บทที่ 5

สรุปผลการวิจัยและการอภิปรายผล

จากการทดลองศึกษาเพื่อค้นหาสถิติทดสอบที่มีความแกร่งสำหรับทดสอบความเท่ากันของค่าเฉลี่ยประชากร โดยใช้สถิติทดสอบ 3 วิธีภายใต้ลักษณะการแจกแจงของประชากรแบบต่าง ๆ ขนาดตัวอย่าง อัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน และอัตราส่วนของค่าเฉลี่ยตามที่กำหนดไว้ในขอบเขตของการวิจัย ได้ผลสรุปดังนี้

5.1 ผลสรุปของความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1

จากการทดลองหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของสถิติทดสอบแบบ ANOVA F-TEST สถิติทดสอบแบบ Trimmed W และสถิติทดสอบแบบ Trimmed F แล้วนำมาเปรียบเทียบกับค่า α ที่กำหนด โดยใช้เกณฑ์ในการพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของคอคแครน (Cochran) ได้ผลสรุปดังนี้

1. สถิติทดสอบแบบ ANOVA F-TEST สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีมาก เมื่อประชากรมีการแจกแจงเป็นแบบปกติ ที่มีอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานของประชากรเท่ากัน ทั้งกรณีที่ $\alpha = 0.05$ และ 0.01 และทุกระดับของขนาดตัวอย่างที่ทำการศึกษา แต่เมื่อประชากรมีลักษณะการแจกแจงเป็นแบบปกติปลอมปน สถิติทดสอบแบบ ANOVA F-TEST สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้บางในกรณีที่ประชากรมีตัวอย่างขนาดใหญ่ [(50, 50, 50) และ (30, 40, 50)] โดยสามารถควบคุมได้ทั้งกรณีที่ $\alpha = 0.05$ และ 0.01

สำหรับกรณีที่ประชากรมีค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานแตกต่างกันแบบ 1:2:3 สถิติทดสอบแบบ ANOVA F-TEST ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เลย ในทุกรูปแบบของการแจกแจงและทุกขนาดตัวอย่างที่ทำการศึกษาที่ $\alpha = 0.01$ แต่เมื่อ $\alpha = 0.05$ สถิติทดสอบแบบ ANOVA F-TEST จะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ในกรณีที่ขนาดตัวอย่างใหญ่ (50, 50, 50) โดยสามารถควบคุมได้ทั้งกรณีที่การแจกแจง

เป็นแบบปกติและแบบปกติปลอมปน แต่เมื่อขนาดตัวอย่างเป็น (30, 40, 50) สถิติทดสอบแบบ ANOVA F-TEST สามารถควบคุมได้เฉพาะกรณีที่ประชากรแจกแจงแบบปกติเท่านั้น ส่วนที่ขนาดตัวอย่างเล็ก (10, 10, 10) และ (5, 10, 15) สถิติทดสอบแบบ ANOVA F-TEST ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เลยในทุกรูปแบบของการแจกแจงและทุกระดับ α ที่ศึกษา

สำหรับกรณีที่ประชากรมีค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานแตกต่างกันแบบ 3:2:1 นั้น สถิติทดสอบแบบ ANOVA F-TEST ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เลยทั้งกรณีที่ขนาดตัวอย่างเป็น (5, 10, 15) และ (30, 40, 50) ในทุกระดับของค่า α

2. สถิติทดสอบแบบ Trimmed W สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อประชากรมีการแจกแจงแบบปกติ ทั้งกรณีที่มีอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานเท่าและไม่เท่ากัน ในทุกขนาดตัวอย่างที่ทำการศึกษา ส่วนกรณีที่ประชากรมีลักษณะการแจกแจงเป็นแบบปกติปลอมปนทั้ง 4 รูปแบบนั้น สถิติทดสอบแบบ Trimmed W ก็สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ทั้งนี้จะต้องพิจารณาค่าร้อยละของการตัดค่าสังเกตที่ปลายหางทั้ง 2 ด้านของการแจกแจง (g) เปอร์เซ็นต์การปลอมปนและค่าสเกลแฟคเตอร์ของการแจกแจงแบบปกติปลอมปนด้วย นอกจากนี้ระดับนัยสำคัญก็มีส่วนในการพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของสถิติทดสอบดังกล่าวอีกด้วย ซึ่งจากผลการทดลองพบว่าภายใต้สถานการณ์เดียวกัน (ลักษณะการแจกแจง ขนาดตัวอย่าง อัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน) พบว่าที่ระดับนัยสำคัญ 0.05 สถิติทดสอบแบบ Trimmed W จะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ โดยใช้เปอร์เซ็นต์การตัดข้อมูลที่ปลายหางของการแจกแจง (g) น้อยกว่าที่ระดับ $\alpha = 0.01$ เป็นส่วนใหญ่

3. สถิติทดสอบแบบ Trimmed F สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อประชากรมีการแจกแจงแบบปกติ ทั้งกรณีที่ประชากรมีค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานเท่าและไม่เท่ากัน ที่ระดับนัยสำคัญ 0.05 และ 0.01 ส่วนกรณีที่ประชากรมีการแจกแจงแบบปกติปลอมปนทั้ง 4 รูปแบบนั้น สถิติทดสอบแบบ Trimmed F ก็สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ทั้งนี้ต้องพิจารณาค่าร้อยละของการตัดค่าสังเกตที่ปลายหางทั้ง 2 ด้านของการแจกแจง (ค่า g) เปอร์เซ็นต์การปลอมปน และค่าสเกลแฟคเตอร์ของการ

แจกแจงแบบปกติปลอมปนด้วย นอกจากนี้ระดับนัยสำคัญก็มีส่วนในการพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของสถิติทดสอบดังกล่าวอีกด้วย ซึ่งจากผลการทดลองพบว่า ภายใต้สถานการณ์เดียวกัน (ลักษณะการแจกแจง ขนาดตัวอย่าง อัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน) พบว่า ที่ระดับนัยสำคัญ 0.05 สถิติทดสอบแบบ Trimmed F จะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ โดยใช้ค่า g น้อยกว่าที่ระดับ 0.01 เป็นส่วนใหญ่

5.2 ผลสรุปของอำนาจของการทดสอบ

จากผลการทดลองเพื่อหาข้อสรุปเกี่ยวกับอำนาจของการทดสอบของสถิติทดสอบทั้ง 3 วิธี ภายใต้ลักษณะการแจกแจงแบบต่าง ๆ ได้ผลสรุปดังนี้คือ

1. ภายใต้การแจกแจงแบบปกติ สถิติทดสอบแบบ ANOVA F-TEST ให้ค่าอำนาจของการทดสอบสูงสุดเมื่อประชากรมีค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานของประชากรเท่ากัน ในทุกขนาดตัวอย่างที่ทำการศึกษา แต่ในกรณีที่ประชากรมีค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานต่างกันแบบ 1:2:3 สถิติทดสอบแบบ Trimmed F จะให้ค่าอำนาจของการทดสอบสูงกว่าสถิติทดสอบแบบ Trimmed W เมื่อ $\mu_1 : \mu_2 : \mu_3$ มีค่า 1:1:2 และให้ผลกลับกันเมื่อ $\mu_1 : \mu_2 : \mu_3$ มีค่า 2:1:1 หรือกล่าวอีกนัยหนึ่งก็คือ ประชากรใดมีค่าส่วนเบี่ยงเบนมาตรฐานสูง และค่าเฉลี่ยสูงด้วยพบว่าสถิติทดสอบแบบ Trimmed F จะให้ค่าอำนาจของการทดสอบสูงกว่าสถิติทดสอบแบบ Trimmed W ทั้งกรณีที่ $\alpha = 0.05$ และ 0.01 ในทุกขนาดตัวอย่างที่ทำการศึกษา

2. ภายใต้การแจกแจงแบบปกติปลอมปนทั้ง 4 รูปแบบ สถิติทดสอบแบบ Trimmed W แบบ Trimmed F ต่างก็ให้ค่าอำนาจของการทดสอบดีแตกต่างกันไปตามสถานการณ์ เช่น กรณีที่ประชากรมีค่าส่วนเบี่ยงเบนมาตรฐานสูง และค่าเฉลี่ยสูงนั้นพบว่าสถิติทดสอบแบบ Trimmed F จะให้ค่าอำนาจของการทดสอบสูงกว่า สถิติทดสอบแบบ Trimmed W และกรณีที่ประชากรมีค่าส่วนเบี่ยงเบนมาตรฐานสูงแต่ค่าเฉลี่ยต่ำ สถิติทดสอบแบบ Trimmed W จะให้ค่าอำนาจของการการทดสอบสูงกว่าสถิติทดสอบแบบ Trimmed F ทั้งนี้จะต้องพิจารณาลักษณะการแจกแจงด้วยว่ามีลักษณะเป็นหางยาวมากหรือน้อย ซึ่งจากผลการทดลองจะเห็นว่า กรณีที่ประชากรมีลักษณะการแจกแจงที่มีหางยาวมากขึ้น ค่าร้อยละของการตัดค่าสังเกตที่ปลายหางทั้ง

2 ด้านของการแจกแจง (ค่า g) จะเพิ่มมากขึ้น เช่น ที่การแจกแจงแบบ C(10, 5) และ C(10, 10) สถิติทดสอบทั้ง 2 แบบ (Trimmed W และ Trimmed F) จะใช้ค่า g ประมาณ 10-15% แต่ที่การแจกแจงแบบ C(20, 5) และ C(20, 10) สถิติทดสอบทั้ง 2 ประเภท จะใช้ค่า g ประมาณ 15-20% เป็นต้น

3. การเพิ่มขนาดตัวอย่าง หรือเพิ่มความแตกต่างของส่วนเบี่ยงเบนมาตรฐาน หรือค่าเฉลี่ยของแต่ละประชากรมากขึ้น จะมีผลทำให้ค่าอำนาจการทดสอบของสถิติทดสอบมีค่ามากขึ้น นอกจากนี้ ความยาวของหางของการแจกแจงจะมีผลต่อค่าร้อยละของการตัดค่าสังเกตที่ปลายหางทั้ง 2 ด้านของการแจกแจงอีกด้วย

5.3 การอภิปรายผล

จากผลการวิจัย เมื่อพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของสถิติทดสอบทั้ง 3 วิธี ภายใต้ลักษณะการแจกแจงแบบปกติและการแจกแจงแบบปกติปลอมปนนั้น พบว่า ลักษณะการแจกแจงของประชากร อัตราส่วนของส่วนเบี่ยงเบนมาตรฐานและขนาดตัวอย่าง จะมีผลต่อความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของสถิติทดสอบ เช่น ในกรณีที่ข้อมูลไม่ได้มีการแจกแจงแบบปกติที่มีค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานเท่ากันแล้ว สถิติทดสอบแบบ ANOVA F-TEST จะไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เลยในทุกขนาดตัวอย่างที่ทำการศึกษา ซึ่งสอดคล้องกับการทดลองของ Morton B. Brown and Alan B. Forsythe (1974) สำหรับการพิจารณาค่าอำนาจของการทดสอบของสถิติทดสอบภายใต้ลักษณะการแจกแจงแบบปกติและการแจกแจงแบบปกติปลอมปนนั้น พบว่า ลักษณะการแจกแจงของประชากร ขนาดตัวอย่าง อัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน และอัตราส่วนของค่าเฉลี่ยของประชากร จะมีผลต่อค่าอำนาจของการทดสอบของสถิติทดสอบดังรายละเอียดที่ได้เสนอไว้ในการสรุปผลของการวิจัยข้างต้น

5.4 ข้อเสนอแนะ

จากการวิจัยครั้งนี้ ได้ผลสรุปว่า ในการเลือกใช้สถิติทดสอบที่ทดสอบความเท่ากันของค่าเฉลี่ยประชากรในกรณีที่มีจำนวนประชากรมากกว่า 2 ชุดนั้น สถิติทดสอบแบบ ANOVA F-TEST หรือการวิเคราะห์ความแปรปรวน เป็นการทดสอบที่ควรนำมาใช้ เมื่อประชากรมีการแจกแจงแบบปกติที่มีค่าส่วนเบี่ยงเบนมาตรฐาน (หรือค่าความแปรปรวน) เท่ากัน ทุกประชากร

แต่ถ้าลักษณะการแจกแจงของประชากรไม่เป็นไปตามข้อกำหนดข้างต้น หากแต่มีการแจกแจงที่เป็นแบบล่มมาตรหางยาว ซึ่งในที่นี้ศึกษากรณีที่เป็นการแจกแจงแบบปกติปลอมปนแล้ว ควรเลือกใช้สถิติทดสอบแบบ Trimmed W หรือ Trimmed F ทั้งนี้ขึ้นอยู่กับค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน และค่าอัตราส่วนของค่าเฉลี่ยของประชากรด้วย เช่น ภายใต้ลักษณะการแจกแจงแบบเดียวกัน ประชากรที่มีค่าส่วนเบี่ยงเบนมาตรฐานสูง และมีค่าเฉลี่ยสูง (เช่น $\sigma_1 : \sigma_2 : \sigma_3 = 1:2:3$ คู่กับ $\mu_1 : \mu_2 : \mu_3 = 1:1:2$ หรือ $\sigma_1 : \sigma_2 : \sigma_3 = 3:2:1$ คู่กับ $\mu_1 : \mu_2 : \mu_3 = 2:1:1$) ควรเลือกใช้สถิติทดสอบแบบ Trimmed F ซึ่งจะให้ค่าอำนาจของการทดสอบสูง ส่วนประชากรใดที่มีค่าส่วนเบี่ยงเบนมาตรฐานต่ำแต่มีค่าเฉลี่ยสูง (เช่น $\sigma_1 : \sigma_2 : \sigma_3 = 1:2:3$ คู่กับ $\mu_1 : \mu_2 : \mu_3 = 2:1:1$ หรือ $\sigma_1 : \sigma_2 : \sigma_3 = 3:2:1$ คู่กับ $\mu_1 : \mu_2 : \mu_3 = 1:1:2$) ควรเลือกใช้สถิติทดสอบแบบ Trimmed W ทั้งนี้จะต้องพิจารณาค่า g ที่เหมาะสมอีกด้วย นอกจากนี้ในการคำนวณค่าระดับความเป็นอิสระ (degree of freedom:df) ของสถิติทดสอบทั้ง 2 วิธี (Trimmed W และ Trimmed F) ซึ่งในบางกรณีอาจได้ค่า df ที่ไม่ตรงกับค่า df ในตาราง F Hyunshik Lee และ Karen Yuen Fung (1982) ได้แนะนำให้ใช้วิธีเทียบบัญญัติไตรยางค์เพื่อหาค่าที่เหมาะสมได้ นอกจากนี้จากผลการทดลองเกี่ยวกับค่าอำนาจการทดสอบของสถิติทดสอบทั้ง 3 วิธี (กรณี $\alpha = 0.05$ ศึกษาได้จากตารางที่ 4.1-4.5 ส่วนกรณีที่ $\alpha = 0.01$ ศึกษาได้จากตารางที่ 4.7-4.11) เมื่อกำหนดให้ปัจจัยหนึ่งเปลี่ยนไป แต่ปัจจัยอื่น ๆ คงที่ พบว่าปัจจัยที่มีผลกระทบต่อค่าอำนาจการทดสอบมากที่สุดคือ อัตราส่วนของส่วนเบี่ยงเบนมาตรฐานของประชากร ส่วนปัจจัยที่มีผลกระทบรองลงมาคือ ลักษณะการแจกแจงของประชากร ขนาดตัวอย่าง เปอร์เซ็นต์การตัดข้อมูลที่ปลายหางของการแจกแจงและระดับนัยสำคัญตามลำดับ อย่างไรก็ตามข้อเสนอแนะดังกล่าวกำหนดภายใต้ขอบเขตของการวิจัยที่ได้เสนอไว้ในบทที่ 1 สำหรับการวิจัยในครั้งนี้เท่านั้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5.1 แสดงผลการทดสอบที่มีอำนาจการทดสอบสูงที่สุด ณ ค่าแกนการณ์ต่าง ๆ ซึ่งจำแนกตามการแจกแจงของประชากร ขนาดตัวอย่าง อัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน อัตราส่วนของค่าเฉลี่ย และระดับนัยสำคัญ

ขนาดตัวอย่าง	การแจกแจง		ปกติ		C(10,5)		C(10,10)		C(20,5)		C(20, 10)	
	$\sigma_1:\sigma_2:\sigma_3$	$\mu_1:\mu_2:\mu_3$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
10, 10, 10	1:1:1	1:1:2	AF	AF	TW(10)	TW(10)	TW(10)	TW(10)	TW(10)	TW(20)	TW(20)	TW(20)
		1:2:3	AF	AF	TW(10)	TW(10)	TF(20)	TW(10)	TF(20)	TW(20)	TW(20)	TW(20)
	1:2:3	1:1:2	TF(0)	TF (0)	TF(10)	TF(10)	TF(10)	TF(15)	TF(20)	TF(20)	TF(20)	TF(20)
		2:1:1	TW(0)	TW (0)	TW(10)	TW(10)	TW(10)	TW(10)	TW(10)	TW(20)	TW(20)	TW(20)
50, 50, 50	1:1:1	1:1:2	AF TW(0),TF(0)	AF TW(0),TF(0)	TW(10)	TW(10)	TW(10)	TW(10)	TW(10)	TW(20)	TW(15)	TW(20)
		1:2:3	AF TW(0),TF(0)	AF TW(0),TF(0)	TW(0)	TW(10)	TW(10)	TW(10)	TW(10)	TW(20)	TW(15)	TW(20)
	1:2:3	1:1:2	TF(0)	TF(5)	TF(0)	TF(0)	TW(10)	TF(5)	TW(10)	TF(0)	TF(15)	TF(15)
		2:1:1	TW(0)	TW(0)	TW(0)	TW(10)	TW(10)	TW(5)	TW(10)	TW(5)	TW(15)	TW(20)
5, 10, 15	1:1:1	1:1:2	AF	AF	TF(10)	TF(15)	TF(20)	TF(20)	TF(20)	TF(20)	TW(20)	TF(20)
		1:2:3	AF	AF	TF(10)	TF(15)	TF(20)	TF(20)	TW(15)	TF(20)	TW(20)	TF(20)
	1:2:3	1:1:2	TF(0)	TF(0)	TF(0)	TF(5)	TF(15)	TF(10)	TF(15)	TF(15)	TF(20)	TF(20)
		2:1:1	TW(0)	TW(0)	TW(10)	TW(20)	TW(20)	TW(20)	TW(20)	TW(20)	TW(20)	TF(20)
	3:2:1	1:1:2	TF(0)	TW(0)	TW(10)	TW(5)	TW(15)	TW(15)	TW(15)	TW(15)	TF(20)	TF(20)
		2:1:1	TF(0)	TF(0)	TF(15)	TF(15)	TF(20)	TF(15)	TF(20)	TF(15)	TF(20)	TF(20)
30, 40, 50	1:1:1	1:1:2	AF TW(0),TF(0)	AF TW(0),TF(0)	TW(0)	TW(10)	TW(10)	TW(10)	TW(10)	TW(20)	TW(20)	TW(20)
		2:1:1	AF TW(0),TF(0)	AF TW(0),TF(0)	TW(0)	TW(0)	TW(10)	TW(10)	TW(10)	TW(20)	TW(20)	TW(20)
	1:2:3	1:1:2	TF(0)	TF(0)	TF(0)	TF(0)	TF(10)	TF(10)	TF(5)	TF(5)	TF(15)	TF(10)
		2:1:1	TW(0)	TW(0)	TW(0)	TW(0)	TW(10)	TW(10)	TW(10)	TW(10)	TW(20)	TF(15)
	3:2:1	1:1:2	TW(0)	TW(0)	TW(10)	TW(10)	TW(15)	TW(10)	TW(15)	TW(10)	TW(20)	TW(20)
		2:1:1	TF(0)	TF(0)	TF(0)	TF(5)	TF(10)	TF(10)	TF(5)	TF(5)	TF(15)	TF(15)

จากตารางที่ 5.1 ซึ่งเป็นตารางสรุปผลอำนาจการทดสอบที่มีค่าสูงที่สุดของสถิติทดสอบทั้ง 3 วิธี ภายใต้สถานการณ์ต่าง ๆ โดยจำแนกตามลักษณะการแจกแจงของประชากร ขนาดตัวอย่าง อัตราส่วนของส่วนเบี่ยงเบนมาตรฐาน อัตราส่วนของค่าเฉลี่ย และระดับนัยสำคัญ ซึ่งจากตารางดังกล่าว นักวิจัยสามารถเลือกการทดสอบไปใช้งานได้ตามสถานการณ์ที่มีอยู่

อนึ่ง สำหรับการนำสถิติทดสอบดังกล่าวไปใช้งานในทางปฏิบัติจริงนั้น ขั้นตอนง่าย ๆ ในการจัดการกับข้อมูลที่มีอยู่ และวิธีการเลือกสถิติทดสอบที่จะใช้ในการวิเคราะห์เพื่อให้มีความเหมาะสมมากที่สุด มีขั้นตอนดังนี้

1. ทำการตรวจสอบข้อมูลที่มีอยู่ ซึ่งอาจจะใช้วิธีการทางสถิติในการทดสอบสมมุติฐานเกี่ยวกับลักษณะการแจกแจงของประชากรว่ามีการแจกแจงเป็นไปตามที่สงสัยหรือไม่ หรืออาจใช้โปรแกรมสำเร็จรูป SPSS^{X} ช่วยในการ plot ข้อมูลเพื่อดูลักษณะการกระจายของข้อมูลชุดนั้น ๆ

2. เลือกสถิติทดสอบที่เหมาะสมกับลักษณะการกระจายของข้อมูล เมื่อต้องการทดสอบสมมติฐานเกี่ยวกับความเท่ากันของค่าเฉลี่ยของประชากร โดยถ้าข้อมูลมีลักษณะการกระจายเป็นแบบปกติที่มีค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานเท่ากันทุกประชากรแล้ว ในที่นี้ $\sigma_1 : \sigma_2 : \sigma_3 = 1:1:1$ ควรเลือกใช้สถิติทดสอบแบบ ANOVA F-TEST ซึ่งจะให้อำนาจของการทดสอบสูงในทุกขนาดตัวอย่างและทุกระดับ α ที่ทำการศึกษา แต่ถ้าประชากรมีค่าอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานแตกต่างกันควรเลือกใช้สถิติทดสอบแบบ Trimmed W หรือ Trimmed F ทั้งนี้ขึ้นอยู่กับอัตราส่วนของส่วนเบี่ยงเบนมาตรฐานและอัตราส่วนของค่าเฉลี่ยด้วย เช่น ประชากรใดมีค่าส่วนเบี่ยงเบนมาตรฐานสูง และค่าเฉลี่ยสูง (เช่น $\sigma_1 : \sigma_2 : \sigma_3 = 1:2:3$ คู่กับ $\mu_1 : \mu_2 : \mu_3 = 1:1:2$ หรือ $\sigma_1 : \sigma_2 : \sigma_3 = 3:2:1$ คู่กับ $\mu_1 : \mu_2 : \mu_3 = 2:1:1$) ควรเลือกใช้สถิติทดสอบแบบ Trimmed F ซึ่งจะให้ค่าอำนาจของการทดสอบสูงสุด ส่วนประชากรใดที่มีค่าส่วนเบี่ยงเบนมาตรฐานต่ำแต่ค่าเฉลี่ยสูง (เช่น $\sigma_1 : \sigma_2 : \sigma_3 = 1:2:3$ คู่กับ $\mu_1 : \mu_2 : \mu_3 = 2:1:1$ หรือ $\sigma_1 : \sigma_2 : \sigma_3 = 3:2:1$ คู่กับ $\mu_1 : \mu_2 : \mu_3 = 1:1:2$) ควรใช้สถิติทดสอบแบบ Trimmed W ซึ่งจะให้ผลเหมือนกันเมื่อข้อมูลมีการแจกแจงเป็นแบบปกติปลอมปนด้วย ทั้งนี้จะต้องพิจารณาลักษณะความยาวที่ปลายหางของการแจกแจงด้วย

ว่ายาวมากหรือน้อย เช่น ถ้าการแจกแจงมีลักษณะเป็นหางยาวมากก็จะใช้ค่าร้อยละของการตัดค่าสังเกตที่ปลายหาง (g) มากขึ้น ซึ่งรายละเอียดได้เล่นอไว้แล้วในบทที่ 4

3. เมื่อเลือกสถิติทดสอบที่เหมาะสมกับลักษณะการแจกแจงของข้อมูลได้แล้วก็ทำการคำนวณค่าสถิติทดสอบที่เลือกได้ในข้อ 2

4. เปรียบเทียบค่าที่คำนวณได้ในข้อ 3 กับค่าวิกฤตที่เปิดได้จากตารางและทำการตัดสินใจตามเกณฑ์ที่ได้เล่นอไว้แล้วในบทที่ 2

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย