การอธิบายและพยากรณ์จำนวนผู้ป่วยโรค วัณโรค คอตีบ ไอกรน บาดทะยัก และโปลิโอ ด้วยตัวแบบสถิติ

นางสาวกิตติยา เลียงเครือ
นางสาวอรุณี ผลมณี
นายอุดม ธรรมอมฤตกุล

ปัญหาพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต
ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2536

Describing And Forecasting Number of Tuberculosis, Diphtheria, Pertussis, Tetanus and Poliomyelitis patients by Suitable Statistical Model

Miss Kittiya Liangkhrua

Miss Arunee Ponmanee

Mr. Udom Dhamamarutkul

A Special Problem Submitted in Partial Fulfillment of the

Requirement for the Degree of Bachelor of Science

Department of the Applied Statistics

Faculty of Science

King Mongkut's Institute of Technology Ladkrabang

1993

หัวข้อปัญหาพิเศษ	การอธิบายและพยากรณ์จำนวนผู้ป่วยโรค วัณโรค คอตีบ ไอกรน บาดทะยัก และโปลิโอ ด้วยตัวแบบสถิติ
โดย	นางสาวกิตติยา เลียงเครือ
	นางสาวอรุณี ผลมณี
	นายอุดม ธรรมอมฤตกุล
ภาควิชา	สถิติประยุกต์
อาจารย์ที่ปรึกษา	ผศ.หัทยา เชี่ยววัฒกี

ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
อนุมัติให้โครงงานพิเศษฉบับนี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรวิทยาศาสตรบัณฑิต

______________________ หัวหน้าภาควิชา

(อาจารย์วีรศักดิ์ สุรพัฒน์)

คณะกรรมการโครงงานพิเศษ

หัทยา เชี่ยววัฒกี ______________________ ประธานกรรมการ

(ผศ.หัทยา เชี่ยววัฒกี)

______________________ กรรมการ

(อาจารย์จุฑาธิป ตัณสถิตย์)

______________________ กรรมการ

(อาจารย์นวลสวาท หิรัญสกลวงศ์)

หัวข้อปัญหาพิเศษ : การอธิบายและพยากรณ์จำนวนผู้ป่วยโรค วัณโรค คอตีบ ไอกรน บาดทะยัก และโปลิโอ ด้วยตัวแบบสถิติ

นักศึกษา นางสาวกิตติยา เลียงเครือ

นางสาวอรุณี ผลมณี

นายอุดม ธรรมอมฤตกุล

อาจารย์ที่ปรึกษา ผศ.หัทยา เชี่ยววัฒก์

ภาควิชา สถิติประยุกต์

ปีการศึกษา 2536

บทคัดย่อ

วัตถุประสงค์ของปัญหาพิเศษนี้คือ การหาตัวแบบทางสถิติที่เหมาะสมสำหรับการพยากรณ์จำนวนผู้ป่วยด้วยโรค วัณโรค คอตีบ ไอกรน บาดทะยัก และโปลิโอ ทั้งประเทศ ทุกภาค และในกรุงเทพมหานครโดยได้ข้อมูลจากกองระบาดวิทยา ตั้งแต่เดือนมกราคม พ.ศ.2530 ถึง เดือนธันวาคม พ.ศ.2534 จะทำการหาตัวแบบทางสถิติที่เหมาะสมสำหรับแต่ละโรค โดยทำการแยกข้อมูลออกเป็นข้อมูลที่ไม่มีความผันแปรฤดูกาลกับข้อมูลที่มีความผันแปรฤดูกาล สำหรับข้อมูลที่ไม่มีความผันแปรฤดูกาล จะวิเคราะห์อนุกรมเวลาด้วยเทคนิคดังต่อไปนี้ วิธีทำให้เรียบแบบใช้ค่าเฉลี่ยเคลื่อนที่ การปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียว การปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้ง ด้วยวิธีของบราวน์และโฮลท์ การปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้ง และเทคนิคการวิเคราะห์อนุกรมเวลาแบบบอกซ์และเจนกินส์ หากข้อมูลมีความผันแปรฤดูกาล จะวิเคราะห์อนุกรมเวลาด้วยการปรับแนวโน้ม และฤดูกาลของวินเตอร์ และเทคนิคการวิเคราะห์อนุกรมเวลาแบบบอกซ์และเจนกินส์

ผลการวิเคราะห์สรุปได้ว่า การพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในภาคตะวันออกเฉียงเหนือ โรคคอตีบในภาคเหนือ โรคบาดทะยักในภาคตะวันออกเฉียงเหนือ และภาคใต้ เหมาะสมที่จะใช้วิธีการปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียว

การพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในภาคใต้ เหมาะสมที่จะใช้วิธีการปรับแบบเอกซ์โปเนนเชียลซ้ำสองครั้งด้วยวิธีของบราวน์

การพยากรณ์จำนวนผู้ป่วยโรคคอตีบในภาคตะวันออกเฉียงเหนือ เหมาะสมที่จะใช้วิธีการปรับแนวโน้มและฤดูกาลของวินเตอร์

การพยากรณ์จำนวนผู้ป่วยโรคไอกรนทั้งประเทศ โรคไอกรนทุกภาค และ ในกรุงเทพมหานคร โรคโปลิโอทั้งประเทศ ในกรุงเทพมหานคร ภาคกลาง ภาคเหนือ โรคบาดทะยักในกรุงเทพมหานคร ในภาคกลาง ภาคเหนือ โรควัณโรคทั้งประเทศ ในกรุงเทพมหานคร และ โรควัณโรคในทุกภาค เหมาะสมที่จะใช้เทคนิคการวิเคราะห์อนุกรมเวลาแบบบอกซ์และเจนกินส์

Special Problem Title : Describing And Forcasting Number of Tuberculosis , Diphtheria , Pertussis, Tetanus and Poliomyelitis Patients By Suitable statistical model .

Name Miss Kittiya Liangkhura

Miss Arunee Ponmanee

Mr. Udom Dhamamrutkul

Special Problem Advisor Assistant Professor Hattaya Chiowattakee

Department Applied Statistics

Academic Year 1993

ABSTRACT

The purpose of the study is to find the proper statistical models for describing and forecasting the number of patients with Tuberculosis, Diphtheria, Pertussis, Tetanus and Poliomyelitis in the whole kingdom, in every regions and in Bangkok. The data was provided from the Division of Epidemiology Ministry of Public Health from January,1987 to December,1991 . The original data was grouped into nonseasonal variation and seasonal variation data. For the nonseasonal variation data we used Moving average, Single Exponential, Brown's Double Exponential, Holt's Double Exponential, Triple Exponential and Box-Jenkins method to analyse.

For the seasonal variation data. We used Winter's Linear and Seasonal Exponential Smoothing and Box-Jenkins method to analyse.

We found that the proper technique for forecasting the number of poliomyelitis patients in North East, the number of Diphtheria patients in North, the number of Tetanus patients in North East and South is Single Exponential.

The proper technique for forcasting the number of over all Tetanus patients is Brown's Double Exponential.

The proper technique for forcasting the number of Diphtheria patients in North East is Winter's Linear and Seasonal Exponential Smoothing.

The proper technique for forcasting the number of others is Box-Jenkins method .

กิตติกรรมประกาศ

ขอกราบขอบพระคุณ ผศ.หัทยา เชี่ยววัฒก์ อาจารย์ที่ปรึกษาปัญหาพิเศษซึ่งให้ความช่วยเหลือ สละเวลาอันมีค่าให้คำแนะนำและคำปรึกษาต่าง ๆ ตลอดจนหนังสืออ้างอิงต่าง ๆ มากมาย ที่ท่านอาจารย์ได้กรุณาแนะนำและมอบให้ค้นคว้าหาความรู้เพิ่มเติมประกอบในการทำปัญหาพิเศษนี้ จนกระทั่งปัญหาพิเศษนี้ลุล่วงไปได้ดี และขอกราบขอบคุณอาจารย์จุฑาธิป ตัณสถิตย์ อาจารย์นวลสวาท หิรัญสกลวงศ์ และคณาจารย์ทุกท่านที่อบรมสั่งสอน ตลอดจนให้คำแนะนำต่าง ๆ มาโดยตลอด

ขอขอบคุณกองระบาดวิทยา กระทรวงสาธารณสุข ที่กรุณามอบข้อมูลจำนวนผู้ป่วยโรคต่าง ๆ ตั้งแต่ พ.ศ.2530 ถึง พ.ศ.2534 รวมทั้งเจ้าหน้าที่ธุรการประจำภาควิชาสถิติประยุกต์ ที่ให้บริการและเป็นธุระช่วยเหลือในเรื่องต่างๆ

และขอบคุณเพื่อนๆ ที่คอยให้กำลังใจ และช่วยเหลือในด้านต่างๆมาด้วยดีโดยตลอด

นางสาวกิตติยา เลียงเครือ
นางสาวอรุณี ผลมณี
นายอุดม ธรรมอมฤตกุล

สารบัญ

หน้า

สารบัญรูป

รูปที่ หน้า

รูปที่ หน้า

รูปที่ หน้า

รูปที่ หน้า

สารบัญตาราง

ตารางที่		หน้า

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในสมัยที่ระบบต่างๆทางการแพทย์ยังไม่มีการพัฒนาเท่าที่ควร คนในสมัยก่อนจึงมักเจ็บป่วยด้วยโรคติดเชื้อต่างๆกันมาก และโรคติดเชื้อต่างๆเหล่านี้ก็เป็นสาเหตุที่สำคัญอย่างหนึ่งที่ทำให้ทรัพยากรมนุษย์ของเราต้องเกิดการทุพพลภาพและถึงแก่ชีวิต แต่ในปัจจุบันเทคโนโลยีทางการแพทย์ได้เจริญรุดหน้าไปอย่างรวดเร็ว จึงมีผู้คิดค้นวัคซีนเพื่อใช้ในการป้องกันและควบคุมการติดต่อของโรคติดเชื้อบางอย่างได้แล้ว แต่นั่นก็ไม่ได้หมายความว่าจะไม่มีการเจ็บป่วยของโรคเกิดขึ้น ดังเช่นโรคคอตีบ บาดทะยัก ไอกรน วัณโรค และโปลิโอ ซึ่งมีโครงการรณรงค์ให้ภูมิคุ้มกัน แต่ก็ยังมีผู้ป่วยปรากฎอยู่

โรคคอตีบ (Diphtheria) ยังเป็นปัญหาสำคัญสำหรับประเทศที่กำลังพัฒนาและยังไม่สามารถให้ภูมิคุ้มกันโรคแก่ประชากรได้อย่างทั่วถึง สำหรับประเทศไทย จากสถิติโรคติดต่อของกระทรวงสาธารณสุข พบว่าในปีพ.ศ.2519 มีผู้ป่วยด้วยโรคคอตีบทั้งประเทศ 2,345 ราย และตาย 117 ราย จากจำนวนผู้ป่วยนี้เป็นผู้ป่วยในเขตกรุงเทพมหานครถึง 209 ราย ปีพ.ศ.2534 กองระบาดวิทยาได้รับรายงานผู้ป่วยไข้คอตีบทั่วประเทศจำนวน 53 ราย จาก 32 จังหวัด มีผู้ป่วยเสียชีวิต 11 ราย หากจะคิดเป็นอัตราป่วยตาย ซึ่งเป็นสัดส่วนระหว่างจำนวนผู้ป่วยซึ่งเสียชีวิตต่อจำนวนผู้ป่วยทั้งหมด อัตราป่วยตายจากโรคคอตีบจะเท่ากับร้อยละ 20.8 ซึ่งนับเป็นอัตราป่วยตายที่สูงที่สุดในรอบสิบปี จากเดิมซึ่งมีอัตราป่วยตายอยู่ระหว่างร้อยละ 4.1 ถึง 15.0 โรคคอตีบจะเกิดจากการได้รับเชื้อแบคทีเรียชนิดหนึ่งจากพาหะนำโรค ซึ่งอาจจะเป็นผู้ใหญ่หรือเด็กที่เคยได้รับวัคซีนป้องกันโรคคอตีบครบแล้ว ในกรณีนี้ผู้ที่เป็นพาหะนำโรคจะไม่แสดงอาการเจ็บป่วย นอกจากนี้อาจได้รับเชื้อจากผู้ป่วยที่เป็นโรคโดยตรง เช่นการไอหรือจามรดกัน หรืออาจจะมีเชื้อโรคปนเปื้อนมากับนมและอาหาร นอกจากนี้แล้ว ผู้ที่ได้รับวัคซีนครบแล้วก็อาจเป็นโรคคอตีบได้แต่อาการมักไม่

รุนแรง ซึ่งการเกิดโรคนี้ในชุมชนจะเกิดจากมีการให้ภูมิคุ้มกันไม่ทั่วถึง และการอยู่กันอย่างแออัด การป่วยเป็นโรคคอตีบจะส่งผลต่อต่อมทอนซิล ลำคอ กล่องเสียง หรือ ผิวหนัง จะมีอาการเจ็บคอ อาจมีต่อมน้ำเหลืองที่คอโต มีการอุดตันของทางเดินหายใจ ทำให้เป็นแผลเรื้อรังที่ผิวหนัง อาจร้ายแรงถึงกับทำให้เกิดอัมพาตของประสาทสมอง ประสาทควบคุมการเคลื่อนไหวและความรู้สึก และกล้ามเนื้อหัวใจอักเสบ ซึ่งอาการนี้จะพบบ่อยและเป็นสาเหตุการตายมากที่สุด

โรคบาดทะยัก (Tetanus) เป็นโรคที่เกิดจากเชื้อแบคทีเรีย ซึ่งปรากฏอยู่ทั่วไปตามพื้นดิน ซึ่งปะปนอยู่กับอุจจาระของสัตว์เลี้ยงต่างๆ โดยเฉพาะอย่างยิ่งม้า นอกจากนี้ยังอาจพบเชื้อบาดทะยักได้ในน้ำทะเล เชื้อบาดทะยักสามารถทนน้ำเดือด 100 องศาเซลเซียสได้เป็นเวลานานถึง 60 นาที ในสภาพที่ไร้แสงและความร้อนมันสามารถอยู่ได้นานถึง 11 ปี นอกจากนั้นยังทนต่อน้ำยาฆ่าเชื้อโรคได้ดีกว่าแบคทีเรียชนิดอื่นๆอีกด้วย โดยทั่วไปเชื้อบาดทะยักจะเข้าสู่ร่างกายทางบาดแผลซึ่งปนเปื้อนด้วยสิ่งแปลกปลอม เช่น แผลที่เปื้อนดินโคลน แผลถูกยิง แผลถูกของแหลมตำ หรือแผลที่มีเนื้อเยื่อถูกทำลายมาก อย่างไรก็ตามการติดเชื้ออาจเกิดได้ในระหว่างการทำแท้ง การฉีดยา การผ่าตัด หรืออาจเกิดจากฟันผุ หูน้ำหนวกก็ได้ ที่โรงพยาบาลรามาธิบดีพบว่าร้อยละ 50 ของคนไข้ที่เป็นโรคบาดทะยักจะพบว่ามีบาดแผล ร้อยละ 15 มีอาการหูน้ำหนวก และอีกร้อยละ 35 ไม่พบบาดแผลที่แน่นอน

โรคบาดทะยักทำให้มีอาการเจ็บ และเสียวที่บาดแผล ทำให้มีอาการกระวนกระวาย ปวดศีรษะ กล้ามเนื้อที่คาง คอ หดเกร็งและเจ็บปวด ขากรรไกรค้าง ใบหน้ามองดูเหมือนแสยะยิ้ม กลืนอาหารลำบาก ต่อไปจะมีอาการเกร็งและชักกระตุกของกล้ามเนื้อตามร่างกาย โดยเฉพาะอย่างยิ่งเวลาได้รับความกระทบกระเทือน ได้ยินเสียงดังหรือแม้แต่กระทบแสงสว่าง ขณะที่ชักกล้ามเนื้อเกี่ยวกับการหายใจจะหดเกร็ง ทำให้หายใจขัดและขาดอ๊อกซิเจนได้ อาจมีไข้เล็กน้อย นอกจากนี้ ในทารกแรกเกิด บาดทะยักยังเป็นสาเหตุอันดับหนึ่งที่ทำให้ทารกเสียชีวิตในช่วงอายุ 28 วันแรก ทารกเกิดใหม่อาจได้รับเชื้อบาดทะยักเข้าทางสายสะดือที่ตัดเป็นแผลด้วยมีดหรือไม้รวกที่สกปรก หรือรักษาบาดแผลไม่สะอาด ทารกจะกวนมารดาแล้วไม่ยอมดูดนม ขากรรไกรแข็ง ต่อมาจะชักตัวแข็ง โดยมากทารกจะถึงแก่กรรมในระยะนี้ การเจ็บป่วยเนื่องจากโรคบาดทะยักนี้พบได้ทั่วโลก แต่พบ

มากในประเทศเกษตรกรรมและด้อยพัฒนาที่ประชาชนต้องสัมผัสกับมูลสัตว์ และไม่ค่อยได้รับภูมิคุ้มกันโรค นับเป็นสาเหตุการตายที่สำคัญอันหนึ่งของประเทศทางแถบเอเชีย อาฟริกา และอเมริกาใต้ สำหรับประเทศไทยบาดทะยักยังเป็นสาเหตุการตายที่สำคัญของทารก ทั้งในชนบท และในเมือง

โรคไอกรน (Pertussis) เป็นโรคติดเชื้อของระบบทางเดินหายใจ พบมากในเด็กแรกเกิด จนกระทั่งถึงเด็กโต ในเด็กเล็กโดยเฉพาะเด็กแรกเกิดจะมีอันตรายจากการเป็นโรคไอกรนมากที่สุด เนื่องจากไม่มีภูมิคุ้มกันจากมารดาผ่านรก โรคไอกรนเป็นโรคที่พบบ่อยในเด็ก ทั้งในประเทศร้อน และประเทศหนาว สำหรับชุมชนที่ให้ภูมิคุ้มกันแก่ประชาชน และมีการรักษาที่ดี การเจ็บป่วยและตายจะลดลงมาก แต่ในชุมชนที่ยังไม่สู้เจริญนัก การเจ็บป่วยและการตายยังคงมีจำนวนมากในเด็กอายุต่ำกว่า 1 ปี

การป้องกันโรคคอตีบ บาดทะยัก ไอกรน นี้ทำได้โดยการฉีดวัคซีน DTP (Diphtheria,Tetanus Toxoid and Pertussis Vaccine) โดยเริ่มฉีดในเด็กอายุ 2-3 เดือน ฉีด 3 ครั้ง ห่างกันครั้งละ 2 เดือน แล้วฉีดซ้ำกระตุ้นอีกหลังครั้งสุดท้ายประมาณ 1-1 ปีครึ่ง กับเวลาที่เด็กเข้าโรงเรียน เวลาที่มีการระบาดของโรค และทุกครั้งเมื่อเป็นแผลถูกของตำหรือถูกสัตว์กัด

จากงานวิจัยเรื่อง การศึกษาการตอบสนองของแอนติบอดีต่อโรคไอกรนในเด็กที่ได้รับวัคซีนป้องกันโรคคอตีบ บาดทะยัก และไอกรนครบ 3 ครั้ง ของพิพัฒน์ ลักษมีจรัลกุล และ เสาวลักษณ์ ลักษมีจรัลกุล (2528) พบว่า จากโครงการสร้างเสริมภูมิคุ้มกันโรคโดยการให้วัคซีนป้องกันโรคคอตีบ บาดทะยัก และ ไอกรน ซึ่งเริ่มตั้งแต่ พ.ศ.2520 แต่ปัจจุบันก็ยังมีการเจ็บป่วยเนื่องจากโรคนี้ทุกปี

สำหรับ**โรควัณโรค** (Tuberculosis) เป็นโรคติดต่อเรื้อรังที่เป็นสาเหตุการตายในอันดับต้นๆสำหรับประเทศไทย ทั้งนี้เพราะสภาพภาวะเศรษฐกิจ สังคม และการศึกษาของประชาชนโดยทั่วไปเป็นสิ่งสำคัญที่ทำให้การกำจัดโรคนี้เป็นไปได้ยาก เนื่องจากวัณโรคเป็นโรคที่ทำให้เกิดการเจ็บป่วยได้กับทุกระบบ และทุกอวัยวะของร่างกายประการหนึ่ง และเป็นโรคที่พบบ่อยอีกประการหนึ่ง จึงเป็นโรคที่ควรคำนึงถึง อาการของโรคที่ตรวจพบได้แก่ ไอเรื้อรัง ไอเป็นเลือด หอบเหนื่อย ต่อมน้ำเหลืองโต ที่พบบ่อยคือต่อมน้ำเหลืองบริเวณคอ ต่อมน้ำเหลืองอักเสบเรื้อรัง ตับและม้ามอาจโต หงุดหงิด อ่อนเพลีย

ไม่มีแรง ปวดศีรษะ คลื่นไส้ อาเจียน ชัก หมดสติปวดข้อ ข้ออักเสบ หลังโกง สันหลังคด กล้ามเนื้อหดเกร็ง ท่าเดินผิดแปลกไป ปวดท้องเป็น ๆ หาย ๆ เรื้อรัง เบื่ออาหาร เติบโตช้ากว่าวัย ท้องอืด ถ่ายปัสสาวะบ่อย ปัสสาวะขัดและมีเลือดปน ซึ่งจะมีอาการเจ็บปวด อาจทำให้เยื่อหุ้มสมองอักเสบ ทำให้เกิดอัมพาตของประสาทสมอง การติดโรคมักเกิดจากการหายใจเอาละอองเสมหะ น้ำมูก น้ำลายของผู้ป่วยเข้าสู่ปอด หรือจากแมลงนำเชื้อโรค การใช้ของใช้ร่วมกับผู้ป่วย การหายใจเอาฝุ่นละอองที่มีเชื้อวัณโรคเข้าสู่ปอด การติดต่อกับผู้ป่วยเสมอ ๆ ทำให้เกิดโรคได้ง่าย การรับประทานเนื้อสัตว์หรือผลิตภัณฑ์จากสัตว์ที่เป็นโรค โดยไม่ได้รับการฆ่าเชื้อโรคเสียก่อน

โรคโปลิโอ (Poliomyelitis) จะทำให้ผู้ป่วยพิการ แขนขาลีบ กระดูกแขนขาลีบเล็กผิดรูปร่าง กระดูกข้อต่อปูดโปนบิดเบี้ยว

สาเหตุที่ทำให้ผู้จัดทำปัญหาพิเศษสนใจศึกษา โรคคอตีบ บาดทะยัก และไอกรน ก็เนื่องจากอันตรายของโรค และการให้วัคซีนป้องกันโรคร่วมกัน (Vaccine DTP) แต่ก็ยังมีการเจ็บป่วยอยู่เป็นจำนวนมาก สำหรับสาเหตุที่ทำให้สนใจศึกษาโรควัณโรค ก็คือ วัณโรคเป็นโรคติดต่อเรื้อรังที่เป็นสาเหตุการตายในอันดับต้นๆ สำหรับประเทศไทย และจากการศึกษาระบาดวิทยาของโรคโปลิโอในเขตกรุงเทพมหานคร พ.ศ.2526 ถึง 2527 ของอะเคื้อ อุณหเลขกะ และคณะ พบว่า อุบัติการณ์ของจำนวนผู้ป่วยโรคโปลิโอ ในกรุงเทพมหานครลดลง แต่จากรายงานของกองระบาดวิทยาประจำปี พ.ศ.2534 พบว่า ผู้ป่วยที่เป็นโรคโปลิโอในประเทศไทยมีจำนวนเพิ่มขึ้นจึงทำให้สนใจที่จะศึกษาโรคโปลิโอนี้ด้วย

1.2 วัตถุประสงค์ของการศึกษา

1.หาตัวแบบทางสถิติที่เหมาะสมสำหรับอธิบายและพยากรณ์จำนวนผู้ป่วยทั้งประเทศที่ป่วยเป็นโรค วัณโรค คอตีบ ไอกรน บาดทะยัก และโปลิโอ

2.หาตัวแบบทางสถิติที่เหมาะสมสำหรับอธิบาย และพยากรณ์จำนวนผู้ป่วยด้วยโรควัณโรค คอตีบ ไอกรน บาดทะยัก และโปลิโอ โดยแบ่งแยกออกเป็นภาคตามที่กองระบาดวิทยาได้แบ่งไว้ คือ ภาคกลาง ภาคเหนือ ภาคตะวันออกเฉียงเหนือ ภาคใต้ และกรุงเทพมหานคร

1.3 ขอบเขตของการศึกษา

ศึกษาจำนวนผู้ป่วยโรค

-วัณโรค (Tuberculosis)

-คอตีบ (Diphtheria)

-ไอกรน (Pertussis)

-บาดทะยัก (Tetanus)

-โปลิโอ (Poliomyelitis)

จากข้อมูลรายเดือน ระยะเวลา 5 ปี เริ่มตั้งแต่ เดือนมกราคม พ.ศ.2530 ถึง เดือนธันวาคม พ.ศ.2534 รวม 60 เดือน เพื่อหาตัวแบบทางสถิติที่เหมาะสมสำหรับการพยากรณ์จำนวนผู้ป่วยโรค วัณโรค คอตีบ ไอกรน บาดทะยัก และ โปลิโอ

1.4 แหล่งข้อมูล

สรุปรายงานการเฝ้าระวังโรคประจำปี พ.ศ.2530 - พ.ศ.2534

กองระบาดวิทยา สำนักงานปลัดกระทรวง

กระทรวงสาธารณสุข

กรุงเทพมหานคร 10200

โทร.2821887,2825824

1.5 วิธีการดำเนินงาน

1.ศึกษาลักษณะของข้อมูล เพื่อดูลักษณะการเคลื่อนไหวของข้อมูลทั้ง 5 โรค ว่ามีแนวโน้มเป็นอย่างไร มีความผันแปรเป็นฤดูกาลหรือไม่

2.ทำการวิเคราะห์ข้อมูลเพื่อหาตัวแบบที่เหมาะสมในการพยากรณ์จำนวนผู้ป่วยโดยใช้ วิธีการทำให้เรียบ (Smoothing method) ได้แก่ การเฉลี่ยเคลื่อนที่ การทำให้เรียบแบบเอกซ์โปเนนเชียลครั้งเดียว สองครั้ง สามครั้ง วิธีการปรับแนวโน้มและฤดูกาลของวินเตอร์ และเทคนิคการวิเคราะห์อนุกรมเวลาแบบบอกซ์และเจนกินส์

3.วิเคราะห์ผลด้วยโปรแกรมสำเร็จรูปทางสถิติ FORECAST PLUS และศึกษาการใช้โปรแกรม STATGRAPHIC, MINITAB และ TSP โดยใช้เครื่อง Microcomputer

4.เปรียบเทียบเทคนิคต่างๆที่ใช้ในการพยากรณ์ เพื่อสรุปหาตัวแบบที่เหมาะสมที่สุด สำหรับอธิบายและพยากรณ์จำนวนผู้ป่วยโรคต่างๆ

1.6 ประโยชน์ที่ได้รับจากการศึกษานี้

1.ได้นำวิชาสถิติที่ศึกษามาประยุกต์ใช้ในการทำงานจริงๆ ซึ่งจะเป็นประโยชน์ในการทำงานต่อไป

2.ผลการวิเคราะห์ที่ได้สามารถใช้พยากรณ์จำนวนผู้ป่วย โรควัณโรค คอตีบ ไอกรน บาดทะยัก และ โปลิโอ ในอนาคตได้

3.ได้ศึกษาการใช้โปรแกรมสำเร็จรูปทางสถิติ FORECAST PLUS,STATGRAPHIC MINITAB และ TSP

บทที่ 2
ทฤษฎีการวิเคราะห์ข้อมูล

2.1 ทฤษฎีการวิเคราะห์ข้อมูล

การวิเคราะห์อนุกรมเวลามีวัตถุประสงค์เพื่อให้ได้ตัวแบบจำลอง (Model) ที่สามารถอธิบายลักษณะของอนุกรมเวลา โดยมีข้อสมมติว่า รูปแบบของข้อมูลในอดีตจะดำเนินต่อไปในอนาคต ซึ่งเป็นข้อกำหนด (Assumption) ที่เกี่ยวกับความต่อเนื่อง (Continue) ซึ่งอาจกล่าวได้ว่า เทคนิคการพยากรณ์ไม่สามารถให้ผลพยากรณ์ที่ดีได้ หากข้อกำหนดไม่เป็นจริง

การเลือกเทคนิคการพยากรณ์สำหรับอนุกรมเวลาแต่ละชุดนั้น ต้องพิจารณาองค์ประกอบ (Factor) ดังต่อไปนี้

- รูปแบบการพยากรณ์ที่ต้องการ (The forecast form desired)
- กรอบเวลา (The time frame)
- รูปแบบของข้อมูล (The pattern of data)
- ความแม่นยำที่ต้องการ (The accuracy desired)
- ข้อมูลที่หาได้ (The availability of data)
- ความเข้าใจในวิธีต่างๆ ที่ทำการวิเคราะห์ (The ease of operation and understanding)
- ค่าใช้จ่าย (The cost of forecasting)

วิธีการตรวจสอบเพื่อเลือกตัวแบบในการพยากรณ์

1. พิจารณาจากค่า Box-Pierce Chi-Square Statistics

เมื่อทำการวิเคราะห์ข้อมูลด้วยวิธีการใดวิธีการหนึ่งแล้ว ในการพิจารณาว่าตัวแบบจากวิธีการนั้น ๆ สามารถใช้พยากรณ์ข้อมูลได้หรือไม่ จะตรวจสอบตัวแบบโดยวิธีวิเคราะห์เศษตกค้าง(Analysis of residual) คือพิจารณาความแตกต่างระหว่างข้อมูลที่สังเกตได้กับค่าพยากรณ์ที่กำหนดด้วยตัวแบบที่เลือกแล้ว

วิธีตรวจสอบที่มีประสิทธิภาพที่สุดเป็นวิธีตรวจสอบที่ เรียกว่า สถิติบอกซ์-เพียร์ซ ไคสแควร์ (Box-Pierce Chi-Square statistics) ซึ่งจะเป็นตัวที่จะบอกว่า สหสัมพันธ์ในตัวเองของเศษตกค้าง k ค่าแรกนั้นเพียงพอ สำหรับตัวแบบหรือยัง ค่า X^2 คำนวณจากสูตร ดังนี้

$$X^2 = (n-d) \sum_{L=1}^{k} r_L^2(e)$$

โดย n คือ จำนวนค่าสังเกตในอนุกรมเวลาเดิม

d คือ องศาของความแตกต่าง (degree of differencing) ที่ใช้ในการแปลงอนุกรมเวลาเดิมให้เป็นอนุกรมเวลาที่อยู่ในสภาวะสมดุลย์

$r_L^2(e)$ คือ กำลังสองของ $r_L(e)$ ซึ่งคือสหสัมพันธ์ในตัวเองของตัวอย่างของเศษตกค้างที่ระยะห่าง L หน่วยเวลา

k คือ จำนวนสหสัมพันธ์ในตัวเองของเศษตกค้างที่ใช้ในการคำนวณ X^2 มีค่าเท่ากับเท่าใดก็ได้ แต่มักใช้ k = 12 (อาจเป็น 24 ,36 ก็ได้)

ถ้า $X^2 < X^2_{0.05}$ ที่องศาแห่งความเป็นอิสระ (degree of freedom) k-np แล้ว โดย np เป็นจำนวนพารามิเตอร์ที่ต้องประมาณค่าในตัวแบบ แสดงว่า ตัวแบบเหมาะสม สามารถใช้พยากรณ์ข้อมูลได้

หากมีตัวแบบที่สามารถใช้พยากรณ์ข้อมูลได้หลายตัวแบบจะต้องพิจารณาจากค่าความคลาดเคลื่อน MSE และ MAPE ซึ่งอธิบายในหัวข้อถัดไป

2. พิจารณาจากค่าความคลาดเคลื่อน ต่อไปนี้

1.1 Mean Square Error (MSE)

$$MSE = \sum_{t=1}^{m} (e_{n+t})^2 / m$$

1.2 Mean Absolute Percent Error (MAPE)

$$MAPE = (100/m)*\{\sum_{t=1}^{m} \left| e_{n+t}/Z_{n+t} \right|\}$$

เมื่อ e_{n+t} คือ ค่าความคลาดเคลื่อนจากการพยากรณ์ค่าสังเกตในอนาคต

Z_{n+t} คือ ค่าสังเกตจริง

m คือ จำนวนเทอมของค่าความคลาดเคลื่อนที่ใช้คำนวณ

มีวิธีพิจารณาดังนี้

- เลือกวิธีการพยากรณ์ที่มีค่า MSE และ MAPE น้อยที่สุด

- ถ้าค่า MSE และ MAPE ของวิธีพยากรณ์แต่ละวิธีมีค่าใกล้เคียงกันให้เลือกวิธีพยากรณ์ที่มีการประมาณค่าพารามิเตอร์จำนวนน้อยกว่า

- ถ้าวิธีพยากรณ์ของวิธีหนึ่งมีค่า MSE สูง ค่า MAPE ต่ำ ส่วนอีกวิธีหนึ่งมีค่า MSE ต่ำ ค่า MAPE สูง ให้พิจารณาค่า MSE เป็นหลักโดยเลือกวิธีที่มีค่า MSE ต่ำ

การตรวจสอบการผันแปรตามฤดูกาลของข้อมูลอนุกรมเวลา

ในที่นี้ใช้ 2 วิธี คือ

วิธีที่ 1 การเปรียบเทียบค่าข้อมูลกับค่าเฉลี่ย

วิธีนี้เป็นการเปรียบเทียบ ค่าข้อมูลในแต่ละเดือนกับค่าเฉลี่ยในแต่ละปีนั้น ๆ จะเห็นได้ชัดเจนเมื่อพิจารณากราฟ โดยพิจารณาจากค่าสังเกตว่ามีการขึ้นลงรอบค่าเฉลี่ยในแต่ละปีเป็นลักษณะเดียวกันหรือไม่ ถ้าเป็นลักษณะเดียวกัน จะสรุปว่าข้อมูลนั้นมีการผันแปรตามฤดูกาล

วิธีที่ 2 การใช้ค่าสหสัมพันธ์ในตัวเอง (Autocorrelation)

นำค่าสังเกตมาหาค่าสหสัมพันธ์ในตัวเอง เขียนกราฟค่าสหสัมพันธ์ในตัวเอง พิจารณาค่าสหสัมพันธ์ที่มีค่าสูง ว่ามีการเกิดห่างกันเป็นจำนวนเท่าของคาบเวลาที่ล่ากว่ากันจากจุด

แรกหรือไม่ เช่น ถ้าค่าสหสัมพันธ์ในตัวเองของข้อมูลที่ล่ากว่ากัน 12 คาบเวลามีค่าสูง และค่าสังเกตทุกหน่วยเวลาที่ล่ากว่ากันเป็นจำนวนเท่าของ 12 ก็มีค่าสูง แสดงว่าข้อมูลชุดนั้นมีความผันแปรแบบฤดูกาล

เทคนิคในการวิเคราะห์อนุกรมเวลา

การเฉลี่ยเคลื่อนที่

การเฉลี่ยเคลื่อนที่(Moving average) เป็นการเฉลี่ยน้ำหนักข้อมูล หรือ ค่าสังเกตแบบเลขคณิต โดยให้น้ำหนักแก่แต่ละข้อมูลเท่ากันหลังจากที่ได้กำหนดจำนวนเทอมที่จะเฉลี่ยแล้ว และจะตัดข้อมูลที่เก่าที่สุดออกไป เมื่อมีข้อมูลใหม่เกิดขึ้น

การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีการเฉลี่ยเคลื่อนที่ครั้งเดียว (Single Moving Average)

เป็นกระบวนการปรับให้เรียบ ที่ใช้พยากรณ์ล่วงหน้าในระยะสั้นๆ ตัวแบบของการเฉลี่ยเคลื่อนที่ครั้งเดียวเป็นตัวแบบคงที่ ซึ่งการวิเคราะห์อนุกรมเวลาด้วยวิธีการเฉลี่ยเคลื่อนที่ครั้งเดียวนั้น การเคลื่อนไหวของข้อมูลควรจะไม่มีแนวโน้ม โดยกระจายรอบค่าคงที่ค่าหนึ่ง มีข้อเสียคือต้องมีข้อมูลเก็บไว้เป็นจำนวนมากกว่าวิธีการเฉลี่ยอย่างง่าย และวิธีการเฉลี่ยเคลื่อนที่ครั้งเดียวไม่สามารถจัดการกับแนวโน้ม หรือ การเปลี่ยนแปลงฤดูกาลได้ ซึ่งมีสิ่งที่ต้องคำนึงถึงคือ ถ้ากระบวนการของค่าสังเกตค่อนข้างจะคงที่ (stable) ก็ควรใช้ N มาก แต่ถ้ากระบวนการของค่าสังเกตเปลี่ยนไปเร็ว ควรใช้ N น้อย เพื่อจะทราบการเปลี่ยนแปลงได้อย่างรวดเร็ว

จากสมการ $S_t = \dfrac{Y_{t-1}+Y_t+Y_{t+1}}{N}$

ซึ่งกำหนดโดย $F_t = S_{t-1}$

หลักในการกำหนดค่า N อีกประการหนึ่ง โดยการพิจารณาค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนในการพยากรณ์ แทนด้วย E โดย

$$E = \sum_{i=1}^{t} e_t^2 \; / \text{ จำนวนเทอมของค่าคลาดเคลื่อน}$$

และเปรียบเทียบ E ที่ได้จาก N ขนาดต่างๆกัน โดยค่า N ใดให้ค่า E ต่ำสุด เป็นค่า N ที่ดีที่สุดในการเฉลี่ยเคลื่อนที่

การวิเคราะห์อนุกรมเวลาด้วยวิธีการเฉลี่ยเคลื่อนที่ซ้ำ 2 ครั้ง

(Double Moving Average)

ในการวิเคราะห์อนุกรมเวลาด้วยวิธีการเฉลี่ยเคลื่อนที่ครั้งเดียวสำหรับข้อมูลที่มีแนวโน้มนั้น อาจทำให้เกิดค่าความคลาดเคลื่อน จึงจะต้องทำการขจัดค่าความคลาดเคลื่อนจากการพยากรณ์ให้หมดไป ด้วยวิธีการเฉลี่ยเคลื่อนที่ซ้ำ 2 ครั้ง ซึ่งก็คือการเฉลี่ยเคลื่อนที่ครั้งแรกด้วยจำนวน M เทอม และเฉลี่ยเคลื่อนที่ครั้งเดียวซ้ำด้วยจำนวน N เทอม โดยตัวแบบเชิงเส้นนิยามโดย

$$X_t = \beta_0 + \beta_1 t + e_t$$

ทำการคำนวณค่าต่างๆ จากสูตรต่อไปนี้

$$M_t^{(2)} = \frac{M_t + M_{t-1} + \ldots + M_{t-N+1}}{N}$$

ค่าที่คำนวณจากการเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง จะคลาดเคลื่อนจากค่าที่คำนวณจากการเฉลี่ยเคลื่อนที่ครั้งเดียว เท่ากับ ค่าที่คำนวณด้วยวิธีเฉลี่ยเคลื่อนที่ครั้งเดียวคลาดเคลื่อนจากค่าสังเกตที่เกิดขึ้นจริง

เมื่อให้ a_t และ b_t เป็นค่าประมาณของ β_0 และ β_1 ตามลำดับ ดังนั้น ค่าพยากรณ์ของข้อมูล ณ คาบเวลา t+T พยากรณ์ ณ เวลา t คือ

$$F_{t+T} = a_t + b_t T$$

โดย a_t เป็นฐานสำหรับปรับจุดเริ่มต้นของค่าพยากรณ์ เพื่อลดความคลาดเคลื่อนที่จุดเริ่มต้นของค่าพยากรณ์ เป็นค่าที่ได้จากการเพิ่มค่าที่คำนวณได้จากการเฉลี่ยเคลื่อนที่ครั้งเดียวเข้าไปในผลต่าง ระหว่างค่าเฉลี่ยเคลื่อนที่ครั้งเดียวและสองครั้ง ทำให้ค่าพยากรณ์ที่ได้ไม่ตามหลังหรือล่า (lag) กว่าข้อมูลจริง

$$a_t = 2M_t - M_t^{(2)}$$

$$b_t = \frac{2}{N-1}(M_t - M_t^{(2)})$$

b_t คือ ความชันของเส้นตรง

การวิเคราะห์อนุกรมเวลาด้วยวิธีการปรับเรียบแบบ เอกซโปเนนเชียลครั้งเดียว (Single Exponential Smoothing)

ลักษณะของข้อมูลที่เหมาะสมกับการวิเคราะห์ด้วยวิธีนี้ ควรมีการกระจายรอบค่าคงที่ค่าหนึ่ง คือ ตัวแบบอนุกรมเวลาเป็นแบบคงที่ ไม่มีแนวโน้ม เช่นเดียวกับตัวแบบอนุกรมเวลาพยากรณ์ด้วยวิธีเฉลี่ยเคลื่อนที่ครั้งเดียว แต่จะขจัดข้อจำกัดที่มีอยู่ในวิธีเฉลี่ยเคลื่อนที่ครั้งเดียว ที่ต้องเก็บข้อมูลจำนวนมากพอควรเพื่อใช้ในการหาค่าเฉลี่ย และข้อมูลทุกค่ามีความสำคัญเท่ากันหมด ซึ่งไม่เหมาะสำหรับการพยากรณ์ที่ข้อมูลตัวล่าสุด มีอิทธิพลต่อค่าพยากรณ์มากกว่าข้อมูลที่ถัดไป

นิยามทางคณิตศาสตร์สำหรับการทำให้เรียบแบบเอกซโปเนนเชียลครั้งเดียว คือ

$$S_t = \alpha X_t + (1-\alpha) S_{t-1}$$

$$S_0 = Y_1$$

S_t คือ ค่าของการทำให้เรียบแบบเอกโปเนนเชียลครั้งเดียว ณ คาบเวลา t ของข้อมูล $x_t, x_{t-1}, \ldots$ เรียกค่านี้ว่า ค่าประมาณทำให้เรียบ หรือ สถิติทำให้เรียบ (smoothed estimate statistic)

α คือ ค่าคงที่ปรับเรียบ เป็นค่าที่กำหนดน้ำหนักของการเฉลี่ย โดย $0 < \alpha < 1$

ค่าพยากรณ์ ณ คาบเวลา t+1 คือ

$$F_t = S_{t+1}$$

การวิเคราะห์อนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซโปเนนเชียลซ้ำสองครั้ง (Double Exponential Smoothing)

การปรับเรียบแบบเอกซโปเนนเชียลซ้ำสองครั้ง มีตัวแบบเป็นแบบเดียวกับการเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง คือ

$$x_t = \beta_0 + \beta_1 t + e_t$$

ซึ่งถ้าคำนวณด้วยวิธีปรับเรียบแบบเอกซโปเนนเชียลครั้งเดียว ค่าจะคลาดเคลื่อนจากข้อมูลจริง ความคลาดเคลื่อนนี้ขจัดได้ด้วยการปรับเรียบแบบเอกซโปเนนเชียลซ้ำสองครั้ง แต่ค่าที่คำนวณด้วยวิธีปรับเรียบแบบเอกซโปเนนเชียลซ้ำสองครั้งนี้ยังคงตามหลังค่าที่เกิดขึ้นจริง เนื่องจากข้อมูลมีแนวโน้ม จึงต้องปรับฐานของจุดที่เริ่มพยากรณ์ โดยการนำค่าที่คำนวณจากการปรับเรียบแบบเอกซโปเนนเชียลครั้งเดียวบวกกับผลต่างของค่าที่คำนวณได้จากการปรับเรียบแบบเอกซโปเนนเชียลครั้งเดียว และซ้ำสองครั้ง และต้องนำค่าความชันมาคำนวณในการพยากรณ์ด้วย สรุปเป็นสูตรต่าง ๆ ดังนี้

$$S_t^{(2)} = \alpha S_t + (1-\alpha) S_{t-1}^{(2)}$$

เมื่อ S_t คือ ค่าปรับเรียบแบบเอกซโปเนนเชียลครั้งเดียว ณ คาบเวลา t

$S_t^{(2)}$ คือ ค่าปรับเรียบแบบเอกซโปเนนเชียลซ้ำสองครั้ง ณ คาบเวลา t

α คือ ค่าคงที่ปรับเรียบที่กำหนดน้ำหนักของการเฉลี่ย โดย $0 < \alpha < 1$

เมื่อให้ a_t และ b_t เป็นค่าประมาณของ β_0 และ β_1 ณ คาบเวลา t ตามลำดับ โดย a_0 เป็นจุดที่เริ่มพยากรณ์ และ b_t เป็นความชันของเส้นตรง

$$a_t = 2S_t - S_t^{(2)}$$

$$b_t = \frac{\alpha}{1-\alpha}(S_t - S_t^{(2)})$$

เมื่อ $t=0$ จะได้

$$b_0 = \frac{n\Sigma t x_t - \Sigma xy\Sigma t}{n\Sigma t^2 - (\Sigma t)^2}$$

$$a_0 = \frac{\Sigma x_t}{n} - \frac{b_t \Sigma t}{n}$$

ค่าพยากรณ์ ณ คาบเวลา $t+T$ คือ

$$F_{t+T} = a_t + b_t T$$

การปรับเรียบเอกซโปเนนเชียลซ้ำสองครั้ง วิธีของโฮลท์

(Double Exponential Smoothing : Holt's Two-Parameter Method)

วิธีทำให้เรียบเอกซโปเนนเชียลเชิงเส้นของโฮลท์มีหลักคล้ายกับวิธีของบราวน์ (Brown's) ต่างกันที่วิธีของโฮลท์ไม่ได้ใช้สูตรทำให้เรียบซ้ำสองครั้งโดยตรง แต่จะแยกทำให้ค่าแนวโน้มเรียบ โดยใช้ค่าคงที่ทำให้เรียบ (Smoothing constants) สองตัวที่มีค่าระหว่าง 0 ถึง 1 และใช้สมการสามสมการ คือ

$$S_t = \alpha Y_t + (1-\alpha)(S_{t-1} + b_{t-1}) \qquad \text{........(2.1)}$$

$b_t = \delta(S_t - S_{t-1}) + (1-\delta)b_{t-1}$ ....(2.2)

$F_{t+1} = S_t + b_t$ ....(2.3)

เมื่อ Y_t คือ ค่าสังเกต ณ คาบเวลาที่ t

S_t คือ ค่าปรับเรียบของข้อมูล ณ คาบเวลา t

b_t คือ ค่าปรับเรียบของแนวโน้ม ณ คาบเวลาที่ t (Smoothing of trend)

α คือ ค่าคงที่ในการปรับเรียบ มีค่าระหว่าง 0 ถึง 1

δ คือ ค่าคงที่ในการปรับเรียบของการปรับค่าแนวโน้ม มีค่าระหว่าง 0 ถึง 1

สมการ (2.1) ปรับ S_t โดยตรง สำหรับแนวโน้มของคาบเวลาก่อนคาบนี้ (b_{t-1}) โดยการบวก กับค่าทำให้เรียบหลังสุด S_{t-1}

สมการ (2.2) จะปรับแนวโน้มให้เป็นปัจจุบัน ซึ่งแสดงในรูปความแตกต่างระหว่างค่าทำให้เรียบสองค่าสุดท้าย เพราะว่า ค่าใหม่ควรจะสูงหรือต่ำกว่าค่าที่เกิดก่อนค่าใหม่นี้ เมื่อข้อมูลมีแนวโน้ม และเนื่องจากอาจจะมีความสุ่มเหลืออยู่ จึงต้องทำให้แนวโน้มคาบเวลาสุดท้าย ($S_t - S_{t-1}$) เรียบด้วย δ และบวกค่านี้กับค่าประมาณแนวโน้มคาบเวลา ก่อนหน้านี้คูณด้วย $(1-\delta)$

สมการ (2.3) ใช้สำหรับพยากรณ์ล่วงหน้า 1 คาบเวลา

ขบวนการทำให้เรียบเอกซโปเนนเชียลเชิงเส้นของโฮลท์ ต้องมีตัวประมาณค่า 2 ค่า ค่าหนึ่งเพื่อใช้เป็นค่าประมาณสำหรับค่าทำให้เรียบค่าแรก คือ $S_1 = x_1$ ส่วน b_1 ประมาณค่าได้หลายแบบ คือ ให้ $b_1 = x_2 - x_1$

หรือ $$b_1 = \frac{(x_2 - x_1) + (x_3 - x_2) + (x_4 - x_3)}{3}$$

หรือ กำหนด b_1 ด้วยตา หลังจากลงจุดข้อมูลแล้ว

การวิเคราะห์อนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซโปเนนเชียลซ้ำสามครั้ง

(Triple Exponential Smoothing or Brown Exponential Smoothing)

การปรับเรียบแบบเอกซโปเนนเชียลซ้ำสามครั้งเป็นวิธีพยากรณ์อนุกรมเวลาที่มีแนวโน้มเป็นรูปยกกำลังสอง คือมีตัวแบบอนุกรมเวลานิยามดังนี้

$$x_t = \beta_0 + \beta_1 + \beta_2 t^2 + e_t$$

แต่วิธีปรับเรียบแบบเอกซโปเนนเชียลซ้ำสามครั้ง จะให้ตัวแบบอนุกรมเวลา นิยามด้วย

$$x_t = \beta_0 + \beta_1 + (1/2)\beta_2 t^2 + e_t$$

เพื่อทำให้สมการประมาณค่าอยู่ในรูปที่ง่ายขึ้น

การปรับเรียบแบบเอกซโปเนนเชียลซ้ำสามครั้ง มีสถิติทำให้เรียบสามตัว ซึ่งนิยามโดยสมการ

$$S_t = \alpha X_t + (1-\alpha)S_{t-1}$$

$$S_t^{(2)} = \alpha S_t + (1-\alpha)S_{t-1}^{(2)}$$

$$S_t^{(3)} = \alpha S_t^{(2)} + (1-\alpha)S_{t-1}^{(3)}$$

ณ คาบเวลา ใดๆ ให้ a_t, b_t, c_t เป็นค่าประมาณของ $\beta_0, \beta_1, \beta_2$ ตามลำดับ กำหนดโดยสมการต่อไปนี้

$$a_t = 3S_t - 3S_t^{(2)} + S_t^{(3)}$$

$$b_t = \alpha \, [(6-5\alpha)S_t - (10-8\alpha)S_t^{(2)} + (4-3\alpha)S_t^{(3)}$$

$$c_t = \left[\frac{\alpha}{1-\alpha}\right]^2 (S_t - 2S_t^{(2)} + S_t^{(3)})$$

เมื่อ S_t คือ ค่าปรับเรียบแบบเอกโปเนนเชียลครั้งเดียว ณ คาบเวลา t

$S_t^{(2)}$ คือ ค่าปรับเรียบแบบเอกโปเนนเชียลซ้ำสองครั้ง ณ คาบเวลา t

$S_t^{(3)}$ คือ ค่าปรับเรียบแบบเอกโปเนนเชียลซ้ำสามครั้ง ณ คาบเวลา t

โดยให้ค่าสังเกตค่าแรกเป็นค่า $S_0, S_0^{(2)}$ และ $S_0^{(3)}$

ค่าพยากรณ์ ณ คาบเวลา t+T คือ

$$F_{t+T} = a_t + b_t T + (1/2) c_t T^2$$

การปรับเรียบอนุกรมเวลาแบบฤดูกาลที่มีแนวโน้มแบบเส้นตรง

จากวิธีที่ผ่านมา ได้แก่ วิธีเฉลี่ยเคลื่อนที่,วิธีปรับเรียบแบบต่าง ๆ ใช้กับข้อมูลที่ไม่มีฤดูกาล ซึ่งเป็นข้อมูลที่เกิดแนวโน้มในลักษณะสุ่ม และไม่เกิดซ้ำกันในระยะเวลาเดียวกัน แต่ถ้าข้อมูลมีฤดูกาล ซึ่งเป็นข้อมูลที่เกิดซ้ำ ๆ กัน ภายในระยะเวลาหนึ่ง ๆ เช่น เทศกาลซื้อของขวัญ ยอดขายจะเพิ่มขึ้นในเดือนมกราคมของทุกปี ปริมาณน้ำจะเพิ่มขึ้นในเดือนตุลาคมของทุกปี เป็นต้น ข้อมูลลักษณะนี้ ใช้วิธีต่าง ๆ ที่กล่าวมาแล้วได้ไม่ดีนัก เพราะจะทำให้มีความผิดพลาดเกิดขึ้น ซึ่งสามารถขจัดข้อผิดพลาดนี้ได้โดยใช้วิธีการปรับแนวโน้ม และฤดูกาลแบบวินเตอร์ (Winter's Linear and Seasonal Exponential Smoothing) ซึ่งเหมาะสำหรับพยากรณ์อนุกรมเวลาที่มีแนวโน้มเป็นเส้นตรง ทั้งฤดูกาลแบบบวกและแบบคูณ

ลักษณะการเกิดฤดูกาล จำแนกได้ 2 ลักษณะ ได้แก่

1. ฤดูกาลแบบบวก (Additive Seasonal) ดูจากลักษณะกราฟจะพบว่า ข้อมูลจะมีการเพิ่มขึ้นหรือลดลงในอัตราคงที่ ไม่ว่าแนวโน้มจะเพิ่มขึ้นหรือลดลงก็ตาม
2. ฤดูกาลแบบคูณ (Multiplicative Seasonal)

ข้อมูลมีการเพิ่มขึ้นหรือลดลงแบบเป็นอัตราส่วนกันกับแนวโน้ม ไม่ว่าข้อมูลจะมีแนวโน้มลดลงหรือเพิ่มขึ้น

ตัวแบบอนุกรมเวลาที่มีฤดูกาลแบบคูณ

$$x_t = (\beta_0 + \beta_1 t) * SN_t + e_t$$

ตัวแบบอนุกรมเวลาที่มีฤดูกาลแบบบวก

$$x_t = \beta_0 + \beta_1 t + SN_t + e_t$$

วิธีพยากรณ์ของวินเตอร์เมื่อข้อมูลมีฤดูกาลแบบคูณ (Multiplicative seasonal)

ตัวแบบที่มีแนวโน้มเป็นเส้นตรงและฤดูกาลแบบคูณนิยามดังนี้

$$x_t = T * SN_t$$
$$= (\beta_0 + \beta_1 t) * SN_t + e_t$$

เมื่อ x_t เป็นข้อมูลหรือค่าสังเกต ณ เวลา t

T เป็นแนวโน้ม

SN_t คือ ปัจจัยฤดูกาล (Seasonal factor) เป็นปัจจัยที่ปรับสำหรับฤดูกาล

e_t คือ ความคลาดเคลื่อนสุ่ม (มีค่าเท่ากับ 0 ถ้าข้อมูลจริงยังไม่เกิด)

S_t , b_t และ I_t เป็นค่าประมาณของ β_0 , β_1 และ SN_t ตามลำดับ

วิธีการของวินเตอร์ประกอบด้วยสมการในการปรับเรียบ 3 สมการ ได้แก่

1.สมการปรับเรียบโดยส่วนรวม

$$S_t = \frac{\alpha x_t}{I_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}) \qquad \text{........(1)}$$

2.สมการปรับเรียบแนวโน้ม คือ

$$b_t = \gamma(S_t + S_{t-1}) + (1-\gamma)b_{t-1} \qquad \text{........(2)}$$

3.สมการปรับเรียบฤดูกาล เป็นการหาปัจจัยฤดูกาล คือ

$$I_t = \frac{\beta x_t}{S_t} + (1-\beta)I_{t-L} \qquad \text{........(3)}$$

สมการสำหรับพยากรณ์ ณ เวลา t นิยามดังนี้

$$F_{t+m} = (S_t + b_t m) I_{t-L+m}$$

เมื่อ L คือ ความยาวของฤดูกาล

b_t คือ ความชันของเส้นตรง

I_t คือ ปัจจัยฤดูกาล ณ เวลา t

F_{t+m} คือ ค่าพยากรณ์ ณ เวลา t โดยพยากรณ์ล่วงหน้า m คาบเวลา

α, β, δ คือ ค่าคงที่ปรับเรียบ มีค่าระหว่าง 0 ถึง 1

สมการ (1) คล้ายสมการปรับเรียบ single exponential ต่างกันที่ ได้ปรับความชันของคาบเวลาล่าสุด (b_{t-1}) โดยการนำความชันบวกกับค่าที่พยากรณ์ด้วยวิธีปรับเรียบแบบ single exponential ซึ่งมีผลทำให้ปรับค่าที่พยากรณ์ได้ให้ขึ้นมาอยู่ บนฐานเดียวกับข้อมูลค่าล่าสุด และเป็นค่าที่ไม่ตามหลังข้อมูลที่เกิดขึ้น นอกจากนี้ยังได้ขจัดอิทธิพลของฤดูกาลออกจาก x_t โดยการหารด้วยปัจจัยฤดูกาล (I_{t-L})

สมการ (2) เป็นการปรับค่าแนวโน้ม โดยถือว่า ในกรณีที่ข้อมูลมีแนวโน้มเป็นเส้นตรง ค่าที่พยากรณ์ได้ล่าสุดควรจะอยู่สูงกว่าหรือต่ำกว่าค่าที่มาก่อน และเนื่องจากอาจมีอิทธิพลของสิ่งที่ไม่คาดคิดรวมอยู่ด้วย จึงขจัดด้วยการนำ δ มาคูณ และคูณค่าความชันในช่วงเวลาถัดมาด้วย $(1-\delta)$

สมการ (3) เป็นการหาปัจจัยฤดูกาลที่เวลา t ซึ่งใช้หลักการปรับเรียบแบบ exponential ทั่วไป คือ ให้น้ำหนักความสำคัญต่อค่า x_t ด้วย β และให้น้ำหนักปัจจัยฤดูกาล ตัวล่าสุดด้วย $(1-\beta)$ ค่าปัจจัย x_t/S_t ด้วย β และให้น้ำหนักปัจจัยฤดูกาลตัวล่าสุดด้วย $(1-\beta)$ ค่าปัจจัยฤดูกาลล่วงหน้านี้ คำนวณสำหรับช่วงเวลา t-L เนื่องจาก L เป็นระยะเวลาของความเป็นฤดูกาล

ในการพยากรณ์อนุกรมเวลาสำหรับคาบเวลา t+m จะใช้ค่าประมาณของปัจจัยฤดูกาลที่คำนวณ ณ คาบเวลา t+m-1 คือ I_{t+m-L} ถ้าการพยากรณ์นั้นคำนวณสำหรับคาบ

เวลาที่มากกว่า L ในอนาคต ดัชนี $t+m-L$ จะหมายถึงปัจจัยฤดูกาล ที่ยังไม่ได้คำนวณ ในกรณีเช่นนี้ ควรใช้ค่าประมาณล่าสุดของปัจจัยฤดูกาลที่เหมาะสมในการพยากรณ์

ค่าประมาณเริ่มต้น

จากสมการปรับเรียบโดยส่วนรวม สมการปรับเรียบแนวโน้ม และสมการปรับเรียบฤดูกาล พบว่า จะต้องหาค่าประมาณเริ่มต้น S_t, b_t และ I_t ณ เวลาเริ่มต้น 0 สำหรับ $t = 1,2,......,L$ โดย L เป็นจำนวนฤดูกาลที่แตกต่างกัน ที่จะกล่าวต่อไปนี้ ฤดูกาลเกิดซ้ำทุก ๆ ปี ดังนั้น L เป็นจำนวนฤดูกาลใน 1 ปี สมมติว่ามีข้อมูลในอดีตจำนวน m ปี

ให้ x_i เป็นค่าเฉลี่ยของค่าสังเกตในปีที่ i โดย i มีค่าเท่ากับ $1,2,......,m$

ค่าประมาณเริ่มต้นของส่วนประกอบแนวโน้ม b_t คือ

$$b_0 = \frac{x_m - x_1}{(m - 1)L}$$

เมื่อ x_m เป็นค่าเฉลี่ยค่าสังเกตในปีที่ m

ค่าประมาณเริ่มต้นของ S_t คือ

$$S_0 = x_1 - (L/2)\ b_0$$

และค่าประมาณเริ่มต้นสำหรับปัจจัยฤดูกาล L ค่า คำนวณจากนิพจน์

$$I_t = \frac{x_t}{x_1 - \left[\frac{L + 1}{2} - j\right] b_0}$$

ซึ่งแสดงปัจจัยฤดูกาลและความคลาดเคลื่อนสุ่ม

เมื่อ $\bar{x}_i$ เป็นค่าเฉลี่ยของค่าสังเกตสำหรับปีที่ฤดูกาล t ปรากฏ

ดังนั้น $i = 1$ เมื่อ $1 \leq t \leq L$

และ $i = 2$ เมื่อ $L+1 \leq t \leq 2L$

ฯลฯ

j แทนตำแหน่งของฤดูกาล t ภายในปี ถ้าอนุกรมเวลาประกอบด้วยข้อมูลรายเดือน ดังนั้น j = 1 สำหรับเดือนมกราคม j = 2 สำหรับเดือนกุมภาพันธ์ เป็นต้น

I_t ในสมการ จะให้ค่าประมาณ m ค่าสำหรับแต่ละปัจจัยฤดูกาลที่ต่าง ๆ กัน ค่าหนึ่งสำหรับปีหนึ่ง ถ้าอนุกรมเวลาเป็นรายเดือน จะได้ค่าประมาณปัจจัยฤดูกาล m ค่า สำหรับแต่ละเดือน ค่าประมาณ m ค่านี้ จะเฉลี่ยได้ค่าประมาณหนึ่งค่าสำหรับแต่ละฤดูกาลที่ต่างกัน คือ

$$\bar{I}_t = (1/m) \sum_{k=0}^{m-1} I_{t+kL} \quad \text{สำหรับ } t = 1,2,\ldots,L$$

ซึ่งเป็นดัชนีฤดูกาลเฉลี่ยสำหรับแต่ละฤดูกาลที่แตกต่างกัน

ปัจจัยฤดูกาลที่คำนวณได้นี้ จะต้องปรับค่าให้ผลบวกเท่ากับ L คือ

$$I_t\,(\text{ที่เวลาเริ่มต้น}) = \frac{\bar{I}_t\,[\,L\,]}{\sum_{t=1}^{L} \bar{I}_t}$$

การกำหนดค่า α , β และ δ มีหลักดังต่อไปนี้ คือ α , β และ δ ที่ให้ค่าผลบวกกำลังสองค่าคลาดเคลื่อนในการพยากรณ์ต่ำสุด จะเป็นค่า α , β และ δ ที่ดีที่สุด ในการหา α , β และ δ จะเริ่มด้วยการคำนวณค่าเริ่มต้น S_t , b_t และ I_t สำหรับ

$t = 1,2,......,\alpha$ โดยใช้ข้อมูล m_1 ปีแรก จากที่มีอยู่ m ปี การกำหนดค่า α , β , δ นี้ ทำโดยการลองผิดลองถูก ซึ่งต้องใช้เวลาและค่าใช้จ่ายสูง

ในการประมาณค่าเริ่มต้นสำหรับค่าสถิติปรับเรียบ ปัจจัยฤดูกาล และค่าแนวโน้มนั้น เมื่อไม่มีข้อมูลในอดีต จะใช้ค่าต่อไปนี้เป็นค่าเริ่มต้น คือ

$$S_{L+1} = x_{L+1}$$

เมื่อ L คือความยาวของฤดูกาล และค่าประมาณเริ่มต้นของปัจจัยฤดูกาล คือ

$$I_1 = X_1/X$$
$$I_2 = X_2/X$$
$$\vdots$$
$$I_L = X_L/X$$

$$\text{โดย } x = \sum_{i=1}^{L} x_i/L$$

และ $b_{L+1} = \{(x_{L+1} - x_1) + (x_{L+2} - x_2) + (x_{L+3} - x_3)\} / 3(L)$
เนื่องจากจำนวนข้อมูลที่เพียงพอสำหรับการประมาณค่าเริ่มต้นอยู่ระหว่าง 8 ถึง 3(L)

เทคนิคการวิเคราะห์อนุกรมเวลาแบบบอกซ์และเจนกินส์
(Box and Jenkins Method)

การวิเคราะห์อนุกรมเวลาด้วยเทคนิคของบอกซ์และเจนกินส์ (Box and Jenkins) แตกต่างจากเทคนิคทำให้เรียบ คือ ไม่ได้กำหนดตัวแบบก่อนการวิเคราะห์ แต่เลือกตัวแบบจากการพิจารณาลักษณะของสหสัมพันธ์ในตนเอง (Autocorrelation) และสหสัมพันธ์ในตนเองส่วนย่อย (Partial autocorrelation) ของอนุกรมเวลาที่อยู่ในสภาวะสเตชันนารี (Stationary) โดยบอกซ์และเจนกินส์สร้างตัวแบบ 3 ตัวแบบ คือ ตัวแบบการถดถอยในตนเอง (Autoregressive model) ตัวแบบเฉลี่ยเคลื่อนที่ (Moving average model) และตัวแบบผสมการถดถอยในตนเองกับเฉลี่ยเคลื่อนที่ (Mixed autoregressive moving average model)

การแปลงอนุกรมเวลาให้เป็นกระบวนการสเตชันนารี

กระบวนการสเตชันนารี มีค่าเฉลี่ยและความแปรปรวนคงที่ ซึ่งเป็นคุณสมบัติที่จำเป็นสำหรับการวิเคราะห์อนุกรมเวลาด้วยเทคนิคบอกซ์และเจนกินส์ ถ้าขาดคุณสมบัติดังกล่าวจะต้องแปลงกระบวนการให้เป็นแบบสเตชันนารี ซึ่งมี 2 กรณี คือ

1.กรณีค่าเฉลี่ยไม่คงที่เปลี่ยนไปตามเวลา

ถ้าอนุกรมเวลามีค่าเฉลี่ยเปลี่ยนไปตามเวลาโดยที่ความแปรปรวนคงที่ และอนุกรมเวลานี้ไม่มีฤดูกาล สามารถแปลงอนุกรมเวลานี้ให้เป็นอนุกรมเวลาสเตชันนารี โดยการหาค่าแตกต่าง ดังนี้

ค่าแตกต่างครั้งที่หนึ่ง (First differences) แทนด้วย ∇x_t และ

$$\nabla x_t = x_t - x_{t-1} \quad \text{สำหรับ } t = 1,2,\ldots,n$$

ค่าแตกต่างครั้งที่สอง (second differences) แทนด้วย $\nabla^2 x_t$ และ

$$\begin{aligned}\nabla^2 x_t &= \nabla(\nabla x_t)\\ &= (x_t - x_{t-1}) - (x_{t-1} - x_{t-2})\\ &= x_t - 2x_{t-1} + x_{t-2} \quad \text{สำหรับ } t = 3,4,\ldots,n\end{aligned}$$

$\nabla^{d}x_{t}$ แทนค่าแตกต่างครั้งที่ d ได้จากการหาค่าแตกต่างเป็นลำดับจากค่าแตกต่างครั้งก่อน ๆ

ถ้าอนุกรมเวลามีฤดูกาลด้วย การหาค่าแตกต่างฤดูกาล จะหาค่าแตกต่างของค่าอนุกรมเวลาที่อยู่ห่างกันเท่ากับความยาวของฤดูกาล s และแทนด้วย $\nabla_{s}x_{t}$

ค่าแตกต่างฤดูกาลครั้งที่หนึ่ง คือ

$$\nabla_{s}x_{t} = x_{t} - x_{t-s}$$

เช่น ข้อมูลรายงวด s = 4

ค่าแตกต่างครั้งที่หนึ่ง คือ $\nabla_{4}x_{t} = x_{t} - x_{t-4}$

ค่าแตกต่างครั้งที่สอง คือ
$$\begin{aligned}\nabla_{4}^{2}X_{t} &= \nabla_{4}(x_{t} - x_{t-4}) \\ &= (x_{t} - x_{t-4}) - (x_{t-4} - x_{t-8}) \\ &= x_{t} - 2x_{t-4} + x_{t-8}\end{aligned}$$

และค่าแตกต่างฤดูกาลครั้งที่ D แทนด้วย $\nabla_{s}^{D}x_{t}$

ในบางครั้งอาจต้องหาค่าแตกต่างแบบไม่มีฤดูกาลและแบบมีฤดูกาลด้วย เมื่อค่าเฉลี่ยไม่มีฤดูกาลและมีฤดูกาลมีการเปลี่ยนแปลงตามเวลาที่เปลี่ยนแปลงไป กรณีเช่นนี้จะต้องหาค่าแตกต่างทั้งแบบไม่มีฤดูกาลและแบบมีฤดูกาล โดยจะหาค่าแตกต่างแบบใดก่อนก็ได้ ซึ่งแสดงให้เห็นได้ดังนี้

$$\begin{aligned}\nabla_{4}X_{t} &= \nabla_{4}(x_{t} - x_{t-1}) \\ &= (x_{t} - x_{t-1}) - (x_{t-4} - x_{t-5}) \\ &= (x_{t} - x_{t-4}) - (x_{t-1} - x_{t-5}) \\ &= \nabla\nabla_{4}x_{t}\end{aligned}$$

ค่าแตกต่างครั้งที่ d แบบไม่มีฤดูกาล เมื่อหาร่วมกันกับค่าแตกต่างอย่างมีฤดูกาลครั้งที่ D แทนด้วยสัญลักษณ์ $\nabla_{s}^{D}\nabla^{d}x_{t}$

เมื่อ D แทนอันดับของค่าแตกต่างฤดูกาล

d แทนอันดับของค่าแตกต่างไม่มีฤดูกาล

s แทนความยาวของฤดูกาล

สัญลักษณ์ที่ใช้แทน ∇ อีกสัญลักษณ์หนึ่งคือ 1 - B เรียก B ว่า Back shift operator โดย

$$Bx_t = x_{t-1}$$
$$B^2x_t = B(Bx_t) = Bx_{t-1} = x_{t-2}$$

เมื่อย้อนหลังไป k หน่วยเวลา จะแทนด้วย

$$B^k x_t = x_{t-k}$$

จาก
$$\begin{aligned} \nabla x_t &= x_t - x_{t-1} \\ &= x_t - Bx_t \\ &= x_t(1-B) \\ \nabla &= 1 - B \end{aligned}$$

และค่าแตกต่างครั้งที่ d เท่ากับ $(1 - B)^d$ คือ

$$\nabla^d x_t = (1 - B)^d x_t$$

2. กรณีความแปรปรวนเปลี่ยนไปตามเวลา

การแปลงกระบวนการที่ความแปรปรวนไม่คงที่ แปลงได้หลายวิธี ขึ้นกับลักษณะการเปลี่ยนแปลงของความแปรปรวน ถ้าความแปรปรวนเป็นสัดส่วนกับค่าเฉลี่ยของอนุกรมเวลา โดยที่ค่าเฉลี่ยอนุกรมเวลาเพิ่มขึ้น หรือ ลดลงอย่างคงที่ ก็ควรจะแปลงด้วยล็อกการิทึม ($\ln x_t$) วิธีอื่น ๆ ที่จะแปลงให้ค่าความแปรปรวนคงที่ เช่น แปลงด้วยรากที่สอง, แปลงด้วยการกลับเศษส่วน, แปลงด้วยรากที่สี่ เป็นต้น

ในอนุกรมเวลาชุดหนึ่ง ๆ อาจมีค่าเฉลี่ยและค่าความแปรปรวนไม่คงที่ทั้งสองอย่าง ควรจะแปลงให้ความแปรปรวนคงที่ก่อนจะแปลงค่าเฉลี่ยให้คงที่

การวิเคราะห์อนุกรมเวลาด้วยเทคนิคบ๊อกซ์และเจนกินส์ ประกอบด้วย 4 ขั้น คือ

ขั้นที่ 1 ขั้นการค้นหาตัวแบบ (Identification) เป็นการกำหนดตัวแบบที่เหมาะกับอนุกรมที่ต้องการพยากรณ์

ขั้นที่ 2 ขั้นการประมาณค่า(Estimation)เป็นการประมาณค่าพารามิเตอร์ของตัวแบบ

ขั้นที่ 3 ขั้นตรวจสอบตัวแบบ (Diagnostic checking) เป็นการตรวจว่าตัวแบบเหมาะสมกับข้อมูลหรือไม่

ขั้นที่ 4 ขั้นพยากรณ์ (Forecasting) ใช้ตัวแบบที่เหมาะสมพยากรณ์ค่าของข้อมูลในอนาคต

<u>ขั้นที่ 1</u> การค้นหาตัวแบบ

ตัวแบบทั่วไป คือ ARIMA(p,d,q)

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Z_t = \theta_0 + (1 - \theta_1 B - \dots - \theta_q B^q)a_t$$

สามารถแยกย่อยได้อีก 3 ขั้นตอน ดังนี้

1. เขียนกราฟอนุกรมเวลาและเลือกรูปแบบการแปลงค่า

นำข้อมูลมาเขียนกราฟ เพื่อพิจารณาว่าอนุกรมชุดนั้นประกอบด้วยส่วนประกอบแนวโน้ม, ฤดูกาล, รูปแบบไม่ปกติ อนุกรมอยู่ในสภาวะสมดุลย์หรือไม่ จากนั้นเลือกวิธีการแปลงที่เหมาะสม เพื่อให้ความแปรปรวนคงที่

2. ถ้าค่าสหสัมพันธ์ในตัวเองลดลงอย่างช้า ๆ และค่าสหสัมพันธ์ในตัวเองบางส่วนตัดออกได้หลังคาบเวลาที่ล่ากว่ากัน 1 ต้องหาผลต่างอันดับหนึ่ง ถ้ายังไม่สมดุลย์ต้องหาผลต่างอันดับสูงขึ้นไปอีก

3. คำนวณค่าสหสัมพันธ์ในตัวเองและสหสัมพันธ์ในตัวเองบางส่วนของอนุกรมที่อยู่ในสภาวะสมดุลย์ เพื่อนำมาพิจารณาในการเลือกอันดับ p,q

ตาราง 2.1 แสดงหลักเกณฑ์ในการเลือกตัวแบบ ARIMA(p,d,q)

กระบวนการ	ACF	PACF
AR(1)	ลดลงแบบเอกซโปเนนเชียล หรือคลื่นรูป sine	ตัดออกได้หลังคาบเวลาที่ล่ากว่ากัน p
MA(q)	ตัดออกหลังคาบเวลาที่ล่ากว่ากัน	ลดลงแบบเอกซโปเนนเชียล หรือคลื่นรูป sine
ARMA(p,q)	ตัดออกหลังคาบเวลาที่ล่ากว่ากัน (p-q)	ตัดออกหลังคาบเวลาที่ล่ากว่ากัน (p-q)

สูตรการหาฟังก์ชันสหสัมพันธ์ในตัวเอง

$$r_k = \frac{\sum_{t=a}^{n-k} (z_t - z)(z_{t+k} - z)}{\sum_{t=a}^{n} (z_t - z)^2}$$

เมื่อ z เป็นค่าเฉลี่ยของอนุกรมเวลา $z_1, z_2, z_3, \ldots, z_n$

$$z = \frac{\sum_{t=a}^{n} z_t}{n-a+1}$$

สูตรการหาฟังก์ชันในตัวเองบางส่วน

$$r_{kk} = \begin{cases} r_1 & k=1 \\ \dfrac{r_k - \sum_{j=1}^{k-1} r_{k-1,j}\, r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j}\, r_j} & k=2,3,\ldots \end{cases}$$

โดย $r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j}$ สำหรับ $j = 1,2,\ldots,k-1$

ขั้นที่ 2 การประมาณค่าพารามิเตอร์

เมื่อได้รูปแบบจากขั้นที่ 1 แล้ว จากนั้นก็ทำการประมาณค่าพารามิเตอร์ในตัวแบบ นิยมใช้วิธีกำลังสองน้อยที่สุด

ขั้นที่ 3 การตรวจสอบความเหมาะสมของตัวแบบ

หลังจากที่ประมาณค่าพารามิเตอร์แล้ว จำเป็นที่จะต้องทำการตรวจสอบว่าตัวแบบที่เลือกนั้นสามารถใช้เป็นตัวแบบในการพยากรณ์ได้ดีหรือไม่

มีวิธีตรวจสอบดังนี้

1. ดูจากค่าความคลาดเคลื่อน

ค่าความคลาดเคลื่อนภายหลังการพยากรณ์ ควรจะเป็นอิสระอย่างแท้จริง (Random noise) โดยใช้สถิติบอกซ์-เพียร์ซไคสแควร์ (Box-Pierce Chi-Square Statistics) เป็นสถิติการแจกแจง แบบไคสแควร์ของค่าสหสัมพันธ์ในตัวเองของความคลาดเคลื่อน (Autocorrelation of residual)

2. ศึกษาค่าสถิติตัวอย่างจากวิธีที่ดีที่สุดนั้น เพื่อดูว่าจะสามารถทำตัวแบบให้ง่ายขึ้นกว่านั้นหรือไม่

สมมติฐานของการทดสอบ

H_0 : ไม่มีสหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อนที่ล่ากว่ากัน k คาบเวลา หรือ สหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อนเท่ากับศูนย์

H_1 : มีสหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อนที่ล่ากว่ากัน k คาบเวลา หรือ สหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อนไม่เท่ากับศูนย์

คำนวณ

สัญลักษณ์ Q

$$Q = (N-d) \sum_{i=1}^{k} r_i^2 (e_i)$$

ถ้าข้อมูลมีฤดูกาล

$$Q = (N-d-DL) \sum_{i=1}^{k} r_i^2 (e_i)$$

เมื่อ N คือ จำนวนข้อมูลที่ใช้ในการวิเคราะห์อนุกรมเวลา

d คือ จำนวนอันดับของการหาค่าความแตกต่างที่ทำให้อนุกรมอยู่ในสภาวะสมดุลย์

k คือ จำนวนคาบเวลาที่ล่ากว่ากันของสหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อน

$r_i^2 (e_i)$ คือ กำลังสองของสหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อน

L คือ จำนวนระยะเวลาของฤดูกาลของฤดูกาลในแต่ละปี

D คือ จำนวนอันดับของการหาค่าความแตกต่างฤดูกาลที่ทำให้อนุกรมอยู่ในสภาวะสมดุลย์

หากค่า Q ที่คำนวณได้น้อยกว่าที่องศาแห่งความเป็นอิสระเท่ากับ k-np จะยอมรับสมมติฐาน H_0 แสดงว่าไม่มีสหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อนที่ล่ากว่ากัน k

กัน k คาบเวลา แสดงว่าตัวแบบที่ใช้เหมาะสมกับอนุกรมเวลาชุดนั้น เมื่อ np คือจำนวนพารามิเตอร์ในตัวแบบ

<u>ขั้นที่ 4</u> การพยากรณ์ข้อมูล

เมื่อได้ตัวแบบที่เหมาะสมแล้วจึงทำการพยากรณ์ สำหรับวิธีบอกซ์และเจนกินส์ใช้ได้ดีสำหรับการพยากรณ์ล่วงหน้า 1 คาบเวลา

ตัวแบบสำหรับการวิเคราะห์อนุกรมเวลาแบบบอกซ์และเจนกินส์

- อนุกรมเวลาไม่มีการผันแปรตามฤดูกาล
- อนุกรมเวลามีการผันแปรตามฤดูกาล

จำแนกตัวแบบอนุกรมเวลาได้ 6 ประเภท คือ

1. กระบวนการถดถอยในตนเองแบบไม่มีฤดูกาลอันดับ p (Nonseasonal Autoregressive Process of Order p (AR(p))) เป็นกระบวนการอนุกรมเวลาที่ค่าปัจจุบัน x_t แทนได้ด้วยฟังก์ชันเชิงเส้นของค่าในอดีตกับค่าความรบกวนสุ่ม a_t โดยที่อนุกรมเวลา $\{x_t\}$ เป็นอนุกรมเวลาที่อยู่ในสภาวะสเตชันนารี

เมื่อกำหนดให้ $z_t = x_t - \mu$

จะได้กระบวนการหรือตัวแบบ (model) การถดถอยในตนเองอันดับ p คือ

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + e_t$$

เมื่อ $\phi_1, \phi_2, \dots, \phi_p$ คือ พารามิเตอร์ของการถดถอยในตัวเองแบบไม่มีฤดูกาล (Nonseasonal Autoregressive Parameter)

e_t คือ ค่าความคลาดเคลื่อน ณ คาบเวลา t

z_t คือ อนุกรมเวลาที่อยู่ในสภาวะสเตชันนารีทั้งความแปรปรวนและค่าเฉลี่ย

2.กระบวนการเฉลี่ยเคลื่อนที่แบบไม่มีฤดูกาลอันดับ q (Nonseasonal Moving Average Process of Order q (MA(q)))เป็นการแสดงความสัมพันธ์ของข้อมูลปัจจุบันกับค่าความคลาดเคลื่อนของอนุกรมเวลาในอดีต ซึ่งหมายถึงว่า ความแปรปรวนใน q หน่วยเวลาในอดีต จะมีผลกระทบต่อค่าปัจจุบันของอนุกรมเวลา ซึ่งมีรูปแบบดังนี้

เมื่อกำหนดให้ $z_t = x_t - \mu$

จะได้กระบวนการหรือตัวแบบ (model) การเฉลี่ยเคลื่อนที่แบบไม่มีฤดูกาลอันดับ q คือ

$$z_t = \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} + e_t$$

เมื่อ $\theta_1, \theta_2, \dots, \theta_q$ คือ พารามิเตอร์การเฉลี่ยเคลื่อนที่แบบปกติไม่มีฤดูกาล (Nonseasonal Moving Average Parameter)

e_t คือ ค่าความคลาดเคลื่อน ณ คาบเวลา t

z_t คือ อนุกรมเวลาที่อยู่ในสภาวะสเตชันนารีทั้งความแปรปรวนและค่าเฉลี่ย

3.กระบวนการผสมการถดถอยในตัวเองแบบไม่มีฤดูกาลอันดับ p กับการเฉลี่ยเคลื่อนที่แบบไม่มีฤดูกาลอันดับ q ซึ่งมีรูปแบบดังนี้

เมื่อกำหนดให้ $z_t = x_t - \mu$

จะได้กระบวนการหรือตัวแบบ (model) การถดถอยในตนเองอันดับ p กับการเฉลี่ยเคลื่อนที่แบบไม่มีฤดูกาลอันดับ q คือ

$$z_t = \phi_1 z_{t-1} - \phi_2 z_{t-2} - \dots - \phi_p z_{t-p} - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} + e_t$$

เมื่อ $\theta_1, \theta_2, \dots, \theta_q$ คือ พารามิเตอร์ของการเฉลี่ยเคลื่อนที่แบบไม่มีฤดูกาล (Nonseasonal Moving Average Parameter)

$\phi_1, \phi_2, \ldots, \phi_p$ คือ พารามิเตอร์ของการถดถอยในตัวเองแบบไม่มีฤดูกาล (Nonseasonal Autoregressive Parameter)

e_t คือ ค่าความคลาดเคลื่อน ณ คาบเวลา t

z_t คือ อนุกรมเวลาที่อยู่ในสภาวะสเตชันนารีทั้งความแปรปรวนและค่าเฉลี่ย

4.กระบวนการถดถอยในตนเองแบบมีฤดูกาลอันดับ P (Nonseasonal Autoregressive Process of Order P (AR(P)$_s$) เป็นกระบวนการอนุกรมเวลาที่ค่าปัจจุบัน x_t แทนได้ด้วยฟังก์ชันเชิงเส้นของค่าในอดีตกับค่าความรบกวนสุ่ม a_t โดยที่อนุกรมเวลา $\{x_t\}$ เป็นอนุกรมเวลาที่อยู่ในสภาวะสเตชันนารี

เมื่อกำหนดให้ $z_t = x_t - \mu$

จะได้กระบวนการหรือตัวแบบ (model) การถดถอยในตนเองอันดับ P คือ

$$z_t = \Phi_1 z_{t-1} + \ldots + \Phi_P z_{t-P} + e_t$$

เมื่อ $\Phi_1, \ldots, \Phi_P$ คือ พารามิเตอร์ของการถดถอยในตัวเองแบบมีฤดูกาล (Seasonal Autoregressive Parameter)

e_t คือ ค่าความคลาดเคลื่อน ณ คาบเวลา t

z_t คือ อนุกรมเวลาที่อยู่ในสภาวะสเตชันนารีทั้งความแปรปรวนและค่าเฉลี่ย

5.กระบวนการเฉลี่ยเคลื่อนที่แบบมีฤดูกาลอันดับ Q (Seasonal Moving Average Process of Order Q (MA(Q)$_s$)) เป็นการแสดงความสัมพันธ์ของข้อมูลปัจจุบันกับค่าความคลาดเคลื่อนของอนุกรมเวลาในอดีต หมายความว่า ความแปรปรวนใน Q หน่วยเวลาในอดีตจะมีผลกระทบต่อค่าปัจจุบันของอนุกรมเวลา ซึ่งมีรูปแบบดังนี้

เมื่อกำหนดให้ $z_t = x_t - \mu$

จะได้กระบวนการหรือตัวแบบ (model) การเฉลี่ยเคลื่อนที่แบบมีฤดูกาลอันดับ Q คือ

$$z_t = \Theta_1 a_{t-1} - \ldots - \Theta_Q a_{t-Q} + e_t$$

เมื่อ $\Theta_1, \ldots, \Theta_Q$ คือ พารามิเตอร์ของการเฉลี่ยเคลื่อนที่แบบมีฤดูกาล (Seasonal Moving Average Parameter)

e_t คือ ค่าความคลาดเคลื่อน ณ คาบเวลา t

z_t คือ อนุกรมเวลาที่อยู่ในสภาวะสเตชันนารีทั้งความแปรปรวนและค่าเฉลี่ย

6.กระบวนการผสมการถดถอยในตัวเองแบบมีฤดูกาลอันดับ P กับการเฉลี่ยเคลื่อนที่แบบมีฤดูกาลอันดับ Q ซึ่งมีรูปแบบดังนี้

เมื่อกำหนดให้ $z_t = x_t - \mu$

จะได้กระบวนการหรือตัวแบบ (model) การถดถอยในตนเองอันดับ P กับการเฉลี่ยเคลื่อนที่อันดับ Q แบบมีฤดูกาล คือ

$$\begin{aligned} z_t = {} & \Phi_1 z_{t-1} - \Phi_2 z_{t-2} - \ldots - \Phi_P z_{t-P} - \Theta_1 a_{t-1} \\ & - \Theta_2 a_{t-2} - \ldots - \Theta_Q a_{t-Q} + a_t \end{aligned}$$

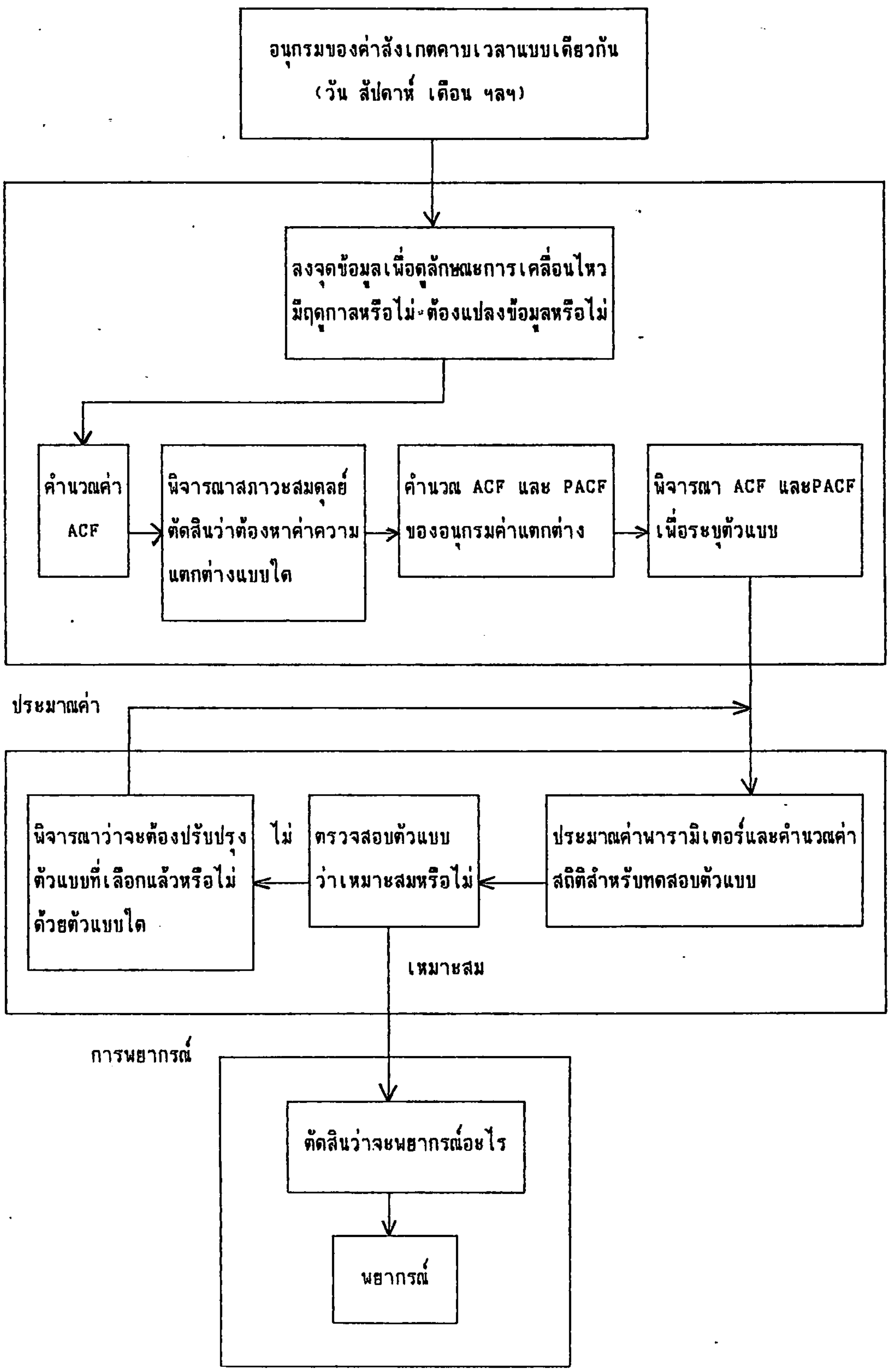
เมื่อ $\Phi_1, \ldots, \Phi_P$ คือ พารามิเตอร์ของการถดถอยในตัวเองแบบมีฤดูกาล (Seasonal Autoregressive Parameter)

$\Theta_1, \ldots, \Theta_Q$ คือ พารามิเตอร์ของการเฉลี่ยเคลื่อนที่แบบมีฤดูกาล (Seasonal Moving Average Parameter)

a_t คือ ค่าความคลาดเคลื่อน ณ คาบเวลา t

z_t คือ อนุกรมเวลาที่อยู่ในสภาวะสเตชันนารีทั้งความแปรปรวนและค่าเฉลี่ย

แผนผังวิเคราะห์อนุกรมเวลาด้วยเทคนิคของบ๊อกซ์และเจนกินส์



2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องนี้จะครอบคลุมถึง เอกสารและงานวิจัยที่ทำการศึกษาเกี่ยวกับการป่วยด้วยโรค วัณโรค คอตีบ ไอกรน บาดทะยัก และโปลิโอ และเอกสารและงานวิจัยซึ่งใช้วิธีการวิเคราะห์อนุกรมเวลาในการวิเคราะห์ข้อมูล

<u>เอกสารและงานวิจัยที่ทำการศึกษาเกี่ยวกับการป่วยด้วยโรค วัณโรค คอตีบ ไอกรน บาดทะยัก และโปลิโอ</u>

<u>กองระบาดวิทยา</u> รายงานว่าในปี พ.ศ.2534 มีผู้ป่วยไข้คอตีบทั่วประเทศจำนวน 53 ราย จาก 32 จังหวัด มีผู้ป่วยเสียชีวิต 11 ราย หากจะคิดเป็นอัตราป่วยตาย ซึ่งเป็นสัดส่วนระหว่างจำนวนผู้ป่วยซึ่งเสียชีวิตต่อจำนวนผู้ป่วยทั้งหมด อัตราป่วยตายจากโรคคอตีบจะเท่ากับร้อยละ 20.8 ซึ่งนับเป็นอัตราป่วยตายที่สูงที่สุดในรอบสิบปี จากเดิมซึ่งมีอัตราป่วยตายอยู่ระหว่าง 4.1 ถึง 15.0 สำหรับโรควัณโรค ได้รับรายงานจำนวนผู้ป่วยทั้งสิ้น 21,107 ราย ซึ่งเพิ่มขึ้นจากปี 2533 เพียงเล็กน้อย ส่วนจำนวนผู้ป่วยโรคโปลิโอในประเทศไทยพบว่ามีจำนวนเพิ่มขึ้น และกองระบาดวิทยาได้รับรายงานจำนวนผู้ป่วยโรคบาดทะยัก 166 ราย มีผู้เสียชีวิตจำนวน 27 ราย

<u>พิพัฒน์ ลักษมีจรัลกุล และ เสาวลักษณ์ ลักษมีจรัลกุล</u> (2528) ทำการศึกษาการตอบสนองแอนติบอดีต่อโรคไอกรนในเด็กที่ได้รับวัคซีนป้องกันโรคคอตีบ บาดทะยัก และไอกรน ครบ 3 ครั้ง พบว่า โครงการสร้างเสริมภูมิคุ้มกันโรคไอกรน โดยให้วัคซีนป้องกันโรคคอตีบ บาดทะยัก และไอกรน(วัคซีน DTP) เริ่มตั้งแต่ พ.ศ.2520 แต่ก็ยังมีการป่วยด้วยโรคไอกรนทุกปี

<u>อะเคื้อ อุณหเลขกะ และคณะ</u> (2528) ได้ทำการศึกษาระบาดวิทยาของโรคโปลิโอในกรุงเทพมหานคร พ.ศ.2526-2527 พบว่า อุบัติการณ์ของจำนวนผู้ป่วยโรคโปลิโอในกรุงเทพมหานครลดลง

<u>เอกสารและงานวิจัยซึ่งใช้วิธีการวิเคราะห์อนุกรมเวลาในการวิเคราะห์ข้อมูล</u>

<u>กฤษฎา วาณิชชัชวาลย์ และคณะ</u> (2534) ทำการวิเคราะห์แนวโน้มจำนวนนักท่อง-เที่ยวชาวต่างประเทศที่เข้ามาในประเทศไทย และพยากรณ์จำนวนนักท่องเที่ยวในเดือนกรกฎาคมปี 2534 พบว่า แนวโน้มจำนวนนักท่องเที่ยวชาวแอฟริกาจะมีฤดูกาลแบบคูณ วิธีพยากรณ์ที่เหมาะสมคือ การปรับแนวโน้มและฤดูกาลของวินเตอร์ ค่าพยากรณ์คือ 2,278 คน แนวโน้มจำนวนนักท่องเที่ยวชาวเอเชียตะวันออกจะมีฤดูกาลแบบคูณ วิธีพยากรณ์ที่เหมาะสมคือการปรับแนวโน้มและฤดูกาลของวินเตอร์ ค่าพยากรณ์คือ 258,387 คน แนวโน้มจำนวนนักท่องเที่ยวชาวยุโรปจะมีฤดูกาลแบบบวก วิธีพยากรณ์ที่เหมาะสมคือ เทคนิคการวิเคราะห์อนุกรมเวลาแบบบอกซ์และเจนกินส์ ค่าพยากรณ์คือ 96,715 คน แนวโน้มจำนวนนักท่องเที่ยวชาวตะวันออกกลางจะมีฤดูกาลแบบคูณ วิธีพยากรณ์ที่เหมาะสมคือ การปรับแนวโน้มและฤดูกาลของวินเตอร์ ค่าพยากรณ์คือ 7,744 คน แนวโน้มจำนวนนักท่องเที่ยวชาวเอเชียใต้จะมีฤดูกาลแบบบวก วิธีพยากรณ์ที่เหมาะสมคือ การปรับแนวโน้มและฤดูกาลของวินเตอร์ ค่าพยากรณ์คือ 23,849 คน แนวโน้มจำนวนนักท่องเที่ยวชาวอเมริกัน จะมีฤดูกาลแบบคูณ วิธีพยากรณ์ที่เหมาะสมคือ เทคนิคการวิเคราะห์อนุกรมเวลาแบบ บอกซ์ และเจนกินส์ ค่าพยากรณ์คือ 431,447 คน แนวโน้มจำนวนนักท่องเที่ยวทั้งหมดจะมีฤดูกาลแบบคูณ วิธีพยากรณ์ที่เหมาะสมคือ การปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้งด้วยวิธีของวินเตอร์ ค่าพยากรณ์คือ 415,537 คน

<u>พิภพ สมิทธินันท์ และคณะ</u> (2532) ทำการวิเคราะห์และพยากรณ์ปริมาณความต้องการใช้น้ำมันเชื้อเพลิงประเภทต่างๆในอนาคต ผลการวิเคราะห์สรุปได้ว่าความต้องการใช้น้ำมัน เตา น้ำมันก๊าด และน้ำมันดีเซลหมุนช้าในปี 2533 จะมีลักษณะคงที่ไม่มีฤดูกาล สามารถพยากรณ์โดยใช้การปรับเรียบแบบเอกซโปเนนเชียลครั้งเดียว

บทที่ 3

วิธีดำเนินการวิจัย

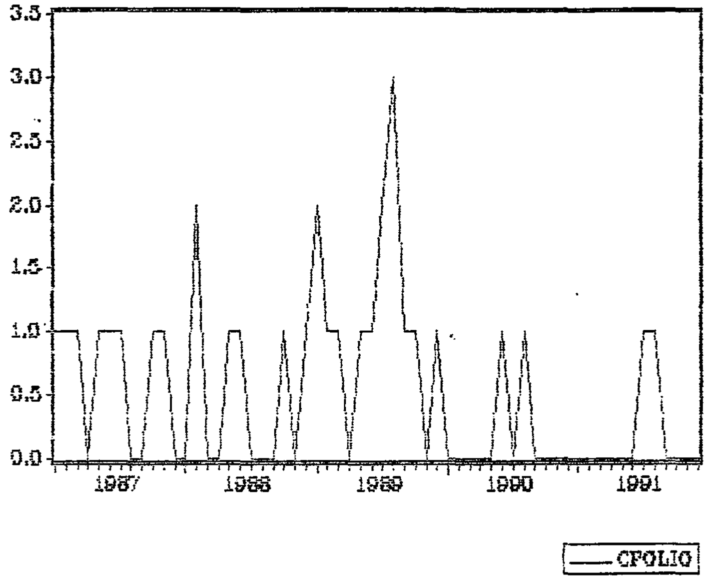
3.1 แหล่งที่มาของข้อมูล

ข้อมูลที่ใช้ในการพยากรณ์จำนวนผู้ป่วยโรคต่างๆ เป็นข้อมูลประเภททุติยภูมิ ซึ่งทางกองระบาดวิทยา กระทรวงสาธารณสุขเป็นหน่วยงานที่ทำหน้าที่รวบรวมข้อมูล โดยจัดเก็บข้อมูลเป็นรายเดือนในแต่ละโรค โดยเก็บข้อมูลรวมทั้งประเทศ และ แยกเป็นภาค คือ ภาคกลาง ภาคเหนือ ภาคตะวันออกเฉียงเหนือ ภาคใต้ และกรุงเทพมหานคร

สาเหตุที่ใช้ข้อมูลทุติยภูมิในปัญหาพิเศษนี้ เพราะว่าการพยากรณ์ข้อมูลในอนาคตนั้น เราจะต้องนำข้อมูลในอดีตมาศึกษาถึงรูปแบบ ปัจจัยต่างๆที่เกี่ยวข้องและแนวโน้มของข้อมูล โดยใช้ข้อมูลตั้งแต่เดือนมกราคม พ.ศ.2530 ถึงเดือนธันวาคม พ.ศ.2534 ซึ่งเป็นข้อมูลรายเดือนจำนวน 60 เดือน

3.2 การวิเคราะห์อนุกรมเวลาด้วยวิธีการเฉลี่ยเคลื่อนที่ครั้งเดียว (Single Moving Average)

การวิเคราะห์อนุกรมเวลาด้วยการเฉลี่ยเคลื่อนที่ครั้งเดียวนั้น ข้อมูลควรจะไม่มีแนวโน้ม โดยมีค่าสังเกตกระจายรอบค่าคงที่ค่าหนึ่ง ดังข้อมูลแสดงจำนวนผู้ป่วยโรคโปลิโอในภาคกลาง ในรูป 3.1



รูปที่ 3.1 แสดงข้อมูลจำนวนผู้ป่วยโรคโปลิโอในภาคกลาง

จากสมการ $S_t = \dfrac{Y_{t-1} + Y_t + Y_{t+1}}{3}$

$$F_t = S_{t-1}$$

นำค่า Y_t ในตารางที่ 3.1 คำนวณค่า F_t

$$F_1 = \frac{114+94+94}{3}$$

$$F_2 = \frac{94+94+96}{3}$$

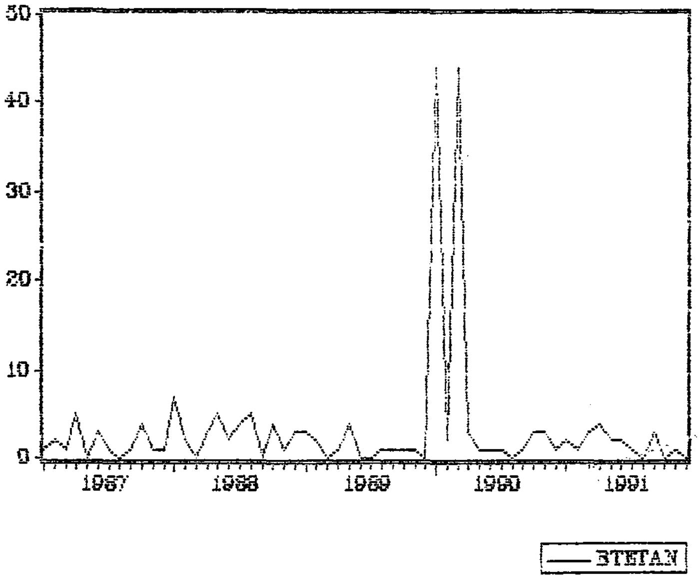
ตารางที่ 3.1 ค่าข้อมูลดิบ ค่าพยากรณ์ และค่าคลาดเคลื่อนของจำนวนผู้ป่วยโรคโปลิโอในภาคกลาง ด้วยวิธีการเฉลี่ยเคลื่อนที่ครั้งเดียว

N = 3

Time	Original Data	Forecast	Error
1	1		
2	1		
3	1		
4	0	1	−1.0
5	1	1	0.3
6	1	1	0.3
7	1	1	0.3
8	0	1	−1.0
9	0	1	−0.7
10	1	0	0.7
11	1	0	0.7
12	0	1	−0.7
13	0	1	−0.7
14	2	0	1.7
15	0	1	−0.7
16	0	1	−0.7
17	1	1	0.3
18	1	0	0.7
19	0	1	−0.7
20	0	1	−0.7
21	0	0	−0.3
22	1	0	1.0
23	0	0	−0.3
24	1	0	0.7
25	2	1	1.3
26	1	1	0.0
27	1	1	−0.3
28	0	1	−1.3
29	1	1	0.3
30	1	1	0.3
31	2	1	1.3
32	3	1	1.7
33	1	2	−1.0
34	1	2	−1.0
35	0	2	−1.7
36	1	1	0.3
37	0	1	−0.7
38	0	0	−0.3
39	0	0	−0.3
40	0	0	0.0
41	0	0	0.0
42	1	0	1.0
43	0	0	−0.3
44	1	0	0.7
45	0	1	−0.7
46	0	0	−0.3
47	0	0	−0.3
48	0	0	0.0
49	0	0	0.0
50	0	0	0.0
51	0	0	0.0
52	0	0	0.0
53	0	0	0.0
54	0	0	0.0
55	1	0	1.0
56	1	0	0.7
57	0	0	−0.7
58	0	1	−0.7
59	0	0	−0.3
60	0	0	0.0
61	0	0	
62		0	
63		0	
64		0	
65		0	
66		0	
67		0	
68		0	
69		0	
70		0	
71		0	
72		0	

3.3 การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียว (Single Exponential Smoothing : SES)

การวิเคราะห์อนุกรมเวลาด้วยการปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียว ข้อมูลควรจะไม่มีแนวโน้ม โดยมีค่าสังเกตกระจายรอบค่าคงที่ค่าหนึ่ง ดังข้อมูลแสดงจำนวนผู้ป่วยโรคบาดทะยักในกรุงเทพมหานคร ในรูป 3.2



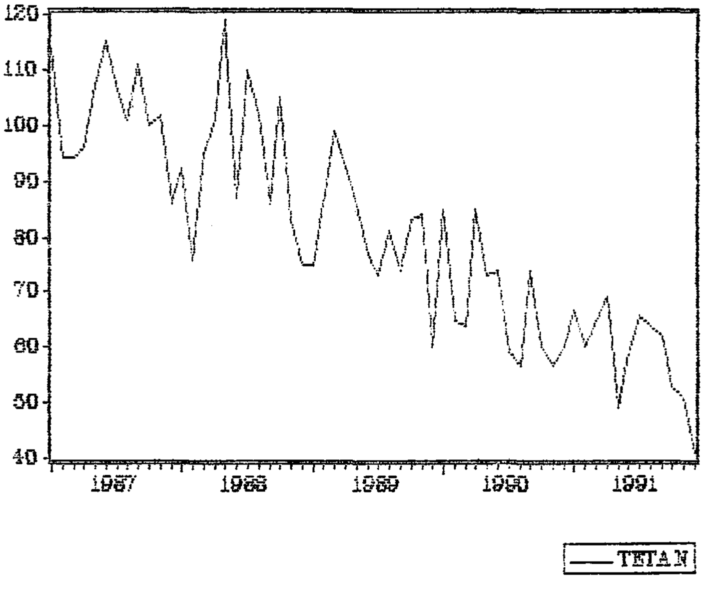
รูปที่ 3.2 แสดงข้อมูลจำนวนผู้ป่วยโรคบาดทะยักในกรุงเทพมหานคร

ตารางที่ 3.2 ค่าข้อมูลดิบ ค่าพยากรณ์ และค่าคลาดเคลื่อนของจำนวนผู้ป่วยโรคบาดทะยักในกรุงเทพมหานคร ด้วยวิธี SES ที่ $\alpha = 0.02$

Time	Original Data	Forecast	Error
1	1	-	
2	2	2	0.2
3	1	2	−78.8
4	5	2	64.5
5	0	2	
6	3	2	40.0
7	1	2	−82.4
8	0	2	
9	1	2	−77.2
10	4	2	56.1
11	1	2	−80.1
12	1	2	−78.5
13	7	2	74.7
14	2	2	6.3
15	0	2	
16	3	2	38.7
17	5	2	62.8
18	2	2	3.8
19	4	2	51.8
20	5	2	60.6
21	0	2	
22	4	2	50.3
23	1	2	−102.8
24	3	2	33.1
25	3	2	32.4
26	2	2	−2.4
27	0	2	
28	1	2	−100.5
29	4	2	50.4
30	0	2	
31	0	2	
32	1	2	−94.5
33	1	2	−92.6
34	1	2	−90.8
35	1	2	−89.0
36	0	2	
37	44	2	95.8
38	2	3	−33.9
39	44	3	93.9
40	3	3	−16.4
41	1	3	−248.1
42	1	3	−243.1
43	1	3	−238.3
44	0	3	
45	1	3	−226.8
46	3	3	−7.4
47	3	3	−7.3
48	1	3	−221.4
49	2	3	−58.5
50	1	3	−214.7
51	3	3	−3.5
52	4	3	22.5
53	2	3	−56.0
54	2	3	−54.9
55	1	3	−207.5
56	0	3	
57	3	3	0.9
58	0	3	
59	1	3	−191.4
60	0	3	
61		3	
62		3	
63		3	
64		3	
65		3	
66		3	
67		3	
68		3	
69		3	
70		3	
71		3	
72		3	

3.4 การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้งของบราวน์ (Brown's Double Exponential Smoothing)

การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้งค่าสังเกตในอดีตควรจะมีแนวโน้มเป็นเส้นตรง ดังเช่นข้อมูลจำนวนผู้ป่วยโรคบาดทะยักทั่วประเทศ ในรูป 3.3



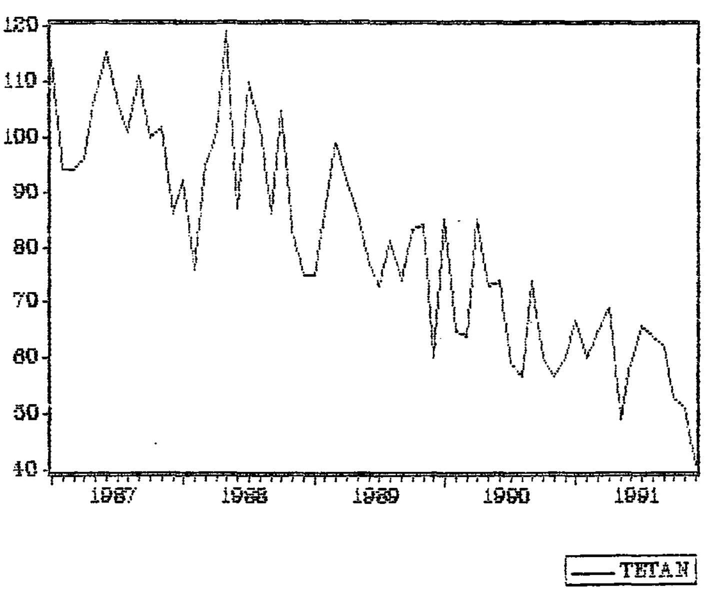
รูปที่ 3.3 แสดงแนวโน้มของจำนวนผู้ป่วยโรคบาดทะยักทั่วประเทศ

ตารางที่ 3.3 ค่าข้อมูลดิบ ค่าพยากรณ์ และค่าคลาดเคลื่อนของจำนวนผู้ป่วยโรคบาดทะยักทั่วประเทศ ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้ง ของบราวน์ ที่ α = 0.11

Time	Original Data	Forecast	Error
1	114		
2	94	105	−11.7
3	94	103	−9.7
4	96	102	−5.8
5	107	101	6.0
6	115	102	11.1
7	106	105	0.6
8	101	106	−4.9
9	111	105	5.1
10	100	107	−7.0
11	102	106	−3.9
12	86	105	−22.6
13	92	102	−10.4
14	76	100	−31.0
15	95	94	0.7
16	101	94	6.7
17	119	95	19.8
18	87	100	−15.4
19	110	98	11.3
20	102	100	1.8
21	86	101	−17.1
22	105	98	7.1
23	83	99	−19.4
24	75	96	−27.4
25	75	91	−21.1
26	86	87	−1.1
27	99	86	13
28	92	88	4
29	86	89	−3.1
30	77	88	−13.8
31	73	85	−16.2
32	81	82	−0.8
33	74	81	−9.1
34	83	79	5.4
35	84	79	6.3
36	60	79	−31.8
37	85	74	12.8
38	65	76	−16.3
39	64	72	−13.2
40	85	70	18.0
41	73	72	1.4
42	74	71	3.6
43	59	71	−20.5
44	57	68	−18.6
45	74	64	13.1
46	60	65	−9.0
47	57	63	−10.9
48	60	61	−1.4
49	67	60	11.2
50	60	60	−0.1
51	65	59	9.2
52	69	59	14.1
53	49	60	−23.4
54	59	57	3.2
55	68	57	14.3
56	64	58	9.9
57	62	58	6.1
58	53	58	−10
59	51	56	−10.6
60	41	54	−32.8
61		51	
62		50	
63		49	
64		48	
65		47	
66		46	
67		45	
68		44	
69		43	
70		42	
71		41	
72		40	

3.5 การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้งของโฮลท์ (Holt's Two-Parameter Exponential Smoothing)

การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้งของโฮลท์ ค่าสังเกตในอดีตควรจะมีแนวโน้มเป็นเส้นตรง เช่น ข้อมูลจำนวนผู้ป่วยโรคบาดทะยักทั่วประเทศ ที่ $\alpha = 0.1$, $\beta = 0.3$



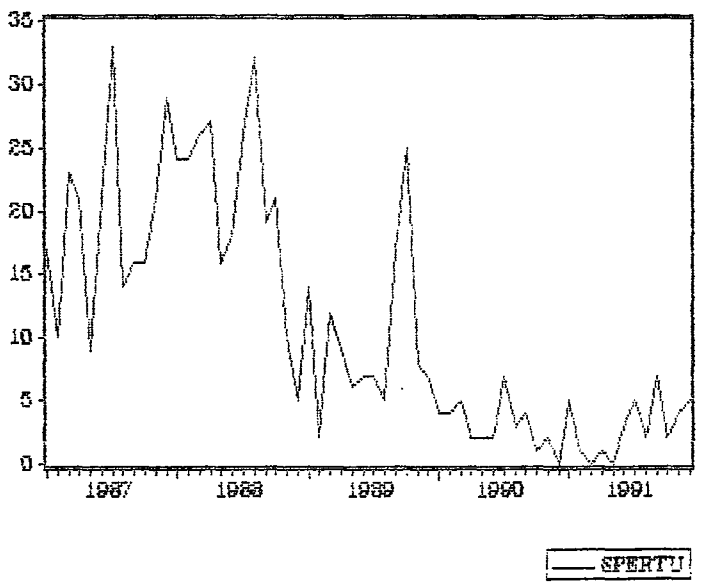
รูปที่ 3.4 แสดงแนวโน้มจำนวนผู้ป่วยโรคบาดทะยักทั่วประเทศ

ตารางที่ 3.4 ค่าข้อมูลดิบ ค่าพยากรณ์ และค่าคลาดเคลื่อนของจำนวนผู้ป่วยโรคบาดทะยักทั่วประเทศ ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้ง ของโฮลท์ ที่ $\alpha = 0.1$, $\beta = 0.3$

Time	Original Data	Forecast	Error
1	114		
2	94	106	−12.4
3	94	103	−9.0
4	96	100	−4.4
5	107	99	7.9
6	115	102	13.3
7	106	106	−0.3
8	101	107	−5.8
9	111	106	5.5
10	100	108	−7.8
11	102	106	−3.8
12	86	105	−18.9
13	92	99	−6.9
14	76	96	−20.3
15	95	89	5.9
16	101	90	11.1
17	119	93	26.4
18	87	101	−13.7
19	110	96	13.6
20	102	101	1.4
21	86	101	−15.2
22	105	96	8.6
23	83	99	−16.0
24	75	94	−18.7
25	75	87	−12.1
26	86	82	3.9
27	99	82	17.0
28	92	86	5.7
29	86	87	−1.4
30	77	86	−9.4
31	73	83	−9.6
32	81	79	2.5
33	74	78	−4.1
34	83	76	7.4
35	84	77	7.2
36	60	78	−18.1
37	85	71	16.7
38	65	74	−9.5
39	64	70	−6.4
40	85	67	18.0
41	73	72	1.5
42	74	71	2.9
43	59	71	−12.2
44	57	66	−9.4
45	74	62	11.8
46	60	65	−4.7
47	57	62	−5.1
48	60	59	0.8
49	67	58	8.9
50	60	60	0.3
51	65	59	6.2
52	69	60	9.2
53	49	62	−12.9
54	59	57	1.9
55	66	57	9.3
56	64	59	5.1
57	62	60	2.1
58	53	60	−7.1
59	51	57	−6.3
60	41	55	−13.6
61		49	
62		48	
63		47	
64		46	
65		44	
66		43	
67		42	
68		41	
69		39	
70		38	
71		37	
72		36	

3.6 การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้ง (Triple Exponential Smoothing)

การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้ง ค่าสังเกตในอดีตควรจะมีแนวโน้มแบบยกกำลังสอง เช่น ข้อมูลจำนวนผู้ป่วยโรคไอกรนภาคใต้ จะทำการวิเคราะห์ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้ง ที่ $\alpha = 0.2$



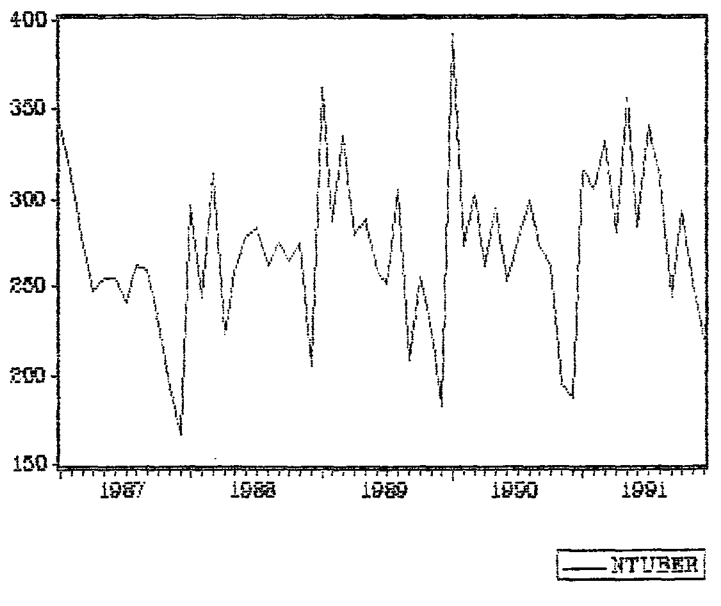
รูปที่ 3.5 แสดงจำนวนผู้ป่วยโรคไอกรนในภาคใต้

ตารางที่ 3.5 ค่าข้อมูลดิบ ค่าพยากรณ์ และค่าคลาดเคลื่อนของจำนวนผู้ป่วยโรคไอกรนในภาคใต้ ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้ง ที่ $\alpha = 0.2$

Time	Original Data	Forecast	Error
1	17		
2	10	18	−80.6
3	23	15	33.1
4	21	21	1.7
5	9	22	−143.4
6	21	15	29.8
7	33	17	48.4
8	14	25	−80.7
9	16	19	−16.1
10	16	15	4.4
11	21	13	37.3
12	29	15	48.7
13	24	21	13.5
14	24	21	11.1
15	26	22	16.9
16	27	23	15.2
17	16	24	−51.5
18	18	18	−2.2
19	26	16	38.3
20	32	20	38.9
21	19	25	−33.8
22	21	21	−1.2
23	10	20	−98.8
24	5	13	−150.4
25	14	5	62.8
26	2	7	−226.1
27	12	1	95.2
28	9	3	62.1
29	6	4	36.2
30	7	3	61.9
31	7	3	58.4
32	5	3	32
33	17	3	83.6
34	25	10	60.5
35	8	19	−139.4
36	7	15	−107.7
37	4	11	−173.5
38	4	7	−72.2
39	5	4	10.7
40	2	4	−86.3
41	2	2	17.6
42	2	1	71.6
43	7	0	98.3
44	3	3	−2.1
45	4	3	34.5
46	1	3	−205.1
47	2	2	20.0
48	0	1	
49	5	0	96.6
50	1	3	−152.5
51	0	2	
52	1	1	45.3
53	0	1	
54	3	0	100.0
55	5	2	69.3
56	2	4	−86.1
57	7	3	53.6
58	2	6	−195.8
59	4	4	−12.1
60	5	5	5.6
61		5	
62		6	
63		6	
64		7	
65		7	
66		8	
67		9	
68		9	
69		10	
70		10	
71		11	
72		11	

3.7 การวิเคราะห์อนุกรมเวลาด้วยวิธีการปรับแนวโน้มและฤดูกาลของวินเตอร์ (Winter's Linear and Seasonal Exponential Smoothing)

วิเคราะห์อนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้ง ข้อมูลที่นำมาวิเคราะห์จะต้องมีความผันแปรฤดูกาล เช่น ข้อมูลจำนวนผู้ป่วยโรควัณโรคในภาคเหนือ มีค่า $\alpha = 0.1$, $\beta = 0.1$, $\gamma = 0.3$ โดยมีความผันแปรฤดูกาลแบบบวก ดังรูป 3.6



รูปที่ 3.6 แสดงจำนวนผู้ป่วยโรควัณโรคในภาคเหนือ

ตารางที่ 3.6 ค่าข้อมูลดิบ ค่าพยากรณ์ และค่าคลาดเคลื่อน จำนวนผู้ป่วยโรควัณโรค ในภาคเหนือ ด้วยวิธีการปรับแนวโน้ม และฤดูกาลของวินเตอร์ ที่ $\alpha = 0.1$, $\beta = 0.1$, $\delta = 0.3$

Time	Original Data	Forecast	Error
1	341	342	−0.2
2	312	312	−0.1
3	274	274	−0.1
4	247	247	−0.1
5	255	255	−0.1
6	255	255	−0.1
7	241	241	−0.1
8	263	263	−0.1
9	260	260	0.0
10	230	230	0.0
11	193	193	0.0
12	166	166	0.0
13	297	353	−19.0
14	244	318	−30.4
15	314	271	13.6
16	223	248	−11.1
17	260	252	3.0
18	278	252	9.4
19	284	240	15.6
20	262	266	−1.4
21	276	262	5.1
22	266	233	12.4
23	276	199	27.7
24	206	181	12.2
25	363	354	2.4
26	288	322	−11.9
27	336	313	6.8
28	280	271	3.1
29	289	290	−0.4
30	260	296	−13.9
31	252	284	−12.8
32	305	291	4.8
33	208	294	−41.5
34	257	261	−1.4
35	225	235	−4.3
36	183	192	−5.1
37	393	362	7.8
38	275	320	−16.2
39	304	323	−6.4
40	262	272	−3.7
41	295	284	3.7
42	253	280	−10.6
43	277	268	3.3
44	299	290	3.2
45	275	264	4.0
46	263	262	0.4
47	195	234	−20.2
48	188	188	−0.2
49	317	367	−15.9
50	305	295	3.4
51	332	310	6.7
52	282	264	6.2
53	356	285	20.0
54	284	277	2.6
55	341	279	18.3
56	312	307	1.6
57	245	283	−15.4
58	294	274	6.9
59	252	238	5.6
60	221	209	5.3
61		389	
62		324	
63		340	
64		294	
65		314	
66		300	
67		302	
68		324	
69		299	
70		295	
71		257	
72		227	

3.8 การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีบอกซ์และเจนกินส์

(Box and Jenkins Method)

1.ค้นหาตัวแบบที่เหมาะสม

- ทำให้อนุกรมเวลาอยู่ในสภาวะสมดุลย์
- ค้นหาตัวแบบโดยพิจารณาจากกราฟของ ACF และ PACF ว่ามีลักษณะอย่างไร

2.ประมาณค่า Parameter ของตัวแบบ

3.การตรวจสอบตัวแบบโดยใช้ค่าสถิติบอกซ์-เพียร์ซไคสแควร์ และพิจารณาค่าความคลาดเคลื่อน

4.การพยากรณ์ เมื่อได้ตัวแบบที่เหมาะสมจึงทำการพยากรณ์

```
Fourth Root of CPERTU - BOX-JENKINS ANALYSIS: IDENTIFICATION

                    AUTOCORRELATION FUNCTION

                  60 Observations in the Working Series
                  Mean of the Working Series = 1.87
                  Standard Deviation of the Working Series = .3624

Lag  Value   T-Value -1.0                     0.0                      +1.0
                     :---------------------------------------------------:
  1   0.74    5.69   :                   [     |*****]************       :
  2   0.72    3.89   :                [        |*********]*********      :
  3   0.68    2.97   :              [          |**********]******        :
  4   0.72    2.79   :            [            |************]*****       :
  5   0.67    2.30   :          [              |**************]**        :
  6   0.63    2.00   :         [               |***************]         :
  7   0.58    1.73   :        [                |*************** ]        :
  8   0.56    1.58   :       [                 |**************   ]       :
  9   0.58    1.58   :       [                 |***************  ]       :
 10   0.58    1.51   :      [                  |**************    ]      :
 11   0.54    1.36   :     [                   |*************      ]     :
 12   0.53    1.29   :     [                   |*************      ]     :
 13   0.60    1.44   :    [                    |***************     ]    :
 14   0.56    1.28   :   [                     |**************       ]   :
 15   0.50    1.13   :   [                     |*************        ]   :
                     :---------------------------------------------------:

                        [ ] = Estimated Two-Standard Error Limits

     Box-Pierce Chi-Square Statistic with 15 Degrees of Freedom = 343.2
```

รูปที่ 3.7 แสดงค่าสหสัมพันธ์ในตัวเองของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง

```
Fourth Root of CPERTU - BOX-JENKINS ANALYSIS: IDENTIFICATION

                   PARTIAL AUTOCORRELATION FUNCTION

                   60 Observations in the Working Series
                   Mean of the Working Series = 1.87
                   Standard Deviation of the Working Series = .3624

Lag  Value  T-Value -1.0                 0.0                  +1.0
                    :--------------------------------------------------:
  1   0.74   5.69   :                  [     |*****]*************      :
  2   0.40   3.10   :                  [     |*****]****               :
  3   0.17   1.32   :                  [     |****  ]                  :
  4   0.30   2.29   :                  [     |*****]*                  :
  5   0.06   0.43   :                  [     |**    ]                  :
  6  -0.03  -0.25   :                  [    *|      ]                  :
  7  -0.07  -0.56   :                  [   **|      ]                  :
  8  -0.06  -0.49   :                  [   **|      ]                  :
  9   0.11   0.85   :                  [     |***   ]                  :
 10   0.12   0.89   :                  [     |***   ]                  :
 11   0.01   0.11   :                  [     |      ]                  :
 12   0.05   0.40   :                  [     |*     ]                  :
 13   0.26   2.03   :                  [     |*****]*                  :
 14  -0.05  -0.40   :                  [    *|      ]                  :
 15  -0.19  -1.44   :                  [*****|      ]                  :
                    :--------------------------------------------------:

                        [ ] = Estimated Two-Standard Error Limits
```

รูปที่ 3.8 แสดงค่าสหสัมพันธ์ในตัวเองบางส่วนของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง

จากรูป 3.8 จะเห็นว่าอนุกรมเวลาอยู่ในสภาวะสเตชันนารีแล้ว จึงทำการค้นหาตัวแบบ พบว่า ตัวแบบที่ดีที่สุดในการพยากรณ์คือ AR(2) ซึ่งมีค่า Box-Pierce Chi-square = 17.4, ค่า Probability = 0.133, ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (MSE) และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ (MAPE) ต่ำสุด คือ 0.0465 และ 45.4197 ตามลำดับ

มีสมการพยากรณ์จำนวนผู้ป่วยโรคไอกรนในภาคกลาง คือ

$$z_t = 0.2563 + 0.4379z_{t-1} + 0.4116z_{t-2} + e_t$$

เมื่อ z_t คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง ณ คาบเวลาที่ t

ตารางที่ 3.7 ค่าข้อมูลดิบ ค่าพยากรณ์ และค่าจำนวนผู้ป่วยโรคไอกรนในภาคกลาง ด้วยวิธีการวิเคราะห์ แบบบอกซ์และเจนกินส์

Time	Original Data	Forecast	Error
1	42		
2	26		
3	30		
4	23	23.9	−0.9
5	16	22.5	−6.5
6	16	17.1	−1.1
7	23	14.6	8.4
8	17	17.3	−0.3
9	18	17.6	0.4
10	20	15.8	4.2
11	20	17.0	3.0
12	13	17.8	−4.8
13	24	14.7	9.3
14	22	16.2	5.8
15	19	20.0	−1.0
16	31	18.1	12.9
17	27	21.3	5.7
18	29	24.5	4.5
19	22	23.9	−1.9
20	22	21.8	0.2
21	21	19.3	1.7
22	34	18.9	15.1
23	16	23.2	−7.2
24	7	20.4	−13.4
25	11	10.3	0.7
26	25	8.8	16.2
27	31	15.5	15.5
28	12	23.9	−11.9
29	24	17.4	6.6
30	23	15.7	7.3
31	22	20.4	1.6
32	26	19.7	6.3
33	13	20.8	−7.8
34	13	16.6	−3.6
35	10	12.2	−2.2
36	12	10.9	1.1
37	7	10.6	−3.6
38	5	9.1	−4.1
39	5	6.2	−1.2
40	14	5.4	8.6
41	2	8.7	−6.7
42	8	6.1	1.9
43	3	4.8	−1.8
44	8	5.4	2.6
45	6	5.5	0.5
46	8	7.1	0.9
47	4	7.2	−3.2
48	3	6.0	−3.0
49	6	4.0	2.0
50	8	4.8	3.2
51	7	7.2	−0.2
52	4	7.6	−3.6
53	4	5.7	−1.7
54	3	4.5	−1.5
55	3	4.0	−1.0
56	3	3.60	−0.6
57	4	3.6	0.4
58	10	4.0	6
59	2	6.8	−4.8
60	5	5.2	−0.2
61		3.8	
62		4.9	
63		4.8	
64		5.3	
65		5.5	
66		5.8	
67		6.0	
68		6.2	
69		6.4	
70		6.6	
71		6.8	
72		6.9	

บทที่ 4

ผลการวิเคราะห์

1. วัณโรค

1.1 วัณโรคทั่วประเทศ

จากการเปรียบเทียบวิธีพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรควัณโรคทั่วประเทศ พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ น้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ARIMA (1,1,0) * $(0,1,1)_{12}$ แปลงข้อมูลด้วยล็อกการิทึม (รายละเอียดอยู่ในตาราง ที่ 4.1)

$$z_t = \delta + \phi_1 z_{t-1} - \Theta_{1,12} e_{t-12} + e_t$$

โดย $Z_t = y_t - y_{t-1} - y_{t-12} - y_{t-13}$

$x_t = \exp(y_t)$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

$\Theta_{1,12}$ คือ พารามิเตอร์การเฉลี่ยเคลื่อนที่แบบมีฤดูกาล (Seasonal Moving Average parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ จำนวนผู้ป่วยโรควัณโรคทั่วประเทศ ณ คาบเวลาที่ t

y_t คือ จำนวนผู้ป่วยที่ผ่านการแปลงข้อมูลด้วยล็อกการิทึม (logarithm) เพื่อปรับความแปรปรวนให้คงที่

x_t คือ จำนวนผู้ป่วยที่แท้จริง

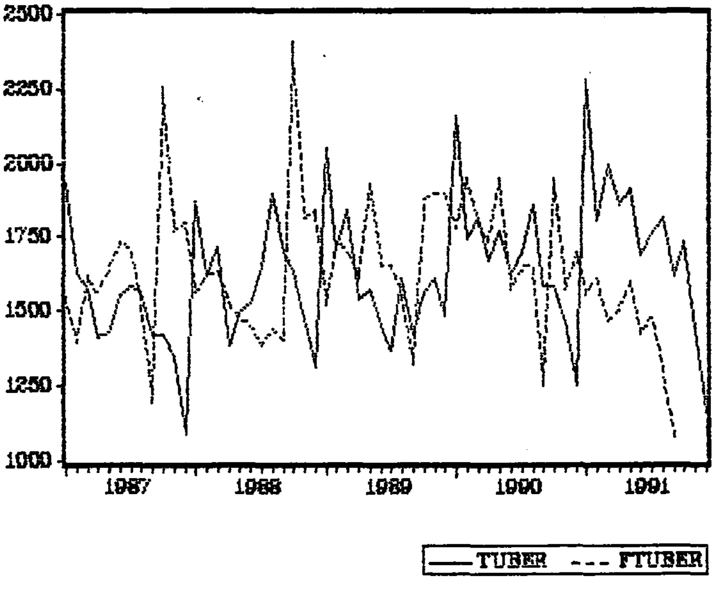
$$z_t = 0.0562z_{t-1} - 0.38e_{t-12} + e_t$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรควัณโรคทั่วประเทศ ณ คาบเวลาที่ t คือ

$$z_t = 0.0562z_{t-1} - 0.38e_{t-12}$$

ตารางที่ 4.1 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรควัณโรคทั่วประเทศ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Winter (ฤดูกาลแบบบวก)	37.0	0.012	14034.0	5.30	ค่า $\alpha = 0.7$, $\beta = 0.1$ $\gamma = 0.9$
Winter (ฤดูกาลแบบคูณ)	70.6	0	16689.0	5.80	ค่า $\alpha = 0.07$, $\beta = 0.1$ $\gamma = 0.3$
ARIMA (1,1,0) * $(0,1,1)_{12}$	20.4	0.085	0.00585	6.48	แปลงด้วย ล็อก (log)



——— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.1 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรควัณโรคทั่วประเทศ

1.2 วัณโรคในกรุงเทพมหานคร

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรควัณโรคในกรุงเทพมหานคร พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) (ดูรายละเอียดในตารางที่ 4.2)

$$z_t = \delta + \phi_1 z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ จำนวนผู้ป่วยโรคโปลิโอในกรุงเทพมหานคร ณ คาบเวลาที่ t

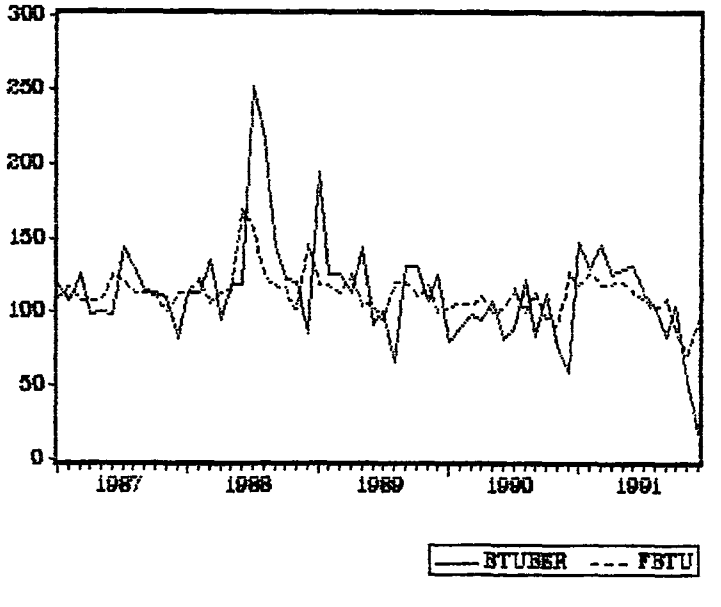
$$z_t = 64.8851 + 0.4226z_{t-1} + e_t$$

ในกรณีนี้ไม่ต้องแปลงข้อมูลเพื่อปรับความแปรปรวนและค่าเฉลี่ยให้คงที่ ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในกรุงเทพมหานคร ณ คาบเวลาที่ t คือ

$$z_t = 64.8851 + 0.4226z_{t-1}$$

<u>ตารางที่ 4.2</u> เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรควัณโรคในกรุงเทพมหานคร

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Average	22.1	0.230	1348.8	29.09	จำนวนเทอมที่ใช้ 3 เทอม
Single Exp.	10.6	0.937	1138.8	27.74	ค่า α = 0.37
Double Exp. (Brown)	11.5	0.905	1209.8	29.73	ค่า α = 0.11
Double Exp. (Holt)	11.3	0.913	1215.5	28.07	ค่า α = 0.4 , β = 0.1
Triple Exp. (Brown)	11.1	0.921	1302.5	29.58	ค่า α = 0.14
AR(1)	8.6	0.803	1047.7	25.66	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.2 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรควัณโรคในกรุงเทพมหานคร

1.3 วัณโรคในภาคกลาง

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรควัณโรคในภาคกลาง พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองค่าคลาดเคลื่อนและค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือวิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ARIMA(0,0,1) * $(0,1,1)_{12}$ แปลงข้อมูลด้วยล็อกการิทึม (ดูรายละเอียดในตารางที่ 4.3)

$$z_t = \delta + e_t - \theta_1 e_{t-1} - \Theta_{1,12} e_{t-12} + \theta_1 \Theta_{1,12} e_{t-13}$$

โดย

$$z_t = y_t - y_{t-12}$$

$$x_t = \exp(y_t)$$

เมื่อ δ คือ ค่าคงที่

θ_1 คือ พารามิเตอร์การเฉลี่ยเคลื่อนที่ (Moving Average parameter)

$\Theta_{1,12}$ คือ พารามิเตอร์การเฉลี่ยเคลื่อนที่แบบมีฤดูกาล (Seasonal Moving Average parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ จำนวนผู้ป่วยโรควัณโรคในภาคกลาง ณ คาบเวลาที่ t

y_t คือ จำนวนผู้ป่วยที่ผ่านการแปลงข้อมูลด้วยล็อกการิทึม (logarithm) เพื่อปรับความแปรปรวนให้คงที่

x_t คือ จำนวนผู้ป่วยที่แท้จริง

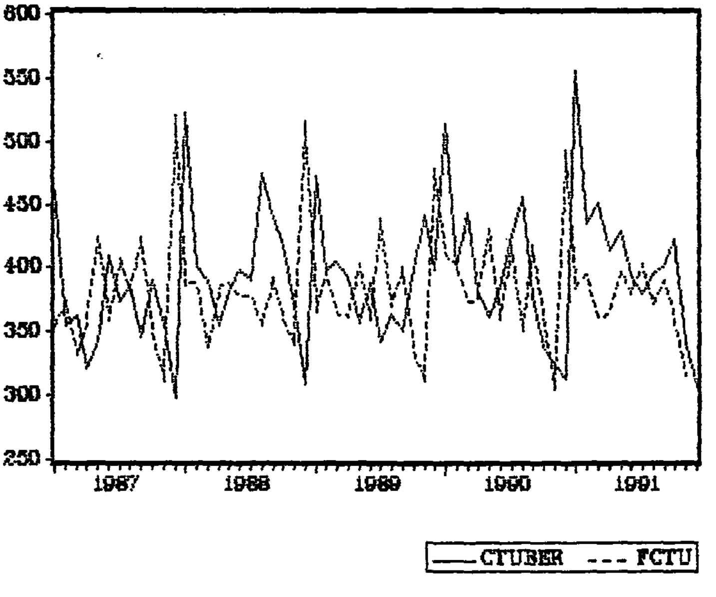
$$z_t = 0.6138e_{t-1} - 0.8176e_{t-12} + 0.5018e_{t-13} + e_t$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในภาคกลาง ณ คาบเวลา t คือ

$$z_t = 0.6138e_{t-1} - 0.8176e_{t-12} + 0.5018e_{t-13}$$

<u>ตารางที่ 4.3</u> เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรควัณโรคในภาคกลาง

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Winter (ฤดูกาลแบบบวก)	69.9	0	1356.0	7.00	ค่า α = 0.1 , β = 0.1 γ = 0.1
Winter (ฤดูกาลแบบคูณ)	58.1	0	1473.0	7.20	ค่า α = 0.1 , β = 0.1 γ = 0.1
ARIMA (0,0,1) * $(0,1,1)_{12}$	21.7	0.06	0.0067	7.17	แปลงด้วย ล็อก (log)



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.3 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรควัณโรคในภาคกลาง

1.4 วัณโรคในภาคเหนือ

จากตารางเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรควัณโรคในภาคเหนือ พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ARIMA(0,1,1)$_{12}$ แปลงข้อมูลด้วยรากที่สอง (ดูรายละเอียดในตารางที่ 4.4)

$$z_t = \delta + e_t - \Theta_{1,12} e_{t-12}$$

โดย

$$z_t = y_t - y_{t-12}$$
$$x_t = (y_t)^2$$

เมื่อ δ คือ ค่าคงที่

$\Theta_{1,12}$ คือ พารามิเตอร์การเฉลี่ยเคลื่อนที่ (Moving Average parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ จำนวนผู้ป่วยโรควัณโรคในภาคเหนือ ณ คาบเวลาที่ t

y_t คือ จำนวนผู้ป่วยที่ผ่านการแปลงข้อมูลด้วยรากที่สองเพื่อปรับความแปรปรวนให้คงที่

x_t คือ จำนวนผู้ป่วยที่แท้จริง

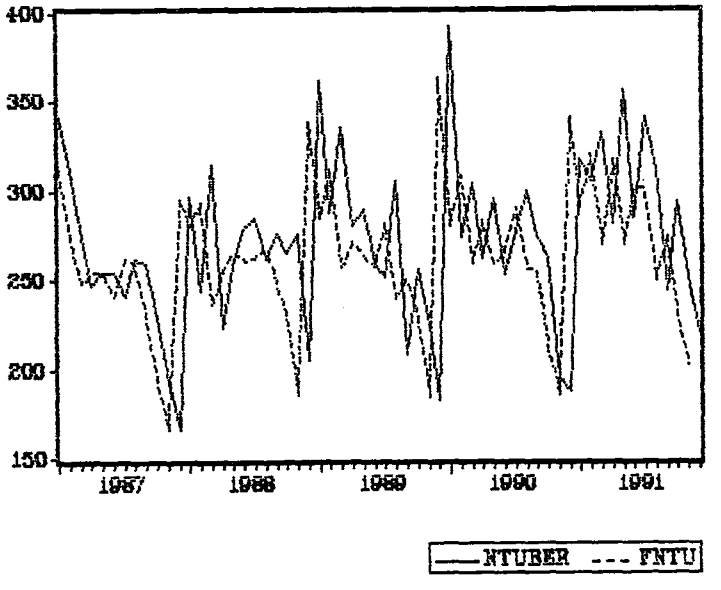
$$z_t = -0.5234e_{t-12} + e_t$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในภาคเหนือ ณ คาบเวลา t คือ

$$z_t = -0.5234e_{t-12}$$

ตารางที่ 4.4 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรควัณโรคในภาคเหนือ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Winter (ฤดูกาลแบบบวก)	18.4	0.561	906.0	7.60	ค่า α = 0.1 , β = 0.1 δ = 0.3
Winter (ฤดูกาลแบบคูณ)	17.9	0.594	953.0	7.90	ค่า α = 0.1 , β = 0.1 δ = 0.4
$ARIMA(0,1,1)_{12}$	17.0	0.254	0.9231	10.46	แปลงด้วยรากที่สอง



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.4 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรควัณโรคในภาคเหนือ

1.5 วัณโรคในภาคตะวันออกเฉียงเหนือ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรควัณโรคในภาคตะวันออกเฉียงเหนือ พบว่า วิธีพยากรณ์ ที่มีค่าผลบวกกำลังสองค่าคลาดเคลื่อนและค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ น้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ARIMA(0,1,1) * (0,1,1)$_{12}$ แปลงข้อมูลด้วยล็อกการิทึม (ดูรายละเอียดในตาราง 4.5)

$$z_t = \delta + e_t - \theta_1 e_{t-1} - \Theta_{1,12} e_{t-12} + \theta_1 \Theta_{1,12} e_{t-13}$$

โดย

$$z_t = y_t - y_{t-1} - y_{t-12} + y_{t-13}$$

$$x_t = \exp(y_t)$$

เมื่อ δ คือ ค่าคงที่

θ_1 คือ พารามิเตอร์การเฉลี่ยเคลื่อนที่แบบไม่มีฤดูกาล (Moving Average parameter)

$\Theta_{1,12}$ คือ พารามิเตอร์การเฉลี่ยเคลื่อนที่แบบมีฤดูกาล (Seasonal Moving Average parameter)

e_t คือ การรบกวนสุ่ม (Random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ จำนวนผู้ป่วยโรควัณโรคในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาที่ t

y_t คือ จำนวนผู้ป่วยที่ผ่านการแปลงข้อมูลด้วยล็อกการิทึมเพื่อปรับความแปรปรวนให้คงที่

x_t คือ จำนวนผู้ป่วยที่แท้จริง

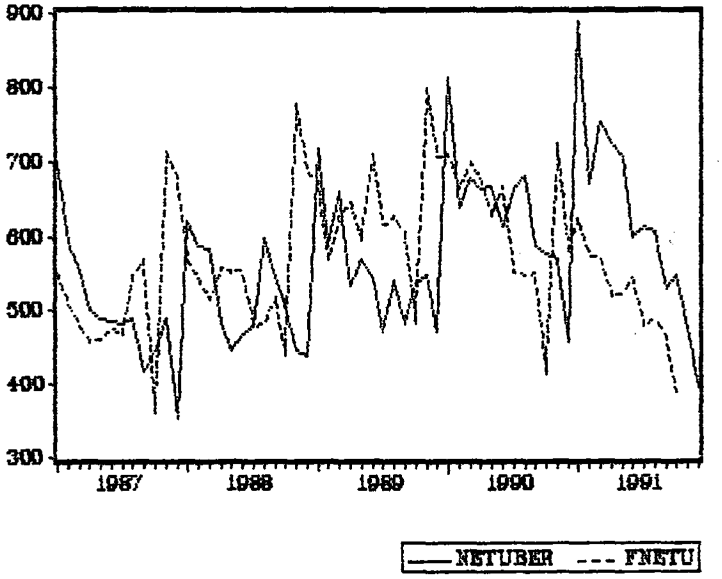
$$z_t = -0.4039e_{t-1} - 0.5746e_{t-12} + 0.2321e_{t-13} + e_t$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในภาคตะวันออกเฉียงเหนือ ณ คาบเวลา t คือ

$$z_t = -0.4039e_{t-1} - 0.5746e_{t-12} + 0.2321e_{t-13}$$

ตารางที่ 4.5 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรควัณโรคในภาคตะวันออกเฉียงเหนือ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Winter (ฤดูกาลแบบบวก)	15.6	0.741	2332.0	6.40	ค่า $\alpha = 0.4$, $\beta = 0.1$ $\delta = 0.4$
Winter (ฤดูกาลแบบคูณ)	13.5	0.854	2415.0	6.50	ค่า $\alpha = 0.2$, $\beta = 0.1$ $\delta = 0.4$
ARIMA(0,1,1) * $(0,1,1)_{12}$	12.4	0.494	1.2457	7.85	แปลงด้วยล็อก (log)



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.5 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์ และเจนกินส์ ของโรควัณโรคในภาคตะวันออกเฉียงเหนือ

1.6 วัณโรคในภาคใต้

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรควัณโรคในภาคใต้ พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือวิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสมคือ ตัวแบบการถดถอยในตนเองอันดับที่ 1 (First Order Autoregressive model) (ดูรายละเอียดในตารางที่ 4.6)

$$z_t = \delta + \phi_1 z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (Random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ จำนวนผู้ป่วยโรคโปลิโอในภาคเหนือ ณ คาบเวลาที่ t

ในกรณีนี้ไม่ต้องแปลงข้อมูลเพื่อปรับความแปรปรวนและค่าเฉลี่ยให้คงที่

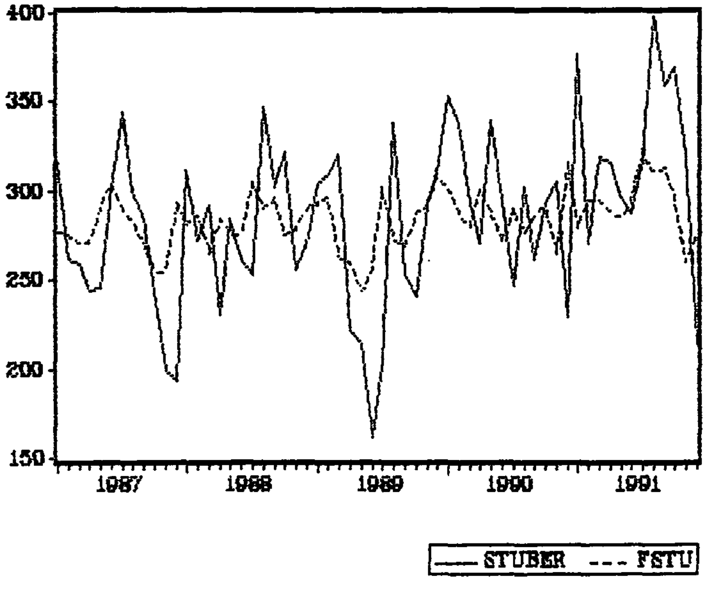
$$z_t = 188.4071 + 0.3377z_{t-1} + e_t$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในภาคใต้ ณ คาบเวลาที่ t คือ

$$z_t = 188.4071 + 0.3377z_{t-1}$$

ตารางที่ 4.6 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรควัณโรคในภาคใต้

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Average	27.5	0.070	2901.2	15.58	จำนวนเทอมที่ใช้ 3 เทอม
Single Exp.	28.1	0.082	2309.8	14.39	ค่า α = 0.08
Double Exp. (Brown)	29.5	0.059	2219.0	14.27	ค่า α = 0.01
Double Exp. (Holt)	27.1	0.102	2371.4	14.53	ค่า α = 0.1 , β = 0.1
Triple Exp. (Brown)	20.8	0.347	2740.1	15.22	ค่า α = 0.1
AR(1)	19.5	0.109	2100.4	13.90	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.6 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรควัณโรคในภาคใต้

2.คอตีบ

2.1 คอตีบทั่วประเทศ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคคอตีบทั่วประเทศพบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองของค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดคือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยตัวเองอันดับที่ 1 (First Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สี่ (ดูรายละเอียดในตารางที่ 4.7)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม(random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคคอตีบทั่วประเทศ ณ คาบเวลาที่ t

$$Z_t = 0.2957 + 0.8243\ Z_{t-1} + e_t$$

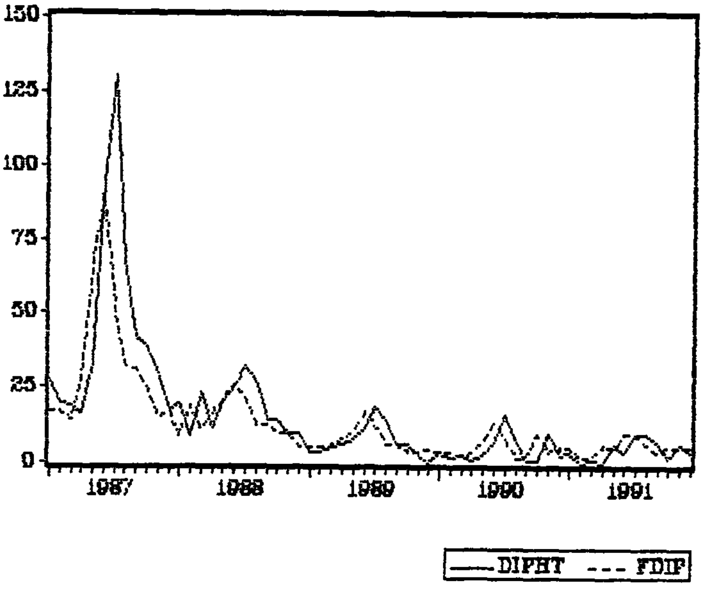
ดังนั้น สมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคคอตีบทั่วประเทศ ณ คาบเวลา t คือ

$$Z_t = 0.2957 + 0.8243\ Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสี่จึงจะได้จำนวนผู้ป่วยโรคคอตีบทั่วประเทศอย่างแท้จริง

ตารางที่ 4.7 เปรียบเทียบวิธีพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคคอตีบทั่วประเทศ

เทคนิคการวิเคราะห์	Box-Pierce Chi-Square	P-value	MSE	MAPE	หมายเหตุ
Winter's Additive	78.7	0	661	236.1	ค่า $\alpha=1.0, \beta=0.1, \delta=0.1$
AR(1)	19.5	0.109	0.08	396.1	แปลงด้วย รากที่สี่



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.7 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคคอตีบทั่วประเทศ

2.2 คอตีบในกรุงเทพมหานคร

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วย วิธีต่าง ๆ ของโรคคอตีบในกรุงเทพมหานครพบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองของค่าคลาดเคลื่อนน้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สี่ (ดูรายละเอียดในตารางที่ 4.8)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม(random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคคอตีบในกรุงเทพมหานคร ณ คาบเวลาที่ t

$$Z_t = 0.5911Z_{t-1} + e_t$$

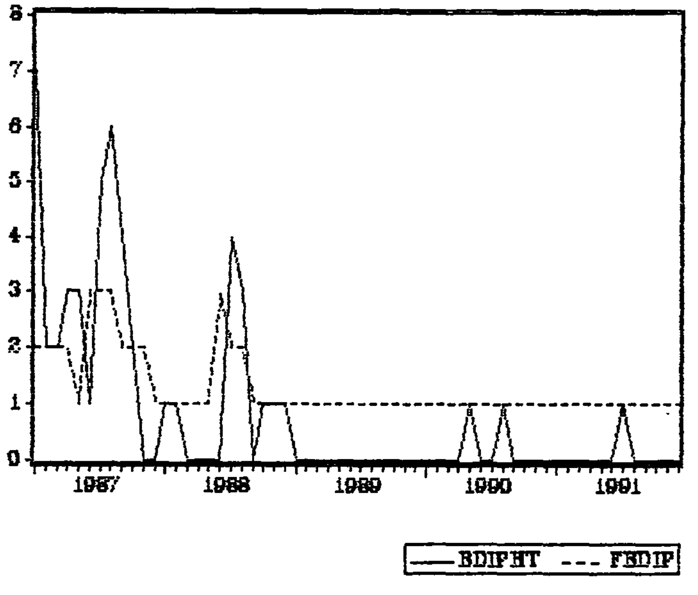
ดังนั้น สมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคคอตีบในภาคกลาง ณ คาบเวลา t คือ

$$Z_t = 0.5911\ Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสี่ และบวกกับค่าพยากรณ์ เมื่อ 12 คาบก่อนหน้านี้ จึงจะได้จำนวนผู้ป่วยโรคคอตีบในกรุงเทพมหานครอย่างแท้จริง

ตารางที่ 4.8 เปรียบเทียบวิธีพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคคอตีบใน
กรุงเทพมหานคร

เทคนิคการวิเคราะห์	Box-Pierce Chi-Square	P-value	MSE	MAPE	หมายเหตุ
Winter's Additive	25.6	0.181	4	88.6	α =0.1, β=0.2, δ=0.8
AR(1)	11.7	0.631	0.008	98.5	แปลงด้วยรากที่สี่



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.8 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์
และเจนกินส์ ของโรคคอตีบในกรุงเทพมหานคร

2.3 คอตีบในภาคกลาง

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคคอตีบในภาคกลาง พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองของค่าคลาดเคลื่อนน้อยที่สุดคือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สี่ (ดูรายละเอียดในตารางที่ 4.9)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม(random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคคอตีบในภาคกลาง ณ คาบเวลาที่ t

$$Z_t = 0.4499 + 0.6465\ Z_{t-1} + e_t$$

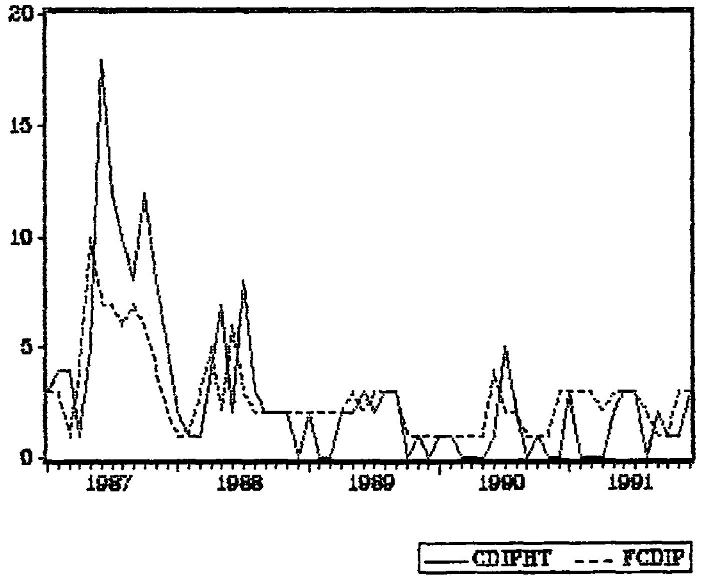
ดังนั้น สมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคคอตีบในภาคกลาง ณ คาบเวลา t คือ

$$Z_t = 0.4499 + 0.6465\ Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้วต้องนำมาทำการยกกำลังสี่จึงจะได้จำนวนผู้ป่วยโรคคอตีบในภาคกลางอย่างแท้จริง

ตารางที่ 4.9 เปรียบเทียบวิธีพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคคอตีบในภาคกลาง

เทคนิคการวิเคราะห์	Box-Pierce Chi-Square	P-value	MSE	MAPE	หมายเหตุ
Winter's Additive	18.4	0.559	12	93.7	ค่า α=0.1, β=0.2, γ=0.9
AR(1)	19.9	0.097	0.037	271.4	**แปลงด้วย รากที่สี่**



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.9 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคคอตีบในภาคกลาง

2.4 คอตีบในภาคเหนือ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคคอตีบในภาคเหนือ พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองของค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสมบูรณ์น้อยที่สุด รวมทั้งค่าความน่าจะเป็นในการยอมรับสูงสุดจากตัวแบบที่ใช้ได้ คือ ตัวแบบการปรับเรียบแบบเอกซโปเนนเชียลครั้งเดียว โดยใช้วิธีของบราวน์ (ดูรายละเอียดในตารางที่ 4.10)

$$X_t = \beta_t + e_t$$

เมื่อ X_t คือ ค่าสังเกต หรือข้อมูลอนุกรมเวลา

β_t คือ ค่าคงที่ทำให้เรียบ

e_t คือ ความคลาดเคลื่อนสุ่ม (error)

$$S_t = \alpha X_t + (1-\alpha)S_{t-1}$$

เมื่อ α คือ ค่าประมาณของ β_t

S_t คือ ค่าประมาณทำให้เรียบ เป็นค่าที่กำหนดน้ำหนักของการเฉลี่ย

โดยปกติจะกำหนดให้ $S_0 = X_0$

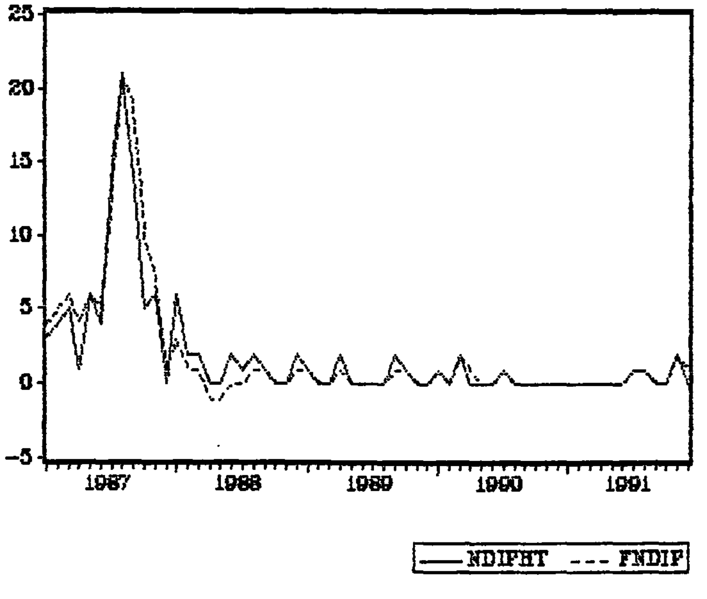
$$F_{t+1} = S_t$$

ดังนั้น สมการพยากรณ์จำนวนผู้ป่วยโรคคอตีบในภาคเหนือ ณ คาบเวลา t คือ

$$S_t = 0.88\ X_t + 0.22\ S_{t-1}$$

ตารางที่ 4.10 เปรียบเทียบวิธีพยากรณ์จำนวนผู้ป่วยวิธีต่างๆของโรคคอตีบในภาคเหนือ

เทคนิคการวิเคราะห์	Box-Pierce Chi-Square	P-value	MSE	MAPE	หมายเหตุ
Simple moving avg	57.4	0	10.9	79.34	N=3
Single Exp.	**30.7**	**0.057**	**7.9**	**41.01**	**α= 0.88**
Double Exp.(Brown)	31.4	0.036	9.9	43.51	α= 0.39
Double Exp.(Holt)	30.5	0.045	8.8	45.77	α = 0.9, β = 0.1



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.10 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียวของบราวน์ α =0.88 ของโรคคอตีบในภาคเหนือ

2.5 คอตีบในภาคตะวันออกเฉียงเหนือ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคคอตีบในภาคตะวันออกเฉียงเหนือพบว่า วิธีพยากรณ์ปรับแนวโน้มและฤดูกาลของวินเตอร์ เป็นวิธีที่สามารถใช้ได้เพียงวิธีเดียวในทุกวิธีที่ได้ทดสอบมา ซึ่งประกอบด้วยสมการปรับเรียบ 3 สมการ คือ สมการปรับเรียบโดยส่วนรวม สมการปรับสำหรับแนวโน้ม และสมการปรับสำหรับฤดูกาล (ดูรายละเอียดในตารางที่ 4.11) โดยมีตัวแบบคือ

$$X_t = (\beta_0 + \beta_1 t)SN_t + e_t.$$

เมื่อ X_t คือ ค่าสังเกต ณ เวลา t

β_0 คือ พารามิเตอร์การปรับระดับ

β_1 คือ พารามิเตอร์การปรับแนวโน้ม

SN_t คือ ปัจจัยฤดูกาล

ซึ่งมีตัวประมาณค่า 3 ตัว กำหนดโดยสมการ ดังนี้

$$S_t = \alpha \frac{X_t}{I_{t-L}} + (1-\alpha)(S_{t-1}+b_{t-1})$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1}$$

$$I_t = \beta \frac{X_t}{S_t} + (1-\beta)I_{t-L}$$

เมื่อ S_t คือ ตัวประมาณค่าของ β_0

b_t คือ ตัวประมาณค่าของ β_1

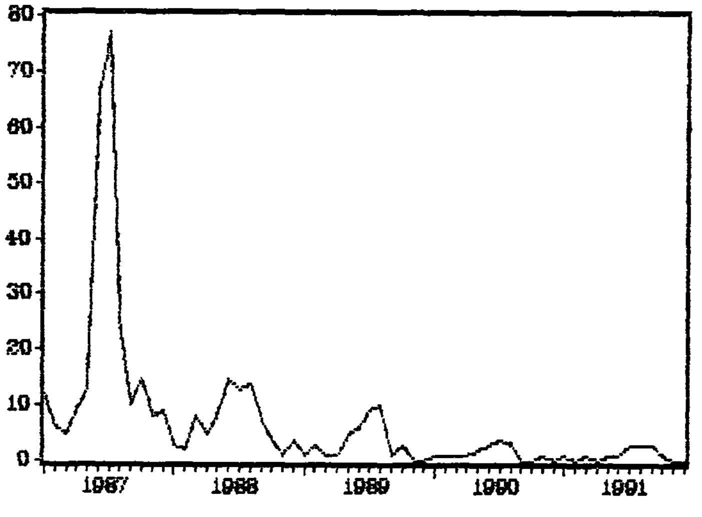
I_t คือ ตัวประมาณค่าของ SN_t

ดังนั้น สมการพยากรณ์จำนวนผู้ป่วยโรคคอตีบในภาคตะวันออกเฉียงเหนือ ณ คาบเวลา t+m คือ

$$F_{t+m} = (S_t + b_t m) I_{t-L+m}$$

ตารางที่ 4.11 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคคอตีบในภาคตะวันออกเฉียงเหนือ

เทคนิคการวิเคราะห์	Box-Pierce Chi-Square	P-value	MSE	MAPE	หมายเหตุ
Winter's Additive	**34.2**	**0.064**	**127**	**259.7**	**α =0.1, β=0.1, δ=1.0**



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.11 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีปรับแนวโน้มและฤดูกาลของวินเทอร์ α=0.1, β=0.1, δ=1.0 ของโรคคอตีบในภาคตะวันออกเฉียงเหนือ

2.6 คอตีบในภาคใต้

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคคอตีบในภาคใต้ พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองของค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 แบบมีฤดูกาล (First Order Seasonal Autoregressive Model) (ดูรายละเอียดในตารางที่ 4.12)

$$Z_t = \phi_1 Z_{t-8} + e_t$$

เมื่อ $\phi_{1,8}$ คือ พารามิเตอร์ถดถอยในตัวเองแบบมีฤดูกาล (Seasonal Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (Random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors) ณ คาบเวลา t

$$Z_t = 0.4893Z_{t-8} + e_t$$

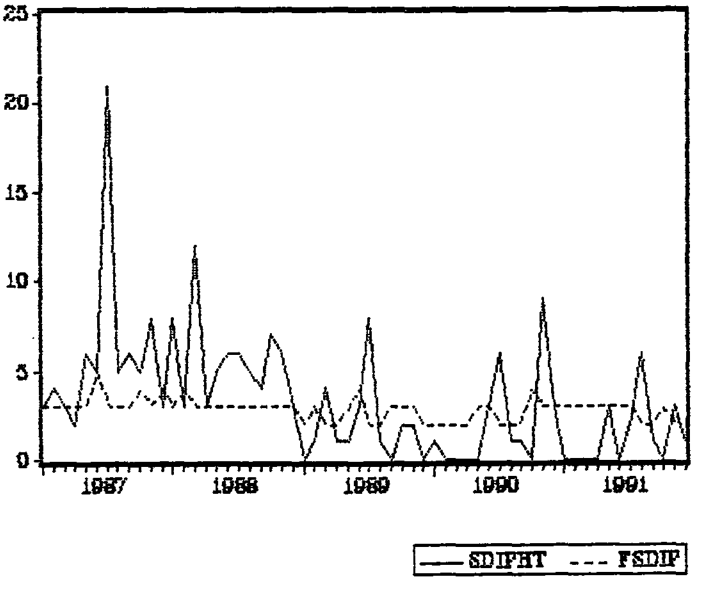
ดังนั้น สมการพยากรณ์จำนวนผู้ป่วยโรคคอตีบในภาคใต้ ณ คาบเวลา t คือ

$$Z_t = 0.4893Z_{t-8}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องบวกกับค่าพยากรณ์ เมื่อ 8 คาบก่อนหน้านี้ จึงจะได้จำนวนผู้ป่วยโรคคอตีบในภาคใต้อย่างแท้จริง

ตารางที่ 4.12 เปรียบเทียบวิธีพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคคอตีบในภาคใต้

เทคนิคการวิเคราะห์	Box-Pierce Chi-Square	P-value	MSE	MAPE	หมายเหตุ
Winter's Additive	16.2	0.704	15	67.1	α =0.1, β=0.1, δ=1.0
SMA(1)	9.3	0.815	10.09	350.01	แปลงด้วยรากที่สี่
SAR(1)	16.3	0.295	7.79	309.40	แปลงด้วยรากที่สี่



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.12 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคคอตีบในภาคใต้

3.ไอกรน

3.1 ไอกรนทั้งประเทศ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคไอกรนทั้งประเทศ พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสมคือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สอง (ดูรายละเอียดในตารางที่ 4.13)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_t คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนทั้งประเทศ ณ คาบเวลาที่ t

$$Z_t = 1.3172 + 0.8289\ Z_{t-1} + e_t$$

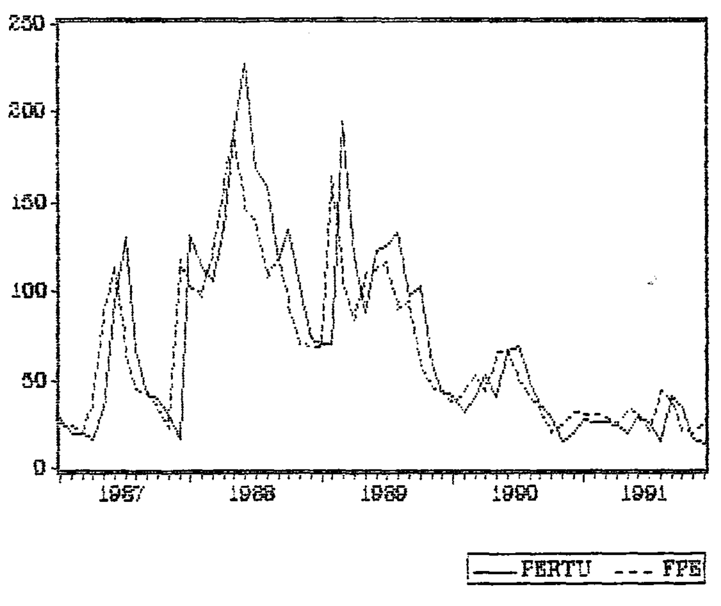
ดังนั้นสมการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนทั้งประเทศ ณ คาบเวลา t คือ

$$Z_t = 1.3172 + 0.8289\ Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสองจึงจะได้จำนวนผู้ป่วยโรคไอกรนทั้งประเทศที่แท้จริง

ตารางที่ 4.13 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคไอกรนทั่วประเทศ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	54	0	1458.4	44.9044	N= 3
Single Exp.	20.4	0.374	1110.7	34.6872	ค่า α = 0.83
Double Exp(Brown)	37.7	0.006	1349.1	42.8276	ค่า α = 0.35
Double Exp(Holt)	22	0.286	1197.3	37.0783	ค่า α = 0.8, β=0.1
Triple Exp(Brown)	33.3	0.022	1601.3	46.6151	ค่า α = 0.26
AR(1)	12.5	0.49	3.088	30.9873	แปลงข้อมูลด้วยรากที่2



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.13 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคไอกรนทั่วประเทศ

3.2 ไอกรนในกรุงเทพมหานคร

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคไอกรนในกรุงเทพมหานครพบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive Model) โดยทำการแปลงข้อมูลด้วยรากที่สี่ (ดูรายละเอียดในตารางที่ 4.14)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในกรุงเทพมหานคร ณ คาบเวลาที่ t-1

$$Z_t = 0.8356 + 0.4801\ Z_{t-1} + e_t$$

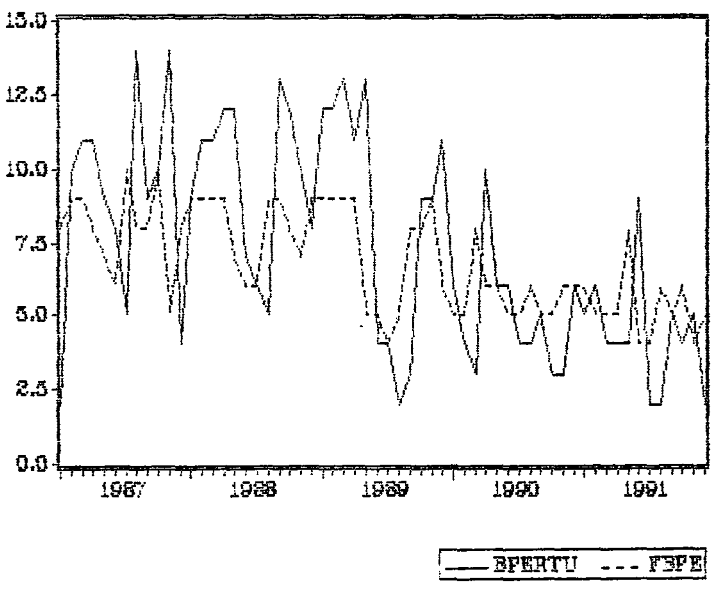
ดังนั้นสมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในกรุงเทพมหานคร ณ คาบเวลา t คือ

$$Z_t = 0.8356 + 0.4801\ Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสี่จึงจะได้จำนวนผู้ป่วยโรคไอกรนในกรุงเทพมหานครที่แท้จริง

ตารางที่ 4.14 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคไอกรนในกรุงเทพมหานคร

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	64.7	0	11.7	52.2677	N= 3
Single Exp.	72.9	0	14.0	75.2032	ค่า α = 0.01
Double Exp(Brown)	311.5	0	102.8	206.8025	ค่า α = 0.01
Double Exp(Holt)	36.2	0.01	11.1	51.5888	ค่า α = 0.1, β=0.1
Triple Exp(Brown)	45.9	0.001	13.1	47.2835	ค่า α = 0.18
AR(1)	21.9	0.057	0.037	44.2392	แปลงข้อมูลด้วยรากที่สี่



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.14 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคไอกรนในกรุงเทพมหานคร

3.3 ไอกรนในภาคกลาง

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคไอกรนในภาคกลาง พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนน้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสมคือ ตัวแบบการถดถอยในตัวเองอันดับที่ 2 (Second Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สี่ ซึ่งถึงแม้ว่าค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์ ของตัวแบบนี้จะไม่น้อยที่สุด แต่ก็ต่างจากค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์ที่น้อยที่สุดเพียงเล็กน้อย (ดูรายละเอียดในตารางที่ 4.15)

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเองอันดับที่ 1 (first order Autoregressive parameter)

ϕ_2 คือ พารามิเตอร์การถดถอยในตัวเองอันดับที่ 2 (second order Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง ณ คาบเวลาที่ t-1

Z_{t-2} คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง ณ คาบเวลาที่ t-2

$$Z_t = 0.2563 + 0.4379Z_{t-1} + 0.4116Z_{t-2} + e_t$$

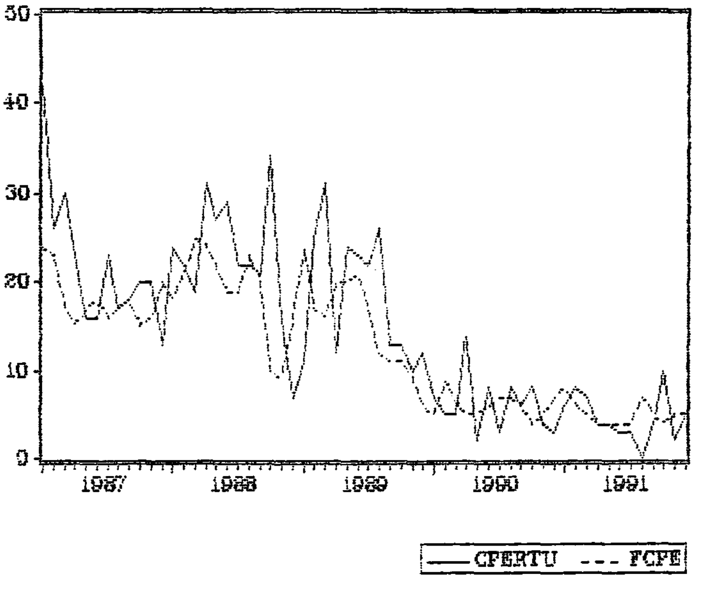
ดังนั้นสมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง ณ คาบเวลา t คือ

$$Z_t = 0.2563 + 0.4379\,Z_{t-1} + 0.4116Z_{t-2}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสี่ จึงจะได้จำนวนผู้ป่วยโรคไอกรนในภาคกลางที่แท้จริง

ตารางที่ 4.15 ตารางเปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคไอกรนในภาคกลาง

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	31.8	0.023	40.1	50.7530	N= 3
Single Exp.	20.5	0.364	34.7	50.2928	ค่า α = 0.34
Double Exp(Brown)	19.4	0.431	39.3	44.8928	ค่า α = 0.22
Double Exp(Holt)	18.9	0.464	39.0	44.7044	ค่า α =0.4, β =0.1
Triple Exp(Brown)	20.3	0.378	49.3	51.1358	ค่า α = 0.25
AR(2)	17.4	0.133	0.047	45.4197	แปลงข้อมูลด้วยรากที่สี่



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.15 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคไอกรนในภาคกลาง

3.4 ไอกรนในภาคเหนือ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคไอกรนในภาคเหนือพบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนน้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสมคือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สอง (ดูรายละเอียดในตารางที่ 4.16)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเองอันดับที่ 1 (first order Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคเหนือ ณ คาบเวลาที่ t-1

$$Z_t = 0.8023 + 0.7109Z_{t-1} + e_t$$

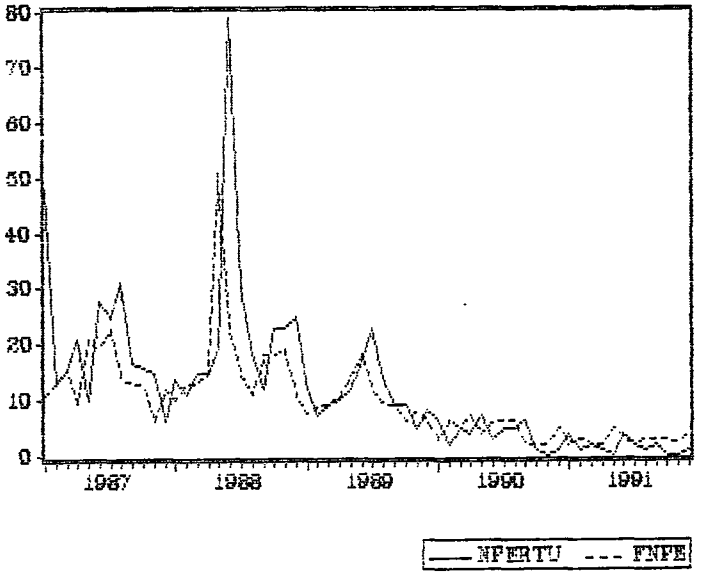
ดังนั้นสมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคเหนือ ณ คาบเวลา t คือ

$$Z_t = 0.8023 + 0.7109Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสอง จึงจะได้จำนวนผู้ป่วยโรคไอกรนในภาคเหนือ ที่แท้จริง

<u>ตารางที่ 4.16</u> เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคไอกรนในภาคเหนือ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	20.9	0.285	124.2	64.7957	N= 3
Single Exp.	17.3	0.573	106.0	73.9047	ค่า α =0.20
Double Exp(Brown)	23.3	0.225	109.3	65.5117	ค่า α =0.07
Double Exp(Holt)	34.4	0.017	113.6	74.2240	ค่า α =0.1, β =0.1
Triple Exp(Brown)	14.8	0.737	148.3	65.3492	ค่า α =0.19
AR(1)	13.5	0.336	1.11	คำนวณไม่ได้	แปลงข้อมูลด้วยรากที่2



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.16 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคไอกรนในภาคเหนือ

3.5 ไอกรนในภาคตะวันออกเฉียงเหนือ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคไอกรนในภาคตะวันออกเฉียงเหนือพบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนและค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สี่ (ดูรายละเอียดในตารางที่ 4.17)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาที่ t-1

$$Z_t = 0.4753 + 0.7905\ Z_{t-1} + e_t$$

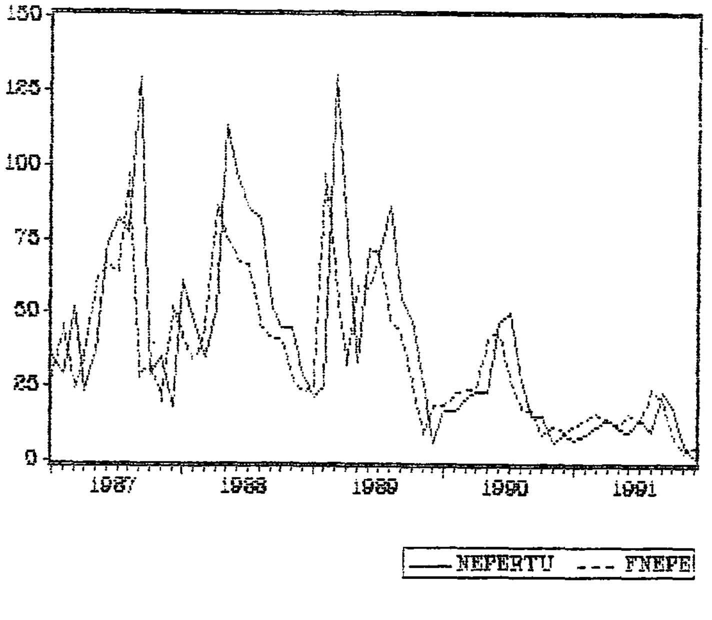
ดังนั้นสมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วย โรคไอกรน ในภาคตะวันออกเฉียงเหนือ ณ คาบเวลา t คือ

$$Z_t = 0.4753 + 0.7905\ Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสี่จึงจะได้จำนวนผู้ป่วยโรคไอกรนในภาคตะวันออกเฉียงเหนือ ที่แท้จริง

ตารางที่ 4.17 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคไอกรนในภาคตะวันออกเฉียงเหนือ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	52.4	0	792.5	92.1325	N= 3
Single Exp.	36.4	0.01	681.0	77.7904	ค่า α = 0.51
Double Exp(Brown)	56.4	0	759.8	87.4828	ค่า α = 0.18
Double Exp(Holt)	31.7	0.034	746.2	74.6837	ค่า α =0.6, β =0.10
Triple Exp(Brown)	42	0.02	868.4	86.3785	ค่า α =0.18
AR(1)	16.8	0.21	0.124	60.635	**แปลงข้อมูลด้วยรากที่สี่**



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.17 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคไอกรนในภาคตะวันออกเฉียงเหนือ

3.6 ไอกรนในภาคใต้

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่างๆของโรคไอกรนในภาคใต้ พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดคือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสมคือ ตัวแบบการถดถอยในตัวเองอันดับที่ 2 (Second Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สอง (ดูรายละเอียดในตารางที่ 4.18)

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเองอันดับที่ 1 (first order Autoregressive parameter)

ϕ_2 คือ พารามิเตอร์การถดถอยในตัวเองอันดับที่ 2 (second order Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคใต้ ณ คาบเวลาที่ t-1

Z_{t-2} คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคใต้ ณ คาบเวลาที่ t-2

$$Z_t = 0.4261 + 0.5264Z_{t-1} + 0.3205Z_{t-2} + e_t$$

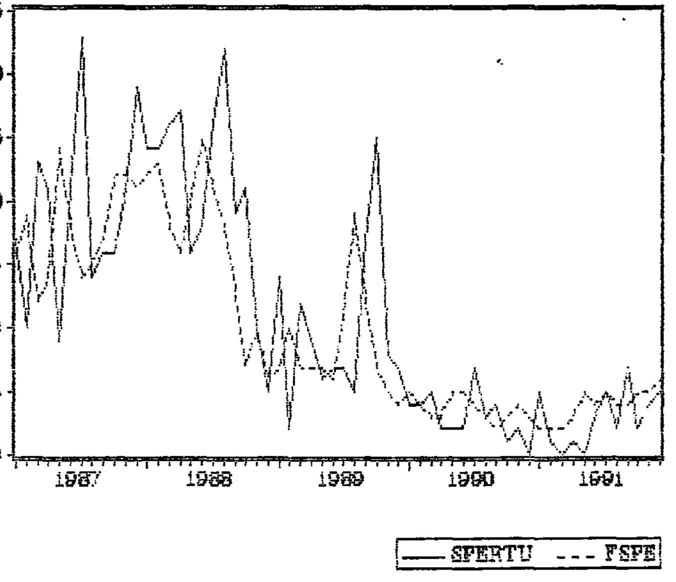
ดังนั้นสมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง ณ คาบเวลา t คือ

$$Z_t = 0.4261 + 0.5264Z_{t-1} + 0.3205Z_{t-2}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสอง จะได้จำนวนผู้ป่วยโรคไอกรนในภาคใต้ที่แท้จริง

ตารางที่ 4.18 เปรียบเทียบวิธีพยากรณ์ต่างๆของโรคไอกรนในภาคใต้

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	41.4	0.001	37.8	67.0077	N= 3
Single Exp.	23.3	0.224	36.0	67.9196	ค่า α =0.36
Double Exp(Brown)	21.5	0.307	37.9	58.4959	ค่า α =0.17
Double Exp(Holt)	21.4	0.314	38.0	58.4411	ค่า α =0.3, β =0.1
Triple Exp(Brown)	26.8	0.11	45.2	61.5342	ค่า α =0.20
AR(2)	15.1	0.234	0.77	1.1699	แปลงข้อมูลด้วยรากที่2



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.18 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคไอกรนในภาคใต้

4. บาดทะยัก

4.1 บาดทะยักทั่วประเทศ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคบาดทะยักทั่วประเทศ พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดคือ วิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้งของวินเตอร์ แต่เนื่องจากต้องประมาณค่าพารามิเตอร์ 3 ค่า จึงเลือกวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้ง ของบราวน์ เพราะมีค่าผลบวกกำลังสองค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์มากกว่าวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้งของวินเตอร์เพียงเล็กน้อย แต่ประมาณค่าพารามิเตอร์น้อยกว่า คือ ต้องประมาณค่าพารามิเตอร์เพียง 1 ค่า (ดูรายละเอียดในตารางที่ 4.19) ตัวแบบที่ใช้ได้ คือ

$$x_t = \beta_0 + \beta_1 t + e_t$$

เมื่อ β_0 , β_1 คือ ค่าคงที่ทำให้เรียบ

e_t คือ ความคลาดเคลื่อนสุ่ม (error)

x_t คือ จำนวนผู้ป่วยโรคบาดทะยักทั่วประเทศ ณ คาบเวลาที่ t

ดังนั้นสมการในการพยากรณ์ค่าสังเกตในอนาคต ณ คาบเวลาที่ $t + T$ คือ

$$F_{t+T} = a_t + b_t T$$

คำนวณ a_t และ b_t ดังนี้

$$a_t = 2S_t - S_t^{(2)}$$

$$b_t = \frac{\alpha}{1 - \alpha} (S_t - S_t^{(2)})$$

$$S_t = \alpha x_t + (1 - \alpha)S_{t-1}$$

$$S_t^{(2)} = \alpha S_t + (1 - \alpha)S_{t-1}^{(2)}$$

คำนวณค่าได้ดังนี้ $S_t = 60$, $S_t^{(2)} = 68$

$$a_t = 2(60) - 68$$

$$= 52$$

$$b_t = \frac{0.11}{1 - 0.11}(60 - 68)$$

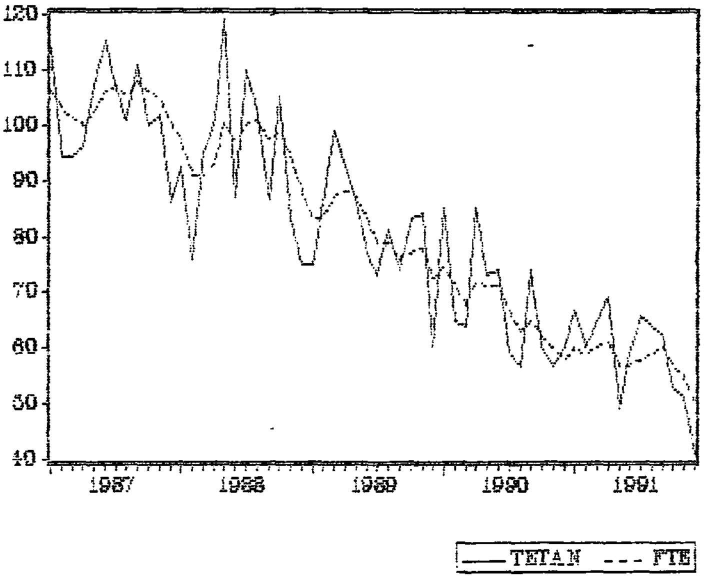
$$= -0.9888$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรคบาดทะยักทั่วประเทศ ณ คาบเวลาที่ $t + T$ คือ

$$F_{t+T} = 52 - 0.9888\ T$$

ตารางที่ 4.19 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคบาดทะยักทั่วประเทศ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Average	23.2	0.182	121.2	12.10	จำนวนเทอมที่ใช้ 3 เทอม
Single Exp.	26.5	0.117	112.8	11.76	ค่า α = 0.31
Double Exp. (Brown)	**24.5**	**0.178**	**106.7**	**11.28**	**ค่า α = 0.11**
Double Exp. (Holt)	24.9	0.165	102.7	11.12	ค่า α = 0.1 , β = 0.3
Triple Exp. (Brown)	23.4	0.219	109.2	11.39	ค่า α = 0.08
ARIMA (0,1,1)	13.2	0.507	105.4	11.56	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.19 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้งของบราวน์ α =0.11 ของโรคบาดทะยักทั่วประเทศ

4.2 บาดทะยักในกรุงเทพมหานคร

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคบาดทะยักในกรุงเทพมหานคร พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนน้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 โดยทำการปรับค่าเฉลี่ยให้คงที่ด้วยการหาค่าแตกต่างแรก (First Order Autoregressive with First Difference model) (ดูรายละเอียดในตารางที่ 4.20)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเองอันดับที่ 1 (first order Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ $x_t - x_{t-1}$

x_t คือ จำนวนผู้ป่วยโรคบาดทะยักในกรุงเทพมหานคร ณ คาบเวลาที่ t

$$Z_t = -0.7317Z_{t-1} + e_t$$

ดังนั้นสมการพยากรณ์ค่า Z_t คือ

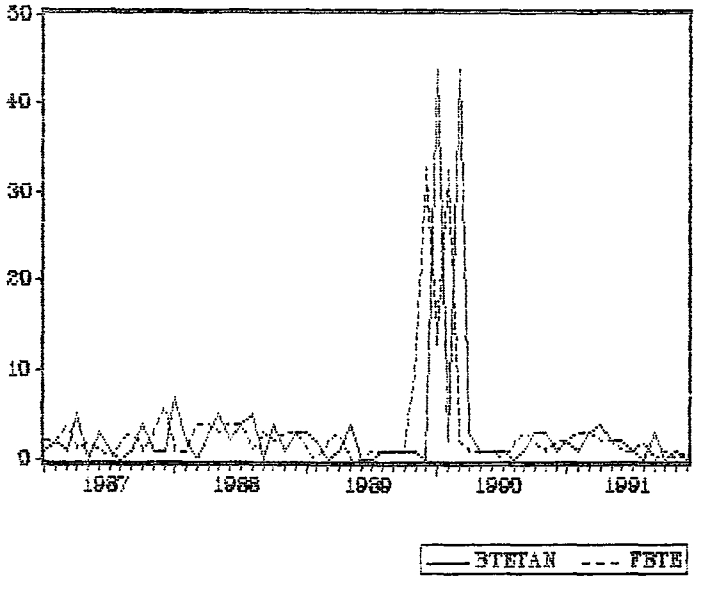
$$Z_t = -0.7317Z_{t-1}$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรคบาดทะยักในกรุงเทพมหานคร ณ คาบเวลาที่ t คือ

$$x_t = x_{t-1} - 0.7317Z_{t-1}$$

ตารางที่ 4.20 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคบาดทะยักในกรุงเทพมหานคร

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	21	0.277	74.1	156.5817	N= 3
Single Exp.	14.5	0.753	62.5	65.2226	ค่า α =0.02
Double Exp(Brown)	14.6	0.746	62.1	93.6690	ค่า α =0.01
Double Exp(Holt)	14.8	0.734	68.1	138.6092	ค่า α =0.1, β =0.1
Triple Exp(Brown)	15.6	0.681	72.5	174.5718	ค่า α =0.08
ARIMA(1,1,0)	9.7	0.781	58.64	คำนวณไม่ได้	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.20 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคบาดทะยักในกรุงเทพมหานคร

4.3 บาดทะยักในภาคกลาง

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคบาดทะยักในภาคกลาง พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนน้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สอง (ดูรายละเอียดในตารางที่ 4.21) ถึงแม้ว่าค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์จะไม่ใช่ค่าที่น้อยที่สุด แต่ก็มีค่าใกล้เคียงกับค่าที่น้อยที่สุด ดังนั้นจึงเลือกใช้ตัวแบบนี้ในการพยากรณ์

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคบาดทะยักในภาคกลาง ณ คาบเวลาที่ t-1

$$Z_t = 3.5730 + 0.1129\ Z_{t-1} + e_t$$

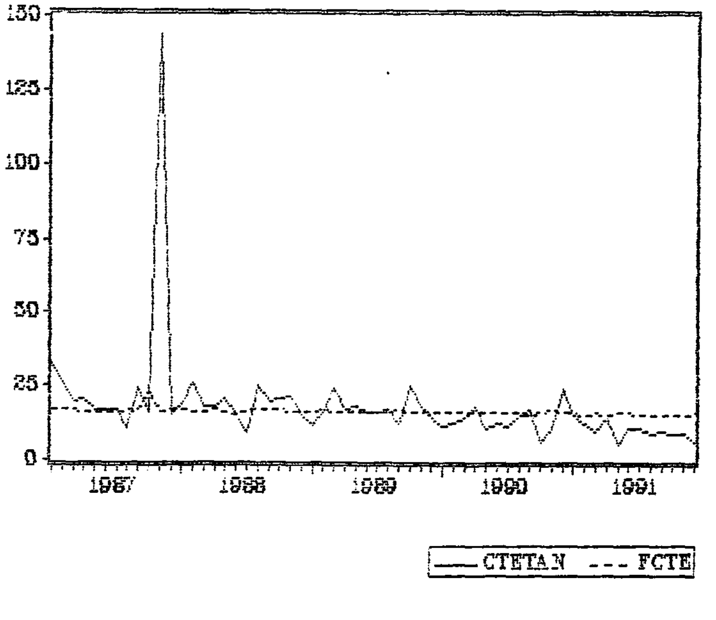
ดังนั้นสมการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคบาดทะยักในภาคกลาง ณ คาบเวลา t คือ

$$Z_t = 3.5730 + 0.1129\ Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสองจึงจะได้จำนวนผู้ป่วยโรคบาดทะยักในภาคกลางที่แท้จริง

ตารางที่ 4.21 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคบาดทะยักในภาคกลาง

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	21.3	0.263	389.0	38.6760	N= 3
Single Exp.	7.1	0.994	298.3	44.4511	ค่า α =0.06
Double Exp(Brown)	8.6	0.98	317.9	26.9638	ค่า α =0.07
Double Exp(Holt)	7.3	0.993	312.1	25.3932	ค่า α =0.1, β =0.1
Triple Exp(Brown)	6.5	0.996	436.8	51.6681	ค่า α =0.16
AR(1)	14.3	0.352	1.54	34.198	แปลงข้อมูลด้วยรากที่2



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.21 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคบาดทะยักในภาคกลาง

4.4 บาดทะยักในภาคเหนือ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคบาดทะยักในภาคเหนือ พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสมคือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สอง (ดูรายละเอียดในตารางที่ 4.22)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคบาดทะยักในภาคเหนือ ณ คาบเวลาที่ t-1

$$Z_t = 1.9677 + 0.5576\ Z_{t-1} + e_t$$

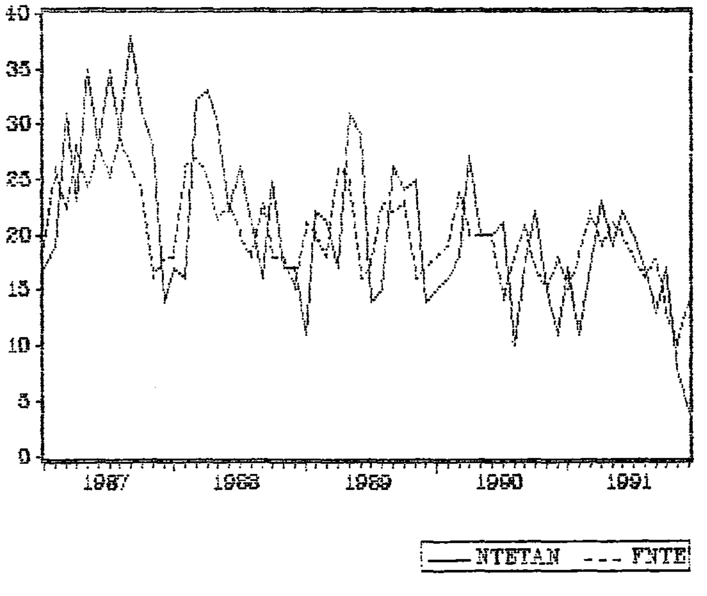
ดังนั้นสมการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคบาดทะยักในภาคเหนือ ณ คาบเวลา t คือ

$$Z_t = 1.9677 + 0.5576\ Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสองจึงจะได้จำนวนผู้ป่วยโรคบาดทะยักในภาคเหนือที่แท้จริง

ตารางที่ 4.22 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคบาดทะยักในภาคเหนือ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	49.1	0	46.7	33.2965	N= 3
Single Exp.	34.9	0.014	40.3	33.0755	ค่า α =0.20
Double Exp(Brown)	33.2	0.023	46.2	33.2797	ค่า α = 0.21
Double Exp(Holt)	30.5	0.046	45.4	32.4060	ค่า α = 0.5,β=0.1
Triple Exp(Brown)	25	0.161	53.2	31.7089	ค่า α = 0.22
AR(1)	16.7	0.213	0.51	29.1657	แปลงข้อมูลด้วยรากที่2



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.22 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคบาดทะยักในภาคเหนือ

4.5 บาดทะยักในภาคตะวันออกเฉียงเหนือ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคบาดทะยักในภาคตะวันออกเฉียงเหนือ พบว่า มีค่าผลบวกกำลังสองค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดคือ วิธีการปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียวของบราวน์ (ดูรายละเอียดในตารางที่ 4.23) ดังนั้นตัวแบบที่ใช้ได้ คือ

$$x_t = \beta_t + e_t$$

เมื่อ β_t คือ ค่าคงที่ทำให้เรียบ

e_t คือ ความคลาดเคลื่อนสุ่ม (error)

x_t คือ จำนวนผู้ป่วยโรคบาดทะยักในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาที่ t

สมการในการพยากรณ์ค่าสังเกตในอนาคต ณ คาบเวลาที่ t + 1 คือ

$$\begin{aligned} F_{t+1} &= B_t = S_t \\ &= \alpha x_t + (1 - \alpha)S_{t-1} \\ &= (0.14)x_t + (1 - 0.14)S_{t-1} \\ &= (0.14)x_t + (0.86)S_{t-1} \end{aligned}$$

โดย B_t เป็นค่าประมาณของ β_t

ดังนั้น สมการในการพยากรณ์จำนวนผู้ป่วยในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาที่ t + 1 คือ

$$F_{t+1} = (0.14)x_t + (0.86)S_{t-1}$$

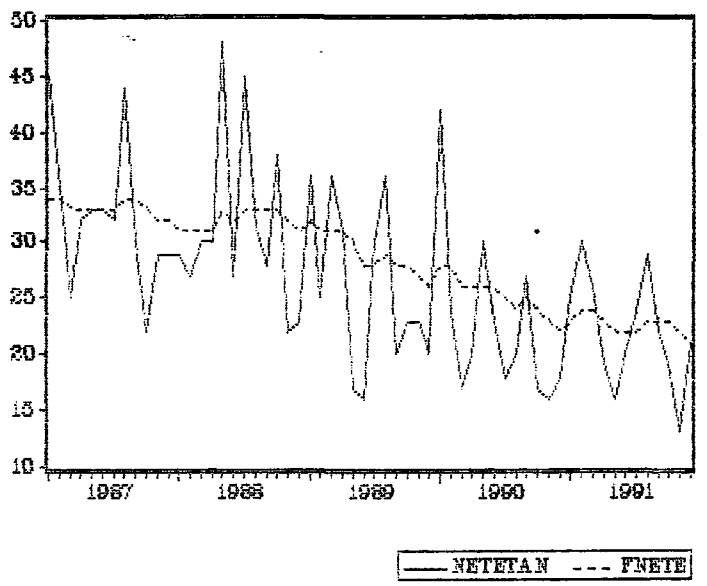
เมื่อ x_t คือ จำนวนผู้ป่วยโรคบาดทะยักในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาที่ t

S_{t-1} คือ ค่าของการทำให้เรียบแบบเอกซโปเนนเชียลครั้งเดียว ณ คาบเวลาที่ t-1

F_{t+1} คือ ค่าพยากรณ์ ณ คาบเวลา t+1

ตารางที่ 4.23 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคบาดทะยักในภาคตะวันออกเฉียงเหนือ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Average	29.5	0.043	62.1	26.29	จำนวนเทอมที่ใช้ 3 เทอม
Single Exp.	**15.5**	**0.689**	**49.1**	**23.66**	**ค่า α = 0.14**
Double Exp. (Brown)	14.2	0.771	52.3	20.31	ค่า α = 0.09
Double Exp. (Holt)	12.5	0.865	52.0	20.63	ค่า α = 0.1 , β = 0.1
Triple Exp. (Brown)	13.5	0.812	56.3	24.77	ค่า α = 0.09
AR(1)	16.1	0.242	54.7	26.52	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.23 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียวของบราวน์ α =0.14 ของโรคบาดทะยักในภาคตะวันออกเฉียงเหนือ

4.6 บาดทะยักในภาคใต้

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคบาดทะยักในภาคใต้ พบว่า วิธีพยากรณ์ที่มีค่าผลบวกกำลังสองค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดคือ วิธีการปรับเรียบแบบเอกซโปเนนเชียลครั้งเดียว โดยใช้วิธีของบราวน์ (ดูรายละเอียดในตารางที่ 4.24) ดังนั้นตัวแบบที่ใช้ได้ คือ

$$x_t = \beta_t + e_t$$

เมื่อ x_t คือ จำนวนผู้ป่วยโรคบาดทะยักในภาคใต้ ณ คาบเวลาที่ t

β_t คือ ค่าคงที่ทำให้เรียบ

e_t คือ ความคลาดเคลื่อนสุ่ม (error)

สมการในการพยากรณ์ค่าสังเกตในอนาคต ณ คาบเวลาที่ $t + 1$ คือ

$$\begin{aligned} F_{t+1} &= a_t = S_t \\ &= \alpha x_t + (1 - \alpha)S_{t-1} \\ &= (0.22)x_t + (1 - 0.22)S_{t-1} \\ &= (0.22)x_t + (0.78)S_{t-1} \end{aligned}$$

โดย a_t เป็นค่าประมาณของ β_t

ดังนั้น สมการในการพยากรณ์จำนวนผู้ป่วยในภาคใต้ ณ คาบเวลาที่ $t + 1$ คือ

$$F_{t+1} = (0.22)x_t + (0.78)S_{t-1}$$

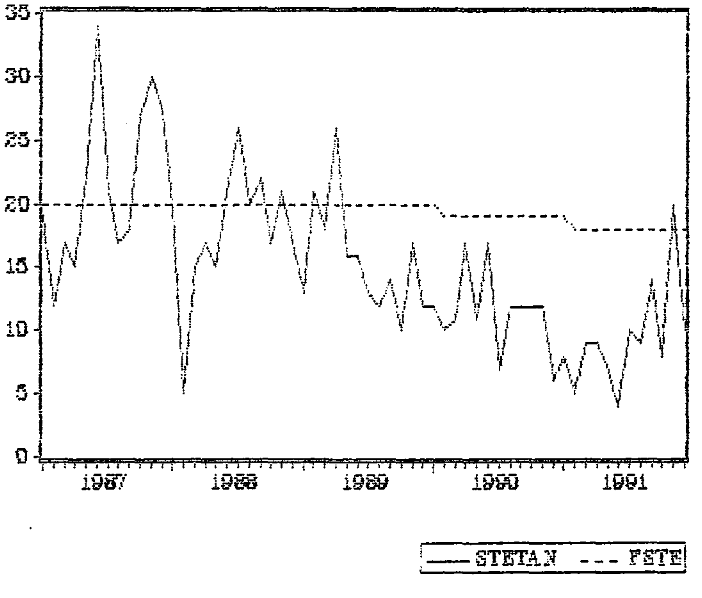
เมื่อ x_t คือ จำนวนผู้ป่วยโรคบาดทะยักในภาคใต้ ณ คาบเวลาที่ t

S_{t-1} คือ ค่าของการทำให้เรียบแบบเอกซโปเนนเชียลครั้งเดียว ณ คาบเวลาที่ $t-1$

F_{t+1} คือ ค่าพยากรณ์ ณ คาบเวลา $t+1$

ตารางที่ 4.24 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคบาดทะยักในภาคใต้

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Average	35.7	0.008	33.2	35.68	จำนวนเทอมที่ใช้ 3 เทอม
Single Exp.	21.1	0.330	27.1	33.20	ค่า α = 0.22
Double Exp. (Brown)	21.3	0.319	29.3	33.22	ค่า α = 0.16
Double Exp. (Holt)	21.5	0.311	29.3	33.01	ค่า α = 0.3 , β = 0.1
Triple Exp. (Brown)	19.7	0.410	32.6	33.43	ค่า α = 0.16
ARMA(1,1)	10.7	0.557	26.0	31.48	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.24 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียวของบราวน์ α =0.22 ของโรคบาดทะยักในภาคใต้

5.โปลิโอ

5.1 โปลิโอทั้งประเทศ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคโปลิโอทั้งประเทศพบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนน้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสมคือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) (ดูรายละเอียดในตารางที่ 4.25)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ จำนวนผู้ป่วยโรคโปลิโอทั้งประเทศ ณ คาบเวลาที่ t-1

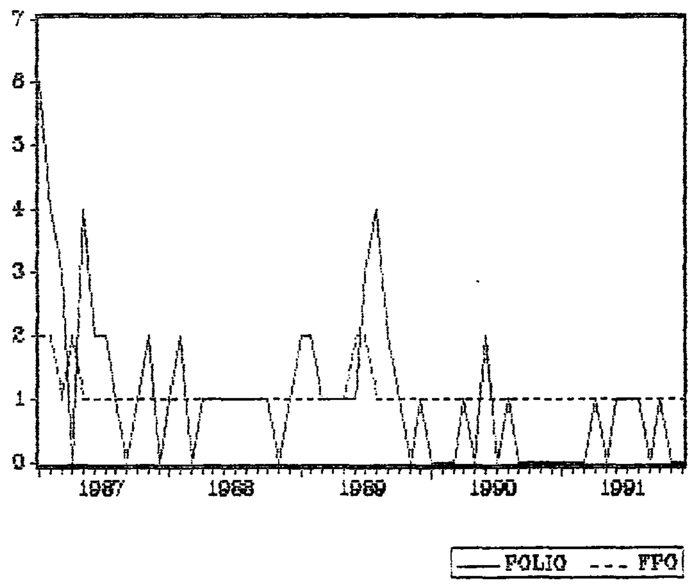
$$Z_t = 0.5180 + 0.5460\ Z_{t-1} + e_t$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอทั้งประเทศ ณ คาบเวลา t คือ

$$Z_t = 0.5180 + 0.5460\ Z_{t-1}$$

ตารางที่ 4.25 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคโปลิโอทั่วประเทศ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	10.7	0.905	1.2	43.8095	N= 3
Single Exp.	12.5	0.864	1.1	24.6363	ค่า α = 0.29
Double Exp(Brown)	10.6	0.935	1.1	36.7952	ค่า α =0.23
Double Exp(Holt)	110.5	0.94	1.1	37.9674	ค่า α =0.4, β =0.1
Triple Exp(Brown)	11.6	0.901	1.3	36.7081	ค่า α =0.22
AR(1)	14.4	0.788	1.0	คำนวณไม่ได้	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.25 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคโปลิโอทั่วประเทศ

5.2 โปลิโอในกรุงเทพมหานคร

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคโปลิโอในกรุงเทพมหานคร พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนน้อยที่สุด คือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสม คือ ตัวแบบการถดถอยในตัวเอง อันดับที่ 1 (First Order Autoregressive model) (ดูรายละเอียดในตารางที่ 4.26)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ จำนวนผู้ป่วยโรคโปลิโอในกรุงเทพมหานคร คาบเวลาที่ t-1

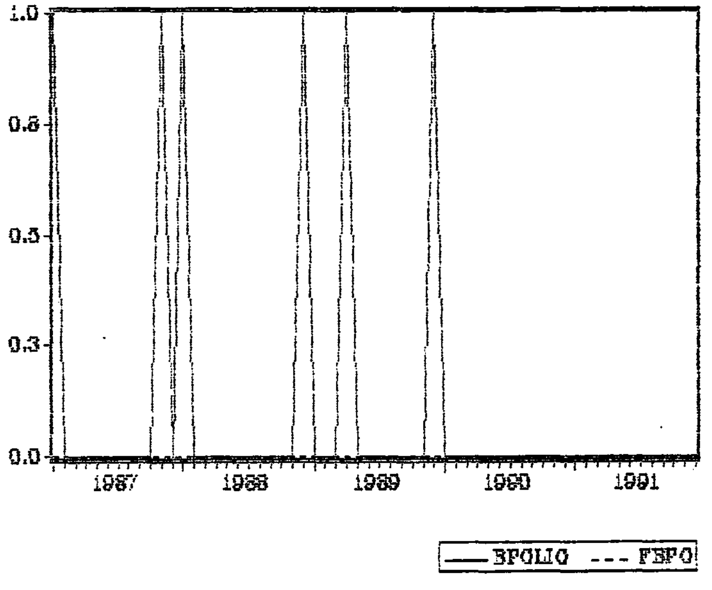
$$Z_t = 0.0943 - 0.0943\ Z_{t-1} + e_t$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในกรุงเทพมหานคร ณ คาบเวลา t คือ

$$Z_t = 0.0943 - 0.0943\ Z_{t-1}$$

ตารางที่ 4.26 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคโปลิโอในกรุงเทพมหานคร

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	27.2	0.075	0.1	93.3333	N= 3
Single Exp.	18.2	0.508	0.1	7.5682	ค่า α = 0.01
Double Exp(Brown)	20.8	0.347	0.1	8.8432	ค่า α =0.12
Double Exp(Holt)	20.9	0.343	0.1	8.6878	ค่า α =0.2, β =0.1
Triple Exp(Brown)	18.8	0.468	0.2	6.7664	ค่า α =0.23
AR(1)	14.2	0.357	0.078	คำนวณไม่ได้	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.26 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคโปลิโอในกรุงเทพมหานคร

5.3 โปลิโอในภาคกลาง

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคโปลิโอในภาคกลาง พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดคือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสมคือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สอง (ดูรายละเอียดในตารางที่ 4.27)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคโปลิโอในภาคกลาง ณ คาบเวลาที่ t-1

Z_t คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคโปลิโอในภาคกลาง ณ คาบเวลาที่ t

$$Z_t = 0.6210 + 0.4071\ Z_{t-1} + e_t$$

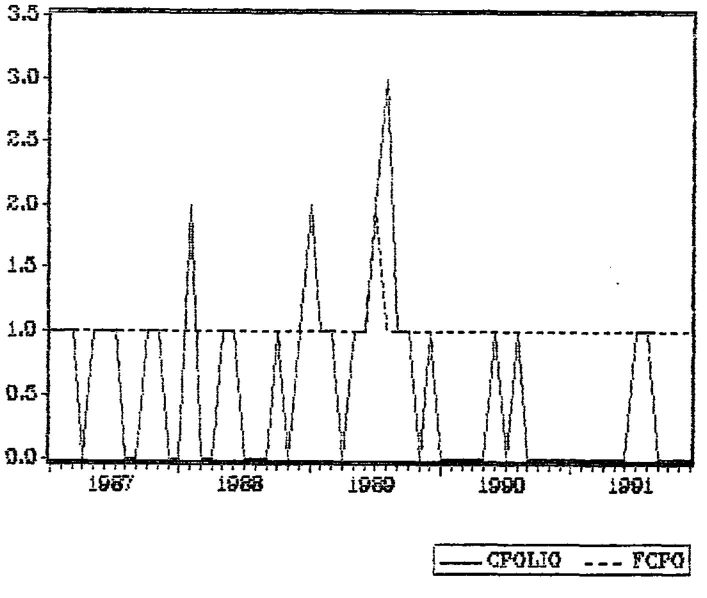
ดังนั้นสมการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคโปลิโอในภาคกลาง ณ คาบเวลา t คือ

$$\hat{Z}_t = 0.6210 + 0.4071\ Z_{t-1}$$

เมื่อได้ค่าพยากรณ์มาแล้ว ต้องนำมาทำการยกกำลังสองจึงจะได้จำนวนผู้ป่วยโรคโปลิโอในภาคกลางที่แท้จริง

ตารางที่ 4.27 เปรียบเทียบวิธีพยากรณ์ต่าง ๆ ของโรคโปลิโอในภาคกลาง

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	18.6	0.417	0.5	59.95	N= 3
Single Exp.	14.8	0.733	0.4	19.57	ค่า α = 0.18
Double Exp(Brown)	13.4	0.816	0.5	24.25	ค่า α = 0.11
Double Exp(Holt)	14.6	0.751	0.5	24.23	ค่า α =0.3, β =0.1
Triple Exp(Brown)	14.6	0.747	0.5	23.51	ค่า α =0.13
AR(1)	10.6	0.64	0.018	0.1506	แปลงข้อมูลด้วยรากที่สอง



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.27 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคโปลิโอในภาคกลาง

5.4 โปลิโอในภาคเหนือ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคโปลิโอในภาคเหนือ พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อน และค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดคือ วิธีของบอกซ์และเจนกินส์ โดยตัวแบบที่เหมาะสมคือ ตัวแบบการถดถอยในตัวเองอันดับที่ 1 (First Order Autoregressive model) (ดูรายละเอียดในตารางที่ 4.28)

$$Z_t = \delta + \phi_1 Z_{t-1} + e_t$$

เมื่อ δ คือ ค่าคงที่

ϕ_1 คือ พารามิเตอร์การถดถอยในตัวเอง (Autoregressive parameter)

e_t คือ การรบกวนสุ่ม (random shock) หรือ ความคลาดเคลื่อนสุ่ม (errors)

Z_{t-1} คือ จำนวนผู้ป่วยโรคโปลิโอในภาคเหนือ ณ คาบเวลาที่ t-1

Z_t คือ จำนวนผู้ป่วยโรคโปลิโอในภาคเหนือ ณ คาบเวลาที่ t

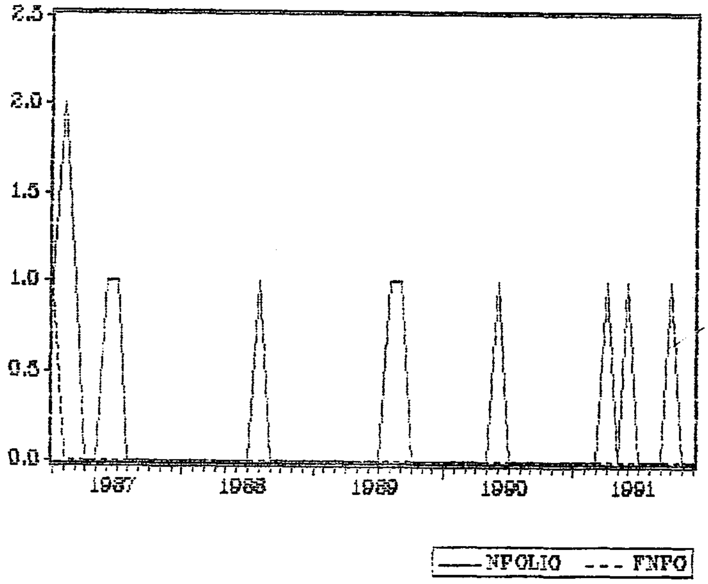
$$Z_t = 0.1424 + 0.2766\ Z_{t-1} + e_t$$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในภาคเหนือ ณ คาบเวลา t คือ

$$\hat{Z}_t = 0.1424 + 0.2766\ Z_{t-1}$$

ตารางที่ 4.28 เปรียบเทียบวิธีพยากรณ์ต่างๆของโรคโปลิโอในภาคเหนือ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	18.4	0.432	0.2	85.18	N= 3
Single Exp.	15.7	0.68	0.2	13.0147	ค่า α = 0.17
Double Exp(Brown)	17.8	0.537	0.2	14.9929	ค่า α = 0.18
Double Exp(Holt)	14.7	0.741	0.2	14.4675	ค่า α =0.2, β =0.2
Triple Exp(Brown)	19.3	0.44	0.2	13.2896	ค่า α =0.14
AR(1)	20.5	0.083	0.183	คำนวณไม่ได้	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.28 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีบอกซ์และเจนกินส์ ของโรคโปลิโอในภาคเหนือ

5.5 โปลิโอในภาคตะวันออกเฉียงเหนือ

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคโปลิโอในภาคตะวันออกเฉียงเหนือ พบว่าวิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนและค่าเฉลี่ยเปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์น้อยที่สุด คือ ตัวแบบจากวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียวของบราวน์ โดยมีค่าที่กำหนดน้ำหนักการปรับเรียบแบบเอกซ์โปเนนเชียล (α) เท่ากับ 0.19 (ดูรายละเอียดในตารางที่ 4.29)

$$x_t = \beta_t + e_t$$

เมื่อ β_t คือ ค่าคงที่ทำให้เรียบ

e_t คือ ความคลาดเคลื่อนสุ่ม (error)

x_t คือ จำนวนผู้ป่วยโรคโปลิโอในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาที่ t

สมการในการพยากรณ์ค่าสังเกตในอนาคต ณ คาบเวลาที่ $t + 1$ คือ

$$\begin{aligned} F_{t+1} &= a_t = S_t \\ &= \alpha x_t + (1 - \alpha)S_{t-1} \\ &= (0.19)x_t + (1 - 0.19)S_{t-1} \\ &= (0.19)x_t + (0.81)S_{t-1} \end{aligned}$$

โดย a_t คือ ค่าประมาณของ β_t

α คือ ค่าที่กำหนดน้ำหนักของการปรับเรียบแบบเอกซ์โปเนนเชียล โดย $0 < \alpha < 1$

S_t คือ ค่าของการทำให้เรียบแบบเอกซโปเนนเชียลครั้งเดียว ณ คาบเวลาที่ t

ดังนั้น สมการในการพยากรณ์จำนวนผู้ป่วยโปลิโอในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาที่ t + 1 คือ

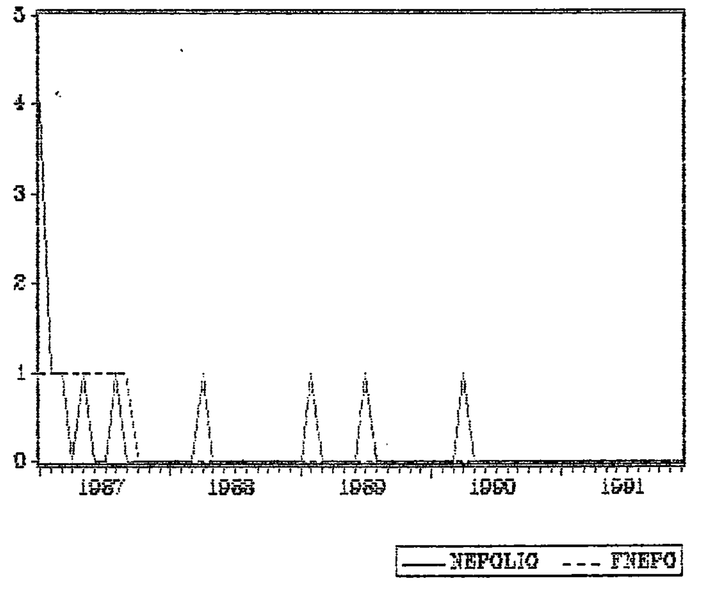
$$F_{t+1} = (0.19)x_t + (0.81)S_{t-1}$$

เมื่อ x_t คือ จำนวนผู้ป่วยโรคโปลิโอในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาที่ t

S_{t-1} คือ ค่าของการทำให้เรียบแบบเอกซโปเนนเชียลครั้งเดียว ณ คาบเวลาที่ t-1

ตารางที่ 4.29 เปรียบเทียบวิธีพยากรณ์ต่างๆของโรคโปลิโอในภาคตะวันออกเฉียงเหนือ

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	15.2	0.649	0.2	83.3333	N= 3
Single Exp.	24.5	0.178	0.2	8.0161	ค่า = 0.19
Double Exp(Brown)	19.4	0.433	0.2	15.0542	ค่า α = 0.23
Double Exp(Holt)	29.1	0.065	0.2	14.2823	ค่า α =0.3,β =0.2
Triple Exp(Brown)	16.4	0.631	0.3	13.9351	ค่า α =0.29
Box-Jenkins					ค่าพารามิเตอร์เป็นศูนย์



——— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.29 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียวของบราวน์ α =0.19 ของโรคโปลิโอในภาคตะวันออกเฉียงเหนือ

5.6 โปลิโอในภาคใต้

จากการเปรียบเทียบวิธีการพยากรณ์จำนวนผู้ป่วยวิธีต่าง ๆ ของโรคโปลิโอในภาคใต้ พบว่า วิธีพยากรณ์ที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนน้อยที่สุด คือ ตัวแบบจากวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้งของโฮลท์ (ดูรายละเอียดในตารางที่ 4.30) ซึ่งต้องประมาณค่าพารามิเตอร์ถึง 2 ค่า พิจารณาวิธีที่มีค่าเฉลี่ยผลบวกกำลังสองของค่าคลาดเคลื่อนใกล้เคียงกัน พบว่า ตัวแบบที่มีค่าคลาดเคลื่อน ใกล้เคียงกันแต่ประมาณค่าพารามิเตอร์เพียงตัวเดียวคือ ตัวแบบจากวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้งของบราวน์ โดยมีค่าที่กำหนดน้ำหนักการปรับเรียบแบบเอกซ์โปเนนเชียล (α) เท่ากับ 0.01 ดังนั้นตัวแบบที่ใช้ได้ คือ

$$X_t = \beta_0 + \beta_1 t + (1/2)\beta_2 t^2 + e_t$$

เมื่อ $\beta_0, \beta_1, \beta_2$ คือ ค่าคงที่ทำให้เรียบ

e_t คือ ความคลาดเคลื่อนสุ่ม (error)

X_t คือ จำนวนผู้ป่วยโรคโปลิโอในภาคใต้ ณ คาบเวลาที่ t

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในภาคใต้ ณ คาบเวลา $t + T$ คือ

$$F_{t+T} = a_t + b_t T + (1/2)c_t T^2$$

คำนวณค่า a_t, b_t และ c_t ดังนี้

$$a_t = 3S_t - 3S_t^{(2)} + 3S_t^{(3)}$$
$$b_t = \alpha/2(1-\alpha)^2\{(6-5\alpha)S_t - (10-8\alpha)S_t^{(2)} + (4-3\alpha)S_t^{(3)}\}$$
$$c_t = (\alpha/1-\alpha)^2\{S_t - 2S_t^{(2)} + S_t^{(3)}\}$$
$$S_t = \alpha X_t + (1-\alpha)S_{t-1}$$
$$S_t^{(2)} = \alpha S_t + (1-\alpha)S^{(2)}_{t-1}$$

$S_t^{(3)} = \alpha S^{(2)}_{t-1} + (1-\alpha) S^{(3)}_{t-1}$

α คือ ค่าที่กำหนดน้ำหนักการปรับเรียบแบบเอกซ์โปเนนเชียล โดย $0 < \alpha < 1$

$S_t = 0$

$S_t^{(2)} = 0$

$S_t^{(3)} = 0$

ดังนั้น

$a_t = 0$

$b_t = 0$

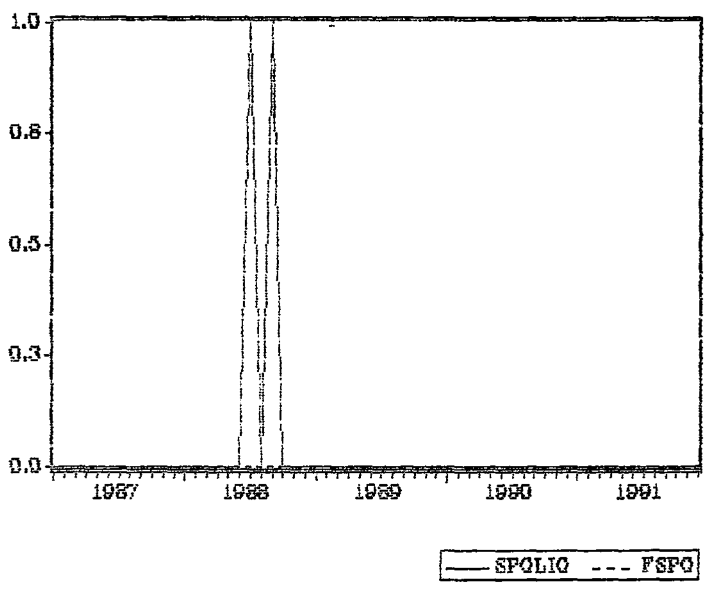
$c_t = 0$

ดังนั้นสมการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในภาคใต้ คือ

$F_{t+T} = 0$

ตารางที่ 4.30 เปรียบเทียบวิธีพยากรณ์ต่างๆของโรคโปลิโอในภาคใต้

เทคนิคการวิเคราะห์	Box-Pierce	P-value	MSE	MAPE	หมายเหตุ
Simple moving Avg	22.7	0.203	0	83.3333	N= 3
Single Exp.	14.6	0.746	0	3.3731	ค่า α =0.01
Double Exp(Brown)	14.4	0.761	0	3.3564	ค่า α =0.01
Double Exp(Holt)	16.4	0.628	0	3.2069	ค่า α =0.10, β=0.10
Triple Exp(Brown)	**14.3**	**0.769**	**0**	**3.34**	**ค่า α =0.01**
AR(1)	14.6	0.33	0.033	คำนวณไม่ได้	
MA(1)	14.6	0.33	0.033	คำนวณไม่ได้	



—— ค่าจริง

---- ค่าพยากรณ์

รูปที่ 4.30 แสดงค่าการพยากรณ์เปรียบเทียบกับค่าจริง ด้วยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้งของบราวน์ α =0.01 ของโรคโปลิโอในภาคใต้

บทที่ 5
สรุปผลการวิเคราะห์และข้อเสนอแนะ

สรุปผลการวิเคราะห์

ผลการวิเคราะห์ เพื่อหาตัวแบบทางสถิติที่เหมาะสมสำหรับการอธิบายและพยากรณ์จำนวนผู้ป่วยทั้งประเทศและรายภาคที่เกิดจากโรค วัณโรค คอตีบ ไอกรน บาดทะยัก และโปลิโอ เป็นดังนี้

1. ตัวแบบทางสถิติที่เหมาะสมสำหรับการอธิบายและพยากรณ์จำนวนผู้ป่วยทั้งประเทศ

1.1 โรควัณโรค พยากรณ์โดยใช้วิธีของ บอกซ์และเจนกินส์ ใช้ตัวแบบ ARIMA $(1,1,0)*(0,1,1)_{12}$ หมายถึง ต้องใช้จำนวนผู้ป่วยโรควัณโรคทั้งประเทศในอดีต ณ ตำแหน่งที่ไม่ใช่ฤดูกาลย้อนหลัง 1 คาบเวลา และหาค่าความแตกต่างแบบปกติไม่มีฤดูกาล 1 ครั้ง ,ใช้ค่าความคลาดเคลื่อนจากการพยากรณ์จำนวนผู้ป่วยโรควัณโรคทั่วประเทศ ณ ตำแหน่งฤดูกาลย้อนหลัง 1 คาบเวลาฤดูกาล และหาค่าความแตกต่างแบบมีฤดูกาล 1 ครั้ง ในการพยากรณ์จำนวนผู้ป่วยโรควัณโรคทั้งประเทศ ณ คาบเวลาถัดไป สมการพยากรณ์คือ

$$Z_t = 0.0562Z_{t-1} - 0.38e_{t-12}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรควัณโรคทั้งประเทศที่ผ่านการแปลงข้อมูลด้วยล็อกการิทึมเพื่อให้ มีความแปรปรวนคงที่ , หาค่าความแตกต่างแบบปกติไม่มีฤดูกาล 1 ครั้ง และหาค่าความแตกต่างแบบมีฤดูกาล 1 ครั้ง เพื่อให้มีค่าเฉลี่ยคงที่ ณ คาบเวลา t

e_{t-12} คือ ค่าคลาดเคลื่อนจากการพยากรณ์ ณ คาบเวลาที่ตำแหน่งฤดูกาล

1.2 โรคคอตีบ พยากรณ์โดยใช้วิธีของ บอกซ์และเจนกินส์ ใช้ตัวแบบ AR(1) หมายถึง ต้องใช้จำนวนผู้ป่วยโรคคอตีบทั้งประเทศในอดีตย้อนหลัง 1 คาบเวลา ในการพยากรณ์จำนวนผู้ป่วยโรคคอตีบทั้งประเทศ ณ คาบเวลาต่อไป สมการพยากรณ์ คือ

$$Z_t = 0.2957 + 0.8243\ Z_{t-1}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรคคอตีบทั้งประเทศที่ผ่านการแปลงข้อมูลด้วยรากที่สี่เพื่อให้มีความแปรปรวนคงที่ ณ คาบเวลาที่ t

1.3 โรคไอกรน พยากรณ์โดยใช้วิธีของ บอกซ์และเจนกินส์ ใช้ตัวแบบ AR(1)หมายถึง ต้องใช้จำนวนผู้ป่วยโรคไอกรนทั่วประเทศย้อนหลัง 1 คาบเวลา ในการพยากรณ์จำนวนผู้ป่วยโรคไอกรนทั้งประเทศ ณ คาบเวลาถัดไป สมการพยากรณ์ คือ

$$Z_t = 1.3172 + 0.8289Z_{t-1}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรคไอกรนทั่วประเทศ ที่ผ่านการแปลงด้วยรากที่สอง เพื่อให้มีความแปรปรวนคงที่ณ คาบเวลา t

1.4 โรคบาดทะยัก พยากรณ์โดยใช้การปรับเรียบแบบเอกโปเนนเชียลซ้ำสองครั้ง ที่ α=0.11 สมการพยากรณ์ คือ

$$F_{t+\mathcal{T}} = 52 - 0.9888\,\mathcal{T}$$

เมื่อ $F_{t+\mathcal{T}}$ คือ ค่าพยากรณ์จำนวนผู้ป่วยในอนาคต ณ คาบเวลา t+1

$\mathcal{T}$ คือ จำนวนคาบเวลาที่ต้องการพยากรณ์ล่วงหน้า

1.5 โปลิโอ พยากรณ์โดยใช้วิธีของ บอกซ์และเจนกินส์ ใช้ตัวแบบ AR(1) หมายถึง ต้องใช้จำนวนผู้ป่วยโรคโปลิโอทั่วประเทศในอดีตย้อนหลัง 1 คาบเวลา

ในการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอทั่วประเทศ ณ คาบเวลาถัดไป สมการพยากรณ์ คือ

$$Z_t = 0.5180 + 0.5460\ Z_{t-1}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรคโปลิโอทั่วประเทศ ณ คาบเวลา t

2. ตัวแบบทางสถิติที่เหมาะสมสำหรับการอธิบายและพยากรณ์จำนวนผู้ป่วยแยกตามรายภาค

2.1 โรควัณโรคในกรุงเทพมหานคร พยากรณ์โดยใช้วิธีของบอกซ์ และเจนกินส์ ใช้ตัวแบบใช้ AR(1) หมายถึง ต้องใช้จำนวนผู้ป่วยโรควัณโรคในกรุงเทพมหานคร ซึ่งเป็น ค่าในอดีตย้อนหลัง 1 คาบเวลา ในการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในกรุงเทพมหานคร สมการพยากรณ์คือ

$$Z_t = 64.8851 + 0.4226Z_{t-1}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรควัณโรคในกรุงเทพมหานคร ที่มีความแปรปรวน และค่าเฉลี่ยคงที่ และเป็นจำนวนผู้ป่วยที่แท้จริง ณ คาบเวลา t

2.2 โรควัณโรคในภาคกลาง พยากรณ์โดยใช้วิธีของ บอกซ์และเจนกินส์ ใช้ตัวแบบ $ARIMA(0,0,1)*(0,1,1)_{12}$ หมายถึง ต้องใช้ค่าคลาดเคลื่อนจากการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในภาคกลางในอดีต ณ ตำแหน่งที่ไม่ใช่ฤดูกาลย้อนหลัง 1 คาบเวลา , ใช้ค่าความคลาดเคลื่อนจากการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในภาคกลางในอดีต ณ ตำแหน่งฤดูกาลย้อนหลัง 1 คาบเวลา และหาค่าความแตกต่างแบบมีฤดูกาล 1 ครั้ง ในการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในภาคกลาง สมการพยากรณ์ คือ

$$Z_t = 0.6138e_{t-1} - 0.8176e_{t-12} - 0.5018e_{t-13}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรควัณโรคในภาคกลาง ที่ผ่านการแปลงข้อมูลด้วยล็อกการิทึม และ หาค่าความแตกต่างแบบมีฤดูกาล 1 ครั้ง

e_t คือ ค่าคลาดเคลื่อนจากการพยากรณ์ ณ คาบเวลาที่ t

2.3 โรควัณโรคในภาคเหนือ พยากรณ์โดยใช้วิธีของบอกซ์และเจนกินส์ ใช้ตัวแบบ $ARIMA(0,1,1)_{12}$ หมายความว่า ในการพยากรณ์ต้องใช้ค่าคลาดเคลื่อนจากการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในภาคเหนือในอดีต ณ ตำแหน่งที่เป็นฤดูกาลย้อนหลัง 1 คาบเวลา สมการพยากรณ์คือ

$$Z_t = -0.5234e_{t-12}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรควัณโรคในภาคเหนือ ที่ผ่านการแปลงข้อมูลด้วยรากที่สอง และหาค่าความแตกต่างแบบมีฤดูกาล 1 ครั้ง

e_t คือ ค่าคลาดเคลื่อนจากการพยากรณ์ ณ คาบเวลาที่ t

2.4 โรควัณโรคในภาคตะวันออกเฉียงเหนือ พยากรณ์โดยใช้วิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้คือ $ARIMA(0,1,1)*(0,1,1)_{12}$ หมายความว่า ในการพยากรณ์ต้องใช้ค่าคลาดเคลื่อนในอดีตจากการพยากรณ์จำนวนผู้ป่วยโรควัณโรคในภาคเหนือ ณ ตำแหน่งที่ไม่ใช่ฤดูกาลย้อนหลัง 1 คาบเวลา และหาค่าความแตกต่างแบบปกติไม่มีฤดูกาล 1 ครั้ง , ต้องใช้ค่าคลาดเคลื่อนในอดีตจากการพยากรณ์จำนวนผู้ป่วย ณ ตำแหน่งที่เป็นฤดูกาล 1 คาบเวลา และหาค่าความแตกต่างแบบมีฤดูกาล 1 ครั้ง สมการพยากรณ์คือ

$$Z_t = -0.4039e_{t-1} - 0.5746e_{t-12} - 0.2321e_{t-13}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรควัณโรคในภาคตะวันออกเฉียงเหนือที่ผ่านการแปลงข้อมูลด้วยล็อกการิทึม และหาค่าความแตกต่างแบบมีฤดูกาลและไม่มีฤดูกาล

e_t คือ ค่าคลาดเคลื่อนจากการพยากรณ์ ณ คาบเวลา t

2.5 โรควัณโรคในภาคใต้ พยากรณ์โดยใช้วิธีของ บอกซ์และเจนกินส์ ใช้ตัวแบบ AR(1) หมายความว่า ในการพยากรณ์ต้องใช้จำนวนผู้ป่วยโรควัณโรคในภาคใต้ที่เป็นค่าในอดีตย้อนหลัง 1 คาบเวลา สมการพยากรณ์ คือ

$$Z_t = 188.4071 + 03377Z_{t-1}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรควัณโรคภาคใต้ ณ คาบเวลา ที่ t

2.6 โรคคอตีบในกรุงเทพมหานคร พยากรณ์โดยใช้วิธีของ บอกซ์และเจนกินส์ ใช้ตัวแบบ AR(1) หมายความว่า ในการพยากรณ์ต้องใช้จำนวนผู้ป่วยโรคคอตีบในกรุงเทพมหานครที่เป็นค่าในอดีตย้อนหลัง 1 คาบเวลา สมการพยากรณ์ คือ

$$Z_t = 0.5911\ Z_{t-1}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรคคอตีบในกรุงเทพมหานคร ที่ผ่านการแปลงข้อมูลด้วยรากที่สี่ ณ คาบเวลาที่ t

2.7 โรคคอตีบในภาคกลาง พยากรณ์โดยใช้วิธีของ บอกซ์และเจนกินส์ ซึ่งได้ใช้ตัวแบบ AR(1) หมายความว่า ในการพยากรณ์ต้องใช้จำนวนผู้ป่วยโรคคอตีบในภาคกลางที่เป็นค่าในอดีตย้อนหลัง 1 คาบเวลา สมการพยากรณ์ คือ

$$Z_t = 0.4499 + 0.6465\ Z_{t-1}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรคคอตีบในภาคกลาง ที่ผ่านการแปลงข้อมูลด้วยรากที่สี่ ณ คาบเวลาที่ t

2.8 โรคคอตีบในภาคเหนือ พยากรณ์โดยใช้วิธีปรับเรียบเอกซ์โปเนนเชียลครั้งเดียว หมายความว่า จำนวนผู้ป่วยในอนาคตไม่มีแนวโน้มเพิ่มขึ้นหรือลดลง มีการเคลื่อนรอบค่าเฉลี่ยค่าหนึ่ง โดยสมการพยากรณ์ คือ

$$F_{t+1} = S_t$$

โดย

$$S_t = 0.88x_t + 0.22S_{t-1}$$

เมื่อ F_{t+1} คือ ค่าพยากรณ์จำนวนผู้ป่วยในอนาคต ณ คาบเวลา t+1

x_t คือ จำนวนผู้ป่วยโรคคอตีบในภาคเหนือ ณ คาบเวลา t

S_t คือ ค่าประมาณทำให้เรียบ เป็นค่าที่กำหนดน้ำหนักของการเฉลี่ย

S_{t-1} คือ ค่าของการทำให้เรียบแบบเอกโปเนนเชียลครั้งเดียว ณ คาบเวลาที่ t-1

2.9 โรคคอตีบในภาคตะวันออกเฉียงเหนือ พยากรณ์โดยวิธีการปรับแนวโน้ม และฤดูกาลของวินเตอร์แบบคูณ หมายความว่า จำนวนผู้ป่วยในอนาคตมีการเพิ่มขึ้นหรือลดลงอย่างเป็นอัตราส่วนกับแนวโน้ม ใช้ตัวแบบ

$$X_t = (\beta_0 + \beta_1 t)SN_t + e_t$$

เมื่อ X_t คือ ค่าสังเกต ณ เวลา t

β_0 คือ พารามิเตอร์การปรับระดับ

β_1 คือ พารามิเตอร์การปรับแนวโน้ม

SN_t คือ ปัจจัยฤดูกาล

ซึ่งมีตัวประมาณค่า 3 ตัว กำหนดโดยสมการ ดังนี้

$$S_t = \frac{\alpha X_t}{I_{t-L}} + (1-\alpha)(S_{t-1}+b_{t-1})$$

$$b_t = \delta(S_t-S_{t-1}) + (1-\delta)b_{t-1}$$

$$I_t = \frac{\beta X_t}{S_t} + (1-\beta)I_{t-L}$$

ซึ่งค่า $\alpha = 0.1, \beta = 0.1, \delta = 1.0$

เมื่อ S_t คือ ตัวประมาณค่าของ β_0

b_t คือ ตัวประมาณค่าของ β_1

I_t คือ ตัวประมาณค่าของ SN_t

ดังนั้น สมการพยากรณ์ที่ได้คือ

$$F_{t+m} = (S_t + b_t m)I_{t-L+m}$$

2.10 โรคคอตีบในภาคใต้ พยากรณ์โดยวิธีของ บอกซ์และเจนกินส์ ใช้ตัวแบบ $AR(1)_8$ หมายความว่า ในการพยากรณ์ต้องใช้จำนวนผู้ป่วยโรคคอตีบในภาคใต้ในอดีตย้อนหลัง 1 คาบเวลาฤดูกาล สมการพยากรณ์ คือ

$$Z_t = 0.4893Z_{t-8}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรคคอตีบในภาคใต้ ที่หาค่าความแตกต่างแบบมีฤดูกาล 1 ครั้ง ณ คาบเวลาที่ t

2.11 โรคไอกรนในกรุงเทพมหานคร พยากรณ์โดยวิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้คือ AR(1) โดยทำการแปลงข้อมูลด้วยรากที่สี่ หมายความว่า ต้องใช้ค่า

รากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในกรุงเทพหนึ่งคาบเวลาย้อนหลัง ในการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในกรุงเทพ ณ คาบเวลาถัดไป สมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในกรุงเทพ คือ

$$Z_t = 0.8356 + 0.4801Z_{t-1}$$

เมื่อ Z_t คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในกรุงเทพ ณ คาบเวลา t

$x_t = Z_t^4$

x_t คือ จำนวนผู้ป่วยโรคไอกรนในกรุงเทพ ณ คาบเวลา t

2.12 โรคไอกรนในภาคกลาง พยากรณ์โดยวิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้คือ AR(2) โดยทำการแปลงข้อมูลด้วยรากที่สี่ หมายความว่า ต้องใช้ค่ารากที่สี่ ของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง 2 ค่าในการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง ณ คาบเวลาต่อไป สมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง คือ

$$Z_t = 0.2563 + 0.4379Z_{t-1} + 0.4116Z_{t-2}$$

เมื่อ Z_t คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง ณ คาบเวลา t

$x_t = Z_t^4$

x_t คือ จำนวนผู้ป่วยโรคไอกรนในภาคกลาง ณ คาบเวลา t

2.13 โรคไอกรนในภาคเหนือ พยากรณ์ด้วยวิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้คือ AR(1) โดยทำการแปลงข้อมูลด้วยรากที่สอง หมายความว่า ต้องใช้ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคกลางย้อนหลัง 1 คาบเวลาในการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคเหนือ ณ คาบเวลาต่อไปสมการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคเหนือ คือ

$$Z_t = 0.8023 + 0.7109Z_{t-1}$$

เมื่อ Z_t คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคเหนือ ณ คาบเวลา t

$X_t = Z_t^2$

X_t คือ จำนวนผู้ป่วยโรคไอกรนในภาคเหนือ ณ คาบเวลา t

2.14 โรคไอกรนในภาคตะวันออกเฉียงเหนือ พยากรณ์ด้วยวิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้คือ AR(1) โดยทำการแปลงข้อมูลด้วยรากที่สี่ หมายความว่าต้องใช้ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคตะวันออกเฉียงเหนือย้อนหลัง 1 คาบเวลา ในการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาถัดไป สมการพยากรณ์ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคตะวันออกเฉียงเหนือ คือ

$$Z_t = 0.4753 + 0.7905Z_{t-1}$$

เมื่อ Z_t คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคตะวันออกเฉียงเหนือ ณ คาบเวลา t

$X_t = Z_t^4$

X_t คือ จำนวนผู้ป่วยโรคไอกรนในภาคตะวันออกเฉียงเหนือ ณ คาบเวลา t

2.15 โรคไอกรนในภาคใต้พยากรณ์ด้วยวิธีของ บอกซ์และเจนกินส์ ใช้ตัวแบบ AR(2) โดยทำการแปลงข้อมูลด้วยรากที่สอง หมายความว่า ต้องใช้ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคใต้ย้อนหลัง 2 คาบเวลาในการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคใต้ ณ คาบเวลาต่อไป สมการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคใต้ คือ

$$Z_t = 0.4261 + 0.5264Z_{t-1} + 0.3205Z_{t-2}$$

เมื่อ Z_t คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคไอกรนในภาคใต้ ณ คาบเวลา t

$X_t = Z_t^2$

x_t คือ จำนวนผู้ป่วยโรคไอกรนในภาคใต้ ณ คาบเวลา t

2.16 โรคบาดทะยักในกรุงเทพมหานคร พยากรณ์ด้วยวิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้ คือ ตัวแบบการถดถอยในตนเองอันดับที่ 1 โดยทำการปรับค่าเฉลี่ยให้คงที่ด้วยการหาค่าแตกต่างแรก (first order Autoregressive, first difference model) หมายความว่า ต้องใช้ค่าแตกต่างแรกของจำนวนผู้ป่วยโรคบาดทะยักในกรุงเทพมหานครในอดีต 1 คาบเวลา ในการพยากรณ์ค่าแตกต่างแรกของจำนวนผู้ป่วยโรคบาดทะยักในกรุงเทพมหานคร ณ คาบเวลาถัดไป สมการพยากรณ์จำนวนผู้ป่วยโรคบาดทะยักในกรุงเทพมหานครคือ

$$Z_t = -0.7317\ Z_{t-1}$$

เมื่อ Z_t คือ อนุกรมเวลาที่มีความแปรปรวนและค่าเฉลี่ยคงที่

2.17 โรคบาดทะยักในภาคกลาง พยากรณ์ด้วยวิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้คือ AR(1) โดยทำการแปลงข้อมูลด้วยรากที่สอง หมายความว่าต้องใช้ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคกลางย้อนหลัง 1 คาบเวลา ในการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง ณ คาบเวลาถัดไป สมการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคกลาง คือ

$$Z_t = 3.5730 + 0.1129Z_{t-1}$$

เมื่อ Z_t คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคบาดทะยักในภาคกลาง ณ คาบเวลา t

$X_t = Z_t^2$

x_t คือ จำนวนผู้ป่วยโรคบาดทะยักในภาคกลาง ณ คาบเวลา t

2.18 โรคบาดทะยักในภาคเหนือ พยากรณ์ด้วยวิธีของ บอกซ์และเจนกินส์ ตัวแบบที่ใช้คือ AR(1) โดยทำการแปลงข้อมูลด้วยรากที่สอง หมายความว่า ต้องใช้ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคเหนือย้อนหลัง 1 คาบเวลา ในการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคเหนือ ณ คาบเวลาต่อไป สมการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคไอกรนในภาคเหนือ คือ

$$Z_t = 1.9677 + 0.5576Z_{t-1}$$

เมื่อ Z_t คือ ค่ารากที่สี่ของจำนวนผู้ป่วยโรคบาดทะยักในภาคเหนือ ณ คาบเวลา t

$X_t = Z_t^2$

X_t คือ จำนวนผู้ป่วยโรคบาดทะยักในภาคเหนือ ณ คาบเวลา t

2.19 โรคบาดทะยักในภาคตะวันออกเฉียงเหนือ พยากรณ์โดยใช้วิธีการปรับเรียบแบบเอกโปเนนเชียลครั้งเดียว วิธีของบราวน์ (Brown's) ที่ $\alpha = 0.14$ โดยใช้สมการพยากรณ์คือ

$$F_{t+1} = 0.14x_t + 0.86S_{t-1}$$

เมื่อ F_{t+1} คือ ค่าพยากรณ์ล่วงหน้า ณ คาบเวลา t+1

x_t คือ จำนวนผู้ป่วยโรคบาดทะยัก ในภาคตะวันออกเฉียงเหนือ ณ คาบเวลา t

S_{t-1} คือ ค่าของการปรับให้เรียบแบบเอกโปเนนเชียลครั้งเดียว ณ คาบเวลา t-1

2.20 โรคบาดทะยักในภาคใต้ พยากรณ์โดยใช้วิธีการปรับเรียบแบบ

เอกโป่เนนเชียลครั้งเดียววิธีของบราวน์ ที่ $\alpha = 0.22$ โดยใช้สมการพยากรณ์คือ

$$F_{t+1} = 0.22x_t + 0.78S_{t-1}$$

เมื่อ F_{t+1} คือ ค่าพยากรณ์ล่วงหน้า ณ คาบเวลา t+1

x_t คือ จำนวนผู้ป่วยโรคบาดทะยัก ในภาคใต้ ณ คาบเวลา t

S_{t-1} คือ ค่าของการปรับให้เรียบแบบเอกโป่เนนเชียลครั้งเดียว ณ คาบเวลา t-1

2.21 โรคโปลิโอในกรุงเทพมหานคร พยากรณ์ด้วยวิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้ คือ ตัวแบบการถดถอยในตนเองอันดับที่ 1 (first order Autoregressive model) หมายความว่าต้องใช้จำนวนผู้ป่วยโรคโปลิโอในกรุงเทพมหานครในอดีตหนึ่งค่าในการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในกรุงเทพมหานคร ณ คาบเวลาถัดไป สมการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอทั่วประเทศ คือ

$$Z_t = 0.0943 - 0.0943\ Z_{t-1}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรคโปลิโอในกรุงเทพมหานคร ณ คาบเวลา t

2.22 โรคโปลิโอในภาคกลาง พยากรณ์ด้วยวิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้คือ ตัวแบบการถดถอยในตนเองอันดับที่ 1 (first order Autoregressive model) โดยทำการแปลงข้อมูลด้วยรากที่สอง หมายความว่า ต้องใช้ค่ารากที่สองของจำนวนผู้ป่วยโรคโปลิโอในภาคกลางในอดีตหนึ่งค่าในการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคโปลิโอในภาคกลาง สมการพยากรณ์ค่ารากที่สองของจำนวนผู้ป่วยโรคโปลิโอในภาคกลาง คือ

$$Z_t = 0.6210 + 0.4071\ Z_{t-1}$$

เมื่อ Z_t คือ ค่ารากที่สองของจำนวนผู้ป่วยโรคโปลิโอในภาคกลาง ณ คาบเวลา t

$X_t = Z^2_t$

X_t คือ จำนวนผู้ป่วยโรคโปลิโอในภาคกลาง ณ คาบเวลา t

2.23 โรคโปลิโอในภาคเหนือ พยากรณ์ด้วยวิธีของบอกซ์และเจนกินส์ ตัวแบบที่ใช้ คือ ตัวแบบการถดถอยในตนเองอันดับที่ 1 (first order Autoregressive model) หมายความว่า ต้องใช้จำนวนผู้ป่วยโรคโปลิโอในภาคเหนือในอดีตย้อนหลัง 1 คาบเวลา ในการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในภาคเหนือ ณ คาบเวลาถัดไป สมการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในภาคเหนือ คือ

$$Z_t = 0.1424 + 0.2766\ Z_{t-1}$$

เมื่อ Z_t คือ จำนวนผู้ป่วยโรคโปลิโอในภาคเหนือ ณ คาบเวลา t

2.24 โรคโปลิโอในภาคตะวันออกเฉียงเหนือ พยากรณ์โดยวิธีปรับเรียบแบบเอกซ์โปเนนเชียลครั้งเดียวของบราวน์ โดยมีค่าที่กำหนดน้ำหนักการปรับเรียบแบบเอกซ์โปเนนเชียล (α) เท่ากับ 0.19 สมการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในภาคตะวันออกเฉียงเหนือ ณ คาบเวลาที่ t+1 คือ

$$F_{t+1} = (0.19)x_t + (0.81)S_{t-1}$$

เมื่อ x_t คือ จำนวนผู้ป่วยโรคโปลิโอในภาคตะวันออกเฉียงเหนือ ณ คาบเวลา t

S_{t-1} คือ ค่าของการทำให้เรียบแบบเอกซ์โปเนนเชียลครั้งเดียว ณ คาบเวลา t-1

2.25 โรคโปลิโอในภาคใต้พยากรณ์ โดยวิธีปรับเรียบแบบเอกซ์โปเนนเชียลซ้ำสามครั้ง วิธีของบราวน์ โดยมีค่ากำหนดที่น้ำหนักการปรับเรียบแบบเอกซ์โปเนนเชียล (α) เท่ากับ 0.01 สมการพยากรณ์จำนวนผู้ป่วยโรคโปลิโอในภาคใต้ ณ คาบเวลาที่ $t+T$ คือ

$$F_{t+T} = a_t + b_tT + (1/2)C_tT^2$$

ซึ่ง $a_t = 0$

$b_t = 0$

$C_t = 0$

ดังนั้น

$F_{t+T} = 0$

ตารางที่ 5.1 แสดงค่าพยากรณ์เปรียบเทียบกับค่าจริง ของจำนวนผู้ป่วยโรคต่าง ๆ ทั่วประเทศ ตั้งแต่เดือนมกราคม ถึง ธันวาคม พ.ศ.2535

เดือน	ชื่อโรค	เทคนิคที่ใช้	ค่าพยากรณ์	ค่าสังเกตจริง
มกราคม	วัณโรค	บอกซ์ และเจนกินส์	1860	2223
กุมภาพันธ์			1543	2000
มีนาคม			1624	1979
เมษายน			1392	1772
พฤษภาคม			1463	1710
มิถุนายน			1436	1837
กรกฎาคม			1564	1914
สิงหาคม			1651	1901
กันยายน			1470	1897
ตุลาคม			1475	1703
พฤศจิกายน			1346	1470
ธันวาคม			1162	1125
มกราคม	คอตีบ	บอกซ์ และเจนกินส์	5	5
กุมภาพันธ์			5	3
มีนาคม			6	3
เมษายน			6	1
พฤษภาคม			6	-
มิถุนายน			7	3
กรกฎาคม			7	4
สิงหาคม			7	8

ตาราง 5.1 (ต่อ)

เดือน	ชื่อโรค	เทคนิคที่ใช้	ค่าพยากรณ์	ค่าสังเกตจริง
กันยายน	คอตีบ	บอกซ์และเจนกินส์	7	4
ตุลาคม			7	1
พฤศจิกายน			7	4
ธันวาคม			8	3
มกราคม	ไอกรน	บอกซ์และเจนกินส์	20	22
กุมภาพันธ์			25	24
มีนาคม			30	35
เมษายน			34	28
พฤษภาคม			38	27
มิถุนายน			41	37
กรกฎาคม			44	56
สิงหาคม			46	54
กันยายน			49	53
ตุลาคม			50	34
พฤศจิกายน			52	25
ธันวาคม			53	18
มกราคม	บาดทะยัก	ปรับแนวโน้มและ	51	48
กุมภาพันธ์		ฤดูกาลของวินเตอร์	50	47
มีนาคม			49	68
เมษายน			48	51
พฤษภาคม			47	57

ตาราง 5.1 (ต่อ)

เดือน	ชื่อโรค	เทคนิคที่ใช้	ค่าพยากรณ์	ค่าสังเกตจริง
มิถุนายน	บาดทะยัก	ปรับแนวโน้มและ	46	58
กรกฎาคม		ฤดูกาลของวินเตอร์	45	50
สิงหาคม			44	49
กันยายน			43	45
ตุลาคม			42	53
พฤศจิกายน			41	31
ธันวาคม			40	31
มกราคม	โปลิโอ	บอกซ์และเจนกินส์	1	-
กุมภาพันธ์			1	1
มีนาคม			1	-
เมษายน			1	-
พฤษภาคม			1	1
มิถุนายน			1	1
กรกฎาคม			1	1
สิงหาคม			1	1
กันยายน			1	1
ตุลาคม			1	1
พฤศจิกายน			1	1
ธันวาคม			1	-

ข้อเสนอแนะ

1. วิธีวิเคราะห์อนุกรมเวลาของบอกซ์และเจนกินส์ เมื่อได้ข้อมูลจริงเพิ่มขึ้น จะต้องนำมาปรับปรุงค่าพยากรณ์อยู่เสมอ จะทำให้ค่าพยากรณ์ที่ได้ ใกล้ค่าจริงมากยิ่งขึ้น

2. ค่าพยากรณ์ที่ได้จากตัวแบบที่เลือกใช้ อาจจะมีค่าต่างจากค่าจริงค่อนข้างมาก สาเหตุก็คือ ลักษณะข้อมูลในปัจจุบันอาจไม่เหมือนกับลักษณะของข้อมูลในอดีต มีการเปลี่ยนแปลงสภาวะแวดล้อม ดังนั้น ไม่ควรนำตัวแบบที่ได้ใช้ในการพยากรณ์ในระยะเวลาที่ห่างจากคาบเวลาที่มีข้อมูลอยู่มากนัก และควรจะต้องติดตามผลที่เกิดขึ้นจริง เป็นระยะ ๆ ถ้ามีความแตกต่างก็ควรจะต้องมีการปรับปรุง แก้ไข ต่อไป

3. หากมีผู้สนใจทำการวิจัยต่อไป มีโรคที่น่าสนใจคือ โรคไวรัสตับอักเสบชนิดเอ และบี โดยเฉพาะไวรัสตับอักเสบชนิดบี เป็นโรคที่เป็นแล้วต้องใช้เวลาพักฟื้นนาน และรักษาไม่หาย ต้องระวัง รักษาสุขภาพของตนเองตลอดเวลา ถ้าหากรู้แนวโน้มการเกิดโรคอาจจะหาทางป้องกันได้ดีขึ้น

4. โปรแกรมสำเร็จรูป STATGRAPHIC ไม่เหมาะกับข้อมูลที่จะใช้วิธีที่มีค่าพารามิเตอร์หลายตัว แต่เหมาะกับข้อมูลที่ใช้วิธีที่มีค่าพารามิเตอร์น้อย ๆ โปรแกรมสำเร็จรูป MINITAB เหมาะกับข้อมูลที่จะใช้วิธีของบอกซ์และเจนกินส์ ส่วนโปรมแกรมสำเร็จรูป FORECAST PLUS เหมาะกับข้อมูลทุกประเภท

ภาคผนวก

ภาคผนวก ก

การใช้โปรแกรมสำเร็จรูป FORECAST PLUS ในการวิเคราะห์อนุกรมเวลา

ความรู้เบื้องต้น

โปรแกรมสำเร็จรูป FORECAST PLUS เป็นโปรแกรมสำเร็จรูปที่มีเทคนิค การพยากรณ์ ให้เลือกใช้ 13 วิธี โดยสามารถใช้โปรแกรมนี้ในการพิจารณาส่วนประกอบของข้อมูลอนุกรมเวลา ลักษณะข้อมูล วิเคราะห์ข้อมูล และการหาค่าพยากรณ์

ขั้นตอนเบื้องต้นในการใช้โปรแกรม FORECAST PLUS

สำหรับคอมพิวเตอร์ที่มี drive มากกว่าสอง และได้เก็บโปรแกรม FORECAST PLUS ไว้ใน drive C (hard disk) ให้ทำดังนี้

1.เปิดเครื่อง และเข้าสู่ drive C

2.ใส่ key disk ใน drive A

3.กด CD FORECAST <enter>

4.เมื่อขึ้น C:FORECAST แล้วให้กด FORECAST <enter> อีกครั้ง

5.จะปรากฏ Main Menu ของ FORECAST PLUS

ถ้าคอมพิวเตอร์เป็นระบบ สอง drive ให้ทำดังนี้

1.ใส่แผ่น DOS ใน drive A แล้ว set คอมพิวเตอร์ใหม่

2.ใส่ key disk ใน drive A

3.ใส่แผ่นข้อมูล ใน drive B

4.กด FP

5.จะปรากฏ Main menu ของ FORECAST PLUS

ส่วนประกอบของโปรแกรม FORECAST PLUS

โปรแกรมสำเร็จรูป FORECAST PLUS ประกอบด้วย 4 ส่วน คือ โปรแกรมสำหรับเตรียมข้อมูล โปรแกรมสำหรับการพิจารณาลักษณะข้อมูล โปรแกรมสำหรับการวิเคราะห์ข้อมูลและหาค่าพยากรณ์ และโปรแกรม Batch Mode เลือกใช้โปรแกรมย่อย 4 โปรแกรมนี้จากโปรแกรมหลัก (Main Menu) ซึ่งจะปรากฏบนจอดังนี้

THE MAIN MENU

Forecast Plus is devided into four components. The main menu selection looks like this :

1. Data Management
2. Exploratory Package
3. Forecasting Analysis
4. Batch Mode
5. End the Program

which ?

เลือกส่วนต่างๆด้วยการพิมพ์เลขที่ส่วนของโปรแกรมที่ต้องการ แล้วกด <enter> เช่น ต้องการสร้างไฟล์ข้อมูล ให้พิมพ์เลข 1 <enter> ต้องการพยากรณ์ข้อมูล ให้พิมพ์เลข 3 <enter> เป็นต้น และพิมพ์เลข 5 เมื่อต้องการเลิกใช้โปรแกรม FORECAST PLUS และกด <enter>

1. Data Management

โปรแกรมย่อยส่วนแรกของโปรแกรม FORECAST PLUS คือ Data Management ใช้สำหรับจัดเตรียมไฟล์ข้อมูล ประกอบด้วยโปรแกรมย่อย 9 โปรแกรมกำกับด้วยเลข 1 ถึง 9 ซึ่งปรากฏในจอ เมื่อได้เลือกเลข 1 ใน Main Menu แล้ว ดังนี้

Data file name :(NONE)

*** Data Management Menu ***

1. Enter new data
2. Edit an existing data file
3. Edit an existing labels file
4. Print a data file
5. Transform a data file
6. Trading day adjustment
7. Restructure/merge data files
8. Read/Write DIF format files
9. Change parameter table
10. Batch editer
11. Return to main menu

which ?

เลือกเลขที่กำกับหน้าโปรแกรมที่ต้องการ แล้ว <enter> เช่น ต้องการสร้างไฟล์ข้อมูลใหม่ให้กดเลข 1 <enter> ต้องการให้พิมพ์ผลทางหน้าจอหรือเครื่องพิมพ์ ให้กดเลข 9 (พิมพ์ CON เพื่อให้ผลแสดงออกทางจอ และพิมพ์ PRN เพื่อให้ผลออกทางเครื่องพิมพ์ หลังจากกดเลข 9 แล้ว) กด 11 แล้ว <enter> เมื่อต้องการกลับไปที่ Main Menu เพื่อเลือกใหม่ เป็นต้น

2. Exploratory Package

เป็นส่วนที่ใช้สำหรับการพิจารณาลักษณะข้อมูล เมื่อกดเลข 2 ใน Main Menu จะปรากฏหน้าจอด้วยทางเลือก 7 ทาง คือ

*** Exploratory Package ***

1. Time Plot
2. 4253HT Robust Smoothing
3. Box Plot (trend/cycle variation)
4. Aggregate Box Plot (seasonal variation)
5. Spread vs Level Plot
6. Autocorrelation Function
7. Return to Main Menu

Which ?

ส่วนนี้ประกอบด้วยโปรแกรมย่อย 7 โปรแกรม เมื่อต้องการดูลักษณะการเคลื่อนไหวของข้อมูลให้กดหมายเลข 1 ถ้าต้องการดูลักษณะความแปรปรวนของข้อมูลให้กดเลข 5 เป็นต้น

3. **Forecasting Analysis**

เป็นส่วนที่ใช้ในการวิเคราะห์ข้อมูลและพยากรณ์ หลังจากได้ตรวจดูลักษณะข้อมูลแล้ว กลับไปที่ Main Menu เพื่อเลือกวิธีวิเคราะห์ข้อมูล ตรวจสอบตัวแบบที่ใช้ (กรณีวิเคราะห์ผลด้วยวิธี Box-Jenkins) และพยากรณ์ค่า FORECAST PLUS จะเข้ามาในส่วน Forecasting Analysis เมื่อกดหมายเลข 3 ใน Main Menu หน้าจอจะปรากฏวิธีวิเคราะห์ผลแบบต่างๆ ให้เลือกดังนี้

*** Forecasting Analysis ***

1. Simple Moving Average (unweighted)
2. Single Exponential Smoothing
3. Double Exponential Smoothing
4. Holt's Two-Parameter Smoothing
5. Harrison's Harmonic Smoothing
6. Brown's Quadratic Exponential Smoothing
7. Winter's Seasonal Smoothing
8. Robust Decomposition
9. Census X-11 Decomposition
10. Regression Trend Analysis
11. Multiple Regression Analysis
12. Generalized Adaptive Filtering
13. Box-Jenkins Analysis
14. Return to Main Menu

Which ?

การเตรียมไฟล์ข้อมูล (Data File Management)

เมื่อต้องการสร้างไฟล์ข้อมูลใหม่ให้กดเลข 1 ใน Main Menu แล้ว กด <enter> และกดเลข 1 ใน Data Menagement Menu คือโปรแกรม Enter new data แล้วกด <enter>

ใน FORECAST PLUS มีโปรแกรม data entry สองโปรแกรมแยกจากกัน สำหรับกรณีหนึ่งตัวแปร(singlevariable)และกรณีหลายตัวแปร (multiple variable) การเตรียมไฟล์ข้อมูล ทำได้ดังนี้

1.การกำหนดชื่อ ไฟล์

เมื่อเลือก Enter new data โปรแกรมจะถามชื่อ data file ที่ต้องการเพิ่มถ้าใส่ชื่อไฟล์ที่มีอยู่แล้ว หมายความว่า ต้องการเพิ่มข้อมูลต่อจากข้อมูลตำแหน่งสุดท้ายของไฟล์นั้น ถ้าเป็นชื่อไฟล์ใหม่ที่ต้องการสร้างแต่ยังไม่มีใน directory เครื่องจะถามว่า ต้องการสร้างไฟล์ใหม่หรือไม่ เมื่อตอบว่า ใช่ เครื่องจะให้กำหนดชื่อไฟล์ ซึ่งใช้ได้ทั้งตัวอักษรและตัวเลข แต่ต้องไม่เกิน 8 ตัวอักษรที่เรียงติดต่อกัน ไม่ต้องเว้นวรรค และกำกับด้วยอักษรแสดง drive B เช่น B:filename หมายความว่า แผ่นที่เก็บไฟล์ข้อมูลนี้อยู่ใน drive B แต่ถ้าไม่กำกับ drive หน้าชื่อไฟล์ใด ๆ โปรแกรมจะกำกับไฟล์นั้นด้วย drive ที่กำหนดใน parameter table

ตัวอย่างเช่น สร้างไฟล์ข้อมูลชื่อ TUBER.FC ให้เก็บไว้ในแผ่นที่อยู่ใน drive B จะเป็นดังนี้

B:TUBER.FC

2.การสร้างไฟล์ข้อมูลใหม่

ใน FORECAST PLUS แต่ละไฟล์ข้อมูลจะมีได้ 2 ไฟล์ โดย ไฟล์หนึ่งบรรจุข้อมูล อีกไฟล์หนึ่งบรรจุข้อกำหนดในการ label (labeling information) เช่น สร้างไฟล์ชื่อ STUBER.FC เครื่องจะสร้างไฟล์ชื่อ STUBER.LBL ไปพร้อมกันด้วย ซึ่ง STUBER.LBL นี้คือ label file

ตัวอย่างไฟล์ข้อมูลชื่อ STUBER.FC ใน drive B มี label file ชื่อ B:STUBER.LBL ส่วน information ที่เก็บไว้ใน .LBL file นี้ จะปรากฏบนจอให้กำหนด คือ จำนวนตัวแปร, ชนิดของ label (เป็น วัน เดือน ปี), ปีที่เริ่มต้น (เมื่อ label เป็นปี), ค่าสังเกตเบื้องต้น (เช่นเริ่มปีที่ 1), จำนวนค่าสังเกตระหว่าง label คือ จากปีที่1 ถึง ปีที่ 2 มีค่าสังเกตกี่ค่า (ปีที่อยู่ติดกันมีค่าสังเกตกี่ค่า), label เพิ่มขึ้นขนาดเท่ากับเท่าใด เช่น ใน 1 ปี มี 12 เดือน จำนวนค่าสังเกตเพิ่มขึ้นเดือนละ 1 ค่า เป็นต้น

สรุปผลที่ปรากฏบนหน้าจอ ในการสร้างข้อมูลดังนี้

1. กด 1 <enter> ใน Main Menu เข้าสู่ Data Management Menu
2. กด 1 <enter> ใน Data Management Menu เข้าสู่ Enter New Data
3. ตอบคำถามต่างๆ ใน Data Entry ดังนี้

*** DATA ENTRY ***

What data file would you like to add to ? B:TUBER.FC

The data file B:TUBER.FC does not exist.

Do you want to create it (Y/N) ?

ให้ตอบ Y <enter> หน้าจอจะปรากฏข้อความต่อไปนี้ โดยสดมภ์ขวาสุดเป็นข้อความที่ต้องการใส่เข้าไปในไฟล์ข้อมูล

*** Labels File Creation ***

Creating labels file B:TUBER.LBL:

Number of variables :	1
Type of labeling :	year
Starting label :	1987
Starting observation :	1
Number of observations between each label :	12
Increment of the labels :	1
	<enter>

Press ENTER to accept : <enter>

4.เมื่อ <enter> แล้ว จะปรากฏหน้าจอดังนี้

*** Labels File Creation ***

Variable 1 label : TUBER <enter>

Press ENTER to accept : <enter>

พิมพ์ชื่อตัวอักษรที่ต้องการศึกษา คือ TUBER แล้วกด enter จะปรากฏหน้าจอดังนี้

Record 1 : 1987(1) > 1936 < พิมพ์ข้อมูลตัวที่ 1 <enter>

Record 2 : 1987(2) > 1627 < พิมพ์ข้อมูลตัวที่ 2 <enter>

. .

. .

. .

. .

. .

Record 60: 1991(12) > 1158 < พิมพ์ข้อมูลตัวสุดท้าย <ESC>

ให้กด ESC เมื่อพิมพ์ข้อมูลตัวสุดท้ายแล้ว

5.จะปรากฏหน้าจอดังนี้

---> ESCAPE EXIT

Do you want to :

1. Save the data on disk

2. Re-enter the entry/edit mode

3. Exit without saving the changes

Which ?

ให้กดเลข 1 เพื่อต้องการ save ข้อมูลลงในแผ่น disk แล้ว <enter>

6. จะปรากฏข้อความ

Current data file name = B:TUBER.FC

Enter the new name for data file or

Just press <enter> to save the existing name:B:TUBER.FC

---> Returning to Data Management Menu

กด enter โดยไม่ต้องพิมพ์ข้อความใด ๆ โปรแกรมจะเก็บข้อมูลไว้ที่ B:TUBER.FC โดยอัตโนมัติ คือเก็บข้อมูลไว้ในไฟล์ ชื่อ TUBER.FC ที่ disk drive B

7. เมื่อสร้างไฟล์ข้อมูลเสร็จแล้ว เครื่องจะกลับไปที่ Data Management Menu ให้กดเลข 11 ใน Menu นี้แล้ว <enter> โปรแกรมจะกลับไปที่ Main Menu เพื่อให้เลือกต่อไปว่า จะตรวจลักษณะข้อมูล (กดเลข2) หรือจะวิเคราะห์ผล (กดเลข3) หรือจะออกจาก FORECAST PLUS (กดเลข5) แล้วกด <enter>

3.การแก้ไขไฟล์ข้อมูลที่สร้างด้วยโปรแกรม FORECAST PLUS

เมื่อต้องการแก้ไขไฟล์ข้อมูล เช่น การลบออก การเพิ่มข้อมูล การแก้ไขข้อมูลที่ผิด ที่ record ต่าง ๆ การลดและการเพิ่มข้อมูล ทำได้เฉพาะส่วนท้ายของไฟล์ข้อมูลเท่านั้น ไม่สามารถแทรกข้อมูลระหว่าง record ได้ ถ้าต้องการแทรกข้อมูล จะต้องใช้ editor ตัวอื่น เช่น SPSS หรือ CW

การแก้ไขข้อมูลด้วย FORECAST PLUS ทำได้ดังนี้

1. กด 1 ใน Main Menu <enter> เพื่อเข้าสู่ Data Management

2. กด 2 <enter> ใน Data Management Menu เพื่อ Edit an existing data file

*** Data Editor ***

What data file would you like to edit ?

ให้พิมพ์ชื่อไฟล์ที่ต้องการแก้ไข คือ B:TUBER.FC <enter>

3.จะปรากฏข้อมูลที่จอภาพดังนี้

```
Record  1 : 1988(1)   >1936<
Record  2 : 1988(2)   >1627<
           .         .
           .         .
           .         .
Record 57 : 1992(9)   >1621<
Record 58 : 1992(10)  >1739<
Record 59 : 1992(11)  >1444<
Record 60 : 1992(12)  >1158<
```

4.ถ้าต้องการลบข้อมูลตัวที่ 60 ให้เลื่อน cursor ไปอยู่ที่ตำแหน่งที่ต้องการลบ แล้วกด Back space หรือ Delete ลบทีละตัวอักษร

5.ถ้าต้องการเพิ่มข้อมูลตัวที่ 61-65 ให้เลื่อน cursor ไปอยู่ที่ตำแหน่งที่ต้องการเพิ่ม และพิมพ์ข้อมูลทีละบรรทัด

6.เมื่อลดหรือเพิ่มข้อมูลเรียบร้อยแล้วให้กด Esc เพื่อ save ข้อมูล และหน้าจอจะเป็นดังนี้

```
---> ESCAPE NEXT

Do you want to :

     1. Save the data on disk

     2. Re-enter the entry/edit mode

     3. Exit without saving the changes


        which ?
```

กด 1 <enter> จะปรากฏข้อความดังนี้

```
Current data file name = B:TUBER.FC

Enter the name for the data file

-or- Just press <enter> to use the existing name :
```

7. กด <enter> โปรแกรมจะทำงานโดยอัตโนมัติในการเก็บไฟล์ข้อมูล TUBER.FC ที่ disk drive B หรือจะตั้งชื่อไฟล์ที่แก้ไขใหม่ก็ได้ แล้ว <enter> ชื่อใหม่นั้นจะเก็บไว้ที่ disk drive B

8.เมื่อ save ข้อมูลที่แก้ไขแล้ว โปรแกรมจะกลับไปที่ Data Management Menu ให้กดเลข 11 <enter> เพื่อกลับไปที่ Main Menu

4. การเรียกไฟล์ข้อมูลที่สร้างโดยโปรแกรม SPSS/PC และ CW เข้ามาใช้งานในโปรแกรม FORECAST PLUS

ไฟล์ข้อมูลที่นำมาใช้ในโปรแกรม FORECAST PLUS เป็นได้ 2 แบบ คือ ไฟล์ข้อมูลที่สร้างด้วยโปรแกรม FORECAST PLUS ดังกล่าวแล้ว หรือเป็นไฟล์ข้อมูลที่สร้างด้วยโปรแกรมอื่นๆมาใช้ ทำได้ดังนี้

1. กดเลข 1 <enter> ใน Main Menu เพื่อเข้าสู่ Data Management Menu

2. กดเลข 1 <enter> ใน Data Management Menu เพื่อเข้าสู่ Data Entry

*** DATA ENTRY ***

What data file would you like to add to ?

ให้พิมพ์ชื่อไฟล์ที่สร้างจากโปรแกรมอื่นที่ต้องการเรียกมาใช้ เช่น B:TETAN.SK <enter> เป็นการเรียกไฟล์ชื่อ TETAN.SK ที่อยู่ใน disk drive B

3. หน้าจอจะปรากฏข้อความ

*** DATA ENTRY ***

What data file would you like to add to ? B:TETAN.SK

The tables file B:TETAN.LBL does not exist.

Doyou want to create it (Y/N) ?

พิมพ์ Y <enter>

4.ปรากฏข้อความดังนี้

*** Labels File Creation ***

Creating labels file B:TETAN.LBL

Number of variables :	1
Type of label :	YEAR
Starting labels :	1987
Starting observation :	1
Number of observation between each label :	12
Increment of the labels :	1
	<enter>

Press Enter to accept : <enter>

พิมพ์ label ไฟล์ ตามตัวอย่างสดมภ์ขวาสุด แล้ว <enter>

5. หน้าจอจะปรากฏข้อความต่อไปนี้

*** Labels File Creation ***

variable 1 label :

พิมพ์ชื่อตัวแปรที่ต้องการศึกษา สำหรับไฟล์ข้อมูล TETAN.SK ในที่นี้ตัวแปรคือ NUMBER <enter>

6. ปรากฏข้อความหน้าจอ

```
*** Labels File Creation ***

variable 1 label   : NUMBER    <enter>
   Press ENTER to accept :     <enter>
```

7. ปรากฏข้อมูลหน้าจอ ดังนี้

```
Record  1  :  1    >130<
Record  2  :  2    >136<
        .          .
        .          .
        .          .
Record 100 :  100  >141<
```

8. กด Esc เพื่อ save ข้อมูลที่เรียกมาใช้ในโปรแกรม FORECAST PLUS หน้าจอจะปรากฏข้อความ

```
----> ESCAPE EXIT

      Do you want to :
          1. Save the data on disk
          2. Re-enter the entry/edit mode
          3. Exit without saving the changes
             Which ?
            กด 1 <enter>
```

9. จะมีข้อความเพิ่มขึ้นที่จอภาพ ดังนี้

current data file name B:TETAN.FC

Enter the new name for the data file

-or- Just press <enter> to use the existing name :B:TETAN.FC

---> Returning to Data Management Menu

กด <enter> เครื่องจะ save ข้อมูลชื่อ TETAN.FC ไว้ที่ disk B

10. กดเลข 11<enter> ใน Data Management Menu เพื่อกลับไปที่ Main Menu

พิจารณาลักษณะข้อมูล (Exploratory Data)

หลังจากสร้างไฟล์ข้อมูลเสร็จแล้ว หรือเรียกข้อมูลที่สร้างด้วยโปรแกรมอื่นมาใช้ใน FORECAST PLUS ให้ตรวจลักษณะข้อมูลด้วยการเลือกหมายเลข 2 ใน Main Menu จะปรากฏหน้าจอดังนี้

*** Exploratory Package ***

1. Time Plot
2. 4253HT Robust Smoothing
3. Box Plot (Trend/Cycle variation)
4. Aggregate Box Plot (seasonal variation)
5. Spread vs Level Plot
6. Autocorrelation Function
7. Return to Main Menu

Which ?

เมื่อเลือกหมายเลข 1 หน้าจอจะปรากฏ แผนภาพการกระจายแสดงการเคลื่อนไหวของข้อมูล ผู้วิเคราะห์จะพิจารณาจากภาพว่าข้อมูลมีแนวโน้มฤดูกาล หรือวัฏจักรหรือไม่

หมายเลข 2 เป็นการพิจารณาข้อมูลเมื่อทำข้อมูลให้เรียบแล้ว ด้วยการแปลงข้อมูลด้วย (การกลับเศษเป็นส่วน (Reciprocal) = เลข 1, 1/รากที่2 (1/Sqrt) = เลข2, 1/รากที่4 (1/4th root) = เลข 3, ล็อกการิทึม (Log) = เลข 4, รากที่ 4 (4th root) = เลข 5, รากที่ 2 (Sqrt) = เลข 6 และไม่แปลง (None) = เลข 7) เมื่อเลือกการแปลงข้อมูลด้วยวิธีใดวิธีหนึ่ง (หมายเลข1-7) แล้วจะปรากฏกราฟของข้อมูลเศษตกค้าง (Residual) ซึ่งเป็นผลต่างของข้อมูลเดิมกับข้อมูลที่ทำให้เรียบแล้ว

*** 4253HT Robust Smoothing ***

Transformation to original data series : 7

(1=Reciprocal,2=1/Sqrt,3=1/4th root,

4=Log,5=4th root,6=Sqrt,7=None)

Time Plot of rough : Y

Press ENTER to accept :

หมายเลข 3 แสดงการกระจายของค่าสังเกตเป็นกลุ่มๆ แต่ละกลุ่มอาจจะประกอบด้วยค่าสังเกต 12 ค่าหรือ 7 ค่า ฯลฯ แสดงด้วยรูปสี่เหลี่ยมผืนผ้า มีขอบเขตจากควอไทล์ที่ 1-3 ค่าเฉลี่ยของกลุ่มอยู่ภายในกรอบสี่เหลี่ยมนี้ จุดปลายสุดสองด้านนอกกรอบสี่เหลี่ยม คือ ค่าต่ำสุด และสูงสุดของข้อมูลแต่ละกลุ่ม ดังรูปตัวอย่าง

NUMBER - EXPLORATORY DATA ANALYSIS

BOX PLOT

```
         1094                         1631                                2287
       :-^-----------------------------^------------------------------------^-:
       :                                                                      :
       :                                                                      :
       :                 |---------|                                          :
1987   : x            |--|    *    |-|               x                        :
       :                 |---------|                                          :
       :                                                                      :
       :                    |-------------|                                   :
1988   :            |-------|      *      |--------|                          :
       :                    |-------------|                                   :
       :                                                                      :
       :                    |-----------|                                     :
1989   :               |----|.   *      |--------|            x               :
       :                    |-----------|                                     :
       :                                                                      :
       :                         |-----------|                                :
1990   :         |---------------|     *     |--|                   x         :
       :                         |-----------|                                :
       :                                                                      :
       :                            |-------------|                           :
1991   :    x              |--------|      *      |----|                   x  :
       :                            |-------------|                           :
       :                                                                      :
       :----------------------------------------------------------------------:
```

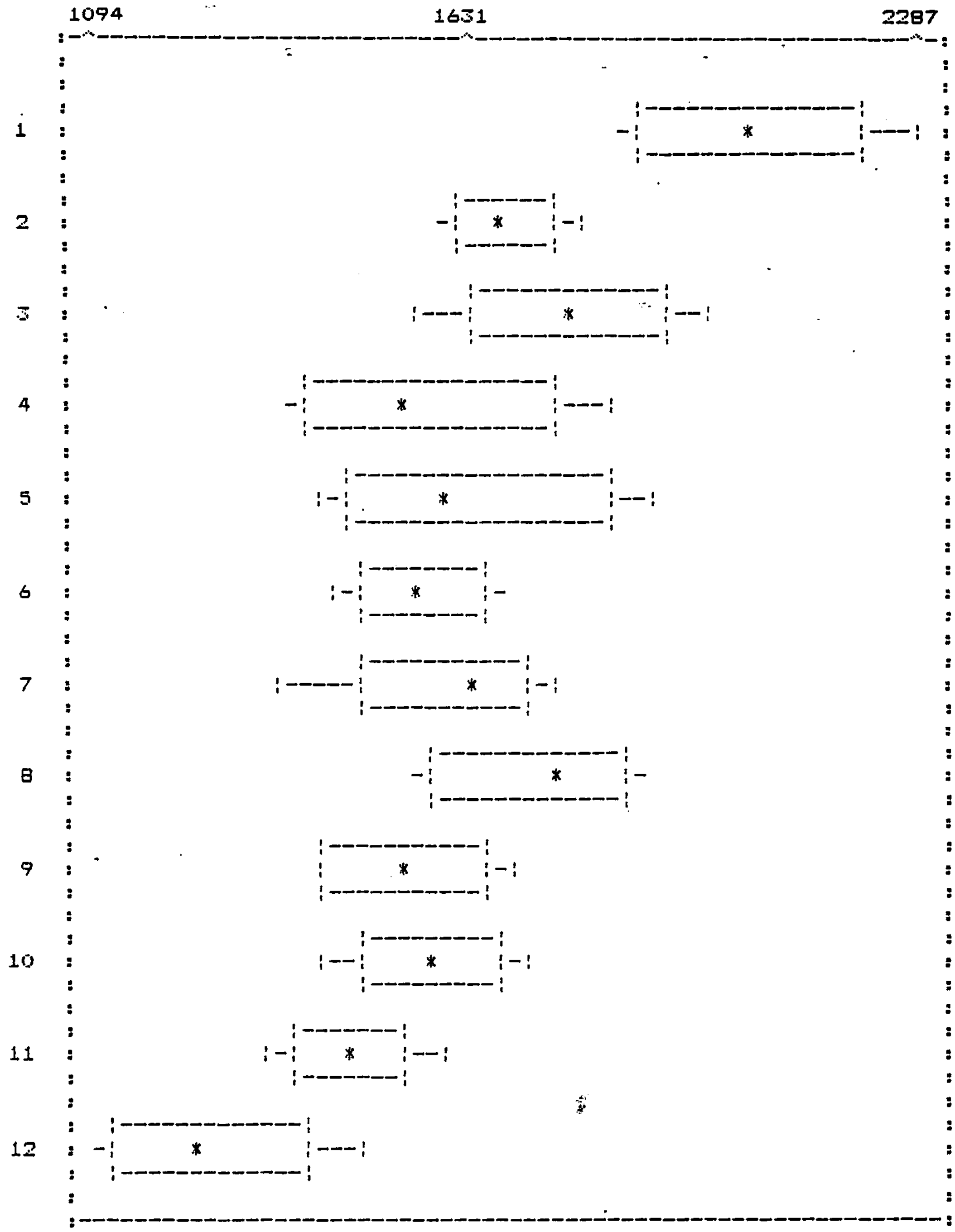
* = Midmean
x = Outlier
Each box summarizes 12 observations

ใช้หมายเลข 3 ในการพิจารณาว่าข้อมูลมีแนวโน้ม หรือวัฏจักรหรือไม่

หมายเลข 4 ใช้สำหรับพิจารณาความผันแปรตามฤดูกาล แสดงข้อมูลเป็นกลุ่ม แต่ละกลุ่มประกอบด้วยค่าสังเกตจำนวนเท่ากับความยาวของฤดูกาล เช่น 4 ค่าสังเกต , 12 ค่าสังเกต เป็นต้น กรอบสี่เหลี่ยมแต่ละรูปมีขอบเขตจากควอไทล์ที่ 1-3 ค่าเฉลี่ยแต่ละกลุ่ม อยู่ภายในกรอบสี่เหลี่ยม และแสดงค่าต่ำสุดกับค่าสูงสุดของแต่ละกลุ่ม ดังรูป

NUMBER - EXPLORATORY DATA ANALYSIS

AGGREGATE BOX PLOT



* = Midmean
Each box summarizes 5 observations

หมายเลข 5 ใช้ในการดูว่าความแปรปรวนของข้อมูลคงที่หรือไม่ หมายเลข 5 ใน Menu Exploratory Package จะปรากฏข้อความ

*** Spread vs Level Plot ***

Transformation to original data series : 7

(1=Reciprocal,2=1/Sqrt,3=1/4th root,

4=Log,5=4th root,6=Sqrt,7=None)

Number of observation in each subset of data : 12

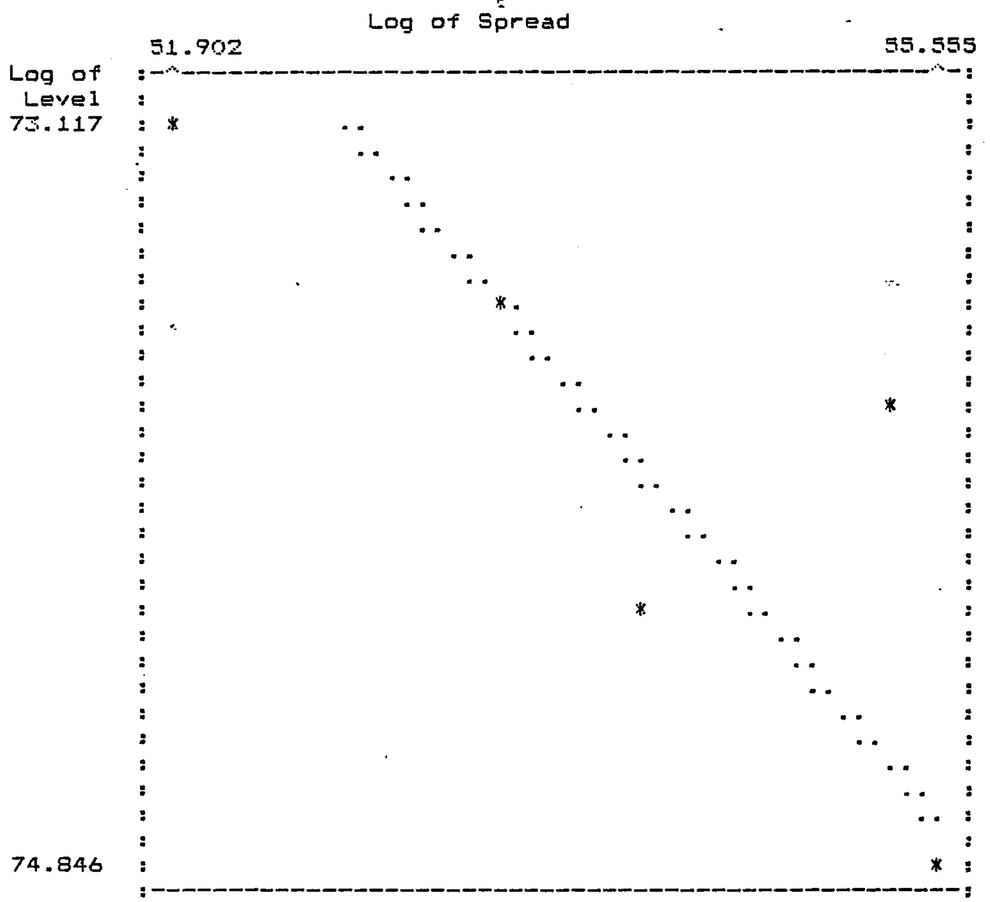
Press ENTER to accept :

โดยผู้วิเคราะห์จะเลือกข้อความใส่ในสดมภ์ขวามือ เช่น เลือกใส่ 7 และ 12 หมายความว่า ไม่ต้องแปลงข้อมูล และให้เซ็ทย่อยของข้อมูลประกอบด้วยค่าสังเกต 12 ค่า แทนด้วย * ในกราฟ เมื่อ <enter> จะปรากฏตำแหน่งของ * และท้ายรูปจะแนะนำว่า ควรจะแปลงข้อมูลด้วยวิธีใด

NUMBER - EXPLORATORY DATA ANALYSIS

SPREAD VS LEVEL PLOT

All Data Values are Scaled by 10^1



* Summarizes Subsets of 12 Consecutive Observations
Least Squares Estimate of Slope = 1.8017
Approximate Transformation to Stabilize Spread = Reciprocal

หมายเลข 6 แสดงฟังก์ชั่นสหสัมพันธ์ของข้อมูลที่ห่างกัน 1 คาบเวลา 2 คาบเวลา.... ฯลฯ เพื่อพิจารณาว่าข้อมูล ณ คาบเวลาต่าง ๆ มีสหสัมพันธ์ในตัวเองจนถึงคาบเวลาใด ลักษณะของสหสัมพันธ์ในตัวเองจะค่อย ๆ ลดลงอย่างช้า ๆ หรือลดลงอย่างรวดเร็ว เพื่อเลือกตัวแบบสำหรับข้อมูลชุดนั้น ๆ เมื่อกดหมายเลข 6 จะปรากฏข้อความดังนี้

```
*** Autocorrelation Function ***

Transformation to original data series :        7
(1=Reciprocal,2=1/Sqrt,3=1/4th root,
4=Log,5=4th root,6=Sqrt,7=None)
Degree of regular differencing:                 0
Degree of seasonal differencing:                0
Length of seasonality:                          12
Number of lags to print:                        A

                    Press ENTER to accept  :
```

โดยข้อความสดมภ์ขวาเป็นตัวเลขที่ผู้วิเคราะห์เลือกใส่ เมื่อ <enter> แล้วจะปรากฏรูปฟังก์ชันสหสัมพันธ์ในตัวเองดังนี้

```
NUMBER - EXPLORATORY DATA ANALYSIS

     AUTOCORRELATION FUNCTION

60 Observations in the Working Series
Mean of the Working Series = 1631.23
Standard Deviation of the Working Series =  227.4213

Lag  Value  T-Value -1.0                   0.0                    +1.0
                    :-------------------------------------------------:
  1   0.20    1.57  :                  [     |*****]                  :
  2   0.21    1.61  :                  [     |*****]                  :
  3  -0.05   -0.36  :                  [    *|     ]                  :
  4  -0.09   -0.67  :                  [   **|     ]                  :
  5   0.12    0.91  :                  [     |***  ]                  :
  6   0.07    0.57  :                  [     |**   ]                  :
  7   0.12    0.95  :                  [     |***  ]                  :
  8  -0.16   -1.24  :                  [ ****|     ]                  :
  9  -0.22   -1.71  :                  [*****|     ]                  :
 10  -0.14   -1.09  :                  [ ****|     ]                  :
 11  -0.23   -1.78  :                  [*****|     ]                  :
 12   0.74    5.73  :                  [     |*****]************      :
 13   0.02    0.19  :                  [     |*    ]                  :
 14   0.21    1.59  :                  [     |*****]                  :
 15  -0.05   -0.37  :                  [    *|     ]                  :
 16  -0.10   -0.81  :                  [  ***|     ]                  :
 17   0.11    0.88  :                  [     |***  ]                  :
 18   0.04    0.30  :                  [     |*    ]                  :
 19   0.15    1.16  :                  [     |**** ]                  :
 20  -0.15   -1.19  :                  [ ****|     ]                  :
                    :-------------------------------------------------:

                 [ ] = Estimated Two-Standard Error Limits

Box-Pierce Chi-Square Statistic with 20 Degrees of Freedom = 56.3
Probability = 0
```

วิเคราะห์ข้อมูลและหาค่าพยากรณ์

โปรแกรม FORECAST PLUS ส่วนที่ใช้วิเคราะห์ผล และหาค่าพยากรณ์ ประกอบด้วย โปรแกรมย่อยวิธีพยากรณ์ 13 วิธี เข้าสู่โปรแกรม Forecasting analysis ด้วยการเลือกหมายเลข 3 ใน Main Menu จะปรากฏวิธีต่าง ๆ ให้เลือกบนจอดังนี้

*** Forecasting Analysis ***

1. Simple Moving Average (unweighted)
2. Single Exponential Smoothing
3. Double Exponential Smoothing
4. Holt's Two Parameter Smoothing
5. Harrison's Harmonic Smoothing
6. Brown's Quadratic Exponential Smoothing
7. Winter's Seasonal Smoothing
8. Robust Decomposition
9. Census X-11 Decomposition
10. Regression Trend Analysis
11. Multiple Regression Analysis
12. Generalized Adaptive Filtering
13. Box-Jenkins Analysis
14. Return to Main Menu

Which ?

1. Simple Moving Average (unweighted)

เป็นวิธีเฉลี่ยเคลื่อนที่อย่างง่าย ไม่ถ่วงน้ำหนัก เมื่อเลือกเลข 1 <enter> จะปรากฏข้อความบนจอภาพ สดมภ์ขวามือเป็นข้อความที่ผู้วิเคราะห์เลือกเช่น ความยาว

หรือจำนวนที่จะเฉลี่ยเคลื่อนที่ หากต้องการเฉลี่ยครั้งละ 3 ค่า ก็พิมพ์เลข 3 แทนที่ A หากไม่เลือกเอง เครื่องจะเลือกให้โดยอัตโนมัติ หากต้องการให้แสดงข้อมูลเดิม ค่าพยากรณ์ และค่าคลาดเคลื่อน ให้เปลี่ยน N เป็น Y เลือกเสร็จแล้ว <enter>

*** Simple Moving Average ***

Length of moving average :	A
Lead time :	1
Number of forecasts :	1
Time plot of original data, forcasts, and error :	Y
Residual autocorrelation function :	Y
Table of original data, forecasts, and error :	N
Summary statistics:	Y

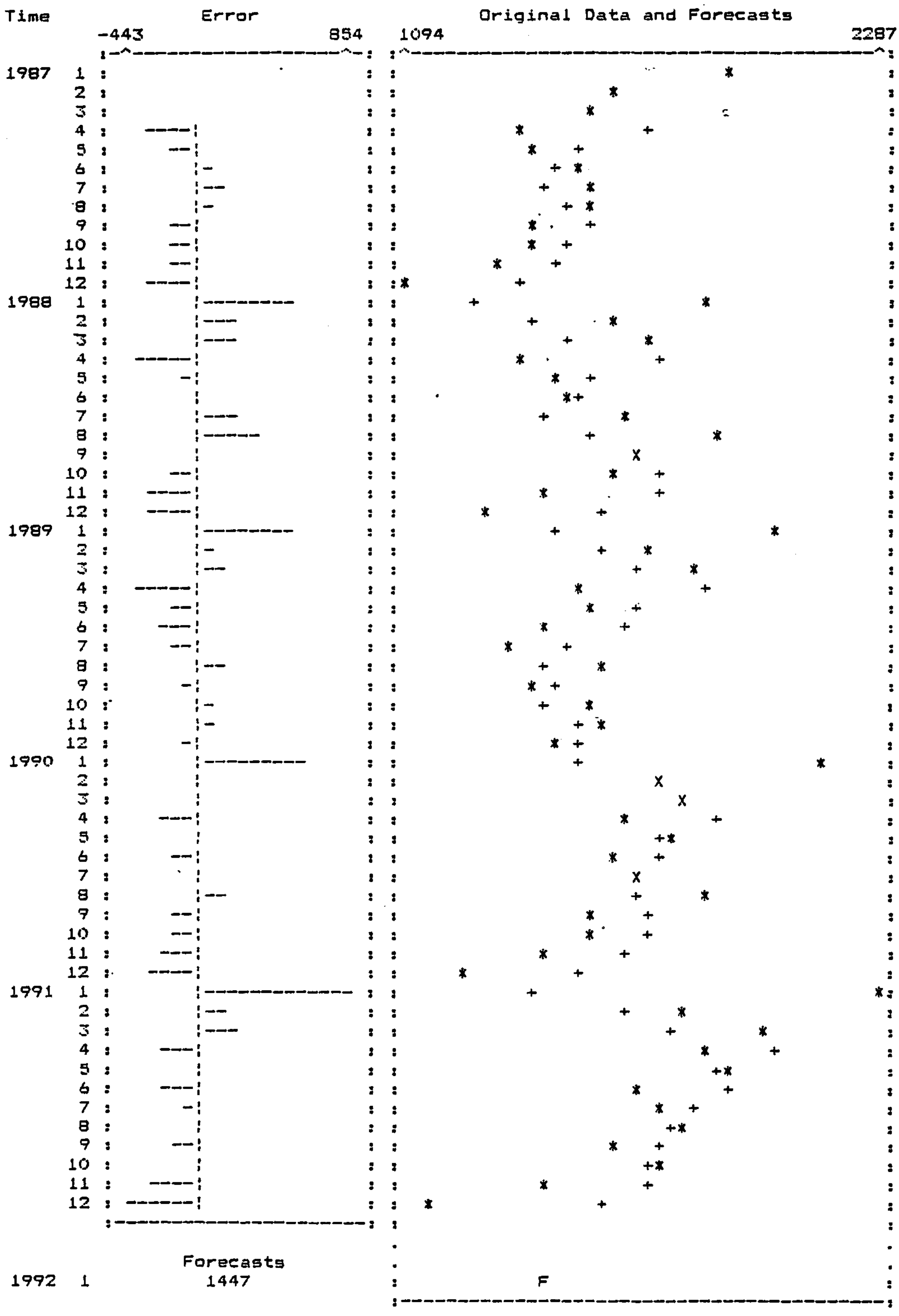
Press ENTER to accept :

รูปต่อไป แสดงค่าคลาดเคลื่อน ข้อมูลเดิม และค่าพยากรณ์ของข้อมูลชุดหนึ่ง (แสดงเพียงปีเดียวเป็นตัวอย่าง) ที่ใช้วิธีเฉลี่ยเคลื่อนที่ครั้งละ 3 เทอม

NUMBER - SIMPLE MOVING AVERAGES

Length of Moving Average = 3   Lead Time = 1

TIME PLOT OF ORIGINAL DATA, FORECASTS, AND ERROR



* = Observed Data Value
+ = 1 - Step Ahead Forecast
F = Forecast From Origin Period 57
X = Overlap

NUMBER - SIMPLE MOVING AVERAGES

Length of Moving Average = 3   Lead Time = 1

RESIDUAL AUTOCORRELATION FUNCTION

Mean of the Residuals =-10.64
Standard Deviation of the Residuals = 250.6562

```
Lag  Value  T-Value -1.0                    0.0                        +1.0
                    :-----------------------------------------------------:
 1    0.00    0.03  :                [      |       ]                     :
 2   -0.03   -0.23  :                [     *|       ]                     :
 3   -0.40   -2.97  :             ***[******|       ]                     :
 4   -0.21   -1.58  :                [ *****|       ]                     :
 5   -0.07    0.55  :                [      |**     ]                     :
 6    0.15    1.12  :                [      |****   ]                     :
 7    0.22    1.65  :                [      |******]                      :
 8   -0.09   -0.68  :                [    **|       ]                     :
 9   -0.42   -3.11  :             ***[******|       ]                     :
10   -0.15   -1.09  :                [  ****|       ]                     :
11   -0.15   -1.13  :                [  ****|       ]                     :
12    0.70    5.21  :                [      |******]**********            :
13   -0.07   -0.52  :                [    **|       ]                     :
14    0.06    0.41  :                [      |*      ]                     :
15   -0.23   -1.73  :                [******|       ]                     :
16   -0.13   -0.94  :                [   ***|       ]                     :
17    0.10    0.74  :                [      |**     ]                     :
18    0.09    0.67  :                [      |**     ]                     :
                    :-----------------------------------------------------:
```

[ ] = Estimated Two-Standard Error Limits

Box-Pierce Chi-Square Statistic with 18 Degrees of Freedom = 60.8
Probability = 0

NUMBER - SIMPLE MOVING AVERAGES

Length of Moving Average = 3   Lead Time = 1

SUMMARY STATISTICS

Mean of the Original Data = 1631.23
Standard Deviation of the Original Data = 227.421

Mean of the Residuals =-10.64
Standard Deviation of the Residuals = 250.656

	Mean Absolute % Error	Mean % Error	Mean Square Error
Entire Series	11.	-2.	61839.5

2. Single Exponential Smoothing

เป็นวิธีทำให้เรียบแบบเอกโปเนนเชียลครั้งเดียว ด้วยการเลือกหมายเลข 2 ใน Forecasting Analysis Menu จะปรากฏบนจอภาพ ดังนี้

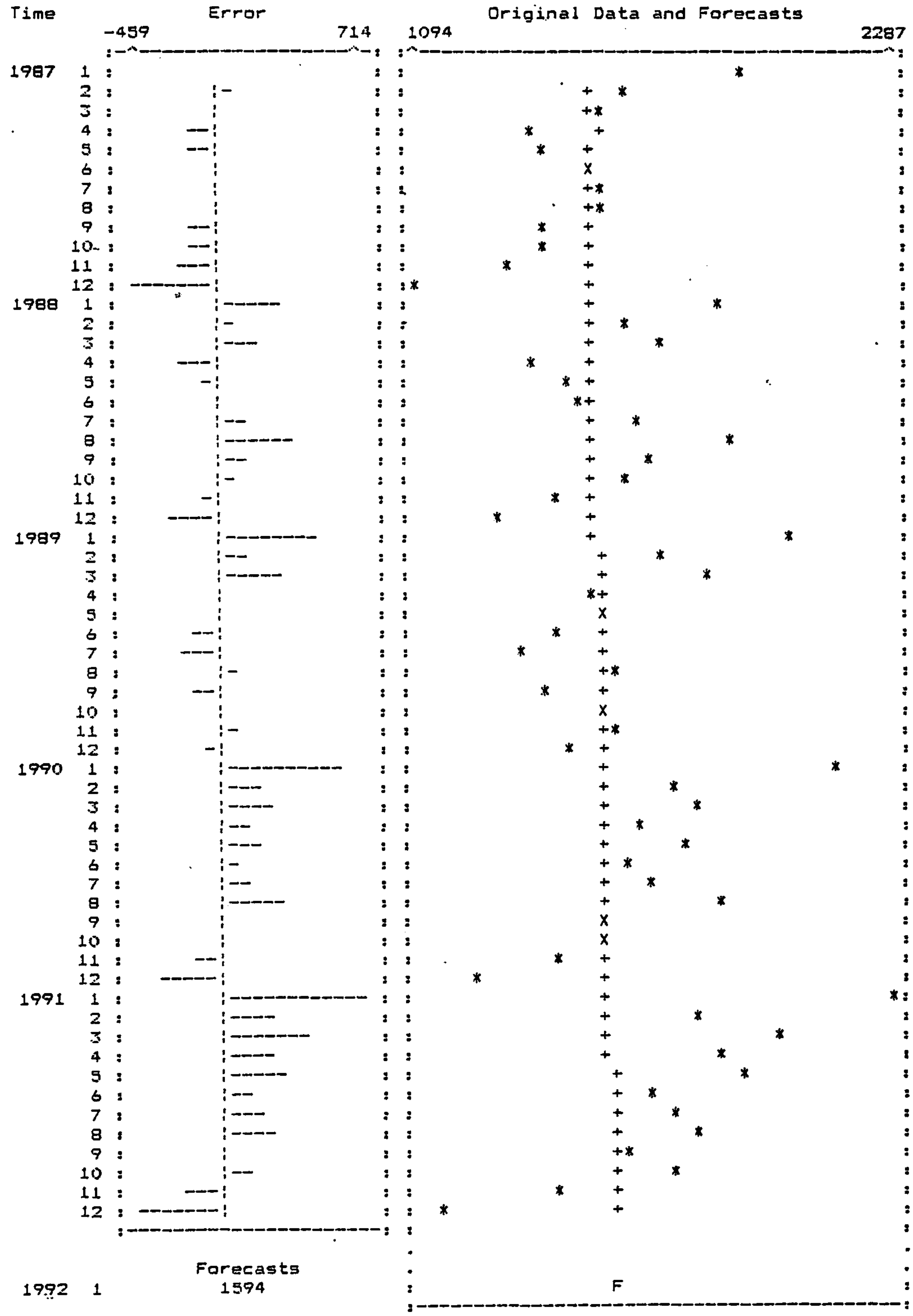
```
*** Single Exponential Smoothing ***

Smoothing constant :                                      A
Starting value [SO] :                                     A
Lead time :                                               1
Number of forecasts :                                     1
Time plot of original data, forecasts, and error :        Y
Residual autocorrelation function :                       Y
Table of original data, forecasts, and error :            N
Summary statistics :                                      Y
Final smoothed statistics :                               N


                              Press ENTER to accept :
```

เลือกค่าที่เหมาะสมสำหรับสดมภ์ขวามือเครื่องจะแสดงรูป และค่าความคลาดเคลื่อนข้อมูลดิบ และค่าพยากรณ์ แสดงฟังก์ชั่นสหสัมพันธ์ในตัวเองของเศษตกค้าง สรุปค่าสถิติ และสถิติทำให้เรียบชุดสุดท้าย ในทำนองเดียวกันกับวิธีเฉลี่ยเคลื่อนที่อย่างง่าย

NUMBER - SINGLE EXPONENTIAL SMOOTHING

Smoothing Constant = 0.01   Lead Time = 1

TIME PLOT OF ORIGINAL DATA, FORECASTS, AND ERROR



* = Observed Data Value
\+ = 1 - Step Ahead Forecast
F = Forecast From Origin Period 60
X = Overlap

NUMBER - SINGLE EXPONENTIAL SMOOTHING

Smoothing Constant = 0.01 Lead Time = 1

RESIDUAL AUTOCORRELATION FUNCTION

Mean of the Residuals = 59.03
Standard Deviation of the Residuals = 224.0661

```
Lag  Value  T-Value  -1.0                    0.0                     +1.0
                     :--------------------------------------------------:
  1   0.16    1.20   :                  [     |****  ]                  :
  2   0.22    1.66   :                  [     |***** ]                  :
  3  -0.04   -0.32   :                  [    *|      ]                  :
  4  -0.05   -0.38   :                  [    *|      ]                  :
  5   0.14    1.06   :                  [     |***   ]                  :
  6   0.08    0.61   :                  [     |**    ]                  :
  7   0.17    1.33   :                  [     |****  ]                  :
  8  -0.10   -0.78   :                  [  ***|      ]                  :
  9  -0.13   -1.01   :                  [  ***|      ]                  :
 10  -0.07   -0.50   :                  [   **|      ]                  :
 11  -0.02   -0.13   :                  [     |      ]                  :
 12   0.65    5.00   :                  [     |******]*********         :
 13   0.00    0.04   :                  [     |      ]                  :
 14   0.19    1.48   :                  [     |***** ]                  :
 15  -0.00   -0.02   :                  [     |      ]                  :
 16  -0.00   -0.02   :                  [     |      ]                  :
 17   0.17    1.33   :                  [     |****  ]                  :
 18   0.06    0.46   :                  [     |**    ]                  :
 19   0.17    1.30   :                  [     |****  ]                  :
                     :--------------------------------------------------:
```

[ ] = Estimated Two-Standard Error Limits

Box-Pierce Chi-Square Statistic with 19 Degrees of Freedom = 40.5
Probability = .003

NUMBER - SINGLE EXPONENTIAL SMOOTHING

Smoothing Constant = 0.01 Lead Time = 1

SUMMARY STATISTICS

Mean of the Original Data = 1631.23
Standard Deviation of the Original Data = 227.421

Mean of the Residuals = 59.03
Standard Deviation of the Residuals = 224.066

	Mean Absolute % Error	Mean % Error	Mean Square Error
Optimization Period	10.	-5.	32852.0
Remainder of Series	11.	4.	59653.8
Entire Series	11.	2.	52839.8

3. **Double Exponential Smoothing**

เลือกหมายเลข 3 ใน Forecasting Analysis Menu จอภาพจะปรากฏจอความให้เลือก หลังจากเลือกแล้ว <enter> เครื่องจะแสดงค่าต่างๆ ดังรูปตัวอย่างต่อไปนี้

*** Double Exponential Smoothing ***

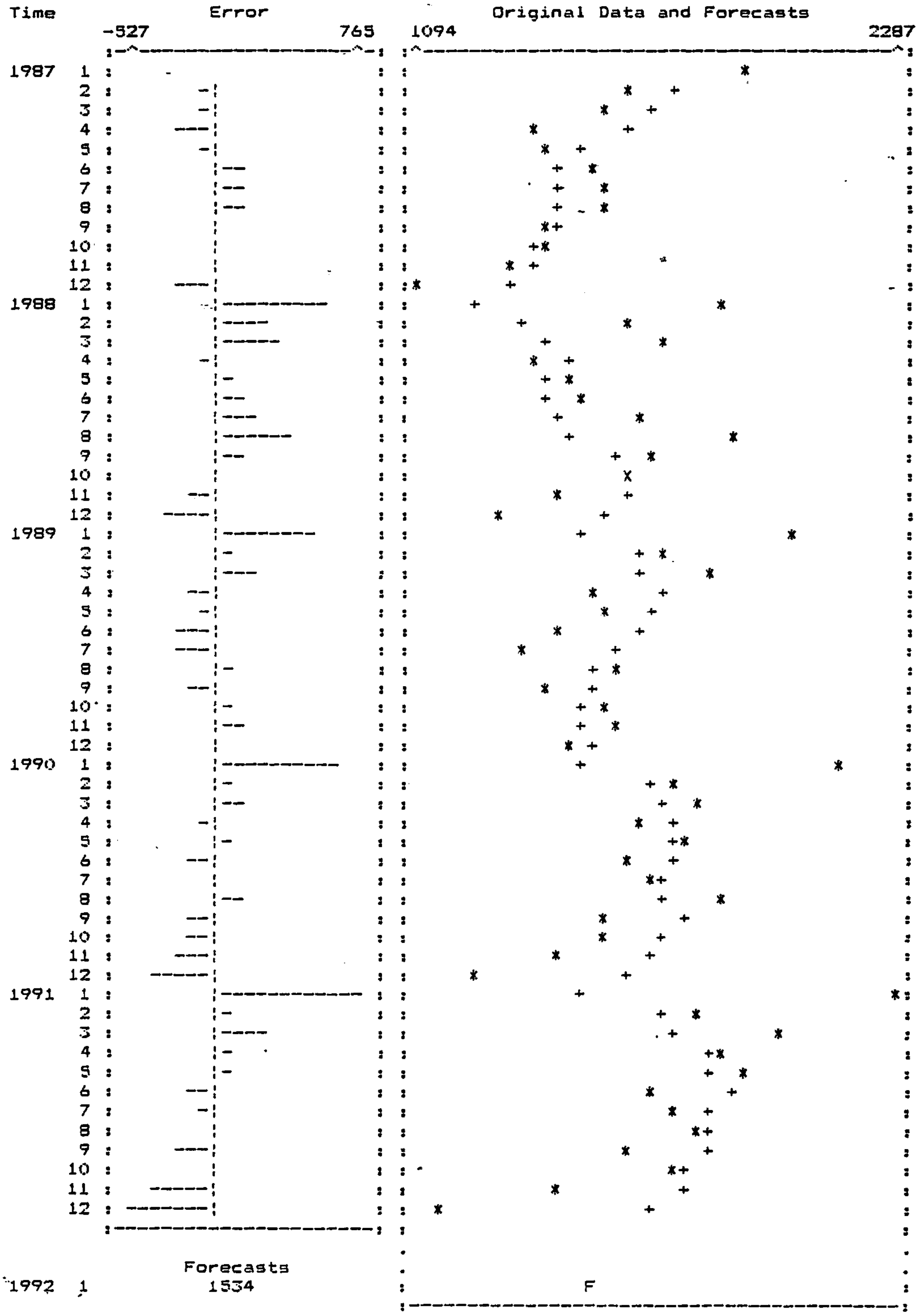
Smoothing constant :	A
Strating values [SO(1),SO(2)]:	A
Lead time :	1
Number of forecasts :	1
Time plot of original data, forecasts, and error :	Y
Residual autocorrelation function :	Y
Table of original data, forecasts, and error :	N
Summary statistics :	Y
Final smoothing statistics :	N

Press ENTER to accept :

NUMBER - DOUBLE EXPONENTIAL SMOOTHING

Smoothing Constant = 0.14 Lead Time = 1

TIME PLOT OF ORIGINAL DATA, FORECASTS, AND ERROR



* = Observed Data Value
+ = 1 - Step Ahead Forecast
F = Forecast From Origin Period 60
X = Overlap

NUMBER - DOUBLE EXPONENTIAL SMOOTHING

Smoothing Constant = 0.14   Lead Time = 1

RESIDUAL AUTOCORRELATION FUNCTION

Mean of the Residuals = 12.34
Standard Deviation of the Residuals = 241.9483

```
Lag  Value  T-Value -1.0                  0.0                          +1.0
                    :-----------------------------------------------------:
  1   0.04   0.33   :                  [      |*     ]                    :
  2   0.13   1.00   :                  [      |***   ]                    :
  3  -0.16  -1.20   :                  [  ****|      ]                    :
  4  -0.16  -1.19   :                  [  ****|      ]                    :
  5   0.07   0.54   :                  [      |**    ]                    :
  6   0.02   0.13   :                  [      |      ]                    :
  7   0.14   1.09   :                  [      |****  ]                    :
  8  -0.16  -1.24   :                  [  ****|      ]                    :
  9  -0.19  -1.45   :                  [ *****|      ]                    :
 10  -0.13  -0.98   :                  [   ***|      ]                    :
 11  -0.13  -0.97   :                  [   ***|      ]                    :
 12   0.59   4.53   :                  [      |******]********            :
 13  -0.10  -0.77   :                  [   ***|      ]                    :
 14   0.13   0.99   :                  [      |***   ]                    :
 15  -0.10  -0.75   :                  [    **|      ]                    :
 16  -0.10  -0.78   :                  [   ***|      ]                    :
 17   0.09   0.65   :                  [      |**    ]                    :
 18  -0.03  -0.22   :                  [     *|      ]                    :
 19   0.13   0.99   :                  [      |***   ]                    :
                    :-----------------------------------------------------:
```

[ ] = Estimated Two-Standard Error Limits

Box-Pierce Chi-Square Statistic with 19 Degrees of Freedom = 35.8
Probability = .011

NUMBER - DOUBLE EXPONENTIAL SMOOTHING

Smoothing Constant = 0.14   Lead Time = 1

SUMMARY STATISTICS

Mean of the Original Data = 1631.23
Standard Deviation of the Original Data = 227.421

Mean of the Residuals = 12.34
Standard Deviation of the Residuals = 241.948

	Mean Absolute % Error	Mean % Error	Mean Square Error
Optimization Period	10.	1.	48245.6
Remainder of Series	11.	-2.	60921.9
Entire Series	11.	-1.	57699.1

4. Holt's Two-Parameter Smoothing

หมายเลข 4 ใน Forecasting Analysis Menu แสดงการวิเคราะห์ข้อมูลและหาค่าพยากรณ์ด้วยวิธีของ Holt ซึ่งเป็นวิธีทำให้เรียบแบบเอกซ์โปเนนเชียล ที่มีค่าคงที่ทำให้เรียบสองตัวซึ่งผู้วิเคราะห์สามารถกำหนดค่าได้ในสดมภ์ขวามือ และเครื่องจะแสดงค่าต่างๆ ตามที่ปรากฏในจอภาพ ตามข้อความด้านซ้ายมือ

*** Holt's Two-Parameter Smoothing ***

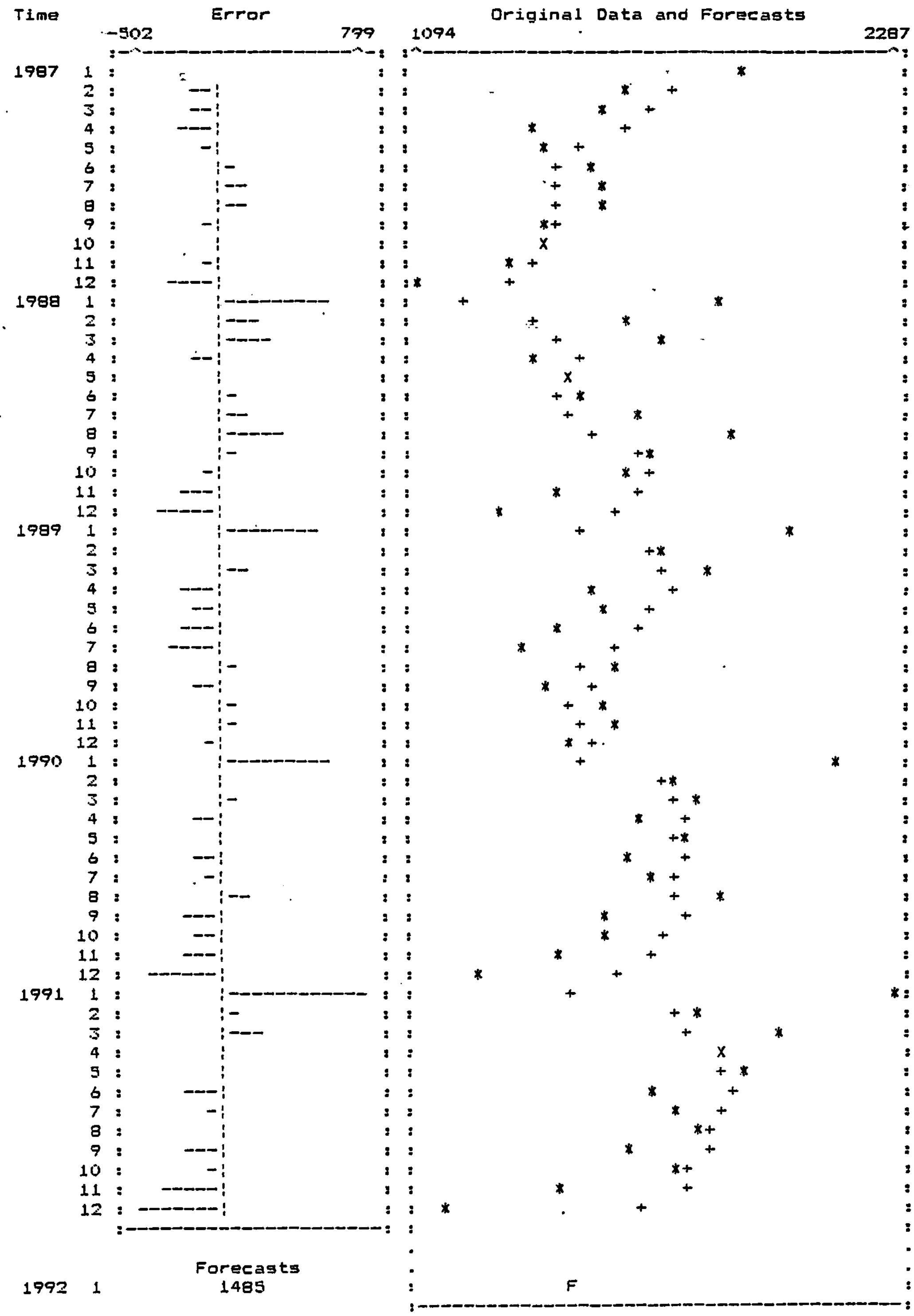
Smoothing constant for level :	A
Smoothing constant for trend :	A
Initial values for level and trend :	A
Lead time :	1
Number of forecasts :	1
Time plot of original data, forecasts, and error :	Y
Residual autocorrelation function :	Y
Table of original data, forecasts, and error :	N
Summary statistics :	Y
Final smoothed statistics :	N

Press ENTER to accept :

NUMBER - HOLT'S TWO-PARAMETER EXPONENTIAL SMOOTHING

Smoothing Constants: Alpha = 0.30 Beta = 0.10 Lead Time = 1

TIME PLOT OF ORIGINAL DATA, FORECASTS, AND ERROR



* = Observed Data Value
+ = 1 - Step Ahead Forecast
F = Forecast From Origin Period 60
X = Overlap

NUMBER - HOLT'S TWO-PARAMETER EXPONENTIAL SMOOTHING

Smoothing Constants: Alpha = 0.30 Beta = 0.10   Lead Time = 1

RESIDUAL AUTOCORRELATION FUNCTION

```
                    Mean of the Residuals = 1.04
                    Standard Deviation of the Residuals =  242.5481

Lag  Value  T-Value -1.0                 0.0                   +1.0
                    :-------------------------------------------------:
  1   0.01   0.11   :                 [     :      ]                  :
  2   0.11   0.85   :                 [     :***   ]                  :
  3  -0.18  -1.39   :                 [ *****:     ]                  :
  4  -0.18  -1.35   :                 [  ****:     ]                  :
  5   0.06   0.49   :                 [     :**    ]                  :
  6   0.01   0.10   :                 [     :      ]                  :
  7   0.15   1.12   :                 [     :****  ]                  :
  8  -0.17  -1.30   :                 [  ****:     ]                  :
  9  -0.20  -1.53   :                 [ *****:     ]                  :
 10  -0.14  -1.05   :                 [   ***:     ]                  :
 11  -0.15  -1.15   :                 [  ****:     ]                  :
 12   0.61   4.70   :                 [     :******]********          :
 13  -0.11  -0.88   :                 [   ***:     ]                  :
 14   0.13   1.01   :                 [     :***   ]                  :
 15  -0.10  -0.79   :                 [   ***:     ]                  :
 16  -0.11  -0.83   :                 [   ***:     ]                  :
 17   0.09   0.69   :                 [     :**    ]                  :
 18  -0.03  -0.21   :                 [     *:     ]                  :
 19   0.14   1.06   :                 [     :***   ]                  :
                    :-------------------------------------------------:

                      [ ] = Estimated Two-Standard Error Limits
```

Box-Pierce Chi-Square Statistic with 19 Degrees of Freedom = 39.2
Probability = .004

NUMBER - HOLT'S TWO-PARAMETER EXPONENTIAL SMOOTHING

Smoothing Constants: Alpha = 0.30 Beta = 0.10   Lead Time = 1

SUMMARY STATISTICS

Mean of the Original Data = 1631.23
Standard Deviation of the Original Data = 227.421

Mean of the Residuals = 1.04
Standard Deviation of the Residuals = 242.548

	Mean Absolute % Error	Mean % Error	Mean Square Error
Optimization Period	11.	1.	48561.8
Remainder of Series	11.	-2.	60994.4
Entire Series	11.	-1.	57833.6

5. **Brown's Quadratic Exponential Smoothing**

เมื่อเลือกที่จะใช้เทคนิคของบราวน์ในการวิเคราะห์ผล ให้กดหมายเลข 5 โปรแกรม FORECAST PLUS จะจัดการกับข้อมูลตามรายละเอียดที่ปรากฏตามข้อความด้านซ้ายมือในจอภาพ เมื่อเลือกใส่ข้อความลงในสดมภ์ขวามือแล้ว ดังนี้

<u>หมายเหตุ</u> Brown's Quadratics Exponential Smoothing ใน Forecast Plus ก็คือ Triple Exponential Smoothing Technique

*** Brown's Quadratic Exponential Smoothing ***

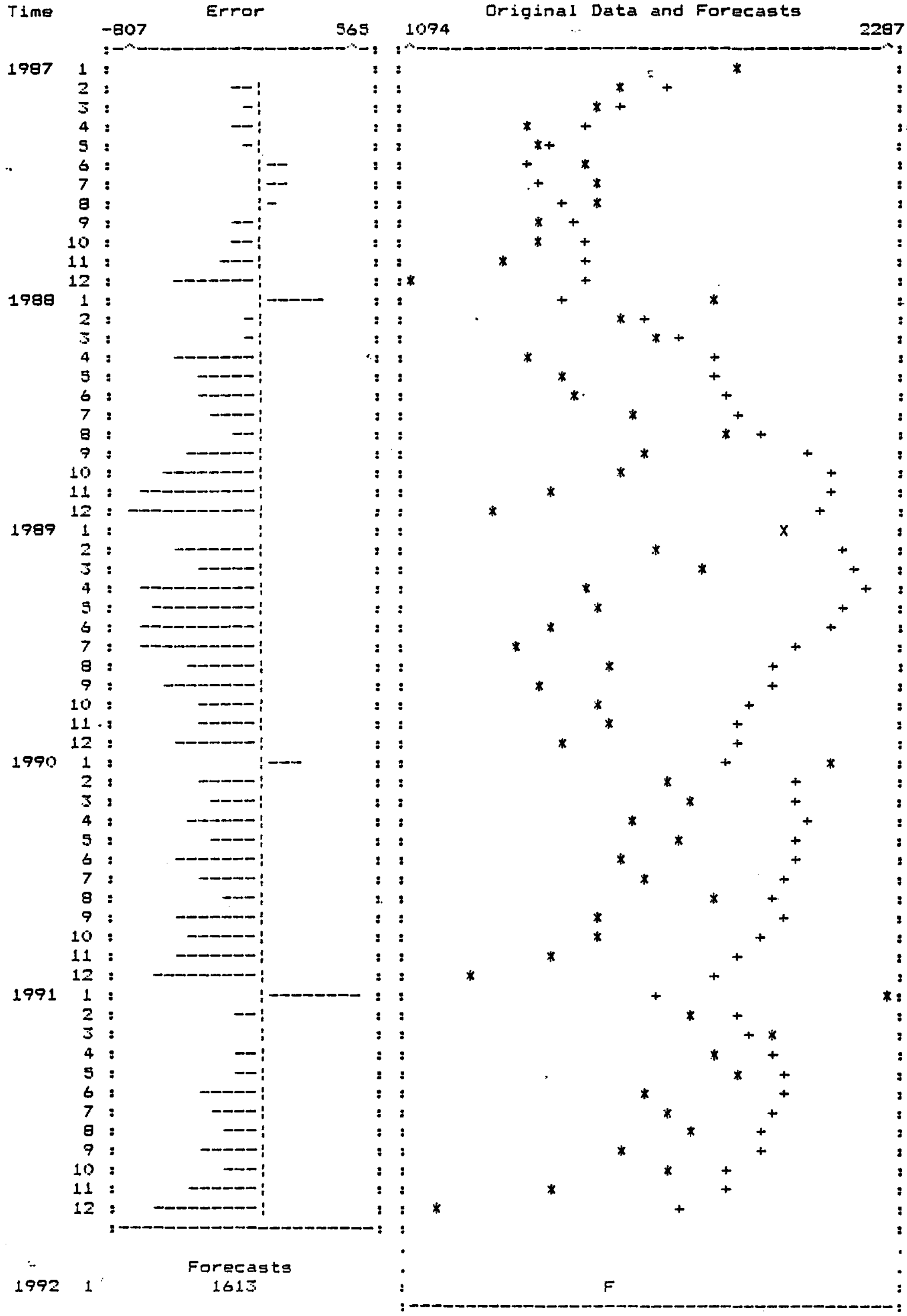
Smoothing constant :	A
Starting values [SO(1),SO(2),SO(3)] :	A
Lead time :	1
Number of forecasts :	1
Time plot of original data, forecasts, and error :	Y
Residual autocorrelation function :	Y
Table of original data, forecasts, and error :	N
Summary statistics :	Y
Final smoothed statistics :	N

Press ENTER to accept :

NUMBER - QUADRATIC EXPONENTIAL SMOOTHING

Smoothing Constant = 0.10   Lead Time = 1

TIME PLOT OF ORIGINAL DATA, FORECASTS, AND ERROR



* = Observed Data Value
\+ = 1 - Step Ahead Forecast
F = Forecast From Origin Period 60
X = Overlap

NUMBER - QUADRATIC EXPONENTIAL SMOOTHING

Smoothing Constant = 0.10   Lead Time = 1

RESIDUAL AUTOCORRELATION FUNCTION

```
                     Mean of the Residuals =-275.58
                     Standard Deviation of the Residuals =  277.8269

Lag  Value  T-Value  -1.0                     0.0                     +1.0
                     :---------------------------------------------------:
  1   0.24    1.85   :                  [      |******]                  :
  2   0.30    2.28   :                  [      |******]                  :
  3   0.08    0.59   :                  [      |**    ]                  :
  4   0.06    0.43   :                  [      |*     ]                  :
  5   0.20    1.55   :                  [      |***** ]                  :
  6   0.14    1.05   :                  [      |***   ]                  :
  7   0.20    1.56   :                  [      |***** ]                  :
  8  -0.07   -0.53   :                  [    **|      ]                  :
  9  -0.13   -1.02   :                  [   ***|      ]                  :
 10  -0.11   -0.87   :                  [   ***|      ]                  :
 11  -0.18   -1.38   :                  [  ****|      ]                  :
 12   0.33    2.55   :                  [      |******]*                 :
 13  -0.27   -2.05   :                  [******|      ]                  :
 14  -0.12   -0.96   :                  [   ***|      ]                  :
 15  -0.31   -2.35   :                 *[******|      ]                  :
 16  -0.35   -2.66   :                **[******|      ]                  :
 17  -0.22   -1.72   :                  [******|      ]                  :
 18  -0.34   -2.63   :                **[******|      ]                  :
 19  -0.27   -2.09   :                  [******|      ]                  :
                     :---------------------------------------------------:

                          [ ] = Estimated Two-Standard Error Limits
```

Box-Pierce Chi-Square Statistic with 19 Degrees of Freedom = 57.5
Probability = 0

NUMBER - QUADRATIC EXPONENTIAL SMOOTHING

Smoothing Constant = 0.10   Lead Time = 1

SUMMARY STATISTICS

Mean of the Original Data = 1631.23
Standard Deviation of the Original Data = 227.421

Mean of the Residuals =-275.58
Standard Deviation of the Residuals = 277.827

	Mean Absolute % Error	Mean % Error	Mean Square Error
Optimization Period	12.	-6.	50149.3
Remainder of Series	25.	-23.	186488.8
Entire Series	22.	-19.	151826.2

6. **Winters' Seasonal Smoothing**

เป็นวิธีพยากรณ์ข้อมูลที่มีการเปลี่ยนแปลงตามฤดูกาลและมีแนวโน้ม จะต้องกำหนดค่าคงที่ทำให้เรียบ 3 ค่า คือค่าคงที่ทำให้เรียบสำหรับระดับ สำหรับแนวโน้ม และสำหรับฤดูกาล ตัวแบบของ Winter มี 2 แบบ คือ แบบบวก (additive model) และแบบคูณ (multiplicative model) ดังนั้นเมื่อเข้าสู่โปรแกรมย่อย Forecasting analysis และเลือกหมายเลข 7 <enter> แล้ว จะปรากฏข้อความบนจอภาพให้เลือก และจะแสดงผลต่าง ๆ ตามข้อความด้านซ้ายมือในจอภาพ

*** Winter's Seasonal Smoothing ***

Additive (1) or Multiplicative (2) model :	2
Length of seasonality :	12
Starting values [A/U] :	A
Level Smoothing constant :	A
Trend smoothing constant :	A
Seasonal smoothing constant :	A
Lead time :	1
Number of forecasts :	1
Time plot of original data, forecasts, and error :	Y
Residual autocorrelation function :	Y
Table of original data, forecasts, and error :	N
Summary statistics :	Y
Final smoothed statistics :	N

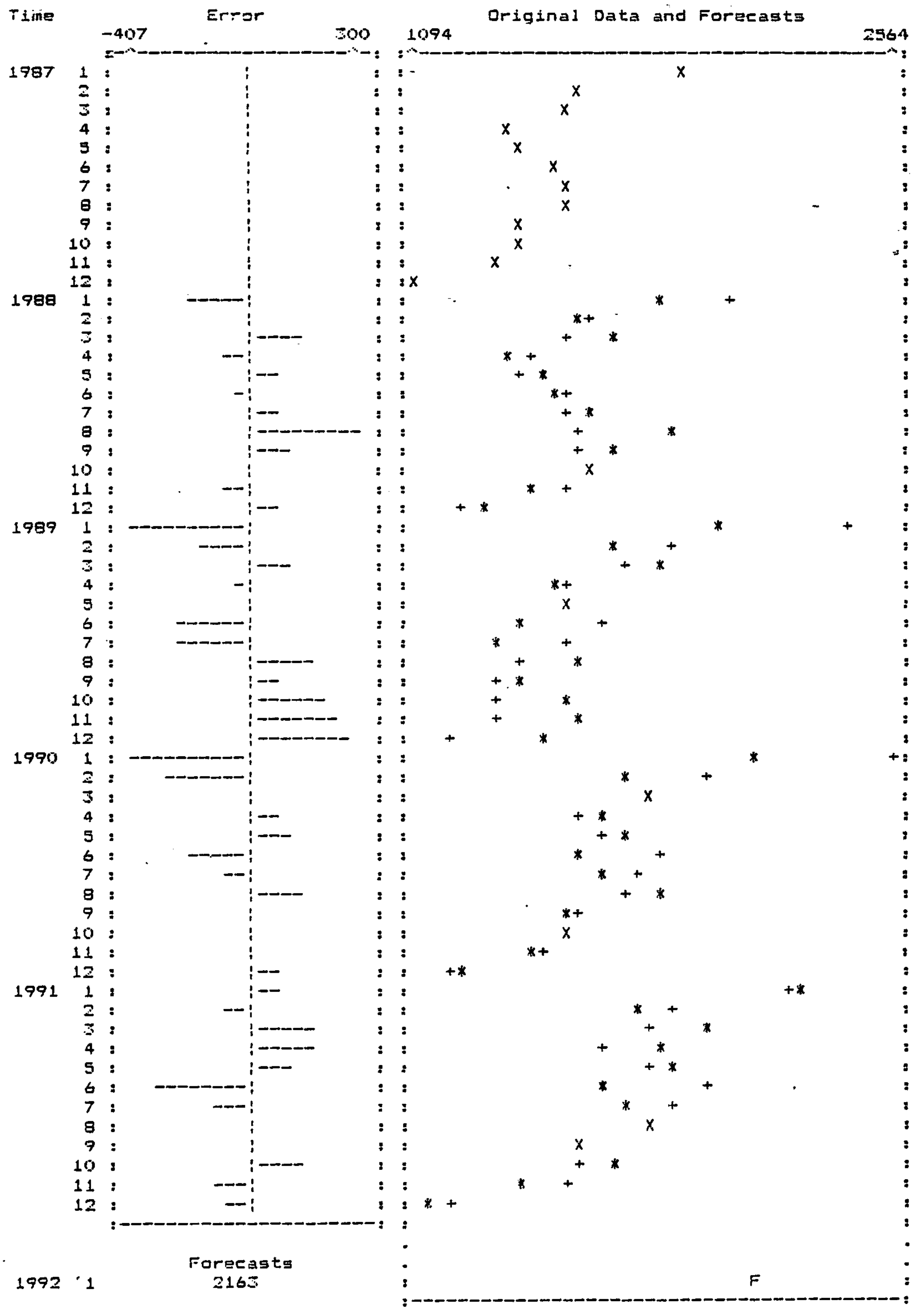
Press ENTER to accept :

หากไม่เลือกค่าเริ่มต้น ให้พิมพ์ A เครื่องจะกำหนดค่าให้โดยอัตโนมัติ แต่ถ้าต้องการใช้ค่าใดค่าหนึ่งเป็นค่าเริ่มต้น (Starting value) ให้พิมพ์ค่านั้นลงไปในสดมภ์ขวามือ คือ เปลี่ยนจาก A เป็นค่าที่ต้องการ สำหรับค่าคงที่ทำให้เรียบ 3 ค่า ถ้าใช้ A เครื่องจะเลือกใช้ค่าคงที่จากข้อมูล n/4 ค่าแรก ผู้วิเคราะห์สามารถกำหนดจำนวนข้อมูลที่ใช้เป็น n/2 ค่าแรก หรือ n/2 ค่าสุดท้าย หรือทั้ง n ค่าก็ได้ ค่าคงที่ทำให้เรียบที่เครื่องเลือกให้จะให้ค่าพยากรณ์ที่มีค่า MSE น้อยที่สุด

NUMBER – WINTERS' MULTIPLICATIVE SEASONAL EXPONENTIAL SMOOTHING

Smoothing Constants: Alpha = 0.50 Beta = 0.10 Gamma = 0.10 Lead Time = 1

TIME PLOT OF ORIGINAL DATA, FORECASTS, AND ERROR



* = Observed Data Value
+ = 1 - Step Ahead Forecast
F = Forecast From Origin Period 60
X = Overlap

NUMBER - WINTERS' MULTIPLICATIVE SEASONAL EXPONENTIAL SMOOTHING

Smoothing Constants: Alpha = 0.50 Beta = 0.10 Gamma = 0.10 Lead Time = 1

RESIDUAL AUTOCORRELATION FUNCTION

Mean of the Residuals =-13.02
Standard Deviation of the Residuals = 144.6827

```
Lag  Value  T-Value -1.0                        0.0                         +1.0
                    :-----------------------------------------------------------:
 1    0.17   1.33   :                    [     |****  ]                         :
 2   -0.23  -1.77   :                    [*****|      ]                         :
 3   -0.20  -1.55   :                    [*****|      ]                         :
 4   -0.11  -0.87   :                    [  ***|      ]                         :
 5    0.05   0.37   :                    [     |*     ]                         :
 6   -0.11  -0.81   :                    [  ***|      ]                         :
 7   -0.01  -0.07   :                    [     |      ]                         :
 8    0.00   0.02   :                    [     |      ]                         :
 9   -0.08  -0.63   :                    [   **|      ]                         :
10   -0.36  -2.80   :                 ***[*****|      ]                         :
11   -0.17  -1.32   :                    [ ****|      ]                         :
12    0.46   3.57   :                    [     |*****]******                    :
13    0.11   0.89   :                    [     |***   ]                         :
14   -0.06  -0.43   :                    [    *|      ]                         :
15   -0.00  -0.00   :                    [     |      ]                         :
16    0.09   0.72   :                    [     |**    ]                         :
17    0.31   2.42   :                    [     |*****]**                        :
18   -0.08  -0.62   :                    [   **|      ]                         :
19   -0.08  -0.61   :                    [   **|      ]                         :
20   -0.07  -0.54   :                    [   **|      ]                         :
                    :-----------------------------------------------------------:
```

[ ] = Estimated Two-Standard Error Limits

Box-Pierce Chi-Square Statistic with 20 Degrees of Freedom = 39.9
Probability = .005

NUMBER - WINTERS' MULTIPLICATIVE SEASONAL EXPONENTIAL SMOOTHING

Smoothing Constants: Alpha = 0.50 Beta = 0.10 Gamma = 0.10 Lead Time = 1

SUMMARY STATISTICS

Mean of the Original Data = 1631.23
Standard Deviation of the Original Data = 227.421

Mean of the Residuals =-13.02
Standard Deviation of the Residuals = 144.683

	Mean Absolute % Error	Mean % Error	Mean Square Error
Optimization Period	3.1	0.5	8356
Remainder of Series	8.0	-1.6	29019
Entire Series	6.1	-0.8	20754

7. **Box-Jenkins Analysis**

เป็นเทคนิคการวิเคราะห์อนุกรมเวลาวิธีสุดท้ายใน FORECAST PLUS เมื่อเลือกหมายเลข 13 ใน Forecasting Analysis Menu จะปรากฏข้อความบนจอให้เลือกดังนี้

<u>ขั้นที่ 1</u>

*** Box-Jenkins Analysis : Main Menu ***

1) Identification

2) Estimation/Forecasting

3) Return to Forecasting Menu

Which ?

หมายเลข 1 ใช้สำหรับตรวจดูลักษณะข้อมูลเพื่อเลือกตัวแบบที่เหมาะสมสำหรับข้อมูลในการพยากรณ์ต่อไป

ในโปรแกรมย่อยหมายเลข 1 นี้ จะให้ผู้วิเคราะห์เลือกว่าจะแปลงข้อมูลหรือไม่ ถ้าแปลงจะแปลงด้วยวิธีใด เช่น หารากที่สอง แปลงข้อมูลด้วยล็อกการิทึม เป็นต้น และให้กำหนดว่าจะหาค่าแตกต่างปกติอันดับที่เท่าใด หาค่าแตกต่างฤดูกาลอันดับที่เท่าใด ความยาวของฤดูกาลมีค่าเท่าใด โปรแกรมนี้จะแสดงการลงจุดฟังก์ชั่นสหสัมพันธ์ในตัวเอง และสหสัมพันธ์ในตัวเองบางส่วน และจำนวนคาบเวลาที่ห่างกันใน ACF

<u>ขั้นที่ 2</u> เมื่อเลือก 1 ในขั้นที่หนึ่งแล้วจะเข้าสู่ส่วนของ Identification

*** Box-Jenkins Analysis : Identification ***

```
Transformation to original data series :          7
(1=Reciprocal,2=1/Sqrt,3=1/4th root,
4=Log,5=4th root,6=Sqrt,7=None)
Degree of Regular Differencing  :                 0
Degree of Seasonal Differencing :                 0
Length of Seasonal Period :                       12
Plot of Autocorrelation Function :                Y
Plot of Patial Autocorrelation Function :         Y
Number of lags to print in ACFs :                 12

                        Press ENTER to accept  :
```

ผู้วิเคราะห์เลือกค่าต่างๆใส่ในสดมภ์ขวามือ เช่น ตัวอย่างเลข 7 หมายความว่า จะไม่แปลงข้อมูล เลข 0 คือ ไม่มีการหาค่าแตกต่างทั้งแบบปกติ และแบบมีฤดูกาล กำหนดความยาวฤดูกาลเท่ากับ 12 ให้พิมพ์ฟังก์ชั่นสหสัมพันธ์ในตัวเอง และฟังก์ชั่นสหสัมพันธ์ในตัวเองบางส่วน กำหนดให้จำนวนคาบที่ห่างกัน 12 คาบเวลา พิมพ์ใน ACF

NUMBER - BOX-JENKINS ANALYSIS: IDENTIFICATION

AUTOCORRELATION FUNCTION

```
                    60 Observations in the Working Series
                    Mean of the Working Series = 1631.23
                    Standard Deviation of the Working Series = 227.4213

Lag  Value  T-Value -1.0                     0.0                      +1.0
                    :-------------------------------------------------:
  1   0.20    1.57  :                   [      |*****]                 :
  2   0.21    1.61  :                   [      |*****]                 :
  3  -0.05   -0.36  :                   [     *|     ]                 :
  4  -0.09   -0.67  :                   [    **|     ]                 :
  5   0.12    0.91  :                   [      |***  ]                 :
  6   0.07    0.57  :                   [      |**   ]                 :
  7   0.12    0.95  :                   [      |***  ]                 :
  8  -0.16   -1.24  :                   [ ****|      ]                 :
  9  -0.22   -1.71  :                   [*****|      ]                 :
 10  -0.14   -1.09  :                   [ ****|      ]                 :
 11  -0.23   -1.78  :                   [*****|      ]                 :
 12   0.74    5.73  :                   [      |*****]************     :
 13   0.02    0.19  :                   [      |*    ]                 :
 14   0.21    1.59  :                   [      |*****]                 :
 15  -0.05   -0.37  :                   [     *|     ]                 :
                    :-------------------------------------------------:

                      [ ] = Estimated Two-Standard Error Limits

     Box-Pierce Chi-Square Statistic with 15 Degrees of Freedom = 52
     Probability = 0
```

NUMBER - BOX-JENKINS ANALYSIS: IDENTIFICATION

PARTIAL AUTOCORRELATION FUNCTION

```
                    60 Observations in the Working Series
                    Mean of the Working Series = 1631.23
                    Standard Deviation of the Working Series = 227.4213

Lag  Value  T-Value -1.0                     0.0                      +1.0
                    :-------------------------------------------------:
  1   0.20    1.57  :                   [      |*****]                 :
  2   0.17    1.35  :                   [      |**** ]                 :
  3  -0.12   -0.97  :                   [   ***|     ]                 :
  4  -0.10   -0.79  :                   [   ***|     ]                 :
  5   0.20    1.58  :                   [      |*****]                 :
  6   0.06    0.46  :                   [      |*    ]                 :
  7   0.01    0.08  :                   [      |     ]                 :
  8  -0.23   -1.81  :                   [*****|      ]                 :
  9  -0.17   -1.31  :                   [ ****|      ]                 :
 10   0.04    0.29  :                   [      |*    ]                 :
 11  -0.17   -1.31  :                   [ ****|      ]                 :
 12   0.94    7.27  :                   [      |*****]**************** :
 13  -1.00   -7.75  :******************[*****|      ]                 :
 14  -0.92   -7.09  :   ***************[*****|      ]                 :
 15  -1.00   -7.75  :******************[*****|      ]                 :
                    :-------------------------------------------------:

                      [ ] = Estimated Two-Standard Error Limits
```

ขั้นที่ 3 หลังจากพิจารณาลักษณะของ ACF และ PACF แล้ว ผู้วิเคราะห์อาจกำหนดตัวแบบสำหรับอนุกรมเวลาชุดนี้ได้เลย หรือเห็นว่า ควรจะต้องแปลงข้อมูลเพื่อให้ข้อมูลอยู่ในสภาพ stationary ก่อน แล้วจึงกำหนดตัวแบบ ถ้าเป็นกรณีที่ 1 ผู้วิเคราะห์จะให้ FORECAST PLUS ประมาณค่าของพารามิเตอร์ของตัวแบบที่กำหนดขึ้น โดยกลับไปที่ Forecasting Menu แล้วเลือกหมายเลข 2 (Estimation / Forecasting) จะปรากฏข้อความบนจอภาพดังนี้

```
*** Box-Jenkins Analysis : Estimation ***

Transformation to original data series :            4
(1=Reciprocal,2=1/Sqrt,3=1/4th root,
4=Log,5=4th root,6=Sqrt,7=None)

Degree of Regular Differencing  :                   1
Degree of Seasonal Differencing :                   1
Length of Seasonal Period :                         12

Regular Autoregressive Terms :                      1
Regular Moving Average Terms :                      0
Seasonal Autoregressive Terms :                     0
Seasonal Moving Average Terms :                     1
Include constant term (Y/N/A) :                     A
Number of Backforecasts :                           0
Maximum Iterations for Parameter Estimates :        50
Convergence Tolerance :                             .001
Initial estimates from data (A) or user (U) :       A
Plot of Residual Autocorrelation Function :         Y
Number of lags to print in ACFs :                   12

                          Press ENTER to accept  :
```

ตามตัวอย่างนี้ ผู้วิเคราะห์จะแปลงข้อมูลเดิม และจะหาค่าแตกต่างแรกทั้งสำหรับกรณีปกติ และสำหรับฤดูกาล (เลข 1 คือ หาค่าแตกต่างแรก) ความยาวของฤดูกาลเท่ากับ 12 คาบเวลา ในตัวแบบจะประกอบด้วย การถดถอยในตัวเอง 1 เทอม และการเฉลี่ยเคลื่อนที่แบบมีฤดูกาล 1 เทอมคือตัวแบบเป็น ARIMA $(1,1,0)*(0,1,1)_{12}$ โดยจะให้โปรแกรมกำหนดค่าคงที่ให้ในตัวแบบ (พิมพ์ A ถ้าไม่ต้องการค่าคงที่ให้พิมพ์ N) จำนวนค่าพยากรณ์ย้อนหลัง = 0 คือไม่แสดงค่าพยากรณ์ย้อนหลัง จำนวนรอบสูงสุดในการประมาณค่า = 50 ค่าประมาณเริ่มต้น จะเป็นไปได้โดยอัตโนมัติ ถ้าพิมพ์ A หรือผู้วิเคราะห์จะกำหนดเป็นค่าอื่นก็ได้ ให้ลงจุดฟังก์ชั่นสหสัมพันธ์ในตัวเองของเศษตกค้างด้วยจำนวนคาบของ ACF ที่จะพิมพ์ออกมาเท่ากับ 12 คาบเวลา จะปรากฏผลดังนี้

NUMBER - BOX-JENKINS ANALYSIS

PARAMETER ESTIMATES AND STATISTICS

47 Observations Used For Parameter Estimates
Degree of Regular Differencing = 1
Degree of Seasonal Differencing = 1
Seasonal Period = 12

Sum of Squared Errors = 656121
Residual Variance = 11312.42

Parameter	Value	T-Statistic
AR(1)	-0.0449	-0.33
SMA(1)	0.3992	2.51

Correlation Matrix

	AR(1)	SMA(1)
AR(1)	1.00	
SMA(1)	-0.24	1.00

NUMBER - BOX-JENKINS ANALYSIS

RESIDUAL AUTOCORRELATION FUNCTION

46 Observations in the Residual Series
Mean of the Residual Series =-6.07
Standard Deviation of the Residual Series = 118.4813

```
Lag  Value  T-Value -1.0                    0.0                    +1.0
                    :--------------------------------------------------:
  1   0.01    0.08  :                  [     |     ]                   :
  2  -0.11   -0.72  :                  [  ***|     ]                   :
  3  -0.08   -0.56  :                  [   **|     ]                   :
  4  -0.26   -1.74  :                  [*****|     ]                   :
  5  -0.11   -0.75  :                  [  ***|     ]                   :
  6   0.02    0.17  :                  [     |*    ]                   :
  7  -0.00   -0.01  :                  [     |     ]                   :
  8   0.19    1.29  :                  [     |*****]                   :
  9   0.15    1.05  :                  [     |**** ]                   :
 10  -0.34   -2.32  :                **[*****|     ]                   :
 11  -0.30   -2.05  :                 *[*****|     ]                   :
 12  -0.01   -0.09  :                  [     |     ]                   :
 13  -0.08   -0.51  :                  [   **|     ]                   :
 14   0.22    1.52  :                  [     |*****]                   :
 15   0.42    2.87  :                  [     |*****]****               :
                    :--------------------------------------------------:
```

[ ] = Estimated Two-Standard Error Limits

Box-Pierce Chi-Square Statistic with 13 Degrees of Freedom = 27.6
Probability = .01

พิจารณาจาก T-Statistic และ Box-Pierce Chi-Square Statistic Probability = 0.01 แสดงว่า ตัวแบบที่เลือกไม่เหมาะสม ไม่ควรเป็นตัวแบบ ARIMA(1,1,0)*(0,1,1)$_{12}$ จะต้องกลับไปพิจารณาลักษณะข้อมูลใหม่ และเลือกตัวแบบใหม่ ในที่นี้จะสมมติว่า ตัวแบบเหมาะสม เพื่อแสดงการพยากรณ์ ดังนั้นจะพิมพ์เลข 1 เพื่อทำการพยากรณ์

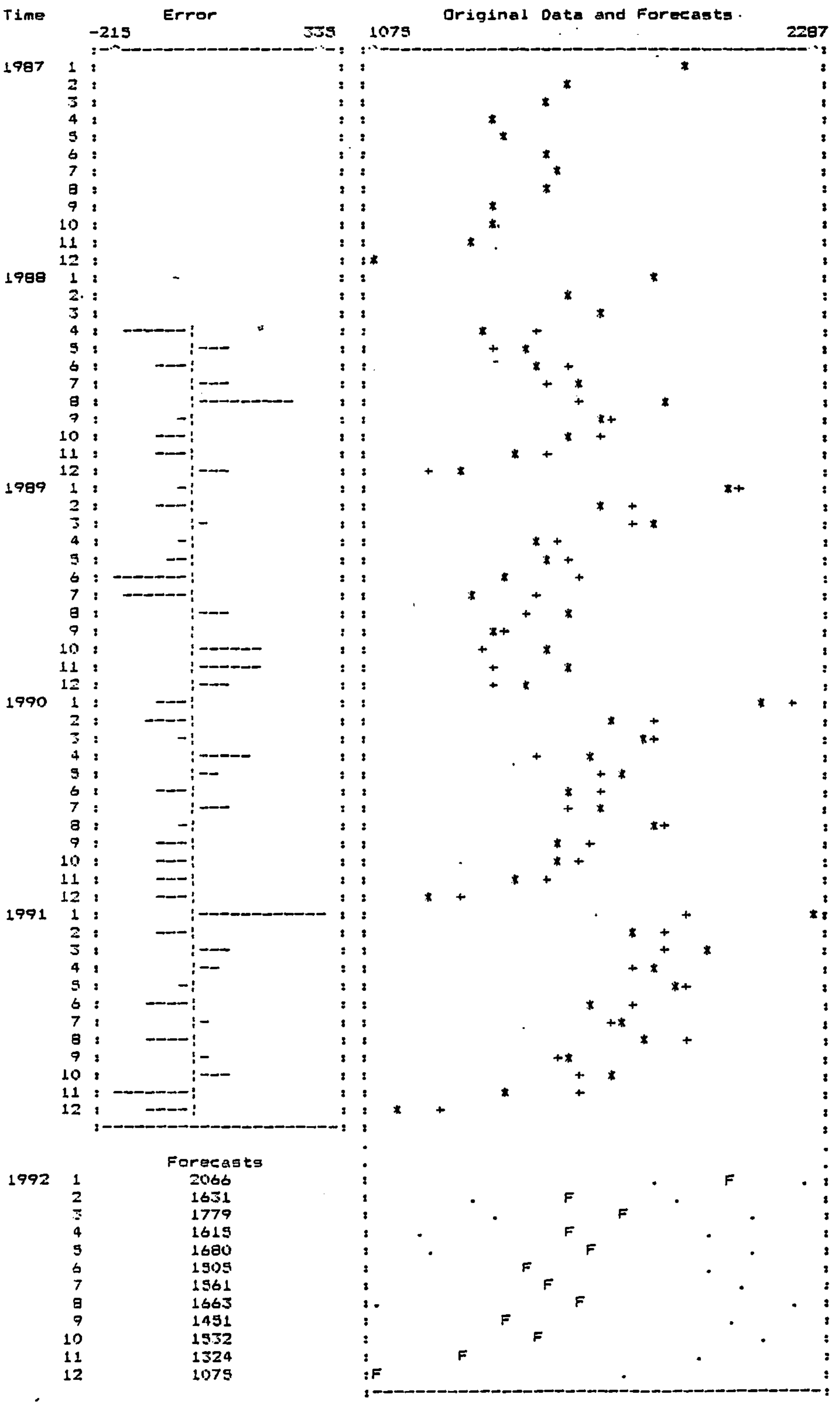
<u>ขั้นที่ 4</u> พยากรณ์ค่า เมื่อพิมพ์หมายเลข 1 ต่อท้ายคำถามว่า ต้องการทำอะไรต่อไป แล้วจะปรากฏข้อความบนจอภาพดังนี้ โดยตัวเลขสดมภ์ขวามือ เป็นค่าที่ผู้วิเคราะห์เลือกใช้

```
*** Box-Jenkins Analysis : Forecasting ***

Number of forecasts :                                12
Origin for forecasts :                               51
Percent forcast confidence interval :                95
Time plot of original data, forecasts, and error :   Y
Table of original data, forecasts, and error :       Y
```

เมื่อ <enter> แล้วจะปรากฏผลต่าง ๆ ดังต่อไปนี้

NUMBER - BOX-JENKINS ANALYSIS

TIME PLOT OF ORIGINAL DATA, FORECASTS, AND ERROR



* = Observed Data Value
+ = 1-Step Ahead Forecast
F = Forecast From Origin Period 60
. . = Approximate 95% Confidence Interval

NUMBER -BOX -JENKINS ANALYSIS

TABLE OF ORIGINAL DATA, FORECASTS, AND ERROR

Time		Original Data	Forecast	95% Confidence Limits Lower	95% Confidence Limits Upper	Error
1987	1	1936				
	2	1627				
	3	1577				
	4	1413				
	5	1434				
	6	1553				
	7	1588				
	8	1570				
	9	1426				
	10	1426				
	11	1346				
	12	1094				
1988	1	1867				
	2	1619				
	3	1717				
	4	1386	1546	1338	1755	-160
	5	1493	1414	1206	1623	79
	6	1528	1608	1400	1817	-80
	7	1659	1567	1358	1775	92
	8	1899	1637	1428	1845	262
	9	1704	1743	1535	1952	-39
	10	1635	1706	1498	1915	-71
	11	1472	1558	1350	1767	-86
	12	1314	1224	1015	1432	90
1989	1	2051	2083	1874	2291	-32
	2	1709	1805	1596	2013	-96
	3	1848	1811	1603	2020	37
	4	1542	1579	1371	1788	-37
	5	1574	1617	1408	1825	-43
	6	1446	1644	1436	1853	-198
	7	1370	1547	1339	1756	-177
	8	1612	1515	1306	1723	97
	9	1426	1433	1224	1641	-7
	10	1566	1385	1177	1594	181
	11	1614	1428	1219	1636	186
	12	1489	1410	1202	1619	79
1990	1	2157	2237	2029	2446	-80
	2	1742	1856	1648	2065	-114
	3	1817	1846	1637	2054	-29
	4	1674	1529	1320	1737	145
	5	1772	1716	1507	1924	56
	6	1629	1720	1512	1929	-91
	7	1703	1625	1416	1833	78
	8	1859	1899	1691	2108	-40
	9	1585	1680	1471	1888	-95
	10	1581	1657	1448	1865	-76
	11	1470	1561	1353	1770	-91
	12	1248	1321	1112	1529	-73
1991	1	2287	1952	1744	2161	335
	2	1810	1901	1692	2109	-91
	3	1999	1899	1691	2108	100
	4	1860	1793	1584	2001	67
	5	1916	1935	1727	2144	-19
	6	1695	1811	1603	2020	-116
	7	1762	1741	1533	1950	21
	8	1816	1934	1726	2143	-118
	9	1621	1584	1376	1793	37
	10	1739	1644	1435	1852	95
	11	1444	1659	1450	1867	-215
	12	1158	1259	1051	1468	-101
1992	1		2066	1858	2275	
	2		1631	1343	1920	
	3		1779	1428	2129	
	4		1615	1211	2018	
	5		1680	1230	2130	
	6		1505	1012	1997	
	7		1561	1030	2093	
	8		1663	1096	2231	
	9		1451	850	2053	
	10		1532	898	2166	
	11		1324	660	1989	
	12		1075	381	1769	

Confidence Limits before Period 60 are 1-Step Ahead Limits

ภาคผนวก ข
การใช้โปรแกรมสำเร็จรูป MICROTSP

โปรแกรม MICROTSP เป็นโปรแกรมทางสถิติที่มีขีดความสามารถในการคำนวณ regression และ forecasting บนเครื่องคอมพิวเตอร์ส่วนบุคคล IBM ซึ่งสามารถใช้ในประเภทงานต่อไปนี้

- การพยากรณ์การขาย (Sales forecasting)
- การวิเคราะห์และพยากรณ์ราคาขาย (Cost analysis and forecasting)
- การวิเคราะห์การเงิน (Financial analysis)
- การพยากรณ์เศรษฐกิจแบบมหภาค (Macroeconomic forecasting)
- การจำลองตัวแบบ (Simulation)
- การวิเคราะห์ค่าข้อมูลทางวิทยาศาสตร์และการตีราคา (Scientific data analysis and evaluation)

ในที่นี้จะกล่าวถึงส่วนของโปรแกรมที่ใช้งานในปัญหาพิเศษนี้เท่านั้น ซึ่งในการวิเคราะห์ผลของปัญหาพิเศษใช้โปแกรมเพื่อจุดประสงค์ในการสร้างกราฟค่าของข้อมูล และการค้นหาค่าพารามิเตอร์ที่ดีที่สุดของการวิเคราะห์อนุกรมเวลาแบบมีฤดูกาลของ Winter คือ ค่าของอัลฟา (alpha) , เบต้า (beta) และแกมมา (gamma)

การเข้าใช้งานโปรแกรม TSP เมื่อโปรแกรมอยู่ใน Harddisk ทำได้โดย

1. เมื่ออยู่ในเครื่องหมาย c:\ ให้พิมพ์ cd tsp ได้ดังนี้

```
c:\cd tsp   <enter>
c:\tsp>
```

2. พิมพ์ qtsp เพื่อเข้าสู่โปรแกรม ได้ดังนี้

```
c:\TSP\qtsp   <enter>
```

รูปที่ 2.1

3. ได้ดังรูปที่ 2.1 จะปรากฏอยู่สักครู่หนึ่ง แล้วจะเข้าสู่หน้าต่างคำสั่ง ดังรูปที่ 2.2

รูปที่ 2.2

4. การเรียกใช้คำสั่งฟังก์ชันต่าง ๆ ของโปรแกรมทำได้ 2 วิธีคือ

4.1 การพิมพ์ชื่อคำสั่งฟังก์ชันที่หน้าต่างฟังก์ชัน

4.2 การกดแป้นพิมพ์ฟังก์ชัน (F keys) แล้วเลื่อนแถบสว่างไปยังคำสั่งที่ต้องการ แป้นฟังก์ชันที่ใช้มีดังนี้

- การกดแป้น F3 เรียกส่วนคำสั่งประเภทเกี่ยวกับไฟล์
- การกดแป้น F4 เรียกส่วนคำสั่งประเภทการจัดการข้อมูล
- การกดแป้น F5 เรียกส่วนคำสั่งประเภทการคำนวณทางสถิติ
- การกดแป้น F6 เรียกส่วนคำสั่งควบคุมโปรแกรม TSP

1 <u>การนำข้อมูลเข้าสู่โปรแกรม TSP</u>

ขั้นตอนการนำข้อมูลเข้าสู่โปรแกรม TSP โดยที่แฟ้มข้อมูลที่ต้องการนำเข้าเป็นแฟ้มข้อมูลตัวอักษร (ASCII File) มีดังนี้

3.1.1 กำหนดขนาดของแฟ้มข้อมูลที่ต้องการนำเข้าในโปรแกรม ด้วยคำสั่ง Create โดยต้องกำหนดประเภทของข้อมูลและจำนวนความยาวแถวของข้อมูล ดังรูปที่ 2.3

```
 range 1987.01 - 1991.12 | series: current=0  maximum=322 | output LPT1:
                         No series in work file
 current SMPL 1987.01 - 1991.12 | path C:                    print POFF

>create
 Frequency // Monthly
 Starting date // 1987.01
 Ending date // 1991.12

>







F1-BREAK  F2-LAST COMMAND | F3-FILES  F4-DATA  F5-STATISTICS  F6-TSP CONTROL
```

รูปที่ 2.3

3.1.2 อ่านข้อมูลจากดิสก์เข้าสู่โปรแกรมด้วยคำสั่ง read ซึ่งต้องกำหนดชื่อและ disk drive ของแฟ้มข้อมูล , ประเภทลักษณะของข้อมูลในแฟ้มข้อมูล และชื่อที่ใช้เรียกแฟ้มข้อมูลในโปรแกรม (Series list) ดังรูปที่ 2.4 ซึ่งสามารถทราบว่าโปรแกรมมีการอ่านแฟ้มข้อมูลใดบ้างด้วยการสังเกตชื่อที่ปรากฏในกรอบหน้าต่างส่วนบนของหน้าจอโปรแกรม

```
 range 1987.01 - 1991.12 | series: current-1 maximum-322 |   output LPT1:
TUBER
 current SMPL 1987.01 - 1991.12 | path C:                    print POFF

>create
 Frequency // Monthly
 Starting date // 1987.01
 Ending date // 1991.12

>read
 Data File Format // Data ordered by series
 File name // A:TUBER.FC
 Series list // TUBER

>




 F1-BREAK  F2-LAST COMMAND | F3-FILES  F4-DATA  F5-STATISTICS  F6-TSP CONTROL
```

รูปที่ 2.4

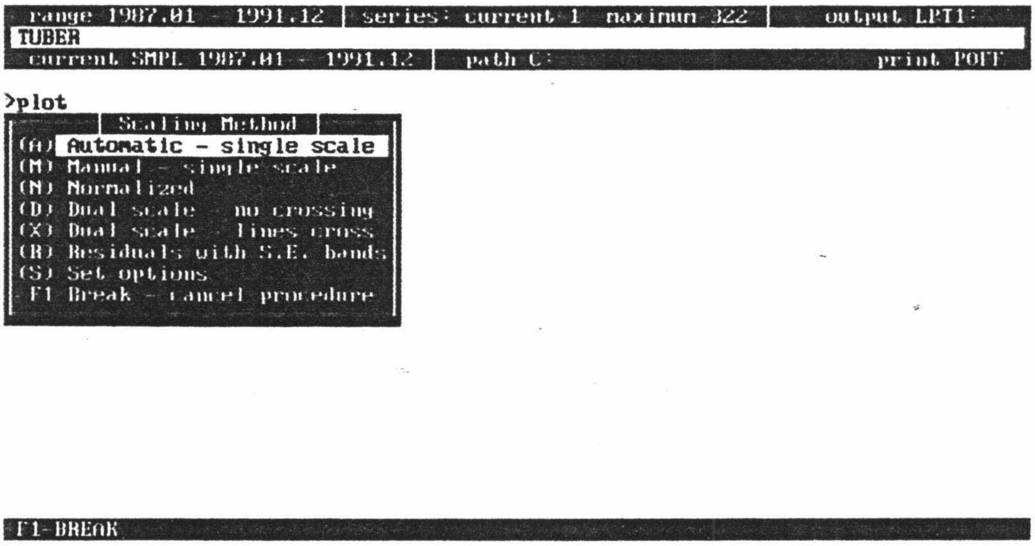
3.1.3 บันทึกข้อมูลที่เรียกเข้ามาด้วยคำสั่ง save ดังรูปที่ 2.5

```
 range 1987.01 - 1991.12 | series: current 1  maximum 322 |   output LPT1
TUBER
 current SMPL 1987.01  1991.12 | path C:                      print POFF

>save
 File name // C:TUBER.TSP
 Dropping file extension.

>




F1-BREAK  F2-LAST COMMAND | F3-FILES  F4-DATA  F5-STATISTICS  F6-TSP CONTROL
```

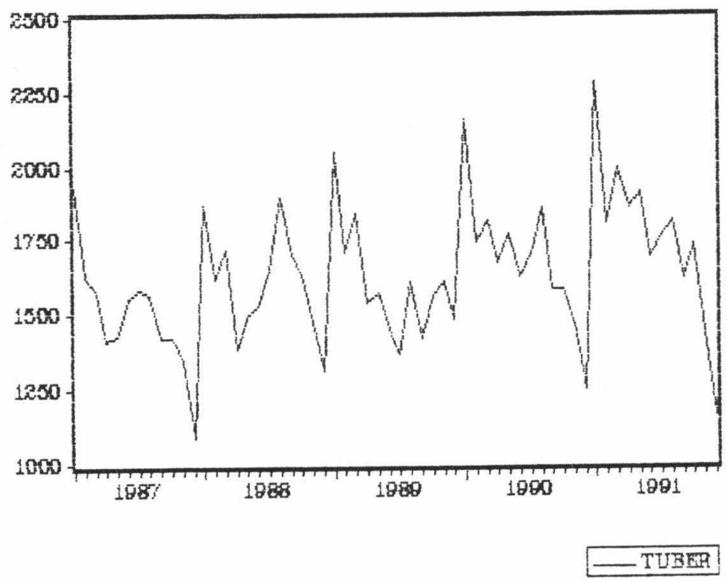
รูปที่ 2.5

2. <u>การสร้างกราฟข้อมูล</u>

การสร้างกราฟข้อมูลด้วยวิธีการเรียกคำสั่ง Plot ซึ่งต้องกำหนดชื่อของแฟ้มข้อมูลที่ใช้ในโปรแกรม (Series list) จะปรากฏเมนูดังรูปที่ 2.6 ซึ่งสามารถกำหนดตัวเลือกต่าง ๆ ของกราฟในคำสั่ง Set Options ถ้าเป็นกราฟปกติทั่ว ๆ ไป ใช้คำสั่ง Automatic แล้วจะปรากฏรูปลักษณะกราฟดังรูปที่ 2.7 ซึ่งสามารถพิมพ์ออกเครื่องพิมพ์ด้วยการกดแป้นพิมพ์ P และกดแป้นพิมพ์ X เพื่อออกจากโปรแกรม


range 1987.01 - 1991.12
series: current 1 maximum 322
output LPT1:
TUBER
current SMPL 1987.01 - 1991.12
path C:
print POFF
>plot
Scaling Method
(A) Automatic - single scale
(M) Manual - single scale
(N) Normalized
(D) Dual scale - no crossing
(X) Dual scale - lines cross
(R) Residuals with S.E. bands
(S) Set options
F1 Break - cancel procedure
F1-BREAK

รูปที่ 2.6


2500
2250
2000
1750
1500
1250
1000
1987
1988
1989
1990
1991
TUBER

รูปที่ 2.7

ภาคผนวก ค

การใช้โปรแกรมสำเร็จรูป MINITAB

MINITAB เป็นโปรแกรมสำเร็จรูป ที่ใช้ในการวิเคราะห์อนุกรมเวลาโดยใช้เทคนิคบอกซ์และเจนกินส์ และใช้วิเคราะห์โดยวิธีอื่น ๆ ได้อีก แต่ในปัญหาพิเศษนี้จะใช้เฉพาะเทคนิคบอกซ์และเจนกินส์

1. เมื่ออยู่ในเครื่องหมาย c:\ ให้พิมพ์ cd minitab ได้ดังนี้

   c:\cd minitab <enter>

   c:\MINITAB\

2. ให้พิมพ์ minitab เพื่อเข้าสู่โปรแกรม

   c:\MINITAB\minitab <enter>

   ได้ดังรูป

M I N I T A B

Data Analysis Software
Release 6.1.1 — Standard Version
Copyright(C) Minitab, Inc. 1987

This software is licensed to:
Microcomputer Center TU.

Serial # 611-0058-000457

You may use Minitab under the terms of the License Agreement enclosed with this program; please read it. This License entitles: a) one user to run this copy of Minitab on any number of computers; b) more than one user to run this copy of Minitab on a single computer, BUT it is a violation of the License to run this copy of Minitab on more than one computer simultaneously. Government users see HELP FGU.

Press any key to continue.

3. อ่านข้อมูลจากไฟล์ที่ได้ทำการบันทึกเอาไว้แล้ว ด้วยคำสั่ง

   Read data [from'FILENAME'] into C...C

   เช่น MTB> read data [from 'a:b_tuber.fc'] into C1

   MTB>

ซึ่งจะเป็นการอ่านข้อมูล พร้อมทั้งเก็บไว้ในคอลัมน์ที่ 1 ถึง n ใด ๆ ที่ได้กำหนดขึ้นมา

4. จากคำสั่งข้างต้น โปรแกรมจะแสดงข้อมูลเริ่มต้น 4 ค่า ดังนี้

C1

119 107 126 99 ...

5. ทำการทดสอบดูสหสัมพันธ์ในตัวเอง และสหสัมพันธ์ในตัวเองบางส่วนโดยใช้คำสั่ง

ACF [with up to K lag] for series in C [put into C]

PACF [with up to K lag] for series in C [put into C]

เช่น

ACF c1 <enter>

PACF c1 <enter>

ซึ่งจะเป็นการทดสอบดูสหสัมพันธ์ในตัวเอง และสหสัมพันธ์ในตัวเองบางส่วนในคอลัมน์ c1

6. โปรแกรมจะแสดงรูปสหสัมพันธ์ในตัวเอง และสหสัมพันธ์ในตัวเองบางส่วนตามลำดับ ดังนี้

```
ACF of C1

           -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
             +----+----+----+----+----+----+----+----+----+----+
   1   0.334                          XXXXXXXXX
   2   0.168                          XXXXX
   3   0.005                          X
   4   0.010                          X
   5   0.109                          XXXX
   6   0.084                          XXX
   7   0.007                          X
   8   0.065                          XXX
   9  -0.039                         XX
  10   0.091                          XXX
  11  -0.031                         XX
  12   0.014                          X
  13   0.039                          XX
  14   0.030                          XX
  15   0.089                          XXX
  16   0.069                          XXX
  17  -0.029                         XX
  18   0.006                          X
  19  -0.089                        XXX
MTB >
```

```
PACF of C1

           -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
             +----+----+----+----+----+----+----+----+----+----+
   1   0.334                          XXXXXXXXX
   2   0.064                          XXX
   3  -0.078                        XXX
   4   0.022                          XX
   5   0.128                          XXXX
   6   0.011                          X
   7  -0.062                        XXX
   8   0.093                          XXX
   9  -0.078                        XXX
  10   0.107                          XXXX
  11  -0.095                        XXX
  12   0.035                          XX
  13   0.045                          XX
  14   0.002                          X
  15   0.068                          XXX
  16   0.015                          X
  17  -0.060                         XX
  18   0.000                          X
  19  -0.064                        XXX
MTB >
```

7. ทำการหาตัวแบบในวิธี Box-Jenkins

กรณีข้อมูลไม่มีฤดูกาล ใช้คำสั่ง

```
ARIMA p=K, d=K, q=K, data in C [put residuals in C
     [predicteds in C [coefficients in C]]]
```

กรณีข้อมูลมีฤดูกาล ใช้คำสั่ง

```
ARIMA  p=K, d=K, q=K, P=K ,D=K ,Q=K ,S=K data in C
     [put residuals in C [predicteds in C
          [coefficients in C]]]
```

เมื่อ p คือ จำนวนพารามิเตอร์การถดถอยในตัวเองแบบปกติไม่มีฤดูกาล

d คือ จำนวนครั้งของการหาค่าความแตกต่างแบบปกติไม่มีฤดูกาล

q คือ จำนวนพารามิเตอร์ของการเฉลี่ยเคลื่อนที่แบบปกติไม่มีฤดูกาล

P คือ จำนวนพารามิเตอร์การถดถอยในตัวเองแบบมีฤดูกาล

D คือ จำนวนครั้งของการหาค่าความแตกต่างแบบมีฤดูกาล

Q คือ จำนวนพารามิเตอร์ของการเฉลี่ยเคลื่อนที่แบบมีฤดูกาล

C คือ ข้อมูลในคอลัมน์ที่ใช้วิเคราะห์

เช่น

```
ARIMA 1 1 0 C1
```

ซึ่งเป็นการทดสอบตัวแบบ ARIMA โดยกำหนดให้ พารามิเตอร์การถดถอยในตัวเอง มีค่าเท่ากับ 1 ค่าพารามิเตอร์ของการเฉลี่ยเคลื่อนที่ มีค่าเท่ากับ 0 และค่าความแตกต่างแรกเป็น 1 ดังค่าต่าง ๆ ต่อไปนี้

```
Estimates at each iteration
Iteration         SSE      Parameters
    0           133294     0.100
    1           120953    -0.050
    2           113551    -0.200
    3           111093    -0.342
    4           111087    -0.349
    5           111087    -0.350
    6           111087    -0.350
Relative change in each estimate less than  0.0010

Final Estimates of Parameters
Type       Estimate     St. Dev.  t-ratio
AR   1     -0.3499      0.1117     -3.13

Differencing: 1 regular difference
No. of obs.:  Original series 83, after differencing 82
Residuals:    SS =  111071  (backforecasts excluded)
              MS =    1371  DF = 81

Modified Box-Pierce chisquare statistic
Lag                   12            24            36            48
Chisquare   12.9(DF=11)   17.1(DF=23)   23.5(DF=35)   43.3(DF=47)
MTB >
```

ภาคผนวก ง

การใช้โปรแกรมสำเร็จรูป STATGRAPHIC วิเคราะห์อนุกรมเวลา

การเข้าใช้งาน " โปรแกรมสำเร็จรูป STATGRAPHIC " นั้นมีโปรแกรม STATGRAPHIC ใน Harddisk แล้ว สามารถเรียกใช้โปรแกรมตามขั้นตอน ดังนี้

1. เมื่ออยู่ในเครื่องหมาย c:\ ให้พิมพ์ cd statgraf ได้ดังนี้

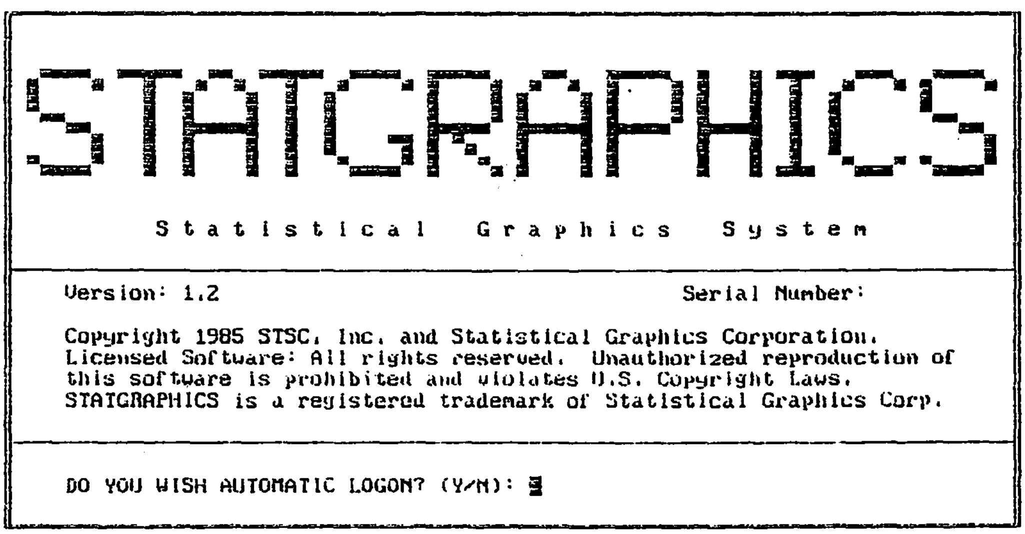
   c:\cd statgraf <Enter>

   c:\statgraf>

2. พิมพ์ statgraf เพื่อเข้าสู่โปรแกรม ได้ดังนี้

   c:\statgraf\statgraf <Enter>

3. ได้ดังรูปที่ 1.1 สามารถกดแป้นตัวอักษร "Y" หรือ <Enter> ได้ดังรูปที่ 1.2 เพื่อใช้ค่ากำหนดติดตั้งที่เกี่ยวกับ Hardware เดิม หากต้องการแก้ไขค่าติดตั้งเหล่านี้ใช้การกดแป้น "N" แล้วทำการกำหนดค่าติดตั้งนี้ให้เหมาะสมกับการใช้งาน



รูปที่ 1.1

4. กด <Enter> ได้ดังรูปที่ 1.3

5. กด <Enter> ได้ดังรูปที่ 1.4 ซึ่งเป็นเมนูหลักของโปรแกรม

```
STATGRAPHICS

Statistical Graphics
        System

     Press ENTER to begin.
```

รูปที่ 1.2

```
WORKAREA AVAILABLE: 168 KB
MATH CO-PROCESSOR AVAILABLE: NO

PLEASE PLACE:

  -  EITHER STATGRAPHICS PROGRAM DISK INTO DRIVE C
  -  YOUR DATA DISK INTO DRIVE C

Then press ENTER to continue.
```

รูปที่ 1.3

```
              STATGRAPHICS Statistical Graphics System

DATA HANDLING AND SYSTEM UTILITIES          TIME SERIES PROCEDURES
  A. Data Management                          L. Forecasting
  B. System Environment                       M. Quality Control
  C. Report Writer and Graphics Replay        N. Smoothing
  D. Plotter Interface                        O. Time Series Analysis

PLOTTING AND DESCRIPTIVE STATISTICS         ADVANCED PROCEDURES
  E. Plotting Functions                       P. Categorical Data Analysis
  F. Descriptive Methods                      Q. Multivariate Methods
  G. Estimation and Testing                   R. Nonparametric Methods
  H. Distribution Functions                   S. Sampling
  I. Exploratory Data Analysis                T. Experimental Design

ANOVA AND REGRESSION ANALYSIS               MATHEMATICAL PROCEDURES
  J. Analysis of Variance                     U. Numerical Analysis
  K. Regression Analysis                      V. Mathematical Functions

        Use cursor keys to highlight desired section.  Then press ENTER.

 1HELP  2LABEL  3SAVSC  4RECORD 5     6     7     8     9REVIEW 10QUIT
INPUT:  SUN MAR 14  1993 01:46:00 PM  VERSION 1.2                REC:OFF
```

รูปที่ 1.4

จะสังเกตเห็นได้ว่า ในเมนูหลักนั้นมีฟังก์ชันให้เลือก ตั้งแต่ A-V โดยทำการเลือกฟังก์ชันทำได้โดย การกดแป้นพิมพ์ตัวอักษร A-V ตามฟังก์ชันที่ต้องการหรือใช้แป้นพิมพ์ลูกศรเพื่อเคลื่อนแถบสว่างไปยังตำแหน่งของฟังก์ชันที่ต้องการ

ในภาคผนวกอธิบายเฉพาะฟังก์ชันที่ใช้ในปัญหาพิเศษนี้เท่านั้น

1. การจัดการเกี่ยวกับข้อมูล

การจัดการเกี่ยวกับข้อมูลสำหรับ STATGRAPHIC นั้นสามารถนำเข้าข้อมูลที่เป็นไฟล์ข้อความ (ASCII FILE) ที่ถูกสร้างขึ้นจาก Editor ทั่วไปได้ เช่น SK, QEDIT เป็นต้น เพื่อใช้งานในโปรแกรม หรือ ข้อมูลที่สร้างจาก Worksheet ของโปรแกรม Lotus รวมทั้งสร้างข้อมูลจากโปรแกรมเองได้ ในที่นี้จะอธิบายเฉพาะการนำเข้าข้อมูลที่เป็นไฟล์ข้อความ (ASCII FILE) เท่านั้น โดยทำตามขั้นตอนดังต่อไปนี้

1.1 จากเมนูหลัก ดังรูปที่ 1.4 เลื่อนแถบสว่างไปที่

A. Data Management

กด <Enter> ได้ดังรูปที่ 1.5

1.2 ครั้งแรกของการใช้ข้อมูลชุดนั้น ต้องทำการนำข้อมูลเข้ามาในส่วนของโปรแกรมก่อน เมื่ออยู่ในหน้าจอดังรูปที่ (1.5) เลื่อนแถบสว่างไปที่

5. Import Data from ASCII Data File

กด <Enter> ได้ดังรูปที่ 1.6

1.4 ใส่ข้อมูลตามที่โปรแกรมต้องการ เมื่อให้ข้อมูลครบแล้วได้ดังรูปที่ 1.8

1.6 เมื่อกด <Enter> จะปรากฏหน้าจอดังรูปที่ 1.5

หมายเหตุ : ขั้นตอน 1.2-1.6 ทำเฉพาะครั้งแรกที่ใช้ข้อมูลชุดนั้น ในการทำงานครั้งต่อไปกับข้อมูลชุดนั้นไม่จำเป็นต้องทำให้ข้ามไปทำในขั้นตอน 1.7)

1.7 เลื่อนแถบสว่างไปที่

3. Read Variable Definitions from SG File

กด <Enter> ได้ดังรูปที่ 1.10

1.8 เลือกไฟล์ที่ต้องการวิเคราะห์ เช่น เลื่อนแถบแสงไปที่ K291 แล้วกด <Enter> ได้ดังรูปที่ 1.11

1.9 เลื่อนแถบแสงลงมาที่ชื่อไฟล์ที่ต้องการวิเคราะห์ เช่น เลื่อนแถบแสงไปที่ K291 แล้วกด <A> เพื่อทำการอ่านข้อมูลทั้งหมด (A=READ ALL VARIABLES)

1.10 เมื่อสิ้นสุดการนำค่าตัวแปรข้อมูลเข้าสู่โปรแกรม สามารถกด <F10> เพื่อออกจากเมนูฟังก์ชันย่อยที่ใช้งานจนกระทั่งถึงเมนูหลัก

1.11 หากต้องการดูรายชื่อตัวแปรที่ถูกอ่านเข้าในโปรแกรมแล้ว รวมทั้งดูข้อมูลที่เกิดจากการอ่านและ การคำนวณของโปรแกรม เลื่อนแถบสว่างไปที่

1. Manipulate Defined Variables

กด <Enter> ได้ดังรูปที่ 1.12

1.12 เลือกชื่อตัวแปรที่ต้องการวิเคราะห์ เลื่อนแถบสว่างลงมาที่ตัวแปรที่ต้องการแล้วกดแป้นตัวอักษร D เพื่อทำการแสดงข้อมูลทั้งหมด (D=DISPLAY) ได้ดังรูปที่ 1.13

```
DATA MANAGEMENT

 1.  Manipulate Defined Variables
 2.  Full-Screen Data Editor
 3.  Read Variable Definitions from SG File
 4.  Write to STATGRAPHICS File
 5.  Import Data from ASCII Data File
 6.  Export Data to ASCII Data File
 7.  Import Data from DIF File
 8.  Export Data to DIF File
 9.  Import Lotus 1-2-3 Worksheet
10.  Export Lotus 1-2-3 Worksheet
11.  Recode Missing Values

Use cursor keys to highlight desired procedure.
Then press ENTER.

1HELP  2LABEL  3SAVSC  4RECORD 5   6   7   8   9REVIEW 10QUIT
INPUT   SUN MAR 14  1993 01:50:00 PM  VERSION 1.2      REC:OFF
```

รูปที่ 1.5

```
DO YOU WISH TO READ AS A FORMATTED OR UNFORMATTED FILE (F/U):

1HELP  2LABEL  3SAVSC  4RECORD 5   6   7   8   9REVIEW 10QUIT
INPUT   SUN MAR 14  1993 02:00:00 PM  VERSION 1.2      REC:OFF
```

รูปที่ 1.6

```
                              Import ASCII Files
--------------------------------------------------------------------------------

Enter the name of the ASCII file to read (Display Files): k291
Enter the maximum record width (Quit): 3
Enter the STATGRAPHICS File name (Up to 8 Characters): k291








 1HELP   2LABEL  3SAVSC  4RECORD 5       6       7       8       9REVIEW 10QUIT
INPUT    SUN MAR 14  1993 02:06:00 PM  VERSION 1.2                      REC:OFF
```

รูปที่ 1.7

```
                          Import ASCII Files
          File: C:K291.DAT   Record Size: 3   STATGRAPHICS File: K291
           Name      Type Start End           Name      Type Start End
           k291       N       1   3

Enter field information, F5 to display, F6 to extract names from first record.
*1HELP  2      3      4      5DISPLY 6XTRACT 7      8      9       10QUIT
INPUT    SUN MAR 14  1993 02:08:00 PM  VERSION 1.2                      REC:OFF
```

รูปที่ 1.8

```
                    Import ASCII Files
------------------------------------------------------------------
File processing complete.  STATGRAPHICS File Directory is:
VARIABLE NAME   TYPE  RANK  LENGTH       DATE      TIME  COMMENT
k291              N      1   49           3/14/93  14:10
Press ENTER to continue.




1HELP  2LABEL  3SAVSC  4RECORD 5     6     7     8     9REVIEW 10QUIT
INPUT    SUN MAR 14  1993 02:10:00 PM  VERSION 1.2              REC OFF
```

รูปที่ 1.9

```
THE FOLLOWING FILES ARE AVAILABLE:

FILENAME    DISK     FILENAME    DISK
------------------   ------------------
C2           C       K291         C
C290         C       K294         C
C291         C       K295         C
C294         C       K6           C
C295         C       K8           C
CCHART       C       TESTQC       C
CUTUSA       C       TSDATA       C
IRONORE      C       TUSA         C
K0           C       UCHART       C
K1           C       USA          C
K2           C


USE CURSOR KEYS TO HIGHLIGHT DESIRED FILE. THEN PRESS ENTER.
1HELP  2LABEL  3SAVSC  4RECORD 5     6     7     8     9REVIEW 10QUIT
INPUT    SUN MAR 14  1993 02:16:00 PM  VERSION 1.2              REC OFF
```

รูปที่ 1.10

```
THE FOLLOWING VARIABLES ARE CURRENTLY IN FILE C:K291:

VARIABLE NAME    TYPE  RANK  LENGTH       DATE     TIME  COMMENT
--------------------------------------------------------------------------------
£DIRECTORY                                3/14/93 14:10 FILE DIRECTORY
K291               N      1   49          3/14/93 14:10




         USE CURSOR KEYS TO HIGHLIGHT DESIRED VARIABLE, THEN PRESS:
  A=READ ALL VARIABLES    D=DISPLAY SELECTED VARIABLE    R=READ SELECTED VARIABLE
 1HELP  2LABEL  3SAVSC  4RECORD 5       6        7         8        9REVIEW 10QUIT
INPUT   SUN MAR 14  1993 02:19:00 PM  VERSION 1.2                           REC:OFF
```

รูปที่ 1.11

```
THE FOLLOWING VARIABLES ARE CURRENTLY DEFINED:

VARIABLE NAME      DISK  FILENAME  TYPE  RANK  LENGTH
------------------------------------------------------
K291                 C    K291       N     1    49




         USE CURSOR KEYS TO HIGHLIGHT DESIRED VARIABLE, THEN PRESS:
A=ASSIGN C=COPY D=DISPLY E=ERASE F=CHAR M=NUM N=NEW R=RGNAME S=RESHAPE V=VERIFY
 1HELP  2LABEL  3SAVSC  4RECORD 5       6        7         8        9REVIEW 10QUIT
INPUT   SUN MAR 14  1993 07:06:00 PM  VERSION 1.2                           REC:OFF
```

รูปที่ 1.12

```
k291
27 38 24 35 27 26 31 33 38 38 34 32 32 42 38 47 29 39 32 36 24 46 43 48 34 39
      32 43 34 32 31 38 28 39 36 33 41 26 22 33 29 39 31 27 18 38 31 27 -32768
```

```
1HELP 2LABEL 3SAVSC 4RECORD 5      6      7      8      9REVIEW 10QUIT
INPUT  SUN MAR 14  1993 07:09:00 PM  VERSION 1.2                  REC:OFF
```

รูปที่ 1.13

2. <u>การพยากรณ์ (Forecasting)</u>

จากเมนูหลัก ดังรูปที่ 1.4 เลื่อนแถบสว่างไปที่

L. Forecasting

กด <Enter> ได้ดังรูปที่ 1.14

```
FORECASTING

1.  Brown's Exponential Smoothing
2.  Holt's Linear Exponential Smoothing
3.  Winter's Seasonal Smoothing
4.  Simple Trend Analysis
5.  Exponential Power Curve
6.  Life-Cycle Fitting
7.  Seasonal Decomposition

Use cursor keys to highlight desired procedure.
Then press ENTER.

1HELP  2LABEL  3SAVSC  4RECORD 5  6  7  8  9REVIEW 10QUIT
INPUT  SUN MAR 14  1993 02:21:08 PM  VERSION 1.2  REC:OFF
```

รูปที่ 1.14

จากรูปที่ 1.14 แสดงให้เห็นถึงวิธีที่ใช้ในการพยากรณ์ของโปรแกรมสำเร็จรูป Statgraphic ที่สามารถทำได้

เมื่อต้องการพยากรณ์อนุกรมด้วยอนุกรมเวลาด้วยวิธีใดทำได้โดยเลื่อนแถบสว่างไปที่หัวข้อที่สนใจ กด <Enter> ในที่นี้กล่าวถึงเฉพาะวิธีที่ใช้ในปัญหาพิเศษนี้ ซึ่ง

ตัวอย่างหน้าจอของวิธีต่าง ๆ ที่ปรากฏ

-ดังรูปที่ 1.15 "Brown's Exponential Smoothing"

-ดังรูปที่ 1.16 "Holt's Exponential Smoothing"

-ดังรูปที่ 1.17 "Winter Seasonal Smoothing"

```
ENTER THE NAME OF THE VARIABLE CONTAINING YOUR DATA:  K291
ENTER THE NUMBER OF FORECASTS DESIRED (12):
ENTER A VALUE FOR THE SMOOTHING CONSTANT 0<ALPHA<1 (0.1):
SIMPLE, LINEAR, OR QUADRATIC SMOOTHING? (0/1/2):



1HELP  2LABEL  3SAVSC  4RECORD 5      6      7      8      9REVIEW 10QUIT
INPUT  SUN MAR 14  1993 02:26:00 PM  VERSION 1.2                 REC:OFF
```

รูปที่ 1.15

```
ENTER THE NAME OF THE VARIABLE CONTAINING YOUR DATA:  K291
ENTER THE NUMBER OF FORECASTS DESIRED (12):
ENTER A VALUE FOR THE SMOOTHING CONSTANT 0<ALPHA<1 (0.1):
ENTER A VALUE FOR THE LINEAR CONSTANT 0<BETA<1 (0.1):
ENTER A VALUE FOR THE SEASONAL CONSTANT 0<GAMMA<1 (0.1):
ENTER THE LENGTH OF SEASONALITY (12):



1HELP  2LABEL  3SAVSC  4RECORD 5      6      7      8      9REVIEW 10QUIT
INPUT  SUN MAR 14  1993 02:34:00 PM  VERSION 1.2                 REC:OFF
```

รูปที่ 1.16

```
ENTER THE NAME OF THE VARIABLE CONTAINING YOUR DATA:  k291
ENTER THE NUMBER OF FORECASTS DESIRED (12):
ENTER A VALUE FOR THE SMOOTHING CONSTANT 0<ALPHA<1 (0.1):
ENTER A VALUE FOR THE LINEAR CONSTANT 0<BETA<1 (0.1):









 1HELP  2LABEL  3SAVSC  4RECORD 5      6      7      8      9REVIEW 10QUIT
INPUT   SUN MAR 14  1993 02:30:00 PM  VERSION 1.2                   REC:OFF
```

รูปที่ 1.17

บรรณานุกรม

Anderson,O.D. Time Series Analysis and Forecasting The Box-Jenkins approach, pp.6-133, Butterworth, London, 1976.

David,S.Walonick. FORCAST PLUS,1986 (copy).

Bruce,L. Bowerman & Richard,T.O' Connell. Time Series and Forecasting, pp.1-450, Duxbury Press, Massachusetts, 1979.

Granger,C.W.J. & Newbold,P. Forecasting Economic Time Series, pp.1-42, Academic press, New York, 1977.

Johnson,L.A. & Montgomery,D.C. Forecasting the Time Series Analysis, pp.43-75, McGraw-Hill, New York, 1976.

Wheelright,S.C. & Makridakis,S. Forecasting Methods for Management, pp.22-204, John Wiley & Sons, New York, 1976.

กฤษฎา วาณิชชัชวาลย์ และ คณะ. การวิเคราะห์แนวโน้มจำนวนนักท่องเที่ยวชาวต่างประเทศที่เข้ามาในประเทศไทย และอื่นๆที่เกี่ยวข้อง, วิทยานิพนธ์วิทยาศาสตรบัณฑิต ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2534.

จันทรนิวัทธ์ เกษมสันต์ และ บุญชอบ พงษ์พาณิชย์. กุมารเวชศาสตร์, พิมพ์ครั้งที่ 1, หน้าที่ 65-285, โครงการตำราศิริราช, 2522

พิพัฒน์ ลักษมีจรัลกุล และ เสาวลักษณ์ ลักษมีจรัลกุล. "การตอบสนองของแอนติบอดีต่อโรคไอกรน ในเด็กที่ได้รับวัคซีนป้องกันโรคคอตีบ บาดทะยัก และ ไอกรน ครบ 3 ครั้ง" <u>รวมบทคัดย่อการวิจัยทางการแพทย์ มหิดล</u> (2528):193-194.

พิภพ สมิทธินันท์ และ คณะ. <u>การวิเคราะห์แนวโน้มเพื่อพยากรณ์ความต้องการใช้น้ำมันเชื้อเพลิงของประเทศไทย ในปีพุทธศักราช 2533</u> วิทยานิพนธ์วิทยาศาสตรบัณฑิต ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2532.

สมชาย สุพันธ์วณิช. <u>หลักระบาดวิทยา</u>, หน้า 55-190, โรงพิมพ์สามมิตร, 2524.

วิชิต หล่อจีระชุณห์กุล และ คณะ. <u>เทคนิคการพยากรณ์เชิงสถิติ</u>, พิมพ์ครั้งที่ 1, หน้าที่ 157-186, โรงพิมพ์เรือนแก้วการพิมพ์, 2524.

อะเคื้อ อุณหะเลขกะ และ คณะ. "ระบาดวิทยาของโรคโปลิโอในเขตกรุงเทพมหานคร พ.ศ.2526-2527" <u>รวมบทคัดย่อการวิจัยทางการแพทย์ มหิดล</u> (2528):225-231.

ประวัตินักศึกษาผู้จัดทำ

ชื่อ นางสาวกิตติยา เลียงเครือ

วัน เดือน ปี เกิด 12 มีนาคม พ.ศ. 2515

สำเร็จมัธยมต้นจากโรงเรียน สาธิตวิทยาลัยครูพระนคร เมื่อปีการศึกษา 2529

สำเร็จมัธยมปลายจากโรงเรียน สาธิตวิทยาลัยครูพระนคร เมื่อปีการศึกษา 2532

ชื่อ นางสาวอรุณี ผลมณี

วัน เดือน ปี เกิด 17 ตุลาคม พ.ศ. 2514

สำเร็จมัธยมต้นจากโรงเรียน ท่าใหม่ "พูลสวัสดิ์ราษฎร์นุกูล" เมื่อปีการศึกษา 2529

สำเร็จมัธยมปลายจากโรงเรียน ท่าใหม่ "พูลสวัสดิ์ราษฎร์นุกูล" เมื่อปีการศึกษา 2532

ชื่อ นายอุดม ธรรมอมฤตกุล

วัน เดือน ปี เกิด 9 เมษายน พ.ศ. 2515

สำเร็จมัธยมต้นจากโรงเรียน ชลราษฎรอำรุง เมื่อปีการศึกษา 2529

สำเร็จมัธยมปลายจากโรงเรียน ชลราษฎรอำรุง เมื่อปีการศึกษา 2532