บทที่ 2
การตรวจเอกสาร

แนวคิด และทฤษฎี

2.1 การพยากรณ์ (forecasting)

การพยากรณ์เป็นเทคนิคทางสถิติอย่างหนึ่ง ที่ใช้เป็นเครื่องมือในการคาดเดาเหตุการณ์หรือสิ่งที่จะเกิดขึ้นในอนาคต การพยากรณ์มีบทบาทสำคัญในการวางแผนหรือตัดสินใจ ทำให้สามารถกำหนดนโยบาย หรือเตรียมการให้สอดคล้องกับความต้องการและสถานการณ์ที่จะเกิดขึ้นได้ในอนาคต ซึ่งจะช่วยให้แผนงานที่กำหนดมีประสิทธิภาพยิ่งขึ้น

วิธีการพยากรณ์ โดยทั่วไปการพยากรณ์สามารถแบ่งได้เป็น 2 วิธี คือ

1.วิธีการพยากรณ์เชิงคุณลักษณะ (Qualitative Forecasting Methods)

วิธีการพยากรณ์เชิงคุณลักษณะหรือเชิงคุณภาพ เป็นการพยากรณ์ที่ใช้ความเชื่อความรู้สึกประสบการณ์หรือความคิดเห็นส่วนตัวของผู้เชียวชาญ หรือผู้มีประสบการณ์คาดคะเนเหตุการณ์ต่างๆที่จะเกิดขึ้น วิธีการนี้ใช้เมื่อไม่มีข้อมูล หรือมีข้อมูลน้อย รวมถึงข้อมูลไม่น่าเชื่อถือ วิธีนี้เหมาะสำหรับการพยากรณ์ในระยะปานกลางหรือระยะยาว เช่นวิธีการเดลไฟ (Delphi Method) เป็นต้น

2.วิธีการพยากรณ์เชิงปริมาณ (Quantitative Forecasting Methods)

เป็นการพยากรณ์ที่อาศัยข้อมูลข่าวสารเกี่ยวกับเหตุการณ์ที่เกิดขึ้นแล้วในอดีต แล้วนำข้อมูลเหล่านี้มาศึกษาหารูปแบบ หรือตัวแบบพยากรณ์ โดยวิธีการทางสถิติและคณิตศาสตร์ ซึ่งสามารถจะทำได้เมื่อมีข้อมูลในอดีตอยู่ในรูปของตัวเลข หรือสามารถแปลงเป็นตัวเลขได้ สมมุติว่ารูปแบบการแปรผันของข้อมูลที่ผ่านมาจะมีแนวโน้มเป็นลักษณะคล้ายคลึงกันเช่นนั้นด้วยในอนาคต (โดยเฉพาะในระยะสั้น)

วิธีการเชิงปริมาณหรือตัวแบบเชิงปริมาณ สามารถจำแนกออกได้เป็น

- ตัวแบบภายใต้ภาวะคงที่ หรือที่เรียกว่าตัวแบบเชิงกำหนด (Deterministic Models) ตัวแบบประเภทนี้จะไม่มีองค์ประกอบของความไม่แน่นอนหรือไม่มีความคลาดเคลื่อนหรือมีความคลาดเคลื่อนน้อยมากที่สามารถจะละไว้ได้ โดยทั่วไปจะพบในเรื่องของวิทยาศาสตร์กายภาพ (Physical Sciences)

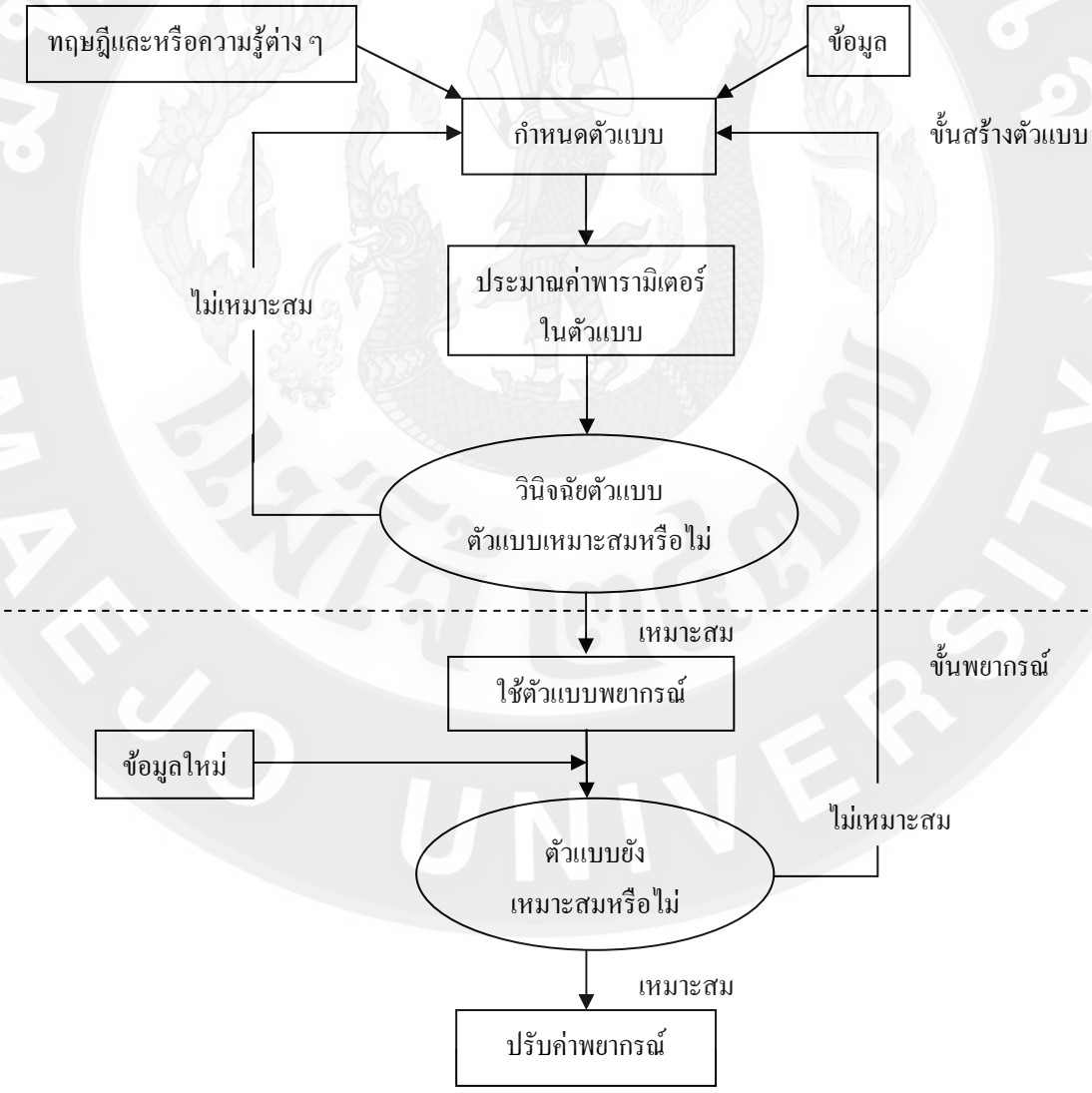
- ตัวแบบภายใต้ภาวะไม่คงที่ หรือที่เรียกว่าตัวแบบความน่าจะเป็น (Probabilistic Models) หรือตัวแบบเฟ้นสุ่ม (Stochastic Models) ตัวแบบประเภทนี้จะเป็นการจำลองความสัมพันธ์ เนื่องจากไม่อาจจะระบุความสัมพันธ์ระหว่างตัวแปรที่ถูกต้องสมบูรณ์ได้ ฉะนั้น ตัว

แบบที่จำลองความสัมพันธ์เหล่านี้ย่อมจะมีค่าผิดพลาดหรือความคลาดเคลื่อนเป็นองค์ประกอบ และค่าผิดพลาดมีการแจกแจงความน่าจะเป็น หรือค่าผิดพลาดเป็นตัวแปรสุ่ม ดังนั้นเรียกตัวแบบที่สร้างขึ้นว่า ตัวแบบความน่าจะเป็น เช่น ตัวแบบการถดถอย (Regression Models) และตัวแบบอนุกรมเวลา (Time Series Models) เป็นต้น

ในการศึกษาครั้งนี้จะทำการวิเคราะห์ข้อมูล โดยใช้ ตัวแบบการถดถอย (Regression Models) และตัวแบบอนุกรมเวลา (Time Series Models) ในการพยากรณ์พื้นที่เพาะปลูก และราคาของข้าวเจ้านาปี ข้าวเหนียวนาปี และข้าวนาปรัง

โครงสร้างของระบบงานพยากรณ์

ระบบงานพยากรณ์จะแบ่งออกเป็น 2 ขั้นตอนใหญ่ๆ คือ ขั้นตอนการสร้างตัวแบบพยากรณ์ และขั้นตอนการพยากรณ์ โดยมีรายละเอียดดังแผนภูมิ (รูปที่ 2.1)



รูปที่ 2.1 แผนภูมิแสดงโครงสร้างของระบบงานพยากรณ์

1.งานขั้นสร้างตัวแบบ

จะเริ่มด้วยการกำหนดตัวแบบทดลอง เป็นตัวแบบเริ่มต้นที่คาดว่าจะเป็นตัวแบบที่ใช้ได้ โดยอาศัยความรู้ และทฤษฎีต่าง ๆ ที่เกี่ยวข้อง รวมถึงการวิเคราะห์รูปแบบข้อมูลเบื้องต้น เช่น การใช้กราฟ เพื่อหาความสัมพันธ์ระหว่างตัวแปร จากนั้นทำการประมาณค่าพารามิเตอร์ในตัวแบบ และทำการตรวจสอบเพื่อความเพียงพอหรือความถูกต้องเหมาะสมในเชิงสถิติ โดยทำการตรวจสอบข้อสมมติ หรือคุณสมบัติต่าง ๆ ที่เกี่ยวข้อง ในเชิงสถิติรวมทั้งรูปแบบความสัมพันธ์ของตัวแบบ ทั้งนี้เพราะ ตัวแบบที่กำหนดขึ้นครั้งแรกนั้นอาจจะยังไม่เหมาะสมเพียงพอ ถ้าพบว่าตัวแบบที่กำหนดไม่สอดคล้องข้อสมมติในเชิงสถิติ หรือยังมีรูปแบบไม่เหมาะสม จะทำการปรับแก้ตัวแบบใหม่ และประมาณค่าพารามิเตอร์ของตัวแบบใหม่ และตรวจสอบความเพียงพอและรูปแบบของตัวแบบใหม่อีก กรรมวิธีจะวนเวียนเช่นนี้ จนกว่าจะพบว่า ตัวแบบพยากรณ์ผ่านการทดสอบ มีความเหมาะสมเพียงพอเชิงสถิติ เมื่อผ่านขั้นนี้แล้วก็สามารถจะใช้ตัวแบบพยากรณ์ค่าที่ต้องการ ซึ่งเข้าสู่ขั้นพยากรณ์

2.งานขั้นพยากรณ์

จากตัวแบบพยากรณ์ที่ได้ เรานำไปใช้พยากรณ์ ตัวแบบพยากรณ์นั้นอาจใช้งานได้ในระยะเวลาหนึ่ง เมื่อเวลาผ่านไปเราได้ข้อมูลใหม่เพิ่มขึ้น ข้อมูลที่เกิดขึ้นใหม่นี้ควรนำมาใช้ตรวจสอบตัวแบบพยากรณ์ว่ายังคงมีความเหมาะสมเพียงพอในเชิงสถิติหรือไม่ ถ้าพบว่าตัวแบบไม่เหมาะสม ควรปรับแก้ตัวแบบใหม่โดยกลับเข้าสู่งานขั้นสร้างตัวแบบดังกล่าวข้างต้น

การวินิจฉัยตัวแบบพยากรณ์

การวินิจฉัยตัวแบบพยากรณ์ คือ การตรวจสอบคุณสมบัติของค่าคลาดเคลื่อนของค่าพยากรณ์ หรือที่เรียกว่า “เศษตกค้าง (Residuals)” เพื่อพิจารณาตัวเลขพยากรณ์ที่สร้างขึ้นมีความเหมาะสมเพียงพอหรือไม่ ในเชิงสถิติ การตรวจสอบจะตรวจสอบว่าส่วนตกค้างมีคุณสมบัติตามข้อสมมติหรือเงื่อนไขต่าง ๆ ของตัวแบบหรือเทคนิคพยากรณ์หรือไม่ รวมทั้งตรวจสอบรูปแบบของตัวแบบที่กำหนดขึ้น ถ้าพบว่ามีคุณสมบัติไม่เป็นไปตามข้อสมมติบางข้อ หรือทั้งหมด นักพยากรณ์ควรพิจารณาปรับแก้ตัวแบบพยากรณ์นั้น คุณสมบัติพื้นฐานที่จะตรวจสอบสำหรับวิธีการพยากรณ์เชิงปริมาณโดยทั่วไป ได้แก่ ความไม่มีสหสัมพันธ์ ค่าเฉลี่ย เป็นศูนย์ ความแปรปรวนคงที่ และการแจกแจงปกติ

การเปรียบเทียบตัวแบบพยากรณ์

ในงานพยากรณ์ โดยปกตินักพยากรณ์จะพิจารณาสร้างเส้นตัวแบบพยากรณ์มากกว่าหนึ่งตัวแบบสำหรับการพยากรณ์เรื่องหนึ่งๆ เพื่อคัดเลือกตัวแบบที่คาดว่าจะให้ค่าพยากรณ์มีความคลาดเคลื่อนต่ำสุดในการคัดเลือกตัวแบบ และนำค่าที่ได้มาเปรียบเทียบกัน และควรจะเลือกค่าวัดความคลาดเคลื่อนต่ำสุด ซึ่งในการศึกษาครั้งนี้ผู้วิจัยจะใช้ค่าเฉลี่ยของเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์ (Mean absolute Percentage Error) เป็นเกณฑ์ในการตัดสินใจเลือกตัวแบบพยากรณ์

2.2 การทดสอบ Unit Root

โดยวิธี Dickey-Fuller Test (DF) หรือ Augmented Dickey-Fuller Test (ADF) การทดสอบหา Unit Root เป็นการทดสอบตัวแปรอนุกรมเวลาที่ใช้ในการศึกษาเพื่อทดสอบความนิ่งของข้อมูล (Stationary) [I(0) : Integrated of Order 0] หรือความไม่นิ่งของข้อมูล (Non-Stationary) [I(d);d>0 : Integrated of Order d] โดยสามารถเขียนรูปแบบสมการได้เป็น 3 รูปแบบคือ

ถ้า X_t เป็นแนวเดินเชิงสุ่ม (Random Walk) จะได้แบบจำลองดังนี้

$$X_t = \rho X_{t-1} + \varepsilon_t \qquad (2.1)$$

ถ้า X_t เป็นแนวเดินเชิงสุ่มซึ่งมีความโน้มเอียงทั่วไปรวมอยู่ด้วย (Random Walk with Drift) จะได้แบบจำลองดังนี้

$$X_t = \alpha + \rho X_{t-1} + \varepsilon_t \qquad (2.2)$$

ถ้า X_t เป็นแนวเดินเชิงสุ่มซึ่งมีความโน้มเอียงทั่วไปรวมอยู่ด้วย (Random Walk with Drift) และมีแนวโน้มตามเวลาเชิงเส้น (Linear Time Trend) จะได้แบบจำลองดังนี้

$$X_t = \alpha + \beta t + \rho X_{t-1} + \varepsilon_t \qquad (2.3)$$

โดยที่ X_t และ X_{t-1} คือตัวแปรที่ใช้ในการศึกษา ณ เวลาที่ t และ t-1

α , ρ และ β คือค่าคงที่

t คือ แนวโน้มเวลา

ε_t คืออนุกรมตัวแปรสุ่ม ที่มีการแจกแจงแบบปกติที่เป็นอิสระต่อกันและเหมือนกัน (independent and identical distribution) โดยมีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนคงที่ สามารถเขียนแทนด้วยสัญลักษณ์ $\varepsilon_t \sim iid\ (0, \sigma_\varepsilon^2)$

ในการทดสอบ X_t มีลักษณะนิ่งของตัวแปร (Stationary Process) (X_t~I(0)) หรือไม่ สามารถทำการทดสอบได้โดยการแปลงสมการที่ *(2.1)* *(2.2)* และ *(2.3)* ให้อยู่ในรูปของ First Differencing (ΔX_t) โดยนำ X_{t-1} ลบออกทั้ง 2 ข้างของสมการ *(2.1)* *(2.2)* และ *(2.3)* จะได้

$$\Delta X_t = \theta X_{t-1} + \varepsilon_t \qquad (2.4)$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t \quad (2.5)$$

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \varepsilon_t \quad (2.6)$$

โดยที่ $\theta = \rho - 1$

โดยมีสมมติฐานว่าง (Null Hypothesis) ในการทดสอบคือ $\theta = 0$ ในขณะที่สมมติฐานทางเลือก (Alternative Hypothesis) ในการทดสอบคือ $\theta < 0$ โดยทำการเปรียบเทียบค่าสถิติ t (t-statistic) ที่คำนวณได้กับค่าที่เหมาะสมที่อยู่ในตาราง Dickey-Fuller (Dickey-Fuller Tables) หรือกับค่าวิกฤต MacKinnon (MacKinnon Critical Values) ในกรณีที่ยอมรับสมมติฐานหลักแสดงว่าตัวแปรนั้นมีลักษณะเป็น Non-Stationary ถ้าปฏิเสธสมมติฐานหลักยอมรับสมมติฐานรองแสดงว่าตัวแปรนั้นมีลักษณะเป็น Stationary

ในกรณีที่เกิดปัญหา Autocorrelation เราจะใช้วิธี Augmented Dickey-Fuller Test (ADF) Test โดยเพิ่ม lagged change $\left[\sum_{j=1}^{p}\phi_j \Delta X_{t-j}\right]$ เข้าไปในสมการทางขวามือ ของสมการ *(2.4)* *(2.5)* และ *(2.6)* ซึ่งสามารถทดสอบหาค่า Unit Root ได้ดีกว่าโดยใช้แบบจำลองดังต่อไปนี้

$$\Delta X_t = \theta X_{t-1} + \left[\sum_{j=1}^{p}\phi_j \Delta X_{t-j}\right] + \varepsilon_t \quad (2.7)$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \left[\sum_{j=1}^{p}\phi_j \Delta X_{t-j}\right] + \varepsilon_t \quad (2.8)$$

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \left[\sum_{j=1}^{p}\phi_j \Delta X_{t-j}\right] + \varepsilon_t \quad (2.9)$$

โดยที่ p = จำนวนของ lag ที่ใส่เข้าไปเพื่อแก้ปัญหา Autocorrelation ในตัวแปรสุ่ม

ϕ = ค่าสัมประสิทธิ์

โดยจะมีการทดสอบเช่นเดียวกับวิธีการของ Dickey and Fuller เพราะค่าสถิติทดสอบมีการแจกแจงเชิงเส้นกำกับที่เหมือนกัน ดังนั้นจึงสามารถใช้ค่าวิกฤตแบบเดียวกันได้

กรณีที่ผลการทดสอบสมมติฐานพบว่า X_t มียูนิทรูทนั้นต้องนำค่า ΔX_t มาทำ Differencing ไปเรื่อยๆ จนสามารถปฏิเสธสมมติฐานที่ว่า X_t เป็น Non – Stationary Process ได้ เพื่อทราบ Order of Integration (d) ว่าอยู่ในระดับใด [$X_t \sim I(d)$; $d > 0$]

ถ้าหากพบว่าข้อมูลดังกล่าวมีลักษณะไม่นิ่งและมีอันดับความสัมพันธ์ของข้อมูล (Order of Integration) ที่มากกว่า 0 [ทดสอบว่า $X_t \sim I(d)$] หรือไม่ ซึ่งจะทำการทดสอบตามรูปแบบสมการดังต่อไปนี้

$$\Delta^{d+1}X_t = \alpha + \beta_t + (\rho - 1)\Delta^{d}X_{t-1} + \sum_{i=1}^{\rho}\phi\Delta^{d+1}X_{t-j} + \varepsilon_t \qquad (2.10)$$

ภายหลังจากทราบค่า d (Order of Integration) แล้วต้องทำการ Differencing ตัวแปร (เท่ากับ d+1 ครั้ง) ตามกระบวนการของ Box –Jenkins Method ก่อนที่จะนำตัวแปรดังกล่าวมาทำการหาสมการถดถอย เพื่อหลีกเลี่ยงปัญหา Spurious Regression ซึ่งคือการที่ความสัมพันธ์ระหว่างตัวแปรอยู่ในรูปแบบความสัมพันธ์ที่ไม่แท้จริง ถึงแม้วิธีนี้จะได้รับความนิยมใช้กันอย่างแพร่หลาย แต่การกระทำดังกล่าวจะทำให้แบบจำลองที่ได้จากการประมาณขาดข้อมูลในส่วนของการปรับตัวแปรต่างๆ เพื่อเข้าสู่ดุลยภาพระยะยาว

2.3 วิธีการวิเคราะห์การถดถอย (Regression Analysis)

การวิเคราะห์การถดถอย เป็นเทคนิคเชิงสถิติหนึ่งสำหรับการศึกษาวิเคราะห์และจำลองรูปแบบความสัมพันธ์พึ่งพิงเชิงคณิตศาสตร์ระหว่างตัวแปรสองกลุ่ม ตัวแปรกลุ่มหนึ่งเรียกว่า “ตัวแปรตาม” หรือ “ตัวแปรผล” (Response variable) มีหนึ่งตัวแปร เป็นตัวแปรที่นักสถิติ หรือนักพยากรณ์ สนใจที่จะศึกษาลักษณะการเปลี่ยนแปลง หรือพยากรณ์ค่าหรือควบคุม โดยศึกษาวิเคราะห์หารูปแบบความสัมพันธ์กับตัวแปรอีกกลุ่มหนึ่ง เรียกว่าตัวแปรในกลุ่มนี้ว่า “ตัวแปรอิสระ” (independent variables) หรือ “ตัวแปรให้ค่าพยากรณ์หรือค่าทำนาย” (predictor variable) ตัวแปรในกลุ่มนี้อาจมีหนึ่งหรือมากกว่าตัวแปรหนึ่งตัวแปร และรูปแบบความสัมพันธ์เชิงคณิตศาสตร์หรือเชิงสถิติที่ได้ เรียกว่า “ตัวแบบการถดถอย” หรือ “สมการถดถอย” จากสมการถดถอย สามารถอธิบายลักษณะการเปลี่ยนแปลงของตัวแปรตาม หรือ พยากรณ์ค่าของตัวแปรตาม หรือใช้ในการควบคุมตัวแปรตาม โดยใช้รูปแบบสมการและค่าของตัวแปรอิสระ

ตัวแบบการถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Models)

ในการวิเคราะห์รูปแบบความสัมพันธ์พึ่งพิงเชิงเส้นของตัวแปรตาม γ บนตัวอิสระมากกว่าหนึ่งตัวแปร $X_1, X_2, \ldots, X_k$ เราเรียกรูปแบบคามสัมพันธ์นั้นว่า *“ตัวแบบการถดถอยเชิงเส้นพหุคุณ”* หรือ *“สมการถดถอยเชิงเส้นพหุคูณ”* และมีรูปแบบทั่วไปดังนี้

$$Y = E(Y) + \varepsilon$$

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \qquad (2.11)$$

โดยที่ค่าเฉลี่ย $E(Y)$ หรือ $E(Y \mid x_1, x_2, \ldots, x_k)$ ของ Y เท่ากับ $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$ เมื่อกำหนดตัวแปรอิสระ $x_1, x_2, \ldots, x_k$

สำหรับตัวอย่างสุ่ม $(X_{i1},X_{i2},...,X_{ik},Y_i)$, i= 1,2,….,n นั่นคือ

$X_1= X_{i1}$, $x_2= X_{i2}$, …… $x_k= X_{ik}$, $Y=Y_i$ และให้ $e=e_i$ สำหรับ i= 1,2,….,n เขียนตัวแบบ *(2.11)* ได้ใหม่ดังนี้

$$Y = E(Y) + \varepsilon_i$$

$$= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \quad , i= 1,2,\ldots,n \qquad (2.12)$$

และ ε_i , i= 1,2,….,n มีข้อสมมุติดังนี้

1. $E(\varepsilon_i) = 0$, i= 1,2,...,n
2. $V(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2$, i= 1,2,…,n
3. $Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i, \varepsilon_j) = 0$, $i \neq j: i , j = 1,2,...,n$
4. $\varepsilon_i \sim N(0, \sigma^2)$, i= 1,2,….,n

ด้วยวิธี OLS ตัวประมาณแบบกำลังสองน้อยที่สุด $B_0,B_1,...,B_k$ ของ $\beta_0, \beta_1,..., \beta_k$ *ตามลำดับ* คำนวณหาได้โดยการแก้สมการ

$$\frac{\partial Q}{\partial \beta i}\Big|\beta_0 = B_0, \beta_1 = B_1, \ldots \beta_k = B_k \qquad = 0, \ i=0,1,2,...,k$$

โดยที่

$$Q = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{i1} - \ldots - \beta_k X_{ik})^2$$

หรือนั่นคือ โดยการแก้ระบบสมการ *(2.13)* :

$$nB_0 + B_1 \sum_{i=1}^{n} X_{i1} + B_2 \sum_{i=1}^{n} X_{i2} + \ldots + B_k \sum_{i=1}^{n} X_{ik} = \sum_{i=1}^{n} Y_i$$

$$B_0 \sum_{i=1}^{n} X_{i1} + B_1 \sum_{i=1}^{n} X_{i1}^2 + B_2 \sum_{i=1}^{n} X_{i1}X_{i2} + \ldots + B_k \sum_{i=1}^{n} X_{i1}X_{ik} = \sum_{i=1}^{n} X_{i1}Y_i$$

$$B_0 \sum_{i=1}^{n} X_{i2} + B_1 \sum_{i=1}^{n} X_{i2}X_{i1} + B_2 \sum_{i=1}^{n} X_{i2}^2 + \ldots + B_k \sum_{i=1}^{n} X_{i2}X_{ik} + \sum_{i=1}^{n} X_{i2}Y_i \ldots\ldots$$

$$B_0 \sum_{i=1}^{n} X_{ik} + B_1 \sum_{i=1}^{n} X_{ik}X_{i1} + B_2 \sum_{i=1}^{n} X_{ik}X_{i2} + \ldots + B_k \sum_{i=1}^{n} X_{ik}^2 \sum_{i=1}^{n} X_{ik}Y_i \qquad (2.13)$$

โดยทั่วไป การวิเคราะห์การถดถอยเชิงพหุคูณ จะใช้เมตริกซ์เป็นเครื่องมือ ซึ่งจะทำให้การวิเคราะห์สะดวกมากขึ้น ฉะนั้นให้

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \qquad \underline{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & & & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$$\underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \qquad \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_k \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_k \end{bmatrix}$$

ดังนั้น $\underline{Y}$ *เป็นเมตริกซ์ขนาด* nx1 *ของตัวแปรสุ่ม หรือเวกเตอร์ขนาด* n *ของตัวแปรสุม*

$\underline{\beta}$ เป็นเวกเตอร์ขนาด *k+1* ของพารามิเตอร์

$\underline{\varepsilon}$ เป็นเวกเตอร์ขนาด *n* ของตัวแปรสุ่มค่าผิดพลาด

$\underline{B}$ เป็นเวกเตอร์ขนาด *k+1* ของตัวประมาณของค่าพารามิเตอร์

และ $\underline{X}$ เป็นเมตริกซ์ขนาด *nx(k+1)* ของค่าคงที่ 1 และค่าของตัวแปรอิสระ

โดยการใช้สัญลักษณ์เมตริกซ์ข้างต้น เขียนระบบสมการ *(2.12)* ได้ใหม่ดังนี้

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\varepsilon} \qquad (2.14)$$

และเขียนข้อสมมติของ $\underline{\varepsilon}$ สามารถเขียนได้สั้น ๆ ดังนี้

$$\underline{\varepsilon} \sim N_n(\underline{0}, \underline{I}\sigma^2) \qquad (2.15)$$

ซึ่งหมายความว่า $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ เป็นอิสระกัน และต่างมีการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และความแปรปรวน σ^2 *(*$N(0,\sigma^2)$*)*

ในทำนองเดียวกัน สามารถเขียนข้อสมมติของ Y_i, i=1,2,…,n ได้ดังนี้

$$\underline{Y} \sim N_n(\underline{X}\beta, \underline{I}\sigma^2) \qquad (2.16)$$

จากระบบสมการปกติ *(2.13)* เขียนในเทอมของเมตริกซ์ได้สั้น ๆ เป็น

$$(\underline{X}'\underline{X})\underline{B} = \underline{X}'\underline{Y} \qquad (2.17)$$

ซึ่ง $\underline{X}'$ *หมายถึงเมตริกซ์ที่สับเปลี่ยน (*transposed matrix*) ของเมตริกซ์* $\underline{X}$

$$\underline{X}' = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & & & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

การแก้สมการ *(2.17)* หา $\underline{B}$ จะสมมติว่าหาเมตริกซ์ผกผันได้ $(\underline{X}'\underline{X})$ ของเมตริกซ์ $\underline{X}'\underline{X}$ ได้ ซึ่งเป็นจริงโดยทั่วไปในทางปฏิบัติ เพราะฉะนั้น ตัวประมาณแบบกำลังสองน้อยที่สุดสามัญคือ

$$\underline{B} = (X'X)^{-1}\underline{X}'\underline{Y} \qquad (2.18)$$

และตัวแบบพยากรณ์ค่า Y หรือตัวประมาณของค่าเฉลี่ย $E(Y)$ ของ Y เมื่อกำหนด $X_1 = X_{01}$, $X_2 = X_{02}$,..., $X_k = X_{0k}$ คือ

$$\hat{Y} = \underline{X}_0'\underline{B}$$
$$= B_0 + B_1X_{01} + B_2X_{02} + \cdots + B_kX_{0k}$$

ซึ่ง $\underline{X}_0' = (X_{01}, X_{02}, \ldots, X_{0k})$

ขั้นตอนการสร้างตัวแบบการถดถอยเชิงเส้น

กรรมวิธีการสร้างตัวแบบการถดถอยเชิงเส้น มีขั้นตอนที่ควรดำเนินการดังนี้

1. กำหนดรูปแบบความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ

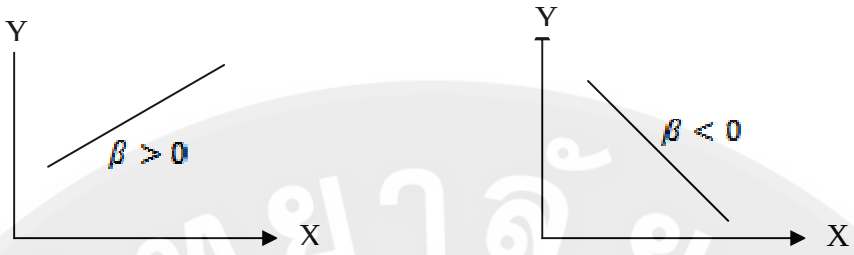
การกำหนดรูปแบบความสัมพันธ์ อาจใช้ทฤษฎีที่เกี่ยวข้อง เช่น ทฤษฎีเศรษฐศาสตร์มีการกำหนดรูปแบบความสัมพันธ์ที่เรียกว่าตัวแบบเศรษฐศาสตร์ ซึ่งอาจจะนำมาประยุกต์ได้กับเรื่องที่ศึกษา ในกรณีที่ไม่สามารถหาทฤษฎีใดมาประยุกต์ได้ นักพยากรณ์จะพิจารณารูปแบบความสัมพันธ์ โดยอาศัยข้อมูลที่มีอยู่ของตัวแปรตาม และของตัวแปรอิสระ ซึ่งสามารถพิจารณาโดยใช้กราฟดังนี้

1.1 โดยทั่วไปจะมีข้อสมมติว่าตัวแปรตาม Y มีการแจกแจงความน่าจะเป็นแบบปกติ ฉะนั้น ควรเขียนกราฟแผนภาพแบบจุด หรือแผนภาพฮิสโทแกรม (histogram diagram) เพื่อดูลักษณะการกระจาย หรือการแจกแจงของ Y ว่าเข้ารูปลักษณะแบบสมมาตรหรือไม่ ถ้าพบว่ามีลักษณะไม่สมมาตร โดยเบ้ไปทางซ้ายหรือทางขวามาก ควรที่จะแปลงข้อมูลของ Y เพื่อให้เข้าลักษณะการแจกแจงแบบสมมาตร วิธีการแปลงค่าของ Y อาจจะทดลองด้วยแบบต่างๆ เช่น $\sqrt{Y}$, $1/\sqrt{Y}$, 1/Y , lnY , หรือ logY เป็นต้น

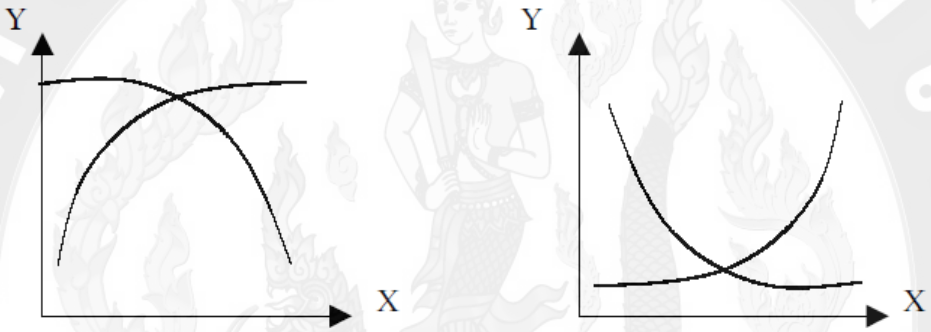
1.2 เขียนกราฟระหว่างตัวแปรตาม กับตัวแปรอิสระทีละตัว เพื่อพิจารณากำหนดรูปแบบความสัมพันธ์เป็นคู่ ๆ ระหว่างตัวแปรตามและตัวแปรอิสระ

ตัวอย่างรูปแบบฟังก์ชันแสดงความสัมพันธ์ระหว่างตัวแปรตาม Y และตัวแปรอิสระ x ใด ๆ โดยละองค์ประกอบที่เป็นค่าผิดพลาด ε ในฟังก์ชัน

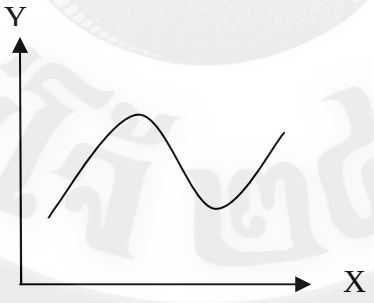
(1) ตัวแบบเชิงเส้น : Y=$\alpha + \beta x$



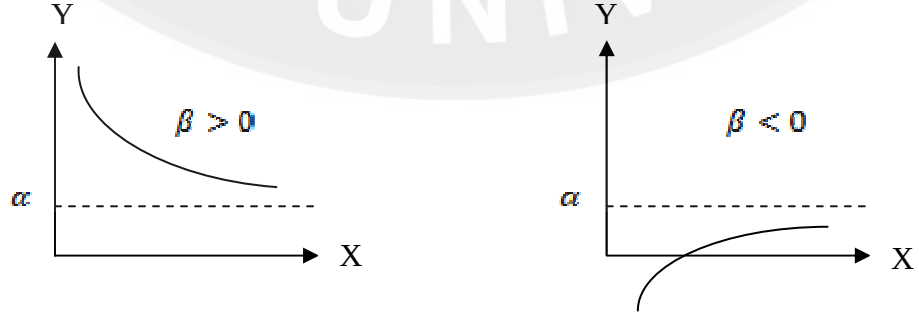
(2) ตัวแบบกำลังสอง (quadratic model) หรือตัวแบบพหุนามอันดับสอง (2nd –order polynomial model) : Y= $\beta_0+\beta_1 X + \beta_2 X^2$



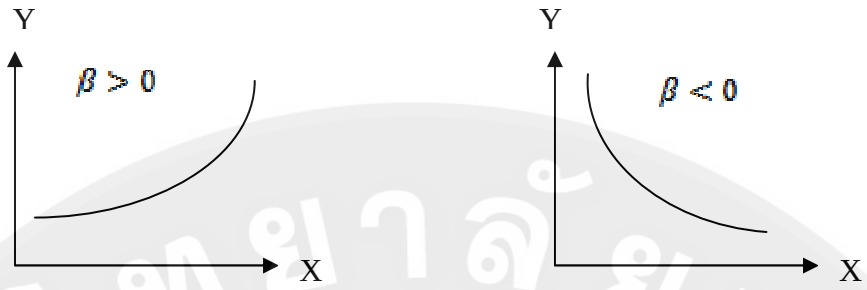
(3) ตัวแบบพหุนามอันดับสาม : Y=$\beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$



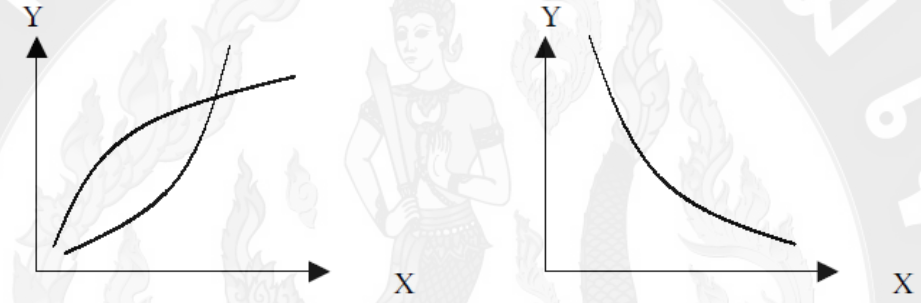
(4) ตัวแบบส่วนกลับใน *x*: Y = $\alpha + \beta\left(\frac{1}{X}\right), X \neq 0$



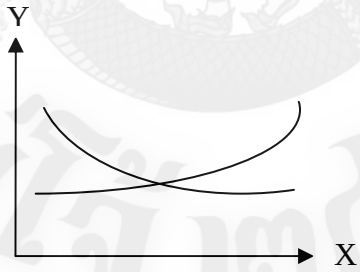
(5)ตัวแบบเซมิล็อค (Semi log model): $\ln Y = \alpha + \beta x$



(6) ตัวแบบล็อค-ล็อค (*Log-log model*): $\ln Y = \alpha + \beta \ln x$



(7) ตัวแบบส่วนกลับ $Y = \frac{1}{\alpha+\beta x}$



1.3 ในการแปลงตัวแบบการถดถอยไม่เป็นเชิงเส้นให้เป็นตัวแบบการถดถอยเชิงเส้น

ตัวแบบเริ่มแรกที่นักพยากรณ์เลือก อาจไม่อยู่ในลักษณะเหมือนตัวแบบทั่วไป *(2.11)* ซึ่งอยู่ในรูปแบบเชิงเส้นทั้งในเทอมของพารามิเตอร์ และในเทอมของตัวแปร ในกรณีของตัวแปรเราสามารถแปลงให้อยู่ในแบบเชิงเส้นได้ง่ายดังตัวอย่างต่อไปนี้

ตัวอย่างการแปลงตัวแปร

(1) $Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$ ให้ $x_1 = x$ และ $x_2 = x^2$ ได้ตัวแบบการถดถอยเชิงเส้นพหุคูณ : $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$

(2) $Y = \alpha + \beta\left(\frac{1}{x}\right) + \varepsilon$ ให้ $x_1 = 1/x$ ได้ตัวแบบ

$$Y = \alpha + \beta x_1 + \varepsilon$$

(3) $\ln Y = \alpha + \beta x + \varepsilon$

$$Y^{*} = \alpha + \beta x + \varepsilon$$

(4) $\ln Y = \alpha + \beta \ln X + \varepsilon$ ให้

$Y^{*} = \ln Y$ และ $X_1 = \ln X$ ได้ตัวแบบ

$$Y^{*} = \alpha + \beta X_1 + \varepsilon$$

(5) $Y = \frac{1}{\alpha+\beta x+\varepsilon}$ ให้ $Y^{*} = \frac{1}{Y}$ ได้ตัวแบบ

(6) $Y^{*} = \alpha + \beta X + \varepsilon$

$Y = \beta_0 + \beta_1\left(\frac{1}{X_1}\right) + \beta_2 X_2 + \beta_3 X_1 X_2 + \varepsilon$ ให้ $X^{*}_1 = \frac{1}{X_1}, X^{*}_2 = X_2, X^{*}_3 = X_1 X_2$

ได้ตัวแบบ

$$Y = \beta_0 + \beta_1 x_1^{*} + \beta_2 x_2^{*} + \beta_3 x_3^{*} + \varepsilon$$

การแปลงตัวแบบการถดถอยให้เป็นตัวแบบการถดถอยเชิงเส้นนั้นอาจกระทำไม่ได้ ดังเช่นตัวอย่างตัวแบบ

$$Y = \alpha + \beta e^{k} + \varepsilon$$

ตัวแบบดังกล่าว เรียกว่า “ตัวแบบการถดถอยไม่เป็นเชิงเส้น” และใช้เทคนิคการประมาณไม่เป็นเชิงเส้น (*Nonlinear estimation)* เช่นวิธีการเกาส์-นิวตัน (*Guass-Newton method)* ในการประมาณค่าพารามิเตอร์ของตัวแบบ

ในกรณีที่มีหนึ่งตัวแปรอิสระ และพบว่าข้อมูลของ Y และ X ปรากฏความสัมพันธ์ไม่เป็นเชิงเส้น นักพยากรณ์อาจใช้วิธีการแปลงตัวแปรตาม Y และ / หรือตัวแปรอิสระ X ให้มีความสัมพันธ์เชิงเสน และการแปลงที่บ่อยครั้งใช้ได้ผลดี คือ

ให้ $Y^{*} = \ln Y$ หรือ $\log_{10} Y$ $\quad X^{*} = \ln X$ หรือ $\log_{10} X$

$Y^{*} = \sqrt{Y}$ $\quad X^{*} = \sqrt{X}$

$Y^{*} = 1/Y$ $\quad X^{*} = 1/X$

2.การประมาณค่าพารามิเตอร์ของตัวแบบการถดถอย

เมื่อนักพยากรณ์กำหนดตัวแบบการถดถอยเป็นตัวแบบทดลองได้แล้ว ซึ่งอาจจะมีมากกว่าหนึ่งตัวแบบ ขั้นตอนต่อไปก็คือ ประมาณค่าพารามิเตอร์ในตัวแบบ วิธีแบบฉบับที่ใช้กันทั่วไปคือวิธี *OLS* ในการประมาณค่าพารามิเตอร์ และใช้วิธีวิเคราะห์สถิติมาตรฐานทั่วไปในการอนุมานเกี่ยวกับพารามิเตอร์ตลอดจนค่าพยากรณ์แบบช่วง

3.การวินิจฉัยความเพียงพอของตัวแบบการถดถอย

งานขั้นกำหนดรูปแบบของตัวแบบ และงานประมาณค่าพารามิเตอร์ของตัวแบบเป็นเพียงงานครึ่งหนึ่งของกรรมวิธีสร้างตัวแบบพยากรณ์ ตัวแบบพยากรณ์ที่ได้จากขั้น 1 และ 2 อาจจะยังไม่เหมาะสมหรือยังไม่เพียงพอที่จะใช้พยากรณ์ นักพยากรณ์จึงควรตรวจสอบและทำการเปรียบเทียบคัดเลือกตัวแบบพยากรณ์ ถ้าตรวจสอบพบว่า ตัวแบบที่กำลังพิจารณายังขาดความเหมาะสม จะกลับไปทำงานในขั้นที่ 1 ถึง 3 ซ้ำ ๆ จนกว่าจะได้ตัวแบบที่เหมาะสมเพียงพอในเชิงสถิติที่จะใช้พยากรณ์ค่าต่อไป

เนื่องจากการอนุมานเชิงสถิติเกี่ยวกับพารามิเตอร์ ซึ่งเกี่ยวกับการทดสอบข้อสมมติฐานต่าง ๆ และการประมาณค่าแบบช่วงของค่าพารามิเตอร์ ตลอดจนการทดสอบและการประมาณค่าแบบช่วงของค่าเฉลี่ย *E(Y)* ของ *Y* และของค่าพยากรณ์ของค่าจริง *Y* โดยทั่วไปการทดสอบและการประมาณค่าดังกล่าว จะกระทำภายใต้ข้อสมมติของ ε_i ดังนั้น จึงจำเป็นต้องตรวจสอบความเหมาะสม หรือความเพียงพอของตัวแบบด้วยการตรวจสอบคุณสมบัติของค่าผิดพลาด ε_i แต่เนื่องจากไม่ทราบค่าจริง ε_i ฉะนั้นจะตรวจสอบคุณสมบัติของค่าเศษเหลือตกค้าง e_i ซึ่งกำหนด $e_i = y_t - \hat{y}_t$ ซึ่งเป็นค่าประมาณของ ε_i เมื่อตัวแบบถูกต้องเพียงพอ และตรวจสอบว่าค่า e_i, (i=1,2,…,n) มีคุณสมบัติสอดคล้องหรือไม่ นั่นคือ มีค่าเฉลี่ยเท่ากับศูนย์ มีความแปรปรวนคงที่ ไม่มีอัตสหสัมพันธ์ และมีการแจกแจงแบบปกติหรือไม่ นอกจากนี้อาจพบว่า ตัวแบบไม่เหมาะสมเนื่องมาจากตัวแบบมีรูปแบบยังไม่ถูกต้องเหมาะสม วิธีการตรวจสอบนักพยากรณ์อาจเลือกใช้วิธีกราฟ หรือวิธีการทดสอบเชิงสถิติ ซึ่งวิธีเชิงสถิติเป็นวิธีที่มีระเบียบ (*formal*) หรือมีทฤษฎีสนับสนุน อย่างไรก็ตาม ถึงแม้ว่าวิธีกราฟอาจจะไม่เป็นวิธีเชิงระเบียบทฤษฎี แต่โดยทั่วไปก็เป็นวิธีการที่เพียงพอที่จะใช้วินิจฉัยตัวแบบ และเป็นวิธีที่ง่ายที่ใช้โดยทั่วไป ฉะนั้นในที่นี้จะกล่าวถึงเฉพาะวิธีกราฟ

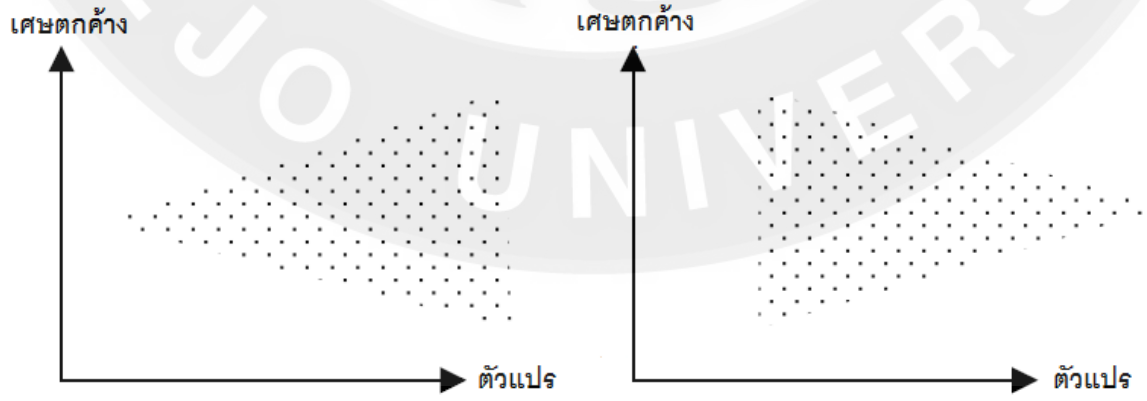
โดยวิธีกราฟ เราจะเขียนกราฟของค่าเศษตกค้าง e_i หรือค่าเศษตกค้างมาตรฐาน (*Standardized residuals*) $e_i/\sqrt{MSE}$ กับตัวแปรต่างๆ:

1. ตัวแปร $\hat{y}_t$

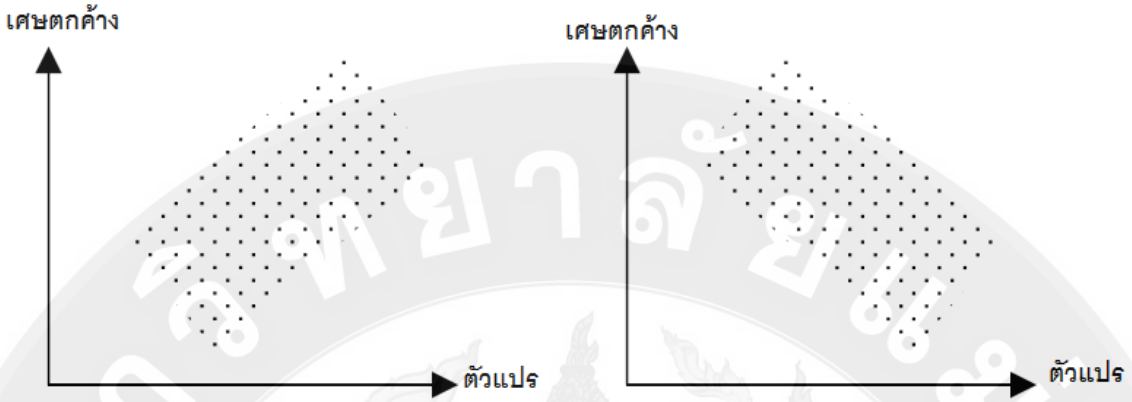
2. ตัวแปรอิสระ $\mathbf{X_i}$ แต่ละตัว
3. ตัวแปรเวลา ถ้าข้อมูลเป็นอนุกรมเวลา
4. เขียนกราฟความน่าจะเป็นแบบปกติ (*normal probability plot)* หรือแผนภาพฮิสโทแกรมของ e_i หรือ $e_i/\sqrt{MSE}$

ตัวอย่างกราฟรูปที่ *2.2* ถึง *2.5* แสดงลักษณะความสัมพันธ์ที่เป็นไปได้ระหว่าง e_i (หรือ$e_i/\sqrt{MSE}$) กับตัวแปร $\hat{y}_i$, $\mathbf{X_i}$หรือเวลา

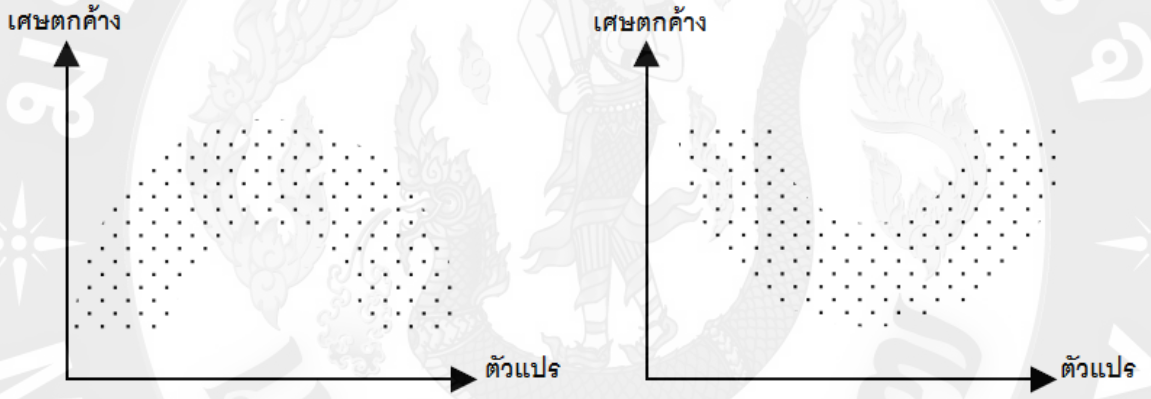
รูปที่ 2.2



รูปที่ 2.3



รูปที่ 2.4



รูปที่ 2.5

ถ้ากราฟระหว่าง e_i (หรือ $e_i/\sqrt{MSE}$) และ $\hat{y}_i$ และตัวแปรอิสระแต่ละตัว และกับเวลา (ถ้าเป็นอนุกรมเวลา) ทั้งหมดมีรูปแบบการกระจายของจุดเป็นแนวขนานดังรูปที่ 2.2 แสดงว่าตัวแบบการถดถอยเชิงเส้นมีรูปแบบเหมาะสม ค่าผิดพลาดมีค่าเฉลี่ยศูนย์ และความแปรปรวนคงที่ นอกจากนี้ในกรณีที่เป็นกราฟ e_i กับเวลา แสดงด้วยว่า ความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระไม่แปรเปลี่ยนตามเวลา หรือไม่มีอิทธิพลของเวลา และค่าคลาดเคลื่อนสุ่มไม่มีอัตสหสัมพันธ์กัน แต่ถ้ากราฟมีรูปแบบไม่เป็นแนวขนาน ดังเช่นรูปที่ 2.3 ถึง 2.5 แสดงว่า ค่าผิดพลาดไม่สอดคล้องคุณสมบัติบางข้อหรือทุกข้อในคุณสมบัติของ e_i หรือตัวแบบยังมีรูปแบบไม่เหมาะสม

ถ้ารูปกราฟปรากฏดังรูป 2.3 แสดงว่า ความแปรปรวนของค่าผิดพลาดไม่คงที่ การแก้ปัญหา อาจใช้วิธีแปลงค่าของตัวแปรตาม *Y* (เช่น ทดลองด้วยการแปลงเป็น $\ln Y, \frac{1}{Y}, \sqrt{Y}$ หรือ $1/\sqrt{Y}$ เป็นต้น) หรือใช้วิธีแก้ปัญหาเฉพาะในเรื่องนี้ เช่นใช้วิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนัก (Weighted Least Square Method)

ในกรณีที่รูปกราฟปรากฏดังรูปที่ 2.4 หรือ 2.5 แสดงว่า ตัวแบบการถดถอยเชิงเส้นยังมีรูปแบบไม่เหมาะสม ในรูป 2.4 แสดงว่า ควรมีเทอมค่าคงที่ หรือองค์ประกอบเชิงเส้นในตัวแบบ และในรูป 2.5 แสดงว่า ควรมีเทอมที่มีกำลังสูงขึ้นในตัวแบบ เช่น ควรมีเทอมกำลังสอง αX^2ของตัวแปรอิสระ x ในตัวแบบ และถ้าข้อมูลเป็นอนุกรมเวลากราฟรูป 2.4 แสดงว่า ควรมีเทอมเชิงเส้นหรือเทอมอันอับหนึ่งของเวลาในตัวแบบ และกราฟรูป 2.5 แสดงว่าควรมีเทอมอันดับหนึ่งและอันดับสองของเวลาในตัวแบบด้วย

การตรวจสอบค่าผิดพลาดมีสหสัมพันธ์กันหรือไม่ โดยปกติจะตรวจสอบเมื่อข้อมูลเป็นอนุกรมเวลา และวิธีการตรวจสอบมีหลายวิธี เช่น ใช้วิธีพิจารณาค่าของฟังก์ชันอัตสหสัมพันธ์ตัวอย่าง (Sample Autocorrelation Function (SACF) γ_k ของ $e_1, e_2, \ldots, e_n$ ที่คาบเวลาห่างกัน *k, (k=1,2,...)* เปรียบเทียบกับค่าตัดสินเชิงสถิติหรือค่าวิกฤต (*Critical value)* $2/\sqrt{n}$ โดยประมาณที่ระดับนัยสำคัญ (*significant level)* 0.05 หรือใช้ค่าของตัวสถิติ *Durbin-Watson* สำหรับตรวจสอบอัตสหสัมพันธ์ที่คาบเวลาห่างกัน 1 (*k=1)* เมื่อพบว่าค่าผิดพลาดมีสหสัมพันธ์กัน ในการวิเคราะห์การถดถอยเชิงเส้น จะมีวิธีการแก้ปัญหานี้โดยเฉพาะ เช่นวิธีการแปลงตัวแปรและวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนัก เป็นต้น

การเปรียบเทียบตัวแบบพยากรณ์

โดยปกตินักพยากรณ์จะสร้างตัวแบบมากกว่าหนึ่งตัวแบบ สำหรับในการพยากรณ์เรื่องหนึ่ง เพื่อคัดเลือกตัวแบบที่เหมาะสมกับข้อมูลในเรื่องนั้น ๆ โดยพิจารณาจากค่าความคลาดเคลื่อนว่าตัวแบบใดให้ค่าพยากรณ์ที่มีความคลาดเคลื่อนต่ำสุด

2.4 วิธีบอกซ์-เจนกินส์ (*Box and Jenkins 'Method)*

วิธีบอกซ์-เจนกินส์ เป็นวิธีการสร้างตัวแบบพยากรณ์ โดยพิจารณาสหสัมพันธ์ระหว่าง *Y* ที่ตำแหน่งเวลา หรือคาบเวลา *t* (Y_t) และ *Y* ที่ตำแหน่งเวลา หรือคาบเวลาต่าง ๆ ที่ผ่านมา (Y_{t-1}, Y_{t-2}, ...) เมื่อได้ตัวแบบแล้ว ตัวแบบนี้จะแสดงความสัมพันธ์ระหว่าง Y_t กับ Y_{t-1}, Y_{t-2}, ... และจะใช้ตัวแบบนี้พยากรณ์ค่า $Y_{t+1}, Y_{t+2}, \ldots$ ในอนาคต วิธีนี้จะเหมาะสำหรับการพยากรณ์ในระยะสั้น

หรือระยะปานกลาง และขนาดของข้อมูลที่นำมาวิเคราะห์ควรมีข้อมูลอย่างน้อย 50 จำนวน เพราะถ้าข้อมูลมีจำนวนน้อยเกินไป อาจจะทำให้ไม่เห็นอิทธิพล หรือรูปแบบของฤดูกาล

นอกจากนี้ วิธีบอกซ์-เจนกินส์ จะใช้กับข้อมูลอนุกรมเวลาที่อยู่ในสภาพนิ่ง หรือคงที่ (*Stationary data series)* ซึ่งหมายถึง คงที่ในค่าเฉลี่ย ในค่าความแปรปรวน และในค่าความแปรปรวนร่วม (หรือสหสัมพันธ์คงที่) ไม่แปรผันตามเวลา

ลักษณะตัวแบบบอกซ์-เจนกินส์

แนวคิดของการพัฒนาตัวแบบบอกซ์-เจนกินส์ มาจากการศึกษาวิเคราะห์กระบวนการเชิงเส้น หรือตัวกรองเชิงเส้น (*Linear filter):*

$$Y_t = \mu + \alpha_t + \varphi_1\alpha_{t-1} + \varphi_2\alpha_{t-2} + \cdots \qquad (2.19)$$

นั่นคือ พิจารณาอนุกรมเวลา หรือค่าสังเกต Y_t เกิดจากผลบวกเชิงเส้นของตัวแปรสุ่ม $\alpha_t, \alpha_{t-1}, \ldots$ ที่ไม่มีสหสัมพันธ์ เราเรียกตัวแปรสุ่ม $\alpha_t, \alpha_{t-1}, \ldots$ ค่าผิดพลาดสุ่ม หรือเรียกว่ากระตุกสุ่ม (*Random shocks)* และสมมติว่าแต่ละตัวมีค่าเฉลี่ยเท่ากับศูนย์ มีความแปรปรวนคงที่ และโดยทั่วไปสมมติด้วยว่ามีการแจกแจงปกติ

ในตัวกรองเชิงเส้น หรือตัวแบบเชิงเส้น *(2.19)* พารามิเตอร์ μ คือ ค่าระดับค่าเฉลี่ยของ Y_t เมื่ออนุกรมเวลาอยู่ในสภาพคงที่ และพารามิเตอร์ $\varphi_1, \varphi_2, \ldots$ เป็นน้ำหนักที่ให้กับตัวแปรสุ่ม $\alpha_{t-1}, \alpha_{t-2}, \ldots$

กระบวนการ หรือตัวแบบเชิงเส้น *(2.19)* จะไม่ให้ประโยชน์ถ้าพารามิเตอร์มีจำนวนอนันต์ (จำนวนไม่รู้จบ) เพราะฉะนั้น จะสร้างตัวแบบที่ประกอบด้วยพารามิเตอร์จำนวนจำกัด และเพียงพอที่จะอธิบายอนุกรมเวลาที่พิจารณา

1. ตัวแบบภายใต้ภาวะคงที่ (*Stationary Models)*

รูปแบบของอนุกรมเวลาที่ใช้ในการพยากรณ์โดยวิธีบอกซ์-เจนกินส์ เมื่ออนุกรมเวลามีคุณสมบัติคงที่ มีดังนี้

1.1 ตัวแบบอัตถดถอยอันดับที่ *p (Autoregressive Model of Order p) : AR(p)* ซึ่ง *p* คือ อันดับที่ของตัวแบบอัตถดถอย มีรูปแบบทั่วไปดังนี้

$$Y_t - \mu = \emptyset_1(Y_{t-1} - \mu) + \emptyset_2(Y_{t-2} - \mu) + \emptyset_p(Y_{t-p} - \mu) + \alpha_t$$

หรือ

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + \alpha_t$$

หรือ

$$Y_t = C + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + \alpha_t \qquad (2.20)$$

โดยให้ $Z_t = Z_t - \mu, Z_{t-1} - \mu, \ldots$ และ
$c = \mu(1 - \emptyset_1 - \emptyset_2 - \cdots - \emptyset_p)$ และ $\emptyset_1, \emptyset_2, \ldots, \mu$ เป็นพารามิเตอร์ ซึ่งโดยทั่วไปไม่ทราบค่า จะต้องประมาณค่าจากข้อมูล

ตัวอย่าง เมื่อ *AR(p)* ที่มีค่า *p=1* และ *p=2* มีรูปแบบดังนี้

ตัวแบบ *AR (1) :*

$$Y_t = C + \emptyset_1 Y_{t-1} + \alpha_t \; ; \; |\emptyset_1| < 1$$

โดยมีเงื่อนไขที่ทำให้อนุกรมเวลาเป็นตัวแบบคงที่ $-1 < \emptyset_1 < 1$

ตัวแบบ AR(2) :

$$Y_t = C + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \alpha_t$$

โดยมีเงื่อนไขที่ทำให้อนุกรมเวลาเป็นตัวแบบคงที่

$$\emptyset_1 + \emptyset_2 < 1$$
$$\emptyset_2 - \emptyset_1 < 1$$
$$-1 < \emptyset_2 < 1$$

1.2 ตัวแบบค่าเฉลี่ยเคลื่อนที่อันดับที่ *q* (*Moving Average Model of Order q) : MA (q)* ซึ่ง *q* คือ อันดับที่ของรูปแบบค่าเฉลี่ยเคลื่อนที่ โดยมีรูปแบบดังนี้

$$Y_t = \mu - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \cdots - \theta_q \alpha_{t-q} + \alpha_t \qquad (2.21)$$

ตัวแบบบอกซ์-เจนกินส์ ยังมีเงื่อนไขที่ต้องสอดคล้องอีกหนึ่งเงื่อนไขนอกเหนือจากเงื่อนไขคงที่ (*Stationary)* คือ เงื่อนไข “ผกผันได้” (*invertibility)* ซึ่งพบว่า ตัวแบบ *AR(p), P*$<\infty$มีคุณสมบัติผกผันได้เสมอ แต่อาจไม่คงที่ ขณะที่ตัวแบบ *MA(q)* , *q*$<\infty$ มีคุณสมบัติคงที่เสมอ แต่อาจจะผกผันไม่ได้ เพราะฉะนั้น ต้องตรวจสอบคุณสมบัติคงที่ในตัวแบบ *AR* และตรวจสอบคุณสมบัติผกผันได้ในตัวแบบ *MA*

ตัวอย่าง เมื่อ *MA (q)* มีค่า *q=1* และ *q=2* มีรูปแบบดังนี้

ตัวแบบ *MA (1) :*

$$Y_t = \mu - \theta_1 Y_{t-1} + \alpha_t; \; |\theta_t| < 1$$

โดยมีเงื่อนไขผกผันได้ $-1 < \theta_1 < 1$

ตัวแบบ *MA(2) :*

$$Y_t = \mu - \theta_1 Y_{t-1} - \theta_2 Y_{t-2} + \alpha_t$$

โดยมีเงื่อนไขผกผันได้

$$\theta_1 + \theta_2 < 1$$

$$\theta_2 - \theta_1 < 1$$
$$-1 < \theta_2 < 1$$

1.3ตัวแบบผสมอัตถดถอย – ค่าเฉลี่ยเคลื่อนที่ ที่มีอันดับ *p* และ *q (Autoregressive-Moving Average Model of order p and q) : ARMA(p,q)* มีรูปแบบดังนี้

$$Y_t = \mu + \emptyset_1(Y_{t-1} - \mu) + \emptyset_2(Y_{t-2} - \mu) + \cdots + \emptyset_p(Y_{t-p} - \mu) - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2} - \cdots - \theta_q\alpha_{t-q} + \alpha_t \quad (2.22)$$

ตัวอย่าง : เมื่อตัวแบบ *ARMA (p,q)* ที่มี *p=1,q=1* มีรูปแบบดังนี้

$$Y_t = c + \emptyset_1 Y_{t-1} - \theta_1 Y_{t-1} + \alpha_t$$

โดยมีเงื่อนไขที่ทำให้อนุกรมเวลาเป็นตัวแบบคงที่ และผกผันได้

$$-1 < \emptyset_1 < 1 \text{ และ } -1 < \theta_t < 1$$

2. ตัวแบบภายใต้ภาวะไม่คงที่ (*Nonstationary Models)* และตัวแบบ *ARIMA*

ถ้าข้อมูลอนุกรมเวลาไม่อยู่ในสภาพคงที่ในค่าเฉลี่ย และ/หรือ ความแปรปรวน จะต้องแปลงอนุกรมเวลาให้อยู่ในสภาพคงที่ก่อนพิจารณากำหนดตัวแบบ

ในกรณีอนุกรมเวลาไม่คงที่ในค่าเฉลี่ย การหารูปแบบที่เหมาะสมให้อนุกรมเวลานั้นจะต้องแปลงอนุกรมเวลาให้คงที่ในค่าเฉลี่ย โดยการหาผลต่าง (*Differencing)* ของอนุกรมเดิม ถ้าผลต่างครั้งที่ 1 ของอนุกรมเวลา ทำให้อนุกรมเวลามีค่าเฉลี่ยคงที่ ก็สามารถนำอนุกรมเวลานี้ไปหาตัวแบบที่เหมาะสมต่อไปได้ แต่ถ้าผลต่างครั้งที่ 1 ของอนุกรมเวลานั้น อนุกรมเวลายังไม่คงที่ในค่าเฉลี่ย จะต้องหาผลต่างครั้งที่ 2,3,... ต่อไปจนกว่าคุณสมบัติของอนุกรมเวลาคงที่ในค่าเฉลี่ย ซึ่งโดยทั่วไป ถ้าอนุกรมเวลามีแนวโน้ม มักจะทำผลต่างสองครั้งจึงจะคงที่ การทำให้ผลต่างไม่ควรทำหลายครั้งมากเกินความจำเป็น เพราะจะมีผลทำให้ค่าพยากรณ์มีความคลาดเคลื่อนสูง

ในกรณีที่อนุกรมเวลาไม่คงที่ในความแปรปรวน หรือมีการเคลื่อนไหวเป็นเส้นโค้ง วิธีการแปลงอนุกรมเวลาที่ใช้กันมาก คือ ใส่ *ln* ในอนุกรม Y_t วิธีนี้มักจะใช้เมื่อความแปรปรวนแปรผันตามค่าเฉลี่ย บางกรณีการใช้ *ln* อาจไม่ได้ผล ก็ควรทดลองใช้วิธีอื่น เช่น ใช้วิธีหารากที่สอง

เมื่ออนุกรมเวลามีสภาพไม่คงที่ หรือไม่เคลื่อนไหวรอบค่าเฉลี่ยคงที่ค่าหนึ่งค่าเดียว จะต้องแปลงข้อมูลดังที่กล่าวไปแล้ว ฉะนั้น ถ้ามีการทำผลต่าง *d* ครั้ง จะเขียนตัวแบบผสมเป็น *ARIMA(p,d,q) (Autoregressive Integrated Moving Average Model)* ซึ่งมีรูปแบบทั่วไป :

ตัวแบบ *ARIMA(p,d,q) :*

$$\phi_p(B)(1-B)^d Y_t = \delta + \theta_q(B)\alpha_t \qquad (2.23)$$

หรือ

$$\phi_p(B)W_t = \delta + \theta_q(B)\alpha_t \qquad (2.24)$$

ซึ่งให้ $W_t = (1-B)^d Y_t$ และ δ (อาจมีค่าเท่ากับศูนย์) เป็นพารามิเตอร์แสดงระดับค่าเฉลี่ยคงที่ของอนุกรม W_t

ตัวอย่าง *: ARIMA(p,d,q)* เมื่อ *p=1,d=1,q=1* มีรูปแบบดังนี้

$$(1-\phi_1 B)(1-B)Y_t = \delta + (1-\theta_t B)\alpha_t$$

หรือ

$$Y_t = \delta + (1+\phi_1)Y_{t-1} - \phi_2 Y_{t-2} + \alpha_t - \theta_1\alpha_{t-1}$$

หรือ

$$W_t + \delta + \phi_1 W_{t-1} + \alpha_t - \theta_1\alpha_{t-1}, W_t = Y_t - Y_{t-1}$$

3. ตัวแบบ *ARIMA* เมื่อมีองค์ประกอบฤดูกาล

ถ้าอนุกรมเวลาที่พิจารณามีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง จะได้ว่าลักษณะของอนุกรมเวลาประเภทนี้ เปลี่ยนแปลงขึ้นลงเลียนแบบกันตามช่วงเวลา เราเรียกรูปแบบอนุกรมเวลาชนิดนี้ว่า รูปแบบอนุกรมเวลาที่มีฤดูกาล *(Seasonal Time Series Model)* โดยรูปแบบอนุกรมเวลาจะคงลักษณะของตัวแบบที่ไม่ใช่ฤดูกาล นั่นคือ จะมีตัวแบบ *ARIMA* ด้วยอันดับ *(P,Q,D),* ซึ่ง *P* คือ อันดับในส่วนของกระบวนการ *AR,Q* คือ อันดับในส่วนของกระบวนการ *MA,* และ *D* คือ จำนวนครั้งทำผลต่างอนุกรมเวลาห่างกัน *s* คาบเวลา

เมื่อนำองค์ประกอบในส่วนที่ไม่ใช่ฤดูกาล และส่วนที่เป็นฤดูกาลมาผนวกเข้าด้วยกันจะได้ตัวแบบ *ARIMA* ที่แสดงส่วนประกอบทั้งสอง และตัวแบบทั่วไปตัวแบบหนึ่งคือ ตัวแบบในรูปผลคูณ *ARIMA (p,d,q)(P,D,Q)* มีรูปแบบดังนี้

ตัวแบบ *ARIMA (p,d,q) (P,D,Q)* มีรูปแบบดังนี้

$$\phi_p(B)\Phi_p(B^\delta)(1-B)^d(1-B^\delta)^D Y_t = \delta + \theta_q(B)\Theta_Q(B^\delta)a_t$$

โดยที่ $\Phi_p(B^\delta) = 1 - \Phi_\delta B^\delta - \Phi_{2\delta}B^{2\delta} - \ldots - \Phi_p B^{p\delta}$

$$\Theta_Q(B^\delta) = 1 - \Theta_\delta B^\delta - \Theta_{2\delta}B^{2\delta} - \ldots - \Theta_{Q\delta}B^{Q\delta}$$

ตัวอย่าง *: ARIMA(p,d,q)(P,D,Q)* เมื่อ *p=0,d=1,q=0,P=0,D=1,Q=1* มีรูปแบบดังนี้

ARIMA (0,1,0)(0,1,1) :

$$(1-B)(1-B^4)Y_t = \delta + (1-\Theta_4 B^4)a_t$$

หรือ

$$W_t = \delta - \Theta_4 a_{t-4} + a_t \ , \ W_t = (1-B)(1-B^4)Y_t$$

ARIMA $(1,1,1)(1,1,0)_{12}$ *:*

$$(1-\phi_1 B)(1-\Phi_{12}B^{12})(1-B)(1-B^{12})Y_t = \delta + (1-\theta_1 B)a_t$$

หรือ

$$W_t = \delta + a_t + \phi_1 W_{t-1} + \Phi_{12} W_{t-12} - \phi_1 \Phi_{12} W_{t-13} - \theta_1 a_{t-1}$$

โดยที่ $W_t = (1-B)(1-B^{12})Y_t$

ขั้นตอนวิธีสร้างตัวแบบบอกซ์-เจนกินส์ หรือตัวแบบ *ARIMA*

1. การกำหนดตัวแบบ *ARIMA (p,d,q)(P,D,Q)*

การกำหนดตัวแบบ *ARIMA* จะต้องพิจารณากำหนดอันดับ *p,d,q* และจะต้องกำหนดอันดับ *P,D,Q* และ *s* ด้วย ถ้าตรวจสอบว่าข้อมูลอนุกรมเวลามีองค์ประกอบฤดูกาลด้วยคาบฤดูกาล *s*

อันดับ *p* และ *q* คือ อันดับของกระบวนการ *AR* และ *MA* ในส่วนที่ไม่ใช่ฤดูกาล และ *P* และ *Q* คือ อันดับของกระบวนการ *AR* และ *MA* ในส่วนที่เป็นฤดูกาล สำหรับ *d* คือ จำนวนครั้งทำผลต่างอนุกรมเวลา เมื่ออนุกรมในส่วนที่ไม่ใช่ฤดูกาลมีสภาพไม่คงที่ในค่าเฉลี่ย

การพิจารณากำหนดอันดับ (p,d,q) และ (P,D,Q) จะพิจารณาแยกจากกัน แต่ใช้หลักการพิจารณาเหมือนกัน

กระบวนการ AR และ MA ต่างมีรูปแบบโครงสร้างเฉพาะสำหรับอันดับ pและ q ของฟังก์ชันอัตสหสัมพันธ์ ACF (Autocorrelation Function) แทนด้วย p_k และโครงสร้างของฟังก์ชันอัตสหสัมพันธ์ย่อย PACF (Partial Autocorrelation Function) แทนด้วย $\emptyset_{kk}$ ซึ่ง k หมายถึงคาบเวลาห่างระหว่างอนุกรม และเรียกคาบเวลานี้ว่า “แล็ก k” (lag k) ฉะนั้น p_1 หมายถึง อัตสหสัมพันธ์ของอนุกรมเวลาที่ห่างกัน 1 หน่วย หรือ 1 คาบเวลา (Y_t, Y_{t+1}), t=1,2,… ซึ่งวัดความสัมพันธ์เชิงเส้นระหว่างอนุกรมเวลาที่ห่างกัน 1 คาบเวลา และ p_2 คือ อัตสหสัมพันธ์ของอนุกรมเวลาที่ห่างกัน 2 หน่วย หรือ 2 คาบเวลา (Y_t, Y_{t+2}), t=1,2,… สำหรับ $\emptyset_{kk}$ เป็นอัตสหสัมพันธ์ของอนุกรมเวลาที่ห่างกัน k หน่วย หรือ k คาบเวลา (Y_t, Y_{t+k}) โดยพิจารณาจากผลกระทบจากอนุกรมเวลา $Y_{t+1}, Y_{t+2}, \dots, Y_{t+k}$ เข้ามาด้วย ค่าของ p_k และ $\emptyset_{kk}$ ต่างมีค่าอยู่ระหว่าง -1 และ 1 ซึ่งจะมีตัวอย่างสูตรของฟังก์ชันเหล่านี้ เช่น

กระบวนการ AR(1) : $Y_t = \delta + \emptyset Y_{t-1} + \alpha_1$

$p_k = \emptyset^k$ สำหรับ k = 0 , 1 , 2 , …

$$\emptyset_1 = p_1 = 0, \emptyset_{kk} = 0 \qquad \text{สำหรับ } k = 2, 3, \ldots$$

กระบวนการ MA(1) : $Y_t = \mu + \alpha_t - \emptyset\alpha_{t-1}$

$$p_1 = \frac{-\theta}{1+\theta^2}, p_k = 0 \qquad \text{สำหรับ } k = 2, 3, \ldots$$

$$\emptyset_{kk} = \frac{-\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}} \qquad \text{สำหรับ } k=1,2,\ldots$$

เพราะฉะนั้นการกำหนดอันดับ จะประมาณค่า p_k และ $\emptyset_{kk}$ โดยใช้ข้อมูลอนุกรมเวลาที่นำมาวิเคราะห์ แทนค่าประมาณด้วย $\hat{p}_k$ และ $\widehat{\emptyset_{kk}}$ และเรียกว่า “ฟังก์ชันอัตสหสัมพันธ์ตัวอย่าง” SACF(Sample Autocorrelation Function) และ “ฟังก์ชันอัตสหสัมพันธ์ย่อยตัวอย่าง” SPACF (Sample Partial Autocorrelation Function)

ค่าประมาณ $\hat{p}_k$ และ $\widehat{\emptyset_{kk}}$ ซึ่งจะคำนวณค่าโดยใช้ข้อมูลอนุกรมเวลา มีสูตรทั่วไปดังนี้

$$\hat{p}_k = \frac{\sum_{t=1}^{n-k}(y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2} \text{ สำหรับ } k = 0,1,2,\ldots$$

$$\widehat{\emptyset}_{11} = \hat{p}_1$$

$$\widehat{\emptyset}_{kk} = \frac{\hat{p}_k - \sum_{j=1}^{k-1}\widehat{\emptyset}_{k-1,i}(\widehat{p_{k-j}})}{1 - \sum_{j=1}^{k-1}\widehat{\emptyset}_{k-1,j}(\hat{p}_j)} \text{ สำหรับ } k = 2,3,\ldots$$

ซึ่ง $\widehat{\emptyset}_{k,j} = \widehat{\emptyset}_{k-1,j} - \widehat{\emptyset}_{kk}\widehat{\emptyset}_{k-1,k-j} \; (k = 3,4,\ldots; j = 1,2,\ldots,k-1)$

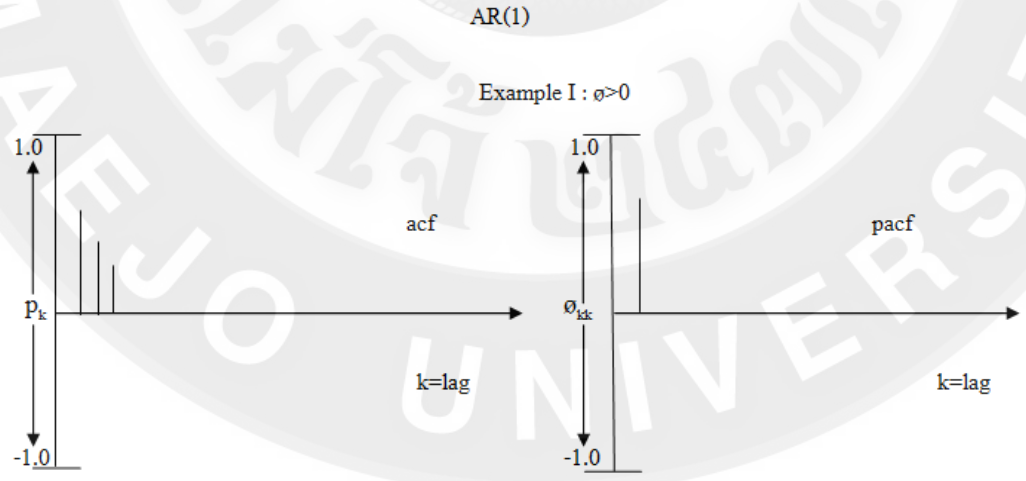
Y_t = *ข้อมูลที่เวลา* t

สรุปลักษณะการแปรผันของ ACF และ PACF ของกระบวนการอนุกรมเวลาคงที่สำหรับกระบวนการพื้นฐานดังนี้

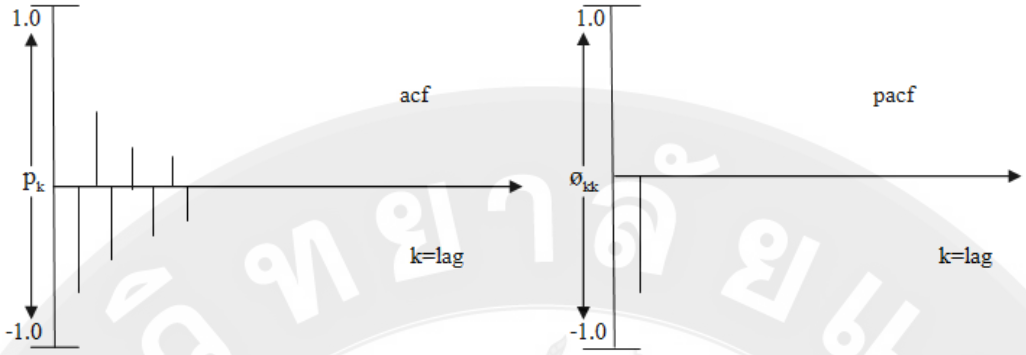
ตารางที่ 2.1 แสดงลักษณะ ACF และ PACF สำหรับตัวแบบ ARMA ต่างๆ

ตัวแบบ	ACF	PACF
AR(1)	ค่า p_k ลดลงอย่างรวดเร็ว ขณะที่ K>1	ค่า $ø_{kk}$ จะมีค่าสูงที่ k=1 และเท่ากับ 0 เมื่อ K>1
AR(2)	ค่า p_k ลดลงอย่างรวดเร็ว ขณะที่ K>1	ค่า $ø_{kk}$ จะมีค่าสูงที่ k=1,2 และเท่ากับ 0 เมื่อ K>2
MA(1)	ค่า p_k จะมีค่าสูงที่ k=1 และเท่ากับ 0 เมื่อ K>1	ค่า $ø_{kk}$ ลดลงอย่างรวดเร็ว ขณะที่ K>1
MA(2)	ค่า p_k จะมีค่าสูงที่ k=1,2 และเท่ากับ 0 เมื่อ K>2	ค่า $ø_{kk}$ ลดลงอย่างรวดเร็ว ขณะที่ K>1
ARMA(1,1)	ค่า p_k ลดลงอย่างรวดเร็ว หลังจาก k=1	ค่า p_k ลดลงอย่างรวดเร็ว หลังจาก k=1

และมีรูปแบบความสัมพันธ์ (เชิงทฤษฎี) ของ AFC และ PACF ของกระบวนการอนุกรมเวลาคงที่ สำหรับกระบวนการพื้นฐาน แสดงเป็นกราฟคอเรลโลแกรมได้ดังนี้

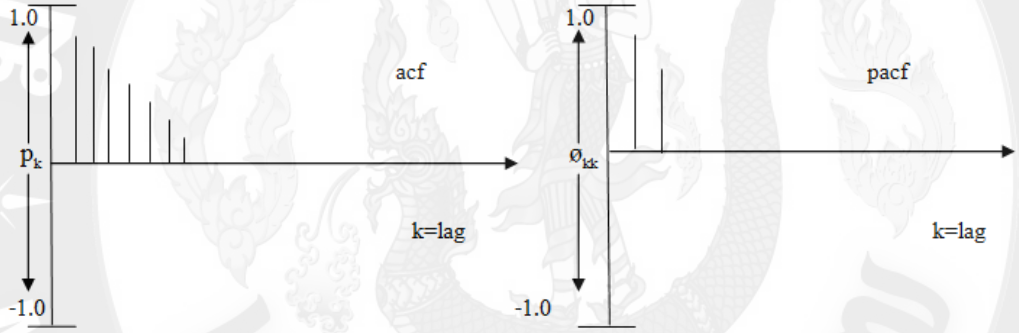


Example II : ø>0

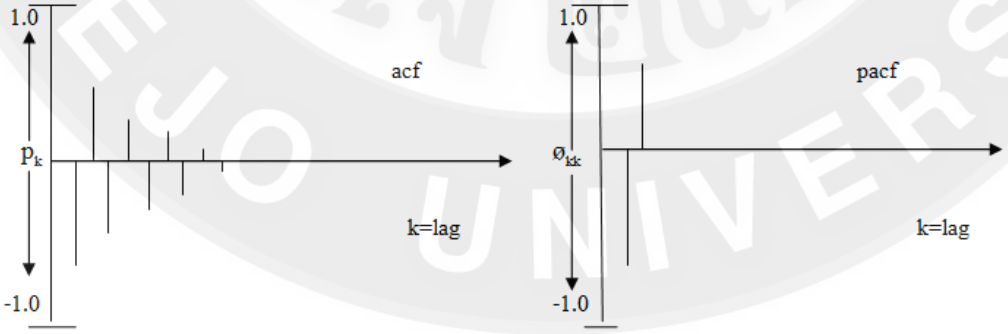


AR(2)

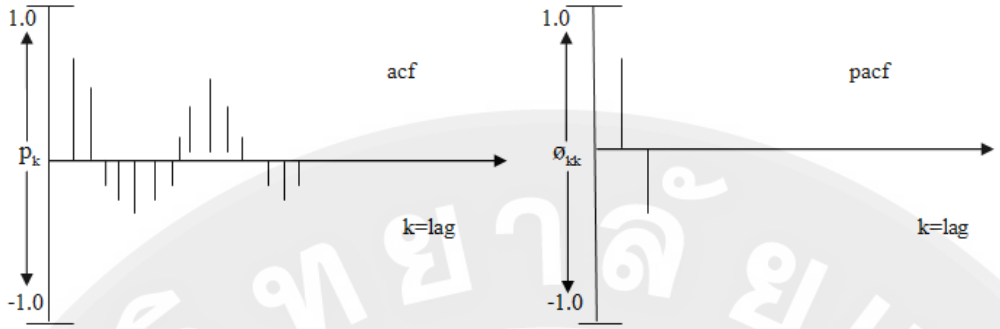
Example I



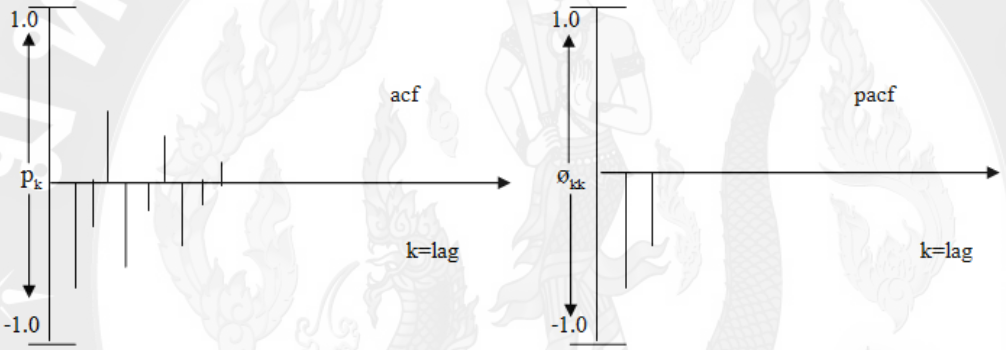
Example II



Example III

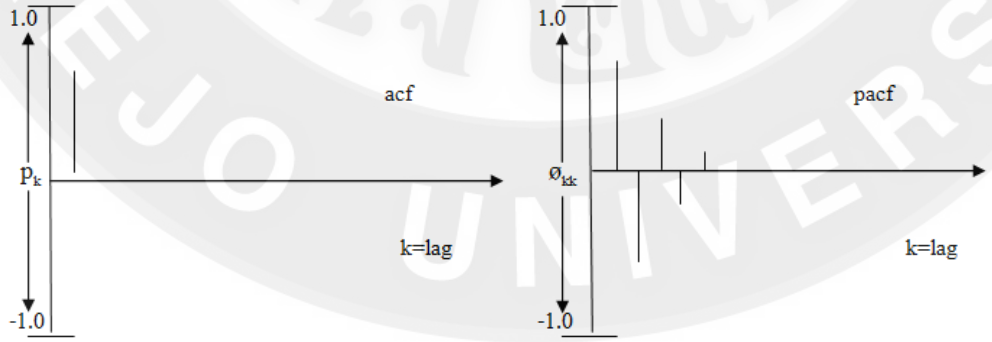


Example IV

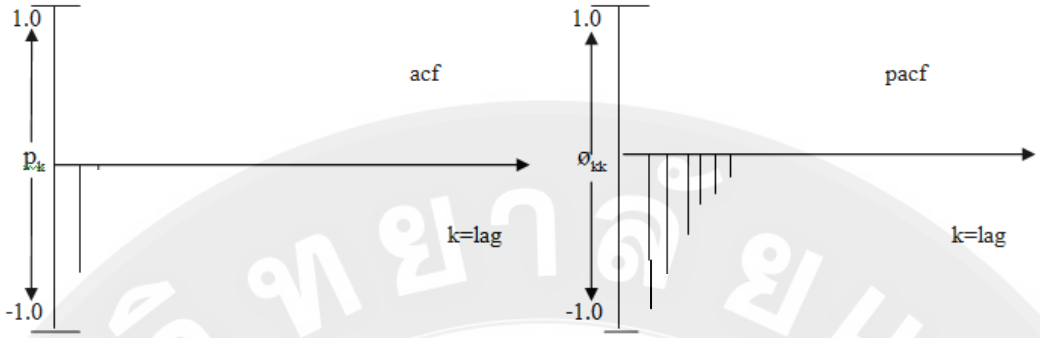


MA(1)

Example I : ø>0

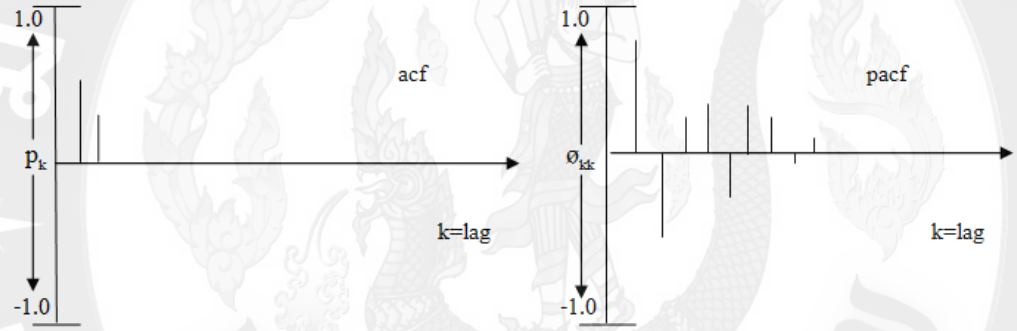


Example II

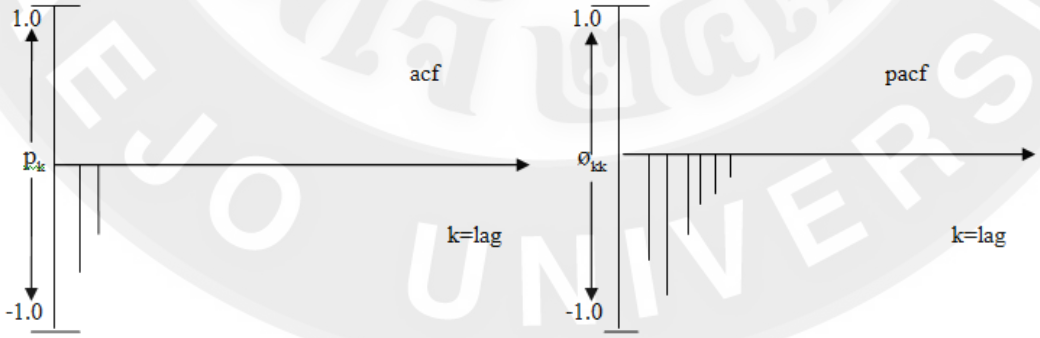


MA(2)

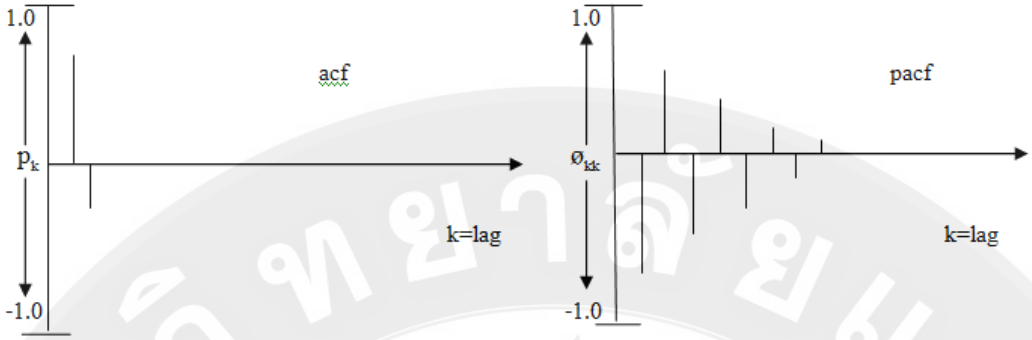
Example I



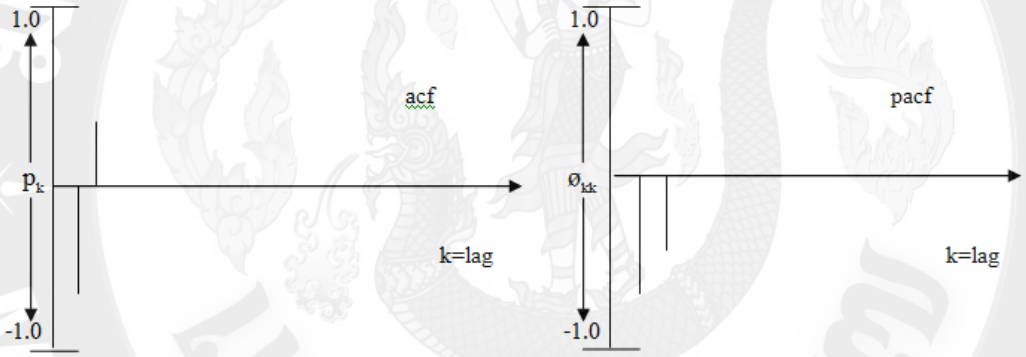
Example II



Example III

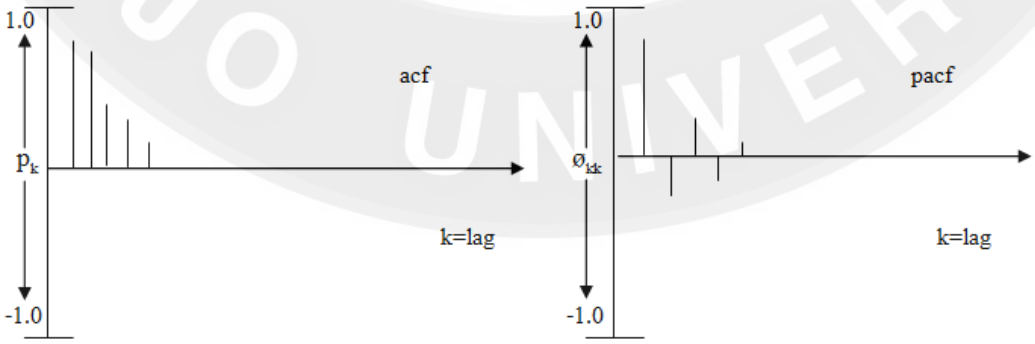


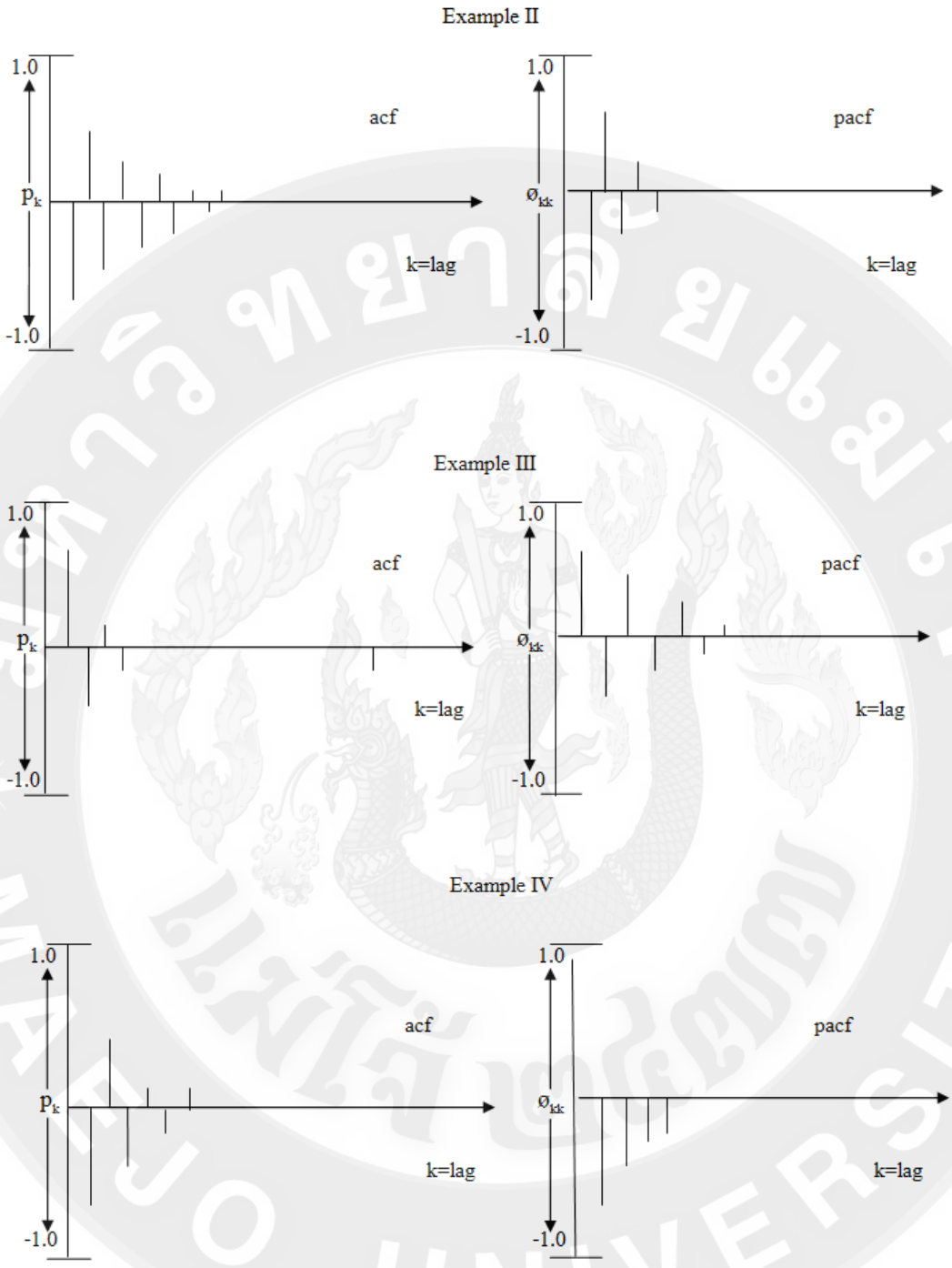
Example IV



AR(1,1)

Example I





ในองค์ประกอบที่เป็นฤดูกาลมีคาบฤดูกาล S การกำหนดอันดับ P และ Q พิจารณาทำนองเดียวกันกับองค์ประกอบที่ไม่เป็นองค์ประกอบที่ไม่เป็นฤดูกาล โดยพิจารณาโครงสร้างแปรผันของอัตสหสัมพันธ์ $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ ที่เล็ก ฤดูกาล s, 2s, 3s, เปรียบเทียบโครงสร้าง ρ_k และ ϕ_{kk} ทางทฤษฎีซึ่งมีลักษณะตามที่กล่าวไปแล้วข้างต้น

2. การประมาณค่าพารามิเตอร์ในตัวแบบ ARIMA

เมื่อนักพยากรณ์เลือกตัวแบบ ARIMA ทดลองได้แล้ว ขั้นต่อไปคือ ประมาณค่าของพารามิเตอร์ที่ปรากกฏในตัวแบบ วิธีการประมาณค่าพารามิเตอร์ที่นิยมใช้กันมาก คือ วิธีกำลังสองน้อยที่สุดแบบไม่เป็นเชิงเส้น (nonlinear least – squares method)

เมื่อใช้โปรแกรมคอมพิวเตอร์ในเรื่อง ARIMA (ซึ่งในการศึกษาครั้งนี้ใช้โปรแกรม SPSS for Windows) ในขั้นประมาณค่าพารามิเตอร์ นอกจากจะได้ค่าประมาณของพารามิเตอร์แล้ว จะมีค่าของตัวสถิติต่างๆ ปรากฏออกมาด้วย ซึ่งให้ประโยชน์ในการทดสอบเชิงสถิติว่า องค์ประกอบหรือพารามิเตอร์นั้นควรมีอยู่ในตัวแบบหรือไม่ ซึ่งนั่นคือ หนทางหนึ่งในการพิจารณาว่า ตัวแบบที่พิจารณานั้นเหมาะสมเพียงพอหรือไม่ในเชิงสถิติ

3. การวินิจฉัยตัวแบบ ARIMA

ภายหลังที่ประมาณค่าพารามิเตอร์ของตัวแบบ ARIMA แล้วนักพยากรณ์ควรจะตรวจสอบตัวแบบก่อนที่จะตัดสินใจนำตัวแบบนั้นไปใช้พยากรณ์ เนื่องจากตัวแบบที่พิจารณาคัดเลือกในขั้นแรกนั้น อาจยังเลือกไม่ถูกต้องเหมาะสม จึงควรวินิจฉัย และถ้าพบว่ายังไม่เหมาะสมควรกลับไปขั้นที่ 1 พิจารณาปรับปรุงแก้ไขตัวแบบใหม่ และดำเนินการขั้นที่ 2 ประมาณค่าและวินิจฉัยในขั้นที่ 3 กรรมวิธีจะกระทำซ้ำๆ เช่นนี้จนกว่าจะตัวแบบที่เหมาะสมเพียงพอในเชิงสถิติ

การวินิจฉัยตัวแบบ จะทำการตรวจสอบคุณสมบัติเชิงสถิติของค่าผิดพลาดสุ่ม $\{ a_t \}$ และการทดสอบว่า ค่าผิดพลาดสุ่มมีอัตสหสัมพันธ์หรือไม่ จะเป็นการตรวจสอบที่สำคัญมากที่สุดในการวินิจฉัยความเพียงพอในเชิงสถิติของตัวแบบ ARIMA ฉะนั้นการทดสอบ จะคำนวณค่า SACF และ SPACF ของค่าเศษเหลือตกค้าง $e_t = y_t - \hat{y}_t$ ซึ่งเป็นค่าประมาณของ a_t ที่เล็ก k ต่างๆ และทดสอบด้วยค่าของตัวสถิติที (t) สำหรับทดสอบว่า ค่าผิดพลาดสุ่มมีอัตสหสัมพันธ์หรือไม่ที่แต่ละแล็ก k = 1,2,3,…,m และทดสอบอัตสหสัมพันธ์รวมหรือพร้อมกัน k แล็ก ด้วยตัวสถิติไคกำลังสอง (Chi – squared test) ว่าค่าผิดพลาดไม่มีอัตสหสัมพันธ์ k แล็กแรก นอกจากการวินิจฉัยตัวแบบด้วยการทดสอบเชิงสถิติแล้ว นักพยากรณ์อาจตรวจสอบด้วยวิธีการอื่นๆ ด้วยเช่น การเขียนกราฟของเศษเหลือตกค้างกับแกนเวลา ถ้าพบว่าของเศษเหลือตกค้างกระจายเป็นแนวในลักษณะขนานรอบค่าเฉลี่ยศูนย์ แสดงเหตุผลได้ว่า ค่าผิดพลาดมีค่าเฉลี่ยศูนย์ และมีความแปรปรวนคงที่ แต่ถ้าการกระจายของค่าเศษเหลือตกค้างมีรูปแบบต่างไปจากแนวขนาน ควรพิจารณาปรับปรุงแก้ไขตัวแบบซึ่งอาจจะพบว่าความแปรปรวนยังไม่คงที่ด้วยวิธีการแปลงข้อมูล เป็นต้น

ผลจากการวินิจฉัยตัวแบบ นอกจากจะช่วยตรวจสอบว่าตัวแบบที่กำลังพิจารณาเหมาะสมเพียงพอในเชิงสถิติหรือไม่แล้ว ยังให้แนวทางในการปรับปรุงแก้ไขตัวแบบด้วย ถ้าพบว่าตัวแบบยังไม่เหมาะสม กล่าวคือ จากลักษณะของ SACF และ SPACF ของเศษเหลือตกค้าง อาจพบว่า ควร

เพิ่มองค์ประกอบ MA เข้าในตัวแบบ ถ้ายังไม่มีองค์ประกอบ MA หรือ เพิ่มอันดับของ MA ให้มากขึ้น หรืออาจ พบว่าควรเพิ่มองค์ประกอบ AR หรืออันดับของ AR ในตัวแบบ เป็นต้น

2.5 วิธีการปรับให้เรียบแบบเอ็กซโพเนนเชียล (Exponential Smoothing Method)

วิธีการปรับให้เรียบแบบเอ็กซโพเนนเชียล เป็นอีกวิธีการหนึ่งของการวิเคราะห์ข้อมูลอนุกรมเวลา เพื่อสร้างตัวแบบพยากรณ์ วิธีการนี้ใช้หลักการหาค่าเฉลี่ยถ่วงน้ำหนัก (Weighted Average) ของข้อมูลในอดีต โดยให้น้ำหนัก (ความสำคัญ) มากสุดกับข้อมูลปัจจุบัน และจากนั้น ให้น้ำหนักลดลงเรื่อยๆ กับข้อมูลในอดีต (เนื่องจากมีแนวคิดว่าผลกระทบของข้อมูล หรือค่าสังเกตปัจจุบันที่มีต่อค่าในอนาคต จะมากกว่า ผลกระทบของข้อมูลในอดีต) ยิ่งข้อมูลถอยหลังไปมากๆ จะมีน้ำหนักลดลงมาก ซึ่งลักษณะการให้น้ำหนัก จะให้น้ำหนักลดลงแบบเรขาคณิต (Geometric) หรือแบบเอ็กซโพเนนเชียล (Exponential) น้ำหนัก จะมีค่าอยู่ระหว่าง 0 และ 1 และรวมกันเท่ากับ 1 [ตัวอย่างเช่น $Y_t = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + ..., (0 < \alpha < 1)$]

1. วิธีการปรับให้เรียบครั้งเดียวแบบเอ็กซโพเนนเชียล (Simple Exponential Smoothing Method)

เป็นวิธีการหนึ่งในกลุ่มวิธีการปรับให้เรียบแบบเอ็กซโพเนนเชียล วิธีการนี้เหมาะที่จะใช้กับอนุกรมเวลาที่มีระดับค่าเฉลี่ยไม่คงที่ โดยมีระดับเคลื่อนไหวช้าๆ ไม่มีแนวโน้ม ไม่มีวัฏจักร และไม่มีฤดูกาล

การพยากรณ์ค่าของ Y ที่เวลา t+1 จากตำแหน่งเวลาปัจจุบัน t ด้วยวิธีการปรับให้เรียบครั้งเดียวแบบเอ็กซโพเนนเชียล มีสูตรพยากรณ์ ดังนี้

$$F_t = \alpha Y_t + (1-\alpha)F_{t-1} \qquad \text{สำหรับ t=1, 2, …}$$

โดยที่

F_t = ค่าพยากรณ์สำหรับค่าของ Y ที่เวลา t+1 จากเวลาปัจจุบัน t, 1 = 1,2,3,…

F_{t-1} = ค่าพยากรณ์สำหรับค่าของ Y ที่เวลา t-1+1 จากเวลาปัจจุบัน t-1, 1 = 1,2,3,…

Y_t = ค่าสังเกต (ค่าจริง) ของ Y ที่เวลา t

α = ตัวประกอบปรับให้เรียบ (Smoothing Factor) มีค่าอยู่ระหว่าง 0 กับ 1

จากสูตรพยากรณ์ข้างต้น พบว่าการหาค่าพยากรณ์ F_t จะต้องทราบค่า F_{t-1} จะต้องทราบค่า F_{t-2} … จะต้องทราบค่า F_t และสุดท้ายจะต้องทราบค่า F_0 เมื่อทราบค่าเริ่มต้น F_0 จะหาค่า $F_1, F_2, ..., F_{t-1}$ และสุดท้ายหาค่า F_t ได้ โดยใช้สูตรข้างต้นต่อเนื่องกัน หนทางหนึ่งในการ

กำหนดค่าเริ่มต้น F_0 คือ ให้ F_0 เท่ากับค่าเฉลี่ยของข้อมูลอนุกรมเวลา $Y_1, Y_2, ..., Y_t$ ที่มีอยู่ของ Y นั่นคือ ให้ $F_0 = \frac{1}{t}(Y_1 + Y_2 + ... + Y_t)$

อีกหนทางหนึ่งในการเลือกค่า α คือทดลองแปรเปลี่ยนค่า α เช่น เริ่มจาก α=0.01 ต่อไปเป็น 0.02, 0.03,… และแต่ละค่า α คำนวณค่า F_t และหาค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสอง (MSE) จากนั้น เปรียบเทียบค่า MSE ทั้งหมด และเลือกค่า α ที่ให้ MSE ต่ำสุด

2. วิธีการปรับให้เรียบสองครั้งแบบเอ็กซโพเนนเชียล (Double Exponential Smoothing Method)

จากแนวคิดการปรับให้เรียบครั้งเดียวแบบเอ็กซโพเนนเชียล นำมาขยายผลใช้กับข้อมูลอนุกรมเวลาที่มีแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วงเวลา T วิธีนี้เหมาะสำหรับข้อมูลที่มีการเคลื่อนไหวแบบแนวโน้มเชิงเส้นตรง (linear trend data) และไม่มีการเคลื่อนไหวแบบฤดูกาล (seasonal data) เหมาะสมกับการพยากรณ์ในระยะสั้นจนถึงการพยากรณ์ในระยะปานกลางวิธีนี้มีสูตรการพยากรณ์ค่าจริง T_{t+1} ที่เวลา $T+1$ จากเวลาปัจจุบัน T ดังนี้

$$\hat{Y}_T(l) = \left(2 + \frac{\alpha l}{1-\alpha}\right) S_T^{[1]} - \left(1 + \frac{\alpha l}{1-\alpha}\right) S_T^{[2]}$$

$$S_T^{[1]} = \alpha Y_t + (1-\alpha) S_{T-1}^{[1]}$$

$$S_T^{[2]} = \alpha S_T^{[1]} + (1-\alpha) S_{T-1}^{[2]}$$

การคำนวณ $S_T^{[1]}$ และ $S_T^{[2]}$ ต้องการทราบค่า $S_{T-2}^{[1]}, S_{T-2}^{[2]}, S_{T-3}^{[1]}, S_{T-3}^{[2]}, ..., S_0^{[1]}, S_0^{[2]}$ และดังนั้นต้องเริ่มด้วยค่า $S_0^{[1]}$ และ $S_0^{[2]}$ เราประมาณค่า $S_0^{[1]}$ และ $S_0^{[2]}$ ได้ดังนี้

$$S_0^{[1]} = \hat{\beta}_0 - \frac{(1-\alpha)}{\alpha}\hat{\beta}_1$$

$$S_0^{[2]} = \hat{\beta}_0 - 2\frac{(1-\alpha)}{\alpha}\hat{\beta}_1$$

ซึ่ง

$$\hat{\beta}_0 = \overline{Y} - \frac{(T+1)}{2}\hat{\beta}_1$$

$$\hat{\beta}_1 = \frac{12\sum_{t=1}^{T}(t - (T+1)/2)Y_t}{T^3 - T}$$

3. วิธีพารามิเตอร์สองตัวของโฮลท์ (Holt's Two – Parameter Method)

วิธีการของโฮลท์ มีลักษณะคล้ายกับวิธีการปรับให้เรียบสองครั้งแบบเอ็กซโพเนนเชียล แต่มีลักษณะทั่วไปมากกว่า วิธีนี้เหมาะสำหรับข้อมูลที่มีการเคลื่อนไหวแบบแนวโน้มเชิงเส้น มีสูตรการพยากรณ์ดังนี้

$$\hat{Y}_t = S_t + l\hat{\beta}_t$$

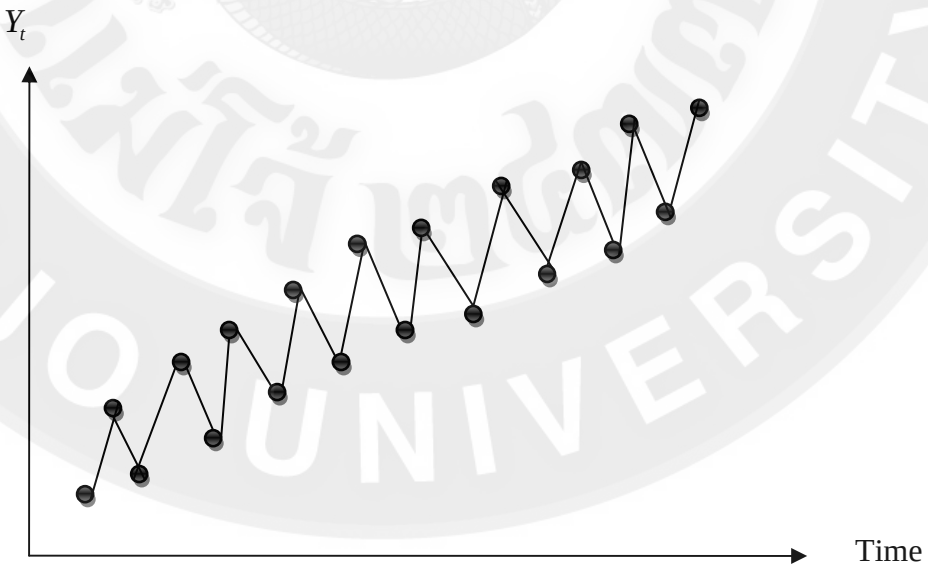
ตัวสถิติปรับระดับ $S_t = \alpha Y_t + (1-\alpha)(S_{t-1} + \hat{\beta}_{t-1})$

ตัวสถิติปรับแนวโน้ม $\hat{\beta}_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\hat{\beta}_{t-1}$

จะเห็นได้ว่าวิธีการของโฮลท์ใช้พารามิเตอร์ปรับให้เรียบสองตัว คือ α ,$(0<\alpha<1)$ และ γ ,$(0<\gamma<1)$ ซึ่งนักพยากรณ์จะต้องกำหนดค่าทั้งสองนี้ และกำหนดค่าเริ่มต้น S_1 และ $\hat{\beta}_1$

4. วิธีการพยากรณ์ของวินเตอร์ (Winters' Forecast Method)

วิธีนี้เหมาะสำหรับข้อมูลที่มีการเคลื่อนไหวมีแนวโน้มและมีฤดูกาล (ดูรูป 2.6) วิธีการพยากรณ์ของวินเตอร์ เป็นการขยายผลของวิธีการของโฮลท์ โดยเพิ่มพารามิเตอร์หรือค่าคงที่ปรับให้เรียบอีกหนึ่งตัวรวมเป็นสามตัวคือ ค่าคงที่ปรับให้เรียบสำหรับระดับ (α_1) ค่าคงที่ปรับให้เรียบสำหรับแนวโน้มหรือความชัน (α_2) และค่าคงที่ปรับให้เรียบสำหรับฤดูกาล (α_3) ค่าทั้งสามมีค่าอยู่ระหว่าง 0 และ 1



รูปที่ 2.6 แนวโน้มและฤดูกาล

ตัวแบบอนุกรมเวลาตัวแบบหนึ่งของวินเตอร์ ซึ่งประกอบด้วยองค์ประกอบแนวโน้มและฤดูกาลมีสมการดังนี้ และมีชื่อเรียกว่าตัวแบบผลคูณของวินเตอร์

$$Y_t = (\mu_t + \beta_t t)I_t + \varepsilon_t$$

ซึ่ง μ_t, β_t, I_t เป็นพารามิเตอร์แสดงระดับ ความชัน และฤดูกาล ของอนุกรมเวลาตามลำดับและ ε_t คือ ความคลาดเคลื่อนสุ่ม โดยมีข้อสมมติพื้นฐานคือ มีค่าเฉลี่ยศูนย์ ความแปรปรวนคงที่ และไม่มีสหสัมพันธ์กัน

ตัวแบบข้างต้นเหมาะสมกับอนุกรมเวลามีการแกว่ง หรือการผันแปรของฤดูกาลเป็นสัดส่วนโดยตรงกับระดับของอนุกรม กล่าวคือ การแกว่งจะมากขึ้นขณะที่ระดับของอนุกรมเพิ่มขึ้น (รูป 2.6) ส่วนการผันแปรของ ε_t ไม่ขึ้นอยู่กับระดับอนุกรม

จากตัวแบบข้างต้น ได้สูตรพยากรณ์ดังนี้

$$\hat{Y}_t(l) = (\hat{\mu} + l\hat{\beta}_t)\hat{I}_{t+l-m}, \qquad t = m, m+l, \ldots$$

ซึ่ง

$$\hat{\mu}_t = \alpha_1(Y_t / \hat{I}_{t-m}) + (1-\alpha_1)(\hat{\mu}_{t-1} + \hat{\beta}_{t-1})$$
$$\hat{\beta}_t = \alpha_2(\hat{\mu}_t - \hat{\mu}_{t-1}) + (1-\alpha_2)\hat{\beta}_{t-1}$$
$$\hat{I}_t = \alpha_3(Y_t / \hat{\mu}_t) + (1-\alpha_3)\hat{I}_{t-m}$$

m = ความยาวของคาบฤดูกาล เช่น m = 12 (สำหรับอนุกรมเวลารายเดือน)
หรือ m = 4 (สำหรับอนุกรมเวลารายสามเดือน)

การคำนวณค่าพยากรณ์ $\hat{Y}_t(l)$ ต้องกำหนดค่าเริ่มต้นของ $\hat{\mu}_t, \hat{\beta}_t, \hat{I}_t$ นอกเหนือจากการกำหนดค่าคงที่ปรับให้เรียบ $\alpha_1, \alpha_2, \alpha_3$ และหนทางหนึ่งในการกำหนดค่าเริ่มต้น คือให้

$$\hat{\mu}_m = (Y_1 + Y_2 + \ldots + Y_m)/m$$
$$\hat{I}_t = Y_t / \hat{\mu}_m, \qquad t = 1,2,\ldots,m$$
$$\hat{\beta}_m = 0$$

2.6 วิธีอัตถดถอย (Autoregressive Method)

วิธีอัตถดถอย เป็นวิธีการที่จัดอยู่ในกลุ่มการวิเคราะห์อนุกรมเวลา เพื่อหาตัวแบบพยากรณ์ค่าของตัวแปร Y ที่ต้องการเช่น ตัวแปร Y ที่เป็นราคาผลผลิตการเกษตร พื้นที่เพาะปลูกผลผลิตการเกษตร เป็นต้น ตัวแบบจะเป็นตัวแบบการถดถอยที่แสดงความสัมพันธ์ระหว่างค่า Y_t

ของ Y ณ เวลา t กับค่าของ Y ในอดีตให้เป็น $Y_{t-1}, Y_{t-2}, ..., Y_{t-k}$ ได้เป็นตัวแบบอัตถดถอย และมีรูปแบบทั่วไปดังนี้

$$Y_t = \theta_0 + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + ... + \theta_k Y_{t-k} + \varepsilon_t$$

ซึ่ง สมมติว่า ค่าคลาดเคลื่อนสุ่ม ε_t มีค่าเฉลี่ยเท่ากับศูนย์ มีความแปรปรวนคงที่ ไม่มีอัตสหสัมพันธ์ และมีการแจกแจงปกติ สำหรับเวลาถอยหลัง k โดยทั่วไป จะเลือกกำหนด $k = 13$ แต่ถ้าข้อมูลอนุกรมเวลา เป็นข้อมูลรายเดือนที่มีฤดูกาล ควรกำหนด $k = 25$ จากตัวแบบข้างต้นได้ตัวแบบพยากรณ์

$$Y_t = c_0 + c_1 Y_{t-1} + c_2 Y_{t-2} + ... + c_k Y_{t-k}$$

ซึ่ง $c_0, c_1, ..., c_k$ เป็นค่าประมาณของ $\theta_0, \theta_1, ..., \theta_k$ ตามลำดับ ทั้งนี้ พิจารณานัยสำคัญของ $c_0, c_1, ..., c_k$ ด้วยซึ่งถ้าค่าใดไม่มีนัยสำคัญ จะตัดค่านั้นออกจากตัวแบบ

2.7 วิธีวิเคราะห์อนุกรมเวลาแบบคลาสสิค (Classical Time Series Analysis)

อนุกรมเวลา $\{Y_t\}$ จะมีการเคลื่อนไหวตามกาลเวลาในรูปแบบต่างๆ เช่นมีแนวโน้มและมีฤดูกาล ฉะนั้น หนทางหนึ่งในการวิเคราะห์อนุกรมเวลา เพื่อหาตัวแบบพยากรณ์ คือ วิเคราะห์หาองค์ประกอบในข้อมูลอนุกรมเวลาที่ศึกษามาสร้างเป็นตัวแบบพยากรณ์

องค์ประกอบของอนุกรมเวลา จำแนกได้เป็น 3 องค์ประกอบหลักคือ องค์ประกอบแนวโน้ม-วัฏจักร (Trend – cycle component) องค์ประกอบฤดูกาล (seasonal component) และองค์ประกอบไม่ปกติ หรือองค์ประกอบส่วนเหลือ (irregular or remainder component) และดังนั้นโดยวิธีนี้ จะทำการแยกองค์ประกอบ ซึ่งได้ตัวแบบสำหรับการพยากรณ์ค่า Y_t และเรียกตัวแบบที่ได้ว่า ตัวแบบอนุกรมเวลาแบบคลาสสิค ซึ่งจำแนกได้เป็นสองตัวแบบคือ

1. ตัวแบบเชิงบวก (Additive Model)

$$Y_t = T_t + S_t + \varepsilon_t$$

2. ตัวแบบเชิงบวก (Additive Model)

$$Y_t = T_t \times S_t \times \varepsilon_t$$

โดยที่

Y_t คือค่าของอนุกรมเวลา ณ คาบเวลา t

ε_t คือค่าคลาดเคลื่อนสุ่ม

T_t คือองค์ประกอบแนวโน้ม-วัฏจักร ณ คาบเวลา t ซึ่งโดยทั่วไปจะอยู่ในรูปแบบพหุนามอันดับต่ำ (low-order polynomial) เช่น $T_t = \beta_0 + \beta_1 t$ และ $T_t = \beta_0 + \beta_1 t + \beta_2 t^2$ เป็นต้น

S_t คือองค์ประกอบฤดูกาล ณ คาบเวลา t ซึ่งอาจจะอยู่ในรูปแบบของตัวแปรบ่งชี้ฤดูกาล (seasonal indicators) เช่น

$S_t = \alpha_1 I_{1,t} + \alpha_2 I_{2,t} + ... + \alpha_{11} I_{11,t}$ ($I_{t,t}$ เป็นตัวแปรบ่งชี้ฤดูกาล)

หรือ อยู่ในรูปแบบฟังก์ชันตรีโกณ เช่น

$$S_t = \phi_1 \sin\left(\frac{2\pi t}{12}\right) + \phi_2 \cos\left(\frac{2\pi t}{12}\right) + \phi_3 \sin\left(\frac{4\pi t}{12}\right) + \phi_4 \cos\left(\frac{4\pi t}{12}\right)$$

ตัวแบบเชิงบวก เหมาะสำหรับข้อมูลอนุกรมเวลาที่มีความแปรปรวนคงที่ ไม่แปรผันตามเวลา หรือการแกว่งของฤดูกาลไม่แปรผันตามระดับของอนุกรมเวลา มิฉะนั้น ตัวแบบเชิงคูณจะเหมาะสมกว่า อย่างไรก็ตาม เราสามารถใช้วิธีการแปลงข้อมูลอนุกรมเวลาที่ไม่คงที่ในความแปรปรวน หรือแปลงตัวแบบเชิงคูณเป็นตัวแบบเชิงบวกได้ โดยการใส่ ln ซึ่งจะได้ตัวแบบ

$$\ln Y_t = \ln T_t + \ln S_t + \ln \varepsilon_t$$

นั่นคือ ใช้ตัวแบบเชิงบวกกับข้อมูลที่ใส่ ln

ตัวอย่างตัวแบบอนุกรมเวลาแบบคลาสสิค :

$$Y_t = b_0 + b_1 t + b_2 I_{1,t} + b_3 I_{2,t} + ... + b_{12} I_{11,t}$$

โดยที่ t แทนคาบเวลา

$I_{t,t} (t = 1,2,...,11)$ แทนตัวบ่งชี้ สำหรับเดือนที่ t ในคาบเวลา t (ให้ $I_{t,t} \approx 1$ สำหรับเดือนที่ i ในคาบเวลา t มิฉะนั้น ให้ $I_{t,t} = 0$)

2.8 ตัวแบบการถดถอยที่มีค่าคลาดเคลื่อนในรูปแบบ AR (Regression Model with AR errors)

จากตัวแบบการถดถอยในรูปแบบทั่วไป

$$Y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + ... + \beta_k x_{k,t} + \varepsilon_t$$

โดยทั่วไป จะมีข้อสมมติข้อหนึ่งที่ว่า ค่าคลาดเคลื่อนสุ่ม ε_t ไม่มีอัตสหสัมพันธ์ (no autocorrelation) แต่ถ้า Y_t เป็นข้อมูลอนุกรมเวลา มักจะพบว่า ε_t มีอัตสหสัมพันธ์ และรูปแบบอัตสหสัมพันธ์รูปแบบหนึ่งที่พบบ่อยเสมอ คือ รูปแบบ AR(1) (first – order autoregressive model) ซึ่งถ้า ε_t มีอัตสหสัมพันธ์ AR(1) จะเขียนตัวแบบการถดถอยข้างต้นได้เป็น

$$Y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + ... + \beta_k x_{k,t} + \varepsilon_t$$

$$\varepsilon_t = \phi\varepsilon_{t-1} + e_t$$

ซึ่งมีข้อสมมติ e_t มีค่าเฉลี่ยเท่ากับศูนย์ มีความแปรปรวนคงที่ ไม่มีอัตสหสัมพันธ์ และมีการแจกแจงปรกติ

ตัวแบบพยากรณ์เมื่อค่าคลาดเคลื่อนสุ่ม ε_t มีอัตสหสัมพันธ์ AR (1) เขียนได้ดังนี้

$$Y_t = b_0 + b_1 x_{1,t} + b_2 x_{2,t} + ... + b_k x_{k,t} + \varepsilon_t$$

$$\varepsilon_t = c\varepsilon_{t-1}$$

โดยที่ $b_0, b_1, b_2, ..., b_k, c$ เป็นค่าประมาณของ $\beta_0, \beta_1, \beta_2, ..., \beta_k$ และ ϕ ตามลำดับ

งานวิจัยที่เกี่ยวข้อง

คติวิช กันธา (2548) ศึกษาการแพร่กระจายของพื้นที่ และปัจจัยที่สนับสนุนและเป็นอุปสรรคในการปลูกยางพารา เปรียบเทียบการยอมรับและปัจจัยที่มีอิทธิพลต่อการยอมรับและไม่ยอมรับการปลูกยางพาราของเกษตรกรบ้านน้ำมิน ตำบลแม่ลาว อำเภอเชียงคำ และบ้านสา ตำบลทุ่งกล้วย กิ่งอำเภอภูซาง จังหวัดพะเยา และศึกษาการประยุกต์ใช้ระบบสารสนเทศภูมิศาสตร์และข้อมูลระยะไกลในการวิเคราะห์หาพื้นที่ที่มีศักยภาพเหมาะสมสำหรับปลูกยางพาราในจังหวัดพะเยา โดยใช้วิธีการศึกษาในเชิงประวัติและเชิงพื้นที่ ข้อมูลที่ใช้ได้รวบรวมข้อมูลทางสถิติของเกษตรกรผู้ปลูกยางพาราในจังหวัดพะเยา ตั้งแต่ปี พ.ศ. 2534 – 2546 และเกษตรกรที่จะปลูกยางพาราในปี พ.ศ. 2547 – 2549 ที่เข้าร่วมโครงการปลูกยางพารากับรัฐบาล เก็บแบบสอบถามจากเกษตรกร 2 กลุ่ม ได้แก่ เกษตรกรที่ปลูกและไม่ได้ปลูกยางพารา และรวบรวมข้อมูลจากกรมส่งเสริมคุณภาพสิ่งแวดล้อม กรมพัฒนาที่ดิน กรมอุตุนิยมวิทยา และกรมป่าไม้ ในการหาพื้นที่ที่มีศักยภาพเหมาะสมสำหรับการปลูกยางพาราของจังหวัดพะเยาในอนาคต ผลการศึกษา พบว่า รูปแบบการแพร่กระจายของพื้นที่ปลูกยางพาราในจังหวัดพะเยาเป็นแบบสุ่ม เนื่องจากพื้นที่ปลูกยางพาราไม่ได้ขยายออกไปยังอำเภออื่นๆ ที่อยู่ติดกับอำเภอเมือง ซึ่งเป็นพื้นที่เริ่มต้นของการปลูกยางพารา แต่มีการขยายออกไปยังกิ่งอำเภอภูซางที่อยู่ไกลกว่า ซึ่งเป็นผลมาจากการยอมรับการส่งเสริมการปลูกยางพาราของผู้นำชุมชนที่มาจากพ่อค้าภายในจังหวัด ส่วนรูปแบบการแพร่กระจายในระดับหมู่บ้านก็พบว่าเป็นแบบสุ่มเช่นเดียวกัน สำหรับปัจจัยที่สนับสนุนการแพร่กระจายของพื้นที่ปลูกยางพาราเป็นปัจจัยทางด้านเศรษฐกิจ เช่น การที่ราคาของผลผลิตพืชชนิดอื่นตกต่ำ และความคาดหวังว่าจะมีรายได้ที่เพิ่มมากขึ้นจากราคายางพาราในอนาคต ส่วนปัจจัยที่เป็นอุปสรรคต่อการแพร่กระจายของพื้นที่ปลูกยางพาราเป็นปัจจัยทางด้านเศรษฐกิจเหมือนกัน เช่น การขาดแคลนที่ดินสำหรับการปลูกยางพารา และที่ดินไม่มีเอกสารสิทธิ์ ส่วนการยอมรับการปลูกยางพารา

หลังจากการรับรู้ข้อมูลข่าวสารครั้งแรกในเรื่องยางพาราภายในจังหวัดพะเยา เกษตรกรบ้านน้ำมินมีระยะเวลาในการยอมรับที่เร็วมากกว่าเกษตรกรบ้านสา แต่แหล่งข้อมูลข่าวสารที่ได้รับครั้งแรก ระดับความรู้ก่อนการปลูกยางพารา และการค้นคว้าข้อมูลเพิ่มเติมหลังการปลูกยางพาราของเกษตรกรทั้ง 2 หมู่บ้านไม่มีความแตกต่างกัน สำหรับอิทธิพลต่อการยอมรับการปลูกยางพาราของเกษตรกรเป็นปัจจัยทางด้านเศรษฐกิจ เช่น ความคาดหวังว่าจะได้รับผลตอบแทนสูงจากราคายางพารา และการมีตลาดรองรับในอนาคต และปัจจัยทางด้านการเมือง/องค์กรการแพร่กระจาย เช่น นโยบายการขยายพื้นที่ปลูกยางพาราของรัฐบาล และการมีหน่วยงานเข้ามาให้ความรู้ในชุมชน ส่วนปัจจัยที่มีอิทธิพลต่อการไม่ยอมรับการปลูกยางพาราของเกษตรกรเป็นปัจจัยทางด้านเศรษฐกิจเช่นกัน เช่น การขาดแคลนเงินทุน และขาดแคลนที่ดินสำหรับการปลูกยางพารา ปัจจัยทางด้านกายภาพ คือ ความลำบากในการเดินทางเข้าไปในพื้นที่ และปัจจัยในระดับบุคคล คือ เกษตรกรไม่ชอบทดลองปลูกพืชใหม่ สำหรับการวิเคราะห์หาพื้นที่ที่มีศักยภาพเหมาะสมสำหรับการปลูกยางพาราในจังหวัดพะเยา ได้ใช้ชั้นข้อมูลสารสนเทศภูมิศาสตร์ 7 ชั้น คือ พื้นที่ป่าไม้ การใช้ประโยชน์ที่ดิน ชุดดิน ปริมาณน้ำฝน ระดับความสูง ความลาดชัน และพื้นที่การเกิดไฟป่า พบว่าพื้นที่ส่วนใหญ่ของจังหวัดพะเยามีความเหมาะสมในการปลูกยางพาราเท่ากับ 1,802,299.59 ไร่ คิดเป็นร้อยละ 46.62 ของพื้นที่จังหวัด ซึ่งอำเภอที่มีศักยภาพเหมาะสมที่สุด ได้แก่ อำเภอเชียงคำ อำเภอเมือง และกิ่งอำเภอภูซาง

จตุพล ทองบุตร (2553) ศึกษาปัจจัยที่มีผลต่อการตัดสินใจปลูกยางพารา และเปรียบเทียบการตัดสินใจปลูกยางพาราของเกษตรกรในจังหวัดศรีสะเกษ จำแนกตาม ระดับการศึกษา รายได้เฉลี่ยต่อคนต่อปี จำนวนแรงงานในครัวเรือน ประสบการณ์ในการปลูกยางพารา และขนาดสวนยางพารา กลุ่มตัวอย่างเป็นเกษตรกรที่เข้าร่วมโครงการปลูกยางพารา 1 ล้านไร่ ที่ขึ้นทะเบียนขอรับกล้ายางพารากับสำนักงานกองทุนสงเคราะห์ยางพารา จังหวัดศรีสะเกษ จาก 11 อำเภอ กลุ่มตัวอย่างกำหนดขนาดตามตารางเครจซี่และมอร์แกน (Krejcie and Morgan) ได้กลุ่มตัวอย่างจำนวน 367 คน ใช้การสุ่มแบบชั้นภูมิตามเขตอำเภอ แต่ละชั้นภูมิสุ่มตัวอย่างอย่างง่ายตามสัดส่วน เครื่องมือที่ใช้ในการวิจัยคือแบบสอบถามซึ่งแบ่งเป็น 2 ตอน และสถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ร้อยละ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์สหสัมพันธ์ ผลการวิจัย พบว่า เกษตรกรส่วนใหญ่เป็นเพศชาย ร้อยละ 92.64 เพศหญิง ร้อยละ 7.36 ส่วนใหญ่มีอายุระหว่าง 36 ปี ถึง 45 ปี ร้อยละ 49.05 รองลงมาคืออายุ 46 ปี ถึง 55 ปี ร้อยละ 26.43 และมีสถานภาพสมรสแล้ว ร้อยละ 88.56 รองลงมาคือโสด ร้อยละ6.27 มีระดับการศึกษาชั้นประถมศึกษา ร้อยละ 56.13 รองลงมาคือระดับมัธยมศึกษาตอนต้นและมัธยมศึกษาตอนปลาย ร้อยละ 16.62 มีรายได้มากกว่า 35,000 บาท ร้อยละ 39.78 รองลงมา มีรายได้ 23,000- 35,000 ร้อยละ 34.06 นอกจากนั้นส่วนใหญ่มีสมาชิกในครัวเรือน

4 คน ร้อยละ 26.16 รองลงมาคือ 3 คน ร้อยละ 25.07 มีแรงงานในครัวเรือนส่วนใหญ่ 1 คน ร้อยละ 40.33 รองลงมา 2 คน ร้อยละ 35.15 ส่วนใหญ่มีประสบการณ์ในการปลูกยางพารา 6 -10 ปี ร้อยละ 47.68 รองลงมาคือ 1-5 ปี ร้อยละ 27.25 นอกจากนั้นส่วนใหญ่มีขนาดสวนยางพาราขนาดกลาง (16-30 ไร่) ร้อยละ 70.03 รองลงมาคือมีสวนยางพาราขนาดเล็ก (2-15ไร่) ร้อยละ 18.80 เกษตรกรส่วนใหญ่ร้อยละ 96.19 ได้แหล่งเงินทุนจากธนาคารเพื่อการเกษตรและสหกรณ์(ธกส.) มากที่สุด รองลงมา ได้แก่ สหกรณ์การเกษตรและอื่นๆ คิดเป็นร้อยละ 1.63 เกษตรกรได้รับข่าวสารยางพาราจากผู้นำในท้องถิ่นมากที่สุด คือ ร้อยละ 47.68 รองลงมาได้แก่ เจ้าหน้าที่ส่งเสริม ร้อยละ 37.61 อาชีพเดิมของเกษตรกรส่วนใหญ่ ร้อยละ 48.50 ประกอบอาชีพเกษตรกรรม รองลงมาได้แก่ อาชีพรับราชการ ร้อยละ 27.79 เกษตรกร ร้อยละ 85.01 ปลูกยางพารา พันธุ์ RRIM 600 รองลงมาคือพันธุ์ RRIT 251 ร้อยละ 10.08 ปัจจัยที่มีผลต่อการตัดสินใจปลูกยางพาราของเกษตรกรในจังหวัดศรีสะเกษทั้ง 4 ด้านนั้น มีระดับที่มีผลต่อการตัดสินใจปานกลาง และเมื่อพิจารณาแต่ละด้านพบว่าปัจจัยด้านเทคนิคและด้านกายภาพมีผลต่อการ ตัดสินใจของเกษตรกร จากการวิเคราะห์ความสัมพันธ์ของปัจจัยที่ส่งผลต่อการตัดสินใจปลูกยางพาราใน จังหวัดศรีสะเกษพบว่า ระดับการศึกษา รายได้เฉลี่ย จำนวนสมาชิกในครอบครัว แรงงานในครอบครัว ประสบการณ์ในการปลูกยางพารา ปัจจัยด้านเศรษฐกิจ มีความสัมพันธ์กับขนาดสวนยางพาราอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 การเปรียบเทียบการตัดสินใจปลูกยางพาราของเกษตรกรในจังหวัดศรีสะเกษ จำแนกตาม ระดับการศึกษา รายได้เฉลี่ยต่อคนต่อปี จำนวนแรงงานในครัวเรือน ประสบการณ์ในการปลูกยางพารา และขนาดสวนยางพารา พบว่า ข้อมูลพื้นฐานของเกษตรกร ด้านระดับการศึกษา ไม่มีผลต่อการตัดสินใจปลูกยางพารา

ทรัพย์ไพศาล จันทร์สระบัว (2551) ศึกษาลักษณะด้านสังคม และเศรษฐกิจของเกษตรกรที่ทำสวนยางพารา และองค์ประกอบภายนอกชุมชนอำเภอบ้านกรวด จังหวัดบุรีรัมย์ วิเคราะห์ความแตกต่างของผลกระทบด้านสังคม และด้านเศรษฐกิจต่อเกษตรกร ก่อนและหลังการทำสวนยางพารา และศึกษาวิเคราะห์ความแตกต่างของผลกระทบด้านสังคม กับผลกระทบด้านเศรษฐกิจต่อเกษตรกร หลังการทำสวนยางพารา ประชากรที่ศึกษา ได้แก่ สำนักงานกองทุนสงเคราะห์การทำสวนยางบุรีรัมย์ ที่พ้นสงเคราะห์และเปิดกรีดยางแล้ว ในปี 2540 ถึงปี 2550 อำเภอบ้านกรวด จังหวัดบุรีรัมย์ จำนวน 291 คน ได้มาโดยวิธีการสุ่มตัวอย่างอย่างง่าย การเก็บข้อมูลโดยใช้เครื่องมือที่ผ่านการตรวจสอบความเที่ยงตรงจากผู้เชี่ยวชาญ วิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูปทางสถิติ สถิติที่ใช้ได้แก่ ค่าร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนแบบมาตรฐาน ค่าสูงสุด และค่าต่ำสุด สถิติวิเคราะห์ความแตกต่าง และสถิติวิเคราะห์ความสัมพันธ์อย่างง่าย ผลการศึกษาลักษณะด้านสังคมของเกษตรกร พบว่า ส่วนใหญ่เป็นเพศชาย อายุมากกว่า 50 ปี จบระดับการศึกษาประถมศึกษาปีที่ 6

จำนวนสมาชิกในครัวเรือน 5 คน มีการติดต่อกับเจ้าหน้าท่าแลทุนสงเคราะห์การทำสวนยางจังหวัดบุรีรัมย์ 2-3 ครั้งต่อปี แหล่งที่เปิดรับข่าวสารเกี่ยวกับการทำสวนยางพาราคือ สำนักงานกองทุนสงเคราะห์การทำสวนยางจังหวัดบุรีรัมย์ มีความรู้ความเข้าใจในระบบก้อนทำสวนยางพาราร้อยละ 76.2 และมีประสบการณ์ในการทำสวนยางพาราร้อยละ 63 และเข้าฝึกอบรมการปลูก การดูแลรักษาและกรีดยางร้อยละ 94 จำนวนเข้าฝึกอบรมคนละ 1-2 ครั้ง ร้อยละ 83.6 เมื่อเกษตรกรมีปัญหาเกี่ยวกับทำสวนยางพาราจะเลือกปรึกษาเจ้าหน้าที่สำนักงานลงทุนสงเคราะห์การทำสวนยางจังหวัดบุรีรัมย์ ร้อยละ 61.5 ด้านเศรษฐกิจของเกษตรกร พบว่า ส่วนใหญ่มีอาชีพหลักทำสวนยางพาราร้อยละ 89.2 อาชีพเสริมทำเกษตรกรรมร้อยละ 4.2 รายได้ครัวเรือนในปี 2549 มากกว่า 100,000 บาท ร้อยละ 58.4 เกษตรกรยังมีหนี้สินเดิมน้อยกว่า 40,000 บาท ร้อยละ 46.2 มีพื้นที่ดินถือครองทางการเกษตรมากกว่า 30 ไร่ ร้อยละ 63.3 พื้นที่สวนยางพารา ร้อยละ 87.4 โรคที่เกิดในสวนยางพาราเป็นโรคเปลือกแห้งร้อยละ 83.6 มีการใช้วิธีการป้องกัน ทายารักษาร้อยละ 64.1 ด้านองค์ประกอบภายนอกภายนอกชุมชน พบว่า การส่งเสริมด้านเงินทุน และสินเชื่อ การส่งเสริมช่วยเหลือด้านการตลาดและการส่งเสริมข่าวสารที่เกี่ยวกับการทำสวน ยางพารา มีความแตกต่างกัน เมื่อเปรียบเทียบก่อนและหลังการทำสวนยางพารา อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ด้านผลกระทบทางสังคมต่อเกษตรกร พบว่า แรงงานในครอบครัวไปทำงานนอกหมู่บ้านแรงงานในครอบครัวกลับมาทำงานที่บ้าน แรงงานรับจ้างที่มาจากถิ่นอื่น โอกาสทางการศึกษาของสมาชิกในครอบครัว โอกาสในการพักผ่อนของผู้ทำสวนยางพารา การไปร่วมกิจกรรมต่าง ๆ ของหมู่บ้าน การไปร่วมงานประเพณีของเพื่อนบ้าน การพบปะพูดคุยกับเพื่อนในหมู่บ้าน การพบปะพูดคุยกับเพื่อนต่างหมู่บ้าน โอกาสไปท่องเที่ยว โอกาสในการไปใช้สิทธิ์ทางการเมือง โอกาสในการดื่มสุรา โอกาสในการเล่นการพนันต่าง ๆ มีความแตกต่างกัน เมื่อเปรียบเทียบก่อนและหลังการทำสวนยางพารา อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 การไปร่วมทำบุญที่วัด และการไปร่วมประชุมในหมู่บ้าน ไม่มีความแตกต่างกันเมื่อเปรียบเทียบก่อนและหลังการทำสวนยางพารา อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ซึ่งไม่มีผลกระทบต่อการทำสวนยางพารา ด้านผลกระทบทางเศรษฐกิจต่อเกษตรกร พบว่า รายได้จากการทำสวนยางพารารายได้ในภาคเกษตรกรรม รายได้นอกภาคเกษตรกรรม รายจ่ายในการทำสวนยางพารา รายจ่ายในภาคการเกษตรกรรม รายจ่ายนอกภาคการเกษตรกรรม การจ้างแรงงานในการทำสวนยางพาราการจ้างแรงงานในการการเกษตรกรรม หนี้สินในการทำสวนยางพารา หนี้สินในภาคเกษตรกรรม หนี้สินนอกภาคเกษตรกรรม หนี้สินในระบบหนี้สินที่ค้างจ่ายด้านทรัพย์สินบางประการของเกษตรกร และประมาณมูลค่าทรัพย์ทั้งหมด มีความแตกต่างกันเมื่อเปรียบเทียบก่อนและหลังการทำสวนยางพารา อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ด้านผลกรทบทางสังคมต่อเกษตรกรกับผลกระทบทางเศรษฐกิจต่อเกษตรกร พบว่า ไปร่วม

กิจกรรมต่าง ๆ ของหมู่บ้านกับ การจ้างแรงงานในภาคการเกษตรกรรม โอกาสในการดื่มสุรากับรายจ่ายในการทำสวนยางพาราโอกาสในการเล่นการพนันต่าง ๆ กับรายจ่ายในการทำสวนยางพารา การไปร่วมประชุมในหมู่บ้านกับรายจ่ายนอกภาคการเกษตรกรรม โอกาสในการไปใช้สิทธิ์ทางการเมืองกับจำนวนหนี้สินที่ค้างจ่ายมีความสัมพันธ์ ไปในทิศทางเดียวกัน เมื่อเปรียบเทียบหลังการทำสวนยางพารา อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โอกาสในการพักผ่อนของเกษตรกรกับรายได้นอกภาคการเกษตรกรรม โอกาสในการไปท่องเที่ยวกับรายได้นอกภาคการเกษตรกรรม การไปร่วมกิจกรรมต่าง ๆ ของหมู่บ้านกับการจ้างแรงงานในการทำสวนยางพาราโอกาสในการไปท่องเที่ยวกับราย จ่ายนอกภาคการเกษตรกรรม โอกาสในการพักผ่อนของเกษตรกรกับจำนวนหนี้ในการทำสวนยางพารา โอกาสในการไปใช้สิทธิ์ทางการเมืองกับจำนวนหนี้ในการทำสวนยางพารา โอกาสในการไปใช้สิทธิ์ทางการเมืองกับจำนวนหนี้สินที่ค้างจ่าย มีความสัมพันธ์ไปในทิศทางตรงกันข้าม เมื่อเปรียบเทียบหลังการทำสวนหนี้สินที่ค้างจ่าย มีความสัมพันธ์ไปในทิศทางตรงกันข้าม เมื่อเปรียบเทียบหลังการทำสวนยางพารา อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 สำหรับด้านอื่น ๆ ไม่พบความสัมพันธ์

เนตินัย พระไตรยะ (2551) ศึกษาการวิเคราะห์หาแนวขอบเขตพรมแดนเชิงเฟ้นสุ่ม และวิเคราะห์หาประสิทธิภาพการผลิตยางพารา รวมถึงผลกระทบของปัจจัยการผลิตต่อประสิทธิภาพการผลิตยางพาราในพื้นที่ภาคเหนือ การเลือกกลุ่มตัวอย่างใช้วิธีสุ่มตัวอย่างอย่างง่าย และใช้วิธีการศึกษาเป็นวิธีเส้นพรมแดนการผลิตเชิงเฟ้นสุ่ม จากการศึกษาและวิเคราะห์ผลของลักษณะทั่วไปของเกษตรกร พบว่า ส่วนใหญ่มีอายุระหว่าง 46 – 50 ปี ประกอบอาชีพเกษตรกรรมร่วมกับการปลูกยางพารา การศึกษาอยู่ในระดับประถมศึกษา โดยส่วนมากตัดสินใจปลูกยางจากการแนะนำของผู้ที่ปลูกยางมาก่อน และการส่งเสริมของบริษัทเอกชน โดยเริ่มปลูกยางในปี พ.ศ. 2537 พื้นที่ส่วนใหญ่เป็นพื้นที่ปลูกยางพาราไม่เกิน 30 ไร่ เกษตรกรใช้เวลาในการดูแลสวนยางก่อนเปิดกรีดยางประมาณ 6 – 7 ปี ปัจจุบันเกษตรกรเปิดกรีดยางมาแล้ว 5 – 6 ปี ซึ่งส่วนใหญ่ไม่มีประสบการณ์ด้านการปลูก การดูแลรักษา และการกรีดยางมาก่อน และได้เข้ารับการฝึกอบรมในเรื่องดังกล่าว จากการรวมกลุ่มกันเองในท้องถิ่น บริษัทเอกชน และหน่วยงานรัฐที่เกี่ยวข้อง พื้นที่ส่วนใหญ่มีสภาพเป็นเชิงเขาลาดชัน มีการเลือกใช้ระยะปลูกแบบ 3 × 7 หรือ 76 – 80 ต้นต่อไร่ โดยใช้พันธุ์ RRIM600 ในช่วง 1 – 4 ปีแรก จะปลูกพืชอื่นแซมระหว่างแถวยางพารา หลังจากปีที่ 4 ไม่นิยมปลูกพืชแซม สำหรับระยะเวลากรีดยางประมาณ 8 – 9 เดือนในรอบ 1 ปี มีระยะหน้ากรีดยางเฉลี่ย 26 เซนติเมตร ใช้ระบบกรีดแบบ 2 วัน เว้น 1 วันทั้งในฤดูแล้ง และฤดูฝน งดกรีดในวันที่ฝนตก และช่วงผลัดใบของต้นยาง ปริมาณน้ำฝนในรอบปี พ.ศ. 2549 เฉลี่ย 1,160 มิลลิลิตรต่อปี เกษตรกรส่วนใหญ่ใช้ปุ๋ยอินทรีย์และปุ๋ยเคมีร่วมกัน โดยจะใส่ปุ๋ยอินทรีย์อัตรา 180 กิโลกรัมต่อไร่ต่อปี และปุ๋ยเคมีสูตร 30-

5-18 ร่วมด้วยในอัตรา 98 กิโลกรัมต่อไร่ต่อปี ซึ่งจะใส่ปุ๋ยห่างกัน 15 วันตามลำดับ แรงงานที่ใช้ดูแลประมาณ 23 วันต่อไร่ต่อปี ส่วนใหญ่ใช้การถางหญ้ามากกว่าการใช้สารเคมีหรือยาปราบศัตรูพืชในการกำจัดวัชพืช และส่วนใหญ่ต้นยางพาราเป็นโรคราสีชมพูมากที่สุด ส่วนการผลิต พบว่า สมการการผลิตมีรูปแบบฟังก์ชันการผลิตเป็นสมการแบบ Cobb-Douglas และมีสมการความไม่มีประสิทธิภาพทางเทคนิคของผู้ผลิต เมื่อวิเคราะห์ค่าสัมประสิทธิ์สมการยางพารา พบว่า จำนวนเดือนที่กรีดในรอบปี ปริมาณการใช้ปุ๋ยอินทรีย์ และแรงงานที่ใช้ดูแล เป็นปัจจัยที่มีผลกระทบต่อปริมาณผลผลิต โดยมีความสัมพันธ์ในทิศทางเดียวกันกับปริมาณผลผลิต ส่วนการใช้สารเคมีหรือยาปราบศัตรูพืช และโรคของต้นยางพารา มีความสัมพันธ์ในทิศทางตรงข้ามกับปริมาณผลผลิต สำหรับค่าสัมประสิทธิ์ของสมการความไม่มีประสิทธิภาพ พบว่า แรงงานที่ใช้ในการดูแลยางพารามีค่าสัมประสิทธิ์เป็นลบ ส่วนระดับการศึกษา ประสบการณ์ในการดูแล และขนาดของพื้นที่ มีค่าสัมประสิทธิ์เป็นบวก การศึกษาประสิทธิภาพการผลิตของเกษตรกรกลุ่มตัวอย่างคำนวณด้วยวิธี stochastic มีค่าประสิทธิภาพเฉลี่ย เท่ากับ 0.94131 โดยสวนยางพาราส่วนใหญ่มีระดับประสิทธิภาพในการผลิตสูงถึงสูงมาก

บุญมาก พ่ออามาตย์ (2552) ศึกษาสภาพ และแนวทางการผลิต การตลาด และการเพิ่มรายได้ ของเกษตรกรผู้ปลูกยางพาราในอำเภอนิคมคำสร้อย จังหวัดมุกดาหาร จำนวน 155 คน โดยใช้แบบสัมภาษณ์แบบมีโครงสร้างเป็นคำถามทั้งแบบปลายปิด แบบปลายเปิด และแบบมาตราส่วนประมาณค่า 5 ระดับ สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ผลการศึกษา พบว่า สถานภาพทางสังคม และเศรษฐกิจของเกษตรกรในพื้นที่ส่วนใหญ่เป็นเพศชาย อายุระหว่าง 30 - 40 ปี สำเร็จการศึกษาระดับประถมศึกษา มีสมาชิกในครัวเรือน 4 คน มีรายได้จากยางพารามากกว่า 100,000 บาทต่อปี คิดเป็น ร้อยละ 26.45 รองลงมามีรายได้ระหว่าง 50,000 – 60,000 บาทต่อปี คิดเป็นร้อยละ 23.87 สภาพการผลิตยางพาราของเกษตรกร ส่วนใหญ่มีพื้นที่ปลูกยางพาราระหว่าง 10 – 15 ไร่ มีพื้นที่เปิดกรีดแล้วต่ำกว่า 10 ไร่ คิดเป็นร้อยละ 43.87 กรีดมาแล้ว 3 ปี มีแรงงานกรีดยางในครัวเรือน จำนวน 2 คน แรงงานจ้าง จำนวน 1 คน โดยแบ่งค่าจ้างกรีด 60 : 40 ส่วนใหญ่ปลูกยางพันธุ์อาร์อาร์ไอเอ็ม 600 ปุ๋ยเคมีที่ใส่หลังเปิดกรีดยาง ได้แก่สูตร 30 – 5 – 18 แบ่งใส่ 2 ครั้งต่อปี อัตรา 1 กิโลกรัมต่อต้นต่อปี เกษตรกรจะกรีดยางเมื่ออายุ 7 ปี โดยใช้ระบบกรีดครึ่งลำต้นกรีด 2 วันเว้น 1 วัน กรีดในช่วงเวลา 01.00 – 06.00 น. และกรีดน้อยกว่า 15 วันต่อเดือน ได้ผลผลิตระหว่าง 301 – 310 กิโลกรัมต่อไร่ต่อปี คิดเป็นร้อยละ 32.90 รองลงมาได้ผลผลิตระหว่าง 291 – 300 กิโลกรัม และ 281 – 290 กิโลกรัมต่อไร่ต่อปี คิดเป็นร้อยละ 21.93 และ 19.35 ตามลำดับ สภาพการตลาดยางพาราเกษตรกรส่วนใหญ่จะขายยางให้จุดรับซื้อยางในอำเภอในรูปยางแผ่นดิบราคา 40 – 50 บาทต่อกิโลกรัม คิดเป็นร้อยละ 47.09 รองลงมา 51 – 70 บาทต่อกิโลกรัม คิดเป็นร้อยละ 30.96 ส่วนการเพิ่มรายได้ของเกษตรกร

ในพื้นที่จะอยู่ในรูปของการปลูกพืชแซมยาง และพืชร่วมยาง ได้แก่ พืชผัก ข้าวโพด ไม้ดอก ไม้ประดับ และหวาย สำหรับการเลี้ยงสัตว์ในสวนยาง ได้แก่ แพะ แกะ และจิ้งหรีด

พูลศักดิ์ อินทรโยธา (2552) ศึกษาปัจจัยที่มีผลต่อการเข้าร่วมโครงการสงเคราะห์ปลูกแทนของเกษตรกรชาวสวนยางอย่างเป็นระบบ เก็บรวบรวมข้อมูลจากการสัมภาษณ์ตามแบบสอบถาม เกษตรกรเจ้าของสวนยาง จำนวน 265 ราย จากเป้าหมาย 240 ราย คิดเป็นร้อยละ 110.42 ในพื้นที่ทั้งภาคใต้ และภาคตะวันออก ซึ่งเป็นแหล่งปลูกยางเดิมที่ สกย. ให้การส่งเสริมสงเคราะห์ปลูกแทนตามพระราชบัญญัติกองทุนสงเคราะห์การทำสวนยาง ตามมาตรา และมาตรา 8 โดยมีจังหวัดตัวอย่างได้แก่ สุราษฎร์ธานี พังงา สงขลา สตูล จันทบุรี และตราด ผลการศึกษาพบว่า เกษตรกรเจ้าของสวนยางส่วนใหญ่จะเป็นเพศชาย และอยู่ในวัยสูงอายุ ส่วนใหญ่นับถือศาสนาพุทธ มีการศึกษาไม่ค่อยสูงมากนัก และมีสถานภาพสมรส มีอาชีพหลักเป็นเกษตรกรชาวสวนยาง มีรายได้ส่วนตัว และรายได้ในครอบครัวเฉลี่ยต่อเดือนมากกว่า 10,000 บาท ส่วนใหญ่มีสมาชิกในครอบครัว 4 คน และมีบ้านที่อยู่อาศัยเป็นของตัวเอง การเข้าร่วมโครงการสงเคราะห์ปลูกแทนกับ สกย. เกษตรกรส่วนใหญ่รู้จัก สกย. และเคยยื่นสงเคราะห์ปลูกแทน โดยได้รับเงินสนับสนุนในการปลูกยางจาก สกย. ในการรับวัสดุเช่นพันธุ์ยาง เกษตรกรไม่ได้รับจาก สกย. เนื่องจากการมีนโยบายให้เกษตรกรผู้รับการสงเคราะห์ทำการหาพันธุ์ยางมาปลูกเอง เกษตรกรบางส่วนจึงได้ต้นพันธุ์ยางที่มีคุณสมบัติไม่ได้มาตรฐาน และมีราคาสูงกว่าที่ สกย. กำหนด การรับวัสดุปลูกเช่น ปุ๋ยรองก้นหลุมและปุ๋ยบำรุง สกย. ไม่มีปุ๋ยมาจ่ายให้เกษตรกร มีบ้างที่ ต้องรวมกลุ่มกันผลิตเอง ยาปราบวัชพืชและยาป้องกันรักษาโรคและศัตรูยาง เกษตรกรไม่ได้รับจาก สกย. การใส่ปุ๋ยบำรุงต้นยางตามงวดงานเกษตรกรได้ใส่ปุ๋ยตรงตามงวดงานทุกครั้ง มีการซ่อมต้นยางที่ตายไป มีการปราบวัชพืชตามงวดงานทุกครั้ง มีการตัดแต่งกิ่งตามงวดงานทุกครั้งและมีการทำแนวป้องกันไฟในฤดูแล้งของทุกปี และส่วนใหญ่ไม่เคยมีไฟไหม้สวนยาง มีการปลูกพืชแซม เช่น ข้าวไร่ ฯลฯ และปลูกพืชคลุมดินในสวนยาง ส่วนปัจจัยด้านความชัดเจนของโครงการสงเคราะห์ปลูกแทน พบว่า เกษตรกรเจ้าของสวนยางส่วนใหญ่มีความเห็นว่าโครงการสงเคราะห์ปลูกแทนได้ให้ความช่วยเหลือเกษตรกรด้วยความชัดเจนมากอยู่ในระดับสูง การกำหนดขั้นตอนการปฏิบัติขอรับการสงเคราะห์มีความชัดเจนมาก โครงการสงเคราะห์ปลูกแทนมีความสำคัญต่อเกษตรกรชาวสวนยางเป็นอย่างมาก โครงการสงเคราะห์ปลูกแทนมีข้อจำกัดในเรื่องหลักฐานที่ดิน และมีบ้างที่เกษตรกรมีข้อจำกัดในเรื่องการเลือกปฏิบัติของพนักงาน สกย. เกษตรกรเจ้าของสวนยางส่วนใหญ่มีความคิดเห็นว่าการให้ความช่วยเหลือของโครงการปลูกแทนในพื้นที่มีความชัดเจนมาก มีเกษตรกรสนใจเข้าร่วมโครงการจำนวนมาก และได้รับการอนุมัติให้เข้าร่วมโครงการมากเช่นกัน และโครงการสงเคราะห์ปลูกแทนทำให้เกิดการรวมกลุ่มสร้างเครือข่ายของคนในชุมชนและชุมชนใกล้เคียงมาก

เกษตรกรเจ้าของสวนยางส่วนใหญ่ให้เหตุผลที่มีน้ำหนักมากที่สุดได้แก่ การไม่มีกรรมสิทธิ์ในการถือครองที่ดินที่จะนำมายื่นขอรับการสงเคราะห์ ทั้งๆ ที่ได้รับข่าวสารจากการประชาสัมพันธ์ในการยื่นขอรับการสงเคราะห์จาก สกย. เกษตรกรเข้าใจในการยื่นขอรับการสงเคราะห์ รวมทั้งมีความเข้าใจในคำอธิบายของเจ้าหน้าที่ สกย. ในการยื่นขอรับทุนอยู่ในเกณฑ์สูง เหตุผลที่ไม่ได้มายื่นขอรับทุนสงเคราะห์ปลูกแทนก็เพราะว่าในช่วงที่ขอทุนฯ ไม้ยางมีราคาต่ำจึงทำให้ชะลอการตัดสินใจยื่นขอรับทุน และในช่วงขอทุนฯ ราคายางอยู่ในระดับสูง ทำให้ตัดสินใจขอรับทุนสงเคราะห์ปลูกแทนต้องชะลอไว้ก่อนเช่นกัน ส่วนเรื่องการคมนาคมขนส่งไม่สะดวก ไม่ใช่สาเหตุที่ทำให้ไม่เดินทางมายื่นขอทุน ทั้งนี้ผู้นำชุมชน เจ้าหน้าที่จากส่วนราชการอื่น และอบต. ไม่ได้ให้ความช่วยเหลือในการยื่นขอรับการสงเคราะห์ เจ้าหน้าที่ สกย. ให้ความช่วยเหลือในการยื่นขอรับการสงเคราะห์ ประเด็นที่ทำให้เกษตรกรยังไม่ตัดสินใยเข้าร่วมโครงการฯ อีกสาเหตุหนึ่งก็เพราะ สกย. ไม่มีนโยบายจัดหาวัสดุปลูก เช่น พันธุ์ยาง ปุ๋ย ยาปราบวัชพืช ยาป้องกันและรักษาโรคศัตรูพืช และเหตุผลที่เข้าร่วมสงเคราะห์ปลูกแทน เพราะได้เปล่าทั้งเงินและวัสดุปลูก ด้านปัญหาและอุปสรรคในการยื่นขอรับการสงเคราะห์ปลูกแทนกับ สกย. พบว่า

1. ด้านการคมนาคม มีบ้างที่การคมนาคมไม่สะดวก ระยะทางไกล สวนยางตั้งอยู่บนเนินเขา

2. ด้านเอกสารสิทธิ์ในที่ดิน ไม่มีกรรมสิทธิ์ในการถือครองที่ดิน ที่ดินอยู่ในเขตป่าสงวนแห่งชาติ และเขตอุทยานแห่งชาติ

3. ด้านพนักงานของ สกย. มีบ้างที่ผิดนัดในการตรวจสวน คำอธิบายในการยื่นขอรับการสงเคราะห์และการบำรุงรักษาสวนยางยังไม่เป็นที่เข้าใจ

4. ด้านกฎและระเบียบของ สกย. ต้องให้เจ้าของกรรมสิทธิ์ที่ดินไปยื่นขอรับการสงเคราะห์ด้วยตนเอง จะตั้งตัวแทนได้ก็จะต้องได้รับการอนุมัติก่อน

5. ในการยื่นขอรับการสงเคราะห์สิ่งที่เกษตรกรต้องการ เช่น ขอให้ขยายเวลาการให้สงเคราะห์เป็น 7½ ปี เงื่อนไขเกี่ยวกับอายุต้นยางในการขอรับการสงเคราะห์ ปุ๋ยบำรุงให้ใส่ในปริมาณน้อย ปุ๋ยบำรุงให้ซื้อหาเองมีราคาสูง และไม่แน่ใจว่าจะตรงตามสูตร เงินสงเคราะห์ในปัจจุบันยังคงน้อยอยู่ ต้นยางตาย สกย.ไม่มียาป้องกันรักษาโรคและศัตรูยาง

มนฤดี เกิดสมบุญ (2542) ศึกษาเรื่องการพยากรณ์พื้นที่เพาะปลูก ผลผลิต และราคาสินค้าเกษตร ในกรณีข้าวนาปี ข้าวนาปรัง ข้าวโพดเลี้ยงสัตว์ ถั่วเขียว และถั่วเหลือง โดยใช้เทคนิคการพยากรณ์เชิงสถิติ ได้แก่ วิธีการวิเคราะห์การถดถอย วิธี Box-Jenkins วิธีการปรับให้เรียบแบบเอ็กซโพเนนเชียล วิธีอัตถดถอย วิธีการวิเคราะห์อนุกรมเวลาแบบคลาสสิค และเปรียบเทียบกับวิธีของศูนย์สารสนเทศการเกษตร และใช้ค่าเฉลี่ยของเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เป็นเกณฑ์ใน

การเปรียบเทียบวิธีการพยากรณ์ทั้ง 6 วิธี ซึ่งข้อมูลที่นำมาใช้เป็นข้อมูลทุติยภูมิ โดยรวบรวมมาจากศูนย์สารสนเทศการเกษตร กรมเศรษฐกิจการพาณิชย์ สำนักงานคณะกรรมการพัฒนาเศรษฐกิจและสังคมแห่งชาติ สำนักงานสถิติแห่งชาติ และกรมชลประทาน ผลการศึกษาพบว่า จากการเปรียบเทียบตัวแบบพยากรณ์ที่ได้จากวิธีต่างๆ ทั้ง 6 วิธี โดยพิจารณาจากค่า MAPE พบว่า ตัวแบบการถดถอยจะเหมาะกับการพยากรณ์พื้นที่เพาะปลูกข้าวนาปี ข้าวนาปรัง ปริมาณผลผลิตข้าวนาปี ข้าวนาปรัง ตัวแบบ Box-Jenkins จะเหมาะกับการพยากรณ์ราคาข้าวเปลือกเจ้านาปี 5% และตัวแบบอัตถดถอยจะเหมาะกับการพยากรณ์ราคาข้าวเปลือกเจ้านาปรังความชื้น 14 – 15% จากนั้นนำตัวแบบมาพยากรณ์พื้นที่เพาะปลูกและผลผลิตค่าล่วงหน้าอีก 4 ปี คือ ปี พ.ศ. 2540 – 2543 และราคาสินค้าเกษตรล่วงหน้าอีก 2 ปี คือ ปี พ.ศ. 2542 – 2543 จากผลการพยากรณ์ สรุปได้ดังนี้ ข้าวนาปี : ในปี พ.ศ. 2543 พื้นที่เพาะปลูกจะมีแนวโน้มลดลงจากปี พ.ศ. 2542 ประมาณ 1.12% สำหรับปริมาณผลผลิต และราคาข้าวเปลือกเจ้านาปี 5% จะมีแนวโน้มเพิ่มขึ้นจากปี พ.ศ. 2542 ประมาณ 0.74% และ 3.55% ตามลำดับ ข้าวนาปรัง : พื้นที่เพาะปลูก ผลผลิต และราคาข้าวเปลือกเจ้านาปรังความชื้น 14 – 15% จะมีแนวโน้มลดลงจากปี พ.ศ. 2542 ประมาณ 5.61% 0.68% และ 13.67% ตามลำดับ

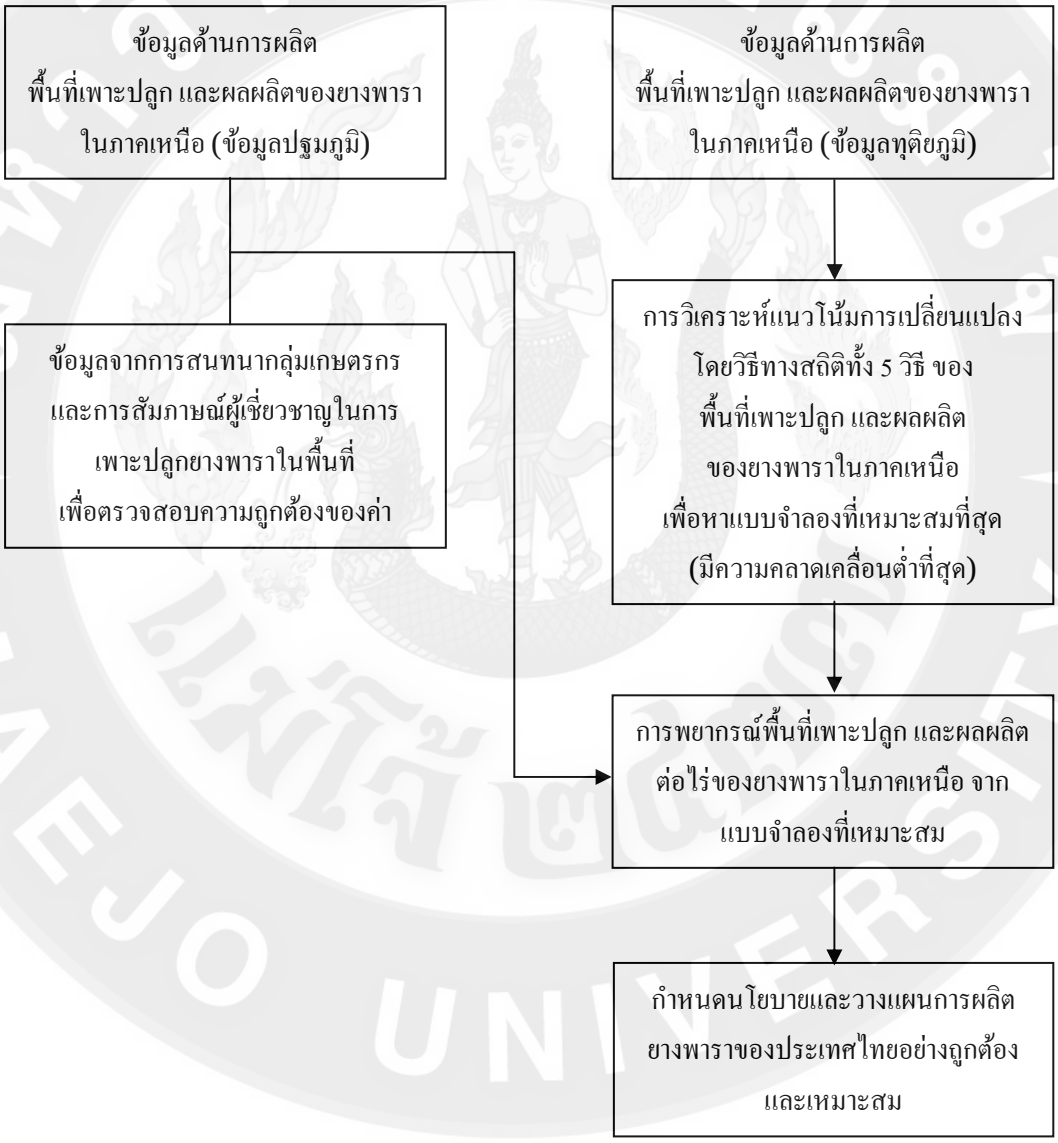
สัตตพงษ์ บุตรโยจันโท (2547) ศึกษาสภาพวิถีชีวิตของครอบครัวเกษตรกรสวนยางพารา ด้านสังคม เศรษฐกิจ ภาวะสุขภาพ และสิ่งแวดล้อม รวมทั้งวิเคราะห์ผลการดำเนินวิถีชีวิตของครอบครัวเกษตรกรสวนยางพารา ด้านสังคม เศรษฐกิจ ภาวะสุขภาพ และสิ่งแวดล้อม จากครอบครัวที่เริ่มปลูกยางพาราเป็นกลุ่มแรกเมื่อ พ.ศ.2532 ในพื้นที่บ้านกกดู่ หมู่ที่ 8 และหมู่ที่10 อำเภอเมือง จังหวัดเลย จำนวน 23 ครอบครัว รูปแบบของการวิจัยเป็นการวิจัยเชิงคุณภาพ ซึ่งใช้เทคนิคการเก็บรวบรวมข้อมูลโดยการศึกษาภาคสนามด้วยการใช้วิธีการสังเกตแบบมีส่วนร่วม แอละไม่มีส่วนร่วม การสัมภาษณ์ผู้ให้ข้อมูล 2 กลุ่ม คือ ผู้นำชุมชน และผู้ที่เกี่ยวข้องกับกลุ่มผู้ปลูกยางพารา รวมทั้งใช้ข้อมูลจากเอกสาร ผลการศึกษาด้านสังคม พบว่า สภาพวิถีชีวิตของเกษตรกรสวนยางพาราเปลี่ยนแปลงไปจากเดิม การดำเนินวิถีชีวิตในเรื่องกิจวัตรประจำวันเปลี่ยนแปลง จากการทำงานกลางวันไปเป็นกลางคืน และต่อเนื่องถึงกลางวัน มีความเป็นอยู่ดีขึ้น มีการปรับปรุงหรือสร้างที่อยู่อาศัยใหม่ มีเครื่องใช้เครื่องอำนวยความสะดวกที่ทันสมัย อาหารที่นำมาบริโภคส่วนใหญ่ได้จากการซื้อ บทบาทและความสัมพันธ์ของสมาชิกในครอบครัวส่วนใหญ่ยังเหมือนเดิม ความสัมพันธ์กับเพื่อนบ้านต่างคุ้มต่างละแวก และต่างหมู่บ้านลดลง มีเวลาเข้าร่วมงานสังคมน้อยลง ด้านเศรษฐกิจ พบว่า การทำสวนยางพารามีรายได้ดีกว่าการปลูกพืชเกษตรอย่างอื่น รายจ่ายของครอบครัวส่วนใหญ่เป็นค่าอาหาร เกษตรกรส่วนใหญ่ไม่มีหนี้สิน บางส่วนที่มีหนี้สินเป็นหนี้สินเพื่อการลงทุน และหนี้สินเพื่อการออม มีการใช้แรงงานจากสมาชิกในครอบครัว และจ้างแรงงานจากคนในหมู่บ้าน ด้านสุขภาพ พบว่า เกษตรกรสวนยางพาราได้ออกกำลัง

จากการกรีดยาง ทำให้ร่างกายแข็งแรง การเจ็บป่วยมีบ้างเล็กน้อย เช่น อ่อนเพลีย ไข้หวัด แพ้น้ำกรดทำแผ่นยาง ด้านสิ่งแวดล้อม พบว่า การทำสวนยางพาราทำให้ป่าไม้ธรรมชาติลดลง แต่เกิดป่าไม้เศรษฐกิจทดแทน มีน้ำเสียที่เกิดจากกระบวนการทำแผ่นยางส่งกลิ่นเน่าเหม็น สรุปผลการทำสวนยางพารา พบว่า ด้านดี ครอบครัวเกษตรกรสวนยางพารามีรายได้ทุกวัน ฐานะทางเศรษฐกิจดีขึ้นเมื่อเทียบกับการปลูกพืชเกษตรชนิดอื่น ส่งผลให้วิถีชีวิตความเป็นอยู่ดีขึ้น การกรีดยางเป็นการออกกำลังกายทำให้ร่างกายแข็งแรงมีรูปร่างสมส่วน นอกจากนี้ยังเป็นกานเพิ่มพื้นที่ป่าสีเขียวทดแทนป่าธรรมชาติในรูปป่าเศรษฐกิจ ส่วนด้านเสีย ได้แก่ วิถีการดำเนินชีวิตในเรื่องกิจวัตรประจำวัน จากเวลากลางวันเป็นกลางคืนต่อเนื่องถึงกลางวัน ทำให้ร่างกายอ่อนเพลีย ความสัมพันธ์กับเพื่อนบ้านต่างคุ้มต่างละแวกหมู่บ้านลดลง ระยะเวลาในการเข้าร่วมงานสังคมลดลง มีค่าใช้จ่ายในการซื้ออาหาร และเครื่องอำนวยความสะดวกเพิ่มมากขึ้น และทำให้สิ่งแวดล้อมเป็นพิษจากน้ำเน่าเสีย

สุชาติ สานกิ่งทอง (2550) ศึกษาข้อมูลพื้นฐานอุตสาหกรรมเกษตรยางพารา และความเป็นไปได้ในการลงทุนอุตสาหกรรมเกษตรยางพารา จังหวัดพิษณุโลก โดยได้ดำเนินการศึกษาตามขั้นตอน 2 ขั้นตอนหลัก ได้แก่ ขั้นตอนที่ 1 ทำการศึกษาข้อมูลพื้นฐานเกี่ยวกับอุตสาหกรรมเกษตรยางพารา ขั้นตอนที่ 2 ศึกษาความเป็นไปได้ในการลงทุนอุสาหกรรมเกษตรยางพารา จังหวัดพิษณุโลก ซึ่งผู้วิจัยได้ศึกษาข้อมูลบริบทอุตสาหกรรมยางพาราจากเอกสาร การสัมภาษณ์ หน่วยงานราชการ และผู้ประกอบการอุตสาหกรรมเกษตรยางพาราทั้งใน และนอกจังหวัดพิษณุโลก กลุ่มเกษตรกรสวนยางพารา จังหวัดพิษณุโลก โดยทำการสัมภาษณ์เชิงลึก ในด้านสภาพแวดล้อมโดยทั่วไปของจังหวัดพิษณุโลก ด้านสภาพการปลูกยางพาราในจังหวัดพิษณุโลก ด้านการตลาด ด้านการผลิต ด้านการเงิน และด้านการบริหารจัดการ เพื่อนำเสนอข้อมูลต่างๆ ที่ได้จากการศึกษามาวิเคราะห์ถึงความเป็นไปได้ในการลงทุนอุตสาหกรรมเกษตร ยางพารา จังหวัดพิษณุโลก ผลการวิจัยพบว่า สภาพพื้นที่ภูมิประเทศและภูมิอากาศจังหวัดพิษณุโลก บริเวณเขตอำเภอนครไทย อำเภอชาติตระการ อำเภอเนินมะปราง และบางส่วนของอำเภอวัดโบสถ์ มีความเหมาะสมในการปลูกยางพารา เนื่องจากมีสภาพภูมิประเทศและภูมิอากาศที่เหมาะสมแก่ยางพารา และในปัจจุบันมีการส่งเสริมการปลูกยางพาราในพื้นที่ดังกล่าวเป็นจำนวน 44,777 ไร่ ซึ่งในอนาคตอีก 5 ปีข้างหน้าจังหวัดพิษณุโลกจะมีศักยภาพในการจัดตั้งโรงงานอุตสาหกรรมเกษตร ยางพาราเชิงพาณิชย์ได้โดยเริ่มต้นที่กำลังการผลิต 40 ตันต่อวัน ราคายางพาราแผ่นดิบรมควันชั้น 3 จะ อยู่ที่ราคา 76 – 78 บาทต่อกิโลกรัม และการลงทุนจัดตั้งโรงงานอุตสาหกรรมเกษตรยางพาราในจังหวัดพิษณุโลก จะใช้เงินทุนประมาณ 72,792,000 บาท และในอีก 10 ปีข้างหน้ากำลังการผลิตจะเป็น 100 ตันต่อวัน โดยราคายางพาราแผ่นดิบรมควันชั้น 3 จะอยู่ที่ราคา 80-90 บาทต่อกิโลกรัม และคาดหวังว่าจะต้องใช้เงินลงทุน 12,399,200 บาท โดยมีอัตราผลตอบแทนการลงทุน 20.90% และระยะเวลาการคืนทุน 8.84 ปี

กรอบแนวคิดของการวิจัย

จากการศึกษาทบทวนวรรณกรรม และทฤษฎีที่เกี่ยวข้องกับพื้นที่เพาะปลูก และผลผลิตของยางพารา รวมถึงปัจจัยที่เกี่ยวข้องต่างๆ ทำให้สามารถสรุปเป็นกรอบแนวความคิดของการศึกษาเรื่องการพยากรณ์พื้นที่เพาะปลูก และผลผลิตต่อไร่ของยางพาราในเขตพื้นที่ภาคเหนือได้ดังนี้



รูปที่ 2.7 กรอบแนวคิดการวิจัย

สมมติฐานการวิจัย

1. แบบจำลองการถดถอยเป็นแบบจำลองที่มีความเหมาะสมสำหรับการพยากรณ์พื้นที่เพาะปลูก และผลผลิตต่อไร่ของยางพาราในเขตพื้นที่ภาคเหนือมากที่สุด เนื่องจากเป็นการวิเคราะห์รูปแบบความสัมพันธ์เชิงเส้นระหว่างตัวแปรตามและตัวแปรอิสระมากกว่าหนึ่งตัว ซึ่งเป็นปัจจัยที่ส่งผลกระทบต่อตัวแปรตาม แบ่งออกเป็น

1.1 ผลผลิตยางพาราต่อไร่ (PRO) ราคายางพาราที่เกษตรกรได้รับ (PRI) ปริมาณน้ำฝน (RAIN) และการส่งเสริมจากรัฐบาลให้ภาคเหนือและภาคตะวันออกเฉียงเหนือเป็นพื้นที่เพาะปลูกยางพาราใหม่ (DM) มีความสัมพันธ์ในทิศทางเดียวกับพื้นที่เพาะปลูกยางพารา (AREA) ส่วนราคาน้ำมันดีเซล (OIL) มีความสัมพันธ์ในทิศทางตรงกันข้ามกับพื้นที่เพาะปลูกยางพารา (AREA) อีกทั้งพื้นที่เพาะปลูกยางพาราอาจขึ้นอยู่กับพื้นที่เพาะปลูกยางพาราในปีที่ผ่านมา

1.2 ปริมาณน้ำฝน (RAIN) จำนวนวันฝนตก (RAINDAY) ความยาวนานแสงเฉลี่ย (SUNSHINE) และความชื้นสัมพัทธ์เฉลี่ย (RH) มีความสัมพันธ์ในทิศทางเดียวกับผลผลิตต่อไร่ของยางพารา (PRO) ส่วนอุณหภูมิเฉลี่ย (TEMP) อุณหภูมิสูงสุดเฉลี่ย (MAXTEMP) และอุณหภูมิต่ำสุดเฉลี่ย (MINTEMP) มีความสัมพันธ์ในทิศทางตรงกันข้ามกับผลผลิตต่อไร่ของยางพารา (PRO) อีกทั้งผลผลิตต่อไร่ของยางพาราอาจขึ้นอยู่กับผลผลิตต่อไร่ของยางพาราในปีที่ผ่านมา

2. พื้นที่เพาะปลูก และผลผลิตต่อไร่ของยางพาราในเขตพื้นที่ภาคเหนือมีแนวโน้มเพิ่มขึ้นเมื่อเปรียบเทียบกับปี พ.ศ. 2554