8. การประมาณค่า

1. การประมาณค่าแบบจุด
2. การประมาณค่าแบบช่วง และการสร้างช่วงเชื่อมั่น
3. การสร้างช่วงเชื่อมั่นของค่าเฉลี่ย
4. การสร้างช่วงเชื่อมั่นของสัดส่วนจากตัวอย่างกลุ่มโต
5. การกำหนดขนาดตัวอย่างเพื่อใช้ในการประมาณค่า
6. การประมาณค่าผลต่างของค่าเฉลี่ย 2 กลุ่ม
7. การประมาณค่าผลต่างของสัดส่วน 2 กลุ่ม
8. แบบฝึกหัด

ในบทต้น ๆ โดยเฉพาะบทที่ 7 เป็นความรู้สถิติเบื้องต้น เพื่อนำมาใช้ศึกษาสถิติภาคอนุมาน (statistical inference) ในบทที่ 8-10 สถิติอนุมานเป็นการศึกษาค่าพารามิเตอร์ซึ่งเป็นค่าของประชากร ได้แก่ค่าเฉลี่ยของประชากร ความแตกต่างของค่าเฉลี่ยของประชากร 2 กลุ่ม ค่าสัดส่วนของประชากร และความแตกต่างของสัดส่วนระหว่าง 2 ประชากร และความแปรปรวนของประชากร

การศึกษาสถิติภาคอนุมานจำแนกได้ 2 หัวข้อ คือ

1. การประมาณค่า (estimation, บทที่ 8)
2. การทดสอบสมมติฐาน (hypothesis testing, บทที่ 9, 10)

การประมาณค่าพารามิเตอร์ มี 2 วิธี คือ

1. การประมาณค่าแบบจุด (point estimation)
2. การประมาณค่าแบบช่วง (interval estimation)

---

1. การประมาณค่าแบบจุด

(Point Estimation)

ถ้าใช้ค่าเดี่ยวจากตัวอย่างเพียงค่าเดียวประมาณค่าพารามิเตอร์ จะเรียกวิธีการประมาณนั้นว่า เป็นวิธีประมาณค่าแบบจุด หรือแบบค่าเดี่ยว เช่น การใช้ $\bar{x}$ ประมาณค่า μ เป็นวิธีประมาณค่าแบบจุดเพราะใช้ค่าประมาณเพียง 1 ค่า และเรียก $\bar{x}$ ว่า ตัวประมาณค่าแบบจุด (point estimator) ส่วนค่าที่คำนวณได้ของ $\bar{x}$ เช่น คำนวณอายุเฉลี่ยของนักศึกษารามคำแหงได้ 20.5 ค่า 20.5 เรียกว่า ค่าประมาณแบบจุด (point estimate)

ลักษณะของตัวประมาณค่าที่ดี

(Good Estimators)

ตัวประมาณค่าที่ดีจะต้องมีคุณสมบัติ 4 ประการคือ

1. Consistensy **หรือ ความคงเส้นคงวา**

เมื่อเพิ่มขนาดตัวอย่าง n ให้โตขึ้นเรื่อย ๆ และทำให้ ค่าสถิติจากตัวอย่าง มีค่าใกล้เคียงกับค่าประชากรที่ต้องการประมาณ จะเรียกตัวสถิตินั้นว่ามีคุณสมบัติ "คงเส้นคงวา" ทั้ง x และ p มีคุณสมบัตินี้ เพราะเมื่อเพิ่มขนาดตัวอย่าง $\bar{x}$ จะเข้าไปสู่ μ และ p จะมีค่าเข้าใกล้ π

2. Efficiency **หรือ ความมีประสิทธิภาพ**

ถ้าจะกล่าวว่าตัวสถิติใดมีประสิทธิภาพหรือไม่ จะต้องอยู่ในรูปการเปรียบเทียบ ตัวสถิติที่มาจากการแจกแจงตัวอย่างสุ่มเดียวกัน จะมีประสิทธิภาพถ้ามีค่าเบี่ยงเบนมาตรฐานต่ำกว่า

คุณสมบัติข้อนี้ จำเป็นเพราะตัวประมาณค่าใดถ้ามีค่าเบี่ยงเบนมาตรฐานต่ำ จะมีความน่าจะเป็นสูงที่จะให้ค่าประมาณใกล้เคียงกับค่าพารามิเตอร์ที่ต้องการประมาณ เช่น $\bar{x}$ มีค่าเบี่ยงเบนมาตรฐานต่ำกว่าค่ามัธยฐาน จึงมีประสิทธิภาพสูงกว่าค่ามัธยฐาน

3. Unbiasness **หรือ ความไม่เอียงเฉ**

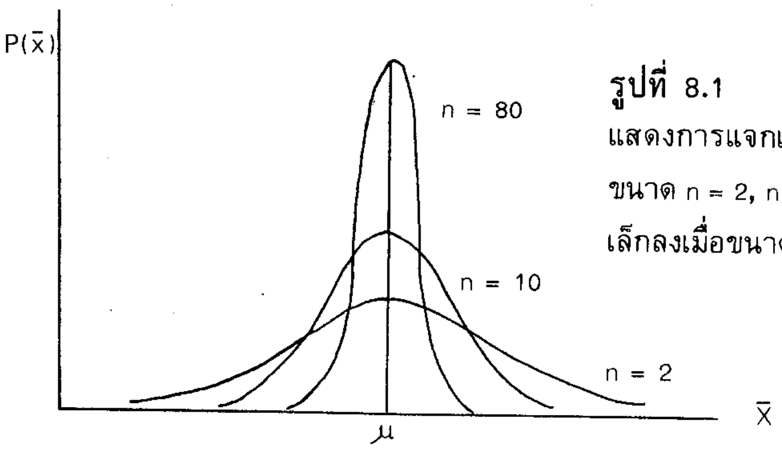
ตัวสถิติที่มีคุณสมบัติข้อนี้ จะมีค่าเฉลี่ยจากการแจกแจงของตัวอย่างสุ่มเป็นค่าเดียวกับค่าพารามิเตอร์ เช่น ในบทที่ 7 ได้แสดงการแจกแจงของ $\bar{x}$ และ $E(\bar{x}) = \mu$ = ค่าเฉลี่ยของประชากร $\bar{x}$ จึงเป็นตัวประมาณค่าที่ **ไม่เอียงเฉ** ในขณะเดียวกัน จากการแจกแจงของ p จะมีค่าเฉลี่ย คือ $E(p) = \pi$ = สัดส่วนของประชากร = พารามิเตอร์ ดังนั้น p จึงเป็นตัวประมาณค่าที่ไม่เอียงเฉของ π

4. Sufficiency **หรือ ความพอเพียง**

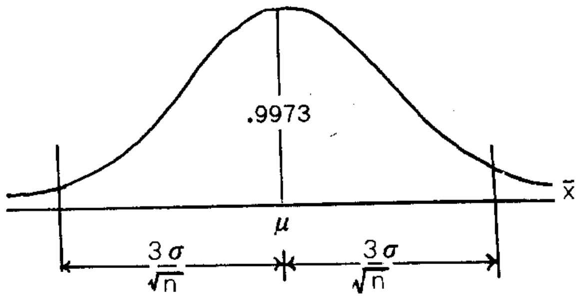
หมายถึงการที่ตัวประมาณค่านั้น ได้รวบรวมข่าวสารจากตัวอย่างไว้ได้มากที่สุด และไม่มีตัวประมาณค่าตัวอื่นใดที่จะนำข่าวสารจากตัวอย่าง ไปอนุมานประชากรได้มากกว่า

ความคลาดเคลื่อนของการประมาณค่า (Error of Estimation)

เราทราบว่าเมื่อใช้ค่าสถิติประมาณค่าพารามิเตอร์ เช่น ใช้ $\bar{x}$ ประมาณค่า μ ส่วนใหญ่ $\bar{x}$ จะไม่ใช่ค่าเดียวกับ μ ความแตกต่างนี้เรียกว่า **ความคลาดเคลื่อนจากตัวอย่างสุ่ม** หรือ sampling error และตามกฎขีดจำกัดเข้าสู่ส่วนกลาง เราทราบว่าเมื่อขนาดตัวอย่างโตพอ $\bar{x}$ จะมีการแจกแจงโดยประมาณแบบโค้งปกติ และมีค่าเฉลี่ย $E(\bar{x}) = \mu$ และความคลาดเคลื่อนมาตรฐาน $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ นั่นคือ $\sigma_{\bar{x}}$ จะมีค่าเล็กลงเมื่อเพิ่มขนาดตัวอย่างดังรูปที่ 8.1 ดังนั้น เมื่อขนาดตัวอย่างโตพอสมควร การแจกแจงของ $\bar{x}$ ส่วนใหญ่ (99.73%) จะอยู่ระหว่าง 3 หน่วย ของความคลาดเคลื่อนมาตรฐานจากแกนกลาง (ค่าเฉลี่ย) ดังรูปที่ 8.2



รูปที่ 8.1
แสดงการแจกแจงของ $\bar{x}$ จากตัวอย่างสุ่มขนาด n = 2, n = 10 และ n = 80 ค่า $\sigma_{\bar{x}}$ จะเล็กลงเมื่อขนาดตัวอย่างโตขึ้น



รูปที่ 8.2
แสดงการแจกแจงของ $\bar{x}$ จากตัวอย่างสุ่มขนาด n จะมีความน่าจะเป็น .9973 ที่ $\bar{x}$ จะอยู่ภายใน 3 หน่วยความคลาดเคลื่อนมาตรฐานจาก μ

ตัวอย่าง 1 น้ำหนักของไก่พันธุ์เนื้อจากฟาร์มแห่งหนึ่งมีน้ำหนักเฉลี่ย μ และค่าเบี่ยงเบนมาตรฐาน σ = 1.2 กิโลกรัม ถ้าสุ่มมา 36 ตัว ได้น้ำหนักเฉลี่ย 2 กิโลกรัม จงหาความคลาดเคลื่อนในการประมาณค่า μ ด้วย $\bar{x}$

$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 1.2/\sqrt{36} = 1.2/6 = 0.2$

ดังนั้น ความคลาดเคลื่อนสูงสุด = 3σ = 3(.0.2) = 0.6 กิโลกรัม

นั่นคือ น้ำหนักเฉลี่ยที่แท้จริง (μ) น่าจะมีค่าอยู่ระหว่าง 1.4 ถึง 2.6 กิโลกรัม

ตัวอย่าง 2 ถ้า 20% ของวัตถุดิบที่โรงงานซื้อมามีคุณภาพต่ำกว่ามาตรฐาน ถ้าในการตรวจรับสินค้า 100 ชิ้น พบสินค้าคุณภาพต่ำกว่ามาตรฐาน 15 ชิ้น จงประมาณสัดส่วนของสินค้าที่มีคุณภาพต่ำกว่ามาตรฐาน และหาความคลาดเคลื่อนของการประมาณค่า π ด้วย p

$$p = 15/100 = .15, \quad \sigma_p = \sqrt{\pi(1-\pi)/n} = \sqrt{(0.2)(0.8)/100} = 0.04$$

ค่าประมาณของ π (สัดส่วนที่แท้จริงของสินค้าคุณภาพต่ำ) คือ p = .15 ความคลาดเคลื่อนสูงสุดในการประมาณค่า π ด้วย p

$= 3\sigma_p = 3(.04) = .12$

นั่นคือ π จะมีค่าส่วนใหญ่อยู่ระหว่าง 3% และ 27%

สาเหตุที่ต้องทราบค่าความคลาดเคลื่อนของค่าประมาณ เพราะวิธีประมาณค่าแบบจุด จะใช้เพียงค่าเดียว ซึ่งเราทราบว่ามักไม่ใช่ค่าเดียวกับพารามิเตอร์ เช่น ให้ประมาณแขกที่จะมางานเลี้ยง เพื่อจะได้จัดเตรียมของได้ถูกต้อง ถ้าบอกว่าจากสถิติเดิมจะมีโดยเฉลี่ย 350 คน ก็ยังถือว่า ตัวสถิตินี้ “ไม่พอเพียง” เพราะไม่ทราบว่า จำนวนจริงจะผิดพลาดไปเท่าใด ถ้าเราทราบว่าค่าที่ได้จะมีความคลาดเคลื่อน 10 คน เราย่อมสบายใจและเต็มใจจัดเตรียมอาหารสำหรับ 350 คน แต่ถ้าค่าประมาณนั้นมีความคลาดเคลื่อน 90 คน เราจะไม่ค่อยยินดีกับค่าประมาณ 350 คนนั้น ดังนั้น ถ้าทราบความคลาดเคลื่อนของค่าประมาณ จะทำให้ค่าประมาณแบบจุดมีประโยชน์มากขึ้น

แบบฝึกหัด

8.1 จงบอกเครื่องมือสำคัญที่ใช้วิเคราะห์สถิติภาคอนุมาน

8.2 ค่าประมาณแบบจุดมีข้อเสียอย่างไร? มีวิธีแก้ไขอย่างไร?

8.3 จงอธิบายความแตกต่างระหว่าง "ตัวประมาณค่า" และ "ค่าประมาณ"

8.4 คุณสมบัติของตัวประมาณค่าที่ดีมีอะไรบ้าง?

8.5 ความคงเส้นคงวา (consistency) มีส่วนในการกำหนดขนาดตัวอย่างอย่างไร?

8.6 อัตราผลตอบแทนการลงทุน (คิดเป็นเปอร์เซ็นต์) ของโรงงานผลิตยาที่สุ่มมา 10 โรงงานมีดังนี้

17.0 25.0 13.0 8.5 27.5 20.0 18.5 17.0 16.0 12.0

จงหาค่าประมาณแบบจุดของ μ และ σ^2 (17.4, 34.1)

8.7 ในการสอบถามพนักงานที่สุ่มมา 500 ของโรงงานทอผ้าแห่งหนึ่ง พบว่า มีคนงาน 284 คน ไม่พอใจการปรับปรุงระบบการทำงาน ผู้บริการจึงอยากทราบสัดส่วนของคนงานทั้งหมดที่เห็นชอบการปรับปรุง จงหาค่าประมาณแบบจุดของสัดส่วนที่แท้จริงดังกล่าว และหาความคลาดเคลื่อนสูงสุด (.066)

8.8 โรงงานผลิตสินค้าต้องการประมาณผลผลิตถัวเฉลี่ยต่อชั่วโมง ของคนงานทั้งหมดที่ทำงานประเภทเดียวกัน เมื่อสุ่มคนงานมา 100 คน ได้ผลผลิตเฉลี่ย 90 หน่วย จงประมาณผลผลิตเฉลี่ยที่แท้จริง และถ้าทราบว่าความแปรปรวนของผลผลิตต่อชั่วโมงเป็น 625 จงหาความคลาดเคลื่อนที่อาจเป็นได้ของการประมาณค่า (7.5)

8.9 โรงงานเคมีภัณฑ์ต้องการประมาณผลผลิตเฉลี่ยต่อสัปดาห์ของเคมีภัณฑ์ชนิดหนึ่ง ซึ่งทราบว่ามีความแปรปรวนต่อสัปดาห์ = 400 จากสถิติการผลิต 64 สัปดาห์ ได้ผลผลิตเฉลี่ย 500 ตัน จงประมาณค่า μ และหาความคลาดเคลื่อนที่อาจเกิดขึ้นของการประมาณ (7.5)

8.10 ผู้จัดการฝ่ายบุคคลของห้างสรรพสินค้าแห่งหนึ่งต้องการประมาณจำนวนวันลาหยุดของพนักงานใน 1 ปี ถ้าจำนวนวันหยุดถัวเฉลี่ยต่อปีของพนักงานที่สุ่มมา 144 คน คือ 3.5 วัน และถ้าทราบว่าความแปรปรวนของวันหยุดคือ 36 จงหาความคลาดเคลื่อนของค่าประมาณ (1.5)

8.11 ผู้ผลิตยาชนิดหนึ่งอ้างว่ายาที่ผลิตใหม่จะมีผลในการรักษาโรคชนิดหนึ่งถึง 90% เมื่อสุ่มผู้ป่วยที่รักษาด้วยยาดังกล่าวมา 100 คน มี 80 คนหายจากโรค จงประมาณสัดส่วนที่แท้จริงของการรักษา และความคลาดเคลื่อนของค่าประมาณ (.80, .09)

8.12 ถ้ามหาวิทยาลัยรามคำแหง แจ้งว่ามีนักศึกษาสนใจการบรรยายทางโทรทัศน์อยู่ 20% ถ้าฝ่ายสถิติสุ่มนักศึกษามา 625 คน พบว่ามี 180 คน ที่รับชมการบรรยายทางโทรทัศน์ด้วย จงประมาณค่าสัดส่วนที่แท้จริงของผู้รับชมทางโทรทัศน์และความคลาดเคลื่อนจากการสุ่มตัวอย่าง (28.8, .048)

8.13 ถ้า X เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงแบบยูนิฟอร์ม (ทุกค่าของ x เกิดขึ้นด้วยโอกาสเท่ากัน = $\frac{1}{k}$) จากประชากร $\{2, 4, 6\}$ และถ้าสุ่มมา 2 ตัว แบบมีการแทนที่ เพื่อหา x จงแสดงว่า

ก. $E(\bar{x}) = \mu$

ข. $E(s^2) = \sigma^2$

8.14 ถ้าปกติเปอร์เซ็นต์สินค้าชำรุดของโรงงานทอผ้าขนหนูคือ 20% จะต้องสุ่มสินค้ามาตรวจสอบกี่หน่วย ถ้าต้องการให้มั่นใจว่า

ก) สัดส่วนชำรุดอยู่ระหว่าง 0.08 ถึง 0.32? (100)

ข) สัดส่วนชำรุดอยู่ระหว่าง 0.14 ถึง 0.26? (400)

ค) สัดส่วนชำรุดอยู่ระหว่าง 0.17 ถึง 0.23? (1,600)

---

2. การประมาณค่าแบบช่วง

ค่าประมาณแบบช่วง คือ พิสัยของค่าพารามิเตอร์ ในบทนี้จะกล่าวถึงพารามิเตอร์ 2 ตัวก่อน คือ μ และ π ในการประมาณค่า μ ต้องระลึกถึงทฤษฎีขีดจำกัดเข้าสู่ส่วนกลาง ซึ่งสรุปการแจกแจงของ $\bar{x}$ จากตัวอย่างสุ่มขนาด n ดังนี้

1. ถ้า x มาจากตัวอย่างซึ่งสุ่มจากประชากรแบบปกติ $\bar{x}$ จะต้องมีการแจกแจงแบบปกติด้วย ไม่ว่าจะมีขนาดตัวอย่างเล็กหรือโต
2. ถ้า x มาจากประชากรที่ไม่ได้มีการแจกแจงแบบปกติ และอาจไม่ทราบการแจกแจงที่แท้จริง แต่ถ้าขนาดตัวอย่างโตพอควร $\bar{x}$ จะมีการแจกแจงโดยประมาณแบบปกติ

นอกจากนั้น นักศึกษาควรจำคุณสมบัติที่สำคัญของโค้งปกติ คือ

1. พื้นที่ภายใต้โค้ง ประมาณ 68.3% อยู่ระหว่าง $\mu \pm \sigma$
2. พื้นที่ภายใต้โค้ง ประมาณ 95.5% อยู่ระหว่าง $\mu \pm 2\sigma$
3. พื้นที่ภายใต้โค้ง ประมาณ 99.7% อยู่ระหว่าง $\mu \pm 3\sigma$

ดังนั้น ถ้าทราบว่าอายุเฉลี่ยการใช้งานของแบตเตอรี่ที่สุ่มมา 200 ลูก คือ 36 เดือน และทราบว่าอายุการใช้งานมีค่าเบี่ยงเบนมาตรฐาน 10 เดือน นั่นคือ $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 10/\sqrt{200} = .71$ เดือน

1. $\bar{X} \pm \sigma_{\bar{X}} = 36 \pm .71 = 35.3, 36.7$ เดือน
2. $\bar{X} \pm 2\sigma_{\bar{X}} = 36 \pm 2(.71) = 34.6, 37.4$ เดือน
3. $\bar{X} \pm 3\sigma_{\bar{X}} = 36 \pm 3(.71) = 33.9, 38.1$ เดือน

ทั้ง 3 ข้อนี้เป็นค่าประมาณแบบช่วงของ μ ซึ่งจะอธิบายได้ดังนี้

ถ้าสุ่มตัวอย่างขนาด n จากประชากรเดียวกันมา 1,000 ชุด ทุก ๆ ชุดคำนวณค่า $\bar{x}$

1. ถ้าหา $\bar{x} \pm \sigma$ ทั้ง 1,000 ชุด จะมีอยู่ 68.3% ของช่วงเหล่านั้น คือ 683 อันที่ครอบคลุมค่า μ
2. ถ้าหา $\bar{x} \pm 2\sigma$ ทั้ง 1,000 ชุด จะมี 95.5% ของช่วงเหล่านั้น คือ 955 อันที่ครอบคลุมค่า μ และ
3. ถ้าหา $\bar{x} \pm 3\sigma$ ทั้ง 1,000 ชุด จะมีอยู่ 99.7% ของช่วงเหล่านั้นคือ 997 อันที่ครอบคลุมค่า μ

ดังนั้น จึงสรุปผลสำหรับเรื่องแบตเตอรี่ได้ดังนี้

1. เรามีความมั่นใจ 68.3% ว่าอายุเฉลี่ยการใช้งานที่แท้จริงของแบตเตอรี่อยู่ระหว่าง 35.3 ถึง 36.7 เดือน
2. เรามีความมั่นใจ 95.5% ว่าอายุเฉลี่ยการใช้งานที่แท้จริงของแบตเตอรี่อยู่ระหว่าง 34.6 ถึง 37.4 เดือน
3. เรามีความมั่นใจ 99.7% ว่า อายุเฉลี่ยการใช้งานที่แท้จริงของแบตเตอรี่อยู่ระหว่าง 33.9 ถึง 38.1 เดือน

แบบฝึกหัด

8.15 กำหนดให้ $\sigma = 0.9$, $n = 36$, $\bar{x} = 9.6$

ก) จงหาความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย (0.15)

ข) จงสร้างค่าประมาณแบบช่วง 1 หน่วยของค่าเบี่ยงเบนมาตรฐานโดยรอบค่าเฉลี่ย

$(9.45 < \mu < 9.75)$

8.16 สุ่มคนงานมา 81 คน มีอายุเฉลี่ย 24.5 ปี และทราบว่า $\sigma = 3.6$

ก) จงหาความคลาดเคลื่อนมาตรฐานของอายุเฉลี่ย (0.4)

ข) จงหาช่วงโดยรอบค่าเฉลี่ยตัวอย่างที่จะครอบคลุมค่าเฉลี่ยของประชากร 95.5% ของจำนวนครั้งทั้งหมด ($23.7 < \mu < 25.3$)

8.17 ถ้า $\sigma^2 = 196$, $n = 49$, $\bar{x} = 210$ จงสร้างค่าประมาณแบบช่วงที่จะครอบคลุมค่าเฉลี่ยที่แท้จริง 68.3% ของจำนวนครั้งทั้งหมด ($208 < \mu < 212$)

8.18 เจ้าของภัตตาคารแห่งหนึ่งได้เก็บสถิติจำนวนลูกค้าแต่ละคืนรวม 25 คืน พบว่ามีลูกค้าโดยเฉลี่ยคืนละ 68 คน และทราบว่า $\sigma = 4.25$

(ก) จงหาค่าประมาณแบบช่วงที่มีความน่าจะเป็น 68.3% ที่จะครอบคลุมค่าเฉลี่ยที่แท้จริง ($67.15 < \mu < 68.85$)

(ข) จงหาค่าประมาณแบบช่วงที่มีความน่าจะเป็น 95,5% ที่จะครอบคลุมค่าเฉลียทีแท้จริง ($66.3 < \mu < 69.7$)

การสร้างช่วงเชื่อมั่น (Confidence Interval)

เราหาค่าประมาณแบบช่วงโดยใช้ค่า σ, 2σ และ 3σ ไปบวกเข้าและลบออก แต่ถ้าเราเปิดตาราง Z จะพบว่า เมื่อ $Z = 1.64$ จะทำให้เหลือพื้นที่ตรงกลางระหว่างจุด $\pm$ 1.64 = 90% พื้นที่ 95% อยู่ระหว่างจุด $\pm$ 1.96 และพื้นที่ 99% อยู่ระหว่างจุด $\pm$ 2.58 ดังนั้น ถ้าเรากำหนดค่าความน่าจะเป็นควบคู่กับค่าประมาณแบบช่วง เราจะเรียกว่าช่วงเชื่อมั่น ซึ่งปกติ ระดับความเชื่อมั่นที่นิยมใช้มีอยู่ 3 ค่า คือ 90%, 95% และ 99% ซึ่งจะมีระยะห่างจากค่าเฉลี่ยเป็น $1.64\ \sigma_{\bar{x}}$ $1.96\ \sigma_{\bar{x}}$ และ $2.58\ \sigma_{\bar{x}}$ ตามลำดับ เช่น $\bar{x} \pm 1.96\sigma_{\bar{x}}$ คือช่วงเชื่อมั่น 95%

$\bar{x} - 1.96\sigma_{\bar{x}}$ คือขีดจำกัดล่าง (lower limit)

$\bar{x} + 1.96\sigma_{\bar{x}}$ คือขีดจำกัดบน (upper limit)

95% คือ ระดับความเชื่อมั่น (Confidence level)

พึงสังเกตว่า ถ้าระดับความเชื่อมั่นต่ำ ช่วงเชื่อมั่นจะแคบ แต่ถ้าระดับความเชื่อมั่นสูงช่วงเชื่อมั่นจะกว้าง ดังแสดงความสัมพันธ์ในตาราง 8.1

ตารางที่ 8.1 แสดงความสัมพันธ์ระหว่างระดับความเชื่อมั่น และช่วงเชื่อมั่น

	ลูกค้าถาม	ผู้จัดการตอบ	หมายถึง ระดับความเชื่อมั่น	หมายถึง ช่วงเชื่อมั่น
1.	ฉันจะได้ตู้เย็นภายใน 1 ปีไหม?	ผมมีความมั่นใจว่าต้องได้แน่นอนครับ	สูงกว่า 99%	1 ปี
2.	ท่านจะจัดส่งตู้เย็นไปยังบ้านพักภายใน- 1 เดือนไหม?	ผมเชื่อแน่ว่าจะจัดส่งได้ภายใน 1 เดือน	อย่างน้อย 95%	1 เดือน
3.	ท่านจะจัดส่งตู้เย็นไปยังบ้านพักภายใน 1 สัปดาห์ไหม?	ผมคิดว่าจะจัดส่งได้ภายในสัปดาห์นี้	ประมาณ 80%	1 สัปดาห์
4.	ท่านจะจัดส่งตู้เย็นไปยังบ้านพักภายในวันพรุ่งนี้ไหม?	ผมไม่แน่ใจว่าพรุ่งนี้จะมีของหรือไม่	ประมาณ 40%	1 วัน
5.	จะมีตู้เย็นรออยู่ที่บ้านเย็นนี้ไหม?	โอกาสน้อยมากที่ตู้เย็นจะไปถึงก่อนท่านกลับบ้าน	1%	1 ชั่วโมง

ข้อควรระวัง ถ้าเรากล่าวว่า "เรามีความมั่นใจ 95% ว่าอายุเฉลี่ยการใช้งานที่แท้จริงของแบตเตอรี่อยู่ระหว่าง 30 ถึง 42 เดือน" ไม่ได้หมายความว่ามีโอกาส .95 ที่อายุการใช้งานเฉลี่ยอยู่ระหว่างช่วงดังกล่าว แต่มีความหมายว่า ถ้าเราใช้ขนาดตัวอย่างเท่าเดิมนี้ และสร้างช่วงเชือมัน 95% จากทุก ๆ ตัวอย่าง จะมีประมาณ 95% ของช่วงเหล่านั้นที่ครอบคลุมค่าเฉลี่ยของประชากร

แบบฝึกหัด

8.19 ถ้าท่านต้องการระดับความเชื่อมั่น 80% จงหาขีดจำกัดบน และขีดจำกัดล่างในรูปค่าเฉลี่ยและความคลาดเคลื่อนมาตรฐาน

8.20 เหตุใดค่าประมาณจะมีความหมายน้อยลง

ก) เมื่อมีระดับความเชื่อมั่นสูง ข) เมื่อมีช่วงเชื่อมั่นแคบ

8.21 จงสร้างช่วงเชื่อมั่นโดยมีระดับความเชื่อมั่นต่าง ๆ ดังนี้

ก) 50% ข) 75% ค) 85% ง) 98%

---

3. การสร้างช่วงเชื่อมั่นของค่าเฉลี่ยของประชากร

1. ถ้าทราบค่า σ

$(1 - \alpha)$ 100% ช่วงเชื่อมั่นของ μ คือ

$\bar{x} - Z_{\alpha/2}\sigma_{\bar{x}} < \mu < \bar{X} + Z_{\alpha/2}\sigma_{\bar{x}}$ โดยที่ $\sigma_{\bar{x}} = \sigma/\sqrt{n}$

นั่นคือ

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

2. ถ้าไม่ทราบค่า σ ต้องใช้ s เป็นค่าประมาณของ σ และต้องพิจารณาขนาดตัวอย่างด้วย ถ้า $n \geqslant 30$ จะประมาณได้ด้วยโค้งปกติ (Z)

2.1 ถ้าสุ่มจากประชากรแบบนับถ้วน คือทราบค่า N

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N - n}{N - 1}}$$

เมื่อไม่ทราบค่า σ ให้ $\hat{\sigma} = s$

$$\hat{\sigma}_{\bar{x}} = \frac{s}{\sqrt{n}}\sqrt{\frac{N - n}{N - 1}}$$

2.2 ถ้าสุ่มจากประชากรแบบนับไม่ถ้วน (infinite)

$$\hat{\sigma}_{\bar{x}} = s/\sqrt{n}$$

3. ถ้าไม่ทราบค่า σ และตัวอย่างขนาดเล็ก คือ $n < 30$

$\hat{\sigma}_{\bar{X}} = s/\sqrt{n}$ **แต่จะมีการแจกแจงแบบ** t **ที่มี** $df = n - 1$

ประวัติ การแจกแจงแบบ t เป็นผลงานของ W.S. Gossett ซึ่งเป็นพนักงานโรงงานกลั่นเบียร์ในเมืองดับบลิน ไอร์แลนด์ ในราวต้นคริสตศตวรรษ 1900 เขาได้ตีพิมพ์ผลงานนี้ภายใต้นามปากกา "Student" เพื่อหลีกเลี่ยงข้อห้ามของนายจ้างที่ห้ามพนักงานเขียนผลงานโดยใช้ชื่อจริง การแจกแจงที่เขาค้นพบ จึงมีชื่อว่า การแจกแจงแบบ t ของ Student หรือ Student's t distribution จะใช้การแจกแจงแบบ t **ในการประมาณค่าเมื่อขนาดตัวอย่างต่ำกว่า 30 และไม่ทราบค่าแท้จริงของ** σ รูปร่างของ t จะคล้ายกับ Z คือเป็นรูประฆังคว่ำ สมมาตรที่แกนกลางคือ $\mu = 0$ มีค่าอยู่ระหว่าง $\pm\infty$ **แต่โค้ง** t **จะมีรูปร่างเปลี่ยนแปลงตาม** degree of freedom (ν) เมื่อ n มีขนาดโตพอควรคือประมาณ 30 ขึ้นไป โค้ง t และ Z จะซ้อนกันพอดี ดังนั้น เราจึงหาค่าของ Z เมื่อมี α ต่าง ๆ จากการเปิดตาราง t ที่ $\nu = \infty$

การเปิดตาราง t

ตารางที่ 2 ท้ายเล่ม คือ ตาราง t จะต่างกับตาราง Z ดังนี้

1) จะกำหนดค่า t เพียงบางค่าที่ทำให้พื้นที่ปลายหางรวมกัน = α , α = 10, 5, 2 และ 1%

2) ตาราง t จะวัดโอกาสที่ค่าประมาณของประชากร จะไม่อยู่ในช่วงเชื่อมั่น เช่น ถ้าสร้างช่วงเชื่อมั่น 90% ต้องเปิดตาราง t ที่ α = 10%

3) การเปิดตาราง t ต้องดู df ด้วย สำหรับการประมาณค่า μ เพียง 1 ตัว ให้ใช้ df $(\nu) = n - 1$

ตัวอย่าง 1

ต้องการประมาณค่าอายุการใช้งานของแผ่นยางติดที่ปัดน้ำฝนบนกระจกรถยนต์ ซึ่งทราบว่า σ = 6 เดือน ด้วยความเชื่อมั่น 95% และกำหนดให้

$n = 100$, $\bar{x} = 21$ เดือน, $\sigma = 6$ เดือน

ดังนั้น $\sigma_{\bar{x}} = 6/\sqrt{100} = .6$ เดือน

$Z_{\alpha/2} = Z_{.025} = 1.96$

ช่วงเชื่อมั่น 95% ของ μ คือ

$= \bar{X} \pm Z_{\alpha/2}\sigma_{\bar{x}}$

$= 21 \pm 1.96\ (0.6)$

$= 21 \pm 1.18$

$= 19.82,\ 22.18$

นั่นคือ

เราประมาณอายุการใช้งานของแผ่นยางปัดน้ำฝนว่ามีอายุระหว่าง 19.82 ถึง 22.18 เดือน ด้วยความเชื่อมั่น 95%

ตัวอย่าง 2 ในอำเภอหนึ่งมี 700 ครอบครัว ต้องการประมาณรายได้ต่อปีต่อครอบครัวของอำเภอนี้ โดยใช้ความเชื่อมั่น 90% และมีข้อมูลดังนี้

$n = 50$, $\bar{X} = 4{,}800$ บาท, $S = 950$ บาท

ไม่ทราบค่า $\sigma_{\bar{x}}$ จึงหา $\hat{\sigma}_{\bar{x}} = s/\sqrt{n} = 950/\sqrt{50}$

แต่เนื่องจากทราบ $N = 700$ (finite population) จึงต้องปรับค่า $\hat{\sigma}_{\bar{x}}$ ด้วย $\sqrt{\frac{N-n}{N-1}}$

ดังนั้น $$\sigma_{\bar{x}} = \frac{950}{\sqrt{50}}\sqrt{\frac{700-50}{700-1}}$$

$$= \frac{950}{\sqrt{50}}\sqrt{\frac{650}{699}} = 134.34\ \sqrt{.9299}$$

$$= (134.37)(.9643) = 129.57 \text{ บาท}$$

ช่วงเชื่อมั่น 90% ของ μ คือ (μ = รายได้ถัวเฉลี่ยต่อครอบครัว)

$\bar{X} + Z_{.05}\ s_{\bar{x}}$

$= 4{,}800 \pm (1.64)(129.57)$

$= 4{,}800 \pm 212.50$

$= 4{,}587.50,\ 5{,}012.50$ บาท

นั่นคือ กล่าวด้วยความเชื่อมั่น 90% ว่า รายได้ถัวเฉลี่ยต่อปีต่อครัวเรือนของอำเภอดังกล่าวอยู่ระหว่าง 4,587.50 ถึง 5,012.50 บาท

ตัวอย่าง 3 ผู้จัดการโรงงานต้องการประมาณจำนวนวัตถุดิบต่อสัปดาห์ ด้วยความเชื่อมั่น 95% และมีข้อมูลดังนี้

$n = 10$ สัปดาห์, $\bar{x} = 11{,}400$ ตัน, $s = 700$ ตัน

กรณีนี้ ไม่ทราบค่า σ จึงประมาณด้วย s

$\hat{\sigma}_{\bar{x}} = s/\sqrt{n} = 700/\sqrt{10} \quad = 221.38$ ตัน

ช่วงเชื่อมั่น 95% ของ μ คือ

$\bar{x} \pm t_{\alpha/2}, \quad s/\sqrt{n}$

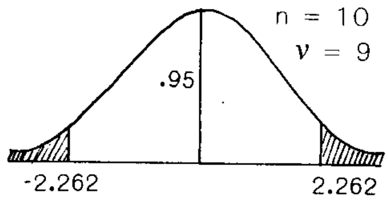
เปิดตาราง t ที่ $\nu = (n - 1) = 9, \quad \alpha = .025$ จะได้ t .05, 9 = 2.262

ดังนั้น ช่วงเชื่อมั่น 95% ของ μ คือ

$11{,}400 \pm (2.262)(221.38)$

$= 11{,}400 \pm 500.76$

$= 10{,}899,\ 11{,}901$ ตัน



จะกล่าวด้วยความเชื่อมั่น 95% ว่า จำนวนวัตถุดิบโดยเฉลี่ยต่อสัปดาห์อยู่ระหว่าง 10,899 ถึง 11,901 ตัน

แบบฝึกหัด

8.22 ฝ่ายการเงินต้องการสำรองเงินค่าใช้จ่ายในการเดินทางสำหรับพนักงานขาย เขาต้องการทราบระยะทางเป็นไมล์ต่อ 1 วัน จากพนักงาน 64 คน ที่สุ่มมาได้ค่าเฉลี่ย 120 ไมล์ต่อวัน และ $\sigma = 12$ จงสร้างช่วงเชื่อมั่น 90% ของค่าเฉลี่ยที่แท้จริง $(117.5325 < \mu < 122.4675)$

8.23 นักประดิษฐ์ไม้ตีกอล์ฟได้ทดลองไม้ชนิดใหม่ที่เขาประดิษฐ์ พบว่า ในจำนวนไม้ 145 อัน ที่นำไปทดลอง สามารถตีลูกไปไกลกว่าเดิมโดยเฉลี่ยถึง 25% และมีค่าเบี่ยงเบนมาตรฐาน 7.2% จงสร้างช่วงเชื่อมั่น 90% ของค่าเฉลี่ยที่แท้จริง $(24.02 < \mu < 25.98)$

8.24 สุ่มตัวอย่าง 49 จำนวนจากประชากร 240 จำนวน ได้ $\bar{x} = 15.8$, $S = 4.2$ จงสร้าง 98% ช่วงเชื่อมั่นของค่าเฉลี่ย $(14.55 < \mu < 17.05)$

8.25 ในการทดลองแรงอัดต่อยาง 81 เส้น พบว่า สามารถรับน้ำหนักได้โดยเฉลี่ย 26 ปอนด์ต่อ 1 ตารางนิ้ว และ $s = 1.8$ ปอนด์ต่อ 1 ตารางนิ้ว จงสร้างช่วงเชื่อมั่น 90% ของค่าเฉลี่ยประชากร $(25.671 < \mu < 26.329)$

8.26 จงหาค่าจากตาราง t สำหรับสร้างช่วงเชื่อมั่นโดยกำหนดข้อมูลให้ดังนี้

ก) n = 5; 99% ง) n = 16 ; 95%

ข) n = 18; 99% จ) n = 18; 95%

ค) n = 27; 95% ฉ) n = 14, 90%

8.27 กำหนดขนาดตัวอย่าง และค่าจากตาราง t จงหาระดับความเชื่อมั่น

ก) n = 20, t = ± 1.729

ข) n = 12, t = ± 2.201

ค) n = 7, t = ± 3.707

8.28 พนักงานการเงินบริษัทหนึ่งต้องการประมาณเวลาที่ใช้สำหรับเรียกเก็บเงินเขาสุ่มมา 24 บัญชี พบว่า ระยะเวลาเรียกเก็บโดยถัวเฉลี่ยคือ 27.3 วัน และ s = 1.9 จงสร้างช่วงเชื่อมั่น 98% ของระยะเวลาเรียกเก็บที่แท้จริง ($26.33 < \mu < 28.27$)

8.29 ให้ n = 8 มีข้อมูลดังนี้

12.1 11.9 12.4 12.3 11.9 12.1 12.4 12.1

จงสร้าง 90% ช่วงเชื่อมั่นของ μ

($12.02 < \mu < 12.28$)

---

4. การสร้างช่วงเชื่อมั่นของสัดส่วนของประชากร จากตัวอย่างขนาดโต

ในบทที่ 7 ได้กล่าวถึง $p = \frac{x}{n}$ ว่าเป็นค่าประมาณของ π, โดยที่ x มีการแจกแจงแบบทวินาม และตามกฎขีดจำกัดเข้าสู่ส่วนกลาง เราทราบว่า เมื่อ n โต ($n\pi \geq 5$, $n(1-\pi) \geq 5$) จะประมาณได้ด้วยการแจกแจงปกติที่มี

$\mu_x = n\pi$ และ $\sigma_x = \sqrt{n\pi(1-\pi)}$

ดังนั้น

$\mu_p = \pi$ และ $\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$

ดังนั้น เมื่อเราต้องการประมาณค่า π ค่าประมาณแบบจุดของ π คือ p นั่นคือ

$\hat{\pi} = p = x/n$

n = จำนวนครั้งที่ทำการทดลอง หรือขนาดตัวอย่าง

X = จำนวนผลสำเร็จ หรือจำนวนลักษณะที่สนใจ

p = สัดส่วนความสำเร็จจากตัวอย่าง (p เป็นค่าสถิติ)

π = สัดส่วนความสำเร็จจากประชากร (π เป็นค่าพารามิเตอร์)

แต่ในการสร้างช่วงเชื่อมั่นของ π จะแบ่งเป็น 2 กรณี คือ

1. ถ้าทราบค่าของ π, $\sigma_p = \sqrt{\pi(1-\pi)/n}$

2) ถ้าไม่ทราบค่า π, ต้องประมาณ π ด้วย p ดังนั้น

$$\hat{\sigma}_p = \sqrt{\hat{\pi}(1-\hat{\pi})/n} = \sqrt{pq/n}$$

ดังนั้น $(1-\alpha)$ 100% ช่วงเชื่อมั่นของ π คือ

1. $p \pm Z_{\alpha/2}\sqrt{\pi(1-\pi)/n}$ เมื่อทราบค่า π

2. $p \pm Z_{\alpha/2}\sqrt{pq/n}$ เมื่อไม่ทราบค่า π

ตัวอย่าง ในการสอบถามพนักงานที่สุ่มมาเป็นตัวอย่าง 75 คน ขององค์กรหนึ่งเกี่ยวกับรูปแบบสวัสดิการ พบว่า มี 40% ที่พอใจรูปแบบที่กำหนดให้ จงหาช่วงเชื่อมั่น 99% ของสัดส่วนของประชากร

$n = 75$ $p = .4$, $q = .6$

ไม่ทราบค่า π, ดังนั้น $\hat{\sigma}_p = \sqrt{pq/n} = \sqrt{(.4)(.6)/75} = .057$

99% ช่วงเชื่อมั่นของ π คือ

$$\begin{aligned} p \pm Z_{.005}\hat{\sigma}_p &= .4 \pm 2.58\ (.057) \\ &= .4 \pm .147 \\ &= .253,\ .547 \end{aligned}$$

กล่าวด้วยความเชื่อมั่น 99% ได้ว่า มีพนักงานเห็นชอบด้วย 25.3 - 54.7%

แบบฝึกหัด

8.30 ในการสอบพนักงานขาย 400 คน พบว่า 64 คน มีลักษณะ "หลงตัวเอง" จงสร้าง 99% ช่วงเชื่อมั่นของสัดส่วนที่แท้จริง (11.5 - 20.6%)

8.31 ผลการสำรวจบัญชีลูกค้าที่กู้เงินไปปลูกบ้าน 120 ราย จากทั้งหมด 1,200 ราย พบว่ามีอยู่ 72% ที่อยู่ในสภาวะ "ไว้วางใจได้"

ก) จงสร้าง 95% ช่วงเชื่อมั่นของสัดส่วน "ที่ไว้วางใจได้" ของประชากร ($.64 < \pi < .80$)

ข) จงหา 95% ช่วงเชื่อมั่นของจำนวนลูกค้า "ที่ไว้วางใจได้" (768 - 960 คน)

8.32 ในการสำรวจร้านอาหาร 64 แห่ง จากทั้งหมด 1,200 แห่ง ในเมืองหนึ่งพบว่า 45% ของร้านอาหารตัวอย่างประสบภาวะขาดทุนเพราะการจัดการผิดพลาด จงหา 95% ช่วงเชื่อมั่นของสัดส่วนที่แท้จริง ($.33 < \pi < .57$)

8.33 จากข้อ 8.33 จงหา 95% ของจำนวนร้านที่จัดการผิดพลาด (396 - 684 ร้าน)

---

5. การกำหนดขนาดตัวอย่างเพื่อใช้ในการประมาณค่า

การกำหนดขนาดตัวอย่างสำหรับประมาณค่าเฉลี่ย

เราทราบว่า ในการประมาณค่า μ ด้วย $\bar{x}$ จะมีความคลาดเคลื่อนจากตัวอย่าง (sampling error) อย่างสูงสุดคือ $\pm 3\sigma_{\bar{x}}$ หรือด้วยความน่าจะเป็น .997 แต่ถ้าเราต้องการใช้ระดับเชื่อมั่นตามปกติที่นิยมใช้ เช่น 90%, 95%, 99% นั่นคือ α = .10, .05 และ .01 ตามลำดับ เราต้องเปิดตารางโค้งปกติที่ $\alpha/2$ คือ $Z_{\alpha/2}$ นั่นเอง ถ้าให้ E คือความคลาดเคลื่อนในการประมาณค่า μ ด้วย $\bar{x}$

$$e = Z_{\alpha/2}\sigma_{\bar{x}}$$

$$e = \frac{Z_{\alpha/2}\sigma}{\sqrt{n}}$$

$$e^2 = \frac{Z^2_{\alpha/2}\sigma^2}{n}$$

$$n = \frac{Z^2_{\alpha/2}\sigma^2}{e^2} \quad \text{หรือ} \quad \left[\frac{Z_{\alpha/2}\sigma}{e}\right]^2$$

ตัวอย่าง

ต้องการประมาณรายได้ต่อเดือนของนักศึกษาที่จบคณะบริหารธุรกิจ โดยทราบว่า σ = 1,000 บาท และต้องการให้ความคลาดเคลื่อนอยู่ภายใน $\pm$ 500 บาท ด้วยความเชื่อมั่น 95% จะต้องสุ่มนักศึกษามาทั้งหมดกี่คน?

$$\alpha = .05,\ Z_{\alpha/2} = Z_{.025} = 1.96,\ e = 500$$

$$n = (Z_{\alpha/2}\sigma/e)^2 = \left(\frac{1.96 \times 1{,}000}{500}\right)^2 = 15.36$$

นั่นคือ ต้องสุ่มมา 16 คน เพื่อสอบถามรายได้ และจะได้ $\bar{x}$ ต่างจาก μ ไม่เกิน 500 บาท ด้วยความเชื่อมั่น 95%

ข้อสังเกต ถ้าไม่ทราบค่า σ และไม่เคยมีการทดลองมาก่อนก็ไม่ทราบว่า σ จะมีค่าประมาณเท่าใด จะใช้ s ประมาณ σ ก็ไม่ได้ เพราะยังไม่ได้สุ่มตัวอย่าง กำลังหาขนาดตัวอย่างอยู่ ก็อาจประมาณคร่าว ๆ ได้ เช่น ถ้าไม่ทราบว่า σ = 1,500 ลองหาพิสัยของค่าเฉลี่ย สมมุติว่ารายได้ต่ำสุดและสูงสุดต่างกัน 5,000 บาท เราทราบว่า พื้นที่ 99.7% อยู่ระหว่าง $\mu \pm 3\sigma$ นั่นคือ ระยะห่างจาก $\mu = 6\sigma'$

$$e = 6\ \sigma = 5{,}000$$

$$\hat{\sigma} \qquad 5{,}000/6 = 833.33 \text{ บาท}$$

การกำหนดขนาดตัวอย่างสำหรับประมาณค่าสัดส่วนของประชากร

ในการประมาณค่า π ด้วย p จะมีความคลาดเคลื่อน $\pm 3\sigma_p$ ด้วยความน่าจะเป็น .997 เมื่อใช้ระดับความน่าจะเป็นเป็น $(1 - \alpha)$

ความคลาดเคลื่อน $\quad e = Z_{\alpha/2}\sigma_p$

$$= Z_{\alpha/2}\sqrt{\pi(1-\pi)/n}$$

นั่นคือ

$$n = \frac{Z^2_{\alpha/2}\pi(1-\pi)}{e^2}$$

แต่ปกติมักไม่ทราบค่า π ถ้าจะประมาณด้วย $p = \frac{x}{n}$ ก็ยังไม่ทราบค่า p อีก เพราะยังไม่ได้สุ่มตัวอย่าง ดังนั้น จะต้องหาค่า $\hat{\pi}(1-\hat{\pi})$ คือ pq ที่โตที่สุด เพื่อจะได้ขนาดตัวอย่างที่โตไว้ก่อน เพื่อความปลอดภัยลองพิจารณาผลคูณของ p และ q เมื่อ p มีค่าต่าง ๆ

ถ้า p = .2, q = .8, pq = .16

ถ้า p = .3, q = .7, pq = .21

$\boxed{\text{ถ้า } p = .5,\ q = .5,\ pq = .25}$

ถ้า p = .6, q = .4, pq = .24

ถ้า p = .9, q = .1, pq = .09

จะเห็นว่าผลคูณ p และ q จะมากที่สุดเมื่อให้ p = .5, q = .5
ดังนั้น หากไม่ทราบค่า π ให้ใช้ $\pi = .5$ หรือ $\frac{1}{2}$

$$n = \frac{Z^2_{\alpha/2}(\frac{1}{2})(\frac{1}{2})}{e^2} = \frac{Z^2_{\alpha/2}}{4e^2}$$

ตัวอย่าง ถ้ามั่นใจว่าสัดส่วนที่แท้จริง = 0.75 จงหาขนาดตัวอย่าง เพื่อใช้ประมาณค่า π โดยให้มีความผิดพลาดเพียง ± .04 (4%) ด้วยความเชื่อมั่น 98%

$$\pi = .75,\ (1-\pi) = .25,\ e = .04,\ Z_{\alpha/2} = Z_{.01} = 2.58$$

$$n = \frac{(2.58)^2(.75)(.25)}{(.04)^2} = 780$$

ถ้าไม่ทราบว่า $\pi = .75$ ให้ $\pi = .5 = \frac{1}{2}$

$$n = \frac{(2.58)^2}{4(.04)^2} = 1040$$

จะเห็นว่า ถ้าไม่ทราบค่า π ต้องสุ่มขนาดโตกว่า เพื่อความมั่นใจ เพราะเราทราบว่า ถ้าขนาดตัวอย่างโตจะทำให้ความคลาดเคลื่อนน้อยลง นั่นคือ p มีโอกาสใกล้ π มากขึ้น

8.34 ถ้า σ = 200, จงหาขนาดตัวอย่างเพื่อใช้ประมาณค่าเฉลี่ยประชากรโดยให้มีความคลาดเคลื่อนภายใน 100 คะแนน ด้วยความเชื่อมั่น 90% (11)

8.35 จงหาขนาดตัวอย่างสำหรับวางสินค้าในตลาดทดลอง เพื่อจะประมาณสัดส่วนความนิยมสินค้าของประชากรให้อยู่ภายใน $\pm$ .03 ด้วยความเชื่อมั่น 95% (1068)

8.36 ถ้า σ = .8 จะต้องใช้ขนาดตัวอย่างเท่าใด เพื่อใช้ประมาณค่าเฉลี่ยให้อยู่ภายใน $\pm$ .25 ด้วยความเชื่อมั่น 98% (56)

แบบฝึกหัดทบทวน

8.37 ถ้าอัตราไหวตัวของหุ้นที่สุ่มมา 49 หุ้น คือ 2.45 บาท ต่อวัน และจากการศึกษาเดิมพบว่า σ = .70 บาท จงสร้างช่วงเชื่อมั่นโดยให้ครอบคลุมค่าเฉลี่ยที่แท้จริง 99.7% ของจำนวนครั้งทั้งหมด ($2.15 < \mu < 2.75$)

8.38 ผู้จัดการธนาคารแห่งหนึ่งพบว่า ผู้ฝากออมทรัพย์ควรจะมีเงินอยู่ในบัญชีโดยถัวเฉลี่ยไม่ต่ำกว่า 1,000 บาท ธนาคารจึงจะไม่ขาดทุน เขาจึงต้องการทราบสัดส่วนที่แท้จริงของบัญชีเงินฝาก 1,000 บาทขึ้นไป เขาจะต้องสุ่มมากี่บัญชีจึงจะให้ค่าที่แท้จริงอยู่ภายใน $\pm$ .04 ด้วยความเชื่อมั่น 95% (601)

8.39 ถ้า 95% ของ μ คือ 84 ถึง 116

75% ของ μ คือ 90.96, 109.04

จงวิจารณ์ข้อดี-เสียของช่วงเชื่อมั่น 2 อันนี้

8.40 จากการสุ่มคนงานมา 81 คนจาก 2,200 คน ในโรงงานทอผ้า พบว่า มีวันหยุดโดยเฉลี่ยเดือนละ 3.2 วัน และมีค่าเบี่ยงเบนมาตรฐาน = 0.9 วัน จงสร้างช่วงเชื่อมั่น 95% ของจำนวนวันหยุดต่อคนต่อเดือน ($3.004 < \mu < 3.396$)

8.41 กำหนดให้ $\bar{X}$ = 96, σ 4.8, n = 36 จงหาระดับความเชื่อมั่นของช่วงเชื่อมั่นต่อไปนี้

ก) (94.4 97.6) ข) (94 98) ค) (95.328 96.672)

(95.5%, 98.76%, 59.9%)

8.42 จงหาระดับความเชื่อมั่นของช่วงเชื่อมั่นต่อไปนี้

ก) $\bar{X} \pm 1.5\sigma_{\bar{X}}$ (86.64 %)

ข) $\bar{X} \pm 1.7\sigma_{\bar{X}}$ (91.08%)

ค) $\bar{X} \pm 2.3\sigma_{\bar{X}}$ (97.86%)

8.43 ให้ X คือยอดขายผลิตภัณฑ์นมของบริษัทหนึ่ง ซึ่งทราบว่า σ = 12.4 ถ้าฝ่ายขายต้องการสุ่มตัวอย่างเพื่อตรวจสอบความนิยม โดยต้องการให้มีความมั่นใจ 98% ว่าค่าประมาณจะต่างจากค่าเฉลี่ยประชากรไม่เกิน ± 3 คะแนน เขาจะต้องใช้ขนาดตัวอย่างเท่าใด? (93)

8.44 โรงงานผลิตแก้วประสบปัญหาผลผลิตมีเปอร์เซ็นต์ชำรุดสูง ฝ่ายเทคนิคได้เสนอวิธีปรับปรุง โดยมั่นใจว่าจะปรับปรุงได้ 75% โรงงานควรทำการทดลองผลิตกี่หน่วยจึงจะมั่นใจ 98% ได้ว่า สัดส่วนชำรุดจากตัวอย่างจะไม่ต่างจากสัดส่วนชำรุดของประชากรเกิน ± .04? (7)

8.45 กรมขนส่งได้สุ่มตัวอย่างรถบัสโดยสารมา 64 คัน พบว่ามีจำนวนผู้โดยสาร โดยเฉลี่ย 3.5 คนต่อ 1 กิโลเมตร และจากการศึกษาก่อนหน้าที่พบว่า σ = 1.6 คน ต่อ 1 กิโลเมตร จงสร้างช่วงเชื่อมั่น 95% ของจำนวนผู้โดยสารโดยเฉลี่ยต่อ 1 กิโลเมตร (3.108, 3.892)

8.46 พนักงานตรวจสอบมาตรฐานสินค้าได้สุ่มสินค้ามา 100 กระสอบ พบว่ามี 35 กระสอบมีคุณภาพต่ำกว่ามาตรฐาน

ก) จงสร้างช่วงเชื่อมั่น 95% ของ π (.258, .444)

ข) ในการสุ่มครั้งต่อไป เขาควรใช้ขนาดตัวอย่างเท่าใด ถ้าต้องการให้ค่าประมาณ p ไม่ต่างจาก π เกิน .05 ด้วยความเชื่อมั่น 95% (350)

8.47 ผู้จัดการโรงงานผลิตเสื้อสำเร็จรูปได้สุ่มบัญชีลูกค้าค้างชำระเงินมา 300 บัญชี พบว่ามีอยู่ 120 บัญชีที่ค้างชำระ 30 วันขึ้นไป จงสร้างช่วงเชื่อมั่น 98% ของสัดส่วนบัญชีที่ค้างชำระเกิน 30 วัน (.3342, .4658)

8.48 บริษัทส่งสินค้าออกได้รับข้อเสนอจากเจ้าของเรือบรรทุกสินค้าว่าจะคิดราคาขนส่งอัตราเดียวกันสำหรับทุก ๆ หีบห่อ โดยมีข้อแม้ว่า ต้องบรรจุหีบห่อให้มีน้ำหนักใกล้เคียงกัน เจ้าของเรือได้สุ่มสินค้ามา 144 หีบ เพื่อหาน้ำหนักที่จะใช้เป็นมาตรฐานต่อไป เขาได้น้ำหนักเฉลี่ย 128.4 ออนซ์ และมีค่าเบี่ยงเบนมาตรฐาน 0.6 ออนซ์ จงหาช่วงโดยรวมค่าเฉลี่ยตัวอย่างซึ่งจะครอบคลุมค่าเฉลี่ยของประชากร 95.5% ของจำนวนครั้งทั้งหมด (128.3, 128.5)

8.49 เจ้าของภัตตาคารต้องการเปลี่ยนเฟอร์นิเจอร์ใหม่ เขาต้องการประมาณรายรับว่าจะคุ้มกับการเปลี่ยนแปลงหรือไม่ เขาอยากทราบรายรับโดยถัวเฉลี่ยต่อลูกค้า 1 คน จากใบเสร็จรับเงินที่สุ่มมาของลูกค้า 8 ราย ได้ค่าเฉลี่ย 105 บาท และมีส่วนค่าเบนมาตรฐาน 25 บาท จงสร้าง 95% ช่วงเชื่อมั่นของรายรับต่อลูกค้า 1 คน (84.1, 125.9)

6. การประมาณค่าผลต่างของค่าเฉลี่ย 2 กลุ่ม

มีบ่อยครั้งที่เราต้องการเปรียบเทียบของ 2 อย่าง เช่น เปรียบเทียบผลผลิตระหว่างการใช้ปุ๋ย (ก) และ (ข) เปรียบเทียบอายุการใช้งานของหลอดไฟ 2 ชนิด เปรียบเทียบความนิยมระหว่างสินค้า 2 ชนิด เป็นต้น ในการเปรียบเทียบระหว่างของ 2 กลุ่มนี้ เราจะต้องเก็บข้อมูลจากตัวอย่างสุ่มมา ดังนั้น ถ้าข้อมูลจากตัวอย่างสุ่ม 2 กลุ่ม ด้วยขนาด n_1 และ n_2 ตามลำดับ ทำให้เราได้ค่าสถิติ $\bar{x}_1$ และ $\bar{x}_2$ เราก็ทราบได้ทันทีว่า ตัวพารามิเตอร์ที่ต้องการประมาณค่าคือ μ_1 และ μ_2 ซึ่งเรามักนำมาเปรียบเทียบกันในรูปผลต่าง คือ $(\mu_1 - \mu_2)$ แต่ถ้าจากตัวอย่างสุ่มขนาด n_1 และ n_2 จากประชากร 2 กลุ่มนั้น ทำให้เราได้ค่าสถิติ $p_1 = \frac{x_1}{n_1}$ และ $p_2 = \frac{x_2}{n_2}$ เราก็ทราบได้ทันทีว่าตัวพารามิเตอร์ที่เราสนใจประมาณค่า คือ $\pi_1 - \pi_2$ ในเบื้องต้นนี้ จะกล่าวถึงวิธีประมาณค่า $\mu_1 - \mu_2$ ก่อน แต่ก่อนอื่นต้องทราบทฤษฎีสำคัญอันหนึ่ง ดังนี้

ถ้า $\bar{x}_1$ และ $\bar{x}_2$ เป็นค่าเฉลี่ยจากตัวอย่างที่เป็น*อิสระ*กัน 2 กลุ่ม ที่มีขนาดตัวอย่าง n_1 และ n_2 ตามลำดับ และเป็นตัวอย่างที่สุ่มมาจากประชากร 2 ประชากร ซึ่งมีความแปรปรวน σ_1^2 และ σ_2^2 ตามลำดับดังนั้น ความคลาดเคลื่อนมาตรฐานของผลต่างค่าเฉลี่ยคือ

$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

ถ้าตัวแปรเชิงสุ่ม X_1, X_2 มาจากประชากรแบบปกติ $\bar{x}_1, \bar{x}_2$ จะมีการแจกแจงแบบปกติด้วย ดังนั้น จึงทำให้ $(\bar{x}_1 - \bar{x}_2)$ มีการแจกแจงแบบปกติด้วย แต่ถ้าไม่ทราบว่า X_1, X_2 มีการแจกแจงแบบปกติ $\bar{x}_1, \bar{x}_2$ จะยังคงมีการแจกแจงแบบปกติโดยประมาณถ้าขนาดตัวอย่าง n_1, n_2 เกิน 30 ดังนั้น $(\bar{x}_1 - \bar{x}_2)$ จึงจะมีการแจกแจงแบบปกติ และมี

ค่าเฉลี่ย $= E(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$

ความแปรปรวน $= \sigma^2_{(\bar{x}_1 - \bar{x}_2)} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

ดังนั้น ในการประมาณค่า $(\mu_1 - \mu_2)$ จะแบ่งเป็น 2 หัวข้อใหญ่ คือ

1) เมื่อทราบค่า σ_1^2 และ σ_2^2

2) เมื่อไม่ทราบค่า σ_1^2, σ_2^2 จึงต้องประมาณจากตัวอย่าง (s_1^2, s_2^2)

สำหรับค่าประมาณแบบจุดของ $(\mu_1 - \mu_2)$ ก็คือ $(\bar{x}_1 - \bar{x}_2)$ ไม่ว่าจะทราบหรือไม่ทราบ σ_1^2, σ_2^2 แต่ค่าประมาณแบบช่วงจะต่างกัน ดังนั้น จะกล่าวถึงการสร้างช่วงเชื่อมั่นของแต่ละกรณี ดังนี้

1. **เมื่อทราบค่า** σ_1^2 , σ_2^2

$(1 - \alpha)$ 100% ช่วงเชื่อมั่นของ $(\mu_1 - \mu_2)$ คือ

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

ตัวอย่าง ต้องการประมาณความแตกต่างของสารนิโคตินในบุหรี่ 2 ชนิด ซึ่งมีข้อมูล ดังนี้

	1	2
ขนาดตัวอย่าง	$n_1 = 100$	$n_2 = 100$
ค่าเฉลี่ยของตัวอย่าง	$\bar{x}_1 = 0.8$	$\bar{x}_2 = 1.0$
ความแปรปรวนของประชากร	$\sigma_1^2 = 0.36$	$\sigma_2^2 = 0.64$

ดังนั้น ช่วงเชื่อมั่น 95% ของผลต่างที่แท้จริงของสารนิโคติน คือ

$$(0.8 - 1.0) \pm 1.96 \sqrt{\frac{0.36}{100} + \frac{0.64}{100}}$$

$$= \quad 0.2 \pm 1.96(0.1) = -0.396, -0.004$$

หรือ $= -0.396 < \mu_1 - \mu_2 < -0.004$

---

2. **เมื่อไม่ทราบค่า** σ_1^2 , σ_2^2 แต่ $n_1 \geq 30$, $n_2 \geq 30$

ให้ประมาณ σ_1^2 , σ_2^2 ด้วย s_1^2 และ s_2^2 ดังนั้น

$(1 - \alpha)$ 100% ช่วงเชื่อมั่นของ $(\mu_1 - \mu_2)$ คือ

$$(\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

จากตัวอย่างเดิม ถ้าไม่ทราบค่า σ_1^2, σ_2^2 และจากตัวอย่างได้ $s_1^2 = .40$, $s_2^2 = .75$

ดังนั้น $\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{.40}{100} + \frac{.75}{100}} = .107$

ดังนั้น 95% ช่วงเชื่อมั่นของ $(\mu_1 - \mu_2)$ คือ

$$(0.8 - 1.0) \pm 1.96\ (.107)$$

$$= -0.2 + .21 = -0.41,\ 0.01$$

หรือ $-0.41 < \mu_1 - \mu_2 < 0.01$

3. เมื่อไม่ทราบค่า σ_1^2, σ_2^2 แต่ขนาดตัวอย่างของกลุ่มใดกลุ่มหนึ่ง หรือทั้ง 2 กลุ่มน้อยกว่า 30

จะต้องมีข้อสมมุติเพิ่มเติมดังนี้

1) ประชากรทั้ง 2 กลุ่ม มีการแจกแจงแบบปกติ หรือมีรูปร่างใกล้เคียงโค้งปกติ

2) ทั้ง 2 ประชากรมีความแปรปรวนเท่ากัน คือ σ^2 (แต่ไม่ทราบค่า)

3) สุ่มตัวอย่างที่เป็นอิสระกันมา 2 กลุ่มด้วยขนาด n_1, n_2

ดังนั้น $\sigma^2_{(\bar{x}_1 - \bar{x}_2)} = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$

ต้องประมาณค่า σ^2 โดย

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

ถ้า $n_1 = n_2 = n$

$$s^2 = \frac{(n - 1)s_1^2 + (n - 1)s_2^2}{2n - 2} = \frac{(n - 1)(s_1^2 + s_2^2)}{2(n - 1)}$$

$$= \frac{s_1^2 + s_2^2}{2}$$

ดังนั้น $(1 - \alpha)$ 100% ช่วงเชื่อมั่นของ $(\mu_1 - \mu_2)$ คือ

$$(\bar{x}_1 - \bar{x}_2) \pm t^{\alpha/2}_{(n_1 + n_2 - 2)} \sqrt{s_p^2 (\frac{1}{n_1} + \frac{1}{n_2})}$$

ตัวอย่าง $n_1 = 17$ $\bar{x}_1 = 545$ $s_1 = 50$ $t_{.025,\ 30} = 2.042$

$n_2 = 15$ $\bar{x}_2 = 495$ $s_2 = 55,$

ดังนั้น $s_p^2 = \dfrac{16(50^2) + 14(55^2)}{16 + 14} = 2745$

$$\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{2745 \left(\frac{1}{17} + \frac{1}{15}\right)} = \sqrt{344.47} = 18.56$$

นั่นคือ 95% ช่วงเชื่อมั่นของ $(\mu_1 - \mu_2)$

$= (545 - 495) \pm (2.042)\ (18.56)$

$= 12.1,\ 87.9$

หรือ $12.1 < \mu_1 - \mu_2 < 87.9$

4. เมื่อข้อมูลที่สุ่มมา 2 กลุ่มไม่เป็นอิสระกัน หรือเป็นข้อมูลแบบจับคู่ (Matched Samples หรือ Paired Observations)

กรณีนี้จะไม่สนใจค่าของ σ_1^2, σ_2^2 ว่าทราบหรือไม่ แต่ปกติมักไม่ทราบ เพราะจะไม่ใช้ค่า σ_1^2, σ_2^2 ทั้งนี้เนื่องจาก X, Y (แทน X_1, X_2) ไม่เป็นอิสระกัน และมีความสัมพันธ์กันเป็นคู่ ๆ รวมได้ n คู่ คือ $(x_1, y_1), (x_2, y_2)....(x_n, y_n)$ จึงเรียกว่า**ข้อมูลแบบจับคู่** ข้อมูลแบบนี้มักเก็บจากหน่วยทดลองอันเดียวกันในรูป ก่อน-หลัง การทดลอง ดังนั้น จะต้องสร้างตัวแปรเชิงสุ่มตัวใหม่ที่เป็นอิสระกัน คือ $d_i = y_i - x_i, i = 1, ..., n$ รวม d_i ได้ n ตัว แล้วหาค่าเฉลี่ยและความแปรปรวนของ d_i ดังนี้

$\bar{d} = \Sigma_1^n\ d_i/n$, $\bar{d}$ จะเป็นค่าประมาณของ $\mu_1 - \mu_2$

และ $s_{\bar{d}}^2 = s_d^2/n$

โดยที่ $s_d^2 = \sum_{i=1}^{n} (d_i - \bar{d})^2/(n - 1)$

$$= \frac{\sum_{i=1}^{n} d_i^2 - (\Sigma d_i)^2/n}{(n - 1)}$$

ดังนั้น $(1 - \alpha)$ 100% ช่วงเชื่อมั่นของ $(\mu_1 - \mu_2)$ คือ

$$\bar{d} \pm t_{\alpha/2, (n-1)} \times s_{\bar{d}}$$

ตัวอย่าง น้ำหนักก่อนและหลังเข้าโปรแกรมลดน้ำหนักของผู้ทดลอง 10 คน มีดังนี้

(x) ก่อนเข้าโปรแกรม	189	202	220	207	194	177	193	202	208	233
(y) หลังเข้าโปรแกรม	170	179	203	192	172	161	174	187	186	204
$d_i = (x_i - y_i)$	19	23	17	15	22	16	19	15	22	29

$$\Sigma d_i = (19 + 23 + \ldots + 29) = 197$$

$$\bar{d} = 197/10 = 19.7$$

$$\Sigma d_i^2 = (19^2 + 23^2 + \ldots + 29^2) = 4055$$

$$s_d^2 = \frac{\Sigma d_i^2 - (\Sigma d_i)^2/n}{n-1} \quad \frac{4055 - (197)^2/10}{9} = 19.34$$

$$s_{\bar{d}} = \sqrt{\frac{s_d^2}{n}} = \sqrt{\frac{19.34}{10}} = \sqrt{1.934} = 1.39, \quad t_{9, .05} = 1.833$$

ดังนั้น 90% ช่วงเชื่อมั่นของ $(\mu_1 - \mu_2)$ คือ

$$19.7 \pm (1.833)(1.39)$$

$$= 19.7 \pm 2.55$$

$$= 17.15 \; ; \; 22.25 \quad \text{หรือ } 17.15 < \mu_1 - \mu_2 < 22.25$$

นั่นคือ กล่าวด้วยความเชื่อมั่น 90% ว่าผู้เข้าโปรแกรมจะสามารถลดน้ำหนักได้โดยเฉลี่ยระหว่าง 17.15 ถึง 22.25 ปอนด์

7. การประมาณค่าผลต่างของสัดส่วน 2 กลุ่ม

เมื่อต้องการเปรียบเทียบระหว่าง 2 อัตรา หรือการเปรียบเทียบเปอร์เซนต์ของ 2 กลุ่มข้อมูลจะมีการแจกแจงแบบทวินาม นั่นคือ เราต้องการเปรียบเทียบพารามิเตอร์ของการแจกแจงแบบทวินาม 2 กลุ่ม คือ $(\pi_1 - \pi_2)$ ในเมื่อ π_1, π_2 คือสัดส่วนความสำเร็จในประชากรที่ 1 และ 2 ตามลำดับ ดังนั้น เราจะต้องทำการทดลองแบบทวินามและเก็บข้อมูลมา และรวบรวมข้อมูลใส่ตาราง ดังนี้

	กลุ่มที่ 1	กลุ่มที่ 2	
ความสำเร็จ	x_1	x_2	$p_1 = \frac{x_1}{n_1}$
ความล้มเหลว	$n_1 - x_1$	$n_2 - x_2$	$p_2 = \frac{x_2}{n_2}$
	n_1	n_2	

$$\mu_{(p_1 - p_2)} = \pi_1 - \pi_2$$

$$\sigma^2_{(p_1 - p_2)} = \sigma^2_{p_1} + \sigma^2_{p_2} \quad = \quad \frac{\pi_1(1 - \pi_1)}{n_1} + \frac{\pi_2(1 - \pi_2)}{n_2}$$

แต่ปกติจะไม่ทราบค่า π_1, π_2 ดังนั้น จึงต้องประมาณด้วย p_1, p_2 ตามลำดับนั่นคือ

$$\hat{\sigma}_{(p_1 - p_2)} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

และ $(1 - \alpha)$ 100% ช่วงเชื่อมั่นของ $(\pi_1 - \pi_2)$ คือ

$$(p_1 - p_2) \pm Z_{\alpha/2}\, \hat{\sigma}_{(p_1 - p_2)}$$

ตัวอย่าง บริษัทหนึ่งเปิดหลักสูตรอบรมพนักงานเกี่ยวกับการจัดการ เมื่อจบหลักสูตรได้สอบถามพนักงานที่สุ่มมาเป็นตัวอย่าง 100 คน มี 50 คน ที่มีความพอใจนายจ้าง และเมื่อใช้คำถามเดียวกันถามพนักงานที่สุ่มมาอีก 100 คน จากอีกบริษัทหนึ่งซึ่งไม่มีหลักสูตรอบรมดังกล่าว มีพนักงาน 40 คน ที่มีความพอใจนายจ้าง จงสร้าง 95% ช่วงเชื่อมั่นของความแตกต่างของสัดส่วนความพอใจนายจ้าง

$$p_1 = \frac{50}{100} = .5, \; p_2 = \frac{40}{100} = .4, \; Z_{.025} = 1.96$$

$$\hat{\sigma}_{(p_1 - p_2)} = \sqrt{\frac{(.5)(.5)}{100} + \frac{(.4)(.6)}{100}} = \sqrt{\frac{.49}{100}} = .07$$

ดังนั้น 95% ช่วงเชื่อมั่นของ $(\pi_1 - \pi_2)$ คือ

$$(.5 - .4) \pm (1.96)(.07) = 0.1 \pm .1372$$

$$= -0.372 \; ; \; 0.2372$$

หรือ $\quad -0.0372 < (\pi_1 - \pi_2) < 0.2372$

แบบฝึกหัด

8.50 เครื่องจำหน่ายน้ำอัดลมแบบอัตโนมัติ 2 เครื่อง ได้รับการติดตั้งให้จำหน่ายน้ำในปริมาณเท่ากัน แต่ลูกค้าหลายรายสังเกตว่าให้ปริมาณต่างกัน ฝ่ายจัดจำหน่ายจึงทำการสุ่มจากทั้ง 2 เครื่อง ได้ข้อมูลดังนี้

	เครื่องที่ 1	เครื่องที่ 2
ขนาดตัวอย่าง	$n_1 = 40$	$n_2 = 60$
ค่าเฉลี่ย	$\bar{x}_1 = 22$	$\bar{x}_2 = 24$
ความแปรปรวน	$\sigma_1^2 = 50$	$\sigma_2^2 = 45$

ก) จงหาค่าประมาณแบบจุดของความแตกต่างที่แท้จริง และความคลาดเคลื่อนสูงสุดจากการประมาณค่า (4.24)

ข) จงหา 95% ช่วงเชื่อมั่นของความแตกต่างของปริมาณเฉลี่ยจาก 2 เครื่อง (−4.77, 0.77)

8.51 ให้ X_1, X_2 เป็นจำนวนผลผลิตรายชั่วโมงของคนงานชายและหญิงตามลำดับ และสมมุติให้มีความแปรปรวนเท่ากัน คือ $\sigma_1^2 = \sigma_2^2 = 64$ จากการสุ่มผลผลิตชายและหญิงมาเพศละ 32 ชั่วโมง ทำงานพบว่า $\bar{x}_1 = 85$ หน่วย, $\bar{x}_2 = 78$ หน่วย

ก) จงหาค่าประมาณแบบจุดของความแตกต่างที่แท้จริง และความคลาดเคลื่อนที่เป็นไปได้ในการประมาณค่า (5.64)

ข) จงสร้าง 95% ช่วงเชื่อมั่นของ $(\mu_1 - \mu_2)$ (3.32, 10.68)

8.52 ให้ X_1, X_2 แทนน้ำหนักของคนงานชายและหญิงตามลำดับ และสมมุติให้ $\sigma_1^2 = \sigma_2^2 = 200$ และ $n_1 = n_2 = 100$ พบว่า $\bar{x}_1 = 70$ กิโลกรัม $\bar{x}_2 = 50$ กิโลกรัม จงสร้าง 95% ช่วงเชื่อมั่น ของ $(\mu_1 - \mu_2)$ (16.08, 23.92)

8.53 เชื่อว่าพนักงานที่ใช้พิมพ์ดีดไฟฟ้าจะพิมพ์งานได้รวดเร็วกว่าพิมพ์ดีดธรรมดาข้อมูลที่ได้จากการทดลอง มีดังนี้

	เครื่องไฟฟ้า	แบบธรรมดา
ขนาดตัวอย่าง	$n_1 = 25$	$n_2 = 25$
ค่าเฉลี่ย (จำนวนคำต่อ 1 นาที)	$\bar{x}_1 = 58$	$\bar{x}_2 = 55$
ความแปรปรวน	$\sigma_1^2 = 78$	$\sigma_2^2 = 66$

สมมุติว่า อัตราการพิมพ์ดีดมีการแจกแจงแบบปกติ จงหา 95% ช่วงเชื่อมั่นของความแตกต่างที่แท้จริงระหว่างอัตราพิมพ์จากเครื่องไฟฟ้า และเครื่องธรรมดา (−1.704, 7.704)

8.54 ซื้อเครื่องจักรชนิดเดียวกันมา 2 เครื่อง เมื่อสำรวจความขัดข้องของเครื่องทั้ง 2 ได้ข้อมูลดังนี้

	เครื่อง I	เครื่อง II
ขนาดตัวอย่าง	$n_1 = 16$	$n_2 = 16$
ค่าเฉลี่ย (นาที)	$\bar{x}_1 = 55$	$\bar{x}_2 = 85$
ความแปรปรวน	$\sigma_1^2 = 1600$	$\sigma_2^2 = 2000$

สมมุติว่าระยะเวลาความขัดข้องมีการแจกแจงแบบปกติ จงสร้าง 98% ช่วงเชื่อมั่นของความแตกต่างที่แท้จริงของเครื่องทั้ง 2 (-4.89, 64.89)

8.55 สุ่มตัวอย่างครัวเรือนมา 300 ครัวเรือน เพื่อประมาณปริมาณไฟฟ้าที่ใช้ต่อครัวเรือนในเดือนเมษายน และได้สุ่มครัวเรือนมาอีก 400 ครัวเรือน ในปีถัดมา เพื่อดูปริมาณไฟฟ้าที่ใช้ในเดือนเมษายน ได้ข้อมูลดังนี้

ปีก่อน	ปีนี้
$n_1 = 300$	$n_2 = 400$
$\bar{x}_1 = 1{,}252$ kwh	$\bar{x}_2 = 1{,}325$ kwh
$s_1 = 257$ kwh	$s_2 = 263$ kwh

ก) จงสร้าง 95% ช่วงเชื่อมั่นของอัตราการเปลี่ยนแปลงการใช้ไฟต่อครอบครัวระหว่าง 2 ปี (34.13, 111.87)

ข) ถ้าในปีนี้สำรวจการใช้ไฟจากครัวเรือนเดิม 300 ครัวเรือนนั้น จะใช้วิธีการประมาณค่าแบบเดิมในข้อ (ก) ได้หรือไม่?

8.56 ถ้า 41% ของคู่สมรสที่หย่ากันจะไม่มีบุตร และถ้าในเมืองหนึ่ง พบว่ามีคู่สมรสที่หย่ากันโดยไม่มีบุตรอยู่ 437 คู่ จากที่สุ่ม 1,000 คู่ จงสร้าง 98% ช่วงเชื่อมั่นของเปอร์เซนต์คู่สมรสที่หย่ากันโดยไม่มีบุตร (.401 $< \pi <$ .473)

8.57 จากข้อ 8.56 ได้มีการสุ่มตัวอย่างคู่สมรสที่หย่ากันในอีกเมืองหนึ่ง จำนวน 800 คู่ พบว่ามี 346 คู่ ไม่มีบุตรด้วยกัน

ก) จงสร้าง 95% ช่วงเชื่อมั่นของ $\pi_1 - \pi_2$ (-.0412, .0502)

ข) จากช่วงเชื่อมั่นในข้อ (ก) แสดงว่าอัตราการหย่าร้างโดยไม่มีบุตรของ 2 เมืองนี้ใกล้เคียงกันไหม?

8.58 จำนวนนักศึกษาที่จบจากคณะบริหารธุรกิจซึ่งเข้าทำงานในหน่วยงานหนึ่งได้ 2 ปี และยังคงปฏิบัติงานอยู่ โดยจำแนกตามระดับปริญญา มีดังนี้

ปริญญา	จำนวนที่จ้าง	จำนวนที่ปฏิบัติงานในปัจจุบัน
ปริญญาตรี	205	123
ปริญญาโท	50	18

ก) จงสร้าง 99% ช่วงเชื่อมั่นของผลต่างของสัดส่วนผู้จบ 2 ระดับปริญญา และยังคงปฏิบัติงานอยู่ภายหลังเข้า 2 ปี **(.044, .436)**

ข) ข้อมูลนี้ได้มาจากประชากรแบบใด?

8.59 ในการเปรียบเทียบรถยนต์ 2 ชนิด ว่ามีระบบควบคุมของเสียต่างกันหรือไม่ได้ข้อมูลดังนี้

	รถยนต์ I	รถยนต์ II
จำนวนวัน	$n_1 = 16$	$n_1 = 16$
ดัชนีของเสีย	$\bar{x}_1 = 60$	$\bar{x}_1 = 55$
	$s_1 = 9$	$s_2 = 9$

จงสร้างช่วงเชื่อมั่น 95% ของความแตกต่างของดัชนีของเสีย (-1.23, 11.23)

8.60 ต้องการทราบความแตกต่างของการเรียนระหว่างนักเรียนชายและหญิงโดยสมมุติว่าคะแนนสอบมีการแจกแจงแบบปกติ ได้ข้อมูล ดังนี้

คะแนนของนักเรียนชาย	2.9	3.1	2.7	3.3	3.0
คะแนนของนักเรียนหญิง	3.6	2.8	3.6	3.2	2.8

จงสร้าง 95% ช่วงเชื่อมั่นของความแตกต่างของคะแนนสอบ (-0.27, 0.67)

8.61 ในการทดสอบเพื่อเปรียบเทียบการเรียนรู้คำศัพท์และการพูดของเด็กซึ่งได้รับวิธีการอบรมที่ต่างกัน 2 กลุ่ม โดยสมมุติว่า คะแนนมีการแจกแจงแบบปกติและมีความแปรปรวนเท่ากัน ได้ข้อมูลดังนี้

	กลุ่มที่ 1	กลุ่มที่ 2
ขนาดตัวอย่าง	$n_1 = 10$	$n_2 = 8$
ค่าเฉลี่ย	$\bar{x}_1$ 95	$\bar{x}_2 = 97$
ความแปรปรวน	$s_1^2 = 40$	$s_2^2 = 36$

ก) จงสร้าง 95% ช่วงเชื่อมั่นของความแตกต่างของการเรียนรู้ระหว่าง 2 กลุ่ม

(-8.22, 4.22)

ข) จากช่วงเชื่อมั่นในข้อ (ก) มีหลักฐานแสดงว่าการเรียนรู้ของเด็ก 2 กลุ่ม มีความแตกต่างอย่างเห็นได้ชัดหรือไม่?

8.62 ต้องการทราบว่าเด็ก 16 คน สามารถเรียนรู้ภาษาอังกฤษ และคณิตศาสตร์ ด้วยความสำเร็จเท่าเทียมกันหรือไม่ ได้คะแนนสอบของ 2 วิชา ดังนี้

นักเรียน	ภาษาอังกฤษ	คณิตศาสตร์
A	84	84
B	55	57
C	85	90
D	98	97
E	80	74
F	55	53
G	80	75
H	64	63
I	91	90
J	85	82
K	90	88
L	94	98
M	75	77
N	86	90
O	91	85
P	92	86

จงสร้าง 95% ช่วงเชื่อมั่นของความแตกต่างของประชากร (.59, 2.59)