บทที่ 2
การว่างงานและวิธีการทางสถิติที่เกี่ยวข้อง

การว่างงาน และวิธีการทางสถิติเกี่ยวข้องในส่วนของบทนี้จะแบ่งออกเป็น 2 ส่วน คือ ส่วนแรกจะกล่าวถึงการว่างงานและสถานการณ์การว่างงานของประเทศไทย ส่วนที่สองจะกล่าวถึงวิธีการทางสถิติที่ใช้ในการศึกษาตัวแบบพยากรณ์อัตรา การว่างงานของประเทศไทย ดังนี้

2.1 การว่างงาน

ในประเทศกำลังพัฒนาเช่นประเทศไทย การว่างงานมักจะพบในรูปแบบที่ โรงงานลดกิจการลงจึงปลดคนงานลง และการว่างงานดังกล่าวถือเป็นปัญหาใหญ่ในภาคอุตสหากรรมมากกว่าในภาคเกษตรกรรม เนื่องจากแรงงานในภาคอุตสาหกรรมนั้นกำลังแรงงานทุกคนทำงานแลกกับเงินเดือนเพื่อเป็นค่าเลี้ยงชีพ ถ้าถูกออกจากงานจะสูญเสียงาน รายได้ เงินออม บ้านและทรัพย์สิน ฝีมือ เพื่อน และความเชื่อถือในตัวเอง ส่วนในภาคเกษตรกรรม ประชากรส่วนใหญ่เป็นเกษตรกรหรือประกอบอาชีพส่วนตัวจึงทำให้ปัญหาการว่างงานไม่ค่อยรุนแรง

จำนวนคนว่างงานสามารถทราบได้โดย 2 วิธี คือ

1) การสำรวจอาจเป็นการสำรวจเฉพาะการว่างงาน หรือโครงการสำรวจทางด้านแรงงานบางโครงการที่มีการสำรวจการว่างงานรวมอยู่ด้วย หรือสำมะโนประชากรทั่วประเทศ ซึ่งประเทศไทยได้ตัวเลขจำนวนคนว่างงานจากการสำมะโนประชากรและการสำรวจแรงงาน

2) ให้คนว่างงานมาจดทะเบียนหางานทำที่สำนักงานจัดหางานและในขณะเดียวกันเพื่อรับเบี้ยประกันการว่างงาน ซึ่งมักกระทำกันในประเทศอุตสาหกรรม แต่มีข้อเสีย คือ ทราบผลได้ช้า ใช้ค่าใช้จ่ายสูง จึงไม่สามารถจัดทำสม่ำเสมอได้

2.1.1ประเภทของการว่างงาน

เพื่อเป็นการจัดกลุ่มคนว่างงาน และแยกการว่างงานตามสาเหตุที่เกิด อาจจำแนกการว่างงานออกได้เป็น 5 ประการคือ

1) การว่างงานเกิดตามปกติหรือเพราะการเปลี่ยนงาน ได้แก่ การว่างงานเกิดขึ้นเพราะคนงานเปลี่ยนจากงานหนึ่งไปทำอีกงานหนึ่ง หรือผู้สำเร็จการศึกษาใหม่กำลังจะเข้าสู่กำลังแรงงาน ต้องว่างงานอยู่ระยะหนึ่ง คนงานลาออกจากงานเดิมโดยสมัครใจเพื่อหางานใหม่

2) การว่างงานตามฤดูกาล การว่างงานที่เกิดในฤดูหนึ่งหรือช่วงเวลาหนึ่ง ในระยะเวลา 1 ปี เรียกการว่างงานตามฤดูกาล เหตุที่เกิดการว่างงานชนิดนี้ เพราะสภาพดินฟ้าอากาศ เช่น งานก่อสร้างจะทำในหน้าแล้ง เพราะหน้าฝนงานก่อสร้างทำได้ลำบาก คนงานก่อสร้างจะว่างงาน นอกจากนี้

ยังเกิดจากนิสัยซื้อของของคน เช่นหน้าคริสมาส คนซื้อของมาก กิจการคึกคัก จ้างคนจำนวนมาก พอหมดคริสมาส คนงานเหล่านี้จะถูกให้ออกจากงาน เกิดการว่างงาน

3). การว่างงานเกิดขึ้นเพราะใช้เทคนิคใหม่ การว่างงานชนิดนี้ เป็นผลมาจากการนำเครื่องจักรชนิดใหม่ วิธีการผลิตแบบใหม่ ความรู้ใหม่ หรือที่เรียกว่าเทคนิคใหม่ เข้ามาใช้แทนแรงคน

4). การว่างงานเนื่องจากโครงสร้างทางเศรษฐกิจ โครงสร้างทางเศรษฐกิจมักเปลี่ยนไปตามรสนิยมของผู้บริโภค เช่น คนไทยเคยกินข้าว ถ้าเปลี่ยนไปกินขนมปัง ชาวนาจะว่างงาน เว้นแต่หันไปปลูกข้าวสาลีได้ เมื่อคนนิยมใช้ถังน้ำพลาสติก จะทำให้คนงานทำถังสังกะสีว่างงาน อุตสาหกรรมพลาสติกเกิดขึ้น จะทำให้อุตสหากรรมทำถังน้ำ เครื่องเด็กเล่น และสิ่งอื่น ๆ ที่ทำด้วยพลาสติกแทนได้ ลดกิจการลง เกิดมีคนว่างงาน

5). การว่างงานเนื่องจากเศรษฐกิจตกต่ำหรือเพราะวัฏจักรธุรกิจ ได้แก่การว่างงานที่เกิดขึ้นเพราะระดับธุรกิจขึ้น ๆ ลง ๆ เมื่อถึงเวลาเศรษฐกิจถดถอยอย่างรุนแรง จะเกิดการว่างงานจำนวนมาก มักจะเกิดจากผลของสงคราม, สงครามกลางเมือง หรือเหตุ ฉุกเฉินอื่น ๆ ทำให้ประเทศขาดความมั่นคงทางเศรษฐกิจ เป็นเหตุให้เกิดการว่างงาน

6) อื่นๆ นอกจากการว่างงานประเภทของการว่างงานที่กล่าวมาข้างต้น ยังมีสาเหตุอื่นๆที่ทำให้เกิดการว่างงาน เช่น

6.1). ผู้เข้าสู่กำลังแรงงานเพิ่มเร็ว ตามการเพิ่มของประชากรที่เพิ่มเร็ว เช่น ไทยประชากรเพิ่มปีละ 3.2% หรือปีละเกือบล้านคน จะทำให้ผู้เข้าสู่กำลังแรงงาน หรือผู้ถึงเกณฑ์อายุและออกทำงานเพิ่มจำนวนตามกันไป มีผู้ประมาณไว้ราวครึ่งล้านคนต่อปี ถ้าตำแหน่งงานไม่เพิ่มขึ้นหรือเพิ่มเพียงนิดหน่อยเป็นเวลาหลายปีก็จะเป็นจำนวนคนว่างงานมากขึ้น

6.2). การลงทุน ไม่เพิ่มขึ้นมาก จึงเป็นเหตุให้การขยายตำแหน่งงานไม่เพิ่ม

6.3). ผู้ประกอบอาชีพด้านเกษตรกรรมกษตรกรรมไม่มีกำลังด้านการเงินพอที่จะขยายเนื้อที่ทำกิน ให้เพียงพอกับลูกหลานเกษตรกรที่เกิดใหม่และเพิ่มขึ้นเรื่อยๆ ทำให้แรงงานในภาคเกษตรเปลี่ยนไปสู่แรงงานในภาคอุตสาหกรรมมากขึ้น ในขณะที่ตำแหน่งงานในภาคอุตสาหกรรมเพิ่มขึ้นไม่เพียงพอ

2.1.2 แนวคิดเกี่ยวกับการมีงานทำและการว่างงาน

ทฤษฎีที่สำคัญและนิยมใช้สำหรับการสำหรับการอธิบายการมีงานทำ และการว่างงานในทางเศรษฐศาสตร์มี 2 ทฤษฎี คือ (สุรักษ์ บุนนาค และ วันรักษ์ มิ่งมณีนาคิน , 2520 , น.41-49)

1.ทฤษฎีการว่างงานของสำนักคลาสสิค (The Classical Theory of Employment) ทฤษฎีนี้ให้ความสนใจในเรื่องของการทำงานเต็มอัตรา (Full Employment) ซึ่งโดยความหมายคือภาวะการนำปัจจัยการผลิตที่มีอยู่ในขณะนั้นมาใช้ในกระบวนการผลิตอย่างมีประสิทธิภาพ เฉพาะ

อย่างยิ่งคนที่สามารถทำงานได้จะมีงานทำทุกคน และยินดีทำงาน ณ ค่าจ้างขณะนั้น สมมุติฐานสำคัญของทฤษฎีนี้ คือ ในขณะหนึ่งขณะใดระบบเศรษฐกิจจะอยู่ในภาวะสมดุล ณ ระดับที่มีการจ้างงานเต็มที่เสมอ การว่างงานอาจเกิดขึ้นได้แต่ก็เพียงชั่วคราวและสามารถแก้ไขได้เอง โดยไม่จำเป็นต้องดำเนินนโยบายแต่อย่างใด

2. ทฤษฎีเศรษฐศาสตร์มหภาคของสำนักงานเคนส์ (John Maynard Keynes) ทฤษฎีนี้ได้รับการยอมรับมากกว่าทฤษฎีแรก เพราะสามารถใช้ในระบบเศรษฐกิจที่มีการจ้างงานเต็มที่ และที่มีการว่างงาน ความเชื่อที่สำคัญของทฤษฎีนี้เชื่อว่า โดยทั่วไปแล้วเศรษฐกิจมักอยู่ต่ำกว่าระดับที่มีการจ้างงานเต็มที่ สมมุติฐานของทฤษฎีนี้เชื่อว่า การว่างงานเกิดขึ้นได้เสมอ ถ้าไม่แก้ด้วยนโยบายและระบบเศรษฐกิจอาจอยู่ในภาวะการจ้างงานไม่เต็มที่เรื่อยไปก็ได้

2.1.3 สถานการณ์การมีงานทำและการว่างงานในประเทศไทย

จากสถานการณ์แรงงานในปี 2539-2540 มีผลกระทบต่อความต้องการแรงงานของประเทศและในช่วงปี 2540-2545 โดยทำให้เกิดการชลอตัวลงหรือมีการเลิกจ้าง และพบว่าภาวะเศรษฐกิจในปี 2540 ดังกล่าว ได้ทำให้แนวโน้มความต้องการแรงงานในภาคอุตสาหกรรมชะลอตัวลง โดยเมื่อเทียบกับภาวะปกติ โดยภาวะการว่างงานเฉลี่ยของปี 2541 มีผู้ว่างงาน 1.41 ล้านคน และในปี 2542 มีผู้ว่างงาน 1.48 ล้านคน ต่อมาในปี 2543 ความต้องการแรงงานเพิ่มสูงขึ้น เนื่องมาจากการปรับตัวทางเศรษฐกิจที่ดีขึ้น ภาวะการว่างงานเฉลี่ยของปี 2543 มีผู้ว่างงาน 1.19 ล้านคนซึ่งได้ลดลงจากปี 2542

จากการที่กระทรวงแรงงานและสวัสดิการสังคม และสำนักงานสถิติแห่งชาติ ได้มีการสำรวจภาวะการมีงานทำ เก็บรวบรวมข้อมูลทางด้านแรงงาน ก็เพื่อให้หน่วยงานราชการที่เกี่ยวข้อง เช่น กระทรวงแรงงาน และสวัสดิการสังคม นำไปประมาณการกำลังแรงงาน การมีงานทำ และการว่างงาน เพื่อใช้เป็นข้อมูลในการวางแผนพัฒนาและตัดสินใจเกี่ยวกับนโยบายด้านแรงงาน และนโยบายด้านอื่นๆที่เกี่ยวข้อง ส่วนกองวางแผนทรัพยากรมนุษย์ สำนักงานคณะกรรมการพัฒนาการเศรษฐกิจและสังคมแห่งชาติ ได้มีการประเมินผลกระทบของวิกฤตการณ์ทางเศรษฐกิจและแนวโน้มของเศรษฐกิจที่ชะลอตัวและนำมาพยากรณ์อัตราการว่างงานในแต่ละปี เพื่อใช้เป็นข้อมูลในการวางแผนด้านกำลังคน และด้านอื่นๆที่เกี่ยวข้อง ดังนั้นการหาตัวแบบพยากรณ์อัตราการว่างงานในภาคอุตสาหกรรมที่เหมาะสมนั้น จะเป็นข้อมูลที่เป็นประโยชน์ต่อกระทรวงแรงงานและสวัสดิการสังคม และกองวางแผนทรัพยากรมนุษย์ สำนักงานคณะกรรมการพัฒนาเศรษฐกิจและสังคมแห่งชาติ เพื่อช่วยให้มีการวางแผนต่างๆ ได้เป็นอย่างดี โดยนำข้อมูลทางด้านเศรษฐกิจ และข้อมูลทางด้านแรงงานมาพิจารณาประกอบกับค่าพยากรณ์อัตราการว่างงานในภาคอุตสาหกรรมที่ได้ จะทำให้หน่วยงานที่เกี่ยวข้องทราบว่าควรส่งเสริมแนวทางการศึกษา แนวทางการลงทุน แก่ประชาชนอย่างไร

2.2 ทฤษฎีที่ใช้ในการพยากรณ์อัตราการว่างงานในภาคอุตสาหกรรม

2.2.1 การพยากรณ์

การพยากรณ์เป็นเทคนิคทางสถิติอย่างหนึ่ง ที่ใช้เป็นเครื่องมือในการคาดการณ์เหตุการณ์ หรือสิ่งที่จะเกิดขึ้นในอนาคต กาะพยากรณ์มีบทบาทสำคัญในการวางแผน หรือการตัดสินใจทำให้สามารถกำหนดนโยบาย หรือเตรียมการให้สอดคล้องกับความต้องการและสถานการณ์ที่จะเกิดขึ้นในอนาคต ซึ่งจะช่วยให้แผนงานที่กำหนดมีประสิทธิภาพยิ่งขึ้น

2.2.1.1 วิธีการพยากรณ์

การพยากรณ์สามารถแบ่งได้เป็น 2 วิธีการใหญ่ๆ คือ

1. วิธีการพยากรณ์เชิงคุณลักษณะ (Qualitative Forecasting Methods) เป็นการพยากรณ์ที่ใช้ความเชื่อความรู้สึก ประสบการณ์ หรือดุลพินิจของผู้เชี่ยวชาญ หรือผู้มีประสบการณ์คาดคะเนเหตุการณ์ต่างๆที่จะเกิดขึ้น วิธีการนี้ใช้เมื่อไม่มีข้อมูล หรือมีข้อมูลน้อย รวมถึงข้อมูลไม่น่าเชื่อถือ วิธีนี้เหมาะสำหรับการพยากรณ์ในระยะปานกลางหรือระยะยาว เช่น วิธีการเดลไฟ (Delphi Method) เป็นต้น

2. วิธีการพยากรณ์เชิงปริมาณ (Quantitative Forecasting Methods) เป็นการพยากรณ์ที่อาศัยข้อมูลข่าวสารเกี่ยวกับเหตุการณ์ที่เกิดขึ้นแล้วในอดีตแล้วนำข้อมูลเหล่านี้มาศึกษาหารูปแบบ หรือตัวแบบพยากรณ์ โดยวิธีการทางสถิติและคณิตศาสตร์ ซึ่งสามารถจะทำได้เมื่อมีข้อมูลในอดีตอยู่ในรูปของตัวเลข หรือสามารถแปลงเป็นตัวเลขได้ และสมมติว่ารูปแบบการแปรผันของข้อมูลที่ผ่านมาจะมีแนวโน้มเป็นลักษณะคล้ายคลึงกันเช่นนั้นด้วยในอนาคต (โดยเฉพาะในระยะสั้น) วิธีการเชิงปริมาณหรือตัวแบบเชิงปริมาณ สามารถจำแนกออกได้เป็น

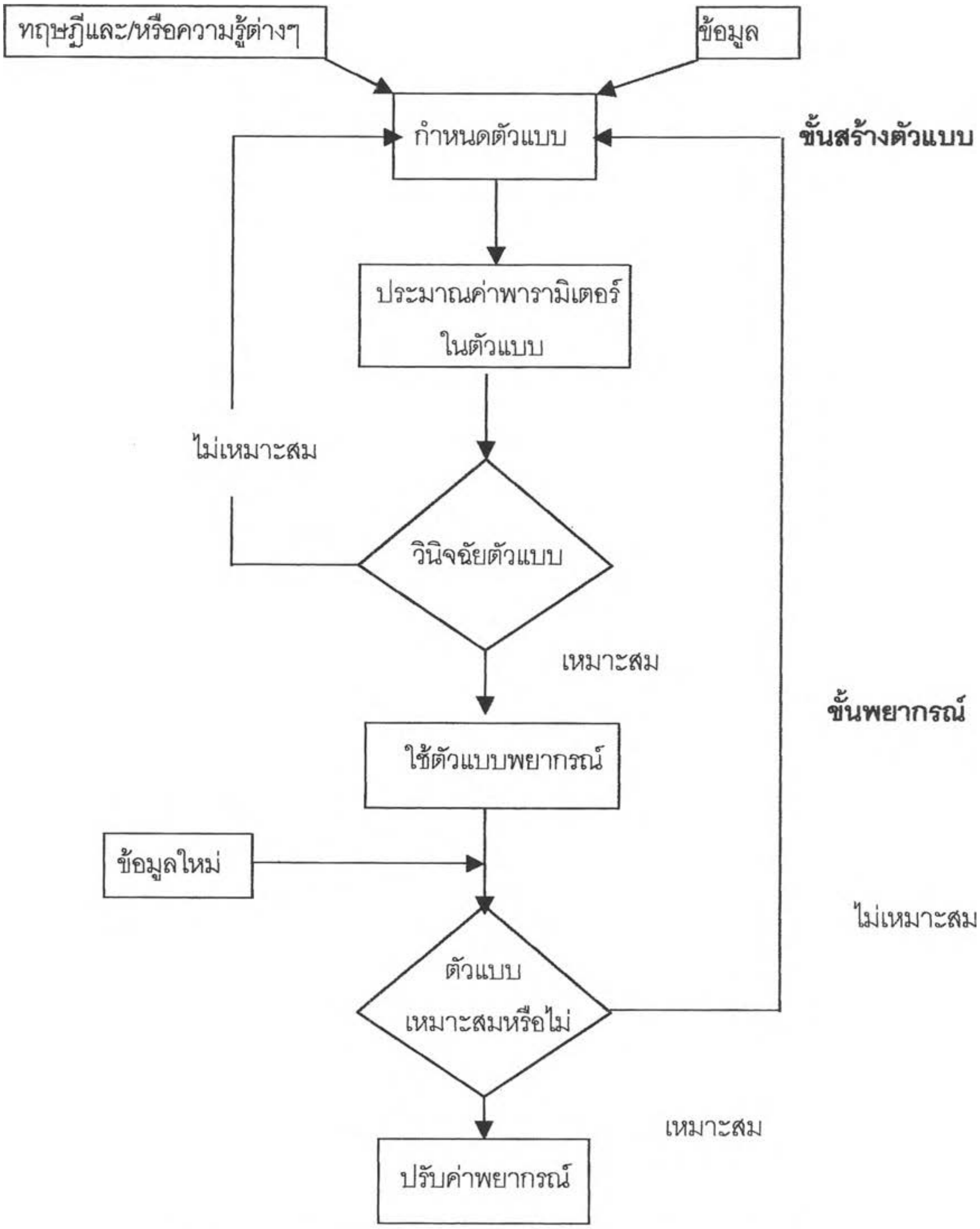
- *ตัวแบบภายใต้ภาวะคงที่ หรือที่เรียกว่าตัวแบบเชิงกำหนด (Deterministic Models)* ตัวแบบประเภทนี้จะไม่มีองค์ประกอบของความไม่แน่นอนหรือไม่มีความคลาดเคลื่อน หรือมีความคลาดเคลื่อนน้อยมากที่สามารถจะละไว้ได้ โดยทั่วไปจะพบในเรื่องของวิทยาศาสตร์กายภาพ (Physical Sciences)

- *ตัวแบบภายใต้ภาวะไม่คงที่ หรือที่เรียกว่าตัวแบบความน่าจะเป็น (Probabilistic Models)* หรือตัวแบบเฟ้นสุ่ม (Stochastic Models) ตัวแบบประเภทนี้จะเป็นการจำลองความสัมพันธ์ เนื่องจากไม่อาจจะระบุความสัมพันธ์ระหว่างตัวแปรที่ถูกต้องสมบูรณ์ได้ ฉะนั้น ตัวแบบที่จำลองความสัมพันธ์เหล่านี้ย่อมจะมีความคลาดเคลื่อนเป็นองค์ประกอบ และค่าคลาดเคลื่อนมีการแจกแจงของความน่าจะเป็น หรือค่าคลาดเคลื่อนเป็นตัวแปรสุ่ม ดังนั้นเรียกตัวแบบที่สร้างขึ้นว่า ตัวแบบความน่าจะเป็น เช่น ตัวแบบการถดถอย (Regression Models) และตัวแบบอนุกรมเวลา (Time Series Models) เป็นต้น

ในการศึกษาครั้งนี้จะทำการวิเคราะห์ข้อมูล โดยใช้ ตัวแบบการถดถอย และตัวแบบอนุกรมเวลา ในการพยากรณ์อัตราการว่างงาน

2.2.1.2 โครงสร้างของระบบงานพยากรณ์

ระบบงานพยากรณ์ จะแบ่งออกเป็น 2 ขั้นตอนใหญ่ๆ คือ ขั้นตอนการสร้างตัวแบบพยากรณ์ และขั้นตอนการพยากรณ์ โดยมีรายละเอียดดังแผนภูมิ (ดูรูป 2.1)



รูป 2.1. แผนภูมิแสดงโครงสร้างของระบบงานพยากรณ์

1. งานขั้นสร้างตัวแบบ

จะเริ่มด้วยการกำหนดตัวแบบทดลอง เป็นตัวแบบเริ่มต้นที่คาดว่าจะเป็นตัวแบบที่ใช้ได้ โดยอาศัยความรู้ และทฤษฎีต่างๆ ที่เกี่ยวข้อง รวมถึงการวิเคราะห์รูปแบบข้อมูลเบื้องต้น เช่น การใช้กราฟ scatter diagram เพื่อหาความสัมพันธ์ระหว่างตัวแปร จากนั้นทำการประมาณค่าพารามิเตอร์ใน ตัวแบบ และทำการตรวจสอบความความถูกต้องเหมาะสมในเชิงสถิติ โดยทำการตรวจสอบข้อสมมติ หรือคุณสมบัติต่างๆ ที่เกี่ยวข้อง ในเชิงสถิติรวมทั้งรูปแบบความสัมพันธ์ของตัวแบบ ทั้งนี้เพราะตัวแบบที่กำหนดขึ้นครั้งแรกนั้นอาจจะยังไม่เหมาะสมเพียงพอ ถ้าพบว่าตัวแบบที่กำหนดไม่สอดคล้องข้อสมมติในเชิงสถิติ หรือยังมีรูปแบบไม่เหมาะสม จะทำการปรับตัวของตัวแบบใหม่อีก กรรมวิธีจะวนเวียนเช่นนี้ จนกว่าจะพบว่าตัวแบบพยากรณ์ผ่านการทดสอบ มีความเหมาะสมเชิงสถิติ เมื่อผ่านขั้นนี้แล้วก็จะใช้ตัวแบบพยากรณ์ค่าที่ต้องการ ซึ่งเข้า สู่ขั้นพยากรณ์

2. งานขั้นพยากรณ์

จากตัวแบบพยากรณ์ที่ได้ เรานำไปใช้พยากรณ์ ตัวแบบพยากรณ์นั้นอาจใช้งานได้ในระยะเวลาหนึ่ง เมื่อเวลาผ่านไปเราได้ข้อมูลใหม่เพิ่มขึ้น ข้อมูลที่เกิดขึ้นใหม่นี้ควรนำมาใช้ตรวจสอบตัวแบบพยากรณ์ว่ายังคงมีความเหมาะสม ในเชิงสถิติหรือไม่ ถ้าพบว่าตัวแบบไม่เหมาะสม ควรปรับแก้ตัวแบบใหม่โดยกลับเข้าสู่งานขั้นสร้างตัวแบบดังกล่าวข้างต้น

การวินิจฉัยตัวแบบพยากรณ์ คือ การตรวจสอบคุณสมบัติของค่าคลาดเคลื่อนของค่าพยากรณ์ หรือที่เรียกว่า “เศษตกค้าง” (residuals) เพื่อพิจารณาว่าตัวแบบพยากรณ์ที่สร้างขึ้นมีความเหมาะสมหรือไม่ ในเชิงสถิติ การตรวจสอบจะตรวจสอบว่าเศษตกค้างมีคุณสมบัติตามข้อสมมติหรือเงื่อนไขต่างๆ ของตัวแบบหรือเทคนิคพยากรณ์หรือไม่รวมทั้งตรวจสอบรูปแบบของตัวแบบที่กำหนดขึ้น ถ้าพบว่ามีคุณสมบัติไม่เป็นไปตามข้อสมมติบางข้อ หรือทั้งหมดนักพยากรณ์ควรพิจารณาปรับแก้ตัวแบบพยากรณ์นั้น คุณสมบัติพื้นฐานที่จะตรวจสอบสำหรับวิธีการพยากรณ์เชิงปริมาณโดยทั่วไป ได้แก่ ความไม่มีสหสัมพันธ์ ค่าเฉลี่ยเป็นศูนย์ ความแปรปรวนคงที่ และการแจกแจงปกติ

การเปรียบเทียบตัวแบบพยากรณ์ ในงานพยากรณ์ โดยปกตินักพยากรณ์จะพิจารณาสร้างตัวแบบพยากรณ์มากกว่าหนึ่งตัวแบบสำหรับการพยากรณ์เรื่องหนึ่ง ๆ เพื่อคัดเลือกตัวแบบที่คาดว่าจะให้ค่าพยากรณ์มีความคลาดเคลื่อนต่ำ ซึ่งในการศึกษาครั้งนี้ผู้วิจัยจะใช้ค่าเฉลี่ยของเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error) เป็นเกณฑ์ในการตัดสินใจเลือกตัวแบบพยากรณ์

ในงานวิจัยนี้จะใช้วิธีการวิเคราะห์หาตัวแบบดังต่อไปนี้

2.2.2 วิธีการวิเคราะห์การถดถอย เป็นเทคนิคเชิงสถิติหนึ่งสำหรับการศึกษาวิเคราะห์และจำลองรูปแบบความสัมพันธ์เชิงคณิตศาสตร์ระหว่างตัวแปรสองกลุ่ม ตัวแปรกลุ่มหนึ่งเรียกว่า “ตัวแปรตาม” (dependent variable) หรือ “ตัวแปรผล” (response variable) มีหนึ่งตัวแปร เป็นตัวแปรที่นักสถิติ หรือนักพยากรณ์ สนใจที่จะศึกษาลักษณะการเปลี่ยนแปลง หรือพยากรณ์ค่าหรือควบคุม โดยศึกษาวิเคราะห์หารูปแบบความสัมพันธ์กับตัวแปรอีกกลุ่มหนึ่ง เรียกตัวแปรในกลุ่มนี้ว่า “ตัวแปรอิสระ” (independent variables) หรือ “ตัวแปรให้ค่าพยากรณ์หรือค่าทำนาย” (predictor variables) ตัวแปรในกลุ่มนี้อาจมีหนึ่งหรือมากกว่าหนึ่งตัวแปร และรูปแบบความสัมพันธ์เชิงคณิตศาสตร์หรือเชิงสถิติที่ได้ เรียกว่า “ตัวแบบการถดถอย” หรือ “สมการการถดถอย” จากสมการการถดถอย สามารถอธิบายลักษณะการเปลี่ยนแปลงของตัวแปรตาม หรือพยากรณ์ค่าของตัวแปรตาม หรือใช้ในการควบคุมตัวแปรตาม โดยใช้รูปแบบสมการและค่าของตัวแปรอิสระ

1 ตัวแบบการถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Models) ในการวิเคราะห์รูปแบบความสัมพันธ์เชิงเส้นของตัวแปรตาม Y ต่อตัวแปรอิสระมากกว่าหนึ่งตัวแปร $x_1,x_2,\ldots\ldots,x_k$ เราเรียกรูปแบบความสัมพันธ์นั้นว่า “ตัวแบบการถดถอยเชิงเส้นพหุคูณ” หรือ “สมการถดถอยเชิงเส้นพหุคูณ” และมีรูปแบบทั่วไปดังนี้

$$\begin{aligned} Y &= E(Y) + \varepsilon \\ &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \qquad (1) \end{aligned}$$

โดยที่ค่าเฉลี่ย $E(Y)$ หรือ $E(Y|x_1,x_2,\ldots,x_k)$ 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 $x_{i1}, x_{i2}, \ldots, x_{ik}, Y_i$, $i = 1,2,\ldots,n$ นั่นคือ $x_1 = x_{i1}, x_2 = x_{i2}, \ldots, x_k = x_{ik}, Y = Y_i$ และให้ $\varepsilon = \varepsilon_i$ สำหรับ $i = 1, 2\ldots,n$ ***เขียน***ตัวแบบ (1) ได้ใหม่ดังนี้

$$\begin{aligned} Y_i &= E(Y_i) + \varepsilon_i \\ &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \quad , i = 1,2,\ldots,n \qquad (2) \end{aligned}$$

และ ε_i , $i = 1,2,\ldots,n$ มีข้อสมมติดังนี้

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

ด้วยวิธี OLS ตัวประมาณแบบกำลังสองน้อยที่สุด $B_0, B_1, \ldots, B_k$ ของ $\beta_0, \beta_1, \ldots, \beta_k$ ตามลำดับ คำนวณได้โดยการแก้สมการ

$$\frac{\partial Q}{\partial \beta_1}\bigg|_{\beta_0 = B_0, \beta_1 = B_1, \ldots, \beta_k = B_k} = 0 \quad , i = 1,2,\ldots,k$$

โดยที่

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

โดยการแก้ระบบสมการจะได้ระบบสมการปกติ:

$$B_0 \sum_{i=1}^{n} x_{ik} + B_1 \sum_{i=1}^{n} x_{ik} x_{i1} + B_2 \sum_{i=1}^{n} x_{ik} x_{i2} + \ldots + B_k \sum_{i=1}^{n} x_{ik}^2 = \sum_{i=1}^{n} x_{ik} Y_i \qquad (3)$$

โดยทั่วไป การวิเคราะห์การถดถอยเชิงพหุคูณ จะใช้เมตริกซ์เป็นเครื่องมือ ซึ่งจะทำให้การวิเคราะห์สะดวกมากขึ้น ฉะนั้นให้

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

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

ดังนั้น $\underline{Y}$ เป็นเมตริกซ์ขนาด $n \times 1$ ของตัวแปรสุ่ม หรือเวกเตอร์ขนาด n ของตัวแปรสุ่ม

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

$\underline{\varepsilon}$ เป็นเวกเตอร์ขนาด n ของตัวแปรสุ่มค่าผิดพลาด

$\underline{B}$ ..เป็นเวกเตอร์ขนาด $k+1$ ของตัวประมาณพารามิเตอร์

และ $\underline{X}$ เป็นเวกเตอร์ขนาด $n \times (k+1)$ ของค่าคงที่ 1 และค่าของตัวแปรอิสระ

โดยการใช้สัญลักษณ์เมตริกซ์ข้างต้น เขียนระบบสมการ (2) ได้ใหม่ดังนี้

$$\underline{Y} = \underline{X}\,\underline{\beta} + \underline{\varepsilon} \qquad (4)$$

และเขียนข้อสมมติของ $\underline{\varepsilon}$ สามารถเขียนสั้นได้ดังนี้

$$\underline{\varepsilon} \sim \mathbf{N_n}(\underline{0}, \underline{I}\sigma^2) \qquad (5)$$

ซึ่งหมายความว่า $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ เป็นอิสระกัน และต่างมีการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 ละความแปรปรวน σ^2

ในทำนองเดียวกัน สามารถเขียนข้อสมมติของ Y_i, $i = 1, 2, \ldots, n$ ได้ดังนี้

$$\underline{Y} \sim \mathbf{N_n}(\underline{X}\underline{\beta}, \underline{I}\sigma^2) \qquad (6)$$

จากระบบสมการปกติ (3) เขียนในเทอมของเมตริกซ์ได้สั้น ๆ เป็น

$$(\underline{X}'\,\underline{X})\underline{B} = \underline{X}'\,\underline{Y} \qquad (7)$$

การแก้สมการ (7) หา จะสมมติว่าหาเมตริกซ์ผกผันได้ $(\underline{X}'\,\underline{X})$ ของเมตริกซ์ $\underline{X}'\,\underline{X}$ ได้ ซึ่งเป็นจริงโดยทั่วไปในทางปฏิบัติ เพราะฉะนั้นตัวประมาณแบบกำลังสองน้อยที่สุดสามัญคือ

$$\underline{B} = \underline{(\underline{X}'\,\underline{X})^{-1}}\,X'\,\underline{Y} \qquad (8)$$

และตัวแบบพยากรณ์ค่า Y หรือตัวประมาณของค่าเฉลี่ย $E(Y)$ ของ Y เมื่อกำหนด $x_1 = x_{01}, x_2 = x_{02}, \ldots, x_k = x_{0k}$ คือ

$$\hat{Y} = \underline{X}'_0\,\underline{B}$$

$$= B_0 + B_1x_{01} + B_2x_{02} + \ldots + B_kx_{0k}$$

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

<u>ขั้นตอนการสร้างตัวแบบการถดถอยเชิงเส้น</u>

กรรมวิธีสร้างตัวแบบการถดถอยเชิงเส้น มีขั้นตอนที่ควรดำเนินการดังนี้

1. กำหนดรูปแบบความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ

การกำหนดรูปแบบความสัมพันธ์ อาจใช้ทฤษฎีที่เกี่ยวข้อง เช่น ทฤษฎีเศรษฐศาสตร์กำหนดรูปแบบความสัมพันธ์ ซึ่งอาจจะนำมาประยุกต์ได้กับเรื่องที่ศึกษา ในกรณีที่ไม่สามารถหาทฤษฎีใดมาประยุกต์ได้ นักพยากรณ์จะพิจารณารูปแบบความสัมพันธ์ โดยอาศัยข้อมูลที่มีอยู่ของตัวแปรตามและของตัวแปรอิสระ ซึ่งสามารถพิจารณาโดยใช้กราฟ

1.1 โดยทั่วไปจะมีข้อสมมติว่าตัวแปรตาม Y มีการแจกแจงความน่าจะเป็นแบบปกติ ฉะนั้น ควรเขียนกราฟแผนภาพแบบจุด หรือแผนภาพฮิสโทแกรม (histogram diagram) เพื่อดูลักษณะกระจาย หรือการแจกแจงของ Y ว่าเข้ารูปลักษณะแบบสมมาตรหรือไม่ ถ้าพบว่ามีลักษณะไม่

สมมาตร โดยเบ้ไปทางซ้ายหรือทางขวามาก ควรที่จะแปลงข้อมูลของ Y เพื่อให้เข้าลักษณะการแจกแจงแบบสมมาตร วิธีการแปลงค่าของ Y อาจจะทดลองด้วยแบบต่างๆ เช่น $\sqrt{Y}, 1/\sqrt{Y}, 1/Y, \ln Y,$ หรือ $\log_{10} Y$ เป็นต้น

1.2 เขียนกราฟระหว่างตัวแปรตาม กับตัวแปรอิสระทีละตัว เพื่อพิจารณากำหนดรูปแบบความสัมพันธ์เป็นคู่ ๆ ระหว่างตัวแปรตามและตัวแปรอิสระ

1.3 แปลงแบบตัวแบบการถดถอยที่ไม่อยู่ในรูปแบบเชิงเส้นให้เป็นตัวแบบการถดถอยเชิงเส้น ตัวอย่างการแปลงตัวแบบการถดถอยให้อยู่ในรูปแบบเชิงเส้น เช่น

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

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

ตัวอย่างที่ 2 $Y = \alpha + \beta(1/x) + \varepsilon$ ให้. $x_1 = 1/x$ ได้ตัวแบบ

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

2. การประมาณค่าพารามิเตอร์ของตัวแบบการถดถอย เมื่อนักพยากรณ์กำหนดตัวแบบการถดถอยเป็นตัวแบบเริ่มต้นได้แล้ว ซึ่งอาจจะมีมากกว่าหนึ่งตัวแบบ ขั้นตอนต่อไปก็คือประมาณค่าพารามิเตอร์ในตัวแบบ ซึ่งในงานวิจัยชิ้นนี้จะใช้วิธีกำลังสองน้อยสุดในการประมาณค่าพารามิเตอร์ เนื่องจากเป็นตัวประมาณที่มีประสิทธิภาพ และใช้วิธีวิเคราะห์ทางสถิติในการอนุมานเกี่ยวกับพารามิเตอร์ตลอดจนค่าพยากรณ์แบบช่วง

3. การวินิจฉัยความเหมาะสมของตัวแบบการถดถอย งานขั้นกำหนดรูปแบบตัวแบบและงานประมาณค่าพารามิเตอร์ของตัวแบบเป็นเพียงงานครึ่งหนึ่งของกรรมวิธีการสร้างตัวแบบพยากรณ์ ตัวแบบพยากรณ์ที่ได้จากขั้น 1 และ 2 อาจยังไม่เหมาะสมหรือยังไม่เพียงพอที่จะใช้พยากรณ์ นักพยากรณ์จึงควรตรวจสอบและทำการเปรียบเทียบคัดเลือกตัวแบบพยากรณ์ ถ้าตรวจสอบพบว่า ตัวแบบที่กำลังพิจารณายังขาดความเหมาะสม จะกลับไปทำงานในขั้นที่ 1 ถึง 3 ซ้ำ ๆ จนกว่าจะได้ตัวแบบที่เหมาะสมเพียงพอในเชิงสถิติที่จะใช้พยากรณ์ค่าต่อไป

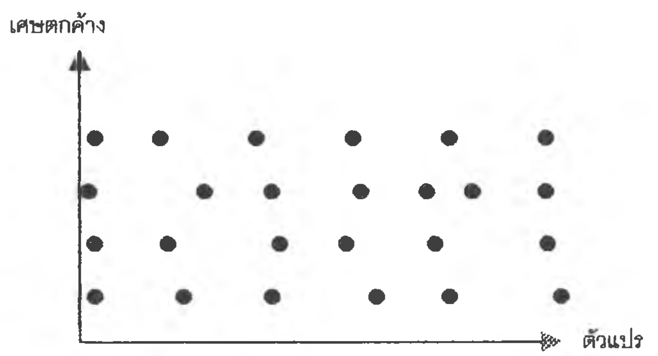
เนื่องจากการอนุมานเชิงสถิติเกี่ยวกับพารามิเตอร์ซึ่งเกี่ยวกับการทดสอบข้อสมมติฐานต่างๆ และการประมาณค่าแบบช่วงของค่าพารามิเตอร์ ตลอดจนการทดสอบและการประมาณค่าแบบช่วงของค่าเฉลี่ย E(Y) ของ Y และของค่าพยากรณ์ของค่าจริง Y โดยทั่วไปการทดสอบและการประมาณค่าดังกล่าว จะกระทำภายใต้ข้อสมมติของ ε_i ดังนั้น จึงจำเป็นต้องตรวจสอบความเหมาะสมของตัวแบบด้วยการตรวจสอบคุณสมบัติของค่าความคลาดเคลื่อนสุ่ม ε_i แต่เนื่องจากไม่ทราบค่าจริง

ε_i ฉะนั้นจะตรวจสอบคุณสมบัติของค่าเศษตกค้าง ซึ่งกำหนด $e_i = y_i - \hat{y}_i$ ซึ่งเป็นค่าประมาณของ ε_i เมื่อตัวแบบถูกต้องเพียงพอ และตรวจสอบว่าค่า e_i (i=1, 2,...,n) มีคุณสมบัติสอดคล้องหรือไม่ นั่นคือ มีค่าเฉลี่ยเท่ากับศูนย์ มีความแปรปรวนคงที่ ไม่มีอัตสหสัมพันธ์และมีการแจกแจงแบบปกติหรือไม่ นอกจากจากนี้อาจพบว่า ตัวแบบไม่เหมาะสม ซึ่งวิธีการตรวจสอบนักพยากรณ์อาจเลือกใช้วิธีกราฟ หรือวิธีการทดสอบเชิงสถิติ ซึ่งวิธีการทดสอบเชิงสถิติเป็นวิธีเชิงระเบียบทฤษฎี แต่โดยทั่วไปวิธีตรวจสอบโดยใช้กราฟก็เป็นวิธีการที่เพียงพอที่จะใช้วินิจฉัยตัวแบบ และเป็นวิธีง่ายๆ ที่ใช้โดยทั่วไป

โดยวิธีกราฟ เราจะเขียนกราฟของค่าเศษตกค้าง e_i หรือค่าเศษตกค้างมาตรฐาน (standardardized residual) $e_i / \sqrt{MSE}$ กับตัวแปรต่างๆ

1. ตัวแปร $\hat{y}_i$
2. ตัวแปรอิสระ x_i แต่ละตัว
3. ตัวแปรเวลา ถ้าข้อมูลเป็นอนุกรมเวลา
4. เขียนกราฟความน่าจะเป็นแบบปกติ (normal probability plot) หรือแผนภาพฮิสโตแกรมของ e_i หรือ $e_i / \sqrt{MSE}$

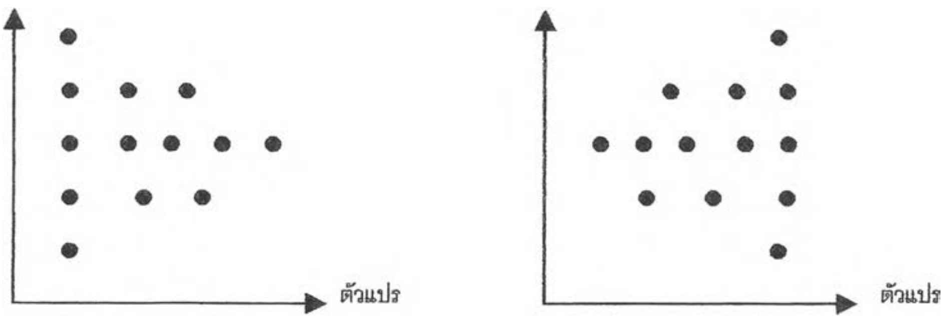
ตัวอย่างกราฟรูป 2.2 ถึง 2.4 แสดงถึงความสัมพันธ์ ที่เป็นไปได้ระหว่าง e_i หรือ $e_i / \sqrt{MSE}$ กับตัวแปร $\hat{y}_i$, x_i



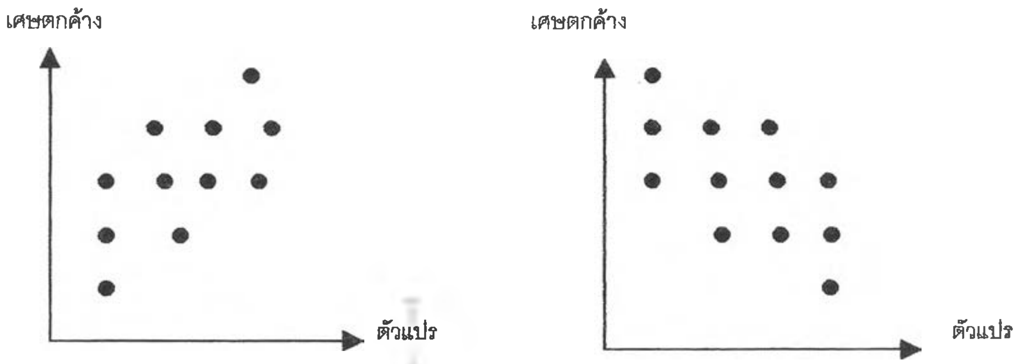
รูป 2.2

เศษตกค้าง

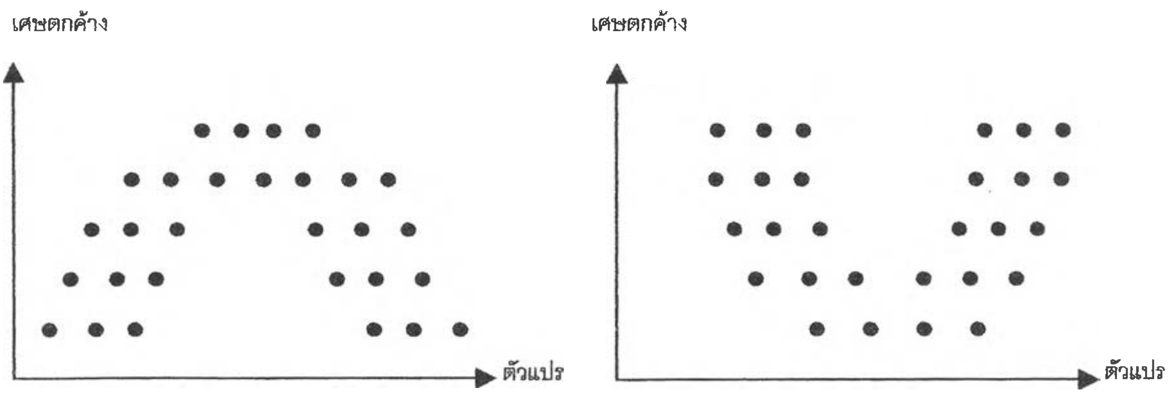
เศษตกค้าง



รูป 2.3



รูป 2.4



รูป 2.5

ถ้ากราฟระหว่าง e_i (หรือ $e_i/\sqrt{MSE}$) และ $\hat{y}_i$ และตัวแปรอิสระแต่ละตัว และกับเวลา (ถ้าเป็นอนุกรมเวลา) ทั้งหมดมีรูปแบบการกระจายของจุดไม่มีรูปแบบ ดังรูป 2.2 แสดงว่าตัวแบบการถดถอยเชิงเส้นมีรูปแบบเหมาะสม ค่าคลาดเคลื่อนสุ่มมีค่าเฉลี่ยศูนย์และความแปรปรวนคงที่ นอกจากนี้ในกรณีที่เป็นกราฟ e_i กับเวลาแสดงด้วยว่าความสัมพันธ์ระหวางตัวแปรตามและตัวแปรอิสระไม่แปรเปลี่ยนตามเวลา หรือไม่มีอิทธิพลของเวลา และค่าคลาดเคลื่อนสุ่มไม่มีอัตสหพันธ์กัน แต่ถ้ากราฟมีรูปแบบไม่เป็นแนวขนาน ดังเช่นรูป 2.3 ถึง รูป 2.5 แสดงว่าค่าคลาดเคลื่อนสุ่มไม่สอดคล้องคุณสมบัติบางข้อหรือทุกข้อ หรือตัวแบบยังมีรูปแบบไม่เหมาะสม

ถ้ารูปกราฟปรากฏดังรูป 2.3 แสดงว่า ความแปรปรวนของค่าคลาดเคลื่อนสุ่มไม่คงที่ การแก้ปัญหาอาจใช้วิธีแปลงค่าของตัวแปรตาม Y (เช่นทดลองด้วยการแปลงเป็น $\ln Y, \sqrt{Y}, 1/Y$ หรือ $1/\sqrt{Y}$ เป็นต้น) หรือใช้วิธีแก้ปัญหาเฉพาะในเรื่องนี้ เช่นใช้วิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนัก (Weighted Least Squares Method)

ในกรณีที่รูปภาพปรากฏดังรูป 2.4 หรือ 2.5 แสดงว่า ตัวแบบการถดถอยเชิงเส้นยังมีรูปแบบไม่เหมาะสม ในรูป 2.4 แสดงว่า ควรมีเทอมค่าคงที่หรือองค์ประกอบเชิงเส้นในตัวแบบ และในรูป 2.4 แสดงว่าควรมีเทอมที่มีกำลังสูงขึ้นในตัวแบบ เช่น ควรมีเทอมกำลังสอง αx^2 ของตัวแปรอิสระ x ในตัวแบบ และถ้าข้อมูลเป็นอนุกรมเวลากราฟรูป 2.4 แสดงว่า ควรมีเทอมเชิงเส้นหรือเทอมอันดับหนึ่งของเวลาในตัวแบบ และกราฟรูป 2.5 แสดงว่าควรมีเทอมอันดับหนึ่งและอันดับสองของเวลาในตัวแบบด้วย

การตรวจสอบค่าคลาดเคลื่อนสุ่มมีสหสัมพันธ์กันหรือไม่ โดยปกติจะตรวจสอบเมื่อข้อมูลเป็นอนุกรมเวลา และวิธีการตรวจสอบมิหลายวิธี เช่น ใช้วิธีพิจารณาค่าของฟังก์ชันอัตสหสัมพันธ์ตัวอย่าง [Sample Autocorrelation Function (SACF)] r_k ของ $e_1, e_2, \ldots, e_n$ ที่คาบเวลาห่างกัน $k, (k = 1, 2, \ldots)$ เปรียบเทียบกับค่าตัดสินเชิงสถิติหรือค่าวิกฤต (Critical value) $2/\sqrt{n}$ โดยประมาณ ที่ระดับนัยสำคัญ (significance level) 0.05 หรือใช้ค่าของตัวสถิติ Durbin-Watson สำหรับตรวจสอบอัตสหสัมพันธ์ที่คาบเวลาห่างกัน 1 ($k = 1$) เมื่อพบว่าค่าผิดพลาดมีสหสัมพันธ์กัน ในการวิเคราะห์การถดถอยเชิงเส้น จะมีวิธีการแก้ปัญหานี้โดยเฉพาะเช่นวิธีการแปลงตัวแปรและวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนัก เป็นต้น

4. การเปรียบเทียบตัวแบบพยากรณ์

โดยปกตินักพยากรณ์จะสร้างตัวแบบมากกว่าหนึ่งตัวแบบ สำหรับการพยากรณ์เรื่องหนึ่ง เพื่อคัดเลือกตัวแบบที่เหมาะสมกับข้อมูลในเรื่องนั้น ๆ โดยพิจารณาจากค่าความคลาดเคลื่อนว่าตัวแบบใดให้ค่าพยากรณ์ที่มีความคลาดเคลื่อนต่ำที่สุด

2.2.3 **ตัวแบบอนุกรมเวลา** (Time Serie Moels) ซึ่งวิธีในการสร้างตัวแบบอนุกรมเวลามีหลายวิธีดังต่อไปนี้

1. วิธีการปรับให้เรียบแบบเอกซโปเนนเชียล (Exponential Smoothing Method)

วิธีการปรับให้เรียบแบบเอกซโปเนนเชียล เป็นอีกวิธีการหนึ่งของการวิเคราะห์ข้อมูลอนุกรมเวลา เพื่อสร้างตัวแบบพยากรณ์ วิธีการนี้ใช้หลักการหาค่าเฉลี่ยถ่วงหน้ำหนัก (Weighted Average) ของข้อมูลในอดีต โดยให้น้ำหนัก (ความสำคัญ) มากสุดกับข้อมูลปัจจุบัน และจากนั้นให้น้ำหนักลดลงเรื่อย ๆ กับข้อมูลในอดีต (เนื่องจากมีแนวคิดว่าผลกระทบของข้อมูล หรือค่าสังเกตปัจจุบันที่มีต่อค่าในอนาคต จะมากกว่าผลกระทบของข้อมูลในอดีต) ยิ่งข้อมูลถอยหลังไปมาก ๆ จะมีน้ำหนักลดลงมาก ซึ่งลักษณะการให้น้ำหนักจะให้น้ำหนักลดลงแบบเรขาคณิต (Geometric) หรือแบบเอกซโพเนนเชียล (Exponential) น้ำหนัก จะมีค่าอยู่ระหว่าง 0 และ 1 และรวมกันเท่ากับ 1 [ตัวอย่างเช่น $Y_t = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \ldots, (0 < \alpha < 1)$]

1.1 วิธีการปรับให้เรียบครั้งเดียวแบบเอกซโปเนนเชียล (Simple Exponential Smoothing Method) เป็นวิธีการหนึ่งในกลุ่มวิธีการปรับให้เรียบแบบเอกซโพเนนเชียล วิธีการนี้เหมาะที่จะใช้กับอนุกรมเวลาที่มีระดับค่าเฉลี่ยไม่คงที่ โดยมีระดับเคลื่อนไหวช้า ๆ โดยไม่มีแนวโน้ม ไม่มีวัฏจักรและไม่มีฤดูกาล การพยากรณ์ค่าของ *Y* ที่เวลา *t+1*จากตำแหน่งเวลาปัจจุบัน *t* ด้วยวิธีการปรับให้เรียบครั้งเดียวแบบเอกซโพเนนเชียล มีสูตรพยากรณ์ ดังนี้

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

โดยที่

F_t = ค่าพยากรณ์สำหรับค่าของ *Y* ที่เวลา *t+1* จากเวลาปัจจุบัน *t* , *l* =1,2,3,...

F_{t-1}= ค่าพยากรณ์สำหรับค่าของ *Y* ที่เวลา *t-1*+1 จากเวลาปัจจุบัน *t*-1 , *l* =1,2,3,...

Y_t = ค่าสังเกต (ค่าจริง) ของ *Y* ที่เวลา *t*

α = ตัวประกอบปรับให้เรียบ (Smoothing Factor) มีค่าอยู่ระหว่าง 0 กับ 1

จากสูตรพยากรณ์ข้างต้นพบว่าการหาค่าพยากรณ์ F_t จะต้องทราบค่า F_{t-1} จะต้องทราบค่า F_{t-2} ...จะต้องทราบค่า F_1 และสุดท้ายจะต้องทราบค่า F_0 เมื่อทราบค่าเริ่มต้น F_0 จะหาค่า $F_1, F_2, \ldots, F_{t-1}$ และสุดท้ายหาค่า F_t ได้ โดยใช้สูตรข้างต้นต่อเนื่องกัน หนทางหนึ่งในการกำหนดค่าเริ่มต้น F_0 คือ ให้ F_0 เท่ากับค่าเฉลี่ยของของข้อมูลอนุกรมเวลา $Y_1, Y_2, \ldots, Y_t$ ที่มีอยู่ของ *Y* นั่นคือ ให้ $F_0 = \frac{1}{t}(Y_1 + Y_2 + \ldots + Y_t)$

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

1.2. วิธีการปรับให้เรียบสองครั้งแบบเอกซโปเนนเชียล (Double Exponential Smoothing Method) จากแนวคิดการปรับให้เรียบครั้งเดียวแบบเอกซโพเนนเชียล นำมาขยายผลใช้กับข้อมูลอนุกรมเวลาที่มีแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วงเวลา T วิธีนี้เหมาะสมสำหรับข้อมูลที่มีการเคลื่อนไหวแบบแนวโน้มเชิงเส้น (linear trend data) และไม่มีการเคลื่อนไหวแบบฤดูกาล (seasonal data) เหมาะสมกับการพยากรณ์ในระยะสั้นจนถึงการพยากรณ์ในระยะปานกลาง วิธีนี้มีสูตรการพยากรณ์ค่าจริง Y_{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 $S_{T-2}^{[1]}, S_{T-2}^{[2]}, S_{T-3}^{[1]}, S_{T-3}^{[2]}, \ldots, S_0^{[1]}, S_0^{[2]}$ ซึ่งต้องเริ่มด้วยค่า $S_0^{[1]}$ และ $S_0^{[2]}$ เราประมาณค่า $S_0^{[1]}$ และ $S_0^{[2]}$ ได้ดังนี้

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

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

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

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

1.3 วิธีพารามิเตอร์สองตัวของโฮลท์ (Holt's Two –parameter Method) วิธีการของโฮลท์ มีลักษณะคล้ายกับวิธีการปรับให้เรียบสองครั้งแบบเอกซโพเนนเชียล แต่มีลักษณะทั่วไปมากกว่าวิธีนี้เหมาะสำหรับข้อมูลที่มีการเคลื่อนไหวแบบแนวเชิงเส้น มีสูตรการพยากรณ์ดังนี้

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

ซึ่ง

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

2.วิธีอัตถดถอย (Autoregressive Method)

วิธีอัตถดถอย เป็นวิธีการที่จัดอยู่ในกลุ่มวิเคราะห์อนุกรมเวลา เพื่อหาตัวแบบพยากรณ์ค่าของตัวแปร Y ที่ต้องการเช่นตัวแปร Y อัตราการว่างงานเป็นต้น ตัวแบบจะเป็นตัวแบบการถดถอยที่แสดงความสัมพันธ์ระหว่างค่า Y_t ของ Y ณ เวลา t กับค่าของ Y ในอดีตให้เป็น $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k}$ ได้เป็นตัวแบบอัตถดถอยและมีรูปแบบทั่วไปดังนี้

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

ซึ่งสมมติว่าค่าคลาดเคลื่อนสุ่ม ε_t มีค่าเฉลี่ยเท่ากับศูนย์ มีค่าแปรปรวนคงที่ ไม่มีอัตสหสัมพันธ์ และมีการแจกแจงปกติ สำหรับเวลาถอยหลัง k โดยทั่วไป ซึ่งในงานวิจัยนี้จะเลือก k=12 จากตัวแบบข้างต้นได้ตัวแบบพยากรณ์

$$Y_t = C_0 + C_1 Y_{t-1} + C_2 Y_{t-2} + \ldots + C_k Y_{t-k}$$

ซึ่ง $C_1, C_2, \ldots, C_k$ เป็นค่าประมาณของ $\theta_0, \theta_1, \theta_2, \ldots, \theta_k$ ตามลำดับ ทั้งนี้ พิจารณานัยสำคัญของ $C_1, C_2, \ldots, C_k$ ด้วย ซึ่งถ้าค่าใดไม่มีนัยสำคัญ จะตัดค่านั้นออกจากตัวแบบ

3.วิธีการวิเคราะห์อนุกรมเวลาแบบคลาสสิค (Classical Time Series Analysis) อนุกรมเวลา $\{Y_t\}$ จะมีการเคลื่อนไหวตามกาลเวลาในรูปแบบต่างๆ เช่นมีแนวโน้มและมีฤดูกาล ฉะนั้นหนทางหนึ่งในการวิเคราะห์อนุกรมเวลา เพื่อหาตัวแบบพยากรณ์ คือ วิเคราะห์องค์ประกอบในข้อมูลอนุกรมเวลาที่ศึกษามาสร้างเป็นตัวแบบพยากรณ์

องค์ประกอบของอนุกรมเวลาจำแนกได้เป็น 3 องค์ประกอบหลัก คือ องค์ประกอบแนวโน้ม-วัฏจักร (trend-cycle component) องค์ประกอบฤดูกาล (seasonal component) และองค์ประกอบไม่ปกติ (irregular or remainder component) ดังนั้นโดยวิธีนี้ จะทำการแยกองค์ประกอบ ซึ่งได้ตัวแบบสำหรับการพยากรณ์ค่า Y_t และเรียกตัวแบบที่ได้ว่า ตัวแบบอนุกรมเวลาแบบคลาสสิค ซึ่งจำแนกได้เป็นสองตัวแบบคือ

1.ตัวแบบเชิงบวก (Additive Model)

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

2.ตัวแบบชิงคูณ (Multiplicative Model)

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

โดยที่

Y_t คือ ค่าของอนุกรมเวลา ณ คาบเวลา t

ε_t คือ ค่าคลาดเคลื่อนสุ่ม

T_t คือ องค์ประกอบแนวโน้ม –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 คือ องค์ประกอบฤดูกาล ณ คาบเวลา t ซึ่งอาจจะอยู่ในรูปแบบของตัวแปรบ่งชี้ฤดูกาล (seasonal indicators) เช่น

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

หรือ อยู่ในรูปแบบฟังก์ชันตรีโกณ เช่น

$$S_t = \phi_1 \sin\left(\frac{2\pi t}{12}\right) + \phi_2 \cos\left(\frac{2\pi t}{12}\right) + \phi_3 \sin\left(\frac{4\pi t}{12}\right) + \varphi_4 \cos\left(\frac{4\pi t}{12}\right)$$

ตัวแบบเชิงบวก เหมาะสำหรับข้อมูลอนุกรมเวลาที่มีความแปรปรวนคงที่ ไม่แปรผันตามเวลา หรือการแกว่งของฤดูกาลไม่แปรผันตามระดับของอนุกรมเวลา มิฉะนั้น ตัวแบบเชิงคูณจะเหมาะสมกว่า อย่างไรก็ตามเราสามารถใช้วิธีการแปลงข้อมูลอนุกรมเวลาที่ไม่คงที่ในความแปรปรวนหรือแปลงตัวแบบเชิงคูณเป็นตัวแบบเชิงบวกได้ โดยการใส่ ln ซึ่งจะได้ตัวแบบ

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

นั่นคือ ใช้ตัวแบบเชิงบวกกับข้อมูลที่ใส่ ln

ตัวอย่างตัวแบบอนุกรมเวลาแบบคลาสิค

$$Y_t = b_0 + b_1 t + b_2 I_{1,t} + b_3 I_{2,t} + \varepsilon_t$$

โดยที่ t แทนคาบเวลา

$I_{i,t}$ แทนตัวบ่งชี้สำหรับไตรมาสที่ i .ในคาบเวลา t (ให้ $I_{i,t} = i$ สำหรับไตรมาสที่ i ในคาบเวลา t มิฉะนั้น ให้ $I_{i,t}$ =0)

4.ตัวแบบการถดถอยที่มีค่าคลาดเคลื่อนในรูปแบบ AR (Regression Model with AR errors) จากตัวแบบการถดถอยในรูปแบบทั่วไป

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

โดยทั่วไป จะมีข้อสมมติข้อหนึ่งที่ว่าค่าคลาดเคลื่อนสุ่ม ε_t ไม่มีอัตสหสัมพันธ์แต่ถ้า Y_t เป็นข้อมูลอนุกรมเวลา มักจะพบว่า ε_t มีอัตสหสัมพันธ์ และรูปแบบอัตสหสัมพันธ์รูปแบบหนึ่งที่พบบ่อยเสมอ คือ รูปแบบ AR (1) (first – order autoregressive model) ซึ่งถ้า ε_t มีอัตสหสัมพันธ์ AR (1) จะเขียนตัวแบบการถดถอยข้างต้นได้เป็น

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

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

ซึ่งมีข้อสมมติ e_t มีค่าเฉลี่ยเท่ากับศูนย์มีความแปรปรวนคงที่ไม่มีอัตสหสัมพันธ์ และมีการแจกแจงปรกติ

ตัวแบบพยากรณ์เมื่อค่าคลาดเคลื่อนสุ่ม ε_t มีอัตสหสัมพันธ์ AR (1) เขียนได้ดังนี้

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

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

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