

การใช้สถิติ ในการวิเคราะห์ข้อมูลการวิจัย

* รศ.ดร. บุญชม ศรีสะอาด

สถิติที่ใช้ในการวิเคราะห์ข้อมูลการวิจัยมีหลายกลุ่มหรือหลายชนิดในที่นี้จะยกกล่าวเป็น 7 กลุ่มดังนี้

1. สถิติที่ใช้อธิบายคุณลักษณะหรือรายละเอียดของกลุ่มที่ศึกษา

1.1 ร้อยละ (Percentage) เป็นสถิติที่นิยมใช้มาก เป็นการเทียบความถี่หรือจำนวนที่ต้องการกับความถี่หรือจำนวนที่เทียบเป็น 100 หากจากสูตร

$$P = \frac{P}{N} \times 100 \dots\dots\dots (1)$$

1.2 การวัดแนวโน้มเข้าสู่ส่วนกลาง (Measures of Central Tendency) เป็นสถิติที่ใช้เป็นตัวแทนของข้อมูลในกลุ่มหรือเซตนั้น ที่นิยมใช้มี ค่าเฉลี่ย มัธยฐาน และฐานนิยม

ค่าเฉลี่ย

ก. ด้านการบรรยาย

1. ที่แจ้งความยากของแบบสอบถาม
2. เปรียบเทียบความสามารถของกลุ่ม
3. ใช้พิจารณา Content Validity เช่นระบบ 0 ถึง 4 ข้อใดมีค่าเฉลี่ย 2.5 ถึง 4.0 นับว่าเป็นข้อที่มีความเที่ยงตรงสูงมาก
4. ใช้ผลการประเมินโครงสร้างเมื่อใช้เครื่องมือประเภท Rating Scale
5. ใช้สำหรับศึกษาระดับความมากน้อยของตัวแปรที่มีในกลุ่มที่จะศึกษาเมื่อข้อมูลอยู่ในมาตราอันดับหรือสัดส่วน เช่น ระดับความต้องการในเรื่องต่าง ๆ ระดับความสนใจในเหตุการณ์ กิจกรรม สิ่งต่าง ๆ ระดับเจตคติต่อเหตุการณ์ บุคคล สิ่งต่าง ๆ ระดับปัญหาในเรื่องต่าง ๆ เป็นต้น ตัวอย่างงานวิจัยในลักษณะนี้ได้แก่

- ศึกษาความต้องการการนิเทศการสอนของครูภาษาไทย
- ขวัญในการทำงานของครูอาจารย์โรงเรียนมัธยมศึกษา
- ปัญหาและความต้องการในการดำเนินงานห้องสมุดกลุ่มโรงเรียนของครูบรรณารักษ์ห้องสมุดโรงเรียน

* รองศาสตราจารย์ ภาควิชาการวัดผล และวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ มหาสารคาม

6. ใช้สำหรับคำนวณหาค่าสถิติอื่น ๆ เช่น หาส่วนเบี่ยงเบนมาตรฐาน หาคะแนน

6. ใช้สำหรับคำนวณค่าสถิติอื่น ๆ เช่น หาส่วนเบี่ยงเบนมาตรฐาน หาคะแนนมาตรฐาน Z-score ฯลฯ Z-score หาสูตร $Z = \frac{x - \bar{x}}{s}$

ข. ด้านสถิติอ้างอิง

ใช้ในการคำนวณค่าสถิติต่าง ๆ เช่น ค่า t-test เป็นต้น

คำมัยฐาน

1. มัยฐานจะเหมาะสมกว่าค่าเฉลี่ยในกรณีที่ข้อมูลชุดนั้นมีการแจกแจงแบบเบ้มาก ๆ หรือกรณีที่ข้อมูลบางค่าผิดปกติมาก ๆ เช่น สูงหรือต่ำกว่าค่าอื่นอย่างมากมาย เช่นในบางท้องที่รายได้ของประชากรต่อเดือนส่วนใหญ่จะใกล้เคียงกัน เช่น ระหว่าง 3,000-5,000 บาท แต่มีเศรษฐีคนหนึ่งมีรายได้ต่อเดือน 300,000 บาท ถ้านำรายได้ต่อเดือนของเศรษฐีผู้นี้ไปรวมกับรายได้ต่อเดือนของคนอื่น ๆ แล้วหาค่าเฉลี่ย ก็จะได้ค่าเฉลี่ยที่ค่อนข้างสูง กรณีเช่นนี้ควรจะใช้คำมัยฐาน

2. กรณีที่ต้องการหาค่าที่เป็นตัวแทนในกลุ่มของค่าสถิติบางชนิด เช่น ค่าสหพันธ์อย่างง่าย ค่าขนาดของผลลบ (Effect size) ฯลฯ ค่าเหล่านี้ไม่เหมาะสมที่จะนำมาบวกลบกันจึงไม่เหมาะสม ที่จะใช้ค่าเฉลี่ยเพื่อเป็นตัวแทน กรณีดังกล่าวนิยมใช้คำมัยฐาน โดยระบุพิสัยร่วมด้วย ดังตัวอย่างรายงานการวิจัยของผู้เขียน (บุญชม ศรีสะอาด 2524 : 45) ที่ศึกษางานวิจัยที่เกี่ยวกับความสัมพันธ์ระหว่างผลการเรียนกับแรงจูงใจใฝ่สัมฤทธิ์ หลังจากที่ยารายงานค่าสหพันธ์ระหว่างตัวแปรทั้งสองจากการศึกษาของหลายคน แล้วสรุปค่าที่เป็นตัวแปรโดยใช้คำมัยฐานดังนี้

“สรุปค่าสหพันธ์ระหว่างผลการเรียนกับแรงจูงใจใฝ่สัมฤทธิ์ จากการศึกษาที่กล่าวมา มีพิสัย .42 ถึง .72 และมัยฐานเท่ากับ .53”

ผลสรุปดังกล่าวนำไปเป็นเหตุผลในการตั้งสมมุติฐานทางการวิจัยของผู้วิจัย ซึ่งทำให้มีความมั่นใจในความถูกต้องของสมมุติฐานมากขึ้น

ฐานนิยม

ฐานนิยมได้รับการนำไปใช้บ้างในสถิติบรรยาย แต่จะได้รับการนำไปใช้น้อยในสถิติอ้างอิง ในด้านการบรรยายนั้นหากอยู่ในรูปของมาตราอันตรภาคสัดส่วนจะใช้ฐานนิยมเมื่อต้องการทราบตัวแทนของข้อมูลอย่างคร่าว ๆ แต่ถ้าข้อมูลมีคุณลักษณะของมาตรานามบัญญัติจะใช้ฐานนิยมมาก ดังตัวอย่างกรณีที่ควรใช้ฐานนิยมดังนี้

- ต้องการทราบความนิยมของกลุ่มใด ๆ เช่น ต้องการทราบว่าในจังหวัด ก. ใ้รรถยนต์ยี่ห้อใดมากที่สุด นักศึกษาในมหาวิทยาลัย ข. ชอบพรรคการเมืองใด
- ต้องการทราบมาตราวัดแนวโน้มนำเข้าสู่ส่วนกลางของการแจกแจงของสิ่งใด ๆ ที่จัดเป็นประเภท ๆ

เช่น ต้องการทราบการแจกแจงของสีผิวหนึ่ง ต้องการทราบว่าผู้ชมรายการโทรทัศน์มีทัศนคติเช่นไรกับรายการต่าง ๆ

1.3 การวัดการกระจาย (Measures of Variability) เป็นสถิติที่ช่วยให้ทราบถึงระดับของการกระจายหรือการแปรผันของคะแนนในกลุ่มนั้น นิยมใช้ส่วนเบี่ยงเบนมาตรฐาน และความแปรปรวน

ส่วนเบี่ยงเบนมาตรฐาน

ก. ด้านการบรรยาย มักใช้ควบคู่กับค่าเฉลี่ย เช่น

1. ในการสร้างแบบทดสอบมาตรฐาน จะต้องพิจารณาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน โดย รายงานค่าดังกล่าวออกมา

2. ใช้สำหรับแสดงการเปรียบเทียบความสามารถของกลุ่มควบคู่กับค่าเฉลี่ย

3. ใช้ร่วมกับค่าเฉลี่ย ในการพิจารณาความเที่ยงตรงตามเนื้อหา (Content validity) ของแบบทดสอบอิงเกณฑ์ โดยนำข้อสอบแบบอิงเกณฑ์ให้ผู้เชี่ยวชาญด้านเนื้อหาวิชาหรือครูผู้สอนวิชานั้นหรือทั้งสองฝ่าย พิจารณาว่าข้อสอบแต่ละข้อวัดจุดประสงค์ตามที่ผู้สร้างข้อสอบได้ระบุไว้หรือไม่ แล้วให้คะแนนความเหมาะสมเป็น 4, 3, 2, 1 หรือ 0 ถ้าเห็นว่าข้อนั้นเหมาะสมสำหรับวัตถุประสงค์ที่ระบุไว้ในระดับมากที่สุด มาก ปานกลาง น้อย หรือไม่เหมาะสมเลย ตามลำดับ ผู้สร้างข้อสอบจะนำเอาผลการตอบมาหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนน แต่ละข้อ ข้อใดมีค่าเฉลี่ย 2.50 ถึง 4.00 และส่วนเบี่ยงเบนมาตรฐานไม่เกิน 1.00 ถือว่าเป็นข้อที่มีความเที่ยงตรงตามเนื้อหา มาก สมควรนำไปใช้วัดในจุดประสงค์นั้นได้

4. ใช้ร่วมกับค่าเฉลี่ยในการชี้ผลการประเมินผลโดยสรุปจะชี้ถึงระดับความคิดเห็นหรือระดับของการประเมินผล ส่วนเบี่ยงเบนมาตรฐานจะชี้ถึงว่าผู้ประเมินผลเหล่านั้นตัดสินใจได้คล้าย ๆ กันหรือแตกต่างกัน

5. ใช้ร่วมกับค่าเฉลี่ยในการศึกษาระดับความมากน้อยของค่าตัวแปร ที่มีในกลุ่มที่ศึกษา เมื่อข้อมูลอยู่ในระดับอันดับหรืออัตราส่วน ส่วนเบี่ยงเบนมาตรฐานจะชี้ถึงว่าสมาชิกในกลุ่มนั้นมีการกระจายในค่าตัวแปรที่ศึกษามากหรือน้อย

6. ใช้สำหรับคำนวณค่าสถิติอื่น ๆ เช่น คำนวณหาคะแนนมาตรฐาน Z-score ซึ่งคำนวณจากสูตร $Z = \frac{x - \bar{X}}{S}$

ข. ด้านสถิติอ้างอิง ใช้ในการคำนวณค่าสถิติต่าง ๆ เช่น ค่า t ใน t-test คำนวณโดยใช้ค่าความแปรปรวนของแต่ละกลุ่ม ความแปรปรวนก็คือกำลังสองของส่วนเบี่ยงเบนมาตรฐาน

2. สถิติที่ใช้หาความสัมพันธ์ระหว่างตัวแปร

สถิติที่หาความสัมพันธ์ (Correlation) ซึ่งก็คืออัตราคนที่บอกทิศทางและขนาด ความ

สัมพันธ์ระหว่างตัวแปรต่าง ๆ สหพันธ์มีหลายชนิด ขึ้นกับจำนวนตัวแปรประเภทหรือลักษณะหรือระดับของการวัดตัวแปร เช่น

สหพันธ์กรณีมีตัวแปรสองตัว มีหลายชนิด แต่ที่นิยมใช้ได้แก่ สหพันธ์อย่างง่ายหรือที่เรียกว่า สหสัมพันธ์แบบ Product-moment ได้รับการนำไปใช้ในการวิจัยหลายรูปแบบ ดังนี้

1. ใช้ในการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรสองตัว ตัวอย่างเช่นผู้เขียนวิจัย รูปแบบของผลการเรียนในโรงเรียน. ในการวิจัยเรื่องนี้มีส่วนหนึ่งที่ผู้วิจัยต้องการค้นหาว่า ตัวแปรอิสระแต่ละตัว อันได้แก่ ความรู้พื้นฐานเดิม ความถนัด แรงจูงใจใฝ่สัมฤทธิ์ มโนภาพเกี่ยวกับตนเอง ฯลฯ มีความสัมพันธ์กับผลการเรียนเพียงใด จึงหาสัมประสิทธิ์สหสัมพันธ์อย่างง่ายระหว่างตัวแปรอิสระแต่ละตัว ดังกล่าวกับผลการเรียน

2. ใช้เป็นค่าพื้นฐานสำหรับการวิเคราะห์ทางสถิติที่มีความซับซ้อนบางตัว เช่น ในการวิเคราะห์องค์ประกอบ (factor analysis) ในการวิเคราะห์สหสัมพันธ์พหุคูณ (Multiple regression) ในการวิเคราะห์ความแปรปรวนพหุคูณ (Multivariate analysis of variance) เป็นต้น

สหสัมพันธ์พหุคูณ (Multiple Correlation) เป็นสหสัมพันธ์ระหว่างตัวแปรตาม 1 ตัวกับกลุ่มตัวแปรอิสระ (มากกว่า 1 ตัว) เขียนแทนด้วยสัญลักษณ์ $R_{y.1.2...K}$ หรือ R ตัวอย่างกรณีที่ใช้สหสัมพันธ์พหุคูณ เช่น ผู้วิจัยต้องการศึกษาความสัมพันธ์ระหว่างผลสัมฤทธิ์ในวิชาภาษาไทย (ตัวแปรตาม) กับกลุ่มตัวแปรอิสระ 4 ตัว ได้แก่ ความถนัด แรงจูงใจใฝ่สัมฤทธิ์ มโนภาพเกี่ยวกับตนเอง คุณภาพของการสอน เป็นต้น

สหสัมพันธ์คาโนนิกอล (Canonical Correlation) เป็นสหสัมพันธ์ระหว่างกลุ่มตัวแปรตาม (มากกว่า 1 ตัว) กับกลุ่มตัวแปรอิสระ (มากกว่า 1 ตัว) ตัวอย่างกรณีที่ใช้สหสัมพันธ์คาโนนิกอล เช่น ผู้วิจัยต้องการศึกษาความสัมพันธ์ระหว่างผลการเรียน (ตัวแปรตาม) ซึ่งมี 2 ประเภท คือ ผลสัมฤทธิ์ และเจตคติที่มีต่อวิชา กับกลุ่มตัวแปรอิสระ 4 ตัว คือ ความถนัด แรงจูงใจใฝ่สัมฤทธิ์ มโนภาพเกี่ยวกับตนเอง คุณภาพในการสอน เป็นต้น

3. การวิเคราะห์การถดถอยพหุคูณ

การวิเคราะห์การถดถอยพหุคูณ (Multiple regression) เป็นเรื่องของการพยากรณ์ตัวเกณฑ์ 1 ตัว โดยมีตัวพยากรณ์ตั้งแต่ 2 ตัวขึ้นไป ตัวอย่างเช่น ผู้วิจัยต้องการพยากรณ์ผลการเรียน (Y) โดยใช้ตัวพยากรณ์ 3 ตัว คือ ความถนัด (X_1) ความสนใจ (X_2) และคุณภาพของการสอน (X_3) เป็นต้น เทคนิคของการวิเคราะห์การถดถอยพหุคูณจะช่วยให้ทราบอิทธิพลของตัวแปรอิสระแต่ละตัวและทั้งกลุ่มที่มีต่อตัวแปรตามในการวิเคราะห์การถดถอยพหุคูณ สิ่งสำคัญที่ต้องการหา คือ

1. สัมประสิทธิ์สหสัมพันธ์พหุคูณ
2. สมการพยากรณ์ในรูปคะแนนดิบ หรือในรูปคะแนนมาตรฐาน หรือทั้งคู่

3. ความคลาดเคลื่อนมาตรฐานในการพยากรณ์

นอกจากนั้นจะต้องทดสอบนัยสำคัญของสัมประสิทธิ์สหสัมพันธ์พหุคูณอาจทดสอบ นัยสำคัญของสัมประสิทธิ์การถดถอย และทดสอบนัยสำคัญของตัวแปรที่เพิ่มเข้ามาในสมการถดถอย

4. การวิเคราะห์สาเหตุหรือการวิเคราะห์เส้นทาง (Path analysis) เป็นเทคนิคที่ไรท์ (Sewall Wright) เป็นผู้คิดขึ้น เทคนิคการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรหลายตัวไม่ว่าจะเป็นการวิเคราะห์ สหสัมพันธ์พหุคูณ (Multiple correlation) การวิเคราะห์ตัวประกอบร่วม (Commonality analysis) หรือการวิเคราะห์สหสัมพันธ์คาโนนิคอล (Canonical correlation) ล้วนแต่ชี้ถึงการมีความสัมพันธ์แบบธรรมชาติระหว่างตัวแปรหรือกลุ่มตัวแปรไม่ได้ยืนยันหรือสนับสนุนถึงความสัมพันธ์ในรูปที่เป็นสาเหตุและผล การยืนยันหรือสนับสนุนความสัมพันธ์ในรูปที่เป็นสาเหตุและผลคือการยืนยันสนับสนุน ว่า ตัวแปรอิสระ (Independent variable) ตัวใด เป็นสาเหตุให้เกิดความแปรปรวนหรือความแตกต่าง ในตัวแปรตาม (Dependent variable) และสาเหตุดังกล่าวนี้เป็นสาเหตุที่เกิด จากตัวแปรอิสระตัวนั้น ๆ โดยตรง หรือเป็นสาเหตุโดยทางอ้อมกล่าวคือไปร่วมกับตัวแปรอื่นในการ ทำให้เกิดความแปรปรวนใน ตัวแปรตาม หรือเป็นไปทั้งสองทาง ความรู้ดังกล่าวนี้นับว่าเป็นความรู้ที่ลึกลงไปและช่วยให้เกิดความกระจ่างชัดได้มากขึ้น เพื่อที่จะทราบความรู้ในลักษณะดังกล่าว ไรท์ คิดเทคนิคของการวิเคราะห์สาเหตุ ที่เรียกว่า Path analysis ขึ้น

ในการวิเคราะห์ด้วย Path analysis นั้น จะต้องมีการสร้างหรือรูปแบบความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ในรูปที่เป็นสาเหตุและผลเสียก่อน โครงสร้างหรือรูปแบบดังกล่าวนี้เป็นโครงสร้างหรือรูปแบบตามสมมุติฐาน ที่จะตั้งสร้างทฤษฎี แนวความคิดและผลการวิจัย มีเหตุผลน่าเชื่อถือกว่าตัวแปรอิสระนั้น ๆ เป็นสาเหตุต่อตัวแปรตาม ตามโครงสร้างหรือรูปแบบที่แท้จริง จากนั้นก็นำค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรมาวิเคราะห์ตามโครงสร้างหรือรูปแบบ ด้วยเทคนิคของ Path analysis ผลการวิเคราะห์จะเป็นการทดสอบความเที่ยงตรง (validity) ของโครงสร้างหรือรูปแบบตามสมมุติฐานนั้น และต้องปรับปรุงโครงสร้างหรือรูปแบบอีกด้วย นั่นคือบอกได้ว่า จากหลักฐานตามทฤษฎีที่ไปสังเกตพบได้ โครงสร้างนี้ สามารถสนับสนุนการอ้างสาเหตุตามโครงสร้างหรือรูปแบบนั้น ๆ ได้หรือไม่ กรณีที่มีการปรับปรุงโครงสร้างหรือรูปแบบแล้วจะสามารถสนับสนุนการอ้างสาเหตุตามโครงสร้างหรือรูปแบบที่ได้นั้นหรือไม่

5. สถิติที่ใช้ทดสอบสมมุติฐานเกี่ยวกับความแตกต่างระหว่างค่าเฉลี่ย ๕ กลุ่ม นิยมใช้ t-test ซึ่งจำแนกได้เป็น ๒ กรณีคือ กรณีกลุ่มตัวอย่างไม่เป็นอิสระ กับกรณีกลุ่มตัวอย่างเป็นอิสระหากเป็นจำแนกกรณีกลุ่มตัวอย่างเป็นอิสระแล้วกันยังจำแนกได้หลายกรณีในที่นี้จะกล่าวถึงกรณีเดียวคือ ความแปรปรวนของประชากรเท่ากัน กับกรณีที่ assume ว่าความแปรปรวนของประชากรไม่เท่า

กันดังนี้

5.1 กรณีกลุ่มตัวอย่างไม่เป็นอิสระแก่กัน (Dependent) การจัดว่ากลุ่มตัวอย่างไม่เป็นอิสระแก่กันก็เนื่องจากผลการวัดในสมาชิก 2 กลุ่ม หรือผลการวัดจากกลุ่มเดิมออกมา 2 ค่านั้น มีความสัมพันธ์กัน ดังกรณีต่อไปนี้

ก. กลุ่มเดิมวัดออกมา 2 ค่า เช่น คะแนนสอบก่อนสอนกับสอบหลังสอน คะแนนบทบาทที่เป็นจริงกับบทบาทที่คาดหวัง ฯลฯ แต่ละคนมีผลการวัดออกมา 2 ค่า จึงถือว่าค่าทั้งสองนั้นไม่เป็นอิสระแก่กัน

ข. กรณีแบ่งกลุ่มตัวอย่างโดยใช้วิธีจับคู่ คนที่มีลักษณะคล้ายกันจัดไปอยู่คนละกลุ่มเป็นคู่ ๆ ไป

ค. กรณีคู่ฝาแฝด แยกออกเป็น 2 กลุ่ม
สูตรที่ใช้ในการคำนวณค่า t ได้แก่

$$t = \frac{\Sigma D}{\sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{(n-1)}}} \dots\dots\dots(2)$$

เมื่อ t แทน ค่าสถิติที่จะใช้เปรียบเทียบกับค่าวิกฤต เพื่อทราบความมีนัยสำคัญ

D แทน ค่าผลต่างระหว่างคู่คะแนน

n แทน จำนวนกลุ่มตัวอย่างหรือจำนวนคู่คะแนน

เมื่อคำนวณค่า t และ จะนำค่า t ที่คำนวณได้ไปเปรียบเทียบกับ ค่าวิกฤตของ t ที่ระดับ นัยสำคัญที่กำหนดไว้ถ้าค่า t ที่คำนวณได้มีค่าเท่ากับหรือมากกว่า ค่าวิกฤตของ t ก็จะปฏิเสธสมมุติฐานหลัก (H_0) ยอมรับสมมุติฐานตรงข้าม (H_A) ซึ่งโดยทั่วไปจะเป็นสมมุติฐานในการวิจัยที่ตั้งขึ้น แต่ถ้าค่า t ที่คำนวณได้มีค่าน้อยกว่าค่าวิกฤตของ t ก็จะยอมรับ H_0

ค่าวิกฤตของ t หาได้จากการเปิดตารางค่าวิกฤตของ t โดยนำค่า df ซึ่งเท่ากับ $n - 1$ และระดับนัยสำคัญไปที่กำหนดขึ้นไปเปิดตารางดังกล่าว ก็จะพบค่าวิกฤตของ t ตามต้องการ

5.2 กรณีกลุ่มตัวอย่างเป็นอิสระแก่กัน (Independent) การจัดว่ากลุ่มตัวอย่างเป็นอิสระแก่กันหรือไม่สัมพันธ์กัน ก็เนื่องจากค่าเฉลี่ยสองค่าไม่ได้ตามกลุ่มเดียวกันถือว่าเป็นกรณีกลุ่มตัวอย่างเป็นอิสระแก่กัน

ก. การสุ่มตัวอย่างหรือผลการวัด จากประชากรที่ต้องศึกษา แล้วจัดอย่างสุ่มออกเป็นสองกลุ่ม คือกลุ่มทดลองและกลุ่มควบคุม สองกลุ่มนี้จัดว่าเป็นอิสระแก่กัน ตัวอย่าง เช่น ประชากรที่ศึกษานักเรียนชั้น ป.6 ของโรงเรียนแห่งหนึ่ง จำนวน 400 คน ผู้วิจัยสุ่มมาศึกษา 70 คน โดยจัด

สุ่มเป็นสองกลุ่ม กลุ่มละ 35 คน

ข. กลุ่มตัวอย่างสองกลุ่ม สุ่มมาจากประชากรสองกลุ่ม ตัวอย่างเช่น ประชากรชาย 200 คน สุ่มมา 30 คน ประชากรหญิง 180 คน สุ่มมา 30 คน สองกลุ่มที่สุ่มมานี้ถือว่าเป็นอิสระแก่กัน

ในที่นี้จะกล่าวถึงกลุ่มตัวอย่างที่เป็นอิสระแก่กันสองกรณีย่อยคือ กรณีที่ assume ว่าความแปรปรวนของประชากรทั้งสองเท่ากัน กับกรณีที่ assume ว่าความแปรปรวนของประชากรทั้งสองไม่เท่ากัน

5.2.1 กรณีที่ assume ว่าความแปรปรวนของประชากรทั้งสองเท่ากัน คำนวณค่า t

จากสูตร

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2} \right) \left(\frac{n_1 + n_2}{n_1 n_2} \right)}} \quad \text{.....(3)}$$

เมื่อ t แทน ค่าสถิติที่จะใช้เปรียบเทียบกับค่าวิกฤตในการแจกแจงแบบ t เพื่อทราบ ความนัยสำคัญ

$\bar{X}_1, \bar{X}_2$ แทน ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1 และกลุ่มที่ 2 ตามลำดับ

S_1^2, S_2^2 แทน ความแปรปรวนของกลุ่มตัวอย่างกลุ่มที่ 1 และกลุ่มที่ 2 ตามลำดับ

n_1, n_2 แทน จำนวนกลุ่มตัวอย่างในกลุ่มที่ 1 กลุ่มที่ 1 และกลุ่มที่ 2 ตามลำดับ

• ค่า Degree of Freedom (df) ในกรณีนี้เท่ากับ $n_1 + n_2 - 2$

5.2.2 กรณีที่ assume ว่าความแปรปรวนของประชากรทั้งสองกลุ่มไม่เท่ากันคำนวณค่า t

จากสูตร

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad \text{.....(4)}$$

กรณีนี้ถ้ากลุ่มตัวอย่างมีน้อยจะใช้ค่า df จากสูตร

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} \right)^2}{\frac{\left(\frac{S_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2} \right)^2}{n_2 - 1}} \quad \text{.....(5)}$$

6. สถิติที่ใช้ทดสอบสมมุติฐานเกี่ยวกับความแตกต่างระหว่างค่าเฉลี่ยของ 3 กลุ่มขึ้นไป

ใช้การวิเคราะห์ความแปรปรวน (Analysis of Variance) ซึ่งใช้ตัวย่อว่า ANOVA หรือ F-test การวิเคราะห์ความแปรปรวนใช้สำหรับทดสอบสมมุติฐานหลัก (H_0) ที่ว่าค่าเฉลี่ยของกลุ่มตัวอย่างทุกกลุ่มมาจากประชากรที่มีค่าเฉลี่ยเท่ากัน การวิเคราะห์ความแปรปรวนมีหลายวิธีขึ้นกับการออกแบบทางเดียว (One-Way ANOVA) กรณีที่มีตัวแปรอิสระสองตัว เรียกว่าการวิเคราะห์ความแปรปรวนสองทาง (Two-Way ANOVA) กรณีที่มีตัวแปรอิสระ 3 ตัวเรียกว่า การวิเคราะห์ความแปรปรวนสามทาง (Three-Way ANOVA)

การวิเคราะห์ความแปรปรวนเป็นการทดสอบสมมุติฐานโดยพิจารณาภาพรวม ไม่ได้พิจารณาคู่ใดคู่หนึ่ง (เว้นแต่กรณีที่ใช้กับกลุ่มซึ่งจะให้ผลเหมือนกับการใช้ t-test) ถ้าผลวิเคราะห์พบว่า F ไม่มีนัยสำคัญทางสถิติ (ค่า F น้อยกว่า ค่าวิกฤต) ก็สรุปได้ว่า ค่าเฉลี่ยของกลุ่มตัวอย่างทุกกลุ่มมาจากประชากรที่มีค่าเฉลี่ยไม่แตกต่างกัน แต่ถ้าพบว่า F มีนัยสำคัญทางสถิติ (ค่า F ที่คำนวณเท่ากับหรือมากกว่าค่าวิกฤต) ก็สรุปได้ว่า ค่าเฉลี่ยของกลุ่มตัวอย่างเหล่านั้นมาจากประชากรที่มีค่าเฉลี่ยแตกต่างกัน เมื่อผู้วิจัยต้องการทราบว่ามีความแตกต่างกันก็ต้องทำการทดสอบด้วยสถิติ ประเภท Multiple Comparison ต่อไป ซึ่งมีหลายวิธีสามารถเลือกใช้ตามความเหมาะสม เช่นวิธีของ Tukey วิธีของ Scheffe วิธีของ Newman-Keuls วิธีของ Duncan เป็นต้น (รศ. นิภา ศรีโพธิ์โรจน์ เคยเขียนหลักการ เลือกใช้ลงในวารสารเมื่อปีที่ 1 ฉบับที่ 1 พ.ศ. 2535)

ในการวิเคราะห์ความแปรปรวน จะคำนวณค่า F และค่า F มีค่าใช้ส่วนตัวไหน ขึ้นกับการออกแบบในการวิจัย

การวิเคราะห์ความแปรปรวนทางเดียวหรือการวิเคราะห์องค์ประกอบเดียว คือการวิเคราะห์ความแปรปรวนในกรณีที่มีตัวแปรอิสระเพียงตัวเดียว และมีตัวแปรตามตัวเดียวตัวแปรอิสระอาจจะเป็นตัวแปรเชิงคุณลักษณะ (Qualitative) เช่น วิธีสอน แบบของการบำบัดทางจิตวิทยา แบบของการอบรมเลี้ยงดู เป็นต้น หรือตัวแปรเชิงปริมาณ (Quantitative) เช่น จำนวนการเสริมแรงทางบวก ปริมาณปุ๋ยที่ให้กับพืช ปริมาณยาที่ให้กับคนป่วยทางจิต เป็นต้น ตัวแปรอิสระนั้นจะมีระดับหรือประเภทย่อยอยู่จำนวนหนึ่ง เช่น ด้านตัวแปรเชิงคุณลักษณะ วิธีสอน อาจมี 3 วิธี แบบของการบำบัดทางจิตวิทยามี 4 แบบ แบบของการอบรมเลี้ยงดูมี 3 แบบ ด้านตัวแปรเชิงปริมาณจำนวนการเสริมแรงทางบวกอาจมี 3 ขนาด ปริมาณปุ๋ยที่ให้กับพืชมีต่างกัน 5 ขนาด ปริมาณยาที่ให้กับคนป่วย ทางจิตมี 4 ระดับ เป็นต้น

ตัวอย่างลักษณะการวิจัยที่ใช้การวิเคราะห์ความแปรปรวนทางเดียว

การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนวิชา.....ของกลุ่มที่เรียนด้วยวิธีสอนที่ต่างกัน
.....วิธี (นิยมใช้ตั้งแต่ 3 วิธีขึ้นไป)

การเปรียบเทียบคุณลักษณะต่าง ๆ เช่น ความซื่อสัตย์ ความพร้อมในการอ่าน

เจตคติต่อเรื่องใดเรื่องหนึ่ง วินัยในตนเอง ฯลฯ (โดยเลือกคุณลักษณะอย่างใดอย่างหนึ่ง) ระหว่างกลุ่ม..... (เช่นกลุ่มที่ได้รับการอบรมเลี้ยงดูต่างกัน 3 แบบ)

การวิเคราะห์ความแปรปรวนทางเดียวมีหลายวิธีที่สำคัญได้แก่ กรณีการออกแบบการวิจัยแบบ CR-K, RB-K และ RM-K

การวิเคราะห์ความแปรปรวนสองทางคือ การวิเคราะห์ความแปรปรวนในกรณีที่มีตัวแปรอิสระหรือตัวแปรจัดกระทำ (Treatment) หรือปัจจัย (Factor) สองตัว ตัวแปรตามมีตัวเดียว ข้อดีของการวิเคราะห์แบบนี้ก็คือ นอกจากจะสามารถประเมินผลตัวแปรอิสระสองตัวไปพร้อม ๆ กันแล้วยังทราบผลรวมหรือที่เรียกว่า ปฏิสัมพันธ์ (Interaction) ระหว่างตัวแปรทั้งสองด้วย

ปฏิสัมพันธ์ระหว่างตัวแปรหมายถึง การที่ผลของตัวแปรหนึ่งจะเกิดขึ้นอย่างชัดเจนกับระดับหรือประเภทของอีกตัวแปรหนึ่ง เช่น ข้าวพันธุ์หนึ่งให้ผลสูงกว่าอีกพันธุ์หนึ่งในปุ๋ยชนิดหนึ่งมากกว่าปุ๋ยชนิดหนึ่ง หรือวิธีสอนที่ต่างกันต่อเด็กแต่ละประเภท เป็นต้น

ตัวอย่างลักษณะการวิจัยที่ใช้การวิเคราะห์ความแปรปรวนสองทาง

ก. การเปรียบเทียบความคิดสร้างสรรค์ของนักเรียนที่มีความสามารถทางการเรียนต่างกัน และได้รับการสอนด้วยวิธีที่ต่างกัน

การวิจัยนี้มีตัวแปรอิสระสองตัวคือ ความสามารถทางการเรียน (อาจแบ่งเป็น 3 ระดับ เช่น ความสามารถทางการเรียนสูง ปานกลาง ต่ำ) และวิธีสอน (อาจใช้ 3 วิธี เช่น วิธีสอนแบบระดมความคิด วิธีสอนแบบกลุ่มสัมพันธ์ วิธีสอนแบบสืบเสาะ) ตัวแปรตามมีตัวเดียวคือ ความคิดสร้างสรรค์ ตัวแปรตามอาจเป็นตัวอื่น เช่น ผลสัมฤทธิ์ทางการเรียน วินัยในตนเอง เจตคติต่อวิชาที่เรียน ฯลฯ อย่างใดอย่างหนึ่ง

ข. การเปรียบเทียบวินัยในตนเอง ของนักเรียนที่มาจากครอบครัวที่ผู้ปกครองประกอบอาชีพต่างกัน และมีฐานะทางสังคมต่างกัน

การวิจัยนี้มีตัวแปรอิสระสองตัวคือ อาชีพผู้ปกครอง (อาจมี 4 ประเภท เช่น อาชีพค้าขาย รับราชการ เกษตร รับจ้าง) และฐานะทางสังคม (อาจมี 3 ระดับ คือสูง ปานกลาง ต่ำ) ตัวแปรตามมีตัวเดียวคือวินัยในตนเอง

ตัวแปรตามอาจเป็นตัวอื่น เช่น ผลสัมฤทธิ์ทางการเรียน ความคิดสร้างสรรค์แรงจูงใจใฝ่สัมฤทธิ์ มโนคติเกี่ยวกับตนเอง ฯลฯ อย่างใดอย่างหนึ่ง

การวิเคราะห์ความแปรปรวนสองทางมีหลายวิธี เช่นแบบ CRF-a, b และ RMF-a, b เป็นต้น

7. สถิติที่ใช้ทดสอบสมมติฐานเกี่ยวกับความแตกต่างและความสัมพันธ์ กรณีข้อมูลอยู่ในรูปของความถี่ นิยมใช้ไคสแควร์ (Chi - Square) ซึ่งมีหลายกรณีดังจะกล่าวถึง 2 กรณี คือกรณีกลุ่มตัวอย่างเดียวกันกรณีกลุ่มตัวอย่างตั้งแต่ 2 กลุ่มขึ้นไป ที่เป็นอิสระแก่กัน สัญลักษณ์ของไคสแควร์คือ

7.1 กรณีกลุ่มตัวอย่างกลุ่มเดียว คำนวณจากสูตร

$$X^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \dots\dots\dots(6)$$

เมื่อ X^2 แทน ค่าสถิติที่ใช้เปรียบเทียบค่าวิกฤตในการแจกแจงแบบไคสแควร์เพื่อ
 ทราบความนัยสำคัญ

O_i แทน ความถี่ที่เกิดขึ้นจริง (Observed Frequency) ในกลุ่มหรือประเภท
 ที่ i

E_i แทน ความถี่ตามทฤษฎี หรือความถี่ที่คาดหวัง (Expected Frequency)
 ในกลุ่มหรือประเภทที่ i

K แทนจำนวนประเภทย่อยของกลุ่มนั้นกรณีนี้มี $df = K - 1$

7.2 กรณีกลุ่มตัวอย่างตั้งแต่ 2 กลุ่มขึ้นไป คำนวณสูตร

$$X^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \dots\dots\dots(7)$$

เมื่อ X^2 แทน ค่าสถิติที่ใช้เปรียบเทียบค่าวิกฤตในการแจกแจงแบบไคสแควร์
 เพื่อทราบความนัยสำคัญ

O_{ij} แทน ความถี่ที่เกิดขึ้นจริง (Observed Frequency) ในแถวที่ i คอลัมน์ที่
 j

E_{ij} แทน ความถี่ตามทฤษฎี หรือความถี่ที่คาดหวัง (Expected Frequency)
 ในแถวที่ i คอลัมน์ที่ j

r แทน จำนวนแถวหรือจำนวนกลุ่ม

c แทน จำนวนคอลัมน์หรือประเภทของกลุ่มตัวอย่าง

หลังจากคำนวณ X^2 จากสูตรนี้แล้ว จะนำค่า X^2 นี้ไปเปรียบเทียบกับค่าวิกฤตของไคสแควร์ จาก
 ตาราง โดยใช้ df เท่ากับ $(r-1)(c-1)$ ถ้า X^2 ที่คำนวณได้มีค่าเท่ากับหรือมากกว่าค่าวิกฤต แสดงว่า
 X^2 นั้นมีนัยสำคัญ ก็จะปฏิเสธ H_0 ยอมรับ H_A

๗ เอกสารอ้างอิง

- บุญชม ศรีสะอาด การวิจัยเบื้องต้น มหาสารคาม : ภาควิชาการวัดผลและวิจัยการศึกษา
คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ มหาสารคาม, 2536, 168 หน้า.
- “_____” วิธีการทางสถิติสำหรับการวิจัย เล่ม 1 มหาสารคาม : ภาควิชาวัดผลและวิจัย
การศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ มหาสารคาม, 2532, 368 หน้า.
- ไพเราะ อ่อนเกตุผล และคณะ. การพัฒนาหลักสูตร : ทฤษฎีและการปฏิบัติ, แปลจาก Curriculum
Development : Theory & Practice ของ Hilda Taba, เอกสารประกอบการเรียน ทฤษฎี
หลักสูตร. 2536.
- บุหงา วัฒนะ และคณะ. ความเข้าใจหลักสูตร แปลจาก Curriculum : A Comprehensive Introduction
ของ John D. McNeil, เอกสารประกอบการเรียน ทฤษฎีหลักสูตร, 2528.
- บุญมี พันธุ์ไชย และคณะ. หลักสูตร : การออกแบบการพัฒนา แปลจาก Curriculum : Design and
Development ของ David Pratt, เอกสารประกอบการเรียนทฤษฎีหลักสูตร, 2538.
- สุमितต คุณานุการ. หลักสูตรและการสอน, พระนคร กรุงเทพมหานครการพิมพ์, 2518.
- Bocon, Francis L. and Others. The Administration of The Modern Secondary School.
New York : The Macmillan Co. 1953.
- Taba, Hilda, Curriculum Development : Theory and Practice. New York : Harcourt Brace &
World Inc. 1962.
- Cronbach, Lee J. Designing evaluations of education and social program. San Francisco
Jossey-Bass 1983.
- Tyler, Ralph, Basic Principles of Curriculum and Instruction. University of Chicago Press,
1970.