

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการวิจัยเพื่อเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบ Z_r , Z_k และ Z_v ในการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์เมื่อข้อมูลมีการแจกแจงปกติทวิ และแกมมาทวิ โดยที่ผู้วิจัยกำหนดสถานการณ์ต่างๆ ดังนี้

1. การแจกแจงของประชากรเป็นการแจกแจงปกติทวิ ซึ่งเรากำหนดให้ค่า $\mu_1 = \mu_2 = 0$ และ $\sigma_1^2 = \sigma_2^2 = 1$
2. การแจกแจงของประชากรเป็นการแจกแจงแกมมาทวิ ซึ่งเรากำหนดให้ค่า $\alpha_1 = \alpha_2 = 5, 7$ และ $\beta_1 = \beta_2 = 1$ ตามลำดับ
3. ขนาดตัวอย่างเท่ากับ 10, 15 และ 20 ตามลำดับ
4. การทดสอบประกอบด้วยกรณีทีวิเคราะห์ข้อมูลทั้งหมด และกรณีที่ข้อมูลถูกคัดทิ้งทางขวา 10% และ 20%
5. ค่าสัมประสิทธิ์สหสัมพันธ์ที่ใช้ในการทดสอบเท่ากับ 0, 0.05, 0.10, 0.15, 0.3, 0.5 และ 0.8 ตามลำดับ
6. กำหนดระดับนัยสำคัญ $\alpha = 0.05$ และ 0.10

การสรุปว่าตัวสถิติทดสอบใดมีความเหมาะสมสำหรับในการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์เมื่อข้อมูลมีการแจกแจงปกติทวิและแกมมาทวิ จะพิจารณาจากความสามารถในการควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจในการทดสอบของตัวสถิติทดสอบ ซึ่งสรุปได้ดังนี้

การเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของตัวสถิติทดสอบ Z_r , Z_k และ Z_v

ก) ข้อมูลสมบูรณ์

ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของตัวสถิติทดสอบทั้ง 3 ตัวเมื่อข้อมูลมีการแจกแจงปกติทวิและแกมมาทวิจะให้ผลสรุปของการควบคุมความคลาดเคลื่อนประเภทที่ 1 แตกต่างกันดังนี้

กรณีที่ 1 การแจกแจงปกติทวิ

ตัวสถิติทดสอบ Z_r และ Z_x จะสามารถควบคุมค่าความคลาดเคลื่อนประเภทที่ 1 ได้ทุกกรณี ส่วนตัวสถิติทดสอบ Z_v จะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ไม่ได้เมื่อค่าสัมประสิทธิ์สหสัมพันธ์มีค่าสูง

กรณีที่ 2 การแจกแจงแกมมาทวิ

ตัวสถิติทดสอบ Z_r และ Z_x จะสามารถควบคุมค่าความคลาดเคลื่อนประเภทที่ 1 ได้ทุกกรณี ส่วนตัวสถิติทดสอบ Z_v จะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ไม่ได้เมื่อค่า $\alpha_1 = \alpha_2 = 7$ ค่าสัมประสิทธิ์สหสัมพันธ์มีค่าสูงและขนาดตัวอย่างมีค่าน้อย

บ) ข้อมูลที่ถูกคัดทิ้งทางขวา

ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของตัวสถิติทดสอบทั้ง 3 ตัวเมื่อข้อมูลมีการแจกแจงปกติทวิและแกมมาทวิจะให้ผลสรุปของการควบคุมความคลาดเคลื่อนประเภทที่ 1 แตกต่างกันดังนี้

กรณีที่ 1 การแจกแจงปกติทวิ

ตัวสถิติทดสอบ Z_r จะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ทุกค่าสัมประสิทธิ์สหสัมพันธ์ที่มีขนาดตัวอย่าง = 10 ส่วนตัวสถิติทดสอบ Z_x จะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ เมื่อค่าสัมประสิทธิ์สหสัมพันธ์มีค่าสูง ส่วนตัวสถิติทดสอบ Z_v จะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้โดยทั่วไปและจะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ไม่ได้เมื่อค่าสัมประสิทธิ์สหสัมพันธ์มีค่าสูง

กรณีที่ 2 การแจกแจงแกมมาทวิ

ตัวสถิติทดสอบ Z_r จะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ไม่ได้กรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์มีค่าสูง และเกือบทุกค่าสัมประสิทธิ์สหสัมพันธ์เมื่อใช้ขนาดตัวอย่าง = 10 ส่วนตัวสถิติทดสอบ Z_x จะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ในบางกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์มีแนวโน้มสูง , เมื่อใช้ขนาดตัวอย่าง = 15 , 20 ส่วนตัวสถิติทดสอบ Z_v จะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้โดยทั่วไปและจะควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ในบางกรณีที่ค่าสัมประสิทธิ์สหสัมพันธ์มีค่าสูง

ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของตัวสถิติทดสอบทั้ง 3 ตัว โดยทั่วไปจะมีแนวโน้มลดลงเมื่อขนาดตัวอย่าง(กรณีข้อมูลสมบูรณ) ค่าสเกลพารามิเตอร์ และสัมประสิทธิ์สหสัมพันธ์มีค่าเพิ่มขึ้น ส่วนการเพิ่มขนาดตัวอย่างในกรณีข้อมูลถูกตัดทิ้งทางขวา(๑)จะทำให้ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เกิดขึ้น 2 กรณี คือกรณีแรก เมื่อ $\alpha = 10\%$ การเพิ่มขนาดตัวอย่างจะทำให้ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีแนวโน้มลดลงยกเว้นเมื่อข้อมูลมีการแจกแจงแกมมาทวิและมีค่าสเกลพารามิเตอร์เท่ากับ 5 โดยทั่วไปค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จะมีแนวโน้มเพิ่มขึ้น กรณีที่สอง เมื่อ $\alpha = 20\%$ การเพิ่มขนาดตัวอย่างจะทำให้ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีแนวโน้มเพิ่มขึ้น ส่วนการเพิ่มจำนวนข้อมูลถูกตัดทิ้งทางขวาจะทำให้ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เกิดขึ้นสองกรณี คือกรณีแรก เมื่อขนาดตัวอย่างมีค่าน้อยการเพิ่มค่า α จะส่งผลให้ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีค่าลดลง กรณีที่สอง เมื่อขนาดตัวอย่างเท่ากับ 20 การเพิ่มค่า α จะส่งผลให้ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีค่าเพิ่มขึ้น สำหรับการเพิ่มระดับนัยสำคัญ α จะส่งผลให้ตัวสถิติทดสอบทั้งหมดควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ดีขึ้น

การเปรียบเทียบค่าอำนาจการทดสอบของตัวสถิติทดสอบ Z_r , Z_k และ Z_v

ค่าอำนาจการทดสอบของตัวสถิติทดสอบทั้ง 3 ตัว เมื่อข้อมูลมีการแจกแจงปกติทวิและแกมมาทวิจะให้อันดับของค่าอำนาจการทดสอบของตัวสถิติทดสอบสูงสุดต่ำสุดซึ่งมีลักษณะคล้ายกันดังนี้

ก) ข้อมูลสมบูรณ

ตัวสถิติทดสอบ Z_r และ Z_k จะให้ค่าอำนาจการทดสอบใกล้เคียงกันโดยทั่วไป และจะให้ค่าอำนาจการทดสอบสูงสุดเมื่อค่าสัมประสิทธิ์สหสัมพันธ์(ρ) มีค่าน้อยถึงปานกลาง ส่วนตัวสถิติทดสอบ Z_v จะให้ค่าอำนาจการทดสอบสูงสุดเมื่อค่า $\rho = 0$

ข) ข้อมูลที่ถูกตัดทิ้งทางขวา

ตัวสถิติทดสอบ Z_k จะให้ค่าอำนาจการทดสอบสูงที่สุดโดยทั่วไปเมื่อ ρ มีค่าน้อยถึงปานกลาง ส่วนตัวสถิติทดสอบ Z_v จะให้ค่าอำนาจการทดสอบสูงสุดเมื่อ $\rho = 0$

ปัจจัยที่มีอิทธิพลต่อค่าอำนาจการทดสอบของตัวสถิติทดสอบทั้ง 3 ตัว คือ

1. ค่าอำนาจการทดสอบจะแปรผันตามขนาดตัวอย่างและระดับนัยสำคัญ

2. ค่าอำนาจการทดสอบจะแปรผกผันกับจำนวนข้อมูลที่ถูกตัดทิ้งทางขวาและค่าสเกลพารามิเตอร์ในการแจกแจงแกมมาทวิ

3. เมื่อกำหนดให้ค่าสัมประสิทธิ์สหสัมพันธ์ในสมมติฐานว่างน้อยกว่าค่าสัมประสิทธิ์สหสัมพันธ์ในสมมติฐานทางเลือกอื่น การเพิ่มค่าสัมประสิทธิ์สหสัมพันธ์ในสมมติฐานว่างจะทำให้ค่าอำนาจการทดสอบเพิ่มขึ้นยกเว้นบางกรณีของตัวสถิติทดสอบ Z_{α} ซึ่งที่ทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าน้อยถึงปานกลาง, ข้อมูลถูกตัดทิ้งทางขวาและขนาดตัวอย่างมีค่าน้อย

ข้อเสนอแนะ

ผู้วิจัยมีข้อเสนอแนะเกี่ยวกับงานวิจัยดังนี้

1. ผู้วิจัยคิดว่าควรมีการศึกษาในกรณีที่กำหนดให้สเกลพารามิเตอร์แต่ละกลุ่มประชากรมีค่าไม่เท่ากันข้อมูลที่มีลักษณะเช่นนี้ส่วนใหญ่จะเป็นข้อมูลทางการแพทย์เช่น การหาความสัมพันธ์ระหว่างอายุของคนที่เกี่ยวข้องกับปริมาณคอเลสเตอรอลในร่างกาย

2. ผู้วิจัยคิดว่าควรจะมีการศึกษาในกรณีที่ข้อมูลมีค่าถูกตัดทิ้งทางซ้าย (left censored data) และข้อมูลถูกตัดทิ้งแบบสุ่ม (random censored data) เพื่อสรุปว่ากรณีข้อมูลแบบไหนที่ทำให้ตัวสถิติทดสอบมีอำนาจการทดสอบสูงซึ่งข้อมูลที่มีลักษณะที่ถูกตัดทิ้งทางซ้ายและถูกตัดทิ้งแบบสุ่มส่วนใหญ่จะพบในทางการแพทย์ เช่นกรณีคนไข้บางรายขาดการติดต่อกับโรงพยาบาลซึ่งกรณีนี้จะเป็นข้อมูลถูกตัดทิ้งแบบสุ่ม

แผนผังสรุปการเลือกตัวสถิติทดสอบเพื่อใช้ในการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์เมื่อข้อมูลมีการแจกแจงปกติทวิและแกมมาทวิ

