บทที่ ๑

บทนำ

สมมติว่าเรานำพืชชนิดหนึ่ง r พันธุ์ไปปลูกในที่ดิน c แปลง เพื่อเปรียบเทียบดูความแตกต่างของพันธุ์ต่าง ๆ และของที่ดินแปลงต่าง ๆ ในการนี้เราอาจทำการปลูกพืชทุกพันธุ์ลงในที่ดินทุกแปลง โดยอาศัยการวิเคราะห์ความแปรปรวน (Analysis of Variance) เราก็อาจวิเคราะห์ผลของการทดลองดูความแตกต่างที่ต้องการได้

ถ้าให้ x_{ij} เป็นปริมาณของผลิตผลที่ได้รับจากพืชพันธุ์ที่ i ซึ่งปลูกในที่ดินแปลงที่ j เราก็ถือได้ว่า x_{ij} ทั้งหลายนี้เป็นตัวแปรสุ่มซึ่งไม่ขึ้นต่อกัน เราจะใช้การวิเคราะห์ความแปรปรวนได้ก็ต่อเมื่อ x_{ij} เหล่านี้มีการแจกแจงปกติ และมีความแปรปรวน σ^2 เดียวกัน สมมติให้มัชฌิมเลขคณิตของ x_{ij} เป็น η_{ij}

ให้

$$(๑.๑) \qquad \mu = \frac{\sum_{j=1}^{c}\sum_{i=1}^{r} \eta_{ij}}{rc}$$

$$(๑.๒) \qquad \eta_{i.} = \frac{\sum_{j=1}^{c} \eta_{ij}}{c}$$

$$(๑.๓) \qquad \eta_{.j} = \frac{\sum_{i=1}^{r} \eta_{ij}}{r}$$

เราจะได้ว่า

$$(๑.๔) \qquad \alpha_i = \eta_{i.} - \mu$$

คือ main effect ของผลผลิตของพืชพันธุ์ที่ i

และ

$$(๑.๕) \qquad \beta_j = \eta_{.j} - \mu$$

คือ main effect ของผลผลิตของพืชที่ปลูกในที่ดินแปลงที่ j

เราได้ว่า

(๑.๖) $$\sum_{i} \alpha_{i} = 0$$

และ

(๑.๗) $$\sum_{j} \beta_{j} = 0$$

เพราะว่า

$$\eta_{ij} = \mu + (\eta_{i.} - \mu) + (\eta_{.j} - \mu) + (\eta_{ij} - \eta_{i.} - \eta_{.j} + \mu)$$

$$= \mu + \alpha_{i} + \beta_{j} + \gamma_{ij}$$

โดยที่

$$\gamma_{ij} = \eta_{ij} - \eta_{i.} - \eta_{.j} + \mu$$

คือ interaction ของพืชพันธุ์ที่ i และที่ดินแปลงที่ j

เราได้ว่า

(๑.๘) $$\sum_{i} \gamma_{ij} = 0$$

และในทำนองเดียวกัน

(๑.๙) $$\sum_{j} \gamma_{ij} = 0$$

ถ้าเรามีพืชเพียงสองพันธุ์และมีที่ดินเพียงสองแปลง ตัวแปร i, j ย่อมมีค่าเพียงสองค่า คือ 1 กับ 2 ในกรณีเช่นนี้ เราจะได้ว่า

$$\gamma_{11} = \frac{1}{4} (\eta_{11} - \eta_{12} - \eta_{21} + \eta_{22})$$

$$\gamma_{22} = \frac{1}{4} (\eta_{22} - \eta_{21} - \eta_{12} + \eta_{11})$$

$$\gamma_{12} = \frac{1}{4} (\eta_{12} - \eta_{11} - \eta_{22} + \eta_{21})$$

$$\gamma_{21} = \frac{1}{4} (\eta_{21} - \eta_{22} - \eta_{11} + \eta_{12})$$

เราจะเห็นว่า ถ้ากรณีใดกรณีหนึ่งในสองกรณี ต่อไปนี้

(๑) $\eta_{11} > \eta_{12}, \quad \eta_{11} > \eta_{21}, \quad \eta_{22} > \eta_{12}, \quad \eta_{22} > \eta_{21}$

(๒) $\eta_{11} < \eta_{12}, \quad \eta_{11} < \eta_{21}, \quad \eta_{22} < \eta_{12}, \quad \eta_{22} < \eta_{21}$

เกิดขึ้นแล้วค่าของ γ_{ij} จะต่างจากศูนย์ ดังนั้นการที่เราจะดูว่ามี interaction หรือไม่นั้น เราจึงอาจดูได้โดยอาศัยอสมการ (๑) หรือ (๒)

<u>นิยาม</u> ให้

X_{11}	X_{12}
X_{21}	X_{22}

เป็นตัวแปรสุ่ม ซึ่งไม่ขึ้นต่อกัน

ให้

$$I = \begin{cases} 1 & \text{ถ้า } X_{11} > X_{12},\ X_{11} > X_{21},\ X_{22} > X_{12},\ X_{22} > X_{21} \\ 0 & \text{ในกรณีนอกนั้น} \end{cases}$$

และ

$$J = \begin{cases} 1 & \text{ถ้า } X_{11} < X_{12},\ X_{11} < X_{21},\ X_{22} < X_{12},\ X_{22} < X_{21} \\ 0 & \text{ในกรณีนอกนั้น} \end{cases}$$

ตัวแปรสุ่ม I กับ J ข้างบนนี้เป็นตัวแปรสุ่ม ซึ่งเราอาจใช้ค่าของมันเป็นเครื่องวัดเพื่อดูว่ามี interaction ระหว่าง แถวกับคอลัมน์หรือไม่ ถ้าในเซลล์หนึ่ง ๆ มี observation มากกว่าหนึ่ง observation เช่น

ถ้าให้

X_1 $\vdots$ X_P	Y_1 $\vdots$ Y_Q
Z_1 $\vdots$ Z_R	W_1 $\vdots$ W_S

เป็น observation ที่ได้จากการทดลองปลูกพืชสองพันธุ์ในที่ดินสองแปลง โดยที่ $X_1, \ldots\ldots, X_P$ เป็น observation ที่ได้จากพืชพันธุ์ที่หนึ่งซึ่งปลูกในที่ดินแปลงที่หนึ่ง

$Y_1, \ldots\ldots, Y_Q$ เป็น observation ที่ได้จากพืชพันธุ์ที่หนึ่งซึ่งปลูกในที่ดินแปลงที่สอง

$Z_1, \ldots\ldots, Z_R$ เป็น observation ที่ได้จากพืชพันธุ์ที่สองซึ่งปลูกในที่ดินแปลงที่หนึ่ง

และ $W_1, \ldots\ldots, W_S$ เป็น observation ที่ได้จากพืชพันธุ์ที่สองซึ่งปลูกในที่ดินแปลงที่สอง

ในที่นี้ X_p, Y_q, Z_r, W_s แต่ละชุดก็จะให้ค่า I ชุดละค่า และทำนองเดียวกันสำหรับค่า J เราจะนิยามตัวแปรสุ่ม I, J สำหรับกรณีที่เรามี observation มากกว่าหนึ่ง observation ในแต่เซลล์เสียใหม่ดังนี้

<u>นิยาม</u> ให้ $X(p), p = 1, \ldots\ldots, P, Y(q), q = 1, \ldots Q, Z(r), r=1, \ldots, R$ และ $W(s), s = 1, \ldots S$ เป็นตัวแปรสุ่มซึ่งไม่ขึ้นต่อกัน

ให้

$$(๑.๑๐) \quad I(p,q,r,s) = \begin{cases} 1 & \text{ถ้า } X(p) > Y(q), X(p) > Z(r), W(s) > Y(q), W(s) > Z(r) \\ 0 & \text{ในกรณีนอกนั้น} \end{cases}$$

$$(๑.๑๑) \quad J(p,q,r,s) = \begin{cases} 1 & \text{ถ้า } X(p) < Y(q), X(p) < Z(r), W(s) < Y(q), W(s) < Z(r) \\ 0 & \text{ในกรณีนอกนั้น} \end{cases}$$

$$(๑.๑๒) \qquad I = \Sigma\, I(p, q, r, s)$$

$$(๑.๑๓) \qquad J = \Sigma\, J(p, q, r, s)$$

โดยที่ในผลบวกทางขวามือของ (๑.๑๒) และ (๑.๑๓) นั้น ตัวแปร p, q, r, s แปรค่าจาก 1 ถึง P, Q, R, S ตามลำดับ

ในการทำวิทยานิพนธ์นี้ เรามีจุดมุ่งหมายที่จะศึกษาหาการกระจายของ I กับ J เนื้อหาสำคัญของผลลัพธ์ที่ได้อยู่ในบทที่ ๒ และบทที่ ๓ ในบทที่ ๒ เราหา joint probability distribution กับโมเมนต์ อันดับที่หนึ่ง และอันดับที่สองของ I(p,q,r,s) กับ J(p', q', r', s') ในบทที่ ๓ เราพิสูจน์ว่าเมื่อ P,Q,R,S มีค่ามากนั้น เราอาจประมาณ joint distribution function ของ I กับ J ได้ด้วย bivariate normal distribution.