บทที่ 5
การประมาณค่า

หลังจากเก็บรวบรวมข้อมูลเรียบร้อยแล้ว การอธิบายลักษณะของข้อมูลจากตัวอย่างจะอธิบายจากค่าสถิติที่เกี่ยวข้อง ซึ่งใช้วิธีการวิเคราะห์ข้อมูลในส่วนของสถิติพรรณนา ทั้งที่จริงแล้วสิ่งที่เราต้องการเป็นการอธิบายลักษณะของข้อมูลในประชากร ซึ่งจะอธิบายจากค่าพารามิเตอร์ที่เกี่ยวข้อง ดังนั้นจึงต้องนำค่าสถิติที่เกี่ยวข้องกลับมาสรุปค่าพารามิเตอร์ที่สนใจ ซึ่งวิธีการในการนำค่าสถิติมาสรุปหรืออ้างอิงค่าพารามิเตอร์นั้นใช้วิธีการวิเคราะห์ข้อมูลในส่วนของสถิติอ้างอิง และถ้าวัตถุประสงค์ต้องการนำค่าสถิติมาสรุปหรืออ้างอิงว่าค่าพารามิเตรอร์มีค่าประมาณเท่าไร วิธีการในการวิเคราะห์ข้อมูลนั้นคือ การประมาณค่า (estimation)

ความหมายของการประมาณค่า

พิจารณาตัวอย่าง 5.1 และ 5.2 เพื่ออธิบายความหมายของคำว่าการประมาณค่า

ตัวอย่าง 5.1 ในการศึกษาเกี่ยวกับน้ำหนักการบรรจุนมกระป๋องของเครื่องจักร A ประชากรที่ศึกษาคือนมกระป๋องที่บรรจุโดยเครื่องจักร A ตัวอย่างคือนมกระป๋องที่บรรจุโดยเครื่องจักร A จำนวน 50 กระป๋อง และข้อมูลที่เก็บรวบรวมคือน้ำหนักการบรรจุนมกระป๋อง (g)

ถ้าต้องการทราบว่าเครื่องจักร A บรรจุนมกระป๋องหนักกระป๋องละกี่กรัม ค่าที่อธิบายได้คือค่าเฉลี่ย หรือ พารามิเตอร์ μ แต่ในทางปฏิบัติทราบเพียงค่าสถิติ $\overline{X}$

ถ้าต้องการทราบว่าเครื่องจักร A บรรจุนมกระป๋องแต่ละกระป๋องหนักแตกต่างกันกี่กรัม ค่าที่อธิบายได้คือค่าส่วนเบี่ยงเบนมาตรฐาน หรือ พารามิเตอร์ σ แต่ในทางปฏิบัติทราบเพียงค่าสถิติ s

ถ้าต้องการทราบว่าเครื่องจักร A บรรจุนมกระป๋องที่มีน้ำหนักน้อยกว่า 10 กรัมร้อยละเท่าใด ค่าที่อธิบายได้คือค่าสัดส่วน หรือ พารามิเตอร์ p แต่ในทางปฏิบัติทราบเพียงค่าสถิติ $\hat{p}$

ตัวอย่าง 5.2 ถ้าต้องการศึกษาเกี่ยวกับน้ำหนักการบรรจุนมกระป๋องของเครื่องจักร A และ B ประชากรที่ศึกษากลุ่มที่ 1 คือนมกระป๋องที่บรรจุโดยเครื่องจักร A ตัวอย่าง คือนมกระป๋องที่บรรจุโดยเครื่องจักร A จำนวน 50 กระป๋อง และข้อมูลที่เก็บรวบรวมคือน้ำหนักการบรรจุนมกระป๋อง (g)

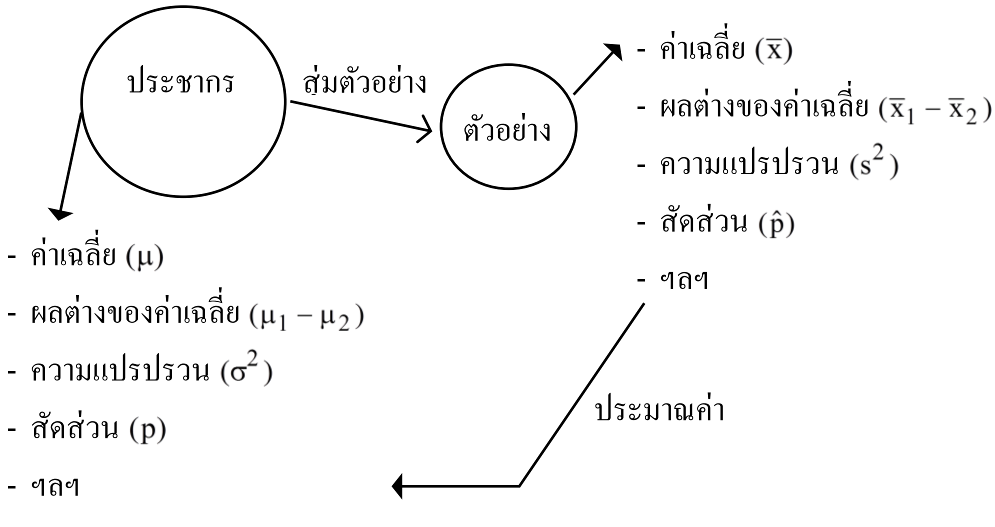
ประชากรที่ศึกษากลุ่มที่ 2 คือนมกระป๋องที่บรรจุโดยเครื่องจักร B ตัวอย่าง คือนมกระป๋องที่บรรจุโดยเครื่องจักร B จำนวน 80 กระป๋อง และข้อมูลที่เก็บรวบรวมคือน้ำหนักการบรรจุนมกระป๋อง (g)

ถ้าต้องการทราบว่าเครื่องจักร A และ B บรรจุนมกระป๋องมีน้ำหนักหนักต่างกันกระป๋องละกี่กรัม ค่าที่อธิบายได้คือผลต่างของค่าเฉลี่ย หรือ พารามิเตอร์ $\mu_1 - \mu_2$ แต่ในทางปฏิบัติทราบเพียงค่าสถิติ $\bar{x}_1 - \bar{x}_2$

ถ้าต้องการทราบว่าความแปรปรวนของน้ำหนักนมกระป๋องที่บรรจุโดยเครื่องจักร A และ B มีค่าต่างกันเท่าใด ค่าที่อธิบายได้คืออัตราส่วนของความแปรปรวน หรือ พารามิเตอร์ $\frac{\sigma_1^2}{\sigma_2^2}$ แต่ในทางปฏิบัติทราบเพียงค่าสถิติ $\frac{s_1^2}{s_2^2}$

ถ้าต้องการทราบว่าเครื่องจักร A และ B ที่บรรจุนมกระป๋องที่มีน้ำหนักน้อยกว่า 10 กรัมต่างกันร้อยละเท่าใด ค่าที่อธิบายได้คือค่าผลต่างของสัดส่วน หรือ พารามิเตอร์ $p_1 - p_2$ แต่ในทางปฏิบัติทราบเพียงค่าสถิติ $\hat{p}_1 - \hat{p}_2$

จากตัวอย่าง 5.1 และ 5.2 จะเห็นว่าค่าพารามิเตอร์ที่สามารถอธิบายและตอบคำถามในประชากรที่ศึกษามีมากมาย ขึ้นอยู่กับว่าเรามีคำถามอะไร และเมื่อต้องการทราบค่าพารามิเตอร์ใดในทางปฏิบัติเราจะหาค่าสถิติที่เกี่ยวข้อง และนำมาประมาณค่าพารามิเตอร์นั้น ดังนี้



จากรูปข้างต้นสามารถสรุปได้ว่า

การประมาณค่า หมายถึงการนำค่าสถิติกลับมาสรุป หรืออ้างอิงเกี่ยวกับค่าพารามิเตอร์ที่สนใจว่ามีค่าเท่ากับเท่าไร หรืออาจกล่าวได้ว่าการประมาณค่าก็คือการสรุปค่าพารามิเตอร์ที่สนใจโดยอาศัยค่าจริงจากตัวอย่าง หรือที่เรียกว่าค่าสถิตินั่นเอง

ประเภทของการประมาณค่า

ในการประมาณค่าพารามิเตอร์ต่าง ๆ นั้นจะมีหลักการ และวิธีการต่าง ๆ ที่เหมือนกัน จะแตกต่างกันที่สูตรในการประมาณค่าซึ่งขึ้นอยู่กับการแจกแจงของค่าสถิติ ดังนั้นในส่วนนี้เราจะศึกษาเกี่ยวกับหลักการและวิธีการในการประมาณค่าพารามิเตอร์ซึ่งมี 2 วิธี คือ

1. การประมาณค่าแบบจุด (point estimation)
2. การประมาณค่าแบบช่วง (interval estimation)

1. การประมาณค่าแบบจุด

การประมาณค่าแบบจุดเป็นการประมาณค่าพารามิเตอร์ที่สนใจด้วยค่าสถิติเพียงค่าเดียวเท่านั้น นั่นคือ

พารามิเตอร์ μ จะมีค่าประมาณเท่ากับค่าสถิติ $\bar{x}$
พารามิเตอร์ σ^2 จะมีค่าประมาณเท่ากับค่าสถิติ s^2
พารามิเตอร์ σ จะมีค่าประมาณเท่ากับค่าสถิติ s
พารามิเตอร์ p จะมีค่าประมาณเท่ากับค่าสถิติ $\hat{p}$
พารามิเตอร์ $\mu_1 - \mu_2$ จะมีค่าประมาณเท่ากับค่าสถิติ $\bar{x}_1 - \bar{x}_2$
พารามิเตอร์ $p_1 - p_2$ จะมีค่าประมาณเท่ากับค่าสถิติ $\hat{p}_1 - \hat{p}_2$ เป็นต้น

และเรียก $\bar{x}$ เป็นตัวประมาณค่าแบบจุดของ μ
s^2 เป็นตัวประมาณค่าแบบจุดของ σ^2
s เป็นตัวประมาณค่าแบบจุดของ σ
$\hat{p}$ เป็นตัวประมาณค่าแบบจุดของ p
$\bar{x}_1 - \bar{x}_2$ เป็นตัวประมาณค่าแบบจุดของ $\mu_1 - \mu_2$
$\hat{p}_1 - \hat{p}_2$ เป็นตัวประมาณค่าแบบจุดของ $p_1 - p_2$ เป็นต้น

จะเห็นว่าการประมาณค่าวิธีนี้ ประมาณค่าพารามิเตอร์ให้เท่ากับค่าสถิติที่คำนวณจากตัวอย่างกลุ่มหนึ่ง ดังนั้นถ้ากลุ่มตัวอย่างเปลี่ยนไปค่าสถิติจะเปลี่ยนไปด้วย การประมาณค่าแบบจุดนี้จึงเป็นการประมาณค่าพารามิเตอร์ที่มีโอกาสเกิดความคลาดเคลื่อนในการประมาณค่ามาก เพราะเป็นไปได้ยากที่ค่าพารามิเตอร์จะมีค่าเท่ากับค่าสถิติ เช่น ผู้ว่าราชการจังหวัดต้องการทราบราคาน้ำมันในจังหวัดนครปฐมว่าในขณะนี้ราคาประมาณกี่บาท จึงให้ลูกน้องออกไปสำรวจข้อมูลมา ลูกน้องออกไปสำรวจราคาน้ำมันของปั๊มน้ำมัน 10 แห่ง แล้วหาราคาน้ำมันเฉลี่ย ปรากฏว่าได้ 32.30 บาท จึงนำข้อมูลนี้ไปเสนอผู้ว่าราชการจังหวัดว่า ในขณะนี้ราคาน้ำมันของจังหวัดนครปฐมประมาณ 32.30 บาท ซึ่งในความเป็นจริงนั้นเป็นไปได้ยากที่ราคาน้ำมันจะมีค่าเท่ากับ 32.30 บาท จะเห็นได้ว่าการประมาณค่าแบบนี้มีโอกาสเกิดความคลาดเคลื่อนมาก จึงควรใช้กับข้อมูลที่มีการกระจายน้อย เพื่อให้การประมาณค่ามีความคลาดเคลื่อนน้อยลงจึงมีการประมาณค่าอีกแบบ ดังนี้

2. การประมาณค่าแบบช่วง

การประมาณค่าแบบช่วงเป็นการประมาณค่าพารามิเตอร์ด้วยช่วงของจำนวน 2 จำนวนซึ่งมีตัวประมาณค่าแบบจุดอยู่กึ่งกลางระหว่างจำนวน 2 จำนวนนั้น

ตัวอย่าง 5.3 จากการประมาณค่าราคาน้ำมันแบบจุดได้เท่ากับ 32.30 บาท ถ้าใช้การประมาณค่าแบบช่วงอาจนำเสนอข้อมูลให้ผู้ว่าราชการจังหวัดทราบว่า ราคาน้ำมันของจังหวัดนครปฐมมีค่าตั้งแต่ 31.30 ถึง 33.30 บาท ซึ่งคำนวณได้จาก

$$31.30 \xleftarrow[32.30-1.00]{} 32.30 \xrightarrow{32.30+1.00} 33.30 = 32.30 \pm 1.00 \text{ บาท}$$

จะเห็นว่าถ้าใช้การประมาณค่าแบบช่วงโอกาสที่ราคาน้ำมันจะอยู่ในช่วง 11.30 ถึง 13.30 บาทมีมากกว่าการประมาณด้วยตัวประมาณแบบจุด จึงทำให้ความคลาดเคลื่อนในการประมาณค่าน้อยลง

จากตัวอย่าง 5.1 สามารถเขียนช่วงในการประมาณค่าพารามิเตอร์ที่สนใจได้ดังนี้

$$L \le \mu \le U$$
$$L \le \sigma^2 \le U$$
$$L \le \sigma \le U$$
$$L \le p \le U$$
$$L \le \mu_1 - \mu_2 \le U$$
$$L \le \hat{p}_1 - \hat{p}_2 \le U$$

เมื่อ L คือ ค่าต่ำสุด และ U คือค่าสูงสุด ถ้าค่าของ L และ U แตกต่างกันมากก็หมายความว่าพารามิเตอร์ที่สนใจมีค่าที่เป็นไปได้มากในช่วง L ถึง U และถ้าค่าของ L และ U แตกต่างกันน้อยก็หมายความว่าพารามิเตอร์ที่สนใจมีค่าที่เป็นไปได้น้อยในช่วง L ถึง U ซึ่งค่า L และ U จะแตกต่างกันมากหรือน้อยขึ้นอยู่กับค่า ๆ หนึ่งที่กำหนดซึ่งเรียกว่าระดับความเชื่อมั่น (confidence level)

ระดับความเชื่อมั่น

ระดับความเชื่อมั่น หมายถึง ความน่าจะเป็นที่พารามิเตอร์จะมีค่าอยู่ในช่วงประมาณที่สร้างขึ้น เขียนแทนด้วยสัญลักษณ์ $1-\alpha$ นั่นคือ

$$P(L \le parameter \le U) = 1-\alpha$$

เช่น ถ้ากำหนดระดับความเชื่อมั่น 95% หรือ $1-\alpha = 0.95$ หมายความว่าถ้าทำการประมาณค่าพารามิเตอร์โดยการสุ่มตัวอย่าง 100 กลุ่ม และสร้างช่วงในการประมาณค่าพารามิเตอร์ 100 ช่วง จะมีค่าพารามิเตอร์อยู่ในช่วงที่สร้างขึ้น 95 ช่วง หรือโอกาสที่พารามิเตอร์จะมีค่าอยู่ในช่วงประมาณที่สร้างขึ้นเท่ากับ 95%

การประมาณค่าของพารามิเตอร์ที่ใช้กันทั่วไปในการวิเคราะห์ข้อมูลคือ การประมาณค่าเฉลี่ย ผลต่างของค่าเฉลี่ย สัดส่วน และผลต่างของสัดส่วน เป็นต้น ซึ่งจะแตกต่างกันที่สูตรในการคำนวณหาค่า L และ U

การประมาณค่าพารามิเตอร์ค่าเฉลี่ยของประชากร 1 กลุ่ม (μ)

1. การประมาณค่าแบบจุดของค่าเฉลี่ย

ตัวประมาณแบบจุดของ μ คือ $\bar{x} = \dfrac{\sum x}{n}$

2. การประมาณค่าแบบช่วงของค่าเฉลี่ย

ที่ระดับความเชื่อมั่น 100(1-α)% ช่วงในการประมาณค่า μ คือ

$$\bar{x} - t_{\alpha/2\,,\,n-1}\, {}^{s}\!/\!_{\sqrt{n}} \le \mu \le \bar{x} + t_{\alpha/2\,,\,n-1}\, {}^{s}\!/\!_{\sqrt{n}}$$

เมื่อ $\bar{x}$ แทนค่าเฉลี่ย

s แทนส่วนเบี่ยงเบนมาตรฐาน

$t_{\alpha/2,n-1}$ แทนค่าจากตารางการแจกแจงแบบที ขึ้นอยู่กับค่า $1-\alpha$

n แทนจำนวนตัวอย่าง

ในบางครั้งอาจจะเขียนช่วงความเชื่อมั่นของ μ ในรูปของ

$$\bar{x} \pm t_{\alpha/2\,,\,n-1} \, {}^{s}\!/\!_{\sqrt{n}} \quad \text{หรือ} \quad \left[\bar{x} - t_{\alpha/2\,,\,n-1} \, {}^{s}\!/\!_{\sqrt{n}} \,, \bar{x} + t_{\alpha/2\,,\,n-1} \, {}^{s}\!/\!_{\sqrt{n}}\right]$$

ตัวอย่าง 5.4 สำนักงานสถิติเก็บข้อมูลเกี่ยวกับอายุของผู้ใช้แรงงานในจังหวัดนครปฐมเพื่อใช้เป็นข้อมูลในการกำหนดอัตราค่าแรงต่อไป จึงทำการสุ่มตัวอย่างผู้ที่ใช้แรงงานมา 50 คน เพื่อหาอายุเฉลี่ย ปรากฏข้อมูลดังนี้

45	38	22	58	40	42	43	32	34	19
19	21	33	16	49	29	30	43	37	62
65	57	60	41	28	35	37	51	37	26
43	26	27	31	33	24	34	28	39	38
33	27	42	40	31	34	38	35	29	33

จงหาช่วงความเชื่อมั่น 90 % ของอายุเฉลี่ยของผู้ใช้แรงงาน

วิธีทำ จากข้อมูลค่าประมาณแบบจุดของ μ คือ $\bar{x}$ และค่าประมาณแบบจุดของ σ^2 คือ s^2

$$\begin{aligned}
\bar{x} &= \frac{\sum x}{n} \\
&= \frac{45+38+22+\ldots+33}{50} \\
&= 36.28 \\
s &= \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}} \\
&= \sqrt{\frac{(45-36.28)^2 + (38-36.28)^2 + \ldots + (33-36.28)^2}{49}} \\
&= 11.13
\end{aligned}$$

กำหนด μ แทนอายุเฉลี่ยของผู้ใช้แรงงานในจังหวัดนครปฐม

ดังนั้นช่วงความเชื่อมั่น 90 % ของอายุเฉลี่ยของผู้ใช้แรงงาน คือ

$$\bar{x} \pm z_{\alpha/2}\, {}^{s}\!/\!_{\sqrt{n}} \;=\; 36.28 \pm (1.645)({}^{11.13}\!/\!_{\sqrt{50}})$$
$$= \; 36.38 \pm 2.589$$

หรือ $33.691 \le \mu \le 38.869$ นั่นคือที่ระดับความเชื่อมั่น 90 % อายุเฉลี่ยของผู้ใช้แรงงานในจังหวัดนครปฐมอยู่ในช่วง 33.691 ปี ถึง 38.869 ปี

หมายเหตุ จะเห็นว่าที่ระดับความเชื่อมั่นสูงขึ้น ช่วงในการประมาณค่าพารามิเตอร์จะกว้างขึ้น นั่นหมายถึงค่าพารามิเตอร์ที่เป็นไปได้จะมากขึ้นด้วย การนำผลการประมาณค่าไปใช้จะต้องพิจารณาลำบากขึ้นด้วย และถ้าระดับความเชื่อมั่นน้อยลง ช่วงในการประมาณค่าพารามิเตอร์จะแคบลง นั่นหมายถึงค่าพารามิเตอร์ที่เป็นไปได้น้อยลง การนำผลการประมาณค่าไปใช้ก็พิจารณาได้ง่ายขึ้น ดังนั้นการกำหนดระดับความเชื่อมั่นในการประมาณค่าจึงขึ้นอยู่กับผู้ประมาณค่า หรือผู้ใช้ผลการประมาณค่าเป็นหลัก

การประมาณค่าพารามิเตอร์ผลต่างของค่าเฉลี่ยของประชากร 2 กลุ่ม $(\mu_1 - \mu_2)$

1. การประมาณค่าแบบจุดของผลต่างของค่าเฉลี่ย $\mu_1 - \mu_2$ คือ $\bar{x}_1 - \bar{x}_2$

เมื่อ $\bar{x}_1 = \dfrac{\sum_{i=1}^{n_1} x_{1i}}{n_1}$ และ $\bar{x}_2 = \dfrac{\sum_{i=1}^{n_2} x_{2i}}{n_2}$

2. การประมาณค่าแบบช่วงของผลต่างของค่าเฉลี่ย

แบ่งออกเป็น 3 กรณี คือ

1. ประชากร 2 กลุ่มเป็นอิสระกัน
2. ประชากร 2 กลุ่มไม่เป็นอิสระกัน

2.1 ประชากร 2 กลุ่มเป็นอิสระกัน ในกรณีที่ไม่ทราบ σ^2 และตัวอย่างมีขนาดเล็กถูกสุ่มมาจากประชากรที่ประมาณได้ด้วยการแจกแจงแบบปกติ การแจกแจงค่าจากตัวอย่างของ $\bar{X}_1 - \bar{X}_2$ จะมีการแจกแจงแบบ t ด้วยองศาแห่งความเป็นอิสระเท่ากับ n-1

2.1.1 ในกรณีที่ไม่ทราบ σ^2 แต่ทราบว่า $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 1

ที่ระดับความเชื่อมั่น100(1-α)% ช่วงในการประมาณค่า $\mu_1 - \mu_2$ คือ

$$(\bar{x}_1 - \bar{x}_2) - t_{(\alpha/2, n_1+n_2-2)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu \le (\bar{x}_1 - \bar{x}_2) + t_{(\alpha/2, n_1+n_2-2)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

เมื่อ $\bar{x}_1$ แทนค่าเฉลี่ยของตัวอย่างกลุ่มที่ 1

$\bar{x}_2$ แทนค่าเฉลี่ยของตัวอย่างกลุ่มที่ 2

$t_{\alpha/2, n_1+n_2-2}$ แทนค่าจากตารางการแจกแจงแบบที ขึ้นอยู่กับค่า $1-\alpha$

n_1 แทนจำนวนตัวอย่างกลุ่มที่ 1

n_2 แทนจำนวนตัวอย่างกลุ่มที่ 2

s_p แทนส่วนเบี่ยงเบนมาตรฐานร่วม และ $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$

ในบางครั้งอาจจะเขียนช่วงความเชื่อมั่นของ $\mu_1 - \mu_2$ ในรูปของ

$$(\bar{x}_1 - \bar{x}_2) \pm t_{(\alpha/2, n_1+n_2-2)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$ หรือ

$$\left[(\bar{x}_1 - \bar{x}_2) - t_{(\alpha/2, n_1+n_2-2)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \;,\; (\bar{x}_1 - \bar{x}_2) + t_{(\alpha/2, n_1+n_2-2)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right]$$

2.2.2 ในกรณีที่ไม่ทราบ σ^2 แต่ทราบว่า $\sigma_1^2 \ne \sigma_2^2$

ที่ระดับความเชื่อมั่น100(1-α)% ช่วงในการประมาณค่า $\mu_1 - \mu_2$ คือ

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

เมื่อ $\bar{x}_1$ แทนค่าเฉลี่ยของตัวอย่างกลุ่มที่ 1

$\bar{x}_2$ แทนค่าเฉลี่ยของตัวอย่างกลุ่มที่ 2

s_1^2 แทนส่วนเบี่ยงเบนมาตรฐานของตัวอย่างกลุ่มที่ 1

s_2^2 แทนส่วนเบี่ยงเบนมาตรฐานของตัวอย่างกลุ่มที่ 2

$t_{\alpha/2}, \nu$ แทนค่าจากตารางการแจกแจงแบบปกติมาตรฐาน ขึ้นอยู่กับค่า $1-\alpha$

n_1 แทนจำนวนตัวอย่างกลุ่มที่ 1

n_2 แทนจำนวนตัวอย่างกลุ่มที่ 2

ในบางครั้งอาจจะเขียนช่วงความเชื่อมั่นของ $\mu_1 - \mu_2$ ในรูปของ

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

$$\left[(\bar{x}_1 - \bar{x}_2) - t_{(\alpha/2,\nu)}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\ ,(\bar{x}_1 - \bar{x}_2) + t_{(\alpha/2\,,\nu)}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right]$$

ตัวอย่าง 5.5 ในการศึกษาวิธีการสอน 2 แบบกับนักศึกษามหาวิทยาลัยราชภัฏนครปฐมจึงทำการทดลองสอนกับนักศึกษา 2 กลุ่ม โดยกลุ่มแรกใช้วิธีการสอน A กับนักศึกษาจำนวน 12 คน และกลุ่มที่ 2 ใช้วิธีการสอน B กับนักศึกษาจำนวน 10 คน จากการให้ทดสอบด้วยข้อสอบเดียวกันปรากฏว่า กลุ่มแรกได้คะแนนเฉลี่ย 85 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 4 คะแนน ส่วนกลุ่มที่ 2 ได้คะแนนเฉลี่ย 81 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 5 คะแนน จงหาช่วงความเชื่อมั่น 90% ในการประมาณผลต่างระหว่างคะแนนเฉลี่ยของนักศึกษามหาวิทยาลัยราชภัฏนครปฐมทั้ง 2 กลุ่ม โดยสมมติว่าประชากรทั้งสองกลุ่มมีการแจกแจงแบบปกติด้วยความแปรปรวนเท่ากัน

วิธีทำ ให้ μ_1 เป็นคะแนนเฉลี่ยของนักศึกษาทั้งหมดที่ใช้วิธีการสอน A

μ_2 เป็นคะแนนเฉลี่ยของนักศึกษาทั้งหมดที่ใช้วิธีการสอน B

$$\bar{x}_1 = 85 \quad \bar{x}_2 = 81 \quad s_1 = 4 \quad s_2 = 5 \quad n_1 = 12 \quad n_2 = 10$$

ในที่นี้ $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ดังนั้นจะประมาณ σ^2 ด้วย s_p^2 โดยที่

$$\begin{aligned} s_p^2 &= \frac{(n_1-1)s_1^2 + (n_2-1)s_1^2}{n_1+n_2-2} \\ &= \frac{(11)(16)+(9)(25)}{12+10-2} \\ &= 20.05 \\ s_p &= 4.478 \end{aligned}$$

ช่วงความเชื่อมั่น 90% หมายความว่า $\alpha = 0.1 \;\therefore\; t_{(\alpha/2,n_1+n_2-2)} = t_{(0.05,20)} = 1.725$

ดังนั้นช่วงความเชื่อมั่น 90%ของผลต่างระหว่างคะแนนเฉลี่ยของนักศึกษามหาวิทยาลัยราชภัฏนครปฐมทั้ง 2 กลุ่ม คือ

$$\begin{aligned} (\bar{x}_1 - \bar{x}_2) \pm t_{(\alpha/2,n_1+n_2-2)}\, s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} &= (85-81) \pm (1.725)(4.478)\sqrt{\frac{1}{12}+\frac{1}{10}} \\ &= 4 \pm (7.72)(0.49) \\ &= 4 \pm 3.78 \end{aligned}$$

นั่นคือ $0.22 < \mu < 7.78$ หมายความว่าผลต่างระหว่างคะแนนเฉลี่ยของนักศึกษามหาวิทยาลัยราชภัฏนครปฐมที่ใช้วิธีการสอน A กับ B อยู่ในช่วง 0.22 คะแนน กับ 7.78 คะแนน ที่ระดับความเชื่อมั่น 90%

ตัวอย่าง 5.6 จากตัวอย่าง 5.5 ถ้าข้อสมมติที่ว่าความแปรปรวนของประชากรทั้ง 2 กลุ่มเท่ากัน ไม่เป็นจริงแล้ว จงหาช่วงความเชื่อมั่น 95% ของ $\mu_1 - \mu_2$

องศาแห่งความเป็นอิสระ คือ

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

$$= \frac{(16/12 + 10)^2}{\frac{(16/12)^2}{12-1} + \frac{(25/10)^2}{9-1}}$$

$$= 14.96 \approx 15$$

ดังนั้น $t_{(\alpha/2,\nu)} = t_{(0.025,15)} = 2.131$

จะได้ช่วงความเชื่อมั่น 95% ของ $\mu_1 - \mu_2$ คือ

$$\bar{x}_1 - \bar{x}_2 \pm t_{(\alpha/2\ ,\ \nu)}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (85-81) \pm (2.131)\sqrt{\frac{16}{12} + \frac{25}{9}}$$

$$= 4 \pm (2.131)(2.03)$$

$$= 4 \pm 4.33$$

หมายความว่าผลต่างระหว่างคะแนนเฉลี่ยของนักศึกษามหาวิทยาลัยราชภัฏนครปฐมที่ใช้วิธีการสอน A กับ B อยู่ในช่วง –0.33 คะแนน กับ 8.33 คะแนน ที่ระดับความเชื่อมั่น 95 %

2.2 ประชากร 2 กลุ่มไม่เป็นอิสระกัน ในหัวข้อที่ผ่านมาเราได้กล่าวถึงการประมาณค่าแบบช่วงของผลต่างของค่าเฉลี่ยของประชากร 2 กลุ่มที่เป็นอิสระกัน นั้นคือตัวอย่างที่สุ่มมาจากประชากรกลุ่มที่ 1 เป็นอิสระจากตัวอย่างที่สุ่มมาจากประชากรกลุ่มที่ 2 ซึ่งเมื่อรวมขนาดของตัวอย่างแล้วจะมีขนาด $n_1 + n_2$ แต่ในการทดลองบางประเภท เช่นการควบคุมอาหารให้ถูกวิธีจะช่วยลดน้ำหนักได้อย่างมีประสิทธิภาพเพียงใด โดยเปรียบเทียบน้ำหนักก่อน และหลังควบคุม หรือการเปรียบเทียบผลการเรียนของนักศึกษาในปีสุดท้ายก่อนที่จะจบ กับปีแรกที่เริ่มเข้าเป็นนักศึกษาใหม่ว่าแตกต่างกันหรือไม่ เป็นต้น การทดลองเหล่านี้จะใช้ขนาดตัวอย่างเพียงแค่ n เท่านั้น

สำหรับการทดลอง ซึ่งแต่ละตัวอย่างจะมีข้อมูล 2 ค่า เช่น น้ำหนักก่อนการควบคุม และน้ำหนักหลังการควบคุม ผลการเรียนของนักศึกษาในปีสุดท้ายก่อนที่จะจบ และปีแรกที่เริ่มเข้าเป็นนักศึกษาใหม่

ข้อมูลในลักษณะเช่นนี้จะเป็นข้อมูลของตัวอย่างที่มาจากประชากร 2 กลุ่มที่ไม่เป็นอิสระกัน เรียกว่า ข้อมูลคู่ (pair data) การประมาณค่าผลต่างของค่าเฉลี่ยของประชากร 2 กลุ่มในลักษณะนี้เรียกว่าการประมาณค่าเฉลี่ยของผลต่างระหว่าง 2 ประชากรโดยวิธีการจับคู่

กำหนดให้ x_{1i} และ x_{2i} เป็นข้อมูลของตัวอย่างกลุ่มที่ 1 (ก่อน) และกลุ่มที่ 2 (หลัง) และ $d_i = x_{1i} - x_{2i}$ เป็นผลต่างของข้อมูลแต่ละคู่ เมื่อ $i = 1,\ldots, n$ ลักษณะของข้อมูลเป็นดังนี้

กลุ่มที่ 1 (ก่อน) x_{1i}	**กลุ่มที่ 2 (หลัง)** x_{2i}	**ผลต่าง** $d_i = x_{1i} - x_{2i}$
x_{11}	x_{21}	$d_1 = x_{11} - x_{21}$
x_{12}	x_{22}	$d_2 = x_{12} - x_{22}$
x_{13}	x_{23}	$d_3 = x_{13} - x_{23}$
$\vdots$	$\vdots$	$\vdots$
x_{1n}	x_{2n}	$d_n = x_{1n} - x_{2n}$

ที่ระดับความเชื่อมั่น100(1-α)% ช่วงในการประมาณค่า μ_d คือ

$$\bar{d} - t_{\alpha/2\,,\,n-1}\, {s_d}/{\sqrt{n}} \le \mu_d \le \bar{d} + t_{\alpha/2\,,\,n-1}\, {s_d}/{\sqrt{n}}$$

เมื่อ $\bar{d}$ แทนค่าเฉลี่ยของผลต่าง และ $\bar{d} = \dfrac{\sum_{i=1}^{n} d_i}{n}$

s_d แทนส่วนเบี่ยงเบนมาตรฐานของผลต่าง และ $s_d = \sqrt{\dfrac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n-1}}$

$t_{\alpha/2, n-1}$ แทนค่าจากตารางการแจกแจงแบบที ขึ้นอยู่กับค่า $1-\alpha$

n แทนจำนวนคู่ตัวอย่าง

ในบางครั้งอาจจะเขียนช่วงความเชื่อมั่นของ μ_d ในรูปของ

$$\bar{d} \pm t_{\frac{\alpha}{2}\,,\,n-1} \frac{s_d}{\sqrt{n}} \text{ หรือ } \left[\bar{d} - t_{\alpha/2\,,\,n-1}\, {s_d}/{\sqrt{n}}\ ,\ \bar{d} + t_{\alpha/2\,,\,n-1}\, {s_d}/{\sqrt{n}}\right]$$

ตัวอย่าง 5.7 นักจิตวิทยาผู้หนึ่งต้องการเปรียบเทียบวิธีการวัดระดับ IQ 2 วิธี จึงทำการสุ่มตัวอย่างเด็กมา 8 คน แล้วทดสอบวัด IQ ด้วยวิธีการวัดแบบที่ 1 เว้นระยะเวลาพอประมาณแล้วทดสอบวัด IQ ด้วยวิธีการวัดแบบที่ 2 กับเด็กกลุ่มเดิม ได้ผลดังตารางด้านล่าง จงประมาณค่าเฉลี่ยของความแตกต่างของระดับ IQ ที่ระดับความเชื่อมั่น 99 %

คนที่	ระดับ IQ		ผลต่าง
	วิธีที่ 1	วิธีที่ 2	
1	77	72	5
2	74	68	6
3	82	76	6
4	73	68	5
5	87	84	3
6	69	68	1
7	66	61	5
8	80	76	4
			$\sum d = 35$

$$\bar{d} = \frac{\sum d_i}{n} = \frac{35}{8} = 4.375$$

$$s_{\bar{d}}^2 = \frac{\sum_{i=1}^{n}(d_i - \bar{d})^2}{n-1} = \frac{(5-4.375)^2 + (6-4.375)^2 + \ldots + (4-4.375)^2}{7} = 2.839$$

วิธีทำ กำหนด μ_d แทนค่าเฉลี่ยของความแตกต่างของระดับ IQ

ดังนั้นช่วงความเชื่อมั่น 99% ของ μ_d คือ

$$\bar{d} \pm t_{\frac{\alpha}{2}, n-1} \frac{s_d}{\sqrt{n}} = 4.375 \pm t_{0.005,7} \frac{1.685}{\sqrt{8}}$$

$$= 4.375 \pm (3.499)(0.596)$$

$$= 4.375 \pm 2.085$$

ดังนั้น $2.29 < \mu_d < 6.46$ หมายความว่าที่ระดับความเชื่อมั่น 99 % วิธีการวัดระดับ IQ วิธีที่ 1 วัดระดับ IQ สูงกว่าวิธีที่ 2 อยู่ในช่วง 2.29 ถึง 6.46 คะแนน

การใช้โปรแกรมสำเร็จรูป SPSS

1. การประมาณค่าเฉลี่ยของประชากร 1 กลุ่ม

โปรแกรม SPSS สามารถประมาณค่าพารามิเตอร์ค่าเฉลี่ยของประชากร 1 กลุ่ม ด้วยคำสั่ง Analyze / Descriptive Statistics / Explore และคำสั่ง Analyze / Compare Means / One-Sample T Test

ตัวอย่าง 5.8 จงประมาณค่าเฉลี่ยของตัวแปร x (อายุการใช้งานของหลอดไฟฟ้า) ในแฟ้มข้อมูล exsample6.sav ที่ระดับความเชื่อมั่น 95% ด้วยคำสั่ง Analyze / Descriptive Statistics / Explore

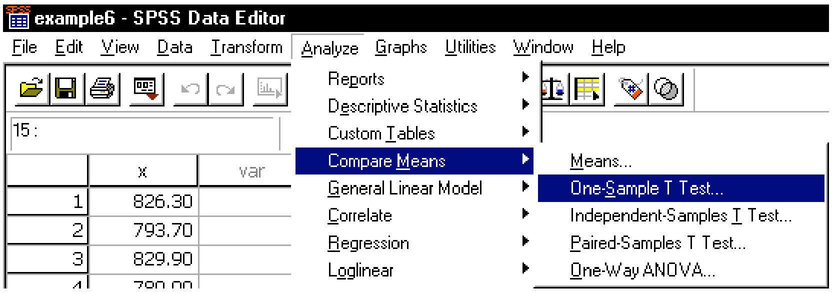
ขั้นที่ 1 เปิดแฟ้มข้อมูล exsample6.sav เลือกคำสั่ง Analyze / Descriptive Statistics / Explore เลือกตัวแปร x ไว้ในช่อง Dependent List เลือก OK ได้ผลลัพธ์ดังนี้

Descriptives

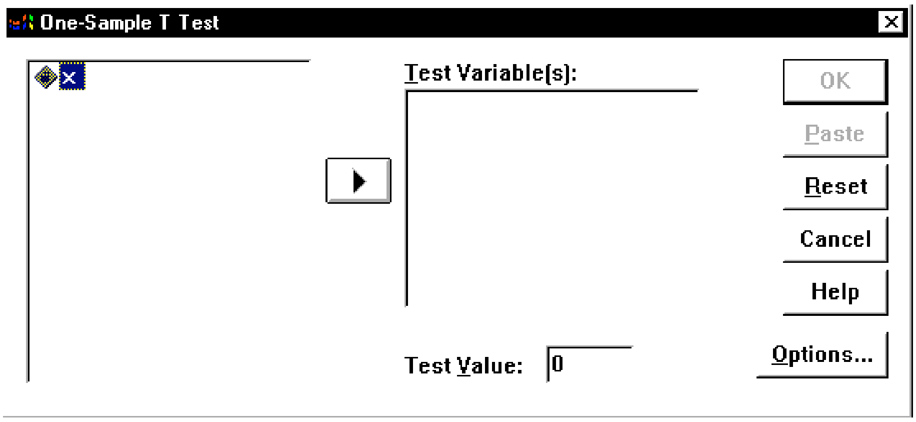
			Statistic	Std. Error
X	Mean		780.0000	7.30434
	95% Confidence Interval for Mean	Lower Bound	765.0610	
		Upper Bound	794.9390	
	5% Trimmed Mean		780.3722	
	Median		786.6000	
	Variance		1600.601	
	Std. Deviation		40.00751	
	Minimum		717.80	
	Maximum		835.80	
	Range		118.00	
	Interquartile Range		86.0500	
	Skewness		-.274	.427
	Kurtosis		-1.473	.833

ตัวอย่าง 5.9 จากตัวอย่าง 5.8 จงประมาณค่าอายุการใช้งานเฉลี่ยของหลอดไฟฟ้าที่ระดับความเชื่อมั่น 95% ด้วยคำสั่ง Analyze / Compare Means / One-Sample T Test

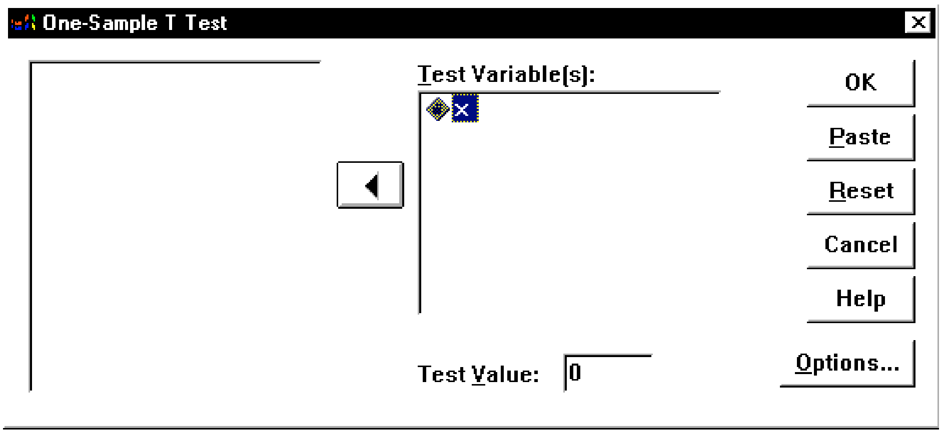
ขั้นที่ 1 เปิดแฟ้มข้อมูล example6.sav เลือกคำสั่ง Analyze / Compare Means / One-Sample T Test ดังนี้



ขั้นที่ 3.คลิกที่ One-Sample T Test จอภาพจะขึ้นเมนูย่อยของคำสั่ง One-Sample T Test



ขั้นที่ 4.เลือกตัวแปร x มาไว้ที่ช่อง Test Variable(s)



ขั้นที่ 5. คลิก OK จะได้ผลลัพธ์ดังนี้

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
X	30	780.0000	40.00751	7.30434

One-Sample Test

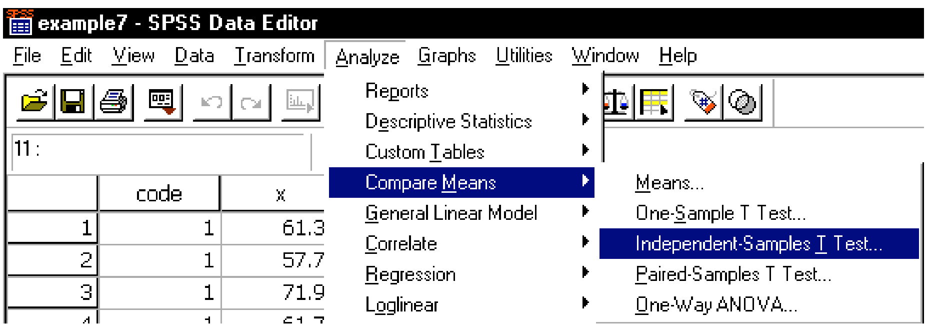
	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
X	106.786	29	.000	780.0000	765.0610	794.9390

2. การประมาณค่าผลต่างค่าเฉลี่ยของประชากร 2 กลุ่ม

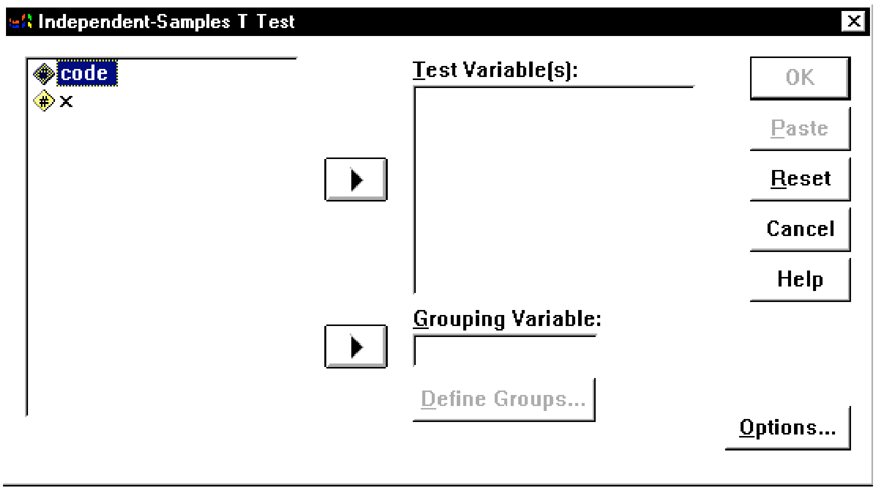
2.1 กรณีประชากร 2 กลุ่มเป็นอิสระ สามารถประมาณค่าผลต่างค่าเฉลี่ยด้วยคำสั่ง Analyze / Compare Means / Independent-Samples T Test

ตัวอย่าง 5.10 จงประมาณค่าผลต่างของค่าเฉลี่ยของตัวแปร x ในแฟ้มข้อมูล example7.sav ที่ระดับความเชื่อมั่น 95%

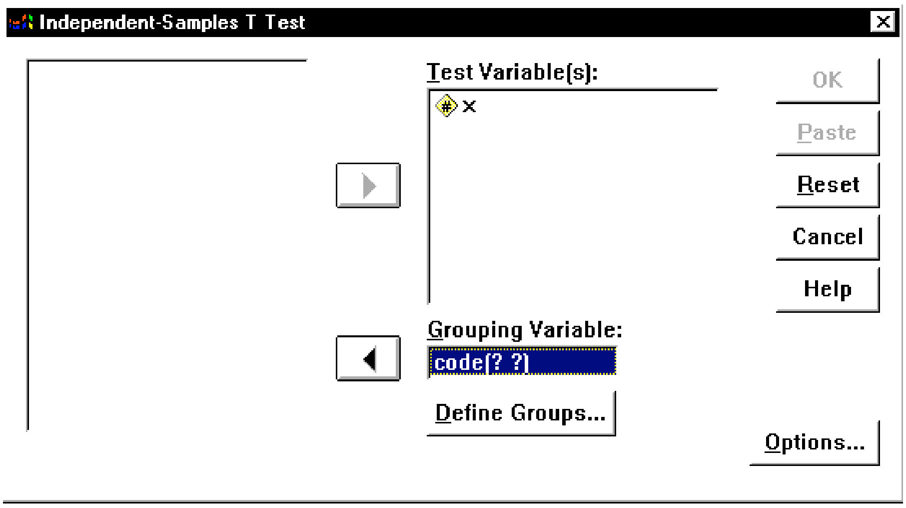
ขั้นที่ 1 เปิดแฟ้มข้อมูล example7.sav เลือกคำสั่ง Analyze / Compare Means / Independent-Samples T Test ดังนี้



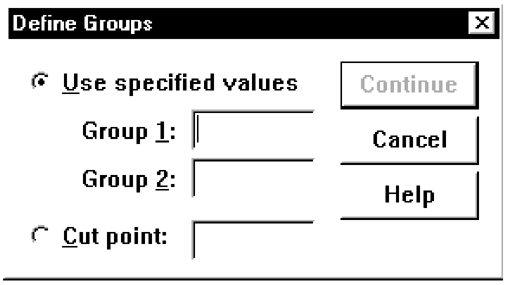
ขั้นที่ 2.คลิกที่ Independent-Samples T Test จะได้เมนูย่อยของคำสั่ง Independent-Samples T Test ดังนี้



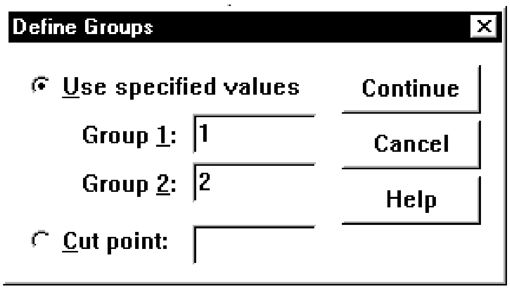
ขั้นที่ 3. เลือกตัวแปร x มาไว้ที่ช่อง Test Variable(s) และเลือกตัวแปร code มาไว้ที่ช่อง Grouping Variable ดังนี้



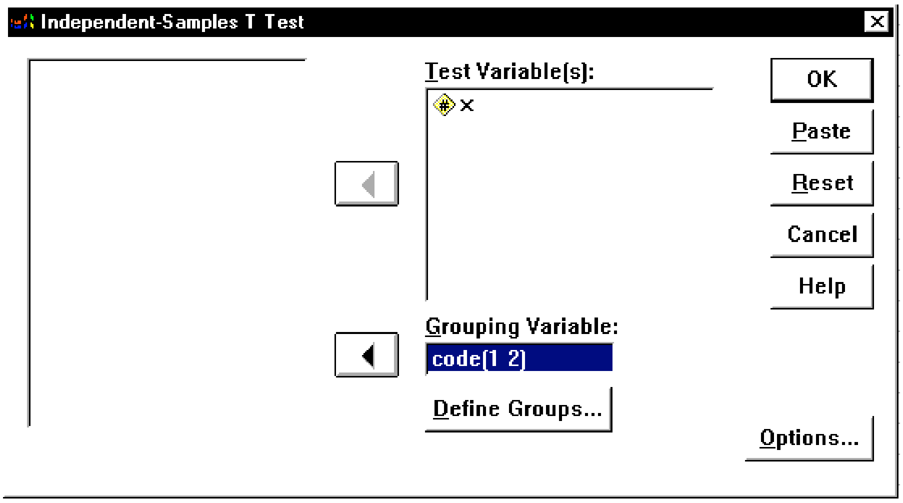
ขั้นที่ 4 การเลือกหมายเลขของกลุ่มในตัวแปร code ที่ต้องการวิเคราะห์ข้อมูลให้คลิกที่ Defined Groups จะได้เมนูย่อยของการเลือกหมายเลขกลุ่มเป็นดังนี้



ขั้นที่ 5 ให้นำเมาส์มาคลิกที่ช่อง Group 1 และ พิมพ์หมายเลข 1 ในช่อง Group 1 นำเมาส์มาคลิกที่ช่อง Group 2 และ พิมพ์หมายเลข 2 ในช่อง Group 2



ขั้นที่ 6. คลิก Continue จะกลับมาที่เมนูย่อย Independent-Samples T Test ที่ช่อง Grouping Variable ที่ตัวแปร Code จะเปลี่ยนเป็น Code[1 2]



ขั้นที่ 7 เลือก OK จะได้ผลลัพธ์ดังนี้

Group Statistics

	CODE	N	Mean	Std. Deviation	Std. Error Mean
X	1	9	64.0067	5.98773	1.99591
	2	16	58.9950	5.00084	1.25021

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
X	Equal variances assumed	.800	.380	2.242	23	.035	5.0117	2.23531	.38758	9.63575
	Equal variances not assumed			2.128	14.333	.051	5.0117	2.35514	-.02863	10.05196

การนำผลการคำนวณของ SPSS ไปใช้งานต้องเลือกให้เหมาะสมกับข้อกำหนดของประชากร

กรณีที่ 1. ภายใต้เงื่อนไขว่าประชากรมีการแจกแจงปกติและมีความแปรปรวนเท่ากัน ต้องใช้ผลสรุปใน Equal variances assumed

ช่วงความเชื่อมั่น 95% ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ $0.3876 < \mu_1 - \mu_2 < 9.6357$

กรณีที่ 2. ภายใต้เงื่อนไขว่าประชากรมีการแจกแจงปกติและมีความแปรปรวนไม่เท่ากันต้องใช้ผลสรุปใน Equal variances not assumed

ช่วงความเชื่อมั่น 95% ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ $0.0286 < \mu_1 - \mu_2 < 10.0520$

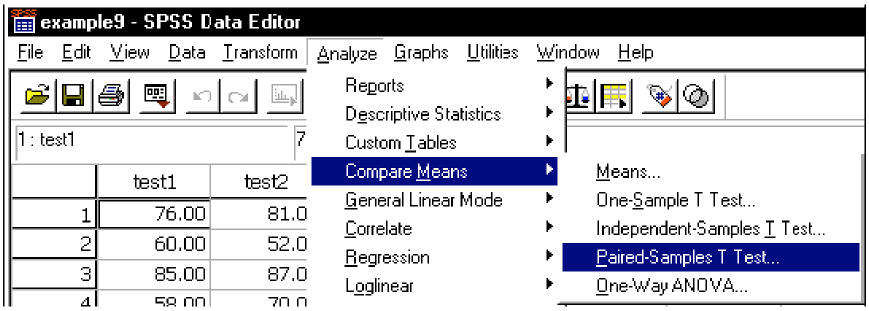
หมายเหตุ ถ้าไม่มีการกำหนดว่าความแปรปรวนเท่ากัน หรือไม่เท่ากัน การสรุปผลทางด้านสถิติเกี่ยวกับความแปรปรวนของประชากรทั้ง 2 ประชากรให้ดูจากค่าสถิติ levene ในตาราง Independent Samples Test ถ้า Sig. ของค่า Leven's Test for Equality of Variances มีค่าน้อยกว่า α แล้วเราสรุปได้ว่าประชากร 2 ชุดมีค่าความแปรปรวนไม่เท่ากัน โดยมีระดับนัยสำคัญ $\alpha = 0.05$

Levene's Test for Equality of Variances	F	.800
	Sig.	.380

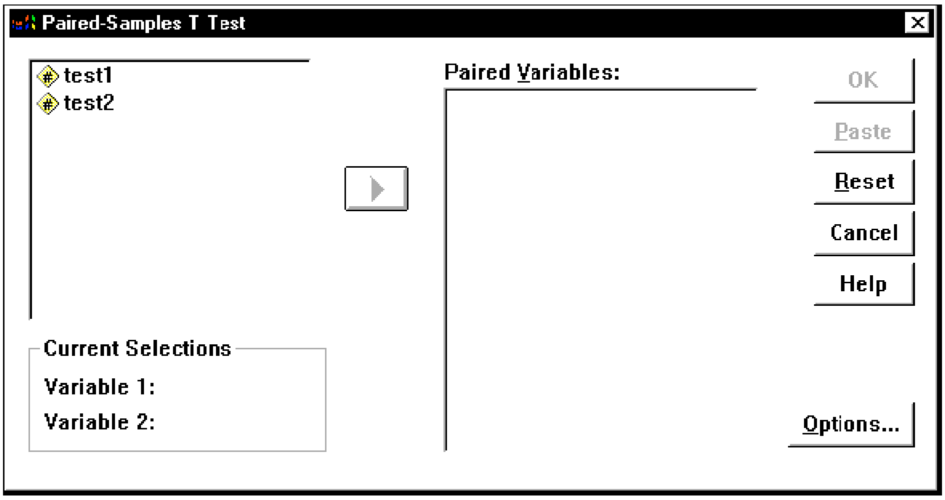
2.2 กรณีประชากร 2 กลุ่มไม่เป็นอิสระ สามารถประมาณค่าผลต่างของค่าเฉลี่ยด้วยคำสั่ง Analyze / Compare Means / Paired–Samples T Test

ตัวอย่าง 5.10 จงประมาณค่าผลต่างของค่าเฉลี่ยของตัวแปร test1 (คะแนนสอบย่อยครั้งที่ 1) และ test2 (คะแนนสอบย่อยครั้งที่ 2) ในแฟ้มข้อมูล example9.sav ที่ระดับความเชื่อมั่น 98%

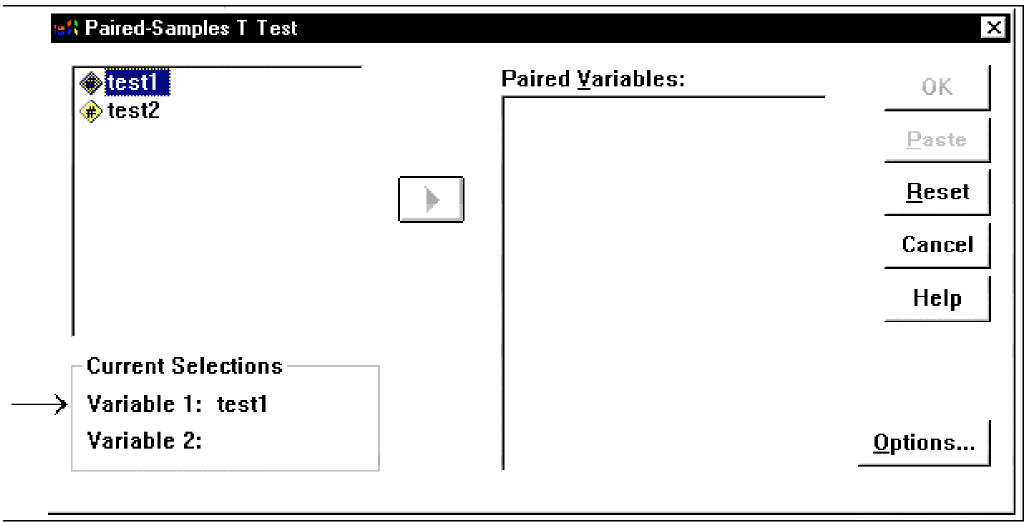
ขั้นที่ 1 เปิดแฟ้มข้อมูล example9.sav เลือกคำสั่ง Analyze / Compare Means / Paired–Samples T Test ดังนี้



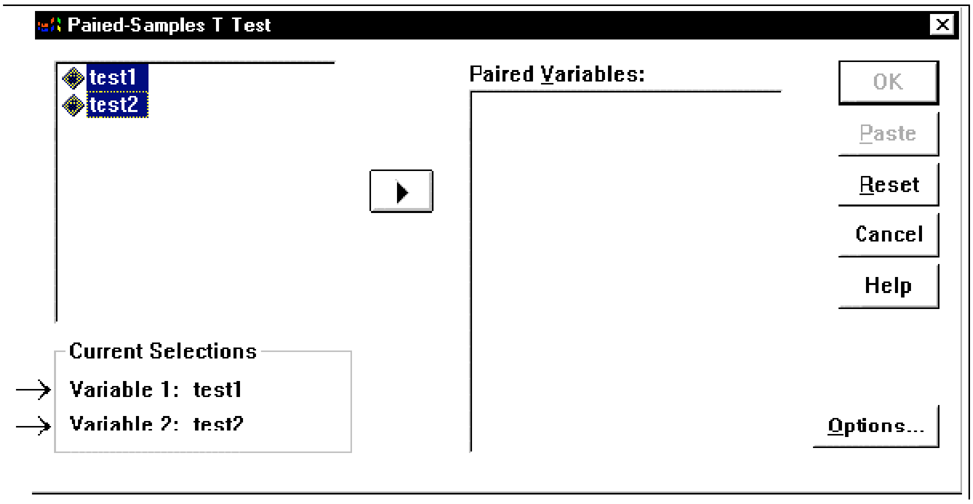
ขั้นที่ 2.คลิกที่ Paired–Samples T Test จะได้เมนูย่อยของคำสั่ง Paired–Samples T Test ดังนี้

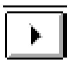


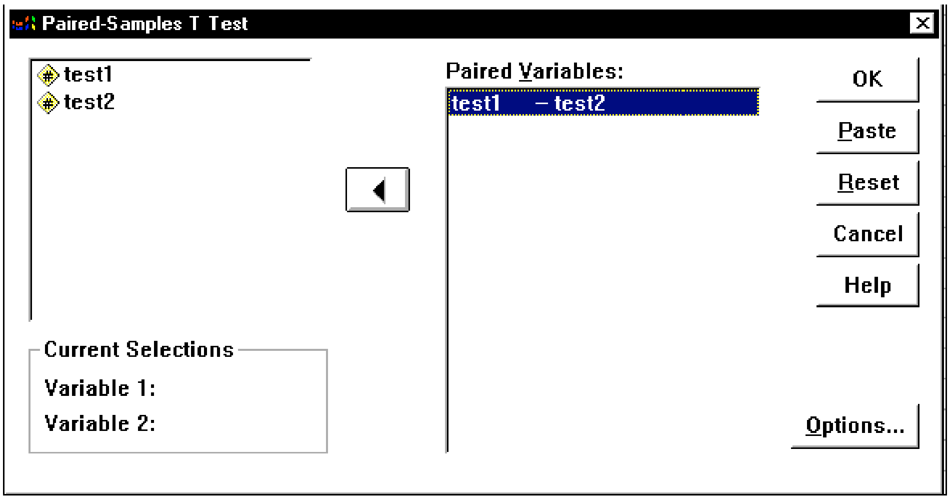
ขั้นที่ 3. การเลือกตัวแปรคลิกที่ตัวแปร test1 ตัวแปร test1 จะมาอยู่ที่ตำแหน่ง Variable 1 ก่อนเพื่อรอการเลือก Variable 2



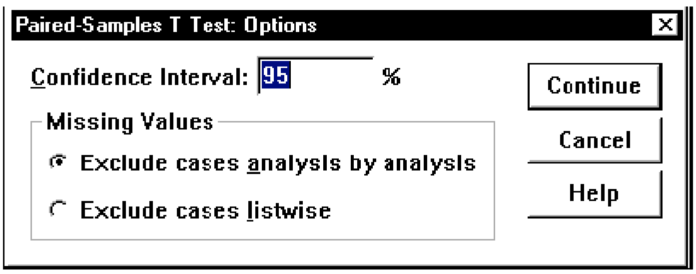
ขั้นที่ 4 คลิกที่ตัวแปร test2 จะได้ Variable 2: test2



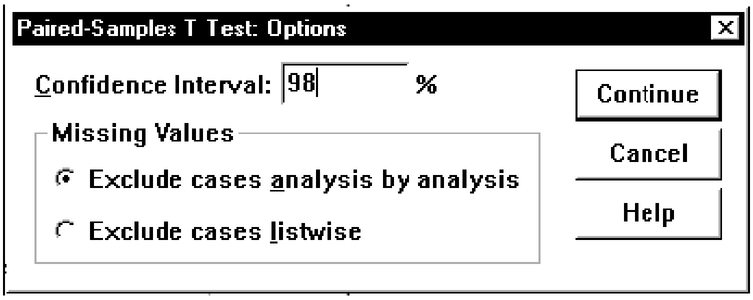
ขั้นที่ 5 คลิกที่  จะได้คู่ของตัวแปรที่ต้องการในช่อง Paired Variables: test1– test2



ขั้นที่ 6 คลิกปุ่ม Option เพื่อเปลี่ยนค่าระดับความเชื่อมั่น จะได้เมนูย่อย Paired Samples T Test Options ดังนี้



ขั้นที่ 7 เปลี่ยนระดับความเชื่อมั่นในช่อง Confidence Interval จาก 95% เป็น 98%



ขั้นที่ 8 เลือก Continue เลือก OK ได้ผลลัพธ์ ดังนี้

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	TEST1	75.8000	10	11.64092	3.68118
	TEST2	77.4000	10	12.17648	3.85054

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	TEST1 & TEST2	10	.857	.002

Paired Samples Test

			Pair 1 TEST1 - TEST2
Paired Differences	Mean		-1.6000
	Std. Deviation		6.38053
	Std. Error Mean		2.01770
	95% Confidence Interval of the Difference	Lower	-6.1644
		Upper	2.9644
t			-.793
df			9
Sig. (2-tailed)			.448

การใช้โปรแกรมสำเร็จรูป MS EXCEL

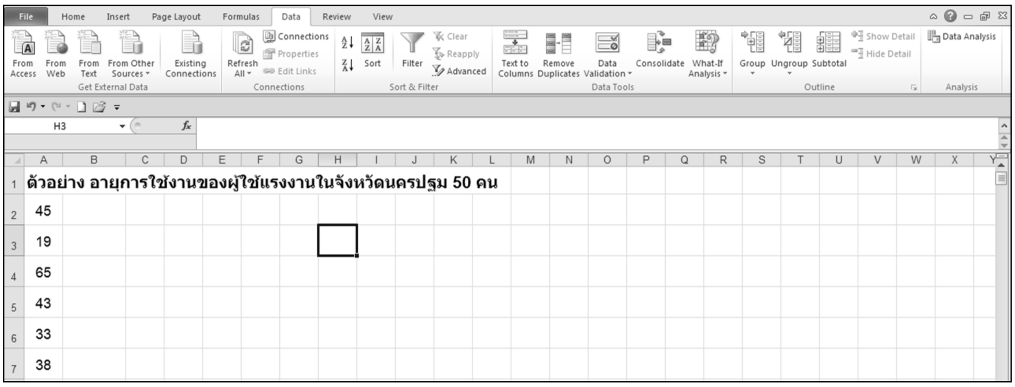
1. การประมาณค่าเฉลี่ยของประชากร 1 กลุ่ม

วิเคราะห์ด้วยชุดคำสั่ง Data Analysis

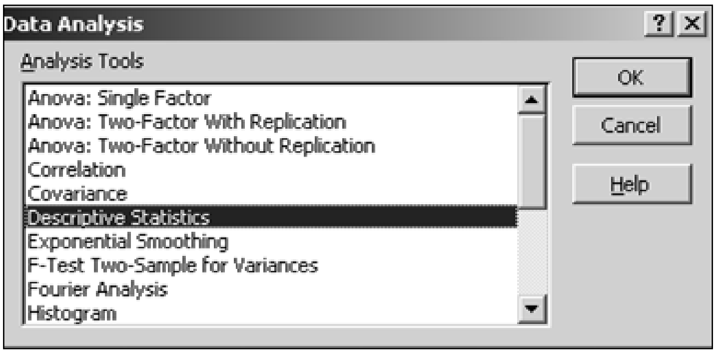
ตัวอย่าง 5.11 สำนักงานสถิติเก็บข้อมูลเกี่ยวกับอายุของผู้ใช้แรงงานในจังหวัดนครปฐมเพื่อใช้เป็นข้อมูลในการกำหนดอัตราค่าแรงต่อไป จึงทำการสุ่มตัวอย่างผู้ที่ใช้แรงงานมา 50 คน เพื่อหาอายุเฉลี่ย ปรากฏข้อมูลดังตาตารางด้านล่าง จงหาช่วงความเชื่อมั่น 90 % ของอายุเฉลี่ยของผู้ใช้แรงงานด้วย Microsoft Excel

45	38	22	58	40	42	43	32	34	19
19	21	33	16	49	29	30	43	37	62
65	57	60	41	28	35	37	51	37	26
43	26	27	31	33	24	34	28	39	38
33	27	42	40	31	34	38	35	29	33

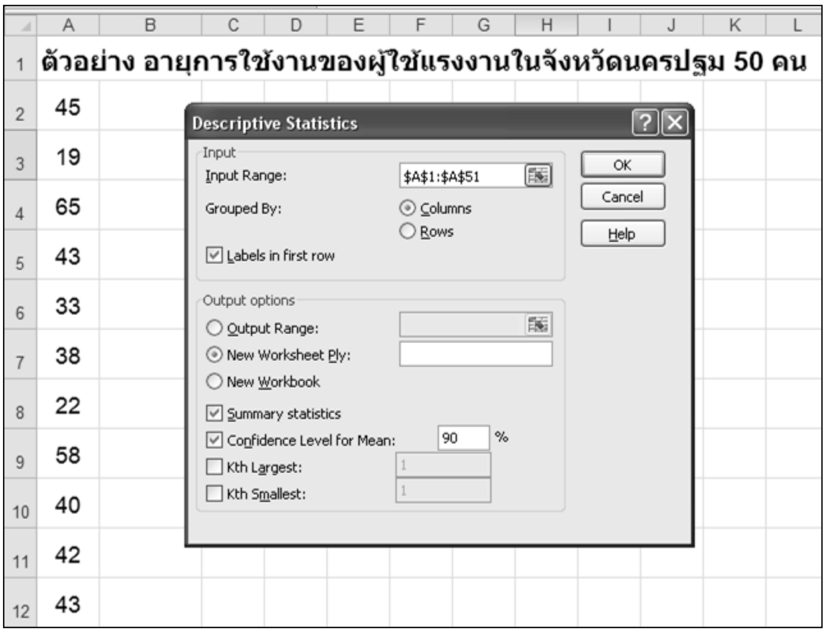
ขั้นตอนที่ 1 ใส่ข้อมูลใน Microsoft Excel เลือกเมนู Data เลือก Data Analysis



ขั้นตอนที่ 2 ในหน้าต่าง Data Analysis เลือก Descriptive Statistics



ขั้นตอนที่ 3 ใส่รายละเอียดในส่วนต่าง ๆ ดังรูป



ขั้นตอนที่ 4 จะได้ผลลัพธ์ ดังรูป

อายุการใช้งานของผู้ใช้แรงงานในจังหวัดนครปฐม	
Mean	**36.28**
Standard Error	**1.574293**
Median	**34.5**
Mode	**33**
Standard Deviation	**11.13194**
Sample Variance	**123.92**
Kurtosis	**0.490552**
Skewness	**0.726771**
Range	**49**
Minimum	**16**
Maximum	**65**
Sum	**1814**
Count	**50**
Confidence Level(90.0%)	**2.639383**

กำหนด μ คือ อายุเฉลี่ยของผู้ใช้แรงงานในจังหวัดนครปฐม

ที่ระดับความเชื่อมั่น 90 % ช่วงในการประมาณค่า μ คือ

$$\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}} = 36.28 \pm 2.64$$

นั่นคือ $33.64 \le \mu \le 38.92$ หมายความว่าอายุเฉลี่ยของผู้ใช้แรงงานในจังหวัดนครปฐม อยู่ในช่วง 33.64 ถึง 38.92 ปี ที่ระดับความเชื่อมั่น 90%

สรุปท้ายบท

การประมาณค่าเป็นวิธีการแรกของการวิเคราะห์ข้อมูลในส่วนสถิติอ้างอิงซึ่งเป็นการประมาณค่าพารามิเตอร์ที่สนใจจากค่าสถิติของตัวอย่างที่เลือกมาเป็นตัวแทนของประชากร โดยทั่วไปการประมาณค่าพารามิเตอร์ใด ๆ มี 2 วิธี คือการประมาณค่าแบบจุด และการประมาณค่าแบบช่วงมีข้อดีและข้อเสียแตกต่างกันไป การประมาณค่าแบบจุดนั้นสะดวกและง่าย แต่โอกาสที่จะประมาณค่าผิดพลาดจากความเป็นจริงมากในขณะที่การประมาณค่าแบบช่วงนั้นมีโอกาสที่จะประมาณค่าผิดพลาดจากความเป็นจริงน้อยลง แต่ก็ยุ่งยากกว่า ดังนั้นควรระมัดระวังในการเลือกวิธีการประมาณค่า

แบบฝึกหัดท้ายบท

1. จงประมาณค่าเฉลี่ยของค่าใช้จ่ายรายวันของนักศึกษา (บาท) ในระดับอุดมศึกษา 80 คน ที่ระดับความเชื่อมั่น 95% และ 98%

53	57	59	60	60	60	61	61	62	62	62	62	63	63	65	65	65	65	67	67
68	68	68	69	71	71	71	72	72	73	73	73	73	74	74	74	75	75	75	75
75	75	75	76	76	76	76	77	77	78	78	78	78	78	79	79	79	81	81	82
82	83	84	85	85	85	86	87	88	88	88	89	90	93	93	94	95	95	96	97

2. ร้านขายยาแผนปัจจุบันร้านหนึ่งสอบถามอายุ (ปี) ของผู้ที่เข้ามาซื้อยาในร้านในวันหนึ่งจำนวน 40 ราย ปรากฏว่าได้ผลดังนี้ 53 58 61 64 40 34 21 12 40 37 24 13 42 38 28 13 43 38 31 16 10 21 33 39 52 52 39 33 21 10 49 7 19 20 31 32 39 39 47 48 จงประมาณค่าเฉลี่ยของอายุผู้มาซื้อยาที่ระดับความเชื่อมั่น 90%