บทที่ 2

ทฤษฎีที่ใช้ในการวิจัย

วิธีประมาณค่าสูญหายที่สนใจนำมาเปรียบเทียบในวิทยานิพนธ์เรื่องนี้มี 4 วิธี คือ

1. วิธีที่ใช้ค่าเฉลี่ย
2. วิธีวิเคราะห์ความถดถอยพหุเชิงเส้น
3. วิธีวิเคราะห์ความถดถอยพหุเชิงเส้นตัดแปลง
4. วิธีวิเคราะห์ส่วนประกอบหลัก

รายละเอียดเกี่ยวกับวิธีประมาณแต่ละวิธีเป็นดังนี้

2.1 วิธีที่ใช้ค่าเฉลี่ย

ประมาณค่าสูญหายของแต่ละตัวแปรโดยใช้ค่าเฉลี่ยจากตัวอย่างที่ข้อมูลไม่สูญหายของแต่ละตัวแปรนั้น ๆ

$$\bar{x}_i = \frac{\sum_{j=1}^{n-d_i} x_{ji}}{n - d_i}$$

โดยที่ $\bar{x}_i$ คือ ค่าเฉลี่ยจากตัวอย่างที่ข้อมูลไม่สูญหายของตัวแปรที่ i, $i = 1, 2, \ldots, p$

d_i คือ จำนวนค่าสูญหายของตัวแปรที่ i

n คือ จำนวนตัวอย่างทั้งหมดที่ใช้ในการเก็บรวบรวมข้อมูล

p คือ จำนวนตัวแปรทั้งหมดที่ใช้วิเคราะห์

x_{ij} คือ ค่าข้อมูลจากตัวอย่างที่ไม่สูญหายของตัวแปรที่ i ค่าสังเกตที่ j

2.2 วิธีวิเคราะห์ความถดถอยพหุเชิงเส้น

การวิเคราะห์ความถดถอยพหุเชิงเส้น เป็นวิธีวิเคราะห์เพื่อหาความสัมพันธ์ระหว่างตัวแปร 2 ตัวขึ้นไป โดยมีตัวแปรหนึ่งเป็นตัวแปรตาม (Y) และตัวแปรอีกตัวหนึ่งเป็นตัวแปรอิสระ (X_i) ส่วนการแสดงความสัมพันธ์ระหว่างตัวแปรทั้งสองนำไปใช้ในการประมาณหรือพยากรณ์ตัวแปร

ตามได้ สำหรับสมการแสดงความสัมพันธ์ระหว่างตัวแปร คือ

$$Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \dots\dots\dots\dots + \beta_p X_p + \varepsilon \qquad (2.2.1)$$

โดยที่

β_o เป็นค่าคงที่

β_i เป็นค่าสัมประสิทธิ์การถดถอยของ Y เทียบกับ X_i เมื่อ $X_1, X_2, \dots\dots, X_{i-1}, X_{i+1}, \dots\dots, X_p$ คงที่ นั่นคือ เมื่อ X_i เปลี่ยนแปลง 1 หน่วย จะมีผลทำให้ Y เปลี่ยนแปลง β_i หน่วย เมื่อตัวแปรอิสระอื่น ๆ ที่เหลือคงที่

ε เป็นค่าความแตกต่างระหว่างค่าประมาณที่ได้จากสมการเส้นถดถอยกับค่าจริง ซึ่งเป็นตัวแปรสุ่มที่ไม่ทราบค่า

จาก (2.2.1) สามารถเขียนให้อยู่ในรูปเมตริกซ์ได้ดังนี้

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \qquad (2.2.2)$$

โดยที่

$$\underset{\sim}{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{n \times 1}$$

$$X = \begin{pmatrix} 1 & X_{11} & X_{12} \cdots\cdots X_{1P} \\ 1 & X_{21} & X_{22} \cdots\cdots X_{2P} \\ 1 & & \\ \vdots & & \\ 1 & X_{n1} & X_{n2} \cdots\cdots X_{nP} \end{pmatrix}$$

ในการวิจัยครั้งนี้ตัวแปรที่มีค่าสูญหายจะเป็นตัวแปรตาม ส่วนตัวแปรอื่น ๆ ที่เหลือเป็นตัวแปรอิสระ เพื่อที่จะประมาณค่าสูญหายด้วยวิธีวิเคราะห์ความถดถอยพหุเชิงเส้น

2.3 วิธีวิเคราะห์ความถดถอยพหุเชิงเส้นดัดแปลง

คล้ายกับวิธีวิเคราะห์ความถดถอยพหุเชิงเส้น จะแตกต่างกันในส่วนของข้อมูลเริ่มต้นคือ ทุก ๆ ค่าสูญหายถูกแทนด้วยค่าเฉลี่ยของแต่ละตัวแปรก่อนแล้วจึงทำการประมาณค่าสูญหายแต่ละค่า ด้วยวิธีวิเคราะห์ความถดถอยพหุเชิงเส้น

2.4 วิธีวิเคราะห์ส่วนประกอบหลัก

การวิเคราะห์ส่วนประกอบหลัก เป็นการวิเคราะห์เพื่อหาความสัมพันธ์ระหว่างตัวแปรตาม (Y) และตัวแปรอิสระ (X) โดยที่ตัวแปรตามเป็นข้อมูลคุณภาพ สามารถเปรียบเทียบขนาดของสิ่งที่สนใจว่าสิ่งไหนมีขนาดที่ใหญ่กว่าแต่ไม่สามารถบอกถึงความแตกต่างเป็นปริมาณที่แน่นอนได้ ซึ่งมีรูปแบบสมการดังนี้คือ

ส่วนประกอบที่ 1 $Y_1 = \beta_{11} X_1 + \beta_{12} X_2 + \ldots\ldots\ldots + \beta_{1p} X_p$

ส่วนประกอบที่ 2 $Y_2 = \beta_{21} X_1 + \beta_{22} X_2 + \ldots\ldots\ldots + \beta_{2p} X_p$

$\vdots$

ส่วนประกอบที่ p $Y_p = \beta_{p1} X_1 + \beta_{p2} X_2 + \ldots\ldots\ldots \quad \beta_{pp} X_p$ (2.2.3)

โดยที่

β_{ji} เป็นค่าสัมประสิทธิ์จากการแยกตัวประกอบของ Y_i ต่อ X_j

จาก (2.2.3) หาค่า β_i ได้จาก

$$(R - \lambda_i I)\, \beta_i = 0 \qquad (2.2.4)$$

โดยที่

β_i เป็นค่าสัมประสิทธิ์ถ่วงน้ำหนักจากการแยกตัวประกอบหรือเรียกว่าไอเกนเวคเตอร์ (eigen vector) ; $i = 1, 2, \ldots, p$

λ_i เป็นค่าไอเกน (eigen value) ซึ่งคือค่าความแปรปรวนของส่วนประกอบหลักที่ i

R เป็นเมตริกซ์ของความสัมพันธ์

$$= \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots\cdots & \rho_{1p} \\ \rho_{21} & 1 & & & \\ \rho_{31} & & 1 & & \\ \vdots & & & & \\ \rho_{p1} & & & & 1 \end{bmatrix}$$

จากทฤษฎี จะได้ว่า

$$\sum_{j=1}^{p} \beta_{ij}^{2} = 1.00 \qquad (2.2.5)$$

$$\prod_{i=1}^{p} \lambda_{i} = |R| \qquad (2.2.6)$$

โดยที่ $|R|$ คือ determinant ของ R

$$\sum_{j=1}^{p} \lambda_{j} = \sum_{j=1}^{p} (\gamma_{jj}) \qquad (2.2.7)$$

การวิเคราะห์ส่วนประกอบหลัก โดยทั่ว ๆ ไปถือว่าส่วนประกอบที่น่าสนใจจะเป็นส่วนประกอบที่มีค่าไอเกน (eigen value) หรือค่าความแปรปรวนตั้งแต่ 1 ขึ้นไปเท่านั้น และสำหรับส่วนประกอบที่ 1 จะมีค่าไอเกนหรือค่าความแปรปรวนมากที่สุด สำหรับสัมประสิทธิ์ถ่วงน้ำหนักที่ได้จากแกนที่ 1 สามารถนำไปวิเคราะห์เปรียบเทียบขนาดของสิ่งที่สนใจได้ อีกทั้งสามารถบอกถึงความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระแต่ละตัวได้ด้วย ในที่นี้ จะประมาณค่าสูญหายแต่ละค่าโดยใช้สัมประสิทธิ์ถ่วงน้ำหนักที่ได้จากแกนที่ 1 คือ $q = (q_1, q_2, \ldots\ldots, q_p)$ แล้วจึงหาขนาดของตัวแปรตาม, a_r โดยที่ $a_r = \sum_{j=1}^{p} y_{jr} q_j$; $r = 1, 2 \ldots\ldots, n$ จากนั้นประมาณค่าสูญหายนั้นด้วย y_{ir} โดยที่ $y_{jr} = a_r q_j$

2.5 <u>การแจกแจงแบบปกติของตัวแปรพหุ</u> (Multivariate Normal Distribution)

$$X = (X_1, X_2, \ldots\ldots\ldots, X_n)'$$

โดยที่ X_i เป็นตัวแปรสุ่ม (Random Variable)

X_i จะมีการแจกแจงแบบปกติ เมื่อมี p.d.f. (probability density function) ดังนี้

$$f_X(x) = \frac{1}{(2\pi)^{\frac{p}{2}} (|\Sigma|)^{\frac{1}{2}}} \exp - \tfrac{1}{2} (X-\mu)' \Sigma^{-1}(X-\mu)$$

โดยที่ $-\infty < x_i < \infty$; $i = 1, 2, \ldots\ldots, p$

เขียนได้เป็น $X \sim N(\underset{\sim}{\mu}, \Sigma)$

เมื่อ $\underset{\sim}{\mu} = (\mu_1, \mu_2, \ldots\ldots, \mu_p)$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \cdots\cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \sigma_{2p} \\ \vdots & \vdots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \sigma_{pp} \end{bmatrix}$$

Σ เป็น p x p positive definite และสมมาตร (Symmetric) ซึ่งคือเมตริกซ์ของความแปรปรวนร่วม (Covariance Matrix)

สามารถสร้าง $X \sim N(0, \Sigma)$ ได้โดย

ให้ $Y \sim N(0, I_p)$

$$X = CY \qquad (2.5.1)$$

จาก (2.5.1)

$$\begin{aligned} \mathrm{Var}\ (X) &= \mathrm{Var}\ (CY) \\ &= C\ \mathrm{Var}\ (Y)\ C' \\ &= C\ I_p\ C' \\ &= C\ C' \end{aligned}$$

ดังนั้น $X \sim N\ (O,\ C\ C')$

แต่เราต้องการสร้างให้ $X \sim N(O,\ \Sigma)$ นั่นก็หมายความว่า Σ จะต้องมีค่าเท่ากับ $C\ C'$ จึงจะทำให้ $X = CY$ ดังนั้น จะต้องหาค่า C ซึ่งได้ว่า C เป็น lower trianqular Matrix

$$C = \begin{pmatrix} c_{11} & 0 & 0 \cdots 0 \\ c_{21} & c_{22} & \vdots \\ \vdots & & \\ c_{p1} & & c_{pp} \end{pmatrix}$$

โดยที่

$$c_{i1} = \sigma_{i1} / \sqrt{\sigma_{11}} \qquad 1 \leq i \leq p$$

$$c_{ii} = \sqrt{\sigma_{ii} - \sum_{k=1}^{i-1} c_{ik}^2} \qquad 1 < i \leq p$$

$$c_{ij} = \left[\sigma_{ij} - \sum_{k=1}^{j-1} c_{ik}\ c_{jk}\right] / c_{jj} \qquad 1 < j < i \leq p$$

$$c_{ij} = 0 \qquad i < j \leq p$$

ซึ่งจะได้

$$X_i = \sum_{j=1}^{i} c_{ij}\ Y_j\ ;\ i=1,\ 2,\ \ldots\ldots,\ p$$

เมื่อ $Y_1,\ Y_2,\ \ldots\ldots,\ Y_p \sim N(0,\ 1)$

ตามหลักเกณฑ์และขั้นตอนต่าง ๆ ที่กล่าวแล้วข้างต้นสามารถนำมาสร้างข้อมูลที่มีการแจกแจงแบบปกติของตัวแปรพหุได้

2.6 เกณฑ์ที่ใช้ในการเปรียบเทียบวิธีประมาณค่าสูญหาย

เกณฑ์ที่ใช้ในการเปรียบเทียบวิธีประมาณค่าสูญหาย คือ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และลำดับที่ของคะแนนรวมของวิธีทั้ง 4 ที่คิดจากจำนวนครั้งของวิธีทั้ง 4 ที่ได้ลำดับ 1 หรือ 2 หรือ 3 หรือ 4 จากการเรียงค่า MSE จากน้อยไปหามากถ่วงน้ำหนักด้วยคะแนน 4, 3, 2, 1 ตามลำดับ (TOTAL WEIGHTED SCORE หรือ TWS)

2.6.1 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย คำนวณจาก

$$MSE = \frac{\sum_{i=1}^{d} (X_i - \hat{X}_i)^2}{d}$$

โดยที่

X_i คือ ค่าข้อมูลจริงที่สูญหายตัวที่ i จากแต่ละตัวอย่าง

$\hat{X}_i$ คือ ค่าประมาณข้อมูลที่สูญหายตัวที่ i จากแต่ละตัวอย่าง

d คือ จำนวนข้อมูลสูญหายจากแต่ละตัวอย่าง

2.6.2 คะแนนรวมถ่วงน้ำหนัก จำแนกตามวิธีที่ใช้ในการประมาณค่า เป็นดังนี้

วิธี	จำนวนครั้งที่ได้				TWS
	อันดับ 1	อันดับ 2	อันดับ 3	อันดับ 4	
1	A	B	C	D	X
2	E	F	G	H	Y
3	I	J	K	L	Z
4	M	N	O	P	S

โดยที่

A, E, I, M คือ จำนวนครั้งที่ได้อันดับที่ 1 ใน 1000 ครั้งของวิธีที่ 1, 2, 3, 4 ตามลำดับ

B, F, J, N คือ จำนวนครั้งที่ได้อันดับที่ 2 ใน 1000 ครั้งของวิธีที่ 1, 2, 3, 4 ตามลำดับ

C, G, K, O คือ จำนวนครั้งที่ได้อันดับที่ 3 ใน 1000 ครั้งของวิธีที่ 1, 2, 3, 4 ตามลำดับ

D, H, L, P คือ จำนวนครั้งที่ได้อันดับที่ 4 ใน 1000 ครั้งของวิธีที่ 1, 2, 3, 4 ตามลำดับ

X, Y, Z, S คือ คะแนนรวมถ่วงน้ำหนักของวิธีที่ 1, 2, 3, 4 ตามลำดับ ซึ่งสามารถคำนวณได้จาก

$$X = (4xA)+(3xB)+(2xC)+(1xD)$$

$$Y = (4xE)+(3xF)+(2xG)+(1xH)$$

$$Z = (4xI)+(3xJ)+(2xK)+(1xL)$$

$$S = (4xM)+(3xN)+(2xO)+(1xP)$$

2.7 <u>การวิเคราะห์ความแปรปรวน</u> (Analysis of Variance)

การวิเคราะห์ความแปรปรวนนี้แบ่งออกเป็น 2 ประเภทใหญ่ ๆ คือ การวิเคราะห์ความแปรปรวนกรณีที่มีตัวแปรตามหนึ่งตัว (Univariate Analysis of Variance) และการวิเคราะห์ความแปรปรวนกรณีที่มีตัวแปรตามหลายตัว (Multivariate Analysis of Variance)

ในที่นี้จะทำการวิเคราะห์ความแปรปรวนกรณีที่มีตัวแปรตามหลายตัว เพื่อที่จะทำการทดสอบผลของปัจจัยทั้ง 3 ว่ามีผลต่อค่าความคลาดเคลื่อนของแต่ละวิธีหรือไม่ โดยที่ตัวแปรตามตัวที่ 1 2 3 และ 4 คือค่าความคลาดเคลื่อนเฉลี่ยของวิธีการประมาณค่าสูญหายวิธีที่ 1 2 3 และ 4 ตามลำดับ และจะทำการวิเคราะห์ความแปรปรวนกรณีที่มีตัวแปรตามหนึ่งตัว เพื่อที่จะทำการทดสอบว่าวิธีทั้ง 4 ให้ค่าความคลาดเคลื่อนเฉลี่ยแตกต่างกันหรือไม่ โดยที่ตัวแปรตามก็คือค่าความคลาดเคลื่อนเฉลี่ยนั่นเอง

GENERAL LINEAR MODEL

$$Y = X\beta + \varepsilon$$

โดยที่

Y เป็นเมตริกซ์ขนาด : nxp ของค่าสังเกต

X เป็นเมตริกซ์ขนาด nxq ของตัวแปรอิสระ

β เป็นเมตริกซ์ขนาด qxp ของสัมประสิทธิ์ของความสัมพันธ์

ε เป็นเมตริกซ์ขนาด nxp ของความคลาดเคลื่อน

สามารถเขียนได้เป็น

$$\begin{pmatrix} Y_{11} & Y_{12}\cdots Y_{1p} \\ Y_{21} & Y_{22}\cdots Y_{2p} \\ \vdots & \vdots \\ Y_{n1} & Y_{np} \end{pmatrix} = \begin{pmatrix} X_{11} & X_{12}\cdots & X_{1q} \\ X_{21} & X_{22} & X_{2q} \\ \vdots & & \\ X_{n1} & & X_{nq} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12}\cdots\beta_{1p} \\ \beta_{21} & \beta_{2p} \\ \vdots & \vdots \\ \beta_{q1} & \beta_{qp} \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12}\cdots & \varepsilon_{1p} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{2p} \\ \vdots & & \\ \varepsilon_{n1} & & \varepsilon_{np} \end{pmatrix}$$

รูปแบบที่ใช้ในการวิเคราะห์

$$Y_{\ell m} = \sum_{j=1}^{q} X_{\ell j}\,\beta_{jm} + \varepsilon_{\ell m} \;;\; \varepsilon_{\ell m} \sim N(O,\Sigma)$$

สมมุติฐานที่ใช้ในการทดสอบ คือ

$$H_o \; : \; C\beta = 0$$

$$H_1 \; : \; C\beta \neq 0$$

สำหรับวิธีการทดสอบ สามารถทำได้ 2 วิธี คือ

วิธีที่ 1 ใช้ UNION INTERSECTION TEST ของ ROY

ตัวสถิติที่ใช้ในการทดสอบคือ

$$Q_S = \frac{C_S}{1+C_S}$$

เมื่อ C_S เป็น Characteristic root ที่มากที่สุดของ $|H - \lambda E| = 0$

โดยที่

$$H = Y'(Q(Q'Q)^{-1}Q')Y \quad \text{โดยที่} \quad Q = X(X'X)^{-1}C'$$

$$E = Y'(I-X(X'X)^{-1}X')Y$$

เกณฑ์ที่ใช้ในการตัดสินใจ คือ จะปฏิเสธ H_o ณ ระดับนัยสำคัญ α ถ้า $Q_S \geq X_{\alpha,S,m,n^*}$

เมื่อ $X_{\alpha,\,S,\,m,\,n^*}$ เป็นค่าที่ได้จากตารางของ HECK

α คือ ความน่าจะเป็นที่จะเกิด Type I error

S คือ min (q, p)

m คือ $\frac{1}{2}$ ($|q-p|-1]$

n^* คือ $\frac{1}{2}$ $[n-q-p-1]$

วิธีที่ 2 ใช้ LIKELIHOOD RATIO PROCEDURE ของ WILK

ตัวสถิติที่ใช้ในการทดสอบ คือ

$$U = \frac{|E|}{|H+E|}$$

เมื่อ $H = Y'|Q(Q'Q)^{-1}Q'|\, Y$

$E = Y'|I-X(X'X)^{-1}X'|\, Y$

เกณฑ์ที่ใช้ในการตัดสินใจ คือ จะปฏิเสธ H_o ณ ระดับนัยสำคัญ α ถ้า $U_{cal} < U_{table}$ โดยที่ U_{cal} คือ ค่าสถิติที่ได้จากการคำนวณ

U_{table} คือ ค่าที่ได้จากเปิดตาราง U

แต่เพื่อหลีกเลี่ยงการใช้ตาราง U ซึ่งค่อนข้างหายากเราสามารถแปลงค่าสถิติ U นี้ให้เป็นค่าสถิติ F ได้ดังนี้คือ

$$F = \frac{n+1}{p+1} \cdot C$$

degree of freedom 2p + 2 และ 2n + 2

โดยที่ $C = \operatorname{tr} HE^{-1}$

n คือ จำนวนตัวอย่าง

p คือ จำนวนตัวแปร

ตารางแสดงแบบแผนการทดลองแบบ 3 ทาง สำหรับตัวแปรตาม 4 ตัวแปร

			ค่า MSE ของ			
			วิธีที่ 1	วิธีที่ 2	วิธีที่ 3	วิธีที่ 4
P = 3	n = 30	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 50	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 70	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 100	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 200	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
P = 5	n = 30	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 50	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 70	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 100	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 200	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
P = 7	n = 30	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 50	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 70	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 100	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				
	n = 200	ρ = 0.2				
		ρ = 0.5				
		ρ = 0.8				

ตัวแปรตาม 4 ตัว คือ วิธีที่ 1 วิธีที่ 2 วิธีที่ 3 วิธีที่ 4

ปัจจัยที่สนใจศึกษา คือ จำนวนตัวแปร p 3 ระดับ

ขนาดตัวอย่าง n 5 ระดับ

ความสัมพันธ์ ρ 3 ระดับ

การวิเคราะห์ว่าวิธีการประมาณค่าสูญหายทั้ง 4 วิธีให้ค่าความคลาดเคลื่อนเฉลี่ยแตกต่างกันหรือไม่นั้น ได้วางแผนการทดลองเป็นแบบ RBD (Randomized Complete Block Design) ซึ่งเป็นแผนการทดลองที่หน่วยทดลองมีความแตกต่างกัน 1 ทาง โดยมีทรีทเมนต์ (Treatment) 4 ทรีทเมนต์ คือ วิธีทั้ง 4 นั่นเอง Block ก็คือสถานการณ์ต่าง ๆ ตามค่า N, P, RHO มีทั้งหมด 106 สถานการณ์

<u>รูปแบบที่ใช้ในการวิเคราะห์</u>

$$Y_{ij} = \mu + \beta_i + T_j + \Sigma_{ij}$$

$i = 1, 2, \ldots, r$; r = จำนวน block

$j = 1, 2, \ldots, n$; n = จำนวนระดับของ factor

H_o : วิธีการประมาณค่าสูญหายทั้ง 4 วิธีให้ค่าความคลาดเคลื่อนเฉลี่ยไม่แตกต่างกัน

H_1 : วิธีการประมาณค่าสูญหายอย่างน้อย 1 วิธีแตกต่างไปจากวิธีอื่น ๆ

สถิติที่ใช้ทดสอบคือ F โดยกำหนดระดับนัยสำคัญ 0.05 ดังนั้นจะปฏิเสธสมมุติฐานถ้าค่า F จากการคำนวณมีค่ามากกว่า $F_{3,315\ (0.05)} = 2.60$

ตารางแสดงแผนการทดลอง RBD สำหรับตัวแปรตาม 1 ตัว

Block	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3	วิธีที่ 4
1				
2				
3				
4				
5				
.				
.				
.				
.				
.				
.				
.				
.				
.				
.				
106				