



บทที่ 1

บทนำ

1.1 ความสำคัญและความเป็นมาของปัญหา

ในการวิจัยเชิงทดลอง เกี่ยวกับการเปรียบเทียบตัวอย่างกลุ่ม 2 ชุด ซึ่งเป็นอิสระต่อกันนั้นอาจมีจุดมุ่งหมายที่แตกต่างกันไปตามวัตถุประสงค์ของงานวิจัยแต่ละเรื่อง เช่น งานวิจัยบางเรื่องต้องการทดสอบความเท่ากันของค่าเฉลี่ยระหว่างประชากร 2 ชุด ซึ่งก็คือการศึกษาความแตกต่างอย่างมีนัยสำคัญของประชากรอันเป็นที่มาของกลุ่มตัวอย่างทั้งสอง โดยจะเริ่มด้วยการทดสอบความเท่ากันของความแปรปรวนระหว่างประชากร 2 ชุดนี้ก่อน และผลจากการทดสอบสมมติฐานดังกล่าว ถ้าหากตกลงยอมรับว่าความแปรปรวนของทั้งสองประชากรเท่ากันแล้ว จึงทำการทดสอบความเท่ากันของค่าเฉลี่ยต่อไป นอกจากนี้ยังมีงานวิจัยอีกไม่น้อยที่สนใจจะอ้างอิงเกี่ยวกับความแปรปรวนของประชากรโดยตรง ซึ่งในกรณีเหล่านี้ เรามักจะเลือกใช้สถิติทดสอบพื้นฐาน (Classical test Statistics) เช่นการทดสอบที (t test) และการทดสอบเอฟ (F test) สำหรับหาคำตอบให้แก่ปัญหาเหล่านี้ ซึ่งโดยทฤษฎีแล้ว การทดสอบหรือสถิติทดสอบเหล่านี้ จะต้องอยู่ภายใต้ข้อตกลงเบื้องต้นจำนวนหนึ่ง ซึ่งจะสามารถยอมรับได้ว่า ผลการวิจัยนั้นเป็นที่เชื่อถือได้

สำหรับการทดสอบสมมติฐานเกี่ยวกับความเท่ากันระหว่างความแปรปรวนของประชากรสองชุดที่เป็นอิสระกันนั้น สถิติทดสอบ (test Statistics) ที่นิยมใช้กันมากที่สุด คือ การทดสอบเอฟ ซึ่งการใช้ การทดสอบวิธีนี้ จะต้องอยู่ภายใต้ข้อตกลงเบื้องต้นว่า ตัวอย่างกลุ่มทั้งสองชุดนั้นจะต้องเป็นอิสระกันและ ประชากรที่นำมาทดสอบนั้นจะต้องมีการแจกแจงเป็นแบบปกติ ซึ่งถ้าประชากรที่นำมาทดสอบนั้นมิได้มีการแจกแจงเป็นแบบปกติแล้ว การทดสอบเอฟนี้จะมีความไว (Sensitive) มากต่อผลการทดสอบ เช่นในการทดสอบเกี่ยวกับความแปรปรวน โดยที่ข้อมูลมาจากประชากรที่มีการแจกแจงแบบปกติ และในกรณีนี้ เราสนใจทดสอบเฉพาะความแปรปรวนเท่านั้น การแจกแจงของประชากร เป็นปัจจัยที่ไม่สนใจต้องการทดสอบ หากว่าผลของการทดสอบสรุปว่าต้องไม่ปฏิเสธ H_0 แต่เมื่อเปลี่ยนการแจกแจงของข้อมูลจากการแจกแจงแบบปกติไปเป็นการแจกแจงแบบอื่นแล้วผลของการทดสอบกลับเปลี่ยนแปลงไปเป็นปฏิเสธ

H_0 แทนที่จะยังคงไม่ปฏิเสธ H_0 อยู่อย่างเดิม ในกรณีเช่นนี้ จะเรียกว่าการทดสอบนี้มีความไวต่อการแจกแจงของประชากรอันเป็นที่มาของข้อมูล ซึ่งเป็นปัจจัยที่ไม่ต้องการทดสอบ เป็นต้น ดังนั้นถ้าหากข้อมูลที่นำมาศึกษาไม่เป็นไปตามข้อตกลง เบื้องต้นดังกล่าวทาง เลือกที่เป็นไปได้สำหรับการทดสอบสมมติฐานคือ

1. ใช้การทดสอบเอฟเดิม โดยที่ข้อมูลยังคงมีคุณสมบัติที่ไม่เป็นไปตามข้อตกลง เบื้องต้น
2. ใช้การทดสอบที่ไม่ใช่พารามิเตอร์ (Nonparametric)
3. ใช้การทดสอบวิธีอื่นที่เหมาะสมกับข้อมูลที่มีคุณสมบัติไม่เป็นไปตามข้อตกลง เบื้องต้น

สำหรับทางเลือกที่หนึ่งนั้น เป็นทางเลือกที่ไม่สู้จะดีนัก เนื่องจากผลสรุปของการวิจัยที่ได้จากการใช้การทดสอบที่ไม่เหมาะสมกับข้อมูลนั้นไม่เป็นที่ยอมรับกันโดยทั่วไป ดังเช่น Scheffe (1959:337-345) ได้กล่าวไว้ว่า "ผลของการกระจายของประชากรที่ไม่เป็นไปตามข้อตกลง เบื้องต้นนั้นจะมีผลเพียงเล็กน้อยต่อการอ้างอิงหรือการทดสอบเกี่ยวกับค่าเฉลี่ย แต่จะมีผลอย่างมากต่อการอ้างอิง หรือการทดสอบเกี่ยวกับค่าความแปรปรวน" นอกจากนี้ Cochran และ Cox (Kirk:1969:60) ยังได้กล่าวว่า "ข้อมูลที่ได้จากการวิจัยนั้นอาจจะไม่ตรงตามข้อตกลง เบื้องต้นที่กำหนดไว้ ดังนั้นค่าต่าง ๆ ที่นำมาวิเคราะห์นั้นจึงเป็นเพียงค่าประมาณมากกว่าจะเป็นค่าที่แท้จริง และการที่ข้อตกลงเบื้องต้นไม่เป็นจริงนี้ยังมีผลต่อระดับนัยสำคัญของการทดสอบและความไวของสถิติทดสอบด้วย" จากคำกล่าวเหล่านี้แสดงให้เห็นว่าผลสรุปที่ได้จากการทดสอบสมมติฐานในทาง เลือกที่หนึ่งจะบิดเบือนหรือผิดเพี้ยนไปจากความเป็นจริง

ส่วนทางเลือกที่สองนั้นเป็นทางเลือกที่ผู้วิจัยสามารถเลือกใช้ได้ เนื่องจากเป็นการทดสอบที่ไม่มีข้อจำกัดเกี่ยวกับลักษณะการแจกแจงของประชากร ซึ่งในระยะเวลา 25 ปี ที่ผ่านมามีการพัฒนาวิธีการทดสอบทางสถิติขึ้นใหม่ โดยเฉพาะเมื่อการทดลองอยู่ในลักษณะที่ใช้ตัวอย่างขนาดเล็ก และรูปแบบของประชากรไม่เป็นแบบปกติ (Non-normal Distribution) แต่จะทำให้อำนาจในการทดสอบลดลงไป

สำหรับทางเลือกที่สามจะเป็นทางเลือกที่น่าสนใจในการนำมาใช้แก้ไขปัญหาข้างต้น กล่าวคือ ในกรณีที่ข้อมูลที่มีอยู่นั้นมิได้มีการแจกแจงเป็นแบบปกติ หากแต่เป็นการแจกแจงแบบสมมาตรหางยาว (Long tail distribution, Heavy tail distribution) หรือ

เป็นการแจกแจงแบบเบ้ (Skew distribution) การหาการทดสอบที่มีคุณสมบัติที่ดีและเหมาะสมมาทำการทดสอบแทนการทดสอบเอฟนั้น จะทำให้การวิเคราะห์และการแปลผลมีความถูกต้องมากขึ้น

จากพื้นฐานความเชื่อที่ว่า การที่ข้อตกลงเบื้องต้นผิดพลาดไปนั้นจะมีผลกระทบเพียงเล็กน้อยต่อการแจกแจงจากตัวอย่างของสถิติทดสอบ (Sampling distribution of the test statistics) ซึ่งถ้าความเชื่อนี้เป็นจริง จะเรียกสถิติทดสอบนั้นว่ามีความแกร่ง (Robust) อันเป็นคำที่ Box (1953) ได้นำมาใช้ หรืออีกนัยหนึ่ง คือ สถิติทดสอบที่มีความแกร่งนั้นจะต้องไม่ทำให้ ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (type I error) หรือความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 2 (type II error) มากขึ้น ถ้าข้อตกลงเบื้องต้นเกิดผิดพลาดไป (Sawat Pratoomraj 1970:1) ดังนั้นคุณสมบัติที่จะชี้ให้เห็นว่าสถิติทดสอบนั้นมีความเหมาะสมหรือไม่ก็คือความแกร่ง (Robustness) และอำนาจของการทดสอบ (Power of the test) นั่นเอง

สำหรับการทดสอบที่ใช้ทดสอบความเท่ากันของความแปรปรวนที่มีความแกร่ง เมื่อประชากรมีการแจกแจงที่ไม่ใช่แบบปกตินั้นมีหลายตัวด้วยกัน เช่น การทดสอบแจคไนฟ์ (Jackknife test) การทดสอบโคลล์แควร์ที่เล่นอโดยแลยาร์ด (Layard Chisquare test) การทดสอบเลเวนเนที่ใช้ค่าเฉลี่ยเลขคณิตเป็นค่าวัดแนวโน้มสู่ส่วนกลาง (Levene test: W_0) และการทดสอบที่ปรับปรุงจาก การทดสอบเลเวนเน โดยใช้ค่ามัธยฐานและค่าเฉลี่ยที่เกิดจากการตัดค่าสูง เกตที่ปลายทั้งสองข้างของชุดตัวอย่างออกเป็นขนาด 10% และ 20% ตามลำดับ (Median, 10% trimmed mean, 20% trimmed mean) เป็นค่าวัดแนวโน้มสู่ส่วนกลางแทนค่าเฉลี่ยเลขคณิต (Modified Levene test : W_{50} , W_{10} , W_{20}) เป็นต้น โดยที่การทดสอบกลุ่มหลังสุดนี้อาจจะเป็นสถิติทดสอบที่เหมาะสมกับข้อมูลที่มาจากประชากรที่มีการแจกแจงแบบสมมาตรหางยาวและแบบเบ้มากกว่า

เมื่อข้อตกลงเบื้องต้นเกี่ยวกับลักษณะการแจกแจงของประชากรนั้นไม่เป็นจริงก็ไม่ควรที่จะเลือกใช้การทดสอบเอฟ ทั้งนี้เพราะว่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 จะมีค่าสูง ซึ่งหมายความว่า ผู้วิจัยจะเสี่ยงต่อการปฏิเสธสมมติฐาน H_0 เมื่อสมมติฐาน H_0 ถูกในกรณีนี้ผู้วิจัยอาจจะพิจารณาการทดสอบอื่นที่แกร่งกว่าการทดสอบเอฟ เช่น การทดสอบแบบไม่ใช้พารามิเตอร์ (Nonparametric test) แต่ในกรณีที่ไม่ต้องการใช้การ

ทดสอบดังกล่าวนี้ก็จะเลือกใช้การทดสอบทางด้านพาราเมตริก (Parametric Test) ได้ ซึ่งในกรณีที่เลือกใช้การทดสอบทางด้านพาราเมตริกนั้นจะมีสถิติทดสอบหลายตัวที่น่าสนใจ ดังนั้น สิ่งเป็นสิ่งที่น่าศึกษาเพื่อจะได้มีข้อสรุปว่า การทดสอบที่มีความแกร่งตัวใดเหมาะสมกับการแจกแจงแบบสัณฐานมาตรหางยาวและการแจกแจงแบบเบ้ รวมทั้งเป็นแนวทางในการเลือกใช้สถิติทดสอบที่เหมาะสมสำหรับการทดสอบความเท่ากันของความแปรปรวนระหว่างประชากรสองชุด และด้วยเหตุว่าการคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจของการทดสอบเมื่อลักษณะของการแจกแจงไม่เป็นไปตามข้อตกลงเบื้องต้นโดยตรงนั้น กระทำได้ยาก ดังนั้นผู้วิจัยจึงเลือกศึกษาโดยวิธีมอนติคาร์โล (Monte Carlo method) ซึ่งเป็นวิธีที่ศึกษาในรูปของการจำลอง (Simulation) โดยอาศัยตัวเลขสุ่มเทียม (Pseudo random number) และสามารถระบุขนาดของกลุ่มตัวอย่าง ความแปรปรวน ค่าเฉลี่ย และลักษณะการแจกแจงของประชากรได้ตามที่ผู้วิจัยต้องการ

1.2 วัตถุประสงค์ของการวิจัย

ในการวิจัยครั้งนี้ต้องการศึกษา เปรียบเทียบการทดสอบที่มีความแกร่งที่ใช้ทดสอบความเท่ากันของความแปรปรวนระหว่างประชากรสองชุด 7 วิธีคือ

1. การทดสอบเอฟ (F)
 2. การทดสอบแจคไนฟ์ (J)
 3. การทดสอบโคสแควร์ที่เสนอโดยเลয়ারด์ (CS)
 4. การทดสอบเลเวนเน (W₀)
 5. การทดสอบที่ปรับปรุงมาจากการทดสอบเลเวนเนโดยใช้ค่ามัธยฐานเป็นค่าวัดแนวโน้มสู่ส่วนกลางแทนค่าเฉลี่ย (W₅₀)
 6. การทดสอบที่ปรับปรุงมาจากการทดสอบเลเวนเนโดยใช้ค่าเฉลี่ยที่เกิดจากการตัดค่าสังเกตุที่ปลายทั้งสองข้างของชุดตัวอย่างออกเป็นขนาด 10% (W₁₀)
 7. การทดสอบที่ปรับปรุงมาจากการทดสอบเลเวนเนโดยใช้ค่าเฉลี่ยที่เกิดจากการตัดค่าสังเกตุที่ปลายทั้งสองข้างของชุดตัวอย่างออกเป็นขนาด 20% (W₂₀)
- เพื่อหาข้อสรุปเกี่ยวกับความเหมาะสมของการทดสอบดังกล่าว ภายใต้อัตลักษณ์การแจกแจงของประชากรแบบปกติ (Normal Distribution) แบบที (t Distribution) ซึ่งเป็นการ

แจกแจงที่มีลักษณะลัมมาตรงยาว และการแจกแจงแบบไคส์แควร์ (Chisquare Distribution) แบบไวบูลล์ (Weibull Distribution) ซึ่งเป็นการแจกแจงแบบเบ้ขวา นอกจากนี้ยังต้องการศึกษาว่า เมื่อประชากรทั้งสองชุดมีการแจกแจงแตกต่างกันแต่มีลักษณะคล้ายกันแล้วจะยังคงมีสถิติทดสอบที่เหมาะสมที่จะนำไปทดสอบความเท่ากันของความแปรปรวนระหว่างประชากรสองชุดนั้นหรือไม่โดยศึกษากรณีที่มีการแจกแจงของประชากร เป็นแบบลัมมาตรงคือ แบบปกติกับที่ และแบบเบ้คือ แบบไคส์แควร์กับแบบไวบูลล์

1.3 ลัมมาฐานของการวิจัย

ในการทดสอบความเท่ากันของความแปรปรวนระหว่างประชากรสองชุด ที่ไม่ได้มีการแจกแจงแบบปกตินั้น การทดสอบที่ปรับปรุงจากการทดสอบเลเวนเนทัง 3 วิธี (Modified Levene test) จะมีความเหมาะสมมากที่สุด

1.4 ข้อกำหนดของการวิจัย

การวิจัยครั้งนี้ถือว่า ความแกร่งและอำนาจของการทดสอบเป็นดัชนีที่ผู้วิจัยจะใช้เป็นเกณฑ์ในการเลือกการทดสอบที่เหมาะสมในการนำไปใช้กับแต่ละการแจกแจงของประชากรที่กำหนดในวัตถุประสงค์ของการวิจัย

1.5 ขอบเขตของการวิจัย

1. ศึกษาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจของการทดสอบเอฟ การทดสอบแควนไนท์ การทดสอบไคส์แควร์ที่เสนอโดยเลয়ারด์ การทดสอบเลเวนเน และการทดสอบที่ปรับปรุงจากการทดสอบเลเวนเนทัง 3 วิธี เมื่อประชากรทั้งสองชุดมีการแจกแจงแบบปกติ แบบที่ แบบไคส์แควร์ และแบบไวบูลล์ รวมทั้งเมื่อประชากรทั้งสองชุดมีการแจกแจงคล้ายกันคือ ประชากรหนึ่งเป็นแบบปกติกับอีกประชากรหนึ่งเป็นแบบที่ และแบบที่เป็นไคส์แควร์กับแบบไวบูลล์

2. กำหนดค่าเฉลี่ยของประชากรทั้งสองชุดเท่ากัน

3. อัตราส่วนของความแปรปรวนของประชากรทั้งสอง ($\sigma_1^2 ; \sigma_2^2$) มีค่าเป็น 1:4

1:2 2:1 และ 4:1

4. ชุดตัวอย่างที่ใช้ในการวิจัยครั้งนี้แบ่งเป็น 2 กรณี คือ ขนาดตัวอย่างทั้งสองชุดเท่ากัน และไม่เท่ากัน กรณีที่ขนาดของชุดตัวอย่างเท่ากันจะศึกษากรณีขนาดของตัวอย่าง เป็น 10, 40 และ 100 ส่วนในกรณีที่ขนาดของชุดตัวไม่เท่ากันจะศึกษากรณีขนาดของตัวอย่าง เป็น 10:20, 30:50, 80:100

5. การวิจัยครั้งนี้ทำการทดลองด้วย เครื่องคอมพิวเตอร์ IBM 370/3010 ด้วยเทคนิคมอนติคาร์โล โดยใช้โปรแกรมที่เขียนด้วยภาษาฟอร์แทรน (FORTRAN)

6. กำหนดระดับนัยสำคัญของการทดสอบที่ α เท่ากับ 0.01 และ 0.05

7. การทดสอบจะกระทำซ้ำกัน 1,000 ครั้ง ในแต่ละสถานการณ์ของการทดลองที่กำหนดไว้ในบทที่ 3

1.6 คำจำกัดความ

ความแกร่ง (Robustness) ของการทดสอบ หมายถึง คุณสมบัติของการทดสอบที่ไม่แสดงความไวต่อการเปลี่ยนแปลงของปัจจัยอื่นที่ไม่ใช่ปัจจัยที่ต้องการทดสอบ เช่น การเบี่ยงเบนไปจากข้อตกลงเบื้องต้นของการทดสอบอันจะมีผลต่อความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 2

อำนาจของการทดสอบ (Power of the test) หมายถึงความน่าจะเป็น (Probability) ที่จะปฏิเสธสมมติฐานว่าง (Null Hypothesis) เมื่อสมมติฐานว่างนั้นผิด ซึ่งจะมีค่าเท่ากับ $1-\beta$ เมื่อ β คือความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 2 (Kirk 1969 : 555, Minium 1978:364)

ความคลาดเคลื่อนประเภทที่ 1 (Type I error) เป็นความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างนั้นถูก

ความคลาดเคลื่อนประเภทที่ 2 (Type II error) เป็นความผิดพลาดที่เกิดจากการไม่ปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างนั้นผิด

1.7 ประโยชน์ของการวิจัย

1. ได้สถิติทดสอบที่มีความแกร่งและเหมาะสมกับลักษณะของข้อมูล ให้ผู้ทำวิจัยในสาขาต่าง ๆ ใ้ตีนำไปใช้ทดสอบความเท่ากันของความแปรปรวนระหว่างประชากรสองชุด สำหรับกรณีที่ลักษณะของการแจกแจงของประชากรที่นำมาทดสอบนั้น มีลักษณะการแจกแจงแบบเบ้ขวาและลิมิตหางยาว

2. เป็นแนวทางในการวิจัยหาสถิติทดสอบใหม่ต่อไป