การเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในตัวแบบออโตรีเกรซซีฟอันดับ 1 เมื่อเกิดค่านอกระบบแบบ Innovation Outilers

A Comparative Study on the Parameter Estimation Problem for AR(1) Model with Innovation Outliers

สมเกียรติ กนอมวงษ์ *

บทคัดย่อ

งานวิจัยนี้เป็นการศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในตัวแบบออโตรีเกรซซีฟอันดับ 1 เมื่อเกิดค่านอกระบบแบบ Innovation Outliers(IO) โดยวิธีการประมาณค่าพารามิเตอร์ 2 วิธี คือ ตัวประมาณเอ็มเอ็มแอล(Modified Maximum Likelihood Estimator: MML) และตัวประมาณกำลังสองน้อยที่สุด (Least SquaresEstimator: LS) การเปรียบเทียบกระทำภายใต้เงื่อนไขของขนาดตัวอย่างเท่ากับ 30 และ 60 ค่าพารามิเตอร์ของตัวแบบเท่ากับ 0.2 และ 0.8 การแจกแจงของการรบกวนสุ่มที่มีการแจกแจงแบบแกมมา การแจกแจงของค่านอกระบบแบบ IO ที่มีการแจกแจงแบบไวบูลและลอคนอร์มัล และสัดส่วนของค่านอกระบบแบบ IO เท่ากับ 30%, 50% และ 70% สำหรับข้อมูลที่ใช้ในการวิจัยครั้งนี้ได้จากการจำลองข้อมูลด้วยโปรแกรม MINITAB ทำซ้ำ 100 ครั้งในแต่ละสถานการณ์ที่กำหนด และใช้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) เป็นเกณฑ์ในการเปรียบเทียบ ซึ่งตัวประมาณที่มีประสิทธิภาพดีที่สุดจะมีค่า MSE ต่ำที่สุด ผลการวิจัยสรุปได้ว่าตัวประมาณเอ็มเอ็มแอลจะมีประสิทธิภาพดีกว่าตัวประมาณกำลังสองน้อยที่สุดเกือบทุกสถานการณ์ สำหรับปัจจัยที่มีผลกระทบต่อประสิทธิภาพในการประมาณค่าพารามิเตอร์ พบว่า ประสิทธิภาพของตัวประมาณทั้งสองมีแนวโน้มเพิ่มมากขึ้นเมื่อขนาดตัวอย่างเปลี่ยนจาก 30 เป็น 60 ค่าพารามิเตอร์ของตัวแบบเปลี่ยนจาก 0.2 เป็น 0.8 และค่านอกระบบแบบ IO ที่มีการแจกแจงแบบลอคนอร์มัล เมื่อสเกลพารามิเตอร์เปลี่ยนจาก 0.1 เป็น 0.5 ที่โลเคชันพารามิเตอร์เท่ากับ 1 ในขณะที่ประสิทธิภาพของตัวประมาณทั้งสองมีแนวโน้มลดลง เมื่อการรบกวนสุ่มที่มีการแจกแจงแบบแกมมาเมื่อสเกลพารามิเตอร์เปลี่ยนจาก 3 เป็น 4 ที่โลเคชันพารามิเตอร์ เท่ากับ 1 ค่านอกระบบแบบ IO ที่มีการแจกแจงแบบไวบูล เมื่อสเกลพารามิเตอร์เปลี่ยนจาก 3 เป็น 4 ที่โลเคชันพารามิเตอร์เท่ากับ 1 และสัดส่วนของค่านอกระบบแบบ IO เพิ่มขึ้น

Abstract

A Comparative Study on the Parameter Estimation Problem for AR(1) Model with Innovation Outliers was performed by using two estimators : Modified Maximum Likelihood

* อาจารย์ประจำสำนักส่งเสริมมาตรฐานการศึกษา มหาวิทยาลัยศรีปทุม

estimator (MML) and Least Squares estimator(LS). The comparison was done under the sample of sizes 30 and 60, the parameter of model were 0.2 and 0.8 whereas the distribution of the error a_t was Gamma and the distribution of Innovation Outliers e_t were Weibull and Lognormal having the percent of Innovation Outliers at 30%, 50% and 70%. The data of this experiment were generated by MINITAB Program. The experiment was repeated 100 times under each situation. The mean square error (MSE) value was used as the criterion for which the estimator having lowest MSE value was the most efficient estimator.

From the study we can conclude that the MML estimator is more efficient than the LS estimator in almost all situations. We also found that the efficiency of both estimators tends to be increased when the sample size changed from 30 to 60, the parameter of model changed from 0.2 to 0.8 and the scale parameter of Lognormal distribution of Innovation Outliers changed from 0.1 to 0.5 with location parameter was 1. On the other hand, the efficiency of both estimators tends to be decreased when the scale parameter of Gamma distribution of the error changed from 3 to 4 with location parameter was 1, the scale parameter of Weibull distribution of Innovation Outliers changed from 3 to 4 with location parameter was 1 and the percent of Innovation Outliers increased.

บทนำ

โดยทั่วไปการดำเนินชีวิตของคนส่วนใหญ่ในแต่ละวันเกี่ยวข้องกับการพยากรณ์อยู่ตลอดเวลา เช่น การคาดคะเนเวลาที่จะไปถึงสถานที่ทำงานการคาดคะเนยอดขายสินค้าหรือจำนวนผู้มาใช้บริการ เป็นต้น โดยเฉพาะในการประกอบกิจการหรือการบริหารงานที่ต้องมีการวางแผนล่วงหน้า การพยากรณ์จึงเป็นเครื่องมืออย่างหนึ่งที่ช่วยทำให้ได้ข้อมูลในอนาคตเพื่อประกอบการวางแผนนั้น ๆ ตัวอย่างเช่น การพยากรณ์ปริมาณสินค้าส่งออกชนิดหนึ่ง การพยากรณ์จำนวนนักท่องเที่ยวที่เดินทางเข้าประเทศและการพยากรณ์รายได้ประชาชาติ เป็นต้น จึงเป็นที่ยอมรับกันว่าการพยากรณ์มีบทบาทสำคัญทั้งในงานของภาครัฐและภาคเอกชน

การวิเคราะห์อนุกรมเวลา (Time Series Analysis) เป็นวิธีการทางสถิติวิธีหนึ่งที่สามารถพยากรณ์หรือประมาณค่าในอนาคตและแสดงให้เห็นถึงการเปลี่ยนแปลงของสิ่งที่เราสนใจในแต่ละช่วงเวลาที่เปลี่ยนไป โดยอาศัยข้อมูลในอดีตที่ใช้กันอยู่ประเภทหนึ่ง คือ ข้อมูลอนุกรมเวลา ซึ่งการวิเคราะห์อนุกรมเวลานี้มีบทบาทสำคัญมากในการวางแผนการดำเนินงานต่าง ๆ โดยเฉพาะด้านธุรกิจและเศรษฐศาสตร์ เช่น การวางแผนการผลิต การวางแผนการตลาด การวางแผนการจัดการบุคลากร และการวางแผนการจัดการสินค้าคงคลัง เป็นต้น การนำวิธีการวิเคราะห์อนุกรมเวลามาใช้เพื่อให้ได้ค่าพยากรณ์ที่มีความน่าเชื่อถือสูงจำเป็นต้องเลือกวิธีการพยากรณ์ที่เหมาะสมกับข้อมูลนั้น ๆ ซึ่งวิธีการวิเคราะห์อนุกรมเวลามีอยู่หลายวิธี เช่น เทคนิคการทำให้เรียบ (Smoothing Techniques) วิธีอนุกรมเวลาแบบบ็อกซ์-เจนกินส์ (Box-Jenkins Method) การกรองแบบปรับได้ (Adaptive Filtering) และวิธีอนุกรมเวลาแบบคลาสสิค (Clas-

sical Decomposition Method) เป็นต้น

การกำหนดตัวแบบที่เหมาะสมกับข้อมูลอนุกรมเวลานั้นเป็นขั้นตอนหนึ่งของการวิเคราะห์อนุกรมเวลา ซึ่งตัวแบบที่นิยมใช้กันมาก คือ ตัวแบบที่อธิบายอนุกรมเวลาเป็นฟังก์ชันเชิงเส้นของค่าสังเกตในอดีตและการรบกวนสุ่มทั้งในอดีตและปัจจุบันมีชื่อเรียกว่าตัวแบบ ARIMA หรือที่เรียกกันโดยทั่วไปว่า ตัวแบบบ็อกซ์เจนกินส์ (Box-Jenkins Model) หลังจากกำหนดตัวแบบที่เป็นไปได้ ขั้นตอนต่อไปก็คือการประมาณค่าพารามิเตอร์ในตัวแบบที่กำหนดข้างต้น ซึ่งการประมาณค่าพารามิเตอร์นั้นได้มีผู้ศึกษาวิธีการประมาณอย่างกว้างขวางจากอดีตจนถึงปัจจุบัน ภายใต้ข้อสมมติเบื้องต้นแตกต่างกันออกไป อาทิเช่น ในปี ค.ศ.1999 ติกู วอง และไบอัน (Tiku, Wong and Bian,1999) ได้นำเสนอวิธีการประมาณค่าพารามิเตอร์ในตัวแบบออโตรีเกรซซีฟ (Autoregressive Models) ชนิดที่การรบกวนสุ่มมีการแจกแจงต่างไปจากการแจกแจงแบบปกติ โดยวิธีการประมาณค่าพารามิเตอร์ที่เสนอ คือ ตัวประมาณเอ็มเอ็มแอล (Modified Maximum Likelihood Estimator: MML) รวมทั้งศึกษาทฤษฎี คุณสมบัติต่าง ๆ และเปรียบเทียบประสิทธิภาพของตัวประมาณเอ็มเอ็มแอลกับตัวประมาณกำลังสองน้อยที่สุด (Least Squares Estimator: LS)

นอกจากวิธีการประมาณค่าพารามิเตอร์โดยวิธีข้างต้นแล้ว ยังมีวิธีการประมาณวิธีอื่น ๆ อีก เช่น ตัวประมาณจีเอ็มเอ็ม (Generalized Method of Moments Estimator : GMM) และตัวประมาณของโกว (Guo's Estimator) เป็นต้น โดยแต่ละวิธีจะมีหลักการในการหาตัวประมาณที่แตกต่างกันไป

นอกเหนือจากการเลือกใช้วิธีการประมาณที่เหมาะสมแล้ว ข้อมูลในอดีตก็เป็นอีกปัจจัยหนึ่งที่ส่งผลถึงความถูกต้องและความน่าเชื่อถือของค่าพยากรณ์ เนื่องจากการวิเคราะห์อนุกรมเวลาจะต้องอาศัยข้อมูลหรือค่าสังเกตในอดีตที่ผ่านมา แต่ผู้วิเคราะห์อาจประสบปัญหาเกี่ยวกับข้อมูล เช่น ข้อมูลบางช่วงมีค่าสูงหรือต่ำกว่าปกติ เป็นต้น ซึ่งลักษณะข้อมูลที่มีค่าผิดปกตินี้อาจเป็นผลเนื่องมาจากลักษณะตามธรรมชาติของค่าสังเกตที่ไม่สามารถหลีกเลี่ยงได้ ถึงแม้ว่าเราจะควบคุมการวัดค่าสังเกตเป็นอย่างดี เช่น การเปลี่ยนแปลงนโยบายทางเศรษฐกิจที่มีผลต่อราคาสินค้า ผลกระทบจากการลดค่าของเงิน ผลกระทบจากการเกิดสงคราม ผลกระทบจากการนัดหยุดงานของพนักงานในบริษัทที่มีผลต่อปริมาณการผลิต ผลกระทบจากการเกิดภัยทางธรรมชาติอย่างรุนแรง ฯลฯ เหตุการณ์เหล่านี้ล้วนส่งผลกระทบต่อข้อมูลอนุกรมเวลาทำให้ข้อมูลที่เก็บรวบรวมไว้ในช่วงที่เกิดเหตุการณ์มีค่าผิดปกติไปจากข้อมูลช่วงอื่น ๆ เราเรียกข้อมูลที่ได้รับผลกระทบต่าง ๆ นี้ว่าค่านอกระบบ(Outlier)

แนวคิดเริ่มแรกในการศึกษาเกี่ยวกับค่านอกระบบเกิดขึ้นในปี ค.ศ.1972 ฟอกซ์ (Fox, 1972) ศึกษาเกณฑ์อัตราส่วนภาวะน่าจะเป็น (Likelihood Ratio Criteria) ที่ใช้ในการทดสอบการเกิดค่านอกระบบในข้อมูลอนุกรมเวลา โดยวิธีอัตราส่วนภาวะน่าจะเป็น (Likelihood Ratio Method) สำหรับค่านอกระบบที่ศึกษามีด้วยกัน 2 ชนิด คือ Innovation Outliers (IO) และ Additive Outliers (AO) ในปี ค.ศ.1979 เดนบีและมาร์ติน (Denby and Martin, 1979) ศึกษาความแกร่งของตัวประมาณค่าพารามิเตอร์ในตัวแบบออโตรีเกรซซีฟ อันดับ 1 เมื่อเกิดค่านอกระบบแบบ IO และ AO โดยเปรียบเทียบความเอนเอียงและค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณจีเอ็ม (Generalized M-Estimator : GM) กับตัวประมาณกำลังสองน้อยที่สุด จากบทความที่กล่าวมาข้างต้นเป็นวิธีการประมาณค่าพารามิเตอร์ที่มีข้อสมมติเบื้องต้นแตกต่างกันออกไป สำหรับการวิจัยครั้งนี้เป็นการศึกษาเพิ่มเติมจากเอกสารของติกู วอง และไบอัน (Tiku, Wong and Bian, 1999)

ในกรณีที่เกิดค่านอกระบบแบบ IO โดยศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณเอ็มเอ็มแอลกับตัวประมาณกำลังสองน้อยที่สุด ในตัวแบบออโตรีเกรซซีฟอันดับ 1 ชนิดที่การรบกวนสุ่มมีการแจกแจงแบบแกมมา (Gamma Distribution) ว่าตัวประมาณใดจะมีประสิทธิภาพดีกว่ากัน ซึ่งเกณฑ์ที่ใช้เปรียบเทียบประสิทธิภาพของทั้งสองวิธี คือ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) ของค่าพารามิเตอร์

วัตถุประสงค์การวิจัย

การวิจัยนี้มีวัตถุประสงค์ ประการแรกเพื่อศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณเอ็มเอ็มแอลกับตัวประมาณกำลังสองน้อยที่สุดในการประมาณค่าพารามิเตอร์ สำหรับตัวแบบออโตรีเกรซซีฟอันดับ 1 ชนิดที่การรบกวนสุ่มมีการแจกแจงแบบแกมมา เมื่อเกิดค่านอกระบบแบบ IO และประการที่สองเพื่อศึกษาปัจจัยต่าง ๆ ที่มีผลต่อประสิทธิภาพของการประมาณค่าพารามิเตอร์ทั้งสองวิธี

สมมติฐานการวิจัย

ตัวประมาณเอ็มเอ็มแอลมีประสิทธิภาพดีกว่าตัวประมาณกำลังสองน้อยที่สุดในตัวแบบออโตรีเกรซซีฟอันดับ 1 ชนิดที่การรบกวนสุ่มมีการแจกแจงแบบแกมมา เมื่อเกิดค่านอกระบบแบบ IO

ขอบเขตการวิจัย

การศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณเอ็มเอ็มแอลกับตัวประมาณกำลังสองน้อยที่สุดในการประมาณค่าพารามิเตอร์ มีเงื่อนไขดังต่อไปนี้

ตัวแบบ AR(1) คือ $Y_t = \phi Y_{t-1} + a_t$; t = 1, 2, 3, ..., n เมื่อ a_t แทนการรบกวนสุ่ม ณ เวลา t ที่มีการแจกแจงแบบแกมมา; $\Gamma(b, \sigma)$ ค่าพารามิเตอร์ของตัวแบบเท่ากับ 0.2 และ 0.8 การรบกวนสุ่ม (a_t) ที่มีการแจกแจงแบบแกมมา; $\Gamma(b, \sigma)$ กำหนดให้ b = 3, 4 และ σ = 1 ค่านอกระบบแบบ IO (e_t) มีการแจกแจง 2 การแจกแจง ได้แก่ การแจกแจงแบบไวบูล ; $W(\lambda_1, \lambda_2)$ กำหนดให้ $\lambda_1 = 3, 4$ และ $\lambda_2 = 1$ การแจกแจงแบบลอคนอร์มัล ; $LN(\theta_1, \theta_2)$ กำหนดให้ $\theta_1 = 1$ และ $\theta_2 = 0.1, 0.5$ ขนาดตัวอย่างเท่ากับ 30 และ 60 สัดส่วนของค่านอกระบบแบบ IO เท่ากับ 30%, 50% และ 70% ของขนาดตัวอย่าง สำหรับเกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพ คือ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

ประโยชน์ที่ได้รับจากการวิจัย

เพื่อเป็นแนวทางให้ผู้วิจัยสามารถเลือกใช้วิธีการประมาณค่าพารามิเตอร์ของข้อมูลอนุกรมเวลาได้อย่างเหมาะสมและมีประสิทธิภาพ เมื่อข้อมูลเกิดค่านอกระบบแบบ IO และสามารถพยากรณ์หรือประมาณค่าในอนาคตได้อย่างถูกต้องแม่นยำ ซึ่งการวิเคราะห์อนุกรมเวลาจะมีบทบาทความสำคัญในการวางแผนการดำเนินงานต่างๆ

การดำเนินการวิจัย

วิธีการดำเนินการวิจัยครั้งนี้ ใช้วิธีการจำลองข้อมูลและคำนวณค่าตัวประมาณทั้งสองด้วยโปรแกรม MINITAB และเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ยด้วยโปรแกรมสำเร็จรูป Microsoft Excel นอกจากนี้กำหนดให้มีการทำซ้ำ 100 ครั้ง ในแต่ละสถานการณ์ ผู้วิจัยขอแสดงขั้นตอนการดำเนินการวิจัยตามลำดับต่อไปนี้

1. การจำลองข้อมูลอนุกรมเวลา เมื่อเกิดค่านอกระบบแบบ IO จำลองข้อมูลอนุกรมเวลา Y_t ; t = 1, 2, ..., n เมื่อ n = 30 และ 60 ตามตัวแบบ AR(1) ที่มีค่า

พารามิเตอร์ของตัวแบบเท่ากับ 0.2 และ 0.8 เมื่อเกิดค่านอกระบบแบบ IO

1.1 กำหนดให้ $a_0 = E(a_t) = b\sigma$ ตามฟังก์ชันการแจกแจงของการรบกวนสุ่มนั้น

1.2 คำนวณ $Y_0 = \frac{a_0}{\sqrt{1-\phi^2}}$

1.3 สร้าง a_t ; t = 1, 2, ..., n ให้มีการแจกแจงแบบแกมมา ; $\Gamma(b, \sigma)$ กำหนดให้ b = 3, 4 และ σ = 1

1.4 สร้าง Y_t ; t = 1, 2, ..., n ให้มีรูปแบบความสัมพันธ์คือ $Y_t = \phi Y_{t-1} + a_t$

1.5 สร้างค่านอกระบบแบบ IO (e_t) ที่มีฟังก์ชันการแจกแจงดังต่อไปนี้ การแจกแจงแบบไวบูล (Weibull Distribution) ; $W(\lambda_1, \lambda_2)$ กำหนดให้ λ_1 = 3, 4 และ λ_2 = 1 และการแจกแจงแบบลอคนอร์มัล (Lognormal Distribution) ; $LN(\theta_1, \theta_2)$ กำหนดให้ θ_1 = 1 และ θ_2 = 0.1, 0.5

และกำหนดสัดส่วนของค่านอกระบบแบบ IO เท่ากับ 30%, 50% และ 70% ของขนาดตัวอย่าง

1.6 สร้าง W_t เมื่อเกิดค่านอกระบบแบบ IO ; t = 1, 2, ..., n ให้มีรูปแบบความสัมพันธ์ ดังนี้

$$W_t = Y_t \; ; \; t < t_o$$
$$W_t = Y_t + e_t \; ; \; t = t_o, t_o+1, t_o+2, \ldots, n$$

2. การประมาณค่าพารามิเตอร์โดยตัวประมาณเอ็มเอ็มแอลและตัวประมาณกำลังสองน้อยที่สุด

2.1 ตัวประมาณเอ็มเอ็มแอล (Tiku, M.L., Wong, W.K. and Bian, G,1999)

$$\hat{\phi}_{MML} = K - D\hat{\sigma}$$

โดยที่

$$\hat{\sigma} = \left\{ -J + \sqrt{J^2 + 4nC} \right\} / 2n,$$

$$J = (b-1)\sum_{t=1}^{n} M_t \left(W_{[t]} - KW_{[t]-1} \right),$$

$$C = (b-1)\sum_{t=1}^{n} \beta_t \left(W_{[t]} - KW_{[t]-1} \right)^2,$$

$$K = \frac{\sum_{t=1}^{n} \beta_t W_{[t]} W_{[t]-1}}{\sum_{t=1}^{n} \beta_t W_{[t]-1}^2}$$

$$M_t = \left(\alpha_t - \frac{1}{b-1} \right),$$

$$D = \frac{\sum_{t=1}^{n} M_t W_{[t]-1}}{\sum_{t=1}^{n} \beta_t W_{[t]-1}^2}$$

เมื่อค่าสังเกต $(W_{[t]}, W_{[t]-1})$ แทน (W_t, W_{t-1})

ในรูป $Z_{(t)} = \frac{W_{[t]} - \phi W_{[t]-1}}{\sigma}$

2.2 ตัวประมาณกำลังสองน้อยที่สุด (Tiku, M.L., Wong, W.K. and Bian, G,1999)

$$\hat{\phi}_{LS} = \frac{U}{V}$$

โดยที่

$$U = \sum_{t=1}^{n} W_t \left(W_{t-1} - \overline{W} \right),$$

$$V = \sum_{t=1}^{n} \left(W_{t-1} - \overline{W} \right)^2,$$

$$\overline{W} = \frac{\sum_{t=1}^{n} W_t}{n}$$

3. ทำซ้ำ 100 ครั้ง

4. ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(MSE) ของตัวประมาณ

$$\text{MSE} = \frac{\sum_{j=1}^{N}(\hat{\phi}_j - \phi)^2}{N}$$

เมื่อ ϕ แทน ค่าพารามิเตอร์ของตัวแบบที่กำหนด

$\hat{\phi}_j$ แทน ค่าประมาณจากตัวประมาณจากการทำซ้ำที่ j ; j = 1, 2, ..., N

N แทน จำนวนการทำซ้ำ

สรุปผลการวิจัย

จากการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าพารามิเตอร์ทั้งสองวิธีได้ข้อสรุปดังนี้

1. การเปรียบเทียบประสิทธิภาพตัวประมาณเอ็มเอ็มแอลกับตัวประมาณกำลังสองน้อยที่สุด

ตัวประมาณเอ็มเอ็มแอลมีประสิทธิภาพดีกว่าตัวประมาณกำลังสองน้อยที่สุดในแทบทุกสถานการณ์ ยกเว้นสถานการณ์ที่ข้อมูลอนุกรมเวลาตัวแบบ AR(1) มีค่าพารามิเตอร์ของตัวแบบเท่ากับ 0.2 ชนิดที่การรบกวนสุ่มมีการแจกแจงแกมมา; $\Gamma(b, \sigma)$ กำหนดให้ $b = 3$ และ $\sigma = 1$ เมื่อเกิดค่านอกระบบแบบ IO ที่มีการแจกแจงแบบลอคนอร์มัล; $LN(\theta_1, \theta_2)$ กำหนดให้ $\theta_1 = 1$ และ $\theta_2 = 0.1, 0.5$ ที่ขนาดตัวอย่างเท่ากับ 60 ในทุกสัดส่วนของค่านอกระบบแบบ IO และอีกกรณีที่การรบกวนสุ่มมีการแจกแจงแกมมา; $\Gamma(b, \sigma)$ กำหนดให้ $b = 4$ และ $\sigma = 1$ เมื่อเกิดค่านอกระบบแบบ IO ที่มีการแจกแจงแบบลอคนอร์มัล; $LN(\theta_1, \theta_2)$ กำหนดให้ $\theta_1 = 1$ และ $\theta_2 = 0.5$ ที่ขนาดตัวอย่างเท่ากับ 60 ในทุกสัดส่วนของค่านอกระบบแบบ IO

ตารางที่ 1 แสดงค่า MSE ที่การรบกวนสุ่มมีการแจกแจงแบบแกมมา เมื่อเกิดค่านอกระบบแบบ IO

a_t	e_t	ϕ	ตัวประมาณ	$n = 30$			$n = 60$		
				30%	50%	70%	30%	50%	70%
$\Gamma(3,1)$	$W(3,1)$	0.2	MML	0.01701	0.02387	0.02976	0.01918	0.02132	0.02563
			LS	0.21208	0.28619	0.28613	0.16719	0.22670	0.26734
		0.8	MML	0.00078	0.00091	0.00129	0.00094	0.00106	0.00123
			LS	0.00811	0.02004	0.01990	0.02056	0.02632	0.02673
	$W(4,1)$	0.2	MML	0.02163	0.03765	0.04815	0.02655	0.03412	0.04768
			LS	0.21427	0.28975	0.29554	0.16961	0.22457	0.27105
		0.8	MML	0.00110	0.00157	0.00209	0.00141	0.00166	0.00245
			LS	0.00783	0.01911	0.02016	0.02069	0.02565	0.02679

a_t	e_t	ϕ	ตัว	$n=30$			$n=60$		
			ประมาณ	30%	50%	70%	30%	50%	70%
$\Gamma(3,1)$	$LN(1,0.1)$	0.2	MML	0.03558	0.05207	0.08055	0.02545	0.03028	0.04630
			LS	0.04327	0.05876	0.08470	0.02518	0.02741	0.03439
		0.8	MML	0.00269	0.00445	0.00849	0.00286	0.00455	0.00603
			LS	0.27525	0.43010	0.54559	0.09441	0.16078	0.26457
	$LN(1,0.5)$	0.2	MML	0.02904	0.03257	0.03699	0.02130	0.02402	0.02917
			LS	0.04241	0.03613	0.03709	0.01640	0.01678	0.01898
		0.8	MML	0.00215	0.00249	0.00310	0.00147	0.00168	0.00225
			LS	0.25316	0.31180	0.23420	0.08717	0.12161	0.11898
$\Gamma(4,1)$	$W(3,1)$	0.2	MML	0.01148	0.01475	0.02185	0.01790	0.01860	0.02337
			LS	0.30849	0.37106	0.38040	0.24382	0.32208	0.35129
		0.8	MML	0.00036	0.00039	0.00056	0.00064	0.00065	0.00081
			LS	0.01834	0.03370	0.03613	0.03033	0.03400	0.03375
	$W(4,1)$	0.2	MML	0.01661	0.02638	0.03990	0.02725	0.03812	0.04642
			LS	0.30833	0.37247	0.38651	0.24544	0.32129	0.36377
		0.8	MML	0.00060	0.00076	0.00117	0.00110	0.00136	0.00179
			LS	0.01828	0.03350	0.03660	0.03660	0.03367	0.03372
	$LN(1,0.1)$	0.2	MML	0.06263	0.08839	0.14207	0.04023	0.05411	0.08099
			LS	0.10361	0.13261	0.15453	0.06011	0.08722	0.10793
		0.8	MML	0.00426	0.00625	0.01121	0.00275	0.00374	0.00593
			LS	0.14403	0.30863	0.69916	0.02159	0.02452	0.09259
	$LN(1,0.5)$	0.2	MML	0.04360	0.04230	0.04027	0.03130	0.03694	0.03497
			LS	0.06465	0.06009	0.04319	0.02798	0.03081	0.02784
		0.8	MML	0.00285	0.00280	0.00318	0.00222	0.00223	0.00257
			LS	0.18284	0.28865	0.38709	0.05360	0.06196	0.08737

2. การศึกษาปัจจัยต่างๆ ที่มีผลต่อประสิทธิภาพในการประมาณค่าพารามิเตอร์

2.1 ขนาดตัวอย่าง

ขนาดตัวอย่างเป็นปัจจัยที่ผลต่อประสิทธิภาพในการประมาณค่าพารามิเตอร์ เมื่อทราบสถานการณ์ต่าง ๆ ทุกสถานการณ์ ภายใต้เงื่อนไขการวิจัยครั้งนี้พบว่า เมื่อขนาดตัวอย่างเพิ่มขึ้น ค่า MSE ของตัวประมาณทั้งสองจะลดลง

ตารางที่ 2 แสดงค่า MSE กรณีขนาดตัวอย่าง

ขนาดตัวอย่าง	วิธีการประมาณค่าพารามิเตอร์	
	MML	LS
30	0.02218	0.18542
60	0.01768	0.10798

2.2 ค่าพารามิเตอร์ของตัวแบบ

เมื่อทราบสถานการณ์ ขนาดตัวอย่าง การแจกแจงของการรบกวนสุ่ม การแจกแจงของค่านอกระบบแบบ IO และสัดส่วนของค่านอกระบบแบบ IO ทุกสถานการณ์ภายใต้เงื่อนไขการวิจัยครั้งนี้ กรณีเฉพาะค่าพารามิเตอร์ของตัวแบบ เมื่อค่าพารามิเตอร์ของตัวแบบเพิ่มขึ้นค่า MSE ของตัวประมาณทั้งสองจะลดลง

ตารางที่ 3 แสดงค่า MSE กรณีค่าพารามิเตอร์ของตัวแบบ

ค่าพารามิเตอร์ของตัวแบบ	วิธีการประมาณค่าพารามิเตอร์	
	MML	LS
0.2	0.03738	0.17124
0.8	0.00248	0.12217

2.3 การแจกแจงของการรบกวนสุ่ม

เมื่อทราบสถานการณ์ ขนาดตัวอย่างค่าพารามิเตอร์ของตัวแบบ การแจกแจงของค่านอกระบบแบบ IO และสัดส่วนของค่านอกระบบแบบ IO ทุกสถานการณ์ ภายใต้เงื่อนไขการวิจัยครั้งนี้ กรณีเฉพาะการแจกแจงของการรบกวนสุ่ม เมื่อการรบกวนสุ่มมีการแจกแจงแบบแกมมา ที่มีลักษณะหางยาวเพิ่มขึ้น ค่า MSE ของตัวประมาณทั้งสองจะเพิ่มขึ้น

ตารางที่ 4 แสดงค่า MSE กรณีการแจกแจงของการ รบกวนสุ่ม

การแจกแจงของการรบกวนสุ่ม	วิธีการประมาณค่าพารามิเตอร์	
	MML	LS
$\Gamma(3,1)$	0.01789	0.13515
$\Gamma(4,1)$	0.02197	0.15825

2.4 การแจกแจงของค่านอกระบบแบบ IO

เมื่อทราบสถานการณ์ ขนาดตัวอย่าง ค่าพารามิเตอร์ของตัวแบบ การแจกแจงของการรบกวนสุ่มและสัดส่วนของค่านอกระบบแบบ IO ทุกสถานการณ์ภายใต้เงื่อนไขการวิจัยครั้งนี้ กรณีเฉพาะการแจกแจงของค่านอกระบบแบบ IO เป็นที่น่าสังเกตว่าการแจกแจงของค่านอกระบบแบบ IO มีผลต่อการประมาณค่าพารามิเตอร์ กล่าวคือ ค่านอกระบบแบบ IO มีการแจกแจงแบบไวบูล ที่มีลักษณะหางยาวเพิ่มขึ้น ค่า MSE ของตัวประมาณทั้งสองจะเพิ่มขึ้น ส่วนค่านอกระบบแบบ IO มีการแจกแจงแบบลอคนอร์มัล ที่มีลักษณะหางยาวเพิ่มขึ้น ค่า MSE ของตัวประมาณทั้งสองจะลดลง

ตารางที่ 5 แสดงค่า MSE กรณีการแจกแจงของค่านอกระบบแบบ IO

การแจกแจงของค่านอกระบบแบบ IO	วิธีการประมาณค่าพารามิเตอร์	
	MML	LS
W(3,1)	0.01059	0.15544
W(4,1)	0.01781	0.15663
LN(1,0.1)	0.03358	0.16570
LN(1,0.5)	0.01773	0.10903

2.5 สัดส่วนของค่านอกระบบแบบ IO

เมื่อทราบสถานการณ์ ขนาดตัวอย่าง ค่าพารามิเตอร์ของตัวแบบ การแจกแจงของการรบกวนสุ่ม และการแจกแจงของค่านอกระบบแบบ IO ทุกสถานการณ์ ภายใต้เงื่อนไขการวิจัยครั้งนี้ กรณีเฉพาะสัดส่วนของค่านอกระบบแบบ IO เมื่อสัดส่วนของค่านอกระบบแบบ IO เพิ่มขึ้น ค่า MSE ของตัวประมาณทั้งสองจะเพิ่มขึ้น

ตารางที่ 6 แสดงค่า MSE กรณีสัดส่วนของค่านอกระบบแบบ IO

สัดส่วนของค่านอกระบบแบบ IO	วิธีการประมาณค่าพารามิเตอร์	
	MML	LS
30%	0.01484	0.10999
50%	0.01893	0.15013
70%	0.02602	0.17999

ข้อเสนอแนะ

งานวิจัยต่อไป คือ ศึกษาในกรณีที่ข้อมูลอนุกรมเวลา มีผลกระทบจากปัจจัยอื่น ๆ เช่น Additive Outliers เป็นต้น หรือศึกษาในกรณีที่อนุกรมเวลามีรูปแบบอื่น ๆ เช่น AR(2), AR(3), ..., AR(p)

สำหรับการประมาณค่าพารามิเตอร์ครั้งนี้ ศึกษาเฉพาะวิธีการประมาณค่าพารามิเตอร์ 2 วิธีดังกล่าวข้างต้น ยังมีวิธีอื่น ๆ ที่น่าสนใจ เช่น ตัวประมาณของโกว (Guo's Estimator) ฯลฯ ที่จะทำการศึกษาต่อไป

บรรณานุกรม

ขวัญวงศ์ วัชโรทน. **การศึกษาวิธีการประมาณค่าพารามิเตอร์ของตัวแบบออโตรีเกรซซีฟอันดับที่ 1 และตัวแบบออโตรีเกรซซีฟอันดับที่ 2 เมื่อเกิดค่านอกระบบ.** วิทยานิพนธ์มหาบัณฑิต ภาควิชาสถิติประยุกต์ บัณฑิตวิทยาลัย มหาวิทยาลัยธรรมศาสตร์, 2543.

รัชนี วงค์อนันต์กิจ. **ผลกระทบของการเกิด additive outlier ต่อค่าประมาณจากตัวแบบ aggregated และ disaggregated ARMA(1,1).** สารนิพนธ์มหาบัณฑิต ภาควิชาสถิติประยุกต์ บัณฑิตวิทยาลัย มหาวิทยาลัยธรรมศาสตร์, 2536.

สุพร ฉัตรแก้วรัตนกุล. **การประมาณค่าพารามิเตอร์ของตัวแบบอนุกรมเวลา เมื่อข้อมูลมีค่าผิดปกติ.** วิทยานิพนธ์มหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2538.

Abraham, B. and Ledolter, J. (1942). **Statistical Methods for Forcasting.** New York : John Wiley & Sons.

Akkaya, A.D. and Tiku, M.L. (2001). "CORRIGENDUM : Time Series Models with Asymmetric Innovations." **Communication Statistics - Theory Meth.,** 30(10), 2227-2230.

Akkaya, A.D. and Tiku, M.L. (2001). "Estimating Parameters in Autoregressive Models in Non-normal Stuations : Asymmetric Innovations." **Communication Statistics-Theory Meth.,** 30(3), 517-536.

David, H.A. (1981). **Order Statistics.** New York : John Wiley & Sons.

Denby, L. and Martin, R.D. (1979). "Robust Estimation of the First Order Autoregressive Parameter." **Journal of the American Statistical Association,** 30, 193-204.

Fox, A.J. (1972). "Outliers in Time Series." **Journal of the Royal Statistical Society Series B,** 3, 350-363.

Guo, J.H. (2000). "Robust Estimation for the Coefficient of a First Order Autoregressive Processes." **Communication Statistics-Theory Meth.,** 29(1), 55-66.

Mood, A.M., Graybill, F.A. and Boes, D.C.(1974). **Introduction to Theory of Statistics.** New York : McGraw Hill.

Tiku, M.L., Tan, W.Y. and Balakrishnan, N. (1986). **Robust Inference.** New York : Marcel Dekker.

Tiku, M.L., Wong, W.K. and Bian, G. (1999). "Time Series Models with Asymmetric Innovations." **Communication Statistics-Theory Meth.,** 28(6), 1131-1160.

Vougas, D.V. (2000). "A comparison of LS/ML and GMM Estimation in a Simple AR(1) Model." **Communication Statistics-Simula.,** 29(1), 239-258.