บทที่ 4
การประมาณค่า

วัตถุประสงค์ที่สำคัญวัตถุประสงค์หนึ่งของการวิจัย คือการสรุปผลข้อมูลเพื่อนำไปใช้ประโยชน์ในด้านต่างๆต่อไป ซึ่งการสรุปผลดังกล่าวส่วนมากเกิดจากการวิเคราะห์ข้อมูลจากกลุ่มตัวอย่างทำให้การสรุปผลจากข้อมูลที่ได้นั้นจำเป็นต้องขยายผลสู่ประชากรเป้าหมาย ดังนั้นการใช้สถิติเพื่อการวิจัยโดยการเก็บรวบรวมข้อมูลจากการสุ่มตัวอย่างมีความจำเป็นต้องสรุปผลไปสู่ประชากร กระบวนการทางสถิติที่สำคัญ คือการใช้สถิติเชิงอนุมาน (inferential statistics) และหนึ่งในการวิเคราะห์ข้อมูลทางสถิติเชิงอนุมานนั้น คือ **การประมาณค่า** (estimation) เพื่อนำผลจากการวิเคราะห์ข้อมูลจากกลุ่มตัวอย่างสรุปผลไปสู่ประชากรที่ต้องการศึกษา ดังนั้นผู้วิจัยต้องใช้ข้อมูลเหล่านั้นเพื่อสรุปผลไปสู่คุณลักษณะของประชากร

4.1 ความหมายของการประมาณค่า

การประมาณค่าเป็นวิธีการหนึ่งที่ใช้ค่าสถิติ (statistics) ที่คำนวณค่าได้จากกลุ่มตัวอย่างเพื่อเป็นตัวประมาณค่าพารามิเตอร์ (parameter) ของประชากร ค่าสถิติที่ผู้วิจัยนิยมใช้ ได้แก่ ค่าเฉลี่ย ค่าสัดส่วน และค่าส่วนเบี่ยงเบนมาตรฐาน ค่าเหล่านั้นจะเป็นตัวแทนพื้นฐานของข้อมูลให้กับผู้วิจัยในการค้นหาคำตอบจากวัตถุประสงค์ของการวิจัยได้ เช่น ต้องการทราบรายได้เฉลี่ยของกลุ่มเป้าหมาย ค่าสัดส่วนของแร่ธาตุโปแตสเซียมในร่างกายของคนที่มีอายุมากกว่า 50 ปีขึ้นไป หรือยอดรวมโดยเฉลี่ยของสินค้าที่ผลิตได้ในไตรมาสที่ 3 เป็นต้น ซึ่งค่าประมาณเหล่านี้ไม่ใช่ค่าที่แท้จริงของข้อมูล เป็นเพียงการคาดคะเนของผู้วิจัยที่ต้องการทราบ ค่าต่างๆนั้นจะใกล้เคียงข้อเท็จจริงมากที่สุดขึ้นอยู่กับวิธีการเลือกกลุ่มตัวอย่างและจำนวนตัวอย่างเป็นสำคัญ ถ้าใช้วิธีการสุ่มตัวอย่างที่เหมาะสมและจำนวนข้อมูลที่เก็บรวบรวมมาพอเหมาะ ค่าประมาณที่ได้ก็จะใกล้เคียงค่าแท้จริงมากขึ้น

4.2 รูปแบบของการประมาณค่า

ในการวิจัย ผู้วิจัยสามารถทำการประมาณค่าข้อมูล ได้ 2 แบบ คือ

4.2.1 การประมาณค่าแบบจุด

การประมาณค่าแบบจุด (point estimation) เป็นวิธีการประมาณค่าลักษณะต่างๆของประชากรที่ผู้วิจัยต้องการทราบจากกลุ่มตัวอย่างด้วยค่าเพียงค่าเดียว ได้แก่

1) ค่าเฉลี่ย $(\overline{X})$ จากกลุ่มตัวอย่างเป็นตัวประมาณค่าเฉลี่ยของประชากร(μ)

2) ค่าสัดส่วน ($\hat{p}$) จากกลุ่มตัวอย่างเป็นตัวประมาณค่าสัดส่วนของประชากร (P)

3) ค่าส่วนเบี่ยงเบนมาตรฐาน (S) เป็นตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานของประชากร (σ)

ตัวอย่างเช่น ผู้วิจัยต้องการทราบค่าใช้จ่ายเฉลี่ยต่อเดือนของนักศึกษาชั้นปีที่ 1 ของมหาวิทยาลัยราชภัฏสวนสุนันทา เพื่อใช้เป็นข้อมูลในการตัดสินใจให้ผู้บริหารของสถานศึกษา ถ้าผู้วิจัยสุ่มตัวอย่างนักศึกษามา 150 คนพบว่าค่าใช้จ่ายเฉลี่ยต่อเดือนเป็น 6,000 บาท และนำผลที่ได้ไปสรุปอ้างอิงถึงประชากรทั้งหมด กรณีนี้แสดงให้เห็นว่าผู้วิจัยให้ความสำคัญกับค่าใช้จ่ายโดยเฉลี่ยมากเกินไป ซึ่งมีโอกาสที่จะผิดพลาดสูง

หรือการประมาณค่าความพึงพอใจในการให้บริการของพนักงานขายเครื่องสำอางยี่ห้อหนึ่ง สุ่มตัวอย่างพนักงานขาย 100 คน พบว่าความพึงพอใจเฉลี่ยในการให้บริการเท่ากับ 4.5 ถ้าผู้วิจัยนำผลนี้ไปสรุปกับการให้บริการของพนักงานขายเครื่องสำอางยี่ห้อนี้ทั้งหมดว่าลูกค้ามีความพึงพอใจสูงมากแสดงให้เห็นว่าผู้วิจัยมีความมั่นใจในผลการวิจัยความพึงพอใจโดยเฉลี่ยมากเกินไป โอกาสที่จะเกิดความคลาดเคลื่อนมีได้มาก

สรุปได้ว่า การประมาณค่าแบบจุดจึงมีโอกาสความเสี่ยงในการผิดพลาดสูง เพราะการประมาณแบบนี้ไม่ได้กำหนดขนาดของความผิดพลาดและไม่ได้บอกความเชื่อมั่นว่าการประมาณค่าใกล้เคียงกับค่าแท้จริงเพียงใด ในการวิจัยจึงไม่ค่อยนิยมใช้ กันมากนัก

4.2.2 การประมาณค่าแบบช่วง

การประมาณค่าแบบช่วง (interval estimation) เป็นการประมาณค่าของประชากรที่คำนวณตัวเลขได้ออกมาเป็นช่วง โดยมีการกำหนดขนาดของความผิดพลาดหรือระดับความเชื่อมั่นไว้ เช่น การประมาณค่าใช้จ่ายเฉลี่ยต่อเดือนของนักศึกษาทั้งมหาวิทยาลัย อยู่ในช่วง 5,500 บาท ถึง 6,700 บาท ที่ระดับนัยสำคัญ 0.05

การประมาณค่าแบบช่วง หมายถึง การประมาณค่าพารามิเตอร์ออกมาเป็นช่วงตัวเลข ($a < \theta < b$) คือ มีค่าต่ำสุด (lower confidence limit = a) และค่าสูงสุด (upper confidence limit = b) ที่คาดว่าจะคลุมค่าพารามิเตอร์ที่จะประมาณค่า โดยที่ช่วงของการประมาณค่าอาจจะกว้างหรือแคบก็ได้

ถ้าช่วงของการประมาณค่าแคบ โอกาสที่ค่าพารามิเตอร์จะตกอยู่ในช่วงนั้นย่อมมีน้อย แต่ผลที่ได้รับจากการประมาณค่าจะมีความแม่นยำสูงและย่อมมีผลต่อการนำไปใช้ประโยชน์ได้ดี ถ้าช่วงของการประมาณค่ากว้างจะทำให้โอกาสที่ค่าพารามิเตอร์ตกอยู่ในช่วงนั้นมีมากขึ้น แต่ผลที่ได้รับจะมีความแม่นยำต่ำ มีประโยชน์ต่อการวิเคราะห์น้อย หรืออาจจะใช้ไม่ได้เลยถ้าช่วงนั้นกว้างเกินไป นอกจากนั้นถ้าเราใช้ตัวประมาณค่าต่างกัน แม้ว่าจะใช้ช่วงประมาณเท่ากันก็จะได้ค่าของการประมาณต่างกัน บางช่วงของการประมาณจะครอบคลุมค่า θ และบางช่วงจะไม่ครอบคลุมค่า θ Sample Distribution ของ θ จะช่วยให้สามารถหาค่าของ a และ b ที่ครอบคลุมค่า θ ได้ และการใช้ตัวประมาณค่าที่เท่ากันแต่ใช้ช่วงความเชื่อมั่นแตกต่างกันก็จะได้ช่วงการประมาณค่าแตกต่างกันด้วย ดังนั้นในการประมาณค่าแบบช่วงจึงต้องพิจารณาในเรื่องต่อไปนี้

1) การเลือกตัวประมาณค่า (estimation) ที่ดีที่มีคุณสมบัติครบทั้ง 4 ประการ ที่จะได้กล่าวต่อไป

2) การกำหนดสัมประสิทธิ์แห่งความเชื่อมั่น (coefficient of confidence)

กำหนดให้ช่วง (a , b) เป็นช่วงที่มีค่าที่แท้จริงของพารามิเตอร์ θ ตกอยู่ การหาค่า a , b นี้จะกำหนดในลักษณะเป็นเปอร์เซ็นต์เรียกว่า **ระดับความเชื่อมั่น** (confidence level) เช่น กำหนดระดับความเชื่อมั่น 95% หมายความว่า ถ้ากำหนดความน่าจะเป็นที่ θ จะอยู่ในช่วงนี้เป็น 0.95 แล้ว ช่วงนี้จะมีค่าอยู่ระหว่าง a และ b ดังนั้นค่าความน่าจะเป็นที่ค่าพารามิเตอร์ θ จะอยู่ในช่วง a กับ b นี้เรียกว่า **สัมประสิทธิ์แห่งความเชื่อมั่น** (coefficient of confidence) ระดับความเชื่อมั่นอาจเขียนอยู่ในรูป $(1-\alpha)$ 100% ในเมื่อ $0 < \alpha < 1$ และ $1-\alpha$ ก็คือ สัมประสิทธิ์แห่งความเชื่อมั่น โดยทั่วไปจะเขียน ระดับความเชื่อมั่น อยู่ในรูป $P(a < \theta < b) = 1-\alpha$

โดยที่ a = ขีดจำกัดความเชื่อมั่นล่าง (lower confidence limit)

b = ขีดจำกัดความเชื่อมั่นบน (upper confidence limit)

$1-\alpha$ = สัมประสิทธิ์แห่งความเชื่อมั่น (coefficient of confidence)

4.3 คุณสมบัติของตัวประมาณค่าที่ดี

ตัวประมาณค่า (estimator) จะมีคุณสมบัติ ดังนี้

4.3.1 ตัวประมาณค่าที่ไม่เอนเอียง

ถ้า $\hat{\theta}$ เป็นค่าสถิติที่ใช้เป็นตัวประมาณค่าของ θ เราจะเรียก θ ว่าเป็นตัวประมาณค่าที่ไม่เอนเอียง (unbiased estimator) ของ θ เมื่อ $E(\hat{\theta}) = \theta$ โดยที่ $E(\hat{\theta})$ หมายถึง ค่าคาดหวังทางคณิตศาสตร์ของ $\hat{\theta}$ ตัวอย่างเช่น ค่าสถิติ $\overline{X}$ เป็นตัวประมาณค่าที่ไม่เอนเอียงของพารามิเตอร์ μ เพราะ $E(\overline{X}) = \mu$

4.3.2 ตัวประมาณค่าที่คงเส้นคงวา

$\hat{\theta}$ เป็นตัวประมาณค่าที่คงเส้นคงวา(consistent estimator) ก็ต่อเมื่อ $\lim\limits_{n\to\alpha} p\left[\left|\hat{\theta}_n - \theta\right| > \in\right] = 0$ เมื่อ $\in$ เป็นจำนวนจริงบวก หรือกล่าวอีกนัยหนึ่งว่า ถ้าขนาดตัวอย่างเพิ่มขึ้น $\hat{\theta}$ จะมีค่าเข้าใกล้ θ มากขึ้น

4.3.3 ตัวประมาณค่าที่มีประสิทธิภาพ

$\hat{\theta}$ เป็นตัวประมาณค่าที่มีประสิทธิภาพ (efficient estimator) ก็ต่อเมื่อความแปรปรวน $\hat{\theta}$ มีน้อยกว่าตัวประมาณอื่น ๆ ของ θ หรือกล่าวได้ว่า การวัดความมีประสิทธิภาพวัดได้จากความแปรปรวน คือ ตัวประมาณใดจะเป็นตัวประมาณที่ดีหากความคลาดเคลื่อนต่ำ และความคลาดเคลื่อนของตัวประมาณดูได้จากความแปรปรวน เช่น X และ Y ต่างเป็นตัวประมาณค่าที่ไม่เอนเอียงของ θ จะสรุปว่า X มีประสิทธิภาพมากกว่า (most efficient estimator) ก็ต่อเมื่อความแปรปรวนของ X มีค่าน้อยกว่าความแปรปรวนของ y $\left(\sigma_x^2 < \sigma_y^2\right)$

4.3.4 ตัวประมาณค่าที่มีความเพียงพอ

ถ้า $\hat{\theta}$ เป็นตัวประมาณค่าของ θ แล้ว $\hat{\theta}$ เป็นตัวประมาณค่าที่มีความเพียงพอ (sufficient estimator) เมื่อ $\hat{\theta}$ ได้มาจากข้อมูลทุกตัวในตัวอย่าง เช่น ในการคำนวณหาค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานนั้น ใช้ค่าทุกค่าจากกลุ่มตัวอย่างในการคำนวณ จากสูตร $\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$ และ $S^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{X})^2}{n-1}$ ดังนั้น $\overline{X}$ และ S^2 เป็นตัวประมาณค่าที่มีความเพียงพอ

4.4 การประมาณค่าเฉลี่ยของประชากร

การประมาณค่าเฉลี่ยของประชากร ทำได้ 3 แบบ คือ

4.4.1 การประมาณค่าเฉลี่ยของประชากรแบบจุด

ใช้ค่าเฉลี่ย $(\overline{X})$ จากกลุ่มตัวอย่างเป็นตัวประมาณค่าเฉลี่ยของประชากร (μ) โดยคำนวณได้จากสูตร

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{………………..} \ (4.1)$$

เมื่อ x_i แทน ข้อมูลจากกลุ่มตัวอย่าง

n แทน จำนวนกลุ่มตัวอย่าง

ตัวอย่างที่ 4.1 สุ่มน้ำมะเขือเทศกระป๋องยี่ห้อหนึ่งมา 10 กระป๋อง ชั่งน้ำหนักได้ดังนี้ 14.5 , 16.3 , 15.1 , 15.7, 17.0 , 17.3 , 14.9 , 16.2 , 16.5 , 15.8 กรัม จงประมาณค่าน้ำหนักมะเขือเทศกระป๋องโดยเฉลี่ยแบบจุด

วิธีทำ

ให้ $\overline{X}$ เป็นค่าเฉลี่ยของน้ำหนักน้ำมะเขือเทศกระป๋องยี่ห้อหนึ่งที่สุ่มมา 10 กระป๋อง

สูตร $$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

จากโจทย์ $x_1, x_2, x_3, \dots, x_{10} = 14.5, 16.3, 15.1, \dots, 15.8$, $n = 10$

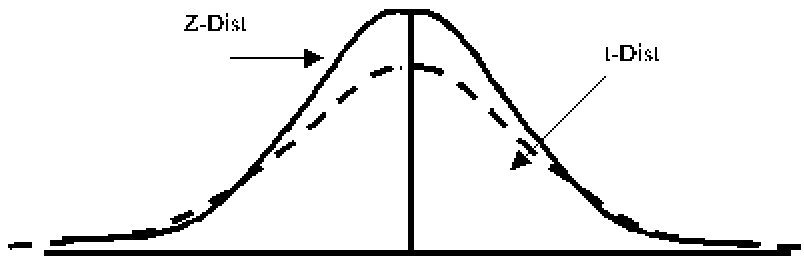
แทนค่า $$\overline{X} = \frac{14.5 + 16.3 + 15.1 + \dots + 15.8}{10}$$

$$= \frac{159.3}{10} = 15.93$$

นั่นคือ น้ำหนักเฉลี่ยของน้ำมะเขือเทศกระป๋องยี่ห้อหนึ่งจำนวน 10 กระป๋องเท่ากับ 15.93 กรัม แสดงว่าน้ำหนักเฉลี่ยของน้ำมะเขือเทศกระป๋องยี่ห้อนี้ประมาณ 15.93 กรัม

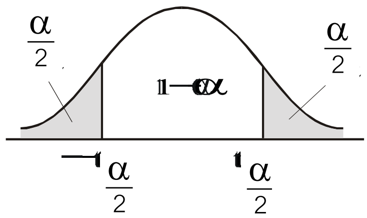
4.4.2 การประมาณค่าเฉลี่ยของประชากร 1 กลุ่มแบบช่วง

ในการประมาณค่าเฉลี่ยของประชากรกลุ่มเดียวแบบช่วง โดยปกติผู้วิจัยจะไม่ทราบค่าความแปรปรวนของประชากร และการแจกแจงของประชากรมีการแจกแจงแบบปกติหรือใกล้เคียงปกติ ดังนั้นในกรณีนี้การทดสอบโดยใช้สถิติ t จะช่วยในการประมาณค่าของข้อมูลได้ เนื่องจากการแจกแจงแบบ t จะมีลักษณะคล้ายหรือใกล้เคียงการแจกแจงแบบ Z ดังภาพที่ 4.1



ภาพที่ 4.1 เปรียบเทียบการแจกแจงของ Z และ t

เมื่อผู้วิจัยประมาณค่า σ^2 ด้วย S^2 จะได้ว่า $t=\dfrac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ มีการแจกแจงแบบที (t-distribution) ด้วยระดับขั้นความเสรี (df) เท่ากับ n-1



ภาพที่ 4.2 แสดง $P(-t_{\frac{\alpha}{2}} < t < t_{\frac{\alpha}{2}})$

จะได้ว่า
$$\begin{aligned}1-\alpha &= P(-t_{\frac{\alpha}{2}} < t < t_{\frac{\alpha}{2}})\\ &= P\left(-t_{\frac{\alpha}{2}} < \frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} < t_{\frac{\alpha}{2}}\right)\\ &= P\left(-t_{\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n}} < \overline{X}-\mu < t_{\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n}}\right)\\ &= P\left(\overline{X}-t_{\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n}} < \mu < \overline{X}+t_{\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n}}\right)\end{aligned}$$

ดังนั้น เมื่อประชากรมีการแจกแจงแบบปกติ หรือใกล้เคียงปกติ ไม่ทราบค่าความแปรปรวนของประชากร (σ^2) และ $n < 30$ ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ ค่าประมาณของค่าเฉลี่ยประชากร คือ

$$\overline{X}-t_{\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n}} < \mu < \overline{X}+t_{\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n}} \text{ หรือ } \mu=\overline{X}\pm t_{\frac{\alpha}{2}}\cdot\frac{S}{\sqrt{n}} \quad \ldots\ldots\ldots\ldots\ldots.. (4.2)$$

ตัวอย่างที่ 4.2 จากตัวอย่างที่ 4.1 จงประมาณช่วงความเชื่อมั่น 99 % ของน้ำหนักมะเขือเทศกระป๋องโดยเฉลี่ย

วิธีทำ

ให้ μ เป็นน้ำหนักมะเขือเทศกระป๋องโดยเฉลี่ย

สูตร $\mu = \bar{X} \pm t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$

จากโจทย์ $n = 10$, $\bar{X} = 15.93$, $\alpha = 0.01$, $df = 9$

และ คำนวณ S จากสูตร $S = \sqrt{\frac{\sum x^2}{n-1} - \frac{n\bar{x}^2}{n-1}}$

$$= \sqrt{\frac{2545.07}{9} - \frac{10(15.93)^2}{9}}$$

$$= \sqrt{282.79 - 281.96} = \sqrt{0.83} = 0.91$$

แทนค่า $\mu = \bar{X} \pm t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$

$$= 15.93 \pm 3.25(\frac{0.91}{\sqrt{10}})$$

$$= 15.93 \pm 3.25(0.29) = 15.93 \pm 0.94$$

$$= 14.99,\ 16.87$$

นั่นคือ น้ำหนักมะเขือเทศกระป๋องโดยเฉลี่ยของยี่ห้อนี้อยู่ในช่วง 14.99 กรัม และ 16.87 กรัม ที่ระดับนัยสำคัญ 0.01

ตัวอย่างที่ 4.3 จากการสอบถามแม่บ้าน 25 คนในหมู่บ้านจัดสรรแห่งหนึ่งเกี่ยวกับจำนวนวันที่ไปซื้อของในห้างสรรพสินค้า ปรากฏว่าได้จำนวนวันเฉลี่ยในการไปซื้อของเท่ากับ 14 วัน และส่วนเบี่ยงเบนมาตรฐาน 2 วัน จงประมาณแบบช่วงจำนวนวันเฉลี่ยของแม่บ้านที่ไปซื้อของในห้างสรรพสินค้าที่ระดับนัยสำคัญ .05

วิธีทำ

ให้ μ เป็นจำนวนวันเฉลี่ยของแม่บ้านที่ไปซื้อของในห้างสรรพสินค้า

สูตร $\mu = \bar{X} \pm t_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$

จากโจทย์ $n = 25$, $\bar{X} = 14$, $S = 2$, $\alpha = 0.05$, $df = 24$

แทนค่า $\mu = 14 \pm t_{.025} \frac{2}{\sqrt{25}} = 14 \pm 2.06 \frac{2}{\sqrt{25}}$

$$= 14 \pm 0.824 = 14.824,\ 13.0176$$

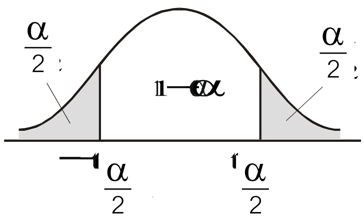
นั่นคือ จำนวนวันของแม่บ้านที่ไปซื้อของในห้างสรรพสินค้าโดยเฉลี่ยอยู่ในช่วง 13.02 วัน ถึง 14.82 วัน ที่ระดับนัยสำคัญ .05

4.4.3 การประมาณค่าเฉลี่ยของประชากร 2 กลุ่ม แบบช่วง

ในการประมาณค่าเฉลี่ยของประชากร 2 กลุ่มนิยมประมาณค่าเฉลี่ยโดยการหาค่าความแตกต่างของค่าเฉลี่ยของประชากรทั้งสองค่าว่ามีความแตกต่างกันมากเพียงใด โดยการประมาณผลต่างค่าเฉลี่ยแบบช่วงของประชากร 2 กลุ่ม $(\mu_1 - \mu_2)$ ในกรณีที่ไม่ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2) ผู้วิจัยสามารถแบ่งการคำนวณได้ 2 แบบ คือ

1) ไม่ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2) แต่ทราบว่า $\sigma_1^2 = \sigma_2^2$ เมื่อสุ่มตัวอย่างขนาด n_1 และ n_2 มาโดยอิสระกันจาก 2 ประชากร ที่มีการแจกแจงแบบปกติ มีค่าเฉลี่ย μ_1 และ μ_2 ซึ่งต่างไม่ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2) โดย n_1 และ $n_2 < 30$

ถ้ากำหนดให้ $\sigma_1^2 = \sigma_2^2 = \sigma_0^2$ โดยประมาณค่า σ_0^2 จากความแปรปรวนร่วมของกลุ่มตัวอย่าง (pooled sample variance) ใช้สัญลักษณ์ว่า S_p^2 ซึ่ง S_p^2 เป็นตัวประมาณค่าที่ไม่เอนเอียงของ σ_0^2 โดยที่ $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$ และจะมีการแจกแจงแบบ t จากสูตร $t = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$



ภาพที่ 4.3 แสดง $P(-t_{\frac{\alpha}{2}} < t < t_{\frac{\alpha}{2}})$

จากภาพที่ 4.3 จะได้ว่า $1\text{-}\alpha = P(-t_{\frac{\alpha}{2}} < t < t_{\frac{\alpha}{2}})$

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

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

ดังนั้น ค่าประมาณแบบช่วงของผลต่างระหว่างค่าเฉลี่ยประชากรทั้ง 2 กลุ่ม $(\mu_1 - \mu_2)$ เมื่อกำหนดว่า ไม่ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2) แต่ทราบว่า $\sigma_1^2 = \sigma_2^2$ และตัวอย่างมีขนาดเล็ก $n_1 < 30, n_2 < 30$ ที่ระดับความเชื่อมั่น $(1\text{-}\alpha)100\%$ คือ

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

หรือ $$\mu_1 - \mu_2 = (\overline{X}_1 - \overline{X}_2) \pm t_{\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$ ………………………...(4.3)

โดยที่ $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 1)S_2^2}{n_1 + n_2 - 2}$ และ t มีค่าระดับขั้นความเสรี $(df) = n_1 + n_2 - 2$

ตัวอย่างที่ 4.4 จากการสุ่มตัวอย่างคะแนนวิชาสถิติธุรกิจของนักศึกษาการจัดการสาขานิเทศศาสตร์และสาขาการบริหารการจัดการ จำนวนสาขาละ 28 คน พบว่าคะแนนเฉลี่ยของทั้ง 2 สาขาเป็น 54 คะแนน และ 60 คะแนน ตามลำดับ และค่าส่วนเบี่ยงเบนมาตรฐานของคะแนนเป็น 7 และ 10 คะแนน ตามลำดับ จงประมาณผลต่างของคะแนนเฉลี่ยของนักศึกษาทั้ง 2 สาขาที่ระดับความเชื่อมั่น 99 % ถ้าทราบว่าค่าความแปรปรวนของคะแนนวิชาสถิติธุรกิจมีค่าไม่แตกต่างกัน

วิธีทำ

ให้ μ_1 และ μ_2 เป็นคะแนนเฉลี่ยวิชาสถิติธุรกิจของนักศึกษาการจัดการสาขานิเทศศาสตร์และสาขาการบริหารการจัดการ

และ μ - μ_2 เป็นผลต่างคะแนนเฉลี่ยวิชาสถิติธุรกิจของนักศึกษาการจัดการสาขานิเทศศาสตร์และสาขาการบริหารการจัดการ

สูตร $$\mu_1 - \mu_2 = (\overline{X}_1 - \overline{X}_2) \pm t_{\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

จากโจทย์ $n_1 = 28, n_2 = 28, \overline{X}_1 = 54, \overline{X}_2 = 60, S_1^2 = 49, S_2^2 = 100,$

$\alpha = .01$ และ $\sigma_1^2 = \sigma_2^2$

$$\because \quad S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(27)(49) + (27)(100)}{28 + 28 - 2} = 74.5$$

$$\therefore \quad S_p = 8.6313$$

$$\therefore \quad \mu_1 - \mu_2 = (54 - 60) \pm t_{.005,\,54}(8.6313)\sqrt{\frac{1}{28} + \frac{1}{28}}$$

$$= -6 \pm (2.75)(8.6313)(0.26726) = -12.3437, 0.3437$$

นั่นคือ ผลต่างคะแนนเฉลี่ยวิชาสถิติธุรกิจของนักศึกษาการจัดการสาขานิเทศศาสตร์และสาขาการบริหารการจัดการ อยู่ในช่วง 0.34 คะแนน ถึง 12.34 คะแนน ที่ระดับความเชื่อมั่น 99 %

2) ไม่ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2) แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$ สุ่มตัวอย่างขนาดเล็ก $(n_1 < 30, n_2 < 30)$ อย่างเป็นอิสระต่อกันจากประชากร 2 กลุ่ม ซึ่งต่างก็มีการแจกแจงแบบปกติ และไม่ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2) แต่มีลักษณะการกระจายต่างกัน จึงทำให้ค่า $\sigma_1^2 \neq \sigma_2^2$ และประมาณค่า σ_1^2 ด้วย S_1^2, σ_2^2 ด้วย S_2^2 จะได้การแจกแจงแบบ t จากสูตร

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

โดยมีค่าระดับขั้นความเสรี (ν)

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_1}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$$

ดังนั้น ค่าประมาณแบบช่วงของผลต่างระหว่างค่าเฉลี่ยประชากรทั้ง 2 กลุ่ม $(\mu_1 - \mu_2)$ เมื่อกำหนดว่า ไม่ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2) แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$ และตัวอย่างมีขนาดเล็ก $(n_1 < 30, n_2 < 30)$ ที่ระดับความเชื่อมั่น $(1\text{-}\alpha)100\%$ คือ

$$(\overline{X}_1 - \overline{X}_2) - t_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \; < \; \mu_1 - \mu_2 \; < \; (\overline{X}_1 - \overline{X}_2) + t_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

หรือ

$$\mu_1 - \mu_2 = (\overline{X}_1 - \overline{X}_2) \pm t_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (4.4)$$

โดยมีค่า

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$$

ตัวอย่างที่ 4.5 จากตัวอย่างที่4.4 จงประมาณผลต่างของคะแนนเฉลี่ยของนักศึกษาทั้ง 2 สาขาที่ระดับความเชื่อมั่น 99% ถ้ากำหนดค่าความแปรปรวนของคะแนนวิชาสถิติธุรกิจมีค่าแตกต่างกัน

วิธีทำ

ให้ μ_1 และ μ_2 เป็นคะแนนเฉลี่ยวิชาสถิติธุรกิจของนักศึกษาการจัดการสาขานิเทศศาสตร์และสาขาการบริหารการจัดการ

และ $\mu_1 - \mu_2$ เป็นผลต่างคะแนนเฉลี่ยวิชาสถิติธุรกิจของนักศึกษาการจัดการสาขานิเทศศาสตร์และสาขาการบริหารการจัดการ

สูตร $$\mu_1-\mu_2=(\overline{X}_1-\overline{X}_2)\pm t_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

จากโจทย์ $n_1=28, n_2=28, \overline{X}_1=54, \overline{X}_2=60, S_1=7, S_2=10, \alpha=.01$

และ $\sigma_1^2 \neq \sigma_2^2$

$$\nu=\frac{(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2})^2}{\frac{(\frac{S_1^2}{n_1})^2}{n_1-1}+\frac{(\frac{S_2^2}{n_2})^2}{n_2-1}}$$

$$=\frac{(\frac{7^2}{28}+\frac{10^2}{28})^2}{\frac{(\frac{7^2}{28})^2}{27}+\frac{(\frac{10^2}{28})^2}{27}} \quad =48.34\approx 49$$

แทนค่า $$\mu_1-\mu_2=(54-60)\pm t_{.005,\,49}\sqrt{\frac{7^2}{28}+\frac{10^2}{28}}$$

$$=-6\pm(2.75)(2.3068) \quad =-6\pm 6.3437$$

$$=-12.3437, 0.3437$$

นั่นคือ ผลต่างคะแนนเฉลี่ยวิชาสถิติธุรกิจของนักศึกษาการจัดการสาขานิเทศศาสตร์และสาขาการบริหารการจัดการ อยู่ในช่วง 0.34 คะแนน ถึง 12.34 คะแนน ที่ระดับความเชื่อมั่น 99 %

4.5 การประมาณค่าสัดส่วนของประชากร (P)

ในการประมาณค่าบางเรื่องนิยมประมาณค่าในรูปเชิงเปรียบเทียบ เช่น ประชาชนในเขตกทม.นิยมใช้แชมพูยี่ห้อสมุนไพรA คิดเป็น 35 % นั่นคือสัดส่วนของตลาดที่คนในกทม.นิยมใช้แชมพูยี่ห้อสมุนไพร A เท่ากับ 0.35 ทำได้ 2 แบบ

4.5.1 การประมาณค่าสัดส่วนประชากรแบบจุด

ใช้ค่าสัดส่วนของกลุ่มตัวอย่าง ($\hat{p}$) เป็นตัวประมาณค่าของค่าสัดส่วนประชากร (P) โดยคำนวณได้จากสูตร

$$\hat{p}=\frac{x}{n} \quad \text{..........................(4.5)}$$

เมื่อ x คือ จำนวนหน่วยตัวอย่างที่สนใจ

n คือ จำนวนกลุ่มตัวอย่าง

ตัวอย่างที่ 4.6 ผู้วิจัยสำรวจตลาดในเขตชานเมืองเพื่อตรวจสอบพฤติกรรมการเลือกซื้อผักปลอดสารพิษ จากกลุ่มตัวอย่าง 500 คน พบว่านิยมเลือกซื้อผักปลอดสารพิษ 295 คน จงประมาณค่าสัดส่วนของคนที่มีพฤติกรรมการเลือกซื้อผักปลอดสารพิษ

วิธีทำ

ให้ $\hat{p}$ เป็นสัดส่วนของคนที่มีพฤติกรรมการเลือกซื้อผักปลอดสารพิษจำนวน 500 คน

สูตร $$\hat{p} = \frac{x}{n}$$

จากโจทย์ $x = 295$, $n = 500$

จะได้ $$\hat{p} = \frac{295}{500} = 0.59$$

นั่นคือ สัดส่วนของคนที่มีพฤติกรรมการเลือกซื้อผักปลอดสารพิษจากสุ่มตัวอย่าง 500 คน เท่ากับ 0.5 แสดงว่า สัดส่วนของคนที่มีพฤติกรรมการเลือกซื้อผักปลอดสารพิษทั้งหมด เท่ากับ 0.59

4.5.2 การประมาณค่าสัดส่วนประชากรแบบช่วงของประชากร 1 กลุ่ม

จากทฤษฎีลิมิตเข้าสู่ส่วนกลาง (central limit theorem) จะได้ว่า ถ้าตัวอย่างมีขนาดใหญ่ ค่าสัดส่วนตัวอย่าง $(\hat{p})$ จะมีการแจกแจงโดยประมาณแบบปกติ โดยมีค่าเฉลี่ย P และความแปรปรวน $\frac{PQ}{n}$ เนื่องจากไม่ทราบค่า P ต้องการประมาณค่า จึงใช้ $\hat{p}\hat{q}$ แทน PQ ในเทอมของความแปรปรวน จะได้ว่า $$Z = \frac{\hat{p} - P}{\sqrt{\frac{\hat{p}\hat{q}}{n}}}$$

ดังนั้น ค่าประมาณแบบช่วงของค่าสัดส่วนประชากร (P) ที่ระดับความเชื่อมั่น $(1-\alpha)$ 100% คือ

$$\hat{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < P < \hat{p} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \text{ หรือ } P = \hat{p} \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} \quad \ldots\ldots\ldots\ldots (4.6)$$

ตัวอย่างที่ 4.7 จากการทดสอบประสิทธิภาพของยาชนิดหนึ่งในการรักษาโรค คนไข้ 800 คน ได้รับการรักษาโรคโดยให้ยาชนิดนี้ และหายจากโรค ในเวลาอันรวดเร็ว จำนวน 512 คน จงหาช่วงความเชื่อมั่น 95% ของสัดส่วนของคนไข้ที่หายจากโรคโดยใช้ยาประเภทนี้

วิธีทำ

ให้ P เป็นสัดส่วนของคนไข้ที่หายจากโรคโดยใช้ยาประเภทนี้

และ $\hat{p}$ เป็นสัดส่วนของคนไข้ที่หายจากโรคโดยใช้ยาประเภทนี้จากคนไข้ 800 คน

สูตร $$\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < P < \hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

จากโจทย์ $$\hat{p}=\frac{512}{800}=0.64,\ \alpha=.05,\ n=800$$

แทนค่า $$0.64-Z_{.025}\sqrt{\frac{(0.64)(0.36)}{800}} < P < 0.64+Z_{.025}\sqrt{\frac{(0.64)(0.36)}{800}}$$

$$0.64-1.96\sqrt{\frac{(0.64)(0.36)}{800}} < P < 0.64+1.96\sqrt{\frac{(0.64)(0.36)}{800}}$$

$$0.64-0.0333 < P < 0.64+0.0333$$

$$0.6067 < P < 0.6733$$

นั่นคือ สัดส่วนของคนไข้ที่หายจากโรคโดยใช้ยาประเภทนี้อยู่ในช่วง 0.61 ถึง 0.67 ด้วยความเชื่อมั่น 95%

4.5.3 การประมาณค่าผลต่างระหว่างค่าสัดส่วนแบบช่วงประชากร2 กลุ่ม

เนื่องจากการประมาณค่าสัดส่วนประชากรจะต้องใช้ขนาดตัวอย่างใหญ่ ดังนั้นจากทฤษฎีลิมิตสู่ส่วนกลาง (central limit theorem) จะได้ว่า $\hat{p}_1$ มีการแจกแจงโดยประมาณแบบปกติด้วยค่าเฉลี่ย P_1 ค่าความแปรปรวน $\frac{P_1Q_1}{n_1}$ ในทำนองเดียวกัน $\hat{p}_2$ จะมีการแจกแจงแบบปกติด้วยค่าเฉลี่ย P_2 ค่าความแปรปรวน $\frac{P_2Q_2}{n_2}$

$$\therefore \quad P_1-P_2 \sim \text{normal } (P_1-P_2, \frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2})$$

หรือ $$Z=\frac{(\hat{p}_1-\hat{p}_2)-(P_1-P_2)}{\sqrt{\frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2}}} \sim \text{normal } (0,1)$$

แต่เราไม่ทราบค่า P_1,P_2 จึงต้องประมาณค่าความแปรปรวน $\frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2}$ ด้วย $\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}$

ดังนั้น ค่าประมาณแบบช่วงของผลต่างระหว่างค่าสัดส่วนของประชากร 2 กลุ่ม (P_1-P_2) ที่ระดับความเชื่อมั่น (1-α)100% คือ

$$(\hat{p}_1-\hat{p}_2)-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}} < P_1-P_2 < (\hat{p}_1-\hat{p}_2)+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}} \quad \text{หรือ}$$

$$P_1-P_2=(\hat{p}_1-\hat{p}_2)\pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}} \quad \ldots\ldots\ldots\ldots\ldots (4.7)$$

ตัวอย่างที่ 4.8 สุ่มตัวอย่างครอบครัวในจังหวัดเชียงใหม่และกทม. มาอย่างละ 500 ครอบครัว เพื่อตรวจสอบจำนวนโทรศัพท์มือถือที่ใช้ในแต่ละครอบครัว ปรากฏว่าในจังหวัดเชียงใหม่และกทม. มีครอบครัวที่มีโทรศัพท์ใช้เกิน 3 เครื่องอยู่ 215 และ 350 ครอบครัว ตามลำดับ จงประมาณผลต่างของสัดส่วนดังกล่าวนี้ ที่ช่วงความเชื่อมั่น 90 %

วิธีทำ

ให้ P_1 และ P_2 เป็นสัดส่วนของครอบครัว ที่มีโทรศัพท์ใช้เกิน 3 เครื่องในจังหวัดเชียงใหม่และกทม.

และ $P_1 - P_2$ เป็นผลต่างระหว่างสัดส่วนของครอบครัวที่มีโทรศัพท์ใช้เกิน 3 เครื่องใน 2 จังหวัด

สูตร $$P_1 - P_2 = (\hat{p}_1 - \hat{p}_2) \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

จากโจทย์ $$n_1 = n_2 = 500, \hat{p}_1 = \frac{215}{500} = 0.43, \hat{q}_1 = 1 - 0.43 = 0.57,$$

$$\hat{p}_2 = \frac{350}{500} = 0.70, \hat{q}_2 = 0.30, \alpha = .10$$

แทนค่า $$P_1 - P_2 = (0.43 - 0.70) \pm Z_{.05}\sqrt{\frac{(0.43)(0.57)}{500} + \frac{(0.7)(0.3)}{500}}$$

$$= -0.27 \pm 1.645(0.032) \qquad = -0.27 \pm 0.05$$

$$= -0.32,\ 0.22$$

นั่นคือ ผลต่างระหว่างสัดส่วนของครอบครัวที่มีโทรศัพท์ใช้เกิน 3 เครื่องในจังหวัดเชียงใหม่และกทม.อยู่ในช่วง 0.22 ถึง 0.32 ที่ช่วงความเชื่อมั่น 90 %

4.6 การใช้โปรแกรมคอมพิวเตอร์ช่วยในการวิเคราะห์ข้อมูล

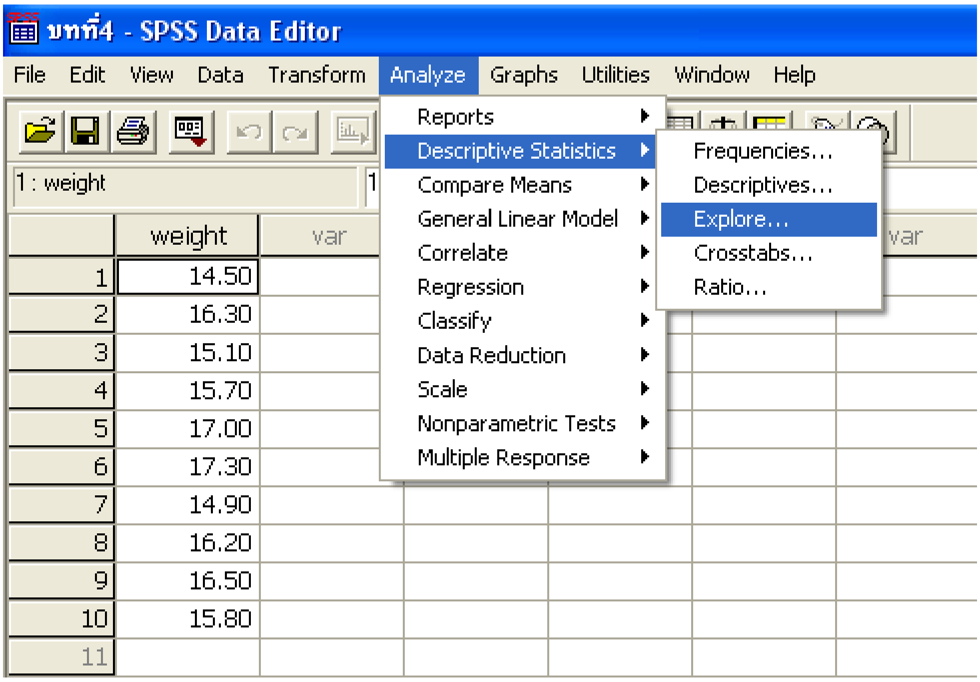
การวิเคราะห์ข้อมูลจากการสุ่มตัวอย่างเพื่อประมาณค่าประชากรนั้นบางการวิจัยมักนิยมทำการสุ่มตัวอย่างเนื่องจากประชากรมีจำนวนมาก และตัวอย่างที่สุ่มก็มีจำนวนมากเช่นกัน ดังนั้นการใช้เครื่องมือช่วยในการคำนวณจึงมีความจำเป็นในการใช้โปรแกรมสำเร็จรูป ซึ่งมีอยู่มากมาย โดยในที่นี้จะกล่าวถึงการใช้โปรแกรมสำเร็จรูป SPSS ช่วยในการวิเคราะห์ข้อมูล และการประมาณค่าสามารถใช้คำสั่งคอมพิวเตอร์วิเคราะห์ข้อมูลดังนี้

คำสั่ง Analyze ⟶ Descriptive Statistics ⟶ Explore….

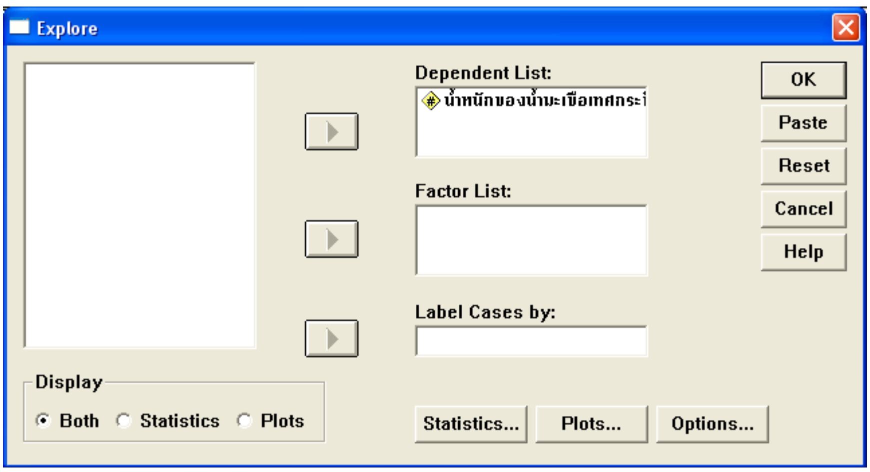
ตัวอย่าง 4.9 สุ่มน้ำมะเขือเทศกระป๋องยี่ห้อหนึ่งมา 10 กระป๋อง ชั่งน้ำหนักได้ดังนี้ 14.5 , 16.3 , 15.1 , 15.7, 17.0 , 17.3 , 14.9 , 16.2 , 16.5 , 15.8 กรัม จงประมาณค่าแบบจุดและช่วงความเชื่อมั่น 99 % ของน้ำหนักมะเขือเทศกระป๋องโดยเฉลี่ย

วิธีทำ โจทย์ให้ประมาณค่าแบบจุดและช่วงความเชื่อมั่น 99 % ของน้ำหนักมะเขือเทศกระป๋องโดยเฉลี่ย มีขั้นตอนดังต่อไปนี้

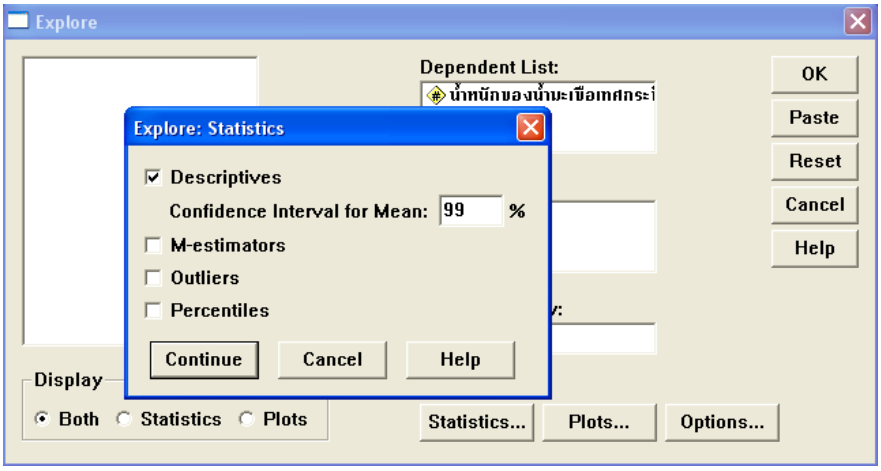
1) ที่เมนูคำสั่งเลือก Analyze ⟶ Descriptive Statistics ⟶ Explore….



2) เลือกตัวแปรน้ำหนักของน้ำมะเขือเทศกระป๋อง (Weight) ไปใส่ในช่อง Dependent List: และช่อง Display คลิก Both



3) คลิกปุ่ม Statistics เลือก Descriptive และกำหนด Confidence Interval for Mean: เป็น 99% แล้วกด Continue



4) เมื่อกำหนดเงื่อนไขต่างๆตามต้องการแล้ว คลิก OK

ผลลัพธ์จากโปรแกรมสำเร็จรูป

<u>ส่วนที่ 1</u>

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
น้ำหนักของน้ำมะเขือเทศกระป๋องโดยเฉลี่ย	10	100.0%	0	.0%	10	100.0%

จากตารางจำนวนน้ำมะเขือเทศกระป๋องมี 10 กระป๋อง คิดเป็น 100 % (ดูจากค่าValid) ส่วนค่าสูญหาย (missing) เท่ากับ 0 แสดงว่าไม่มีน้ำหนักน้ำมะเขือเทศใดเลยที่นำมาคิดคำนวณไม่ได้

<u>ส่วนที่ 2</u>

Descriptives

			Statistic	Std. Error
น้ำหนักของน้ำมะเขือเทศกระป๋องโดยเฉลี่ย	Mean		15.9300	.28715
	99% Confidence Interval for Mean	Lower Bound	14.9968	
		Upper Bound	16.8632	
	5% Trimmed Mean		15.9333	
	Median		16.0000	
	Variance		.825	
	Std. Deviation		.90805	
	Minimum		14.50	
	Maximum		17.30	
	Range		2.80	
	Interquartile Range		1.5750	
	Skewness		-.102	.687
	Kurtosis		-.889	1.334

สรุปผลจากตารางข้างต้นได้ว่า น้ำหนักเฉลี่ยของน้ำมะเขือเทศกระป๋องยี่ห้อหนึ่ง เท่ากับ 15.93 กรัม แบบจุด (mean) และแบบช่วงอยู่ระหว่าง 14.9968กรัม (lower bound) ถึง 16.8632 กรัม (upper bound) ที่ความเชื่อมั่น 99 %

4.7 สรุป

การประมาณค่าเป็นวิธีการหนึ่งทางสถิติเพื่อหาค่าตัวแทนของข้อมูลที่ผู้วิจัยเก็บรวบรวมมาได้ และสรุปผลนำไปสู่คุณลักษณะที่ต้องการของประชากร โดยมีคุณสมบัติที่ดี 4 ประการ คือ เป็นตัวประมาณค่าที่ไม่เอนเอียง มีความคงที่คงเส้นคงวา มีประสิทธิภาพ และมีความเพียงพอ การประมาณค่าทำได้ 2 รูปแบบ คือ แบบจุด และแบบช่วง ทั้งนี้ขึ้นอยู่กับวัตถุประสงค์ของการวิจัย รวมทั้งสามารถประมาณค่าเฉลี่ย ค่าสัดส่วนของประชากร 1 กลุ่ม และ 2 กลุ่มได้

4.8 คำถามทบทวน

1. สมาคมกอล์ฟท้องถิ่นแห่งหนึ่งรายงานผลคะแนนการแข่งขันโดยสุ่มคะแนน 10 ครั้ง ได้ผลดังนี้ 72 76 81 75 73 85 79 73 74 82 ถ้าค่าส่วนเบี่ยงเบนมาตรฐานของคะแนนเป็น 4 จงประมาณช่วงความเชื่อมั่น 99 % ของคะแนนเฉลี่ยในการแข่งขันกอล์ฟท้องถิ่นแห่งนี้

2. จากการวิจัยการขาดสารอาหารของเด็กปฐมวัยในชนบทแห่งหนึ่ง ผู้วิจัยได้สุ่มตัวอย่างเด็กปฐมวัยจำนวน 150 คน จากเด็กก่อนเรียนทั้งหมด พบว่ามีเด็ก 70 คน เป็นโรคขาดสารอาหาร จงประมาณสัดส่วนของเด็กปฐมวัยที่ขาดสารอาหารในชนบทแห่งนี้ ที่ระดับความเชื่อมั่น 95 %

3. บริษัทแห่งหนึ่งทำการประเมินผลการทำงานของพนักงาน โดยทำการประเมินผลงานของแต่ละบุคคลเป็นแต้มคะแนน ถ้าสุ่มพนักงานของบริษัทแห่งนี้จำนวน 18 คน พบว่าแต้มคะแนนเฉลี่ยเท่ากับ 115 คะแนน และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 13 คะแนน จงคำนวณช่วงความเชื่อมั่น 99 % ของแต้มคะแนนเฉลี่ยของพนักงานในบริษัทแห่งนั้น

4. โรงงานผลิตอุปกรณ์กีฬาแห่งหนึ่ง ได้ทำการสุ่มตัวอย่างลูกขนไก่มาจำนวน 90 ลูก พบว่ามีลูกขนไก่ที่มีรอยตำหนิอยู่ 15 ลูก จงประมาณค่าสัดส่วนของลูกเทนนิสของโรงงานแห่งนี้แบบจุด

5. จากการสุ่มสัมภาษณ์ประชาชนในห้างสรรพสินค้าแห่งหนึ่งเกี่ยวกับการโฆษณาสินค้าชนิดหนึ่ง ปรากฏว่าผู้ถูกสัมภาษณ์จำนวน 85 คน จาก 250 คน สามารถจดจำสโลแกนของสินค้าชนิดนี้ได้ จงหาช่วงความเชื่อมั่น 95 % ของสัดส่วนของผู้ที่จดจำสโลแกนของสินค้าได้

6. ผู้ผลิตหลอดไฟฟ้ายี่ห้อหนึ่งต้องการประมาณราคาขายปลีกเฉลี่ยที่ร้านค้าในจังหวัดภูเก็ตรับไปจำหน่าย จึงสุ่มตัวอย่างร้านค้าที่รับหลอดไฟฟ้าของบริษัทแห่งนี้ไปจำหน่ายมา 20 ร้านปรากฏว่าได้ราคาขายปลีกดังนี้ 85 82 80 90 93 95 79 81 78 78 91 85 84 81 87 87 85 80 97 80 จงหาค่าประมาณแบบจุดของราคาขายปลีกหลอดไฟฟ้าของร้านค้าในจังหวัดภูเก็ต

7. คณะกรรมาธิการป้องกันอาชญากรรมของสภาผู้แทนราษฎร์ได้ประมาณการเกิดอาชญากรรมด้วยอาวุธปืนในพื้นที่ความเสี่ยงสูงของประเทศ โดยได้สุ่มข้อมูลจาก 300 แฟ้มข้อมูล พบว่าเป็นอาชญากรรมจากอาวุธปืน 180 แฟ้มข้อมูล จงประมาณสัดส่วนการเกิดอาชญากรรมด้วยอาวุธปืนในพื้นที่ความเสี่ยงสูง

8. จากการสำรวจการลดน้ำหนักของลูกค้าที่เข้ามาใช้บริการในสถานเสริมความงามแห่งหนึ่ง พบว่าในช่วงเวลา 4 สัปดาห์ของการเข้าโปรแกรมลดน้ำหนัก ลูกค้า 25 คนสามารถลดน้ำหนักได้ 3 5 4 3 4 3 6 8 3 4 4 6 7 6 5 9 8 8 6 3 6 7 6 5 5 กิโลกรัม ถ้าความแปรปรวนของน้ำหนักเท่ากับ 81 กิโลกรัม จงหาช่วงความเชื่อมั่น 95 % ของน้ำหนักเฉลี่ยที่หายไปในช่วงเวลา 4 สัปดาห์ของลูกค้าในสถานเสริมความงามแห่งนี้

9. โรงงานผลิตกระเป๋าแห่งหนึ่ง ซึ่งผลิตกระเป๋า 2 แบบ คือ แบบ ก และแบบ ข มีความสนใจในราคาขายปลีกของกระเป๋า 2 แบบนี้ จึงทำการสุ่มตัวอย่างร้านขายปลีก 15 ร้านที่ขายกระเป๋าแบบ ก ปรากฏว่าได้ราคาขายปลีกกระเป๋าเฉลี่ยใบละ 210 บาท โดยมีส่วนเบี่ยงเบนมาตรฐาน 25 บาท และสุ่มอีก 10 ร้านที่ขายกระเป๋าแบบ ข พบว่าราคาขายปลีกกระเป๋าเฉลี่ยใบละ 280 บาท โดยมีส่วนเบี่ยงเบนมาตรฐาน 20 บาท ภายใต้ข้อสมมุติที่ว่า ถ้าราคากระเป๋านี้มีการแจกแจงแบบปกติไม่ทราบค่าความแปรปรวนของประชากรแต่ทราบว่าเท่ากัน จงหา 99 % ช่วงความเชื่อมั่นในการประมาณผลต่างของราคาขายปลีกของกระเป๋า 2 แบบนี้

10. จากการสำรวจสภาพที่อยู่อาศัยใน 2 จังหวัด พบว่าในจังหวัดแรก 58 ครอบครัวจาก 90 ครอบครัวเช่าบ้านอยู่ อีกจังหวัดหนึ่งพบว่า 85 ครอบครัวจาก 120 ครอบครัวเช่าบ้านอยู่เช่นกัน จงประมาณแบบจุดและแบบช่วงของผลต่างระหว่างสัดส่วนของครอบครัวที่เช่าบ้านอยู่ใน 2 จังหวัดนี้ที่ระดับนัยสำคัญ 0.05

11. จอภาพคอมพิวเตอร์ที่ผลิตจากบริษัท A และ B มีอายุการใช้งานที่ใกล้เคียงการแจกแจงแบบปกติ โดยมีความแปรปรวนของอายุการใช้งานของทั้งสองบริษัทไม่เท่ากัน ถ้าสุ่มตัวอย่างจอภาพคอมพิวเตอร์ที่ผลิตจากบริษัท A และ B มา 15 และ 20 จอภาพ จากการทดสอบพบว่าอายุการใช้งานเฉลี่ยของจอภาพคอมพิวเตอร์ที่ผลิตจากบริษัท A และ B เท่ากับ 8 และ 7.5 ปี และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.8 และ 0.9 ปี ตามลำดับ จงประมาณค่าความแตกต่างของจอภาพคอมพิวเตอร์ที่ผลิตจากสองบริษัท โดยใช้ความเชื่อมั่น 95 %

12. บรรณาธิการหนังสือพิมพ์ธุรกิจฉบับหนึ่งต้องการประมาณราคาห้องพักโดยเฉลี่ยต่อ 1 คืน จึงได้สุ่มถามราคาห้องพัก 32 แห่ง จาก 1000 แห่ง ได้ข้อมูลดังนี้ 286 260 225 298 167 187 193 123 251 188 199 239 234 259 191 227 228 146 165 158 278 147 153 187 153 146 170 169 162 125 130 165 จงประมาณค่าราคาห้องพักโดยเฉลี่ยต่อ 1 คืน แบบจุดและแบบช่วงที่ระดับนัยสำคัญ 0.05 โดยใช้การวิเคราะห์จากโปรแกรมคอมพิวเตอร์

13. ข้อมูลที่ได้มาจากการสุ่ม โดย $n = 64, \bar{x} = 22.5$ และ $S = 3.4$ จงประมาณแบบจุดและแบบช่วงของข้อมูลนี้ที่ระดับนัยสำคัญ 0.05

14. ข้อมูลที่ได้มาจากการสุ่ม โดย $n = 81, \bar{x} = 44.25$ และ $S = 4.5$ จงประมาณแบบจุดและแบบช่วงของข้อมูลนี้ที่ระดับนัยสำคัญ 0.05

15. ในการชำระภาษีเงินได้ของครอบครัวในชนบท โดยสุ่มครอบครัวที่ชำระภาษีเงินได้มา 36 ครอบครัว จาก 1990 ครอบครัว ได้ดังนี้ 2160 1001 7200 9280 3060 6300 1300 2150 2700 8400 1900 4450 2560 1000 3220 4530 2550 2400 2790 7280 4500 7600 2750 1300 5000 5820 9500 2860 9700 6170 1810 1540 2780 9300 1510 1600 1650 จงประมาณการชำระภาษีเงินได้โดยเฉลี่ยของครอบครัวในชนบท แบบจุดและแบบช่วงที่ระดับนัยสำคัญ 0.05 โดยใช้การวิเคราะห์จากโปรแกรมคอมพิวเตอร์

เอกสารอ้างอิง

Cochran, W.G. (1953). **Sampling techniques.** New York: John Wiley & Sons.

Dixon, W. J. and Massey, F. J. (1969). **Introduction to statistical analysis** (3rd ed.). Tokyo: McGraw – Hill Kogakusha.

Kerlinger, F. N. (1973). **Foundations of behavioral research** (2nd ed.). New York: Holt. Rinehart & Winston.

Kohler, H. (2002). **Statistics for business and economics** (2nd ed.). New York : Thomson Learning.

Mann, P.S. (1995). **Introductory statistics** (2nd ed.). New York: John Wiley & Sons.