บทที่ 3

ระเบียบวิธีวิจัย

ในบทที่ 3 จะกล่าวถึงระเบียบวิธีวิจัยที่ใช้ในการศึกษาเรื่อง "การวิเคราะห์ความเป็นไปได้ทางเศรษฐกิจในการพัฒนาอำเภอเบตง จังหวัดยะลา ให้เป็นแหล่งท่องเที่ยวเชิงสุขภาพ" โดยในการวิจัยครั้งนี้จะใช้การวิเคราะห์เชิงปริมาณ (Quantitative Method) โดยจะแบ่งออกเป็น 2 ส่วนคือ ส่วนแรกในปีงบประมาณ 2562 จะเป็นการประเมินอุปสงค์ของการใช้บริการการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา ส่วนที่สองในปีงบประมาณ 2563 เพื่อประเมินสถานะของการให้บริการการท่องเที่ยวเชิงสุขภาพของสถานประกอบการและหน่วยงานที่เกี่ยวข้องกับการท่องเที่ยวในอำเภอเบตง จังหวัดยะลา นอกจากนี้ยังจัดทำเส้นทางการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา ซึ่งในส่วนที่สองนี้เป็นการประเมินความพร้อมทางด้านอุปทานของบริการการท่องเที่ยวเชิงสุขภาพ สำหรับการดำเนินการวิจัยในส่วนแรกนั้นจะตอบวัตถุประสงค์การวิจัยดังต่อไปนี้

1) เพื่อศึกษาความต้องการของนักท่องเที่ยวชาวไทยและชาวมาเลเซียเกี่ยวกับการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา
2) เพื่อศึกษาค่าใช้จ่ายของนักท่องเที่ยวทั้งชาวไทยและชาวมาเลเซียที่เต็มใจจะจ่ายเพื่อต้องการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา เพื่อนำรายได้ที่เกิดขึ้นไปพัฒนาแหล่งท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา
3) เพื่อศึกษาว่า สถานการณ์ความไม่สงบที่เกิดขึ้นในอำเภอเบตง จังหวัดยะลา มีส่งผลต่ออุปสงค์ความต้องการท่องเที่ยวของนักท่องเที่ยวชาวไทยและชาวมาเลเซียที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัดยะลาหรือไม่

3.1 การเก็บรวบรวมข้อมูลเพื่อการศึกษา

จากวัตถุประสงค์ในการศึกษาข้างต้น ผู้วิจัยจำเป็นต้องใช้ข้อมูลในการศึกษา 2 ส่วนด้วยกัน ได้แก่ 1) ข้อมูลปฐมภูมิ เพื่อใช้ศึกษาวิเคราะห์ตามวัตถุประสงค์ที่ 1) และ 2) และข้อมูลทุติยภูมิ เพื่อศึกษาวิเคราะห์ตามวัตถุประสงค์ที่ 3) รายละเอียดในการเก็บข้อมูลทั้งสองประเภท มีดังนี้

3.1.1 การเก็บข้อมูลปฐมภูมิด้วยแบบสอบถาม

3.1.1.1 การคำนวณกลุ่มตัวอย่าง (Sample) ที่ต้องการศึกษา

จากวัตถุประสงค์ทั้ง 2 ข้อแรก ทางผู้วิจัยจะต้องทราบจำนวนกลุ่มตัวอย่าง ในการเก็บข้อมูลความต้องการของและค่าใช้จ่ายของนักท่องเที่ยวในการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา โดยคำนวณจาก 2 ส่วนคือนักท่องเที่ยวชาวไทยและนักท่องเที่ยวชาวมาเลเซีย ในส่วนของนักท่องเที่ยวชาวมาเลเซีย ทางผู้วิจัยทราบจำนวนประชากรนักท่องเที่ยวชาวมาเลเซียที่เดินทางเข้ามาท่องเที่ยวอำเภอเบตง จังหวัดยะลา โดยสามารถเก็บข้อมูลได้จากสำนักงานตรวจคนเข้าเมืองเบตง (Betong Immigration Bureau) ในส่วนของนักท่องเที่ยวชาวไทย ทางผู้วิจัยได้สอบถามเจ้าหน้าที่สำนักงานการท่องเที่ยวและกีฬา จังหวัดยะลา ทางเจ้าหน้าที่แจ้งว่า ทางสำนักงานการท่องเที่ยวและกีฬา จังหวัดยะลา ไม่ได้มีการเก็บข้อมูลนักท่องเที่ยวชาวไทยที่เดินทางมายังอำเภอเบตง จังหวัดยะลาโดยตรง แต่เจ้าหน้าที่แจ้งว่า การเก็บข้อมูลนักท่องเที่ยวชาวไทยที่เดินทางเข้ามาท่องเที่ยวในจังหวัดยะลาร้อยละ 95 เป็นนักท่องเที่ยวที่เดินทางไปท่องเที่ยวที่อำเภอเบตง ดังนั้น ผู้วิจัยจึงประมาณการว่า จำนวนนักท่องเที่ยวไทยที่เดินทางมาท่องเที่ยวในจังหวัดยะลา ร้อยละ 95 เป็นนักท่องเที่ยวที่ตั้งใจเดินทางไปท่องเที่ยว ณ อำเภอเบตง จังหวัดยะลา มีเพียงร้อยละ 5 เท่านั้นที่เดินทางมาท่องเที่ยวในบริเวณอื่นๆของจังหวัดยะลา

ดังนั้น เมื่อทางผู้วิจัยทราบจำนวนประชากรที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัดยะลา ทำให้ผู้วิจัยสามารถเลือกใช้สูตรในการคำนวณกลุ่มตัวอย่างได้อย่างถูกต้อง ในกรณีนี้ ทางผู้วิจัยทราบจำนวนประชากร ดังนั้นทางผู้วิจัยจึงใช้สูตร Taro Yamane ในการคำนวณกลุ่มตัวอย่าง ตามสมการทางคณิตศาสตร์ ดังนี้

$$n = \frac{N}{1 + NE^2}$$

โดยกำหนดให้ n คือ จำนวนกลุ่มตัวอย่าง N คือขนาดประชากร E คือค่าความคาดเคลื่อนที่ยอมรับได้ ซึ่งในการศึกษาครั้งนี้ ใช้จำนวนนักท่องเที่ยวทั้งชาวไทยและชาวมาเลเซียที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบ

ตง จังหวัดยะลา ซึ่งในปี 2560 มีจำนวนทั้งสิ้น 609,055 คน (กระทรวงการท่องเที่ยวและกีฬา, 2561) เป็นขนาดประชากรที่ต้องการศึกษา โดยแบ่งเป็นนักท่องเที่ยวชาวไทยจำนวน 133,589 คนและนักท่องเที่ยวชาวมาเลเซีย 475,466 คนและให้ค่าความคาดเคลื่อนที่ยอมรับได้เท่ากับร้อยละ 5 ในการสุ่มตัวอย่าง จากสูตรการคำนวณกลุ่มตัวอย่างของ Taro Yamane สามารถคำนวณกลุ่มตัวอย่างได้ดังนี้

$$n = \frac{609{,}055}{1 + 609{,}055(0.05)^2}$$

$$n = 399.74$$

ดังนั้น ในการวิจัยครั้งนี้จะเก็บตัวอย่างจำนวน 400 คน โดยแบบสอบถามที่ใช้มีทั้งภาษาไทยและภาษาอังกฤษ สำหรับแบบสอบถามภาษาอังกฤษ ต้องผ่านการแปลกลับเป็นภาษาไทยเพื่อทดสอบความเที่ยงตรงของแบบสอบถาม ซึ่งแบบสอบถามดังกล่าวแสดงไว้ในภาคผนวกที่ 1 ในส่วนของการสุ่มตัวอย่าง (Sampling) จะใช้วิธีการเลือกกลุ่มตัวอย่างแบบโควตา (Quota sampling) เป็นการเลือกกลุ่มตัวอย่างโดยคำนึงถึงสัดส่วนองค์ประกอบของประชากร โดยเมื่อต้องการกลุ่มตัวอย่าง 400 คน จะแบ่งการเก็บข้อมูลเป็นนักท่องเที่ยวชาวไทย 100 คน นักท่องเที่ยวชาวมาเลเซีย 300 คน (จำนวนนักท่องเที่ยวไทยคิดเป็นร้อยละ 21.93 หรือคำนวณจำนวนการเก็บแบบสอบถามได้เป็น 88 ชุด ส่วนนักท่องเที่ยวมาเลเซียคิดเป็นร้อยละ 78.07 หรือคำนวณเป็นจำนวนการเก็บแบบสอบถามได้เป็น 312 ชุด แต่เพื่อให้ง่ายและสะดวกจึงเก็บแบบสอบถามจากนักท่องเที่ยวชาวไทย 100 ชุดและนักท่องเที่ยวชาวมาเลเซียจำนวน 300 ชุด ตามลำดับ)

ในส่วนของการเลือกเก็บข้อมูลกลุ่มตัวอย่างจะใช้วิธีการเลือกแบบบังเอิญ (Accidental sampling) เพื่อให้ได้จำนวนตามต้องการโดยไม่มีหลักเกณฑ์ กลุ่มตัวอย่างจะเป็นใครก็ได้ที่มีอายุตั้งแต่ 20 ปีขึ้นไป ซึ่งเป็นกลุ่มคนที่มีความสามารถหรือกำลังซื้อในสินค้าที่เกี่ยวข้องกับการท่องเที่ยวเชิงสุขภาพ ในกรณีที่ผู้ให้สัมภาษณ์ไม่ต้องการให้สัมภาษณ์หรือหยุดการให้สัมภาษณ์กลางคัน ทางผู้วิจัยจะหยุดสัมภาษณ์และเลือกผู้ให้สัมภาษณ์ที่เต็มใจให้สัมภาษณ์รายใหม่ ทดแทนคนเดิมที่ไม่ให้ข้อมูล

3.1.2 การเก็บข้อมูลทุติยภูมิจำนวนนักท่องเที่ยว

ในการวิเคราะห์สถานการณ์ความไม่สงบที่เกิดขึ้นว่ามีผลต่อการอุปสงค์ความต้องการท่องเที่ยวของนักเที่ยวชาวไทยที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัดยะลาหรือไม่นั้น ผู้วิจัยจะเก็บรวบรวมข้อมูลจำนวนนักท่องเที่ยวชาวไทย จาก สำนักงานการท่องเที่ยวและกีฬาจังหวัดยะลา ส่วนข้อมูลทุติยภูมิจำนวนนักท่องเที่ยวชาวมาเลเซียที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัดยะลา ทางผู้วิจัยจะเก็บรวบรวม

ข้อมูลจำนวนนักท่องเที่ยวชาวมาเลเซียจากสำนักงานตรวจคนเข้าเมืองเบตงและจากเว็บไซด์ของศูนย์วิจัยทางด้านการตลาดการท่องเที่ยว

3.2 การศึกษาความต้องการของนักท่องเที่ยวชาวไทยและชาวมาเลเซียเกี่ยวกับการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา เพื่อบรรลุวัตถุประสงค์ข้อที่ 1 ด้วยแบบจำลอง Logit

แบบจำลอง Logit เป็นแบบจำลองสองทางเลือก (Binary Choice Models) ที่สร้างขึ้นมาเพื่อแสวงหาคำตอบถึง ความน่าจะเป็น (probability) ในการตัดสินใจที่จะเลือกระหว่างสองทางเลือก ว่าแต่ละทางเลือกมีความน่าจะเป็นมากน้อยเพียงใด และขึ้นอยู่กับปัจจัยที่เป็นตัวกำหนดใดบ้าง นอกจากแบบจำลอง Logit แล้ว แบบจำลอง Probit ก็สามารถใช้วิเคราะห์ความน่าจะเป็นระหว่างสองทางเลือกได้เช่นเดียวกัน แต่ค่าสัมประสิทธิ์ของแบบจำลอง Probit นั้นแปลความยากกว่าค่าสัมประสิทธิ์ของแบบจำลอง Logit อีกทั้งผลที่ได้จากแบบจำลอง Logit และ Probit ก็ให้ข้อสรุปทางสถิติที่ใกล้เคียงกัน ดังนั้นเมื่อได้ข้อมูลจากแบบสอบถามในขั้นตอนที่ 1 มาแล้ว ทางคณะผู้วิจัยจะใช้แบบจำลอง Logit เพื่อวิเคราะห์ความต้องการใช้บริการการท่องเที่ยวเชิงสุขภาพในพื้นที่อำเภอเบตง จังหวัดยะลา

การวิเคราะห์ด้วยแบบจำลอง Logit จะต้องอาศัยตัวแบบที่ต่างจากตัวแบบเศรษฐมิติทั่วไปที่ใช้วิเคราะห์ที่มีตัวแปรตามเป็นตัวแปรเชิงปริมาณ (Quantitative variable) และมีลักษณะต่อเนื่อง แต่ตัวแบบสองทางเลือกมีตัวแปรตามคือ”การตัดสินใจ” เป็นตัวแปรเชิงคุณภาพ (qualitative variable) ที่มีค่าเพียง 2 ค่าหรือสองทางเลือก ซึ่งใช้ในการวิเคราะห์ อธิบายการตัดสินใจเลือกในกรณีต่างๆ เช่น ประชาชนตัดสินใจที่จะซื้อสินค้าหรือไม่ซื้อสินค้า ตัดสินใจที่จะซื้อหุ้นเซ็นทรัลพัฒนา (CPN) หรือไม่ซื้อ การตัดสินใจเลือกดังกล่าวจะมีสองทางเลือก ซึ่งจะถูกกำหนดโดยปัจจัยที่อยู่เบื้องหลังของผู้ตัดสินใจเลือก เช่น รายได้ อายุ ระดับการศึกษาเป็นต้น การหาความสัมพันธ์ระหว่างการตัดสินใจและปัจจัยกำหนด โดยใช้วิธีการหาสมการถดถอยโดยปกตินั้นไม่เหมาะสม เนื่องจากการตัดสินใจในการเลือกซึ่งเป็นตัวแปรตามมีลักษณะเป็นค่าที่แยกออกจากกัน (Dichotomous) กล่าวอีกนัยเป็นค่าที่ไม่ต่อเนื่อง (Continuous) ทำให้สมมติฐานที่อยู่เบื้องหลังตัวแบบสมการถดถอย (Normal basic assumptions) ในเรื่องการแจกแจงของตัวแปรตามว่า จะต้องมีการแจกแจงแบบปกติ (Normal distribution) ไม่เป็นจริง ดังนั้นจึงไม่สามารถใช้แบบจำลองตัวแบบความน่าจะเป็นเส้นตรง (The Linear Probability Model) ได้ จึงเป็นอีกเหตุผลหนึ่งที่ทำให้ทางผู้วิจัยต้องใช้แบบจำลอง Logit Model ในการวิเคราะห์ข้อมูล (ถวิล นิลใบ, 2555, น.1-2)

แบบจำลองโลจิก (The Logit Model)

ถวิล นิลใบ (2555, น.8-35) ได้กล่าวไว้ว่า แบบจำลอง Logit เป็นแบบจำลองที่ใช้แก้ปัญหาการพยากรณ์ความน่าจะเป็นที่ได้จากตัวแบบความน่าจะเป็นเชิงเส้นตรงที่อยู่นอกของเขต 0 และ 1 การแก้ไขคือต้องปรับเปลี่ยนตัวแบบเดิมที่จะทำให้มีคุณสมบัติที่จะสามารถเปลี่ยนค่าตัวแปรอิสระซึ่งมีค่าใดๆ ให้เป็นค่าความน่าจะเป็นที่อยู่ในช่วงระหว่าง 0 และ 1 ซึ่งความสัมพันธ์ระหว่างค่าความน่าจะเป็นและตัวแปรอิสระอาจจะอยู่ในลักษณะทิศทางเดียวกัน หรือในทางตรงกันข้ามก็ได้ ตัวแบบโลจิกแสดงได้โดยสมการ

$$P_i = E(Y = 1|X_i) = \frac{1}{1+e^{-(\beta_1+\beta_2 X_i)}}$$

หรือ $$P_i = \frac{1}{1+e^{-Z_i}}$$

โดยที่ $$Z_i = \beta_1 + \beta_2 X_i \qquad \text{.......................(3.1)}$$

ค่า e เป็นฐานของ natural logarithms ซึ่งมีค่าประมาณ 2.718 ส่วนค่า P_i คือค่าความน่าจะเป็นที่บุคคลตัดสินใจเลือกสถานการณ์ที่สนใจ ภายใต้ค่า X ณ ระดับต่างๆ ที่กำหนด

การคำนวณตัวแบบโลจิก

เนื่องจากตัวแบบโลจิกที่แสดงในสมการที่ (3.1) ไม่ได้อยู่ในรูปความสัมพันธ์เชิงเส้นตรง กล่าวคือค่าตัวแปรตามมีความสัมพันธ์แบบไม่เป็นเชิงเส้นตรงกับค่าตัวแปรอิสระและค่าพารามิเตอร์ ดังนั้น จำเป็นต้องแปลงตัวแบบนี้ให้อยู่ในรูปความสัมพันธ์เชิงเส้นตรง ซึ่งมีขั้นตอนดังนี้

ถ้า P แสดงความน่าจะเป็นในความต้องการการบริการเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา (1-P) จะแสดงความน่าจะเป็นที่ ไม่ต้องการการบริการเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา

$$1 - P_i = \frac{1}{1+e^{Z_i}} \qquad \text{..................... (3.2)}$$

นำสมการที่ (3.2) ไปหารสมการที่ (3.1) ผลที่ได้คือ

$$\frac{P_i}{1-P_i} = \frac{1+e^{Z_i}}{1+e^{-Z_i}} = e^{Z_i} \qquad \text{...................... (3.3)}$$

ทั้งนี้ $\frac{P_i}{1-P_i}$ คือ อัตราส่วนระหว่างความน่าจะเป็นที่ต้องการการบริการเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา และความน่าจะเป็นที่ไม่ต้องการการบริการเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา (สัดส่วนนี้ภาษาวิชาการเรียกว่า "odds") ถ้าเราแปลงสมการที่ (3.3) ให้อยู่ในรูปของ natural log ได้ผลดังนี้

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = Z_i = \beta_1 + \beta_2 X_i \qquad \text{.......................} (3.4)$$

L ที่ปรากฏในสมการ (3.4) เป็นค่า log ของอัตราส่วนความน่าจะเป็น (หรือ natural log ของ odds) ซึ่งเรียกว่า "Logit" จึงทำให้เรียกตัวแบบนี้ว่า Logit model และเพื่อวัตถุประสงค์การคำนวณ เราจะแปลงสมการ (3.4) ให้อยู่ในรูปของ Stochastic relation โดยใส่ค่าตัวรบกวน (Disturbance term) ลงในสมการ

$$L_i = \beta_1 + \beta_2 X_i + \varepsilon_i \qquad \text{.......................} (3.5)$$

จะเห็นว่า สมการ (3.5) ตัวแปรตามคือ "Logit" (L) มีความสัมพันธ์เชิงเส้นตรงทั้งกับตัวแปรอิสระและค่าพารามิเตอร์ ซึ่งจะสะดวกและง่ายที่นำไปใช้ในการคำนวณหาค่าพารามิเตอร์ สมการ (3.5) เป็นรูปแบบของสมการถดถอยอย่างง่าย (Simple regression) ถ้าเขียนในรูป ของตัวแบบสมการถดถอยพหุคูณ (Multiple regression) เขียนได้ดังนี้

$$\ln\left(\frac{P_i}{1-P_i}\right) = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots\ldots + \beta_k X_k + u_i \qquad \text{.....................} (3.6)$$

ตัวแปรตาม คือสัดส่วนความน่าจะเป็นแสดงในรูปของ natural log (log ฐาน e) ส่วนค่าตัวแปรอิสระ (X) อาจจะเป็นตัวแปรเชิงปริมาณหรือเชิงคุณภาพก็ได้ การคำนวณค่าพารามิเตอร์ของสมการที่ (3.6) สามารถดำเนินการได้ 2 วิธีคือ Ordinary least squares (OLS) และ Maximum Likelihood (ML) วิธี ML เป็นวิธีที่เหมาะสมและสอดคล้องกับรูปแบบการเก็บข้อมูลที่เก็บจากแต่ละบุคคล (individual) ซึ่งเป็นข้อมูลภาคตัดขวาง (Cross section data) สำหรับวิธีกำลังสองน้อยที่สุดไม่สามารถใช้ได้กับข้อมูลที่เก็บจากแต่ละบุคคล หากแต่จะใช้ข้อมูลที่เก็บเป็นกลุ่ม (Grouped data) ซึ่งปกติงานวิจัยทั่วไปมักไม่ค่อยเก็บข้อมูลในลักษณะนี้ ดังนั้นในงานวิจัยครั้งนี้จะใช้วิธีการคำนวณแบบ Maximum Likelihood ซึ่งจะแสดงรายละเอียดของการคำนวณในส่วนถัดไป

การคำนวณแบบจำลอง Logit model ด้วยวิธี Maximum Likelihood

ถวิล นิลใบ (2555, น.15-17) กล่าวไว้ว่า วิธี Maximum likelihood (ML) มีข้อดีหลายประการ ประการแรก สามารถคำนวณตัวแบบที่มีลักษณะไม่ใช่เส้นตรงได้ เช่นกรณีตัวแบบโลจิก ประการที่สอง ขั้นตอนการคำนวณยังสามารถจัดการกับปัญหาความแปรปรวนของค่าตัวรบกวนไม่คงที่ (heteroscedasticity)

ประการที่สาม สามารถใช้ได้กับข้อมูลทั้งที่เป็นกลุ่มและแต่ละบุคคลได้ สำหรับข้อด้อยของวิธีนี้คือมีความยุ่งยากและใช้เวลาในการคำนวณมาก เมื่อเทียบกับวิธีกำลังสองน้อยที่สุด แต่ปัจจุบันข้อด้อยนี้ ไม่มีปัญหาเนื่องจากปัจจุบันมีโปรแกรมคอมพิวเตอร์ช่วยในการคำนวณ จึงทำให้วิธีนี้ได้รับ ความนิยมอย่างมาก

หลักการสำคัญของวิธี ML คือจะต้องสร้างฟังก์ชั่นที่แสดงความน่าจะเป็นร่วม (Joint probability) ของค่าสังเกตของตัวอย่างที่สุ่มมา ฟังก์ชั่นดังกล่าวนี้มีชื่อเรียกว่า “Likelihood function” รูปแบบทั่วไปของฟังก์ชั่นนี้ เขียนได้ดังนี้

$$L(X_1, X_2, \dots\dots, X_n; \theta_1, \theta_2, \dots\dots . \theta_k)$$

โดยที่ X_i คือตัวแปรที่สังเกต θ_i คือค่าพารามิเตอร์ที่เราต้องการคำนวณ จากนั้นคำนวณหาค่าพารามิเตอร์ที่ทำให้ค่าฟังก์ชั่นความน่าจะเป็นนี้มีค่าสูงสุด (จึงเรียกวิธีคำนวณนี้ว่า “maximum likelihood”) ซึ่งดำเนินการตามขั้นตอนการหาค่าสูงสุดของฟังก์ชั่นคือ หาค่าอนุพันธ์บางส่วน (Partial derivative) ของฟังก์ชั่นความน่าจะเป็นเทียบกับค่าพารามิเตอร์แต่ละตัว กำหนดให้ค่าอนุพันธ์เท่ากับศูนย์แล้วแก้ระบบสมการหาค่าพารามิเตอร์ ตัวอย่างสำหรับ likelihood function ของตัวแบบโลจิกของค่าสังเกตจำนวน n ค่า (มีตัวแปรอิสระ 1 ตัว) แสดงในรูปของ natural log คือ

$$L = \sum_{i=1}^{n} [y_i \ln(P_i) + (1 - y_i) \ln(1 - P_i)]$$

แทนค่า P_i ด้วย $E(Y_i)$ ในสมการ $P_i = \beta_1 + \beta_2 X_i$

$$\iota = \log \text{likelihood} = \sum_{i=1}^{n} [y_i (\alpha + \beta X_i) \ln(1 - e^{\alpha + \beta X_i})]$$

$$\frac{\partial l}{\partial \alpha} = \sum_{i=1}^{n} \left[y_i - \frac{1}{(1 + e^{-\alpha - \beta X_i})}\right] = 0$$

$$\frac{\partial l}{\partial \beta} = \sum_{i=1}^{n} \left[y_i X_i - \frac{1}{(1 + e^{-\alpha - \beta X_i})}\right] = 0$$

เนื่องจากฟังก์ชั่นอนุพันธ์ที่ได้ไม่ได้อยู่ในรูปของสมการเส้นตรง จึงไม่สามารถเขียนสูตรของการคำนวณหาค่าพารามิเตอร์สองค่าคือ α และ β ทำให้การคำนวณของวิธีนี้จึงยุ่งยากเมื่อเทียบกับวิธี LS

3.3 การศึกษาค่าใช้จ่ายของนักท่องเที่ยวทั้งชาวไทยและชาวมาเลเซียที่เต็มใจจะจ่ายเพื่อต้องการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา เพื่อบรรลุวัตถุประสงค์ข้อที่ 2 ด้วยแบบจำลองสมการถดถอยเชิงเส้นหลายตัวแปร

การศึกษาการใช้จ่ายของนักท่องเที่ยว (Tourist expenditure) ทั้งขาวไทยและชาวมาเลเซีย ที่เต็มใจจ่ายเพื่อการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา มีสมมุติฐานว่า ค่าใช้จ่ายที่นักท่องเที่ยวเต็มใจจ่ายเพื่อบริการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา จะมีค่าเท่ากับ ค่าใช้จ่ายที่นักท่องเที่ยวเคยใช้จ่ายเพื่อบริการท่องเที่ยวเชิงสุขภาพในลักษะเดียวกัน ตามหลักการประเมินความเต็มใจจ่ายจากตลาดที่มีอยู่เดิม (Willingness to pay – revealed preference - market data)

ลำดับต่อมา ผู้วิจัยการศึกษาปัจจัยที่กำหนดการใช้จ่ายเพื่อบริการท่องเที่ยวเชิงสุขภาพของนักท่องเที่ยว เพื่อค้นหาปัจจัยที่เป็นตัวกำหนดการใช้จ่ายเพื่อบริการท่องเที่ยวเชิงสุขภาพของนักท่องเที่ยว ด้วยแบบจำลองการใช้จ่ายของนักท่องเที่ยว (Tourist expenditure model) ที่มีแนวคิดตามทฤษฎีอุปสงค์ทางอ้อม (Indirect demand) ที่ว่าการใช้จ่ายของนักท่องเที่ยวจะขึ้นอยู่กับรายได้และปัจจัยกำหนดอื่นๆ โดยปัจจัยที่กำหนดการใช้จ่ายเพื่อบริการท่องเที่ยวเชิงสุขภาพของนักท่องเที่ยวแบ่งออกเป็น 4 กลุ่ม ประกอบด้วย

1) ตัวแปรทางด้านประชากร ได้แก่ สัญชาติ เพศ อายุ ศาสนา ระดับการศึกษา อาชีพ และรายได้ต่อเดือน
2) ตัวแปรทางด้านพฤติกรรมการท่องเที่ยว ได้แก่ ช่องทางการรับข้อมูล ลักษณะการเดินทางท่องเที่ยว ระยะเวลาที่ใช้ในการท่องเที่ยว ที่พักแรม ค่าใช้จ่ายที่ใช้ในการท่องเที่ยว
3) ตัวแปรทางด้านประสบการณ์การใช้บริการท่องเที่ยวเชิงสุขภาพ
4) ตัวแปรทางด้านการใช้บริการท่องเที่ยวเชิงสุขภาพ

3.3.1 วิเคราะห์ความแปรปรวน

ผู้วิจัยใช้การวิเคราะห์ความแปรปรวน (ANOVA) เพื่อวิเคราะห์ค่าเฉลี่ยของค่าใช้จ่ายที่นักท่องเที่ยวเต็มใจจ่ายเพื่อบริการท่องเที่ยวเชิงสุขภาพที่จำแนกตามประเภทของตัวแปรทางประชากร ตัวแปรพฤติกรรมการท่องเที่ยว ตัวแปรประสบการณ์การใช้บริการท่องเที่ยวเชิงสุขภาพ และตัวแปรการใช้บริการท่องเที่ยวเชิงสุขภาพในอำเภอเบตง จังหวัดยะลา โดยมีสมมุติฐานหลักว่า ค่าใช้จ่ายที่นักท่องเที่ยวเต็มใจจ่ายเพื่อบริการท่องเที่ยวเชิงสุขภาพ ซึ่งจำแนกตามประเภทของแต่ละตัวแปรนั้นมีค่าเฉลี่ยไม่แตกต่างกัน หากค่าความน่าจะเป็นที่ได้จากการทดสอบความแปรปรวนมีค่าน้อยกว่า 0.05 หรือ 0.1 จะสามารถปฏิเสธสมมุติฐานหลัก และ

ยอมรับสมมุติฐานรองที่ว่า ค่าใช้จ่ายที่นักท่องเที่ยวเต็มใจจ่ายเพื่อบริการท่องเที่ยวเชิงสุขภาพ ซึ่งจำแนกตามประเภทของแต่ละตัวแปรนั้นมีค่าเฉลี่ยแตกต่างกัน ณ ระดับความเชื่อมั่นร้อยละ 95 และร้อยละ 90 ตามลำดับ และตัวแปรที่มีนัยสำคัญทางสถิติจะทำไปใช้ในการจัดทำแบบจำลองสมการถดถอยเชิงเส้นหลายตัวแปรต่อไป

3.3.2 แบบจำลองสมการถดถอยเชิงเส้นหลายตัวแปร

ผู้วิจัยใช้แบบจำลองสมการถดถอยเชิงเส้นหลายตัวแปร (Multiple linear regression model) เป็นแบบจำลองการใช้จ่ายของนักท่องเที่ยวเพื่อบริการท่องเที่ยวเชิงสุขภาพ โดยมีรูปแบบสมการดังต่อไปนี้

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

โดยที่ Y คือ ค่าใช้จ่ายที่นักท่องเที่ยวเต็มใจจ่ายเพื่อบริการท่องเที่ยวเชิงสุขภาพ

X_1 ถึง X_n คือ ตัวแปรที่ใช้ในการศึกษาทั้ง 4 กลุ่มรวม 30 ตัวแปร

α, β_1 และ β_n คือ ค่าสัมประสิทธิ์คงที่ และค่าสัมประสิทธิ์ของตัวแปร X_1 ถึง X_n

ในเลือกตัวแปรที่ใช้ในแบบจำลอง ผู้วิจัยเลือกตัวแปรที่มีค่าความน่าจะเป็นของการวิเคราะห์ความแปรปรวนน้อยกว่า 0.1 มาใช้ในแบบจำลองสมการถดถอยเชิงเส้นหลายตัวแปร (Multiple Linear Regression Model)

ผู้วิจัยใช้วิธีกำลังสองน้อยที่สุด (Ordinary Least Square Method) ในการประมาณค่าสัมประสิทธิ์ของแบบจำลอง และใช้วิธี Stepwise ในการคัดเลือกตัวแปรอิสระค่าสัมประสิทธิ์ที่ได้ของแบบจำลอง

3.4. การศึกษาว่า สถานการณ์ความไม่สงบที่เกิดขึ้นในอำเภอเบตง จังหวัดยะลา มีส่งผลต่ออุปสงค์ความต้องการท่องเที่ยวของนักท่องเที่ยวชาวไทยและชาวมาเลเซียที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัดยะลาหรือไม่ เพื่อบรรลุวัตถุประสงค์ข้อที่ 3 ด้วยแบบจำลอง SARIMA with Intervention

ในการวิเคราะห์เพื่อหาคำตอบว่า สถานการณ์ความไม่สงบในพื้นที่ในพื้นที่ 3 จังหวัดชายแดนใต้มีผลต่อการตัดสินใจเดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัด โดยใช้แบบจำลองทางเศรษฐมิติคือ SARIMA with intervention ในการวิเคราะห์ และใช้ข้อมูลทุติยภูมิในการประมาณการจำนวนนักท่องเที่ยวที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัดยะลา

ในส่วนของข้อมูลสถานการณ์ความไม่สงบจะเป็นข้อมูลสถานการณ์ความไม่สงบที่เกิดขึ้นในอำเภอเบตง อ.ธารโต อ.บันนังสตา และเหตุการณ์ที่สำคัญหรือที่เกิดขึ้นในอำเภอเมือง จังหวัดยะลา ซึ่งข้อมูลดังกล่าว

ได้มาจากศูนย์เฝ้าระวังสถานการณ์ความไม่สงบชายแดนใต้ เมื่อได้ข้อมูลมาเรียบร้อยแล้ว นำมาวิเคราะห์ด้วยแบบจำลอง SARIMA Intervention โดยมีขั้นตอนต่างๆดังนี้

ขั้นตอนที่ 1 การทดสอบความนิ่งของข้อมูล (Unit Root Test)

ก่อนที่จะนำข้อมูลไปใช้ในการทดสอบ SARIMA Intervention จะต้องนำข้อมูลจำนวนนักท่องเที่ยวชาวมาเลเซีย ซึ่งข้อมูลมีลักษณะเป็นข้อมูลอนุกรมเวลา โดยมีจำนวนข้อมูลทั้งหมด 177 เดือน มาปรับทดสอบความนิ่งของข้อมูลก่อน ซึ่งถ้าข้อมูลหนึ่ง ก็สามารถนำข้อมูลดังกล่าวไปใช้ในการประมาณค่าได้ทันที แต่ถ้าข้อมูลดังกล่าวมีลักษณะที่ไม่นิ่ง (Non-stationary) ทางผู้วิจัยจำเป็นต้องปรับข้อมูลให้อยู่ในลักษณะ natural logarithm ดังสมการข้างล่างนี้

$$Y_t = 100 * (\log(g_t) - \log(g_{t-1}))$$

หรืออาจจะปรับข้อมูลเป็นผลต่าง (Difference) เมื่อกำหนดให้ Δ คือผลต่างครั้งที่ 1 และ Δ^d คือผลต่างครั้งที่ d ดังนั้นจะได้สมการข้างล่างนี้

$$\Delta Y_t = Y_t - Y_{t-1}$$

$$\Delta^d = \Delta^{d-1} Y_t - \Delta^{d-1} Y_{t-1}$$

โดยที่ Y_t คือ จำนวนนักท่องเที่ยวมาเลเซียที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัดยะลา

หลังจากที่ปรับข้อมูลให้อยู่ในรูปแบบ natural logarithm หรือในรูปของผลต่างแล้ว จะต้องนำข้อมูลได้ไปทดสอบความนิ่ง (Stationary) โดยอัครพงศ์ อั้นทอง (2555, น.29) กล่าวว่า การทดสอบความนิ่งข้อมูลอนุกรมเวลาจำเป็นต้องทำเป็นอันดับแรก ก่อนที่นำมาใช้ในแบบจำลองเพื่อใช้ในการประมาณค่า ซึ่งโดยทั่วไปมักพบว่า ข้อมูลอนุกรมเวลามีลักษณะไม่นิ่ง (Non-stationary) แตกต่างจากข้อสมมุติฐานดั้งเดิมของนักเศรษฐศาสตร์คลาสสิก ที่กล่าวว่า ข้อมูลมีลักษณะนิ่ง (Stationary) ค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) มีค่าคงที่ ดังนั้นก่อนที่จะนำข้อมูลมาใช้ จะต้องนำข้อมูลเหล่านี้ มาทดสอบคุณสมบัติว่า มีความนิ่งหรือไม่

ในการศึกษาครั้งนี้ เลือกทดสอบความนิ่งของข้อมูลอนุกรมเวลาโดยใช้วิธี Augmented Dickey-Fuller (ADF) test ที่เสนอโดย Dickey and Fuller 1979 และ 1981 เนื่องจาก เป็นวิธีที่ได้รับการยอมรับและเป็นที่นิยมอย่างแพร่หลาย ในการศึกษาความนิ่งของข้อมูลอนุกรมเวลา และหากผลการทดสอบที่ได้มีความไม่นิ่ง หมายความว่า ชุดของข้อมูลเหล่านี้มีการเคลื่อนไหวไปตามแนวโน้มที่เพิ่มขึ้นตามกาลเวลา (Time

Trend) และความแปรปรวนวิ่งห่างออกจากเดิมไปเรื่อยๆตามแนวโน้มของระยะเวลาที่เพิ่มขึ้น ดังนั้นเพื่อให้การประมาณค่ามีความถูกต้องน่าเชื่อถือ ข้อมูลเหล่านี้จะถูกนำมาปรับให้นิ่งโดยการทำผลต่างลำดับที่ 1 (First difference) หรือลำดับที่สูงขึ้นไปจนกว่าข้อมูลจะมีความนิ่ง แล้วจึงนำไปใช้ในการประมาณค่าในแบบจำลองต่อไป เนื่องจากในปัจจุบันเป็นที่ยอมรับกันอย่างกว้างขวางว่า การใช้ข้อมูลที่มีความไม่นิ่ง ไปใช้ประมาณค่าในแบบจำลองต่าง ๆ อาจส่งผลให้เกิดปรากฏการณ์ของความสัมพันธ์ปลอม (Spurious Regression) โดยใช้สมการต่อไปนี้

$$\Delta y_t = \mu + \gamma y_{t-1} + \sum_{i=2}^{\infty} \beta_i \Delta y_{t-i+1} + \varepsilon_t \qquad \text{.........................} (3.7)$$

และในกรณีที่นำแนวโน้มของเวลา (Time Trend) เข้ามาร่วมพิจารณาด้วยจะได้สมการที่ (3.8)

$$\Delta y_t = \mu + \gamma y_{t-1} + \mu_2 t + \sum_{i=2}^{\infty} \beta_i \Delta y_{t-i+1} + \varepsilon_t \text{ ....} (3.8)$$

โดยที่ y_t คือ ข้อมูลอนุกรมเวลาที่ต้องการทดสอบ

ε_t คือ ตัวคลาดเคลื่อน (Error term)

ภายใต้สมมติฐาน

$H_0 : \gamma = 0$ (Non-stationary)

$H_1 : \gamma \neq 0$ (Stationary)

จากสมการที่ (3.11) และ (3.12) ถ้าค่า $\gamma = 0$ แสดงว่าชุดข้อมูลนี้มีความไม่นิ่ง ค่าสถิติที่คำนวณได้ไม่สามารถปฏิเสธสมมุติฐานหลักหรือ $H_0: \gamma = 0$ ได้

ขั้นตอนที่ 2 การกำหนดรูปแบบ (Identification)

อัครพงศ์ อั้นทอง (2555, น.108-114) กล่าวว่า การหารูปแบบ AR และ MA ที่เหมาะสมกับข้อมูลอนุกรมเวลา โดยพิจารณาจากคอเรลโลเกรมของค่าสัมประสิทธิ์สหสัมพันธ์ (Autocorrelation function: ACF) และค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (Partial correlation function: PACF) ของข้อมูลจำนวนนักท่องเที่ยวชาวมาเลเซียที่เดินทางมายังอำเภอเบตง จังหวัดยะลา ซึ่งจากคอเรลโลเกรม ถ้าปรากฏว่ายังมีค่าสัมประสิทธิ์สหสัมพันธ์ และค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน แสดงว่าไม่สามารถใช้แบบจำลอง ARMA (Autoregressive Moving Average) ได้ ดังนั้นจึงจำเป็นต้องใช้แบบจำลอง ARIMA (Autoregressive Integrated Moving Average) แทน ในกรณีที่มีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง โดยเฉพาะเรื่องของจำนวนนักท่องเที่ยวมาเลเซียที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัดยะลา ซึ่ง

สามารถสังเกตได้จากคอเรลโลเกรมเช่นเดียวกัน ก็จำเป็นต้องใช้ Seasonal ARIMA หรือ SARIMA แทนแบบจำลอง ARIMA

ดังนั้นแบบจำลอง SARIMA พัฒนาจากข้อมูลอนุกรมเวลาที่ปราศจาก Intervention ซึ่งในกรณีกรณีนี้คือข้อมูลนักท่องเที่ยวมาเลเซียที่เดินทางเข้ามาท่องเที่ยวอำเภอเบตง จังหวัดยะลา ตั้งแต่เดือนมกราคม พ.ศ. 2547 – กันยายน พ.ศ.2561 รวมทั้งสิ้น 177 เดือน สำหรับการกำหนดรูปแบบของแบบจำลอง SARIMA (p,d,q) สามารถพิจารณาจากกราฟค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเอง (ACF) และค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองบางส่วน (PACF) แล้วนำรูปแบบดังกล่าวไปตรวจสอบความเหมาะสมด้วยการพิจารณากราฟของค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองของค่าคลาดเคลื่อน และการทดสอบด้วยวิธี Box and Pierce (Box and Ljung) จนในที่สุดจะได้รูปแบบของแบบจำลอง SARIMA (p,d,q) ที่เหมาะสมสำหรับการใช้ในการพยากรณ์

ขั้นตอนที่ 3 ประมาณค่าสัมประสิทธิ์ของแบบจำลอง (Estimation)

ซึ่งเป็นการประมาณสัมประสิทธิ์ของรูปแบบที่กำหนด ซึ่งอาจมีมากกว่า 1 รูปแบบ วิธีที่นิยมใช้คือ OLS และ MLE เนื่องจากโดยส่วนใหญ่ข้อมูลอนุกรมเวลาจะไม่ค่อยมีลักษณะเป็นเส้นตรง ดังนั้น โดยส่วนใหญ่การวิเคราะห์อนุกรมเวลาจะไม่สามารถใช้การวิเคราะห์แบบ OLS ได้ ต้องใช้ MLE ในการวิเคราะห์

ขั้นตอนที่ 4 ตรวจสอบรูปแบบ (Diagnostic checking)

เป็นการตรวจสอบแบบจำลองของข้อมูลมีความเหมาะสมหรือไม่ และค่าคลาดเคลื่อนมีลักษณะเป็น white noise ตามสมมติฐานและเงื่อนไขของแบบจำลองบ๊อกซ์และเจนกินส์หรือไม่ โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองของค่าคลาดเคลื่อนหรือค่าสถิติ Ljung-Box statistics: LB) การทดสอบค่าพารามิเตอร์ด้วย t-test และการพิจารณา Goodness of fit ของแบบจำลองด้วยค่า AIC และ BIC

ขั้นตอนที่ 5 นำข้อมูลที่ได้มาวิเคราะห์ภายใต้แบบจำลอง SARIMA with intervention

จากการศึกษางานวิจัยเกี่ยวกับผลกระทบที่มีต่ออุปสงค์การท่องเที่ยว โดยการใช้แบบจำลองการแทรกแซงหรือ Intervention model ซึ่งแบบจำลองการแทรกแซงที่ใช้ในการวิเคราะห์ผลนั้นประกอบด้วย 4 รูปแบบด้วยกันคือ ARIMA model, SARIMA model, ARIMA with intervention model and SARIMA with intervention model ซึ่งในการศึกษาวิจัยเกี่ยวกับสถานการณ์ความไม่สงบที่เกิดขึ้นในอำเภอเบตง จังหวัดยะลา เพื่อเป็นการพิจารณาเหตุการณ์ความในพื้นที่4 จังหวัดชายแดนใต้ มีส่งผลต่ออุปสงค์ความต้องการท่องเที่ยวของนักท่องเที่ยวชาวมาเลเซียที่เดินทางเข้ามาท่องเที่ยวในอำเภอเบตง จังหวัดยะลา ซึ่งแต่

ละเหตุการณ์มีระยะเวลาของผลกระทบที่แตกต่างกันออกไป ดังนั้นทางผู้วิจัยได้เลือกใช้แบบจำลอง SARIMA with intervention model โดยมีขั้นตอนในการศึกษาดังนี้

แบบจำลอง Seasonal - ARIMA with intervention model

ในกรณีที่ข้อมูลอนุกรมเวลามีค่าสุดโต่ง (Outlier) สามารถเพิ่มตัวแปร Outlier เข้าไปในแบบจำลอง ARIMA และ SARIMA ในฐานะของตัวแปรถดถอย โดยรูปแบบของค่าสุดโต่งที่พบอยู่เสมอ คือ Additive outliers และ Level shifts กลายเป็นแบบจำลอง SARIMA with intervention (p, d, q) (P, D, Q)S ซึ่งเขียนในรูปสมการต่อไปนี้

SARIMA with intervention model = Intervention function + SARIMA noise model Intervention function เขียนสมการในรูปแบบดังนี้

$$y_t = f(I_t) + N_t$$

โดยที่ y_t คือ ค่าสังเกต ณ เวลาที่ t จากอนุกรมเวลาที่เป็น Stationary

$f(I_t)$ คือ เป็นฟังก์ชันของตัวแปรหุ่น(Dummy) ที่แสดงอิทธิพลของ Intervention ณ เวลาที่ t

ถ้า $I_t = 1$ เมื่อเกิด Intervention

$I_t = 0$ เมื่อไม่เกิด Intervention

N_t คือ เป็น Noise series ก่อนเกิด Intervention ที่มีรูปแบบ ARIMA(p, d, q)

โดยทั่วไปอิทธิพลของ Intervention ที่มีต่ออนุกรมเวลามี 2 ลักษณะคือ ผลกระทบที่คงอยู่ตลอดไป (Step function) และผลกระทบที่เกิดขึ้นเฉพาะบางช่วงของเวลาแล้วหมดไป (Pulse function) ความแตกต่างดังกล่าวทำให้การกำหนดรูปแบบของตัวแปรหุ่นแตกต่างกันดังนี้

ก. Step function เป็นกรณีที่ผลกระทบของ Intervention เกิดขึ้น ณ เวลาที่ t และคงอยู่ตลอดไป สามารถกำหนดตัวแปรหุ่นที่แสดงอิทธิพลของ Intervention ได้ดังนี้

$$S_t^T = \begin{cases} 0, t < T \\ 1, t \geq T \end{cases}$$

ข. Pulse functionเป็นกรณีที่ผลกระทบของInterventionเกิดขึ้นแล้วคงอยู่เพียงช่วงเวลาใดเวลาหนึ่ง ในกรณีนี้ตัวแปรหุ่นที่แสดงถึงอิทธิพลของ Intervention มีรูปแบบดังนี้

$$P_t^T = \begin{cases} 0, t \neq T \\ 1, t = T \end{cases}$$

ขั้นตอนที่ 6 การตรวจสอบความเหมาะสมของแบบจำลอง SARIMA with Intervention

ค่า AIC เป็นค่าที่แสดงความแปรปรวนของค่าคลาดเคลื่อนที่เกิดจากการประมาณค่า ดังนั้นจึงควรเลือกให้ค่า AIC มีค่าต่ำที่สุด เนื่องจากให้ค่าคลาดเคลื่อนมีความแปรปรวนน้อยที่สุด และ ค่า log-likelihood เทียบได้กับค่าสัมประสิทธิ์สหสัมพันธ์ในตัวเองของค่าคลาดเคลื่อน (Residuals) เพราะถ้าค่า log-likelihood ยิ่งมีค่ามาก แสดงว่าสมการหรือแบบจำลองที่สร้างขึ้นมีคุณภาพหรือกลมกลืนกับข้อมูลมาก

ขั้นตอนที่ 7 การพยากรณ์ (Forecasting)

นำแบบจำลองที่ผ่านการตรวจสอบไปพยากรณ์ค่าในอนาคต สามารถทำได้ทั้งการพยากรณ์แบบจุด (Point Forecast) และการพยากรณ์แบบช่วง (Interval Forecast)