การพัฒนาวิธีประมาณค่าพารามิเตอร์ของตัวแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง เมื่อข้อมูลมีค่าสูญหาย

The Development of the Parameter Estimation Methods in First-Order Autoregressive Model with Missing Values

วราฤทธิ์ พานิชกิจโกศลกุล

ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต ปทุมธานี 12121

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาและเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ (ρ) ของตัวแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง เมื่อข้อมูลมีค่าสูญหาย ด้วยวิธีการประมาณ 3 วิธีคือ วิธีกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิด (RM) วิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิด (RMD) และวิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิดปรับปรุง (IRMD) ร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 5 และ 10 ขนาดตัวอย่าง 4 ระดับ คือ 25, 50, 100 และ 250 ในการวิจัยครั้งนี้ใช้วิธีการจำลองแบบมอนติคาร์โล และทำการทดลองซ้ำ ๆ กัน 10,000 ครั้งในแต่ละสถานการณ์ เพื่อคำนวณค่าเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ผลการวิจัยสรุปได้ดังนี้

สำหรับทุกระดับของขนาดตัวอย่าง และทุกระดับของร้อยละของค่าสูญหาย วิธี RM ให้ค่า $|Bias|$ และ MSE ต่ำที่สุด เมื่อค่า ρ มีค่าน้อย ($\approx$ 0.1 ถึง 0.4) วิธี RMD ให้ค่า $|Bias|$ และ MSE ต่ำที่สุด เมื่อค่า ρ มีค่าปานกลาง ($\approx$ 0.5 ถึง 0.7) และ วิธี IRMD ให้ค่า $|Bias|$ และ MSE ต่ำที่สุด เมื่อค่า ρ มีค่ามาก ($\approx$ 0.8 ถึง 0.9)

คำสำคัญ: การประมาณค่าพารามิเตอร์ ตัวแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง ค่ามัธยฐานเวียนเกิด ค่าสูญหาย

Abstract

The objective of this research is to develop and to compare the parameter (ρ) estimation methods in first-order autoregressive model with missing values. The methods are recursive mean OLS method (RM), recursive median OLS method (RMD), and improved recursive median OLS method (IRMD). The percentages of missing values are 5% and 10%. The sample sizes are 25, 50, 100, and 250. This research uses the Monte

Carlo simulation method. The experiment was repeated 10,000 times for each condition to calculate the absolute of bias ($|Bias|$) and the mean squared error (MSE). Results of the research are as follows:

For all sample sizes and all percentages of missing values, the $|Bias|$ and the MSE of an RM method are the lowest when ρ is small (≈ 0.1 to 0.4). The $|Bias|$ and the MSE of an RMD method are the lowest when ρ is moderate (≈ 0.5 to 0.7). Finally, the $|Bias|$ and the MSE of an IRMD method are the lowest when ρ is large (≈ 0.8 to 0.9).

Keywords: Parameter Estimation, First-Order Autoregressive Model, Recursive Median, Missing Values

1. บทนำ

การพยากรณ์อนุกรมเวลา (Time series forecasting) เป็นวิธีการพยากรณ์วิธีหนึ่งที่นิยมใช้กันมาก วิธีนี้จะใช้ข้อมูลในอดีต โดยจะศึกษาถึงลักษณะการเปลี่ยนแปลงของข้อมูลเมื่อเวลาเปลี่ยนไปว่ามีลักษณะเช่นไร และทำการกำหนดรูปแบบของการแปรเปลี่ยนที่เหมาะสมที่สุดสำหรับข้อมูลนั้น โดยจะอยู่ในรูปของความสัมพันธ์กับเวลา การพยากรณ์อนุกรมเวลามีอยู่ด้วยกันหลายวิธี เช่น เทคนิคการทำให้เรียบ (Smoothing techniques) การกรองแบบปรับได้ (Adaptive filtering) วิธีอนุกรมเวลาแบบคลาสสิค (Classical time series methods) และวิธีอนุกรมเวลาแบบบอกซ์-เจนกินส์ (Box-Jenkins methods) เป็นต้น

นอกเหนือจากการเลือกใช้วิธีการพยากรณ์ที่เหมาะสม และการกำหนดตัวแบบอนุกรมเวลาแล้ว วิธีการประมาณค่าพารามิเตอร์ของตัวแบบอนุกรมเวลาก็เป็นอีกปัจจัยหนึ่งที่ส่งผลถึงความแม่นยำและเชื่อถือได้ของค่าพยากรณ์ [1] วิธีการประมาณค่าพารามิเตอร์ของตัวแบบอนุกรมเวลามีอยู่หลายวิธี เช่น วิธีกำลังสองน้อยที่สุด (Ordinary least squares method: OLS) วิธีการประมาณความควรจะเป็นสูงสุด (Maximum likelihood estimation method: MLE) เป็นต้น โดยแต่ละวิธีการจะมีหลักการในการหาตัวประมาณค่าแตกต่างกันไป อีกทั้งยังให้ผลลัพธ์ที่แตกต่างกันออกไปด้วย ซึ่งวิธีการประมาณค่าพารามิเตอร์นั้นได้มีผู้ศึกษาจากอดีตถึงปัจจุบันภายใต้ข้อสมมติเบื้องต้นที่แตกต่างกันออกไป อาทิ ในปี ค.ศ. 1999 โช และชิน [2] ได้พัฒนาวิธีกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิด (Recursive Mean OLS method) สำหรับตัวแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง (First-order autoregressive Model: AR(1)) ในปี ค.ศ. 2000 วาวกัส [3] ได้พัฒนาวิธีการประมาณจีเอ็มเอ็ม (Generalized method of moment method) สำหรับตัวแบบ AR(1) ในปี ค.ศ. 2004 สอาด [4] ได้ปรับปรุงวิธีการประมาณจีเอ็มเอ็มของวาวกัส โดยใช้แนวคิดของโชและชิน [2]

เมื่อเลือกวิธีการประมาณค่าพารามิเตอร์ที่เหมาะสมแล้ว ข้อมูลในอดีตก็เป็นอีกปัจจัยหนึ่งที่ส่งผลถึงความถูกต้องแม่นยำและความเชื่อถือได้ของค่าพยากรณ์ เนื่องจากการวิเคราะห์อนุกรมเวลาจะอาศัยข้อมูลในอดีตที่ผ่านมา แต่ผู้วิเคราะห์อาจประสบปัญหาเกี่ยวกับข้อมูล เช่น ข้อมูลบางค่าสูญหาย (Missing values) ทำให้ผลการวิเคราะห์และการพยากรณ์คลาดเคลื่อน หรืออาจนำไปสู่การตัดสินใจหรือวางแผนที่ผิดพลาดได้ ดังนั้นผู้วิจัยจึงพัฒนาวิธีการประมาณค่าพารามิเตอร์ของตัวแบบ

อัตตสหสัมพันธ์อันดับที่หนึ่ง (AR(1)) เมื่อข้อมูลมีค่าสูญหาย ซึ่งวิธีการที่พัฒนาขึ้นใหม่มี 2 วิธี ได้แก่ วิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิด (Recursive median OLS method) และวิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิดปรับปรุง (Improved recursive median OLS method) นอกจากนั้นจะเปรียบเทียบประสิทธิภาพของวิธีการที่พัฒนาขึ้นใหม่กับวิธีกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิด (Recursive mean OLS method) ซึ่งเสนอโดยโซ และชิน [2]

ข้อมูลอนุกรมเวลาที่ศึกษาในครั้งนี้เป็นอนุกรมเวลาตัวแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง AR(1) เมื่อข้อมูลมีค่าสูญหาย จากนั้นทำการเปรียบเทียบประสิทธิภาพของทุกวิธีด้วยค่าเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean square error: MSE) ของตัวประมาณพารามิเตอร์

2. ขอบเขตของการวิจัย

2.1 อนุกรมเวลา $\{Y_t\}$ เป็นอนุกรมเวลาตัวแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง AR(1) เขียนตัวแบบได้ดังนี้ [5]

$$Y_t - \mu \quad = \quad \rho(Y_{t-1} - \mu) + a_t$$

ในที่นี้กำหนดค่าเฉลี่ยของอนุกรมเวลา (μ) เท่ากับ 0

2.2 กำหนดค่าพารามิเตอร์ (ρ) ของตัวแบบ 9 ระดับ คือ 0.1 ถึง 0.9 เพิ่มขึ้นครั้งละ 0.1

2.3 การแจกแจงของความคลาดเคลื่อนสุ่ม a_t มีการแจกแจงแบบปกติ (Normal distribution) โดยมีฟังก์ชันความหนาแน่นของความน่าจะเป็น

$$f(a) \quad = \quad \frac{1}{\sigma_a\sqrt{2\pi}}\exp(-\frac{1}{2\sigma_a^2}(a-\mu_a)^2)$$

ในที่นี้กำหนดให้ $\mu_a = 0$ และ $\sigma_a = 1$

2.4 ขนาดตัวอย่าง (n) ที่ใช้ในการศึกษา มี 4 ระดับ คือ 25, 50, 100 และ 250

2.5 ร้อยละของค่าสูญหาย (p) เท่ากับ ร้อยละ 5 และ 10 และตำแหน่งของค่าสูญหายเป็นไปอย่างสุ่ม ซึ่งอยู่ระหว่างเวลาที่ t = 2 ถึง t = n-1

2.6 โปรแกรมที่ใช้ในการวิจัยครั้งนี้ทั้งหมด เขียนด้วยโปรแกรม R เวอร์ชัน 2.8.0 ซึ่งทำการทดลองซ้ำ 10,000 ครั้งในแต่ละสถานการณ์ของการทดลอง

3. วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้มีวิธีดำเนินการวิจัยดังนี้

3.1 จำลองค่าความคลาดเคลื่อนสุ่ม a_t

การสร้างการแจกแจงของความคลาดเคลื่อนสุ่มที่ต้องการศึกษา มีการแจกแจงแบบปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

3.2 จำลองข้อมูลอนุกรมเวลา AR(1) ที่มีค่าสูญหาย

สร้าง Y_0 ให้มีการแจกแจงแบบปกติ ซึ่งมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ $\frac{\sigma_a^2}{1-\rho^2} = \frac{1}{1-\rho^2}$ และสร้าง a_t ; t = 1, 2, …, n ให้มีการแจกแจงแบบปกติ ซึ่งมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 จากนั้นสร้าง Y_t; t = 1, 2, …, n ให้มีรูปแบบความสัมพันธ์คือ

$$Y_t = \rho Y_{t-1} + a_t$$

เมื่อสร้างข้อมูลอนุกรมเวลา AR(1) แล้ว ต่อไปจะทำการสร้างข้อมูลอนุกรมเวลา AR(1) ที่มีค่าสูญหาย โดยสุ่มเลือกตำแหน่งของค่าสูญหายระหว่างเวลาที่ t = 2 ถึง t = n-1 จากนั้นบันทึกค่า Y_t ในคาบเวลานั้นให้เป็นค่าสูญหาย

3.3 ประมาณค่าสูญหาย

สมมติว่า Y_T เป็นค่าสูญหาย ประมาณค่า Y_T ด้วย $E(Y_T \mid Y_1, Y_2, ..., Y_{T-1}) = Y_{T-1}$ [6] นั่นคือ ค่าประมาณของ Y_T คือ Y_{T-1}

3.4 ประมาณค่าพารามิเตอร์ของข้อมูลอนุกรมเวลาทั้ง 3 วิธี

3.4.1 วิธีกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิด โซ และชิน [2] ได้พัฒนาวิธีกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิด โดยมีหลักการเช่นเดียวกับวิธีกำลังสองน้อยที่สุด คือการทำให้ผลรวมค่ากำลังสองของความคลาดเคลื่อนภายใต้พารามิเตอร์ของตัวแบบมีค่าต่ำที่สุด แต่วิธีกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิดจะประมาณค่าเฉลี่ยของอนุกรมเวลา (μ) ด้วยค่าเฉลี่ยแบบเวียนเกิด (Recursive mean: $\bar{Y}_t$) แทนค่าเฉลี่ยเลขคณิต (Arithmetic mean: $\bar{Y}$) ตัวประมาณกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิด คือ

$$\hat{\rho}_{RM} = \frac{\sum_{t=2}^{n}(Y_t - \bar{Y}_{t-1})(Y_{t-1} - \bar{Y}_{t-1})}{\sum_{t=2}^{n}(Y_{t-1} - \bar{Y}_{t-1})^2}$$

โดยที่ $\bar{Y}_t = \frac{\sum_{i=1}^{t} Y_i}{t}$

3.4.2 วิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิด ผู้วิจัยได้พัฒนาวิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิด โดยใช้แนวคิดของโซ และชิน [2] กล่าวคือ วิธีนี้มีหลักการเช่นเดียวกับวิธีกำลังสองน้อยที่สุด และประมาณค่าเฉลี่ยของอนุกรมเวลา (μ) ด้วยค่ามัธยฐานเวียนเกิด (Recursive median: $\tilde{Y}_t$) แทนค่าเฉลี่ยเลขคณิต (Arithmetic mean: $\bar{Y}$) ตัวประมาณกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิด คือ

$$\hat{\rho}_{RMD} = \frac{\sum_{t=2}^{n}(Y_t - \tilde{Y}_t)(Y_{t-1} - \tilde{Y}_{t-1})}{\sum_{t=2}^{n}(Y_{t-1} - \tilde{Y}_{t-1})^2}$$

โดยที่ $\tilde{Y}_t = median\{Y_1, Y_2, ..., Y_t\}$

3.4.3 วิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิดปรับปรุง ผู้วิจัยมีแนวคิดที่ปรับปรุงประสิทธิภาพของวิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิด โดยนำค่ามัธยฐานเวียนเกิด $\tilde{Y}_t$ มาหาค่าเฉลี่ยแบบเวียนเกิดอีกครั้งหนึ่ง ซึ่งตัวประมาณกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิดปรับปรุง คือ

$$\hat{\rho}_{IRMD} = \frac{\sum_{t=2}^{n}(Y_t - \ddot{Y}_t)(Y_{t-1} - \ddot{Y}_{t-1})}{\sum_{t=2}^{n}(Y_{t-1} - \ddot{Y}_{t-1})^2}$$

โดยที่ $\ddot{Y}_t = \frac{\sum_{i=1}^{t} \tilde{Y}_i}{t}$

และ $\tilde{Y}_t = median\{Y_1, Y_2, ..., Y_t\}$

3.5 คำนวณค่าเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณพารามิเตอร์ แล้วทำการเปรียบเทียบ

$$MSE = E(\hat{\rho} - \rho)^2$$

$$= \frac{\sum_{m=1}^{M}((\hat{\rho}_K)_m - \bar{\hat{\rho}}_K)^2}{M-1} + Bias^2$$

$$Bias = \frac{\sum_{m=1}^{M}(\hat{\rho}_K)_m}{M} - \rho$$

เมื่อ $\hat{\rho}_K$ แทน ตัวประมาณค่าพารามิเตอร์

K แทน วิธีการประมาณค่าพารามิเตอร์ ซึ่งในที่นี้คือ RM, RMD และ IRMD

M แทน จำนวนรอบของการทำซ้ำ

4. ผลการวิจัย

ในการนำเสนอผลการวิจัยเพื่อความสะดวก จะใช้สัญลักษณ์ต่าง ๆ แทนความหมาย ดังนี้

RM หมายถึง วิธีกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิด

RMD หมายถึง วิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิด

IRMD หมายถึง วิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิดปรับปรุง

$|Bias|$ หมายถึง ค่าเอนเอียงสัมบูรณ์

MSE หมายถึง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย

* หมายถึง วิธีการประมาณค่าพารา - มิเตอร์ที่ให้ค่า $|Bias|$ ต่ำที่สุด หรือค่า MSE ต่ำที่สุด

ผลการเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ ทั้ง 3 วิธี นำเสนอดังตารางที่ 1-2 ในที่นี้จะพิจารณาค่าเอนเอียงสัมบูรณ์ และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ดังนี้

4.1 ค่าเอนเอียงสัมบูรณ์

ผลการเปรียบเทียบค่าเอนเอียงสัมบูรณ์ จำแนกตามขนาดตัวอย่าง สรุปรายละเอียดได้ดังนี้

<u>เมื่อขนาดตัวอย่างเท่ากับ 25</u>

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 5 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าเท่ากับ 0.1 และ 0.2 วิธี RM ให้ค่า $|Bias|$ ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าตั้งแต่ 0.3 ถึง 0.5 วิธี RMD ให้ค่า $|Bias|$ ต่ำที่สุด และกรณีที่ค่า ρ มีค่าตั้งแต่ 0.6 ถึง 0.9 วิธี IRMD ให้ค่า $|Bias|$ ต่ำที่สุด

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 10 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.3 วิธี RM ให้ค่า $|Bias|$ ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าเท่ากับ 0.4 และ 0.5 วิธี RMD ให้ค่า $|Bias|$ ต่ำที่สุด และกรณีที่ค่า ρ มีค่าตั้งแต่ 0.6 ถึง 0.9 วิธี IRMD ให้ค่า $|Bias|$ ต่ำที่สุด

<u>เมื่อขนาดตัวอย่างเท่ากับ 50</u>

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 5 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.3 วิธี RM ให้ค่า $|Bias|$ ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าตั้งแต่ 0.4 ถึง 0.6 วิธี RMD ให้ค่า $|Bias|$ ต่ำที่สุด และกรณีที่ค่า ρ มีค่าตั้งแต่ 0.7 ถึง 0.9 วิธี IRMD ให้ค่า $|Bias|$ ต่ำที่สุด

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 10 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.5 วิธี RM ให้ค่า $|Bias|$ ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าเท่ากับ 0.6 และ 0.7 วิธี RMD ให้ค่า $|Bias|$ ต่ำที่สุด และกรณีที่ค่า ρ มีค่าเท่ากับ 0.8 และ 0.9 วิธี IRMD ให้ค่า $|Bias|$ ต่ำที่สุด

<u>เมื่อขนาดตัวอย่างเท่ากับ 100</u>

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 5 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.5 วิธี RM ให้ค่า $|Bias|$ ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าเท่ากับ 0.6 และ 0.7 วิธี RMD ให้ค่า $|Bias|$ ต่ำที่สุด และกรณีที่ค่า ρ มีค่าเท่ากับ 0.8 และ 0.9 วิธี IRMD ให้ค่า $|Bias|$ ต่ำที่สุด

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 10 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.6 วิธี RM ให้ค่า $|Bias|$ ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าเท่ากับ 0.7 และ 0.8 วิธี RMD ให้ค่า $|Bias|$ ต่ำที่สุด และกรณีที่ค่า ρ มีค่าเท่ากับ 0.9 วิธี IRMD ให้ค่า $|Bias|$ ต่ำที่สุด

<u>เมื่อขนาดตัวอย่างเท่ากับ 250</u>

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 5 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.6 วิธี RM ให้ค่า $|Bias|$ ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าตั้งแต่ 0.7 ถึง 0.9 วิธี RMD ให้ค่า $|Bias|$ ต่ำที่สุด

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 10 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.7 วิธี RM ให้ค่า $|Bias|$ ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าเท่ากับ 0.8 และ 0.9 วิธี RMD ให้ค่า $|Bias|$ ต่ำที่สุด

4.2 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย

ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ย จำแนกตามขนาดตัวอย่าง สรุปรายละเอียดได้ดังนี้

<u>เมื่อขนาดตัวอย่างเท่ากับ 25</u>

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 5 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าเท่ากับ 0.1 วิธี RM ให้ค่า MSE ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าตั้งแต่ 0.2 ถึง 0.5 วิธี RMD ให้ค่า MSE ต่ำที่สุด และกรณีที่ค่า ρ มีค่าตั้งแต่ 0.6 ถึง 0.9 วิธี IRMD ให้ค่า MSE ต่ำที่สุด

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 10 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าเท่ากับ 0.1 และ 0.2 วิธี RM ให้ค่า MSE ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าตั้งแต่ 0.3 ถึง 0.5 วิธี RMD ให้ค่า MSE ต่ำที่สุด และกรณีที่ค่า ρ มีค่าตั้งแต่ 0.6 ถึง 0.9 วิธี IRMD ให้ค่า MSE ต่ำที่สุด

<u>เมื่อขนาดตัวอย่างเท่ากับ 50</u>

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 5 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.3 วิธี RM ให้ค่า MSE ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าตั้งแต่ 0.4 ถึง 0.7 วิธี RMD ให้ค่า MSE ต่ำที่สุด และกรณีที่ค่า ρ มีค่าเท่ากับ 0.8 และ 0.9 วิธี IRMD ให้ค่า MSE ต่ำที่สุด

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 10 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.5 วิธี RM ให้ค่า MSE ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าตั้งแต่ 0.6 ถึง 0.8 วิธี RMD ให้ค่า MSE ต่ำที่สุด และกรณีที่ค่า ρ มีค่าเท่ากับ 0.9 วิธี IRMD ให้ค่า MSE ต่ำที่สุด

<u>เมื่อขนาดตัวอย่างเท่ากับ 100</u>

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 5 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.6 วิธี RM ให้ค่า MSE ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าเท่ากับ 0.7 และ 0.8 วิธี RMD ให้ค่า MSE ต่ำที่สุด และกรณีที่ค่า ρ มีค่าเท่ากับ 0.9 วิธี IRMD ให้ค่า MSE ต่ำที่สุด

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 10 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.6 วิธี RM ให้ค่า MSE ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าเท่ากับ 0.7 และ 0.8 วิธี RMD ให้ค่า MSE ต่ำที่สุด และกรณีที่ค่า ρ มีค่าเท่ากับ 0.9 วิธี IRMD ให้ค่า MSE ต่ำที่สุด

<u>เมื่อขนาดตัวอย่างเท่ากับ 250</u>

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 5 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.6 วิธี RM ให้ค่า MSE ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าเท่ากับ 0.7 และ 0.8 วิธี RMD ให้ค่า MSE ต่ำที่สุด และกรณีที่ค่า ρ มีค่าเท่ากับ 0.9 วิธี IRMD ให้ค่า MSE ต่ำที่สุด

เมื่อร้อยละของค่าสูญหาย เท่ากับ ร้อยละ 10 สรุปได้ดังนี้ ในกรณีที่ค่า ρ มีค่าตั้งแต่ 0.1 ถึง 0.8 วิธี RM ให้ค่า MSE ต่ำที่สุด ส่วนกรณีที่ค่า ρ มีค่าเท่ากับ 0.9 วิธี RMD ให้ค่า MSE ต่ำที่สุด

5. สรุปผลการวิจัย

ผลการวิจัยครั้งนี้ เมื่อพิจารณาโดยรวมสรุปได้ดังนี้

วิธีกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิด (วิธี RM) ให้ค่า $|Bias|$ และ MSE ต่ำที่สุด เมื่อค่า ρ มีค่าน้อย ($\approx$ 0.1 ถึง 0.4)

วิธีกำลังสองน้อยที่สุดแบบมัยฐานเวียนเกิด (วิธี RMD) ให้ค่า $|Bias|$ และ MSE ต่ำที่สุด เมื่อค่า ρ มีค่าปานกลาง ($\approx$ 0.5 ถึง 0.7)

วิธีกำลังสองน้อยที่สุดแบบมัยฐานเวียนเกิดปรับปรุง (วิธี IRMD) ให้ค่า $|Bias|$ และ MSE ต่ำที่สุด เมื่อค่า ρ มีค่ามาก ($\approx$ 0.8 ถึง 0.9)

6. ข้อเสนอแนะ

จากผลการวิจัยพบว่าวิธีการประมาณค่าพารามิเตอร์ที่ให้ค่า $|Bias|$ และ MSE ต่ำที่สุดจะขึ้นอยู่กับระดับของพารามิเตอร์ แต่ในทางปฏิบัติเราจะไม่ทราบค่าพารามิเตอร์ ก่อนทำการประมาณค่าพารามิเตอร์ได้ ดังนั้นเพื่อให้สามารถเลือกวิธีการประมาณค่าพารามิเตอร์ที่เหมาะสมได้ ควรทำการประมาณค่าพารามิเตอร์เบื้องต้น (Preliminary Estimation or Initial Estimation) ซึ่งจะทำให้ทราบค่าประมาณพารามิเตอร์เบื้องต้น จากนั้นสามารถเลือกใช้วิธีการประมาณค่าพารามิเตอร์ได้อย่างเหมาะสมและมีประสิทธิภาพในสถานการณ์ต่าง ๆ ต่อไปได้ ซึ่งสูตรในการประมาณค่าพารามิเตอร์เบื้องต้นโดยวิธีกำลังสองน้อยที่สุด มีดังนี้

$$\hat{\rho} = \frac{\sum_{t=2}^{n}(Y_t-\bar{Y})(Y_{t-1}-\bar{Y})}{\sum_{t=2}^{n}(Y_{t-1}-\bar{Y})^2}$$

ตารางที่ 1 ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณค่าพารามิเตอร์ของตัวแบบ AR(1) เมื่อร้อยละของค่าสูญหาย (p) เท่ากับร้อยละ 5

n	ρ	$\|Bias\|$			MSE		
		RM	RMD	IRMD	RM	RMD	IRMD
25	0.1	0.0210*	0.0444	0.1354	0.0415*	0.0423	0.0653
	0.2	0.0075*	0.0270	0.1164	0.0401	0.0397*	0.0579
	0.3	0.0143	0.0002*	0.0899	0.0396	0.0384*	0.0501
	0.4	0.0261	0.0158*	0.0726	0.0386	0.0373*	0.0441
	0.5	0.0437	0.0368*	0.0490	0.0371	0.0358*	0.0373
	0.6	0.0621	0.0588	0.0226*	0.0354	0.0346	0.0302*
	0.7	0.0764	0.0759	0.0011*	0.0351	0.0350	0.0266*
	0.8	0.0969	0.0978	0.0283*	0.0365	0.0366	0.0238*
	0.9	0.1157	0.1191	0.0572*	0.0372	0.0385	0.0225*
50	0.1	0.0261*	0.0445	0.1009	0.0206*	0.0219	0.0355
	0.2	0.0164*	0.0315	0.0906	0.0196*	0.0203	0.0325
	0.3	0.0030*	0.0155	0.0773	0.0187*	0.0188	0.0285
	0.4	0.0086	0.0011*	0.0648	0.0173	0.0173*	0.0248
	0.5	0.0180	0.0110*	0.0526	0.0165	0.0164*	0.0214
	0.6	0.0279	0.0231*	0.0402	0.0152	0.0150*	0.0176
	0.7	0.0368	0.0335	0.0282*	0.0137	0.0137*	0.0139

ตารางที่ 1 ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณค่าพารามิเตอร์ของตัวแบบ AR(1) เมื่อร้อยละของค่าสูญหาย (p) เท่ากับร้อยละ 5 (ต่อ)

n	ρ	$\lvert Bias\rvert$			MSE		
		RM	RMD	IRMD	RM	RMD	IRMD
	0.8	0.0479	0.0462	0.0114*	0.0124	0.0124	0.0100*
	0.9	0.0600	0.0594	0.0087*	0.0118	0.0120	0.0071*
100	0.1	0.0386*	0.0515	0.0945	0.0115*	0.0128	0.0216
	0.2	0.0278*	0.0389	0.0833	0.0105*	0.0114	0.0191
	0.3	0.0171*	0.0259	0.0712	0.0094*	0.0097	0.0161
	0.4	0.0061*	0.0136	0.0614	0.0084*	0.0087	0.0140
	0.5	0.0009*	0.0051	0.0529	0.0076*	0.0077	0.0119
	0.6	0.0083	0.0041*	0.0439	0.0068*	0.0068	0.0096
	0.7	0.0155	0.0122*	0.0331	0.0058	0.0058*	0.0073
	0.8	0.0231	0.0208	0.0206*	0.0051	0.0051*	0.0051
	0.9	0.0289	0.0275	0.0066*	0.0038	0.0038	0.0027*
250	0.1	0.0380*	0.0453	0.0720	0.0055*	0.0061	0.0101
	0.2	0.0291*	0.0350	0.0625	0.0046*	0.0051	0.0085
	0.3	0.0206*	0.0256	0.0539	0.0040*	0.0043	0.0073
	0.4	0.0131*	0.0173	0.0462	0.0035*	0.0037	0.0062
	0.5	0.0067*	0.0103	0.0395	0.0031*	0.0032	0.0052
	0.6	0.0023*	0.0051	0.0343	0.0025*	0.0026	0.0042
	0.7	0.0043	0.0021*	0.0265	0.0021	0.0021*	0.0031
	0.8	0.0077	0.0060*	0.0200	0.0016	0.0016*	0.0021
	0.9	0.0115	0.0103*	0.0106	0.0011	0.0011	0.0011*

ตารางที่ 2 ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณค่าพารามิเตอร์ของตัวแบบ AR(1) เมื่อร้อยละของค่าสูญหาย (p) เท่ากับร้อยละ 10

n	ρ	$\lvert Bias\rvert$			MSE		
		RM	RMD	IRMD	RM	RMD	IRMD
25	0.1	0.0508*	0.0747	0.1648	0.0437*	0.0458	0.0739
	0.2	0.0320*	0.0517	0.1413	0.0421*	0.0430	0.0649
	0.3	0.0091*	0.0237	0.1131	0.0385	0.0378*	0.0529
	0.4	0.0099	0.0009*	0.0880	0.0360	0.0350*	0.0449
	0.5	0.0298	0.0232*	0.0606	0.0348	0.0339*	0.0367
	0.6	0.0507	0.0474	0.0334*	0.0346	0.0338	0.0313*
	0.7	0.0722	0.0706	0.0053*	0.0340	0.0332	0.0256*
	0.8	0.0879	0.0897	0.0201*	0.0328	0.0332	0.0216*
	0.9	0.1139	0.1175	0.0554*	0.0361	0.0373	0.0215*

ตารางที่ 2 ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณค่าพารามิเตอร์ของตัวแบบ AR(1) เมื่อร้อยละของค่าสูญหาย (p) เท่ากับร้อยละ 10 (ต่อ)

n	ρ	$\|Bias\|$			MSE		
		RM	RMD	IRMD	RM	RMD	IRMD
50	0.1	0.0748*	0.0938	0.1503	0.0264*	0.0297	0.0484
	0.2	0.0535*	0.0700	0.1302	0.0221*	0.0239	0.0407
	0.3	0.0323*	0.0456	0.1073	0.0198*	0.0208	0.0343
	0.4	0.0142*	0.0246	0.0886	0.0174*	0.0178	0.0282
	0.5	0.0006*	0.0078	0.0733	0.0156*	0.0157	0.0233
	0.6	0.0168	0.0112*	0.0524	0.0141	0.0141*	0.0180
	0.7	0.0291	0.0255*	0.0358	0.0127	0.0125*	0.0136
	0.8	0.0432	0.0418	0.0164*	0.0118	0.0118*	0.0098
	0.9	0.0563	0.0558	0.0053*	0.0112	0.0113	0.0069*
100	0.1	0.0782*	0.0914	0.1348	0.0163*	0.0187	0.0312
	0.2	0.0590*	0.0702	0.1152	0.0131*	0.0146	0.0252
	0.3	0.0426*	0.0519	0.0982	0.0106*	0.0115	0.0205
	0.4	0.0280*	0.0353	0.0832	0.0090*	0.0095	0.0169
	0.5	0.0125*	0.0186	0.0660	0.0077*	0.0079	0.0132
	0.6	0.0005*	0.0051	0.0519	0.0066*	0.0067	0.0103
	0.7	0.0102	0.0068*	0.0391	0.0055	0.0055*	0.0075
	0.8	0.0206	0.0183*	0.0223	0.0047	0.0047*	0.0049
	0.9	0.0280	0.0266	0.0076*	0.0036	0.0036	0.0026*
250	0.1	0.0810*	0.0884	0.1152	0.0106*	0.0119	0.0183
	0.2	0.0634*	0.0695	0.0966	0.0078*	0.0087	0.0140
	0.3	0.0472*	0.0526	0.0810	0.0058*	0.0064	0.0109
	0.4	0.0334*	0.0380	0.0670	0.0044*	0.0047	0.0084
	0.5	0.0211*	0.0248	0.0538	0.0033*	0.0034	0.0063
	0.6	0.0113*	0.0142	0.0432	0.0026*	0.0027	0.0048
	0.7	0.0022*	0.0045	0.0326	0.0020*	0.0020	0.0034
	0.8	0.0046	0.0028*	0.0232	0.0015*	0.0015	0.0022
	0.9	0.0100	0.0088*	0.0120	0.0010	0.0010*	0.0010

7. เอกสารอ้างอิง

[1] วราฤทธิ์ พานิชกิจโกศลกุล, การเปรียบเทียบการประมาณค่าพารามิเตอร์ของตัวแบบอนุกรมเวลา, วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ, 2545.

[2] So, B.S. and Shin, D.W., Recursive Mean Adjustment in Time Series Inferences, Statistics and Probability Letters, Vol. 43, pp.65-73, 1999.

[3] Vougas, D.V., A Comparison of LS/ML and GMM Estimation in a Sample AR(1) Model, Communications in Statistics: Simulation and Computation, Vol. 29, No.1, pp.239-258, 2000.

[4] Niwitpong, S., Improved GMM Estimator of AR(1) Process Near Unit Root, Journal of Applied Science-King Mongkut's Institute of Technology North Bangkok, Vol. 3, No.1, pp.21-28, 2004.

[5] Wei, W.W.S., Time Series Analysis: Univariate and Multivariate Methods, Boston, Addison-Wesley, 2006.

[6] Shin, D.W. and Sarkar, S., Testing for a Unit Root in an AR(1) Time Series Using Irregularly Observed Data, Journal of Time Series Analysis, Vol. 17, pp.309-321, 1996.