หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองเพื่อศึกษาเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบ 4 ตัว ที่ใช้ทดสอบการแจกแจงการอยู่รอดของ 2 กลุ่มตัวอย่าง ในกรณีที่มีค่าสังเกตไม่สมบูรณ์ โดยใช้ตัวสถิติทดสอบ Logrank Test (Lr) ตัวสถิติทดสอบ Generalized Wilcoxon Test (GW) ตัวสถิติทดสอบ Cox - Mantel Test (CM) และตัวสถิติทดสอบ Modified Logrank Test (MLr) ในสถานการณ์ต่าง ๆ ที่กำหนดขึ้นดังนี้

1) การแจกแจงของประชากรเป็นแบบเอกซ์โปเนนเชียล แบบลอกนอร์มอลและแบบไวบูลล์

2) ขนาดตัวอย่างที่สุ่มมาจากประชากร คือ 20 และ 50

3) กำหนดระดับนัยสำคัญของการทดสอบคือ 0.05 และ 0.10

4) การทดสอบตัวสถิติประกอบด้วยกรณีที่วิเคราะห์ข้อมูลเมื่อมีค่าสังเกตไม่สมบูรณ์แบบสุ่ม (random cersored data) และเมื่อมีค่าสังเกตไม่สมบูรณ์ทางขวา (right censored data) โดยพิจารณาเปอร์เซนต์ของค่าสังเกตไม่สมบูรณ์เป็น 10%, 20% และ 30% ของขนาดตัวอย่างในแต่ละกลุ่มดังนี้

- เปอร์เซนต์ของค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 เป็น 10% และในกลุ่มที่ 2 เป็น 10% ของขนาดตัวอย่าง
- เปอร์เซนต์ของค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 เป็น 20% และในกลุ่มที่ 2 เป็น 10% ของขนาดตัวอย่าง
- เปอร์เซนต์ของค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 เป็น 20% และในกลุ่มที่ 2 เป็น 20% ของขนาดตัวอย่าง
- เปอร์เซนต์ของค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 เป็น 30% และในกลุ่มที่ 2 เป็น 20% ของขนาดตัวอย่าง

การวิจัยครั้งนี้ใช้การจำลองด้วยเทคนิคมอนติคาร์โลกับเครื่อง AMDHAL 3860 โดยใช้โปรแกรมฟอร์แทรน 77 เพื่อสร้างข้อมูลและสถานการณ์ต่าง ๆ ตามแผนการทดลอง โดยทำซ้ำทั้งหมด 500 ครั้ง ในแต่ละสถานการณ์ เมื่อขนาดตัวอย่างเป็น 20 และทำซ้ำ 200 ครั้งในแต่ละสถานการณ์ เมื่อขนาดตัวอย่างเป็น 50

5.1 สรุปผลการวิจัย

การสรุปผลว่าตัวสถิติทดสอบใดมีความเหมาะสมสำหรับทดสอบการแจกแจงการอยู่รอดของ 2 กลุ่มตัวอย่าง ในกรณีที่มีค่าสังเกตไม่สมบูรณ์ จะพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลองเป็นอันดับแรก แล้วจึงพิจารณาอำนาจการทดสอบของตัวสถิติทดสอบเป็นอันดับต่อไป ซึ่งผลสรุปทั้ง 2 ขั้นตอนเป็นดังนี้

5.1.1 เปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของตัวสถิติทดสอบ 4 ตัว ณ.ระดับนัยสำคัญ 0.05 และ 0.10

1) กรณีค่าสังเกตไม่สมบูรณ์แบบสุ่ม

พบว่าส่วนใหญ่สถิติทดสอบ Lr จะไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อเปอร์เซนต์ค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 และ 2 เป็น 20% และเมื่อมีการแจกแจงแบบเอกซ์โปเนนเชียลพบว่า สถิติทดสอบ MLr, CM และ GW สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ใกล้เคียงกัน และเมื่อมีการแจกแจงแบบลอกนอร์มอลและแบบไวบูลล์ สถิติทดสอบ GW สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้มากที่สุด ส่วนสถิติทดสอบ CM และ MLr จะควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ใกล้เคียงกัน

2) กรณีค่าสังเกตไม่สมบูรณ์ทางขวา

พบว่าส่วนใหญ่แล้วเมื่อเปอร์เซนต์ค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 และ 2 เท่ากับ 20% แล้วสถิติทดสอบ CM, MLr และ Lr ไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ นอกจากนั้นพบว่าในกรณีทั่ว ๆ ไป แล้ว สถิติทดสอบ GW CM และ MLr สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ใกล้เคียงกัน

5.1.2 การเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบ 4 ตัว ที่ใช้ทดสอบการเปรียบเทียบการแจกแจงการอยู่รอดใน 2 กลุ่ม เมื่อมีค่าสังเกตไม่สมบูรณ์

1. การวิเคราะห์ข้อมูลกรณีที่มีค่าสังเกตไม่สมบูรณ์แบบสุ่ม

1.1 เมื่อประชากรมีการแจกแจงแบบเอกซ์โปเนนเชียล

โดยทั่วไปพบว่า สถิติทดสอบ MLr มีอำนาจการทดสอบสูงสุด และมีค่าใกล้เคียงกับสถิติทดสอบ CM

1.2 เมื่อประชากรมีการแจกแจงแบบลอกนอร์มอล

โดยทั่วไปพบว่า สถิติทดสอบ MLr มีอำนาจการทดสอบสูงสุด ยกเว้นเมื่อประชากรมีขนาดเท่ากับ 50 และเปอร์เซนต์ค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 และ 2 เป็น 20% และ 10% ตามลำดับ พบว่าสถิติทดสอบ GW มีอำนาจการทดสอบสูงสุด

1.3 เมื่อประชากรมีการแจกแจงแบบไวบูลล์

โดยทั่วไปพบว่า สถิติทดสอบ MLr มีอำนาจการทดสอบสูงสุด ยกเว้นที่ระดับนัยสำคัญ 0.05 เมื่อประชากรมีขนาดเท่ากับ 50 และเปอร์เซนต์ค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 และ 2 เป็น 30% และ 20% ตามลำดับ พบว่าสถิติทดสอบ Lr มีอำนาจการทดสอบสูงสุด รองลงมาคือ สถิติทดสอบ GW

2. การวิเคราะห์ข้อมูลกรณีที่มีค่าสังเกตไม่สมบูรณ์ทางขวา

2.1 เมื่อประชากรมีการแจกแจงแบบเอกซ์โปเนนเชียล

พบว่า สถิติทดสอบ GW จะมีอำนาจการทดสอบสูงสุด เมื่อ $H_a : S_1 > S_2$ และเมื่อ $H_a : S_1 < S_2$ พบว่า สถิติทดสอบ MLr จะมีอำนาจการทดสอบสูงที่สุด ยกเว้นเมื่อ R_1 และ R_2 เท่ากับ 20%สถิติทดสอบ GW จะมีอำนาจการทดสอบสูงสุด

2.2 เมื่อประชากรมีการแจกแจงแบบลอกนอร์มอล

โดยทั่วไปพบว่า สถิติทดสอบ MLr มีอำนาจการทดสอบสูงสุดและใกล้เคียงกับสถิติทดสอบ CM ยกเว้นเมื่อ R_1 และ R_2 เท่ากับ 20% สถิติทดสอบ GW จะมีอำนาจการทดสอบสูงสุด

2.3 เมื่อประชากรมีการแจกแจงแบบไวบูลล์

โดยทั่วไปพบว่า สถิติทดสอบ MLr มีอำนาจการทดสอบสูงสุดและใกล้เคียงกับสถิติทดสอบ CM ยกเว้นเมื่อ R_1 และ R_2 = 20% และ R_1 = 30%, R_2 = 20% สถิติทดสอบ GW จะมีอำนาจการทดสอบสูงสุด

5.4 ปัญหาที่เกิดขึ้นในการวิจัย

การหาอำนาจการทดสอบของตัวสถิติทดสอบทั้ง 4 ตัว ในกรณีที่ขนาดตัวอย่างในแต่ละกลุ่มเป็น 50 นั้น ในการประมวลผลต้องใช้เวลานานมาก การวิจัยครั้งนี้จึงทำการทดลองซ้ำ เมื่อขนาดตัวอย่างแต่ละกลุ่มเท่ากับ 50 เพียง 200 รอบเท่านั้น

5.3 ข้อเสนอแนะ

1) ด้านการเลือกตัวสถิติทดสอบไปใช้ จะพิจารณาตัวสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ และมีอำนาจการทดสอบสูงสุด

2) ด้านการวิจัย

ก. ควรศึกษาวิจัยตัวสถิติทดสอบอื่นที่ใช้ในการทดสอบการเปรียบเทียบการแจกแจงการอยู่รอด เช่น Renyi - type Test เมื่อนำมาเปรียบเทียบกับตัวสถิติที่ผู้วิจัยได้ศึกษาไว้

ข. ควรศึกษาเพิ่มเติมกรณีที่ขนาดตัวอย่างมีขนาดใหญ่ ($n > 100$) ทั้งนี้เพราะการศึกษาทางด้านการแพทย์นั้น ควรใช้ตัวอย่างขนาดใหญ่ ถ้าใช้ตัวอย่างขนาดเล็กและมีค่าสังเกตไม่สมบูรณ์เป็นจำนวนมาก อาจจะทำให้ผลการวิเคราะห์ผิดพลาดได้ และจำนวนค่าสังเกตไม่สมบูรณ์ไม่ควรมากกว่า 30 %

3) ในการวิจัยครั้งนี้จะพิจารณาถึงการนำสถิติทดสอบไปใช้ในกรณีทั่ว ๆ ไป จึงใช้เกณฑ์ของ Bladley เป็นเกณฑ์ในการตัดสินการควบคุมความคลาดเคลื่อนประเภทที่ 1 ซึ่งถ้าต้องการให้ได้สถิติทดสอบที่มีความเข้มในการทดสอบสูง ควรใช้เกณฑ์ของ Cochran เป็นเกณฑ์ในการตัดสินการควบคุมความคลาดเคลื่อนประเภทที่ 1

4) เนื่องจากในการวิจัยครั้งนี้บางกรณีที่ไม่สามารถนำผลการวิจัยไปประยุกต์ใช้ได้ เช่น กรณีที่การเลือกค่าพารามิเตอร์ในการแจกแจงไม่เหมาะสม หรือจำนวนครั้งของการกระทำซ้ำน้อยไป ดังนั้นหากมีการศึกษาในรายละเอียดอื่นต่อไปควรเลือกค่าพารามิเตอร์ของการแจกแจงอื่น ๆ ที่เหมาะสม กล่าวคือ ควรศึกษาในรูปแบบของการแจกแจง ซึ่งใกล้เคียงกับการแจกแจงแบบเอกซ์โปเนนเชียล

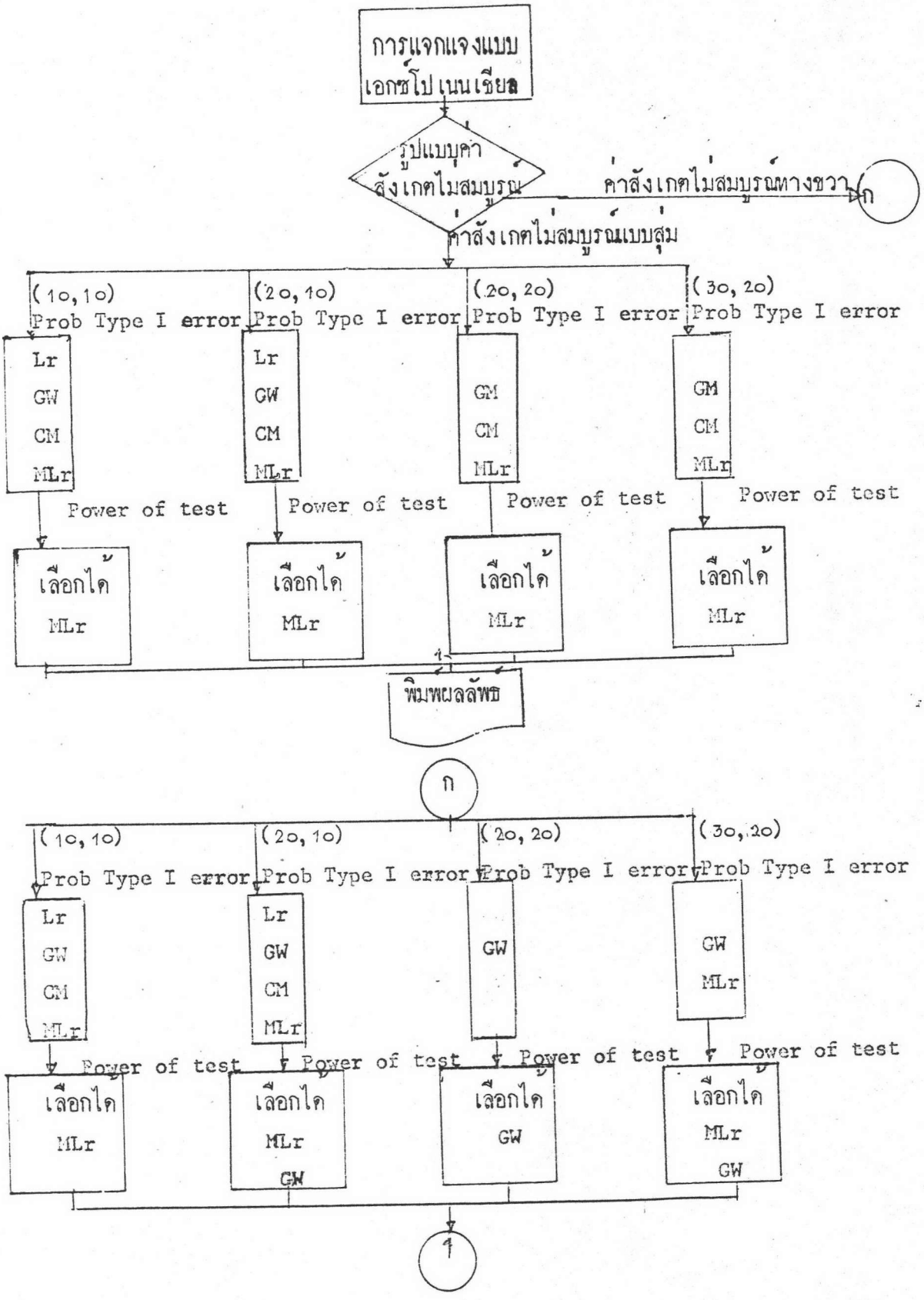
5.4 ข้อสังเกต

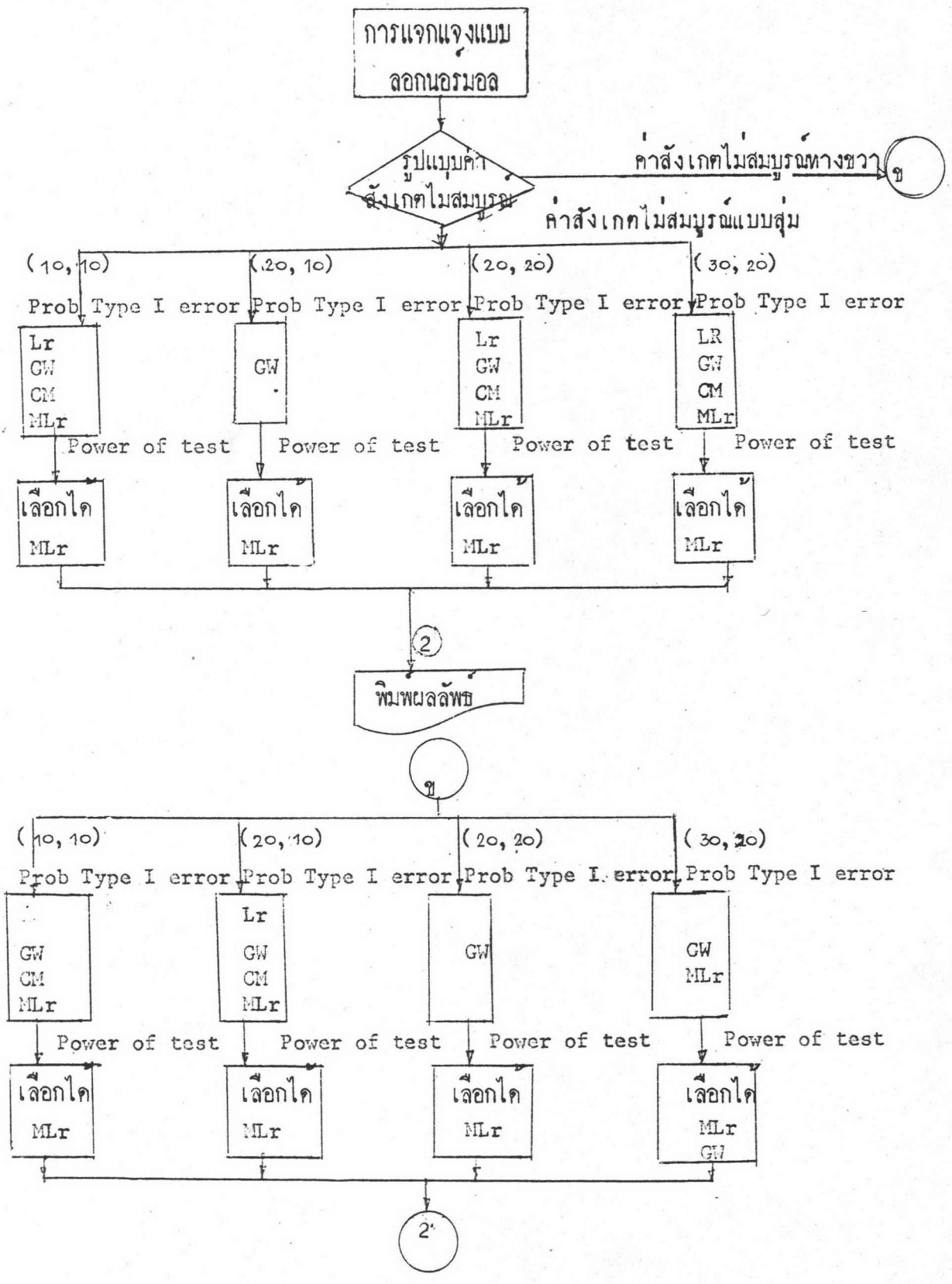
1) จากการวิจัยพบว่าในกรณีที่มีค่าสังเกตไม่สมบูรณ์ทางขวา เมื่อ $r_1 = 20\ \%$ และ $r_2 = 20\ \%$ โดยส่วนมากแล้วจะไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ซึ่งอาจเนื่องจากจำนวนของค่าสังเกตสมบูรณ์มีจำนวนน้อย และค่าของอันดับของค่าสังเกตมีค่าน้อย ทำให้ค่าสถิติทดสอบที่คำนวณได้ได้ค่าที่ไม่เหมาะสม

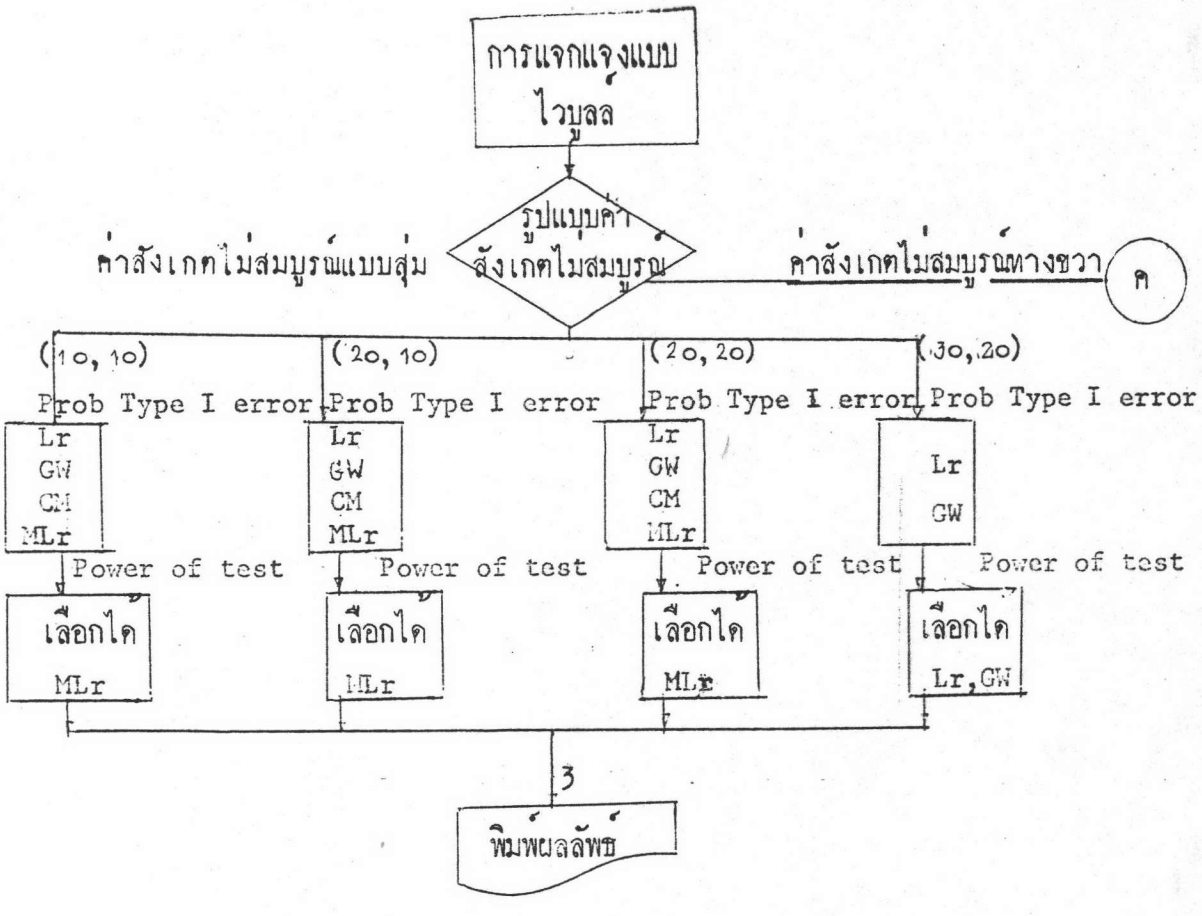
2) จากการวิจัยพบว่าเมื่อการแจกแจงมีรูปแบบใกล้เคียงกับการแจกแจงเอกซ์โปเนนเชียล จะทำให้อำนาจการทดสอบมีค่าสูง

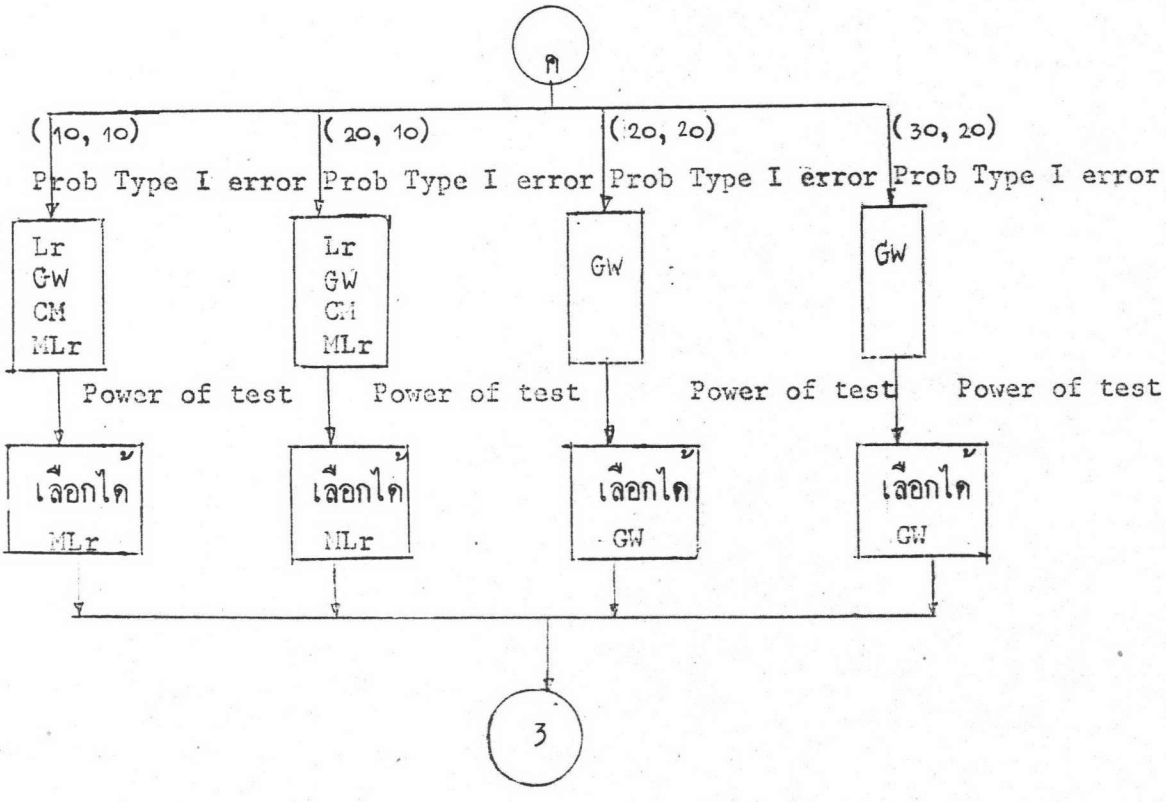
3) ค่าสังเกตไม่สมบูรณ์ทางขวาจะมีผลต่ออำนาจการทดสอบมากกว่าค่าสังเกตไม่สมบูรณ์แบบสุ่ม และพบว่าเมื่อจำนวนค่าสังเกตไม่สมบูรณ์มีค่าเท่ากันแล้ว พบว่าอำนาจการทดสอบในกรณีที่มีค่าสังเกตไม่สมบูรณ์ทางขวาจะมีค่าน้อยกว่าในกรณีที่มีค่าสังเกตไม่สมบูรณ์แบบสุ่ม

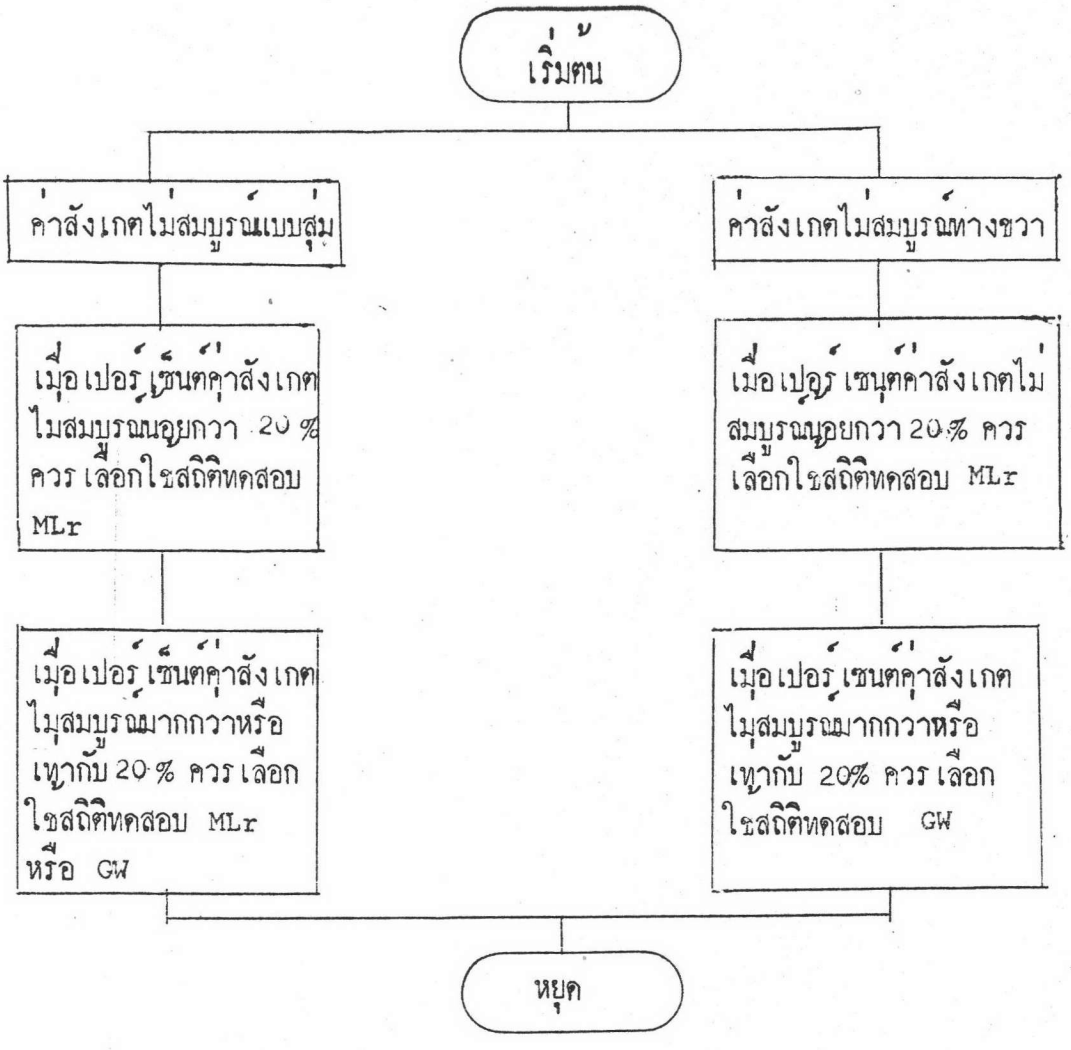
สรุปการเลือกตัวสถิติทดสอบที่ให้อำนาจการทดสอบสูงสุด ในสถานการณ์ต่าง ๆ ดังนี้




การแจกแจงแบบ
ลอกนอรมอล
รูปแบบค่า
สังเกตไม่สมบูรณ์
ค่าสังเกตไม่สมบูรณ์ทางขวา
ข
ค่าสังเกตไม่สมบูรณ์แบบสุ่ม
(10, 10)
(20, 10)
(20, 20)
(30, 20)
Prob Type I error
Prob Type I error
Prob Type I error
Prob Type I error
Lr
GW
CM
MLr
GW
Lr
GW
CM
MLr
LR
GW
CM
MLr
Power of test
Power of test
Power of test
Power of test
เลือกได
MLr
เลือกได
MLr
เลือกได
MLr
เลือกได
MLr
2
พิมพผลลัพธ
ข
(10, 10)
(20, 10)
(20, 20)
(30, 20)
Prob Type I error
Prob Type I error
Prob Type I error
Prob Type I error
GW
CM
MLr
Lr
GW
CM
MLr
GW
GW
MLr
Power of test
Power of test
Power of test
Power of test
เลือกได
MLr
เลือกได
MLr
เลือกได
MLr
เลือกได
MLr
GW
2'


การแจกแจงแบบไวบูลล
รูปแบบค่าสังเกตไม่สมบูรณ์
ค่าสังเกตไม่สมบูรณ์แบบสุ่ม
ค่าสังเกตไม่สมบูรณ์ทางขวา
ค
(10, 10)
(20, 10)
(20, 20)
(30, 20)
Prob Type I error
Lr
GW
CM
MLr
Power of test
เลือกได้
MLr
Lr, GW
3
พิมพ์ผลลัพธ์


ค
(10, 10)
(20, 10)
(20, 20)
(30, 20)
Prob Type I error
Lr
GW
CM
MLr
GW
Power of test
เลือกได้
MLr
GW
3


เริ่มต้น
ค่าสังเกตไม่สมบูรณ์แบบสุ่ม
ค่าสังเกตไม่สมบูรณ์ทางขวา
เมื่อเปอร์เซ็นต์ค่าสังเกตไม่สมบูรณ์น้อยกว่า 20 % ควรเลือกใช้สถิติทดสอบ MLr
เมื่อเปอร์เซ็นต์ค่าสังเกตไม่สมบูรณ์น้อยกว่า 20% ควรเลือกใช้สถิติทดสอบ MLr
เมื่อเปอร์เซ็นต์ค่าสังเกตไม่สมบูรณ์มากกว่าหรือเท่ากับ 20 % ควรเลือกใช้สถิติทดสอบ MLr หรือ GW
เมื่อเปอร์เซ็นต์ค่าสังเกตไม่สมบูรณ์มากกว่าหรือเท่ากับ 20% ควรเลือกใช้สถิติทดสอบ GW
หยุด