หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีการดำเนินการทดลอง

ในการวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบตัวสถิติทดสอบที่ใช้ในการทดสอบการแจกแจงการอยู่รอดของ 2 กลุ่มตัวอย่าง ในกรณีที่มีค่าสังเกตไม่สมบูรณ์ 4 วิธี คือ 'Log rank test, Generalized Wilcoxon test, Cox-Mantel test และ Modified Log rank test โดยจะศึกษาถึงอำนาจการทดสอบและความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบทั้ง 4 วิธีดังกล่าว เมื่อประชากรมีรูปแบบการแจกแจงเป็นแบบ เอกซ์โปเนนเชียล (Exponential distribution), ไวบูลล์ (Weibull distribution) และลอกนอร์มอล (Lognormal distribution) สำหรับขนาดตัวอย่าง 20 และ 50 และกำหนดให้มีเปอร์เซนต์ของค่าสังเกตไม่สมบูรณ์เป็น 10, 20 และ 30 เปอร์เซนต์ของขนาดตัวอย่างในแต่ละกลุ่ม โดยพิจารณาเป็น 2 กรณี คือ

1. เมื่อค่าสังเกตไม่สมบูรณ์เป็นค่าสังเกตไม่สมบูรณ์ทางขวา (Right-Censored observation) และ

2. เมื่อค่าสังเกตไม่สมบูรณ์เป็นค่าสังเกตไม่สมบูรณ์แบบสุ่ม (Random-Censored observation) ทั้งนี้เทคนิคที่ใช้ในการหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบ คือ วิธีมอนติคาร์โล (Monte Carlo Method) ซึ่งเป็นเทคนิคอย่างหนึ่งที่ใช้แก้ปัญหาในการคำนวณทางคณิตศาสตร์

3.1 วิธีมอนติคาร์โล (Monte Carlo Method)

เทคนิคในการจำลองตัวแบบทางคณิตศาสตร์มีอยู่หลายวิธี วิธีมอนติคาร์โลเป็นวิธีหนึ่งที่นิยมใช้กันอย่างแพร่หลายในปัจจุบัน ซึ่งหลักการของวิธีมอนติคาร์โลนี้เป็นการจำลองตัวเลขสุ่ม (Random Number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา

ในการวิจัยครั้งนี้จะใช้วิธีมอนติคาร์โลดังกล่าว ในการสร้างข้อมูลที่มีสภาพการแจกแจงตามที่ต้องการ ซึ่งขั้นตอนของวิธีมอนติคาร์โลที่ใช้กันอยู่ในปัจจุบันแบ่งเป็น 3 ขั้นตอน ดังนี้

3.1.1 การสร้างตัวเลขสุ่ม

การใช้ตัวเลขสุ่มเป็นสิ่งสำคัญมากในวิธีมอนติคาร์โล เพราะหลักการของวิธีมอนติคาร์โลนั้นจะใช้ตัวเลขมาช่วยในการหาคำตอบของปัญหา ซึ่งลักษณะของตัวเลขสุ่มจะมีการแจกแจงแบบสม่ำเสมอ (Uniform distribution) ใช้ช่วง (0,1) สำหรับวิธีการสร้างตัวเลขสุ่มนี้ มีผู้เสนอไว้หลายวิธี แต่วิธีหนึ่งที่ดีคือ วิธีที่ไวท์และชมิดท์ (White and Schemidt 1975 : 421) เสนอไว้ กล่าวคือ ลักษณะของตัวเลขสุ่มที่เกิดขึ้น มีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) และเป็นอิสระซึ่งกันและกัน

3.1.2 การประยุกต์ปัญหาที่ต้องการศึกษาเพื่อใช้กับตัวเลขสุ่ม

ซึ่งขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา บางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรง แต่อาจมีขั้นตอนอื่นอีกหลาย ๆ ขั้นตอน ซึ่งบางขั้นตอนต้องใช้ตัวเลขสุ่ม

3.1.3 การทดลองกระทำซ้ำ (replication)

เมื่อประยุกต์ปัญหาให้ใช้กับตัวเลขสุ่มได้แล้ว ขั้นตอนต่อไปก็คือ การทดลองโดยใช้กระบวนการสุ่ม (random process) มากระทำในลักษณะซ้ำ ๆ กัน เพื่อหาคำตอบของปัญหาที่ต้องการศึกษา

3.2 แผนการทดลอง

การวิจัยครั้งนี้กำหนดสถานการณ์ต่าง ๆ ที่ต้องการศึกษาโดยสุ่มตัวอย่างจากประชากรที่มีการแจกแจงแบบเดียวกันทั้ง 2 ประชากร ซึ่งลักษณะการแจกแจงจะเป็นแบบเบ้ ได้แก่ การแจกแจงเอกซ์โปเนนเชียล, การแจกแจงลอคนอร์มอล และการแจกแจงไวบูลล์

สำหรับการแจกแจงแบบเอกซ์โปเนนเชียลจะศึกษาที่ค่า $\beta = 0.5, 1$ และ 2

สำหรับการแจกแจงแบบลอคนอร์มอลจะศึกษาที่ค่า $\mu = 0$, $\sigma^2 = 0.1$, 0.3 และ 0.5

สำหรับการแจกแจงแบบไวบูลล์จะศึกษาที่ค่า $\beta = 1$, $\alpha = 1$, 2 และ 4 เมื่อค่า β คือ Scale parameter และค่า α คือ Shape parameter

การแจกแจงทั้งหมดดังกล่าวจะใช้ขนาดตัวอย่างเป็น 20 และ 50 โดยมีเปอร์เซนต์ของค่าสังเกตไม่สมบูรณ์เป็น 10, 20 และ 30 ของขนาดตัวอย่าง ดังแสดงในตารางได้ดังนี้

ให้ r_1 เป็น จำนวนค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1

r_2 เป็น จำนวนค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 2

กรณีที่ *	$N_1 = N_2 = 20$		$N_1 = N_2 = 50$	
	r_1	r_2	r_1	r_2
1	2	2	5	5
2	4	2	10	5
3	4	4	10	10
4	6	4	15	10

* หมายเหตุ

กรณีที่ 1 หมายถึง เมื่อค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 เป็น 10% และในกลุ่มที่ 2 เป็น 10% ของขนาดตัวอย่าง

กรณีที่ 2 หมายถึง เมื่อค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 เป็น 20% และในกลุ่มที่ 2 เป็น 10% ของขนาดตัวอย่าง

กรณีที่ 3 หมายถึง เมื่อค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 เป็น 20% และในกลุ่มที่ 2 เป็น 20% ของขนาดตัวอย่าง

กรณีที่ 4 หมายถึง เมื่อค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 เป็น 30% และในกลุ่มที่ 2 เป็น 20% ของขนาดตัวอย่าง

3.3 ขั้นตอนในการทดลอง

ขั้นตอนในการทดลองแบ่งเป็น 2 ขั้นตอน คือ

1. การสร้างโปรแกรมย่อย (Subroutine) สำหรับการสร้างการแจกแจงของประชากร เมื่อมีค่าพารามิเตอร์ตามที่กำหนดไว้ และกำหนดค่าเปอร์เซนต์ค่าสังเกตไม่สมบูรณ์ ในทั้ง 2 ประชากรตามที่ต้องการศึกษา

2. การหาค่าอำนาจการทดสอบและค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ซึ่งรายละเอียด สำหรับแต่ละขั้นตอนเป็นดังนี้

3.3.1 การสร้างโปรแกรมย่อยสำหรับสร้างการแจกแจงของประชากรตามที่ต้องการศึกษา

การสร้างลักษณะการแจกแจงของประชากรทุกรูปแบบที่ต้องการศึกษานั้น ใช้โปรแกรมภาษาฟอร์แทรน 77 โดยใช้กับเครื่อง AMDAHL 3860 ซึ่งการสร้างลักษณะการแจกแจงแบบต่าง ๆ นั้นจะต้องใช้ตัวเลขสุ่มเป็นพื้นฐานในการสร้าง สำหรับรายละเอียดในการสร้างการแจกแจงแบบต่าง ๆ เป็นดังนี้

3.3.1.1 การแจกแจงแบบเอกซ์โปเนนเชียล (Exponential distribution)

การผลิตเลขสุ่มที่มีการแจกแจงแบบเอกซ์โปเนนเชียลนี้ จะอาศัยวิธี Inverse transformation methods จากฟังก์ชั่นความหนาแน่น และฟังก์ชั่นการแจกแจงสะสมของเอกซ์โปเนนเชียล มีดังนี้

$$f(x) = \frac{1}{\beta} e^{-x/\beta} \quad ; \quad \beta > 0, \; x \geq 0$$

$$F(x) = 1 - e^{-x/\beta} \quad ; \quad \beta > 0, \; x \geq 0$$

จากวิธี Inverse transfomation จะได้ว่า $F(x)$ จะเท่ากับ $1 - e^{-x/\beta}$ และเมื่อได้เลขสุ่มจากการแจกแจงแบบสม่ำเสมอ (R) แล้วจึงกำหนดให้

$$R = F(x) = 1 - e^{-x/\beta}$$

Inverse Transformation ในที่นี้ คือ

$$1 - R = e^{-x/\beta}$$

โดยที่ (1-R) ยังเป็นเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอ ซึ่งมีขอบเขตจำกัดคือ (0,1)

ดังนั้น จึงอาจแทน 1-R ด้วย R และสมการข้างต้นจะเปลี่ยนแปลง

$$R = e^{-x/\beta}$$

$$x = -\beta \times \ln R$$

ดังนั้น $F^{-1}(R) = -\beta \times \ln R = x$

ในการวิจัยครั้งนี้ได้เขียนโปรแกรมด้วยภาษาฟอร์ทรน 77 โดยเขียนเป็น SUBROUTINE มีชื่อว่า EXPO (BETA, XX) ดังแสดงไว้ในรายละเอียดภาคผนวก ข.

3.3.1.2 <u>การแจกแจงแบบลอคนอร์มอล</u> (Lognormal Distribution)

การแจกแจงแบบลอคนอร์มอล หาได้จากการหา Exponential ของฟังก์ชั่น SNORM (DMEAN, SIGMA) ซึ่งเป็นโปรแกรมเพื่อสร้างข้อมูลที่มีการแจกแจงแบบปกติ โดยมีค่าเฉลี่ยและค่าความแปรปรวนเป็น DMEAN และ SIGMA ตามลำดับ สำหรับรายละเอียดแสดงไว้ในภาคผนวก ข.

3.3.1.3 <u>การแจกแจงแบบไวบูลล์</u> (Weibull Distribution)

การแจกแจงแบบไวบูลล์สร้างจากฟังก์ชั่น WEIBUL (ALPHA 1, BETA 1) ซึ่งฟังก์ชั่นนี้สร้างโดยวิธี Inverse transformation ดังรายละเอียดแสดงไว้ในภาคผนวก ข. ฟังก์ชั่น WEIBUL นี้ จะกำหนดค่า ALPHA1 ; BETA1 เพื่อให้ได้ค่าเฉลี่ยและค่าความแปรปรวนตามต้องการ ผลลัพธ์คือ ค่า WEIBUL จะมีการแจกแจงแบบไวบูลล์ที่มีค่าเฉลี่ยเป็น BETA1 *(1 + 1/ALPHA1) และความแปรปรวนเป็น $(BETA1)^2 * [\Gamma(1 + 2/ALPHA1) - \Gamma^2(1 + 1/ALPHA1)]$

3.3.2 การหาค่าความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติทดสอบ

เขียนโปรแกรมคอมพิวเตอร์เพื่อหาอัตราความคลาดเคลื่อนประเภทที่ 1 เมื่อประชากรทั้ง 2 กลุ่ม มีการแจกแจงเดียวกัน โดยที่ค่าของ Shape parameter และ scale parameter มีค่าเท่ากัน ทั้ง 2 กลุ่มตัวอย่าง และหาอำนาจของการทดสอบของตัวสถิติทดสอบทั้ง 4 วิธี เมื่อค่าของ shape parameter ของ 2 กลุ่มตัวอย่าง มีค่าแตกต่างกัน ดังนี้

ก. เมื่อมีการแจกแจงแบบเอกซ์โปเนนเชียล จะทดลองหาค่าอำนาจของการทดสอบเมื่อ

ก.1) ค่าพารามิเตอร์ (ในที่นี้คือ ค่าβ) ในกลุ่มที่ 1 เป็น 0.5 และค่าพารามิเตอร์ ในกลุ่มที่ 2 เป็น 1 และ 2

ก.2) ค่าพารามิเตอร์ในกลุ่มที่ 1 เป็น 1 และค่าพารามิเตอร์ในกลุ่มที่ 2 เป็น 0.5 และ 2

ก.3) ค่าพารามิเตอร์ในกลุ่มที่ 1 เป็น 2 และค่าพารามิเตอร์ในกลุ่มที่ 2 เป็น 0.5 และ 1

ข. เมื่อมีการแจกแจงแบบลอคนอร์มอล จะทดลองหาค่าอำนาจของการทดสอบเมื่อค่าของ Scale parameter (ในที่นี้คือ ค่าμ) ของประชากรทั้ง 2 กลุ่ม มีค่าเท่ากัน คือ ศูนย์ และค่าของ Shape parameter (ในที่นี้คือ σ^2) ของประชากรทั้ง 2 กลุ่มมีค่าต่าง ๆ กัน ดังนี้คือ

ข.1) เมื่อค่า σ^2 ในกลุ่มที่ 1 เป็น 0.1 และ ค่า σ^2 ในกลุ่มที่ 2 เป็น 0.3 และ 0.5

ข.2) เมื่อค่า σ^2 ในกลุ่มที่ 1 เป็น 0.3 และค่า σ^2 ในกลุ่มที่ 2 เป็น 0.1 และ 0.5

ข.3) เมื่อค่า σ^2 ในกลุ่มที่ 1 เป็น 0.5 และ ค่า σ^2 ในกลุ่มที่ 2 เป็น 0.1 และ 0.3

ค. เมื่อมีการแจกแจงแบบไวบูลล์ จะทดลองหาค่าอำนาจของการทดสอบเมื่อค่าของ Scale parameter(ในที่นี้คือ β) ของประชากรทั้ง 2 กลุ่ม มีค่าเท่ากับ คือ 1 และ ค่าของ Shape parameter (ในที่นี้คือ α) ของประชากรทั้ง 2 กลุ่ม มีค่าต่าง ๆ กัน ดังนี้ คือ

ค.1) เมื่อค่า α ในกลุ่มที่ 1 เป็น 1 และ ค่า α ในกลุ่มที่ 2 เป็น 2 และ 4

ค.2) เมื่อค่า α ในกลุ่มที่ 2 เป็น 2 และ ค่า α ในกลุ่มที่ 2 เป็น 1 และ 4

ค.3) เมื่อค่า α ในกลุ่มที่ 1 เป็น 4 และ ค่า α ในกลุ่มที่ 2 เป็น 1 และ 2

นอกจากนั้นยังพิจารณาถึงเปอร์เซนต์ของค่าสังเกตไม่สมบูรณ์ในแต่ละกลุ่มของประชากรดังนี้ คือ

เมื่อ เปอร์เซ็นต์ของค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 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

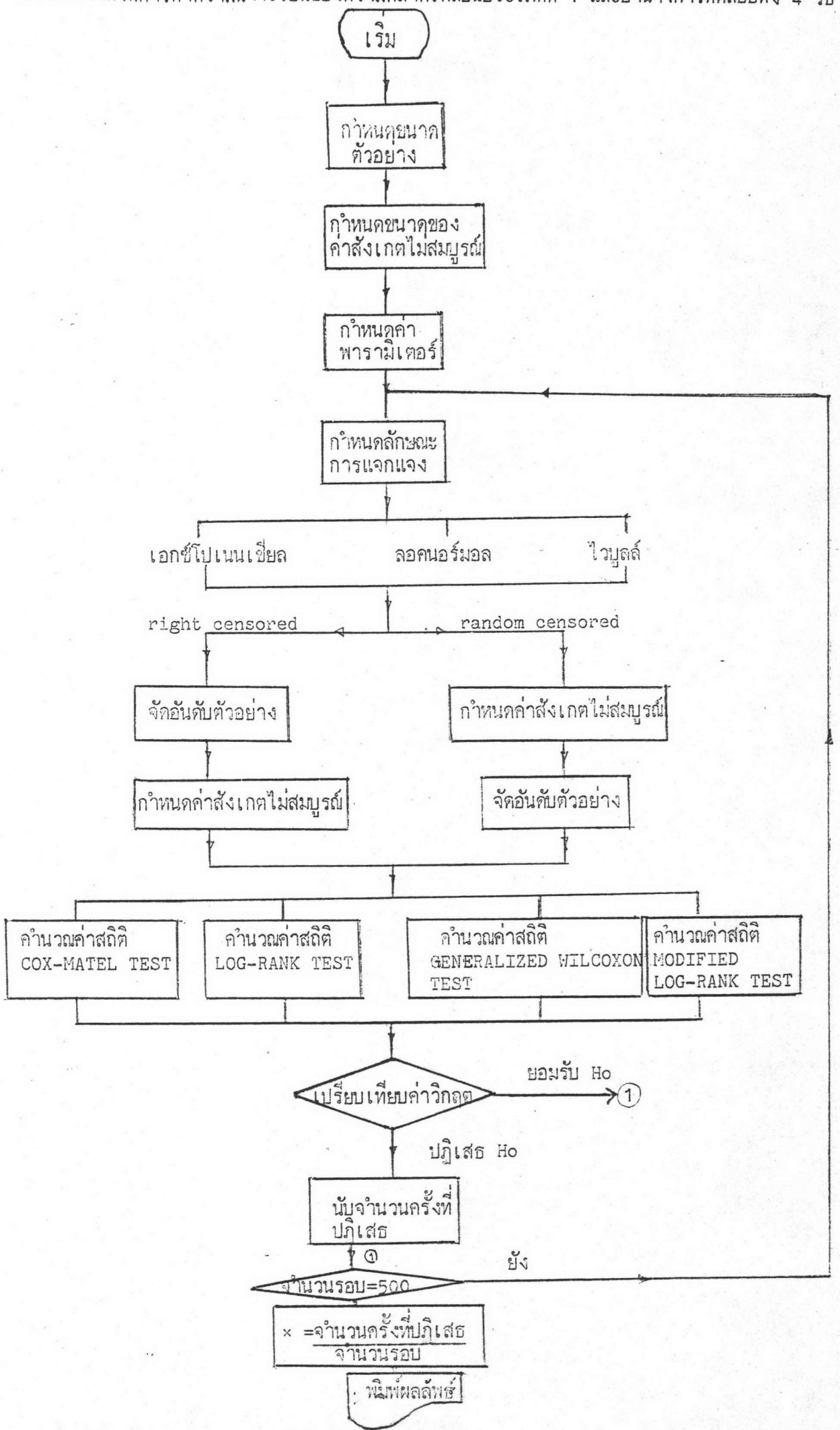
เปอร์เซนต์ของค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1 เป็น 30% และในกลุ่มที่ 2 เป็น 20% ของขนาดตัวอย่าง

ซึ่งในที่นี้พิจารณาเมื่อขนาดตัวอย่างเป็น 20 และ 50 และลักษณะของค่าสังเกตไม่สมบูรณ์เป็นแบบ Singly Censored และที่แบบผสมของ Singly failure และ multiply time censored ดังกล่าวแล้วในตอนที่ 1.1

โปรแกรมที่ใช้ในการวิจัย

โปรแกรมที่ใช้ในการวิจัยทั้งหมดเขียนด้วยภาษาฟอร์แทรน 77 โดยใช้กับเครื่อง AMDAHL3860 โปรแกรมที่ใช้คำนวณอำนาจของการทดสอบ และความคลาดเคลื่อนประเภทที่ 1 ในโปรแกรมหนึ่ง ๆ จะคำนวณเมื่อมีขนาดตัวอย่าง และมีการแจกแจงของประชากรอย่างใดอย่างหนึ่งเท่านั้น ดังแผนผังงานในการหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจของการทดสอบทั้ง 4 วิธี ดังนี้

แสดงผังงานในการหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบทั้ง 4 วิธี



รายละเอียดของแผนผังโปรแกรมในการคำนวณความน่าจะเป็นของค่าคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติที่ใช้ในการเปรียบเทียบการแจกแจงการอยู่รอด จะเสนอเป็นขั้นตอนต่อไปนี้

ขั้นที่ 1

กำหนดขนาดตัวอย่างและเปอร์เซนต์ค่าสังเกตไม่สมบูรณ์ ตามหัวข้อ 3.2

โดยที่ N_1 คือ จำนวนข้อมูลในกลุ่มที่ 1

N_2 คือ จำนวนข้อมูลในกลุ่มที่ 2

r_1 คือ จำนวนค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 1

r_2 คือ จำนวนค่าสังเกตไม่สมบูรณ์ในกลุ่มที่ 2

ขั้นที่ 2

สร้างประชากรให้มีการแจกแจงลักษณะต่าง ๆ ดังนี้

การแจกแจงแบบเอกซโปเนนเชียล โดยใช้คำสั่ง SUBROUTINE EXPO (BETA, xx)

การแจกแจงแบบลอคนอร์มอล โดยใช้คำสั่ง EXP (SNORM(DMEAN, SEGMA))

การแจกแจงแบบไวบูลล์ โดยใช้คำสั่ง FUNCTION WEIBUL (ALPHA1, BETA1)

โดยที่กำหนดค่าพารามิเตอร์ของแต่ละการแจกแจง ตามหัวข้อ 3.2 ซึ่งการกำหนดขนาดตัวอย่าง เปอร์เซนต์ค่าสังเกตไม่สมบูรณ์ และการคำนวณค่าสถิติทดสอบ 4 ตัว ดังกล่าวใน 2.3 จะทำโดยการเลือกประชากรครั้งละ 1 ประชากร และจะหยุดการทดลองเมื่อทำการทดลองครบทุกกรณีที่กำหนดของแต่ละประชากร

ขั้นที่ 3

กรณีที่มีค่าสังเกตไม่สมบูรณ์แบบสุ่ม

1) กำหนดตำแหน่งค่าสังเกตไม่สมบูรณ์จากข้อมูลในขั้นที่ 2 ให้ได้จำนวนค่าสังเกตไม่สมบูรณ์ เท่ากับจำนวนที่กำหนดในขั้นที่ 1

2) จัดเรียงข้อมูลทั้ง 2 กลุ่มตัวอย่างรวมกัน โดยเรียงข้อมูลจากน้อยไปมาก โดยใช้คำสั่ง SUBROUTINE SORT (IP, DATA)

กรณีที่มีค่าสังเกตไม่สมบูรณ์ทางขวา

1) จัดเรียงข้อมูลในแต่ละกลุ่ม ซึ่งได้จากขั้นที่ 2 โดยเรียงข้อมูลจากน้อยไปมาก โดยใช้คำสั่ง SUBROUTINE SORT (IP, DATA)

2) กำหนดตำแหน่งค่าสังเกตไม่สมบูรณ์จากข้อมูลใน 1) ให้ได้จำนวนค่าสังเกตไม่สมบูรณ์เท่ากับจำนวนที่กำหนดในขั้นที่ 1

3) จัดเรียงข้อมูลทั้ง 2 กลุ่มตัวอย่างรวมกัน โดยเรียงข้อมูลจากน้อยไปมาก โดยใช้คำสั่ง SUBROUTINE SORT (IP, DATA)

ขั้นที่ 4

คำนวณค่าสถิติทดสอบทั้ง 4 ตัว ตามลำดับ

ขั้นที่ 5

ทำการเปรียบเทียบตัวสถิติทดสอบที่คำนวณได้ในขั้นที่ 4 กับค่าวิกฤตที่กำหนดขึ้น ถ้าการเปรียบเทียบปรากฏผลว่า ปฏิเสธสมมติฐาน Ho ให้นับจำนวนครั้งที่ปฏิเสธสมมติฐานเพิ่มขึ้น 1 ครั้ง ของตัวสถิติแต่ละตัว และทำการทดลองซ้ำตั้งแต่ขั้นที่ 3 ถึงขั้นที่ 5 500 ครั้ง ในแต่ละสถานการณ์ เพื่อหาผลลัพธ์ของค่าความน่าจะเป็นของค่าความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติทดสอบทั้ง 4 ตัว