บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในการเก็บรวบรวมข้อมูล (data) เพื่อใช้ในการวิเคราะห์ตามวัตถุประสงค์ของการศึกษาใด ๆ นั้นสามารถทำได้สองลักษณะ คือ การเก็บรวบรวมข้อมูลจากการสังเกต และการเก็บรวบรวมข้อมูลจากการทดลอง (Experimental Study) ในการเก็บรวบรวมข้อมูลจากการสังเกตนั้นเราจะสามารถระบุประชากรเป้าหมาย (Identification of Target Population) และสามารถที่จะอนุมาน (Inference) กลับสู่ประชากรได้กว้าง แต่เราจะไม่สามารถควบคุมตัวแปรรบกวน (Nuisance Variable) ได้ เหมือนกับข้อมูลที่เก็บมาจากการวางแผนการทดลอง และการวิเคราะห์ข้อมูลที่เก็บจากการสังเกตสามารถวิเคราะห์ได้เฉพาะความสัมพันธ์ของข้อมูล แต่ข้อมูลที่เก็บจากการวางแผนการทดลองสามารถวิเคราะห์เหตุ และผลของข้อมูลได้

ดังนั้นการวางแผนการทดลอง (Experimental Design) จึงเป็นที่นิยมใช้กันมากในหลายสาขา เช่น ทางด้านการแพทย์ การเกษตร จิตวิทยา การศึกษา วิศวกรรม ฯลฯ โดยวัตถุประสงค์ของการทดลอง (Experiment) จะเป็นการค้นคว้าเพื่อที่จะได้ความรู้ใหม่เกิดขึ้น หรือเพื่อยืนยันความรู้เดิมที่มีอยู่แล้วว่ามีความถูกต้อง ผลจากการทดลองจะทำให้เกิดประโยชน์ต่อการตัดสินใจได้อย่างถูกต้อง การออกแบบการทดลองให้เหมาะสมจึงจำเป็นอย่างยิ่งในการวางแผนทดลองเพื่อที่จะได้ข้อสรุปที่ถูกต้อง แผนแบบการทดลองมีหลายแผนแบบขึ้นอยู่กับ ลักษณะงานที่ต้องการทำการทดลอง และลักษณะของหน่วยทดลอง (Experimental Unit) แผนการทดลองหนึ่งที่นิยมใช้กันมาก คือ แผนการทดลองแบบจัตุรัสละติน (Latin Square Design : LSD) เป็นแผนแบบการทดลองที่ต้องการทดสอบความแตกต่างระหว่างอิทธิพลของปัจจัยทดลอง(Treatment Factor) โดยจะใช้แผนแบบการทดลองนี้เมื่อลักษณะของหน่วยทดลองไม่มีความสม่ำเสมอหรือหน่วยทดลองไม่คล้ายคลึงกัน 2 ลักษณะ โดยจะทำการจัดหน่วยทดลองที่คล้ายคลึงกันอยู่ในกลุ่มเดียวกันเรียกว่าบล็อก (Block) ซึ่งแต่ละระดับ (Level) ในบล็อกจะได้รับปัจจัยทดลองทุกระดับ ดังนั้นจะได้จำนวนระดับของวิธีการทดลองเท่ากับระดับของปัจจัยแบ่งบล็อก ทำให้ความแปรปรวนลดลง และจะได้ว่าผลการทดสอบนั้นแตกต่างกันจากอิทธิของปัจจัยทดลองอย่างแท้จริง

เมื่อทำการทดลองแล้วเราไม่แน่ใจว่าผลการทดลองนั้นถูกต้องจริงหรือไม่ก็อาจทำการเก็บรวบรวมข้อมูลเพิ่มเติมโดยทำการทดลองซ้ำ แต่บางครั้งหน่วยทดลองมีจำกัด หรือถ้าเราเปลี่ยนหน่วยทดลองทำให้ต้องเสียค่าใช้จ่ายในการทดลองสูงขึ้น ดังนั้นทางเลือกหนึ่งที่เรานิยมทำ คือ ทำการวัดข้อมูลจากหน่วยทดลองเดิมซ้ำ (Repeated Measure) โดยใช้แผนแบบการทดลองแบบเดิม

โดยการเก็บข้อมูลลักษณะนี้เป็นการเก็บข้อมูลระยะยาว (Longitudinal Data) ซึ่ง (Diggle J.P : 1995) ได้รวบรวมการวิเคราะห์ข้อมูลสำหรับข้อมูลระยะยาว

การศึกษาข้อมูลระยะยาวนั้นจะแตกต่างจากการศึกษาข้อมูลตัดตามขวาง(Cross-sectional Data) คือ ข้อมูลตัดตามขวางทำการวัดเพียงครั้งเดียว และข้อมูลระยะยาวแตกต่างจากข้อมูลอนุกรมเวลา (Time Series Data) โดยข้อมูลอนุกรมเวลาจะเก็บมาศึกษาเพียงหน่วยทดลองเดียวหลายระยะเวลา และเก็บในระยะเวลายาวกว่าข้อมูลระยะยาว ซึ่งการเก็บข้อมูลระยะยาวนี้มีการใช้อย่างแพร่หลายในการทดลองทางการแพทย์ และ การเกษตร

เมื่อเราทำการทดลอง และเก็บรวบรวมข้อมูลเสร็จแล้วขั้นตอนต่อไป คือ การวิเคราะห์ข้อมูล ซึ่งการวิเคราะห์ข้อมูลตามแผนแบบการทดลองจัตุรัสละตินเมื่อข้อมูลเป็นข้อมูลระยาวนั้นต้องใช้การวิเคราะห์ที่แตกต่างจากเดิม เพราะข้อมูลจากหน่วยทดลองเดิมที่เก็บมาอาจมีความสัมพันธ์กัน เรียกสถานการณ์นี้ว่า อัตตสหสัมพันธ์ (Autocorrelation) ซึ่งรูปแบบที่พบโดยทั่วไปจะเป็นอัตตสหสัมพันธ์อันดับที่หนึ่ง (First Order Autoregressive : AR(1))

ในการวิเคราะห์ข้อมูลที่ได้จากการทดลองนั้นมีข้อสมมติ (Assumption) ว่าความคลาดเคลื่อน (Error) มีการแจกแจงแบบปกติ (Normal Distribution) ที่ค่าเฉลี่ย (Mean) 0 และความแปรปรวน (Variance) เท่ากับ σ^2 และวิธีการประมาณค่าที่นิยมใช้กันโดยทั่วไปคือ วิธีกำลังสองต่ำสุดแบบสามัญ (Ordinary Least Square Method) เนื่องจากเป็นวิธีที่ทำให้ตัวประมาณมีคุณสมบัติตัวประมาณที่ไม่เอนเอียงเชิงเส้นที่ดีที่สุด (Best Linear Unbiased Estimator : BLUE) เมื่อข้อมูลเป็นไปตามข้อตกลง แต่จากข้อมูลที่เก็บตามระยะเวลาจะเห็นว่าข้อมูลจะไม่เป็นไปตามข้อตกลงเบื้องต้นกล่าวคือ ความคลาดเคลื่อนที่เกิดขึ้นของข้อมูลจะไม่เป็นอิสระกัน ทำให้ตัวประมาณด้วยวิธีกำลังสองต่ำสุดแบบสามัญที่ได้ไม่เป็นตัวประมาณเชิงเส้นที่ดีที่สุด แม้ว่าจะเป็นตัวประมาณไม่เอนเอียง

ดังนั้นผู้วิจัยจึงสนใจที่จะทำการศึกษาการประมาณค่าพารามิเตอร์ในแผนแบบการทดลองจัตุรัสละตินเมื่อเก็บข้อมูลระยะยาว และรูปแบบความสัมพันธ์เป็นอัตตสัมพันธ์อันดับที่หนึ่ง โดยใช้วิธีการประมาณ 2 วิธี คือ

1. การประมาณด้วยวิธีความควรจะเป็นสูงสุด (Maximum Likelihood Estimation)
2. การประมาณด้วยวิธีสองขั้น (Two-Stage Estimation)

โดยการประมาณค่าด้วยวิธีความควรจะเป็นสูงสุดนั้นจะดูลักษณะการแจกแจงของข้อมูลค่าสังเกต แต่ในการประมาณค่าด้วยวิธีการประมาณแบบสองขั้นนั้น จะพิจารณาจากความคลาดเคลื่อนที่เกิดขึ้นของข้อมูลค่าสังเกต

1.2 วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในแผนแบบการทดลองจัตุรัสละตินที่มีอิทธิพลปัจจัยคงที่ เมื่อข้อมูลที่นำมาวิเคราะห์เป็นข้อมูลระยะยาว และความคลาดเคลื่อนมีอัตตสหสัมพันธ์อันดับที่หนึ่ง ทำการประมาณค่าพารามิเตอร์ 2 วิธี คือ

1. วิธีความควรจะเป็นสูงสุด (Maximum Likelihood Method)
2. วิธีการประมาณแบบสองขั้น (Two-Stage Method)

1.3 สมมติฐานของการวิจัย

การประมาณค่าพารามิเตอร์ในแผนแบบการทดลองจัตุรัสละตินที่มีอิทธิพลปัจจัยคงที่ เมื่อข้อมูลที่นำมาวิเคราะห์เป็นข้อมูลระยะยาว โดยความคลาดเคลื่อนมีอัตตสัมพันธ์อันดับที่หนึ่ง การประมาณค่าพารามิเตอร์ด้วยวิธีความควรจะเป็นสูงสุดเหมาะสมกับการประมาณพารามิเตอร์มากกว่าวิธีประมาณแบบสองขั้น

1.4 ข้อตกลงเบื้องต้น

1. ศึกษาภายใต้แผนแบบการทดลองจัตุรัสละตินซึ่งไม่อันตรกิริยา (Interaction) ระหว่างเทอมปัจจัยทดลองและปัจจัยแบ่งบล็อกแรก ปัจจัยทดลองและปัจจัยแบ่งบล็อกที่สอง ปัจจัยแบ่งบล็อกแรกและปัจจัยแบ่งบล็อกที่สอง ปัจจัยทดลองและปัจจัยแบ่งบล็อกแรกและปัจจัยแบ่งบล็อกที่สอง
2. ข้อมูลที่เก็บซ้ำมีระยะห่างเวลาเท่า ๆ กัน
3. ปัจจัยทั้งหมดเป็นแบบคงที่ (Fixed Factor)
4. ในการศึกษาครั้งนี้จะให้ตัวแบบที่มีรูปแบบสมการดังนี้

$$Y_{h(\ell)} = \mu_{h(\ell)} + e_{h(\ell)} \qquad (1.4.1)$$

หรือ

$$Y_{ijk(\ell)} = \mu + \tau_i + \alpha_j + \beta_k + \gamma_\ell + (\tau\gamma)_i + e_{ijk(\ell)} \qquad (1.4.2)$$

$$i = j = k = 1, 2, \cdots, p$$

$$\ell = 1, 2, \cdots, b$$

$Y_{ijk(\ell)}$ หมายถึง ค่าสังเกตปัจจัยทดลอง i ปัจจัยแบ่งบล็อกแรก j ปัจจัยแบ่งบล็อกที่สอง k เก็บซ้ำเวลาที่ ℓ

μ หมายถึง พารามิเตอร์ค่าเฉลี่ยรวม

τ_i หมายถึง พารามิเตอร์ผลกระทบปัจจัยทดลอง i

α_j หมายถึง พารามิเตอร์ผลกระทบปัจจัยแบ่งบล็อกแรกระดับที่ j

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

β_k หมายถึง พารามิเตอร์ผลกระทบปัจจัยแบ่งบล็อกสองระดับที่ k

γ_ℓ หมายถึง พารามิเตอร์ผลกระทบจากการวัดซ้ำเวลาที่ ℓ

$(\tau\gamma)_{i\ell}$ หมายถึง พารามิเตอร์ผลกระทบร่วมระหว่างปัจจัยทดลองที่ i และการวัดซ้ำระยะเวลาที่ ℓ

$e_{ijk(\ell)}$ หมายถึง ความคลาดเคลื่อนสุ่มของหน่วยทดลอง ปัจจัยทดลอง i ปัจจัยแบ่งบล็อกแรก j ปัจจัยแบ่งบล็อกที่สอง k และการวัดซ้ำเวลาที่ ℓ

$Y_{h(\ell)}$ หมายถึง ค่าสังเกตจากหน่วยทดลอง h ในการวัดซ้ำครั้งที่ ℓ โดยมีจำนวนหน่วยทดลองทั้งหมด $N = p^2$ หน่วยทดลอง

$\mu_{h(\ell)}$ หมายถึง พารามิเตอร์ผลกระทบคงที่ทั้งหมดในหน่วยทดลอง h ในการวัดซ้ำระยะเวลาที่ ℓ

$e_{h(\ell)}$ หมายถึง ความคลาดเคลื่อนจากหน่วยทดลอง h ในการวัดซ้ำระยะเวลาที่ ℓ

5. ความคลาดเคลื่อนสุ่มของหน่วยทดลองในการวัดซ้ำเวลาที่ ℓ $e_{h(\ell)}$ มีอัตตสหสัมพันธ์อันดับที่หนึ่ง AR(1) โดยมีรูปแบบความสัมพันธ์ดังสมการ

$$e_{h(\ell)} = \rho e_{h(\ell-1)} + \varepsilon_{h(\ell)}$$

ความคลาดเคลื่อน $\varepsilon_{h(\ell)}$, $\varepsilon_{h(\ell')}$ เป็นอิสระกัน (Independent) และมีการแจกแจงแบบปกติที่ค่าเฉลี่ย 0 และความแปรปรวน σ_ε^2

6. ความคลาดเคลื่อนสุ่มของหน่วยทดลอง $e_{h(\ell)}$, $e_{h'(\ell)}$ เมื่อ $h \neq h' = 1, 2, \cdots, N$ $\ell = 1, 2, \cdots, b$ เป็นอิสระกัน

7. กำหนด $\sum_{i=1}^{p} \tau_i = 0 \quad \sum_{j=1}^{p} \alpha_j = 0 \quad \sum_{k=1}^{p} \beta_k = 0$

$$\sum_{\ell=1}^{b} \gamma_\ell = 0 \qquad \sum_{i=1}^{p}\sum_{\ell=1}^{b} (\tau\beta)_{i\ell} = 0$$

8. ซึ่งจากสมการ (1.4.2) เราสามารถเขียนรูปสมการในรูปเมทริกซ์ได้ดังนี้

$$\underset{\sim h}{Y} = X_h \underset{\sim}{\beta} + \underset{\sim h}{e} \qquad (1.4.3)$$

เมื่อ $\underset{\sim h}{y} = \left(y_{h(1)}, y_{h(2)}, \cdots, y_{h(b)}\right)'$ เป็นเวกเตอร์ค่าสังเกตของตัวแปรสุ่มจากหน่วยทดลองในปัจจัยทดลอง i ปัจจัยแบ่งบล็อกแรก j ปัจจัยแบ่งบล็อกที่สอง k มีขนาด b x 1 ซึ่งเวกเตอร์มีทั้งหมดจำนวน $N = p^2$ เวกเตอร์

$\underset{\sim}{\beta}$ คือ เวกเตอร์พารามิเตอร์ขนาด r x 1 เมื่อ $r = (3p-3) + b + (p-1)(b-1)$

$\underset{\sim h}{e}$ คือ เวกเตอร์ความคลาดเคลื่อนสุ่มขนาด b x 1

Σ_h คือ เมทริกซ์ความแปรปรวนร่วมของ $e_{h(\ell)}$ และ $e_{h(\ell')}$ ซึ่งมีรูปแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง ขนาด b x b

หอสมุดกลาง ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

9. จากสมการ (1.4.3) เมื่อพิจารณาการแจกแจงของ $\underset{\sim}{Y}_h$ จะได้ว่าเป็นการแจงแจงแบบปกติหลายตัวแปร $\underset{\sim}{Y}_h \sim N_b\left(X_h \underset{\sim}{\beta},\ \Sigma_h\right)$ และอาจเขียน (1.4.2) ในรูปสมการ

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{e} \qquad (1.4.4)$$

เมื่อ $\underset{\sim}{Y} = \left(\underset{\sim 1}{y'}, \underset{\sim 2}{y'}, \cdots, \underset{\sim N}{y'}\right)$ เป็นเวกเตอร์ค่าสังเกตทั้งหมดขนาด $(p^2\ b)\,x1$

X คือ เมทริกซ์ของผลกระทบคงที่ขนาด $(p^2\ b)\ x\ r$

$\underset{\sim}{\beta}$ คือ เวกเตอร์พารามิเตอร์ขนาด $r\,x\,1$

$\underset{\sim}{e}$ คือ เวกเตอร์ความคลาดเคลื่อนสุ่มขนาด $(p^2\ b)\,x1$

จากสมการ (1.4.4) เมื่อพิจารณาเฉพาะความคลาดเคลื่อนของ $\underset{\sim}{Y}$ จะได้ว่า

$$\begin{aligned} Var\left(\underset{\sim}{y}\right) &= Var(X\underset{\sim}{\beta} + \underset{\sim}{e}) \\ &= Var(X\underset{\sim}{\beta}) + Var(\underset{\sim}{e}) \\ &= Var(\underset{\sim}{e}) = \Sigma \end{aligned}$$

เมื่อ $\Sigma = \begin{bmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_N \end{bmatrix}_{NbxNb}$

Σ_h คือ เมทริกซ์ความแปรปรวนร่วมของความคลาดเคลื่อนของหน่วยทดลอง $h(\ell)$ และ $h(\ell')$ ซึ่งมีรูปแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง AR(1)

เมื่อ $\Sigma_{h(\ell'\ell)} = \sigma_\varepsilon^2\, \rho^{|\ell'-\ell|} \sum_{L=0}^{b-1} \rho^{2L}$; $h = 1, 2, \cdots, N$; $b = \min(\ell, \ell')$

เช่น กรณีเก็บข้อมูลซ้ำ 3 ระยะเวลา

$$\Sigma_h = \begin{bmatrix} \sigma_\varepsilon^2 & \rho\sigma_\varepsilon^2 & \rho^2\sigma_\varepsilon^2 \\ \rho\sigma_\varepsilon^2 & (1+\rho^2)\sigma_\varepsilon^2 & \rho(1+\rho)\sigma_\varepsilon^2 \\ \rho^2\sigma_\varepsilon^2 & \rho(1+\rho)\sigma_\varepsilon^2 & (1+\rho^2+\rho^4)\sigma_\varepsilon^2 \end{bmatrix}_{3x3} = \sigma_\varepsilon^2\, V_h$$

1.5 ขอบเขตของการวิจัย

1. ศึกษาภายใต้ความคลาดเคลื่อนของการเก็บซ้ำแต่ละหน่วยทดลองมีอัตตสหสัมพันธ์อันดับที่หนึ่ง AR(1) โดยความคลาดเคลื่อนมีรูปแบบความสัมพันธ์ดังสมการนี้

$$e_{h(\ell)} = \rho e_{h(\ell-1)} + \varepsilon_{h(\ell)} \qquad (1.5.1)$$

ความคลาดเคลื่อน $\varepsilon_{h(\ell)}$ เป็นอิสระกัน (Independent) และมีการแจกแจงแบบปกติที่ค่าเฉลี่ย 0 และความแปรปรวน σ_ε^2

2. กำหนดระดับปัจจัยเป็น 3 4 5 ระดับ หรือ $i = j = k = 3,\ 4,\ 5$
3. กำหนดระยะเวลาการเก็บข้อมูลซ้ำเป็น 4 ระดับ คือ 3 4 6 9 ตามลำดับ
4. กำหนดค่าพารามิเตอร์ในตัวแบบถดถอยลำดับที่หนึ่งทิศทางเดียวกันเป็น 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 ทิศทางตรงข้ามเป็น –0.1 -0.2 -0.3 -0.4 -0.5 -0.6 -0.7 -0.8 -0.9
5. กำหนดค่าสัมประสิทธิ์ความผันแปร (C.V.) เป็น 10% 20% 30%
6. ศึกษาการประมาณค่าพารามิเตอร์ภายใต้ $\tau_i\ \alpha_j\ \beta_k\ \gamma_\ell\ (\tau\gamma)_{i\ell}$ เป็นแบบคงที่ (Fixed Effect) โดยใช้เทคนิคมอนติคาร์โล (Monte Carlo Simulation Method)

1.6 เกณฑ์ในการตัดสินใจ

พิจารณาวิธีการประมาณค่าทั้งสองวิธีโดยทำการหาค่ารากที่สองของความคลาดเคลื่อน-กำลังสองเฉลี่ย (Root Mean Sum Square Error) สำหรับเปรียบเทียบค่าประมาณสำหรับพารามิเตอร์ในตัวแบบอิทธิพลปัจจัยคงที่

$$RMSE = \frac{\sum_{j=1}^{500} \sqrt{\left(\underset{\sim}{y}_j^{*} - \underset{\sim}{\hat{y}}_j^{*}\right)' \left(\underset{\sim}{y}_j^{*} - \underset{\sim}{\hat{y}}_j^{*}\right) \Big/ M - r}}{500} \qquad \text{เมื่อ} \quad \underset{\sim}{y}^{*} = \hat{C}^{-1}\, \underset{\sim}{y}$$

เกณฑ์การพิจารณา คือ วิธีประมาณใดที่ให้ค่าต่ำสุดแสดงว่าวิธีการประมาณวิธีนั้นเหมาะสมกับการประมาณค่าพารามิเตอร์

1.7 คำจำกัดความที่ใช้ในการวิจัย

1. ข้อมูลระยะยาว (Longitudinal Data) หมายถึง ข้อมูลที่เก็บจากหน่วยทดลองเดิมมากกว่าหนึ่งครั้ง ในช่วงเวลาที่แตกต่างกันโดยทำการเก็บแต่ละครั้งหน่วยทดลองจะได้รับอิทธิพลจากปัจจัยทดลองเดิมและจากปัจจัยแบ่งบล็อกเดิม
2. อัตตสหสัมพันธ์ (Autocorrelation) หมายถึง เหตุการณ์ที่ความคลาดเคลื่อนสุ่มจากหน่วยทดลองเดิม ปัจจัยแบ่งบล็อกเดิมแต่ต่างช่วงระยะเวลา $e_{h(\ell)}$ และ $e_{h(\ell')}$ มีความสัมพันธ์กัน

1.8 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถวิเคราะห์ข้อมูลจากแผนการทดลองแบบจัตุรัสละตินที่มีอิทธิพลปัจจัยคงที่เมื่อทำการเก็บข้อมูลระยะยาว

2. เป็นแนวทางพัฒนาวิธีการประมาณค่าพารามิเตอร์ กรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบรูปแบบอื่น
3. เป็นแนวทางในการศึกษาการประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณอื่น
4. เป็นแนวทางในการประมาณค่าพารามิเตอร์เมื่อใช้แผนการทดลองแบบอื่น

1.9 สัญลักษณ์ในการวิจัย

เนื่องจากในการจัดหน่วยทดลองให้กับวิธีการและบล็อกนั้นสามารถจัดได้หลายวิธีขึ้นอยู่กับผู้ทำการทดลองว่าจะจัดอย่างไรดังนั้นการเขียนสัญลักษณ์ต่าง ๆ ก็อาจเกิดการสับสนได้ ดังนั้นเพื่อความเข้าใจลักษณะการจัดหน่วยทดลองและทฤษฎี ได้สรุปสัญลักษณ์ที่ใช้ดังนี้

i	คือ หน่วยทดลองได้รับวิธีการทดลองระดับที่ i
j	คือ หน่วยทดลองอยู่บล็อกแรกในระดับที่ j
k	คือ หน่วยทดลองอยู่บล็อกที่สองระดับที่ k
ℓ	คือ หน่วยทดลองนั้นมาจากการเก็บซ้ำระยะเวลา ℓ
h	คือ หน่วยทดลองที่ h มีจำนวน N หน่วยทดลอง
$N = p^2$	คือ จำนวนหน่วยทดลองทั้งหมดที่นำมาทดลอง
$\Sigma_h = \sigma_\varepsilon^2 V_h$	คือ เมทริกซ์ความแปรปรวนร่วมของหน่วยทดลอง
$\underset{\sim h}{e}$	คือ เวกเตอร์ความคลาดเคลื่อนของหน่วยทดลอง h
M	คือ จำนวนค่าสังเกตทั้งหมดมีจำนวน $Nxb = p^2xb$ ค่าสังเกต
r	คือ จำนวนพารามิเตอร์คงที่ ในการประมาณ
$\underset{\sim h}{y}$	คือ เวกเตอร์ของค่าสังเกต ของหน่วยทดลอง h
$\underset{\sim}{y}$	คือ เวกเตอร์ค่าสังเกตทั้งหมด จากการทดลอง
σ_ε^2	คือ ความแปรปรวนของความคลาดเคลื่อน
ρ	คือ สัมประสิทธิ์ในตัวแบบอัตตสหสัมพันธ์ของความคลาดเคลื่อน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย