



รายงานผลการวิจัยฉบับสมบูรณ์

การเลือกฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มเพื่อสร้างตัวแบบ
สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาหลายชุดพร้อมกันที่มีค่า อัตตสหสัมพันธ์
แนวโน้ม และค่าผิดปกติ ด้วยวิธีการของเบย์: กรณีศึกษาราคาหุ้นที่มีความผันผวนของ
ตลาดหลักทรัพย์แห่งประเทศไทย

**Selection of Cumulative Distribution Function for Trend in Model Formation
for the Analysis of Simultaneous Multiple Time Series Data with
Autoregression, Trend and Outliers Using Bayesian Method: Case Study in
Fluctuate Stock price in The Stock Exchange of Thailand**

นายพิษณุ ทองขาว

ผู้ช่วยศาสตราจารย์นิศยา บุญสิทธิ

นายกฤษฎา เหล็กดี

นางสาวอรศิริ จันทรเมือง

งานวิจัยนี้ได้รับทุนสนับสนุนจากงบประมาณ ปี พ.ศ. 2557

มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

ชื่อเรื่อง : การเลือกฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มเพื่อสร้าง
 ตัวแบบ สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาหลายๆตัวพร้อมกันที่มีค่า อัตราสหสัมพันธ์
 แนวโน้ม และค่าผิดปกติ ด้วยวิธีการของเบย์:กรณีศึกษาราคาหุ้นที่มีความผันผวนของ
 ตลาดหลักทรัพย์แห่งประเทศไทย

ผู้วิจัย : พิษณุ ทองขาว นิตยา บุญลธิธิ์ กฤษฎา เหล็กดี และอรศิริ จันทร์เมือง
พ.ศ. : 2557



บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อเลือกฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสม สำหรับ
 แนวโน้มที่เหมาะสมในแบบการพยากรณ์ด้วยวิธีการของเบย์ เพื่อวิเคราะห์ข้อมูลอนุกรมเวลา
 หลายๆตัวพร้อมกันโดยมีส่วนประกอบของการวิเคราะห์ค่า อัตราสหสัมพันธ์ แนวโน้ม และค่า
 ผิดปกติ และประยุกต์ใช้กับข้อมูลอนุกรมเวลาซึ่งเป็นราคาปิดของหุ้นจากตลาดหลักทรัพย์แห่ง
 ประเทศไทยจำนวน 5 ตัว โดยเขียนอัลกอริทึมในโปรแกรม OpenBUGS และประเมิน
 ประสิทธิภาพของตัวแบบต่างๆด้วยการเขียนโปรแกรมจำลองสถานการณ์ด้วยโปรแกรม R
 หลังจากนั้นนำตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มต่างๆมา
 เปรียบเทียบกัน ได้แก่ฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเวย์บูล กับ
 ฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียล โดยดูจากค่า
 RMSE MAPE และ MAE ผลการวิจัยพบว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็น
 สดสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลให้ค่า RMSE MAPE และ MAE ต่ำสุดของหุ้น 4 ตัว
 ได้แก่ หุ้น EA NWR KTB และ ADVANCE ทั้งการหาตัวแบบที่เหมาะสม (Model Fitting) และ
 การตรวจสอบความถูกต้องของตัวแบบ (Model Validation) ส่วนตัวแบบเบย์ที่มีค่าฟังก์ชันการแจก
 แจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเวย์บูลให้ค่า RMSE MAPE และ MAE ต่ำสุด 1 หุ้น
 คือ หุ้น PTT ทั้งการหาตัวแบบที่เหมาะสม และการตรวจสอบความถูกต้องของตัวแบบ

คำสำคัญ: ฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้ม อนุกรมเวลา
 อัตราสหสัมพันธ์ ค่าผิดปกติ วิธีการของเบย์

Title : Selection of Cumulative Distribution Function for Trend in Model Formation for the Analysis of Simultaneous Multiple Time Series Data with Autocorrelation, Trend and Outliers Using Bayesian Method: Case Study in Fluctuate Stock Price in the Stock Exchange of Thailand

Researcher: Pitsanu Tongkhaw, Nittaya Boonsith, Krisada Lekdee, and Onsiri Junmuang

Year : 2014

Abstract

The objective of this research is to select a proper cumulative function for trend in an analysis of simultaneous multiple time series data with autocorrelation, trend, and outliers using a Bayesian method. The proposed model is applied to fluctuate stock price data in the stock exchange of Thailand. The data consist of the closing prices of 5 stocks. The algorithms are written in OpenBUGS and the performance of the model is evaluated via a simulation study in R. The cumulative distribution functions considered here are a Weibull cumulative distribution function and an exponential cumulative distribution. They are compared using RMSE, MAPE, and MAE as criteria. The research found that an exponential cumulative distribution function for trends gives smaller values of the three criteria in both model fitting and model validation for the 4 stocks, EA, NWR, and KTB. However, a Weibull cumulative distribution function gives smaller values of three criteria for the PTT in both model fitting and model validation.

Keyword: Cumulative distribution function for trend, Time series, Autocorrelation, Outliers, Bayesian method

กิตติกรรมประกาศ

การวิจัยเรื่องนี้สำเร็จลงด้วยดี ด้วยความช่วยเหลือจากหลายท่าน คณะผู้วิจัยขอขอบพระคุณ รศ. สุภัทรา โกไศยกานนท์ รักษาการแทนอธิการบดีมหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร ผศ. ดร. วัลลภ ภูผา คณบดีคณะวิศวกรรมศาสตร์ ผศ.จุฑามาศ พิรพัชระ ผู้อำนวยการสถาบันวิจัยและพัฒนา ที่ให้การสนับสนุนการทำงานวิจัยของอาจารย์ในมหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร มาตั้งแต่เริ่มต้น ทำนี้คณะผู้วิจัยขอขอบคุณมหาวิทยาลัยเทคโนโลยีราชมงคลพระนครที่ได้ให้ทุนสนับสนุน จนกระทั่งงานวิจัยฉบับนี้สำเร็จลงด้วยดี

คณะผู้วิจัย



สารบัญ

	หน้า
กิตติกรรมประกาศ	ก
บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ค
สารบัญ	ง
บทที่ 1 บทนำ	1
ความเป็นมาและความสำคัญของปัญหา	1
วัตถุประสงค์ของการวิจัย	3
ขอบเขตของการวิจัย	3
สมมุติฐานในการวิจัย	3
ประโยชน์ที่คาดว่าจะได้รับ	4
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	5
ทฤษฎีที่เกี่ยวข้องกับเทคนิคการพยากรณ์	5
ทฤษฎีที่เกี่ยวข้องกับหุ่น	22
งานวิจัยที่เกี่ยวข้อง	27
บทที่ 3 วิธีดำเนินการวิจัย	30
ประชากรและกลุ่มตัวอย่าง	30
เครื่องมือที่ใช้ในการศึกษา	30
วิธีการดำเนินศึกษา	31
สร้างตัวแบบด้วยวิธีของเบย์	31

การประเมินผลวิจัย	32
บทที่ 4 ผลการวิจัย	33
ผลการวิเคราะห์ลักษณะของข้อมูล	33
ผลการสร้างตัวแบบ	41
ผลของการประเมินประสิทธิภาพของตัวแบบ	47
ผลของการประมาณค่าพารามิเตอร์	49
ผลของการเปรียบเทียบตัวแบบพยากรณ์	54
บทที่ 5 สรุปผลการวิจัย และข้อเสนอแนะ	59
สรุปผลการวิจัย	59
ข้อเสนอแนะในการวิจัยครั้งต่อไป	62
บรรณานุกรม	63
ภาคผนวก	64
ประวัติคณะผู้วิจัย	66



บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ข้อมูลอนุกรมเวลามีส่วนประกอบที่สำคัญได้แก่ แนวโน้ม (trend) อัตตสหสัมพันธ์ (autocorrelation) ฤดูกาล (seasonal) วัฏจักร (cycle) และเหตุการณ์ผิดปกติ (irregular variation or outlier) สำหรับการเลือกฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มของข้อมูลอนุกรมเวลาที่เหมาะสมกับข้อมูลที่มีความผันผวน หรือเป็นตัวแปรสุ่มนั้นจะช่วยทำให้การวิเคราะห์ข้อมูลอนุกรมเวลาให้มีความแม่นยำขึ้นอีกด้านหนึ่งจากส่วนประกอบทั้งหมด ซึ่งจะต้องกำหนดฟังก์ชันต่างๆของการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มโดยการปรับให้พอดีกับข้อมูลให้มากที่สุด (fitting) หรือมีความผิดพลาดน้อยที่สุดเมื่อเปรียบเทียบกับข้อมูลจริง ฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มที่นิยมใช้กันอยู่ทั่วไปได้แก่ ฟังก์ชันการแจกแจงความน่าจะเป็นแบบเวย์บูล (Weibull cumulative distribution function) แบบเอ็กซ์โพเนนเชียล (exponential cumulative distribution function) แบบปกติ (normal cumulative distribution function) เป็นต้น แต่ปัญหาที่ตามมาคือจะทำให้ตัวแบบสำหรับการวิเคราะห์ข้อมูลอนุกรมเวลามีความซับซ้อนมากขึ้นและการหาคำตอบด้วยวิธีทางคณิตศาสตร์ทั่วไปไม่สามารถทำการประมาณพารามิเตอร์ได้ ซึ่งจะต้องประมาณพารามิเตอร์โดยการจำลองสถานการณ์ และถ้าใส่ฟังก์ชันครบทุกส่วนของข้อมูลอนุกรมเวลาจากที่ได้กล่าวมาข้างต้นวิธีการจำลองสถานการณ์ก็ยังมีขีดจำกัดในการประมาณพารามิเตอร์เพื่อให้ใกล้เคียงกับข้อมูลจริงมากที่สุด วิธีการหนึ่งซึ่งสามารถประมาณค่าพารามิเตอร์โดยใช้การจำลองสถานการณ์กับตัวแบบที่มีฟังก์ชันหลายๆฟังก์ชันรวมอยู่ด้วยกันซึ่งเรียกว่าวิธีการของเบย์ สำหรับวิธีการของเบย์ ในประเทศไทยยังถือว่ามีการใช้กันอยู่น้อยมาก และที่มีการใช้ส่วนใหญ่ ก็ยังเป็นตัวแบบที่วิเคราะห์กันทั่วไป ส่วนใหญ่ก็ยังใช้กันอยู่ในตัวแบบที่มีข้อจำกัดที่เป็นตัวแบบง่ายๆไม่สามารถวิเคราะห์ข้อมูลที่มีส่วนประกอบของฟังก์ชันต่างๆหลายฟังก์ชันรวมกันได้ ซึ่งก็จะทำให้ได้ผลการวิเคราะห์ไม่แตกต่างจากการวิเคราะห์ด้วยตัวแบบอื่นๆที่นิยมใช้กันอยู่ในปัจจุบัน ข้อดีของวิธีการของเบย์คือ สามารถนำไปประยุกต์ใช้ในการประมาณค่าพารามิเตอร์ในตัวแบบที่มีความซับซ้อนได้ดี และการที่พารามิเตอร์มิได้หลายระดับมองเป็นฟังก์ชันต่างๆย่อมสอดคล้องกับสภาพความเป็นจริงในชีวิตประจำวันที่ว่าทุกสิ่งทุกอย่างย่อมขึ้นอยู่กับกันและกันและไม่เป็นอิสระกัน และสถานการณ์ในปัจจุบันส่วนใหญ่ก็มักจะไม่แน่นอน ซึ่งความไม่แน่นอนของทุกสิ่งทุกอย่างต้องมีการขึ้นอยู่กับกันและกันอย่างเป็นขั้นๆดังที่กล่าวมา เช่น ราคาขึ้นอยู่กับความต้องการ ดินฟ้าอากาศ ภัยธรรมชาติ นโยบายทางการบริหารประเทศ เป็นต้น และ

ความต้องการก็ยังขึ้นอยู่กับปัจจัยทางด้านอื่นๆอีก ทำให้เกิดการขึ้นอยู่กับสิ่งต่างๆหลายๆชั้นจึงทำให้เกิดความซับซ้อนมากขึ้นไปอีกสำหรับการวิเคราะห์ ดังนั้นตัวแบบทางคณิตศาสตร์ที่ใช้ในการวิเคราะห์จึงมีความซับซ้อนตามไปด้วย และต้องบรรจุฟังก์ชันของความเป็นต่างๆที่เหมาะสมกับข้อมูลในการประมาณค่าพารามิเตอร์หรือตัวแปรไว้ในตัวแบบ ซึ่งตัวแบบประเภทนี้เรียกว่าตัวแบบเชิงสโตแคสติก วิธีของเบย์กับตัวแบบที่วิเคราะห์ข้อมูลเชิงลึกหลายชั้นในปัจจุบันจึงมีผู้นำไปประยุกต์ใช้ในงานวิจัยด้านต่างๆกันอย่างแพร่หลายส่วนใหญ่จะอยู่ในต่างประเทศ

สำหรับข้อมูลอนุกรมเวลาที่มีความผันผวนส่วนใหญ่จะแบ่งไปด้วยอิทธิพลจาก แนวโน้ม (Trend) ฤดูกาล (Seasonal) วัฏจักร (Cycle) และเหตุการณ์ผิดปกติ (Irregular Variation) ต่างๆ เป็นต้น และปัจจุบันได้มีการใช้ตัวแบบหลากหลายวิธีที่จะสอดคล้องตามติดข้อมูลเหล่านี้เพื่อให้ได้ความใกล้เคียงและแม่นยำมากที่สุด เช่น วิธีปรับเรียบแบบ Exponential (Exponential Smoothing, EXPS), วิธี Autoregressive Integrated Moving Average (ARIMA) และวิธี Seasonal Autoregressive Integrated Moving Average (SARIMA) วิธีโครงข่ายประสาทเทียม (Artificial Neural Network, ANN) วิธีจีเนติก อัลกอริทึม (Genetic Algorithms, GA) วิธีตรรกศาสตร์คลุมเครือ (Fuzzy logic) เป็นต้น แต่ตัวแบบที่จะสอดคล้องตามติดข้อมูลอนุกรมเวลาเหล่านั้น หรือนิยมเรียกว่าวิธีการคาดการณ์ หรือพยากรณ์ ที่ได้กล่าวมาข้างต้นนั้น ยังมีข้อจำกัดในการประมาณค่าพารามิเตอร์และยังไม่ครอบคลุมการวิเคราะห์ข้อมูลที่มีความผันผวนสูง และข้อมูลอนุกรมเวลาหลายตัวพร้อมๆกัน ที่มีอัตราสหสัมพันธ์ แนวโน้ม และค่าผิดปกติ ดังนั้นจึงทำให้เกิดความคลาดเคลื่อนในการสอดคล้องตาม และทำให้การวางแผนตัดสินใจในกระบวนการต่างๆของระบบผิดพลาดตามไปด้วย ซึ่งเราจะพบข้อมูลประเภทนี้บ่อยมากในชีวิตประจำวัน สำหรับตัวแบบลักษณะนี้ในประเทศไทยยังไม่มีใครเคยทำวิจัยมาก่อนด้วยเพราะเป็นตัวแบบที่มีความซับซ้อนมากและต้องใช้เทคนิคการประมาณค่าพารามิเตอร์โดยการเขียนโปรแกรมทางคณิตศาสตร์ขั้นสูงตามไปด้วย

ด้วยเหตุนี้ผู้วิจัยจึงสนใจที่จะศึกษา และสร้างตัวแบบวิเคราะห์ข้อมูลเชิงลึกหลายชั้น กับฟังก์ชันการแจกแจงความเป็นอิสระสำหรับแนวโน้ม สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาหลายๆตัวพร้อมกันที่มีค่า อัตราสหสัมพันธ์ แนวโน้ม และค่าผิดปกติ รวมอยู่ด้วย เพื่อให้สอดคล้องกับสภาพความเป็นจริงในชีวิตประจำวัน ด้วยวิธีการของเบย์ และนำไปประยุกต์ใช้กับข้อมูลราคาปิดหุ้นที่มีความผันผวนของตลาดหลักทรัพย์แห่งประเทศไทย ซึ่งจะใช้หลักการประมาณค่าพารามิเตอร์ที่อยู่ในรูปของกระบวนการสโตแคสติกสำหรับข้อมูลอนุกรมเวลาหลายๆตัวพร้อมๆกัน และเมื่อเราสร้างตัวแบบได้แล้วจะต้องมีการประเมินค่าความน่าเชื่อถือของตัวแบบโดยใช้วิธีจำลองสถานการณ์และเขียนโปรแกรมในโปรแกรม R ซึ่งเหตุจูงใจที่เลือกใช้กับราคาหุ้นก็เพราะว่าราคาหุ้นหลายๆตัวมีความผันผวนสูง และมีเหตุการณ์ต่างๆเกิดขึ้น เช่น สถานะทางเศรษฐกิจ

นโยบายทางการเมืองภายในประเทศ การปั่นหุ้น หรือปัจจัยต่างๆที่เกิดจากในและนอกประเทศ ได้ตลอด (ข้อมูลจากศูนย์ส่งเสริมการพัฒนาความรู้ตลาดทุน) และตัวแบบที่ได้จากงานวิจัยนี้ก็สามารถนำไปใช้ในการประมาณการเพื่อพยากรณ์ข้อมูลอนุกรมเวลาอื่นๆที่มีความผันผวนสูงได้เช่นกัน

วัตถุประสงค์ของการวิจัย

- 1) เพื่อสร้างตัวแบบการพยากรณ์ที่สามารถพยากรณ์ข้อมูลราคาหลายๆตัวพร้อมกันให้มีความแม่นยำมากขึ้น โดยเพิ่มการเลือกฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มที่เหมาะสมเข้าไปในตัวแบบ ภายใต้ข้อมูลที่มี อัตราสหสัมพันธ์ แนวโน้ม และค่าผิดปกติ และนำตัวแบบการพยากรณ์ที่หาได้ ไปประยุกต์ใช้กับราคาหุ้นที่มีความผันผวนสูงหลายๆตัวพร้อมกัน
- 2) เพื่อใช้หลักการสร้างตัวแบบ และโปรแกรมที่เขียนได้ไปประยุกต์ใช้กับข้อมูลอนุกรมเวลาด้านอื่นๆที่มีความผันผวนสูง และเกิดเหตุการณ์ ผิดปกติต่างๆได้

สมมุติฐานของการวิจัย

กำหนดฟังก์ชันการแจกแจงความน่าจะเป็นก่อนหน้าต่างๆของราคาปิดของหุ้นต่างๆ (prior distribution) และเมื่อนำตัวแบบที่สร้างได้ใหม่ไปเปรียบเทียบกับตัวแบบเบย์ของงานวิจัยก่อนแล้วได้ค่าผิดพลาดที่ต่ำกว่า

ขอบเขตของโครงการวิจัย

เลือกข้อมูลราคาปิดของหุ้นรายวัน จำนวน 6 เดือน ที่มีความผันผวนซึ่งวิเคราะห์จากบริษัทหลักทรัพย์เอเซียพลัสโดยพิจารณาจากหุ้นที่มีค่า Beta (ค่าความผันผวนของราคาหุ้นเทียบกับตลาดหุ้น) สูงกว่า 1 และค่าพี/อี ที่สูงกว่า 20 เท่า จำนวน 2 ตัว ได้แก่ บมจ.พลังงานบริสุทธิ์ (EA) บมจ.เนาวรัตน์พัฒนาการ (NWR) และหุ้นที่เป็นที่รู้จักกันทั่วไปอีก 3 ตัว ได้แก่ บมจ. ปตท (PTT) บมจ.ธนาคารกรุงไทย (KTB) บมจ.แอดวานซ์ อินโฟร์ เซอร์วิส (ADVANCE) ข้อมูลจากตลาดหลักทรัพย์แห่งประเทศไทย

ประโยชน์ที่คาดว่าจะได้รับ

- 1) ได้ตัวแบบการพยากรณ์ข้อมูลอนุกรมเวลาหลายๆตัวพร้อมกันที่สามารถนำไปใช้ในการพยากรณ์ข้อมูลอนุกรมเวลาที่มีความผันผวนสูงเช่นราคาหุ้น และข้อมูลอื่นๆได้ความแม่นยำขึ้น

- 2) สามารถใช้หลักการประมาณพารามิเตอร์เชิงสโตแคสติกสำหรับตัวแบบวิเคราะห์เชิงลึกที่มีความซับซ้อนกับข้อมูลอนุกรมเวลาหลายๆตัวพร้อมกันกับข้อมูลอนุกรมเวลาชนิดอื่นๆได้
- 3) ค่าพยากรณ์ที่ได้สามารถนำไปใช้สำหรับวางแผนและตัดสินใจต่างๆได้ เช่น การวางแผนการตัดสินใจลงทุนขยายงานต่อที่ได้จากข้อมูลการพยากรณ์ราคาหุ้นของบริษัท หรือ การตัดสินใจวางแผนในโรงงานอุตสาหกรรมจากข้อมูลราคา หรือปริมาณความต้องการสินค้าที่มีความผันผวนสูง
- 4) นำตัวแบบการพยากรณ์ที่ได้ไปประยุกต์ใช้กับการพยากรณ์ด้านอื่นๆได้
- 5) ใช้เป็นพื้นฐานในการพัฒนาตัวแบบต่อไป



บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การเลือกฟังก์ชันการแจกแจงความน่าจะเป็นเหมาะสมสำหรับแนวโน้มเพื่อสร้างตัวแบบ สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาหลายๆตัวพร้อมกันที่มีค่า อัตราสหสัมพันธ์ แนวโน้ม และค่าผิดปกติ ด้วยวิธีการของเบย์:กรณีศึกษาราคาหุ้นที่มีความผันผวนของตลาดหลักทรัพย์แห่งประเทศไทย ในครั้งนี้ผู้ศึกษาได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องตามหัวข้อดังต่อไปนี้

ทฤษฎีที่เกี่ยวข้อง
งานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้องกับเทคนิคการพยากรณ์

ก่อนที่จะทำการตัดสินใจเลือกวิธีการพยากรณ์ใดๆควรจะพิจารณาถึงลักษณะรูปแบบของข้อมูลที่กำลังตัดสินใจว่ามีความสอดคล้องกับลักษณะของวิธีการพยากรณ์ใดที่ต้องการจะเลือกใช้ สำหรับการพยากรณ์โดยทั่วไป มีหลักเกณฑ์ในการพิจารณาความรู้พื้นฐานที่เกี่ยวข้องดังต่อไปนี้

2.1.1 เทคนิคการพยากรณ์ (Box et al.,1994)

เป็นวิธีการที่ใช้คาดการณ์ข้อมูลในอนาคต โดยคาดว่าจะมีลักษณะเช่นเดียวกับข้อมูลในปัจจุบันหรืออดีต เช่น ข้อมูลยอดขายหรืออุปสงค์ในความเป็นจริง ซึ่งได้รับอิทธิพลจากแนวโน้ม (Trend) ฤดูกาล (Seasonal) วัฏจักร (Cycle) และเหตุการณ์ผิดปกติ (Irregular Variation) ต่างๆ เป็นต้น การพยากรณ์ (Forecasting) มีการนำไปใช้อย่างแพร่หลาย ทั้งทางด้านเศรษฐกิจและสังคม ยกตัวอย่างเช่น การพยากรณ์ความต้องการ ราคา ผลผลิต หรือแม้กระทั่งการพยากรณ์อากาศ อุณหภูมิ ปริมาณน้ำฝน การพยากรณ์เป็นสิ่งที่จำเป็นอย่างมากในการวางแผนในด้านต่างๆ ทั้งนี้ ความแม่นยำ รวดเร็ว เป็นปัจจัยสำคัญในการวัดคุณภาพของการพยากรณ์ เทคนิคหรือวิธีการพยากรณ์ก็ได้มีการพัฒนามาอย่างต่อเนื่องเพื่อตอบสนองต่อความต้องการดังกล่าว ซึ่งเทคนิคการพยากรณ์สามารถแบ่งได้เป็นสองประเภทหลักๆคือ เทคนิคการพยากรณ์เชิงคุณภาพ (Qualitative Forecasting Techniques) และ เทคนิคการพยากรณ์เชิงปริมาณ (Quantitative Forecasting Techniques)

ในส่วนของเทคนิคการพยากรณ์เชิงปริมาณ สามารถแบ่งออกเป็นสองกลุ่มหลักๆ ได้แก่ การพยากรณ์แบบอนุกรมเวลา (Time Series Forecasting) เช่น วิธีนาอิว (Naive) วิธีปรับเรียบ (Exponential Smoothing) วิธี ARIMA วิธีของเบย์ (Bayesian) และ การพยากรณ์แบบวิเคราะห์ความสัมพันธ์ (Causal Forecasting) เช่น วิธีวิเคราะห์ความถดถอย (Regression Analysis) วิธี Econometric วิธี Input-Output นอกจากนี้ก็ยังมีเทคนิคการพยากรณ์สมัยใหม่

อีกหลายวิธีที่ได้รับความนิยม เช่น วิธีโครงข่ายประสาทเทียม (Artificial Neural Network, ANN) วิธีจีเนติก อัลกอริทึม (Genetic Algorithms, GA) วิธีตรรกศาสตร์คลุมเครือ สำหรับในงานวิจัยนี้จะขออธิบายเทคนิคการพยากรณ์ที่นิยมใช้ในปัจจุบัน และเทคนิคที่ใช้เป็นพื้นฐานความรู้ในการศึกษาและทำวิจัย

1) วิธีการพยากรณ์แบบถดถอยเชิงพหุ (Multiple Regression) (Yan,2009)

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \varepsilon_t \quad (1)$$

เมื่อ t แทนเวลา โดยที่ $t = 1, \dots, n$

Y_t แทน ตัวแปรตาม ณ เวลา t และ X_{it} แทนตัวแปรอิสระ ณ เวลา t โดยที่ $i = 1, \dots, k$

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$ แทน ค่าคงที่ และ ε_t แทน ค่าความผิดพลาด ณ เวลา t และเป็นอิสระกัน โดย ε_t จะมีการแจกแจงแบบปกติ มีค่าเฉลี่ย เท่ากับ 0 และค่าความแปรปรวน คือ σ^2 สำหรับ

$$\text{การพยากรณ์ ณ เวลา } t \text{ จะได้ } \hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 X_{1t} + \hat{\beta}_2 X_{2t} + \dots + \hat{\beta}_k X_{kt} + \varepsilon_t \quad (2)$$

เมื่อ $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ ได้จากการประมาณค่าโดยวิธี maximum likelihood

2) วิธีการพยากรณ์แบบค่าเฉลี่ยเคลื่อนที่ (Moving Average (MA)) (Bisgaard,2011)

$$F_{t+1} = \frac{y_t + y_{t-1} + y_{t-2} + \dots + y_{t+1-n}}{N} \quad (3)$$

เมื่อ N = ขนาดของ moving average โดยที่ $N = 1, \dots, n$

ตัวอย่างเช่น

$$F_4 = \frac{y_3 + y_2 + y_1}{3}$$

$$F_5 = \frac{y_4 + y_3 + y_2 + y_1}{4}$$

เมื่อ

$$N = 1 \rightarrow F_2 = y_1, F_3 = y_2, \dots$$

$$N = n \rightarrow F_{n+1} = \bar{y}$$

ถ้า N เป็นฤดูกาล ดังนั้น จะสามารถกำจัดฤดูกาลไปได้ในตัว

สำหรับวิธี Moving Average ที่ได้กล่าวมาแล้ว ยังมีวิธี double MA, triple MA และ centered MA อีก เช่น 2x2 MA, 3x5 MA , 3x3x5 MA เป็นต้น

กรณีมี trend ในข้อมูล

ให้ใช้วิธี MA และ double MA โดยกำหนดให้มี moving average length ที่เท่ากันดังนี้

$$F_{t+m} = a_t + b_t m_t \quad \text{โดยที่ } m = 1, 2, 3, \dots \quad (4)$$

เมื่อ

$$\begin{aligned} a_t &= 2S'_t - S''_t \\ b_t &= \frac{2}{N-1}(S'_t - S''_t) \\ S'_t &= \frac{1}{N}(y_t + y_{t-1} + y_{t-2} + \dots + y_{t+1-n}) \\ S''_t &= \frac{1}{N}(S'_t + S'_{t-1} + S'_{t-2} + \dots + S'_{t+1-n}) \end{aligned}$$

3) วิธีการพยากรณ์แบบ Exponential Smoothing (EXPS) (Montgomery, 2008)

Exponential smoothing เป็นการพยากรณ์โดยกำหนดน้ำหนักให้ค่าสังเกตในปัจจุบันมีค่ามากกว่าน้ำหนักของค่าสังเกตก่อนหน้านั้น มี 3 แบบคือ

1. Single exponential smoothing

ตัวแบบการพยากรณ์แบบ Single exponential smoothing นิยามดังนี้

ให้ S_t แทนค่าพยากรณ์ของค่าสังเกต Y_t ณ เวลา t เมื่อ $t = 1, \dots, n$

$$S_t = \alpha Y_t + (1 - \alpha) S_{t-1}, \quad 0 \leq \alpha \leq 1 \quad (5)$$

ค่าพยากรณ์ S_1 ไม่เกิดขึ้นเนื่องจากไม่มีเทอม S_0 นอกเสียจากจะกำหนดค่าเริ่มต้นให้ S_0 เทคนิคนี้ใช้สำหรับข้อมูลที่ ไม่มี trend และ seasonal ถ้าข้อมูลมี trend จะใช้ Double exponential smoothing และถ้ามีทั้ง trend และ seasonal จะใช้ Triple Exponential Smoothing

2. Double exponential smoothing หรือ Holt's method

ตัวแบบการพยากรณ์แบบ Double exponential smoothing ได้จาก ขยายตัว

แบบของ Single exponential smoothing ออกไปดังนี้

$$S_t = \alpha Y_t + (1 - \alpha)(S_{t-1} + A_{t-1}), \quad 0 \leq \alpha \leq 1 \quad (6)$$

$$\text{โดยที่ } A_t = \beta(S_t - S_{t-1}) + (1 - \beta)A_{t-1}, \quad 0 \leq \beta \leq 1 \quad (7)$$

เมื่อ β คือค่าคงที่แสดง trend

3. Triple Exponential Smoothing หรือ Holt-Winters method

ตัวแบบการพยากรณ์แบบ Triple exponential smoothing ได้จากขยายตัวแบบของ Double exponential smoothing ออกไป มี 2 แบบคือ

3.1 Multiplicative Seasonal Model

$$S_t = \alpha(Y_t / B_{t-s}) + (1 - \alpha)(S_{t-1} + A_{t-1}), \quad 0 \leq \alpha \leq 1, \quad (8)$$

$$A_t = \beta(S_t - S_{t-1}) + (1 - \beta)A_{t-1}, \quad 0 \leq \beta \leq 1, \quad (9)$$

$$B_t = \gamma(Y_t / S_t) + (1 - \gamma)B_{t-s}, \quad 0 \leq \gamma \leq 1, \quad (10)$$

เมื่อ γ คือค่าคงที่แสดง seasonal และ s คือความยาวช่วงของ seasonal ใช้ตัวแบบการพยากรณ์นี้เมื่อข้อมูลมี seasonal เป็นแบบการคูณ (multiplicative seasonality) คือในเวลาเดียวกันของแต่ละฤดูกาล ค่าสังเกตจะเพิ่มขึ้นหรือลดลงเป็นค่าร้อยละ ตัวอย่างเช่นข้อมูลรายเดือน 5 ปี ความยาวช่วงของ seasonal คือ 12 และข้อมูลปรากฏให้เห็นว่าในเดือนธันวาคมของแต่ละปี ค่าสังเกตจะเพิ่มขึ้นร้อยละ 40 ไม่ใช่เพิ่มขึ้นหรือลดลงเป็นค่าคงที่

3.2 Additive Seasonal Model

$$S_t = \alpha(Y_t - B_{t-s}) + (1 - \alpha)(S_{t-1} + A_{t-1}), \quad 0 \leq \alpha \leq 1, \quad (11)$$

$$A_t = \beta(S_t - S_{t-1}) + (1 - \beta)A_{t-1}, \quad 0 \leq \beta \leq 1, \quad (12)$$

$$B_t = \gamma(Y_t - S_t) + (1 - \gamma)B_{t-s}, \quad 0 \leq \gamma \leq 1, \quad (13)$$

เมื่อ γ คือค่าคงที่แสดง seasonal และ s คือความยาวช่วงของ seasonal ใช้ตัวแบบการพยากรณ์นี้เมื่อข้อมูลมี seasonal เป็นแบบการบวก (additive seasonality) คือในเวลาเดียวกันของแต่ละฤดูกาล ค่าสังเกตจะเพิ่มขึ้นหรือลดลงเป็นค่าคงที่ ในทางปฏิบัติเราไม่ทราบว่าจะต้องกำหนดค่า moving average length เท่าไรดี จึงจะทำให้ ค่า error หรือ mean square error (MSE) หรือ ค่าวัดความผิดพลาดอื่นๆ ให้ค่าต่ำที่สุด ดังนั้นเราควรจะต้องกำหนดค่า α ให้เหมาะสม

4) วิธีการพยากรณ์แบบ Autoregressive Integrated Moving Average Model (ARIMA) (Montgomery, 2008)

แบบจำลอง ARIMA เป็นแบบจำลองที่ได้รับความนิยม และเป็นวิธีที่ให้ค่าพยากรณ์ในระยะสั้นที่ดี อีกทั้งในการจัดทำสมการและการพยากรณ์ยังมีขั้นตอนที่ยุ่งยาก และซับซ้อนน้อยกว่าแบบมหภาคที่อยู่ในลักษณะระบบสมการหลายชั้น สำหรับแบบจำลอง ARIMA เป็นแบบจำลองที่พัฒนาโดย George E.P.Box และ Gwilym M. Jenkins ในปี ค.ศ. 19702 โดยพื้นฐานแล้วแบบจำลอง ARIMA เป็นวิธีที่ให้ค่าพยากรณ์ในระยะสั้นที่ดี หรือเหมาะกับการพยากรณ์ไปข้างหน้าในช่วงเวลาสั้นๆ และต้องมีช่วงของข้อมูลที่ยาวพอสมควร แบบจำลอง ARIMA(p,d,q) ประกอบด้วย 3 ส่วนหลักๆ ได้แก่ แบบจำลอง Auto Regressive (AR(p)) กระบวนการ Integrated (I(d)) และแบบจำลอง Moving Average (MA(q)) โดยรายละเอียดของแต่ละส่วนมีดังนี้

1. แบบจำลอง Auto Regressive (AR(p))

แบบจำลอง Auto Regressive เป็นรูปแบบที่แสดงว่าค่าสังเกต y_t ถูกกำหนดจากค่าของ $y_t, \dots, y_{t-p}$ หรือ ค่าสังเกตที่เกิดขึ้นก่อนหน้า p โดยกระบวนการหรือระบบ AR(p) คือ กระบวนการหรือระบบ Auto Regressive ที่มีอันดับที่ p ซึ่งเขียนอยู่ในรูปสมการได้ดังนี้

$$AR(p) \text{ คือ } x_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (14)$$

โดยที่

μ คือ ค่าคงที่ (Constant Term)

ϕ_j คือ พารามิเตอร์ตัวที่ j

ε_t คือ ความคลาดเคลื่อน ณ เวลา t

ในกรณี ของ AR(1) สามารถเขียนรูปแบบสมการได้ดังนี้

$$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t \quad (15)$$

หรือ

$$y_t - \phi_1 y_{t-1} = \mu + \varepsilon_t \quad (16)$$

หรือ

$$(1 - \phi_1 \beta) x_t = \mu + \varepsilon_t \quad (17)$$

เมื่อ β คือ backward shift operation

และในกรณี ของ AR(2) สามารถเขียนรูปแบบสมการได้ดังนี้

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t \quad (18)$$

หรือ

$$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} = \mu + \varepsilon_t \quad (19)$$

หรือ

$$(1 - \phi_1 \beta - \phi_2 \beta^2) y_t = \mu + \varepsilon_t \quad (20)$$

2. แบบจำลอง Moving Average (MA(q))

แบบจำลอง Moving Average (MA) เป็นรูปแบบที่แสดงว่าค่าสังเกต y_t ถูกกำหนดจากค่าความคลาดเคลื่อน $\varepsilon_{t-1}, \dots, \varepsilon_{t-p}$ หรือค่าความคลาดเคลื่อนที่อยู่ก่อนหน้า โดยกระบวนการหรือระบบ MA(q) คือกระบวนการหรือระบบ Moving Average ที่มีอันดับ q ซึ่งเขียนในรูปของ MA (q) ได้ดังนี้

$$\text{MA (q) คือ } y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (21)$$

โดยที่

μ คือ ค่าคงที่ (Constant Term)

θ_j คือ พารามิเตอร์ตัวที่ j

ε_t คือ ความคลาดเคลื่อน ณ เวลา t

ในกรณี MA(1) สามารถเขียนรูปแบบสมการได้ดังนี้

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (22)$$

หรือ

$$y_t = \mu + (1 - \theta_1 \beta) \varepsilon_t \quad (23)$$

และในกรณี MA(2) สามารถเขียนรูปแบบสมการได้ดังนี้

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \quad (24)$$

หรือ

$$y_t = \mu + (1 - \theta_1 \beta - \theta_2 \beta^2) \varepsilon_t \quad (25)$$

3. แบบจำลอง Auto Regressive Moving Average (ARMA(p,q))

เป็นแบบจำลองที่นำเอากระบวนการ Auto Regressive และ Moving Average มาใช้รวมกัน โดยกระบวนการหรือระบบ ARMA(p,q) คือกระบวนการหรือระบบ Auto Regressive ที่มีอันดับที่ p และ Moving Average ที่มีอันดับ q ซึ่งเขียนอยู่ในรูปสมการได้ดังนี้

$$\Delta_t y_t = \delta + \phi y_{t-1} + \phi y_{t-2} + \dots + \phi y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (26)$$

โดยที่

y_t คือ ค่าสังเกตในอนุกรมเวลา ณ เวลา t

p คือ อันดับของ Autoregressive

q คือ อันดับของ Moving Average

δ คือ ค่าคงที่ (Constant Term)

t คือ เวลา

ϕ คือ พารามิเตอร์ของ Auto Regressive

θ คือ พารามิเตอร์ของ Moving Average

ε_t คือ กระบวนการ white noise ซึ่งก็คือ ค่าความคลาดเคลื่อน ณ เวลา t

4. กระบวนการ Integrated (I (d))

กระบวนการ Integrated (I(d)) เป็นการหาผลต่างของอนุกรมเวลาระหว่างข้อมูล ณ ปัจจุบันกับข้อมูลย้อนหลังไป d คาบเวลา โดยสาเหตุที่ต้องทำการหาผลต่างของอนุกรมเวลาเนื่องจากแบบจำลอง ARIMA ต้องใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีคุณสมบัติคงที่ (Stationary) เท่านั้น โดยในกรณีข้อมูลอนุกรมเวลาที่ใช้ในการวิเคราะห์มีคุณสมบัติไม่คงที่ (Nonstationary) จะต้องทำการแปลงข้อมูลดังกล่าวให้เป็นข้อมูลที่มีคุณสมบัติคงที่ก่อน โดยการหาผลต่างของข้อมูลอนุกรมเวลาก่อนที่นำไปสร้างแบบจำลอง ARIMA ซึ่งโดยทั่วไปแล้วถ้าต้องหาผลต่างอันดับที่ d สามารถเขียนในรูปของ I(d) ได้ดังนี้

$$I(d) \text{ คือ } \Delta_d x_t = \Delta_{d-1}(x_t - x_{t-1}) \text{ หรือ } (1-\beta)^d x_t \quad (27)$$

ในกรณี I (1) สามารถเขียนรูปแบบได้ดังนี้

$$I(1) \text{ คือ } \Delta x_t = (x_t - x_{t-1}) \text{ หรือ } (1-\beta)x_t$$

ในกรณี I (2) สามารถเขียนรูปแบบได้ดังนี้

$$I(2) \text{ คือ } \Delta_2 x_t = \Delta(x_t - x_{t-1}) \text{ หรือ } (1-\beta)^2 x_t$$

โดยที่

ε_t คือ พจน์ความคลาดเคลื่อน ณ เวลา t

$(1-\beta)^d x_t$ คือ ผลต่างอันดับที่ d

β คือ Backward shift operation

จากรายละเอียดต่างๆ ที่กล่าวข้างต้นถ้านำแบบจำลอง Auto Regressive แบบจำลอง Moving Average และ กระบวนการ Integrated มาพิจารณารวมกันสามารถนำมากำหนดเป็นรูปแบบทั่วไปของแบบจำลอง ARIMA ที่ใช้ในการประมาณการคือ

$$\Delta_t y_t = \delta + \phi \Delta_d y_{t-1} + \phi \Delta_d y_{t-2} + \dots + \phi \Delta_d y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (28)$$

โดยที่

y_t คือ ค่าสังเกตในอนุกรมเวลา ณ เวลา t

d คือ จำนวนครั้งของการหาผลต่างเพื่อให้อนุกรมเวลามีคุณสมบัติคงที่ (Stationary)

p คือ อันดับของ Autoregressive

q คือ อันดับของ Moving Average

δ คือ ค่าคงที่ (Constant Term)

t คือ เวลา

Δ_d คือ ผลต่างอันดับที่ d

$\phi_1, \dots, \phi_p$ คือ พารามิเตอร์ของ Auto Regressive

$\theta_1, \dots, \theta_q$ คือ พารามิเตอร์ของ Moving Average

ε_t คือ กระบวนการ white noise ซึ่งก็คือ ค่าความคลาดเคลื่อน ณ เวลา t

ภายใต้ข้อสมมติที่ว่าความคลาดเคลื่อนที่คนละเวลาเป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน

กันโดยมีการแจกแจงปกติที่มีค่าเฉลี่ยเป็นศูนย์ และความแปรปรวนคงที่

ดังนั้นเพื่อให้ง่ายขึ้นสำหรับการหาค่าพยากรณ์ Y_t ของ ARIMA สามารถหาได้

จากนิยามที่สรุปไว้ต่อไปนี้

$$(1 - \phi_1 \beta - \phi_2 \beta^2 - \dots - \phi_p \beta^p)(1 - \beta)^d Y_t = (1 - \omega_1 \beta - \omega_2 \beta^2 - \dots - \omega_q \beta^q) \varepsilon_t \quad (29)$$

$$\text{เมื่อ } \beta Y_t = Y_{t-1}, \beta^2 Y_t = Y_{t-2}, \beta^3 Y_t = Y_{t-3}, \dots \quad (30)$$

$$\beta \varepsilon_t = \varepsilon_{t-1}, \beta^2 \varepsilon_t = \varepsilon_{t-2}, \beta^3 \varepsilon_t = \varepsilon_{t-3}, \dots \quad (31)$$

ตัวอย่างเช่น

ARIMA(0,1,0) คือ

$$\begin{aligned} (1 - \beta)Y_t &= \varepsilon_t \\ Y_t - \beta Y_t &= \varepsilon_t \\ Y_t - Y_{t-1} &= \varepsilon_t \\ Y_t &= Y_{t-1} + \varepsilon_t \end{aligned} \quad (32)$$

ARIMA(2,1,2) คือ

$$\begin{aligned}
 (1 - \phi_1\beta - \phi_2\beta^2)(1 - \beta)Y_t &= (1 - \omega_1\beta - \omega_2\beta^2)\varepsilon_t \\
 (1 - \phi_1\beta - \phi_2\beta^2)(Y_t - \beta Y_t) &= \varepsilon_t - \omega_1\beta\varepsilon_t - \omega_2\beta^2\varepsilon_t \\
 (Y_t - \phi_1\beta Y_t - \phi_2\beta^2 Y_t) - (\beta Y_t - \phi_1\beta^2 Y_t - \phi_2\beta^3 Y_t) &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \\
 Y_t - \phi_1\beta Y_t - \phi_2\beta^2 Y_t - \beta Y_t + \phi_1\beta^2 Y_t + \phi_2\beta^3 Y_t &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \\
 Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - Y_{t-1} + \phi_1 Y_{t-2} + \phi_2 Y_{t-3} &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \\
 Y_t - \phi_1 Y_{t-1} - Y_{t-1} - \phi_2 Y_{t-2} + \phi_1 Y_{t-2} + \phi_2 Y_{t-3} &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \\
 Y_t - (\phi_1 - 1)Y_{t-1} - (\phi_2 - \phi_1)Y_{t-2} + \phi_2 Y_{t-3} &= \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2} \\
 Y_t &= (\phi_1 - 1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} - \phi_2 Y_{t-3} + \varepsilon_t - \omega_1\varepsilon_{t-1} - \omega_2\varepsilon_{t-2}
 \end{aligned} \tag{33}$$

ตัวแบบ ARIMA ใช้วิเคราะห์ข้อมูลได้ทั้งแบบ stationary และ nonstationary ข้อมูล stationary คือข้อมูลที่ mean และ variance ไม่เปลี่ยนแปลงไปตามเวลา ข้อมูล nonstationary คือข้อมูลที่ mean หรือ variance เปลี่ยนแปลงไปตามเวลา ค่า variance ทำให้ stationary ได้โดยการแปลงข้อมูลด้วยฟังก์ชันลอการิทึม (Log transform) ซึ่งต้องทำก่อนใช้ตัวแบบ A สำหรับค่า mean ทำให้ stationary ได้โดยการทำผลต่างข้อมูล (differencing) ซึ่งสามารถทำในตัวแบบ ARIMA ได้ โดยการกำหนดค่าให้กับ d ส่วนระดับของ autocorrelation (p) และระดับของ moving average (q) พิจารณาได้จากกราฟ (Autocorrelation function) ACF และ Partial autocorrelation function (PACF) (Montgomery, 2008)

5. Autocorrelation Function (ACF)

เป็นฟังก์ชันของการวัดสหสัมพันธ์ระหว่างข้อมูล ณ เวลา t (x_t) และ ข้อมูล ณ เวลา $t-k$ (x_{t-k}) ของช่วงเวลาห่างกัน k หน่วย ซึ่งแทนด้วยสัญลักษณ์ ρ_k หรือ r_k ในกรณีสหสัมพันธ์ในตัวเองของตัวอย่าง ซึ่งสามารถคำนวณได้ดังนี้

$$\rho_k \text{ หรือ } r_k = \frac{\sum_{t=k+1}^n (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2} \tag{34}$$

เมื่อ $\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t$ และ $k = 0, 1, 2, 3, \dots$

โดยความคลาดเคลื่อนมาตรฐานของ r_k (Standard Error of r_k) ซึ่งมีสูตรในการคำนวณดังนี้

$$se_{r_k} = \frac{1}{\sqrt{n}} \tag{35}$$

สหสัมพันธ์ในตัวเองของข้อมูลสุ่ม (random data) มีการแจกแจงเชิงตัวอย่างที่สามารถประมาณได้ โดยการแจกแจงปกติด้วยค่าเฉลี่ย (mean) เท่ากับศูนย์ และความคลาดเคลื่อนมาตรฐานเท่ากับ $\frac{1}{\sqrt{n}}$

ในการศึกษาจะใช้สหสัมพันธ์ในตัวเองเป็นเครื่องมือที่สำคัญสำหรับสืบค้นคุณสมบัติของข้อมูลอนุกรมเวลาเชิงประจักษ์ โดยมี 2 วิธีสำหรับทดสอบว่าค่า r_k มีค่าแตกต่างไปจากศูนย์หรือไม่โดยใช้การแจกแจงปกติมาตรฐาน (Standard Normal Distribution) หรือ ใช้ค่าสถิติ Box-Pierce Q statistic ซึ่งมีรูปแบบดังนี้

การแจกแจงปกติมาตรฐาน (Standard Normal Distribution)

$$r_k \sim N(0, \frac{1}{\sqrt{n}})$$

ค่าสถิติ Box-Pierce Q statistic

$$Q = n \sum_{k=1}^m r^2 \sim \chi^2(m-p-q)$$

โดยที่ m คือค่าล่าหรือค่าล่าหลังสูงสุด (Maximum Lag) ที่พิจารณา

6. Partial Autocorrelation Function (PACF)

เป็นการพิจารณาสหสัมพันธ์ระหว่างตัวแปร x_t กับ x_{t-k} อาจเป็นไปได้ว่าสหสัมพันธ์ดังกล่าวเป็นผลเนื่องมาจากสหสัมพันธ์ระหว่าง 2 ตัวแปรนี้กับตัวแปร $x_{t-1}, \dots, x_{t-k+1}$ ดังนั้นเพื่อที่จะได้สหสัมพันธ์ระหว่าง x_t กับ x_{t-k} ที่ได้ขจัดความเกี่ยวข้องระหว่างตัวแปรทั้งสองตัวนี้กับตัวแปร $x_{t-1}, \dots, x_{t-k+1}$ ดังกล่าว จึงต้องทำการวัดสหสัมพันธ์ของทั้งสองตัวแปรในรูปแบบของการสหสัมพันธ์แบบมีเงื่อนไข $Corr(x_t, x_{t-k} | x_{t-1}, \dots, x_{t-k+1})$ ซึ่งเรียกว่า Partial Autocorrelation โดยแทนด้วยสัญลักษณ์ ϕ_{kk} แต่ถ้านำสหสัมพันธ์ในตัวเองบางส่วนมาพิจารณาในรูปแบบฟังก์ชัน จะเรียกว่า Partial Autocorrelation Function (PACF) ซึ่ง ϕ_{kk} สามารถคำนวณได้ดังนี้

$$\phi_{kk} = \frac{Cov[(x_t - \hat{x}_t), (x_{t-k} - \hat{x}_{t-k})]}{\sqrt{Var(x_t - \hat{x}_t)} \sqrt{Var(x_{t-k} - \hat{x}_{t-k})}} \quad (36)$$

$$\text{โดยที่} \quad \hat{x}_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_k x_{t-k+1} \quad (37)$$

5) วิธีการพยากรณ์แบบ Seasonal Autoregressive Integrated Moving Average Model (SARIMA) (Montgomery, 2008), (Bisgaard, 2011)

วิธีการพยากรณ์แบบ SARIMA ถูกพัฒนามาจากวิธีการพยากรณ์ ARIMA (p,d,q) โดยได้เพิ่ม (P,D,Q) ของ Seasonal เข้าไปอีก SARIMA หรือ Seasonal ARIMA แทนด้วย

ARIMA (p, d, q)(P, D, Q) และถ้ามีตัวแปรร่วม (covariate) X_{1t} และ X_{2t} สามารถนิยามได้ดังต่อไปนี้

$$\phi(\beta)[\Delta(Y_t - c_1 X_{1t} - c_2 X_{2t}) - \mu] = \theta(\beta)\varepsilon_t \quad (38)$$

เมื่อ c_1 และ c_2 คือสัมประสิทธิ์การถดถอย

$\phi(\beta) = \phi_p(\beta)\phi_p(\beta^s)$ และ $\theta(\beta) = \theta_q(\beta)\theta_q(\beta^s)$ โดยที่

$\phi_p(\beta) = 1 - \phi_1\beta - \dots - \phi_p\beta^p$, $\theta_q(\beta) = 1 - \theta_1\beta - \dots - \theta_q\beta^q$,

$\phi_p(\beta^s) = 1 - \phi_1\beta^s - \dots - \phi_p\beta^{sp}$, และ $\theta_q(\beta^s) = 1 - \theta_1\beta^s - \dots - \theta_q\beta^{sq}$. Δ คือ

differencing operator, $\Delta = (1 - \beta)^d(1 - \beta^s)^D$. β คือ backshift operator

และ $\beta Y_t = Y_{t-1}$, $\beta^2 Y_t = Y_{t-2}$ เป็นแบบนี้ต่อไป s คือ seasonal lag และ ε_t คือ ความผิดพลาด

(error)ที่มีการแจกแจงแบบปกติมีค่าเฉลี่ยเป็น 0 และความแปรปรวน σ^2 . ϕ 's และ θ 's คือ

พารามิเตอร์ autoregressive แบบ seasonal และ non-seasonal ตามลำดับ θ 's และ ϕ 's คือ

พารามิเตอร์ moving average แบบ seasonal และ non-seasonal ตามลำดับ p และ q คือ

ลำดับ (order) ของพารามิเตอร์ autoregressive แบบ non-seasonal และ พารามิเตอร์

moving average แบบ non-seasonal ตามลำดับโดยที่ P และ Q คือ ลำดับ (order) ของ

พารามิเตอร์ autoregressive แบบ seasonal และ พารามิเตอร์ moving average แบบ

seasonal ตามลำดับ d และ D แทน non-seasonal และ seasonal differencesตามลำดับ

7) ตัวแบบการพยากรณ์แบบโครงข่ายประสาทเทียม (Artificial Neural Network=ANN)

(Anderson, 1997), (Hagan, 1996)

โครงข่ายประสาทเทียม หรือเรียกสั้นๆว่าข่ายงานประสาท (Neural Network) คือ การใช้ตัวแบบทางคณิตศาสตร์ในการประมวลผลสารสนเทศแบบเดียวกับสมองของมนุษย์ ANN เป็น Nonlinear model เป้าหมายของ ANN เหมือนกับตัวแบบ Regression คือต้องการประมาณค่าพารามิเตอร์เพื่อให้เกิดค่าความผิดพลาด (error) น้อยที่สุด สำหรับการหาคำตอบของตัวแบบ ANN นั้นจะใช้การวนซ้ำ

ANN โครงข่ายแบบป้อนไปข้างหน้า (Feed Forward Network) คือสารสนเทศ หรือข้อมูลจะถูกส่งผ่านจากชั้นที่ซ่อนอยู่ (Hidden Layer) หนึ่ง ไปยังชั้นที่ซ่อนอยู่ต่อไป (ซึ่งในหนึ่งโครงข่ายจะมีชั้นที่ซ่อนอยู่กี่ชั้นก็ได้แต่โดยปกติจะมีชั้นที่ซ่อนอยู่ประมาณ 2-5 ชั้น ตัวแบบนี้ถึงจะทำงานให้ผลที่ดี) สำหรับการส่งข้อมูลจะถูกส่งผ่านไปในทิศทางเดียวกัน ไม่มีการส่งย้อนกลับ โดยกระบวนการ (Process) จะเริ่มจาก ตัวแปรพยากรณ์ (Predictive Variables) ในชั้นของ input ตัวแปรนี้คือข้อมูลที่เรารวบรวมได้ เช่น ราคาสินค้า ปริมาณผลผลิต เป็นต้น โดยข้อมูลนี้

จะถูกส่งต่อไปยังชั้นที่ซ่อนอยู่ (Hidden Layer) ซึ่งชั้นนี้จะบรรจุเครื่องมือที่ใช้ทำงานของ ANN โดยจะเรียกว่า activation function คือฟังก์ชันที่แสดงความสัมพันธ์ของ input ที่เข้ามา กับ Output ที่จะออกไป เช่นฟังก์ชัน logistic หรือฟังก์ชัน Nonlinear อื่นๆ

สำหรับวิธีการประมาณค่าพารามิเตอร์นั้นมีได้หลายวิธี แต่จะขอยกตัวอย่างวิธีที่ได้รับ ความนิยมน้อยอย่างแพร่หลายคือวิธีการแพร่กระจายแบบย้อนกลับ (Back-propagation) ซึ่งเป็น วิธีการใช้อัลกอริทึมการแพร่กระจายแบบย้อนกลับ โดยจะทำการปรับปรุงน้ำหนักคะแนนของ เครือข่ายเพื่อให้ได้ผลรวมมากกว่าค่าที่ตั้งเป้าไว้ (Threshold) หลังจากนั้นจึงจะสามารถส่งข้อมูล ไปยังชั้นขาออก (Output layer) ได้ แล้วจึงทำการคำนวณหาค่าความผิดพลาด ซึ่งค่าความ ผิดพลาดนี้ก็จะถูกส่งกลับเข้าสู่เครือข่ายเพื่อใช้แก้ไขค่าน้ำหนักคะแนนต่อไป จนได้ค่าความ ผิดพลาดต่ำที่สุด

8) วิธีการพยากรณ์แบบ เบย์ (Bayesian) (Robert, 2001), (Congdon,2006),(West,1997)

ตัวแบบเบย์สร้างจาก Likelihood, $p(\mathbf{Y}|\boldsymbol{\theta})$, และ Prior, $\pi(\boldsymbol{\theta})$, เมื่อ $\mathbf{Y}$ คือตัวแปรสุ่ม ที่สังเกตค่าได้ และ $\boldsymbol{\theta}$ คือค่าพารามิเตอร์ที่สังเกตค่าไม่ได้

การแจกแจงร่วม (Joint Distribution) ของ $\boldsymbol{\theta}$ กับ $\mathbf{Y}$ สามารถเขียนอยู่ในรูป

$$\pi(\boldsymbol{\theta}, \mathbf{Y}) = p(\mathbf{Y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \text{ และ Posterior ที่สร้างจากกฎของเบย์คือ}$$

$$\pi(\boldsymbol{\theta}|\mathbf{Y}) = \frac{p(\boldsymbol{\theta}, \mathbf{Y})}{p(\mathbf{Y})} = \frac{p(\mathbf{Y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{p(\mathbf{Y})}$$

(39)

โดยที่

$$p(\mathbf{Y}) = \sum_{\boldsymbol{\theta}} p(\mathbf{Y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \text{ เมื่อ } \boldsymbol{\theta} \text{ เป็นตัวแปรสุ่มชนิดไม่ต่อเนื่อง (Discrete) และ}$$

$$p(\mathbf{Y}) = \int p(\mathbf{Y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} \text{ เมื่อ } \boldsymbol{\theta} \text{ เป็นตัวแปรสุ่มชนิดต่อเนื่อง (Continuous)}$$

เนื่องจาก $p(\mathbf{Y})$ เป็นฟังก์ชันของ $\mathbf{Y}$ ซึ่งไม่ขึ้นอยู่กับการเลือก $\boldsymbol{\theta}$ จึงถูกพิจารณาว่าเป็นค่าคงที่และ สามารถเขียน $\pi(\boldsymbol{\theta}|\mathbf{Y})$ อยู่ในรูป $\pi(\boldsymbol{\theta}|\mathbf{Y}) \propto p(\mathbf{Y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$ นั่นคือ $\pi(\boldsymbol{\theta}|\mathbf{Y})$ เป็นสัดส่วนกับผล คูณของ จาก Likelihood กับ Prior

ตัวแบบที่ซับซ้อนสามารถใช้ตัวแบบเบย์แก้ปัญหาได้ เช่นใช้ตัวแบบเบย์ที่มี 3 ชั้น ได้แก่ขั้นตอนที่ 1 ระบุการแจกแจงของตัวแปรสุ่มที่สังเกตค่าได้เมื่อกำหนดพารามิเตอร์ให้ขั้นตอนที่ 2 ระบุ การแจกแจงของพารามิเตอร์เมื่อกำหนดไฮเปอร์พารามิเตอร์ให้ และขั้นตอนที่ 3 ระบุการแจก แจงของไฮเปอร์พารามิเตอร์ในทำนองเดียวกัน จำนวนขั้นตอนอาจมีมากกว่า 3 ได้

ตัวแบบเบย์สามารถเพิ่มความแกร่ง (Robustness) ให้กับตัวประมาณแบบเบย์ได้ เนื่องจากความไม่แน่นอน (Uncertainty) ถูกนำมาคิดไว้ในขั้นตอนของการแจกแจงของ Prior นอกจากนั้นวิธีการของเบย์ยังทำให้การประมาณค่าพารามิเตอร์ใน Posterior ง่ายขึ้น โดยใช้การ

การจำลองสถานการณ์ (Simulation) การจำลองสถานการณ์ที่ใช้กันอย่างแพร่หลายคือ วิธีเชิ่งตัวเลข Markov Chain Monte Carlo (MCMC)

สำหรับตัวอย่างเพื่อให้เห็นภาพรวมของวิธีเบย์ จะขอยกตัวอย่างตัวแบบที่มีความซับซ้อนจึงใช้วิธีการของเบย์ในการแก้ปัญหา ตัวแบบคือ

$$Y_{it} = \beta_0 + \beta_1 Z_{it,1} + \beta_2 Z_{it,2} + \dots + \beta_p Z_{it,p} + \gamma_i \Delta W(t | \alpha_i, \delta_i) + \gamma_i X_{it} + \varepsilon_{it}$$

เมื่อ Y_{it} แทนราคา หรือปริมาณผลผลิตของพืชชนิดที่ i ในช่วงเวลา t , $i=1, \dots, m$ และ $t=1, \dots, T_i$

โดยที่

$$\varepsilon_{it} \sim N(0, [\gamma_i(1+3\zeta_{it})\sigma_\varepsilon]^2)$$

$$Y_{it} \sim N(\beta_0 + \beta_1 Z_{it,1} + \beta_2 Z_{it,2} + \dots + \beta_p Z_{it,p} + \gamma_i \Delta W(t | \alpha_i, \delta_i) + \gamma_i X_{it}, [\gamma_i(1+3\zeta_{it})\sigma_\varepsilon]^2)$$

Prior (มีหลายระดับ) คือ

$$1. \quad p(\sigma_\varepsilon) \propto \text{Unif}(0, \infty), \quad p(\beta_i) \propto \text{constant}$$

2. ข้อมูลผิดปกติ (Outliners)

$$\zeta_{it} \sim \text{Bern}(0.05)$$

3. ราคาหรือปริมาณผลผลิตรวมทุกช่วงเวลา

$$\gamma_i \sim N(g_i, \sigma_\gamma^2), \quad p(\sigma_\gamma) \propto \text{Unif}(0, \infty)$$

$$g_j \sim N(\mu_g, \sigma_g^2), \quad p(\mu_g) \propto 1, \quad p(\sigma_g) \propto \text{Unif}(0, \infty)$$

$$S_i \sim N(\gamma_i, [0.2\gamma_i]^2)$$

4. Autoregression ที่ซ่อนเร้นอยู่ (Latent Autoregression)

$$X_{it} \sim N(\lambda_{i1}X_{it-1} + \lambda_{i2}X_{it-2}, \sigma_x^2), \quad \sigma_x = 0.8\sigma_\varepsilon$$

$$(\lambda_{i1}, \lambda_{i2})^T \sim N(\mu_\lambda, \Sigma_\lambda), \quad p(\mu_\lambda, \Sigma_\lambda) \propto |\Sigma_\lambda|^{-2}$$

$$(X_{i0}, X_{i,-1})^T \sim N(\mu_{x_0}, \Sigma_{x_0}), \quad \mu_{x_0} = (0, 0)^T, \quad \Sigma_{x_0} = \text{diag}(2, 2).$$

5. พารามิเตอร์อื่นๆ

$$\alpha_i \sim N(a_i, \sigma_\alpha^2), \quad p(\sigma_\alpha) \propto \text{Unif}(0, \infty)$$

$$\delta_i \sim N(d_i, \sigma_\delta^2), \quad p(\sigma_\delta) \propto \text{Unif}(0, \infty)$$

$$a_j \sim N(\mu_a, \sigma_a^2), \quad p(\mu_a) \propto 1, \quad p(\sigma_a) \propto \text{Unif}(0, \infty)$$

$$d_j \sim N(\mu_d, \sigma_d^2), \quad p(\mu_d) \propto 1, \quad p(\sigma_d) \propto \text{Unif}(0, \infty)$$

การประมาณค่าพารามิเตอร์ในแบบฉบับดังกล่าวมีขั้นตอนดังนี้

- ก. สร้าง Likelihood จากการแจกแจงของ Y_{it}
- ข. สร้าง Posterior จากผลคูณของ Likelihood กับ prior ทุกตัว
- ค. จำลองสถานการณ์ด้วยวิธีการ MCMC โดยการใช้การเขียนโปรแกรม ใน Open bugs และ R หรือโปรแกรมคณิตศาสตร์ต่างๆ

9) Markov Chain Monte Carlo (MCMC) และ Gibbs sampling (Robert, 2004)

MCMC เป็นวิธีเชิงตัวเลขที่ใช้สำหรับสร้างข้อมูลจากการแจกแจงที่มีมิติขนาดใหญ่ ในตัวแบบเบย์ เป้าหมายหลักคือการสร้าง $\theta^{(0)}, \theta^{(1)}, \theta^{(2)}, \dots$ ของ Posterior จากห่วงโซ่มาร์คอฟ (Markov Chain) โดยเริ่มจาก Initial state $\theta^{(0)}$ และเมื่อห่วงโซ่คงที่ในการวนซ้ำรอบที่ T เซตของ $\theta^{(0)}, \dots, \theta^{(T)}$ จะถูกตัดทิ้ง เรียกว่า ช่วงของการ burn-in และ $\theta^{(T+1)}, \theta^{(T+2)}, \theta^{(T+3)}, \dots$ เป็นห่วงโซ่ที่คงที่ (Stationary) แล้ว ที่สร้างมาจาก Posterior มีหลายวิธีในการสร้าง MCMC แต่วิธีที่ใช้กันอย่างแพร่หลายคือ Gibbs sampling

Gibbs sampling (Geman and Geman, 1984)

เป็นวิธีการสร้าง MCMC จากการสุ่มตัวอย่างแบบวนซ้ำจากการแจกแจงแบบมีเงื่อนไขของแต่ละพารามิเตอร์เมื่อรู้พารามิเตอร์ที่เหลือทั้งหมดและข้อมูล สมมติว่า Posterior คือ $\pi(\theta | Y)$ ที่มีมิติขนาด k โดยที่ Y แทนข้อมูลที่สังเกตค่าได้ และสำหรับแต่ละ θ_i ของ θ การแจกแจงแบบมีเงื่อนไขของแต่ละพารามิเตอร์เมื่อรู้พารามิเตอร์ที่เหลือทั้งหมดและข้อมูลคือ $\pi(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k, Y) = \pi(\theta_i | \theta_{-i}, Y)$ Gibbs sampling เป็นกระบวนการวนซ้ำมีขั้นตอนดังนี้

$$\text{- กำหนดค่าเริ่มต้น } \theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)}) \quad (40)$$

- รอบที่ i จะเป็นการเปลี่ยนสถานะจาก $\theta^{(i)}$ ไปเป็น $\theta^{(i+1)}$ มีขั้นตอนดังนี้

1. สุ่ม $\theta_1^{(i)}$ จาก $\pi(\theta_1 | \theta_2^{(i-1)}, \theta_3^{(i-1)}, \dots, \theta_k^{(i-1)}, Y)$

2. สุ่ม $\theta_2^{(i)}$ จาก $\pi(\theta_2 | \theta_1^{(i)}, \theta_3^{(i-1)}, \dots, \theta_k^{(i-1)}, Y)$

.

.

3. สุ่ม $\theta_k^{(i)}$ จาก $\pi(\theta_k | \theta_1^{(i)}, \theta_2^{(i)}, \dots, \theta_{k-1}^{(i)}, Y)$

ลำดับของการสุ่ม $\theta_1^{(1)}, \theta_2^{(2)}, \dots, \theta^{(T)}$ เป็นสถานะต่อเนื่องกันของ Markov Chain

10) การวัดความถูกต้องแม่นยำของการพยากรณ์ (Najafi and

Tarazkar, 2006), (Yelland, 2010)

การวัดความถูกต้องแม่นยำของการพยากรณ์นั้นเป็นการเลือกวิธีการพยากรณ์ที่เหมาะสมกับข้อมูลอนุกรมเวลาในแต่ละชุด วิธีวัดความถูกต้องแม่นยำของการพยากรณ์นั้นมีหลายวิธีแต่วิธีที่ใช้กันมากก็คือ เราจะใช้การพิจารณาจากค่าวัดความถูกต้อง ซึ่งต่างเป็นฟังก์ชันของค่าความคลาดเคลื่อน e_t โดยที่ e_t เป็นผลต่างของค่าจริง (Y_t) กับค่าพยากรณ์ ($\hat{Y}_t$) ณ เวลา t ดังนี้

1. Mean Squared Error (MSE)

$$\frac{\sum_{t=1}^n e_t^2}{n} = \frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{n} \quad (41)$$

วิธี MSE เป็นวิธีที่ใช้กันทั่วไป ข้อเสียของวิธีนี้คือไม่มีฐานการเปรียบเทียบ และถ้า MSE มีค่าสูงอาจเป็นเพราะมีความคลาดเคลื่อนสูง หรือขึ้นอยู่กับขนาดของข้อมูล

2. Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{\sum_{t=1}^n |e_t / Y_t|}{n} \times 100 \quad (42)$$

วิธี MAPE เป็นหนึ่งในวิธีที่ถูกยอมรับ และที่ใช้ในการเปรียบเทียบมากที่สุดสำหรับอนุกรมเวลา

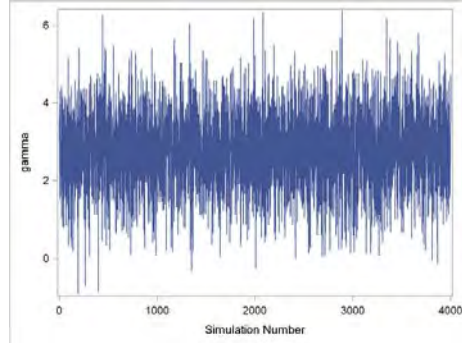
3. Mean Absolute Error (MAE)

$$MAE = \frac{\sum_{t=1}^n |e_t|}{n} = \frac{\sum_{t=1}^n |Y_t - \hat{Y}_t|}{n} \quad (43)$$

เมื่อค่า MSE (Mean Squared Error) MAPE (Mean Absolute Percentage Error) และ MAE (Mean Absolute Error) มีค่าต่ำ แสดงถึง วิธีการพยากรณ์นั้นมีความถูกต้องมาก

11) การตรวจสอบการลู่เข้าของ MCMC โดยดูที่ Trace plots

Trace plots เป็นตัวชี้วัดตัวหนึ่งที่สามารถตรวจสอบการลู่เข้าของ MCMC ซึ่งมันจะช่วยให้บอกว่า chain ของพารามิเตอร์แต่ละตัวลู่หรือยังไม่ลู่เข้าสู่ stationary distribution และมันยังเป็นตัวช่วยบอกว่าอีกนานเท่าไรมันถึงจะลู่เข้า trace plots ยังสามารถบอกเราว่ามันลู่เข้าดีหรือยังไม่ดีอีกด้วย ดังตัวอย่างในภาพที่ 1 ที่แสดง trace plots ที่ลู่เข้าสู่ stationary distribution ที่ดี (SAS, 2011)



ภาพที่ 1 Trace plots ที่ลู่เข้าสู่ stationary distribution ที่ดีของ gamma

12) การจำลองสถานการณ์สำหรับประเมินประสิทธิภาพของตัวแบบ (Bernd, 2004)

ประสิทธิภาพของตัวประมาณจะถูกประเมินจาก 3 ตัว ที่นิยมใช้กันอยู่ทั่วไปซึ่งได้แก่ Relative Bias (RB) Mean Squared Error (MSE) และ the coverage probability (CP) ซึ่งตัวประเมินแต่ละตัวจะคำนวณจากเซตของข้อมูลที่เป็นอิสระกันจากการจำลองสถานการณ์ที่มาจากกระบวนการของ MCMC โดยที่มี T เซต โดยที่ $T = T_1, \dots, T_B$ และ S มีจำนวนที่ใหญ่พอ สามารถแสดงสูตรของตัวประเมินแต่ละตัวดังสมการข้างล่าง

$$\bar{mean} = S^{-1} \sum_{s=1}^S T_s^{(k)} = \bar{T}^{(k)}, \quad (44)$$

$$\bar{bias} = \bar{T}^{(k)} - \mu \quad (45)$$

$$\bar{SD} = \sqrt{(S-1)^{-1} \sum_{s=1}^S (T_s^{(k)} - \bar{T}^{(k)})^2}, \quad (46)$$

$$\bar{MSE} = S^{-1} \sum_{s=1}^S (T_s^{(k)} - \mu)^2 \approx \bar{SD}^2 + \bar{bias}^2 \quad (47)$$

$$RB(\hat{\mu}_Y) = \frac{1}{B} \sum_{b=1}^B \frac{\hat{\mu}_Y^{(b)} - \mu}{\mu} \quad (48)$$

$$MSE(\hat{\mu}_Y) = \frac{1}{B} \sum_{b=1}^B (\hat{\mu}_Y^{(b)} - \mu)^2 \quad (49)$$

$$CP(\hat{\mu}_Y) = \frac{1}{B} \sum_{b=1}^B I(\hat{\mu}_L^{(b)} < \mu < \hat{\mu}_U^{(b)}) \quad (50)$$

2.2 ทฤษฎีที่เกี่ยวข้องกับหุ้น (ตลาดหลักทรัพย์แห่งประเทศไทย, 2557)

ค่าสถิติเบื้องต้นของหลักทรัพย์ (Security) หมวดธุรกิจ (Sector) กลุ่มอุตสาหกรรม (Industry) และค่าตลาดรวม (Market) ที่ตลาดหลักทรัพย์ เผยแพร่ผ่านระบบข้อมูลหลักทรัพย์มีทั้งหมด 8 ตัวได้แก่

- 2.2.1 อัตราส่วนราคาปิดต่อกำไรต่อหุ้น (Price-Earning Ratio: P/E)
- 2.2.2 อัตราส่วนราคาปิดต่อมูลค่าหุ้นทางบัญชี (Price-Book Value: P/BV)
- 2.2.3 มูลค่าหุ้นทางบัญชี (Book Value Per Share: BVPS)
- 2.2.4 อัตราส่วนราคาปิดต่อมูลค่าทรัพย์สินสุทธิ (Price-Net Asset Value: P/NAV)
- 2.2.5 มูลค่าทรัพย์สินสุทธิ (Net Asset Value : NAV)
- 2.2.6 อัตราส่วนเงินปันผลต่อผลตอบแทน (Dividend Yield: DIY)
- 2.2.7 มูลค่าหลักทรัพย์ตามราคาตลาด (Market Capitalization)
- 2.2.8 อัตราหมุนเวียนการซื้อขายหลักทรัพย์ (Turnover Ratio) แสดงรายละเอียดดังตารางที่ 1-8

ทั้งนี้ตลาดหลักทรัพย์ฯ ได้กำหนดเงื่อนไขในการคำนวณค่าสถิติหลักๆ ดังนี้

- 1) ในการคำนวณค่าสถิติของหมวดธุรกิจ กลุ่มอุตสาหกรรม และตลาดรวม ทางตลาดหลักทรัพย์ฯ จะเลือกใช้วิธีการเฉลี่ยถ่วงน้ำหนัก (Weighted Average) ตามมูลค่าหลักทรัพย์ตามราคาตลาดรวม (Market Capitalization)
- 2) ค่าสถิติของตลาดหลักทรัพย์ฯ จะถูกคำนวณบนพื้นฐานของผลการดำเนินงานในรอบ 1 ปีที่ผ่านมา โดยจะอิงจากตัวเลขที่ประกาศในงบการเงินล่าสุด

ตารางที่ 1 อัตราส่วนราคาปิดต่อกำไรต่อหุ้น (Price-Earning Ratio: P/E)

ความหมาย	อัตราส่วนเปรียบเทียบระหว่างราคาตลาดของหุ้นสามัญต่อกำไรสุทธิต่อหุ้นสามัญ (EPS) ที่บริษัทนั้นทำได้ในรอบระยะเวลา 12 เดือนหรือในรอบ 1 ปีล่าสุด
ประเภทหลักทรัพย์ที่ถูกคำนวณ	✓ หลักทรัพย์ (เฉพาะประเภทหุ้นสามัญ ยกเว้น หลักทรัพย์ที่จัดอยู่ใน หมวดธุรกิจ กองทุนรวม อสังหาริมทรัพย์) ✓ หมวดธุรกิจ ✓ กลุ่มอุตสาหกรรม ✓ ตลาดรวม (SET, mai) , SET50 และ SET100
สูตรในการคำนวณค่าของหลักทรัพย์	- สำหรับข้อมูลในอดีตถึงวันที่ 30 มิถุนายน พ.ศ. 2547 $\frac{\text{ราคาปิดของหุ้นสามัญ} \times (\text{จำนวนหุ้นสามัญ} - \text{จำนวนหุ้นซื้อคืน})}{\text{กำไรงวด 12 เดือนล่าสุด}}$ - สำหรับข้อมูลตั้งแต่วันที่ 2 กรกฎาคม พ.ศ. 2547 เป็นต้นไป $\frac{\text{ราคาปิดของหุ้นสามัญ} \times [(\text{จำนวนหุ้นสามัญ} + \text{จำนวนหุ้นบริมสิทธิ์}) - \text{จำนวนหุ้นซื้อคืน}]}{\text{กำไรงวด 12 เดือนล่าสุด}}$

ตารางที่ 2 อัตราส่วนราคาปิดต่อมูลค่าหุ้นทางบัญชี (Price-Book Value: P/BV)

ความหมาย	มูลค่าตามราคาตลาดโดยรวมของหลักทรัพย์จดทะเบียน ซึ่งเป็นค่าที่คำนวณจากการนำราคาปิดของหลักทรัพย์จดทะเบียน คูณกับจำนวนหลักทรัพย์จดทะเบียนปัจจุบัน (Listed Shares)
ประเภทหลักทรัพย์ที่ถูกคำนวณ	✓ หลักทรัพย์ (เฉพาะประเภทหุ้นสามัญ, หุ้นบุริมสิทธิ์, Warrant, หน่วยลงทุน) ✓ หมวดธุรกิจ ✓ กลุ่มอุตสาหกรรม ✓ ตลาดรวม (SET, mai) , SET50 และ SET100
สูตรในการคำนวณค่าของหลักทรัพย์	$\text{ราคาปิดของหุ้น} \times \text{ปริมาณหุ้นจดทะเบียนกับตลาดหลักทรัพย์}$
หมายเหตุ	การคำนวณค่า Market Capitalization ของตลาดรวม (SET, mai) ตลท.ฯ จะเลือกหลักทรัพย์ประเภทหุ้นสามัญ, หุ้นบุริมสิทธิ์และ Warrant มาใช้ในการคำนวณ การคำนวณค่า Market Capitalization ของ SET50, SET100 และกลุ่ม/หมวดอุตสาหกรรม ตลท.ฯ จะเลือกหลักทรัพย์ประเภทหุ้นสามัญเท่านั้น มาใช้ในการคำนวณ

ตารางที่ 3 มูลค่าหุ้นทางบัญชี (Book Value Per Share: BVPS)

ความหมาย	อัตราส่วนการหมุนเวียนการซื้อขาย ซึ่งใช้วัดปริมาณการซื้อขายหลักทรัพย์เมื่อเทียบกับปริมาณหุ้นจดทะเบียน (Listed Shares) โดยคำนวณค่าเป็นเปอร์เซ็นต์	
ประเภท หลักทรัพย์ที่ถูก คำนวณ	✓ หลักทรัพย์ (เฉพาะประเภทหุ้นสามัญ) ✓ หมวดธุรกิจ ✓ กลุ่มอุตสาหกรรม ✓ ตลาดรวม (SET, mai) , SET50 และ SET100	
สูตรในการ คำนวณค่าของ หลักทรัพย์	รายวัน	$\frac{\text{ปริมาณการซื้อขายหลักทรัพย์ ณ วันนั้น} \times 100}{\text{ปริมาณหุ้นจดทะเบียนกับตลาดหลักทรัพย์ ณ วันนั้น}}$
	รายเดือน รายไตรมาส รายปี	$\frac{\text{ผลรวมปริมาณการซื้อขายหลักทรัพย์ในช่วงเวลานั้น} \times 100}{\text{ค่าเฉลี่ยปริมาณหุ้นจดทะเบียนกับตลาดหลักทรัพย์ในช่วงเวลานั้น}}$
หมายเหตุ	ปริมาณการซื้อขายหลักทรัพย์จะใช้ผลรวมของทุกกระดาน คือ Main, Foreign, Odd Lot, Big Lot และรวมปริมาณซื้อขายในช่วง Off-hour ด้วย	

ตารางที่ 4 อัตราส่วนราคาปิดต่อมูลค่าทรัพย์สินสุทธิ (Price-Net Asset Value: P/NAV)

ความหมาย	อัตราส่วนเปรียบเทียบระหว่างราคาปิดของ หลักทรัพย์ที่จัดอยู่ใน Property Fund ต่อ มูลค่าทรัพย์สินสุทธิต่อหน่วยที่รายงานผ่านงบการเงิน ณ สิ้นงวด
ประเภท หลักทรัพย์ที่ถูก คำนวณ	✓ หลักทรัพย์ (เฉพาะหลักทรัพย์ที่จัดอยู่ใน Property Fund Sector และหน่วยลงทุน (Unit Trusts)) ✓ หมวดธุรกิจ ✓ กลุ่มอุตสาหกรรม ✓ ตลาด SET, mai , SET50 และ SET100
สูตรในการ คำนวณค่าของ หลักทรัพย์	■ สำหรับข้อมูลตั้งแต่วันที่ 30 มีนาคม พ.ศ. 2552 เป็นต้นไป $\frac{(\text{ราคาปิดของหุ้น} \times \text{จำนวนหุ้นสามัญ})}{\text{สินทรัพย์สุทธิ}}$

ตารางที่ 5 มูลค่าทรัพย์สินสุทธิ (Net Asset Value : NAV)

ความหมาย	▪ มูลค่าทรัพย์สินสุทธิต่อหน่วย (ที่รายงานผ่านงบการเงิน)
ประเภท หลักทรัพย์ที่ถูก คำนวณ	✓ หลักทรัพย์ (หุ้นสามัญเฉพาะที่จัดอยู่ใน Property Fund Sector และหลักทรัพย์ประเภท หน่วยลงทุน) ✗ หมวดธุรกิจ ✗ กลุ่มอุตสาหกรรม ✗ ตลาดรวม (SET,mai) , SET50 และ SET100
สูตรในการ คำนวณค่าของ หลักทรัพย์	▪ สำหรับข้อมูลตั้งแต่วันที่ 30 มีนาคม พ.ศ. 2552 เป็นต้นไป สินทรัพย์สุทธิ จำนวนหน่วยลงทุนที่จำหน่ายแล้ว

ตารางที่ 6 อัตราส่วนเงินปันผลตอบแทน (Dividend Yield: DIY)

ความหมาย	▪ อัตราเปรียบเทียบเงินปันผลจ่ายต่อหุ้นสามัญเทียบกับราคาตลาดของหุ้นสามัญเพื่อดูผลตอบแทนว่าหากลงทุนซื้อหุ้น ณ ระดับราคาตลาดปัจจุบัน จะมีโอกาสได้รับเงินปันผลคิดเป็นอัตราร้อยละเท่าไรของราคาหุ้น
ประเภท หลักทรัพย์ที่ถูก คำนวณ	✓ หลักทรัพย์ (เฉพาะประเภทหุ้นสามัญ) ✓ หมวดธุรกิจ ✓ กลุ่มอุตสาหกรรม ✓ ตลาดรวม (SET,mai) , SET50 และ SET100
สูตรในการ คำนวณค่าของ หลักทรัพย์	▪ สำหรับข้อมูลในอดีตถึงวันที่ 30 มิถุนายน พ.ศ. 2547 มูลค่าเงินปันผลรวมที่ประกาศจ่ายล่าสุด 12 เดือน ราคาปิดของหุ้นสามัญ X (จำนวนหุ้นสามัญ - จำนวนหุ้นซื้อคืน) ▪ สำหรับข้อมูลตั้งแต่วันที่ 2 กรกฎาคม พ.ศ. 2547 เป็นต้นไป มูลค่าเงินปันผลรวมยึดตามรอบผลประกอบการประจำปี (ล่าสุด) ราคาปิดของหุ้นสามัญ X (จำนวนหุ้นสามัญ - จำนวนหุ้นซื้อคืน)

ตารางที่ 7 มูลค่าหลักทรัพย์ตามราคาตลาด (Market Capitalization)

ความหมาย	มูลค่าตามราคาตลาดโดยรวมของหลักทรัพย์จดทะเบียน ซึ่งเป็นค่าที่คำนวณจากการนำราคาปิดของหลักทรัพย์จดทะเบียน คูณกับจำนวนหลักทรัพย์จดทะเบียนปัจจุบัน (Listed Shares)
ประเภทหลักทรัพย์ที่ถูกคำนวณ	✓ หลักทรัพย์ (เฉพาะประเภทหุ้นสามัญ, หุ้นบุริมสิทธิ, Warrant, หน่วยลงทุน) ✓ หมวดธุรกิจ ✓ กลุ่มอุตสาหกรรม ✓ ตลาดรวม (SET,mai) , SET50 และ SET100
สูตรในการคำนวณค่าของหลักทรัพย์	ราคาปิดของหุ้น X ปริมาณหุ้นจดทะเบียนกับตลาดหลักทรัพย์
หมายเหตุ	การคำนวณค่า Market Capitalization ของตลาดรวม (SET, mai) ตลท.ฯ จะเลือกหลักทรัพย์ประเภทหุ้นสามัญ, หุ้นบุริมสิทธิและ Warrant มาใช้ในการคำนวณ การคำนวณค่า Market Capitalization ของ SET50, SET100 และกลุ่ม/หมวดอุตสาหกรรม ตลท.ฯ จะเลือกหลักทรัพย์ประเภทหุ้นสามัญเท่านั้น มาใช้ในการคำนวณ

ตารางที่ 8 อัตราหมุนเวียนการซื้อขายหลักทรัพย์ (Turnover Ratio)

ความหมาย	อัตราหมุนเวียนการซื้อขาย ซึ่งใช้วัดปริมาณการซื้อขายหลักทรัพย์เมื่อเทียบกับปริมาณหุ้นจดทะเบียน (Listed Shares) โดยคำนวณค่าเป็นเปอร์เซ็นต์	
ประเภทหลักทรัพย์ที่ถูกคำนวณ	✓ หลักทรัพย์ (เฉพาะประเภทหุ้นสามัญ) ✓ หมวดธุรกิจ ✓ กลุ่มอุตสาหกรรม ✓ ตลาดรวม (SET,mai) , SET50 และ SET100	
สูตรในการคำนวณค่าของหลักทรัพย์	รายวัน	$\frac{\text{ปริมาณการซื้อขายหลักทรัพย์ ณ วันนั้น} \times 100}{\text{ปริมาณหุ้นจดทะเบียนกับตลาดหลักทรัพย์ ณ วันนั้น}}$
	รายเดือน รายไตรมาส รายปี	$\frac{\text{ผลรวมปริมาณการซื้อขายหลักทรัพย์ในช่วงเวลานั้น} \times 100}{\text{ค่าเฉลี่ยปริมาณหุ้นจดทะเบียนกับตลาดหลักทรัพย์ในช่วงเวลานั้น}}$
หมายเหตุ	ปริมาณการซื้อขายหลักทรัพย์จะใช้ผลรวมของทุกกระดาน คือ Main, Foreign, Odd Lot, Big Lot และรวมปริมาณซื้อขายในช่วง Off-hour ด้วย	

งานวิจัยที่เกี่ยวข้อง

ตัวแบบการพยากรณ์ของไทยในปัจจุบันใช้ตัวแบบการพยากรณ์ส่วนใหญ่ยังใช้ตัวแบบการพยากรณ์แบบดั้งเดิม ทั่วๆไป (Classical model) เช่น Multiple Regression, Moving Average (MA), Exponential Smoothing (EXPS), ARIMA และ Seasonal ARIMA เป็นต้น ตัวแบบโครงข่ายประสาทเทียม (Artificial Neural Network) ตัวแบบ Semiparametric Multiple Regression เป็นต้น ตัวแบบที่มีความซับซ้อนมากขึ้น เช่นตัวแบบที่มีการนำความสัมพันธ์ของข้อมูลอนุกรมเวลามาคิดคำนวณด้วย และตัวแบบที่กำหนดให้พารามิเตอร์ในตัวแบบมีความไม่แน่นอน (uncertainty) ยังมีน้อยมาก ส่วนใหญ่จะอยู่ในต่างประเทศ สำหรับในงานวิจัยนี้ได้ศึกษางานวิจัยที่เกี่ยวข้องที่ใช้ตัวแบบการพยากรณ์ที่นำมาใช้กับด้านอุตสาหกรรม และนำไปใช้กับงานด้านอื่นๆ ดังนี้

ตัวแบบการพยากรณ์แบบดั้งเดิม (Classical Models) ในการพยากรณ์ผลลัพธ์ที่จะเกิดขึ้นในอนาคตของเหตุการณ์ด้านต่างๆ นั้น Wright (1986) นำตัวแบบ Simple Exponential และ Holt ไปใช้กับข้อมูลอนุกรมเวลาที่ไม่สม่ำเสมอ (Irregular) Deetee (1991) ประยุกต์ใช้ตัวแบบ Box Jenkins (ARIMA) ในการพยากรณ์ราคาข้าว พบว่าตัวแบบนี้มีประสิทธิภาพอย่างเห็นได้ชัด และดีกว่าตัวแบบ Decomposition เช่นเดียวกันกับที่ Kerdsoomboon (1999) พบว่า ตัวแบบ Box Jenkins พยากรณ์ราคาข้าวได้ดีกว่าตัวแบบสถิติเบื้องต้น Cipraและคณะ (1995) ใช้ตัวแบบ Holt-Winter กับข้อมูลอนุกรมเวลาที่มีข้อมูลสูญหาย (Missing) Hyndman (2002) พบว่าการใช้ตัวแบบ Single Exponential Smoothing ในกรอบการทำงาน (Framework) ของตัวแบบ State-space มีประสิทธิภาพสูง โดยนำไปใช้กับข้อมูลอนุกรมเวลา M-Competition ซึ่งเป็นข้อมูลทางด้านเศรษฐศาสตร์ การเงิน อุตสาหกรรม ประชากรศาสตร์ และอื่นๆ ในขณะที่ Sangpattaranate (2005) ที่พบว่า Box Jenkins เป็นตัวแบบที่ใช้ในการพยากรณ์ราคาข้าวได้ดีกว่าตัวแบบ Holt-Winter และตัวแบบการถดถอย Iqbal และคณะ (2005) ใช้ตัวแบบ ARIMA พยากรณ์ผลผลิตและพื้นที่เพาะปลูกข้าวสารในประเทศไทยปาล์มน้ำมัน เพื่อใช้เป็นข้อมูลให้กับรัฐบาลในการกำหนดนโยบาย Mishra และ Desai (2005) ใช้ตัวแบบ SARIMA ในการพยากรณ์ภัยแล้ง (Drought) โดยใช้ข้อมูลอนุกรมเวลาเป็นค่าดัชนีมาตรฐานของหยาดน้ำที่ตกมาจากชั้นบรรยากาศ (Precipitation) สำหรับตัวแบบการพยากรณ์แบบ exponential smoothing นั้นมีการนำไปใช้อย่างแพร่หลายนั้น และมีการยืดขยายเป็น Simple exponential smoothing, Holt, Holt-Winters และ double exponential smoothing Cipra (2006) ใช้ตัวแบบ double exponential smoothing กับข้อมูลอนุกรมเวลาที่ไม่สม่ำเสมอ boosarawongse และคณะ (2007) ได้ใช้วิธีการพยากรณ์ Box-Jenkins(ARIMA)กับ Artificial Neural Network สำหรับพยากรณ์ commodity prices การส่งออกข้าวของไทย 4 ชนิด ผลพบว่าวิธีการพยากรณ์ทั้งสองให้ผลที่ดี แต่วิธีการพยากรณ์ Artificial Neural Network ให้ผลการพยากรณ์ที่แม่นยำข้าวที่ดีกว่า 3 ชนิด โดยเปรียบเทียบจากค่าความคลาดเคลื่อนต่างๆ Sumer และคณะ (2009) ศึกษาการใช้ตัวแบบ ARIMA, SARIMA และ ตัวแบบการถดถอย (Regression Model) ที่มีฤดูกาล (Seasonal) เป็นตัวแปรซ่อนเร้น (Latent variable) ในการ

พยากรณ์ ปริมาณความต้องการกระแสไฟฟ้าพบว่าตัวแบบการถดถอยที่มีฤดูกาลเป็นตัวแปรซ่อนเร้น พยากรณ์ได้แม่นยำกว่าARIMA และSARIMA Kahforoushan, Zarif and Mashahir (2010) ศึกษาการพยากรณ์ผลผลิตทางการเกษตร ซึ่งได้แก่ การปลูกพืช การเลี้ยงสัตว์ การประมง และการปลูกป่า โดยใช้วิธีการพยากรณ์ 4 วิธี ได้แก่ วิธีปรับเรียบเอ็กโปเนนเชียลโดยวิธี Holt-Winters แบบไม่มีฤดูกาล (Holt-Winters (no seasonal) Exponential Smoothing Model) วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins Model) วิธีโครงข่ายประสาทเทียม (Artificial Neural Network Model) และ วิธีARIMA (ARIMA Model) และใช้ ค่า MAE MSE และMAPE เปรียบเทียบผลการพยากรณ์แต่ละวิธี ผลการศึกษาพบว่า วิธีโครงข่ายประสาทเทียมเหมาะสมในการประมาณค่าพารามิเตอร์ในตัวแบบ (Learn Stage) วิธีบ็อกซ์-เจนกินส์เหมาะสมในการประเมินความถูกต้องในตัวแบบ (Model Validation) แต่วิธีปรับเรียบเอ็กโปเนนเชียลโดยวิธี Holt-Winters แบบไม่มีฤดูกาล ให้ค่า MAPE ต่ำสุดในการประมาณค่าพารามิเตอร์ในตัวแบบ (Model Fitting) และการประเมินความถูกต้องในตัวแบบ

นอกจากตัวแบบการพยากรณ์แบบเดิมที่กล่าวมาแล้วนั้น ยังมีอีกตัวแบบหนึ่งที่กำลังได้รับความนิยมเพิ่มมากขึ้นเรื่อยๆ คือตัวแบบการพยากรณ์แบบเบย์ ตัวแบบนี้เหมาะสำหรับกรณีที่กำหนดให้พารามิเตอร์ในตัวแบบมีความไม่แน่นอน (Uncertainty) และยังสามารถนำความสัมพันธ์ของข้อมูลอนุกรมเวลามาคิดคำนวณได้ด้วย ดังปรากฏในงานของ Monahan (1983) ใช้วิธีของเบย์ในการประมาณค่าพารามิเตอร์ในตัวแบบ ARMA และ Broemeling และ Land (1984) ใช้วิธีของเบย์ในการประมาณค่าพารามิเตอร์ในตัวแบบ AR(p) Liu (1994) ใช้วิธีของเบย์ในการประมาณค่าพารามิเตอร์ในตัวแบบ AR(p) ที่มีตัวแปรภายนอก (Exogenous) รวมอยู่ด้วย Neelamegham และ Chintagunta (1999) ประยุกต์ใช้ตัวแบบเบย์ในการพยากรณ์การจำนวนผู้เข้าชมภาพยนตร์ใหม่ ที่เข้าฉายในโรงภาพยนตร์ในสัปดาห์แรกทั้งภายในประเทศและบางประเทศในต่างประเทศ ซึ่งมีประโยชน์ต่อผู้ที่เกี่ยวข้องเช่น เจ้าของโรงภาพยนตร์ ผู้แทนจำหน่าย และผู้จัดทำโฆษณา เป็นต้น จำนวนผู้เข้าชมภาพยนตร์เป็นจำนวนนับ ที่มีการแจกแจงแบบปัวซอง ผลการศึกษาพบว่า ตัวแบบเบย์พยากรณ์ได้แม่นยำในระดับประเทศ และเมื่อพิจารณาค่ารากที่สองของค่าเฉลี่ยกำลังสองของความผิดพลาด (Root Mean Square Error) และค่าเฉลี่ยความผิดพลาดสัมบูรณ์ (Mean Absolute Error) พบว่ามีค่าต่ำกว่าตัวแบบของผู้วิจัยอื่นๆ คือ ตัวแบบของ Sawhney และ Eliashberg (1996) Fourth และ Woodlock (1960) และตัวแบบ Naïve (Logged) OLS และ Poisson Maximum Likelihood de Alba และ Mendoza(2006) ศึกษาการพยากรณ์โดยใช้วิธีเบย์ เมื่อมีข้อมูลจำนวนน้อย ค่าที่พยากรณ์เป็นค่าสะสมของตัวแปรต่อเนื่องที่เป็นค่าบวกโดยทราบค่าสะสมของข้อมูลมาส่วนหนึ่งแล้ว ตัวแบบที่ถูกรวบรวมเสนอเป็นการอธิบายความสัมพันธ์ระหว่างค่ารวมทั้งหมดกับค่ารวมมาแล้วบางส่วนของตัวแปรภายใต้อิทธิพลของฤดูกาลแบบคงที่ (Stable Seasonality) ผลการศึกษาพบว่าตัวแบบที่นำเสนอเหมาะสมเมื่อมีข้อมูลจำนวนน้อย และตัวแบบมาตรฐานทั่วไปไม่เหมาะสม Pedroza (2006) ใช้ตัวแบบเบย์ในการพยากรณ์อัตราการเสียชีวิต ของชายชาวสหรัฐอเมริกา ใช้ MCMC ในการประมาณค่าพารามิเตอร์ และใช้ Gibbs sampling ในการสุ่มตัวอย่างจาก Posterior กลุ่มตัวอย่างเป็นข้อมูลการเสียชีวิตของชายชาวสหรัฐอเมริกา เป็นการพยากรณ์

อัตราการเสียชีวิตในช่วงปี 1990-1999 โดยใช้ข้อมูลปี 1959-1989 การพยากรณ์นี้เมื่อนำไปเปรียบเทียบกับค่าสังเกตจริง และวิธีการของ Lee-Carter พบว่าวิธีการของเบย์เหมาะสมกว่า de Alba และ Mendoza(2007) ที่นำตัวแบบการพยากรณ์แบบเบย์ไปใช้กับข้อมูลอนุกรมเวลาที่มีอิทธิพลของฤดูกาลรวมอยู่ด้วย Yelland (2009) ใช้วิธีของเบย์ประมาณค่าพารามิเตอร์ในตัวแบบ state-space 3 ประเภทคือ Adjusted Gaussian Dynamic Linear Model (AG), Poisson Dynamic Log-Linear Model (PL) และ Gamma-Poisson Local Level Model (GP) รวมทั้ง ตัวแบบ Climatological Baseline Model (Cm) กับข้อมูลปริมาณความต้องการซื้อสินค้า พบว่า ตัวแบบ GP ดีที่สุด Yelland (2010) นำเสนอมีความเหมาะสมกับข้อมูลปริมาณความต้องการซื้อชิ้นส่วนอุปกรณ์คอมพิวเตอร์ และเมื่อเปรียบเทียบกับตัวแบบมาตรฐานอื่นๆ ได้แก่ Exponential smoothing (ExpS) และ Judgmental Methods (Judg) พบว่าวิธีเบย์ มีความเหมาะสมมากกว่า จากที่กล่าวมาทั้งหมดนั้น



บทที่ 3

วิธีการดำเนินงานวิจัย

ในการศึกษาการเลือกฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มเพื่อสร้างตัวแบบ สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาหลายๆตัวพร้อมกันที่มีค่า อัตราสหสัมพันธ์ แนวโน้ม และค่าผิดปกติ ด้วยวิธีการของเบย์: กรณีศึกษาราคาหุ้นที่มีความผันผวนของตลาดหลักทรัพย์แห่งประเทศไทย นี้ผู้ศึกษาได้กำหนดวิธีการศึกษาไว้ตามขั้นตอนดังนี้

- 3.1 ประชากรและกลุ่มตัวอย่าง
- 3.2 เครื่องมือที่ใช้ในการศึกษา
- 3.3 วิธีดำเนินการศึกษา
- 3.4 การสร้างตัวแบบพยากรณ์ด้วยวิธีของเบย์
- 3.5 ประเมินผลและข้อเสนอแนะ

3.1 ประชากรและกลุ่มตัวอย่าง

3.1.1 ประชากร

ประชากรคือ ราคาปิดของหุ้นรายวัน จำนวน 6 เดือน ตั้งแต่วันที่ 1 เมษายน 2555 ถึง 19 กันยายน 2557 จากตลาดหลักทรัพย์แห่งประเทศไทย

3.1.2 กลุ่มตัวอย่าง

ข้อมูลราคาปิดของหุ้นรายวัน จำนวน 6 เดือน ตั้งแต่วันที่ 1 เมษายน 2557 ถึง 19 กันยายน 2557 ที่มีความผันผวนซึ่งวิเคราะห์จากบริษัทหลักทรัพย์เอเชียพลัส โดยพิจารณาจากหุ้นที่มีค่า Beta (ค่าความผันผวนของราคาหุ้นเทียบกับตลาดหุ้น) สูงกว่า 1 และค่าพี/อี ที่สูงกว่า 20 เท่า จำนวน 2 ตัว ได้แก่ บมจ.พลังงานบริสุทธิ์ (EA) บมจ.เนาวรัตน์พัฒนาการ (NWR) และหุ้นที่เป็นที่รู้จักกันทั่วไป อีก 3 ตัว ได้แก่ บมจ. ปตท (PTT) บมจ.ธนาคารกรุงไทย (KTB) บมจ.แอดวานซ์ อินโฟร์ เซอร์วิส (ADVANCE) จากตลาดหลักทรัพย์แห่งประเทศไทย

3.2 เครื่องมือที่ใช้ในการศึกษา

3.2.1 เครื่องคอมพิวเตอร์ Notebook

3.2.2 โปรแกรมสำเร็จรูปทางสถิติ SPSS สำหรับวิเคราะห์ข้อมูล

3.2.3 โปรแกรมคอมพิวเตอร์ OpenBUGS และ R สำหรับสร้าง และวิเคราะห์ตัวแบบในงานวิจัยนี้

3.3 วิธีการดำเนินการศึกษา

3.3.1 เก็บรวบรวมข้อมูล

เก็บรวบรวมข้อมูลราคาปิดของหุ้นรายวัน จำนวน 6 เดือน ตั้งแต่วันที่ 1 เมษายน 2557 ถึง 19 กันยายน 2557 ที่มีความผันผวนซึ่งวิเคราะห์จากบริษัทหลักทรัพย์เอเซียพลัสโดยพิจารณาจากหุ้นที่มีค่า Beta (ค่าความผันผวนของราคาหุ้นเทียบกับตลาดหุ้น) สูงกว่า 1 และค่าพี/อี ที่สูงกว่า 20 เท่า จำนวน 2 ตัว ได้แก่ บมจ.พลังงานบริสุทธิ์ (EA) บมจ.เนาวรัตน์พัฒนาการ (NWR) และหุ้นที่เป็นที่รู้จักกันทั่วไปอีก 3 ตัว ได้แก่ บมจ. ปตท (PTT) บมจ.ธนาคารกรุงไทย (KTB) บมจ. แอควานซ์ อินโฟร เซอร์วิส(ADVANCE) จากตลาดหลักทรัพย์แห่งประเทศไทยโดยจะใช้ข้อมูลตั้งแต่วันที่ 1 เมษายน ถึง 29 สิงหาคม 2557 (99 วัน) สำหรับหาตัวแบบที่เหมาะสม (Fitting Model) และใช้ข้อมูลตั้งแต่วันที่ 1 กันยายน ถึง 19 กันยายน 2557 (15 วัน) สำหรับเปรียบเทียบเพื่อตรวจสอบความถูกต้อง ของตัวแบบ (Validation Model)

3.3.2 วิเคราะห์ข้อมูล

นำข้อมูลข้างต้นที่รวบรวมมาได้มาวิเคราะห์คุณลักษณะของข้อมูลว่ามีส่วนประกอบอะไรบ้าง เช่น แนวโน้ม (Trend) อัตตสหสัมพันธ์ (Autocorrelation) ข้อมูลผิดปกติ (Outlier) และฤดูกาล (Seasonal) หลังจากนั้นสร้างตัวแบบการพยากรณ์ข้อมูลด้วยวิธีของเบย์ โดยเขียนโปรแกรมใน OpenBUGS และตรวจเช็คประสิทธิภาพของตัวแบบด้วยโปรแกรม R

3.4 สร้างตัวแบบด้วยวิธีของเบย์

นำข้อมูลที่ได้จากการวิเคราะห์นำข้อมูลที่ได้จากการวิเคราะห์ในหัวข้อ 3.3.1 และ 3.3.2 ไปสร้างตัวแบบเบย์ที่มีฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวย์บูล กับตัวแบบเบย์ที่มีฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียล สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาหลายๆตัวพร้อมกันที่มีค่า อัตตสหสัมพันธ์ แนวโน้ม และค่าผิดปกติ ด้วยวิธีการประมาณค่าพารามิเตอร์ใช้วิธี MCMC โดยใช้อัลกอริทึมของการสุ่มตัวอย่างแบบ Gibbs ทำการเขียนโปรแกรม ใน OpenBUGS และประเมินประสิทธิภาพของตัวแบบด้วยการ

เขียนโปรแกรมใน R เพื่อเปรียบเทียบวิธีการพยากรณ์ทั้งสองวิธีที่สร้างขึ้นมามากด้วย

3.5 ประเมินผลการวิจัย

นำผลการวิจัยที่ได้มาวิเคราะห์ และพิจารณาเพื่อหาข้อสรุป และข้อเสนอแนะ

สถานที่เก็บรวบรวมข้อมูล

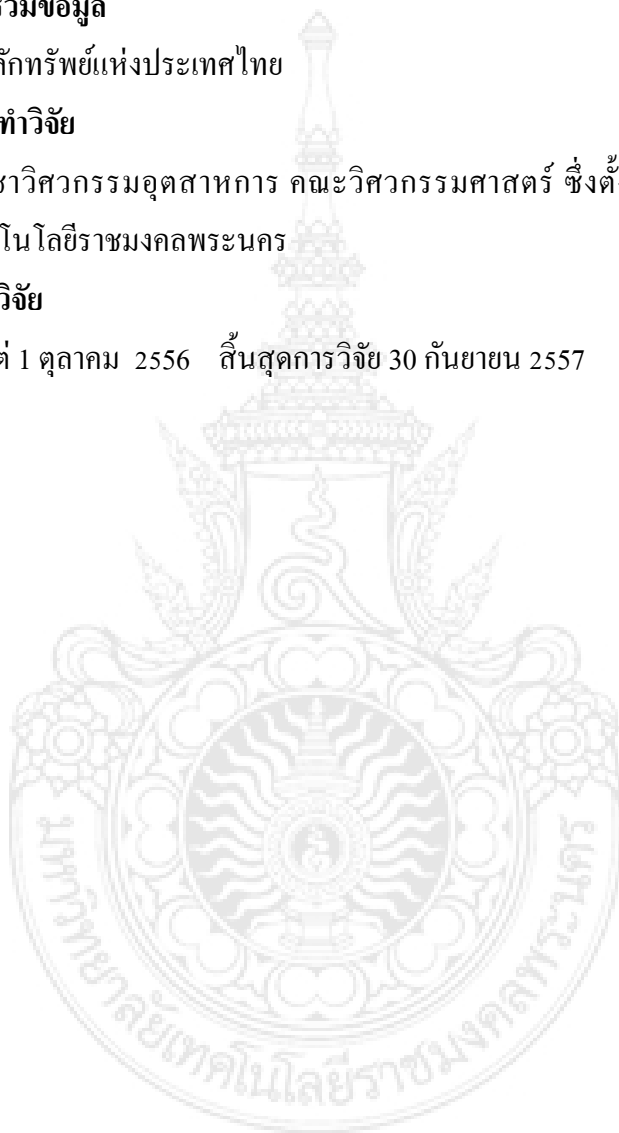
ตลาดหลักทรัพย์แห่งประเทศไทย

สถานที่ใช้ในการทำวิจัย

สาขาวิชาวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์ ซึ่งตั้งอยู่ที่ศูนย์พระนครเหนือ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

ระยะเวลาในการวิจัย

เริ่มตั้งแต่ 1 ตุลาคม 2556 สิ้นสุดการวิจัย 30 กันยายน 2557



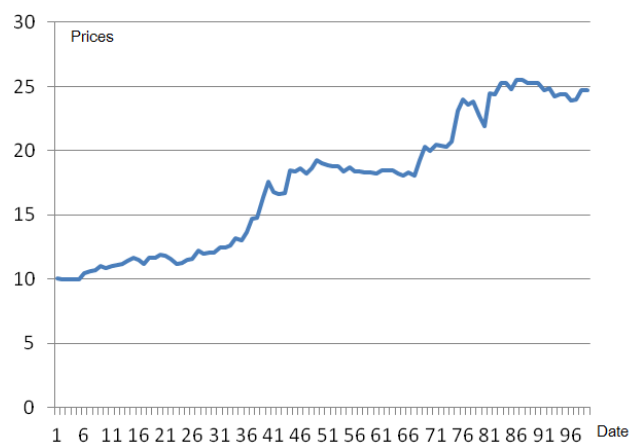
บทที่ 4

ผลการวิจัย

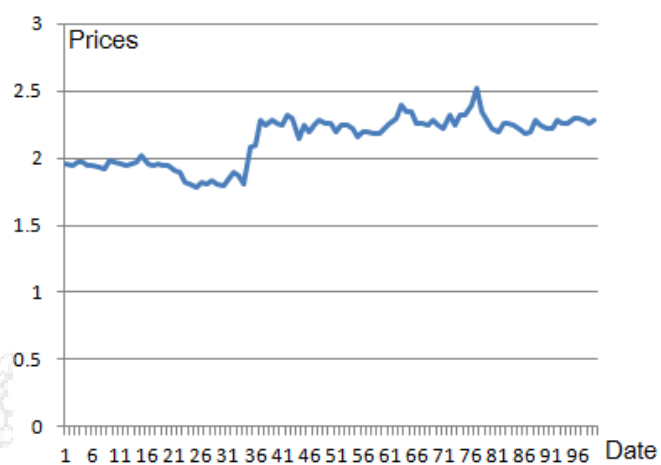
ผลการศึกษาการเลือกฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มเพื่อสร้างตัวแบบ สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาหลายๆตัวพร้อมกันที่มีค่า อัตตสหสัมพันธ์ แนวโน้ม และค่าผิดปกติ ด้วยวิธีการของเบย์:กรณีศึกษาราคาหุ้นที่มีความผันผวนของตลาดหลักทรัพย์แห่งประเทศไทย ในงานวิจัยนี้ได้ศึกษาข้อมูลเฉพาะข้อมูลราคาปิดของหุ้นรายวัน จำนวน 6 เดือน ตั้งแต่วันที่ 1 เมษายน 2557 ถึง 19 กันยายน 2557 ที่มีความผันผวนซึ่งวิเคราะห์จากบริษัทหลักทรัพย์ เอเซียพลัสโดยพิจารณาจากหุ้นที่มีค่า Beta (ค่าความผันผวนของราคาหุ้นเทียบกับตลาดหุ้น) สูงกว่า 1 และค่าพีอี ที่สูงกว่า 20 เท่า จำนวน 2 ตัว ได้แก่ บมจ.พลังงานบริสุทธิ์ (EA) บมจ.เนาวรัตน์พัฒนาการ (NWR) และหุ้นที่เป็นที่รู้จักกันทั่วไปอีก 3 ตัว ได้แก่ บมจ. ปตท (PTT) บมจ. ธนาคารกรุงไทย (KTB) บมจ.แอดวานซ์ อินโฟร์ เซอร์วิส(ADVANCE) จากตลาดหลักทรัพย์แห่งประเทศไทย โดยสร้างตัวแบบพยากรณ์ และประมาณค่าพารามิเตอร์ดังที่ได้แสดงในบทที่ 3 เมื่อได้ค่าพารามิเตอร์ต่างๆแล้ว จึงนำมาใช้พยากรณ์ไปข้างหน้ากับข้อมูลที่เก็บไว้จำนวนอีก 15 วันสุดท้าย เพื่อเปรียบเทียบกับข้อมูลจริงด้วยวิธีการพยากรณ์แบบเบย์ที่มีฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเวกซ์บูลที่เคยเสนอในงานวิจัยไปแล้วกับแบบเอ็กซ์โพเนนเชียล ซึ่งมีรายละเอียดดังนี้

4.1 ผลการวิเคราะห์ลักษณะของข้อมูล

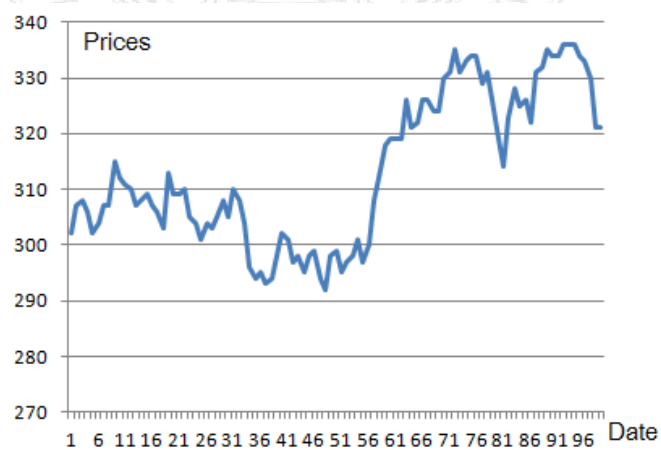
ภาพที่ 5-9 แสดงลักษณะของข้อมูลหุ้นแต่ละตัว ตั้งแต่วันที่ 1 เมษายน พ.ศ. 2557 ถึง วันที่ 29 สิงหาคม พ.ศ. 2557 ภาพที่ 10-14 แสดงการวัดสหสัมพันธ์ระหว่างข้อมูล ด้วยกราฟ Autocorrelation Function (ACF) และ Partial Autocorrelation Function (PACF)



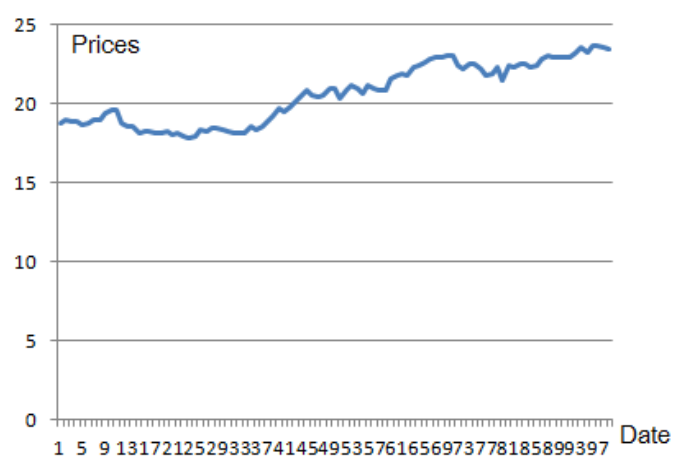
ภาพที่ 2 กราฟแสดงลักษณะของหุ้น EA



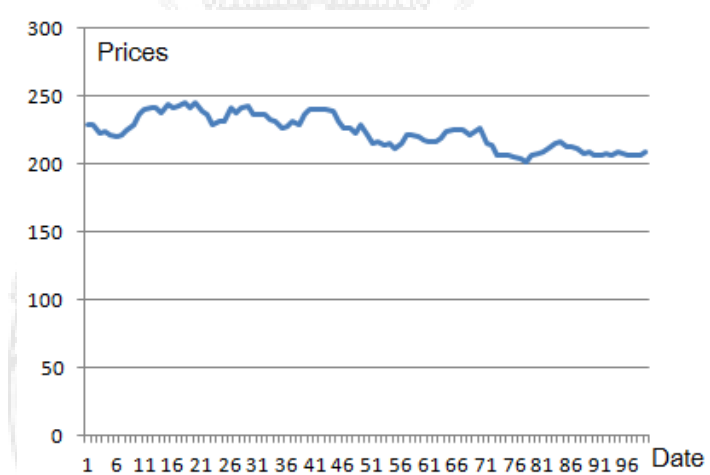
ภาพที่ 3 กราฟแสดงลักษณะของหุ้น NWR



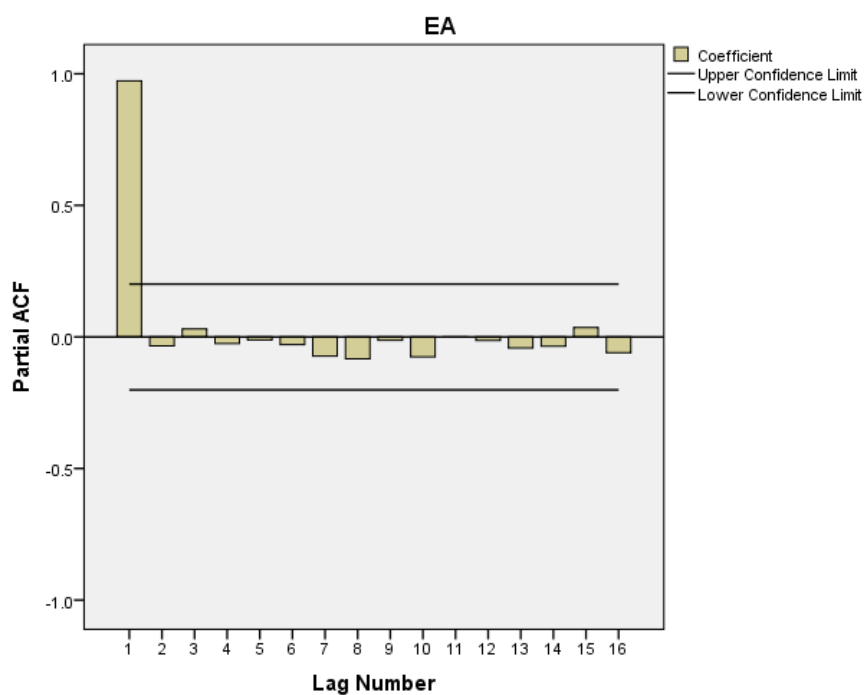
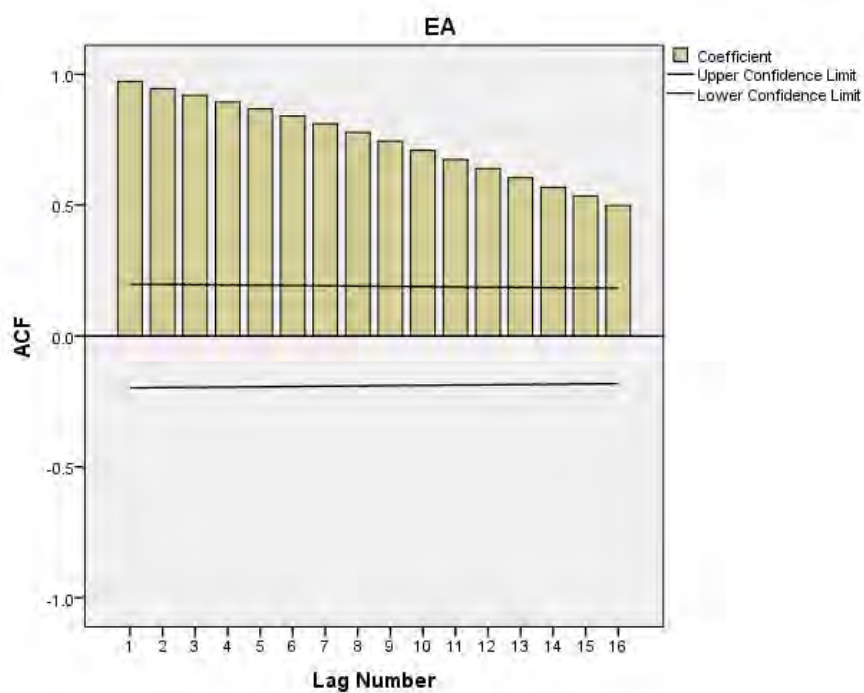
ภาพที่ 4 กราฟแสดงลักษณะของหุ้น PTT



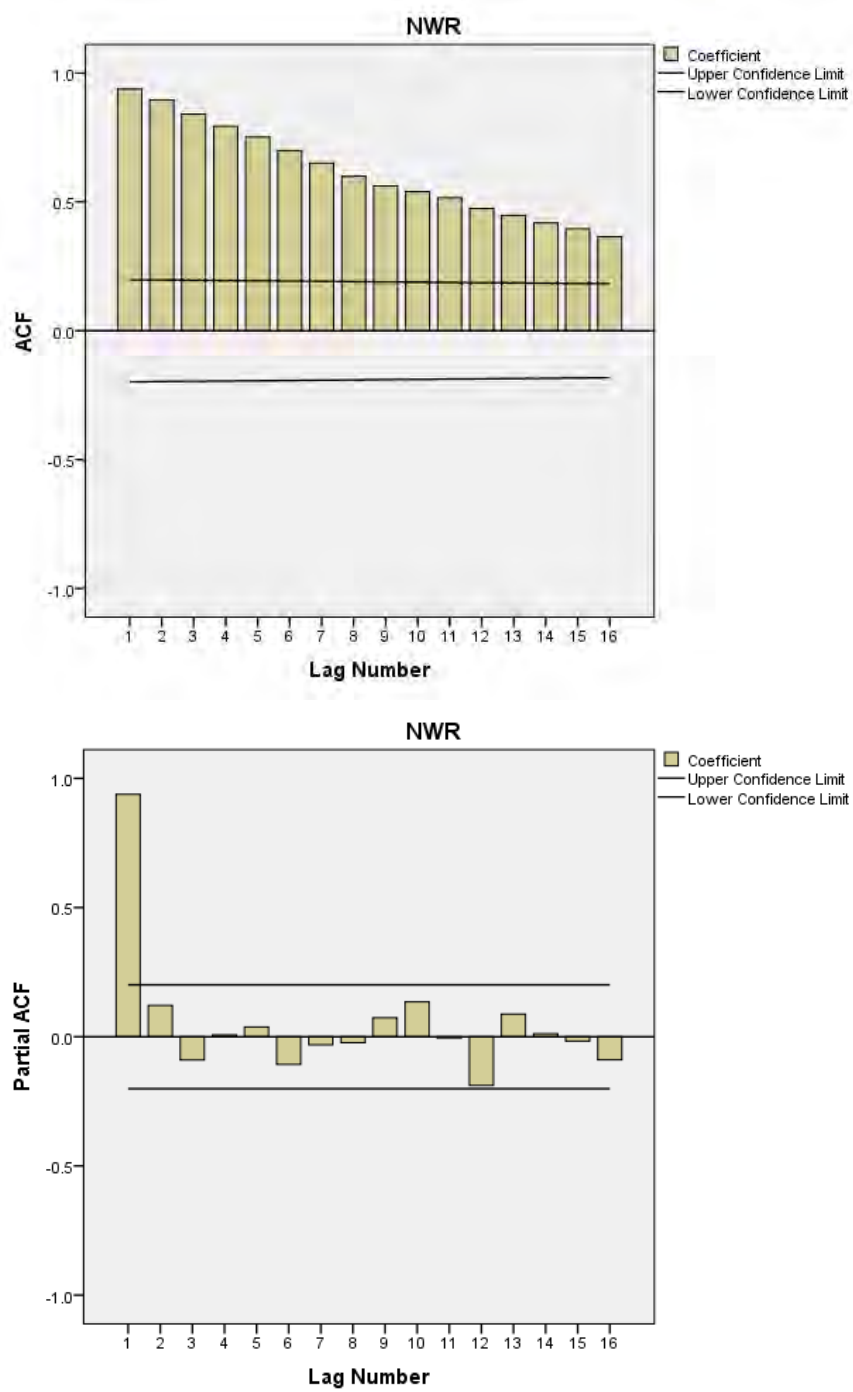
ภาพที่ 5 กราฟแสดงลักษณะของหุ้น KTB



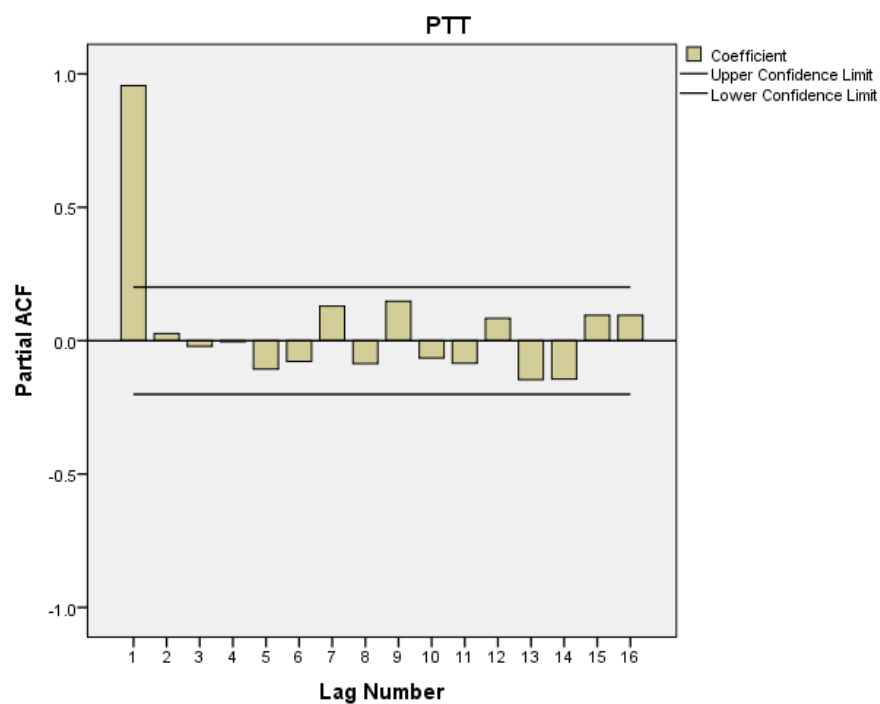
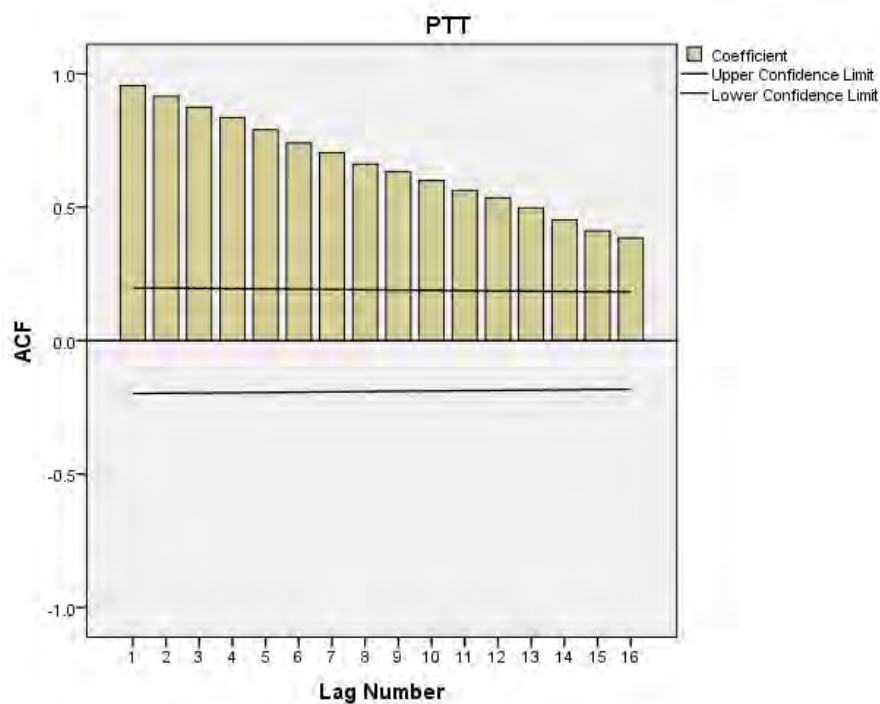
ภาพที่ 6 กราฟแสดงลักษณะของหุ้น ADVANCE



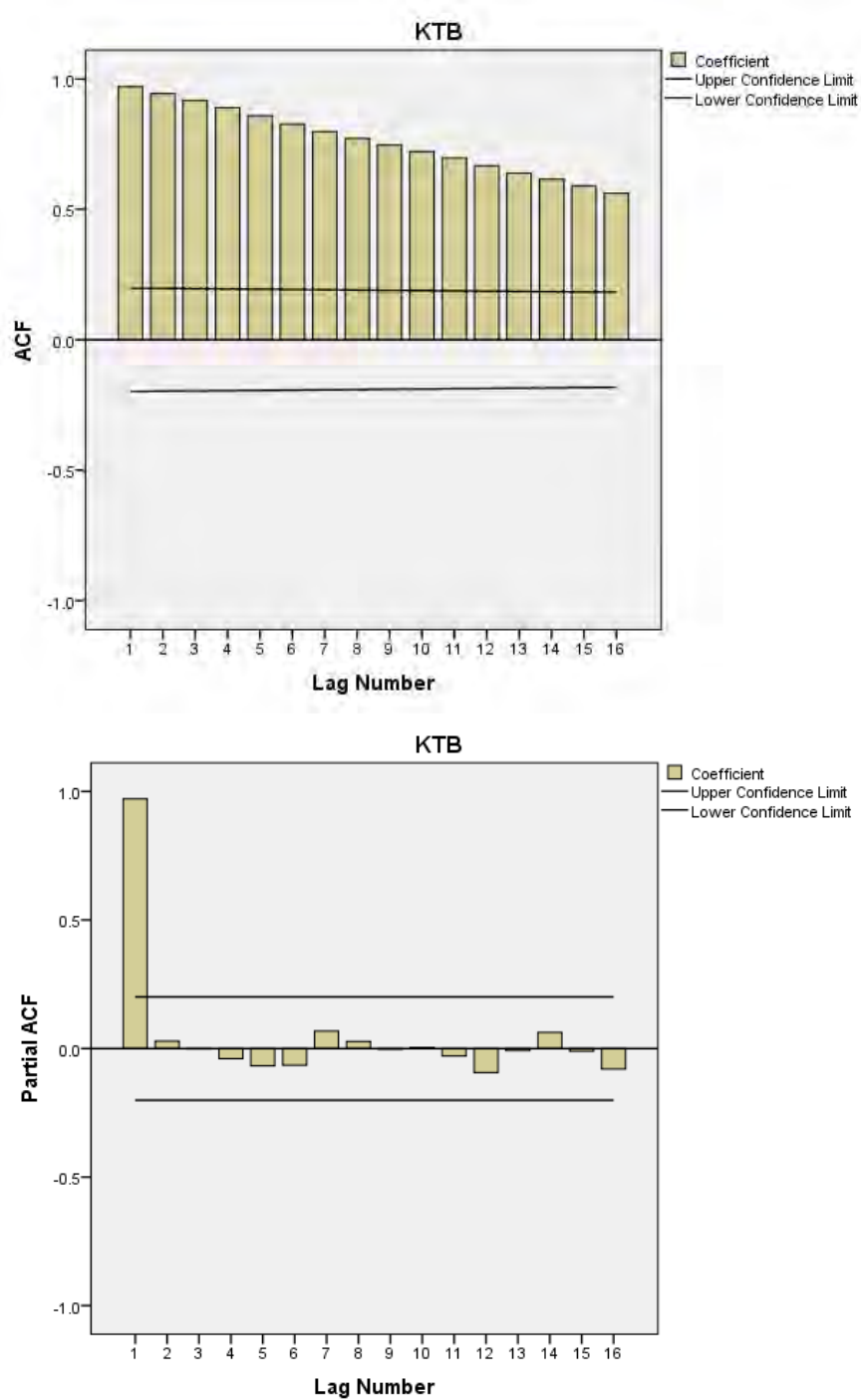
ภาพที่ 7 แสดงการวัดสหสัมพันธ์ข้อมูลหุ้น EA



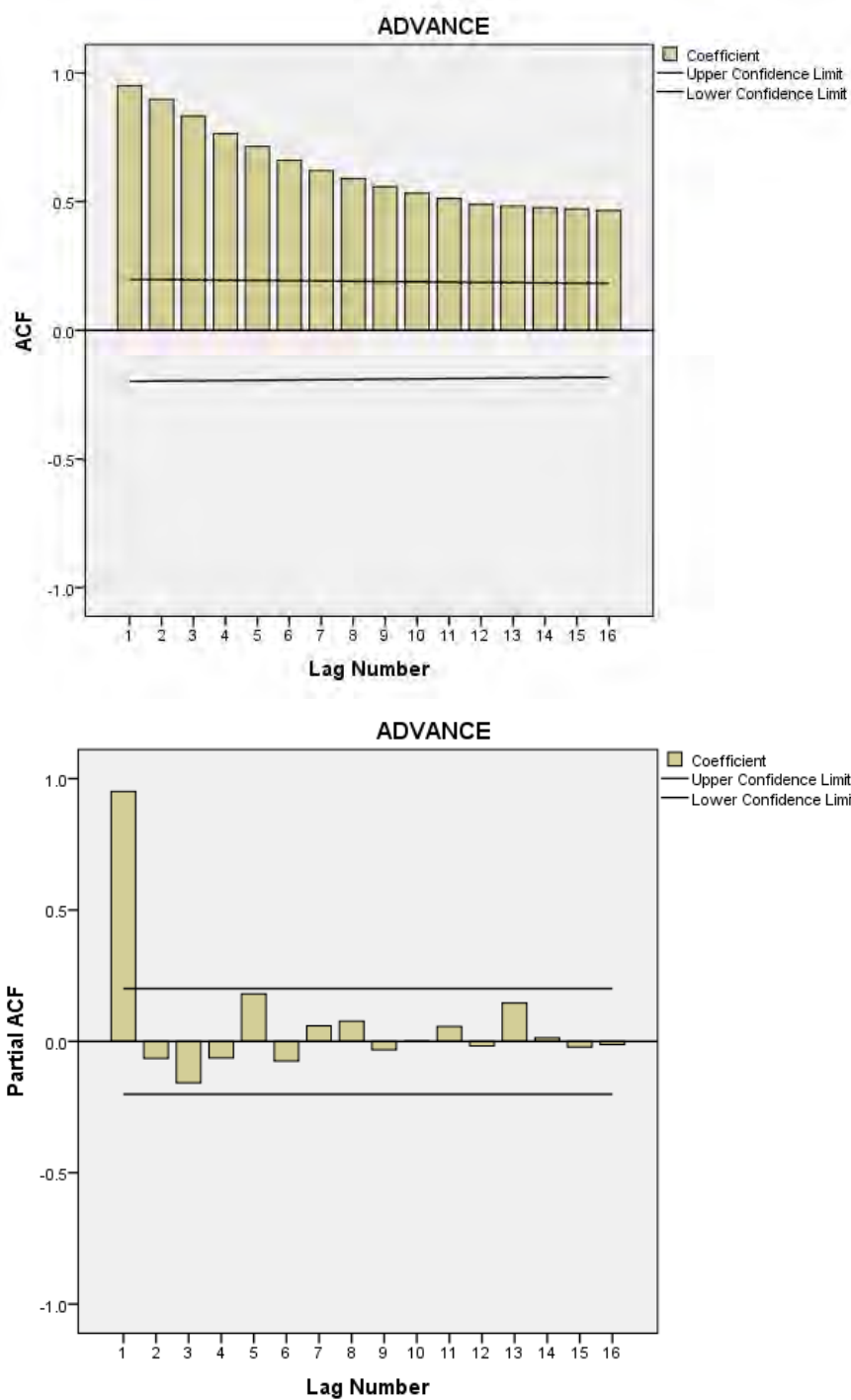
ภาพที่ 8 แสดงการวัดสหสัมพันธ์ข้อมูลหุ่น NWR



ภาพที่ 9 แสดงการวัดสหสัมพันธ์ข้อมูลหุ้น PTT



ภาพที่ 10 แสดงการวัดสหสัมพันธ์ข้อมูลหุ้น KTB



ภาพที่ 11 แสดงการวัดสหสัมพันธ์ข้อมูลหุ้น ADVANCE

จากภาพที่ 5-9 แสดงการวิเคราะห์ลักษณะของหุ้นแต่ละตัว และแสดงให้เห็นว่าข้อมูลมีแนวโน้ม มีข้อมูลบางตัวสูง และต่ำผิดปกติจากตัวอื่นๆ แสดงว่าข้อมูลมีบางตัวผิดปกติ ภาพที่ 10-14

แสดงให้เห็นว่าข้อมูลมีความสัมพันธ์กันอย่างน้อย 1 lag แต่เลือกใช้ที่ 1 lag เพราะมีอยู่ 1 แท่งที่สูงกว่าแท่งอื่น หลังจากนั้นเมื่อวิเคราะห์ลักษณะของข้อมูลได้แล้วจึงนำลักษณะของข้อมูลไปสร้างตัวแบบและได้ตัวแบบในหัวข้อต่อไป

4.2 ผลของสร้างตัวแบบ

ผลที่ได้จากการวิเคราะห์ข้อมูลข้างต้นมาสร้างตัวแบบ มีรายละเอียดดังต่อไปนี้

$$Y_{it} \sim N\left(\gamma_i(\Delta W(t|\alpha_i, \delta_i) + A_{it}), [\gamma_i(3 + \zeta_{it})\sigma_y]^2\right)$$

เมื่อ i คือราคาปิดของหุ้นแต่ละตัว โดยที่ $i = 1, 2, \dots, 5$

t คือราคาปิดของหุ้นแต่ละตัวที่เก็บมาตามเวลา โดยที่ $t = 1, 2, \dots, n$

เมื่อค่าเฉลี่ยของ Y_{it} คือ

$$E(Y_{it}) = \gamma_i(\Delta W(t|\alpha_i, \delta_i) + A_{it}) \quad (51)$$

ค่าความแปรปรวนของ Y_{it} คือ

$$Var(Y_{it}) = [\gamma_i(3 + \zeta_{it})\sigma_y]^2 \quad (52)$$

เมื่อ γ_i คือค่าคาดหวังของราคาปิดของหุ้นแต่ละตัวของ Z_i ซึ่ง Z_i ก็คือผลรวมของข้อมูลอนุกรมเวลาทั้งหมดในช่วงเวลาที่ศึกษาของราคาปิดหรือวิเคราะห์ของหุ้นแต่ละตัว $W(t|\alpha_i, \delta_i)$ คือฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของเวย์บูล (Weibull) หรือ ฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของแบบเอ็กซ์โพเนนเชียล (exponential) A_{it} อัตราสหสัมพันธ์ที่ซ่อนเร้นอยู่ในช่วงเวลา t ของหุ้นแต่ละตัว ζ_{it} คือข้อมูลผิดปกติในช่วงเวลา t ของหุ้นแต่ละตัว และ σ_y^2 ความแปรปรวนของ Y_{it}

4.2.1 prior distribution ของตัวแบบในงานวิจัยนี้คือ

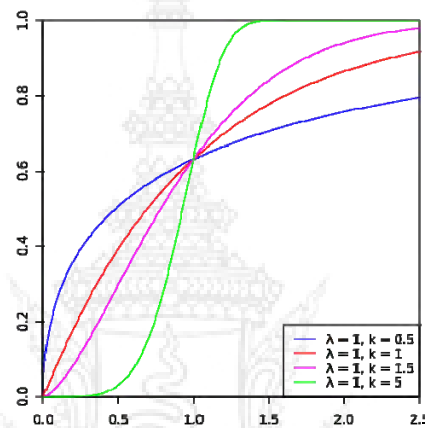
1) Prior distribution ของแนวนอน

$$\Delta W(t|\alpha_i, \delta_i) = W(t|\alpha_i, \delta_i) - W(t-1|\alpha_i, \delta_i), \quad \begin{matrix} W(t|\alpha_i, \delta_i) \square CDF(Weibul) \text{ or} \\ W(t|\alpha_i) \square CDF(exp) \end{matrix}$$

$CDF(Weibul)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของเวย์บูลคือ

$$F(y; \alpha, \delta) = \begin{cases} 1 - e^{-(y/\delta)^\alpha}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

และ แสดงด้วยภาพกราฟเมื่อกำหนดให้พารามิเตอร์มีค่าต่าง ๆ กันดังภาพที่ 15

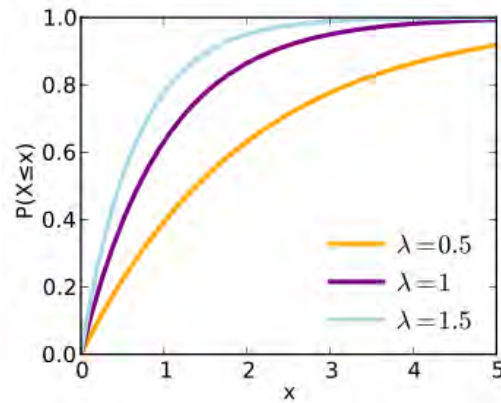


ภาพที่ 12 กราฟฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของเวย์บูลเมื่อกำหนดค่าต่างๆของพารามิเตอร์ (ค่า $k = \alpha$ และ $\lambda = \delta$ สำหรับในงานวิจัยนี้)

$CDF(exp)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของเอ็กซ์โพเนนเชียล คือ

$$F(y; \alpha) = \begin{cases} 1 - e^{-\alpha y}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

และ แสดงด้วยภาพกราฟเมื่อกำหนดให้พารามิเตอร์มีค่าต่าง ๆ กันดังภาพที่ 16



ภาพที่ 13 กราฟฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของเอ็กซ์โพเนนเชียล เมื่อกำหนดค่าต่างๆของพารามิเตอร์ (ค่า $\lambda = \alpha$ สำหรับในงานวิจัยนี้)

$$\begin{aligned}\alpha_i &\sim N_{[0,\infty)}(\mu_{\alpha_i}, \sigma_{\alpha}^2), p(\mu_{\alpha}) \sim N(0, 1.0E09), \\ p(\sigma_{\alpha}^2) &\sim \text{InvGamma}(0.1, 0.001) \\ \delta_i &\sim N_{[0,\infty)}(\mu_{\delta_i}, \sigma_{\delta}^2), p(\mu_{\delta}) \sim N(0, 1.0E09), \\ p(\sigma_{\delta}^2) &\sim \text{InvGamma}(0.1, 0.0001)\end{aligned}$$

2) Prior distribution ของอัตตสหสัมพันธ์ที่ซ่อนเร้น: AR(1):

$$\begin{aligned}A_{it} &\sim N(\lambda_{i1}A_{it-1}, \sigma_A^2), p(\sigma_A^2) \sim \text{InvGamma}(0.1, 0.0001) \\ \lambda_{i1} &\sim N(0, 1.0E09), A_{i0} = 0\end{aligned}$$

3) Prior distribution ของข้อมูลผิดปกติ:

$$\zeta_{it} \sim \text{Bern}(0.05)$$

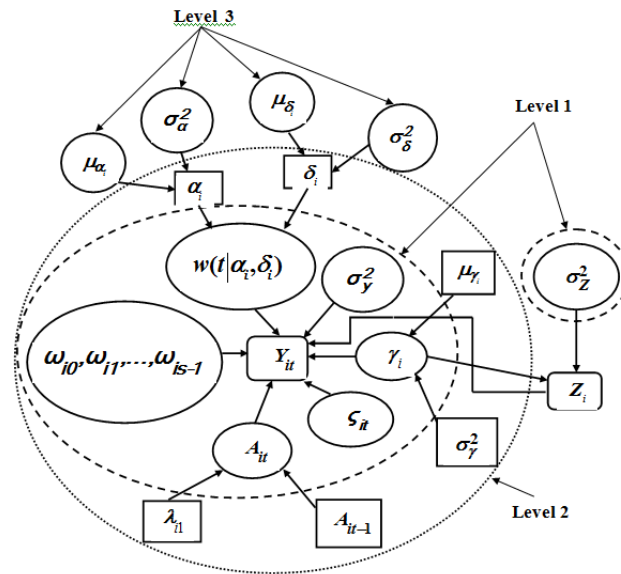
4) Prior distribution ของค่าคาดหวังของผลรวมของข้อมูลอนุกรมเวลา:

$$\begin{aligned}\gamma_i &\sim N_{[0,\infty)}(\mu_{\gamma_i}, \sigma_{\gamma}^2), p(\mu_{\gamma}) \sim N(0, 1.0E09) \\ p(\sigma_{\gamma}^2) &\sim \text{InvGamma}(0.1, 0.0001)\end{aligned}$$

5) Prior distribution ของผลรวมของข้อมูลอนุกรมเวลาทั้งหมดในช่วงเวลาที่ศึกษา

$$Z_i \sim N(\gamma_i, \sigma_Z^2), p(\sigma_Z^2) \sim \text{InvGamma}(0.1, 0.0001)$$

สำหรับโครงสร้างของตัวแบบในงานวิจัยนี้สามารถแสดงได้ในภาพที่ 17



ภาพที่ 14 โครงสร้างระดับชั้นต่างๆของพารามิเตอร์ในตัวแบบที่นำเสนอ

สำหรับสมการพยากรณ์ของตัวแบบเบย์ในงานวิจัยนี้แสดงได้ดังสมการ 53

$$p(\hat{Y}_{it+1} | Y_{it}, Y_{it-1}, \dots, Y_{i1}) = \int \dots \int p(\hat{Y}_{it+1} | \theta) p(\theta | Y_{it}, Y_{it-1}, \dots, Y_{i1}) d\theta \quad (53)$$

$$p(\hat{Y}_{it+1} | Y_{it}, Y_{it-1}, \dots, Y_{i1}) \propto \int \dots \int p(\hat{Y}_{it+1} | \theta) p(Y_{it}, Y_{it-1}, \dots, Y_{i1} | \theta) p(\theta) d\theta$$

จากสมการที่ 46 เราสามารถประมาณวิธีการหาคำตอบได้โดยใช้การสุ่มตัวอย่างแบบ Gibbs

ตามสมการเพื่อประมาณค่า $\hat{Y}_{it+1}$

4.2.2 การประมาณค่าพารามิเตอร์โดยวิธี MCMC

จากตัวแบบข้างต้น และกำหนด priors ให้กับพารามิเตอร์ทั้งหมดแล้ว เราจะใช้วิธี Markov Chain Monte Carlo (MCMC) สำหรับประมาณค่าพารามิเตอร์ โดยใช้วิธีการสุ่มตัวอย่างแบบ Gibbs ซึ่งวิธีการสุ่มตัวอย่างแบบ Gibbs นี้จะเหมาะสมกับฟังก์ชันการแจกแจงแบบมีเงื่อนไขเมื่อจำนวนรอบของการสุ่มตัวอย่างมากๆ มันก็จะเข้าสู่ฟังก์ชันการแจกแจงความน่าจะเป็นร่วมของ posterior ใดๆ (joint posterior distribution) สามารถเขียน likelihood ของตัวแบบได้ดังสมการ

$$\begin{aligned}
& f(Y_{i1}, \dots, Y_{in} | \gamma_i, w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{i1}, \dots, \xi_{in}, \sigma_Y^2, \\
& \lambda_{i1}, \mu_{\alpha_i}, \sigma_\alpha^2, \mu_{\delta_i}, \sigma_\delta^2, \mu_{\gamma_i}, \sigma_\gamma^2, \sigma_Z^2) \\
& = \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{it}, \sigma_Y^2, \lambda_i, \mu_{\alpha_i}, \sigma_\alpha^2, \\
& \mu_{\delta_i}, \sigma_\delta^2, \mu_{\gamma_i}, \sigma_\gamma^2, \sigma_Z^2)
\end{aligned} \tag{54}$$

สามารถเขียนผลคูณ prior distributions ของพารามิเตอร์ทั้งหมดได้ดังนี้

$$\begin{aligned}
& [p(\gamma_i | \mu_{\gamma_i}, \sigma_\gamma^2) p(\mu_{\gamma_i}) p(\sigma_\gamma^2) p(w(t) | \alpha_i, \delta_i) p(\alpha | \mu_{\alpha_i}, \sigma_\alpha^2) p(\mu_{\alpha_i}) p(\sigma_\alpha^2) p(\delta | \mu_{\delta_i}, \sigma_\delta^2) p(\mu_{\delta_i}) \\
& p(\sigma_\delta^2) p(A_{i1}, \dots, A_{in} | \lambda_i, \sigma_A^2) p(\lambda_i) p(\sigma_A^2), p(\omega_{i1}), \dots, p(\omega_{is-1}) p(\xi_{i1}), \dots, p(\xi_{in}) \\
& p(\sigma_Y^2)]
\end{aligned}$$

สามารถเขียน posterior distribution ซึ่งเกิดจากผลคูณของ likelihood กับ ผลคูณ prior distributions ของพารามิเตอร์ทั้งหมด ได้ดังนี้

$$\begin{aligned}
& p(\gamma_i, w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{i1}, \dots, \xi_{in}, \sigma_Y^2, \lambda_i, \mu_{\alpha_i}, \\
& \sigma_\alpha^2, \mu_{\delta_i}, \sigma_\delta^2, \mu_{\gamma_i}, \sigma_\gamma^2, \sigma_Z^2 | Y_{i1}, \dots, Y_{in}) \\
& = \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{it}, \sigma_Y^2, \lambda_i, \mu_{\alpha_i}, \\
& \sigma_\alpha^2, \mu_{\delta_i}, \sigma_\delta^2, \mu_{\gamma_i}, \sigma_\gamma^2, \sigma_Z^2) [p(\gamma_i | \mu_{\gamma_i}, \sigma_\gamma^2) p(\mu_{\gamma_i}) p(\sigma_\gamma^2) p(w(t) | \alpha_i, \delta_i) \\
& p(\alpha | \mu_{\alpha_i}, \sigma_\alpha^2) p(\mu_{\alpha_i}) p(\sigma_\alpha^2) p(\delta | \mu_{\delta_i}, \sigma_\delta^2) p(\mu_{\delta_i}) p(\sigma_\delta^2) p(A_{i1}, \dots, A_{in} | \lambda_i, \sigma_A^2) \\
& p(\lambda_i) p(\sigma_A^2) p(\omega_{i1}), \dots, p(\omega_{is-1}) p(\xi_{i1}), \dots, p(\xi_{in}) p(\sigma_Y^2)]
\end{aligned} \tag{55}$$

สำหรับวิธีการประมาณค่าพารามิเตอร์ด้วยวิธี MCMC โดยอัลกอริทึมของการสุ่มตัวอย่างแบบ Gibbs จะทำการสร้าง The full conditional distributions ให้กับพารามิเตอร์แต่ละตัว ซึ่ง The full conditional distribution ของพารามิเตอร์แต่ละตัว เกิดจาก ผลคูณของ the likelihood กับ all priors ที่เกี่ยวข้องกับพารามิเตอร์ของตัวเอง ตัวอย่างเช่น the full conditional distributions ของพารามิเตอร์หลักๆ $(\gamma_i, \alpha_i, \delta_i, A_{it}, \xi_{it}, \mu_{\alpha_i})$ แสดงดังต่อไปนี้ ส่วนตัวอื่นๆจะแสดงไว้ที่ภาคผนวก

The full conditional distribution ของ γ_i คือ

$$\begin{aligned}
& p(\gamma_i | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{i1}, \dots, \xi_{in}, \sigma_Y^2, \\
& \quad \lambda_i, \mu_{\alpha_i}, \sigma_{\alpha_i}^2, \mu_{\delta_i}, \sigma_{\delta_i}^2, \mu_{\gamma_i}, \sigma_{\gamma_i}^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\
& = \prod_{t=1}^n f(Y_{it} | \gamma_t, w(t), \alpha_t, \delta_t, A_{it}, \omega_{t1}, \dots, \omega_{ts-1}, \xi_{it}, \sigma_Y^2, \\
& \quad \lambda_t, \mu_{\alpha_t}, \sigma_{\alpha_t}^2, \mu_{\delta_t}, \sigma_{\delta_t}^2, \mu_{\gamma_t}, \sigma_{\gamma_t}^2, \sigma_Z^2) \\
& \quad [p(\gamma_i | \mu_{\gamma_i}, \sigma_{\gamma_i}^2) p(\mu_{\gamma_i}) p(\sigma_{\gamma_i}^2)]
\end{aligned} \tag{56}$$

The full conditional distribution for α_i is

$$\begin{aligned}
& p(\alpha_i | w(t), \gamma_i, \delta_i, A_{i1}, \dots, A_{in}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{i1}, \dots, \xi_{in}, \sigma_Y^2, \\
& \quad \lambda_i, \mu_{\alpha_i}, \sigma_{\alpha_i}^2, \mu_{\delta_i}, \sigma_{\delta_i}^2, \mu_{\gamma_i}, \sigma_{\gamma_i}^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\
& = \prod_{t=1}^n f(Y_{it} | \gamma_t, w(t), \alpha_t, \delta_t, A_{it}, \omega_{t1}, \dots, \omega_{ts-1}, \xi_{it}, \sigma_Y^2, \\
& \quad \lambda_t, \mu_{\alpha_t}, \sigma_{\alpha_t}^2, \mu_{\delta_t}, \sigma_{\delta_t}^2, \mu_{\gamma_t}, \sigma_{\gamma_t}^2, \sigma_Z^2) \\
& \quad [p(\alpha_i | \mu_{\alpha_i}, \sigma_{\alpha_i}^2) p(\mu_{\alpha_i}) p(\sigma_{\alpha_i}^2)]
\end{aligned} \tag{57}$$

The full conditional distribution ของ δ_i คือ

$$\begin{aligned}
& p(\delta_i | w(t), \gamma_i, \alpha_i, A_{i1}, \dots, A_{in}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{i1}, \dots, \xi_{in}, \sigma_Y^2, \\
& \quad \lambda_i, \mu_{\alpha_i}, \sigma_{\alpha_i}^2, \mu_{\delta_i}, \sigma_{\delta_i}^2, \mu_{\gamma_i}, \sigma_{\gamma_i}^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\
& = \prod_{t=1}^n f(Y_{it} | \gamma_t, w(t), \alpha_t, \delta_t, A_{it}, \omega_{t1}, \dots, \omega_{ts-1}, \xi_{it}, \sigma_Y^2, \\
& \quad \lambda_t, \mu_{\alpha_t}, \sigma_{\alpha_t}^2, \mu_{\delta_t}, \sigma_{\delta_t}^2, \mu_{\gamma_t}, \sigma_{\gamma_t}^2, \sigma_Z^2) [p(\delta_i | \mu_{\delta_i}, \sigma_{\delta_i}^2) p(\mu_{\delta_i}) p(\sigma_{\delta_i}^2)]
\end{aligned} \tag{58}$$

The full conditional distribution ของ A_{it} คือ

$$\begin{aligned}
& p(A_{it} | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{i1}, \dots, \xi_{in}, \sigma_Y^2, \\
& \quad \lambda_i, \mu_{\alpha_i}, \sigma_{\alpha_i}^2, \mu_{\delta_i}, \sigma_{\delta_i}^2, \mu_{\gamma_i}, \sigma_{\gamma_i}^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\
& = \prod_{t=1}^n f(Y_{it} | \gamma_t, w(t), \alpha_t, \delta_t, A_{it}, \omega_{t1}, \dots, \omega_{ts-1}, \xi_{it}, \sigma_Y^2, \lambda_t, \mu_{\alpha_t}, \\
& \quad \sigma_{\alpha_t}^2, \mu_{\delta_t}, \sigma_{\delta_t}^2, \mu_{\gamma_t}, \sigma_{\gamma_t}^2, \sigma_Z^2) [p(A_{it} | \lambda A_{it-1}, \sigma_A^2) p(\lambda) p(A_{it} | \lambda A_{it-1}, \\
& \quad \sigma_A^2) p(\sigma_A^2)]
\end{aligned} \tag{59}$$

เมื่อ $A_{i-t} = A_{i1}, \dots, A_{it-1}, A_{it+1}, \dots, A_{in}$.

The full conditional distribution ของ ξ_{it} คือ

$$\begin{aligned}
& p(\xi_{it} | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{i-t}, \sigma_Y^2, \\
& \quad \lambda_i, \mu_{\alpha_i}, \sigma_\alpha^2, \mu_{\delta_i}, \sigma_\delta^2, \mu_{\gamma_i}, \sigma_\gamma^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\
& = \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{i-t}, \sigma_Y^2, \\
& \quad \lambda_i, \mu_{\alpha_i}, \sigma_\alpha^2, \mu_{\delta_i}, \sigma_\delta^2, \mu_{\gamma_i}, \sigma_\gamma^2, \sigma_Z^2) [p(\xi_{it})]
\end{aligned} \tag{60}$$

เมื่อ $\xi_{i-t} = \xi_{i1}, \dots, \xi_{it-1}, \xi_{it+1}, \dots, \xi_{in}$.

The full conditional distribution ของ μ_{α_i} คือ

$$\begin{aligned}
& p(\mu_{\alpha_i} | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{i1}, \dots, \xi_{in}, \sigma_Y^2, \\
& \quad \lambda_i, \mu_{\alpha_i}, \sigma_\alpha^2, \mu_{\delta_i}, \sigma_\delta^2, \mu_{\gamma_i}, \sigma_\gamma^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\
& = \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \omega_{i1}, \dots, \omega_{is-1}, \xi_{it}, \sigma_Y^2, \\
& \quad \lambda_i, \mu_{\alpha_i}, \sigma_\alpha^2, \mu_{\delta_i}, \sigma_\delta^2, \mu_{\gamma_i}, \sigma_\gamma^2, \sigma_Z^2) [p(\mu_{\alpha_i})]
\end{aligned} \tag{61}$$

4.3 ผลของการประเมินประสิทธิภาพของตัวแบบ

การประมาณค่าพารามิเตอร์ด้วยวิธี MCMC โดยใช้อัลกอริทึมของการสุ่มตัวอย่างแบบ Gibbs โดยทำการเขียนอัลกอริทึมในโปรแกรม OpenBUGS เมื่อได้ค่าพารามิเตอร์ที่ประมาณได้จากข้อมูลจริงมาแล้ว หลังจากนั้นจะใช้ค่าพารามิเตอร์ทุกตัวมาทำการจำลองสถานการณ์สร้างชุดข้อมูลมาใหม่อีก 1000 ชุดซึ่งจะถูกเขียนอัลกอริทึมในโปรแกรม R และประเมินประสิทธิภาพของตัวแบบโดยการประเมินจากพารามิเตอร์ที่สำคัญของแต่ละตัวของฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของแนวโน้มนัยแบบเอ็กซ์โพเนนเชียลได้ผลดังตารางที่ 1

ตารางที่ 9 ประสิทธิภาพของตัวแบบโดยการประเมินจากพารามิเตอร์ของหุ่น EA

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.021	0.411	0.941	μ_γ	0.012	0.216	0.975
σ_Z^2	0.011	0.521	0.952	σ_γ^2	0.011	0.319	0.971
γ	0.004	0.124	0.957	μ_α	0.011	0.133	0.964
λ	0.031	0.874	0.959	σ_α^2	0.016	0.364	0.964
α	0.009	0.885	0.960	μ_δ	0.015	0.522	0.971
δ	0.008	0.755	0.952	σ_δ^2	0.011	0.494	0.977

ตารางที่ 10 ประสิทธิภาพของตัวแบบโดยการประเมินจากพารามิเตอร์ของหุ้น NRW

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.033	0.416	0.436	μ_γ	0.008	0.416	0.972
σ_z^2	0.031	0.533	0.789	σ_γ^2	0.006	0.429	0.971
γ	0.012	0.143	0.765	μ_α	0.021	0.326	0.968
λ	0.011	0.765	0.944	σ_α^2	0.019	0.344	0.969
α	0.008	0.908	0.958	μ_δ	0.015	0.617	0.971
δ	0.017	0.675	0.678	σ_δ^2	0.014	0.481	0.970

ตารางที่ 11 ประสิทธิภาพของตัวแบบโดยการประเมินจากพารามิเตอร์ของหุ้น PTT

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.011	0.446	0.949	μ_γ	0.012	0.216	0.971
σ_z^2	0.013	0.549	0.960	σ_γ^2	0.001	0.329	0.973
γ	0.009	0.156	0.962	μ_α	0.011	0.126	0.965
λ	0.011	0.824	0.957	σ_α^2	0.017	0.363	0.969
α	0.019	0.817	0.956	μ_δ	0.015	0.597	0.979
δ	0.018	0.793	0.954	σ_δ^2	0.015	0.483	0.978

ตารางที่ 12 ประสิทธิภาพของตัวแบบโดยการประเมินจากพารามิเตอร์ของหุ้น KTB

P	RB	MSE	CP	P	RB	MSE	CP
σ_Y^2	0.015	0.339	0.941	μ_γ	0.011	0.216	0.871
σ_z^2	0.003	0.668	0.960	σ_γ^2	0.011	0.329	0.977
γ	0.009	0.097	0.962	μ_α	0.011	0.126	0.965
λ	0.021	0.657	0.968	σ_α^2	0.012	0.164	0.889
α	0.017	0.863	0.956	μ_δ	0.014	0.597	0.979
δ	0.010	0.733	0.954	σ_δ^2	0.011	0.483	0.758

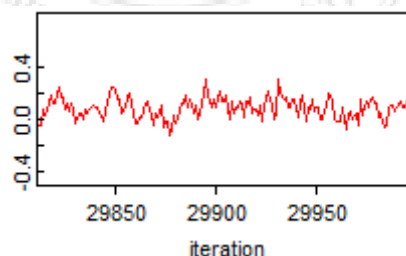
ตารางที่ 13 ประสิทธิภาพของตัวแบบโดยการประเมินจากพารามิเตอร์ของหุ่น ADVANCE

P	RB	MSE	CP	P	RB	MSE	CP
σ_y^2	0.021	0.543	0.848	μ_γ	0.002	0.616	0.961
σ_z^2	0.013	0.546	0.860	σ_γ^2	0.001	0.821	0.963
γ	0.009	0.556	0.762	μ_α	0.011	0.426	0.961
λ	0.011	0.524	0.867	σ_α^2	0.011	0.464	0.965
α	0.019	0.517	0.756	μ_δ	0.019	0.588	0.978
δ	0.011	0.592	0.858	σ_δ^2	0.012	0.583	0.978

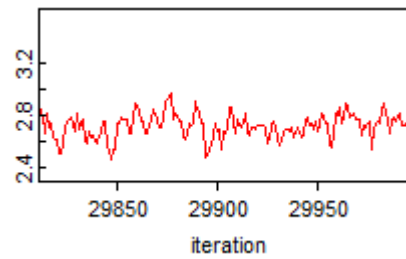
จากตารางที่ 9-13 พบว่าค่า RB และค่า MSE ของพารามิเตอร์แต่ละตัวของฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของแบบเอ็กซ์โพเนนเชียลมีค่าต่ำมาก และพบว่าค่า CP ของพารามิเตอร์แต่ละตัวมีค่าสูงมาก ส่วนพารามิเตอร์แต่ละตัวของตัวแบบที่มีฟังก์ชันการแจกแจงความน่าจะเป็นสะสมของแบบเวย์บูลก็ให้ผลใกล้เคียงกันกับแบบเอ็กซ์โพเนนเชียลซึ่งได้แสดงไว้ในงานวิจัยครั้งที่แล้วสำหรับในงานวิจัยนี้ได้แสดงผลการทดลองเฉพาะเอ็กซ์โพเนนเชียลเท่านั้น ดังนั้นจึงสรุปได้ว่าประสิทธิภาพของตัวแบบทั้งสองตัวในงานวิจัยนี้อยู่ในเกณฑ์ที่ดีมาก

4.4 ผลของการประมาณพารามิเตอร์

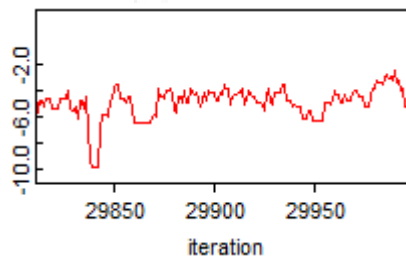
เมื่อได้ตัวแบบที่เหมาะสมแล้วจึงเลือกใช้ค่าพารามิเตอร์ที่ประมาณได้จากข้อมูลจริงจำนวน 99 วัน ตัวอย่างผลของ trace plot ที่แสดงการการลู่เข้าสู่การแจกแจงใดการแจกแจงหนึ่ง (stationary distribution) ของพารามิเตอร์บางตัว (γ_i) ของหุ่นต่างๆของทั้งสองตัวแบบแสดงได้ดังภาพที่ 14-18 ส่วนพารามิเตอร์ตัวอื่นที่ไม่ได้นำมาแสดงก็จะมีผลการลู่เข้าสู่การแจกแจงใดการแจกแจงหนึ่งเช่นกัน



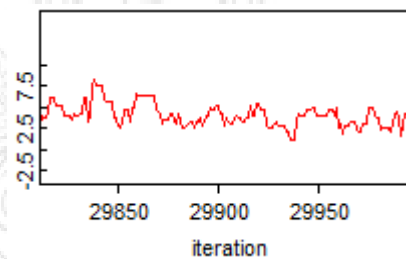
ภาพที่ 15 Trace ของ γ_1 ของหุ่น EA ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเวย์บูล



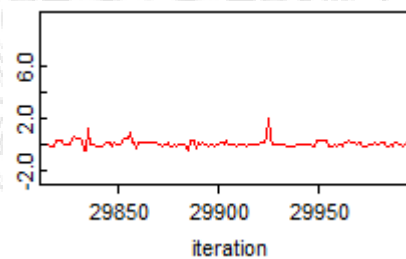
ภาพที่ 16 Trace ของ γ_2 ของหุ่น NWR ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเวทย์บูล



ภาพที่ 17 Trace ของ γ_1 ของหุ่น EA ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล



ภาพที่ 18 Trace ของ γ_4 ของหุ่น KTB ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล



ภาพที่ 19 Trace ของ γ_5 ของหุ่น ADVANCE ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล

จากภาพ 14-18 แสดงการลู่เข้าของ MCMC โดยพิจารณาจากกราฟของ Trace plot พบว่า MCMC ลู่เข้าสู่การแจกแจงใดการแจกแจงหนึ่งเมื่อมีการสุ่มตัวอย่าง แบบกิบส์ จำนวน 30000 รอบ

โดยตัด 10000 รอบแรกทิ้ง Trace plot ในภาพที่ 14-18 ไม่แสดงว่ามีแนวโน้มเกิดขึ้น แสดงถึงการลู่เข้าของ MCMC สำหรับพารามิเตอร์ $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5$ ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล และฟังก์ชันแนวโน้มแบบเวย์บูล ส่วนภาพ Trace plot ตัวอื่นๆของทั้งสองตัวแบบที่ไม่ได้นำมาแสดงพบว่า MCMC ลู่เข้าสู่การแจกแจงใดการแจกแจงหนึ่งเช่นเดียวกัน ส่วนผลการประมาณค่าพารามิเตอร์แสดงได้ดังตารางที่ 14-23

ตารางที่ 14 ค่าพารามิเตอร์ของหุ่น EA ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเวย์บูล

Parameter	Value	Parameter	Value
α	1232.50	μ_γ	933.50
λ	20.78	σ_Y^2	-0.40
δ	273.91	σ_α^2	-206.71
γ	556.35	σ_δ^2	-473.65
μ_α	87.92	σ_γ^2	-26.96
μ_δ	4488.63	σ_z^2	4311.37

ตารางที่ 15 ค่าพารามิเตอร์ของหุ่น NWR ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเวย์บูล

Parameter	Value	Parameter	Value
α	179.50	μ_γ	1113.00
λ	40.59	σ_Y^2	40.19
δ	270.31	σ_α^2	63.60
γ	545.00	σ_δ^2	71.35
μ_α	87.44	σ_γ^2	60.48
μ_δ	118.63	σ_z^2	4430.00

ตารางที่ 16 ค่าพารามิเตอร์ของหุ้น PTT ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเวกซ์บูล

Parameter	Value	Parameter	Value
α	1262.50	μ_γ	316.94
λ	50.78	σ_Y^2	22.45
δ	303.91	σ_α^2	509.46
γ	586.35	σ_δ^2	1091.80
μ_α	117.92	σ_γ^2	121.77
μ_δ	4518.63	σ_z^2	187.90

ตารางที่ 17 ค่าพารามิเตอร์ของหุ้น KTB ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเวกซ์บูล

Parameter	Value	Parameter	Value
α	149.50	μ_γ	1083.00
λ	10.59	σ_Y^2	10.19
δ	240.31	σ_α^2	33.60
γ	515.00	σ_δ^2	41.35
μ_α	57.44	σ_γ^2	30.48
μ_δ	88.63	σ_z^2	4400.00

ตารางที่ 18 ค่าพารามิเตอร์ของหุ้น ADVANCE ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเวกซ์บูล

Parameter	Value	Parameter	Value
α	346.84	μ_γ	3483.69
λ	24.57	σ_Y^2	125.79
δ	557.52	σ_α^2	199.07
γ	1194.80	σ_δ^2	223.33
μ_α	133.26	σ_γ^2	189.30
μ_δ	205.62	σ_z^2	13865.90

ตารางที่ 19 ค่าพารามิเตอร์ของหุ่น EA ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล

Parameter	Value	Parameter	Value
α	208.68	μ_γ	1023.82
λ	69.77	σ_Y^2	-48.99
δ	299.49	σ_α^2	-25.58
γ	574.18	σ_δ^2	-17.83
μ_α	116.62	σ_γ^2	-28.70
μ_δ	147.81	σ_z^2	4340.82

ตารางที่ 20 ค่าพารามิเตอร์ของหุ่น NWR ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล

Parameter	Value	Parameter	Value
α	447.87	μ_γ	784.63
λ	108.96	σ_Y^2	-88.18
δ	338.68	σ_α^2	-64.77
γ	613.37	σ_δ^2	-57.02
μ_α	155.81	σ_γ^2	-67.89
μ_δ	187.43	σ_z^2	4301.63

ตารางที่ 21 ค่าพารามิเตอร์ของหุ่น PTT ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล

Parameter	Value	Parameter	Value
α	184.63	μ_γ	1047.87
λ	-178.18	σ_Y^2	198.96
δ	-154.77	σ_α^2	428.68
γ	-147.02	σ_δ^2	703.37
μ_α	-157.89	σ_γ^2	245.81
μ_δ	4211.63	σ_z^2	277.00

ตารางที่ 22 ค่าพารามิเตอร์ของหุ่น KTB ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล

Parameter	Value	Parameter	Value
α	1481.99	μ_γ	548.49
λ	139.18	σ_Y^2	139.58
δ	162.59	σ_α^2	369.30
γ	170.34	σ_δ^2	643.99
μ_α	159.47	σ_γ^2	186.43
μ_δ	4528.99	σ_z^2	217.62

ตารางที่ 23 ค่าพารามิเตอร์ของหุ่น ADVANCE ของตัวแบบที่มีฟังก์ชันแนวโน้มแบบเอ็กซ์โพเนนเชียล

Parameter	Value	Parameter	Value
α	684.01	μ_γ	398.99
λ	-118.80	σ_Y^2	150.17
δ	-95.39	σ_α^2	609.61
γ	-87.64	σ_δ^2	1158.99
μ_α	-98.51	σ_γ^2	243.87
μ_δ	4271.01	σ_z^2	306.25

จากตารางที่ 14-23 แสดงถึงตัวอย่างของค่าพารามิเตอร์ที่สำคัญบางตัวของทั้งสองตัวแบบที่เข้าสู่การแจกแจงโคแจกแจงหนึ่ง แต่ค่าพารามิเตอร์ตัวอื่นๆที่ไม่ได้แสดงในงานวิจัยนี้ได้เข้าสู่การแจกแจงโคแจกแจงหนึ่งเหมือนกันในภาพที่ 14-18 ที่ประมาณ 30000 รอบขึ้นไป เช่นกัน ซึ่งสรุปได้ว่าค่าพารามิเตอร์แต่ละตัวจะมีค่าที่ไม่เปลี่ยนแปลงแล้วจึงยอมรับผลของการประมาณค่าพารามิเตอร์ที่ได้ในตารางที่ 14-23 และตัวอื่นๆที่ไม่ได้นำมาแสดงเพื่อนำไปใช้พยากรณ์ค่า y_{it} ของหุ่นแต่ละตัวในหัวข้อต่อไป

4.5 ผลของการเปรียบเทียบตัวแบบ

เมื่อได้ตัวแบบที่เหมาะสมแล้วจึงเลือกใช้ค่าพารามิเตอร์ที่ประมาณได้จากข้อมูลจริงจำนวน 99 วัน มาเปรียบเทียบกับวิธีการพยากรณ์ทั้งสองวิธี คือตัวแบบเบย์ที่มีฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวย์บูล กับตัวแบบเบย์ที่มีฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โพเนนเชียลเพื่อหาตัวแบบที่เหมาะสม (Fitting Model)

และเปรียบเทียบเพื่อตรวจสอบความถูกต้อง ของตัวแบบ (Validation Model) จากข้อมูลที่เหลือ อีก 15 วัน แสดงได้ดังตารางที่ 24-33

ตารางที่ 24 การเปรียบเทียบหาตัวแบบที่เหมาะสม (Fitting Model) ของหุ้น EA

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
99 วัน	1. วิธีเบย์แบบเวย์บูล	1.311	5.644	1.209
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	1.011	5.244	1.119

ตารางที่ 25 การเปรียบเทียบหาตัวแบบที่เหมาะสม (Fitting Model) ของหุ้น NWR

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
99 วัน	1. วิธีเบย์แบบเวย์บูล	1.996	6.132	1.032
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	1.592	5.931	1.735

ตารางที่ 26 การเปรียบเทียบหาตัวแบบที่เหมาะสม (Fitting Model) ของหุ้น PTT

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
99 วัน	1. วิธีเบย์แบบเวย์บูล	5.996	10.932	5.932
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	6.096	11.432	6.232

ตารางที่ 27 การเปรียบเทียบหาตัวแบบที่เหมาะสม (Fitting Model) ของหุ้น KTB

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
99 วัน	1. วิธีเบย์แบบเวย์บูล	1.822	5.846	1.817
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	1.011	5.244	1.119

ตารางที่ 28 การเปรียบเทียบหาตัวแบบที่เหมาะสม (Fitting Model) ของหุ่น ADVANCE

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
99 วัน	1. วิธีเบย์แบบเวย์บูล	8.711	14.186	5.686
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	7.796	13.132	4.932

ตารางที่ 29 การเปรียบเทียบตรวจสอบความถูกต้องของตัวแบบ (Validation Model) ของหุ่น EA

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
15 วัน	1. วิธีเบย์แบบเวย์บูล	14.17	26.19	10.94
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	13.55	25.57	10.32

ตารางที่ 30 การเปรียบเทียบตรวจสอบความถูกต้องของตัวแบบ (Validation Model) ของหุ่น NWR

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
15 วัน	1. วิธีเบย์แบบเวย์บูล	19.95	31.97	16.72
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	18.18	30.2	14.95

ตารางที่ 31 การเปรียบเทียบตรวจสอบความถูกต้องของตัวแบบ (Validation Model) ของหุ่น PTT

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
15 วัน	1. วิธีเบย์แบบเวย์บูล	14.34	26.36	10.98
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	14.50	26.52	11.27

ตารางที่ 32 การเปรียบเทียบเพื่อตรวจสอบความถูกต้องของตัวแบบ (Validation Model) ของหุ้น KTB

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
15 วัน	1. วิธีเบย์แบบเวย์บูล	17.376	12.712	14.512
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	15.606	10.942	12.742

ตารางที่ 33 การเปรียบเทียบเพื่อตรวจสอบความถูกต้องของตัวแบบ (Validation Model) ของหุ้น ADVANCE

ข้อมูล	วิธีการพยากรณ์	การวัดค่าความผิดพลาด		
		RMSE	MAPE	MAE
15 วัน	1. วิธีเบย์แบบเวย์บูล	15.946	11.282	13.082
	2. วิธีเบย์แบบเอ็กซ์โปเนนเชียล	15.066	10.402	12.202

จากตารางที่ 24-28 ได้จากการที่นำข้อมูลจริงจำนวน 99 วัน มาประมาณค่าพารามิเตอร์เพื่อการเปรียบเทียบตัวแบบที่เหมาะสม (Fitting Model) จากค่าพยากรณ์ของทั้งสองวิธี กับข้อมูลจริงพบว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นเหมาะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลให้ค่า RMSE MAPE และ MAE ต่ำสุดของหุ้น 4 ตัว ได้แก่ หุ้น EA NWR KTB และ ADVANCE สำหรับการหาตัวแบบที่เหมาะสม โดยหุ้น EA มีค่า RMSE MAPE และ MAE เท่ากับ 1.011 5.244 และ 1.119 ตามลำดับ หุ้น NWR มีค่า RMSE MAPE และ MAE เท่ากับ 1.592 5.931 และ 1.735 ตามลำดับ หุ้น KTB มีค่า RMSE MAPE และ MAE เท่ากับ 1.011 5.244 และ 1.119 ตามลำดับ หุ้น ADVANCE มีค่า RMSE MAPE และ MAE เท่ากับ 7.796 13.132 และ 4.932 ตามลำดับ สำหรับหุ้น PTT ตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นเหมาะสมสำหรับแนวโน้มแบบเวย์บูล ให้ค่า RMSE MAPE และ MAE ต่ำสุดกว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นเหมาะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลเพียงเล็กน้อย มีค่า RMSE MAPE และ MAE เท่ากับ 5.996 10.932 และ 5.932 ตามลำดับ ตารางที่ 29-33 ได้จากการที่นำข้อมูลจริงที่เหลืออีก จำนวน 15 วัน มาประมาณค่าพารามิเตอร์เพื่อเปรียบเทียบการตรวจสอบความถูกต้องของตัวแบบ (Validation Model) จากค่าพยากรณ์ของทั้งสองวิธี พบว่าตัวแบบเบย์ที่มีค่า

ฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลให้ค่า RMSE MAPE และ MAE ต่ำสุดของหุ้น 4 ตัว ได้แก่ หุ้น EA NWR KTB และ ADVANCE สำหรับการหาตัวแบบที่เหมาะสม โดยหุ้น EA มีค่า RMSE MAPE และ MAE เท่ากับ 13.55 25.57 และ 10.32 ตามลำดับ หุ้น NWR มีค่า RMSE MAPE และ MAE เท่ากับ 18.18 30.2 และ 14.95 ตามลำดับ หุ้น KTB มีค่า RMSE MAPE และ MAE เท่ากับ 15.606 10.942 และ 12.742 ตามลำดับ หุ้น ADVANCE มีค่า RMSE MAPE และ MAE เท่ากับ 15.066 10.402 และ 12.202 ตามลำดับ สำหรับหุ้น PTT ตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวย์บูล ให้ค่า RMSE MAPE และ MAE ต่ำสุดกว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลเพียงเล็กน้อย มีค่า RMSE MAPE และ MAE เท่ากับ 14.34 26.36 และ 10.98 ตามลำดับเช่นกัน

บทที่ 5

สรุปผลการวิจัย และข้อเสนอแนะ

การศึกษการเลือกฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มเพื่อสร้างตัวแบบ สำหรับวิเคราะห์ข้อมูลอนุกรมเวลาหลายๆตัวพร้อมกันที่มีค่า อัตราสหสัมพันธ์ แนวโน้ม และค่าผิดปกติ ด้วยวิธีการของเบย์:กรณีศึกษาราคาหุ้นที่มีความผันผวนของตลาดหลักทรัพย์แห่งประเทศไทย สามารถสรุปผลการวิจัย และข้อเสนอแนะได้ดังนี้

สรุปผลการวิจัย

5.1 ลักษณะทั่วไปของประชากร และกลุ่มตัวอย่าง

ประชากร และกลุ่มตัวอย่างคือ ข้อมูลราคาปิดของหุ้นรายวัน จำนวน 6 เดือน ตั้งแต่ วันที่ 1 เมษายน 2557 ถึง 19 กันยายน 2557 ที่มีความผันผวนซึ่งวิเคราะห์จากบริษัทหลักทรัพย์ เอเชียพลัสโดยพิจารณาจากหุ้นที่มีค่า Beta (ค่าความผันผวนของราคาหุ้นเทียบกับตลาดหุ้น) สูงกว่า 1 และค่าพีอี ที่สูงกว่า 20 เท่า จำนวน 2 ตัว ได้แก่ บมจ.พลังงานบริสุทธิ์ (EA) บมจ.เนาวรัตน์ พัฒนาการ (NWR) และหุ้นที่เป็นที่รู้จักกันทั่วไปอีก 3 ตัว ได้แก่ บมจ. ปตท (PTT) บมจ. ธนาคารกรุงไทย (KTB) บมจ.แอดวานซ์ อินโฟร์ เซอร์วิส(ADVANCE) จากตลาดหลักทรัพย์แห่งประเทศไทยโดยจะใช้ข้อมูลตั้งแต่วันที่ 1 เมษายน ถึง 29 สิงหาคม 2557 (99 วัน) สำหรับหาตัวแบบที่เหมาะสม (Fitting Model) และใช้ข้อมูลตั้งแต่วันที่ 1 กันยายน ถึง 19 กันยายน 2557 (15 วัน) สำหรับเปรียบเทียบเพื่อตรวจสอบความถูกต้อง ของตัวแบบ (Validation Model)

5.2 การสร้างตัวแบบการพยากรณ์โดยวิธีของเบย์

การสร้างตัวแบบในงานวิจัยนี้เริ่มจากตรวจสอบลักษณะของข้อมูลจากการพล็อตกราฟ ดังภาพที่ 5-9 และการวัดสหสัมพันธ์ระหว่างข้อมูล ด้วยกราฟ Autocorrelation Function (ACF) และ Partial Autocorrelation Function (PACF) ภาพที่ 10-14 เพื่อทราบส่วนประกอบต่างๆของข้อมูลอนุกรมเวลาจากประชากรตัวอย่าง กลุ่มตัวอย่างคือ ข้อมูลราคาปิดของหุ้นรายวันจำนวน 5 ตัว ซึ่งก็มีในส่วนของ แนวโน้ม มีค่าผิดปกติ และมีความสัมพันธ์กัน หลังจากนั้นนำส่วนประกอบต่างๆที่ไปป้อนแบบสร้างตัวแบบโดยวิธีของเบย์ ตัวแบบที่ได้คือ
$$Y_{it} \sim N\left(\gamma_i(\Delta W(t|\alpha_i, \delta_i) + A_{it}), [\gamma_i(3 + \zeta_{it})\sigma_y]^2\right)$$

โดยเขียนอัลกอริทึมในโปรแกรม OpenBUGS และประเมินประสิทธิภาพของตัวแบบต่างๆ ด้วยการเขียนโปรแกรมจำลองสถานการณ์ด้วยโปรแกรม R หลังจากนั้นนำตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มต่างๆมาเปรียบเทียบกัน ได้แก่ ตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเวย์บูล กับตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นที่เหมาะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียล โดยดูจากค่า RMSE MAPE และ MAE และสมการพยากรณ์ ได้แก่

$$p(\hat{Y}_{t+1} | Y_t, Y_{t-1}, \dots, Y_1) = \int \dots \int p(\hat{Y}_{t+1} | \theta) p(\theta | Y_t, Y_{t-1}, \dots, Y_1) d\theta$$

5.3 การประเมินประสิทธิภาพของตัวแบบ

จากประมาณค่าพารามิเตอร์ด้วยวิธี MCMC โดยใช้อัลกอริทึมของการสุ่มตัวอย่างแบบ Gibbs โดยทำการเขียนอัลกอริทึมในโปรแกรม OpenBUGS เมื่อได้ค่าพารามิเตอร์ที่ประมาณได้จากข้อมูลจริงมาแล้ว หลังจากนั้นจะใช้ค่าพารามิเตอร์ทุกตัวมาทำการจำลองสถานการณ์สร้างชุดข้อมูลมาใหม่อีก 1000 ชุด ซึ่งจะถูกเขียนอัลกอริทึมในโปรแกรม R และประเมินประสิทธิภาพของตัวแบบโดยดูจากค่า RB ค่า MSE และค่า CP ของพารามิเตอร์แต่ละตัว ซึ่งค่า RB และค่า MSE มีค่าต่ำมาก และค่า CP ของพารามิเตอร์แต่ละตัวมีค่าสูงมาก สรุปได้ว่าประสิทธิภาพของตัวแบบในงานวิจัยนี้อยู่ในเกณฑ์ที่ดีมาก

5.4 การประมาณค่าพารามิเตอร์

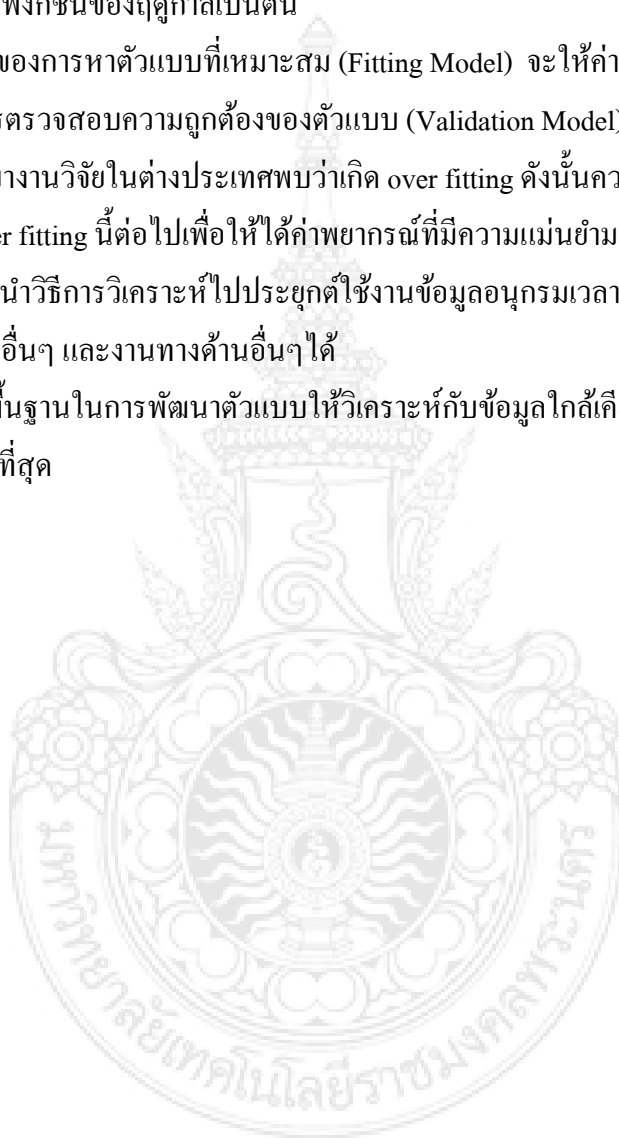
สำหรับการประมาณค่าพารามิเตอร์จะพิจารณาจาก กราฟ trace plot ของพารามิเตอร์แต่ละตัวว่าเข้าสู่การแจกแจงใดแจกแจงหนึ่ง (stationary distribution) หรือยัง สำหรับค่าพารามิเตอร์ ทุกตัว ในงานวิจัยนี้เข้าสู่การแจกแจงใดแจกแจงหนึ่ง ที่ประมาณ 30000 รอบ และเบิร์นอินข้อมูลช่วงแรกทิ้งไปที่ 1000 รอบ โดยค่าพารามิเตอร์ที่ประมาณทั้งสองตัวแบบได้แสดงดังตารางที่ 14-23

5.5 การเปรียบเทียบตัวแบบ

ตัวแบบที่นำมาเปรียบเทียบในงานวิจัยนี้ได้แก่ ตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวย์บูล กับตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียล โดยเปรียบเทียบจากค่า RMSE MAPE และ MAE ซึ่งทำการเปรียบเทียบสองส่วนคือ ส่วนที่ 1 นำข้อมูลจริง จำนวน 99 วัน มาประมาณค่าพารามิเตอร์เพื่อการเปรียบเทียบหาตัวแบบที่เหมาะสม (Fitting Model) จากค่าพยากรณ์ของทั้งสองตัวแบบ ส่วนที่ 2 นำข้อมูลจริง จำนวน 15 วันที่เหลือ มาประมาณค่าพารามิเตอร์เพื่อเปรียบเทียบตรวจสอบความถูกต้องของตัวแบบ (Validation Model) จากค่าพยากรณ์ของทั้งสองตัวแบบเช่นกัน ผลการวิจัยพบว่า การเปรียบเทียบในส่วนของการหาตัวแบบที่เหมาะสม (Fitting Model) และการตรวจสอบความถูกต้องของตัวแบบ (Validation Model) ของงานวิจัยนี้ตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลให้ค่า RMSE MAPE และ MAE ต่ำสุดของหุ้น 4 ตัว ได้แก่ หุ้น EA NWR KTB และ ADVANCE สำหรับหุ้น PTT ตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวย์บูล ให้ค่า RMSE MAPE และ MAE ต่ำสุดกว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลทั้งในส่วนของการหาตัวแบบที่เหมาะสม (Fitting Model) และการตรวจสอบความถูกต้องของตัวแบบ (Validation Model) จึงสรุปได้ว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเอ็กซ์โปเนนเชียลมีความเหมาะสมมากกว่าตัวแบบเบย์ที่มีค่าฟังก์ชันการแจกแจงความน่าจะเป็นสะสมสำหรับแนวโน้มแบบเวย์บูล สามารถนำไปใช้ในการพยากรณ์ข้อมูลอนุกรมเวลาต่างๆ ได้อย่างดี

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ในการวิจัยครั้งต่อไป วิธีการของเบย์สามารถเพิ่มส่วน หรือลดประกอบต่างๆในข้อมูลอนุกรมเวลาได้เพื่อให้เหมาะสมกับการวิเคราะห์ข้อมูลแม่นยำเพิ่มขึ้น เช่น ฟังก์ชันของค่าผิดปกติ ฟังก์ชันของฤดูกาล เป็นต้น
2. ในส่วนของการหาตัวแบบที่เหมาะสม (Fitting Model) จะให้ค่า error ต่ำมาก แต่ในส่วนของการตรวจสอบความถูกต้องของตัวแบบ (Validation Model) มีค่า error สูง ซึ่งจากการศึกษางานวิจัยในต่างประเทศพบว่าเกิด over fitting ดังนั้นควรจะมีการศึกษาในส่วน of over fitting นี้ต่อไปเพื่อให้ได้ค่าพยากรณ์ที่มีความแม่นยำมากขึ้น
3. สามารถนำวิธีการวิเคราะห์ไปประยุกต์ใช้งานข้อมูลอนุกรมเวลาอื่นๆกับอุตสาหกรรมทางด้านอื่นๆ และงานทางด้านอื่นๆได้
4. ใช้เป็นพื้นฐานในการพัฒนาตัวแบบให้วิเคราะห์กับข้อมูลใกล้เคียงกับโลกแห่งความเป็นจริงมากที่สุด



บรรณานุกรม

- ตลาดหลักทรัพย์แห่งประเทศไทย.2557. ข้อมูลการซื้อขาย. <http://www.set.or.th/th/index.html>.
เข้าถึง 19 กันยายน 2557
- Anderson, J. A.1997. **An Introduction to Neural Networks** ,MIT Press. Chapter 1-
17: 1-585
- Berger,D.E. **Introduction to Multiple Regression**.Claremont Graduate University.
- Bernd A. Berg. 2004. **Markov Chain Monte Carlo Simulations and Their Statistical
Analysis**. Singapore, World Scientific.
- Bisgaard.S., Kulahci,M.2011. **Time Series Analysis and Forecasting by Example**. John Wiley
& Sons. Inc.
- Boonsarawongse, R. and H.C. Co. 2007. Forecasting Thailand's rice export: Statistical
techniques vs artificial neural network. **Computers & industrial engineering**. 53:610-
627.
- Box, George E.P.,G.M.Jenkins and G. C. Reinsel.1994.**Time Series Analysis Forecasting and
Control**. 3rd ed.,Prentice-Hall,Inc,USA.
- Brace,N., Kemp,R., Snelgar,R.,2006. **Spss for Psychologists: A Guide to Data
Analysis Using Spss for Windows**.Palgrave Macmillan, Houndmills, UK
- Broemeling, L. and M. Land. 1984. On forecasting with univariate autoregressive process: A
Bayesian approach. **Comm. Statist. Theory Methods**. 13:1305-1320.
- Cipra, T., J. Trujillo and A. Rubio. 1995. Holt-Winters method with missing observations.
Manag. Sci. 41:174-78.
- Cipra, T. 2006. Exponential smoothing for irregular data. **Appl. Math**. 51:597-604.
- Congdon, P. 2006. **Bayesian Statistical Modelling**, 2nd ed., John Wiley & Sons:
New York, pp. 1-56.
- Deetae, N.1991. **A comparative study of forecasting techniques on rice, cassava and
mungbean. Farm Gate Prices**. Master Thesis, Kasetsart University, Thailand.63-64.
- Evans, M., N. Hastings and J.B. Peacock. 2000. **Statistical Distribution**.3rd ed., John Wiley
and Sons, Inc., New York.

Fourth, L.A. and J.W. Woodlock. 1960. Early prediction of market success for grocery products.

Journal of Marketing. 61: 68-78.

Fundamentals of Neural Networks, blog, 10 May, 2012, www.myreaders.info/html/artificial_intelligence.html.

Gnanadesikan, R. and J.R. Kettenring. 1972. Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data. **Biometrics**. 28: 81-124.

Hagan, M.T., Demuth, H.B., and Beale, M.H. 1996. **Neural Network Design**, PWS Publ. Company.

Haykin, S. S. 1999. **Neural Networks: A Comprehensive Foundation**. Prentice Hall, Chapter 1-15: 1-889.

Hyndman, R.J., Koehler, A.B., Snyder, R.D. and Grose, S. 2002. A state space framework for automatic forecasting using exponential smoothing methods. **Int. J. Forecasting**. 18: 439-454.

Iqbal, N., K. Bakhsh, A. Maqbool and A.S. Ahmad. 2005. Use of the ARIM Model for Forecasting Wheat Area and Production in Pakistan. **Journal of Agriculture & Social Sciences**. 1(2): 120-122.

Kahforoushan, E., Zarif, M., and Mashahir, E.B. 2010. Prediction of added value of agricultural subsectors using artificial neural networks: Box-Jenkins and Holt-Winters methods. **Journal of Development and Agricultural Economics**. 2(4): 115-121.

Kerdsomboon, M. 1999. **Forecasting of agricultural products and prices**. Master Thesis, Chulalongkorn University, Thailand. 197-200.

Lee, R.D. and L.R. Carter. 1992. Modeling and forecasting U.S. mortality (discussion 671-675). **Journal of the American Statistician Association**. 87: 659-671.

Lui, S. 1994. Mutiperiod Bayesian forecasts for AR models. **Ann. Inst. Statist. Mat.** 46(3): 429-452.

Mishra, A.K. and V.R. Desai. 2005. Drought forecasting using stochastic models. **Stochastic Environmental Research and Risk Assessment**. 19(5): 326-339.

Monahan, J.F. 1983. Fully Bayesian analysis of ARMA time series models. **J. Econometrics**. 21: 307-3631.

Montgomery, D.C., Jennings, C.L., Kulahci, M. 2008. **Introduction to Time Series Analysis and**

- Forecasting**A. John Wiley & Sons. Inc.
- National News Bureau of Thailand. **Thai agricultural area and path towards the kitchen of the world.** blog, 3 Jan, 2011, <http://thainews.prd.go.th/view>
- Neelamegham, R. and P. Chintagunta. 1999. A Bayesian model to forecast new product performance in domestic and international markets. **Marketing science**. 18(2): 115-136.
- Pedroza, C. 2006. A Bayesian forecasting model: predicting U.S. male mortality. **Biostatistics**. 7(4): 530-550.
- Pena, D. and F.J. Prieto. 2001. Multivariate Outlier Detection and Robust Covariance Matrix Estimation. **Technometrics**. 43: 286-300.
- Rencher, A.C. 1998. **Multivariate Statistical Inference and Applications**. John Wiley & Sons, Inc., New York.
- Robert, C. 2001. **The Bayesian Choice: From Decision-Theoretic Foundation**, 2nd ed., Springer: New York, pp. 8-31.
- Robert, C.P. and G. Casella. 2004. **Monte Carlo Statistical Methods**, 2nd ed., Springer: New York, pp. 267-291.
- Sangpattaranate, P. 2005. **Forecasting of rice prices in Thailand**. Master Thesis, Kasetsart University, Thailand. 149-153.
- SAS. 2011. SAS/STAT®9.3 User's Guide. SAS Institute Inc., North Carolina.
- Sumer, K.K., O. Goktas and A. Hepsag. 2009. The application of seasonal latent variable in forecasting electricity demand as an alternative method. **Energy Policy**. 37(4): 1317-1322.
- Tsay, R. S. **TSMODEL Algorithms**. The University of Chicago, USA.
- West, M., Harrison, J. 1997. **Bayesian forecasting and dynamic models**, 2nd ed., Springer. New York
- Wright, D.J. 1986. Forecasting data published at irregular time intervals using extension of Holt's method. **Manag. Sci.** 32: 499-510.
- Yan, X., and Su, X.G. 2009. **Linear Regression Analysis Theory and Computing**. World Scientific Publishing Company.

- Yelland, P. 2009. Bayesian forecasting for low-count time series using stat-space models: An empiricalevaluation for inventory management. **International Journal of Production Economics**. 118: 95-103.
- Yelland, P.M. 2010. Bayesian Forecasting of Part Demand. **International Journal of Forecasting**. 26: 374-396.



ภาพผนวก

The full conditional distribution ของ σ_Y^2 คือ

$$\begin{aligned} & p(\sigma_Y^2 | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \xi_{i1}, \dots, \xi_{in}, \lambda_i, \mu_{\alpha_i}, \sigma_{\alpha}^2, \mu_{\delta_i}, \sigma_{\delta}^2, \mu_{\gamma_i}, \sigma_{\gamma}^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\ &= \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \xi_{it}, \sigma_Y^2, \lambda_i, \mu_{\alpha_i}, \sigma_{\alpha}^2, \mu_{\delta_i}, \sigma_{\delta}^2, \mu_{\gamma_i}, \sigma_{\gamma}^2, \sigma_Z^2) [p(\sigma_Y^2)] \end{aligned}$$

The full conditional distribution ของ σ_{α}^2 คือ

$$\begin{aligned} & p(\sigma_{\alpha}^2 | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \xi_{i1}, \dots, \xi_{in}, \lambda_i, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\delta_i}, \sigma_{\delta}^2, \mu_{\gamma_i}, \sigma_{\gamma}^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\ &= \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \xi_{it}, \sigma_Y^2, \lambda_i, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\delta_i}, \sigma_{\delta}^2, \mu_{\gamma_i}, \sigma_{\gamma}^2, \sigma_Z^2) [p(\sigma_{\alpha}^2)] \end{aligned}$$

The full conditional distribution ของ λ_i คือ

$$\begin{aligned} & p(\lambda_i | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \xi_{i1}, \dots, \xi_{in}, \sigma_{\alpha}^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\delta_i}, \sigma_{\delta}^2, \mu_{\gamma_i}, \sigma_{\gamma}^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\ &= \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \xi_{it}, \sigma_Y^2, \sigma_{\alpha}^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\delta_i}, \sigma_{\delta}^2, \mu_{\gamma_i}, \sigma_{\gamma}^2, \sigma_Z^2) [p(\lambda_i)] \end{aligned}$$

The full conditional distribution ของ μ_{γ_i} คือ

$$\begin{aligned} & p(\mu_{\gamma_i} | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \xi_{i1}, \dots, \xi_{in}, \sigma_{\alpha}^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\delta_i}, \sigma_{\delta}^2, \lambda_i, \sigma_{\gamma}^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\ &= \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \xi_{it}, \sigma_Y^2, \sigma_{\alpha}^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\delta_i}, \sigma_{\delta}^2, \lambda_i, \sigma_{\gamma}^2, \sigma_Z^2) [p(\mu_{\gamma_i})] \end{aligned}$$

The full conditional distribution ของ μ_{δ_i} คือ

$$\begin{aligned} & p(\mu_{\delta_i} | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \xi_{i1}, \dots, \xi_{in}, \sigma_{\alpha}^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\gamma_i}, \sigma_{\delta}^2, \lambda_i, \sigma_{\gamma}^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in}) \\ &= \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \xi_{it}, \sigma_Y^2, \sigma_{\alpha}^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\gamma_i}, \sigma_{\delta}^2, \lambda_i, \sigma_{\gamma}^2, \sigma_Z^2) [p(\mu_{\delta_i})] \end{aligned}$$

The full conditional distribution ของ σ_δ^2 คือ

$$p(\sigma_\delta^2 | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \xi_{i1}, \dots, \xi_{in}, \sigma_\alpha^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\gamma_i}, \mu_{\delta_i}, \lambda_i, \sigma_\gamma^2, \sigma_Z^2, Y_{i1}, \dots, Y_{in})$$

$$= \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \xi_{it}, \sigma_Y^2, \sigma_\alpha^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\gamma_i}, \mu_{\delta_i}, \lambda_i, \sigma_\gamma^2, \sigma_Z^2) [p(\sigma_\delta^2)]$$

The full conditional distribution ของ σ_Z^2 คือ

$$p(\sigma_Z^2 | w(t), \alpha_i, \delta_i, A_{i1}, \dots, A_{in}, \xi_{i1}, \dots, \xi_{in}, \sigma_\alpha^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\gamma_i}, \sigma_\delta^2, \lambda_i, \sigma_\gamma^2, \mu_{\delta_i}, Y_{i1}, \dots, Y_{in})$$

$$= \prod_{t=1}^n f(Y_{it} | \gamma_i, w(t), \alpha_i, \delta_i, A_{it}, \xi_{it}, \sigma_Y^2, \sigma_\alpha^2, \mu_{\alpha_i}, \sigma_Y^2, \mu_{\gamma_i}, \sigma_\delta^2, \lambda_i, \sigma_\gamma^2, \mu_{\delta_i}) [p(\sigma_Z^2)]$$



ประวัติผู้วิจัย

ชื่อ-สกุล: นายพิษณุ ทองขาว

(Mr. Pitsanu Tongkhaw)

ตำแหน่ง: อาจารย์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร
คณะวิศวกรรมศาสตร์ สาขาวิศวกรรมอุตสาหการ

การศึกษา: ว.ศ.ด. วิศวกรรมอุตสาหการ มหาวิทยาลัยเกษตรศาสตร์

ชื่อ-สกุล: ผศ. นิตยา บุญสิทธิ์

(Assist. Prof. Nittaya Boonsith)

ตำแหน่ง: ผู้ช่วยศาสตราจารย์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร
คณะวิทยาศาสตร์และเทคโนโลยี สาขาวิชาคณิตศาสตร์และสถิติ

การศึกษา: ค.ม (คณิตศาสตร์) จุฬาลงกรณ์มหาวิทยาลัย

ชื่อ-สกุล: นายกฤษฎา เหล็กดี

(Mr. Krisada Lekdee)

ตำแหน่ง: อาจารย์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร
คณะวิทยาศาสตร์และเทคโนโลยี สาขาวิชาคณิตศาสตร์และสถิติ

การศึกษา: ปริญญาตรี (สถิติ) มหาวิทยาลัยเกษตรศาสตร์

ชื่อ-สกุล: นางสาวอรศิริ จันทร์เมือง

(Miss. Onsiri Chanmuang)

ตำแหน่ง: อาจารย์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร
คณะวิศวกรรมศาสตร์ สาขาวิศวกรรมอุตสาหการ

การศึกษา: ว.ศ.ม. วิศวกรรมการจัดการอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระ
จอมเกล้าพระนครเหนือ