

# แบบจำลองค่าสุดขีดของอัตราการป่วยโรคไข้หวัดใหญ่ ในภาคตะวันออกเฉียงเหนือของประเทศไทย

## Extreme Value Modeling for the Rate of Influenza Patients in the Northeast of Thailand

วิชุดา เห็นเจริญ\* และ ปิยภัทร บุษบาบดินทร์

Wichuda Hencharoen\* and Piyapatr Busababodhin

หน่วยวิจัยสถิติและสถิติประยุกต์ ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยมหาสารคาม

Statistics and Applied Statistics Research Unit, Mathematics Department, Faculty of Science, Mahasarakham University

Received : 26 May 2017

Accepted : 24 July 2017

Published online : 15 September 2017

### บทคัดย่อ

ไข้หวัดใหญ่เป็นโรคติดต่อที่มีการแพร่กระจายอยู่ทั่วทั้งประเทศไทย ซึ่งสถานการณ์การเจ็บป่วยและการเสียชีวิตด้วยโรคไข้หวัดใหญ่นับว่าเป็นปัญหาสำคัญในงานสาธารณสุขของประเทศไทย โดยเฉพาะอย่างยิ่งอัตราการป่วยโรคไข้หวัดใหญ่ในภาคตะวันออกเฉียงเหนือของประเทศไทยได้เพิ่มขึ้นอย่างต่อเนื่อง ผู้วิจัยมีความสนใจทฤษฎีค่าสุดขีด (Extreme Value Theory: EVT) มาใช้วิเคราะห์หารูปแบบที่เหมาะสมภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป (Generalized Extreme Value Distribution: GEVD) ในกระบวนการไม่คงที่ กล่าวคือเมื่อพารามิเตอร์ระดับตำแหน่ง ( $\mu$ ) และพารามิเตอร์ระดับขนาด ( $\sigma$ ) มีการเปลี่ยนแปลงโดยขึ้นอยู่กับเวลา สำหรับงานวิจัยนี้ผู้วิจัยได้คาดการณ์ความน่าจะเป็นของการระบาดของโรคไข้หวัดใหญ่ในภาคตะวันออกเฉียงเหนือของประเทศไทย โดยใช้ข้อมูลจำนวนผู้ป่วยโรคไข้หวัดใหญ่สะสมสูงสุดรายสัปดาห์ทั้งสิ้น 20 จังหวัดในภาคตะวันออกเฉียงเหนือของประเทศไทย ตั้งแต่เดือนมกราคม พ.ศ. 2546 ถึงเดือนธันวาคม พ.ศ. 2558 จากสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข มาวิเคราะห์ข้อมูลด้วยแพ็คเกจ "extRemes" ในโปรแกรม R ผลปรากฏว่า ในภาคตะวันออกเฉียงเหนือของประเทศไทยมีจังหวัดที่ควรมีการเฝ้าระวังอัตราการป่วยโรคไข้หวัดใหญ่ในอนาคต ได้แก่ จังหวัดนครราชสีมา จังหวัดนครพนม และจังหวัดบึงกาฬ เนื่องจากมีค่าประมาณระดับการเกิดซ้ำ (Return level) ในรอบการเกิดซ้ำ 5 ปี 10 ปี และ 20 ปี สูงที่สุดเป็นสามอันดับแรก ซึ่งผลของการวิเคราะห์นี้ในครั้งนี้อาจเป็นแนวทางนำไปสู่การวางแผนงานทางด้านสาธารณสุขเพื่อรับมือกับเหตุการณ์ของโรคไข้หวัดใหญ่ในแต่ละกลุ่มจังหวัดในภาคตะวันออกเฉียงเหนือของประเทศไทยต่อไป

**คำสำคัญ :** ทฤษฎีค่าสุดขีด การแจกแจงค่าสุดขีดวงนัยทั่วไป กระบวนการไม่คงที่ ไข้หวัดใหญ่

\*Corresponding author. E-mail : wichuda.hencharoen@gmail.com

## Abstract

Influenza is a contagious disease with high transmissibility to spread around Thailand. Recently, the considerable morbidity and mortality of Influenza presents an enormous burden on Thailand public health. Especially, the rate of influenza infective in the Northeast of Thailand has been continually increased. Therefore, we develop statistical models using Extreme Value Theory (EVT); the Generalized Extreme Value Distribution (GEVD) in non-stationary process, where is location parameter ( $\mu$ ) and scale parameter ( $\sigma$ ) depending on the time. We forecast to the probability of outbreak of highly pathogenic influenza in the Northeast of Thailand by using the incidence data of highly pathogenic influenza of 20 provinces in the Northeast of Thailand from January 2003 to December 2015 which are collected from the Data center bureau of epidemiology department of disease control ministry of public health. The package "extRemes" in R programming is used to analyze data and model. The results show that in the Northeast of Thailand has three provinces that should be monitored for influenza in the future includes Nakhon Ratchasima, Nakhon Phanom, and Bung Kan Because there is the return level estimate in 5 years, 10 years, and 20 years has highest is the first three. The results of this analysis can be applied to public health planning to cope with influenza events in each of provinces in the Northeast of Thailand.

**Keywords :** extreme value theory, generalized extreme value distribution, non-stationary process, influenza

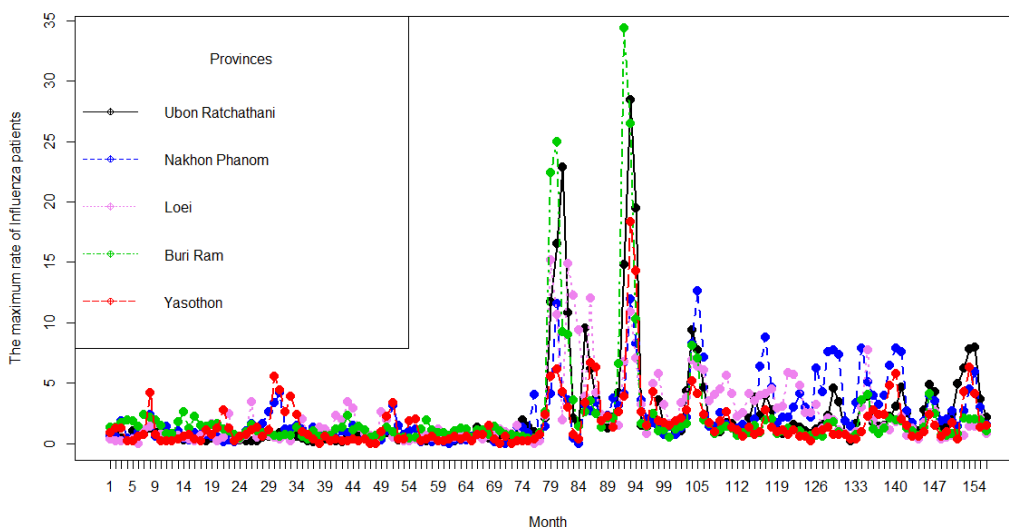
## บทนำ

ในแต่ละปีมีการประมาณว่ามีผู้ติดเชื้อไข้หวัดใหญ่ทั่วโลกสูงถึงร้อยละ 10-15 ของประชากรทั้งหมด การที่ไวรัสมีการแพร่ระบาดได้อย่างกว้างขวางนี้เนื่องมาจากลักษณะที่สำคัญของไวรัสชนิดนี้ที่มีการเปลี่ยนแปลงลักษณะของโปรตีนสำคัญที่เป็นเป้าหมายในการโจมตีของระบบภูมิคุ้มกันของร่างกายได้ง่าย เมื่อลักษณะของโปรตีนดังกล่าวเปลี่ยนแปลงไปก็ทำให้ภูมิคุ้มกันในร่างกายของผู้ที่เคยติดเชื้อไม่สามารถป้องกันการติดเชื้อของไวรัสสายพันธุ์ใหม่ได้จึงทำให้สามารถติดเชื้อซ้ำได้อีก ดังนั้นเมื่อไวรัสมีการเปลี่ยนแปลงไปครั้งหนึ่งก็มักจะมีการระบาดตามมาเพราะไม่มีภูมิคุ้มกันต่อไวรัสสายพันธุ์ใหม่อยู่เลย จึงทำให้มีไวรัสสายพันธุ์ต่างๆอยู่มากมาย โดยไวรัสที่กำลังแพร่ระบาดอยู่ในยุโรปและอเมริกาส่วนใหญ่เป็น Influenza A(H3N2) แต่ก็มีรายงานพบ Influenza A(H1N1) ปนอยู่บ้างในบางประเทศรวมทั้ง Influenza B ที่มีลักษณะคล้าย B/Yamanashi/166/98 (Ueawaragun, 2010)

สำหรับสถานการณ์ของโรคไข้หวัดใหญ่ในประเทศไทย ผู้วิจัยได้ทำการสืบค้นจำนวนผู้ป่วยและอัตราการป่วยโรคไข้หวัดใหญ่จากระบบรายงานการเฝ้าระวังโรค 506 ของสำนักโรคติดต่อวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข พบว่าจาก 77 จังหวัดของประเทศไทยในปี พ.ศ. 2559 พบผู้ป่วยโรคไข้หวัดใหญ่จำนวน 169,362 ราย คิดเป็นอัตราป่วย 258.86 ต่อแสนประชากร สำหรับในปี พ.ศ. 2558 พบผู้ป่วยโรคไข้หวัดใหญ่จำนวน 77,926 ราย คิดเป็นอัตราป่วย 119.66 ต่อแสนประชากร สำหรับในปี พ.ศ. 2557 พบผู้ป่วยโรคไข้หวัดใหญ่จำนวน 74,134 ราย คิดเป็นอัตราป่วย 114.13 ต่อแสนประชากร และในปี พ.ศ. 2556 พบผู้ป่วยโรคไข้หวัดใหญ่จำนวน 43,866 ราย คิดเป็นอัตราป่วย 68.46 ต่อแสนประชากร ซึ่งจากข้อมูลดังกล่าว

แสดงให้เห็นว่าจำนวนของผู้ป่วยโรคไข้หวัดใหญ่ในปี พ.ศ. 2559 ได้เพิ่มสูงขึ้นกว่า 2 เท่าจากปี พ.ศ. 2557-2558 และได้เพิ่มสูงขึ้นกว่า 3.5 เท่าจากปี พ.ศ. 2556

สำหรับสถานการณ์ของโรคไข้หวัดใหญ่ในแต่ละภูมิภาคของประเทศไทยนั้น พบว่า อัตราการป่วยโรคไข้หวัดใหญ่ในภาคตะวันออกเฉียงเหนือของประเทศไทยได้เพิ่มสูงขึ้นอย่างต่อเนื่องในทุกปี โดยผู้วิจัยได้นำข้อมูลจำนวนผู้ป่วยโรคไข้หวัดใหญ่สะสมสูงสุดรายสัปดาห์ จากสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข ทั้งสิ้น 20 จังหวัดในภาคตะวันออกเฉียงเหนือของประเทศไทย ระหว่างปี พ.ศ. 2546-2558 ทั้งหมด 13 ปี มาทำการหาจำนวนผู้ป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือน และนำข้อมูลที่ได้ในข้างต้นมาหาอัตราการป่วย (Morbidity rate) ต่อแสนประชากร ผลที่ได้แสดงให้เห็นว่า จังหวัดที่มีผลรวมของอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือน 5 อันดับแรกในภาคตะวันออกเฉียงเหนือของประเทศไทย ได้แก่ จังหวัดอุบลราชธานี จังหวัดนครพนม จังหวัดเลย จังหวัดบุรีรัมย์ และจังหวัดยโสธร ตามลำดับ ซึ่งตั้งแต่เดือนกรกฎาคม ปี พ.ศ. 2552 (ตำแหน่ง Index ที่ 79) จนถึงปัจจุบันพบอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนเพิ่มสูงขึ้นอย่างต่อเนื่อง แสดงดังภาพที่ 1



**ภาพที่ 1** อัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือน พ.ศ. 2546-2558 ในจังหวัดที่มีผลรวมของอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือน 5 อันดับแรกในภาคตะวันออกเฉียงเหนือของประเทศไทย

จากสถานการณ์ของโรคไข้หวัดใหญ่นี้แสดงให้เห็นว่า การคาดการณ์การเกิดโรคและภัยสุขภาพที่จะเกิดขึ้นในอนาคตมีความสำคัญอย่างยิ่ง ซึ่งผลของการคาดการณ์ทำให้เกิดการเฝ้าระวังและทราบถึงโอกาสของการเกิดโรคในอนาคต นำไปสู่การวางแผนเตรียมการป้องกันและการรักษาได้อย่างมีประสิทธิภาพ ดังนั้นในการใช้แบบจำลองที่ไม่เหมาะสมกับข้อมูลจะส่งผลให้การพยากรณ์มีความคลาดเคลื่อนสูง ทำให้การตัดสินใจในการวางแผนไม่สอดคล้องกับสถานการณ์ที่เกิดขึ้นในอนาคต จากการทบทวนวรรณกรรมที่เกี่ยวข้องได้มีนักวิจัยนำเสนอวิธีการที่หลากหลาย ซึ่งให้ประสิทธิภาพและผลการพยากรณ์ที่แตกต่างกัน อาทิเช่น Suttachana *et al.*, (2012) ได้พยากรณ์การเกิดโรคไข้หวัดใหญ่ ประเทศไทย ปี พ.ศ. 2555 โดยใช้ข้อมูลจำนวนผู้ป่วยไข้หวัดใหญ่จำแนกรายเดือนตั้งแต่วันที่ 1 มกราคม พ.ศ. 2553 ถึง 31 กรกฎาคม พ.ศ. 2555 โดย

วิธีการปรับเรียบแบบเอ็กซ์โปเนนเชียล Winters' multiplicative สำหรับ Seesomboon *et al.*, (2013) ได้พยากรณ์การเกิดโรคไข้หวัดใหญ่ พื้นที่ 8 จังหวัดภาคเหนือตอนบน ปี พ.ศ. 2557 โดยใช้ข้อมูลโรคไข้หวัดใหญ่ในพื้นที่ 8 จังหวัดภาคเหนือตอนบน จากฐานข้อมูลรายงาน 506 ปี พ.ศ. 2552 – 2556 โดยใช้วิธี Holt-Winters additive seasonal และ Chanwirat *et al.*, (2015) ได้พยากรณ์โรคและภัยสุขภาพ กรณีโรคไข้หวัดใหญ่ในประเทศไทย ปี พ.ศ. 2559 โดยศึกษาขนาดวิทยาของการเกิดโรค การกระจายของโรคไข้หวัดใหญ่ในประเทศไทย ระหว่างปี พ.ศ. 2553-2558 ด้วยวิธีการวิเคราะห์อนุกรมเวลาโดยใช้เทคนิคการปรับเรียบแบบเอ็กซ์โปเนนเชียล

ความแม่นยำในการคาดการณ์ และการทราบถึงลักษณะของข้อมูลทางด้านระบาดวิทยา เป็นสิ่งสำคัญอย่างมากในการอธิบายการระบาดของโรค จากการศึกษางานวิจัยที่เกี่ยวข้องที่ใช้วิธีการพยากรณ์ข้อมูลอนุกรมเวลากับข้อมูลทางด้านระบาดวิทยา มักจะพบปัญหาความคลาดเคลื่อนไม่เป็นไปตามข้อตกลงเบื้องต้นในการใช้งาน ซึ่งเมื่อพบข้อมูลที่มีจำนวนผู้ป่วยและผู้เสียชีวิตที่สูงมาก ๆ นั้น ข้อมูลในลักษณะดังกล่าวถูกมองว่าเป็นข้อมูลผิดปกติ เมื่อนำมาวิเคราะห์จึงถูกตัดออก แต่ข้อมูลจำนวนผู้ป่วยสูงสุดนั้นเป็นข้อมูลสำคัญอย่างมาก ในการวางแผนเตรียมความพร้อมของทรัพยากรเพื่อป้องกันและรับมือกับการรักษาได้อย่างมีประสิทธิภาพสูงสุด ดังนั้นการนำทฤษฎีค่าสุดขีด (Extreme value theory: EVT) มาประยุกต์ใช้กับข้อมูลลักษณะดังกล่าวด้วยการแจกแจงค่าสุดขีดวงนัยทั่วไป (Generalized extreme value distribution: GEVD) จะสามารถคาดการณ์อัตราการป่วยสูงสุดและอัตราการเสียชีวิตสูงสุด จากตัวอย่างงานวิจัย อาทิเช่น การศึกษาของ Chen *et al.*, (2015) ได้สร้างแบบจำลองการแจกแจงค่าสุดขีดวงนัยทั่วไป (GEVD) และการแจกแจงพาราโตวงนัยทั่วไป (GPD) ในการคาดการณ์ความรุนแรงของสถานการณ์การเกิดโรคไข้หวัดใหญ่ในจังหวัดเจ้อเจียง ประเทศจีน และการศึกษาของ Thomas *et al.*, (2016) ได้เสนอการแจกแจงค่าสุดขีดวงนัยทั่วไป ในการคาดการณ์จำนวนผู้เสียชีวิตสูงสุดด้วยโรคไข้หวัดใหญ่และปอดบวมในประเทศฝรั่งเศส ซึ่งวิธีการนี้สามารถบอกได้ว่าในอีก 30 ปีข้างหน้าความน่าจะเป็นที่จะมีผู้เสียชีวิตเกิน 12 ต่อแสนประชากรอยู่ที่ร้อยละ 3 และในอีก 90 ปีจะมีผู้เสียชีวิตเกิน 14 ต่อแสนประชากรอยู่ที่ร้อยละ 1.1

จากความสำคัญข้างต้น ทำให้ผู้วิจัยตระหนักได้ว่าการเกิดโรคไข้หวัดใหญ่เป็นปัญหาทางสาธารณสุขที่สำคัญ และการคาดการณ์โดยเลือกใช้ตัวแบบที่เหมาะสมจะทำให้เกิดการเฝ้าระวังและทราบถึงโอกาสของการเกิดโรคในอนาคตได้อย่างแม่นยำ ดังนั้นผู้วิจัยจึงสนใจศึกษาการวิเคราะห์ทางสถิติสำหรับค่าสุดขีดด้วยการประยุกต์ใช้ทฤษฎีค่าสุดขีด (EVT) นำเสนอการแจกแจงค่าสุดขีดวงนัยทั่วไป (GEVD) ในการสร้างแบบจำลองเพื่อใช้คาดการณ์อัตราการป่วยโรคไข้หวัดใหญ่ในรอบการเกิดซ้ำ 5 ปี 10 ปี และ 20 ปี กรณีศึกษาอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนในภาคตะวันออกเฉียงเหนือของประเทศไทย

### วิธีดำเนินการวิจัย

ในงานวิจัยนี้ใช้ข้อมูลทุติยภูมิ (Secondary data) เป็นข้อมูลจำนวนผู้ป่วยโรคไข้หวัดใหญ่สะสมสูงสุดรายสัปดาห์จากสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข ทั้งสิ้น 20 จังหวัดในภาคตะวันออกเฉียงเหนือ ระหว่างปี พ.ศ. 2546-2558 ทั้งหมด 13 ปี โดยพื้นที่ในการศึกษาประกอบด้วย จังหวัดมหาสารคาม จังหวัดร้อยเอ็ด จังหวัดขอนแก่น จังหวัดกาฬสินธุ์ จังหวัดนครราชสีมา จังหวัดบุรีรัมย์ จังหวัดสุรินทร์ จังหวัดศรีสะเกษ จังหวัดอุบลราชธานี จังหวัดอำนาจเจริญ จังหวัดมุกดาหาร จังหวัดยโสธร จังหวัดนครพนม จังหวัดสกลนคร จังหวัดหนองคาย จังหวัดบึงกาฬ จังหวัดเลย จังหวัดชัยภูมิ จังหวัดหนองบัวลำภู และจังหวัดอุดรธานี ซึ่งข้อมูลดังกล่าวจะนำมาหาจำนวนผู้ป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือน

ของแต่ละจังหวัด และนำข้อมูลที่ได้ในข้างต้นมาหาอัตราการป่วย (Morbidity rate) ต่อแสนประชากร จากนั้นจึงนำมาวิเคราะห์หาแบบจำลองค่าสุดขีดวงนัยทั่วไปโดยใช้แพ็คเกจ "extRemes" ในโปรแกรม R

**1. ทฤษฎีและแนวคิดของค่าสุดขีด**

1.1 ทฤษฎีบทขีดจำกัดส่วนกลาง (Central limit theory)

ถ้า  $\bar{X}$  เป็นค่าเฉลี่ยของตัวอย่างสุ่ม  $X_1, X_2, \dots, X_n$  ที่มีขนาด  $n$  และมาจากการแจกแจงใดๆ ซึ่งมีค่าเฉลี่ยเป็น  $\mu$  และมีความแปรปรวนเป็น  $\sigma^2$  แล้วการแจกแจงของ  $\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$  จะลู่เข้าสู่การแจกแจงปกติมาตรฐาน  $(N(0,1))$  เมื่อขนาดตัวอย่างสุ่มมีขนาดใหญ่  $(n \rightarrow \infty)$

1.2 ทฤษฎีค่าสุดขีด (Extreme value theory)

การศึกษาทฤษฎีค่าสุดขีดจะใช้ค่าสูงสุดของตัวอย่างสุ่ม โดยกำหนดให้  $M_n$  แทนค่าสูงสุด (Coles & Nadaraja, 2001) แทนด้วย

$$M_n = \max(X_1, X_2, \dots, X_n) \tag{1}$$

โดยสามารถหาค่าความน่าจะเป็นและฟังก์ชันการแจกแจงสะสมของ  $M_n$  ได้ดังนี้

$$\begin{aligned} \Pr(M_n \leq z) &= \Pr(X_1 \leq z, \dots, X_n \leq z) \\ &= \Pr(X_1 \leq z) \cdot \dots \cdot \Pr(X_n \leq z) = \{F(z)\}^n \end{aligned} \tag{2}$$

ซึ่งในความเป็นจริงแล้วฟังก์ชัน  $F(z)$  เป็นฟังก์ชันที่ไม่ทราบค่าที่แท้จริงได้ แต่ในทางปฏิบัติพยายามใช้เทคนิคพื้นฐานในการประมาณค่าจากข้อมูลที่เก็บรวบรวมได้เพื่อนำไปสู่การประมาณค่าฟังก์ชันความหนาแน่นของความน่าจะเป็นของ  $M_n$  ในสมการข้างต้น ซึ่งยากที่จะทำให้ฟังก์ชันที่ประมาณได้เข้าใกล้ฟังก์ชันที่แท้จริง ดังนั้นแนวทางที่สามารถประมาณฟังก์ชันที่แท้จริงได้คือวิธีเทียบเคียงกับทฤษฎีบทขีดจำกัดส่วนกลางกล่าวคือ พิจารณาการหาแจกแจงของ

$$\frac{M_n - b_n}{a_n} \text{ เมื่อ } a_n \text{ และ } b_n \text{ เป็นค่าคงที่ ซึ่งจะนำไปสู่การหาการแจกแจงของ } \left\{ F\left(\frac{M_n - b_n}{a_n}\right) \right\}^n \text{ พบว่าเมื่อ } (n \rightarrow \infty)$$

ไม่ได้ลู่เข้าสู่การแจกแจงหนึ่งเหมือนกรณีทฤษฎีบทขีดจำกัดส่วนกลางที่  $\frac{\bar{X} - \mu_n}{\sigma_n} \rightarrow N(0,1)$  แต่

$$\left\{ F\left(\frac{M_n - b_n}{a_n}\right) \right\}^n \rightarrow G(z) \text{ เมื่อ } a_n > 0 \text{ และ } b_n > 0 \text{ เป็นค่าคงที่ โดยที่ } G(z) \text{ ไม่ขึ้นอยู่กับขนาดของตัวอย่างสุ่ม } (n)$$

นำไปสู่ทฤษฎีบทของรูปแบบค่าสุดขีด (External types theorem)

## 2. การแจกแจงค่าสุดขีดวงนัยทั่วไป

การแจกแจงค่าสุดขีดวงนัยทั่วไป (Generalized extreme value distribution: GEVD) เหมาะสำหรับวิเคราะห์ค่าสุดขีดในช่วงคาบเวลาที่น่าสนใจ เช่น รายปี รายเดือน รายไตรมาส หรือรายสัปดาห์ เป็นต้น โดยจะเลือกข้อมูลที่สูงสุดในแต่ละช่วงคาบเวลาที่ผู้วิเคราะห์สนใจ

ถ้าให้  $x$  เป็นตัวอย่างสุ่มที่มีการแจกแจงค่าสุดขีดวงนัยทั่วไปที่เขียนแทนด้วย  $X \sim GEVD(\mu, \sigma, \xi)$  ที่มีฟังก์ชันการแจกแจงสะสม ฟังก์ชันความหนาแน่นของความน่าจะเป็น และฟังก์ชันควอนไทล์ เป็นดังนี้ (Coles & Nadaraja, 2001)

ฟังก์ชันการแจกแจงสะสม

$$F(x) = \exp \left\{ - \left( 1 + \xi \frac{x - \mu}{\sigma} \right)^{-1/\xi} \right\} \quad (3)$$

ฟังก์ชันความหนาแน่นของความน่าจะเป็น

$$f(x) = \frac{1}{\sigma} \left[ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right]^{(-1/\xi)-1} \exp \left\{ - \left( 1 + \xi \frac{x - \mu}{\sigma} \right)^{-1/\xi} \right\} \quad (4)$$

เมื่อ  $1 + \xi \left( \frac{x - \mu}{\sigma} \right) > 0$  โดยที่  $\mu$  แทน พารามิเตอร์ระบุตำแหน่ง (Location parameter)  $\sigma$  แทน พารามิเตอร์ระบุขนาด (Scale parameter) และ  $\xi$  แทนพารามิเตอร์ระบุรูปร่าง (Shape parameter)

และฟังก์ชันควอนไทล์

$$x(F) = \begin{cases} \mu - \frac{\sigma}{\xi} \left[ 1 - \{-\log(1-p)\}^{-\xi} \right] & , \quad \xi \neq 0 \\ \mu - \sigma \log \{-\log(1-p)\} & , \quad \xi = 0 \end{cases} \quad (5)$$

ซึ่ง GEVD ได้แบ่งตามพารามิเตอร์ระบุรูปร่าง (Shape:  $\xi$ ) ได้แก่ กรณี  $\xi = 0$  เรียกว่า การแจกแจงกัมเบล (Gumbel distribution) สำหรับกรณี  $\xi > 0$  เรียกว่า การแจกแจงฟร็ีเชท (Fréchet distribution) และกรณี  $\xi < 0$  เรียกว่า การแจกแจงไวบูลล์ (Weibull distribution) ทั้งนี้ต้องพิจารณาช่วงความเชื่อมั่นของพารามิเตอร์ด้วย

### 3. รูปแบบที่เหมาะสมภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป

หาการแจกแจงและรูปแบบที่เหมาะสมภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป โดยพิจารณาภายใต้เงื่อนไขดังนี้

#### 3.1 ภายใต้กระบวนการคงที่ (Stationary process)

รูปแบบ 1:  $\mu$ ,  $\sigma$  และ  $\xi$  เป็นค่าคงที่

#### 3.2 ภายใต้กระบวนการไม่คงที่ (Non-Stationary process)

รูปแบบ 2:  $\mu(t) = \beta_0 + \beta_1 t$  โดยที่  $\sigma$  และ  $\xi$  เป็นค่าคงที่

รูปแบบ 3:  $\sigma(t) = \alpha_0 + \alpha_1 t$  โดยที่  $\mu$  และ  $\xi$  เป็นค่าคงที่

รูปแบบ 4:  $\sigma(t) = \exp(\alpha_0 + \alpha_1 t)$  โดยที่  $\mu$  และ  $\xi$  เป็นค่าคงที่

รูปแบบ 5:  $\mu(t) = \beta_0 + \beta_1 t$ ,  $\sigma(t) = \exp(\alpha_0 + \alpha_1 t)$  โดยที่  $\xi$  เป็นค่าคงที่

### 4. การวิเคราะห์ความแตกต่างของรูปแบบ

วิเคราะห์ความแตกต่างของรูปแบบโดยใช้การทดสอบอัตราส่วนภาวะน่าจะเป็น (Likelihood ratio test) ซึ่งในการประมาณค่าพารามิเตอร์หรือฟังก์ชันของพารามิเตอร์จะใช้วิธีภาวะน่าจะเป็นสูงสุด (Maximum likelihood method) ส่วนในการทดสอบสมมติฐานจะใช้ทฤษฎีบทประกอบของเนย์แมนและเพียร์สัน (Neyman – Pearson lemma) ในการสร้างการทดสอบ (Sinsomboonthong, 2015)

ให้  $X_1, \dots, X_n$  เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น  $f(x; \theta)$  โดยที่  $\theta \in \Omega$  และให้  $\omega$  เป็นเซตย่อยของปริภูมิพารามิเตอร์  $\Omega$

ในการทดสอบสมมติฐานหลัก  $H_0: \theta \in \omega$  เทียบกับสมมติฐานรอง  $H_1: \theta \in \Omega - \omega$  ให้  $L(\theta; x_1, \dots, x_n)$  เป็นฟังก์ชันภาวะน่าจะเป็น (Likelihood function) ของตัวอย่างสุ่ม  $X_1, \dots, X_n$

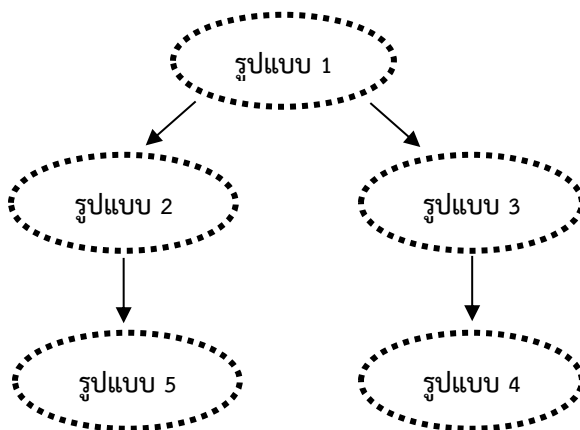
หาอัตราส่วนภาวะน่าจะเป็นทั่วไป ได้ดังนี้

$$\lambda(x_1, \dots, x_n) = \frac{\sup_{\theta \in \omega} L(\theta; x_1, \dots, x_n)}{\sup_{\theta \in \Omega} L(\theta; x_1, \dots, x_n)} \quad (6)$$

โดยที่  $\sup_{\theta \in \omega} L(\theta; x_1, \dots, x_n)$  คือ ค่าสูงสุดของฟังก์ชันภาวะน่าจะเป็น  $L(\theta; x_1, \dots, x_n)$  เมื่อ  $\theta \in \omega$

$\sup_{\theta \in \Omega} L(\theta; x_1, \dots, x_n)$  คือ ค่าสูงสุดของฟังก์ชันภาวะน่าจะเป็น  $L(\theta; x_1, \dots, x_n)$  เมื่อ  $\theta \in \Omega$

โดยการทดสอบอัตราส่วนภาวะน่าจะเป็น เพื่อหารูปแบบที่ดีที่สุดภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป มีขั้นตอนการเปรียบเทียบรูปแบบรายคู่ตามลำดับดังภาพที่ 2



ภาพที่ 2 ขั้นตอนการเปรียบเทียบรูปแบบรายคู่ตามลำดับ

## 5. การตรวจสอบความเหมาะสมของรูปแบบ

การศึกษาค้นคว้าได้ใช้เทคนิคทางกราฟ 2 ชนิด ในการพิจารณาความเหมาะสมของรูปแบบ

### 5.1 Quantile plot

กราฟนี้เป็นการเปรียบเทียบค่าควอนไทล์ (Quantile) ของแบบจำลองกับค่าของข้อมูลจริง หมายถึงการพล็อตระหว่างจุดควอนไทล์ของค่าคาดหวังของการแจกแจงที่ได้จากการสร้างแบบจำลอง กับจุดควอนไทล์ของข้อมูลจริงที่ต้องการทดสอบว่าการแจกแจงมีความเหมาะสมหรือไม่ (Busababodin, 2016)

$$\left\{ \left( \hat{F}^{-1} \left( \frac{i}{n+1} \right), x_{(i)} \right) \right\}; i = 1, \dots, n \quad (7)$$

ซึ่งการประมาณฟังก์ชันการแจกแจง  $\hat{F}$  ถ้าข้อมูลที่ศึกษามีการแจกแจงที่เหมาะสม Quantile plot จะอยู่ใกล้เส้นทแยงมุม (กราฟจะเป็นเส้นตรง)

### 5.2 Density plot

กราฟนี้มีไว้เพื่อพิจารณาว่าแบบจำลองของค่าสุดขีดวางนัยทั่วไปเหมาะสมกับชุดข้อมูลจริงหรือไม่ สร้างกราฟโดยพล็อตค่าข้อมูลจริง กับค่าความหนาแน่นของความน่าจะเป็นของแบบจำลอง ซึ่งถ้าเส้นทั้งสองมีลักษณะคล้ายกัน แสดงว่าแบบจำลองของค่าสุดขีดวางนัยทั่วไปมีความเหมาะสมกับชุดข้อมูลจริง



## 6. ระดับการเกิดซ้ำ

ระดับการเกิดซ้ำ (Return level) ค่าของข้อมูล ณ ตำแหน่งความน่าจะเป็นที่สนใจ เมื่อกำหนดให้  $p$  คือความน่าจะเป็นของเหตุการณ์ที่  $Z > \hat{Z}_p$  อาจกล่าวว่า ระดับการเกิดซ้ำ คือค่าคาดหวังที่จะเกิดเหตุการณ์  $Z > \hat{Z}_p$  โดยเฉลี่ย 1 ครั้งในทุกๆ  $T$  ปี เมื่อ  $T$  คือรอบปีการเกิดซ้ำที่มีความสัมพันธ์กับความน่าจะเป็นของเหตุการณ์โดยที่  $T = \frac{1}{p}$

$$\hat{Z}_T = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left\{ 1 - \left[ -\log \left( 1 - \frac{1}{T} \right) \right]^{-\hat{\xi}} \right\} \quad (8)$$

เมื่อ  $\hat{\mu}$  แทน ค่าประมาณของพารามิเตอร์ระบุตำแหน่ง (Location parameter)

$\hat{\sigma}$  แทน ค่าประมาณของพารามิเตอร์ระบุขนาด (Scale parameter)

$\hat{\xi}$  แทน ค่าประมาณของพารามิเตอร์ระบุรูปร่าง (Shape parameter)

$T$  แทน รอบปีการเกิดซ้ำ (Return period)

## ผลการวิจัยและวิจารณ์ผล

ผู้วิจัยได้นำข้อมูลจำนวนผู้ป่วยโรคไข้หวัดใหญ่สะสมสูงสุดรายสัปดาห์ มาทำการหาจำนวนผู้ป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนของแต่ละจังหวัดในภาคตะวันออกเฉียงเหนือของประเทศไทย และนำข้อมูลที่ได้ในข้างต้นมาหาอัตราการป่วย (Morbidity rate) ต่อแสนประชากร รายละเอียดของข้อมูลเบื้องต้นแสดงดังตารางที่ 1

## 7. ผลการวิเคราะห์ข้อมูลเบื้องต้น

จากตารางที่ 1 พบว่า ข้อมูลอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนของทุกจังหวัด มีค่าต่ำสุดหรือไม่มีอัตราการป่วย ได้แก่ จังหวัดชัยภูมิ จังหวัดมหาสารคาม จังหวัดหนองบัวลำภู จังหวัดอำนาจเจริญ จังหวัดกาฬสินธุ์ จังหวัดยโสธร จังหวัดบึงกาฬ จังหวัดร้อยเอ็ด จังหวัดนครพนม จังหวัดมุกดาหาร จังหวัดเลย และจังหวัดหนองคาย มีค่าควอนไทล์ที่ 50 อยู่ในช่วง 0.2294 ถึง 1.4203 ต่อแสนประชากร มีค่าควอนไทล์ที่ 90 อยู่ในช่วง 0.8364 ถึง 6.0434 ต่อแสนประชากร และมีค่าสูงสุดเท่ากับ 34.3810 ต่อแสนประชากร คือจังหวัดบุรีรัมย์ โดยข้อมูลอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนของแต่ละจังหวัดในภาคตะวันออกเฉียงเหนือของประเทศไทย ทุกจังหวัดจะมีจำนวนเดือนอยู่ที่ 156 เดือน แต่จะมีเพียงจังหวัดบึงกาฬที่มีจำนวนเดือนอยู่ที่ 60 เดือน เพราะจังหวัดบึงกาฬนั้นเริ่มก่อตั้งขึ้นเมื่อปี พ.ศ. 2554 ซึ่งแยกออกมาจากจังหวัดหนองคาย

**ตารางที่ 1** ข้อมูลเบื้องต้นของอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนโดยจำแนกตามจังหวัด

จังหวัด	ค่าต่ำสุด (ต่อแสนประชากร)	ควอนไทล์ที่ 50 (ต่อแสนประชากร)	ควอนไทล์ที่ 90 (ต่อแสนประชากร)	ค่าสูงสุด (ต่อแสนประชากร)
ชัยภูมิ	0.0000	0.4436	2.0309	15.3038
บุรีรัมย์	0.3273	1.1706	2.5946	34.3810
มหาสารคาม	0.0000	0.6336	1.4353	7.6761
ศรีสะเกษ	0.1366	0.7627	2.1984	9.9039
หนองบัวลำภู	0.0000	0.3978	1.6075	13.7940
อำนาจเจริญ	0.0000	0.5346	1.8851	11.0372
กาฬสินธุ์	0.0000	0.3078	0.9203	6.5348
ยโสธร	0.0000	0.9262	4.1186	18.3607
บึงกาฬ	0.0000	0.8449	4.5847	9.4106
ร้อยเอ็ด	0.0000	0.2294	0.8364	5.4294
ขอนแก่น	0.0572	0.5966	2.2629	17.2240
นครพนม	0.0000	1.4203	6.0434	12.6406
นครราชสีมา	0.0391	0.4678	4.0258	13.3435
มุกดาหาร	0.0000	0.4388	2.3395	27.4489
เลย	0.0000	1.2866	5.6421	15.1710
สกลนคร	0.0895	0.5359	2.2697	5.9969
สุรินทร์	0.1442	0.9034	2.4118	9.5156
หนองคาย	0.0000	0.5869	3.5019	9.0101
อุดรธานี	0.0641	0.4595	1.7063	17.4333
อุบลราชธานี	0.1124	1.0348	4.7722	28.4226

#### 8. รูปแบบของค่าประมาณพารามิเตอร์ที่เหมาะสมของข้อมูลภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป

จากตารางที่ 2 พบว่า รูปแบบของค่าประมาณพารามิเตอร์ที่เหมาะสมของข้อมูลอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป รูปแบบ 1 คิดเป็นร้อยละ 30 ได้แก่ จังหวัดชัยภูมิ จังหวัดบุรีรัมย์ จังหวัดมหาสารคาม จังหวัดศรีสะเกษ จังหวัดหนองบัวลำภู และจังหวัดอำนาจเจริญ สำหรับรูปแบบ 2 คิดเป็นร้อยละ 10 ได้แก่ จังหวัดกาฬสินธุ์ และจังหวัดยโสธร สำหรับรูปแบบ 3 คิดเป็นร้อยละ 10 ได้แก่ จังหวัดบึงกาฬ และจังหวัดร้อยเอ็ด และรูปแบบ 5 คิดเป็นร้อยละ 50 ได้แก่ จังหวัดขอนแก่น จังหวัดนครพนม จังหวัดนครราชสีมา จังหวัดมุกดาหาร จังหวัดเลย จังหวัดสกลนคร จังหวัดสุรินทร์ จังหวัดหนองคาย จังหวัดอุดรธานี และจังหวัดอุบลราชธานี

ตารางที่ 2 รูปแบบของค่าประมาณพารามิเตอร์ที่เหมาะสมของข้อมูลอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป จำแนกตามจังหวัด

จังหวัด	รูปแบบของค่าประมาณพารามิเตอร์ที่เหมาะสม
ชัยภูมิ	รูปแบบ 1: $\hat{\mu} = 0.3218, \hat{\sigma} = 0.2928, \hat{\xi} = 0.6775$
บุรีรัมย์	รูปแบบ 1: $\hat{\mu} = 0.9147, \hat{\sigma} = 0.5430, \hat{\xi} = 0.6139$
มหาสารคาม	รูปแบบ 1: $\hat{\mu} = 0.5026, \hat{\sigma} = 0.3384, \hat{\xi} = 0.2893$
ศรีสะเกษ	รูปแบบ 1: $\hat{\mu} = 0.5888, \hat{\sigma} = 0.3932, \hat{\xi} = 0.5150$
หนองบัวลำภู	รูปแบบ 1: $\hat{\mu} = 0.2582, \hat{\sigma} = 0.2900, \hat{\xi} = 0.5758$
อำนาจเจริญ	รูปแบบ 1: $\hat{\mu} = 0.3234, \hat{\sigma} = 0.3766, \hat{\xi} = 0.5690$
กาฬสินธุ์	รูปแบบ 2: $\mu(t) = 0.2537+(-0.0715t), \hat{\sigma} = 0.1935, \hat{\xi} = 0.4012$
ยโสธร	รูปแบบ 2: $\mu(t) = 0.6908+0.2620t, \hat{\sigma} = 0.6481, \hat{\xi} = 0.5938$
บึงกาฬ	รูปแบบ 3: $\sigma(t) = 0.6360+0.1772t, \hat{\mu} = 0.6318, \hat{\xi} = 0.7973$
ร้อยเอ็ด	รูปแบบ 3: $\sigma(t) = 0.1632+(-0.0374t), \hat{\mu} = 0.1751, \hat{\xi} = 0.5090$
ขอนแก่น	รูปแบบ 5: $\mu(t) = 0.4794+0.3103t, \sigma(t) = \exp(-0.0071+0.3532t), \hat{\xi} = 0.4764$
นครพนม	รูปแบบ 5: $\mu(t) = 1.0631+0.6601t, \sigma(t) = \exp(-0.1023+1.0717t), \hat{\xi} = 0.2679$
นครราชสีมา	รูปแบบ 5: $\mu(t) = 0.3520+0.1669t, \sigma(t) = \exp(-0.0560+0.4053t), \hat{\xi} = 0.7472$
มุกดาหาร	รูปแบบ 5: $\mu(t) = 0.3470+0.1647t, \sigma(t) = \exp(0.2455+0.1420t), \hat{\xi} = 0.6561$
เลย	รูปแบบ 5: $\mu(t) = 0.9786+0.7614t, \sigma(t) = \exp(-0.0651+0.9100t), \hat{\xi} = 0.5077$
สกลนคร	รูปแบบ 5: $\mu(t) = 0.4252+0.2932t, \sigma(t) = \exp(0.0312+0.2831t), \hat{\xi} = 0.5062$
สุรินทร์	รูปแบบ 5: $\mu(t) = 0.7718+(-0.2880t), \sigma(t) = \exp(0.5880+(-0.1311t)), \hat{\xi} = 0.3416$
หนองคาย	รูปแบบ 5: $\mu(t) = 0.5412+0.5040t, \sigma(t) = \exp(-0.0515+0.4660t), \hat{\xi} = 0.4515$
อุดรธานี	รูปแบบ 5: $\mu(t) = 0.3759+(-0.1686t), \sigma(t) = \exp(0.4097+(-0.1014t)), \hat{\xi} = 0.6769$
อุบลราชธานี	รูปแบบ 5: $\mu(t) = 0.8122+0.7124t, \sigma(t) = \exp(0.0578+0.5212t), \hat{\xi} = 0.6322$

### 9. ผลการวิเคราะห์รูปแบบของค่าประมาณพารามิเตอร์ที่เหมาะสมภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป และค่าประมาณระดับการเกิดซ้ำของจังหวัดนครราชสีมา

จากการวิเคราะห์รูปแบบของค่าประมาณพารามิเตอร์ที่เหมาะสมภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไปและค่าประมาณระดับการเกิดซ้ำ ของทุกจังหวัดในภาคตะวันออกเฉียงเหนือของประเทศไทยนั้น ผู้วิจัยพบว่าจังหวัดนครราชสีมา มีค่าประมาณระดับการเกิดซ้ำของอัตราการป่วยโรคใช้หัวใจใหญ่ในรอบ 5 ปี 10 ปี และ 20 ปี สูงที่สุดในภาคตะวันออกเฉียงเหนือของประเทศไทย ดังนั้นผู้วิจัยจึงขอนำเสนอผลการวิเคราะห์ของจังหวัดนครราชสีมาแสดงดังตารางที่ 3 ถึง 5 ดังนี้

**ตารางที่ 3** จำนวนพารามิเตอร์ (P) และค่า Negative log-likelihood (NLL) ของรูปแบบข้อมูลอัตราการป่วยโรคใช้หัวใจใหญ่สูงสุดรายเดือนในจังหวัดนครราชสีมา

รูปแบบ	P	NLL
รูปแบบ 1: $\mu, \sigma, \xi$ เป็นค่าคงที่	3	177.1748
รูปแบบ 2: $\mu(t) = \beta_0 + \beta_1 t$ โดยที่ $\sigma$ และ $\xi$ เป็นค่าคงที่	4	174.7546
รูปแบบ 3: $\sigma(t) = \alpha_0 + \alpha_1 t$ โดยที่ $\mu$ และ $\xi$ เป็นค่าคงที่	4	166.7883
รูปแบบ 4: $\sigma(t) = \exp(\alpha_0 + \alpha_1 t)$ โดยที่ $\mu$ และ $\xi$ เป็นค่าคงที่	4	162.2317
รูปแบบ 5: $\mu(t) = \beta_0 + \beta_1 t, \sigma(t) = \exp(\alpha_0 + \alpha_1 t)$ โดยที่ $\xi$ เป็นค่าคงที่	5	155.2067

จากตารางที่ 3 ค่า Negative log-likelihood (NLL) ของรูปแบบข้อมูลอัตราการป่วยโรคใช้หัวใจใหญ่สูงสุดรายเดือนในจังหวัดนครราชสีมา เรียงลำดับจากน้อยไปมากได้ดังนี้ รูปแบบ 1 (177.1748), รูปแบบ 2 (174.7546), รูปแบบ 3 (166.7883), รูปแบบ 4 (162.2317) และรูปแบบ 5 (155.2067) ตามลำดับ

**ตารางที่ 4** การวิเคราะห์ความแตกต่างของรูปแบบโดยใช้การทดสอบอัตราส่วนภาวะน่าจะเป็น (Likelihood ratio test) ในการคัดเลือกรูปแบบของพารามิเตอร์ของจังหวัดนครราชสีมา

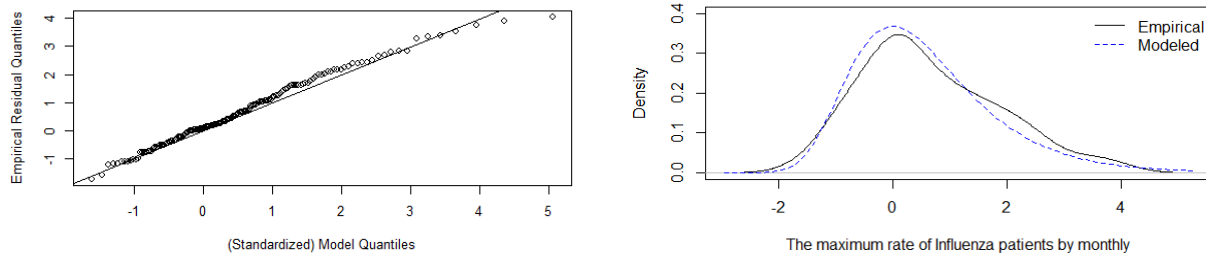
เปรียบเทียบรูปแบบ	อัตราส่วนภาวะน่าจะเป็น	ไคสแควร์	p-value	รูปแบบที่ถูกเลือก
1 กับ 3	20.7730	3.8410	5.171e-06*	3
3 กับ 4	9.1132	0.0000	2.2e-16*	4
1 กับ 2	4.8406	3.8410	0.0278*	2
2 กับ 5	39.0957	3.8410	4.035e-10*	5
4 กับ 5	14.0500	3.8410	0.0002*	5

\* ที่ระดับนัยสำคัญ 0.05

**ตารางที่ 5** ค่าประมาณพารามิเตอร์และช่วงความเชื่อมั่น 95% ของรูปแบบข้อมูลอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนในจังหวัดนครราชสีมา

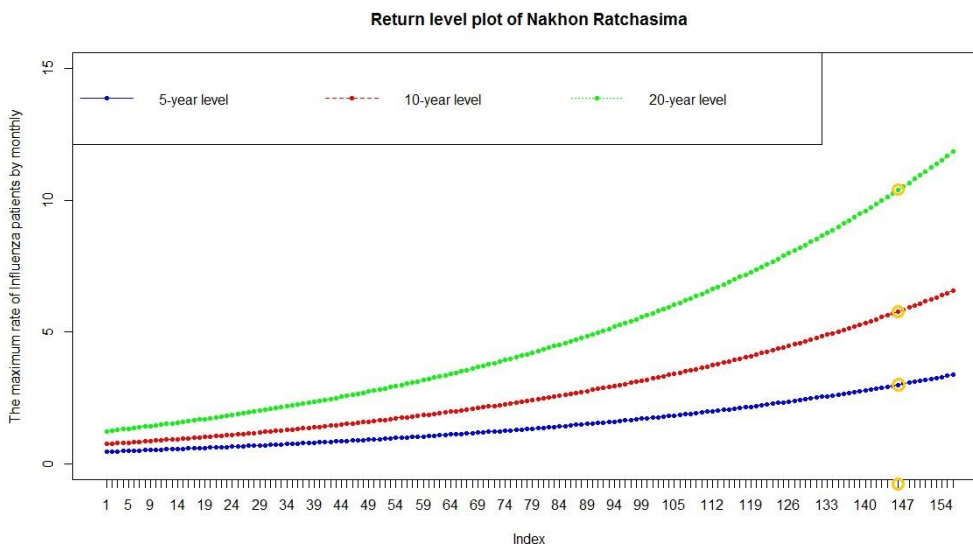
รูปแบบ	ค่าประมาณพารามิเตอร์	ขอบล่าง	ขอบบน
รูปแบบ 5: $\mu(t) = \beta_0 + \beta_1 t$ , $\sigma(t) = \exp(\alpha_0 + \alpha_1 t)$ โดยที่ $\xi$ เป็นค่าคงที่	$\hat{\beta}_0 = 0.3520$ $\hat{\beta}_1 = 0.1669$ $\hat{\alpha}_0 = -0.0560$ $\hat{\alpha}_1 = 0.4053$ $\hat{\xi} = 0.7472$	0.2829 0.0682 -0.1397 0.2736 0.5657	0.4212 0.2656 0.0277 0.5369 0.9287

จากตารางที่ 4 และ 5 พบว่า รูปแบบ 5 ถูกเลือกให้เป็นรูปแบบที่ดีที่สุดของข้อมูลอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนในจังหวัดนครราชสีมา ดังนั้นสามารถเขียนสมการอธิบายความหมายได้ ดังนี้  $X \sim GEV(\mu(t) = 0.3520 + 0.1669t, \sigma(t) = \exp(-0.0560 + 0.4053t), \xi = 0.7472)$  และพิจารณาความเหมาะสมของรูปแบบที่ได้ในข้างต้นแสดงดังภาพที่ 3



**ภาพที่ 3** Quantile plot (ซ้าย) และ Density plot (ขวา) ของแบบจำลองค่าสุดขีดวงนัยทั่วไปกับข้อมูลอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนในจังหวัดนครราชสีมา

จากภาพที่ 3 Quantile plot (ซ้าย) พบว่า จุดควอนไทล์ของค่าคาดหวังของการแจกแจงที่ได้จากการสร้างแบบจำลองกับจุดควอนไทล์ของข้อมูลจริง เส้นกราฟอยู่ในรูปแนวเส้นตรง  $y = x$  แสดงว่าข้อมูลทั้งสองมีการแจกแจงเหมือนกัน และ Density plot (ขวา) พบว่า พล็อตค่าข้อมูลจริง กับค่าความหนาแน่นของความน่าจะเป็นของแบบจำลอง ทั้งสองเส้นมีลักษณะคล้ายกัน จึงสรุปได้ว่าแบบจำลองของค่าสุดขีดวงนัยทั่วไปมีความเหมาะสมกับชุดข้อมูลนี้ ดังนั้นผู้วิจัยจึงนำค่าประมาณพารามิเตอร์ที่วิเคราะห์ได้ไปทำการหาค่าประมาณระดับการเกิดซ้ำ (Return level) ของอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนในรอบการเกิดซ้ำ 5 ปี 10 ปี และ 20 ปี ของจังหวัดนครราชสีมา แสดงดังภาพที่ 4

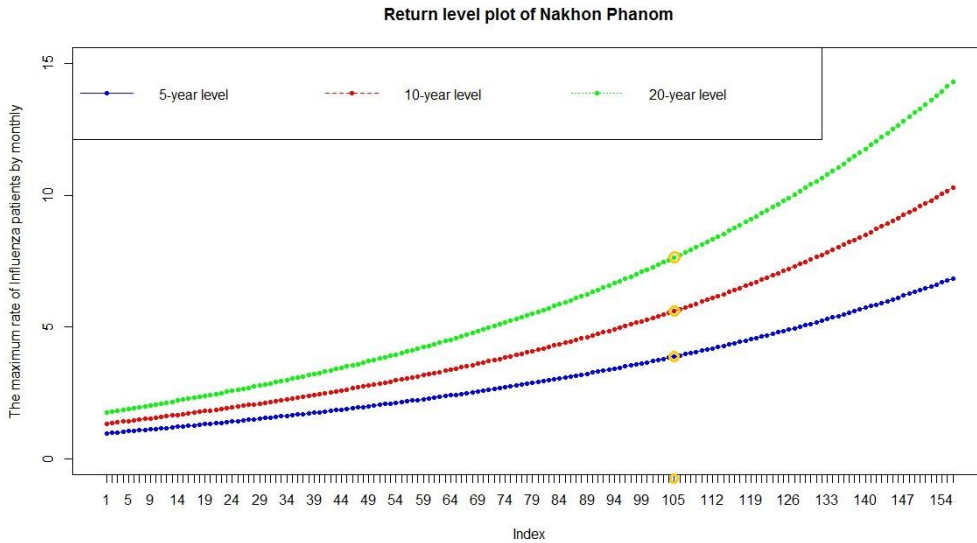


**ภาพที่ 4** กราฟระดับการเกิดซ้ำ (Return level plot) ของจังหวัดนครราชสีมา

จากภาพที่ 4 ผู้วิจัยได้เปรียบเทียบช่วงเวลาที่เกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนย้อนหลัง 13 ปี (พ.ศ. 2546 – 2558) จากข้อมูลจริงในจังหวัดนครราชสีมา พบว่าอัตราการป่วยที่สูงที่สุดตรงกับเดือนกุมภาพันธ์ ปี พ.ศ. 2558 ซึ่งอยู่ในตำแหน่ง Index ที่ 146 จะได้ว่า ค่าประมาณระดับการเกิดซ้ำของจังหวัดนครราชสีมา ในรอบการเกิดซ้ำ 5 ปี จะเกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนที่ 2.9898 ต่อแสนประชากร ด้วยความน่าจะเป็น 0.2 สำหรับรอบการเกิดซ้ำ 10 ปี จะเกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนที่ 5.7725 ต่อแสนประชากร ด้วยความน่าจะเป็น 0.1 และรอบการเกิดซ้ำ 20 ปี จะเกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนที่ 10.3907 ต่อแสนประชากร ด้วยความน่าจะเป็น 0.05

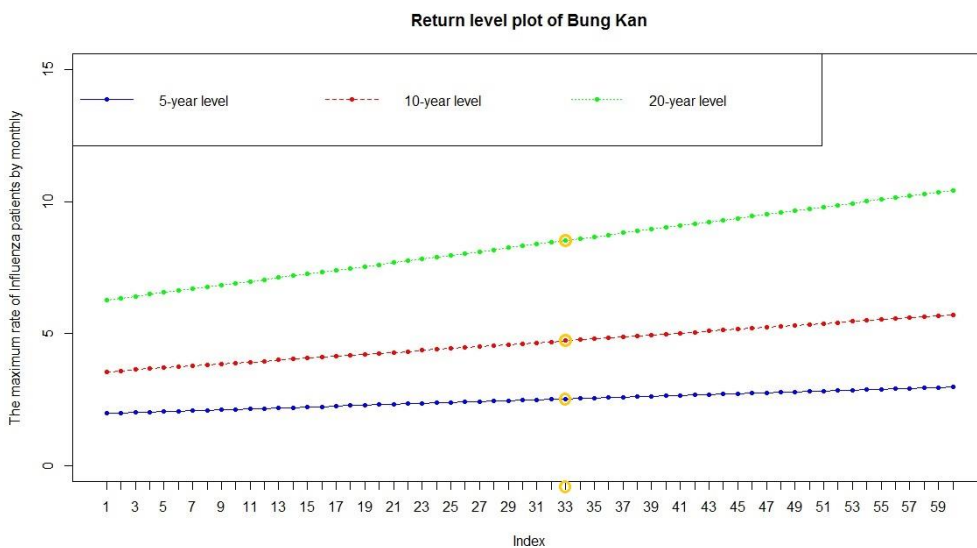
#### 10. ผลการวิเคราะห์ค่าประมาณระดับการเกิดซ้ำของจังหวัดนครพนม และจังหวัดบึงกาฬ

จากการวิเคราะห์รูปแบบของค่าประมาณพารามิเตอร์ที่เหมาะสมภายใต้การแจกแจงค่าสุดขีดวางนัยทั่วไปและค่าประมาณระดับการเกิดซ้ำ ของทุกจังหวัดในภาคตะวันออกเฉียงเหนือของประเทศไทย ผู้วิจัยพบว่าจังหวัดนครพนมและจังหวัดบึงกาฬ มีค่าประมาณระดับการเกิดซ้ำของอัตราการป่วยโรคไข้หวัดใหญ่ในรอบ 5 ปี 10 ปี และ 20 ปี สูงรองลงมาจากจังหวัดนครราชสีมา ซึ่งนับว่าเป็นอีกสองจังหวัดในภาคตะวันออกเฉียงเหนือที่ควรมีการเฝ้าระวังการเกิดโรคไข้หวัดใหญ่ในอนาคต โดยค่าประมาณระดับการเกิดซ้ำแสดงดังภาพที่ 5 และ 6 ดังนี้



ภาพที่ 5 กราฟระดับการเกิดซ้ำ (Return level plot) ของจังหวัดนครพนม

จากภาพที่ 5 ผู้วิจัยได้เปรียบเทียบช่วงเวลาที่เกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนย้อนหลัง 13 ปี (พ.ศ. 2546 – 2558) จากข้อมูลจริงในจังหวัดนครพนม พบว่าอัตราการป่วยที่สูงที่สุดตรงกับเดือนกันยายน ปี พ.ศ. 2554 ซึ่งอยู่ในตำแหน่ง Index ที่ 105 จะได้ว่า ค่าประมาณระดับการเกิดซ้ำของจังหวัดนครพนม ในรอบการเกิดซ้ำ 5 ปี จะเกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนที่ 3.8698 ต่อแสนประชากร ด้วยความน่าจะเป็น 0.2 สำหรับรอบการเกิดซ้ำ 10 ปี จะเกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนที่ 5.6084 ต่อแสนประชากร ด้วยความน่าจะเป็น 0.1 และรอบการเกิดซ้ำ 20 ปี จะเกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนที่ 7.6388 ต่อแสนประชากร ด้วยความน่าจะเป็น 0.05

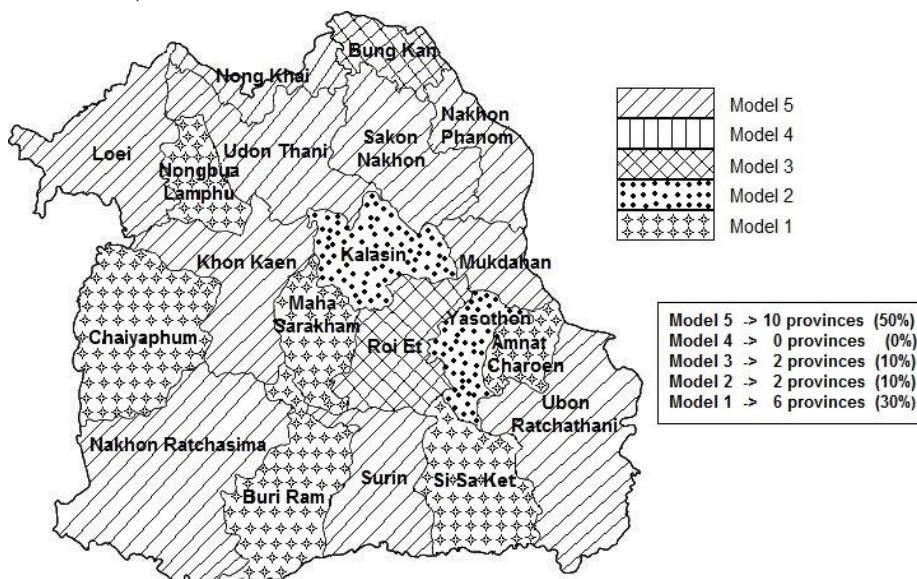


ภาพที่ 6 กราฟระดับการเกิดซ้ำ (Return level plot) ของจังหวัดบึงกาฬ

จากภาพที่ 6 ผู้วิจัยได้เปรียบเทียบช่วงเวลาที่เกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนย้อนหลัง 5 ปี (พ.ศ. 2554 – 2558) จากข้อมูลจริงในจังหวัดบึงกาฬ พบว่าอัตราการป่วยที่สูงที่สุดตรงกับเดือนกันยายน ปี พ.ศ. 2556 ซึ่งอยู่ในตำแหน่ง Index ที่ 33 จะได้ว่า ค่าประมาณระดับการเกิดซ้ำของจังหวัดบึงกาฬ ในรอบการเกิดซ้ำ 5 ปี จะเกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนที่ 2.5138 ต่อแสนประชากร ด้วยความน่าจะเป็น 0.2 สำหรับรอบการเกิดซ้ำ 10 ปี จะเกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนที่ 4.7236 ต่อแสนประชากร ด้วยความน่าจะเป็น 0.1 และรอบการเกิดซ้ำ 20 ปี จะเกิดอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนที่ 8.5283 ต่อแสนประชากร ด้วยความน่าจะเป็น 0.05

### สรุปผลการวิจัย

การสร้างแบบจำลองภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป (GEVD) กับข้อมูลอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนในภาคตะวันออกเฉียงเหนือของประเทศไทย ผลลัพธ์ที่ได้แสดงดังภาพที่ 7 พบว่า รูปแบบ 1 เป็นกลุ่มจังหวัดขนาดเล็ก มีจำนวนประชากรที่ค่อนข้างคงที่ได้แก่ จังหวัดชัยภูมิ จังหวัดบุรีรัมย์ จังหวัดมหาสารคาม จังหวัดศรีสะเกษ จังหวัดหนองบัวลำภู และจังหวัดอำนาจเจริญ สำหรับรูปแบบ 2 เป็นกลุ่มจังหวัดที่มีบริเวณที่ตั้งอยู่ในตอนกลางของภาคตะวันออกเฉียงเหนือ ได้แก่ จังหวัดกาฬสินธุ์ และจังหวัดยโสธร สำหรับรูปแบบ 3 เป็นกลุ่มจังหวัดขนาดเล็กที่มีการเจริญเติบโตของเศรษฐกิจและจำนวนประชากรอย่างต่อเนื่อง ได้แก่ จังหวัดบึงกาฬ และจังหวัดร้อยเอ็ด และรูปแบบ 5 เป็นกลุ่มจังหวัดทางเศรษฐกิจที่สำคัญของภาคตะวันออกเฉียงเหนือ รวมทั้งมีจำนวนประชากรและจำนวนผู้ป่วยที่ค่อนข้างสูง ได้แก่ จังหวัดขอนแก่น จังหวัดนครพนม จังหวัดนครราชสีมา จังหวัดมุกดาหาร จังหวัดเลย จังหวัดสกลนคร จังหวัดสุรินทร์ จังหวัดหนองคาย จังหวัดอุดรธานี และจังหวัดอุบลราชธานี



ภาพที่ 7 รูปแบบของค่าประมาณพารามิเตอร์ที่เหมาะสมกับข้อมูลอัตราการป่วยโรคไข้หวัดใหญ่สูงสุดรายเดือนในภาคตะวันออกเฉียงเหนือของประเทศไทย ภายใต้การแจกแจงค่าสุดขีดวงนัยทั่วไป

ภายใต้สภาพภูมิอากาศ และสถานการณ์ทั่วไปต่างๆ ที่มีอิทธิพลต่อการเกิดโรคไข้หวัดใหญ่ไม่เปลี่ยนแปลงไปจากเดิมมากนัก ในภาคตะวันออกเฉียงเหนือของประเทศไทยมีจังหวัดที่ควรมีการเฝ้าระวังการเกิดโรคไข้หวัดใหญ่ในอนาคต ได้แก่



จังหวัดนครราชสีมา จังหวัดนครพนม และจังหวัดบึงกาฬ เนื่องจากมีค่าประมาณระดับการเกิดซ้ำ (Return level) ในรอบการเกิดซ้ำ 5 ปี 10 ปี และ 20 ปี สูงที่สุดเป็นสามอันดับแรก ซึ่งผลของการวิเคราะห์ในครั้งนี้จึงอาจเป็นแนวทางนำไปสู่การวางแผนงานทางด้านสาธารณสุขเพื่อรับมือกับเหตุการณ์ของโรคไข้หวัดใหญ่ในแต่ละกลุ่มจังหวัดในภาคตะวันออกเฉียงเหนือของประเทศไทยต่อไป ดังนั้นการนำทฤษฎีค่าสุดขีด (EVT) มาประยุกต์ใช้กับข้อมูลทางด้านระบาดวิทยาจึงเป็นอีกทางเลือกหนึ่งที่น่าสนใจ อย่างไรก็ตามข้อจำกัดของวิธีการนี้คืออิทธิพลความไม่แน่นอนของปัจจัยต่างๆ อาทิเช่น อุณหภูมิ ความชื้น ความหนาแน่นของประชากรในพื้นที่ จำนวนประชากรแฝง และการได้รับวัคซีนป้องกันโรคไข้หวัดใหญ่ ที่จะส่งผลต่อค่าความน่าจะเป็นของอัตราการป่วยและการเสียชีวิตที่คาดการณ์ได้ จึงจำเป็นต้องใช้วิธีการวิเคราะห์และแบบจำลองอื่นที่เหมาะสมต่อไป

### กิตติกรรมประกาศ

โครงการวิจัยนี้ได้รับทุนสนับสนุนโครงการวิจัยของนิสิตปริญญาโท มหาวิทยาลัยมหาสารคาม งบประมาณแผ่นดินปี 2560 และข้อมูลจำนวนผู้ป่วยโรคไข้หวัดใหญ่สะสมสูงสุดรายสัปดาห์ จากสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุข

### เอกสารอ้างอิง

- Busababodin, P. (2016). *Extremes Value Analysis with R*. Maha Sarakham: Maha Sarakham University.
- Chanwirat, V., Gankan, A., & Chimin, M. (2015). The Report of Forecasting and Health risk in Influenza cases in year 2016. Retrieved October 21, 2016, from <http://odpc11.ddc.moph.go.th/newsdata8pic482.pdf>
- Chen, J., Lei, X., Zhang, L., & Peng, B. (2015). Using Extreme Value Theory Approaches to Forecast the Probability of Outbreak of Highly Pathogenic Influenza in Zhejiang, China. *PLoS ONE*, 10(2).
- Coles, S., & Nadaraja, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Great Britain: Springer-Varlag London Limited.
- Epidemiology Department of Disease Control Ministry of Public Health (2016). The Report on the Disease Surveillance System 506. Retrieved March 17, 2016, from [http://www.boe.moph.go.th/boedb/d506\\_1/ds\\_wk2pdf.php?ds=15&yr=58](http://www.boe.moph.go.th/boedb/d506_1/ds_wk2pdf.php?ds=15&yr=58)
- Thomas, M., Lemaitre, M., Wilson, M.L., Viboud, C., Yordanov, Y., Wackernagel, H., & Carrat, F. (2016). Applications of Extreme Value Theory in Public Health. *PLoS ONE*, 11(7).
- Seesomboon, N., Sayrat, W., & Tachina, P. (2013). Forecasting the Occurrence of Influenza in 8 provinces in the Northern of Thailand in year 2014. Retrieved September 21, 2016, from [http://www.interfetpthailand.net/forecast/files/report\\_2014/report\\_2014\\_no14.pdf](http://www.interfetpthailand.net/forecast/files/report_2014/report_2014_no14.pdf)
- Sinsomboonthong, S. (2015). *Mathematical Statistics 2*. Bangkok: Jamjureeproduct.
- Suttachana, S., Kaewnokkao, W., Suangtoh, P., Sayumpoorujinan, S., & Kongyoo, S. (2012). Forecasting the Occurrence of Influenza in Thailand in year 2012. Retrieved September 21, 2016, from [http://www.interfetpthailand.net/forecastfilesreport\\_2012report\\_2012\\_11\\_no20.pdf](http://www.interfetpthailand.net/forecastfilesreport_2012report_2012_11_no20.pdf)

Ueawaragun, P. (2010). Influenza. Retrieved October 6, 2016, from <http://www.si.mahidol.ac.th/sidoctor/e-pl/article/detail.asp?id=311>