คำนำ

เอกสารนี้เป็นการรวบรวมเนื้อหาวิชาสถิติที่ใช้ในงานวิจัยทางด้านการเกษตร โดยแบ่งออกเป็น 2 เล่ม เล่มที่ 1 ประกอบด้วย หลักวิชาการสถิติ การวางแผนงานทดลองแบบต่างๆ สหสัมพันธ์และรีเกรสชัน ส่วนเล่มที่ 2 ประกอบด้วย ข้อมูลที่มีปัญหา เทคนิคทางสถิติในการปฏิบัติงานทดลอง การสุ่มเก็บตัวอย่างและการบันทึกข้อมูล การวิเคราะห์รวมและการนำเสนอผลงานทดลองแบบต่างๆ

กลุ่มวิจัยและวิเคราะห์สถิติการเกษตร (ฝ่ายวิชาการสถิติ) หวังว่า เอกสารเล่มนี้จะเป็นประโยชน์ในการปฏิบัติงานวิจัยของท่าน

กลุ่มวิจัยและวิเคราะห์สถิติการเกษตร

ศูนย์สารสนเทศ

2553

สารบัญ

หลักวิชาการสถิติ

วิจิตรา พลเยี่ยม

หลักวิชาสถิติ

ความหมายของสถิติ

สถิติมีความหมายอย่างกว้าง ๆ 2 ประการ คือ

1. หมายถึงยอดตัวเลขที่ได้จากการรวบรวม เพื่อใช้บรรยายเหตุการณ์ข้อเท็จจริงต่าง ๆ เช่น สถิติผลผลิตข้าวโพดของประเทศไทย สถิติการเกิดการตาย สถิติน้ำฝน เป็นต้น

2. หมายถึงศาสตร์ที่ว่าด้วยหลักการและระเบียบและวิธีการทางสถิติที่ใช้ในการเก็บรวบรวม การนำเสนอ การวิเคราะห์ และการตีความหมายของตัวเลขเหล่านั้น

ตัวเลขเพียงตัวเดียวหรือความจริงเพียงข้อเดียวไม่ก่อให้เกิดสถิติ เช่น ในการแข่งขันวิ่ง 100 เมตร ผู้ชนะทำสถิติได้ 10 วินาที ไม่ถือว่าเป็นสถิติตามคำนิยามแต่เป็นข้อมูลประเภททะเบียนในลักษณะของ “best record”

สถิติแบ่งออกเป็น 2 ชนิดใหญ่ คือ

1. <u>สถิติเชิงพรรณนา</u> (Descriptive) ได้แก่ วิธีการทางสถิติ ที่เกี่ยวข้องกับการเก็บรวบรวมข้อมูลจัดให้อยู่ในรูปแบบที่สามารถสรุปคุณสมบัติของประชากรที่เราศึกษาอยู่ เช่น ทำตารางเสนอเป็นแผนภาพ เป็นต้น

2. <u>สถิติเชิงอ้างอิง</u> (Inferential) ได้แก่ วิธีทางสถิติที่ใช้วิเคราะห์หาข้อสรุปเกี่ยวกับประชากรโดยอาศัยผลของการศึกษาจากตัวอย่างบางส่วนมาอ้างอิงว่าเป็นคุณสมบัติของประชากรนั้น

ประโยชน์ของสถิติ

ประโยชน์ของสถิติพอสรุปได้ดังนี้

1. เป็นสิ่งชี้ให้เห็นถึงข้อเท็จจริงของเหตุการณ์ หรือ เรื่องราวที่กำลังสนใจศึกษาโดยใช้หลักเกณฑ์ของการมีเหตุผลที่ดี
2. เป็นเครื่องมือในการวางแผนของโครงการหรืองานต่างๆ
3. เป็นระเบียบวิธีสำหรับการวิเคราะห์ในงานวิจัยทั่ว ๆ ไป
4. เป็นเครื่องมือสำหรับการประเมินผลงานที่ได้กระทำไปแล้ว

ศัพท์สถิติ

<u>ค่าสังเกต</u> (Observation) หมายถึงค่าที่วัดได้ของลักษณะใดลักษณะหนึ่ง เช่น ในการทดลองเปรียบเทียบพันธุ์ข้าว ค่าสังเกตอาจเป็นผลผลิตข้าวเป็นกิโลกรัมต่อไร่ ซึ่งแตกต่างกันไปในแต่ละแปลง

<u>ข้อมูล</u> (Data) หมายถึงกลุ่มของค่าสังเกตที่ได้จากการทดลองใดทดลองหนึ่ง ข้อมูลจำแนกออกได้เป็น 2 ประเภท คือ

1. <u>ข้อมูลเชิงปริมาณ</u> (Quantitative Data) ได้แก่ ข้อเท็จจริงเกี่ยวกับตัวเลข เช่น ความสูง น้ำหนัก อายุ ฯลฯ

2. <u>ข้อมูลเชิงคุณภาพ</u> (Qualitative Data) ได้แก่ ข้อเท็จจริงที่ไม่เกี่ยวกับตัวเลข เช่น พันธุ์พืช เพศ ฯลฯ

<u>ตัวแปร</u> (Variable) หมายถึงลักษณะบางอย่างที่เราสนใจศึกษา โดยที่ลักษณะนั้น ๆ สามารถเปลี่ยนค่าได้ เช่น ในการศึกษาคัดพันธุ์ปอแก้ว ตัวแปรอาจจะเป็นน้ำหนักเส้นใยแห้งและความสูง ตัวแปร มี 2 ชนิด คือ ชนิดที่หนึ่ง ตัวแปรอิสระ (Independent Variable) ซึ่งเป็นเหตุทำให้เกิดการเปลี่ยนแปลงให้ตัวแปรชนิดที่สอง ซึ่งเรียกว่า ตัวแปรตาม (Dependent Variable)

ตัวแปรมี 2 ลักษณะ คือ

1. <u>ตัวแปรต่อเนื่อง</u> (Continuous Variable) ได้แก่ ตัวแปรที่ให้ค่าเป็นเลขที่ต่อเนื่องหรือมีจุดทศนิยม เช่น ความสูง น้ำหนัก

2. <u>ตัวแปรไม่ต่อเนื่อง</u> (Discrete Variable) ได้แก่ ตัวแปรที่ให้ค่าเป็นตัวเลขไม่ต่อเนื่อง เช่น จำนวนรวงข้าวต่อกอ เป็นต้น ตัวแปรชนิดนี้มักให้ค่าที่เกิดจากตัวเลขจำนวนเต็มเสมอ

<u>ประชากร</u> (Population) หมายถึง กลุ่มของสิ่งที่เรากำลังสนใจศึกษาทั้งหมด ณ เวลาใดเวลาหนึ่ง ค่าที่วัดได้จากประชากร เรียกว่า <u>พารามิเตอร์</u> (Parameter)

<u>ตัวอย่าง</u> (Sample) หมายถึง ส่วนหนึ่งของประชากรที่เลือกขึ้นมาเพื่อเป็นตัวแทนของประชากรนั้น ๆ ตัวอย่างที่เลือกมาโดยการสุ่มเรียก ตัวอย่างสุ่ม (Random Sample) ค่าที่วัดได้จากตัวอย่างสุ่มนี้เรียกว่า <u>ค่าสถิติ</u> (Statistic)

สัญลักษณ์ (Notation)

N	จำนวนข้อมูลในประชากร
n	จำนวนข้อมูลในตัวอย่าง
X_i	ค่าของข้อมูล (ค่าสังเกต) แต่ละค่า
μ	ค่าเฉลี่ยของประชากร
$\bar{X}$	ค่าเฉลี่ยของตัวอย่าง
σ^2	ความแปรปรวนของประชากร
S^2	ความแปรปรวนของตัวอย่าง
Σ	ผลรวม

การจัดและอธิบายข้อมูล

ข้อมูลสถิติที่รวบรวมได้ ถ้ายังไม่ได้มีการเปลี่ยนแปลงเพื่อความประสงค์อย่างหนึ่งอย่างใด เรียกว่า ข้อมูลดิบ (Raw Data) ข้อมูลดิบนี้ยังใช้ประโยชน์ไม่ได้มากนักควรจะได้มีการจัดข้อมูลให้อยู่ในรูปที่เข้าใจง่าย เช่น การจัดลำดับและการแจกแจงความถี่

การจัดลำดับ (Arrays) คือ การจัดเรียงค่าสังเกตจากค่ามากไปหาน้อยหรือจากค่าน้อยไปหามาก การจัดเรียงนี้จะทำให้เห็นลักษณะของข้อมูลได้ดีขึ้น แม้จะยังไม่ได้ทำการวิเคราะห์ผล ก็พอจะมองเห็นว่าค่าใดเป็นค่าสูงสุด ค่าใดเป็นค่าต่ำสุด นอกจากนี้ยังพอจะเห็นการกระจายของข้อมูลชุดนี้ว่ามีมากน้อยเพียงใด

<u>ตัวอย่างที่ 1</u> คะแนนสอบวิชาสถิติเบื้องต้นของนักเรียน 60 คน (คะแนนเต็ม 200) เป็นดังนี้

161	88	149	198	95	64	134	137	102	123
116	87	101	46	176	109	88	102	70	109
117	174	118	96	76	99	151	77	102	52
70	145	83	80	77	44	71	98	131	128
145	74	76	153	38	76	186	137	171	128
154	53	90	151	107	131	138	97	79	158

จากการจัดลำดับคะแนนจากน้อยไปหามาก จะได้

38	70	76	87	97	102	118	134	149	161
44	70	77	88	98	107	123	137	151	171
46	71	77	88	99	109	128	137	151	174
52	74	79	90	101	109	128	138	153	176
53	76	80	95	102	116	131	145	154	186
64	76	83	96	102	117	131	145	158	198

ทำให้เห็นได้โดยง่ายว่าคะแนนต่ำสุดที่นักเรียนทำได้ คือ 38 คะแนน ขณะที่คะแนนสูงสุดเป็น 198 คะแนน

การแจกแจงความถี่ (Frequency Distribution) คือ การจัดแบ่งค่าสังเกตออกเป็นพวกหรือชั้น (Class) และนับจำนวนค่าสังเกตที่ตกอยู่ในแต่ละชั้น การแจกแจงความถี่นี้จะแสดงด้วยตารางหรือกราฟก็ได้ เช่น

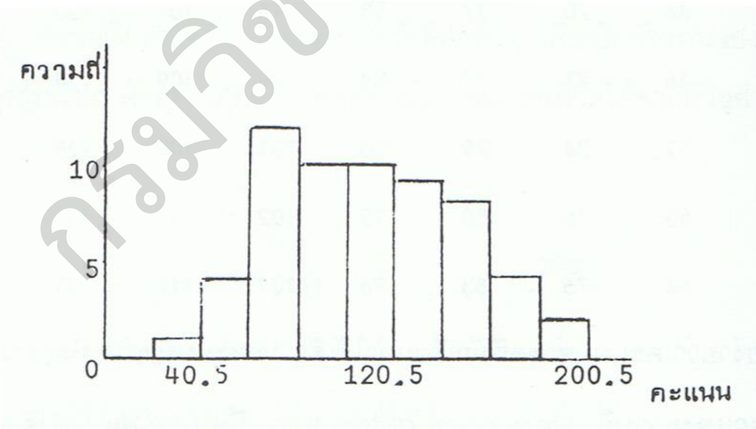
1. **ตารางแจกแจงความถี่** เป็นการจัดกลุ่มและนำเสนอในรูปตาราง

ตารางที่ 1 แสดงคะแนนสอบวิชาสถิติของนักเรียน จากคะแนนเต็ม 200 คะแนน

คะแนน	ความถี่
20 – 40	1
41 – 60	4
61 – 80	12
81 – 100	10
101 – 120	10
121 – 140	9
141 – 160	8
161 – 180	4
181 – 200	2
รวม	60

2. **ฮิสโตแกรม** (Histogram) เป็นการแสดงด้วยรูปสี่เหลี่ยมผืนผ้าหลายรูปต่อเนื่องกัน และให้พื้นที่ของสี่เหลี่ยมผืนผ้าแต่ละรูปแทนความถี่ของแต่ละช่วง

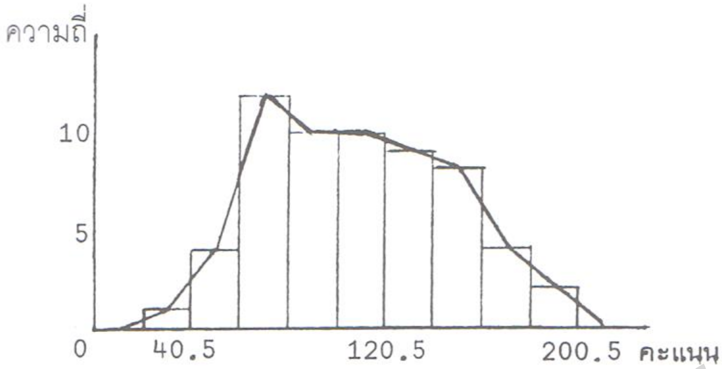
รูปที่ 1 ฮิสโตแกรมแสดงคะแนนสอบวิชาสถิติของนักเรียน



3. **รูปหลายแหลี่ยมแห่งความถี่** (Polygon) เป็นการลากเส้นตรงเชื่อมจุดกึ่งกลางยอดของแต่ละรูปสี่เหลี่ยมผืนผ้าในฮิสโตแกรม

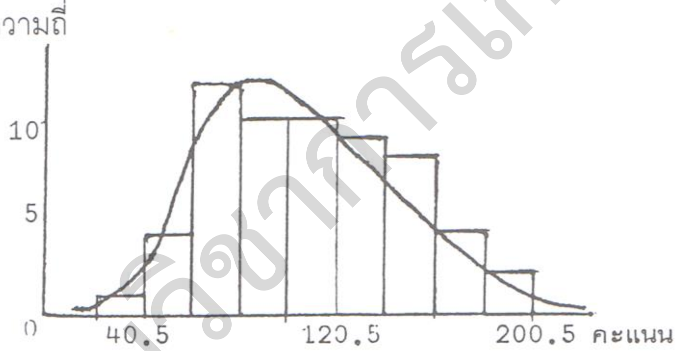
<u>รูปที่ 2</u> รูปหลายเหลี่ยมแห่งความถี่ แสดงคะแนนสอบวิชาสถิติของนักเรียน

อาจแยกชนิดได้ตามรูปร่าง



4. **เส้นโค้งของความถี่** (Frequency Curve) เป็นการปรับเส้นตรงจากรูปหลายเหลี่ยมแห่งความถี่ให้เป็นเส้นโค้ง เส้นโค้งนี้จะแสดงลักษณะการแจกแจง (Distribution) ของข้อมูล

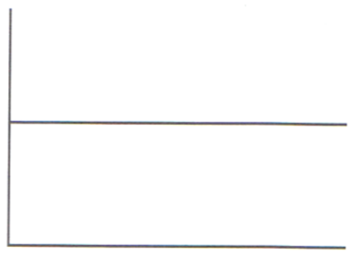
<u>รูปที่ 3</u> เส้นโค้งของความถี่ แสดงคะแนนสอบวิชาสถิติ



การแจกแจงความถี่มีหลายชนิด อาจแยกชนิดได้ตามรูปร่างของกราฟที่ได้ ที่พบมาก มีดังนี้

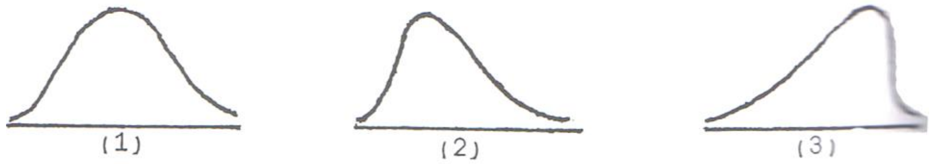
1. รูปสี่เหลี่ยมผืนผ้า ถ้าความถี่ของแต่ละชั้นมีค่าเท่ากัน

<u>รูปที่ 4</u> ตัวอย่างการแจกแจงรูปสี่เหลี่ยมผืนผ้า



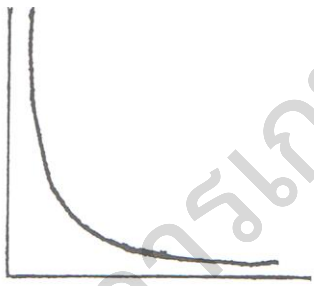
2. รูประฆัง ถ้าความถี่ของคะแนนตอนกลาง ๆ มากกว่าคะแนนหัวท้าย การแจกแจงรูประฆังนี้มี 3 แบบ

<u>รูปที่ 5</u> ตัวอย่างเส้นโค้งรูประฆัง (1) เส้นโค้งปกติ (2) เส้นโค้งเบ้ทางขวา (3) เส้นโค้งเบ้ทางซ้าย



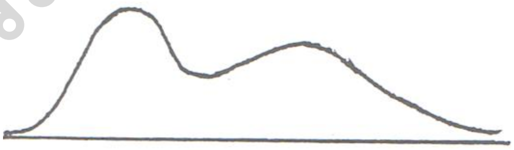
3. รูปตัว J กลับ ถ้าความถี่สูงของคะแนนน้อยๆเท่านั้น

<u>รูปที่ 6</u> ตัวอย่างเส้นโค้งรูปตัว J กลับ



4. รูป Bimodal เป็นการแจกแจงคล้ายการแจกแจงรูประฆัง แต่มียอดระฆังสองยอด มักพบมากในกรณีที่มีข้อมูล 2 ชุดผสมกัน

<u>รูปที่ 7</u> ตัวอย่างเส้นโค้งรูป Bimodal



การนำเสนอข้อมูลในรูปตารางหรือกราฟ เพื่อใช้ในการสรุปและแสดงปริมาณของข้อมูล เป็นการอธิบายที่ยังไม่ละเอียดพอ จำเป็นจะต้องใช้การคำนวณเข้ามาช่วยเพื่อหาค่า ๆ หนึ่งหรือมากกว่าสำหรับใช้เป็นสถิติหรืออธิบายลักษณะของข้อมูลให้ละเอียดยิ่งขึ้นอีก การวัดมูลฐานนั้นได้แก่ การวัดแนวโน้มเข้าสู่ส่วนกลาง (Measures of central tendency) และการวัดการกระจาย (Measures of dispersion or variation)

การวัดแนวโน้มเข้าสู่ส่วนกลาง เป็นวิธีการที่จะแสดงให้เห็นถึงศูนย์กลางของลักษณะนั้น ๆ การวัดแนวโน้มเข้าสู่ส่วนกลางในวิชาสถิติมีหลายแบบ แบบใหญ่ ๆ มีดังนี้

1. **ฐานนิยม** (Mode) เป็นวิธีง่ายที่สุดในการชี้ให้เห็นถึงลักษณะการกระจายของข้อมูล ฐานนิยมจะเป็นค่าข้อมูลที่เกิดบ่อยที่สุด (มีความถี่สูงสุด) ในข้อมูลชุดหนึ่ง ๆ อาจจะไม่มีฐานนิยมเลยก็ได้ ถ้าค่าของข้อมูลในชุดนั้น ๆ แตกต่างกันทุกตัว หรืออาจจะมีฐานนิยมมากกว่า 1 ค่า ถ้ามีข้อมูลที่ซ้ำกันในจำนวนที่เท่ากัน

อย่างไรก็ตามฐานนิยมไม่ได้ช่วยทำให้เห็นภาพพจน์เกี่ยวกับลักษณะของข้อมูลมากนัก เพราะบอกได้แต่เพียงค่าของข้อมูลที่มีความถี่มากที่สุดแต่ไม่ได้บอกถึงตำแหน่งและความสัมพันธ์กับข้อมูลตัวอื่น ๆ

2. **มัธยฐาน** (Median) คือ ค่าของข้อมูลที่อยู่ตรงกลางของกลุ่มข้อมูลครึ่งหนึ่ง (50%) ของจำนวนข้อมูลมีค่าสูงกว่าค่ามัธยฐาน และอีกครึ่งหนึ่งมีค่าต่ำกว่า ในกรณีที่จำนวนข้อมูลเป็นเลขคู่แล้ว มัธยฐาน จะเท่ากับค่าเฉลี่ยของข้อมูล 2 ค่าที่อยู่ตรงกลาง

ค่ามัธยฐานอธิบายลักษณะการกระจายของกลุ่มข้อมูลได้ดีกว่าฐานนิยมเล็กน้อยกล่าวคือ มัธยฐานบอกเราถึงค่าข้อมูลที่อยู่กึ่งกลาง แต่ยังคงไม่ได้พิจารณาค่าข้อมูลที่อยู่สูงกว่าหรือต่ำกว่า

3. **ค่าเฉลี่ย** (Mean หรือ Average : μ, $\bar{X}$) หรือค่าเฉลี่ยเลขคณิตได้จากการรวมทุกค่าของข้อมูลแล้วหารด้วยจำนวนข้อมูลทั้งหมด

สูตรที่ใช้คำนวณหา Mean :

ของประชากร : $$\mu = \sum_{i=1}^{N} X_i / N$$

ของตัวอย่าง : $$\bar{X} = \sum_{i=1}^{n} X_i / n$$

จะเห็นได้ว่าค่าเฉลี่ยคำนวณมาจากค่าของข้อมูลทุกตัว ขณะที่ฐานนิยมพิจารณาจากความถี่และมัธยฐานพิจารณาจากตำแหน่งเท่านั้น โดยปกติค่าเฉลี่ยจะเป็นตัวแทนของการวัดแนวโน้มเข้าสู่ส่วนกลางที่ดีที่สุด

คุณสมบัติบางประการของค่าตัวกลาง

1. การกระจายของข้อมูลเท่ากันทั้งสองข้าง (Symmetry) ค่าเฉลี่ยมัธยฐานและฐานนิยมจะเท่ากัน

2. ถ้าการกระจายของข้อมูลทั้งสองข้างไม่เท่ากัน คือ เบ้ไปข้างใดข้างหนึ่ง (Skew) ค่าเฉลี่ย มัธยฐาน และฐานนิยมจะแตกต่างกัน ค่าเฉลี่ยจะถูกกระทบกระเทือน จากข้อมูลที่มีค่าสูงสุดหรือต่ำสุด ในกรณีนี้มัธยฐานหรือฐานนิยมจะเป็นตัวแทนที่ดีกว่าค่าเฉลี่ย

3. เมื่อต้องการหาว่าข้อมูลชุดนั้นมีการกระจายอย่างไร จะต้องหาค่าเฉลี่ย เพราะค่าเฉลี่ยจะเป็นค่าหนึ่งที่บอกลักษณะของการกระจาย (อีกค่าคือความแปรปรวน)

4. ผลรวมของผลต่างระหว่างค่าของข้อมูลแต่ละตัวกับค่าเฉลี่ยแล้วยกกำลังสอง (Sum of square of deviation from mean : $\sum\left(X_i - \overline{X}\right)^2$ จะมีค่าน้อยที่สุด ถ้าใช้ตัวกลางอื่น ๆ แทนจะได้ค่ามากกว่านี้

5. ถ้าเลือกตัวอย่างขนาดใหญ่หลาย ๆ ชุด จากประชากรชุดเดียวกัน ค่าเฉลี่ยค่อนข้างจะคงที่กว่ามัธยฐานและฐานนิยม ฉะนั้นในทางทฤษฎีจึงนิยมใช้ค่าเฉลี่ยเป็นค่าตัวกลาง

<u>ตัวอย่างที่ 2</u> ข้อมูลชุดหนึ่งประกอบด้วยค่าสังเกต จำนวน 11 ค่า ดังนี้

11 14 17 15 17 17 19 10 13 16 14

จากการเรียงลำดับข้อมูลจะได้

10 11 13 14 14 15 16 17 17 17 19

และการวัดแนวโน้มเข้าสู่ส่วนกลางทั้ง 3 แบบได้ค่าดังนี้

Mode = 17

Median = 15

$$\text{Mean} = \frac{10 + 11 + 13 + \ldots + 19}{11} = 14.8$$

การวัดการกระจาย ค่าเฉลี่ยเพียงอย่างเดียว ไม่พอเพียงสำหรับอธิบายลักษณะเฉพาะของข้อมูลชุดนั้น ๆ เพราะการแจกแจงความถี่ของข้อมูลชุดต่าง ๆ อาจจะมีค่าเฉลี่ยเท่ากัน แต่รูปของการแจกแจงความถี่ต่างกันมาก เช่น

<u>ตัวอย่างที่ 3</u>	ข้อมูลชุดที่ 1	7	25	58	74	86	ค่าเฉลี่ย = 50
	ข้อมูลชุดที่ 2	47	48	49	51	55	ค่าเฉลี่ย = 50

ข้อมูลทั้งสองชุดมีค่าเฉลี่ยเท่ากันคือ 50 แต่ข้อมูลชุดที่ 1 ประกอบด้วยค่าน้อยและค่ามากคละกันไป ขณะที่ชุดที่ 2 ข้อมูลมีค่าไล่เลี่ยกัน ถ้าดูแต่ค่าเฉลี่ยแต่เพียงลำพังแล้ว อาจจะทำให้เข้าใจผิดได้ว่าข้อมูลทั้งสองชุดมีลักษณะคล้ายกัน จึงนับว่าค่าเฉลี่ยแต่เพียงอย่างเดียวให้ประโยชน์ไม่มากนัก

เพียงให้ทราบว่าข้อมูลแต่ละชุดมีการกระจายมากน้อยเพียงใด จึงมีการวัดการกระจายของข้อมูลขึ้น หากข้อมูลชุดใดมีการกระจายมากแสดงว่าค่าของข้อมูลมีใหญ่และเล็กคละกันมาก แต่ถ้ามีการกระจายน้อยก็แสดงว่าข้อมูลในชุดนั้นมีค่าไล่เลี่ยกัน

การวัดการกระจายของข้อมูลที่สำคัญมี 4 แบบ คือ

1. **พิสัย** (Range) คือ ผลต่างของค่าสูงสุดกับต่ำสุด จากตัวอย่างที่ 3 พิสัยของข้อมูลชุดที่หนึ่งเท่ากับ 86 – 7 = 79

พิสัยเป็นการวัดการกระจายอย่างคร่าว ๆ เท่านั้น ข้อเสียของพิสัยก็คือ ไม่ได้ใช้ข้อมูลทั้งหมด แต่ใช้เพียงค่าสูงสุดกับต่ำสุดเท่านั้น

2. **ส่วนเบี่ยงเบนเฉลี่ย** (Average deviation หรือ Mean absolute deviation : M.D.) ในการวัดส่วนเบี่ยงเบนเฉลี่ยนี้ ข้อมูลทุกตัวจะถูกนำมาคำนวณด้วย ถ้าข้อมูลมีการกระจายมากส่วนเบี่ยงเบนเฉลี่ยจะมีขนาดใหญ่ แต่ถ้าข้อมูลมีการกระจายน้อยส่วนเบี่ยงเบนเฉลี่ยจะมีขนาดเล็ก ส่วนเบี่ยงเบนเฉลี่ยหาได้จากสูตร

$$\text{M.D.} = \frac{|X_1 - \bar{X}| + |X_2 - \bar{X}| + \ldots + |X_N - \bar{X}|}{N}$$

$$= \frac{\sum_{i=1}^{N} |X_i - \bar{X}|}{N}$$

เมื่อ $|\ \ |$ = absolute value = ค่าของตัวเลขที่ตัดเครื่องหมายทิ้ง

<u>ตัวอย่างที่ 4</u> จากตัวอย่างที่ 2 ข้อมูลชุดที่ 1 ประกอบด้วยค่าสังเกต 5 ค่า ดังนี้

7 25 58 74 และ 86

ค่าเฉลี่ย ($\bar{X}$) $= \dfrac{7 + 25 + 58 + 74 + 86}{5} = 50$

ค่า mean deviation

$$\text{M.D.} = \frac{|7 - 50| + |25 - 50| + \ldots + |86 - 50|}{5}$$

$$= \frac{136}{5} = 27.2$$

3. **ส่วนเบี่ยงเบนมาตรฐาน** (Standard deviation : σ, S) เป็นการวัดการกระจายโดยใช้ผลรวมของกำลังสองของส่วนเบี่ยงเบนจากค่าเฉลี่ย แล้วหาราก (Root) ที่ 2

สูตรหา Standard deviation :

ของประชากร : $\sigma = \sqrt{\dfrac{\sum_{i=1}^{N} (X_i - \mu)^2}{N}}$

ของตัวอย่าง : $S = \sqrt{\dfrac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n - 1}}$

มีอีกสูตรที่ใช้ในการคำนวณ (Working formula) คือ

$$S = \sqrt{\frac{\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2/n}{n-1}}$$

กำลังสองของส่วนเบี่ยงเบนมาตรฐานเรียกว่า ความแปรปรวน (Variance : σ^2, S^2)

โดยการใช้ส่วนเบี่ยงเบนมาตรฐาน จะสามารถแก้ปัญหาเกี่ยวกับการตัดเครื่องหมายลบทิ้งโดยใส่เครื่องหมาย | | เพราะการยกกำลังสองของส่วนเบี่ยงเบนแต่ละค่าจะได้เครื่องหมายเป็นบวกทั้งหมด นอกจากนี้ส่วนเบี่ยงเบนมาตรฐานยังใช้ในสมการใช้วิชาสถิติขั้นสูงด้วย จึงนับว่าส่วนเบี่ยงเบนมาตรฐานเป็นมาตราวัดการกระจายที่ดีที่สุดในจำนวนมาตราวัดการกระจายทั้งหมด

<u>ตัวอย่างที่ 5</u> ค่าความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานจากข้อมูลในตัวอย่างที่ 4 คือ

ความแปรปรวน (Variance) :

$$S^2 = \frac{(7-50)^2 + (25-50)^2 + \ldots + (86-50)^2}{(5-1)}$$

$$= 1{,}102.50$$

และส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)

$$S = \sqrt{1{,}102.50} = 33.20$$

4. **สัมประสิทธิ์ความแปรปรวน** (Coefficient of variation : C.V.) เป็นค่าสถิติที่แสดงการกระจายของข้อมูลแบบ Relative dispersion นั่นคือ มีหน่วยในรูปร้อยละ เนื่องจาก C.V. ไม่ขึ้นกับหน่วยที่วัด จึงใช้เปรียบเทียบการกระจายของข้อมูล 2 ชุด ได้โดยไม่จำเป็นต้องเป็นหน่วยเดียวกัน เมื่อพอจะทราบค่า C.V.จากประสบการณ์แล้วเราจะใช้ประกอบการคำนวณหาขนาดตัวอย่างที่เหมาะสมได้นอกจากนี้ ค่า C.V. ยังใช้ประเมินค่าความถูกต้องของค่าประมาณหรือ Treatment effect อีกด้วย ค่า C.V. คำนวณได้โดย

$$C.V. = \frac{S}{\bar{X}} \times 100\%$$

<u>ตัวอย่างที่ 6</u> จากข้อมูลชุดที่ 2 ในตัวอย่างที่ 3 ประกอบด้วยค่าสังเกต 5 ค่า ดังนี้

47 48 49 50 51 และ 55

$\bar{X}$ = 50

S^2 = 10 S = 3.1623

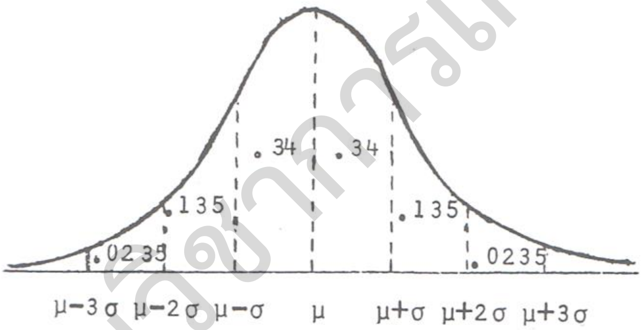
และ C.V. $= \frac{3.1623}{50} \times 100 = 6.3\%$

การแจกแจงแบบปกติ (Normal distribution)

การแจกแจงแบบปกติ เป็นการแจกแจงที่สำคัญที่สุดในการศึกษาสถิติ พบในตัวแปรส่วนใหญ่ประชากรที่มีการแจกแจงแบบปกติ เรียกว่า ประชากรปกติ (Normal population) ซึ่งมีตัวแปรเป็นตัวแปรต่อเนื่อง (Continuous variable) มีค่าระหว่าง $-\infty$ ถึง ∞ และเส้นโค้งแสดงการแจกแจงแบบปกติ เรียกโค้งปกติ ซึ่งมีลักษณะ ดังนี้

1. เป็นรูประฆังคว่ำ เส้นโค้งทั้งสองข้างเหมือนกัน (Symmetric) มีความโด่งพอเหมาะ คือมีความโด่งแบบ Meso kurtic
2. ข้อมูลที่มีการกระจายแบบปกติ จะมีค่า Mean Median และ Mode เท่ากัน และอยู่ตรงกลาง
3. พื้นที่ภายใต้โค้งทั้งหมด หมายถึง Probability ของเหตุการณ์ทั้งหมด (Sample space) ซึ่งมีค่าเท่ากับ 1 หรือ 100%

รูปที่ 8 การแจกแจงแบบปกติ และ Probability ที่สัมพันธ์กัน



การแจกแจงแบบ t (t-distribution)

การแจกแจงแบบ t เป็นการแจกแจงที่สำคัญอีกอย่างหนึ่งในวิชาสถิติ สามารถใช้แทนการแจกแจงปกติในกรณีที่ไม่ทราบค่าความแปรปรวน ค่าสถิติ t คำนวณได้โดย

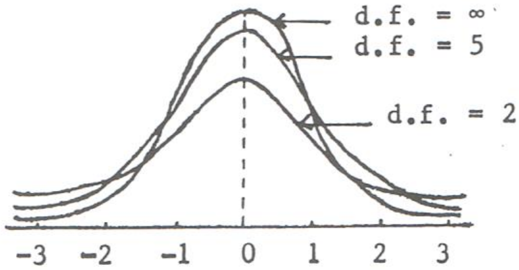
$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} \quad \text{หรือ} \quad \frac{\bar{X} - \mu}{S_{\bar{X}}}$$

คุณสมบัติของการแจกแจงแบบ t มีดังนี้

1. เส้นโค้งเป็นรูประฆังคว่ำ มีลักษณะสมมาตรรอบจุดศูนย์กลาง และมีความโด่งแบบ Platy kurtic ซึ่งจะมากหรือน้อยขึ้นอยู่กับ Degree of Freedom (d.f.)
2. มีค่าเฉลี่ยอยู่ที่จุดศูนย์
3. ตัวแปร t เป็นตัวแปรลักษณะต่อเนื่อง มีค่าตั้งแต่ $-\infty$ ถึง ∞

4. พื้นที่ภายใต้โค้งทั้งหมดหมายถึง Probability ของเหตุการณ์ทั้งหมด ซึ่งมีค่าเท่ากับ 1 หรือ 100 %
5. เส้นโค้งการแจกแจงแบบ t จะใกล้เคียงกับการแจกแจงแบบปกติเมื่อค่า(df) เข้าใกล้ ∞

<u>รูปที่ 9</u> การแจกแจงแบบ t



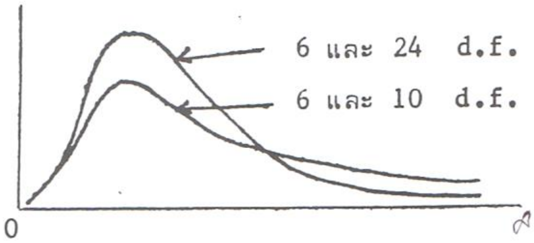
การแจกแจงแบบ F (F-distribution)

ถ้า S^2_1 และ S^2_2 เป็นค่าความแปรปรวนของตัวอย่างขนาด n_1 และ n_2 ซึ่งต่างถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ และตัวอย่างทั้งสองชุดเป็นอิสระต่อกันแล้ว $(S^2_1 / \sigma^2_1)/(S^2_2 / \sigma^2_2)$ จะมีการแจกแจงแบบ F ซึ่งมี Degree of Freedom $(n_1 - 1)$ และ $(n_2 - 1)$

คุณสมบัติของการแจกแจงแบบ F มีดังนี้

1. เส้นโค้งเบ้ ในลักษณะเบ้ขวา และทุกขณะที่ขนาดตัวอย่างเพิ่มขึ้น เส้นโค้ง F จะเบ้น้อยลง
2. ตัวแปร F เป็นตัวแปรลักษณะต่อเนื่อง มีค่าระหว่าง 0 ถึง ∞
3. พื้นที่ภายใต้โค้ง F หมายถึง Probability ของเหตุการณ์ทั้งหมด ซึ่งมีค่าเท่ากับ 1 หรือ 100 %

<u>รูปที่ 10</u> การแจกแจงแบบ F



การแจกแจงแบบไคสแควร์ (χ^2 - Distribution)

ไคสแควร์ (Chi – Square หรือ χ^2 ก็คือกำลังสองของ Standard normal (Z)

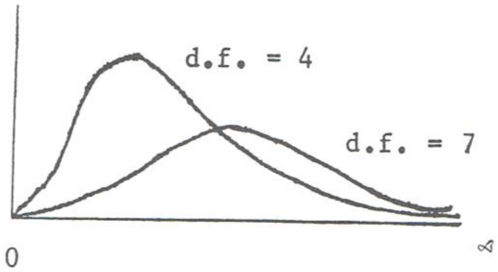
คุณสมบัติของการแจกแจงแบบ χ^2 มีดังนี้

1. เป็นโค้งเบ้ในลักษณะเบ้ขวา n มีขนาดใหญ่ขึ้น โค้ง χ^2 จะเบ้น้อยลง

2. ตัวแปร χ^2 เป็นตัวแปรลักษณะต่อเนื่อง มีค่าระหว่าง 0 ถึง ∞

3. พื้นที่ภายใต้โค้ง χ^2 ทั้งหมดหมายถึง Probability ของเหตุการณ์ทั้งหมด ซึ่งมีค่าเท่ากับ 1 หรือ 100 %

<u>รูปที่ 11</u> การแจกแจงแบบ χ^2



การแจกแจงตัวอย่าง (Sampling distribution)

ตามจำกัดความ Sampling distribution หมายถึงการแจกแจงโอกาสที่ค่าสถิติ (Statistic) จะเป็นไปได้ทั้งหมดถ้าเลือกตัวอย่างทุกชุด (ที่จะเป็นไปได้ทั้งหมด) ที่มีขนาดคงที่จาก ประชากร(Population) ที่กำหนดให้

ตัวอย่าง เช่น ถ้า Population มีขนาด N=3 ประกอบด้วย 1,2,3 และเราเลือก Sample ขนาด n=2 แบบใส่คืนแล้วมีโอกาสที่จะได้ตัวอย่างแบบต่างๆ ถึง 9 ชุด ดังนี้

ชุดที่	ตัวอย่าง	ค่าสถิติ	
		$\overline{X}$	S^2
1	1,1	1	0
2	1,2	1.5	0.5
3	1,3	2	2
4	2,1	1.5	0.5
5	2,2	2	0
6	2,3	2.5	0.5
7	3,1	2	2
8	3,2	2.5	0.5
9	3,3	<u>3</u>	<u>0</u>
	รวม	<u>18</u>	<u>6</u>

ค่า Parameter :

$$\mu = \frac{1+2+3}{3} = 2$$

$$\sigma^2 = \frac{(1+2)^2 + (2+2)^2 + (3 \quad 2)^2}{3} = \frac{2}{3}$$

จะเห็นได้ว่า 1. $\bar{X}$ ทุกค่าไม่เท่ากับ μ (มีบางค่าของ $\bar{X}$ เท่านั้นที่เท่ากับ μ)

แต่ $\bar{X} = \mu$

$$\text{เมื่อ} \quad \bar{X} = \frac{\sum X^2}{\text{จำนวนตัวอย่างที่เป็นไปได้}} = \frac{18}{9} = 2$$

2. $S^2 \neq \sigma^2$ แต่ = $\bar{S}^2 = \sigma^2$

$$\text{เมื่อ} \quad \bar{S}^2 = \frac{\sum S^2}{\text{จำนวนตัวอย่างที่เป็นไปได้}} = \frac{6}{9} = \frac{2}{3}$$

เมื่อ $\bar{X} \neq \mu$ ค่าสถิติที่วัดความแตกต่างระหว่างค่า $\bar{X}$ กับ μ ก็คือ $\sigma_{\bar{X}}$ ที่เรียกว่า standard error of sample means ดังนั้น $\sigma_{\bar{X}}$ จึงคำนวณจาก

$$\sigma_{\bar{X}} = \sqrt{\frac{\Sigma(\bar{X} - \mu)^2}{\text{จำนวนตัวอย่างที่เป็นไปได้}}}$$

$$\text{จากตัวอย่าง} \quad \sigma_{\bar{X}} = \sqrt{\frac{(1-2)^2 + (1.5-2)^2 + \ldots + (3-2)^2}{9}} = \sqrt{\frac{3}{9}} = .5774$$

แต่ในทางปฏิบัติเราจะเลือกตัวอย่างมาเพียง 1 ชุด เท่านั้น และ $\sigma_{\bar{X}}$ จะคำนวณจากสูตร

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

$$\text{จากตัวอย่าง} \quad \sigma_{\bar{X}} = \sqrt{\frac{2}{3 \times 2}} = .5774$$

โดยทั่วไปแล้วเราใช้ $S_{\bar{X}}$ เป็นตัวสถิติที่ใช้ประมาณ $\sigma_{\bar{X}}$ ของประชากร โดยค่า $S_{\bar{X}}$ คำนวณจาก

$$S_{\bar{X}} = \frac{S}{\sqrt{n}}$$

การประมาณค่าและการทดสอบ

ในการวิเคราะห์ข้อมูล ปกติจะไม่วิเคราะห์ข้อมูลประชากร (Population) เนื่องจากจำนวนประชากรมักจะมีจำนวนมาก เราต้องสิ้นเปลืองเงินและเวลา ดังนั้น จึงมักจะวิเคราะห์ข้อมูลตัวอย่าง (Sample) ซึ่งเป็นส่วนหนึ่งของประชากรแล้วนำผลวิเคราะห์นั้นมาอ้างอิงว่าเป็นคุณสมบัติของประชากรค่าวิเคราะห์ต่าง ๆ ทางสถิติจากตัวอย่างนี้เรียกว่า ค่าสถิติ (Statistic) ซึ่งโดยขบวนการทางสถิติอย่างเดียวกันถ้าวิเคราะห์โดยใช้ข้อมูลของประชากรค่าวิเคราะห์เหล่านี้เรียกว่าพารามิเตอร์ (Parameter)

วิธีการที่จะใช้อ้างอิงถึงประชากร โดยอาศัยผลลัพธ์ซึ่งได้จากตัวอย่างที่สุ่มจากประชากรนั้น ๆ เรียกว่า สถิติเชิงอ้างอิง (Inference statistics) ซึ่งเกี่ยวข้องกับปัญหาในสองด้าน คือ การประมาณค่าและการทดสอบสมมุติฐาน

การประมาณค่า (Estimation) เป็นการนำค่าสถิติไปใช้ประมาณค่าพารามิเตอร์ที่เกี่ยวข้องกัน การประมาณค่ามี 2 ชนิด ได้แก่

1. **การประมาณค่าเดี่ยว** (Point estimation) ได้แก่ การใช้ตัวเลขเพียงค่าเดียวในการประมาณพารามิเตอร์ของประชากรที่สอดคล้องกัน เช่น ใช้ค่าสถิติ $\bar{X}$ ประมาณค่า μ เป็นต้น

2. **การประมาณแบบช่วง** (Interval estimation) เป็นการประมาณพารามิเตอร์โดยใช้ค่าสองค่าบอกช่วงการประมาณ ซึ่งในการประมาณแบบนี้ต้องบอกระดับความเชื่อมั่น ในการประมาณด้วย เช่น Prob. $(a< \mu <b) = .95$ ซึ่งหมายถึงค่า μ ที่ถูกประมาณโดยช่วง a และ b ด้วยระดับความเชื่อมั่น 95%เมื่อ a และ b คือตัวเลขใดๆ ช่วงระหว่าง a และ b นี้เรียกว่าช่วงความเชื่อมั่น (Confidence interval : C.I.)

การทดสอบสมมุติฐาน (Test of hypothesis) เป็นระเบียบวิธีทางสถิติที่จะช่วยในการตัดสินใจว่าค่าสถิติที่ได้มานั้นสอดคล้องกับค่าที่แท้จริงหรือพารามิเตอร์หรือไม่เพียงใด

ก. ข้อสมมุติฐาน (Hypothesis) หมายถึงข้อที่เราสมมุติขึ้นเพื่อทดสอบว่าสิ่งที่เราพิจารณาอยู่นั้น เป็นไปตามที่คาดไว้หรือไม่ ข้อสมมุติฐานมี 2 ชนิด คือ

1. Simple hypothesis ได้แก่ ข้อสมมุติฐานที่กำหนดค่าแน่นอนตายตัวเพียงค่าเดียว เช่น $\mu = 4$

2. Composite hypothesis ได้แก่ ข้อสมมุติฐานที่กำหนดค่าที่คาดว่าจะเป็นนั้นเป็นช่วง เช่น $\mu > 4$ หรือ $3 \leq \mu \leq 6$

ในการทดสอบสมมุติฐาน จะต้องกำหนดข้อสมมุติฐานก่อนทำการทดสอบ ซึ่งจะประกอบด้วย

1. **Null hypothesis** (H_o) ได้แก่ ข้อสมมุติฐานที่ตั้งขึ้นเพื่อจะบอกค่าของพารามิเตอร์ที่ใช้ในการทดลองเพื่อใช้เป็นหลักในการปฏิเสธ (Reject) H_o

2. **Alternative hypothesis** (H_a) ได้แก่ ข้อมูลสมมุติฐานที่ผู้วิเคราะห์กำหนดขึ้นตามสิ่งที่เขาคาดว่าน่าจะเป็น

ข. ประเภทของความผิดพลาด (Types of error) ในการใช้ค่าสถิติไปสรุปตีความเกี่ยวกับพารามิเตอร์นั้น อาจจะมีความผิดพลาดเกิดขึ้นได้ ความผิดพลาดนี้เป็นไปได้ 2 แบบ คือ

1. Type I error แทนด้วย α คือ ความน่าจะเป็นในการปฏิเสธ H_o ทั้ง ๆ ที่ H_o เป็นจริง

2. Type II error แทนด้วย β คือความน่าจะเป็นในการยอมรับ H_o ทั้ง ๆ ที่ H_o เท็จ

ในการทดสอบสมมุติฐานระดับนัยสำคัญที่กำหนดก็คือ ค่า α นั่นเอง ซึ่งจะบ่งให้ทราบถึงขนาดของขอบเขตปฏิเสธ (Rejection area) โดยทั่วไปแล้วทางด้านการเกษตรมักจะกำหนดให้ $\alpha = .05$ หรือ .01 ค่า $(1 - \alpha)$ เรียกว่า เรียกว่า **ระดับความเชื่อมั่น**

ค. การทดสอบแบบหางเดียวและสองหาง จากการที่ Alternative hypothesis (H_a) ถูกตั้งในรูปของ Composite hypothesis ซึ่งกำหนดค่าเป็นช่วงของพารามิเตอร์ที่ต้องการทดสอบนั้น รูปของการทดสอบแบ่งออกเป็น 2 แบบ คือ

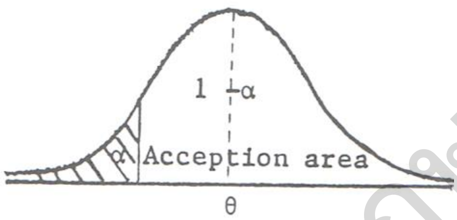
1. **การทดสอบแบบหางเดียว** (One-tailed test) เมื่อกำหนดให้ H_a มีช่วงอยู่สูงกว่าหรือต่ำกว่าค่าพารามิเตอร์ มี 2 แบบ ได้แก่

การทดสอบแบบหางเดียวข้างมาก ข้อสมมุติฐาน คือ

H_o : $\mu \leq \theta$ VS H_a : $\mu > \theta$

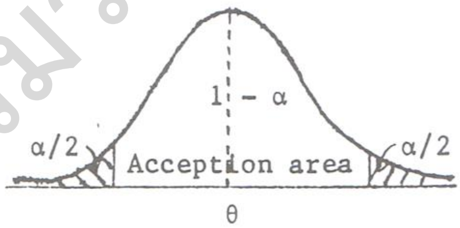
การทดสอบแบบหางเดียวข้างน้อย ข้อสมมุติฐาน คือ

H_o : $\mu \geq \theta$ VS H_a : $\mu < \theta$



2. **การทดสอบแบบสองหาง** (Two-tailed test) เมื่อกำหนดให้ H_a ช่วงของพารามิเตอร์อยู่ทั้งสองหาง ข้อสมมุติฐาน คือ

H_o : $\mu = \theta$ VS H_a : $\mu \neq \theta$



ง. หลักการทดสอบสมมุติฐาน การทดสอบสมมุติฐานมีหลักดำเนินการเป็นขั้นๆ ดังต่อไปนี้

1. กำหนดสมมุติฐาน Null hypothesis H_o และ Alternative hypothesis H_a
2. กำหนดระดับนัยสำคัญ (α)
3. พิจารณา Critical region
4. เลือกตัวสถิติที่จะใช้ทดสอบและคำนวณหาค่าสถิติจากข้อมูล
5. พิจารณาสรุปผล

การทดสอบค่าเฉลี่ยประชากร

ถ้า X_1 , X_2 , ... , X_n เป็นตัวอย่างสุ่มซึ่งมาจากประชากรที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ และความแปรปรวน σ^2 แล้ว ต้องการจะทดสอบว่าค่าเฉลี่ยซึ่งได้มาจากตัวอย่าง ($\bar{X}$) มาจากประชากรที่สุ่มมา ซึ่งมี $\mu = \mu_0$ ($H_0 : \mu = \mu_0$) การทดสอบอาจแบ่งเป็น 2 กรณี คือ เมื่อทราบค่าความแปรปรวน (σ^2) และเมื่อไม่ทราบค่าความแปรปรวน

การทดสอบค่าเฉลี่ยเมื่อทราบค่าความแปรปรวน เมื่อทราบค่าความแปรปรวนของประชากรแล้วจะทดสอบค่าเฉลี่ยได้ โดยใช้ Z-test ซึ่งค่าสถิติ Z คำนวณได้ดังนี้

$$Z = \frac{\bar{X} - \mu_o}{\sigma_{\bar{X}}}$$

เมื่อ $\sigma_{\bar{X}} = \sigma/\sqrt{n}$

เปรียบเทียบค่า |Z| ที่คำนวณได้ กับค่า Z ที่เปิดได้จากตาราง Normal ที่ $Z_{1-\alpha/2}$ ถ้า $H_a : \mu \neq \mu_o$ ที่ $Z_{1-\alpha}$ ถ้า $H_a : \mu > \mu_o$ และ ที่ $-Z_{1-\alpha}$ ถ้า $H_a : \mu < \mu_o$ ช่วงความเชื่อมั่นคำนวณได้โดย

$$100(1-\alpha)\%\ \text{C.I.} = \bar{X} \pm Z_{\alpha/2}.\sigma_{\bar{X}}$$

การทดสอบค่าเฉลี่ยเมื่อไม่ทราบค่าความแปรปรวน เมื่อไม่ทราบค่าความแปรปรวนของประชากรจะต้องประมาณค่า σ^2 โดยใช้ S^2 แล้วทดสอบค่าเฉลี่ยโดยใช้ t-test ค่าสถิติ t คำนวณ ดังนี้

$$t = \frac{\bar{X} - \mu_o}{S_{\bar{X}}}$$

เมื่อ $S_{\bar{X}} = S/\sqrt{n}$

เปรียบเทียบค่า t ที่คำนวณได้กับค่า t ที่เปิดได้จากตาราง t ที่ $t_{(n-1,\alpha/2)}$ ถ้า $H_a : \mu \neq \mu_o$ และที่ $t_{(n-1,\alpha)}$ โดยไม่สนใจเครื่องหมาย ถ้า $H_a : \mu > \mu_o$ หรือ $\mu < \mu_o$ เมื่อ (n-1) คือ จำนวน Degree of Freedom (d.f.)

ค่า C.I. คำนวณดังนี้

$$100(1-\alpha)\%\text{C.I.} = \bar{X} \pm t_{\alpha/2,(n\ \ 1)d.f.}.S_{\bar{X}}$$

<u>ตัวอย่างที่ 7</u> เชื่อกันว่าเมื่อฉีดฮอร์โมน ก ให้กับแม่ไก่แล้ว จะทำให้น้ำหนักของไข่ไก่เพิ่มขึ้น 130 กรัม ดังนั้น จึงมีการสุ่มไข่ไก่จำนวน 30 ฟอง จากแม่ไก่ที่ได้รับการฉีดฮอร์โมน ก ชั่งน้ำหนักดูพบว่าน้ำหนักที่เพิ่มขึ้นเฉลี่ยแล้วเป็น 175 กรัมต่อฟอง และมีความเบี่ยงเบนมาตรฐานเป็น 87.5 จากการทดลองนี้พอจะเชื่อถือได้หรือไม่ว่าน้ำหนักเพิ่มขึ้นเฉลี่ย 130 กรัมจริง

$\bar{X}$ = 175 กรัม S = 87.5

$H_o : \mu \leq 130$

$H_o : \mu > 130$

α = .05

$$t = \frac{\bar{X} - \mu_o}{S/\sqrt{n}}$$

$$= \frac{175 - 130}{87.5/\sqrt{30}} = 2.82^*$$

α = .05 และ n-1 = 29 d.f.

$t_{(29,.95)}$ = 1.70

ปฏิเสธ H_o : μ = 130 เนื่องจาก t ที่คำนวณได้ >t จากตาราง แปลผลได้ว่า ยังไม่มีหลักฐานเพียงพอที่จะกล่าวว่าน้ำหนักไข่เพิ่มขึ้น เฉลี่ยแล้วน้อยกว่า 130 กรัม นั่นคือ ไข่จากแม่ไก่ที่ได้รับการฉีดฮอร์โมนมีน้ำหนักเฉลี่ยเพิ่มจากปกติมากกว่า 130 กรัม

การทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากรสองกลุ่ม

ถ้า X_{11} , X_{12} , ...X_{1n_1} เป็นตัวอย่างสุ่มจากประชากรกลุ่มที่ 1 ซึ่งมีการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ_1 และความแปรปรวน σ_1^2 แล้ว X_{21} , X_{22} , ...X_{1n_2} เป็นตัวอย่างสุ่มจากประชากรกลุ่มที่ 2 ซึ่งมีการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ_2 และความแปรปรวน σ_2^2 แล้วประชากรทั้งสองกลุ่มเป็นอิสระซึ่งกันและกันต้องการทดสอบความแตกต่างของค่าเฉลี่ยของประชากรทั้งสองกลุ่ม ($H_o : \mu_1 = \mu_2$) การทดสอบจะแบ่งเป็น 2 กรณี คือ เมื่อทราบค่าความแปรปรวน (σ^2) และเมื่อไม่ทราบค่าความแปรปรวน

การทดสอบความแตกต่างของค่าเฉลี่ยเมื่อทราบค่าความแปรปรวน เมื่อทราบค่าความแปรปรวนของประชากรแล้ว จะทดสอบความแตกต่างของค่าเฉลี่ยได้ โดยใช้ Z-test ค่าสถิติ Z คำนวณดังนี้

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}}$$

ถ้า $\sigma_1^2 = \sigma_2^2 = \sigma^2$: $\sigma_{\bar{X}_1 - \bar{X}_2} = \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

ถ้า $\sigma_1^2 \neq \sigma_2^2$: $\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

เปรียบเทียบค่า Z ที่คำนวณได้กับค่าZจากตารางNormal ที่ $Z_{\alpha/2}$ และ $Z_{1-\alpha/2}$ ถ้า $H_a : \mu_1 \neq \mu_2$ ที่ $Z_{1-\alpha}$ ถ้า $H_a : \mu_1 > \mu_2$ และ ที่ Z_{α} ถ้า $H_a : \mu_1 < \mu_2$ ค่า C.I. คำนวณได้ดังนี้

$100(1-\alpha)\%C.I. = (\bar{X}_1 - \bar{X}_2) \pm Z_{\alpha/2}.\sigma_{\bar{X}_1 - \bar{X}_2}$

การทดสอบความแตกต่างของค่าเฉลี่ยเมื่อไม่ทราบค่าความแปรปรวน เมื่อไม่ทราบความแปรปรวนของประชากรจะต้องประมาณค่า σ_1^2 และ โดนใช้ S^2_1 และ S^2_2 แล้วทดสอบความแตกต่างของค่าเฉลี่ยทั้งสองโดยใช้ t-test ค่าสถิติ t คำนวณ ดังนี้

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_{\bar{X}_1-\bar{X}_2}}$$

$S_{\bar{X}_1-\bar{X}_2}$ นี่คือ Standard error of difference between two means หรือใช้สัญลักษณ์อีกแบบหนึ่งที่ว่า $S_{\bar{d}}$

<u>กรณีที่ 1</u> ถ้าทดสอบได้ว่าความแปรปรวนของทั้งสองประชากรเท่ากัน ($\sigma_1^2 = \sigma_2^2 = \sigma^2$) จะต้องใช้ค่าประมาณความแปรปรวนร่วม (Pooled variance : S^2_p) ค่า S^2_p จะคำนวณได้โดย

$$S^2_p = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{\Sigma x_1^2 + \Sigma x_2^2}{n_1 + n_2 - 2}$$

เมื่อ $$\Sigma x_1^2 = \sum_{i=1}^{n_1} X_{1i}^2 - (\sum_{i=1}^{n_1} X_{1i})^2/n_1$$

$$\Sigma x_2^2 = \sum_{i=1}^{n_2} X_{2i}^2 - (\sum_{i=1}^{n_2} X_{2i})^2/n_2$$

n_1 และ n_2 คือจำนวนตัวอย่างจากประชากรกลุ่มที่ 1 และกลุ่มที่ 2 ตามลำดับ

$$S_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

ดังนั้น ค่าสถิติ t คำนวณได้โดย

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

เปรียบเทียบค่า t ที่คำนวณได้กับ ค่า t จากตาราง t ที่ $(n_1 + n_2 - 2)$ d.f. และ α ที่กำหนด

ค่า C.I. คำนวณดังนี้

$$100(1-\alpha)\%\text{C.I.} = (\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2,(n_1+n_2-2)} \cdot S_{\bar{X}_1-\bar{X}_2}$$

<u>กรณีที่ 2</u> ถ้าทดสอบได้ว่าความแปรปรวนของทั้งสองประชากรไม่เท่ากัน ($\sigma_1^2 \neq \sigma_2^2$) จะประมาณค่า σ_1^2 ด้วย S_1^2 และ σ_2^2 ด้วย S_2^2

ค่า $S_{\bar{X}_1-\bar{X}_2}$ คำนวณได้โดย

$$S_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

ดังนั้น ค่าสถิติ t คำนวณได้โดย

$$t = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)}} \quad \text{และ} \quad V = \frac{(S_1^2/n_1+S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1-1}+\frac{(S_2^2/n_2)^2}{n_2-1}}$$

เปรียบเทียบค่า t ที่คำนวณได้ กับค่า t จากตาราง t ที่ (ν) d.f. และ α ที่กำหนด ถ้า $n_1 = n_2$ แต่ถ้า $n_1 \neq n_2$ แล้วเทียบค่า t ที่คำนวณได้กับ t_w ค่า t_w คำนวณ ได้ดังนี้

$$t_W = \frac{W_1t_1 + W_2t_2}{W_1 + W_2}$$

เมื่อ $W_1 = S_1^2/n_1$, $W_2 = S_2^2/n_2$

t_1 = ค่า t จากตาราง $t(n_1\text{-}1)$ d.f. และ

t_2 = ค่า t จากตาราง $t(n_2\text{-}1)$ d.f. ตามระดับนัยสำคัญที่กำหนดค่า

C.I. คำนวณได้โดย

$$100(1-\alpha)\%\text{C.I.} = (\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2,(\nu)} \cdot S_{\bar{X}_1-\bar{X}_2} \quad \text{ถ้า } n_1 = n_2$$

$$100(1-\alpha)\%\text{C.I.} = (\bar{X}_1 - \bar{X}_2) \pm t_W \cdot S_{\bar{X}_1-\bar{X}_2} \quad \text{ถ้า } n_1 \neq n_2$$

การทดสอบความแตกต่างระหว่างความแปรปรวนของประชากรสองกลุ่ม

ในกรณีที่เราไม่แน่ใจว่า ค่าความแปรปรวนจากสองประชากรที่เราสุ่มมามีค่าเท่ากัน เพื่อใช้ประกอบการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของสองประชากรแล้ว จะทดสอบได้โดยใช้ F-test

ถ้าเลือกตัวอย่างขนาด n_1 จากประชากรกลุ่มที่ 1 ซึ่งมีความแปรปรวน σ_1^2 และขนาด n_2 จากประชากรกลุ่มที่ 2 ซึ่งมีความแปรปรวน σ_2^2 แล้ว ต้องการทดสอบความแตกต่างระหว่างความแปรปรวนทั้งสองค่า ($H_o: \sigma_1^2 = \sigma_2^2$) ค่า F-statistic คำนวณได้โดย

$$F = \frac{\text{large } S^2}{\text{small } S^2} = \frac{S_1^2}{S_2^2} \quad \text{เมื่อ } S_1^2 > S_2^2$$

เปรียบเทียบค่า F ที่คำนวณได้กับค่า F จากตารางแบบสองหาง (Two-tailed test) ที่ $(n_1\text{-}1)$ และ $(n_2\text{-}1)$ d.f. และระดับนัยสำคัญที่กำหนด

<u>ตัวอย่างที่ 8</u> ในการทดลองใช้ปุ๋ยเพื่อเพิ่มผลผลิตข้าว ผู้ทดลองปลูกข้าวพันธุ์ กข 23 จำนวน 56 แปลง แล้วเลือกใส่ปุ๋ย Rock phosphate อัตรา 100 กิโลกรัม ต่อไร่ จำนวน 28 แปลง เมื่อต้นข้าวอายุได้ 60 วัน สุ่มวัดความสูงจากทุกแปลง ๆ ละ 10 กอ ข้อมูลความสูง เฉลี่ยต่อกอ (เซนติเมตร) เป็นดังนี้

แปลงที่	ใส่ปุ๋ย (X_1)	ไม่ใส่ปุ๋ย (X_2)
1	63.0	66.0
2	65.3	63.7
3	69.3	61.0
4	66.7	59.7
5	67.7	64.3
6	62.3	63.3
7	71.7	60.7
8	59.3	65.7
9	61.7	56.0
10	63.0	53.0
11	70.7	64.7
12	63.0	57.0
13	65.3	53.7
14	65.3	55.0
15	71.0	65.7
16	62.0	61.0
17	60.0	57.7
18	57.3	67.3
19	67.0	68.3
20	68.3	58.7
21	57.7	61.3
22	66.0	67.7
23	65.7	63.7
24	65.7	63.3
25	66.7	63.3
26	71.3	64.3
27	63.3	65.0
28	68.7	61.7
$\bar{X}$	65.18	61.89
ΣX	1,825	1,732.8
ΣX^2	119,377.6	107,715.4
n	28	28
$\Sigma x^2 = \Sigma X^2 - \frac{(\Sigma X)^2}{n}$	426.7072	479.8343
$S^2 = \frac{\Sigma x^2}{n \quad 1}$	15.8040	17.7716

ต้องการทดสอบความแตกต่างของความสูงที่อายุ 60 วัน ของข้างพันธุ์ กข 23 เมื่อใส่ปุ๋ย Rock phosphate กับไม่ใส่ปุ๋ย ที่ระดับนัยสำคัญ .05 และ .01

$H_o : \mu_1 = \mu_2$ หรือ $\mu_1 - \mu_2 = 0$

$H_a : \mu_1 \neq \mu_2$

α = .05 และ .01

ทดสอบความแปรปรวนของทั้งสองประชากรว่าเท่ากันหรือไม่ ($H_o : \sigma_1^2 = \sigma_2^2$ VS $H_a : \sigma_1^2 \neq \sigma_2^2$)
โดยใช้ F-test

$$F = \frac{\text{large } S^2}{\text{small } S^2} = \frac{17.7716}{15.8040} = 1.12^{ns}$$

$$F_{.025,(27,27)} = 2.17$$

นั่นคือ ความแปรปรวนของทั้งสองประชากรไม่ต่างกัน จึงต้องหา Pooled variance

$$S_p^2 = \frac{\Sigma x_1^2 + \Sigma x_2^2}{n_1 + n_2 - 2} = \frac{426.7072 + 479.8343}{28 + 28 - 2} = 16.7878$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(65.18 - 61.89) - 0}{\sqrt{16.7878 \left(\frac{1}{28} + \frac{1}{28}\right)}} = 3.01^{**}$$

ค่า $t_{.025,54} = 2.00$ และ $t_{.005,54} = 2.67$

ปฏิเสธสมมุติฐาน $H_o : \mu_1 = \mu_2$

แปลผลได้ว่า ปุ๋ย Rock Phosphate มีผลทำให้ความสูงที่อายุ 60 วันของข้าวพันธุ์ กข23 แตกต่างจากเมื่อไม่ใส่ปุ๋ย

<u>ตัวอย่างที่ 9</u> นายดำสงสัยว่าถ้าปลูกข้าวโพดด้วยระยะปลูกที่แตกต่างกัน 2 ระยะ คือ ระยะระหว่างแถว x หลุม = 75 x 25 ซ.ม. และ 75 x 75 ซ.ม. และปล่อยให้วัชพืชขึ้นโดยไม่มีการกำจัดแล้วผลผลิตข้าวโพดที่ได้จะต่างกันหรือไม่เขาจึงทดลองปลูกข้าวโพด 2 ระยะปลูกดังกล่าว ปล่อยให้วัชพืชขึ้นจนถึงเวลาเก็บเกี่ยว ปรากฏว่าข้อมูลน้ำหนักเมล็ดข้าวโพดเป็นกิโลกรัมต่อไร่มีความชื้น 15% ได้เป็นดังนี้

แปลงที่	ระยะปลูก 75 x 25 ซม. X_1	ระยะปลูก 75 x 75 ซม. X_2
1	411	448
2	472	178
3	422	546
4	536	474
5	567	536
6	427	485
7	462	578
8	528	471
9	520	559
10	490	493
11	451	490
12	425	274
13	511	502
14	380	488
15	430	494
$\bar{X}$	468.8	467.7
n	15	15
ΣX	7,032	7,016
ΣX^2	3,338,078	3,438,516
$\Sigma x^2 = \Sigma X^2 - (\Sigma X)^2/n$	41,476.4	156,898.9
$s^2 = \Sigma x^2/(n-1)$	2,962.6	11,207.1

ต้องการทดสอบว่าน้ำหนักเมล็ดข้าวโพดที่ระยะปลูก 75 x 25 ซ.ม. แตกต่างจากระยะ 75 x 75 ซ.ม. หรือไม่เมื่อไม่มีการกำจัดวัชพืช

$H_o : \mu_1 = \mu_2$

$H_a : \mu_1 \neq \mu_2$

$\alpha = .05$

ทดสอบความแปรปรวนของประชากรทั้งสองว่าเท่ากันหรือไม่ ($H_o : \sigma_1^2 = \sigma_2^2$ VS $H_a : \sigma_1^2 \neq \sigma_2^2$) โดยใช้ F-test

$$F = \frac{\text{large } S^2}{\text{small } S^2}$$

$$= \frac{11,207.1}{2,962.6} = 3.78^*$$

$$F_{.025,(14,14)} = 3.00$$

ความแปรปรวนของทั้งสองประชากรแตกต่างกัน ดังนั้น

$$S_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

$$= \sqrt{\frac{2,962.6}{15}+\frac{11,207.1}{15}} = 30.7351$$

$$t = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_{\bar{X}_1-\bar{X}_2}}$$

$$= \frac{(468.8-467.7)-0}{30.7351} = 0.036^{ns}$$

เปิดตาราง t ที่ df เท่ากับ ν ดูสูตรหน้า 20

$$t_{.025,14} = 2.145$$

แปลผลได้ว่า ยังไม่มีหลักฐานเพียงพอที่จะกล่าวว่าเมื่อไม่มีการกำจัดวัชพืชแล้วผลผลิตน้ำหนักเมล็ดข้าวโพดที่ปลูกระยะ 75 x 25 ซ.ม. ต่างจากระยะ 75 x 75 ซ.ม. นั่นคือ เมื่อไม่มีการกำจัดวัชพืชแล้วผลผลิตน้ำหนักเมล็ดข้าวโพดที่ได้ เมื่อปลูกระยะ 75x25 ซ.ม. ไม่ต่างจากเมื่อปลูกระยะ 75 x 75 ซ.ม.

การทดสอบค่าเฉลี่ยของประชากรสองกลุ่มแบบจับคู่ (Paired comparison)

เป็นการทดสอบค่าเฉลี่ยของประชากรสองกลุ่มซึ่งมาจากตัวอย่างเดียวกันหรือไม่เป็นอิสระต่อกันในการทดสอบแบบนี้ข้อมูลจะถูกจัดเป็นคู่ สิ่งทดลองที่จับคู่ได้จะต้องเหมือนกันหรือคล้ายคลึงกันมาก เช่น ฝาแฝด ซีกซ้ายขวาของใบไม้ใบเดียวกัน ฯลฯ นี้จะช่วยลดการแปรปรวนจากแหล่งอื่น ๆ เพื่อว่าความแตกต่างที่เกิดขึ้นจะเป็นเนื่องจากทรีตเมนต์เกือบทั้งหมด

ถ้าตัวอย่างสุ่ม n คู่ แต่ละคู่ประกอบด้วยค่าสังเกต (X_{1i} , X_{2i}) เราต้องการทดสอบว่าไม่มีความแตกต่างระหว่างค่าเฉลี่ยของประชากรทั้งสอง ($H_o : \mu_d = 0$) แล้ว จะทดสอบได้โดยใช้ t-test ค่าสถิติ t คำนวณได้ดังนี้

$$t = \frac{\bar{d}-\mu_d}{S_{\bar{d}}}$$

เมื่อ d_i = ความแตกต่างของแต่ละคู่ หรือเท่ากับ X_{1i}-X_{2i} โดย i = 1,2,…..,n

$$\bar{d} = \sum_{i=1}^{n} d_i/n \text{ ซึ่งเท่ากับ } \bar{X}_1 - \bar{X}_2$$

$$S_d^2 = \frac{\sum_{i=1}^{n} d_i^2 - (\sum_{i=1}^{n} d_i)^2/n}{n \quad 1} \text{ และ}$$

$$S_{\bar{d}} = \sqrt{\frac{S_d^2}{n}}$$

เปรียบเทียบค่า t ที่คำนวณได้ กับค่า t จากตารางที่ (n-1) d.f. และระดับนัยสำคัญที่กำหนด

<u>ตัวอย่างที่ 10</u> ในการเปรียบเทียบวิธีการนับปริมาณเซลไรโซเบียม ที่เลี้ยงในอาหารเหลว 2 วิธี ผู้ทดลองเลี้ยงไรโซเบียมใน Broth ทิ้งไว้ 6-7 วัน แล้วนับปริมาณโคโลนีไรโซเบียมที่ปรากฏบนอาหารนั้น Log ฐาน 10 ของจำนวน <u>Rhizobium japonicum</u> ที่นับได้ต่อมิลลิลิตร เป็นดังนี้

ตัวอย่าง	**วิธี ก** X_1	**วิธี ข** X_2	**ผลแตกต่าง** $d=X_1-X_2$
1	9.66	8.60	1.06
2	9.62	9.37	0.25
3	9.50	9.27	0.23
4	9.46	8.21	1.25
5	9.15	8.49	0.66
6	9.82	8.57	1.25
7	9.66	8.32	1.34
8	8.90	8.72	0.18
9	8.99	8.95	0.04
10	9.08	9.09	-0.01

คำนวณ $\bar{X}_1 = 9.384 \quad \bar{X}_2 = 8.759 \quad \bar{d} = \bar{X}_1 - \bar{X}_2 = 0.625$

$$\sum d_i = 6.25 \qquad \sum d_i^d = 6.6293 \qquad n = 10$$

$$S_d^2 = \frac{\sum d^2 - (\sum d)^2/n}{n-1} = \frac{6.6293 - (6.25)^2/10}{10-1} = 0.3026$$

$$S_{\bar{d}}^2 = \frac{S_d^2}{n} = \frac{0.3026}{10} = 0.03026$$

$S^2_{\bar{d}} = \sqrt{S^2_{\bar{d}}} = \sqrt{0.0303} = 0.1739$

$H_o : \bar{D} = 0$

$H_a : \bar{D} \neq 0$

$\alpha = .05$

$$t = \frac{\bar{d} - \bar{D}}{S_{\bar{d}}}$$

$$= \frac{0.6250 - 0}{0.1739} = 3.59^*$$

กำหนดให้ระดับนัยสำคัญ (α) = .05

เทียบค่า t ที่คำนวณได้กับค่า t จากตารางที่ α = .05 และ (10-1=9) d.f.

$t_{.025,9} = 2.26$

ปฏิเสธสมมุติฐาน $H_o : \bar{D} = 0$

<u>แปลผลได้ว่า</u> การนับปริมาณโคโลนีไรโซเบียมโดยวิธีการ ก ให้ผลแตกต่างจากการนับโดยวิธีการ ข

ทดสอบความแตกต่างระหว่างความแปรปรวนของประชากรหลายกลุ่ม

เมื่อต้องการทดสอบความคล้ายคลึงของความแปรปรวนที่มาจากหลายประชากร (Test for homogeneity of variances) แล้ว วิธีทดสอบที่เป็นที่รู้จักและใช้กันอย่างแพร่หลายก็คือ วิธีของ Bartlett ซึ่งรู้จักกันในนาม Bartlett's Chi-square Test

ถ้า $S^2_1, S^2_2, \ldots\ldots, S^2_k$ เป็นความแปรปรวนจาก k ตัวอย่าง ซึ่งมีขนาด $n_1, n_2, \ldots\ldots n_k$ แล้ว การทดสอบความแตกต่างระหว่างความแปรปรวน k ค่า

$(H_o : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2)$

<u>กรณีที่ 1</u> ถ้า $n_1 = n_2 = \ldots\ldots = n_k = n$ แล้ว ค่าสถิติ χ^2 คำนวณได้ดังนี้

$$\chi^2 = \frac{(2.3026)(f)(k \log S_p^2 - \sum_{i=1}^{k} \log S_i^2)}{1 + ((k+1)/3kf)}$$

เมื่อ $f = (n-1)$

S_1^2 = ค่าความแปรปรวนจากประชากร (i = 1,2,…,k)

$S_p^2 = \sum_{i=1}^{k} S_i^2 / k$

โดยที่ k เป็นจำนวนค่าความแปรปรวนที่ต้องการทดสอบ

เปรียบเทียบค่า χ^2 ที่คำนวณได้ χ^2 กับค่า χ^2 จากตารางที่ (k-1) d.f. และระดับนัยสำคัญที่กำหนด

<u>ตัวอย่างที่ 11</u> ค่าความแปรปรวนของเปอร์เซ็นต์เมล็ดไม่สมบูรณ์อันเกิดจากการวิธีการปฏิบัติ 8 วิธี เป็นดังนี้

วิธีการ	S^2	n	Log S^2
1	5.5092	4	0.7411
2	1.2158	4	0.0849
3	0.3667	4	-0.4357
4	6.4200	4	0.8075
5	1.5633	4	0.1941
6	0.9100	4	-0.0410
7	0.4067	4	-0.3908
8	0.4425	4	-0.3541
รวม	16.8342		0.6060

$$S_p^2 = \sum_{i=1}^{k} S_i^2/k = \frac{16.8342}{8} = 2.1043$$

$$\chi^2 = \frac{(2.3026)(4-1)(8\log 2.1043 - 0.6060)}{1+((8+1)/3(8)(3))}$$

$$= \frac{13.6696}{1.125} = 12.1507^{ns}$$

$$\chi^2_{.05,7} = 14.07$$

นั่นคือ ไม่มีหลักฐานเพียงพอที่จะปฏิเสธว่าค่าความแปรปรวนจาก 8 วิธีการแตกต่างกัน

<u>กรณีที่ 2</u> ขนาดของตัวอย่างไม่เท่ากันในทุกค่า ค่าสถิติ χ^2 คำนวณได้ดังนี้

$$\chi^2 = \frac{2.3026((A)(\log S_p^2) - B)}{1+\frac{1}{3(k-1)}(C-\frac{1}{A})}$$

$$A = \sum_{i=1}^{k}(n_i - 1)$$

$$B = \sum_{i=1}^{k}(n_i - 1)\log S_i^2$$

$$C = \sum_{i=1}^{k}\frac{1}{(n_i - 1)}$$

$$S_p^2 = \frac{\sum_{i=1}^{k}(n_i - 1)S_i^2}{\sum_{i=1}^{k}(n_i - 1)}$$

เปรียบเทียบค่า χ^2 ที่คำนวณได้กับค่า χ^2 จากตารางที่ (k-1) d.f. และระดับนัยสำคัญที่กำหนด

ตัวอย่างที่ 12 ค่าความแปรปรวนของจำนวนสปอร์ต่อน้ำหนักดิน 100 กรัม ของเชื้อราที่นับได้จากวิธีการปฏิบัติ 4 วิธีเป็นดังนี้

วิธีการ	S^2	n	$(n-1)S^2$	Log S^2	(n-1)log S^2	1/(n-1)
1	121.0	6	605.0	2.0828	10.4139	0.20
2	8,680.9	4	26,042.7	3.9386	11.8157	0.333
3	9,205.5	4	27,616.5	3.9640	11.8921	0.333
4	9,413.5	4	28,240.5	3.9738	11.9213	0.333
	รวม		82,504.7		46.0430	

k = 4

A = 14   B = 46.0430   C = 1.20 S^2_p = 5,893.2

$$\chi^2 = \frac{2.3026((14)(\log 5{,}893.2) - 46.0430)}{1 + \frac{1}{3(4-1)}(1.20 - \frac{1}{14})}$$

$$= \frac{15.5239}{1.1254} = 13.79^{**}$$

$\chi^2_{.05,3}$ = 7.81 และ $X^2_{.01,3}$ = 11.34

ปฏิเสธสมมุติฐาน $H_o : \sigma^2_1 = \sigma^2_2 = \sigma^2_3 = \sigma^2_4$ นั่นคือ มีความแตกต่างระหว่างความแปรปรวนทั้ง 4 ค่านี้

หลักสำคัญในการวางแผนการทดลอง

สง่า ดวงรัตน์

หลักสำคัญในการวางแผนงานทดสอบ

การค้นคว้าวิจัยทางการเกษตร ส่วนใหญ่มักเป็นการทดลองในศูนย์วิจัย / สถานีทดลอง และไร่นาเกษตรกร จึงมักมีงานทดลอง โดยการใช้แบบแผนการทดลอง (Experimental design) แบบต่างๆ แต่ไม่ใช่หมายความว่า งานวิจัยทางการเกษตรทุกสาขาวิชาจำเป็นจะต้องเริ่มด้วย การเลือกใช้ Design แบบใดแบบหนึ่งเสมอไป ทั้งนี้ก็เนื่องด้วยงานทดลองด้านการเกษตร แบ่งได้เป็น 3 ประเภท คือ

1. การทดลองเบื้องต้น ต้องการทราบผลกว้าง ๆ เช่น การศึกษารวบรวมพันธุ์พืชที่ปลูกจากแหล่งต่าง ๆ นำมารวบรวมและคัดเลือกโดยการสังเกตการณ์ทดลองเปรียบเทียบเช่นนั้นมักทำเพียงซ้ำเดียว ข้อมูลที่เก็บได้นำมาวิเคราะห์หาค่าเฉลี่ยของความสูงผลผลิตหรืออัตราร้อยละเท่านั้น ไม่มีการเปรียบเทียบระหว่างพันธุ์พืชชุดนั้นว่ามีความแตกต่างกันทางสถิติหรือไม่

หรือตัวอย่างการศึกษาหาชนิดของเชื้อราที่เกิดในพันธุ์พืชต่างๆ ที่ขายในท้องตลาด ผู้วิจัยเพียงแต่ไปรวบรวมพันธุ์พืชเหล่านั้นมาจากหลายๆ แหล่ง แล้วทำการวิเคราะห์หาชนิดของเชื้อราว่ามีกี่ประเภท การทดลองแบบนี้ถือว่าเป็นการทดลองเบื้องต้น ซึ่งยังไม่จำเป็นต้องใช้ Design แบบใด จากการทดลองเบื้องต้นนี้ผลจะชี้ให้เห็นว่าเราควรจะทดสอบในขั้นต่อไปอย่างไรบ้าง

2. การทดลองขั้นตัดสินใจ เป็นขั้นที่ได้กลั่นกรองมาจากการทดลองเบื้องต้น โดยการคัดเลือกสายพันธุ์ที่มีผลผลิตและคุณภาพดีนำมาเปรียบเทียบกัน ซึ่งในขั้นนี้จำเป็นต้องนำ Design แบบต่าง ๆ ที่เหมาะสมมาใช้ เพื่อเปรียบเทียบสิ่งที่ต้องการทดสอบหรือทรีตเมนต์ มักทำหลายซ้ำ หลายท้องที่ หลายฤดูกาล เป็นต้น การทดลองแบบนี้มักเริ่มด้วยการเลือกแบบแผนการทดลองปัจจัยเดียวก่อน แล้วติดตามด้วยแบบแผนการทดลองมากกว่าหนึ่งปัจจัย ทั้งนี้ก็เพื่อจะให้ได้ผลการวิจัยที่มีประสิทธิภาพ คุ้มค่ากับเวลาแรงงานและงบประมาณ

การทดลองขั้นนี้ มีนักวิชาการเป็นผู้ดำเนินงาน ทำในห้องทดลอง ในสถานีทดลองเป็นส่วนใหญ่หรือทำในไร่นาเกษตรกร แล้วแต่ความเหมาะสม มักจะทำในแปลงย่อยขนาดเล็ก ผลที่ได้จากการทดลองนี้จะต้องนำไปทดสอบในสภาพไร่นาของเกษตรกรก่อนที่จะเผยแพร่ให้นำไปปฏิบัติตามต่อไป

3. การทดลองแบบกึ่งทดลองกึ่งสาธิต เป็นการทดลองที่นักวิชาการหรือนักส่งเสริมจัดทำในรูปของ Package ต่าง ๆ ที่นักวิชาการแต่ละเรื่องค้นคว้าวิจัยมาแล้ว โดยผ่านการทดลองขั้นตัดสินใจมาก่อนโดยมีการจัดพันธุ์พืชที่มีคุณภาพดี ให้ผลผลิตสูงมีวิธีการปลูกดูแลรักษา และการป้องกันและกำจัดศัตรูพืชที่เหมาะสม นำมาทดสอบพร้อมกันในรูปของ Package ต่าง ๆ โดยนำมาเปรียบเทียบให้เห็นชัดในแปลงทดสอบผืนใหญ่ ในขั้นตอนนี้มักจะศึกษาทางด้านเศรษฐศาสตร์ควบคู่ไปด้วยเพื่อจะได้ศึกษาถึงต้นทุนรายจ่ายและผลตอบแทนคุ้มค่าที่เกษตรกรควรได้รับ การดำเนินงานทดลองแบบนี้ถ้านักส่งเสริมและเกษตรกรทำร่วมกันเรียกว่า "การทำแปลงทดสอบพืช

คำศัพท์ต่าง ๆ ที่ควรทราบ

หน่วยทดลอง (Experiment unit) หมายถึง หน่วยที่เล็กที่สุดของวัตถุถือวัสดุในการทดลองที่ได้รับทรีตเมนต์อย่างเดียวกัน หน่วยการทดลองแต่ละหน่วยอาจเป็นหน่วยเดียวหรืออาจเป็นกลุ่มของวัสดุการทดลอง ตัวอย่าง เช่น

หน่วยเดียว

การนับจำนวนต้นถั่วที่เป็นโรควิสาของลูกผสมหลายพันธุ์ ซึ่งทดลองปลูกกระถางละ 1 ต้นในที่นี้ต้นถั่วแต่ละกระถาง คือ 1 หน่วยทดลอง

หน่วย / กลุ่ม

ในการทดลองเกี่ยวกับพืช มักจะหมายถึงพื้นดินแปลงเล็ก ๆ ที่มีรูปร่างสี่เหลี่ยมผืนผ้าค่อนข้างยาว เรียกว่า แปลงย่อย (Plot) หนึ่งแปลงย่อยอาจเท่ากับการปลูกข้าว 4 แถว ๆละ 20 ต้นเพราะฉะนั้น 1 หน่วยการทดลอง มีต้นข้าวเท่ากับ 80 ต้น เป็นต้น

ทรีตเมนต์ (Treatment) สิ่งที่ต้องการเปรียบเทียบข้อแตกต่างระหว่างลักษณะใดของพืช เรียกว่าสิ่งทดลองหรือทรีตเมนต์ (Treatment) ซึ่งอาจให้คำนิยามทั่วๆไปได้ว่า

สิ่งทดลองหรือทรีตเมนต์ หมายถึงวิธีการต่าง ๆ ซึ่งกระทำหรือเกิดขึ้นกับหน่วยทดลองเพื่อนำมาวัดผลเปรียบเทียบกันว่ามีความแตกต่างระหว่างทรีตเมนต์ที่ต้องการศึกษาวิจัยหรือไม่ ตัวอย่าง เช่น ในการศึกษา

1. เปรียบเทียบความแตกต่างของข้าวลูกผสม 2 พันธุ์ ในที่นี้มีทรีตเมนต์ คือ ข้าว 2 พันธุ์
2. เปรียบเทียบผลตอบสนองของปุ๋ย 5 ตำหรับต่อถั่วเหลือง เรียกว่ามีทรีตเมนต์ที่ต้องกาเปรียบเทียบ 5 ชนิด

ความคลาดเคลื่อนของการทดลอง (Experimental error) ในการทดลองเกี่ยวกับพืช ถึงแม้เราจะให้แต่ละหน่วยการทดลองได้รับทรีตเมนต์อย่างเดียวกัน มีการดูแลรักษาเหมือนกันหมดทุกอย่าง ซึ่งเราหวังว่าผลลัพธ์ที่ได้ เช่น ผลผลิตควรจะมีค่าเท่ากัน แต่มักจะไม่เป็นไปตามคิด ตัวอย่างเช่น เมื่อปลูกพันธุ์พืชชนิดเดียวกัน ในสิ่งแวดล้อมเหมือนกัน เช่น ปลูกถั่วเหลือง สจ.1 สองแปลง อยู่ติดกัน ระยะปลูกเท่ากัน ขนาดแปลงเท่ากัน ปลูกพร้อมกัน มีการดูแลปฏิบัติเหมือนกันทุกอย่าง

สจ. 1	สจ. 1

โดยหวังว่าควรจะได้ผลผลิต (หรือผลตอบสนองอย่างอื่น เช่น ความสูง จำนวนต้น จำนวนฝัก และลักษณะอื่น ๆ) เหมือนกัน แต่ความเป็นจริงมักจะเป็นไปไม่ได้ ทั้งนี้เพราะมีปัจจัยหลายอย่าง

เช่น ความแตกต่างทั้งของพันธุ์พืชนั้น ๆ อิทธิพลของดิน อุณหภูมิ ความชื้น แสงสว่าง รวมทั้งสาเหตุอื่น ๆ ทำให้ผลผลิตของทั้งสองแปลงแตกต่างกัน

ความแตกต่างที่เกิดจากผลผลิตของทั้งสองแปลงที่มีการดูแลปฏิบัติเหมือนกันทุกอย่างนี้ เรียกว่าค่าความแปรปรวนของการทดลองหรือค่าความคลาดเคลื่อนของการทดลอง (Experimental error)

การดำเนินงานทดลองใดๆ ก็ตาม จำต้องมีการควบคุมค่าความแปรปรวนของการทดลอง (Experimental error) ให้มีค่าน้อยที่สุด เพื่อที่จะมั่นใจได้ว่า สิ่งที่ต้องการทดสอบหรือ ทรีตเมนต์นั้นมีความแตกต่างกันจริง ไม่ใช่แตกต่างกันเนื่องจากสิ่งแวดล้อมหรือบังเอิญเป็นไปเท่านั้น

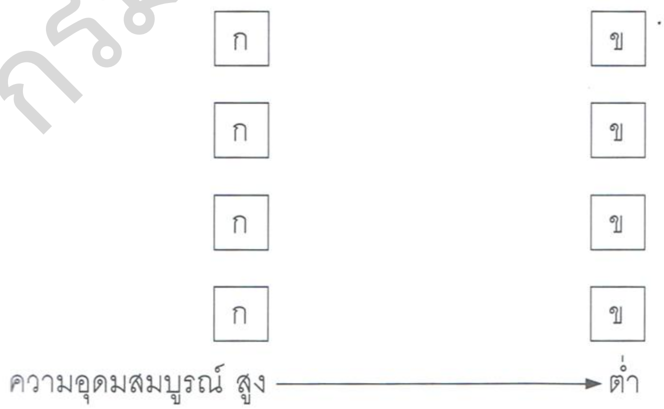
หลักสำคัญในการวางแผนงานทดลอง ต้องประกอบด้วย

1. การทำซ้ำ (Replication) ควรทำการทดสอบ treatment ใดก็ตามมากกว่า 1 ครั้ง (1 ซ้ำ) เพื่อใช้ประมาณค่าของ Experimental error เพราะค่านี้คำนวณได้จากความแตกต่างที่เกิดจากผลที่ได้ของอย่างน้อยสองหน่วยทดลองที่มีการดูแลปฏิบัติเหมือนกันทุกอย่าง

2. การสุ่มเลือก (Randomization) กำหนดให้ treatment ที่ทำการศึกษาได้รับการสุ่มเลือกในโอกาสที่เท่า ๆ กัน รวมทั้งได้รับการปฏิบัติที่เท่า ๆกัน เช่น ถ้าทุกแปลงมีการใส่ปุ๋ยก็ควรได้รับการใส่ในอัตราที่เท่ากัน และในเวลาเดียวกันเป็นต้น ทั้งนี้เพื่อ

2.1 การขจัดความลำเอียง (Bias) ที่เกิดขึ้นโดยไม่ตั้งใจ ทั้งก่อให้เกิดความแน่ใจว่าทุกทรีตเมนต์ที่ใช้ทดสอบไม่ได้เสียเปรียบกัน ยกตัวอย่างเช่น ในการปลูกพืช 2 ชนิด สมมุติให้เป็นพันธุ์ ก และพันธุ์ ข

สมมุติว่า ผู้ทดลองทำการปลูกพันธุ์ ก และพันธุ์ ข พันธุ์ละ 4 แปลง (หรือ 4 ซ้ำ) ดังในรูป



ถ้าบังเอิญแปลงที่ปลูกพันธุ์ ก ไปตกอยู่ในดินที่มีความอุดมสมบูรณ์สูง ส่วนแปลงที่ปลูกพันธุ์ ข ไปตกอยู่ในดินที่มีความอุดมสมบูรณ์ค่อนข้างต่ำกว่า ดังในรูป

ถ้าผลของการทดลองชุดนี้ปรากฏ ผลผลิตของพันธุ์ ก สูงกว่าพันธุ์ ข อาจทำให้ผู้ทดลองไม่สามารถสรุปได้ว่า โดยความเป็นจริงแล้ว พันธุ์ ก ดีกว่าพันธุ์ ข จริงหรือไม่ ทั้งนี้เพราะพันธุ์ ก บังเอิญได้ปลูกในดินที่มีความอุดมสมบูรณ์สูงกว่า ย่อมได้เปรียบกว่าพันธุ์ ข

วิธีที่ถูกต้อง คือ ต้องมีการสุ่มเลือก โดยให้โอกาสพันธุ์ ก และพันธุ์ ข ตกอยู่ในดินที่มีความอุดมสมบูรณ์ของดินดีหรือเลวในโอกาสที่เท่า ๆ กัน เช่น หลังจากการสุ่มเลือกแล้วอาจเป็นดังนี้

เมื่อได้ทำการปลูกพันธุ์พืชตามแผนผังที่ได้สุ่มเลือกถูกต้องแล้ว ถ้าผลจากการทดลองสรุปออกมาว่าพันธุ์ ก ดีกว่าพันธุ์ ข ผู้ทดลองย่อมมีความมั่นใจว่า พันธุ์ ก ดีกว่าพันธุ์ ข จริง ๆ ไม่ใช่ บังเอิญเป็นไปเหมือนดังกรณีแรก

2.2 เพื่อให้ค่า Experimental error ที่คำนวณได้เป็นค่าประมาณที่ถูกต้อง

2.3 ทำให้ผู้ทดลองสามารถสรุปผลได้ด้วยความมั่นใจว่าทรีตเมนต์ที่ทำการทดสอบนั้นมีความแตกต่างกันจริงที่ระดับความเชื่อมั่น 95 % หรือ 99 % (นิยมใช้สัญลักษณ์ * หรือ ** แทนตามลำดับ)

ระดับความเชื่อมั่น 95 % (หรือระดับความเป็นไปได้ .05) หมายถึง ในการทดลอง 100 ครั้ง โอกาสที่จะสรุปสมมุติฐานผิดมีเพียง 5 ครั้ง

ระดับความเชื่อมั่น 99 % (หรือระดับความเป็นไปได้ .01) หมายถึง ในการทดลอง 100 ครั้งโอกาสที่จะสรุปสมมุติฐานผิดมีเพียง 1 ครั้ง

3. มีการควบคุมหรือกำจัดความแปรปรวนให้เหลือน้อยที่สุด จะกล่าวถึงในหัวข้อเทคนิคทางสถิติในการดำเนินงานทดลอง ซึ่งจะอยู่ในเล่มที่ 2

การวางแผนงานทดลอง
สำหรับปัจจัยเดียว

สุทธีราภรณ์ สิริสิงห์

แบบแผนการทดลองปัจจัยเดียว

การทดลองใดๆ ที่ศึกษาถึงผลตอบสนองของปัจจัยหนึ่งโดยเฉพาะ และมีการควบคุมปัจจัยที่เกี่ยวข้องอื่นๆ ให้อยู่คงที่ หรือเป็นอย่างเดียวเหมือนกันตลอดไม่ว่าปัจจัยแรกจะเปลี่ยนแปลงไปอย่างใดก็ตาม เรียกว่า การทดลองปัจจัยเดียว ตัวอย่างเช่น การทดลองเปรียบเทียบพันธุ์อ้อยในเขตชลประทานซึ่งมีวัตถุประสงค์ที่จะทราบว่ามีพันธุ์ใดบ้างที่ให้ผลดีในสภาพที่ใช้น้ำชลประทานโดยควบคุมเกี่ยวกับการดูแลรักษาตั้งแต่ปลูกจนกระทั่งเก็บเกี่ยวเหมือนกันหมดทุกพันธุ์เช่นนี้เรื่องของพันธุ์จึงเป็นปัจจัยที่ทำการศึกษาเพียงอย่างเดียว ส่วนปัจจัยอื่น ๆ ในการดูแลรักษาได้แก่ วิธีการปลูก การให้น้ำ การใส่ปุ๋ย ฯลฯ เป็นปัจจัยคงที่ซึ่งไม่ว่าพันธุ์ใดๆ ก็ต้องได้รับเหมือนกันทุกพันธุ์

แผนการทดลองปัจจัยเดียวที่เป็นรูปแบบมาตรฐาน มี 3 แบบ คือ แผนการทดลองแบบสุ่มตลอด (Completely Randomized Design) แบบสุ่มในบล็อคสมบูรณ์ (Randomized Complete Block Design) และแบบละตินสแควร์ (Latin Square Design)

แผนการทดลองแบบสุ่มตลอด (CRD)

เป็นแผนการทดลองที่มีลักษณะง่าย และสะดวกทั้งในแง่การปฏิบัติและการคำนวณ เหมาะสำหรับหน่วยทดลองที่มีความสม่ำเสมอกันมากไม่มีความแตกต่างเนื่องจากปัจจัยอื่น ๆ เช่น อายุ น้ำหนัก สภาพแวดล้อม ต่างๆ ฯลฯ หรือถ้ามีก็น้อยมาก หน่วยทดลองแต่ละหน่วยมีโอกาสได้รับทรีตเมนต์เท่ากันจึงมักใช้กับการทดลองในห้องปฏิบัติการหรือเรือนทดลอง ซึ่งนอกจากจะกำหนดให้หน่วยทดลองสม่ำเสมอแล้วยังสามารถควบคุมไม่ให้เกิดความแปรปรวนในปัจจัยอื่นๆได้ด้วย

การสุ่มและแผนผังการทดลอง

การสุ่มอาจทำได้ง่าย ด้วยการจับฉลาก หรือการใช้ตารางเลขสุ่ม

ตัวอย่าง การสุ่มโดยวิธีจับฉลาก

การทดลองเปรียบเทียบ 5 ทรีตเมนต์ A, B, C, D, E, ทำ 3 ซ้ำ จะต้องมีหน่วยทดลองทั้งหมด 15 หน่วย (สมมุติให้หน่วยทดลองเป็นกระถาง) ให้เบอร์ประจำกระถางทุกกระถาง เขียนชื่อทรีตเมนต์ลงในกระดาษแผ่นเล็ก ชื่อละ 3 แผ่น รวมทั้งหมด 15 แผ่น ใส่ในกล่อง คลุกให้ทั่วแล้วสุ่มหยิบขึ้นมาทีละแผ่น สมมุติว่าได้ C ให้กับกระถางที่ 1 หยิบครั้งที่ 2 ได้ A ในกระถางที่ 2 ทำเช่นนี้ไปจนครบ 15 ครั้ง จะได้ผลของการสุ่มดังนี้ คือ

กระถางที่	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ทรีตเมนต์ที่สุ่มได้	C	A	B	E	C	D	D	A	E	B	A	E	C	B	D

ตัวอย่างการสุ่มโดยใช้ตารางการสุ่ม

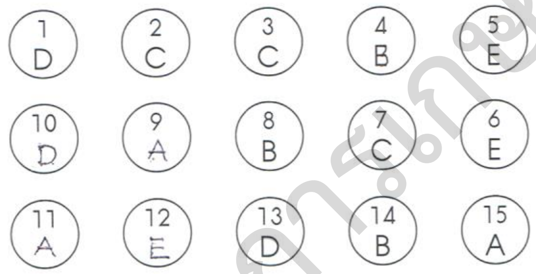
การทดลองในตัวอย่างแรก เมื่อให้เบอร์ประจำกระถางทั้ง 15 กระถางแล้ว จึงทำการสุ่มจากตารางเลขสุ่ม สมมุติสุ่มได้ Column ที่ 21 Row ที่ 21 ใช้เลข 2 หลัก นับลงทาง Column ให้ได้ 15 ตัว ไม่ซ้ำกัน จะได้เลขดังนี้

84 78 39 85 66 87 94 32 08 50 96 43 76 32 89

เรียงลำดับเลขที่สุ่มได้จากน้อยไปมากเพื่อใช้แทนเบอร์ประจำกระถางพร้อมทั้งเรียงชื่อทรีตเมนต์ควบคู่กันไปดังนี้

เลขที่สุ่ม	84	78	39	85	66	87	94	32	08	50	96	43	76	92	89
เรียงลำดับเลขที่สุ่มได้	9	8	3	10	6	11	14	2	1	5	15	4	7	13	12
เรียงชื่อทรีตเมนต์	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E

เมื่อจัดกระถางเป็น 3 แถว ๆ ละ 5 กระถางจะได้แผนผังการทดลองดังนี้



การวิเคราะห์ผลการทดลอง

1. การวิเคราะห์วาเหรียนซ์ เป็นการหาความแปรปรวนซึ่งเกิดจากแหล่งต่าง ๆ ภายในการทดลองสำหรับ CRD มีแหล่งความแปรปรวน 2 ชนิด คือ ความแปรปรวนที่เกิดจากทรีตเมนต์ และความคลาดเคลื่อนของการทดลอง ซึ่งเรียกว่า Experimental Error

ขั้นตอนของการคำนวณมีดังนี้

<u>ตัวอย่าง</u> การทดลอง 5 ทรีตเมนต์ 4 ซ้ำ (t = 5 , r = 4)

(1) จัดกลุ่มตัวเลขของแต่ละทรีตเมนต์ลงในรูปตาราง

ทรีตเมนต์/ซ้ำ	1	2	3	4	รวม	ค่าเฉลี่ย
1	X_{11}	X_{12}	X_{13}	X_{14}	T_1	$\bar{T}_1$
2	X_{21}	X_{22}	X_{23}	X_{24}	T_2	$\bar{T}_1$
3	X_{31}	X_{32}	X_{33}	X_{34}	T_3	T_3
4	X_{41}	X_{42}	X_{43}	X_{44}	T_4	T_4
5	X_{51}	X_{52}	X_{53}	X_{54}	T_5	T_5
รวม					G	$\bar{X}$

(2) คำนวณผลรวม ค่าเฉลี่ยของแต่ละทรีตเมนต์ ผลรวมทั้งหมด (G) และค่าเฉลี่ยทั้งหมด ($\bar{X}$)

(3) คำนวณ Degree of Freedom (df)

Total df	=	t x r - 1	= 5 x 4 - 1	= 19
Treatment df	=	t - 1	= 5 - 1	= 4
Error	=	(tr – 1) – (t – 1)	= 19 - 4	= 15

(4) คำนวณค่าความแปรปรวนที่เกิดจากแหล่งต่าง ๆ โดยใช้ตารางในข้อ (1)

- ค่า Sum of Square (SS)

$$\text{Correction Factor (C.F.)} = \frac{G^2}{t \times r}$$

$$\text{Total SS} = X^2_{11} + X^2_{12} + \ldots\ldots X^2_{54} - \text{C.F.}$$

$$\text{Treatment SS} = \frac{T_1^2 + T_2^2 + \ldots\ldots T_5^2 - \text{CF}}{r}$$

$$\text{Error SS} = \text{Total SS} - \text{Treatment SS}$$

- ค่า Mean Square (MS)

$$\text{Treatment MS} = \frac{\text{Treatment SS}}{\text{Treatment df}}$$

$$\text{Error MS} = \frac{\text{Error SS}}{\text{Error df}}$$

- ค่า F - value

$$F = \frac{\text{Treatment MS}}{\text{Error df}}$$

(5) เปิดตาราง F ที่ DF ของ Treatment (4) และ Error (15) จากตัวเลขแถวบน และด้านซ้ายสุดตามลำดับ จะได้ตัวเลข 2 ค่า คือ 3.06 และ 4.89

(6) ตารางสรุปผลการวิเคราะห์โดยกรอกค่าความแปรปรวนทั้งหมดที่คำนวณได้ลงในตารางวิเคราะห์วาเหรียนซ์

ANALYSIS OF VARIANCE (ANOV)

Source of Variation (SOV)	Degree of Freedom (DF)	Sum of Square (SS)	Mean Square (MS)	F-Value	
				Cal	Table
					5% 1%
Total	19				
Treatment	4				
Error	15				

(7) คำนวณค่า Coefficient of Variation (C.V.) ใส่ท้ายตาราง ANOV เพื่อชี้ให้เห็นถึงระดับความน่าเชื่อถือของการทดลองนั้น ๆ

$$C.V. \quad = \quad \frac{\sqrt{\text{Error MS}}}{\overline{X}} \times 100$$

(8) เปรียบเทียบค่า F ที่คำนวณได้กับค่า F จากตาราง

- ถ้าค่า F ที่วิเคราะห์ได้ มากกว่า F จากตารางที่ 1 % ใส่เครื่องหมาย ** บนค่า F ที่คำนวณได้ แสดงว่าค่าเฉลี่ยของทรีตเมนต์อย่างน้อย 1 คู่ มีความแตกต่างกันที่ระดับความเชื่อมั่น 99 %
- ถ้าค่า F ที่คำนวณได้ เท่ากับหรือมากกว่าค่า F จากตารางที่ 5 % แต่ไม่มากกว่าที่ 1 % ใส่เครื่องหมาย * บนค่า F ที่คำนวณได้แสดงว่าค่าเฉลี่ยของทรีตเมนต์อย่างน้อย 1 คู่ มีความแตกต่างกันที่ระดับความเชื่อมั่น 95 %
- ถ้าค่า F ที่คำนวณได้ น้อยกว่า F จากตาราง ใส่เครื่องหมาย ns ไว้บนค่า F แสดงว่ายังไม่มีหลักฐานเพียงพอที่จะยอมรับว่าทรีตเมนต์เหล่านั้นแตกต่างกัน

2. การเปรียบเทียบความแตกต่างของค่าเฉลี่ยของทรีตเมนต์

จากตาราง ANOV ถึงแม้ค่า F จึงชี้ให้เห็นว่าทรีตเมนต์มีความแตกต่างกันก็ตาม ในกรณีที่ F Cal > F Table แต่ก็ไม่ได้บอกชัดเจนว่ามีทรีตเมนต์คู่ใดบ้างที่แตกต่างกัน หรือมีทรีตเมนต์ใดบ้างที่มีค่าเฉลี่ยสูงกว่า Control จนถือว่าแตกต่างกัน จึงต้องทำการคำนวณค่าสถิติที่ใช้เปรียบเทียบค่าเฉลี่ยของทรีตเมนต์ ด้วยวิธีใดวิธีหนึ่งจาก 2 วิธีที่นิยมใช้กัน คือ LSD หรือ DMRT

(1) Least Significant Difference (LSD) ใช้เมื่อการทดลองมีไม่เกิน 5 ทรีตเมนต์ หรือ การทดลองที่มีทรีตเมนต์เป็น Check รวมอยู่ด้วย ซึ่งอาจจะมีเกิน 5 ทรีตเมนต์ก็ได้ และต้องใช้เปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยของทรีตเมนต์ใด ๆ กับ Check

$$LSD \quad = \quad t \times S_{\overline{d}}$$

เมื่อ t = ค่าจากตาราง student t ที่ df ของ Error เปิดได้ 2 ค่า คือ ที่ P.05 และ P.01

$$S_{\overline{d}} \quad = \quad \sqrt{\frac{2\text{xError MS}}{r}}$$

เปรียบเทียบค่าแตกต่างของค่าเฉลี่ยทรีตเมนต์คู่ที่ต้องการทราบ (d) กับค่า LSD ถ้าค่า d มากกว่า LSD แสดงว่าค่าเฉลี่ยทรีตเมนต์ทั้งสองแตกต่างกัน ถ้าค่า d น้อยกว่า LSD ก็แสดงว่าค่าเฉลี่ยของทั้งสองทรีตเมนต์ไม่แตกต่างกัน

(2) Duncan's New Multiple Range Test (DMRT) ใช้เปรียบเทียบค่าเฉลี่ยทั้งของการทดลองที่มีจำนวนมากกว่า 5 ทรีตเมนต์และไม่มี Check ที่ได้ค่า F Cal สูงกว่า F Table หรือการทดลองที่ได้ค่า F Cal > 1 แต่ ns. และต้องการเปรียบเทียบค่าเฉลี่ยโดยคำนวณดังนี้

$$LSR = SSR \times S_{\bar{x}}$$

เมื่อ SSR = ค่าจากตาราง Duncan ที่ df ของ Error จำนวน 4 ตัว คือ เมื่อ P = 2, 3, 4, 5 จะได้ SSR_2, SSR_3, SSR_4, SSR_5

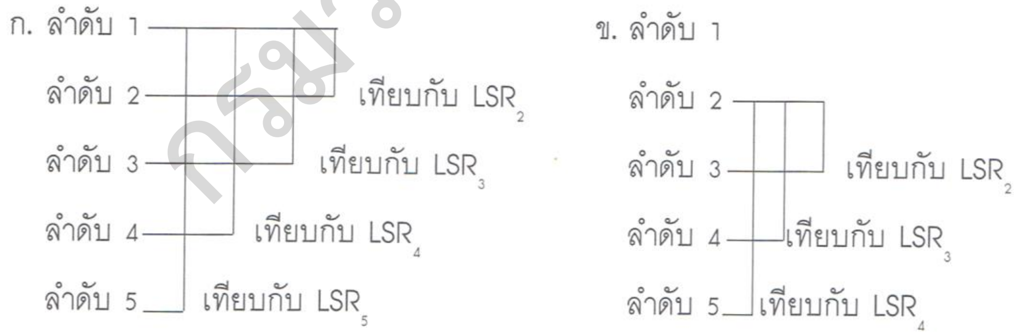
$$S_{\bar{x}} = \sqrt{\frac{\text{Error MS}}{r}}$$

ค่า LSR จะมีจำนวนเท่ากับจำนวนทรีตเมนต์ลบ 1 เช่น มี 5 ทรีตเมนต์ จะต้องคำนวณค่า LSR 4 ค่า คือ LSR_2 , LSR_3 , LSR_4 และ LSR_5 ดังนี้

$LSR_2 = SSR_2 \times S_{\bar{x}}$ $\qquad$ $LSR_3 = SSR_3 \times S_{\bar{x}}$

$LSR_4 = SSR_4 \times S_{\bar{x}}$ $\qquad$ $LSR_5 = SSR_5 \times S_{\bar{x}}$

- เรียงลำดับค่าเฉลี่ยจากมากไปหาน้อยหรือกลับกัน แล้วแต่ลักษณะข้อมูลที่ทำการศึกษา
- เปรียบเทียบค่าแตกต่างของค่าเฉลี่ยแต่ละคู่กับค่า LSR โดยเทียบค่าแตกต่างของค่าเฉลี่ยลำดับที่ 1 กับลำดับที่ 5 กับ LSR_5 ดังแผนผัง ก. ถ้าค่าแตกต่างมากกว่าค่า LSR แสดงว่าค่าเฉลี่ยลำดับที่ 1 และ ลำดับที่ 5 แตกต่างกัน



เปรียบเทียบค่าเฉลี่ยลำดับที่ 1 ลำดับที่ 4 กับ LSR_4 ทำเช่นนี้จนกว่าค่าแตกต่างน้อยกว่าค่า LSR จึงหยุด แล้วทำซ้ำเช่นเดิม โดยเปลี่ยนเป็นใช้ค่าแตกต่างระหว่างค่าเฉลี่ยลำดับที่ 2 กับลำดับที่ 5 และเทียบกับค่า LSR_4 ตามแผนผัง ข.

- การเปรียบเทียบค่าเฉลี่ยโดยวิธี DMRT มักใช้ระดับความเชื่อมั่น 95 % โดยใช้ตาราง SSR ที่ P . 05 และใช้ตัวอักษรหรือกลุ่มตัวอักษรแสดงความแตกต่างหรือไม่แตกต่างของค่าเฉลี่ย

<u>ตัวอย่างที่ 1 การทดลองที่มีจำนวนซ้ำเท่ากัน</u>

เรื่อง การใช้ฮอร์โมน NAA และ IAA เร่งการเจริญเติบโตของท่อนพันธุ์มันสำปะหลัง

ทรีตเมนต์ (t) : T_1 = จุ่มน้ำ T_4 = IAA อัตราที่ 1

T_2 = NAA อัตราที่ 1 T_5 = IAA อัตราที่ 2

T_3 = NAA อัตราที่ 2

จำนวนซ้ำ (r) = 4

แผนผังการทดลอง

T_2	T_3	T_3	T_4
T_4	T_1	T_5	T_1
T_5	T_3	T_4	T_2
T_1	T_2	T_1	T_5
T_5	T_4	T_2	T_3

ขั้นตอนการวิเคราะห์ผลการทดลองมีดังนี้

1. การวิเคราะห์วาเหรียนซ์ ดำเนินการตามขั้นตอนที่ 1 (1) ถึง (8) ดังนี้
   - จัดกลุ่มข้อมูลลงในตาราง

ตารางแสดงความสูง (ซ.ม.) ของต้นมันสำปะหลังพันธุ์ระยอง 3 อายุ 3 เดือน ภายหลังการใช้ฮอร์โมน

ทรีตเมนต์\ซ้ำ	**1**	**2**	**3**	**4**	**รวม**	**ค่าเฉลี่ย**
1 จุ่มน้ำ	26.0	25.1	20.2	24.0	95.3	23.8
2 NAA อัตรา 1	29.4	32.0	22.5	27.9	111.8	28.
3 NAA อัตรา 2	28.5	25.5	25.6	26.3	105.9	26.5
4 IAA อัตรา 1	30.0	28.1	26.3	28.1	112.5	28.1
5 IAA อัตรา 2	34.6	27.4	29.0	30.2	121.2	30.3
รวม					546.7	27.34

- คำนวณผลรวม ค่าเฉลี่ยของทรีตเมนต์ ผลรวมทั้งหมด และค่าเฉลี่ย
- คำนวณ df : Total df. = 5 x 4 - 1 = 19

  Treatment df. = 5 - 1 = 4

  Error df. = 19 - 4 = 15

- คำนวณค่าความแปรปรวนของแหล่งต่าง ๆ ต่อไปนี้

$$\text{C.F.} = \frac{(546.7)^2}{5\text{x}4} = 14{,}944.0445$$

$$\text{Total} = 26.0^2+25.1^2+\ldots+30.2^2 - \text{C.F.} = 200.4455$$

$$\text{Treatment SS} = \frac{95.3^2 + 111.8^2 + \ldots + 121.2^2}{4} \quad \text{C.F.} = 91.4130$$

$$\text{Error SS} = 200.4455-91.4130 = 109.0325$$

$$\text{Treatment MS} = \frac{91.4130}{15} = 7.2688$$

$$\text{Error MS} = \frac{109.0325}{15} = 7.2688$$

$$\text{F} = \frac{22.8532}{7.2688} = 3.14$$

- เปิดตาราง F ที่ df ของ Treatment กับ Error คือ 4, 15 ได้ค่า F .05 = 3.06 และ .01 = 4.89
- กรอกค่าที่วิเคราะห์ได้ในตาราง ANOV

SOV	df	SS	MS	F-Value		
				Cal.	5%	1%
Total	19	200.4455				
Treatment	4	91.4130	22.8532	3.14*	3.06	4.89
Error	15	109.0329	7.2688			

- คำนวณค่า C.V. ใส่ท้ายตาราง ANOV

$$\text{C.V.} = \frac{\sqrt{7.2688}}{27.34}\text{x}100 = 9.86\%$$

2. การคำนวณค่าสถิติเปรียบเทียบค่าเฉลี่ยของทรีตเมนต์

จากตาราง ANOV ค่า F จากการคำนวณมากกว่า F จากตารางที่ 5 % แต่ไม่มากกว่าที่ 1 % หมายความว่ามีค่าเฉลี่ยของทรีตเมนต์อย่างน้อย 1 คู่ แตกต่างกันที่ระดับความเชื่อมั่น 95 % เพื่อจะทดสอบว่าค่าเฉลี่ยของทรีตเมนต์คู่ไหนแตกต่างกันบ้าง จึงคำนวณค่าสถิติอย่างใดอย่างหนึ่ง คือ LSD หรือ DMRT

วิธี LSD

$$\text{LSD 5\%} = t_{.05}\text{ x }\sqrt{\frac{2\text{x}7.2688}{4}} = 2.131 \text{ x } 1.9064 = 4.1$$

$$\text{LSD 1\%} = t_{.01}\text{ x }\sqrt{\frac{2\text{x}7.2688}{4}} = 2.947 \text{ x } 1.9064 = 5.6$$

- เปรียบเทียบความแตกต่างของความสูงของต้นมันสำปะหลังระหว่างคู่ใดคู่หนึ่งของทรีตเมนต์

	T_1(จุ่มน้ำ)	$T_2(NAA_1)$	$T_3(NAA_2)$	$T_4(IAA_1)$	$T_5(IAA_2)$
ค่าเฉลี่ย	23.8	28.0	26.5	28.1	30.3
T_1 23.8	-	4.2*	2.7^{ns}	4.3*	6.5**
T_2 28.0			1.5^{ns}	0.1^{ns}	2.3^{ns}
T_3 26.5				1.6^{ns}	3.8^{ns}
T_4 28.1				-	2.2^{ns}

**แตกต่างกันที่ระดับความเชื่อมั่น 99 % *แตกต่างกันที่ระดับความเชื่อมั่น 95 %

ns ไม่แตกต่างกันทางสถิติ

การแปลผล

ฮอร์โมนที่ให้ผลในการเร่งการเจริญเติบโตของท่อนมันสำปะหลัง คือ IAA อัตรา 2, IAA อัตรา 1 และ NAA อัตรา 1 ทำให้ต้นมันเมื่ออายุ 3 เดือน สูงกว่าไม่ใช้ฮอร์โมนจนแสดงความแตกต่างอย่างมีนัยสำคัญ โดยเฉพาะ IAA อัตรา 2 ทำให้ได้ต้นสูง 30.3 ซ.ม. แต่ในระหว่างการใช้ฮอร์โมนด้วยกันไม่มีผลแตกต่างกันในเรื่องของความสูง

<u>ตัวอย่างที่ 2</u> การทดลองมีจำนวนซ้ำไม่เท่ากัน

การทดลองแบบ CRD บางครั้งจำเป็นต้องทำจำนวนซ้ำไม่เท่ากัน อาจเนื่องจากมีหน่วยทดลองไม่เพียงพอ หรือเริ่มทำด้วยซ้ำเท่า ๆ กัน แต่เกิดความเสียหายกับหน่วยทดลองในระหว่างการปฏิบัติ ซึ่งความเสียหายนั้นไม่ได้เกิดจากผลของทรีตเมนต์ และไม่สามารถเก็บข้อมูลได้ครบทุกครั้งซ้ำ การวางแผนแบบนี้มีข้อได้เปรียบคือ สามารถวิเคราะห์ทางสถิติได้ โดยไม่ต้องเสียการทดลองนี้ไปทั้งหมด

เรื่อง การศึกษาผลของสารฆ่าแมลงบางชนิดต่อแตนเบียนไข่หนอนเจาะลำต้นข้าวโพดในห้องปฏิบัติการ

ทรีตเมนต์ = สารฆ่าแมลง 10 ชนิด

จำนวนซ้ำ = 2 - 4 ซ้ำ

มีขั้นตอนการคำนวณเช่นเดียวกับการวิเคราะห์ CRD ที่มีจำนวนซ้ำเท่ากัน

1. การวิเคราะห์วาเหรียนซ์

- จัดกลุ่มข้อมูลลงในตาราง หาผลรวมและค่าเฉลี่ยของแต่ละทรีตเมนต์ ผลรวมทั้งหมด (G) และค่าเฉลี่ยของทั้งหมด ($\bar{X}$)

ตารางแสดง % ไข่หนอนที่ถูกแตนเบียน (ที่รอดจากการพ่นสารเคมี) เข้าทำลาย

ทรีตเมนต์\ซ้ำ	1	2	3	4	รวม	ค่าเฉลี่ย
1 Atabron	55.06	60.94	62.31	49.95	228.26	57.065
2 Z-killer	37.23	70.81	62.51	52.42	222.97	55.742
3 น้ำ (check)	65.35	80.72	65.88	51.88	263.83	65.958
4 Ripcord	66.42	71.00	74.66		212.08	70.693
5 Kumiphos	71.19	55.92	57.29		184.40	61.467
6 Karate	78.46	56.42	80.72		215.60	71.867
7 Sumicidin	59.93	63.51	74.11		197.55	65.850
8 Pay-off	59.74	62.37	67.70		189.81	63.270
9 Thiodan	57.99	56.42	42.19		156.60	52.200
10 Lorsban	39.99	35.24			75.23	37.615
รวม					1,946.33(G)	60.823 $(\bar{X})$

- คำนวณ df : total df = (3x4)+(3x6)+(2x1)-1 = 31

  Treatment df = 10 – 1 = 9

  Error df = 31 – 9 = 22

- คำนวณค่าความแปรปรวนของแหล่งต่างๆ

$$\text{C.F.} = \frac{1{,}946.33^2}{32} = 118{,}381.2647$$

$$\text{Total SS} = (55.06^2 + 60.94^2 + \ldots + 35.24^2) - \text{C.F.} = 4{,}300.6432$$

$$\text{Treatment SS} = \frac{228.26^2}{4} + \frac{222.97^2}{4} + \ldots \frac{212.08^2}{3} + \ldots \frac{75.23^2}{2} - \text{C.F.}$$

$$= 2{,}318.6561$$

$$\text{Error SS} = 4{,}300.6432 - 2{,}318.6561 = 1{,}981.9871$$

$$\text{Treatment MS} = \frac{2{,}318.6561}{9} = 257.6284$$

$$\text{Error MS} = \frac{1{,}981.9871}{22} = 90.0903$$

$$\text{F Value} = \frac{257.6284}{90.0903} = 2.85$$

- เปิดตาราง F ที่ df ของ treatment = 9 และ Error = 22 ได้ค่า F table
- คำนวณค่า C.V.

$$\text{C.V.} = \frac{\sqrt{90.0903}}{60.823} \text{x}100 = 15.6\%$$

- กรอกค่าที่วิเคราะห์ได้ลงในตาราง ANOV

SOV	df	SS	MS	F-Value		
				Cal	5%	1%
Total	31	4,300.6432				
Treatment	9	2,318.6561	257.6284	2.85*	2.35	3.35
Error	22	1,981.9871	90.0903			

C.V. = 15.6%

2. เปรียบเทียบค่าแตกต่างของค่าเฉลี่ยด้วยวิธี DMRT โดยเรียงค่าเฉลี่ยจากมากไปหาน้อยและมีการคำนวณค่า LSR เพื่อเปรียบเทียบค่าเฉลี่ยระหว่างทรีตเมนต์ที่มีจำนวนซ้ำแตกต่างกัน คือ 4 ซ้ำ VS 3 ซ้ำ, 4 ซ้ำ VS 2 ซ้ำ, 3 ซ้ำ VS 2 ซ้ำ โดยใช้สูตรการคำนวณ $S_{\bar{X}}$ ดังนี้

$$S_{\bar{X}} = \sqrt{\frac{1}{2}\left(\frac{1}{r_i}+\frac{1}{r_j}\right)Error\ MS}$$

เมื่อ r_i และ r_j เป็นจำนวนซ้ำที่ไม่เท่ากัน

$$4\text{ VS }3 = \sqrt{\frac{1}{2}\left(\frac{1}{4}+\frac{1}{3}\right)90.0903}$$

$$4\text{ VS }2 = \sqrt{\frac{1}{2}\left(\frac{1}{4}+\frac{1}{2}\right)90.0903}$$

$$3\text{ VS }2 = \sqrt{\frac{1}{2}\left(\frac{1}{3}+\frac{1}{2}\right)90.0903}$$

และดำเนินการเปรียบเทียบโดยใช้หลักการในข้อ 2 (2) ได้ผลการเปรียบเทียบดังนี้

ลำดับที่	ทรีตเมนต์	ค่าเฉลี่ย	
1	6 Karate	71.867	a
2	4 Ripcord	70.993	ab
3	3 น้ำ	65.956	bc
4	7 Sumicidin	65.850	abc
5	8 Pay-off	63.270	abc
6	5 Kumiphos	61.467	abc
7	1 Atabron	57.065	abc
8	2 Z-killer	55.7425	bc
9	9 Thiodan	52.200	cd
10	10 Lorban	37.615	d

การแปลผล

สารฆ่าแมลงที่เป็นพิษต่อแตนเบียนน้อยที่สุด มีผลทำให้แตนเบียนรอดชีวิตจนมีโอกาสเข้าทำลายไข่หนอนเจาะลำต้นข้าวโพดมากที่สุด คือ Karate (ไข่ หนอนถูกทำลายถึง 71.9 %) ผลนี้ไม่แตกต่างไปจากการใช้น้ำ Ripcord Sumicidin Pay-off Kumiphos และ Atabron แต่แตกต่างกับ t-killer Thiodan และ Lorsban โดยที่ Losrsban และ Thiodan เป็นพิษต่อแตนเบียนมากจนเหลือความสามารถในการทำลายไข่หนอนได้เพียง 37.6 และ 52.2 % ตามลำดับเท่านั้น

แผนการทดลองแบบสุ่มในบล็อคสมบูรณ์ (RCB)

โดยธรรมชาติหน่วยทดลองทั่ว ๆ ไปโดยเฉพาะในไร่มักไม่มีความสม่ำเสมอกันมากนัก ถ้าทำการทดลองโดยไม่จัดการกับหน่วยทดลองให้มีความใกล้เคียงกันหรือแตกต่างกันน้อยที่สุดแล้ว ย่อมเกิดความแปรปรวนในการทดลองมากจนกระทั่งไม่สามารถตรวจสอบว่าทรีตเมนต์ต่างๆ ในการทดลองนั้นให้ผลแตกต่างกันอย่างไรหรือไม่ จึงจำเป็นต้องทำการจัดกลุ่มของหน่วยทดลองให้ภายในกลุ่มมีความแตกต่างกันน้อยที่สุด แต่แตกต่างกันมากที่สุดระหว่างกลุ่ม เรียกว่า บล็อค (Block)

การจัดบล็อค

การจัดบล็อคมีวัตถุประสงค์เพื่อลดความแปรปรวนของการทดลองให้น้อยลง มีหลักการที่ต้องพิจารณา คือ

1. ต้นเหตุของความแปรปรวนในหน่วยทดลอง เช่น ความอุดมสมบูรณ์ของดิน หรือทิศทางการระบาดของแมลงในงานทดลองด้านอารักขาพืช หรือความลาดเอียงของพื้นที่เป็นต้น
2. การจัดรูปร่างและทิศทางของบล็อค ยกตัวอย่างเฉพาะเรื่องความอุดมสมบูรณ์ของดินเท่านั้นซึ่งสามารถนำไปใช้กับความแปรปรวนอันเกิดจากต้นเหตุอื่น ๆ ได้
    - เมื่อดินมีทิศทางของความอุดมสมบูรณ์ของดินทางเดียว ให้จัดบล็อคเป็นรูปร่างแคบและยาวให้ด้านยาวของบล็อค ตั้งฉากกับทิศทางของความอุดมสมบูรณ์ของดิน
    - เมื่อดินมีความอุดมสมบูรณ์ 2 ทิศทาง ทางหนึ่งมากกว่าอีกทางหนึ่ง จัดบล็อคเป็นรูปร่างแคบและยาว ให้ด้านยาวของบล็อคตั้งฉากกับทิศทางที่มีความอุดมสมบูรณ์มากกว่า
    - เมื่อดินมีความอุดมสมบูรณ์แตกต่างกันเป็น 2 ทิศทางตั้งฉากกัน ใช้วิธีจัดบล็อคให้เป็นสี่เหลี่ยมจัตุรัส หรือใช้แผนการทดลองแบบละตินสแควร์อย่างใดอย่างหนึ่ง
    - เมื่อไม่รู้ทิศทางอุดมสมบูรณ์ของดิน ควรจัดบล็อคให้เป็นรูปสี่เหลี่ยมจัตุรัสมากที่สุด

การสุ่มและแผนผังการทดลอง

ทำการสุ่มแยกกันทีละบล็อค ในแต่ละบล็อคจะต้องมีครบทุกทรีตเมนต์ เช่นทำการทดลอง 5 ทรีตเมนต์ : A B C D E ทำ 4 ซ้ำ ให้จัดบล็อคตามหลักการข้างต้น 4 บล็อค ในแต่ละบล็อคแบ่งเป็นแปลงย่อย 5 แปลง ให้เบอร์ประจำแปลงย่อย แล้วสุ่มทรีตเมนต์ด้วยวิธีการใดวิธีการหนึ่งในเรื่อง CRD ให้กับแปลงย่อยทั้ง 5 แปลง ในบล็อคที่ 1 และทำการสุ่มเช่นเดียวกันนี้กับบล็อคที่ 2, 3 และ 4 ตามลำดับจะได้แผนผังการทดลองดังนี้

Block 1	2	3	4
C	A	E	D
B	E	C	A
A	C	A	B
D	B	D	C
E	D	B	E

การวิเคราะห์การทดลอง

การวิเคราะห์ผลในแผนการทดลองแบบสุ่มในบล็อคสมบูรณ์ จะมีความแปรปรวนจาก 3 แหล่ง คือ ความแปรปรวนที่เกิดจากบล็อค ทรีตเมนต์ และความคลาดเคลื่อนที่เกิดจากการปฏิบัติ (Experimental Error)

มีขั้นตอนการคำนวณต่อไปนี้

<u>ตัวอย่าง</u> การทดลอง 5 ทรีตเมนต์ A B C D E ทำ 4 บล็อค(ซ้ำ) (t = 5 , r = 4)

1. กรอกตัวเลขลงในตาราง

ทรีตเมนต์\ซ้ำ	1	2	3	4	รวม	ค่าเฉลี่ย
T_1 = A	X_{11}	X_{12}	X_{13}	X_{14}	T_1	$\bar{X}_1$
T_2 = B	X_{21}	X_{22}	X_{23}	X_{24}	T_2	$\bar{X}_2$
T_3 = C	X_{31}	X_{32}	X_{33}	X_{34}	T_3	$\bar{X}_3$
T_4 = D	X_{41}	X_{42}	X_{43}	X_{44}	T_4	$\bar{X}_4$
T_5 = E	X_{51}	X_{52}	X_{53}	X_{54}	T_5	$\bar{X}_5$
รวม	R_1	R_2	R_3	R_4	G	$\bar{X}$

2. คำนวณผลรวม ค่าเฉลี่ยของแต่ทรีตเมนต์ ผลรวมของซ้ำ ผลรวมทั้งหมดและค่าเฉลี่ย

3. คำนวณ Degree of Freedom (df)

Total df	=	t x r - 1	=	5 x 4 - 1	=	19
Replication df	=	r - 1	=	4 - 1	=	3
Trratment df	=	t - 1	=	5 - 1	=	4
Error df	=	(t – 1) (r – 1)	=	3 x 4	=	12

4. คำนวณค่าความแปรปรวนที่เกิดจากแหล่งต่าง ๆ โดยใช้ข้อมูลในตารางข้อ 1

- ค่า Sum of Square (SS)

Correction Factor (C.F.) $= \dfrac{G^2}{txr}$

Total SS $= (X^2_{11} + X^2_{21} + \ldots + X^2_{54}) - C.F.$

Replication SS $= \dfrac{(R_1^2 + R_2^2 + \ldots + R_4^2)}{t} - C.F.$

Treatment SS $= \dfrac{(T_1^2 + T_2^2 + \ldots + T_5^2)}{r} - C.F.$

Error SS = Total SS - Replication SS – Treatment SS

- ค่า Mean Square (MS)

Replication MS $= \dfrac{\text{Replication SS}}{r - 1}$

Treatment MS $= \dfrac{\text{Treatment SS}}{t - 1}$

Error MS $= \dfrac{\text{Error SS}}{(t - 1)(r - 1)}$

- ค่า F - Value

$$F = \frac{\text{Treatment MS}}{\text{Error MS}}$$

5. เปิดตาราง F ที่ (t - 1) = 4 และ (t - 1)(r - 1) = 12 จะได้ค่าตัวเลข 2 ค่าคือ 3.26 และ 5.41

6. ทำตารางสรุปผลการวิเคราะห์ (ANOV) โดยกรอกค่าที่คำนวณได้ทั้งหมดในตาราง

SOV	df	SS	MS	F-Value		
				Cal	5%	1%
Total	19					
Replication	3					
Treatment	4					
Error	12					

7. คำนวณค่า C.V. ใส่ท้ายตาราง ANOV

8. เปรียบเทียบค่า F ที่คำนวณได้กับค่า F ในตารางโดยใช้หลักการเช่นเดียวกับ CRD

9. คำนวณค่าสถิติเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยของทรีตเมนต์

<u>ตัวอย่าง</u> เรื่อง การศึกษาผลของความถี่ในการฉีด NAA ต่อการเพิ่มผลผลิตและคุณภาพของแตงกวา

ทรีตเมนต์ (T) : T_1 = ไม่พ่น NAA

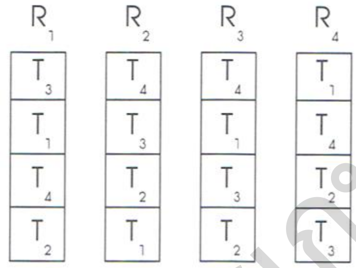
T_2 = พ่น 3 ครั้ง

T_3 = พ่น 6 ครั้ง

T_4 = พ่น 9 ครั้ง

จำนวนซ้ำ (r)

แผนผังการทดลอง



- วิเคราะห์วาเหรียนซ์ของข้อมูลในตาราง

ตารางแสดงน้ำหนักผลผลิตแตงกวา (กรัม)

ทรีตเมนต์\ซ้ำ	**1**	**2**	**3**	**4**	**รวม**	**ค่าเฉลี่ย**
1 ไม่พ่น	312.70	325.91	340.91	393.87	1,373.39	343.3475
2 พ่น 3 ครั้ง	490.76	509.01	560.48	515.96	2,076.21	519.0525
3 พ่น 6 ครั้ง	579.60	647.23	637.95	583.64	2,448.42	612.1050
4 พ่น 9 ครั้ง	394.11	381.20	468.37	469.53	1,713.21	428.3025
รวม	1,777.17	1,863.35	2,007.71	1,963.00	7,611.23	475.7019

- คำนวณค่า df

Total df = 4 x 4 - 1 = 15

Replication = 4 - 1 = 3

Treatment df = 4 - 1 = 3

Error df = 15 - 3 - 3 = 9

- คำนวณค่าความแปรปรวนของแหล่งต่าง ๆ คือ

$$\text{C.F.} = \frac{7{,}611.23^2}{16} = 3{,}620{,}676.382$$

$$\text{Total SS} = 312.70^2+325.91^2+\ldots+469.53^2 - \text{C.F.} = 177{,}888.2456$$

$$\text{Replication} = \frac{1{,}777.17^2 + \ldots + 1{,}963.00^2}{4} - \text{C.F.} = 7{,}992.3368$$

$$\text{Treatment SS} = \frac{1{,}373.39^2 + \ldots + 1{,}713.21^2}{4} = 160{,}997.8821$$

$$\text{Error SS} = 177{,}888.2456 - 7{,}992.3368 - 160{,}997.8821 = 8{,}898.0267$$

$$\text{Replication MS} = \frac{7{,}992.3368}{3} = 2{,}664.1123$$

$$\text{Treatment MS} = \frac{160{,}997.8821}{3} = 53{,}665.9607$$

$$\text{Error MS} = \frac{8{,}898.0267}{9} = 988.6696$$

$$F = \frac{53{,}665.9607}{988.6696} = 54.28$$

- เปิดตาราง F ที่ df (3 , 9) ได้ค่า F ที่ . 05 = 3.86 และที่ .01 = 6.99
- คำนวณค่า C.V. เพื่อใส่ท้ายตาราง

$$\text{C.V.} = \frac{\sqrt{988.6696}}{475.7019} \times 100$$

- กรอกค่าที่วิเคราะห์ได้ทั้งหมดลงในตาราง ANOV

SOV	df	SS	MS	F-Value		
				Cal	5%	1%
Total	15	177,888.2456				
Replication	3	7,992.3368	2,664.1123			
Treatment	3	160,997.8821	53,665.9607	54.28**	3.86	6.99
Error	9	8,896.0267	988.6696			

C.V. = 6.6%

- คำนวณค่า LSD เพื่อเปรียบเทียบความแตกต่างระหว่างการพ่น NAA จำนวนครั้งต่าง ๆ กับการไม่พ่น

$$\text{LSD 5\%} = t_{.05} \times \sqrt{\frac{2 \times 988.6696}{4}} = 2.262 \times 22.29$$

$$= 50.29 \text{ กรัม}$$

$$\text{LSD 1\%} = t_{.01} \times \sqrt{\frac{2 \times 988.6696}{4}} = 3.250 \times 22.2336$$

$$= 72.26 \text{ กรัม}$$

ทรีตเมนท์	น้ำหนักผลผลิตเฉลี่ย (กรัม)	ค่าแตกต่างจาก Check
1. ไม่พ่น NAA (Check)	343.35	-
2. พ่น 3 ครั้ง	519.05	175.70**
3. พ่น 6 ครั้ง	612.10	268.76**
4. พ่น 9 ครั้ง	428.30	84.96**

** แตกต่างกันที่ระดับความเชื่อมั่น 99 %

การแปลผล

การพ่น NAA มีผลต่อการเพิ่มผลผลิตแตงกวา โดยให้ผลผลิตสูงกว่าที่ไม่พ่นจนแตกต่างกันอย่างมีนัยสำคัญยิ่ง การปลูกแตงกวาโดยไม่ใช้ฮอร์โมน NAA เลยได้ผลผลิตเพียง 343.35 กรัม เมื่อพ่น NAA 3 – 6 ครั้ง ช่วยให้ผลผลิตสูงขึ้นถึง 519.05 และ 612.10 กรัม ตามลำดับแต่เมื่อพ่น 9 ครั้ง กลับทำให้ผลผลิตลดลง

เปรียบเทียบประสิทธิภาพกับ CRD (Relative Efficiency)

การทดลองแบบ RCB สามารถเปรียบเทียบประสิทธิภาพกับ CRD โดยคำนวณจากสูตร

$$RE = \frac{(r-1)E_b + r(t-1)E_e}{(rt-1)E_e}$$

เมื่อ E_b คือ Replication MS และ E_e คือ error MS ในตารางวิเคราะห์ของ RCB แต่ในกรณีที่ error df น้อยกว่า 20 ค่า RE จะต้องคูณค่าคงที่ k ตามสูตร

$$k = \frac{((r-1)(t-1)+1)(t(r-1)+3)}{((r-1)(t-1)+3)(t(r-1)+1)}$$

จากตัวอย่างข้างต้น มี error df เพียง 9 จึงต้องคำนวณดังนี้

$$\begin{aligned} RE_{adj} &= RE \times k \\ &= \frac{(r-1)E_b + r(t-1)E_e}{(et-1)E_e} \; \frac{\{(r-1)(t-1)+1\}\{t(r-1)+3\}}{\{(r-1)(t-1)+3\}\{t(r-1)+1\}} \\ &= \frac{3(2,664.1123) + 4(3)(988.6696)}{(4x4-1)(988.6696)} \; \frac{(3x3+1)(4x3+3)}{(3x3+3)(4x3+1)} \\ &= 1.29 \end{aligned}$$

แสดงให้เห็นว่าการทำ RCB มีประสิทธิภาพดีกว่า CRD 29%

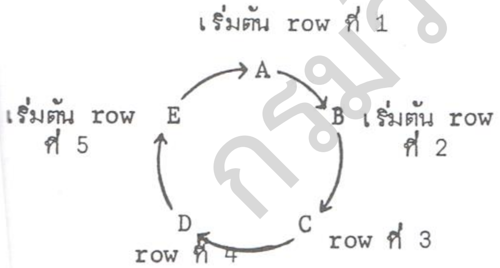
แผนการทดลองแบบละตินสแคสร์ (LS)

งานทดลองเกี่ยวกับการเกษตรที่ทำในไร่ มักพบว่ามีความแปรปรวนของหน่วยทดลองอยู่เสมอ ส่วนใหญ่ความแปรปรวนนั้นมีเพียงทิศทางเดียว จึงต้องทำการจัดบล็อคให้ด้านยาวของบล็อคตั้งฉากกับทิศทางของความแปรปรวนนั้น แต่บางครั้งความแปรปรวนอาจเป็นได้ 2 ทิศทางตั้งฉากกัน เช่น พื้นที่มีความลาดเอียง 2 ทาง จึงจึงต้องทำบล็อคเป็น 2 ทิศทางตั้งฉากกับทิศทางของความแปรปรวนที่มีอยู่ คือ บล็อคทางแนวนอน (Row blocking) และบล็อคทางแนวตั้ง (Column blocking) การบล็อคเช่นนี้เป็นหลักการสำคัญในการวางแผนในการทดลองแบบละตินสแควร์ แต่ละบล็อคในแนวนอนหรือแนวตั้งจะต้องมีครบทุกทรีตเมนต์ ดังนั้นจึงเป็นข้อบังคับของ LS ว่า จำนวนทรีตเมนต์ จำต้องเท่ากับจำนวน Row และจำนวน Column ถ้ามีจำนวนทรีตเมนต์การทดลองก็จะใหญ่เพราะจะต้องทำซ้ำมากให้เท่ากันตรงข้ามถ้าจำนวนทรีตเมนต์
น้อย ก็จะมีความน่าเชื่อถือน้อยด้วย จำนวนทรีตเมนต์ที่เหมาะสมจึงไม่ควรน้อยกว่า 4 และไม่เกิน 8

การสุ่มและแผนผังการทดลอง

สมมุติต้องการเปรียบเทียบ 5 ทรีตเมนต์ A B C D E ในพื้นที่ที่มีความลาดเอียง 2 ทิศทางตั้งฉากกัน การวางแผนผังการทดลอง จะต้องทำเป็นแปลงรูปร่างสี่เหลี่ยมจัตุรัสที่มีแปลงย่อยเป็น 5 x 5 แปลงมีขั้นตอนดังนี้ คือ

1. **ทำแบบมาตรฐาน** (Standard plan) โดยการใช้แผนผังวงกลมเรียงตัวอักษร A B C D E ไม่ให้ซ้ำกันทุกแถว ไม่ว่าแถวนอนหรือแถวตั้ง ทุกแถวจะมีตัวอักษรครบทั้ง 5 ตัว ถ้าเขียนตามแถวนอน ใน Row ที่ 1 เริ่มต้นด้วยอักษร A ตามด้วย B C ถึง E ใน Row ที่ 2 จะเริ่มต้นด้วยอักษร B แล้วตาม C D E จบด้วย A ทำเช่นนี้ไปจนถึง Row ที่ 5 ซึ่งจะเริ่มต้นอักษร E แล้วตามด้วย A ไปจบที่ D จะได้แผนผังซึ่งเรียกว่าแบบมาตรฐานดังนี้



row ที่ 1	A	B	C	D	E
row ที่ 2	B	C	D	E	A
row ที่ 3	C	D	E	A	B
row ที่ 4	D	E	A	B	C
row ที่ 5	E	A	B	C	D

การสุ่ม

การจัดตัวอักษรตามแบบมาตรฐาน ทุกแถวไม่ว่าแถวนอนหรือแถวตั้งต่างก็มีครบทุกทรีตเมนต์ไม่ซ้ำกัน แต่เรียงกันเป็นระเบียบ ทรีตเมนต์เดียวกันจะอยู่ในเส้นทแยงมุมจากล่างซ้ายไปบนขวา ไม่มีการสุ่ม จึงต้องทำการสุ่มด้วยวิธีการต่อไปนี้

1. กำหนดเลขประจำ Row เป็น 1, 2, 3, 4, 5 แล้วทำการสุ่ม Row หรือ Block ตามแนวนอน โดยวิธีจับฉลากสมมุติได้เลข 3, 5, 2, 1, 4 ตามลำดับ ให้สลับ Row เอา Row ที่ 3 ขึ้น เรียงไว้ด้านบนถัดลงมา คือ Row ที่ 5, 2, 1 และ 4 จะได้แผนผังใหม่ ดังนี้

row					
3	C	D	E	A	B
5	E	A	B	C	D
2	B	C	D	E	A
1	A	B	C	D	E
4	D	E	A	B	C

2. กำหนดเลขประจำ Column กับแผนผังใหม่เป็น 1, 2, 3, 4, 5 แล้วทำการสุ่ม Column ด้วยวิธีเดียวกัน สมมุติได้ 2 1 4 3 5 แล้วเรียง Column ใหม่ต่อจากข้อ 1 ขึ้นต้นด้วย Column 2 แล้ว ตามด้วย Column 1, 4, 3, 5 ตามลำดับ จนได้แผนผังที่สมบูรณ์ ดังนี้

column	2	1	4	3	5
	D	C	A	E	B
	A	E	C	B	D
	C	B	E	D	A
	B	A	D	C	E
	E	D	B	A	C

แผนผังการทดลอง

จัดแปลงทดลองตามแผนผังข้างต้นโดยแบ่งพื้นที่แปลงทั้งหมดเป็น 5 x 5 แปลงย่อย คือ

ลาดเอียง 1

ลาดเอียง 2

	1	2	3	4	5
1	D	C	A	E	B
2	A	E	C	B	D
3	C	B	E	D	A
4	B	A	D	C	E
5	E	D	B	A	C

การวิเคราะห์ผลการทดลอง

1. การวิเคราะห์ความแปรปรวน

เนื่องจากการทดลองแบบ LS มีแหล่งความแปรปรวนทั้งหมด 4 แหล่ง คือ Row Column ทรีตเมนต์ และความคลาดเคลื่อนจากการปฏิบัติงานทดลอง (Experimental Error) จึงต้องเตรียมตารางข้อมูลเพื่อการคำนวณ 2 ตาราง โดยกรอกข้อมูลที่ได้ในแผนผังการทดลองที่ทำจริง 1 ตาราง คำนวณผลรวมทาง Row , Column และผลรวมทั้งหมด และตารางแสดงผลรวม ค่าเฉลี่ยของทรีตเมนต์อีก 1 ตาราง

<u>ตัวอย่าง</u> การปรับปรุงคุณภาพของวัสดุที่มีอยู่โดยเพิ่มลดปริมาณธาตุอาหารในวัสดุผลิตเห็ดฟาง

ทรีตเมนต์ : วัสดุเห็ดฟาง 4 ตำรับ

แผนผังการทดลอง :

row \ column	1	2	3	4
1	4	2	1	3
2	2	4	3	1
3	3	1	2	4
4	1	3	4	2

ขั้นตอนการวิเคราะห์

1) กรอกข้อมูลการทดลองในแผนผังการทดลองและคำนวณผลรวมของ Row Column ผลรวมทั้งหมด และทำตารางผลรวม ค่าเฉลี่ยของแต่ละทรีตเมนต์

ตารางแสดงน้ำหนักเห็ดฟางจากการใช้วัสดุในการเพาะตำรับต่าง ๆ (กก.)

Row/column(t_r)	1	2	3	4	รวม
1	3.275 (4)	1.890 (2)	3.886 (1)	2.470 (3)	11.521
2	2.120 (2)	4.800 (4)	2.615 (3)	2.960 (1)	12.495
3	3.805 (3)	3.750 (1)	3.725 (2)	6.010 (4)	17.290
4	5.000 (1)	3.975 (3)	5.995 (4)	4.450 (2)	19.420
รวม	14.200	14.415	16.221	15.890	60.726

ทรีตเมนต์	ผลรวม	ค่าเฉลี่ย
1	15.596	3.899
2	12.185	3.046
3	12.865	3.216
4	20.080	5.020
		$\overline{X}$ = 3.7954

2) คำนวณค่า Degree of freedom (df)

กำหนดให้จำนวนทรีตเมนต์ = จำนวน row = จำนวน Column = t

Total df = $t^2 - 1 = 4^2 - 1 = 15$

Row df = Column df = Treatment df = t - 1

= 4 - 1 = 3

Error df = (t – 1) (t – 2) = (4 – 1) (4 – 2) = 6

3) คำนวณค่าความแปรปรวนของแหล่งต่าง ๆ

Sum of Square (SS)

$$\text{C.F} = \frac{60.726^2}{16} = 230.4779$$

$$\text{Total SS} = (3.275^2 + 1.890^2 + \ldots + 4.450^2) - \text{C.F.} = 23.2353$$

$$\text{Row SS} = \frac{(11.521^2 + \ldots + 19.420^2)}{4} \quad \text{C.F.} = 10.7568$$

$$\text{Column SS} = \frac{(14.200^2 + \ldots + 15.890^2)}{4} \quad \text{C.F.} = 0.7833$$

$$\text{Treatment SS} = \frac{(15.596^2 + \ldots + 20.080^2)}{4} \quad \text{C.F.} = 9.6281$$

Error SS = Total SS-Row SS-Column SS-Treatment SS

= 23.2353-10.7568-0.7833-9.62 = 2.0671

Mean Square (MS)

$$\text{Row MS} = \frac{\text{Row SS}}{\text{Row df}} = \frac{10.7568}{3} = 3.5856$$

$$\text{Column MS} = \frac{\text{Column SS}}{\text{Column df}} = \frac{0.7833}{3} = 0.2611$$

$$\text{Treatment MS} = \frac{\text{Treatment SS}}{\text{Treatment df}} = \frac{9.6281}{3} = 3.2094$$

$$\text{Error MS} = \frac{\text{Error SS}}{\text{Error df}} = \frac{2.0671}{6} = 0.3445$$

$$F = \frac{\text{Treatment MS}}{\text{Error MS}} = \frac{3.2094}{0.3445} = 9.32$$

4) เปิดตาราง F ที่ df (t – 1) และ (t – 1) (t – 2) ได้ค่า F (3,6 df) ที่ 5% และ 1% เป็น 4.76, 9.78 ตามลำดับ

5) คำนวณค่า C.V. ใส่ท้ายตารางสรุปผลความแปรปรวน

$$\text{C.V.} = \frac{\sqrt{\text{Error MS}}}{\overline{X}} \times 100$$

$$= \frac{\sqrt{0.3445}}{3.7954} \times 100$$

$$= 15.46\%$$

6) กรอกค่าที่คำนวณได้ทั้งหมดลงในตารางสรุปผลความแปรปรวน (ANOV)

SOV	df	SS	MS	F-Value		
				Cal	5%	1%
Total	15	23.2353				
Row	3	10.7568	3.5856	10.41**		
Column	3	0.7833	0.2611	<1		
Treatment	3	9.6281	3.2094	9.32*	4.76	6.78
Error	6	2.0671	0.3445			

C.V. = 15.46%

2. การเปรียบเทียบค่าเฉลี่ยของทรีตเมนต์

เนื่องจากค่าของ F แสดง Significant ซึ่งหมายความว่าการทดลองนี้จะต้องมีวัสดุสำหรับเพาะอย่างน้อย 2 ตำรับ ที่ให้ผลผลิตของเห็ดแตกต่างกัน จึงต้องคำนวณค่าสถิติเพื่อใช้ตัดสินความแตกต่างของค่าเฉลี่ย แต่เนื่องจากทรีตเมนต์ที่กำหนดการทดลองมีเพียง 4 ทรีตเมนต์ จึงควรใช้ค่า LSD ที่ระดับความเชื่อมั่น 5% เป็นตัวตัดสิน ดังนี้

$$\text{LSD } 5\% = t_{.05}\sqrt{\frac{2\text{xError MS}}{t}}$$

$$= 2.447\sqrt{\frac{2\text{x}0.3445}{4}}$$

$$= 1.0156$$

การแปลความหมาย

จากการทดลองปรับปรุงคุณภาพวัสดุเพาะเห็ด 4 ตำรับ ปรากฏผลว่า การเพาะเห็ดด้วยวัสดุตำรับที่ 4 ให้ผลผลิตสูงสุดและแตกต่างจากตำรับอื่น ๆ อย่างมีนัยสำคัญทางสถิติ คือให้ผลผลิตเฉลี่ย 5.0 กก. ขณะที่ตำรับที่ 1, 2, 3 ให้ผลผลิตเฉลี่ย 3.9 3.0 และ 3.2 กก. ตามลำดับ ซึ่งผลผลิตจากทั้ง 3 ตำรับหลังนี้ใกล้เคียงกันมากจนไม่ถือว่ามีความแตกต่างกันทางสถิติ

เปรียบเทียบประสิทธิภาพของ LS กับ RCB และ CRD

กรณีที่หน่วยทดลองมีความแตกต่างกันด้วยสาเหตุอย่างเดียว เช่น มีความแปรปรวนไปในทิศทางเดียวกัน การจัดหน่วยทดลองให้เป็นกลุ่มทางวิธีการของการบล็อคในแผ่นการทดลองแบบ RCB ย่อมมีประสิทธิภาพสูงกว่าการใช้แผนการทดลองแบบ CRD แต่สำหรับการทดลองกับหน่วยทดลองที่มีหรือคาดว่ามีความแปรปรวนด้วย 2 สาเหตุ หรือไม่รู้ความแปรปรวนการใช้แผนการทดลองแบบ LS ซึ่งทำบล็อค 2 ทิศทาง คือ บล็อคทางแนวนอนและทางแนวตั้ง ก็ควรจะมีประสิทธิภาพสูงกว่าการไม่มีบล็อค หรือทำบล็อคด้านเดียว โดยทำการเปรียบเทียบประสิทธิภาพได้ดังขั้นตอน ต่อไปนี้

1. ก่อนเปรียบเทียบประสิทธิภาพของ LS กับ CRD หรือ RCB จะต้องตรวจดูค่าความแปรปรวนซึ่งเกิดขึ้นทางด้าน Row และ Column ก่อน

- คำนวณค่า F ของ Row และ Column

$$F\text{ (row)} = \frac{\text{Row MS}}{\text{Error MS}}$$

$$F\text{ (Column)} = \frac{\text{Column MS}}{\text{Error MS}}$$

จากตัวอย่าง $$F\text{ (row)} = \frac{3.5856}{0.3445} = 10.41$$

$$F\text{ (Column)} = \frac{0.2611}{0.3445} = <1$$

- เปรียบเทียบค่า F จากการคำนวณค่า F จากตารางที่ df ของ Row หรือ Column กับ df ของ Error จากตัวอย่าง Row df = Column df = 3 และ Error df = 6 เปิดค่า F ที่ 5% และ 1% ได้ 4.76 และ 9.78 ตามลำดับ จากการคำนวณ F (Row) มีค่ามากกว่า F จากตารางทั้ง 5% และ 1% แต่ F (Column) มีค่าน้อยกว่า 1 จึงชี้ให้เห็นว่าการบล็อคทางแนวนอน ฝทำได้เหมาะสม สามารถลดความแปรปรวนของผลการทดลองได้แต่การบล็อคทางแนวตั้งไม่มีผล

2. คำนวณค่าเปรียบเทียบประสิทธิภาพของ LS กับ CRD และ RCB

- เปรียบเทียบ LS กับ CRD

$$RE\text{ (CRD)} = \frac{E_r + E_C + (t-1)E_e}{(t+1)E_e}$$

เมื่อ E_r = Row MS

E_c = Column MS

E_e = Error MS

จากตัวอย่าง $$RE\text{ (CRD)} = \frac{3.5856 + 0.2611 + 3(0.3445)}{5\text{x}0.3445} = 2.83$$

- เปรียบเทียบ LS กับ RCB

$$\text{RE (RCB,Col)} = \frac{E_c + (t-1)E_e}{t(E_e)}$$

$$\text{RE (RCB,Row)} = \frac{E_r + (t-1)E_e}{t(E_e)}$$

สำหรับการทดลองตามแผน LS ที่มีค่า Error df น้อยกว่า 20 จะต้องปรับค่าคงที่ k

ตามสูตร $$k = \frac{((t-1)(t-2)+1)((t-1)^2+3)}{((t-1)(t-2)+3)((t-1)^2+1)}$$

จากตัวอย่างมีค่า Error df = 6 จึงต้องคูณด้วยค่า k ดังนี้

$$\text{RE (RCB,Row)} = \frac{(3.5856 + 3(0.3445))}{(4(0.3445))}\frac{(3x2+1)(3^2+3)}{(3x2+3)(3^2+1)} = 3.13$$

$$\text{RE (RCB,Col)} = \frac{(0.2661 + 3(0.3445))}{(4(0.3445))}\frac{(3x2+1)(3^2+3)}{(3x2+3)(3^2+1)} = 0.88$$

ผลการเปรียบเทียบชี้ให้เห็นว่าการใช้แผน LS โดยทำบล็อคทางแนวนอนและแนวตั้งในการทดลองนี้มีประสิทธิภาพสูงกว่าการใช้แผนการทดลองแบบ CRD ถึง 183% ซึ่งหมายความว่าถ้าทำการทดลองแบบ CRD จะต้องทำถึง 11 ซ้ำ จึงจะทดสอบความแตกต่างของทรีตเมนต์ได้ผลเช่นเดียวกับใช้แผนแบบ LS นอกจากนี้ยังชี้ให้เห็นอีกด้วยว่า ถ้าใช้แผนการทดลองแบบ LS โดยทำบล็อคทางแนวนอนเพิ่มขึ้นในการทดลองเดิมที่เป็น RCB และมี Column เป็น Block เพียงอย่างเดียว สามารถเพิ่มประสิทธิภาพการทดลองขึ้นได้อีก 213% แต่ตรงกันข้ามถ้าเดิมทำ RCB ที่มีบล็อคทางแนวนอนอยู่แล้วไม่จำเป็นต้องทำ LS โดยบล็อคทางแนวตั้งอีก เพราะไม่ได้ช่วยให้เพิ่มประสิทธิภาพขึ้นแต่อย่างใด

การวางแผนงานทดลอง
สำหรับหลายปัจจัย

เสาวนีย์ พิสิฏฐพันธ์

การทดสอบมากกว่าปัจจัยเดียว

สิ่งมีชีวิตทุกชนิดไม่ว่าจะเป็นสัตว์หรือพืช มีการตอบสนองต่อสิ่งแวดล้อมต่างๆ ไม่เหมือนกัน ดังนั้นการทดลองที่ทำเกี่ยวกับปัจจัยเดียวในบางครั้งจึงไม่เพียงพอ เพราะว่าการตอบสนองของสิ่งมีชีวิตต่อปัจจัยที่ทดสอบจะเปลี่ยนแปลงไปถ้ามีปัจจัยอื่นๆ มาเกี่ยวข้องด้วย การทำการทดสอบมากกว่าปัจจัยเดียวพร้อมๆ กันจะช่วยให้สามารถทราบเรื่องราวของสิ่งที่กำลังศึกษาให้กว้างขวางขึ้น

ปัจจัยต่างๆ ในทางเกษตรแบ่งได้เป็น 2 พวก

1. ปัจจัยเชิงคุณภาพ (Qualitative factors) หมายถึงปัจจัยที่แต่ละระดับไม่มีความต่อเนื่องกัน เมื่อนำระดับเหล่านั้นมาเรียงกันบนแกนแนวนอน จะสลับที่กันได้ นั่นคือตำแหน่งบนแกนนอนไม่ใช่ค่าของระดับนั้นๆ อย่างแท้จริง แต่เป็นตัวแทนของระดับหนึ่งๆ เพียงชั่วขณะนั้น เช่น

- พันธุ์ข้าว กข1 ขาวดอกมะลิ 105....................
- พันธุ์ถั่วเหลือง สจ1 สจ2 สจ4....................
- สารเคมีกำจัดแมลง ฟูราดาน คารโบฟูราน........................
- ปุ๋ยชนิดต่างๆ แอมโมเนียมซัลเฟต ยูเรีย...........................

  ฯลฯ

2. ปัจจัยเชิงปริมาณ (Quantitative factors) หมายถึงปัจจัยที่แต่ละระดับมีความต่อเนื่องกัน ตำแหน่งบนแกนแนวนอนคือ ค่าของระดับนั้นๆ ไม่อาจจะสลับที่กันได้ เช่น

- ปุ๋ยไนโตรเจน 6, 12, 18 และ 24 กก./ไร่
- ฟูราดาน อัตรา 2, 3, 4 และ 8 กก./ไร่
- การปลูกพืชโดยใช้ระยะห่างระหว่างแถว 50 ซม. และใช้ระยะห่างระหว่างต้นในแถวเป็น 10, 15, 20 และ 25 ซม.
- การปลูกพืชจำนวน 25, 36, 80 ต้น/ไร่

ให้สังเกตว่าปัจจัยเชิงปริมาณนี้อาจจะมีระยะห่างระหว่างระดับ (space) ต่างๆ เท่ากันหรือไม่เท่ากันได้

การศึกษาสองปัจจัยพร้อมๆ กัน เช่น

ปัจจัยที่ 1 คือ พันธุ์ข้าว 2 พันธุ์ ได้แก่ พันธุ์ กข5 และกข7

ปัจจัยที่ 2 คือ ปุ๋ยไนโตรเจน 2 ระดับ ได้แก่ ไม่ใส่ปุ๋ยและใส่ปุ๋ย 6 กก./ไร่

การรวมตัวระหว่างแต่ละระดับของ 2 ปัจจัย เป็น 4 แบบ ดังนี้

กข5 ไม่ใส่ปุ๋ย

กข5 ใส่ปุ๋ย 6 กก./ไร่

กข7 ไม่ใส่ปุ๋ย

กข7 ใส่ปุ๋ย 6 กก./ไร่

ทั้ง 4 แบบนี้เป็น 4 Treatment ที่จะทำการเปรียบเทียบ แต่เรียกชื่อใหม่ว่า Treatment Combination หรือ Factorial Combination และทั้งสองปัจจัยนี้ถือว่าเป็น factorial ซึ่งกันและกัน

การทำการศึกษาสองปัจจัยในการทดลองเดียวกัน นอกจากจะศึกษาถึงผลการตอบสนองของแต่ละปัจจัย (main effect) ที่ทดสอบแล้วยังสามารถจะศึกษาถึง Interaction ระหว่างปัจจัยทั้งสองนั้นด้วย ทั้งสองปัจจัยไม่มี Interaction ต่อกันเมื่อผลการตอบสนองต่อปัจจัยที่ 1 เหมือนกันไม่ว่าปัจจัยที่สองจะเปลี่ยนแปลงไปอย่างไร แต่ถ้าการเปลี่ยนแปลงของปัจจัยที่สองทำให้การตอบสนองต่อปัจจัยที่ 1 เปลี่ยนแปลงไป แสดงว่าทั้งสองปัจจัยมี Interaction ต่อกัน

<u>ตาราง 1</u> ค่าสมมติเพื่อแสดงการตอบสนองของพันธุ์ข้าว กข5 และ กข7 ต่อปุ๋ยไนโตรเจน

ผลผลิต ถัง/ไร่

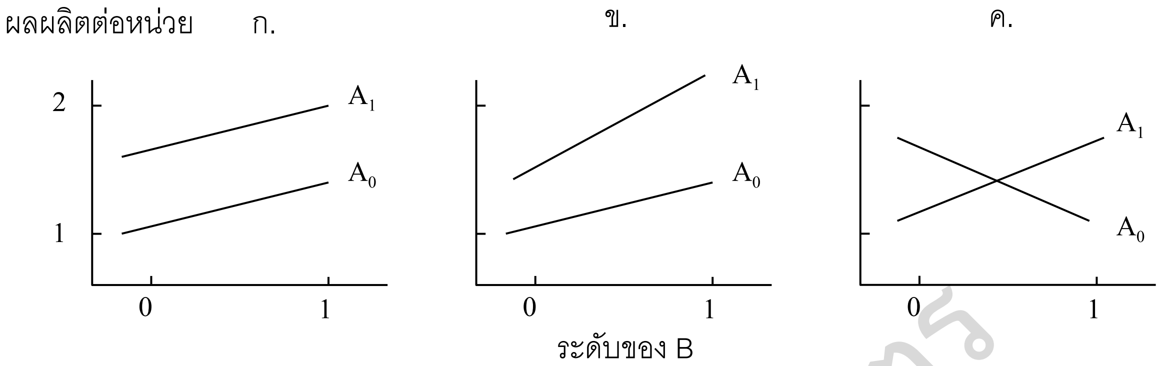
พันธุ์	ไนโตรเจน กก./ไร่		เฉลี่ย
	0	6	
ก. <u>ไม่มี Interaction</u>			
กข5	10	30	20
กข7	20	40	30
เฉลี่ย	15	35	
ข. <u>มี Interaction</u>			
กข5	10	30	20
กข7	20	20	20
เฉลี่ย	15	25	

ในตาราง 1 ก. จะเห็นว่า การเพิ่มปุ๋ยไนโตรเจน 6 กก./ไร่ จะให้ผลผลิตเพิ่มขึ้น 20 หน่วย ทั้ง กข5 และ กข7 หรือถ้าจะพิจารณาอีกด้านหนึ่ง คือ พันธุ์ กข7 จะให้ผลผลิตสูงกว่า กข5 10 หน่วย ไม่ว่าจะใช้ปุ๋ยหรือไม่ แสดงว่าทั้งพันธุ์และปุ๋ยไม่มี Interaction ต่อกัน ในกรณีเช่นนี้ สามารถจะเปรียบเทียบการตอบสนองของปัจจัยใดปัจจัยหนึ่งได้โดยใช้ค่าเฉลี่ย นั่นคือ สามารถศึกษา main effect ได้

จากตาราง 1 ข. จะเห็นว่า การเพิ่มปุ๋ยไนโตรเจน 6 กก./ไร่ ทำให้ผลผลิตของ กข5 เพิ่มขึ้น 10 หน่วย แต่ปุ๋ยจำนวนเดียวกัน ไม่สามารถเพิ่มผลผลิตของ กข7 ได้ ดังนั้น ถ้าเราสรุปตามค่าเฉลี่ยของพันธุ์ทั้งสองว่าปุ๋ย 6 กก./ไร่ เพิ่มผลผลิตได้ 10 หน่วย จึงไม่เป็นความจริงสำหรับ

พันธุ์ทั้งคู่ ในทำนองเดียวกัน ถ้าเปรียบเทียบผลผลิตของทั้งสองพันธุ์ เมื่อไม่มีการใส่ปุ๋ย จะเห็นว่าพันธุ์ กข7 ให้ผลผลิตสูงกว่า กข5 อยู่ 10 หน่วย แต่ถ้าใช้ปุ๋ย 6 กก./ไร่ พันธุ์ที่ให้ผลผลิตสูงกว่ากลับเป็นพันธุ์ กข5 แสดงว่าการตอบสนองต่อปุ๋ยของพันธุ์ทั้งสองไม่เหมือนกัน ดังนั้นจะสรุปผลตาม main effect ไม่ได้

เพื่อจะให้เข้าใจ Interaction ได้อย่างชัดเจนจึงใช้อธิบายด้วยรูปภาพ ดังนี้



ก. ไม่มี Interaction การเพิ่มของ A_0 เมื่อ B เปลี่ยนจาก 0 เป็น 1 เหมือนกับการเพิ่มของ A_1

ข. มี Interaction เล็กน้อย การเพิ่มของ A_1 เมื่อ B เปลี่ยนจาก 0 เป็น 1 มากกว่าการเพิ่มของ A_0 เมื่อ B เปลี่ยนแบบเดียวกัน

ค. มี Interaction มาก เมื่อเปลี่ยน B จาก 0 เป็น 1 A_1 เท่านั้นที่เพิ่ม ส่วน A_0 จะลดลง

ถึงแม้ว่าการศึกษาอิทธิพลของหลายปัจจัยพร้อมๆ กัน รวมทั้ง Interaction ของแต่ละปัจจัยจะเป็นประโยชน์อย่างยิ่งสำหรับการทดลองด้านเกษตร แต่ก็มีปัญหากล่าวคือ นอกจากจะเกิดความยุ่งยากของการตีความหมายและแปลผลการทดลองมากกว่าการทดลองปัจจัยเดียวแล้ว อุปสรรคอีกประการหนึ่งคือ จำนวน Combination จะเพิ่มขึ้นอย่างรวดเร็ว ในรูปของผลคูณ ทำให้ต้องทำการทดลองขนาดใหญ่เป็นการสิ้นเปลืองงบประมาณและขนาดการทดลองที่ใหญ่มาก อาจจะทำให้ดูแลการทดลองไม่ได้ดีเท่าที่ควร การทำการศึกษาหลายปัจจัยจึงควรจะทำเมื่อปัจจัยเหล่านั้นได้รับการศึกษามาแล้วครั้งหนึ่งก่อน และเลือกเฉพาะระดับที่มีความสำคัญจริงๆ เพื่อให้จำนวนของ Treatment Combination พอเหมาะกับการทำการทดลอง

การทดลองแฟคตอเรียล 2 ปัจจัย (2-Factors Factorial Experiment)

เมื่อผู้ทดสอบมีความสนใจในปัจจัยทั้งสองเท่าๆ กันต้องการทดสอบด้วยความเที่ยงตรงเท่าๆ กัน หรือมีความสะดวกในการปฏิบัติเกี่ยวกับการทดลองไม่ต้องลงทุนจนมากเกินไป ก็อาจจะใช้การทดลองชนิดนี้ได้ แต่ก่อนอื่นต้องเข้าใจเสียก่อนว่า Factorial Experiment หมายถึง การทดลองที่มีกลุ่มของ treatment ที่ทำการทดลองเป็นการรวมตัวของปัจจัยที่ต้องการทดสอบ ตามที่อธิบายไว้ข้างต้น แต่ไม่ได้บ่งบอกถึงลักษณะของการจัด treatment เมื่อเวลาทำการทดลอง นั่นคือ

ไม่ได้บอกให้ทราบว่าเป็นแผนการทดลองแบบใด ดังนั้นจึงต้องทำการจัดกลุ่มของ treatment ที่ต้องการทดลองให้เข้าลักษณะของแผนการทดลองพื้นฐานอันใดอันหนึ่งใน 3 แบบ ได้แก่

Factorial Experiment ที่จัดในรูป Completely Randomized Design (CRD)

Factorial Experiment ที่จัดในรูป Randomized Complete Block Design (RCB)

Factorial Experiment ที่จัดในรูป Latin Square Design (LTS)

การจัดรูปของ Factorial Experiment จะเป็นแบบใดขึ้นอยู่กับคุณสมบัติของหน่วยทดลองเหมือนกับใน Design พื้นฐานดังที่ได้กล่าวมาแล้วข้างต้น แต่ในงานวิจัยทางเกษตร Factorial Experiment ที่จัดในรูป RCB จะเป็นแบบที่นิยมใช้มากที่สุด

<u>ตัวอย่าง 1</u> การทดลองแฟคทอเรียล 2 ปัจจัยที่วางแผนแบบ RCB

เรื่อง การศึกษาช่วงระยะเวลาปลูกข้าวสำหรับนาปรังในท้องที่ที่มีสภาพหนาวเย็นในภาคเหนือ

วัตถุประสงค์ เพื่อหาช่วงเวลาเหมาะสมในการปลูกข้าวนาปรังในท้องที่ที่มีสภาพหนาวเย็นที่จะได้ผลผลิตสูงสุดของพันธุ์ต่างๆ

สถานที่ทำการทดลอง สถานีทดลองข้าวแพร่

ฤดูปลูก พฤศจิกายน 2523 – กุมภาพันธ์ 2524

ปัจจัยที่ทดสอบ - ปัจจัย A พันธุ์ข้าว 2 พันธุ์

A_1 กข7

A_2 BKN 6625-109-4-3-1-3

- ปัจจัย B เวลาปักดำ 6 เวลา

B_1 ปักดำรุ่นที่ 1 (25 พ.ย. 2523)

B_2 ปักดำรุ่นที่ 2 (10 ธ.ค. 2523)

B_3 ปักดำรุ่นที่ 3 (25 ธ.ค. 2523)

B_4 ปักดำรุ่นที่ 4 (9 ม.ค. 2524)

B_5 ปักดำรุ่นที่ 5 (26 ม.ค. 2524)

B_6 ปักดำรุ่นที่ 6 (10 ก.พ. 2524)

Treatment ที่ทำการทดสอบ

$T_1 = A_1B_1$: กข7 ปักดำรุ่นที่ 1 (กข7/1)

$T_2 = A_1B_2$: กข7 ปักดำรุ่นที่ 2 (กข7/2)

$T_3 = A_1B_3$: กข7 ปักดำรุ่นที่ 3 (กข7/3)

$T_4 = A_1B_4$: กข7 ปักดำรุ่นที่ 4 (กข7/4)

$T_5 = A_1B_5$: กข7 ปักดำรุ่นที่ 5 (กข7/5)

$T_6 = A_1B_6$: กข7 ปักดำรุ่นที่ 6 (กข7/6)

T_7 = A_2B_1 : BKN 6625 - ปักดำรุ่นที่ 1 (6625/1)

T_8 = A_2B_2 : BKN 6625 - ปักดำรุ่นที่ 2 (6625/2)

T_9 = A_2B_3 : BKN 6625 - ปักดำรุ่นที่ 3 (6625/3)

T_{10} = A_2B_4 : BKN 6625 - ปักดำรุ่นที่ 4 (6625/4)

T_{11} = A_2B_5 : BKN 6625 - ปักดำรุ่นที่ 5 (6625/5)

T_{12} = A_2B_6 : BKN 6625 - ปักดำรุ่นที่ 6 (6625/6)

ขนาดแปลงปลูก 3x4 เมตร

ระยะปลูก 25x25 ซม. ปักดำ 3 ต้น/กอ

ขนาดพื้นที่เก็บเกี่ยว 2x3 เมตร

วางแผนการทดลองแบบ Factorial in RCB 3 ซ้ำ

แผนผังการทดลอง (Lay out)

T_1	T_{12}	T_7	T_3	T_9	T_{11}	Block 1
T_8	T_4	T_5	T_2	T_6	T_{10}	
T_{10}	T_9	T_{12}	T_2	T_3	T_5	Block 2
T_{11}	T_1	T_6	T_7	T_4	T_8	
T_9	T_3	T_{12}	T_{11}	T_2	T_1	Block 3
T_{10}	T_6	T_5	T_4	T_7	T_8	

ข้อสังเกต : ใช้หลักการสุ่ม เช่นเดียวกับ RCB

1. ตารางข้อมูล ผลผลิตข้าว (กก./ไร่)

ตารางที่ 1

Treatment	ปัจจัย A	ปัจจัย B	Block 1	2	3	ผลรวม
1	กข7	รุ่นที่ 1	312	335	293	940
2		รุ่นที่ 2	368	324	368	1060
3		รุ่นที่ 3	382	271	265	918
4		รุ่นที่ 4	531	438	493	1462
5		รุ่นที่ 5	566	630	546	1742
6		รุ่นที่ 6	493	540	508	1541
7	BKN 6625	รุ่นที่ 1	431	359	416	1206
8		รุ่นที่ 2	466	432	486	1384
9		รุ่นที่ 3	479	478	414	1371
10		รุ่นที่ 4	613	584	534	1731
11		รุ่นที่ 5	649	462	508	1619
12		รุ่นที่ 6	645	522	474	1641
ผลรวม			5935	5375	5305	16615

ตารางที่ 2

A	B 1	2	3	4	5	6	ผลรวม (A)
กข7	940	1060	918	1462	1742	1541	7663
BKN 6625	1206	1384	1371	1731	1619	1641	8952
ผลรวม (3)	2146	2444	2289	3193	3361	3182	16615

ข้อสังเกต : การเตรียมข้อมูล เพื่อส่งวิเคราะห์ด้วยเครื่องคอมพิวเตอร์ เตรียมเฉพาะข้อมูลดิบของตารางที่ 1 โดยไม่ต้องมีผลรวมทั้งด้านแนวตั้งและแนวนอน

2. คำนวณ Degree of freedom (df)

ระดับปัจจัย A	=	2		
ระดับปัจจัย B	=	6		
จำนวน Block	=	3		
Total df	=	(2x6x3) – 1	=	35
Block df	=	3 – 1	=	2
A df	=	2 – 1	=	1
B df	=	6 – 1	=	5
(A x B) df	=	(2-1)(6-1)	=	5
Error df	=	(12-1)(3-1)	=	22

3. คำนวณค่า Sum of Square (SS)

จากตารางที่1

Correction factor (C.F.) $= \frac{16615^2}{2x6x3}$

$= 7,668,284.03$

Total SS $= 312^2+335^2+\ldots+474^2 - C.F.$

$= 8,050,641 - 7,668,284.03$

$= 382,356.97$

Block SS $= \frac{5935^2 + 5375^2 + 5305^2}{2x6} - C.F.$

$= 7,688,156.25-7,668,284.03$

$= 19,872.22$

Treatment SS $= \frac{940^2 + 1060^2 + ... + 1641^2}{3} - C.F.$

$= 7,985,516.34 - 7,668,284.03$

$= 317,232.31$

จากตารางที่ 2

$$\text{A SS} = \frac{7663^2 + 8952^2}{6\text{x}3} - \text{C.F.}$$

= 7,714,437.39 – 7,668,284.03

= 46,153.36

$$\text{B SS} = \frac{2146^2 + 2444^2 + ... + 3182^2}{2\text{x}3} - \text{C.F.}$$

= 7,905,777.84 – 7,668,284.03

= 237,493.81

A x B SS = 317,232.31 - 46,153.36 - 237,493.81

= 33,585.14

Error SS = 382,356.97 - 19,872.22 - 317,232.31

= 45,252.44

4. คำนวณค่า Mean Square

$$\text{Block MS} = \frac{19{,}872.22}{2}$$

= 9936.11

A MS = 46,153.36

$$\text{B MS} = \frac{237{,}493.81}{5}$$

= 47,498.76

$$\text{A x B MS} = \frac{33{,}585.14}{5}$$

= 6,717.03

$$\text{Error MS} = \frac{45{,}252.44}{22}$$

= 2,056.93

5. คำนวณค่า F

$$F_{Block} = \frac{9,936.11}{2,056.93} = 4.83$$

$$F_{A} = \frac{46,153.36}{2,056.93} = 22.43$$

$$F_{B} = \frac{47,498.76}{2,056.93} = 23.09$$

$$F_{AB} = \frac{6,717.03}{2,056.93} = 3.26$$

6. เปิดค่า F จากตาราง F ดังนี้

F_{Block} ; df ของ Error = 22(f_2) df ของ Block = 2(f_1)

$F_{.05}$ = 3.44

$F_{.01}$ = 5.72

F_{A} ; df ของ Error = 22(f_2) df ของ A = 1(f_1)

$F_{.05}$ = 4.30

$F_{.01}$ = 7.94

F_{B} ; df ของ Error = 22(f_2) df ของ B = 5(f_1)

$F_{.05}$ = 2.66

$F_{.01}$ = 3.99

F_{AB} ; เหมือน F ของ B เพราะว่าค่า df ของทั้งคู่เท่ากัน

7. การเปรียบเทียบค่า F ที่คำนวณได้กับค่า F ในตาราง

ค่า F ที่คำนวณได้ของ Block และของ A x B มีค่าสูงกว่า $F_{.05}$ จากตาราง แต่ไม่สูงกว่า .01 ให้ใส่ * ลงบนค่า F ที่คำนวณได้

ค่า F ที่คำนวณได้ของ A และ B มีค่าสูงกว่า $F_{.01}$ จากตาราง ดังนั้นให้ใส่ **ลงบนค่า F ที่คำนวณได้

8. คำนวณค่า C.V.

$$\text{Mean} = \frac{16{,}615}{36}$$

$$= 461.5$$

$$\text{C.V.} = \frac{\sqrt{2{,}056.93}}{461.5} \times 100$$

$$= 9.8\%$$

9. ตารางวิเคราะห์แวเหรียนซ์ของผลผลิต (กก./ไร่)

Source of Variation	Degree of freedom	Sum of Square	Mean Square	F
Total	35	382,357		
Block	2	19,872	9,936.11	4.83*
A	1	46,153	46,153.36	22.43**
B	5	237,494	47,498.76	23.09**
A x B	5	33,585	6,717.03	3.26*
Error	22	45,252	2,056.93	

C.V. = 9.8%

10. ทำตารางเสนอผลค่าเฉลี่ยของ Treatment จากตารางที่ 2 ในข้อ 1 คำนวณ

ค่าเฉลี่ย A $= \frac{7663}{6\text{x}3} =$ 426…………………

B $= \frac{2146}{2\text{x}3} =$ 358…………………

A x B $= \frac{940}{3} =$ 313…………………

ตารางค่าเฉลี่ย (เฉลี่ยจากค่าสังเกต 3 ซ้ำ)

ปัจจัย A	ปัจจัย B รุ่น						เฉลี่ย (A)
	1	2	3	4	5	6	
กข7	313	353	306	487	581	514	426
BKN 6625	402	461	457	577	540	547	497
เฉลี่ย (B)	358	407	381	532	560	530	

11. การคำนวณค่าสถิติเพื่อใช้เทียบความแตกต่างของค่าเฉลี่ย

ดูตารางวิเคราะห์แวเหรียนซ์จะเห็นว่า A x B มีค่า F* แสดงว่าในการศึกษาครั้งนี้จะต้องเปรียบเทียบผลผลิตของข้าวตาม A x B นั่นคือ ข้าวทั้งสองพันธุ์ต้องการเวลาปลูกที่เหมาะสมที่จะให้ผลผลติสูงสุดคนละเวลาหรือกล่าวอีกอย่างหนึ่งว่าบางช่วงเวลาเหมาะสำหรับใช้พันธุ์ กข7 แต่บางเวลา BKN 6625 เหมาะสมกว่า

ก. เพื่อศึกษาว่าพันธุ์ข้าวแต่ละพันธุ์จะให้ผลผลิตสูงที่สุดในช่วงเวลาใด จะเห็นว่ามีการทดลองถึง 6 ช่วงเวลา คือ 6 รุ่น ดังนั้นควรจะใช้ DMRT ในการตัดสินใจ

$$LSR = SSR.S_{\bar{x}}$$

$$S_{\bar{x}} = \sqrt{\frac{\text{Error MS}}{r}}$$

$$= \sqrt{\frac{2{,}056.93}{3}}$$

$$= 26.158$$

df ของ Error = 22

ระดับความเชื่อมั่น 95% หรือที่ความเป็นไปได้ .05

P	2	3	4	5	6
SSR_p	2.93	3.08	3.17	3.24	3.29
LSR_p	76.7	80.6	83.0	84.8	86.2

การเปรียบเทียบให้เปรียบค่าเฉลี่ยของรุ่นที่ปักดำต่างๆ ทีละพันธุ์ ดังนี้

รุ่นที่ปักดำ	กข7
1	313 c
2	353 c
3	306 c
4	487 b
5	581 a
6	514 ab

รุ่นที่ปักดำ	BKN 6625-109-4-3-1-3
1	402 b
2	461 b
3	457 b
4	577 a
5	540 a
6	547 a

ข. เพื่อศึกษาว่าในการปักดำแต่ละเวลาพันธุ์ใดจะให้ผลผลิตดีกว่ากัน เนื่องจากว่ามีเพียง 2 พันธุ์ จึงใช้ LSD ได้

(df ของ Error = 22)

$$LSD = t.S_{\overline{d}}$$

$$S_{\overline{d}} = \sqrt{\frac{2x2{,}056.93}{3}}$$

$$= 37.031$$

$$t_{.05} = 2.074$$

$$t_{.01} = 2.819$$

LSD.05 = 76.8

LSD.01 = 104.4

จะเขียนค่า LSD ใส่ไว้ใต้ตารางค่าเฉลี่ยก็ได้ แต่เพื่อให้เห็นชัดเจน อาจจะใช้เปรียบเทียบ ดังนี้

รุ่นปักดำ	กข7	BKN 6625	ค่าต่างกัน
1	313	402	89*
2	353	461	108**
3	306	457	151**
4	487	577	90*
5	581	540	41^{ns}
6	514	547	33^{ns}

ใช้ค่า LSD ที่คำนวณได้เปรียบเทียบกับค่าที่แตกต่าง ถ้าลบกันแล้วผลต่างที่ได้

น้อยกว่าค่า	LSD.05	ใส่ ns
มากกว่า	LSD.05	ใส่ *
มากกว่า	LSD.01	ใส่ **

12. ตารางเสนอผล อาจจะเสนอในรูป ดังนี้

การศึกษาช่วงระยะเวลาปลูกข้าวนาปรัง ในท้องที่ที่มีสภาพหนาวเย็นในภาคเหนือ แพร่ ฤดูแล้ง 2524

ตารางค่าเฉลี่ย ผลผลิต (กก./ไร่)

	รุ่นที่ปักดำ	กข7 1/	BKN 6625 1/	ค่าแตกต่าง 2/
1.	25 พ.ย. 23	313 c	402 b	89*
2.	10 ธ.ค. 23	353 c	461 b	108**
3.	25 ธ.ค. 23	306 c	457 b	151**
4.	9 ม.ค. 24	487 b	577 b	90*
5.	26 ม.ค. 24	581 a	540 a	41 ns
6.	10 ก.พ. 24	514 ab	547 a	33 ns

1/ ตัวเลขที่ตามหลังด้วยตัวอักษรเหมือนกันในแต่ละพันธุ์ ไม่แตกต่างกันทางสถิติ ใช้ DMRT ที่ระดับความเชื่อมั่น 95%

2/ * แตกต่างกันโดยเทียบกับ LSD.05

** แตกต่างกันโดยเทียบกับ LSD.01

ns ไม่แตกต่างกันทางสถิติ

การแปลผล

ข้าวพันธุ์ กข7 ช่วงปักดำที่ให้ผลผลิตดีที่สุด คือ ประมาณปลายเดือนมกราคม (581 กก./ไร่) ส่วนพันธุ์ BKN 6625 ปลูกได้ตั้งแต่ต้นมกราคม ถึง ต้นกุมภาพันธ์ โดยให้ผลผลิตตั้งแต่ 540 ถึง 577 กก./ไร่ ตั้งแต่ปลายมกราคมจะปลูกพันธุ์ใดในระหว่าง 2 พันธุ์ก็ได้ ผลผลิตจะไม่แตกต่างกัน แต่ระหว่างปลายเดือนพฤศจิกายน ถึงต้นเดือนมกราคม พันธุ์ BKN 6625 จะให้ผลผลิตดีกว่า กข7 จาก 89 ถึง 151 กก./ไร่

ตัวอย่าง 2 การทดลองการตอบสนองของข้าวต่อปุ๋ยไนโตรเจนและอินทรีย์วัตถุต่างๆ ดำเนินการทดลองโดยแผนกดินปุ๋ย

ปัจจัยที่ 1 คือ อินทรีย์วัตถุ 3 ชนิด ได้แก่ ไม่ใส่อินทรีย์วัตถุ ใส่ฟางข้าวและใส่ปุ๋ยเทศบาล

ปัจจัยที่ 2 คือ ปุ๋ยไนโตรเจน 4 ระดับ คือ 0, 6, 12 และ 18 กก./ไร่

Treatment ที่ 1 ใส่ฟางข้าวอย่างเดียว (R-N_0)

ที่ 2 ใส่ฟางข้าวและไนโตรเจน 6 กก./ไร่ (R-N_6)

ที่ 3 ใส่ฟางข้าวและไนโตรเจน 12 กก./ไร่ (R-N_{12})

ที่ 4 ใส่ฟางข้าวและไนโตรเจน 18 กก./ไร่ (R-N_{18})

ที่ 5 ปลูกข้าวโดยไม่ใส่ปุ๋ย (N_0)

ที่ 6 ใส่ไนโตรเจน 6 กก./ไร่ (N_6)

ที่ 7 ใส่ไนโตรเจน 12 กก./ไร่ (N_{12})

ที่ 8 ใส่ไนโตรเจน 18 กก./ไร่ (N_{18})

ที่ 9 ใส่ปุ๋ยเทศบาลอย่างเดียว (C-N_0)

ที่ 10 ใส่ปุ๋ยเทศบาลและไนโตรเจน 6 กก./ไร่ (C-N_6)

ที่ 11 ใส่ปุ๋ยเทศบาลและไนโตรเจน 12 กก./ไร่ (C-N_{12})

ที่ 12 ใส่ปุ๋ยเทศบาลและไนโตรเจน 18 กก./ไร่ (C-N_{18})

ทั้ง 12 Treatment นี้ ทำการทดลองแบบ Randomized Complete block 3 ซ้ำ ดังแผนผังการทดลองต่อไปนี้

Lay Out

Rep.I	Rep.II	Rep.III
R-N_0	N_{18}	N_{12}
N_6	R-N_6	N_0
C-N_{12}	N_6	R-N_{18}
N_{18}	C-N_{18}	C-N_{12}
R-N_6	C-N_0	R-N_0
N_{12}	N_{12}	C-N_6
C-N_6	R-N_{18}	C-N_{18}
R-N_{18}	N_0	N_6
C-N_{18}	R-N_{12}	C-N_0
N_0	C-N_6	R-N_{12}
R-N_{12}	C-N_{12}	N_{18}
C-N_0	R-N_0	R-N_6

C = City compost
R = Rice straw
} 800 กก./ไร่

ขนาดแปลงย่อย 4x5 ม.

ขนาดแปลงเก็บเกี่ยว 2x3 ม.

<u>ตารางที่ 1</u> ข้อมูลผลผลิต (กก./ไร่)

Organic Matter	Nitrogen Level	Rep. 1	Rep. 2	Rep. 3	Total
Rice Straw	0	541	595	519	1655
	6	703	704	654	2061
	12	763	844	721	2328
	18	764	801	725	2290
CK	0	520	503	537	1560
	6	597	582	603	1782
	12	686	696	701	2083
	18	774	765	678	2217
City Compost	0	473	389	427	1289
	6	647	628	565	1840
	12	770	600	681	2051
	18	850	768	823	2441

วิธีคำนวณ

C.F. = $(541+595+\ldots823)^2/3x4x3$

Total SS = $(541^2+595^2+\ldots823^2)$ – C.F.

Block SS = $\left((541+703+\ldots+850)^2+(595+704+\ldots+768)^2 +(519+654+\ldots+823)^2\right)/3x4$ – C.F.

Treatment SS = $(1655^2+2061^2+\ldots2441^2)/3$ – C.F.

Organic SS = $\left((1655+2061+\ldots+2290)^2+(1560+1782+\ldots+2217)^2 +(1289+1840+\ldots+2441)^2\right)/3x4$ – C.F.

Nitrogen SS = $\left((1655+1560+1289)^2+(2061+1782+1840)^2+\ldots +(2290+2217+2441)^2\right)/3x3$ – C.F.

O x N = Treatment SS – Organic SS – Nitrogen SS

Error = Total SS – Block SS – Treatment SS.

ตารางที่ 2 ตารางวิเคราะห์แวเหรียนซ์ของผลผลิต (กก./ไร่)

SOV	df	SS	MS	F
Total	35	488,526.97		
Block	2	8,599.06	4,299.53	2.45
Treatment	11	441,400.31	40,127.30	22.91**
Organic matter (O)	2	27,435.39	13,717.70	7.83**
Nitrogen (N)	3	378,894.53	126,298.17	72.12**
O x N	6	35,070.39	5,845.06	3.34*
Error	22	38,527.61	1,751.26	

C.V. = 6.3%

เพราะว่าผลของการวิเคราะห์ แสดงว่า Interaction มีนัยสำคัญทางสถิติ ดังนั้น จะต้องศึกษาการตอบสนองต่อปุ๋ยไนโตรเจนตามแต่ละชนิดของปุ๋ยอินทรีย์

ตารางที่ 3 Table of means (Kg/rai)

Organic Matter	Level of N (Kg/rai)				O - mean
	0	6	12	18	
Rice Straw	552	687	776	763	694
None (CK)	520	594	694	739	637
City Compost	430	613	984	814	635
N - Mean	500	631	718	772	

For Comparison of 2 Treatment mean LSD.05 = 71 Kg/rai

.01 = 96 Kg/rai

การแปลผล

จากตารางที่ 2 และ 3 แสดงให้เห็นว่า การใช้ฟางข้าวเป็นอินทรีย์วัตถุร่วมกับปุ๋ยไนโตรเจนจะทำให้ผลผลิตสูงกว่าการใช้ปุ๋ยไนโตรเจนอย่างเดียว ทุกระดับ แต่มีความแตกต่างทางสถิติเพียง 2 ระดับ คือ เมื่อใส่ N_6 (687 VS 594) และ N_{12} (776 VS 694) เท่านั้น เมื่อเทียบกับการใส่ปุ๋ยเทศบาลซึ่งให้ผลผลิตสูงกว่าการใส่ไนโตรเจนเพียงอย่างเดียว ที่ระดับ N_{18} (814 VS 739) ถ้าใส่ปุ๋ยเทศบาลอย่างเดียวโดยไม่ใส่ไนโตรเจน ผลผลิตจะลดลงถึง 90 กก./ไร่ (430 VS 520)

Trend Comparison

เนื่องจากปัจจัยที่ศึกษา คือ ไนโตรเจน เป็นปัจจัยเชิงปริมาณ จึงควรทำการศึกษาว่าการตอบสนองต่อปุ๋ยไนโตรเจนของข้าว ในสภาพแวดล้อมทั้ง 3 สภาพ คือ ใช้ไนโตรเจนล้วนๆ ใช้ปุ๋ยเทศบาลและใช้ฟางข้าว จะเป็นอย่างไร โดยวิธีแยกค่าของ Interaction SS ออกเป็นส่วนๆ (Partitioning Sum of Square) ดังแสดงไว้ในตารางที่ 4

<u>ตารางที่ 4</u> การตอบสนองต่อไนโตรเจนในปุ๋ยอินทรีย์แต่ละชนิด (Contribution of nitrogen to Linear, Quadratic and Cubic Sum Square)

Condition	Linear	Quadratic	Cubic
Rice Straw	78626**	16,428**	459
CK (N alone)	86033**	645	1009
City Compost	224115**	2,160	4489

วิธีคำนวณ

ใช้ค่า total จากตารางที่ 1 และเนื่องจากระยะระหว่างระดับเท่ากัน (Equal Space) จึงใช้ค่าสัมประสิทธิ์ Orthogonal Polynomial จากตารางได้

	N_0	N_6	N_{12}	N_{18}	
Rice Straw	1655	2061	2328	2290	(รวมจาก 3 ซ้ำ)
CK	1560	1782	2083	2217	
City Compost	1289	1840	2051	2441	
สัมประสิทธิ์ (n = 4)					
Linear (X_1)	-3	-1	+1	+3	ตัวหาร 20
Quadratic (X_2)	+1	-1	-1	+1	ตัวหาร 4
Cubic (X_3)	-1	+3	-3	+1	ตัวหาร 20

$$\text{Rice Straw N Linear} = \frac{\left(1655(-3)+2061(-1)+2328(+1)+2290(+3)\right)^2}{20x3}$$

$$= 78626.4$$

$$\text{N Quadratic} = \frac{\left(1655(+1)+2061(-1)+2328(-1)+2290(+1)\right)^2}{4x3}$$

$$= 16428.0$$

$$\text{N Cubic} = \frac{\left(1655(-1)+2061(+3)+2328(-3)+2290(+1)\right)^2}{20x3}$$

$$= 459.3$$

ฯลฯ

ข้อควรจำ ทั้ง L Q และ C Sum Square มี df ส่วนละ =1 ฉะนั้น Sum Square จะมีค่าเท่ากับ Mean Square การทำ Significant Test ใช้ค่า Error (22 df) เป็นตัวทดสอบแล้วเทียบกับตาราง F เหมือนปกติ จากผลการวิเคราะห์ในตารางที่ 4 และใช้ค่าเฉลี่ยในตารางที่ 3 ไปคำนวณสมการได้ดังนี้

1. เมื่อใช้ฟางข้าว ข้าวจะตอบสนองต่อ N เป็นรูป Quadratic

   $\hat{Y} = 549.20 + 30.53\ N – 1.0277\ N^2 (R^2 = 0.998)$

2. เมื่อใช้ N อย่างเดียว ไม่มีปุ๋ยอินทรีย์ การตอบสนองเป็นรูป Linear

   $\hat{Y} = 523.20 + 12.62\ N\ (r^2 = 0.981)$

3. ใช้ N ร่วมกับปุ๋ยเทศบาลการตอบสนองเป็นรูป Linear

   $\hat{Y} = 451.8 + 20.38\ N\ (r^2 = 0.972)$

Factorial ที่มีทรีตเมนต์เพิ่มเติม

Factorial Experiment ที่กล่าวมาทั้ง 2 ตัวอย่าง เป็น Complete factorial คือ ทดลองเฉพาะทรีตเมนต์ที่เกิดจากการรวมตัวของแต่ละระดับของสองปัจจัยที่ต้องการศึกษาเท่านั้น แต่บางครั้งอาจจะมีทรีตเมนต์อื่นๆ นอกเหนือจาก Complete factorial รวมอยู่ด้วยคือ เป็นทรีตเมนต์เพิ่มเติม ซึ่งส่วนมากมักจะเป็นทรีตเมนต์ที่เป็น Control หรือ Check เช่น การทดลองการตอบสนองต่อปุ๋ยไนโตรเจนและฟอสเฟตอย่างละ 3 ระดับ Factorial combination ที่เกิดขึ้นเท่ากับ 9 คือ N_1P_1 N_1P_2 N_1P_3 N_2P_1 N_2P_2 N_2P_3 N_3P_1 N_3P_2 และ N_3P_3 แต่ผู้ทดลองมักจะเพิ่ม N_0P_0 ไม่ใช้ปุ๋ยทั้ง 2 ชนิด ถือว่าเป็น check เพื่อใช้วัดความอุดมสมบูรณ์ของพื้นที่ทดลอง ถึงจะมีทรีตเมนต์เพิ่มเติมก็ยังคงนับว่าเป็น Factorial experiment และ N_0P_0 ก็จะถูกสุ่มรวมไปกับส่วนที่เป็น Complete factorial การเพิ่มเติมนี้อาจจะเพิ่มมากกว่าหนึ่งทรีตเมนต์ก็ได้ ถ้าใช้ทดลอง RCB กับตัวอย่างนี้ แผนผังการทดลองจะเป็นดังนี้

N_2P_2	N_3P_1
N_1P_1	N_3P_3
N_0P_0	N_1P_2
N_3P_2	N_2P_3
N_1P_3	N_2P_1

Block 1

N_1P_3	N_2P_3
N_3P_2	N_3P_1
N_1P_2	N_2P_1
N_3P_3	N_0P_0
N_2P_2	N_1P_1

Block 2

N_3P_2	N_1P_3
N_2P_1	N_0P_0
N_3P_1	N_2P_2
N_1P_2	N_1P_1
N_2P_3	N_3P_3

Block 3

3 x 3 + 1 Factorial Experiment ที่วางแผนแบบ RCB

ตารางวิเคราะห์แวเหรียนซ์

SOV	df	SS	MS	F
Total	29			
Block	2			
Treatment	9			
CK VS Treated	(1)			
N	(2)			
P	(2)			
N x P	(4)			
Error	18			

C.V. %

<u>ตัวอย่าง 3</u> โครงการวิจัยศึกษาผลตอบสนองของปุ๋ยไนโตรเจนและฟอสเฟตต่อการเจริญเติบโตและผลผลิตเส้นใยป่านรามี

ทำการทดลองเมษายน 2530 – ตุลาคม 2531 ที่ไร่กสิกร อำเภอเชียงคาน จังหวัดเลย

วัตถุประสงค์ เพื่อศึกษาผลการใช้ปุ๋ยไนโตรเจนและฟอสเฟตที่มีต่อการเจริญเติบโตและผลผลิตเส้นใยของป่านรามี

ปัจจัยที่ 1 ไนโตรเจน 3 ระดับ 0, 20, 40 กก./ไร่

ปัจจัยที่ 2 ฟอสเฟต 3 ระดับ 0, 10, 20 กก./ไร่

จำนวนทรีตเมนต์ = 3 x 3 + 1

(ทุกทรีตเมนต์ใส่โปแตช 20 กก./ไร่ ยกเว้น check)

ใช้แผนการทดลองแบบ RCB

แผนผังการทดลอง

40-0 26	0-0 27	CK 28	40-20 29	20-0 30	Rep.3
20-20 21	0-20 22	0-10 23	40-10 24	20-10 25	
0-10 16	20-0 17	20-10 18	CK 19	0-20 20	Rep.2
20-20 11	40-0 12	0-0 13	40-10 14	40-20 15	
20-0 6	0-10 7	0-20 8	20-20 9	40-20 10	Rep.1
20-10 1	CK 2	40-0 3	0-0 4	40-10 5	

ขนาดแปลงย่อย 4 x 6 เมตร ระยะปลูก .80 x .60 เมตร

ขนาดแปลงเก็บเกี่ยว 2.4 x 3.6 เมตร

ตารางที่ 1 ความสูงของป่านรามีขณะเก็บเกี่ยว (ซม.)

ไนโตรเจน	ฟอสเฟต	ซ้ำที่ 1	2	3
0	0	99	91	120
0	10	99	94	122
0	20	109	105	108
20	0	117	122	151
20	10	131	123	149
20	20	126	117	128
40	0	160	133	164
40	10	152	162	161
40	20	163	169	162
Check		114	94	108

การคำนวณค่าวิเคราะห์แวเหรียนซ์ คำนวณเหมือนแฟคตอเรียลในตัวอย่างที่ 2 ในส่วนของ Total SS Block SS Treatment และ Error SS แต่ส่วนของรายละเอียดในการแยกค่า Treatment SS ทำดังนี้

$$\text{C.F.}_{\cdot 1} = (99+91+\ldots+108)^2/3 \times 10$$

$$\text{C.F.}_{\cdot 2} = (99+91+\ldots+162)^2/3 \times 9$$

$$\text{CK VS Treated} = (114+94+108)^2/3 + \text{C.F.}_{\cdot 2} - \text{C.F.}_{\cdot 1}$$

$$\text{N SS} = \Big((99+91+120+99+\ldots108)^2+(117+\ldots+128)^2 +(160+\ldots+162)^2\Big)/3\text{x}3-\text{C.F.}_{\cdot 2}$$

$$\text{P SS} = \Big((99+91+120+117+122+151+160+133+164)^2 +(99+\ldots+161)^2+(109+\ldots+162)^2\Big)/3\text{x}3-\text{C.F.}_{\cdot 2}$$

$$\text{NxP SS} = \Big((99+91+120)^2+(99+94+122)^2+\ldots(163+169+162)^2\Big) /3-\text{C.F.}_{\cdot 2}-\text{N.SS}-\text{P.SS}$$

ตารางที่ 2 ตารางวิเคราะห์แวเหรียนซ์ความสูงขณะเก็บเกี่ยว (ซม.)

SOV	df	SS	MS	F
Total	29	17857		
Block	2	1359	679.50	
Treatment	9	14988	1665.34	19.85**
CK VS Treatment	1	1779	1779.00	21.20**
N	2	12784	6392.11	76.19**
P	2	83	41.33	<1
N x P	4	342	85.61	1.02^{ns}
Error	18	1510	83.89	

C.V. = 7.13%

ตารางที่ 3 ความสูงขณะเก็บเกี่ยว (ซม.)

ไนโตรเจน	ฟอสเฟต			ค่าเฉลี่ย N
	0	10	20	
0	103	105	107	105
20	130	134	124	129
40	152	158	165	158
ค่าเฉลี่ย P	129	133	132	131

ค่าเฉลี่ยของ Check 105

LSD (สำหรับเปรียบเทียบค่าเฉลี่ยของ N 2 ค่า) .05 = 9 ซม.

การแปลผล

จากตารางที่ 2 จะเห็นว่าทรีตเมนต์ที่ทดลองมีความแตกต่างกันแต่เมื่อแยกความแปรปรวนของทรีตเมนต์ออกมาเป็นส่วนๆ แล้วจะเห็นว่าความแตกต่างเนื่องมาจาก

1. CK VS Treated นั่นคือ การใช้ปุ๋ยทุกตำหรับจะทำให้ผลผลิตแตกต่างกับการไม่ใช้ปุ๋ย ดูจากตารางที่ 3 จะเห็นว่า ค่าเฉลี่ยของ CK 105 ซม. ต่ำกว่าความสูงเฉลี่ยจากการใช้ปุ๋ย N P ร่วมกับ K ซึ่งวัดได้ 131 ซม.

2. ในกลุ่มของการใช้ปุ๋ย N และ P ร่วมกับ K ที่อัตราคงที่ (20 กก./ไร่) พบว่า N มีอิทธิพลต่อความสูงของป่านรามี แต่ยังไม่หลักฐานชัดเจนว่าค่า P ทำให้ความสูงของป่านเปลี่ยนแปลงไป และเนื่องจาก N เป็นปัจจัยเชิงปริมาณ จึงทำการศึกษาการตอบสนองของป่านรามีต่อไนโตรเจน ให้ผลดังนี้

N	0	20	40
ผลรวม	947	1164	1426
สัมประสิทธิ์ Linear	-1	0	+1 (ตัวหาร = 2)
N Linear SS	= $(-947+0+1426)^2/2 \times 9$ =		12746.7
เมื่อใช้ค่า MS Error (83.89) ทดสอบได้ค่า F		=	151.95**

แสดงว่าการตอบสนองต่อปุ๋ยไนโตรเจนในช่วงระยะที่ศึกษายังเป็นเส้นตรงไม่ถึงจุดตอบสนองสูงสุด และเมื่อใช้ค่าเฉลี่ย N จากตารางที่ 3 ไปคำนวณสมการได้ดังนี้

$$\hat{Y} \quad = \quad 104.17 + 1.325^{**}N \qquad (r^2 = 0.997^*)$$

Split plot Design

Two – Factors Factorial Experimental ที่กล่าวมาแล้ว ถ้านำมาจัดรูปร่างหรือตำแหน่งเสียใหม่ให้เป็นหน่วยของการทดลอง (plot size) 2 ขนาด ก็จะกลายเป็น Design ใหม่ที่เรียกว่า Split plot design มีหน่วยทดลอง 2 ขนาดสำหรับปัจจัย 2 ชนิด ที่ทำการศึกษา หน่วยทดลองอันที่มีขนาดใหญ่เรียกว่า Main plot ในหนึ่งซ้ำจำนวน Main plot จะต้องเท่ากับจำนวนระดับของปัจจัยแรก หน่วยทดลองที่มีขนาดเล็กกว่าเรียกว่า Sub plot จำนวน Sub plot ในแต่ละ Main plot จะต้องเท่ากับจำนวนระดับของปัจจัยหลัง

การมีหน่วยทดลอง 2 ขนาด จะทำให้ความแม่นยำในการวัดการตอบสนองของแต่ละปัจจัยไม่เท่ากัน ปัจจัยที่มีขนาดของหน่วยทดลองเล็กจะถูกวัดอย่างละเอียดกว่าปัจจัยที่มีหน่วยทดลองขนาดใหญ่ ดังนั้นจึงมีความแม่นยำสูงกว่า

จะเลือกใช้ Design นี้เมื่อ

1. ในเชิงปฏิบัติหน่วยทดลองของปัจจัยหนึ่งต้องใหญ่กว่าอีกปัจจัยหนึ่ง เช่น การทดสอบจำนวนวันของการให้น้ำแก่ฝ้ายพันธุ์ต่างๆ ควรจะใช้การให้น้ำเป็น Main plot และสายพันธุ์ของฝ้ายเป็น Sub plot หรือ

2. เมื่อเรามีความสนใจที่จะทดสอบปัจจัยใดปัจจัยหนึ่งมากกว่าอีกปัจจัยหนึ่ง และเป็นไปได้ในทางปฏิบัติก็ให้ใช้ปัจจัยที่สนใจน้อยกว่าเป็น Main plot และปัจจัยที่ต้องการจะศึกษาอย่างละเอียดเป็น Sub plot

Randomization และ Layout

การ Random สำหรับ design นี้ ต้องทำ 2 ครั้ง ครั้งแรกทำการสุ่มสำหรับ Main plot ด้วยวิธีการแบบเดียวกับ Design พื้นฐานแบบใดแบบหนึ่งใน 3 แบบ คือ CRD, RCB และ LS หลังจากนั้นจึงทำการสุ่มครั้งที่สอง สำหรับการสุ่ม Sub plot นั้น ใช้วิธีการสุ่มเช่นเดียวกับ RCB คือต้องสุ่ม Sub plot ในแต่ละ Main plot อย่างเป็นอิสระ ซึ่งกันและกันเหมือนกับการสุ่ม treatment ในแต่ละ Block และทุก Main plot จะต้องมี Sub plot ให้ครบ

ตัวอย่าง ศึกษาการตอบสนองของข้าว 3 พันธุ์ ต่อปุ๋ยแอมโมเนียมซัลเฟตระยะยาว

วัตถุประสงค์ เพื่อศึกษาว่าข้าวพันธุ์ กข5, กข7, และ กข9 มีการตอบสนองต่อปุ๋ยแอมโมเนียมซัลเฟตอย่างไร

Main plot พันธุ์ข้าว กข5, กข7, กข9 วางแผนแบบ RCB 4 ซ้ำ

Sub plot ปุ๋ยแอมโมเนียมซัลเฟต 6 ระดับ 0, 3, 6, 9, 12 และ 18 กก./ไร่ (มี P และ K รองพื้นอย่างละ 6 กก./ไร่) และ Check คือ ไม่ใส่ปุ๋ยรวมเป็น 7 treatment

Randomization

ชั้นที่ 1 ทำการสุ่ม Main plot ทั้ง 3 ลงในแต่ละ Block อย่างเป็นอิสระกัน

Block 1	Block 2
RD7	RD9
RD9	RD5
RD5	RD7

Block 3	Block 4
RD5	RD9
RD7	RD7
RD9	RD5

ชั้นที่ 2 ในแต่ละ Main plot แบ่งเป็น 7 Sub plot เพื่อบรรจุ Sub plot treatment ด้วย วิธีสุ่มเป็นอิสระกันทุก Main plot

Treatment	ที่ 1	Check	
Treatment	ที่ 2	N-P-K	0-6-6
Treatment	ที่ 3	N-P-K	3-6-6
Treatment	ที่ 4	N-P-K	6-6-6
Treatment	ที่ 5	N-P-K	9-6-6
Treatment	ที่ 6	N-P-K	12-6-6
Treatment	ที่ 7	N-P-K	18-6-6

6	5	2	3	1	4	7	RD7
2	1	5	6	7	3	4	RD9
3	6	7	4	5	2	1	RD5

Block 1

3	1	5	2	6	7	4	RD9
5	7	6	4	1	2	3	RD5
3	6	7	4	5	2	1	RD7

Block 2

4	3	1	2	5	7	6	RD5
7	6	2	5	3	4	1	RD7
6	7	3	4	2	1	5	RD9

Block 3

3	6	5	1	4	2	7	RD9
5	3	6	2	7	1	4	RD7
2	7	4	3	1	5	6	RD5

Block 4

<u>ตัวอย่าง 1</u>

เรื่อง การศึกษาการตอบสนองของข้าวต่อการใช้ซิลิกอนไดออกไซด์และซิงค์ออกไซด์ (SiO_2, ZnO)

วัตถุประสงค์ เพื่อศึกษาว่าพันธุ์ข้าวที่ทดลอง 6 พันธุ์ มีการตอบสนองต่อซิลิกอนและสังกะสีอย่างไร

สถานที่ทำการทดลอง สถานีทดลองข้าวสุรินทร์

ฤดูปลูก กค. – ตค. 2524

ปัจจัยที่ทดลอง main plot 4 ระดับ (ทุกระดับใส่ NPK 6-6-6 กก./ไร่)

<u>ระดับ</u>	<u>SiO_2</u>	<u>ZnO</u>	
M_1	0	0	กก./ไร่
M_2	5	0	กก./ไร่
M_3	0	3	กก./ไร่
M_4	50	3	กก./ไร่

Sub plot 6 พันธุ์

S_1	NSPT	เหนียวสันป่าตอง
S_2	KDML 105	ขาวดอกมะลิ 105
S_3	RD1	กข 1
S_4	RD7	กข 7
S_5	BKN-6721-5-39-1	(สายพันธุ์)
S_6	RD8	กข 8

ขนาดแปลงย่อย ปลูก 3 x 5 เมตร ระยะปลูก 25 x 25 ซม.

ขนาดแปลงเก็บเกี่ยว 2 x 4 เมตร

วางแผนการทดลองแบบสพลิทพล็อต main plot จัดในรูป RCB 4 ซ้ำ

แผนผังแปลงทดลอง

M_1	M_3	M_2	M_4
S_3	S_4	S_6	S_1
S_5	S_2	S_4	S_3
S_2	S_3	S_5	S_2
S_4	S_1	S_2	S_5
S_1	S_6	S_3	S_4
S_6	S_5	S_1	S_6

Block 1

M_2	M_4	M_3	M_1
S_1	S_4	S_5	S_6
S_2	S_3	S_4	S_5
S_5	S_1	S_3	S_4
S_3	S_6	S_1	S_2
S_6	S_2	S_6	S_3
S_4	S_5	S_2	S_1

Block 2

M_3	M_2	M_1	M_4
S_2	S_1	S_3	S_5
S_1	S_6	S_4	S_4
S_5	S_2	S_6	S_3
S_3	S_4	S_1	S_2
S_6	S_3	S_5	S_1
S_4	S_5	S_2	S_6

M_3	M_1	M_4	M_2
S_3	S_6	S_5	S_1
S_6	S_2	S_3	S_6
S_2	S_5	S_6	S_4
S_5	S_1	S_4	S_2
S_4	S_3	S_1	S_5
S_1	S_4	S_2	S_3

Block 3 Block 4

1.1 เตรียมข้อมูล (ผลผลิต กก./ไร่)

ตารางที่ 1

Main plot	Sub plot	Block				ผลรวม
A	B	1	2	3	4	
1	1	430	371	349	359	1509
	2	412	439	444	442	1737
	3	472	500	475	463	1910
	4	318	423	426	380	1547
	5	263	375	420	395	1453
	6	408	498	455	434	1795
2	1	439	407	361	383	1590
	2	415	450	504	442	1811
	3	462	503	485	451	1901
	4	452	312	437	331	1532
	5	414	487	537	392	1830
	6	482	440	494	452	1868
3	1	374	380	378	371	1503
	2	384	464	480	469	1797
	3	495	454	475	491	1915
	4	453	405	409	449	1716
	5	421	443	474	461	1799
	6	409	432	447	431	1719
4	1	412	477	402	433	1724
	2	491	490	473	472	1926
	3	531	500	489	532	2052
	4	352	445	501	455	1753
	5	479	509	472	483	1943
	6	492	519	501	483	1995
ผลรวม		10260	10723	10888	10454	42325

ตารางที่ 2 (A x B)

Main plot (A)	Sub plot (B) 1	2	3	4	5	6	ผลรวม
1	1509	1737	1910	1547	1453	1795	9951
2	1590	1811	1901	1532	1830	1868	10532
3	1503	1797	1915	1716	1799	1719	10449
4	1724	1926	2052	1753	1943	1995	11393
ผลรวม	6326	7271	7778	6548	7025	7377	42325

ตารางที่ 3

Main plot (A)	Block 1	2	3	4	ผลรวม
1	2303	2606	2569	2473	9951
2	2664	2599	2818	2451	10532
3	2536	2578	2663	2672	10449
4	2757	2940	2838	2858	11393
ผลรวม	10260	10723	10888	10454	42325

ข้อสังเกต ถ้าใช้คอมพิวเตอร์ในการวิเคราะห์ เตรียมเฉพาะข้อมูลดิบของตารางที่ 1 (โดยไม่ต้องมีผลรวมทั้งแนวตั้ง และแนวนอน) เพียงตารางเดียว

1.2 คำนวณ Degree of freedom (df)

Total df = 4 x 6 x 4 - 1

= 95

Block df = 4 - 1

= 3

A df = 4 - 1

= 3

Error a df = (4 – 1) (4 – 1)

= 9

B df = 6 – 1

= 5

A x B df = (4-1) (6-1)

= 15

Error df = 4 (6 – 1) (4 – 1)

= 60

1.3 คำนวณ Sum of Square (SS) ต่างๆ

จากตารางที่ 1

Correction Factor (C.F.)= $\frac{42325^2}{4x6x4}$

= 18,660,475.26

Total SS = $(430^2+371^2+\ldots+483^2)$ – C.F.

= 18,923,037.00 – 18,660,475.26

= 262,5621.74

Block SS = $\frac{10260^2 + \ldots + 10454^2}{4x6}$ - C.F.

= 18,970,207.88 – 18,660,475.26

= 9,732.62

จากตารางที่ 2

A SS = $\frac{9,951^2 + \ldots + 11,393^2}{6x4}$ - C.F.

= 18,705,311.46 – 18,660,475.26

= 44,836.20

B SS = $\frac{6,326^2 + \ldots + 7,377^2}{4x4}$ - C.F.

= 18,751,878.69 – 18,660,475.26

= 91,403.43

A x B SS = $\frac{1,509^2 + 1,737^2 + \ldots + 1,995^2}{4}$ - C.F.-A.SS-B.SS

= 18,821,965.75 – 18,660,475.26 – 44,836.20

– 91,403.43

= 25,250.86

จากตารางที่ 3

Error (a) SS = $\frac{2,303^2 + ... + 2,858^2}{6x4}$ - C.F.-A.SS-Block SS

= 18,731,057.84-18,660475.26-44,836.20-9,732.62

= 16,013.76

Error (b) SS = Total SS-(Block SS)-(A.SS)-(B.SS)-(AxB SS)

- (Error a SS)

= 262,561.74-9,732.62-44,836.20-91,403.43

-25,250.86-16,013.76

= 75,324.88

คำนวณค่า Mean Square (MS)

Block MS = $\frac{9,732.62}{3}$

= 3,244.20

A MS = $\frac{44,836.20}{3}$

= 14,945.40

B MS = $\frac{91,403.43}{5}$

= 18,280.69

A x B MS = $\frac{25,250.86}{15}$

= 1,683.39

Error (a) MS = $\frac{16,013.76}{9}$

= 1,779.31

Error (b) MS = $\frac{75,324.88}{60}$

= 1,255.42

1.5 คำนวณค่า F

$$F_{Block} = \frac{3,244.20}{1,779.31} = 1.82$$

$$F_A = \frac{14,945.40}{1,779.31} = 8.40$$

$$F_B = \frac{18,280.69}{1,255.42} = 14.56$$

$$F_{AxB} = \frac{1,683.39}{1,255.42} = 1.34$$

1.6 เปิดค่า F จากตาราง

F_{Block} และ F_A ดูที่ f_2 = 9 และ f_1 = 3

$F_{.05}$ = 3.86

$F_{.01}$ = 6.99

F_B และ F_{AB} ดูที่ f_2 = 60 f_1 = 5 และ 15 ตามลำดับ

B : $F_{.05}$ = 2.37

$F_{.01}$ = 3.34

AB : $F_{.05}$ = ระหว่าง 1.86 และ 1.81

$F_{.01}$ = ระหว่าง 2.40 และ 2.32

1.7 เปรียบเทียบค่า F จากตารางและ F จากการคำนวณ

F_{Block} คำนวณได้น้อยกว่า F จากตารางใส่ ns

F_A คำนวณได้มากกว่า $F_{.01}$ จากตารางใส่ **

F_B คำนวณได้มากกว่า $F_{.01}$ จากตารางใส่ **

F_{AB} คำนวณได้น้อยกว่า F จากตารางใส่ ns

1.8 คำนวณค่า C.V.

$$\text{mean} = \frac{42,325}{96}$$

$$= 440.89$$

$$\text{C.V. (a)} = \frac{\sqrt{1,779.31}}{440.89} \text{x}100$$

$$= 9.6\%$$

$$\text{C.V. (b)} = \frac{\sqrt{1,225.42}}{440.89} \text{x}100$$

$$= 8.0\%$$

1.9 ทำตารางวิเคราะห์แวเหรียนซ์ โดยกรอกตัวเลขที่คำนวณได้ทั้งหมดลงในตาราง

ตารางวิเคราะห์แวเหรียนซ์ของผลผลิต (กก./ไร่)

Source of Variation	Degree of Freedom	Sum of Square	Mean Square	F
Total	95	262,562		
Block	3	9,733	3,244.20	1.83^{ns}
A	3	44,836	14,945.41	8.40**
Error (a)	9	16,014	1,779.31	
B	5	91,403	18,280.69	14.56**
A x B	15	25,251	1,683.39	1.34^{ns}
Error (b)	60	75,325	1,255.42	

C.V. (a) = 9.6%

C.V. (b) = 8.0%

1.10 ทำตารางค่าเฉลี่ยจากตารางที่ 2 ในข้อ 1

$$A_1 = \frac{9,951}{6\text{x}4} = 414.6\ldots\ldots\ldots$$

$$B_1 = \frac{6,326}{4\text{x}4} = 395.4\ldots\ldots\ldots$$

$$A_1B_1 = \frac{1,509}{4} = 377.2\ldots\ldots\ldots$$

ตารางค่าเฉลี่ยของผลผลิต (เฉลี่ยจาก 4 ซ้ำ)

Sub plot		Main plot				เฉลี่ย
		1	2	3	4	(พันธุ์)
1	NSPT	377	398	376	431	395
2	KDML 105	434	453	449	482	454
3	RD 1	478	475	479	513	486
4	RD 7	387	383	429	438	409
5	BKN – 6721	363	458	450	486	439
6	RD 8	449	467	430	499	461
เฉลี่ย (A)		415	439	435	475	

1.11 คำนวณตัวสถิติที่จะใช้เปรียบเทียบ

จากตารางวิเคราะห์แวเหรียนซ์ จะเห็นว่า ส่วนที่มีนัยสำคัญทางสถิติ คือ A และ B เท่านั้น ส่วน Interaction ไม่มีนัยสำคัญ แสดงว่าระดับสารเคมีมีอิทธิพลต่อผลผลิตและอิทธิพลนี้จะคล้ายคลึงกันในแต่ละพันธุ์ นั่นคือ ศึกษาเพียงอิทธิพลของสารเคมีและพันธุ์ก็เพียงพอแล้ว

และเนื่องจากพันธุ์ที่ใช้ในการทดสอบมีถึง 6 พันธุ์ จึงควรใช้ Duncan Multiple Range Test (DMRT) เป็นตัวตัดสิน ซึ่งควรใช้สถิติตัวเดียวกันนี้ในการเปรียบเทียบระดับสารเคมีด้วย

$$LSR_P = SSR_P \times S_{\bar{x}}$$

ก. เปรียบเทียบค่าเฉลี่ยของ (main plot) ระดับสารเคมี

$$S_{\bar{x}} = \sqrt{\frac{1,779.31}{4x6}}$$

$$= 8.61$$

สารเคมีทั้งหมด 4 ระดับ

P	2	3	4 (df=9)
SSR_P	3.20	3.34	3.41
LSR_P	27.55	28.76	29.36

เมื่อเปรียบเทียบค่าเฉลี่ยของระดับสารเคมีตามวิธีเปรียบเทียบของ DMRT ได้ผลดังนี้

ระดับ	SiO_2	ZnO	ค่าเฉลี่ย
1	0	0	415 b
2	5	0	439 b
3	0	3	435 b
4	50	3	475 a

ข. เปรียบเทียบค่าเฉลี่ยของ Sub plot สายพันธุ์

$$S_{\bar{x}} = \sqrt{\frac{1,255.42}{4x4}}$$

$$= 8.86$$

พันธุ์ทั้งหมดมี 6 พันธุ์

P	2	3	4	5	6 (df =60)
SSR_P	2.83	2.98	3.08	3.14	3.20
LSR_P	25.1	26.4	27.3	27.8	28.3

เมื่อเปรียบเทียบค่าเฉลี่ยของพันธุ์

พันธุ์	ค่าเฉลี่ย
1 NSPT	395 c
2 KDML 105	454 b
3 RD1	486 a
4 RD7	409 c
5 BKN 6721	439 b
6 RD8	461 b

1.12 การทำตารางเสนอผลการทดลอง

การศึกษาการตอบสนองของข้าวต่อการใช้ซิลิกอนไดออกไซด์และซิงค์ออกไซด์

สุรินทร์ กค. – ตค. 2523

ผลผลิต[1/] กก./ไร่ (เฉลี่ยจาก 4 ซ้ำ)

สายพันธุ์		ระดับสารเคมี (SiO_2 – ZnO)				เฉลี่ยพันธุ์
		1	2	3	4	
1	NSPT	377	398	376	431	395 c
2	KDML 105	434	453	449	482	454 b
3	RD1	478	475	479	513	486 a
4	RD7	387	383	429	438	409 c
5	BKN 6721	363	458	450	486	439 b
6	RD8	449	467	430	499	461 b
เฉลี่ย (สารเคมี)		415 b	439 b	435 b	475 a	

1/ ตัวเลขที่ตามหลังด้วยตัวอักษรเหมือนกันไม่ต่างกันทางสถิติ โดยใช้ DMRT ที่ระดับความเชื่อมั่น 95%

การแปลผล

ระดับสารเคมีมีอิทธิพลต่อผลผลิต กล่าวคือ การใช้ SiO_2 ร่วมกับ ZnO ระดับ 50 และ 3 กก./ไร่ ตามลำดับ จะให้ผลผลิตสูงที่สุดถึง 475 กก./ไร่ ส่วนการไม่ใช้สารทั้งสองหรือใช้เพียงอย่างใดอย่างเดียวจะได้ผลผลิตที่ไม่แตกต่างกัน 415, 439 และ 435 กก./ไร่ ตามลำดับ

ในด้านของพันธุ์ที่ใช้ทดสอบทั้ง 6 สายพันธุ์ แบ่งเป็นกลุ่มได้ดังนี้ RD1 ให้ผลผลิตสูงที่สุด 486 กก./ไร่ รองลงมาคือ KDML 105 และ RD8 ได้ 454 และ 461 กก./ไร่ ตามลำดับ สองพันธุ์นี้ให้ผลผลิตไม่แตกต่างกับ BKN 6721 (439 กก./ไร่) ส่วน RD7 และ NSPT ให้ผลผลิตต่ำที่สุด (409 และ 395 กก./ไร่)

ตัวอย่างที่ 2 Top dressing of Nitrogen and compound fertilizers for rice. Wet season 1973. สถานีทดลองข้าวควนกุฎ พัทลุง

Main plot – Variety : RD1 and RD5 (จัดในรูป RCB)

Sub plot – ปุ๋ย 10 ตำหรับ ดังนี้

Basal + Topdress		Active Ingredient	
		N	P_2O_5
1	Check	-	-
2	AP 10 + AS 4	(1.6+0.8)	2.0
3	AP 20 + AS 8	(3.2+1.6)	4.0
4	AP 30 + AS 12	(4.8+2.4)	6.0
5	AP 10 + U 1.8	(1.6+0.8)	2.0
6	AP 20 + U 3.6	(3.2+1.6)	4.0
7	AP 30 + U 5.4	(4.8+2.4)	6.0
8	AP 10 + AP 5	(1.6+0.8)	3.0
9	AP 20 + AP 10	(3.2+1.6)	6.0
10	AP 30 + AP 15	(4.8+2.4)	9.0

Replication : 4

ตารางที่ 1 Grain Yield (Kg/rai)

Mainplot Subplot	Rep.				Total
	1	2	3	4	
RD1					
1. ไม่ใส่ปุ๋ย	251	201	235	217	904
2. AP 10 + AS 4	331	283	326	289	1229
3. AP 20 + AS 8	396	360	361	368	1485
4. AP 30 + AS 12	416	400	468	385	1669
5. AP 10 + U 1.8	302	295	338	256	1191
6. AP 20 + U 3.6	440	347	375	334	1496
7. AP 30 + U 5.4	496	359	428	389	1672
8. AP 10 + AP 5	335	313	285	233	1166
9. AP 20 + AP 10	452	406	370	356	1584
10. AP 30 + AP 15	417	406	503	376	1702
Sub Total	3836	3370	3689	3203	14098
RD5					
1. ไม่ใส่ปุ๋ย	243	235	291	232	1001
2. AP 10 + AS 4	233	236	277	254	999
3. AP 20 + AS 8	350	268	271	261	1150
4. AP 30 + AS 12	309	294	319	338	1260
5. AP 10 + U 1.8	308	314	265	282	1169
6. AP 20 + U 3.6	298	260	296	309	1163
7. AP 30 + U 5.4	294	295	330	308	1227
8. AP 10 + AP 5	366	309	265	264	1204
9. AP 20 + AP 10	316	235	273	267	1091
10. AP 30 + AP 15	372	320	297	353	1342
Sub total	3089	2766	2884	2867	11606
Total	6925	6136	6573	6070	25704

วิธีการคำนวณ

Correction factor (C)	=	$(25,704)^2/80$		
Total SS	=	$(251^2+\ldots+353^2)$-C	=	345,232.80
Rep.SS	=	$(6,925^2+\ldots+6,070^2)$/10x2-C	=	24,072.30
Var.SS	=	$(14,098^2+11,606^2)$/4x10-C	=	77,625.80
Error (a)	=	$(3,836^2+\ldots+2,867^2)$/10-C-Rep.SS-Var.SS	=	6,561.50
Fert.SS	=	$(904+1,001)^2+\ldots+(1,702+1,342)^2$/4x2-C	=	141,032.05
Var. x Fert.SS	=	$(904^2+\ldots+1,342^2)$/4-C-Fert.SS-Var.SS	=	50,537.45
Error (b)	=	Total SS-Rep.SS-Var.SS-E(a)-Fert.SS-Var.x Fert.SS		
	=	45,403.70		

แล้วจึงนำค่าที่คำนวณได้ไปสร้างตารางวิเคราะห์แวเหรียนซ์ต่อไป

ตารางที่ 2 ตารางวิเคาะห์แวเหรียนซ์ของผลผลิต (กก./ไร่)

SOV	df	SS	MS	F
Total	79	345,232.80		
Rep.	3	24,072.30	8,024.10	
Var.	1	77,625.80	77,625.80	35.49**
Error (a)	3	6,561.50	2,187.17	
Fertilizer	9	141,032.05	15,670.33	18.63**
Var. x Fert	9	50,537.45	5,615.27	6.67**
Error (b)	54	45,403.70	840.81	

C.V. (a) = 14.5%

C.V. (b) = 9.0%

ตารางที่ 3 ค่าเฉลี่ย กก./ไร่

Fert.	Var.		Fert.mean
	RD1	RD5	
1 (CK)	226	250	238
2	307	250	278
3	371	287	329
4	417	315	366
5	298	292	295
6	374	291	332
7	418	307	362
8	291	301	296
9	396	273	334
10	425	335	380
Var.mean	352	290	

LSD	(สำหรับเปรียบเทียบค่าเฉลี่ยของระดับปุ๋ย 2 ระดับที่ใช้พันธุ์เดียวกัน)	.05	=	41 กก./ไร่
		.01	=	55 กก./ไร่
LSD	(สำหรับเปรียบเทียบค่าเฉลี่ยของ 2 พันธุ์ ที่อยู่ในปุ๋ยระดับเดียวกันหรือต่างระดับ)	.05	=	50 กก./ไร่
		.01	=	75 กก./ไร่

เนื่องจากจำนวนของ Sub plot มีหลายระดับ การใช้เทคนิคของ Partitioning SS มาช่วยในการศึกษา main effect ของปุ๋ยและ Interaction ระหว่างพันธุ์กับปุ๋ยจะช่วยให้เข้าใจดียิ่งขึ้น โดยแยก treatment ออกเป็น 4 กลุ่ม เพื่อทำการเปรียบเทียบระหว่างกลุ่ม (Group Comparison) ดังนี้

กลุ่มของ CK คือ Treatment ที่ 1

กลุ่มของพวกใส่ปุ๋ย AP รองพื้น 10 กก. ได้แก่ Treatment 2, 5, 8

กลุ่มของพวกใส่ปุ๋ย AP รองพื้น 20 กก. ได้แก่ Treatment 3, 6, 9

กลุ่มของพวกใส่ปุ๋ย AP รองพื้น 30 กก. ได้แก่ Treatment 4, 7, 10

สัมประสิทธิ์สำหรับใช้ในการเปรียบเทียบแสดงไว้ในตารางต่อไปนี้

Group Comparison (ของปุ๋ย)

Comparison	Coefficient of Comparison										Divisor
	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	
1. Check VS Treated	-9	+1	+1	+1	+1	+1	+1	+1	+1	+1	90
2. Among AP_{10}											
(T_2 & T_5) VS T_8	0	+1	0	0	+1	0	0	-2	0	0	6
T_2 VS T_5	0	+1	0	0	-1	0	0	0	0	0	2
3. Among AP_{20}											
(T_3 & T_6) VS T_9	0	0	+1	0	0	+1	0	0	-2	0	6
T_3 VS T_6	0	0	+1	0	0	-1	0	0	0	0	2
4. Among AP_{30}											
(T_4 & T_7) VS T_{10}	0	0	0	+1	0	0	+1	0	0	-2	6
T_4 VS T_7	0	0	0	+1	0	0	-1	0	0	0	2

สำหรับการคำนวณทั้ง main effect (ปุ๋ย) และ Interaction (พันธุ์ x ปุ๋ย) ใช้สัมประสิทธิ์ชุดเดียวกัน เพียงแต่มีตัวหารต่างกัน ใช้ค่าผลรวมของ Treatment จากตารางที่ 1 มาคำนวณดังนี้

Main plot	Sub plot										รวม
	1	2	3	4	5	6	7	8	9	10	
RD1	904	1229	1485	1669	1191	1496	1672	1166	1584	1702	14098
RD5	1001	999	1150	1260	1169	1163	1227	1204	1091	1342	11606
รวม	1905	2228	2635	2929	2360	2659	2899	2370	2675	3044	25704

วิธีคำนวณ

Main effect

$$\text{CK VS Treated SS} = \frac{\left((-9)(1{,}905) + (1)(2{,}228) + \ldots + (1)(3{,}044)\right)^2}{90 \times 4 \times 2}$$

$$= 61{,}494.05$$

$$(T_2 \,\&\, T_5) \text{ VS } T_8 \text{ SS} = \frac{\left((1)(2{,}228) + (1)(2{,}360) + (-2)(2{,}370)\right)^2}{6 \times 4 \times 2}$$

$$= 481.33$$

ฯลฯ

Interaction

$$\text{Var. x (CK VS Treated) SS} = \frac{\left((-9)(904) + (1)(1{,}229) + \ldots + (1)(1{,}702)\right)^2}{90 \times 4} + \frac{\left((-9)(1{,}001) + (1)(999) + \ldots + (1)(1{,}342)\right)^2}{90 \times 4} - \text{(CK VS treated) SS}$$

$$= 16{,}646.45$$

ฯลฯ

ตารางที่ 4 Fertilizer (main effect ของ Sub plot จากตารางที่ 2)

Source	df	SS	MS	F
Fert	9	141,032.05	15,670.23	18.63**
CK VS treated	(1)	61,494.05	61,494.05	73.14**
Among treated	(8)	79,538.00	9,942.25	11.82**
Among ($AP_{10,20,30}$)	2	76,401.75	38,200.88	45.43**
Among AP_{10}				
T_2&T_5 VS T_8	1	481.33	481.33	<1
T_2 VS T_5	1	1,089.00	1,089.00	1.29^{ns}
Among AP_{20}				
T_3&T_6 VS T_9	1	65.33	65.33	<1
T_3 VS T_6	1	36.00	36.00	<1
Among AP_{30}				
T_4&T_7 VS T_{10}	1	1,408.33	1,408.33	1.67^{ns}
T_4 VS T_7	1	56.25	56.25	<1

ตารางที่ 5 Variety x Fertilizer (มาจากตารางที่ 2)

Source	df	SS	MS	F
Var. x Fert.	9	50,537.45	5,615.27	6.67**
Var x (CK VS treated)	(1)	16,646.45	16,646.45	19.80**
Var x (Among treated)	(8)	33,891.00	4,236.38	5.04**
Var x (Among $AP_{10,20,30}$)	2	26,383.58	13,191.79	15.69**
Var x (Among AP_{10})	2			
Var x (T_2 & T_5 VS T_8)	(1)	2,241.33	2,241.33	2.67^{ns}
Var x (T_2 VS T_5)	(1)	2,704.00	2,704.00	3.21^{ns}
Var x (Among AP_{20})	2			
Var x (T_3 & T_6 VS T_9)	(1)	2,106.75	2,106.75	2.50^{ns}
Var x (T_3 VS T_6)	(1)	0.25	0.25	<1
Var x (Among AP_{30})	2			
Var x (T_4 & T_7 VS T_{10})	(1)	374.08	374.08	<1
Var x (T_4 VS T_7)	(1)	81.00	81.00	<1
Error (b)	54	45,403.70	840.81	

ตารางที่ 6 ค่าเฉลี่ย กก./ไร่ (จากตารางที่ 3)

Var.	Fertilizer			Fertilizer mean	CK-mean
	AP_{10}	AP_{20}	AP_{30}		
RD1	299	378	420	366	226
RD5	281	284	319	295	250
Average	290	331	370	330	238

LSD (สำหรับเปรียบเทียบค่าเฉลี่ยจากปุ๋ย 2 ระดับในพันธุ์เดียวกัน) .05 = 24 kg/rai

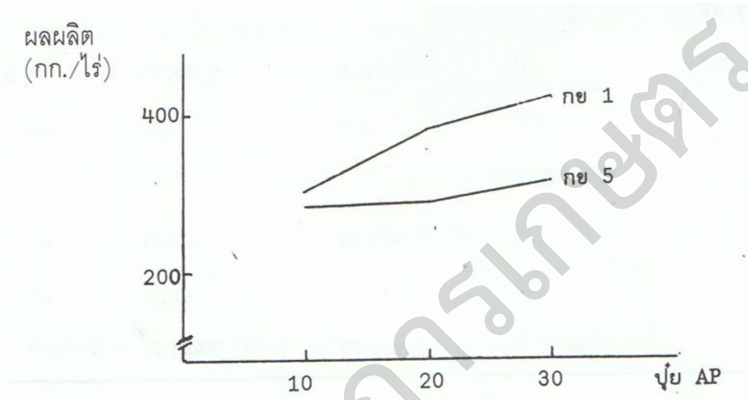
.01 = 32 kg/rai

การแปลผล โดยพิจารณาตารางที่ 4, 5 และ 6 พร้อมๆ กัน

เพราะว่า Var x Fert. significant จึงต้องศึกษาการตอบสนองต่อปุ๋ยของแต่ละพันธุ์

1. CK VS. Treated** (ตารางที่ 4) หมายความว่า การใส่ปุ๋ยทำให้ผลผลิตเปลี่ยนแปลง และ Var.x(CK VS. Trested)** (ตารางที่ 5) แสดงว่าการเปลี่ยนแปลงนั้น ต่างกันไปในแต่ละพันธุ์ ดูที่ค่าเฉลี่ย (ตารางที่ 6) จะพบว่า กข1 ตอบสนองต่อปุ๋ยทุกตำหรับโดยเฉลี่ยถึง 140 กก. (366 VS 226) และ กข5 ตอบสนองเพียง 45 กก. (295 VS 250)

2. Among AP 10, 20, 30** (ตารางที่ 4) แสดงว่าการรองพื้นด้วยปุ๋ย AP อัตราต่างๆ กัน ทำให้ผลผลิตของข้าวแตกต่างกัน ยิ่งกว่านั้น Var x Among AP 10, 20, 30** (ตารางที่ 5) แสดงให้เห็นว่า ความแตกต่างของการตอบสนองต่อปุ๋ยดังกล่าวไม่เหมือนกันในสองพันธุ์ที่ทำการทดลอง จากตารางเฉลี่ย (ตารางที่ 6) จะเห็นว่า กข1 มีการตอบสนองสูงกว่า กข5 ดังรูป



กข1 มีการตอบสนองต่อปุ๋ยที่เพิ่มขึ้นอย่างมีนัยสำคัญยิ่ง (378-299=79** และ 420-378=42**) ส่วน กข5 นั้นไม่มีการตอบสนองที่มีนัยสำคัญ เมื่อเพิ่มปุ๋ยระดับกลาง (284-281=3 และ 319-284=35**)

3. ในแต่ละหมู่ของปุ๋ยรองพื้นคือหมู่ 10, 20, 30 พบว่าไม่ว่าจะ Top dress ด้วย N ล้วนๆ หรือ N+P ไม่ทำให้ผลผลิตของทั้งสองสายพันธุ์เปลี่ยนแปลงเลย เช่น

Among AP_{10} (จากตาราง4 และ ตาราง5)

T_2 & T_5 VS T_8^{ns} และ Var x (T_2 & T_5 VS T_8)^{ns} นั่นคือการ Top dress ด้วย N ล้วนๆ ไม่แตกต่างกับ Top dress ด้วย N+P ทั้งสองสายพันธุ์

T_2 VS T_5^{ns} และ Var x (T_2 VS T_5)^{ns} นั่นคือการ Top dress ด้วย N ล้วนๆ ไม่ว่าจะเป็น N จากปุ๋ยชนิดใดก็ไม่ทำให้เกิดการเปลี่ยนแปลงของผลผลิตทั้งสองพันธุ์

สรุปผล

ข้าวพันธุ์ กข1 ตอบสนองต่อปุ๋ยได้ดีกว่า กข5 โดยเฉลี่ยประมาณ 3 เท่า (140 VS 45) และไม่ว่าจะใส่ปุ๋ยรองพื้นเพิ่มจาก 10 เป็น 20 หรือเป็น 30 กก. กข1 ก็เพิ่มผลผลิตอย่างมีนัยสำคัญ คือ เพิ่มระหว่าง 79 – 121 กก./ไร่ ส่วน กข5 ผลผลิตจะเพิ่มเมื่อใช้ปุ๋ยรองพื้นระดับสูงๆ เท่านั้นและ

เพิ่มเพียง 38 กก./ไร่ การใช้ปุ๋ยแต่งหน้า ไม่ว่าจะเป็นปุ๋ยประเภทไนโตรเจนล้วนๆ หรือมีฟอสเฟต ด้วย ก็ไม่ทำให้ผลผลิตของข้าวทั้ง กข1 และ กข5 มีการเปลี่ยนแปลงแต่อย่างใด

<u>ตัวอย่าง 3</u> ผลการตอบสนองของข้าวต่อการใส่ปุ๋ยแอมโมเนียซัลเฟตระยะยาว สถานีทดลองข้าวชัยนาท นาปี 2519 (ดูแผนผังจากตัวอย่างการสุ่ม)

ตารางที่ 1 ผลผลิต (กก./ไร่)

Var.	Tr.No.	Fertilizer	Replication				Total
			1	2	3	4	
RD5	1	CK	731	621	692	620	2664
	2	0-6-6	686	663	711	693	2753
	3	3-6-6	636	695	725	724	2780
	4	6-6-6	744	720	813	708	2985
	5	9-6-6	706	777	822	711	3016
	6	12-6-6	657	709	739	751	2856
	7	18-6-6	536	529	603	724	2392
RD7	1	CK	692	792	678	593	2755
	2	0-6-6	785	705	602	622	2714
	3	3-6-6	794	838	737	729	3098
	4	6-6-6	856	868	819	754	3297
	5	9-6-6	838	836	825	882	3381
	6	12-6-6	899	904	916	833	3552
	7	18-6-6	1021	944	952	913	3830
RD9	1	CK	567	599	600	573	2339
	2	0-6-6	689	610	650	694	2643
	3	3-6-6	735	719	785	656	2895
	4	6-6-6	784	763	788	737	3072
	5	9-6-6	838	685	893	797	3213
	6	12-6-6	868	798	807	835	3308
	7	18-6-6	880	819	832	832	3363
Average	Total		15942	15594	15989	15381	62906

ตารางที่ 2 Two-way table (Variety x Fertilizer)

Fertilizer	Variety			Total
	RD5	RD7	RD9	
CK	2664	2755	2339	7758
0-6-6	2753	2714	2643	8110
3-6-6	2780	3098	2895	8773
6-6-6	2985	3297	3072	9354
9-6-6	3016	3381	3213	9610
12-6-6	2856	3552	3308	9716
18-6-6	2392	3830	3363	9585
Total	19446	22627	20833	62906

ตารางที่ 3 สำหรับคำนวณ Error (a)

Variety	Block				Total
	1	2	3	4	
RD5	4696	4714	5105	4931	19446
RD7	5885	5887	5529	5326	22627
RD9	5361	4993	5355	5124	20833
Total	15942	15594	15989	15381	62906

ใช้ค่าในตารางที่ 1, 2 และ 3 คำนวณค่าต่างๆ ดังนี้

วิธีคำนวณ

Correction factor C.F. = $(62906)^2/84 = 47{,}109{,}105.1904$

Total SS = $(731^2+\ldots+832^2) - \text{C.F.} = 885257$

Replication SS = $\frac{1}{21}(15{,}942^2+\ldots+15{,}381^2) - \text{C.F.} = 12{,}013.00$

Variety SS = $\frac{1}{28}(19{,}466^2+\ldots+20{,}833^2) - \text{C.F.} = 181{,}678.17$

Error (a) = $\frac{1}{7}(4{,}696^2+\ldots+5{,}124^2)$-C.F.-Rep SS-Variety SS

= 51,186.49

Fertilizer SS = $\frac{1}{12}(7{,}758^2+\ldots+9{,}585^2)$-C.F. = 311,435.64

Var. x Fert. SS = $\frac{1}{4}(2{,}664^2+\ldots+3{,}363^2)$-C.F.-Var. SS-Fert SS

= 217,822.51

Error (b) = total SS-Rep.SS-Var. SS-E(a) - Fert SS-(Var. x Fert.) SS

= 111,121.00

แล้วคำนวณค่า degree of freedom และ mean square ของแต่ละส่วนแลค่า C.V. ทำเป็นตารางดังนี้

ตารางที่ 4 ตารางวิเคราะห์แวเหรียนซ์ผลผลิต (กก./ไร่)

SOV	df	SS	MS	F
Total	83	885257		
Replication	3	12013	4004.33	
Variety	2	181678	90839.00	10.64**
Error (a)	6	51186	8531.00	
Fertilizer	6	311436	51906.00	25.22**
Variety x Fert.	12	217823	18151.92	8.82**
Error (b)	54	111121	2057.80	

C.V. (a) = 12%

C.V. (b) = 6%

ตารางที่ 5 ค่าเฉลี่ยผลผลิต (กก./ไร่)

Fertilizer	Variety			Fert-mean
	RD5	RD7	RD9	
CK	666	689	585	646
0-6-6	688	678	661	676
3-6-6	695	774	724	731
6-6-6	746	824	768	780
9-6-6	754	845	803	801
12-6-6	714	888	827	810
18-6-6	598	958	841	799
Var.mean	694	808	744	

For Comparison of	LSD.05	LSD.01	
2 Fert. means of the same var.	64	86	Kg/rai
2 Var. mean at same or different level of fert.	84	120	Kg/rai

จากตารางที่ 4 จะเห็นว่าพันธุ์และปุ๋ยมี Interaction ต่อกัน นั่นคือพันธุ์ข้าวที่ทดลองทั้งสามพันธุ์มีการตอบสนองต่อปุ๋ยแตกต่างกัน แต่เนื่องจากปุ๋ยที่ใช้ทดลองมีถึง 7 ตำหรับ การแปลผลการทดลองจึงควรจะใช้เทคนิคทางสถิติช่วยเพื่อให้การแปลผลสะดวกและถูกต้องยิ่งขึ้น แทนที่จะใช้ค่า LSD อย่างเดียว โดยใช้ทั้ง Group และ Trend comparison สำหรับตัวอย่างนี้นักวิชาการเจ้าของเรื่องทราบว่าในกลุ่มของสายพันธุ์ที่ทำการทดลอง ข้าวพันธุ์ กข7 และ กข9 มีอายุใกล้เคียงกันมาก (128 และ 120 วัน ตามลำดับ) ส่วนพันธุ์ กข5 เป็นข้าวอายุค่อนข้างหนัก (150 วัน) จึงจะทำการแยก SS ของ main plot ออกเป็นกลุ่มของ กข7 และ 9 กลุ่มหนึ่งกับ กข5 อีกกลุ่มหนึ่ง ส่วน Sub plot นั้น เป็นการเปรียบเทียบปุ๋ยไนโตรเจนระดับต่างๆ ซึ่งเป็นลักษณะเชิงปริมาณ (Quantitative) สามารถแยกการเปรียบเทียบได้ดังนี้

1. เปรียบเทียบระหว่างการไม่ใส่ปุ๋ยและการใส่ปุ๋ย
2. ในพวกที่ใส่ปุ๋ยเทียบว่ามีแนวโน้มของการตอบสนองต่อปุ๋ยเป็นอย่างไร (ในชั้นนี้ต้องระวังว่าระดับของปุ๋ยไนโตรเจนที่ใช้ในการทดลอง มีระยะที่ห่างเท่ากันหรือไม่ นั่นคือการเพิ่มปริมาณของปุ๋ย จากระดับหนึ่งไปอีกระดับหนึ่งเท่ากันหรือไม่ ถ้าระยะห่างไม่เท่ากันจะต้องคำนวณสัมประสิทธิ์ที่จะใช้เอง ใช้สัมประสิทธิ์จากตารางไม่ได้ สัมประสิทธิ์ที่ใช้ครั้งนี้ คือ

ตำหรับปุ๋ย	N_L	N_Q
0	-8	5
3	-5	0
6	-2	-3
9	1	-4
12	4	-3
18	10	5
ตัวหาร	210	84

และเพราะว่าการทดลองครั้งนี้ Interaction มีนัยสำคัญทางสถิติ ดังนั้นจึงต้องทำการแบ่งแยก SS ของ Interaction ออกเป็นส่วนๆ ตาม main effect ซึ่งอาจจะแบ่งตามตัวใดตัวหนึ่งหรือทั้งสองตัวก็ได้ สำหรับตัวอย่างนี้ทำตามแบบหลัง

เมื่อได้คำนวณ SS แต่ละส่วนตามที่วางแผนไว้แล้วก็จะได้ผลเป็นตารางที่ 6 ดังนี้

ตารางที่ 6 Partitioning Treatment SS

SOV	df	MS	F
1. Variety	2	90,839.00	10.64**
RD5 VS (RD7 & RD9)	(1)	124,206.00	14.56**
RD7 VS RD9	(1)	57,472.00	6.74**
Error	6	8,531.00	
2. Fertilizer	6	51,906.00	25.22**
CK VS treated	(1)	146,746.00	71.31**
N Linear	(1)	112,949.00	54.89**
N Quadratic	(1)	51,072.00	24.82**
N Cubic	(1)	317.00	<1
Residual	(2)	176.00	<1

3. Variety x Fertilizer	12	18,151.92	8.82**
V x CK VS treated	(2)	20,964.50	10.19**
(RD5 VS RD7 & RD9)xN_L	(1)	149,755.00	72.77**
(RD7 VS RD9)xN_L	(1)	8,510.00	4.14*
(RD5 VS RD7 & RD9)xN_Q	(1)	9,651.00	4.69*
(RD7 VS RD9)xN_Q	(1)	556.00	<1
V x N_C	(2)	1,936.00	<1
V x Residual	(4)	887.50	<1
Error (b)	54	2,057.80	

จากตารางที่ 6 จะเห็นว่าแบ่งเป็น 3 ส่วน คือ ส่วนของ main effect (1) , (2) และส่วนของ Interaction (3) ซึ่งเป็นส่วนที่จะต้องเอาใจใส่มากกว่า 2 ส่วนแรก

ส่วนที่ 1 แสดงว่าโดยเฉลี่ยของทุกระดับปุ๋ยพันธุ์ทั้งสามให้ผลผลิตแตกต่างกัน

ส่วนที่ 2 แสดงว่าโดยเฉลี่ยจากทุกสายพันธุ์ การใส่ปุ๋ยทุกตำหรับ เฉลี่ยแล้วทำให้ผลผลิตแตกต่างจากการไม่ใส่ปุ๋ย และการตอบสนองต่อปุ๋ยเป็นไปในรูป Quadratic คือ เมื่อใส่ปุ๋ยผลผลิตเพิ่มอัตราคงที่ระยะหนึ่งแล้วอัตราการเพิ่มจะค่อยๆ ลดลง เป็นผลให้การตอบสนองต่อปุ๋ยเป็นเส้นโค้ง

ส่วนที่ 3 แสดงว่าการศึกษาการตอบสนองต่อปุ๋ยต้องกระทำแยกกันในแต่ละสายพันธุ์โดยที่การแปลผลนี้ต้องใช้ตารางที่ 5 ควบคู่ไปด้วย

V x CK VS. Treated มีนัยสำคัญแสดงว่า การใส่ปุ๋ยจะทำให้ผลผลิตทั้งสามสายพันธุ์เพิ่มไม่เท่ากัน จากตารางที่ 5 คำนวณได้ ดังนี้

	สายพันธุ์		
	กข5	กข7	กข9
ไม่ใส่ปุ๋ย	666	689	585
ใส่ปุ๋ย	699	828	770
เพิ่ม	33	139	185

นั่นคือ ปุ๋ยทุกตำหรับทำให้ผลผลิตของ กข9 เพิ่มจากการไม่ใส่ปุ๋ยมากที่สุดถึง 185 กก./ไร่ รองลงมา ได้แก่ กข7 และ กข5 เพิ่มน้อยที่สุด

(RD5 VS RD7 & RD9) x N_L และ (RD5 VS RD7 & RD9) x N_Q มีนัยสำคัญแสดงว่า เมื่อเฉลี่ยผลผลิตของ กข7 และ กข9 แล้ว เปรียบเทียบกับผลผลิตของ กข5 การตอบสนองต่อปุ๋ยจะแตกต่างกันทั้งในช่วงที่เป็นเส้นตรงและเส้นโค้ง

(RD7 VS RD9)xN_L* และ (RD7 VS RD9)xN_Q^{ns}

แสดงให้เห็นว่าช่วงการตอบสนองต่อปุ๋ยที่เป็นเส้นตรงของทั้งสองพันธุ์ต่างๆ กัน

เพื่อให้เห็นการตอบสนองชัดจึงได้ทำการคำนวณดังต่อไปนี้

Contribution to Linear and Quadratic

	N_L	N_Q
RD5	14,868*	44,390**
RD7	529,660**	5,432^{ns}
RD9	82,246**	11,457*

จากตารางข้างต้น ทำให้เห็นได้ว่าการตอบสนองต่อปุ๋ยไนโตรเจนช่วงระยะ 1 – 18 กก./ไร่ ของ กข7 เป็นเส้นตรง ส่วนของ กข5 และ กข9 เป็นเส้นโค้ง แสดงว่า กข7 มีการตอบสนองต่อปุ๋ยได้สูงกว่าอีก 2 สายพันธุ์ ดังสมการ ดังต่อไปนี้

กข7 : $\hat{Y}$ = 712+14.45* N (r^2 = 0.9489**)

กข5 : $\hat{Y}$ = 675+18.8* N-1.28*N^2 (R^2 = 0.9372**)

กข9 : $\hat{Y}$ = 663+21.49** N-0.64**N^2 (R^2 = 0.9995**)

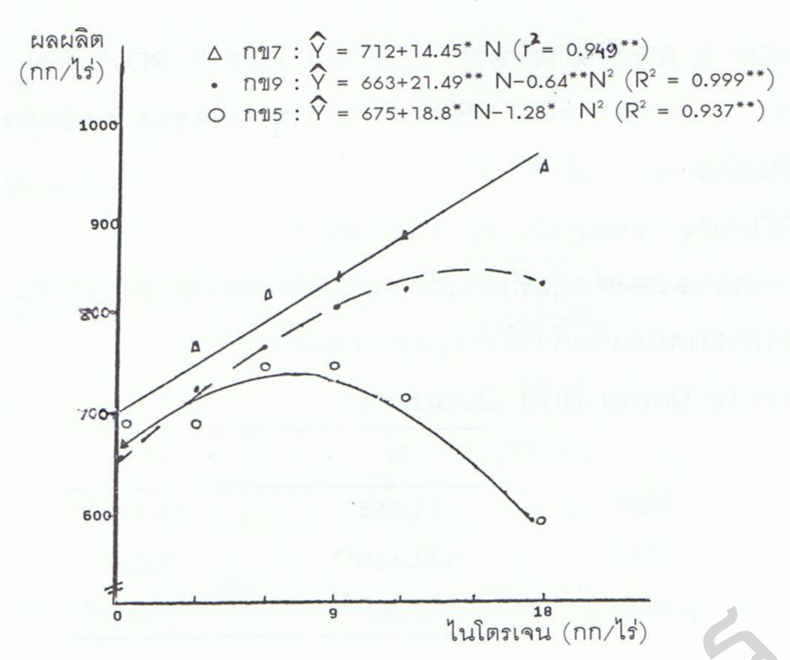
สรุปผล

กข5, กข7 และ กข9 มีการตอบสนองต่อปุ๋ยต่างกัน การใส่ปุ๋ยตำหรับต่างๆ ที่ทำการทดลอง จะทำให้ กข9 เพิ่มผลผลิตสูงกว่าการไม่ใช้ปุ๋ยแตกต่างกับอีกสองสายพันธุ์ คือ เฉลี่ยแล้ว กข9 เพิ่มถึง 185 กก./ไร่ สำหรับ กข7 เพิ่ม 139 กก./ไร่ และ กข5 เพิ่มเพียง 30 กก./ไร่ แต่การตอบสนองต่อปุ๋ยของ กข7 จะเป็นการตอบสนองในเชิงเส้นตรง สำหรับช่วง จาก 0-18 กก./ไร่ กข7 จะให้ผลผลิตเพิ่มคงที่ในอัตรา 14.45 กก./ไร่ ในขณะที่พันธุ์ กข5 และ กข9 จะมีการตอบสนองเป็นเส้นโค้งดังรูป

△ กข7 : $\hat{Y}$ = 712+14.45* N (r^2 = 0.9489**)

● กข9 : $\hat{Y}$ = 663+21.49** N-0.64**N^2 (R^2 = 0.9995**)

○ กข5 : $\hat{Y}$ = 675+18.8* N-1.28*N^2 (R^2 = 0.9372**)



รูปแสดงความสัมพันธ์ระหว่างปุ๋ยไนโตรเจนและผลผลิตของข้าว กข5 กข7 และ กข9

วิธีคำนวณสัมประสิทธิ์ตำหรับ Trend Comparison ที่มีระยะระหว่างระดับไม่เท่ากัน

ตำหรับปุ๋ย	Coded (X_1)	สัมประสิทธิ์ Linear (L)	สัมประสิทธ์ Quadratic (Q)
0	0	a	a
3	1	a + 1	a + b + 1
6	2	a + 2	a + 2b + 4
9	3	a + 3	a + 3b + 9
12	4	a + 4	a + 4b + 16
18	6	a + 6	a + 6b + 36

สมการ Linear $a + X_i$

สมการ Quadratic $a + bX_i + X_i^2$

กฎที่ 1 สัมประสิทธิ์ Linear รวมกันมีค่าเท่ากับ 0

$$a + (a+1)+(a+2)+(a+3)+(a+4)+(a+6) = 0$$

$$6a + 16 = 0$$

$$a = \frac{-16}{6}$$

แทนค่า a ในสัมประสิทธิ์ Linear

ตำหรับปุ๋ย	สัมประสิทธิ์	Linear (L)	(X_6)	(÷2)
0	-16/6	= -16/6	= -16	-8
3	-16/6+1	= -10/6	= -10	-5
6	-16/6+2	= -4/6	= -4	-2
9	-16/6+3	= 2/6	= 2	1
12	-16/6+4	= 8/6	= 8	4
18	-16/6+6	= 20/6	= 20	10

กฎที่ 2 1) สัมประสิทธิ์ Quadratic รวมกันเท่ากับ 0 และ

2) สัมประสิทธิ์ Linear x สัมประสิทธิ์ Quadratic รวมกัน เท่ากับ 0 ดังนั้น

1. $a+(a+b+1)+(a+2b+4)+\ldots\ldots+(a+6b+36) = 0$

$6a+16b+66 = 0$

2. $-8(a)+(-5)(a+b+1)+\ldots\ldots+10(a+6b+36) = 0$

$70b+420 = 0$

$b = -6$

แทนค่า b ใน 1 $6a+16(-6)+66 = 0$

$a = 5$

แทนค่า a และ b ในสัมประสิทธิ์ Quadratic (Q)

ตำหรับปุ๋ย	สัมประสิทธิ์ Quadratic (Q)		
0	a	= 5	= 5
3	a+b+1	= 5+(-6)+1	= 0
6	a+2b+4	= 5+2(-6)+4	= -3
9	a+3b+9	= 5+3(-6)+9	= -4
12	a+4b+16	= 5+4(-6)+16	= -3
18	a+6b+36	= 5+6(-6)+36	= 5

Split plot แบบพิเศษ จากตัวอย่างของ split plot ที่ได้กล่าวมาแล้วจะเห็นว่าเป็นการนำเอา Factorial Experiment มาจัดรูปใหม่ sub plot ในทุก main plot จะต้องเหมือนกัน แต่ในบางครั้ง Sub plot อาจจะไม่ใช่ Factorial combination ก็ได้ เช่น การเปรียบเทียบสายพันธุ์พืชที่มาจากคู่ผสมต่างๆ กัน เปรียบเทียบพันธุ์ข้าวที่สามารถแบ่งเป็นกลุ่มๆ โดยใช้ลักษณะต่างๆ ได้แก่ความสูงหรืออายุเป็นตัวแบ่งกลุ่มโดยมีข้อแม้ว่า จำนวนสายพันธุ์ หรือพันธุ์ในแต่ละคู่ผสมหรือแต่ละกลุ่มต้องมีจำนวนเท่ากัน ดังนี้

(ก) กลุ่มสูง	(ค) กลุ่มเตี้ย	(ข) กลุ่มปานกลาง	
พันธุ์ 1	16	7	
2	17	11	
4	13	9	Block 1
5	15	12	
3	18	8	
6	14	10	

(ค)	(ข)	(ก)	
18	9	3	
14	7	1	
15	12	5	Block 2
13	10	6	
16	11	4	
17	8	2	

(ข)	(ก)	(ค)	
9	2	14	
12	5	16	
11	4	15	Block 3
8	1	17	
10	6	18	
7	3	13	

(ข)	(ค)	(ก)	
8	13	4	
10	18	6	
12	15	5	Block 4
7	17	2	
11	14	3	
9	16	1	

ตารางวิเคราะห์แวเหรียนซ์จะเป็นดังนี้

SOV	df
Total (3x4x6-1)	71
Block (4-1)	3
Group (3-1)	2
E (a)	6
Within Group ก (6 – 1)	5
Within Group ข (6 – 1)	5
Within Group ค (6 – 1)	5
E (b)	45

C.V.(a) %

C.V.(b) %

ตารางค่าเฉลี่ยจะมีลักษณะดังนี้

	กลุ่ม ก.		กลุ่ม ข.		กลุ่ม ค.	
	สายพันธุ์	ค่าเฉลี่ย	สายพันธุ์	ค่าเฉลี่ย	สายพันธุ์	ค่าเฉลี่ย
	1		7		13	
	2		8		14	
	3		9		15	
	.		.		.	
	.		.		.	
	.		.		.	
	6	______	12	______	18	______
ค่าเฉลี่ย	ก	______	ข	______	ค	______

$S_{\bar{d}}$ ของสายพันธุ์ในกลุ่มเดียวกัน $= \sqrt{2Eb/r}$

$S_{\bar{d}}$ ของสายพันธุ์ที่ต่างกลุ่มกัน $= \sqrt{\frac{2}{br}\left(Ea + (b-1)Eb\right)}$

$S_{\bar{d}}$ ของกลุ่ม $= \sqrt{2Ea/br}$

(r = จำนวนซ้ำ b = จำนวนสายพันธุ์ในกลุ่ม)

Split Block (or Strip Plot) Design

เป็นการทดลองแบบหนึ่งที่ดัดแปลงไปจาก Split plot design ที่มี main plot จัดในรูป RCB โดยการจัด Sub plot เดียวกันของแต่ละ main plot ให้ตรงกัน เป็น strip ดังนี้

A_1	B_2	B_4	B_3	B_1
A_2	B_2	B_4	B_3	B_1
A_3	B_2	B_4	B_3	B_1

Block 1

ด้วยการจัดแบบนี้จะทำให้เกิดขนาดของแปลง (plot size) ขึ้น 3 ขนาด คือ แปลงสำหรับปัจจัย A แปลงสำหรับปัจจัย B และ แปลงสำหรับการทดสอบ Interaction AxB

เมื่อเทียบกับ Design แบบ Split plot จะมีความแตกต่างกันดังนี้

1. ความสำคัญของปัจจัยที่ทดสอบ ทั้งปัจจัย A และ B จะถูกทดสอบด้วยน้ำหนักเท่ากัน และต่างก็ได้รับพิจารณาเหมือนว่าเป็น Main plot ทั้งคู่

2. Randomization ของ Lay Out ใน Block หนึ่ง จะทำการสุ่มปัจจัย A ลงก่อนแล้วสุ่มปัจจัย B ครั้งเดียว ไม่เหมือนกับ Split plot ที่ต้องสุ่ม Sub plot ในทุก main plot แต่ยังคงต้องสุ่มใหม่ทุก Block เหมือนกันทั้งสอง Design

3. ในการวิเคราะห์แวเหรียนซ์ของ Split Block design จะมี Mean Square Error 3 ตัว เนื่องจากมีขนาดแปลง 3 ขนาด ดังนั้นจึงมี C.V. ทั้งหมด 3 ค่า C.V.(a) และ C.V.(b) จะมีขนาดใหญ่กว่า C.V.(c) ถ้าเป็นการวางแผนอย่างมีประสิทธิภาพ ระหว่าง C.V.(a) และ C.V.(b) ค่าใดจะใหญ่กว่าขึ้นอยู่กับลักษณะของปัจจัยที่ทดสอบ

Design นี้ จะถูกใช้เมื่อ

1. ปัจจัยที่ทดสอบทั้งสองต้องทำในแปลงขนาดใหญ่ทั้งคู่
2. เป็นไปได้ในทางปฏิบัติ
3. ในระหว่างการทดลองต้องการจะแสดงให้เห็นความแตกต่างระหว่างระดับต่างๆ ของแต่ละปัจจัยอย่างชัดเจน

ตารางวิเคราะห์แวเหรียนซ์ของ Split Block Design ที่มีจำนวนระดับปัจจัย A=a ระดับปัจจัย B=b และทำการทดลอง r Block จะเป็นดังนี้

ANOV

SOV	df	SS	MS	F
Total	abr-1			
Block	r-1			
A	a-1			
Error (a)	(a-1)(r-1)			
B	b-1			
Error (b)	(b-1)(r-1)			
A x B	(a-1)(b-1)			
Error (c)	(a-1)(b-1)(r-1)			

$$\text{C.V. (a)} = \frac{\sqrt{\text{MS Error(a)}}}{\text{mean}} \times 100\%$$

$$\text{C.V. (b)} = \frac{\sqrt{\text{MS Error(b)}}}{\text{mean}} \times 100\%$$

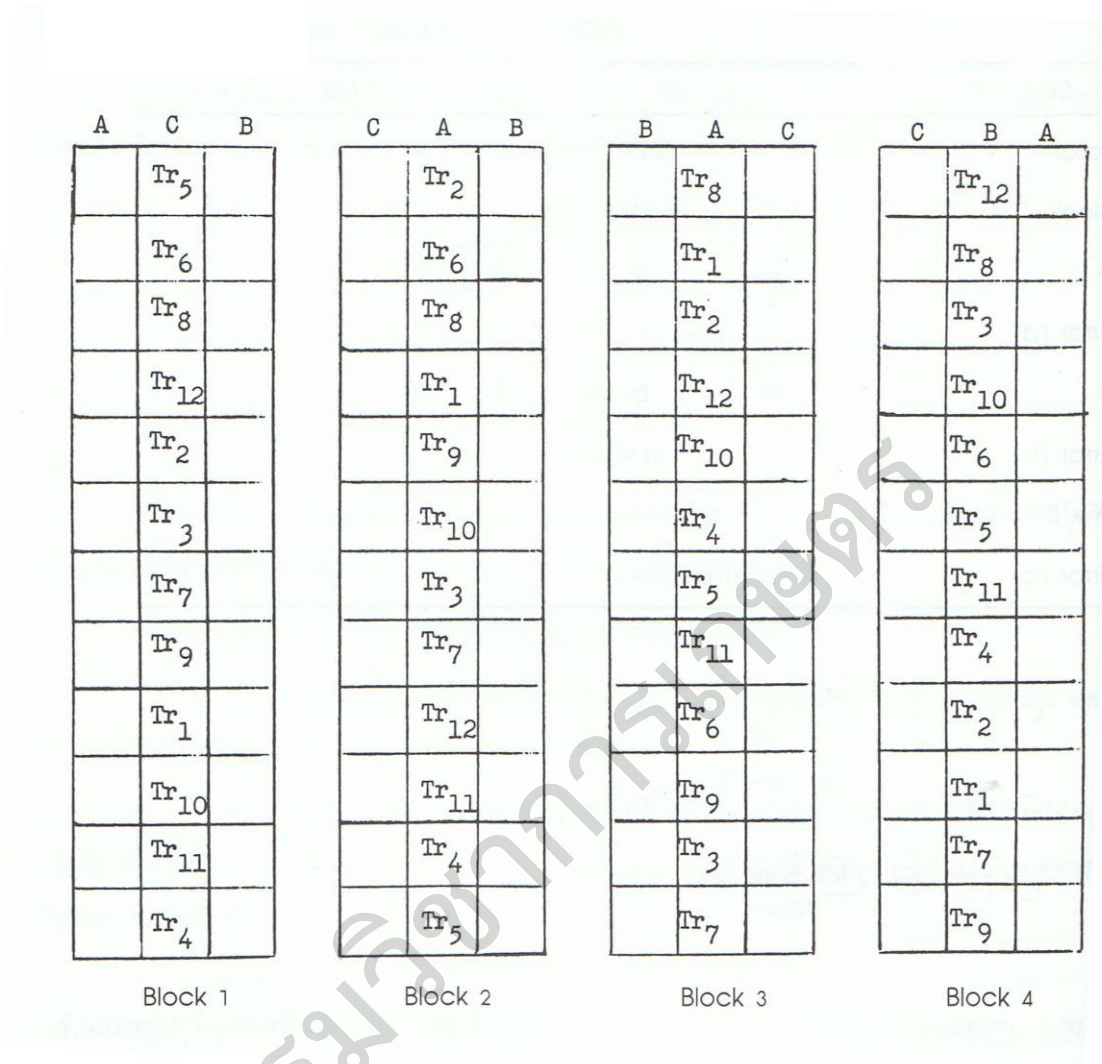
$$\text{C.V. (c)} = \frac{\sqrt{\text{MS Error(c)}}}{\text{mean}} \times 100\%$$

ตัวอย่างการทดลองการแบ่งใส่ปุ๋ยไนโตรเจนสองระดับกับข้าวสามสายพันธุ์ที่สถานีทดลองข้าว สุพรรณบุรี ในฤดูนาปี 2516

ปัจจัยที่ 1 ข้าว 3 พันธุ์ ปัจจัยที่ 2 วิธีการใส่ปุ๋ย 12 ตำหรับ

ข้าว 3 พันธุ์	วิธีการใส่ปุ๋ย	
A = LPT 123	Tr 1 = $\frac{6}{0}$-12-6	Tr 7 = $\frac{12}{0}$-12-6
B = RD5	2 = $\frac{4}{2}$-12-6	8 = $\frac{8}{4}$-12-6
C = KDML 105	3 = $\frac{3}{3}$-12-6	9 = $\frac{6}{6}$-12-6
	4 = $\frac{2}{4}$-12-6	9 = $\frac{4}{8}$-12-6
	5 = $\frac{0}{6}$-12-6	9 = $\frac{0}{12}$-12-6
	6 = 0-0-0	9 = 0-12-6

แผนผังแปลงทดลอง



ตารางที่ 1 ผลผลิต (กก./ไร่)

Treatment	Variety	Block 1	Block 2	Block 3	Block 4
(6,0)-12-6	KDML105	583	557	537	603
	RD5	580	568	534	576
	LPT123	594	531	508	596
(4,2)-12-6	KDML105	613	621	573	742
	RD5	576	573	623	762
	LPT123	592	611	582	716
(3,3)-12-6	KDML105	430	579	629	569
	RD5	590	593	673	502
	LPT123	583	593	612	570
(2,4)-12-6	KDML105	545	535	614	586
	RD5	558	617	557	590
	LPT123	543	637	609	713
(0,6)-12-6	KDML105	608	618	692	649
	RD5	616	649	766	660
	LPT123	579	690	760	779
0-0-0(CK)	KDML105	385	224	277	574
	RD5	330	227	278	580
	LPT123	321	269	370	605
(12,0)-12-6	KDML105	647	598	620	650
	RD5	582	611	546	639
	LPT123	489	544	496	571
(8,4)-12-6	KDML105	769	716	760	715
	RD5	665	773	752	604
	LPT123	643	672	612	709
(6,6)-12-6	KDML105	730	686	689	698
	RD5	661	662	782	634
	LPT123	694	651	658	728
(4,8)-12-6	KDML105	782	720	678	688
	RD5	763	741	637	684
	LPT123	759	750	653	670
(0,12)-12-6	KDML105	562	543	587	547
	RD5	677	579	830	764
	LPT123	605	725	777	756
0-12-6	KDML105	507	435	364	531
	RD5	480	405	381	488
	LPT123	472	512	457	541

I. จากตารางที่ 1 คำนวณตารางต่างๆ ดังนี้

ตารางที่ 2 A x B

ปุ๋ย (B)	พันธุ์ (A)			รวม (B)
	KDML105	RD5	LPT123	
(6,0)-12-6	2,280	2,258	2,229	6,767
(4,2)-12-6	2,549	2,534	2,501	7,584
(3,3)-12-6	2,207	2,358	2,358	6,923
(2,4)-12-6	2,280	2,322	2,502	7,104
(0,6)-12-6	2,567	2,691	2,808	8,066
0-0-0	1,460	1,415	1,565	4,440
(12,0)-12-6	2,515	2,378	2,100	6,993
(8,4)-12-6	2,960	2,794	2,636	8,390
(6,6)-12-6	2,803	2,699	2,731	8,233
(4,8)-12-6	2,868	2,825	2,832	8,525
(0,12)-12-6	2,239	3,030	2,863	8,132
0-12-6	1,837	1,754	1,982	5,573
(A)	28,565	29,058	29,107	86,730

ตัวอย่างKDML 105 ระดับ (6,0)-12-6 = 583+557+537+603 = 2,280

ตารางที่ 3 A x Block

พันธุ์ (A)	Block				รวม (A)
	1	2	3	4	
KDML 105	7,161	6,832	7,020	7,552	28,565
RD 5	7,078	7,138	7,359	7,483	29,058
LPT 123	6,874	7,185	7,094	7,954	29,104
รวม (Block)	21,113	21,155	21,473	22,989	86,730

ตัวอย่างKDML 105 Block 1 = 583+613+430+...+507 = 7,161

ตารางที่ 4 B x Block

ปุ๋ย (B)	Block				รวม (B)
	1	2	3	4	
(6,0)-12-6	1,757	1,656	1,579	1,775	6,767
(4,2)-12-6	1,781	1,805	1,778	2,220	7,584
(3,3)-12-6	1,603	1,765	1,914	1,641	6,923
(2,4)-12-6	1,646	1,789	1,780	1,889	7,104
(0,6)-12-6	1,803	1,957	2,218	2,088	8,066
0-0-0	1,036	720	925	1,759	4,440
(12,0)-12-6	1,718	1,753	1,662	1,860	6,993
(8,4)-12-6	2,077	2,161	2,124	2,028	8,390
(6,6)-12-6	2,085	1,959	2,129	2,060	8,233
(4,8)-12-6	2,304	2,211	1,968	2,042	8,525
(0,12)-12-6	1,844	2,027	2,194	2,067	8,132
0-12-6	1,459	1,352	1,202	1,560	5,573
รวม (Block)	21,113	21,155	21,473	22,989	86,730

ตัวอย่าง (6,0)-12-6 Block 1 = 583+580+594 = 1,757

II. คำนวณค่าวิเคราะห์วาเหรียนซ์ต่างๆ จากตารางที่ 1-4 ดังนี้

Mean = 86,730/3x12x4 = 602

Correction Factor (C.F.) = $(86,7300)^2$/3x12x4

จากตารางที่ 1

Total SS = $583^2+557^2+\ldots+541^2$ - C.F.

= 2,050,354

จากตารางที่ 3

Block SS = $\left((21,113^2+21,155^2+21,473^2+22,989^2)/3\text{x}12\right)$ - C.F.

= 65,373

(A) SS = $\left((28,565^2+29,058^2+29,107^2)/4\text{x}12\right)$ - C.F.

= 3,745

Error (a) SS = $\left((7,161^2+6,832^2+\ldots+7,954^2)/12\right)$-C.F.-(A) SS-Block SS

= 67,547

จากตารางที่ 4

(A) SS = $\left((6,767^2+7,584^2+....+5,573^2)/3x4\right)$ – C.F.

= 1,381,399

Error (a) SS = $\left((1,757^2+1,656^2+...+1,560^2)/3\right)$-C.F.-(B) SS-Block SS

= 338,924

จากตารางที่ 2

(A x B) SS = $(2,280^2+2,258^2+...+1,982^2)/4$-C.F.-(A) SS-(B) SS

= 148,599

Error (C) = Total SS-Block SS-(A) SS-(B) SS-(AxB) SS-Error (a) -Error(b)

= 44,767

ตารางที่ 5 ตารางวิเคราะห์วาเหรียนซ์ของผลผลิต (กก./ไร่)

SOV	df	SS	MS	F
Total	143	2,050,354	21,791	
Block	3	65,373	1,872	
พันธุ์ (A)	2	3,745	11,258	<1
Error (a)	6	67,547		
ปุ๋ย (B)	11	1,381,399	125,582	12.23**
Error (b)	33	338,924	10,270	
A x B	22	148,599	6,755	9.96**
Error (c)	66	44,767	678	

<u>คำนวณ C.V.</u>

C.V. (a) = $(\sqrt{11,258}/602)x100$

= 17.6%

C.V. (b) = $(\sqrt{10,270}/602)x100$

= 16.8%

C.V. (c) = $(\sqrt{678}/602)x100$

= 4.3%

III. ทำตารางค่าเฉลี่ยจากตารางที่ 2 และ คำนวณค่า LSD สำหรับเปรียบเทียบปัจจัย B และ Interaction ดังนี้

1. <u>สำหรับเปรียบเทียบปัจจัย A ในปัจจัย B เดียวกัน</u>

$$S_{\bar{d}} = \sqrt{2\left((b-1)E_c + E_a\right)/rb}$$

$$t' = \frac{(b-1)E_c t_c + E_a t_a}{(b-1)E_c + E_a}$$

t.05,(.01) = 2.447, (3.707)

df6

t.05,(.01) = 1.998, (2.656)

df66

$$t'.05 = \frac{(12-1)(678)(1.998) + (11{,}258)(2.447)}{(12-1)(678) + 11{,}258}$$

$$= 2.268$$

$$t'.01 = \frac{(12-1)(678)(2.656) + (11{,}258)(3.707)}{(12-1)(678) + 11{,}258}$$

$$= 3.288$$

$$S_{\bar{d}} = \sqrt{2\left((12-1)(678) + 11{,}258\right)/4x12}$$

$$= 27.9255$$

$$LSD = t' \times S_{\bar{d}}$$

สำหรับเปรียบเทียบค่าเฉลี่ยเนื่องจากปัจจัย A ในปัจจัย B เดียวกัน

LSD.05 = 27.9255 x 2.268

= 63.34

LSD.01 = 27.9255 x 3.288

= 91.819

2. <u>สำหรับเปรียบเทียบปัจจัย B ในปัจจัย A เดียวกัน</u>

$S_{\bar{d}} = \sqrt{2\left((a-1)E_c + E_b\right)/ra}$

$t' = \dfrac{(a-1)E_c t_c + E_b t_b}{(a-1)E_c + E_b}$

t.05,(.01) = 2.036, (2.736)

df33

t.05,(.01) = 1.998, (2.656)

df66

$t'.05 = \dfrac{(3-1)(678)(1.998) + (10,270)(2.036)}{(3-1)(678) + 10,270}$

= 2.032

$t'.01 = \dfrac{(3-1)(678)(2.656) + (10,270)(2.736)}{(3-1)(678) + 10,270}$

= 2.727

$S_{\bar{d}} = \sqrt{2\left((3-1)(678) + 10,270\right)/4x3}$

= 44.0189

สำหรับเปรียบเทียบค่าเฉลี่ยเนื่องจากปัจจัย B ในปัจจัย A เดียวกัน

LSD.05 = 2.032 x 44.0189

= 89.45

LSD.01 = 2.727 x 44.0189

= 120.04

ตารางที่ 6 ค่าเฉลี่ยผลผลิต (กก./ไร่)

Fert	Variety			Fert-mean
N-P-K	KDML 105	RD5	LPT 123	
(6+0)-12-6	570	564	557	564
(4+2)-12-6	637	633	625	632
(3+3)-12-6	551	589	589	577
(2+4)-12-6	570	580	625	592
(0+6)-12-6	642	673	702	672
CK(0-0-0)	365	354	391	370
(12+0)-12-6	629	594	525	583
(8+4)-12-6	740	698	659	699
(6+6)-12-6	701	675	683	686
(4+8)-12-6	717	706	708	710
(0+12)-12-6	560	757	716	678
0-12-6	459	438	495	464
Var-mean	595	605	606	

$LSD_{.05}$ (สำหรับเปรียบเทียบค่าเฉลี่ยของพันธุ์ 2 พันธุ์ ที่ใช้ปุ๋ยระดับเดียวกัน) .05 = 63 กก./ไร่
$LSD_{.01}$ (สำหรับเปรียบเทียบค่าเฉลี่ยของพันธุ์ 2 พันธุ์ ที่ใช้ปุ๋ยระดับเดียวกัน) .01 = 92 กก./ไร่
$LSD_{.05}$ (สำหรับเปรียบเทียบค่าเฉลี่ยของระดับปุ๋ย 2 ระดับในพันธุ์เดียวกัน) .05 = 89 กก./ไร่
$LSD_{.01}$ (สำหรับเปรียบเทียบค่าเฉลี่ยของระดับปุ๋ย 2 ระดับในพันธุ์เดียวกัน) .01 = 120 กก./ไร่

จากตารางที่ 5 จะเห็นว่า F ของ B และ AxB มีนัยสำคัญยิ่งทางสถิติ แต่เนื่องจากทั้งสองส่วนมี df สูง คือ 11 และ 22 จึงทำให้การแปลผลการทดลองในตารางที่ 6 ยาก ควรจะแยก df และ SS ของทั้งสองส่วนออกเป็นส่วนย่อยๆ ตามลักษณะที่ต้องการเปรียบเทียบ เมื่อพิจารณาระดับปุ๋ยที่ทดสอบอาจจะแยกได้ดังนี้

1. การเปรียบเทียบระหว่างการใช้และไม่ใช้ไนโตรเจนกลุ่ม (0,0,0) และ (0-12-6) VS (กลุ่ม N_6 และกลุ่ม N_{12}) เพื่อดูว่าการใช้ไนโตรเจนจะทำให้ผลผลิตต่างกับการไม่ใช้ไนโตรเจนเท่าใด

2. ในกลุ่มของTreatment ที่ไม่ใช้ไนโตรเจน เปรียบเทียบ 0-0-0 และ 0-12-6 เพื่อดูอิทธิพลของ P และ K

3. เปรียบเทียบ treatment กลุ่มที่ใช้ไนโตรเจนระดับ 6 และ กลุ่มที่ใช้ไนโตรเจนระดับ 12 ว่าจะทำให้ผลผลิตต่างกันอย่างไร

4. ในกลุ่มของ treatment ที่ใช้ไนโตรเจนเท่ากัน (คือกลุ่มของ 6 หรือกลุ่มของ 12) การแบ่งใส่ไนโตรเจนเป็นอัตราส่วนต่างกันจะทำให้ผลผลิตต่างกันหรือไม่

จากการแยก df และ SS ตามที่อธิบายไว้ข้างต้นจะให้ผลดังที่แสดงไว้ในตารางที่ 7 และตารางค่าเฉลี่ยอีก 1 ตาราง คือ ตารางที่ 8 ดังนี้

ตารางที่ 7 ตารางวิเคราะห์วาเหรียนซ์ผลผลิต (กก./ไร่)

Source	df	SS	MS	F
Total	143	2,050,354		
Block	3	65,373	21,791	1.94^{ns}
Variety (V)	2	3,745	1,872	<1
Error (a)	6	67,547	11,258	
Fertilizer (F)	11	1,381,399	125,582	12.23**
(CK+0-12-6) VS (N_6+N_{12})	1	896,568	896,568	87.30**
CK VS 0-12-6	1	53,487	53,487	5.21**
N_6 VS N_{12}	1	122,177	122,177	11.90**
Among N_6	4	94,285	23,571	2.30^{ns}
Among N_{12}	4	124,882	31,220	3.04*
Error (b)	33	338,924	10,270	
Variety x Fertilizer (VxF)	22	148,599	6,755	9.96**
Vx(CK+0-12-6) VS (N_6+N_{12})	2	10,322	5,161	7.61**
VxCK VS 0-12-6	2	380	190	<1
VxN_6 VS N_{12}	2	9,890	4,945	7.29**
V x Among N_6	8	11,998	1,500	2.21*
V x Among N_{12}	8	116,009	14,501	21.39**
Error (c)	66	44,767	678	

C.V. (a) = 17.6% C.V. (b) = 16.8% C.V. (c) = 4.3%

ตารางที่ 8 ค่าเฉลี่ยผลผลิต (กก./ไร่)

Fert N-P-K	Variety KDML 105	RD5	LPT 123	mean
N_6	594	615	620	609
N_{12}	669	684	658	670
CK (0-0-0)	365	354	391	370
0-12-6	459	438	495	464

$LSD_{.05}$ (For comparison of N_6 to N_{12} within the same Variety) .05 = 40 kg/rai

$LSD_{.01}$ (For comparison of N_6 to N_{12} within the same Variety) .01 = 54 kg/rai

การแปลผล

จากตารางที่ 7 ให้พิจารณาส่วนของ F และ VxF พร้อมๆ กัน

1. (CK+0-12-6) VS (N_6+N_{12})F = 87.30** แสดงว่าการใช้ปุ๋ยไม่ว่าจะเป็น N_6 หรือ N_{12} จะให้ผลผลิตเฉลี่ยต่างกับการไม่ใช้ปุ๋ยไนโตรเจนจากตารางที่ 8 จะเห็นว่าผลผลิตของการใช้ปุ๋ย N_6 และ N_{12} เฉลี่ยได้ = $\frac{1}{2}(609+670)$ ส่วนการไม่ใช้ปุ๋ยไนโตรเจนได้ผลผลิตเฉลี่ย = $\frac{1}{2}(370+464)$

V x (CK+0-12-6) VS (N_6+N_{12}) F = 7.61** แสดงว่า การใช้ปุ๋ยไนโตรเจน (N_6 กับ N_{12}) กับการไม่ใช้ปุ๋ยไนโตรเจน (CK กับ 0-12-6) จะทำให้ผลผลิตแตกต่างกันและความแตกต่างกันนี้ต่างกันในระหว่างพันธุ์ที่ใช้ทดสอบ

จากตารางที่ 8

KDML 105	ใช้ปุ๋ยได้ผลผลิตเฉลี่ย	=	$\frac{1}{2}(594+669)$	=	632
	เทียบกับไม่ใช้ปุ๋ย N	=	$\frac{1}{2}(365+459)$	=	412
RD5	ใช้ปุ๋ย N	=	$\frac{1}{2}(615+684)$	=	650
	ไม่ใช้ปุ๋ย N	=	$\frac{1}{2}(354+438)$	=	396
LPT 123	ใช้ปุ๋ย N	=	$\frac{1}{2}(620+658)$	=	639
	ไม่ใช้ปุ๋ย N	=	$\frac{1}{2}(391+495)$	=	443

2. CK VS 0-12-6 F = 5.21* แสดงว่า การไม่ใช้ P และ K จะทำให้ผลผลิตต่างกับการใช้ (370 ต่างกับ 464)

Vx CK VS 0-12-6 ค่า F น้อยกว่า 1

แสดงว่าทุกพันธุ์จะให้การตอบสนองต่อ P และ K ไม่ต่างกัน

KDML ; 365 น้อยกว่า 459

RD5 ; 354 น้อยกว่า 438

LPT 123 ; 391 น้อยกว่า 495

3. N_6 VS N_{12} F = 11.90** แสดงว่าการใช้ไนโตรเจนระดับ 6 ทำให้ผลผลิตต่างกับระดับ 12 609 น้อยกว่า 670

V x N_6 VS N_{12} ค่า F = 7.29** แสดงว่าผลผลิตที่ได้จากการใช้ N_6 และ N_{12} ในทุกพันธุ์มีความแตกต่างไม่เหมือนกัน

KDML 594 เทียบกับ 669

RD5 615 เทียบกับ 684

LPT 123 620 เทียบกับ 658

เมื่อนำค่าเปรียบเทียบนี้ไปทดสอบกับค่า LSD ที่ให้ไว้ในตารางที่ 8 นี้ จะเห็นว่า LPT 123 ใช้ N_6 และ N_{12} ผลผลิตไม่มีความแตกต่างกัน

4. ในกลุ่มของ Treatment ที่ใช้ไนโตรเจนเท่ากัน

4.1 Among N_6 F = 2.30^{ns} แสดงว่าในการแบ่งใส่ไนโตรเจน 6 กก./ไร่ ด้วยอัตราส่วนต่างๆ กัน ไม่ทำให้ผลผลิตโดยเฉลี่ยต่างกัน แต่ V x Among N_6 F = 2.21* แสดงว่าที่ว่าโดยเฉลี่ย ไม่ต่างกันนั้นไม่จริงเสมอไป บางพันธุ์ อาจจะตอบสนองต่ออัตราส่วนของการใช้ปุ๋ยต่างๆ ในขณะที่บางพันธุ์อาจจะไม่มีการตอบสนอง ดูค่าเฉลี่ยในตารางที่ 6 5 treatment แรก คือ (6+0), (4+2), (3+3), (2+4), (0+6)

	KDML 105	RD5	LPT123
6+0	570	564	557
4+2	637	633	625
3+3	551	589	589
2+4	570	580	625
0+6	642	673	702

โดยใช้ค่า LSD ที่แสดงไว้ท้ายตารางที่ 6

LSD สำหรับเปรียบเทียบผลผลิตของตำหรับปุ๋ย 2 ตำหรับ ในพันธุ์เดียวกัน (89,120 กก./ไร่) เป็นตัวเปรียบเทียบ

4.2 Among N_{12} และ V x Among N_{12} จะแปลผลและใช้ค่าต่างๆ ในการประกอบการแปลผลเช่นเดียวกับ 4.1

เตือนความจำ

สำหรับการทดลอง 2 ปัจจัย ที่มีปัจจัย A มีจำนวนระดับ = a

ปัจจัย B มีจำนวนระดับ = b

ทำการทดลองด้วยจำนวนซ้ำ = r

ก. Factorial Experiment

ในการเปรียบเทียบความแตกต่างระหว่างทรีตเมนต์แต่ละคู่ ใช้ Standard Error of Difference ($S_{\bar{d}}$) ดังนี้

$$A : S_{\bar{d}} = \sqrt{2E/br}$$

$$B : S_{\bar{d}} = \sqrt{2E/ar}$$

A x B : (A ในแต่ละระดับของ B หรือ B ในแต่ละระดับของ A)

$$S_{\bar{d}} = \sqrt{2E/r}$$

(E = Error)

ข. Split Plot Design ที่มี A เป็น main plot

$$A : S_{\bar{d}} = \sqrt{2E_a/br}$$

$$B : S_{\bar{d}} = \sqrt{2E_b/ar}$$

A x B : มี $S_{\bar{d}}$ 2 ตัว

B ใน A ระดับเดียวกัน $S_{\bar{d}} = \sqrt{2E_b/r}$

A ใน B ระดับเดียวกันหรือคนละระดับ

$$S_{\bar{d}} = \sqrt{\frac{2}{br}\left((b-1)E_b + E_a\right)}$$

และ $$t' = \frac{(b-1)E_b t_b + E_a t_a}{(b-1)E_b + E_a}$$

ค. Strip plot Design

$$A \quad : \quad S_{\bar{d}} = \sqrt{2E_a/br}$$

$$B \quad : \quad S_{\bar{d}} = \sqrt{2E_b/ar}$$

$$A \text{ ใน } B \text{ ระดับเดียวกัน } S_{\bar{d}} = \sqrt{\frac{2}{br}\left((b-1)E_c + E_a\right)}$$

$$B \text{ ใน } A \text{ ระดับเดียวกัน } S_{\bar{d}} = \sqrt{\frac{2}{ar}\left((a-1)E_c + E_b\right)}$$

3 Factors Factorial Experiment

ทำการทดลอง 3 ปัจจัยพร้อมๆ กัน เช่น การศึกษาการตอบสนองของข้าว กข7 ต่อปุ๋ยไนโตรเจน (N) ฟอสเฟต (P) และ โปแตช (K) อย่างละ 3 ระดับ (N 3, 6, 12 กก./ไร่ P 0, 6, 12 กก./ไร่ K 0, 6, 12 กก./ไร่)

Treatment Combination ที่จะใช้ในการทดสอบจะมีจำนวน 3x3x3 = 27

Treatment คือ		N-P-K		N-P-K		N-P-K
	T_1	3–0–0	T_{10}	6–0–0	T_{19}	12–0–0
	T_2	3–0–6	T_{11}	6–0–6	T_{20}	12–0–6
	T_3	3–0–12	T_{12}	6–0–12	T_{21}	12–0–12
	T_4	3–6–0	T_{13}	6–6–0	T_{22}	12–6–0
	T_5	3–6–6	T_{14}	6–6–6	T_{23}	12–6–6
	T_6	3–6–12	T_{15}	6–6–12	T_{24}	12–6–12
	T_7	3–12–0	T_{16}	6–12–0	T_{25}	12–12–0
	T_8	3–12–6	T_{17}	6–12–6	T_{26}	12–12–6
	T_9	3–12–12	T_{18}	6–12–12	T_{27}	12–12–12

เหมือนกับ 2 Factors Factorial Experiment คือ จำนวน Treatment Combination เหล่านี้จะจัดการทดลองตามแบบการทดลองพื้นฐานแบบใดแบบหนึ่งใน 3 แบบ คือ CRD หรือ RCB หรือ LS ก็ได้ โดยใช้หลักและวิธีการ Random เหมือนกับหลักของ Design นั้นๆ

เมื่อทำการทดลอง 3 ปัจจัยพร้อมๆ กัน จะทำการศึกษา :-

Main effect คือ การตอบสนองต่อปัจจัยหนึ่ง หรือการเปลี่ยนแปลงของผลผลิต เมื่อเกิดการเปลี่ยนแปลงระดับของแต่ละปัจจัย main effect ในที่นี้มี 3 ตัวคือ N P และ K

Interaction การศึกษา 3 ปัจจัยจะมี Interaction 2 หมู่

First order interaction ได้แก่ interaction ที่เกิดระหว่างสองปัจจัยคู่ใดคู่หนึ่ง คือ NxP, NxK และ PxK หรือเรียกอีกอย่างหนึ่งว่า Two-factor interaction

Second order interaction ได้แก่ interaction ที่เกิดระหว่างสามปัจจัยนั้นๆ คือ NxPxK หรือ three-factor interaction.

<u>ตัวอย่าง</u> 3 factor Factorial Experiment

โครงการ วิธีการใช้ปุ๋ยแบบผสมผสานเพื่อเพิ่มประสิทธิภาพการใส่ปุ๋ยเคมี

เรื่อง การใช้ปุ๋ยกับข้าวโพดฝักอ่อน ที่ปลูกต่อเนื่องกันในสภาพดินไร่ที่มีการชลประทาน

สถานที่ทดลอง ไร่กสิกร จังหวัดราชบุรี

ปัจจัยที่ศึกษา

A 2 ระดับ
- ตัดต้นออกจากแปลง
- สับกลบลงดิน

B 2 ระดับ
- ไม่ใส่มูลวัว
- ใส่มูลวัว 1 ตัน/ไร่/ปี

C 3 ระดับ
- ปุ๋ย N 0, 10, 12 กก./ไร่ (โดยใส่ P, K คงที่อย่างละ 5 กก./ไร่)

วัตถุประสงค์ เพื่อศึกษาถึงสัดส่วนของปุ๋ยไนโตรเจนและปุ๋ยคอกที่เหมาะสมต่อการปลูกข้าวโพดฝักอ่อนระยะยาวโดยมีการสับกลบต้นข้าวโพดลงดินและตัดทิ้ง

แผนผังแปลงทดลอง

$A_1B_1C_3$ 6	$A_1B_2C_2$ 12	$A_1B_2C_3$ 18	$A_1B_1C_1$ 24	$A_2B_1C_1$ 30	$A_1B_1C_3$ 36
$A_2B_2C_1$ 5	$A_2B_2C_2$ 11	$A_2B_1C_3$ 17	$A_1B_2C_1$ 23	$A_2B_2C_2$ 29	$A_2B_2C_1$ 35
$A_2B_1C_1$ 4	$A_2B_1C_2$ 10	$A_2B_2C_3$ 16	$A_2B_1C_1$ 22	$A_1B_1C_1$ 28	$A_2B_2C_3$ 34
$A_1B_1C_2$ 3	$A_1B_1C_1$ 9	$A_2B_2C_1$ 15	$A_2B_2C_2$ 21	$A_2B_1C_2$ 27	$A_1B_2C_3$ 33
$A_2B_2C_3$ 2	$A_1B_2C_1$ 8	$A_1B_1C_2$ 14	$A_2B_1C_2$ 20	$A_1B_2C_1$ 26	$A_1B_2C_2$ 32
$A_2B_1C_3$ 1	$A_1B_2C_3$ 7	$A_1B_2C_2$ 13	$A_1B_1C_3$ 19	$A_2B_1C_3$ 25	$A_1B_1C_2$ 31
← I →		← II →		← III →	

ขนาดแปลงย่อย 5 x 6 เมตร (10 แถวๆ ละ 12 ต้น)

ขนาดพื้นที่เก็บเกี่ยว 3 x 4 เมตร (6 แถวๆ ละ 8 ต้น)

การทดลองนี้ทำต่อเนื่องกันปีละ 4 ครั้ง ผลของการปลูกครั้งที่ 3 น้ำหนักฝักข้าวโพด (กก./ไร่) เป็นดังนี้

DATA SET.NO.13 OPERATOR NAME : TATSANEE

FACTORIAL 3 FACTORS IN RCB

NO. OF A = 2 NO. OF B = 2 NO.OF C = 3 NO. OF REPLICATION = 3

RAW DATA

			R1	R2	R3
A1	B1	C1	265.0000	89.0000	269.0000
		C2	1,009.0000	1,289.0000	788.0000
		C3	1,085.0000	1,111.0000	941.0000
	B2	C1	660.0000	347.0000	447.0000
		C2	908.0000	1,117.0000	916.0000
		C3	1,153.0000	988.0000	1,029.0000
A2	B1	C1	775.0000	592.0000	743.0000
		C2	1,125.0000	1,147.0000	1,047.0000
		C3	1,123.0000	1,139.0000	1,164.0000
	B2	C1	1,063.0000	847.0000	489.0000
		C2	1,071.0000	1,057.0000	1,132.0000
		C3	1,243.0000	1,269.0000	863.0000

ANALYSIS OF VARIANCE

SOV	df	SS	MS	F
REPLICATION	2	120,059.5555	60,029.777	
A	1	336,013.4444	336,013.4444	16.88**
B	1	22,400.1111	22,400.1111	1.12
C	2	2,195,246.8888	1,097,623.4444	55.14*
AB	1	8,100.0000	8,100.0000	0.40
AC	2	202,600.2222	101,300.1111	5.08*
BC	2	85,628.2222	42,814.1111	2.15^{ns}
ABC	2	17,584.6666	8,792.3333	0.44
ERROR	22	437,909.1111	19,904.9595	
TOTAL	35	3,425,542.2222		

C.V. = 15.7%

		5%	1%
LEVEL OF SIGNIFICANCE		5%	1%
T-VALUE		2.0739	2.8188
LSD FOR			
	A	97.5320	132.5634
	B	97.5320	132.5634
	C	119.4518	162.3563
	AB	137.9311	187.4729
	AC	168.9304	229.6065
	BC	168.9304	229.6065
	ABC	238.9036	324.7127

ABC-MEANS		C1	C2	C3
A1	B1	207.6666	1,028.6666	1,045.6666
	B2	484.6666	980.3333	1,056.6666
A2	B1	703.3333	1,106.3333	1,142.0000
	B2	799.6666	1,086.6666	1,125.0000

AB-MEANS	B1	B2
A1	760.6666	840.5555
A2	983.8888	1,003.7777

AC-MEANS	C1	C2	C3
A1	346.1666	1,004.5000	1,051.1666
A2	751.5000	1,096.5000	1,133.5000

BC-MEANS	C1	C2	C3
B1	455.5000	1,067.5000	1,093.8333
B2	642.1666	1,033.5000	1,090.8333

A-MEANS	
A1	800.6111
A2	993.8333

B-MEANS	C1
B1	872.2777
B2	922.1666

C-MEANS

C1	548.8333
C2	1,050.5000
C3	1,092.3333

GRAND MEAN = 897.2222

** END OF PROGRAM **

การแปลผล

จากตารางวิเคราะห์แวเหรียนซ์ พบว่า ปัจจัย A, C และ AxC มีนัยสำคัญแต่ B คือมูลวัวและ Interaction ระหว่างมูลวัวกับปัจจัยอื่นๆ ให้ค่า F ที่น้อยกว่า F ในตารางทั้งสิ้น แสดงว่าในการทดลองครั้งนี้ยังไม่มีหลักฐานเพียงพอที่จะกล่าวได้ว่า การใช้มูลวัวมีอิทธิพลต่อผลผลิตของข้าวโพดฝักอ่อนหรือไม่ แต่การสับกลบต้นข้าวโพดลงในแปลงและการใช้ปุ๋ยไนโตรเจนจะทำให้ผลผลิตของข้าวโพดฝักอ่อนเปลี่ยนแปลง สำหรับกรณีนี้จะต้องดูค่าเฉลี่ยที่ AC-MEANS พร้อมทั้งใช้ค่า LSD ที่โปรแกรมคำนวณไว้ให้เป็นตัวเปรียบเทียบดังนี้

ตาราง AC-MEANS

	C1	C2	C3
A1	346.2	1,004.5	1,051.2
A2	751.5	1,096.5	1,131.5
LSD FOR			
AC	168.9	229.6	

ถ้าเปรียบเทียบผลผลิตของข้าวโพดเมื่อไม่ใช้ปุ๋ยไนโตรเจนเลย (C_1) พบว่าการสับกลบต้น (A_2) จะให้ผลผลิตสูงกว่าการตัดต้นทิ้ง (A_1) (751.5-346.2 = 405.3**) แต่ถ้ามีการใช้ปุ๋ยไนโตรเจนไม่ว่าจะเป็น 10 หรือ 20 กก./ไร่ ร่วมกับการใช้ต้นสับกลบไม่ทำให้ผลผลิตข้าวโพดเปลี่ยนแปลง (1,096.5-1,004.5 = 92^{ns} หรือ 1,131.5-1,051.2 = 80.3^{ns})

3-factor factorial combination นี้อาจจะนำไปจัดเป็นแผนการทดลองแบบต่างๆ ได้อีกตามวัตถุประสงค์ของผู้ทดลอง เช่น

1. Split – Split Plot Design ถ้าผู้ทดลองสนใจปัจจัยทั้ง 3 ไม่เท่ากัน หรือในทางปฏิบัติบางปัจจัยอาจจะต้องการขนาดแปลงย่อยที่ใหญ่กว่าปัจจัยอื่นๆ

ถ้า A เป็นปัจจัยที่ผู้ทดลองสนใจน้อยที่สุด

B เป็นปัจจัยที่สนใจมากกว่า A

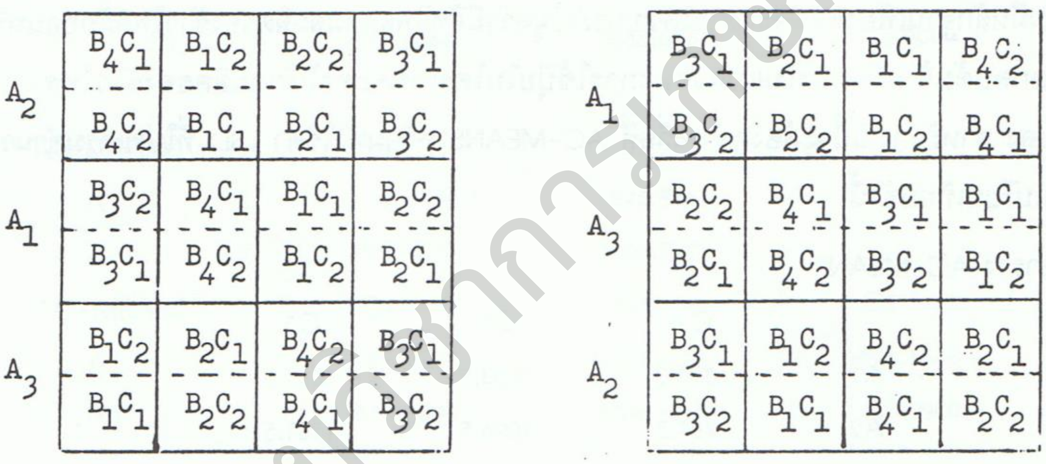
C เป็นปัจจัยที่สนใจและอยากจะทราบรายละเอียดที่สุด เช่นนี้ก็ควรจัด 3 factor factorial ให้อยู่ในรูปของ Split-Split plot Design โดยให้

A เป็น Main plot (หน่วยทดลองใหญ่)

B เป็น Sub plot (หน่วยทดลองกลาง)

C เป็น Sub sub plot (หน่วยทดลองเล็ก)

A อาจจะจัดในรูปของ Design พื้นฐานแบบใดก็ได้ ปกติทางเกษตรมักจะใช้ RCB ถ้า A B และ C มี 3, 4 และ 2 ระดับตามลำดับ แบบแผนการทดลองจะเป็นดังนี้



Block 1 Block 2

ถ้าทำการทดลอง 4 ซ้ำ ตารางวิเคราะห์เป็นดังนี้

SOV	df
Total	(3x4x2x4) - 1 = 95
Block	4 – 1 = 3
A	3 – 1 = 2
E (a)	(4 – 1)(3 – 1) = 6
B	4 – 1 = 3
AB	(3 - 1)(4 - 1) = 6
E (b)	3 (4 – 1)(4 – 1) = 27
C	2 – 1 = 1
AC	(3 – 1)(2 – 1) = 2
BC	(4 – 1)(2 – 1) = 3
ABC	(3 – 1)(4 – 1)(2 – 1) = 6
E (c)	3x4 (4-1)(2-1) = 36

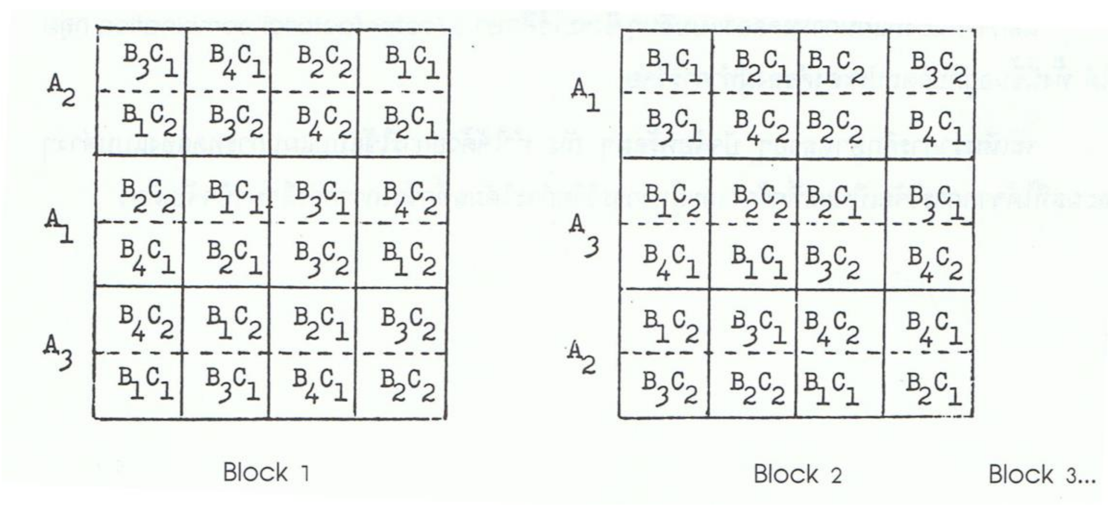
C.V. (a)

C.V. (b)

C.V. (c)

2. Split plot ที่มีปัจจัยเดี่ยวเป็น main plot และ 2 factor factorial เป็น Sub plot หรือกลับกันแล้วแต่วัตถุประสงค์ของผู้ทำการทดลอง เช่น

ถ้าผู้ทดลองสนใจ B และ C เท่าๆ กัน แต่สนใจ A น้อยที่สุด ก็อาจจะวางแผนการทดลองแบบ Split Plot โดยมี A เป็น main plot และ B x C เป็น Sub plot ดังนี้



ตารางวิเคราะห์

SOV	df
Total	95
Block	3
A	2
E (a)	6
(Sub plot 7)	
B	3
C	1
BC	3
(Interaction 14)	
AB	6
AC	2
ABC	6
E (b)	63

C.V. (a)

C.V. (b)

นอกจากนี้ยังมีแผนการทดลองแบบอื่นๆ ซึ่งจะใช้ศึกษา 3 factor factorial combination กลุ่มนี้ได้ ทั้งนี้ขึ้นอยู่กับวัตถุประสงค์ของผู้ทำการวิจัย

จะเห็นว่าการศึกษาหลายๆ ปัจจัยพร้อมๆ กัน ทำได้ด้วยการใช้แบบแผนการทดลองแบบต่างๆและผลที่ได้จากการวิจัยเป็นผลที่กว้างและผู้ทำการวิจัยก็จะได้ผลที่กว้างกว่าการศึกษาปัจจัยเดียว

เอกสารประกอบการเขียน

Anderson, R.L. and T.A. Bancroft (1952) Statistical theory in Research. Mc Graw-Hill Book Co, Inc. New york. 399pp.

Cochran, W.G. and G.M. Cox.(1957) Experimental Design (2nd E.d.)John Wiley and Sons, Inc. New york. 611pp.

Federer., W.T. (1955) Experimental Design. The Macmillan Company, New york. 543pp.

Gomez., K.A. and A.A. Gomez. (1984) Statistical Procedures for Agricultural Research. (2nd Ed.) John Wiley and Sons, Inc. USA. 680pp.

Snedecor, G.W. and W.G. Cochran (1969) Statistical Method (6th Ed. 3rd Printing). The Iowa State University Press Ames, Iowa. 592pp.

เอกสารประกอบคำบรรยายการฝึกอบรมสถิติ หลักสูตร การวางแผนงานทดลอง (2527) ดำเนินงานโดยฝ่ายวิเคราะห์ทางสถิติ กองแผนงานและวิชาการ การวางแผนงานทดลองสำหรับหลายปัจจัย (หน้า 55 - 139)

สหสัมพันธ์และรีเกรสชั่น

วิจิตรา พลเยี่ยม

สหสัมพันธ์ และรีเกรสชัน

(Correlation and Regression)

นอกจากการวิเคราะห์ และทดสอบสมมติฐานของแต่ละตัวแปรแล้ว บางครั้งเราอาจสนใจตัวแปร (Variables) ที่มีอยู่นั้นว่ามีความสัมพันธ์ (relationship) กันอย่างไร และจะช่วยตัดสินใจในการคาดคะเน (predict) และวางแผนงานได้อย่างไร จะเห็นว่าเราต้องการ หลักการบางอย่างมาช่วยวิเคราะห์เกี่ยวกับความสัมพันธ์ระหว่างข้อมูล 2 ชุดหรือมากกว่า วิธีการทางสถิติที่สามารถนำมาแก้ไขปัญหานี้ก็คือ Regression and Correlation Analysis

Regression Analysis

Regression เป็นวิธีการหนึ่งที่จะใช้ช่วยตรวจหาลักษณะ (nature) ของความสัมพันธ์ระหว่างตัวแปรตั้งแต่สองตัวขึ้นไป ตัวแปรเหล่านี้แยกออกได้เป็น 2 ประเภท คือ **ตัวแปรอิสระ** (independent variable) ซึ่งสามารถกำหนดค่าได้ตามที่ต้องการ เช่น ปริมาณปุ๋ยที่ใส่ต่อ 1 ไร่ของข้าวโพด หรือ ตัวแปรซึ่งมีค่าตามที่สังเกตได้โดยที่ไม่ถูกควบคุม เช่น ความชื้น ผลที่เกิดจากการเปลี่ยนแปลงของตัวแปรอิสระนี้จะกระทบกระเทือนตัวแปรอีกตัวหนึ่ง ซึ่งเรียกว่า**ตัวแปรตาม** (dependent variable) ตัวอย่าง เช่น ปริมาณปุ๋ยที่ใส่ต่อไร่ อาจจะมีผลทำให้ผลผลิตข้าวโพดเพิ่มขึ้นหรือลดลง ผลผลิตข้าวโพดจะถือว่าเป็นตัวแปรตามในกรณีนี้

ถ้าพิจารณาแต่เพียงปัจจัยเดียว คือ ตัวแปรอิสระเพียงตัวเดียวที่มีผลต่อตัวแปรตามแล้ว เราเรียก Regression นั่นว่า "**Simple Regression**" เช่น ปริมาณปุ๋ยต่อไร่ (X) ซึ่งมีผลต่อผลผลิตข้าวโพด (Y) ตามความเป็นจริงแล้วผลผลิตข้าวโพดมิได้ขึ้นอยู่กับปุ๋ย (X_1) เพียงอย่างเดียว ยังมีอิทธิพลของตัวแปรอิสระอื่นๆ เช่น พันธุ์ (X_2) การดูแลรักษา (X_3) ฯลฯ เข้ามาเกี่ยวข้อง และเมื่อเราพิจารณาอิทธิพลของปัจจัยมากกว่าหนึ่งตัว Regression นั้นเรียกว่า "Multiple Regression"

ความสัมพันธ์ระหว่าง X และ Y อาจเป็นรูปเส้นตรง (linear) หรือไม่ใช่เส้นตรง (non linear) ก็ได้ ดังนั้นรูปแบบของ Regression Analysis แบ่งออกได้เป็น 4 แบบ ดังนี้

1. Simple Linear Regression
2. Multiple Linear Regression
3. Simple NonLinear Regression
4. Multiple NonLinear Regression

Simple Linear Regression เกี่ยวข้องกับการวิเคราะห์ความสัมพันธ์ของตัวแปร 2 ชุด คือ ตัวแปรอิสระ X และตัวแปรตาม Y โดยเราจะศึกษาว่า ถ้า X เปลี่ยนไป 1 หน่วย จะทำให้ค่า Y เปลี่ยนไปกี่หน่วย

ในการวิเคราะห์ Simple Linear Regression สมการแสดงรูปแบบความสัมพันธ์ ระหว่าง X และ Y ของประชากร (population) คือ

$$Y \quad = \quad \alpha + \beta x + \varepsilon$$

เมื่อ α = intercept หรือ จุดตัดของเส้นตรงบนแกน Y

β = Slope หรือ ความชันของเส้น

ε = ความคลาดเคลื่อนสุ่ม

ประมาณค่า α ด้วย a β ด้วย b และ ε ด้วย e แล้ว จะได้รูปสมการเป็น

$$Y \quad = \quad a + bX + e$$

ให้ y – e เป็น $\hat{Y}$ หมายถึงค่าประมาณของ Y แล้ว

$$\hat{Y}_L \quad = \quad a + bX$$

ซึ่งก็คือสมการของ Simple linear Regression

ถ้ามีข้อมูลจำนวน n คู่ คือ $(X_1, Y_1), (X_2, Y_2), \ldots\ldots, (X_n, Y_n)$ โดย X เป็นตัวแปรอิสระ และ Y เป็นตัวแปรตาม และ X กับ Y มีความสัมพันธ์กันในรูปเส้นตรง (Y = a+bX) แล้ว โดยวิธี least square (วิธีที่ถือหลักว่าผลรวมของกำลังสองของความคลาดเคลื่อนน้อยที่สุดหรือ $\Sigma\varepsilon^2$ มีค่าน้อยที่สุด) จะคำนวณ a และ b ได้ดังนี้

$$b \quad = \quad \Sigma xy \,/\, \Sigma x^2$$

$$a \quad = \quad \bar{Y} - b\bar{X}$$

เมื่อ

$$\bar{X} \quad = \quad \sum_{i=1}^{n} X_i/n$$

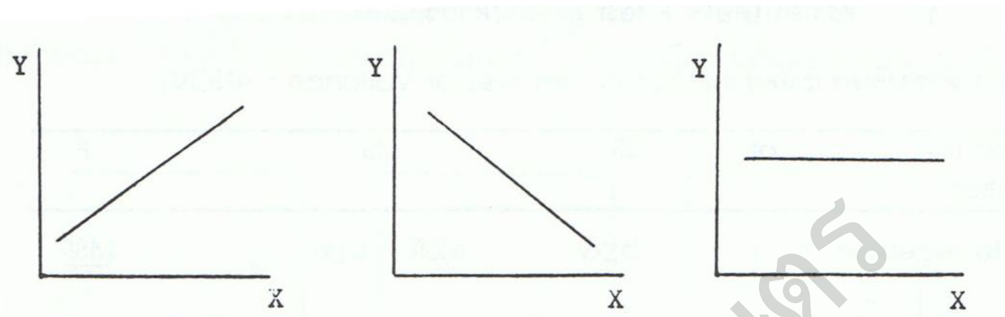
$$\bar{Y} \quad = \quad \sum_{i=1}^{n} Y_i/n$$

$$\Sigma x^2 \quad = \quad \sum_{i=1}^{n} X_i^2 \quad (\sum_{i=1}^{n} X_i)^2/n$$

$$\Sigma y^2 \quad = \quad \sum_{i=1}^{n} Y_i^2 \quad (\sum_{i=1}^{n} Y_i)^2/n$$

$$\Sigma xy \quad = \quad \sum_{i=1}^{n} X_i Y_i \quad (\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i)/n$$

ค่า a เป็นค่า Y – Intercept และ b เป็นค่า Slope ของเส้น regression ค่า b นี้เรียกว่า **“Coefficient of Regression”** ซึ่งอาจจะเป็นทั้งบวก ลบ หรือศูนย์ ถ้า b มีค่าเป็นบวก แสดงว่าค่าของ X มีผลทำให้ Y มีค่าเปลี่ยนแปลงไปในทางเดียวกัน กล่าวคือ ถ้า X เพิ่ม Y ก็เพิ่มตาม (รูปที่ 1) ถ้า b มีค่าเป็นลบแสดงว่าค่าของ X มีผลทำให้ Y มีค่าเปลี่ยนแปลงไปในทางตรงข้าม กล่าวคือ ถ้า X มีค่าเพิ่ม Y จะมีค่าลด (รูปที่ 2) แต่ถ้า b เท่ากับศูนย์ แสดงว่า X จะไม่มีผลต่อ Y เลย เส้น regression จะเป็นดังรูปที่ 3



รูปที่ 1 รูปที่ 2 รูปที่ 3

จากเส้น regression ; $\hat{Y} = a+bX$ เมื่อต้องการประมาณค่า $Y(\hat{Y})$ จะทำได้โดยแทนค่า X ลงในสมการ และจากสมการนี้จะสรุปได้ว่า เมื่อ X เปลี่ยนแปลงไป 1 หน่วย Y จะเปลี่ยนแปลงไป b หน่วย

ค่าความแปรปรวน จาก regression ($S^2_{y.x}$) คำนวณได้โดย

$$S^2_{y.x} = \frac{\Sigma y^2 - \frac{(\Sigma xy)^2}{\Sigma x^2}}{n-2}$$

ความแปรปรวนของ b คือ

$$S^2_b = \frac{S^2_{y.x}}{\Sigma x^2}$$

และความแปรปรวนของค่าประมาณของ $Y(\hat{Y})$ คือ

$$V(\hat{Y}) = S^2_{y-x}\left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{\Sigma x^2}\right)$$

การทดสอบสมมติฐาน เพื่อทดสอบความสัมพันธ์เชิงเส้นตรงระหว่าง X กับ Y โดยสมมติฐานเป็นดังนี้

Ho : $\beta = 0$ (X และ Y ไม่มีความสัมพันธ์เชิงเส้นตรงต่อกัน)

Ha : $\beta \neq 0$ (X และ Y มีความสัมพันธ์เชิงเส้นตรงต่อกันจริง)

ตัวสถิติที่ใช้ในการทดสอบจะเป็น t หรือ F ก็ได้

ก. **ทดสอบโดยใช้** t-test ค่าสถิติ t คำนวณได้โดย

$$t \quad = \quad \frac{b - \beta o}{S_b}$$

เปรียบเทียบค่า t ที่คำนวณได้กับค่า t จากตารางที่ (n-2) degree of freedom และ $\alpha/2$

เมื่อ α คือ ระดับนัยสำคัญที่กำหนด

ข. **ทดสอบโดยใช้** F-test สูตรการคำนวณแสดงในตารางที่ 1

<u>ตารางที่ 1</u> ตารางวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOV)

Source of Variation	df	SS	MS	F
Due to regression	1	$b\Sigma xy$	$MSR = b\Sigma xy$	$\frac{MSR}{S^2_{y.x}}$
residual	n-2	$\Sigma y^2 - b\Sigma xy$	$S^2_{y.x} = \frac{\Sigma y^2 - b\Sigma xy}{n-2}$	
Total	n-1	Σy^2		

เปรียบเทียบค่า F ที่คำนวณได้กับค่า F จากตารางที่ (1, n-2) d.f. และระดับนัยสำคัญที่กำหนด

ช่วงเชื่อมั่นสำหรับค่าประมาณ

ก. สำหรับค่าประมาณของ Y($\hat{Y}$)

$$(1 - \alpha)\,\%CI \quad = \quad \hat{Y} \pm t_{\alpha/2(n-2)} \sqrt{S^2_{y.x}\left(\frac{1}{n} + \frac{(x_i + \bar{x}^2)}{\Sigma x^2}\right)}$$

ข. สำหรับค่าประมาณของ β (b)

$$(1 - \alpha)\,\%CI \quad = \quad b \pm t_{\alpha/2(n\ \ 2)} S_b$$

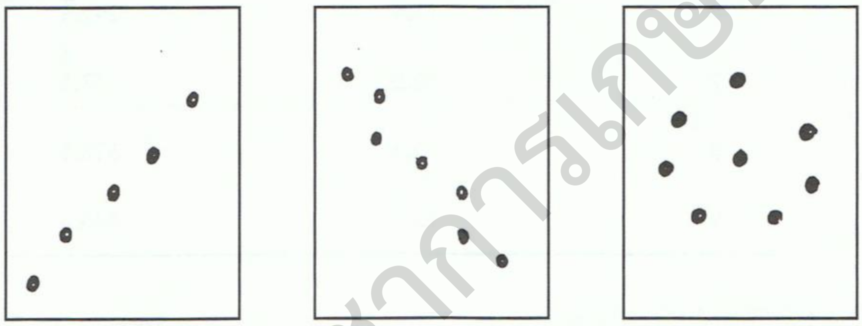
Correlation Analysis

Correlation เป็นวิธีการวัดขนาดความสัมพันธ์ระหว่างตัวแปรต่างๆ โดยจะถือตัวแปรใดเป็นตัวแปรอิสระหรือตัวแปรตามก็ได้ Correlation ของตัวแปรมีหลายแบบอาจจะเป็นแบบเส้นตรงหรือเส้นโค้งก็ได้

Simple Linear Correlation เป็นการวัดขนาดความสัมพันธ์ระหว่างตัวแปร 2 ตัว เมื่อตัวแปร 2 ตัวนี้มีความสัมพันธ์กันแบบเส้นตรง ขนาดของความสัมพันธ์วัดได้จากค่า **สัมประสิทธิ์สหสัมพันธ์** (Correlation coefficient) ซึ่งใช้สัญลักษณ์ ρ แทน Population Correlation Coefficient และ r แทน sample correlation coefficient ค่า r นี้จะคำนวณได้โดย

$$r = \frac{\Sigma xy}{\sqrt{(\Sigma x^2)(\Sigma y^2)}}$$

r จะมีค่าอยู่ระหว่าง -1 และ +1 ($-1 \le r \le +1$) ถ้าหาก r มีค่า = +1 หรือเกือบเท่ากับ +1 หมายความว่าข้อมูลทั้งสอง (X,Y) มีความสัมพันธ์กันมากแสดงว่าถ้าข้อมูล X เปลี่ยนไป ข้อมูล Y จะเปลี่ยนตามไปในทางเดียวกัน กล่าวคือ X เพิ่ม Y เพิ่ม X ลด Y ก็ลดตาม (รูปที่ 4) แต่ถ้า r มีค่าเท่ากับ -1 หรือเกือบเท่ากับ -1 ข้อมูล X และ Y มีความสัมพันธ์กันมากแต่เป็นไปในทางด้านตรงข้ามกัน กล่าวคือ ถ้า Variable หนึ่งเพิ่มจะเป็นผลให้อีก Variable หนึ่งลด (รูปที่ 5) และถ้าหาก r มีค่า = 0 หรือเกือบเท่ากับศูนย์ แสดงให้เห็นว่าข้อมูล 2 ชุดนั้นไม่มีความสัมพันธ์กันเลยหรือเกือบไม่มีความสัมพันธ์กันเลย (รูปที่ 6)



รูปที่ 4 รูปที่ 5 รูปที่ 6

การทดสอบว่า X และ Y มีความสัมพันธ์กันหรือไม่ (Ho : $\rho = 0$ VS Ha : $\rho \neq 0$) ทำได้โดยคำนวณค่า r แล้วเทียบค่าที่คำนวณได้กับค่าจากตาราง r ที่ (n-2) degree of freedom และระดับนัยสำคัญที่กำหนดโดยที่ n คือ จำนวนคู่ของค่าสังเกต

Coefficient of Determination

คือ กำลังสอง Coefficient of Correlation ค่า Coefficient of Determination แบบ simple linear ใช้สัญลักษณ์ r^2 แสดงให้เห็นว่าสมการ regression ที่ได้นั้นเหมาะสมกับข้อมูลเพียงใด

<u>ตัวอย่างที่ 1</u> ข้อมูลผลผลิตและจำนวนรวงต่อกอของข้าวนาสวน ปี 2531 จำนวน 9 แปลง เป็นดังนี้

แปลงที่	จำนวนรวงต่อกอ (X)	ผลผลิตข้าว (กก./ไร่) Y
1	10.0	287.8
2	9.3	286.8
3	13.0	358.0
4	10.5	278.3
5	11.8	344.8
6	11.5	292.5
7	12.0	317.5
8	10.8	313.5
9	15.5	440.3

$n = 9$

$\Sigma X = 104.4 \qquad \bar{X} = \frac{\Sigma X}{n} = \frac{104.4}{9} = 11.6$

$\Sigma Y = 2{,}919.5 \qquad \bar{Y} = \frac{\Sigma Y}{n} = \frac{2{,}919.5}{9} = 324.4$

$\Sigma x^2 = 1{,}238.12 \qquad \Sigma Y^2 = 968{,}093.85 \qquad \Sigma XY = 34{,}574.23$

$$\Sigma x^2 = \Sigma X^2 - \frac{(\Sigma X)^2}{n} = 1{,}238.12 - \frac{(104.4)^2}{9} = 27.0800$$

$$\Sigma y^2 = \Sigma Y^2 - \frac{(\Sigma Y)^2}{n} = 968{,}093.85 - \frac{(2{,}919.5)^2}{9} = 21{,}040.2875$$

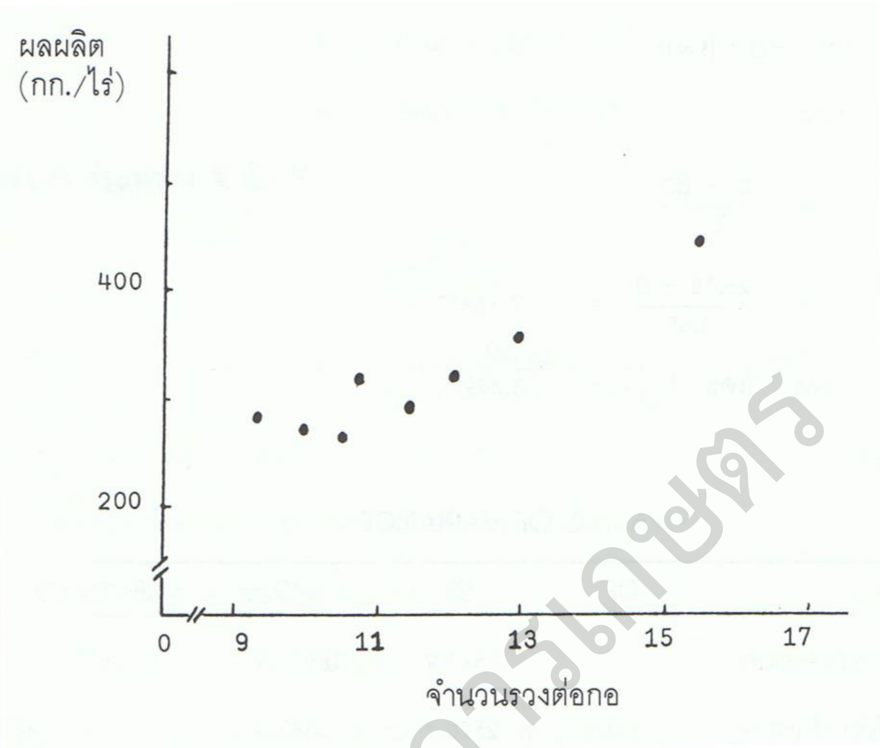
$$\Sigma xy = \Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{n} = 34{,}574.23 - \frac{(104.4)(2{,}919.5)}{9} = 708.0269$$

$$S_{y.x}^2 = \frac{\Sigma Y^2 \quad {(\Sigma xy)^2}/{\Sigma x^2}}{n-2} = \frac{21{,}040.2875 - \frac{(708.0269)^2}{27.0800}}{9-2} = 361.1991$$

$$S_b^2 = \frac{S_{Y.X}^2}{\Sigma X^2} = \frac{361.1991}{27.0800} = 13.3382$$

$$S_b = \sqrt{S_b^2} = \sqrt{13.3382} = 3.6521$$

<u>รูปที่ 7</u> Scatter diagram ระหว่างผลผลิตกับจำนวนรวงต่อกอ



จาก Scatter diagram จะเห็นว่ารูปแบบความสัมพันธ์ระหว่างจำนวนรวงต่อกอ (X) และผลผลิต (Y) มีแนวโน้มจะเป็นเส้นตรงซึ่งมีรูปแบบสมการ

$$\hat{Y} = a + bX$$

คำนวณค่า a และ b

$$b = \frac{\sum xy}{\sum x^2} = \frac{708.0269}{27.0800} = 26.1457$$

$$a = \bar{Y} - b\bar{X} = 324.4 - (26.15)(11.6) = 21.06$$

สมการเส้นตรงที่ประมาณได้ คือ

$$\hat{Y} = 21.06 + 26.15X$$

<u>ทดสอบสมมติฐาน</u>เกี่ยวกับพารามิเตอร์

Ho : β = 0 VS Ha : $\beta \neq 0$

ก. ใช้ t – test

$$t = \frac{b - \beta o}{S_b}$$

$$t = \frac{26.15 - 0}{3.65} = 7.164^{**}$$

$t_{.025,7}$ = 2.365 และ $t_{.005,7}$ = 3,499

ข. ใช้ F - test

ANALYSIS OF VARIANCE

SOV	DF	SS	MS	F
Regression	1	18,511.9	18,511.9	51.25**
Residual	7	2,528.4	361.2	
Total	8	21,040.3		

$F_{.05,(1,7)}$ = 5.59 และ $F_{.01,(1,7)}$ = 12.25

นั่นคือ ปฏิเสธสมมติฐานที่ว่า β = 0 แสดงว่าจำนวนรวงเพิ่มขึ้นจะเป็นผลทำให้ได้ผลผลิตข้าวเพิ่มขึ้นในรูปเส้นตรง โดยที่เมื่อจำนวนรวงต่อกอเพิ่มขึ้น 1 รวงจะเป็นผลทำให้ผลผลิตข้าวเพิ่มขึ้น ประมาณ 26 กิโลกรมต่อไร่

ช่วงความเชื่อมั่น (Confidence interval) ของ β

$$95\%\ C.I. = b \pm t_{\alpha/2_{(n-2)}}.S_b$$

$$= 26.15 \pm (2.365)(3.6521)$$

$$= 26.15 \pm 8.64$$

$$= (17.51, 34.79)$$

แสดงว่าที่ความเชื่อมั่น 95% เมื่อจำนวนรวงต่อกอเพิ่มขึ้น 1 รวง ปริมาณผลผลิตข้าวจะเพิ่มขึ้นระหว่าง 17.51 ถึง 34.79 กิโลกรัมต่อไร่

การประมาณค่า Y ถ้าจำนวนรวงต่อกอ (X) เป็น 12 รวง แล้วน่าจะได้ผลผลิตข้าว

$$\hat{Y} = 21.06 + (26.15)(12)$$

$$= 334.9 \text{ กิโลกรัมต่อไร่}$$

ความสัมพันธ์ระหว่าง X กับ Y

$$r = \frac{\Sigma xy}{\sqrt{(\Sigma x^2)(\Sigma y^2)}}$$

$$= \frac{708.0269}{\sqrt{(27.0800)(21,040.2875)}} = 0.938^{**}$$

$r_{.05,7}$ = .666 และ $r_{.01,7}$ = .798

แสดงว่า จำนวนรวงมีความสัมพันธ์กับผลผลิตข้าว

Coefficient of Determination

$$r^2 = (0.938)^2 = 0.88$$

แสดงว่า เส้น regression Y = 21.06+26.15x สามารถอธิบาย Variation ที่เกิดขึ้นทั้งหมดใน Y ได้ 88% หรืออีกนัยหนึ่งคือ 88% ของความแปรปรวนทั้งหมด อธิบายได้ด้วยเส้น regression นี้