



การตรวจจับค่านอกกลุ่มโดยใช้แผนภูมิควบคุมที่มีความเคร่ง

สำนักหอสมุดกลาง



โดย

นางสาวอรรรณ เหลืองสีเพชร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติประยุกต์

ภาควิชาสถิติ

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2556

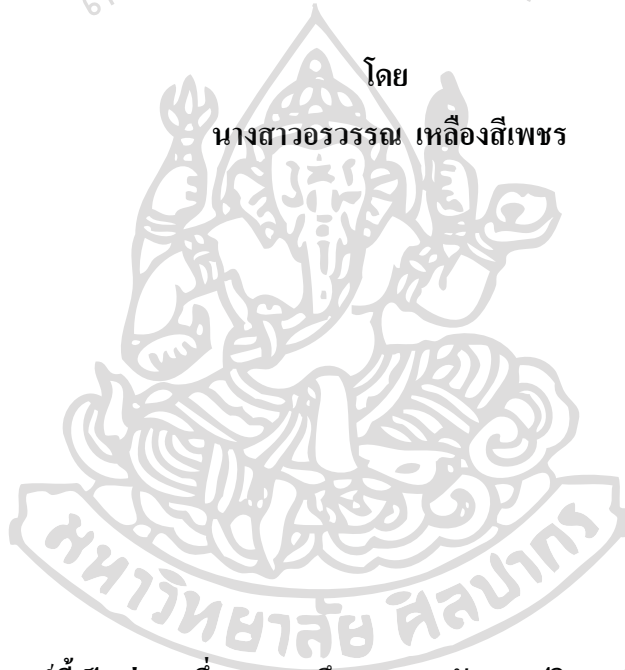
ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

การตรวจจับค่านอกกลุ่มโดยใช้แผนภูมิควบคุมที่มีความแกร่ง

สำนักหอสมุดกลาง

โดย

นางสาวอรรรณ เหลืองสีเพชร



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติประยุกต์

ภาควิชาสถิติ

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2556

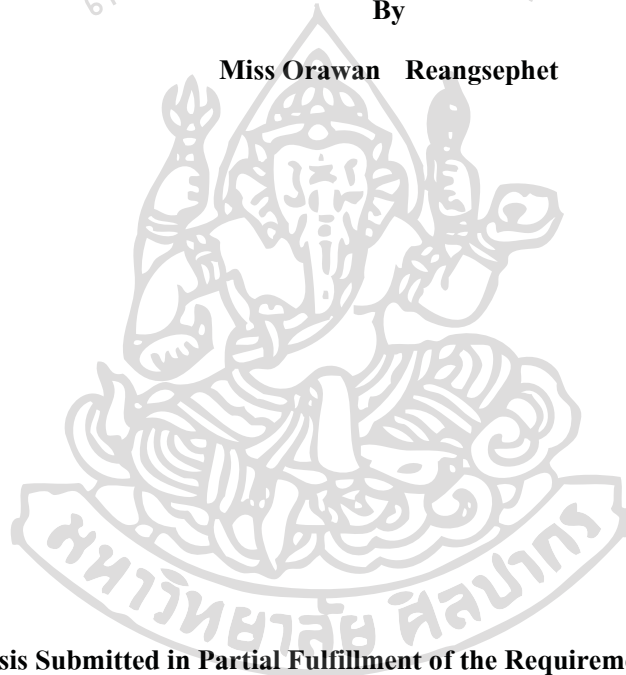
ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

OUTLIER DETECTION USING ROBUST CONTROL CHART

สำนักหอสมุดกลาง

By

Miss Orawan Reangsephet



A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

Master of Science Program in Applied Statistics

Department of Statistics

Graduate School, Silpakorn University

Academic Year 2013

Copyright of Graduate School, Silpakorn University

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง “ การตรวจจับค่านอก
กลุ่มโดยใช้แผนภูมิควบคุมที่มีความแกร่ง ” เสนอโดย นางสาวอรรพรรณ เหลืองสีเพชร เป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์

.....

(รองศาสตราจารย์ ดร.ปานใจ ชารัทสนวงศ์)

คณบดีบัณฑิตวิทยาลัย

วันที่.....เดือน..... พ.ศ.....

อาจารย์ที่ปรึกษาวิทยานิพนธ์

รองศาสตราจารย์ ดร.สุดา ตระการเถลิงศักดิ์

คณะกรรมการตรวจสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. กุศยา ปลั่งพงษ์พันธ์)

...../...../.....

..... ผู้ทรงคุณวุฒิ
(ผู้ช่วยศาสตราจารย์ ดร. ปราณิ นิลกรณ์)

...../...../.....

..... กรรมการ

(รองศาสตราจารย์ ดร.สุดา ตระการเถลิงศักดิ์)

...../...../.....

54304206 : สาขาวิชาสถิติประยุกต์

คำสำคัญ : แผนภูมิควบคุม/ค่านอกกลุ่ม/การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล

อรรถกร เหลืองสีเพชร : การตรวจจับค่านอกกลุ่มโดยใช้แผนภูมิควบคุมที่มีความ
แกร่ง. อาจารย์ที่ปรึกษาวิทยานิพนธ์ : รศ.ดร.สุดา ตระการเถลิงศักดิ์. 63 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) เสนอแผนภูมิควบคุมที่มีความแกร่งสำหรับตรวจจับค่านอกกลุ่ม
ในอนุกรมเวลาซึ่งใช้วิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่ง 2) เปรียบเทียบ
ประสิทธิภาพของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความ
แกร่งกับวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมาตรฐาน และ 3) ศึกษาผลกระทบของค่านอก
กลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับวิธีการพยากรณ์ที่มีความแกร่งและวิธีมาตรฐาน ใน
การศึกษาครั้งนี้ใช้การจำลองแบบตัวแบบอนุกรมเวลาซึ่ง โดยศึกษาตัวแบบ AR(1) ค่าพารามิเตอร์ ϕ เท่ากับ
0.1, 0.5 และ 0.9 ขนาดตัวอย่างเท่ากับ 100 และ 200 โดยแบ่งตัวอย่างออกเป็น 2 ส่วนเท่าๆกันคือ 50% ของค่า
สังเกตแรกเป็นตัวอย่างใน training period และ 50% ของค่าสังเกตถัดมาเป็นตัวอย่างใน test period ค่านอก
กลุ่มประเภท AO ด้วยขนาด 3, 5 และ 10 จำนวน 10% ของตัวอย่างใน training period และ/หรือ test period
โดยวัดประสิทธิภาพของแผนภูมิควบคุมจาก 1) ความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 2) กำลังการ
ทดสอบ และ 3) อัตราการเตือนผิด แต่ละพารามิเตอร์ทำซ้ำ 1,000 ครั้ง

ผลการวิจัยพบว่า

1. การวัดประสิทธิภาพ

กรณีไม่มีค่านอกกลุ่มเกิดขึ้นในอนุกรมเวลาทั้งใน training period และ test period แผนภูมิ
ควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานมีประสิทธิภาพดีกว่า แต่เมื่อเกิดค่านอกกลุ่มขึ้นในอนุกรมเวลาส่วน
training period และ/หรือ test period แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่งมี
ประสิทธิภาพดีกว่า

2. ผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ

กรณีไม่มีค่านอกกลุ่มเกิดขึ้นในอนุกรมเวลาวิธีพยากรณ์มาตรฐานและวิธีการพยากรณ์ที่มีความ
แกร่งมีค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ใกล้เคียงกัน แต่เมื่อเกิดค่านอกกลุ่มในอนุกรมเวลาค่าประมาณ
พารามิเตอร์ที่ปรับให้เรียบสำหรับวิธีการพยากรณ์มาตรฐานจะเกิดความเอนเอียงสู่ศูนย์ ในขณะที่ค่าประมาณ
พารามิเตอร์ที่ปรับให้เรียบสำหรับวิธีการพยากรณ์ที่มีความแกร่งมีค่าใกล้เคียงกัน

ภาควิชาสถิติ

ลายมือชื่อนักศึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2556

54304206 : MAJOR : APPLIED STATISTICS

KEY WORD : CONTROL CHART/OUTLIER/EXPONENTIAL SMOOTHING

ORAWAN REANGSEPHET : OUTLIER DETECTION USING ROBUST CONTROL CHART. THESIS ADVISOR : ASST. PROF. SUDA TAKANTASAK, Ph.D . 63 pp.

The objectives of this research are (1) To present a robust control chart for detection of outliers in stationary time series based on the robust exponential smoothing forecasts. (2) To compare the performance between control chart based on the robust exponential smoothing forecasts and control chart based on the standard exponential smoothing forecasts. (3) To study the outlier effect on the estimated smoothing parameter (λ) for the robust forecast and the standard forecast. A simulation study has been conducted on stationary model, the AR(1) with parameter ϕ 0.1, 0.5 and 0.9 . The sample size is set at 100 and 200 , including the training period of length 50% of the first generated observations and the next 50% of the generated observations constitute the test period. The magnitudes of additive outliers are set to 3, 5 and 10 and the percentage of outliers in the training sample and/or the test sample are taken as 0% (clean) and 10% . The three performance measures considered are 1) probability of type I error , 2) power of the test and 3) false alarm rate. Each combinations is carried out 1,000 replications.

The results of the study indicate that

1. The performance measures

For time series without outliers in training period and test period, the control chart based on the standard exponential smoothing forecast is outperform the robust version. However, for the presence of outliers in the training period and/or test period, the control chart based on the robust exponential smoothing performs better than the standard approach.

2. The outlier effect on the estimated smoothing parameter(λ)

For time series without outliers, the estimated smoothing parameter of both methods is quite similar. For time series with outliers, the estimated smoothing parameter for standard forecast is biased to zero. Contrast to the estimated smoothing parameter for robust forecast is stability

Department of Statistics

Graduate School, Silpakorn University

Student's signature

Academic Year 2013

Thesis Advisor's signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้เป็นอย่างดีเนื่องจากได้รับความอนุเคราะห์จากรองศาสตราจารย์ ดร. สุดา ตระการเถลิงศักดิ์ ซึ่งเป็นอาจารย์ที่ปรึกษาและควบคุมวิทยานิพนธ์ ที่กรุณาให้คำแนะนำ คำปรึกษา รวมทั้งช่วยตรวจทานแก้ไขข้อบกพร่องต่าง ๆ เป็นอย่างดียิ่งด้วยความเมตตาตลอดเวลาในการทำ วิทยานิพนธ์นี้ตั้งแต่เริ่มต้นจนสำเร็จเรียบร้อย ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างยิ่งในความกรุณาของท่าน ไว้ ณ ที่นี้ด้วย

นอกจากนี้ผู้วิจัยขอกราบขอบพระคุณผู้ช่วยศาสตราจารย์ ดร . กุศยา ปลั่งพงษ์พันธ์ ประธาน กรรมการตรวจสอบวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร . ปราณี นิลกรณ์ ผู้ทรงคุณวุฒิ สำหรับคำแนะนำ ตรวจสอบ และคำชี้แนะ ทำให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น ตลอดจนขอกราบขอบพระคุณคณาจารย์ ภาควิชาสถิติ มหาวิทยาลัยศิลปากรทุกท่านที่ช่วยประสิทธิ์ประสาทองค์ความรู้ ความช่วยเหลือ และ คำแนะนำต่างๆ มากมายในด้านการเรียนและคำสอนที่เป็นแนวทางในการดำเนินชีวิตต่อผู้วิจัยตลอด ระยะเวลาการศึกษา

กราบขอบพระคุณบิดา มารดา พี่ชาย และเสียที่เป็นแรงสนับสนุนทั้งกำลังทรัพย์ และกำลังใจ ให้กับผู้วิจัยมาตลอดการศึกษา จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จและสมบูรณ์

ขอบคุณพี่ๆ เพื่อนๆ และน้องๆ สาขาวิชาสถิติทุกท่านที่ให้ความช่วยเหลือและเป็นแรงผลักดัน ตลอดมา และขอบคุณพี่น้องลัทธิ เอียวเจริญ ที่ให้ความช่วยเหลือและอำนวยความสะดวกงานด้านเอกสาร แก่ผู้วิจัย

สารบัญ

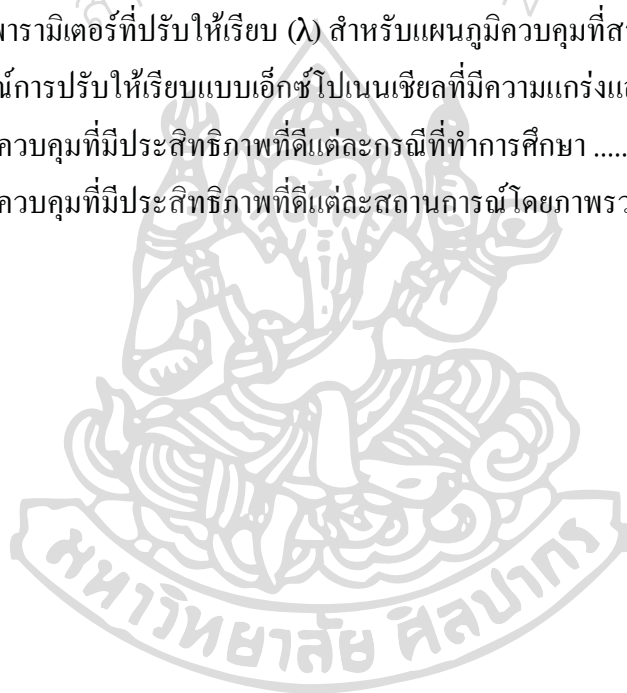
	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญตาราง	ฅ
สารบัญภาพ	ญ
บทที่	
1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
วัตถุประสงค์ของการศึกษา.....	2
ขอบเขตของการศึกษา.....	3
ขั้นตอนของงานวิจัย.....	4
ประโยชน์ที่คาดว่าจะได้รับ.....	5
นิยามศัพท์เฉพาะ.....	5
2 วรรณกรรมที่เกี่ยวข้อง.....	7
ทฤษฎีที่เกี่ยวข้อง.....	7
ตัวแบบอัตรดอดอยเฉลี่ยเคลื่อนที่.....	8
ค่านอกกลุ่ม.....	10
แผนภูมิควบคุม.....	12
งานวิจัยที่เกี่ยวข้อง.....	16
3 วิธีดำเนินการวิจัย.....	27
4 ผลการวิเคราะห์ข้อมูล.....	32
การวัดประสิทธิภาพของแผนภูมิควบคุม.....	32
ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1.....	32
กำลังการทดสอบ.....	34
อัตรการเตือนผิด.....	38
ผลกระทบของค่านอกกลุ่มที่มีต่อค่าพารามิเตอร์ที่ปรับให้เรียบ.....	40

บทที่	หน้า
5	
สรุป อภิปรายผล และข้อเสนอแนะ	44
การวัดประสิทธิภาพของแผนภูมิควบคุม.....	44
ผลกระทบของค่านอกกลุ่มที่มีต่อค่าพารามิเตอร์ที่ปรับให้เรียบ	47
อภิปรายผล	48
ข้อเสนอแนะ	50
รายการอ้างอิง	51
ภาคผนวก	53
ภาคผนวก ก	
โปรแกรมคอมพิวเตอร์ที่ใช้ในการจำลองและประมวลผลข้อมูล.....	54
ประวัติผู้วิจัย	63



สารบัญตาราง

ตารางที่		หน้า
1	ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน	33
2	กำลังการทดสอบสำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน	35
3	อัตราการเตือนผิดสำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน	39
4	ค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน	40
5	สรุปแผนภูมิควบคุมที่มีประสิทธิภาพที่ดีแต่ละกรณีที่ทำการศึกษา	45
6	สรุปแผนภูมิควบคุมที่มีประสิทธิภาพที่ดีแต่ละสถานการณ์โดยภาพรวม	47



สารบัญญภาพ

ภาพที่		หน้า
1	เปรียบเทียบค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแข็งแกร่งและวิธีมาตรฐานสำหรับขนาดตัวอย่างเท่ากับ 100 จำแนกตามขนาดของค่านอกกลุ่ม (w_1) และพารามิเตอร์สหสัมพันธ์ (ϕ).....	41
2	เปรียบเทียบค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแข็งแกร่งและวิธีมาตรฐานสำหรับขนาดตัวอย่างเท่ากับ 200 จำแนกตามขนาดของค่านอกกลุ่ม (w_1) และพารามิเตอร์สหสัมพันธ์ (ϕ).....	42
3	เปรียบเทียบค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแข็งแกร่งและวิธีมาตรฐานระหว่างขนาดตัวอย่างเท่ากับ 100 และ 200 จำแนกตามขนาดของค่านอกกลุ่ม (w_1) และพารามิเตอร์สหสัมพันธ์ (ϕ).....	43
4	สรุปผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแข็งแกร่งและวิธีมาตรฐาน.....	48

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การตรวจจับค่านอกกลุ่ม (outlier) มีความสำคัญอย่างมากในการวิเคราะห์อนุกรมเวลา เนื่องจากเมื่อค่านอกกลุ่มปรากฏขึ้นในอนุกรมเวลาจะส่งผลกระทบต่อวิเคราะห์ข้อมูลเป็นอย่างมาก อาจทำให้เกิดความยุ่งยากในการกำหนดตัวแบบหรือตัวแบบที่สร้างขึ้นไม่ถูกต้องเหมาะสม เกิดความเอนเอียงในการประมาณค่าพารามิเตอร์ของตัวแบบ ส่งผลให้ขาดความน่าเชื่อถือในการพยากรณ์อนุกรมเวลา

แผนภูมิควบคุม (Control Chart) เป็นเครื่องมือหนึ่งที่ใช้ในการตรวจจับการเปลี่ยนแปลงของข้อมูลหรือค่านอกกลุ่ม โดยอาศัยหลักการทางสถิติเพื่อให้สามารถจับความผิดปกติที่เกิดขึ้นได้อย่างทันท่วงที แผนภูมิควบคุมที่นิยมใช้สำหรับค่าค่าหนึ่งคือ Individual Control Chart หรือ EWMA Control Chart แต่เนื่องจากแผนภูมิควบคุมเหล่านี้มักมีข้อสมมติว่าข้อมูลจะต้องเป็นอิสระกัน ดังนั้นหากนำแผนภูมิควบคุมเหล่านี้ไปตรวจจับค่านอกกลุ่มในอนุกรมเวลาที่ข้อมูลมีสหสัมพันธ์กันจะส่งผลให้ประสิทธิภาพของแผนภูมิควบคุมลดลง เพื่อแก้ปัญหากรณีข้อมูลมีสหสัมพันธ์ในตัวเอง (Autocorrelation) ในงานวิจัยนี้จึงนำเสนอแผนภูมิควบคุมที่จัดอยู่ในประเภทของแผนภูมิควบคุมสาเหตุเฉพาะ (Special - Cause Control Chart : SCC) ที่เสนอโดย Alwan และ Robert (1988) ซึ่งสร้างแผนภูมิควบคุมจากความคลาดเคลื่อนในการพยากรณ์ (forecast error) และตรวจจับความผิดปกติจากค่าความคลาดเคลื่อนในการพยากรณ์ เนื่องจากในการวิเคราะห์อนุกรมเวลา (Time series Analysis) มีข้อสมมติว่าความคลาดเคลื่อนที่ได้จากการพยากรณ์เป็นอิสระกัน โดยถ้าค่าสังเกตมีค่าแตกต่างจากค่าพยากรณ์มากจะส่งผลให้ค่าความคลาดเคลื่อนมีค่ามากหรือน้อยเกินกว่าที่ควรจะเป็น ค่าสังเกตนั้นจะได้รับการตรวจจับและถูกจัดเป็นค่านอกกลุ่ม (Outliers) ดังนั้นสิ่งที่จำเป็นอีกอย่างหนึ่งในการสร้างแผนภูมิควบคุมที่มีประสิทธิภาพและน่าเชื่อถือคือวิธีการพยากรณ์ข้อมูลที่ต้องแม่นยำและมีความน่าเชื่อถือ

วิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing) เป็นวิธีการพยากรณ์อย่างง่าย (Naïve prediction method) ที่นิยมใช้ทั่วไปและมีประสิทธิภาพในการพยากรณ์อนุกรมเวลาบางที่ แต่ข้อเสียของวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลคือเป็นวิธีที่มีความไวต่อค่านอกกลุ่มเนื่องจากค่าพยากรณ์ล่วงหน้าจะขึ้นอยู่กับค่าสังเกตในปัจจุบันและค่าพยากรณ์ในอดีต เมื่อค่านอก

กลุ่มเกิดขึ้นจึงส่งผลกระทบต่อค่าพยากรณ์ลำดับถัด ๆ มา ถ้าค่าสังเกตในปัจจุบันเป็นค่านอกกลุ่มจะส่งผลกระทบต่อให้ค่าพยากรณ์ได้รับอิทธิพลจากค่านอกกลุ่มทำให้เกิดความคลาดเคลื่อนในการพยากรณ์เพิ่มมากขึ้น นอกจากนี้การเกิดค่านอกกลุ่มยังส่งผลให้การเลือกพารามิเตอร์ที่ปรับให้เรียบเอนเอียง เหตุนี้เนื่องจากค่าพารามิเตอร์ที่ปรับให้เรียบจะเลือกจากค่าที่ทำให้ความคลาดเคลื่อนในการพยากรณ์มีค่าต่ำที่สุด ยิ่งไปกว่านั้นถ้านำความคลาดเคลื่อนจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลไปใช้ในการสร้างแผนภูมิควบคุมจะส่งผลให้แผนภูมิควบคุมที่สร้างขึ้นไม่มีความน่าเชื่อถือและอาจเกิดลำดับของการเตือนผิด (sequence of false alarms) หรือตรวจจับค่านอกกลุ่มไม่พบก็ได้

ในงานวิจัยนี้ได้เสนอแผนภูมิควบคุมสาเหตุเฉพาะที่มีความแกร่งโดยใช้วิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลตามที่ Gelper , Fried และ Croux (2010) ได้เสนอไว้ซึ่งเป็นวิธีการพยากรณ์ที่มีความแกร่งต่อค่านอกกลุ่ม โดยใช้แนวคิดการแทนที่อนุกรมเวลา ที่มีค่านอกกลุ่มด้วยอนุกรมเวลาที่ไม่มีค่านอกกลุ่ม (Clean series) และทำการเปรียบเทียบประสิทธิภาพของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งกับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมาตรฐานในการตรวจจับค่านอกกลุ่มในอนุกรมเวลาคงที่ และศึกษาผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบสำหรับวิธีการพยากรณ์ของแผนภูมิควบคุมทั้งสองแบบ

วัตถุประสงค์ของการศึกษา

1. เพื่อเสนอแผนภูมิควบคุมที่มีความแกร่งและมีประสิทธิภาพในการตรวจจับค่านอกกลุ่มในอนุกรมเวลาคงที่
2. เพื่อเปรียบเทียบประสิทธิภาพของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งกับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมาตรฐาน
3. เพื่อศึกษาผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบสำหรับวิธีการพยากรณ์ของแผนภูมิควบคุมทั้งสองแบบ

ขอบเขตของการศึกษา

ในงานวิจัยนี้จะเสนอแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแข็งแกร่ง เปรียบเทียบประสิทธิภาพในการตรวจจับค่านอกกลุ่มระหว่างแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแข็งแกร่งกับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมาตรฐาน และศึกษาผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบสำหรับวิธีการพยากรณ์ของแผนภูมิควบคุมทั้งสองแบบ โดยกำหนดขอบเขตของการศึกษาครั้งนี้ดังนี้

1. อนุกรมเวลา z_t เป็นอนุกรมเวลาคงที่ (Stationary)
2. ตัวแบบของอนุกรมเวลาที่ศึกษาคือ AR(1) กรณีค่าคงที่ในอนุกรมเวลาเท่ากับ 0 นั่นคือ $z_t = \phi z_{t-1} + \varepsilon_t$ โดยกำหนดค่าพารามิเตอร์ $\phi = 0.1, 0.5$ และ 0.9 เพื่อศึกษากรณีที่อนุกรมเวลามีสหสัมพันธ์ในตัวเองระดับน้อย ปานกลาง และมาก
3. ความคลาดเคลื่อนสุ่ม ε_t มีข้อสมมติว่าเป็นอิสระกันและมีการแจกแจงแบบปกติมาตรฐาน นั่นคือ $\varepsilon_t \sim \text{i.i.d. } N(0,1)$
4. ขนาดตัวอย่างที่ใช้ในการศึกษาคือ $T=100$ และ 200 โดยแบ่งตัวอย่างออกเป็น 2 ช่วงคือ training period และ test period โดย
 - 4.1 ขนาดตัวอย่าง training period คือ $n = 50, 100$ ตามลำดับ โดยตัวอย่าง 10 ตัวอย่างแรกจะใช้เป็น start period เพื่อใช้ในการกำหนดค่าเริ่มต้น
 - 4.2 ขนาดตัวอย่าง test period คือ 50 และ 100 ตามลำดับ
5. ค่านอกกลุ่มที่ทำการศึกษาเป็นค่านอกกลุ่มประเภท AO คือค่านอกกลุ่มประเภทที่ส่งผลกระทบต่อค่าสังเกตเพียงค่าเดียวเท่านั้น
6. ขนาด(magnitudes)ของค่านอกกลุ่มที่ใช้ในการศึกษาคือ 3, 5 และ 10 ซึ่งให้ค่านอกกลุ่มมีขนาดเป็น 3 เท่า 5 เท่า และ 10 เท่าของส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน เพื่อศึกษากรณีค่านอกกลุ่มขนาดเล็ก ปานกลาง และใหญ่
7. กำหนดระดับนัยสำคัญ $\alpha = 0.05$
8. จำนวนค่านอกกลุ่มคือ 10% ของจำนวนตัวอย่างในช่วงเวลา training period หรือ test period

9. ตำแหน่งของค่านอกกลุ่มเป็นแบบสุ่ม

10. การเปรียบเทียบประสิทธิภาพของแผนภูมิควบคุมวัดโดยพิจารณาจากความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 (Probability of type I error) กำลังการทดสอบ (power of the test) และ อัตราการเตือนผิด (false alarm rate)

ขั้นตอนของงานวิจัย

ขั้นตอนในการดำเนินงานวิจัยมีดังนี้

1. จำลองข้อมูลอนุกรมเวลาโดยใช้ตัวแบบคือ AR(1) ด้วยพารามิเตอร์และขนาดของอนุกรมเวลาตามที่กำหนด
2. แบ่งข้อมูลออกเป็น 2 ช่วงเวลาคือ training period และ test period โดยตัวอย่าง 10 ตัวอย่างแรกจะใช้เป็น start period
3. สร้างค่านอกกลุ่มจำนวน 10% ของจำนวนตัวอย่างในช่วงเวลา training period และ/หรือ test period ที่มีขนาดคือ 3, 5 และ 10
4. สร้างแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งกับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมาตรฐาน โดยพิจารณาอนุกรมเวลาในช่วง training period
5. นำแผนภูมิควบคุมที่สร้างขึ้นจากทั้งสองวิธีมาตรวจจับค่านอกกลุ่มในอนุกรมเวลาช่วง test period
6. ทำซ้ำขั้นตอนที่ 1- 5 ในแต่ละพารามิเตอร์ที่กำหนดจำนวน 1,000 ครั้ง
7. กำหนดค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 (Probability of type I error) กำลังการทดสอบ (power of the test) และอัตราการเตือนผิด (false alarm rate)
8. เปรียบเทียบประสิทธิภาพของแผนภูมิควบคุมทั้งสอง
9. วิเคราะห์และรายงานผล

ประโยชน์ที่คาดว่าจะได้รับ

แผนภูมิควบคุมที่สร้างขึ้นมีความแกร่งและมีประสิทธิภาพในการตรวจจับค่านอกกลุ่ม และเป็นวิธีที่ง่ายและไม่ซับซ้อนเมื่อนำไปใช้งาน

นิยามศัพท์เฉพาะ

1. training period คือส่วนของอนุกรมเวลาที่ใช้ในการกำหนดขีดจำกัดควบคุม
2. test period คือส่วนของอนุกรมเวลาที่ใช้ในการตรวจจับค่านอกกลุ่มและวัดประสิทธิภาพของแผนภูมิควบคุม
3. ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 (Probability of type I error) คือความน่าจะเป็นที่ค่าสังเกตใน test period จะตกออกนอกขีดจำกัดควบคุมเมื่อกระบวนการอยู่ในการควบคุม (in control) โดยที่

$$\text{ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1} = \sum \left(\frac{\text{จำนวนตัวอย่างที่ถูกตรวจจับเป็นค่านอกกลุ่มในกรณี test period ไม่มีค่านอกกลุ่ม}}{\text{จำนวนตัวอย่างทั้งหมดใน test period}} \right) / 1000$$

4. กำลังการทดสอบ (power of the test) คือความน่าจะเป็นที่ค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับ นั่นคือค่านอกกลุ่มที่สร้างขึ้นตกออกนอกขีดจำกัดควบคุม โดยที่

$$\text{กำลังการทดสอบ} = \sum \left(\frac{\text{จำนวนค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับได้}}{\text{จำนวนค่านอกกลุ่มที่สร้างขึ้นทั้งหมดใน test period}} \right) / 1000$$

5. อัตราการเตือนผิด (false alarm rate) คือร้อยละของจำนวนตัวอย่างใน test period ที่ไม่ได้เป็นค่านอกกลุ่มที่สร้างขึ้นแต่ตกออกนอกขีดจำกัดควบคุมและชี้ว่าเป็นค่านอกกลุ่มเมื่อกระบวนการไม่ได้อยู่ในการควบคุม (out-of-control)

$$\text{อัตราการเตือนผิด} = \sum \left(\frac{\text{จำนวนตัวอย่างที่ไม่ได้เป็นค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับได้}}{\text{จำนวนตัวอย่างทั้งหมดใน test period ที่ไม่ได้เป็นค่านอกกลุ่มที่สร้างขึ้น}} \right) / 1000$$

6. ขนาดของค่านอกกลุ่ม (w_1, w_2) คือค่าสังเกตที่เป็นค่านอกกลุ่มจะมีค่าเพิ่มขึ้นเป็น w_1, w_2 เท่าของส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนตามลำดับ

โดยที่ w_1 แทนขนาดของค่าออกกลุ่มที่เกิดขึ้นใน training period
 w_2 แทนขนาดของค่าออกกลุ่มที่เกิดขึ้นใน test period

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

อนุกรมเวลา (Time series) คือเซตของค่าสังเกต (observation) หรือข้อมูล (data) ที่จัดเก็บรวบรวมตามลำดับเวลา เช่นราคาหุ้นรายสัปดาห์ ยอดขายรถยนต์รายเดือน ปริมาณการส่งออกสินค้าประเภทสิ่งทอและเครื่องนุ่งห่มรายปี เป็นต้น ลักษณะของอนุกรมเวลาโดยทั่วไปประกอบด้วย 4 องค์ประกอบหลักคือ

1. แนวโน้ม (Trend) เป็นการเปลี่ยนแปลงของข้อมูลที่อาจเพิ่มขึ้นหรือลดลง ลักษณะการเปลี่ยนแปลงอาจจะเป็นเชิงเส้น เส้นโค้ง พาราโบลา หรือเส้นโค้งเอ็กซ์โปเนนเชียล (Exponential) ก็เป็นไปได้ ซึ่งเป็นการเปลี่ยนแปลงในระยะเวลายาว ๆ มากกว่า 1 ปี

2. ฤดูกาล (Seasonal) เป็นการเปลี่ยนแปลงของข้อมูลที่มีลักษณะขึ้นลงซ้ำ ๆ กันเมื่อถึงเวลาเดิมในฤดูกาลหนึ่ง ๆ อาจจะเป็นรายวัน รายสัปดาห์ รายไตรมาส หรือรายเดือนก็ได้ การเปลี่ยนแปลงลักษณะนี้จะเป็นการเปลี่ยนแปลงในระยะเวลาดสั้น ๆ ไม่เกิน 1 ปี

3. วัฏจักร (Cycle) เป็นการเปลี่ยนแปลงของข้อมูลที่มีลักษณะขึ้นลงของข้อมูลเคลื่อนที่ซ้ำ ๆ กันคล้ายฤดูกาล แต่เป็นการเคลื่อนที่อย่างช้า ๆ ใช้เวลานานหลายปีในการเปลี่ยนแปลง

4. เหตุการณ์ที่ผิดปกติ (Irregular variation) เป็นการเปลี่ยนแปลงที่เกิดจากปัจจัยอื่น ๆ นอกเหนือจาก แนวโน้ม ฤดูกาล และวัฏจักร ซึ่งเป็นเหตุการณ์ที่ไม่สามารถคาดคะเนได้ล่วงหน้าและไม่ได้เกิดขึ้นบ่อยเช่นแผ่นดินไหว การปฏิวัติ คลื่นสึนามิ หรือการปิดโรงงาน เป็นต้น

การวิเคราะห์อนุกรมเวลา (Time Series Analysis) เป็นวิธีการทางสถิติที่ใช้ศึกษาลักษณะการเปลี่ยนแปลงของอนุกรมเวลา กำหนดตัวแบบเพื่อใช้ในการพยากรณ์หรือทำนายค่าของอนุกรมเวลาในอนาคต ตัวแบบที่นิยมใช้อย่างมากในการวิเคราะห์อนุกรมเวลาคือตัวแบบอัตถถดถอยเคลื่อนที่ (Autoregressive – Moving average model : ARMA) ซึ่งเป็นการวิเคราะห์อนุกรมเวลาแบบบอซซ์ -

เจนคินส์ (Box-Jenkins Analysis) อนุกรมเวลาที่จะนำมาใช้ในการวิเคราะห์อนุกรมเวลาโดยวิธีวิธีบ็อกซ์-เจนคินส์จะต้องเป็นอนุกรมเวลาคงที่ (Stationary)

อนุกรมเวลาคงที่ (Stationary) คืออนุกรมเวลาที่มีคุณสมบัติทางสถิติไม่เปลี่ยนแปลงเมื่อเวลาเปลี่ยนแปลงไป นั่นคือการแจกแจงความน่าจะเป็นร่วมของค่าสังเกต m ค่า $z_{t_1}, z_{t_2}, z_{t_3}, \dots, z_{t_m}$ ที่เกิดขึ้น ณ เวลา $t_1, t_2, t_3, \dots, t_m$ จะมีการแจกแจงเหมือนกับการแจกแจงความน่าจะเป็นร่วมของค่าสังเกต m ค่า $z_{t_1+s}, z_{t_2+s}, z_{t_3+s}, \dots, z_{t_m+s}$ ที่เกิดขึ้น ณ เวลา $t_1 + s, t_2 + s, t_3 + s, \dots, t_m + s$ แต่ในทางปฏิบัติเป็นไปได้ยากที่จะทราบการแจกแจงความน่าจะเป็นร่วมของค่าสังเกตในอนุกรมเวลา ดังนั้นในการวิเคราะห์อนุกรมเวลาจึงนิยมใช้คุณสมบัติของการคงที่อย่างอ่อน (Weak Stationarity) คืออนุกรมเวลา $z_t, t=1, 2, \dots, n$ เป็นอนุกรมเวลาคงที่ ถ้ามีคุณสมบัติทั้ง 3 ข้อดังนี้

1. ค่าเฉลี่ยคงที่ นั่นคือ $E(z_t) = \mu, t=1, 2, \dots, n$
2. ความแปรปรวนคงที่ นั่นคือ $V(z_t) = \sigma^2, t=1, 2, \dots, n$
3. ความแปรปรวนร่วม (covariance) คงที่ไม่ขึ้นอยู่กับเวลา t แต่ขึ้นอยู่กัช่วงห่างของเวลา k

(หรือเรียกว่า lag k) นั่นคือ $\text{cov}(z_t, z_{t+k}) = \rho_k$ ทุกค่า k

อนุกรมเวลาที่ไม่สอดคล้องกับคุณสมบัติที่กล่าวมาข้างต้นนี้จะเรียกอนุกรมเวลานี้ว่าอนุกรมเวลาไม่คงที่ (Non-stationary) ในกรณีอนุกรมเวลาไม่คงที่สามารถทำให้เป็นอนุกรมเวลาคงที่ได้ สำหรับกรณีอนุกรมเวลามีค่าเฉลี่ยไม่คงที่หรืออนุกรมได้รับอิทธิพลจากแนวโน้มหรือฤดูกาลสามารถทำเป็นอนุกรมเวลาที่คงที่ได้โดยการหาผลต่าง ในกรณีอนุกรมเวลาที่มีความแปรปรวนไม่คงที่สามารถปรับให้เป็นอนุกรมเวลาที่มีความแปรปรวนคงที่โดยการใส่ค่าลอการิทึม (Natural logarithm) ของอนุกรมเวลา นั่นคืออยู่ในรูป $\ln(z_t)$ แทนการใส่อนุกรมเวลา z_t ถ้าการแปลงข้อมูลโดยใช้ค่าลอการิทึมยังไม่ทำให้อนุกรมเวลาที่อยู่ในรูปลอการิทึมมีความแปรปรวนคงที่ ก็อาจใช้ Power Transformation อื่นแทน สำหรับกรณีที่อนุกรมเวลามีค่าเฉลี่ยและความแปรปรวนไม่คงที่ จะต้องทำการแปลงอนุกรมเวลาให้มีความแปรปรวนคงที่ก่อนแล้วจึงปรับให้อนุกรมเวลามีค่าเฉลี่ยที่คงที่โดยการหาผลต่างที่เหมาะสม

ตัวแบบอัตถดถอยเฉลี่ยเคลื่อนที่ (Autoregressive – Moving average model)

ตัวแบบอัตถดถอยเฉลี่ยเคลื่อนที่ที่ใช้สัญลักษณ์ว่า ARMA ตัวแบบอัตถดถอยเฉลี่ยเคลื่อนที่อันดับ p, q หรือ ARMA(p, q) สามารถเขียนได้ดังนี้

$$\phi_p(B)z_t = \delta + \theta_q(B)\varepsilon_t$$

$$z_t = \delta + \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

หรือเขียนอยู่ในรูป

$$z_t = \frac{\delta + \theta_q(B)\varepsilon_t}{\phi_p(B)}$$

โดยที่ δ คือค่าคงที่

z_t คือค่าสังเกตของอนุกรมเวลา ณ เวลา t

$\phi_p(B)$ คือ Autoregressive operator of order p ซึ่ง

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

$\theta_q(B)$ คือ Moving average operator of order q ซึ่ง

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

p คืออันดับของตัวแบบอัตถถดถอย (Autoregressive model)

q คืออันดับของตัวแบบเฉลี่ยเคลื่อนที่ (Moving average model)

B คือ Backshift operator นั่นคือ $B^j z_t = z_{t-j}$

และ ε_t คือความคลาดเคลื่อนสุ่มหรือเรียกว่า White Noise process โดยมีข้อสมมติเบื้องต้นว่าความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติ ค่าเฉลี่ยเท่ากับศูนย์ ความแปรปรวนคงที่ และเป็นอิสระกัน นั่นคือ $\varepsilon_t \sim i.i.d. N(0, \sigma_\varepsilon^2)$

สำหรับตัวแบบ ARMA ที่มีอันดับของตัวแบบเฉลี่ยเคลื่อนที่ q เท่ากับ 0 จะเป็นตัวแบบอัตถถดถอยซึ่งเป็นตัวแบบที่ค่าสังเกตปัจจุบันของอนุกรมเวลาจะขึ้นอยู่กับค่าสังเกตในอดีตกับความคลาดเคลื่อนปัจจุบัน รูปแบบทั่วไปของตัวแบบอัตถถดถอยอันดับ p (Autoregression model order p) ใช้สัญลักษณ์แทนว่า AR(p) คือ

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + \varepsilon_t$$

โดยที่ δ คือค่าคงที่

ϕ_i คือพารามิเตอร์อัตถถดถอยอันดับที่ i , $i=1,2,\dots,p$

และ ε_t คือความคลาดเคลื่อนสุ่ม โดยที่ $\varepsilon_t \sim i.i.d. N(0, \sigma_\varepsilon^2)$

หรือเขียนในรูป Backshift operator ได้ดังนี้

$$\phi_p(B)z_t = \delta + \varepsilon_t \quad \text{โดยที่} \quad \phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

สำหรับตัวแบบ ARMA ที่มีอันดับของตัวแบบอัตโนมัติ p เท่ากับ 0 จะเป็นตัวแบบค่าเฉลี่ยเคลื่อนที่ ซึ่งเป็นตัวแบบที่ค่าสังเกตปัจจุบันของอนุกรมเวลาจะขึ้นอยู่กับความคลาดเคลื่อนปัจจุบันและความคลาดเคลื่อนในอดีต รูปแบบทั่วไปของตัวแบบค่าเฉลี่ยเคลื่อนที่อันดับ q (Moving average model order q) ใช้สัญลักษณ์แทนว่า MA(q) คือ

$$z_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

โดยที่ δ คือค่าคงที่

θ_i คือพารามิเตอร์ค่าเฉลี่ยเคลื่อนที่อันดับที่ i , $i=1,2,\dots,q$

และ ε_t คือความคลาดเคลื่อนสุ่ม โดยที่ $\varepsilon_t \sim \text{i.i.d. } N(0, \sigma_\varepsilon^2)$

หรือเขียนในรูป Backshift operator ได้ดังนี้

$$z_t = \delta + \theta_q(B)\varepsilon_t \quad \text{โดยที่} \quad \theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

นอกจากการวิเคราะห์อนุกรมเวลา ARMA แล้ววิธีที่นิยมใช้ทั่วไปสำหรับการพยากรณ์อนุกรมเวลาคงที่คือวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing) และวิธีที่นิยมใช้ทั่วไปสำหรับการพยากรณ์อนุกรมเวลาที่มีแนวโน้มและฤดูกาลคือวิธีการพยากรณ์ Holt - Winter ซึ่งทั้งสองวิธีนี้เป็นวิธีการพยากรณ์อย่างง่าย (Naïve prediction method) และมีประสิทธิภาพในการพยากรณ์และเป็นวิธีการพยากรณ์ที่ไม่ต้องการการกำหนดตัวแบบ

ค่าผิดปกติ (Outlier)

ค่าผิดปกติ (Outlier) คือค่าสังเกตที่มีค่าแตกต่างจากค่าสังเกตอื่น ๆ อย่างมากในอนุกรมเวลานั้นๆ ซึ่งอาจได้รับอิทธิพลต่างๆ ทำให้ค่าสังเกตมีค่าสูงกว่าปกติหรือต่ำกว่าปกติ เช่นการบันทึกข้อมูลผิดพลาด อุปกรณ์ที่ใช้ในการผลิตชำรุด หรือเกิดสถานการณ์ที่ผิดปกติต่างๆ เป็นต้น การปรากฏของค่าผิดปกติในอนุกรมเวลาส่งผลกระทบต่อการวิเคราะห์ข้อมูลเป็นอย่างมาก เช่นอาจสร้างความยุ่งยากในการกำหนดตัวแบบ หรือตัวแบบที่สร้างขึ้นไม่ถูกต้อง เกิดความเอนเอียงในการประมาณค่าพารามิเตอร์ของตัวแบบทำให้ขาดความน่าเชื่อถือในการพยากรณ์ ดังนั้นการตรวจจับค่าผิดปกติจึงมีความสำคัญอย่างมากในการวิเคราะห์อนุกรมเวลา

Fox (1972) ได้แบ่งค่าผิดปกติที่ปรากฏในอนุกรมเวลาออกเป็น 2 ประเภท คือ Type I outliers และ Type II outliers

Type I outliers คือค่านอกกลุ่มประเภทที่ส่งผลกระทบต่อค่าสังเกตเพียงค่าเดียวเท่านั้น นั่นคือ ณ เวลาที่เกิดค่านอกกลุ่มประเภทนี้จะส่งผลทำให้ค่าสังเกตของอนุกรมเวลา ณ เวลานั้นมีการเปลี่ยนแปลงอย่างมากเพียงค่าเดียวเท่านั้น หลังจากนั้นอนุกรมเวลาจะกลับสู่สภาวะปกติ ตัวแบบของอนุกรมเวลาที่มีค่านอกกลุ่มประเภทนี้ปรากฏอยู่คือ

$$x_t = \begin{cases} z_t & \text{เมื่อ } t \neq T' \\ z_t + w_I & \text{เมื่อ } t = T' \end{cases}$$

โดยที่ z_t คือค่าสังเกตของตัวแบบ ARMA(p,q)

w_I คือขนาด (magnitude) ของ Type I outliers

T' คือเวลาที่เกิด Type I outliers

หรือสามารถเขียนเป็น

$$x_t = z_t + w_I I_t^{(T')}$$

โดยที่ $I_t^{(T')}$ คือฟังก์ชันตัวชี้ (Indicator function)

$$I_t^{(T')} = \begin{cases} 1 & \text{เมื่อ } t = T' \\ 0 & \text{เมื่อ } t \neq T' \end{cases}$$

Type II outliers คือค่านอกกลุ่มประเภทที่นอกจากจะส่งผลกระทบต่ออนุกรมเวลา ณ ตำแหน่งเวลาที่เกิดค่านอกกลุ่มแล้วยังส่งผลกระทบต่อค่าสังเกตอื่น ๆ ถัดมาอีกด้วย นั่นคือค่านอกกลุ่มประเภทนี้ จะส่งผลกระทบต่อค่าสังเกตที่เวลา $T', T'+1, T'+2, \dots$ เมื่อ T' คือตำแหน่งเวลาที่เกิดค่านอกกลุ่ม ตัวแบบอนุกรมเวลาที่มีค่านอกกลุ่มประเภทนี้ปรากฏอยู่คือ

$$x_t = z_t + \frac{\theta(B)}{\phi(B)} w_{II} I_t^{(T')}$$

โดยที่ w_{II} คือขนาด (magnitude) ของ Type II outliers

$I_t^{(T')}$ คือฟังก์ชันตัวชี้ (Indicator function)

$$I_t^{(T')} = \begin{cases} 1 & \text{เมื่อ } t = T' \\ 0 & \text{เมื่อ } t \neq T' \end{cases}$$

ต่อมา Tsay (1986) ได้ให้ชื่อกับค่าผิดปกติที่ Fox (1972) ได้แบ่งประเภทไว้ใหม่คือ Additive outlier (AO) แทน Type I outliers และ Innovational outlier (IO) แทน Type II outliers และได้จัดประเภทของค่าผิดปกติเพิ่มขึ้นจากที่ Fox (1972) ได้จัดไว้จากเดิมอีก 3 ประเภทคือ

Level shift (LS) คือ ค่าผิดปกติประเภทที่ทำให้ค่าเฉลี่ยของอนุกรมเวลามีการเปลี่ยนแปลงไปอย่างถาวร โดยส่งผลกระทบต่อค่าสังเกต ณ เวลา T' และค่าสังเกตอื่น ๆ ถัดมา เมื่อ T' คือตำแหน่งเวลาที่เกิดค่าผิดปกติ ตัวแบบของอนุกรมเวลาที่มีค่าผิดปกติประเภท LS ปรากฏอยู่คือ

$$x_t = z_t + \frac{1}{(1-B)} w_{LS} I_t^{(T')}$$

โดยที่ w_{LS} คือขนาด (magnitude) ของค่าผิดปกติประเภท LS

Temporary change (TC) คือค่าผิดปกติที่มีลักษณะคล้ายกับค่าผิดปกติประเภท AO แต่ทว่าผลกระทบนั้นยังส่งผลกระทบต่อค่าสังเกตอื่น ๆ ถัดมาอีกด้วยคล้ายค่าผิดปกติประเภท LS แต่ผลกระทบของค่าผิดปกติประเภทนี้จะไม่คงอยู่ถาวรซึ่งจะมีลักษณะค่อยๆ ลดลงแบบเอ็กซ์โปเนนเชียล (exponential) ด้วยอัตรา (rate) δ ตัวแบบของอนุกรมเวลาที่มีค่าผิดปกติประเภท TC ปรากฏอยู่คือ

$$x_t = z_t + \frac{1}{(1-\delta B)} w_{TC} I_t^{(T')}$$

โดยที่ w_{TC} คือขนาด (magnitude) ของค่าผิดปกติประเภท TC

δ คือค่าคงที่ เมื่อ $0 < \delta < 1$

Variance change (VC) คือค่าผิดปกติที่ส่งผลกระทบต่อค่าความแปรปรวนของค่าสังเกตในข้อมูลอนุกรมเวลา ณ เวลา T' ซึ่งเปลี่ยนแปลงจาก σ_ϵ^2 ไปเป็น $(1 + w_{VC})\sigma_\epsilon^2$ ตัวแบบของอนุกรมเวลาที่มีค่าผิดปกติประเภท VC ปรากฏอยู่คือ

$$x_t = z_t + \frac{\theta(B)}{\phi(B)} w_{VC} \epsilon_t$$

โดยที่ w_{VC} คือขนาด (magnitude) ของค่าผิดปกติประเภท VC

แผนภูมิควบคุม (Control Chart)

แผนภูมิควบคุมเป็นเครื่องมือในการตรวจสอบและเฝ้าระวังการเปลี่ยนแปลงของกระบวนการที่อาศัยหลักการทางสถิติเพื่อให้สามารถจับความผิดปกติได้อย่างทันที่ ก่อนที่ความผิดปกตินั้นจะส่งผลกระทบต่อความเสียหายของกระบวนการ นอกจากนี้แผนภูมิควบคุมจะทำหน้าที่ในการตรวจสอบและเฝ้าระวังการเปลี่ยนแปลงกระบวนการผลิตแล้วยังสามารถให้รายละเอียดในส่วนอื่น ๆ ด้วย เช่น การประเมินค่าเฉลี่ย

ของคุณภาพ การประเมินการกระจายของคุณภาพ ความสามารถของกระบวนการ และช่วยป้องกันความเสียหายอันเกิดจากการผิดปกติของกระบวนการผลิตได้อีกด้วย

ประเภทของแผนภูมิควบคุมสามารถจำแนกตามลักษณะของข้อมูลออกเป็น 3 ประเภทหลัก ๆ ดังนี้

1. แผนภูมิควบคุมสำหรับตัวแปร (Control Chart for Variable) เป็นแผนภูมิควบคุมที่ใช้เพื่อควบคุมกระบวนการผลิตสำหรับลักษณะคุณภาพที่วัดค่าเป็นค่าต่อเนื่อง เช่น เส้นผ่านศูนย์กลางของลูกสูบ เครื่องยนต์ ปริมาตรเครื่องคั้นน้ำอัดลม หรืออุณหภูมิของตู้เย็น เป็นต้น ซึ่งคุณลักษณะคุณภาพชนิดนี้เรียกว่าตัวแปร (variable) แผนภูมิประเภทนี้ที่นิยมกันมากที่สุด คือแผนภูมิควบคุมคุณภาพค่าเฉลี่ย (\bar{X}) หรือ Individual chart และแผนควบคุมที่วัดการกระจายด้วยส่วนเบี่ยงเบนมาตรฐาน (σ) ค่าพิสัย (R)

2. แผนภูมิควบคุมสำหรับคุณลักษณะ (Control Chart for Attributes) เป็นแผนภูมิควบคุมที่ใช้เพื่อควบคุมกระบวนการผลิตสำหรับคุณลักษณะ (Attribute) ซึ่งคุณลักษณะนี้เป็นการตรวจสอบผลิตภัณฑ์ว่ามีลักษณะคุณภาพสอดคล้อง (conform) หรือไม่สอดคล้อง (non-conform) กับข้อกำหนด (specification) แผนภูมิประเภทนี้ประกอบด้วยแผนภูมิควบคุมสัดส่วนของเสีย (p control chart) แผนภูมิควบคุมร้อยละของของเสีย (PD control chart) แผนภูมิควบคุมจำนวนของเสีย (np control chart) แผนภูมิควบคุมรอยตำหนิ (c control chart) และแผนภูมิควบคุมรอยตำหนิต่อหน่วยที่ตรวจสอบ (u control chart)

3. แผนภูมิสำหรับตัวแปรชนิดอื่นๆ (Other Control Charts for Variables) จากแผนภูมิทั้ง 2 ประเภทที่กล่าวมาข้างต้นโดยส่วนใหญ่แล้วจะเกี่ยวข้องกับการควบคุมค่าเฉลี่ยและการกระจายของกระบวนการผลิต ซึ่งยังมีแผนภูมิควบคุมอื่นๆ ซึ่งประยุกต์ใช้ในบางสถานการณ์ เช่น แผนภูมิควบคุมค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบเอ็กซ์โปเนนเชียล (Exponential Weighted Moving Average Control Chart : EWMA)

ส่วนประกอบของแผนภูมิควบคุม (Component of Control Chart) แบ่งออกเป็น 5 ส่วน คือ

1. แกนนอน (Horizontal axis) แสดงลำดับของกลุ่มย่อย (Subgroup number) หรือเวลาที่เก็บรวบรวมข้อมูล
2. แกนตั้ง (Vertical axis) แสดงค่าสถิติ (Statistics) เช่น ค่าเฉลี่ย (Mean) มัชยฐาน (Median) หรือส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) เป็นต้น
3. จุดทึบ (Solid cycle) แสดงค่าสถิติแต่ละกลุ่มย่อย หรือค่าสังเกตของข้อมูล
4. เส้นกลาง (Central line : CL) แสดงมาตรฐานหรือเป้าหมายของกระบวนการ
5. ขีดจำกัดควบคุม (Control limit) แบ่งออกเป็นขีดจำกัดควบคุมล่าง (Lower control limit : LCL) และขีดจำกัดควบคุมบน (Upper control limit : UCL)

ลักษณะเส้นกลางและขีดจำกัดควบคุมทั่วไปของแผนภูมิควบคุมคือ

$$UCL = \mu + L\sigma$$

$$CL = \mu$$

$$LCL = \mu - L\sigma$$

เมื่อ μ แทนค่าเฉลี่ย

σ แทนส่วนเบี่ยงเบนมาตรฐาน

และ L แทนค่าคงที่ (constant)

หลังจากที่สร้างแผนภูมิควบคุมแล้ว ถ้าไม่มีจุดใดตกออกนอกขีดจำกัดควบคุมจะเรียกกระบวนการนั้นว่ากระบวนการอยู่ในการควบคุม (in control) แต่ถ้ามีจุดใดจุดหนึ่งตกออกนอกขีดจำกัดควบคุมจะเรียกกระบวนการนั้นว่ากระบวนการอยู่นอกการควบคุม (out-of-control) และเรียกจุดที่ตกออกนอกขีดจำกัดควบคุมนี้ว่าค่าผิดปกติหรือค่านอกกลุ่ม (Outliers)

แผนภูมิควบคุมดั้งเดิม (Traditional control chart) เช่น Shewhart Control Chart (เช่น แผนภูมิควบคุมคุณภาพค่าเฉลี่ย (\bar{X}) แผนภูมิควบคุมที่วัดการกระจายด้วยส่วนเบี่ยงเบนมาตรฐาน (σ) หรือค่าพิสัย (R)) หรือ EWMA chart มักมีข้อสมมติว่าข้อมูลจะต้องมีการแจกแจงปกติเหมือนกันและเป็นอิสระกัน แต่เมื่อกระบวนการไม่ได้มีการแจกแจงแบบปกติ (non-normality) และ/หรือไม่เป็นอิสระกันนั้นคือเกิดสหสัมพันธ์ในตัวเอง (Autocorrelation) จะพบว่าการใช้แผนภูมิควบคุมดั้งเดิมมาตรวจสอบกระบวนการจะมีประสิทธิภาพลดลง เนื่องจากการไม่ได้มีการแจกแจงแบบปกติ (non-normality) และ/หรือเกิดสหสัมพันธ์ในตัวเอง (Autocorrelation) หรือเมื่อเกิดค่านอกกลุ่มขึ้น จะส่งผลให้การประมาณค่าพารามิเตอร์ของกระบวนการเอนเอียง อัตราการเตือนผิด (false alarm rate) สูงขึ้น และทำให้การตรวจจับการเปลี่ยนแปลงของกระบวนการหรือค่านอกกลุ่มช้าลง ดังนั้นเมื่อเกิดปัญหาดังกล่าวข้างต้นจึงไม่เหมาะสมที่จะใช้แผนภูมิควบคุมดั้งเดิม

การสร้างแผนภูมิควบคุมเมื่อกระบวนการไม่ได้มีการแจกแจงแบบปกติ (non-normality) และ/หรือเกิดสหสัมพันธ์ในตัวเอง (Autocorrelation) หรือเมื่อเกิดค่านอกกลุ่มขึ้นสามารถทำได้หลายแนวทาง เช่น

- ทำการปรับขีดจำกัดควบคุมให้เหมาะสมกับลักษณะข้อมูล
- ใช้ตัวประมาณที่มีความแกร่ง (Robust Estimator) ซึ่งเป็นตัวประมาณมีประสิทธิภาพสูง Breakdown point สูง ตัวอย่างเช่น ค่าเบี่ยงเบนสมบูรณ์ของมัธยฐาน (Median absolute

deviation : MAD) เป็นตัวประมาณการกระจายที่มีความแปรปรวนซึ่งมี Breakdown point เท่ากับ 50% นั่นคือตัวประมาณ MAD จะเปลี่ยนแปลงไปเมื่อมีการเปลี่ยนแปลงของข้อมูล หรือมีจำนวน ค่านอกกลุ่มมากถึง 50% ในขณะที่ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) สำหรับข้อมูล n ข้อมูลมี Breakdown point เพียง $\frac{1}{n}$ นั่นคือตัวประมาณส่วนเบี่ยงเบนมาตรฐานจะเปลี่ยนแปลงไปเมื่อเกิดค่านอกกลุ่มเพียงค่าเดียว (Abu-Shawiesh :2008)

- สร้างแผนภูมิควบคุมบนพื้นฐานของความคลาดเคลื่อนในการพยากรณ์ (Special - Cause Control Chart หรือ Chart of residuals) เป็นการสร้างแผนภูมิควบคุมโดยค่าที่ใช้ในการควบคุมเป็นความคลาดเคลื่อนในการพยากรณ์แทนการใช้ค่าสังเกตหรือข้อมูลเดิม เนื่องจากความคลาดเคลื่อนในการพยากรณ์มักสอดคล้องกับข้อสมมติว่ามีการแจกแจงแบบปกติเหมือนกันและเป็นอิสระกัน โดยที่ถ้าค่าสังเกตมีค่าแตกต่างไปจากค่าที่พยากรณ์มาก จะส่งผลให้ความคลาดเคลื่อนของการพยากรณ์มีขนาดใหญ่ผิดปกติและค่าสังเกตนั้นจะถูกจัดเป็นค่านอกกลุ่ม (Alwan และ Robert : 1988)

สำหรับเกณฑ์ในการพิจารณาการกำหนดแผนภูมิควบคุมมีดังนี้

1. จำนวนการเตือนผิด (false alarms) แผนภูมิควบคุมที่มีประสิทธิภาพไม่ควรเกิดกรณีการเตือนว่ากระบวนการอยู่นอกการควบคุม (Out of control) ใดๆที่กระบวนการอยู่ในการควบคุม
2. ความไวต่อการตรวจจับค่านอกกลุ่ม แผนภูมิควบคุมควรตรวจจับค่านอกกลุ่มได้ทันทั่วทั้งที่มีค่านอกกลุ่มปรากฏขึ้น
3. ความง่ายสำหรับผู้ใช้ในการอธิบายความหมาย

โดยทั่วไปแล้วเมื่อพยายามลดอัตราการเตือนผิด (false alarms rate) จะเป็นการลดความไวของแผนภูมิควบคุมในการตรวจจับค่านอกกลุ่มด้วย ดังนั้นความสามารถในการตรวจจับค่านอกกลุ่มที่เกิดขึ้นจริงจะต่ำลง (Bisgaard : 2005)

งานวิจัยที่เกี่ยวข้อง

Alwan และ Robert (1988) ได้เสนอ Special - Cause Control Chart (SCC) สำหรับข้อมูลที่มีสหสัมพันธ์ในตัวเอง โดยที่ SCC จะใช้ความคลาดเคลื่อนในการพยากรณ์เป็นค่าที่ใช้ควบคุม เทคนิคการพยากรณ์ที่นิยมกันมากคือการวิเคราะห์อนุกรมเวลาแบบบ็อกซ์และเจนกินส์ (Box-Jenkins Analysis) เนื่องจากถ้าตัวแบบที่กำหนดถูกต้องและเหมาะสม ส่วนเหลือ (residual) จะมีข้อสมมติตรงกับข้อสมมติของแผนภูมิควบคุมดั้งเดิมคือเป็นอิสระกันและมีการแจกแจงแบบปกติ

Wright และ HU (2003) ได้ทำการเปรียบเทียบวิธีการตรวจจับค่านอกกลุ่ม joint estimation (JE) และวิธีค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบเอ็กซ์โปเนนเชียล (EWMA) ในการตรวจจับค่านอกกลุ่มสำหรับข้อมูลที่มีสหสัมพันธ์ในตัวเองระยะสั้น เมื่อค่านอกกลุ่มเป็นค่าสุดท้ายของอนุกรมเวลาและเป็นค่านอกกลุ่มประเภท AO เนื่องจากเมื่อค่านอกกลุ่มเกิดขึ้นเป็นค่าแรกค่านอกกลุ่มทุกประเภทจะเป็นค่านอกกลุ่มประเภท AO โดยวิธี JE เสนอโดย Chen และ Lui (1993) และสำหรับวิธี EWMA เสนอโดย Ng และ Case ซึ่งเป็นวิธีสำหรับอนุกรมเวลาระยะสั้นสำหรับค่าสังเกตค่าหนึ่ง

พิจารณาอนุกรมเวลา z_t ; $t = 1, 2, \dots, T$

กำหนดให้ moving range คือ $MR_t = |z_t - z_{t-1}|$; $t = 2, 3, \dots, T$

ตัวสถิติ EWMA คือ $EWMA_t = \lambda MR_t + (1 - \lambda)EWMA_{t-1}$

เมื่อ ค่าเริ่มต้น $EWMA_1$ กำหนดจากค่าเฉลี่ยของ moving range

$$EWMA_1 = (\sum_{t=2}^T MR_t) / (T - 1) = \overline{MR}$$

ดังนั้นขีดจำกัดควบคุมแผนภูมิควบคุมของ EWMA คือ

$$UCL = \overline{MR} \left(1 + 3 \frac{d'_3}{d'_2} \sqrt{\frac{\lambda}{2-\lambda}} \right)$$

$$CL = \overline{MR}$$

$$LCL = \overline{MR} \left(1 - 3 \frac{d'_3}{d'_2} \sqrt{\frac{\lambda}{2-\lambda}} \right)$$

เมื่อ d'_2 และ d'_3 ได้จากตาราง Factors for Constructing Variables Control Charts ของ Montgomery (2009) เมื่อ n แทนจำนวนค่าสังเกตที่ทำการ moving นั้นคือ $n=2$ โดยถ้า $EWMA_t$ ตกออกนอกขีดจำกัดควบคุม ค่านั้นจะถือเป็นค่านอกกลุ่ม

การจำลองการศึกษาด้วยตัวแบบ ARMA แต่ละอนุกรมจะสร้างจาก seed ที่ต่างกันเพื่อป้องกันการมีสหสัมพันธ์ระหว่างอนุกรม และในแต่ละพารามิเตอร์จะจำลองอนุกรมเวลา 40 อนุกรมโดย $\phi = 0 = 0.2, 0.5$ และ 0.7 สำหรับอนุกรมอัตถคดอยอันดับ 1 (AR(1)) และอนุกรมค่าเฉลี่ยเคลื่อนที่อันดับ 1 (MA(1)) ใช้ขนาดตัวอย่างสำหรับแต่ละอนุกรม 9, 11, 13 และ 15 และขนาดของค่านอกกลุ่มเท่ากับ 2, 3

และ 4 พิจารณาประสิทธิภาพของทั้งสองวิธีจาก F_{1L} แทนร้อยละของการตรวจจับค่านอกกลุ่มในตำแหน่งแรกและ F_{FA} แทนร้อยละของการเตือนผิดนั้นคือร้อยละของการตรวจจับค่านอกกลุ่มเมื่อไม่มีค่านอกกลุ่มปรากฏ จากการจำลองการศึกษาพบว่าวิธีการ EWMA มีประสิทธิภาพที่ดีเทียบเท่ากับ JE สำหรับ AR(1) เมื่อพิจารณาจาก F_{1L} และมีประสิทธิภาพดีกว่า JE สำหรับ AR(1) เมื่อพิจารณาจาก F_{FA} และสำหรับ MA(1) วิธี EWMA ดีกว่าวิธี JE เมื่อพิจารณาจากทั้ง F_{1L} และ F_{FA} จากงานวิจัยนี้แสดงให้เห็นว่าวิธีที่มีความซับซ้อนมากกว่าไม่จำเป็นต้องมีประสิทธิภาพที่ดีกว่า นอกจากนี้ข้อดีของวิธี EWMA คือไม่ต้องกำหนดรูปแบบของอนุกรมเวลา

Bisgaard (2005) ได้พิจารณาปัญหาสหสัมพันธ์ในตัวเองที่ส่งผลกระทบต่อวิธีการควบคุมกระบวนการทางสถิติ โดย Bisgaard ได้กล่าวว่าสหสัมพันธ์ในตัวเองจะส่งผลกระทบต่อแผนภูมิควบคุมและโดยส่วนใหญ่แล้วในกระบวนการเชิงอุตสาหกรรมจะพบว่าสหสัมพันธ์ในตัวเองเป็นเชิงบวก (positive autocorrelation) โดยได้พิจารณาจากข้อมูลจริงซึ่งเป็นข้อมูลของการควบคุมอุณหภูมิที่เก็บทุกๆ ชั่วโมง ติดต่อกัน 80 ค่า z_t ; $t = 1, 2, \dots, 80$ และอนุกรมเวลานี้อยู่ในการควบคุม (in control) ลักษณะของข้อมูลมีการเคลื่อนที่เป็นแบบคลื่น (wavy pattern) ไม่เป็นแบบสุ่ม (random pattern) จากการกำหนดตัวแบบพบว่าอนุกรมเวลานี้เป็น AR(2)

แผนภูมิควบคุมแรกที่น่ามาพิจารณาคือ individuals control chart แผนภูมิควบคุมนี้มีข้อสมมติว่าข้อมูลจะต้องเป็นอิสระ ในที่นี้จะใช้ขีดจำกัดควบคุมแบบ 3 - sigma และคำนวณความผันแปรของกระบวนการจาก moving range คือ $MR_t = |z_t - z_{t-1}|$ ดังนั้นขีดจำกัดควบคุมคือ

$$UCL = \bar{z} + 3 \frac{\overline{MR}}{d_2}$$

$$CL = \bar{z}$$

$$LCL = \bar{z} - 3 \frac{\overline{MR}}{d_2}$$

เมื่อ \overline{MR} คือค่าเฉลี่ยของ moving range และ d_2 ดูจาก motgomery(2001) เมื่อ $n=2$ และนำขีดจำกัดนี้ไปควบคุมข้อมูลอุณหภูมิพบว่ากระบวนการอยู่นอกการควบคุม (out-of-control)

แผนภูมิควบคุมที่สองที่น่ามาพิจารณาคือ EWMA control chart แผนภูมิควบคุมนี้มีข้อสมมติว่าข้อมูลจะต้องเป็นอิสระและมีการแจกแจงแบบปกติ ในที่นี้จะใช้ขีดจำกัดควบคุมแบบ 3 - sigma

$$\text{โดยตัวสถิติ EWMA คือ } EWMA_t = \lambda z_t + (1 - \lambda)EWMA_{t-1}$$

ดังนั้นขีดจำกัดควบคุมแผนภูมิควบคุมของ EWMA คือ

$$UCL = \bar{z} + 3 \frac{\overline{MR}}{d_2} \sqrt{\frac{\lambda}{2-\lambda} [1 - (1 - \lambda)^{2t}]}$$

$$CL = \bar{z}$$

$$LCL = \bar{z} - 3 \frac{MR}{d_2} \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2t}]}$$

ถ้า $EWMA_t$ ตกลงนอกขีดจำกัดควบคุม ค่ามันจะถือเป็นค่านอกกลุ่ม จากนั้นนำขีดจำกัดนี้ไปควบคุมข้อมูลอุณหภูมิพบว่ากระบวนการอยู่นอกการควบคุม (out-of-control)

จากแผนภูมิควบคุมทั้งสองที่กล่าวมาข้างต้นจะพบว่ากระบวนการอยู่นอกการควบคุม (out-of-control) ทั้ง ๆ ที่กระบวนการอยู่ในการควบคุม (in control) สาเหตุหลักมาจากอนุกรมเวลาที่พิจารณานี้มีสหสัมพันธ์ในตัวเองทำให้ขีดจำกัดควบคุมที่สร้างขึ้นไม่เหมาะสม

Bisgaard จึงได้เสนอแนวทางการแก้ปัญหาเมื่อเกิดปัญหาสหสัมพันธ์ในตัวเองแนวทางแรกคือการใช้ขีดจำกัดที่กว้างขึ้น (inflated limits) สำหรับ Individuals control chart วิธีการนี้จะเป็นการคำนวณความแปรปรวนของกระบวนการจาก variance inflation factor โดยสำหรับ AR(2) จะได้ว่าความแปรปรวนของ z_t คือ

$$\sigma_z^2 = \left(\frac{1 - \phi_2}{1 + \phi_2} \right) \frac{\sigma_a^2}{\{(1 - \phi_2)^2 - \phi_1^2\}}$$

โดยที่ σ_a^2 แทนความแปรปรวนของส่วนเหลือ

และขีดจำกัดควบคุมคือขีดจำกัดควบคุมคือ

$$UCL = \bar{z} + 3\hat{\sigma}_z$$

$$CL = \bar{z}$$

$$LCL = \bar{z} - 3\hat{\sigma}_z$$

จากนั้นนำขีดจำกัดนี้ไปควบคุมข้อมูลอุณหภูมิพบว่ากระบวนการอยู่ในการควบคุม (in control) เนื่องจากขีดจำกัดควบคุมที่สร้างขึ้นกว้างกว่าขีดจำกัดควบคุมของ individual control chart แต่ขีดจำกัดควบคุมที่กว้างขึ้นอาจนำไปสู่การตรวจจับที่ช้าลงเมื่อเกิดค่านอกกลุ่มขึ้น

อีกแนวทางหนึ่งคือการใช้แผนภูมิ EWMA โดยใช้ส่วนเหลือ (e_t) ที่ได้จากการพยากรณ์เป็นค่าที่ใช้ในการควบคุม

โดยตัวสถิติ EWMA คือ $EWMA_t = \lambda e_t + (1 - \lambda)EWMA_{t-1}$

ดังนั้นขีดจำกัดควบคุมแผนภูมิควบคุมของ EWMA คือ

$$UCL = 3\hat{\sigma}_e \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2t}]}$$

$$CL = 0$$

$$LCL = 3\hat{\sigma}_e \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2t}]}$$

ซึ่งพบว่ากระบวนการอยู่ในการควบคุม (in control)

Gelper , Fried และ Croux (2010) ได้กล่าวไว้ว่าวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลเป็นวิธีการพยากรณ์อย่างง่าย (naïve prediction method) ที่นิยมใช้ในการพยากรณ์อนุกรมเวลาที่มีลักษณะคงที่ และวิธีการพยากรณ์ Holt-Winters เป็นวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสองชั้น (Double exponential smoothing) ซึ่งเป็นวิธีที่นิยมใช้มากในอนุกรมเวลาที่ประกอบด้วยแนวโน้ม (Trend) และฤดูกาล (Seasonal) โดยทั้งสองวิธีนี้ไม่ต้องการตัวแบบสำหรับการพยากรณ์ แต่วิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลและ Holt-Winters เป็นวิธีที่มีความไวต่อค่าผิดปกติ (outliers) นั่นคือเมื่อเกิดค่าผิดปกติขึ้นจะส่งผลกระทบต่อวิธีการพยากรณ์ทั้งสองวิธีนี้ 2 ทางคือค่าพยากรณ์ที่ปรับให้เรียบเนื่องจากค่าพยากรณ์จะขึ้นอยู่กับค่าในอดีต และการกำหนดค่าพารามิเตอร์ที่ปรับให้เรียบเนื่องจากการกำหนดค่าพารามิเตอร์จะเลือกจากค่าพารามิเตอร์ที่ทำให้ผลบวกกำลังสองของความคลาดเคลื่อนในการพยากรณ์มีค่าต่ำที่สุด พวกเขาจึงได้เสนอวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (exponential smoothing) ที่มีความแกร่งและ Holt-Winters ที่มีความแกร่ง (Robust Holt-Winters : RHW) โดยใช้แนวคิดของการแทนที่อนุกรมเวลาที่มีค่าผิดปกติด้วยอนุกรมเวลาที่ไม่มียค่าผิดปกติ (cleaning data) ซึ่งให้นำหนักกับค่าผิดปกติลดลงบนพื้นฐานของกฎ 2 ซิกมา (2-sigma rule)

สำหรับวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (exponential smoothing) เมื่อพิจารณาอนุกรมเวลา $z_t ; t = 1, 2, \dots, T$ ค่าพยากรณ์ล่วงหน้า 1 คาบเวลาคือ $\hat{z}_{t|t-1}$ สามารถคำนวณได้จาก

$$\hat{z}_{t|t-1} = \lambda z_{t-1} + (1 - \lambda)\hat{z}_{t-1|t-2} \quad 1)$$

เมื่อ λ แทนพารามิเตอร์ที่ปรับให้เรียบ โดยที่ $0 \leq \lambda \leq 1$ สำหรับค่าเริ่มต้น \hat{z}_m กำหนดจากค่าเฉลี่ยของค่าสังเกต m ค่าแรก ; z_1, z_2, \dots, z_m

สำหรับวิธีการพยากรณ์ Holt-Winters เมื่อพิจารณาอนุกรมเวลา $z_t ; t = 1, 2, \dots, T$ ค่าพยากรณ์ล่วงหน้า h คาบเวลาคือ $\hat{z}_{t+h|t}$ สามารถคำนวณได้จาก

$$\hat{z}_{t+h|t} = \hat{\alpha}_t + h\hat{\beta}_t \quad 2)$$

โดยที่

$$\hat{\alpha}_t = \lambda_1 z_t + (1 - \lambda_1)(\hat{\alpha}_{t-1} + \hat{\beta}_{t-1}) \quad 3)$$

$$\hat{\beta}_t = \lambda_2(\hat{\alpha}_t - \hat{\alpha}_{t-1}) + (1 - \lambda_2)\hat{\beta}_{t-1}$$

$\hat{\alpha}_t$ คือระดับ (local) ณ ตำแหน่งเวลาของอนุกรมเวลา

$\hat{\beta}_t$ คือแนวโน้ม (trend) ณ ตำแหน่งเวลาของอนุกรมเวลา

และ λ_1, λ_2 แทนพารามิเตอร์ที่ปรับให้เรียบ โดยที่ $0 \leq \lambda \leq 1$ สำหรับค่าเริ่มต้น $\hat{\alpha}_m$ และ $\hat{\beta}_m$ กำหนดจาก ordinary least square (OLS) ของเส้นแนวโน้มเชิงเส้น นั่นคือการถดถอย z_t บนเวลา t ; $t=1,2,\dots,m$ จะได้ว่า $\hat{\alpha}_m = \hat{\alpha}_0 + \hat{\beta}_0 m$ และ $\hat{\beta}_m = \hat{\beta}_0$ เมื่อ $\hat{\alpha}_0$ แทน intercept และ $\hat{\beta}_0$ แทนความชัน (slope)

วิธีที่มีความแกร่งมีแนวคิดคือการแทนที่ค่าสังเกต z_t ด้วยรูปแบบของอนุกรมเวลาที่ไม่มีความนอกกลุ่ม (clean version) z_t^* สำหรับวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (exponential smoothing) ที่มีความแกร่งจะได้ว่า

$$\hat{z}_{t|t-1} = \lambda z_{t-1}^* + (1 - \lambda)\hat{z}_{t-1} \quad 4)$$

และคล้ายคลึงกันสำหรับวิธีการพยากรณ์ Holt-Winters ที่มีความแกร่งจะได้ว่า

$$\hat{z}_{t+h|t} = \hat{\alpha}_t + h\hat{\beta}_t$$

โดยที่

$$\hat{\alpha}_t = \lambda_1 z_t^* + (1 - \lambda_1)(\hat{\alpha}_{t-1} + \hat{\beta}_{t-1}) \quad 5)$$

$$\hat{\beta}_t = \lambda_2(\hat{\alpha}_t - \hat{\alpha}_{t-1}) + (1 - \lambda_2)\hat{\beta}_{t-1}$$

สำหรับทั้งวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลและ Holt-Winters อนุกรมเวลาที่ไม่มีความนอกกลุ่ม (clean series) z_t^* คือ

$$z_t^* = \psi_k \left(\frac{z_t - \hat{z}_{t|t-1}}{\hat{\sigma}_t} \right) \hat{\sigma}_t + \hat{z}_{t|t-1} \quad 6)$$

เมื่อ $\hat{\sigma}_t$ แทนส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในการพยากรณ์

$\psi_k(x)$ คือ Huber ψ -function ประยุกต์ใช้กับค่ามาตรฐานของความคลาดเคลื่อนในการพยากรณ์ล่วงหน้า 1 คาบเวลา

$$\text{โดยที่ } \psi_k(x) = \begin{cases} x & ; |x| < k \\ \text{sign}(x)k & ; \text{อื่นๆ} \end{cases}$$

Huber ψ -function เป็นฟังก์ชันที่ใช้ลดอิทธิพลของค่านอกกลุ่ม โดยเป็นการแทนที่ค่าที่มีค่าสูงมากเกินไปและต่ำมากเกินไปหรือค่านอกกลุ่มด้วยค่าที่เหมาะสม นั่นคือถ้าค่าสังเกต z_t แตกต่างจากค่า

พยากรณ์ $\hat{z}_{t|t-1}$ น้อย ๆ แล้ว z_t^* จะเท่ากับ z_t แต่ในทางตรงกันข้ามถ้าค่าสังเกต z_t แตกต่างจากค่าพยากรณ์ $\hat{z}_{t|t-1}$ มาก ๆ ค่าสังเกตนั้นจะพิจารณาเป็นค่านอกกลุ่มแล้วแทนที่ z_t ด้วย z_t^* ซึ่งเป็นค่าที่มีขอบเขต (boundary value) ที่ขึ้นอยู่กับค่า k ซึ่งค่า k นั้นเป็นค่าที่ใช้ระบุค่านอกกลุ่ม โดยทั่วไปแล้วค่า k จะกำหนดให้เท่ากับ 2 (2-sigma rule) เมื่อสมมติให้ความคลาดเคลื่อนในการพยากรณ์ล่วงหน้า 1 คาบเวลา $e_t = z_t - \hat{z}_{t|t-1}$ มีการแจกแจงแบบปกติ และถ้ากำหนดให้ k เท่ากับค่าไม่จำกัด (infinity) จะได้ว่าวิธีที่มีความแกร่งจะลดรูปเป็นวิธีมาตรฐาน

สำหรับ $\hat{\sigma}_t$ แทนส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนในการพยากรณ์ ณ เวลา t จะประมาณด้วยวิธีที่มีความแกร่ง นั่นคือค่าเบี่ยงเบนสมบูรณ์ของมัธยฐาน (Median Absolute Deviation : MAD) กำหนดโดย

$$\hat{\sigma}_t = \text{med}_{1 \leq t \leq n} |e_t - \text{med}_{1 \leq t \leq n}(e_t)|$$

แต่อย่างไรก็ตาม MAD เป็นการวัดการกระจายที่เป็นค่าคงที่ ดังนั้นสำหรับรูปแบบที่ยอมให้การกระจายมีความแตกต่างกันไปอย่างช้า ๆ (slowly) สามารถกำหนดได้จาก

$$\hat{\sigma}_t^2 = \lambda_\sigma \rho_k \left(\frac{e_t}{\hat{\sigma}_{t-1}} \right) \hat{\sigma}_{t-1}^2 + (1 - \lambda_\sigma) \hat{\sigma}_{t-1}^2 \quad 7)$$

เมื่อ λ_σ เป็นพารามิเตอร์ที่ปรับให้เรียบ

ρ เป็นฟังก์ชันความสูญเสียแบบมีขอบเขต (bounded loss-function) โดยเลือกใช้ biweight ρ -function กำหนดโดย

$$\rho_k(x) = \begin{cases} c_k(1 - (1 - (x/k)^2)^3) & ; |x| < k \\ c_k & ; \text{อื่น ๆ} \end{cases} \quad 8)$$

เมื่อ c_k แทนค่าคงที่ที่ทำให้พารามิเตอร์ของความแปรปรวนมีคุณสมบัติคงเส้นคงวา (consistency) สำหรับความคลาดเคลื่อนที่มีการแจกแจงแบบปกติ โดยทั่วไปถ้า $k=2$ จะมีค่า $c_k=2.52$

จากวิธีที่มีความแกร่งจะเห็นว่าค่านอกกลุ่มไม่ได้ส่งผลกระทบต่อค่าที่ปรับให้เรียบและพารามิเตอร์ที่ปรับให้เรียบ นอกจากนี้ยังพิจารณาวิธีการกำหนดค่าเริ่มต้นที่มีความแกร่งอีกด้วย สำหรับค่าเริ่มต้นของวิธีการปรับให้เรียบแบบเอ็กโปเนนเชียลใช้มัธยฐานของค่าสังเกต m ค่าแทนที่ค่าเฉลี่ย และสำหรับค่าเริ่มต้นของวิธี Holt-Winters จะแทนที่ ordinary least square (OLS) ของเส้นแนวโน้มเชิงเส้นด้วยตัวประมาณมัธยฐานแบบทวนซ้ำ (repeat median estimator) เมื่อ

$$\hat{\beta}_0 = \text{med}_i \left(\text{med}_{i \neq j} \frac{z_i - z_j}{i - j} \right)$$

และ $\hat{\alpha}_0 = \text{med}_i (z_i - \hat{\beta}_0 i)$ สำหรับ $i, j = 1, 2, \dots, m$

สำหรับวิธีการเลือกพารามิเตอร์ที่ปรับให้เรียบ λ สำหรับวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล และ λ_1, λ_2 สำหรับวิธี Holt-Winters โดยปกติแล้วมักจะเลือกพารามิเตอร์ที่ปรับให้เรียบที่ทำให้ความคลาดเคลื่อนในการพยากรณ์เฉลี่ย (mean squared forecast error : MSFE) มีค่าต่ำที่สุด

$$\text{MSFE}(e_1, \dots, e_T) = \sum_{t=1}^T e_t^2 \quad (9)$$

ซึ่ง MSFE จะมีค่าเปลี่ยนแปลงไปเมื่อมีค่าความคลาดเคลื่อนในการพยากรณ์ขนาดใหญ่เพียงค่าเดียว และมักจะนำไปสู่การเลือกพารามิเตอร์ที่ปรับให้เรียบที่เอนเอียงสู่ศูนย์ จึงเป็นเหตุผลที่จะใช้ MSFE ในรูปแบบที่มีความแกร่ง โดยใช้ตัวประมาณ τ^2 (τ^2 -estimator)

$$\tau^2 = s_T^2 \frac{1}{T} \sum_{t=1}^T \gamma_k \left(\frac{e_t}{s_T} \right) \quad (10)$$

$$\text{เมื่อ } s_T = 1.48 \text{ med}_t |e_t|$$

ตัวประมาณ τ^2 เป็นตัวประมาณที่ให้น้ำหนักที่ลดลง (downweights) กับความคลาดเคลื่อนในการพยากรณ์ที่มีขนาดใหญ่การลดลงทำโดยฟังก์ชันการสูญเสีย

$$\gamma_k(y) = \min(k^2, y^2) = \psi_k(y)^2 \quad (11)$$

ด้วย $k=2$ และเลือกพารามิเตอร์ที่ปรับให้เรียบที่ทำให้ τ^2 มีค่าต่ำที่สุด

หลังจากได้เสนอวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและ Holt-Winters ที่มีความแกร่ง (RHW) และได้ทำการเปรียบเทียบการพยากรณ์วิธี Holt-Winters ที่มีความแกร่ง (RHW) กับ Holt-Winters มาตรฐาน (HW) โดยใช้การจำลองการศึกษาจากตัวแบบแนวโน้มเชิงเส้น (linear trend model) กำหนดโดย

$$z_t = \alpha_t + e_t, \quad e_t \sim N(0, \sigma_e^2) \quad (12)$$

เมื่อ α_t แทนระดับที่ t ตำแหน่งเวลา t

β_t แทนแนวโน้มที่ t ตำแหน่งเวลา t

$$\text{โดยที่} \quad \begin{aligned} \alpha_t &= \alpha_{t-1} + \beta_{t-1} + \eta_t, & \eta_t &\sim N(0, \sigma_\eta^2) \\ \beta_t &= \beta_{t-1} + v_t, & v_t &\sim N(0, \sigma_v^2) \end{aligned}$$

เทอมความคลาดเคลื่อน e_t , η_t และ v_t เป็นอิสระกันและไม่มีสหสัมพันธ์ในตัวเอง ในการจำลองการศึกษาจะให้ $\sigma_e=1$ และ $\sigma_\eta=\sigma_v=0.1$ โดยพิจารณารูปแบบการจำลอง 4 รูปแบบดังนี้

Clean data : CD	อนุกรมเวลาไม่มีค่านอกกลุ่ม	$e_t \sim \text{iid } N(0,1)$
Symmetric outliers : SO	ความน่าจะเป็นที่จะเกิดค่านอกกลุ่มที่มีค่ามากเกินไปหรือน้อยเกินไปเท่ากัน	$e_t \sim \text{iid } \varepsilon N(0,1) + (1-\varepsilon)N(0,20)$
Asymmetric outliers : AO	เกิดค่านอกกลุ่มที่มีค่าบวก (positive) เท่านั้น	$e_t \sim \text{iid } \varepsilon N(0,1) + (1-\varepsilon)N(20,1)$
Fat-tailed data : FT	อนุกรมเวลาไม่มีค่านอกกลุ่ม	$e_t \sim \text{iid } t_3$

โดยที่ ε แทนความน่าจะเป็นที่จะเกิดค่านอกกลุ่มเท่ากับ 0.05

t_3 แทนการแจกแจงแบบ t ที่องศาความเป็นอิสระเท่ากับ 3

พิจารณาค่า MSFE และ τ^2 สำหรับในกรณีอนุกรมเวลาไม่มีค่านอกกลุ่มพบว่าวิธีการพยากรณ์ Holt-Winters มีค่า MSFE และ τ^2 ต่ำกว่าวิธีการพยากรณ์ Holt-Winters ที่มีความแกร่งแต่ไม่มากนัก แต่เมื่อเกิดค่านอกกลุ่ม (SO,AO) ขึ้นจะพบว่าวิธีการพยากรณ์ Holt-Winters ที่มีความแกร่งมีค่า MSFE และ τ^2 ต่ำกว่าวิธีการพยากรณ์ Holt-Winters อย่างชัดเจนนั่นคือวิธีการพยากรณ์ Holt-Winters ที่มีความแกร่งสามารถพยากรณ์อนุกรมเวลาได้ถูกต้องมากกว่าวิธีการพยากรณ์ Holt-Winters

เมื่อพิจารณาผลกระทบของค่านอกกลุ่มที่มีต่อค่า MSFE และ τ^2 สำหรับวิธีการพยากรณ์ Holt-Winters ในกรณีที่ไม่มีค่านอกกลุ่มพบว่าค่า MSFE และ τ^2 มีค่าใกล้เคียงกัน แต่ในกรณีที่เกิดค่านอกกลุ่ม (SO,AO) ขึ้นพบว่า MSFE จะมีค่าสูงกว่า τ^2 มาก และสำหรับวิธีการพยากรณ์ Holt-Winters ที่มีความแกร่งทั้งในกรณีที่ไม่มีค่านอกกลุ่มและกรณีที่เกิดค่านอกกลุ่ม (SO,AO) พบว่า MSFE และ τ^2 มีค่าใกล้เคียงกัน

สำหรับกรณีการแจกแจงไม่ได้มีการแจกแจงแบบปกติ (FT) จะพบว่าวิธีการพยากรณ์ Holt-Winters ที่มีความแกร่งมีค่า MSFE และ τ^2 ต่ำกว่าวิธีการพยากรณ์ Holt-Winters ค่อนข้างมาก และพบว่าทั้งสองวิธีการพยากรณ์นี้ให้ค่า MSFE สูงกว่า τ^2 มาก

จากการจำลองผลลัพธ์ที่สังเกตได้อีกอย่างหนึ่งนั่นคือการเลือกพารามิเตอร์ที่ปรับให้เรียบ (λ_1 , λ_2) พบว่าเมื่อเกิดค่านอกกลุ่ม (SO,AO) ขึ้นสำหรับวิธีการพยากรณ์ Holt-Winters (λ_1 , λ_2) จะมีค่าลดลงสู่ศูนย์เมื่อเกิดค่านอกกลุ่ม ในขณะที่วิธีการพยากรณ์ Holt-Winters ที่มีความแกร่ง (λ_1 , λ_2) ของทั้งกรณีที่มีค่านอกกลุ่มและไม่มีค่านอกกลุ่มจะมีค่าใกล้เคียงกัน และในกรณีการแจกแจงไม่ได้มีการแจกแจงแบบปกติ

(FT) จะไม่สามารถเปรียบเทียบกับพารามิเตอร์ที่ปรับให้เรียบกรณีไม่มีค่านอกกลุ่ม, มีค่านอกกลุ่ม SO และค่านอกกลุ่ม AO ได้โดยตรงเนื่องจากอนุกรมเวลาที่สร้างขึ้นไม่ได้เป็นอนุกรมเวลาที่มีค่านอกกลุ่มแต่มีลักษณะการแจกแจงที่แตกต่างออกไป

นอกจากนี้ยังได้เปรียบเทียบประสิทธิภาพของวิธีการพยากรณ์ Holt-Winters กับวิธีการพยากรณ์ Holt-Winters ที่มีความแกร่งด้วยตัวแบบ ARIMA(1,1,1) และตัวแบบ deterministic trend

$$\text{ตัวแบบ ARIMA}(1,1,1) : (1 - 0.2B)(1 - B)z_t = (1 - 0.5)e_t \text{ เมื่อ } e_t \sim \text{iid } N(0,1)$$

$$\text{ตัวแบบ deterministic trend} : z_t = 0.5t + e_t$$

$$\text{โดยที่ } (1 - 0.2B)e_t = (1 - 0.5)\delta_t \text{ เมื่อ } \delta_t \sim \text{iid } N(0,1)$$

พิจารณาค่า MSFE และ τ^2 ผลลัพธ์ของทั้งสองตัวแบบนี้มีลักษณะเช่นเดียวกันกับตัวแบบแนวโน้มเชิงเส้น(สมการ 12) ดังนั้นเมื่อเกิดค่านอกกลุ่มในอนุกรมเวลาวิธีการพยากรณ์ที่มีความแกร่งจึงเป็นที่ต้องการ เพื่อให้การพยากรณ์มีความถูกต้องและน่าเชื่อถือ

Croux , Gelper และ Mahieu (2011) ได้เสนอแผนภูมิควบคุมที่มีความแกร่งต่อค่านอกกลุ่มในอนุกรมเวลา เพื่อป้องกันผลกระทบจากค่านอกกลุ่มในการประมาณขีดจำกัดควบคุม โดยใช้แผนภูมิควบคุมที่จัดอยู่ในประเภท Special - Cause Control Chart (SCC) (Alwan&Roberts :1988) นั่นคือ ความคลาดเคลื่อนการพยากรณ์ล่วงหน้า 1 คาบเวลา (one-step-ahead) เป็นค่าที่ใช้ควบคุมในเทคนิคแผนภูมิควบคุม เพื่อขจัดปัญหาของข้อมูลที่มีสหสัมพันธ์ในตัวเอง และใช้วิธีการพยากรณ์ Holt-Winters ซึ่งเป็นที่นิยมใช้กันอย่างแพร่หลายและเป็นวิธีง่าย ๆ ในการพยากรณ์อนุกรมเวลาที่ประกอบด้วยแนวโน้มและสหสัมพันธ์ในตัวเอง ในการศึกษานี้ได้แบ่งอนุกรมเวลาออกเป็น 2 ส่วนนั่นคือ “training period” เป็นส่วนของอนุกรมเวลาที่ใช้ในการกำหนดขีดจำกัดควบคุม และ “test period” เป็นส่วนของอนุกรมเวลาที่ใช้ในการทดสอบขีดจำกัดควบคุมที่สร้างขึ้นใน training period มาควบคุม อนุกรมเวลาส่วนนี้ โดยที่ค่านอกกลุ่มอาจปรากฏขึ้นในอนุกรมเวลาทั้ง 2 ส่วน เมื่อเกิดค่านอกกลุ่มขึ้นจะส่งผลให้การประมาณพารามิเตอร์เอนเอียง และค่าพยากรณ์ถัดมาหลังจากเกิดค่านอกกลุ่มจะได้รับอิทธิพลของค่านอกกลุ่มส่งผลทำให้เกิดความคลาดเคลื่อนในการพยากรณ์มีค่าเพิ่มขึ้น เนื่องจากวิธีการพยากรณ์ Holt-Winters จะขึ้นอยู่กับค่าสังเกตในอดีต ซึ่งในงานวิจัยนี้ได้ใช้ความคลาดเคลื่อนในการพยากรณ์มาควบคุม ดังนั้นเมื่อเกิดค่านอกกลุ่มใน training period จะส่งผลให้ความคลาดเคลื่อนในการพยากรณ์มีค่าเพิ่มขึ้นและทำให้ขีดจำกัดควบคุมกว้างขึ้น และเมื่อเกิดค่านอกกลุ่มใน test period จะทำให้เกิดอัตราการเตือนผิดที่สูงขึ้น จากที่กล่าวมาจะพบว่าการสร้างแผนภูมิควบคุม SCC โดยวิธีการพยากรณ์ Holt-Winters ได้รับผลกระทบจากค่านอกกลุ่มค่อนข้างมาก ดังนั้นพวก

เขาจึงได้เสนอวิธีการพยากรณ์ Holt-Winters ที่มีความแกร่งตามที่ Gelper , Fried และ Croux (2010) ได้เสนอไว้ และนำมาสร้างแผนภูมิควบคุม หลังจากนั้นได้ทำการเปรียบเทียบประสิทธิภาพของแผนภูมิควบคุมวิธี Holt-Winters มาตรฐานและวิธี Holt-Winters ที่มีความแกร่ง โดยพิจารณาประสิทธิภาพจากความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 , กำลังและอัตราการตรวจจับที่ผิดพลาดของแผนภูมิควบคุม เมื่อกำหนด $\alpha=0.05$ โดยการจำลองการ ศึกษาจากตัวแบบแนวโน้มเชิงเส้น ซึ่ง ตัวแบบนี้สร้างอนุกรมเวลาแบบไม่คงที่ ประกอบด้วยแนวโน้ม และมีสหสัมพันธ์กัน ความคลาดเคลื่อนในการพยากรณ์ของวิธี Holt-Winters สร้างอนุกรม iid จากการแจกแจงแบบปกติ จำลอง 1,000 รอบ พิจารณา ขนาดของตัวอย่างใน training $n=50$ และ $n=100$ และค่าสังเกต 200 ค่าถัดมาเป็นตัวอย่างทดสอบ ค่านอกกลุ่มสร้างขึ้นโดยบวกค่า ko เข้ากับค่าสังเกตที่สุ่มมาจากอนุกรมเวลา

จากการจำลองการศึกษาเมื่อพิจารณาความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 พบว่า สำหรับกรณีที่ไม่มีความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ของแผนภูมิควบคุมทั้งสองวิธีมีค่าใกล้เคียงกันและมีการบิดเบือนเล็กน้อยไปจาก 0.05 และเมื่อตัวอย่างขนาดใหญ่ขึ้นความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ของการจำลองจะมีค่าใกล้เคียง 5% มากขึ้น ในกรณีค่านอกกลุ่มปรากฏ 2% และ 5% ในตัวอย่าง training พบว่าความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ของแผนภูมิที่ไม่มีความแกร่งจะเข้าใกล้ค่า 0 ขณะที่ ความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ของวิธีที่มีความแกร่งมีการบิดเบือนเล็กน้อยไปจาก 0.05 เมื่อ พิจารณากำลังจากตัวอย่าง training ไม่มีค่านอกกลุ่ม ทั้งสองวิธีจะมีกำลังในการตรวจจับค่านอกกลุ่มที่สร้างขึ้นใกล้เคียงกัน แต่เมื่อมีค่านอกกลุ่ม 2% หรือ 5% ปรากฏในตัวอย่าง training พบว่า Holt-Winters ที่มีความแกร่งมีประสิทธิภาพดีกว่าวิธีมาตรฐานอย่างชัดเจน ความแตกต่างของกำลังประมาณ 10% และเมื่อพิจารณาอัตราการตรวจจับที่ผิดพลาดของทั้ง 2 แผนภูมิควบคุม เมื่อข้อมูล training ไม่มีค่านอกกลุ่มเพื่อกำหนดให้ α ใกล้เคียง 0.05 และตัวอย่างทดสอบมีค่านอกกลุ่ม 10 % ของข้อมูลโดยกำหนดขนาดค่านอกกลุ่ม $w=5,10,15$ และ 20 พบว่าอัตราการตรวจจับที่ผิดพลาดไม่ขึ้นอยู่กับจำนวนตัวอย่างใน training period นอกจากนี้สำหรับ $w=5$ วิธีที่มีความแกร่งมีประสิทธิภาพดีกว่าวิธีที่ไม่มีความแกร่งเพียงเล็กน้อย แต่เมื่อขนาดของค่านอกกลุ่ม (w) มีค่ามากขึ้น วิธีที่มีความแกร่งจะมีอัตราการเตือนที่ผิดพลาดต่ำกว่ามาก

นอกจากนี้ได้พิจารณาวิธีที่นำเสนอเกี่ยวกับตัวอย่างข้อมูลจริงซึ่งเป็นข้อมูลรายเดือนเกี่ยวกับจำนวนที่อยู่อาศัยที่กำลังก่อสร้างและจำนวนที่อยู่อาศัยที่สร้างสมบูรณ์แล้ว สำหรับจำนวนที่อยู่อาศัยที่กำลังก่อสร้างมีค่านอกกลุ่มขนาดใหญ่ 2 ค่าใน training period และมีค่านอกกลุ่มขนาดใหญ่ 1 ค่าใน test period พบว่าวิธีที่มีความแกร่งพยากรณ์ได้ดีกว่าวิธีมาตรฐานเมื่อพิจารณาจากความคลาดเคลื่อนในการพยากรณ์ one-step-ahead กำลังสองเฉลี่ย (MSFE) และพบว่าขีดจำกัดควบคุมของแผนภูมิควบคุมที่ไม่มีความแกร่งกว้าง

กว่าแผนภูมิควบคุมวิธีที่มีความแรงทำให้มีเพียงแผนภูมิควบคุมวิธีที่มีความแรงเท่านั้นที่สามารถตรวจจับค่านอกกลุ่มได้ และสำหรับจำนวนที่อยู่อาศัยที่สร้างสมบูรณ์ไม่ได้แสดงค่านอกกลุ่มขนาดใหญ่ในทั้ง 2 ส่วนของอนุกรมเวลา พบว่าวิธีที่มีความแรงและวิธีมาตรฐานมีประสิทธิภาพในการพยากรณ์อนุกรมเวลาได้ใกล้เคียงกันเมื่อพิจารณาจากความคลาดเคลื่อนในการพยากรณ์ one-step-ahead กำลังสองเฉลี่ย (MSFE) และยังพบว่าแผนภูมิควบคุมมาตรฐานและที่มีความแรงมีลักษณะคล้ายคลึงกัน เนื่องจากอนุกรมเวลาไม่ได้ประกอบด้วยค่านอกกลุ่มขนาดใหญ่ แต่อย่างไรก็ตามแผนภูมิควบคุมที่มีความแรงมีขีดจำกัดควบคุมแคบกว่าเล็กน้อย โดยร้อยละของค่าสังเกตที่อยู่นอกขีดจำกัดควบคุมสำหรับวิธีมาตรฐานเท่ากับ 2% ซึ่งดูเหมือนว่าจะดีเกินไป และสำหรับวิธีที่มีความแรงเท่ากับ 6.1% ซึ่งใกล้เคียงกับความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 คือ $\alpha = 5\%$ ดังนั้นวิธีที่มีความแรงที่สร้างขึ้นเป็นวิธีที่มีความน่าเชื่อถือมากกว่าวิธีที่ไม่มีความแรงเมื่อมีค่านอกกลุ่มปรากฏ โดยเฉพาะเมื่อมีค่านอกกลุ่มปรากฏในช่วง training

บทที่ 3

วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแผนภูมิควบคุมที่มีความแข็งแกร่งและมีประสิทธิภาพในการตรวจจับค่าผิดปกติในอนุกรมเวลาจริงที่เปรียบเทียบประสิทธิภาพในการตรวจจับค่าผิดปกติระหว่างแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแข็งแกร่งกับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมาตรฐาน และศึกษาผลกระทบของค่าผิดปกติที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบสำหรับวิธีการพยากรณ์ของแผนภูมิควบคุมทั้งสองแบบ ภายใต้สถานการณ์ที่แตกต่างกันของตัวแบบอนุกรมเวลา ARMA ที่คงที่ ค่าวัดประสิทธิภาพของแผนภูมิประกอบด้วยความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1, กำลังการทดสอบ (power of the test) และอัตราการเตือนผิด (false alarm rate) โดยใช้โปรแกรม R เวอร์ชัน 3.0.1 สำหรับการจำลองข้อมูล การพยากรณ์ข้อมูลตลอดจนการสร้างแผนภูมิควบคุมเพื่อตรวจจับค่าผิดปกติ ผู้วิจัยมีขั้นตอนในการดำเนินงานวิจัยดังนี้

1. จำลองข้อมูลอนุกรมเวลาจริงที่ใช้ตัวแบบคือ AR(1) ด้วยพารามิเตอร์ $\phi=0.1, 0.5$ และ 0.9 และความคลาดเคลื่อนสุ่ม $\varepsilon_t \sim \text{iid}N(0,1)$ ขนาดของอนุกรมเวลา $T=100$ และ 200
2. แบ่งข้อมูลออกเป็น 2 ช่วงเวลาคือ training period จำนวน $n = 50, 100$ โดยใช้ตัวอย่าง 10 ค่าแรกเป็น start period เพื่อใช้ในการกำหนดค่าเริ่มต้นและ test period จำนวน 50 และ 100 ตามลำดับ
3. สุ่มตำแหน่งเพื่อหาตำแหน่งของค่าผิดปกติ จำนวน 10% ของจำนวนตัวอย่างในช่วงเวลา training period หรือ test period เมื่อได้ตำแหน่งของค่าผิดปกติแล้วพิจารณาค่าสังเกต ณ ตำแหน่งที่ถูกสุ่มสร้างค่าผิดปกติที่มีขนาด 3, 5 และ 10 ถ้าค่าสังเกตเป็นค่าบวกค่าผิดปกติจะถูกสร้างขึ้นโดยค่าสังเกตบวกด้วยขนาดค่าผิดปกติ และถ้าค่าสังเกตเป็นค่าลบค่าผิดปกติจะถูกสร้างขึ้นโดยค่าสังเกตลบด้วยขนาดของค่าผิดปกติ เพื่อไม่ให้ขนาดของค่าผิดปกติลดลง
4. สร้างแผนภูมิควบคุมโดยพิจารณาอนุกรมเวลาในช่วง training period ; $t = 1, 2, 3, \dots, n$

4.1 แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล
มาตรฐาน

- (1) จำนวนค่าเริ่มต้น \hat{z}_{start} จากค่าเฉลี่ยของตัวอย่าง 10 ค่าแรก ; z_1, z_2, \dots, z_{10}
- (2) ทำการพยากรณ์อนุกรมเวลาล่วงหน้า 1 คาบเวลาโดยไม่พิจารณาช่วงเวลาเริ่มต้น $t = 1, 2, 3, \dots, 10$ โดยค่าพยากรณ์ ณ เวลา $t+1$ จำนวนได้จาก

$$\hat{z}_{t+1|t} = \lambda z_t + (1 - \lambda)\hat{z}_{t|t-1} ; t = 10, 11, \dots, n-1$$

- (3) จำนวนส่วนเหลือ $e_t = z_t - \hat{z}_{t|t-1}$
- (4) เลือกค่าพารามิเตอร์ที่ปรับให้เรียบ λ ที่ทำให้ผลบวกของส่วนเหลือกำลังสองมีค่าต่ำที่สุดนั่นคือ

$$\lambda^{opt} = \operatorname{argmin}_{\lambda} \sum_{t=11}^n e_t^2$$

- (5) หลังจากเลือกพารามิเตอร์ที่ปรับให้เรียบได้แล้วทำการพยากรณ์และคำนวณส่วนเหลือแล้วสร้างขีดจำกัดควบคุม

$$UCL = z_{\alpha/2} * S$$

$$CL = 0$$

$$LCL = -z_{\alpha/2} * S$$

โดยที่ S คือส่วนเบี่ยงเบนมาตรฐานของส่วนเหลือ เมื่อ $S^2 = \frac{\sum_{t=11}^n e_t^2}{n-10}$

$z_{\alpha/2}$ คือควอนไทล์ $(1 - \alpha/2)$ สำหรับการแจกแจงแบบปกติมาตรฐาน

ในงานวิจัยนี้กำหนดให้

$$\alpha=0.05 \text{ จะได้ว่า } z_{0.025} = 1.96$$

4.2 แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มี
ความแกร่ง

- (1) จำนวนค่าเริ่มต้น \hat{z}_{start} คำนวณจากมัธยฐานของตัวอย่าง 10 ตัวอย่างแรก ; z_1, z_2, \dots, z_{10} แทนการใช้ค่าเฉลี่ยของตัวอย่างเนื่องจากมัธยฐานเป็นตัวประมาณที่มีความแกร่ง

$$\text{และ } \hat{\sigma}_{start} = 1.4826 \operatorname{med}_{1 \leq t \leq 10} |z_t - \operatorname{med}_{1 \leq t \leq 10}(z_t)|$$

- (2) ทำการพยากรณ์อนุกรมเวลาล่วงหน้า 1 คาบเวลาโดยการแทนที่อนุกรมเวลา z_t ด้วยอนุกรมเวลาที่ไม่มีย่านออกกลุ่ม z_t^* บนพื้นฐานของ Huber ψ - function ซึ่ง

เป็นการแทนค่าสังเกตที่มีค่ามากเกินไปหรือค่าน้อยเกินไปด้วยค่าเหมาะสมกว่า
โดยค่าพยากรณ์ ณ เวลา $t+1$ คำนวณได้จาก

$$\hat{z}_{t+1|t} = \lambda z_t^* + (1 - \lambda)\hat{z}_{t|t-1} ; t=10, 11, \dots, n-1$$

โดยที่ $z_t^* = \psi_k \left(\frac{e_t}{\hat{\sigma}_t} \right) \hat{\sigma}_t + \hat{z}_{t|t-1}$

เมื่อ $\hat{\sigma}_t$ แทนตัวประมาณการกระจายโดยใช้ตัวประมาณค่าเบี่ยงเบน
สัมบูรณ์ของมัธยฐาน (MAD) โดยที่

$$\hat{\sigma}_t = 1.4826 \operatorname{med}_{11 \leq s \leq t} |e_s - \operatorname{med}_{10 \leq s \leq t}(e_s)|$$

และ $\psi_k(x)$ แทน Huber ψ - function ประยุกต์ใช้กับค่ามาตรฐานของ
ความคลื่อนในการพยากรณ์ล่วงหน้า 1 คาบเวลา

$$\text{โดยที่ } \psi_k(x) = \begin{cases} x & ; |x| < k \\ \operatorname{sign}(x)k & ; \text{ อื่นๆ} \end{cases}$$

ค่า k แทนค่าคงที่ (constant) โดยจะพบว่าถ้าความแตกต่างระหว่างค่า
สังเกต z_t กับค่าพยากรณ์ \hat{z}_t มีค่าน้อยจะได้ว่าค่าสังเกตที่ไม่มีค่านอกกลุ่ม (clean
value) z_t^* จะเท่ากับค่าสังเกต z_t แต่ถ้าความแตกต่างระหว่างค่าสังเกต z_t กับค่า
พยากรณ์ \hat{z}_t มีค่ามาก ค่าสังเกตนั้นจะพิจารณาเป็นค่านอกกลุ่ม และจะถูกแทนที่
ด้วยค่าที่มีขอบเขต (boundary value) ที่ขึ้นอยู่กับค่า k โดยในงานวิจัยนี้กำหนดค่า k
เท่ากับ 2 ตามกฎ 2 ซิกมา (Gelper, Fried and Croux :2010)

(3) คำนวณส่วนเหลือ $e_t = z_t - \hat{z}_{t|t-1}$

(4) เลือกค่าพารามิเตอร์ที่ปรับให้เรียบ λ ที่ทำให้ $s_0^2 \sum_{t=11}^n \gamma_k \left(\frac{e_t}{s_0} \right)$ มีค่าต่ำที่สุดนั้น
คือ

$$\lambda^{\text{opt}} = \operatorname{argmin}_{\lambda} s_0^2 \sum_{t=11}^n \gamma_k \left(\frac{e_t}{s_0} \right)$$

เมื่อ $s_0 = 1.4826 \operatorname{med}_{11 \leq t \leq n} |e_t|$

$$\gamma_k(y) = \min(k^2, y^2)$$

- (5) หลังจากเลือกพารามิเตอร์ที่ปรับให้เรียบได้แล้วทำการพยากรณ์และคำนวณส่วนเหลือแล้วสร้างขีดจำกัดควบคุม

$$UCL = z_{\alpha/2} * \tau$$

$$CL = 0$$

$$LCL = -z_{\alpha/2} * \tau$$

โดยที่ $\tau^2 = \frac{1}{n-10} s_0^2 \sum_{t=11}^n Y_k \left(\frac{e_t}{s_0} \right)$

$z_{\alpha/2}$ คือควอนไทล์ $(1 - \alpha/2)$ สำหรับการแจกแจงแบบปกติ

มาตรฐาน

ในงานวิจัยนี้กำหนดให้ $\alpha=0.05$ จะได้ว่า $z_{0.025} = 1.96$

สำหรับวิธีที่มีความแข็งแกร่งจะใช้ตัวประมาณ τ^2 แทนตัวประมาณ S ในการสร้างขีดจำกัดควบคุม เนื่องจาก τ^2 เป็นตัวประมาณที่มีความแกร่ง โดยที่จะให้น้ำหนักกับค่าความคลาดเคลื่อนในการพยากรณ์ที่มีค่ามาก ๆ ลดลง หรืออีกนัยก็คือผลกระทบของความคลาดเคลื่อนในการพยากรณ์ที่มีค่าสูงอันเนื่องมาจากอิทธิพลของค่านอกกลุ่มนั้นจะถูกจำกัด (limit) ซึ่งส่งผลให้ขีดจำกัดควบคุมที่สร้างขึ้นนี้ไม่ได้รับอิทธิพลจากค่านอกกลุ่ม

10. นำแผนภูมิควบคุมที่สร้างขึ้นทั้งสองวิธีมาตรวจจับค่านอกกลุ่มในอนุกรมเวลาช่วง test period

11. ทำซ้ำขั้นตอนที่ 1- 5 ในแต่ละพารามิเตอร์ที่กำหนดจำนวน 1,000 ครั้ง และคำนวณค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ค่าเฉลี่ยของความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1, กำลังของการทดสอบ (power of the test) และ อัตราการเตือนผิด (false alarm rate) สำหรับในแต่ละพารามิเตอร์ โดยที่

$$\text{ค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ} = \sum (\text{ค่าประมาณพารามิเตอร์ที่ปรับให้เรียบของครั้งที่ } i) / 1000$$

$$\text{ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1} = \sum \left(\frac{\text{จำนวนตัวอย่างที่ถูกตรวจจับเป็นค่านอกกลุ่มในกรณี test period ไม่มีค่านอกกลุ่ม}}{\text{จำนวนตัวอย่างทั้งหมดใน test period}} \right) / 1000$$

$$\text{กำลังของการทดสอบ} = \sum \left(\frac{\text{จำนวนค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับได้}}{\text{จำนวนค่านอกกลุ่มที่สร้างขึ้นทั้งหมดใน test period}} \right) / 1000$$

$$\text{อัตราการเตือนผิด} = \sum \left(\frac{\text{จำนวนตัวอย่างที่ไม่ได้เป็นค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับได้}}{\text{จำนวนตัวอย่างทั้งหมดใน test period ที่ไม่ได้เป็นค่านอกกลุ่มที่สร้างขึ้น}} \right) / 1000$$

12. เปรียบเทียบประสิทธิภาพของแผนภูมิควบคุมทั้งสองวิธีจากความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1, กำลังการทดสอบ (power of the test) และ อัตราการเตือนผิด (false alarm rate) สำหรับในแต่ละพารามิเตอร์ และศึกษาผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบสำหรับวิธีการพยากรณ์ของแผนภูมิควบคุมทั้งสองแบบ

บทที่ 4

ผลการวิเคราะห์ข้อมูล

งานวิจัยนี้มีจุดประสงค์เพื่อเสนอแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่ง เปรียบเทียบประสิทธิภาพในการตรวจจับค่านอกกลุ่มระหว่างแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งกับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมาตรฐาน โดยพิจารณาดังนี้

1. การวัดประสิทธิภาพ

ภายใต้เกณฑ์

- 1.1 ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 (Probability of type I error)
- 1.2 กำลังการทดสอบ (power of the test)
- 1.3 อัตราการเตือนผิด (false alarm rate)

2. ผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ(λ)

ผลการวิจัยพบว่า

1. การวัดประสิทธิภาพของแผนภูมิควบคุม

1.1 ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 (Probability of type I error)

ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 (Probability of type I error) คือความน่าจะเป็นที่ค่าสังเกตใน test period จะตกออกนอกขีดจำกัดควบคุมเมื่อใน test period ไม่มีค่านอกกลุ่มเกิดขึ้น นั่นคือสถานการณ์ที่ไม่มีค่านอกกลุ่มเกิดขึ้นใน test period ซึ่งแผนภูมิควบคุมที่มีประสิทธิภาพควรมีค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ใกล้เคียงกับ α

โดยแบ่งผลการวิเคราะห์ออกเป็น 2 กรณีคือ

- 1.1.1 กรณีไม่มีค่านอกกลุ่มใน training period
- 1.1.2 กรณีมีค่านอกกลุ่มใน training period

ตารางที่ 1 ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน

ขนาดตัวอย่าง	w_1	วิธีมาตรฐาน			วิธีที่มีความแกร่ง		
		$\phi = 0.1$	$\phi = 0.5$	$\phi = 0.9$	$\phi = 0.1$	$\phi = 0.5$	$\phi = 0.9$
100	ไม่มีค่านอกกลุ่ม	0.06514	0.06620	0.06364	0.08740	0.08844	0.08410
	3	0.00510	0.00672	0.00894	0.04134	0.04112	0.03918
	5	0.00030	0.00080	0.00430	0.04110	0.04186	0.03950
	10	0.00000	0.00004	0.00086	0.03928	0.04271	0.03698
200	ไม่มีค่านอกกลุ่ม	0.05701	0.05914	0.05681	0.07350	0.07612	0.07099
	3	0.00354	0.00415	0.00320	0.03008	0.02924	0.02240
	5	0.00010	0.00021	0.00118	0.02896	0.02964	0.02088
	10	0.00000	0.00000	0.00035	0.02866	0.02875	0.02902

ตารางที่ 1 แสดงความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ที่ระดับนัยสำคัญ $\alpha = 0.05$ สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน จำแนกตามขนาดตัวอย่าง ขนาดของค่านอกกลุ่ม (w_1) และพารามิเตอร์สหสัมพันธ์ (ϕ) พบว่า

1.1.1 กรณีไม่มีค่านอกกลุ่มใน training period

แผนภูมิควบคุมที่สร้างจากวิธีมาตรฐานมีความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ใกล้เคียงกับ α มากกว่าแผนภูมิควบคุมที่สร้างจากวิธีที่มีความแกร่ง สำหรับทุกพารามิเตอร์ ϕ (0.1 , 0.5 และ 0.9) โดยในแต่ละแผนภูมิควบคุมจะมีความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ใกล้เคียงกันสำหรับทุกพารามิเตอร์ ϕ (0.1 , 0.5 และ 0.9) นอกจากนี้ยังจะเห็นว่าแผนภูมิควบคุมทั้งสองแบบมีความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 มีค่าสูงกว่า α และเมื่อจำนวนตัวอย่างในอนุกรมเวลามากขึ้นจะทำให้ทั้งสองแผนภูมิควบคุมมีค่าของความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ใกล้เคียงกับ α มากขึ้น

1.1.2 กรณีมีค่านอกกลุ่มใน training period

แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่งมีความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ใกล้เคียงกับ α มากกว่าแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานสำหรับทุกพารามิเตอร์ ϕ (0.1 , 0.5 และ 0.9) และสำหรับทุกขนาดของค่านอกกลุ่ม (w_1) และเมื่อขนาดตัวอย่างในอนุกรมเวลาเพิ่มขึ้นจะพบว่าแผนภูมิควบคุมทั้งสองแบบมีความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ลดลงสู่ศูนย์มากขึ้น

สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานพบว่าเมื่อขนาดค่านอกกลุ่ม (w_1) เพิ่มขึ้นจะส่งผลให้ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 มีค่าลดลงเข้าใกล้ศูนย์ซึ่งเป็นลักษณะเช่นเดียวกันสำหรับทุกพารามิเตอร์ ϕ (0.1 , 0.5 และ 0.9) นอกจากนี้สำหรับแต่ละขนาดค่านอกกลุ่ม (w_1) เมื่อระดับของสหสัมพันธ์ในตัวเองของอนุกรมเวลาเพิ่มขึ้นพบว่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 มีค่าเพิ่มขึ้น

ในขณะที่แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่งพบว่าสำหรับทุกพารามิเตอร์ ϕ (0.1 , 0.5 และ 0.9) ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 จะมีค่าใกล้เคียงกันสำหรับขนาดค่านอกกลุ่ม (w_1) หนึ่งๆ และในแต่ละขนาดตัวอย่าง 100 และ 200

1.2 กำลังการทดสอบ (power of the test)

กำลังการทดสอบ (power of the test) คือความน่าจะเป็นที่ค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับได้นั้นคือสถานการณ์ที่มีค่านอกกลุ่มเกิดขึ้นใน test period

โดยแบ่งผลการวิเคราะห์ข้อมูลออกเป็น 3 กรณีคือ

1.2.1 ภาพรวม

1.2.2 กรณีไม่มีค่านอกกลุ่มใน training period

1.2.3 กรณีมีค่านอกกลุ่มใน training period

ตารางที่ 2 กำลังการทดสอบสำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบ
เอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน

ขนาด ตัวอย่าง	w ₁	w ₂	วิธีมาตรฐาน			วิธีที่มีความแกร่ง		
			$\phi = 0.1$	$\phi = 0.5$	$\phi = 0.9$	$\phi = 0.1$	$\phi = 0.5$	$\phi = 0.9$
100	ไม่มีค่า นอกกลุ่ม	3	0.9972	0.9558	0.8746	0.9990	0.9838	0.903
		5	0.9996	0.9858	0.9406	0.9988	0.9998	0.9889
		10	1.0000	0.9952	0.9484	1.0000	1.0000	1.0000
	3	3	0.8630	0.7810	0.5984	0.9862	0.9520	0.8138
		5	0.9988	0.9836	0.9280	1.0000	0.9994	0.9788
		10	1.0000	0.9966	0.9474	1.0000	1.0000	1.0000
	5	3	0.3440	0.3468	0.2306	0.9858	0.9526	0.7996
		5	0.9852	0.9648	0.8016	1.0000	0.9984	0.9782
		10	0.9996	0.9984	0.9500	1.0000	1.0000	1.0000
	10	3	0.0076	0.0058	0.0192	0.9844	0.9526	0.7846
		5	0.1708	0.1598	0.1444	0.9996	0.9986	0.9832
		10	0.9982	0.9932	0.9494	1.0000	1.0000	1.0000
200	ไม่มีค่า นอกกลุ่ม	3	0.9990	0.9496	0.8717	0.9998	0.9804	0.9019
		5	1.0000	0.9770	0.9247	1.0000	0.9999	0.9913
		10	1.0000	0.9982	0.9344	1.0000	1.0000	1.0000
	3	3	0.8846	0.7857	0.5725	0.9954	0.9555	0.8032
		5	0.9994	0.9815	0.9165	1.0000	0.9992	0.9827
		10	1.0000	0.9990	0.9285	1.0000	1.0000	0.9996
	5	3	0.3213	0.2923	0.1724	0.9956	0.9545	0.7792
		5	0.9945	0.9723	0.8066	0.9999	0.9996	0.9787
		10	1.0000	0.9989	0.9313	1.0000	1.0000	0.9999
	10	3	0.0007	0.0013	0.0092	0.9969	0.9557	0.7943
		5	0.0842	0.0094	0.1067	1.0000	0.9991	0.9785
		10	0.9994	0.9974	0.9425	1.0000	1.0000	1.0000

ตารางที่ 2 แสดงกำลังการทดสอบ สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับ ให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน จำแนกตามขนาดตัวอย่าง ขนาดของค่านอกกลุ่ม (w_1, w_2) และพารามิเตอร์สหสัมพันธ์ (ϕ) พบว่า

1.2.1 ภาพรวม

แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่งมีกำลังการทดสอบสูงกว่า แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐาน สำหรับทุกกรณีที่ทำการศึกษา นั่นคือแผนภูมิ ควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่งสามารถตรวจจับค่านอกกลุ่มได้ดีกว่าแผนภูมิ ควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานในทุก ๆ กรณี ยกเว้น กรณีขนาดตัวอย่าง 100 ไม่มีค่า นอกกลุ่มใน training period ขนาดของค่านอกกลุ่ม test period ขนาดปานกลาง ($w_2=5$) และ สหสัมพันธ์ในตัวเองระดับน้อย ($\phi=0.1$)

สำหรับแผนภูมิควบคุมทั้งสองแบบพบว่าถ้าระดับของสหสัมพันธ์ในตัวเองของอนุกรม เวลาเพิ่มขึ้นจะส่งผลให้กำลังการทดสอบลดลงสำหรับทุก ๆ กรณี ยกเว้นสำหรับแผนภูมิควบคุมที่ สร้างจากวิธีการพยากรณ์มาตรฐาน สำหรับขนาดตัวอย่าง 100 กรณีขนาดของค่านอกกลุ่มใน training period ขนาดใหญ่ ($w_1=10$) และขนาดของค่านอกกลุ่ม test period ขนาดเล็ก ($w_2=3$) และสำหรับขนาดตัวอย่าง 200 กรณีขนาดของค่านอกกลุ่มใน training period ขนาดใหญ่ ($w_1=10$) และขนาดของค่านอกกลุ่ม test period ขนาดเล็กและปานกลาง ($w_2=3,5$)

สำหรับแต่ละพารามิเตอร์ ϕ (0.1 , 0.5 และ 0.9) และแต่ละขนาดของค่านอกกลุ่มใน training period (ไม่มีค่านอกกลุ่มหรือ $w_1=3,5$ และ 10) เมื่อขนาดของค่านอกกลุ่ม test period (w_2) มีขนาดใหญ่ขึ้นพบว่าแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ทั้งสองวิธีมีกำลังการทดสอบ เพิ่มขึ้น

1.2.2 กรณีไม่มีค่านอกกลุ่มใน training period

แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐาน

พบว่ามีกำลังการทดสอบสูงกว่า 90% ยกเว้นกรณีที่ค่านอกกลุ่มใน test period มี ขนาดเล็ก ($w_2=3$) และอนุกรมมีสหสัมพันธ์ในตัวเองมาก ($\phi = 0.9$) จะมีกำลังการทดสอบเพียง 87%

แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่ง

พบว่ามีการทดสอบสูงกว่า 98% ยกเว้นในกรณีค่านอกกลุ่มใน test period มีขนาดเล็ก ($w_2=3$) และระดับของสหสัมพันธ์ในตัวเองของอนุกรมเวลาอยู่ในระดับมาก ($\phi = 0.9$) จะมีการทดสอบเพียง 90%

1.2.3 กรณีมีค่านอกกลุ่มใน training period

แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐาน

พบว่าเมื่อเกิดค่านอกกลุ่มใน training period จะส่งผลให้กำลังการทดสอบลดลง ซึ่งถ้าขนาดค่านอกกลุ่มใน training period เพิ่มมากขึ้นจะส่งผลให้กำลังการทดสอบลดลงมากขึ้น โดยเฉพาะถ้าขนาดของค่านอกกลุ่มใน training period (w_1) มีขนาดใหญ่กว่าขนาดของค่านอกกลุ่มใน test period (w_2) จะทำให้กำลังการทดสอบลดลงอย่างมาก ยิ่งไปกว่านั้นเมื่อเกิดค่านอกกลุ่มใน training period ขนาดใหญ่ ($w_1=10$) และเกิดค่านอกกลุ่มใน test period ขนาดเล็กและปานกลาง ($w_2=3,5$) จะพบว่ากำลังการทดสอบมีค่าต่ำมากนั่นคือแทบจะไม่สามารถตรวจจับค่านอกกลุ่มได้เลย

เมื่อสหสัมพันธ์ในตัวเองของอนุกรมเวลาอยู่ในระดับน้อย ($\phi = 0.1$) พบว่าถ้าขนาดตัวอย่างมากขึ้นจะส่งผลให้กำลังการทดสอบมีมากขึ้น ยกเว้นกรณีที่ขนาดของค่านอกกลุ่มใน training period (w_1) มีขนาดใหญ่กว่าขนาดของค่านอกกลุ่มใน test period (w_2) ($w_1=5$ และ $w_2=3$ กับ $w_1=10$ และ $w_2=3,5$) กรณีนี้จะพบว่ากำลังการทดสอบจะลดลงเมื่อขนาดตัวอย่างเพิ่มมากขึ้น

เมื่อสหสัมพันธ์ในตัวเองของอนุกรมเวลาอยู่ในระดับปานกลาง ($\phi = 0.5$) พบว่ากรณีเกิดค่านอกกลุ่มใน test period มีขนาดใหญ่ ($w_2=10$) กำลังการทดสอบจะเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ขณะที่เมื่อเกิดค่านอกกลุ่มใน test period มีขนาดปานกลาง ($w_2=5$) จะมีการทดสอบจะลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

และเมื่อสหสัมพันธ์ในตัวเองของอนุกรมเวลาอยู่ในระดับมาก ($\phi = 0.9$) สำหรับทุก ๆ ขนาดของค่านอกกลุ่ม (training period และ test period) พบว่าเมื่อขนาดตัวอย่างเพิ่มขึ้นจะทำให้กำลังการทดสอบลดลง ยกเว้นกรณีที่ค่านอกกลุ่ม training period และ test period มีขนาดปานกลาง ($w_1, w_2=5$)

แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่ง

พบว่าค่านอกกลุ่มใน training period ไม่ได้ส่งผลกระทบต่อกำลังการทดสอบ โดยสำหรับทุกๆขนาดของค่านอกกลุ่มใน training period และ test period ($w_1=3,5$ และ 10 , $w_2=3,5$ และ 10) และทุกพารามิเตอร์ ϕ (0.1 , 0.5 และ 0.9) จะมีกำลังการทดสอบมากกว่า 95% ยกเว้นในกรณีค่านอกกลุ่มใน test period มีขนาดเล็ก ($w_2=3$) และระดับของสหสัมพันธ์ในตัวเองของอนุกรมเวลาอยู่ในระดับมาก ($\phi = 0.9$) จะมีกำลังการทดสอบเพียง 77% ถึง 81%

1.3 อัตราการเตือนผิด (false alarm rate)

อัตราการเตือนผิด (false alarm rate) คือร้อยละของจำนวนตัวอย่างใน test period ที่ไม่ได้เป็นค่านอกกลุ่มที่สร้างขึ้นแต่ถูกตรวจจับเป็นค่านอกกลุ่มเมื่อมีค่านอกกลุ่มเกิดขึ้นใน test period โดยในกรณีวัดประสิทธิภาพจากอัตราการเตือนผิดจะกำหนดให้อนุกรมเวลาในส่วน training period ไม่เกิดค่านอกกลุ่มขึ้น เพื่อให้ระดับนัยสำคัญของแผนภูมิควบคุมทั้งสองแบบมีค่าใกล้เคียงกับ $\alpha=0.05$

ตารางที่ 3 แสดงอัตราการเตือนผิด กำหนดระดับนัยสำคัญ $\alpha = 0.05$ สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน จำแนกตามขนาดตัวอย่าง ขนาดของค่านอกกลุ่ม (w_2) และพารามิเตอร์สหสัมพันธ์ (ϕ) พบว่า

ตารางที่ 3 อัตราการเตือนผิดสำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบ
เอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน

ขนาดตัวอย่าง	w_2	วิธีมาตรฐาน			วิธีที่มีความแกร่ง		
		$\phi = 0.1$	$\phi = 0.5$	$\phi = 0.9$	$\phi = 0.1$	$\phi = 0.5$	$\phi = 0.9$
100	3	0.07500	0.11307	0.13529	0.09158	0.10938	0.12149
	5	0.09998	0.15429	0.16091	0.09124	0.10836	0.12669
	10	0.13964	0.20956	0.26796	0.09489	0.10449	0.12642
200	3	0.06556	0.10452	0.12952	0.07658	0.09171	0.11120
	5	0.07651	0.14710	0.14996	0.07624	0.08883	0.11324
	10	0.11079	0.20811	0.25776	0.07338	0.09131	0.11791

ในกรณีที่มีค่านอกกลุ่มใน test period สำหรับขนาดตัวอย่างเท่ากับ 100 และ 200 พารามิเตอร์ $\phi = 0.1, 0.5$ และ 0.9 และ ขนาดของค่านอกกลุ่ม $w_2 = 3, 5$ และ 10 พบว่าอัตราการเตือนผิดของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานสูงกว่าอัตราการเตือนผิดของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่ง (ยกเว้นกรณี $\phi = 0.1$ และ $w_2 = 3$)

เมื่อขนาดของค่านอกกลุ่ม (w_2) มีขนาดใหญ่ขึ้นจะส่งผลให้แผนภูมิควบคุมทั้งสองมีอัตราการเตือนผิดแตกต่างกันมากขึ้น นั่นคือสำหรับแต่ละพารามิเตอร์ ϕ ($0.1, 0.5$ และ 0.9) จะเห็นว่าเมื่อค่านอกกลุ่มมีขนาดเพิ่มมากขึ้นจะส่งผลให้อัตราการเตือนผิดของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานมีค่าสูงขึ้น ในขณะที่อัตราการเตือนผิดของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่งมีค่าใกล้เคียงกันในทุกขนาดของค่านอกกลุ่ม (w_2)

เมื่อขนาดตัวอย่างเพิ่มมากขึ้นจาก 100 เป็น 200 แผนภูมิควบคุมที่สร้างขึ้นจากทั้ง 2 วิธีต่างมีอัตราการเตือนผิดลดลง นอกจากนี้ยังพบว่าเมื่ออนุกรมเวลามีระดับของสหสัมพันธ์ในตัวเองสูงขึ้น แผนภูมิควบคุมทั้ง 2 แบบจะมีอัตราการเตือนผิดเพิ่มสูงขึ้น

2. ผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ

ค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับแผนภูมิควบคุมทั้งสองวิธีจะถูกกำหนดขึ้นจากอนุกรมเวลาในช่วง training period และนำค่าพารามิเตอร์ที่ปรับให้เรียบที่เหมาะสมที่ได้นำมาใช้ในการพยากรณ์อนุกรมเวลาทั้งในช่วง training period และ test period ดังนั้นถ้าค่าประมาณพารามิเตอร์ที่ปรับให้เรียบเป็นค่าประมาณที่มีความเอนเอียง อาจส่งผลให้ค่าพยากรณ์ที่ได้ไม่เหมาะสมกับข้อมูล และเมื่อนำความคลาดเคลื่อนในการพยากรณ์มาสร้างแผนภูมิควบคุมและใช้ในการตรวจจับค่านอกกลุ่มอาจทำให้เกิดความผิดพลาด

ดังนั้นการศึกษาผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับแผนภูมิควบคุมทั้งสองแบบจะช่วยให้เห็นถึงประสิทธิภาพของแผนภูมิควบคุมมากยิ่งขึ้น

ตารางที่ 4 ค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน

ขนาดตัวอย่าง	w_1	วิธีมาตรฐาน			วิธีที่มีความแกร่ง		
		$\phi = 0.1$	$\phi = 0.5$	$\phi = 0.9$	$\phi = 0.1$	$\phi = 0.5$	$\phi = 0.9$
100	ไม่มีค่านอกกลุ่ม	0.14105	0.57445	0.9209	0.19565	0.59535	0.92585
	3	0.11725	0.35665	0.84770	0.16755	0.46485	0.87915
	5	0.10195	0.23255	0.76400	0.15345	0.45360	0.87085
	10	0.09630	0.01436	0.45285	0.14065	0.42730	0.87445
200	ไม่มีค่านอกกลุ่ม	0.09190	0.60080	0.93140	0.12465	0.61105	0.93530
	3	0.07670	0.35050	0.88560	0.11465	0.47100	0.90100
	5	0.07003	0.20830	0.80345	0.10405	0.45250	0.89505
	10	0.06515	0.10510	0.49905	0.09385	0.44090	0.89590

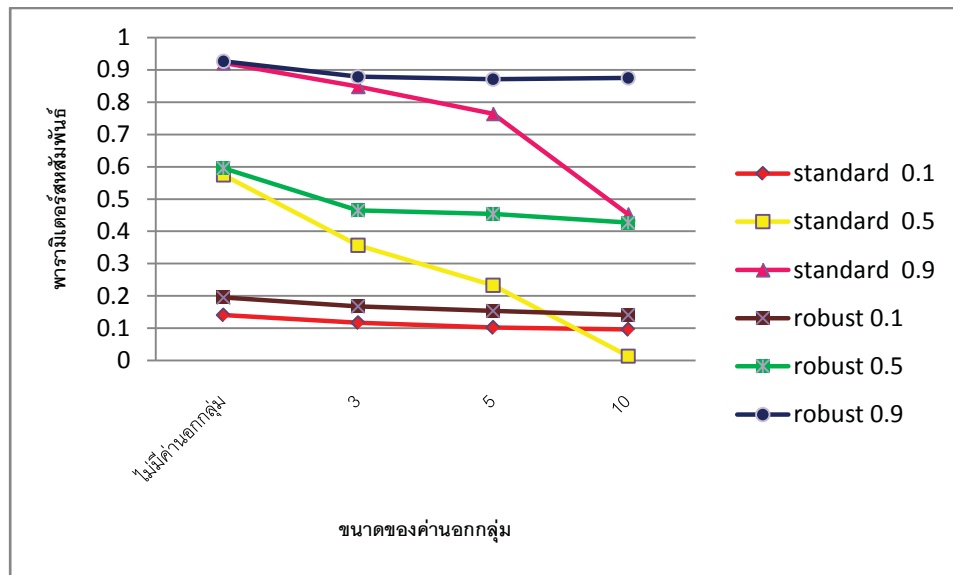
ตารางที่ 4 แสดงค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน จำแนกตามขนาดตัวอย่าง ขนาดของค่านอกกลุ่ม (w_1) และพารามิเตอร์สหสัมพันธ์ (ϕ) พบว่า

กรณีไม่มีค่านอกกลุ่มใน training period

สำหรับขนาดตัวอย่างเท่ากับ 100 พบว่าค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมทั้งสองแบบมีค่าใกล้เคียงกัน โดยค่าประมาณพารามิเตอร์ที่ปรับให้เรียบของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแรงมีค่าสูงกว่าเล็กน้อย

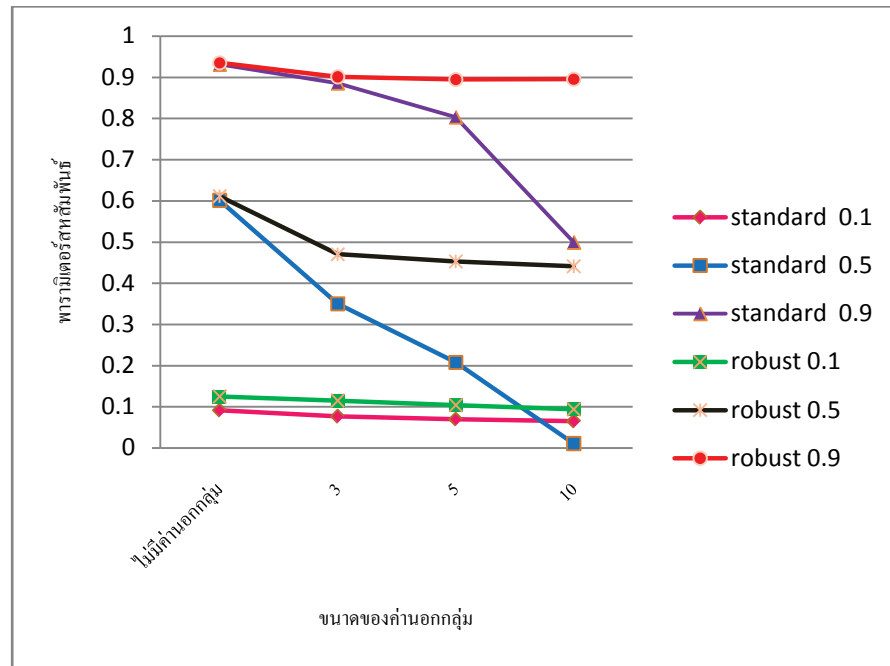
กรณีที่มีค่านอกกลุ่มใน training period

สำหรับขนาดตัวอย่างเท่ากับ 100 พบว่าค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานสำหรับแต่ละพารามิเตอร์ ϕ (0.1 , 0.5 และ 0.9) จะมียาลดลงสู่ศูนย์เมื่อขนาดของค่านอกกลุ่ม (w_1) มีขนาดใหญ่ขึ้น(ภาพที่1) ในขณะที่ค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแรงสำหรับแต่ละพารามิเตอร์ ϕ (0.1 , 0.5 และ 0.9) มีค่าใกล้เคียงกันและใกล้เคียงกันกับค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ในกรณีไม่มีค่านอกกลุ่มปรากฏใน training period (ภาพที่ 1)



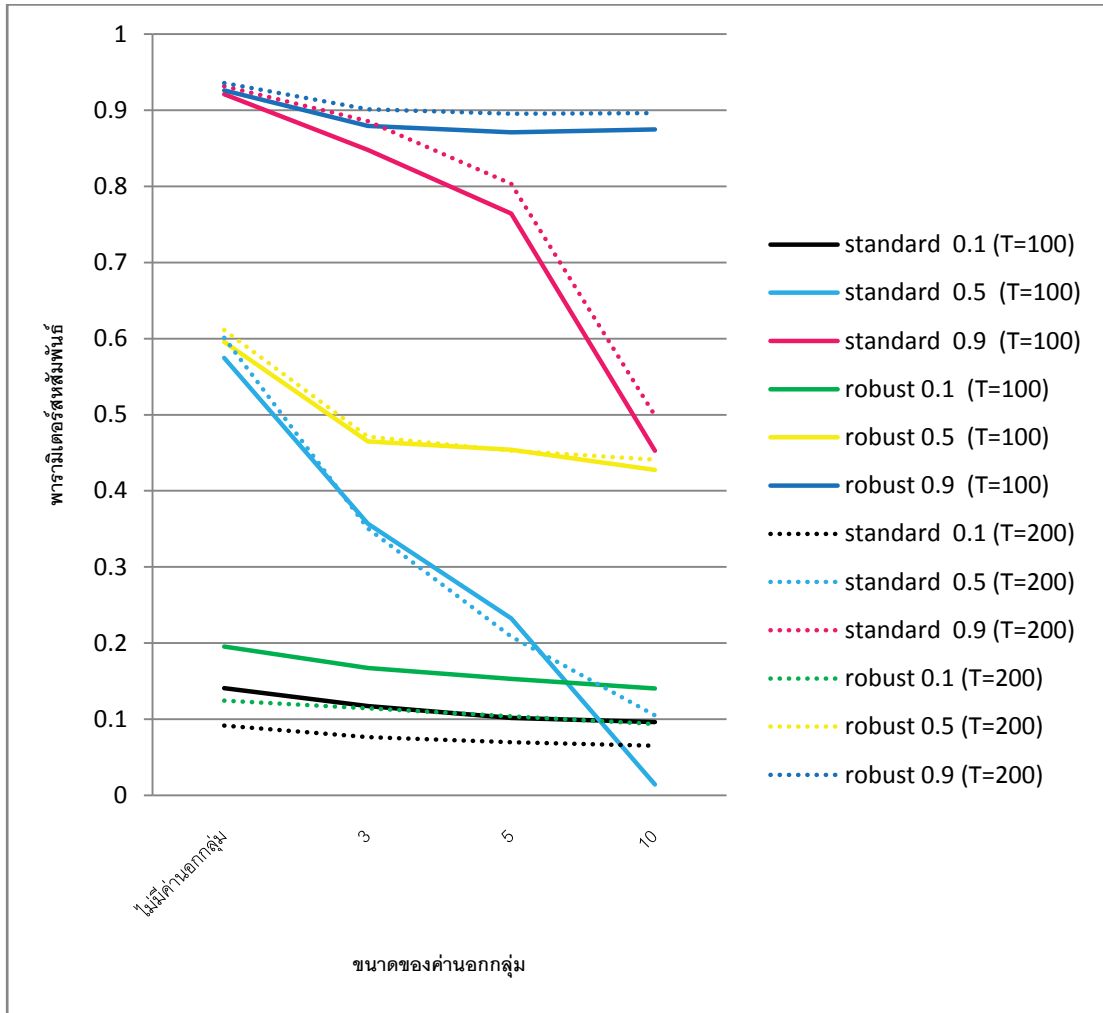
ภาพที่ 1 เปรียบเทียบค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแรงและวิธีมาตรฐาน สำหรับขนาดตัวอย่างเท่ากับ 100 จำแนกตามขนาดของค่านอกกลุ่ม (w_1) และพารามิเตอร์สหสัมพันธ์ (ϕ)

สำหรับขนาดตัวอย่างเท่ากับ 200 พบว่าค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) มีลักษณะเช่นเดียวกับกรณีขนาดตัวอย่างเท่ากับ 100 (ภาพที่ 2)



ภาพที่ 2 เปรียบเทียบค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน สำหรับขนาดตัวอย่างเท่ากับ 200 จำแนกตามขนาดของค่าผิดปกติ (w_1) และพารามิเตอร์สหสัมพันธ์ (ϕ)

นอกจากนี้ยังพบว่า สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานและวิธีที่มีความแกร่งสำหรับขนาดตัวอย่างของอนุกรมเวลา 100 และ 200 มีค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ค่าใกล้เคียงกันสำหรับแต่ละพารามิเตอร์ ϕ (0.5 และ 0.9) ยกเว้นในกรณี $\phi=0.1$ ค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) มีค่าลดลง เมื่อขนาดตัวอย่างของอนุกรมเวลาเพิ่มขึ้น ทั้งในกรณีที่มีค่าผิดปกติและไม่มีค่าผิดปกติใน training period (ภาพที่ 3)



ภาพที่ 3 เปรียบเทียบค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน ระหว่างขนาดตัวอย่างเท่ากับ 100 และ 200 จำแนกตามขนาดของค่านอกกลุ่ม(w_1) และพารามิเตอร์สหสัมพันธ์ (ϕ)

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเสนอแผนภูมิควบคุมที่มีความแข็งแกร่งและมีประสิทธิภาพในการตรวจจับค่านอกกลุ่มในอนุกรมเวลาคงที่ เพื่อใช้เป็นเครื่องมือในการตรวจจับค่านอกกลุ่มที่มีประสิทธิภาพ โดยเป็นวิธีที่ง่ายและไม่ซับซ้อนเมื่อนำไปใช้งาน และเปรียบเทียบประสิทธิภาพของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแข็งแกร่งกับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมาตรฐาน ผลการวิจัย

1. การวัดประสิทธิภาพของแผนภูมิควบคุม

ภายใต้เกณฑ์

- 1.1 ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 (Probability of type I error)
- 1.2 กำลังการทดสอบ (power of the test)
- 1.3 อัตราการเตือนผิด (false alarm rate)

จากเกณฑ์การวัดประสิทธิภาพข้างต้น แผนภูมิควบคุมที่มีประสิทธิภาพที่ดีสำหรับแต่ละกรณีที่ทำการศึกษา จำแนกตามเกณฑ์การวัดประสิทธิภาพ ขนาดตัวอย่าง ขนาดค่านอกกลุ่ม (w_1, w_2) และพารามิเตอร์ของระดับสหสัมพันธ์ในตัวเอง (ϕ) สามารถสรุปได้ตามตารางที่ 5 ดังนี้

ตารางที่ 5 สรุปแผนภูมิควบคุมที่มีประสิทธิภาพที่ดีแต่ละกรณีที่ทำการศึกษา

เกณฑ์การวัด ประสิทธิภาพ	ขนาด ตัวอย่าง	w ₁	w ₂	Φ		
				0.1	0.5	0.9
ความน่าจะเป็น ที่จะเกิดความ คลาดเคลื่อน ประเภทที่ 1	100	ไม่มีค่านอกกลุ่ม	ไม่มีค่านอกกลุ่ม	standard	standard	Standard
		3		robust	robust	robust
		5		robust	robust	robust
		10		robust	robust	robust
	200	ไม่มีค่านอกกลุ่ม	ไม่มีค่านอกกลุ่ม	standard	standard	Standard
		3		robust	robust	robust
		5		robust	robust	robust
		10		robust	robust	robust
กำลัง การทดสอบ	100	ไม่มีค่านอกกลุ่ม	3	robust	robust	robust
			5	standard	robust	robust
			10	standard,robust	robust	robust
		3	3	robust	robust	robust
			5	robust	robust	robust
			10	standard,robust	robust	robust
		5	3	robust	robust	robust
			5	robust	robust	robust
			10	robust	robust	robust
	10	3	robust	robust	robust	
		5	robust	robust	robust	
		10	robust	robust	robust	

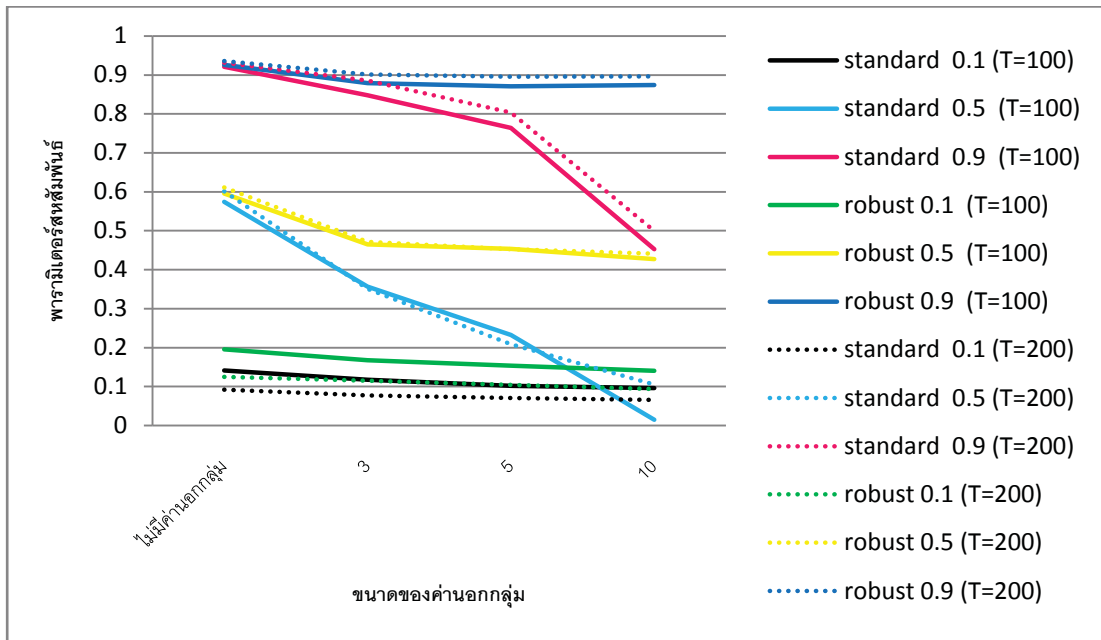
เกณฑ์การวัด ประสิทธิภาพ	ขนาด ตัวอย่าง	w ₁	w ₂	φ				
				0.1	0.5	0.9		
กำลัง การทดสอบ	200	ไม่มีค่านอกกลุ่ม	3	robust	robust	robust		
			5	standard, robust	robust	robust		
			10	standard, robust	robust	robust		
		3	3	robust	robust	robust		
			5	robust	robust	robust		
			10	standard, robust	robust	robust		
		5	3	robust	robust	robust		
			5	robust	robust	robust		
			10	standard, robust	robust	robust		
		10	3	robust	robust	robust		
			5	robust	robust	robust		
			10	robust	robust	robust		
		อัตรา การเตือนผิด	100	ไม่มีค่านอกกลุ่ม	3	standard	robust	robust
					5	robust	robust	robust
					10	robust	robust	robust
200	ไม่มีค่านอกกลุ่ม		3	standard	robust	robust		
			5	robust	robust	robust		
			10	robust	robust	robust		

จากตารางที่ 5 สามารถสรุปได้ดังตารางที่ 6
 ตารางที่ 6 สรุปแผนภูมิควบคุมที่มีประสิทธิภาพที่ดีแต่ละสถานการณ์โดยภาพรวม

ค่านอกกลุ่ม		Φ		
Training period	Test period	0.1	0.5	0.9
ไม่มี	ไม่มี	standard	standard	Standard
ไม่มี	มี	robust	robust	robust
		ยกเว้นกรณีขนาดค่านอกกลุ่มใน test period ($W_2=3$) และ $\Phi = 0.1$		
มี	ไม่มี	robust	robust	robust
มี	มี	robust	robust	robust

2. ผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ

จากการศึกษาผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สามารถสรุปได้ดังภาพที่ 4 ซึ่งจะเห็นว่ากรณีไม่มีค่านอกกลุ่มเกิดขึ้นในอนุกรมเวลา ค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับแผนภูมิควบคุมทั้งสองแบบมีค่าใกล้เคียงกัน แต่เมื่อเกิดค่านอกกลุ่มขึ้นค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐาน จะเกิดความเอนเอียงลดลงสู่ศูนย์ ในขณะที่ค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่ง คูเหมือนจะไม่ได้รับผลกระทบจากการเกิดค่านอกกลุ่มซึ่งมีค่าประมาณพารามิเตอร์ที่ปรับให้เรียบค่อนข้างคงที่



ภาพที่ 4 สรุปผลกระทบของค่านอกกลุ่มที่มีต่อค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่งและวิธีมาตรฐาน

อภิปรายผล

จากการวิจัยครั้งนี้พบว่าในกรณีที่ไม่มีค่านอกกลุ่มเกิดขึ้นในอนุกรมเวลา (ทั้งใน training period และ test period) อาจกล่าวได้ว่าแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานมีประสิทธิภาพดีกว่าแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่ง ซึ่งสอดคล้องกับ Croux , Gelper และ Mahieu (2011) เนื่องจากแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแกร่งใช้ตัวประมาณการกระจาย τ แทนตัวประมาณ S นำมาสร้างขีดจำกัดควบคุมซึ่งเป็นการลดอิทธิพลของความคลาดเคลื่อนในการพยากรณ์ที่มีค่าสูงหรือต่ำเกินไปถูกจำกัด ส่งผลให้ขอบเขตควบคุมที่สร้างขึ้นแคบกว่าขอบเขตควบคุมของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐาน และส่งผลให้ทำการตรวจจับค่าสังเกตที่ไม่ได้เป็นค่านอกกลุ่มว่าเป็นค่านอกกลุ่มมีจำนวนมากกว่า นอกจากนี้จะเห็นว่าแผนภูมิควบคุมทั้งสองแบบมีค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) ใกล้เคียงกัน

แต่เมื่อเกิดค่านอกกลุ่มขึ้นในอนุกรมเวลาไม่ว่าจะเป็นค่านอกกลุ่มที่เกิดขึ้นใน training period และ/หรือ test period พบว่าแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์แกร่งมีประสิทธิภาพดีกว่าแผนภูมิ

ควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐาน และค่าประมาณพารามิเตอร์ที่ปรับให้เรียบ (λ) สำหรับ แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานเกิดความเอนเอียงสู่ศูนย์ สอดคล้องกับ Gelper , Fried และ Croux (2010) ซึ่งสาเหตุที่เป็นเช่นนี้เนื่องจากเมื่อเกิดค่านอกกลุ่มขึ้นก็จะส่งผลกระทบต่อแผนภูมิ ควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานดังนี้

ประการแรกคือค่าพารามิเตอร์ที่ปรับให้เรียบ (λ) เกิดความเอนเอียง เนื่องจากค่าพารามิเตอร์ที่ปรับให้เรียบ (λ) จะถูกเลือกจากค่าที่ทำให้ความคลาดเคลื่อนในการพยากรณ์มีค่าต่ำที่สุด เมื่อเกิดค่านอกกลุ่มขึ้นจะส่งผลให้ค่าพยากรณ์ล่วงหน้าขึ้นอยู่กับค่าปัจจุบัน(ที่มีค่านอกกลุ่มประกอบอยู่ด้วย)น้อยลง เพื่อลดอิทธิพลของค่านอกกลุ่มที่จะส่งผลต่อค่าพยากรณ์ล่วงหน้าทำให้ความคลาดเคลื่อนในการพยากรณ์ลดน้อยลง ทำให้พารามิเตอร์ที่ปรับให้เรียบที่เหมาะสมมีความเอนเอียงสู่ศูนย์

ประการที่สองคือแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐานไม่สามารถตรวจจับค่า นอกกลุ่มได้ สาเหตุ เนื่องจากการกำหนดขีดจำกัดควบคุมของแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ มาตรฐานขึ้นอยู่กับส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ของความคลาดเคลื่อนในการพยากรณ์ ซึ่งส่วนเบี่ยงเบนมาตรฐานเป็นตัวประมาณการกระจายที่ไม่มีความแปรผัน ตามที่ Abu-Shawiesh (2008) กล่าวไว้ ดังนั้น เมื่อเกิดค่านอกกลุ่มแม้เพียงค่าเดียวก็จะส่งผลให้ตัวประมาณการกระจายมีค่าสูงขึ้น และทำให้ขอบเขตควบคุมกว้างขึ้น ทำให้ไม่สามารถตรวจจับค่านอกกลุ่มอย่างทันทั่วทั้งที่

และประการสุดท้ายคือเกิดการเตือนผิดสูง เนื่องจาก ค่านอกกลุ่มจะส่งผลต่อค่าพยากรณ์ล่วงหน้าจึงอาจส่งผลให้เกิดความคลาดเคลื่อนมีค่าสูง/ต่ำมากเกินไปจนเกินความเป็นจริง เมื่อนำค่าความคลาดเคลื่อนไปใช้ในการตรวจจับค่านอกกลุ่มจะส่งผลให้เกิดลำดับการเตือนผิดและทำให้อัตราการเตือนผิดสูงมากขึ้น โดยเฉพาะเมื่อขนาดของค่านอกกลุ่มมากขึ้นจะทำให้ค่าพยากรณ์ได้รับอิทธิพลจากค่านอกกลุ่มเพิ่มมากขึ้น

ในขณะที่แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์ที่มีความแปรผัน ค่าพยากรณ์จะขึ้นอยู่กับค่าสังเกตในปัจจุบันที่อิทธิพลของค่านอกกลุ่มถูกจำกัดจึงทำให้ค่าพยากรณ์ได้รับอิทธิพลจากค่านอกกลุ่มที่เกิดขึ้นน้อย และการกำหนดขีดจำกัดควบคุมใช้ตัวประมาณการกระจาย τ ที่มีความแปรผัน นำมาสร้างขีดจำกัดควบคุมซึ่งเป็นการลดอิทธิพลของความคลาดเคลื่อนในการพยากรณ์ที่ได้รับอิทธิพลจากค่านอกกลุ่ม ทำให้มีค่าสูง/ต่ำเกินไปถูกจำกัด ส่งผลให้ขอบเขตควบคุมที่สร้างขึ้นไม่กว้างเกินกว่าที่ควรจะเป็น ทำให้สามารถตรวจจับค่านอกกลุ่มได้อย่างทันทั่วทั้งที่ และเมื่อค่าพยากรณ์ได้รับอิทธิพลจากค่านอกกลุ่มที่เกิดขึ้นน้อย ก็จะส่งผลให้ค่าความคลาดเคลื่อนในการพยากรณ์มีค่าสูง /ต่ำมากเกินไปจนเกินความเป็นจริง ทำให้อัตราการเตือนผิดด้วย และยังพบว่าค่าพารามิเตอร์ที่ปรับให้เรียบ (λ) มีค่าใกล้เคียงกันเมื่อมีค่านอกกลุ่มเกิดขึ้นใน training period และยังมีค่าใกล้เคียงกับกรณีไม่เกิดค่านอกกลุ่มขึ้นใน training period อีกด้วย

นอกจากนี้ยังพบว่าระดับของสหสัมพันธ์ในตัวเองของอนุกรมเวลาก็ยังส่งผลต่อประสิทธิภาพของแผนภูมิควบคุมด้วยเช่นกัน นั่นคือเมื่อระดับของสหสัมพันธ์ในตัวเองเพิ่มขึ้นจะส่งผลให้ประสิทธิภาพของแผนภูมิควบคุมทั้งสองแบบลดลง แต่ไม่ว่าระดับของสหสัมพันธ์ในตัวเองของอนุกรมเวลาจะมากหรือน้อย ก็ยังพบว่าเมื่อเกิดค่านอกกลุ่มขึ้นใน training period และ/หรือ test period แผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์แกร่งก็มีประสิทธิภาพดีกว่าแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์มาตรฐาน

ข้อเสนอแนะ

1. ในงานวิจัยนี้มุ่งเน้นศึกษาอนุกรมเวลาทางด้านกระบวนการทางอุตสาหกรรมซึ่งเป็นกระบวนการคงที่และมีสหสัมพันธ์เชิงบวก ตามที่ Bisgaard (2005) ได้กล่าวไว้ ซึ่งถ้าต้องการตรวจจับค่า นอกกลุ่มในอนุกรมเวลาลักษณะอื่น ๆ เช่น อนุกรมเวลาที่มีแนวโน้ม ฤดูกาล สามารถนำแนวคิดดังกล่าวนี้ไปประยุกต์ใช้ร่วมกับวิธีการพยากรณ์ที่เหมาะสมกับลักษณะของอนุกรมเวลานั้นได้
2. ในงานวิจัยนี้ได้ใช้ตัวประมาณค่าเบี่ยงเบนสัมบูรณ์ของมัธยฐาน (MAD) เป็นตัวประมาณการกระจาย ตามที่ Gelper , Fried และ Croux (2010) ได้เสนอไว้ ในงานวิจัยต่อไปอาจใช้ตัวประมาณการกระจายที่มีความแกร่งตัวอื่น ๆ เช่น Interquatile range : IQR (Huber: 1981)

รายการอ้างอิง

- จรงค์ ศรีทิพย์. (2548). “ การตรวจหาค่านอกกลุ่มในอนุกรมเวลา โดยกระบวนการทำซ้ำและการตรวจสอบแบบตัดออก k ค่า.” วิทยานิพนธ์ปริญญาโทมหาบัณฑิต สาขาวิชาสถิติประยุกต์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร.
- ไพโรจน์ ขาวสิทธิ์ชัย. “The R Statistical Programming Language.” เอกสารประกอบการสอนรายวิชา 515571 ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร.
- สายชล สันสมบูรณ์ทอง. (2554). การควบคุมคุณภาพเชิงสถิติและวิศวกรรม. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร: จามจุรีโปรดักท์.
- Alwan, Layth C., Robert, Herry V. (1988). “Time-Series Modeling for Statistical Process Control.” **Journal of Business & Econometric Statistics** 6(January): 87-95.
- Bisgaard, Soren. (2005). “Quality Quandaries : The Effect of Autocorrelation on Statistical Process Control Procedures.” **Quality Engineering** 17: 481-489.
- Chen C, Liu L. (1993). “Joint estimation of model parameters and outlier effects in time series.” **Journal of the American Statistical Association** 33: 284–97.
- Croux, Christophe., Gelper, Sarah., and Mahiou, Koen. (2011). “Robust control charts for time series data.” **Expert Systems with Applications** 38: 13810-13815.
- Fox, A.J. (1972). “Outlier in Time Series.” **Journal of the Royal Statistical Society Series B** 34(May): 350-363.
- Gelper, Sarah., Fried, Roland., and Croux, Christophe. (2010). “Robust Forecasting with Exponential and Holt-Winters Smoothing.” **Journal of Forecasting** 29(June): 285-300.
- Huber, Peter J. (1981). **Robust Statistics**. York: John Wiley.
- Montgomery, Douglas. (2005). **Introduction to Statistical Quality Control**. 5th ed. New York: John Wiley and Sons.
- Moustafa Omar Ahmed Abu-Shawiesh. (2008). “A simple Robust Control Chart Based on MAD.” **Journal of Mathematics and Statistics** 4: 102-107.
- Tsay, Ruey S. (1988). “Outliers, Level Shift, and Variance Changes in time series.” **Journal of Forecasting** 7: 1-20.

Wright, C.M., and HU, M.Y. (2003). "A note detecting outliers in short autocorrelated data using joint estimation and exponentially weighted moving average." **Journal of Management Science** 31(March): 319-326.

ภาคผนวก
โปรแกรมที่ใช้ในงานวิจัย

โปรแกรมคอมพิวเตอร์ที่ใช้ในการจำลองและประมวลผลข้อมูล

โปรแกรมที่ใช้ R เวอร์ชัน 3.0.1

ตัวแปรที่ใช้ในโปรแกรมประกอบด้วย

T	แทนขนาดตัวอย่างที่ใช้ในการศึกษา
n	แทนขนาดตัวอย่างใน training period
correlate	แทนค่าพารามิเตอร์ ϕ
e	แทนความคลาดเคลื่อนสุ่ม
z	แทนอนุกรมเวลาคงที่
wtrain	แทนขนาดของค่านอกกลุ่มที่เกิดขึ้นใน training period
wtest	แทนขนาดของค่านอกกลุ่มที่เกิดขึ้นใน test period
locationtrain	แทนตำแหน่งของอนุกรมเวลาที่ถูกรวมเพื่อสร้างค่านอกกลุ่มใน training period
locationtest	แทนตำแหน่งของอนุกรมเวลาที่ถูกรวมเพื่อสร้างค่านอกกลุ่มใน test period

ตัวแปรสำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมาตรฐาน

zhat	แทนค่าพยากรณ์ล่วงหน้า 1 คาบเวลาใน training period
residual	แทนค่าความคลาดเคลื่อนในการพยากรณ์ล่วงหน้า 1 คาบเวลาใน training period
Zhat	แทนค่าพยากรณ์ล่วงหน้า 1 คาบเวลาใน test period
res	แทนค่าความคลาดเคลื่อนในการพยากรณ์ล่วงหน้า 1 คาบเวลาใน test period
SSE	แทนผลบวกของส่วนเหลือกำลังสองมีค่าต่ำที่สุด
Lam	แทนค่าพารามิเตอร์ที่ปรับให้เรียบ (λ) ที่เหมาะสมที่สุด
UCL	แทนขีดจำกัดควบคุมบนของแผนภูมิควบคุม
LCL	แทนขีดจำกัดควบคุมล่างของแผนภูมิควบคุม
all	แทนจำนวนค่านอกกลุ่มที่ตรวจจับได้ทั้งหมดทั้งที่เป็นค่านอกกลุ่มที่สร้างขึ้นและค่าที่ไม่ได้เป็นค่านอกกลุ่มแต่ได้รับการตรวจจับ
outlier	แทนจำนวนของค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับได้
false	แทนจำนวนการตรวจจับผิด (ค่าที่ไม่ได้เป็นค่านอกกลุ่มแต่ได้รับการตรวจจับ)

result แทนค่าเฉลี่ยของ λ , จำนวนค่านอกกลุ่มที่ตรวจจับได้ทั้งหมด , จำนวนของค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับได้ และจำนวนการตรวจจับผิดตามลำดับ จากการทำซ้ำ 1,000 ครั้ง

ตัวแปรสำหรับแผนภูมิควบคุมที่สร้างจากวิธีการพยากรณ์การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีความแกร่ง

zhatrb แทนค่าพยากรณ์ล่วงหน้า 1 คาบเวลาของอนุกรมเวลา

residualrb แทนค่าความคลาดเคลื่อนในการพยากรณ์ล่วงหน้า 1 คาบเวลาของอนุกรมเวลา

scale แทนตัวประมาณการกระจาย σ_t

stderror แทนค่ามาตรฐานของความคลาดเคลื่อนในการพยากรณ์ล่วงหน้า 1 คาบเวลา

clean แทนอนุกรมเวลาที่ไม่มีค่านอกกลุ่ม z_t^*

SSErb แทนผลบวกของส่วนเหลือกำลังสองมีค่าต่ำที่สุด

Lamrb แทนค่าพารามิเตอร์ที่ปรับให้เรียบ (λ) ที่เหมาะสมที่สุด

UCLrb แทนขีดจำกัดควบคุมบนของแผนภูมิควบคุม

LCLrb แทนขีดจำกัดควบคุมล่างของแผนภูมิควบคุม

allrb แทนจำนวนค่านอกกลุ่มที่ตรวจจับได้ทั้งหมดทั้งที่เป็นค่านอกกลุ่มที่สร้างขึ้นและค่าที่ไม่ได้เป็นค่านอกกลุ่มแต่ได้รับการตรวจจับ

outlierrb แทนจำนวนของค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับได้

falserb แทนจำนวนการตรวจจับผิด (ค่าที่ไม่ได้เป็นค่านอกกลุ่มแต่ได้รับการตรวจจับ)

resultrb แทนค่าเฉลี่ยของ λ , จำนวนค่านอกกลุ่มที่ตรวจจับได้ทั้งหมด , จำนวนของค่านอกกลุ่มที่สร้างขึ้นใน test period ถูกตรวจจับได้ และจำนวนการตรวจจับผิดตามลำดับ จากการทำซ้ำ 1,000 ครั้ง

โปรแกรมหลัก

```

sumLamda<-0
sumAll <-0
sumFalse <-0
sumoutlier <-0
sumLamdarb<-0
sumAllrb <-0
sumFalserb <-0
sumoutlierrb <-0
for(j in 1:1000)
{
T<-100
correlate<-0.1
wtrain<-3
wtest<-0
controlchart<- Control.Chart(T,correlate, wtrain, wtest)
sumLamda<-sumLamda+ controlchart$Lamda
sumAll<-sumAll+ controlchart$all
sumFalse<-sumFalse+ controlchart$False
sumoutlier <-sumoutlier + controlchart$outlier
sumLamdarb<-sumLamdarb+ controlchart$Lamdarb
sumAllrb<-sumAllrb+ controlchart$allrb
sumFalserb<-sumFalserb+ controlchart$Falserb
sumoutlierrb <-sumoutlierrb + controlchart$outlierrb
}
MeanLamda<- (sumLamda/1000)
MeanAll<- (sumAll/(0.5*T*1000))
MeanFalse<-(sumFalse/((0.9*(0.5*T))*1000))
Meanoutlier <-(sumoutlier /((0.1*(0.5*T))*1000))

```

```

MeanLamdarb<- (sumLamdarb/1000)
MeanAllrb<- (sumAllrb/(0.5*T*1000))
MeanFalserb<- (sumFalserb/((0.9*(0.5*T))*1000))
Meanoutlierrb <- (sumoutlierrb /((0.1*(0.5*T))*1000))
result<-c(MeanLamda,MeanAll,MeanFalse,Meanoutlier)
result
resultrb<-c(MeanLamdarb,MeanAllrb,MeanFalserb,Meanoutlierrb)
resultrb

```

โปรแกรมสำหรับการสร้างและวัดประสิทธิภาพของแผนภูมิควบคุมทั้งสองแบบ

```

Control.Chart <-function(T,correlate, wtrain, wtest)
{
n=0.5*T
z=c()
e=c()
e<-rnorm(T,0,1)
z[1]<-e[1]
  for(j in 2:T)
  {
    z[j]<-(correlate*z[j-1])+e[j]
  }
locationtrain<-sample(n,0.1*n)
z[locationtrain]<-z[locationtrain]+(sign(z[locationtrain])*wtrain)
locationtest<-sample(n+1:n,0.1*(T-n))
z[locationtest]<-z[locationtest]+(sign(z[locationtest])*wtest)
zhat=c()
residual=c()
Lamda<-0
SSE<-100000000

```

```

for(j in seq(0,1,0.05))
{
lam=j
  {
      zstart<- mean(z[1:10])
      zhat [1]<- (lam* z[10])-((1-lam)*zstart)
      residual[1]<-z[11]-zhat[1]
      for(i in 2:(n-10))
          {
              zhat [i]<-(lam*z [i+9])-((1-lam)* zhat [i-1])
              zhat [i]
              residual[i]<- z [i+10]- zhat [i]
          }
      sse<-sum(residual^2)
  }
  ifelse(sse<SSE,
    {SSE<-sse
      Lamda<-lam}
    ,{SSE<-SSE
      Lamda<-Lamda}
  )
SSE
Lamda
}
UCL=1.96*sqrt(SSE/(n-10))
LCL=-1.96*sqrt(SSE/(n-10))
Zhat=c()
res=c()
Lam<-Lamda

```

```

Zhat [1]<- (Lam* z[10])-((1-Lam)*zstart)
res [1]<-z[11]-Zhat[1]
  for(i in 2:(T-10))
    {
      Zhat [i]<-(Lam*z [i+9])-((1-Lam)* Zhat [i-1])
      res[i]<- z [i+10]- Zhat[i]
    }
outlier<-0
all<-0
for(i in c(locationtest-10))
  {
    if(res[i]>=UCL|| res[i]<=LCL)
      outlier <- outlier +1
  }
for(i in (n-9):(T-10))
  {
    if(res [i]>=UCL|| res [i]<=LCL) all<- all+1
  }
false <- (all-outlier)
false
zhatrb=c()
residualrb=c()
scale=c()
clean=c()
stderror=c()
medianres=c()
cleanresidualrb=c()
SSErb<-1000000000
Lamdarb<-0

```

```

for(j in seq(0,1,0.05))
{
lamrb<-j
{
zhatstart<- median(z[1:10])
scalestart<-mad(z[1:10])
error<-z[10]- zhatstart
stderrorstart<- error/ scalestart
if(abs(stderrorstart)>2) stderrorstart <- sign(stderrorstart)*2
cleanstart<-(stderrorstart*scalestart)+ zhatstart
zhatrb[1]<- (lamrb* cleanstart) -((1-lamrb)* zhatstart)
for(i in 1:(n-11))
{
residualrb[i]<- z [i+10]- zhatrb[i]
medianres[i]<-median(c(residualrb[1:i],error))
scale[i]<-(1.4826)*median(abs(residualrb[1:i]-
medianres[i]))
stderror[i]<- residualrb[i]/ scale[i]
if(abs(stderror[i])>2) stderror[i]<-sign(stderror[i])*2
clean[i]<- (stderror[i]* scale[i])+ zhatrb[i]
zhatrb[i+1]<- ((lamrb* clean[i]) -((1-lamrb)* zhatrb[i]))
}
residualrb[n-10]<- z [n]- zhatrb[n-10]
S<-1.4826*median(abs(residualrb[1:(n-10)]))
cleanresidualrb <-(residualrb/S) ^2
for(l in 1:( n-10))
{
if(cleanresidualrb[l] >4) cleanresidualrb[l] <- 4
}
sserb<-(S^2) *sum(cleanresidualrb)
}
}

```

```

    }
    ifelse(sserb<SSErb,
    {SSErb<-sserb
    Lamdarb<-lamrb}
    ,{SSErb<-SSErb
    Lamdarb<-Lamdarb}
    )
SSErb
Lamdarb
}
SSErb
Lamdarb
UCLrb=1.96*sqrt(SSErb/(n-10))
LCLrb=-1.96*sqrt(SSErb/(n-10))

Lamrb<- Lamdarb
zhatrb[1]<- (Lamrb * cleanstart) -((1- Lamrb)* zhatstart)
  for(i in 1:(T-11))
    {
      residualrb[i]<- z [i+10]- zhatrb[i]
      medianres[i]<-median(c(residualrb[1:i],error))
      scale[i]<-(1.4826)*median(abs(residualrb[1:i]- medianres[i]))
      stderror[i]<- residualrb[i]/ scale[i]
      if(abs(stderror[i])>2) stderror[i]<-sign(stderror[i])*2
      clean[i]<- stderror[i]* scale[i]+ zhatrb[i]
      zhatrb[i+1]<- (Lamrb* clean[i]) -((1-Lamrb)* zhatrb[i])
    }
  residualrb[T-10]<- z [T]- zhatrb[T-10]

outlierrb<-0

```



```

    allrb<-0
for(i in c(locationtest-10))
{
    if(residualrb[i]> UCLrb|| residualrb[i]< LCLrb) outlierrb <- outlierrb +1
}
for(i in (n-9):(T-10))
{
    if(residualrb[i]>= UCLrb|| residualrb[i]<=LCLrb) allrb<- allrb+1
}
falserb <- (allrb-outlierrb)
list(Lamda=Lam,all=all,False=false, outlier = outlier ,Lamdarb=Lamrb,allrb=allrb,Falserb=falserb,
outlierrb = outlierrb)
}

```

หมายเหตุ โปรแกรมนี้กำหนดให้อนุกรมเวลาที่จำลองมีขนาดตัวอย่าง 100 ตัวอย่าง โดยที่ $\phi=0.1$ ซึ่งมีค่านอกกลุ่มขนาดเล็กเกิดขึ้นใน training period และไม่มีค่านอกกลุ่มเกิดขึ้นใน test period ($w_1=3, w_2=0$) สำหรับอนุกรมเวลาอื่น ๆ สามารถกำหนดค่าของตัวแปรต่าง ๆ ตามขอบเขตที่กำหนดไว้ โดยทำการเปลี่ยนแปลงในส่วน of โปรแกรมหลัก

ประวัติผู้วิจัย

ชื่อ – สกุล
ที่อยู่

นางสาวอรรณณ เหลืองสีเพชร
บ้านเลขที่ 254 หมู่ 6 ตำบลประสาธสิทธิ์ อำเภอดำเนินสะดวก
จังหวัดราชบุรี

ประวัติการศึกษา

พ.ศ. 2554

สำเร็จการศึกษาระดับปริญญาวิทยาศาสตรบัณฑิต สาขาสถิติ
มหาวิทยาลัยศิลปากร และเข้าศึกษาต่อระดับปริญญาวิทยาศาสตร
มหาบัณฑิต สาขาสถิติประยุกต์ มหาวิทยาลัยศิลปากร

ประวัติการทำงาน

พ.ศ. 2554 – ปัจจุบัน

ผู้ช่วยสอนรายวิชาสถิติพื้นฐาน ภาควิชาสถิติ มหาวิทยาลัยศิลปากร