บทที่ 3

วิธีดำเนินการวิจัย

ในการวิจัยเรื่องการพยากรณ์แนวโน้มของการเกิดโรคที่ต้องเฝ้าระวังทางระบาดวิทยาของจังหวัดฉะเชิงเทรานี้ ต้องการศึกษาเปรียบเทียบวิธีการพยากรณ์ โดยใช้การวิเคราะห์อนุกรมเวลา (Time Series Analysis) ซึ่งวิธีการพยากรณ์ที่นำมาใช้ในการวิจัยครั้งนี้ ประกอบด้วย วิธีปรับให้เรียบแบบเลขชี้กำลังครั้งเดียว วิธีปรับให้เรียบแบบเลขชี้กำลังสองครั้ง วิธีการของโฮลท์ วิธีการของวินเตอร์ วิธีการบอกซ์และเจนกินส์ การวิเคราะห์อนุกรมเวลาแบบคลาสสิค และวิธีการพยากรณ์ที่เป็นค่าพยากรณ์ร่วม โดยการให้ค่าเฉลี่ยน้ำหนัก ดังต่อไปนี้ วิธีการให้น้ำหนักที่เท่ากัน และวิธีการหาค่าพยากรณ์ร่วมด้วยวิธีค่าสัมบูรณ์ต่ำสุด ซึ่งผู้วิจัยจะนำผลการพยากรณ์ดังวิธีข้างต้น โดยพิจารณาจากการเปรียบเทียบของค่าเฉลี่ยร้อยละของค่าความคลาดเคลื่อนสัมบูรณ์ต่ำสุด วิธีนั้นก็จะเป็นวิธีที่เหมาะสมกับข้อมูลอนุกรมเวลาชุดนั้น ๆ

3.1 กลุ่มตัวอย่าง

กลุ่มตัวอย่าง คือ จำนวนผู้ป่วยด้วยโรคที่ต้องเฝ้าระวังทางระบาดวิทยา ที่สำนักงานสาธารณสุขจังหวัดฉะเชิงเทรา ตามระบบการเฝ้าระวังโรคทางระบาดวิทยา ซึ่งในการวิจัยครั้งนี้เลือกเฉพาะโรคที่ต้องเฝ้าระวังทางระบาดวิทยา ที่เป็นปัญหาสาธารณสุขจังหวัดฉะเชิงเทรา และในปัจจุบันยังไม่สามารถควบคุมให้ลดจำนวนลงได้ในช่วงเวลาต่าง ๆ จำนวน 10 โรค ด้วยกัน ได้แก่ โรคอุจจาระร่วง, โรคตาแดง, โรคไข้หวัดใหญ่, โรคอาหารเป็นพิษ, โรคตับอักเสบ, โรคไข้เลือดออก, โรคปอดบวม, โรควัณโรค, โรคหัด และโรคสุกใส

3.2 เครื่องมือที่ใช้ในการวิจัย

3.2.1 เครื่องไมโครคอมพิวเตอร์ โปรเซสเซอร์ 486 DX 2 (หน่วยความจำหลัก 8 เมกกะไบต์) พร้อมอุปกรณ์ จำนวน 1 ชุด

3.2.2 โปรแกรมสำเร็จรูป SPSS$^+$ for windows, Minitab และQSB$^+$

3.3 การเก็บรวบรวมข้อมูล

3.3.1 การเก็บรวบรวมข้อมูลการป่วย ด้วยโรคที่ต้องเฝ้าระวังทางระบาดวิทยาจากแบบบันทึกจำนวนผู้ป่วย เป็นรายเดือน(E.2, E.3) ของจังหวัดฉะเชิงเทราตั้งแต่ปี พ.ศ. 2524 – 2541 ซึ่งเป็นข้อมูลที่เรียบเรียงจากบัตรรายงานผู้ป่วย(รง.506) และบัตรเปลี่ยนแปลงการรายงานผู้ป่วย (รง.507) จากกองระบาดวิทยา สาธารณสุขจังหวัดฉะเชิงเทรา

3.3.2 บันทึกข้อมูลการป่วย ด้วยโรคที่ต้องเฝ้าระวังทางระบาดวิทยา เป็นรายเดือน ลงในแผ่นดิสก์

3.4 การวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ประกอบด้วยข้อมูลอนุกรมเวลาจำนวน 10 ตัวแบบ โดยใช้ข้อมูลการป่วยรายเดือนทั้งหมด 180 หน่วย (ข้อมูลในปี พ.ศ. 2526 - 2540) ซึ่งเป็นส่วนที่นำมาวิเคราะห์หาตัวแบบที่เหมาะสม ซึ่งจากการคำนวณเมื่อได้ตัวแบบที่ใช้พยากรณ์ข้อมูลก็จะทำการพยากรณ์ล่วงหน้า 12 คาบเวลา (พยากรณ์ล่วงหน้าในปี พ.ศ. 2541) จากนั้นนำข้อมูลที่พยากรณ์ได้นี้ไปเปรียบเทียบกับข้อมูลในส่วนที่เหลือไว้จำนวน 12 คาบเวลา (ข้อมูลในปี พ.ศ. 2541) จากนั้นนำมาตรวจสอบความแม่นยำในการพยากรณ์ โดยการหาค่าเฉลี่ยร้อยละของความคลาดเคลื่อนจากการพยากรณ์ (MAPE) ซึ่งการรวบรวมข้อมูลของจำนวนผู้ป่วย ของโรคที่ต้องเฝ้าระวังทางระบาดวิทยามาจากกองระบาดวิทยา สาธารณสุขจังหวัดฉะเชิงเทรา และเมื่อได้ตัวแบบพยากรณ์ที่เหมาะสมแล้ว จึงทำการพยากรณ์จำนวนผู้ป่วยเป็นรายเดือน และอัตราการป่วยเป็นรายปี โดยพยากรณ์ล่วงหน้าตั้งแต่ ปี พ.ศ. 2542 - 2544 ซึ่งข้อมูลอนุกรมเวลาประกอบด้วย

ตัวแบบที่ 1 จำนวนผู้ป่วยโรคอุจจาระร่วง (Diarrhoea)
ตัวแบบที่ 2 จำนวนผู้ป่วยโรคอาหารเป็นพิษ (Food poisoning)
ตัวแบบที่ 3 จำนวนผู้ป่วยโรคตับอักเสบ(Virus B hepatitis)
ตัวแบบที่ 4 จำนวนผู้ป่วยโรคตาแดง (Haemorrhagic conjunctivitis)
ตัวแบบที่ 5 จำนวนผู้ป่วยโรคไข้หวัดใหญ่ (Influenza)
ตัวแบบที่ 6 จำนวนผู้ป่วยโรคสุกใส (Chickenpox)
ตัวแบบที่ 7 จำนวนผู้ป่วยโรคหัด (Measles)
ตัวแบบที่ 8 จำนวนผู้ป่วยโรคไข้เลือดออก (Dengue Haemorrhagic Fever)
ตัวแบบที่ 9 จำนวนผู้ป่วยโรคปอดบวม (Pneumonia)

<u>ตัวแบบที่ 10</u> จำนวนผู้ป่วยโรควัณโรค (Tuberculosis, Pulmonary T.B)

วิธีการพยากรณ์อนุกรมเวลาโดยวิธีบอกซ์และเจนกินส์ เทคนิคปรับให้เรียบ การวิเคราะห์อนุกรมเวลาแบบคลาสสิค และวิธีการพยากรณ์ร่วม ในการวิจัยนี้มีขนาดของข้อมูลที่นำมาวิเคราะห์ด้วยวิธีการพยากรณ์ต่าง ๆ ข้างต้น จำนวน 192 คาบเวลา โดยแบ่งข้อมูลเป็น 2 ส่วน ในส่วนที่ 1 มีจำนวน 180 คาบเวลา (ข้อมูลในปี พ.ศ. 2526 - 2540) ซึ่งเป็นส่วนที่นำมาวิเคราะห์หาตัวแบบที่เหมาะสม ซึ่งจากการคำนวณเมื่อได้ตัวแบบที่ใช้พยากรณ์ข้อมูลก็จะทำการพยากรณ์ล่วงหน้า 12 คาบเวลา (พยากรณ์ล่วงหน้าในปี พ.ศ. 2541) จากนั้นนำข้อมูลที่พยากรณ์ได้นี้ไปเปรียบเทียบกับข้อมูลในส่วนที่เหลือไว้จำนวน 12 คาบเวลา (ข้อมูลในปี พ.ศ. 2541) จากนั้นนำมาตรวจสอบความแม่นยำในการพยากรณ์ โดยการหาค่าเฉลี่ยร้อยละความคลาดเคลื่อนจากการพยากรณ์ (MAPE)

3.4.1 วิธีบอกซ์และเจนกินส์

วิธีการพยากรณ์ของบอกซ์และเจนกินส์จะต้องตรวจสอบข้อมูลอนุกรมเวลาอยู่ในสภาพคงที่ (Stationary Data Series) ซึ่งสภาพคงที่ หมายถึงการคงที่ในค่าเฉลี่ย ความแปรปรวน และความแปรปรวนร่วม ซึ่งการทำให้ข้อมูลอนุกรมเวลาคงที่ทำได้โดยการแปลงข้อมูล ซึ่งการแปลงข้อมูลอนุกรมเวลาให้คงที่ในค่าเฉลี่ยจะใช้วิธีการทำผลต่างโดยนำข้อมูลมาลบกันได้เป็นข้อมูลอนุกรมเวลาชุดใหม่ และในกรณีอนุกรมเวลาไม่คงที่ในความแปรปรวนจะแปลงข้อมูลอนุกรมเวลาโดยการใส่ ln ในอนุกรมเวลา Y_t สำหรับอนุกรมเวลาที่องค์ประกอบฤดูกาลโดยมีค่าเวลาของฤดูกาลก็ทำการหาผลต่างเช่นเดียวกันเพื่อให้อนุกรมเวลาคงที่ก่อนที่จะนำมาวิเคราะห์ตัวแบบ

วิธีการของบอกซ์และเจนกินส์มีตัวแบบทั่วไป คือ ตัวแบบ ARIMA (Autoregressive Integrated Moving Average Models : ARIMA)

<u>ขั้นตอนวิธีการสร้างตัวแบบ ARIMA</u>

1. นำข้อมูลอนุกรมเวลาเหล่านี้มาตรวจสอบดูว่าข้อมูลอนุกรมเวลามีลักษณะอยู่ในสภาพคงที่ในค่าเฉลี่ย และความแปรปรวนหรือไม่ โดยดูว่าค่า ACF มีลักษณะลดลงอย่างรวดเร็วหรือถูกตัดออกในช่วงเวลาใดหรือไม่ ถ้าค่า ACF ไม่มีลักษณะดังกล่าวแสดงว่าอนุกรมเวลาไม่อยู่ในสภาพคงที่จะต้องแปลงข้อมูลอนุกรมเวลานั้นให้เป็นอนุกรมเวลาชุดใหม่ให้อยู่ในสภาพคง

ที่ทั้งในค่าเฉลี่ย และความแปรปรวน โดยทำการแปลงข้อมูลให้คงที่ในค่าเฉลี่ยด้วยการทำผลต่างระหว่างข้อมูลที่อยู่ติดกัน ส่วนการแปลงข้อมูลให้คงที่ในความแปรปรวน ทำได้โดยการใส่ ***ln*** ให้กับข้อมูล สำหรับอนุกรมเวลาที่มีองค์ประกอบฤดูกาลไม่คงที่ในค่าเฉลี่ยก็จะทำการหาผลต่างเช่นเดียวกัน เพื่อให้อนุกรมเวลาอยู่ในสภาพคงที่ก่อนพิจารณากำหนดตัวแบบ

2. กำหนดรูปแบบเบื้องต้นของข้อมูลจำนวนผู้ป่วย โดยการนำข้อมูลจำนวนผู้ป่วยที่อยู่ในสภาพคงที่แล้ว มากำหนดรูปแบบเบื้องต้นโดยเปรียบเทียบกัน ดูว่าใกล้เคียงกับรูปแบบใดของตัวแบบ ARIMA

3. ประมาณค่าพารามิเตอร์ เมื่อกำหนดรูปแบบเบื้องต้น และคำนวณค่าพารามิเตอร์ต่าง ๆ ของรูปแบบที่กำหนด

4. ตรวจสอบความเหมาะสมของรูปแบบ ดูว่าเป็นรูปแบบที่เหมาะสมกับอนุกรมเวลาชุดนั้น ๆ จริงหรือไม่ โดยพิจารณาความคลาดเคลื่อน (e_t) ; t=1,2,3,...,T ที่คำนวณได้ แทนค่าอนุกรมเวลา และค่าพารามิเตอร์ในรูปแบบนั้นแล้ว ทดสอบจนได้เทอมความคลาดเคลื่อนที่เป็นอิสระกัน รูปแบบนั้นจึงจะเป็นรูปแบบที่เหมาะสม ถ้าไม่เป็นไปตามข้อตกลงดังกล่าว ให้กลับไปในขั้นตอนที่ 1 อีกครั้ง

5. นำตัวแบบพยากรณ์จำนวนผู้ป่วยในอนาคต

3.4.2 เทคนิคการปรับให้เรียบ

วิธีการปรับให้เรียบสำหรับการพยากรณ์ข้อมูลอนุกรมเวลามีหลายวิธีด้วยกัน สำหรับการวิจัยในครั้งนี้ เทคนิคที่นำมาใช้คือ วิธีการพยากรณ์ปรับให้เรียบแบบวินเตอร์ เนื่องจากข้อมูลโดยส่วนใหญ่แล้วเป็นข้อมูลแบบอนุกรมเวลาที่องค์ประกอบฤดูกาล ส่วนวิธีการทำให้เรียบแบบเลขชี้กำลังอย่างง่าย วิธีการแบบเลขชี้กำลังสองครั้ง วิธีการทำให้เรียบแบบโฮลท์เป็นวิธีรองลงมาที่นำมาใช้กับข้อมูลที่ไม่มีองค์ประกอบในส่วนของฤดูกาล ซึ่งวิธีการใดจะเหมาะสมมากน้อยเพียงใด ผู้วิจัยจะต้องนำมาพิจารณาจากค่าพยากรณ์ที่ได้จากวิธีนั้น ๆ

ในการวิจัยครั้งนี้เนื่องจากข้อมูลอนุกรมเวลาทุกโรคมีการเคลื่อนไหวแบบแนวโน้มและมีองค์ประกอบฤดูกาล ฉะนั้นควรจะใช้วิธีการพยากรณ์ที่พิจารณาองค์ประกอบฤดูกาลร่วมด้วย ซึ่งเหมาะกับวิธีการพยากรณ์ของวินเตอร์ โดยจะดำเนินการหาค่าคงที่โดยใช้โปรแกรมสำเร็จรูปคำนวณเพื่อให้ได้ค่าพยากรณ์ที่มีความคลาดเคลื่อนต่ำที่สุด จากนั้นทำการวินิจฉัยความเพียงพอของตัวแบบ ดังนั้นการจะเลือกใช้วิธีใดจึงขึ้นอยู่กับลักษณะของข้อมูลอนุกรมเวลาในแต่ละชุดนั้น ซึ่งแต่ละวิธีจะเหมาะสมเพียงใดจะพิจารณาจากความคลาดเคลื่อนจากการพยากรณ์

3.4.3 การวิเคราะห์อนุกรมเวลาแบบคลาสสิค

ในการวิเคราะห์อนุกรมเวลาแบบคลาสสิคนี้เป็นการประยุกต์ใช้การวิเคราะห์การถดถอยกับการวิเคราะห์อนุกรมเวลาเข้าด้วยกัน โดยในการวิเคราะห์ได้ทำการแยกอนุกรมเวลาออกเป็นส่วนประกอบต่าง ๆ โดยมีตัวแบบพื้นฐานอยู่ 2 ตัวแบบ คือ ตัวแบบเชิงคูณ และตัวแบบเชิงบวก ซึ่งในการวิเคราะห์อนุกรมเวลาแบบคลาสสิคนี้ ได้นำวิธีการวิเคราะห์ถดถอยมาประยุกต์ใช้ โดยใช้วิธี ENTER ในการคัดเลือกสมการพยากรณ์ ซึ่งการประยุกต์ใช้การวิเคราะห์การถดถอยกับการวิเคราะห์อนุกรมเวลานั้นจะกำหนดตัวแปรบ่งชี้แนวโน้ม (Trend Indicator) และตัวแปรบ่งชี้ฤดูกาล (Seasonal Indicator) ซึ่งในงานวิจัยนี้มีข้อมูลเป็นรายเดือน จึงกำหนดตัวแปรบ่งชี้ฤดูกาลจำนวน 11 ตัว โดยมีขั้นตอนดังนี้

<u>ขั้นตอนวิธีการสร้างตัวแบบโดยการวิเคราะห์อนุกรมเวลาแบบคลาสสิค</u>

1. พิจารณาการเคลื่อนไหวของอนุกรมเวลาว่ามีรูปแบบการรวมตัวเป็นรูปแบบคูณหรือไม่ ถ้าเป็นรูปแบบคูณต้องทำการแปลงข้อมูลอนุกรมเวลาเหล่านั้นให้มีลักษณะการรวมตัวเป็นรูปแบบบวกก่อน โดยการใส่ ln ให้กับอนุกรมเวลาชุดใหม่

2. ทำการประมาณค่าพารามิเตอร์โดยวิธีการวิเคราะห์การถดถอย โดยใช้วิธี ENTER ซึ่งจะพิจารณาตัวแปรบ่งชี้แนวโน้มและตัวแปรบ่งชี้ฤดูกาลทุกตัว เนื่องจากถ้าตัวแปรบ่งชี้ฤดูกาลใด ๆ ไม่ถูกนำเข้ามาในสมการแล้ว ฤดูกาลนั้นก็จะไม่สามารถหาค่าพยากรณ์ได้

3. ทำการวินิจฉัยตัวแบบ โดยพิจารณาจากฟังก์ชันก์อัตตสหสัมพันธ์ และฟังก์ชันอัตตสหสัมพันธ์ย่อยของเศษเหลือตกค้างว่าตัวแบบพยากรณ์ที่ได้นั้นมีความเหมาะสมเพียงพอหรือไม่ ถ้าลักษณะกราฟของเศษเหลือตกค้างมีลักษณะเข้ากับรูปแบบของ AR(1) จะต้องทำการประมาณค่าพารามิเตอร์ใหม่ โดยใช้การวิเคราะห์อนุกรมเวลาโดยวิธี Autoregression โดยวิธีดังกล่าวนี้รวมเอาเหอมของเศษเหลือตกค้างที่มีรูปแบบ AR(1) เข้ากับตัวแบบเป็นตัวแบบใหม่ หลังจากนั้นวินิจฉัยตัวแบบใหม่ โดยพิจารณาฟังก์ชันก์อัตตสหสัมพันธ์ และฟังก์ชันอัตตสหสัมพันธ์ย่อยของเศษเหลือตกค้างของตัวแบบใหม่อีกครั้ง ถ้าพบว่าตัวแบบที่กำหนดยังไม่เหมาะสมอีก จึงหยุดทำการวินิจฉัย

3.4.4 วิธีการพยากรณ์ร่วม

การนำเทคนิคการพยากรณ์ไปใช้ ผู้ใช้ควรเลือกวิธีการพยากรณ์ให้เหมาะสมกับลักษณะของข้อมูล โดยเลือกใช้วิธีการพยากรณ์ที่ทำให้ได้ค่าพยากรณ์ที่ใกล้เคียงกับค่าจริง หรือมีความคลาดเคลื่อนพยากรณ์ต่ำสุด เพื่อให้ได้ค่าพยากรณ์ที่ถูกต้องและเชื่อถือได้สูง ดังนั้นการหาค่าพยากรณ์ร่วม(Combined Forecast) จึงเป็นอีกทางเลือกหนึ่งที่นำมาศึกษาในหารวิจัยครั้งนี้ โดยการรวมวิธีการพยากรณ์เข้าด้วยกันโดยการให้น้ำหนักแต่ละวิธีการพยากรณ์ ซึ่งมีวิธีการให้น้ำหนักดังต่อไปนี้ วิธีการให้น้ำหนักที่เท่ากัน(Simple Average Method) และวิธีการหาค่าพยากรณ์ร่วมด้วยวิธีค่าสัมบูรณ์ต่ำสุด(Least Absolute Value Method)

<u>ขั้นตอนวิธีการสร้างตัวแบบโดยวิธีการพยากรณ์ร่วม</u>

1. หารูปแบบสมการพยากรณ์ด้วยวิธีการพยากรณ์ต่าง ๆ ดังต่อไปนี้ คือ
   1.1 การวิเคราะห์อนุกรมเวลาแบบคลาสสิค
   1.2 วิธีการของวินเตอร์ (Winter's Method)
   1.3 วิธีการบอกซ์และเจนกินส์ (Box – Jenkins Method)

2. หาค่าเฉลี่ยถ่วงน้ำหนักของแต่ละวิธีการ

   2.1 วิธีการให้น้ำหนักของแต่ละวิธีเท่ากัน (Simple Average Method) มีสูตรการคำนวณดังนี้

$$W_j = \frac{1}{m}$$

โดยที่ W_j คือ ค่าเฉลี่ยน้ำหนักของวิธีการพยากรณ์ที่ j

m คือ จำนวนวิธีการพยากรณ์ที่นำมารวมกัน

   2.2 วิธีค่าสัมบูรณ์ต่ำสุด (Least Absolute Value Method)

เป็นวิธีการคำนวณหาค่าเฉลี่ยถ่วงน้ำหนัก โดยอาศัยเทคนิคโปรแกรมเชิงเส้น ภายใต้เงื่อนไขที่จะทำให้ผลรวมของค่าสัมบูรณ์ของความคลาดเคลื่อนต่ำสุด

<u>ขั้นตอนที่ 1</u> จากข้อมูลอนุกรมเวลา (Y_t) และค่าพยากรณ์จากสมการพยากรณ์ที่ได้จากวิธีการพยากรณ์ต่าง ๆ ในข้อ 1 $(\hat{Y}_{jt}$ นำมาแปลงให้อยู่ในรูปสมการเชิงเส้น

<u>ขั้นตอนที่ 2</u> *คำนวณหาค่าถ่วงน้ำหนักด้วยวิธีค่าสัมบูรณ์ต่ำสุด* โดยนำสมการเชิงเส้นที่ได้ในขั้นตอนที่ 1 มาแก้สมการหาค่าถ่วงน้ำหนัก โดยใช้โปรแกรมสำเร็จรูป QSB$^+$ ช่วยในการคำนวณ

3. นำค่าถ่วงน้ำหนักที่ได้ในข้อ 2.1 และ 2.2 มาหาค่าพยากรณ์ในอนาคต

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย