

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

1. สรุปผลการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงจำลอง (Simulation Research) มีวัตถุประสงค์เพื่อศึกษาความแข็งแกร่งของสัมประสิทธิ์สหสัมพันธ์ เมื่อพบข้อมูลมีค่าผิดปกติจากกลุ่ม โดยใช้ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์จำนวน 4 แบบ คือ สหสัมพันธ์แบบเพียร์สัน สหสัมพันธ์แบบสเปียร์แมน สหสัมพันธ์แบบเคนดอลล์ และ สหสัมพันธ์แบบถ่วงน้ำหนัก ศึกษาความแข็งแกร่งของสัมประสิทธิ์สหสัมพันธ์ในสถานการณ์ต่างๆ ที่แตกต่างกันดังนี้ ขนาดตัวอย่างมี 3 ระดับ คือ 20, 50 และ 100 กำหนดระดับความสัมพันธ์ที่ใช้ในการทดสอบ 5 ระดับคือ 0, 0.20, 0.50, 0.80 และ 1.0 กำหนดจำนวนของค่าผิดปกติจากกลุ่มซึ่งเป็นค่าผิดปกติจากกลุ่มระดับปานกลาง ด้านมีค่ามากให้มีจำนวนเท่ากับ 0%, 5%, 10%, 20% และ 30% ของขนาดตัวอย่าง ตามลำดับ โดยแยกศึกษา การเกิดค่าผิดปกติจากกลุ่มที่ละตัวแปร ระดับนัยสำคัญในการทดสอบคือ 0.05 ทำการจำลองข้อมูลด้วยเทคนิค มอนติคาร์โล บนเครื่องคอมพิวเตอร์ส่วนบุคคลโดยใช้โปรแกรมภาษาซีชาร์ป (C#) ชื่อโปรแกรมคือ Robust Correlation Test ในแต่ละสถานการณ์ทำการทดลองซ้ำ 1,000 ครั้ง เกณฑ์ที่ใช้ในการเปรียบเทียบความแข็งแกร่งของสัมประสิทธิ์สหสัมพันธ์มี 2 แบบ คือ ความแข็งแกร่งในการทดสอบทางสถิติ และ ความแข็งแกร่งในการประมาณค่า ผลการวิจัยมีข้อสรุปดังนี้

ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์ทั้ง 4 แบบ ได้รับผลกระทบจากค่าผิดปกติจากกลุ่มที่ต่างกัน เมื่อพิจารณาจากค่าประสิทธิภาพสัมพัทธ์ พบว่า ส่วนใหญ่ค่าประสิทธิภาพสัมพัทธ์จากการทดสอบในแต่ละสถานการณ์มีค่าน้อยกว่า 1 แสดงให้เห็นว่า ค่าผิดปกติจากกลุ่มจะทำให้ผลการทดสอบทางสถิติมีการปฏิเสธ สมมติฐานหลัก เมื่อสมมติฐานหลักไม่เป็นจริงน้อยลง โดยตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน จะได้รับผลกระทบจากค่าผิดปกติจากกลุ่มมากที่สุด เนื่องจากค่าประสิทธิภาพสัมพัทธ์มีความผันแปรจาก 1 มากที่สุด ส่วนตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน สหสัมพันธ์แบบเคนดอลล์และสหสัมพันธ์แบบถ่วงน้ำหนักจะได้รับผลกระทบน้อยกว่าสหสัมพันธ์แบบเพียร์สัน เมื่อพิจารณาจากเกณฑ์ที่ใช้ในการ เปรียบเทียบความแข็งแกร่งของสัมประสิทธิ์สหสัมพันธ์ที่ใช้ในการศึกษา สามารถสรุปได้ดังนี้

1.1 ความแข็งแกร่งในการทดสอบทางสถิติ

ความแข็งแกร่งในการทดสอบทางสถิติจะพิจารณาผลการทดสอบใน 3 ลักษณะคือ ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 อำนาจในการทดสอบทางสถิติ และค่าประสิทธิภาพสัมพัทธ์ สามารถสรุปผลการศึกษาได้ดังนี้

1.1.1 ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1

สถานการณ์ส่วนใหญ่ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์ทั้ง 4 แบบ สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ยกเว้น เมื่อตัวอย่างมีขนาดเล็ก ($n = 20$) และมีค่าผิดปกติจากกลุ่มใน ตัวอย่างเท่ากับ 0% และ 10% ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ส่วนตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบเคนดอลล์ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เมื่อมีค่าผิดปกติจากกลุ่ม 20 % ของขนาดตัวอย่าง เมื่อตัวอย่างมีขนาดใหญ่ ($n \geq 50$) และมีค่าผิดปกติจากกลุ่มจำนวน 0% - 30% ของขนาดตัวอย่าง ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์ทั้ง 4 แบบ สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี รายละเอียดดังตารางที่ 1

ตารางที่ 1 สรุปผลความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์ทั้ง 4 แบบ จำแนกตามจำนวนค่าผิดปกติจากกลุ่ม

ค่าผิดปกติ จากกลุ่ม	r_{xy}	r_s	τ	r_b
0%	ควบคุมไม่ได้ เมื่อ $n = 20$	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์
5 %	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์
10 %	ควบคุมไม่ได้ เมื่อ $n = 20$	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์
20 %	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์	ควบคุมไม่ได้ เมื่อ $n = 20$	ควบคุมได้ทุก สถานการณ์
30 %	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์	ควบคุมได้ทุก สถานการณ์

1.1.2 อำนาจการทดสอบทางสถิติ

กรณีที่ไม่มีค่าผิดปกติจากกลุ่มในตัวอย่าง และตัวอย่างมีขนาดเล็ก ($n=20$) ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมนและตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบถ่วงน้ำหนัก มีอำนาจในการทดสอบสูงที่สุด แต่เมื่อตัวอย่างมีขนาดใหญ่ ($n \geq 50$) ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันมีอำนาจในการทดสอบทางสถิติสูงที่สุด

กรณีที่มียค่าผิดปกติจากกลุ่มตั้งแต่ 5% - 20% ของขนาดตัวอย่าง ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบถ่วงน้ำหนักและสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมนมีอำนาจในการทดสอบทางสถิติสูงที่สุดและมีค่าใกล้เคียงกัน เมื่อมีค่าผิดปกติจากกลุ่มจำนวนมากคือ 30% ของขนาดตัวอย่าง ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมนมีอำนาจในการทดสอบทางสถิติสูงที่สุด แต่ถ้าค่าสัมประสิทธิ์สหสัมพันธ์ของประชากรมีค่ามากแล้ว ($\rho \geq 0.80$) การมีค่าผิดปกติจากกลุ่มในตัวอย่างจะไม่ส่งผลกระทบต่ออำนาจในการทดสอบทางสถิติของตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์ทั้ง 4 แบบ รายละเอียดดังตารางที่ 2

ตารางที่ 2 ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์ที่มีอำนาจการทดสอบสูงที่สุดในแต่ละสถานการณ์ จำแนกตามจำนวนค่าผิดปกติจากกลุ่ม ขนาดตัวอย่าง และขนาดความสัมพันธ์

ค่าผิดปกติจากกลุ่ม	n	$\rho = 0.20$	$\rho = 0.50$	$\rho = 0.80$	$\rho = 1.00$
0%	20	r_s	r_b	r_b	r_{xy}, r_s, r_b, T
	50	r_{xy}	r_{xy}	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T
	100	r_{xy}	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T
5 %	20	r_s	r_b	r_s, r_b	r_{xy}, r_s, r_b, T
	50	r_b	r_b	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T
	100	r_b	r_s, r_b	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T
10 %	20	r_s, r_b	r_s	r_b	r_{xy}, r_s, r_b, T
	50	r_b	r_s, r_b	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T
	100	r_b	r_s, r_b, T	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T
20 %	20	r_s	r_s, r_b	r_b	r_{xy}, r_s, r_b, T
	50	r_b	r_s, r_b	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T
	100	r_s	r_s, r_b	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T
30 %	20	r_s	r_s	r_s	r_{xy}, r_s, r_b, T
	50	r_s	r_s	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T
	100	r_s	r_s	r_{xy}, r_s, r_b, T	r_{xy}, r_s, r_b, T

1.2 ความแกร่งในการประมาณค่า

การพิจารณาความแกร่งในการประมาณค่า จะพิจารณาจากค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (MSE) โดยตัวประมาณค่าที่ให้ค่า MSE ต่ำกว่าจะมีความแกร่งในการประมาณค่า จากการศึกษาสรุปได้ว่า ค่า MSE จะแปรผกผันกับค่าสัมประสิทธิ์สหสัมพันธ์ของประชากรและขนาดตัวอย่าง แต่จะแปรผันตามค่าผิดปกติจากกลุ่ม เมื่อมีค่าผิดปกติจากกลุ่มในตัวอย่างตั้งแต่ 5% - 10% ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบถ่วงน้ำหนัก จะมีค่า MSE ต่ำที่สุด รองลงมาคือ ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน แต่เมื่อจำนวนค่าผิดปกติจากกลุ่มในตัวอย่างมากกว่า 10% ของขนาดตัวอย่าง ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน จะมีค่า MSE ต่ำที่สุด ดังนั้นตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบถ่วงน้ำหนักจะมีความแกร่งในการประมาณค่าเมื่อมีค่าผิดปกติจากกลุ่มในตัวอย่างไม่เกิน 10% และ ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน มีความแกร่งในการประมาณค่าเมื่อมีค่าผิดปกติจากกลุ่มในตัวอย่างมากกว่า 10% รายละเอียดดังตารางที่ 3

ตารางที่ 3 ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่า MSE ต่ำที่สุดในแต่ละสถานการณ์ จำแนกตามจำนวนค่าผิดปกติจากกลุ่ม ขนาดตัวอย่าง และขนาดความสัมพันธ์

ค่าผิดปกติ จากกลุ่ม	n	$\rho = 0.00$	$\rho = 0.20$	$\rho = 0.50$	$\rho = 0.80$	$\rho = 1.00$
5 %	20	r_b	r_b	r_b	r_b	r_s
	50	r_b	r_b	r_b	r_b	r_s
	100	r_b	r_b	r_b	r_b	r_s
10 %	20	r_b	r_b	r_b	r_s	r_s
	50	r_b	r_b	r_b	r_b	r_s
	100	r_b	r_b	r_b	r_b	r_s
20 %	20	r_b	r_s	r_b	r_s	r_s
	50	r_s	r_s	r_s	r_s	r_s
	100	r_s	r_b	r_s	r_s	r_s
30 %	20	r_s	r_s	r_s	r_s	r_s
	50	r_s	r_s	r_s	r_s	r_s
	100	r_s	r_s	r_s	r_s	r_s

ดังนั้นเมื่อพิจารณาจากเกณฑ์ที่ใช้ในการเปรียบเทียบความแกร่งของสัมประสิทธิ์สหสัมพันธ์ที่ใช้ในการศึกษา และ พิจารณาคุณสมบัติของสถิติที่มีความแกร่งตามเกณฑ์ของ Mosteller และ Tukey (1977) ที่กล่าวว่า สถิติที่มีความแกร่งควรมีคุณสมบัติที่สำคัญ 2 ประการคือ

1) ความต้านทาน หมายถึง คุณสมบัติของตัวประมาณค่าหรือสถิติทดสอบที่ไม่ทำให้ผลการประมาณค่าและผลการทดสอบทางสถิติเปลี่ยนแปลงมากเมื่อข้อมูลเดิมเปลี่ยนแปลงเพียงเล็กน้อยหรือเปลี่ยนแปลงไปอย่างมาก

2) ความแกร่งด้านประสิทธิภาพ หมายถึง สถิติที่มีประสิทธิภาพ สูงกว่าแบบอื่นในทุกสถานการณ์ที่มีความผันแปร โดยเปรียบเทียบในสถานการณ์เดียวกัน

จากผลการศึกษา สรุปได้ว่า

1) เมื่อไม่มีค่าผิดปกติจากกลุ่ม และตัวอย่างมีขนาดเล็ก ($n < 50$) ตัวประมาณค่าสหสัมพันธ์แบบสเปียร์แมนมีความแกร่งมากที่สุด แต่ถ้าตัวอย่างมีขนาดใหญ่ ($n \geq 50$) ตัวประมาณค่าสหสัมพันธ์แบบเพียร์สันมีความแกร่งมากที่สุด

2) เมื่อมีค่าผิดปกติจากกลุ่ม ไม่เกิน 10% ของขนาดตัวอย่าง ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบถ่วงน้ำหนัก และ ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมนมีความแกร่งมากที่สุดและใกล้เคียงกัน ในทุกขนาดตัวอย่างที่ทำการทดสอบ

3) เมื่อมีค่าผิดปกติจากกลุ่ม มากกว่า 10% ของขนาดตัวอย่าง ตัวประมาณค่าสหสัมพันธ์แบบสเปียร์แมน มีความแกร่งมากที่สุด ในทุกขนาดตัวอย่างที่ทำการทดสอบ

4) ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน จะมีความแกร่งน้อยที่สุดเมื่อพบว่าข้อมูลมีค่าผิดปกติจากกลุ่ม ไม่ว่ากรณีใดๆ

2. ข้อเสนอแนะ

2.1 ข้อเสนอแนะที่ได้จากการวิจัย

ในการศึกษาความสัมพันธ์ระหว่างสองตัวแปรที่เป็นตัวแปรแบบต่อเนื่อง และมีมาตรวัดแบบอันตรภาค (Interval Scale) ขึ้นไป โดยทั่วไปแล้วนักวิจัยไม่ทราบความสัมพันธ์ของประชากรที่ศึกษา ดังนั้นผู้วิจัยจึงมีข้อเสนอแนะในการเลือกใช้ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์ โดยพิจารณาตามเกณฑ์ความแรงของสัมประสิทธิ์สหสัมพันธ์ที่ได้จากการศึกษาครั้งนี้ ดังนี้

2.1.1 เมื่อไม่มีค่าผิดปกติจากกลุ่ม และตัวอย่างมีขนาดเล็ก ($n < 50$) การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ควรใช้วิธีการประมาณค่าสหสัมพันธ์แบบสเปียร์แมน แต่ถ้าตัวอย่างมีขนาดใหญ่ ($n \geq 50$) ควรใช้วิธีการประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

2.1.2 ในทุกขนาดตัวอย่าง เมื่อมีค่าผิดปกติจากกลุ่ม ไม่เกิน 10% ของขนาดตัวอย่าง การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ควรใช้วิธีการประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบถ่วงน้ำหนัก หรือสหสัมพันธ์แบบสเปียร์แมน ซึ่งในกรณีนี้ตัวประมาณค่าทั้งสองมีความแรงใกล้เคียงกัน

2.1.3 ในทุกขนาดตัวอย่าง เมื่อมีค่าผิดปกติจากกลุ่ม มากกว่า 10% ของขนาดตัวอย่าง การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ควรใช้วิธีการประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน

2.1.4 ไม่ควรใช้ตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน เมื่อพบว่าข้อมูลมีค่าผิดปกติจากกลุ่ม ไม่ว่ากรณีใดๆ

2.2 ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

2.2.1 การศึกษาครั้งนี้ ผู้วิจัยกำหนดให้มีค่าผิดปกติจากกลุ่มเฉพาะในตัวแปรใดตัวแปรหนึ่งเท่านั้น จึงควรมีการศึกษาความแรงของสัมประสิทธิ์สหสัมพันธ์เมื่อเกิดค่าผิดปกติจากกลุ่ม ทั้งในตัวแปร X และ Y

2.2.2 เกณฑ์ที่ใช้ในการเปรียบเทียบความแรงของสัมประสิทธิ์สหสัมพันธ์ ควรมีการศึกษาเพิ่มเติมในส่วนของการความสามารถในการควบคุมทิศทางความสัมพันธ์ร่วมด้วย

2.2.3 การกำหนดจำนวนค่าผิดปกติจากกลุ่ม ควรเพิ่มความละเอียดสัดส่วนในแต่ละระดับให้มากขึ้น เช่น ค่าผิดปกติจากกลุ่มเท่ากับ 5%, 6%, 7%, ... ของขนาดตัวอย่าง เพื่อให้เห็นแนวโน้มการเปลี่ยนแปลงความแรงของตัวประมาณค่าสัมประสิทธิ์สหสัมพันธ์ที่ชัดเจนขึ้น

