

$r$  คือ ค่าคงที่ที่ทำให้เรียบสำหรับข้อมูล มีค่าอยู่ระหว่าง 0 ถึง 1

ถ้า  $r$  ที่มีค่ามาก ค่าน้ำหนักจะลดลงอย่างรวดเร็ว ส่วน  $r$  ที่มีค่าน้อย ค่าน้ำหนักจะลดลงอย่างช้าๆ นั่นคือ จะใช้ค่า  $r$  สูงเมื่อต้องการถ่วงน้ำหนักให้กับค่าล่าสุดมากกว่าค่าก่อนหน้า ดังแสดงในตารางที่ 3

ตารางที่ 3 แสดงน้ำหนักที่ให้กับค่าสังเกตด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว

เวลา	น้ำหนัก	r				
		0.1	0.3	0.5	0.6	0.9
1	r	0.1	0.3	0.5	0.6	0.9
2	$r(1-r)$	0.09	0.21	0.25	0.24	0.09
3	$r(1-r)^2$	0.081	0.147	0.125	0.096	0.009
4	$r(1-r)^3$	0.0729	0.1029	0.0625	0.0384	0.0009
5	$r(1-r)^4$	0.06561	0.07203	0.03125	0.01536	0.00009
6	$r(1-r)^5$	0.059049	0.050421	0.015625	0.006144	0.000009
7	$r(1-r)^6$	0.0531441	0.0352947	0.0078125	0.0024576	0.0000009
8	$r(1-r)^7$	0.0478297	0.0247063	0.00390625	0.0009830	0.00000009
9	$r(1-r)^8$	0.0430467	0.0172944	0.00195312	0.0003932	0.000000009

ที่มา: ทรงศิริ (2549)

### 3.2 วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงโดยวิธีของโฮลต์ (Holt's exponential smoothing method)

วิธีการพยากรณ์นี้เหมาะสำหรับข้อมูลอนุกรมเวลาที่มีการเคลื่อนไหวของแนวโน้มเส้นตรงและไม่มีควมผันแปรเนื่องจากฤดูกาล วิธีของโฮลต์ใช้ค่าปรับให้เรียบ 2 ค่า ได้แก่  $r$  เป็นค่าคงที่สำหรับปรับแนวโน้ม และ  $x$  เป็นค่าคงที่สำหรับปรับความลาดชัน ซึ่งมีค่าอยู่ระหว่าง 0 ถึง 1 ดังนั้นสมการพยากรณ์  $p$  ช่วงเวลาล่วงหน้าของโฮลต์ ณ เวลา  $t$  คือ

$$\hat{Y}_{t+p} = \hat{T}_t + p\hat{S}_t$$

โดยที่  $\hat{Y}_{t+p}$  คือ ค่าพยากรณ์  $p$  ช่วงเวลาล่วงหน้า ณ เวลา  $t$

$\hat{T}_t$  คือ ค่าปรับให้เรียบสำหรับแนวโน้ม

$\hat{S}_t$  คือ ค่าปรับให้เรียบสำหรับความลาดชัน

$p$  คือ ช่วงเวลาที่ต้องการพยากรณ์ไปข้างหน้า

$$\text{ซึ่ง } \hat{T}_t = r Y_t + (1-r)(\hat{T}_{t-1} + \hat{S}_{t-1})$$

$$\hat{S}_t = x(\hat{T}_t - \hat{T}_{t-1}) + (1-x)\hat{S}_{t-1}$$

โดยที่  $r$  คือ ค่าคงที่ปรับค่าให้เรียบสำหรับแนวโน้ม

$x$  คือ ค่าคงที่ปรับค่าให้เรียบสำหรับความลาดชัน

### 3.3 วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลของวินเทอร์ (Winter's exponential smoothing method)

วิธีการพยากรณ์นี้เหมาะสำหรับข้อมูลอนุกรมเวลาที่มีการเคลื่อนไหวของแนวโน้มเส้นตรงและมีความผันแปรของฤดูกาล วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลของวินเทอร์จะใช้ค่าปรับให้เรียบ 3 ค่า คือ  $r$ ,  $x$  และ  $u$  สำหรับค่าแนวโน้ม ค่าความลาดชันของแนวโน้มเชิงเส้น และค่าวัฏจักรของฤดูกาล ตามลำดับ ซึ่งมีค่าอยู่ระหว่าง 0 ถึง 1 วิธีการพยากรณ์นี้มี 2 รูปแบบ โดยจะแบ่งตามอิทธิพลของฤดูกาล คือ อนุกรมเวลาที่มีการแปรผันตามฤดูกาลแบบบวก และ อนุกรมเวลาที่มีการแปรผันตามฤดูกาลแบบคูณ ซึ่งมีรูปแบบดังนี้

#### 3.3.1 ตัวแบบอนุกรมเวลาที่มีการแปรผันตามฤดูกาลแบบบวก (additive seasonality)

มีลักษณะของการแปรผันตามฤดูกาลที่มีลักษณะความผันแปรคงที่เมื่ออนุกรมเวลามีแนวโน้มเพิ่มขึ้นหรือแตกต่างจากเดิมไม่มากนัก ณ เวลาต่างๆ กัน โดยมีสมการพยากรณ์  $p$  ช่วงเวลาล่วงหน้าของวินเทอร์ ณ เวลา  $t$  ดังนี้

$$\hat{Y}_{t+p} = \hat{T}_t + p\hat{S}_t + \hat{S}_{t+p-s}$$

$$\text{ซึ่ง } \hat{T}_t = r(Y_t - \hat{S}_{t-s}) + (1-r)(\hat{T}_{t-1} - \hat{S}_{t-1})$$

$$\hat{S}_t = x(\hat{T}_t - \hat{T}_{t-1}) + (1-x)\hat{S}_{t-1}$$

$$\hat{S}_t = u(Y_t - \hat{T}_t) + (1-u)\hat{S}_{t-s}$$

โดยที่  $\hat{Y}_{t+p}$  คือ ค่าพยากรณ์  $p$  ช่วงเวลาล่วงหน้า ณ เวลา  $t$   
 $\hat{T}_t$  คือ ค่าปรับให้เรียบสำหรับแนวโน้ม  
 $r$  คือ ค่าคงที่ปรับค่าให้เรียบสำหรับแนวโน้ม  
 $\hat{S}_t$  คือ ค่าปรับให้เรียบสำหรับค่าความลาดชัน  
 $x$  คือ ค่าคงที่ปรับค่าให้เรียบสำหรับค่าความลาดชัน  
 $\hat{S}_t$  คือ ค่าประมาณค่าชันฤดูกาล ณ เวลา  $t$   
 $u$  คือ ค่าคงที่ปรับค่าให้เรียบสำหรับอิทธิพลของฤดูกาล  
 $p$  คือ ช่วงเวลาที่ต้องการพยากรณ์ไปข้างหน้า  
 $s$  คือ ความยาวของฤดูกาล

### 3.3.2 ตัวแบบอนุกรมเวลาที่มีการแปรผันตามฤดูกาลแบบคูณ (multiplicative seasonality)

อนุกรมเวลาที่มีการผันแปรตามฤดูกาลเป็นสัดส่วนกับความลาดชันของอนุกรมเวลา เช่น เมื่อความลาดชันของอนุกรมเวลาเพิ่มขึ้น การแปรผันตามฤดูกาลก็จะเพิ่มขึ้นด้วย ซึ่งมีสมการพยากรณ์  $p$  ช่วงเวลาล่วงหน้าของวินเทอร์ ณ เวลา  $t$  ดังนี้

$$\hat{Y}_{t+p} = (\hat{T}_t + p\hat{S}_t)\hat{S}_{t+p-s}$$

ซึ่ง 
$$\hat{T}_t = r \left( \frac{Y_t}{\hat{S}_{t-s}} \right) + (1-r)(\hat{T}_{t-1} - \hat{S}_{t-1})$$

$$\hat{S}_t = x(\hat{T}_t - \hat{T}_{t-1}) + (1-x)\hat{S}_{t-1}$$

$$\hat{S}_t = u \left( \frac{Y_t}{\hat{T}_t} \right) + (1-u)\hat{S}_{t-s}$$

สัญลักษณ์ตัวแปรต่าง ๆ ใช้การอธิบายทำนองเดียวกับตัวแบบอนุกรมเวลาที่มีการแปรผันตามฤดูกาลแบบบวก

### เกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการพยากรณ์

ค่าวัดความคลาดเคลื่อนของการพยากรณ์ เป็นฟังก์ชันของค่าความคลาดเคลื่อนของการพยากรณ์โดยไม่คำนึงถึงทิศทาง นั่นคือ พิจารณาค่าสัมบูรณ์ของค่าความคลาดเคลื่อน ค่ากำลังสองของค่าความคลาดเคลื่อน หรือสัดส่วนของค่าความคลาดเคลื่อนกับค่าจริง ซึ่งฟังก์ชันของค่าความคลาดเคลื่อนที่ช่วยในการวัดความคลาดเคลื่อนของการพยากรณ์ มีดังต่อไปนี้

1. MAD หรือ ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย เป็นค่าวัดความคลาดเคลื่อนของการพยากรณ์ที่วัดจากขนาดของค่าความคลาดเคลื่อนของการพยากรณ์โดยไม่คำนึงถึงทิศทางของความคลาดเคลื่อนที่แทนด้วยค่าสัมบูรณ์ของค่าความคลาดเคลื่อนของการพยากรณ์ ซึ่งค่า MAD จะมีหน่วยเดียวกันกับค่าสังเกต

$$MAD = \frac{\sum_{i=1}^n |e_i|}{n}$$

2. MSE หรือ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เป็นค่าวัดความคลาดเคลื่อนของการพยากรณ์ที่วัดจากขนาดของค่าความคลาดเคลื่อนของการพยากรณ์ที่ได้จากกำลังสองของค่าความคลาดเคลื่อน ค่า MSE จะมีหน่วยวัดเป็นกำลังสองของหน่วยวัดของค่าสังเกต ซึ่งค่า MSE จะไวต่อค่าความคลาดเคลื่อนที่มีขนาดใหญ่ บางครั้งผู้พยากรณ์อาจจะใช้รากกำลังสองของค่า MSE ที่เรียกว่า RMSE (root mean square error) แทน ซึ่งมีหน่วยเดียวกันกับค่าสังเกต

$$MSE = \frac{\sum_{i=1}^n e_i^2}{n}$$

3. MAPE หรือ เปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย เป็นค่าวัดความคลาดเคลื่อนของการพยากรณ์ที่วัดจากขนาดของความคลาดเคลื่อนของการพยากรณ์เทียบกับค่าจริง ค่าวัดความคลาดเคลื่อนนี้เป็นค่าที่ไม่มีหน่วย จึงเหมาะที่จะใช้ในการเปรียบเทียบอนุกรมเวลาหลายชุดที่อาจมีหน่วยของการวัดต่างกันเมื่อใช้วิธีการพยากรณ์เดียวกัน

$$\text{MAPE} = \frac{\sum_{i=1}^n \frac{|e_i|}{Y_i}}{n}$$

ในการเปรียบเทียบวิธีการพยากรณ์ที่เหมาะสมที่สุด จะต้องเป็นวิธีที่ให้ค่า MAD MSE และ MAPE น้อยที่สุด

### การตรวจสอบค่าพยากรณ์

สมการพยากรณ์แต่ละวิธีจะเหมาะกับระยะของการพยากรณ์ที่ต่างกัน ได้แก่ ระยะสั้น ระยะกลาง และระยะยาว สมการพยากรณ์ที่สร้างขึ้นและผ่านการตรวจสอบแล้วว่าดีที่สุดจะนำไปหาค่าพยากรณ์ในอนาคต อย่างไรก็ตามแม้ว่าจะพยากรณ์จากสมการพยากรณ์ที่ดีที่สุดแล้ว แต่เมื่อพยากรณ์ไปช่วงเวลาหนึ่งค่าพยากรณ์นั้นอาจจะผิดพลาด นั่นคือ ค่าพยากรณ์ต่างจากค่าจริงที่เก็บรวบรวมมาใหม่มาก สาเหตุสำคัญ ได้แก่ เหตุการณ์ที่เกิดในอดีตมีการเปลี่ยนแปลง หรือปัจจัยต่าง ๆ ที่มีผลต่อการเคลื่อนไหวของอนุกรมเวลาที่เปลี่ยนไป นอกจากนั้นการพยากรณ์ที่ผิดพลาดมักจะทำให้เกิดความเสียหายต่อการปฏิบัติที่อาจจะประเมินค่าไม่ได้ ซึ่งเครื่องมือที่ช่วยตรวจสอบค่าพยากรณ์จะนำค่าพยากรณ์จากสมการพยากรณ์ที่สร้างขึ้นไปเปรียบเทียบกับค่าจริงที่เก็บรวบรวมมาใหม่ ซึ่งค่าจริงดังกล่าวยังไม่ได้นำมาใช้ในการสร้างสมการพยากรณ์ ถ้าค่าจริงและค่าพยากรณ์ไม่ต่างกันมากแสดงว่าค่าพยากรณ์อยู่ในการควบคุม แต่ถ้าค่าพยากรณ์ห่างจากค่าจริงมากแสดงว่าค่าพยากรณ์ไม่อยู่ในการควบคุม ค่าพยากรณ์ที่ไม่อยู่ในการควบคุมจะเป็นสัญญาณเตือนแก่ผู้พยากรณ์ว่าค่าพยากรณ์เริ่มไม่ถูกต้องหรือสมการพยากรณ์ที่สร้างขึ้นเริ่มไม่เหมาะสม การสร้างสมการพยากรณ์ใหม่จะทำด้วยวิธีใดวิธีหนึ่งโดยนำค่าสังเกตใหม่มาร่วมในการสร้างสมการพยากรณ์ใหม่ ดังนี้

1. กำหนดรูปแบบอนุกรมเวลาเดิมและใช้วิธีการพยากรณ์เดิมแต่นำค่าสังเกตที่เก็บรวบรวมมาใหม่ไปปรับสมการพยากรณ์เดิม ค่าประมาณของพารามิเตอร์ที่ได้จะแตกต่างจากค่าประมาณเดิม
2. กำหนดรูปแบบอนุกรมเวลาใหม่ตามลักษณะการเคลื่อนไหวของอนุกรมเวลาใหม่ที่พบแต่ใช้วิธีการพยากรณ์เดิม
3. กำหนดรูปแบบอนุกรมเวลาเดิมแต่ใช้วิธีการพยากรณ์ที่เหมาะสมกว่า

#### 4. กำหนดรูปแบบอนุกรมเวลาใหม่ตามลักษณะการเคลื่อนไหวของอนุกรมเวลาใหม่ที่พบ และใช้วิธีการพยากรณ์ใหม่

เพื่อให้ขั้นตอนของการพยากรณ์มีความสมบูรณ์จึงควรตรวจสอบการพยากรณ์ค่าในอนาคตทุกระยะ การตรวจสอบความถูกต้องของค่าพยากรณ์จะใช้ค่าวัดความคลาดเคลื่อนของการพยากรณ์เป็นเกณฑ์ในการพิจารณา เช่น ค่าความคลาดเคลื่อน (individual error :  $e_t$ ) ผลรวมของค่าความคลาดเคลื่อน (simple cusum :  $\sum e_t$ ) ผลรวมเคลื่อนที่ของค่าความคลาดเคลื่อน (moving total error :  $MTE_t$ ) และค่าปรับให้เรียบของค่าความคลาดเคลื่อน (smoothed error :  $E_t$ ) ค่าสถิติเหล่านี้เป็นตัวประมาทที่มีความผันแปรเข้ามาเกี่ยวข้อง ในการตรวจสอบจะนำค่าวัดความผันแปรของค่าสถิติที่เรียกว่าค่าคลาดเคลื่อนมาตรฐานมาปรับค่าวัดความคลาดเคลื่อนของการพยากรณ์เพื่อให้ค่าวัดความคลาดเคลื่อนของการพยากรณ์เป็นค่าวัดความคลาดเคลื่อนมาตรฐาน จะเรียกค่าวัดความคลาดเคลื่อนมาตรฐานว่า tracking signal ซึ่งรายละเอียดของค่าวัดความคลาดเคลื่อนมีดังนี้

1. ค่าความคลาดเคลื่อนที่เวลา  $t$  หรือ  $e_t$  เป็นผลต่างของค่าจริงและค่าพยากรณ์ที่เวลา  $t$  ค่าคลาดเคลื่อนมาตรฐานของ  $e_t$  หรือ  $t_{e_t}$  มีค่าคงที่เท่ากับ  $t_e$  สำหรับทุกค่าของ  $t$  นั่นคือ  $t_{e_t} = t_e$
2. ผลรวมของค่าความคลาดเคลื่อนที่เวลา  $t$  หรือ  $\sum e_t$  เป็นผลรวมของค่าความคลาดเคลื่อนของการพยากรณ์นับตั้งแต่เริ่มการพยากรณ์จนถึงเวลา  $t$  โดย

$$\sum e_t = e_1 + e_2 + \dots + e_t = e_t + \sum e_{t-1}$$

ค่าคลาดเคลื่อนมาตรฐานของ  $\sum e_t$  หรือ  $t_{\sum e_t}$  เป็นฟังก์ชันของเวลา  $t$  นั่นคือ  $t_{\sum e_t} = t_e \sqrt{t}$

3. ผลรวมเคลื่อนที่ของค่าความคลาดเคลื่อน  $k$  ค่า ที่เวลา  $t$  หรือ  $MTE_t$  เป็นผลรวมของค่าความคลาดเคลื่อนของการพยากรณ์ครั้งละ  $k$  ค่า ที่เวลา  $t$  โดยเริ่มตั้งแต่เวลา  $t-k+1$  จนถึง  $t$  โดย

$$MTE_t = e_t + e_{t-1} + \dots + e_{t-k+1} = (e_t - e_{t-k}) + MTE_{t-1}$$

ค่าคลาดเคลื่อนมาตรฐานของ  $MTE_t$  หรือ  $t_{MTE_t}$  เท่ากับ  $t_{MTE}$  สำหรับทุกค่าของ  $t$  นั่นคือ

$$t_{MTE_t} = t_{MTE} = t_e \sqrt{k}$$

4. ค่าปรับให้เรียบของค่าความคลาดเคลื่อนที่เวลา  $t$  หรือ  $E_t$  เป็นค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนของการพยากรณ์นับตั้งแต่เริ่มการพยากรณ์จนถึงเวลา  $t$  โดยค่าถ่วงน้ำหนักที่ให้กับค่าความคลาดเคลื่อนที่เวลา  $t, t-1$  จะลดลงแบบเอ็กซ์โปเนนเชียล นั่นคือ ให้ค่าถ่วงน้ำหนัก  $y, y(1-y), y(1-y)^2, \dots$  กับค่าความคลาดเคลื่อน  $e_t, e_{t-1}, e_{t-2}, \dots$  ตามลำดับ

$$E_t = ye_t + y(1-y)e_{t-1} + y(1-y)^2e_{t-2} + \dots = ye_t + (1-y)E_{t-1}$$

ค่าปรับให้เรียบ  $y$  มีค่าอยู่ระหว่าง 0 ถึง 1 และมีผลรวมของค่าถ่วงน้ำหนักเท่ากับ 1

ค่าคลาดเคลื่อนมาตรฐานของ  $E_t$  หรือ  $t_{E_t}$  ขึ้นกับเวลา  $t$  เขียน  $t_{E_t}$  ในเทอมของ  $t_{e_t}$  ได้เป็น  $t_{e_t} \sqrt{\frac{y}{2-y}}$  ซึ่ง  $t_{e_t} = 1.25MAD_t$  มี  $MAD_t = y|e_t| + (1-y)MAD_{t-1}$  นั่นคือ  $t_{E_t} = 1.25MAD_t \sqrt{\frac{y}{2-y}}$  ในการคำนวณค่า  $E_t$  ต้องทราบค่า  $E_{t-1}$  ในทางปฏิบัติต้องกำหนดค่าเริ่มต้นให้กับ  $E_t$  ที่เรียก  $E_{init}$  ซึ่งเป็นทำนองเดียวกันกับการกำหนดค่าเริ่มต้นให้กับค่าต่าง ๆ ด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล

ตารางที่ 4 ค่าวัดความคลาดเคลื่อนและค่าคลาดเคลื่อนมาตรฐาน

ค่าวัดความคลาดเคลื่อนเดิม		ค่าวัดความคลาดเคลื่อนใหม่	
ค่าวัด	ค่าคลาดเคลื่อนมาตรฐาน	ค่าวัด	ค่าคลาดเคลื่อนมาตรฐาน
$e_t$	$t_{e_t} = t_e$	$e_t$	$t_e$
$\sum E_t$	$t_{\sum e_t} = t_e \sqrt{t}$	$\frac{\sum E_t}{\sqrt{t}}$	$t_e$
$MTE_t$	$t_{MTE_t} = t_e \sqrt{k}$	$MTE_t$	$t_e \sqrt{k}$
$E_t$	$t_{E_t} = 1.25MAD_t \sqrt{\frac{y}{2-y}}$	$\frac{E_t}{MAD_t}$	$1.25 \sqrt{\frac{y}{2-y}}$

ค่าวัดความคลาดเคลื่อนที่ใช้ในการตรวจสอบค่าพยากรณ์และค่าคลาดเคลื่อนมาตรฐานที่สัมพันธ์กันแสดงในตารางที่ 4 กรณีที่ค่าคลาดเคลื่อนมาตรฐานขึ้นกับเวลา  $t$  มีการปรับค่าวัดความคลาดเคลื่อนของการพยากรณ์เพื่อความสะดวกในการตรวจสอบค่าพยากรณ์หรือปรับค่าวัดความคลาดเคลื่อนให้มีค่าคลาดเคลื่อนมาตรฐานที่ไม่เป็นฟังก์ชันของ  $t$  นั่นคือ จากค่าวัดความ

คลาดเคลื่อน  $\sum E_t$  ที่มีค่าคลาดเคลื่อนมาตรฐาน  $t_{\sum e_t}$  ซึ่ง  $t_{\sum e_t} = t_e \sqrt{t}$  กำหนดค่าวัดความคลาดเคลื่อนใหม่เป็น  $\frac{\sum E_t}{\sqrt{t}}$  ที่มีค่าคลาดเคลื่อนมาตรฐาน  $t_e$  และจากค่าวัดความคลาดเคลื่อน  $E_t$  มีค่าคลาดเคลื่อนมาตรฐาน  $t_{E_t}$  ซึ่ง  $t_{E_t} = 1.25MAD_t \sqrt{\frac{y}{2-y}}$  กำหนดค่าวัดความคลาดเคลื่อนใหม่เป็น  $\frac{E_t}{MAD_t}$  มีค่าคลาดเคลื่อนมาตรฐาน  $1.25 \sqrt{\frac{y}{2-y}}$  สำหรับค่าวัดความคลาดเคลื่อนที่มีค่าคลาดเคลื่อนมาตรฐานไม่เป็นฟังก์ชันของ  $t$  ได้แก่  $e_t, \frac{\sum E_t}{\sqrt{t}}, MTE_t$  และ  $\frac{E_t}{MAD_t}$  ต่อไปจะแทนด้วย  $TS_1, TS_2, TS_3$  และ  $TS_4$  ซึ่ง  $TS_1, TS_2, TS_3$  และ  $TS_4$  มีขีดจำกัดของขอบเขตการยอมรับที่ไม่ขึ้นกับเวลา  $t$  ดังตารางที่ 5

ตารางที่ 5 tracking signal และช่วงควบคุมเมื่อค่าความคลาดเคลื่อนเป็นอิสระกัน

Tracking signal (TS)	ช่วงควบคุม
$TS_{1t} = e_t$	$\pm 2t_{TS_1}$ หรือ $2t_e$
$TS_{2t} = \frac{\sum E_t}{\sqrt{t}}$	$\pm 2t_{TS_2}$ หรือ $2t_e$
$TS_{3t} = MTE_t$	$\pm 2t_{TS_3}$ หรือ $2t_e \sqrt{k}$
$TS_{4t} = \frac{E_t}{MAD_t}$	$\pm 2t_{TS_4}$ หรือ $2(1.25) \sqrt{\frac{y}{2-y}}$

1. จากอนุกรมเวลาในอดีตขนาด  $n$  ที่มีค่าความคลาดเคลื่อนของการพยากรณ์  $e_1, e_2, \dots, e_n$  ประมาณค่าคลาดเคลื่อนมาตรฐาน  $t_e$  จากรากที่สองของค่าเฉลี่ยของค่าความคลาดเคลื่อนของการ

พยากรณ์กำลังสอง นั่นคือ  $s = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}}$

2. จากค่าความคลาดเคลื่อนของการพยากรณ์ที่ได้จากการพยากรณ์แล้ว  $t$  ช่วงเวลาหรือจาก  $e_1, e_2, \dots, e_t$  ประมาณค่าคลาดเคลื่อนมาตรฐาน  $t_e$  ด้วยค่าต่าง ๆ ได้แก่

ก. รากที่สองของค่าเฉลี่ยของค่าความคลาดเคลื่อนของการพยากรณ์กำลังสองที่เวลา  $t$

$$\text{นั่นคือ } s_{e_t} = \sqrt{\frac{\sum_{i=1}^t e_i^2}{t}}$$

ข. รากที่สองของค่าเฉลี่ยถ่วงน้ำหนักของค่าความคลาดเคลื่อนกำลังสองที่เวลา  $t$  โดยค่าถ่วงน้ำหนักแบบเอ็กซ์โปเนนเชียล  $y$  มีค่าอยู่ระหว่าง 0 ถึง 1 นั่นคือ  $s_{e_t} = \sqrt{\text{MSE}_t}$  ซึ่ง  $\text{MSE}_t = ye_t^2 + (1-y)\text{MSE}_{t-1}$

ค. ขนาด 1.25 เท่าของค่าเฉลี่ยถ่วงน้ำหนักของค่าสัมบูรณ์ของค่าความคลาดเคลื่อนของการพยากรณ์ โดยค่าถ่วงน้ำหนักแบบเอ็กซ์โปเนนเชียล  $y$  มีค่าอยู่ระหว่าง 0 ถึง 1 นั่นคือ  $s_{e_t} = 1.25\text{MAD}_t$  ซึ่ง  $\text{MSE}_t = y|e_t| + (1-y)\text{MAD}_{t-1}$

### ผลงานวิจัยที่เกี่ยวข้อง

ออมพิไล (2548) ศึกษาการทำนายเครือข่ายการควบคุมยีนจากข้อมูลไมโครอาร์เรย์ของดีเอ็นเอด้วยเทคนิคเครือข่ายเบย์เซียน โดยการพิจารณาจากกลุ่มยีน HAP2, HAP3, HAP4 และ CYC1 ซึ่งเกี่ยวข้องกับกระบวนการไดออกซิซิฟท์ของยีสต์ *Saccharomyces Cerevisiae* มีจุดประสงค์ในการทำวิจัยเพื่อทำนายรูปแบบการควบคุมกันระหว่างยีน จากข้อมูลการแสดงออกของยีนเหล่านี้ภายใต้สถานะไดออกซิซิฟท์ที่มีการเปลี่ยนแปลงการใช้แหล่งพลังงานจากกลูโคสเป็นเอทานอล โดยใช้วิธีเบย์เซียนเพื่อประมาณค่าความน่าจะเป็นอย่างมีเงื่อนไข และคะแนนสำหรับทุกเครือข่ายการควบคุมกันที่เป็นไปได้ จากการทดลองผู้วิจัยพบว่ามีเครือข่ายทั้งหมด 8 เครือข่ายที่ให้คะแนนสูงสุดเท่ากัน โดยเครือข่ายที่พบทั้งหมดมีลักษณะบางประการคล้ายกัน มีลักษณะร่วมกัน คือ ทุกเครือข่ายมีความสัมพันธ์จากยีน HAP2 และ HAP3 ไปยังยีน CYC1 ซึ่งสอดคล้องกับรายงานทางชีววิทยาที่กล่าวว่ายีน HAP2, HAP3 และ HAP4 มีลักษณะการทำงานที่ร่วมกันไปควบคุมยีน CYC1 โดยที่ยีน HAP2 และ HAP3 เป็นโปรตีนควบคุมที่เข้าไปจับกับส่วนโปรโมเตอร์ที่เฉพาะเจาะจงของยีน CYC1 และยีน HAP4 นั้นมีรายงานว่าเป็นตัวกระตุ้นเพื่อให้ CYC1 เริ่มทำงานได้

วราฤทธิ์ (2550) ศึกษาการพยากรณ์ดัชนีราคาผู้บริโภครายเดือนของกรุงเทพมหานครตั้งแต่เดือนมกราคม 2541 ถึงเดือนธันวาคม 2549 โดยใช้ปี 2545 เป็นปีฐาน แบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดที่หนึ่งใช้สำหรับการวิเคราะห์เพื่อสร้างตัวแบบพยากรณ์ เป็นข้อมูลดัชนีราคาผู้บริโภครายเดือนตั้งแต่เดือนมกราคม 2541 ถึงเดือนธันวาคม 2548 ระยะเวลา 96 เดือน และชุดที่สองเป็นดัชนีราคา

ผู้บริโภครายเดือนตั้งแต่เดือนมกราคม 2549 ถึงเดือนธันวาคม 2549 ระยะเวลา 12 เดือน ใช้เป็นข้อมูลทดสอบ สำหรับการตรวจสอบตัวแบบพยากรณ์ที่สร้างขึ้นด้วยข้อมูลชุดที่หนึ่ง ทำการเปรียบเทียบวิธีการพยากรณ์ 3 วิธี คือ วิธีบ็อกซ์-เจนกินส์ วิธีการของโฮลด์ และวิธีการพยากรณ์รวมโดยการใช้การวิเคราะห์การถดถอย ซึ่งจะเปรียบเทียบประสิทธิภาพของแต่ละวิธีด้วยค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ต่ำที่สุด จากการศึกษาพบว่า วิธีการของโฮลด์เป็นวิธีการที่ให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ต่ำที่สุด จึงเป็นวิธีที่เหมาะสมที่สุดสำหรับข้อมูลดัชนีราคาผู้บริโภครายเดือนของกรุงเทพมหานคร

สมพล (2551) ศึกษาการเรียนรู้โครงสร้างเครือข่ายเบย์เซียนด้วย Genetic Algorithm (GA) โดยใช้โครโมโซมรูปแบบ Directed Acyclic Graph (DAG) เพื่อทำการเรียนรู้โครงสร้างเครือข่ายเบย์เซียนที่ดีที่สุดรวมทั้งนำเสนอวิธีการ Crossover และ Mutation ของประชากร เพื่อให้มีคุณสมบัติของ DAG โดยไม่ต้องใช้วิธีการแก้ไข เพื่อปรับปรุงประชากรให้กลับเป็น DAG ดั้งเดิม ซึ่งในการวิจัยผู้วิจัยได้นำเสนอวิธีการประยุกต์ใช้ Frequency Count Tree เพื่อใช้ในการหาค่าความเหมาะสมของประชากรแทนการติดต่อกับฐานข้อมูลโดยตรง ซึ่งทำให้ GA มีประสิทธิภาพที่สูงขึ้น จากผลการศึกษาพบว่า อัลกอริทึมที่นำเสนอให้ประสิทธิภาพดีกว่าการใช้รูปโครโมโซมในรูปแบบกราฟทั่วไป เนื่องจากโครโมโซมในรูปแบบ DAG ไม่ต้องทำการแก้ไขประชากรที่เป็นกราฟทั่วไปให้มีคุณสมบัติของ DAG

ประการ (2552) ศึกษาการวิเคราะห์ความสัมพันธ์ระหว่างราคาน้ำมันในตลาดโลก อัตราดอกเบี้ยเงินกู้ ดัชนีราคาวัสดุก่อสร้าง ดัชนีราคาที่อยู่อาศัย และสินเชื่อสหกรณ์การเกษตรที่มีผลต่อผลประกอบการของบริษัทพัฒนาอสังหาริมทรัพย์ที่จดทะเบียนในตลาดหลักทรัพย์แห่งประเทศไทย โดยผลประกอบการที่ผู้วิจัยศึกษาในครั้งนี้ ได้แก่ รายได้ จัดเป็นตัวแปร 4 แบบ คือ แบบที่ 1 รายได้รวมของบริษัทที่ประกอบธุรกิจขายบ้านและคอนโด แบบที่ 2 รายได้รวมของบริษัทที่ประกอบธุรกิจขายบ้าน แบบที่ 3 คือ รายได้รวมของบริษัทที่ประกอบธุรกิจขายคอนโด และแบบที่ 4 คือ รายได้รวมของทั้ง 6 บริษัทที่ประกอบธุรกิจพัฒนาอสังหาริมทรัพย์ที่ทำการศึกษา โดยใช้ข้อมูลในช่วงระยะเวลาเป็นรายไตรมาส ตั้งแต่ไตรมาสที่ 1 ปี 2544 จนถึงไตรมาสที่ 2 ปี 2552 รวมระยะเวลา 34 ไตรมาส ซึ่งในการศึกษาครั้งนี้ผู้วิจัยใช้การวิเคราะห์เชิงพรรณนาโดยอาศัยค่าสถิติในรูปร้อยละ ใช้การวิเคราะห์ความสัมพันธ์ เทคนิคการวิเคราะห์ปัจจัย และเทคนิคการสร้างสมการความถดถอย ผลจากการศึกษาพบว่า ราคาน้ำมันในตลาดโลก ดัชนีราคาวัสดุก่อสร้าง ดัชนีราคาที่อยู่อาศัย และสินเชื่อสหกรณ์การเกษตร มีความสัมพันธ์กันและมีความสัมพันธ์กับตัวแปรตามด้านรายได้ของธุรกิจอสังหาริมทรัพย์ทั้ง 4 แบบ ส่วนอัตราดอกเบี้ยเงินกู้ไม่มีความสัมพันธ์กับตัวแปร

ใดๆ จึงนำตัวแปรต้นทั้ง 4 ตัวแปรมาวิเคราะห์ปัจจัยทำให้ได้ตัวแปรใหม่ที่มีชื่อว่า ปัจจัยด้านรายได้ของธุรกิจพัฒนาอสังหาริมทรัพย์ ซึ่งดัชนีราคาวัสดุก่อสร้างมีน้ำหนักต่อตัวแปรปัจจัยธุรกิจพัฒนาอสังหาริมทรัพย์มากที่สุด จึงนำไปสร้างสมการพยากรณ์แนวโน้ม จากการทดสอบและวิเคราะห์แนวโน้มพบว่า สภาวะของธุรกิจอสังหาริมทรัพย์มีแนวโน้มด้านรายได้ที่เพิ่มมากขึ้นและมีทิศทางไปในทางเดียวกันทั้งหมด เนื่องจากสภาวะธุรกิจอยู่ในช่วงระยะการฟื้นตัว

ธัญพร (2553) ได้ศึกษาการพยากรณ์อัตราเงินเฟ้อของประเทศไทย เพื่อหาตัวแบบที่มีความเหมาะสม โดยทำการเปรียบเทียบวิธีการพยากรณ์ทั้ง 5 วิธี ได้แก่ วิธีการพยากรณ์ของโฮลด์ วิธีบ็อกซ์-เจนกินส์ วิธีทรานสเฟอร์ฟังก์ชัน ตัวแบบเวกเตอร์ออโตรีเกรสซีฟ (VAR) และการพยากรณ์รวมด้วยวิธีการวิเคราะห์การถดถอย ซึ่งข้อมูลอนุกรมเวลาที่ใช้ศึกษาประกอบด้วย ดัชนีราคาผู้บริโภค อัตราแลกเปลี่ยนเงินตราต่างประเทศ มูลค่าการนำเข้าสินค้า ระดับราคาน้ำมันเบนซิน อัตราการว่างงานและผลิตภัณฑ์มวลรวมภายในประเทศ ตั้งแต่เดือนมกราคม 2541 ถึงเดือนธันวาคม 2550 รวมระยะเวลา 120 เดือน โดยเกณฑ์ที่ใช้เปรียบเทียบจะพิจารณาจากค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย (MAD) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่า Akaike Information Criterion (AIC) ผลจากการศึกษา พบว่า ตัวแบบทรานสเฟอร์ฟังก์ชันที่มีรูปแบบ  $\hat{Z}_t = 0.0014 + 0.0545Z_t^x + (1 - 0.3048B)^{-1}e_t$  ให้ค่า MAPE MAD MSE และ AIC น้อยที่สุด จึงมีความเหมาะสมสำหรับการพยากรณ์ดัชนีราคาผู้บริโภคของประเทศไทยมากที่สุด แต่วิธีนี้เหมาะสำหรับพยากรณ์ในช่วงระยะเวลาสั้นๆ 1-3 เดือน หากต้องการพยากรณ์หลายช่วงเวลามากขึ้น เช่น 12 เดือน อาจพิจารณาวิธีของโฮลด์ที่ให้ค่าความคลาดเคลื่อนของค่าพยากรณ์ที่น้อยกว่าวิธีอื่นๆ

โรชนานี และ ปริดาภรณ์ (2553) ศึกษาการพยากรณ์มูลค่าการส่งออกปลาหูน้ำกระป๋องของประเทศไทยโดยการเปรียบเทียบการวิเคราะห์ข้อมูลอนุกรมเวลาโดยวิธีบ็อกซ์-เจนกินส์และวิธีการถดถอยที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียลกำลังสองในช่วงระยะเวลา 6 ปี ตั้งแต่เดือนมกราคม 2545 ถึงเดือนมีนาคม 2552 จากผลการศึกษาพบว่า ตัวแบบที่เหมาะสมสำหรับการพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์ได้รูปแบบการพยากรณ์ คือ ARIMA(0,1,1) โดยที่มีสมการพยากรณ์ดังนี้  $\hat{Y}_t = Y_{t-1} + 45.126241 - 0.427354e_{t-1}$  ส่วนการพยากรณ์ด้วยวิธีการถดถอยที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียลกำลังสอง มีสมการพยากรณ์ดังนี้  $\hat{Y}_t = (2434.486)(1.098677)^t (1.00002)^{t^2}$  ซึ่งเมื่อได้รูปแบบสมการพยากรณ์จากทั้ง 2 วิธีแล้ว จะเปรียบเทียบวิธีทั้ง 2 วิธี โดยการพิจารณาจากค่าความคลาดเคลื่อนกำลังสองเฉลี่ยหรือ MSE พบว่า การพยากรณ์โดยวิธีการถดถอยที่มีแนวโน้ม

แบบเอ็กซ์โปเนนเชียลกำลังสองเป็นวิธีการที่มีความเหมาะสมสำหรับข้อมูลมากกว่าวิธีบ็อกซ์-เจนกินส์ เนื่องจากให้ค่า MSE ต่ำกว่า

Gnophanxay (2011) ศึกษาการผสมผสานกันระหว่างเครือข่ายความเชื่อเบย์เซียนและขั้นตอนวิธีเชิงพันธุกรรม สำหรับพยากรณ์ความสำเร็จในการเรียนของนักศึกษาคณะวิทยาศาสตร์ของมหาวิทยาลัยแห่งชาติลาว โดยศึกษาจากผลสัมฤทธิ์ทางการเรียนของนักศึกษาจากแผนการศึกษาในช่วงปี 2547 ถึงปี 2551 จากผลการศึกษาพบว่า จำนวนร้อยละ 54 ของนักศึกษาที่ไม่สัมฤทธิ์ผลทางการเรียนเนื่องจากผลกระทบทางการเรียนวิชาพื้นฐานส่งผลไปสู่การศึกษาที่ต่อเนื่องกัน ซึ่งปัญหานี้จำเป็นต้องได้รับการศึกษาเพื่อหาปัจจัยที่มีอิทธิพลต่อกัน โดยการหาความสัมพันธ์ระหว่างวิชาพื้นฐานและวิชาต่อเนื่อง ผู้วิจัยจึงใช้วิธีการผสมผสานกันระหว่างเครือข่ายความเชื่อเบย์เซียนและขั้นตอนวิธีเชิงพันธุกรรม โดยที่เครือข่ายความเชื่อเบย์เซียนมีรูปแบบลักษณะเป็นกราฟแบบมีทิศทางและไม่วนเป็นวงกลม ซึ่งจะแสดงถึงชุดของตัวแปรสุ่มและความน่าจะเป็นอย่างมีเงื่อนไขของกราฟที่ใช้ในการตัดสินใจ ส่วนขั้นตอนวิธีเชิงพันธุกรรมเป็นวิธีการหาค่าที่เหมาะสม สามารถใช้ค้นหาโครงสร้างเครือข่ายความเชื่อเบย์เซียนที่เหมาะสมเพื่อใช้สำหรับพยากรณ์ ผู้วิจัยได้ดำเนินการพัฒนาปรับปรุงวิธีการกลายพันธุ์ของขั้นตอนวิธีเชิงพันธุกรรมเพื่อให้มีประสิทธิภาพดีขึ้น โดยวัดความถูกต้องในการพยากรณ์ด้วยเครือข่ายความเชื่อเบย์เซียน ผลจากการศึกษาพบว่า เมื่อใช้วิธีการกลายพันธุ์ 2 จุดจะให้ประสิทธิภาพที่ดีกว่าเดิม

Zuo and Kita (2012) กล่าวว่า การพยากรณ์อนุกรมเวลาแสดงให้เห็นถึงการแจกแจงของค่าความคลาดเคลื่อนว่าเป็นการแจกแจงปกติ จากการศึกษาที่ผ่านมาพบว่า การแจกแจงของความผันแปรราคาหุ้นไม่ได้มีการแจกแจงปกติ โดยเฉพาะอย่างยิ่งการวิเคราะห์ข้อมูลหุ้นตามจริงแสดงถึงความผันแปรรอบค่า  $\pm t$  และ  $\pm 3t$  ซึ่งมีค่ากว้างกว่าการแจกแจงปกติ บางครั้งวิธีการแจกแจงปกติอาจจะพยากรณ์ราคาหุ้นได้ไม่แม่นยำนัก ดังนั้น การพยากรณ์ราคาหุ้นนี้จะใช้วิธีเครือข่ายเบย์เซียนในการพยากรณ์แทน ซึ่งเครือข่ายเบย์เซียนนี้เป็นกราฟความน่าจะเป็นที่แสดงถึงเซตของตัวแปรสุ่มและเงื่อนไขโดยอาศัยกราฟที่ไม่มีการวนเป็นวงกลม เครือข่ายเบย์เซียนสามารถพยากรณ์ราคาหุ้นโดยไม่ใช้รูปแบบ White Noise แต่เนื่องจากราคาหุ้นเป็นค่าที่ต่อเนื่องจึงจำเป็นที่จะต้องแปลงให้เป็นค่าที่ไม่ต่อเนื่อง จากนั้นจะใช้เครือข่ายเบย์เซียนในการพยากรณ์โดยอาศัยค่าตัวเลขจากราคาหุ้นก่อนหน้า โดยศึกษาการพยากรณ์ราคาหุ้นของบริษัท NIKKEI เปรียบเทียบกับบริษัท TOYOTA MOTOR โดยใช้วิธีเครือข่ายเบย์เซียนเปรียบเทียบกับวิธีการวิเคราะห์อนุกรมเวลา ซึ่งวิธีเครือข่ายเบย์เซียนเป็นกราฟความน่าจะเป็นที่แสดงถึงเซตของตัวแปรสุ่มและการขึ้นต่อกันอย่างมีเงื่อนไขโดยอาศัยกราฟที่ไม่มีการวนเป็นวงกลม โดยในขั้นตอนแรกนั้นผู้วิจัยทำการแปลงข้อมูลอัตราส่วน

ราคาต่อกำไรให้เป็นเซตของข้อมูลที่เป็นจำนวนนับ จากนั้นจึงใช้วิธีเครือข่ายเบย์เซียนเพื่อพยากรณ์ราคาหุ้นของทั้งสองบริษัท จากผลการศึกษาพบว่า ทั้งบริษัท NIKKEI และบริษัท TOYOTA MOTOR ที่วิเคราะห์โดยวิธีเครือข่ายเบย์เซียนจะให้ค่า RMSE ต่ำกว่าวิธีการวิเคราะห์อนุกรมเวลานั้นคือ บริษัท NIKKEI มีค่า RMSE เท่ากับ 2.3065 บริษัท TOYOTA MOTOR มีค่า RMSE เท่ากับ 3.2669 และถ้าวิเคราะห์ด้วยวิธีอนุกรมเวลาจะให้ค่า RMSE ของบริษัท NIKKEI เท่ากับ 2.7331 และค่า RMSE ของบริษัท TOYOTA MOTOR เท่ากับ 4.0669 และค่าสัมประสิทธิ์สหสัมพันธ์โดยวิธีเครือข่ายเบย์เซียนของบริษัท NIKKEI และบริษัท TOYOTA MOTOR คือ 0.8028 และ 0.6284 ตามลำดับ ซึ่งมีค่าสูงกว่าวิธีการวิเคราะห์อนุกรมเวลาซึ่งให้ค่า CC ของบริษัท NIKKEI เท่ากับ 0.6974 และค่า CC ของบริษัท TOYOTA MOTOR เท่ากับ 0.4692

## อุปกรณ์และวิธีการ

### อุปกรณ์

1. เครื่องไมโครคอมพิวเตอร์ของภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
2. โปรแกรม Microsoft Excel Version 2007 และโปรแกรมทางสถิติ Matlab R2010b, SAS Version 9.3 และ WEKA Version 3.6
3. ข้อมูลดัชนีเซท 50 ดัชนีราคาวัสดุก่อสร้าง ดัชนีราคาที่อยู่อาศัย ดัชนีความเชื่อมั่นทางธุรกิจ อัตราดอกเบี้ย และดัชนีราคาที่ดิน ซึ่งข้อมูลเหล่านี้เป็นข้อมูลรายเดือนตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

### วิธีการ

#### การเก็บข้อมูลและการเตรียมข้อมูล

ข้อมูลที่ใช้ในการศึกษาครั้งนี้เป็นข้อมูลทุติยภูมิ (secondary data) ประกอบด้วยดัชนีราคา หุ่นกลุ่มอสังหาริมทรัพย์และก่อสร้าง ข้อมูลดัชนีเซท 50 ข้อมูลดัชนีราคาวัสดุก่อสร้าง ดัชนีราคาที่อยู่อาศัย ดัชนีความเชื่อมั่นทางธุรกิจ อัตราดอกเบี้ย และดัชนีราคาที่ดิน ซึ่งข้อมูลเหล่านี้เป็นข้อมูลรายเดือนตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556 รวมระยะเวลา 60 เดือน โดยที่ข้อมูลดัชนีราคาหุ่นกลุ่มอสังหาริมทรัพย์และก่อสร้าง ข้อมูลดัชนีเซท 50 ได้ข้อมูลจากตลาดหลักทรัพย์แห่งประเทศไทย ข้อมูลดัชนีราคาวัสดุก่อสร้าง ได้ข้อมูลจากสำนักดัชนีเศรษฐกิจการค้า ข้อมูลดัชนีราคาที่อยู่อาศัย ดัชนีความเชื่อมั่นทางธุรกิจ อัตราดอกเบี้ย และดัชนีราคาที่ดิน ได้ข้อมูลจากธนาคารแห่งประเทศไทย

วิธีการวิจัยสามารถลำดับเป็นขั้นตอนได้ดังต่อไปนี้

1. ศึกษาปัจจัยทางเศรษฐกิจที่สำคัญที่ส่งผลกระทบต่อ การเปลี่ยนแปลงดัชนีราคาหุ้นกลุ่ม อสังหาริมทรัพย์และก่อสร้างในประเทศไทย ได้แก่ ดัชนีเซท 50 ดัชนีราคาวัสดุก่อสร้าง ดัชนีราคา ที่อยู่อาศัย ดัชนีความเชื่อมั่นทางธุรกิจ อัตราดอกเบี้ย และดัชนีราคาที่ดิน

2. จัดเตรียมข้อมูลอนุกรมเวลา

3. ศึกษาการเคลื่อนไหวของข้อมูลอนุกรมเวลาว่ามีลักษณะการเคลื่อนไหวในรูปแบบใด

4. นำข้อมูลอนุกรมเวลารายเดือนตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556 มา สร้างสมการพยากรณ์ด้วยวิธีเครือข่ายเบย์เซียน วิธีบ็อกซ์-เจนกินส์ และวิธีการปรับให้เรียบแบบ เอ็กซ์โปเนนเชียล เพื่อหารูปแบบการพยากรณ์ที่เหมาะสมที่สุดสำหรับข้อมูล

#### 4.1 วิธีเครือข่ายเบย์เซียน

ในขั้นแรกของการวิเคราะห์ข้อมูลต้องทำการแปลงค่าข้อมูลให้เป็นค่าที่ไม่ต่อเนื่อง จากนั้นจึงคำนวณเครือข่ายเบย์เซียนโดยใช้ K2 Algorithm เพื่อหาว่าตัวแปรใดบ้างที่มีอิทธิพลต่อตัวแปรเป้าหมายที่สนใจศึกษา พร้อมทั้งหาตารางความน่าจะเป็นอย่างมีเงื่อนไขของตัวแปรแต่ละตัวแปร จากนั้นหารูปแบบสมการพยากรณ์ที่เหมาะสมสำหรับข้อมูล เมื่อได้รูปแบบแล้วทำการพยากรณ์ค่าในอนาคต

#### 4.2 วิธีบ็อกซ์-เจนกินส์

เป็นวิธีการพยากรณ์ที่สามารถใช้ได้กับข้อมูลที่มีการเคลื่อนไหวทุกประเภท และ จะต้องมีการตรวจสอบข้อมูลอนุกรมเวลา เพื่อพิจารณาว่าเป็นอนุกรมเวลาที่เป็นสเตชันนารีหรือไม่ จากนั้นกำหนดรูปแบบอนุกรมเวลา ทำการประมาณค่าพารามิเตอร์ที่ได้ในรูปแบบ ตรวจสอบความเหมาะสมของรูปแบบโดยการพิจารณาลักษณะการเคลื่อนไหวของอนุกรมเวลาของค่าคลาดเคลื่อน เมื่อพบรูปแบบที่เหมาะสมแล้ว จึงทำการพยากรณ์ค่าในอนาคต

#### 4.3 วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล

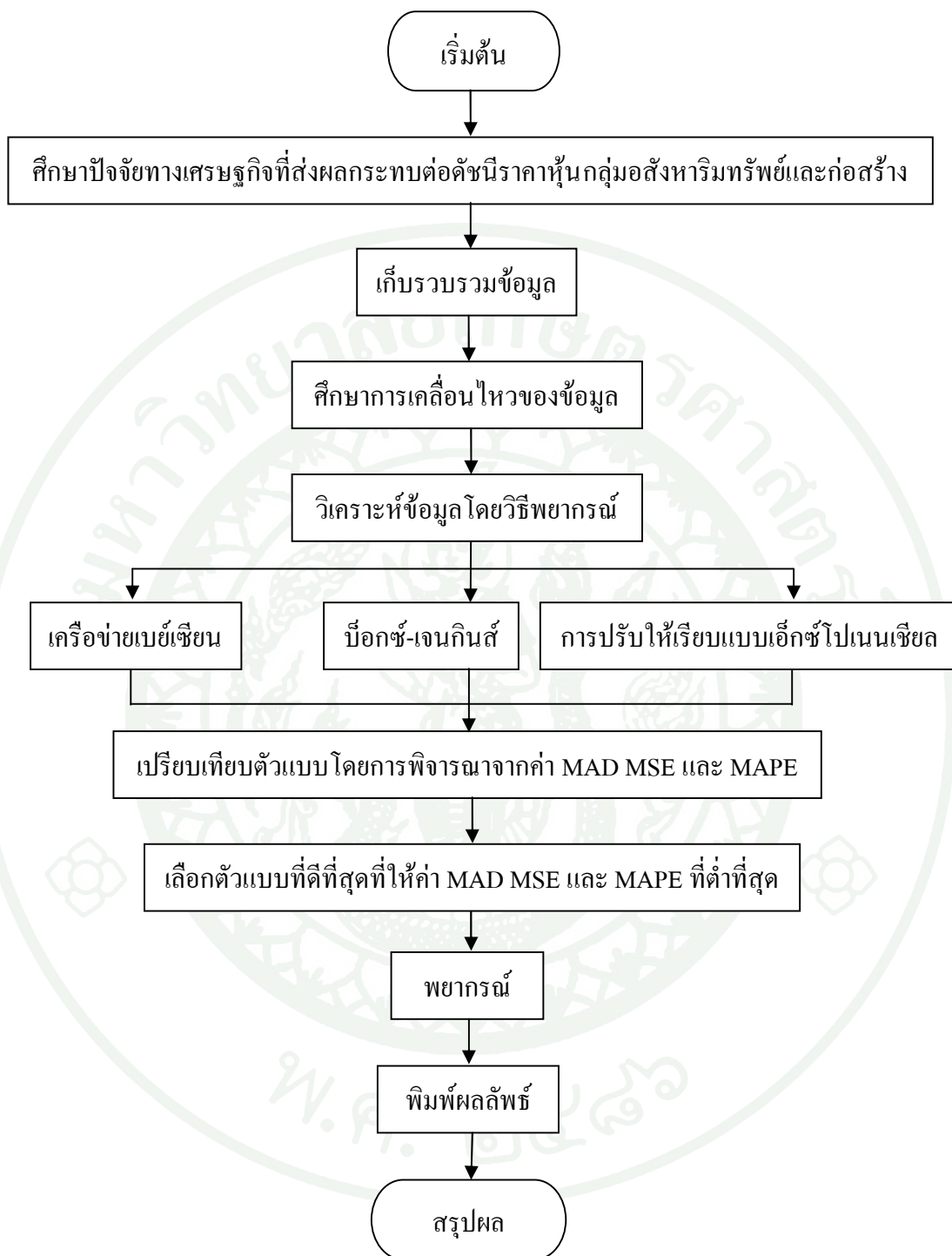
เป็นวิธีการใช้ค่าสังเกตในอดีตในการสร้างสมการพยากรณ์โดยน้ำหนักที่ให้กับค่าสังเกตแต่ละค่าแตกต่างกัน วิธีการปรับให้เรียบนี้เหมาะกับข้อมูลที่มีลักษณะการเคลื่อนไหวของอนุกรมเวลาที่มีแบบแผนเปลี่ยนแปลงอยู่ตลอดเวลา โดยทุกครั้งที่มีการสังเกตใหม่เข้ามาจะนำค่าสังเกตใหม่ดังกล่าวไปปรับสมการพยากรณ์ จากนั้นหารูปแบบของวิธีการปรับให้เรียบที่เหมาะสมกับข้อมูล และทำการพยากรณ์ค่าในอนาคต

5. นำค่าพยากรณ์ที่ได้จากข้อ 4 มาคำนวณค่าวัดความคลาดเคลื่อนของการพยากรณ์โดยพิจารณาจากค่า MAD MSE และ MAPE

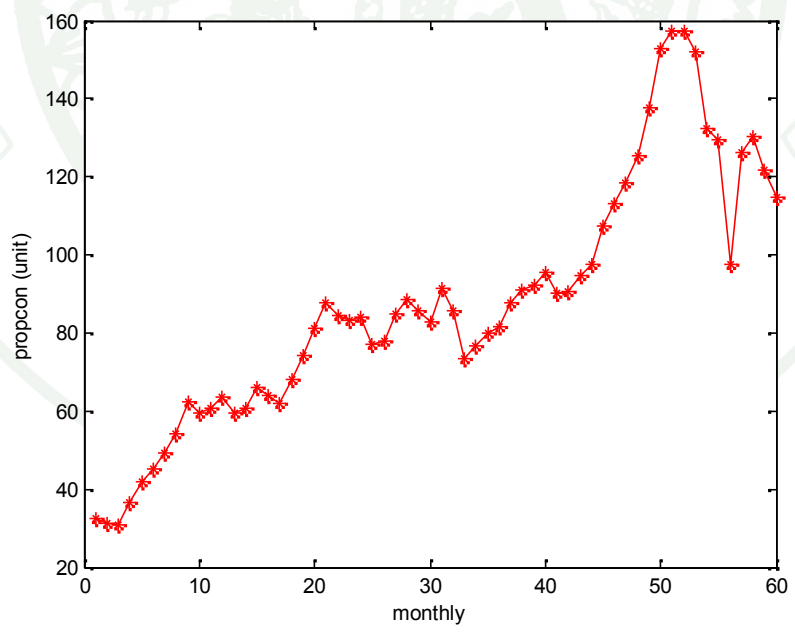
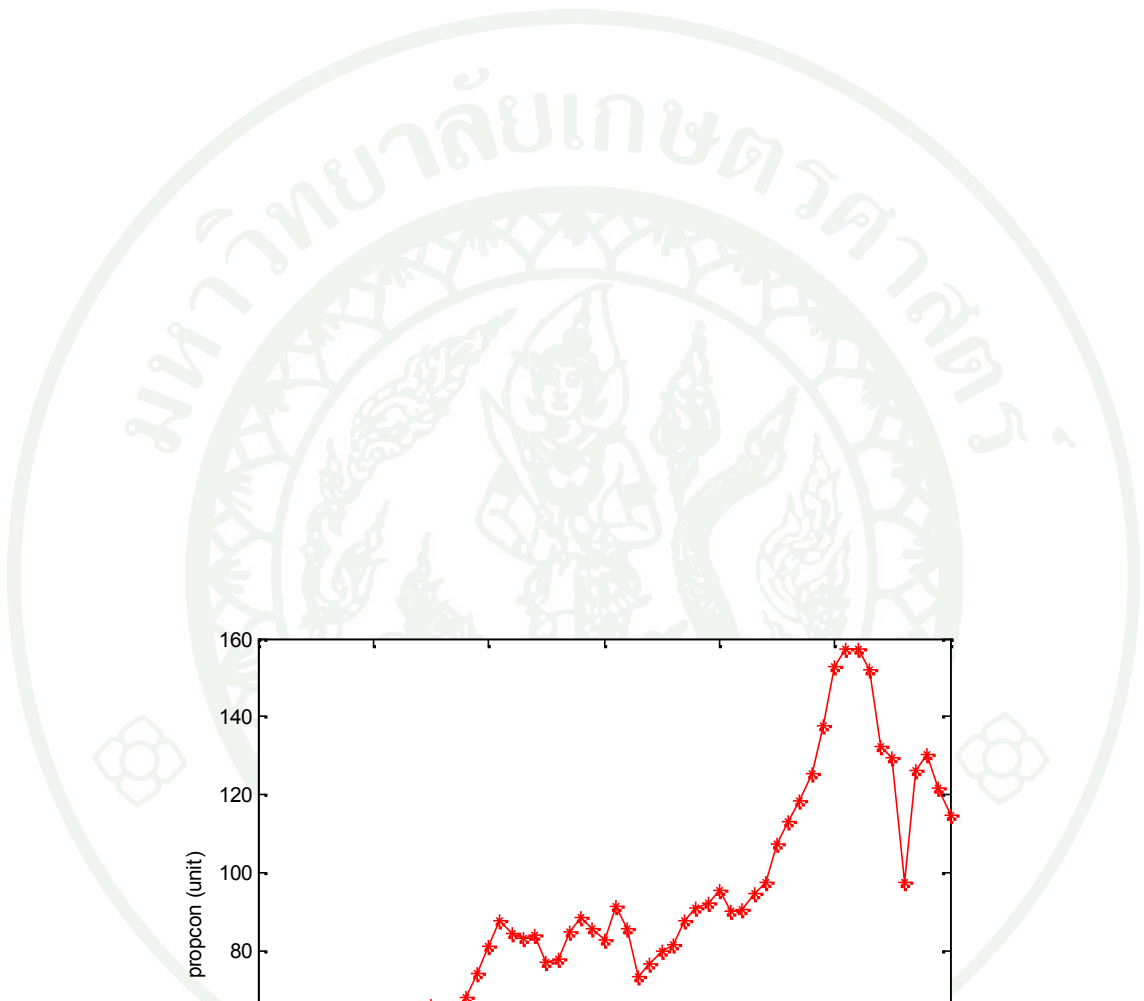
6. นำค่า MAD MSE และ MAPE ที่ได้จากอนุกรมเวลา มาพิจารณาเปรียบเทียบว่าการพยากรณ์โดยวิธีการใดเหมาะสมที่สุดสำหรับข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง

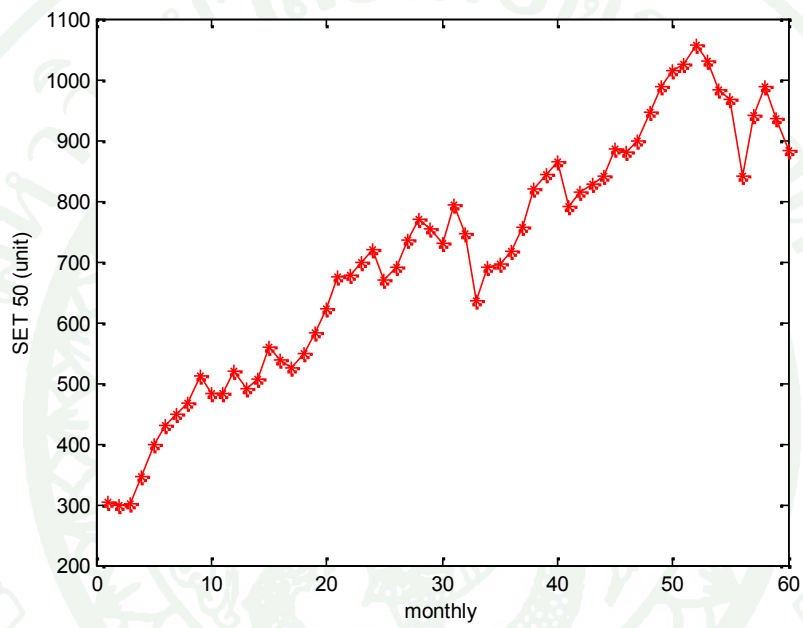
7. เมื่อได้วิธีที่เหมาะสมที่สุดจากข้อ 6 ทำการพยากรณ์ค่าในอนาคต

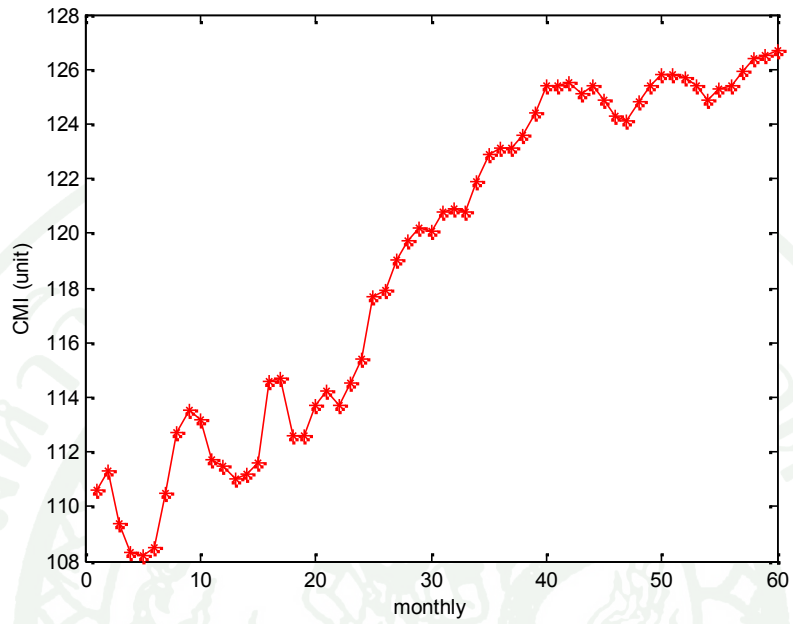
8. สรุปผลการศึกษา

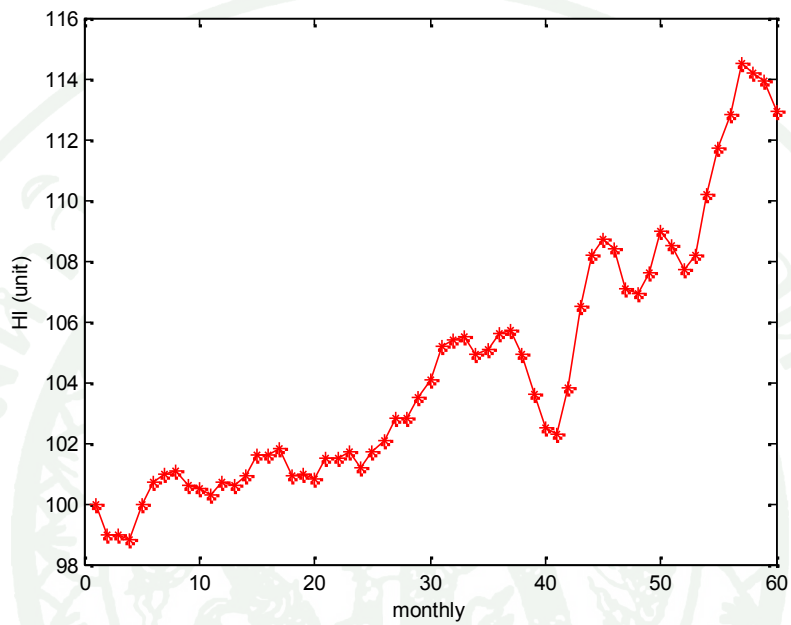


ภาพที่ 11 แผนภาพสรุปขั้นตอนการดำเนินการวิจัย







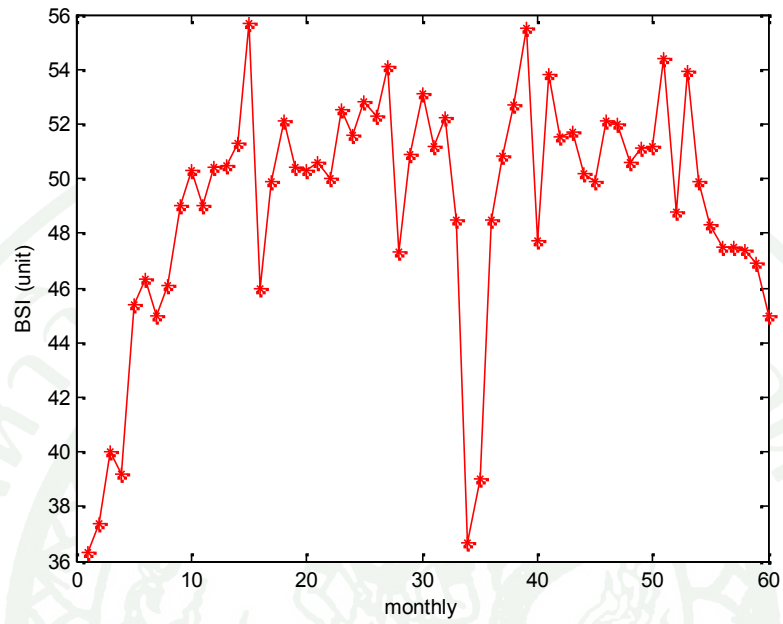


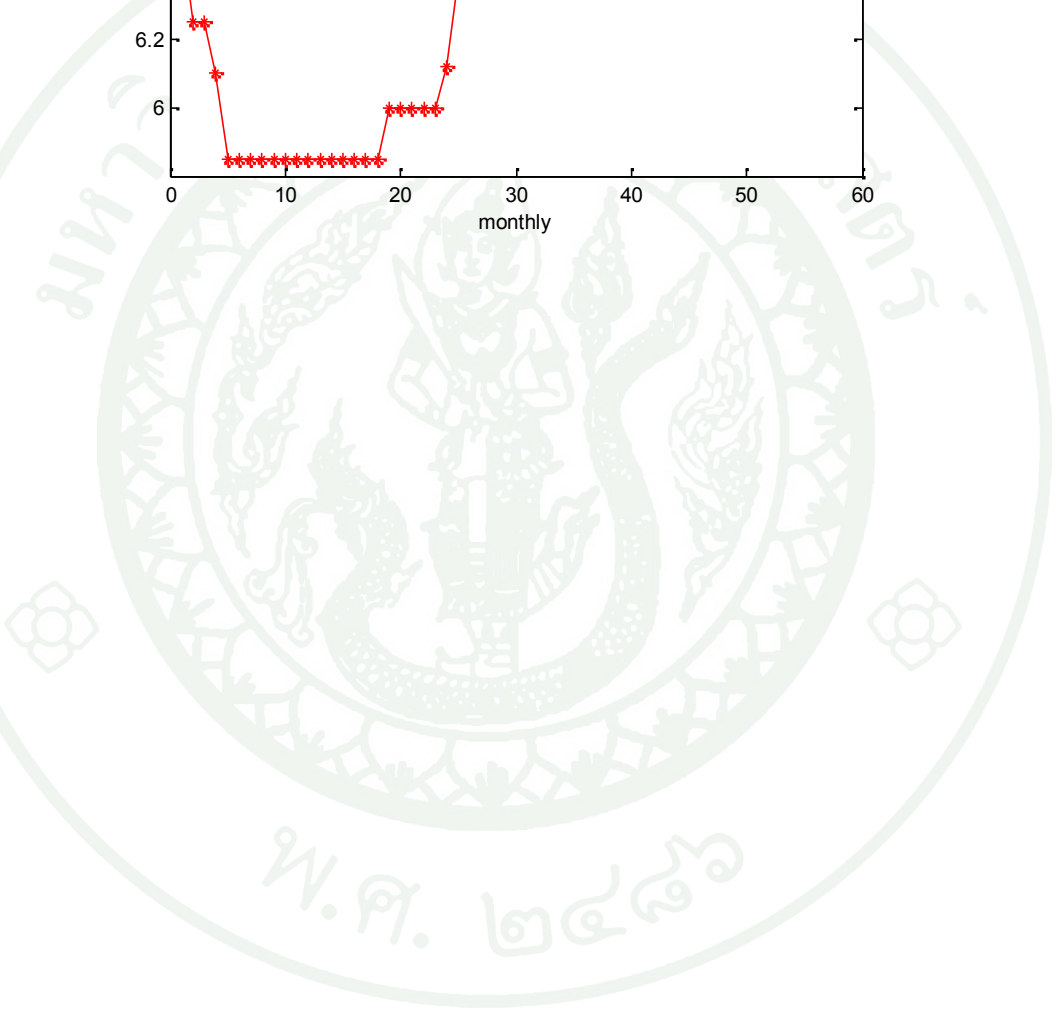
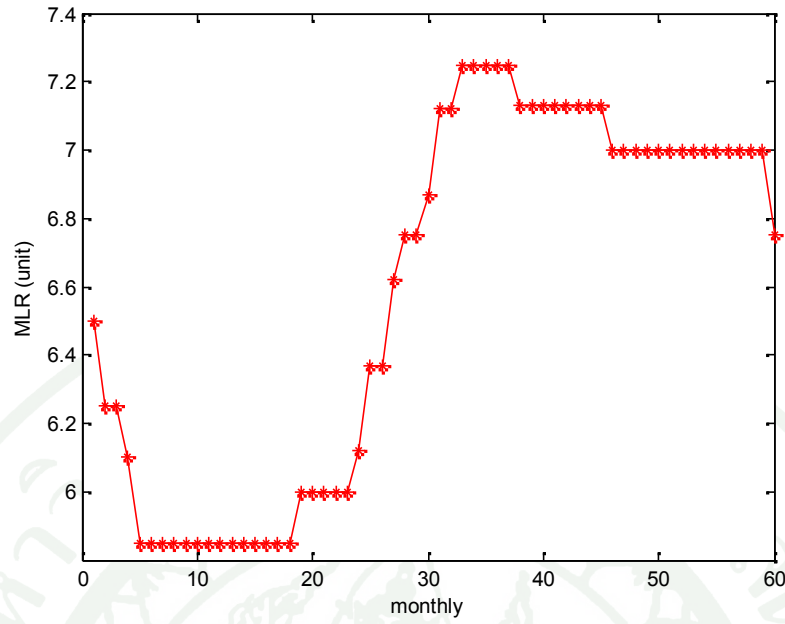
ต้นปี 2554 ดัชนีราคาที่อยู่อาศัยทรุดตัวลงอย่างมาก แต่หลังจากครึ่งปีหลังของปีเดียวกัน บ้านเดี่ยวและทาวน์เฮาส์ปรับตัวขึ้นส่งผลให้ดัชนีราคาที่อยู่อาศัยค่อยๆปรับตัวสูงขึ้น แต่หากพิจารณาในพื้นที่กรุงเทพมหานครจะพบว่าทาวน์เฮาส์ปรับตัวขึ้นในอัตราที่สูงกว่าบ้านเดี่ยว เพราะช่วงเวลาที่ผ่านมาทาวน์เฮาส์มีราคาต่ำกว่าบ้านเดี่ยว ซึ่งปัจจัยหลักที่ส่งผลให้ดัชนีราคาที่อยู่อาศัยสูงขึ้นนี้ เนื่องมาจากต้นทุนราคาค่าก่อสร้าง ราคาที่ดินเป็นตัวแปรสำคัญที่ทำให้ราคาที่อยู่อาศัยเพิ่มขึ้นตาม แต่ทั้งนี้ก็ยังมีขึ้นอยู่กับการที่ต้นทุนบางส่วนเกิดจากภาวะการแข่งขันสูงทำให้ผู้ประกอบการส่วนใหญ่เลือกที่จะขยับราคาบ้านขึ้นในอัตราน้อยกว่าต้นทุนที่เพิ่มขึ้น ไม่เช่นนั้นดัชนีราคาทาวน์เฮาส์จะต้องปรับเพิ่มขึ้นอีกเป็นเท่าตัว (ธนาคารแห่งประเทศไทย, 2555)

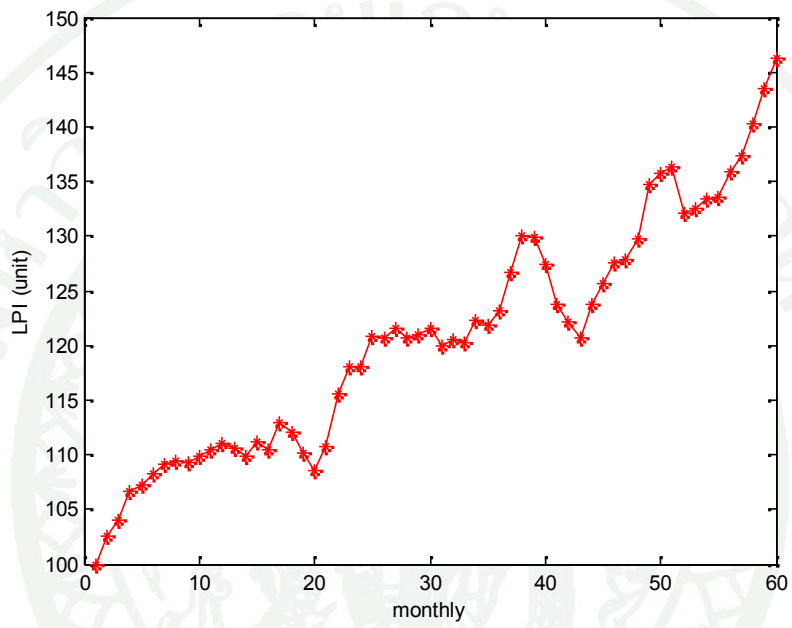
ราคาที่อยู่อาศัยในเขตกรุงเทพฯ และปริมณฑลใน 9 เดือนแรกของปี 2555 ปรับตัวเพิ่มขึ้นทุกประเภท กล่าวคือ ดัชนีราคาคอนโดมิเนียม ดัชนีราคาที่ดิน ดัชนีราคาทาวน์เฮาส์พร้อมที่ดิน และดัชนีราคาบ้านเดี่ยวพร้อมที่ดินขยายตัวร้อยละ 7.6, 4.0, 3.0 และ 1.2 จากช่วงเดียวกันของปีก่อน ตามลำดับ และสูงกว่าค่าเฉลี่ยในช่วงเดียวกันของปี 2551 ถึงปี 2554 อยู่ที่ร้อยละ 19.8, 15.9, 9.8 และ 5.2 ตามลำดับ

อย่างไรก็ดีปัจจัยที่อาจจะส่งผลกระทบต่อการเติบโตของตลาดที่อยู่อาศัย คือ

1. ความเสี่ยงของเศรษฐกิจโลก โดยเฉพาะประเทศคู่ค้าหลัก คือ สหรัฐ ยุโรป และญี่ปุ่นยังไม่ฟื้นตัวชัดเจน ประกอบกับเศรษฐกิจจีนมีแนวโน้มชะลอตัวแม้การชะลอตัวของเศรษฐกิจโลกจะไม่มีผลกระทบต่อธุรกิจอสังหาริมทรัพย์โดยตรง แต่เป็นผลทางอ้อม โดยเฉพาะต่อการส่งออกซึ่งเป็นรายได้หลักของประเทศ
2. มาตรการจำกัดวงเงินสินเชื่อของธนาคารแห่งประเทศไทยมีผลบังคับใช้ในกลุ่มโครงการบ้านแนวราบในปี 2556
3. การปรับขึ้นราคาประเมินที่ดินทั่วประเทศ เพื่อใช้ในการจดทะเบียนสิทธิ และนิติกรรมรอบบัญชีปี 2555 ถึงปี 2558 บังคับใช้ตั้งแต่ 1 กรกฎาคม 2555
4. ต้นทุนปรับตัวสูงขึ้นจากการปรับขึ้นค่าแรงขั้นต่ำเป็น 300 บาท ทั่วประเทศในปี 2556 รวมถึงราคาวัสดุก่อสร้าง และราคาที่ดินที่เพิ่มสูงขึ้น







มั่นคงของสถานภาพทางการเงินของตนเอง ผู้ประกอบการจึงมีมาตรการดึงดูดผู้ซื้อบ้านด้วยการลดราคาเพื่อแข่งขันกับคู่แข่งในตลาดทั้งรายใหญ่และรายย่อย (ศูนย์ข้อมูลอสังหาริมทรัพย์, 2553)

## ส่วนที่ 2 การกำหนดตัวแบบที่เหมาะสมของวิธีการพยากรณ์ทั้ง 3 วิธี

### 1. วิธีเครือข่ายเบย์เซียน

ในงานวิจัยนี้ผู้วิจัยเลือกใช้ข้อมูลปัจจัยทางเศรษฐกิจที่สำคัญที่ส่งผลกระทบต่อการเปลี่ยนแปลงดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง ซึ่งจะให้โปรแกรมทำการเรียนรู้โครงสร้างของความสัมพันธ์ที่เป็นไปได้ แล้วเลือกโครงสร้างที่ดีที่สุด นั่นก็คือ โครงสร้างที่ให้คะแนนสูงที่สุดซึ่งอาจเป็นหนึ่งเครือข่ายหรือกลุ่มของเครือข่ายก็ได้ จากนั้นจึงนำเครือข่ายที่ได้ไปใช้ร่วมกับข้อมูลที่มีเพื่อหาค่าและใช้เป็นตัวแทนความน่าจะเป็นอย่างมีเงื่อนไขของเครือข่ายที่ได้

#### ขั้นตอนการวิจัย

1. เลือกข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างที่สนใจศึกษาตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

2. ทำการวิเคราะห์ข้อมูลจากโปรแกรม ซึ่งมีการทำงานตามลำดับดังต่อไปนี้

#### 2.1 แปลงข้อมูลที่ได้มาให้เป็นข้อมูลชนิดที่ไม่ต่อเนื่อง

ให้ ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง แทนด้วยตัวแปร “ $X_1$ ”  
 ข้อมูลดัชนีเซท 50 แทนด้วยตัวแปร “ $X_2$ ”  
 ข้อมูลดัชนีราคาวัสดุก่อสร้าง แทนด้วยตัวแปร “ $X_3$ ”  
 ข้อมูลดัชนีราคาที่อยู่อาศัย แทนด้วยตัวแปร “ $X_4$ ”  
 ข้อมูลดัชนีความเชื่อมั่นทางธุรกิจ แทนด้วยตัวแปร “ $X_5$ ”  
 ข้อมูลอัตราดอกเบี้ย แทนด้วยตัวแปร “ $X_6$ ”  
 ข้อมูลดัชนีราคาที่ดิน แทนด้วยตัวแปร “ $X_7$ ”

2.2 นำข้อมูลที่ได้จากขั้นตอน 2.1 มาเป็นข้อมูลเข้าของการวิจัย

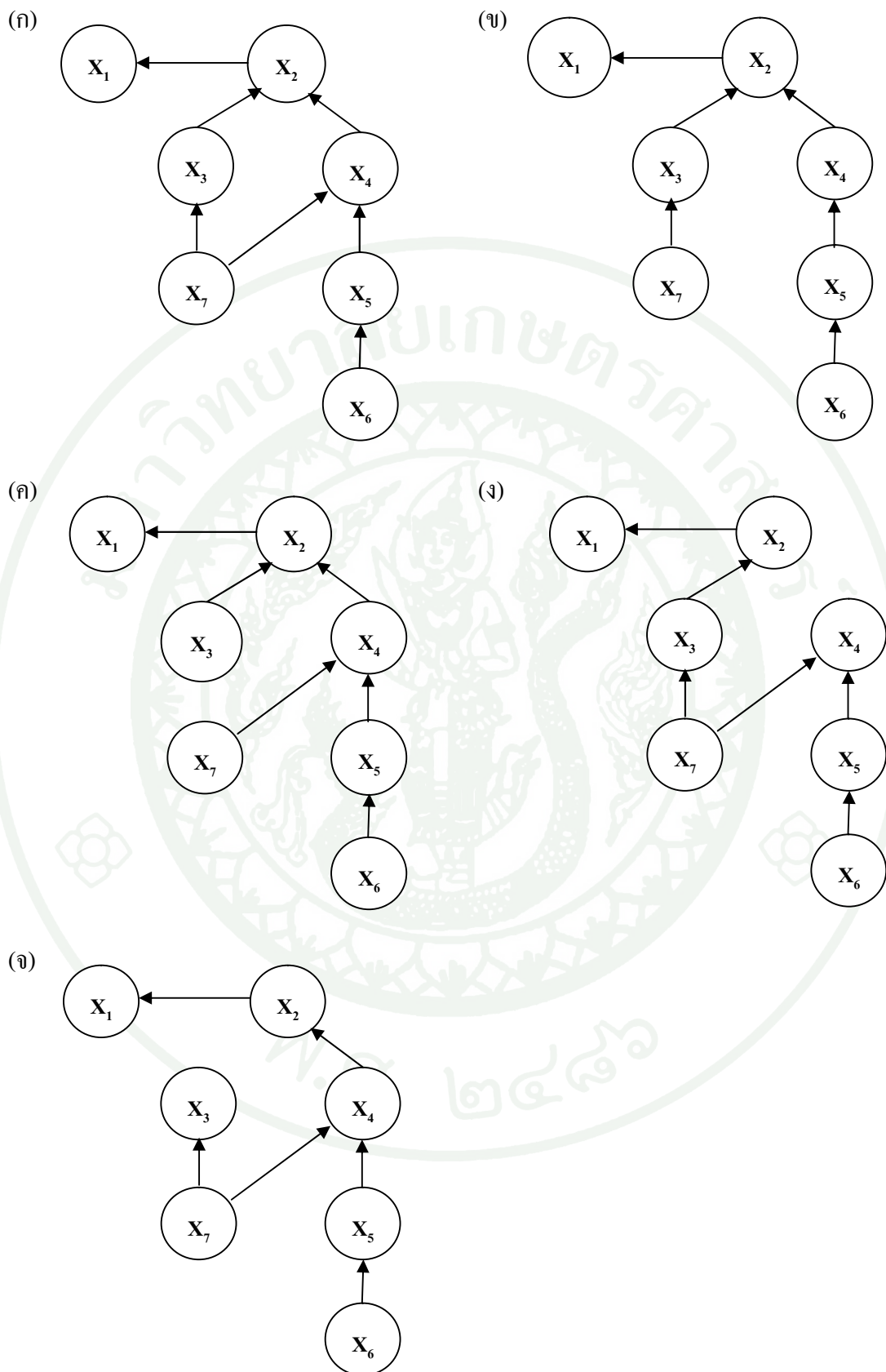
### 2.3 คำนวณค่าคะแนน K2 Algorithm ให้กับเครือข่ายทั้งหมดที่เป็นไปได้

### 2.4 เขียนกราฟสำหรับเครือข่ายเบย์เซียนที่ได้คะแนนสูงสุดดังภาพที่ 18

ในที่นี้ การแปลงข้อมูลจากชนิดต่อเนื่องเป็นข้อมูลชนิดไม่ต่อเนื่อง กำหนดเกณฑ์ดังต่อไปนี้

- ถ้าดัชนีราคาหุ้นมีราคาต่ำมาก จะแทนด้วยสัญลักษณ์ “A”
- ถ้าดัชนีราคาหุ้นมีราคาต่ำ จะแทนด้วยสัญลักษณ์ “B”
- ถ้าดัชนีราคาหุ้นมีค่อนข้างต่ำ จะแทนด้วยสัญลักษณ์ “C”
- ถ้าดัชนีราคาหุ้นมีราคาคงที่ จะแทนด้วยสัญลักษณ์ “D”
- ถ้าดัชนีราคาหุ้นมีราคาค่อนข้างสูง จะแทนด้วยสัญลักษณ์ “E”
- ถ้าดัชนีราคาหุ้นมีราคาสูง จะแทนด้วยสัญลักษณ์ “F”
- ถ้าดัชนีราคาหุ้นมีราคาสูงมาก จะแทนด้วยสัญลักษณ์ “G”

ได้เครือข่ายที่ให้คะแนนสูงที่สุดทั้งหมด 5 เครือข่ายดังต่อไปนี้



ภาพที่ 19 เครือข่ายผลลัพธ์ทั้ง 5 เครือข่ายที่ให้คะแนนสูงที่สุดโดยวิธีเครือข่ายเบย์เซียน

จากภาพที่ 19 (ก) แสดงให้เห็นว่า ปัจจัยทางเศรษฐกิจทั้ง 6 ปัจจัยส่งผลกระทบต่อ การเปลี่ยนแปลงของดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง โดยที่อัตราดอกเบี้ยมีผลกระทบต่อดัชนีความเชื่อมั่นทางธุรกิจ และส่งผลไปยังดัชนีราคาที่อยู่อาศัย เพราะเมื่อดัชนีความเชื่อมั่นทางธุรกิจมีค่าสูงขึ้น ดัชนีราคาที่อยู่อาศัยจะสูงขึ้นตามไปด้วย ส่วนดัชนีราคาที่ดินส่งผลให้ดัชนีราคาวัสดุก่อสร้างและดัชนีราคาที่อยู่อาศัยเพิ่มสูงขึ้น ซึ่งเมื่อดัชนีราคาวัสดุก่อสร้างและดัชนีราคาที่อยู่อาศัยเกิดการเปลี่ยนแปลงจะส่งผลกระทบต่อดัชนีเซท 50 และทำให้ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างเกิดการเคลื่อนไหวไปในทิศทางเดียวกัน ซึ่งเครือข่ายอีก 4 เครือข่ายจะให้ผลในทำนองเดียวกัน

จะเห็นว่าจากเครือข่ายผลลัพธ์ทั้ง 5 เครือข่ายที่ได้ ไม่ว่าจะปัจจัยทางเศรษฐกิจปัจจัยใดปัจจัยหนึ่งจะส่งผลกระทบต่อปัจจัยอื่น แต่โดยภาพรวมแล้ว ปัจจัยทางเศรษฐกิจทั้ง 6 ปัจจัยจะมีผลทำให้ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างเกิดการเปลี่ยนแปลงในที่สุด เมื่อได้เครือข่ายผลลัพธ์แล้ว จะได้ค่าความน่าจะเป็นของแต่ละเครือข่ายดังตารางต่อไปนี้

ตารางที่ 6 ค่าความน่าจะเป็น  $P(X_7 = x)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)

x	A	B	C	D	E	F	G
$P(X_7 = x)$	0.0500	0.3000	0.0500	0.2833	0.1500	0.1167	0.0500

ตารางที่ 7 ค่าความน่าจะเป็น  $P(X_6 = x)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)

x	A	B	C	D	E	F	G
$P(X_6 = x)$	0.3167	0.0667	0.0333	0.0333	0.0500	0.2500	0.2500

ตารางที่ 8 ค่าความน่าจะเป็น  $P(X_5 | X_6)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)

Probability distribution over $X_5$	$X_6$						
	A	B	C	D	E	F	G
$P(A   X_6)$	-	0.0167	-	0.0167	-	-	0.0333
$P(B   X_6)$	-	0.0333	-	-	-	-	-
$P(D   X_6)$	0.0833	-	-	-	0.0333	0.0167	-

ตารางที่ 8 (ต่อ)

Probability distribution over $X_5$	$X_6$						
	A	B	C	D	E	F	G
$P(E   X_6)$	0.0667	-	-	-	-	0.1000	0.0667
$P(F   X_6)$	0.1500	0.0167	0.0333	-	0.0167	0.0833	0.1167
$P(G   X_6)$	0.0167	-	-	0.0167	-	0.0500	0.0333

ตารางที่ 9 ค่าความน่าจะเป็น  $P(X_4 | X_5, X_7)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)

$P(X_4   X_5, X_7)$	$X_5$						
	A	B	C	D	E	F	G
$P(A   X_5, A)$	0.0333	0.0167	-	-	-	-	-
$P(A   X_5, B)$	-	0.0167	-	0.0500	0.0333	0.1167	-
$P(B   X_5, B)$	-	-	-	0.0333	0.0167	0.0167	0.0167
$P(B   X_5, C)$	-	-	-	-	0.0167	0.0333	-
$P(B   X_5, D)$	-	-	-	0.0167	-	0.0333	0.0333
$P(B   X_5, E)$	-	-	-	-	0.0167	-	-
$P(C   X_5, D)$	0.0333	-	-	-	0.0167	0.0667	0.0167
$P(C   X_5, E)$	-	-	-	-	-	0.0167	0.0167
$P(D   X_5, D)$	-	-	-	0.0167	0.0167	0.0167	-
$P(D   X_5, E)$	-	-	-	-	0.0167	0.0333	-
$P(D   X_5, F)$	-	-	-	-	-	0.0167	-
$P(E   X_5, D)$	-	-	-	-	0.0167	0.0167	-
$P(E   X_5, E)$	-	-	-	-	-	0.0167	0.0167
$P(E   X_5, F)$	-	-	-	-	0.0167	0.0167	0.0167
$P(F   X_5, F)$	-	-	-	-	0.0167	-	-
$P(G   X_5, F)$	-	-	-	-	0.0333	-	-
$P(G   X_5, G)$	-	-	-	0.0333	0.0167	-	-

ตารางที่ 10 ค่าความน่าจะเป็น  $P(X_3 | X_7)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)

Probability distribution over $X_3$	$X_7$						
	A	B	C	D	E	F	G
$P(A   X_7)$	0.0333	0.0667	-	-	-	-	-
$P(B   X_7)$	0.0167	0.1500	-	-	-	-	-
$P(C   X_7)$	-	0.0833	0.0500	-	-	-	-
$P(D   X_7)$	-	-	-	0.0333	-	-	-
$P(E   X_7)$	-	-	-	0.1167	-	-	-
$P(F   X_7)$	-	-	-	0.0500	0.0333	-	-
$P(G   X_7)$	-	-	-	0.0833	0.1167	0.1167	0.0500

ตารางที่ 11 ค่าความน่าจะเป็น  $P(X_2 | X_3, X_4)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)

$P(X_2   X_3, X_4)$	$X_3$						
	A	B	C	D	E	F	G
$P(A   X_3, A)$	0.0667	0.0167	-	-	-	-	-
$P(B   X_3, A)$	0.0333	0.0667	0.0167	-	-	-	-
$P(B   X_3, B)$	-	0.0167	-	-	-	-	-
$P(C   X_3, A)$	-	0.0500	0.0167	-	-	-	-
$P(C   X_3, B)$	-	0.0167	0.0333	-	-	-	-
$P(D   X_3, B)$	-	-	0.0667	0.0333	-	-	-
$P(D   X_3, C)$	-	-	-	-	0.0333	0.0333	-
$P(D   X_3, D)$	-	-	-	-	-	0.0167	-
$P(E   X_3, B)$	-	-	-	-	0.0333	-	0.0167
$P(E   X_3, C)$	-	-	-	-	0.0500	0.0167	0.0167
$P(E   X_3, D)$	-	-	-	-	-	0.0167	0.0167
$P(F   X_3, B)$	-	-	-	-	-	-	0.0167
$P(F   X_3, C)$	-	-	-	-	-	-	0.0167
$P(F   X_3, D)$	-	-	-	-	-	-	0.0333
$P(F   X_3, E)$	-	-	-	-	-	-	0.0500

ตารางที่ 11 (ต่อ)

$P(X_2   X_3, X_4)$	$X_3$						
	A	B	C	D	E	F	G
$P(F   X_3, G)$	-	-	-	-	-	-	0.0667
$P(G   X_3, D)$	-	-	-	-	-	-	0.0333
$P(G   X_3, E)$	-	-	-	-	-	-	0.0667
$P(G   X_3, F)$	-	-	-	-	-	-	0.0167
$P(G   X_3, G)$	-	-	-	-	-	-	0.0167

ตารางที่ 12 ค่าความน่าจะเป็น  $P(X_1 | X_2)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ก)

Probability distribution over $X_1$	$X_2$						
	A	B	C	D	E	F	G
$P(A   X_2)$	0.0833	0.0167	-	-	-	-	-
$P(B   X_2)$	-	0.1167	0.0667	-	-	-	-
$P(C   X_2)$	-	-	0.0500	0.1667	-	-	-
$P(D   X_2)$	-	-	-	0.0167	0.1667	0.0667	-
$P(E   X_2)$	-	-	-	-	-	0.0500	-
$P(F   X_2)$	-	-	-	-	-	0.0667	0.0500
$P(G   X_2)$	-	-	-	-	-	-	0.0833

ค่าความน่าจะเป็น  $P(X_7 = x)$  ดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ข) ได้ตารางเช่นเดียวกันกับตารางที่ 6

ค่าความน่าจะเป็น  $P(X_6 = x)$  ดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ข) ได้ตารางเช่นเดียวกันกับตารางที่ 7

ค่าความน่าจะเป็น  $P(X_5 | X_6)$  ดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ข) ได้ตารางเช่นเดียวกันกับตารางที่ 8

ตารางที่ 13 ค่าความน่าจะเป็น  $P(X_4 | X_5)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ข)

Probability distribution over $X_4$	$X_5$						
	A	B	C	D	E	F	G
$P(A   X_5)$	0.0333	0.0333	-	0.0500	0.0333	0.1167	-
$P(B   X_5)$	-	-	-	0.0500	0.0500	0.0833	0.0500
$P(C   X_5)$	0.0333	-	-	-	0.0167	0.0833	0.0333
$P(D   X_5)$	-	-	-	-	0.0333	0.0833	-
$P(E   X_5)$	-	-	-	-	0.0333	0.0500	0.0333
$P(F   X_5)$	-	-	-	-	0.0167	-	-
$P(G   X_5)$	-	-	-	0.0333	0.0500	-	-

ค่าความน่าจะเป็น  $P(X_3 | X_5)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ข) ได้ตารางเช่นเดียวกันกับตารางที่ 10

ค่าความน่าจะเป็น  $P(X_2 | X_3, X_4)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ข) ได้ตารางเช่นเดียวกันกับตารางที่ 11

ค่าความน่าจะเป็น  $P(X_1 | X_2)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ข) ได้ตารางเช่นเดียวกันกับตารางที่ 12

ค่าความน่าจะเป็น  $P(X_7 = x)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ค) ได้ตารางเช่นเดียวกันกับตารางที่ 6

ค่าความน่าจะเป็น  $P(X_6 = x)$  ดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ค) ได้ตารางเช่นเดียวกันกับตารางที่ 7

ค่าความน่าจะเป็น  $P(X_5 | X_6)$  ดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ค) ได้ตารางเช่นเดียวกันกับตารางที่ 8

ค่าความน่าจะเป็น  $P(X_4 | X_5, X_7)$  คำนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ค) ได้ตารางเช่นเดียวกับ ตารางที่ 9

**ตารางที่ 14** ค่าความน่าจะเป็น  $P(X_3 = x)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ค)

x	A	B	C	D	E	F	G
$P(X_3 = x)$	0.1000	0.1667	0.1333	0.0333	0.1167	0.0833	0.3667

ค่าความน่าจะเป็น  $P(X_2 | X_3, X_4)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ค) ได้ตารางเช่นเดียวกับ ตารางที่ 11

ค่าความน่าจะเป็น  $P(X_1 | X_2)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ค) ได้ตารางเช่นเดียวกับ ตารางที่ 12

ค่าความน่าจะเป็น  $P(X_7 = x)$  คำนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ง) ได้ตารางเช่นเดียวกับ ตารางที่ 6

ค่าความน่าจะเป็น  $P(X_6 = x)$  คำนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ง) ได้ตารางเช่นเดียวกับ ตารางที่ 7

ค่าความน่าจะเป็น  $P(X_5 | X_6)$  คำนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ง) ได้ตารางเช่นเดียวกับ ตารางที่ 8

ค่าความน่าจะเป็น  $P(X_4 | X_5, X_7)$  คำนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ง) ได้ตารางเช่นเดียวกับ ตารางที่ 9

ค่าความน่าจะเป็น  $P(X_3 | X_7)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ง) ได้ตารางเช่นเดียวกับ ตารางที่ 10

ตารางที่ 15 ค่าความน่าจะเป็น  $P(X_2 | X_3)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ง)

Probability distribution over $X_2$	$X_3$						
	A	B	C	D	E	F	G
$P(A   X_3)$	0.0667	0.0167	-	-	-	-	-
$P(B   X_3)$	0.0333	0.0833	0.0167	-	-	-	-
$P(C   X_3)$	-	0.0667	0.0500	-	-	-	-
$P(D   X_3)$	-	-	0.0667	0.0333	0.0333	0.0500	-
$P(E   X_3)$	-	-	-	-	0.0833	0.0333	0.0500
$P(F   X_3)$	-	-	-	-	-	-	0.1833
$P(G   X_3)$	-	-	-	-	-	-	0.1333

ค่าความน่าจะเป็น  $P(X_1 | X_2)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (ง) ได้ตารางเช่นเดียวกันกับตารางที่ 12

ค่าความน่าจะเป็น  $P(X_7 = x)$  ดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (จ) ได้ตารางเช่นเดียวกันกับตารางที่ 6

ค่าความน่าจะเป็น  $P(X_6 = x)$  ดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (จ) ได้ตารางเช่นเดียวกันกับตารางที่ 7

ค่าความน่าจะเป็น  $P(X_5 | X_6)$  ดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (จ) ได้ตารางเช่นเดียวกันกับตารางที่ 8

ค่าความน่าจะเป็น  $P(X_4 | X_5, X_7)$  ดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (จ) ได้ตารางเช่นเดียวกันกับตารางที่ 9

ค่าความน่าจะเป็น  $P(X_3 | X_7)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (จ) ได้ตารางเช่นเดียวกันกับตารางที่ 10

ตารางที่ 16 ค่าความน่าจะเป็น  $P(X_2 | X_1)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (จ)

Probability distribution over $X_1$	$X_2$						
	A	B	C	D	E	F	G
$P(A   X_2)$	0.0833	-	-	-	-	-	-
$P(B   X_2)$	0.1167	0.0167	-	-	-	-	-
$P(C   X_2)$	0.0667	0.0500	-	-	-	-	-
$P(D   X_2)$	-	0.1000	0.0667	0.0167	-	-	-
$P(E   X_2)$	-	0.0500	0.0833	0.0333	-	-	-
$P(F   X_2)$	-	0.0167	0.0167	0.0333	0.0500	-	0.0667
$P(G   X_2)$	-	-	-	0.0333	0.0667	0.0167	0.0167

ค่าความน่าจะเป็น  $P(X_1 | X_2)$  ของดัชนีราคาหุ้นของเครือข่ายในภาพที่ 19 (จ) ได้ตารางเช่นเดียวกันกับตารางที่ 12

จากวิธีเครือข่ายเบย์เซียนนี้ สามารถหาค่าวัดความคลาดเคลื่อนของการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556 ซึ่งแสดงดังตารางที่ 17 และได้สมการพยากรณ์ได้ดังนี้

$$Y_t = \arg \left[ \max_{r_t} (P(r_t | B)) \right]$$

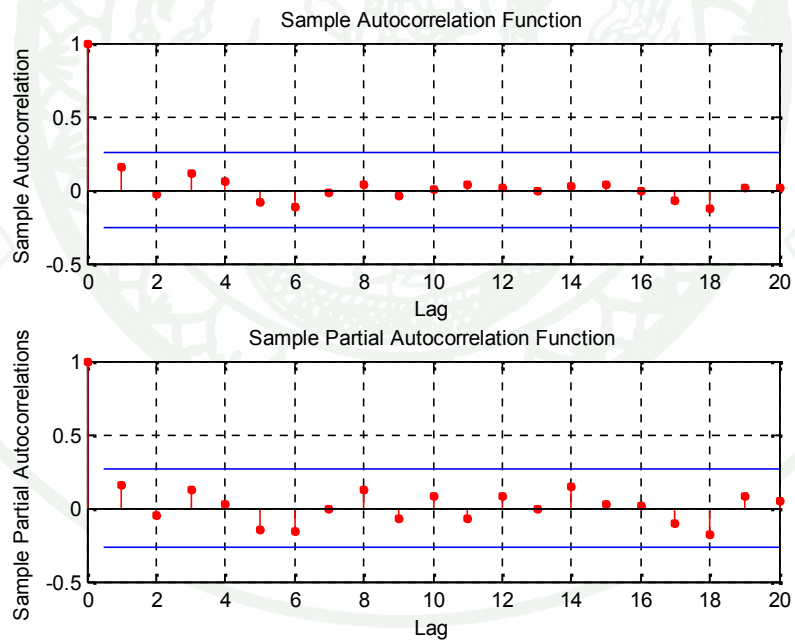
โดยที่  $r_t$  แทนเซตของอัตราส่วนราคาต่อกำไร เมื่อ  $l = 1, 2, \dots, L$

ตารางที่ 17 แสดงค่าความคลาดเคลื่อนจากการพยากรณ์ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างด้วยวิธีเครือข่ายเบย์เซียน

MAD	MSE	MAPE
9.3856	123.3816	13.1546

$\{Z_t\}$

$$Z_t = \frac{1}{\sqrt{Y_t}}$$



เนื่องจากข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างเป็นอนุกรมเวลาที่ไม่มีความสัมพันธ์เชิงเส้น จึงต้องทำการแปลงอนุกรมเวลาเดิมให้เป็นอนุกรมเวลาใหม่ที่มีความสัมพันธ์เชิงเส้น ซึ่งข้อมูลราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างนี้มีลักษณะการเคลื่อนไหวเนื่องจากอิทธิพลของแนวโน้ม ดังนั้นจึงกำจัดอิทธิพลของแนวโน้มเชิงเส้น โดยการหาผลต่างได้ออนุกรมใหม่  $\{W_t\}$  ดังนี้

$$W_t = \nabla^d Z_t = Z_t - Z_{t-1}$$

## 2.2 การประมาณค่าพารามิเตอร์จากรูปแบบได้ค่าประมาณดังตารางที่ 18

ตารางที่ 18 แสดงการประมาณค่าพารามิเตอร์จากรูปแบบ White Noise ของดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง

ค่าประมาณพารามิเตอร์				
Type	Estimate	Standard Error	T	P
Mu	1.3912	1.0555	1.32	0.1927

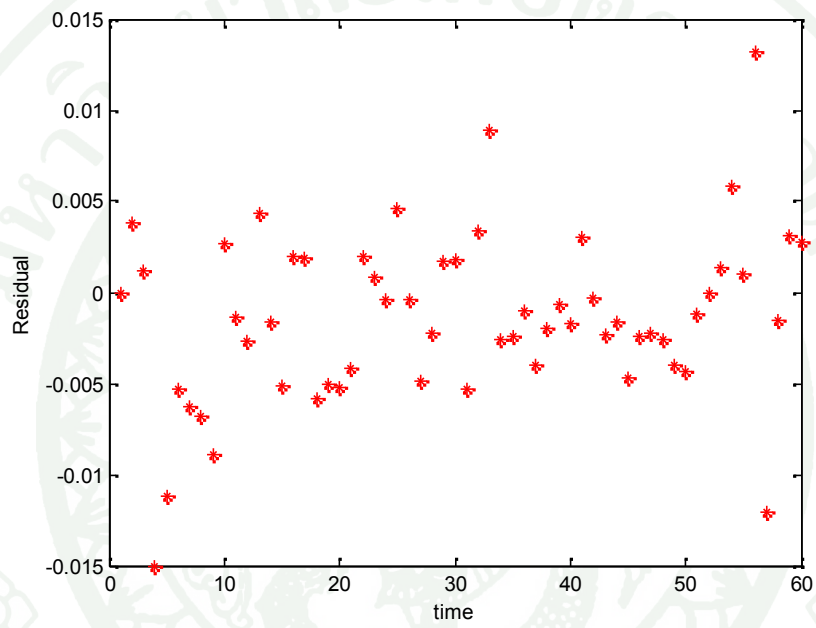
รูปแบบ White Noise มีสมการพยากรณ์ คือ  $W_t = u + v_t$  แต่จากตารางที่ 18 เมื่อทดสอบความเหมาะสมของตัวแบบโดยการทดสอบนัยสำคัญของพารามิเตอร์ พบว่า ค่า P-value ของค่าเฉลี่ยมีค่าเท่ากับ 0.1927 ซึ่งมีค่ามากกว่าระดับนัยสำคัญ 0.05 ซึ่งไม่มีความเหมาะสมที่จะอยู่ในตัวแบบจึงต้องทำการตัดทิ้งค่าคงที่ทิ้งไป จึงได้รูปแบบใหม่ คือ รูปแบบ White Noise สามารถเขียนสมการพยากรณ์ได้ดังนี้

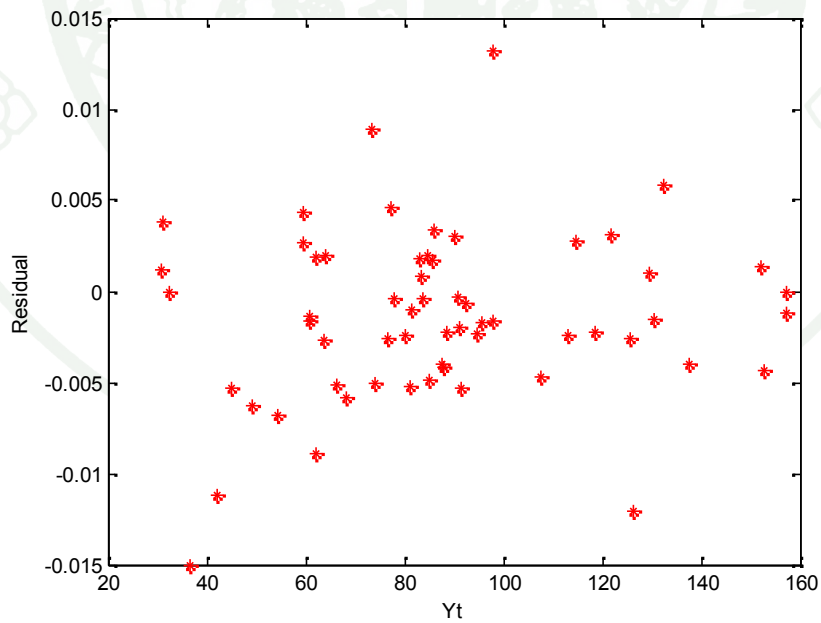
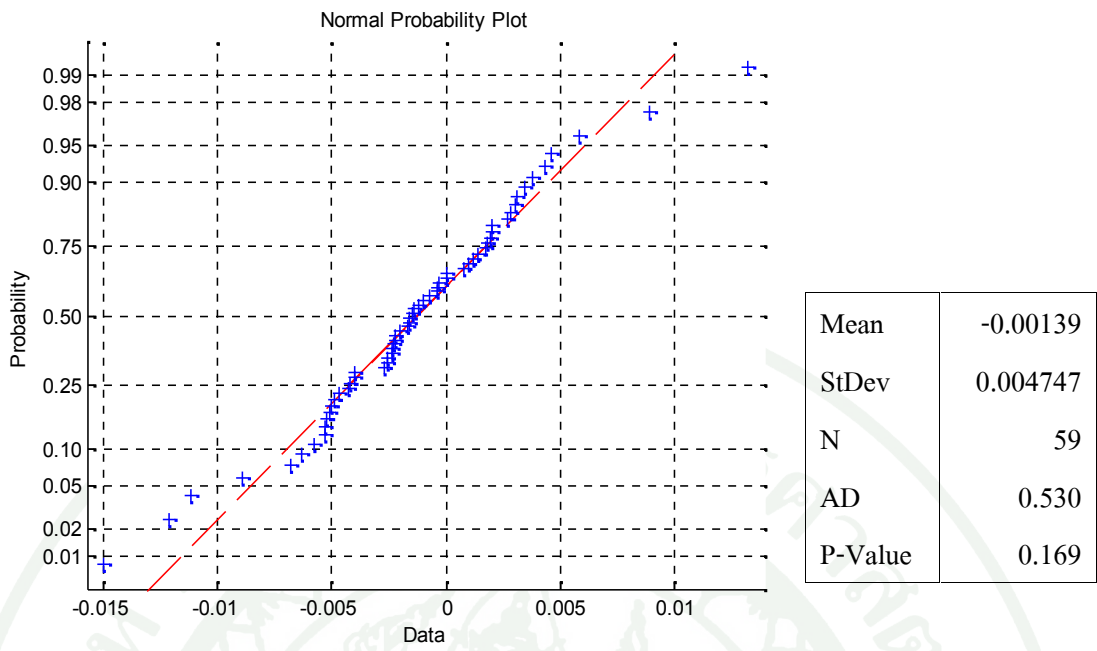
$$W_t = v_t$$

โดยที่  $W_t = Z_t - Z_{t-1}$

$$Z_t = \frac{1}{\sqrt{Y_t}}$$

$$v_t = 0$$





เมื่อพิจารณาจากกราฟ Normal Probability Plot แสดงในภาพที่ 22 พบว่า ค่าความคลาดเคลื่อนมีการแจกแจงปกติ เนื่องจากมีค่าสถิติทดสอบ Anderson Darling เท่ากับ 0.530 และค่า p-value มีค่าเท่ากับ 0.169 ซึ่งมีค่ามากกว่าระดับนัยสำคัญ 0.05 ส่วนการทดสอบอนุกรมเวลาของค่าคลาดเคลื่อนที่แต่ละเวลาว่ามีลักษณะการเคลื่อนไหวที่เป็นอิสระจากกันจะทดสอบจากกราฟแผนภาพการกระจายระหว่างค่าคลาดเคลื่อน ( $v_t$ ) กับค่าจริง ( $Y_t$ ) ดังแสดงในภาพที่ 23 พบว่า ค่าคลาดเคลื่อนมีการกระจายอย่างสุ่ม มีการขึ้นลงเรีวรอบๆ เส้นศูนย์ นั่นคือ ค่าคลาดเคลื่อนเป็นอิสระจากกัน และสำหรับการทดสอบค่าคลาดเคลื่อนโดยการทดสอบสัมประสิทธิ์สหสัมพันธ์ในตัวเองของค่าคลาดเคลื่อนจากตัวสถิติ Box-Pierce พบว่า ค่าคลาดเคลื่อนของการพยากรณ์ที่อยู่ห่างกัน 1, 2, 3, ..., m ช่วงเป็นอิสระจากกันแสดงดังตารางที่ 19

**ตารางที่ 19** แสดงค่าสถิติของ Box-Pierce จากตัวแบบ White Noise ของดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง

To Lag	Chi-Square	DF	P-value	Autocorrelations					
6	6.78	6	0.3419	0.232	0.070	0.185	0.110	-0.009	-0.045
12	8.72	12	0.7269	0.044	0.087	0.022	0.065	0.085	0.070
18	10.64	18	0.9092	0.049	0.070	0.083	0.045	-0.019	-0.081
24	16.16	24	0.8823	0.042	0.047	-0.055	0.091	0.183	-0.084

ดังนั้น สามารถสรุปได้ว่า ตัวแบบ  $\frac{1}{\sqrt{Y_t}} = \frac{1}{\sqrt{Y_{t-1}}} + \epsilon_t$  มีความเหมาะสม โดยมีค่าวัดความคลาดเคลื่อนของการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556 ซึ่งแสดงดังตารางที่ 20

**ตารางที่ 20** แสดงค่าความคลาดเคลื่อนจากการพยากรณ์ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างด้วยวิธีบ็อกซ์-เจนกินส์

MAD	MSE	MAPE
5.7307	66.5540	6.5117

### 3. วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์

เนื่องจากข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างมีการเคลื่อนไหวเนื่องจากอิทธิพลของแนวโน้มเชิงเส้นตรง ดังนั้นผู้วิจัยจึงเลือกวิธีการพยากรณ์โดยวิธีการของโฮลต์ โดยการพยากรณ์ต้องกำหนดค่าคงที่ปรับค่าให้เรียบที่เหมาะสม ซึ่งจะต้องมีค่าอยู่ระหว่าง 0 ถึง 1 จากการใช้โปรแกรม Microsoft Excel ในการวิเคราะห์ได้ค่าคงที่สำหรับปรับแนวโน้ม และค่าคงที่สำหรับปรับความลาดชันที่ดีที่สุดคือ  $r = 1$  และ  $x = 0$  ตามลำดับ ซึ่งค่าปรับให้เรียบนี้จะให้ค่าผลรวมกำลังสองของความคลาดเคลื่อนที่มีค่าต่ำที่สุด ซึ่งสามารถเขียนสมการพยากรณ์  $p$  ช่วงเวลาล่วงหน้าของโฮลต์ ณ เวลา  $t$  ได้ดังนี้

$$\hat{Y}_{t+p} = \hat{T}_t + p\hat{S}_t$$

โดยที่  $\hat{T}_t = Y_t$

$$\hat{S}_t = \hat{S}_{t-1}$$

จากรูปแบบสมการพยากรณ์ข้างต้น เมื่อนำมาตรวจสอบข้อสมมติฐานเบื้องต้นโดยใช้สถิติทดสอบ พบว่า ความคลาดเคลื่อนมีการแจกแจงปกติ และค่าประมาณสัมประสิทธิ์สหสัมพันธ์แบบออโตของค่าความคลาดเคลื่อนของการพยากรณ์ด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์เป็นอิสระต่อกัน แสดงว่ารูปแบบสมการพยากรณ์ที่กำหนดขึ้นมีความเหมาะสมอย่างมีนัยสำคัญที่ระดับนัยสำคัญ 0.05 จึงนำค่าไปหาค่าพยากรณ์ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างระหว่างเดือนมกราคม 2552 ถึงเดือนธันวาคม 2556 โดยมีค่าความคลาดเคลื่อนจากการพยากรณ์แสดงดังตารางที่ 21

**ตารางที่ 21** แสดงค่าความคลาดเคลื่อนจากการพยากรณ์ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์ และก่อสร้างด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์

MAD	MSE	MAPE
5.4168	64.2380	6.3403

ส่วนที่ 3 คือ การเปรียบเทียบประสิทธิภาพของตัวแบบที่ได้และการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทยจากวิธีพยากรณ์ที่เหมาะสมที่สุด

จากค่าวัดความคลาดเคลื่อนของการพยากรณ์ 3 ค่า คือ MAD MSE และ MAPE ที่ได้จากวิธีการพยากรณ์ของดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง โดยทำการพยากรณ์ในช่วงเดือนมกราคม 2552 ถึงธันวาคม 2556 สำหรับการพยากรณ์ในช่วงดังกล่าวนี้ พบว่า วิธีวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์ให้ค่าวัดความคลาดเคลื่อนน้อยที่สุด รองลงมาคือ วิธีบ็อกซ์-เจนกินส์ และวิธีเครือข่ายเบย์เซียนเป็นวิธีที่ให้ค่าความคลาดเคลื่อนมากที่สุด ดังแสดงค่าในตารางที่ 22

ตารางที่ 22 ค่าวัดความคลาดเคลื่อนของการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างช่วงเดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

วิธีการพยากรณ์	ค่าวัดความคลาดเคลื่อน		
	MAD	MSE	MAPE
1.วิธีเครือข่ายเบย์เซียน	9.3856	123.3816	13.1546
2.วิธีบ็อกซ์-เจนกินส์	5.7307	66.5540	6.5117
3.วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์	<b>5.4168</b>	<b>64.2380</b>	<b>6.3403</b>

จากผลการหาตัวแบบที่เหมาะสมทั้ง 3 วิธีที่ทำการศึกษาของข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย ผลการเปรียบเทียบดังตารางที่ 14 มีดังนี้

พิจารณาจากค่า MAD พบว่า วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์ให้ค่าต่ำที่สุดเท่ากับ 5.4168 วิธีบ็อกซ์-เจนกินส์ให้ค่าเท่ากับ 5.7307 และวิธีเครือข่ายเบย์เซียนให้ค่าสูงที่สุดเท่ากับ 9.3856

พิจารณาจากค่า MSE พบว่า วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์ให้ค่าต่ำที่สุดเท่ากับ 64.2380 วิธีบ็อกซ์-เจนกินส์ให้ค่าเท่ากับ 66.5540 และวิธีเครือข่ายเบย์เซียนให้ค่าสูงที่สุดเท่ากับ 123.3816

พิจารณาจากค่า MAPE พบว่า วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์ให้ค่าต่ำที่สุดเท่ากับ 6.3403 วิธีบ็อกซ์-เจนกินส์ให้ค่าเท่ากับ 6.5117 และวิธีเคลื่อนย้ายเบย์เซียนให้ค่าสูงที่สุดเท่ากับ 13.1546

จากสมการพยากรณ์ของวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์นำไปคำนวณหาค่าพยากรณ์ เพื่อตรวจสอบว่าค่าพยากรณ์จากสมการพยากรณ์ที่สร้างขึ้นนั้นแตกต่างจากค่าจริงหรือไม่ โดยการเปรียบเทียบค่าพยากรณ์ด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์กับค่าจริงตั้งแต่เดือนมกราคม 2557 ถึงเดือนธันวาคม 2557 แสดงในตารางที่ 23

**ตารางที่ 23** แสดงการเปรียบเทียบค่าพยากรณ์ ค่าจริง และค่าคลาดเคลื่อนด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์ ตั้งแต่เดือนมกราคมถึงเดือนธันวาคม 2557

เดือน	ค่าพยากรณ์ ( $\hat{Y}_t$ )	ค่าจริง ( $Y_t$ )	$e_t = Y_t - \hat{Y}_t$	$\% \frac{ e_t }{Y_t}$
มกราคม	116.26	112.15	-4.11	3.67
กุมภาพันธ์	117.91	118.94	1.03	0.86
มีนาคม	119.56	121.08	1.52	1.25
เมษายน	121.22	125.77	4.55	3.62
พฤษภาคม	122.87	127.20	4.33	3.41
มิถุนายน	124.52	135.29	10.77	7.96
กรกฎาคม	126.17	136.20	10.03	7.36
สิงหาคม	127.82	141.14	13.32	9.43
กันยายน	129.47	144.09	14.62	10.15
ตุลาคม	131.12	145.17	14.05	9.68
พฤศจิกายน	132.77	145.32	12.54	8.63
ธันวาคม	134.43	136.18	1.75	1.29

จากตารางที่ 23 พบว่า ตั้งแต่เดือนเมษายนค่าพยากรณ์เริ่มมีค่าห่างจากค่าจริงมากขึ้นเรื่อย ๆ ซึ่งเป็นสัญญาณเตือนว่าพยากรณ์ที่สร้างขึ้นนั้นเริ่มไม่เหมาะสม จึงต้องมีการตรวจสอบค่าพยากรณ์

ด้วย tracking signal ได้แก่  $TS_1$ ,  $TS_2$ ,  $TS_3$  และ  $TS_4$  การตรวจสอบค่าพยากรณ์ ถ้าค่า  $TS_t$  อยู่ในช่วงควบคุมอธิบายได้ว่าค่าพยากรณ์มีค่าใกล้เคียงกับค่าจริง จึงหาค่าพยากรณ์ต่อไปได้ แต่ถ้าค่า  $TS_t$ อยู่นอกช่วงควบคุมอธิบายได้ว่าค่าพยากรณ์มีค่าต่างจากค่าจริงต้องปรับสมการพยากรณ์ใหม่ซึ่งอาจจะเป็นการกำหนดรูปแบบอนุกรมเวลาใหม่หรือใช้วิธีการพยากรณ์ใหม่โดยเพิ่มขนาดของอนุกรมเวลาแสดงดังตารางที่ 24

ตารางที่ 24 แสดงการคำนวณค่า  $TS_1$ ,  $TS_2$ ,  $TS_3$  และ  $TS_4$

t	$Y_t$	$TS_1$	$TS_2$		$TS_3$	$TS_4$		
		$e_t$	$\sum e_t$	$\frac{\sum e_t}{\sqrt{t}}$	$MTE_t$	$E_t$	$MAD_t$	$\frac{E_t}{MAD_t}$
1	112.15	-4.11	-4.11	-4.1100		-0.4110	5.2861	-0.0778
2	118.94	1.03	-3.08	-2.1779		-0.2669	4.8605	-0.0549
3	121.08	1.52	-1.56	-0.9007		-0.0882	4.5265	-0.0195
4	125.77	4.55	2.99	1.4950	2.9900	0.3756	4.5288	0.0829
5	127.20	4.33	7.32	3.2736	11.4300*	0.7710	4.5089	0.1710
6	135.29	10.77	18.09	7.3852	21.1700	1.7709	5.1350	0.3449
7	136.20	10.03	28.12	10.6287	29.6809	2.5969	5.6246	0.4617
8	141.14	13.32	41.44	14.6514	38.4504	3.6692	6.3941	0.5738*
9	144.09	14.62	56.06	18.6862*	48.7387	4.7641	7.2165	0.6602
10	145.17	14.05	70.11	22.1693	52.0156	5.6924	7.8996	0.7206
11	145.32	12.54	82.65	24.9203	54.5304	6.3777	8.3642	0.7625
12	136.18	1.75	84.41	24.3658	42.9651	5.9154	7.7032	0.7679

ช่วงควบคุม :  $TS_1$ , (-16.0297,16.0297);  $TS_2$ , (-16.0297,16.0297);  $TS_3$ , (-8.0149,8.0149);

$TS_4$ , (-0.5735,0.5735)

หมายเหตุ \* เตือน

พิจารณาค่า  $TS_1$ ,  $TS_2$ ,  $TS_3$  และ  $TS_4$  ดังตารางที่ 24 โดยเทียบกับช่วงควบคุมที่สัมพันธ์กัน พบว่า  $TS_2$  มีค่าอยู่นอกช่วงควบคุมเมื่อ  $t = 9$  นั่นคือ ณ เวลา  $t = 9$  จะมีสัญญาณเตือนว่าควรมีการปรับปรุงสมการพยากรณ์เพราะค่าพยากรณ์เริ่มห่างจากค่าจริงมาก  $TS_3$  มีค่าอยู่นอกช่วงควบคุมเมื่อ  $t = 5$  นั่นคือ ณ เวลา  $t = 5$  จะมีสัญญาณเตือนว่า ควรมีการปรับปรุงสมการพยากรณ์ เพราะค่าพยากรณ์เริ่ม

ห่างจากค่าจริงมาก และ  $TS_4$  มีค่าอยู่นอกช่วงควบคุมเมื่อ  $t=8$  นั่นคือ ณ เวลา  $t=8$  จะมีสัญญาณเตือนว่าควรมีการปรับปรุงสมการพยากรณ์เพราะค่าพยากรณ์เริ่มห่างจากค่าจริงมาก จากผลที่ได้แสดงว่าการใช้ tracking signal ที่ต่างกันอาจจะให้สัญญาณเตือนในเวลาที่แตกต่างกัน

## วิจารณ์

การวิเคราะห์ข้อมูลอนุกรมเวลาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย โดยใช้วิธีการพยากรณ์ 3 วิธี คือ วิธีเครือข่ายเบย์เซียน วิธีบ็อกซ์-เจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล มีการวิจารณ์ดังรายละเอียดต่อไปนี้

1. วิธีเครือข่ายเบย์เซียน เป็นวิธีที่ใช้หาเหตุผลภายใต้ความไม่แน่นอน โดยอาศัยการเรียนรู้โครงสร้างเครือข่ายเบย์เซียนจากข้อมูลดิบซึ่งบ่งบอกถึงเหตุการณ์ที่จะเกิดขึ้น และเวลาที่ใช้ในการเรียนรู้โครงสร้างดังกล่าวนี้ขึ้นอยู่กับตัวแปรและขนาดของข้อมูล ยิ่งข้อมูลมีจำนวนตัวแปรมากจะทำให้รูปแบบที่ต้องพิจารณามากขึ้นตามไปด้วย ในงานวิจัยนี้ผู้วิจัยได้นำโปรแกรม WEKA ซึ่งเป็นโปรแกรมสำเร็จรูปที่สามารถวิเคราะห์หารูปแบบโครงสร้างเครือข่ายเบย์เซียน มาใช้ร่วมกับโปรแกรม Matlab ที่สามารถวิเคราะห์ข้อมูลอนุกรมเวลา นำทั้งสองโปรแกรมมาประยุกต์ใช้ในการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง แต่อย่างไรก็ตามก่อนที่จะนำโปรแกรมทั้งสองมาทำการวิเคราะห์ข้อมูลควรต้องมีความรู้ ความเข้าใจในกระบวนการเรียนรู้ของเครือข่ายเบย์เซียนนี้ด้วย เพื่อสามารถใช้ประโยชน์ได้อย่างเต็มที่และมีประสิทธิภาพสูงสุด แต่จากการวิเคราะห์ข้อมูลอนุกรมเวลา วิธีนี้มีความยุ่งยากและซับซ้อนในการวิเคราะห์ข้อมูลไม่ว่าจะเป็นการหาความสัมพันธ์ของตัวแปรแต่ละตัว อีกทั้งยังมีวิธีการหารูปแบบเครือข่ายหลายขั้นตอนซึ่งยากต่อการอธิบาย และยังไม่ใช่วิธีที่ดีที่สุดเพราะให้ค่าความคลาดเคลื่อนสูงกว่าวิธีอื่น สาเหตุเนื่องมาจากวิธีนี้อาจจะเหมาะสมสำหรับข้อมูลที่เป็นข้อมูลเชิงคุณภาพมากกว่าข้อมูลเชิงปริมาณที่ทำการศึกษานั้นเอง

2. วิธีบ็อกซ์-เจนกินส์เป็นการพยากรณ์ที่นำอนุกรมเวลาในอดีตมาพยากรณ์อนุกรมเวลาในอนาคตโดยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองและสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วนเพื่อหาตัวแบบที่เหมาะสม ซึ่งข้อดีของการวิเคราะห์ด้วยวิธีนี้ คือ วิธีนี้มีความเหมาะสมสำหรับการพยากรณ์ในระยะเวลานั้น ๆ ซึ่งข้อมูลที่น่ามาวิเคราะห์ในงานวิจัยนี้เป็นข้อมูลรายเดือนระยะเวลา 60 เดือนซึ่งถือเป็นข้อมูลที่มีระยะเวลาไม่ยาวนานนัก เมื่อทำการวิเคราะห์โดยวิธีบ็อกซ์-

เงินกินสี่จึงให้ผลที่ค่อนข้างแม่นยำและมีประสิทธิภาพ แต่ยังไม่ใช่วิธีที่เหมาะสมที่สุด เนื่องจากค่าวัดความคลาดเคลื่อนที่ได้จากวิธีนี้ยังไม่ใช่ว่าต่ำที่สุด

3. วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ข้อมูลอนุกรมเวลาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างมีการเคลื่อนไหวของแนวโน้มเส้นตรงและไม่มีความผันแปรเนื่องจากฤดูกาล จึงใช้การวิเคราะห์อนุกรมเวลาโดยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์ ผลการศึกษาพบว่า วิธีนี้เป็นวิธีที่ให้ค่าความคลาดเคลื่อนต่ำที่สุดจึงเหมาะสมที่สุดสำหรับนำไปพยากรณ์ดัชนีราคาหุ้นในอนาคต ข้อดีของการวิเคราะห์ด้วยวิธีนี้ คือ เป็นวิธีการพยากรณ์ที่ง่ายและไม่มีความยุ่งยากมากนัก ช่วยประหยัดเวลาและค่าใช้จ่ายในการพยากรณ์ ซึ่งวิธีนี้เหมาะสำหรับข้อมูลที่มีลักษณะการเคลื่อนไหวของอนุกรมเวลาที่มีแบบแผนที่เปลี่ยนแปลงอยู่ตลอดเวลา ดังนั้นเมื่อมีค่าสังเกตใหม่เข้ามาจะต้องนำค่าสังเกตเหล่านั้นไปปรับสมการพยากรณ์ และทำการตรวจสอบเพื่อหารูปแบบที่เหมาะสมต่อไป และจากการพยากรณ์ล่วงหน้าเพื่อหาช่วงการพยากรณ์ที่เหมาะสม โดยการพิจารณาค่าวัดความคลาดเคลื่อน MAD MSE และ MAPE พบว่า การพยากรณ์ล่วงหน้าในระยะเวลาสั้น ๆ 1 - 4 เดือนเหมาะสมที่สุดสำหรับการพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างในประเทศไทย

## สรุปและข้อเสนอแนะ

### สรุป

การศึกษาในครั้งนี้ผู้วิจัยใช้ข้อมูลดัชนีเซท 50 ดัชนีราคาวัสดุก่อสร้าง ดัชนีราคาที่อยู่อาศัย ดัชนีความเชื่อมั่นทางธุรกิจ อัตราดอกเบี้ย และดัชนีราคาที่ดิน ซึ่งเป็นข้อมูลรายเดือนตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556 ทำการวิเคราะห์ห้แนวโน้มดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์ และก่อสร้างในประเทศไทย โดยใช้วิธีการพยากรณ์ 3 วิธี คือ วิธีเครือข่ายเบย์เซียน วิธีบ็อกซ์-เจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล เมื่อได้รูปแบบที่เหมาะสมที่สุดแล้ว จะนำรูปแบบดังกล่าวมาหาช่วงการพยากรณ์ล่วงหน้า 1 เดือนถึง 12 เดือน เพื่อเปรียบเทียบหาช่วงการพยากรณ์ล่วงหน้าที่เหมาะสมที่สุด ซึ่งเกณฑ์ที่ใช้ในการวัดความคลาดเคลื่อนของการพยากรณ์ จะพิจารณาจากค่า MAD MSE และ MAPE ซึ่งจากการวิเคราะห์ห้อนุกรมเวลา พบว่า อนุกรมเวลาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างมีลักษณะการเคลื่อนไหวของแนวโน้มเส้นตรงแต่ไม่มีอิทธิพลของฤดูกาล ผลการศึกษาสามารถสรุปผลได้ดังตารางที่ 25

ตารางที่ 25 แสดงการเปรียบเทียบวิธีการพยากรณ์ทั้ง 3 วิธี

วิธีการพยากรณ์	ตัวแบบ	ค่าวัดความคลาดเคลื่อนของการพยากรณ์		
		MAD	MSE	MAPE
1. วิธีเครือข่ายเบย์เซียน	$Y_t = \arg \left[ \max_r (P(r   B)) \right]$	9.3856	123.3816	13.1546
2. วิธีบ็อกซ์-เจนกินส์	$\frac{1}{\sqrt{Y_t}} = \frac{1}{\sqrt{Y_{t-1}}} + \epsilon_t$	5.7307	66.5540	6.5117
3. วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลต์	$\hat{Y}_{t+p} = \hat{T}_t + p\hat{S}_t$ โดยที่ $\hat{T}_t = Y_t$ และ $\hat{S}_t = \hat{S}_{t-1}$	<b>5.4168</b>	<b>64.2380</b>	<b>6.3403</b>

จากการตรวจสอบค่าพยากรณ์ที่ได้โดยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโสด์ เมื่อพิจารณาจากค่า  $TS_1$ ,  $TS_2$ ,  $TS_3$  และ  $TS_4$  โดยเทียบกับช่วงควบคุมที่สัมพันธ์กัน พบว่า  $TS_2$ ,  $TS_3$  และ  $TS_4$  มีค่าอยู่นอกช่วงควบคุมเมื่อ  $t = 9$ ,  $t = 5$  และ  $t = 8$  ตามลำดับ นั่นคือ ณ เวลา  $t = 9$ ,  $t = 5$  และ  $t = 8$  จะมีสัญญาณเตือนว่าควรมีการปรับปรุงสมการพยากรณ์เพราะค่าพยากรณ์เริ่มห่างจากค่าจริงมาก

จากการวิเคราะห์เพื่อหาช่วงการพยากรณ์ล่วงหน้าที่เหมาะสมที่สุด พบว่า การพยากรณ์ล่วงหน้าในระยะเวลาสั้น ๆ 1 - 4 เดือน เหมาะสมที่สุดสำหรับอนุกรมเวลาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง

### ข้อเสนอแนะ

1. ในการศึกษาครั้งนี้ ผู้วิจัยได้ศึกษาตัวแปรที่มีผลกระทบต่อดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างจากผลงานของผู้วิจัยหลาย ๆ ท่านที่ทำการศึกษาเกี่ยวกับปัจจัยที่มีผลกระทบต่อดัชนีราคาหุ้นในกลุ่มต่าง ๆ ของประเทศไทย ซึ่งอาจจะมียางตัวแปรที่ผู้วิจัยไม่ได้นำมาพิจารณาด้วย เช่น อัตราเงินเฟ้อ ผลิตภัณฑ์มวลรวมในประเทศ (GDP) ดัชนีการลงทุนภาคเอกชน เป็นต้น ดังนั้นในการ ศึกษาครั้งต่อไปอาจนำตัวแปรเหล่านี้มาพิจารณาร่วมด้วย
2. เนื่องจากการลงทุนในอสังหาริมทรัพย์นั้นมีความเสี่ยงสูง ดังนั้น ผู้ที่จะลงทุนในธุรกิจประเภทนี้ควรที่จะทำการศึกษาข้อมูลอย่างระมัดระวัง รวมทั้งทำความเข้าใจในปัจจัยที่อาจส่งผลกระทบต่อด้วย เพื่อช่วยให้นักลงทุนทำการตัดสินใจลงทุนได้อย่างมีประสิทธิภาพสูงสุดและลดความเสี่ยงให้น้อยลง
3. เมื่อเวลาผ่านไปทำให้ข้อมูลที่น่ามาศึกษาในครั้งนี้มีการปรับเปลี่ยนค่าสังเกตให้เป็นปัจจุบันมากขึ้น ดังนั้น ควรมีการตรวจสอบความเหมาะสมของตัวแบบที่ใช้ในการพยากรณ์อีกครั้งว่ามีความเหมาะสมกับข้อมูลชุดนั้นหรือไม่ เพื่อจะได้้นำตัวแบบที่เหมาะสมไปใช้ประโยชน์ในคราวต่อไป

## เอกสารและสิ่งอ้างอิง

ข่าวอสังหาริมทรัพย์. 2553. ศูนย์ข้อมูลอสังหาริมทรัพย์. แหล่งที่มา:

<http://www.reic.or.th/Default.aspx>, 12 ธันวาคม 2556.

ตลาดหลักทรัพย์แห่งประเทศไทย. 2557. **SET Index Highlight Report**. แหล่งที่มา:

[http://www.set.or.th/th/products/index/setindex\\_p1.html](http://www.set.or.th/th/products/index/setindex_p1.html), 8 มกราคม 2557.

ทรงศิริ แต่สมบัติ. 2549. การพยากรณ์เชิงปริมาณ. พิมพ์ครั้งที่ 1. สำนักพิมพ์  
มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

ธนาคารแห่งประเทศไทย. 2557. **ดัชนีความเชื่อมั่นทางธุรกิจ**. แหล่งที่มา:

[http://www.bot.or.th/Thai/EconomicConditions/Thai/Index/Pages/BSI\\_Series.aspx](http://www.bot.or.th/Thai/EconomicConditions/Thai/Index/Pages/BSI_Series.aspx), 30  
มกราคม 2557.

ธนาคารแห่งประเทศไทย. 2557. **ดัชนีราคาที่ดิน**. แหล่งที่มา:

<http://www2.bot.or.th/statistics/ReportPage.aspx?reportID=680&language=th>, 14 มีนาคม  
2557.

ธนาคารแห่งประเทศไทย. 2557. **ดัชนีราคาที่อยู่อาศัย**. แหล่งที่มา:

<http://www2.bot.or.th/statistics/ReportPage.aspx?reportID=680&language=th>, 7  
กุมภาพันธ์ 2557.

ธนาคารแห่งประเทศไทย. 2557. **อัตราดอกเบี้ย**. แหล่งที่มา:

[http://www.bot.or.th/thai/statistics/financialmarkets/interestrates/\\_layouts/application/inter  
est\\_rate/IN\\_Rate.aspx](http://www.bot.or.th/thai/statistics/financialmarkets/interestrates/_layouts/application/interest_rate/IN_Rate.aspx), 9 กุมภาพันธ์ 2557.

ธัญพร อริยะเสถณี. 2553. การพยากรณ์อัตราเงินเฟ้อของประเทศไทย. วิทยานิพนธ์ปริญญาโท,  
มหาวิทยาลัยเกษตรศาสตร์.

ประกาศ มาถาวร. 2552. ปัจจัยที่มีผลต่อผลประกอบการของบริษัทพัฒนาอสังหาริมทรัพย์ที่จดทะเบียนในตลาดหลักทรัพย์แห่งประเทศไทย. สารนิพนธ์การศึกษาค้นคว้าด้วยตนเอง, มหาวิทยาลัยนเรศวร.

โรชนานี หะยีสสา และปริดาภรณ์ กาญจนสำราญวงศ์. 2553. การเปรียบเทียบการวิเคราะห์ข้อมูลอนุกรมเวลาโดยวิธีบ็อกซ์-เจนกินส์ และวิธีการถดถอยที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียลกำลังสอง กรณีศึกษามูลค่าการส่งออกปลาทูน่ากระป๋องของประเทศไทย. วารสารมหาวิทยาลัยทักษิณ 13(2): 1-10.

วราฤทธิ์ พานิชกิจ โกศลกุล. 2550. การพยากรณ์ดัชนีราคาผู้บริโภครายเดือนของกรุงเทพมหานคร: กรณีศึกษาเปรียบเทียบโดยวิธีการของบ็อกซ์-เจนกินส์ วิธีการของโฮสต์ และวิธีการพยากรณ์รวม. การประชุมวิชาการมหาวิทยาลัยศรีปทุม. (คณิตศาสตร์และสถิติ). มหาวิทยาลัยศรีปทุม, กรุงเทพฯ.

สมพล สุนัยรัตนภรณ์. 2551. การเรียนรู้โครงสร้างเบย์เซียนเน็ตเวิร์คด้วย Genetic Algorithm โดยใช้รูปแบบ Directed Acyclic Graph. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์) ภาควิชาวิทยาการคอมพิวเตอร์, มหาวิทยาลัยเกษตรศาสตร์.

สาชน โยบายการเงิน. 2556. รายงานเศรษฐกิจและการเงินประจำปี 2552-2556. ธนาคารแห่งประเทศไทย.

สำนักดัชนีเศรษฐกิจการค้า. 2557. ดัชนีราคาวัสดุก่อสร้าง. แหล่งที่มา:

[http://www.indexpr.moc.go.th/price\\_present/csi/stat/other/conyear.asp](http://www.indexpr.moc.go.th/price_present/csi/stat/other/conyear.asp), 3 กุมภาพันธ์ 2557.

ออมพิไล มโนรัตน์. 2548. การทำนายเครือข่ายการควบคุมยีนจากข้อมูลไมโครอาร์เรย์ของดีเอ็นเอด้วยเทคนิคเครือข่ายเบย์เซียน. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเชียงใหม่.

Ben-Gal, I. 2007. **Bayesian networks**, pp.1-6. In F. Ruggeri, R. S. Kennett, & F. W. Faltin (Eds.). *Encyclopedia of statistics in equity and reliability* (Vol. 1). John Wiley & Sons.

Gnophanxay, C. 2011. **A Combination of Bayesian Belief Network and Genetic Algorithm for Student Achievements of Faculty of Science, National University of Laos**. M.S.

Thesis, King Mongkut's University of Technology North Bangkok.

Gregory F. C., E. Herskovits. 1993. **A Bayesian Method for the Introduction of Probabilistic Networks from Data**. *Machine Learning*, 9: 309-347.

Jensen, F. V. 1996. **An introduction to Bayesian networks**. Aalborg University, Denmark.

Ruiz, C. 2005. **Illustration of the K2 Algorithm for Learning Bayes Net Structures**.

Massachusetts, U.S.:Department of Computer Science, Worcester Polytechnic Institute.

Available Source: [http://web.cs.wpi.edu/~cs539/s05/Projects/k2\\_algorithm.pdf](http://web.cs.wpi.edu/~cs539/s05/Projects/k2_algorithm.pdf), 24

February 2014.

Zuo, Y. and E. Kita. 2012. Stock price forecast using Bayesian network. **Expert Systems with Applications** 39:6729-6737. Cited Pearl, J., and S. Russell. 2002. **Handbook of brain theory and neural networks**. 157-160.



ภาคผนวก



ภาคผนวก ก  
ข้อมูลปัจจัยทางเศรษฐกิจต่าง ๆ ที่ทำการศึกษาตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

ตารางผนวกที่ ก1 ข้อมูลดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

เดือน	ปี				
	2552	2553	2554	2555	2556
มกราคม	32.53	59.44	77.28	87.67	137.61
กุมภาพันธ์	31.15	60.93	77.85	91.09	152.66
มีนาคม	30.72	66.14	85.03	92.38	157.16
เมษายน	36.54	64.01	88.51	95.49	157.22
พฤษภาคม	42.06	62.11	85.79	90.17	151.94
มิถุนายน	45.13	68.17	82.97	90.65	132.30
กรกฎาคม	47.23	74.11	91.53	94.78	129.36
สิงหาคม	54.32	81.26	85.83	97.77	97.77
กันยายน	62.18	87.75	73.29	107.50	126.21
ตุลาคม	59.16	84.52	76.66	113.05	130.50
พฤศจิกายน	60.90	83.28	80.05	118.46	121.79
ธันวาคม	63.55	83.87	81.43	125.55	114.61

ที่มา: ตลาดหลักทรัพย์แห่งประเทศไทย (2557)

ตารางผนวกที่ ก2 ข้อมูลดัชนีเซท 50 ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

เดือน	ปี				
	2552	2553	2554	2555	2556
มกราคม	303.60	490.77	668.58	757.45	988.21
กุมภาพันธ์	299.58	507.67	690.89	818.78	1014.03
มีนาคม	300.20	558.28	735.98	842.91	1024.92
เมษายน	346.81	538.94	770.69	863.82	1055.62
พฤษภาคม	398.35	524.14	754.08	791.54	1029.01
มิถุนายน	430.35	547.47	729.78	814.38	982.99
กรกฎาคม	449.92	583.83	794.01	828.67	966.03
สิงหาคม	467.27	622.66	747.07	840.32	840.32
กันยายน	511.82	674.40	636.30	886.58	940.88
ตุลาคม	483.66	677.54	691.97	879.06	987.06
พฤศจิกายน	483.55	698.45	697.33	898.81	935.22
ธันวาคม	520.69	720.20	718.40	945.39	883.40

ที่มา: ตลาดหลักทรัพย์แห่งประเทศไทย (2557)

ตารางผนวกที่ ก3 ข้อมูลดัชนีราคาวัสดุก่อสร้าง ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

เดือน	ปี				
	2552	2553	2554	2555	2556
มกราคม	110.60	111.00	117.70	123.10	125.40
กุมภาพันธ์	111.30	111.20	117.90	123.60	125.80
มีนาคม	109.40	111.60	119.90	124.40	125.80
เมษายน	108.30	114.60	119.70	125.40	125.70
พฤษภาคม	108.20	114.70	120.20	125.40	125.40
มิถุนายน	108.50	112.60	120.10	125.50	124.90
กรกฎาคม	110.50	112.60	120.80	125.10	125.30
สิงหาคม	112.70	113.70	120.90	125.40	125.40
กันยายน	113.50	114.20	120.80	124.90	125.90
ตุลาคม	113.20	113.70	121.90	124.30	126.40
พฤศจิกายน	111.70	114.50	122.90	124.10	126.50
ธันวาคม	111.50	115.40	123.10	124.80	126.70

ที่มา: สำนักดัชนีเศรษฐกิจการค้า (2557)

ตารางผนวกที่ ก4 ข้อมูลดัชนีราคาที่อยู่อาศัย ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

เดือน	ปี				
	2552	2553	2554	2555	2556
มกราคม	100.00	100.60	101.70	105.70	107.60
กุมภาพันธ์	99.00	100.90	102.10	104.90	109.00
มีนาคม	99.00	101.60	102.80	103.60	108.50
เมษายน	98.80	101.60	102.80	102.50	107.70
พฤษภาคม	100.00	101.80	103.50	102.30	108.20
มิถุนายน	100.70	100.90	104.10	103.80	110.20
กรกฎาคม	101.00	101.00	105.20	106.50	111.70
สิงหาคม	101.10	100.80	105.40	108.20	112.80
กันยายน	100.60	101.50	105.50	108.70	114.50
ตุลาคม	100.50	101.50	104.90	108.40	114.20
พฤศจิกายน	100.30	101.70	105.10	107.10	113.90
ธันวาคม	100.70	101.20	105.60	106.90	112.90

ที่มา: ธนาคารแห่งประเทศไทย (2557)

ตารางผนวกที่ ก5 ข้อมูลดัชนีความเชื่อมั่นทางธุรกิจ ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

เดือน	ปี				
	2552	2553	2554	2555	2556
มกราคม	36.30	50.47	52.80	50.80	51.10
กุมภาพันธ์	37.40	51.30	52.30	52.70	51.20
มีนาคม	40.00	55.70	54.10	55.50	54.40
เมษายน	39.20	46.00	47.30	47.70	48.80
พฤษภาคม	45.40	49.90	50.90	53.80	53.90
มิถุนายน	46.30	52.10	53.10	51.50	49.90
กรกฎาคม	45.00	50.40	51.20	51.70	48.30
สิงหาคม	46.10	50.30	52.20	50.20	47.50
กันยายน	49.00	50.60	48.50	49.90	47.50
ตุลาคม	50.30	50.00	36.70	52.10	47.40
พฤศจิกายน	49.00	52.50	39.00	52.00	46.90
ธันวาคม	50.40	51.60	48.50	50.60	45.00

ที่มา: ธนาคารแห่งประเทศไทย (2557)

ตารางผนวกที่ 6 ข้อมูลอัตราดอกเบี้ย ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

เดือน	ปี				
	2552	2553	2554	2555	2556
มกราคม	6.50	5.85	6.37	7.25	7.00
กุมภาพันธ์	6.25	5.85	6.37	7.13	7.00
มีนาคม	6.25	5.85	6.62	7.13	7.00
เมษายน	6.10	5.85	6.75	7.13	7.00
พฤษภาคม	5.85	5.85	6.75	7.13	7.00
มิถุนายน	5.85	5.85	6.87	7.13	7.00
กรกฎาคม	5.85	6.00	7.12	7.13	7.00
สิงหาคม	5.85	6.00	7.12	7.13	7.00
กันยายน	5.85	6.00	7.25	7.13	7.00
ตุลาคม	5.85	6.00	7.25	7.00	7.00
พฤศจิกายน	5.85	6.00	7.25	7.00	7.00
ธันวาคม	5.85	6.12	7.25	7.00	6.75

ที่มา: ธนาคารแห่งประเทศไทย (2557)

ตารางผนวกที่ ก7 ข้อมูลดัชนีราคาที่ดิน ตั้งแต่เดือนมกราคม 2552 ถึงเดือนธันวาคม 2556

เดือน	ปี				
	2552	2553	2554	2555	2556
มกราคม	100.00	110.60	120.80	126.70	134.70
กุมภาพันธ์	102.50	109.80	120.70	130.00	135.70
มีนาคม	104.00	111.20	121.60	129.90	136.30
เมษายน	106.70	110.40	120.70	127.40	132.10
พฤษภาคม	107.30	112.90	121.00	123.80	132.50
มิถุนายน	108.30	112.00	121.50	122.20	133.40
กรกฎาคม	109.10	110.10	119.90	120.70	133.60
สิงหาคม	109.40	108.60	120.50	123.70	135.90
กันยายน	109.30	110.70	120.30	125.70	137.40
ตุลาคม	109.90	115.60	122.30	127.60	140.30
พฤศจิกายน	110.40	118.60	121.80	127.80	143.50
ธันวาคม	111.10	118.10	123.20	129.70	146.20

ที่มา: ธนาคารแห่งประเทศไทย (2557)



## ตัวอย่างการวิเคราะห์อนุกรมเวลา

ในงานวิจัยนี้ ศึกษาการพยากรณ์อนุกรมเวลา 3 วิธี ได้แก่ วิธีเครือข่ายเบย์เซียน วิธีบ็อกซ์-เจนกินส์ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮลด์ เพื่อพยากรณ์ดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้างของประเทศไทย ซึ่งในที่นี่จะแสดงตัวอย่างการวิเคราะห์รูปแบบการพยากรณ์สำหรับแต่ละวิธี ดังรายละเอียดต่อไปนี้

### 1. วิธีเครือข่ายเบย์เซียน

เริ่มต้นด้วยการแปลงค่าข้อมูลให้เป็นค่าที่ไม่ต่อเนื่อง และกำหนดเครือข่ายเบย์เซียนโดยการคำนวณ K2 Algorithm โดยใช้โปรแกรม Matlab R2010b

```
propcon = stock(:,1);
set50 = stock(:,2);
CMI = stock(:,3);
HI = stock(:,4);
BSI = stock(:,5);
MLR = stock(:,6);
LPI = stock(:,7);
y1 = lagmatrix(propcon,1);
y2 = lagmatrix(set50,1);
y3 = lagmatrix(CMI,1);
y4 = lagmatrix(HI,1);
y5 = lagmatrix(BSI,1);
y6 = lagmatrix(MLR,1);
y7 = lagmatrix(LPI,1);
x1 = int8((log(propcon)-log(y1))*100)
x2 = int8((log(set50)-log(y2))*100)
x3 = int8((log(CMI)-log(y3))*100)
x4 = int8((log(HI)-log(y4))*100)
x5 = int8((log(BSI)-log(y5))*100)
x6 = int8((log(MLR)-log(y6))*100)
x7 = int8((log(LPI)-log(y7))*100)
function LGObj = constructLGObj(OriginalSample);
OriginalSample = [0,0,0,0,0,0,0;...;-6,-6,0,-1,-4,-4,2];
LGObj.VarNumber = size( OriginalSample,2 );
LGObj.CaseLength = size( OriginalSample,1 );
LGObj.VarSample = OriginalSample;
[ LGObj.VarRange,LGObj.VarRangeLength ] = DimensionRangeValue(
OriginalSample,1 : LGObj.VarNumber );
function [ Dim,DimLength ] = DimensionRangeValue(OriginalData,Vec);
CountNumber = 0 ;
D=size(Vec,2);
if size(Vec,1)>1, Vec=Vec';
end;
DimLength = zeros( 1,D );
t=0;
for q=1 : D;
TempVector = unique( OriginalData(:,Vec(q))' );
if TempVector(1) == -1 ,TempVector(1)=[];
```

```

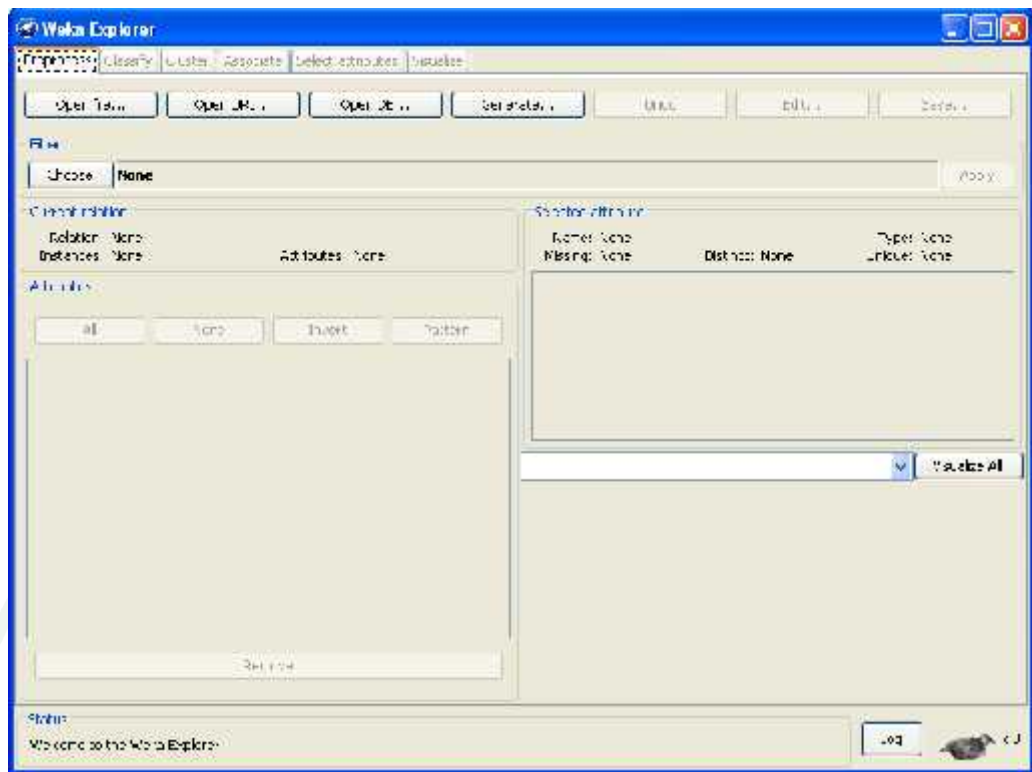
end;
RangeNumber = size(TempVector,2);
DimLength( q ) = RangeNumber;
t=t+1;
if CountNumber == 0;
    CountNumber = RangeNumber;
    Dim = zeros( D ,CountNumber );
    Dim(t,:) = TempVector ;
elseif CountNumber >= RangeNumber;
    Dim(t,:) = [TempVector, zeros(1,CountNumber - RangeNumber)];
elseif CountNumber < RangeNumber;
    Dim = [Dim, zeros( D,RangeNumber - CountNumber )];
    CountNumber = RangeNumber;
    Dim(t,:) = TempVector;
end;
end;
end
function GFunValue = gClosedFun( LG, X, PAX );
if size(PAX,1)>1;
    PAX = PAX';
end
GFunValue = 0;
LG = struct( LGObj );
N = LG.CaseLength;
X = [1,2,3,4,5,6,7];
PAX = [1,2,3,4,5,6,7];
UsedSample = DiscardNoneExist(LG.VarSample,[X PAX]);

DimX = LG.VarRangeLength( X );
RangeX = LG.VarRange( X,: );
ri = DimX;
OriginalData = LG.VarSample;
d = 1;
while d <= N;
    Frequency = zeros( 1,DimX );
    while d <= N && UsedSample( d ) == 1;
        d=d+1;
    end
    if d > N,break;end
    for t1 = 1:DimX;
        if RangeX(t1) == OriginalData( d,X ), break; end
    end
    Frequency( t1 ) = 1;
    UsedSample( d )= 1;
    ParentValue = OriginalData( d, PAX );
    d = d + 1;
    if d > N, break;end
    for k = d : N;
        if UsedSample( k )== 0 ;
            if ParentValue == OriginalData( k, PAX );
                t1 = 1;
                while RangeX( t1 )~=OriginalData( k,X ),t1=t1+1;
                    end
                Frequency( t1 ) = Frequency( t1 ) + 1;
                UsedSample( k ) = 1;
            end
        end
    end
end
Sum = sum( Frequency );

```



→



## 2. วิธีบ็อกซ์-เจนกินส์

กำหนดรูปแบบ การประมาณค่าพารามิเตอร์ และการตรวจสอบความเหมาะสมของตัวแบบ โดยใช้โปรแกรม SAS Version 9.3 มีรายละเอียดดังนี้

```

data propcon;
input y @@;
time = _n_;
datalines;
32.53
31.15
30.72
36.54
:
130.50
121.79
114.61
;
proc plot data = propcon;
plot y*time;
proc arima data = propcon;
identify var=y(1);
estimate printall plot;
forecast lead=6 out=fcast1;
run;

```

### 3. วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮสต์

กำหนดรูปแบบเริ่มต้นด้วยการวิเคราะห์รูปแบบการถดถอยโดยใช้โปรแกรม Microsoft Excel Version 2007 ดังนี้

**ตารางผนวกที่ ข1** การวิเคราะห์ความแปรปรวนสำหรับอนุกรมเวลาดัชนีราคาหุ้นกลุ่มอสังหาริมทรัพย์และก่อสร้าง

	Coefficients	Standard Error	T Stat	P-value
Intercept	37.0255	3.4317	10.7891	1.72E-15
X Variable	1.6513	0.0978	16.8770	4.78E-24

เมื่อได้รูปแบบเริ่มต้นและกำหนดค่าปรับให้เรียบที่ดีที่สุด คือ  $r = 1$  และ  $x = 0$  จากนั้นวิเคราะห์หารูปแบบโดยใช้โปรแกรม Microsoft Excel Version 2007 ดังตารางผนวกที่ ข2

**ตารางผนวกที่ ข2** การวิเคราะห์รูปแบบด้วยวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเชิงเส้นตรงของโฮสต์

Time	Y	Level	Growth Rate	Forecast
1	32.53	32.5294	1.6513	38.6768
2	31.15	31.1500	1.6513	34.1807
3	30.72	30.7200	1.6513	32.8013
4	36.54	36.5397	1.6513	32.3713
5	42.06	42.0601	1.6513	38.1910
6	45.13	45.1292	1.6513	43.7114
7	49.23	49.2309	1.6513	46.7805
8	54.32	54.3193	1.6513	50.8822
⋮	⋮	⋮	⋮	⋮
59	121.79	121.7900	1.6513	132.1513
60	114.61	114.6100	1.6513	123.4413

## ประวัติการศึกษาและการทำงาน

ชื่อ	นางสาวเมธินี กวินภาส
เกิดวันที่	12 กันยายน 2532
สถานที่เกิด	จังหวัดกรุงเทพมหานคร
ประวัติการศึกษา	วท.บ. (คณิตศาสตร์) จุฬาลงกรณ์มหาวิทยาลัย
ตำแหน่งปัจจุบัน	-
สถานที่ทำงานปัจจุบัน	-
ผลงานดีเด่นและ/หรือรางวัลทางวิชาการ	-
ทุนการศึกษาที่ได้รับ	ได้รับทุนผู้ช่วยสอนจากบัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์ (พ.ศ.2555 และ พ.ศ.2556)