เอกสาร
การพัฒนาบุคลากรในสังกัด
ด้วยวิธีการ Coaching

เรื่อง
“หลักสำคัญในการวางแผนการทดลอง”

โดย

นายเมธิน ศิริวงศ์
ผู้เชี่ยวชาญด้านวางระบบการพัฒนาที่ดิน
สำนักงานพัฒนาที่ดินเขต 8

พ.ศ. 2565

บทนำ

เอกสาร “ขั้นตอนในการวางแผนการทดลอง” และ “หลักสำคัญในการวางแผนการทดลอง” เป็นเอกสารที่ใช้ในการ Coaching ของผู้เชี่ยวชาญด้านวางระบบการพัฒนาที่ดิน สพข.๘ อันเนื่องมาจากการที่กรมฯ มีนโยบายในการขับเคลื่อนการเขียนโครงการวิจัยของทุกหน่วยงาน โดยกำหนดเป็นตัวชี้วัดระดับหน่วยงาน ปี ๒๕๖๕ ที่จะต้องมีโครงการวิจัยที่นำเสนอผ่านความเห็นชอบจากคณะทำงานวิชาการระดับหน่วยงาน (คณะ ๑) ส่งต่อคณะทำงานกลั่นกรองโครงการวิจัยตามสาขาวิชา (คณะ ๒) ไม่น้อยกว่า ๓ เรื่องต่อหน่วยงาน และจากประสบการณ์ของ ผชช.สพข.๘ ในการเป็นกรรมการในการพิจารณาโครงการวิจัยของกรม และกรรมการประเมินผลงานของนักวิจัย พบว่า งานวิจัยบางส่วนยังดำเนินการไม่ถูกต้องตามหลักวิชาการสถิติ ส่งผลให้เกิดความเสียหายทั้งงบประมาณ เวลา และที่สำคัญคือผลการวิจัยไม่ถูกต้อง ส่งผลกระทบเสียหายต่อไปถึงการนำผลการวิจัยไปถ่ายทอดหรือการต่อยอดงานวิจัยที่จะล้มเหลวตามมา รวมถึงงานวิจัยในเรื่องดังกล่าว ไม่สามารถนำกลับมาของบประมาณดำเนินการได้อีก

ดังนั้น เพื่อเป็นการแก้ปัญหาดังกล่าว จึงได้ทำการรวบรวมเอกสารการดำเนินงานวิจัยที่เกี่ยวข้องกับงานพัฒนาที่ดิน จากตำรา “สถิติ การวางแผนการทดลองขั้นต้น” ของอาจารย์สุรพล อุปดิสสกุล และเอกสารประกอบการบรรยายการฝึกอบรมสถิติหลักสูตรต่างๆ ของกรมพัฒนาที่ดิน และกรมวิชาการเกษตร ให้แก่นักวิชาการเกษตรในอดีตที่ผ่านมา โดยสรุปเฉพาะสาระที่เกี่ยวข้องกับงานของกรมพัฒนาที่ดิน และเป็นพื้นฐานเบื้องต้นสำหรับการใช้ประโยชน์อย่างง่ายสำหรับผู้ปฏิบัติ เพื่อให้นักวิจัยที่จะส่งข้อเสนอโครงการวิจัยได้มีกรอบวิชาการทางสถิติที่ถูกต้อง เพื่อประสิทธิผลของโครงการวิจัยในที่สุด

สุดท้ายนี้ ขอขอบพระคุณ อาจารย์สุรพล อุปดิสสกุล คุณสง่า ดวงรัตน์ และคุณสุขาวดี นาคะทัต ที่ผู้รวบรวมได้ใช้ต้นฉบับเอกสารทางสถิติของท่านเพื่อจัดทำเอกสารฉบับนี้

หลักสำคัญในการวางแผนการทดลอง
(PRINCIPLES OF EXPERIMENTAL DESIGN)

โดย นายเมธิน ศิริวงศ์
ผู้เชี่ยวชาญด้านวางระบบการพัฒนาที่ดิน สพข.8

หลักการใหญ่ในการวางแผนการทดลองประกอบด้วยการทำซ้ำ (replication) การสุ่ม (randomization) และการควบคุมความคลาดเคลื่อนโดยการจัดกลุ่มหน่วยการทดลอง (local control or error control by grouping of experiment units)

1. การทำซ้ำ (replication)

การทำซ้ำหมายถึง การที่ผู้ทดลองจัดสิ่งทดลอง (treatment) หนึ่งๆ ลงในหน่วยทดลอง (experimental unit) มากกว่า 1 หน่วย ทั้งนี้ไม่จำเป็นว่าทุกๆ สิ่งทดลองจะต้องมีการทำซ้ำเท่ากันหมด อาจจะใช้จำนวนซ้ำต่างกันก็ได้ขึ้นอยู่กับความเหมาะสมและการพิจารณาของผู้ทดลอง แต่ปกติมักใช้จำนวนซ้ำเท่ากันเพื่อความสะดวกในการวิเคราะห์ผลทางสถิติ เหตุผลที่จะต้องมีการทำซ้ำมีดังนี้

1.1 เพื่อสามารถประมาณค่าความคลาดเคลื่อนของการทดลอง (experimental error) ดังที่กล่าวมาแล้วว่าความคลาดเคลื่อนของการทดลองจะประมาณได้จากความแปรปรวน (variability) ของหน่วยการทดลองต่างๆ ที่ได้รับสิ่งทดลองอย่างเดียวกันหรือเรียกว่า variance (s^2) ความคลาดเคลื่อนนี้จะใช้ทดสอบความแตกต่างระหว่างสิ่งทดลอง (treatment differences) ถ้าหากทำการทดลองเพียงซ้ำเดียว (ในบางกรณี) ผู้ทดลองไม่สามารถจะรู้ได้ว่าความแตกต่างระหว่างสิ่งทดลองนั้นเกิดจากผลหรืออิทธิพลของสิ่งทดลองเอง หรือเกิดเนื่องจากความแปรปรวนซึ่งมีอยู่แล้วระหว่างหน่วยทดลอง ในแผนการทดลองบางชนิด เช่น randomized complete block แบบธรรมดา ถ้าไม่มีการทำซ้ำก็จะไม่สามารถทดสอบความแตกต่างระหว่างสิ่งทดลองได้เลย เพราะว่าไม่มีตัวความคลาดเคลื่อน (error term)

1.2 เพื่อเพิ่มความเที่ยงตรง (accuracy) ของการทดลอง ในการทดลองต่างๆ นั้น ผู้ทดลองจะต้องตระหนักว่าข้อมูลหรือค่าต่างๆ ที่วัดผลหรือบันทึกได้จากการทดลองนั้น ปกติเป็นค่าของตัวอย่าง (sample) ซึ่งเป็นตัวประเมินค่า (estimator) ของค่าที่แท้จริงของประชาการ (population) เท่านั้น เช่น ความแปรปรวนของการทดลอง (sample variance, s^2) เป็นตัวประเมินค่าของความแปรปรวนของประชาการ (population variance,∂^2) หรือค่าเฉลี่ยของตัวอย่าง (sample mean, $\bar{y}$) เป็นตัวประเมินค่าของค่าเฉลี่ยของประชากร (population mean, μ) เป็นต้น ดังนั้นเมื่อจำนวนซ้ำมากขึ้นหรือจำนวนตัวอย่างที่นำมาใช้คำนวณค่าทางสถิติต่างๆ มีมากขึ้น โอกาสที่จะได้ค่าทางสถิติจากตัวอย่างแต่ละค่าก็ย่อมใกล้เคียงกับค่านั้นๆ ของประชากรมากขึ้น ทั้งนี้เนื่องจากในการทดลองต่างๆ นั้นตามปกติผู้ทดลองจะไม่สามารถนำประชากรทั้งหมดมาทำการทดลองได้ ดังนั้นการสุ่มตัวอย่างมาจากประชากรให้มากที่สุด เท่าที่จะทำได้ก็จะเป็นการเพิ่มความเที่ยงตรงของการทดลอง หรือได้ค่าทางสถิติต่างๆ ที่ไม่

ลำเอียงหรือไม่อคติ (unbias) โอกาสที่จะสรุปผลของสิ่งที่ศึกษาทดลองในประชากรนั้นๆ ได้ถูกต้องโดยไม่มีอคติก็จะมีมากขึ้น จึงนับว่าเป็นการทดลองที่เชื่อถือได้

1.3 เพื่อเพิ่มความแน่นอน (precision) ของการทดลอง การเพิ่มจำนวนซ้ำนอกจากจะทำให้ความเที่ยงตรง (accuracy) เพิ่มขึ้นแล้ว ความแน่นอนก็เพิ่มขึ้นด้วย เช่น การหาค่าเฉลี่ยของตัวอย่าง ($\bar{y}$) ของประชากรใดประชากรหนึ่งจากตัวอย่างหลายๆ กลุ่ม ค่าเฉลี่ยของตัวอย่างเหล่านี้จะมีค่าใกล้เคียงระหว่างกันและกันหากว่าจำนวนสมาชิกของตัวอย่างในแต่ละกลุ่มมีมาก ดังนั้นในการทดลองที่มีจำนวนซ้ำสูงพอแล้ว หากว่ามีการทดลองแบบเดียวกันอีก (repetition) ผลที่ได้ก็ย่อมใกล้เคียงกัน หรือกล่าวได้ว่าผลการทดลองนี้มีความแน่นอนสูง นอกจากนี้เมื่อเพิ่มจำนวนซ้ำแล้ว ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยของตัวอย่าง (standard deviation of sample means บางทีเรียกสั้นๆ ว่า standard error, $s_{\bar{y}}$) จะลดลง เนื่องจาก $s_{\bar{y}} = s/\sqrt{n}$ ซึ่ง s คือความเบี่ยงเบนมาตรฐานของตัวอย่าง (standard deviation of samples) ซึ่งเท่ากับ $\sqrt{s^2}$ n คือจำนวนซ้ำ ดังนั้นจะเห็นได้ว่า $s_{\bar{y}}$ จะเป็นสัดส่วนกลับกับรากที่สองของจำนวนซ้ำ กล่าวคือ เมื่อจำนวนซ้ำ หรือ n เพิ่มขึ้น ค่า $s_{\bar{y}}$ ก็จะลดลง

ค่า $s_{\bar{y}}$ แสดงถึงความแปรปรวนของค่าเฉลี่ยของตัวอย่าง ($\bar{y}$) ของสิ่งทดลองต่างๆ ดังนั้นเมื่อ $s_{\bar{y}}$ ลดลง ค่าเฉลี่ยของตัวอย่างของสิ่งทดลองที่ได้จึงมีความแน่นอนสูง เป็นผลให้การทดลองนั้นมีความแน่นอนสูงตามไปด้วย

1.4 เพื่อลดความคลาดเคลื่อนของการทดลอง (experimental error) การเพิ่มจำนวนซ้ำจะทำให้ degree of freedom (d.f.) ของตัวความคลาดเคลื่อน (error term) มากขึ้น ดังนั้นการเพิ่มจำนวนซ้ำเท่ากับทำให้ค่าของความแปรปรวน (variance) ของความคลาดเคลื่อนลดลง ซึ่งมีผลให้การทดลองมีความว่องไวในการวัดความแตกต่าง (sensitivity) ได้ดี หรือมีความว่องไวในการเปรียบเทียบสูงขึ้น จะสามารถวัดความแตกต่างเพียงเล็กน้อยระหว่างสิ่งทดลองได้ดีขึ้น

อีกนัยหนึ่งการเพิ่มจำนวนซ้ำจะทำให้ความคลาดเคลื่อนแบบที่ 1 และแบบที่ 2 (type I error and type II error) ของการทดลองลดลง ซึ่งเป็นผลให้ความว่องไวในการวัดผลหรือ sensitivity เพิ่มขึ้นดังกล่าวมาแล้ว

1.5 การทำซ้ำจะทำให้สรุปผลการทดลองได้กว้างขึ้น ทั้งนี้เพราะว่าในการทำซ้ำโดยเฉพาะอย่างยิ่งการทดลองในไร่นานั้นจะต้องใช้พื้นที่มากขึ้น อันเป็นผลให้ได้ครอบคลุมหรือได้ตัวแทนของดิน (ซึ่งมีความแปรปรวนต่างกันอยู่) ดีขึ้น นอกจากนี้การทำซ้ำหลายๆ ปี ตลอดจนหลายๆ ท้องที่ (location) จะยิ่งทำให้ได้ผลที่แน่นอนยิ่งขึ้น เพราะว่าแต่ละปีและท้องที่ย่อมมีความแปรปรวนหรือมีความแตกต่างกันอยู่ ถ้าหากใช้ข้อมูลจากที่เดียวหรือฤดูเดียวอาจจะทำให้การสรุปผลการทดลองเป็นไปอย่างแคบๆ เฉพาะการทดลองนั้นๆ เท่านั้น แม้กระทั่งการทดลองในห้องปฏิบัติการก็เช่นกัน หากว่ามีการทำซ้ำโดยบุคคลอื่นหรือสภาพอื่นก็ย่อมทำให้สรุปผลดีขึ้น

2. การพิจารณาจำนวนซ้ำ

การพิจารณาจำนวนซ้ำอาจอาศัยหลักต่อไปนี้ประกอบการพิจารณาคือ

2.1 ความเที่ยงตรง (accuracy) และความแน่นอน (precision) ที่ต้องการในงานวิจัยนั้น หากว่าปริมาณความแตกต่างของสิ่งทดลองมีน้อย หรือผู้ทดลองต้องการวัดผลที่แตกต่างกันเพียงเล็กน้อย ก็ต้องใช้จำนวนซ้ำมากขึ้น ดังที่กล่าวมาแล้วในข้อ 1.2 – 1.4

2.2 คุณสมบัติของหน่วยการสุ่ม (sampling unit) หรือขนาดหรือจำนวนที่สุ่มวัดข้อมูลและหน่วยทดลอง (experimental unit)ถ้าหากว่าหน่วยทั้งสองหรืออย่างใดอย่างหนึ่งมีความแปรปรวนเนื่องจากตัวของมันเองมาก เช่นพืชผสมข้าม (cross pollinated crops) หรือดินที่มีความแปรปรวนแตกต่างกันมาก ก็ควรใช้จำนวนซ้ำมาก เพื่อที่จะได้ตัวแทนของแต่ละสิ่งทดลองอย่างแท้จริง หรือสามารถวัดผลความแตกต่างระหว่างสิ่งทดลองอย่างแท้จริง มิใช่เกิดจากปัจจัยอื่น

2.3 จำนวนสิ่งทดลองที่ใช้ ถ้าหากว่าสิ่งทดลองมีน้อยจำนวนซ้ำก็ต้องสูงพอที่จะให้จำนวน degree of freedom (d.f.) ของความคลาดเคลื่อนเหมาะสม ปกติ d.f. ของความคลาดเคลื่อนควรจะไม่ต่ำกว่า 12 ทั้งนี้หากว่าค่า d.f. ของความคลาดเคลื่อนต่ำ ประสิทธิภาพในการตรวจสอบหาความแตกต่างระหว่างสิ่งทดลองก็ต่ำ เนื่องจากว่าค่า F-ratio จากตารางทางสถิติ (ซึ่งใช้เป็นตัวเปรียบเทียบกับค่า F-ratio ที่คำนวณได้) จะมีค่าสูงมาก แต่ถ้า d.f. ของความคลาดเคลื่อนค่อนข้างสูงจะลดความผิดพลาดเนื่องจากการสุ่ม (randomization) อันเป็นผลให้ความแปรปรวน (variance) ของความคลาดเคลื่อนลดลง หรือตัวหารในการคำนวณค่า F-ratio ลดลงซึ่งมีผลให้ค่า F-ratio ที่คำนวณได้สูงขึ้น

2.4 ชนิดของแผนการทดลอง รวมทั้งขนาดและรูปร่างของหน่วยการทดลอง ทั้งนี้แผนการทดลองแต่ละชนิดมีหลักการแยกความแปรปรวนจากแหล่งต่างๆ และหลักการคำนวณหา d.f. ของความคลาดเคลื่อนต่างกัน หากว่าแผนการทดลองดี ขนาดและรูปร่างของหน่วยการทดลองเหมาะสม (โดยเฉพาะงานวิจัยในไร่นา ขนาดและรูปร่างของหน่วยการทดลองสำคัญมาก) จำนวนซ้ำก็อาจไม่ต้องสูงนัก

2.5 ปัจจัยอื่นๆ เช่น พื้นที่ในการทดลอง เงินทุนในการวิจัย แรงงาน ตลอดจนวัตถุหรือวัสดุในการทดลอง เป็นต้น ปัจจัยเหล่านี้จะมีข้อจำกัดในตัวมันเองว่าควรจะมีจำนวนซ้ำได้ไม่เกินจำนวนเท่าใด

3. การคำนวณหาจำนวนซ้ำ

นอกจากหลักการกว้างๆ ที่กล่าวว่า d.f. ของตัวความคลาดเคลื่อน (error term) ควรจะไม่ต่ำเกินไปแล้วนั้น ผู้ทดลองอาจมีวิธีคำนวณหาจำนวนซ้ำได้หลายวิธี แต่ที่ง่ายและนิยมกันก็โดยวิธี statistic t

W.T.Federer (1955) ได้เสนอสูตรการคำนวณ ดังนี้

$r = (2t_{\alpha}^2 s^2)/d^2$

r = จำนวนซ้ำ

t_{α} = ค่า t ซึ่งดูได้จากตารางทางสถิติของ t (t-table) ที่ระดับความเป็นไปได้ α โดยใช้ d.f. ของ s^2

s^2 = ค่าประเมินของ error variance (∂^2) ซึ่งผู้ทดลองดูได้จากการทดลองก่อนๆ ที่มีลักษณะคล้ายๆ กัน

d = ความแตกต่างระหว่างค่าเฉลี่ยของสิ่งทดลอง (treatment) ที่ผู้ทดลองต้องการตรวจสอบให้ละเอียดมากน้อยแค่ไหน ถ้าหากต้องการวัดให้ละเอียดมากคือ สามารถวัดความแตกต่างเพียงเล็กน้อยได้ ก็ใช้ค่า d น้อย ค่า d นี้ผู้ทดลองกำหนดขึ้นเอง

W.H.Hathaway (1961) ได้ดัดแปลงสูตรที่แนะนำโดย W.G.Cochran และ G.M.Cox (1957) ซึ่งมีลักษณะคล้ายๆ กับสูตรของ Federer ดังนี้

$$r \geq [2C^2(t1+t2)^2]/d^2$$

ซึ่ง r = จำนวนซ้ำ

C = coefficient of variation (C.V.) ซึ่งผู้ทดลองดูได้จาก การทดลองก่อนๆ ที่มีลักษณะคล้ายๆ กัน

t1 = ค่า t จากตารางทางสถิติ (t- table) ที่ระดับความเป็นไปได้ (α) ตามต้องการ โดยใช้ d.f. ของตัวคลาดเคลื่อน (error term)

t2 = ค่า t จากตาราง จะเปิดโดยใช้ระดับความเป็นไปได้ (α) = 2(1-P) และใช้ d.f. ของตัวคลาดเคลื่อน โดยที่ P = propability ที่ทำการทดลองจะให้ผลของการทดลองแสดงความแตกต่างทางสถิติ (significant result)

d = ปริมาณความแตกต่างระหว่างสิ่งทดลองที่ต้องการทดสอบแสดงอยู่ในรูปร้อยละ (percent) ของค่าเฉลี่ย (mean)

ทั้งวิธีของ Federer และ Hathaway ที่กล่าวมาแล้วนั้น มีปัญหาในการคำนวณคล้ายๆ กัน กล่าวคือผู้ทดลองรู้ค่าต่างๆ ในสมการยกเว้น 2 ค่า คือค่า r และค่า t เนื่องจากจำนวน d.f. ของตัวคลาดเคลื่อน (error term) ที่จะใช้เปิดค่า t จากตารางนั้นขึ้นอยู่กับจำนวนซ้ำหรือค่า r ดังนั้นขั้นตอนในการคำนวณจะมีดังนี้

3.1 สมมติค่า d.f. ของ error

3.2 เปิดค่า t จากตารางทางสถิติ (t-table) ตามระดับความเป็นไปได้ (α) ตามความต้องการเช่น α = .05 โดยใช้ d.f. ที่สมมติขึ้นในข้อ 3.1

3.3 แทนค่า t ที่เปิดได้ในสมการ จะได้ค่าขั้นต้นของจำนวนซ้ำ หรือ r_1 ออกมา

3.4 ใช้ค่า r_1 นี้หา d.f. ของ error อีก (ตามชนิดหรือวิธีการหาค่า d.f. error ของแผนการทดลองนั้น)

3.5 เปิดค่า t ที่เปิดได้ในสมการ โดยใช้ d.f. ที่ได้จากข้อ 3.4

3.6 แทนค่า t จากข้อ 3.5 ในสมการ คำนวณได้ค่า r_2

3.7 เปรียบเทียบค่า r_1 และ r_2 หากว่า $r_1 = r_2$ ก็สรุปจำนวนซ้ำได้ทันที หากว่า $r_1 \neq r_2$ ก็คำนวณหา r_3 โดยวิธีการคล้ายๆ กับการหาค่า r_2 (คือข้อ 3.4-3.6)

3.8 ให้เปรียบเทียบค่า r ที่คำนวณได้กับตัวเดิมเช่นนี้เรื่อยๆ จนได้ค่า r เท่ากัน เช่น $r_3 = r_2$ เป็นต้น ปกติการหาจำนวนซ้ำนี้จะใช้การคาดคะเนประมาณ 3 ครั้ง (r_3) ก็จะได้จำนวนซ้ำที่เหมาะสมกับการทดลองนั้น

4. การสุ่ม (randomization)

การสุ่มเป็นวิธีหนึ่ง (หรืออาจเป็นวิธีที่ดีที่สุด) ที่จะประกันได้ว่าแต่ละสิ่งทดลองจะถูกจัดลงในหน่วยการทดลองอย่างไม่มีอคติหรือเที่ยงธรรม ดังนั้นอิทธิพลจากปัจจัยต่างๆ อันเป็นแหล่งของความคลาดเคลื่อนต่อสิ่งทดลองหนึ่งๆ จะเป็นไปได้โดยสุ่ม วิธีการสุ่มและขอบเขต หรือโอกาสที่จะใช้การสุ่มขึ้นอยู่กับชนิดของการทดลอง

ประโยชน์ของการสุ่มพอจะสรุปได้ดังนี้

4.1 การสุ่มช่วยเพิ่มประสิทธิภาพในการประมาณค่าความคลาดเคลื่อนของการทดลอง (experimental error) ทั้งนี้ความคลาดเคลื่อนในการทดลองขึ้นอยู่กับความแตกต่างหรือความแปรปรวนระหว่างหน่วยทดลอง (experimental unit) ต่างๆ ที่ได้รับสิ่งทดลองอย่างเดียวกัน ดังนั้นการประมาณค่าคลาดเคลื่อนนี้จะมีประสิทธิภาพเมื่อหน่วยการทดลองที่ได้รับสิ่งทดลองอย่างเดียวกันกระจายอยู่ทั่วไปในการทดลองนั้น ทั้งนี้เพราะว่าโดยเฉลี่ยแล้วหน่วยการทดลองที่อยู่ใกล้กันหรือเหมือนกันมีแนวโน้มที่จะให้ผลคล้ายกันมากกว่าหน่วยการทดลองที่อยู่ห่างกันหรือแตกต่างกัน เช่น ลูกคนแรกกับลูกคนสุดท้าย ดังนั้นเพื่อตัดปัญหาเรื่องนี้ ผู้ทดลองควรใช้การสุ่มจัดสิ่งทดลองลงในหน่วยการทดลองต่างๆ เพื่อให้แต่ละสิ่งทดลองจะอยู่ที่ไหนก็ได้ภายในขอบเขตของแผนการทดลองชนิดนั้น (ทั้งนี้ก็เพื่อพยายามลดความแปรปรวนจากแหล่งที่ไม่สามารถควบคุมหรือทราบสาเหตุ)

4.2 การสุ่มจะขจัดอคติหรือความลำเอียง (bias) ในการจัดสิ่งทดลองต่างๆ ลงในหน่วยการทดลอง โดยปกติแล้วถ้าหากไม่มีการสุ่ม ความลำเลียงหรืออคติจะเกิดขึ้นเสมอ ซึ่งอาจเรียกว่า ความอคติเนื่องจากการจัดระบบ (systematic bias) เช่น ลูกคนแรกรับเงินไปก่อน ส่วนลูกคนสุดท้องได้ส่วนที่เหลือ ดังนั้น เพื่อที่จะให้ทุกสิ่งทดลองมีโอกาสเท่ากันหมดที่จะอยู่ในหน่วยการทดลองใดๆ ก็ได้ ผู้ทดลองจะต้องใช้การสุ่ม โดยเฉพาะอย่างยิ่งการทดลองที่มีการทำซ้ำ การสุ่มจะลดอคติหรือความลำเอียงลงไปได้มาก

อย่างไรก็ดีผู้ทดลองจะต้องเข้าใจเกี่ยวกับการสุ่มให้ดี เพราะงานวิจัยบางอย่างนั้นไม่สามารถใช้การสุ่มอย่างสมบูรณ์ได้ หรือใช้ไม่ได้เหมาะ ตัวอย่างเช่น การศึกษาระบบรากพืชที่ใช้ระยะปลูกต่างๆ กัน โดยใช้เครื่องเจาะดินที่มีรากพืชนั้นมาศึกษาดูว่ารากจะมีการเจริญเติบโตแผ่ขยายไปทางด้านใดบ้าง (ทั้งทางแนวดิ่งและแนวนอน) จะเห็นว่าไม่สามารถสุ่มเจาะดินได้ แต่จะต้องทำการเจาะอย่างมีระบบ คือแต่ละหลุมจะต้องห่างจากหลุมอื่นของแต่ละพื้นที่ทดลอง

W.G. Cochran และ G.M. Cox (1957) ได้กล่าวไว้พอสรุปได้ว่า “การสุ่มนั้นเปรียบเสมือนการประกันชีวิตหรือประกันภัย คือผู้ทำประกันทำไว้เพื่อกันเหตุการณ์ที่ไม่คาดฝันเท่านั้น ไม่จำเป็นที่ผู้ประกันจะต้องประสบเหตุนั้นๆ เสมอไป และหากเกิดเหตุการณ์นั้นๆ ขึ้นผลที่ตามมาก็อาจไม่รุนแรงนัก”

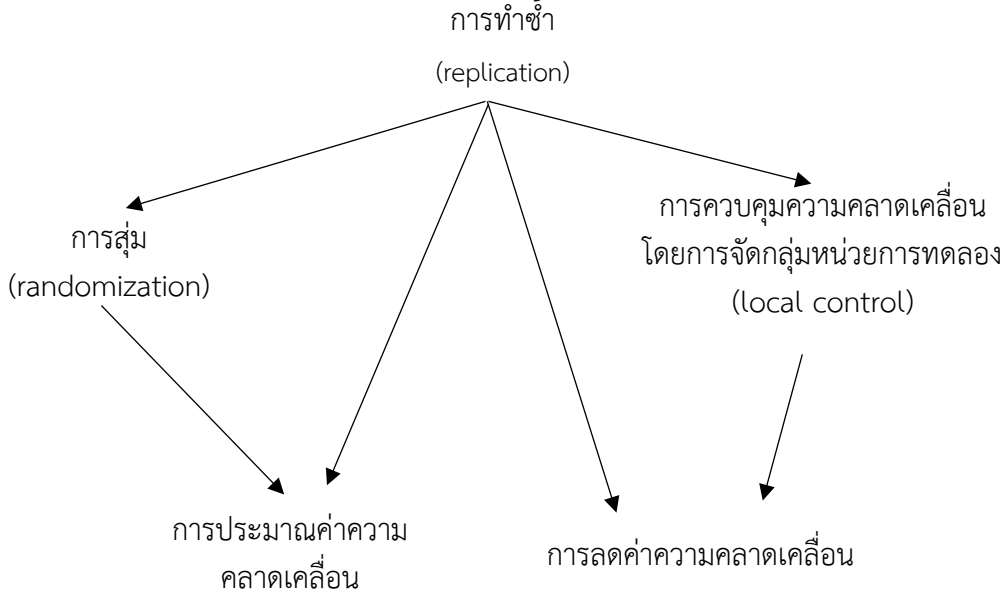
5. การควบคุมความคลาดเคลื่อนโดยการจัดกลุ่มหน่วยการทดลอง (local control or error control by grouping of experimental units)

การควบคุมความคลาดเคลื่อน (error) โดยการจัดกลุ่มหน่วยการทดลอง ในที่นี้หมายถึงการจัดกลุ่มของหน่วยการทดลองโดยมีการตั้งข้อจำกัดบางอย่างในการสุ่มสิ่งทดลองลงในหน่วยการทดลอง ทั้งนี้ก็เพื่อที่จะสามารถควบคุมหรือทราบแหล่งที่ก่อให้เกิดความแปรปรวน และแยกออกจากความคลาดเคลื่อนในการทดลอง ให้เหลือเป็นความคลาดเคลื่อนของการทดลองที่แท้จริง (experimental error) โดยไม่พัวพันกับปัจจัยอื่นที่มีผลทำให้เกิดความแปรปรวน ทำให้การทดลองมีประสิทธิภาพสูงขึ้นซึ่งถ้าพิจารณาให้ดีจะเห็นว่าการจัดกลุ่มนี้จะคล้อยตามแผนการทดลอง ซึ่งมีลักษณะการจัดสิ่งทดลองลงในหน่วยการทดลองแบบต่างๆ กันไป บางท่านกล่าวว่า local control นี้เป็นคำพ้องกับแผนการทดลอง (experimental design) นั่นเอง

การจัดกลุ่มของหน่วยการทดลองนี้เกี่ยวข้องกับการจัดบล็อค (block) ซึ่ง “บล็อค” หมายถึงกลุ่มหรือหมู่ของหน่วยการทดลองกลุ่มหนึ่ง การจัดบล็อคนี้มีจุดมุ่งหมายให้หน่วยการทดลองภายในแต่ละบล็อคนั้นมีความสม่ำเสมอมากที่สุด หรือแตกต่างกันน้อยที่สุด เช่น ในไร่นา อาจเป็นพื้นที่ที่ดินเป็นลักษณะเดียวกัน หรืออยู่ในความลาดเทระดับเดียวกัน ส่วนระหว่างบล็อคให้มีความแตกต่างกันได้ ดังนั้นความแตกต่างระหว่างหน่วยการทดลองภายในบล็อคหนึ่งๆ ก็เนื่องจากปัจจัยบางอย่างที่ไม่สามารถควบคุมหรือทราบจนแยกแยะออกมาได้จริงๆ ซึ่งจะรวมอยู่ในความคลาดเคลื่อนของการทดลอง สำหรับในการจัดกลุ่มที่ดีจะพบว่าความคลาดเคลื่อนของการทดลองจะลดลงมากกว่ากรณีที่จัดกลุ่มไม่ถูกต้อง ในการวิเคราะห์ผลทางสถิติโดยวิธีวิเคราะห์ความแปรปรวน (analysis of variance) อิทธิพลของบล็อคก็จะถูกแยกออกไปต่างหากจากอิทธิพลของสิ่งทดลองและความคลาดเคลื่อนของการทดลอง ซึ่งจะมีผลให้ผู้ทดลองสามารถควบคุมหรือแยกแยะความคลาดเคลื่อนได้ดีขึ้น

ในการจัดสิ่งทดลองลงในหน่วยการทดลองของแต่ละบล็อคนั้น จะต่างกันไปขึ้นอยู่กับแผนการทดลอง (experimental design) ดังที่กล่าวมาแล้ว ในแผนการทดลองแบบง่ายๆ เช่น randomized complete block (RCB) ในแต่ละบล็อคจะต้องมีครบทุกสิ่งทดลอง บล็อคแบบนี้เรียกว่า “บล็อคสมบูรณ์” (complete block) เป็นซ้ำหนึ่งที่มีครบทุกสิ่งทดลอง (treatment) แผนการทดลองบางอย่างจะมีบางสิ่งทดลองอยู่ในแต่ละบล็อคหรือแต่ละบล็อคมีสิ่งทดลองไม่ครบ บล็อคชนิดหลังนี้เรียกว่า “บล็อคไม่สมบูรณ์” (incomplete block) ซึ่งการใช้บล็อคไม่สมบูรณ์นี้ก็มีเหตุผลที่จะไม่ให้ขนาดของบล็อคใหญ่เกินไป อันเป็นผลให้การควบคุมความคลาดเคลื่อนไม่มีประสิทธิภาพ

6. ความสัมพันธ์ระหว่างการทำซ้ำ (replication) การสุ่ม (randomization) และการควบคุมความคลาดเคลื่อนโดยการจัดกลุ่มหน่วยการทดลอง (local control) ที่มีต่อความคลาดเคลื่อนในการทดลอง แสดงได้ดังนี้



ภาพแสดงความสัมพันธ์ของการทำซ้ำ การสุ่มและการควบคุมความคลาดเคลื่อน โดยการจัดกลุ่มหน่วยทดลอง ต่อความคลาดเคลื่อน

จะเห็นได้ว่าการทำซ้ำมีบทบาทสำคัญมาก คือทั้งสามารถใช้ในการประมาณค่าความคลาดเคลื่อนอย่างมีประสิทธิภาพและลดความคลาดเคลื่อน นอกจากนี้ก็มีส่วนสัมพันธ์ต่อการสุ่มและการจัดกลุ่มสิ่งทดลองด้วย ดังนั้นผู้ทดลองจะต้องพิจารณาเรื่องการทำซ้ำให้มากเป็นพิเศษ อย่างไรก็ดีหากว่าการเพิ่มจำนวนซ้ำให้มากจะต้องใช้วัสดุหรือวัตถุในการทดลอง (experimental material) ที่ไม่ค่อยสม่ำเสมอ หรือทำให้การปฏิบัติไม่สะดวกแล้วผลที่ได้ก็ไม่มีประสิทธิภาพเท่าการใช้จำนวนซ้ำที่พอเหมาะหรือไม่มากนัก