# ขนาดตัวอย่างที่น้อยที่สุดสำหรับการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร

**วราฤทธิ์ พานิชกิจโกศลกุล**[1]

มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต คลองหลวง ปทุมธานี 12121

*รับเมื่อ 18 กรกฎาคม 2548 ตอบรับเมื่อ 30 พฤศจิกายน 2548*

## บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาขนาดตัวอย่างที่น้อยที่สุดที่เหมาะสมในการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากรโดยใช้ตัวสถิติ Z และตัวสถิติ t เมื่อประชากรมีการแจกแจงสมมาตรและการแจกแจงไม่สมมาตร ในการหาขนาดตัวอย่างจะทำการเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง $(1 - \hat{\alpha})$ กับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด $(1 - \alpha)$ โดยกำหนดค่า $1 - \alpha$ เท่ากับ 0.90, 0.95 และ 0.99 ดังนั้นเมื่อ $1 - \hat{\alpha}$ เท่ากับ $1 - \alpha$ อย่างมีนัยสำคัญแล้ว จะส่งผลให้ได้ขนาดตัวอย่างที่เหมาะสม การวิจัยครั้งนี้ใช้วิธีการจำลองแบบมอนติคาร์โล และทำการทดลองซ้ำๆ กัน 10,000 ครั้งในแต่ละสถานการณ์ ผลการวิจัยสรุปได้ดังนี้

เมื่อประชากรมีการแจกแจงสมมาตร ขนาดตัวอย่างที่น้อยที่สุดซึ่งครอบคลุมทุกกรณีที่ศึกษา เท่ากับ $n_1 \geq 17$ และ $n_2 \geq 30$ เมื่อประชากรมีการแจกแจงแบบเบ้ขวา ขนาดตัวอย่างที่น้อยที่สุดซึ่งครอบคลุมทุกกรณีที่ศึกษา เท่ากับ $n_1 \geq 18$ และ $n_2 \geq 40$ และเมื่อประชากรมีการแจกแจงแบบเบ้ซ้าย ขนาดตัวอย่างที่น้อยที่สุดซึ่งครอบคลุมทุกกรณีที่ศึกษา เท่ากับ $n_1 \geq 17$ และ $n_2 \geq 42$

---

[1] *อาจารย์ ภาควิชาคณิตศาสตร์และสถิติ*

# The Least Sample Sizes for Interval Estimation of the Difference between the Means of Two Populations

**Wararit Panichkitkosolkul**[1]
Thammasat University, Rangsit Campus, Klong Laung, Pathum Thani 12121

*Received 18 July 2005 ; accepted 30 November 2005*

## Abstract

The objective of this research is to find the least sample sizes for interval estimation of the difference between the means of two populations using z and t statistics. Populations specified in this research are symmetry and non-symmetry distributions. In order to obtain the sample sizes, confidence coefficients estimation ($1 - \hat{\alpha}$) are compared with the specified confidence coefficients ($1 - \alpha$): 0.90, 0.95 and 0.99. When $1 - \hat{\alpha}$ is significantly equal to $1 - \alpha$, then the sample sizes is appropriate. This research used the Monte Carlo Simulation method. The experiment was repeated 10,000 times for each condition. Results of the study are as follows:

When the population has symmetry distribution, the least sample sizes that cover all cases are $n_1 \geq 17$ and $n_2 \geq 30$ . When the population has positively skewed distribution, the minimum sample sizes that cover all cases are $n_1 \geq 18$ and $n_2 \geq 40$. When the population has negatively skewed distribution, the minimum sample sizes that cover all cases are $n_1 \geq 17$ and $n_2 \geq 42$.

[1] *Lecturer, Department of Mathematics and Statistics*

## 1. บทนำ

ในการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร จะต้องอาศัยตัวประมาณแบบจุดของ $\mu_1-\mu_2$ และการแจกแจงของตัวประมาณนั้น ซึ่งทราบกันดีว่า $\bar{X}_1 - \bar{X}_2$ เป็นตัวประมาณแบบจุดของ $\mu_1-\mu_2$ เมื่อสุ่มตัวอย่างขนาด $n_1$ จากประชากรที่ 1 ซึ่งมีการแจกแจงปกติด้วยค่าเฉลี่ย $\mu_1$ และความแปรปรวน $\sigma_1^2$ และสุ่มตัวอย่างขนาด $n_2$ จากประชากรที่ 2 ซึ่งมีการแจกแจงปกติด้วยค่าเฉลี่ย $\mu_2$ และความแปรปรวน $\sigma_2^2$ โดยที่ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ การแจกแจงของประชากรทั้งสองเป็นอิสระต่อกัน จะได้ว่า $\bar{X}_1-\bar{X}_2$ มีการแจกแจงปกติด้วยค่าเฉลี่ย $E(\bar{X}_1-\bar{X}_2)=\mu_1-\mu_2$ และมีความแปรปรวน $Var(\bar{X}_1-\bar{X}_2)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ สามารถเขียนได้ว่า $\bar{X}_1-\bar{X}_2 \sim N(\mu_1-\mu_2, \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2})$ ดังนั้นช่วงความเชื่อมั่น (1 - $\alpha$)100% ของ $\mu_1-\mu_2$ คือ [1]

$$(\bar{X}_1-\bar{X}_2)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<(\bar{X}_1-\bar{X}_2)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

สำหรับประชากรทั้งสองที่ไม่ได้มีการแจกแจงปกติ เมื่อขนาดตัวอย่าง $n_1$ และ $n_2$ ใหญ่ การประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากรจะกระทำบนพื้นฐานของทฤษฎีบทขีดจำกัดส่วนกลาง (Central limit theorem) เพื่อทำให้ผลต่างระหว่างค่าเฉลี่ยตัวอย่างมีการแจกแจงใกล้เคียงแบบปกติ (Approximate normal) ดังนั้นจึงส่งผลให้ช่วงความเชื่อมั่น (1 - $\alpha$)100% ของ $\mu_1-\mu_2$ สามารถประมาณได้จากสูตรข้างต้น

ในทางปฏิบัตินั้นการทราบค่าความแปรปรวนของประชากรนั้นเป็นเรื่องที่ยาก ดังนั้นสามารถประมาณค่าความแปรปรวนของประชากรทั้งสองได้ด้วยความแปรปรวนตัวอย่าง $S_1^2$ และ $S_2^2$ ดังนั้นกรณีที่ประชากรมีการแจกแจงปกติหรือใกล้เคียง ไม่ทราบความแปรปรวนของประชากร และตัวอย่างขนาดเล็ก จะได้ว่า

$$t'=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$$

มีการแจกแจงใกล้เคียงกับการแจกแจงที โดยมี

องศาเสรี ($\nu$) เท่ากับ $\dfrac{(S_1^2/n_1+S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1-1}+\frac{(S_2^2/n_2)^2}{n_2-1}}$ ดังนั้นช่วงความเชื่อมั่น (1 - $\alpha$)100% ของ $\mu_1-\mu_2$

คือ $$(\bar{X}_1-\bar{X}_2)-t_{\alpha/2,\nu}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}<\mu_1-\mu_2<(\bar{X}_1-\bar{X}_2)+t_{\alpha/2,\nu}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

นอกจากนี้ในบางกรณีที่ไม่ทราบความแปรปรวนของประชากร แต่ทราบเพียงว่าความแปรปรวนของสองประชากรเท่ากัน ($\sigma_1^2 = \sigma_2^2 = \sigma^2$) ดังนั้นจะนำความแปรปรวนตัวอย่าง $S_1^2$ และ $S_2^2$ มาร่วมกันประมาณค่า $\sigma^2$ โดยที่ $S_p^2 = \dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ จะได้ว่า $t = \dfrac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ มีการแจกแจงที โดยมีองศาเสรีเท่ากับ $n_1 + n_2 - 2$ ดังนั้นช่วงความเชื่อมั่น $(1 - \alpha)100\%$ ของ $\mu_1-\mu_2$ คือ [2]

$$(\bar{X}_1-\bar{X}_2)-t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<(\bar{X}_1-\bar{X}_2)+t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

ในทางปฏิบัติการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากรจะต้องพิจารณาถึงปัจจัยหลายอย่าง เช่น ขนาดตัวอย่าง การแจกแจงของประชากร และความแปรปรวนของประชากร เป็นต้น ดังนั้นตามที่ได้กล่าวไว้แล้วข้างต้น ถ้าทราบความแปรปรวนของประชากรโดยทั่วไปจะประมาณผลต่างระหว่างค่าเฉลี่ยของสองประชากรโดยใช้ตัวสถิติ Z แต่ถ้าไม่ทราบความแปรปรวนของประชากรจะประมาณผลต่างระหว่างค่าเฉลี่ยของสองประชากรโดยใช้ตัวสถิติ t

สำหรับการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากรโดยใช้ตัวสถิติ Z และตัวสถิติ t เมื่อพิจารณาถึงรูปแบบการแจกแจงของประชากร ถ้าเข้าใกล้ปกติมาก จะส่งผลให้ขนาดตัวอย่างที่ใช้ในการประมาณมีขนาดเล็ก แต่เมื่อไกลจากปกติมาก จะส่งผลให้ขนาดตัวอย่างที่ใช้ในการประมาณมีขนาดใหญ่ขึ้น ฉะนั้นเมื่อใดจึงควรใช้ตัวสถิติ Z และตัวสถิติ t จึงเป็นเรื่องที่น่าสนใจศึกษา และที่กล่าวว่าขนาดตัวอย่างใหญ่หรือขนาดตัวอย่างเล็กนั้น ขนาดตัวอย่างจำนวนเท่าใดจึงถือว่ามีขนาดใหญ่ หรือเมื่อใดถือว่ามีขนาดเล็กนั้น จึงเป็นสิ่งที่น่าสนใจศึกษา

## 2. ขอบเขตของการวิจัย

ในการวิจัยครั้งนี้ผู้วิจัยสนใจศึกษาขนาดตัวอย่างที่น้อยที่สุดสำหรับการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร มีขอบเขตของการวิจัยดังนี้

**1) กำหนดการแจกแจงของประชากร** ประชากรทั้งสองมีการแจกแจงแบบเดียวกัน มีค่าเฉลี่ยเท่ากับ 100 และความแปรปรวนเท่ากับ 100 โดยมีการแจกแจงดังต่อไปนี้

(1) การแจกแจงสมมาตร (Symmetry distribution)

การแจกแจงสมมาตร มีค่าสัมประสิทธิ์ความเบ้เท่ากับ 0 และค่าสัมประสิทธิ์ความโด่งเท่ากับ 3 เมื่อกำหนดค่าพารามิเตอร์เป็นค่าใดๆ ก็ตาม ค่าสัมประสิทธิ์ความเบ้และค่าสัมประสิทธิ์ความโด่งจะคงที่

(2) การแจกแจงแบบเบ้ขวา (Positively skewed distribution)

การแจกแจงแบบเบ้ขวา กำหนดค่าสัมประสิทธิ์ความเบ้ และค่าสัมประสิทธิ์ความโด่ง ดังนี้

| ความเบ้ | 0.5 | 1.0 | 1.5 | 2.0 |
|---|---|---|---|---|
| | 2.4 | 4.2 | 5.8 | 9.0 |
| ความโด่ง | 3.6 | 5.4 | 7.0 | 12.6 |
| | 6.6 | 8.4 | 10.0 | 15.6 |

(3) การแจกแจงแบบเบ้ซ้าย (Negatively skewed distribution)

การแจกแจงแบบเบ้ซ้าย กำหนดค่าสัมประสิทธิ์ความเบ้ และค่าสัมประสิทธิ์ความโด่ง ดังนี้

| ความเบ้ | - 0.5 | - 1.0 | - 1.5 | - 2.0 |
|---|---|---|---|---|
| | 2.4 | 4.2 | 5.8 | 9.0 |
| ความโด่ง | 3.6 | 5.4 | 7.0 | 12.6 |
| | 6.6 | 8.4 | 10.0 | 15.6 |

**2) กำหนดสูตร** สูตรที่ใช้ในการหาช่วงความเชื่อมั่นของผลต่างระหว่างค่าเฉลี่ยของสองประชากรโดยใช้ตัวสถิติ Z และตัวสถิติ t ในกรณีต่างๆ ดังนี้

- สูตรที่ใช้ในการหาช่วงความเชื่อมั่นโดยใช้ตัวสถิติ Z กรณีทราบความแปรปรวนประชากร

$$(\bar{x}_1-\bar{x}_2)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

- สูตรที่ใช้ในการหาช่วงความเชื่อมั่นโดยใช้ตัวสถิติ Z กรณีไม่ทราบความแปรปรวนประชากร แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$

$$(\bar{x}_1-\bar{x}_2)-z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

- สูตรที่ใช้ในการหาช่วงความเชื่อมั่นโดยใช้ตัวสถิติ t กรณีไม่ทราบความแปรปรวนประชากร แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$

$$(\bar{x}_1-\bar{x}_2)-t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

● สูตรที่ใช้ในการหาช่วงความเชื่อมั่นโดยใช้ตัวสถิติ t กรณีไม่ทราบความแปรปรวนประชากร แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$

$$(\bar{X}_1-\bar{X}_2)-t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<(\bar{X}_1-\bar{X}_2)+t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

กรณีที่ความแปรปรวนของประชากรทั้งสองไม่เท่ากัน ( $\sigma_1^2 \neq \sigma_2^2$ ) จะกำหนดความแตกต่างของอัตราส่วนของความแปรปรวนโดยใช้ค่านอนเซนทรัลลิตี้พารามิเตอร์ $\phi$ (Noncentrality parameter) [3] เป็นเกณฑ์วัดความแตกต่างของความแปรปรวนของประชากร โดยที่

$$\phi = \left( \frac{\sum_{i=1}^{k}\left(\sigma_i^2 - \bar{\sigma}^2\right)^2 / 2}{\sigma_1^2} \right)^{1/2}$$

เมื่อ $\sigma_1^2$ เป็นค่าความแปรปรวนของประชากรที่มีค่าต่ำที่สุด

$\sigma_i^2$ เป็นค่าความแปรปรวนของประชากรที่ i โดย i = 1 , 2

$\bar{\sigma}^2$ เป็นค่าเฉลี่ยความแปรปรวนของประชากรทั้ง 2 กลุ่ม โดยมีรายละเอียดดังนี้

| **ระดับความแตกต่างของความแปรปรวน** | **อัตราส่วนความแปรปรวน** | $\phi$ |
|---|---|---|
| มีความแตกต่างกันน้อย ($0 < \phi < 1.5$) | 1 : 1.5 | 0.25 |
| มีความแตกต่างกันปานกลาง ($1.5 \leq \phi < 3.0$) | 1 : 5 | 2.0 |
| มีความแตกต่างกันมาก ($\phi \geq 3.0$) | 1 : 10 | 4.5 |

3) กำหนดค่าสัมประสิทธิ์ความเชื่อมั่น 1-$\alpha$ เท่ากับ 0.99 0.95 และ 0.90 หรือระดับนัยสำคัญ $\alpha$ เท่ากับ 0.01 0.05 และ 0.10

4) ระดับนัยสำคัญในการเปรียบเทียบความแตกต่างระหว่างค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง (1 - $\hat{\alpha}$) กับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด (1 - $\alpha_0$) เท่ากับ 0.05

5) โปรแกรมที่ใช้ในการวิจัยครั้งนี้ทั้งหมดเขียนด้วยภาษาฟอร์แทรนเพาวเวอร์สเตชัน โดยใช้เครื่องไมโครคอมพิวเตอร์ ทำการทดลองซ้ำ 10,000 ครั้ง ในแต่ละสถานการณ์ของการทดลอง

## 3. วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีขั้นตอนดังต่อไปนี้

### 3.1 การสร้างข้อมูลให้มีการแจกแจงแบบต่างๆ

ในการสร้างข้อมูลให้มีการแจกแจงแบบต่าง ๆ ตามที่กำหนดในขอบเขตของการวิจัย สามารถทำได้โดยการจำลองโดยใช้เทคนิคมอนติคาร์โล

### 3.2 คำนวณช่วงความเชื่อมั่นของผลต่างระหว่างค่าเฉลี่ยของสองประชากร

กำหนดให้ขนาดตัวอย่างของประชากรที่ 1 ($n_1$) เท่ากับ 2 และขนาดตัวอย่างของประชากรที่ 2 ($n_2$) เท่ากับ 2($n_2$) จากนั้นนำข้อมูลตัวอย่างที่สร้างขึ้นโดยวิธีการจำลองแบบมอนติคาร์โลคำนวณช่วงความเชื่อมั่นของผลต่างระหว่างค่าเฉลี่ยของสองประชากรโดยใช้ตัวสถิติ Z และตัวสถิติ t ในกรณีต่างๆ ดังที่กล่าวไว้ในหัวข้อ 2)

### 3.3 คำนวณค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง

นำช่วงความเชื่อมั่นที่คำนวณได้มาพิจารณาว่าครอบคลุมค่าผลต่างของค่าเฉลี่ยของสองประชากรหรือไม่ ถ้าช่วงความเชื่อมั่นใดครอบคลุมค่าผลต่างของค่าเฉลี่ยของสองประชากรจะทำการนับจำนวนครั้งและบวกสะสมค่าไว้ เมื่อทำครบทุกช่วงแล้ว ก็จะนำค่าสะสมที่ได้ของช่วงที่ครอบคลุมค่าผลต่างของค่าเฉลี่ยของสองประชากรมาหารด้วยจำนวนรอบ ซึ่งจะเรียกค่าที่ได้ว่า ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง (1 - $\hat{\alpha}$) สามารถเขียนได้ดังนี้

$$1-\hat{\alpha} = \frac{\text{จำนวนครั้งทั้งหมดที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์ } \mu_1-\mu_2}{10{,}000}$$

### 3.4 เปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองกับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

ทำการเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง (1 - $\hat{\alpha}$) ว่ามีค่าเท่ากับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด (1 - $\alpha_0$) อย่างมีนัยสำคัญหรือไม่ สามารถเขียนสมมุติฐานในการทดสอบได้ดังนี้

$H_0 : P = P_0$

$H_0 : P \neq P_0$

จะปฏิเสธสมมุติฐาน $H_0$ เมื่อ $\dfrac{\hat{P}-P_0}{\sqrt{\dfrac{P_0(1-P_0)}{M}}} \geq Z_{\alpha^*/2}$ หรือ $\dfrac{\hat{P}-P_0}{\sqrt{\dfrac{P_0(1-P_0)}{M}}} \leq -Z_{\alpha^*/2}$

นั่นคือ $\hat{P} \geq P_0 + Z_{\alpha^*/2}\sqrt{\dfrac{P_0(1-P_0)}{M}}$ หรือ $\hat{P} \leq P_0 - Z_{\alpha^*/2}\sqrt{\dfrac{P_0(1-P_0)}{M}}$

เมื่อ P คือ สัมประสิทธิ์ความเชื่อมั่น หรือ P = 1 - $\alpha$

$P_0$ คือ สัมประสิทธิ์ความเชื่อมั่นที่กำหนด หรือ $P_0 = 1 - \alpha_0$

$\hat{P}$ คือ สัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง หรือ $\hat{P} = 1 - \hat{\alpha}$

M คือ จำนวนรอบของการทดลอง ในที่นี้เท่ากับ 10,000 รอบ

$\alpha^*$ คือ ระดับนัยสำคัญของการทดสอบ ในที่นี้กำหนดเท่ากับ 0.05

ดังนั้น เกณฑ์ในการยอมรับขนาดตัวอย่างที่ใช้ในการประมาณค่า คือ

- ถ้า $1 - \alpha_0 = 0.99$ จะยอมรับขนาดตัวอย่างที่ใช้ในการประมาณค่าเมื่อค่า

$$0.98805 \leq 1 - \hat{\alpha} \leq 0.99195$$

- ถ้า $1 - \alpha_0 = 0.95$ จะยอมรับขนาดตัวอย่างที่ใช้ในการประมาณค่าเมื่อค่า

$$0.94573 \leq 1 - \hat{\alpha} \leq 0.95427$$

- ถ้า $1 - \alpha_0 = 0.90$ จะยอมรับขนาดตัวอย่างที่ใช้ในการประมาณค่าเมื่อค่า

$$0.89412 \leq 1 - \hat{\alpha} \leq 0.90588$$

ถ้าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง $(1 - \hat{\alpha})$ มีค่าไม่เท่ากับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด $(1 - \alpha_0)$ อย่างมีนัยสำคัญ ให้เพิ่มขนาดตัวอย่างของประชากรที่ 2 ขึ้นไปอีก 1 หน่วย แล้วย้อนกลับไปทำในข้อ 3.1 จนกระทั่งค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองเท่ากับค่าสัมประสิทธิ์ความเชื่อมั่นที่แท้จริงอย่างมีนัยสำคัญ แต่ถ้าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง $(1 - \hat{\alpha})$ มีค่าเท่ากับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด $(1 - \alpha_0)$ แสดงว่าขนาดตัวอย่างที่ระดับนั้นเหมาะสมสำหรับการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากรโดยใช้ตัวสถิตินั้นๆ เมื่อขนาดตัวอย่างของประชากรที่ 1 $(n_1)$ เท่ากับ 2 จากนั้นทำการเพิ่มขนาดตัวอย่างของประชากรที่ 1 ขึ้นไปอีก 1 หน่วย แล้วย้อนกลับไปทำในข้อ 3.1 จนครบทุกกรณีที่กำหนด

## 4. ผลการวิจัย

ในการศึกษาหาขนาดตัวอย่างที่น้อยที่สุดที่เหมาะสมในการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากรโดยใช้ตัวสถิติ Z และตัวสถิติ t ทั้งกรณีทราบความแปรปรวนของสองประชากร และกรณีไม่ทราบความแปรปรวนของสองประชากร ผลการวิจัยจำแนกตามการแจกแจงของประชากรมีดังนี้

### 4.1 ประชากรมีการแจกแจงสมมาตร

ในกรณีที่ทราบค่าความแปรปรวนของสองประชากร การประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร โดยใช้ตัวสถิติ Z สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 2$ และ $n_2 \geq 3$

ในกรณีที่ไม่ทราบค่าความแปรปรวนของสองประชากร แต่ทราบว่าความแปรปรวนของประชากรทั้งสองไม่เท่ากัน $(\sigma_1^2 \neq \sigma_2^2)$ โดยใช้ตัวสถิติ Z สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 17$ และ $n_2 \geq 30$ และโดยใช้ตัวสถิติ t สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 3$ และ $n_2 \geq 10$

ในกรณีที่ไม่ทราบค่าความแปรปรวนของสองประชากร แต่ทราบว่าความแปรปรวนของประชากรทั้งสองเท่ากัน $(\sigma_1^2 = \sigma_2^2)$ โดยใช้ตัวสถิติ t เมื่อข้อมูลมีค่าความแปรปรวนของประชากรทั้งสองมีค่าเท่ากัน สำหรับ

ขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 2$ และ $n_2 \geq 2$ และเมื่อข้อมูลมีค่าความแปรปรวนของประชากรทั้งสองแตกต่างกันน้อย ปานกลาง และมาก ขนาดตัวอย่างที่น้อยที่สุดที่เหมาะสมคือ ขนาดตัวอย่างของประชากรที่ 1 ควรมีค่าใกล้เคียงกับขนาดตัวอย่างของประชากรที่ 2 ผลการวิจัยแสดงดังตารางที่ 1

**ตารางที่ 1** ขนาดตัวอย่างที่น้อยที่สุด ($n_1$ , $n_2$) ในการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร เมื่อประชากรมีการแจกแจงสมมาตร

| กรณี | ตัวสถิติ | อัตราส่วนความแปรปรวนของประชากรที่ 1 และ 2 | ขนาดตัวอย่างที่น้อยที่สุด |
|---|---|---|---|
| ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ | Z | ทุกอัตราส่วน | ( 2 , 3 ) |
| ไม่ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$ | Z | ทุกอัตราส่วน | ( 17 , 30 ) |
| | t | ทุกอัตราส่วน | ( 3 , 10 ) |
| ไม่ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ แต่ทราบว่า $\sigma_1^2 = \sigma_2^2$ | t | 1 : 1 | ( 2 , 2 ) |
| | | 1 : 1.5 , 1 : 5 และ 1 : 10 | $n_1 \approx n_2$ |

### 4.2 ประชากรมีการแจกแจงแบบเบ้ขวา

ในกรณีที่ทราบค่าความแปรปรวนของสองประชากร การประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร โดยใช้ตัวสถิติ Z สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 4$ และ $n_2 \geq 9$

ในกรณีที่ไม่ทราบค่าความแปรปรวนของสองประชากร แต่ทราบว่าความแปรปรวนของประชากรทั้งสองไม่เท่ากัน ($\sigma_1^2 \neq \sigma_2^2$) โดยใช้ตัวสถิติ Z สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 18$ และ $n_2 \geq 40$ และโดยใช้ตัวสถิติ t สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 3$ และ $n_2 \geq 13$

ในกรณีที่ไม่ทราบค่าความแปรปรวนของสองประชากร แต่ทราบว่าความแปรปรวนของประชากรทั้งสองเท่ากัน ($\sigma_1^2 = \sigma_2^2$) โดยใช้ตัวสถิติ t สรุปได้ดังนี้

1) เมื่อข้อมูลมีค่าความแปรปรวนของประชากรทั้งสองมีค่าเท่ากัน สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 2$ และ $n_2 \geq 6$

2) เมื่อข้อมูลมีค่าความแปรปรวนของประชากรทั้งสองแตกต่างกันน้อย กรณีที่ค่าสัมประสิทธิ์ความเบ้เท่ากับ 0.5 และ 1.0 ขนาดตัวอย่างที่น้อยที่สุดที่เหมาะสมคือ ขนาดตัวอย่างของประชากรที่ 1 ควรมีค่าใกล้เคียงกับขนาดตัวอย่างของประชากรที่ 2 ส่วนกรณีที่ค่าสัมประสิทธิ์ความเบ้เท่ากับ 1.5 และ 2.0 ขนาดตัวอย่างที่น้อยที่สุดที่เหมาะสมคือ ขนาดตัวอย่างของประชากรที่ 2 ควรมีค่ามากกว่าครึ่งหนึ่งของขนาดตัวอย่างของประชากรที่ 1

3) เมื่อข้อมูลมีค่าความแปรปรวนของประชากรทั้งสองแตกต่างกันปานกลางและมาก ขนาดตัวอย่างที่น้อยที่สุดที่เหมาะสมคือ ขนาดตัวอย่างของประชากรที่ 1 ควรมีค่าใกล้เคียงกับขนาดตัวอย่างของประชากรที่ 2 ผลการวิจัยแสดงดังตารางที่ 2

**ตารางที่ 2** ขนาดตัวอย่างที่น้อยที่สุด ($n_1$ , $n_2$) ในการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร เมื่อประชากรมีการแจกแจงแบบเบ้ขวา

| กรณี | ตัวสถิติ | อัตราส่วนความแปรปรวนของประชากรที่ 1 และ 2 | ขนาดตัวอย่างที่น้อยที่สุด |
|---|---|---|---|
| ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ | Z | ทุกอัตราส่วน | ( 4 , 9 ) |
| ไม่ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$ | Z | ทุกอัตราส่วน | ( 18 , 40 ) |
| | t | ทุกอัตราส่วน | ( 3 , 13 ) |
| ไม่ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ แต่ทราบว่า $\sigma_1^2 = \sigma_2^2$ | t | 1 : 1 | ( 2 , 6 ) |
| | | 1 : 1.5 และ $\alpha_3$ = 0.5 , 1.0 | $n_1 \approx n_2$ |
| | | 1 : 1.5 และ $\alpha_3$ = 1.5 , 2.0 | $n_2 > \frac{n_1}{2}$ |
| | | 1 : 5 และ 1 : 10 | $n_1 \approx n_2$ |

### 4.3 ประชากรมีการแจกแจงแบบเบ้ซ้าย

ในกรณีที่ทราบค่าความแปรปรวนของสองประชากร การประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร โดยใช้ตัวสถิติ Z สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 4$ และ $n_2 \geq 16$

ในกรณีที่ไม่ทราบค่าความแปรปรวนของสองประชากร แต่ทราบว่าความแปรปรวนของประชากรทั้งสองไม่เท่ากัน ($\sigma_1^2 \neq \sigma_2^2$) โดยใช้ตัวสถิติ Z สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 17$ และ $n_2 \geq 42$ และโดยใช้ตัวสถิติ t สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 3$ และ $n_2 \geq 10$

ในกรณีที่ไม่ทราบค่าความแปรปรวนของสองประชากร แต่ทราบว่าความแปรปรวนของประชากรทั้งสองเท่ากัน ($\sigma_1^2 = \sigma_2^2$) โดยใช้ตัวสถิติ t สรุปได้ดังนี้

1) เมื่อข้อมูลมีค่าความแปรปรวนของประชากรทั้งสองมีค่าเท่ากัน สำหรับขนาดตัวอย่างที่น้อยที่สุดเท่ากับ $n_1 \geq 3$ และ $n_2 \geq 5$

2) เมื่อข้อมูลมีค่าความแปรปรวนของประชากรทั้งสองแตกต่างกันน้อย กรณีที่ค่าสัมประสิทธิ์ความเบ้เท่ากับ -0.5 และ -1.0 ขนาดตัวอย่างที่น้อยที่สุดที่เหมาะสมคือ ขนาดตัวอย่างของประชากรที่ 1 ควรมีค่าใกล้เคียงกับขนาดตัวอย่างของประชากรที่ 2 ส่วนกรณีที่ค่าสัมประสิทธิ์ความเบ้เท่ากับ -1.5 และ -2.0 ขนาดตัวอย่างที่น้อยที่สุดที่เหมาะสมคือ ขนาดตัวอย่างของประชากรที่ 2 ควรมีค่ามากกว่าครึ่งหนึ่งของขนาดตัวอย่างของประชากรที่ 1

3) เมื่อข้อมูลมีค่าความแปรปรวนของประชากรทั้งสองแตกต่างกันปานกลางและมาก ขนาดตัวอย่างที่น้อยที่สุดที่เหมาะสมคือ ขนาดตัวอย่างของประชากรที่ 1 ควรมีค่าใกล้เคียงกับขนาดตัวอย่างของประชากรที่ 2 ผลการวิจัยแสดงดังตารางที่ 3

**ตารางที่ 3** ขนาดตัวอย่างที่น้อยที่สุด ($n_1$ , $n_2$) ในการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร เมื่อประชากรมีการแจกแจงแบบเบ้ซ้าย

| กรณี | ตัวสถิติ | อัตราส่วนความแปรปรวนของประชากรที่ 1 และ 2 | ขนาดตัวอย่างที่น้อยที่สุด |
|---|---|---|---|
| ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ | Z | ทุกอัตราส่วน | ( 4 , 16 ) |
| ไม่ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$ | Z | ทุกอัตราส่วน | ( 17 , 42 ) |
| | t | ทุกอัตราส่วน | ( 3 , 10 ) |
| ไม่ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ แต่ทราบว่า $\sigma_1^2 = \sigma_2^2$ | t | 1 : 1 | ( 3 , 5 ) |
| | | 1 : 1.5 และ $\alpha_3$ = -0.5 , -1.0 | $n_1 \approx n_2$ |
| | | 1 : 1.5 และ $\alpha_3$ = -1.5 , -2.0 | $n_2 > \frac{n_1}{2}$ |
| | | 1 : 5 และ 1 : 10 | $n_1 \approx n_2$ |

## 5. สรุปผลการวิจัย

ในการหาขนาดตัวอย่างที่น้อยที่สุดสำหรับการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร เมื่อประชากรมีการแจกแจงสมมาตร ขนาดตัวอย่างที่น้อยที่สุดซึ่งครอบคลุมทุกกรณีที่ศึกษา เท่ากับ $n_1 \geq 17$ และ $n_2 \geq 30$ เมื่อประชากรมีการแจกแจงแบบเบ้ขวา ขนาดตัวอย่างที่น้อยที่สุดซึ่งครอบคลุมทุกกรณีที่ศึกษา เท่ากับ $n_1 \geq 18$ และ $n_2 \geq 40$ และเมื่อประชากรมีการแจกแจงแบบเบ้ซ้าย ขนาดตัวอย่างที่น้อยที่สุดซึ่งครอบคลุมทุกกรณีที่ศึกษา เท่ากับ $n_1 \geq 17$ และ $n_2 \geq 42$ ซึ่งจะเห็นได้ว่าผลที่ได้แตกต่างจากตำราสถิติโดยทั่วไป ซึ่งกำหนดว่าในกรณีที่ไม่ทราบค่า $\sigma_1^2$ และ $\sigma_2^2$ แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$ ถ้า $n_1 \geq 30$ และ $n_2 \geq 30$ สามารถประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากรโดยใช้สถิติ Z แต่ถ้า $n_1 < 30$ และ $n_2 < 30$ ใช้สถิติ t ในการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร

ดังนั้นในทางปฏิบัติเมื่อประชากรมีการแจกแจงสมมาตร ขนาดตัวอย่างของประชากรที่ 1 ควรมีค่าไม่ต่ำกว่า 17 หน่วย และขนาดตัวอย่างของประชากรที่ 2 ควรมีค่าไม่ต่ำกว่า 30 หน่วย ส่วนกรณีที่ประชากรมีการแจกแจงแบบเบ้ขวา ขนาดตัวอย่างของประชากรที่ 1 ควรมีค่าไม่ต่ำกว่า 18 หน่วย และขนาดตัวอย่างของประชากรที่ 2 ควรมีค่าไม่ต่ำกว่า 40 หน่วย และกรณีที่ประชากรมีการแจกแจงแบบเบ้ซ้าย ขนาดตัวอย่างของประชากรที่ 1 ควรมีค่าไม่ต่ำกว่า 17 หน่วย และขนาดตัวอย่างของประชากรที่ 2 ควรมีค่าไม่ต่ำกว่า 42 หน่วย

## 6. กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนสนับสนุนจากโครงการวิจัยเสริมหลักสูตร มหาวิทยาลัยธรรมศาสตร์ ประจำปี 2547 ผู้วิจัยจึงใคร่ขอขอบพระคุณมหาวิทยาลัยธรรมศาสตร์ มา ณ โอกาสนี้ด้วย

## 7. เอกสารอ้างอิง

1. Hahn, G.J. and Meeker, W.Q., 1991, *Statistical Intervals : A Guide for Practitioners,* New York, John Wiley & Sons.

2. Dennis, D.W., William, M., and Richard, L.S., 2002, *Mathematical Statistics with Applications,* New York, Duxbury.

3. Games, P.A., Winkler, H.B., and Probert, D.A., 1972, "Robust Tests for Homogeneity of Variance," *Educational and Psychological Measurement*, Vol. 32, pp. 887-909.