หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

ทฤษฎีและแนวความคิด

2.1 การเดินทางที่เกิดขึ้น (Trip Generation)

ความต้องการในการเดินทาง จะมีความสัมพันธ์กับองค์ประกอบที่เกี่ยวข้องโดยตรงกับกิจกรรมของมนุษย์ซึ่งจะเป็นในรูปแบบใดจะขึ้นอยู่กับลักษณะการใช้ที่ดิน (Land use) รวมไปถึงตำแหน่งที่ตั้งและลักษณะการใช้ประโยชน์ของกิจกรรมประเภทนั้น ๆ ซึ่งการเดินทางนี้จะมีลักษณะของ Trip Attraction เมื่อเกิดการเดินทางเข้าสู่แต่ละกิจกรรมในช่วงเวลาที่เกิดขึ้นๆ โดยไม่คำนึงถึงจุดเริ่มต้น (Origin) และหลังจากเข้ามาใช้บริการของกิจกรรมดังกล่าวแล้วจะเดินทางกลับซึ่งจะก่อให้เกิดลักษณะการเดินทางของ Trip Production ในช่วงเวลาที่เกิดและไม่คำนึงถึงจุดหมายปลายทาง (Destination) ช่วงเวลาที่เกิดนั้น ได้แก่ ช่วงเวลา 1 สัปดาห์ ช่วงเวลา 1 วัน หรือช่วงชั่วโมงเร่งด่วน เป็นต้น โดยการวัดจะอยู่ในรูปของปริมาณการเดินทางของยานพาหนะ (Vehicle trips)

ความสัมพันธ์ระหว่างปริมาณการเดินทางที่เกิดขึ้นกับลักษณะการใช้ที่ดิน สามารถสร้างในรูปแบบจำลองทางคณิตศาสตร์ (Mathematical Model) ที่เกี่ยวกับสัญญลักษณ์ (Symbolic) ที่ให้ความหมายแน่นอน ดังรูปแบบจำลองสำหรับหาปริมาณการเดินทางดังนี้คือ

แบบจำลองการเดินทางที่เกิดขึ้น (Trip Generation Model) ในรูปแบบจำลองทั่ว ๆ ไป (General Model)

$$G = a_o + a_1X_1 + a_2X_2 + ... + a_nX_n$$

โดยที่

G = ปริมาณการเดินทางที่เกิดขึ้น (Generated trips)

$a_o, a_1, ...a_n$ = พารามิเตอร์ (parameters) ของแบบจำลอง

$x_1, x_2, ...x_n$ = ตัวแปรอิสระ (Independent variable) แสดงคุณลักษณะการใช้ที่ดิน สภาพทางสังคมและเศรษฐกิจต่าง ๆ ที่เกี่ยวข้อง (Socio-economic Characteristics)

ในการวิเคราะห์แบบจำลองการเดินทางที่เกิดขึ้นจะใช้วิธีการวิเคราะห์การถดถอย (Regression technique)และใช้วิธีการของ Stepwise ในการหาสมการที่ดีที่สุดนั่นคือ ค่า $a_o, a_1, ..., a_n$ ที่เหมาะสมสำหรับความสัมพันธ์ระหว่างปริมาณการเดินทางกับข้อมูลทางลักษณะการใช้ที่ดิน

การวิเคราะห์การถดถอย (Regression Analysis) เป็นวิธีการทางสถิติที่ใช้หารูปแบบคณิตศาสตร์ ความสัมพันธ์ระหว่างตัวแปรที่ศึกษานี้ กำหนดให้ตัวแปรหนึ่งมีค่าแปรตามตัวแปรอื่น เรียกตัวแปรนั้นว่าตัวแปรตาม (Dependent Variable,Y) หรือตัวแปรตอบสนอง (Response Variable) ตัวแปรอื่นจะเป็นตัวกำหนดค่าของตัวแปรตาม เรียกว่าตัวแปรอิสระ (Independent Variable,X) หรือตัวแปรถดถอย (Regressor Variable) หรือ predictor ในการวิเคราะห์การถดถอยจะทำการสร้าง Model เป็นสมการถดถอย (Regression Equation) ขึ้นมาเพื่อพยากรณ์ตัวแปรตาม โดยมีตัวแปรอิสระเป็นตัวกำหนด ดังนั้นการวิเคราะห์การถดถอยจะทำการศึกษาความสัมพันธ์ในรูปแบบต่าง ๆ ได้แก่ สมการเส้นตรง สมการเส้นโค้งแบบต่าง ๆ เพื่อให้ได้แบบจำลองที่เหมาะสมและดีที่สุด ดังนี้

1. รูปแบบเส้นตรง (LINEAR FORM)

$$Y = a_o + a_1 x_1 + a_2 x_2 + ...$$

2. รูปแบบเส้นโค้ง GEOMETRIC

$$Y = a_o x_1^{a_1} x_2^{a_2} .....$$

ซึ่งสามารถแสดงในรูปแบบเส้นตรงที่แปลงจาก GEOMETRIC FORM ได้ดังนี้

$$lnY = lna_o + a_1 lnx_1 + a_2 lnx_2 +.....$$

3. รูปแบบเส้นโค้ง EXPONENTIAL

$$Y = a_o \; a_1^{x_1} \; a_2^{x_2} .....$$

ซึ่งสามารถแสดงในรูปแบบเส้นตรงที่แปลงจาก EXPONENTIAL FORM ได้ดังนี้

$$lnY = lna_o+(lna_1)x_1+(lna_2)x_2+......$$

สามารถสรุปค่าต่าง ๆ ของแบบจำลองได้ดังนี้

1. สัมประสิทธิ์เส้นถดถอย (Regression Coefficient;a_o, a_1) โดยที่ a_1 เป็นความชันของเส้นถอถอย คือค่าที่แสดงถึงการเปลี่ยนแปลงในค่าเฉลี่ยของ Y เมื่อ X เพิ่มขึ้นหนึ่งหน่วย a_o เป็นค่าของ Y-intercept ของเส้นถดถอย ถ้าขอบเขตของรูปแบบได้รวมค่า X = 0 แล้ว a_o จะเป็นค่าเฉลี่ยของ Y เมื่อ X = 0 แต่ถ้าขอบเขตของรูปแบบไม่ได้รวม X = 0 แล้ว a_o ก็จะไม่มีความหมายนอกจากเป็นพจน์ (term) ที่ประกอบกันขึ้นเป็นรูปแบบการถดถอย ในการประมาณค่า a_o และ a_1 จะใช้วิธีกำลังสองน้อยที่สุด (method of least squares) ดังสูตรการคำนวณต่อไปนี้

$$a_o = \bar{Y} - a_1\bar{X}$$

$$a_1 = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}$$

โดยที่

$$\bar{Y} = \frac{\Sigma Y}{n}$$

$$\bar{X} = \frac{\Sigma X}{n}$$

สัมประสิทธิ์เส้นถดถอยสามารถเป็นตัวตรวจสอบความสัมพันธ์ระหว่างตัวแปรทั้งสองจะมีต่อกันหรือไม่ ซึ่งค่านี้จะบอกให้ทราบเกี่ยวกับ

1.1 ความสัมพันธ์ระหว่างตัวแปร นั่นคือ ถ้าค่าของสัมประสิทธิ์เส้นถดถอยเท่ากับศูนย์ แสดงว่าตัวแปรทั้งสองไม่มีความสัมพันธ์กัน

1.2 อัตราการเปลี่ยนแปลงของ Y เมื่อ X เปลี่ยนแปลงไปแต่ละหน่วยนั่นคือ เมื่อ x เปลี่ยนแปลงไปหนึ่งหน่วย ค่า Y จะเปลี่ยนแปลงไป a_1 หน่วยเป็นต้น

2. ค่าการประมาณการคลาดเคลื่อน (THE STANDARD ERROR OF ESTIMATE;SEE) เป็นค่าที่จะบอกให้ทราบว่าข้อมูลที่รวบรวมมาได้นั้นมีการกระจายไปจากเส้นถดถอยมากน้อยเพียงใด ถ้าค่าของ Y แต่ละค่าที่รวบรวมได้แตกต่างจากเส้นถดถอยที่ประมาณขึ้น Y_e มากการประมาณค่าของ Y โดยอาศัยเส้นถดถอยก็จะไม่ดีเท่าที่ควร แต่ถ้าค่าที่รวบรวมได้แต่ละค่าอยู่ใกล้ ๆ หรือ

อยู่บนเส้นถดถอยค่าที่ประมาณได้จากเส้นถดถอยนี้จะให้ความเชื่อถือได้มากขึ้น ซึ่งจะมีลักษณะคล้ายๆ การวัดค่าเบี่ยงเบนมาตรฐาน (STANDARD DEVIATION) จากค่าเฉลี่ย (MEAN) เช่นกรณีของสมการเส้นถดถอยอย่างง่ายคือมีเพียงตัวแปรเดียว จะหาค่าได้จากความสัมพันธ์ดังนี้

$$SEE = \sqrt{\frac{\Sigma(Y-Y_e)^2}{n-k}}$$

โดยที่ n = จำนวนคู่ของข้อมูล

k = จำนวนค่าพารามิเตอร์ในสมการเส้นถดถอย

Y = ค่าของตัวแปรตาม (Dependent Vaviable)

Y_e = ค่าของตัวแปรตามที่คำนวนได้จากสมการเส้นถดถอย

3. การทดสอบความมีนัยสำคัญเกี่ยวกับสัมประสิทธิ์เส้นถดถอย (t-test) เป็นการตรวจสอบความสัมพันธ์ระหว่างตัวแปรว่ามีค่านัยสำคัญ (Significant level;α) หรือความเชื่อ (confidence Value =1-α) มากน้อยแค่ไหนเพื่อเป็นการยืนยันความสัมพันธ์ระหว่างตัวแปรได้แน่นอน ซึ่งสามารถหาได้จากความสัมพันธ์ สำหรับสมการเส้นถดถอยอย่างง่าย ซึ่งกำหนดสมมติฐานเพื่อทดสอบดังนี้

B แทนสัมประสิทธิ์เส้นถดถอยแท้จริงในประชากร

H_o : B = 0 หมายความว่า ไม่มีความสัมพันธ์ระหว่างตัวแปร

H_1 : B ≠ 0 หมายความว่า x มีอิทธิพลต่อการเปลี่ยนแปลงค่า Y

$$t_{n-2} = \frac{(a_n-B)}{SEE}\sqrt{\Sigma(X-X)^2}$$

โดยที่ x = ค่าตัวแปรอิสระ (Indepent variable)

$\bar{x}$ = ค่าเฉลี่ยของตัวแปรอิสระ

เปรียบเทียบค่า t_{n-2} ที่ได้กับค่าวิกฤต $t_{\alpha/2,(n-2)}$ ถ้าค่า t_{n-2} อยู่ในบริเวณวิกฤต (critical region) จะปฏิเสธ H_o และยอมรับ H_1

4. ค่าสัมประสิทธิการตัดสินใจ (The coefficient of determination;R^2) เป็นดัชนีที่ใช้บอกความสัมพันธ์ระหว่าง Y กับ x ซึ่งทำให้สามารถตัดสินใจได้ว่าควรนำ x มาพิจารณาในการวิเคราะห์เกี่ยวกับ Y หรือไม่ ถ้าค่า R^2 เข้าใกล้ศูนย์ แสดงความสัมพันธ์ระหว่าง x กับ Y ในแบบเส้นตรง(Linear correlation) มีน้อยมาก แต่ถ้าค่า R^2 เข้าใกล้ 1 แสดงว่า x มีอิทธิพลหรือมีความสัมพันธ์กับ Y มาก และเป็นไปในแบบเส้นตรง ค่า R^2 สามารถหาได้จากความสัมพันธ์ดังนี้

$$R^2 = \frac{\Sigma(Y_e-\bar{Y})^2}{\Sigma(Y-\bar{Y})^2}$$

$$= 1- \frac{\Sigma(Y-Y_e)^2}{\Sigma(Y-\bar{Y})^2}$$

ดังนั้น $0 \leq R^2 \leq 1$

5. ค่าสัมประสิทธิ์สหสัมพันธ์ (The coefficient of correlation,R) เป็นตัวบอกระดับความสัมพันธ์แบบเส้นตรงระหว่าง x กับ Y ซึ่งหาจากรากที่สองของค่าสัมประสิทธิการตัดสินใจ (R^2)

ดังนั้น $-1 \leq R \leq 1$ และเครื่องหมายของ R เหมือนกับเครื่องหมายของสัมประสิทธิ์สันถดถอย($a_1,a_2....a_n$) ค่าของ R จะแสดงความสัมพันธ์ดังนี้

5.1 R = -1 แสดงว่ามีความสัมพันธ์กันอย่างเต็มที่ระหว่าง X กับ Y ในแบบเส้นตรง แต่เป็นความสัมพันธ์แบบผกผัน

5.2 R = 1 แสดงว่ามีความสัมพันธ์กันอย่างเต็มที่ระหว่าง X กับ Y ในแบบเส้นตรง

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

และในการเลือกสมการที่มีหลายตัวแปรอิสระให้ได้สมการที่เหมาะสมที่สุดของแต่ละแบบ จะใช้วิธีการพิจารณาของโปรแกรมที่เรียกว่า STEPWISE REGRESSION ซึ่งมีขั้นตอนการคัดเลือกตัวแปรอิสระดังนี้

1. จากค่าสัมประสิทธิ์สหสัมพันธ์ (R) จะเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามที่ต้องการศึกษามากที่สุดนั่นคือ มีค่า R ใกล้ 1 มากกว่าตัวแปรอิสระตัวอื่น

2. ตรวจสอบออกมาว่าตัวแปรอิสระนั้นพอเพียงสำหรับการวิเคราะห์เกี่ยวกับปัจจัยที่ต้องการหรือไม่ ดูได้จากค่าสัมประสิทธิ์การตัดสินใจ (R^2) นั่นคือถ้าR^2 มีค่าเข้าใกล้ 1 แสดงว่า x มีความสัมพันธ์กับ Y มาก และเป็นไปในแบบเส้นตรงและดูจากค่าสัมประสิทธิ์การตัดสินใจส่วนย่อย (The coefficient of partial determination) สำหรับข้อมูลที่มีตัวแปรอิสระมากกว่าหนึ่งตัว

3. ถ้าต้องการตัวแปรอิสระอีก ให้ตรวจสอบตัวแปรที่มีความสัมพันธ์สูงสุดกับตัวแปรตาม จากตัวแปรอิสระที่เหลือเข้ามา

4. ตรวจสอบสมการถดถอยโดยการวิเคราะห์ความแปรปรวนว่า เป็นเส้นถดถอยที่ดีหรือไม่โดยวิธี t-test ซึ่งทดสอบจากเส้นถดถอยโดยส่วนรวม และทดสอบสัมประสิทธิ์เส้นถดถอย (Regression coefficient ,a_n) แต่ละค่า