การพยากรณ์จำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่ด้วยวิธีบ็อกซ์-เจนกินส์

ธันวา เจริญศิริ (วท.บ.)[1], ชนาธิป โสภณพิมล (วท.บ.)[1], ดนุสรณ์ ธนะปาละ (วท.บ.)[1],
วีระชัย ขันทองคำ (วท.บ.)[1]

[1]ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ชลบุรี

บทคัดย่อ

วัตถุประสงค์ เพื่อกำหนดตัวแบบที่เหมาะที่สุดสำหรับการพยากรณ์จำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่
วิธีการศึกษา เป็นการศึกษาระบาดวิทยา โดยข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลทุติยภูมิซึ่งเก็บรวบรวมจากรายงานของสำนักระบาดวิทยากรมควบคุมโรค กระทรวงสาธารณสุข กรุงเทพมหานคร ซึ่งบันทึกข้อมูลเป็นรายเดือนตั้งแต่เดือนพฤษภาคม ค.ศ. 2009 ถึงเดือนสิงหาคม ค.ศ. 2015 และวิเคราะห์ข้อมูลด้วยวิธีบ็อกซ์-เจนกินส์ เกณฑ์ที่ใช้ในการวัดความแม่นยำของการพยากรณ์ คือ ค่ารากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย ค่าร้อยละของค่าสัมบูรณ์ของค่าคลาดเคลื่อนเฉลี่ยและค่าสัมบูรณ์ของค่าคลาดเคลื่อนเฉลี่ย
ผลการศึกษา ตัวแบบที่เหมาะสำหรับการพยากรณ์จำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่คือตัวแบบ ARIMA$(0,1,0)(1,1,1)_{12}$
สรุปผล ตัวแบบวิธีบ็อกซ์-เจนกินส์นี้สามารถนำไปใช้พยากรณ์จำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่ นอกจากนี้ยังสามารถนำไปประกอบการวางแผนเพื่อหาแนวทางป้องกันและควบคุมโรคที่จะเกิดขึ้นได้ในอนาคต

คำสำคัญ : การพยากรณ์ โรคสุกใส วิธีบ็อกซ์-เจนกินส์

ผู้นิพนธ์ที่รับผิดชอบ: ธันวา เจริญศิริ
ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ชลบุรี ประเทศไทย
E-mail: charoensiri_t@hotmail.com

Forecasting The Number of Chicken pox in Chiang Mai with Box-Jenkins Method

Thanva Charoensiri (B.Sc.)[1], Chanatip Soponpimol (B.Sc.)[1], Danusorn Thanapala (B.Sc.)[1], Weerachai Khanthongkhamv (B.Sc.)[1]

[1]Department of Mathematics, Faculty of Science, Burapha University, Chonburi

Abstract

Objective this reseach was objected to determine the optimal model for forecasting the number of Chicken pox in Chiang Mai.

Methods Data to be used in this study was secondary data which monthly collected from the Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health, Bangkok. since May, 2009 to August, 2015 and analyzed with Box-Jenkins method. Criteria to be measured the accuracy of forecasting was the root mean square error (RMSE), the mean absolute percentage error (MAPE) and the mean absolute error (MAE).

Results The result of study showed that the optimal model for forecasting the number of Chicken pox in Chiang Mai was the ARIMA$(0,1,0)(1,1,1)_{12}$.

Conclusions The Box-Jenkins model capably forecasted the number of Chicken pox in Chiang Mai also assembly designed to protect and control the diseases in the future.

Keywords : Forecasting, Chickenpox, Box-Jenkins method

Corresponding author : Thanva Charoensiri**.** Department of Mathematics, Faculty of Science, Burapha University, Chonburi, Thailand.
E-mail: charoensiri_t@hotmail.com

บทนำ

การพยากรณ์โรคและภัยสุขภาพเป็นมาตรการหนึ่งที่ถูกบรรจุไว้ในแผนยุทธศาสตร์ของกรมควบคุมโรค เพื่อคุ้มครองสุขภาพของประชาชนในลักษณะของการคาดการณ์หรือการทำนายล่วงหน้าเกี่ยวกับโรคและภัยสุขภาพต่าง ๆ ที่จะเกิดขึ้นในรอบปี ซึ่งมีเป้าหมายหลักเพื่อให้ทุกหน่วยงานที่เกี่ยวข้องทั้งส่วนกลางและภูมิภาคสามารถพยากรณ์โรคและภัยสุขภาพล่วงหน้าได้ โดยเลือกใช้เทคนิคการพยากรณ์สำหรับโรคที่เกิดขึ้นตามฤดูกาล โรคระบาดที่เกิดขึ้นได้ตลอดทั้งปี ซึ่งจะสามารถพยากรณ์จำนวนผู้ป่วยที่จะเกิดขึ้นมากน้อยเท่าไหร่ จะทำให้สามารถวางแผนป้องกันและเตรียมรับมือกับโรคที่จะเกิดขึ้นได้เพื่อลดจำนวนผู้ป่วยที่จะเกิดขึ้นได้การพยากรณ์จำนวนผู้ป่วยที่จะเกิดในช่วงเวลาต่าง ๆ เป็นเครื่องมือที่ช่วยพยากรณ์จำนวนผู้ป่วย ในอนาคตได้เนื่องจากการพยากรณ์นั้นมีประโยชน์และความสำคัญต่อการวางแผนดำเนินงานเกี่ยวกับโรคต่าง ๆ มาก ตัวแบบพยากรณ์ที่ใช้เทคนิคการวิเคราะห์อนุกรมเวลามีหลายวิธีถูกนำมาใช้ในการพยากรณ์จำนวนผู้ป่วย ซึ่งการพิจารณาประสิทธิภาพของแต่ละวิธีจะพิจารณาจากค่าคลาดเคลื่อนน้อยที่สุดของการพยากรณ์ และหนึ่งในวิธีที่ผู้วิจัยส่วนใหญเลือกใช้คือ วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins) ซึ่งเป็นวิธีที่ใช้ง่าย ง่ายต่อการเข้าใจ

จากการวิเคราะห์สถานการณ์โรคสุกใสระหว่างปี พ.ศ.2552 ถึงปี พ.ศ.2558 แนวโน้มการเกิดโรคสุกใสจะเป็นโรคที่มีแนวโน้มเกิดเพิ่มขึ้นอย่างต่อเนื่องเป็น 1.3 เท่า ภาคที่มีอัตราการเป็นโรคสุกใสสูงสุดคือ ภาคเหนือ รองลงมาคือภาคกลางภาคใต้ และภาคตะวันออกเฉียงเหนือ การถ่ายทอดโรคสุกใสนั้นจะเกิดในสถานที่ที่มีคนอยู่รวมกันเป็นจำนวนมาก ๆ ดังนั้นวิธีป้องกันการถ่ายทอดโรคจึงต้องพยายามแยกผู้ป่วยออกจากผู้อื่นโดยเร็ว เนื่องจากไม่มีระบบการให้วัคซีนป้องกันโรคสุกใสเพื่อป้องกันการแพร่เชื้อซึ่งการควบคุมโรคจะได้ผลดีหรือไม่นั้นจะต้องอาศัยระบบเฝ้าระวังโรคที่มีคุณภาพโดยจะต้องมีการรายงานหรือระบบการแจ้งข่าวที่รวดเร็ว มิฉะนั้นแล้วจะไม่สามารถยับยั้งการแพร่ระบาดของโรคได้ ถึงแม้โรคสุกใสจะไม่ก่อให้เกิดความรุนแรงในคนทั่ว ๆ ไป แต่ในคนบางกลุ่มถือว่าเป็นกลุ่มเสี่ยงต่อการทำให้เกิดภาวะแทรกซ้อนซึ่งอาจจะเป็นอันตรายถึงแก่ชีวิตได้ กลุ่มเสี่ยงดังกล่าวได้แก่ ผู้ที่มีภูมิคุ้มกันบกพร่องแต่กำเนิด ผู้ที่กินยากดภูมิต้านทาน ทารกแรกเกิด สตรีมีครรภ์ ผู้สูงอายุที่ป่วยเป็นโรคเบาหวาน ผู้ป่วยโรคมะเร็ง ผู้ป่วยที่อยู่ระหว่างการรักษาด้วยเคมีบำบัดหรือรังสีรักษา ผู้ป่วยโรคภูมิคุ้มกันบกพร่อง กลุ่มเสี่ยงเหล่านี้ควรจะได้รับการดูแลเป็นพิเศษทั้งก่อนการติดเชื้อหรือเมื่อติดเชื้อไปแล้ว (สำนักระบาดวิทยา, 2558)

ดังนั้น เพื่อให้ได้วิธีการพยากรณ์ที่เหมาะและง่ายต่อการนำไปใช้สำหรับการพยากรณ์จำนวนผู้ป่วยที่เป็นโรคสุกใสในจังหวัดเชียงใหม่ เนื่องจากเป็นหัวเมืองใหญ่ของประเทศไทยและมีอัตราการเป็นโรคสุกใสสูงสุดจึงนำไปสู่ความสนใจของผู้วิจัยที่จะทำการหาตัวแบบสมการที่เหมาะที่สุดด้วยวิธีบ็อกซ์-เจนกินส์ เพื่อให้ได้ตัวแบบสมการที่เหมาะที่มีความแม่นยำ ซึ่งปัจจุบันยังไม่เคยมีการพยากรณ์จำนวนผู้ป่วยด้วยวิธีดังกล่าว เพื่อให้ทราบถึงจำนวนผู้ป่วยโรคสุกใสที่จะเกิดขึ้นในอนาคต โดยสามารถนำข้อมูลที่ได้ไปใช้ประกอบการวางแผนเพื่อหาแนวทางป้องกันและควบคุมโรคที่จะเกิดขึ้น

วัตถุประสงค์การวิจัย

เพื่อหาตัวแบบพยากรณ์ที่เหมาะสำหรับการพยากรณ์จำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่ด้วยวิธีบ็อกซ์-เจนกินส์

วิธีการศึกษา

ลำดับการศึกษามีขั้นตอนดังนี้

การจัดเตรียมข้อมูล

การวิจัยนี้มุ่งเน้นเพื่อทำศึกษาวิธีการพยากรณ์อนุกรมเวลาด้วยวิธีบ็อกซ์-เจนกินส์กับข้อมูลจำนวนผู้ป่วยที่เป็นโรคสุกใสในจังหวัดเชียงใหม่ ข้อมูลที่นำ

มาทำการวิเคราะห์เป็นข้อมูลจำนวนผู้ป่วยย้อนหลังจากรายงานของสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข กรุงเทพมหานคร ลักษณะข้อมูลจำแนกเป็นรายเดือนตั้งแต่เดือนพฤษภาคม พ.ศ.2552 ถึงเดือนสิงหาคม พ.ศ.2558 และทำการพยากรณ์จำนวนผู้ป่วยในอนาคตอีก 3 เดือนคือ เดือนกันยายน ถึงเดือนพฤศจิกายน เทียบกับจำนวนผู้ป่วยจริง

การศึกษาความเคลื่อนไหวของอนุกรมเวลา

เพื่อพิจารณาเบื้องต้นว่าอนุกรมนั้นมีลักษณะแบบใด ทำการทดสอบองค์ประกอบแนวโน้มและความแปรผันตามฤดูกาลโดยพิจารณาจากกราฟอัตสหสัมพันธ์ (Auto Correlation Function: ACF) และกราฟอัตสหสัมพันธ์บางส่วน (Partial Auto Correlation Function: PACF)

การวิเคราะห์ข้อมูลทางสถิติ

การวิเคราะห์ในงานวิจัยนี้ทำโดยการพยากรณ์อนุกรมเวลาด้วยการสร้างตัวแบบพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์ วิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูป MINITAB 14

วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins)

ตัวแบบบ็อกซ์-เจนกินส์เป็นตัวแบบที่อธิบายอนุกรมเวลาเป็นฟังก์ชันเชิงเส้นของค่าสังเกตในอดีตและการรบกวนทั้งในอดีตและปัจจุบัน ซึ่งการพยากรณ์โดยใช้ตัวแบบ ARIMA เป็นวิธีการสร้างสมการพยากรณ์อนุกรมเวลาโดยการกำหนด ARIMA ซึ่งเป็นรูปแบบที่อธิบายลักษณะการเคลื่อนไหวของอนุกรมเวลาที่ค่าสังเกตมีสหสัมพันธ์กันแบบออโตและสเตชันนารี ซึ่งสามารถใช้ได้กับข้อมูลที่มีลักษณะเคลื่อนไหวทุกประเภทและให้ผลการพยากรณ์ที่ค่อนข้างแม่นยำกว่าวิธีการพยากรณ์แบบอื่น ๆ สำหรับขั้นตอนการวิเคราะห์อนุกรมเวลาของวิธีบ็อกซ์-เจนกินส์ (คณินทร์ ธีรภาพโอฬาร, 2558) มีดังนี้

ขั้นตอนที่ 1 ตรวจสอบข้อมูลเพื่อพิจารณาว่าอนุกรมเวลาเป็นสเตชันนารีและพิจารณาอนุกรมเวลาว่าได้รับผลกระทบจากอิทธิพลของฤดูกาลหรือไม่ โดยจะพิจารณาจากกราฟฟังก์ชันอัตสหสัมพันธ์ (Auto Correlation Function: ACF) และกราฟฟังก์ชันอัตสหสัมพันธ์บางส่วน (Partial Auto Correlation Function: PACF) สำหรับอนุกรมเวลาที่ได้รับผลกระทบจากอิทธิพลฤดูกาล อนุกรมเวลาจะมีการเปลี่ยนแปลงขึ้นลงตามช่วงเวลาต่าง ๆ ที่สอดคล้องกับการเปลี่ยนแปลงที่เกิดขึ้นในอดีต

ขั้นตอนที่ 2 การกำหนดตัวแบบของการพยากรณ์ (Model Identification) ซึ่งเป็นขั้นตอนการพิจารณาว่าตัวแบบใดเหมาะที่สุดกับข้อมูลของข้อมูลที่นำมาวิเคราะห์ ซึ่งสามารถพิจารณาจากค่า ACF และ PACF ซึ่งจะได้ตัวแบบคือ ARIMA (p,d,q)(P,D,Q)L

ขั้นตอนที่ 3 การประมาณค่าพารามิเตอร์ (Parameter Estimation) ของรูปแบบที่กำหนดด้วยวิธีความควรจะเป็นสูงสุด (Maximum Likelihood)

ขั้นตอนที่ 4 การตรวจสอบความเหมาะสมของตัวแบบ (Diagnostic Checking) จะมีการตรวจสอบความเหมาะสมว่าสอดคล้องกับข้อสมมุติเบื้องต้นหรือไม่ ซึ่งจะมีการตรวจสอบข้อสมมุติดังนี้

1. ตรวจสอบการแจกแจงปรกติ
2.ตรวจสอบค่าคลาดเคลื่อนเป็นอิสระ
3. ตรวจสอบค่าคลาดเคลื่อนมีความแปรปรวนคงที่
4. ตรวจสอบผลรวมค่าคลาดเคลื่อน

แต่ถ้าหากตัวแบบที่กำหนดไม่มีเหมาะสมจะต้องมีการกำหนดตัวแบบในขั้นตอนที่ 2 ใหม่

ขั้นตอนที่ 5 การพยากรณ์ (Forecasting) เมื่อตัวแบบอนุกรมที่เลือกผ่านการตรวจสอบข้อสมมุติสำหรับความเหมาะสมของตัวแบบเรียบร้อยแล้ว ซึ่งจะได้สมการพยากรณ์ค่าประมาณพารามิเตอร์สำหรับตัวแบบทั่วไปของการพยากรณ์นั่นคือ Autoregressive

Integressive Moving Average :ARIMA (p, d, q)(P, D, Q) ซึ่งสามารถแสดงได้ดังสมการที่ 1 (Bowerman and O'Connell, 2005)

$$\Phi_P\left(B^s\right)\phi_p\left(B\right)\left(1-B\right)^d\left(1-B^s\right)^D X_t = \theta_q\left(B\right)\Theta_Q\left(B^s\right)u_t \quad (1)$$

โดยที่ X_t แทนอนุกรมเวลา ณ เวลา t

u_t แทนอนุกรมเวลาของค่าคลาดเคลื่อนที่มีการแจกแจงปรกติและเป็นอิสระกันด้วยค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_u^2

$\delta = \mu\phi_p\left(B\right)\Phi_P\left(B^S\right)$ แทนค่าคงตัว โดยที่ μ แทนค่าเฉลี่ยของอนุกรมเวลาที่มีลักษณะคงที่ (Stationary)

$\phi_p\left(B\right) = 1-\phi_1 B-\phi_2 B^2 \ldots -\phi_p B^p$ แทนตัวดำเนินการดำเนินการสหสัมพันธ์ในตัวเองแบบไม่มีฤดูกาลอันดับที่ p (Non-Seasonal Autoregressive Operator of Order p: AR(p))

$\theta_q\left(B\right) = 1-\theta_1 B-\theta_2 B^2 \ldots -\theta_q B^q$ แทนตัวดำเนินการสหสัมพันธ์ในตัวเองแบบมีฤดูกาลอันดับที่ P (Seasonal Autoregressive Operator of Order p: SAR(P))Ø

$\Phi_P\left(B^s\right) = 1-\Phi_1 B-\Phi_2 B^{2s} \ldots -\Phi_P B^{PS}$ แทนตัวดำเนินการเฉลี่ยเคลื่อนที่แบบไม่มีฤดูกาลอันดับที่q (Non-Seasonal Moving Average Operator of Order p: MA(q))

$\Theta_Q\left(B^s\right) = 1-\Theta_1 B-\Theta_2 B^{2s} \ldots -\Theta_Q B^{QS}$ แทนตัวดำเนินการเฉลี่ยเคลื่อนที่แบบมีฤดูกาลอันดับที่ Q (Non-Seasonal Moving Average Operator of Order p: SMA(Q))

t แทนคาบเวลา ซึ่งมีค่าตั้งแต่ 1 ถึง n_1 แทนจำนวนข้อมูลในอนุกรมเวลาชุดที่ 1

s แทนจำนวนคาบของฤดูกาล

d และ D แทนผลต่างภายในและระหว่างฤดูกาล ตามลำดับ

B แทนตัวดำเนินการถอยหลัง (Backward Operator) โดยที่ $B^s X_t = X_{t-s}$

เกณฑ์การวัดความแม่นยำของการพยากรณ์

การพิจารณาตัวแบบพยากรณ์ที่เหมาะที่สุดสำหรับการพยากรณ์จำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่ โดยทั่วไปจะใช้เกณฑ์วัดความแม่นยำของการพยากรณ์ 3 เกณฑ์คือค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) ร้อยละของค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ย (MAPE) และค่าสัมบูรณ์ของค่าคลาดเคลื่อนเฉลี่ย (MAE) ดังนี้ (คณินทร์ ธีรภาพโอฬาร, 2558) (มุกดา แม้นมินทร์, 2549)

1. ค่ารากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) คำนวนได้ดังนี้

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n} e_t^2}$$

2. ร้อยละของค่าสัมบูรณ์ของค่าคลาดเคลื่อนเฉลี่ย (Mean Absolute Percentage Error: MAPE) เป็นการวัดความคลาดเคลื่อนระหว่างค่าจริงและค่าพยากรณ์ค่า MAPE ไม่มีหน่วยและมีค่าเป็นบวกเสมอ คำนวนได้ดังนี้

$$MAPE = \frac{1}{N}\sum_{i=1}^{n}\left|\frac{e_t}{X_t}\right|100\%$$

3. ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) คำนวนได้ดังนี้

$$MAE = \frac{1}{N}\sum_{i=1}^{n}\left|e_t\right|$$

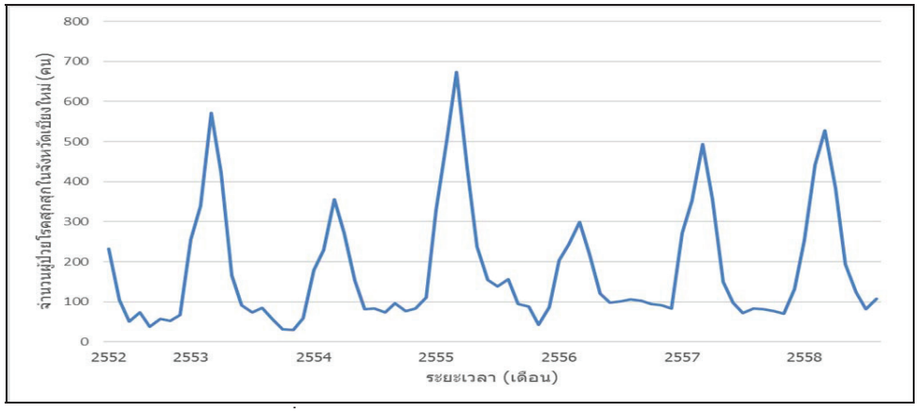
โดยที่ e_t เป็นผลต่างระหว่างค่าจริงกับค่าพยากรณ์ ณ เวลา t

X_t เป็นค่าจริงของข้อมูลอนุกรมเวลา ณ เวลา t

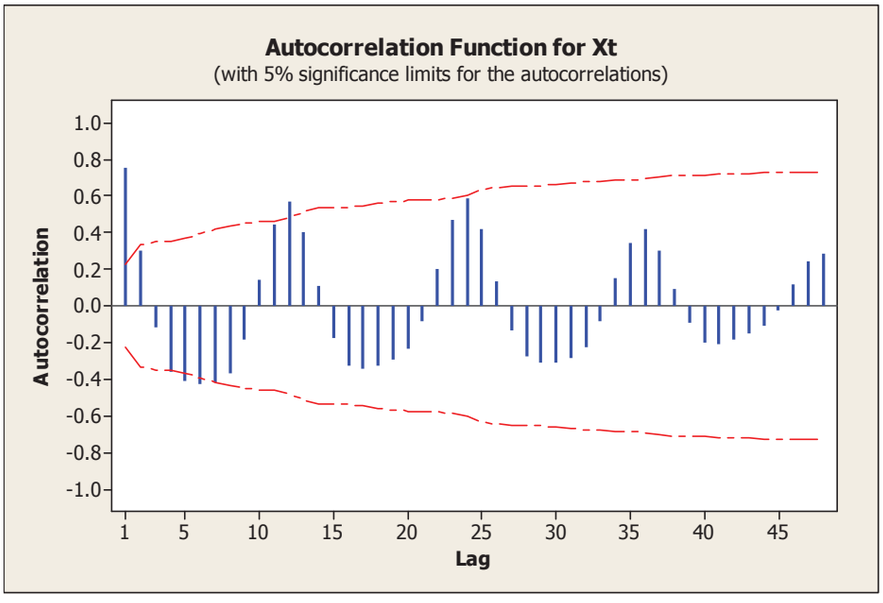
N เป็นจำนวนค่าสังเกตช่วงเวลาทั้งหมด

ผลการวิจัย

ข้อมูลจำนวนผู้ป่วยที่เป็นโรคสุกใสในจังหวัดเชียงใหม่ จากรายงานของสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข กรุงเทพมหานคร นำมาวิเคราะห์กราฟอนุกรมเวลาของจำนวนผู้ป่วยโรคสุกใส ในจังหวัดเชียงใหม่จะพบว่าข้อมูลจำนวนผู้ป่วยได้รับผลกระทบโดยตรงจากอิทธิพลของฤดูกาล ซึ่งข้อมูลมีการเปลี่ยนแปลงขึ้นลงตามช่วงเวลาต่างๆจึงสรุปได้ว่าข้อมูลจำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่มีรูปแบบฤดูกาล แสดงดังภาพ ที่ 1



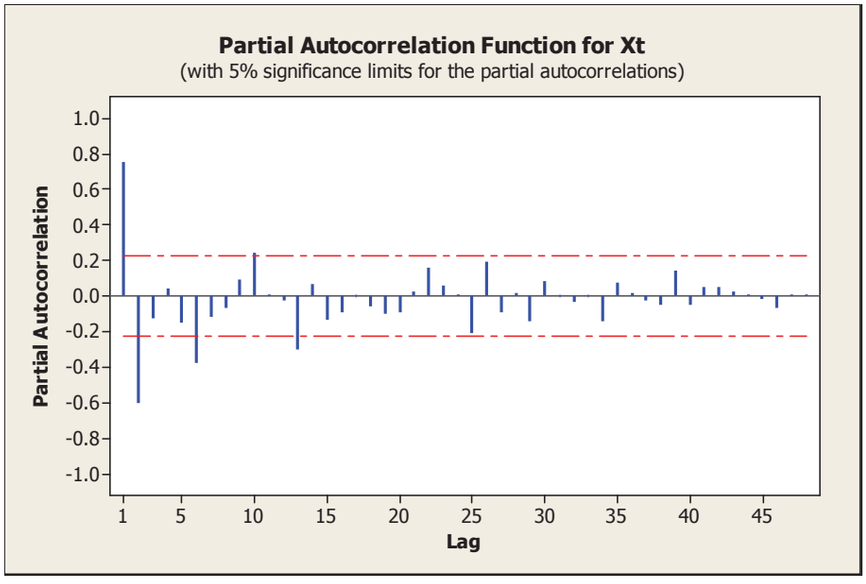
ภาพที่ 1 กราฟจำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่



ภาพที่ 2 กราฟฟังก์ชันอัตสหสัมพันธ์

สามารถยืนยันผลการได้รับผลกระทบโดยตรงจากอิทธิพลของฤดูกาลด้วยวิธีการวิเคราะห์อนุกรมเวลาด้วยวิธีบ็อกซ์-เจนกินส์ (Box-Jenkins) โดยพิจารณาจากกราฟฟังก์ชันอัตสหสัมพันธ์และกราฟฟังก์ชันอัตสหสัมพันธ์บางส่วน แสดงดังภาพที่ 2 และภาพที่ 3

จากกราฟฟังก์ชันอัตสหสัมพันธ์และกราฟฟังก์ชันอัตสหสัมพันธ์บางส่วน จะพบว่าข้อมูลอนุกรมเวลามีองค์ประกอบแนวโน้มและฤดูกาลมีระยะห่างเวลาตามความยาวของฤดูกาลกันคือ



ภาพที่ 3 กราฟฟังก์ชันอัตสหสัมพันธ์บางส่วน

12, 24, 36, ... หน่วย จึงทำการกำหนดตัวแบบพยากรณ์ที่เหมาะที่สุดสำหรับข้อมูลชุดนี้คือ ARIMA(0,1,0)$(1,1,1)_{12}$ ไม่มีพจน์ของค่าตัว จะได้ตัวแบบคือ

$$X_t = X_{t-12} + X_{t-1} - X_{t-13} + \phi_1 X_{t-12} - \phi_1 X_{t-24} - \phi_1 X_{t-13} + \phi_1 X_{t-25} + u_t - \Theta_1 u_{t-12}$$

และประมาณค่าพารามิเตอร์จะได้คือ และดังตารางที่ 1 ซึ่งได้ตัวแบบการพยากรณ์จำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่คือ

$$X_{t-12} + X_{t-1} - X_{t-13} - 0.6778X_{t-12} + 0.6778X_{t-24} + 0.6778X_{t-13} - 0.6778X_{t-25} - 0.8553u_{t-12}$$

ซึ่งจากการเลือกตัวแบบแล้วจึงทำการตรวจสอบความเหมาะสมของตัวแบบ พบว่าค่าคลาดเคลื่อนมีการแจกแจงปกติ (AD=0.17, p-value=0.92) มีค่าเฉลี่ยมีค่าเท่ากับศูนย์ (T=0.15, p-value=0.88) มีความแปรปรวนคงที่ (T=-0.39, p-value=0.70) และเป็นอิสระต่อกันหรือไม่มีอัตสหสัมพันธ์ (lag12, lag 24, lag36 และlag 48 ไม่มีนัยสำคัญในทุก ๆ lag)

ดังนั้น ตัวแบบของอนุกรมเวลาของจำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่ เมื่อทำการตรวจสอบความเหมาะสมของตัวแบบแล้วผ่านทุกข้อสมมุตินั้นคือ ARIMA(0,1,0)x$(1,1,1)_{12}$

ตารางที่ 1 แสดงค่าประมาณของพารามิเตอร์สำหรับตัวแบบ ARIMA(0,1,0)(1,1,1)$_{12}$

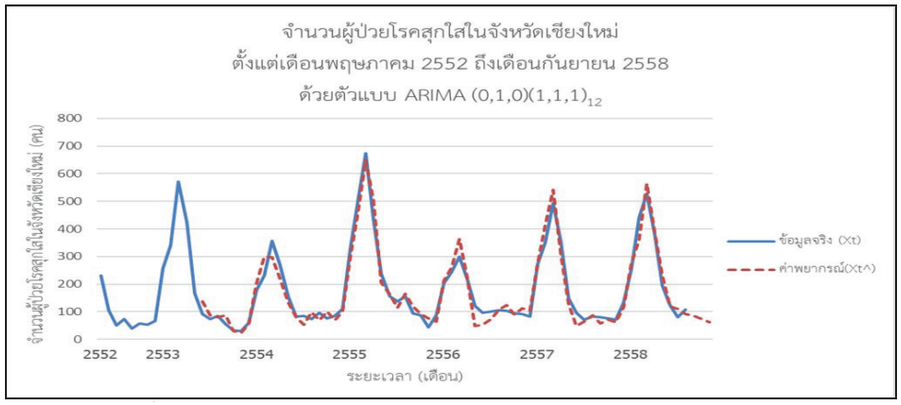
พารามิเตอร์	สัมประสิทธิ์	ส่วนเบี่ยงเบนของสัมประสิทธิ์	สถิติทดสอบ T	P-value
Φ_1	-0.6778	0.1075	-0.630	< 0.001
Θ_1	0.8553	0.1115	7.67	< 0.001

ผลการพยากรณ์จำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่

จากจำนวนผู้ป่วยโรคสุกใสพบว่า จังหวัดเชียงใหม่มีอัตราผู้ป่วยสูงเป็นอันดับ 5 ของประเทศ ประมาณ 137.46 ต่อแสนประชากร (ข้อมูลเฝ้าระวังโรค ตั้งแต่ 1 ม.ค. – 15 พ.ย. 58) และผลการศึกษาหาตัวแบบที่เหมาะที่สุดด้วยวิธีบ็อกซ์-เจนกินส์ คือ ARIMA(0,1,0)(1,1,1)$_{12}$ และได้ค่าพยากรณ์ตั้งแต่เดือนกันยายนถึงพฤศจิกายน พ.ศ.2558 และการเปรียบเทียบประสิทธิภาพวิธีการพยากรณ์ แสดงตารางที่ 2 และผลจากการเปรียบเทียบค่าข้อมูลจริงกับค่าพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์คือ ARIMA(0,1,0)(1,1,1)$_{12}$ แสดงดังภาพที่ 4

ตารางที่ 2 แสดงค่าพยากรณ์และค่าเปรียบเทียบประสิทธิภาพวิธีการพยากรณ์

ช่วงเวลา	จำนวนผู้ป่วยโรคสุกใสจริง (คน)	ค่าพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์ (คน)	เปรียบเทียบประสิทธิภาพของวิธีการพยากรณ์		
			RMSE	MAPE	MAE
กันยายน 58	70	84.34			
ตุลาคม 58	56	72.62	33.17	18.25%	25.64
พฤศจิกายน 58	-	61.93			



ภาพที่ 4 แสดงการเปรียบเทียบค่าข้อมูลจริงกับค่าพยากรณ์ด้วยวิธีบ็อกซ์-เจนกินส์

วิจารณ์

จากการศึกษานำเสนอวิธีการพยากรณ์อนุกรมเวลาด้วยวิธีบ็อกซ์-เจนกินส์ที่เหมาะที่สุดกับอนุกรมเวลาจำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่ ซึ่งปัจจุบันยังไม่เคยมีการพยากรณ์จำนวนผู้ป่วยด้วยวิธีดังกล่าว โดยใช้ข้อมูลตั้งเดือนพฤษภาคม พ.ศ. 2552 จนถึงเดือนกันยายน พ.ศ. 2558 จำนวน 76 ค่า จากรายงานของสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข กรุงเทพมหานคร ซึ่งจากวิธีบ็อกซ์-เจนกินส์ตัวแบบที่เหมาะที่สุดกับข้อมูลชุดนี้ คือ ARIMA$(0,1,0)(1,1,1)_{12}$ มีความแม่นยำที่สุดซึ่งเมื่อเปรียบเทียบประสิทธิภาพวิธีการพยากรณ์ด้วยวิธีทั้ง 3 นั่นคือ ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) มีค่าเท่ากับ 33.17 ร้อยละของค่าสัมบูรณ์ของความคลาดเคลื่อนเฉลี่ย (MAPE) มีค่าเท่ากับร้อยละ 18.25 และค่าสัมบูรณ์ของค่าคลาดเคลื่อนเฉลี่ย (MAE) มีค่าเท่ากับ 25.64 ดังนั้น การลดจำนวนผู้ป่วยที่เกิดโรคสุกใสในพื้นที่จังหวัดเชียงใหม่ให้มีประสิทธิภาพ ควรดำเนินการป้องกันควบคุมโรคล่วงหน้าเพื่อลดจำนวนผู้ป่วยก่อนถึงช่วงที่มีการระบาดของโรคสุกใสในพื้นที่ให้มีประสิทธิภาพ

บทสรุป

งานวิจัยนี้ได้ตัวแบบการพยากรณ์จากวิธีบ็อกซ์-เจนกินส์ คือ ARIMA$(0,1,0)(1,1,1)_{12}$ ซึ่งตัวแบบนี้สามารถนำไปใช้พยากรณ์จำนวนผู้ป่วยที่จะเกิดขึ้นและสามารถนำไปประกอบการวางแผนเพื่อหาแนวทางป้องกันและควบคุมโรคที่จะเกิดขึ้นในอนาคตได้ อีกทั้งนำไปประยุกต์ใช้พยากรณ์จำนวนผู้ป่วยที่มีอนุกรมเวลาลักษณะการเคลื่อนไหวที่คล้ายคลึงกัน

เอกสารอ้างอิง

1. คณินทร์ ธีรภาพโอฬาร. เอกสารการเรียนการสอน “รายวิชาเทคนิคการพยากรณ์”. วิชาเทคนิคการพยากรณ์ สาขาสถิติ. ชลบุรี: คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา; 2558.
2. ดาว สงวนรังศิริกุล, หรรษา เชี่ยวอนันตวานิช และมณีรัตน์ แสงเกษม. การศึกษาเปรียบเทียบเพื่อหาตัวแบบที่เหมาะสมสำหรับการพยากรณ์จำนวนผู้ป่วยที่เป็นโรคเฝ้าระวังทางระบาดวิทยาในกรุงเทพมหานคร. วารสารวิจัยและพัฒนา มจธ. 2558; 38: 35-55.
3. มุกดา แม้นมินทร์. อนุกรมเวลาและการพยากรณ์. กรุงเทพฯ: โรงพิมพ์ บริษัท โฟร์พริ้นติ้ง จำกัด; 2549.
4. วรางคณา กีรติวิบูลย์. การเปรียบเทียบวิธีการพยากรณ์ระหว่างวิธีบ็อกซ์-เจนกินส์ วิธีการทำให้เรียบแบบเอกซ์โพเนนเชียลที่มีฤดูกาลอย่างง่าย และวิธีการพยากรณ์รวม สำหรับการพยากรณ์อุณหภูมิเฉลี่ยต่อเดือนในเขตกรุงเทพมหานคร. วารสารวิทยาศาสตร์บูรพา 2556; 18: 149-60.
5. สำนักโรคระบาด. จำนวนผู้ป่วย-เสียชีวิตรายเดือนโรคสุกใส[อินเตอร์เน็ต]. 2558. [เข้าถึงเมื่อ 3 พฤศจิกายน 2558]. เข้าถึงได้จากhttp://www.boe.moph.go.th/boedb/surdata/disease.php6.
6. Bowerman BL, O’Connell RT. Koehler AB. Forecasting, time series and regression: An applied approach. 4thed. The United States of America: Thomson Books; 2005.