หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการวิจัยและการอภิปรายผล

จากการทดลองศึกษาเกี่ยวกับสถิติทดสอบที่มีความแกร่งสำหรับทดสอบความเท่ากันของความแปรปรวนระหว่างประชากรสองชุดทั้ง 7 ตัวดังที่ได้กล่าวไว้ในบทที่ 2 ภายใต้การแจกแจงแบบต่าง ๆ ขนาดตัวอย่างและสัดส่วนของความแปรปรวนระดับต่าง ๆ ที่กำหนดไว้ในบทที่ 3 เพื่อหาผลสรุปว่าสถิติทดสอบใดมีความเหมาะสมที่จะนำไปใช้ทำการทดสอบข้อมูล เมื่อข้อมูลดัง-กล่าวมาจากประชากรที่มีลักษณะสมมาตร สมมาตรชนิดหางยาว และลักษณะเบ้ นั้น สามารถสรุปผลได้ดังนี้

5.1 ผลสรุปของความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1

จากการทดลองหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทด-สอบเอฟ การทดสอบแจคไนฟ์ การทดสอบไคสแควร์ที่เสนอโดยเลยาร์ด การทดสอบเลเวนเน และการทดสอบที่ปรับปรุงจากการทดสอบเลเวนเนทั้ง 3 วิธี เมื่อนำมาเปรียบเทียบกับค่า α ที่กำหนด โดยใช้เกณฑ์พิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของคอคแครน (Cochran) และเกณฑ์ของแบรดเลย์ (Bradley) ผลสรุปคือ

1. การทดสอบเอฟ สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก เมื่อประชากรมีการแจกแจงเป็นแบบปกติทั้งกรณีที่ α = 0.01 และ 0.05 และทุกระดับของขนาดตัวอย่าง แต่เมื่อการแจกแจงมีรูปแบบอื่น ๆ การทดสอบเอฟนี้ไม่สามารถควบคุมความคลาด-เคลื่อนประเภทที่ 1 ได้เลยไม่ว่า α และขนาดตัวอย่างจะอยู่ในระดับใดก็ตาม

2. การทดสอบแจคไนฟ์ สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อประชากรมีการแจกแจงแบบปกติที่ทุกระดับขนาดตัวอย่างยกเว้นเมื่อ α = 0.01 และขนาดตัวอย่างเท่ากับ 40:40 และจะควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้บ้างเมื่อการแจกแจงแจงเป็นแบบไวบูลล์ แต่เมื่อการแจกแจงเป็นแบบไคสแควร์และแบบทีนั้นมันจะควบคุมความคลาด-เคลื่อนประเภทที่ 1 ไม่ได้เลยเมื่อ α = 0.01 ในทุกสถานการณ์ที่ได้ทำการศึกษาในการวิจัยนี้

แต่เมื่อ α = 0.05 จะมีบางกรณีที่การทดสอบแจคไนฟ์สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เมื่อการแจกแจงเป็นแบบไคสแควร์และแบบที คือกรณีตัวอย่างมีขนาดเป็น (30:50)

3. การทดสอบไคสแควร์ที่เสนอโดยเลยาร์ด สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ทุกระดับของขนาดตัวอย่างเมื่อ α = 0.05 และการแจกแจงเป็นแบบปกติ สำหรับการแจกแจงแบบอื่น ๆ ในทุกสถานการณ์ที่ทำการศึกษานั้นมันจะควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เฉพาะกรณีที่ตัวอย่างมีขนาดกลางและขนาดใหญ่เท่านั้น

4. การทดสอบเลเวนเน สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ทุกสถานการณ์ที่ทำการศึกษาเมื่อการแจกแจงเป็นแบบปกติและแบบที ส่วนกรณีการแจกแจงแบบไคสแควร์การทดสอบนี้จะไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เลยแม้แต่สถานการณ์เดียว สำหรับการแจกแจงแบบไวบูลล์ เมื่อ α = 0.01 และ 0.05 และตัวอย่างมีขนาดใหญ่การทดสอบเลเวนเนจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้

5. การทดสอบที่ปรับปรุงจากการทดสอบเลเวนเน W_{50} สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ในเกือบทุกสถานการณ์ที่ศึกษา ยกเว้น กรณีที่ α = 0.01 และขนาดตัวอย่างเท่ากับ 10:10 เท่านั้นที่ควบคุมความคลาดเคลื่อนประเภทที่ 1 ไม่ได้

6. การทดสอบที่ปรับปรุงจากการทดสอบเลเวนเน W_{10} และ W_{20} สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ทุกระดับนัยสำคัญและขนาดตัวอย่างเมื่อการแจกแจงเป็นแบบปกติและแบบที แต่เมื่อการแจกแจงเป็นแบบไคสแควร์และแบบไวบูลล์การทดสอบทั้ง 2 นี้จะควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เฉพาะกรณีตัวอย่างขนาดกลางและใหญ่ เมื่อ α = 0.01 และ 0.05

7. สำหรับการทดสอบทั้ง 7 วิธีนี้ จะไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ แม้แต่สถานการณ์เดียวเมื่อประชากรที่นำมาศึกษานั้นมิได้มีการแจกแจงเป็นแบบเดียวกัน

8. การเพิ่มขนาดของตัวอย่างและการเปลี่ยนแปลงรูปแบบของการแจกแจงของประชากรมีผลต่อการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบทั้ง 7 วิธี

5.2 ผลสรุปของอำนาจของการทดสอบ

จากการทดลองเพื่อหาค่าอำนาจของการทดสอบทั้ง 7 วิธี เมื่อประชากรมีการแจกแจงแบบปกติ แบบไคสแควร์ แบบไวบูลล์ และแบบที สรุปผลได้ดังนี้

1. เมื่อการแจกแจงเป็นแบบปกติ การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าการทดสอบอื่น ๆ ในทุก ๆ สถานการณ์ที่ทำการศึกษา โดยที่อำนาจของการทดสอบนี้จะเพิ่มขึ้นเมื่อสัดส่วนของความแปรปรวนสูงขึ้น หรือขนาดตัวอย่างเพิ่มขึ้น หรือระดับนัยสำคัญเพิ่มขึ้น

2. เมื่อการแจกแจงเป็นแบบไคสแควร์ ซึ่งเป็นการแจกแจงแบบเบ้ขวานั้น อำนาจของการทดสอบที่ปรับปรุงจากการทดสอบเลเวนเน W_{50} มีค่าสูงที่สุดเมื่อตัวอย่างมีขนาดเล็กและ $\alpha = 0.01$ และ 0.05 ในกรณีตัวอย่างขนาดกลาง ณ ระดับนัยสำคัญทั้ง 2 ระดับ การทดสอบที่ปรับปรุงจากการทดสอบเลเวนเน W_{10} จะมีอำนาจของการทดสอบสูงที่สุด ส่วนกรณีตัวอย่างขนาดใหญ่ ณ ระดับนัยสำคัญทั้ง 2 ระดับนั้น การทดสอบที่ปรับปรุงจากการทดสอบเลเวนเน W_{20} มีอำนาจของการทดสอบสูงที่สุด ซึ่งการทดสอบที่มีอำนาจของการทดสอบสูงที่สุดเหล่านี้จะมีอำนาจของการทดสอบเพิ่มสูงขึ้นถ้าสัดส่วนของความแปรปรวนเพิ่มมากขึ้น

3. เมื่อการแจกแจงเป็นแบบไวบูลล์ ซึ่งมีลักษณะเบ้ขวาอีกการแจกแจงหนึ่งนั้น อำนาจของการทดสอบที่ปรับปรุงจากการทดสอบเลเวนเน W_{50} ยังคงมีค่าสูงที่สุดเมื่อตัวอย่างมีขนาดเล็ก ณ $\alpha = 0.01$ และ $\alpha = 0.05$ เมื่อตัวอย่างมีขนาดไม่เท่ากันส่วนกรณีขนาดตัวอย่างเท่ากันเมื่อ $\alpha = 0.05$ นี้การทดสอบแลคไนพ์มีอำนาจของการทดสอบสูงที่สุด , ในกรณีตัวอย่างขนาดกลาง เมื่อ $\alpha = 0.01$ และตัวอย่างมีขนาดเท่ากัน การทดสอบแลคไนพ์และการทดสอบที่ปรับปรุงจากการทดสอบเลเวนเน W_{20} มีอำนาจของการทดสอบสูงกว่าการทดสอบอื่น ๆ แต่เมื่อตัวอย่างมีขนาดไม่เท่ากัน การทดสอบที่ปรับปรุงจากการทดสอบเลเวนเน W_{10} จะมีอำนาจของการทดสอบสูงที่สุด และเมื่อ $\alpha = 0.05$ อำนาจของการทดสอบที่ปรับ-ปรุงจากการทดสอบเลเวนเน W_{10} มีค่ามากที่สุดเมื่อขนาดตัวอย่างเท่ากัน แต่อำนาจของการทดสอบเลเวนเน (W_0) จะมีค่าสูงที่สุดเมื่อขนาดตัวอย่างไม่เท่ากัน , สำหรับกรณีตัวอย่างขนาดใหญ่ ณ α ทั้ง 2 ระดับและตัวอย่างมีขนาดเท่ากันหรือไม่เท่ากัน นั้นการทดสอบเลเวนเนและการทดสอบที่ปรับปรุงจากการทดสอบเลเวนเนทั้ง 3 วิธี W_{10}, W_{20} W_{50} นั้นมีอำนาจของการทดสอบสูงพอ ๆ กัน ซึ่งอำนาจของการทดสอบดังกล่าวจะเพิ่มมากขึ้นเมื่อสัดส่วนของความแปรปรวนเพิ่มขึ้น

4. เมื่อการแจกแจงเป็นแบบที ซึ่งมีลักษณะสมมาตรหางยาวนั้น ในกรณีที่ตัวอย่างมีขนาดเล็ก เมื่อขนาดตัวอย่างเท่ากัน ณ α ทั้ง 2 ระดับ นั้นอำนาจของการทดสอบเลเวนเนมีค่าสูงที่สุด แต่เมื่อขนาดตัวอย่างไม่เท่ากัน ณ $\alpha = 0.01$ อำนาจของการทดสอบที่ปรับปรุงจากการทดสอบเลเวนเน W_{10} มีค่าสูงที่สุด และ ณ $\alpha = 0.05$ อำนาจของการทดสอบไคสแควร์และการทดสอบเลเวนเนจะมีค่าสูงที่สุด, สำหรับกรณีตัวอย่างขนาดกลาง ณ $\alpha = 0.01$ และ 0.05 เมื่อขนาดตัวอย่างเท่ากัน การทดสอบเลเวนเนมีอำนาจของการทดสอบสูงที่สุด และเมื่อขนาดตัวอย่างไม่เท่ากัน อำนาจของการทดสอบไคสแควร์ และการทดสอบเลเวนเนมีค่าสูงที่สุด, ส่วนในกรณีตัวอย่างขนาดใหญ่ ณ α ทั้ง 2 ระดับและขนาดตัวอย่างทั้งที่เท่ากันและไม่เท่ากันนั้น การทดสอบเลเวนเนและการทดสอบที่ปรับปรุงจากการทดสอบเลเวนเนทั้ง 3 วิธี มีอำนาจของการทดสอบสูงพอ ๆ กัน โดยที่เมื่อสัดส่วนของความแปรปรวนเพิ่มขึ้นอำนาจของการทดสอบต่าง ๆ เหล่านี้จะมีค่าเพิ่มมากขึ้น

5. การเพิ่มสัดส่วนของความแปรปรวน การเพิ่มระดับนัยสำคัญ การเพิ่มขนาดตัวอย่าง นั้นมีผลทำให้อำนาจของการทดสอบแต่ละวิธีเพิ่มสูงขึ้น

5.3 การอภิปรายผล

จากการวิจัยครั้งนี้เมื่อพิจารณาถึงความแกร่งหรือความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบในกรณีที่ประชากรมีการแจกแจงแบบต่าง ๆ ในลักษณะที่สมมาตร สมมาตรหางยาว และเบ้ขวา นั้นปรากฏว่าลักษณะการกระจายของข้อมูลมีผลต่อความแกร่งของการทดสอบ เช่น เมื่อลักษณะการกระจายของข้อมูลไม่เป็นแบบปกติแล้ว การทดสอบ F จะไม่แกร่ง ซึ่งสอดคล้องกับการทดลองของมิลเลอร์(1968) อย่างไรก็ตามมิลเลอร์ได้ศึกษาการทดสอบต่าง ๆ ภายใต้การแจกแจงที่แตกต่างไปจากการวิจัยครั้งนี้ คือศึกษาการทดสอบ บ๊อกซ์ และแอนเดอร์สัน แจคไนฟ์ เลเวนเน บ๊อกซ์ และโมเสส ภายใต้การแจกแจงแบบปกติ ยูนิฟอร์ม ดับเบิ้ลเอ็กซ์โพเนนเชียล สคิวเนสเอ็กซ์โพเนนเชียล และซิคท์เพาเวอร์ (Sixth Powes) แต่ในการวิจัยครั้งนี้ศึกษาการทดสอบ F, J, CS, W_o, W_{10}, W_{50} และ W_{20} ภายใต้การแจกแจงแบบปกติ ไคสแควร์ ไวบูลล์และที ซึ่งจากผลการวิจัย มีเพียงกรณีการทดสอบ F เท่านั้นที่สามารถเปรียบเทียบผลกับการวิจัยของมิลเลอร์ได้ ส่วนรายละเอียดของผลสรุปในกรณีอื่น ๆ ได้นำเสนอไว้แล้วในการสรุปผลการวิจัยข้างต้น จากผลสรุปของการวิจัยจะเห็นว่า

นอกจากรูปแบบการแจกแจงของประชากรและขนาดตัวอย่างจะมีผลต่อความแกร่งของการทดสอบแล้ว อำนาจของการทดสอบก็มีปัจจัยที่จะทำให้มีการเปลี่ยนแปลงได้เหมือนกันเช่น ขนาดของตัวอย่าง สัดส่วนของความแปรปรวน $(\sigma_1^2 : \sigma_2^2)$ และระดับนัยสำคัญที่กำหนดเป็นต้น

5.4 ข้อเสนอแนะ

ในการเลือกใช้สถิติทดสอบความเท่ากันของความแปรปรวนระหว่างประชากรสองชุดนั้น การทดสอบเอฟ เป็นการทดสอบที่ควรเลือกใช้ เมื่อประชากรมีการแจกแจงเป็นแบบปกติ สำหรับกรณีที่ประชากรมิได้มีการแจกแจงแบบปกติ หากแต่มีการแจกแจงอื่นที่มีรูปแบบมาตรฐานแน่นอน เช่น ลักษณะข้อมูลเป็นแบบสมมาตรหางยาวมีการแจกแจงแบบที หรือลักษณะข้อมูลเป็นแบบเบ้ขวามีการแจกแจงแบบไคสแควร์ ไวบูลล์แล้วหากจะยังคงใช้การทดสอบเอฟ ในการทดสอบข้อมูลอยู่นั้น ผลจากการวิเคราะห์จะไม่เป็นที่น่าเชื่อถือ เนื่องจากในกรณีนี้ถือได้ว่าไม่เป็นไปตามข้อตกลงเบื้องต้น ซึ่งจะเป็นสาเหตุทำให้ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีค่าสูงขึ้น อันมีความหมายว่าผู้วิจัยจะต้องเสี่ยงต่อการปฏิเสธ H_o เมื่อ H_o ถูก ดังนั้นโดยอาศัยผลจากการวิจัยครั้งนี้จะสรุปในรูปของตาราง เพื่อเป็นแนวทางสำหรับการเลือกสถิติทดสอบในการนำไปใช้งาน ดังตารางที่ 5.1

ตารางที่ 5.1 แสดงตัวสถิติที่มีอำนาจของการทดสอบสูงที่สุด ณ สถานการณ์ต่าง ๆ ซึ่งจำแนกตามการแจกแจงของประชากร ขนาดตัวอย่าง สัดส่วนของความแปรปรวนและระดับนัยสำคัญ

ขนาดตัวอย่างของการวิจัยครั้งนี้	การแจกแจง			ปกติ		ไคสแควร์		ไวบูลล์		ที	
	ระดับนัยสำคัญ			0.01	0.05	0.01	0.05	0.01	0.05	0.01	0.05
เล็ก (10-20)	$n_1=n_2$	$\sigma_1^2:\sigma_2^2$	2:1	F	F	W_{50}	W_{50}	W_{50}	J	W_0	W_0
			4:1	F	F	W_{50}	W_{50}	W_{50}	J	W_0	W_0
	$n_1<n_2$		2:1	F	F	W_{50}	W_{50}	W_{50}	W_{50}	W_{10}	CS
			4:1	F	F	W_{50}	W_{50}	W_{50}	W_{50}	W_{10}	CS
			1:2	F	F	W_{50}	W_{50}	W_{50}	W_{50}	W_{10} W_{20}	W_0
			1:4	F	F	W_{50}	W_{50}	W_{50}	W_{50}	W_{10} W_{20}	W_0
กลาง (30-50)	$n_1=n_2$	$\sigma_1^2:\sigma_2^2$	2:1	F	F	W_{10}	W_{10}	J	W_{10}	W_0	W_0
			4:1	F	F	W_{10}	W_{10}	W_{20}	W_{10}	W_0	W_0
	$n_1<n_2$		2:1	F	F	W_{10}	W_{10}	W_{50}	W_0	CS	CS
			4:1	F	F	W_{10}	W_{10}	W_{50}	W_0	CS	CS
			1:2	F	F	CS, W_{10}	W_{10}	W_{50}	W_0	W_0	W_0
			1:4	F	F	W_{10}	W_{10} W_{20}	W_0	W_0	W_0	W_0
ใหญ่ (80-100)	$n_1=n_2$	$\sigma_1^2:\sigma_2^2$	2:1	F	F	W_{20}	W_{10}	W_{10}	W_0	W_0	W_0
			4:1	F	F	W_{50} W_{20}	W_{10} W_{20} W_{50}	W_{10} W_{20} W_{50}	CS W_0 W_{10} W_{20} W_{50}	W_0 W_{10} W_{20}	W_0 W_{10} W_{20} W_{50}
	$n_1<n_2$		2:1	F	F	W_{20}	W_{20}	W_0	W_0	W_0	W_0
			4:1	F	F	W_{50} W_{20}	CS W_{50} W_{20}	W_0 W_{10} W_{20} W_{50}	W_0 W_{10} W_{20} W_{50}	W_{10} W_{20} W_{50}	W_0 W_{10} W_{20} W_{50}
			1:2	F	F	W_{20}	W_{20}	W_0	W_0	W_0	W_0
			1:4	F	F	W_{50} W_{20}	CS W_{50} W_{20}	W_0 W_{10} W_{20} W_{50}	W_0 W_{10} W_{20} W_{50}	W_0 W_{10} W_{20} W_{50}	W_0 W_{10} W_{20} W_{50}

จากตารางที่ 5.1 นักวิเคราะห์สามารถเลือกการทดสอบไปใช้งานได้ตามสถานการณ์ต่าง ๆ ที่ขึ้นอยู่กับการแจกแจงของประชากรขนาดตัวอย่าง สัดส่วนของความแปรปรวนและระดับนัยสำคัญที่ต้องการทดสอบได้

อนึ่ง สำหรับการนำสถิติทดสอบดังกล่าวไปใช้ในทางปฏิบัติจริงนั้น ขั้นตอนง่าย ๆ ในการจัดการกับข้อมูลที่มีอยู่ในมือและการเลือกสถิติทดสอบที่จะใช้ในการวิเคราะห์เพื่อให้มีความเหมาะสมมากที่สุดพอจะอธิบายได้เป็นขั้นตอนดังนี้

1. ทำการตรวจสอบข้อมูลที่มีอยู่ ซึ่งอาจจะใช้โปรแกรมสำเร็จรูป $SPSS^X$ ช่วยในการ plot ข้อมูลเพื่อดูการกระจายของข้อมูลชุดนั้น ๆ

2. เลือกสถิติทดสอบที่เหมาะสมกับข้อมูล เมื่อต้องการทดสอบเกี่ยวกับความเท่ากันของความแปรปรวนโดยที่

2.1 ถ้าข้อมูลมีการกระจายในลักษณะสมมาตร ควรเลือกใช้การทดสอบเอฟ ทุกระดับของขนาดตัวอย่าง

2.2 ถ้าข้อมูลมีการกระจายในลักษณะสมมาตรแต่มีค่าของข้อมูลบางค่ากระจายออกห่างจากกลุ่มมาก (ลักษณะหางยาว) สำหรับการทดสอบเมื่อตัวอย่างมีขนาดเล็กควรเลือกใช้การทดสอบ W_o หรือ W_{10} แต่เมื่อขนาดของตัวอย่างเพิ่มมากขึ้น จะมีการทดสอบที่สามารถเลือกไปใช้งานได้หลายตัวเช่น W_o, W_{10}, W_{20}, W_{50} และ CS เป็นต้น

2.3 ถ้าข้อมูลมีการกระจายในลักษณะเบ้ขวา เมื่อต้องการทดสอบข้อมูลที่มีขนาดเล็ก สถิติทดสอบที่เหมาะสมในการนำมาใช้งานคือ W_{50} ส่วนในการทดสอบเมื่อขนาดตัวอย่างเพิ่มมากขึ้น การทดสอบที่จะเลือกมาใช้งานได้คือ W_{10}, W_{20} และ W_{50}

เนื่องจากในการวิจัยครั้งนี้มิได้ศึกษาในกรณีที่ข้อมูลกระจายในลักษณะเบ้ซ้าย ดังนั้นผู้วิจัยจึงขอเสนอแนะว่า ถ้าข้อมูลมีการกระจายในลักษณะดังกล่าวแล้ว การทดสอบ W_{50} จะเป็นการทดสอบที่คาดว่าจะเหมาะสมที่สุดที่จะนำมาใช้งานได้เพาะเหตุว่า W_{50} นั้นได้ใช้ค่ามัธยฐานเป็นค่าวัดแนวโน้มสู่ส่วนกลางซึ่งถือว่าเป็นตัวแทนของข้อมูลที่ดีสำหรับการกระจายของข้อมูลแบบเบ้ ซึ่งผลของการทดสอบน่าจะได้ผลดีเหมือนกรณีที่ข้อมูลมีการกระจายแบบเบ้ขวา

3. ทำการคำนวณค่าของสถิติทดสอบที่เลือกได้ในข้อ 2

4. เปรียบเทียบค่าที่คำนวณได้ในข้อ 3 กับค่าวิกฤตที่เปิดได้จากตารางและทำการตัดสินใจตามเกณฑ์ที่ได้เสนอไว้แล้วในบทที่ 2