บทที่ 1

บทนำ

ความเป็นมาและความสำคัญปัญหา

ในการวางแผนงาน ข้อมูลเป็นสิ่งสำคัญที่นำมาใช้ประกอบการตัดสินใจ ซึ่งการเก็บข้อมูลทำได้หลายวิธี วิธีที่นิยมใช้มากวิธีหนึ่ง คือ การเก็บบันทึกข้อมูลเป็นตัวเลขตามลำดับของเวลาในช่วงต่าง ๆ เช่น จำนวนนักเรียนรายปี จำนวนครูรายปี ค่าใช้จ่ายรายเดือน เป็นต้น ข้อมูลตัวเลขชุดที่เก็บรวบรวมในลักษณะเช่นนี้เรียกว่า อนุกรมเวลา จากการวิเคราะห์อนุกรมเวลาจะทำให้ทราบถึงรูปแบบการเปลี่ยนแปลงของข้อมูลเมื่อเวลาเปลี่ยนไป และในกรณีที่สถานการณ์และสิ่งแวดล้อมไม่เปลี่ยนแปลงมากนัก สิ่งที่เกิดขึ้นในอดีตย่อมชี้ให้เห็นสิ่งที่จะเกิดขึ้นในอนาคต กล่าวคือรูปแบบการเปลี่ยนแปลงของข้อมูลอนุกรมเวลาจะมีประโยชน์ในการพยากรณ์ค่าในอนาคต ดังนั้นเทคนิคการพยากรณ์จึงมีความจำเป็นอย่างยิ่งต่อการวางแผนและการตัดสินใจ การพยากรณ์ที่เหมาะสมและแม่นยำจะช่วยลดความเสี่ยงหรือความผิดพลาดในการตัดสินใจหรือการวางแผน

การพยากรณ์ แบ่งได้ 2 กลุ่ม คือ การพยากรณ์เชิงคุณภาพ (qualitative forecasting) และการพยากรณ์เชิงปริมาณ (quantitative forecasting) การพยากรณ์เชิงคุณภาพเป็นการพยากรณ์ที่ต้องใช้ความรู้ ความสามารถ และประสบการณ์ในเรื่องที่จะพยากรณ์และเรื่องที่เกี่ยวข้อง ส่วนการพยากรณ์เชิงปริมาณจะเน้นการใช้รายละเอียดของข้อมูลในอดีตมาเป็นแนวทางในการพยากรณ์ โดยใช้เทคนิคและวิธีการพยากรณ์ต่าง ๆ ที่ได้มีผู้พัฒนาและเสนอแนะไว้ ปัจจุบันมีผู้ใช้การพยากรณ์แพร่หลายมากขึ้น วิธีการพยากรณ์ดังกล่าว ได้แก่ วิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนน เชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) วิธีการวิเคราะห์การถดถอย (Regression analysis: REG) วิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) และวิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) เป็นต้น

มีงานวิจัยในประเทศไทยที่ใช้เทคนิคการพยากรณ์เชิงปริมาณทางการศึกษา ที่ใช้โมเดลทางเศรษฐศาสตร์มาประยุกต์ใช้ในการพยากรณ์ความต้องการกำลังคน ได้แก่ นงลักษณ์ วิรัชชัย (2513), พรรณมาศ ตันฉาย (2513), บุญธรรม กิจปรีดาพิสุทธิ์ (2513), นิตยา ภัสสรศิริ (2514), สมหวัง พิธิยานุวัฒน์ (2514), อังคณา พัฒนผลไพบูลย์ (2531) และเงนุช อินทรวงษ์โชติ (2538) งานวิจัยที่ใช้วิธีง่าย (naive) เพื่อพยากรณ์ความต้องการครูระดับประถมและมัธยมศึกษา ได้แก่ ณัตติณา วัฒนาชยากูล (2513), นภาพร สิงหทัต (2518), วรรณพร วิเชียรวงศ์ (2521) และกานต์ กุณาศล (2525) งานวิจัยที่ใช้การพยากรณ์ด้วยวิธีบ็อกซ์-เจนกิ้นส์ เพื่อพยากรณ์ข้อมูลอนุกรมเวลาทางเศรษฐศาสตร์และข้อมูลอนุกรมเวลาทางการศึกษา ได้แก่

วันพร เหลืองอาภาพงศ์ (2520), บุษบา พิกุลผล (2522), เกศิณี กมลรัตน์ (2530), ธิดารัตน์ จันทวี (2539), บำเพ็ญ ปิดชิด (2540), เอกภพ ยานะวิมุติ (2543) และอรุณี หงษ์ศิริวัฒน์ (2543)

จากผลการวิจัยที่ใช้เทคนิคการพยากรณ์ข้อมูลอนุกรมเวลาในประเทศไทย พบว่า วิธีบ็อกซ์-เจนกิ้นส์ เป็นวิธีการพยากรณ์ที่ให้ค่าพยากรณ์ที่มีค่าความถูกต้อง (accuracy) ค่อนข้างสูงกว่าวิธีอื่นในการพยากรณ์ระยะสั้น (short term forecasting) และไม่มีการกำหนดโมเดลความสัมพันธ์ขึ้นมาก่อนการวิเคราะห์ (บุษบา พิกุลผล; 2522, เกศิณี กมลรัตน์; 2530, ธิดารัตน์ จันทวี; 2539, บำเพ็ญ ปิดชิด; 2540, เอกภพ ยานะวิมุติ; 2543 และอรุณี หงษ์ศิริวัฒน์; 2543) แต่อย่างไรก็ตาม วิธีบ็อกซ์-เจนกิ้นส์ เป็นวิธีที่ค่อนข้างยุ่งยากในการวิเคราะห์ วิธีการคำนวณสลับซับซ้อน เหมาะกับข้อมูลอนุกรมเวลาที่มีจำนวนมาก

ในทางปฏิบัติข้อมูลอนุกรมเวลาทางการศึกษามีการจัดเก็บไว้ค่อนข้างน้อย ซึ่งส่วนใหญ่เก็บไว้ไม่เกิน 5 ปี กรณีที่มีข้อมูลไม่มากเทคนิคการพยากรณ์ที่เหมาะสม ได้แก่ วิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (single exponential smoothing) การแยกส่วนประกอบ (decomposition) อย่างไรก็ตามวิธีการเหล่านี้มีข้อจำกัดเกี่ยวกับข้อตกลงเบื้องต้นว่า ต้องมีการกำหนดโมเดลหรือแบบแผนแนวโน้มของอนุกรมขึ้นมาก่อนที่จะทำการวิเคราะห์ ถ้ามีการกำหนดโมเดลไม่เหมาะสม จะทำให้เกิดความคลาดเคลื่อนในการพยากรณ์ได้สูง

จากการศึกษาความแม่นยำของเทคนิคการพยากรณ์ โดย Makridakis และคณะ (1982) โดยใช้เทคนิคการพยากรณ์ 24 วิธี ในการพยากรณ์ข้อมูลอนุกรมเวลา 1,001 ชุด ที่เรียกว่า M-Competition อนุกรมเวลาทั้งหมดเป็นข้อมูลอุตสาหกรรม รายเดือน รายไตรมาส และรายปี ทำการพยากรณ์ล่วงหน้า 1, 2, 3, 6, 9, 12, 15 และ 18 ช่วงเวลา จากการศึกษานี้แสดงให้เห็นว่า ไม่มีวิธีใดที่ดีที่สุดสำหรับข้อมูลทุกชุดและการพยากรณ์ล่วงหน้าไปทุกช่วงเวลา บางวิธีอาจมีความแม่นยำกว่าบางวิธีในระยะสั้น แต่บางวิธีอาจมีความแม่นยำในระยะยาว และสังเกตได้ว่าวิธีที่ง่ายบางครั้งมีความแม่นยำเทียบเท่าวิธีที่ยาก ๆ ดังนั้นจึงต้องอาศัยประสบการณ์และความชำนาญเพื่อเลือกวิธีที่ดีที่สุดและเหมาะสมที่สุดในการพยากรณ์ ทำให้ผู้พยากรณ์ที่ไม่มีประสบการณ์และความชำนาญเพียงพอเกิดความยุ่งยากและลำบากในการตัดสินใจที่จะเลือกวิธีพยากรณ์ว่าควรใช้วิธีใดจึงจะเหมาะสมและมีความถูกต้องมากที่สุด และจากงานวิจัยนี้ Makridakis และคณะ ได้ใช้เทคนิคการพยากรณ์วิธีหนึ่งที่น่าสนใจ คือ combining forecasts ซึ่งเป็นการรวมหลาย ๆ วิธีพยากรณ์เข้าด้วยกัน แล้วพบว่าวิธีดังกล่าวมีความสามารถในการพยากรณ์ได้แม่นยำขึ้น ดังนั้นจึงน่าจะเป็นอีกแนวทางเลือกหนึ่งสำหรับผู้พยากรณ์

จากงานวิจัยของ Winkler และ Makridakis (1983) ได้ทำการศึกษาวิธีพยากรณ์ร่วม โดยให้น้ำหนักของ Newbold และ Granger (1974) 5 แบบ จากวิธีการพยากรณ์ 10 วิธี คือ วิธีง่าย (naive) วิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) วิธีการวิเคราะห์การถดถอย

(Regression analysis: REG) วิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) วิธีการปรับให้เรียบแบบปรับอัตราส่วน (adaptive response rate exponential smoothing: ARRES) และวิธี Automatic AEP โดยใช้ข้อมูล 1,001 ชุด (series) จากแหล่งข้อมูลต่าง ๆ หลายแหล่งและหลายประเภท มีทั้งข้อมูลที่เป็นรายปี รายไตรมาส และรายเดือน พบว่า วิธีการพยากรณ์ร่วมโดยการให้น้ำหนักเฉลี่ยของ Newbold และ Granger แบบที่ 1 และแบบที่ 3 เป็นวิธีที่ดีที่สุดในการพยากรณ์ โดยมีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) น้อยที่สุด และในปีเดียวกัน Makridakis และ Winkler ได้ทำการศึกษาเพื่อตรวจสอบความถูกต้องของการพยากรณ์ร่วม โดยใช้ข้อมูลอนุกรมเวลา 111 ชุด และ 1,001 ชุด ทั้งประเภทรายเดือน รายไตรมาส และรายปี ใช้วิธีการพยากรณ์ร่วมโดยให้น้ำหนักแบบ simple everage จากวิธีการพยากรณ์เดี่ยว 14 วิธี ได้แก่ วิธีง่าย (naive) วิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) วิธีการปรับให้เรียบแบบปรับอัตราส่วน (adaptive response rate exponential smoothing: ARRES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) วิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) วิธีการวิเคราะห์การถดถอย (Regression analysis: REG) วิธี Automatic AEP วิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) วิธี Lewandowski's FORSYS system วิธี Parzen's ARIMA methodology และวิธี Bayesian Forecasting โดยใช้ MAPE เป็นเกณฑ์ในการตัดสิน พบว่าความถูกต้องของการพยากรณ์ร่วมขึ้นอยู่กับจำนวนวิธีที่นำมารวมกัน ถ้าใช้จำนวนวิธีการพยากรณ์ในการนำมารวมกันเพิ่มมากขึ้น ค่าความถูกต้องก็จะเพิ่มตาม

และต่อมาทัศนีย์ อินทนู (2543) ได้นำผลการวิจัยของ Winkler และ Makridakis (1983) มาทำการศึกษาเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์เดี่ยว 8 วิธี ได้แก่ วิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) วิธีการวิเคราะห์การถดถอย (Regression analysis: REG) วิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) และวิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) และวิธีพยากรณ์ร่วม (combined forecast) จากวิธีพยากรณ์เดี่ยวดังที่กล่าวมา ยกเว้นวิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) โดยใช้น้ำหนักเฉลี่ยของ Newbold และ Granger 2 แบบที่ดีที่สุดจากผลการวิจัยของ Winkler และ Makridakis (1983) และใช้ค่าความคลาดเคลื่อน 2 แบบ คือ เปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) เป็นเกณฑ์ ฐานข้อมูลที่ใช้ มี 2 ฐาน ฐานแรก คือ ข้อมูลอนุกรมเวลารายเดือน 3 ชุด ได้แก่ ปริมาณการยืมหนังสือ

ภาษาไทย ภาษาอังกฤษ และวิทยานิพนธ์ 65 ช่วงเวลา ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เป็นข้อมูลที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ฐานที่ 2 คือ ข้อมูลอนุกรมเวลารายปี 2 ชุด ได้แก่ จำนวนครูวิทยาศาสตร์ และคณิตศาสตร์ 14 ปี ของกรมสามัญศึกษา กระทรวงศึกษาธิการ เป็นข้อมูลที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล พบว่า วิธีการพยากรณ์ที่เหมาะสมกับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ได้แก่ ปริมาณการยืมหนังสือภาษาไทย คือ วิธีการพยากรณ์ร่วม ที่ได้จากการรวมของวิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) กับวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) ซึ่งเป็นวิธีที่เหมาะสมกับข้อมูลทั้งสองวิธี โดยให้น้ำหนักแบบที่ 2 และวิทยานิพนธ์ คือ วิธีการพยากรณ์ร่วม ที่ได้จากการรวมของวิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) กับวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) ซึ่งเป็นวิธีที่เหมาะสมกับข้อมูลทั้งสองวิธี โดยให้น้ำหนักแบบที่ 1 สำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ได้แก่ จำนวนครูวิทยาศาสตร์และจำนวนครูคณิตศาสตร์ คือ วิธีการพยากรณ์ร่วม ที่ได้จากการนำวิธีการพยากรณ์ทุกวิธีมารวมกัน โดยให้น้ำหนักแบบที่ 1 และแบบที่ 2 ตามลำดับ

จากการศึกษาวิธีการพยากรณ์ร่วมพบว่ามีข้อดีหลายประการ และจากเหตุผลที่ว่า วิธีการพยากรณ์ที่จะนำมารวมกัน (combine) ยิ่งใช้วิธีมากขึ้นเท่าใด ค่าความถูกต้องก็จะเพิ่มมากตามไปด้วย และค่าความแปรปรวนของความคลาดเคลื่อนก็จะยิ่งลดลง แต่จะเห็นได้ว่าการที่นำวิธีพยากรณ์หลาย ๆ วิธีมารวมกันมากยิ่งขึ้นเท่าใดก็จะเกิดความยุ่งยากมากขึ้นเช่นเดียวกัน ทำให้พยากรณ์ได้ช้าและมีข้อจำกัดของข้อมูลตามวิธีการพยากรณ์ด้วย

ในปัจจุบันเทคนิคการพยากรณ์ได้พัฒนาไปมาก มีเทคนิคการพยากรณ์เกิดขึ้นหลายวิธี เทคนิคการพยากรณ์ที่น่าสนใจวิธีหนึ่ง คือ วิธีของเบย์ (Bayesian method) เป็นการนำเอาทฤษฎีความน่าจะเป็นเข้ามาช่วยในการพยากรณ์ ซึ่งเป็นทฤษฎีที่ได้ถูกนำมาใช้กันอย่างแพร่หลายในงานสถิติต่าง ๆ รูปแบบการพยากรณ์ด้วยวิธีของเบย์ มีขั้นตอนที่ทำให้ข้อมูลเป็นปัจจุบัน (update) อยู่เสมอ ไม่จำเป็นต้องใช้ข้อมูลขนาดใหญ่เหมือนกับวิธีบ็อกซ์-เจนกิ้นส์ และจากการศึกษางานวิจัยของ de Alba และ Mendoza (2001) ได้ทำการพยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีของเบย์ เปรียบเทียบกับวิธีบ็อกซ์-เจนกิ้นส์ โดยใช้ข้อมูล 2 ฐาน ฐานแรก คือ ข้อมูลปริมาณการใช้ไฟฟ้าของเมืองไอโอวา เป็นข้อมูลรายเดือนตั้งแต่เดือนมกราคม 1971 ถึงเดือนตุลาคม 1978 ซึ่งเป็นข้อมูลที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล (seasonal) และฐานที่สอง คือ ค่าใช้จ่ายในการจัดการธนาคาร Maxican เป็นข้อมูลรายเดือนตั้งแต่เดือนมกราคม 1987 ถึง เดือนตุลาคม 1994 ซึ่งเป็นข้อมูลที่มีอิทธิพลของแนวโน้ม (trend) โดยใช้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) เป็นเกณฑ์ พบว่า ในกรณีที่ข้อมูลอนุกรมเวลามีขนาดเล็ก วิธีของเบย์มีค่า MAPE ต่ำกว่าวิธีบ็อกซ์-เจนกิ้นส์ สำหรับข้อมูลทั้งสองชุด

จากการศึกษาการพยากรณ์ด้วยวิธีของเบย์ในงานวิจัยของ de Alba และ Mendoza (2001) และจากการค้นคว้างานวิจัยในประเทศไทย พบว่า ยังไม่มีการนำวิธีพยากรณ์นี้มาใช้พยากรณ์ทั้งข้อมูลอนุกรมเวลาทางเศรษฐศาสตร์และข้อมูลอนุกรมเวลาทางการศึกษา ผู้วิจัยจึงมีความสนใจที่จะนำวิธีการพยากรณ์นี้มาใช้พยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษา และนำมาเปรียบเทียบกับวิธีบ็อกซ์-เจนกิ้นส์ และวิธีการพยากรณ์

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ร่วม ซึ่งถือเป็นวิธีที่มีประสิทธิภาพ ว่าวิธีใดมีค่าความคลาดเคลื่อนในการพยากรณ์ต่ำที่สุด หรือมีความถูกต้องมากที่สุดในกรณีที่ใช้ข้อมูลขนาดเล็กเท่ากัน

ในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ผู้วิจัยใช้ข้อมูลจากฐานข้อมูลของทัศนีย์ อินทนู (2543) คือ ปริมาณการยืมสิ่งพิมพ์ ซึ่งแยกเป็นหนังสือภาษาไทย หนังสือภาษาอังกฤษ และวิทยานิพนธ์ ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ซึ่งมีลักษณะเป็นข้อมูลรายเดือน นอกจากนี้ผู้วิจัยได้เก็บรวบรวมข้อมูลอนุกรมเวลาชุดนี้เพิ่มเติมเพื่อให้มีความเป็นปัจจุบันมากขึ้น และทำการพยากรณ์ล่วงหน้า 12 ช่วงเวลา เพื่อเป็นประโยชน์ต่อการวางแผนและการบริหารงานของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ส่วนข้อมูลทางการศึกษาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ผู้วิจัยใช้ข้อมูลจากฐานข้อมูลของบำเพ็ญ ปิดชิด (2540) ได้แก่ จำนวนนักเรียนในระดับประถมศึกษา และระดับมัธยมศึกษา มีลักษณะข้อมูลเป็นรายปี และผู้วิจัยได้เก็บรวบรวมข้อมูลอนุกรมเวลาชุดนี้เพิ่มเติมเพื่อให้มีความเป็นปัจจุบันมากขึ้น และทำการพยากรณ์ล่วงหน้า 5 ช่วงเวลา เพื่อประโยชน์ในการวางแผนเพื่อพัฒนาการศึกษาให้สอดคล้องกับการปฏิรูปการศึกษาตามพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542

วัตถุประสงค์ของการวิจัย

1.เพื่อเปรียบเทียบความคลาดเคลื่อนระหว่างวิธีพยากรณ์ด้วยวิธีของเบย์ (Bayesian method) วิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) และวิธีการพยากรณ์ร่วม (combined forecasting) ในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่มีการเปลี่ยนแปลงและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล โดยใช้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) เป็นเกณฑ์

2.เพื่อพยากรณ์การยืมสิ่งพิมพ์ ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยเป็นการพยากรณ์ระยะปานกลาง 12 เดือนล่วงหน้า และพยากรณ์จำนวนนักเรียนในระดับประถมศึกษาและระดับมัธยมศึกษา ซึ่งเป็นการพยากรณ์ระยะยาว 5 ปีล่วงหน้า

สมมติฐานการวิจัย

จากผลการวิจัยของ de Alba และ Mendoza (2001) พบว่า ในกรณีที่ข้อมูลอนุกรมเวลามีขนาดเล็ก วิธีของเบย์มีค่า MAPE ต่ำกว่าวิธีบ็อกซ์-เจนกิ้นส์ สำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล และจากผลการวิจัยของทัศนีย์ อินทนู (2543) พบว่า วิธีการพยากรณ์ที่เหมาะสมกับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล คือ วิธีการพยากรณ์ร่วม ที่ได้จากการรวมของวิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) กับวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) ผู้วิจัยจึงตั้งสมมติฐานการวิจัยข้อแรก ดังนี้

1. ในกรณีข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลสำหรับข้อมูลขนาดเล็ก วิธีการพยากรณ์ที่มีความคลาดเคลื่อนในการพยากรณ์น้อยที่สุดหรือมีความถูกต้องในการพยากรณ์มากที่สุด คือ วิธี

ของเบย์ (Bayesian method) รองลงมาคือ วิธีการพยากรณ์ร่วม (combined forecasting) และวิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) ตามลำดับ

จากผลการวิจัยของ de Alba และ Mendoza (2001) พบว่า ในกรณีที่ข้อมูลอนุกรมเวลามีขนาดเล็ก วิธีของเบย์มีค่า MAPE ต่ำกว่าวิธีบ็อกซ์-เจนกิ้นส์ สำหรับข้อมูลอนุกรมเวลาที่มีแนวโน้ม แต่ยังไม่มีผลงานวิจัยเกี่ยวกับวิธีพยากรณ์ร่วม ในกรณีนี้ ดังนั้นผู้วิจัยจึงตั้งสมมติฐานข้อสอง ดังนี้

2. ในกรณีข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลสำหรับข้อมูลขนาดเล็ก วิธีการพยากรณ์ทั้งสามวิธีมีความคลาดเคลื่อนในการพยากรณ์หรือมีความถูกต้องในการพยากรณ์ไม่แตกต่างกัน

ขอบเขตของการวิจัย

การวิจัยในครั้งนี้ใช้วิธีการพยากรณ์ด้วยวิธีของเบย์ (Bayesian method) วิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) และวิธีการพยากรณ์ร่วม (combined forecasting) 2 วิธี ได้แก่ วิธีที่ 1 ได้จากการนำค่าพยากรณ์จากวิธีพยากรณ์เดี่ยวที่เหมาะสมกับลักษณะของข้อมูลมารวมกันโดยให้น้ำหนักตามแบบของ Newbold และ Granger แบบที่ 1 และวิธีที่ 2 ได้จากการนำค่าพยากรณ์จากวิธีพยากรณ์เดี่ยวที่เหมาะสมกับลักษณะของข้อมูลมารวมกันโดยให้น้ำหนักตามแบบของ Newbold และ Granger แบบที่ 2 โดยพิจารณาเลือกวิธีพยากรณ์เดี่ยวที่เหมาะสมกับลักษณะของข้อมูลจาก 3 กรณี คือ กรณีแรก เมื่อข้อมูลไม่มีอิทธิพลของแนวโน้มและฤดูกาลจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) และวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) กรณีที่สอง เมื่อข้อมูลมีอิทธิพลของแนวโน้มเพียงอย่างเดียวจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสำหรับข้อมูลที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียล (exponential smoothing for exponential trend: ESE) และวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) และกรณีที่สาม เมื่อข้อมูลมีอิทธิพลของแนวโน้มและฤดูกาลจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) และวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) โดยใช้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่าความคลาดเคลื่อนเฉลี่ย (ME) เป็นเกณฑ์ในการเปรียบเทียบ

ข้อจำกัดของการวิจัย

การพยากรณ์ด้วยวิธีของเบย์ ในขั้นการกำหนดโมเดลสำหรับข้อมูลที่มีแนวโน้มและฤดูกาลสามารถกำหนดได้ 2 โมเดล คือ

1) $y_t = b_1 + b_2 t + b_3 \sin\frac{2\pi t}{L} + b_4 \cos\frac{2\pi t}{L} + \varepsilon_t$

และ 2) $y_t = b_1 + b_2\, t + b_3 \sin\frac{2\pi t}{L} + b_4 \cos\frac{2\pi t}{L} + b_5 \sin\frac{4\pi t}{L} + b_6 \cos\frac{4\pi t}{L} + \varepsilon_t$

ในงานวิจัยนี้จะเลือกใช้เพียงโมเดลที่ 1 เพราะวิธีการคำนวณไม่ยุ่งยากและสะดวกต่อการนำมาใช้มากกว่าโมเดลที่ 2 ซึ่งมีความสลับซับซ้อนมากในการคำนวณ ไม่สะดวกต่อการนำมาใช้

ข้อตกลงเบื้องต้นของการวิจัย

โครงสร้างการเปลี่ยนแปลงของข้อมูลในอดีต ปัจจุบัน และอนาคต เป็นแบบเดียวกัน

นิยามศัพท์เฉพาะที่ใช้ในการวิจัย

การพยากรณ์ร่วม หมายถึง การนำวิธีการพยากรณ์หลาย ๆ วิธีมารวมกัน โดยให้น้ำหนักตามแบบของ Newbold และ Granger 2 แบบ คือ

1. $$w_i = \frac{\left(\sum_{t=n-v}^{n-1} e_t^{(i)2}\right)^{-1}}{\sum_{j=1}^{p}\left(\sum_{t=n-v}^{n-1} e_t^{(j)2}\right)^{-1}}$$

2. $$w_i = \beta\left[\frac{\left(\sum_{t=n-v-1}^{n-2} e_t^{(i)2}\right)^{-1}}{\sum_{j=1}^{p}\left(\sum_{t=n-v-1}^{n-2} e_t^{(j)2}\right)^{-1}}\right] + (1-\beta)\left[\frac{\left(\sum_{t=n-v}^{n-1} e_t^{(i)2}\right)^{-1}}{\sum_{j=1}^{p}\left(\sum_{t=n-v}^{n-1} e_t^{(j)2}\right)^{-1}}\right], \qquad 0<\beta<1$$

ค่าการพยากรณ์ คือ $\hat{y}_t = \sum_{i=1}^{p} w_i \hat{y}_t^{(i)}$

โดยที่ $e_t^{(i)} = \frac{(y_t - \hat{y}_t^{(i)})}{y_t}$, i = 1,2,3,...,p

เมื่อ n = จำนวนข้อมูล

$\hat{y}_t^{(i)}$ = ค่าพยากรณ์ ณ เวลาที่ t วิธีพยากรณ์ที่ i

$\hat{y}_t$ = ค่าพยากรณ์ ณ เวลาที่ t

w_i = น้ำหนักของวิธีการพยากรณ์ที่ i

p = จำนวนวิธีการพยากรณ์

ข้อมูลอนุกรมเวลา หมายถึง ข้อมูลที่ได้จากค่าสังเกตชุดหนึ่ง ซึ่งถูกจัดเรียงกันตามลำดับการเกิดขึ้นก่อนหลัง โดยที่ข้อมูลชุดนี้จะถูกเก็บมา ณ ช่วงเวลาที่ห่างเท่า ๆ กัน เช่น เก็บเป็นรายเดือน รายไตรมาส หรือรายปี

สิ่งพิมพ์ หมายถึง หนังสือภาษาไทย หนังสือภาษาอังกฤษ และวิทยานิพนธ์ ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่บริการให้ผู้ใช้ยืมออกจากศูนย์บรรณสารสนเทศทางการศึกษาได้

ข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล หมายถึง ข้อมูลที่ได้จากค่าสังเกตชุดหนึ่งที่มีการเคลื่อนไหวขึ้นลงตามระยะเวลาเป็นช่วงที่แน่นอน และลักษณะการเคลื่อนไหวในระยะเวลาหนึ่งจะคล้ายกับช่วงเวลาอื่นซ้ำ ๆ กัน ในการวิจัยครั้งนี้ใช้ข้อมูลปริมาณการยืมสิ่งพิมพ์ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ได้แก่ หนังสือภาษาไทย หนังสือภาษาอังกฤษ และวิทยานิพนธ์

ข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล หมายถึง ข้อมูลที่ได้จากค่าสังเกตชุดหนึ่งที่มีการเคลื่อนไหวขึ้นลงตามระยะเวลาเป็นช่วงที่ไม่แน่นอน ในการวิจัยครั้งนี้ใช้ข้อมูลจำนวนนักเรียนในระดับประถมศึกษา และระดับมัธยมศึกษา

ข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงที่แสดงแนวโน้ม หมายถึง ข้อมูลที่ได้จากค่าสังเกตชุดหนึ่งที่มีการเคลื่อนไหวเปลี่ยนแปลงไปตามระยะเวลา อาจเป็นเส้นตรง เส้นโค้งแบบครอดราติก หรือเส้นโค้งแบบเอ็กซ์โปเนนเชียล

แนวโน้มแบบเส้นโค้ง หมายถึง การเปลี่ยนแปลงที่แสดงแนวโน้มเส้นโค้งแบบควอดราติก

ข้อมูลขนาดเล็ก หมายถึง ข้อมูลอนุกรมเวลาที่มีขนาดน้อยกว่า 30 ช่วงเวลา

ข้อมูลขนาดใหญ่ หมายถึง ข้อมูลอนุกรมเวลาที่มีขนาดมากกว่าหรือเท่ากับ 30 ช่วงเวลา

ค่าจริง หมายถึง ค่าของข้อมูลอนุกรมเวลาที่เกิดขึ้นจริง

ค่าพยากรณ์ หมายถึง ค่าผลการพยากรณ์จากวิธีของเบย์ วิธีบ็อกซ์-เจนกิ้นส์ และวิธีการพยากรณ์ร่วม

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อเป็นแนวทางในการเลือกใช้เทคนิคในการพยากรณ์ที่เหมาะสมกับชุดอนุกรมเวลาทางการศึกษาที่มีและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล

2. เพื่อเป็นแนวทางในการวางแผนการจัดสรรทรัพยากรทางการศึกษาของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ได้อย่างเหมาะสม และเป็นแนวทางในการจัดการศึกษาให้เหมาะสมกับจำนวนนักเรียนทั้งระดับประถมศึกษา และระดับมัธยมศึกษา ของกระทรวงศึกษาธิการ