บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การวิจัยเป็นการศึกษาค้นคว้าหาความจริงในสาขาวิชาต่างๆอย่างมีเป้าหมาย มีวัตถุประสงค์ที่แน่นอน และเป็นการกระทำอย่างมีระบบแบบแผนโดยถูกต้องตามหลักวิชาเพื่อให้ได้ข้อความรู้ที่เชื่อถือได้ การทำวิจัยจึงจำเป็นที่จะต้องวางแผนและตรวจสอบข้อมูลทุกขั้นตอน และขั้นตอนสำคัญขั้นตอนหนึ่งคือการวิเคราะห์ข้อมูลซึ่งงานวิจัยส่วนใหญ่จำเป็นต้องอาศัยระเบียบวิธีการทางสถิติช่วยในการสรุปผล และตอบคำถามตามที่ตั้งจุดมุ่งหมายไว้

การวิเคราะห์ข้อมูลในการวิจัยทางสังคมศาสตร์ มักที่จะตอบคำถามการวิจัยเพื่อบรรยายลักษณะของตัวแปรใน 4 ลักษณะคือ เพื่ออธิบายลักษณะหรือสภาพของตัวแปรเพื่อแสดงให้เห็นความสัมพันธ์ระหว่างตัวแปรต่างๆ เพื่ออธิบายถึงความสัมพันธ์เชิงเหตุและผลระหว่างตัวแปร และเพื่อเปรียบเทียบความแตกต่างระหว่างกลุ่มประชากรต่างๆ สถิติวิเคราะห์ที่ใช้ในการวิจัยตามหน้าที่ดังกล่าวนั้นได้รับการพัฒนาเพื่อใช้ให้เหมาะสมกับสถานการณ์ต่างๆกันซึ่งผู้วิจัยจำเป็นต้องรู้จักเลือกใช้สถิติทดสอบให้เหมาะสมกับข้อมูล

สถิติทดสอบ(test statistics)ที่ใช้เป็นพื้นฐานกันมากก็มี Z-test,t-test, X^2-test และ F- test ซึ่งสถิติทดสอบเหล่านี้เป็นสถิติที่ใช้ในการทดสอบสมมติฐานเพื่อที่จะอ้างอิงสรุปผลค่าสถิติจากกลุ่มตัวอย่างไปยังค่าพารามิเตอร์ของประชากร การทดสอบสมมติฐานเป็นการทดสอบเพื่อหาสรุปค่าพารามิเตอร์ของกลุ่มประชากรต่างๆ โดยส่วนมากแล้วจะเป็นการทดสอบเกี่ยวกับค่าเฉลี่ย ทดสอบเกี่ยวกับสัดส่วน ทดสอบเกี่ยวกับการแจกแจงความถี่ ทดสอบเกี่ยวกับความแปรปรวน และทดสอบเกี่ยวกับความสัมพันธ์ เป็นต้น

สถิติทดสอบแต่ละแบบจะมีข้อตกลงเบื้องต้น (Assumption) ในการใช้แตกต่างกัน ซึ่งการฝ่าฝืนข้อตกลงเบื้องต้นของสถิติทดสอบ อาจทำให้ได้ผลสรุปผิดพลาดไม่ตรงกับความจริง สถิติทดสอบแต่ละแบบมีข้อตกลงเบื้องต้นไว้น้อยบ้างมากบ้าง และจากการศึกษาค้นคว้า พบว่าการฝ่าฝืนข้อตกลงบางข้อของสถิติทดสอบไม่ได้ทำให้เกิดผลกระทบต่อการวิเคราะห์ข้อมูลมากนัก

แต่บางข้อของข้อตกลงเบื้องต้นผู้วิจัยจะฝ่าฝืนไม่ได้เพราะจะทำให้เกิดความคลาดเคลื่อนในการสรุปผล โดยเฉพาะอย่างยิ่งการใช้สถิติวิเคราะห์ในการทดสอบสมมติฐาน ความสำคัญของการใช้สถิติวิเคราะห์จึงมีผลต่อความเชื่อถือได้ในการอ้างอิงไปยังประชากร

ในงานวิจัยที่ต้องการทดสอบสมมติฐานเพื่อเปรียบเทียบค่าเฉลี่ยของประชากร โดยใช้ค่าสังเกตจากกลุ่มตัวอย่าง การเปรียบเทียบค่าเฉลี่ยของกลุ่มตัวอย่าง 2 กลุ่ม ส่วนมากนิยมใช้สถิติทดสอบ ที (t- test) ถ้าสำหรับกรณีกลุ่มตัวอย่างมากกว่าสองกลุ่ม เทคนิคการทดสอบสมมติฐานทางสถิติที่นำมาใช้กันมากคือ สถิติทดสอบเอฟ (F-test) โดยผ่านระเบียบวิธีวิเคราะห์ที่เรียกกันว่า การวิเคราะห์ความแปรปรวน(Analysis of Variance : ANOVA) และเทคนิคการวิเคราะห์ความแปรปรวนดังกล่าว มีผู้ให้มโนทัศน์หลายๆด้าน ดังนี้

การวิเคราะห์ความแปรปรวนเป็นวิธีการทางคณิตศาสตร์ในการแบ่งผลรวมกำลังสองทั้งหมด (Partition a total sum of squares)ออกเป็นส่วนประกอบย่อยๆ ซึ่งสัมพันธ์กับแหล่งของความแปรปรวน เพื่อใช้ประโยชน์ในการวิเคราะห์ข้อมูลที่ได้รับในเชิงปริมาณ (Steel, 1960)

การวิเคราะห์ความแปรปรวน จะกล่าวถึงทฤษฎีและวิธีการสำหรับวิเคราะห์ข้อมูลที่ได้จากการทดลองรวมขั้นตอนต่างๆในการวิเคราะห์ จุดมุ่งหมายทั่วไปของ ANOVA ก็เพื่อตัดสินว่าองค์ประกอบใดของการทดลองมีผล(effect)มากที่สุดต่อข้อมูลในการทดลอง และเป็นการให้ข้อมูลเชิงปริมาณเกี่ยวกับความสำคัญขององค์ประกอบต่างๆ และระดับของ effect ที่ส่งให้เกิดความแตกต่างกันของตัวแปร (Lee, 1975)

การวิเคราะห์ความแปรปรวนที่ใช้กันมากนั้น เป็นวิธีการวิเคราะห์เพื่อเปรียบเทียบค่าเฉลี่ยหลายๆกลุ่มที่มาจากประชากรซึ่งมีการแจกแจงแบบปกติ เมื่อไม่ทราบค่าส่วนเบี่ยงเบนมาตรฐาน โดยการพิจารณาประชากร k กลุ่ม ที่มีการแจกแจงแบบปกติ มีค่าเฉลี่ยเป็น $1,2,\ldots,k$และมีส่วนเบี่ยงเบนมาตรฐานเท่ากัน (Common Standard deviation σ^2) (Berman, 1971)

Kirk(1982)ได้จำแนกข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนไว้ 2 ชุด คือ ชุดที่ 1 เป็นข้อตกลงเบื้องต้นของสถิติทดสอบเอฟ(F - Assumption) ได้แก่

1. ค่าสังเกตมาจากประชากรที่มีการแจกแจงปกติ
2. ค่าสังเกตเป็นกลุ่มตัวอย่างที่ได้รับการสุ่มมาจากประชากร
3. เศษและส่วนของอัตราส่วน F เป็นค่าประมาณจากประชากรที่มีความแปรปรวนเท่ากัน
4. เศษและส่วนของอัตราส่วน F เป็นอิสระต่อกัน

ชุดที่ 2 เป็นข้อตกลงเบื้องต้นของโมเดล (Model Assumption)จากโมเดลของสมการ $X_{ij} = \mu + \alpha_j + e_{ij}$ โมเดลนี้แสดงว่าคะแนนของ observe X_{ij} นั้นรวมความแปรปรวนทุกแหล่งที่จะมีผลทำให้ค่าที่แท้จริงของ X_{ij} แปรเปลี่ยนไป โดยถือว่าแหล่งของความแปรปรวนจะมีเพียง 2 แหล่ง คือ Systematic error(α_j) และ Random error (e_{ij}) และโมเดลของการใช้การวิเคราะห์ความแปรปรวน ต้องมีข้อตกลงเบื้องต้นเช่นเดียวกันกับข้อตกลงเบื้องต้นสถิติทดสอบ F

ในการใช้สถิติทดสอบเอฟสำหรับวิเคราะห์ความแปรปรวน มักพบว่าข้อมูลที่ได้จากการศึกษาอาจจะไม่เป็นไปตามข้อตกลงเบื้องต้นข้อใดข้อหนึ่ง ดังที่ Cochran และ Cox (1957) กล่าวว่าในทางปฏิบัติจริงเรามักจะไม่แน่ใจว่า ข้อมูลที่ได้มาจะเป็นไปตามข้อตกลงเบื้องต้นทั้งหมด และบ่อยครั้งที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นบางข้อ ซึ่งลักษณะดังกล่าวนี้ Snedecor(1973) กล่าวว่ามีคำถามที่จะต้องตอบอยู่ 2 ข้อ คือ ข้อมูลที่ใช้ทดสอบตรงตามข้อตกลงเบื้องต้นหรือไม่ และถ้าไม่เป็นไปตามข้อตกลงเบื้องต้นแล้วการใช้สถิติทดสอบเอฟจะเป็นอย่างไร สำหรับการตอบคำถามกรณีที่ขาดข้อตกลงเบื้องต้นด้านการแจกแจงของประชากร มีผู้ศึกษาและกล่าวไว้ดังนี้

Scheffé (1970) กล่าวว่า ผลของการฝ่าฝืนข้อตกลงเบื้องต้นของความเป็นปกติของการแจกแจงมีผลเพียงเล็กน้อยต่อการอ้างอิงด้วยค่าเฉลี่ย แต่จะมีผลมากต่อการอ้างอิงด้วยค่าความแปรปรวน

สำหรับกรณีข้อมูลที่ใช้ในการทดสอบ ไม่เป็นไปตามข้อตกลงเบื้องต้นด้านความเป็นเอกพันธ์ของความแปรปรวน (Homogeneity of Variance) มีผู้ศึกษาและกล่าวไว้ดังนี้

Cochran และ Cox (1957)กล่าวว่า ถ้าสถิติทดสอบเอฟไม่เป็นไปตามข้อตกลงเบื้องต้นจะเกิดผลกระทบทั้งระดับความมีนัยสำคัญและความไวของการทดสอบ(Sensitivity) ถ้าผู้วิจัยคิดว่าจะทำการทดสอบ α =.05 แต่ความเป็นจริงแล้วค่าของ α อาจจะเป็น .08

Glass,Peckham และ Sanders (1972) ได้สรุปผลจากการศึกษา เมื่อความแปรปรวนแตกต่างกันและขนาดของกลุ่มตัวอย่างไม่เท่ากัน ในการใช้สถิติทดสอบเอฟ จากการศึกษาของ Hsu (1938),Scheffe (1959), Pratt (1964),Box (1954),Boneau (1960),Vamder Vaart (1961) ดังนี้

1. เมื่อกลุ่มตัวอย่างที่มีจำนวนน้อย ถูกสุ่มมาจากประชากรที่มีความแปรปรวนมาก ระดับความมีนัยสำคัญที่เป็นจริง จะมากกว่าระดับความมีนัยสำคัญที่กำหนด (α)

2. เมื่อกลุ่มตัวอย่างที่มีจำนวนน้อย ถูกสุ่มมาจากประชากรที่มีความแปรปรวนน้อย ระดับนัยความมีสำคัญที่เป็นจริง จะน้อยกว่าระดับความมีนัยสำคัญที่กำหนด (α)

Kirk (19828) กล่าวว่า การแจกแจงเอฟจะเปลี่ยนแปลงไป เมื่อฝ่าฝืนข้อตกลงเบื้องต้น เกี่ยวกับความเป็นเอกพันธ์ของความแปรปรวน (Homogeneity ofVariance) สำหรับขนาดของกลุ่มตัวอย่างเท่ากัน (Box,1954 a,b; Cochran,1947;Norton as Cited by Lindquist, 1953) เมื่อความแปรปรวนของประชากรแตกต่างกันมาก Roganและ Keselman (1977) พบว่า ค่าระดับความมีนัยสำคัญที่แท้จริง อาจจะใหญ่กว่าค่าระดับความมีนัยสำคัญที่กำหนด สำหรับขนาดของกลุ่มตัวอย่างไม่เท่ากันการฝ่าฝืนข้อตกลงเบื้องต้นข้อนี้จะมีผลอย่างมากต่อการกำหนดระดับความมีนัยสำคัญ

Glass (1970)ได้สรุปถึงเรื่องวิวิธพันธ์ของความแปรปรวน(Heterogeneous of Variance) จากการศึกษาของ(Hsu, 1938; scheffe, 1959;Lindquist, 1953; Boneau, 1960; Cochran, 1947; Godard และ Lindquist,1940;Horsnell, 1953; Welch, 1937) ไว้ดังนี้

1. เมื่อขนาดของกลุ่มตัวอย่างเท่ากัน ผลจากการไม่เท่ากันของความแปรปรวนของประชากรแต่ละกลุ่มตัวอย่างที่มีต่อระดับความมีนัยสำคัญของสถิติทดสอบเอฟ อาจมีเพียงเล็กน้อย

2. เมื่อขนาดของกลุ่มตัวอย่างไม่เท่ากัน และความแปรปรวนของประชากรแต่ละกลุ่มตัวอย่างไม่เท่ากันจะเกิดผลดังนี้

2.1 เมื่อกลุ่มตัวอย่างมีขนาดเล็ก และความแปรปรวนของประชากรมีขนาดใหญ่ ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (Type I Error) จะใหญ่กว่า α หรือกล่าวได้ว่า ผลกระทบของกรณีนี้จะทำให้การแจกแจงของเอฟเลื่อนไปทางขวา

2.2 เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ และความแปรปรวนของประชากรมีขนาดเล็ก ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (Type I Error) จะเล็กกว่า α หรือกล่าวได้ว่า ผลกระทบของกรณีนี้จะทำให้การแจกแจงของเอฟเลื่อนไปทางซ้าย

จากการศึกษาการใช้สถิติทดสอบ F ดังกล่าว แสดงให้เห็นว่าเมื่อขนาดของกลุ่มตัวอย่างแต่ละกลุ่มเท่ากัน ถ้ามีการฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ของความแปรปรวนแล้วสามารถที่จะใช้สถิติทดสอบ F ได้ ทั้งนี้เพราะสถิติทดสอบ F จะยังมีความแกร่ง (Robustness)ต่อการฝ่าฝืนข้อตกลงเบื้องต้นข้อนี้ แต่ถ้าขนาดของกลุ่มตัวอย่างในแต่ละกลุ่มไม่เท่ากันแล้วการฝ่าฝืนข้อตกลงเบื้องต้นข้อนี้จะมีผลอย่างมากในการใช้สถิติทดสอบ F ดังนั้นในงานการวิจัยเชิงทดลอง จึงมีการวางแผนการวิจัยเพื่อให้ขนาดของกลุ่มตัวอย่างแต่ละกลุ่มเท่ากันเพราะสามารถที่จะใช้สถิติทดสอบเอฟได้ โดยไม่ต้องมีการทดสอบการเท่ากันของความแปรปรวน แต่ในงานการวิจัยเชิงทดลองบางครั้งอาจจะมีการสูญหายของข้อมูลทำให้ขนาดของกลุ่มตัวอย่างแต่ละกลุ่มไม่เท่ากัน และในปัจจุบันได้มีการนำการวิเคราะห์ความแปรปรวนมาใช้

ในงานการวิจัยเชิงบรรยายเป็นจำนวนมาก ซึ่งการวิจัยเชิงบรรยายขนาดของกลุ่มตัวอย่างแต่ละกลุ่มมักจะไม่เท่ากัน ซึ่ง Box (1953)ได้เสนอแนะไว้ว่าก่อนที่จะใช้วิธีการวิเคราะห์ความแปรปรวน เพื่อทดสอบการเท่ากันของค่าเฉลี่ยของประชากรที่มากกว่า 2 กลุ่ม ควรจะมีการตรวจสอบเกี่ยวกับการเท่ากันของความแปรปรวนเสียก่อนเพื่อที่จะไม่เกิดข้อสรุปที่ผิด ดังนั้นในการวิจัยต่างๆก่อนที่จะใช้สถิติทดสอบเอฟ เมื่อขนาดของกลุ่มตัวอย่างไม่เท่ากัน ควรจะได้มีการทดสอบข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกพันธ์ของความแปรปรวน (Homogeneity of Variance) เสียก่อน

Game และคณะ (1972) ได้กล่าวว่า ถ้านักวิจัยไม่สนใจที่จะตรวจสอบการเท่ากันของความแปรปรวนแล้ว เมื่อผลการวิเคราะห์สรุปได้ว่า ค่าเฉลี่ยของประชากรแตกต่างกันแล้ว และผู้วิจัยต้องการที่จะทดสอบภายหลังโดยการเปรียบเทียบพหุ (Multiple Comparison) โดยใช้วิธีการต่างๆ เช่น วิธีของ Tukey วิธีของ Newman - Keuls วิธีของ Scheffe หรือวิธีของ Duncan ซึ่งวิธีการเปรียบเทียบดังกล่าวนี้มีข้อตกลงเบื้องต้นเกี่ยวกับการเท่ากันของความแปรปรวนของประชากรที่นำมาทดสอบ แต่เนื่องจากยังไม่มีหลักฐานยืนยันเกี่ยวกับความเท่ากันของความแปรปรวนของประชากร ดังนั้นผลของการเปรียบเทียบพหุเหล่านี้ อาจได้ผลสรุปที่ผิดพลาดได้เช่นกัน

สถิติทดสอบที่ใช้ในการทดสอบการเท่ากันของความแปรปรวนของประชากรที่มากกว่า 2 กลุ่ม มีสถิติที่ใช้ได้หลายแบบ สถิติพาราเมตริก (Parametric Statistics) ที่รู้จักกันมากและนิยมใช้ในการทดสอบได้แก่ สถิติทดสอบบาร์ตเลต (Bartlett's test) ซึ่ง Bartlett (1937) ได้พัฒนามาจากสถิติทดสอบ Neyman ที่มีการแจกแจงแบบไคสแควร์ Glass และ Hopkin (1984) ได้กล่าวถึงสถิติทดสอบบาร์เลตว่าเป็นสถิติทดสอบที่ให้ผลการทดสอบที่ถูกต้อง ถ้ากลุ่มของประชากรมีลักษณะการแจกแจงแบบปกติ แต่ถ้ามีการแจกแจงที่มีค่าความโด่งมากกว่า 3 หรือมีลักษณะการแจกแจงที่มีความเบ้แบบต่างๆแล้ว ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1ของสถิติทดสอบบาร์ตเลต จะมีค่ามากกว่าระดับนัยสำคัญ α ที่กำหนด ประชากรมีลักษณะการแจกแจงแบบที่มีความโด่งน้อยกว่า 3 ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จะมีค่าน้อยกว่าระดับความมีนัยสำคัญที่กำหนด และ Glass ได้แนะนำการใช้สถิติทดสอบบาร์ตเลตว่า ควรจะใช้ทดสอบเมื่อขนาดของกลุ่มตัวอย่างแต่ละกลุ่มมีค่ามากกว่า 5 Gartside(1972)ได้ศึกษาอำนาจการทดสอบของสถิติทดสอบที่ใช้ทดสอบการเท่ากันของความแปรปรวนของประชากรที่มากกว่า 2 กลุ่มโดยศึกษาจาก สถิติทดสอบบาร์ตเลต สถิติทดสอบบาร์ตเลตที่ปรับปรุง สถิติทดสอบโคแครน สถิติทดสอบฮาร์ตเลย์ สถิติทดสอบแคดเวล สถิติทดสอบ Log ANOVA I,Log ANOVA II และสถิติทดสอบ Log ANOVA III ผลการวิจัย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สรุปว่า สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงกว่าสถิติทดสอบอื่นๆ เมื่อประชากรมีลักษณะการแจกแจงแบบปกติและขนาดของกลุ่มตัวอย่างทุกกลุ่มเท่ากับ 16 ซึ่งสอดคล้องกับผลการวิจัยของ อารยา กุลานุช(2528) ที่ศึกษาอำนาจการทดสอบของสถิติทดสอบบาร์ตเลต สถิติทดสอบแคดเวล และสถิติทดสอบเลอวีนที่ปรับปรุง โดยศึกษาจากประชากรที่มีการแจกแจงแบบปกติและไม่เป็นแบบปกติ ขนาดของกลุ่มตัวอย่างทุกกลุ่มเท่ากันคือเท่ากับ 5,11,21,31 และ 61 ผลการวิจัยพบว่า สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงกว่าสถิติทดสอบอื่นๆ และ Glass (1970) ได้กล่าวว่า สถิติทดสอบบาร์ตเลตเป็นสถิติทดสอบที่มีความไวเมื่อมีการฝ่าฝืนข้อตกลงเบื้องต้นของประชากรที่มีการแจกแจงเป็นแบบปกติแต่ถ้าประชากรมีการแจกแจงไม่เป็นแบบปกติแล้วจะมีค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มากกว่า α ที่ผู้วิจัยกำหนดและ Glass ยังได้กล่าวถึงสถิติทดสอบบาร์ตเลต ว่าเป็นสถิติทดสอบที่มีความยุ่งยากซับซ้อนในการคำนวณมาก เพราะจะต้องมีการคำนวณจากค่าlogarithmของความแปรปรวนของกลุ่มตัวอย่าง ดังนั้นจึงได้เสนอแนะให้ใช้สถิติทดสอบเลอวีน (Levene's test) ซึ่งเป็นสถิติที่มีลักษณะการแจกแจงแบบเอฟ(F-Distributiion)

Levene (1960) ได้เสนอสถิติทดสอบการเท่ากันของความแปรปรวนโดยมีวิธีการคำนวณเช่นเดียวกับ การวิเคราะห์ความแปรปรวนทางเดียว (one way ANOVA) โดยการแทนค่าสังเกตแต่ละค่าด้วยค่าสัมบูรณ์ของส่วนเบี่ยงเบนระหว่างค่าสังเกตแต่ละค่ากับค่าเฉลี่ยของกลุ่มตัวอย่างนั้นๆ สถิติทดสอบเลอวีนเป็นสถิติทดสอบที่มีความแกร่งเมื่อประชากรมีการแจกแจงไม่เป็นแบบปกติ และจะมีอำนาจการทดสอบที่น่าเชื่อถือได้เมื่อขนาดของกลุ่มตัวอย่างแต่ละกลุ่มไม่น้อยกว่า 10

Brown และ Forsyth (1974) ได้ศึกษาความแกร่งของการทดสอบ(Robust) โดยศึกษาจากประชากร 2 กลุ่ม และลักษณะการแจกแจงของประชากร 4 ประเภท ของสถิติทดสอบ Levene, Jackknife และ X^2- test พบว่า สถิติทดสอบของเลอวีนเป็นสถิติทดสอบที่มีความแกร่งภายใต้การแจกแจงของประชากรที่ไม่เป็นแบบปกติ

Conover (1978) ได้เสนอแนะสถิตินอนพาราเมตริกสำหรับการทดสอบการเท่ากันของความแปรปรวนของประชากร 2 กลุ่ม หรือมากกว่า 2 กลุ่ม โดยใช้ผลรวมกำลังสองของอันดับ ซึ่งเรียกสถิติทดสอบนี้ว่า สถิติทดสอบสแควร์แรงค์ (Squared Ranks test) ซึ่งเป็นสถิติทดสอบที่มีลักษณะการแจกแจงไคสแควร์แบบนอนพาราเมตริก

ในการทดสอบการเท่ากันของความแปรปรวนของประชากร2กลุ่ม สมชัย ยืนนาน (2528)ได้ศึกษาเปรียบเทียบอำนาจของการทดสอบของสถิติทดสอบเอฟ มู้ด ซีเกลทูกี กลอทซ์ และสแควร์แรงค์ โดยศึกษาจากประชากร 2 กลุ่ม ได้ผลการวิจัยว่า เมื่อประชากรมีการแจกแจงเป็นแบบปกติ สถิติทดสอบเอฟ จะมีอำนาจการทดสอบสูงเมื่อเทียบกับสถิติ

แล้ว สถิติทดสอบสแควร์แรงค์ จะมีอำนาจการทดสอบสูงกว่าสถิติทดสอบเอฟ เมื่ออัตราส่วนของความแปรปรวนของประชากรมากกว่า 1 แต่กรณีที่อัตราส่วนของความแปรปรวนของประชากรน้อยกว่า 1 สถิติทดสอบเอฟมีอำนาจการทดสอบสูงที่สุด ถ้าประชากรมีการแจกแจงแบบยูนิฟอร์ม สถิติทดสอบสแควร์แรงค์ สถิติทดสอบกลอทซ์ มีอำนาจการทดสอบสูงเมื่อเทียบกับสถิติทดสอบอื่นๆ และในกรณีที่ประชากรมีการแจกแจงแบบเบ้ ขนาดของกลุ่มตัวอย่างเท่ากัน สถิติทดสอบเอฟ มีอำนาจการทดสอบสูงกว่าสถิติทดสอบสแควร์แรงค์ สำหรับกรณีที่ขนาดของกลุ่มตัวอย่างไม่เท่ากัน สถิติทดสอบสแควร์แรงค์ มีอำนาจการทดสอบสูงกว่า สถิติทดสอบเอฟ เมื่ออัตราส่วนของความแปรปรวนของประชากรมากกว่า 1แต่เมื่ออัตราส่วนของความแปรปรวนของประชากรน้อยกว่า 1 สถิติทดสอบเอฟมีอำนาจการทดสอบสูงกว่า และเมื่อพิจารณาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เมื่อประชากรมีการแจกแจงแบบเบ้ ปรากฏว่าความเบ้ของประชากร และขนาดของกลุ่มตัวอย่าง มีผลต่ออำนาจการทดสอบของสถิติทดสอบสแควร์แรงค์

จากผลการศึกษาและข้อเสนอแนะ การใช้สถิติทดสอบการเท่ากันของความแปรปรวนยังไม่สามารถหาหลักฐานและข้อสรุปที่ชัดเจนเกี่ยวกับการเลือกใช้สถิติทดสอบบาร์ตเลต เลอวีน และสแควร์แรงค์ เมื่อจำนวนกลุ่มตัวอย่างเป็น 3,4 และ5กลุ่ม ควรจะเลือกใช้สถิติทดสอบใดในการทดสอบการเท่ากันของความแปรปรวนของประชากร ซึ่งในการพิจารณาเลือกใช้สถิติทดสอบเราจำเป็นต้องพิจารณาถึง ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 แล้วจึงพิจารณาถึงอำนาจการทดสอบเป็นลำดับต่อไป โดยมีขั้นตอนดังนี้คือให้ความน่าจะเป็นที่จะยอมให้เกิดความคลาดเคลื่อนประเภทที่ 1ไม่เกินระดับความมีนัยสำคัญ α ที่กำหนดไว้ และเมื่อเป็นไปตามเงื่อนไขดังกล่าวแล้ว สิ่งที่จะพิจารณาเลือกสถิติทดสอบอีกก็คือ เลือกสถิติทดสอบที่มีโอกาสน้อยที่สุดที่จะยอมรับสมมติฐาน H_0 เมื่อสมมติฐาน H_0 นั้นผิด ซึ่งหมายความว่าให้มีอำนาจการทดสอบสูงสุด (Neyman 1950 cited by Derek Srisukho 1974) ซึ่งสามารถศึกษาโดยวิธีซิมูเลท (Simulate) จะทำให้ได้ข้อสรุปที่เด่นชัดภายใต้สถานการณ์คล้ายการทดลอง คือสามารถที่จะระบุหรือกำหนดขนาดของกลุ่มตัวอย่าง ความแปรปรวนและลักษณะของการแจกแจงของประชากรได้ด้วยเครื่องคอมพิวเตอร์ โดยใช้เทคนิคที่เรียกว่ามอนติคาร์โล (Monte Carlo) ซึ่งเป็นวิธีที่สามารถนำมาดำเนินการเพื่อหาผลสรุปในการทดลองแก้ปัญหาได้ในปัจจุบัน

วัตถุประสงค์ของการวิจัย

1. เพื่อเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 จากผลการทดลองของสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน และสถิติทดสอบสแควร์แรงค์ กับระดับความมีนัยสำคัญที่กำหนด

2. เพื่อเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน สถิติทดสอบสแควร์แรงค์

สมมติฐานการวิจัย

จากการศึกษาของ Gartside (1972) และ Games (1972) ได้แนะนำให้ใช้สถิติทดสอบบาร์ตเลต สำหรับกลุ่มตัวอย่างที่ได้มาอย่างสุ่มจากประชากรที่มีการแจกแจงแบบปกติ และมีขนาดของกลุ่มตัวอย่างเท่ากันทุกกลุ่ม และ Brown และ Fosythe (1974)ได้ศึกษาความแกร่งของสถิติทดสอบที่ใช้ทดสอบการเท่ากันของความแปรปรวนของประชากร 2 กลุ่ม ที่ถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ ได้เสนอแนะว่าสถิติทดสอบเลอวีนเป็นสถิติทดสอบที่แกร่งที่สุด สมชัย ยืนนาน (2528)ได้ศึกษาอำนาจการทดสอบของสถิติทดสอบเอฟกับสถิติทดสอบสแควร์แรงค์ ที่ใช้ในการทดสอบการเท่ากันของความแปรปรวนของประชากร 2 กลุ่ม ได้ข้อสรุปว่า ถ้าประชากรมีการแจกแจงแบบปกติ และขนาดของกลุ่มตัวอย่างแต่ละกลุ่มไม่เท่ากันสถิติทดสอบสแควร์แรงค์มีอำนาจการทดสอบสูงกว่าสถิติทดสอบเอฟ จากการศึกษาดังกล่าว ผู้วิจัยจึงคาดว่าสถิติทดสอบบาร์เลต เลอวีนและสแควร์แรงค์ ควรจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เท่ากับระดับนัยสำคัญที่กำหนด แต่เนื่องจากสถิติทดสอบทั้ง 3 แบบ มีลักษณะการแจกแจงและวิธีการคำนวณที่แตกต่างกัน โดยสถิติทดสอบบาร์ตเลตคำนวณจากข้อมูลเดิมหรือค่าที่สังเกตได้โดยตรง สถิติทดสอบเลอวีนคำนวณจากค่าสัมบูรณ์ของข้อมูลเดิมลบกับค่าเฉลี่ยของกลุ่มตัวอย่างนั้น สถิติทดสอบสแควร์แรงค์คำนวณจากอันดับของค่าสัมบูรณ์ของผลต่างของข้อมูลเดิมกับค่าเฉลี่ยของกลุ่ม ซึ่งการแปลงข้อมูลน่าจะทำให้อำนาจการทดสอบทั้ง 3 แบบแตกต่างกัน และจากการศึกษาของ อารยา กุลานุช (2528)ได้ศึกษาอำนาจของการทดสอบของสถิติทดสอบบาร์ตเลต เลอวีน และแคดเวล โดยศึกษาในกรณีที่กลุ่มตัวอย่าง 3 กลุ่มที่สุ่มมาจากประชากรมีการแจกแจงแบบปกติ และขนาดของกลุ่มตัวอย่างเท่ากันทุกกลุ่ม ได้ข้อสรุปว่า สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงสุด ดังนั้นจึงตั้งสมมติฐานการวิจัยดังนี้

1. สถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน และสถิติทดสอบสแควร์แรงค์สามารถควบคุมค่าความคลาดเคลื่อนประเภทที่ 1 ได้เท่ากับระดับความมีนัยสำคัญที่กำหนด

2. สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงกว่าสถิติทดสอบเลอวีน และสถิติทดสอบสแควร์แรงค์ สถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงกว่าสถิติทดสอบสแควร์แรงค์

ข้อตกลงเบื้องต้น

1. การวิจัยครั้งนี้เลือกใช้โปรแกรมย่อยสับรูทีน (Subprogram Subroutine) ที่มีหลักฐานและการศึกษาในรูปแบบการแปลงข้อมูลซึ่งมีอยู่แล้วสร้าง Random Number เพื่อสร้างลักษณะของประชากรที่ต้องการศึกษา

2. ใช้เกณฑ์ของ Cochran (1954:Cited by Ramsey 1980) เป็นเกณฑ์ในการตัดสินความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 กับระดับนัยสำคัญที่กำหนดดังนี้ ถ้าค่าความคลาดเคลื่อนประเภทที่ 1 ของสถิติทดสอบใดมีค่าตั้งแต่ 0.04 ถึง 0.06 สำหรับ α =.05 และมีค่าตั้งแต่ 0.007 ถึง 0.015 สำหรับ α =.01 จะถือว่าสถิติทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เท่ากับระดับความมีนัยสำคัญ .05 และ .01

ข้อจำกัดของการวิจัย

การวิจัยครั้งนี้ได้กำหนดขนาดของกลุ่มตัวอย่างที่ศึกษาให้มีค่าตั้งแต่ 10 ถึง 50 เพราะจากการศึกษาของ Games และคณะ (1972) ได้เสนอแนะว่า ควรใช้สถิติทดสอบเลอวีน เมื่อขนาดของกลุ่มตัวอย่างมีค่าตั้งแต่ 10 ขึ้นไป เพราะจะทำให้อำนาจการทดสอบของสถิติทดสอบเลอวีนเป็นที่เชื่อถือได้ และในงานการวิจัยเชิงทดลองทางสังคมศาสตร์มักจะวิจัยกับนักเรียนเป็นห้องเรียนซึ่งห้องเรียนห้องหนึ่งๆส่วนมากจะมีนักเรียนประมาณ 50 คน ดังนั้นการวิจัยครั้งนี้จึงกำหนดให้ขนาดของกลุ่มตัวอย่างมีค่าตั้งแต่ 10 ถึง 50 และการกำหนดให้ขนาดของกลุ่มตัวอย่างแต่ละกลุ่มไม่เท่ากัน ไม่สามารถที่จะศึกษาได้ทุกกรณี จึงศึกษาจากกรณีที่ขนาดของกลุ่มตัวอย่างแต่ละกลุ่มมีค่าแตกต่างกันกลุ่มละ 5 ดังนี้ 10,15,20,25,30,35,40,45 และ 50 ตามลำดับ เพื่อที่จะได้ผลสรุปที่ครอบคลุม

ขอบเขตของการวิจัย

1. การวิจัยครั้งนี้ศึกษาตัวแปรดังต่อไปนี้

1.1 ตัวแปรอิสระ (Independent Variables) คือ

ก. สถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน และสถิติทดสอบสแควร์แรงค์

ข. ความแปรปรวนของประชากร

ค. ขนาดของกลุ่มตัวอย่าง

1.2 ตัวแปรตาม (Dependent Variables) คือ

ก. ค่าความคลาดเคลื่อนประเภทที่ 1 (Actual Type I Error) ของสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน และสถิติทดสอบสแควร์แรงค์

ข. อำนาจของการทดสอบ ของสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน และสถิติทดสอบสแควร์แรงค์

2. ศึกษาเฉพาะกรณีที่การแจกแจงของประชากรเป็นแบบปกติ

3. ศึกษาเฉพาะกรณีจำนวนตัวอย่าง 3, 4 และ 5 กลุ่ม และศึกษาจากกรณีที่ขนาดของกลุ่มตัวอย่างไม่เท่ากัน ดังนี้

3.1 กลุ่มตัวอย่าง 3 กลุ่ม ขนาดของกลุ่มไม่เท่ากันมี 7 กรณีได้แก่
(10,15,20),(15,20,25), (20,25,30),(25,30,35),
(30,35,40),(35,40,45), (40,45,50),

3.2 กลุ่มตัวอย่าง 4 กลุ่ม ขนาดของกลุ่มไม่เท่ากันมี 6 กรณี ได้แก่
(10,15,20,25),(15,20,25,30),(20,25,30,35),
(25,30,35,40),(30,35,40,45),(35,40,45,50)

3.3 กลุ่มตัวอย่าง 5 กลุ่ม ขนาดของกลุ่มไม่เท่ากันมี 5 กรณีได้แก่
(10,15,20,25,30),(15,20,25,30,35),(20,25,30,35,40)
(25,30,35,40,45),(30,35,40,45,50)

4. ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบ ของสถิติทดสอบบาร์ตเลต เลอวีน และ สแควร์แรงค์ จะเปรียบเทียบเฉพาะเมื่อ α มีค่าเป็น .01 และ .05 เท่านั้น

5. กำหนดค่าพารามิเตอร์ μ = 500 และ σ^2 =100 สำหรับค่าความแปรปรวนที่น้อยที่สุด การกำหนดค่าความแปรปรวนของประชากรแต่ละกลุ่ม กำหนดให้ความแปรปรวนของประชากรทุกกลุ่มมีค่าเท่ากับ 100 สำหรับการพิจารณาค่าความคลาดเคลื่อนประเภทที่ 1 แต่ถ้าพิจารณาอำนาจการทดสอบ จะกำหนดค่าความแปรปรวนที่แตกต่างกันโดยการกำหนดความแตกต่างตามวิธีการของ Game และ Probert (1972) โดยใช้ค่านอนเซ็นทรัลลิตี้พารามิเตอร์(Noncentrality Parameter)หรือ ฟี (ϕ) เป็นตัววัดความแตกต่างของความแปรปรวนของประชากร ดังนี้

ตารางที่ 1 แสดงค่าฟี และความแปรปรวนของประชากรในกลุ่มต่างๆ

ϕ	ความแปรปรวนของประชากร(Population Variances)		
	กลุ่มตัวอย่าง 3 กลุ่ม	กลุ่มตัวอย่าง 4 กลุ่ม	กลุ่มตัวอย่าง 5 กลุ่ม
0	100,100,100	100,100,100,100	100,100,100,100,100
1	100,105,123	100,105,110,127	100,110,115,120,130
2	100,150,200	100,105,175,187	100,110,115,120,210
3	100,260,300	100,108,250,300	100,105,110,115,328
4	100,280,480	100,150,200,500	100,110,115,120,500
5	100,570,600	100,140,570,600	100,110,120,580,600
6	100,690,850	100,180,790,810	100,110,120,775,800
7	100,990,1000	100,110,970,990	100,105,978,980,990
8	100,1000,1430	100,105,1100,1330	100,110,1100,1200,1370
9	100,1330,1700	100,105,1250,1700	100,110,1330,1500,1700

6. ในแต่ละกรณีของการทดลองครั้งนี้ จะทำการทดลองซ้ำ 4000 ครั้ง

7. การวิจัยนี้จำลองการทดลองขึ้นในคอมพิวเตอร์ โดยอาศัยเทคนิคมอนติคาร์โล ซิมูเลชั่น (Monte Carlo Simulation Technique)

คำจำกัดความ

1. ความแกร่ง(Robustness)ของการทดสอบ หมายถึง คุณสมบัติของการทดสอบทางสถิติที่ไม่แสดงถึงความไว(Sensitive)ต่อการเบี่ยงเบน หรือการฝ่าฝืนข้อตกลงเบื้องต้นการทดสอบที่มีผลต่อโอกาสของการเกิดความคลาดเคลื่อนประเภทที่ 1 และโอกาสของการเกิดความคลาดเคลื่อนประเภทที่ 2

2. ความคลาดเคลื่อนประเภทที่ 1(Type I Error)หมายถึง ความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมติฐานสูญ(H_o) เมื่อสมมติฐานสูญ(H_o)เป็นจริง โอกาสที่จะทำให้เกิดความคลาดเคลื่อนประเภทที่ 1 กำหนดด้วย α

3. ค่าความคลาดเคลื่อนประเภทที่ 1 จากผลการทดลอง หมายถึงความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่เกิดขึ้นจริงในการทดสอบสมมติฐานของการทดสอบตามแผนการทดลองครั้งนี้ ซึ่งหาได้จากสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานสูญ (H_o) เมื่อสมมติฐานสูญนั้นเป็นจริง กับจำนวนครั้งของการทดลองทั้งหมด

4. อำนาจการทดสอบ (Power of Test) หมายถึง ความน่าจะเป็นที่จะปฏิเสธสมมติฐานสูญ(H_o) เมื่อสมมติฐานสูญ(H_o)นั้นผิด ซึ่งการวิจัยครั้งนี้จะถือว่าเป็นสัดส่วนของของจำนวนครั้งที่ปฏิเสธH_o เมื่อสมมติฐาน H_oนั้นผิด ต่อจำนวนครั้งของการทดลองทั้งหมด

6. ฟี (ϕ) หมายถึงดัชนีที่บอกถึงระดับของความแตกต่างกันของความแปรปรวนของประชากร ซึ่งคำนวณจากสูตร

$$\phi = \sqrt{\sum_{i=1}^{k} (\sigma_i^2 - \sigma^2)^2 / k\sigma_i^2}$$

เมื่อ $\sigma_i^2 = (\prod_{i=1}^{k} \sigma_i^2)^{1/k}$

=ค่าเฉลี่ยเรขาคณิตของความแปรปรวนของประชากร

<u>ประโยชน์ของการวิจัย</u>

ผลการวิจัยจะช่วยให้ผู้ใช้สถิติมีข้อสรุป และหลักฐานในการเลือกใช้สถิติทดสอบสำหรับการทดสอบการเท่ากันของความแปรปรวน เมื่อจะใช้ทดสอบจากกลุ่มตัวอย่าง 3,4 และ 5 กลุ่มที่ถูกสุ่มมาจากประชากรที่การแจกแจงแบบปกติ ได้อย่างมีประสิทธิภาพและเหมาะสมกับลักษณะของข้อมูล