

แผนบริหารการสอนประจำบทที่ 7

หัวข้อเนื้อหาประจำบท

1. ความสำคัญของแผนการทดลอง
2. แผนการทดลองที่ให้ข้อมูลจากการทดลองที่สามารถนำไปวิเคราะห์ผลทางสถิติได้
3. แผนการทดลองแบบสุ่มตลอด (สุ่มอย่างสมบูรณ์)
4. ข้อดีของแผนการทดลองแบบสุ่มอย่างสมบูรณ์
5. ข้อเสียของแผนการทดลองแบบสุ่มอย่างสมบูรณ์
6. วิธีสุ่ม (Randomized Procedure) ตามแบบของ CRD
7. การบันทึกข้อมูล
8. การวิเคราะห์ข้อมูล (Data analysis)
9. แนวคิดเชิงทฤษฎี ของการวิเคราะห์ความแปรปรวน

วัตถุประสงค์เชิงพฤติกรรม

เพื่อให้ให้นักศึกษาสามารถ

1. บอกความสำคัญของแผนการทดลองได้
2. บอกแผนการทดลองแบบสุ่มตลอด ได้
3. บอกข้อดี ข้อเสียของแผนการทดลองแบบสุ่มอย่างสมบูรณ์ได้

วิธีการสอนและการจัดกิจกรรมการเรียนการสอนประจำบท

1. ศึกษาเอกสารประกอบการสอน เรื่อง แผนการทดลองแบบสุ่มสมบูรณ์
2. ศึกษาค้นคว้าเพิ่มเติมเรื่อง โจทย์วิจัย
3. ตั้งโจทย์วิจัยของตนเอง

สื่อการเรียนการสอน

1. เอกสารประกอบการสอน
2. สไลด์ Microsoft Power Point เรื่อง แผนการทดลองแบบสุ่มสมบูรณ์

การวัดผลและการประเมินผล

1. สังเกตจากความสนใจ ความตั้งใจเรียน
2. ตรวจสอบคำตอบจากการค้นคว้าเพิ่มเติมเรื่องโจทย์วิจัย
3. ตรวจสอบคำตอบจากโจทย์วิจัยของตนเอง

บทที่ 7

แผนการทดลองแบบสุ่มสมบูรณ์

แผนการทดลอง(Experimental design) ที่ใช้ในการทดลองทางชีววิทยา (Biometrical approach) มีหลายแบบตามรายละเอียดของการทดลองแต่ละประเภท ซึ่งอาจเป็นทั้งแผนการทดลองอย่างง่ายและแผนการทดลองที่ซับซ้อน แต่แผนการทดลองที่จัดว่าเป็นแผนการทดลองขั้นพื้นฐานที่นิยมใช้โดยทั่วไป มีอยู่ 3 แผนการทดลองได้แก่

1. แผนการทดลองแบบ Completely randomized design (CRD)
2. แผนการทดลอง Randomized complete block design (RCBD)
3. แผนการทดลองแบบ Latin square design (LSD)

ซึ่งเหตุผลที่เรียกแผนการทดลองทั้ง 3 นี้ว่า เป็น Basic design ก็เพราะว่า แผนการทดลองทั้ง 3 แผนนี้จะใช้เป็นพื้นฐานของแผนการทดลองขั้นสูงอื่นๆที่จำเป็นต้องอาศัยหลักและวิธีการของแผนการทดลองทั้ง 3 แผนนี้เป็นพื้นฐาน ดังนั้นแผนการทดลองทั้ง 3 นี้จึงมีความสำคัญต่อผู้เรียนชั้นนี้เป็นอย่างมาก ทั้งนี้เพราะเมื่อได้ศึกษาขั้นสูงขึ้นก็สามารถใช้ความรู้พื้นฐานจากแผนการทดลองขั้นต้นนี้ได้เป็นอย่างดี

ความสำคัญของแผนการทดลอง

ในการค้นหาความจริงที่มีอยู่แล้วในธรรมชาติ โดยอาศัยวิธีการวิจัยที่เป็นการวิจัยเชิงทดลอง (Experimental research) นั้น แผนการทดลองที่จะเลือกมาใช้ให้เหมาะสมกับคุณสมบัติของหน่วยทดลอง (Experimental unit) ที่มีหรือที่หาได้ในสภาพนั้นๆ นับว่ามีความสำคัญอย่างมากต่อความถูกต้องในการสรุปผลการทดสอบสมมุติฐาน การวิจัยหรือการทดลองใดที่เลือกใช้แผนการทดลองที่เหมาะสมกับหน่วยทดลองที่ใช้ทดลองมากกว่าก็จะให้ความแม่นยำในการทดสอบที่สูงกว่า เนื่องจากมีความคลาดเคลื่อนทดลอง (experimental error) ที่ต่ำกว่า ทั้งนี้เพราะงานวิจัยส่วนใหญ่จะใช้วิธีวิเคราะห์ความแปรปรวน (analysis of variance: ANOVA) เป็นเครื่องมือในการบอกถึง **ความมีหรือความไม่มีอิทธิพลของ Treatment** ซึ่งถ้าหากผู้วิจัยสามารถแยก(analyzed) สาเหตุของความผันแปรที่เกิดขึ้นในการทดลองนั้นๆได้อย่างถูกต้องตรงตามความผันแปรของหน่วยทดลองที่ใช้จริงได้มากกว่า ก็จะทำให้ค่าประมาณของความผันแปรอันเนื่องมาจากอิทธิพลของ Treatment หรือค่า Treatment mean of square: Tr MS ที่สูงกว่า ในขณะที่เดียวกันก็จะให้ค่าประมาณของความผันแปร

แปรอันเนื่องมาจากอิทธิพลของความคลาดเคลื่อน(error mean of square: Error MS) ที่ต่ำกว่า ซึ่งการทดลองใดที่ค่า Tr MS สูง และมีค่า Error MS ที่ต่ำ ก็จะมีโอกาสมากกว่าในการตรวจสอบพบอิทธิพลของ Treatment ได้ โดยใช้ระดับความเชื่อมั่น (Level of confidence) ในการสรุปผลการทดสอบที่สูงกว่าได้ ซึ่งทั้งความแม่นยำและความเชื่อมั่นในการสรุปผลการทดลองที่สูงเป็นสิ่งที่ผู้วิจัยทุกคนต้องการ

ดังนั้นแผนการทดลองที่เหมาะสมกับธรรมชาติของหน่วยทดลองที่ใช้ จึงมีความสำคัญเป็นอย่างมากต่อการได้มาซึ่งข้อสรุปที่เป็นจริง และการมีแผนการทดลองแบบต่างๆหลายๆแบบไว้ใช้ในการทดลองก็เพื่อให้สามารถเลือกใช้แผนการทดลองที่เหมาะสมกับธรรมชาติของหน่วยทดลองที่มี ซึ่งแผนการทดลองที่ดีจะต้องให้ความคลาดเคลื่อนทดลอง(experimental error) ที่ต่ำที่สุด (minimized error) แต่ให้ความมั่นใจในการ สรุปผลได้สูงที่สุด คือมีโอกาสของการสรุปผลผิดพลาดต่ำที่สุดนั่นเอง

ฉะนั้นเหตุผลสำคัญของการเลือกใช้แผนการทดลองที่ต่างกันจึงอยู่ที่ความผันแปรตามธรรมชาติของหน่วยทดลอง (Experimental unit) ก่อนการให้ Treatment กล่าวคือถ้าหากว่าหน่วยทดลองมีความสม่ำเสมอ กันก่อนการให้ Treatment ก็ใช้แผนการทดลองแบบหนึ่ง แต่ถ้าก่อนการให้ Treatment ผู้วิจัยทราบว่ามี ความผันแปร อันเนื่องมาจากปัจจัยบางอย่าง ก็สามารถจัดหน่วยทดลองออกเป็นกลุ่มต่างๆเสียก่อนตามสาเหตุของความผันแปรนั้นๆ ซึ่งแผนการทดลองที่เหมาะสมกับธรรมชาติของหน่วยทดลองแบบนี้ก็ควรจะเป็นอีกแผนหนึ่ง หรือกรณีที่ก่อนการให้ Treatment แก่หน่วยทดลองผู้ทดลองทราบว่ามีความแตกต่างกันใน 2 ทิศทางหรือ 2 สาเหตุอยู่ก่อนแล้ว จึงจำเป็นที่จะต้องจัดกลุ่มของหน่วยทดลองเสียก่อนทั้งใน 2 ทิศทางตามสาเหตุของความผันแปรนั้นๆ ก่อนที่จะให้ได้รับ Treatment ก็จะใช้แผนการทดลองอีกแผนหนึ่งที่เหมาะสมกับธรรมชาติของหน่วยทดลองแบบนี้ซึ่งจะกล่าวรายละเอียดในลำดับต่อไป

อย่างไรก็ตามสำหรับนักวิจัยที่เริ่มต้นใหม่อาจจะขาดความมั่นใจในการที่จะสรุปว่าหน่วยทดลองที่มีหรือหน่วยทดลองที่นำมาใช้ในการทดลองนั้นมีความสม่ำเสมอหรือไม่ก่อนการทดลอง ทั้งนี้เพื่อให้สามารถเลือกใช้แผนการทดลองที่เหมาะสมกับหน่วยทดลองที่ใช้ทดลองได้ ดังนั้นความสม่ำเสมอ(uniformity) ของหน่วยทดลองจึงจำเป็นต้องมีหลักพิจารณาที่ดี ซึ่งโดยทั่วไปแล้วการที่จะพิจารณาว่าหน่วยทดลองนั้นๆสม่ำเสมอหรือไม่นั้น มีหลักพิจารณาสำคัญอยู่ตรงที่ว่าผู้วิจัยต้องรู้ความเป็นจริงในสภาพธรรมชาติของสิ่งและลักษณะที่ศึกษานั้นๆว่า การแสดงออกของลักษณะที่ศึกษานั้นๆมีปัจจัยอะไรมากระทบบ้าง และถ้าหากปัจจัยที่มากกระทบต่อลักษณะนั้นๆก่อนการให้

Treatment ที่มีต่อหน่วยทดลองทุกๆหน่วยเหมือนกันแล้วลักษณะที่ปรากฏออกมาของหน่วยทดลองเหมือนกัน ก็ถือว่าหน่วยทดลองนั้นสม่ำเสมอ

ดังนั้นโดยสรุปแล้วในการทดลอง การที่ผู้ทดลองจะเลือกใช้แผนการทดลองแบบใดก็อยู่ที่ธรรมชาติของความผันแปรของหน่วยทดลองที่มีหรือที่สามารถหามาใช้ได้ในการทดลองนั้น ซึ่งจะต้องเข้าใจถึงธรรมชาติของหน่วยทดลองที่มีว่าเป็นอย่างไร แล้วจึงเลือกใช้แผนการทดลองที่เหมาะสมกับธรรมชาติของหน่วยทดลองเหล่านั้น หลังจากนั้นจึงดำเนินการทดลองตามแผนการทดลองตามขั้นตอนต่างๆจนเสร็จสิ้นกระบวนการทดลอง และนำผลที่ได้จากการทดลองที่เป็นตัวเลขใช้เป็นหลักฐานเพื่อนำไปใช้ในการพิสูจน์สมมติฐานที่กำหนดไว้ว่า สมมติฐานที่ตั้งไว้นั้นเป็นจริงหรือไม่ต่อไป

แผนการทดลองที่ให้ข้อมูลจากการทดลองที่สามารถนำไปวิเคราะห์ผลทางสถิติได้ (Valid experimental design)

ตามที่ได้ทราบมาแล้วว่า ในการทดลองนั้นจะเกิดความคลาดเคลื่อนจากการทดลอง (Experimental error) ขึ้นได้เสมอๆ ซึ่งนักวิจัยที่ติดจะต้องหาทางควบคุมความคลาดเคลื่อนดังกล่าวให้น้อยที่สุด โดยที่วิธีการควบคุมความคลาดเคลื่อนนั้นอาจทำได้ทั้งวิธีการควบคุมโดยตรง ซึ่งได้แก่การมีเทคนิคที่ละเอียดอ่อนในการทดลอง(refining technique) ได้แก่การปฏิบัติต่อหน่วยทดลองที่เหมือนกันในเวลาเดียวกัน รวมทั้งการใช้เทคนิคที่เหมาะสมในการทดลองนั้นๆ นอกจากนี้อาจจะใช้วิธีการควบคุมโดยวิธีอ้อม ซึ่งได้แก่การเลือกใช้แผนการทดลองที่เหมาะสมกับสภาพในธรรมชาติของหน่วยทดลองที่มี ซึ่งแผนการทดลองที่จะให้ข้อมูลที่มีคุณภาพและสามารถใช้วิเคราะห์ทางสถิติได้จะต้องมีลักษณะที่สำคัญ (Essential requirement) ดังต่อไปนี้ คือ

ต้องมีการทำซ้ำ (Replication) เพื่อให้สามารถประมาณค่า Error ได้

จะต้องมีการสุ่ม (Randomization) เพื่อขจัดความลำเอียง (bias)

จะต้องมีการควบคุมความคลาดเคลื่อนในการทดลอง (Error Control) ซึ่งสามารถทำได้โดย

- เลือกใช้แผนการทดลอง (experimental design) ให้เหมาะสม
- ใช้ตัวแปรร่วม (Covariate หรือ Concomitant Variable) ซึ่งเรียกการวิเคราะห์ โดยวิธีนี้ว่า การวิเคราะห์ความแปรปรวนร่วม (Analysis of Covariance: ANCOVA)
- เลือกขนาดของหน่วยทดลองให้เหมาะสม

ดังนั้นในทุกๆ แผนการทดลองจะต้องประกอบด้วยลักษณะที่สำคัญทั้ง 3 ประการนี้เป็นสำคัญ ยกเว้นแต่แผนการทดลองบางประเภทที่พิเศษออกไปที่ไม่จำเป็นต้องมี replication ก็ได้

ซึ่งเป็นลักษณะพิเศษของแผนการทดลองนั้นๆ แต่แผนการทดลองทุกๆ ไป จะต้องประกอบด้วยลักษณะ 3 ประการนี้ที่สำคัญ

แผนการทดลองแบบสุ่มตลอด (สุ่มอย่างสมบูรณ์) (Completely Randomized Design: CRD)

CRD เป็นแผนการทดลองที่จัดเป็น Basic design และเป็นแผนการทดลองที่ง่ายที่สุด โดยที่ Treatment ที่จะให้กับหน่วยทดลองทั้งหมดที่ใช้ในการทดลองเป็นไปโดยสุ่มอย่างสมบูรณ์ คือหน่วยทดลองทุกหน่วยมีโอกาสเท่าๆกันที่จะได้รับ Treatment ใดๆ ก็ได้ ดังนั้น แผนการทดลองแบบนี้จึงมีความเหมาะสมกับการทดลองในกรณีที่หน่วยทดลองทั้งหมดที่ใช้ในการทดลองมีความสม่ำเสมอทั้งนี้เพราะเมื่อหน่วยทดลองแต่ละหน่วยสม่ำเสมอจริง ไม่ว่าจะสุ่มให้หน่วยทดลองนั้นๆ ให้ได้รับ Treatment ใดๆ ก็ถือว่าหน่วยทดลองนั้นๆ มีความสม่ำเสมอ ดังนั้นความแตกต่างที่จะเกิดขึ้นจากการทดลองลงก็จะเป็นผลมาจากอิทธิพลของ Treatment และความคลาดเคลื่อนในการทดลอง (experimental error) เท่านั้น ซึ่งสามารถทดสอบต่อไปได้ว่า ความแปรปรวน(variance) ของทั้ง 2 ส่วนนี้แตกต่างกันหรือไม่ และสามารถใช้ค่าสัดส่วนของค่าทั้งสองนี้เป็นเกณฑ์ในการบอกถึงความแตกต่างระหว่างอิทธิพลของ Treatment ได้ว่า มีจริงหรือไม่ โดยอาศัยวิธีวิเคราะห์ความแปรปรวน (Analysis of variance: ANOVA) ต่อไป

ข้อดีของแผนการทดลองแบบสุ่มอย่างสมบูรณ์

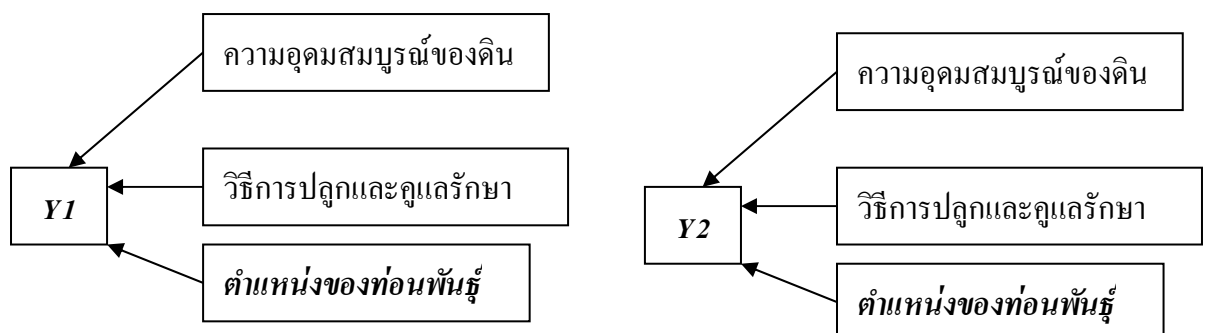
1. ใช้กับการทดลองที่มี Treatment จำนวนมากๆ ได้ดี โดยที่ในแต่ละ Treatment อาจจะมีจำนวนซ้ำไม่เท่ากันก็ได้ แต่โดยทั่วไปแล้วเพื่อให้การคำนวณเป็นไปโดยสะดวกจึงนิยมกำหนดให้จำนวนซ้ำเท่ากัน
2. เป็นแผนการทดลองที่ทำได้ง่าย และเมื่อมีข้อมูลสูญหายหรือไม่สามารถบันทึกได้ (missing data) ก็ไม่มีผลกระทบต่อการศึกษามากนัก
3. การวิเคราะห์ทางสถิติทำได้ง่าย และให้ค่า degree of freedom ของ Error (Error df) ที่มากกว่าแผนการทดลองอื่นๆ ดังนั้น ถ้าหากจะพิจารณาความแม่นยำในการทดลองโดยใช้จำนวน df of error เป็นเกณฑ์แล้วพบว่า แผนการทดลองแบบ CRD เป็นแผนการทดลองให้ความแม่นยำมากที่สุด (เนื่องจากการที่มีจำนวน df of error สูงทำให้ค่า error mean of square ต่ำ และการมี error mean of square ต่ำ ก็แปลว่ามีความแม่นยำสูงนั่นเอง) แต่จริงๆ แล้วเราไม่ได้ใช้เฉพาะมีจำนวน df of error เท่านั้นในการพิจารณาแต่เรายังมีข้อพิจารณาอื่นๆ อีกหลายอย่างจึงสามารถบอกได้ว่าแผนการทดลองแบบใดมีความแม่นยำกว่า

ข้อเสียของแผนการทดลองแบบสุ่มอย่างสมบูรณ์

1. เป็นแผนการทดลองที่ควบคุมความคลาดเคลื่อน (Error control) ได้น้อยกว่าแผนการทดลองอื่นๆ
2. แผนการทดลองแบบนี้จะใช้ได้ดีก็ต่อเมื่อหน่วยทดลองมีความสม่ำเสมอ (homogeneous) กันจริงๆ ซึ่งเกี่ยวกับความสม่ำเสมอของหน่วยทดลองที่ใช้ในการทดลองว่ามีความสม่ำเสมอหรือไม่นั้น ต้องพิจารณาว่าปัจจัยต่างๆ ที่มากระทบกับลักษณะที่เราศึกษานั้นเหมือนกันหรือสม่ำเสมอหรือไม่ ซึ่งถ้าหากว่าปัจจัยที่มากระทบเหล่านั้นเหมือนกัน หรือสม่ำเสมอ เราก็ถือว่าหน่วยทดลองนั้นมีความสม่ำเสมอ นั่นคือ การที่จะพิจารณาว่าหน่วยทดลองที่นำมาใช้นั้นมีความสม่ำเสมอหรือไม่ ต้องทราบก่อนว่า

1. อะไรคือลักษณะที่เราศึกษา
2. ลักษณะเหล่านั้นมีปัจจัยอะไรมากระทบบ้าง
3. ปัจจัยที่มากระทบต่อลักษณะนั้นๆ มีความสม่ำเสมอหรือไม่ เช่น ในการทดลองทางด้านพืชอาหารสัตว์ เพื่อต้องการทดสอบอิทธิพลของสายพันธุ์ที่มีต่อผลผลิตของหญ้าเนเปียร์

ซึ่งในกรณีนี้เราพอจะรู้ว่าลักษณะที่เราศึกษาคือผลผลิตของหญ้าเนเปียร์ (yield=Y) และเราพอจะรู้ว่าปัจจัยที่จะมากระทบต่อผลผลิต ที่ได้จากแต่ละหน่วยทดลองได้แก่อะไรบ้างซึ่งได้แก่

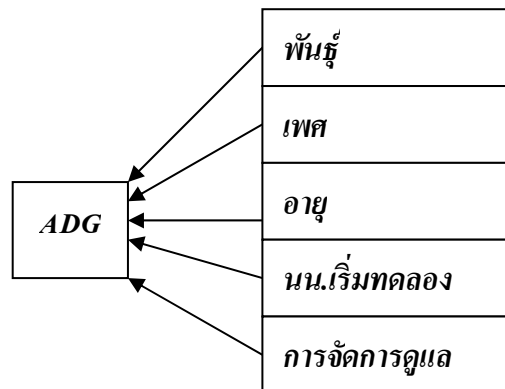


นั่นคือ ผู้วิจัยจะต้องรู้ว่า ดินหรือแปลงที่ใช้ปลูกหญ้าเนเปียร์ทั้ง 2 กลุ่ม มีความสม่ำเสมอหรือไม่ ทั้งในด้านชนิดของดิน(คุณสมบัติทางเคมี ฟิสิกส์ ซึ่งได้แก่ ชนิดของเนื้อดิน ความอุดมสมบูรณ์ของดิน

ความเป็น กรด-ด่าง โครงสร้าง ความพรุน ความสามารถในการการดูดซับน้ำและคุณสมบัติอื่นๆที่เกี่ยวข้องกับการให้ผลผลิต(Y) ของหญ้าเนเปียร์ของดินรวมทั้งการจัดการต่างๆ ที่ให้แก่ Treatment แต่ละ Treatment ว่ากระทำเหมือนกันหรือไม่ เช่น การรดน้ำ พรวนดิน ใส่ปุ๋ย รวมทั้งวิธีการปลูก เช่นปลูกตามแนวตะวัน หรือวางแสงตะวัน ตลอดจนแหล่งของเมล็ดพันธุ์ที่ใช้ปลูก ว่ามาจากแหล่งเดียวกันหรือไม่ เป็นต้น ซึ่งถ้าหากว่าปัจจัยต่างๆ เหล่านี้ที่มีผลกระทบต่อการให้ผลผลิต (Y) ที่เหมือนกันหรือไม่ เป็นต้น ซึ่งถ้าหากว่าปัจจัยต่างๆ เหล่านี้ที่มีผลกระทบต่อการให้ผลผลิต (Y) ที่เหมือนกันในแต่ละหน่วยทดลอง ก็ถือว่าหน่วยทดลองที่ใช้ในการทดลองนั้นสม่ำเสมอ

หรือในการทดลองทางสัตว์เช่นในโค ผู้วิจัยต้องการจะศึกษาอิทธิพลของชนิดของแป้ง 3 ชนิดที่มีการเจริญเติบโต ของสัตว์ ซึ่งในกรณีนี้เราต้องรู้ว่า

1. ลักษณะที่ต้องศึกษาคือการเจริญเติบโตของโคซึ่งอาจจะวัดออกมาในรูปของน้ำหนักเพิ่มเฉลี่ย/วัน (ADG)
2. ปัจจัยที่มีผลต่อการเจริญเติบโตของโคได้แก่



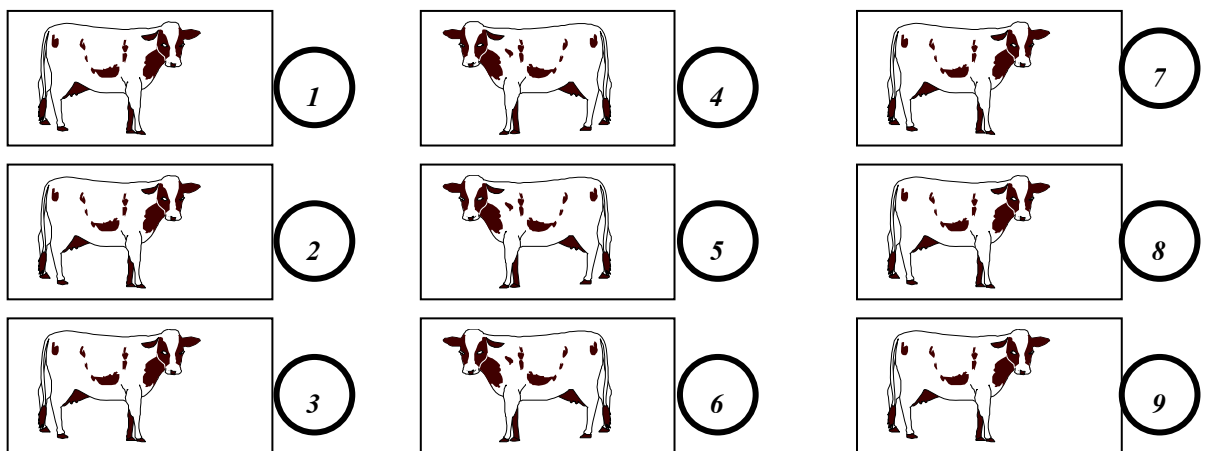
ซึ่งถ้าหากว่าปัจจัยต่างๆ เหล่านี้ที่มากกระทบเหมือนกัน ก็ถือว่าโคทุกตัวมีความสม่ำเสมอ ก็สามารถใช้แผนการทดลองแบบCRD ได้

โดยทั่วไปแล้วนักวิจัยในแต่ละสาขาจะเป็นผู้เข้าใจว่าหน่วยทดลองที่นำมาใช้ในการทดลองนั้น สม่ำเสมอหรือไม่ เพราะนักวิจัยได้คลุกคลีอยู่กับสิ่งเหล่านั้นอยู่ตลอดเวลา ดังนั้นในขั้นนี้พอจะโดยสรุปได้ว่า CRD เป็นแผนการทดลองที่เลือกใช้เมื่อหน่วยทดลองมีความสม่ำเสมอก่อนการให้ Treatment ทั้งนี้เพราะว่าหน่วยทดลองที่สม่ำเสมอจะมี inherent variability ที่น้อยที่สุดโดยหวังให้ความแตกต่างที่เกิดขึ้นนั้นมีผลมาจาก Treatment จริงๆมากที่สุด

วิธีสุ่ม (Randomized Procedure) ตามแบบของ CRD

โดยใช้วิธีการสุ่มตลอด (Completely at random) คือหน่วยทดลองทุกๆ หน่วยทดลอง มีโอกาสเท่าๆ กันที่จะได้รับ Treatment ใดๆ ก็ได้และการที่หน่วยทดลองใดจะได้รับ Treatment ใดในการทดลองนั้นๆ จะต้องเป็นไปโดยการสุ่มอย่างสมบูรณ์ซึ่งอาจจะใช้วิธีต่อไปนี้คือ

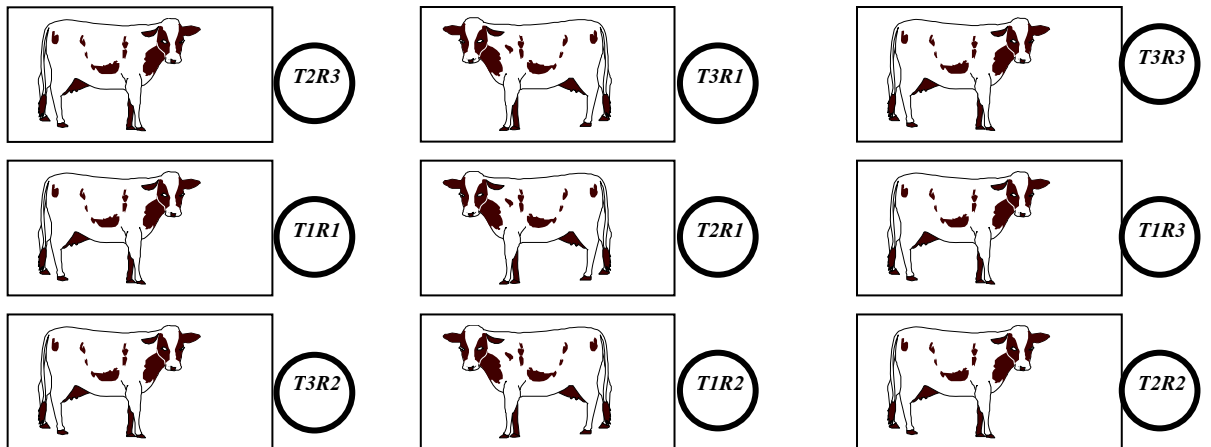
โดยการจับสลาก ทำได้โดยการเขียนเบอร์ลงในแผ่นกระดาษแผ่นเล็กๆ ที่มีวนได้ โดยให้มีจำนวนเท่ากับจำนวนหน่วยทดลองที่มีในการทดลองนั้นๆ เช่นในการทดลองที่เกี่ยวกับโคนมที่มี 3 Treatment และมีจำนวนซ้ำเป็น 3 ดังนั้นการทดลองนี้จึงมีจำนวนหน่วยทดลองทั้งหมดเท่ากับ $3 \times 3 = 9$ หน่วยทดลอง ซึ่งมี Treatment และซ้ำดังนี้คือ T_1R_1 T_1R_2 T_1R_3 T_2R_1 T_2R_2 T_2R_3 T_3R_1 T_3R_2 T_3R_3 จนครบทั้ง 9 ใบ ขณะเดียวกัน หน่วยทดลอง(ในที่นี้หน่วยทดลองคือโค 1 ตัว ในคอกขังเดี่ยวขนาด 2×4 เมตร) ที่ใช้ทั้ง 9 หน่วยทดลองก็ต้องระบุลำดับหรือตำแหน่งให้ชัดเจนเพื่อให้รู้ได้ว่าหน่วยทดลองนั้นๆ จะได้รับ Treatment ใดในการจับสลากตามลำดับที่ของการหยิบที่ตรงกับลำดับของ หน่วยทดลองนั้นๆ ดังนี้



เมื่อระบุลำดับเบอร์ให้แก่หน่วยทดลองจนครบ 9 เบอร์แล้วและได้เขียน Combination ของ Treatment และ replication ที่เป็นไปได้ทั้งหมด (possible combination) ลงในกระดาษที่ทำเป็นสลากเสร็จแล้ว จึงม้วนกระดาษเป็นชิ้นเล็กๆ เพื่อให้สามารถคละกันได้ดีขึ้น แล้วนำไปใส่ภาชนะซึ่งอาจจะเป็นกล่องเล็กๆ แล้วเขย่าให้แต่ละเบอร์เข้ากันดีแล้วจึงทำการสุ่มจับสลากออกมาโดยอาจจะกำหนดไว้ว่า ในการหยิบที่ละ 1 ใบนั้น ใบแรกที่หยิบได้ ไม่ว่าจะได้ combination ใดก็จะจัดให้หน่วยทดลองที่ 1 และในการหยิบครั้งที่ 2 ไม่ว่าจะ ได้ combination ใดก็จะจัดให้หน่วยทดลองที่ 2 และจับอย่างนี้ไปจนทุกหน่วยทดลองทุกหน่วยได้รับ Treatment ครบหมดแล้ว จากนั้นจึงวาดแผนผังทั้งหมดที่ได้จากการสุ่ม Treatment ให้แก่หน่วยทดลองทุกๆ หน่วยให้ชัดเจน และเราเรียกแผนผังของหน่วย

ทดลองทั้งหมดที่ระบุ Treatment และลำดับที่การซ้ำเรียบร้อยแล้วว่าผังการทดลอง (Experimental layout) ดังภาพข้างล่างนี้

ผังการทดลอง
(Experimental layout)



โดยวิธีการนี้จะเห็นได้ว่าหน่วยทดลองใดมีโอกาสที่จะได้รับ Treatment ใดๆนั้นจะต้องเป็นไปโดยการสุ่ม ซึ่งผู้ทดลองไม่สามารถจะกำหนดไว้ล่วงหน้าได้ว่าจะให้หน่วยทดลองใดได้รับ Treatment ใด จึงเป็นการขจัดความลำเอียง (bias) ออกไปได้ เพราะทุกอย่างที่เกิดขึ้นนั้น เป็นไปโดยการสุ่มไม่มีการได้เปรียบเสียเปรียบระหว่างหน่วยทดลอง

ซึ่งข้อเสียของจัด Treatment ให้แก่หน่วยทดลองโดยใช้วิธีการจับสลากนี้คือ ในกรณีที่มี Treatment และจำนวนซ้ำมากๆ วิธีการนี้อาจจะไม่สะดวกในทางปฏิบัติ ซึ่งอาจจะมีทางเลือกอื่นได้อีกหลายวิธี

การบันทึกข้อมูล

ในระหว่างการทดลองหรือเมื่อสิ้นสุดการทดลองแล้วจะต้องมีการบันทึกข้อมูลจากการทดลองซึ่งข้อมูลที่บันทึกนั้น อาจจะมีการบันทึกในหลายๆ เรื่อง

เช่น ในทางสัตวใหญ่เช่นในโค อาจจะมีการบันทึกอายุ และน้ำหนักเมื่อเริ่มทดลองของโคแต่ละตัว บันทึกปริมาณอาหารชั้นและอาหารหยาบที่กิน/ตัว/วัน ฯลฯ ซึ่งข้อมูลที่ได้จากการบันทึกประจำวันดังกล่าวจะต้องนำไปรวบรวมเป็นรายสัปดาห์หรือ รายเดือนหรือตามช่วงเวลาอื่นๆที่มีความหมายต่อการศึกษา เช่น ทุกๆ 2 สัปดาห์เนื่องจากต้องชั่ง น้ำหนักตัวสัตว์ทุกๆ 2 สัปดาห์

จากนั้นจึงนำมาสรุปเป็นตารางบันทึกข้อมูลของแต่ละรายการหรือแต่ละประเภทที่ศึกษาตลอดการทดลอง เพื่อใช้สำหรับการวิเคราะห์ข้อมูลทางสถิติ

ตารางที่ 1 ตัวอย่างบันทึกอาหารประจำ 2 สัปดาห์ (ตัดจากบันทึกจริง)

ตัวอย่าง ตารางสรุปบันทึกปริมาณอาหารที่กินตลอดการทดลอง

Replication	Treatment			
	1	2	3	4
1
2
3
4
Total				
Average				

การจัดเก็บข้อมูลในลักษณะนี้ก็เพื่อให้สามารถใช้ประโยชน์จากข้อมูลเพื่อการวิเคราะห์ต่างๆ ได้ง่ายขึ้น แต่ก่อนที่จะได้ข้อมูลมาอยู่ในรูปลักษณะนี้ได้ต้องมีตารางบันทึกละเอียดรายวัน(ตารางที่3.2) รายสัปดาห์ (ตารางที่3.3) มาก่อน ซึ่งตารางดังกล่าวนั้นควรจะมีรายละเอียดต่างๆที่ครบถ้วนเพื่อให้ได้ข้อมูลที่ดี มีคุณภาพ และสามารถใช้อรรถประโยชน์จากข้อมูลในด้านต่างๆได้ดีที่สุด

การวิเคราะห์ข้อมูล (Data analysis)

การวิเคราะห์ผลทางสถิติที่ใช้ในการทดลองทางวิทยาศาสตร์ทั่วไป ได้แก่การวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) ซึ่งการวิเคราะห์ความแปรปรวนนี้เป็นวิธีการทางสถิติที่ใช้ในการเปรียบเทียบค่าเฉลี่ยตั้งแต่3 ค่าขึ้นไปพร้อมๆกัน (การเปรียบเทียบค่าเฉลี่ย 2 ค่าใช้ T-test) โดยที่บุคคลที่ได้นำเอาวิธีการวิเคราะห์ความแปรปรวนมาใช้เพื่อเปรียบเทียบค่าเฉลี่ยได้แก่ Sir Ronald A. Fisher ดังนั้นเพื่อเป็นเกียรติแก่เขา จึงใช้สัญลักษณ์ F ในการทดสอบนัยสำคัญ และเรียก F นี้ว่า Fisher F

อย่างไรก็ตามข้อมูลที่ใช้ในการวิเคราะห์ความแปรปรวนจะต้องเป็นข้อมูลที่ได้จากการวัดด้วยมาตราอัตราส่วน (Ratio Scale) และมาตราอันดับ (Interval scale) เป็นอย่างน้อย ส่วนข้อมูลที่ได้จากการวัดด้วยมาตราเรียงลำดับ (Ordinal scale) เช่นลำดับที่1 ที่ 2 ที่ 3 นั้น ไม่สามารถนำมาวิเคราะห์โดยวิธีนี้

แนวคิดเชิงทฤษฎี ของการวิเคราะห์ความแปรปรวน

จากการที่ได้ศึกษาถึงอิทธิพลของปัจจัยต่างที่มีผลต่อลักษณะต้องการศึกษา เช่น ในการศึกษาถึง อิทธิพลของระดับปุ๋ยยูเรียที่มีต่อปริมาณผลผลิตของหญ้าเนเปียร์ ซึ่งสมมุติว่าใช้ปุ๋ยแตกต่างกัน 4 ระดับ คือ 0 50 100 และ 150 กิโลกรัม/ไร่ และได้ผลผลิตเฉลี่ย 2,000 2,400 3,000 และ 3,000 กิโลกรัม/ไร่ ตามลำดับ จากค่าเฉลี่ยที่ได้พอจะบอกให้ทราบว่าระดับปุ๋ยระดับใดที่ให้ผลผลิตดีกว่ากัน แต่ในการสรุปผลโดยใช้เพียงแต่ค่าเฉลี่ยอย่างเดียวมาเป็นเครื่องตัดสินใจอาจจะเป็นการสรุปที่ไม่ถูกต้อง ทั้งนี้เพราะเราทราบแล้วว่าในการทดลองทุกครั้งจะต้องมีความคลาดเคลื่อนอยู่ด้วยเสมอ ฉะนั้น ค่าความแตกต่างที่แท้จริง ตามตัวเลขจึงอาจจะไม่แตกต่างกันมากพอ จนสามารถสรุปได้ว่าแตกต่างกันจริง

ดังนั้นในการเปรียบเทียบกัน จึงจำเป็นต้องพิจารณาถึงความกระจายหรือความแปรปรวนของข้อมูลรอบๆ ค่าเฉลี่ยมาประกอบด้วย ซึ่งในการพิจารณาความแตกต่างระหว่างค่าเฉลี่ยของ Treatment โดยการพิจารณาถึงความกระจายของข้อมูลประกอบ นับเป็น **หลักขั้นพื้นฐาน** ของการวิเคราะห์ความแปรปรวน ซึ่งเป็นการนำความแปรปรวนของข้อมูลทั้งหมดที่มีมารวมกัน จากนั้นจึงแยกความแปรปรวนออกเป็นส่วนๆ ตามแหล่งที่มาหรือสาเหตุ ซึ่งในกรณีของการวิเคราะห์ความแปรปรวนโดยใช้แผนการทดลองแบบสุ่มตลอด หรือสุ่มสมบูรณ์ (Completely Randomized Design) ซึ่งเป็นการวิเคราะห์ความแปรปรวนแบบทิศทางเดียว (one way analysis) คือเป็นการวิเคราะห์หรือแยกสาเหตุของความผันแปรที่มีสาเหตุมาจากปัจจัยเพียงปัจจัยเดียว คือจาก Treatment แต่เพียงอย่างเดียว แต่อย่างไรก็ตามในการทดลองทุกครั้ง ย่อมต้องมีความคลาดเคลื่อน (error) เกิดขึ้นด้วยเสมอ ดังนั้นหากทำการทดลองโดยใช้แผนการทดลองแบบนี้ จึงมีสาเหตุของความแปรปรวนอยู่ 2 สาเหตุคือ Treatment และ error นั่นคือในการวิเคราะห์ข้อมูลโดยวิธีวิเคราะห์ความแปรปรวน เมื่อใช้แผนการทดลองแบบสุ่มตลอดนั้น ความแปรปรวนทั้งหมด (total mean of square: Total MS) ที่มีจะถูกแยกออกเป็น 2 ส่วนตามสาเหตุที่มาของมันได้แก่

1. ความแปรปรวนอันเนื่องมาจากอิทธิพลของ Treatment (Treatment mean of square: Tr MS) หรืออาจจะเรียกว่า ความแปรปรวนระหว่างกลุ่ม (Between mean of square)
2. ความแปรปรวนของความคลาดเคลื่อน (Error mean of square: Error MS) หรืออาจจะเรียกว่าความแปรปรวนภายในกลุ่ม (Within mean of square)

โดยที่ความแปรปรวนระหว่างกลุ่มทดลอง หรือความแปรปรวนเนื่องจากอิทธิพลของ Treatment (Treatment mean of square) นั้น เป็นสิ่งสำคัญที่เราต้องการจะศึกษา ส่วนความ

แปรปรวนภายในกลุ่ม คือความแปรปรวนคลาดเคลื่อน (Error mean of square) นั้น เป็นความแปรปรวนที่เกิดขึ้นทุกครั้งที่มีการทดลอง ซึ่งถ้าหากว่าความแปรปรวนทั้ง 2 แหล่งนี้มีค่าเท่ากัน หรือใกล้เคียงกัน หรือไม่ต่างกันมาก ก็แสดงว่าค่าเฉลี่ยของ Treatment ที่เรานำมาเปรียบเทียบกับนั้นไม่แตกต่างกัน (non significant) แต่ถ้าหากว่าความแปรปรวนของทั้ง 2 แหล่งแตกต่างกันมาก หรือไม่ใกล้เคียงกัน หรือไม่เท่ากัน ก็แสดงว่าค่าเฉลี่ยของ Treatment ที่เรานำมาเปรียบเทียบกับดังกล่าวแตกต่างกันอย่างมีนัยสำคัญ (significant)

นั่นคือในการทดสอบความมีนัยสำคัญทางสถิติโดยวิธีวิเคราะห์ความแปรปรวนเป็นการทดสอบโดยใช้สัดส่วน F (F-ratio) ซึ่งเป็นสัดส่วนของค่า 2 ค่า คือค่าของตัวเศษ (ตัวที่อยู่ข้างบน: Numerator) ที่เป็นความแปรปรวนระหว่างกลุ่ม (Between mean of square) หรือเรียกง่าย ๆ ว่าค่าความแปรปรวนอันเนื่องมาจาก Treatment (Treatment mean of square) และค่าตัวส่วน (ตัวที่อยู่ข้างล่าง: Denominator) ที่เป็นความแปรปรวนภายในกลุ่ม (within mean of square) หรือเรียกว่าความแปรปรวนอันเนื่องมาจากความคลาดเคลื่อน (error mean of square) ดังนั้น

$$F\text{-ratio} = \frac{\text{Treatment mean of square}}{\text{Error mean of square}}$$

ซึ่งในการวิเคราะห์ความแปรปรวนนี้เราใช้ ค่า Treatment mean of square เป็นตัวประมาณค่า (estimator) ของความแปรปรวนของ Treatment (Variation of treatment: $\sigma^2 \tau$) และเราใช้ค่า error mean of square เป็นตัวประมาณค่าของความแปรปรวนของความคลาดเคลื่อน (Variation of error: $\sigma^2 \epsilon$) ดังนั้นจึงสามารถเขียนใหม่ได้ว่า

$$F\text{-ratio} = \frac{\sigma^2 \tau}{\sigma^2 \epsilon}$$

ซึ่งถ้าหากว่าความแปรปรวนอันเนื่องมาจาก Treatment มีค่าเท่ากับความแปรปรวนอันเนื่องมาจาก Error ค่า F ที่ได้ก็จะเท่ากับ 1 (เศษเท่ากับส่วน สัดส่วน = 1 เสมอ)

แต่ถ้าเศษมีค่ามากกว่าส่วนมากๆ F-ratio ก็มีค่าสูงกว่า 1 มากขึ้นตามไปด้วย และยิ่ง F-ratio มากกว่า 1 มากขึ้นเท่าไร โอกาสที่เราจะสรุปผลการทดสอบสมมติฐาน ที่สรุปว่าปฏิเสธ Null hypothesis: H_0 ยิ่งมากขึ้นด้วย และการปฏิเสธ Null hypothesis (H_0) ของเราก็คือปฏิเสธว่าจากการที่เราได้กำหนด หรือที่ระบุไว้ว่า $H_0: \tau_i = 0$ หรือ $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ (ซึ่งแปลเป็นภาษาพูดว่า “ไม่มีอิทธิพลอันเนื่องมาจาก Treatment”) นั้นไม่เป็นความจริง หรือไม่ถูกต้อง หรือผิด และเมื่อ H_0 : ที่ตั้งไว้ผิดเราก็มีสิทธิปฏิเสธ และเมื่อปฏิเสธ H_0 : แล้วก็จะไปยอมรับ H_1 ที่ระบุไว้ว่า $\tau_i \neq 0$ หรือ

$\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$ (ซึ่งแปลเป็นภาษาพูดได้ว่า “มีอิทธิพลอันเนื่องมาจาก Treatment” หรือ “ค่าเฉลี่ยของแต่ละ Treatment นั้นไม่เท่ากัน”)

ในทางปฏิบัติเราจะปฏิเสธ Null hypothesis เมื่อ F-ratio ที่คำนวณได้(calculated F หรือ F-cal.) มีค่ามากกว่า ค่า F จากตาราง (Table F หรือ F-tab.) โดยที่ค่า F จากตารางนั้นดูได้จาก ตารางเปอร์เซ็นต์ไทล์ของการกระจายแบบ F ซึ่งต้องเข้าใจการใช้ตารางที่ถูกต้อง

แต่ในการใช้โปรแกรมสำเร็จรูปนั้น โปรแกรมมักจะให้ค่าวิกฤต ซึ่งจะบอกเป็นค่า $P > F$ มาให้ ซึ่งก็คือค่า Probability ของการปฏิเสธ H_0 คือถ้า $P > F$ มีค่า **ต่ำกว่า** ระดับนัยสำคัญที่กำหนดไว้ (0.05 หรือ 0.01) เราก็จะสรุปผลการทดสอบว่าปฏิเสธ H_0 (H_0 ตั้งไว้ว่าไม่มีอิทธิพลเนื่องจาก Treatment) ซึ่งการปฏิเสธก็คือการไม่ยอมรับว่าเป็นจริง และเมื่อไม่ยอมรับ H_0 คือ **ไม่ยอมรับว่า ไม่มีอิทธิพลเนื่องจาก Treatment** ก็มีความหมายว่าเรายอมรับว่า **มีอิทธิพลของ Treatment** จริง หรือสรุปได้ว่า **มีความแตกต่างระหว่างอิทธิพลของ Treatment** จริง และเรียกเป็นภาษาอังกฤษว่า **significant** ส่วนจะเป็นที่นัยสำคัญระดับใดนั้นก็ต้องอ่านจากตัวเลขที่ให้มาอีกครั้งหนึ่ง เช่น ถ้าผลการวิเคราะห์ให้มาว่า $P > F = 0.158$ เราจะสรุปว่ายอมรับ H_0 คือยอมรับว่า **ไม่มีอิทธิพลเนื่องจาก Treatment** จริง หรือ Non significant แต่ถ้าผลการวิเคราะห์ให้มาว่า $P > F = 0.045$ เราจะสรุปว่าปฏิเสธ H_0 ซึ่งมีความหมายว่าอิทธิพลของ Treatment มีจริงจึงทำให้ค่าเฉลี่ยนั้นแตกต่างกัน และถ้าผลการวิเคราะห์ให้มาว่า $P > F = 0.0001$ เราจะสรุปว่าปฏิเสธ H_0 ซึ่งมีความหมายว่าอิทธิพลของ Treatment มีจริงเช่นกัน แต่ระดับความแตกต่างนั้นต่ำกว่าที่ระดับ 0.01 แสดงว่าความแตกต่างนั้นมี **นัยสำคัญยิ่งทางสถิติ** ซึ่งเรียกว่า Highly significant

ดังนั้น การวิเคราะห์ความแปรปรวนจึงเป็นการรวมเอาความแปรปรวนหรือความคลาดเคลื่อนทั้งหมดที่มีจากการทดลอง แล้วมาทำการแยกออกเป็นส่วนๆ ตามสาเหตุของความแปรปรวน แล้วมาพิจารณาว่าความแปรปรวนจากปัจจัยหรือสิ่งที่เราศึกษามีจริงหรือไม่โดยดูจากค่า F ซึ่งถ้ามีจริงก็ต้องศึกษาต่อไปว่ามี Treatment ใดต่างไปจาก Treatment ใดบ้าง ซึ่งต้องอาศัยวิธีการทางสถิติอย่างอื่นคือ การเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยของ Treatment ซึ่งมีหลายวิธี แต่ในขั้นนี้มุ่งที่จะให้นักศึกษาเข้าใจหลักคิดของ การวิเคราะห์ความแปรปรวนก่อนเท่านั้น

ดังนั้นในระดับนี้ให้เข้าใจว่าเมื่อได้ทดสอบแล้วและพบว่าความแตกต่างระหว่าง Treatment นั้นมีนัยสำคัญเราก็ทราบเพียงแต่ว่า **อย่างน้อยที่สุดมี 1 Treatment ที่มีค่าเฉลี่ยแตกต่างไปจาก Treatment อื่นๆ** แต่เรายังไม่รู้ว่า ค่าเฉลี่ยของคู่ใดแตกต่างกันบ้าง ซึ่งจะรู้ได้ก็ต่อเมื่อได้ทำการเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยโดยวิธีต่างๆ เสียก่อน ซึ่งจะกล่าวถึงในลำดับต่อไป

การวิเคราะห์ข้อมูลตามการทดลองแบบสุ่มตลอด (C.R.D.)

การวิเคราะห์ข้อมูลของแผนการทดลองแบบสุ่มตลอด (Completely Randomized Design; C.R.D.) ใช้การวิเคราะห์ความแปรปรวน (Analysis of variance ; ANOVA) ดังนี้

1. นำข้อมูลทั้งหมดที่ได้จากตาราง
2. หาผลรวมแต่ละทรีทเมนต์ (T)
3. หาค่าเฉลี่ยแต่ละทรีทเมนต์ (\bar{T})
4. หาผลรวมทั้งหมด (Grand Total ; G.T.)
5. หาค่าเฉลี่ยทั้งหมด (Grand Mean ; G.M.)
6. หาค่า C.F. จากสูตร

$$C.F. = \frac{(\text{ผลรวมทั้งหมดในการทดลอง})^2}{\text{จำนวนหน่วยทดลองทั้งหมด}}$$

$$C.F. = \frac{(G.T.)^2}{4}$$

$$T = \text{จำนวนทรีทเมนต์}$$

$$r = \text{จำนวนซ้ำ}$$

7. หาค่า Sum of Square (SS) จากสูตร ดังนี้

ก. Total SS = ผลบวกของกำลังสองของข้อมูลจากแต่ละหน่วยทดลอง - C.F.

ข. Treatment SS = $\frac{\text{ผลรวมของกำลัง (ผลรวมของแต่ละทรีทเมนต์)}}{\text{จำนวนซ้ำ}} - C.F.$

ค. Error SS = Total. SS - Treatment SS

8. หาค่า degree of freedom (d.f.) จากสูตรดังนี้

ก. d.f. treatment = t - 1

ข. d.f. error = t (r - 1)

ค. d.f. total = r - 1

9. หาค่า Mean Squares (MS) จากสูตรดังนี้

$$\text{ก. Treatment MS} = \frac{\text{Treatment SS}}{t-1}$$

$$\text{ข. Error MS} = \frac{\text{Error SS}}{t(r-1)}$$

10. หาค่า $F\text{-value}$ ($\frac{F}{C}$) ได้จากสูตรดังนี้

$$F_C (\text{Treatment}) = \frac{\text{Treatment MS}}{\text{Error MS}}$$

ซึ่งมี d.f. ของตัวตั้ง คือ $t-1$ และ d.f. ของตัวหาร คือ $t(r-1)$

11. เปิดตารางจากตาราง 5% and 1% P of the D of F เพื่อเปรียบเทียบค่า $F\text{-value}$ (F_C)

11.1 ถ้าค่า F_C น้อยกว่าค่า $F\text{-table}$ ที่ระดับ 0.05 สรุปว่า ไม่มีความแตกต่างกันทางสถิติ

(non sign) ระหว่าง ทรีทเมนต์ ในเครื่องหมาย ns เหนือค่า F_C

11.2 ถ้าค่า F_C มากกว่าหรือเท่ากับ $F\text{-table}$ ที่ระดับ 0.05 แต่น้อยกว่าค่า $F\text{-table}$ ที่ระดับ

0.01 สรุปว่าทรีทเมนต์มีความแตกต่างกันอย่างมีนัยสำคัญ (sign) ทางสถิติใส่เครื่องหมายดอกจัน (*) 1 ดอกเหนือ F_C

11.3 ถ้าค่า F_C มากกว่าหรือเท่ากับ $F\text{-table}$ ที่ระดับ 0.01 แต่น้อยกว่าค่า $F\text{-table}$ ที่ระดับ

0.01 สรุปว่าทรีทเมนต์มีความแตกต่างกันอย่างมีนัยสำคัญยิ่ง (highly sign) ทางสถิติใส่เครื่องหมายดอกจัน (**) 2 ดอกเหนือ F_C

12. หาค่า C.V. ได้จากสูตรดังนี้

$$\text{C.V.} = \frac{\sqrt{\text{Error MS}}}{\text{G.M.}} \times 100 \%$$

13. ตรวจสอบความแตกต่างของค่าเฉลี่ยเพื่อหาความมั่นใจว่าทรีทเมนต์ใด แตกต่างกัน จาก ทรีทเมนต์ใด ดังเนื้อหาบทต่อไป

ตัวอย่างที่ 1 สมมุติ นายดำทำการทดลองเปรียบเทียบหญ้า 4 พันธุ์ คือ หญ้ากีนนี่ หญ้าบัพเฟิล หญ้าซอกัน หญ้ากรีนแพนนิค โดยวางแผนการทดลองแบบ CRD จัดปลูก 1 ต้นต่อ 1 กระจ่าง ทำ + ซ้ำ ในเรือนกระจก ในการเปรียบเทียบครั้งนี้มีการแตกกอ 1 ต้นต่อ 1 กระจ่าง ดังตารางดังนี้

ตารางแสดงการแตกกอ 1 ต้นต่อ 1 กระจ่าง ของหญ้า 4 พันธุ์ (หน่วย: ต้น)

พันธุ์ (Treatment)	ผลผลิตจำนวนต้นต่อกระจ่าง (ซ้ำ)				รวมผลผลิต (T)	ค่าเฉลี่ย (\bar{T})
	1	2	3	4		
หญ้ากีนนี่	7	5	10	8	$T_1 = 30$	$\bar{T}_1 = 7.5$
หญ้าบัพเฟิล	5	4	6	7	$T_2 = 22$	$\bar{T}_2 = 5.5$
หญ้าซอกัน	2	2	2	4	$T_3 = 10$	$\bar{T}_3 = 2.5$
หญ้ากรีนแพนนิค	3	2	3	4	$T_4 = 12$	$\bar{T}_4 = 3.0$
					G.T. = 74	G.M. = 4.63

หาค่า C.F.

$$\text{จากสูตร } C.F. = \frac{(\text{ผลรวมทั้งหมดในการทดลอง})^2}{\text{จำนวนหน่วยทดลองทั้งหมด}}$$

$$\begin{aligned} \text{แทนค่าในสูตรจะได้ } C.F. &= \frac{74^2}{16} \\ &= 342.25 \end{aligned}$$

หา Sam of Squares (SS)

ก. Total SS

$$\begin{aligned} \text{จากสูตร Total SS} &= \text{ผลบวกของกำลังสองของข้อมูลจากแต่ละหน่วยทดลอง} - C.F. \\ &= (7^2 + 5^2 + 2^2 + 3^2 + \dots + 4^2) - 342.25 \end{aligned}$$

18

$$\begin{aligned} &= 430 - 342.25 \\ &= 87.75 \end{aligned}$$

ข. Treatment SS

$$\text{Treatment SS} = \frac{\text{ผลบวกของกำลังสองของผลรวมของแต่ละที่พืชมณฑล}}{\text{จำนวนซ้ำ}}$$

C. F.

$$\begin{aligned} \text{Treatment SS} &= \frac{302 + 222 + 102 + 122}{4} - 342.25 \\ &= 407 - 342.25 \\ &= 64.75 \end{aligned}$$

$$\begin{aligned} \text{ค. Error SS} &= \\ \text{Error SS} &= \text{Totaal SS} - \text{Treatment SS} \\ &= 87.75 - 64.75 \\ &= 23 \end{aligned}$$

ตาราง degree of freedom (d.f.)

$$\begin{aligned} \text{ก. d.f. Treatment} &= t - 1 \\ &= 4 - 1 \\ &= 3 \\ \text{ข. d.f. error} &= t(r - 1) \\ &= 4(4 - 1) \\ &= 12 \\ \text{ค. d.f. total} &= tr - 1 \\ &= (4 \times 4) - 1 \\ &= 15 \end{aligned}$$

Mean Squares (MS)

$$\begin{aligned}
 \text{ก. MS Treatment} &= \frac{\textit{Treatment SS}}{\textit{d.f.Treatment}} \\
 &= \frac{64.75}{3} \\
 &= 21.58
 \end{aligned}$$

$$\begin{aligned}
 \text{ข. Error MS} &= \frac{\textit{Error SS}}{\textit{d.f.error}} \\
 &= \frac{23}{12} \\
 &= 1.92
 \end{aligned}$$

ค่า *F-Value* (F_2)

$$\begin{aligned}
 \text{จากสูตร } F_C &= \frac{\textit{TreatmentMS}}{\textit{ErrorMS}} \\
 \textit{Treatment MS} &= 21.58 ; \textit{d.f.} = 3 \\
 \textit{Error MS} &= 1.91 ; \textit{d.f.} = 12
 \end{aligned}$$

$$\begin{aligned}
 \text{ดังนั้น } F_C (3:12) &= \frac{21.58}{1.91} \\
 &= 11.24
 \end{aligned}$$

เปิดตารางจากตาราง 5% and 1% Points for the Distribution of *F*

$$F_{.05(3,12)} = 3.49$$

$$F_{.01(3,12)} = 5.95$$

เปรียบเทียบกับค่า F_c พบว่า

$$F_c > F_{.01(3,12)} > F_{.05(3,12)}$$

แสดงว่าทรีทเมนต์ มีความแตกต่างกันจริงทางสถิติอย่างมีนัยสำคัญยิ่งที่ระดับความเป็นไปได้ .01 หรือมีความเชื่อมั่น 99%

ค่า C.V.

$$\begin{aligned} \text{C.V.} &= \frac{\sqrt{\text{ErrorMS}}}{\text{G.M.}} \times 100\% \\ &= \frac{\sqrt{1.92}}{4.63} \times 100\% \\ &= 29.92 \end{aligned}$$

ผลการวิเคราะห์ความแปรปรวน (ANOVA) ของหญ้า 4 พันธุ์

Source of variation (S.O.V.)	df.	SS	MS	$F_{(3,120)}$	
				F_c	F-Table
Treatment	3	64.75	21.58	11.29**	3.49
Error	12	23	1.92		5.95
Total	15	87.75			

$$\text{C.V.} = 29.80\%$$

ตัวอย่างที่ 2 ตารางบันทึกความสูงของถั่วเขียวภายหลังการเพาะด้วยวัสดุที่ต่างกัน 5 ชนิดเป็นเวลา 3 วัน (ชม.)

ซ้ำ(Replication)	ชนิดของวัสดุเพาะ(Treatments)				
	ทราย(control)	ขี้เถ้า	แกลบ	ขุยมะพร้าว	ขี้เลื่อย
1	5.5	5.6	4.2	5.6	4.5
2	4.5	5.7	4.2	6	4.3
3	6	5.4	3.8	5.9	4.2
4	5.4	4.8	3.5	5.4	5.2
5	5.8	4.9	2.9	6.2	5
6	5.4	5.8	2.9	6.1	5.1
รวม	(T ₁) 32.6	(T ₂) 32.20	(T ₃) 21.50	(T ₄) 35.20	(T ₅) 28.30
เฉลี่ย	(\bar{T}_1) 5.43	(\bar{T}_2) 5.36	(\bar{T}_3) 3.58	(\bar{T}_4) 5.86	(\bar{T}_5) 4.72

Grand Total:G.T.=149.80

ขั้นตอนการวิเคราะห์ความแปรปรวน

Grand Mean:G.M.=4.99

1. ขั้นตอนการคำนวณค่าตัวปรับ (Correction Term: CT หรือ Correction Factor: CF)

$$\begin{aligned}
 CF &= \frac{(\text{ผลรวมของข้อมูลทุกค่าจากการทดลอง})^2}{\text{จำนวนค่าสังเกตทั้งหมดที่ศึกษา}} \\
 &= \frac{(\text{Grand Total})^2}{\text{No. of observations}} = \frac{(G.T.)^2}{(T) \times (R)} \\
 &= \frac{(149.80)^2}{30} \\
 &= 748
 \end{aligned}$$

2. ขั้นตอนการคำนวณค่าผลรวมของกำลัง 2 ของแหล่งที่เป็นสาเหตุของความผันแปรต่างๆ

2.1 ผลรวมของกำลัง 2 ของข้อมูลทั้งหมด (Total sum of square: Total SS) ซึ่งคำนวณได้จากสูตรดังต่อไปนี้

$$\begin{aligned}
 \text{Total SS} &= \text{ผลรวมของ (ข้อมูลแต่ละค่าที่ได้จากการศึกษา)}^2 - CF \\
 &= (5.5)^2 + (4.5)^2 + \dots + (5.0)^2 + (5.1)^2 - 748 \\
 &= 772.36 - 748 \\
 &= 24.36
 \end{aligned}$$

2.2 ผลรวมของกำลัง 2 ของ Treatment (Treatment sum of square: Treatment SS) ซึ่งคำนวณได้จากสูตรดังต่อไปนี้

$$\begin{aligned} \text{Treatment SS} &= \frac{\text{ผลรวมของ (ผลบวกของแต่ละ treatment)}^2}{\text{จำนวนซ้ำของแต่ละ treatment}} - \text{CF} \\ &= \frac{(T_1^2) + (T_2^2) + \dots + (T_5^2)}{\text{Replication}} - \text{CF} \\ &= \frac{(32.6)^2 + (32.2)^2 + \dots + (28.3)^2}{6} - 748 \\ &= 766.96 - 748 \\ &= 18.96 \end{aligned}$$

2.3 ผลรวมของกำลัง 2 ของความคลาดเคลื่อนทดลอง (Error sum of square: Error SS) ซึ่งคำนวณได้จากสูตรดังต่อไปนี้

$$\begin{aligned} \text{Error SS} &= \text{Total SS} - \text{Treatment SS} \\ &= 24.36 - 18.96 \\ &= 5.40 \end{aligned}$$

เมื่อนักศึกษาได้ค่า Sum of square ของปัจจัยต่างๆครบตามแผนการทดลองนั้นแล้ว ก็สามารถนำค่าต่างๆทั้งหมดไปใส่ในตารางการวิเคราะห์ความแปรปรวนได้ดังนี้ (ในกรณีเมื่อใช้แผนการทดลองแบบ CRD มีแค่ SS ของ Total SS และ Treatment SS แต่ในการทดลองทุกครั้งจะต้องมี Error เกิดขึ้นเสมอจึงสามารถคำนวณ Error SS ได้ว่ามีเท่าไรได้ดังจะแสดงในลำดับต่อไป)

Analysis of variance: ANOVA

Source of Variation	df	SS	MS	F-ratio
Treatment	n-1			
Error	$[(t)(r)-1]-(n-1)$			
Total	$(t)(r)-1$			

จากนั้นก็นำค่าต่างๆมาใส่ในตารางซึ่งปรากฏผลดังนี้

Source of Variation	df	SS	MS	F-ratio
Treatment	4	18.96	4.74	21.94** →
Error	25	5.40	0.216	└─┘
Total	29	24.36	0.84	

ซึ่งจากผลการทดสอบโดยพิจารณาจากค่า F (F-test) ปรากฏว่าค่าเฉลี่ยที่นำมาเปรียบเทียบกันนั้นมีความแตกต่างกันอย่างมีนัยสำคัญยิ่งทางสถิติ($P < 0.01$) เราจึงปฏิเสธ H_0 และ ยอมรับ H_1

ตัวอย่างที่ 3 การศึกษาเปรียบเทียบสารประกอบฟอสเฟต 3 ชนิด คือ Sodium Tripolyphosphate (STPP), Hexametaphosphate (SHMP) และ Sodium Acid Pyrophosphate (SAPP) ที่ระดับ 0.3 % โดยต้องการทราบว่าสารประกอบฟอสเฟตทั้ง 3 มีผลต่อการสูญเสียน้ำหนักของผลิตภัณฑ์แห้งเมื่อทำการหมักที่ 30 องศาเซลเซียส นาน 4 ชั่วโมง หรือไม่

ในการทดลองใช้เนื้อไก่พันธ์เดียวกันและจากตลาดแหล่งผลิตเดียวกัน มาทำการผลิตแห้งโดยใช้ Treatment ตามที่กำหนด ได้ผลิตภัณฑ์แห้งทั้งสิ้น 24 ตัวอย่าง และนำไปหมักที่อุณหภูมิและเวลาที่กำหนด ทำการหาน้ำหนักที่สูญเสีย (%) ได้ผลดังนี้

การสูญเสียน้ำหนัก (%)			
Tr1 (control)	Tr2(STPP)	Tr3(SHMP)	Tr4(SAPP)
5.42	3.20	4.56	2.32
5.79	3.98	4.62	2.48
6.12	3.62	4.39	2.56
5.92	3.78	4.21	2.36
5.68	3.18	4.92	2.61
5.72	3.92	4.45	2.26

การตั้งสมมุติฐาน : โดยเฉลี่ยแล้วการใช้สารฟอสเฟต 3 ชนิด และ control ทำให้การสูญเสียน้ำหนักเท่ากัน

$H_0 : \text{mean}_1 = \text{mean}_2 = \text{mean}_3 = \text{mean}_4$

$H_A : \text{มี mean อย่างน้อย 1 ตัวมีค่าไม่เท่ากัน}$

Table A.6 Values of F

Denominator df	Probability of a larger F	Numerator df								
		1	2	3	4	5	6	7	8	9
1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	.050	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
	.025	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3
	.010	4052	4999.5	5463	5825	5764	5859	5928	5982	6022
	.005	16211	20000	21615	22500	23056	23437	23715	23925	24091
2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	.025	38.51	39.00	39.17	39.25	39.30	39.33	39.35	39.37	39.39
	.010	98.50	99.00	99.17	99.25	99.30	99.33	99.35	99.37	99.39
	.005	198.5	199.0	199.2	199.3	199.3	199.3	199.4	199.4	199.4
3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
	.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	.005	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88
4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	.005	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	.025	10.01	8.43	7.76	7.38	7.15	6.98	6.85	6.76	6.68
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	.005	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77
6	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
	.010	13.75	10.92	9.78	9.15	8.73	8.47	8.26	8.10	7.98
	.005	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39
7	.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
	.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	.005	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51
8	.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	.005	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34
9	.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
	.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	.005	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54
10	.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
	.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	.005	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97
11	.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59
	.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	.005	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54
12	.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
	.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	.005	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20
13	.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	.025	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31
	.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	.005	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94
14	.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	.025	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
	.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	.005	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72

Table A.6 Values of F (Continued)

		Numerator df																		
		10	12	15	20	24	30	40	60	120	∞	P	df							
1	.100	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33	.100	1							
	.050	241.0	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3	.050								
	.025	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018	.025								
	.010	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366	.010								
	.005	24224	24426	24630	24836	24940	25044	25148	25253	25359	25465	.005								
2	.100	9.99	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49	.100	2							
	.050	19.40	19.41	19.43	19.45	19.45	19.46	19.46	19.47	19.48	19.49	.050								
	.025	39.41	39.43	39.45	39.46	39.46	39.47	39.47	39.48	39.49	39.50	.025								
	.010	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50	.010								
	.005	199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	199.5	199.5	.005								
3	.100	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13	.100	3							
	.050	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90	.050								
	.025	27.23	27.23	27.23	27.23	27.23	27.23	27.23	27.23	27.23	27.23	.025								
	.010	43.69	43.39	43.08	42.78	42.62	42.47	42.31	42.15	42.00	41.83	.010								
	.005	55.55	55.55	55.55	55.55	55.55	55.55	55.55	55.55	55.55	55.55	.005								
4	.100	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.78	.100	4							
	.050	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63	.050								
	.025	8.94	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26	.025								
	.010	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46	.010								
	.005	20.97	20.70	20.44	20.17	20.03	19.89	19.75	19.61	19.47	19.32	.005								
5	.100	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10	.100	5							
	.050	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36	.050								
	.025	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02	.025								
	.010	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02	.010								
	.005	13.62	13.38	13.15	12.90	12.78	12.66	12.53	12.40	12.27	12.14	.005								
6	.100	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72	.100	6							
	.050	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67	.050								
	.025	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85	.025								
	.010	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88	.010								
	.005	10.25	10.03	9.81	9.59	9.47	9.36	9.24	9.12	9.00	8.88	.005								
7	.100	2.70	2.67	2.																

Table A.6 Values of F (Continued)

Denominator df	Probability of a larger F	Numerator df								
		1	2	3	4	5	6	7	8	9
15	.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	.005	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54
16	.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	.025	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
	.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	.005	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38
17	.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	.025	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
	.010	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
	.005	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25
18	.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	1.98
	.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	.025	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93
	.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	.005	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14
19	.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	.025	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88
	.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	.005	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04
20	.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
	.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	.005	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96
21	.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	.025	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80
	.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	.005	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88
22	.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	.025	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76
	.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	.005	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81
23	.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	.025	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73
	.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	.005	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75
24	.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	.005	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69
25	.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	.025	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68
	.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.13
	.005	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64
26	.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	.025	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65
	.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	.005	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60
27	.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.87	1.85
	.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	.025	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63
	.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	.005	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61

Table A.6 Values of F (Continued)

		Numerator df																		
		10	12	15	20	24	30	40	60	120	∞	P	df							
15	.100	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76	.100	15							
	.050	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07	.050								
	.025	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40	.025								
	.010	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87	.010								
	.005	4.42	4.25	4.07	3.88	3.79	3.69	3.58	3.48	3.37	3.26	.005								
16	.100	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72	.100	16							
	.050	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01	.050								
	.025	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32	.025								
	.010	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75	.010								
	.005	4.27	4.10	3.92	3.73	3.64	3.54	3.44	3.33	3.22	3.11	.005								
17	.100	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69	.100	17							
	.050	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96	.050								
	.025	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25	.025								
	.010	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65	.010								
	.005	4.14	3.97	3.79	3.61	3.51	3.41	3.31	3.21	3.10	2.98	.005								
18	.100	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66	.100	18							
	.050	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92	.050								
	.025	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19	.025								
	.010	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57	.010								
	.005	4.03	3.86	3.68	3.50	3.40	3.30	3.20	3.10	2.99	2.87	.005								
19	.100	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63	.100	19							
	.050	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88	.050								
	.025	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13	.025								
	.010	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49	.010								
	.005	3.93	3.76	3.59	3.40	3.31	3.21	3.11	3.00	2.89	2.78	.005								
20	.100	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61	.100	20							
	.050	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84	.050								
	.025	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09	.025								
	.010	3.46	3.32	3.17	3.02	2.94	2.86	2.78	2.69	2.61	2.52	.010								
	.005	3.96	3.85	3.68	3.50	3.42	3.32	3.22	3.12	3.02	2.92	.005								
21	.100	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59	.100	21							
	.050	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81	.050								
	.025	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04	.025								
	.010	3.31	3.17																	

การวิเคราะห์ข้อมูลทางสถิติด้วยโปรแกรมสำเร็จรูป SAS (Statistical Analysis System)

จากการที่นักศึกษาได้ทำการวิเคราะห์ข้อมูลที่ได้จากการทดลองโดยใช้แผนทดลองทั้ง CRD และ RCBD มาแล้วโดยใช้เครื่องคิดเลขช่วยในการคำนวณ นักศึกษาหลายๆคนคงมีความเห็นว่าเป็นเรื่องที่ยุ่งยากและมีความซับซ้อน เนื่องจากจะต้องกระทำด้วยความถูกต้องในทุกๆขั้นตอน และมีการคำนวณข้อมูลชุดเดียวกันซ้ำๆกันหลายครั้ง และเนื่องจากภายหลังปี 1980 เป็นต้นมา วิทยาการทางด้าน computer ได้ถูกนำมาใช้ในกิจกรรมต่างๆในชีวิตมนุษย์ที่หลากหลายมาก โดยเฉพาะในกิจกรรมใดๆที่ทำซ้ำแล้วซ้ำอีก จึงได้มีผู้ค้นคิดและพัฒนาโปรแกรมสำเร็จรูปที่ใช้เพื่อการวิเคราะห์ทางสถิติขึ้น และได้มีใช้มาตั้งแต่ยุคแรกๆที่ คอมพิวเตอร์ส่วนบุคคล (Personal computer: PC) เริ่มได้รับความนิยมมากขึ้น โดยที่โปรแกรมสำเร็จรูปในระยะแรกๆนั้นเป็นโปรแกรมที่มีขนาดเล็ก เช่น Micro stat, Banana stat, IRRI stat หรือโปรแกรมสำเร็จรูปอย่างอื่นซึ่งเขียนขึ้นโดยภาษาคอมพิวเตอร์ที่ง่ายเช่นภาษาเบสิก แต่ก็มีโปรแกรมสำเร็จรูปบางโปรแกรมที่เป็นโปรแกรมขนาดใหญ่ เช่น SAS และ SPSS แต่ต่อมาเมื่อระบบปฏิบัติการ windows ได้รับการพัฒนามาใช้แทนเมื่อระบบปฏิบัติการ DOS ที่ใช้มาแต่เดิม โปรแกรมวิเคราะห์ข้อมูลทางสถิติก็ได้รับการพัฒนามากขึ้น และมีใช้กันอย่างกว้างขวางขึ้น ตัวอย่างของโปรแกรมสำเร็จรูปที่เป็นโปรแกรมขนาดใหญ่และใช้กับระบบปฏิบัติการ windows และเป็นที่ยอมรับในปัจจุบันคือ SAS, SPSS, Stat graphic, Sigma Stat และโปรแกรมสำเร็จรูปอื่นๆ

ในบรรดาโปรแกรมสำเร็จรูปเหล่านี้ SAS จัดเป็นโปรแกรมที่ได้รับความนิยมสูงสุดที่ใช้ในการวิเคราะห์ข้อมูลที่ได้จากการวิจัยทางด้านวิทยาศาสตร์ทุกสาขา ขณะที่ SPSS ได้รับความนิยมมากกว่าในการใช้กับการวิเคราะห์ข้อมูลทางสังคมศาสตร์ อย่างไรก็ตามข้อจำกัดสำคัญของการใช้โปรแกรมสำเร็จรูปเหล่านี้คือค่าลิขสิทธิ์ของโปรแกรมที่มีราคาค่อนข้างแพง จึงนับเป็นอุปสรรคของการใช้โปรแกรมสำเร็จรูปทั้ง 2 นี้สำหรับองค์กร สถาบัน หรือหน่วยงานที่มีข้อจำกัดด้านงบประมาณ

โปรแกรมสำเร็จรูป SAS ที่ใช้กับระบบปฏิบัติการ windows ในระยะแรกๆ เป็นโปรแกรมสำเร็จรูป SAS Version 6.12 ซึ่งเป็นลิขสิทธิ์ของ SAS Institute Incorporation, Cary, North Carolina, USA. ซึ่งมีค่าลิขสิทธิ์ต่อปีเป็นเงินเป็นแสนบาท ดังนั้นหน่วยงานที่ซื้อลิขสิทธิ์มาจึงควรคำนึงถึงความคุ้มค่าของการใช้กับค่าลิขสิทธิ์ที่ต้องจ่ายไป แต่ถึงแม้ว่าค่าลิขสิทธิ์ของโปรแกรม SAS จะแพงแต่โปรแกรมนี้อีกยังคงได้รับความนิยมอย่างกว้างขวาง ทั้งนี้เนื่องจาก SAS เป็นโปรแกรมที่ประสิทธิภาพสูง มีความยืดหยุ่นค่อนข้างดี ใช้ร่วมกับโปรแกรมอื่นๆ โดยเฉพาะโปรแกรมในชุด Microsoft office ทั้ง Microsoft word และ Microsoft excel ได้เป็นอย่างดี โดยเฉพาะ SAS version 6.12 นี้ เป็น version ที่ใช้ได้ง่ายเพราะใช้กับระบบ Windows ที่เมื่อผู้ใช้ไม่เข้าใจในเรื่องใดๆ ก็สามารถเรียกตัวช่วย (Help) ในเรื่องนั้นๆเพื่อทำความเข้าใจได้ได้ตามแบบของโปรแกรมที่ทำงานโดยอาศัยระบบปฏิบัติการ Windows เป็นโปรแกรมควบคุมการทำงาน เหมือนโปรแกรมทั่วไป

ความรู้ทั่วไปเกี่ยวกับโปรแกรมสำเร็จรูป SAS

SAS เป็นคำย่อของ Statistical Analysis System เป็นโปรแกรมสำเร็จรูปที่พัฒนาขึ้นโดย SAS Institute incorporation แห่งเมือง Cary ในมลรัฐ North Carolina ประเทศสหรัฐอเมริกา ซึ่งได้มีการพัฒนามาแล้วหลายๆ version แต่ version ที่นักศึกษาใช้อยู่นี้เป็น version 6.12 ซึ่งเป็น version ที่ทำงานภายใต้ระบบปฏิบัติการ Windows ซึ่งมีลักษณะต่างๆไปของโปรแกรมหลายอย่างแต่จะขอกล่าวถึงเฉพาะที่สำคัญคือ

ชุดคำสั่งของ SAS (SAS command)

คำสั่งของ SAS มี 2 ขั้นตอน (steps) คือตอนที่ป้อนข้อมูลของ SAS (Data step) และ ตอนที่เป็นการประมวลผล (procedure step หรือ Proc step) ซึ่งสามารถประมวลผลได้คราวเดียวกันหลายๆ procedure เช่น Proc print, Proc sort, Proc ANOVA, Proc GLM, Proc Plot ฯลฯ ซึ่งถือได้ว่าเป็นลักษณะที่เด่นที่สุดของ SAS

ตัวอย่างการใช้ SAS ในการวิเคราะห์ข้อมูลทางสถิติ (มทชย 2544)

PROC ANOVA for CRD

ในการทดสอบความแตกต่างของค่าเฉลี่ยโดยใช้ t-test (PROC TTEST) นั้นใช้ได้กรณีที่มีตัวอย่างเพื่อการเปรียบเทียบเพียง 2 กลุ่มหรืองานทดลองที่มีเพียง 2 ทรีทเมนต์ หากต้องการเปรียบเทียบค่าเฉลี่ยมากกว่านี้หรือมีอิทธิพลอื่นเข้ามาพัวพันกับค่าสังเกตหลายปัจจัย ควรใช้วิธีการวิเคราะห์ความแปรปรวน (Analysis of Variance) เพื่อทดสอบโดยรวม (over all test) ของความแตกต่างของค่าเฉลี่ยจากทุกทรีทเมนต์เสียก่อน จากนั้นจึงจะทดสอบโดยละเอียดถึงความแตกต่างของค่าเฉลี่ยแต่ละทรีทเมนต์โดยวิธี multiple comparison หรือ orthogonal contrast ต่อไป

จาก model ของ Completely Randomized Design

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

```
PROC ANOVA;
CLASS ตัวแปร1;
MODEL ตัวแปร1ค่าสังเกต = ตัวแปร1;
MEANS ตัวแปร1;
RUN;
```

Ex5.5

จากการทดลองอาหารชั้นจาก 3 บริษัท (A, B, C) ที่มีผลต่อการให้นมของโคนมพันธุ์แท้โฮลสไตน์เฟรียเนียน ผู้ทดลองนำโคที่มีระยะการให้นมเดียวกันและความใกล้เคียงกันในเรื่องอายุและน้ำหนักและความสมบูรณ์อื่น ๆ จำนวน 12 ตัวเข้าทดลอง โดยโคนมตัวใดจะได้รับอาหารชั้นจากบริษัทใดนั้นเป็นไปอย่างสุ่ม ทำการทดลองเป็นเวลา 3 เดือน ได้ข้อมูลปริมาณน้ำนมเฉลี่ยต่อวันเป็นกิโลกรัมดังนี้

อาหารบริษัท

A	B	C
16.5	19.0	22.5
15.5	18.5	18.5
17.0	20.0	23.0
18.0	19.5	19.0

```

DATA CRD;
  DO REP = 1 TO 4;
    DO FEED = 'A','B','C';
      INPUT MILK @@;
      OUTPUT;
    END;
  END;
CARDS;
16.5 19.0 22.5
15.5 18.5 18.5
17.0 20.0 23.0
18.0 19.5 19.0
;
PROC ANOVA;
  CLASS FEED;
  MODEL MILK = FEED;
  MEANS FEED /DUNCAN;
RUN;

```

Note :

- PROC ANOVA จะสร้างตารางวิเคราะห์ความแปรปรวนตามตัวแปรที่ระบุในคำสั่ง MODEL และคำสั่ง MEANS จะคำนวณค่าเฉลี่ยของแต่ละทริทเมนต์ ส่วน /DUNCAN เป็น option ในการสั่งให้ทำการทดสอบความแตกต่างของค่าเฉลี่ยโดยวิธี Duncan's New Multiple Range Test (DMRT) ที่ระดับนัยสำคัญ 0.05

- ♣ Option สำหรับการตรวจสอบความแตกต่างของค่าเฉลี่ยอื่นๆ ที่นิยมใช้ได้แก่ LSD, TUKEY, SCHEFFE, SNK, BON, DUNNET
- ♣ หากต้องการทดสอบที่ระดับนัยสำคัญอื่นต้องระบุเพิ่ม ซึ่งสามารถกำหนดได้อีก 2 ระดับคือ 0.01 และ 0.10 ตัวอย่างเช่น /DUNCAN ALPHA = 0.01;
- ♣ ในกรณีที่ข้อมูลสูญหายหรือจำนวนซ้ำไม่เท่ากันในแต่ละทรีทเมนต์ให้ใส่เครื่องหมายจุด (.) ที่ข้อมูลว่างนั้น ดังตัวอย่าง 5.6

Ex5.6

ในการเปรียบเทียบคุณภาพซากของสุกร 4 สายพันธุ์ ได้แก่ ลาร์จไวท์ (LW), แลนเรซ (LR), ดุรีอค (DR) และเหมยซาน (MS) โดยมีข้อมูลความหนาแน่นสันหลัง (back fat thickness) ดังนี้

LW	LR	DR	MS
1.3	1.2	1.4	1.8
1.4	1.4	1.2	1.4
1.1	1.0	1.1	1.6
.	1.3	.	1.9

```
DATA CRDMISS;
DO REP = 1 TO 4;
DO BREED = 'LW','LR','DR','MS';
INPUT BACKFAT @@;
OUTPUT;
END;
END;
CARDS;
1.3 1.2 1.4 1.8
1.4 1.4 1.2 1.4
1.1 1.0 1.1 1.6
. 1.3 . 1.9
;
PROC ANOVA;
```

The SAS System 17:05 Monday, July 19, 2009 1

The ANOVA Procedure

Class Level Information

Class	Levels	Values
feed	3	A B C

Number of observations 18

The SAS System 17:05 Monday, July 19, 2009 2

The ANOVA Procedure

Dependent Variable: milk

	Sum of				
Source	DF	Squares	Mean Square	F Value	Pr > F
Model	2	0.23207778	0.11603889	13.48	0.0004
Error	15	0.12915000	0.00861000		
Corrected Total	17	0.36122778			

R-Square	Coeff Var	Root MSE	milk Mean
0.642469	0.853155	0.092790	10.87611

Source	DF	Anova SS	Mean Square	F Value	Pr > F
feed	2	0.23207778	0.11603889	13.48	0.0004

The ANOVA Procedure

Duncan's Multiple Range Test for milk

NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.00861
Number of Means	2 3
Critical Range	.1142 .1197

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	feed
A	11.03667	6	C
B	10.79833	6	A
B			
B	10.79333	6	B