

บทที่ 2

ทฤษฎีที่ใช้ในการวิจัย

ในการวิจัยนี้ได้ทบทวนหลักทฤษฎีต่างๆที่จำเป็นในการหาปฏิสัมพันธ์ระหว่างการคมนาคมขนส่งและการใช้ที่ดิน และการสร้างแบบจำลองอันได้แก่

- ความสัมพันธ์ระหว่างการคมนาคมขนส่งและการใช้ที่ดิน
- แบบจำลองอรรถประโยชน์
- การวิเคราะห์ความสัมพันธ์แบบถดถอย

2.1 ความสัมพันธ์ระหว่างการคมนาคมขนส่งและการใช้ที่ดิน

ในการวิจัยนี้จะอ้างแนวคิดและทฤษฎีที่มีการศึกษาเกี่ยวกับความสัมพันธ์ของตัวแปรทางด้านการคมนาคมขนส่งและการใช้ที่ดินอันได้แก่ ราคาที่ดิน ความหนาแน่นของประชากรและตำแหน่งงาน ระยะห่างจากศูนย์กลาง ค่าขนส่งซึ่งอาจแสดงได้ทั้งในรูประยะทางหรือระยะเวลาในการเดินทาง ความสามารถเข้าถึงและ ความห่างไกลศูนย์กลาง

2.1.1 ความสัมพันธ์ระหว่างค่าใช้จ่ายที่เกิดจากการขนส่ง และราคาที่ดิน

ในการศึกษาปฏิสัมพันธ์ระหว่างการคมนาคมขนส่งและการใช้ที่ดินของ Kau (1976) ได้อ้างผลการศึกษาของ Wingo (1961) ที่ได้ทำการศึกษากิจกรรมคมนาคมขนส่งและที่ดินในเขตเมืองซึ่งชี้ให้เห็นว่าค่าใช้จ่ายในการคมนาคมขนส่งมีผลกระทบต่อกำไรของที่ดินและความต้องการพื้นที่ที่จะใช้เป็นที่อยู่อาศัย โดยค่าของที่ดินจะเพิ่มสูงขึ้นเท่ากับค่าใช้จ่ายในการคมนาคมที่สามารถจะลดลงได้หมายความว่าพื้นที่ใดที่อยู่ใกล้ศูนย์กลางเมืองซึ่งสามารถเดินทางเข้าสู่ศูนย์กลางด้วยระยะทางที่สั้นกว่าก็จะมีราคาสูง ฟังก์ชันความต้องการพื้นที่อยู่อาศัยขึ้นกับค่าของที่ดิน ซึ่งค่าของที่ดินจะสัมพันธ์กับค่าใช้จ่ายในการคมนาคมขนส่งซึ่งเป็นฟังก์ชันของระยะทางจากศูนย์กลางเมืองและความหนาแน่นประชากร และJames ยังได้อ้างถึงการศึกษา “City and Housing” ของ Muth’s (1969)

ที่ได้คิดหลักการ Household Equilibrium ที่อธิบายด้วยสมการดังนี้

$$q(k)p'(k) + T_k = 0 \quad (2.1)$$

เมื่อ

$q(k)$ = ความต้องการ Housing Service ต่อครัวเรือน

$p'(k)$ = อนุพันธ์ของราคา Housing Service ที่ขึ้นกับระยะห่าง k จากศูนย์กลางเมือง

T_k = ค่าใช้จ่ายในการเดินทาง ต่อระยะทาง 1 ไมล์ ที่ระยะห่าง k

ตามหลักการผู้บริโภคมต้องการการบริการที่มีราคาต่ำที่สุดเท่าที่จะเป็นไปได้ ฉะนั้น ความต้องการ Housing Service ย่อมเป็นส่วนกลับกับราคาของ Housing Service สมการที่ (2.1) จึงเขียนได้ใหม่เป็น

$$Ap(k)^{-1} p'(k) + T_k = 0 \quad (2.2)$$

ให้ A และ T_k คงที่แล้วอินทิเกรตสมการจะได้

$$p(k) = pe^{-rk} \quad (2.3)$$

เมื่อ $r = T_k / A$

ค่าของ Housing Service จะลดลงตามระยะทางในรูปแบบฟังก์ชันเอกซ์โปเนนเชียล ส่วนราคาของ Housing Service จะสัมพันธ์กับจำนวนประชากรและจำนวนครัวเรือน

การศึกษาของ Kau (1976) เองได้หาความสัมพันธ์ของค่าของที่ดินกับค่าใช้จ่ายในการคมนาคมขนส่ง โดยเขาพบว่าเมื่อเกิดการเปลี่ยนแปลงขึ้นในระบบขนส่งไม่ว่าจะเป็นเส้นทางหรือการสร้างโครงสร้างพื้นฐานทางด้านคมนาคมขนส่งขึ้นใหม่ จะเกิดผลกระทบต่อการใช้ที่ดินซึ่งแสดงในรูปแบบฟังก์ชันของราคาที่ดินที่ใช้เป็นที่ย่ออาศัยกับระยะรัศมีความห่างไกลจากศูนย์กลางเมือง (ซึ่งใช้แทนค่าใช้จ่ายในการคมนาคมขนส่ง) กล่าวคือการเปลี่ยนแปลงในระบบขนส่งทำให้ค่าใช้จ่ายในการคมนาคมลดลง เป็นผลให้รายได้ในส่วนที่จะมาจ่ายเพื่อเป็นเจ้าของที่ดินมีเพิ่มขึ้น ฟังก์ชันนี้อยู่ในรูปแบบเอกซ์โปเนนเชียล

$$R = R_0 e^{-rd} \quad (2.4)$$

โดยที่

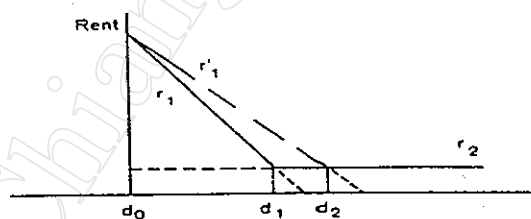
R = ค่าเฉลี่ยของที่ดินในพื้นที่นี้

R_0 = ค่าของที่ดินในเขตศูนย์กลางเมือง

r = rent gradient

d = ระยะห่างจากศูนย์กลางเมือง (ใช้แทนค่าใช้จ่ายในการคมนาคมขนส่ง)

และเขายังได้เปรียบเทียบความสัมพันธ์ระหว่างพฤติกรรมในการเสียเงินเพื่อเป็นเจ้าของที่ดินอยู่อาศัยของคนในเขตเมืองและคนในเขตชานเมือง(ซึ่งมีความแตกต่างกันในด้านค่าใช้จ่ายในการคมนาคมขนส่ง)ดังแสดงในรูปที่1 เมื่อเส้นกราฟ r_1 คือแนวโน้มการเปลี่ยนแปลงของค่าของที่ดินในเขตเมือง d_0 คือศูนย์กลางเมือง ระยะจาก d_0 ถึง d_1 คือระยะห่างที่ผู้อยู่อาศัยยอมจ่ายเงินแพงกว่าเพื่อที่จะได้อยู่ใกล้เมือง ราคาที่ดินในเขตชานเมืองหรือระยะจาก d_1 ออกไปแทนด้วยกราฟเส้น r_2 ความเปลี่ยนแปลงของโครงสร้างพื้นฐานทางด้านการคมนาคมขนส่งทำให้ค่าใช้จ่ายในการเดินทางลดลงส่งผลให้ความชันของกราฟค่าของที่ดินลดลงจาก r_1 เป็น r'_1 ซึ่งก็หมายความว่าผู้อยู่อาศัยต้องเสียเงินเพิ่มขึ้นเนื่องจากราคาที่ดินเพิ่มสูงขึ้นจึงเป็นส่วนผลักดันให้ต้องอยู่ห่างออกมาจากตัวเมืองไกลมากขึ้น(เท่ากับระยะระหว่าง d_1 กับ d_2)ถ้าจะรักษาในส่วนค่าใช้จ่ายให้คงที่



รูป 2.1 แสดงความสัมพันธ์ของราคาที่ดินระหว่างผู้อยู่อาศัยในเขตเมืองและผู้อยู่อาศัยในเขตชานเมือง
ที่มา : Kau (1976)

เมื่อ

r_1 = กราฟแสดงการเปลี่ยนแปลงค่าของที่ดินเริ่มแรก

r'_1 = กราฟค่าของที่ดินเมื่อค่าใช้จ่ายในการเดินทางลดลง

r_2 = กราฟค่าของที่ดินในเขตพื้นที่ชานเมือง

d = ระยะทางจากศูนย์กลางเมือง

ในกรณีที่การเปลี่ยนแปลงของราคาที่ดินไม่ได้เป็นเส้นตรง ถ้าหากมีการลดลงของค่าใช้จ่ายในการเดินทางค่าของที่ดิน การเปลี่ยนแปลงของราคาที่ดินจะไม่ได้ขึ้นกับระยะทางจากศูนย์กลางเพียงอย่างเดียว ยังขึ้นกับปัจจัยอื่นๆ เช่น ค่าคมนาคมขนส่งที่เปลี่ยนแปลงไปเป็นต้น จากหลักการดังกล่าวจึงต้องมีการศึกษาปัจจัยอื่นๆ ที่มีผลต่อค่าของที่ดินเพิ่มเติม



รูป 2.2 แสดงความสัมพันธ์ของราคาที่ดินระหว่างผู้อาศัยในเขตเมืองกับผู้อาศัยในเขตชานเมือง
กรณีที่การเปลี่ยนแปลงราคาที่ดินไม่ได้เป็นแบบเส้นตรง
ที่มา : Kua (1976)

2.1.2 ความสัมพันธ์ระหว่างระยะห่างจากศูนย์กลางและความหนาแน่นของประชากรและแรงงาน

Clark (1951) สร้างความสัมพันธ์ระหว่างความหนาแน่นของที่อยู่อาศัยกับระยะห่างจากศูนย์กลางเมืองที่วัดในลักษณะรัศมี (r) ดังนี้

$$D^P = a * \exp(-\beta r) \quad (2.5)$$

แปลงเป็น

$$\ln D^P = \ln a - \beta r = K - \beta r \quad (2.6)$$

พบว่าสมการของ Clark ไม่เหมาะสมในพื้นที่บริเวณศูนย์กลางเมือง เนื่องจากความหนาแน่นของที่อยู่อาศัยในบริเวณดังกล่าวจะลดลง ความหนาแน่นของประชากรน่าจะถูกแทนด้วยความหนาแน่นทั้งหมด ซึ่งหมายถึงรวมเอาความหนาแน่นของแรงงานหรือคนทำงานเข้าไปด้วย แต่อย่างไรก็ตามสมการนี้ก็ยังมีข้อจำกัดคือจะใช้ได้เฉพาะเมืองที่มีศูนย์กลางเพียงแห่งเดียว

$$D = D^P + \alpha D^E \quad (2.7)$$

โดยที่ D = ความหนาแน่นทั้งหมด
 D^P = ความหนาแน่นของประชากร
 D^E = ความหนาแน่นของแรงงาน
 α = พารามิเตอร์ที่ต้องประมาณค่า

2.1.3 ความสัมพันธ์ความสามารถเข้าถึงและความหนาแน่นของประชากรและแรงงาน

ความสามารถเข้าถึงเป็นหลักการที่ใช้ได้ดีในการศึกษาผลกระทบระหว่างคมนาคมขนส่งและการใช้ที่ดิน เนื่องจากสามารถประเมินค่าออกมาเป็นค่าค่าเดียวในแต่ละพื้นที่ และมีการเปลี่ยนแปลงที่เห็นได้อย่างชัดเจนเมื่อระบบคมนาคมขนส่งเกิดการเปลี่ยนแปลงอย่างใดอย่างหนึ่ง (K.B.Davidson, 1977) ความสามารถเข้าถึงเป็นตัวบอกความได้เปรียบของแหล่งที่ตั้งที่จะติดต่อสัมพันธ์กับพื้นที่อื่นๆเพื่อประกอบกิจกรรม กล่าวคือเป็นวิธีการวัดประโยชน์ของครัวเรือนหรือสถานประกอบการที่ตั้งอยู่ในพื้นที่ว่ามีความสะดวกในการใช้โครงสร้างพื้นฐานด้านคมนาคมขนส่งมากน้อยเพียงใด อันที่จริงแล้วความสามารถเข้าถึงและดัชนีวัดค่าความสามารถเข้าถึงต้องกำหนดหรือเลือกใช้เองตามปัญหาที่จะศึกษา เนื่องจากไม่มีนิยามหรือดัชนีตัวใดเป็นตัวดีที่สุด (ลำดวน ศรีศักดิ์, 2532) ในงานวิจัยนี้จะนิยามความเข้าถึงโดยคำนึงถึงโอกาสของบุคคลที่อยู่ ณสถานที่หนึ่งมีในการเข้าไปมีส่วนในกิจกรรมและกำหนดดัชนีความสามารถเข้าถึงซึ่งชี้ให้เห็นความใกล้-ไกลโดยในที่นี้จะวัดระยะทางและเวลาที่ใช้ในการเดินทาง

Hansen เป็นคนแรกที่สร้างหลักการอธิบายความสามารถเข้าถึง ในปี 1959

$$X_i = \sum_j S_j f(c_{ij}) \quad (2.8)$$

โดยที่ X_i = ความสามารถเข้าถึงของโซน i
 S_j = จำนวนโอกาสสำหรับวัตถุประสงค์อย่างใดอย่างหนึ่งที่โซน j
 $f(c_{ij})$ = ฟังก์ชันแทนการลดคุณค่าของโอกาสอันเนื่องมาจากสิ่งที่ต้องเสียไปเพื่อการเดินทางระหว่างโซน i และ j
 c_{ij} = ค่าใช้จ่ายที่วัดออกมาได้ในการเดินทางระหว่างโซน i และ j อาจอยู่ในรูปของเวลาที่ใช้ในการเดินทาง หรือระยะทางที่ใช้ในการเดินทาง

ถ้ายกเอาสมการของ Hansen มาอธิบายความสัมพันธ์ระหว่าง ความหนาแน่นและความสามารถเข้าถึง จะได้ว่า

$$\phi(\text{Land use Intensity}) = \text{Accessibility}$$

$$\phi(D_i) = X_i = \sum_j S_j f(c_{ij})$$

โดยที่ $\phi(D_i)$ เป็นฟังก์ชันในการวัดค่าความหนาแน่นในโซน i

Patton and Clark N (1970) มีแนวคิดว่าจะมีตัวที่อธิบายความสัมพันธ์กับความหนาแน่นได้ดีกว่าระยะห่างจากศูนย์กลาง(ซึ่งเป็นตัวแทนค่าใช้จ่ายในการเดินทาง) จึงได้พัฒนาทฤษฎีความสัมพันธ์ระหว่างความหนาแน่นทั้งหมดกับค่าใช้จ่ายในการเดินทางไปประกอบกิจกรรมต่างๆ ในศูนย์กลางเมือง โดยสร้างตัวแปรความห่างไกลศูนย์กลางซึ่งจะใช้อธิบายแทนค่าใช้จ่ายในการเดินทาง โคนินยามใหม่ของความห่างไกลศูนย์กลางมาจากนิยามของความสามารถเข้าถึง

$$\text{Centrality}, Y_i = \left(\frac{\sum_j S_j}{\sum_j S_j c_{ij}^{-n}} \right)^{1/n} = \left(\frac{\sum_j S_j}{X_{(n)i}} \right)^{1/n} \quad (2.9)$$

โดยที่ $X_{(n)i}$ คือ accessibility ที่คำนวณจาก power impedance function c_{ij}^{-n} ฉะนั้นสามารถเขียนนิยามของ accessibility ได้อีกแบบคือ

$$X_i = \sum_j S_j f(c_{ij}) = \left(\sum_j S_j \right) f(Y_i) \quad (2.10)$$

จากสมการด้านบนแสดงว่า $f(c_{ij})$ มีรูปแบบของความสัมพันธ์กับความสามารถเข้าถึงเช่นเดียวกันกับรูปแบบความสัมพันธ์ของความห่างไกลศูนย์กลางกับความสามารถเข้าถึง ค่าใช้จ่ายในการเดินทางที่แปรรูปมาจากแบบจำลอง gravity มีรูปแบบดังนี้

$$M_i = \sum_j c_{ij} T_{ij} / \sum_i T_{ij} \quad (2.11)$$

โดยที่ M_i คือ ค่าใช้จ่ายในการเดินทางของโซน i

T_{ij} คือ เกี่ยวการเดินทางจากโซน i ไป j

และนิยามของ T_{ij}

$$T_{ij} = P_i S_j c_{ij}^{-n} \quad (2.12)$$

เมื่อ P_i คือ เกี่ยวการเดินทางทั้งหมดที่ออกจากโซน i
ดังนั้น

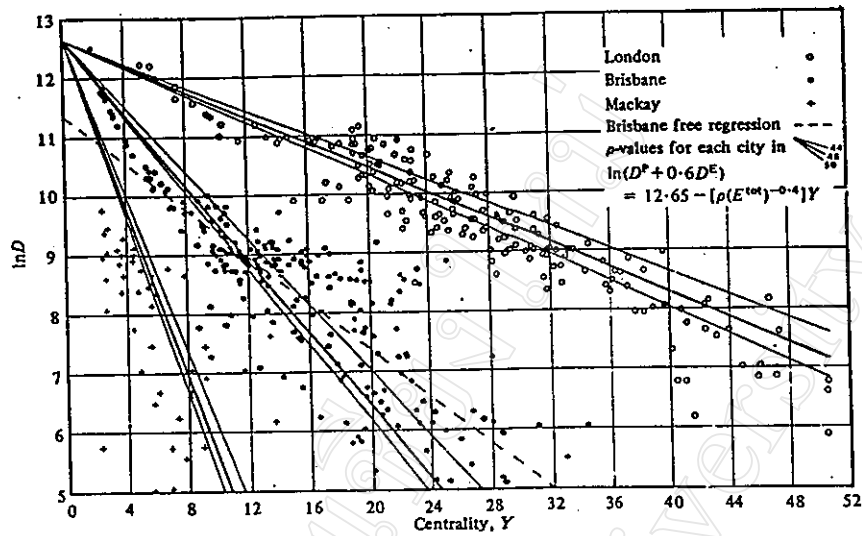
$$M_i = \frac{\sum_j c_{ij} P_i S_j c_{ij}^{-n}}{\sum_j P_i S_j c_{ij}^{-n}} = \frac{P_i \sum_j S_j c_{ij}^{1-n}}{P_i \sum_j S_j c_{ij}^{-n}} = \frac{X_{(n-1)i}}{X_{(n)i}} \quad (2.13)$$

จากสมการแสดงความสัมพันธ์ระหว่างความหนาแน่นของประชากร D^P กับระยะห่างจากศูนย์กลาง r ของ Clark (1951) แทนค่า ความหนาแน่นทั้งหมด D ลงใน D^P และแทนค่าความห่างไกลศูนย์กลางและ ค่าใช้จ่ายในการเดินทางลงใน r จะได้

$$\ln D_i = K - \beta Y_i = K - \beta \left(\frac{\sum_j S_j}{X_{(n)i}} \right)^{1/n} \quad (2.14)$$

$$\text{และ } \ln D_i = K - \beta M_i = K - \beta \frac{X_{(n-1)i}}{X_{(n)i}} \quad (2.15)$$

จากการเก็บข้อมูลการเดินทางโดยระบบขนส่งมวลชนในเขตเมือง London, Brisbane และ Mackay (Queensland) นำมาพล็อตกราฟและหาความสัมพันธ์ของ Centrality-Density ซึ่งรูปแบบของความสัมพันธ์คือ $\ln(D^P + 0.6D^E) = 12.65 - [\rho(E^{tot})^{-0.4}]Y$ โดยที่ E^{tot} ในคือจำนวนแรงงานทั้งหมดของแต่ละเมือง ในส่วนของ ρ จะมีค่าแตกต่างกันในแต่ละเมืองขึ้นกับจำนวนแรงงาน ($\rho = \beta (\sum_j S_j)^{1/n}$)



รูป 2.3 แสดงความสัมพันธ์ระหว่างความห่างไกลศูนย์กลางและความหนาแน่นในเมืองLondon, Brisbane และ Mackay

ที่มา : Clark (1951)

2.2 แบบจำลองอรรถประโยชน์ (Utility Model)

แบบจำลองที่ใช้หลักทฤษฎีอรรถประโยชน์ เป็นแบบจำลองที่นำเอาพฤติกรรมการตัดสินใจของคนมาเป็นพื้นฐานในการคิด โดยคนจะตัดสินใจเลือกสิ่งจะให้ประโยชน์แก่ตัวเองสูงสุด หรือเลือกสิ่งที่ทำให้เกิดความพึงพอใจสูงสุด เช่น ในการเสียค่าใช้จ่ายเพื่อเป็นเจ้าของหรือครอบครองที่ดินที่มีขนาดเท่ากันย่อมต้องการพื้นที่ที่จะสามารถเข้าถึงศูนย์กลางเมืองได้รวดเร็วหรือเสียค่าใช้จ่ายในการคมนาคมถูกกว่า เป็นต้น

2.2.1 ชนิดของแบบจำลองอรรถประโยชน์

แบบจำลองนี้สามารถแบ่งตามค่าอรรถประโยชน์ ได้เป็น 2 ประเภท คือ

ก. แบบจำลองแบบดีเทอร์มินิสติก (Deterministic Utility Model)

เป็นแบบจำลองที่ผู้กำหนดที่ตั้งสามารถกำหนดค่าอรรถประโยชน์ได้แน่นอน โดยผู้เลือกจะประมาณค่าอรรถประโยชน์ของการใช้ที่ดินตามปัจจัยที่มีผลต่อประโยชน์ที่คาดว่าจะ

ได้รับ แล้วใช้กฎการตัดสินใจ (Decision Rule) เลือกพื้นที่ที่ให้ค่าอรรถประโยชน์สูงสุดหรือความพึงพอใจสูงสุด นิยามฟังก์ชันอรรถประโยชน์อยู่ในรูป

$$U = F(\theta_k, X_k) \quad (2.16)$$

โดยอรรถประโยชน์หรือความพึงพอใจที่ได้รับจะแปรเปลี่ยนตามตัวแปรอิสระ ความสัมพันธ์ระหว่างความพึงพอใจกับตัวแปรอิสระมักถูกกำหนดให้เป็นความสัมพันธ์เชิงเส้นตรง ดังนี้

$$U_{in} = \sum_{k=1}^K \beta_k X_{ink} \quad (2.17)$$

เมื่อ U_{in} = อรรถประโยชน์หรือความพึงพอใจที่กลุ่มคนในพื้นที่ n เลือกการใช้ประโยชน์ที่ดินประเภท i
 X_k = ตัวแปรตัวที่ k ที่มีอิทธิพลต่อความพึงพอใจที่กลุ่มคนที่อาศัยในพื้นที่ n จะเลือกประเภทการใช้ที่ดิน i
 β_k = สัมประสิทธิ์ที่แสดงถึงอิทธิพลของตัวแปรตัวที่ k ที่มีต่อระดับความพึงพอใจ โดยในที่นี้ได้สมมุติว่าสัมประสิทธิ์ของแต่ละตัวแปรจะไม่แปรเปลี่ยนไปตามทางเลือก ทั้งนี้เพื่อความสะดวกในการนำเสนอ แต่ในการปฏิบัติ สัมประสิทธิ์ในแต่ละทางเลือกอาจจะแตกต่างกัน
 k = อันดับที่ 1, 2, 3, ..., K
 K = จำนวนของตัวแปรทั้งหมดที่นำมาพิจารณาในการกำหนดค่าอรรถประโยชน์

ยกตัวอย่างเพื่อให้เข้าใจได้ง่าย ดังนี้

$$U = \beta_1 X_1 + \beta_2 X_2 \quad (2.18)$$

เมื่อ X_1 = ค่าใช้จ่ายในการเดินทาง
 X_2 = เวลาในการเดินทาง

β_1, β_2 = พารามิเตอร์ของ X_1, X_2 ตามลำดับ

รูปแบบจำลองรูปแบบนี้มีข้อจำกัดหลายประการ เช่น ตัวแปรบางอย่างกำหนดลงไปเป็นแบบจำลองได้ยาก เช่น ตัวแปรของความสะดวกสบาย และแบบจำลองแบบนี้อธิบายได้เพียงการเลือกหรือไม่เลือกอย่างใดอย่างหนึ่งเท่านั้น ทำให้แบบจำลองใช้ได้กับบางกรณีเท่านั้น (สุทธิพงษ์, 2536)

ข. แบบจำลองแบบสโตแคสติก (Stochastic Utility Model)

เป็นแบบจำลองที่สร้างขึ้นโดยอาศัยความน่าจะเป็นที่กลุ่มคนในพื้นที่ที่จะกำหนดประเภทของการใช้ที่ดิน โดยสมมติว่าผู้กำหนดที่ตั้งไม่สามารถนิยามค่าอรรถประโยชน์ได้แน่นอน อย่างไรก็ตามถือว่าค่าอรรถประโยชน์ส่วนหนึ่งสามารถกำหนดได้ชัดเจน อีกส่วนหนึ่งมีค่าเชิงสุ่ม นิยามคือ

$$U_{in} = V_{in} + \varepsilon_{in} \quad (2.19)$$

เมื่อ U_{in} = ค่าอรรถประโยชน์หรือความพึงพอใจที่กลุ่มคนในพื้นที่ n จะได้รับเมื่อกำหนดการใช้ประโยชน์ที่ดินประเภท i

V_{in} = องค์ประกอบทางระบบ (Systematic Component) เป็นส่วนที่วัดค่าได้

ε_{in} = องค์ประกอบแบบสุ่ม (Random Component) เป็นส่วนที่รวมค่าความไม่แน่นอนเอาไว้

กฎในการตัดสินใจเลือกประเภทการใช้ที่ดินยังคงสมมติว่าผู้เลือกจะเลือกประเภทการใช้ที่ดินที่ให้ค่าอรรถประโยชน์สูงกว่า เพราะฉะนั้นความน่าจะเป็นที่กลุ่มคนในพื้นที่ n จะเลือกประเภทการใช้ที่ดิน i ขึ้นกับความน่าจะเป็นที่ $U_{in} \geq U_{jn}$ หรือ

$$P_n(i) = \text{Prob}(U_{in} \geq U_{jn}, \text{ for all } j \in M_n, j \neq i) \quad (2.20)$$

เมื่อ U_{in}, U_{jn} = ค่าอรรถประโยชน์ที่กลุ่มคนในพื้นที่ n ให้กับประเภทการใช้ที่ดิน i และ j ตามลำดับ

M_n = จำนวนทางเลือกที่มีอยู่

เนื่องจาก

$$U_{in} = V_{in} + \varepsilon_{in}$$

และ

$$U_{jn} = V_{jn} + \varepsilon_{jn}$$

ดังนั้น

$$P_n(i) = \text{Prob}[(V_{in} + \varepsilon_{in}) \geq (V_{jn} + \varepsilon_{jn}), \text{for all } j \in M_n, j \neq i]$$

$$= \text{Prob}[(V_{in} - V_{jn}) \geq (\varepsilon_{jn} - \varepsilon_{in}), \text{for all } j \in M_n, j \neq i] \quad (2.21)$$

กำหนดให้องค์ประกอบสุ่ม (Random Component) มีการกระจายแบบ Gumbel ที่เป็นอิสระต่อกัน จะได้

$$f[\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{jn}] = f(\varepsilon_{1n})f(\varepsilon_{2n}) \dots f(\varepsilon_{jn}) \quad (2.22)$$

จะได้ Probability Density Function :PDF

$$f(\varepsilon) = \exp[-\varepsilon - \exp(-\varepsilon)] \quad (2.23)$$

ได้ Cumulative Density Function : CDF

$$F(\varepsilon) = \exp[-\exp(\varepsilon)] \quad (2.24)$$

ให้ $f[\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{jn}]$ เป็นฟังก์ชันความน่าจะเป็นร่วม (Joint Probability Function) ขององค์ประกอบสุ่ม (random component) จะได้ความน่าจะเป็นของการเลือกในพื้นที่ที่ 1

$$P_n(1) = \int_{\varepsilon_{1n}=-\infty}^{\infty} \int_{\varepsilon_{2n}=-\infty}^{V_{1n}-V_{2n}+\varepsilon_{1n}} \dots \int_{\varepsilon_{jn}=-\infty}^{V_{1n}-V_{jn}+\varepsilon_{1n}} f[\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{jn}] d\varepsilon_{jn} \dots d\varepsilon_{1n} \quad (2.25)$$

แทน (2.22) ใน (2.25)

$$P_n(1) = \int_{\varepsilon_{1n}=-\infty}^{\infty} f(\varepsilon_{1n}) \int_{\varepsilon_{2n}=-\infty}^{V_{1n}-V_{2n}+\varepsilon_{1n}} f(\varepsilon_{2n}) \dots \int_{\varepsilon_{jn}=-\infty}^{V_{1n}-V_{jn}+\varepsilon_{1n}} f(\varepsilon_{jn}) d(\varepsilon_{jn}) \dots d(\varepsilon_{1n}) \quad (2.26)$$

$$\begin{aligned}
P_n(1) &= \int_{\varepsilon_{1n}=-\infty}^{\infty} f(\varepsilon_{1n}) * F[V_{1n} - V_{2n} + \varepsilon_{1n}] * \dots * F[V_{1n} - V_{jn} + \varepsilon_{in}] d(\varepsilon_{in}) \\
&= \int_{\varepsilon_{1n}=-\infty}^{\infty} f(\varepsilon_{1n}) \prod_{j=2}^J F[V_{in} - V_{jn} + \varepsilon_{in}] d(\varepsilon_{in})
\end{aligned} \tag{2.27}$$

หรือเขียนในรูปทั่วไป

$$P_n(1) = \int_{\varepsilon_{in}=-\infty}^{\infty} f(\varepsilon_{in}) \prod_{j \neq i}^J F[V_{in} - V_{jn} + \varepsilon_{in}] d(\varepsilon_{in}) \tag{2.28}$$

จากสมการ (2.23) และ (2.24) พิจารณาเทอม

$$\begin{aligned}
f(\varepsilon_{in}) \prod_{j \neq i}^J F[V_{in} - V_{jn} + \varepsilon_{in}] &= \prod_{j \neq i}^J \exp[-e^{-\varepsilon_{in} - V_{in} + V_{jn}}] \exp[-\varepsilon_{in}] = e^{-\varepsilon_{in}} \\
&= \exp[-\varepsilon_{in} - e^{-\varepsilon_{in}} (1 + \sum_{j \neq i}^J \frac{e^{V_{jn}}}{e^{V_{in}}})]
\end{aligned} \tag{2.29}$$

$$\text{กำหนดให้ } \lambda_m = \ln[1 + \sum_{j \neq i}^J \frac{e^{V_{jn}}}{e^{V_{in}}}] = \ln[\sum_{j \neq i}^J \frac{e^{V_{jn}}}{e^{V_{in}}}]$$

แทนค่า λ_m ลงในสมการ (2.29)

$$\begin{aligned}
f(\varepsilon_{in}) \prod_{j \neq i}^J F[V_{in} - V_{jn} + \varepsilon_{in}] &= \exp[-\varepsilon_{in} - e^{-\varepsilon_{in}} + e^{\lambda_m}] \\
&= \exp[-\varepsilon_{in} - e^{-(\varepsilon_{in} - \lambda_m)}]
\end{aligned} \tag{2.30}$$

แทนค่าสมการ (2.30) ใน (2.28)

$$P_n(i) = \int_{-\infty}^{\infty} \exp[-\varepsilon_{in} - e^{-(\varepsilon_{in} - \lambda_m)}] d(\varepsilon_{in})$$

$$P_n(i) = \exp(-\lambda_{in}) \int_{-\infty}^{\infty} \exp[-\varepsilon^* - e^{-\varepsilon^*}] d(\varepsilon^*) \quad (2.31)$$

เมื่อ

$$\varepsilon^* = \varepsilon_{in} - \lambda_{in}$$

$$P_n(i) = \exp(-\lambda_{in})$$

$$= \frac{e^{V_{in}}}{\sum_{j \neq i}^J e^{V_{jn}}} \quad (2.32)$$

โดยที่ $P_n(i)$ = ความน่าจะเป็นที่กลุ่มคนในพื้นที่ n จะเลือกประเภทการใช้ประโยชน์ที่ดิน i
 V_{in} = ค่าอรรถประโยชน์ในการเลือกประเภทการใช้ที่ดิน i ของกลุ่มคนในพื้นที่ n
 V_{jn} = ค่าอรรถประโยชน์ในการเลือกประเภทการใช้ที่ดิน j ของกลุ่มคนในพื้นที่ n
 (เป็นทางเลือกอื่นนอกเหนือจากตัวที่กำลังพิจารณา)
 J = จำนวนทางเลือกประเภทการใช้ที่ดินทั้งหมดที่มี

แบบจำลองวิเคราะห์ความน่าจะเป็นที่จะเลือกทางเลือกใดทางเลือกหนึ่งดังในสมการที่ 2.32 นั้นโดยทั่วไปเรียกว่า แบบจำลองประเภทโลจิต (Logit Model) ในกรณีที่ทางเลือกนั้นมีเพียง 2 ทางเลือก นิยมเรียกว่า Binary Logit Model (BNL)

$$P_n(i) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}} \quad (2.33)$$

$$\text{และ} \quad P_n(j) = 1 - P_n(i) \quad (2.34)$$

โดยที่ $V_{in} = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$
 $X_1 \dots X_p$ = ตัวแปรตัวที่ 1 ถึงตัวที่ P ที่มีอิทธิพลต่อการเลือกของกลุ่มคนในพื้นที่ n

$$\beta_1 \dots \beta_k = \text{สัมประสิทธิ์ที่แสดงถึงอิทธิพลของตัวแปรตัวที่ 1 ถึงตัวที่ P ที่มีต่อระดับความพึงพอใจ}$$

$$P_n(j) = \text{ความน่าจะเป็นที่กลุ่มคนในพื้นที่ n จะเลือกประเภทการใช้ประโยชน์ที่ดิน j}$$

และในกรณีที่แบบจำลองมีหลายทางเลือกนั้นจะเรียกว่า Multinomial Logit Model (MNL)

2.2.2 การประมาณค่าพารามิเตอร์โดยวิธี Maximum Likelihood

ในการนำวิธีทางสถิติมาประยุกต์ใช้ในการประมาณค่าพารามิเตอร์หรือค่าสัมประสิทธิ์ที่สะท้อนถึงอิทธิพลของตัวแปรต่างๆที่มีต่อระดับความพึงพอใจ นิยมใช้วิธี Maximum Likelihood วิธีดังกล่าวเป็นวิธีการประมาณค่าพารามิเตอร์โดยตรงที่นิยมใช้มากที่สุด เนื่องจากค่าพารามิเตอร์ที่ได้จากการประมาณโดยวิธีนี้จะทำให้ข้อมูลตัวอย่างในสมการมีค่าใกล้เคียงเดิมมาก (Ben – Akiva and Lerman, 1993)

กลไกการวิเคราะห์สามารถอธิบายได้ดังนี้ โดยการสุ่มตัวอย่างกลุ่มคนที่อยู่ในพื้นที่ N พื้นที่ และในแต่ละพื้นที่เราจะทราบว่ามีการใช้ประโยชน์ที่ดินแบบใด หากกำหนดให้ T_n เป็นทางเลือกที่ถูกเลือกจริง ดังนั้นโอกาสที่เราจะสุ่มเลือกกลุ่มคนที่อยู่ในพื้นที่ขึ้นมา N พื้นที่แล้วพบว่าการตัดสินใจเลือกประเภทการใช้ที่ดินของคนในพื้นที่ใดๆจะสอดคล้องกับที่ได้จากสำรวจมีค่าเท่ากับ

$$P_1(T_1) * P_2(T_2) * P_3(T_3) \dots P_N(T_N)$$

ซึ่งจะเรียกผลคูณนี้ว่าค่าความเป็นไปได้ (Likelihood)

ถ้าสมมติให้ตัวแปรเป็นอิสระต่อกันและเป็นข้อมูลแบบสุ่มจากประชากรทั้งหมด กล่าวคือตัวแปร x_n ซึ่งเป็นตัวแปรในฟังก์ชันของ y_n เป็น nonstochastic เราจะสามารถเขียนฟังก์ชัน likelihood ของพารามิเตอร์ β ของข้อมูลตัวอย่างได้ดังนี้

$$L^* = \prod_{n=1}^N f(y_n / x_n, \beta) \quad (2.35)$$

โดย Π คือ ผลคูณอันดับ N

เราจะหาค่าพารามิเตอร์ $\bar{\beta}_N$ ที่ให้ค่า L^* มากที่สุด ทำได้โดยการ maximize logarithm L^* ซึ่งจะไม่ทำให้ค่าพารามิเตอร์ที่ต้องการประมาณเปลี่ยนแปลงไปเนื่องจากคุณสมบัติของฟังก์ชันลอการิทึมคือเมื่อมีการเพิ่มขึ้นของค่าฟังก์ชันลอการิทึมจะเพิ่มขึ้นเพียงค่าเดียวไม่กระทบต่อตัวอื่น

$$\max_{\beta_N} \log L^* = \max_{\beta_N} L = \max_{\beta_N} \sum_{n=1}^N \log f(y_n / x_n, \bar{\beta}_N) \quad (2.36)$$

L มีค่าต่อเนื่องบน $\bar{\beta}_N$ ฉะนั้นจะทำการหาอนุพันธ์ย่อยอันดับที่หนึ่งของ L เทียบกับ $\bar{\theta}_N$ จะได้

$$\frac{\partial L}{\partial \bar{\beta}_{Nk}} = 0 \quad \text{สำหรับ } k = 1, \dots, K$$

เมื่อ $\bar{\beta}_{Nk}$ เป็นพารามิเตอร์ตัวที่ k ในเซตของ $\bar{\beta}_N$

$\bar{\beta}_N$ ที่ได้ อาจยังไม่ใช่ตัวที่ทำให้ค่า L มากที่สุด ต้องทำการตรวจสอบโดยหาอนุพันธ์อันดับสองของ L เทียบกับ $\bar{\beta}_N$ เพื่อหาค่าพารามิเตอร์ที่ให้ค่า L ที่มากที่สุด

การประมาณค่าโดยวิธี maximum likelihood มีขั้นตอนดังนี้

ก. จัดทำ likelihood function ให้อยู่ในรูปผลคูณของความน่าจะเป็นในการเลือกประเภทของการใช้ที่ดินของกลุ่มคนที่อาศัยในพื้นที่ n

$$\begin{aligned} L(\beta) &= \prod_{n=1}^N \prod_{i \in C_n} P_n(i)^{Y_{in}} \\ &= \prod_{n=1}^N \prod_{i \in C_n} \left[\frac{\exp(\beta X_{in})}{\sum_{j \in C_n} \exp(\beta X_{jn})} \right]^{Y_{in}} \end{aligned} \quad (2.37)$$

- เมื่อ L = likelihood function
 Π = ผลคูณอันดับ
 n = จำนวนกลุ่มคนที่กำหนดประเภทการใช้ที่ดินที่อาศัยอยู่ในพื้นที่ $1, 2, 3, \dots, N$
 C_n = เซตของทางเลือก (ประเภทการใช้ที่ดิน) ที่มีอยู่ทั้งหมด
 $P_n(i)$ = ความน่าจะเป็นที่กลุ่มคนที่อยู่ในพื้นที่ n เลือกการใช้ที่ดินประเภท i
 Y_{in} = 1 ในกรณีที่กลุ่มคนในพื้นที่ n เลือกประเภทการใช้ประโยชน์ที่ดิน i
 = 0 ในกรณีที่กลุ่มคนในพื้นที่ n เลือกประเภทการใช้ประโยชน์ที่ดินอื่น
 X_{in}, X_{jn} = ตัวแปรของประเภทการใช้ประโยชน์ที่ดิน i และ j ที่กลุ่มคนในพื้นที่ n เลือกตามลำดับ

ข. ทำ likelihood function ให้อยู่ในรูปของ natural logarithm

$$\ln L(\beta) = \sum_{n=1}^N \sum_{i \in C_n} Y_{in} [\beta X_{in} - \ln \sum_{j \in C_n} \exp(\beta X_{jn})] \quad (2.38)$$

ค. หาค่าพารามิเตอร์โดยการ differentiate $L(\beta)$ เทียบกับตัวแปร $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ ให้อยู่เท่ากับศูนย์ ดังนี้

$$\begin{aligned} \frac{\partial L(\beta)}{\partial \beta_k} &= \sum_{n=1}^N \sum_{i \in C_n} Y_{in} \left[X_{ink} - \frac{\sum_{j \in C_n} \exp(\beta X_{jn}) X_{jnk}}{\sum_{j \in C_n} \exp(\beta X_{jn})} \right] = 0 \\ &= \sum_{n=1}^N \sum_{i \in C_n} [Y_{in} - P_n(i)] X_{ink} = 0 \end{aligned} \quad (2.39)$$

ง. หาค่าพารามิเตอร์ต่างๆ โดยวิธี Newton-Raphson โดยมีขั้นตอนวิธีการดังนี้

1) กำหนดเวกเตอร์ของพารามิเตอร์ของฟังก์ชันยูติลิตี้ : X_0 โดยมากจะกำหนดให้ X_0 มีค่าเริ่มต้นเป็นศูนย์

2) แทนค่า X_0 เพื่อหา likelihood function

3) คำนวณหา X_{n+1} จากสมการ

$$X_{n+1} = X_n - (H_f |_{X_n})^{-1} \nabla f |_{X_n} \quad (2.40)$$

เมื่อ $\nabla f|_{X_n}$ = Gradient vector ของ X_n หาได้จาก partial derivation ของ loglikelihood function เทียบกับ β_{ij} ($i = 1, 2, 3, \dots, k$) ,

$$= \frac{\partial L}{\partial \beta_i}$$

$$= \left[\frac{\partial L}{\partial \beta_1} \frac{\partial L}{\partial \beta_2} \frac{\partial L}{\partial \beta_3} \dots \frac{\partial L}{\partial \beta_k} \right]^T$$

$H_f|_{X_n}$ = Hessian matrix ของ X_n หาได้จาก second partial ของ loglikelihood function เทียบกับ β_{ij} ($i, j = 1, 2, 3, \dots, k$)

$$= \frac{\partial^2 L}{\partial \beta_i \partial \beta_j}$$

$$= \begin{bmatrix} \frac{\partial^2 L}{\partial \beta_1 \partial \beta_1} & \frac{\partial^2 L}{\partial \beta_1 \partial \beta_2} & \dots & \frac{\partial^2 L}{\partial \beta_1 \partial \beta_k} \\ \frac{\partial^2 L}{\partial \beta_2 \partial \beta_1} & \frac{\partial^2 L}{\partial \beta_2 \partial \beta_2} & \dots & \frac{\partial^2 L}{\partial \beta_2 \partial \beta_k} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 L}{\partial \beta_k \partial \beta_1} & \frac{\partial^2 L}{\partial \beta_k \partial \beta_2} & \dots & \frac{\partial^2 L}{\partial \beta_k \partial \beta_k} \end{bmatrix}$$

n = อันดับการ iteration ครั้งที่ n

X_n = ค่า vector ของพารามิเตอร์ของ utility function เมื่อ iteration ครั้งที่ n

เนื่องจากวิธีการ Newton – Raphson นั้นเป็นกรรมวิธีกระทำซ้ำ (Iteration Method) ในการประมาณค่าพารามิเตอร์จึงต้องหยุดการกระทำซ้ำด้วยอัตราการลู่เข้า (% Convergence) ได้แก่

- อัตราการลู่เข้าของตัวแปรตาม (% Convergence for Dependent Value) ในที่นี้คือ

$\frac{\partial L(\beta)}{\partial \beta_k}$ ซึ่งหมายความว่า $\frac{\partial L(\beta)}{\partial \beta_k}$ มีค่าต่างจากศูนย์มากน้อยเท่าไร

$$\% \text{ Convergence} = \left[\frac{\partial L(\beta)}{\partial \beta_k} - 0 \right] * 100$$

- อัตราการลู่เข้าสำหรับส่วนเพิ่มของเวกเตอร์ตัวแปร (% Convergence for Attribute Value) คือการหาอัตราการเปลี่ยนแปลงของค่าพารามิเตอร์เทียบกับค่าของพารามิเตอร์นั้นๆ ที่ประมาณค่าขึ้นในแต่ละ iteration นั้นเอง

$$\% \text{ Convergence} = \frac{\Delta\beta_k}{\beta_k} * 100$$

โดยที่ส่วนเพิ่มของเวกเตอร์ตัวแปร (Attribute Value) คำนวณได้จากสูตร

$$\Delta\beta_k = \sum_{k=1}^K \sum_{n=1}^N \frac{\partial L(\beta)}{\partial \beta_n} \left[\frac{\partial^2 L(\beta)}{\partial \beta_k \partial \beta_n} \right]^{-1} \quad (2.41)$$

2.2.3 การทดสอบค่าทางสถิติ

เป็นการตรวจสอบความน่าเชื่อถือของแบบจำลองว่าค่าสัมประสิทธิ์ที่ได้จากการประมาณมีความเป็นเหตุเป็นผลหรือไม่ ซึ่งจะพิจารณาผลที่ได้จากการตรวจสอบหลายวิธี ได้แก่

ก. การตรวจสอบเครื่องหมายของค่าสัมประสิทธิ์

เครื่องหมายของค่าสัมประสิทธิ์แสดงถึงทิศทางของอิทธิพลที่ตัวแปรมีต่อการเลือกใช้ประโยชน์ที่ดินแต่ละประเภท ถ้าสัมประสิทธิ์มีเครื่องหมายเป็นบวกแสดงว่าความพึงพอใจจะสูงขึ้นตามค่าของตัวแปร ในทางกลับกันถ้าสัมประสิทธิ์มีเครื่องหมายเป็นลบแสดงว่าความพึงพอใจจะลดลงหากตัวแปรมีค่าสูงขึ้น

สำหรับค่าคงที่นั้นจะแสดงถึงอิทธิพลของปัจจัยอื่นๆที่ไม่ได้ถูกพิจารณาในแบบจำลอง เช่น ทักษะส่วนบุคคล สิ่งแวดล้อม เป็นต้น

ข. การทดสอบความสัมพันธ์ระหว่างตัวแปร

ในซอฟต์แวร์ที่ใช้ในการประมาณค่าพารามิเตอร์โดยวิธี maximum likelihood จะใช้ Likelihood Ratio Chi – Square เป็นสถิติทดสอบความอิสระของตัวแปรในแบบจำลอง ค่าสถิติทดสอบเป็น

$$\chi^2_{LR} = -2 \sum_i \sum_j O_{ij} \ln(E_{ij} / O_{ij}) \quad (2.42)$$

โดยที่ χ^2_{LR} = ค่าไคสแควร์ซึ่งมีการแจกแจงแบบไคสแควร์ที่องศาอิสระ $(r-1)(c-1)$
 O_{ij} = ความถี่หรือจำนวนของข้อมูลจริง
 E_{ij} = ความถี่หรือจำนวนข้อมูลที่คาดว่าจะเกิดถ้าตัวแปรทั้งสองเป็นอิสระกัน

สมมติฐานของการทดสอบคือ H_0 : ตัวแปรเป็นอิสระต่อกัน H_1 : ตัวแปรมีความสัมพันธ์กัน หากยอมรับสมมติฐานก็แสดงว่าตัวแปร ไม่มีความสัมพันธ์กัน

ค. การตรวจสอบระดับของความสอดคล้อง (Goodness of Fit)

การตรวจสอบระดับความสอดคล้องเป็นการตรวจสอบความสามารถของแบบจำลองที่จะอธิบายพฤติกรรมของการเลือกซึ่งปรากฏอยู่ในชุดข้อมูลที่นำมาใช้ประมาณหาค่าสัมประสิทธิ์ ซึ่งสามารถวัดได้ด้วยดัชนีวัดความสอดคล้อง (Likelihood Ratio Index; ρ^2) ดังนี้

$$\rho^2 = 1 - \frac{LL(\beta^*)}{LL(0)} \quad (2.43)$$

โดยที่ $LL(\beta^*)$ คือ ค่าลอการิทึมของฟังก์ชันความเป็นไปได้ที่ได้จากการประมาณค่าสัมประสิทธิ์
 $LL(0)$ คือ ค่าลอการิทึมของฟังก์ชันความเป็นไปได้ที่ได้ในกรณีที่สมมติให้สัมประสิทธิ์ทุกตัวมีค่าเท่ากับศูนย์

การสมมติให้สัมประสิทธิ์ทุกตัวมีค่าเท่ากับศูนย์จะส่งผลให้ความน่าจะเป็นที่จะเลือกแต่ละทางเลือกมีค่าเท่ากันหมดในทุกทางเลือก

ดัชนีวัดความสอดคล้องของแบบจำลองจะมีค่าอยู่ระหว่าง 0 กับ 1 คล้ายกับค่าสัมประสิทธิ์สหสัมพันธ์ (R^2) ที่นิยมใช้วัดความสามารถในการอธิบายความสัมพันธ์ระหว่างตัวแปร

ของสมการถดถอย (Regression Equation) แต่จะสื่อความหมายต่างกัน กล่าวคือ ค่าสัมประสิทธิ์สหสัมพันธ์จะชี้ถึงสัดส่วนของความผันแปรของตัวแปรตามที่สามารถอธิบายได้ด้วยสมการถดถอย ในขณะที่ดัชนีวัดความสอดคล้องจะแสดงถึงความสามารถของแบบจำลองที่จะอธิบายพฤติกรรมที่เกิดขึ้นจริง ถ้าดัชนีมีค่าเท่ากับ 1 แสดงว่าแบบจำลองสามารถอธิบายพฤติกรรมการเลือกประเภทการใช้ที่ดินได้ถูกต้องตามที่ได้มีการสำรวจจริง แต่ถ้ามีค่าเท่ากับ 0 แสดงว่าเพียงแต่ใช้การคาดเดาว่าจะมีโอกาสเลือกแต่ละทางเลือกเท่ากันก็สามารถอธิบายพฤติกรรมการเลือกได้ถูกต้องเท่ากับแบบจำลอง

2.3 การวิเคราะห์ความสัมพันธ์แบบถดถอย (Regression Analysis)

ในการวิเคราะห์ความสัมพันธ์ของตัวแปรการใช้ที่ดินและตัวแปรด้านการคมนาคมขนส่ง โดยให้ตัวแปรหนึ่งเป็นอิสระ และอีกตัวแปรหนึ่งเป็นตัวแปรตามจะใช้เทคนิคของการวิเคราะห์ความถดถอย (Regression Analysis) การที่จะวิเคราะห์ว่าตัวแปรสองตัวที่กำลังพิจารณามีความสัมพันธ์กันมากน้อยเพียงใด สมการถดถอยสามารถอธิบายความสัมพันธ์ได้ดีหรือไม่จะต้องอาศัยวิธีการทางสถิติที่จะให้ครรถษณินาชีวัด ในการวิจัยนี้ได้ใช้โปรแกรม SPSS ซึ่งเป็นโปรแกรมทางสถิติที่นิยมใช้กันอย่างแพร่หลายมาเป็นเครื่องมือใช้ในการวิเคราะห์

การวิเคราะห์ความสัมพันธ์แบบถดถอยของตัวแปรสองตัวจะต้องทราบก่อนว่าตัวแปรทั้ง 2 นั้นมีความสัมพันธ์กันในรูปแบบใด การวิเคราะห์เพื่อหารูปแบบความสัมพันธ์ในขั้นแรกจะต้องพล็อตกราฟเพื่อหาความสัมพันธ์ระหว่างตัวแปร 2 ตัว หากว่ากราฟแสดงความสัมพันธ์ในรูปเชิงเส้น ให้ใช้การวิเคราะห์ความถดถอยเชิงเส้น แต่ถ้าหากกราฟแสดงความสัมพันธ์ในรูปไม่เชิงเส้น ให้ใช้คำสั่ง Curve Estimation ซึ่งเป็นคำสั่งย่อยของคำสั่ง Regression ในโปรแกรม SPSS เพื่อหารูปแบบความสัมพันธ์ที่เหมาะสมสำหรับชุดข้อมูล

ขั้นตอนทางสถิติจะประกอบด้วย

- การประมาณค่าพารามิเตอร์ของสมการความถดถอยซึ่งจะทำให้ทราบถึงทิศทางความสัมพันธ์ว่ามีความสัมพันธ์แบบตามกันหรือตรงกันข้ามและความสัมพันธ์นั้นมากน้อยเพียงใด

- การทดสอบความเหมาะสมของสมการถดถอย

สถิติที่ใช้ในการทดสอบ คือ F – test หรือ T – test ซึ่งผลการทดสอบจะต้องเหมือนกัน

- สัมประสิทธิ์การตัดสินใจ (Coefficient of Determination R^2) ซึ่งหมายถึงสัดส่วนที่ตัวแปร x สามารถอธิบายการเปลี่ยนแปลงของตัวแปร Y ค่า R^2 จะไม่มีหน่วยและมีค่าอยู่ระหว่าง 0 ถึง 1 โดยถ้ามีค่าเข้าใกล้ 1 แสดงว่าเปอร์เซ็นต์ที่ x สามารถอธิบายการเปลี่ยนแปลงของ Y มีค่ามาก หรือ x และ Y มีความสัมพันธ์กันมาก แต่ถ้ามีค่าเข้าใกล้ 0 แสดงว่า เปอร์เซ็นต์ที่ x สามารถอธิบายการเปลี่ยนแปลงของ Y มีค่าน้อย