

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์เพื่อทำการเปรียบเทียบประสิทธิภาพของการอนุมานสถิติเชิงเบย์สำหรับการสุ่มตัวอย่างแบบฮิตแอนดรีนกับการสุ่มตัวอย่างแบบกิบส์ ในกรณีที่การแจกแจงภายหลังเป็นแบบปกติที่ถูกตัดหาง เมื่อจำนวนตัวแปร(มิติ)เป็น 2, 3, 4, 5, 7 และ 10 ตัวแปร และเมทริกซ์ความแปรปรวนร่วม Σ มีค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) เป็น 0, 0.5 และ 0.9 โดยใช้เกณฑ์ในการวัดประสิทธิภาพ 2 เกณฑ์ คือ ค่าครึ่งช่วงความเชื่อมั่นที่คำนวณจากวิธีค่าเฉลี่ยกลุ่ม และ ค่า MPSRF ของ บรุกซ์-เกลแมน ซึ่งสามารถสรุปผลการวิจัยได้ดังนี้

5.1 สรุปผลการวิจัย

5.1.1 สรุปผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดรีนกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่าครึ่งช่วงความเชื่อมั่นจากวิธีค่าเฉลี่ยกลุ่ม

จากผลการเปรียบเทียบประสิทธิภาพ พบว่า

เมื่อ α มีจำนวนมิติต่ำๆ เช่น 2, 3 และ 4 มิติ การสุ่มตัวอย่างแบบฮิตแอนดรีนจะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์ ทั้งในกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ในเมทริกซ์ความแปรปรวนร่วม Σ มีค่าต่ำ ($\rho = 0$) สัมประสิทธิ์สหสัมพันธ์มีค่าปานกลาง ($\rho = 0.5$) และสัมประสิทธิ์สหสัมพันธ์มีค่าสูง ($\rho = 0.9$)

เมื่อ α เริ่มมีจำนวนมิติสูงขึ้น เช่น 5, 7 และ 10 มิติ การสุ่มตัวอย่างแบบกิบส์จะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนดรีน ทั้งในกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ในเมทริกซ์ความแปรปรวนร่วม Σ มีค่าต่ำ ($\rho = 0$) สัมประสิทธิ์สหสัมพันธ์มีค่าปานกลาง ($\rho = 0.5$) และสัมประสิทธิ์สหสัมพันธ์มีค่าสูง ($\rho = 0.9$) (ยกเว้นกรณี 7 และ 10 มิติ เมื่อ $\rho = 0.9$ ที่ไม่สามารถเปรียบเทียบกันได้)

5.1.2 สรุปผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดรีนกับการสุ่มตัวอย่างแบบกิบส์ โดยวัดจากค่า MPSRF ของ บรุกซ์-เกลแมน

จากผลการเปรียบเทียบประสิทธิภาพพบว่า

เมื่อ α มีจำนวนมิติต่ำๆ เช่น 2, 3 และ 4 มิติ การสุ่มตัวอย่างแบบฮิตแอนดรีนจะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์ ทั้งในกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ใน

เมทริกซ์ความแปรปรวนร่วม Σ มีค่าต่ำ ($\rho = 0$) สัมประสิทธิ์สหสัมพันธ์มีค่าปานกลาง ($\rho = 0.5$) และสัมประสิทธิ์สหสัมพันธ์มีค่าสูง ($\rho = 0.9$)

เมื่อ α เริ่มมีจำนวนมิติสูงขึ้น เช่น 5, 7 และ 10 มิติ การสุ่มตัวอย่างแบบกิบส์จะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนด์รัน ทั้งในกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ในเมทริกซ์ความแปรปรวนร่วม Σ มีค่าต่ำ ($\rho = 0$) สัมประสิทธิ์สหสัมพันธ์มีค่าปานกลาง ($\rho = 0.5$) และสัมประสิทธิ์สหสัมพันธ์มีค่าสูง ($\rho = 0.9$) (ยกเว้นกรณี 10 มิติ เมื่อ $\rho = 0.9$ ที่ไม่สามารถเปรียบเทียบกันได้)

ตารางที่ 5.1 แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 และค่า MPSRF ของบรู๊คซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

k	rho	half width		MPSRF	
		Hit-and-Run	Gibbs	Hit-and-Run	Gibbs
2	0	0.0295	0.0354	1.01	1.02
	0.5	0.0344	0.0541	1.01	1.02
	0.9	0.0413	0.1048	1.01	1.06
3	0	0.0398	0.0527	1.03	1.04
	0.5	0.0522	0.0639	1.01	1.02
	0.9	0.0556	0.1412	1.01	1.08
4	0	0.0394	0.0403	1.02	1.03
	0.5	0.0551	0.0808	1.01	1.02
	0.9	0.0629	0.1050	1.00	1.04

ตารางที่ 5.1 (ต่อ) แสดงค่าครึ่งช่วงความเชื่อมั่นของตัวแปรตัวที่ 1 และค่า MPSRF ของบรูกซ์-เกลแมน จากการสุ่มตัวอย่างแบบฮิตแอนด์รันและแบบกิบส์

k	rho	half width		MPSRF	
		Hit-and-Run	Gibbs	Hit-and-Run	Gibbs
5	0	0.0571	0.0383	1.03	1.01
	0.5	0.0887	0.0861	1.04	1.01
	0.9	0.1323	0.1144	1.17	1.08
7	0	0.0730	0.0497	1.09	1.08
	0.5	0.1314	0.1163	1.07	1.01
	0.9	*0.2099	*0.2816	1.14	1.10
10	0	0.1046	0.0510	1.09	1.02
	0.5	0.2130	0.1650	1.08	1.03
	0.9	*0.2974	*0.3265	*1.31	*1.48

* หมายถึง ทั้งการจำลองด้วยการสุ่มตัวอย่างแบบฮิตแอนด์รัน และแบบกิบส์ ยังไม่ลู่เข้า จึงไม่สามารถทำการเปรียบเทียบในกรณีนั้นได้

5.2 อภิปรายผลการวิจัย

จากผลการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ โดยใช้เกณฑ์ในการเปรียบเทียบ 2 เกณฑ์ คือ วัดจากค่าครึ่งช่วงความเชื่อมั่น จากวิธีค่าเฉลี่ยกลุ่ม และ วัดจากค่า MPSRF ของ บรูกซ์-เกลแมน พบว่าได้ผลการเปรียบเทียบในทำนองเดียวกันทั้ง 2 เกณฑ์ นั่นคือ ในกรณีที่ α มีจำนวนมิติต่ำๆ เช่น 2, 3 และ 4 มิติ การสุ่มตัวอย่างแบบฮิตแอนด์รันจะมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบกิบส์ ซึ่งผลที่ได้นี้สอดคล้องกับผลจากงานวิจัยของ Chen M.-H. and Schmeiser B.W. (1993) ซึ่งได้ทำการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนด์รันกับการสุ่มตัวอย่างแบบกิบส์ในกรณี

ที่เป็นการแจกแจงแบบปกติของ 2 ตัวแปร(ไม่ตัดหาง) ส่วนในกรณีที่ α มีจำนวนมิติเพิ่มสูงขึ้น เช่น 5, 7 และ 10 มิติ จากผลที่ได้กลับพบว่า การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนดรีน ซึ่งสาเหตุที่เป็นเช่นนี้อาจเป็นเพราะ ในกรณีที่เราศึกษานั้น ค่าพารามิเตอร์ที่เป็นค่าคาดหวัง $\alpha = (\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(k)})$ มีการเรียงอันดับอย่างสมบูรณ์ พารามิเตอร์แต่ละตัวจะถูกจำกัดค่าด้วยพารามิเตอร์ตัวที่อยู่ติดกัน สำหรับการสุ่มตัวอย่างแบบฮิตแอนดรีน ในการจำลองแต่ละรอบจะทำการสุ่มเลือกทิศทาง ซึ่งค่าของตัวอย่างสามารถเปลี่ยนพร้อมกันได้ทุกมิติ แต่ในในแต่ละมิติค่าของตัวอย่างจะเปลี่ยนได้ในระยะทางที่จำกัด ยิ่งจำนวนมิติมาก ข้อจำกัดก็จะยิ่งมาก การเปลี่ยนแปลงค่าในการจำลองแต่ละรอบ ถึงแม้จะเปลี่ยนพร้อมๆ กันได้ทุกมิติก็ตาม แต่ในแต่ละมิติจะเปลี่ยนได้เล็กน้อยเท่านั้น เมื่อเทียบกับการสุ่มตัวอย่างแบบกิบส์ ซึ่งการจำลองแต่ละรอบ ค่าของตัวอย่างจะเปลี่ยนได้ครั้งละหนึ่งมิติโดยที่ค่าในมิติอื่นคงที่ แต่ค่าที่เปลี่ยนในมิตินั้นอาจจะเปลี่ยนได้ในระยะทางที่มากกว่า ดังนั้นเมื่อจำนวนมิติมีค่าสูงๆ จึงทำให้การสุ่มตัวอย่างแบบกิบส์มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบฮิตแอนดรีน

5.3 ข้อเสนอแนะ

1. เมื่อต้องการประมาณค่าคาดหวังของผลตอบแทน ในกรณีที่พารามิเตอร์ที่เป็นค่าคาดหวัง มีการเรียงอันดับอย่างสมบูรณ์ และเวกเตอร์ค่าคาดหวังมีจำนวนมิติต่ำกว่า 5 มิติ ควรเลือกใช้วิธีการสุ่มตัวอย่างแบบฮิตแอนดรีน แต่ถ้าเวกเตอร์ค่าคาดหวังมีจำนวนมิติตั้งแต่ 5 มิติขึ้นไป ควรเลือกใช้วิธีการสุ่มตัวอย่างแบบกิบส์
2. ในการวิจัยครั้งนี้ ได้ทำการกำหนดจำนวนรอบในการจำลองข้อมูลที่แน่นอน แล้วจึงทำการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างแบบฮิตแอนดรีนและแบบกิบส์ โดยเปรียบเทียบที่มิติเดียวกัน และจำนวนรอบที่จำลองข้อมูลเท่ากัน ซึ่งในบางกรณี การจำลองข้อมูลด้วยวิธีทั้ง 2 วิธี อาจจะยังไม่ลู่เข้าดีนัก จึงทำให้การเปรียบเทียบประสิทธิภาพในกรณีนั้นอาจเกิดความผิดพลาดขึ้นได้ ซึ่งการแก้ปัญหาดังกล่าวอาจทำได้โดย เพิ่มจำนวนรอบในการจำลองข้อมูลในกรณีนั้น หรืออาจทำการกำหนดเกณฑ์ขั้นต่ำที่ยอมรับได้ว่าลู่เข้า แล้วเปรียบเทียบประสิทธิภาพจากจำนวนรอบในการจำลองข้อมูลจนเข้าเกณฑ์ขั้นต่ำนั้น