บทที่ 2

กรอบแนวคิดทางทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 กรอบแนวคิดทางทฤษฎี

2.1.1 ARIMA Model

แบบจำลอง Autoregressive Integrated Moving Average (ARIMA Model) ได้ถูกพัฒนาขึ้นมาโดยมีคุณสมบัติคือ

1. การจัดจำแนกชนิดและกระบวนการประมาณค่าอย่างมีประสิทธิภาพ
2. การครอบคลุมไปถึงผลลัพธ์ที่ได้รวบรวมเอาอนุกรมเวลาเชิงฤดูกาล
3. การขยายขอบเขตไปเพื่อรวมเอากระบวนการหรือระบบไม่นิ่งเข้าไว้ด้วย

ข้อมูลอนุกรมเวลาส่วนใหญ่แล้วจะมีลักษณะไม่นิ่ง(nonstationary) และลักษณะของ AR และ MA ของแบบจำลอง ARIMA จะหมายถึงข้อมูลอนุกรมเวลาที่มีลักษณะนิ่ง(stationary) ในกรณีที่ข้อมูลมีลักษณะนิ่ง แบบจำลอง ARIMA(p,d,q) จะมีรูปแบบเป็น ARIMA(p,0,q) ซึ่งก็คือ ARMA นั่นคือ AR(p) และ MA(q) สำหรับกรณีของ AR(1) และ MA(1) นั้นสามารถเขียนได้ในรูป ARIMA(1,0,1) ซึ่งเขียนเป็นสมการได้ดังนี้

$$Y_t = \mu' + \phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1} \tag{2.1a}$$

หรือ

$$\underbrace{(1-\phi_1 L)}_{\text{AR(1)}} Y_t = \mu' + \underbrace{(1-\theta_1 L)}_{\text{MA(1)}} e_t \quad ; \text{ โดย L คือ lag operator} \tag{2.1b}$$

สำหรับกรณีของ AR(1) และ MA(1) โดยที่ข้อมูลนั้นมีลักษณะไม่นิ่งและถ้าทำ 1st – different ข้อมูลที่ได้จากผลต่างมีลักษณะนิ่ง รูปแบบของ ARIMA(1,1,1) จึงสามารถเขียนเป็นสมการได้ดังนี้

$$(1-L)(1-\phi_1 L)Y_t = \mu' + (1-\theta_1 L)e_t \qquad (2.2a)$$

หรือ

$$[1-L(1+\phi_1)+\phi_1 L_2]Y_t = \mu' + e_t - \theta_t e_{t-1}$$

$$Y_t = (1+\phi_1)Y_{t-1} - \phi_1 Y_{t-2} + \mu' + e_t - \theta_1 e_{t-1} \qquad (2.2b)$$

2.1.2 Long Memory

แม้ว่าอนุกรมทางเศรษฐศาสตร์โดยส่วนใหญ่จะเป็น nonstationary และต้องทำผลต่าง (difference)ข้อมูล แต่แท้จริงแล้วไม่มีความจำเป็นที่จะต้องใช้ 1st-difference แล้วจากนั้นการใช้แบบจำลอง ARIMA จึงจะเป็นแบบจำลองที่ได้แก้ไขแล้วที่ดีที่สุดเสมอไป ซึ่งในการวิเคราะห์ของ Box and Jenskins(1970) ได้มีการสมมุติว่าถ้าอนุกรมเป็น nonstationary แล้ว 1st-difference จะแสดงคุณสมบัติที่ดีออกมา โดยที่ไม่มีปัจจัยใด ๆ ที่เกี่ยวกับฤดูกาลเข้ามากระทบ

อาจกล่าวได้ว่าอนุกรมเวลา $\{y_t\}$ ที่มีความ stationary จะมี long memory ถ้า d คือ fractional integration หรือ fractional difference parameter ไม่เท่ากับศูนย์ แล้ว $d \in (-0.5, 0.5)$ จนกระทั่ง spectral density ได้เป็นไปตาม power law นั่นคือ $f(\lambda) \sim k\lambda^{-2d}$ เมื่อ $\lambda \rightarrow 0^+$ ดังนั้นเมื่อ $\lambda \rightarrow 0$ แล้ว $f(\lambda)$ ถ้าไม่โน้มเอียงเข้าสู่ ∞ (ถ้า $d > 0$) ก็โน้มเอียงเข้าสู่ 0 (ถ้า $d < 0$) แต่ถ้า $d = 0$ กล่าวได้ว่า $\{y_t\}$ มี short memory ในกรณีนี้ f(0) จะเป็นไปในเชิงบวกและเชิงอันตะ กระบวนการ ARMA ซึ่งสามารถสลับได้ที่นิ่งแล้วทั้งหมดนั้น จะมี short memory

โดยหนึ่งในผู้เสนอ long memory models ไว้สำหรับอนุกรมเวลาต่าง ๆ คือ D.R. Cox ผู้ซึ่งใช้แบบจำลองเพื่ออธิบายการผันผวนในเส้นผ่าศูนย์กลางเส้นด้ายที่ใช้ถักขณะที่ได้เป็นที่ปรึกษาในอุตสาหกรรมสิ่งทอมาก่อน แบบจำลองนั้นได้รับการคิดค้นมาแต่ดั้งเดิม และ ได้ถูกเผยแพร่และเป็นที่นิยมอย่างกว้างขวาง เมื่อไม่นานมานี้ Hosking(1981) ได้ทำการสร้างความสัมพันธ์ระหว่าง long memory และ fractional differencing ซึ่ง Hosking ได้เสนอแบบจำลองประเภทหนึ่งที่เรียกว่า fractional ARIMA ในกรณีซึ่ง degree of differencing สามารถเป็นตัวเลขจำนวนจริงใด ๆ โดยตัวอย่างที่ง่ายที่สุดของแบบจำลองเหล่าเรียกว่า fractionally integrated noise และยังเป็นที่ทราบว่าเป็น fractional ARIMA(0,d,0) model

เพื่อแสดงคุณสมบัติการมี long memory ที่เป็น stationary process ก็ต่อเมื่อฟังก์ชันอัตสหสัมพันธ์ได้แสดงคุณสมบัติ $\rho(k) \rightarrow C_\rho k^{-\alpha}$ เมื่อ $k \rightarrow \infty$ เมื่อ C_ρ คือ ค่าคงที่ในทางบวก และ α คือจำนวนจริงระหว่าง 0 และ 1 (Zivot and Wang, 2002)

ดังนั้นฟังก์ชันอัตสหสัมพันธ์ของกระบวนการ long memory จะมีอัตราการลดลงอย่างช้า ๆ มีลักษณะเป็นไฮเพอร์โบลิค ในความเป็นจริงนั้นจะลดลงช้าเป็นอย่างมากจน อัตสหสัมพันธ์นั้นไม่สามารถรวมกันได้ ดังสมการที่ (2.3)

$$\sum_{k=-\infty}^{\infty} \rho(k) = \infty \tag{2.3}$$

สำหรับ stationary process ฟังก์ชันอัตสหสัมพันธ์ประกอบไปด้วยรายละเอียดเดิมดัง spectral density เดิม โดยเฉพาะ spectral density ได้นิยามไว้ดังนี้

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho(k) e^{ik\omega} \tag{2.4}$$

เมื่อ ω คือ fourier frequency จาก $\rho(k) \rightarrow C_\rho k^{-\alpha}$ เมื่อ $k \rightarrow \infty$ สามารถแสดงให้เห็นได้ดังนี้ (Hamilton, 1994)

$$f(\omega) \rightarrow C_f \omega^{\alpha-1} \quad \text{เมื่อ} \quad \omega \rightarrow 0 \tag{2.5}$$

เมื่อ C_f เป็นค่าคงที่ในทางบวก ดังนั้นสำหรับกระบวนการ long memory ที่ spectral density มีแนวโน้มไปสู่ค่าอนันต์ที่ความถี่เป็นศูนย์ แทนที่จะใช้ α ในทางปฏิบัติใช้

$$H = 1 - \alpha/2 \in (0.5,1) \tag{2.6}$$

ซึ่งดังที่ทราบคือ hurst coefficient เพื่อวัดค่า long memory ใน y_t ยิ่งค่า H มีขนาดใหญ่มากเท่าใดจะทำให้กระบวนการที่นิ่งยิ่งเป็น long memory มากขึ้นเท่านั้น(Hurst, 1951)

โดยใช้คุณสมบัติของ $\rho(k) \rightarrow C_\rho k^{-\alpha}$ เมื่อ $k \rightarrow \infty$ และคุณสมบัติของ frequency domain $f(\omega) \rightarrow C_f \omega^{\alpha-1}$ เมื่อ $\omega \rightarrow 0$ Granger and Joyeux (1980) และ Hosking (1981)

แสดงให้เห็นถึงกระบวนการ long memory ในอนุกรมเวลา y_t ซึ่งสามารถทำให้เป็นแบบจำลองที่มีคุณสมบัติแบบพาราเมตริกซ์ได้ โดยการขยายกระบวนการรวมกันให้เป็น fractionally integrated process โดยเฉพาะการมี fractional integration ในอนุกรมเวลา y_t ตามสมการ (2.7) นี้

$$(1 - L)^d (y_t - \mu) = u_t \tag{2.7}$$

เมื่อ d คือ fractional integration หรือ fractional difference parameter , μ คือ การคาดการณ์หรือค่าเฉลี่ยของ y_t และ u_t คือ ตัวรบกวนเป็น short memory ที่มีความนิ่งโดยค่าเฉลี่ยเป็นศูนย์ ในทางปฏิบัติเมื่อข้อมูลอนุกรมเวลาที่ต้องทดสอบ ถ้าอนุกรมเวลามีลักษณะเป็นไม่นิ่ง เพราะ $d = 1$ และเมื่อทำข้อมูลอนุกรมเวลาให้เป็นผลต่างก็จะมีความนิ่ง อย่างไรก็ตามสำหรับอนุกรมเวลาทางเศรษฐศาสตร์และการเงินที่ได้มีการทดสอบ ปรากฏว่าผลต่างจำนวนเต็มอาจมีค่าเป็นจำนวนมากซึ่งแสดงได้ โดยที่ไม่มี spectral density ที่ความถี่เป็นศูนย์สำหรับอนุกรมเวลาผลต่าง เพื่อให้มี long memory และ หลีกเลี่ยงการใช้ผลต่างจำนวนเต็มของ y_t แล้วจะให้ d เป็น fractional ซึ่ง fractional difference fiter สำหรับค่า $d > -1$ ที่แท้จริงใด ๆ ถูกนิยามดังนี้ (Zivot and Wang, 2002)

$$(1 - L)^d = \sum_{k=0}^{\infty} \begin{pmatrix} d \\ k \end{pmatrix} (-1)^k L^k \tag{2.8}$$

และค่าสัมประสิทธิ์แบบ Binomial คือ

$$\begin{pmatrix} d \\ k \end{pmatrix} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} \tag{2.9}$$

ซึ่งสังเกตเห็นได้คือ fractional difference filter สามารถทำให้เทียบเท่ากับ infinite order autoregressive filter ซึ่งสามารถแสดงให้เห็นว่าเมื่อ $|d| > 0.5$ แล้ว y_t จะไม่นิ่ง นั่นคือถ้า $0 < d < 0.5$ แล้วอนุกรมเวลา y_t จะนิ่ง และมี long memory แต่ถ้า $-0.5 < d < 0$ แล้วอนุกรมเวลา y_t จะนิ่งและมี short memory และบางครั้ง d จะถูกอ้างอิงให้เป็นเสมือน anti-persistent

เมื่อ fractionally integrated series หรืออนุกรมเวลา y_t มี long memory สามารถแสดงได้ว่า

$$d = H - 0.5 \tag{2.10}$$

และดังนั้น d และ H สามารถถูกใช้สลับกันได้โดยการวัดค่าของ long memory Hosking (1981) ได้แสดงถึงคุณสมบัติการวัดค่า ในสมการ $\rho(k) \longrightarrow C_\rho k^{-\alpha}$ เมื่อ $k \longrightarrow \infty$ และ frequency domain ในสมการ $f(\omega) \longrightarrow C_f \omega^{\alpha-1}$ เมื่อ $\omega \rightarrow 0$ ซึ่งเพียงพอต่อเงื่อนไขเมื่อ $0 < d < 0.5$

2.1.3 Autoregressive Fractional Integrated Moving Average (ARFIMA) Model

แบบจำลอง ARFIMA หรือ FARIMAเป็นเครื่องมือที่เป็นประโยชน์ในการวิเคราะห์อนุกรมเวลาต่าง ๆ เช่น ดาราศาสตร์ การศึกษาเกี่ยวกับน้ำ วิทยาศาสตร์คอมพิวเตอร์ และอื่น ๆ มากมาย แบบจำลองนี้สามารถแสดงถึงลักษณะพิเศษของ “long-range dependence or positive memory” เมื่อ $0 < d < 0.5$ และ “intermediate or negative memory” เมื่อ $-0.5 < d < 0$ ซึ่งได้แสดงไว้ในงานของ Beran (1994) ที่ได้มีการประมาณค่าพารามิเตอร์ d ไว้ โดยแบ่งการประมาณออกเป็น 2 กลุ่ม คือ parametric และ semiparametric method

จากการศึกษาในบางเรื่องเกี่ยวกับการจำลองเหตุการณ์ และได้เปรียบเทียบเทคนิคต่างๆ ของการประมาณค่าใน long memory process ซึ่งอาจพบในผลงานของ Taqqu , Teverovsky and Willinger (1995) เป็นต้น โดยทีมงานเหล่านี้ส่วนใหญ่เกี่ยวข้องกับการประมาณค่าของ d เท่านั้น ทำการประมาณค่าพารามิเตอร์ทุกตัวด้วยแบบจำลอง ARFIMA(p,d,q) ด้วยการหาค่า AR และ MA

วิธีการที่เป็นธรรมเนียมปฏิบัติเพื่อสร้างแบบจำลองของอนุกรมเวลา y_t ที่ I(0) เมื่อแบบจำลอง ARIMA (p,d,q) คือ

$$\phi(L)(1-L)^d(y_t - u) = \theta(L)\in_t$$

เมื่อ $\phi(L)$ และ $\theta(L)$ คือ lag polynomials

$$\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i$$

$$\theta(L) = 1 - \sum_{j=1}^{q} \theta_j L^j$$

โดยที่มีรากอยู่นอกวงกลมหนึ่งหน่วย และ $\in_t$ สมมุติให้เป็นตัวแปรสุ่มแบบปกติโดยค่าเฉลี่ยและความแปรปรวนเท่ากับศูนย์

เมื่อแบบจำลอง ARFIMA มีลักษณะดังนี้

$$\phi(L)(1-L)^{\delta}[(1-L)^{m}y_{t}-\mu]=\theta(L)\epsilon_{t} \qquad (2.11)$$

โดยที่ $-0.5 < d < 0.5$, $\phi(L)$ และ $\theta(L)$ ได้นิยามไว้แล้วดังข้างต้นในแบบจำลอง ARIMA ค่า m คือ จำนวนครั้งของการทำผลต่าง y_t เพื่อให้เกิดความนิ่ง(stationary) ค่า δ คือ fractional ดังนั้นพารามิเตอร์ผลต่าง $d=\delta+m$ โดยที่

1.) ถ้า $-0.5 < d < 0$ แล้วข้อมูลจะมีลักษณะเป็น short memory และ stationary

2.) ถ้า $0 < d < 0.5$ แล้วข้อมูลจะมีลักษณะเป็น long memory และ stationary

สมมุติให้ $\theta(L)$ มีรากทั้งหมดอยู่นอกวงกลม 1 หน่วย ดังนั้นส่วนของกระบวนการ MA จะเป็น

$$[\theta(L)]^{-1}\phi(L)(1-L)^{\delta}[(1-L)^{m}y_{t}-\mu]=\epsilon_{t} \qquad (2.12)$$

ส่วน spectral density ($f(\omega)$) เป็นดังสมการ

$$f(\omega)=f_{u}(\omega)(4\sin^{2}(\omega/2))^{-d} \quad , \quad \omega\in[-\pi,\pi] \qquad (2.13)$$

เมื่อฟังก์ชัน $f_u(\omega)$ คือ spectral density ของกระบวนการ ARMA(p,q) และในเทอมของ $(1-L)^d$ สามารถขยายความได้ดังนี้

$$(1-L)^{d}=\sum_{k=0}^{\infty}\begin{pmatrix} d \\ k \end{pmatrix}(-L)^{k}=\sum_{k=0}^{\infty}b_{k}L^{k}, \qquad (2.14)$$

เมื่อ $b_0=1, b_1=-d, b_2=\frac{1}{2}d(1-d), b_j=\frac{1}{j}b_{j-1}(j-1-d), j\geq 3$ ถ้า $|d|<\frac{1}{2}$ จากนั้น $\sum_{k=0}^{\infty}\begin{pmatrix} d \\ k \end{pmatrix}^{2}<\infty$ และ สมการ (2.14) จะเป็นสมการที่นิยามถึง stationary process สำหรับ $d>-\frac{1}{2}$ และ $(1-L)^{d}$ สามารถที่จะเปลี่ยนให้เพิ่มขึ้นได้ และจากที่แสดงในสมการ (2.12)

นั้นสามารถนำมาใช้เพื่อให้ได้ค่าสัมประสิทธิ์ของตัวแทนที่แสดงเป็นเชิงอัตถดถอยที่เป็นอนันต์ของกระบวนการ ARFIMA นั่นคือ

$$(1-\sum_{i=1}^{\infty}\delta_i L_i)y_t = \epsilon_t \tag{2.15}$$

เมื่อ $\delta_i = b_i - \sum_{j=1}^{q}\theta_j\delta_{i-j} + \sum_{j=1}^{p}\rho_j b_{i-j}$ และ $\sum \delta_i^2 < \infty$ เมื่อ $|d| < \frac{1}{2}$

2.1.4 Maximum Likelihood Estimation

เพื่อให้เกิดความเข้าใจวิธีการประมาณค่าความควรจะเป็นสูงสุด(Maximum Likelihood Estimation:MLE) พิจารณาตัวอย่างต่อไปนี้ (เศรษฐมิติประยุกต์สำหรับการตลาดเกษตร,2548)

ให้ y เป็นตัวแปรสุ่มแจกแจงแบบเบอร์นูลีย์ และมีพารามิเตอร์ที่รู้ค่า(ซึ่งใช้สัญลักษณ์ p สมมุติให้ p อาจจะมีค่าเท่ากับ $\frac{1}{4}$ หรือ $\frac{3}{4}$) สมมุติตัวอย่างมีขนาด n=3 ด้วยค่า $y_1=1$, $y_2=1$, $y_3=0$ การประมาณค่า p ใน MLE คือ การหาโอกาสที่สูงสุดที่จะเกิดค่า p (ซึ่งไม่ทราบค่า) โดยการเลือก $p=\frac{1}{4}$ หรือ $\frac{3}{4}$ หรือค่าอื่นใดที่จะทำให้มีความน่าจะเป็นสูงสุด (maximum propability) ที่ค่า y ในตัวอย่างจะเกิดขึ้น ในเมื่อ $f(y|p) = p^y(1-p)^{1-y}$ สำหรับ y = 0 หรือ 1 ค่าความน่าจะเป็น (probability) ของตัวแปรสุ่มจะมี probability density function ของ ค่า y_1 , y_2, และ y_3 ดังนั้น

$$\begin{aligned} f(y_1=1, y_2=1, y_3=0) &= f(1,1,0) \\ &= \prod_{i=1}^{3} p^{y_i}(1-p)^{1-y_i} \\ &= p.p.(1-p) \end{aligned}$$

ความหมายของสมการ p.p(1-p) คือ ฟังก์ชัน p เมื่อให้ค่าสังเกตได้ของตัวอย่าง y ดังนั้นจึงเรียกได้ว่าเป็นฟังก์ชันความควรจะเป็น(likelihood function) โดยใช้สัญลักษณ์ $L(p|y)$ ซึ่งเหมือนกับ probability density function ของตัวอย่างสุ่ม เพียงแต่ตีความหมายให้เป็นฟังก์ชันของตัวพารามิเตอร์ที่ไม่รู้ค่าแทนฟังก์ชันความควรจะเป็น จากตัวอย่างข้างต้นค่าของฟังก์ชันที่คำนวณได้แก่

เมื่อเลือก $p=\frac{1}{4}$: $L\left(\frac{1}{4}\middle|y\right) = 0.046$

และเมื่อ $p=\frac{3}{4}$: $L\left(\frac{3}{4}\middle|y\right) = 0.141$

ดังนั้นโอกาสหรือความควรจะเป็น (probability หรือ likelihood) สูงสุดในการได้ตัวอย่างที่มีค่า (y_1=1 , y_2=1 , y_3=0) คือ เลือกค่าประมาณ $\widetilde{p} = \frac{3}{4}$ ซึ่งเป็น MLE ของ p แปลว่า ค่าของ p ที่มีความเป็นไปได้มากที่สุดในการสร้างข้อมูล y ชุดนี้นั่นเอง

ในความเป็นจริงชุดของพารามิเตอร์มีมากกว่า 2 ค่า สำหรับการสุ่มแจกแจงแบบเบอนูลีย์นั้น ค่าพารามิเตอร์จะมีค่าอยู่ในช่วง 0 ถึง 1 เท่านั้น วิธีการหนึ่งในการหาค่า p คือ คำนวณเพื่อหาค่าสูงสุดของ $L(p|y) = p^2(1-p)$ ด้วยการแทนค่า p ต่าง ๆ ค่าของ $L(p|y)$ ในแกนตั้งและค่า p ในแกนนอนฟังก์ชันจะมีค่าสูงสุดเมื่อ p โดยทั่วไปแล้วการประมาณค่าพารามิเตอร์เป็นการหาค่าที่ทำให้ฟังก์ชัน probability density function ของตัวอย่างสูงสุดทั่วไป(global maximum) ทั้งนี้โดยใช้ค่าตัวอย่างที่สังเกตมา

2.2 ทฤษฎีและการวิเคราะห์ทางสถิติ

2.2.1 การทดสอบทางสถิติของ Long Memory

1) การทดสอบโดยใช้วิธี R/S(Range Over Standard Deviation) Statistic

การทดสอบสำหรับ long memory หรือ long range dependence คือ rescaled rang หรือ rang over standard deviation หรือ simply R/S Statistic ซึ่งได้เคยถูกเสนอไว้ในครั้งแรกโดย Hurst (1951) ต่อมาถูกปรับปรุงโดย Mandelbrot(1972) ค่าสถิติ R/S เป็นช่วงค่าสูงสุดต่ำสุดของผลรวมบางส่วนของการเบี่ยงเบนของอนุกรมเวลาจากค่าเฉลี่ยของอนุกรมเวลานั้น และ R/S statistic และได้ถูกเปลี่ยนแปลงการวัดให้เหมาะสมโดยค่าเบี่ยงเบนมาตรฐานของอนุกรมเวลาโดยเฉพาะอย่างยิ่งเมื่อพิจารณาถึงอนุกรมเวลา y_t สำหรับ $t = 1,\ldots,T$ ดังนั้น สถิติ R/S ได้นิยามไว้เป็นดังนี้

$$Q_T = \frac{1}{s_T}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(y_j - \overline{y}) - \min_{1\le k\le T}\sum_{j=1}^{k}\left(y_j - \overline{y}\right)\right] \qquad (2.16)$$

เมื่อ Q_T = ค่าสถิติ R/S , $\overline{y} = 1/T\sum_{i=1}^{T} y_i$, และ $s_T = \sqrt{1/T\sum_{i=1}^{T}(y_i - \overline{y})^2}$ ถ้า y_t เป็น i.i.d (independent identically distributed) ซึ่งตัวแปรเชิงสุ่มแบบปกตินั้นคือ

$$\frac{1}{\sqrt{T}} Q_T \Rightarrow V$$

เมื่อ $\Rightarrow$ ให้เป็น weak convergence และ V คือ Brownian bridge ในช่วงหนึ่งหน่วย Lo(1991) ได้คิดวิธีควันไทล์ที่จะเลือกค่าของ V

Lo ได้ชี้ให้เห็นว่า สถิติ R/S ใช้ไม่ได้กับ short rang dependence โดยเฉพาะถ้า y_t มีอัตสหสัมพันธ์(หรือมี short memory) แล้วนั้น การจำกัดการกระจายของ $Q_T / \sqrt{T}$ คือ V ที่ถูกวัดค่าโดยรากที่สองของความแปรปรวนในระยะยาวของ y_t เพื่อให้มี short rang dependence ใน y_t Lo จึงได้ปรับปรุงสูตรสถิติ R/S ดังนี้

$$\widetilde{Q}_T = \frac{1}{\hat{\sigma}_T(q)}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(y_j-\bar{y}) - \min_{1\le k\le T}\sum_{j=1}^{k}(y_j-\bar{y})\right] \qquad (2.17)$$

เมื่อค่าเบี่ยงเบนมาตรฐานของตัวอย่างได้ถูกเปลี่ยนไปเป็นโดยรากที่สองของการประมาณค่าแบบ Newey-West ของความแปรปรวนในระยะยาวด้วย bandwidth q^2 และ Lo แสดงให้เห็นว่าถ้ามี short memory แต่ไม่มี long memory ใน y_t แล้ว $\widetilde{Q}_T$ ยังลู่เข้าสู่ค่า V ซึ่งเป็นช่วงสูงสุดต่ำสุดของ Brownian bridge

2) การทดสอบโดยใช้วิธี GPH Test

Geweke and Porter (1983) ได้เสนอวิธี semi-nonparametric ไว้ เพื่อประมาณค่า fractional differencing parameter (d) โดยต้องใช้การแสดงให้เห็นถึง fractionally integrated process ของอนุกรมเวลา long memory ดังสมการ $(1-L)^d(y_t-\mu) = u_t$ เพื่อเป็นการทดสอบ long memory โดยเฉพาะ spectral density ของ fractionally integrated process ถูกกำหนดโดยฟังก์ชัน

$$f(\omega) = [4\sin^2(\tfrac{\omega}{2})]^{-d} f_u(\omega) \qquad (2.18a)$$

เมื่อ ω คือ Fourier frequency และ $f_u(\omega)$ คือ spectral density ที่มีลักษณะเช่นเดียวกันกับ u_t แล้ว spectral density estimator คำนวณได้จาก

$$I(\lambda) = \frac{1}{2\pi N}\left|\sum_{i,j=1}^{N} X_j e^{ij\lambda}\right|^2 \qquad (2.18b)$$

เมื่อ λ คือ ความถี่ , N คือ จำนวนของแต่ละเทอมในอนุกรม , X_j คือ ข้อมูล , $I(\lambda)$ คือ spectral density estimator ซึ่งเป็น periodogram ของค่าสังเกตุที่ความถี่ λ โดยอนุกรม long range dependence ควรมี periodogram ซึ่งเป็นสัดส่วนกับ $|\lambda|^{1-2H}$ เมื่อเข้าใกล้จุดกำเนิด ดังนั้นการถดถอยของ logarithm ของ periodogram ใน logarithm ของความถี่ (λ) ควรให้ค่าสัมประสิทธิ์ 1-2H ซึ่งเป็นการประมาณค่า H สังเกตได้ว่าพารามิเตอร์ d (fractional difference parameter) สามารถประมาณได้โดยการถดถอยดังนี้

$$\ln f(\omega_j) = \beta - d\ln[4\sin^2(\frac{\omega_j}{2})] + e_j \qquad (2.19)$$

สำหรับ $j = 1,\ldots,n_f(T)$ Geweke and Porter ดังที่แสดงให้เห็น โดยการใช้การประมาณแบบ periodogram ของ $f(\omega_j)$ ใช้การประมาณแบบกำลังสองน้อยที่สุดประมาณค่าของ $\hat{d}$ โดยการใช้การถดถอยดังแสดงไว้ข้างต้นที่ถูกกระจายอย่างเป็นปกติใน large sample ถ้า $n_f(T) = T^{\alpha}$ โดยที่ $0 < \alpha < 1$

$$\hat{d} \sim N(d, \frac{\pi^2}{6\sum_{j=1}^{n_f}(U_j - \bar{U})^2})$$

เมื่อ

$$U_j = \ln[4\sin^2(\frac{\omega_j}{2})]$$

และ $\bar{U}$ คือ ค่าเฉลี่ยตัวอย่างของ $U_j, j=1,\ldots,n_f$ ภายใต้สมมุติฐานว่าง คือ ไม่มี long memory $d = 0$, t-statistic คือ

$$t_{d=0} = \hat{d}.(\frac{\pi^2}{6\sum_{j=1}^{n_f}(U_j - \bar{U})^2})^{-1/2} \qquad (2.20)$$

ข้อจำกัด คือ ให้เป็นการกระจายแบบปกติมาตรฐาน

2.2.2 การประมาณค่า Long Memory Parameter

การประมาณค่า long memory parameter โดยการใช้ R/S statistic และ วิธีการที่อยู่บนพื้นฐานของ periodogram เพื่อทดสอบ long memory ของอนุกรมเวลา Cheung (1993) ได้ทำการเปรียบเทียบด้วยวิธี Monte Carlo จากการทดสอบต่าง ๆ ทำให้ได้การประมาณของ long memory parameter H หรือ d การทดสอบ GPH จะสร้างการประมาณค่า d โดยอัตโนมัติ ซึ่ง R/S Statistic ยังสามารถประมาณค่าของสัมประสิทธิ์ Hurst (H) ได้ และจะแสดงถึงวิธีการ การประมาณ long memory parameter ด้วย periodogram method อีกทั้ง fractional difference parameter หรือค่า d ยังสามารถประมาณค่าได้โดยการใช้แบบจำลอง $\mathrm{FARIMA}(p,d,q)$ ซึ่งประกอบไปด้วยวิธีต่างๆ ดังนี้

1) การวิเคราะห์ R/S Statistic

การทดสอบทางสถิติของ long memory ด้วยวิธี Modified R/S Statistic นั้นเมื่อไม่มี long memory ในอนุกรมเวลาที่ stationary จะทำให้ R/S statistic ลู่เข้าสู่ตัวแปรสุ่มที่อัตรา $T^{1/2}$ อย่างไรก็ตาม เมื่อกระบวนการ stationary แล้ว y_t มี long memory นั้น Modified R/S Statistic ลู่เข้าสู่ตัวแปรสุ่มที่อัตรา T^{H} เมื่อ H คือ สัมประสิทธิ์ของ Hurst

อันดับแรกต้องคำนวณ Modified R/S statistic เพื่อประมาณ long memory parameter (H) โดยการใช้ k_1 หรือ ค่าสังเกตต่างๆ ตามลำดับในตัวอย่าง เมื่อ k_1 ควรใหญ่เพียงพอ จากนั้นเพิ่มจำนวนของค่าสังเกตโดย factor ของ f ซึ่งเป็นการคำนวณ R/S statistic โดยการใช้ $k_i = fk_{i-1}$ ค่าสังเกตต่างๆ ตามลำดับ สำหรับ $i = 2,\ldots,s$ สังเกตได้ว่าเพื่อให้ได้ค่า Modified R/S statistic ในค่า k_i หรือ ค่าสังเกตต่าง ๆ ตามลำดับ ค่า k หนึ่งค่าสามารถแบ่งส่วนตัวอย่างภายใน $[T/k_i]$ ได้ และทำให้ได้ค่าผลต่าง $[T/k_i]$ เมื่อ $[.]$ แสดงให้เห็นถึงส่วนของตัวเลขจำนวนเต็มทางคณิตศาสตร์ของจำนวนจริง ที่เห็นได้ชัดเจนคือ ยิ่งมีค่า k_i ที่ใหญ่ขึ้นก็จะยิ่งมีค่า $[T/k_i]$ ที่น้อยลงเท่านั้น

2) Periodogram Method

ในส่วนของการทดสอบทางสถิติสำหรับ long memory แสดงให้เห็นถึงกระบวนการ long memory เนื่องจากการคำนวณ spectral density estimator แสดงดังสมการ

$$I(\lambda) = \frac{1}{2\pi N}\left|\sum_{i,j=1}^{N} X_j e^{ij\lambda}\right|^2$$

เมื่อ λ คือ ความถี่ , N คือ จำนวนของแต่ละเทอมในอนุกรม , X_j คือ ข้อมูล , $I(\lambda)$ คือ spectral density estimator

เมื่อ spectral density เข้าใกล้ $C_f\omega^{1-2H}$ แล้วความถี่ ω เข้าใกล้ศูนย์ เนื่องจาก spectral density สามารถประมาณโดย periodogram ซึ่ง log-log plot ของ periodogram เปรียบเทียบกับความถี่ ควรกระจายแถว ๆ เส้นตรงด้วยความชัน 1-2H สำหรับความถี่ที่เข้าใกล้ศูนย์ วิธีการนี้สามารถนำไปใช้เพื่อให้ได้ค่าประมาณของ long memory parameter (H)

2.3 ผลงานวิจัยที่เกี่ยวข้อง

Baillie , Chung and Tieslau (1996) ศึกษาหัวข้อเรื่อง **Analyzing Inflation by the Fractionally Integrated ARFIMA-GARCH Model** เป็นการศึกษาการประยุกต์ใช้ long-memory process เพื่ออธิบายภาวะเงินเฟ้อ 10 ประเทศ มีการใช้เครื่องมือใหม่เพื่อให้ได้การประมาณค่า maximum likelihood จากการประมาณค่าของ ARFIMA-GARCH process ซึ่งเป็นกระบวนการที่รวมส่วนย่อย(I(d)) ด้วยส่วนประกอบ ARMA ที่นิ่งโดยการเพิ่มเติมในค่าเฉลี่ยที่เป็นเงื่อนไขของกระบวนการ มากกว่านั้น long memory process นี้ ต้องการยอมให้มี heteroscedasticity ที่เป็นเงื่อนแบบ GARCH ในการวิเคราะห์ภาวะเงินเงินเฟ้อ post-World War II CPI ในแต่ละเดือน สำหรับ 10 ประเทศ จากการศึกษาได้พบหลักฐานของ long memory ที่มีพฤติกรรมการเปลี่ยนกลับของค่าเฉลี่ยในทุกประเทศยกเว้นประเทศญี่ปุ่นซึ่งปรากฏว่ามีความนิ่งสำหรับเศรษฐกิจภาวะเงินเฟ้อสูง

Hauser (1998) ศึกษาหัวข้อเรื่อง **Maximum likelihood estimators for ARMA and ARFIMA models** โดยการศึกษาจากการจำลองเหตุการณ์คุณสมบัติของตัวประมาณค่า time domain 2 ตัว และ frequency domain 2 ตัว สำหรับ autoregressive fractionally integrated moving-average Gaussian models (ARFIMA(p,d,q)) ตัวประมาณค่าที่นำมาพิจารณาคือ maximum likelihood ที่แท้จริง สำหรับข้อมูลที่มีลักษณะเกี่ยวเนื่องกับ modified profile likelihood(MPL) และตัวประมาณค่า Whittle ทั้งที่มี(WLT) และไม่มี tapered data (WL) ความยาวของอนุกรมคือข้อมูล 100 ตัว ตัวประมาณค่าจะถูกเปรียบเทียบในเทอมของ pile-up effect , mean square error , bias , และ empirical confidence level เวอร์ชันที่ได้ taper แล้วของ whittle likelihood ทำให้กลายเป็นตัวประมาณค่าที่เชื่อถือได้สำหรับแบบจำลอง ARMA และ ARFIMA ข้อเสียเล็กน้อยของแบบจำลองที่ใช้ปฏิบัติการในกรณีของแบบจำลองที่แสดงคุณสมบัติที่ดีคือการต้องถูกชดเชยอย่างเพียงพอในแบบจำลองที่สร้างได้ยากกว่า MPL เป็นทางเลือกสำหรับ WLT แต่เป็นความต้องการคำนวณมากกว่า ซึ่งถ้าไม่เทียบเท่าได้กับ EML ก็เป็นที่นิยมมากกว่า EML โดยเฉพาะอย่างยิ่ง สำหรับ fractionally integrated models นั้น EML จะเป็นส่วนสำคัญได้อย่างชัดเจน WL มีความขาดแคลนเป็นอย่างมากสำหรับขอบเขตของพารามิเตอร์ที่มีขนาดใหญ่ และดังนั้นจึงไม่สามารถได้รับการสนับสนุนได้โดยทั่วไป ในทางอื่น EML ควรถูกใช้เพียงแค่สำหรับ fractionally integrated models เนื่องจากศักยภาพของ EML ความเอนเอียงในเชิงลบขนาดใหญ่ของ fractional integration parameter โดยทั่วไป EML ควรดำเนินไปด้วยความรอบคอบสำหรับแบบจำลอง ARMA(1,1) ด้วยการยกเลิกรากเกือบทั้งหมด ในกรณีของ EML และ MPL สำหรับการอนุมาน ในจำนวนใกล้เคียงของราก MA ด้วย +1

Martin and Wilkins (1998) ศึกษาหัวข้อเรื่อง **Indirect estimation of ARFIMA and VARFIMA models** โดยวิธีการประมาณค่าทางอ้อมได้นำมาเสนอสำหรับการประมาณค่า ARFIMA เช่นเดียวกับแบบจำลอง VARFIMA ที่ซับซ้อนมากกว่า เค้าโครงโดยทั่วไปสำหรับการประมาณค่าโดยทางอ้อมของ fractional model ได้รับการพัฒนาครอบคลุมวิธีการจำลองเหตุการณ์ทางเลือกของ auxiliary model และการประมาณค่า algorithm โดยความสนใจพิเศษคือการเพื่อเปรียบเทียบ finite sampling properties ของตัวประมาณค่าทางอ้อมด้วยตัวประมาณค่า Sowell's(1992a) exact time domain maximum likelihood กับ ตัวประมาณค่า spectral maximum-likelihood ของตัวประมาณค่า Fox and Taqqu(1986) and Geweke and Porter(1983) spectral regression ตัวประมาณค่าทางอ้อมสามารถคำนวณได้อย่างรวดเร็วกว่าตัวประมาณค่า exact time domain maximum likelihood ในขณะที่คุณสมบัติตัวอย่างขนาดเล็กที่ดูเหมือนคล้ายกันได้เกิดขึ้น ผลได้จากการคำนวณของตัวประมาณค่าโดยทางอ้อมเทียบกับการเพิ่ม maximum likelihood ดังเช่นความซับซ้อนของการเพิ่มกระบวนการการเกิดของข้อมูล

Reisen and Lopes (1999) ศึกษาหัวข้อเรื่อง **Some simulations and applications of forecasting long-memory time series models** โดยการพยากรณ์ด้วยแบบจำลอง ARFIMA(p,d,q) และ ARIMA(p,d,q) จากการจำลองเหตุการณ์ซึ่งเทคนิคของการพยากรณ์ของแบบจำลอง ARIMA(p,d,q) ทั้งยังสามารถนำไปใช้เมื่อ d เป็น fractional นั่นคือ สำหรับแบบจำลอง ARFIMA(p,d,q) และยังได้ทำการศึกษาเพื่อเปรียบเทียบตัวประมาณค่าสองตัวของ d ที่ได้มาด้วยวิธีการถดถอย ซึ่งทดสอบสมมุติฐานเพื่อตัดสินว่าอนุกรมนั้นๆ มีคุณสมบัติ long memory หรือไม่ และได้นำไปเปรียบเทียบบนพื้นฐานของ k-step ahead forecast errors ของค่า d ได้นำคุณสมบัติใดๆ ของ long-memory models ไปตรวจสอบโดยการใช้เซตของข้อมูลที่แท้จริง

Reisen, Jensen, and Francisco (2003) ศึกษาหัวข้อเรื่อง **Long memory inflationary dynamics : the case of Brazil** เรื่องการมี long memory ในพลวัตรการเกิดเงินเฟ้อกรณีศึกษาประเทศบราซิล เนื่องจากมีข้อถกเถียงในเรื่องการมี unit root ในพลวัตรการเกิดเงินเฟ้อ ที่ได้แสดง

ภาวะความเฉื่อยบางอย่าง ตามความจริงแล้ว Cati , et al (วรสาร Applied Econometrics 1999) ได้พบว่าพลวัตรการเกิดเงินเฟ้อในประเทศบราซิลใกล้ที่จะมีความเฉื่อยอย่างเต็มที่แล้ว การศึกษาในครั้งนี้ได้ประมาณ fractional differencing parameter โดยการใช้ข้อจำกัดของ ARFIMA สำหรับอัตราเงินเฟ้อในประเทศบราซิลและผลลัพธ์ที่ได้นั้นคือ พลวัตรการเกิดเงินเฟ้อถูกสร้างแบบจำลองโดย long memory process ได้ดีกว่ากลไก unit root ดังนั้นจึงมีความหมายโดยนัยว่าไม่มีความเฉื่อยในเงินเฟ้อ ตรงกันข้ามกับสิ่งที่ได้เคยพบโดยนักวิจัยท่านอื่น อีกทั้งการศึกษาในครั้งนี้ได้ค้นพบว่าการประมาณของ fractional parameter ได้คงที่ในการทำผลต่างครั้งแรก(first difference)

Doornik and Ooms (2004) ศึกษาหัวข้อเรื่อง **Inference and forecasting for ARFIMA models with an application to US and UK inflation** เป็นการศึกษาลักษณะที่ใช้ได้จริงของ likelihood-based inference และ เกี่ยวกับการพยากรณ์ของอนุกรมด้วย long memory ที่ได้นำมาพิจารณา ซึ่งใช้แบบจำลอง ARFIMA(p,d,q) เป็นตัวทดสอบด้วยตัวถดถอยที่ได้กำหนดไว้ (deterministic regressor) ลักษณะการสุ่มตัวอย่างมีวิธีนำมาเปรียบเทียบกัน คือ วิธี approximate method และ exact first-order asymptotic method การวิเคราะห์ได้ขยายออกไปถึงใช้การวิเคราะห์แบบ modified profile likelihood ซึ่งเป็นวิธี higher-order asymptotic ได้นำเสนอโดย Cox and Reid (1987) ความสัมพันธ์ของผลต่างระหว่าง วิธีการนั้นได้ถูกทดสอบด้วยแบบจำลองต่าง ๆ และการพยากรณ์ต่างๆของเงินเฟ้อราคาผู้บริโภคหลักรายเดือนในประเทศสหรัฐอเมริกาและเงินเฟ้อราคาผู้บริโภคทั้งหมดในประเทศอังกฤษ

Tsay (2008) ศึกษาหัวข้อเรื่อง **Analyzing Inflation by the ARFIMA Model with Markov-Switching Fractional Differencing Parameter** เป็นการศึกษาโดยการใช้แบบจำลอง ARFIMA ซึ่งได้สร้างมาจาก Markov-switching fractional differencing parameter จากการศึกษาได้พบว่าการเกิด shock จากราคาน้ำมันมีความสำคัญโดยมีผลต่อการเปลี่ยนแปลงรูปแบบเส้นทางความผันผวนของเงินเฟ้อประเทศสหรัฐอเมริกา การประมาณค่ายังให้ผลลัพธ์ที่สนับสนุนที่ว่าเงินเฟ้อสหรัฐฯ มี mean-reverting long memory