หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ผลงานวิจัยด้านต่าง ๆ ส่วนมากอาศัยระเบียบวิธีการทางสถิติโดยเฉพาะกระบวนการวิเคราะห์ข้อมูลเพื่อหาข้อสรุปที่เชื่อว่ามีความถูกต้องในระดับที่เป็นที่ยอมรับเชิงทฤษฎี ซึ่งเรียกว่าสถิติอนุมาน (Statistical Inference) ผู้ใช้สถิติในงานวิจัยหรือการศึกษาใด ๆ มักจะมีคำถามอยู่เสมอว่า คำตอบหรือผลสรุปที่ได้ในขั้นสุดท้ายนั้นมีความเชื่อถือได้มากน้อยเพียงใด ซึ่งคำตอบในเรื่องนี้จำเป็นต้องขึ้นอยู่กับองค์ประกอบหลายอย่างได้แก่ การวางแผนการทดลอง เทคนิคการสุ่มตัวอย่าง เทคนิคการวิเคราะห์ข้อมูลโดยเฉพาะอย่างยิ่งการตัดสินใจของผู้วิจัยในการเลือกใช้ตัวสถิติทดสอบ เพื่อให้ได้ผลสรุปที่สมเหตุสมผลตามสมมติฐานของงานวิจัยนั้น

นักวิจัยเชิงทดลองส่วนมากจะใช้การทดสอบทางสถิติเพื่อเปรียบเทียบค่าเฉลี่ย (mean) และความแปรปรวน (variance) ของประชากร (population) โดยใช้ค่าสังเกต (observation) จากตัวอย่าง (sample) การเปรียบเทียบค่าเฉลี่ยสำหรับประชากร 2 กลุ่ม ส่วนมากใช้ตัวสถิติทดสอบที่ (Student t-test) และการทดสอบประชากรมากกว่า 2 กลุ่ม มักจะใช้การวิเคราะห์ความแปรปรวน (Analysis of variances) ส่วนการทดสอบการเท่ากันของความแปรปรวน (Test for homogeneity of variance) สำหรับประชากร 2 กลุ่ม ใช้ตัวสถิติทดสอบเอฟ (F-test) แต่ถ้ามากกว่า 2 กลุ่ม โดยปกติใช้ตัวสถิติทดสอบไคกำลังสอง (Chi-square) ที่เรียกว่าสถิติทดสอบของบาร์ตเลตต์ (Bartlett's statistic) ซึ่งเป็นตัวสถิติที่พัฒนามาจากตัวสถิติทดสอบของเนย์แมน เพียร์สัน (Neyman-Pearson's statistic)

* การทดสอบเกี่ยวกับการเท่ากันของความแปรปรวนของประชากรมักจะถูกนำมาใช้เมื่อ

1.1.1 ผู้วิจัยมีความสนใจที่จะอ้างอิงเกี่ยวกับความแปรปรวนของประชากรโดยตรงตามวัตถุประสงค์ของงานวิจัยนั้น

1.1.2 ผู้วิจัยต้องการตรวจสอบข้อมูลก่อนการวิเคราะห์ความแปรปรวน เมื่อเกิดความไม่แน่ใจว่าประชากรทุกกลุ่มจะมีความแปรปรวนเท่ากันตามข้อตกลงเบื้องต้น (assumption) ของวิธีวิเคราะห์ความแปรปรวนหรือไม่ ซึ่ง Box (1953 : 318-335) ได้เสนอแนะว่าก่อนที่จะใช้

วิธีการวิเคราะห์ความแปรปรวนเพื่อทดสอบการเท่ากันของค่าเฉลี่ย ควรมีการตรวจสอบเกี่ยวกับการเท่ากันของความแปรปรวนเสียก่อนเพื่อมิให้เกิดผลสรุปที่ผิดพลาด นอกจากนี้ Game (1972 : 887-909) ยังได้แสดงความคิดเห็นเช่นเดียวกันด้วยว่า ผลการวิเคราะห์ความแปรปรวนซึ่งใช้ตัวสถิติเอฟ จะยังมีความเชื่อถือได้ เมื่อมีการฝ่าฝืน (violation) ข้อตกลงเบื้องต้นเกี่ยวกับการเท่ากันของความแปรปรวนสำหรับกลุ่มตัวอย่างที่มีขนาดใหญ่คือมากกว่า 10 ขึ้นไป แต่อย่างไรก็ตาม ถ้าผู้วิจัยไม่สนใจและละเลยเกี่ยวกับการทดสอบการเท่ากันของความแปรปรวนแล้ว เมื่อผลการวิเคราะห์สรุปได้ว่าค่าเฉลี่ยของประชากรแตกต่างกันและผู้วิจัยต้องการจะใช้วิธีการเปรียบเทียบพหุคูณ (Multiple Comparison) โดยใช้วิธีการต่าง ๆ เช่น วิธีของทูกี (Tukey Wholly Significant Difference) วิธีของนิวแมน-คูลส์ (Newman-Keuls) วิธีของดันแคน (Duncan Multiple Range) วิธีของฟิชเชอร์ (Least Significant Difference) หรือ วิธีของเชฟเฟย์ (Scheffe Multiple Comparision Tests) ซึ่งวิธีการเปรียบเทียบทั้งหมดดังกล่าวนี้มีข้อตกลงเบื้องต้นเกี่ยวกับการเท่ากันของความแปรปรวนของประชากรที่นำมาทดสอบแต่ยังไม่มีหลักฐานยืนยันเกี่ยวกับความถูกต้อง เชื่อถือได้ของผลการเปรียบเทียบพหุคูณเหล่านี้เมื่อมีการฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับเรื่องนี้

วิธีการทดสอบการเท่ากันของความแปรปรวนของประชากรที่มากกว่าสองกลุ่มขึ้นไปมีตัวสถิติที่ใช้ทดสอบหลายตัว ดังได้กล่าวมาแล้วโดยมากมักนิยมใช้การทดสอบของบาร์ตเลต (Bartlett's test) ที่มีการแจกแจงแบบไคกำลังสอง และมีองศาความเป็นอิสระของการทดสอบเป็น k-1 เมื่อมีการทดสอบประชากร k กลุ่ม การใช้ตัวสถิตินี้มีข้อตกลงเบื้องต้นคือประชากรต้องมีการแจกแจงแบบปกติ ซึ่ง Brown และ Forsythe (1974 : 364-367) ได้กล่าวว่าวิธีการทดสอบของบาร์ตเลตจะมีผลทำให้ความคลาดเคลื่อนชนิดที่ 1 (Type I error) สูงกว่าระดับที่กำหนด (Nominal level of significance) ถ้าประชากรมีการแจกแจงไม่เป็นแบบปกติ

นอกจากนี้ยังมีตัวสถิติทดสอบการเท่ากันของความแปรปรวนอื่น ๆ อีก ซึ่งได้มีการพัฒนาให้มีประสิทธิภาพมากขึ้นเรื่อย ๆ ดังเช่น Levene (1960 : 278-292) ได้เสนอตัวสถิติทดสอบการเท่ากันของความแปรปรวนด้วยการเปรียบเทียบความแปรปรวนแบบเดียวกับการวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA) โดยการแทนค่าสังเกตแต่ละค่าด้วยค่าสัมบูรณ์ของส่วนเบี่ยงเบน (Absolute deviation) ระหว่างค่าสังเกตแต่ละค่ากับค่าเฉลี่ยของกลุ่มตัวอย่างนั้น

แต่ในปี 1968 Miller ได้พิสูจน์ให้เห็นว่าถ้ากลุ่มตัวอย่างมีขนาดเล็กมาก ๆ จะเกิดความสัมพันธ์กันอย่างสูง (High correlation) ระหว่างส่วนเบี่ยงเบนในกลุ่มเดียวกัน ซึ่งทำให้ตัวสถิติทดสอบของ Levene ขาดความเที่ยงตรง ถ้าการแจกแจงของประชากรมีลักษณะเบ้ (skew) เนื่องจากตัวสถิติของ Levene ไม่เป็น asymtotically distribution free นอกจากประชากรจะมีค่าเฉลี่ยเท่ากับมัธยฐาน ซึ่งจะทำให้มีผลต่อความคลาดเคลื่อนชนิดที่ 1 ของการทดสอบเช่นเดียวกับการทดสอบของบาร์ตเลต ดังนั้น Miller จึงได้เสนอให้หาส่วนเบี่ยงเบนโดยใช้มัธยฐานของตัวอย่างแทนค่าเฉลี่ย

Cadwell (1953 : 336-346) ได้ใช้ตัวสถิติทดสอบการเท่ากันของความแปรปรวนโดยพัฒนามาจากตัวสถิติของฮาร์ตเลย์ (Hartley's statistic หรือ F_{max} test) ซึ่งคำนวณค่าสถิติทดสอบโดยใช้อัตราส่วนระหว่างความแปรปรวนที่มากที่สุด (maximum variance) และความแปรปรวนที่น้อยที่สุด (minimum variance) สำหรับเปรียบเทียบตั้งแต่สองประชากรขึ้นไปที่มีการแจกแจงแบบปกติและใช้กลุ่มตัวอย่างขนาดเท่ากัน ตัวสถิติของแคดเวล (Cadwell's statistic) นั้น คำนวณโดยใช้อัตราส่วนระหว่างพิสัยที่มากที่สุด (maximum range) และพิสัยที่น้อยที่สุด (minimum range) ซึ่งเป็นตัวสถิติที่ใช้ทดสอบได้รวดเร็วเพราะไม่ยุ่งยากในการคำนวณ (Leslie และ Brown 1966 : 221-227)

ภายใต้ข้อตกลงเบื้องต้นของการทดสอบการเท่ากันของความแปรปรวนเกี่ยวกับการแจกแจงของประชากรที่ต้องเป็นแบบปกติซึ่งสภาพการณ์ทั่วไปในทางปฏิบัติข้อมูลที่เก็บรวบรวมได้อาจมิได้เป็นไปตามข้อตกลงข้างต้นผู้วิจัยอาจทำการวิเคราะห์ไปโดยฝ่าฝืนข้อตกลง ถึงแม้ว่าจะมีทฤษฎีการโน้มเข้าสู่ส่วนกลาง (Central Limit Theorem) ซึ่งกล่าวว่า ถ้าสุ่มตัวอย่างจากประชากรที่มีค่าเฉลี่ย μ และความแปรปรวน σ^2 โดยที่ขนาดของกลุ่มตัวอย่าง n ใหญ่เพียงพอ (sufficiently large samples) แล้ว การแจกแจงของค่าเฉลี่ยจากกลุ่มตัวอย่าง ($\bar{x}$) จะเข้าใกล้การแจกแจงแบบปกติ มีค่าเฉลี่ยเท่ากับค่าเฉลี่ยของประชากรคือ μ และมีความแปรปรวนของค่าเฉลี่ยคือ σ_x^2/n (Glass และ Stanley 1970 : 275) แต่บางครั้งในการวิจัยเชิงทดลองก็ไม่อาจใช้กลุ่มตัวอย่างขนาดใหญ่พอที่เหมาะสมกับงานวิจัยนั้น ๆ ได้

อย่างไรก็ดีถ้าเกิดการฝ่าฝืนข้อตกลงเบื้องต้นเมื่อต้องการทดสอบการเท่ากันของความแปรปรวนเกี่ยวกับลักษณะการแจกแจงต้องเป็นแบบปกติแล้ว อาจมีผลกระทบต่อผลสรุปของการทดสอบได้ ซึ่ง Game (1972 : 887-909) ได้กล่าวไว้ว่า ตำราส่วนมากมักจะไม่เน้นถึงข้อตกลงเบื้องต้น

เกี่ยวกับการแจกแจงแบบปกติ (Normality assumption) ในการทดลอบเพื่อเปรียบเทียบค่าเฉลี่ยและความแปรปรวนซึ่งมีความแตกต่างกันอย่างมากในการใช้ตัวสถิติเมื่ออ้างอิงถึงค่าความคลาดเคลื่อนมาตรฐาน 2 แบบคือ ถ้าไม่คำนึงถึงรูปแบบการแจกแจงของประชากร $\sigma_{\bar{x}}$ เท่ากับ $\sigma_x/\sqrt{n}$ เมื่อกลุ่มตัวอย่างมีขนาด n แต่ส่วนเบี่ยงเบนมาตรฐานของความแปรปรวนของตัวอย่าง (S^2) เท่ากับ $\sigma_{S^2} = \sigma_x^2 \sqrt{2/(n-1) + \gamma_2/n}$ เมื่อ γ_2 คือ ดัชนีความโด่งของประชากร (kurtosis index of population) โดยที่ $\gamma_2 = E\left[(x_1 - \mu_x)^4\right] / \sigma^4 - 3$ ซึ่ง γ_2 มีค่าตั้งแต่ -2 ถึง ∞ เมื่อประชากรมีการแจกแจงแบบปกติ γ_2 มีค่าเท่ากับ 0 แต่ถ้าการแจกแจงของประชากรมีความโด่งแบบเลปโตเคอร์คิค (leptokurtic) กล่าวคือ $0 < \gamma_2 < \infty$ จะทำให้ค่าที่แท้จริงของ $\sigma^2_{S^2}$ มากกว่ากรณีที่ประชากรมีการแจกแจงแบบปกติ ซึ่งจะทำให้เกิดความคลาดเคลื่อนชนิดที่ 1 ของการทดลอบสูงกว่าระดับที่กำหนด แต่ถ้าประชากรมีการแจกแจงที่มีความโด่งแบบพลาติเคอร์คิค (platykurtic) $-2 < \gamma_2 < 0$ ค่าที่แท้จริงของ $\sigma^2_{S^2}$ จะน้อยกว่ากรณีที่ประชากรมีการแจกแจงแบบปกติ

นอกจากนี้ Seheffe (1959 : 337-345) ยังได้แสดงความเห็นไว้ว่า การฝ่าฝืนข้อตกลงเบื้องต้นเกี่ยวกับการที่ประชากรต้องมีการแจกแจงแบบปกติจะมีผลกระทบน้อยมากสำหรับการทดลอบเพื่อเปรียบเทียบค่าเฉลี่ย แต่จะมีผลอย่างรุนแรงเมื่ออ้างอิงเกี่ยวกับความแปรปรวน เพราะทฤษฎีเกี่ยวกับการแจกแจงอาจกล่าวได้ว่า "เมื่อขนาดของตัวอย่างใหญ่พอ ผลสรุปอาจถูกต้องเกี่ยวกับค่าเฉลี่ย แต่ผลสรุปเกี่ยวกับการกระจายของข้อมูลจะผิดถ้าความโด่งไม่เป็นศูนย์"

จะเห็นได้ว่า การวิจัยจะได้ผลสรุปที่มีคุณภาพดีเพียงใดนั้นขึ้นอยู่กับเทคนิค หรือวิธีการทดลอบข้อมูลโดยการเลือกใช้ตัวสถิติทดลอบที่เหมาะสมที่สุดสำหรับสภาพการณ์ของข้อมูลนั้นคือควรเลือกใช้ตัวสถิติที่มีอำนาจการทดลอบ (Power of the test) สูงสุด ดังนั้นจึงควรมีการเปรียบเทียบอำนาจการทดลอบของตัวสถิติทดลอบเกี่ยวกับการเท่ากันของความแปรปรวนตามเหตุผลดังกล่าว

ในการศึกษาครั้งนี้ผู้วิจัยได้เลือกศึกษาตัวสถิติทดสอบ 3 ประเภท ดังได้กล่าวมาแล้ว เพราะสถิติทดสอบของบาร์ตเลต เป็นตัวสถิติทดสอบที่นิยมใช้ในการทดสอบการเท่ากันของความแปรปรวนของประชากร สถิติทดสอบของแคดเวล เป็นตัวสถิติทดสอบที่สามารถนำมาใช้ในการทดสอบได้อย่างรวดเร็ว เพราะคำนวณได้ง่าย และสถิติทดสอบของเลอวีน เมื่อแทนค่าเฉลี่ยเลขคณิตด้วยมัธยฐานของตัวอย่าง ซึ่งจะทำให้ตัวสถิตินี้มีลักษณะ asymtetically distribution free และเนื่องจากยังไม่มีการศึกษาเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบทั้งสามนี้โดยตรง การวิจัยครั้งนี้จะศึกษาโดยใช้เทคนิคของการซิมูเลชั่นที่เรียกว่า Monte Carlo Simulation Technique ซึ่งเป็นเทคนิคที่จะทำให้ได้ผลสรุปจากสภาพการณ์ที่เป็นการทดลอง คือ สามารถกำหนดขนาดของกลุ่มตัวอย่าง การแจกแจงของประชากรและอัตราส่วนความแปรปรวนได้โดยการใช้เครื่องคอมพิวเตอร์เข้ามาช่วย

1.2 วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบอำนาจการทดสอบการเท่ากันของความแปรปรวนของประชากรโดยใช้ตัวสถิติทดสอบ 3 ประเภทคือ ตัวสถิติทดสอบ χ^2 หรือตัวสถิติทดสอบของบาร์ตเลต ตัวสถิติทดสอบ C หรือตัวสถิติทดสอบของแคดเวล และตัวสถิติทดสอบ W' หรือ ตัวสถิติทดสอบที่ดัดแปลงมาจากวิธีของเลอวีน (modified Levene's test)

1.3 สมมติฐานของการวิจัย

อำนาจการทดสอบการเท่ากันของความแปรปรวนของตัวสถิติทดสอบ χ^2, C และ W' แตกต่างกัน

1.4 ขอบเขตของการวิจัย

1.4.1 ศึกษาเปรียบเทียบความแปรปรวนของประชากร 3 ประชากร โดยใช้กลุ่มตัวอย่างที่มีขนาดเท่ากัน ทั้งนี้เนื่องจากค่าวิกฤต (Critical value) ตัวสถิติ C จำเป็นต้องใช้ตารางของ Harter (1963 cited by Leslie and Brown 1966 : 226-227) ซึ่งกำหนดขนาดของกลุ่มตัวอย่างเท่ากันทุกกลุ่ม

1.4.2 กำหนดพารามิเตอร์ $\mu = 100$ และ ความแปรปรวนของประชากรกลุ่มที่ 1 $\sigma_1^2 = 100$ ในการสร้างข้อมูลที่มีการแจกแจงแบบปกติ โดยใช้โปรแกรมย่อย (Subroutine) ของ Gauss และข้อมูลที่มีการแจกแจงไม่เป็นแบบปกติเมื่อกำหนดความเบ้และความโด่งที่ระดับต่าง ๆ กัน โดยใช้ฟังก์ชั่นในการสร้างข้อมูลของ Ramberg และคณะ (1979 : 201-214)

1.4.3 กำหนดระดับนัยสำคัญ (level of significance) ของการทดสอบที่ α เท่ากับ 0.05 และ 0.01

1.4.4 การศึกษาครั้งนี้จำลองการทดลองขึ้นโดยใช้เครื่องคอมพิวเตอร์ IBM 370/3010 ด้วยเทคนิคมอนติคาร์โลซิมูเลชั่น ใช้โปรแกรมภาษาฟอร์แทรน (FORTRAN) โดยทำการทดลองซ้ำ 1,000 ครั้งในแต่ละสถานการณ์ของการทดลองเพื่อประหยัดเวลาการทำงานของเครื่องคอมพิวเตอร์

1.5 คำจำกัดความของเทอมต่าง ๆ ที่ใช้ในการวิจัย

1.5.1 อำนาจการทดสอบ (Power of the test) หมายถึงความน่าจะเป็นที่จะปฏิเสธ (reject) สมมติฐานว่าง (null hypothesis) เมื่อสมมติฐานว่างนั้นไม่จริง มีค่าเท่ากับ $1-\beta$ เมื่อ β คือความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนชนิดที่ 2 (Probability of Type II error)

1.5.2 ความคลาดเคลื่อนชนิดที่ 1 (Type I error) คือการปฏิเสธสมมติฐานว่างถ้าสมมติฐานว่างนั้นเป็นจริง

1.5.3 ความคลาดเคลื่อนชนิดที่ 2 (Type II error) คือการยอมรับ (accept) สมมติฐานว่างถ้าสมมติฐานว่างนั้นไม่จริง

1.6 ประโยชน์ที่คาดว่าจะได้รับ

เพื่อเป็นแนวทางให้ผู้ทำการวิจัยได้เลือกใช้ตัวสถิติทดสอบการเท่ากันของความแปรปรวนของประชากร ที่เหมาะสม ในกรณีที่ไม่ทราบแน่ชัดเกี่ยวกับลักษณะการแจกแจงของประชากรที่นำมาทดสอบสมมติฐานนั้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย