บทที่ 2

ทฤษฎี

2.1 ความหมายของการถดถอย (REGRESSION)

การถดถอยเป็นวิธีทางสถิติอย่างหนึ่งที่ใช้ในการตรวจหาหรือตรวจสอบลักษณะของความสัมพันธ์ระหว่างข้อมูลตั้งแต่ 2 ชุดขึ้นไป โดยมีวัตถุประสงค์ว่าจะนำเอาลักษณะความสัมพันธ์อันนี้ไปใช้ให้เป็นประโยชน์ในอนาคต การถดถอยสามารถจำแนกได้เป็นชนิดตามจำนวนของตัวแปรดังนี้

2.1.1 การถดถอยเชิงเดียว (SIMPLE REGRESSION)

2.1.2 การถดถอยเชิงซ้อน (MULTIPLE REGRESSION)

2.2 ความหมายของการถดถอยเชิงเดียว (SIMPLE REGRESSION), การถดถอยเชิงซ้อน (MULTIPLE REGRESSION), และการถดถอยแบบเส้นโค้ง (CURVILINEAR REGRESSION)

2.2.1 การถดถอยเชิงเดียว เป็นการศึกษาหาความสัมพันธ์ระหว่างตัวแปร 2 ชุดเท่านั้น โดยตัวแปรชุดหนึ่งเป็นปัจจัยตาม และอีกชุดหนึ่งเป็นปัจจัยอิสระ หรือกล่าวง่าย ๆ ได้ว่าเป็นการหาความสัมพันธ์ของข้อมูลที่พิจารณาให้มีตัวแปรอิสระเพียงตัวเดียวเท่านั้น เขียนในรูปฟังก์ชันได้ว่า

$$Y = f(X)$$

2.2.2 การถดถอยเชิงซ้อน เป็นการศึกษาหาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 3 ชุดขึ้นไป โดยตัวแปรชุดหนึ่งเป็นปัจจัยตาม ส่วนตัวแปรอื่น ๆ เป็นปัจจัยอิสระ กล่าวง่าย ๆ คือมีตัวแปรอิสระตั้งแต่ 2 ตัวขึ้นไป เขียนเป็นฟังก์ชันในรูปสัญลักษณ์ได้ดังนี้

$$Y = f(X_1, X_2, X_3)$$

รูปฟังก์ชั่นทั่ว ๆ ไปของการถดถอยเชิงซ้อน กรณีมีตัวแปรอิสระ k ตัว

$$Y = f(X_1, X_2, X_3, \ldots X_k)$$

2.2.3 การถดถอยแบบเส้นโค้ง ในการศึกษาหาความสัมพันธ์ของตัวแปรไม่ว่าจะเป็นแบบการถดถอยเชิงเดียวหรือเชิงซ้อนก็ตาม อาจมีความสัมพันธ์ได้ทั้งในรูปที่เป็นเส้นตรงและเส้นโค้ง ถ้าความสัมพันธ์ที่หาได้เป็นรูปเส้นตรง เรียกว่า การถดถอยแบบเส้นตรงเชิงเดียว (SIMPLE LINEAR REGRESSION) หรือการถดถอยแบบเส้นตรงเชิงซ้อน (MULTIPLE LINEAR REGRESSION) แต่ลักษณะของข้อมูลบางประเภทอาจไม่เป็นเส้นตรงก็ต้องศึกษาในรูปเส้นโค้ง ซึ่งเรียกว่า การถดถอยแบบเส้นโค้ง (NON LINEAR REGRESSION หรือ CURVILINEAR REGRESSION) ดังนั้นการศึกษาการถดถอยแบบเส้นโค้งจึงมีการศึกษาทั้งที่เป็นแบบการถดถอยเชิงเดียว และแบบการถดถอยเชิงซ้อน

2.3 แผนภาพกระจาย (SCATTER DIAGRAM), เส้นการถดถอย (REGRESSION LINE) และสมการถดถอย (REGRESSION EQUATION)

ในการวิเคราะห์ความสัมพันธ์ของตัวแปรนั้น สิ่งที่นำมาใช้เป็นหลักในการศึกษาก็คือ ข้อมูลที่แสดงความสัมพันธ์เกี่ยวกับตัวแปรนั้น ๆ นั่นเอง เมื่อได้ข้อมูลแล้วจำต้องนำข้อมูลมาพิจารณาอย่างคร่าว ๆ ก่อนว่า ตัวแปรเหล่านี้จะมีความสัมพันธ์กันหรือไม่ และสัมพันธ์กันในรูปใด เส้นตรงหรือเส้นโค้งในการพิจารณาความสัมพันธ์ของตัวแปรอย่างง่าย ๆ นั้น นิยมนำข้อมูลมาพลอทในแผ่นกราฟ โดยแกนตั้งแทนค่าด้วยตัวแปรตาม แกนนอนแทนด้วยค่าตัวแปรอิสระ ซึ่งเรียกในภาษาของการถดถอยว่า พลอทแผนภาพกระจาย (PLOT SCATTER DIAGRAM)

จากแผนภาพกระจาย ทำให้สามารถมองเห็นลักษณะความสัมพันธ์ระหว่างจุดที่พลอทได้บนแผนภาพกระจาย ถ้าลากเส้นผ่านจุดเหล่านี้ก็จะได้เส้นซึ่งถือว่าเป็นตัวแทนที่แสดงลักษณะหรือแนวโน้มของความสัมพันธ์ เส้นนี้เรียกว่า เส้นการถดถอย (REGRESSION LINE) เส้นการถดถอยจะเป็นเส้นตรงหรือเส้นโค้งขึ้นอยู่กับลักษณะจุดที่พลอทได้บนแผนภาพกระจาย ในกรณีเส้นการถดถอยเป็นเส้นตรง ถ้าให้ α เป็นค่าตัดแกน Y (Y-INTERCEPT) ของเส้นตรงนี้ และ β เป็นความลาดชัน (SLOPE) ของเส้นตรง ในทางคณิตศาสตร์สามารถเขียนสมการเส้นตรงแทนเส้นการถดถอยแบบเส้นตรงได้ว่า

$$Y_i = \alpha + \beta X_i \qquad (2.1)$$

สมการนี้เรียกว่า สมการการถดถอยแบบเส้นตรงเชิงเดียว (SIMPLE LINEAR REGRESSION EQUATION)

2.4 การศึกษาการถดถอยแบบเส้นตรงเชิงเดียวกรณีข้อมูลเป็นข้อมูลตัวอย่าง (SAMPLE SIMPLE LINEAR REGRESSION ANALYSIS)

เนื่องจากในทางปฏิบัติไม่สามารถเก็บข้อมูลของค่า Y ได้ทุกค่าในแต่ละประชากรย่อย เพราะปกติจะทราบค่า Y ได้เพียงบางค่าในแต่ละค่าของ X ซึ่งถือว่าเป็นข้อมูลตัวอย่าง ดังนั้น การศึกษาความสัมพันธ์ระหว่าง X และ Y เมื่อทราบข้อมูลเป็นเพียงตัวอย่าง จึงไม่ใช่ความสัมพันธ์ระว่าง X และ Y อย่างแท้จริง เป็นแต่เพียงความสัมพันธ์ที่คาดคะเนเท่านั้น สำหรับการศึกษาการถดถอยในกรณีที่ข้อมูลเป็นเพียงตัวอย่าง เรียกว่า การวิเคราะห์การถดถอยแบบเส้นตรงเชิงเดียวของข้อมูลตัวอย่าง (SAMPLE SIMPLE LINEAR REGRESSION ANALYSIS)

ด้วยเหตุผลข้างต้นความคลาดเคลื่อนของสมการ (2.1) ย่อมเกิดขึ้นได้ดังสมการ (2.2)

$$Y_i = a + b X_i + e_i \qquad (2.2)$$

เมื่อ e_i = ค่าความคลาดเคลื่อน (ERROR TERM)

a = ค่าที่เส้นการถดถอยตัดแกน Y (Y-INTERCEPT)

b = ค่าความลาด (SLOPE) ของเส้นการถดถอย

เรียก b ว่า สัมประสิทธิ์ของการถดถอย (COEFFICIENT OF REGREESSION) สมการ (2.2) เรียกว่า แบบหุ่นการถดถอยแบบเส้นตรงเชิงเดียวของข้อมูลตัวอย่าง (SAMPLE SIMPLE LINEAR REGRESSION MODEL)

ให้ $\hat{Y}_i = Y_i - e_i$

เมื่อ $\hat{Y}_i$ = ค่า Y_i บนเส้นการถดถอยตาม X_i ที่กำหนด

สมการ (2.2) จะกลายเป็น $\hat{Y}_i = a + bX_i \qquad (2.3)$

2.5 การวิเคราะห์หาสมการการถดถอยโดยวิธีกำลังสองน้อยที่สุดอย่างธรรมดา (ORDINARY LEAST SQUARE METHOD = OLSM)

วิธีกำลังสองน้อยที่สุดอย่างธรรมดาเป็นเทคนิคอันหนึ่งสำหรับหาค่าตัวคาดคะเน (ESTIMATOR) ของสมการ $\hat{Y}_i = a + bX_i$ ให้มีคุณสมบัติตามต้องการ , กล่าวคือ

2.5.1 ตัวคาดคะเนไม่มีความคลาดเคลื่อน (UNBIASED ESTIMATOR)

2.5.2 ความแปรปรวน (VARIANCE) ต่ำสุด

เส้นการถดถอยที่คาดคะเนจากวิธี OLSM จะมีคุณสมบัติดังนี้

- $\Sigma e_i = 0$
- Σe_i^2 = มีค่าน้อยที่สุด
- เส้นการถดถอยจะผ่านจุดตัดของ $\bar{X}$ และ $\bar{Y}$ เมื่อ $\bar{X} = \Sigma X/n$ และ $\bar{Y} = \Sigma Y/n$

ค่า a และ b หาได้จาก

$$a = \bar{Y} - b\bar{X} \qquad (2.4)$$

$$b = [\Sigma X_i Y_i - (\Sigma X_i \Sigma Y_i/n)] / [\Sigma X^2{}_i - (\Sigma X_i)^2/n] \qquad (2.5)$$

2.6 สัมประสิทธิ์ของตัวกำหนด (COEFFICIENT OF DETERMINATION) (r^2)

อัตราส่วนระหว่างผลรวมกำลังสองของการกระจายที่อธิบายได้กับผลรวมกำลังสองของการกระจายทั้งหมดเรียกว่า สัมประสิทธิ์ของตัวกำหนด (COEFFICIENT OF DETERMINATION) (r^2)

$$r^2 = [(\Sigma x_i y_i)^2] / [\Sigma x_i^2 . \Sigma y_i^2] \qquad (2.6)$$

2.6.1 คุณสมบัติของ r^2

2.6.1.1 r^2 มีค่าเป็นบวกเสมอ

2.6.1.2 r^2 มีค่าตั้งแต่ 0 ถึง 1 หรือมีค่าตั้งแต่ 0% ถึง 100%

2.6.2 ประโยชน์ของ r^2

2.6.2.1 r^2 เป็นเครื่องวัดความใกล้ชิดระหว่างเส้นการถดถอยกับค่า Y บนแผนภาพการกระจาย หมายความว่า ถ้า

- $r^2 = 1$ ค่า Y จะอยู่บนเส้นการถดถอยทุกจุด
- r^2 มีค่ามาก Y จะอยู่ใกล้ชิดกับเส้นการถดถอย
- r^2 มีค่าน้อย Y จะอยู่ห่างจากเส้นการถดถอย
- r^2 มีค่าเท่ากับศูนย์ Y จะกระจายห่างจากกันกับเส้นการถดถอยมากจนหาแนวโน้มที่ถูกต้องไม่ได้

2.6.2.2 r^2 เป็นเครื่องวัดความเหมาะสมของเส้นการถดถอยว่า จะแสดงแนวโน้มของข้อมูลบนแผนภาพการกระจายได้มากน้อยเพียงใดเช่น กรณีที่ศึกษาถึงการถดถอยของข้อมูลบนเส้นตรง จะได้ว่าถ้า

- r^2 มีค่าใกล้เคียง 1 แสดงว่าแนวโน้มของข้อมูลจะเป็นลักษณะแนวเส้นตรงมากที่สุด
- r^2 มีค่าใกล้เคียงศูนย์ แสดงว่าแนวโน้มของข้อมูลจะมีลักษณะไม่เป็นเส้นตรง ซึ่งอาจจะเป็นเส้นโค้งก็ได้

2.6.2.3 r^2 เป็นเครื่องแสดงอิทธิพลของตัวแปรอิสระ (X) ที่มีค่าตัวแปรตาม (Y) ทั้งนี้เพราะว่าจากค่า r^2 จะบอกให้ทราบว่า การกระจายทั้งหมดของค่า Y นั้นสามารถอธิบายได้จากเส้นการถดถอยกี่เปอร์เซนต์ หรือกล่าวได้ว่า X มีอิทธิพลต่อ Y กี่เปอร์เซนต์ เช่น $r^2 = 0.7958$ หมายความว่าสมการการถดถอยนี้สามารถอธิบายการกระจายของ Y ได้ 79.58% หรือค่า X ที่กำลังพิจารณาอยู่นั้นมีอิทธิพลต่อ Y เท่ากับ 79.58% ส่วนตัวแปรอิสระอื่น ๆ นอกเหนือไปจาก X รวมแล้วจะมีอิทธิพลต่อ Y เท่ากับ 100 - 79.58 = 20.42%

2.7 การวิเคราะห์สหสัมพันธ์แบบเส้นตรงเชิงเดียว

การวิเคราะห์สหสัมพันธ์แบบเส้นตรงเชิงเดียวมุ่งสนใจเฉพาะค่าสัมประสิทธิ์ของสหสัมพันธ์แบบเส้นตรงเชิงเดียว (COEFFICIENT OF CORRELATION) ใช้สัญลักษณ์ r โดยศึกษาถึง

- การวัดหรือตรวจสอบหาความมากน้อยของความสัมพันธ์ระหว่างตัวแปร 2 ชุด โดยไม่ต้องการทราบว่าอะไรควรเป็นเหตุ อะไรควรเป็นผล

- การวัดหรือตรวจสอบว่า สมการการถดถอยที่ศึกษานั้นเป็นสมการที่เหมาะสมกับข้อมูลหรือไม่

สัมประสิทธิ์สหสัมพันธ์ $r = [\Sigma (X_i-\bar{X})(Y_i-\bar{Y})] / [\Sigma (X_i-\bar{X})^2 \Sigma (Y_i-\bar{Y})^2]^{1/2}$

$$= [\Sigma x_i y_i] / [\Sigma x_i^2 \Sigma y_i^2]^{1/2} \qquad (2.7)$$

2.7.1 คุณสมบัติค่า r

r จะมีค่าตั้งแต่ -1 ถึง +1 ถ้า r เป็น + หมายความว่า ตัวแปร 2 ตัวนั้นมีความสัมพันธ์ไปในทางเดียวกัน แต่ถ้าค่า r เป็น - หมายความว่า ตัวแปร 2 ตัวนั้นมีความสัมพันธ์กันไปในด้านตรงกันข้าม

2.8 การวิเคราะห์การถดถอยแบบเส้นตรงในกรณีมีตัวแปรอิสระหลายตัว (MULTIPLE LINEAR REGRESSION ANALYSIS)

2.8.1 แบบหุ่นเส้นตรง (MULTIPLE LINEAR REGRESSION MODEL)

ให้ Y = ตัวแปรตาม

X_1 = ตัวแปรอิสระตัวที่ 1

X_2 = ตัวแปรอิสระตัวที่ 2

$\vdots$ $\quad$ $\vdots$

X_k = ตัวแปรอิสระตัวที่ k

ถ้าสมมุติว่า $X_1, X_2, \ldots\ldots, X_k$ มีอิทธิพลต่อ Y แบบรูปเส้นตรงจะเขียนเป็นฟังก์ชันได้ดังนี้คือ

$$Y = f(X_1, X_2, \ldots\ldots X_k)$$

เขียนความสัมพันธ์เป็นรูปสมการดังนี้

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + e_i \qquad (2.8)$$

α = ค่าที่ตัดแกน Y ของสมการเส้นตรง

β_i = ค่าสัมประสิทธิ์ของการถดถอย (REGRESSION COEFFICIENT) ของ Y ที่มีต่อ X_i เมื่อกำหนดให้ X ตัวอื่น ๆ คงที่ เรียกค่า β_i ว่าสัมประสิทธิ์ของการถดถอยแต่ละตัว (PARTIAL REGRESSION COEFFICIENT)

e_i = ค่าความคลาดเคลื่อน (ERROR TERM)

สมการ (2.8) เรียกว่า แบบหุ่นของการถดถอยแบบเส้นตรงเชิงซ้อน (MULTIPLE LINEAR REGRESSION MODEL)

เมื่อ $\hat{Y}_i = Y_i - e_i$

สมการ (2.8) จะกลายเป็น

$$\hat{Y}_i = a + b_1 X_{1i} + b_2 X_{2i} + \ldots + b_k X_{ki} \qquad (2.9)$$

เรียกสมการ (2.9) ว่า สมการการถดถอยแบบเส้นตรงเชิงซ้อน กรณีที่ข้อมูลได้จากการสุ่ม (SAMPLE MULTIPLE LINEAR REGRESSION EQUATION)

กรณีที่มีตัวแปรอิสระ 2 ตัว ค่า a หาได้จาก

$$a = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2 \qquad (2.10)$$

ค่า b_1, b_2 หาได้จากการแก้สมการ

$$\Sigma x_{1i}y_i = b_1 \Sigma x_{1i}^2 + b_2 \Sigma x_{1i}x_{2i} \qquad (2.11)$$

$$\Sigma x_{2i}y_i = b_1 \Sigma x_{1i}x_{2i} + b_2 \Sigma x_{2i}^2 \qquad (2.12)$$

โดยที่ $\Sigma x_{1i}y_i = [\Sigma X_{1i} Y_i] - [(\Sigma X_{1i} \Sigma Y_i) / n]$

$$\Sigma x^2_{1i} = [\Sigma X^2_{1i}] - [(\Sigma X_{1i})^2 / n]$$

$$\Sigma x_i y_j = [\Sigma X_i X_j] - [(\Sigma X_i \Sigma X_j) / n]$$

$i \neq j$

2.9 ค่าสัมประสิทธิ์ของตัวกำหนด (COEFFICIENT OF DETERMINATION)

ให้ $R^2_{y.123...k}$ เป็นค่าสัมประสิทธิ์ของตัวกำหนดในการถดถอยแบบเส้นตรงเชิงซ้อน

$$R^2_{y.123...k} = \frac{\sum_{j=1}^{k} b_j \left(\sum_{i=1}^{n} x_{ji} y_i \right)}{\Sigma y_i^2} \qquad (2.13)$$

k = จำนวนตัวแปรอิสระ

n = จำนวนข้อมูลที่สุ่มมา

2.10 การวิเคราะห์การถดถอยเส้นโค้ง (CURVILINEAR REGRESSION ANALYSIS)

การวิเคราะห์การถดถอยของข้อมูลเมื่อตัวแปรมีความสัมพันธ์กันแบบเส้นโค้งนั้นมีลักษณะเหมือนกับกรณีที่เป็นเส้นตรงทุกอย่าง จะแตกต่างกันเฉพาะที่รูปร่างของเส้นการถดถอยและตัวแบบหุ่น (MODEL) เท่านั้น ตัวแบบหุ่นการถดถอยสำหรับวิเคราะห์ความสัมพันธ์ของตัวแปรแบบเส้นโค้งเรียกว่า CURVILINEAR MODELS ซึ่งมีอยู่หลายแบบด้วยกัน แต่ละแบบมีลักษณะของเส้นการถดถอยที่ต่างกัน ตัวอย่างแบบหุ่นของการถดถอยแบบเส้นโค้งที่สำคัญมี

2.10.1 แบบหุ่นดับเบิลลอก (DOUBLE LOG MODEL)

แบบหุ่นตัวแทนคือ $Y = a X^b .f(error)$

2.10.2 แบบหุ่นเอกซ์โพเนนเชียล (EXPONENTIAL MODEL)

แบบหุ่นตัวแทนคือ $Y = ab^x .f(error)$

2.10.3 แบบหุ่นเซมิ-ลอก (SEMI-LOG MODEL)

แบบหุ่นตัวแทนคือ $Y = 10^{a+bx+f(error)}$

หรือ $Y = e^{a+bx+f(error)}$

เมื่อ e = ค่า EXPONENTIAL

2.10.4 แบบทุ่นเรซิโปรคอล (A RECIPROCAL MODEL)

แบบทุ่นตัวแทนคือ $Y = 1/[a+bX+f(error)]$

2.11 การวิเคราะห์สมการการถดถอยเส้นโค้ง

การวิเคราะห์ใช้วิธีการเดียวกับการวิเคราะห์การถดถอยแบบเส้นตรง ต่างกันที่ก่อนการวิเคราะห์จะต้องมีการเปลี่ยนหรือแปลงแบบทุ่นให้อยู่ในลักษณะเป็นแบบทุ่นของการถดถอยเส้นตรงเสียก่อน การเปลี่ยนแบบทุ่นการถดถอยเส้นโค้งให้เป็นแบบเส้นตรงมีหลายวิธี ที่สำคัญมีอยู่ 2 วิธี คือ

2.11.1 วิธีที่ 1 ใช้วิธีแปลงรูป (TRANSFORMED) โดยใช้ลอกฐาน 10 หรือลอกฐาน e แบบการถดถอยเส้นโค้งเมื่อแปลงรูปแล้วจะมีลักษณะเป็นเส้นตรง ตัวอย่างเช่น

รูปแบบเดิม : $Y = a X^b . f(error)$.

รูปแบบลอก : $\log Y = \log a + b \log X + \log f(error)$

2.11.2 วิธีที่ 2 ใช้วิธีสมมุติค่า (REDEFINED) ขึ้นมาให้อยู่ในรูปเส้นตรง ตัวอย่างเช่น

$$Y = 1 / [a+bX+f(error)]$$

ให้ $Z = (1/Y) = a+bX+f(error)$ ซึ่งอยู่ในรูปเส้นตรงในตัวเอง

2.12 การพิจารณาความเหมาะสมของสมการการถดถอย

ในการศึกษาการถดถอยของข้อมูลชุดใดก็ตาม สมการการถดถอยจะดีมีประสิทธิภาพ ถ้ามีลักษณะดังนี้

- มีค่าสัมประสิทธิ์ของตัวกำหนด r^2 สูงพอควร
- ค่าสัมประสิทธิ์ของการถดถอย (b_1) ต้องแตกต่างจากศูนย์อย่างมีนัยสำคัญที่ระดับความเชื่อมั่นสูง
- ค่าสัมประสิทธิ์ของการถดถอย (b_1) ต้องมีขนาดและเครื่องหมายสอดคล้องกับทฤษฎี

2.13 การเลือกตัวแปรอิสระ

ในการวิเคราะห์การถดถอยแบบหลายชั้นนั้น การใช้ตัวแปรอิสระหลาย ๆ ตัวบางครั้งทำให้เกิดความยุ่งยากทั้งในด้านการคำนวณ การนำไปใช้ประโยชน์ เพราะตัวแปรอิสระบางตัวที่ใส่เข้าไปในสมการอาจไม่มีนัยสำคัญทางสถิติเลยก็ได้ ดังนั้นในทางปฏิบัติจึงหาทางขจัดตัวแปรอิสระที่ไม่มีความสำคัญต่อตัวแปรตามทิ้งเสีย แล้วจัดลำดับตัวแปรอิสระตามความสำคัญ

วิธีการจัดลำดับตัวแปรอิสระที่มีความสำคัญต่อตัวแปรตามมีหลายวิธี ในที่นี้ขอกล่าวถึงวิธี ออล โพสสิเบิล รีเกรชชัน (ALL POSSIBLE REGRESSION)

2.14 วิธี ออล โพสสิเบิล รีเกรชชัน (ALL POSSIBLE REGRESSION)

เป็นวิธีการกำจัดตัวแปรอิสระวิธีหนึ่ง โดยการนำตัวแปรอิสระทุกตัวที่คาดว่าจะมีความสัมพันธ์ต่อตัวแปรตาม แล้วใช้ค่าสัมประสิทธิ์ของตัวกำหนด (R^2) เป็นตัวตัดสิน ซึ่งมีวิธีการวิเคราะห์เป็นขั้นดังนี้

2.14.1 ขั้นที่ 1 ใส่ตัวแปรอิสระทุกตัวที่คาดว่าจะมีอิทธิพลต่อตัวแปรตามลงในสมการ ทั้งนี้โดยการใส่ตัวแปรอิสระลงไปแต่ละตัวก่อน สมมุติว่ามีตัวแปรอิสระ 4 ตัว

ก. หาสมการที่มีตัวแปรอิสระ 1 ตัวจากโมเดล

$$Y = a + b_1 X_1 + e_1$$
$$Y = a + b_2 X_2 + e_2$$
$$Y = a + b_3 X_3 + e_3$$
$$Y = a + b_4 X_4 + e_4$$

ข. หาสมการประกอบด้วยตัวแปรอิสระ 2 ตัวจากโมเดล

$$Y = a + b_1 X_1 + b_2 X_2 + e_{12}$$
$$Y = a + b_1 X_1 + b_3 X_3 + e_{13}$$
$$Y = a + b_1 X_1 + b_4 X_4 + e_{14}$$
$$Y = a + b_2 X_2 + b_3 X_3 + e_{23}$$
$$Y = a + b_2 X_2 + b_4 X_4 + e_{24}$$
$$Y = a + b_3 X_3 + b_4 X_4 + e_{34}$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ค. หาสมการที่มีตัวแปรอิสระ 3 ตัวจากโมเดล

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + e_{123}$$
$$Y = a + b_1 X_1 + b_2 X_2 + b_4 X_4 + e_{124}$$
$$Y = a + b_1 X_1 + b_3 X_3 + b_4 X_4 + e_{134}$$
$$Y = a + b_2 X_2 + b_3 X_3 + b_4 X_4 + e_{234}$$

ง. หาสมการที่มีตัวแปรอิสระ 4 ตัว จากโมเดล

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 + e_{1234}$$

2.14.2 ขั้นที่ 2 หาค่าสัมประสิทธิ์ของตัวกำหนดของแต่ละสมการในทุกกลุ่ม

ก. R^2_{y1}, R^2_{y2}, R^2_{y3}, R^2_{y4}

ข. $R^2_{y.12}$, $R^2_{y.13}$, $R^2_{y.14}$, $R^2_{y.23}$, $R^2_{y.24}$, $R^2_{y.34}$

ค. $R^2_{y.123}$, $R^2_{y.124}$, $R^2_{y.134}$, $R^2_{y.234}$

ง. $R^2_{y.1234}$

ต่อจากนั้นก็พิจารณาค่า r^2 ในแต่ละกลุ่ม แล้วเลือกค่า r^2 ที่สูงสุดของแต่ละกลุ่มมาศึกษาโดยอาจทำเป็นตารางดังนี้

กลุ่มที่ 1	จำนวนตัวแปรอิสระในสมการ	R^2
ก.	$Y = f(X_4)$	67.50%
ข.	$Y = f(X_1, X_4)$	97.90%
ค.	$Y = f(X_1, X_2, X_4)$	98.234%
ง.	$Y = f(X_1, X_2, X_3, X_4)$	98.237%

ในกรณีนี้ควรเลือกกลุ่ม ข. คือ ตัวแปรอิสระ X_1 X_4 ซึ่งให้ค่า $R^2_{y.14} = 97.90\%$ ทั้งนี้เพราะเมื่อเพิ่มตัวแปรอิสระ X_2 เข้าไปอีก 1 ตัว ก็จะทำให้ค่า $R^2_{y.124}$ สูงขึ้นอีกไม่ถึง 1% หรือถ้าเพิ่ม X_3 เข้าไปอีก 1 ตัว ค่า $R^2_{y.1234}$ สูงขึ้นเพียงเล็กน้อย แต่การคำนวณยุ่งยากมาก ดังนั้นสมการที่น่าจะดีที่สุดคือ

$$\hat{Y} = a + b_1 X_1 + b_4 X_4$$

2.15 ทฤษฎีทางวิศวกรรม

2.15.1 คุณสมบัติพื้นฐาน (BASIC PROPERTIES)

2.15.1.1 ขีดจำกัดแอตเตอร์เบริก (ATTERBERG LIMIT) เป็นวิธีอธิบายปริมาณความชื้นในมวลดินซึ่งจะบ่งถึงสถานภาพของดินนั้น ประกอบด้วย

2.15.1.1.1 ขีดจำกัดเหลว (LIQUID LIMIT), LL คือปริมาณความชื้นในดินที่น้อยที่สุดที่ทำให้ดินอยู่ในสภาวะไหลตัวได้เป็นขีดแบ่งระหว่างสถานะพลาสติก (PLASTIC) กับสถานะเหลว (LIQUID)

2.15.1.1.2 ขีดจำกัดพลาสติก (PLASTIC LIMIT), PL คือปริมาณความชื้นที่ทำให้ดินเปลี่ยนสถานะพลาสติกเป็นสถานะกึ่งแข็ง (SEMI-SOLID)

2.15.1.1.3 ขีดจำกัดหดตัว (SHRINKAGE LIMIT), SL คือปริมาณความชื้นในดินที่มากที่สุดที่ไม่ทำให้ดินเปลี่ยนปริมาตรเมื่อแห้ง

ขีดจำกัดเหลว ขีดจำกัดพลาสติก และขีดจำกัดหดตัว แสดงดังรูปที่ 2.1

2.15.1.2 ดัชนีพลาสติก (PLASTICITY INDEX), PI คือปริมาณความชื้นในสถานะพลาสติกของดิน เป็นผลต่างระหว่างขีดจำกัดเหลวกับขีดจำกัดพลาสติก

$$PI = LL - PL$$

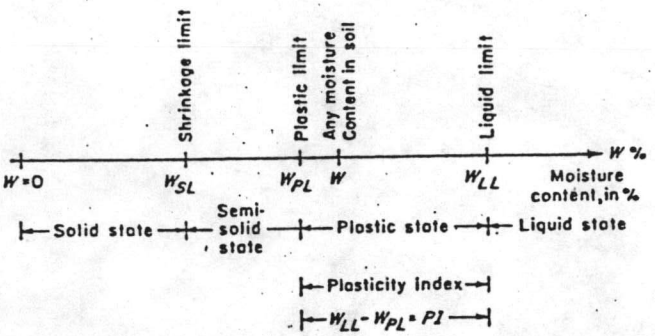
ค่า PI เป็นตัวแสดงถึงความเหนียวของดินและยังแสดงความไวต่อการเปลี่ยนสถานภาพต่อความชื้นของมวลดินนั้น เป็นค่าที่สำคัญ ใช้มากในการจำแนกมวลดิน แสดงดังตารางที่ 2.1

2.15.1.3 ปริมาณความชื้น (WATER CONTENT), w คืออัตราส่วนน้ำหนักของน้ำต่อน้ำหนักของเม็ดดิน ส่วนมากแสดงเป็นร้อยละ

$$w = (W_w / W_s) \times 100\ \%$$

2.15.1.4 หน่วยน้ำหนักรวม (TOTAL UNIT WEIGHT), γ_T คือน้ำหนักทั้งหมดของดินต่อหนึ่งหน่วยปริมาตรของดิน

$$\gamma_T = W/V$$

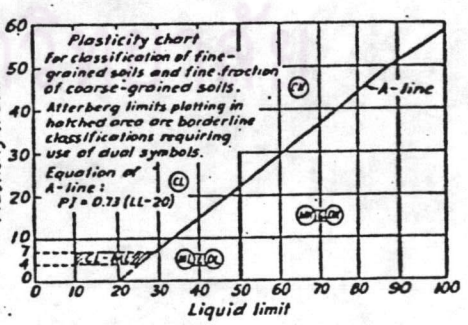


รูปที่ 2.1 แสดงสถานภาพความข้นเหลวของดินเหนียว

Major Divisions			Group Symbols	Typical Names	Classification Criteria
Coarse-Grained Soils More than 50% retained on No. 200 sieve	Gravels 50% or more of coarse fraction retained on No. 4 sieve	Clean Gravels	GW	Well-graded gravels and gravel-sand mixtures, little or no fines	$C_u = D_{60}/D_{10}$ Greater than 4; $C_z = \frac{(D_{30})^2}{D_{10} \times D_{60}}$ Between 1 and 3
			GP	Poorly graded gravels and gravel-sand mixtures, little or no fines	Not meeting both criteria for GW
		Gravels with Fines	GM	Silty gravels, gravel-sand-silt mixtures	Atterberg limits plot below "A" line or plasticity index less than 4
			GC	Clayey gravels, gravel-sand-clay mixtures	Atterberg limits plot above "A" line and plasticity index greater than 7
	Sands More than 50% of coarse fraction passes No. 4 sieve	Clean Sands	SW	Well-graded sands and gravelly sands, little or no fines	$C_u = D_{60}/D_{10}$ Greater than 6; $C_z = \frac{(D_{30})^2}{D_{10} \times D_{60}}$ Between 1 and 3
			SP	Poorly graded sands and gravelly sands, little or no fines	Not meeting both criteria for SW
		Sands with Fines	SM	Silty sands, sand-silt mixtures	Atterberg limits plot below "A" line or plasticity index less than 4
			SC	Clayey sands, sand-clay mixtures	Atterberg limits plot above "A" line and plasticity index greater than 7
Fine-Grained Soils 50% or more passes No. 200 sieve	Silts and Clays Liquid limit 50% or less		ML	Inorganic silts, very fine sands, rock flour, silty or clayey fine sands	
			CL	Inorganic clays of low to medium plasticity, gravelly clays, sandy clays, silty clays, lean clays	
			OL	Organic silts and organic silty clays of low plasticity	
	Silts and Clays Liquid limit greater than 50%		MH	Inorganic silts, micaceous or diatomaceous fine sands or silts, elastic silts	
			CH	Inorganic clays of high plasticity, fat clays	
			OH	Organic clays of medium to high plasticity	
Highly organic soils			Pt	Peat, muck and other highly organic soils	Visual-manual Identification

Classification on basis of percentage of fines
Less than 5% Pass No. 200 sieve: GW, GP, SW, SP
More than 12% Pass No. 200 sieve: GM, GC, SM, SC
5% to 12% Pass No. 200 sieve: Borderline classification requiring use of dual symbols

Atterberg limits plotting in hatched area are borderline classifications requiring use of dual symbols



ตารางที่ 2.1 แสดงการจำแนกดินระบบUNIFIED SOIL CLASSIFICATION(ASTM D-2487)

2.15.1.5 ความถ่วงจำเพาะของดิน (SPECIFIC GRAVITY), G_s คือ อัตราส่วนของหน่วยน้ำหนักของดินต่อหน่วยน้ำหนักของน้ำ (ที่อุณหภูมิ 4° C)

$$G_s = \gamma_s / \gamma_w$$

ค่าความถ่วงจำเพาะหาได้จากการทดลองในห้องปฏิบัติการ ตัวอย่างค่าความถ่วงจำเพาะของดิน แสดงดังในตารางที่ 2.2

2.15.2 คุณสมบัติทางกำลัง (STRENGTH PROPERTIES)

2.15.2.1 UNCONFINED COMPRESSIVE STRENGTH TEST เป็นการทดสอบเพื่อหาค่ากำลังรับแรงอัดของดินเหนียวอย่างง่าย โดยการกดตัวอย่างดินรูปทรงกระบอกทางแนวดิ่งโดยไม่มีความดันหรือการอัดช่วยทางด้านข้าง ค่าที่ได้อยู่ในรูปกำลังรับแรงอัดสูงสุด (q_u) ซึ่งค่ากำลังรับแรงเฉือนแบบอันเดรน (S_u) มีค่า $q_u / 2$ ค่า q_u สามารถใช้จำแนกชนิดของดินได้ตามตารางที่ 2.3

2.15.3 คุณสมบัติทางการยุบอัดตัว (COMPRESSIBILITY PROPERTIES) เป็นการหาคุณสมบัติการยุบอัดตัวของดินที่ได้จากการทดสอบ CONSOLIDATION ค่าที่ใช้พิจารณาการยุบอัดตัวของดินมีดังนี้

2.15.3.1 ดัชนีการอัดแน่น (COMPRESSION INDEX), C_c เป็นตัวที่นิยมใช้ในการชี้บอกความสามารถในการยุบตัวของดิน ซึ่งแสดงในเทอมความชันของกราฟ e - log $\bar{\sigma}$ ในช่วงที่กำหนดความดันให้

$$C_c = (\Delta e) / (\Delta \log \bar{\sigma})$$

SKEMPTON (1944) ได้ประมาณค่าของดัชนีการอัดจากคุณสมบัติของดินดังนี้

สำหรับดินคงสภาพ (UNDISTURBED CLAY)

$C_c = 0.009 (LL - 10\%)$

สำหรับดินเปลี่ยนสภาพ (REMOLDED CLAY)

$C_c = 0.007 (LL - 10\%)$

เมื่อ LL = LIQUID LIMIT OF A SOIL

2.15.3.2 อัตราส่วนการอัดแน่น (COMPRESSION RATIO), CR คือความชันของกราฟ ϵ - log $\bar{\sigma}$ ในช่วงที่กำหนดความดันให้

$$CR = (\Delta\epsilon) / (\Delta \log \bar{\sigma})$$

ความสัมพันธ์ระหว่าง CR และ C_c

$$CR = (C_c) / (1 + e_o)$$

เมื่อ e_o = อัตราส่วนช่องว่างเริ่มแรกของดิน

2.15.3.3 ดัชนีการอัดแน่นซ้ำ (RECOMPRESSION INDEX), C_r เป็นตัวชี้บอกความสามารถในการยุบอัดตัวของดินซึ่งแสดงในเทอมความชันของกราฟ e - log $\bar{\sigma}$ ในช่วงที่กำหนดความดันให้

$$C_r = (\Delta e) / (\Delta \log \bar{\sigma})$$

2.15.3.4 อัตราส่วนการอัดแน่นซ้ำ (RECOMPRESSION RATIO), RR คือความชันของกราฟ ϵ - log $\bar{\sigma}$ ในช่วงที่กำหนดความดันให้

$$RR = (\Delta\epsilon) / (\Delta \log \bar{\sigma})$$

ความสัมพันธ์ระหว่าง RR และ C_r

$$RR = (C_r) / (1 + e_o)$$

วิธีหาค่า C_c, C_r, CR และ RR แสดงดังรูปที่ 2.2

2.15.3.5 หน่วยแรงสูงสุดในอดีต (MAXIMUM PAST PRESSURE OR PRECONSOLIDATION PRESSURE), $\bar{\sigma}_{vm}$ หรือ $\bar{P}_c$ คือหน่วยแรงมากที่สุดที่ดินนั้นถูกกระทำมาก่อนในอดีต สำหรับดินที่มีการทับถมกันตามปกติเรียกว่า "NORMALLY CONSOLIDATED DEPOSIT" หน่วยแรงสูงสุดจะมีค่าเท่ากับหน่วยแรงประสิทธิผลในขณะนั้น ($\bar{\sigma}_{vo}$) ถ้าหน่วยแรงสูงสุดในอดีตมีค่ามากกว่าหน่วยแรงประสิทธิผลในขณะนั้น ดินนั้นเรียกว่า "OVERCONSOLIDATED SOIL" หน่วยแรงสูงสุดในอดีตจะสะท้อนให้ทราบถึงหน่วยแรงมากที่สุดที่เคยกระทำต่อมวลดินในอดีต วิธีการหาค่าPRECONSOLIDATION PRESSURE แสดงดังรูปที่ 2.3

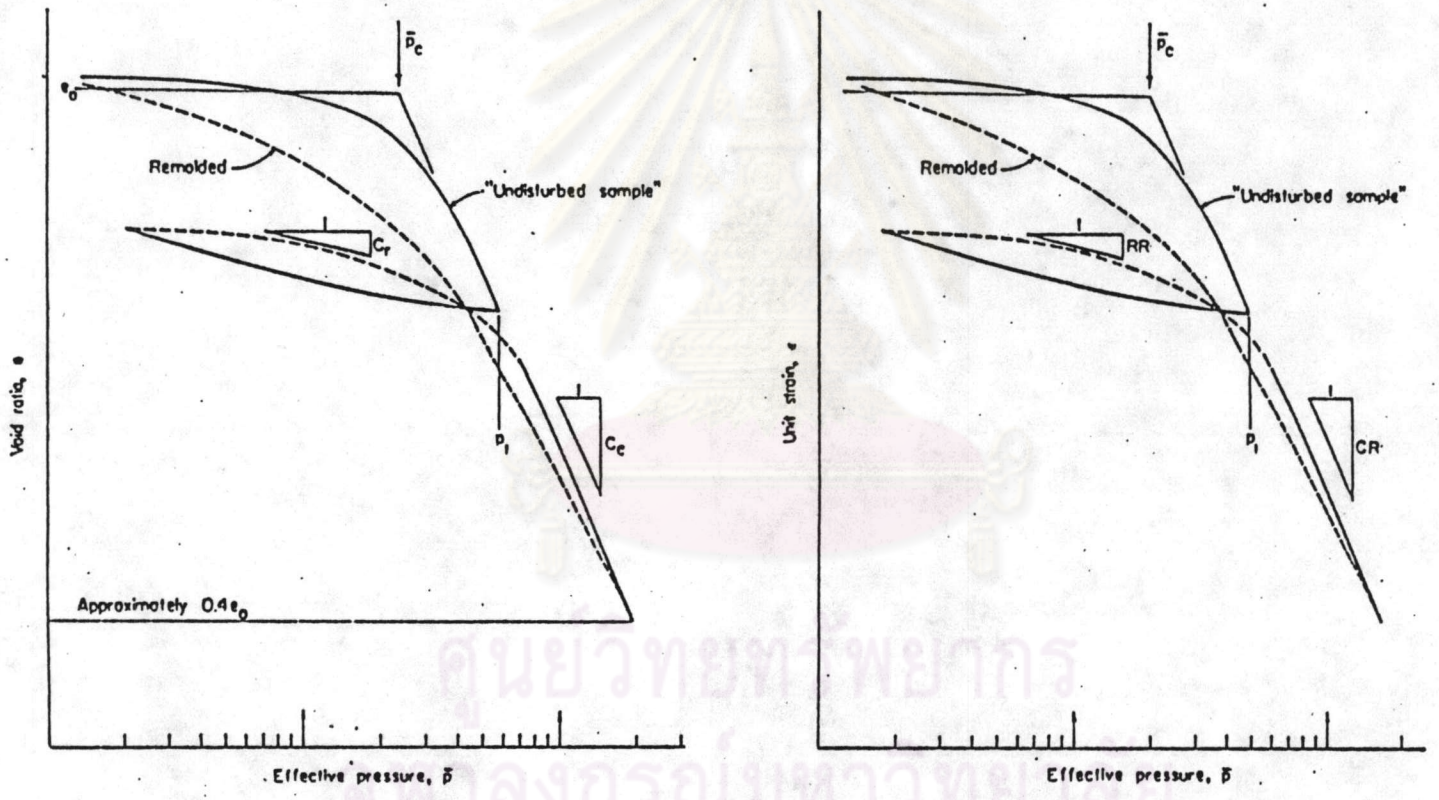
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ชนิดของดิน(Type of soil)	ความถ่วงจำเพาะ , Gs
ทราย(Sand)	2.65-2.67
ทรายแป้ง(Silty sand)	2.67-2.70
ดินเหนียวอนินทรีย์(Inorganic clay)	2.70-2.80
ดินปนไมก้าหรือเหล็ก(Soils with micas or iron)	2.75-3.00
ดินอินทรีย์(Organic soils)	ค่าแปรผันอาจต่ำกว่า 2.00

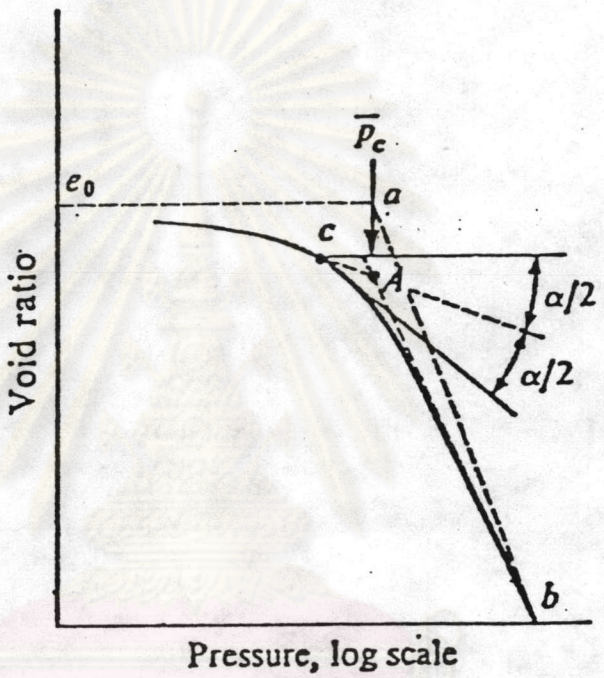
ตารางที่ 2.2 แสดงค่าความถ่วงจำเพาะของดิน

ความอ่อนแข็ง(Consistency)	วิธีจำแนกภาคสนาม(Field Identification)	กำลังรับแรงอัดสูงสุดแกนเดียว (Unconfined Compressive Strength,Qu (ksc))
อ่อนมาก(Very soft)	กดลงอย่างง่ายด้วยกำปั้นมีการยุบตัวประมาณ 2-3 นิ้ว (Easily penetrated several inches by fist)	น้อยกว่า 0.25
อ่อน(Soft)	กดลงอย่างง่ายด้วยนิ้วหัวแม่มือมีการยุบตัวประมาณ 2-3 นิ้ว (Easily penetrated several inches by thumb)	0.25-0.50
ปานกลาง(Medium)	กดลงด้วยนิ้วหัวแม่มือได้ยากใช้แรงมากมีการยุบตัวประมาณ 2-3 นิ้ว (Can be penetrated several inches by thumb with moderate effort)	0.50-1.00
แข็ง(Stiff)	กดลงด้วยนิ้วหัวแม่มือได้ยากใช้แรงมาก(Readily indented by thumb but penetrated only with great effort)	1.00-2.00
แข็งมาก(Very stiff)	เอาเล็บหัวแม่มือจิกได้(Readily indented by thumbnail)	2.00-4.00
แข็งมากมาก(Hard)	เอาเล็บหัวแม่มือจิกได้แต่ต้องใช้แรงมาก(Indented with difficulty by thumbnail)	มากกว่า 4.00

ตารางที่ 2.3 ความอ่อนแข็งของดินเหนียวในเทอมของกำลังรับแรงอัดสูงสุดแกนเดียว



รูปที่ 2.2 แสดงการหาค่าCc,Crจากกราฟ e-log $\bar{P}$ และค่าCR,RRจากกราฟ ε-log $\bar{P}$



รูปที่ 2.3 แสดงการหาค่าพรีคอนโซลิเดชั่นเพรสเชอร์โดยวิธีคลาสซาแกรนเด้(CASAGRANDE)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย