บทที่ 9

ทฤษฎีการประมาณค่า

(Estimation Theory)

เนื้อหาส่วนหนึ่งของสถิติอนุมานจะเกี่ยวกับการประมาณค่าพารามิเตอร์ โดยอาศัยค่าสถิติการประมาณค่า (estimation) เป็นการใช้ค่าสถิติ ซึ่งคำนวณได้จากกลุ่มตัวอย่างสำหรับเป็นตัวประมาณค่าพารามิเตอร์ (parameter) ของประชากร ทั้งนี้ก็เพื่อจะแสดงให้เห็นถึงลักษณะของประชากรที่ต้องการศึกษา

ค่าสถิติที่ใช้ประมาณค่าพารามิเตอร์ เรียกว่า ตัวประมาณค่า (estimator) เช่น $\bar{X}$, S^2, r เป็นตัวประมาณค่าของ μ, σ^2, ρ ตามลำดับ ค่าของพารามิเตอร์ที่ได้จากตัวประมาณค่าเรียกว่า ค่าประมาณ (estimate)

9.1 ประเภทของการประมาณค่าพารามิเตอร์

การประมาณค่าพารามิเตอร์ของประชากร ทำได้ 2 วิธี คือ

1. **การประมาณเป็นจุด** (Point estimation) เป็นการประมาณค่าพารามิเตอร์โดยกำหนดค่าออกมาเป็นตัวเลขโดด ๆ เช่นค่าของ $\bar{X}$ หรือ S^2 ค่าเหล่านี้เป็นค่าที่คำนวณได้จากข้อมูลของกลุ่มตัวอย่างซึ่งสุ่มมาจากประชากร

2. **การประมาณเป็นช่วง** (Interval estimation) เป็นการประมาณค่าพารามิเตอร์ว่าจะตกอยู่ในช่วงใด ในการประมาณเป็นช่วงตัวเลขแบบนี้ จะทำให้ค่าประมาณนั้นมีความเชื่อถือได้มากขึ้น ทั้งนี้เนื่องจากโอกาสที่ค่าพารามิเตอร์ที่แท้จริงจะตกอยู่ภายในช่วงนั้นมีความเป็นไปได้สูง

9.2 คุณสมบัติของตัวประมาณค่าที่ดี (Good estimator)

ตัวประมาณค่าที่ดีควรมีคุณสมบัติ ดังนี้

1. **ไม่ลำเอียง** (Unbias)

ถ้าให้ $\hat{\theta}$ เป็นตัวสถิติซึ่งคำนวณจากค่าของกลุ่มตัวอย่าง

$\hat{\theta}$ จะเป็นตัวประมาณที่ไม่ลำเอียงของ θ ก็ต่อเมื่อ $E(\hat{\theta}) = \theta$

ซึ่งหมายความว่าค่าเฉลี่ยหรือค่าคาดหวังที่คำนวณได้จากกลุ่มตัวอย่างมีค่าเท่ากับพารามิเตอร์

จากนิยามดังกล่าวจะได้ว่า

1. $\bar{X}$ เป็น unbias estimator ของ μ เพราะ $E(\bar{X}) = \mu$

2. $S^2 = \dfrac{\Sigma (X-\bar{X})^2}{n-1}$ เป็น unbias estimate ของ σ^2 เพราะ $E(S^2) = \sigma^2$

3. $S^2 = \dfrac{\Sigma (X-\bar{X})^2}{n}$ เป็น bias estimate ของ σ^2 เพราะ $E(S^2) = \dfrac{n-1}{n}\sigma^2 \neq \sigma^2$

ถ้าเพิ่มกลุ่มตัวอย่าง (n) ให้ใหญ่ขึ้น $E(S^2)$ จะเข้าใกล้ σ^2 ทั้งนี้เพราะ

$$\lim_{n\to\infty} E(S^2) = \left[\lim_{n\to\infty} \frac{n-1}{n}\sigma^2\right] = \sigma^2$$

กรณีเช่นนี้ เรียกว่า S^2 เป็น asymtotically unbiased estimator ของ σ^2

นั่นคือ S^2 เป็นตัวประมาณค่าที่ปราศจากความลำเอียงของ σ^2 ในกรณีที่กลุ่มตัวอย่างมีขนาดใหญ่

2. มีความคงเส้นคงวา (Consistency)

ตัวประมาณค่าที่ดีจะต้องเป็นตัวประมาณค่าที่มีความคงเส้นคงวา กล่าวคือ ควรมีค่า เข้าใกล้ค่าของพารามิเตอร์มากยิ่งขึ้น เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ นั่นคือค่าประมาณที่ได้มีความถูกต้องมากขึ้น เมื่อเพิ่มขนาดของกลุ่มตัวอย่าง นั่นคือ $|\hat{\theta} - \theta|$ จะต้องมีค่าน้อยลง เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ซึ่งสามารถเขียนสัญลักษณ์แทนได้ดังนี้

$$\lim_{n\to\infty} P\left[|\hat{\theta}-\theta| \leqslant \varepsilon\right] = 1 \text{ ทุก ๆ ค่าของ } \varepsilon > 0$$

เมื่อ ε แทนความคลาดเคลื่อนที่ยอมให้เกิดขึ้นในการประมาณค่าพารามิเตอร์

P แทน ความน่าจะเป็น (probability)

3. มีประสิทธิภาพ (Efficiency)

ถ้า $\hat{\theta}$ เป็นตัวประมาณค่าที่ไม่ลำเอียงของ θ และความแปรปรวนของ $\hat{\theta}$ มีค่าน้อยที่สุดเรียก $\hat{\theta}$ ว่าเป็น most efficient estimator ของ θ

นั่นคือถ้ามีสถิติ 2 ตัว ซึ่งสามารถใช้แทนค่าพารามิเตอร์ตัวเดียวกันได้เช่น $\bar{X}$, Mdn สถิติตัวที่มีค่าความคลาดเคลื่อนมาตรฐานน้อยกว่าจะเป็นตัวประมาณค่าที่มีประสิทธิภาพมากกว่าสถิติซึ่งมีความคลาดเคลื่อนมาตรฐานมากกว่า เช่น

ค่าความคลาดเคลื่อนมาตรฐานของมัธยฐาน $(\sigma^2_{\mathrm{Mdn}}) = \dfrac{\pi\sigma^2}{2n}$

ค่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย $(\sigma^2_{\bar{x}}) = \dfrac{\sigma^2}{2n}$

เนื่องจาก $\dfrac{\sigma^2}{2n}$ มีค่าน้อยกว่า $\dfrac{\pi\sigma^2}{2n}$ ดังนั้นค่าเฉลี่ย $(\bar{X})$ จึงเป็นตัวประมาณค่า μ ที่มีประสิทธิภาพดีกว่ามัธยฐาน

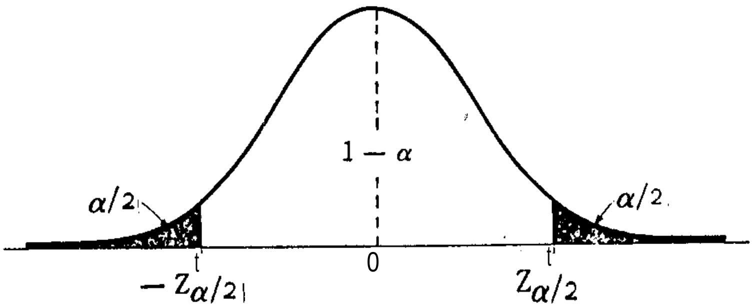
ค่าสถิติใดที่มีคุณสมบัติครบ 3 ประการคือไม่ลำเอียง มีความแปรปรวนน้อยที่สุด และมีความคงเส้นคงวา สถิตินั้นถือว่าเป็นตัวประมาณค่าที่ดี แต่ถ้ามีคุณสมบัติไม่ครบ สถิติใดที่มีคุณสมบัติมากกว่าจะถือว่าเป็นตัวประมาณค่าที่ดีกว่า

9.3 ช่วงความมั่นใจในการประมาณค่า (Level of confidence) หมายถึงความน่าจะเป็นที่ค่าพารามิเตอร์จะตกอยู่ในช่วงที่ได้จากการประมาณค่านั้น ส่วนความน่าจะเป็นที่ค่าพารามิเตอร์จะไม่ตกอยู่ในช่วงนั้น เรียกว่าระดับความมีนัยสำคัญ (level of significance) ซึ่งนิยมเขียนแทนด้วย α ดังนั้นช่วงความมั่นใจจะมีค่าเท่ากับ $1-\alpha$

9.3.1 **การประมาณช่วงของค่าเฉลี่ย (μ)**

ถ้าค่าเฉลี่ยของประชากรเป็น μ ความเบี่ยงเบนมาตรฐานเป็น σ จะใช้ $\bar{X}$ เป็นตัวสถิติที่จะประมาณค่า μ

ถ้ากำหนด confidence coefficient เท่ากับ $1-\alpha$ จากการแจกแจงปกติมาตรฐาน เราสามารถหาค่า Z ซึ่งมีพื้นที่ใต้โค้งเท่ากับ $1-\alpha$ ได้ ดังรูป



$$P\left(-Z_{\alpha/2} < z < Z_{\alpha/2}\right) = 1-\alpha$$

เราทราบว่า $z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ มีการแจกแจงปกติมาตรฐาน ดังนั้น

$$P\left(-Z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < Z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-Z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}} < \bar{X}-\mu < Z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

เอา $\bar{X}$ ลบอสมการในวงเล็บ แล้วคูณตลอดด้วย -1 จะได้

$$P\left(\bar{X}-Z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}+Z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

นั่นคือ $(1-\alpha)$ 100% confidence interval สำหรับ μ จะเป็น

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

ในกรณีที่ไม่รู้ค่า σ และ $n < 30$ สามารถประมาณช่วงของค่าเฉลี่ยได้จาก

$$P\left(-t_{\alpha/2} < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{\alpha/2}\right) = 1-\alpha$$

โดยวิธีการเดียวกับการแจกแจงของ Z สามารถหา $(1-\alpha)$ 100% confidence interval สำหรับ μ ได้ดังนี้

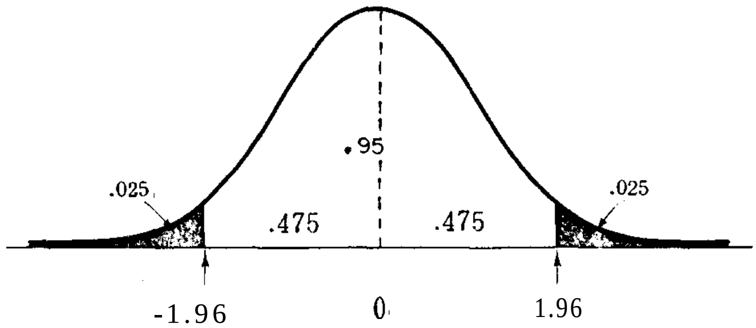
$$\bar{X}-t_{\alpha/2}\cdot\frac{S}{\sqrt{n}} < \mu < \bar{X}+t_{\alpha/2}\cdot\frac{S}{\sqrt{n}}$$

ตัวอย่าง 1 ค่าเกรดเฉลี่ยและความเบี่ยงเบนมาตรฐานของนักศึกษา 36 คนเป็น 2.6 และ 0.3 ตามลำดับ จงหา 95% confidence interval สำหรับเกรดเฉลี่ยของนักศึกษาทั้งหมด

จากโจทย์ $n = 36$, $\bar{X} = 2.6$, $S = 0.3$

เมื่อกำหนด $1-\alpha = .95$ ฉะนั้น $\alpha = .05$

หาค่า $Z_{\alpha/2}$ ได้ดังรูป



เนื่องจากพื้นที่ใต้โค้งจาก 0 ถึง $Z_{.025}$ เท่ากับ .475 (แบ่ง .95 เป็น 2 ส่วน) หาค่า Z จากตารางในภาคผนวกว่า Z จะมีค่าเท่าไร โดยดูตารางย้อนกลับ จะพบว่า $Z_{.025}$ มีค่าเท่ากับ 1.96 ดังนั้น

$$2.6 - (1.96)\left(\frac{0.3}{\sqrt{36}}\right) < \mu < 2.6 + (1.96)\left(\frac{0.3}{\sqrt{36}}\right)$$

จะได้ $2.50 < \mu < 2.70$

นั่นคือ เรามีความมั่นใจถึง 95% ว่าเกรดเฉลี่ยของนักศึกษาทั้งหมด จะมีค่าอยู่ระหว่าง 2.50 ถึง 2.70

ตัวอย่าง 2 ถ้าน้ำหนักเฉลี่ย และความเบี่ยงเบนมาตรฐานของนักเรียนจำนวน 25 คน เป็น 58 ก.ก. และ 10 ก.ก. ตามลำดับ จงหา 95% confidence interval สำหรับน้ำหนักเฉลี่ยของประชากร

จากโจทย์ $n = 25$ $\overline{X} = 58$ $S = 10$

$df = n-1 = 25-1 = 24$

เมื่อ $1-\alpha = .95$ ฉะนั้น $a = .05$

เปิดตารางหาค่า $t_{.025}$ ที่ df = 24 จะได้ค่า **t = 2.064**

ดังนั้นช่วงความเชื่อมั่น **95%** ของ μ คือ

$$58 - (2.064)\left(\frac{10}{\sqrt{25}}\right) < \mu < 58 + (2.064)\left(\frac{10}{\sqrt{25}}\right)$$

จะได้ $53.87 < \mu < 62.13$

นั่นคือเรามีความมั่นใจถึง 95% ว่าน้ำหนักเฉลี่ยของประชากรมีค่าอยู่ระหว่าง 53.87 ถึง 62.13 ก.ก.

9.3.2 การประมาณช่วงผลต่างของค่าเฉลี่ย

ถ้าประชากร 2 กลุ่มมีค่าเฉลี่ยและความเบี่ยงเบนมาตรฐานเป็น μ_1, μ_2 และ σ_1, σ_2 ตามลำดับ สุ่มตัวอย่างมาจำนวน n_1 และ n_2 จากประชากร 2 กลุ่มนี้ค่าสถิติ $\bar{X}_1 - \bar{X}_2$ จะเป็นตัวประมาณค่าที่มีประสิทธิภาพมากที่สุดของผลต่างระหว่าง μ_1 และ μ_2 ดังนั้นเราจึงใช้การแจกแจงของ $\bar{X}_1 - \bar{X}_2$ ในการหา confidence interval

เนื่องจาก $Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$

มีการแจกแจงปกติมาตรฐาน

เมื่อกำหนด α เราสามารถหาค่า $Z_{\alpha/2}$ ซึ่งทำให้

$$P\left(-Z_{\alpha/2} < Z < Z_{\alpha/2}\right) = 1 - \alpha$$

แทนค่า Z ในสมการจะได้

$$P\left(-Z_{\alpha/2} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} < Z_{\alpha/2}\right) = 1 - \alpha$$

จากนี้สามารถหา $(1-\alpha)$ 100% confidence interval สำหรับ $\mu_1 - \mu_2$ ได้ดังนี้

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

ในกรณีที่ไม่รู้ค่า σ แต่ $n \geqslant 30$ เราอาจใช้ S_1, S_2 แทน σ_1 และ σ_2 ได้ ดังนั้นการหา $(1-\alpha)$ 100% confidence interval จะเป็นดังนี้

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

แต่ถ้า n_1 และ n_2 มีค่าน้อย $(n < 30)$ การแจกแจงจะไม่เป็นการแจกแจงปกติมาตรฐาน ซึ่งเราจะใช้ค่าสถิติ Z ในการหา confidence interval ไม่ได้การแจกแจงจะเป็น 2 แบบ ดังนี้

1. ถ้าทราบว่า $\sigma_1 = \sigma_2 = \sigma$ (แต่ไม่ทราบค่า) จะใช้

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

ในการหา confidence interval ของความแตกต่างของประชากร 2 กลุ่ม

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

ดังนั้น $(1-\alpha)$ 100% confidence interval ของ $\mu_1 - \mu_2$ คือ

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

2. ถ้า $\sigma_1 \neq \sigma_2$ จะใช้

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

ในการหา confidence interval ของความแตกต่างของประชากร 2 กลุ่ม degree of freedom ในที่นี้จะเป็น

$$v = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{[(S_1^2/n_1)^2/(n_1 - 1)] + [(S_2^2/n_2)^2/(n_2 - 1)]}$$

ค่า v นี้ต้องปัดให้เป็นจำนวนเต็มที่ใกล้เคียง

และโดยวิธีการเดียวกันสามารถหา $(1-\alpha)$ 100% **confidence interval** ของ $\mu_1 - \mu_2$ ได้ดังนี้

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

ตัวอย่าง 8 แบบทดสอบฉบับหนึ่งนำมาสอบกับนักเรียนหญิง 50 คน และนักเรียนชาย 75 คน นักเรียนหญิงได้คะแนนเฉลี่ย 76 ความเบี่ยงเบนมาตรฐานเป็น 6 นักเรียนชายได้คะแนนเฉลี่ย 82 ความเบี่ยงเบนมาตรฐาน 8 จงหา 96% confidence interval ของผลต่างของคะแนนเฉลี่ยของนักเรียนชายและนักเรียนหญิง

จากโจทย์ $n_1 = 75 \qquad n_2 = 50$

$\bar{X}_1 = 82 \qquad \bar{X}_2 = 76$

$S_1 = 8 \qquad S_2 = 6$

และ $1 - \alpha = .96 \qquad \alpha = .04$

เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ เราใช้ S แทน σ และหาค่าจากการแจกแจงปกติมาตรฐานจากตารางจะได้ $Z_{.02} = 2.054$

ดังนั้น 96% confidence interval ของ $\mu_1 - \mu_2$ คือ

$$(82-76) - 2.054\sqrt{\frac{64}{75}+\frac{36}{50}} < \mu_1 - \mu_2 < (82-76) + 2.054\sqrt{\frac{64}{75}+\frac{36}{50}}$$

จะได้ $3.42 < \mu_1 - \mu_2 < 8.58$

นั่นคือ เรามีความมั่นใจถึง 96% ว่าผลต่างของคะแนนเฉลี่ยของนักเรียนชายและนักเรียนหญิง จะมีค่าอยู่ระหว่าง 3.42 ถึง 8.58

ตัวอย่าง 4 ในการทดลองใช้ตำราเรียน 2 เล่มกับนักเรียน 2 กลุ่ม ๆ ละ 20 คน หลังจากจบการทดลองแล้ว ได้ทำการทดสอบความสามารถในการเรียนของนักเรียน 2 กลุ่มนี้ ปรากฏคะแนนการทดสอบเป็นดังนี้

$\bar{X}_1 = 28 \qquad \bar{X}_2 = 23$

$S_1^2 = 120 \qquad S_2^2 = 80$

จงหา 95% confidence interval ของความแตกต่างของค่าเฉลี่ยของประชากรทั้ง 2 กลุ่ม

จากโจทย์ $df = n_1 + n_2 - 2 = 20 + 20 - 2 = 38$

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

$$= \frac{(19)(120)+(19)(80)}{38}$$

$$= 100$$

$t_{\alpha/2} = t_{.025}$ เปิดตารางหาค่า $t_{.025}$ ที่ degree of freedom 38 ได้ค่า $t = 2.02$

ดังนั้น 95% confidence interval ของ $\mu_1 - \mu_2$ คือ

$$(28-23)-2.02\sqrt{100\left(\frac{1}{20}+\frac{1}{20}\right)} < \mu_1-\mu_2 < (28-23)+2.02\sqrt{100\left(\frac{1}{20}+\frac{1}{20}\right)}$$

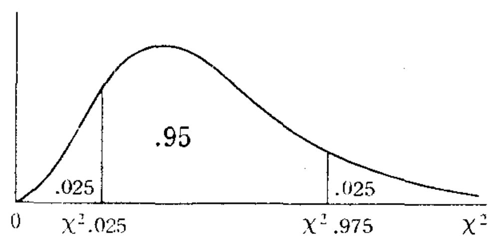
$$-1.39 < \mu_1-\mu_2 < 11.39$$

นั่นคือเรามีความมั่นใจถึง 95% ว่าความแตกต่างของคะแนนเฉลี่ยของประชากร 2 กลุ่มนี้มีค่าอยู่ระหว่าง -1.39 ถึง 11.39

9.3.3 การประมาณช่วงของความแปรปรวน

ในการแจกแจงความน่าจะเป็นของ $\chi^2 = \frac{(n-1)S^2}{\sigma^2}$ เราย่อมทราบได้ว่าความน่าจะเป็นที่ χ^2 จะอยู่ในช่วงหนึ่ง ๆ เป็นเท่าไร โดยดูจากค่าในตารางที่คำนวณไว้ ค่า χ^2_k หนึ่ง ๆ ในตารางจะหมายถึงค่า χ^2 ซึ่งกำหนดช่วง $(0, \chi^2_k)$ ซึ่งความน่าจะเป็น (หรือพื้นที่) ในช่วงนั้นเป็น k

เมื่อเป็นดังนี้เราอาจกล่าวได้ว่า เรามีความมั่นใจที่ระดับใดระดับหนึ่งว่า χ^2 จะอยู่ในช่วงใดช่วงหนึ่ง ตัวอย่างเช่น เรามีความมั่นใจ .95 ว่า χ^2 จะอยู่ในช่วง $(\chi^2_{.025}, \chi^2_{.975})$ ดังรูป



หรือ $\chi^2_{.025} < \chi^2 < \chi^2_{.975}$

$$\chi^2_{.025} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{.975}$$

เอา $(n-1)S^2$ หารตลอด

$$\frac{\chi^2_{.025}}{(n-1)\,S^2} < \frac{1}{\sigma^2} < \frac{\chi^2_{.975}}{(n-1)\,S^2}$$

หรือ $$\frac{(n-1)\,S^2}{\chi^2_{.975}} < \sigma^2 < \frac{(n-1)\,S^2}{\chi^2_{.025}}$$

นี่คือ ช่วงความมั่นใจ .95 ของ σ^2

การหาช่วงความมั่นใจระดับอื่น ๆ ก็ทำโดยนัยเดียวกัน

ตัวอย่าง 5 ถ้าต้องการจะประมาณค่าความแปรปรวนของคะแนนของประชากรในวิชาภาษาไทย ซึ่งมีผู้เข้าสอบ 31 คน และมี $S^2 = 60$ ที่ระดับ confidence interval 90%

$$df = n - 1 = 31 - 1 = 30$$

เนื่องจาก σ^2 มีค่าอยู่ระหว่าง

$$\frac{(n-1)S^2}{\chi^2_{.05}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{.95}}$$

เนื่องจาก $\chi^2_{.05}$ ที่ degree of freedom 30 มีค่า 43.773

และ $\chi^2_{.95}$ ที่ degree of freedom 30 มีค่า 18.493

จะได้

$$\frac{(30)\ (60)}{43.773} < \sigma^2 < \frac{(30)\ (60)}{18.493}$$

$$41.12 < \sigma^2 < 97.33$$

นั่นคือ เรามีความมั่นใจถึง 90% ว่าความแปรปรวนของประชากร จะมีค่าอยู่ระหว่าง 41.12 ถึง 97.33

แบบฝึกหัด 9

1. เมื่ออยากทราบว่าน้ำหนักเฉลี่ยของนักศึกษารามคำแหงทั้งหมดเป็นเท่าใด เราก็สุ่มนักศึกษารามคำแหงมา 100 คน หาน้ำหนักเฉลี่ยของ 100 คนนี้ สมมติว่าได้เท่ากับ 54.2 ก.ก. อยากทราบว่าค่า 54.2 ก.ก. นี้เป็นค่าประมาณที่ลำเอียงหรือไม่ลำเอียง เพราะเหตุใด
2. ค่าประมาณที่ลำเอียงจำเป็นที่จะต้องเป็นค่าประมาณที่ไม่มีประสิทธิภาพหรือไม่ เพราะเหตุใด
3. การกะประมาณแบบจุดดีกว่าการกะประมาณแบบช่วงจริงหรือไม่ เพราะเหตุใด
4. ค่าเฉลี่ยและความแปรปรวนของเกรดเฉลี่ยของนักศึกษารามคำแหง 50 คน เท่ากับ 2.3 และ 0.5 ตามลำดับ จงหาช่วงความเชื่อมั่น 95% ของเกรดเฉลี่ยของนักศึกษาทั้งหมด
5. เหล้าไวน์ยี่ห้อหนึ่งมีค่าเฉลี่ยปริมาตรบรรจุสุทธิต่อขวด 70.5 ซ.ล. ความเบี่ยงเบนมาตรฐาน 5.6 ซ.ล. จงหาช่วงความเชื่อมั่น 99% ของปริมาตรสุทธิของเหล้าไวน์ยี่ห้อนี้ซึ่งสุ่มมา 12 ขวด
6. ทำไมเราจึงต้อง pool variance
7. นักศึกษา 2 กลุ่ม เป็นกลุ่มชาย 30 คน และกลุ่มหญิง 55 คน สอบวิชา MR 371 ได้คะแนนเฉลี่ย 75 และ 61 ความแปรปรวนของคะแนนเป็น 100 และ 90 ตามลำดับ จงหาช่วงความเชื่อมั่น 95% ของผลต่างของคะแนนเฉลี่ยของนักศึกษาชายและนักศึกษาหญิง
8. ปลูกข้าวโพด 2 พันธุ์ พันธุ์ A ได้ข้าวโพดเฉลี่ย 752 ฝักต่อไร่ ความเบี่ยงเบนมาตรฐาน 45 ฝัก พันธุ์ B ได้ข้าวโพดเฉลี่ย 538 ฝักต่อไร่ ความเบี่ยงเบนมาตรฐาน 30 ฝัก จงหาช่วงความเชื่อมั่น 99% ของผลรวมของค่าเฉลี่ยของประชากรคือข้าวโพด 2 พันธุ์
9. สุ่มนักศึกษามา 40 คน พบว่าสมรสแล้ว 10 คน จงหาช่วงความเชื่อมั่น 95% ของอัตราส่วนของนักศึกษาที่สมรสแล้ว
10. สุ่ม I.C. มา 100 ตัว มีน้ำหนักเฉลี่ย 0.5 กรัม ความแปรปรวน .009 กรัม จงหาช่วงความเชื่อมั่น 95% ของความเบี่ยงเบนมาตรฐานของประชากร I.C.