บทที่ 3
วิธีดำเนินการวิจัย

ในบทนี้จะกล่าวถึงวิธีการดำเนินการวิจัย ประกอบด้วย การกำหนดขอบเขตและพื้นที่ศึกษา เครื่องมือที่ใช้ในการวิจัย การเก็บรวบรวมข้อมูล และการออกแบบแบบจำลองการเลือกรูปแบบการเดินทาง โดยมีรายละเอียดดังต่อไปนี้

3.1 ขั้นตอนการดำเนินการวิจัย

การพยากรณ์การเปลี่ยนแปลงรูปแบบการเดินทางระหว่างเมืองในแต่ละช่วงระยะทางหลังการเปิดให้บริการรถไฟความเร็วสูง มีขั้นตอนในการดำเนินการดังนี้

1) ทบทวนทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบจำลองการเลือกรูปแบบการเดินทาง ปัจจัยที่มีผลต่อการเลือกรูปแบบการเดินทางและความสามารถในการแข่งขันและผลกระทบที่อาจเกิดขึ้นกับรูปแบบการเดินทางเดิมในแต่ละช่วงระยะทางหลังการเปิดให้บริการรถไฟความเร็วสูง

2) กำหนดขอบเขตและพื้นที่ศึกษา จากนั้นดำเนินการสำรวจข้อมูลภาคสนามโดยใช้วิธีสัมภาษณ์ ณ จุดสำรวจบริเวณช่วงถนน สถานีรถโดยสาร สถานีรถไฟ และสนามบิน ตามแนวเส้นทางที่ทำการศึกษา

3) ตรวจสอบความถูกต้องของข้อมูลและทำการวิเคราะห์ข้อมูลเบื้องต้นที่ได้จากการสำรวจ ประกอบด้วย ข้อมูลคุณลักษณะทางสังคมประชากรและข้อมูลเกี่ยวกับรายละเอียดการเดินทางของกลุ่มตัวอย่าง

4) พัฒนาแบบจำลองการเลือกรูปแบบการเดินทางโดยใช้โปรแกรม R ในการวิเคราะห์ Multinomial Logit Model และ Mixed Logit Model

5) สรุปผลการวิจัย จัดทำข้อเสนอแนะ และจัดทำรายงานฉบับสมบูรณ์

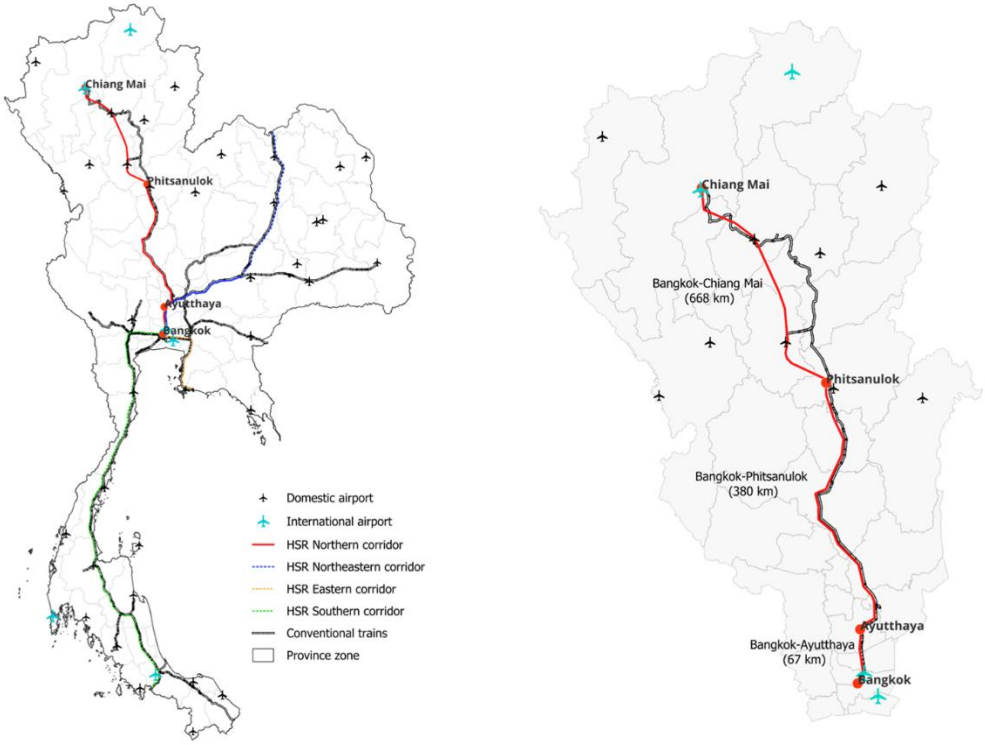
3.2 กำหนดขอบเขตและพื้นที่ศึกษา

การศึกษานี้จะทำการตรวจสอบพฤติกรรมการเลือกรูปแบบการเดินทางจากผู้เดินทางตามแนวเส้นทางรถไฟความเร็วสูงสายกรุงเทพฯ-เชียงใหม่ ซึ่งสามารถแบ่งเป็นช่วงการเดินทางหลัก 3 ช่วงระยะทาง ได้แก่

3.2.1 การเดินทางช่วงกรุงเทพฯ-อยุธยา มีระยะทางประมาณ 76 กิโลเมตร

3.2.2 การเดินทางช่วงกรุงเทพฯ-พิษณุโลก มีระยะทางประมาณ 380 กิโลเมตร

3.2.3 การเดินทางช่วงกรุงเทพฯ-เชียงใหม่ มีระยะทางประมาณ 673 กิโลเมตร



รูปที่ 3.1 แนวเส้นทางรถไฟความเร็วสูง เส้นทาง กรุงเทพฯ-เชียงใหม่

3.3 เครื่องมือที่ใช้ในการวิจัย

ในการวิจัยครั้งนี้มีเครื่องมือหลักที่ใช้ในการศึกษาประกอบด้วยแบบสอบถาม (Questionnaires) และโปรแกรมที่ใช้ในการพัฒนาแบบจำลองการเลือกรูปแบบการเดินทาง

3.3.1 แบบสอบถาม (Questionnaires)

แบบสอบถามที่พัฒนาขึ้นในการศึกษาครั้งนี้ใช้สำหรับสัมภาษณ์ผู้เดินทางตามแนวเส้นทางรถไฟความเร็วสูงสายกรุงเทพฯ-เชียงใหม่ ประกอบด้วย 3 ส่วน ดังนี้

ส่วนที่ 1: ข้อมูลคุณลักษณะทางสังคมประชากรของผู้ตอบแบบสอบถาม

ส่วนที่ 2: ข้อมูลเกี่ยวกับรายละเอียดการเดินทางของผู้ตอบแบบสอบถาม

ส่วนที่ 3: การระบุทางเลือกที่พึงพอใจภายใต้สถานการณ์จำลองเมื่อมีการเปิดให้บริการรถไฟความเร็วสูงเป็นรูปแบบการเดินทางใหม่

3.3.2 โปรแกรมที่ใช้ในการพัฒนาแบบจำลองการเลือกรูปแบบการเดินทาง

แบบจำลองการเลือกรูปแบบการเดินทางที่พัฒนาขึ้นในการศึกษานี้ดำเนินการโดยใช้แพ็กเกจ Apollo ในโปรแกรม R ซึ่งเป็นโปรแกรมที่มีประสิทธิภาพและมีความยืดหยุ่นสูงในการพัฒนาแบบจำลองทางเลือก (Choice Models) (Hess & Palma, 2019)

3.4 การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลเพื่อใช้ประกอบการพัฒนาแบบจำลองมีรายละเอียดการเก็บรวบรวมข้อมูลดังต่อไปนี้

3.4.1 ประชากรและกลุ่มตัวอย่าง

ประชากรและกลุ่มตัวอย่างในการศึกษานี้คือผู้เดินทางในเส้นทางกรุงเทพฯ-เชียงใหม่ โดยแบ่งตามประเภทยานพาหนะของผู้เดินทาง 4 ประเภท ประกอบด้วย รถยนต์ส่วนบุคคล รถโดยสารประจำทาง รถไฟ และเครื่องบิน ขนาดของกลุ่มตัวอย่างขั้นต่ำได้มีการคำนวณโดยใช้สมการ 3.1

$$n = \frac{\pi(1-\pi)z_{\alpha/2}^2}{D^2} \tag{3.1}$$

โดยที่ n คือ จำนวนตัวอย่างขั้นต่ำ, π คือ สัดส่วนทางเลือกจริงของประชากร, D คือ ระดับ error ที่ยอมรับได้ และ $z_{\alpha/2}$ คือ ความเชื่อมั่นที่กำหนด (Malhotra, 2010) การศึกษานี้คำนวณจำนวนตัวอย่างขั้นต่ำที่ระดับความเชื่อมั่น 95% และขอบเขตความคลาดเคลื่อนที่ยอมรับได้ที่ระดับ 5%

จำนวนตัวอย่างถูกพิจารณาจากสัดส่วนการเดินทางจริงของยานพาหนะแต่ละประเภทในปัจจุบันตามแนวเส้นทางกรุงเทพฯ-เชียงใหม่ ในปี พ.ศ. 2562 ซึ่งถือเป็นจำนวนการเดินทางจากสภาวะปกติก่อนเกิดเหตุการณ์การแพร่ระบาดของ COVID-19 ในไทย โดยสถิติปริมาณการเดินทางถูกจัดเก็บโดยหน่วยงานต่าง ๆ ทั้งหน่วยงานภาครัฐและภาคเอกชนภายในประเทศ

จากการคำนวณสัดส่วนการเลือกเดินทางจริงในเส้นทางกรุงเทพฯ-เชียงใหม่ พบว่าสัดส่วนผู้เดินทางด้วยรถยนต์ส่วนบุคคล เท่ากับ 0.45 เครื่องบิน เท่ากับ 0.44 รถโดยสารและรถไฟเท่ากับ 0.08 และ 0.02 ตามลำดับ เมื่อคำนวณตามสมการที่ 3.1 จะได้จำนวนตัวอย่างขั้นต่ำของผู้เดินทางด้วยรถยนต์ส่วนบุคคล เท่ากับ 381 รถโดยสาร 110 รถไฟ 37 และเครื่องบิน 379 ตัวอย่าง แต่ในการศึกษาพฤติกรรมการเลือกรูปแบบการเดินทางจำนวนกลุ่มตัวอย่างขั้นต่ำของผู้มีอำนาจตัดสินใจในแต่ละทางเลือกควรมีอย่างน้อย 50 ตัวอย่าง (Hensher et al., 2015) และหากต้องการเปรียบเทียบความแตกต่างระหว่างกลุ่ม Orme (2010) แนะนำว่าขนาดของกลุ่มตัวอย่างแต่ละกลุ่มควรมีอย่างน้อย 200 ตัวอย่าง ดังนั้น กลุ่มตัวอย่างของผู้เดินทางโดยรถโดยสาร และรถไฟจึงถูกปรับ

เป็นอย่างน้อย 200 ตัวอย่าง จำนวนตัวอย่างขั้นต่ำในการสำรวจผู้เดินทางแต่ละประเภทยานพาหนะแสดงดังตารางที่ 3.1

ตารางที่ 3.1 จำนวนตัวอย่างขั้นต่ำ

รูปแบบการเดินทาง	จำนวนเที่ยวการเดินทางต่อวัน	ส่วนแบ่งการเดินทาง (%)	จำนวนตัวอย่างขั้นต่ำ
รถยนต์ส่วนบุคคล	18,064	45.32	381
รถโดยสาร	3,108	7.80	200
รถไฟ	973	2.44	200
เครื่องบิน	17,718	44.45	379
รวม			1,160

3.4.2 วิธีการสำรวจ

การศึกษาที่เกี่ยวข้องกับการสร้างแบบจำลองพฤติกรรมการเลือกรูปแบบการเดินทาง (Mode Choice Model) มีเทคนิคที่นิยมใช้ในการสัมภาษณ์ผู้เดินทาง 2 เทคนิค คือ Reveal Preference (RP) Technique และ Stated Preference (SP) Technique

1) Reveal Preference (RP) เป็นวิธีการสำรวจข้อมูลของผู้เดินทางภายใต้สถานการณ์ที่เกิดขึ้นจริง

2) Stated Preference (SP) เป็นวิธีการที่นิยมใช้ในการสำรวจข้อมูลภายใต้สถานการณ์จำลองที่ยังไม่เคยเกิดขึ้นมาก่อน

โดยทั้ง RP และ SP เป็นวิธีการที่มีจุดแข็งและจุดอ่อนแตกต่างกัน การสำรวจข้อมูลแบบ RP มีจุดแข็งที่ลักษณะข้อมูลที่ได้รับมีความน่าเชื่อถือเนื่องจากเป็นการสอบถามภายใต้สถานการณ์ที่เกิดขึ้นจริงจากพฤติกรรมจริงของผู้เดินทาง ตรงข้ามกับวิธีการ SP ที่ข้อมูลที่ได้รับมาจากสถานการณ์สมมติที่ยังไม่เคยเกิดขึ้นมาก่อน แต่ข้อมูล RP มีข้อด้อยกว่าข้อมูล SP ที่ไม่มีความยืดหยุ่นในการควบคุมตัวแปรต่าง ๆ ทำให้ไม่สามารถอธิบายความผันแปรของตัวแปรได้เพียงพอ (Swait et al., 1994)

การศึกษานี้ได้นำวิธีการสำรวจข้อมูลโดยใช้เทคนิค Stated Preference (SP) มาใช้ในการสำรวจข้อมูลผู้เดินทาง เนื่องจากการสำรวจในรูปแบบ SP เป็นวิธีการที่มีความยืดหยุ่นในการกำหนดตัวแปร มีค่าใช้จ่ายน้อยและประหยัดเวลามากกว่าวิธีการ RP นอกจากนั้นยังสามารถสร้างสถานการณ์สมมติในการเก็บรวบรวมข้อมูลจากผู้เดินทางเมื่อมีการพิจารณารูปแบบการเดินทางที่ยังไม่เคยเปิดให้ใช้บริการมาก่อน

3.4.3 พื้นที่สำรวจ

การสำรวจข้อมูลกลุ่มตัวอย่างดำเนินการผ่านการสอบถามแบบ Face-to-face บริเวณถนนสายหลัก จุดพักรถหรือสถานีบริการน้ำมันของพื้นที่ศึกษาสำหรับสำรวจผู้เดินทางด้วย

รถยนต์ส่วนบุคคล สถานีขนส่งและสถานีรถไฟหลักในเขตตัวเมืองของแต่ละจังหวัดตามแนวเส้นทาง HSR สำหรับสำรวจผู้เดินทางด้วยรถโดยสารสาธารณะและรถไฟ และบริเวณสนามบินสุวรรณภูมิ สนามบินดอนเมือง สนามบินพิษณุโลก และสนามบินเชียงใหม่ สำหรับสำรวจผู้เดินทางด้วยเครื่องบิน

3.4.4 การออกแบบสถานการณ์จำลอง

การออกแบบสถานการณ์จำลองในการศึกษานี้ สร้างขึ้นโดยพิจารณาจากคุณลักษณะสำคัญ 2 ประการที่มีผลต่อการเลือกรูปแบบการเดินทาง ได้แก่ เวลาในการเดินทาง (Travel Time) และค่าใช้จ่ายในการเดินทาง (Travel Cost) โดยแต่ละทางเลือกในชุดทางเลือกจะถูกกำหนดด้วยระดับที่แตกต่างกัน การออกแบบสถานการณ์จำลองเริ่มต้นด้วยการออกแบบแฟกทอเรียลแบบเต็ม (Full Factorial Design) เพื่อสร้างสถานการณ์จำลองที่เป็นไปได้ทั้งหมด ส่งผลให้ได้สถานการณ์จำนวน 64 สถานการณ์สำหรับเส้นทางกรุงเทพฯ–อยุธยา, 128 สถานการณ์สำหรับเส้นทางกรุงเทพฯ–พิษณุโลก และ 128 สถานการณ์สำหรับเส้นทางกรุงเทพฯ–เชียงใหม่ อย่างไรก็ตาม จำนวนสถานการณ์เหล่านี้ถือว่ามากเกินไป ซึ่งอาจก่อให้เกิดความสับสนและเป็นภาระทางความคิดที่ไม่จำเป็นแก่ผู้ตอบแบบสอบถาม และอาจส่งผลให้คุณภาพของข้อมูลที่รวบรวมได้ลดลง ดังนั้นจึงจำเป็นต้องลดจำนวนสถานการณ์จำลองเหล่านี้ลงโดยใช้การออกแบบเชิงแฟกทอเรียลแบบเศษส่วน (Fractional Factorial Design) ส่งผลให้เหลือ 16 สถานการณ์สำหรับแต่ละช่วงระยะทาง

อย่างไรก็ตาม การให้ผู้ตอบแบบสอบถามประเมินสถานการณ์ครบทั้ง 16 สถานการณ์อาจจะยังสร้างความเหนื่อยล้าและส่งผลต่อคุณภาพของคำตอบ ดังนั้น การออกแบบสถานการณ์จำลองจึงถูกปรับเพิ่มเติมโดยใช้การจัดกลุ่ม (Blocking) ซึ่งเป็นเทคนิคในการแบ่งชุดสถานการณ์หลัก 16 สถานการณ์ออกเป็น 2 กลุ่มย่อย (Blocks) โดยแต่ละกลุ่มจะประกอบด้วยสถานการณ์จำลอง 8 สถานการณ์ จากนั้นผู้ตอบแบบสอบถามในแต่ละช่วงระยะทางจะถูกสุ่มให้ทำแบบสอบถามเพียงหนึ่งกลุ่มเท่านั้นเพื่อให้ภาระงานมีความเหมาะสมและผู้ตอบสามารถจดจ่อได้ดีขึ้น โดยรายละเอียดของระดับคุณลักษณะที่ใช้แสดงดังตารางที่ 3.2

ตารางที่ 3.2 ระดับของคุณลักษณะ (Attribute Levels) ในแบบสอบถาม

แบบสำรวจชุด A: กรุงเทพฯ - อยุธยา		
รูปแบบการเดินทาง	ค่าใช้จ่ายในการเดินทาง (บาท)	เวลาในการเดินทาง (นาที)
รถยนต์ส่วนบุคคล	215	70, 80
รถโดยสาร	120	170, 180
รถไฟแบบดั้งเดิม	105	145, 195
เครื่องบิน	-	-
รถไฟความเร็วสูง	400, 420, 460, 500	85, 170
แบบสำรวจชุด B: กรุงเทพฯ - พิษณุโลก		
รูปแบบการเดินทาง	ค่าใช้จ่ายในการเดินทาง (บาท)	เวลาในการเดินทาง (นาที)
รถยนต์ส่วนบุคคล	750	290, 320
รถโดยสาร	440	420, 450
รถไฟแบบดั้งเดิม	540	340, 490
เครื่องบิน	980, 1500	180
รถไฟความเร็วสูง	780, 860, 1010, 1160	150, 230
แบบสำรวจชุด C: กรุงเทพฯ - เชียงใหม่		
รูปแบบการเดินทาง	ค่าใช้จ่ายในการเดินทาง (บาท)	เวลาในการเดินทาง (นาที)
รถยนต์ส่วนบุคคล	1250	550, 600
รถโดยสาร	750	650, 700
รถไฟแบบดั้งเดิม	900	500, 800
เครื่องบิน	1350, 2100	250
รถไฟความเร็วสูง	1150, 1300, 1550, 1800	200, 300

3.5 การออกแบบแบบจำลองการเลือกรูปแบบการเดินทาง

การศึกษานี้ได้ทำการพัฒนาแบบจำลองประเภท Logit 2 รูปแบบ ได้แก่ Multinomial Logit (MNL) Model และ Mixed Logit (ML) Model โดยอ้างอิงตามทฤษฎีอรรถประโยชน์สูงสุด (Utility Maximization) แบบจำลองที่พัฒนาขึ้นมีวัตถุประสงค์หลักในการตรวจสอบพฤติกรรมการเลือกรูปแบบการเดินทางและปัจจัยที่มีผลต่อการเลือกรูปแบบการเดินทางหลังการเปิดให้บริการรถไฟความเร็วสูง

3.5.1 ตัวแปรที่เกี่ยวข้อง

ตัวแปรหลักในการศึกษานี้แบ่งออกเป็น 3 ประเภท คือ 1) ตัวแปรที่เกี่ยวกับคุณลักษณะของการเดินทาง ได้แก่ ค่าใช้จ่ายในการเดินทาง และเวลาที่ใช้ในการเดินทาง 2) ตัวแปรคุณลักษณะทางสังคมประชากรของผู้เดินทาง ได้แก่ เพศ อายุ รายได้ และ 3) ตัวแปรทดสอบที่เกี่ยวข้องกับการเดินทาง ได้แก่ ระยะการเดินทาง โดยตัวแปรทั้ง 3 ประเภทแสดงดังตารางที่ 3.3

ตารางที่ 3.3 ตัวแปรอิสระที่นำเข้าแบบจำลอง MNL Model และ ML Model

ตัวแปร	ความหมาย	พารามิเตอร์	ประเภทข้อมูล
ตัวแปรคุณลักษณะของการเดินทาง (Attributes Travel)			
Total Travel Cost	ค่าใช้จ่ายในการเดินทาง (บาท)	B_COST	Ratio Scale
Total Travel Time	ระยะเวลาที่ใช้ในการเดินทาง (นาที)	B_TIME	Ratio Scale
ตัวแปรคุณลักษณะทางสังคมประชากรของผู้เดินทาง (Socio-demographic)			
Gender	เพศของผู้ตอบแบบสอบถาม	B_GENDER	Dummy Variable
Age	อายุของผู้ตอบแบบสอบถาม	B_AGE	Ratio Scale
Household Income	รายได้ในครัวเรือนต่อเดือนของผู้ตอบแบบสอบถาม	B_HH_INCOME	Ratio Scale
ตัวแปรทดสอบที่เกี่ยวข้องกับการเดินทาง (Scenario Variables)			
Distance	ระยะการเดินทาง	B_DISTANCE	Dummy Variable

ตัวแปรค่าใช้จ่ายในการเดินทาง (Travel Cost) เป็นตัวแปรที่มีการจัดเก็บข้อมูลด้วยเลขจำนวนเต็ม มีลักษณะของข้อมูลเป็นมาตรวัดอัตราส่วน (Ratio Scale)

ตัวแปรระยะเวลาที่ใช้ในการเดินทาง (Travel Time) เป็นตัวแปรที่มีการจัดเก็บข้อมูลด้วยเลขจำนวนเต็ม มีลักษณะของข้อมูลเป็นมาตรวัดอัตราส่วน (Ratio Scale)

ตัวแปรเพศ (Gender) ของผู้ตอบแบบสอบถาม เป็นตัวแปรที่มีการจัดเก็บข้อมูลโดยกำหนดให้มี 2 ค่า คือ เพศชายและเพศหญิง ซึ่งมีลักษณะของข้อมูลเป็นมาตราวัดนามบัญญัติ (Nominal Scale) การพิจารณาในแบบจำลองได้ทำการแปลงข้อมูลให้อยู่ในรูปของตัวแปรหุ่น (Dummy Variable)

ตัวแปรอายุ (Age) ของผู้ตอบแบบสอบถาม เป็นตัวแปรที่มีการจัดเก็บข้อมูลด้วยเลขจำนวนเต็ม มีลักษณะของข้อมูลเป็นมาตรวัดอัตราส่วน (Ratio Scale)

ตัวแปรรายได้ในครัวเรือนต่อเดือน (Household Income) ของผู้ตอบแบบสอบถาม เป็นตัวแปรที่มีการจัดเก็บข้อมูลโดยแบ่งตามช่วงระดับรายได้ซึ่งมีลักษณะของข้อมูลเป็นมาตรวัดอันดับ (Ordinal Scale) การพิจารณาในแบบจำลองได้ทำการแปลงข้อมูลให้อยู่ในรูปของค่ามัธยฐานของแต่ละอันตรภาคซึ่งมีลักษณะของข้อมูลเป็นมาตรวัดอัตราส่วน (Ratio Scale)

ระยะการเดินทาง (Travel Distance) เป็นตัวแปรที่มีการจัดเก็บข้อมูลด้วยตัวเลข มีลักษณะของข้อมูลเป็นมาตรวัดอัตราส่วน (Ratio Scale) การพิจารณาในแบบจำลองได้ทำการแปลงข้อมูลให้อยู่ในรูปของตัวแปรหุ่น (Dummy Variable) เพื่อตรวจสอบว่าผู้ตอบแบบสอบถามมีการเดินทางในช่วงระยะสั้น ระยะกลาง หรือระยะกลางค่อนไปทางไกล

3.5.2 การพัฒนาแบบจำลองการเลือกรูปแบบการเดินทาง

แบบจำลอง Logit Model ที่พัฒนาขึ้นในการศึกษานี้จะเริ่มจากการกำหนดโครงสร้างแบบจำลองในชุดทางเลือก (Choice Set) ที่ทำการศึกษา 5 ทางเลือก ได้แก่ รถยนต์ส่วนบุคคล รถโดยสาร รถไฟแบบดั้งเดิม เครื่องบิน และรถไฟความเร็วสูง โดยมีฟังก์ชันอรรถประโยชน์ (Utility Function) ดังนี้

แบบจำลองที่ 1 กำหนดให้ค่าสัมประสิทธิ์ของตัวแปรเวลาที่ใช้ในการเดินทาง (B_TIME) และค่าใช้จ่ายในการเดินทาง (B_COST) มีค่าเท่ากันในทุก ๆ ทางเลือก (Generic Parameter) และมีสมมติฐานว่าเพศ (GENDER) ของผู้เดินทางเป็นปัจจัยที่ส่งผลต่อพฤติกรรมการเลือกเดินทางด้วยรถไฟแบบดั้งเดิมและเครื่องบินเมื่อเทียบกับรูปแบบการเดินทางอื่น อายุ (AGE) ของผู้เดินทางเป็นปัจจัยที่ส่งผลต่อพฤติกรรมการเลือกเดินทางด้วยรถโดยสาร เครื่องบิน และรถไฟความเร็วสูงเมื่อเทียบกับรูปแบบการเดินทางอื่น รายได้ในครัวเรือน (HH_INCOME) เป็นปัจจัยที่ส่งผลต่อพฤติกรรมการเลือกเดินทางด้วยรถโดยสาร รถไฟแบบดั้งเดิม เครื่องบิน และรถไฟความเร็วสูงเมื่อเทียบกับรถยนต์ส่วนบุคคล และระยะทาง (DISTANCE) เป็นปัจจัยที่ส่งผลต่อพฤติกรรมการเลือกเดินทางด้วยรถโดยสาร รถไฟแบบดั้งเดิม และรถไฟความเร็วสูงเมื่อเทียบกับรูปแบบการเดินทางอื่น

$$V_{CAR} = B_TIME*TT_CAR + B_COST*TC_CAR \quad (3.2)$$

$$V_{BUS} = ASC_BUS + B_TIME*TT_BUS + B_COST*TC_BUS + B_AGE_BUS*AGE + B_HH_INCOME_BUS*HH_INCOME + B_DISTANCE_BUS*DISTANCE \quad (3.3)$$

$$V_{CR} = ASC_CR + B_TIME*TT_CR + B_COST*TC_CR + B_GENDER_CR*GENDER + B_HH_INCOME_CR*HH_INCOME + B_DISTANCE_CR*DISTANCE \quad (3.4)$$

$$V_{AIR} = ASC_AIR + B_TIME*TT_AIR + B_COST*TC_AIR + B_GENDER_AIR*GENDER + B_AGE_AIR*AGE + B_HH_INCOME_AIR*HH_INCOME \quad (3.5)$$

$$V_{HSR} = ASC_HSR + B_TIME*TT_HSR + B_COST*TC_HSR + B_AGE_HSR*AGE + B_HH_INCOME_HSR*HH_INCOME + B_DISTANCE_HSR*DISTANCE \quad (3.6)$$

แบบจำลองที่ 2 กำหนดให้ค่าสัมประสิทธิ์ของตัวแปรเวลาที่ใช้ในการเดินทาง (B_TIME) และค่าใช้จ่ายในการเดินทาง (B_COST) มีค่าแตกต่างกันไปในแต่ละทางเลือก (Specific Parameter) และมีสมมติฐานเกี่ยวกับเพศ (GENDER) อายุ (AGE) รายได้ในครัวเรือน (HH_INCOME) และระยะการเดินทาง (DISTANCE) เช่นเดียวกับแบบจำลองที่ 1

V_{CAR} = B_TIME_CAR*TT_CAR + B_COST_CAR*TC_CAR + (3.7)

V_{BUS} = ASC_BUS + B_TIME_BUS*TT_BUS + B_COST_BUS*TC_BUS +
B_AGE_BUS*AGE + B_HH_INCOME_BUS*HH_INCOME + (3.8)
B_DISTANCE_BUS*DISTANCE

V_{CR} = ASC_CR + B_TIME_CR*TT_CR + B_COST_CR*TC_CR +
B_GENDER_CR*GENDER + B_HH_INCOME_CR*HH_INCOME + (3.9)
B_DISTANCE_CR*DISTANCE

V_{AIR} = ASC_AIR + B_TIME_AIR*TT_AIR + B_COST_AIR*TC_AIR +
B_GENDER_AIR*GENDER + B_AGE_AIR*AGE + (3.10)
B_HH_INCOME_AIR*HH_INCOME

V_{HSR} = ASC_HSR + B_TIME_HSR*TT_HSR + B_COST_HSR*TC_HSR +
B_AGE_HSR*AGE + B_HH_INCOME_HSR*HH_INCOME + (3.11)
B_DISTANCE_HSR*DISTANCE

โดยที่ V_i คือ อรรถประโยชน์ของทางเลือก i โดยการศึกษานี้ประกอบด้วย 5 ทางเลือก ได้แก่ รถยนต์ส่วนบุคคล (CAR) รถโดยสาร (BUS) รถไฟแบบดั้งเดิม (CR) เครื่องบิน (AIR) และรถไฟความเร็วสูง (HSR)

ASC_i คือ ค่าคงที่ (Alternative Specific Constant) ของทางเลือก i

TT_i คือ เวลาที่ใช้ในการเดินทางด้วยรูปแบบการเดินทาง i

B_TIME คือ ค่าสัมประสิทธิ์ของตัวแปรเวลาที่ใช้ในการเดินทาง (Generic Parameter)

B_TIME_i คือ ค่าสัมประสิทธิ์ของตัวแปรเวลาที่ใช้ในการเดินทางด้วยรูปแบบการเดินทาง i (Alternative Specific Parameter)

TC_i คือ ค่าใช้จ่ายในการเดินทางด้วยรูปแบบการเดินทาง i

B_COST คือ ค่าสัมประสิทธิ์ของตัวแปรค่าใช้จ่ายในการเดินทาง (Generic Parameter)

B_COST_i คือ ค่าสัมประสิทธิ์ของตัวแปรค่าใช้จ่ายในการเดินทางด้วยรูปแบบการเดินทาง i (Alternative Specific Parameter)

B_GENDER_i คือ ค่าสัมประสิทธิ์ของตัวแปรเพศ ของผู้เดินทางด้วยรูปแบบการเดินทาง i

B_AGE_i คือ ค่าสัมประสิทธิ์ของตัวแปรอายุ ของผู้เดินทางด้วยรูปแบบการเดินทาง i

B_HH_INCOME_i คือ ค่าสัมประสิทธิ์ของตัวแปรรายได้ในครัวเรือนต่อเดือน ของผู้เดินทางด้วยรูปแบบการเดินทาง i

B_DISTANCE_i คือ ค่าสัมประสิทธิ์ของตัวแปรระยะทาง ของผู้เดินทางด้วยรูปแบบการเดินทาง i

หลังจากการกำหนดโครงสร้างของแบบจำลองการประมาณค่าพารามิเตอร์ของแบบจำลองการเลือกรูปแบบการเดินทางดำเนินการโดยใช้ชุดข้อมูลสำหรับฝึกสอน (Train Data Set) ซึ่งคิดเป็นร้อยละ 80 ของจำนวนตัวอย่างทั้งหมดในการประมาณค่าสัมประสิทธิ์ของตัวแปรโดยใช้แพ็กเกจ Apollo ในโปรแกรม R จากนั้นจึงทำการคัดเลือกแบบจำลองที่ดีที่สุด โดยพิจารณาจากค่าสถิติที่บ่งชี้ถึงความเหมาะสมของแบบจำลอง ได้แก่ ค่าความน่าจะเป็นสูงสุดของแบบจำลอง (Log-likelihood Value), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) และ Adjusted McFadden R^2 มาทำการทดสอบเพื่อตรวจสอบประสิทธิภาพและความสามารถในการพยากรณ์ของแบบจำลองโดยใช้ชุดข้อมูลสำหรับทดสอบ (Test Data Set) ซึ่งคิดเป็นร้อยละ 20 ของจำนวนตัวอย่างทั้งหมดที่ยังไม่เคยพบมาก่อนในระหว่างการฝึกสอน โดยใช้การตรวจสอบความสอดคล้องนอกกลุ่มตัวอย่าง (Out of Sample Fit)

3.5.3 การตรวจสอบความสอดคล้องนอกกลุ่มตัวอย่าง (Out of Sample Fit)

การตรวจสอบประสิทธิภาพและความสามารถในการพยากรณ์ของแบบจำลองดำเนินการโดยใช้วิธีการตรวจสอบความสอดคล้องนอกกลุ่มตัวอย่าง (Out of Sample Fit) โดยการแบ่งชุดข้อมูลออกเป็นสองส่วน ได้แก่ ข้อมูลส่วนที่ใช้สำหรับการประมาณค่า (Estimation Sample) และข้อมูลส่วนที่ใช้สำหรับการตรวจสอบความถูกต้อง (Validation Sample) เพื่อให้แน่ใจว่าแบบจำลองที่พัฒนาขึ้นไม่ได้เกิดภาวะ Overfitting และเพื่อไม่ให้เกิดความเอนเอียงในการประเมินความสอดคล้องของแบบจำลอง การทำ Cross Validation จึงถูกนำมาใช้ในกระบวนการตรวจสอบ โดยกระบวนการนี้เริ่มต้นด้วยข้อมูลที่ใช้สำหรับการประมาณค่า (Estimation Sample) จะถูกนำมาประมาณค่าพารามิเตอร์ของแบบจำลอง จากนั้นนำพารามิเตอร์ที่ได้จากการประมาณค่าข้างต้นไปใช้ในการคำนวณค่าสถิติเพื่อวัดความสอดคล้องของแบบจำลองโดยวัดจากค่าความน่าจะเป็นสูงสุดของแบบจำลอง (Log-likelihood Value) ของข้อมูลส่วนที่ใช้สำหรับตรวจสอบความถูกต้อง (Validation Sample) โดยการเปรียบเทียบความแตกต่างระหว่าง Log-likelihood ต่อ Observation ของกลุ่มข้อมูลที่เป็น Estimation Sample กับข้อมูล Validation Sample (Hess & Palma, 2019)