

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองเพื่อศึกษาเปรียบเทียบอำนาจการทดสอบของตัวสถิติ 3 ตัวที่ใช้ทดสอบการแจกแจงแบบเอกซ์โปเนนเชียล โดยใช้ตัวสถิติทดสอบ Regression Test (Z) ตัวสถิติทดสอบ Kolmogorov-Smirnov Test (K) และตัวสถิติทดสอบ Gnedenko F test (F) ในสถานการณ์ต่าง ๆ ที่กำหนดขึ้นดังนี้

- 1) การแจกแจงของประชากรเป็นแบบเอกซ์โปเนนเชียล แบบแกมมา แบบไวบูลล์ แบบลอกนอร์มอลและโคสแควร์
- 2) ขนาดตัวอย่างที่สุ่มจากประชากร คือ 30, 50 และ 100
- 3) กำหนดระดับนัยสำคัญของการทดสอบคือ 0.05 และ 0.10
- 4) การทดสอบตัวสถิติประกอบด้วยกรณีที่วิเคราะห์ข้อมูลทั้งหมด (complete data) และการวิเคราะห์ข้อมูลเมื่อมีข้อมูลขาดหาย (censored data) โดยพิจารณาเปอร์เซ็นต์ของข้อมูลขาดหายเป็น 10% และ 20% ทางซ้ายและทางขวาของข้อมูลตามลำดับและเมื่อข้อมูลขาดหายเท่ากันทั้งทางซ้ายและทางขวา 10% และ 20% ตามลำดับ

การวิจัยครั้งนี้ใช้การจำลองด้วยเทคนิคมอนติคาร์โลกับเครื่อง IBM 370/3031 โดยใช้โปรแกรมฟอร์แทรน 77 เพื่อสร้างข้อมูลและสถานการณ์ต่าง ๆ ตามแผนการทดลอง โดยทำซ้ำทั้งหมด 500 ครั้งในแต่ละสถานการณ์

5.1 สรุปผลการวิจัย

การสรุปผลว่าตัวสถิติทดสอบใดมีความเหมาะสมสำหรับทดสอบการแจกแจงแบบเอกซ์โปเนนเชียล พิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลองเป็นอันดับแรก แล้วจึงพิจารณาอำนาจการทดสอบของตัวสถิติทดสอบเป็นอันดับต่อไป ซึ่งผลสรุปทั้ง 2 ขั้นตอนเป็นดังนี้

5.1.1 เปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของตัวสถิติทดสอบ 3 ตัว ณ ระดับนัยสำคัญ 0.05 และ 0.10 ผลการเปรียบเทียบเป็นดังนี้

- 1) เกณฑ์ของ Bladley

ตัวสถิติทดสอบ K ไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อน

ประเภทที่ 1 เพียงกรณีเดียว ณ ระดับนัยสำคัญ 0.05 ขนาดตัวอย่างเท่ากับ 100 และ $r_1 = 0$, $r_2 = 0$ แต่สำหรับกรณีอื่นตัวสถิติทดสอบ 3 ตัวสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้

2) เกณฑ์ของ Cochran

เราสามารถสรุปผลได้ดังนี้

ก. ตัวสถิติทดสอบ K ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งระดับนัยสำคัญ 0.05 และ 0.10 โดยความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าเกณฑ์ที่กำหนดทุกกรณี ซึ่งหมายถึงอำนาจการทดสอบของตัวสถิติทดสอบ K ให้อำนาจการทดสอบต่ำกว่าความเป็นจริงทุกกรณี

ข. ตัวสถิติทดสอบ Z และ F สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ใกล้เคียงกันประมาณ 88% ของการทดลองทั้งหมด โดยกรณีที่ตัวสถิติทดสอบ Z และ F ไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ นั้น ส่วนใหญ่เป็นการวิเคราะห์ข้อมูลที่เกิดข้อมูลขาดหายทางขวา (right censored data) ประมาณ 50% ของการทดลองที่ไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้

5.1.2 การเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบ 3 ตัวที่ใช้ทดสอบการแจกแจงแบบเอกซ์โปเนนเชียล ณ ระดับนัยสำคัญ 0.05 และ 0.10 สรุปผลได้ดังนี้

1) การวิเคราะห์ข้อมูลทั้งหมด (complete data) ตัวสถิติทดสอบ Z มีอำนาจการทดสอบสูงสุด เมื่อประชากรมีการแจกแจงแบบโคสแควร์และการแจกแจงแบบไวบูลส์ ตัวสถิติทดสอบ K มีอำนาจการทดสอบสูงสุดเมื่อประชากรมีการแจกแจงแบบแกมมา ($\alpha = 2, \beta = 1$) และการแจกแจงแบบล็อกนอร์มอล โดยทั่วไปพบว่าตัวสถิติทดสอบ F ให้อำนาจการทดสอบต่ำกว่าตัวสถิติทดสอบ Z และ K

2) การวิเคราะห์ข้อมูลเมื่อมีข้อมูลขาดหาย (censored data) ในการวิเคราะห์ตัวสถิติทดสอบ Z มีอำนาจการทดสอบสูงสุด ยกเว้นกรณีที่ขนาดตัวอย่างเท่ากับ 30 ประชากรมีการแจกแจงแบบแกมมา ($\alpha = 3, \beta = 1$) และการแจกแจงแบบโคสแควร์ที่ระดับความเป็นอิสระเท่ากับ 3 ($r_1 = 5, r_2 = 0$) และ ($r_1 = 6, r_2 = 6$) ซึ่งตัวสถิติทดสอบ F มีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบ Z โดยคิดเป็น 5.36% ของการทดลองทั้งหมด

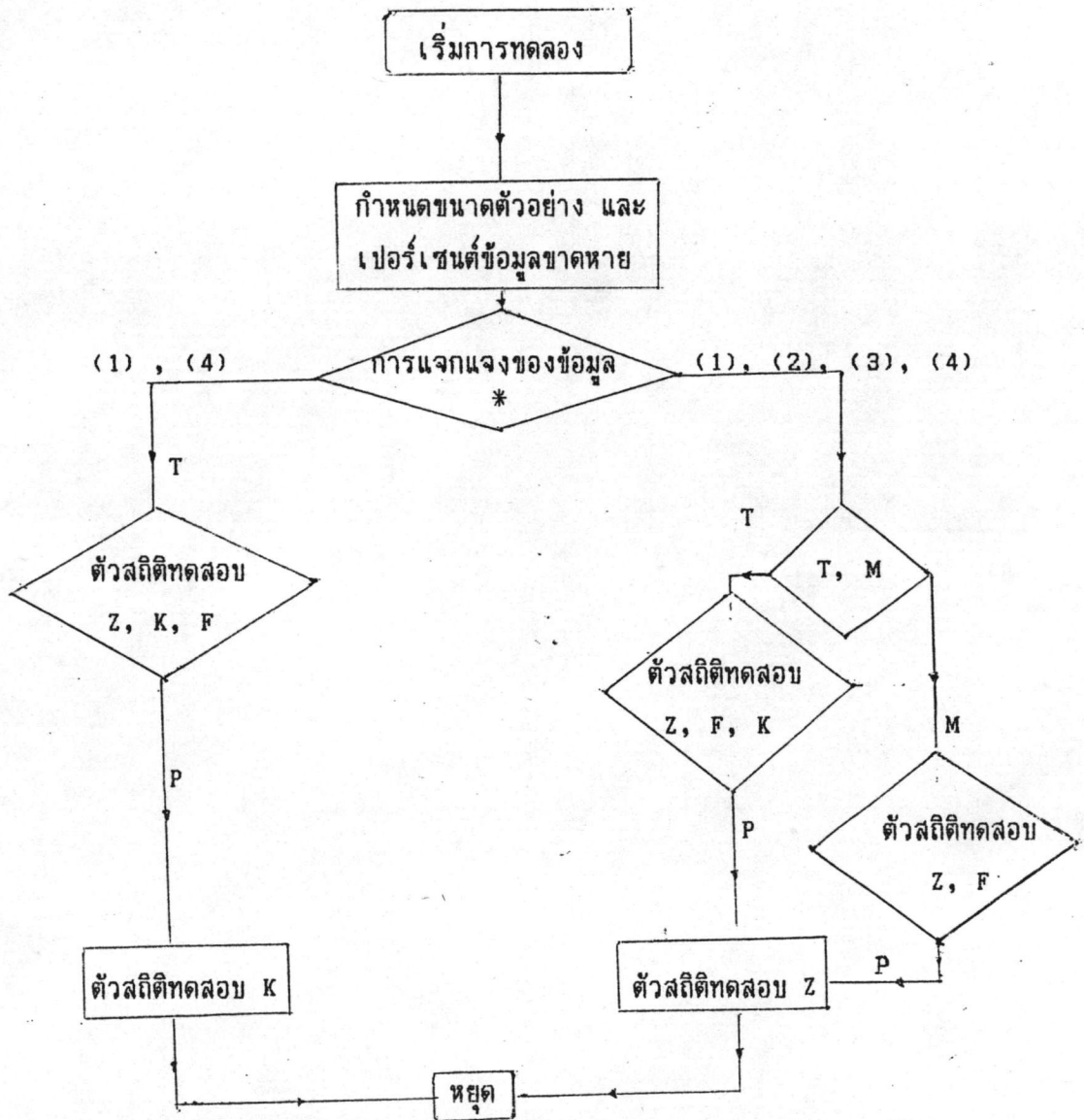
3) การเพิ่มขนาดตัวอย่างและระดับนัยสำคัญมีผลทำให้อำนาจการทดสอบสูงขึ้น

4) การเพิ่มเปอร์เซ็นต์ข้อมูลขาดหาย ในการวิเคราะห์ทางขวาให้อำนาจการทดสอบสูงกว่าการเพิ่มเปอร์เซ็นต์ข้อมูลขาดหายที่เท่ากันทางซ้าย ทุกขนาดตัวอย่างและระดับนัยสำคัญ (ยกเว้นเมื่อประชากรมีการแจกแจงแบบล็อกนอร์มอล $\mu = 0, \sigma^2 = 0.9$) ทั้งนี้เพราะว่าประชากรที่กำหนดในการทดลองครั้งนี้ส่วนใหญ่มีลักษณะเบ้ขวาทำให้ข้อมูลส่วนใหญ่อยู่ทางซ้าย เมื่อมี

ข้อมูลขาดหายเกิดขึ้นทางขวา จึงทำให้อำนาจการทดสอบของตัวสถิติทุกตัวสูงกว่า เมื่อมีข้อมูลขาดหายทางซ้ายในขนาดตัวอย่างที่เท่ากัน

สรุปการเลือกตัวสถิติทดสอบที่ให้อำนาจการทดสอบสูงสุดในสถานการณ์ต่าง ๆ เป็นแผนผัง

ดังนี้



- | | |
|--|-----------------------------|
| T = วิเคราะห์ข้อมูลทั้งหมด | (2) = การแจกแจงแบบไวบูลส์ |
| M = วิเคราะห์ข้อมูล เมื่อมีข้อมูลขาดหาย | (3) = การแจกแจงแบบโคสแควร์ |
| P = ตัวสถิติทดสอบที่ให้อำนาจการทดสอบสูงสุด | (4) = การแจกแจงแบบลอการิธึม |
| (1) = การแจกแจงแบบแกมมา | * = (1), (2), (3), (4) |

5.2 ปัญหาที่เกิดขึ้นในการวิจัย

การหาอำนาจการทดสอบของตัวสถิติทดสอบทั้ง 3 ตัว ในกรณีที่ประชากรมีการแจกแจงแบบแกมมา เมื่อกำหนด shape parameter (α) มากกว่าหรือเท่ากับ 3 ขึ้นไปนั้น การประมวลผลต้องใช้เวลานานมากการวิจัยครั้งนี้จึงทำการทดลองที่ shape parameter เท่ากับ 2 และ 3 เท่านั้น

5.3 ข้อเสนอแนะ

1) ด้านการเลือกตัวสถิติทดสอบไปใช้ จะพิจารณาตัวสถิติที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้มากที่สุด และมีอำนาจการทดสอบสูงสุด ควรเลือกตัวสถิติทดสอบ Regression Test (Z) ทดสอบการแจกแจงแบบเอกซ์โปเนนเชียลกรณีที่ไม่ทราบแน่ชัดเกี่ยวกับลักษณะของประชากร

2) ด้านการวิจัย

ก. ควรศึกษาวิจัยตัวสถิติอื่นที่ใช้ทดสอบการแจกแจงแบบเอกซ์โปเนนเชียล เช่น Gini Statistics, On a Class of Tests of Exponentiality และ Tiku Test (ในกรณีตัวอย่างขนาดเล็ก) เพื่อนำมาเปรียบเทียบกับตัวสถิติที่ผู้วิจัยได้ศึกษาไว้

ข. ควรศึกษาเพิ่มเติมกรณีการเพิ่มเปอร์เซ็นต์ข้อมูลขาดหายในการวิเคราะห์ทั้งทางซ้ายและทางขวา (left and right censored data) โดยเฉพาะข้อมูลขาดหายทางขวา ซึ่งในการวิจัยครั้งนี้มีความสำคัญมากกว่าข้อมูลขาดหายทางซ้ายและพบมากในการปฏิบัติ เช่น ข้อมูลอายุ

ค. ควรศึกษาตัวสถิติทดสอบการแจกแจงแกมมา ซึ่งเป็นกรณีทั่วไปของการแจกแจงแบบเอกซ์โปเนนเชียล ในกรณีที่ shape parameter (α) และ scale parameter (β) เปลี่ยนไป

ง. ในการศึกษางานวิจัยทั่วไป ๆ ในเรื่องข้อมูลขาดหายพบว่า กรณีที่ทำการวิเคราะห์ข้อมูลขนาดใหญ่ ($n > 100$) จำนวนเปอร์เซ็นต์ข้อมูลขาดหายไม่ควรเกิน 30% ของจำนวนข้อมูล และในกรณีการวิเคราะห์ข้อมูลขนาด $n \leq 50$ จำนวนเปอร์เซ็นต์ข้อมูลขาดหายไม่ควรเกิน 20% ของจำนวนข้อมูล