บทที่ 5
การประมาณค่า

หลังจากเก็บรวบรวมข้อมูลเรียบร้อยแล้ว การอธิบายลักษณะของข้อมูลจากตัวอย่าง จะอธิบายจากค่าสถิติที่เกี่ยวข้อง ซึ่งใช้วิธีการวิเคราะห์ข้อมูลในส่วนของสถิติพรรณนา ทั้งที่จริงแล้วสิ่งที่เราต้องการเป็นการอธิบายลักษณะของข้อมูลในประชากร ซึ่งจะอธิบายจากค่าพารามิเตอร์ที่เกี่ยวข้อง ดังนั้นจึงต้องนำค่าสถิติที่เกี่ยวข้องกลับมาสรุปค่าพารามิเตอร์ที่สนใจ ซึ่งการนำค่าสถิติมาสรุปหรืออ้างอิงค่าพารามิเตอร์นั้นสามารถสรุปได้หลายกรณี เช่นประมาณค่า ทดสอบสมมติฐาน ความสัมพันธ์หรือการพยากรณ์ค่าพารามิเตอร์ ดังนั้นวิธีการวิเคราะห์ข้อมูลในส่วนของสถิติอ้างอิงจึงมีหลากหลายวิธี การเลือกใช้วิธีการวิเคราะห์ข้อมูลในงานวิจัยจึงขึ้นอยู่กับวัตถุประสงค์ของงานวิจัย และประเภทของข้อมูล ถ้าวัตถุประสงค์ของงานวิจัยต้องการทราบว่าค่าพารามิเตรอร์มีค่าประมาณเท่าไร วิธีการในการวิเคราะห์ข้อมูลที่เลือกใช้คือ 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 $\bar{x}$

ถ้าต้องการทราบว่าเครื่องจักร A บรรจุนมกระป๋องแต่ละกระป๋องหนักแตกต่างกันกี่กรัม ค่าที่อธิบายได้คือค่าส่วนเบี่ยงเบนมาตรฐาน หรือ พารามิเตอร์ σ แต่ในทางปฏิบัติทราบเพียงค่าสถิติ s

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

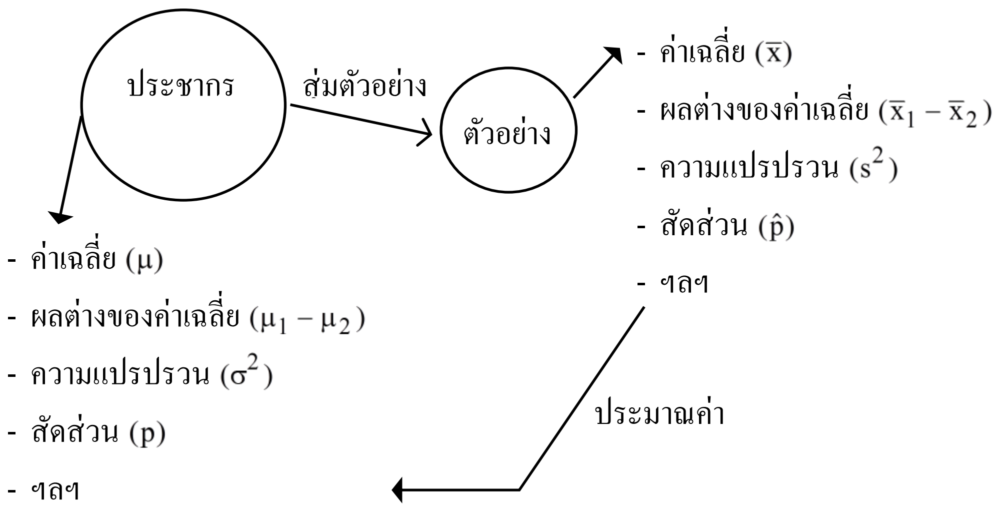
ตัวอย่าง 5.2 ถ้าต้องการศึกษาเกี่ยวกับน้ำหนักการบรรจุนมกระป๋องของเครื่องจักร A และ B ประชากรที่ศึกษากลุ่มที่ 1 คือนมกระป๋องที่บรรจุโดยเครื่องจักร A ตัวอย่าง คือนมกระป๋องที่บรรจุโดยเครื่องจักร A จำนวน 50 กระป๋อง และข้อมูลที่เก็บรวบรวมคือน้ำหนักการบรรจุนมกระป๋อง (g) ประชากรที่ศึกษากลุ่มที่ 2 คือนมกระป๋องที่บรรจุโดยเครื่องจักร 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 และเมื่อต้องการทราบค่าพารามิเตอร์ใด ในทางปฏิบัติเราจะหาค่าสถิติที่เกี่ยวข้อง และนำมาประมาณค่าพารามิเตอร์นั้น ดังนี้



จากรูปข้างต้นสามารถสรุปได้ว่า

การประมาณค่า หมายถึงการนำค่าสถิติกลับมาสรุป หรืออ้างอิงเกี่ยวกับค่าพารามิเตอร์ที่สนใจว่ามีค่าประมาณเท่าไร หรืออาจกล่าวได้ว่าการประมาณค่าก็คือการสรุปค่าพารามิเตอร์ที่สนใจโดยอาศัยค่าจริงจากตัวอย่าง หรือที่เรียกว่าค่าสถิตินั่นเอง

ประเภทของการประมาณค่า

ในการประมาณค่าพารามิเตอร์ต่าง ๆ นั้นจะมีหลักการ และวิธีการต่าง ๆ ที่เหมือนกัน จะแตกต่างกันที่สูตรในการประมาณค่าซึ่งขึ้นอยู่กับการแจกแจงของค่าสถิติ ดังนั้นในส่วนนี้เราจะศึกษาเกี่ยวกับหลักการและประเภทของการประมาณค่าพารามิเตอร์ซึ่งมี 2 ประเภท 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 $\bar{\mathrm{x}}_1 - \bar{\mathrm{x}}_2$

พารามิเตอร์ $\mathrm{p}_1 - \mathrm{p}_2$ จะมีค่าประมาณเท่ากับค่าสถิติ $\hat{\mathrm{p}}_1 - \hat{\mathrm{p}}_2$ เป็นต้น

และเรียก $\bar{x}$ เป็นตัวประมาณค่าแบบจุดของ μ

s^2 เป็นตัวประมาณค่าแบบจุดของ σ^2

s เป็นตัวประมาณค่าแบบจุดของ σ

$\hat{p}$ เป็นตัวประมาณค่าแบบจุดของ p

$\bar{\mathrm{x}}_1 - \bar{\mathrm{x}}_2$ เป็นตัวประมาณค่าแบบจุดของ $\mu_1 - \mu_2$

$\hat{\mathrm{p}}_1 - \hat{\mathrm{p}}_2$ เป็นตัวประมาณค่าแบบจุดของ $\mathrm{p}_1 - \mathrm{p}_2$ เป็นต้น

จะเห็นว่าการประมาณค่าประเภทนี้ ประมาณค่าพารามิเตอร์ให้เท่ากับค่าสถิติที่คำนวณจากตัวอย่างกลุ่มหนึ่ง ดังนั้นถ้ากลุ่มตัวอย่างเปลี่ยนไปค่าสถิติจะเปลี่ยนไปด้วย การประมาณค่าแบบจุดจึงเป็นการประมาณค่าพารามิเตอร์ที่มีโอกาสเกิดความคลาดเคลื่อนในการประมาณค่ามาก เพราะเป็นไปได้ยากที่ค่าพารามิเตอร์จะมีค่าเท่ากับค่าสถิติ เช่น ผู้ว่าราชการจังหวัดต้องการทราบราคาน้ำมันในจังหวัดนครปฐมว่าในขณะนี้ราคาประมาณกี่บาท จึงให้ลูกน้องออกไปสำรวจข้อมูล ลูกน้องไปสำรวจราคาน้ำมันของปั๊มน้ำมัน 10 แห่ง แล้วหาราคาน้ำมันเฉลี่ย ปรากฏว่าได้ 32.30 บาท จึงนำข้อมูลนี้ไปเสนอผู้ว่าราชการจังหวัดว่าในขณะนี้ราคาน้ำมันของจังหวัดนครปฐมประมาณ 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

แบ่งออกเป็น 2 กรณี คือ ทราบ σ^2 และ ไม่ทราบ σ^2

2.1 กรณีทราบ σ^2 ถ้าสุ่มตัวอย่างมาจากประชากรที่มีการแจกแจงแบบปกติ (หรืออาจจะมีการแจกแจงแบบใดก็ได้ ถ้าตัวอย่างมีขนาดใหญ่) เราสามารถสร้างช่วงของความเชื่อมั่น

ของ μ ได้โดยการพิจารณาการแจกแจงค่าจากตัวอย่างของ $\overline{X}$ ซึ่งเราทราบว่าการแจกแจงค่าจากตัวอย่างของ $\overline{X}$ มีการแจกแจงประมาณได้ด้วยการแจกแจงแบบปกติ ด้วยค่าเฉลี่ย $\mu_{\overline{X}} = \mu$ และความแปรปรวน $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$ ดังนั้น $Z = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}}$ จะมีการแจกแจงแบบปกติมาตรฐาน

ดังนั้นที่ระดับความเชื่อมั่น 100(1-α)% ช่วงในการประมาณค่า μ คือ

$$\overline{x} - z_{\alpha/2}\,\sigma/\sqrt{n} \le \mu \le \overline{x} + z_{\alpha/2}\,\sigma/\sqrt{n}$$

เมื่อ $\overline{x}$ แทนค่าเฉลี่ย

σ แทนส่วนเบี่ยงเบนมาตรฐาน

$z_{\alpha/2}$ แทนค่าจากตารางการแจกแจงแบบปกติมาตรฐาน ขึ้นอยู่กับค่า $1-\alpha$

n แทนจำนวนตัวอย่าง

ในบางครั้งอาจเขียนช่วงความเชื่อมั่นของ μ ดังนี้

$$\overline{x} \pm z_{\alpha/2}\,\sigma/\sqrt{n} \quad \text{หรือ} \quad \left[\overline{x} - z_{\alpha/2}\,\sigma/\sqrt{n}\,,\ \overline{x} + z_{\alpha/2}\,\sigma/\sqrt{n}\right]$$

2.2 กรณีไม่ทราบ σ^2 ในบางครั้งนั้นเราไม่ทราบความแปรปรวนของข้อมูลในประชากร ถ้ากรณีที่ตัวอย่างมีขนาดใหญ่ (n$\geq$30) ค่า $s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$ จะเป็นค่าประมาณที่ดีสำหรับค่า σ^2 ดังนั้นเราจึงใช้ s^2 แทน σ^2 ได้ และ $Z = \frac{\overline{x} - \mu}{s/\sqrt{n}}$ ยังคงมีการแจกแจงแบบปกติมาตรฐาน และการประมาณค่าเฉลี่ยของประชากรก็ยังคงเป็นในรูปแบบเดิมเพียงแต่ให้แทนค่า σ ด้วย s ช่วงความเชื่อมั่นในการประมาณค่า 100(1-α)% ของ μ คือ

$$\overline{x} \pm z_{\alpha/2}\,\frac{s}{\sqrt{n}}$$

แต่ถ้าตัวอย่างมีขนาดเล็ก ($n < 30$) ค่า s^2 จะไม่ใช่ค่าประมาณที่ดีของ σ^2 อีกต่อไป และ $\dfrac{(\bar{x}-\mu)}{s/\sqrt{n}}$ ก็จะไม่ใช่ตัวแปรเชิงสุ่ม Z ที่มีการแจกแจงแบบปกติมาตรฐานอีกต่อไป แต่ $\dfrac{(\bar{x}-\mu)}{s/\sqrt{n}}$ จะเป็นตัวแปรเชิงสุ่ม T ที่มีการแจกแจงแบบ t โดยมีองศาแห่งความเป็นอิสระ (degrees of freedom) เท่ากับ n-1

ดังนั้นที่ระดับความเชื่อมั่น 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 เพื่อสอบถามอายุ 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

$$\bar{x} = \frac{\sum x}{n}$$
$$= \frac{45+38+22+\ldots+33}{50}$$
$$= 36.28$$

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

กำหนด μ แทนอายุเฉลี่ยของผู้ใช้แรงงานในจังหวัดนครปฐม

ช่วงความเชื่อมั่น 90 % หมายความว่า 100(1- α)% = 90% ดังนั้น

$$1-\alpha = 0.90$$
$$\alpha = 1 - 0.90$$
$$\alpha = 0.1$$

จากตารางการแจกแจงแบบปกติมาตรฐานได้ค่า $z_{\alpha/2} = 1.645$

ดังนั้นช่วงความเชื่อมั่น 90 % ของอายุเฉลี่ยของผู้ใช้แรงงาน คือ

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

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

ตัวอย่าง 5.5 สุ่มนักศึกษา 36 คนจากนักศึกษาที่ลงทะเบียนเรียนวิชาสถิติ พบว่าผลการเรียนเฉลี่ยเท่ากับ 2.6 และส่วนเบี่ยงเบนมาตรฐานของผลการเรียนเท่ากับ 0.3 จงหาช่วงความเชื่อมั่น 95% และ 99% สำหรับการประมาณผลการเรียนเฉลี่ยของนักศึกษาทั้งหมดที่ลงทะเบียนเรียนวิชาสถิติ

วิธีทำ กำหนด μ แทนคะแนนเฉลี่ยของนักศึกษาที่ลงทะเบียนเรียนวิชาสถิติ

เนื่องจากจำนวนนักศึกษาเท่ากับ 36 มากกว่า 30

ดังนั้นช่วงความเชื่อมั่น 95 % ของผลการเรียนของนักศึกษา คือ

$$\bar{x} \pm z_{\alpha/2}\, {}^{s}\!/\!\sqrt{n} \quad = \quad 2.6 \pm (1.96)({}^{0.3}\!/\!\sqrt{36})$$
$$= \quad 2.6 \pm 0.098$$

หรือ $2.502 \leq \mu \leq 2.698$ นั่นคือผลการเรียนเฉลี่ยของนักศึกษาที่ลงทะเบียนเรียนวิชาหลักสถิติอยู่ในช่วง 2.502 ถึง 2.698 ที่ระดับความเชื่อมั่น 95%

ช่วงความเชื่อมั่น 99 % ของผลการเรียนเฉลี่ยของนักศึกษา คือ

$$\bar{x} \pm z_{\alpha/2}\, {}^{s}\!/\!\sqrt{n} \quad = \quad 2.6 \pm (2.576)({}^{0.3}\!/\!\sqrt{36})$$
$$= \quad 2.6 \pm 0.129$$

หรือ $2.471 \leq \mu \leq 2.729$ นั่นคือคือผลการเรียนเฉลี่ยของนักศึกษาที่ลงทะเบียนเรียนวิชาหลักสถิติอยู่ในช่วง 2.471 ถึง 2.729 ที่ระดับความเชื่อมั่น 95%

หมายเหตุ เมื่อระดับความเชื่อมั่นสูงขึ้น ช่วงในการประมาณค่าพารามิเตอร์จะกว้างขึ้น นั่นหมายถึงค่าพารามิเตอร์ที่เป็นไปได้จะมากขึ้นด้วย การนำผลการประมาณค่าไปใช้จะพิจารณาลำบากขึ้นด้วย และถ้าระดับความเชื่อมั่นน้อยลง ช่วงในการประมาณค่าพารามิเตอร์จะแคบลง นั่นหมายถึงค่าพารามิเตอร์ที่เป็นไปได้น้อยลง การนำผลการประมาณค่าไปใช้ก็พิจารณาได้ง่ายขึ้น ดังนั้นการกำหนดระดับความเชื่อมั่นในการประมาณค่าจึงขึ้นอยู่กับผู้ประมาณค่า หรือผู้ใช้ผลการประมาณค่าเป็นหลัก

ตัวอย่าง 5.6 เจ้าของฟาร์มสุนัขแห่งหนึ่งต้องการทราบระยะเวลาตั้งครรภ์ของสุนัข จึงทำการเลือกตัวอย่างสุนัขมา 15 ตัว และจดบันทึกระยะเวลาเป็นวันที่สุนัขตั้งครรภ์ โดยเริ่มนับตั้งแต่วันแรกที่สุนัขตั้งครรภ์จนถึงวันที่สุนัขคลอดลูกปรากฏผลดังนี้

62.0	61.4	59.8	62.2	60.3	60.4	59.4	60.2	60.4	60.8
61.8	59.2	61.1	60.4	60.9					

จงประมาณระยะเวลาเฉลี่ยที่สุนัขใช้ในการตั้งครรภ์ ที่ระดับความเชื่อมั่น 0.95

วิธีทำ กำหนด μ แทนระยะเวลาเฉลี่ยที่สุนัขตั้งครรภ์

ระดับความเชื่อมั่น 0.95 หมายความว่า $(1-\alpha) = 0.95$ ดังนั้น $\alpha = 0.05$ จากการเปิดตาราง t จะได้ $t_{0.025,14} = 2.145$

ดังนั้นช่วงความเชื่อมั่น 95% ของระยะเวลาเฉลี่ยที่สุนัขใช้ในการตั้งครรภ์ คือ

$$\bar{x} \pm t_{\alpha/2,n-1} \; s/\sqrt{n} = 60.69 \pm (2.145)(0.9/\sqrt{15})$$
$$= 60.69 \pm 0.5$$

หรือ $60.19 \le \mu \le 61.19$ นั่นคือระยะเวลาเฉลี่ยในการตั้งครรภ์ของสุนัขอยู่ในช่วง 60.19 วัน ถึง 61.19 วัน ที่ระดับความเชื่อมั่น 95%

ตัวอย่าง 5.7 การได้รับไนเตรท (nitrate) จากอาหารที่บริโภคในปริมาณสูงอาจทำให้เกิดผลร้ายต่อสิ่งมีชีวิต เช่น วัวจะให้นมได้น้อยลง เพื่อศึกษาเรื่องนี้ผู้ศึกษาได้ใส่ไนเตรทในน้ำดื่มจำนวน 2,000 ppm และให้หนูทดลอง 9 ตัวดื่มน้ำนี้ เปอร์เซ็นต์น้ำหนักของหนูที่เพิ่มขึ้นได้รับการบันทึกเป็นดังนี้

12.7 19.3 20.5 10.5 14.0 10.8 16.6 14.0 17.2

สมมติให้น้ำหนักมีการแจกแจงแบบปกติ จงหาช่วงความเชื่อมั่น 90% ของน้ำหนักเฉลี่ยที่เพิ่มขึ้นของหนู

วิธีทำ กำหนด μ แทนเปอร์เซนต์เฉลี่ยของน้ำหนักหนูที่เพิ่มขึ้น

ดังนั้นช่วงความเชื่อมั่น 90% ของเปอร์เซ็นต์เฉลี่ยของน้ำหนักที่เพิ่มขึ้นมีค่าเป็น

$$\bar{x} \pm t_{\alpha/2,n-1} \; s/\sqrt{n} = 15.067 \pm 1.86 \, (3.558)/\sqrt{9}$$
$$= 15.067 \pm 2.206$$

หรือ $12.861 \le \mu \le 17.273$ หมายความว่าเมื่อได้รับไนเตรทเปอร์เซนต์เฉลี่ยของน้ำหนักที่เพิ่มขึ้นมีค่าตั้งแต่ 12.861% ถึง17.273% ที่ระดับความเชื่อมั่น 90%

ตัวอย่าง 5.8 สถานีวิทยุแห่งหนึ่งต้องการทราบอายุของผู้ฟังรายการในสถานีวิทยุแห่งนี้ จึงทำการเก็บรวบรวมข้อมูลจากผู้ฟังจำนวน 81 คน ผลการวิเคราะห์ข้อมูลเบื้องต้น พบว่าอายุเฉลี่ย 22 ปี และส่วนเบี่ยงเบนมาตรฐาน 3 ปี ที่ระดับความเชื่อมั่น 95 % อายุเฉลี่ยของผู้ฟังรายการวิทยุของสถานีแห่งนี้ประมาณกี่ปี

วิธีทำ μ คือ อายุเฉลี่ยของผู้ฟังรายการในสถานีวิทยุ

ที่ระดับความเชื่อมั่น 95% ช่วงในการประมาณค่า μ คือ

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} = 22 \pm (1.96)(\frac{3}{\sqrt{81}})$$
$$= 22 \pm (1.96)(0.33)$$
$$= 22 \pm 0.65$$

นั่นคือ $21.35 \le \mu \le 22.65$ หมายความว่า ที่ระดับความเชื่อมั่น 95% อายุเฉลี่ยของผู้ฟังรายการในสถานีวิทยุแห่งนี้อยู่ในช่วง 21.35 ถึง 22.65 ปี

ตัวอย่าง 5.9 เจ้าหน้าที่อนามัยคนหนึ่งต้องการทราบน้ำหนักของเด็กหญิงอายุ 15 ปี ในหมู่บ้านที่ตนเองรับผิดชอบ จึงทำการสุ่มตัวอย่างเด็กหญิงอายุ 15 ปี ในหมู่บ้านจำนวน 144 คน แล้วชั่งน้ำหนัก พบว่าน้ำหนักเฉลี่ย 40 กิโลกรัม ส่วนเบี่ยงเบนมาตรฐาน 3 กิโลกรัม ที่ระดับความเชื่อมั่น 95 % น้ำหนักเฉลี่ยของเด็กหญิงอายุ 15 ปี ในหมู่บ้านนี้อยู่ในช่วงกี่กิโลกรัม

วิธีทำ กำหนด μ แทนน้ำหนักเฉลี่ยของเด็กหญิงอายุ 15 ปี

ที่ระดับความเชื่อมั่น 95% ช่วงในการประมาณค่า μ คือ

$$\begin{aligned} \bar{x} \pm t_{\frac{\alpha}{2},\, n-1}\frac{s}{\sqrt{n}} &= 40 \pm t_{0.025,143}\left(\frac{3}{\sqrt{144}}\right) \\ &= 40 \pm (1.96)(0.25) \\ &= 40 \pm 0.49 \end{aligned}$$

นั่นคือ $39.51 \le \mu \le 40.49$ หมายความว่า ที่ระดับความเชื่อมั่น 95% น้ำหนักเฉลี่ยของเด็กหญิงอายุ 15 ปี ในหมู่บ้านนี้อยู่ในช่วง 39.51 ถึง 40.49 กิโลกรัม

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

1. การประมาณค่าแบบจุดของผลต่างของค่าเฉลี่ย

ตัวประมาณค่าแบบจุดของ $\mu_1 - \mu_2$ 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. ทราบ σ_1^2 และ σ_2^2 และประชากร 2 กลุ่มเป็นอิสระกัน
2. ไม่ทราบ σ_1^2 และ σ_2^2 และประชากร 2 กลุ่มเป็นอิสระกัน
3. ไม่ทราบ σ_1^2 และ σ_2^2 และประชากร 2 กลุ่มไม่เป็นอิสระกัน

2.1 ทราบ σ_1^2 และ σ_2^2 ประชากร 2 กลุ่มเป็นอิสระกัน ถ้าสุ่มตัวอย่างที่เป็นอิสระขนาด n_1 และ n_2 โดยที่ n_1 และ n_2 มีขนาดใหญ่จากประชากร 2 กลุ่ม ที่มีการแจกแจงแบบใด ๆ ด้วยค่าเฉลี่ย และความแปรปรวน $\mu_1, \sigma_1^2, \mu_2, \sigma_2^2$ ตามลำดับ จะได้ว่าการแจกแจงค่าจากตัวอย่างของ $\overline{X}_1 - \overline{X}_2$ จะประมาณด้วยการแจกแจงแบบปกติด้วย $\mu_{\overline{X}_1 - \overline{X}_2} = \mu_1 - \mu_2$ และ $\sigma^2_{\overline{X}_1 - \overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ และ $Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ จะมีการแจกแจงแบบปกติมาตรฐาน

ดังนั้นที่ระดับความเชื่อมั่น 100(1-α)% ช่วงในการประมาณ $\mu_1 - \mu_2$ คือ

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

เมื่อ $\bar{x}_1$ แทนค่าเฉลี่ยของตัวอย่างกลุ่มที่ 1

$\bar{x}_2$ แทนค่าเฉลี่ยของตัวอย่างกลุ่มที่ 2

σ_1^2 แทนความแปรปรวนของประชากรกลุ่มที่ 1

σ_2^2 แทนความแปรปรวนของประชากรกลุ่มที่ 2

$z_{\alpha/2}$ แทนค่าจากตารางการแจกแจงแบบปกติมาตรฐาน ขึ้นอยู่กับค่า $1-\alpha$

n_1 แทนจำนวนตัวอย่างกลุ่มที่ 1

n_2 แทนจำนวนตัวอย่างกลุ่มที่ 2

ในบางครั้งอาจจะเขียนช่วงความเชื่อมั่นของ $\mu_1 - \mu_2$ ในรูปของ

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \text{หรือ}$$

$$\left[(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\ ,\ (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

หมายเหตุ ถ้าไม่ทราบความแปรปรวนของประชากรทั้ง 2 กลุ่ม และตัวอย่างที่เลือกมามีขนาดใหญ่ นั่นคือ จำนวนตัวอย่างมากกว่า 30 สามารถใช้ค่า s^2 ของตัวอย่างแต่ละกลุ่ม ประมาณค่า σ^2 ได้ และช่วงในการประมาณค่า $\mu_1 - \mu_2$ คือ $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

ตัวอย่าง 5.10 ข้อสอบมาตรฐานชุดหนึ่งถูกใช้นำไปทดสอบกับนักเรียนหญิง 50 คนและนักเรียนชาย 75 คน นักเรียนหญิงได้คะแนนเฉลี่ย 76 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 6 คะแนน นักเรียนชายได้คะแนนเฉลี่ย 82 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 8 คะแนน จงหาช่วงความเชื่อมั่น 96% ของผลต่างของคะแนนเฉลี่ยของนักเรียนชายกับนักเรียนหญิง

วิธีทำ กำหนด μ_1 แทนคะแนนเฉลี่ยของนักเรียนชาย

μ_2 แทนคะแนนเฉลี่ยของนักเรียนหญิง

ช่วงความเชื่อมั่น 96% หมายความว่า $100(1-\alpha)\% = 96\ \%$

$$1-\alpha = 0.96$$

$$\alpha = 1-0.96$$

$$\alpha = 0.04$$

$$n_1 = 75 \quad \bar{x}_1 = 82 \quad s_1 = 8$$

$$n_2 = 50 \quad \bar{x}_2 = 76 \quad s_2 = 6$$

จากตารางการแจกแจงแบบปกติมาตรฐานจะได้ค่า $z_{\alpha/2} = 2.054$ (โดยการเทียบบัญญัติไตรยางค์)

ดังนั้นช่วงความเชื่อมั่น 96% ของผลต่างของคะแนนเฉลี่ยของนักเรียนชายกับนักเรียนหญิง คือ

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} = (82-76) \pm (2.054)\sqrt{\frac{64}{75}+\frac{36}{50}}$$

$$= 6 \pm 2.58$$

หรือ $3.42 \le \mu_1 - \mu_2 \le 8.58$ นั่นคือนักเรียนชายกับนักเรียนหญิงสอบได้คะแนนแตกต่างกันประมาณ 3.42 ถึง 8.58 คะแนน

2.2 ไม่ทราบ σ_1^2 และ σ_2^2 ประชากร 2 กลุ่มเป็นอิสระกัน ในกรณีที่ไม่ทราบ σ^2 และตัวอย่างมีขนาดเล็กถูกสุ่มมาจากประชากรที่ประมาณได้ด้วยการแจกแจงแบบปกติ การแจกแจงค่าจากตัวอย่างของ $\bar{X}_1 - \bar{X}_2$ จะมีการแจกแจงแบบ t ด้วยองศาแห่งความเป็นอิสระเท่ากับ n-1

2.2.1 ในกรณีที่ไม่ทราบ σ^2 แต่ทราบว่า $\sigma_1^2 = \sigma_2^2 = \sigma^2$ เราจะประมาณ σ^2 ด้วยความแปรปรวนรวมจากตัวอย่าง s_p^2 โดยที่

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

ดังนั้น

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

จะเป็นตัวแปรเชิงสุ่ม T ที่มีการแจกแจงแบบ t ด้วยองศาแห่งความเป็นอิสระ $n_1 + n_2 - 2$ และในทำนองเดียวกันกับการหาช่วงความเชื่อมั่นในกรณีที่ทราบความแปรปรวนของประชากรทั้งสองกลุ่ม

ดังนั้นที่ระดับความเชื่อมั่น100(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

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

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

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

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

โดยที่ T มีการแจกแจงแบบ t ด้วยองศาแห่งความเป็นอิสระ ν ซึ่งมีค่าเป็น

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

ดังนั้นที่ระดับความเชื่อมั่น100(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 5.11 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 กำหนด μ_1 แทนคะแนนเฉลี่ยของนักศึกษาทั้งหมดที่ใช้วิธีการสอน A

μ_2 แทนคะแนนเฉลี่ยของนักศึกษาทั้งหมดที่ใช้วิธีการสอน 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 $0.22 \le \mu \le 7.78$ หมายความว่าผลต่างระหว่างคะแนนเฉลี่ยของนักศึกษามหาวิทยาลัยราชภัฏนครปฐมที่ใช้วิธีการสอน A กับ B อยู่ในช่วง 0.22 คะแนน กับ 7.78 คะแนน ที่ระดับความเชื่อมั่น 90%

ตัวอย่าง 5.12 จากตัวอย่าง 5.11 ถ้าข้อสมมติที่ว่าความแปรปรวนของประชากรทั้ง 2 กลุ่มเท่ากัน ไม่เป็นจริงแล้ว จงหาช่วงความเชื่อมั่น 95% ของ $\mu_1 - \mu_2$

องศาแห่งความเป็นอิสระ คือ

$$\begin{aligned} \nu &= \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1}+\frac{\left(s_2^2/n_2\right)^2}{n_2-1}} \\ &= \frac{(16/12+10)^2}{\frac{(16/12)^2}{12-1}+\frac{(25/10)^2}{9-1}} \\ &= 14.96 \approx 15 \end{aligned}$$

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

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

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

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

$$= 4 \pm 4.33$$

หมายความว่าผลต่างระหว่างคะแนนเฉลี่ยของนักศึกษามหาวิทยาลัยราชภัฏนครปฐม ที่ใช้วิธีการสอน A กับ B อยู่ในช่วง –0.33 คะแนน กับ 8.33 คะแนน ที่ระดับความเชื่อมั่น 95 %

2.3 ไม่ทราบ σ_1^2 และ σ_2^2 ประชากร 2 กลุ่มไม่เป็นอิสระกัน ในหัวข้อที่ผ่านมาเป็นการประมาณค่าแบบช่วงของผลต่างของค่าเฉลี่ยของประชากร 2 กลุ่มที่เป็นอิสระกัน นั้นคือ ตัวอย่างที่สุ่มมาจากประชากรกลุ่มที่ 1 เป็นอิสระจากตัวอย่างที่สุ่มมาจากประชากรกลุ่มที่ 2 ซึ่งเมื่อรวมขนาดของตัวอย่างแล้วจะมีขนาด $n_1 + n_2$ แต่ในการทดลองบางประเภท เช่นการควบคุมอาหารให้ถูกวิธีจะช่วยลดน้ำหนักได้อย่างมีประสิทธิภาพเพียงใด โดยเปรียบเทียบน้ำหนักก่อน และหลังควบคุม หรือการเปรียบเทียบผลการเรียนของนักศึกษาในปีสุดท้ายก่อนที่จะจบ กับปีแรกที่เริ่มเข้าเป็นนักศึกษาใหม่ว่าแตกต่างกันหรือไม่ เป็นต้น การทดลองเหล่านี้จะใช้ขนาดตัวอย่างเพียงแค่ n เท่านั้นสำหรับการทดลอง ซึ่งแต่ละตัวอย่างจะมีข้อมูล 2 ค่า เช่นน้ำหนักก่อนการควบคุม และน้ำหนักหลังการควบคุม ผลการเรียนของนักศึกษาในปีสุดท้ายก่อนที่จะจบ และปีแรกที่เริ่มเข้าเป็นนักศึกษาใหม่ คะแนนก่อนเรียนและหลังเรียน ปริมาณเชื้อแบคทีเรียในอาหารเมื่อเวลาผ่านไป 1 และ 2 ชั่วโมง เป็นต้น

ข้อมูลในลักษณะเช่นนี้จะเป็นข้อมูลของตัวอย่างที่มาจากประชากรเพียงกลุ่มเดียวแต่มีการรวบรวมข้อมูล 2 ครั้ง ซึ่งถือว่าข้อมูล 2 กลุ่มนี้ไม่เป็นอิสระกัน หรือเรียกว่า ข้อมูลคู่ (pair data) การประมาณค่าผลต่างของค่าเฉลี่ยของประชากร 2 กลุ่มในลักษณะนี้เรียกว่าการประมาณค่าเฉลี่ยของผลต่างระหว่าง 2 ประชากรโดยวิธีการจับคู่ หรือการประมาณค่าเฉลี่ยของผลต่างระหว่าง 2 ประชากรที่ไม่เป็นอิสระกัน

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}	$d_n = x_{1n} - x_{2n}$

จะเห็นว่า $d_1, d_2, \ldots, d_n$ เป็นตัวแปรสุ่ม ที่มาจากประชากรที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ_d และไม่ทราบความแปรปรวน σ_d^2 ด้วยวิธีการเช่นเดียวกับการประมาณค่าแบบช่วงของค่าเฉลี่ยของประชากร 1 กลุ่ม จะได้ช่วงความเชื่อมั่นในการประมาณค่า $\mu_d = \mu_1 - \mu_2$ ดังนี้

ที่ระดับความเชื่อมั่น100(1-α)% ช่วงในการประมาณค่า μ_d คือ

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

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

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

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

n แทนจำนวนคู่ตัวอย่าง

ในบางครั้งอาจจะเขียนช่วงความเชื่อมั่นของ μ_d ในรูปของ

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

ตัวอย่าง 5.13 นักจิตวิทยาผู้หนึ่งต้องการเปรียบเทียบวิธีการวัดระดับ IQ 2 วิธี จึงเลือกตัวอย่างเด็กมา 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
			$\sum d = 35$

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

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

วิธีทำ กำหนด μ_d แทนค่าเฉลี่ยของความแตกต่างของระดับ IQ
ช่วงความเชื่อมั่น 99% ของ μ_d คือ

$$\begin{aligned} \bar{d} \pm t_{\frac{\alpha}{2}, n-1} \frac{s_d}{\sqrt{n}} &= 4.375 \pm t_{0.005,7} \frac{1.685}{\sqrt{8}} \\ &= 4.375 \pm (3.499)(0.596) \\ &= 4.375 \pm 2.085 \end{aligned}$$

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

การประมาณค่าสัดส่วนของประชากร 1 กลุ่ม (p)

ในบางครั้งเราอาจสนใจลักษณะของประชากรเกี่ยวกับร้อยละของลักษณะที่สนใจ เช่นร้อยละของผู้ที่นิยมชมชอบการสื่อสารด้วย MSN ร้อยละของลูกค้าที่ชอบใช้ยาสีฟันที่ผสมเกลือ หรือร้อยละผลิตภัณฑ์ที่ไม่ได้มาตรฐานเป็นต้น ซึ่งการจะตอบคำถามเหล่านี้ได้ต้องทราบจำนวนของลักษณะที่สนใจแล้วเทียบกับจำนวนทั้งหมด หรือที่เราเรียกว่าค่าสัดส่วนของสิ่งที่สนใจ โดยทั่วไปการหาค่าสัดส่วนของสิ่งที่สนใจมักใช้กับข้อมูลเชิงคุณภาพมากกว่าข้อมูลเชิงปริมาณ หรือถ้าเป็นข้อมูลเชิงปริมาณก็เป็นข้อมูลเชิงปริมาณที่เก็บเป็นช่วง ๆ

1. การประมาณค่าแบบจุดของสัดส่วน

ตัวประมาณค่าแบบจุดของ p คือ $\hat{p}$ โดยที่

$$\hat{p} = \frac{\text{จำนวนตัวอย่างที่สนใจ}}{\text{จำนวนตัวอย่างทั้งหมด}}$$

2. การประมาณค่าแบบช่วงของสัดส่วน

ช่วงความเชื่อมั่นของ p เมื่อข้อมูลมีการแจกแจงแบบปกติ หรือใกล้เคียงปกติ และตัวอย่างมีจำนวนมากจนทำให้ $n\hat{p}$ และ $n(1-\hat{p})$ มีค่ามากกว่าหรือเท่ากับ 5 เป็นดังนี้

ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ ช่วงในการประมาณค่า p คือ

$$\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

เมื่อ $\hat{p}$ แทนสัดส่วนของตัวอย่างที่สนใจ

$z_{\frac{\alpha}{2}}$ แทนค่าจากตารางการแจกแจงแบบปกติมาตรฐาน มีค่าขึ้นอยู่กับ $1-\alpha$

n แทนจำนวนตัวอย่าง

ในบางครั้งอาจจะเขียนช่วงความเชื่อมั่นของ p ในรูปของ

$$\hat{p} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \text{หรือ} \quad \left[\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

ตัวอย่าง 5.14 ในการเลือกตัวอย่างผู้ใช้รถยนต์ส่วนบุคคลจำนวน 100 คน เพื่อสอบถามยี่ห้อรถคันต่อไปถ้าเขาจะเปลี่ยนรถคันที่ใช้อยู่ในปัจจุบัน ปรากฏว่ามีอยู่ 30 คนที่ต้องการเปลี่ยนไปใช้รถยี่ห้อ A ที่ระดับความเชื่อมั่น 90% จงประมาณค่าสัดส่วนของผู้ใช้รถยนต์ส่วนบุคคลที่จะเปลี่ยนไปใช้รถยี่ห้อ A ในการเปลี่ยนรถครั้งต่อไป

วิธีทำ กำหนด p คือสัดส่วนของผู้ใช้รถยนต์ส่วนบุคคลที่จะเปลี่ยนไปใช้รถยี่ห้อ A
ดังนั้น

$$\hat{p} = \frac{\text{จำนวนผู้ใช้รถยนต์ส่วนบุคคลที่จะเปลี่ยนไปใช้รถยี่ห้อ A}}{\text{จำนวนตัวอย่างทั้งหมด}}$$

$$= \frac{30}{100}$$

$$= 0.30$$

เนื่องจาก $n\hat{p} = (100)(0.30) = 30$; $n(1-\hat{p}) = 100(0.70) = 70$ มากกว่า 5
ดังนั้นที่ระดับความเชื่อมั่น 90% ช่วงในการประมาณค่า p คือ

$$\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.30 \pm (1.645)\sqrt{\frac{(0.30)(0.70)}{100}}$$

$$= 0.30 \pm 0.07$$

นั่นคือ $0.23 \le p \le 0.37$ หมายความว่าที่ระดับความเชื่อมั่น 90% สัดส่วนของผู้ใช้รถยนต์ส่วนบุคคลที่จะเปลี่ยนไปใช้รถยี่ห้อ A ในครั้งต่อไปอยู่ในช่วง 23% ถึง 37%

ตัวอย่าง 5.15 บริษัทขายเครื่องสำอางค์ยี่ห้อ tell me ต้องการทราบว่าผู้หญิงไทยใช้เครื่องสำอางค์ยี่ห้อนี้ประมาณกี่เปอร์เซนต์ จึงสุ่มตัวอย่างผู้หญิงไทยมา 500 คน ปรากฏว่ามีผู้ใช้เครื่องสำอางค์ยี่ห้อนี้ 95 คน ให้ตอบคำถามของบริษัทที่ระดับความเชื่อมั่น 95%

วิธีทำ กำหนด p แทนสัดส่วนของผู้หญิงไทยใช้เครื่องสำอางค์ยี่ห้อ tell me

$$\hat{p} = \frac{\text{จำนวนผู้ใช้เครื่องสำอางค์ยี่ห้อ tell me}}{\text{จำนวนตัวอย่างทั้งหมด}}$$

$$= \frac{95}{500}$$

$$= 0.19$$

เนื่องจาก $n\hat{p} = (500)(0.19) = 95; \; n(1-\hat{p}) = 500(0.81) = 405$ มากกว่า 5
ดังนั้น ที่ระดับความเชื่อมั่น 95% ช่วงในการประมาณค่า p คือ

$$\hat{p} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.19 \pm (1.96)\sqrt{\frac{(0.19)(0.81)}{500}}$$
$$= 0.19 \pm (1.96)(0.0175)$$
$$= 0.19 \pm 0.0343$$

นั่นคือ $0.1557 \le p \le 0.2243$ หมายความว่าที่ระดับความเชื่อมั่น 90% มีผู้หญิงไทยใช้เครื่องสำอางยี่ห้อ tell me ประมาณร้อยละ 15.57 ถึง 22.43

การประมาณค่าผลต่างของสัดส่วนของประชากร 2 กลุ่ม $(p_1 - p_2)$

1. การประมาณค่าแบบจุดของผลต่างของสัดส่วน

ตัวประมาณค่าแบบจุดของ $p_1 - p_2$ คือ $\hat{p}_1 - \hat{p}_2$ โดยที่

$$\hat{p}_1 = \frac{\text{จำนวนตัวอย่างที่สนใจในประชากรกลุ่มที่ 1}}{\text{จำนวนตัวอย่างกลุ่มที่ 1}}$$

$$\hat{p}_2 = \frac{\text{จำนวนตัวอย่างที่สนใจในประชากรกลุ่มที่ 2}}{\text{จำนวนตัวอย่างกลุ่มที่ 2}}$$

2. การประมาณค่าแบบช่วงของผลต่างของสัดส่วน

ช่วงความเชื่อมั่นของ $p_1 - p_2$ เมื่อข้อมูลมีการแจกแจงแบบปกติ หรือใกล้เคียงปกติ และตัวอย่างมีจำนวนมากจนทำให้ $n_1\hat{p}_1, n_1(1-\hat{p}_1)$ และ $n_2\hat{p}_2, n_2(1-\hat{p}_2)$ มีค่ามากกว่าหรือเท่ากับ 5 เป็นดังนี้

ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ ช่วงในการประมาณค่า $p_1 - p_2$ คือ

$$\hat{p}_1 - \hat{p}_2 - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \le p_1 - p_2 \le \hat{p}_1 - \hat{p}_2 + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

เมื่อ $\hat{p}_1$ แทนสัดส่วนของตัวอย่างที่สนใจในตัวอย่างกลุ่มที่ 1

$\hat{p}_2$ แทนสัดส่วนของตัวอย่างที่สนใจในตัวอย่างกลุ่มที่ 2

$z_{\frac{\alpha}{2}}$ แทนค่าจากตารางการแจกแจงแบบปกติมาตรฐาน มีค่าขึ้นอยู่กับ $1-\alpha$

n_1 แทนจำนวนตัวอย่างกลุ่มที่ 1

n_2 แทนจำนวนตัวอย่างกลุ่มที่ 2

บางครั้งอาจเขียนช่วงความเชื่อมั่นของ $p_1 - p_2$ ในรูปของ

$$\hat{p}_1 - \hat{p}_2 \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

หรือ $$\left[\hat{p}_1 - \hat{p}_2 - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}, \hat{p}_1 - \hat{p}_2 + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\right]$$

ตัวอย่าง 5.16 ในการหยั่งเสียงการลงคะแนนของผู้มีสิทธิเลือกตั้งของเขต 1 และเขต 2 ที่มีต่อผู้สมัครเบอร์ 1 จึงทำการสุ่มตัวอย่างผู้มีสิทธิ์เลือกตั้งจากเขต 1 จำนวน 300 คน และเขต 2 จำนวน 200 คน พบว่าในเขต 1 มีผู้ลงคะแนนให้เบอร์ 1 จำนวน 168 คน และเขต 2 จำนวน96 คน ที่ระดับความเชื่อมั่น 90% สัดส่วนของผู้ลงคะแนนให้ผู้สมัครเบอร์ 1 ในเขต 1 ต่างจากเขต 2 ประมาณเท่าไร

วิธีทำ กำหนด p_1 แทนสัดส่วนของผู้ลงคะแนนให้ผู้สมัครเบอร์ 1 ในเขต 1

p_2 แทนสัดส่วนของผู้ลงคะแนนให้ผู้สมัครเบอร์ 1 ในเขต 2

ดังนั้น

$$\hat{p}_1 = \frac{a_1}{n_1} = \frac{168}{300} = 0.56 \quad \text{และ} \quad \hat{p}_2 = \frac{a_2}{n_2} = \frac{96}{200} = 0.48$$

เนื่องจาก $n_1\hat{p}_1 = 168, n_1(1-\hat{p}_1) = 132$ และ $n_2\hat{p}_2 = 96, n_2(1-\hat{p}_2) = 104$ มีค่ามากกว่าหรือเท่ากับ 5 ดังนั้น

ที่ระดับความเชื่อมั่น 90% ช่วงในการประมาณค่า $p_1 - p_2$ คือ

$$\hat{p}_1 - \hat{p}_2 \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$= (0.56 - 0.48) \pm (1.645)\sqrt{\frac{(0.56)(0.44)}{300} + \frac{(0.48)(0.52)}{200}}$$

$$= 0.08 \pm (1.645)(0.045)$$

$$= 0.08 \pm 0.073$$

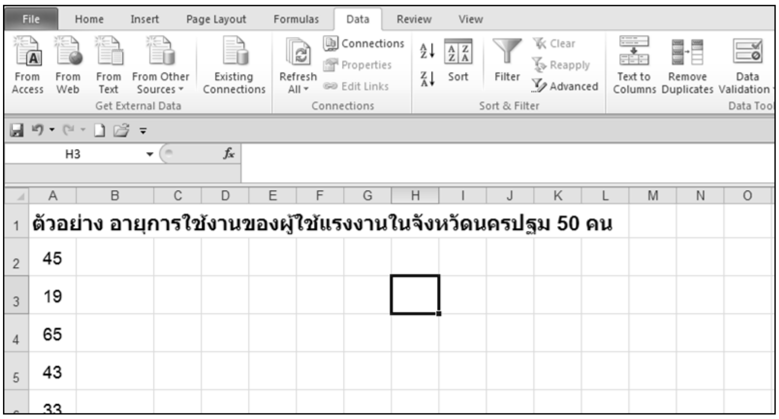
นั่นคือ $0.007 \le p_1 - p_2 \le 0.153$ หมายความว่าสัดส่วนของผู้ลงคะแนนให้ผู้สมัครเบอร์ 1 ในเขต 1 ต่างจากเขต 2 ประมาณ 0.007 ถึง 0.153 ที่ระดับความเชื่อมั่น 90 %

การใช้โปรแกรมสำเร็จรูป

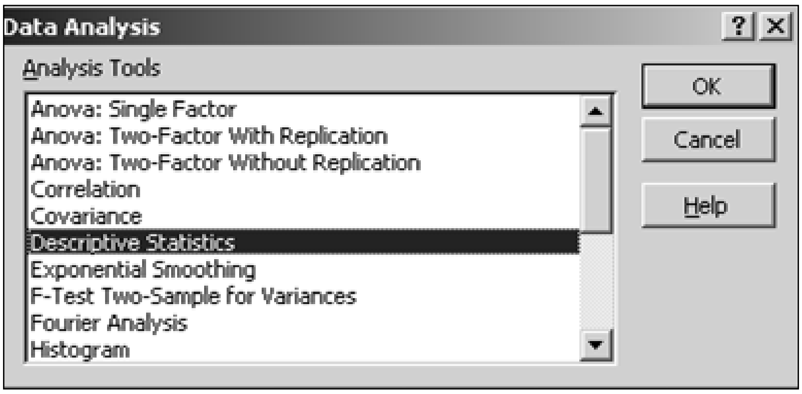
ตัวอย่าง 5.17 สำนักงานสถิติเก็บข้อมูลเกี่ยวกับอายุของผู้ใช้แรงงานในจังหวัดนครปฐมเพื่อใช้เป็นข้อมูลในการกำหนดอัตราค่าแรงต่อไป จึงทำการสุ่มตัวอย่างผู้ที่ใช้แรงงานมา 50 คน เพื่อหาอายุเฉลี่ย ปรากฏข้อมูลดังตาตารางด้านล่าง จงหาช่วงความเชื่อมั่น 90 % ของอายุเฉลี่ยของผู้ใช้แรงงานด้วย Microsoft Excel

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

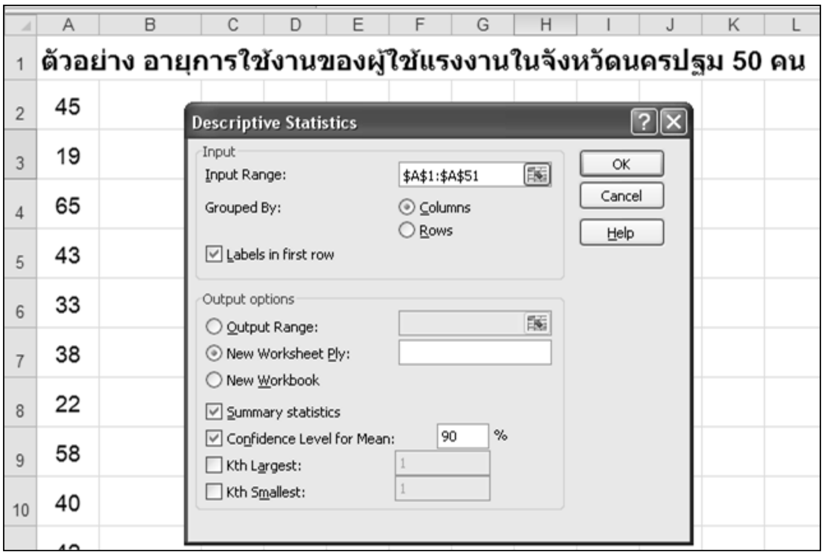
ขั้นตอนที่ 1 ใส่ข้อมูลใน Microsoft Excel เลือกเมนู Data เลือก Data Analysis



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 แทนอายุเฉลี่ยของผู้ใช้แรงงานในจังหวัดนครปฐม
ที่ระดับความเชื่อมั่น 90 % ช่วงในการประมาณค่า μ คือ

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

นั่นคือ $33.64 \leq \mu \leq 38.92$ หมายความว่าอายุเฉลี่ยของผู้ใช้แรงงานในจังหวัดนครปฐมประมาณ 33.64 ถึง 38.92 ปี ที่ระดับความเชื่อมั่น 90%

สรุปท้ายบท

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

แบบฝึกหัดท้ายบท

1. สุ่มตัวอย่างขนาด 9 จากประชากรที่มีการแจกแจงแบบ N(μ,25) จากตัวอย่างที่มี $\bar{x} = 70$ จงประมาณช่วงความเชื่อมั่นของ μ โดยกำหนดระดับความเชื่อมั่น ดังนี้
   1.1 ระดับความเชื่อมั่น 99%
   1.2 ระดับความเชื่อมั่น 95%
   1.3 ระดับความเชื่อมั่น 80%
2. เพื่อประมาณค่าเฉลี่ยของเส้นผ่านศูนย์กลางของลูกปืนของบริษัทแห่งหนึ่ง จึงสุ่มตัวอย่างลูกปืน 200 ลูก ปรากฏว่าได้ค่าเฉลี่ยเป็น0.824 นิ้ว ส่วนเบี่ยงเบนมาตรฐาน 0.042 นิ้ว จงหาช่วงความเชื่อมั่น 95% และ 98%

3. สุ่มตัวอย่างหลอดไฟ 25 หลอด จากบรรดาหลอดไฟ 40 วัตต์ รุ่นหนึ่ง ปรากฏว่าได้อายุเฉลี่ยในการใช้งาน 1,410 ชั่วโมง ส่วนเบี่ยงเบนมาตรฐาน 200 ชั่วโมง สมมติว่าอายุการใช้งานของหลอดไฟมีการแจกแจงแบบปกติ จงประมาณค่าอายุการใช้งานเฉลี่ยของหลอดไฟ 40 วัตต์ ที่ระดับความเชื่อมั่น 0.95
4. ในการผลิตแผ่นเหล็กไร้สนิมให้มีคุณภาพนั้น จะต้องมีการตรวจสอบผลิตภัณฑ์อย่างสม่ำเสมอ เมื่อตรวจสอบเหล็กไร้สนิมชนิดแผ่นกว้าง ชนิดที่มีความกว้าง 660 มม. จำนวน 25 แผ่น ในช่วงสัปดาห์หนึ่ง พบว่าค่าเฉลี่ยของความกว้าง เท่ากับ 662 มม. ความแปรปรวนเท่ากับ 16 จงหา

   4.1 ช่วงความเชื่อมั่น 90% ของ μ

   4.2 โดยไม่ต้องแสดงค่าของช่วงความเชื่อมั่น ถ้าเปลี่ยนจาก 90% เป็น 99% จะมีผลเช่นใดกับช่วงที่สร้างขึ้นใหม่นี้
5. สุ่มตัวอย่างขนาด 16 จากประชากรกลุ่ม 1 ที่มีการแจกแจงแบบปกติ (μ_1,7) โดย $\overline{X}_1 = 42$ สุ่มตัวอย่างขนาด 25 จากประชากรกลุ่ม 2 ที่มีการแจกแจงแบบปกติ (μ_2,9) โดย $\overline{X}_2 = 31$ จงหาช่วงความเชื่อมั่น 90% ของ $\mu_1 - \mu_2$
6. อาจารย์สอนสถิติเบื้องต้นผู้หนึ่งสอนวิชานี้ให้แก่นักเรียน 2 กลุ่ม กลุ่มแรกมีนักเรียน 10 คน กลุ่มที่สอง 13 คน อาจารย์ใช้วิธีการสอบโดยสอบทุกสัปดาห์กับนักเรียนกลุ่มแรก และสอบสอบกลางภาคเพียงครั้งเดียวกับนักเรียนกลุ่มที่ 2 และข้อสอบปลายภาคใช้ชุดเดียวกัน ปรากฏว่าคะแนนเฉลี่ยของนักเรียนกลุ่มแรก คือ 80 คะแนน ค่าเบี่ยงเบนมาตรฐาน 10 คะแนน คะแนนเฉลี่ยของนักเรียนกลุ่มที่ 2 คือ 82 คะแนน ค่าเบี่ยงเบนมาตรฐาน 12 คะแนน ถ้าสมมติว่า ความแปรปรวนของคะแนนของนักเรียนทั้งสองกลุ่มเท่ากัน จงหา

   6.1 ช่วงความเชื่อมั่น 90% ของ $\mu_1 - \mu_2$

   6.2 ถ้าข้อสมมติข้างต้นไม่เป็นจริง จงหาช่วงความเชื่อมั่น 90% ของ $\mu_1 - \mu_2$
7. การตรวจวัดระดับกรดยูริกผู้ป่วยชายในกลุ่มอายุหนึ่งที่เป็นโรคประสาทจำนวน 12 คน ได้ค่าเฉลี่ยระดับกรดยูริกเท่ากับ 4.5 มิลลิกรัมต่อ 100 มิลลิลิตร และจากการตรวจวัดระดับกรดยูริกผู้ป่วยหญิงในกลุ่มอายุเดียวกันกับผู้ป่วยชายที่เป็นโรคประสาทจำนวน 15 คน ได้ค่าเฉลี่ยระดับกรดยูริกเท่ากับ 3.4 มิลลิกรัมต่อ 100 มิลลิลิตร ถ้าปริมาณกรดยูริกในประชากรทั้งสองมีการแจกแจงแบบปกติ โดยที่ประชากรทั้งสองมีความแปรปรวนเท่ากับ 1 เท่ากัน จงสร้างช่วงความเชื่อมั่น 95% ของการประมาณความแตกต่างระหว่างปริมาณระดับกรดยูริกโดยเฉลี่ยในผู้ป่วยชายและผู้ป่วยหญิงที่เป็นโรคประสาทในกลุ่มอายุเดียวกัน

8. การทดลองเพื่อเปรียบเทียบการละลายของสารชนิดหนึ่งด้วยตัวทำละลาย 2 ชนิด คือ A และ B โดยแต่ละครั้งของการทดลองใช้สารทดลอง 100 มิลลิกรัม ตัวทำละลาย 50 เซนติลิตร เมื่อทดลองด้วยตัวทำละลายชนิด A จำนวน 9 ครั้ง พบว่า ได้น้ำหนักสารที่ละลายเฉลี่ย 64 มิลลิกรัม ส่วนเบี่ยงเบนมาตรฐาน 6 มิลลิกรัม ทดลองด้วยตัวทำละลายชนิด B จำนวน 14 ครั้ง พบว่า ได้น้ำหนักสารที่ละลายเฉลี่ย 49 มิลลิกรัม ส่วนเบี่ยงเบนมาตรฐาน 5 มิลลิกรัม ถ้าความแปรปรวนของน้ำหนักสารที่ละลายด้วยตัวทำละลายทั้งสองชนิดเท่ากัน จงหาช่วงความเชื่อมั่น 95% ของการประมาณความแตกต่างของน้ำหนักเฉลี่ยของสารที่ละลายโดยใช้ตัวทำละลายทั้งสองชนิด
9. ผู้ผลิตรองเท้ากีฬารายหนึ่งต้องการทราบว่าลูกค้ามีความนิยมในรองเท้ากีฬาที่ผลิตหรือไม่ จึงทำการสำรวจผู้ซื้อรองเท้ากีฬาทั้งหมด 500 คน พบว่ามีลูกค้า 60% ที่นิยมในรองเท้ากีฬาที่ผลิต อยากทราบว่าที่ระดับความเชื่อมั่น 0.90 สัดส่วนของลูกค้าที่นิยมในรองเท้ากีฬาที่ผลิตเป็นเท่าใด
10. เพื่อทดสอบเชื้อไวรัสชนิดหนึ่งที่ฉีดเข้าไปในกระต่ายว่าทำให้กระต่ายที่ได้รับเชื้อนี้มีโอกาสเป็นเนื้องอกที่หน้าอกในอัตราร้อยละเท่าไร ผู้ทำการวิจัยจึงได้สุ่มตัวอย่างกระต่ายที่ได้รับเชื้อนี้มา 128 ตัว แล้วทำการตรวจสอบ ปรากฏว่ามีอยู่ 72 ตัว ที่เป็นเนื้องอกที่หน้าอก จงประมาณค่าสัดส่วนของกระต่ายทั้งหมดที่ได้รับเชื้อไวรัสนี้แล้วจะเป็นเนื้องอกที่หน้าอก โดยใช้ช่วงความเชื่อมั่น 95%
11. ในการทดลองเพื่อทดสอบประสิทธิภาพของยาที่ใช้รักษาหมู โดยแบ่งหมูที่เป็นโรคชนิดหนึ่งออกเป็น 2 กลุ่ม ๆ ละ 100 ตัว และทดลองให้ยาแก่หมูกลุ่มที่ 1 แต่ไม่ให้ยาแก่หมูกลุ่มที่ 2 นอกจากนั้นได้ควบคุมตัวแปรอื่น ๆ ให้มีความสม่ำเสมอกันทั้งสองกลุ่ม จากการทดลองพบว่า หมูกลุ่มที่ 1 หายจากโรค 75 ตัว ส่วนหมูกลุ่มที่ 2 หายจากโรค 65 ตัว ที่ระดับความเชื่อมั่น 95% สัดส่วนของหมูกลุ่มที่ 1 ที่หายจากการเป็นโรคต่างจากหมูกลุ่มที่ 2 ประมาณเท่าไร
12. จากการสุ่มตัวอย่างผู้ลงคะแนนเสียงเลือกตั้งสมาชิกสภาผู้แทนราษฎรของจังหวัดสงขลาครั้งหนึ่ง จากอำเภอเมือง 300 คน และอำเภอระโนด 200 คน พบว่า มี 56% และ 48% ตามลำดับ ที่เลือกผู้สมัครรับเลือกตั้งเบอร์ 1 จงหาช่วงความเชื่อมั่น 95% ของความแตกต่างระหว่างสัดส่วนที่ผู้ลงคะแนนเลือกผู้สมัครรับเลือกตั้งเบอร์ 1 ของทั้งสองอำเภอนี้