บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง ซึ่งจำลองข้อมูลขึ้นด้วยการทำงานของเครื่องคอมพิวเตอร์โดยใช้เทคนิคการจำลองมอนติคาร์โล (Monte Carlo Simulation Technique) เพื่อเปรียบเทียบประสิทธิภาพของแผนภูมิควบคุมภายใต้สถานการณ์ของกระบวนการผลิตในค่าเฉลี่ย ณ ระดับการเปลี่ยนแปลงของค่าเฉลี่ยและขนาดตัวอย่างต่างๆ โดยจะทำการเปรียบเทียบความยาววิ่งโดยเฉลี่ย(ARL) เพื่อเปรียบเทียบว่าแผนภูมิควบคุมแบบใดให้ความยาววิ่งโดยเฉลี่ยต่ำที่สุด และจะนำเสนอความน่าจะเป็นที่ค่าเฉลี่ยตัวอย่าง($\bar{x}_j$)จะออกนอกขอบเขตควบคุม เมื่อกระบวนการเกิดการเปลี่ยนแปลงในค่าเฉลี่ย ของแต่ละแผนภูมิควบคุมเป็นข้อมูลเพิ่มเติมด้วย สำหรับในบทนี้จะกล่าวถึงรายละเอียดตามหัวข้อต่อไปนี้

3.1 การจำลองโดยใช้เทคนิคมอนติคาร์โล

3.2 ขั้นตอนในการทดลอง

3.3 การผลิตตัวแปรสุ่มให้มีการแจกแจงตามที่กำหนด

3.1 การจำลองโดยใช้เทคนิคมอนติคาร์โล

เทคนิคในการจำลองตัวแบบทางคณิตศาสตร์มีอยู่หลายวิธี เทคนิคมอนติคาร์โลเป็นวิธีการจำลองตัวแบบทางคณิตศาสตร์วิธีหนึ่งที่นิยมใช้กันอย่างมาก ซึ่งหลักการของเทคนิคมอนติคาร์โลนั้นเป็นการจำลองเลขสุ่ม (Random Number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา โดยขั้นตอนของเทคนิคมอนติคาร์โลที่ใช้กันอยู่ในปัจจุบันแบ่งได้เป็น 3 ขั้นตอน ดังนี้

1. <u>การสร้างตัวเลขสุ่ม</u> (Random Number)

สิ่งสำคัญในเทคนิคมอนติคาร์โล คือ การสร้างตัวเลขสุ่ม เพราะต้องนำมาใช้ในการหาคำตอบของปัญหา วิธีการสร้างตัวเลขสุ่มมีหลายวิธี แต่วิธีที่ดีนั้นลักษณะของตัวเลขสุ่มที่สร้างขึ้นมาจะต้องอิสระกันและมีการแจกแจงสม่ำเสมอ (Uniform Distribution) ในช่วง (0,1) ซึ่งควรมีคุณสมบัติของตัวแปรสุ่มที่ดี ดังนี้

(1). ตัวเลขที่ได้มีการกระจายของความน่าจะเป็นแบบยูนิฟอร์มและเป็นอิสระซึ่งกันและกัน

(2). อนุกรมของตัวเลขที่ได้สามารถสร้างซ้ำเดิมได้ (Reproducible)

(3). อนุกรมของตัวเลขไม่ซ้ำเดิมในช่วงที่ต้องการใช้ตัวเลขสุ่ม หมายความว่า ขนาดของความยาวของอนุกรมตัวเลขต้องยาวพอสำหรับการใช้งาน

(4). ใช้เวลาสั้นๆ ในการสร้างเลขสุ่ม

(5). ใช้หน่วยความจำของคอมพิวเตอร์น้อย

2 <u>การประยุกต์ตัวเลขสุ่มให้สามารถใช้กับปัญหาที่ต้องการศึกษา</u>

ในขั้นตอนนี้จะขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา ว่าเราจะนำเลขสุ่มนั้นไปใช้โดยตรงเลยหรือไม่ เพราะบางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรงแต่อาจจะมีบางขั้นตอนที่ต้องการใช้เลขสุ่ม

3. <u>การทดลองกระทำซ้ำ</u> (Replication)

เมื่อเราสามารถประยุกต์ตัวเลขสุ่มให้มีลักษณะที่ต้องการได้แล้ว ขั้นตอนต่อไป เราจะทำการทดลองโดยใช้กระบวนการของการสุ่ม (Random Process) มากระทำในลักษณะที่ซ้ำๆกัน หลายๆ ครั้ง จนกว่าจะได้คำตอบของปัญหาที่เราต้องการศึกษา

3.2 ขั้นตอนในการทดลอง

การวิจัยครั้งนี้จะทำการศึกษาแผนภูมิควบคุม ทั้ง 4 แบบ ของประชากรที่มีการแจกแจงแบบปกติ ซึ่งกำหนดขั้นตอนในการทดลองดังนี้

1. กำหนดขนาดตัวอย่าง และพารามิเตอร์ต่างๆ ตามสถานการณ์ที่กำหนดขึ้น

2. จำลองข้อมูลให้มีลักษณะการแจกแจงแบบปกติ ภายใต้ตัวแบบที่ใช้ในการจำลองข้อมูลเป็นอนุกรมเวลาแบบค่าเฉลี่ยคงที่เฉพาะช่วงเวลา โดยมีตัวแบบ ดังนี้

$$X_t = \mu_0 + \gamma I_t + \varepsilon_t \quad , \quad t = 1,2,3,\ldots \tag{3.1}$$

$$I_t = \begin{cases} 0 & , t \le \ell \\ 1 & , t > \ell \end{cases}$$

การจำลองข้อมูลจะกำหนดระดับการเปลี่ยนแปลงของค่าเฉลี่ยและขนาดตัวอย่างตามขอบเขตของการวิจัย

ตัวอย่างข้อมูลจากการจำลองภายใต้ตัวแบบอนุกรมเวลาแบบค่าเฉลี่ยคงที่ เฉพาะช่วงเวลา (X_t) ดังตารางที่ 3.1 และรูปที่ 3.1

ตารางที่ 3.1 ข้อมูลจากการจำลองภายใต้ตัวแบบอนุกรมเวลาแบบค่าเฉลี่ยคงที่ เฉพาะช่วงเวลา (Xt) เมื่อประชากรมีการแจกแจงแบบปกติ ค่าเฉลี่ยเท่ากับ 100 ความแปรปรวนเท่ากับ 100 กำหนดจำนวนคาบเวลา(Run-In Period) เท่ากับ 15 และระดับการเปลี่ยนแปลงของค่าเฉลี่ย (GAMMA) เท่ากับ 5 10 15 และ 25

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
GAMMA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xt	106.607	99.154	109.199	107.087	102.274	102.181	96.974	88.009	104.072	121.568	101.314	109.189	101.486	94.203	98.228

t	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
GAMMA	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Xt	118.429	96.981	98.601	107.172	103.405	107.984	113.326	97.258	116.134	105.778	102.971	104.834	108.138	124.450	93.601

ตารางที่ 3.1 (ต่อ)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
GAMMA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xt	106.607	99.154	109.199	107.087	102.274	102.181	96.974	88.009	104.072	121.568	101.314	109.189	101.486	94.203	98.228

t	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
GAMMA	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Xt	110.168	99.663	112.540	117.255	104.238	105.506	109.227	108.960	112.434	102.178	110.907	119.548	111.839	114.420	112.653

ตารางที่ 3.1 (ต่อ)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
GAMMA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xt	106.607	99.154	109.199	107.087	102.274	102.181	96.974	88.009	104.072	121.568	101.314	109.189	101.486	94.203	98.228

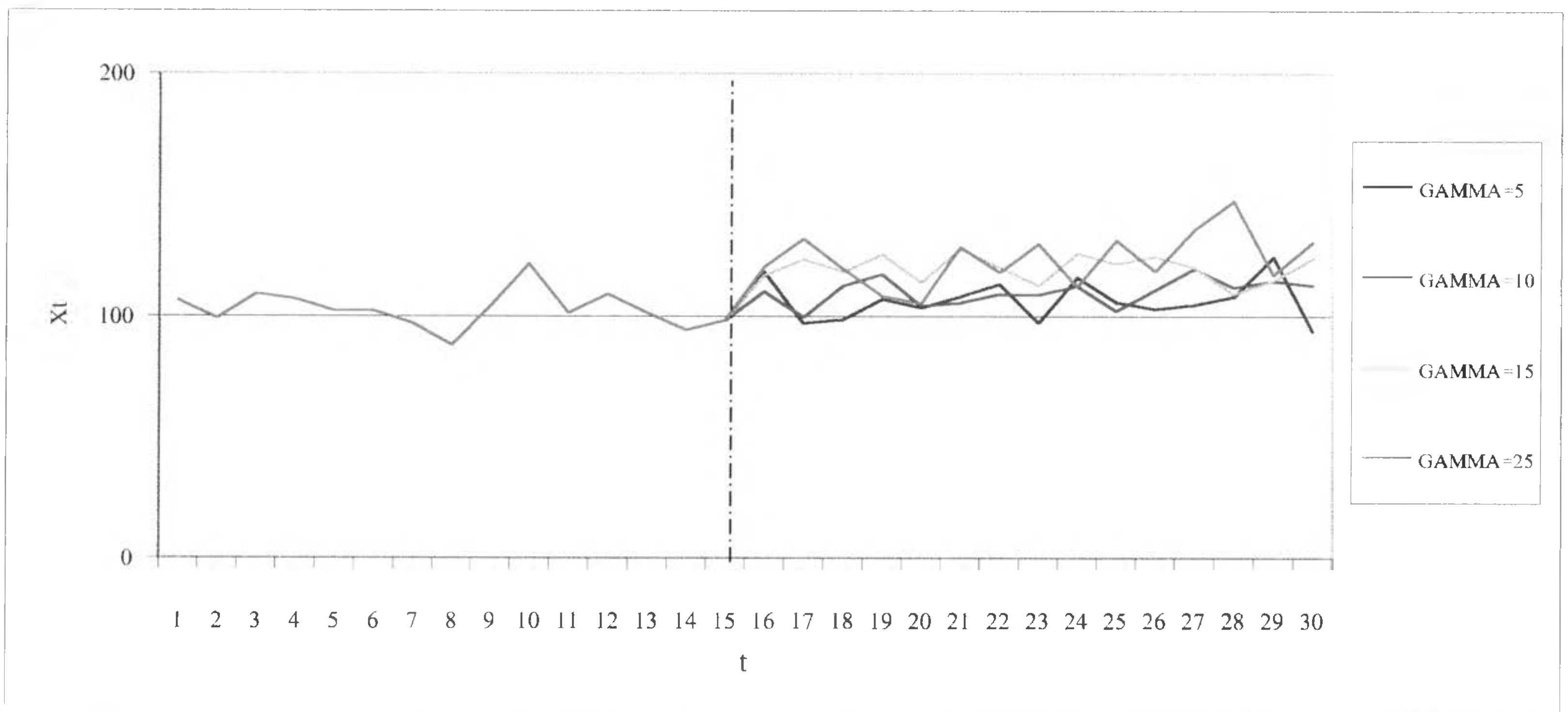
t	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
GAMMA	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
Xt	116.715	123.412	118.474	125.647	114.051	127.865	120.240	112.770	126.004	121.711	124.877	120.030	109.202	114.640	123.902

ตารางที่ 3.1 (ต่อ)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
GAMMA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Xt	106.607	99.154	109.199	107.087	102.274	102.181	96.974	88.009	104.072	121.568	101.314	109.189	101.486	94.203	98.228

t	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
GAMMA	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
Xt	120.536	132.013	119.585	108.278	105.224	128.528	118.170	129.910	112.156	131.343	118.655	135.741	147.693	117.141	130.555

รูปที่ 3.1 กราฟแสดงข้อมูลจากการจำลองภายใต้ตัวแบบอนุกรมเวลาแบบค่าเฉลี่ยคงที่ เฉพาะช่วงเวลา(X_t) เมื่อประชากรมีการแจกแจงแบบปกติ ค่าเฉลี่ยเท่ากับ 100 ความแปรปรวนเท่ากับ 100 กำหนดจำนวนคาบเวลา(Run-In Period) เท่ากับ 15 และระดับการเปลี่ยนแปลงของค่าเฉลี่ย (GAMMA) เท่ากับ 5 10 15 และ 25



3. คำนวณหาขอบเขตควบคุมจากพารามิเตอร์ที่กำหนด สมมติว่าทราบค่าพารามิเตอร์ของแผนภูมิควบคุมทั้ง 4 แบบ เมื่อยังไม่มีการเปลี่ยนแปลงระดับของค่าเฉลี่ย โดยการกำหนดขอบเขตของแต่ละแผนภูมิควบคุม มีดังนี้

(1). แผนภูมิควบคุมค่าเฉลี่ย ($\bar{x}$ Control Chart)

ขอบเขตควบคุมสำหรับแผนภูมิควบคุมค่าเฉลี่ย คือ

$$UCL = \mu_0 + 3\frac{\sigma}{\sqrt{n}}$$
$$LCL = \mu_0 - 3\frac{\sigma}{\sqrt{n}}$$

(2). แผนภูมิควบคุมค่าเฉลี่ยเคลื่อนที่ปรับน้ำหนักแบบเอกซโพเนนเชียล (Exponential Weighted Moving Average Control Chart)

ขอบเขตควบคุมสำหรับค่าเฉลี่ยเคลื่อนที่ปรับน้ำหนักแบบเอกซโพเนนเชียล คือ

$$UCL = \mu_0 + 3\sigma\sqrt{\frac{\lambda}{n(2-\lambda)}}$$
$$LCL = \mu_0 - 3\sigma\sqrt{\frac{\lambda}{n(2-\lambda)}}$$

(3). แผนภูมิควบคุมรวมค่าเฉลี่ยและผลรวมสะสม (Combined $\bar{X}$ – Cumulative Sum control chart)

ขอบเขตของค่าเฉลี่ยและความยาวของช่วงควบคุม สำหรับแผนภูมิควบคุมรวมค่าเฉลี่ยและผลรวมสะสม คือ

$$SCL = \frac{3}{\sqrt{n}}$$
$$H = \frac{5\sigma}{\sqrt{n}}$$

(4). แผนภูมิควบคุมสังเคราะห์ (Synthetic Control Chart)

ขอบเขตควบคุมสำหรับแผนภูมิควบคุมสังเคราะห์ ประกอบด้วย 2 ขอบเขตควบคุม คือ

1. แผนภูมิควบคุม $\bar{X}/S$

$$UCL = \mu_0 + k\frac{\sigma}{\sqrt{n}}$$

$$LCL = \mu_0 - k\frac{\sigma}{\sqrt{n}}$$

2. แผนภูมิควบคุม CRL/S มีขอบเขตควบคุมล่าง LCL_{CRL}

โดยการหาค่าพารามิเตอร์ k ของแผนภูมิควบคุม $\bar{X}/S$ ในข้อ 1 และ LCL_{CRL} ของแผนภูมิควบคุม CRL/S ในข้อ 2 สามารถหาได้จากหัวข้อ 2.2

4. คำนวณหาจำนวนความยาววิ่งโดยเฉลี่ย (ARL) โดยมีขั้นตอนดังนี้

(1). จำลองตัวเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอในช่วง (0,1)

(2). จำลองค่าความคลาดเคลื่อนสุ่ม ณ เวลา t (ε_t) กำหนดให้มีการแจกแจงแบบปกติมีค่าเฉลี่ย(μ) เท่ากับ 0 และค่าความแปรปรวน (σ^2) เท่ากับ 100

(3). กำหนดพารามิเตอร์ค่าเฉลี่ยของประชากร (μ_0) เท่ากับ 100

(4). กำหนดจำนวนคาบเวลาที่เริ่มต้น ℓ เท่ากับ 15

(5). กำหนดพารามิเตอร์ δ และ n

(6). จำลองข้อมูลให้มีลักษณะ ภายใต้ตัวแบบ (3.1)

(7). คำนวณค่าสถิติทดสอบของแต่ละแผนภูมิควบคุม และเริ่มนับจำนวนความยาววิ่งตั้งแต่คาบเวลา $t = \ell+1$ ซึ่งค่าเฉลี่ยของประชากรจะเกิดการเปลี่ยนแปลงจาก μ_0 เป็น $\mu_1 = \mu_0 + \delta\sigma$ โดยทำการเปรียบเทียบกับขอบเขตควบคุม จนกว่าค่าสถิติจะมีค่ามากกว่าขอบเขตควบคุม

(8). ทำตามขั้นตอน (6) – (7) ซ้ำ 1,000 รอบ ($n^* = 1{,}000$) แล้วหาจำนวนความยาววิ่งโดยเฉลี่ย (ARL) ซึ่งคำนวณจากผลรวมของจำนวนความยาววิ่งทั้งหมดหารด้วยจำนวนรอบของการทดลองซ้ำ

เมื่อได้จำนวนความยาววิ่งโดยเฉลี่ย (ARL) ของแต่ละสถานการณ์แล้ว จะทำการเปรียบเทียบกันในเชิงสถิติ โดยแผนภูมิควบคุมใดให้ค่าที่ต่ำกว่าแผนภูมิควบคุมอื่น จะเป็นแผนภูมิควบคุมที่มีประสิทธิภาพมากที่สุด

5. คำนวณหาค่าความน่าจะเป็นที่ค่าเฉลี่ยตัวอย่างจะออกนอกขอบเขตควบคุม เมื่อกระบวนการเกิดการเปลี่ยนแปลงในค่าเฉลี่ย โดยมีขั้นตอนดังนี้

(1). จำลองตัวเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอในช่วง (0,1)

(2). จำลองค่าความคลาดเคลื่อนสุ่ม ณ เวลา t (ε_t) กำหนดให้มีการแจกแจงแบบปกติมีค่าเฉลี่ย(μ) เท่ากับ 0 และค่าความแปรปรวน (σ^2) เท่ากับ 100

(3). กำหนดพารามิเตอร์ค่าเฉลี่ยของประชากร (μ_0) เท่ากับ 100

(4). กำหนดจำนวนคาบเวลาที่เริ่มต้น ℓ เท่ากับ 15

(5). กำหนดพารามิเตอร์ n และ δ

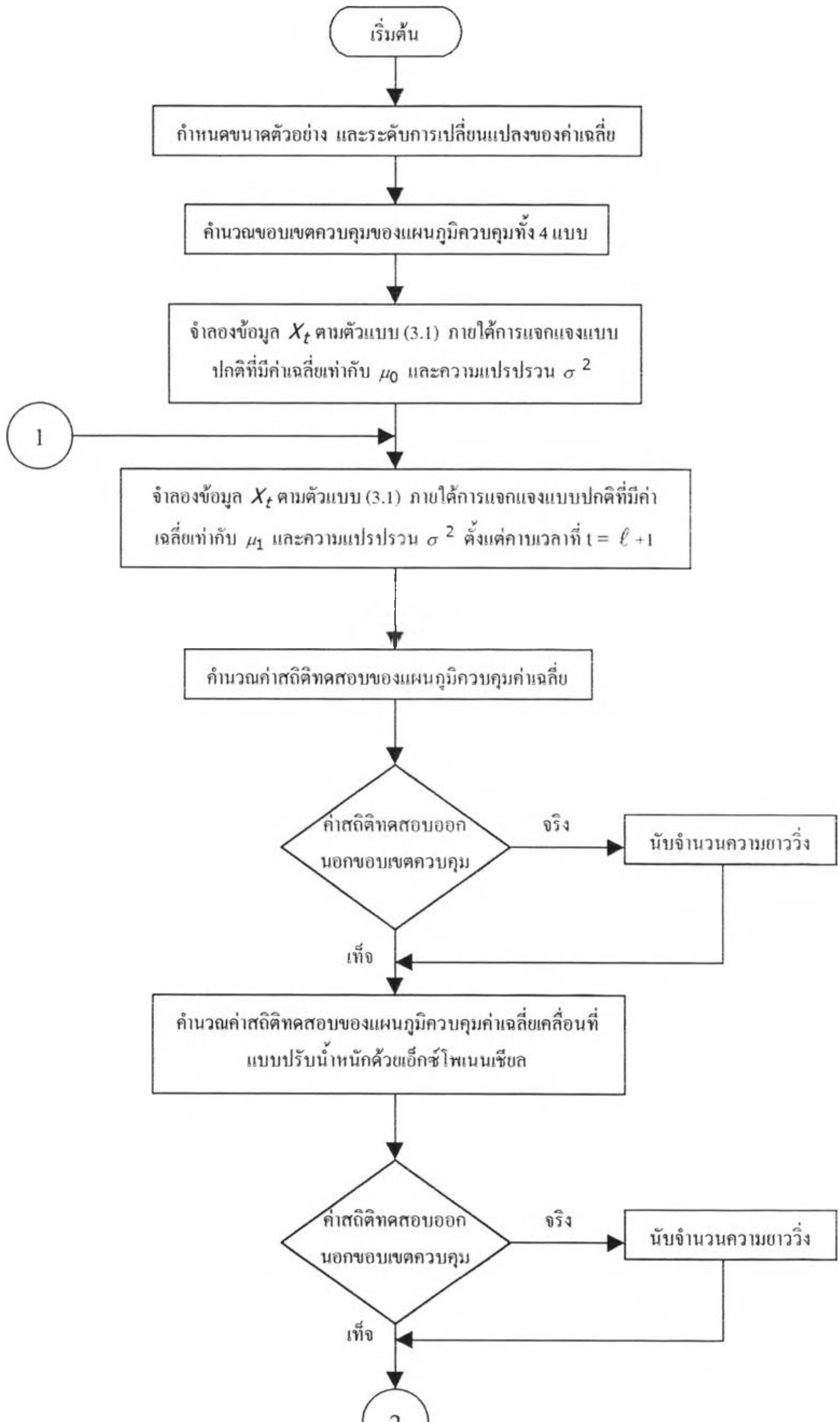
(6). จำลองข้อมูลให้มีลักษณะ ภายใต้ตัวแบบ (3.1)

(7). คำนวณค่าสถิติทดสอบของแต่ละแผนภูมิควบคุม และเริ่มนับจำนวนครั้งของค่าสถิติทดสอบตัวอย่างที่ออกนอกขอบเขตควบคุม ตั้งแต่คาบเวลา $t = \ell+1$ ซึ่งค่าเฉลี่ยของประชากรจะเกิดการเปลี่ยนแปลงจาก μ_0 เป็น $\mu_1 = \mu_0 + \delta\sigma$ โดยทำการเปรียบเทียบกับขอบเขตควบคุม

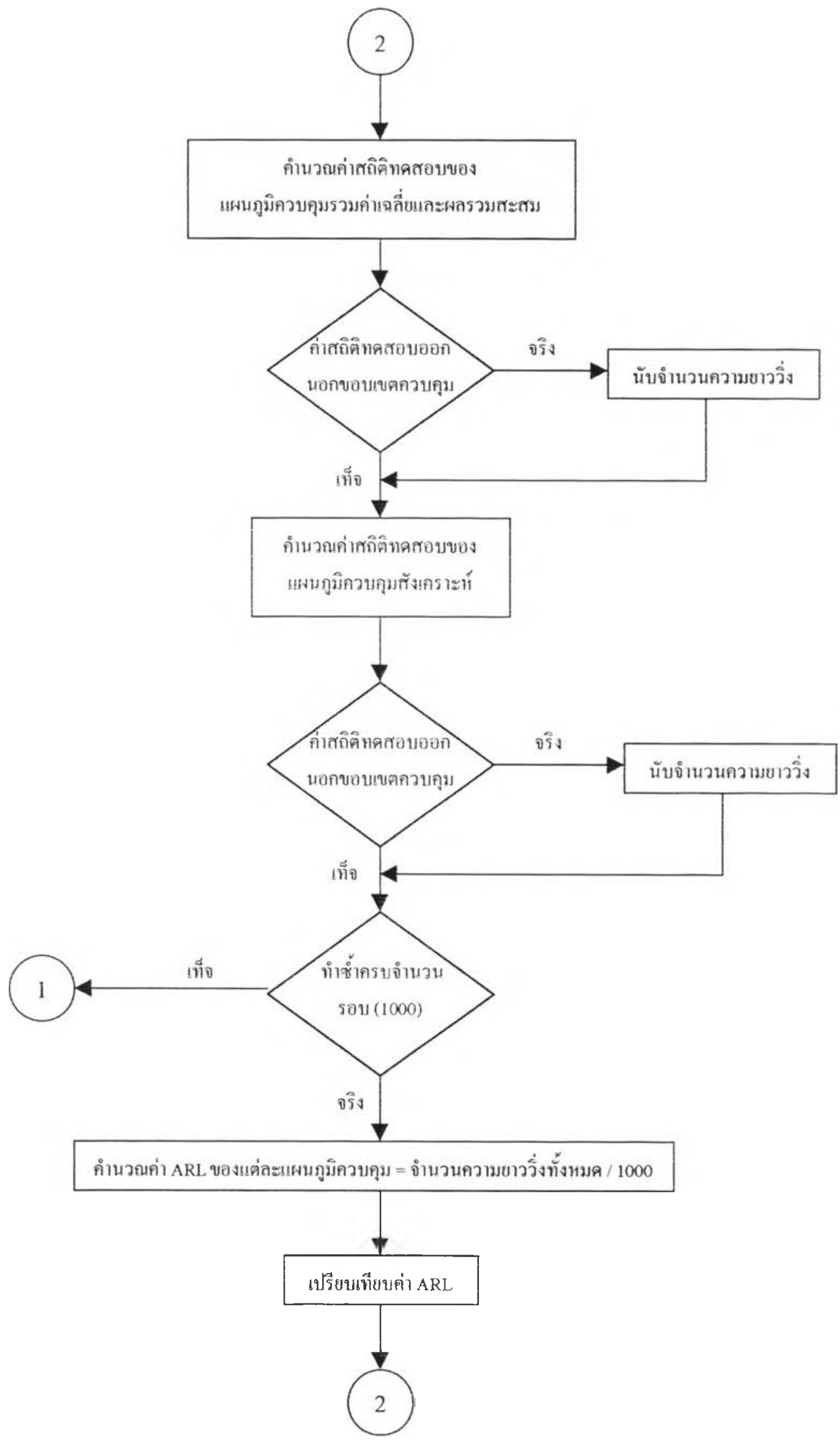
(8). ทำตามขั้นตอน (6) – (7) ซ้ำ 1,000 รอบ ($n^* = 1{,}000$) แล้วหาค่าความน่าจะเป็นที่ค่าเฉลี่ยตัวอย่างจะออกนอกขอบเขตควบคุม เมื่อกระบวนการเกิดการเปลี่ยนแปลงในค่าเฉลี่ยซึ่งคำนวณจากจำนวนครั้งของค่าเฉลี่ยตัวอย่างที่ออกนอกขอบเขตควบคุมทั้งหมดหารด้วยจำนวนรอบของการทดลองซ้ำ

6. ทำตามขั้นตอน 1 - 5 จนครบทุกสถานการณ์

แสดงขั้นตอนการดำเนินงานข้างต้นของการคำนวณหาจำนวนความยาววิ่งโดยเฉลี่ย (ARL) และการคำนวณหาความน่าจะเป็นที่ค่าเฉลี่ยตัวอย่างจะออกนอกขอบเขตควบคุม เมื่อกระบวนการเกิดการเปลี่ยนแปลงในค่าเฉลี่ย โดยแบ่งออกเป็น 2 ส่วน ซึ่งแสดงได้ดังรูปที่ 3.2 และ 3.3 ตามลำดับ

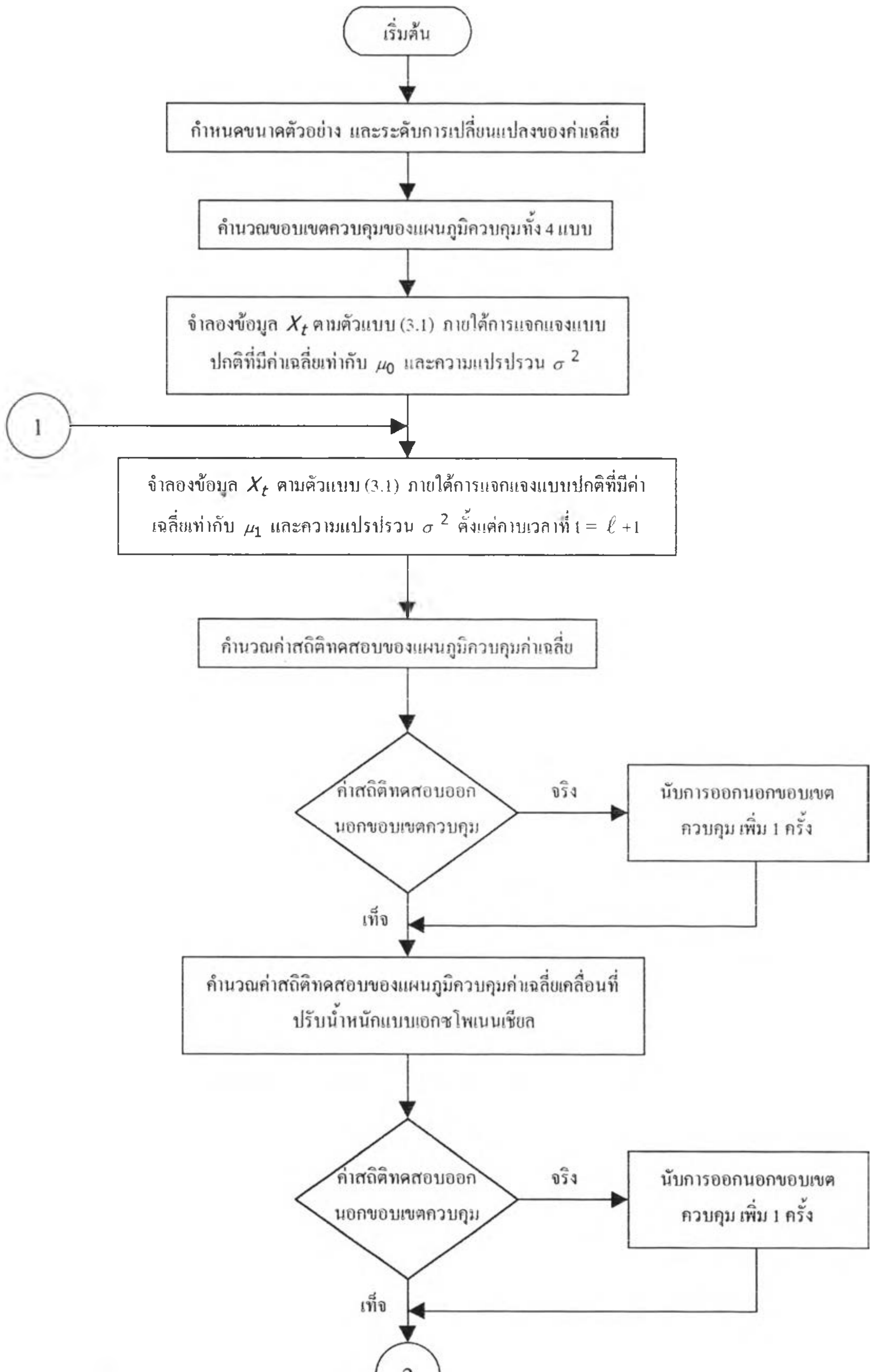


รูปที่ 3.2 แผนผังการหาจำนวนคาบเวลาโดยเฉลี่ยของแผนภูมิควบคุมทั้ง 4 แบบ

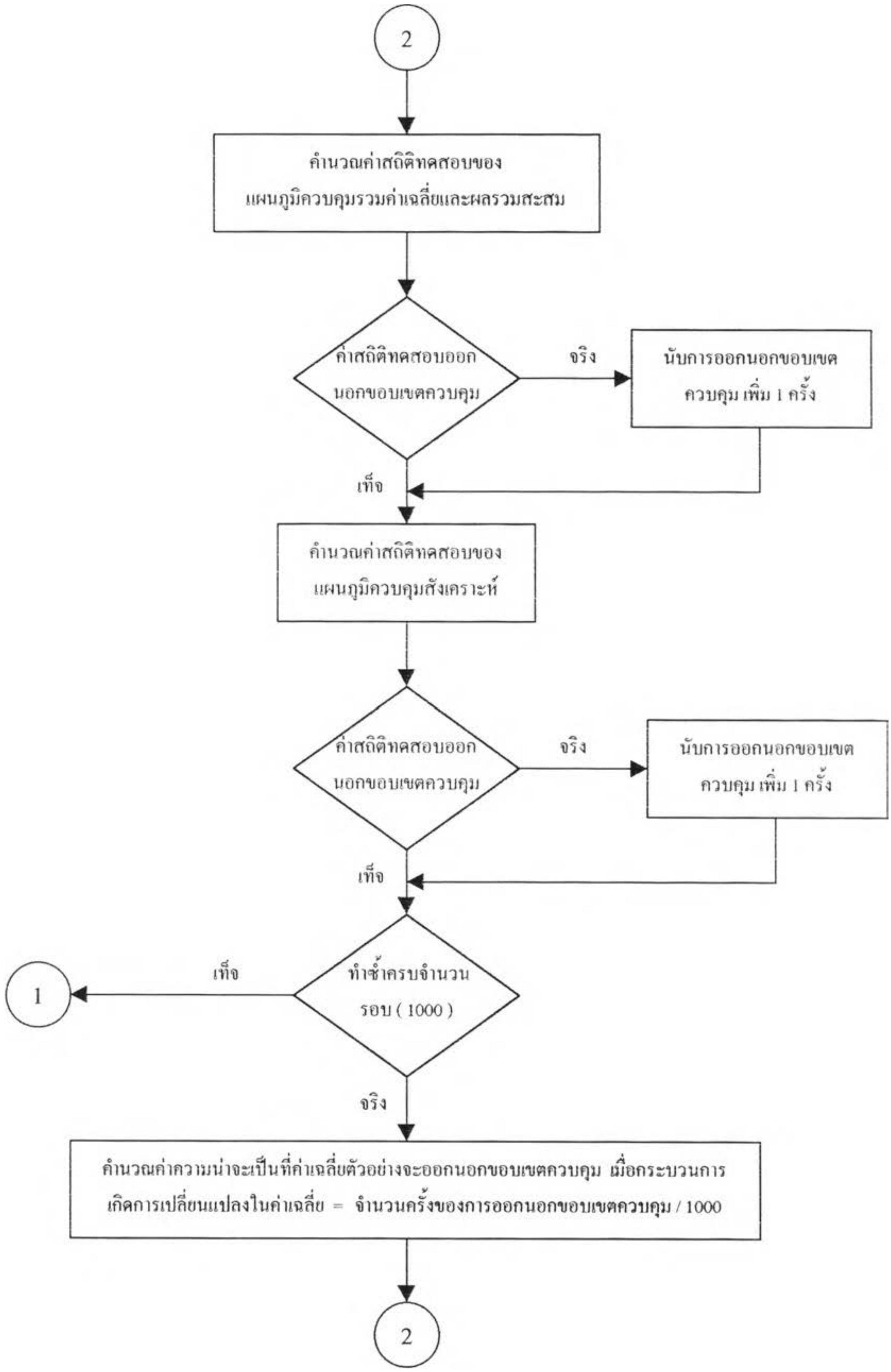

2
คำนวณค่าสถิติทดสอบของ
แผนภูมิควบคุมรวมค่าเฉลี่ยและผลรวมสะสม
ค่าสถิติทดสอบออก
นอกขอบเขตควบคุม
จริง
นับจำนวนความยาววิ่ง
เท็จ
คำนวณค่าสถิติทดสอบของ
แผนภูมิควบคุมสังเคราะห์
ค่าสถิติทดสอบออก
นอกขอบเขตควบคุม
จริง
นับจำนวนความยาววิ่ง
เท็จ
1
เท็จ
ทำซ้ำครบจำนวน
รอบ (1000)
จริง
คำนวณค่า ARL ของแต่ละแผนภูมิควบคุม = จำนวนความยาววิ่งทั้งหมด / 1000
เปรียบเทียบค่า ARL
2

รูปที่ 3.2 (ต่อ)

รูปที่ 3.2 (ต่อ)



รูปที่ 3.3 แผนผังการหาความน่าจะเป็นที่ค่าเฉลี่ยตัวอย่างจะออกนอกขอบเขตควบคุม เมื่อเกิดการเปลี่ยนแปลงในค่าเฉลี่ย ของแผนภูมิควบคุมทั้ง 4 แบบ


2
คำนวณค่าสถิติทดสอบของ
แผนภูมิควบคุมรวมค่าเฉลี่ยและผลรวมสะสม
ค่าสถิติทดสอบออก
นอกขอบเขตควบคุม
จริง
นับการออกนอกขอบเขต
ควบคุม เพิ่ม 1 ครั้ง
เท็จ
คำนวณค่าสถิติทดสอบของ
แผนภูมิควบคุมสังเคราะห์
ค่าสถิติทดสอบออก
นอกขอบเขตควบคุม
จริง
นับการออกนอกขอบเขต
ควบคุม เพิ่ม 1 ครั้ง
เท็จ
ทำซ้ำครบจำนวน
รอบ (1000)
เท็จ
1
จริง
คำนวณค่าความน่าจะเป็นที่ค่าเฉลี่ยตัวอย่างจะออกนอกขอบเขตควบคุม เมื่อกระบวนการ
เกิดการเปลี่ยนแปลงในค่าเฉลี่ย = จำนวนครั้งของการออกนอกขอบเขตควบคุม / 1000
2

รูปที่ 3.3 (ต่อ)

รูปที่ 3.3 (ต่อ)

3.3 การผลิตตัวแปรสุ่มให้มีการแจกแจงตามที่กำหนด

ในการสร้างตัวแปรสุ่มให้มีการแจกแจงตามคุณสมบัติที่ต้องการ สามารถทำได้จากการจำลองข้อมูลด้วยการใช้เทคนิคมอนติคาร์โล โดยการเขียนโปรแกรมคอมพิวเตอร์ภาษาฟอร์แทรนเพาเวอร์สเตชัน (Fortran Power Station)

การสร้างลักษณะการแจกแจงของค่าความคลาดเคลื่อนสุ่ม (ε_t) ตามที่กำหนดไว้ในตัวแบบ ให้มีการแจกแจงแบบปกติ จะต้องใช้ตัวเลขสุ่ม (Random Number) ซึ่งมีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) เป็นพื้นฐานในการสร้าง โดยการเขียนโปรแกรมคอมพิวเตอร์ภาษา ฟอร์แทรน (Fortran) การสร้างการแจกแจงต่างๆดังที่กล่าวมาแล้วมีวิธีการสร้างดังนี้

1. <u>การผลิตเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอในช่วง (0,1)</u>

ชุดตัวเลขที่ผลิตขึ้น ($r_1, r_2, ...$) ต้องมีคุณสมบัติทางสถิติที่สำคัญ 2 ประการ คือ ความเป็นสม่ำเสมอ (Uniform) และความเป็นอิสระ (Independent) ตัวเลขสุ่ม r_i แต่ละตัวจะถูกเลือกอย่างเป็นอิสระหรือสุ่มจากเลขสุ่ม R ที่มีการแจกแจงแบบสม่ำเสมอในช่วง (0 ,1)

วิธีการผลิตเลขสุ่มแบบ Linear Congruential Method เป็นวิธีการผลิตเลขสุ่มที่จะผลิตชุดตัวเลขสุ่มจำนวนเต็ม $X_1, X_2, ...$ มีค่าระหว่าง 0 ถึง M-1 จากสมการตัวผลิต

$$X_i = (aX_{i-1} + c) \bmod M \quad , \; i = 1,2,3,... \tag{3.2}$$

โดยที่ a คือ ค่าคงที่ใดๆ

c คือ ค่าส่วนเพิ่ม (Increment)

X_0 คือ ตัวเลขนำหรือค่าเริ่มต้นของการผลิตเลขสุ่ม

M คือ Modulus

mod หมายความว่า เศษที่เกิดจากการหาร $(aX_{i-1} + c)$ ด้วย M จะเป็นเลขสุ่ม X_i และเป็นเลขสุ่มคล้ายที่จะใช้สุ่มเลขต่อไป

ตัวเลขสุ่มจำนวนเต็ม $X_1, X_2, ...$จากสมการข้างต้นจะเป็นเลขสุ่มที่มีการแจกแจงสม่ำเสมอแบบไม่ต่อเนื่องในช่วง (0 , $M-1$) เพราะฉะนั้นตัวเลขสุ่ม $r_1, r_2, ...$ ที่มีการแจกแจงแบบสม่ำเสมอในช่วง (0 ,1) สามารถผลิตได้จากสมการ

$$r_i = X_i / M \quad , \quad i = 1,2,3,...$$

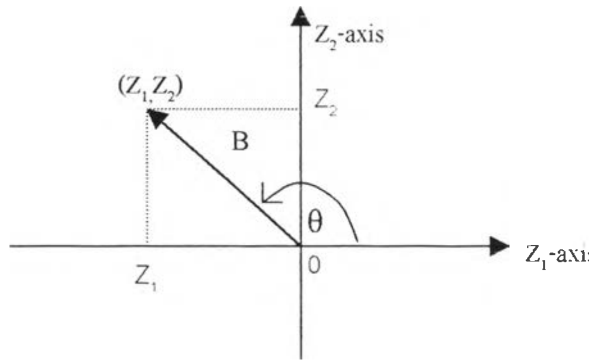
ถ้ากำหนดค่า $c \neq 0$ เรียกตัวผลิตเลขสุ่มนั้นว่า Mix Congruential Method แต่ถ้ากำหนด $c = 0$ เรียกตัวผลิตเลขสุ่มนั้นว่า Multiplicative Congruential Method การกำหนดค่า C, a , M และ X_0 มีความสำคัญมาก เนื่องจากมีผลโดยตรงต่อคุณสมบัติทางสถิติและความยาวของชุดตัวเลขสุ่ม จากสมการ $r_i = X_i / M$ จะได้ว่า r_i มีค่าอยู่ในเซตของ $\{0, 1/M, 2/M, ..., (m-1)/M\}$ ทั้งนี้เพราะค่าของ X_i เป็นจำนวนเต็มที่อยู่ในเซตของ $\{0, 1, 2, ..., M-1\}$ เพราะฉะนั้นค่า r_i จึงมีค่าไม่ต่อเนื่อง แทนที่จะเป็นค่าที่ต่อเนื่องที่มีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) อย่างไรก็ตามจะประมาณความต่อเนื่องได้ โดยการกำหนดให้ M มีค่าใหญ่มากๆ จะมีผลทำให้ช่องว่าง r_i ; $i = 1,2,3,...$ มีค่าเล็กลง ทำให้ได้ค่า r_i ที่มีความต่อเนื่องโดยประมาณลักษณะการกระทำดังกล่าวเป็นการสร้างความหนาแน่น (Density) ในกลุ่มตัวเลขสุ่มให้มีความหนาแน่นสูงในช่วง (0,1) และเพื่อหลีกเลี่ยงชุดตัวเลขสุ่มซ้ำในการใช้งานครั้งหนึ่ง ๆ ตัวผลิตควรมีความยาวของชุดตัวเลขสุ่มมากที่สุดเท่าที่เป็นไปได้ จากการทดสอบมาแล้วเป็นจำนวนมาก วิธีการผลิตเลขสุ่มที่มีคุณสมบัติต่างๆ ดังที่กล่าวไว้ข้างต้น ก็คือวิธี Multiplicative Congruential ที่กำหนด $a = 7^5 = 16807$ การกำหนดค่า M ให้มีขนาดใหญ่มากๆ และเป็นเลขคี่ที่สามารถคำนวณได้จากเครื่องคอมพิวเตอร์โดยที่ $M = 2^b$ เมื่อ b เป็นค่าความยาว 1 word หรือจำนวน bit ใน 1 word ของเครื่องคอมพิวเตอร์ เช่น 32 bit โดย 1 bit สุดท้ายใช้สำหรับแสดงเครื่องหมาย ดังนั้นเลขจำนวนเต็มที่ใหญ่ที่สุดใน 1 word และเป็นเลขคี่ที่คอมพิวเตอร์รับได้ คือ 2^b-1 หรือ 2^{31}-1 มีค่าเท่ากับ 2147483647 นั่นคือค่า M ควรมีค่าเท่ากับ 2147483647

จากค่า a และ M ข้างต้นสามารถเขียนโปรแกรมภาษาฟอร์แทรนที่เป็นโปรแกรมย่อย RAND ดังนี้

```
SUBROUTINE RAND(ISEED , RD)
ISEED = ISEED*16807
IF(ISEED.LT.0) ISEED = ISEED+2147483647+1
RD = ISEED
RD = RD*0.4656613E-9
RETURN
END SUBROUTINE
```

2. <u>การผลิตตัวแปรสุ่มให้มีการแจกแจงแบบปกติ</u>

Box และ Muller(1958) ได้เสนอวิธีสร้างตัวแปรที่มีการแจกแจงแบบปกติมาตรฐานที่มีค่าเฉลี่ยเป็น 0 และค่าความแปรปรวนเป็น 1 พร้อมกัน 2 ค่า Z_1 , Z_2 และแต่ละค่าจะเป็นอิสระซึ่งกันและกัน โดยใช้ตัวผลิต (Generator) Z_1 และ Z_2 ดังรูปต่อไปนี้



จากรูปจะได้ว่า

$$Z_1 = B\cos(\theta) \tag{3.3}$$
$$Z_2 = B\sin(\theta) \tag{3.4}$$

เนื่องจาก $B^2 = Z_1^2 + Z_2^2$ มีการแจกแจงแบบไคกำลังสอง ด้วยระดับขั้นความเสรีเท่ากับ 2 ซึ่งเทียบเท่ากับการแจกแจงชี้กำลัง (Exponential Distribution) ที่มีค่าเฉลี่ยเท่ากับ 2 เราสามารถใช้วิธีการแปลงผกผัน (Inverse Transformation) สร้างเลขสุ่มที่มีการแจกแจงดังต่อไปนี้

$$B = [-2\ln(R)]^{1/2} \tag{3.5}$$

โดยที่ R เป็นเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอในช่วง (0,1)

จากการสมมาตรของการแจกแจงแบบปกติ เราจะได้ว่ามุม θ มีการแจกแจงแบบสม่ำเสมอระหว่าง 0 ถึง 2π เรเดียน และรัศมี B กับ θ เป็นอิสระกัน จากสมการ (3.3) (3.4) และ (3.5) สามารถสร้างเลขสุ่มที่มีการแจกแจงแบบปกติมาตฐานจากเลขสุ่ม 2 ชุด คือ Z_1 และ Z_2 กล่าวคือ

$$Z_1 = [-2\ln(R_1))]^{1/2}\cos(2\pi R_2)$$
$$Z_2 = [-2\ln(R_1))]^{1/2}\sin(2\pi R_2)$$

โดยที่ R_1 และ R_2 เป็นค่าตัวเลขสุ่มที่สร้างจากโปรแกรมย่อย RAND (รูปที่3.3) และเมื่อได้ตัวเลขสุ่มที่มีการแจกแจงแบบปกติมาตรฐานแล้ว เราสามารถสร้างตัวแปรสุ่มให้มีการแจกแจงแบบปกติได้โดยแปลงค่าเลขสุ่มแบบปกติมาตฐานโดยอาศัยฟังก์ชัน

$$X_1 = \mu + \sigma Z_1$$
$$X_2 = \mu + \sigma Z_2$$

จะได้ว่า X_1 และ X_2 มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย $E(X) = \mu$ และ ความแปรปรวน $Var(X) = \sigma^2$ เขียนเป็นสัญลักษณ์ได้ดังนี้ $X_i \sim N(\mu, \sigma^2)$, $i = 1,2,3,\ldots$

โปรแกรมย่อยที่ใช้ในการสร้างเลขสุ่มให้มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย เท่ากับ $E(x) = \mu$ และ ความแปรปรวน $Var(X_i) = \sigma^2$ สามารถเขียนโปรแกรมภาษาฟอร์แทรนที่เป็นโปรแกรมย่อย NORM ดังนี้

```
      SUBROUTINE NORM(ISEED,MEAN,SIGMA,EX1)
      COMMON / ISEED / IX
      REAL  MEAN
      PI=3.1415926
      IF(KK.EQ.1) GOTO 400
410   CALL RAND(ISEED,RD)
      IF((RD.LE.0).OR.(RD.GT.1)) GOTO 410
      RONE=RD
420   CALL RAND(ISEED,RD)
      IF((RD.LE.0).OR.(RD.GT.1)) GOTO 420
      RTWO=RD
      ZONE=SQRT(-2*ALOG(RONE))*COS(2*PI*RTWO)
      ZTWO=SQRT(-2*ALOG(RONE))*SIN(2*PI*RTWO)
      EX1=ZONE*SIGMA+MEAN
      KK=1
```

```
      RETURN
400   EX1=ZTWO*SIGMA+MEAN
      KK=0
      RETURN
      END SUBROUTINE
```