บทที่ 1

บทนำ

1.1 ความสำคัญและความเป็นมาของปัญหา

ในระบบพยากรณ์ วิธีการพยากรณ์เชิงปริมาณจะมีโครงสร้างอยู่ 2 ส่วน ส่วนแรกคือ การสร้างตัวแบบเพื่อใช้ในการพยากรณ์ ส่วนที่สองคือ การนำตัวแบบที่ได้จากส่วนแรกไปใช้ในการพยากรณ์ ในขั้นตอนของการสร้างตัวแบบเพื่อใช้ในการพยากรณ์นั้น จะต้องเริ่มต้นจากการพิจารณาข้อมูลและทฤษฎีหรือความรู้ต่าง ๆ เพื่อกำหนดตัวแบบ จากนั้นจะประมาณค่าพารามิเตอร์ในตัวแบบแล้วทำการวินิจฉัยตัวแบบว่าเพียงพอ (adequacy) หรือไม่ การวินิจฉัยตัวแบบพยากรณ์ สามารถกระทำได้โดยตรวจสอบคุณสมบัติต่าง ๆ ต่อไปนี้ ได้แก่ ความไม่เอนเอียง (unbias) ของค่าพยากรณ์ ความไม่มีสหสัมพันธ์ (Uncorrelation) ของค่าคลาดเคลื่อน (error) ความคงที่ของความแปรปรวนของค่าคลาดเคลื่อน และการแจกแจงแบบปกติ (Normal distribution) ของค่าคลาดเคลื่อน

ถ้าตัวแบบไม่เพียงพอจะต้องทำการกำหนดตัวแบบใหม่ แล้วทำตามขั้นตอนดังที่กล่าวมาแล้วข้างต้น จนกว่าจะได้ตัวแบบที่เพียงพอ จากนั้นจะนำตัวแบบที่ได้นี้ไปใช้ในขั้นตอนของการพยากรณ์ ซึ่งขั้นตอนนี้จะใช้ตัวแบบที่ได้มาพยากรณ์ค่าในอนาคตตามที่ต้องการ

ค่าที่พยากรณ์ไปแล้วนั้น ย่อมจะแตกต่างจากค่าจริงได้เสมอ ค่าจริงที่เกิดขึ้นจะเป็นข้อมูลใหม่ที่นักพยากรณ์ควรนำมาใช้ตรวจสอบความเที่ยงหรือเสถียร (Stability) ของตัวแบบที่ใช้อยู่โดยการตรวจสอบความเอนเอียงของความคลาดเคลื่อนของค่าพยากรณ์ ซึ่งถ้าหากตัวแบบมีความเที่ยงหรือโครงสร้างของตัวแบบและค่าพารามิเตอร์ยังคงไม่เปลี่ยนแปลง จะได้ว่าค่าคลาดเคลื่อนของการพยากรณ์ e_{n+t} , $t = 1,2,...$ จะมาจากการแจกแจง $N(0,\sigma^2)$ และดังนั้นค่าคลาดเคลื่อนประมาณ 95 เปอร์เซ็นต์ จะตกอยู่ภายในช่วง $(-2\sigma , 2\sigma)$ แต่ถ้าหากโครงสร้างของตัวแบบต้องเปลี่ยนไป การแจกแจงจะเคลื่อนหรือเปลี่ยนแปลงไป โดยเฉพาะค่าเฉลี่ยเปลี่ยนทำให้ค่าความคลาดเคลื่อนจะตกอยู่ภายนอกช่วง $(-2\sigma , 2\sigma)$ ในลัดส่วนที่มากขึ้นในกรณีเช่นนี้ นักพยากรณ์ควรปรับปรุงตัวแบบที่ใช้อยู่เสียใหม่

ถึงแม้ว่าเราจะสามารถใช้วิธีเปรียบเทียบค่าคลาดเคลื่อนกับค่าจำกัด $\pm 2\sigma$ แต่ก็อาจจะใช้เวลานานมากจึงจะพบว่าตัวแบบมีความเที่ยงหรือไม่ วิธีที่ใช้ทั่วไปในการตรวจสอบความเอนเอียงของค่าคลาดเคลื่อน หรือความเที่ยงของตัวแบบคือ การตรวจสอบแบบแทร็คกิงซิกแนล (Tracking Signal test)ในแต่ละคาบเวลา แทร็คกิงซิกแนลเป็นตัวสถิติที่ได้จากการหาค่าประมาณของค่าคาดหวังของค่าคลาดเคลื่อนในการพยากรณ์ ด้วยค่าความแปรปรวนของค่าประมาณนั้น ถ้าวิธีการพยากรณ์ให้ค่าพยากรณ์ที่ไม่เอนเอียงแล้ว ค่าแทร็คกิงซิกแนลควรมีค่าไม่ต่างจากศูนย์ในเชิงสถิติ

ในปัจจุบันวิธีพยากรณ์เชิงปริมาณที่นิยมใช้กันมากมีอยู่หลายวิธี เช่น เทคนิคการปรับให้เรียบ (Smoothing Technique) การพยากรณ์แบบการกรองแบบปรับได้ (Adaptive Filtering) การวิเคราะห์อนุกรมเวลาแบบคลาสสิค (Classical Time Series Analysis) การวิเคราะห์การถดถอย (Regression Analysis) และอนุกรมเวลา Box-Jenkins ซึ่งวิธีพยากรณ์เหล่านี้นำไปใช้ประโยชน์ในงานด้านต่าง ๆ มากมาย เช่น การวางแผนการผลิต (Production Planning) การวางแผนด้านการเงิน (Financial Planning) การจัดกำลังคน (Staff Scheduling) ปัญหาทางด้านเศรษฐศาสตร์ และการจัดการสินค้าคงเหลือ (Inventory Management) เป็นต้น

ในการวิจัยนี้ เทคนิคการพยากรณ์ที่ผู้วิจัยสนใจจะนำมาใช้ในการวิจัยเรื่องของแทร็คกิงซิกแนล คือเทคนิคการปรับให้เรียบสำหรับอนุกรมเวลาที่ไม่มีฤดูกาลและจำลองได้ด้วยตัวแบบแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วง (Locally Constant Linear Trend Model) โดยมีวิธีการพยากรณ์ค่าในอนาคตที่เรียกว่า "การปรับให้เรียบแบบเอ็กซ์โพเนนเชียลสองครั้ง" (Double Exponential Smoothing) ตัวแบบดังกล่าวนี้เป็นตัวแบบที่มีใช้กันมากตัวแบบหนึ่งในหลายระบบงาน เช่นระบบการจัดการสินค้าคงเหลือในการพยากรณ์ยอดขายเป็นต้น

ในงานพยากรณ์ เมื่อเลือกวิธีการพยากรณ์ได้เหมาะสมแล้ว ยังจำเป็นต้องมีการตรวจสอบความเที่ยง ของวิธีการพยากรณ์ที่ใช้อยู่เรื่อย ๆ เมื่อเวลาผ่านไปและมีข้อมูลใหม่เพิ่มเข้ามา วิธีการที่ใช้ตรวจสอบก็คือ การตรวจสอบแบบแทร็คกิงซิกแนล ซึ่งมีอยู่ด้วยกันหลายวิธี ในทางปฏิบัติถ้าจะนำวิธีการตรวจสอบนี้ไปใช้ ควรจะเลือกใช้วิธีการตรวจสอบที่สามารถตรวจพบความเอนเอียงของค่าพยากรณ์ได้รวดเร็ว ดังนั้นจึงมีนักสถิติหลายท่านได้ทำการเปรียบเทียบวิธีการตรวจสอบเหล่านี้ แต่เป็นการเปรียบเทียบภายใต้ตัวแบบค่าเฉลี่ยไม่คงที่ตลอดช่วง (Locally

Constant Mean Model) โดยใช้วิธีการพยากรณ์ที่เรียกว่า "การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว" (Single Exponential Smoothing)

สำหรับการวิจัยนี้ จะเป็นการเปรียบเทียบภายใต้ตัวแบบแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วง ซึ่งในการเปรียบเทียบนั้นจะใช้จำนวนคาบเวลา (จำนวนข้อมูล) ที่ต้องใช้ตรวจสอบโดยเฉลี่ย (Average run length) ก่อนที่จะพบว่าค่าสถิติแทร็คกิงซิกแนล มีค่าตกออกนอกขอบเขตควบคุม (Control limit) เป็นเกณฑ์ในการประเมินประสิทธิภาพ (ความรวดเร็วในการตรวจพบความเอนเอียงของค่าพยากรณ์) ของวิธีการตรวจสอบ

1.2 วัตถุประสงค์ของการวิจัย

ในการวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบวิธีการตรวจสอบแบบแทร็คกิงซิกแนล 4 วิธีคือ

1.2.1 วิธีการตรวจสอบแบบผลรวมสะสมอย่างง่าย (Simple Cumulative Sum)

1.2.2 วิธีการตรวจสอบแบบผลรวมสะสมย้อนหลัง (Backward Cumulative Sum)

1.2.3 วิธีการตรวจสอบแบบปรับความคลาดเคลื่อนให้เรียบ (Smoothed Error)

1.2.4 วิธีการตรวจสอบแบบสหสัมพันธ์ในตัวเอง (Autocorrelation)

เพื่อหาข้อสรุปเกี่ยวกับประสิทธิภาพของการตรวจสอบ และหาขอบเขตควบคุมสำหรับแต่ละวิธีการตรวจสอบ เมื่อใช้กับตัวแบบแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วง ภายใต้ระดับแนวโน้มที่เปลี่ยนแปลงไป และภายใต้ค่าพารามิเตอร์ต่าง ๆ ที่ใช้ในวิธีการตรวจสอบ

1.3 สมมติฐานในการวิจัย

สมมติฐานของการวิจัยที่ต้องการจะทดสอบคือ

1.3.1 ในกรณีที่ค่าความแปรปรวน MAD และ MSE ที่ใช้ในวิธีการตรวจสอบมีค่าไม่คงที่ จะมีสมมติฐานดังนี้

1.3.1.1 วิธีการตรวจสอบแบบผลรวมสะสมย้อนหลัง (Backward Cumulative Sum) จะมีประสิทธิภาพดีกว่าวิธีอื่น ๆ

1.3.1.2 วิธีการตรวจสอบแบบสหสัมพันธ์ในตัวเอง (Autocorrelation) และวิธีการตรวจสอบแบบปรับความคลาดเคลื่อนให้เรียบ (Smoothed Error) จะมีประสิทธิภาพดีกว่า วิธีการตรวจสอบแบบผลรวมสะสมอย่างง่าย (Simple Cumulative Sum)

1.3.2 ในกรณีที่ค่าความแปรปรวน MAD และ MSE ที่ใช้ในวิธีการตรวจสอบมีค่าไม่คงที่ จะมีสมมติฐานดังนี้

1.3.2.1 วิธีการตรวจสอบแบบปรับความคลาดเคลื่อนให้เรียบ (Smoothed Error) จะมีประสิทธิภาพดีกว่าวิธีอื่น

1.3.2.2 วิธีการตรวจสอบแบบสหสัมพันธ์ในตัวเอง (Autocorrelation) จะมีประสิทธิภาพดีกว่าวิธีการตรวจสอบแบบผลรวมสะสมอย่างง่าย (Simple Cumulative Sum)

1.4 ข้อตกลงเบื้องต้น

การวิจัยครั้งนี้ จะใช้จำนวนคาบเวลาโดยเฉลี่ย (Average run length) ที่ต้องใช้ในการตรวจพบความเอนเอียงของความคลาดเคลื่อนในการพยากรณ์ของแต่ละวิธีการตรวจสอบเป็นดัชนีสำคัญที่นักวิจัยจะใช้เป็นเกณฑ์ในการประเมินประสิทธิภาพของวิธีการตรวจสอบ

1.5 ขอบเขตของการวิจัย

การดำเนินงานวิจัยครั้งนี้จะมีขอบเขตของการศึกษาคือ

1.5.1 ศึกษาเฉพาะตัวแบบแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วง ซึ่งมีตัวแบบดังนี้

$$Z_t = \beta_o + \beta_1 t + \beta_2 (t-\ell)\ \delta_t + \varepsilon_t \qquad t = 1,2,...$$

$$\delta_t = \begin{cases} 0 & t < \ell \\ 1 & t > \ell \end{cases}$$

t คือคาบเวลาซึ่งเป็นตัวแปรอิสระ

Z_t คือค่าของข้อมูล ณ เวลา t ซึ่งเป็นตัวแปรตาม

β_o , β_1 , β_2 คือพารามิเตอร์ในตัวแบบ

ε_t คือค่าผิดพลาดสุ่ม ณ เวลา t ซึ่งมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับศูนย์มีความแปรปรวนคงที่

ℓ คือจำนวนคาบเวลาหรือจำนวนค่าสังเกตเริ่มต้น (Run-in period) ก่อนที่จะมีการเปลี่ยนระดับแนวโน้ม (ความชัน)

1.5.2 ศึกษาจำนวนคาบเวลาที่ต้องใช้ในการตรวจสอบโดยเฉลี่ย ของวิธีการตรวจสอบทั้ง 4 วิธีที่กล่าวมาแล้ว

1.5.3 ศึกษาในกรณีที่ค่าความแปรปรวนของวิธีการตรวจสอบแบ่งเป็น 2 ลักษณะ คือ

1.5.3.1 ค่าความแปรปรวน MAD และ MSE มีค่าคงที่

1.5.3.2 ค่าความแปรปรวน MAD และ MSE มีค่าไม่คงที่

1.5.4 ศึกษาในกรณีที่แนวโน้ม (β_2) เปลี่ยนแปลงไปเท่ากับ 0.0σ, 0.5σ 1.5σ, 2.0σ , 2.5σ , 3.0σ

1.5.5 ศึกษาในกรณีที่ค่าคงที่ที่ใช้ในการปรับให้เรียบในสูตรพยากรณ์ (α) มีค่าเท่ากับ 0.1 0.2 และ 0.3

1.5.6 ศึกษาในกรณีที่ค่าคงที่ที่ใช้ในการปรับให้เรียบในตัวสถิติแทร็คกิงซิกแนล (γ) มีค่าเท่ากับค่าคงที่ที่ใช้ในการปรับให้เรียบในสูตรพยากรณ์

1.5.7 ในข้อ 1.5.3.2 จะศึกษาเพิ่มเติมในกรณีที่ค่าคงที่ที่ใช้ในการปรับให้เรียบในตัวสถิติแทร็คกิงซิกแนล (γ) มีค่าต่ำกว่าค่าคงที่ที่ใช้ในการปรับให้เรียบในสูตรพยากรณ์ (α) โดยกำหนดให้ $\gamma = 0.05 , 0.1$

1.5.8 ศึกษาในกรณีที่จำนวนข้อมูลในคาบเวลาเริ่มต้นเท่ากับ 20, 40 , 60 และ 100

1.5.9 กำหนดค่าผิดพลาดสุ่มของการพยากรณ์ มีการแจกแจงแบบปกติ มีค่าเฉลี่ยเท่ากับศูนย์ และมีความแปรปรวนเท่ากับ 1 และกำหนดพารามิเตอร์ในตัวแบบแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วง $\beta_o = 0$, $\beta_1 = 1$ ทั้งนี้เพราะการเปลี่ยนแปลงของ β_o , β_1 ตามที่ได้ทดลองแล้วพบว่าไม่มีผลต่อดัชนีที่นักวิจัยใช้เป็นเกณฑ์ในการประเมินประสิทธิภาพของวิธีการตรวจสอบ ในการวิจัยนี้จึงขอเลือกกำหนดค่า β_o และ β_1 เป็นค่าดังกล่าว

1.5.10 ในการวิจัยครั้งนี้จะทำการทดลองด้วยเครื่องคอมพิวเตอร์ IBM 370/3031 โดยใช้เทคนิคมอนติคาร์โล (Monte Carlo Technique) การทดลองจะกระทำซ้ำ ๆ กัน 1,000 ครั้ง ในแต่ละสถานการณ์ของการทดลอง

1.6 คำจำกัดความ

คำจำกัดความที่ใช้ในงานวิจัยมีดังนี้

1.6.1 วิธีการตรวจสอบที่มีประสิทธิภาพดีหมายถึง เมื่อค่าพารามิเตอร์ในตัวแบบแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วงเปลี่ยนแปลงไป วิธีการตรวจสอบนั้นจะต้องมีจำนวนคาบเวลาที่ใช้ตรวจสอบโดยเฉลี่ยน้อย หรือกล่าวได้อีกอย่างหนึ่งว่า สามารถตรวจพบความเปลี่ยนแปลงได้รวดเร็ว

1.6.2 จำนวนคาบเวลาที่ต้องใช้ในการตรวจสอบ (Run length) คือจำนวนคาบที่นับตั้งแต่เริ่มการพยากรณ์หรือเริ่มตรวจสอบจนกระทั่ง วิธีการตรวจสอบนั้นตรวจพบความเอนเอียงของความคลาดเคลื่อนในการพยากรณ์

1.6.3 จำนวนคาบเวลาเริ่มต้นก่อนที่จะมีการตรวจสอบ (Run-in Period) คือจำนวนคาบเวลาหรือจำนวนข้อมูลที่นำมาใช้สร้างตัวแบบในการพยากรณ์

1.7 ประโยชน์ของการวิจัย

1.7.1 เพื่อให้ได้ข้อสรุปจากการทดลองเกี่ยวกับวิธีการตรวจสอบความเอนเอียงของความคลาดเคลื่อนในการพยากรณ์ด้วยตัวแบบแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วง และเพื่อเป็นแนวทางในการเลือกใช้วิธีการตรวจสอบแก่ผู้ที่สนใจต่อไป

1.7.2 เพื่อให้ได้ขอบเขตควบคุมของวิธีการตรวจสอบทั้ง 4 วิธี ที่เหมาะสมที่จะนำไปใช้ในทางปฏิบัติต่อไป