เอกสารการสอน

วิชา 734462 หลักการวางแผนการทดลอง

รศ.อัจฉริยา ปราบอริพ่าย

คณะศิลปศาสตร์และวิทยาศาสตร์

มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน

ภาคต้น ปีการศึกษา 2551

บทที่ 1
แนวคิดพื้นฐานในการวางแผนการทดลอง

การวิจัยเป็นการเสาะหาคำตอบสำหรับปัญหาที่ยังไม่มีข้อสรุป การออกแบบการวิจัยคือการวางแผนและโครงของการสืบค้นเพื่อตอบคำถามของปัญหาของการวิจัย การวางแผนคือรายการทั้งหมดที่ต้องดำเนินการของการวิจัย รวมทั้งกรอบที่ผู้วิจัยจะทำจากสมมติฐานของการวิจัยที่ตั้งขึ้น วิธีที่จะอธิบายกรอบของงานวิจัยได้ดีที่สุดคือ การเขียนเป็นสมการเชิงคณิตศาสตร์เป็นตัวแบบของความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ที่ทำการศึกษา และการแปลความหมายจากการวิเคราะห์ข้อมูลในขั้นสุดท้าย เพื่อให้ได้หลักฐานเชิงประจักษ์ของความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ที่ทำการศึกษา การวางแผนการวิจัยที่เหมาะสมและรัดกุมจะช่วยให้ผลการวิจัยที่ได้น่าเชื่อถือ

ขั้นตอนแรกของการวางแผนการวิจัย คือ การเขียนโครงร่างงานวิจัย ซึ่งประกอบด้วย 3 ส่วน หลัก ได้แก่

1. วัตถุประสงค์ของงานวิจัย อาจอยู่ในรูปของประโยคคำถามที่ต้องการหาคำตอบ สมมติฐานของการทดสอบ หรือผลที่คาดว่าจะเกิดขึ้น ซึ่งประโยคนั้นควรจะชัดเจนแจ่มแจ้งและเฉพาะเจาะจง ในประโยคควรจะบอกพื้นที่ที่ครอบคลุมที่อ้างอิงถึงหรือประชากรที่อ้างอิงถึงซึ่งผู้วิจัยจะต้องกำหนดประชากรที่จะนำผลการทดลองไปใช้ก่อนล่วงหน้าแล้ว และเพื่อควบคุมความแปรปรวนของความคลาดเคลื่อนของการทดลอง
2. รายละเอียดของกรอบของงานวิจัยในการศึกษาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ เหล่านั้น ซึ่งจะบอกว่าควรเก็บข้อมูลอะไร ทำอย่างไรจึงจะเก็บข้อมูลนั้นมาได้ที่ครอบคลุมเกี่ยวกับทรีทเมนต์ ขนาดของการทดลองและวัสดุอุปกรณ์ที่ใช้ในการทดลอง
3. วิธีการวิเคราะห์ข้อมูลที่เป็นตัวเลขเชิงปริมาณที่เป็นตัวแทนของค่าสังเกตนั้น

1. รายละเอียดของการทดลอง

การออกแบบการทดลองเกี่ยวข้องกับการตัดสินใจหลัก ๆ 2 ประเด็น ประเด็นแรกคือ การทดลองนั้นมีวัตถุประสงค์เพื่ออะไร ประเด็นที่สองคือ คำถามหรือข้อสงสัยที่ต้องการคำตอบจากผลการทดลองนี้คืออะไร แล้วพิจารณาว่าควรกำหนดทรีทเมนต์อะไรในการทดลองนั้น เราจะเรียกว่าโครงสร้างของทรีทเมนต์ และพิจารณาเกี่ยวกับอิทธิพลของสภาพแวดล้อม การทำการทดลองซ้ำ การสุ่มหน่วยทดลองให้กับทรีทเมนต์ หรือการสุ่มทรีทเมนต์ให้กับหน่วยทดลอง รวมถึงอาจต้องมีเทคนิคการบล็อกด้วย และเมื่อสิ้นสุดระยะเวลาการทดลองแล้วจึงดำเนินการเก็บข้อมูลเพื่อทำการวิเคราะห์ทางสถิติต่อไป

1.1 โครงสร้างของทรีทเมนต์

ทรีทเมนต์ หมายถึง สิ่งหรือวิธีการปฏิบัติที่ให้แก่หน่วยทดลองแล้ววัดผลการทดลอง หรืออิทธิพลของแต่ละวิธีการมาเปรียบเทียบกัน การเลือกทรีทเมนต์ต่าง ๆ เป็นสิ่งสำคัญในการทดลอง เพื่อให้บรรลุวัตถุประสงค์ของการทดลอง

หน่วยทดลอง หมายถึง หน่วยหรือกลุ่มของวัตถุทดลองที่ได้รับทรีทเมนต์อย่างเดียวกัน อาจเป็น คน สัตว์ ต้นไม้ หรือสิ่งของต่าง ๆ แล้ววัดผลของทรีทเมนต์ที่ได้รับนั้นเป็นข้อมูลหรือค่าสังเกต

โครงสร้างของทรีทเมนต์ ประกอบด้วย

1 คอนโทรล อาจมีอยู่ในการทดลองเพื่อเป็นฐานใช้เปรียบเทียบกับทรีทเมนต์อื่น ๆ ตัวอย่างเช่น ในการศึกษาเกี่ยวกับการเจริญเติบโตและผลผลิตของข้าวพันธ์ใหม่ ผักหรือผลไม้ เราอาจต้องหาพันธ์ที่สามารถใช้เป็นมาตรฐานซึ่งเราทราบข้อมูลอยู่แล้วเพื่อเป็นฐานในการสนับสนุนข้อมูลของพันธ์ใหม่ ซึ่งเป็นพันธ์ที่เกษตรกรรู้จักดี และต้องทราบว่าจะเปรียบเทียบอย่างไร และเป็นพันธ์ที่ต้องการชักชวนให้เปลี่ยนไปเป็นพันธ์ใหม่

2 การไม่ให้ทรีทเมนต์ใด ๆ ในการทดลองทางคลินิคมักเรียกว่า placebo เพื่อตรวจสอบว่าถ้าคนไข้ได้รับยาซึ่งไม่มีทรีทเมนต์ใด ๆ หรือเรียกว่า untreated control และสำหรับการทดลองชนิดอื่น ๆ ก็คือ ไม่ให้ทรีทเมนต์เป็นศูนย์ ในการทดลองทางคลินิก placebo จะเป็นยาที่รูปร่างลักษณะเหมือนยาจริงที่ใช้เป็นทรีทเมนต์ของการทดลอง แต่ทำด้วยวัตถุดิบที่ไม่มีปฏิกิริยา

ใด ๆ ตัวอย่างเช่น การทดลองเกี่ยวกับยาชนิดใหม่ที่ใช้รักษาอาการปวดศีรษะ ซึ่งบางครั้งอาจมีความสับสนเกี่ยวกับคนไข้ว่ามีอาการปวดหรือความเป็นจริงอาจเป็นอาการเงื่องหงอย

3 ทรีทเมนต์ที่มาจากการคอมบิเนชันของระดับของแฟกเตอร์หลาย ๆ แฟกเตอร์ ซึ่งแฟกเตอร์ต่าง ๆ เหล่านี้อาจเป็นอิสระหรือไม่เป็นอิสระกัน หมายถึง อาจไม่มีอิทธิพลร่วมกันหรือมีอิทธิพลร่วมกันก็ได้

4 ทรีทเมนต์ที่มาจากการเพิ่มปริมาณของแฟกเตอร์ตัวหนึ่ง หรือหลายแฟกเตอร์ เพื่อเปรียบเทียบผลการทดลองระหว่างปริมาณแฟกเตอร์ต่าง ๆ นั้น และอาจใช้การพล็อตกราฟข้อมูลจากผลการทดลองที่ระดับปริมาณต่าง ๆ ของแฟกเตอร์หนึ่ง ซึ่งเส้นโค้งนั้นอาจสามารถแสดงรูปแบบหนึ่งของผลที่ได้จากการทดลองนั้น

5 ทรีทเมนต์ที่เป็นกลุ่มต่าง ๆ ที่แตกต่างกัน และเราต้องการเปรียบเทียบระหว่างกลุ่มต่าง ๆ ที่แตกต่างกันนั้น เราสามารถใช้วิธีการของคอนทรัสต์ในการเปรียบเทียบนี้ได้

การเลือกทรีทเมนต์ต่าง ๆ เป็นสิ่งสำคัญในการทำการทดลอง เพื่อให้บรรลุวัตถุประสงค์ของการทดลอง มีประเด็นปัญหาที่เกิดขึ้นในการเลือกทรีทเมนต์คือ

1. การสรุปผลการทดลองตามวัตถุประสงค์การทดลองที่ตั้งไว้ว่าระหว่างทรีทเมนต์ต่าง ๆ นั้น มีอิทธิพลแตกต่างกัน ประเด็นปัญหาคือสรุปผลการทดลองให้ชัดเจนไม่ได้ว่าเนื่องมาจากอิทธิพลของทรีทเมนต์ใด หรือเนื่องมาจากอิทธิพลร่วมของทรีทเมนต์ต่าง ๆ นั้น

2. การกำหนดว่าจะเปรียบเทียบทรีทเมนต์ใดกับทรีทเมนต์ใดเป็นคำถามที่ยาก แต่สามารถบอกได้จากวัตถุประสงค์ของการทดลอง และประชากรที่อ้างอิงถึง

3. บางทรีทเมนต์ที่นำมาทดสอบ อาจเป็นสิ่งที่ไม่ชื่นชอบทำให้ความรู้สึกของคนเข้าไปเกี่ยวข้องกับผลการทดลองด้วย ยกตัวอย่างเช่น การเปรียบเทียบวิธีการสอน 2 วิธีคือ วิธี A และวิธี B ครูบางคนชอบวิธี A ส่วนครูบางคนชอบวิธี B ถ้าแบ่งครูอย่างสุ่มออกเป็น 2 กลุ่ม ให้กลุ่มหนึ่งสอนวิธี A และอีกกลุ่มหนึ่งสอนวิธี B จะมีครูบางคนที่สอนวิธีที่ไม่ชอบ ซึ่งจะมีผลต่อการทดลองทำให้ผลการทดลองคลาดเคลื่อนไป แต่ถ้าไม่ใช้วิธีการสุ่มครู กำหนดให้ครูที่ชอบวิธี A สอนวิธี A และครูที่ชอบวิธี B สอนวิธี B แล้วให้สลับกลุ่มกันให้ครูที่ชอบวิธี A สอนวิธี B และครูที่ชอบวิธี B สอนวิธี A และให้กลุ่มครูที่เป็นกลางอีก 2 กลุ่ม กลุ่มหนึ่งสอนวิธี A อีกกลุ่มหนึ่งสอนวิธี B จะทำให้เกิดทรีทเมนต์ทั้งหมดเท่ากับ 6 ทรีทเมนต์ แทนที่จะเป็น 2 ทรีทเมนต์ ซึ่งในทางปฏิบัติการวางแผนที่ละเอียดละออเช่นนี้ทำได้ยาก เพราะต้องอาศัยทรัพยากรจำนวนมาก

4. การกำหนดให้มีคอนโทรล (control) ในการทดลองสำหรับเปรียบเทียบกับทรีทเมนต์อื่น ๆ ว่าทรีทเมนต์ต่าง ๆ มีอิทธิพลต่อผลการทดลองไม่แตกต่างกัน ยกตัวอย่างเช่น การเปรียบเทียบปุ๋ยไนโตรเจน 3 ชนิด คอนโทรลก็คือทรีทเมนต์ที่ไม่มีไนโตรเจน การวางแผนทรีทเมนต์อาจทำได้ 3 ทาง คือ

1. ทำการเปรียบเทียบปุ๋ยไนโตรเจน 3 ชนิด โดยไม่มีคอนโทรล

2. กรณีที่ดินมีความอุดมสมบูรณ์อยู่แล้ว น่าจะเพิ่มคอนโทรล เพื่อใช้เปรียบเทียบกับปุ๋ยไนโตรเจนทั้ง 3 ชนิด

3. กรณีที่มีการทำการทดลองซ้ำ และคอนโทรลถูกนำมาเปรียบเทียบกับทรีทเมนต์อื่น ๆ ในแต่ละซ้ำ ทำให้คอนโทรลมีจำนวนซ้ำมากกว่าทรีทเมนต์อื่น ๆ ในการทดลอง เช่น การเปรียบเทียบคุณภาพปุ๋ย 4 ชนิด แต่ละชนิดเปรียบเทียบกับคอนโทรลทำให้แต่ละซ้ำมีคอนโทรล 4 ซ้ำ ดังนั้น ถ้าทำการทดลองที่มี 4 ซ้ำ จะทำให้มีคอนโทรลทั้งหมด 16 ซ้ำ

1.2 อิทธิพลของสภาพแวดล้อม

การเปลี่ยนสภาพแวดล้อมในทางชีววิทยา การเกษตร หรือแม้แต่ในโรงงานที่มีการควบคุมสภาพแวดล้อมต่าง ๆ การให้ทรีทเมนต์แก่หน่วยทดลองที่มีลักษณะต่าง ๆ เหมือนกัน แต่ทำการทดลองในเวลาที่แตกต่างกัน ส่วนใหญ่มักจะได้ผลการทดลองแตกต่างกัน หรือในทางการเกษตรที่ทำการทดลองในพื้นที่ที่อยู่ในภูมิภาคแตกต่างกัน หรือการทดลองใช้ยาในโรงพยาบาลต่าง ๆ ที่แตกต่างกัน หรือในการทดลองทางอุตสาหกรรมที่ทำการทดลองในบริษัทต่าง ๆ ที่แตกต่างกัน แต่ในการวิจัยเราสามารถจัดสภาพการณ์ได้โดยทำการทดลองในโรงพยาบาลเดียวกันในเวลาที่แตกต่างกัน และทำการทดลองในบริษัทเดียวกันเป็นการทำการทดลองซ้ำ หรือทำการทดลองทรีทเมนต์เดียวกันมากกว่า 1 การทดลอง เหตุผลข้อหนึ่งที่ทำเช่นนี้เพราะว่าเราไม่สามารถหาหน่วยทดลองที่เหมือนกันจริง ๆ ได้สำหรับการทดลองต่าง ๆ โดยเฉพาะอย่างยิ่งที่เวลาแตกต่างกัน หรือสภาพแวดล้อมอื่น ๆ เช่น อุณหภูมิ ความชื้น ปริมาณแสงสว่าง หรือความเร็วของเครื่องจักรอาจแตกต่างกันได้ ความแตกต่างเพียงเล็กน้อยนี้อาจมีอิทธิพลต่อการทดลองได้

ดังนั้นค่าเฉลี่ยของทรีทเมนต์ต่าง ๆ ที่แตกต่างกันอาจมีอิทธิพลเนื่องจากการทดลองในเวลาที่แตกต่างกัน และสถานที่แตกต่างกัน

ผลการทดลองที่ได้นั้นนอกจากเกิดจากอิทธิพลของทรีทเมนต์ที่สนใจแล้วยังได้รับอิทธิพลจากปัจจัยภายนอกอื่น ๆ ด้วย ทำให้ผลการทดลองเบี่ยงเบนไปจากความเป็นจริง ซึ่งเรียกว่าความ

คลาดเคลื่อนของการทดลอง (experimental errors) แหล่งของความคลาดเคลื่อนของการทดลองมี 2 แหล่งใหญ่ ๆ คือ

1. จากหน่วยทดลองที่ได้รับทรีทเมนต์ เช่น แปลงนา คนไข้ในโรงพยาบาล กลุ่มของหมูในคอก หรือถาดเมล็ดพันธุ์ หน่วยทดลองเหล่านี้แม้ได้รับทรีทเมนต์เดียวกันแต่ให้ผลแตกต่างกัน เนื่องมาจากความแปรปรวนภายในหน่วยทดลองเอง โดยทั่วไปถ้าใช้หน่วยทดลองขนาดใหญ่จะทำให้มีความแปรปรวนน้อยกว่าการใช้หน่วยทดลองขนาดเล็ก

2. จากการดำเนินการทดลองที่ไม่มีมาตรฐานเกี่ยวกับการออกแบบการทดลองและการสุ่ม เป็นความผันแปรภายนอกหน่วยทดลอง เช่นในการเก็บข้อมูลโดยการวัดหรือการนับอาจมีความคลาดเคลื่อนเนื่องจากการใช้เครื่องมือวัดแตกต่างกัน เวลาในการเก็บข้อมูลแตกต่างกัน ตำแหน่งที่วางหน่วยทดลองแตกต่างกัน ตำแหน่งที่เก็บข้อมูลบนหน่วยทดลองแตกต่างกัน เป็นต้น

วิธีที่ช่วยให้ผลการทดลองมีความถูกต้องแม่นยำเพิ่มขึ้น โดยการลดความคลาดเคลื่อนของการทดลองลง มีวิธีการดังนี้

1. การทำการทดลองซ้ำ (Replications) คือให้แต่ละทรีทเมนต์มีจำนวนหน่วยทดลองมากกว่า 1 หน่วยทดลอง

2. เพิ่มการวัดตัวแปรที่รบกวนผลการทดลอง (Additional measurements)

เนื่องจากตัวแปรตาม y ได้รับอิทธิพลจากทรีทเมนต์และยังมีความสัมพันธ์กับตัวแปร x ตัวอย่างเช่น การเปรียบเทียบวิธีการลดน้ำหนัก 3 วิธี ทรีทเมนต์คือวิธีการลดน้ำหนัก การวิเคราะห์ความแปรปรวนธรรมดาอาจไม่สามารถตัดสินใจได้ว่ามีความแตกต่างของอิทธิพลของทรีทเมนต์ทั้ง 3 อย่างมีนัยสำคัญทางสถิติ เพราะความแปรปรวนภายในทรีทเมนต์มีขนาดโตมาก อาจเนื่องมาจากน้ำหนักที่ลดลง y มีความสัมพันธ์เชิงเส้นกับน้ำหนักตอนเริ่มการทดลอง x เราสามารถลดความแปรปรวนภายในทรีทเมนต์ได้โดยการผสมผสานของการวิเคราะห์การถดถอยของ y บน x กับการวิเคราะห์ความแปรปรวนของ y เรียกว่า การวิเคราะห์ความแปรปรวนร่วม ถ้าไม่ใช้การวิเคราะห์ความแปรปรวนร่วมในการทดลองนี้จะต้องเลือกคนที่เข้ารับการทดลองที่มีน้ำหนักเริ่มต้นก่อนการทดลองเท่ากัน หรือแบ่งกลุ่มของคนที่เข้ารับการทดลองออกเป็นกลุ่ม ๆ ตามน้ำหนักเริ่มต้นก่อนการทดลอง ซึ่งเรียกว่าวิธีการบล็อก

3. เทคนิคการบล็อก (Blocking) เมื่อมีความแตกต่างระหว่างหน่วยทดลองมาก ถ้ามีความแตกต่างเพียง 1 ปัจจัย ก็ออกแบบการทดลองโดยการบล็อก 1 ทาง เรียกว่าการออกแบบการ

ทดลองแบบบล็อกสมบูรณ์ ถ้ามีความแตกต่าง 2 ปัจจัย ก็ออกแบบการทดลองโดยการบล็อก 2 ทาง เรียกว่าการออกแบบการทดลองแบบลาตินสแคว เป็นต้น

1.3 การทำการทดลองซ้ำ

จะเห็นได้ว่าเวลาที่มีการทำการทดลองซ้ำ ๆ กัน อิทธิพลของทรีทเมนต์ในการทดลองแต่ละครั้งจะได้ผลการทดลองไม่เหมือนกันในการทดลองแต่ละครั้ง เราเรียกว่ามีความแปรปรวนของผลการทดลอง ถ้าผลการทดลองมีความแปรปรวนมาก ก็จะเป็นตัวบ่งชี้ว่าผลการทดลองมีความไม่แน่นอนสูง การทำการทดลองหลาย ๆ ครั้งหรือที่เรียกว่าการทำซ้ำ (replications) ภายใต้สถานการณ์เดียวกันทุกครั้ง ซึ่งไม่ทราบว่าต้องทำซ้ำกี่ครั้งจึงจะสรุปผลได้ถูกต้อง นี่คือปัญหาของการสรุปผลจากตัวอย่างบางส่วนไปสู่ประชากรทั้งหมด ซึ่งภาษาทางสถิติเรียกว่าสรุปจากตัวอย่างไปสู่ประชากร วิธีการแก้ปัญหานี้ทางสถิติ คือ การหาค่าเฉลี่ยที่เป็นการใช้หลักการทางสถิติช่วยในการแปลความหมายของผลการทดลอง

การประมาณค่าความแปรปรวนของข้อมูลที่สังเกตได้จากแต่ละการทดลองจะได้ค่า σ^2 เนื่อง จากหน่วยทดลองต่าง ๆ ที่ใช้ในการทดลองอาจมีความแตกต่างกัน การประมาณค่านี้ได้มาจากการทำการทดลองทรีทเมนต์หนึ่งบนหน่วยทดลองมากกว่า 1 หน่วยทดลอง เรียกว่า การทำซ้ำ

จำนวนของการทำซ้ำขึ้นกับจำนวนของทรีทเมนต์ และการออกแบบการทดลอง ถ้ามีจำนวนซ้ำมากพอเราจะสามารถมองเห็น outlier ของข้อมูล ถ้าเกิด outlier ขึ้น เราอาจตัดข้อมูลตัวนั้นออกไปจากการวิเคราะห์ หรือถ้าหน่วยทดลองนั้นยังอยู่เราอาจตรวจหน่วยทดลองนั้นอีกครั้ง ในการทดลองขนาดใหญ่มากบางครั้งอาจทำการทดลองเพียงซ้ำเดียวได้ก็จริง แต่นักวิทยาศาสตร์ส่วนมากก็จะไม่สบายใจที่จะสรุปผลการทดลองจากการทดลองเพียงซ้ำเดียวนั้น ถ้าเป็นไปได้จะทำการทดลองซ้ำอีกครั้งในภายหลังในโครงการวิจัยนั้น

การทำซ้ำ หรือจำนวนครั้งของการทดลอง ก็คือจำนวนหน่วยทดลองที่มีในแต่ละทรีทเมนต์ การทำซ้ำมีประโยชน์คือ

1. ช่วยให้ประมาณค่าความคลาดเคลื่อนของการทดลองได้ ซึ่งการประมาณค่าความคลาดเคลื่อนของการทดลองหาได้จากค่าความแปรปรวนของข้อมูลที่สังเกตได้ ที่เกิดจากการทำการทดลองซ้ำหลายครั้งในแต่ละทรีทเมนต์

2. ช่วยให้การประมาณค่าสถิติของทรีทเมนต์มีความถูกต้อง แม่นยำเพิ่มขึ้น ถ้าให้ σ^2 คือ ความแปรปรวนของค่าสังเกต r คือจำนวนซ้ำ และ $\sigma^2_{\bar{y}}$ คือ ความแปรปรวนของประชากร $\bar{y}$

ดังนั้น การประมาณค่า $\sigma^2_{\bar{y}}$ จากสูตร $\sigma^2_{\bar{y}} = \frac{\sigma^2}{r}$ ค่า $\sigma^2_{\bar{y}}$ จะมีค่าน้อยคือความถูกต้องแม่นยำเพิ่มขึ้น ถ้าเพิ่ม จำนวนซ้ำ

สำหรับการวางแผนการทดลองแบบสุ่มสมบูรณ์ภายในบล็อก บล็อกก็คือ จำนวนซ้ำนั่นเอง

3. ขยายขอบเขตของการสรุปผล เช่น ในการทดลองหนึ่ง ถ้าทำซ้ำที่หลาย ๆ สถานีทดลอง คือ ทดลองปลูกข้าวที่รังสิต สระบุรี อุบลราชธานี จะสรุปผลการทดลองได้กว้างขวางกว่าทำเฉพาะสถานีทดลองเดียว

1.4 การสุ่ม (Randomization)

การสุ่ม หมายถึงการที่หน่วยทดลองหน่วยใด ๆ มีโอกาสได้รับทรีทเมนต์ใดทรีทเมนต์หนึ่งโดยสุ่ม อาจใช้วิธีการสุ่มหน่วยทดลองให้แก่ทรีทเมนต์ หรือใช้วิธีการสุ่มทรีทเมนต์ให้แก่หน่วยทดลองก็ได้

จุดประสงค์ของการสุ่มคือ

1. เพื่อให้ได้ค่าประมาณที่ถูกต้องของความคลาดเคลื่อนของการทดลอง โดยทฤษฎีสถิติถือว่าความคลาดเคลื่อนของการทดลองเป็นความแปรปรวนสุ่ม

2. ในการวิเคราะห์ความแปรปรวนมีข้อกำหนดว่าความคลาดเคลื่อนของการทดลองจะต้องเกิดขึ้นอย่างสุ่มและเป็นอิสระกัน ความคลาดเคลื่อนจะต้องมีการกระจายแบบอิสระ โดยการจัดหน่วยทดลองใด ๆ แก่ทรีทเมนต์ใด ๆ โดยการสุ่ม

3. เพื่อไม่ให้เกิดอคติในการจัดหน่วยทดลองให้แก่ทรีทเมนต์

วิธีการสุ่ม ทำได้หลายวิธีคือ 1. จับฉลาก 2. ใช้ตารางเลขสุ่ม 3. สุ่มจาก permutation

สมมติว่าเราต้องการเปรียบเทียบทรีทเมนต์ A และ B และเรามีหน่วยทดลอง 20 หน่วย เราจะให้หน่วยทดลองใดได้รับแต่ละทรีทเมนต์นี้ และแต่ละทรีทเมนต์จำเป็นต้องมีหน่วยทดลองเท่ากันเท่ากับ 10 หน่วยหรือไม่ จริง ๆ แล้วอาจไม่จำเป็น แต่ต้องมีเหตุผลอธิบายได้ เนื่องจากการที่ทรีทเมนต์มีหน่วยทดลองมากกว่าอีกทรีทเมนต์หนึ่งจะเป็นสาเหตุให้ค่าเฉลี่ยของทรีทเมนต์หนึ่งที่ประมาณได้มีความเที่ยงตรงมากกว่าอีกทรีทเมนต์หนึ่ง เนื่องจากความแปรปรวนของค่าเฉลี่ยเท่ากับ σ^2/n เมื่อ n คือจำนวนหน่วยทดลองของทรีทเมนต์นั้น ดังนั้นในที่นี้เราจึงจัดให้ทรีทเมนต์ละ 10 หน่วยทดลอง

วิธีการจัดหน่วยทดลองให้กับทรีทเมนต์ ถ้าเราอาศัยความสะดวกสบาย ตัวอย่างเช่น จัดให้ต้นไม้ 10 แถวแรก สัตว์ทดลอง 10 ตัวแรกที่จับได้ คนไข้ 10 คนแรกที่ป่วย การทำงาน 10 ครั้งแรกของเครื่องจักรที่ดำเนินการในวันปกติที่โรงงานแห่งหนึ่ง เด็กนักเรียน 10 คนแรกในรายการชื่อ ให้ได้รับทรีทเมนต์ A ถ้าเราจัดหน่วยทดลองด้วยวิธีนี้ เราอาจกำลังกำหนดผลการทดลองโดยการเลือกต้นไม้ที่ดีที่สุด 10 แถวแรก ให้ได้รับทรีทเมนต์หนึ่ง และที่เหลือซึ่งคุณภาพไม่ดีให้รับอีกทรีทเมนต์หนึ่ง หรือสัตว์ทดลอง 10 ตัวแรกที่ไม่แข็งแรงถูกจับได้ง่ายให้ทรีทเมนต์หนึ่ง หรือคนไข้ 10 คนแรกที่ได้รับการวินิจฉัยก่อน เนื่องจากป่วยหนักมากชัดเจน หรือช่วงเวลาการทำงานของเครื่องจักรในช่วงแรกซึ่งยังทำงานได้ไม่สมบูรณ์ เป็นการออกแบบการทดลองแบบง่าย ๆ ผลการทดลองที่ได้จากการเปรียบเทียบทรีทเมนต์ 2 ทรีทเมนต์นี้จะได้รับการสงสัยเคลือบแคลง โดยเฉพาะอย่างยิ่งคือ ผลการทดลองที่แตกต่างกันนี้อาจมาจากความแตกต่างของหน่วยทดลองนี้ ซึ่งทราบแล้วตั้งแต่ก่อนทำการทดลอง หรือผลการทดลองนี้อาจไม่ได้รับความเชื่อถือเลย

การสุ่มตัวอย่างให้หน่วยทดลองใด ๆ มีโอกาสได้รับทรีทเมนต์ใด ๆ เป็นไปโดยสุ่ม จะทำให้การเลือกหน่วยทดลองให้กับทรีทเมนต์ใด ๆ ไม่ได้ดำเนินการโดยอาศัยความพึงพอใจ และยังเป็นวิธีการที่ช่วยให้การทดสอบทางสถิติมีความถูกต้องน่าเชื่อถือ ถึงแม้ว่าเทอมเศษตกค้าง (residual terms) ε_{ij} ในตัวแบบเส้นตรงอาจไม่เป็นไปตามข้อตกลงเบื้องต้นเกี่ยวกับความเป็นปกติ

การสุ่มทำเพื่อป้องกันหรือควบคุมตัวแปรรบกวนภายนอกที่อาจเกิดขึ้นในระหว่างการทดลอง เพื่อไม่ให้ทรีทเมนต์ต่าง ๆ ในการทดลองมีความได้เปรียบหรือเสียเปรียบกัน ซึ่งอาจมีผลกระทบต่อผลการทดลอง และเพื่อช่วยให้ความถูกต้องของผลการทดลองเพิ่มขึ้น ยกตัวอย่างเช่น ลำดับที่ของทรีทเมนต์ในการทดลองครั้งหนึ่งหรือในแต่ละซ้ำ อาจมีผลกระทบต่อผลการทดลอง สามารถควบคุมได้โดยการสุ่ม ทำให้แต่ละทรีทเมนต์มีโอกาสเท่ากันในการเกิดขึ้นก่อนหรือหลังในการทดลองแต่ละครั้งหรือแต่ละซ้ำ

หลักพื้นฐานของการออกแบบการทดลองทั้งหมดเกี่ยวข้องกับการสุ่ม ในการออกแบบการทดลองแบบสุ่มสมบูรณ์ ก่อนทำการทดลองต้องมีการสุ่มทรีทเมนต์ให้กับหน่วยทดลอง สำหรับการออกแบบการทดลองแบบสุ่มสมบูรณ์ภายในบล็อก ก่อนทำการทดลองต้องมีการเลือกบล็อกมาอย่างสุ่ม และการกำหนดทรีทเมนต์ให้หน่วยทดลองภายในบล็อกหนึ่งต้องจัดเป็นอันดับอย่างสุ่ม

การสุ่มนี้จะทำให้ความคลาดเคลื่อนของการทดลอง ε ในตัวแบบสถิติของการทดลองนั้นเป็นตัวแปรสุ่ม เพื่อให้เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนที่ว่าในแต่ละ

กลุ่มของทรีทเมนต์ ความคลาดเคลื่อนของการทดลอง ε มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนคงที่เท่ากับ σ^2 และความคลาดเคลื่อนแต่ละตัวเป็นอิสระกัน

ตัวอย่างที่ 1 การศึกษาความเหนียวของใยสังเคราะห์ตัวใหม่ [จาก Montgomery (1997)] ซึ่งจะนำมาตัดเป็นเสื้อเชิ้ตผู้ชาย จากประสบการณ์ผู้วิจัยทราบว่า ความเหนียวของเส้นใยมีผลมาจากน้ำหนักเป็นเปอร์เซ็นต์ของฝ้ายที่อยู่ในใยสังเคราะห์ ผู้วิจัยคาดหมายว่า การเพิ่มน้ำหนักเปอร์เซนต์ของฝ้ายจะเพิ่มความเหนียวของเส้นใย และผู้วิจัยทราบว่าเส้นใยสังเคราะห์ควรประกอบด้วยฝ้าย 10% ถึง 40% ผู้วิจัยตัดสินใจทดสอบใยสังเคราะห์ที่ผลิตจากน้ำหนักเปอร์เซนต์ของฝ้าย 5 ระดับ คือ 15%, 20% , 25%, 30% และ 35% และผู้วิจัยตัดสินใจทดสอบใยสังเคราะห์ 5 ชิ้นสำหรับแต่ละระดับของน้ำหนักเปอร์เซนต์ฝ้ายในใยสังเคราะห์

นี่คือตัวอย่างการทดลองที่มีปัจจัยเดียวซึ่งมี 5 ระดับ และทำการทดลอง 5 ซ้ำ หน่วยทดลองคือชิ้นใยสังเคราะห์มีทั้งหมด 25 ชิ้น ดำเนินการทดลองโดยให้ลำดับการทดลองใยสังเคราะห์ทั้ง 25 ชิ้นเป็นไปอย่างสุ่ม วิธีการสุ่มมีขั้นตอนดังนี้

1. กำหนดตัวเลขลำดับของชิ้นใยสังเคราะห์ที่เป็นหน่วยทดลอง ซึ่งมีน้ำหนักเปอร์เซนต์ฝ้ายระดับต่าง ๆ ได้ดังนี้

ตารางที่ 1.1

น้ำหนักเปอร์เซนต์ฝ้าย	ตัวเลขอันดับของหน่วยทดลอง				
15	1	2	3	4	5
20	6	7	8	9	10
25	11	12	13	14	15
30	16	17	18	19	20
35	21	22	23	24	25

2. สุ่มลำดับที่ของการทดลองซึ่งใช้เครื่องจักรตัวเดียวกันให้กับชิ้นใยสังเคราะห์แต่ละชิ้นโดยสุ่มตัวเลข 1 ถึง 25 สมมุติว่าได้หมายเลข 19 คือ ชิ้นใยสังเคราะห์ที่มีน้ำหนักเปอร์เซนต์ฝ้าย 30% ทำการทดลองเป็นอันดับแรก ทำซ้ำอีกจนครบทั้ง 25 ชิ้น ได้อันดับการทดลองทั้ง 25 ชิ้นดังนี้

ตารางที่ 1.2

ลำดับการทดลอง	ตัวเลขอันดับของชิ้นใยสังเคราะห์ (หน่วยทดลอง)	น้ำหนักเปอร์เซนต์ฝ้าย (ทรีทเมนต์)
1	19	30
2	4	15
3	1	15
4	12	25
5	14	25
6	24	35
7	9	20
8	3	15
9	22	35
10	7	20
11	6	20
12	23	35
13	11	25
14	2	15
15	10	20
16	17	30
17	15	25
18	5	15
19	13	25
20	21	35
21	16	30
22	18	30
23	8	20
24	25	35
25	20	30

การจัดอันดับการทดลองของชิ้นใยสังเคราะห์ทั้งหมดเป็นไปอย่างสุ่มเป็นสิ่งจำเป็นทำเพื่อป้องกันอิทธิพลของตัวแปรภายนอกที่ไม่ทราบว่ามีอยู่หรือไม่ หรือไม่สามารถควบคุมได้ในระหว่างการทดลอง ถ้าเราทำการทดลองโดยไม่ทำให้อันดับของการทดลองเป็นไปอย่างสุ่มคือ ครั้งแรกทำการทดลองกับชิ้นใยสังเคราะห์ที่มีเปอร์เซนต์ฝ้าย 15% ก่อนทั้ง 5 ชิ้น ต่อมาทำการทดลองกับชิ้นใยสังเคราะห์ที่มีเปอร์เซนต์ฝ้าย 20% ต่อไปเรื่อย ๆ จนถึงชิ้นใยสังเคราะห์ที่มีเปอร์เซนต์ฝ้าย

35% การเปิดเครื่องจักรในการทำงานครั้งแรกอาจมีผลต่อการอ่านค่าข้อมูลความเหนียวของชิ้นใยสังเคราะห์

อีกตัวอย่างหนึ่ง การปรุงอาหารเนื้ออบที่อุณหภูมิต่างกัน 4 ระดับด้วยเตาอบ 4 ตู้ ถ้าทำการทดลอง 4 ซ้ำ เราจะไม่ใช้เตาอบตู้เดียวกันที่อุณหภูมิเดียวตลอดทั้ง 4 ซ้ำ เพราะเตาอบแต่ละตู้อาจมีความแตกต่างกัน ซึ่งอาจมีผลกระทบต่อการปรุงอาหาร จึงควรทำการสุ่มอุณหภูมิให้กับ เตาอบในแต่ละซ้ำ สามารถทำการสุ่ม 2 ครั้ง คือ สุ่มอุณหภูมิให้กับเนื้อ และสุ่มเนื้อให้กับเตาอบ หรืออีกวิธีหนึ่งคือ สุ่มเนื้อให้เตาอบ แล้วสุ่มอุณหภูมิให้เตาอบ

การสุ่มช่วยให้ค่าสถิติที่ประมาณได้สามารถสรุปอ้างอิงไปสู่ประชากร

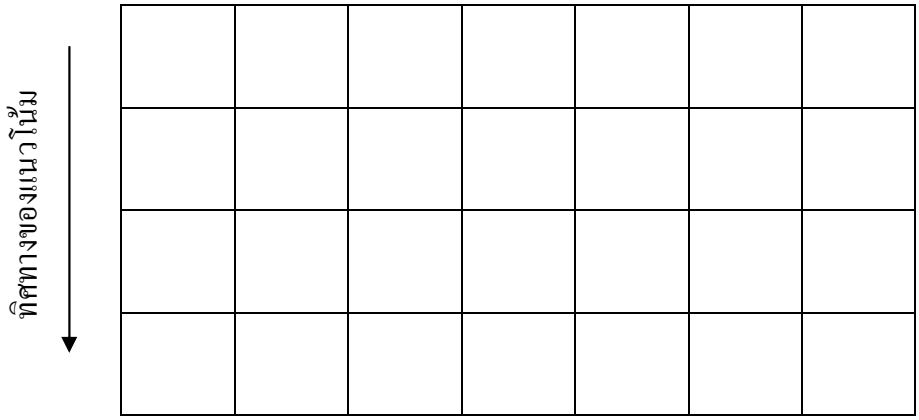
1.5 การบล็อก (Blocking)

การบล็อกเป็นเทคนิคในการแยกอิทธิพลของปัจจัยรบกวนที่เราไม่ต้องการศึกษาแต่ปัจจัยรบกวนนั้นมีผลกระทบกับผลการทดลอง

การบล็อกคือ การรวมกลุ่มหน่วยทดลองที่แตกต่างกันในด้านต่าง ๆ เช่น

1. ลักษณะพื้นที่ของแปลงพืช เช่น ความลาดเทของพื้นที่ ความห่างไกลของพื้นที่ คือ พื้นที่ที่อยู่ติดกันจะมีความเหมือนกันมากกว่าพื้นที่ที่อยู่ห่างไกลกัน หรือในกรีนเฮ้าส์ที่มีอุณหภูมิต่างกัน แสงแดดไม่เท่ากัน บรรยากาศไม่เหมือนกัน น้ำไม่เท่ากัน ความอุดมสมบูรณ์ของแปลง สภาพแหล่งน้ำ
2. คุณลักษณะทางกายภาพของพืชหรือสัตว์ที่ใช้ทดลอง เช่น อายุ น้ำหนัก เพศ ความแข็งแรง พันธุกรรม
3. เวลาในการทดลอง เช่น ช่วงเวลา เช้า-เย็น กลางวัน-กลางคืน วัน สัปดาห์ เดือน ปีที่ทำการทดลองแตกต่างกัน อาจมีผลกระทบต่อการทดลอง

จากตัวอย่างการเปรียบเทียบอิทธิพลของปุ๋ย 6 ชนิดในพื้นที่การทดลองผืนหนึ่ง ซึ่งอาจมีอิทธิพลเกี่ยวกับอากาศของภูมิประเทศนั้น บางครั้งอาจมีอิทธิพลของอุณหภูมิที่เปลี่ยนแปลงไป ความชันของพื้นที่ ความหนาวเย็นบนพื้นที่ที่อยู่บนยอดสุด ความชื้นจากหมอกที่ปกคลุมพื้นที่ที่อยู่ด้านล่างสุดของพื้นที่ที่มีความลาดชัน หรืออาจมีที่กำบังบนพื้นที่การทดลองด้านหนึ่งมากกว่าพื้นที่การทดลองอีกด้านหนึ่งซึ่งเป็นพื้นที่หุบเขา ตัวอย่างเช่น ภาพพื้นที่ต่อไปนี้

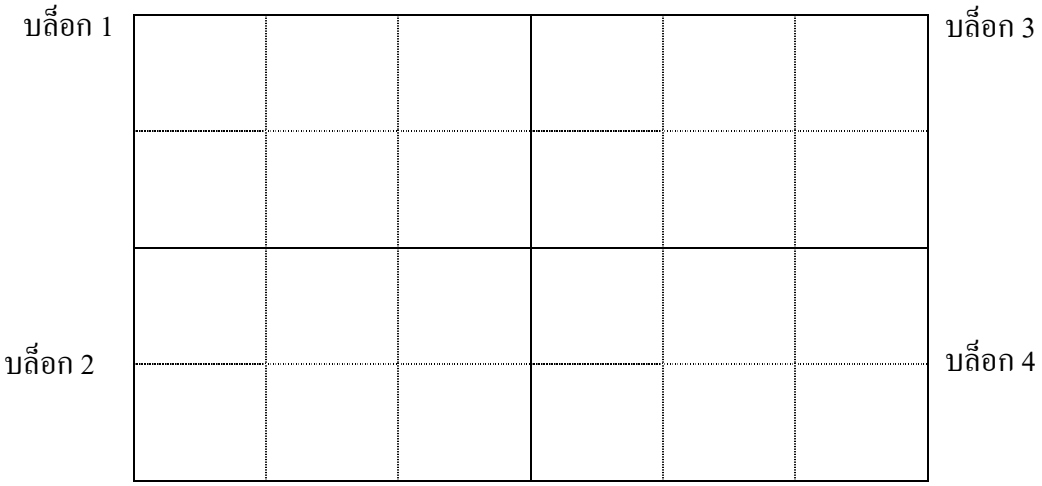


ภาพที่ 1.1

ความแตกต่างของพื้นที่เหล่านี้ทำให้เราต้องออกแบบการทดลองที่สามารถแยกแหล่งของความแปรปรวนเนื่องจากปัจจัยรบกวนเหล่านี้ออกมา เช่น การออกแบบการทดลองแบบบล็อก การออกแบบการทดลองแบบลาตินสแคว เป็นต้น

เมื่อมีการบล็อกในการออกแบบการทดลอง เราจะสามารถแยกแหล่งของความแปรปรวนของปัจจัยรบกวนใด ๆ ทั้งที่ทราบหรือสงสัยว่าอาจเป็นแหล่งของความแปรปรวนของข้อมูล ตัวอย่างเช่น โดยส่วนใหญ่พื้นที่การทดลองมักใช้เวลาในการเก็บเกี่ยวหลายวันเพื่อเก็บข้อมูลเกี่ยวกับผลผลิต ถ้าอากาศไม่ดีก็อาจทำให้การเก็บเกี่ยวยุ่งยากขึ้นอีก ดังนั้นการเก็บเกี่ยวเพื่อเก็บข้อมูลเกี่ยวกับผลผลิตนี้ภายในบล็อกหนึ่งให้เสร็จสมบูรณ์ก่อนที่จะเริ่มทำแบบเดียวกันในบล็อกต่อไปจะทำให้สาเหตุของความแตกต่างเนื่องจากระบบการเก็บเกี่ยวนี้ถูกแยกออกมาอยู่ในเทอมของบล็อก หรือแหล่งของความแปรปรวนเนื่องมาจากบล็อกนี้ในการวิเคราะห์ ถ้าบล็อก 1 ถูกเก็บข้อมูลก่อนที่จะมีพายุฝน ส่วนบล็อกที่เหลืออาจมีความแตกต่างที่เกิดขึ้นอย่างเป็นระบบนี้จากบล็อกที่1 เพราะว่าพายุฝนได้ทำลายผลผลิตบางส่วน หรือเป็นเพราะว่าผลผลิตแก่จัดนิดหน่อยที่วันเก็บเกี่ยวนั้น หรือบางทีแก่เกินไปหรือเป็นเพราะว่าการเก็บเกี่ยวผลผลิตเมื่อตอนที่อากาศแห้งหรือฝนยังไม่ตกมีความสะดวกหรือง่ายกว่าขณะที่พื้นดินยังแฉะในวันที่ 2 หลังฝนตก

เราอาจใช้บล็อกในการแบ่งพื้นที่ทดลองทั้งหมดออกเป็นส่วนๆ ตามเวลาที่ดำเนินการเพาะปลูก เพราะเกษตรกรไม่สามารถดำเนินการเพาะปลูกได้พร้อมกันทั้งหมดของพื้นที่ทดลอง ดังนั้นความแตกต่างของเวลาที่ทำการเพาะปลูกก่อนหลังนี้สามารถจัดให้เป็นบล็อกได้ เราอาจแบ่งบล็อกเพื่อให้ง่ายในการจัดการหน่วยทดลองก็ได้ โดยให้แต่ละบล็อกประกอบด้วยกลุ่มของแปลง ดังภาพ



ภาพที่ 1.2

บล็อกอาจมีรูปร่างแบบใดก็ได้ที่จะทำให้สามารถจัดกลุ่มของหน่วยทดลองที่มีความเหมือนกันอยู่ในแต่ละบล็อกโดยที่มีความแตกต่างที่เกิดขึ้นอย่างเป็นระบบแตกต่างกันระหว่างบล็อกต่าง ๆ

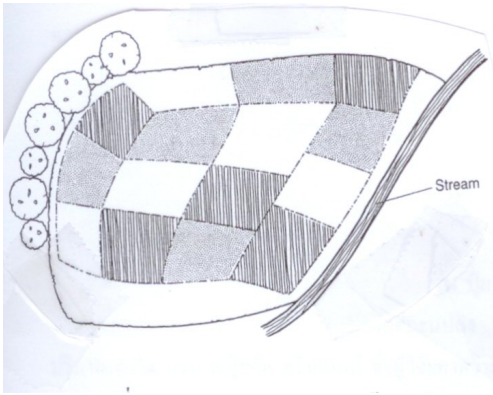
ในพื้นที่ทดลอง ต้นไม้ที่อยู่ใกล้ด้านขอบของพื้นที่อาจจะถูกกำบังมากกว่าพื้นที่ที่อยู่ตรงกลาง เราสามารถจัดแบ่งพื้นที่ออกเป็นบล็อกเพื่อแยกอิทธิพลที่เกิดขึ้นอย่างเป็นระบบนี้ออกมาได้ สำหรับการทดลองในเรือนกระจกที่มีการควบคุม แปลงที่อยู่ใกล้หน้าต่างของเรือนกระจก หรือส่วนริมอาจถูกกระทบจากภายนอกได้มากกว่าส่วนอื่น ๆ ดังนั้นเราควรต้องแยกอิทธิพลเนื่องจากอยู่ใกล้หน้าต่างหรือส่วนริมนี้ออกมาก่อนที่จะให้ทรีทเมนต์แก่หน่วยทดลอง

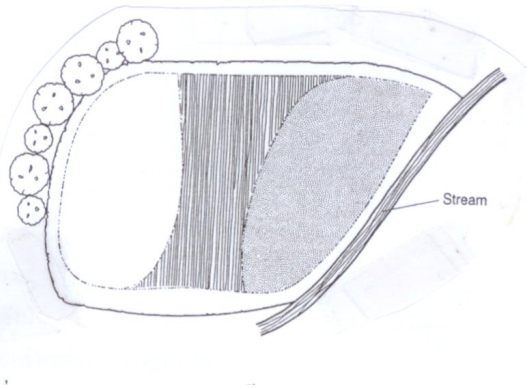
การทดลองในโรงงานอุตสาหกรรมที่ใช้เครื่องจักรหลายตัวในการทำงานแบบเดียวกันก็อาจทำให้เกิดความแตกต่างอย่างเป็นระบบขึ้น ดังนั้นเราอาจจัดให้เครื่องจักรแต่ละตัวเป็นบล็อกหรือผู้ปฏิบัติงานคนละคนที่ควบคุมการทำงานของเครื่องจักรตัวเดียวกันก็อาจทำให้เกิดความแตกต่างอย่างเป็นระบบขึ้นได้ หรือช่วงเวลาการปฏิบัติงานในแต่ละวันหรือวันต่าง ๆ ในสัปดาห์ ก็

อาจจัดเป็นบล็อกได้ หรือเมื่อใช้วัตถุดิบที่ได้มาจากบริษัทแตกต่างกัน เราก็อาจจัดให้วัตถุดิบที่มาจากบริษัทหนึ่งเป็นบล็อกหนึ่ง ซึ่งในแต่ละบล็อกมีหลายหน่วยทดลอง การทดลองเกี่ยวกับการวิจัยตลาด เช่น การขายสินค้าชนิดหนึ่งอาจมีความแตกต่างกันในพื้นที่ต่าง ๆ ของประเทศ เราก็อาจจัดการบล็อกพื้นที่ต่าง ๆ นั้น

2. ข้อมูล

การเก็บข้อมูลจากการทดลองนั้นต้องมีการควบคุมสถานการณ์ในการทดลอง ตัวอย่างเช่น ผู้วิจัยอยากทราบว่าหญ้าชนิดใดเหมาะที่สุดที่จะใช้บนสนามกีฬา การดำเนินการทดลองทำโดยการทดลองปลูกหญ้าชนิดต่าง ๆ บนแปลงที่มีลักษณะคล้ายคลึงกัน แล้วเก็บข้อมูลการเจริญเติบโตและการพัฒนาของหญ้าแต่ละชนิด แปลงที่ปลูกหญ้าแต่ละชนิดนี้ควรเป็นแปลงเล็ก ๆ ที่มีสภาพใกล้เคียงกันมากที่สุด ดังแผนภาพการจัดแปลงทดลองที่ใช้ปลูกหญ้า 3 ชนิด ดังต่อไปนี้





ภาพที่ 1.3 แผนภาพแปลงทดลองหญ้า 3 ชนิด **ภาพที่ 1.4**แผนภาพแปลงทดลองหญ้า 3 ชนิด

หญ้าชนิดที่ 1
หญ้าชนิดที่ 2
หญ้าชนิดที่ 3

ภาพที่ 1.3 แสดงแผนภาพการทดลองปลูกหญ้า 3 ชนิด โดยแบ่งพื้นที่ทดลองออกเป็นแปลงเล็ก ๆ 15 แปลง หญ้าแต่ละชนิดดำเนินการทดลองปลูกในแปลงต่าง ๆ 5 แปลง โดยการสุ่ม

ทำให้หญ้าแต่ละชนิดมีโอกาสปลูกในแปลงที่อยู่ใต้ร่มเงาของต้นไม้ แปลงที่อยู่ใกล้ลำธาร และแปลงที่อยู่ตรงกลางของพื้นที่ทดลอง ซึ่งเป็นการควบคุมตัวแปรสิ่งแวดล้อมที่อาจมีผลต่อการทดลองได้ ภาพที่ 1.4 แบ่งพื้นที่ทดลองออกเป็น 3 แปลง ให้แปลงที่ 1 ปลูกหญ้าชนิดที่ 1 ซึ่งแปลงนี้อยู่ใต้ร่มเงาของต้นไม้ใหญ่ ให้แปลงที่ 2 ปลูกหญ้าชนิดที่ 2 ซึ่งแปลงนี้อยู่ตรงกลางของพื้นที่ทดลอง ให้แปลงที่ 3 ปลูกหญ้าชนิดที่ 3 ซึ่งแปลงนี้อยู่ติดกับลำธาร จะเห็นว่าแต่ละแปลงมีสภาพแวดล้อมแตกต่างกัน

สิ่งแวดล้อมสำหรับการเจริญเติบโตของหญ้าที่ปลูกในสนามกีฬานี้จะเห็นว่าขึ้นกับปัจจัยหลายอย่าง เช่น ชนิดของดิน ได้แก่ ดินทราย ดินเหนียว ดินเปรี้ยว เป็นต้น ความลาดเอียง สภาพดินฟ้าอากาศ ปุ๋ยที่ใช้ก่อนการทดลองและปุ๋ยที่ใช้ระหว่างการทดลอง การกำจัดวัชพืชชนิดอื่น ๆ ในแปลง ทดลอง ปัจจัยต่าง ๆ เหล่านี้บางปัจจัยก็สามารถควบคุมได้โดยง่าย ตัวอย่างเช่น การควบคุมความแตกต่างของดินโดยการทำการทดลองกับดินชนิดหนึ่งโดยเฉพาะ เช่น ดินทราย ผลการทดลองที่ได้เราจะไม่สรุปว่าจะได้ผลการทดลองแบบเดียวกันกับดินชนิดอื่น ๆ แต่ก่อนที่จะสรุปผลการทดลอง ผู้วิจัยควรต้องทำการทดลองซ้ำภายใต้สถานการณ์แบบเดียวกัน สำหรับบางปัจจัยที่เราอาจไม่สามารถควบคุมได้เลย เช่น สภาพดินฟ้าอากาศในแต่ละฤดู ที่เราสามารถทำได้คือสังเกตอิทธิพลของสภาพดินฟ้าอากาศต่อหญ้าชนิดต่าง ๆ หญ้าบางชนิดอาจเจริญเติบโตได้ดีในฤดูร้อน เมื่อเปรียบ เทียบกับหญ้าชนิดอื่น ๆ แต่เจริญเติบโตไม่ดีในฤดูฝน เป็นต้น

ขณะที่บางปัจจัยเราสามารถควบคุมได้ เช่น ปุ๋ย เราอาจใช้ปุ๋ยชนิดต่าง ๆ ในแปลงทดลอง โดยควบคุมได้ว่าจะใช้ปุ๋ยปริมาณเท่าใดในแต่ละแปลง ดังนั้นเราจึงทราบว่าแต่ละแปลงได้รับปุ๋ยปริมาณเท่าใด ปริมาณปุ๋ยคือ ทรีทเมนต์ ซึ่งผู้วิจัยคาดว่ามีอิทธิพลต่อการเจริญเติบโต การทดลองนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความแตกต่างระหว่างทรีทเมนต์ต่าง ๆ

2.1 การวัดข้อมูล

ในการทดลองเราจะวัดข้อมูลที่เป็นผลมาจากอิทธิพลของทรีทเมนต์ในการทดลอง และสามารถนำมาเปรียบเทียบระหว่างทรีทเมนต์ต่าง ๆ ได้ ซึ่งเราต้องคิดให้ดีตั้งแต่ในขั้นตอนการวางแผนการทดลอง ตัวอย่างเช่น ในการทดลองเกี่ยวกับการเพาะปลูกผลไม้ชนิดหนึ่งซึ่งมีวางขายอยู่ในร้านในตลาด จะเห็นว่าสิ่งสำคัญคือ น้ำหนักผลผลิตโดยรวมทั้งหมด ถ้าการเพาะปลูกได้ผลไม้ที่มีน้ำหนักผลผลิตโดยรวมมากก็จะทำให้ผู้ปลูกหรือเกษตรกรมีรายได้ดี แต่อย่างไรก็ตามขนาดของผลไม้แต่ละลูกก็มีความสำคัญมากเช่นเดียวกัน เราสามารถวัดได้ในการเก็บเกี่ยวแต่ละครั้ง โดยการ

สุ่มตัวอย่างผลไม้มา N ลูก ชั่งน้ำหนักแล้วหารน้ำหนักผลผลิตโดยรวมทั้งหมด เพื่อจัดเป็นเกรดต่าง ๆ โดยที่เกรด 1 จะได้ราคาแพงที่สุด ไม่มีโรคหรือตำหนิ ซึ่งต้องมีการตรวจสอบผลไม้กลุ่มตัวอย่างเกี่ยวกับเปอร์เซ็นต์ของผลที่มีตำหนิ หรือถูกทำลายจากเชื้อโรค หรือถ้านำผลผลิตผลไม้ไปทำเป็นน้ำผลไม้ก็ต้องมีการวิเคราะห์ทางเคมี เพื่อวัดเกี่ยวกับวิตามิน

ถ้าในขั้นตอนการวางแผนไม่ได้วางแผนไว้อย่างดี ข้อมูลบางตัวอาจไม่ได้วัด ทำให้ไม่สามารถหาคำตอบสำหรับคำถามที่อยู่ในการทดลองนั้น การวัดข้อมูลจากหน่วยทดลองมักจะเรียกว่า ผล (yield) หรือผลตอบสนอง (response) ของหน่วยทดลองที่ได้รับทรีทเมนต์

2.2 ความผันแปรของข้อมูล

การเก็บข้อมูลจากการทดลองจะพบว่าข้อมูลมีความผันแปรอย่างสุ่มระหว่างหน่วยทดลองต่าง ๆ ที่ผู้วิจัยใช้ในการทดลองนั้น ๆ ดังนั้นในงานวิจัยทางชีววิทยาและทางการเกษตรส่วนมากซึ่งเกี่ยวข้องกับการออกแบบการทดลองและการวิเคราะห์ข้อมูลที่ได้จากการทดลองจึงให้ความสนใจเกี่ยวกับความผันแปรของข้อมูลที่เกิดขึ้นอย่างสุ่มนี้ แม้แต่การทดลองที่ทำในเรือนกระจกที่มีการควบคุมสภาพแวดล้อมต่าง ๆ อย่างเข้มงวดก็ตามก็ยังพบว่าการปลูกต้นไม้ชนิดเดียวกัน 2 ต้น มีการเจริญเติบโตไม่เท่ากันทั้ง ๆ ที่เราได้ต้นไม้ 2 ต้นนี้ที่มาจากเมล็ดพันธ์แหล่งเดียวกัน เตรียมปลูกในดินชนิดเดียวกัน ใส่ปุ๋ยแบบเดียวกัน ปลูกในกระถางแบบเดียวกันและขนาดเท่ากัน วางกระถางต้นไม้ทั้ง 2 นี้ไว้คู่กันบนโต๊ะที่อยู่ตรงกลางเรือนกระจกเพื่อให้ได้รับอุณหภูมิและความชื้นไม่แตกต่างกัน หรืออีกวิธีหนึ่งเราอาจเริ่มต้นจากการเลือกต้นไม้ที่เตรียมปลูกจากเมล็ดพันธ์แหล่งเดียวกันที่มีความสูงของต้นเท่ากัน แล้วนำมาปลูกต่อโดยให้น้ำแบบเดียวกัน และปฏิบัติแบบเดียวกัน แต่หลังจากนั้น 1 เดือน วัดความสูงของต้นไม้ทั้ง 2 ต้นนี้ ก็เป็นไปได้ยากมากที่จะพบว่ามีความสูงเท่ากันพอดี เรามักพบว่าต้นไม้ต้นหนึ่งสูงกว่าอีกต้นหนึ่งอย่างเห็นได้ชัด เป็นความผันแปรที่เกิดขึ้นโดยธรรมชาติ และถ้าทำกับประชากรต้นไม้แบบเดียวกันนี้ก็จะเห็นความผันแปรนี้เช่นเดียวกัน เรามักเรียกความผันแปรนี้ว่า ความคลาดเคลื่อนของการทดลอง (experimental error)

ความคลาดเคลื่อนของการทดลองนี้ทำให้การวิเคราะห์ข้อมูลที่ได้จากการทดลองปลูกต้นไม้เหล่านี้มีความยาก ถ้าความผันแปรของข้อมูลเหล่านี้ไม่มีรูปแบบและถ้าเราไม่ทราบว่ามีความผันแปรมากเพียงใดเกิดขึ้นในการทดลอง แต่ในประชากรส่วนใหญ่มีรูปแบบของความผันแปรนี้อย่างชัดเจน และมักจะมีรูปแบบการแจกแจงเป็นแบบปกติโดยการประมาณ ตัวอย่างเช่น การศึกษาเกี่ยวกับการปลูกข้าวสาลี และข้าวชนิดอื่น ๆ ทำการทดลองโดยการแบ่งพื้นที่ทดลองออกเป็น

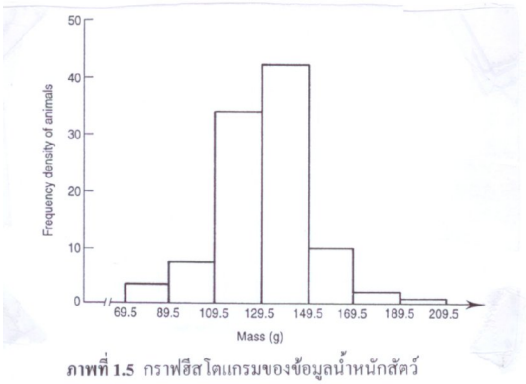
แปลงเล็ก ๆ ขนาด และรูปร่างมาตรฐาน หว่านเมล็ดพันธ์ในแต่ละแปลงด้วยวิธีเดียวกัน ในแต่ละแปลงจะมีต้นข้าวจำนวนมาก เมื่อสิ้นสุดฤดูปลูกเราทำการวัดน้ำหนักผลผลิตข้าวของต้นข้าวแต่ละต้น ซึ่งในทางปฏิบัติทำไม่ได้เพราะในพื้นที่ทดลองทั้งหมดมีต้นข้าวจำนวนมากมาย แต่ถ้าเก็บทีละแปลง เราสามารถทำได้และมักจะใช้เครื่องจักรช่วย การวัดน้ำหนักผลผลิตข้าวในแต่ละแปลงอย่างน้อยก็มีข้อดีอยู่ 2 ข้อ คือ ข้อแรกเราจะได้ตัวเลขที่เล็ก ๆ ที่มีความคลาดเคลื่อนขนาดเล็กซึ่งสัมพันธ์กับขนาดของการวัด (แปลงเล็ก ๆ) ข้อดีข้อสองคือ จากทฤษฎี Central Limit Theorem ทำให้เราคาดได้ว่าการแจกแจงของน้ำหนักผลผลิตข้าวทั้งหมดของแปลงจะเข้าใกล้การแจกแจงแบบปกติ เพราะว่าการเก็บข้อมูลน้ำหนักผลผลิตข้าวทั้งหมดคือผลรวมของผลผลิตข้าวแต่ละต้น จำนวนมาก ดังนั้นในการวิเคราะห์ข้อมูลเราจึงใช้คุณสมบัติของการแจกแจงแบบปกติช่วยในการวิเคราะห์ แต่ถ้าเราไม่สามารถบอกได้ว่าข้อมูลมีการแจกแจงแบบปกติโดยประมาณ เราก็ต้องใช้วิธีการวิเคราะห์แบบไม่อิงพารามิเตอร์ (nonparametric methods) ซึ่งมีอำนาจการทดสอบน้อยกว่า แต่ในความเป็นจริงข้อมูลส่วนมากมีการแจกแจงแบบปกติโดยประมาณ หรือสามารถแปลงข้อมูลให้อยู่ในสเกลการวัดอื่น เพื่อทำให้มีการแจกแจงแบบปกติโดยประมาณ ทำให้ข้อมูลเป็นไปตามข้อตกลงเบื้องต้นเกี่ยวกับความเป็นปกติในการวิเคราะห์ข้อมูลที่ได้จากการทดลองนี้

ตัวอย่างเช่น สัตว์ชนิดหนึ่งมาจากประชากรที่ทำการเพาะพันธ์มาเหมือนกัน 100 ตัว ที่มีเพศเดียวกัน นำมาชั่งน้ำหนักเป็นกรัม สามารถจัดเป็นกลุ่มที่มีช่วงน้ำหนัก 20 กรัม ได้ดังตาราง

ตารางที่ 1.3 น้ำหนัก (กรัม) ของสัตว์ 100 ตัว

น้ำหนัก	70-89	90-109	110-129	130-149	150-169	170-189	190-209
ความถี่	3	7	34	43	10	2	1

จากข้อมูลน้ำหนักเป็นกรัมของสัตว์ 100 ตัวนี้ สามารถเขียนกราฟแบบฮีสโตแกรมได้ดังนี้



ภาพที่ 1.5 กราฟฮีสโตแกรมของข้อมูลน้ำหนักสัตว์

ภาพที่ 1.5 กราฟฮีสโตแกรมของข้อมูลน้ำหนักสัตว์

จากกราฟฮีสโตแกรมของข้อมูลน้ำหนักสัตว์ที่แสดงในภาพที่ 1.5 มีลักษณะโค้งเข้าใกล้แบบสมมาตร โดยมีลักาณะเบ้ขวานิดหน่อย ค่าเฉลี่ยของข้อมูลน้ำหนักเท่ากับ 131.5 และความเบี่ยงเบนมาตรฐานเท่ากับ 20.0 ถ้ากลุ่มตัวอย่างสัตว์ 100 ตัวนี้ถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติที่มี $\mu = 131.5$ และ $\sigma = 20.0$ เราจะสามารถคำนวณหาความถี่คาดหวังได้ เท่ากับ 1.8, 11.8, 32.4, 35.6, 15.5, 2.7 และ 0.2 ตามลำดับ และเมื่อทดสอบสารูปสนิทดี (goodness of fit test) ระหว่างความถี่ที่สังเกตได้กับความถี่คาดหวัง จะสรุปได้ว่ามีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติระหว่างข้อมูลความถี่ที่สังเกตได้กับความถี่คาดหวังที่คำนวณจากประชากรที่มีการแจกแจงแบบปกติที่มี $\mu = 131.5$ และ $\sigma = 20.0$

3. การวิเคราะห์ข้อมูล

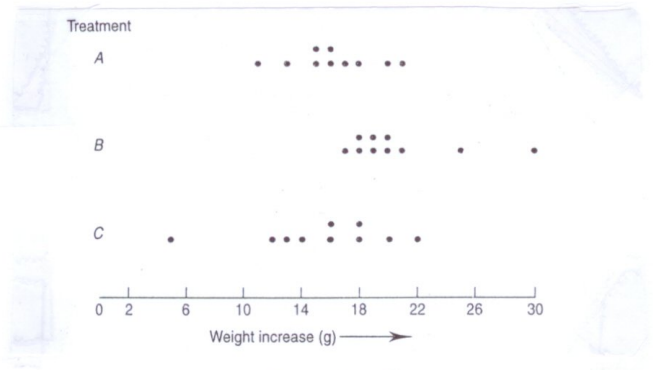
3.1 การวิเคราะห์ข้อมูลขั้นต้น

โดยมากจำนวนข้อมูลที่เก็บมาได้จากการทดลองครั้งหนึ่งมักมีจำนวนน้อยและไม่มีโอกาสที่จะนำมาทดสอบว่ามีการแจกแจงแบบปกติหรือไม่ แต่เราสามารถทำการวิเคราะห์ข้อมูลขั้นต้นได้โดยวิธีการพล็อตกราฟเพื่อตรวจสอบความเป็นปกติของข้อมูลได้ ตัวอย่างเช่น ต้องการเปรียบเทียบอาหารเสริม 3 ชนิด คือ A, B และ C โดยทำการทดลองกับประชากรสัตว์ชนิดหนึ่ง ดำเนินการทดลองโดยแบ่งกลุ่มสัตว์ทดลองออกเป็น 3 กลุ่ม ๆ ละ 10 ตัว ให้แต่ละกลุ่มได้รับอาหารเสริมแต่ละชนิด เก็บข้อมูลน้ำหนักที่เพิ่มขึ้นเมื่อสิ้นสุดการทดลอง ได้ข้อมูลดังตารางที่ 1.4

ตารางที่ 1.4 ข้อมูลน้ำหนักที่เพิ่มขึ้นของสัตว์ทดลอง 3 กลุ่ม ที่ได้รับอาหารเสริมแตกต่างกัน

สัตว์ทดลองตัวที่	1	2	3	4	5	6	7	8	9	10
อาหารเสริม A	16	11	21	15	17	13	20	16	15	18
อาหารเสริม B	18	30	18	17	25	19	20	21	19	20
อาหารเสริม C	14	5	12	16	13	18	16	20	18	22

จากข้อมูลน้ำหนักที่เพิ่มขึ้นของสัตว์ทดลอง 3 กลุ่ม นำมาพล็อตเป็นกราฟจุด (dot – plots) บนสเกลเดียวกัน เพื่อสะดวกในการเปรียบเทียบข้อมูลทั้ง 3 กลุ่มนี้ ในการวิเคราะห์ความแปรปรวนมีข้อตกลงเบื้องต้นว่าข้อมูลแต่ละกลุ่มมีการแจกแจงแบบปกติและมีความแปรปรวนเท่ากัน ได้กราฟจุดดังภาพที่ 1.6



ภาพที่ 1.6 กราฟจุดของข้อมูลน้ำหนักที่เพิ่มขึ้นของอาหารเสริม 3 ชนิด

ภาพที่ 1.6 กราฟจุดของข้อมูลน้ำหนักที่เพิ่มขึ้นของอาหารเสริม 3 ชนิด

จากกราฟจุดพิจารณาได้หรือไม่ว่าข้อมูลแต่ละกลุ่มมีการกระจายแบบสมมาตรอย่างคร่าวๆ ที่จุดค่ากลาง ถ้าไม่มีความสมมาตร ข้อมูลกลุ่มนั้นก็อาจไม่เป็นไปตามข้อตกลงเบื้องต้นเกี่ยวกับความเป็นปกติ

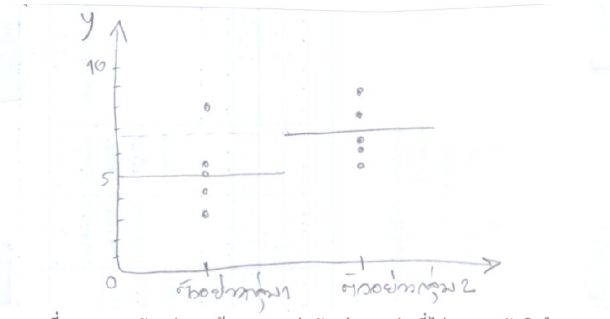
พิจารณาข้อมูลแต่ละกลุ่มว่ามีข้อมูลที่มีค่ากระเด็นออกไปจากกลุ่มหรือไม่ ซึ่งข้อมูลตัวที่กระเด็นออกไปนี้ดูเหมือนว่าไม่อยู่ในประชากรเดียวกันกับข้อมูลตัวอื่น ๆ ถ้ามีเราต้องพิจารณาว่ามีเหตุผลใดที่ทำให้ข้อมูลตัวนั้นมีค่าที่แตกต่างไปจากข้อมูลตัวอื่น ๆ ในกลุ่มเดียวกันนั้น

พิจารณาโดยคร่าวๆ เกี่ยวกับความผันแปรของข้อมูลแต่ละกลุ่มเท่ากันหรือไม่ ถ้าไม่เราก็จะไม่สามารถสรุปได้ว่าข้อมูลทุกกลุ่มมีความแปรปรวนเท่ากัน

จากภาพที่ 1.6 เราจะเห็นได้ว่าข้อมูลในกลุ่มอาหารเสริม A มีการแจกแจงแบบปกติโดยคร่าว ๆ สำหรับข้อมูลในกลุ่ม B มีข้อมูลที่มีค่ามากที่สุด 2 ตัวที่มีค่าห่างออกไปจากข้อมูลตัวอื่น ๆ หรือการแจกแจงของข้อมูลกลุ่มนี้มีความเบ้ สำหรับข้อมูลในกลุ่ม C เราต้องดูว่าสัตว์ทดลองที่มีน้ำหนักเพิ่มน้อยนั้นว่ามีสาเหตุจากการเป็นโรคหรือไม่ซึ่งไม่เกี่ยวข้องกับอาหารเสริมที่ได้รับ ถ้าใช่เราอาจต้องเอาข้อมูลตัวนี้ออกไปจากการวิเคราะห์ แต่ถ้าไม่ใช่เราควรต้องมองหาว่าอะไรทำให้สัตว์ทดลองบางตัวในประชากรนี้น้ำหนักไม่เพิ่มขึ้น ถ้าเป็นไปได้ผู้วิจัยควรติดตามดูสัตว์ทดลองตัวที่ทำให้ข้อมูลห่างออกไปจากประชากร การตัดข้อมูลที่มีค่ามากที่สุด 2 ค่าออกไปจากกลุ่ม B มีผลให้กลุ่ม B มีความผันแปรน้อยกว่ากลุ่ม A หรือ C ถ้าไม่มีเหตุผลที่จะตัดข้อมูล 2 ค่านี้ออกไป ความผันแปรของข้อมูลกลุ่ม B คล้ายกับกลุ่มอื่นๆ

3.2 ตรรกยะของการวิเคราะห์ความแปรปรวน

สำหรับการทดลองที่ศึกษาปัจจัย 1 ปัจจัย ที่มี 2 ระดับ หรือ 2 ทรีทเมนต์ ถ้าเราต้องการเปรียบเทียบค่าเฉลี่ยของ 2 ทรีทเมนต์ ว่ามีความแตกต่างกันหรือไม่ เราอาศัยค่าเฉลี่ยของกลุ่มตัวอย่าง 2 กลุ่ม ที่เป็นอิสระกัน แต่ละกลุ่มมีขนาดตัวอย่างเท่ากันคือ $n_1 = n_2 = 5$ แล้ววัดค่า y ของแต่ละกลุ่มตัวอย่าง นำมาพล็อตจุดได้ดังภาพที่ 1.7

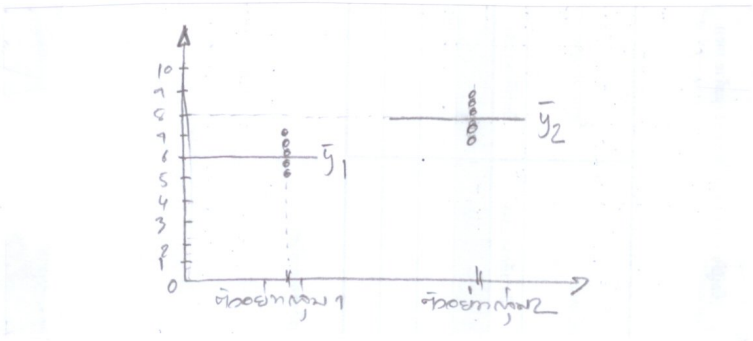


ภาพที่ 1.7 ความผันแปรของข้อมูลของกลุ่มตัวอย่าง 2 กลุ่ม ที่ไม่สามารถตัดสินใจได้ว่าค่าเฉลี่ยของ 2 กลุ่ม แตกต่างกันหรือไม่

ภาพที่ 1.7 ความผันแปรของข้อมูลของกลุ่มตัวอย่าง 2 กลุ่ม ที่ไม่สามารถตัดสินใจได้ว่าค่าเฉลี่ยของ 2 กลุ่ม แตกต่างกันหรือไม่

จากภาพที่ 1.7 ข้อมูล y ของตัวอย่างทั้ง 2 กลุ่ม และค่าเฉลี่ยของตัวอย่างกลุ่มที่ 1 และ 2 เท่ากับ $\bar{y}_1$ และ $\bar{y}_2$ ตามลำดับ เราไม่สามารถบอกได้ว่ามีความแตกต่างกัน

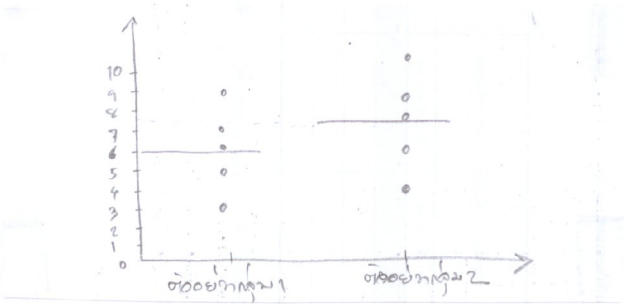
ในภาพที่1.8 (ก.) ค่าของ y ของกลุ่มตัวอย่างทั้ง 2 กลุ่ม มีค่าแยกกันเป็น 2 กลุ่ม ชัดเจน เราอาจบอกได้ว่าค่าเฉลี่ยของตัวอย่างทั้ง 2 กลุ่มนี้แตกต่างกัน



ภาพ 1.8 (ก.) ความผันแปรของข้อมูลของกลุ่มตัวอย่าง 2 กลุ่มที่แยกเป็น 2 กลุ่มชัดเจน

ภาพ 1.8 (ก.) ความผันแปรของข้อมูลของกลุ่มตัวอย่าง 2 กลุ่มที่แยกเป็น 2 กลุ่มชัดเจน

ในภาพที่1.8 (ข.) ค่าของ y ของกลุ่มตัวอย่าง 2 กลุ่ม ทับซ้อนกันส่วนใหญ่เห็นได้ชัดว่าค่าเฉลี่ยของตัวอย่างทั้ง 2 กลุ่ม ไม่แตกต่างกันหรือแตกต่างกันน้อยมาก



ภาพ 1.8 (ข.) ความผันแปรของข้อมูลของกลุ่มตัวอย่าง 2 กลุ่ม ที่ทับซ้อนกันส่วนใหญ่

ภาพ 1.8 (ข.) ความผันแปรของข้อมูลของกลุ่มตัวอย่าง 2 กลุ่ม ที่ทับซ้อนกันส่วนใหญ่

จากความผันแปรของข้อมูลทั้ง 3 กรณีข้างต้น การที่จะบอกว่าค่าเฉลี่ยของตัวอย่าง 2 กลุ่ม แตกต่างกันหรือไม่ พิจารณาโดยการเปรียบเทียบระยะห่างของค่าเฉลี่ย 2 ค่า กับความผันแปรภายในค่า y ของกลุ่มตัวอย่างแต่ละกลุ่ม

เช่นในภาพที่ 1.8 (ก.) ความแตกต่างระหว่างค่าเฉลี่ยของตัวอย่าง 2 กลุ่ม มีค่ามากเมื่อเทียบกับความผันแปรของข้อมูลภายในกลุ่มตัวอย่าง เราจะสรุปว่าค่าเฉลี่ยของประชากร 2 กลุ่ม แตกต่างกัน ในทางกลับกัน ในภาพที่ 1.8 (ข.) ความแตกต่างระหว่างค่าเฉลี่ยของตัวอย่าง 2 กลุ่ม มี

ค่าน้อย เมื่อเทียบกับความผันแปรของข้อมูลภายในกลุ่มตัวอย่าง แสดงให้เห็นว่ามีหลักฐานน้อยมากที่จะสรุปว่าค่าเฉลี่ย 2 ค่า มีความแตกต่างกันอย่างมีนัยสำคัญ

ค่าที่วัดความผันแปรของค่าสังเกตภายในกลุ่มตัวอย่างคือ s^2 ที่คำนวณจากการทดสอบ t-test สำหรับกลุ่มตัวอย่าง 2 กลุ่มที่เป็นอิสระกัน สูตรคือ

ความผันแปรภายในกลุ่ม ; $$s^2 = \frac{\sum_{i=1}^{n_1} (y_{i1} - \bar{y}_1)^2 + \sum_{i=1}^{n2} (y_{i2} - \bar{y}_2)^2}{n_1 + n_2 - 2}$$

$$= \frac{SS_E}{n_1 + n_2 - 2}$$

เมื่อ y_{i1} แทน ค่าสังเกตตัวที่ i ของตัวอย่างกลุ่ม 1

y_{i2} แทน ค่าสังเกตตัวที่ i ของตัวอย่างกลุ่ม 2

เรามักแทนตัวเศษของ s^2 ด้วย SSE เรียกว่า sum of squared errors ซึ่งเป็นค่าของความผันแปรที่อธิบายไม่ได้ของความแตกต่างระหว่างค่าเฉลี่ยของตัวอย่าง 2 กลุ่ม ส่วนค่าของความผันแปรระหว่างกลุ่มตัวอย่าง คำนวณได้จากผลรวมกำลังสองของความเบี่ยงเบนของค่าเฉลี่ยของกลุ่มตัวอย่างแต่ละกลุ่มที่เบี่ยงเบนไปจากค่าเฉลี่ยของค่าสังเกตทั้ง 10 ตัว คือ $\bar{y}$ แล้วหารด้วยจำนวนกลุ่มตัวอย่างลบ 1 สูตรคือ

ความผันแปรระหว่างกลุ่ม ; $$\frac{n_1 (\bar{y}_1 - \bar{y})^2 + n_2 (\bar{y}_2 - \bar{y})^2}{2 - 1} = \frac{SS_{Tr}}{1}$$

เรามักแทนตัวเศษด้วย SSTr เรียกว่า sum of square for treatments ซึ่งเป็นค่าของความผันแปรที่อธิบายได้ของความแตกต่างระหว่างค่าเฉลี่ยของตัวอย่างของทรีทเมนต์ 2 กลุ่ม

และผลรวมของ SSE และ SSTr คือ SST เรียกว่า sum of square for total

$$SST = \sum (y_i - \bar{y})^2$$

สถิติทดสอบที่ใช้ทดสอบความแตกต่างของค่าเฉลี่ยของทรีทเมนต์ต่าง ๆ คือ

$$F = \frac{\text{ความผันแปรระหว่างกลุ่มตัวอย่าง}}{\text{ความผันแปรภายในกลุ่มตัวอย่าง}}$$

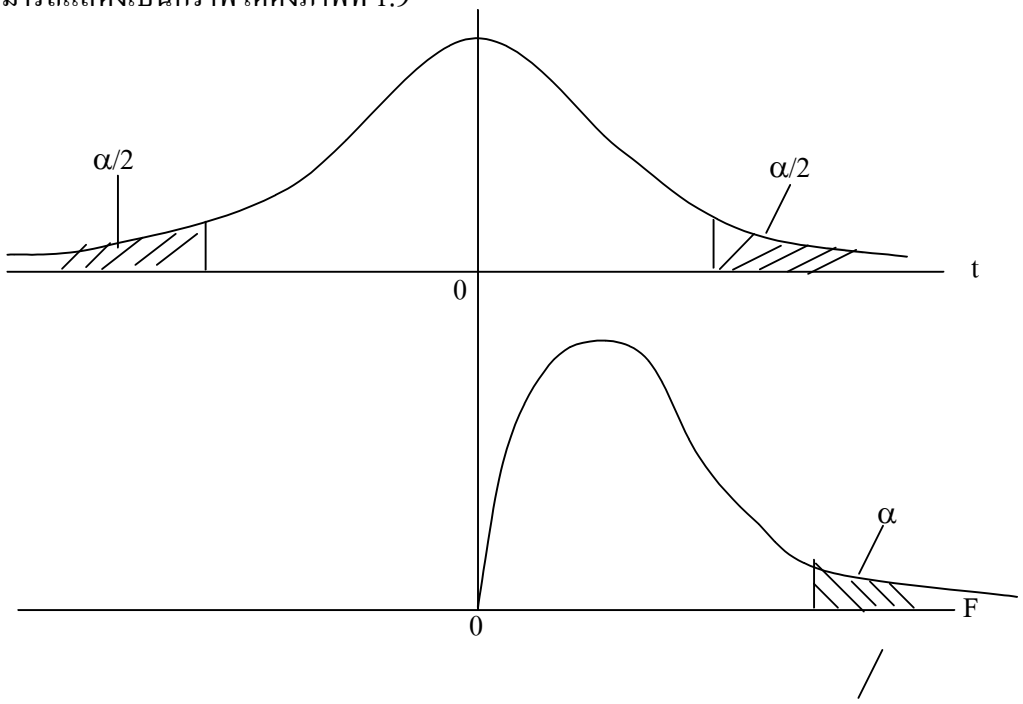
$$= \frac{SS_{Tr} / 1}{SS_E / (n_1 + n_2 - 2)}$$

ซึ่งมีการแจกแจงแบบ F ที่มีจำนวนชั้นอิสระ $df_1 = 1$ และ $df_2 = n_1 + n_2 - 2$ และเราเรียกวิธีการทางสถิตินี้ว่า การวิเคราะห์ความแปรปรวน

การทดสอบ F แบบทางเดียวที่มีจำนวนชั้นอิสระเท่ากับ 1 และ n เทียบได้กับการทดสอบ t – test แบบสองทางที่มีจำนวนชั้นอิสระ เท่ากับ n การทดสอบ t-test นี้ไม่สามารถบอกทิศทางของความแตกต่างระหว่างค่าเฉลี่ยของทรีทเมนต์ 2 ทรีทเมนต์ ในสมมติฐานแย้งได้ ซึ่งเหมือนกับการทดสอบ F แบบทางเดียว ที่ผลของการทดสอบความแตกต่างระหว่างทรีทเมนต์ต่าง ๆ ก็ไม่ได้บอกทิศทาง การทดสอบทั้ง 2 ชนิดนี้มีความเกี่ยวข้องกันดังนี้คือ

$$t^2 = F$$

และสามารถแสดงเป็นกราฟได้ดังภาพที่ 1.9



ภาพที่ 1.9 ความเกี่ยวข้องของการทดสอบ t และการทดสอบ F

3.3 การสรุปอ้างอิงทางสถิติ

การใช้วิธีการทางสถิติเข้ามาช่วยในการแปลความหมายของผลการทดลองนั้น ค่าที่นำมาใช้ทางสถิติเป็นเพียงค่าประมาณ ไม่ใช่ค่าจริง ดังนั้นเราจึงต้องมีขอบเขตของค่าประมาณนี้เป็นค่าต่ำสุดและค่าสูงสุดที่ครอบคลุมค่าที่ประมาณได้ โดยอาศัยความน่าจะเป็นในการคำนวณความน่าจะเป็นนี้ผู้วิจัยเป็นผู้ตัดสินใจเลือกเอง เช่น .95 หมายความว่า ขอบเขตที่คำนวณได้นั้นมี

โอกาสผิดพลาดเพียง .05 หรือ 5% เท่านั้น ขอบเขตที่คำนวณได้จากความน่าจะเป็นนี้เรียกว่า ช่วงความเชื่อมั่น ตัวอย่างเช่น 95% ของช่วงความเชื่อมั่นของค่าเฉลี่ย (μ) หาได้จากสูตร

$$\bar{y} \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

เมื่อ 1.96 ได้มาจากตารางการแจกแจงปกติ

n คือ ขนาดของตัวอย่าง

σ คือ ส่วนเบี่ยงเบนมาตรฐานของประชากร

ถ้าคำนวณค่าเฉลี่ยของตัวอย่าง 100 กลุ่ม เราคาดว่าผลจากการคำนวณ 95 ครั้งใน 100 ครั้ง จะรวมค่า μ อยู่ด้วย และถ้าคำนวณหาค่าเฉลี่ยจากตัวอย่างเพียงกลุ่มเดียวที่มีขนาด n เราสามารถบอกได้ว่า ด้วยความเชื่อมั่น 95% ขอบเขตที่คำนวณได้จะครอบคลุม μ และถ้าไม่ทราบ σ เราจะใช้การแจกแจง t ในการคำนวณช่วงความเชื่อมั่น จากสูตร

$$\bar{y} \pm t_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

เมื่อ s คือ ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง

t มีจำนวนชั้นอิสระ n-1

ขั้นตอนต่อไปที่ต้องพิจารณาคือ การทดสอบสมมติฐานทางการวิจัย ที่ผู้วิจัยต้องการหาคำตอบ โดยอาศัยข้อมูลที่ได้จากการทดลอง แล้วสรุปว่าสมมติฐานเป็นจริงหรือไม่ เราใช้วิธีการทางสถิติที่เรียกว่า การทดสอบนัยสำคัญ (test of significane) เพื่อใช้เป็นกฎในการตัดสินใจว่าจะยอมรับหรือปฏิเสธสมมติฐานทางสถิติ ซึ่งกฎการตัดสินใจนี้เป็นไปได้ 2 ทาง คือ (1) ปฏิเสธหรือไม่ยอมรับสมมติฐานที่เป็นจริงด้วยความน่าจะเป็นหรือโอกาสที่ตัดสินใจผิดพลาดตามที่ผู้วิจัยเป็นผู้กำหนด (2) ปฏิเสธสมมติฐานที่ไม่จริง