# บทที่ 2

## สถิติที่ใช้ในการวิจัย

ตัวแบบถดถอยเชิงเส้นเชิงเดียวโดยทั่วไปและนำมาใช้ในงานวิจัยนี้ มีรูปแบบดังนี้

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \qquad i = 1,2,\dots,n$$

โดยที่ $Y_i$ คือ ตัวแปรตาม

$\beta_0, \beta_1$ คือ พารามิเตอร์ของการถดถอย

$X_i$ คือ ตัวแปรอิสระที่กำหนดเป็นค่าคงที่

$\varepsilon_i$ คือ ความคลาดเคลื่อนสุ่ม ซึ่งเป็นอิสระต่อกัน และมีการแจกแจง $N(0,\sigma^2)$

$n$ คือ ขนาดตัวอย่าง

หรือเขียนได้ในรูปเมทริกซ์

$$\underline{y} = \underline{X}\,\underline{\beta} + \underline{\varepsilon} \tag{2.1}$$

โดยที่

$$\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{pmatrix} \quad \underline{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix} \quad \underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_n \end{pmatrix}$$

งานวิจัยนี้ พิจารณาเป็น 2 ส่วนตามการแจกแจงของตัวแปรตาม ดังนี้

## 2.1 ตัวแปรตามมีการแจกแจงเป็นแบบปกติ : $Y_i \sim N(\underline{\beta}'\underline{X}_i, \sigma^2)$, $\underline{X}_i' = (1, X_i)$, $\underline{\beta}' = (\beta_0, \beta_1)$

### 2.1.1 ตัวประมาณความควรจะเป็นสูงสุด

การประมาณพารามิเตอร์ของการถดถอยด้วยวิธีความควรจะเป็นสูงสุด มีหลักการคือ ตัวประมาณสัมประสิทธิ์การถดถอย $\left(\hat{\underline{\beta}} = (\hat{\beta}_0, \hat{\beta}_1)'\right)$ ที่ทำให้ฟังก์ชันความควรจะเป็นมีค่ามากที่สุด โดยการหาอนุพันธ์ของฟังก์ชันล็อกของฟังก์ชันความควรจะเป็นเทียบกับพารามิเตอร์ $\underline{\beta}$ แล้วให้ผลลัพธ์มีค่าเท่ากับศูนย์ โดยสมมติว่าทราบค่า $\sigma^2$ ดังนี้

ให้ $L(\beta_0, \beta_1 \mid Y_i)$ = ฟังก์ชันความควรจะเป็นของ $\beta_0, \beta_1$ เมื่อกำหนด $\underline{Y}$

= ฟังก์ชันความหนาแน่นร่วมของ $f(\underline{Y})$ ของ $\underline{Y}$

$$L(\beta_0, \beta_1 \mid Y_i) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(Y_i - \beta_0 - \beta_1 X_i)^2\right\}$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2\right\}$$

จะได้ว่า

$$\ln L(\beta_0, \beta_1 \mid Y_i) = -\left(\frac{n}{2}\right)\ln 2\pi - \left(\frac{n}{2}\right)\ln\sigma^2 - \left(\frac{1}{2\sigma^2}\right)\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2$$

$$\frac{\partial \ln L}{\partial \beta_0} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - X_i\beta_1) = 0 \qquad (1)$$

$$\frac{\partial \ln L}{\partial \beta_1} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - X_i\beta_1)X_i = 0 \qquad (2)$$

ให้ $\hat{\beta}_0$ และ $\hat{\beta}_1$ แทนตัวประมาณความควรจะเป็นสูงสุดของ $\beta_0$ และ $\beta_1$ ตามลำดับ

จากสมการที่ (1) ได้

$$\hat{\beta}_0 n + \hat{\beta}_1\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i \qquad (3)$$

จากสมการที่ (2) ได้

$$\hat{\beta}_0\sum_{i=1}^{n} X_i + \hat{\beta}_1\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} Y_i X_i \qquad (4)$$

จากสมการที่ (3) และ (4) หาค่า $\underline{\hat{\beta}} = \left(\hat{\beta}_0, \hat{\beta}_1\right)'$ ได้ดังนี้

ให้ $$\underline{A} = \begin{bmatrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_i^2 \end{bmatrix}$$

และ $$\underline{B} = \begin{bmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{bmatrix}$$

จะได้ $$\underline{A}^{-1} = \frac{1}{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \begin{bmatrix} \sum_{i=1}^{n} X_i^2 & -\sum_{i=1}^{n} X_i \\ -\sum_{i=1}^{n} X_i & n \end{bmatrix}$$

ดังนั้น $$\underline{\hat{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \underline{A}^{-1} \underline{B}$$

$$= \frac{1}{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \begin{bmatrix} \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} X_i Y_i \sum_{i=1}^{n} X_i \\ n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i \end{bmatrix}$$

จะได้ว่า

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} Y_i X_i - \sum_{i=1}^{n} Y_i \left(\frac{\sum_{n=1}^{n} X_i}{n}\right)}{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}} = \frac{\sum_{i=1}^{n} Y_i\left(X_i - \overline{X}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$

และ $$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$$

### 2.1.2. ตัวประมาณเบส์

จากตัวแบบ (2.1) โดยที่ สมมติให้ $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_n$ เป็นอิสระต่อกัน และต่างก็มีการแจกแจงปกติ $N(0, \sigma^2)$ โดยสมมติว่าทราบค่า $\sigma^2$ เพราะฉะนั้น จะได้ว่า $Y_1, Y_2, Y_3, \ldots, Y_n$ เป็นอิสระต่อกัน หรือ $\underline{Y}$ มีการแจกแจงปกติร่วม $n$ ตัวแปร

$$\underline{Y} \sim N(\underline{X}\underline{\beta}, \sigma^2 \underline{I})$$

สมมติ $\underline{\beta}$ เป็นเวกเตอร์สุ่มมีการแจกแจงก่อนสำหรับ $\underline{\beta}$ เมื่อทราบค่า $\sigma^2$ สมมติเป็นการแจกแจงปกติร่วม (joint normal distribution)

$$\underline{\beta} \sim N_2(\underline{\mu}, \underline{V})$$

โดยที่ $\underline{\mu}$ คือ เวกเตอร์ค่าเฉลี่ย ของค่าพารามิเตอร์ $\underline{\beta} = (\beta_0, \beta_1)'$

$$\underline{\mu} = E(\underline{\beta}) = \begin{bmatrix} \mu_0 \\ \mu_1 \end{bmatrix}, \mu_i = E(\beta_i), i = 0,1$$

และ $\underline{V}$ คือ เมทริกซ์ความแปรปรวนร่วม (covariance matrix) ของ $\beta_0$ และ $\beta_1$

$$\underline{V} = \begin{bmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{bmatrix}, \quad \sigma_{ij} = Cov(\beta_i, \beta_j) \quad ,i,j = 0,1$$

จะได้ฟังก์ชันความหนาแน่นก่อน $h(\underline{\beta})$ ของ $\underline{\beta}$ ดังนี้

$$h(\underline{\beta}) = \frac{1}{(2\pi)|\underline{V}|^{1/2}} \exp\left[-\frac{1}{2}(\underline{\beta} - \underline{\mu})' \underline{V}^{-1} (\underline{\beta} - \underline{\mu})\right],$$
$$-\infty < \beta_i < \infty, \quad i = 0,1$$

จากตัวแบบการถดถอย (2.1) สามารถหาตัวประมาณของ $\underline{\beta}$ ด้วยวิธีกำลังสองน้อยสุด (least squares method) ได้ตัวประมาณกำลังสองน้อยสุด $\underline{\tilde{\beta}}$ ของ $\underline{\beta}$ ดังนี้

$$\underline{\tilde{\beta}} = (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{Y} \qquad (2.2)$$

และมีคุณสมบัติว่า $\underline{\tilde{\beta}}$ มีการแจกแจงปกติร่วม 2 ตัวแปร

$$\underline{\tilde{\beta}} \sim N_2\left(\underline{\beta}, \underline{W}\right)$$

โดยที่ $\underline{\beta}$ คือ เวกเตอร์ค่าเฉลี่ย $\underline{\beta} = E\left(\underline{\tilde{\beta}}\right)$

$\underline{W}$ คือ เมทริกซ์ความแปรปรวนร่วม $\underline{W} = Var\left(\underline{\tilde{\beta}}\right) = \sigma^2\left(X^{'}X\right)^{-1}$

เนื่องจาก $\underline{\beta}$ เป็นเวกเตอร์สุ่ม มีฟังก์ชันความหนาแน่นก่อน $h(\underline{\beta})$ ดังนั้น การแจกแจงข้างต้นของ $\underline{\tilde{\beta}}$ เป็นการแจกแจงอย่างมีเงื่อนไข เมื่อกำหนด ค่า $\underline{\beta}$ และเขียนฟังก์ชันความหนาแน่นร่วมแบบมีเงื่อนไข $f\left(\underline{\tilde{\beta}} \mid \underline{\beta}\right)$ ของ $\underline{\tilde{\beta}}$ เมื่อกำหนด $\underline{\beta}$ ได้เป็น

$$f\left(\underline{\tilde{\beta}} \mid \underline{\beta}\right) = \frac{1}{(2\pi)\left|\underline{W}\right|^{1/2}} \exp\left[-\frac{1}{2}\left(\underline{\tilde{\beta}} - \underline{\beta}\right)^{'} \underline{W}^{-1}\left(\underline{\tilde{\beta}} - \underline{\beta}\right)\right],$$

$$-\infty < \beta_i < \infty \quad ,i = 0,1$$

เพราะฉะนั้น ฟังก์ชันความหนาแน่นขอบ (marginal density function) $g\left(\underline{\tilde{\beta}}\right)$ ของ $\underline{\tilde{\beta}}$ หาได้ดังนี้

$$g\left(\underline{\tilde{\beta}}\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\left(\underline{\tilde{\beta}} \mid \underline{\beta}\right) h\left(\underline{\beta}\right) d\beta_0 \, d\beta_1$$

ซึ่งจะได้ผลลัพธ์เป็นฟังก์ชันความหนาแน่นของการแจกแจงปกติร่วม 2 ตัวแปร $N_2 \sim \left(\underline{\mu}, \underline{V} + \underline{W}\right)$[1]

จากฟังก์ชันความหนาแน่นร่วม $f\left(\underline{\tilde{\beta}} \mid \underline{\beta}\right) h\left(\underline{\beta}\right)$ และ ฟังก์ชันความหนาแน่นขอบ $g\left(\underline{\tilde{\beta}}\right)$ หาฟังก์ชันความหนาแน่นภายหลัง (posterior density function) $k\left(\underline{\beta} \mid \underline{\tilde{\beta}}\right)$ ของ $\underline{\beta}$ เมื่อทราบ $\underline{\tilde{\beta}}$ ได้ดังนี้

$$k\left(\underline{\beta} \mid \underline{\tilde{\beta}}\right) = \frac{f\left(\underline{\tilde{\beta}} \mid \underline{\beta}\right) h\left(\underline{\beta}\right)}{g\left(\underline{\tilde{\beta}}\right)}$$

ได้ผลลัพธ์ เป็นฟังก์ชันความหนาแน่นของการแจกแจงปกติร่วม 2 ตัวแปร $N_2 \sim \left(\underline{\mu}^{*}, \underline{V}^{*}\right)$[1] โดยมี

$$Var\left(\underline{\beta} \mid \underline{\tilde{\beta}}\right) = \underline{V}^{*} = \left(\underline{V}^{-1} + \underline{W}^{-1}\right)^{-1} \quad , \underline{W}^{-1} = \frac{1}{\sigma^2}\left(\underline{X}^{'}\underline{X}\right)$$



---

1 Applied Statistical Decision. Raiffa and Schlaifer : 336-338

$$E\left(\underline{\beta} \mid \underline{\widetilde{\beta}}\right) = \underline{\mu}^{*} = \underline{V}^{*}\left(\underline{V}^{-1}\underline{\mu} + \underline{W}\,\underline{\widetilde{\beta}}\right)$$

เพราะฉะนั้น เมื่อกำหนดฟังก์ชันความเสียหาย ( loss function ) ในรูปแบบกำลังสองในการประมาณค่าพารามิเตอร์ จะได้ตัวประมาณเบส์ $\underline{\beta}^{*}$ ของ $\underline{\beta}$ คือ ค่าเฉลี่ย $\underline{\mu}^{*}$ ของการแจกแจงภายหลัง $N_2\left(\underline{\mu}^{*}, \underline{V}^{*}\right)$

$$\underline{\beta}^{*} = \begin{pmatrix} \beta_0^{*} \\ \beta_1^{*} \end{pmatrix} = \underline{\mu}^{*} = \underline{V}^{*}\left(\underline{V}^{-1}\underline{\mu} + \underline{W}\,\underline{\widetilde{\beta}}\right)$$

**2.2. ตัวแปรตามมีการแจกแจงล็อกนอร์มัล :** $Y_i \sim LN\left(\underline{\beta}'\underline{X}_i, \sigma^2\right)$, $\underline{X}_i' = (1, X_i)$, $\underline{\beta}' = (\beta_0, \beta_1)$

กรณีที่ $Y_i$ มีการแจกแจงล็อกนอร์มัล โดยมีค่าเฉลี่ยเท่ากับ $\beta_0 + \beta_1 X_i$ และความแปรปรวนเท่ากับ $\sigma^2$ ซึ่งสมมติว่าทราบค่า $\sigma^2$ โดยจะทำการแปลง $Y_i$ ให้มีการแจกแจงปกติด้วยสูตรการแปลง $\ln Y_i$ ก่อนใช้วิธีความควรจะเป็นสูงสุด และวิธีเบส์ ตามหัวข้อ 2.1 กับค่า $\ln Y_i$