

Joinpoint Regression: ภาพรวมและการประยุกต์ใช้ในการศึกษาวิจัยแนวโน้มหรือสถานการณ์ปัญหาทางวิทยาศาสตร์สุขภาพ

Joinpoint Regression: Overview and Application in Research on Trends or Situations of Health Science Problems

นครินทร์ อสีพงษ์¹, พงษ์เดช สารการ^{2,3*}Nakarin Asiphong¹, Pongdech Sarakarn^{2,3*}

บทคัดย่อ

ในการศึกษาวิจัยทางด้านวิทยาศาสตร์สุขภาพ การวิเคราะห์ข้อมูลในลักษณะของการศึกษาแนวโน้มการเปลี่ยนแปลงของโรครวมไปถึงการพยากรณ์จำนวนผู้ป่วยในอนาคตเป็นรูปแบบการศึกษาหนึ่งที่ได้รับ ความสนใจอย่างมาก ด้วยผลลัพธ์ที่เป็นประโยชน์ต่อหน่วยงานภาครัฐที่รับผิดชอบดูแลด้านสุขภาพของประชาชนในประเทศในการตัดสินใจวางแผนดำเนินงานหรือรับมือกับปัญหาสุขภาพต่าง ๆ วิธีการในการวิเคราะห์ข้อมูลในลักษณะดังกล่าวที่มีการใช้กันอย่างแพร่หลายคือ Joinpoint regression ที่มีการศึกษาแนวโน้ม (Trend) และการเปลี่ยนแปลง (Change) ของข้อมูลผลลัพธ์ทางด้านสุขภาพ เช่น อัตราป่วย อัตราตาย จำนวนผู้ป่วย เป็นต้น และปัจจุบันมีการพัฒนาโปรแกรมสำเร็จรูปในการช่วยให้ นักวิจัยสามารถวิเคราะห์ข้อมูลได้สะดวกมากขึ้น ซึ่งพัฒนาโดย National Cancer Institute นักวิจัยหรือผู้ที่สนใจสามารถเข้าถึงได้โดยไม่มีค่าใช้จ่าย อย่างไรก็ตามเครื่องมือสำเร็จรูปในการช่วยอำนวยความสะดวกในการวิเคราะห์นั้น นับเป็นสิ่งที่ เป็นประโยชน์ แต่สำหรับนักวิจัยแล้วการเลือกใช้วิธีการวิเคราะห์ข้อมูลรวมไปถึงการแปลผลที่ได้จากโปรแกรมสำเร็จรูปได้อย่างมีประสิทธิภาพและถูกต้องนั้นจำเป็นต้องมีความเข้าใจในหลักการ แนวคิด และทฤษฎีที่ได้มาซึ่งผลลัพธ์ของการวิเคราะห์นั้น เพื่อนำไปสู่การเผยแพร่ผลการศึกษาที่ถูกต้องและไม่ก่อให้เกิดความเข้าใจผิดหรือการชี้นำที่ผิด (Misleading)

คำสำคัญ: Joinpoint regression, การวิเคราะห์การเปลี่ยนแปลงของแนวโน้ม, การค้นหาจุดเปลี่ยน

Citation:

Asiphong N, Sarakarn P. Joinpoint Regression: Overview and application in research on trends or situations of health science problems. Health Sci J Thai 2025; 7(4): 25-36. (in Thai); <https://doi.org/10.55164/hsjt.v7i4.268339>

¹ หลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการระบาดและชีวสถิติ หลักสูตรนานาชาติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น 40002

² คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น 40002

³ กลุ่มวิจัยวิทยาการระบาดและป้องกันโรคเมร่งในภูมิภาคอาเซียน (ACEP) ขอนแก่น 40002

¹ Doctor of Philosophy Program in Epidemiology and Biostatistics International Program, Faculty of Public Health, Khon Kaen University, 40002, Thailand

² Faculty of Public Health, Khon Kaen University, 40002, Thailand

³ ASEAN Cancer Epidemiology and Prevention Research Group (ACEP), Khon Kaen, 40002, Thailand

*Corresponding author: E-mail: spondg@kku.ac.th, Tel: 081-3914664

Received: 3 Mar 2024; Revised: 21 Oct 2025; Accepted: 27 Oct 2025

Abstract

Health research is informed by the study of health science, trends, and predicted patient numbers, with the aim of using these results to support government health organizations responsible for public health. It encourages and assists decision-makers in determining whether to plan operations or handle health issues. Joinpoint regression is a commonly used technique for evaluating this type of data. It looks at trends and changes in health outcome data, like the number of patients, mortality rates, and morbidity rates. Additionally, the National Cancer Institute has created a ready-made program to facilitate data analysis for researchers. It is available without charge to researchers and interested parties. However, having ready-made tools to facilitate analysis is a beneficial thing. But for researchers, choosing a data analysis method and interpreting the results obtained from ready-made programs effectively and correctly requires an understanding of the principles, concepts, and theories from which the results of the analysis are derived. This understanding leads to the dissemination of accurate study results and helps prevent misunderstandings or misleading conclusions.

Keyword: Joinpoint regression, Trend change analysis, Finding the change point

บทนำ

ในการศึกษาวิจัยทางด้านวิทยาศาสตร์สุขภาพ ข้อมูลในลักษณะของการศึกษาแนวโน้มและการเปลี่ยนแปลงของโรครวมไปถึงการพยากรณ์จำนวนผู้ป่วยในอนาคตเป็นรูปแบบการศึกษาที่มีประโยชน์ในการวางแผนดำเนินงานหรือรับมือกับปัญหาสุขภาพ ซึ่งวิธีการในการวิเคราะห์ข้อมูลในลักษณะดังกล่าวที่มีการใช้กันอย่างแพร่หลายคือ Joinpoint regression ซึ่งในปัจจุบันวิธีการวิเคราะห์ Joinpoint regression เป็นวิธีการที่ถูกนำมาใช้อย่างแพร่หลายในงานวิจัยทางด้านวิทยาศาสตร์สุขภาพ เพื่อศึกษาแนวโน้ม (Trend) และการเปลี่ยนแปลง (Change) ของข้อมูลผลลัพธ์ทางด้านสุขภาพ เช่น อัตราป่วย อัตราตาย จำนวนผู้ป่วย เป็นต้น⁽¹⁾ โดยมีเวลาเป็นตัวแปรร่วม (Covariate) ซึ่งได้แก่ จุดเวลาที่ทำให้การวัดผลลัพธ์เช่น อัตราป่วยด้วยโรคมะเร็งรายปี หรือจำนวนผู้ป่วยโรคมะเร็งรายปี โดยแนวคิดของการวิเคราะห์ Joinpoint regression นั้นเป็นการค้นหาจุดเปลี่ยน (Change point) ของแนวโน้มของข้อมูล ซึ่งเป็นการพิจารณาจุดเปลี่ยนที่มีนัยสำคัญทางสถิติ จากนั้นทำการพิจารณาจำนวนจุดเปลี่ยนที่เหมาะสมกับชุดข้อมูลมากที่สุดโดยการทดสอบทางสถิติเพื่อให้ได้โมเดลสุดท้ายที่สอดคล้องกับข้อมูลที่สุด และมีการอธิบายการเปลี่ยนแปลงของข้อมูลเป็นช่วง ๆ หรือส่วนย่อย (Segment) โดยอธิบายร่วมกับค่า Annual percent change (APC) เพื่ออธิบายการเปลี่ยนแปลงของแต่ละส่วนย่อย (Segment) นอกจากนี้ยังมีการใช้ค่า Average Annual percent change (AAPC) ในการอธิบายการเปลี่ยนแปลงในภาพรวมของชุดข้อมูลนั้น ๆ ด้วย⁽²⁾ และเนื่องด้วย Joinpoint regression เป็นวิธีการวิเคราะห์ที่สามารถทำให้ทราบถึงตำแหน่งของจุดเปลี่ยนของข้อมูลและจำนวนจุดเปลี่ยนที่เหมาะสมสำหรับข้อมูลชุดนั้น⁽³⁾ จึงทำให้สามารถอธิบายการเปลี่ยนแปลงของข้อมูลได้ชัดเจนและระบุรายละเอียดของการเปลี่ยนแปลงได้

มากขึ้นทำให้ได้มาซึ่งข้อมูลที่มีคุณค่าในการนำไปใช้ประโยชน์ และในปัจจุบันได้มีการพัฒนาโปรแกรมสำเร็จรูปในการช่วยให้นักวิจัยสามารถวิเคราะห์ข้อมูลได้สะดวกมากขึ้น โดยการใช้โปรแกรมสำเร็จรูปสำหรับการวิเคราะห์ Joinpoint regression ที่ชื่อว่า Joinpoint Regression Program ซึ่งพัฒนาโดย National Cancer Institute⁽⁴⁾ ซึ่งนักวิจัยหรือผู้ที่สนใจสามารถเข้าถึงได้โดยไม่มีค่าใช้จ่าย อย่างไรก็ตามการมีเครื่องมือสำเร็จรูปในการช่วยอำนวยความสะดวกในการวิเคราะห์นั้นนับเป็นสิ่งที่ดี แต่สำหรับนักวิจัยแล้วการเลือกใช้วิธีการวิเคราะห์ข้อมูล รวมไปถึงการแปลผลที่ได้จากโปรแกรมสำเร็จรูปได้อย่างมีประสิทธิภาพและถูกต้องนั้นจำเป็นต้องมีความเข้าใจในหลักการ แนวคิด และทฤษฎีที่ได้มาซึ่งผลลัพธ์ของการวิเคราะห์นั้น เพื่อนำไปสู่การเผยแพร่ผลการศึกษาที่ถูกต้องและไม่ก่อให้เกิดความเข้าใจผิด ดังนั้นในบทความนี้จึงมีวัตถุประสงค์เพื่อนำเสนอรายละเอียดวิธีการวิเคราะห์ Joinpoint regression ซึ่งประกอบไปด้วยแนวคิด หลักการ การนำไปใช้ การแปลผล ประโยชน์และข้อพึงระวังในการใช้วิธีการวิเคราะห์ เพื่อเป็นข้อมูลประกอบในการพิจารณาใช้วิธีการวิเคราะห์ด้วย Joinpoint regression

1) แนวคิดของการวิเคราะห์ Joinpoint regression

ในการวิเคราะห์ Joinpoint regression มีแนวคิดมาจากการที่นักวิจัยมีความสนใจที่จะศึกษาและวิเคราะห์ข้อมูลในลักษณะที่เป็นการศึกษาแนวโน้มของเหตุการณ์หรือผลลัพธ์ที่สนใจ ซึ่งแนวคิดของการวิเคราะห์นั้นเป็นการพัฒนามาจากวิธีการของ Simple Linear regression แต่มีการเพิ่มเติมรายละเอียดในการสร้างโมเดลในการทำนายค่าผลลัพธ์ หรือ ตัวแปร Y ในการสร้างโมเดลของ Simple Linear regression นั้นจะเป็นการสร้างโมเดลที่ดีที่สุดสำหรับประมาณค่าข้อมูลทั้งชุดข้อมูลนั้นโดยเส้นสมการที่ได้จะเป็นสมการเส้นตรงเพียงเส้นเดียวที่แสดง

แนวโน้มของผลลัพธ์ในช่วงเวลาทั้งหมดที่นำมาวิเคราะห์ ซึ่งผ่านการพิจารณาแล้วว่าสอดคล้องกับข้อมูลมากที่สุด⁽⁵⁾ โดยใช้การพิจารณาจากโมเดลที่มีความคลาดเคลื่อนน้อยที่สุด⁽⁵⁾ ซึ่งในวิธีการของ Joinpoint regression ได้มีการประยุกต์ใช้แนวคิดดังกล่าวในการวิเคราะห์ข้อมูลโดยมีการพิจารณาว่าในชุดข้อมูลที่นำมาวิเคราะห์ซึ่งมีลักษณะเป็นข้อมูลอนุกรมเวลา นักวิจัยได้คำนึงถึงการเปลี่ยนแปลงของแนวโน้มในช่วงเวลาที่นำมาวิเคราะห์ว่าภายในช่วงเวลาทั้งหมดนั้นอาจพบการเปลี่ยนแปลงของแนวโน้มที่แตกต่างกันเป็นช่วงสั้น ๆ อยู่⁽⁶⁾ ซึ่งสามารถแบ่งออกเป็นช่วงย่อย ๆ (Segment) ได้ โดยนำเอาแนวคิดของการแยกส่วนแต่ละช่วงความชัน (Slope) ของเส้นสมการในชุดข้อมูลนั้น ๆ ซึ่งแต่ละช่วงจะมีเส้น slope ของตัวเอง ทั้งช่วงที่ความชันมีทิศทางเพิ่มขึ้นและช่วงที่ความชันมีทิศทางลดลงเข้ามาใส่ในโมเดล ดังนั้น Joinpoint regression model จึงเป็นโมเดลที่ประกอบไปด้วยเส้นความชันย่อย ๆ ที่ต่อเนื่องกันโดยมีจุดเชื่อมต่อที่เรียกว่า Joinpoint⁽⁷⁾ ซึ่งจากงานวิจัยที่มีการเผยแพร่ในปัจจุบัน จากฐานข้อมูลระดับสากลพบว่ามีการนำวิธีการดังกล่าวไปใช้จำนวนมาก โดยเฉพาะในด้านวิทยาศาสตร์สุขภาพ ศึกษาผลลัพธ์ทางสุขภาพ⁽⁸⁾ เช่น ความชุก (Prevalence)⁽⁹⁾, อุบัติการณ์ (Incidence)⁽⁷⁾ และอัตราการตาย (Mortality rate)^(3,10-12) โดยนักวิจัยมีวัตถุประสงค์ในการศึกษาแนวโน้มของผลลัพธ์นั้น ๆ และใช้วิธี Joinpoint regression เพื่อการสร้างและเลือกโมเดลหรือสมการในการประมาณค่าที่สอดคล้องกับชุดข้อมูลนั้น ๆ มากที่สุด

ขยายความแนวคิดของ Joinpoint regression

จากสมการ Simple linear regression ซึ่งเป็นสมการในการทำนายค่าตัวแปรตาม (y) หนึ่งตัว ที่เป็นผลมาจากตัวแปรอธิบาย (x) หนึ่งตัว โดยแนวคิดของ Joinpoint regression จะมีการอนุมานว่าในกราฟความสัมพันธ์ระหว่าง ตัวแปรตาม (y) และตัวแปรต้น (x) สามารถแบ่งเส้นความชันออกเป็นช่วงย่อย ๆ และในแต่ละส่วนย่อยนั้นจะมีสมการในการประมาณค่าตามสมการของ simple linear regression model ดังนั้น Joinpoint regression model จะประกอบไปด้วยจุดเชื่อม หรือ Joinpoint จำนวน k จุด และ ส่วนย่อยหรือ segment จำนวน $k+1$ ส่วน ซึ่งในแต่ละ segment จะเชื่อมต่อเนื่องกัน⁽⁵⁾

เมื่อให้ y คือตัวแปรผลลัพธ์

x คือตัวแปรเวลา

τ คือ จุดเวลาที่เป็น Joinpoint

$$E[y(x)] = \begin{cases} f_1(x) & \text{for } a \leq x \leq \tau_1 \\ f_2(x) & \text{for } \tau_1 \leq x \leq \tau_2 \\ \vdots & \\ f_j(x) & \text{for } \tau_{j-1} \leq x \leq \tau_j \\ \vdots & \\ f_{k+1}(x) & \text{for } \tau_k \leq x \leq b \end{cases}$$

เมื่อ a และ b คือช่วงของจุดเวลาที่เก็บข้อมูลตัวอย่างเช่น เมื่อนักวิจัยทำการเก็บข้อมูลเป็นรายปี a คือ ปีแรกที่เก็บข้อมูล และ b คือปีสุดท้ายที่เก็บข้อมูล โดยมี τ_j คือ จุด Joinpoint ซึ่งจะอยู่ระหว่าง a และ b ซึ่งเป็นจุดเชื่อมต่อของแต่ละ Segment และในแต่ละ Segment จะมีสมการสำหรับประมาณค่า y ที่ x มีค่าอยู่ระหว่าง τ_{j-1} และ τ_j ดังนี้⁽⁵⁾

$$E[y(x)] = f_j(x) = \beta_{0,j} + \beta_{1,j}x$$

เมื่อ $\beta_{0,j}$ คือ ค่าจุดตัดที่แกน y ของเส้นสมการใน segment ที่ j

$\beta_{1,j}$ คือ ค่าสัมประสิทธิ์ (Coefficient) หรือ ความชันของเส้นสมการใน segment ที่ j

ซึ่งจำนวนของตัวพารามิเตอร์สำหรับการประมาณค่าใน Joinpoint model คือ $p = 2 \cdot (k+1)$

ถ้าค่า x มีค่าอยู่ระหว่าง $\tau_{k-1} < x \leq \tau_k$ ที่ τ_0 สามารถมีค่าเป็นได้ถึง $-\infty$ และ τ_{k-1} มีค่าเป็นได้ถึง ∞ และสมการในการประมาณค่า หรือ $E[y(x)]$ เป็นสมการประมาณค่าอย่างต่อเนื่อง ตั้งแต่ x_0 จนถึง x_n ดังนั้น

$$f_j(\tau_j) = f_{j+1}(\tau_j)$$

$$\beta_{0,j} + \beta_{1,j}\tau_j = \beta_{0,j+1} + \beta_{1,j+1}\tau_j$$

โดยที่ $j = 1, \dots, j$

ในท้ายที่สุดแล้วทางเลือกในการกำหนดพารามิเตอร์สำหรับการประมาณค่าใน Joinpoint model สามารถเขียนเป็นสมการดังนี้⁽⁵⁾

$$E[y(x)] = \beta_{01} + \beta_{11}x + \delta_1(x - \tau_1)^+ + \delta_2(x - \tau_2)^+ + \dots + \delta_j(x - \tau_j)^+ + \dots + \delta_k(x - \tau_k)^+$$

โดยที่ β_{01} คือ ค่าจุดตัดแกน y , β_{11} คือ ค่าความชันของเส้นสมการ regression ของ segment ที่ 1, δ_j คือ ค่าผลต่างของค่าความชันระหว่าง segment j กับ $j+1$ ($\delta_j = \beta_{1(j+1)} - \beta_{1j}$) และ

$$(x - \tau_j)^+ = \begin{cases} 0 & \text{if } (x - \tau_j) < 0 \\ (x - \tau_j) & \text{otherwise} \end{cases}$$

2) หลักการในการวิเคราะห์ Joinpoint regression ประกอบไปด้วย 2 หลักการ คือ

1) ค้นหาจุดเปลี่ยน (Change point) ของแนวโน้ม (Trend) เป็นการวิเคราะห์เส้นกราฟของผลลัพธ์ในแต่ละจุดเวลา ว่ามีการเปลี่ยนแปลงที่จุดเวลาใด โดยการพิจารณาภายใต้เงื่อนไขของการมีนัยสำคัญทางสถิติ กล่าวคือ เป็นการค้นหาจุดเวลาหรือ x ที่เป็นจุดเชื่อมของแต่ละ Segment ที่ดีที่สุดของ Joinpoint model

โดยใช้วิธีที่เรียกว่า Grid search method⁽¹³⁾ ที่มีการสร้างตารางหรือ Grid สำหรับทุกจุดเวลาหรือทุกค่า x ที่มีความเป็นไปได้ที่จะเป็นจุดเชื่อมหรือ Joinpoint แล้วประมาณค่าโดยสมการ Simple linear regression ของช่วง Segment ที่เป็นไปได้ โดยที่จุดเวลาแรกกับจุดเวลาสุดท้ายไม่สามารถเป็นจุดเชื่อมได้ และจุดเชื่อมที่ 2 ต้องตามหลังจากจุดเชื่อมที่ 1 เสมอ

กรณีตัวอย่าง หากนักวิจัยทำการศึกษาค่าข้อมูลในลักษณะ Time series โดยมีจุดเวลาทั้งหมด 9 จุด ($x_1, \dots, x_9$) จุดเวลาที่มีความเป็นไปได้ที่จะเป็นจุดเชื่อมมีอยู่ 7 จุดคือ $x_2, \dots, x_8$ วิธีการในการค้นหาจุดเชื่อมที่เหมาะสมที่สุดสำหรับ Joinpoint regression model เริ่มจากการค้นหาโดยการกำหนดให้โมเดลนั้นมีจุดเชื่อม 1 จุด โดยสมมติให้แต่ละจุดเวลา (x) เป็นจุดเชื่อมหรือ Joinpoint (τ) ของโมเดล แล้วทำการประมาณค่าด้วยสมการ Joinpoint regression model ที่จะได้โมเดลทั้งหมด 7 โมเดล จากนั้นคำนวณค่าความคลาดเคลื่อนของโมเดลด้วยวิธีผลรวมของความคลาดเคลื่อนกำลังสอง (Sum Squared Error : SSE) โดยโมเดลที่มีค่า SSE น้อยที่สุดจะถูกคัดเลือกไว้ในขั้นต้น หลังจากนั้นจะทำการค้นหาจุดเชื่อมด้วยวิธีเดียวกันแต่จะมีการสมมติเพิ่มจุดเชื่อมเป็น 2 จุด แล้วทำการประมาณค่าและคำนวณค่า SSE เปรียบเทียบคัดเลือกโมเดลที่มีค่า SSE น้อยที่สุด และทำซ้ำจนกระทั่งการกำหนดจุดเชื่อมครบตามจำนวนจุดที่เป็นไปได้ จึงเป็นขั้นตอนที่ทำให้เกิดความยากลำบากของนักวิจัยในการคำนวณ เนื่องจากมีการคำนวณซ้ำ ๆ หลายครั้งจึงเป็นภาระหนักของนักวิจัย⁽⁵⁾

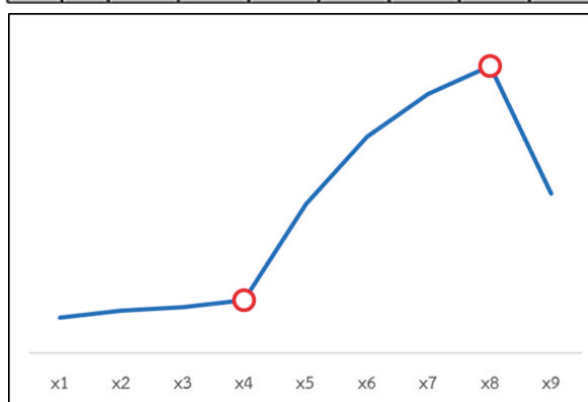
จากรูปภาพเป็นตัวอย่างการค้นหา Joinpoint กรณีกำหนดให้มี 2 จุด ค่าในตารางคือค่า SSE ของโมเดลที่มี Joinpoint 2 จุดที่ต่างกัน จะสังเกตได้ว่า โมเดลที่มีจุดเวลาหรือ x_4 และ x_8 เป็นจุด Joinpoint ที่ 1 (τ_1) และ 2 (τ_2) ตามลำดับ มีค่า SSE น้อยที่สุด ดังนั้นสำหรับ Joinpoint regression model ที่ดีที่สุดสำหรับการมี Joinpoint 2 จุด คือโมเดลที่มีจุดเชื่อมบริเวณจุดเวลาที่ 4 และ 8 ดังแสดงในภาพที่ 1 เมื่อพิจารณาจากกราฟจะพบว่ามี การหักหรือทำมุมของเส้นกราฟ หรือที่เรียกว่า Brake-point ในบริเวณ x_4 และ x_8

2) การคัดเลือกโมเดลที่ดีที่สุด

2.1) permutation test

ทำการทดสอบเปรียบเทียบโมเดลที่ได้มาทั้งหมด หลังจากทำการค้นหา Joinpoint ที่เหมาะสมในแต่ละจำนวน Joinpoint ที่เป็นไปได้แล้ว คำถามต่อมาคือ จำนวน Joinpoint ที่เหมาะสมที่สุดควรมีเท่าไร จึงเป็นที่มาของการเปรียบเทียบโดยใช้วิธีการทางสถิติเข้ามาเพื่อช่วยในการตัดสินใจว่าจำนวน Joinpoint ที่มากที่สุดที่เป็นไปได้หรือที่สอดคล้องกับชุดข้อมูลมากที่สุดควรเป็นเท่าไร⁽¹⁴⁾

		Location of 1st joinpoint						
		1	2	3	4	5	6	7
Location of 2nd joinpoint	2	579						
	3	462	352					
	4	432	389	354				
	5	478	365	397	217			
	6	563	185	238	98	371		
	7	210	147	98	57	172	230	
	8	302	254	231	18	120	345	285



ภาพที่ 1 การหาจุดร่วมด้วยวิธี Grid search method⁽⁵⁾

สถิติที่ใช้ในการทดสอบสมมติฐานในการเปรียบเทียบโมเดลคือ Permutation test โดยใช้ในการเปรียบเทียบระหว่างโมเดลเปล่าหรือ Null model คือโมเดลที่ไม่มี Joinpoint กับโมเดลที่มี Joinpoint หรือเรียกว่า Alternative model โดยการทดสอบนั้นจะเป็นการนำค่า SSE ของ Null model กับ Alternative model มาเปรียบเทียบทำอัตราส่วนกัน ยิ่งค่าอัตราส่วนเข้าใกล้ 1 หมายความว่า 2 โมเดลนั้นไม่แตกต่างกัน หรือ Alternative model ไม่ได้ดีกว่า Null model⁽¹⁵⁾

เพื่อการตัดสินใจที่น่าเชื่อถือบนพื้นฐานของการมีนัยสำคัญทางสถิติ Permutation test ถูกนำมาใช้ในการทดสอบสมมติฐานของการเปรียบเทียบ โดยมีขั้นตอนสมมติฐานสำหรับการเปรียบเทียบดังนี้

$$H_0 : k_a$$

$$H_1 : k_b$$

$$\text{โดย } k_a > k_b$$

ขั้นตอนของการทดสอบนั้นเริ่มต้นจากการเปรียบเทียบระหว่าง Null model กับ Alternative model ที่มีจำนวน Joinpoint มากที่สุด โดยมีสมมติฐานดังนี้

$$H_0 : k = 0$$

$$H_1 : k = \text{Maximum number of joint point}$$

ถ้าผลการทดสอบพบว่า H_0 ถูกปฏิเสธ ในการทดสอบครั้งถัดไปจะเป็นการเปรียบเทียบระหว่าง โมเดลที่มีจำนวน Joinpoint มากที่สุด กับ Model ที่มีจำนวน Joinpoint 1 จุด หรือ

$$H_0 : k = 1$$

$$H_1 : k = \text{Maximum number of joint point}$$

แต่ถ้า ผลการทดสอบพบว่า H_0 ไม่ถูกปฏิเสธ ในการทดสอบครั้งถัดไปจะเป็นการเปรียบเทียบระหว่าง model ที่ไม่มี Joinpoint กับ Model ที่มี Joinpoint มากที่สุดรองลงมา 1 จุด หรือ

$$H_0 : k = 0$$

$$H_1 : k = \text{Maximum number of joint point} - 1$$

ทำการทดสอบเช่นนี้จนในท้ายที่สุดจะได้โมเดลที่ดีที่สุด หรือ โมเดลที่มีจำนวน Joinpoint ที่เหมาะสมที่สุดสำหรับชุดข้อมูลที่ศึกษา^(3,14)

2.2) Bayesian Information Criterion (BIC)

ค่า BIC เป็นวิธีการทางสถิติที่คิดค้นและพัฒนาโดย Gideon E. Schwarz⁽¹⁶⁾ ซึ่งมีความแตกต่างจาก Permutation test เนื่องจาก BIC เป็นการคำนวณโดยคำนึงถึงผลของการเพิ่มขึ้นของตัว Parameter กล่าวคือ วิธี Permutation test จะมุ่งค้นหาโมเดลที่มีจำนวน Joinpoint ที่เหมาะสมที่สุด แต่ BIC จะค้นหาโมเดลที่ดีที่สุดโดยมีการคำนึงถึงผลกระทบของการเพิ่มขึ้นของจำนวนตัวพารามิเตอร์ในสมการด้วย⁽¹⁷⁾ โดยมีแนวคิดว่าการเมื่อทำการสร้างโมเดลที่สอดคล้องกับชุดข้อมูลนั้น ๆ ขึ้นมา มีความเป็นไปได้ที่จะมีการเพิ่มความน่าจะเป็นของโมเดลด้วยการเพิ่มตัวพารามิเตอร์ ซึ่งการกระทำนี้อาจส่งผลในเรื่องของการเกิด Over-fitting วิธีการของ BIC จึงทำการเพิ่มในส่วนที่เรียกว่า penalty term เข้ามาในการคำนวณ ซึ่งมีการนำเข้าไปในส่วนของการคำนวณตัวพารามิเตอร์เข้ามาในสูตรคำนวณของ BIC ซึ่งเป็นการแก้ปัญหาในการที่จะเกิด Over-fitting ของโมเดล นอกจากนี้ค่า BIC ยังมีการตอบสนองต่อการเปลี่ยนแปลงเล็กน้อยของแนวโน้มได้ดีกว่า และเป็นวิธีที่เหมาะสมอีกวิธีการหนึ่ง

เมื่อให้ คือจำนวน Joinpoint ที่สะท้อนให้เห็นในจำนวนของตัวพารามิเตอร์ BIC จะคำนึงถึงความซับซ้อนของโมเดลในการควบคุมการเกิด over-fitting⁽¹⁸⁾ และเป็นวิธีที่สมเหตุสมผลที่จะใช้ในกรณีที่มีข้อจำกัดเกี่ยวกับจำนวน⁽¹⁴⁾ โดยสูตรการคำนวณค่า BIC ดังนี้

$$BIC_k = \ln \frac{Q_k}{n} + \frac{2k + 2}{n} \cdot \ln(n)$$

สำหรับโมเดลที่มีค่า BIC น้อยที่สุดจะถูกพิจารณาเป็นโมเดลที่ดีที่สุด⁽⁵⁾

3) สถิติที่เกี่ยวข้องกับ Joinpoint regression

สำหรับการวิเคราะห์ Joinpoint regression จะทำให้ผู้วิจัยได้ข้อมูลที่เป็นประโยชน์ในการนำไปใช้มากยิ่งขึ้น เนื่องจากได้โมเดลในการทำนายที่มีความสอดคล้องกับข้อมูลจริงมากขึ้นในรายละเอียดของการเปลี่ยนแปลงและขนาดของการเปลี่ยนแปลงในแต่ละจุดเวลา อย่างไรก็ตามโมเดลที่ได้จากการวิเคราะห์นั้นอาจทำให้เกิดความยุ่งยากในการอธิบายให้ผู้อ่านเข้าใจ ดังนั้นจึงมีการคำนวณค่าสถิติที่ช่วยในการสะท้อนให้เห็นถึงการเปลี่ยนแปลงในแต่ละช่วงของการเปลี่ยนแปลงหรือ segment ระหว่าง 2 Joinpoint ที่เรียกว่า ร้อยละการเปลี่ยนแปลงต่อปี (Annual percent change : APC) และค่า ค่าเฉลี่ยร้อยละการเปลี่ยนแปลงต่อปี (Average annual percent change : AAPC)⁽⁸⁾ ในการสะท้อนให้เห็นการเปลี่ยนแปลงในภาพรวมของข้อมูลชุดนั้น โดยรายละเอียดการคำนวณดังต่อไปนี้

3.1 ร้อยละการเปลี่ยนแปลงต่อปี (Annual percent change : APC)

เป็นค่าสถิติที่ใช้ในการอธิบายการเปลี่ยนแปลงของเส้นแนวโน้มย่อย ๆ หรือ Segment แต่ละ Segment ซึ่งจะมีค่า APC แตกต่างกันไป ค่านี้จะใช้อธิบายเพื่อให้ผู้อ่านเข้าใจขนาดของความเปลี่ยนแปลงของแนวโน้มที่ได้จากการวิเคราะห์ Joinpoint regression เป็นการหาค่าร้อยละการเปลี่ยนแปลงระหว่าง Joinpoint ที่ 1 กับ Joinpoint ที่ 2 ซึ่งอยู่ติดกัน ในการประมาณค่า APC สามารถคำนวณโดยการปรับสมการ Linear regression ที่ได้ให้อยู่ในรูปของ Logarithm แล้วคำนวณหาค่าดังนี้^(12,19)

$$\log(y_x) = b_0 + b_1x \text{ และ } y_x = e^{b_0+b_1x}$$

$$\log(y_{x+1}) = b_0 + b_1(x+1) \text{ และ } y_{x+1} = e^{b_0+b_1(x+1)}$$

โดย $\log(y_x)$ และ $\log(y_{x+1})$ คือค่า natural logarithms ของอัตราอุบัติการณ์ในปีที่ x และ $x+1$ APC จากปีที่ y ไปถึงปีที่ $y+1$ มีค่าเท่ากับ

$$= \left[\frac{y_{x+1} - y_x}{y_x} \right] \times 100$$

$$= \left[\frac{e^{b_0+b_1(x+1)} - e^{b_0+b_1x}}{e^{b_0+b_1x}} \right] \times 100$$

$$= (e^{b_1} - 1) \times 100$$

การแปลผลค่าร้อยละการเปลี่ยนแปลงต่อปี จากแนวคิดข้างต้น ค่า APC จะเป็นค่าที่บอกถึงร้อยละของการเปลี่ยนแปลงจากจุดที่ 1 ไปยังจุดที่ 2 กล่าวคือ นำค่าจาก 2 จุดมาลบกัน แล้วเปรียบเทียบจากค่าเดิม (ค่าของจุดที่ 1) ว่าคิดเป็นร้อยละเท่าไร

ดังนั้น ค่า APC จะเป็นค่าร้อยละการเปลี่ยนแปลงของช่วงเวลา จากจุดเวลาที่ 1 ไปยังจุดเวลาที่ 2 ดังภาพที่ 2 จะเห็นว่า จุดเวลาที่ 1 คือ ค.ศ. 2001 และจุดเวลาที่ 2 คือ ค.ศ. 2010 ซึ่งมีค่า APC

เป็นร้อยละ 4 และเป็นค่าบวก สมมติให้ค่าตัวแปรที่ศึกษาเป็น อุบัติการณ์โรคเรื้อรัง อธิบายได้ว่า ในช่วงระหว่างปี ค.ศ.2019-2023 ค่าอุบัติการณ์โรคเรื้อรังนั้นเพิ่มขึ้นร้อยละ 4

Variable	Parallelism test of slopes		Trend 1			Trend 2			APC in 5 recent years	
	Test result (APC)	P value	Year	APC	Confidence interval**	Year	APC	Confidence interval**	APC	Confidence interval**
Age, (Years)										
30-49 and 50-69	Not different	0.2951	2001-2010	4*	(2.7, 5.4)				4*	(2.7, 5.4)

ภาพที่ 2 การรายงานค่าร้อยละการเปลี่ยนแปลงต่อปี⁽⁷⁾

3.2 ค่าเฉลี่ยร้อยละการเปลี่ยนแปลงต่อปี (Average annual percent change : AAPC)

ในการเปลี่ยนแปลงของของอุบัติการณ์ที่ได้จากการวิเคราะห์ Joinpoint regression ในบางกรณีอาจมีได้มากกว่า 1 ความชัน ซึ่งอาจเป็นความชันที่แตกต่างกันหรืออาจตรงกันข้ามกัน ดังนั้น ในการคำนวณค่า APC จะต้องใช้วิธีการคำนวณอีกรูปแบบหนึ่ง คือ average annual percent change (AAPC) ซึ่งเป็นการคำนวณโดยที่แบ่งช่วงของกราฟที่มีความชันแตกต่างกันเป็น ช่วง ๆ เพื่อคำนวณ APC ของแต่ละช่วงแล้วนำมารวมกัน ถ่วงน้ำหนักตามจำนวนปีของแต่ละช่วงเวลา^(12,20) ดังสมการต่อไปนี้

$$APC_i = (e^{b_i} - 1) \times 100$$

$$AAPC = \left\{ e^{\left(\frac{\sum w_i APC_i}{\sum w_i} \right)} - 1 \right\} \times 100$$

ตัวอย่างเช่น เมื่อนักวิจัยได้ทำการศึกษ้อัตราป่วยด้วยโรค A ระหว่างปี ค.ศ. 1999-2015 จากการวิเคราะห์ Joinpoint regression พบว่ามี Joinpoint เกิดขึ้นจำนวน 2 จุด ที่ ปี ค.ศ. 2005 และ ค.ศ. 2009 ดังนั้นสามารถแบ่ง segment ได้ 3 ส่วน คือ 1999-2005, 2005-2009 และ 2009-2015 และมีค่า APC เป็น -0.014, -0.032 และ -0.012 ตามลำดับ และจำนวนปีใน

แต่ละช่วง () เป็น 6,4 และ 6 ปี สูตรคำนวณ AAPC จะเป็นดังนี้

$$AAPC = \left[e^{\frac{-6 \times 0.014 - 4 \times 0.032 - 6 \times 0.012}{6+4+6}} - 1 \right] \times 100$$

$$AAPC = \left[e^{\frac{-0.284}{16}} - 1 \right] \times 100$$

$$AAPC = [e^{-0.01775} - 1] \times 100$$

$$AAPC = [0.982 - 1] \times 100 = -1.8\%$$

การแปลผลค่า AAPC จากภาพที่ 3 จะเห็นว่ามี 2 เส้น แนวโน้ม (Trend) หรือ 2 Segment ค่า AAPC จะเป็นค่าที่แสดงถึงแนวโน้มภาพรวมทั้งหมดจากทุก Segment ทำให้เห็นภาพรวมของแนวโน้มทั้งหมดของกราฟว่ามีลักษณะเป็นอย่างไร แม้ว่ากราฟจะถูกแบ่งเป็น 2 ช่วงที่มีค่า APC ต่างกัน โดยช่วงแรก(ค.ศ.2001-2008)มีค่าเป็นบวกแสดงให้เห็นว่ามีการเพิ่มขึ้น แต่ช่วงที่ 2 (ค.ศ.2008-2010) กลับมีค่าติดลบแสดงให้เห็นว่าการลดลงจากช่วงแรก อย่างไรก็ตามเมื่อพิจารณาค่า AAPC จะเห็นว่าโดยรวมแล้วแนวโน้มของกราฟยังมีทิศทางเพิ่มขึ้นอยู่โดยค่าเฉลี่ยอยู่ที่ร้อยละ 0.8 อธิบายได้ว่า ในช่วงที่ผ่านมาตั้งแต่ ค.ศ.2001-2010 กราฟมีค่าเพิ่มขึ้นและลดลงในบางช่วง แต่ในภาพรวมแล้วยังมีทิศทางเพิ่มขึ้นร้อยละ 0.8

Variable	Parallelism test of slopes		Trend 1			Trend 2			APC in 5 recent years	
	Test result (APC)	P value	Year	APC	Confidence interval**	Year	APC	Confidence interval**	APC	Confidence interval**
Tumor grade										
Well and moderately differentiated	Not different	0.505	2001-2008	14.7*	(9.8, 19.8)	2008-2010	-11.5	(-31, 13.5)	0.8	(-10.1, 13)

ภาพที่ 3 การรายงานค่าเฉลี่ยร้อยละการเปลี่ยนแปลงต่อปี⁽⁷⁾

4) การนำวิธีการ Joinpoint regression ไปใช้ในการศึกษาวิจัยสถานการณ์หรือแนวโน้มปัญหาทางด้านวิทยาศาสตร์สุขภาพ

ปัจจุบันมีการนำวิธีการ Joinpoint regression ไปใช้ในการศึกษาวิจัยและการวิเคราะห์ข้อมูลทางด้านสุขภาพเป็นจำนวนมากขึ้น และมีความหลากหลายในเชิงเนื้อหาของปัญหา

สุขภาพ ไม่ว่าจะเป็นการวิเคราะห์แนวโน้มของโรคมะเร็ง^(2,11) โรคกระดูกพรุน⁽¹²⁾ การฆ่าตัวตาย⁽²¹⁾ หรือแม้กระทั่งโรคติดเชื้อโคโรนาไวรัส 2019^(6,8, 22) ตัวอย่างต่อไปนี้เป็นงานวิจัยที่นำวิธีการวิเคราะห์ Joinpoint regression ไปใช้โดยจะอธิบายตัวอย่างการนำไปใช้และการแปลผลการวิเคราะห์เพื่อให้ผู้อ่านได้ทำความเข้าใจมากขึ้น

ตัวอย่างการวิจัยที่มีการประยุกต์ใช้ Joinpoint regression ในการวิเคราะห์การเปลี่ยนแปลงอุบัติการณ์ของโรคมะเร็งเต้านม ในสตรีอายุ 30-69 ปี ระหว่างปี ค.ศ. 2001-2010⁽⁷⁾ จากภาพจะ เห็นได้ว่านักวิจัยได้ทำการศึกษาโดยวิเคราะห์ข้อมูลจำแนกตาม กลุ่มที่สนใจ มีการนำเสนอค่า APC ของแต่ละช่วงเพื่ออธิบาย การเปลี่ยนแปลงของอัตราอุบัติการณ์ของโรคมะเร็งเต้านม จาก ภาพตารางจะเห็นได้ว่านักวิจัยนำเสนอค่าผลลัพธ์เป็น 2 ช่วง

(Trend) เนื่องจากนักวิจัยพบว่าโมเดลที่ดีที่สุดมี 1 Joinpoint คือเกิดขึ้นที่ปี ค.ศ.1387 ซึ่งเป็นปีที่ถูกวินิจฉัย ดังแสดงในภาพที่ 4 การแปลผลการวิเคราะห์ข้อมูลมีการอธิบายว่า ค่า APC ของกลุ่ม well-differentiated and poorly differentiated tumor grades มีการเพิ่มขึ้นร้อยละ 14.5 ในช่วงแรก(2001-2008) และลดลงร้อยละ 23.2 ในช่วงถัดมา (2008-2010)

Variable	Parallelism test of slopes		Trend 1			Trend 2			APC in 5 recent years	
	Test result (APC)	P value	Year	APC	Confidence interval**	Year	APC	Confidence interval**	APC	Confidence interval**
Age, (Years)										
30-49 and 50-69	Not different	0.2951	2001-2010	4*	(2.7, 5.4)				4*	(2.7, 5.4)
Tumor size (cm)										
≤2 and >2 and ≤5	Different	0.0022							2.8*	(19.8, 7.4)
≤2 and >5	Different	0.0049							39.7*	(19.7, 52.1)
≤2 and >5 and >5	Different	0.0491							11.7*	(3.4, 20.5)
Tumor grade										
Well and moderately differentiated	Not different	0.505	2001-2008	14.7*	(9.8, 19.8)	2008-2010	-11.5	(-31, 13.5)	0.8	(-10.1, 13)
Well and poorly differentiated	Not different	0.526	2001-2008	14.5*	(7.9, 21.5)	2008-2010	-23.2	(-47.1, 5.4)	-6.2	(-21.1, 2.7)
Moderately and poorly differentiated	Not different	0.289	2001-2008	15.6*	(10.3, 21.3)	2008-2010	-16.6	(-36.8, 10)	-7	(-21.4, 10)
Estrogen receptor status										
Negative and positive	Not different	0.069	2001-2010	8.2*	(5.6, 10.9)				8.2*	(5.6, 10.9)
Progesterone receptor status										
Negative and positive	Not different	0.4284	2001-2010	8.1*	(5.1, 8.5)				8.1*	(5.1, 8.5)
Proportion of positive lymph nodes to lymph node surgery (%)										
≤25 and >25	Not different	0.4667	2001-2008	3.8*	(1.2, 6.4)	2008-2010	-11.9	(-25.8, 4.8)	-4.4	(-11.6, 3.5)

*APC and AAPC are significantly different from zero ($P < 0.05$). **Percent, APC = Annual percent change, AAPC = Average annual percent change

ภาพที่ 4 การวิเคราะห์ Joinpoint regression และการเปรียบเทียบค่าเฉลี่ยร้อยละการเปลี่ยนแปลงต่อปีตามลักษณะของเนื้องอก⁽⁷⁾

จากภาพจะเห็นได้ว่านักวิจัยได้ทำการศึกษาโดยวิเคราะห์ ข้อมูลจำแนกตามกลุ่มที่สนใจ มีการนำเสนอค่า APC ของแต่ละ ช่วงเพื่ออธิบายการเปลี่ยนแปลงของอัตราอุบัติการณ์ของโรคมะเร็งเต้านม จากภาพตารางจะเห็นได้ว่านักวิจัยนำเสนอค่า ผลลัพธ์เป็น 2 ช่วง(Trend) เนื่องจากนักวิจัยพบว่าโมเดลที่ดีที่สุด มี 1 Joinpoint คือเกิดขึ้นที่ปี ค.ศ.1387 ซึ่งเป็นปีที่ถูกวินิจฉัย ดังแสดงในภาพที่ 5

นอกจากนี้นักวิจัยมีการนำเสนอสมการสำหรับการ ประมาณค่าโดยการแสดงการแทนค่าในตัวพหุนามดีกรีต่าง ๆ ที่เป็นองค์ประกอบของสมการ ดังแสดงในภาพที่ 6 พิจารณา เฉพาะในกลุ่ม Well differentiated จะได้ว่า

$$\beta_{01} = -187.43$$

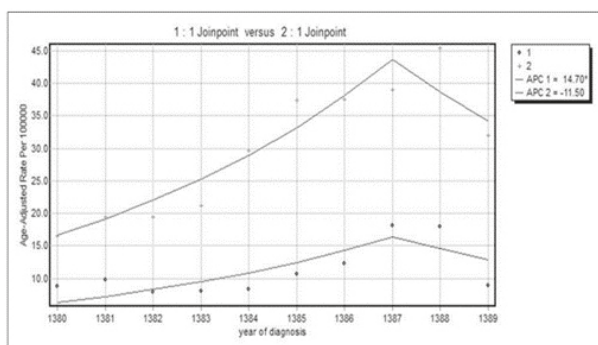
$$\beta_{11} = 0.137$$

$$\delta = -0.259$$

$$\tau = 1387$$

$$y(x) = \beta_{01} + \beta_{11}x + \delta_1(x - \tau_1)^+$$

$$y(x) = -187.43 + 0.137x - 0.259(x - 1387)^+$$



ภาพที่ 5 กราฟแสดงการเปลี่ยนแปลงของอัตราอุบัติการณ์มะเร็งเต้านม⁽⁷⁾

Variable	Classification	Parameter	Estimated parameter
Tumor grade			
		Intercept (group 1)	-187.428
	Group 1: Well differentiated	Intercept (group 2)	-186.451
	Group 2: Moderately differentiated	Slope 1 (combination)	0.137
		Slope 2-slope 1	-0.259
		Intercept (group 1)	-185.094
	Group 1: Well differentiated	Intercept (group 3)	-184.621
	Group 3: Poorly differentiated	Slope 1 (combination)	0.135
		Slope 2-slope 1	-0.398
		Intercept (group 2)	-197.651
	Group 2: Moderately differentiated	Intercept (group 3)	-198.152
	Group 3: Poorly d differentiated	Slope 1 (combination)	0.145
		Slope 2-slope 1	-0.327

*Significantly different from zero ($P < 0.05$)

ภาพที่ 6 การประมาณค่าค่าสัมประสิทธิ์การถดถอยของการเปรียบเทียบค่าเฉลี่ยร้อยละการเปลี่ยนแปลงต่อปี⁽⁷⁾

ตัวอย่างการวิจัยที่ศึกษาในผลลัพธ์เกี่ยวกับผู้ป่วยที่เสียชีวิตด้วยโรคมะเร็งตับ (12) นักวิจัยได้ทำการวิเคราะห์แยกเป็นชาย-หญิงตามกลุ่มอายุที่ต่างกันทุก ๆ 5 ปี มีการนำเสนอค่า APC และ AAPC เพื่ออธิบายสรุปการเปลี่ยนแปลง เมื่อพิจารณาในกลุ่มอายุ 50-54 ปี เพศหญิง พบว่ามีจุด Joinpoint ทั้งหมด 2 จุด คือ ปี ค.ศ. 2005 และปี ค.ศ. 2009 และแบ่ง Segment ได้ 3 ช่วง ดังนี้ 1999-2005,

2005-2009 และ 2009-2015 โดย อัตราการเสียชีวิตด้วยโรคมะเร็งตับในช่วงแรก(1999-2005) ลดลงร้อยละ 1.4 ต่อมาในช่วงที่ 2 (2005-2009) ลดลงร้อยละ 3.2 และในช่วงสุดท้าย (2009-2015) ลดลงอีกร้อยละ 1 และค่า AAPC แสดงให้เห็นว่าตลอดช่วงเวลาปี ค.ศ. 1999-2015 อัตราการเสียชีวิตด้วยโรคมะเร็งตับลดลงเฉลี่ยร้อยละ 1.7 ต่อปี ดังแสดงในภาพที่ 7

Age group, yrs	Period	Women			Period	Men		
		Change year	APC (95% CI)	AAPC (95% CI)		Change year	APC (95%CI)	AAPC (95% CI)
50-54	1999-2015			-1.7 (-2.3, -1.1)	1999-2015			-1.8 (-2.3, -1.3)
	1999-2005	2005	-1.4 (-2.1, -0.7)		1999-2005	2005	-1.4 (-2.1, -0.8)	
	2005-2009	2009	-3.2 (-5.2, -1.1)		2005-2009	2009	-3.2 (-5.0, -1.3)	
	2009-2015		-1.0 (-1.8, -0.3)		2009-2015		-1.2 (-1.8, -0.5)	
55-59	1999-2015	None	+1.2 (-3.5, +6.2)	+1.2 (-3.5, +6.2)	1999-2015	None	-1.5 (-3.8, +1.0)	-1.5 (-3.8, +1.0)
60-64	1999-2015			-0.8 (-1.3, -0.3)	1999-2015	None	-5.1 (-8.9, -1.2)	-5.1 (-8.9, -1.2)
	1999-2002	2002	+2.9 (+0.6, +5.1)					
	2002-2009	2009	-3.4 (-4.1, -2.7)					
	2009-2015		+0.4 (-0.3, +1.2)					
65-69	1999-2015	None	-2.5 (-7.7, +3.1)	-2.5 (-7.7, +3.1)	1999-2015	None	+0.5 (-5.5, +6.9)	+0.5 (-5.5, +6.9)
70-74	1999-2015	None	-4.8 (-10.8, +1.6)	-4.8 (-10.8, +1.6)	1999-2015	None	-0.4 (-4.3, +3.6)	-0.4 (-4.3, +3.6)
75-79	1999-2015	None	-4.7 (-5.7, -3.7)	-4.7 (-5.7, -3.7)	1999-2015	None	-4.2 (-5.2, -3.1)	-4.2 (-5.2, -3.1)
0-84	1999-2015			-4.4 (-5.9, -3.0)	1999-2015			-2.4 (-6.5, +1.8)
	1999-2004	2004	+0.4 (-4.0, +5.0)		1999-2007	2007	+0.1 (-3.4, 3.6)	
	2004-2015		-6.5 (-7.9, -5.2)		2007-2011	2011	-13.0 (-25.7, +1.8)	
					2011-2015		+4.0 (-5.7, +14.7)	
85+	1999-2015	None		-3.4 (-4.2, -2.7)	1999-2015	None		-2.90 (-3.8, -2.0)

ภาพที่ 7 แนวโน้มอัตราการเสียชีวิตจากโรคมะเร็งตับในผู้หญิงและผู้ชายชาวสเปน⁽¹²⁾

จากตัวอย่างที่ได้นำมาแสดงทำให้เห็นว่าการวิเคราะห์ข้อมูลด้วยวิธี Joinpoint regression ช่วยให้นักวิจัยได้ข้อมูลที่ชัดเจนมากยิ่งขึ้นเกี่ยวกับการศึกษาแนวโน้มของผลลัพธ์ที่สนใจ สามารถช่วยให้นักวิจัยทราบถึงการเปลี่ยนแปลงของแนวโน้มว่าเกิดขึ้นที่จุดเวลาใด และสามารถอธิบายได้ว่าในแต่ละช่วงของ

การเปลี่ยนแปลงที่เกิดขึ้นนั้นมีการเพิ่มขึ้นหรือลดร้อยละเท่าใด นอกจากนี้ยังสามารถอธิบายภาพรวมของการเปลี่ยนแปลงทั้งหมดตลอดระยะเวลาที่ทำการศึกษาได้ว่าการเปลี่ยนแปลงเพิ่มขึ้นหรือลดลงเป็นเท่าใด อย่างไรก็ตามนักวิจัยที่มีความสนใจในการวิเคราะห์ข้อมูลด้วย Joinpoint regression ควรมีการศึกษา

รายละเอียดของวิธีการรวมไปถึงข้อควรระวังของการใช้งานอย่างรอบคอบเพื่อลดความผิดพลาดในการวิเคราะห์

5) ข้อพึงระวังในการใช้ Joinpoint regression

5.1 โมเดลที่ดีที่สุดสำหรับการวิเคราะห์ Joinpoint regression ไม่ใช่โมเดลที่ประกอบด้วยจุด Joinpoint มากที่สุด จากที่ได้กล่าวไปในหัวข้อการคัดเลือกโมเดลที่ดีที่สุด จะเห็นได้ว่าข้อพึงระวังที่ควรให้ความสำคัญประการหนึ่งคือ เมื่อโมเดลประกอบด้วยจุด Joinpoint จำนวนมากขึ้นจะทำให้เพิ่มตัวพารามิเตอร์ในสมการ

มากขึ้น จึงทำให้มีโอกาสที่จะเกิดปรากฏการณ์ที่เรียกว่า Over-fitting คือข้อผิดพลาดหนึ่งในการสร้างโมเดลทางสถิติที่ใกล้เคียงกับชุดข้อมูลมากเกินไปส่งผลให้โมเดลที่ได้มาไม่ประโยชน์ในการอ้างอิงสำหรับชุดข้อมูลตั้งต้นเท่านั้นไม่สามารถใช้อ้างอิงในชุดข้อมูลอื่นได้⁽²³⁾ ดังนั้นนักวิจัยควรมีการพิจารณาเลือกโมเดลอย่างเหมาะสมโดยการตัดสินใจเลือกบนพื้นฐานทางสถิติ ตัวอย่างดังแสดงในภาพที่ 8

Standard Parameter						
Cohort	Parameter	Parameter estimates	Standard Error	Z value	Probability	MSE
ISE30	Constant 1	4,683081	0,011339	412,992596	0,000000	0,00008
ISE30	Slope 1	-0,014032	0,004141	-3,388868	0,002526	
ISE30	Slope 2- Slope 1	0,020296	0005071	4,002180	0,000560	
ISE30	Slope 3- Slope 2	-0,014833	0,013417	-1,105525	0,280363	
ISE30	Slope 4- Slope 3	0,013246	0,013417	0,987249	0,333793	
ISE30	Slope 5- Slope 4	0,012979	0,013417	0,967395	0,343410	
ISE30	Slope 6- Slope 5	-0,015035	0,013116	-1,146305	0,263456	
ISE30	Slope 7- Slope 6	-0,013491	0,001913	-7,050813	0,000000	
Test Statistics						
Cohort	Model	Number of Joinpoint	Number of Parameter	Degrees of Freedom	SSE	BIC
ISE 30	#1	0 Joinpoint	2	41	0,0317903	-7,0348560
ISE 30	#2	1 Joinpoint	4	39	0,0165874	-7,5104305
ISE 30	#3	2 Joinpoint(s)	6	37	0,0077217	-8,1001004
ISE 30	#4	3 Joinpoint(s)	8	35	0,0056628	-8,2352841
ISE 30	#5	4 Joinpoint(s)	10	33	0,0036137	-8,5095150
ISE 30	#6	5 Joinpoint(s)	12	31	0,0029564	-8,5353500
ISE 30	#7	6 Joinpoint(s)	14	29	0,0024110	-8,5643233 *
ISE 30	#8	7 Joinpoint(s)	16	27	0,0022394	-8,4632364
ISE 30	#9	8 Joinpoint(s)	18	25	0,0020562	-8,3736408
ISE 30	#10	9 Joinpoint(s)	20	23	0,0018726	-8,2922279
The selected final model has 6 joinpoint.						

ภาพที่ 8 ผลลัพธ์ของการวิเคราะห์ตัวแบบ Joinpoint Regression⁽²⁴⁾

จากภาพจะเห็นได้ว่ามีโมเดลทั้งหมด 10 โมเดลด้วยกัน ซึ่งประกอบไปด้วยโมเดลที่มี Joinpoint ตั้งแต่ 0-9 จุด เมื่อพิจารณาจากค่า SSE จะเห็นว่าโมเดลที่ 10 จะมีค่า SSE น้อยที่สุด นั่นคือโมเดลนี้มีค่าใกล้เคียงกับชุดข้อมูลตั้งต้นมากที่สุดสอดคล้องกับจำนวนตัวพารามิเตอร์ที่มากที่สุดเช่นเดียวกัน ในขณะที่เดียวกันเมื่อพิจารณาค่า BIC จะพบว่า โมเดลที่มีค่า BIC น้อยที่สุดคือ โมเดลที่ 7 ซึ่งประกอบไปด้วย 6 Joinpoint และ 7 Slope หรือ Segment แม้ว่าโมเดลที่ 10 จะมีค่า SSE น้อยที่สุด แต่เนื่องจากเงื่อนไขของความเสถียรที่จะเกิด Over-fitting นักวิจัยจึงพิจารณาเลือกโมเดลที่ดีที่สุดจากการพิจารณาค่า BIC ที่น้อยที่สุด ดังนั้นโมเดลที่ 7 จึงเป็นโมเดลที่ถูกเลือกในครั้งนี้⁽²⁴⁾

5.2 Joinpoint regression ไม่สามารถใช้ได้กับทุกชุดข้อมูล ความเข้าใจขั้นพื้นฐานในการเลือกใช้สถิติแต่ละตัวนี้นักวิจัย

จำเป็นต้องคำนึงถึงคุณลักษณะของข้อมูลที่มีว่าสอดคล้องกับข้อตกลงเบื้องต้นของสถิติตัวนั้น ๆ หรือไม่ สำหรับ Joinpoint regression เช่นเดียวกัน โดยข้อจำกัดของ Joinpoint regression คือ หากชุดข้อมูลที่นำมาวิเคราะห์มีความแตกต่างกันอย่างมากทางด้านของขนาดผลลัพธ์ เช่น อัตราอุบัติการณ์แตกต่างกันมากในแต่ละปี ผลการวิเคราะห์ที่ได้จะค่อนข้างไม่เสถียร^(25,26) ซึ่งมีเหตุมาจากประชากรมีขนาดเล็ก หรือ จุดข้อมูล หรือกล่าวคือจำนวนปีที่นำมาวิเคราะห์มีขนาดไม่เพียงพอ การเปลี่ยนแปลงที่รุนแรงทั้งขนาดหรือทิศทางของแนวโน้ม วิธีการนี้อาจล้มเหลวในการประมาณการระยะยาวหรือแม้กระทั่งในการประมาณการระยะสั้น⁽²⁾ ดังนั้น Joinpoint regression จึงเหมาะสำหรับการวิเคราะห์ข้อมูลระยะยาว⁽²⁷⁾

5.3 สิ่งที่นักวิจัยพึงระวัง⁽²⁷⁾

1) เมื่อกลุ่มตัวอย่างไม่เพียงพอ การรวมจุดข้อมูลหรือจุดเวลาเข้าด้วยกัน เช่น การรวมข้อมูลจาก 3 จุดเป็นจุดเดียว อาจจะเป็นวิธีการทางสถิติทางหนึ่งและสามารถช่วยเพิ่มขนาดตัวอย่างได้ แต่อย่างไรก็ตาม จำนวนของจุดข้อมูลที่น่ามารวมกันแต่ละครั้งควรจะมีจำนวนน้อยที่สุดเท่าที่จะเป็นไปได้ เพื่อให้จุด Joinpoint ที่อาจเกิดขึ้นนั้นสามารถถูกค้นพบได้ การทดสอบสมมติฐานมีแนวโน้มที่ผลลัพธ์จะแตกต่างออกไป แม้ว่าความแปรปรวนของแนวโน้ม หรือ ความชันของข้อมูลรวมจะคล้ายกัน

2) Examination model fitting เป็นสิ่งที่สำคัญ แต่มักถูกมองข้าม ซึ่งวิธีการนี้เป็นการทดสอบโมเดลที่ได้มาว่าสามารถทำนายค่าผลลัพธ์ได้ใกล้เคียงค่าจริงมากน้อยเพียงใด โดยการนำค่า $\hat{y}$ ที่ได้จากการคำนวณตามสมการมาเปรียบเทียบกับค่า $y - \hat{y}$ จากข้อมูลจริง นั่นคือ เรียกว่า Residuals ถ้าโมเดลที่ได้มาเป็นโมเดลที่ดีค่า Residuals จะต้องมีความน้อย นั่นค่า ค่าจริงกับค่าที่ทำนายมีความแตกต่างกันน้อย ซึ่งอีกวิธีการหนึ่งที่จะช่วยในการทดสอบโมเดลคือ การทำ Residual plot ซึ่งเป็นการนำค่า Residual มา plot ในกราฟร่วมกับค่า $\hat{y}$ แล้วสังเกตรูปร่างของกราฟ โดยรายละเอียดสามารถศึกษาเพิ่มเติมได้จากหนังสือ EPIDEMIOLOGY STUDY DESIGN AND DATA ANALYSIS⁽²⁸⁾

สรุป

ปัจจุบันวิธีการ Joinpoint regression ได้รับความสนใจในการนำไปใช้ในงานวิจัยทางด้านวิทยาศาสตร์สุขภาพมากขึ้น ทั้งในต่างประเทศและประเทศไทย โดยลักษณะของงานวิจัยส่วนใหญ่ที่นำไปใช้จะเป็นในลักษณะของการศึกษาสถานการณ์ปัญหาสุขภาพต่าง ๆ เช่น อุบัติการณ์หรือแนวโน้มของการเกิดโรค แนวโน้มปัญหาสุขภาพต่าง ๆ มักพบในกลุ่มงานวิจัยที่ศึกษาเกี่ยวกับอุบัติการณ์ของโรคเรื้อรังโดยเฉพาะทางด้านมะเร็ง เนื่องจากเป็นโรคที่มีการเก็บรวบรวมข้อมูลอย่างต่อเนื่องในลักษณะของข้อมูลระยะยาวและมีคุณลักษณะของข้อมูลที่สามารถทำการวิเคราะห์ด้วยวิธีนี้ได้ อย่างไรก็ตาม วิธีการ Joinpoint regression ยังสามารถนำไปใช้กับข้อมูลสุขภาพด้านอื่น ๆ ได้ ไม่ว่าจะเป็นในเชิงบวกหรือเชิงลบ ขึ้นอยู่กับลักษณะข้อมูลนั้น ๆ ภายใต้เงื่อนไขความเหมาะสมของข้อมูล เช่น ช่วงระยะเวลาของข้อมูล ต้องเป็นข้อมูลที่เป็นลักษณะข้อมูลระยะยาว จำนวนปีหรือจุดเวลา มากพอ การรวบรวมข้อมูลอย่างมีคุณภาพ จะทำให้ได้มาซึ่งผลการวิเคราะห์ที่ถูกต้องและมีประโยชน์ ด้วยผลการวิเคราะห์ที่สามารถบ่งชี้ถึงจุดเปลี่ยนของกราฟอย่างมีนัยสำคัญ ผู้วิจัยจึงนำเสนอประโยชน์ในประเด็นนี้ไปใช้เพื่อการวิเคราะห์หรืออภิปรายสถานการณ์นั้นเพื่อให้ได้ข้อมูลที่เป็นประโยชน์ในเชิงนโยบาย กล่าวคือ เมื่อผู้วิจัยพบจุดเปลี่ยนของกราฟสามารถอภิปรายได้ว่าการเปลี่ยนแปลงของกราฟนั้นเกิดจากการเปลี่ยนแปลงนโยบายทางด้านสาธารณสุขจึงทำให้กราฟ

นั้นเปลี่ยนแปลง ซึ่งสามารถใช้ได้ทั้งในทางบวกหรือทางลบก็ได้ ยกตัวอย่างเช่น กราฟอุบัติการณ์ของโรคเรื้อรังมีการเปลี่ยนแปลงพบจุดเปลี่ยนเกิดขึ้นในปี พ.ศ. 2560 โดยเส้นกราฟพุ่งลงหลังจากปี พ.ศ.2560 อาจเป็นไปได้ว่ากระทรวงสาธารณสุขมีการเพิ่มมาตรการในการคัดกรองมากขึ้น ทำให้ประชาชนได้รับการคัดกรองมากขึ้น ถูกคัดกรองเร็วขึ้น (คัดกรองตั้งแต่อายุน้อย) ได้รับการแนะนำและปรับเปลี่ยนพฤติกรรมมากขึ้น จึงทำให้ประชาชนเป็นโรคเรื้อรังน้อยลง หรือหากพิจารณาในเชิงลบ ตัวอย่างเช่น ใช้ประเมินนโยบายหรือมาตรการที่ผ่านมามีประสิทธิภาพเพียงพอหรือไม่ เช่น ประเมินการให้วัคซีนป้องกันโรค หากพบว่ากราฟมีแนวโน้มคงที่หรือมีจุดเปลี่ยนของกราฟโดยที่เส้นกราฟพุ่งขึ้นไปแสดงให้เห็นว่านโยบายที่ผ่านมานั้นอาจไม่ได้มีประสิทธิภาพมากพอ เหล่านี้สามารถใช้เป็นข้อเสนอในเชิงนโยบายได้ อย่างไรก็ตาม ผู้วิจัยยังคงต้องพึงระวังในการอธิบายและให้ข้อเสนอแนะ เนื่องจากวิธีการวิเคราะห์ Joinpoint regression ก็ยังคงมีข้อจำกัดบางประการและการอภิปรายจุด Change point กับนโยบายหรือมาตรการทางด้านสาธารณสุขนั้นมีความละเอียดอ่อนอย่างมาก เพราะเป็นเพียงการให้ข้อสังเกตเท่านั้น ไม่ใช่เป็นการนำเข้าตัวแล้วผ่านการวิเคราะห์เพื่อยืนยันทางสถิติแต่อย่างใด

เอกสารอ้างอิง

1. Liu B, Kim HJ, Feuer EJ, Graubard BI. Joinpoint Regression Methods of Aggregate Outcomes for Complex Survey Data. *J Surv Stat Methodol*. 2023;11(4):967–89.
2. Li J, Chan NB, Xue J, Tsoi KKF. Time series models show comparable projection performance with joinpoint regression: A comparison using historical cancer data from World Health Organization. *Front Public Heal*. 2022;10.
3. Rea F, Pagan E, Compagnoni MM, Cantarutti A, Pugni P, Bagnardi V, et al. Joinpoint regression analysis with time-on-study as time-scale. Application to three Italian population-based cohort studies. *Epidemiol Biostat Public Heal*. 2017;14(3):e12616-1-e12616-7.
4. National Cancer Institute. Joinpoint Trend Analysis Software [Internet]. 2023 [cited 2024 Jan 20]. Available from: <https://surveillance.cancer.gov/joinpoint/>
5. Bergwall S, Tran D. Joinpoint regression analysis of the COVID-19 epidemic curve in Sweden. *Linköping University*; 2023.

6. Khaleghi S, Nikbakht HA, Khodabandelu S, Khafri S. Incidence and investigation of Covid-19 trend in Babol, northern Iran: A Joinpoint regression analysis. *Casp J Intern Med*. 2022;13(Suppl 3):236–43.
7. Tazhibi M, Babazade S, Dehkordi Z. Application of joinpoint regression in determining breast cancer incidence rate change points by age and tumor characteristics in women aged 30-69 (years) and in Isfahan city from 2001 to 2010. *J Educ Health Promot*. 2014;3(1):115.
8. Chaurasia AR. COVID-19 Trend and Forecast in India: A Joinpoint Regression Analysis. *medRxiv* [Internet]. 2020;(January):2020.05.26.20113399. Available from: <http://medrxiv.org/content/early/2020/06/03/2020.05.26.20113399.abstract>
9. Al-Jamaei AAH, van Dijk BAC, Helder MN, Forouzanfar T, Leemans CR, de Visscher JGAM. A population-based study of the epidemiology of oral squamous cell carcinoma in the Netherlands 1989–2018, with emphasis on young adults. *Int J Oral Maxillofac Surg* [Internet]. 2022;51(1):18–26. Available from: <https://doi.org/10.1016/j.ijom.2021.03.006>
10. Askarishahi M, Malaki moghadam H, Fallahzade H, Lotfi M, Afkhami-Ardekani M. Trend analysis of the rate of mortality due to diabetes mellitus in Iran during the period of 2003-2010: A join point regression analysis. *J Occup Heal Epidemiol*. 2014;3(2):112–6.
11. Olorunfemi G, Libhaber E, Ezechi OC, Musenge E. Population-Based Temporal Trends and Ethnic Disparity in Cervical Cancer Mortality in South Africa (1999–2018): A Join Point and Age–Period–Cohort Regression Analyses. *Cancers (Basel)*. 2022;14(24):1–26.
12. Dragomirescu I, Llorca J, Gómez-Acebo I, Dierssen-Sotos T. A join point regression analysis of trends in mortality due to osteoporosis in Spain. *Sci Rep*. 2019;9(1):1–8.
13. National Cancer Institute. Grid Search Method - Details [Internet]. 2024 [cited 2024 Feb 16]. Available from: <https://surveillance.cancer.gov/help/joinpoint/setting-parameters/method-and-parameters-tab/method/method-grid-search-or-hudsons>
14. Kim H-J, Yu B, Feuer EJ. SELECTING THE NUMBER OF CHANGE-POINTS IN SEGMENTED LINE REGRESSION. *Stat Sin*. 2009 May;19(2):597–609.
15. National Cancer Institute. How Joinpoint Conducts Permutation Testing [Internet]. 2022 [cited 2024 Feb 16]. Available from: <https://surveillance.cancer.gov/help/joinpoint/setting-parameters/method-and-parameters-tab/model-selection-method/permutation-tests/model-selection-method-how-joinpoint-conducts-permutation-testing>
16. Claeskens G, Hjort NL. The Bayesian information criterion. *Model Sel Model Averaging*. 2010;70–98.
17. National Cancer Institute. Bayesian Information Criterion (BIC) Details [Internet]. 2022 [cited 2024 Feb 19]. Available from: <https://surveillance.cancer.gov/help/joinpoint/tech-help/frequently-asked-questions/bayesian-information-criterion-bic-details>
18. Tiwari RC, Cronin KA, Davis W, Feuer EJ. Bayesian model selection for join point regression with application to age-adjusted cancer rates. 2005;54(5): 919–39.
19. Kafle RC. Trend Analysis and Modeling of Health and Environmental Data: Joinpoint and Functional Approach. 2014;
20. Wang J, Zhou L, Yin W, Hu C, Zuo X. Trends of the burden of type 2 diabetes mellitus attributable to high body mass index from 1990 to 2019 in China. *Front Endocrinol (Lausanne)*. 2023;14(May):1–15.
21. Puzo Q, Qin P, Mehlum L. Long-term trends of suicide by choice of method in Norway: A joinpoint regression analysis of data from 1969 to 2012. *BMC Public Health* [Internet]. 2016;16(1):1–9. Available from: <http://dx.doi.org/10.1186/s12889-016-2919-y>
22. Rovetta A. Joinpoint Regression to Determine the Impact of COVID-19 on Mortality in Europe: A Longitudinal Analysis From 2000 to 2020 in 27 Countries. *medRxiv* [Internet]. 2022;2022.01.19. 22269576. Available from: <http://medrxiv.org/content/early/2022/03/01/2022.01.19.22269576.abstract>
23. Chollet F. Machine learning. *Mach Learn* [Internet]. 2017;45(13):40–8. Available from: <https://books.google.ca/books?id=EoYBngEACAAJ&dq=mitchell+machine+learning+1997&hl=en&sa=X&ved=0ahUKewi0mdqfj8TkAhWGsIkKHRcBAt0Q6AEIKjAA>

24. SARAÇLI HTS. Joinpoint Regression Analysis and an Application on Istanbul Stock-Exchange. Alphanumeric J. 2014;2(1).
25. Ries LAG, Wingo PA, Miller DS, Howe HL, Weir HK, Rosenberg HM, et al. The annual report to the nation on the status of cancer, 1973-1997, with a special section on colorectal cancer. Cancer. 2000;88(10): 2398-424.
26. Brenda K. Edwards, PhD: Elizabeth Ward, PhD : Betsy A. Kohler, MPH C. Cancer - 2009 - Edwards - Annual report to the nation on the status of cancer 1975+2006 featuring colorectal cancer.pdf. 2009.
27. Xu J, Lin Y, Yang M, Zhang L. Statistics and pitfalls of trend analysis in cancer research: A review focused on statistical packages. J Cancer. 2020;11(10):2957-61.
28. Woodward M. Epidemiology: Study design and data analysis, third edition. Epidemiology: Study Design and Data Analysis, Third Edition. 2013. 1-821 p.