

บทที่ 5

การประมาณค่าของข้อมูลด้านคุณภาพ (Estimation)

วิธีการที่ใช้ในการประมาณค่า

- การประมาณค่าซึ่งเป็นวิธีการทางสถิติที่ใช้ในการประมาณค่าพารามิเตอร์ ของประชากร เช่น
 - ค่าเฉลี่ย
 - ค่าสัดส่วนด้วยค่าประมาณ (estimate)

ตัวอย่างที่เลือกมาเป็นตัวแทนของประชากรจะมี 2 วิธี

1. การประมาณแบบจุด(Point Estimation)
2. การประมาณแบบช่วง(Interval Estimation)

ค่าพารามิเตอร์และค่าสถิติที่สำคัญ

ค่าวัดลักษณะ	ประชากร (พารามิเตอร์)	ตัวอย่าง (ค่าสถิติ)
ค่าเฉลี่ย	μ	\bar{x}
ส่วนเบี่ยงเบนมาตรฐาน	σ	S.D. หรือ S
ความแปรปรวน	σ^2	S^2
สัดส่วน	P	\hat{p}
จำนวน หรือขนาด	N	n

ระดับความเชื่อมั่น (Confidence Level)

หมายถึง โอกาสที่พารามิเตอร์ของประชากรจะอยู่ในช่วง
ของค่าที่ประมาณได้ แทนด้วย $(1 - \alpha)100\%$
หรือเรียกว่า ช่วงความเชื่อมั่น (Confidence Interval)

เรียก α ว่า ความผิดพลาดที่เกิดขึ้น

ช่วงความเชื่อมั่น 99%

$$\alpha = 0.01$$

ช่วงความเชื่อมั่น 95%

$$\alpha = 0.05$$

$$Z_{\infty/2} = 1.96$$

ช่วงความเชื่อมั่น 90%

$$\alpha = 0.10$$

$$Z_{\infty/2} = 1.645$$

ตาราง ค่า Z สำหรับ $\infty/2$ ที่นิยมใช้ประมาณค่า

ช่วงความเชื่อมั่น	α	พื้นที่ปลาย	ค่า $Z_{\infty/2}$
99 %	0.01	2.58	0.005
95 %	0.05	1.96	0.025
90 %	0.10	1.645	0.05
98 %	0.02	2.33	0.01

การประมาณค่า

1. การประมาณค่าเฉลี่ยประชากร 1 กลุ่ม
2. การประมาณผลต่างค่าเฉลี่ยของประชากร 2 กลุ่ม
3. การประมาณค่าสัดส่วนประชากร 1 กลุ่ม
4. การประมาณผลต่างสัดส่วนของประชากร 2 กลุ่ม

การประมาณช่วงความเชื่อมั่นของค่าเฉลี่ยประชากร 1 กลุ่ม

กรณีที่ 1 ทราบความแปรปรวนของประชากร(ทราบ σ^2)

$$\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

กรณีที่ 2 ไม่ทราบความแปรปรวนของประชากร

แต่ตัวอย่างมีขนาดใหญ่

(ไม่ทราบ σ^2 แต่ $n \geq 30$)

$$\bar{x} - Z_{\alpha/2} \frac{S}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2} \frac{S}{\sqrt{n}}$$

ตัวอย่างที่ 5.1 (จาก ตย.4.18)

น้ำหนักของขนมชนิดหนึ่งที่ผลิตโดยบริษัทผู้ผลิต
ขนมรายหนึ่งมีการแจกแจงใกล้เคียงกับการแจกแจง
ปกติ ด้วยส่วนเบี่ยงเบนมาตรฐาน 1.5 กรัม ถ้าเลือก
ตัวอย่างสุ่มจากขนมที่ผลิตโดยโรงงานแห่งนี้มา
จำนวน 30 ชิ้น ได้ข้อมูลดังนี้
จงประมาณช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยของ
ประชากรของน้ำหนักขนมที่ผลิตจากโรงงานแห่งนี้

ตัวอย่างที่ 5.1 กำหนดให้ $\bar{x} = \frac{896.49}{30} = 29.88$ $Z_{\alpha/2} = 1.96$ $n = 30$ $\sigma = 1.5$

$$\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$29.88 - 1.96 \left(\frac{1.5}{\sqrt{30}} \right) < \mu < 29.88 + 1.96 \left(\frac{1.5}{\sqrt{30}} \right)$$

$$29.88 - 1.96 \left(\frac{1.5}{5.47} \right) < \mu < 29.88 + 1.96 \left(\frac{1.5}{5.47} \right)$$

$$29.88 - 1.96 (0.274) < \mu < 29.88 + 1.96 (0.274)$$

$$29.88 - 0.537 < \mu < 29.88 + 0.537$$

$$29.34... < \mu < ...30.42....$$

ตอบ ดังนั้นค่าเฉลี่ยของน้ำหนักขนมประมาณ 29.34 กรัม ถึง 30.42 กรัม ต่อชิ้น ที่ระดับความเชื่อมั่น 95%

ตัวอย่างที่ 5.2 กำหนดให้ $\bar{x} = \frac{540}{32} = 16.88$ $Z_{\alpha/2} = 1.645$ $n = 32$ $\sigma = 2$

$$\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$16.88 - 1.645 \left(\frac{2}{\sqrt{32}} \right) < \mu < 16.88 + 1.645 \left(\frac{2}{\sqrt{32}} \right)$$

$$16.88 - 1.645 \left(\frac{2}{5.65} \right) < \mu < 16.88 + 1.645 \left(\frac{2}{5.65} \right)$$

$$16.88 - 1.645 (0.354) < \mu < 16.88 + 1.645 (0.354)$$

$$16.88 - 0.582 < \mu < 16.88 + 0.582$$

$$16.29... < \mu < ...30.46....$$

ตอบ ดังนั้นค่าเฉลี่ยของน้ำหนักขนมประมาณ 16.29 กรัม ถึง 30.46 กรัม ต่อชิ้น ที่ระดับความเชื่อมั่น 90%

การประมาณช่วงความเชื่อมั่นของผลต่าง ค่าเฉลี่ยของประชากร 2 กลุ่ม

กรณีที่ 1 ทราบความแปรปรวนของประชากรทั้งสองกลุ่ม
(ทราบ σ_1^2, σ_2^2)

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

กรณีที่ 2 ไม่ทราบความแปรปรวนของประชากรทั้ง

สองกลุ่ม แต่ตัวอย่างมีขนาดใหญ่

(ไม่ทราบ σ_1^2, σ_2^2 และ $n_1, n_2 \geq 30$)

$$(\bar{x}_1 - \bar{x}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

ตัวอย่างที่ 5.3 นักเศรษฐศาสตร์คนหนึ่งทำการสุ่มเลือกร้านค้าในบริเวณห้างสรรพสินค้า ก จำนวน 32 ร้าน พบว่ามีรายได้เฉลี่ยต่อเดือนเท่ากับ 6 ล้านบาท ส่วนเบี่ยงเบนมาตรฐาน 1 ล้านบาท และสุ่มเลือกจากห้างสรรพสินค้า ข จำนวน 30 ร้าน พบว่ามีรายได้เฉลี่ยต่อเดือนเท่ากับ 3.5 ล้านบาท ส่วนเบี่ยงเบนมาตรฐาน 5 ล้านบาท จงประมาณช่วงความเชื่อมั่น 90% ของผลต่างของรายได้เฉลี่ยต่อเดือนของร้านค้าในห้างสรรพสินค้าทั้งสองแห่ง

วิธีทำ จากโจทย์

$$n_1 = 32$$

$$n_2 = 30$$

$$\bar{x}_1 = 6$$

$$\bar{x}_2 = 3.5$$

$$S_1 = 1$$

$$S_2 = 5$$

$$\alpha = 0.10$$

จากข้อมูล จึงใช้กรณีที่ 2 ในการประมาณค่า และจากการเปิดตาราง ได้ค่า $Z_{\frac{\alpha}{2}} = 1.645$

$$(\bar{x}_1 - \bar{x}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$(6 - 3.5) - 1.645 \sqrt{\frac{1^2}{32} + \frac{5^2}{30}} < \mu_1 - \mu_2 < (6 - 3.5) + 1.645 \sqrt{\frac{1^2}{32} + \frac{5^2}{30}}$$

$$2.5 - 1.645 \sqrt{0.03 + 0.83} < u < 2.5 + 1.645 \sqrt{\frac{1}{32} + \frac{25}{30}}$$

$$2.5 - 1.645 \sqrt{0.86} < u < 2.5 + 1.645 \sqrt{0.86}$$

$$2.5 - 1.645 (0.93) < u < 2.5 + 1.645 (0.93)$$

$$2.5 - 1.53 < \mu_1 - \mu_2 < 2.5 + 1.53$$

$$0.97 < \mu_1 - \mu_2 < 4.03$$

ดังนั้น ร้านค้าในห้างสรรพสินค้าทั้งสองจะมีรายได้เฉลี่ยต่างกันประมาณเดือนละ

0.97 ล้านบาท ถึง 4.03 ล้านบาท ที่ระดับความเชื่อมั่น 90%

ตัวอย่างที่ 5.4 จำนวนจุดตำหนิของสินค้าที่เครื่องจักรผลิตได้ในโรงงานทั้งสองเครื่องมีการแจกแจงแบบปกติ โดยเครื่องที่ 1 มีส่วนเบี่ยงเบนมาตรฐาน 0.05 จุด และเครื่องที่สอง มีส่วนเบี่ยงเบนมาตรฐาน 0.1 จุด ทำการสุ่มเลือกสินค้าที่ผลิตโดยเครื่องจักรเครื่องที่ 1 มา 100 ชิ้น แล้วตรวจสอบพบว่ามียุจุดตำหนิเฉลี่ย 0.8 จุด และสุ่มเลือกสินค้าที่ผลิตโดยเครื่องที่ 2 จำนวน 200 ชิ้น พบว่ามีจุดตำหนิเฉลี่ย 0.5 จุด จงประมาณช่วงความเชื่อมั่น 99% ของผลต่างของจำนวนจุดตำหนิเฉลี่ยของสินค้าที่ผลิตโดยเครื่องจักรทั้งสอง

วิธีทำ จากโจทย์

$$\sigma_1 = 0.05$$

$$\sigma_2 = 0.1$$

$$n_1 = 100$$

$$n_2 = 200$$

$$\bar{x}_1 = 0.8$$

$$\bar{x}_2 = 0.5$$

$$\alpha = 0.01$$

จากข้อมูล จึงใช้กรณีที่ 1 ในการประมาณค่า และจาก
การเปิดตาราง ได้ค่า $Z_{\infty/2} = 2.58$

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

$$(0.8 - 0.5) - 2.576 \sqrt{\frac{(0.05)^2}{100} + \frac{(0.1)^2}{200}} < \mu_1 - \mu_2 < (0.8 - 0.5) + 2.576 \sqrt{\frac{(0.05)^2}{100} + \frac{(0.1)^2}{200}}$$

$$0.3 - 0.02 < \mu_1 - \mu_2 < 0.3 + 0.02$$

$$0.28 < \mu_1 - \mu_2 < 0.32$$

ดังนั้นจำนวนจุดจำหน่ายเฉลี่ยของสินค้าที่ผลิตโดยเครื่องจักรทั้งสอง
ต่างกันประมาณ 0.28 ถึง 0.32 จุด ที่ระดับความเชื่อมั่น 99%

การประมาณช่วงความเชื่อมั่นของสัดส่วนประชากร 1 กลุ่ม

ในกรณีที่สุ่มตัวอย่างจำนวน n สิ่ง จากประชากรจำนวน N

ให้ x แทนจำนวนตัวอย่างที่สนใจ

\hat{p} แทนสัดส่วนของกลุ่มตัวอย่างที่สนใจ

\hat{q} แทนสัดส่วนของกลุ่มตัวอย่างที่ไม่สนใจ

ซึ่ง $\hat{p} = \frac{x}{n}$ และ $\hat{q} = 1 - \hat{p}$

ดังนั้น การประมาณช่วงความเชื่อมั่น $(1 - \alpha)100\%$ ของ P คือ

$$\hat{p} - Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} < P < \hat{p} + Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

ตัวอย่างที่ 5.5 จากตัวอย่างสุ่มแก้วน้ำเซรามิกที่ผลิตโดยโรงงานแห่งหนึ่งจำนวน 500 ร้อยใบ พบว่ามีแก้วที่ไม่สามารถพิจารณาให้เป็นสินค้าเกรด A จำนวน 152 ใบ จงหาช่วงความเชื่อมั่น 90% สัดส่วนของแก้วน้ำเซรามิกที่ไม่ใช่เกรด ของโรงงานแห่งนี้

วิธีทำ ให้ \hat{p} แทน สัดส่วนของคนที่ใช้ผงซักฟอกยี่ห้อ A

\hat{q} แทน สัดส่วนของคนที่ไม่ได้ใช้ผงซักฟอกยี่ห้อ A

จะได้

$$\hat{p} = \frac{x}{n} = \frac{152}{500} = 0.304$$

$$\hat{q} = 1 - \hat{p} = 1 - 0.304 = 0.696$$

ที่ $\alpha = 0.05$ เปิดตาราง Z ได้ $Z_{\frac{\alpha}{2}} = 1.645$

แทนค่าสูตร ตัวอย่าง 5.5 (จาก ต.ย.4.19)

$$\hat{p} - Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} < P < \hat{p} + Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.304 - 1.645 \sqrt{\frac{0.304 \times 0.696}{500}} < P < 0.304 + 1.645 \sqrt{\frac{0.304 \times 0.696}{500}}$$

$$0.304 - 0.034 < P < 0.304 + 0.034$$

$$0.27 < P < 0.338$$

ดังนั้นสัดส่วนของแก้วน้ำเซรามิกที่ผลิตโดยโรงงานเหล่านี้มีสัดส่วนการผลิตที่แท้จริงที่ไม่ใช่เกรดA มีประมาณ 0.27 ถึง 0.338 หรือประมาณ ร้อยละ 22 ถึงร้อยละ 33.8 ที่ระดับความเชื่อมั่น 90%

แทนค่าในสูตร ตัวอย่างที่ 5.6 เมื่อ $n = 450$ $Z_{\infty/2} = 1.96$

$$\hat{p} - Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} < P < \hat{p} + Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$\hat{p} = \frac{x}{n} = \frac{200}{450} = 0.44 \text{ ดังนั้น } \hat{q} = 1 - 0.44 = 0.56$$

$$0.44 - 1.96 \sqrt{\frac{(0.44)(0.56)}{450}} < p < 0.44 + 1.96 \sqrt{\frac{(0.44)(0.56)}{450}}$$

$$0.44 - 1.96 \sqrt{0.0005}$$

$$0.44 - 0.43 < p < 0.44 + 0.43$$

$$0.01 < p < 0.87$$

ดังนั้นสัดส่วนของแก้วเซรามิกที่ไม่ใช่เกรด A มีประมาณ 0.01 ถึง

0.87 หรือประมาณร้อยละ 1 ถึงร้อยละ 87 ที่ระดับความเชื่อมั่น

การประมาณช่วงความเชื่อมั่นของผลต่างของสัดส่วน ประชากร 2 กลุ่ม

ให้ \hat{p}_1 แทน สัดส่วนของตัวอย่างที่สนใจกลุ่มที่ 1

\hat{p}_2 แทน สัดส่วนของตัวอย่างที่สนใจกลุ่มที่ 2

$$\hat{p}_1 = \frac{x_1}{n_1}$$

$$\hat{p}_2 = \frac{x_2}{n_2}$$

$$\hat{q}_1 = 1 - \hat{p}_1$$

$$\hat{q}_2 = 1 - \hat{p}_2$$

ดังนั้น การประมาณช่วงความเชื่อมั่น $(1 - \alpha)100\%$ ของ $P_1 - P_2$ คือ

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

ตัวอย่างที่ 5.7 จากการสุ่มตัวอย่างคนเชียงใหม่
ผู้ชาย 400 คน และผู้หญิง 500 คน พบว่า ชอบดู
รายการข่าวการเมือง จำนวน 348 คน และ 280
คน จงประมาณช่วงความเชื่อมั่น 99% ผลต่างของ
สัดส่วนระหว่างผู้ชายกับผู้หญิงที่ชอบดูข่าวการเมือง

วิธีทำ ให้ \hat{p}_1 แทน สัดส่วนของผู้ชายที่ชอบดูข่าวการเมือง

\hat{p}_2 แทน สัดส่วนของผู้หญิงที่ชอบดูข่าวการเมือง

จะได้

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{348}{400} = 0.87$$

$$\hat{q}_1 = 1 - \hat{p}_1 = 1 - 0.87 = 0.13$$

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{280}{500} = 0.56$$

$$\hat{q}_2 = 1 - \hat{p}_2 = 1 - 0.56 = 0.44$$

ที่ $\alpha = 0.01$ เปิดตาราง Z ได้ $Z_{\frac{\alpha}{2}} = 2.576$

แทนค่าในสูตร

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

$$(0.87 - 0.56) - 2.576 \sqrt{\frac{(0.87 \times 0.13)}{400} + \frac{(0.56 \times 0.44)}{500}} < P_1 - P_2 < (0.87 - 0.56) + 2.576 \sqrt{\frac{(0.87 \times 0.13)}{400} + \frac{(0.56 \times 0.44)}{500}}$$

$$0.31 - 0.07 < P_1 - P_2 < 0.31 + 0.07$$

$$0.24 < P_1 - P_2 < 0.38$$

ดังนั้นสัดส่วนของผู้ชายและผู้หญิงที่ชอบดูรายการนี้ต่างกันประมาณ 0.24 ถึง 0.38 หรือคิดเป็นร้อยละ 24 ถึง ร้อยละ 38 ที่ระดับความเชื่อมั่น 99%

$$(0.87 - 0.56) - 2.576 \sqrt{\frac{(0.87 \times 0.13)}{400} + \frac{(0.56 \times 0.44)}{500}} < P_1 - P_2 < (0.87 - 0.56) + 2.576 \sqrt{\frac{(0.87 \times 0.13)}{400} + \frac{(0.56 \times 0.44)}{500}}$$

$$\sqrt{\frac{0.1131}{400}} + \sqrt{\frac{0.2464}{500}}$$

$$\sqrt{0.0003} + \sqrt{0.0005}$$

$$\sqrt{0.0008} = 0.028$$

$$(0.87 - 0.56) - 2.58 (0.028)$$

$$0.31 - 0.07 < p_1 - p_2 < 0.31 - 0.07$$

