บทที่ 8
การประมาณค่าของประชากร

วัตถุประสงค์

เพื่อให้นักศึกษามีความรู้เกี่ยวกับ วิธีการอนุมานทางสถิติ คุณสมบัติของตัวค่าประมาณทางสถิติ การใช้ความน่าจะเป็นในการประมาณค่าแบบช่วง วิธีการประมาณค่าต่าง ๆ ในประชากร โดยอาศัยข้อมูลจากตัวอย่าง

การอนุมานทางสถิติ คือกระบวนการที่เราใช้ค่าสังเกตจากตัวอย่าง เพื่อคาดคะเน ประมาณ หรือทดสอบค่าพารามิเตอร์ของประชากรที่ตัวอย่างถูกเลือกมา

การอนุมานทางสถิติที่อาศัยความน่าจะเป็นมีส่วนเกี่ยวข้องแบ่งได้เป็น 2 ประเภทใหญ่ ๆ คือ

1. การประมาณค่าของประชากร
2. การทดสอบสมมุติฐาน

การประมาณค่า คือการใช้ตัวสถิติจากตัวอย่างเพื่อประมาณค่าพารามิเตอร์ และตัวสถิตินี้แหละเรียกว่าตัวค่าประมาณ อย่างเช่น ค่ามัชฌิมของตัวอย่าง $\bar{X}$ เรียกว่าเป็นตัวค่าประมาณของค่ามัชฌิมของประชากร μ การประมาณค่าอาจแบ่งออกได้เป็น 2 ประเภทคือ

1. การประมาณแบบจุด
2. การประมาณแบบช่วง

การประมาณแบบจุด คือการใช้ค่าค่าเดียวที่ได้จากตัวอย่างประมาณค่าพารามิเตอร์ของประชากรหลักเกณฑ์สำคัญในการพิจารณาว่า จะเป็นตัวค่าประมาณที่ดีนั้นจะต้องเป็นไปดังนี้

1. ความไม่เอนเอียง (Unbiasedness)

นิยาม ถ้าค่าความหวังของตัวค่าประมาณ $\hat{\theta}$ มีค่าเท่ากับค่าจริงของพารามิเตอร์ θ ตัวค่าประมาณที่ใช้นั้นเป็นตัวค่าประมาณที่ไม่เอนเอียง นั้นคือ $E(\hat{\theta}) = \theta$ ตัวอย่างเช่น $E(\bar{x}) = \mu$ $E(\frac{x}{n}) = p$ เป็นต้น

2. *ความแนบนัย (Consistency)*

นิยาม ถ้า $\hat{\theta}$ เป็นตัวค่าประมาณ θ คือค่าพารามิเตอร์ $\hat{\theta}$ จะเป็นตัวค่าประมาณที่แนบนัยกับค่าพารามิเตอร์ θ หากเมื่อขนาดของตัวอย่างเพิ่มขึ้นมาก แล้วค่า $\hat{\theta}$ จะใกล้เคียงกับค่า θ

นั่นคือ ถ้า n คือขนาดของตัวอย่าง $\rightarrow \infty$ แล้ว

$P(\hat{\theta} \rightarrow \theta) \rightarrow 1$

ดังตัวอย่าง ค่า $\bar{x}$ เป็น Consistent estimator ของ μ

s^2 เป็น Consistent estimator ของ σ^2

3. *ประสิทธิภาพ (Efficiency)*

นิยาม ถ้าค่าสถิติ 2 ค่า ใช้ประมาณค่าพารามิเตอร์ตัวเดียวกัน ค่าสถิติตัวที่การแจกแจงตัวอย่างมีความคลาดเคลื่อนมาตรฐานน้อยกว่า เรียกว่าเป็นตัวค่าประมาณที่มีประสิทธิภาพมากกว่า

ตัวอย่าง ถ้าเราใช้ค่าเฉลี่ย $\bar{x}$ จากตัวอย่างกับค่ามัธยฐาน me จากตัวอย่างเพื่อประมาณค่าพารามิเตอร์หรือค่าเฉลี่ยของประชากร μ เราจะพบว่าความคลาดเคลื่อนมาตรฐาน $\bar{x}$ น้อยกว่าของ me

$$V(\bar{x}) = \sigma^2_{\bar{x}} = \frac{\sigma^2}{n} \; ; \; V(me) = \sigma^2_{me} = \frac{\pi \sigma^2}{2n}$$

$$\frac{\sigma^2_{\bar{x}}}{\sigma^2_{me}} = \frac{\frac{\sigma^2}{n}}{\frac{\pi \sigma^2}{2n}} = \frac{2}{\pi} = 0.64$$

ดังนั้น เราจึงกล่าวได้ว่า ประสิทธิภาพของ $\bar{x}$ มีมากกว่า me

4. *ความพอเพียง (Sufficiency)*

นิยาม ค่าสถิติที่มีความพอเพียงสำหรับค่าพารามิเตอร์ θ คือค่าสถิติที่สามารถใช้ประโยชน์จากข้อมูลเกี่ยวกับ θ ทั้งหมด ที่มีอยู่ในตัวอย่าง

ตัวอย่าง หากเรามีตัวอย่างของคะแนนนักศึกษา 30 คน ตัวค่าประมาณคะแนนเฉลี่ยของประชากรอาจจะเป็นคะแนนตัวที่ 1 ในตัวอย่างหรือคะแนนตัวที่ 15 หรือค่าเฉลี่ยของคะแนนตัวที่ 1 กับตัวที่ 30 ก็ได้ แต่ตัวประมาณค่าที่ใช้ประโยชน์จากข้อมูลทั้งหมดในตัวอย่าง คือค่าเฉลี่ยของตัวอย่าง $\bar{x}$ ดังนั้น $\bar{x}$ จึงเป็นตัวค่าประมาณที่พอเพียงสำหรับ μ

การประมาณค่าแบบช่วง

นิยาม กระบวนการที่เราจะเรียกว่าเป็นการประมาณแบบช่วงนี้คือ การนำเอาตัวค่าประมาณแบบจุดมาต่อเติมให้ได้ช่วงของค่าประมาณที่เราเชื่อว่าจะครอบคลุมค่าพารามิเตอร์ในประชากรด้วยความน่าจะเป็นที่เรารู้ได้ เราจะเรียกช่วงค่าประมาณนี้ว่า ช่วงความเชื่อมั่น

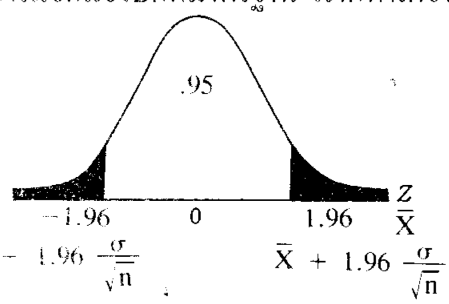
ตัวอย่าง ถ้าจะเปรียบการประมาณแบบจุดก็เหมือนกับการใช้ฉมวกแทงปลา การประมาณแบบช่วงก็เหมือนกับการใช้สุ่มจับปลา

การประมาณแบบช่วงสำหรับค่าเฉลี่ย ตัวอย่างขนาดใหญ่ $(n > 30)$

วิธีการที่เราจะหาช่วงความเชื่อมั่น ทำได้โดยการอาศัยค่าสถิติที่เราทราบว่ามีการแจกแจงแบบปกติมาตรฐาน Z

ในกรณีที่เราจะประมาณค่าเฉลี่ยของประชากร เราทราบว่าการแจกแจงค่าเฉลี่ย $\bar{X}$ ของตัวอย่าง จะมีการแจกแจงแบบปกติ หาก $n > 30$ (มีค่าเฉลี่ย μ ความแปรปรวน $\frac{\sigma^2}{n}$) ถ้าเราต้องการแปลงตัวแปร $\bar{X}$ ให้เป็นตัวแปร Z ที่มีการแจกแจงปกติมาตรฐาน เราก็ทำได้จากสูตร

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$



จากนี้เราก็สามารถหาความน่าจะเป็นของการที่ค่า Z จะมีค่าอยู่ระหว่าง 2 ค่าใดที่กำหนดในรูปของพื้นที่ใต้โค้งปกติระหว่างค่าทั้ง 2 นั้น ในทางกลับกัน หากเรากำหนดความน่าจะเป็นหรือพื้นที่ให้ได้ระหว่างค่า Z 2 ค่า เราก็อาจหาช่วงของค่า Z 2 ค่า ที่จะให้พื้นที่ต้องการได้

จากตารางเราทราบว่า พื้นที่ใต้โค้งปกติเท่ากับ 0.95 ระหว่างค่า $Z = -1.96$ กับ $Z = 1.96$ นั่นคือ

$$P(-1.96 < Z < +1.96) = 0.95$$

หากเราเขียน Z ในรูป $\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$ เพื่อหาว่า ช่วงของค่าอะไรที่ครอบคลุมค่า ด้วยความน่าจะเป็น 0.95 เราจะได้

$$P\left(-1.96 < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < +1.96\right) = 0.95$$

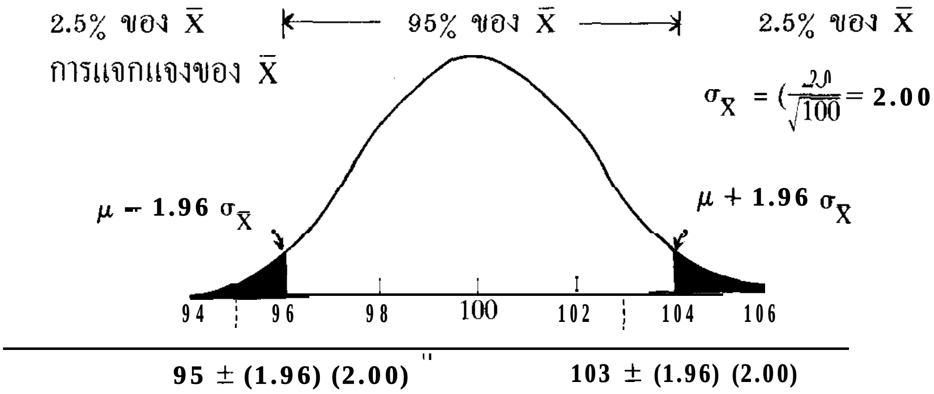
$$P\left(-1.96 \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < +1.96 \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$P\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

นั่นคือ ช่วงค่าประมาณ $\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}$ กับ $\bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$ จะครอบคลุมค่าพารามิเตอร์ μ ด้วยความน่าจะเป็น 0.95 ซึ่งความน่าจะเป็นอันนี้ เราจะเรียกว่า **ระดับความเชื่อมั่น** ส่วนช่วงค่าประมาณที่เราทราบความน่าจะเป็นนี้เรียกว่า **ช่วงความเชื่อมั่น** ระดับความเชื่อมั่นต่าง ๆ ที่พบกันบ่อย ๆ มี 0.90, 0.95 และ 0.99 ในกรณีตัวอย่างขนาดใหญ่ เราอาจใช้ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง s แทน σ ก็ได้

ตัวอย่างที่ 1 แสดงถึงช่วงความเชื่อมั่นที่ครอบคลุมค่า μ ได้อย่างไร ที่เลือกตัวอย่างขนาด 100 จากประชากรซึ่งมี $\mu = 100$, $\sigma = 20$ เราทราบว่าค่าเฉลี่ยเหล่านี้มีการแจกแจงปกติ สมมุติว่าในแต่ละครั้งของการเอาแต่ละค่าเฉลี่ย $\bar{X}$ และอ้างว่า μ จะวางอยู่ บางทีภายในช่วง $\bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}}$ ดังตัวอย่าง หาก $\bar{X} = 103$ เราก็สามารถกล่าวได้ว่า μ วางอยู่บางที่ภายในช่วง $103 \pm (1.96)\left(\frac{20}{\sqrt{100}}\right)$ หรือ 99.08 กับ 106.92 จะเห็นว่า ข้ออ้างของเราถูกต้อง อีกนัยหนึ่ง หาก $\bar{x} = 95$; μ ก็จะวางอยู่ภายในช่วง $95 \pm (1.96)\left(\frac{20}{\sqrt{100}}\right)$ หรือระหว่าง 91.08 กับ 98.92 ครั้งนี้ข้อกล่าวของเราไม่ถูกต้อง สองช่วงเหล่านี้แสดงที่ (ก) กับ (ข) ดังรูปข้างล่าง

หากค่าเฉลี่ย $\bar{x}$ อยู่ไม่ห่างออกไปจาก μ มากกว่า $\pm 1.96\frac{\sigma}{\sqrt{n}}$ ข้ออ้างของเราว่า μ จะอยู่ภายในช่วงจะถูกต้อง นอกจากนั้นจะไม่ถูกต้อง ดังรูปค่าเฉลี่ย $\bar{x}$ ซึ่งจะอยู่ไม่ห่างออกไปจาก μ มากกว่า $\pm 1.96\frac{\sigma}{\sqrt{n}}$ ก็คือส่วนที่ไม่แลเงาของการแจกแจง ประกอบด้วยค่าเฉลี่ย $\bar{x}$ 95% ของทั้งหมด ส่วนที่เหลือ 5% แสดงที่ปลายทั้งสองของการแจกแจงซึ่งนำไปสู่ช่วงที่ไม่ถูกต้อง



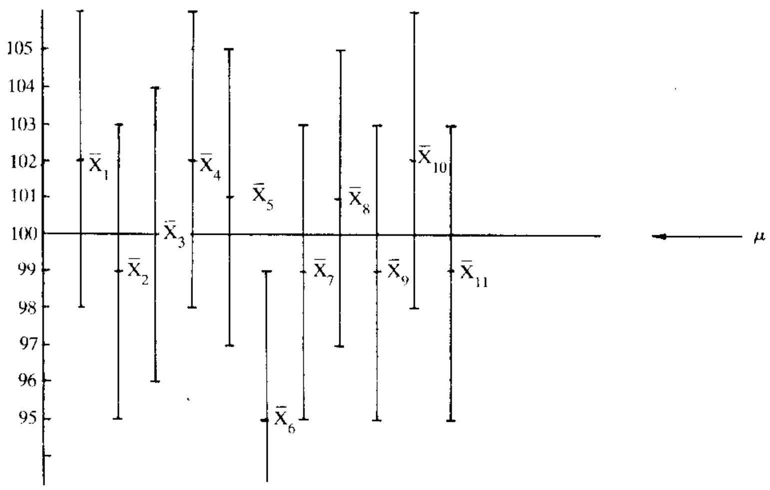
รูปที่ 1 การแจกแจงของ $\bar{x}$ ที่มีขนาดตัวอย่าง n = ๑๐๐ เลือกจากประชากรมี $\mu = 100$, $\sigma = 20$

(ก) ค่าประมาณแบบช่วงของ μ ที่ได้สร้างรอบ $\bar{X}$ = 95 ค่าประมาณไม่ได้ครอบคลุม μ

(ข) ค่าประมาณแบบช่วงของ μ ที่ได้สร้างรอบ $\bar{X}$ = 103 ค่าประมาณครอบคลุม μ

ตัวอย่างที่ 2 แสดงถึงช่วงบางช่วงหากเลือกตัวอย่างโดยสุ่ม

จากประชากรและสร้างช่วงรอบค่าเฉลี่ยของตัวอย่างตามกระบวนการเหมือนตัวอย่างที่ 1 เครื่องหมายที่กึ่งกลางของแต่ละเส้นแทนค่าเฉลี่ย $\bar{X}$ และขอบเขตของเส้นแทนแถบภายใน ซึ่งอ้างถึง μ วางอยู่ ในการแสดงนี้มีอยู่หนึ่งช่วงเท่านั้นที่ไม่ครอบคลุม μ ตามกระบวนที่กำหนดเราอาจคาดว่า 95 ตัวค่าประมาณของ 100 จะรวม μ อยู่ภายในช่วง



รูปที่ 2 การประมาณแบบช่วงของ μ ที่ได้สร้างจากค่าเฉลี่ยของหลาย ๆ ตัวอย่าง สุ่มตามลำดับ

ตัวอย่าง สุ่มตัวอย่างรถบรรทุก 100 คัน ที่วิ่งระหว่างกรุงเทพฯ – นครปฐม มีน้ำหนักเฉลี่ย 45,000 ปอนด์ ส่วนเบี่ยงเบนมาตรฐาน 3,500 ปอนด์

(ก) จงหา 95% ของช่วงความเชื่อมั่นสำหรับน้ำหนักเฉลี่ยของรถบรรทุกทั้งหมดที่วิ่งเส้นทางนี้

(ข) ความคลาดเคลื่อนจะเป็นเท่าใด โดยใช้สัมประสิทธิ์ความเชื่อมั่น 0.95 ที่จะประมาณค่าน้ำหนักเฉลี่ยของรถบรรทุกทั้งหมดเป็น 45,000 ปอนด์

วิธีทำ จากโจทย์เราได้ $\bar{x}$ = 45,000, s = 3,500, n = 100 และ Z = 1.96 เราต้องการประมาณค่า μ แบบช่วง

(ก) ช่วงความเชื่อมั่นสำหรับ μ

$$\bar{x} - Z \frac{s}{\sqrt{n}} < \mu < \bar{x} + Z \frac{s}{\sqrt{n}}$$

แทนค่า

$$45000 - 1.96 \frac{3500}{\sqrt{100}} < \mu < 4500 + 1.96 \frac{3500}{\sqrt{100}}$$

$$45000 - 686 < \mu < 45000 + 686$$

$$44314 < \mu < 45686$$

$\therefore$ น้ำหนักเฉลี่ยของรถบรรทุกทั้งหมดจะอยู่ระหว่าง 44,314 กับ 45,686 ปอนด์ ที่เราเชื่อได้ 95%

(ข) หากเราใช้ $\bar{x} = 45{,}000$ ไปประมาณค่า μ จะทำให้ค่าคลาดเคลื่อนไป $Z \frac{s}{\sqrt{n}} = 1.96 \frac{3500}{\sqrt{100}} = 686$ ด้วยความน่าจะเป็น 0.95

การประมาณค่าแบบช่วงสำหรับค่าเฉลี่ย (ตัวอย่างขนาดเล็ก)

ในกรณีที่เรามีตัวอย่างขนาดเล็ก คือ $n \leq 30$ การแจกแจงของค่าเฉลี่ยจะเป็นการแจกแจงแบบ t ค่าสถิติที่มีการแจกแจงแบบ t คือ ค่าสถิติ

$t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$ คล้ายกับค่าสถิติ Z ต่างกันเพียงขนาดตัวอย่างเท่านั้น ในกรณีที่เราไม่ทราบ σ $(s_{\bar{x}} = \frac{s}{\sqrt{n}})$

พื้นที่ใต้โค้งแจกแจงแบบ t แตกต่างกันไปตามขนาดตัวอย่าง หรือองศาแห่งความเป็นอิสระมีค่าเท่ากับ $n - 1$

องศาแห่งความอิสระ หมายถึงทางเลือกค่าที่เป็นไปได้ เมื่อกำหนดผลบวกให้ อย่างเช่นเมื่อเรากำหนดว่า ผลบวกของเลข 4 ตัวรวมกันได้ 10 จะเห็นได้ว่าเราอาจเลือกเลข 3 ตัวแรกเป็นค่าใด ๆ ก็ได้แต่ตัวที่ 4 จะถูกจำกัดค่าคงที่ไว้ เมื่อได้ 3 ค่าแรก เพราะผลบวกทั้ง 4 ตัวต้องเป็น 10 ในกรณีนี้เรากล่าวว่า องศาแห่งความอิสระเท่ากับ 3

ด้วยเหตุนี้ค่าสถิติ $t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$ นี้จึงเรียกว่ามีการแจกแจงแบบ t โดยมีองศาแห่งความอิสระ $n - 1$ ค่าพื้นที่ใต้โค้งแจกแจง t จะหาได้จากตารางค่า t ตั้งแต่องศาแห่งความอิสระ = 1 ไปถึง 30 แยกตามพื้นที่ใต้โค้งต่าง ๆ กัน

กระบวนการก็เหมือนกันกับกรณีตัวอย่างขนาดใหญ่ เราจะทราบได้จากตารางค่า t ว่าพื้นที่ใต้โค้งระหว่างค่า $-t_{.025}$ ถึง $+t_{.025}$ ที่องศาแห่งความอิสระ $n - 1$ เท่ากับ 0.95 นั่นคือ

$$P(-t_{.025} < t < +t_{.025}) = 0.95$$

$$P(-t_{.025} < \frac{\bar{X} - \mu}{s / \sqrt{n}} < + t_{.025}) = 0.95$$

$$P(\bar{X} - t_{.025} \frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{.025} \frac{s}{\sqrt{n}}) = 0.95$$

$\bar{X} - t_{.025} \frac{s}{\sqrt{n}}$ เรียกขีดจำกัดล่าง $\bar{X} + t_{.025} \frac{s}{\sqrt{n}}$ เรียกขีดจำกัดบน

ความหมายของช่วงความเชื่อมั่นในกรณีตัวอย่างขนาดเล็กก็เหมือนกับในกรณีตัวอย่างขนาดใหญ่

ตัวอย่าง สำนักงานธุรกิจแห่งหนึ่งได้ศึกษาเกี่ยวกับค่าเช่าบ้านที่มี 2 ห้องนอน ในเมืองหนึ่ง ได้สุ่มตัวอย่างขนาด 16 หลังคาเรือน ค่าเช่าเฉลี่ยต่อหลัง 800 บาท ส่วนเบี่ยงเบนมาตรฐาน 120 บาท จงหา 95% ช่วงความเชื่อมั่นสำหรับค่าเช่าเฉลี่ยจริงของบ้าน 2 ห้องนอน

วิธีทำ จากโจทย์เราได้ $\bar{X} = 800, s = 120, n = 16$ และ $t_{.025 : 15} = 2.145$ เราต้องการประมาณค่า แบบช่วงความเชื่อมั่นสำหรับ

$$\bar{X} - t_{.025 ; 15} \frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{.025 : 15} \frac{s}{\sqrt{n}}$$

$$800 - 2.1315 \frac{120}{\sqrt{16}} < \mu < 800 + 2.1315 \frac{120}{\sqrt{16}}$$

$$800 - 63.95 < \mu < 800 + 63.95$$

$$735.65 < \mu < 863.95$$

$\therefore$ ค่าเช่าเฉลี่ยจริงของบ้าน 2 ห้องนอนอยู่ระหว่าง 735.65 กับ 863.95 ที่เราจะเชื่อได้ 95%

การประมาณค่าแบบช่วงสำหรับสัดส่วน

เราได้เห็นในเรื่องการแจกแจงตัวอย่างของสัดส่วน เมื่อ $n > 30$ ว่าการแจกแจงตัวอย่างของสัดส่วน $\frac{x}{n}$ เป็นแบบปกติมีค่าเฉลี่ย p และความแปรปรวน $\frac{p \times q}{n}$ เราอาจหาช่วงความเชื่อมั่นของ p ในประชากรได้ในลักษณะเดียวกับที่ทำการหาช่วงความเชื่อมั่นของค่าเฉลี่ย (ตัวอย่างขนาดใหญ่) อาศัยลักษณะการแจกแจงแบบปกติของค่าสถิติ

$$Z = \frac{\frac{x}{n} - p}{\sqrt{pq / n}}$$

ค่าที่ระดับความเชื่อมั่น 0.95

$$P(-1.96 < Z < +1.96) = 0.95$$

$$P\left(-1.96 < \frac{\frac{x}{n} - p}{\sqrt{\frac{pq}{n}}} < +1.96\right) = 0.95$$

ในทางปฏิบัติเราหาค่า ส่วนเบี่ยงเบนมาตรฐานของประชากร $\sqrt{\frac{pq}{n}}$ ไม่ได้ จึงใช้ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง $\sqrt{\frac{x}{n}(1-\frac{x}{n})/n}$ แทน ดังนั้นค่าที่ระดับความเชื่อมั่น 0.95 สำหรับค่าสัดส่วนในประชากร p ก็คือ

$$p\left(\frac{x}{n} - 1.96\sqrt{\frac{\frac{x}{n}(1-\frac{x}{n})}{n}} < p < \frac{x}{n} + 1.96\sqrt{\frac{\frac{x}{n}(1-\frac{x}{n})}{n}}\right) = 0.95$$

การกำหนดขนาดตัวอย่าง

ปัญหาที่เราจะต้องพบอยู่เสมอในการประมาณค่าก็คือ เราจะต้องใช้ตัวอย่างขนาดเท่าไร เพื่อให้ค่าประมาณมีความแม่นยำใกล้เคียงกับค่าพารามิเตอร์ที่เราต้องการประมาณ เราได้เห็นแล้วว่าขนาดตัวอย่างมีผลต่อความแปรปรวนของการแจกแจงของ $\bar{x}$ หากตัวอย่างยิ่งมีขนาดใหญ่ความแปรปรวนก็ยิ่งน้อย แต่เราต้องคำนึงถึงเวลาและค่าใช้จ่ายในการหาตัวอย่างไว้ด้วย ดังนั้นเราจึงไม่ต้องการที่จะใช้ตัวอย่างขนาดใหญ่เกินความจำเป็น

การกำหนดขนาดตัวอย่างสำหรับประมาณค่าเฉลี่ย

เราทราบแล้วว่า ในการประมาณแบบช่วง $1.96 \frac{\sigma}{\sqrt{n}}$ เป็นค่าที่บอกให้เรารู้ว่าค่าประมาณ $\bar{X}$ คลาดเคลื่อนไปหรือแตกต่างไปจากค่าพารามิเตอร์ μ ได้มากน้อยแค่ไหนด้วยความเชื่อมั่น 95% เพราะฉะนั้นเราจึงสามารถกำหนดขนาดความแม่นยำของค่าประมาณว่าจะให้แตกต่างไปจาก μ ได้ไม่เกินกว่า E ที่เราต้องการได้ด้วยระดับความเชื่อมั่น 0.95 เราจะได้ว่า

$E = 1.96 \frac{\sigma}{\sqrt{n}}$ ซึ่งเราอาจแก้สมการนี้เมื่อเราทราบค่าของ E และ σ เพื่อหาค่าของ n ได้เป็นสูตรดังนี้

$$n = \left(\frac{1.96\,\sigma}{E}\right)^2$$

ค่า n ที่ได้นี้จะเป็นขนาดตัวอย่างที่จะให้ความแม่นยำในระดับที่เราต้องการได้ สูตรทั่ว ๆ ไป $n = \left(\frac{Z\,\sigma}{E}\right)^2$ ขึ้นอยู่กับระดับความเชื่อมั่นที่เราจะกำหนดไว้ก่อน อาจเป็น 0.90, 0.95, 0.98 และ 0.99 ก็ได้ ค่า n ที่ใหญ่กว่านี้ก็จะให้ความแม่นยำดีขึ้นไปอีก แต่อาจจะเกินความจำเป็น ส่วนค่า n ที่เล็กกว่านี้จะไม่ให้ค่าประมาณที่แม่นยำพอ

คำและประโยคที่ควรจำ

การประมาณค่า
ตัวค่าประมาณเชิงจุด
การประมาณค่าเชิงช่วง
ช่วงความเชื่อมั่น
ขอบเขตของความคลาดเคลื่อน
ค่าคลาดเคลื่อนมาตรฐานของ $\bar{X}$
องศาแห่งความอิสระ
ตัวอย่างขนาดเล็ก
การแจกแจงแบบ t

เติมคำลงในช่องว่าง บทที่ 8

จงเติมลงในช่องว่างให้สมบูรณ์ โดยคำตอบที่ถูกต้องอยู่ด้านขวามือ ใช้ไม้บรรทัดปิดคำตอบสำหรับคำถามซึ่งท่านยังไม่ได้ตอบ

ต่อไปนี้เป็นการแสดงปัญหาในทางปฏิบัติเกี่ยวกับการประมาณค่าและขนาดตัวอย่างซึ่งจำเป็นต้องเข้าใจสถิติบ้าง ในสภาวะของการผลิตนี้การค้นหาค่าประมาณสำหรับอายุเฉลี่ยของแบตเตอรี่ที่สร้างขึ้นสำหรับเครื่องคำนวณอิเลคโทรนิคเล็ก ๆ ตัวพารามิเตอร์ที่จะประมาณมีสัญลักษณ์เป็น (1)..........และตัวประมาณเชิงจุดที่ได้มาจากการสุ่มตัวอย่างเป็น (2).........ทฤษฎี (3).........ให้หลักเกณฑ์สำหรับจัดช่วงความเชื่อมั่น สมมติว่า ตัวอย่าง (4)..........ขนาด 100 ของแบตเตอรี่ ให้อายุของการใช้ $\bar{X}$ = 1,232 ชั่วโมง ส่วนเบี่ยงเบนมาตรฐาน 17.5 ชั่วโมง แล้ว ค่าประมาณเชิงจุดสำหรับมัชฌิมจริง ๆ (5)..........เป็น (6)..........ชั่วโมง สำหรับค่าประมาณเชิงช่วงความคลาดเคลื่อนมาตรฐาน $\sigma_{\bar{x}}$ เป็น (7)..........และเลือกระดับความน่าจะเป็น 95.44 เปอร์เซ็นต์ นี้ต้องการ Z = (8)..........แต่ผู้จัดการฝ่ายขายพูดว่าขอบเขตของความคลาดเคลื่อน E มีค่า (9) ....ชั่วโมงใหญ่เกินไปจะต้องกระทำบางสิ่งบางอย่างเพื่อลดความคลาดเคลื่อน มิฉะนั้นผู้ซื้ออาจเริ่มปฏิเสธแบตเตอรี่ หัวหน้าการผลิตพิจารณาความเชื่อมั่นต่ำลงเป็น 90 เปอร์เซ็นต์ ดังนั้น Z = (10)..........และ E = (11)..........ชั่วโมง สำหรับ n เหมือนกัน แต่กลุ่มนักสถิติกล่าวว่า นี้เป็นความคิดที่ผิด "แน่นอนเรามีความคลาดเคลื่อนน้อยลง แต่เราก็มีโอกาส (12)................ (น้อยลง, เหมือนเดิม, มากกว่า) เพื่อที่จะครอบคลุมส่วนเฉลี่ยจริงในช่วงของเรา นี้ไม่ได้เป็นการปรับปรุงที่แท้จริง" ภายหลังจากเขาทั้งหลายได้ตัดสินใจใช้กระบวนการออมชอม โดยการเพิ่มตัวอย่างของเขาเป็น 225 และใช้ 92 เปอร์เซ็นต์ (ลดความเชื่อมั่น) ผลลัพธ์นี้ใน Z = (13)......................และ E = (14)..................... เขายอมรับว่า ความคลาดเคลื่อนและระดับความน่าจะเป็นมีเหตุมีผล เนื่องจากว่าต้นทุนสำหรับการสุ่มตัวอย่างอยู่ภายในงบประมาณของเขา เขาก็ดำเนินตามแผนนี้

(1) μ
(2) $\bar{X}$
(3) ขีดจำกัดส่วนกลาง
(4) สุ่ม
(5) μ
(6) 1,232
(7) 1.75 ชม.
(8) 2.0
(9) 3.5
(10) 1.64
(11) 2.87
(12) น้อยลง
(13) 1.75
(14) 2.04 ชม.

เรามาสรุปจุดบางจุดของตัวอย่างนี้ ขั้นแรก ขอบเขตของความคลาดเคลื่อนหมายความว่า (15) (เพิ่มขึ้น, ใหญ่กว่า, ลดลง) (15) เพิ่มขึ้น

(16)..........ช่วงความเชื่อมั่นสำหรับขนาดตัวอย่างคงที่เพื่อ (16) ความกว้าง หาช่วงที่แคบกว่าสำหรับขนาดตัวอย่างคงที่ หมายความว่าค่าของ E ลดลง ต้องการค่าความน่าจะเป็น (17).........การออมชอมในที่นี้ต้อง (17) ลดลง เอาตัวอย่างสุ่ม (18)..........และในที่นี้ต้องได้ค่าขนาดระหว่างความน่า (18) ใหญ่ขึ้น จะเป็นกับความคลาดเคลื่อน แต่ตัวอย่างใหญ่กว่า ยอมใช้ (19).......... มากกว่า ถ้าหากว่ากำหนดทุนเสีย ทางเลือกอีกอย่างหนึ่งคือต้องใช้สถิติ (19) ทุน มากขึ้น และกระบวนการ (20)..........ยุ่งยากขึ้น ไม่ว่าในกรณีใด ๆ (20) การเลือกตัวอย่าง กระบวนการที่ใช้คะแนน Z ต้องการ $n > (21)$..........

ถึงแม้ว่าการพิจารณานี้ได้เกี่ยวกับประมาณค่าของ μ หรือ (21) 30 p สำหรับในกรณีอื่น ๆ จะพิจารณาในบทต่อไป

ปัญหา ถูกหรือผิด บทที่ 8

ข้อความต่อไปนี้ถูกหรือผิด ถ้าถูกให้ขีดเส้นใต้ที่ตัวอักษร T หน้าข้อ ถ้าผิดให้ขีดเส้นใต้ที่ตัวอักษร F หน้าข้อนั้น ๆ

T F 1. ทฤษฎีขีดจำกัดส่วนกลางเป็นเกณฑ์สำหรับการประมาณค่ามัชฌิมเลขคณิตจากตัวอย่างขนาดใหญ่

T F 2. ระดับความเชื่อมั่นหมายถึงระดับความน่าจะเป็นสำหรับช่วงกลางรอบมัชฌิมเลขคณิต (หรือ p)

T F 3. ในการประมาณค่าทางสถิติ ตัวค่าประมาณที่มีความคลาดเคลื่อนน้อย คำนวณหาได้โดยการสุ่มตัวอย่างที่มีขนาดใหญ่กว่า

T F 4. ในการประมาณค่าของ μ ถ้าหากว่าเพิ่มขนาดตัวอย่างแต่ยึดความเชื่อมั่นคงที่ผลลัพธ์คือช่วงกว้างขึ้น

T F 5. สำหรับการประมาณค่าของ p สัดส่วนจริงของความสำเร็จ ตัวประมาณเชิงจุดร่วม คือ $\hat{p} = X/n$

T F 6. ขอบเขตของความคลาดเคลื่อน E คือความกว้างของช่วงตัวอย่างเช่น $.20 < p < .26$ มี $E = .26 - .20 = .06$

T F 7. การเปลี่ยนช่วงความเชื่อมั่น 95 เปอร์เซ็นต์ไปเป็นระดับความน่าจะเป็น 99 เปอร์เซ็นต์ (สมมติว่าขนาดตัวอย่างไม่มีการเปลี่ยนแปลงภายใต้การสุ่มตัวอย่าง) สิ่งหนึ่งต้องแทน $Z = 1.96$ ด้วย $Z = 2.33$

T F 8. สำหรับค่าประมาณด้วยความน่าจะเป็น 95 เปอร์เซ็นต์ พร้อมด้วย $n = 100$ ให้ $.55 < p < .76$ แล้วค่าของ p คือ .67

T F 9. ในการประมาณค่ามัชฌิมเลขคณิต ตัวอย่างที่ต้องการจากการสุ่มตัวอย่างเป็น $n = 41.8$ ควรจะเป็นตัวอย่างขนาด 42 มากกว่า 41

T F 10. โดยทั่ว ๆ ไปสำหรับการประมาณค่าเชิงช่วงใช้แบบของ Z คือ ค่าประมาณ $\pm$ ($Z \times$ ส่วนเบี่ยงเบนมาตรฐานสำหรับการแจกแจงตัวอย่างของตัวประมาณ)

T F 11. ประชากรของผลลัพธ์ที่อาจจะเป็นไปได้ทั้งหมดเป็นผลมาจากการโยนเหรียญหนึ่งอันสองครั้ง เป็นประชากรจำนวนจำกัด

T F 12. ในการโยนเหรียญหนึ่งอันสองครั้ง จำนวนของหนทางที่อาจจะเป็นไปได้คือสอง

T F 13. เราคาดหวังความผันแปรในตัวสถิติ ถึงแม้ว่าเลือกตัวอย่างเหล่านี้มาจากประชากรเดียวกัน

T F 14. สัดส่วนจริงของหัวและก้อย แสดงคุณลักษณะของเหรียญไม่สามารถทราบได้

T F 15. ถ้าหากว่าเราโยนเหรียญหนึ่งอันหกครั้ง และได้หัวหนึ่งครั้ง และก้อยห้าครั้ง เราอาจสรุปว่าเหรียญเอนเอียง

T F 16. สัมประสิทธิ์ของการกระจายทวินามสมนัยกับความถี่ที่คาดหวังของการเกิดขึ้นของแต่ละผลลัพธ์

T F 17. "องศาแห่งความอิสระ" หมายถึงจำนวนสิ่งของ

T F 18. การแจกแจง t แปรตามฟังก์ชันขององศาแห่งความอิสระ

T F 19. ถ้าหากว่าการแจกแจงหนึ่งที่กำหนดขึ้นเบ้ การแจกแจงของมัชฌิมเลขคณิตของตัวอย่างจะเบ้ด้วย

T F 20. การแจกแจงแบบ t สมมาตรรอบมัชฌิมเลขคณิตที่มีค่าเท่ากับหนึ่ง

T F 21. มัชฌิมเลขคณิตของตัวอย่างที่ประกอบหนึ่งข้อมูลมีค่าใกล้เคียงกับมัชฌิมเลขคณิตของประชากร ถ้าหากว่าการแจกแจงที่กำหนดขึ้นปกติ

T F 22. ค่าประมาณเชิงจุดขึ้นอยู่กับตัวสถิติที่คำนวณได้จากตัวอย่างที่ประกอบด้วยข้อมูล

T F 23. $s^2 = \frac{\Sigma(X - \bar{X})^2}{N}$ มีแนวโน้มเข้าสู่การประมาณความแปรปรวนของประชากรเกินไป

T F 24. ค่าประมาณที่ไม่เอนเอียงหนึ่งเท่ากับ (ค่าเฉลี่ย) ค่าพารามิเตอร์

T F 25. การประมาณที่เลวของ σ ด้วย s โดยเฉพาะตัวอย่างขนาดเล็ก

T F 26. การแจกแจงแบบ t รัดกุมกว่าการแจกแจงปกติ

T F 27. สัดส่วนของพื้นที่ของค่า t บวก เท่ากับสัดส่วนของพื้นที่ของค่า t ลบ ที่สมนัยกัน

T F 28. ขณะที่ d.f เพิ่มขึ้น การแจกแจงแบบ t เข้าใกล้การใกล้การแจกแจงปกติ

T F 29. ในการประมาณค่าเชิงช่วง เรากำหนดช่วงของค่าภายใน ซึ่งเราเชื่อว่าตัวพารามิเตอร์ตกอยู่

T F 30. กำหนดให้ ขีดจำกัดความเชื่อมั่น 95% สำหรับ μ เป็น 101 และ 107 ถ้าหากว่าเราใช้ $\alpha = 0.01$ มีทางที่พอจะเป็นไปได้ว่า เราอาจจะยอมรับ $Ho : \mu = 108$

T F 31. ช่วงภายในซึ่งมัชฌิมเลขคณิตของตัวอย่างวางอยู่ เรียกว่า ช่วงความเชื่อมั่น

เฉลยปัญหาถูกหรือผิด บทที่ 8

1. T	2. T	3. T	4. F	5. T
6. F	7. F	8. F	9. T	10. T
11. F	12. F	13. T	14. T	15. F
16. T	17. F	18. T	19. F	20. F
21. F	22. T	23. F	24. T	25. T
26. F	27. T	28. T	29. T	30. T
31. F