

การพยากรณ์จำนวนผู้ป่วยโรคความดันโลหิตสูงรายใหม่ จังหวัดปทุมธานี

ภาคภูมิ อุนหละขจร^{1*} เจริญชัย อึ้งเจริญสุข² อัษฎาวุฒิ โยธาสุภาพ³ วัฒนา ชยธวัช⁴

Received : January 11, 2024

Revised : July 18, 2024

Accepted : August 15, 2024

บทคัดย่อ

ในประเทศไทย โรคความดันโลหิตสูงเป็นโรคในกลุ่มหลอดเลือดที่เป็นสาเหตุของความเจ็บป่วยด้วยภาวะแทรกซ้อนมากที่สุดและทำให้ผู้ป่วยที่ได้ได้รับการรักษาอย่างทันท่วงทีเสียชีวิต เช่น ภาวะหัวใจล้มเหลว เส้นเลือดในสมองแตก เป็นต้น การพยากรณ์โรคเป็นมาตรการหนึ่งในแผนยุทธศาสตร์ของกรมควบคุมโรคที่ใช้คาดการณ์จำนวนผู้ป่วยที่จะเกิดขึ้นเพื่อนำไปสู่การวางแผนป้องกันด้วยการกำหนดมาตรการหรือนโยบายในรูปของกิจกรรมโครงการเพื่อสร้างเสริมสุขภาพให้เหมาะสมกับจำนวนผู้ป่วยเพื่อลดจำนวนผู้ป่วยในระยะยาว โดยการใช้เทคนิคการพยากรณ์หรือการวิเคราะห์อนุกรมเวลา วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method) ข้อมูลที่ใช้ในการพยากรณ์ คือ ข้อมูลทุติยภูมิผู้ป่วยโรคความดันโลหิตสูง ปีงบประมาณ 2560 - 2566 เป็นข้อมูลเดือนมกราคม 2560 ถึงธันวาคม 2565 จำนวน 72 เดือน ใช้พยากรณ์จำนวนผู้ป่วยรายเดือน พ.ศ. 2566 เปรียบเทียบกับข้อมูลจริงเดือนมกราคม ถึง กันยายน 2566 การสร้างแบบจำลองตามวิธีการบ็อกซ์และเจนกินส์ ARIMA(p,d,q)(P,D,Q)_m ด้วยฟังก์ชัน auto.arima() ใน packages (forecast) จากโมดูล snowCluster โปรแกรมจามาโมวี (jamovi)

ผลการศึกษา พบว่า แบบจำลองอนุกรมเวลา ARIMA (2,1,1)(2,0,0)₁₂ มีความสอดคล้องกับข้อมูลจริง ค่าเฉลี่ยของเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ (MAPE) ของข้อมูล 72 เดือน ของปี พ.ศ. 2560 ถึง 2565 ร้อยละ 11.02 เมื่อนำค่าพยากรณ์ปี พ.ศ. 2566 12 เดือน เทียบกับข้อมูลจริงปี พ.ศ. 2566 มกราคม ถึง กันยายน 2566 พบว่ามีค่าเฉลี่ยของเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ เฉลี่ย 9 เดือน เท่ากับร้อยละ 8.86 ซึ่งอยู่ในเกณฑ์ที่ใช้พยากรณ์ได้ดี สอดคล้องกับผลการศึกษาในประเทศกานาที่ได้นำอนุกรมเวลาพยากรณ์จำนวนผู้ป่วยด้วยอนุกรมเวลาใช้ในการพยากรณ์โรคความดันโลหิตสูง แต่อย่างไรก็ตามผลที่ได้ยังข้อมีจำกัดเนื่องจากไม่ได้มีการนำปัจจัยอื่นๆ เช่น พฤติกรรมการบริโภคอาหาร ความเป็นอยู่ของประชากรในพื้นที่ เข้ามาร่วมในการวิเคราะห์ด้วย ในการศึกษาครั้งถัดไปต้องมีการศึกษาเปรียบเทียบกับวิธีการอื่นเพื่อให้ผลการศึกษาที่มีความน่าเชื่อถือมากขึ้น

คำสำคัญ: การพยากรณ์ ความดันโลหิตสูง ผู้ป่วยรายใหม่ วิธีบ็อกซ์และเจนกินส์

¹ อาจารย์ประจำหลักสูตรสาขาวิชาสาธารณสุขชุมชน คณะสหเวชศาสตร์ มหาวิทยาลัยปทุมธานี อีเมล: Pakpoom.c@ptu.ac.th

² อาจารย์ประจำหลักสูตรสาขาวิชาสาธารณสุขชุมชน คณะสหเวชศาสตร์ มหาวิทยาลัยปทุมธานี อีเมล: Jaroenchai@ptu.ac.th

³ อาจารย์ประจำหลักสูตรสาขาวิชาสาธารณสุข คณะสาธารณสุขศาสตร์ มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ ในพระบรมราชูปถัมภ์ จังหวัดปทุมธานี อีเมล: aussadawut.yo@vru.ac.th

⁴ อาจารย์ประจำหลักสูตรสาขาวิชาการแพทย์แผนไทย คณะสหเวชศาสตร์ มหาวิทยาลัยปทุมธานี อีเมล: vadhana.j@ptu.ac.th

* ผู้นิพนธ์หลัก อีเมล: Pakpoom.C@ptu.ac.th

FORECASTING THE NUMBER OF HIGH BLOOD PRESSURE INCIDENTS IN PATHUM THANI PROVINCE

Pakpoom Ounhalekjit¹ Jaroenchai Ungcharoensuk² Aussadawut Yothasupap³
Vadhana Jayathavaj⁴

Abstract

Hypertension is a leading vascular disease in Thailand, often resulting in severe complications and fatalities, such as heart failure and stroke, when timely treatment is not provided. To address this, the Department of Disease Control employs disease forecasting as part of its strategy to predict the number of potential hypertension cases. This forecasting aids in the development of preventive measures and health promotion activities, aiming to reduce the disease burden over time. In this study, time series analysis using the Box-Jenkins method was utilized for forecasting. The analysis was based on secondary data of hypertension patients from the fiscal years 2017 to 2023, covering the period from January 2017 to December 2022—a total of 72 months. Monthly patient numbers for 2023 were forecasted and subsequently compared with actual data from January to September 2023. The forecasting model was constructed using the Box-Jenkins ARIMA(p,d,q)(P,D,Q)m method, implemented through the `auto.arima()` function from the forecast package in the snowCluster module of the jamovi program.

The study found that the ARIMA (2,1,1)(2,0,0)12 time series model closely matched the actual data. The mean absolute percentage error (MAPE) over the 72 months from 2017 to 2022 was 11.02%. When comparing the forecasted values for the 12 months of 2023 with the actual data from January to September 2023, the MAPE over these 9 months was 8.86%, indicating a high level of forecast accuracy. This finding is consistent with a study in Ghana that used time series analysis to forecast hypertension cases. However, the results have limitations, as other factors such as dietary behaviors and living conditions of the population were not included in the analysis. Future studies should compare this method with others to improve the credibility of the results.

Keywords: Forecasting, High blood pressure, Incidents, Box and Jenkins Method

¹ Lecturer of Department of Public Health, Faculty of Allied Health Sciences, Patumthani University, e-mail: Pakpoom.C@ptu.ac.th

² Lecturer of Department of Public Health, Faculty of Allied Health Sciences, Patumthani University, e-mail: Jaroenchai@ptu.ac.th

³ Lecturer of Department of Public Health, Faculty of Public Health, Valaya Alongkorn Rajabhat University under the Royal Patronage Pathum Thani Province, e-mail: aussadawut.yo@vru.ac.th

⁴ Lecturer of Department of Thai Traditional Medicine, Faculty of Allied Health Sciences, Patumthani University, e-mail: vadhana.j@ptu.ac.th

* Corresponding author e-mail: Pakpoom.C@ptu.ac.th

บทนำ

โรคความดันโลหิตสูงเป็นโรคไม่ติดต่อเรื้อรังที่เป็นปัญหาสำคัญทางสาธารณสุขทั่วโลก โดยเป็นสาเหตุหลักของการเจ็บป่วยและการเสียชีวิตโดยคิดเป็นประมาณ 7.5 ล้านคน หรือประมาณร้อยละ 12.8 ของการเสียชีวิตทั้งหมด (DALYs) ในการวินิจฉัยโรคความดันโลหิตสูงเบื้องต้นจะพิจารณาจากการวัดความดันโลหิตในระดับที่สูงกว่าปกติเรื้อรังอยู่เป็นเวลานาน มากกว่า 140/90 มม.ปรอท ค่าความดันโลหิตตัวบนหรือความดันขณะหัวใจบีบตัว (Systolic blood pressure; SBP) มากกว่า 140 มม.ปรอท และ ความดันโลหิตโลหิตตัวล่างหรือความดันขณะหัวใจคลายตัว (Diastolic blood pressure; SBP) น้อยกว่า 90 มม.ปรอท ซึ่งต้องวัดอย่างน้อย 2 ครั้งในแต่ละครั้ง และวัดซ้ำห่างกัน 2 สัปดาห์ (สมาคมความดันโลหิตสูงแห่งประเทศไทย, 2562) จากรายงานผลการตรวจสุขภาพประชาชนไทยโดยการตรวจร่างกายครั้งที่ 6 พ.ศ. 2562-2563 ผลการคัดกรองระดับความดันโลหิตเฉลี่ยของประชาชนอายุ 15 ปีขึ้นไป เท่ากับ 124.0 และ 72.9 มม.ปรอท ตามลำดับ ผู้ชายจะมีค่าความดันโลหิตโดยเฉลี่ยสูงกว่าเพศหญิง โดยค่าความดันโลหิตตัวบนโดยเฉลี่ยของเพศชายและผู้หญิง มีค่า 127.10 และ 121.10 มม.ปรอท ตามลำดับ ส่วนค่าความดันโลหิตตัวล่างโดยเฉลี่ยของเพศชายและผู้หญิง มีค่า 74.6 และ 71.30 มม.ปรอท ตามลำดับ ค่าความดันโลหิตตัวล่างจะเพิ่มขึ้นตามอายุจนถึงวัย 60 ถึง 69 ปี หลังจากนั้นความดันโลหิตตัวล่างลดลงเมื่ออายุเพิ่มขึ้น ในส่วนของรายงานการวิเคราะห์ความชุกของโรคความดันโลหิตสูงในประชาชนไทยอายุ 15 ปีขึ้นไปมีร้อยละ 25.4 ผู้ชายมีความชุกสูงกว่าของผู้หญิง (วิชัย เอกพลากร, 2564)

จากสถานการณ์ปัญหาโรคไม่ติดต่อเรื้อรังในกลุ่มโรคหัวใจและหลอดเลือดของประชากรไทยในปัจจุบัน มีจำนวนผู้ป่วยที่ป่วยเป็นโรคความดันโลหิตสูงเพิ่มมากขึ้นทุกปี ข้อมูลอัตราป่วยด้วยโรคความดันโลหิตสูง ข้อมูลสถิติจำนวนผู้ป่วยด้วยโรคความดันโลหิตสูง 5 ปีย้อนหลัง มีการรายงานข้อมูลอัตราป่วยโรคความดันโลหิตสูง 5 ปีย้อนหลัง (ต่อประชากรแสนคน) ดังต่อไปนี้ ปี พ.ศ. 2564 มีอัตราป่วย 14,618.0 รายต่อประชากรแสนคน พ.ศ. 2563 มีอัตราป่วย 14,618.0 รายต่อประชากรแสนคน พ.ศ. 2562 มีอัตราป่วย 13,361.1 รายต่อประชากรแสนคน พ.ศ. 2561 มีอัตราป่วย 12,605.5 รายต่อประชากรแสนคน (กองโรคไม่ติดต่อ กรมควบคุมโรค, 2564) จะเห็นได้ว่าเมื่อความดันโลหิตยังคงสูงอยู่ตลอดเวลา ทำให้หัวใจสูบฉีดหนักขึ้นและทำงานมากขึ้น ซึ่งอาจนำไปสู่ภาวะแทรกซ้อนที่อาจเกิดขึ้นได้ เช่น ภาวะหัวใจล้มเหลว โรคหลอดเลือดสมอง และภาวะไตวาย การรักษาความดันโลหิตสูงทำได้โดยการรักษาด้วยยาร่วมกับเปลี่ยนแปลงวิถีชีวิตและปรับเปลี่ยนพฤติกรรมสุขภาพเพื่อลดปัจจัยเสี่ยง ได้แก่ การออกกำลังกาย ลดการรับประทานอาหารที่มีโซเดียมหรืออาหารรสเค็ม และการฝึกควบคุมอารมณ์ จะช่วยให้ผู้ป่วยมีอาการที่ดีขึ้น (MedlinePlus, 2020)

การทำนาย (Prediction) หรือการพยากรณ์ (Forecast) จำนวนผู้ป่วยด้วยโรคเฝ้าระวังที่อาจเกิดขึ้นในอนาคตจะทำให้ได้ข้อมูลการเฝ้าระวังที่เป็นประโยชน์ในการตรวจจับการระบาดของโรค ติดตามสถานการณ์โรค อธิบายธรรมชาติและการกระจายของโรค ประเมินผลมาตรการควบคุมป้องกันโรคและประเมินความเสี่ยงทางสุขภาพที่อาจเกิดขึ้นได้ (วารจกณา เรียนสุทธิ, 2563) ทั้งนี้การพยากรณ์โรคเป็นมาตรการหนึ่งในแผนยุทธศาสตร์ของกรมควบคุมโรค เพื่อคุ้มครองสุขภาพของประชาชนในลักษณะของการคาดการณ์หรือการทำนายล่วงหน้าเกี่ยวกับโรคที่จะเกิดขึ้นซึ่งมีเป้าหมายเพื่อให้หน่วยงานทั้งส่วนกลางและส่วนภูมิภาคสามารถนำข้อมูลที่ได้ไปใช้ใน

การวางแผนป้องกันและเตรียมรับมือกับโรคที่เกิดขึ้นเพื่อลดจำนวนผู้ป่วยในอนาคต เทคนิคการพยากรณ์หรือการวิเคราะห์อนุกรมเวลาที่นำมาใช้ในการพยากรณ์ที่ คือ วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method) (อันวา เจริญศิริ และคณะ, 2558) สามารถอธิบายรูปแบบของการเกิดโรคตามเวลาเพื่อใช้ประเมินผลมาตรการควบคุมป้องกันโรคและความเสี่ยงทางสุขภาพที่อาจเกิดขึ้น เช่น การพยากรณ์การระบาดของโรคไข้หวัดใหญ่ (Tseng & Shih, 2020) โรคไข้เลือดออก (Mekpanyup & Saithanu, 2016) การพยากรณ์จำนวนผู้ป่วยโรคปอดบวม (วารางคณา กิรติวิบูลย์, 2559) แต่ในทางกลับกันกลุ่มโรคไม่ติดต่ออย่างไม่มีหรือมีการศึกษาในไม่มากนักซึ่งในทางระบาดวิทยาของโรคไม่ติดต่อเรื้อรัง (Non-Communicable Diseases; NCDs) สาเหตุของการเกิดโรคมาจากหลายปัจจัย ได้แก่ อายุ เพศ พันธุกรรม สภาพแวดล้อม พฤติกรรมที่ไม่เหมาะสม เช่น การบริโภคอาหารที่ผิด ไม่ออกกำลังกาย เป็นต้น ซึ่งมีความซับซ้อนกว่าโรคติดต่อแต่ด้วยเหตุผลของความสัมพันธ์ต่อการเกิดโรคในเรื่องของเวลาเมื่อประชาชนอายุมากขึ้นแนวโน้มการเกิดโรคดังกล่าวก็พบได้มากขึ้นด้วย

การพยากรณ์อนุกรมเวลาด้วยแบบจำลอง ARIMA (Auto Regressive Integrated Moving Average) เป็นวิธีการแบบสถิติที่ใช้ในการวิเคราะห์และการพยากรณ์ข้อมูลชุดเวลา (Time series data) ประกอบด้วย 3 ส่วนหลัก 1) Autoregression (AR) ค่าถ่วงน้ำหนักของค่าในช่วงเวลาก่อนหน้า 2) Integration (I) ผลต่างระหว่างค่าในช่วงเวลาหนึ่งกับค่าในช่วงเวลาก่อนหน้า เพื่อให้ข้อมูลนิ่ง (Stationary) และ 3) Moving Average (MA) ค่าถ่วงน้ำหนักของค่าความคลาดเคลื่อนของค่าในช่วงเวลาก่อนหน้า แบบจำลอง ARIMA สามารถพยากรณ์แนวโน้มของข้อมูลอนุกรมเวลาจึงมีโครงสร้างที่ซับซ้อน โดยพิจารณาไปจากความสัมพันธ์ระหว่างข้อมูลในช่วงเวลาก่อนหน้า ความเปลี่ยนแปลงที่ผ่านมา เพื่อให้ได้ค่าพยากรณ์ที่มีความแม่นยำและมีประสิทธิภาพ โดยได้มีการนำมาใช้ในเชิงธุรกิจ เช่น ออกแบบการทำนายยอดขาย การวิเคราะห์การลงทุน หรือ วิเคราะห์แนวโน้มของตลาด อย่างไรก็ตาม มีการนำแบบจำลอง ARIMA ไปใช้ในการทำนายจำนวนผู้ป่วยโรคความดันโลหิตสูงและโรคเบาหวานในประเทศกานา โดยใช้ข้อมูลจำนวนผู้ป่วยตั้งแต่เดือนมกราคม 2559 ถึง ธันวาคม 2563 วัดความแม่นยำด้วยค่าเฉลี่ยของเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ (The Mean Absolute Percentage Error - MAPE) (Asante et al., 2022) ส่วนการศึกษาจำนวนผู้ป่วยโรคความดันโลหิตสูงได้ใช้ข้อมูลรายเดือนตั้งแต่เดือนมกราคม 2558 ถึงธันวาคม 2562 (Larty et al., 2020) การพยากรณ์โรคที่เป็นสาเหตุของการเสียชีวิตในประเทศฟิลิปปินส์ใช้ข้อมูลปี 2012 ถึง 2016 ทำการพยากรณ์จำนวนผู้ป่วยปี 2017-2020 (Talirongan et al., 2020)

จากรายงานผลการดำเนินงานประจำปีงบประมาณ 2566 เขตสุขภาพที่ 4 พบว่าประชาชนส่วนใหญ่ป่วยด้วยกลุ่มโรคไม่ติดต่อเรื้อรัง (Non-Communicable diseases : NCDs) โดยเฉพาะโรคความดันโลหิตสูง (กระทรวงสาธารณสุข, 2566) พบว่ามีแนวโน้มผู้ป่วยโรคความดันโลหิตสูงในจังหวัดพุมธานี มีจำนวนสูงขึ้นตั้งข้อมูลเชิงสถิติ 5 ปีย้อนหลัง ตั้งแต่ปี พ.ศ. 2561-2565 มีจำนวนผู้ป่วยโรคความดันโลหิตสูงรายใหม่ ดังนี้ ปี 2561 มีจำนวน 8,635 คน ปี 2562 มีจำนวน 9,219 คน ปี 2563 มีจำนวน 10,109 คน ปี 2564 มีจำนวน 11,674 คน และ ปี 2565 มีจำนวน 9,940 คน จะเห็นว่าแนวโน้มจำนวนผู้ป่วยรายใหม่มีแนวโน้มเพิ่มมากขึ้นด้วยสาเหตุจากพฤติกรรมสุขภาพของตัวผู้ป่วยเอง เช่น การรับประทานอาหารมัน เค็ม และ ไม่ออกกำลังกาย

(เขตสุขภาพที่ 4, 2566) ผู้ป่วยส่วนใหญ่จะเริ่มให้ความสำคัญกับการป้องกันก็เมื่อได้เกิดภาวะแทรกซ้อนกับร่างกายแล้วส่งผลทำให้คุณภาพชีวิตผู้ป่วยไม่ดีขึ้นจนถึงขั้นเสียชีวิตได้ ถ้าทราบแนวโน้มจำนวนผู้ป่วยที่เกิดขึ้นในอนาคตก็จะเป็นเพื่อให้ได้ตัวแบบอนุกรมเวลาที่สามารถพยากรณ์การเพิ่มขึ้นในอนาคตได้อย่างมีประสิทธิภาพเพื่อช่วยในการป้องกันและสามารถนำไปกำหนดกลยุทธ์แก้ไขปัญหาด้านสุขภาพโดยการเตรียมความพร้อมของผู้ปฏิบัติงานด้านการส่งเสริมสุขภาพต่อไป

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาการพยากรณ์ผู้ป่วยโรคความดันโลหิตสูงรายใหม่ของจังหวัดปทุมธานี ปี พ.ศ. 2566
2. เพื่อประเมินประสิทธิภาพความแม่นยำของแบบจำลอง เพื่อใช้ในการพยากรณ์ผู้ป่วยโรคความดันโลหิตสูงรายใหม่ของจังหวัดปทุมธานี ปี พ.ศ. 2567 ต่อไป

วิธีดำเนินการวิจัย

วิธีการพยากรณ์บอกซ์และเจนกินส์

อนุกรมเวลา คือ ข้อมูลที่เรียงลำดับตามเวลา การวิเคราะห์อนุกรมเวลาใช้เพื่อตรวจสอบรูปแบบของการเปลี่ยนแปลงข้อมูลทางสถิติในช่วงเวลาปกติ คาดการณ์รูปแบบเหล่านี้เพื่อให้ได้ค่าประมาณสำหรับอนาคต วิธีการพยากรณ์ทางสถิติทั้งหมดมีลักษณะเป็นการประมาณการณ์ กล่าวคือ เกี่ยวข้องกับการฉายภาพรูปแบบในอดีตหรือความสัมพันธ์ในอนาคต อนุกรมเวลาสามารถเป็นแบบอยู่กับที่ (Stationary) และไม่อยู่กับที่ (Non-stationary) อนุกรมเวลาจะเรียกว่า คงที่ (Stationary) หากมีค่าเฉลี่ยและความแปรปรวนคงที่

Back shift operator หรือ Lag operator B การถอยข้อมูลย้อนกลับไป h

คาบเวลา

$$By_t = y_{t-1} \tag{1}$$

$$B(By_t) = B^2 y_t = y_{t-2}$$

$$B^h y_t = y_{t-h}$$

Differencing ผลต่างของตัวแปรกับคาบเวลาก่อนหน้า

$$y'_t = y_t - y_{t-1} = y_t - By_t = (1 - B)y_t \tag{2}$$

$$y''_t = y_t - y_{t-1} - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$$

$$= (1 - 2B + B^2) y_t = (1 - B)^2 y_t$$

a d^{th} -order difference ผลต่างลำดับของตัวแปรลำดับที่ d (3)

$$y_t^d = (1 - B)^d y_t$$

A first seasonal difference สำหรับ m คาบ ใน 1 รอบฤดูกาล

$$(1 - B)(1 - B^m) y_t = (1 - B - B^m + B^{m+1}) y_t \tag{4}$$

$$= y_t - y_{t-1} - y_{t-m} + y_{t-m-1}$$

An AutoRegressive model of order p - AR(p)

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t \tag{5}$$

$$\begin{aligned} & \text{A Moving Average model of order } q - \text{MA}(q) \\ & y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} \end{aligned} \quad (6)$$

เมื่อนำ Differencing AR(p) และ MA(q) มารวมกัน เป็น ARIMA(p,d,q)

$$y_t^d = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} + e_t \quad (7)$$

$$y_t^d = \nabla^d y_t$$

กรณีที่มีฤดูกาล (seasonal) ก็จะเป็น a multiplicative (p,d,q)(P,D,Q)_m SARIMA model with seasonal period “m” หรือ ARIMA(p,d,q)(P,D,Q)_m

P และ Q เป็นจำนวนเทอมของฤดูกาลเดียวกันย้อนไปในอดีต P และ Q เทอมตามลำดับ

เมื่อ “d” และ “D” หมายถึง จำนวนของความแตกต่าง (an order difference) ที่ไม่ใช่ฤดูกาลและตามฤดูกาล ตามลำดับ (8)

เมื่อ m เป็นช่วงเวลาใน 1 ฤดูกาล แล้ว ARIMA(p,d,q)(P,D,Q)_m คือ

$$\phi(B)\Phi_p(B^m)\nabla^d\nabla_m^D Y_t = \theta(B)\Theta_Q(B^m)\varepsilon_t$$

โดยที่ the autoregressive and moving average operators respectively for the non-seasonal characteristics

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \quad (9)$$

$$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

The autoregressive and moving average operators for the seasonal characteristics

$$\Phi_p(B^m) = (1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p) \quad (10)$$

$$\Theta_Q(B^m) = (1 - \Theta_1 B - \Theta_2 B^2 - \dots - \Theta_Q B^Q)$$

$$\nabla^d \nabla_m^D Y_t = (1 - B)^d (1 - B^m)^D Y_t$$

การพัฒนาเกี่ยวกับทฤษฎีการวิเคราะห์อนุกรมเวลาการพิจารณาเลือกตัวแบบจากสหสัมพันธ์ในตัวเอง (Autocorrelation) กับค่าที่เกิดก่อนหน้า (Lag) จาก ฟังก์ชันสหสัมพันธ์ในตัวเอง (Autocorrelation function; ACF) และ คอเรโลแกรม (Correlogram) (Box & Jenkins, 1976; Madsen, 2008; Makridakis et al., 1982; Meese & Geweke, 1982) การเลือกแบบจำลองที่มีความสอดคล้องกับข้อมูลมากที่สุด คือ การเลือกแบบจำลองที่มีค่า p,d,q ในส่วนที่ไม่มีฤดูกาล, l, และ P,D,Q ส่วนที่มีฤดูกาล m ซึ่งเป็นความยุ่งยากเนื่องจากมีแบบจำลองที่ต้องคำนวณเป็นจำนวนมาก เกณฑ์ที่ใช้ในการเลือกแบบจำลอง คือ Akaike’s Information Criterion (AIC) และ the corrected AIC (AICc) ของแบบจำลองที่มีค่าน้อยที่สุด (Hyndman & Athanasopoulos, 2018)

การเลือกแบบจำลองที่เหมาะสมที่สุดตามวิธีการบอกซ์และเจนกินส์มีความซับซ้อนที่ต้องการผู้ที่ได้รับการฝึกอบรมอย่างเหมาะสมเป็นผู้ดำเนินการ ในต่างประเทศและประเทศไทยมีการใช้วิธีการบอกซ์และเจนกินส์พยากรณ์อนุกรมเวลาในวิทยานิพนธ์ในกลุ่มนักศึกษามหาวิทยาลัยจำนวนมาก (Osarumwense, 2013) แต่ในธุรกิจเป็น

เรื่องปกติที่จะมีสายผลิตภัณฑ์จำนวนมากที่ต้องคาดการณ์อย่างน้อยเดือนละครั้ง ดังนั้น การพัฒนาโปรแกรมที่สามารถเลือกแบบจำลองที่เหมาะสมให้โดยอัตโนมัติ ซึ่งอัลกอริธึมการพยากรณ์เป็นเครื่องมือสำคัญในการประมาณการพารามิเตอร์ของแบบจำลองและคำนวณค่าพยากรณ์โดยไม่มีแทรกแซงจากผู้ใช้ ผู้ใช้โปรแกรมเพียงตัดสินใจเลือกนำค่าพยากรณ์ไปใช้งาน ซึ่งจะนำเสนอการคำนวณตัวแบบมากมายเพื่อเปรียบเทียบตามเกณฑ์การตัดสินใจ ขณะนี้ในปัจจุบันมีการพัฒนาโปรแกรมที่ให้ผลลัพธ์เป็นแบบจำลอง ARIMA(p,d,q)(P,D,Q)m โดยอัตโนมัติ คือ ฟังก์ชัน `auto.arima()` ใน packages (Forecast) ที่พัฒนาด้วยโปรแกรม R ทำงานบน R Studio (Hyndman RJ & Khandakar, 2008) หรือสามารถเรียกใช้ฟังก์ชัน `auto.arima()` จากโมดูล `snowCluster` >> Univariate time series ในโปรแกรมวิเคราะห์ทางสถิติสำเร็จรูป jamovi ได้ (Hyndman, et al., 2022; R Core Team, 2021; Seol, 2023; The jamovi project, 2022; วัฒนา ชยธวัช, 2566)

เกณฑ์ตัดสินความแม่นยำของแบบจำลอง

ค่าข้อมูลจริง y_i จำนวน n ค่า โดย $i = 1, 2, \dots, n$ ส่วน y_i เป็นค่าพยากรณ์ สำหรับคาบเวลาตามข้อมูลที่นำมาใช้พยากรณ์ ส่วนค่าพยากรณ์ไปในอนาคต $i = n + 1, n + 2, \dots$

ค่าเฉลี่ยของเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ (The Mean Absolute Percentage Error - MAPE) (Andrés, 2023)

$$MAPE = \left(\frac{1}{n} \sum_{k=1}^n \left| \frac{y_k - \hat{y}_k}{y_k} \right| \right) \times 100\%$$

MAPE หน่วยเป็นร้อยละ ถ้าน้อยกว่า 10 มีความแม่นยำสูง ระหว่าง 10-20 ใช้พยากรณ์ได้ดี ระหว่าง 20-50 มีเหตุผลพอที่จะใช้พยากรณ์ และ มากกว่า 50 ไม่มีความแม่นยำ (Lewis, 1982)

การพิทักษ์สิทธิ์ของอาสาสมัคร/กลุ่มตัวอย่าง

ข้อมูลที่นำมาวิเคราะห์เป็นข้อมูลรายงานข้อมูลทุติยภูมิจำนวนผู้ป่วยความดันโลหิตสูงรายใหม่โดยรวมเป็นรายเดือน ระหว่างปี พ.ศ. 2560 - 2566 ของจังหวัดปทุมธานี ที่เผยแพร่สาธารณะโดยกระทรวงสาธารณสุข 2566 ไม่สามารถระบุตัวตนเป็นรายบุคคลได้ (ศูนย์ส่งเสริมจริยธรรมการวิจัย มหาวิทยาลัยมหิดล, 2565)

ผลการวิจัย

การรวบรวมข้อมูล

รวบรวมข้อมูลจากระบบ Health Data Center (HDC report) ของกระทรวงสาธารณสุขได้จาก อ่างอิงและ link ที่แนบกระทรวงสาธารณสุข (2566) ดังแสดงในตารางที่ 1

ตารางที่ 1 จำนวนผู้ป่วยความดันโลหิตสูงรายใหม่ จังหวัดปทุมธานี (พ.ศ. 2560-2566)

จำนวนผู้ป่วยความดันโลหิตสูงรายใหม่ (ปี พ.ศ.2560-2566)							
เดือน	2560	2561	2562	2563	2564	2565	2566
1) มกราคม	625	949	760	888	958	992	914
2) กุมภาพันธ์	757	980	713	945	798	795	787
3) มีนาคม	703	866	775	948	941	702	856
4) เมษายน	529	511	654	568	745	724	662
5) พฤษภาคม	645	608	700	576	771	839	742
6) มิถุนายน	786	577	634	686	956	781	805
7) กรกฎาคม	828	673	755	663	1,131	700	682
8) สิงหาคม	759	721	891	795	816	743	739
9) กันยายน	644	683	957	869	850	759	796
10) ตุลาคม	695	752	839	1,225	1,322	1,153	
11) พฤศจิกายน	661	713	748	1,045	1,260	941	
12) ธันวาคม	797	602	793	901	1126	811	
รวม	8,429	8,635	9,219	10,109	11,674	9,940	
ร้อยละ+เพิ่มขึ้น/-ลดลงปีต่อปี		2.44	6.76	9.65	15.48	-14.85	
ค่าเฉลี่ย	702	720	768	842	973	828	776
S.D.	83	140	89	187	186	130	75
ค่าสูงสุด	828	980	957	1,225	1,322	1,153	914
ค่าต่ำสุด	529	511	634	568	745	700	662

ที่มาของข้อมูล : Health Data Center (HDC report)

ตารางที่ 1 ข้อมูลที่รวบรวมได้จาก HDC report ของกระทรวงสาธารณสุขจัดเรียงรายเดือนตามปี พ.ศ. 2561 - 2567 แล้ว จำนวนผู้ป่วยรายใหม่เพิ่มขึ้นปีต่อปีอย่างต่อเนื่องในปี พ.ศ. 2562 - 2565 คือ ร้อยละ 2.44, 6.76, 9.65, และ 15.48 ตามลำดับ ขณะที่ปี พ.ศ. 2566 จำนวนผู้ป่วยโรคความดันโลหิตสูงรายใหม่กลับมีจำนวนลดลงถึงร้อยละ -14.85 อาจมาจากผลกระทบของโรคโควิด-19 เมื่อสถานการณ์โรคโควิด-19 เบาลง

การไปพบแพทย์ก็ลดลงตาม ขณะที่โรคความดันโลหิตสูงเป็นโรคที่ควบคุมได้ ถ้าได้รับการวินิจฉัยและควบคุมอย่างใดก็ตามค่าเฉลี่ยรายเดือนมีแนวโน้มจำนวนผู้ป่วยรายใหม่อยู่ที่ 1,000 ราย ต่อเดือน

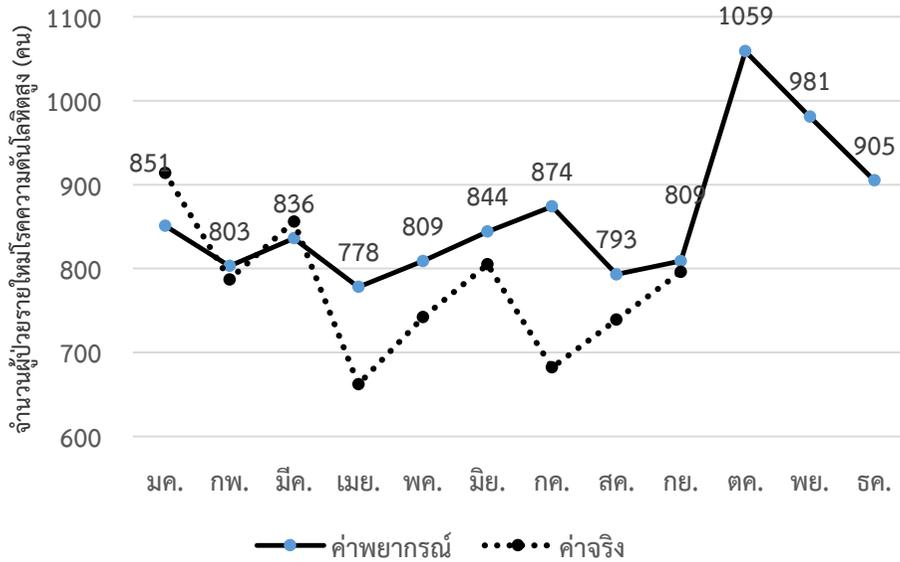
การประมวลผลด้วยโปรแกรม R ฟังก์ชัน auto.arima()

แบบจำลองที่เหมาะสม คือ ARIMA (2,1,1)(2,0,0)₁₂ มีค่าความสอดคล้องกับข้อมูลจริง AIC=894.18, AICc=895.49 และ BIC=907.75 ค่าเฉลี่ยของเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ MAPE ร้อยละ 11.02 ของข้อมูล 72 เดือน ปี พ.ศ. 2560 - 2565 เมื่อนำค่าพยากรณ์ปี พ.ศ. 2566 12 เดือน เทียบกับข้อมูลจริงปี พ.ศ. 2566 มกราคม ถึง กันยายน 2566 พบว่ามี MAPE เฉลี่ย 9 เดือน เท่ากับร้อยละ 8.86 ซึ่งอยู่ในเกณฑ์ที่ใช้พยากรณ์ได้ดี ดังแสดงใน ตารางที่ 2 และภาพที่ 1 ภาพที่ได้จาก โมดูล SnowCluster โปรแกรม Jamovi คือ ภาพรูปแบบข้อมูล พ.ศ. 2560 - 2565 รายเดือน 72 เดือน จำแนกองค์ประกอบอนุกรมเวลา (แนวโน้ม ฤดูกาล และความผันแปรสุ่ม) ภาพแผนผังกล่องข้อมูล มกราคม ถึง ธันวาคม พ.ศ. 2560 - 2565 และ ภาพค่าจริง รายเดือน พ.ศ. 2560 -2565 และค่าพยากรณ์ รายเดือน พ.ศ. 2566 ดังแสดงใน ภาพที่ 2 3 และ 4 ตามลำดับ

ตารางที่ 2 ค่าพยากรณ์จากแบบจำลอง ARIMA(2,1,1)(2,0,0)₁₂

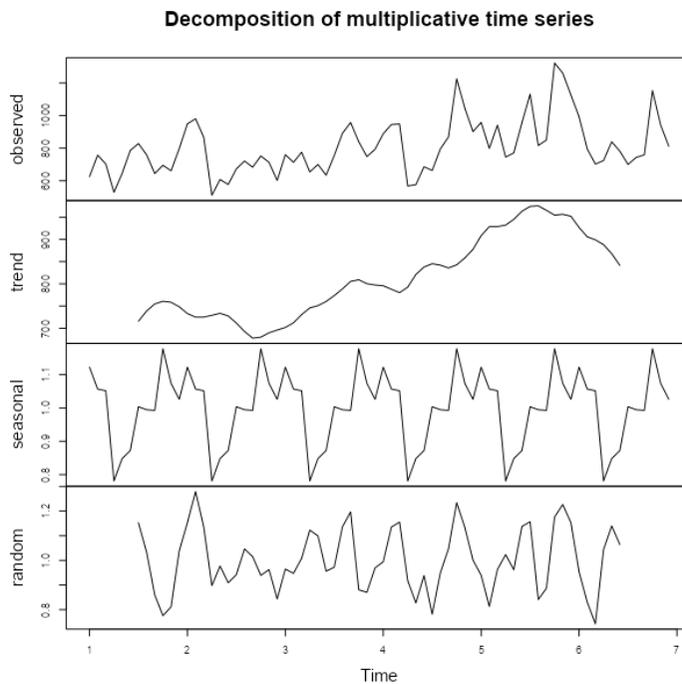
	95% CI		ปี พ.ศ. 2566		
	Lower	Upper	ค่าพยากรณ์	ค่าจริง	APE
1) มกราคม	615	1088	851	914	6.89
2) กุมภาพันธ์	506	1100	803	787	2.03
3) มีนาคม	530	1141	836	856	2.34
4) เมษายน	472	1084	778	662	17.52
5) พฤษภาคม	502	1116	809	742	9.03
6) มิถุนายน	535	1153	844	805	4.84
7) กรกฎาคม	562	1185	874	682	28.15
8) สิงหาคม	478	1108	793	739	7.31
9) กันยายน	490	1127	809	796	1.63
10) ตุลาคม	738	1380	1059		
11) พฤศจิกายน	657	1305	981		
12) ธันวาคม	578	1231	905		
				MAPE	8.86

หมายเหตุ APE คือ ร้อยละความคลาดเคลื่อนสัมบูรณ์



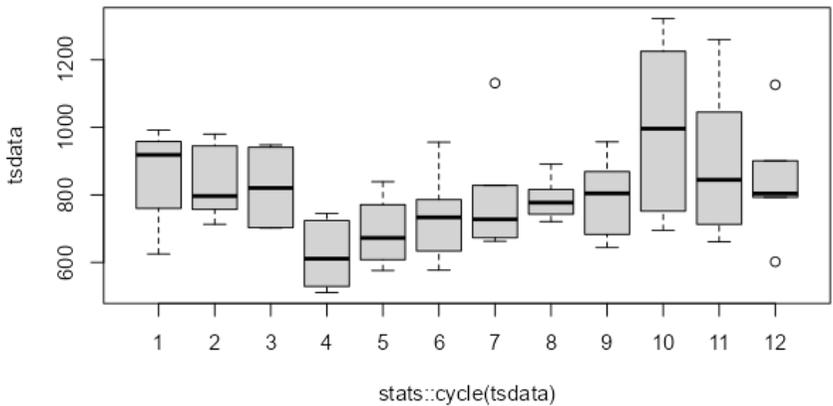
ภาพที่ 1 ค่าจริงและค่าพยากรณ์จำนวนผู้ป่วยรายใหม่โรคความดันโลหิตสูง รายเดือน ปี พ.ศ. 2566

Decomposition



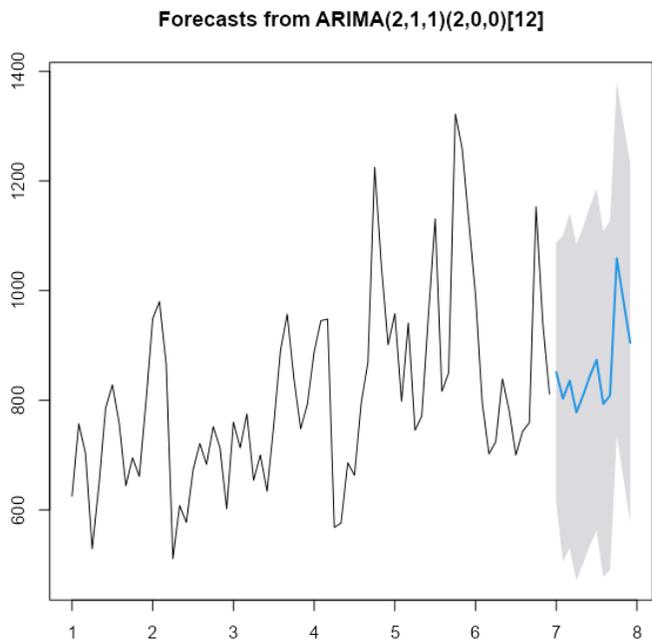
ภาพที่ 2 ภาพรูปแบบข้อมูล 72 เดือน จำแนกองค์ประกอบอนุกรมเวลา

Box plot



ภาพที่ 3 แผนผังกล่องข้อมูล มกราคม ถึง ธันวาคม

Prediction



ภาพที่ 4 ค่าจริงและค่าพยากรณ์

อภิปรายผล

จากการศึกษานำเสนอวิธีการพยากรณ์จำนวนผู้ป่วยโรคความดันโลหิตสูงรายใหม่ ด้วยอนุกรมเวลา บ็อกซ์-เจนกินส์ ด้วยโปรแกรมจาโมวี ใช้ข้อมูลรายเดือนตั้งแต่ปี พ.ศ. 2560 - 2565 จากรายงานของระบบ Health Data Center กระทรวงสาธารณสุข พบว่าแบบจำลองที่เหมาะสม คือ ARIMA (2,1,1)(2,0,0)₁₂ มีค่าความสอดคล้องกับข้อมูลจริง ค่าเฉลี่ยของเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ (MAPE) เท่ากับร้อยละ 11.02 ของข้อมูลจำนวน 72 เดือน ตั้งแต่ปี พ.ศ. 2560 ถึง 2565 และเมื่อนำค่าพยากรณ์ปี พ.ศ. 2566 จำนวน 12 เดือน เทียบกับข้อมูลจริงปี พ.ศ. 2566 มกราคม ถึง กันยายน 2566 พบว่ามี MAPE เฉลี่ย 9 เดือน เท่ากับร้อยละ 8.86 ซึ่งอยู่ในเกณฑ์ที่ใช้พยากรณ์ได้ดี สอดคล้องกับผลการศึกษาของอะซันเตและคณะ (Asante et al., 2022) ที่ได้ทำการศึกษารูปแบบการพยากรณ์จำนวนผู้ป่วยโรคความดันโลหิตสูงและโรคเบาหวาน ผลการวิเคราะห์รูปแบบที่เหมาะสม ARIMA (5, 2, 4) เป็นโมเดลที่เหมาะสมที่สามารถนำไปใช้ในการพยากรณ์ที่จำนวนผู้ป่วยในอนาคตได้ ทั้งนี้เพื่อนำไปใช้ในการวางแผนดูแลผู้ป่วยและการบริหารจัดการ สำหรับการวางแผนกลยุทธ์ที่เหมาะสมและมีประสิทธิภาพ ในประเทศไทยได้มีการนำอนุกรมเวลาด้วยวิธีบ็อกซ์-เจนกินส์ มาใช้ในการทำนายจำนวนผู้ป่วยเบาหวานที่เข้ามารับการรักษาในจังหวัดขอนแก่น ได้แบบจำลองที่เหมาะสม คือ ARIMA (4,0,2) ที่สามารถนำมาทำนายของแนวโน้มจำนวนผู้ป่วยโรคเบาหวานในอนาคตได้ (วรวิทย์ จุลทะทอง และคณะ 2563)

อย่างไรก็ตาม จากผลการศึกษาอนุกรมเวลาที่น่าสนใจในการพยากรณ์จำนวนผู้ป่วยโรคไม่ติดต่อเรื้อรัง (NCDs) ได้แก่ โรคความดันโลหิตสูง โรคเบาหวาน เป็นต้น ยังมีอิทธิพลจากตัวแปรความไม่แน่นอนของปัจจัยภายนอก อาทิเช่น ความหนาแน่นของประชากรในพื้นที่ จำนวนประชากร และพฤติกรรมของประชากร อาจส่งผลต่อค่าความน่าจะเป็นของอัตราจำนวนผู้ป่วยโรคความดันโลหิตสูงรายใหม่ ในจังหวัดปทุมธานี ปัจจัยดังกล่าวเข้ามาเกี่ยวข้องทำให้การพยากรณ์จำนวนผู้ป่วยอาจมีความคลาดเคลื่อนจากความเป็นจริง อย่างไรก็ตาม การผลของการศึกษานี้สามารถนำไปขยายผลการพยากรณ์โรคในกลุ่มไม่ติดต่อเรื้อรังและ ผลข้อมูลที่ได้จะเป็นประโยชน์ต่อผู้บริหารงานทางด้านสาธารณสุข เจ้าหน้าที่สาธารณสุขที่รับผิดชอบในส่วนองงานส่งเสริมสุขภาพในการนำข้อมูลดังกล่าวไปใช้ในการวางแผนงานเพื่อป้องกันโรคความดันโลหิตสูงในระดับของจังหวัด และขยายผลต่อไปในระดับของภูมิภาคโดยการกำหนดมาตรการหรือนโยบายที่เหมาะสมกับสภาพแวดล้อมของพื้นที่ซึ่งจะสามารถช่วยลดจำนวนผู้ป่วยโรคความดันโลหิตสูงได้ในอนาคต

สรุป

การพยากรณ์จำนวนผู้ป่วยโรคความดันโลหิตสูงรายใหม่ล่วงหน้าด้วยอนุกรมเวลาวิธีบ็อกซ์ และเจนกินส์มีความสำคัญในการวางแผนการปฏิบัติงานด้านการสาธารณสุข เพื่อการเตรียมงบประมาณในการดำเนินโครงการรณรงค์เชิงส่งเสริมสุขภาพ ป้องกัน การบำบัดรักษาและฟื้นฟู

ข้อเสนอแนะ

จากผลการศึกษาค้นคว้าครั้งนี้แสดงให้เห็นว่า

การพยากรณ์อนุกรมเวลาใช้เพียงข้อมูลจำนวนผู้ป่วยโรคความดันโลหิตสูงรายใหม่จากฐานข้อมูลกระทรวงสาธารณสุข ปี พ.ศ. 2560-2566 ไม่ได้มีการนำปัจจัยอื่น ๆ เช่น พฤติกรรมการบริโภคอาหาร ความเป็นอยู่ของประชากรในพื้นที่ ตลอดจนผลของการรณรงค์โดยหน่วยงานสาธารณสุขของภาครัฐร่วมด้วยแต่อย่างใด การตรวจสอบผลการพยากรณ์กับข้อมูลจริงจากรายงานของกระทรวงสาธารณสุข เพื่อพัฒนาแบบจำลองที่เหมาะสมต่อไป

ในการศึกษาวิจัยครั้งต่อไปควรเพิ่มข้อมูลปัจจัยที่ครอบคลุมในบริบทของพื้นที่จังหวัดปทุมธานีมาร่วมการวิเคราะห์ด้วย โดยอาจใช้การวิเคราะห์การถดถอยเชิงพหุ เพื่อให้ผลของการพยากรณ์มีความถูกต้องแม่นยำ น่าเชื่อถือมากยิ่งขึ้น

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณ กองยุทธศาสตร์และแผนงาน สำนักงานปลัด กระทรวงสาธารณสุขที่อนุเคราะห์ข้อมูลจำนวนผู้ป่วยโรคความดันโลหิตสูง ในจังหวัดปทุมธานี และคณะสหเวชศาสตร์ มหาวิทยาลัยปทุมธานี ที่เอื้อเฟื้อสถานที่ในการทำวิจัยในครั้งนี้

เอกสารอ้างอิง

- กองโรคไม่ติดต่อ กรมควบคุมโรค กระทรวงสาธารณสุข. (2566). อัตราป่วยรายใหม่ของโรคความดันโลหิตสูงต่อแสนประชากรในปีงบประมาณ จังหวัดปทุมธานี ปีงบประมาณ 2560 ถึง 2566. สืบค้นจาก https://hdcservice.moph.go.th/hdc/reports/report.php?&cat_id=6a1fdf282fd28180eed7d1cfe0155e11&id=29eec762c9591d1f8092da14c7462361
- เขตสุขภาพที่ 4. (2566). รายงานผลการดำเนินงาน ประจำปีงบประมาณ พ.ศ. 2566 เขตสุขภาพที่ 4. สืบค้นจาก https://rh4.moph.go.th/upload_file_section_r4/j11vNzxORx8YtJMDPvp7x9ej0r671e.pdf
- ธันวา เจริญศิริ, ชนาธิป โสภณพิมล, ดนุสรณ์ ธนะपालะ, และวีระชัย ชันทองคำ. (2558). การพยากรณ์จำนวนผู้ป่วยโรคสุกใสในจังหวัดเชียงใหม่ด้วยวิธีบ็อกซ์-เจนกินส์. *บูรพาเวชสาร*, 2(2), 41-49.
- ประกาศมหาวิทยาลัยมหิดล. (2565). เรื่องแนวปฏิบัติสำหรับโครงการวิจัยที่ไม่เข้าข่ายการวิจัยในคน พ.ศ. 2565. สืบค้นจาก <https://sp.mahidol.ac.th/th/LAW/policy/2565-MU-Non-Human.pdf>
- สมาคมความดันโลหิตสูงแห่งประเทศไทย. (2562). แนวทางการรักษาโรคความดันโลหิตสูงในเวชปฏิบัติทั่วไป พ.ศ. 2562. กรุงเทพฯ: ทรिकิงค์.

- ศูนย์ส่งเสริมจริยธรรมการวิจัย มหาวิทยาลัยมหิดล. (2565). **Self-Assessment form whether an activity is human subject research which requires ethical approval**. สืบค้นจาก <https://sp.mahidol.ac.th/th/ethics-human/forms/checklist/2022-Human%20Research%20Checklist-research.pdf>.
- วางแผน กิ่งวิบูลย์. (2559). ตัวแบบการพยากรณ์จำนวนผู้ป่วยโรคปอดอักเสบในประเทศไทย. **วารสารสาธารณสุขมหาวิทยาลัยบูรพา**, 11(1), 24-38.
- วางแผน เรียงสุทธิ. (2563). การพยากรณ์จำนวนผู้ป่วยโรคเฝ้าระวังในประเทศไทย. **Thai Science and Technology Journal**, 28(1), 1-13.
- วรวิทย์ จุลทะทอง, ถิรวัฒน์ นาคคนชม, ประภาวรรณ เสนาเพ็ง, และปิยภัทร บุชบาบดินทร์. (2563). **แบบจำลองอัตราผู้ป่วยโรคเบาหวานที่เข้ารับการรักษาในจังหวัดขอนแก่น**. The National Undergraduate Conference on Statistics 2019, 14 May 2019, Chiang Mai University, Chiang Mai, Thailand, 11-35.
- วิชัย เอกพลากร. (2564). **รายงานการสำรวจสุขภาพประชาชนไทยโดยการตรวจร่างกายครั้งที่ 6 พ.ศ. 2562 - 2563**. คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี มหาวิทยาลัยมหิดล. กรุงเทพฯ: สำนักพิมพ์อักษรกราฟิกแอนดดีไซน์.
- วัฒนา ชยธวัช. (2563). **จาโมวี - โปรแกรมวิเคราะห์ทางสถิติสำเร็จรูปใช้ได้เสรี**. **วารสารกองทุนพัฒนาสื่อปลอดภัยและสร้างสรรค์**, 2(3), 97-122.
- Andrés, D. (2023). **Machine learning pills: Error metrics for time series forecasting**. Retrieved from <https://mlpills.dev/time-series/error-metrics-for-time-series-forecasting/>
- Asante, D.O., Walker, A.N., Seidu, T.A., Kpogo, S.A., & Zou, J. (2022). Hypertension and diabetes in Akasi South District, Ghana: Modeling and forecasting. **Biomed Res Int**, 2022
- Box, G.E.P., & Jenkins, G.M. (1976). **Time series analysis, forecasting and control**. San Francisco: Holden-Day.
- Hyndman, R.J., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., ... & Ysmeen, F. (2024). **Forecast: Forecasting functions for time series and linear models (Version 8.22.0) [R package]**. Retrieved from <https://pkg.robjhyndman.com/forecast>
- Hyndman, R.J., & Athanasopoulos, G. (2018). **Forecasting: principles and practice (2nd ed.)**. Retrieved from <https://otexts.com/fpp2/>
- Hyndman, R.J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. **Journal of Statistical Software**, 27(3), 1-23.

- Lartey, H., Wang, J., Lartey, P.L., Agyei, J., & Aboagye, J.S. (2020). Modeling and prediction of hypertension in Komfo Anokye Teaching Hospital (KATH), Ghana. *Global Research Journal of Public Health and Epidemiology*, 8(6), 13-21.
- Lewis, C.D. (1982). **Industrial and business forecasting methods**. London: Butterworths.
- Madsen, H. (2008). **Time series analysis**. London: Chapman & Hall/CRC.
- Makridakis S., A. Anderson, R.Fields, MD. Jibon, R. Lewandowski, J. Newton, E. Parzen, & R. Winkler. (1982). The accuracy of extrapolation (time-series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1(2), 111-153.
- MedlinePlus. (2020). **High blood pressure**. Retrieved from <https://medlineplus.gov/highbloodpressure.html#summary>.
- Meese, R., & Geweke, J. (1982). **A comparison of autoregressive univariate forecasting procedures for macroeconomic time series**. Unpublished manuscript, University of California, Berkeley.
- Mekpariyup, J., & Saithanu, K. (2016). A new approach to detect epidemic of dengue hemorrhagic fever by combining ARIMA model and adjusted Tukey's control chart with interpretation rules. *Interventional Medicine and Applied Science*, 8(3), 118-120.
- Osarumwense, O.-I. (2013). Applicability of Box-Jenkins SARIMA model in rainfall forecasting: A case study of Port-Harcourt, South-south Nigeria. Canadian. *Journal on Computing in Mathematics, Natural Sciences, Engineering and Medicine*, 4(1), 1-4.
- Talirongan, F.J.B., Talirongan, H., & Orong, M.Y. (2020). Modeling national trends on health in the Philippines using ARIMA. *Journal of Health & Medical Informatics*, 11(1), 1-6. <https://doi.org/10.37421/jhmi.2020.11.342>
- The jamovi project. (2022). **jamovi (Version 2.3)[Computer software]**. Retrieved from <https://www.jamovi.org>
- R Core Team. (2021). **R: A language and environment for statistical computing (Version 4.1) [Computer software]**. Retrieved from <https://cran.r-project.org>
- Seol, H. (2023). **snowCluster: Multivariate analysis (Version 7.1.7) [jamovi module]**. Retrieved from <https://github.com/hyunsooseol/snowCluster>
- Tseng, Y.-J., & Shih, Y.-L. (2020). Developing epidemic forecasting models to assist disease surveillance for influenza with electronic health records. *International Journal of Computers and Applications*, 42(6), 616-621.