บทที่ ๔

ขั้นตอนการวิจัยและผลการวิจัย

๔.๑ การเตรียมข้อมูลสำหรับการวิจัย

ในการวิจัยนี้ได้ศึกษาข้อมูลในลักษณะ empirical study คือจะสร้างประชากรแบบปกติขึ้นมาประชากรหนึ่ง แล้วสุ่มตัวอย่างจากประชากรนี้ไปศึกษาดูการกระจายค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานของตัวอย่างเหล่านั้น ประชากรที่สร้างขึ้นมามี ๒๐๐๐ ตัว การสร้างโดยใช้โปรแกรมชื่อ GAUSS (IX, S, AM, V)[๑] โดยกำหนดค่าพารามิเตอร์ในโปรแกรมดังนี้

IX คือ เลขจำนวนเต็มคี่ใด ๆ ที่มีจำนวนเลขเก้าหลักหรือน้อยกว่าเก้าหลัก (nine or less digits) (ในการวิจัยใช้ IX = ๙๙)

S คือ ค่าความเบี่ยงเบนมาตรฐาน ๗.๕๐๐

AM คือ ค่าเฉลี่ย ๔๐.๕๐๐

V คือ เลขสุ่มที่ต้องการจะมีการแจกแจงแบบปกติ

รายละเอียดของโปรแกรม GAUSS อยู่ในบทที่ ๓ หัวข้อ ๓.๔ และภาคผนวก ก ข้อมูลที่ได้จากการใช้โปรแกรม GAUSS นี้ ข้อมูลทั้งหมดมี ๒๐๐๐ ตัว ตามตารางในภาคผนวก ข (๑)

๔.๒ การตรวจสอบข้อมูลที่จัดเตรียมได้และผลการตรวจสอบ

๔.๒.๑ ข้อมูล ๒๐๐๐ ตัวที่ได้จากโปรแกรม GAUSS นี้ปรากฏมีค่าสูงสุดเท่ากับ ๖๖.๐๗๕ และค่าต่ำสุดเท่ากับ ๑๘.๓๗๒ สำหรับค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานของข้อมูลทั้ง ๒๐๐๐ ตัวนี้ เราใช้โปรแกรม TAB1 คำนวณหา ใช้สูตรดังนี้

$$\text{ค่าเฉลี่ย} = \sum_{i=1}^{๒๐๐๐} V_i/๒๐๐๐$$

$$\text{ค่าความแปรปรวน} = \sum_{i=1}^{๒๐๐๐} (V_i-\bar{V})^2/๒๐๐๐$$

---

[๑] IBM Application program System/360 Scientific Subroutine Package (360A-CM-03X) Version III.

ใช้โปรแกรมชื่อ TAB1 (X, B, NOVER, UBO, FREFQ, PCT, STATS, N, NV)[๑] ทำการคำนวณ ในการวิจัยนี้ใช้พารามิเตอร์เพียงบางส่วนเท่านั้น คือ Y, NOVAR, B, STATS, N, NV

Y คือ เลขสุ่ม ๒๐๐๐ ตัวที่ได้จาก ๔.๑

NOVAR คือ ตัวแปรที่จะถูก Tabulate = 1

NV คือ จำนวนตัวอย่างที่ใช้ในแต่ละตัวแปร = 1

B(I) = 1.0 I = 1, 3

STATS คือ ค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน ค่าสูงสุดและค่าต่ำสุด

N คือ จำนวนตัวแปรทั้งหมด ๒๐๐๐ ตัว

ค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานเมื่อเปรียบเทียบกับค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานที่ต้องการอยู่ในตารางที่ ๔.๑ ส่วนหลักการทำงานของ TAB1 ในการคำนวณหาค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานนั้น ใช้สูตรมาตรฐานทั่ว ๆ ไปดังนี้ $\mu = \sum_{i=1}^{N} y_i/N$, $\sigma = \sqrt{\frac{\sum_{i=1}^{N} (y_i-\mu)^2}{N}}$

ตารางที่ ๔.๑

แสดงค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานของข้อมูลทั้งหมด ๒๐๐๐ ตัวที่ได้จากโปรแกรม GAUSS เปรียบเทียบกับค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานที่ต้องการ

	ค่าเฉลี่ย	ค่าความเบี่ยงเบน	หมายเหตุ
ค่าของพารามิเตอร์ที่ต้องการ	๔๐.๕๐๐ (μ)	๗.๕๐๐ (σ)	กำหนดโดยผู้ทำวิจัย
ค่าของพารามิเตอร์ที่ได้จากข้อมูล ๒๐๐๐ ตัว ซึ่งสร้างขึ้นจากการใช้โปรแกรม GAUSS	๔๐.๖๒๗ (μ)	๗.๔๔๙ (σ)	จากหัวข้อ ๔.๒.๑
ค่าประมาณของพารามิเตอร์ที่ได้จากข้อมูลตัวอย่าง ๒๐๐ ตัว	๓๙.๙๘๗ ($\bar{X}$)	๗.๘๓๐ (S)	จากขั้นตอนที่ ๑ หัวข้อ ๔.๓

[๑] เรื่องเดียวกัน.

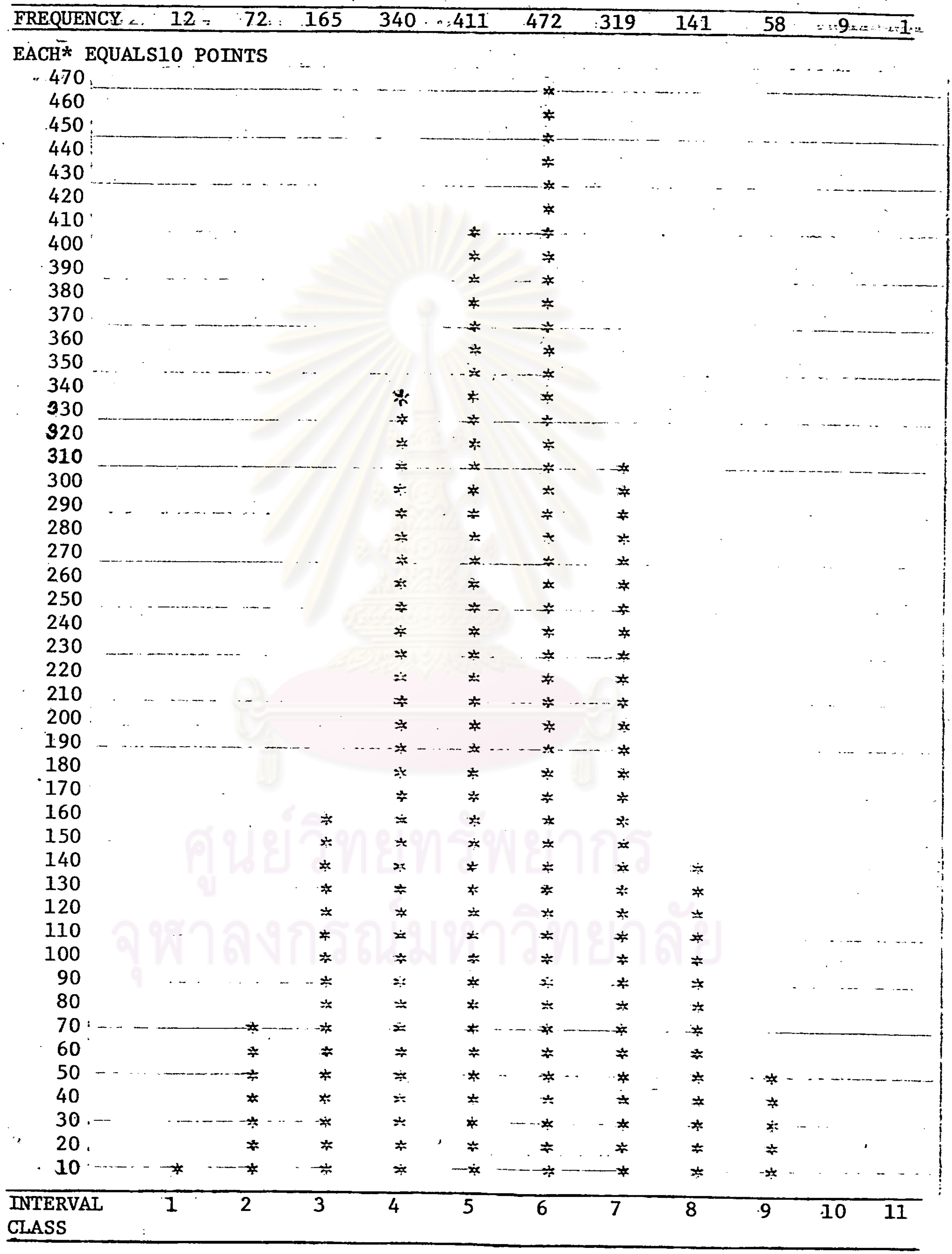
จากตารางจะเห็นว่าค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานของข้อมูล ๒๐๐๐ ตัว ที่ได้จากโปรแกรม GAUSS มีค่าใกล้เคียงกับค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานที่ต้องการ และกราฟฮิสโตแกรมของข้อมูล ๒๐๐๐ ตัว ตามรูป ๔.๑ ก็แสดงลักษณะของฟังก์ชั่น การแจกแจงแบบปกติเป็นที่น่าพอใจ

เลขสุ่ม ๒๐๐๐ ตัวที่สร้างขึ้นมาในหัวข้อ ๔.๑ นี้ สร้างไว้เพื่อช่วยตรวจสอบผลการวิจัยในขั้นต่อไปว่าข้อมูลที่เลือกมาเป็นตัวอย่างสำหรับใช้ในการค้นหาตัวแปรเดียว (Univariate Search) เพื่อทำให้ค่าสถิติทดสอบไคสแควร์ต่ำที่สุดแล้วนั้น ผลของการวิจัยจะยังคงมีลักษณะใกล้เคียงกับลักษณะของประชากรที่ได้สร้างขึ้นในหัวข้อ ๔.๑ หรือไม่ และตัวประมาณค่าพารามิเตอร์ตัวใหม่ที่ได้นี้มีค่าใกล้เคียงกับค่าพารามิเตอร์ของประชากรที่สร้างขึ้นหรือไม่

๔.๒.๒ จากเลขสุ่มในหัวข้อ ๔.๑ จะสุ่มตัวอย่างมาหลาย ๆ ชุดเพื่อทำการศึกษาลักษณะการแจกแจงของข้อมูลตัวอย่างโดยใช้ตัวสถิติทดสอบไคสแควร์ สำหรับการเลือกตัวอย่างจากประชากร ๒๐๐๐ ตัว ในหัวข้อ ๔.๑ นี้ ใช้วิธีการเลือกตัวอย่างแบบมีระบบ (Systemmatic Sampling) ดังได้อธิบายไว้แล้วในหัวข้อ ๓.๔ เพราะสามารถเขียนโปรแกรมใช้คอมพิวเตอร์สุ่มตัวอย่างมาได้ง่าย โปรแกรมที่ใช้คือโปรแกรม SMPLNG (Y, R, N, NI, X)

รูปที่ ๔.๑

ฮิสโตแกรมของเลขสุ่ม ๒๐๐๐ ตัว



โดยที่ Y – คือ เลขสุ่มของประชากร ๒๐๐๐ ตัวที่จะถูกมาเป็นตัวอย่างบางส่วน

R – คือ เลขสุ่มตัวเริ่มต้นที่จะถูกเลือกเป็นตัวอย่างตัวแรกสุดของการเลือกตัวอย่างแต่ละชุด และจะถูกใช้เป็นตัวกำหนดตำแหน่งของเลขสุ่มตัวที่สองที่จะถูกเลือกเป็นตัวอย่าง เช่น R = 1,2,3, ... $\frac{N}{NI}$

โดยใช้สูตร $J = R + (I - 1) \times I1$

J คือ ตำแหน่งของเลขสุ่มที่ถูกเลือกเป็นตัวอย่าง

I1 = N/NI

N คือ จำนวนเลขสุ่มของประชากรทั้งหมดคือ ๒๐๐๐

NI คือ ขนาดของตัวอย่างที่ต้องการ

I คือ ตัวแสดงจำนวนของตัวอย่างมีค่าตั้งแต่ 1 ถึง NI

X คือ ค่าของเลขสุ่มทั้งหมดที่ถูกเลือกเป็นตัวอย่าง

โปรแกรม SMPLNG นี้ ผู้วิจัยได้เขียนขึ้นมาเอง สำหรับข้อมูลตัวอย่างที่เลือกโดยใช้โปรแกรม SMPLNG นี้มีหลายชุดและมีขนาดต่าง ๆ กัน โดยแบ่งออกเป็น ๔ ขนาดคือ

- ตัวอย่างขนาด ๓๐ ที่ใช้ในการวิจัยมี ๑๐ ชุด นั่นคือแต่ละชุดมีขนาด ๑.๕% ของประชากร
- ตัวอย่างขนาด ๕๐ ที่ใช้ในการวิจัยมี ๔๐ ชุด นั่นคือแต่ละชุดมีขนาด ๒.๕% ของประชากร
- ตัวอย่างขนาด ๑๐๐ ที่ใช้ในการวิจัยมี ๒๐ ชุด นั่นคือแต่ละชุดมีขนาด ๕% ของประชากร
- ตัวอย่างขนาด ๒๐๐ ที่ใช้ในการวิจัยมี ๑๐ ชุด นั่นคือแต่ละชุดมีขนาด ๑๐% ของประชากร

ก่อนที่จะพูดถึงผลการวิจัยที่ได้จากตัวอย่างที่เลือกมาศึกษาทั้งหมดรวม ๘๐ ชุด จะขอยกตัวอย่างแสดงผลของการวิจัยทุกขั้นตอนเพียงชุดเดียว ซึ่งมีขนาดของตัวอย่าง ๒๐๐ ตัวอย่าง สำหรับข้อมูลตัวอย่างที่เลือกมานี้รายละเอียดของข้อมูลตัวอย่างอยู่ในภาคผนวก ข (๒)

๔.๓ ขั้นตอนการคำนวณค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐาน ($\bar{X}^*$, S^*) ที่ให้ค่าไคสแควร์ต่ำสุด

ผลการวิจัยในหัวข้อ ๔.๓ นี้เป็นผลการวิจัยของตัวอย่างเพียงตัวอย่างเดียว คือมีขนาด ๒๐๐ ตัวอย่าง เพื่อใช้แสดงขั้นตอนการวิจัยและผลการวิจัยในแต่ละขั้นตอนโดยละเอียด

ขั้นตอนที่ ๑

จากข้อมูลตัวอย่าง ๒๐๐ ตัว หาค่าเฉลี่ย ($\bar{X}$) และค่าความเบี่ยงเบนมาตรฐาน (S) โดยใช้โปรแกรม TAB1 (Y, B, NOVER, UBO, FRFG, PCT, STATS, N, NV)

Y คือ เลขสุ่ม ๒๐๐ ตัว

N คือ จำนวนตัวแปรสุ่มทั้งหมดมีค่า ๒๐๐

สำหรับพารามิเตอร์ตัวอื่น ๆ ใช้เหมือนกับการหา μ และ σ

ได้ค่าเฉลี่ย ($\bar{X}$) = 39.987

ค่าความเบี่ยงเบนมาตรฐาน (S) = 7.830

ขั้นตอนที่ ๒

นำข้อมูล ๒๐๐ ตัววาดกราฟฮิสโตแกรมเพื่อดูลักษณะการแจกแจงของข้อมูลว่าใกล้เคียงฟังก์ชั่นการแจกแจงแบบใด ซึ่งการวาดกราฟฮิสโตแกรมต้องอาศัยค่าความถี่ในแต่ละช่วงของข้อมูล การหาค่าความถี่ในแต่ละช่วงของข้อมูลใช้โปรแกรมชื่อ DIST (N, RANGE, INTER) ซึ่งมีตัวแปรในข้อความ COMMON ๓ ตัว คือ X, JF, CLALT

21.161
18.661
2.500

โดยที่ N คือ จำนวนข้อมูลทั้งหมด = ๒๐๐

RANGE คือ ช่วงห่างในแต่ละช่วง = ๔.๕ 2.5

INTER คือ จำนวนช่วงทั้งหมด = J-1

X คือ ค่าของข้อมูลทั้งหมด ๒๐๐ ตัว

JF คือ ค่าความถี่ในแต่ละช่วง (observed)

CLALT คือ ค่าของช่วงแต่ละช่วง lower, upper limit

โปรแกรม DIST ผู้วิจัยเขียนขึ้นเอง ซึ่งให้ผลตามที่แสดงไว้ในตารางที่ ๔.๒

ตารางที่ ๔.๒

แสดงค่าความถี่ของข้อมูลและค่าคาดหวังของข้อมูลในแต่ละช่วง

CLALT(I) CLALT(I+1) JF(I)

ขอบเขตล่าง (LOWER LIMIT)	ขอบเขตบน (UPPER LIMIT)	ความถี่ (O_i) OBSERVE (FREQUENCY)	ค่าคาดหวัง (E) EXPECTED (FREQUENCY)
18.661	21.161	1.000	0.846
21.161	23.661	4.000	1.844
23.661	26.161	3.000	3.630
26.161	28.661	8.000	6.461
28.661	31.161	13.000	10.392
31.161	33.661	11.000	15.109
33.661	36.161	21.000	19.854
36.161	38.661	14.000	23.581
38.661	41.161	38.000	25.315
41.161	43.661	22.000	24.563
43.661	46.161	23.000	21.543
46.161	48.661	14.000	17.077
48.661	51.161	19.000	12.236
51.161	53.661	2.000	7.924
53.661	56.161	2.000	4.638
56.161	58.661	4.000	2.454
58.661	61.161	0.0	1.173
61.161	63.661	0.0	0.507
63.661	66.161	1.000	0.198

เมื่อได้ค่าความถี่ในแต่ละช่วงแล้ว นำค่าความถี่ในแต่ละช่วงวาดกราฟฮิสโตแกรม โดยใช้โปรแกรม HIST (NU, FREQ, IN)[๑]

NU คือ จำนวนรูปกราฟฮิตโตแกรม

FREQ คือ ค่าความถี่ในแต่ละช่วง (O_i) ~ JF(I)

IN คือ จำนวนช่วงทั้งหมด (INTER)

การดูลักษณะของข้อมูลจากกราฟเพียงอย่างเดียวไม่เป็นการเพียงพอที่จะสรุปว่า ข้อมูลที่กำลังศึกษามีการแจกแจงตามกราฟที่ปรากฏ และมีค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานตามที่หาได้ในตอนต้น

ดังนั้นเราจึงต้องมีการทดสอบว่าข้อมูลชุดนี้มีการแจกแจงตามกราฟที่คาดว่าจะเป็นหรือไม่ โดยการใช้ตัวสถิติทดสอบไคสแควร์ (χ^2) ตามสมการ (2.1) ต่อไป

<u>ขั้นตอนที่ ๓</u>

ทำการทดสอบโดยใช้ตัวสถิติทดสอบไคสแควร์ (χ^2) โดยอาศัยค่าความถี่ในแต่ละช่วงของข้อมูล ค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐาน พร้อมตั้งสมมติฐานในการทดสอบว่าข้อมูลมีการแจกแจงตามลักษณะกราฟที่วาดได้ จากรูป ๔.๒ จะเห็นว่ากราฟหรือฮิตโตแกรมของข้อมูล ๒๐๐ ตัวนี้มีลักษณะการแจกแจงเข้าใกล้ฟังก์ชั่นการแจกแจงแบบปกติ

ดังนั้น H_o : ข้อมูลมีการแจกแจงแบบปกติ

H_1 : ข้อมูลไม่มีการแจกแจงแบบปกติ

ซึ่งเรียกวิธีการทดสอบนี้ว่าการทดสอบภาวะสารูปสนิทดี (goodness of fit test) การคำนวณค่าไคสแควร์ตามสมการ (2.1) นั้น จำเป็นต้องใช้ค่าคาดหวังของข้อมูลในการคำนวณ วิธีการคำนวณหาค่าคาดหวังนั้นได้อธิบายไว้แล้วในบทที่ ๓ หัวข้อ ๓.๑ โดยใช้โปรแกรมชื่อ CHISQ (X, NX, N, W, NI, O, ER, FF) และกราฟฮิสโตแกรมของค่าคาดหวังที่คำนวณได้อยู่ในรูปที่ ๔.๓

---

[๑] เรื่องเดียวกัน

รูปที่ ๔.๒

แสดงฮิสโตแกรมของข้อมูล ๒๐๐ ตัวอย่างที่ใช้โปรแกรม HIST

```
EQUENCY    1    4    3    8   13   11   21   14   38   22   23   14   19    2    2    4    0    0    1
38                                                 *
37                                                 *
36                                                 *
35                                                 *
34                                                 *
33                                                 *
32                                                 *
31                                                 *
30                                                 *
29                                                 *
28                                                 *
27                                                 *
26                                                 *
25                                                 *
24                                                 *
23                                                 *         *
22                                                 *    *    *
21                                       *         *    *    *
20                                       *         *    *    *
19                                       *         *    *    *         *
18                                       *         *    *    *         *
17                                       *         *    *    *         *
16                                       *         *    *    *         *
15                                       *         *    *    *         *
14                                       *    *    *    *    *    *    *
13                             *         *    *    *    *    *    *    *
12                             *         *    *    *    *    *    *    *
11                             *    *    *    *    *    *    *    *    *
10                             *    *    *    *    *    *    *    *    *
9                              *    *    *    *    *    *    *    *    *
8                         *    *    *    *    *    *    *    *    *    *
7                         *    *    *    *    *    *    *    *    *    *
```

รูปที่ ๔.๓

แสดงฮิสโตแกรมที่ได้จากค่าคาดหวังของความถี่ก่อนใช้ UNISER

```
FREQUENCY     0    2    4    7   11   15   20   24   25   24   20   16   11    7    4    2    1    0    0

  25                                                 *
  24                                            *    *    *
  23                                            *    *    *
  22                                            *    *    *
  21                                            *    *    *
  20                                       *    *    *    *    *
  19                                       *    *    *    *    *
  18                                       *    *    *    *    *
  17                                       *    *    *    *    *
  16                                       *    *    *    *    *    *
  15                                  *    *    *    *    *    *    *
  14                                  *    *    *    *    *    *    *
  13                                  *    *    *    *    *    *    *
  12                                  *    *    *    *    *    *    *
  11                             *    *    *    *    *    *    *    *    *
  10                             *    *    *    *    *    *    *    *    *
   9                             *    *    *    *    *    *    *    *    *
   8                             *    *    *    *    *    *    *    *    *
   7                        *    *    *    *    *    *    *    *    *    *    *
   6                        *    *    *    *    *    *    *    *    *    *    *
   5                        *    *    *    *    *    *    *    *    *    *    *
   4                   *    *    *    *    *    *    *    *    *    *    *    *    *
   3                   *    *    *    *    *    *    *    *    *    *    *    *    *
   2              *    *    *    *    *    *    *    *    *    *    *    *    *    *    *
   1              *    *    *    *    *    *    *    *    *    *    *    *    *    *    *    *

INTERVAL      1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18   19
CLASS
```

X คือ ค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐาน

NX คือ ตัวนับจำนวนครั้งที่คำนวณค่าไคสแควร์

N คือ จำนวนช่วงของข้อมูลทั้งหมด = 20

W คือ ค่าของช่วงแต่ละช่วง

NI คือ จำนวนข้อมูลทั้งหมด = 200

O คือ ความถี่ของข้อมูลในแต่ละช่วง

ER คือ ค่าคาดหวังของข้อมูลในแต่ละช่วงที่โปรแกรมคำนวณใช้

FF คือ ค่าไคสแควร์ที่โปรแกรมคำนวณได้

โปรแกรม CHISQ ใช้ในการหาค่าคาดหวังของข้อมูลในแต่ละช่วงและค่าไคสแควร์ ค่าคาดหวังของข้อมูลที่ได้จากโปรแกรม CHISQ อยู่ในตารางที่ ๔.๒ และค่าไคสแควร์ (X^2) ที่ได้จากโปรแกรม CHISQ เท่ากับ ๓๒.๓๒๔ ค่าไคสแควร์จากตารางที่ระดับความเชื่อมั่น ๙๙% และองศาแห่งความเป็นอิสระ ๑๙-๓ = ๑๖ มีค่า ๓๒.๐๐๐ ค่าไคสแควร์ที่ได้จากการคำนวณ (๓๒.๓๒ ๔) มากกว่าไคสแควร์ที่ได้จากตาราง (๓๒.๐๐) ทำให้เราต้องปฏิเสธสมมติฐานที่ว่าข้อมูลที่กำลังศึกษามีการแจกแจงแบบปกติ นั่นคือเราต้องยอมรับว่าข้อมูลนี้ไม่มีการแจกแจงแบบปกติ
ที่มีค่าเฉลี่ยเป็น ๓๙.๙๘๗ และค่าความเบี่ยงเบนมาตรฐานเป็น ๗.๘๓๐

จากผลการทดสอบเราปฏิเสธสมมติฐานในตอนแรก เรามีความจำเป็นต้องตั้งสมมติฐานขึ้นใหม่ แล้วทดสอบสมมติฐานต่อไปจนกว่าจะยอมรับสมมติฐานใดสมมติฐานหนึ่ง ด้วยการเปลี่ยนค่าเฉลี่ยหรือค่าความเบี่ยงเบนมาตรฐานหรือทั้งสองค่า ซึ่งเรียกวิธีการอย่างนี้ว่า Cut and try วิธีการ cut and try ไม่สามารถพิสูจน์ให้เห็นโดยชัดเจนว่าจะได้ค่าไคสแควร์ต่ำสุดเมื่อใด ดังนั้นเราจึงหันไปใช้เทคนิคการค้นหาตัวประมาณค่าพารามิเตอร์ คือค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานตัวใหม่ที่ทำให้ค่าไคสแควร์ต่ำสุดแทนวิธี cut and try โดยใช้เทคนิคการค้นหาค่าอุตมะ (Optimization technique) ที่เรียกกันว่าวิธีการค้นหาตัวแปรเดียว (Univeriate Search) ซึ่งจะได้อธิบายโดยละเอียดในขั้นตอนต่อไป

<u>ขั้นตอนที่ ๔</u>

ใช้เทคนิคการค้นหาค่าอุตมะด้วยวิธีการค้นหาตัวแปรเดียว ประมาณค่าพารามิเตอร์ คือค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานตัวใหม่ โดยมีจุดมุ่งหมายที่จะทำให้ค่าไคสแควร์ต่ำสุด ด้วยการอาศัยค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานจากข้อมูลที่ได้ในตอนแรกสุดเป็นจุดเริ่มต้นในการค้นหา โดยการใช้โปรแกรมชื่อ UNISER ซึ่งวิธีการได้อธิบายไว้ในบทที่ ๒ หัวข้อ ๒.๔ โปรแกรม UNISER ผู้วิจัยเขียนขึ้นมาเอง สัญญลักษณ์ต่าง ๆ ในโปรแกรม "UNISER" มีดังนี้

N = จำนวนค่าประมาณของพารามิเตอร์ที่ต้องการหาค่า

R = ช่วงก้าว (Step length) ในการหาค่าประมาณของพารามิเตอร์

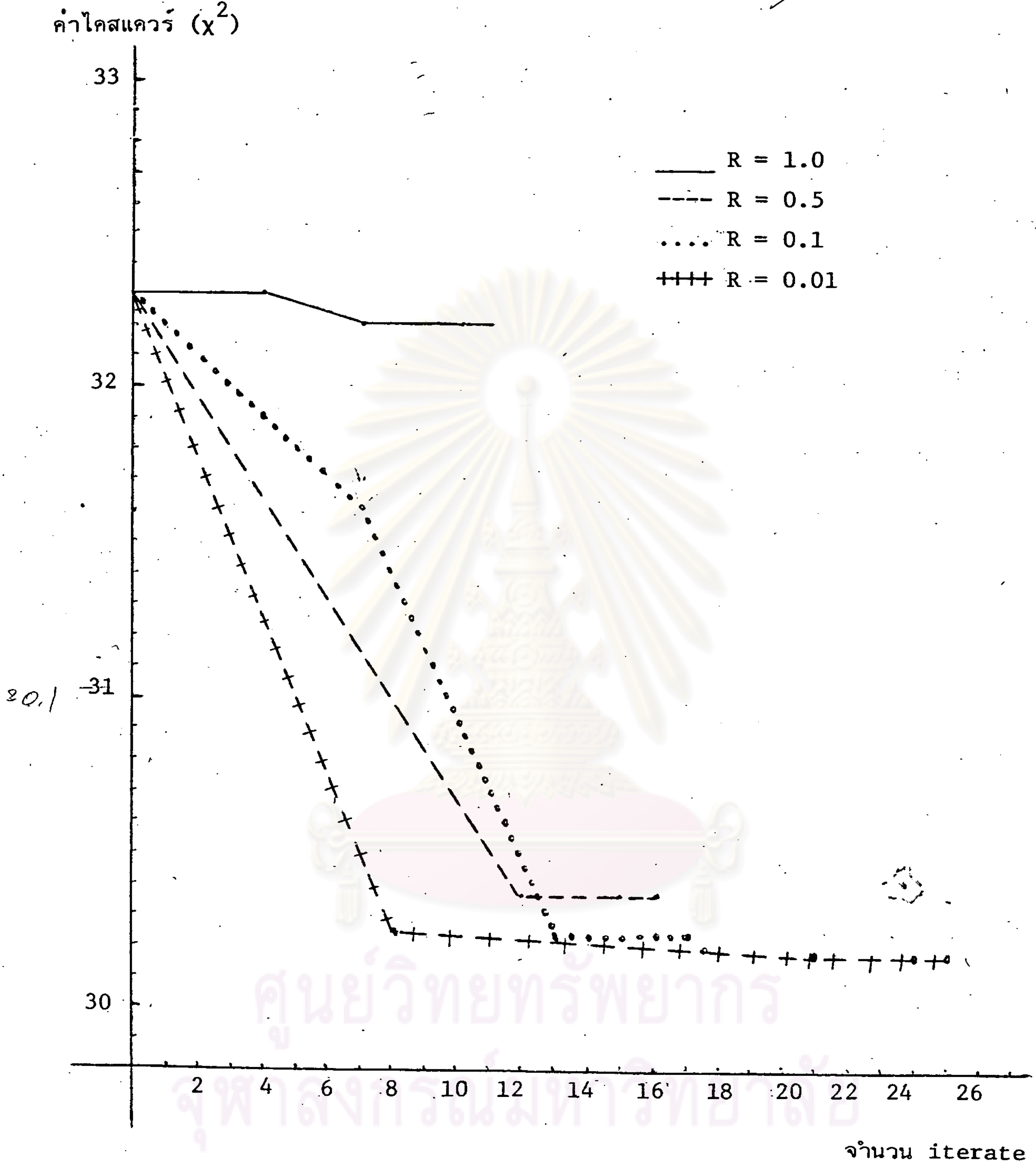
X(1) = ค่าเฉลี่ย ($\bar{X}$) ในฟังก์ชั่นการแจกแจงแบบปกติ

X(2) = ค่าเบี่ยงเบนมาตรฐาน (S) ในฟังก์ชั่นการแจกแจงแบบปกติ

E(1,1) , E(1,2) = ทิศทางอิสระเชิงเส้นทิศทางที่หนึ่งที่จะใช้ในรอบแรกของการค้นหา

E(2,1) , E(2,2) = ทิศทางอิสระเชิงเส้นทิศทางที่สองที่จะใช้ในรอบแรกของการค้นหา

การใช้โปรแกรม UNISER ต้องกำหนด R ซึ่งเป็นช่วงก้าวการค้นหา ถ้ากำหนดค่ามากเกินไปความถูกต้องมีน้อยแต่ไม่เสียเวลา ถ้าใช้ R มีค่าน้อยมาก ความถูกต้องมีมากแต่เสียเวลาในการค้นหา ซึ่งได้แสดงกราฟการค้นหาเมื่อใช้ R ในระดับต่าง ๆ ในรูปที่ ๔.๙



<u>รูปที่ ๔.๔</u>

แสดงการใช้ช่วงก้าว (R) ในระดับต่าง ๆ พร้อมทั้งแสดงจำนวน iterate ในแต่ละระดับของ R ในการค้นหาค่าไคสแควร์ที่ต่ำที่สุด

หลังจากใช้โปรแกรม UNISER โดยใช้ R = 0.01 ผลการหาค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานตัวใหม่ที่ทำให้ค่าไคสแควร์ต่ำสุด แสดงไว้ในตารางที่ ๔.๓

ตารางที่ ๔.๓

แสดงค่าประมาณของพารามิเตอร์ในรอบต่าง ๆ ของการค้นหาค่าอุตมะ

ด้วยวิธีการค้นหาตัวแปรเดียว (R = 0.01)

รอบที่	ทิศทางที่	ค่าเฉลี่ย	ค่าความเบี่ยงเบนมาตรฐาน	ค่าไคสแควร์
จุดเริ่มต้น		๓๙.๙๘๗	๗.๘๓๐	๓๒.๓๒๔
รอบที่ ๑	ทิศทางที่ ๑	๔๐.๔๑๗	๗.๘๓๐	๓๑.๕๙๙
	ทิศทางที่ ๒	๔๐.๔๑๗	๘.๒๕๙	๓๐.๑๖๖
รอบที่ ๒	ทิศทางที่ ๑	๔๐.๓๖๗	๘.๒๕๙	๓๐.๑๖๒
	ทิศทางที่ ๒	๔๐.๓๗๗	๘.๒๔๙	๓๐.๑๖๕

จากตารางที่ ๔.๓ เป็นผลจากการใช้โปรแกรม UNISER เพื่อหาค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานตัวใหม่ ($\bar{X}^*$, S^*) ที่ทำให้ค่าไคสแควร์ต่ำสุด (χ^{*2}) โดยอาศัยค่าเฉลี่ยเท่ากับ ๓๙.๙๘๗ และค่าความเบี่ยงเบนมาตรฐานเท่ากับ ๗.๘๓๐ เป็นจุดเริ่มต้นคำนวณค่าไคสแควร์ได้ ๓๒.๓๒๔

ในรอบแรกในทิศทางที่ ๑ ต้องการหาค่าเฉลี่ยตัวใหม่ที่ทำให้ค่าไคสแควร์ต่ำที่สุด (ต่ำกว่า ๓๒.๓๒๔) โดยให้ค่าความเบี่ยงเบนคงที่ จะพบว่าค่าเฉลี่ยตัวใหม่มีค่าเป็น ๔๐.๔๑๗ โดยได้ค่าไคสแควร์ ๓๑.๕๙๙ หลังจากนั้นก็จะหาค่าความเบี่ยงเบนมาตรฐานตัวใหม่ที่ทำให้ค่าไคสแควร์ต่ำสุด (ต่ำกว่า ๓๑.๕๙๙) โดยให้ถือว่าค่าเฉลี่ยคงที่ ค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานที่คำนวณในแต่ละรอบ จะเป็นจุดเริ่มต้นใหม่ของรอบถัดไปในการคำนวณค่าเฉลี่ย

และค่าความเบี่ยงเบนมาตรฐานตัวใหม่ จะพบว่าค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐาน ในรอบที่ ๒ ทิศทางที่ ๑ เป็นค่าประมาณของพารามิเตอร์ที่ทำให้ค่าไคสแควร์ต่ำที่สุด (χ^{*2}) เท่ากับ ๓๐.๑๖๒ โดยมีค่าเฉลี่ย ($\bar{X}^*$) เท่ากับ ๔๐.๓๖๗ และค่าความเบี่ยงเบนมาตรฐาน (S^*) เท่ากับ ๘.๒๕๙ ค่าไคสแควร์ต่ำสุด (χ^{*2}) นี้คำนวณมาจากค่าคาดหวังที่ได้มาจากค่าเฉลี่ย ($\bar{X}^*$) และค่าความเบี่ยงเบนมาตรฐาน (S^*) ซึ่งแสดงค่าคาดหวังในช่วงต่าง ๆ ในตารางที่ ๔.๔

ตารางที่ ๔.๔

แสดงค่าคาดหวังของความถี่ใหม่หลังจากใช้วิธีการค้นหาตัวแปรเดียว จากข้อมูล ๒๐๐ ตัว และใช้ R=0.01

ขอบเขตล่าง (LOWER LIMIT)	ขอบเขตบน (UPPER LIMIT)	ความถี่ (OBSERVE FREQUENCY) (O)	ค่าคาดหวัง (EXPECTED FREQUENCY) (E)
18.662	21.162	1.000	1.137
21.162	23.662	4.000	2.291
23.662	26.162	3.000	4.213
26.162	28.662	8.000	7.073
28.662	31.162	13.000	10.840
31.162	33.662	11.000	15.168
33.662	36.162	21.000	19.374
36.162	38.662	14.000	22.591
38.662	41.162	38.000	24.049
41.162	43.662	22.000	23.370
43.662	46.162	23.000	20.732
46.162	48.662	14.000	16.790
48.662	51.162	19.000	12.414
51.162	53.662	2.000	8.379
53.662	56.162	2.000	5.162
56.162	58.662	4.000	2.904
58.662	61.162	0.0	1.491
61.662	63.662	0.0	0.699
63.662	66.162	1.000	0.299

<u>ตารางที่ ๔.๕</u>

สรุปผลภายหลังจากการใช้การค้นหาตัวแปรเดียวกับเลขสุ่ม ๒๐๐ ตัว
โดยใช้ R=0.01

	ก่อนใช้ UNISER	หลังใช้ UNISER
ค่าเฉลี่ย	39.987 ($\bar{X}$)	40.367 ($\bar{X}^*$)
ค่าความเบี่ยงเบนมาตรฐาน	7.830 (S)	8.259 (S^*)
ค่าไคสแควร์	32.324 (χ^2)	30.162 (χ^{*2})

และเมื่อนำค่าคาดหวังของความถี่อันใหม่ไปสร้างฮิสโตแกรมจะได้ฮิสโตแกรมดังรูปที่ ๔.๓ ซึ่งแสดงลักษณะใกล้เคียงการแจกแจงแบบปกติมากกว่ารูปที่ ๔.๒ และค่าไคสแควร์ใหม่ที่ได้ก็ต่ำกว่า χ^2 ตัวเดิมหรือดีกว่าเดิม และเมื่อนำค่าไคสแควร์ (χ^{*2}) ใหม่นี้ไปเปรียบเทียบกับค่าไคสแควร์จากตารางที่ระดับนัยสำคัญ ๐.๐๑ จะพบว่าถ้าเราตั้งสมมติฐานในตอนต้นว่า H_o : ข้อมูลมีการแจกแจงแบบ N(๔๐.๓๖๗ , ๘.๒๕๙)

H_1 : ข้อมูลไม่มีการแจกแจงแบบ N(๔๐.๓๖๗ , ๘.๒๕๙)

ซึ่งเมื่อคำนวณค่าไคสแควร์แล้ว ก็คือค่าไคสแควร์ (X^{*2}) เท่ากับ ๓๐.๑๖๒ น้อยกว่าค่าไคสแควร์จากตารางคือ ๓๒.๐๐๐ แสดงว่ายอมรับสมมติฐานที่ว่าข้อมูลมีการแจกแจงแบบ N(๔๐.๓๖๗ , ๘.๒๕๙) นั่นคือ จะเห็นว่าการใช้วิธีการค้นหาตัวแปรเดียวในการค้นหาค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐาน ซึ่งเป็นตัวประมาณค่าพารามิเตอร์ของประชากรที่สนใจศึกษา โดยการทำให้ค่าไคสแควร์ในการทดสอบลักษณะการแจกแจงของข้อมูลต่ำสุดที่ช่วยลดความผิดพลาดแบบที่ ๑ (Type I-error)

<u>ขั้นตอนที่ ๕</u>

ใช้สถิติทดสอบ t-test ทำการทดสอบตัวประมาณค่าเฉลี่ยและตัวประมาณค่าความเบี่ยงเบนมาตรฐานที่ได้จากทั้ง ๓ วิธี ได้แก่

- จากการ optimize χ^2
- จาก conventional method ๒๐๐ ตัวอย่าง

รูปที่ ๔.๕

แสดงกราฟฮิสโตแกรมที่ได้จากค่าคาดหวังของความถี่ใหม่ที่ได้จากการใช้ UNISER

```
FREQUENCY   1   2   4   7  10  15  19  22  24  23  20  16  12   8   5   2   1   0   0
  24                                   *
  23                                   *   *
  22                               *   *   *
  21                               *   *   *
  20                               *   *   *   *
  19                           *   *   *   *   *
  18                           *   *   *   *   *
  17                           *   *   *   *   *
  16                           *   *   *   *   *   *
  15                       *   *   *   *   *   *   *
  14                       *   *   *   *   *   *   *
  13                       *   *   *   *   *   *   *
  12                       *   *   *   *   *   *   *   *
  11                       *   *   *   *   *   *   *   *
  10                   *   *   *   *   *   *   *   *   *
   9                   *   *   *   *   *   *   *   *   *
   8                   *   *   *   *   *   *   *   *   *   *
   7               *   *   *   *   *   *   *   *   *   *   *
   6               *   *   *   *   *   *   *   *   *   *   *
   5               *   *   *   *   *   *   *   *   *   *   *   *
   4           *   *   *   *   *   *   *   *   *   *   *   *   *
   3           *   *   *   *   *   *   *   *   *   *   *   *   *
   2       *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
   1   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

INTERVAL    1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19
CLASS
```

ตารางที่ ๔.๖

ผลการทดสอบตัวประมาณค่าเฉลี่ยที่ได้จากการ Optimize ไคสแควร์

โดยใช้สถิติทดสอบ t-test

Test μ	ระดับนัยสำคัญ = 0.05		
	ใช้ $\bar{X}^*$ ในการคำนวณ	ใช้ $\bar{X}$ ในการคำนวณ	ใช้ $\bar{X}^*$ และ $\bar{X}$ ในการคำนวณ
สมมติฐาน	$H_o : \mu = 40.627$ $H_1 : \mu \neq 40.627$	$H_o : \mu = 40.627$ $H_1 : \mu \neq 40.627$	$H_o : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$
ค่าสถิติทดสอบที่ได้จากการคำนวณ	$t = -0.445$	$t = -1.156$	$t = 0.472$
องศาแห่งความเป็นอิสระ	199	199	∞
ค่าสถิติทดสอบจากตาราง	$\pm$ 1.960	$\pm$ 1.960	$\pm$ 1.960
ผลการทดสอบสมมติฐาน	$-1.96 < t < 1.96$ ยอมรับ H_o	$-1.96 < t < 1.96$ ยอมรับ H_o	$-1.96 < t < 1.96$ ยอมรับ H_o

จากตารางที่ ๔.๖ จะพบว่าตัวประมาณค่าเฉลี่ยที่ได้ใหม่ไม่แตกต่างไปจากค่าเฉลี่ยของประชากร และค่าประมาณของค่าเฉลี่ยที่ได้จากวิธีเดิม

เนื่องจากว่าโดยปกติแล้วถ้าต้องการทราบว่าสถิติตัวไหนเป็นตัวประมาณค่าพารามิเตอร์ที่ดีของประชากรแล้วจะดูได้จาก

- สถิติควรจะได้ค่าที่เที่ยงตรงไม่เอนเอียง (unbiased) หมายความว่าสถิติใดก็ตามเมื่อหาค่าเฉลี่ยของค่าเฉลี่ย ($\bar{\bar{X}}$) ของตัวอย่างทั้งหมดมีค่าเท่ากับ μ (ค่าเฉลี่ยของประชากร) และค่าเฉลี่ยของค่าความเบี่ยงเบนมาตรฐาน (S) ของตัวอย่างทั้งหมดมีค่า

เท่ากับ σ (ค่าความเบี่ยงเบนของประชากร) แล้วแสดงว่าสถิติตัวนั้นเป็นสถิติที่ไม่เอนเอียง (unbias estimater)

- สถิติควรมีประสิทธิภาพ หมายถึงถ้าสถิติไม่เอนเอียงใด ๆ ให้ค่าใกล้เคียงกับค่าจริงมากกว่าสถิติอื่น ๆ เรียกว่ามีประสิทธิภาพ เช่น $V(\bar{X}_1) < V(\bar{X}_2)$ แสดงว่า สถิติ $\bar{X}_1$ มีประสิทธิภาพดีกว่าสถิติ $\bar{X}_2$

เนื่องจากค่าประมาณค่าพารามิเตอร์ตัวใหม่ที่ได้จากการ Optimize ไคสแควร์ไม่สามารถเขียนออกมาเป็นสูตรทางคณิตศาสตร์ได้ จึงใช้วิธีการศึกษาจากตัวอย่างหลาย ๆ ชุด เพื่อศึกษาถึงค่าประมาณค่าพารามิเตอร์ตัวใหม่ว่าเป็นตัวประมาณค่าที่ดีกว่าค่าประมาณของพารามิเตอร์ตัวเก่าหรือไม่ ซึ่งผลการทดลองอยู่ในหัวข้อ ๔.๔

๔.๔ ผลการศึกษาข้อมูลตัวอย่างขนาด ๒๐๐ ตัวอย่าง จำนวน ๑๐ ชุด

ในหัวข้อ ๔.๒.๒ และ ๔.๓ ได้แสดงวิธีการและผลการคำนวณค่าไคสแควร์ (χ^{2^*}) ต่ำสุดของข้อมูลขนาด ๒๐๐ ตัวอย่าง จำนวน ๑ ชุด จะพบว่าค่าไคสแควร์ที่คำนวณได้ใหม่นี้ต่ำกว่าค่าไคสแควร์เดิม และทำการทดสอบตัวประมาณค่าพารามิเตอร์ตัวใหม่ สรุปผลได้ว่าค่าประมาณค่าเฉลี่ยตัวใหม่ไม่แตกต่างจากตัวประมาณค่าเฉลี่ยตัวเดิมอย่างมีนัยสำคัญ และยังมีค่าใกล้เคียงกับค่าเฉลี่ยของประชากรมากกว่า แต่สำหรับค่าประมาณค่าความเบี่ยงเบนมาตรฐานตัวใหม่จะแตกต่างจากตัวประมาณค่าความเบี่ยงเบนมาตรฐานเดิมอย่างมีนัยสำคัญ

ในหัวข้อนี้ เราจะได้ใช้วิธีการตามขั้นตอนที่ได้กล่าวไว้ในหัวข้อ ๔.๒.๒ และทำการศึกษาข้อมูลตัวอย่างขนาด ๒๐๐ จำนวน ๑๐ ชุด ดังได้แสดงผลไว้ในตารางต่อไปนี้

ตารางที่ ๔.๗

เปรียบเทียบค่าสถิติทดสอบแบบไคสแควร์ที่ได้จากวิธี Conventional Method กับวิธี Optimization Method

ตัวอย่างที่	ค่าเฉลี่ย ($\bar{X}$)	ค่าความเบี่ยงเบนมาตรฐาน (S)	ค่าไคสแควร์ (χ^2)	ค่าเฉลี่ยใหม่ ($\bar{X}^*$)	ค่าความเบี่ยงเบนมาตรฐาน (S^*)	ค่าไคสแควร์ที่ Optimal (χ^{2*})
1	40.940	7.672	11.409	41.030	7.982	10.694
2	41.150	6.771	19.299	41.080	7.271	17.255
3	40.854	7.19	8.417	40.804	7.750	6.941
4	39.987	7.83	32.324	40.377	8.250	30.162
5	40.741	7.506	12.930	40.611	7.906	11.953
6	40.264	8.186	21.624	39.834	8.806	19.501
7	40.847	7.399	13.117	40.697	7.909	11.872
8	40.755	6.875	7.544	40.775	7.335	6.252
9	40.478	7.443	9.663	40.308	7.813	8.724
10	40.327	7.337	13.803	40.467	8.577	11.688

จากตารางที่ ๔.๘ จะพบว่าค่าสถิติทดสอบไคสแควร์ที่ได้จากวิธี Conventional Method จะสูงกว่าค่าสถิติทดสอบไคสแควร์ต่ำสุดที่ได้จากวิธี optimization method ทุก ๆ ตัวอย่างที่นำมาทดสอบ

ตารางที่ ๔.๘

เปรียบเทียบผลการยอมรับและปฏิเสธสมมติฐานของค่าไคสแควร์ที่ได้จาก ๒ วิธี

ในระดับความมีนัยสำคัญระดับต่าง ๆ

ตัวอย่างที่	d.f	χ^2	χ^{2^*}	$\alpha = 0.01$		$\alpha = 0.025$		$\alpha = 0.05$	
				ผลจาก χ^2	ผลจาก χ^{2^*}	ผลจาก χ^2	ผลจาก χ^{2^*}	ผลจาก χ^2	ผลจาก χ^{2^*}
1	15	11.409	10.694	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ
2	11	19.299	17.255	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ ปฏิเสธ	ยอมรับ
3	11	8.417	6.941	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ
4	16	32.324	30.162	ปฏิเสธ	ยอมรับ	ปฏิเสธ	ปฏิเสธ	ปฏิเสธ	ปฏิเสธ
5	13	12.930	11.953	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ
6	14	21.624	19.501	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ
7	11	13.117	11.872	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ
8	11	7.544	6.252	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ
9	13	9.663	8.724	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ
10	13	13.803	11.688	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ	ยอมรับ

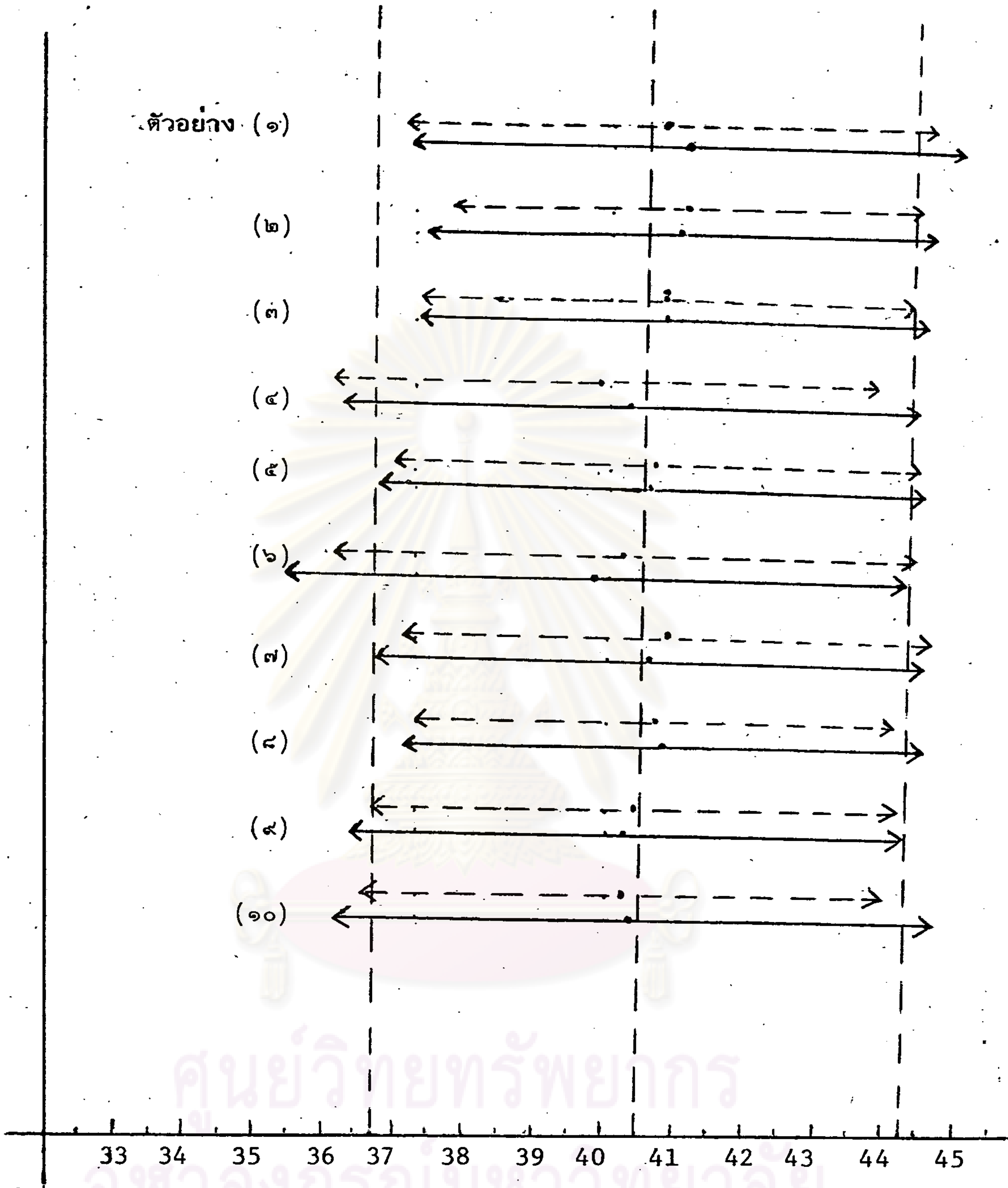
จากตารางที่ ๔.๘ จะพบว่าตัวอย่างที่ ๔ ที่ระดับนัยสำคัญ $\alpha = 0.01$ ถ้าใช้ค่าไคสแควร์ที่มาจากการใช้วิธี Conventional Method ผลการทดสอบปฏิเสธสมมติฐานที่ตั้งไว้ แต่ถ้าใช้ค่าไคสแควร์ที่ได้จากการใช้วิธี Optimization Method ผลการทดสอบยอมรับสมมติฐานที่ตั้งไว้ สำหรับที่ระดับนัยสำคัญ $\alpha = 0.025$ และ $\alpha = 0.05$ ผลการทดสอบปฏิเสธสมมติฐานไม่ว่าจะใช้ค่าไคสแควร์ที่ได้จากวิธีใดก็ตาม สำหรับตัวอย่างอื่น ๆ ผลการทดสอบยอมรับสมมติฐานทั้ง ๒ วิธี ไม่ว่าจะใช้ระดับความมีนัยสำคัญระดับใดก็ตาม

<u>ตารางที่ ๔.๙</u>

เปรียบเทียบค่าเฉลี่ยที่ได้จาก ๒ วิธี

ตัวอย่าง	$\bar{X}$	$\bar{X}^*$	$\mu = 40.627$ $\bar{X}^*$ T-test	$\mu = 40.627$ $\bar{X}$
1	40.940	41.030	0.714	0.577
2	41.150	41.080	0.881	1.092
3	40.854	40.804	0.323	0.446
4	39.987	40.377	0.429	-1.156
5	40.741	40.611	-0.029	0.2.5
6	40.264	39.834	-1.274	-0.627
7	40.847	40.697	0.125	0.421
8	40.755	40.775	0.285	0.263
9	40.478	40.308	-0.577	-0.283
10	40.327	40.467	-0.264	-0.578

ผลการทดสอบโดยใช้ค่าสถิติทดสอบแบบ t จะพบว่าค่าเฉลี่ยที่ได้จากการใช้วิธี Conventional Method และวิธี Optimization Method ให้ผลไม่แตกต่างไปจากค่าเฉลี่ยของประชากร $\mu = 40.627$ ตัว



รูปที่ ๔.๖

แสดงค่าเฉลี่ยและค่าความเบี่ยงเบนมาตรฐานที่ได้จาก ๒ วิธี โดยใช้ตัวอย่างขนาด ๒๐๐ ตัวอย่าง จำนวน ๑๐ ชุด : ←- - - -→ จาก Conventional Methods

←——→ จาก Optimization Methods

๔.๕ ข้อสังเกต

๔.๕.๑ จากผลการวิจัยที่ได้ทดลองกับข้อมูลตัวอย่างขนาด ๒๐๐ จำนวน ๒๐ ชุด ได้ข้อสังเกตดังนี้

ก. ค่าไคสแควร์ที่ได้จากการใช้วิธีการค้นหาตัวแปรเดียว (Univariate search) จะให้ค่าไคสแควร์ต่ำกว่าค่าไคสแควร์เดิมเสมอ

ข. ค่าประมาณค่าเฉลี่ยตัวใหม่ไม่แตกต่างจากตัวประมาณค่าเฉลี่ยตัวเดิม ($\bar{X}$) อย่างมีนัยสำคัญ และค่าประมาณค่าเฉลี่ยตัวใหม่นี้ยังมีค่าใกล้เคียงกับค่าเฉลี่ยของประชากร (๔๐.๖๒๗) มากกว่าค่าประมาณค่าเฉลี่ยตัวเดิม โดยดูจากค่าสถิติทดสอบ t-test ที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยทั้งสองกับค่าเฉลี่ยของประชากร ซึ่งจะพบว่าค่าสถิติ t จากการคำนวณจากค่าเฉลี่ยใหม่เข้าใกล้ศูนย์มากกว่าค่าสถิติ t ที่ได้จากการคำนวณจากค่าเฉลี่ยเดิมเป็นส่วนใหญ่

ค. ค่าความเบี่ยงเบนมาตรฐานที่ได้จากวิธีการค้นหาตัวแปรเดียวจะมีค่ามากกว่าค่าความเบี่ยงเบนมาตรฐานที่ได้จากวิธีเดิม (Conventional Method) เสมอ

๔.๕.๒ ในการวิจัยนี้ทำการทดลองกับตัวอย่างทั้งหมด ๘๐ ชุด ดังที่ได้กล่าวถึงในหัวข้อ ๔.๒.๑ สำหรับผลของตัวอย่างทั้งหมด ๘๐ ชุดนี้ มีแสดงในภาคผนวก ค. โดยใช้วิธีการที่ได้กล่าวไว้แล้วในหัวข้อ ๔.๒.๒ นอกจากจะได้ข้อสังเกตเช่นเดียวกับข้างบนแล้ว มีข้อสังเกตเพิ่มเติมดังนี้ ตามขั้นตอนที่ ๔ หน้า ๖๘ นั้น ผลจากการทดลอง ๘๐ ชุดได้ผลดังนี้

ง. ที่ระดับนัยสำคัญ = 0.01 การทดสอบสมมติฐานเดียวกับลักษณะการแจกแจงของข้อมูล จะให้ผลการยอมรับหรือปฏิเสธสมมติฐานที่ผิดพลาดน้อยกว่าระดับนัยสำคัญ = 0.025 และระดับนัยสำคัญ = 0.05

อย่างไรก็ดีค่าไคสแควร์ที่ได้จากวิธีการค้นหาตัวแปรเดียวที่ระดับนัยสำคัญ = 0.01 ให้ผลการทดสอบไม่มีผิดพลาด ในการทดสอบจากตัวอย่างทั้งหมด ยอมรับว่าข้อมูลมีการแจกแจงแบบปกติ

จ. โดยเฉลี่ยแล้วค่าเฉลี่ยที่ได้จากวิธีเดิม (Conventional Method) จะมีค่า ๔๐.๖๓๔ ซึ่งใกล้เคียงกับค่าเฉลี่ย (๔๐.๖๒๗) ของประชากร

ฉ. โดยเฉลี่ยแล้วค่าเฉลี่ยที่ได้จากวิธีใหม่ (Optimization Method) จะมีค่า ๔๐.๕๙๘ ซึ่งใกล้เคียงกับค่าเฉลี่ย (๔๐.๖๒๗) ของประชากรเช่นเดียวกัน

ช. โดยเฉลี่ยแล้วค่าความเบี่ยงเบนมาตรฐานที่ได้จากวิธีเดิม (Conventional Method) จะมีค่า ๗.๔๒๑ ซึ่งใกล้เคียงกับค่าความเบี่ยงเบนมาตรฐานของประชากรคือ ๗.๔๙๙

ซ. โดยเฉลี่ยแล้วค่าความเบี่ยงเบนมาตรฐานที่ได้จากวิธีใหม่ (Optimization Method) จะมีค่า ๗.๙๕๙ ซึ่งใกล้เคียงกับค่าความเบี่ยงเบนมาตรฐานของประชากรคือ ๗.๔๙๙

๔.๕.๓ ผลจากการวิจัยในครั้งนี้ เราพบว่ายังมีคำถามที่น่าสนใจต่อการค้นคว้าต่อไปในอนาคต เช่น

- บางคนอาจถามว่า ในการกำหนดค่า λ ให้แก่ Univariate Search นั้น เราควรจะเริ่มต้นที่ค่า λ ประมาณเท่าใด? ในวิทยานิพนธ์ฉบับนี้เราได้เลือกค่า $\lambda = 0.01$ ซึ่ง Univariate Search (โปรแกรม UNISER) จะนำค่า $\lambda = 0.01$ ที่เรากำหนดแล้วนี้ไปเปลี่ยนแปร (deviate) ค่า $\bar{X} = 40.627$ และค่า $S = 7.499$ ของโจทย์ เพื่อให้ได้ค่า χ^2 ต่ำสุด ดังนั้นเราก็น่าจะเลือกค่า λ ให้พอประมาณต่อการค้นหาค่า χ^2 ต่ำสุด ให้ได้เร็วที่สุดและได้ accuracy ดีด้วย ในกรณีที่ค่า $\bar{X}$ และ S มี magnitude ต่างกันมาก นั่นคือความสัมพันธ์ของ $\bar{X}$ และ S ต่อ χ^2 แปรปรวนมาก (vary dynamic) Univariate Search จะใช้ไม่ได้ผล โดยทั่วไปเราจึงมักจะหันไปใช้ Optimization Technique ที่มีประสิทธิภาพสูงกว่านี้ ได้แก่ วิธีการของ Powell หรือ Zangwill เป็นต้น

- คำถามที่น่าสนใจอีกคำถามหนึ่งก็คือ ถ้าเรากำหนดจำนวนช่วงของ O_i ให้มากน้อยต่างกัน ย่อมจะมีผลต่อ Optimized χ^2 ด้วย? อย่างไรก็ดีถึงแม้ว่าเราจะได้กำหนดจำนวนช่วงของ O_i ไว้ต่างกันก็ตาม Univariate Search ก็น่าจะให้ค่า Optimized χ^2 ที่มีค่าต่ำกว่า Conventional χ^2 เสมอ

- ตามข้อสังเกตข้อ ๔.๕.๑ (ค) เราได้พบว่า S^* ซึ่งให้ค่า χ^2 ต่ำสุดนั้น <u>จะมีค่าสูงกว่า Conventional S เสมอ</u> จึงเป็นที่น่าสนใจว่า ค่า S^* ที่สูงขึ้นนี้จะมีผลกระทบต่อคุณภาพและประชากรหรือไม่? (เพราะโดยปกติเราเชื่อกันว่าค่าสูง ๆ มักจะไม่ดี)

เนื่องจากวิทยานิพนธ์ฉบับนี้เพียงแต่ได้เสนอแนะวิธีการค้นหาค่า S^* ที่จะให้ค่า χ^2 ต่ำสุดเท่านั้น โดยไม่สนใจคุณภาพของตัวอย่างที่เราศึกษาอยู่ว่าจะมีค่า S^*สูงหรือต่ำเพียงใดหรือไม่ จึงเป็นคำถามที่น่าสนใจสำหรับหัวข้อวิทยานิพนธ์ในฉบับต่อไป

สำหรับผลกระทบต่อคุณภาพของประชากรนั้นเราได้ทดสอบความมั่นใจนั้นด้วย t-test ซึ่งปรากฏผลเป็นที่น่าพอใจ และการที่ S^* ทำให้ได้ค่า χ^{2*} ต่ำสุดนั้น ทำให้ type-I-error ลดลงด้วย