



การประมาณค่าและการทดสอบสมมติฐาน (Estimation and Hypothesis testing)





เนื้อหา

- การประมาณค่า

- การประมาณค่าเฉลี่ยประชากรกลุ่มเดียว

- การทดสอบสมมติฐาน

- ขั้นตอนการทดสอบสมมติฐาน
- การทดสอบสมมติฐานค่าเฉลี่ยประชากรเดียว

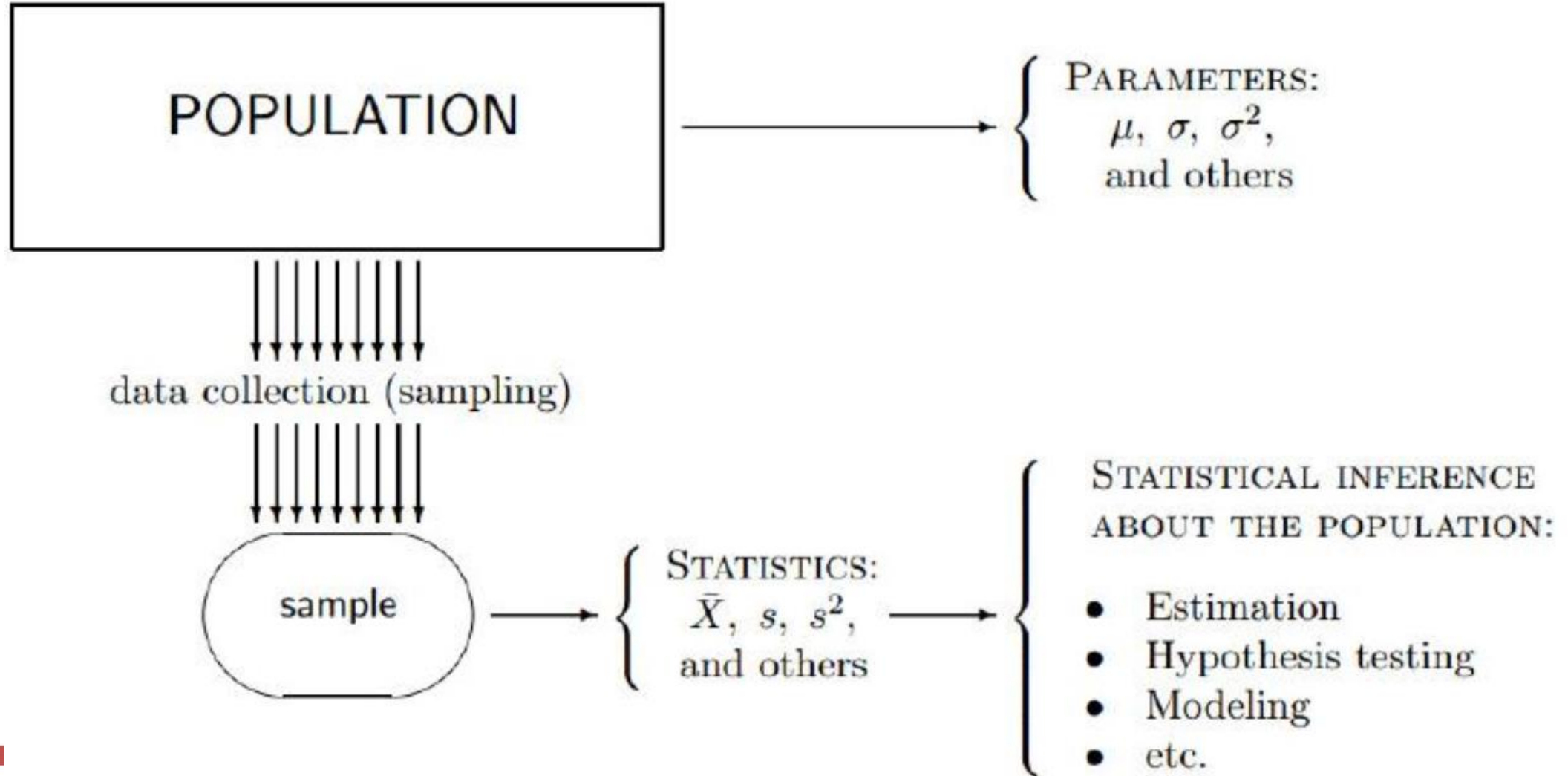




ทบทวนสัญลักษณ์

	ประชากร	ตัวอย่าง
ขนาด	N	n
ค่าเฉลี่ย	$\mu = \frac{\sum_{i=1}^N x_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n - 1}$
ความแปรปรวน	$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$	$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$
ส่วนเบี่ยงเบนมาตรฐาน	$\sigma = \sqrt{\sigma^2}$	$s = \sqrt{s^2}$
สัดส่วน	P	$\hat{p}$

พารามิเตอร์และสถิติ





สถิติเชิงอนุมาน

สถิติเชิงอนุมาน
(Inferential Statistic)

การประมาณค่า
(Estimation)

การทดสอบสมมติฐาน
(Hypothesis Testing)



การประมาณค่า



การประมาณค่า

- การประมาณค่า (Estimation) คือ การประมาณค่าพารามิเตอร์ของประชากร โดยใช้ข้อมูลที่ได้จากชุดตัวอย่าง
- เช่น การประมาณยอดขายเฉลี่ยรายเดือนของบริษัทในปีหน้าเพื่อการวางแผนการผลิต
- ต้องทำอย่างมีหลักเกณฑ์เพื่อให้ค่าประมาณที่ได้มีความคลาดเคลื่อนจากค่าพารามิเตอร์ที่แท้จริงน้อยที่สุด
- ทางสถิติการประมาณค่าพารามิเตอร์สามารถแบ่งได้เป็น 2 ชนิด คือ
 1. การประมาณค่าแบบจุด (Point estimation)
 2. การประมาณค่าแบบช่วง (Interval estimation)



1.การประมาณค่าแบบจุด (Point Estimation)

- การประมาณแบบจุด เป็นการประมาณค่าพารามิเตอร์ด้วยค่าประมาณเพียงค่าเดียว โดยใช้ข้อมูลจากตัวอย่าง

ตัวประมาณแบบจุดของ

ค่าเฉลี่ยของประชากร (μ) $\longleftarrow$ ค่าเฉลี่ยของตัวอย่าง ($\bar{x}$)

ค่าสัดส่วนของประชากร (P) $\longleftarrow$ ค่าสัดส่วนของตัวอย่าง ($\hat{p}$)

ค่าความแปรปรวนของประชากร (σ^2) $\longleftarrow$ ค่าความแปรปรวนของตัวอย่าง (s^2)

ค่าส่วนเบี่ยงเบนมาตรฐานของประชากร (σ) $\longleftarrow$ ค่าส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง (s)



1.การประมาณค่าแบบจุด (Point Estimation)

- ค่าประมาณแบบจุดนี้อาจจะมีค่าเท่าหรือไม่เท่ากับค่าพารามิเตอร์ก็ได้
 - มีโอกาสที่จะคลาดเคลื่อนไปจากค่าพารามิเตอร์ที่แท้จริงได้มาก
 - ถ้าชุดตัวอย่างเป็นตัวแทนที่ไม่ดีของประชากรแล้ว ค่าประมาณที่ได้จะคลาดเคลื่อนไปจากค่าพารามิเตอร์ที่แท้จริงมาก
 - ดังนั้น การประมาณชนิดนี้จึงมีความเสี่ยงสูง
-



ตัวอย่าง

- สุ่มตัวอย่างนักศึกษา 10 คน เพื่อทำการประมาณค่าใช้จ่ายเฉลี่ยต่อวันของนักศึกษาได้ข้อมูลดังนี้

120 145 185 100 95 65 120 100 80 90

จากข้อมูลดังกล่าวต้องการหาค่าเฉลี่ย (μ)



2.การประมาณค่าแบบช่วง (Interval estimation)

การประมาณค่าแบบช่วงเป็นการใช้ช่วงตัวเลขประมาณพารามิเตอร์ของประชากร ช่วงตัวเลข ซึ่งเรียกว่าช่วงความเชื่อมั่น (confidence interval) จะบอกค่าต่ำสุด ค่าสูงสุด ที่คาดว่าจะคลุม พารามิเตอร์นั้น ๆ และจะบอกระดับความเชื่อมั่นของการประมาณไว้ในรูป $P (L < \theta < U) = 1 - \alpha$

ตัวอย่าง

- ซื่ออินเทอร์เน็ตความเร็ว 10MB ต่อวินาที เร็ว 10MB ต่อวินาที จริงหรือไม่
- การพยากรณ์อากาศ แม่นยำมากน้อยเพียงใด
- ความน่าจะเป็นที่ค่านี้จะเข้าใกล้ความจริงมีมากเพียงใด



2.การประมาณค่าแบบช่วง (Interval estimation)

- เมื่อประมาณค่าพารามิเตอร์ θ แล้ว ค่าที่แท้จริงของ θ มีค่าเป็นเท่าใด
- นักสถิติสามารถใช้การประมาณช่วงความเชื่อมั่น ซึ่งจะครอบคลุมค่าพารามิเตอร์จากข้อมูลที่เก็บได้จากตัวอย่าง
- การประมาณค่าพารามิเตอร์แบบช่วง เป็นการใช่วงประมาณที่เรียกว่า ช่วงความเชื่อมั่น (confidence interval)

$$L < \theta < U$$

โดย L คือค่าประมาณต่ำสุด และ U คือค่าประมาณสูงสุดของช่วงการประมาณ



ตัวอย่าง

ค่าใช้จ่ายเฉลี่ยต่อเดือนของนักศึกษาที่มีการแจกแจงปกติ ด้วยค่าส่วนเบี่ยงเบนมาตรฐาน 1600 บาท สุ่มตัวอย่างจำนวนนักศึกษา 1000 คน ค่าใช้จ่ายโดยเฉลี่ยต่อเดือน 3000 บาท จงประมาณค่าใช้จ่ายเฉลี่ยต่อเดือนของนักศึกษากลุ่มนี้ด้วยความเชื่อมั่น 95 %



ระดับความเชื่อมั่นร้อยละ 95

- ความน่าจะเป็นที่ช่วงประมาณที่สร้างขึ้นจะคลุมค่าพารามิเตอร์ 0.95 และไม่คลุมค่าพารามิเตอร์ 0.05



ระดับนัยสำคัญ คือ ค่าโอกาสที่ยอมให้ผิดพลาดได้



ระดับความเชื่อมั่น

ระดับความเชื่อมั่น	ระดับนัยสำคัญ (α)
ร้อยละ 95	$\frac{100 - 95}{100} = 0.05$
ร้อยละ 99	$\frac{100 - 99}{100} = 0.01$
ร้อยละ 92	$\frac{100 - 92}{100} = 0.08$
ร้อยละ 90	$\frac{100 - 90}{100} = 0.10$



สัมประสิทธิ์ความเชื่อมั่น (Confidence coefficient)

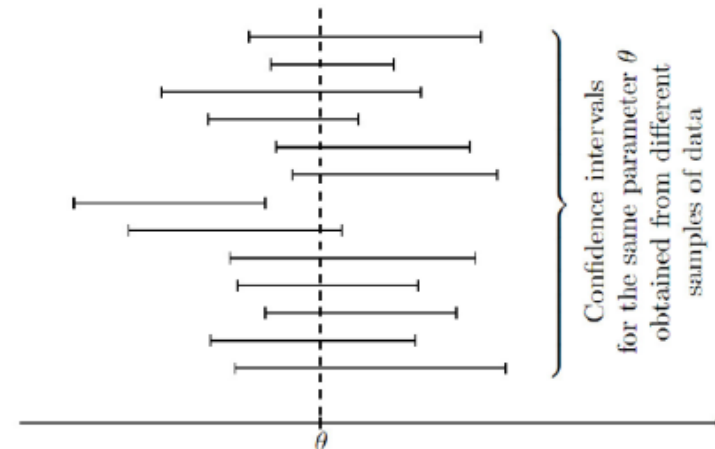
สัมประสิทธิ์ความเชื่อมั่น (Confidence coefficient) หมายถึง ความน่าจะเป็นที่ช่วงประมาณจะครอบคลุมค่าพารามิเตอร์ของประชากร โดยมีค่าเท่ากับ $1 - \alpha$ นั่นคือ $P(L \leq \theta \leq U) = 1 - \alpha$

เช่น ค่าสัมประสิทธิ์ความเชื่อมั่น $(1 - \alpha)$ เท่ากับ 0.90

- โอกาสที่จะครอบคลุมค่าพารามิเตอร์ที่เราสนใจ 90%
- โอกาสที่ไม่ครอบคลุม 10%

หมายความว่า สุ่มตัวอย่าง 100 ครั้ง มี 90 ครั้ง ที่ค่าประมาณแบบช่วงจะคลุม θ และอีก 10 ครั้ง ไม่คลุมค่า θ

และ เรียก α ว่าระดับนัยสำคัญ





แบบฝึกหัด

1. คะแนนทดสอบวิชาสถิติและการวางแผนการทดลองของนักศึกษา จำนวน 12 คน ข้อมูลเป็นดังนี้

16 19 15 14 13 11 13 8 12 17 16 14

จงหาค่าประมาณ คะแนนการสอบโดยเฉลี่ยของนักศึกษากลุ่มนี้





การประมาณค่าเฉลี่ยของประชากร 1 กลุ่ม



ค่าประมาณแบบช่วงของค่าเฉลี่ยประชากรหนึ่งกลุ่ม

Nakhon Pathom Rajabhat University



- กรณีทราบค่าความแปรปรวนของประชากร (σ^2)

$$\mu = \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

- กรณีไม่ทราบค่าความแปรปรวนของประชากร

- ขนาดตัวอย่างใหญ่ ($n \geq 30$)

$$\mu = \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

- ขนาดตัวอย่างเล็ก ($n < 30$)

$$\mu = \bar{x} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$



ตัวอย่าง

กรณี 1

ทราบความแปรปรวนหรือส่วน
เบี่ยงเบนมาตรฐานของประชากร

ในการศึกษาเรื่องน้ำหนักของสุนัขพันธุ์ชิวาวาในประเทศไทย ซึ่งมีค่าส่วนเบี่ยงเบนมาตรฐานของน้ำหนักเท่ากับ 2.16 กิโลกรัม โดยสุ่มสุ่มมา 45 ตัว พบน้ำหนักเฉลี่ยเท่ากับ 4.65 กิโลกรัม จงหา 90% ช่วงความเชื่อมั่นน้ำหนักสุนัขพันธุ์ชิวาวา

กรณี 2

ไม่ทราบความแปรปรวนหรือส่วน
เบี่ยงเบนมาตรฐานของประชากร
 $n \geq 30$

ในการศึกษาเรื่องน้ำหนักของสุนัขพันธุ์ชิวาวาในประเทศไทย โดยสุ่มสุ่มมา 45 ตัว พบน้ำหนักเฉลี่ยเท่ากับ 4.65 กิโลกรัม และส่วนเบี่ยงเบนมาตรฐานของน้ำหนักเท่ากับ 2.16 กิโลกรัม จงหา 90% ช่วงความเชื่อมั่นน้ำหนักสุนัขพันธุ์ชิวาวา

กรณี 3

ไม่ทราบความแปรปรวนหรือส่วน
เบี่ยงเบนมาตรฐานของประชากร
 $n < 30$

ในการศึกษาเรื่องน้ำหนักของสุนัขพันธุ์ชิวาวาในประเทศไทย โดยสุ่มสุ่มมา 19 ตัว พบน้ำหนักเฉลี่ยเท่ากับ 4.65 กิโลกรัม และส่วนเบี่ยงเบนมาตรฐานของน้ำหนักเท่ากับ 2.16 กิโลกรัม จงหา 90% ช่วงความเชื่อมั่นน้ำหนักสุนัขพันธุ์ชิวาวา



ตัวอย่าง

กรณี 1

ทราบความแปรปรวนหรือส่วน
เบี่ยงเบนมาตรฐานของประชากร

- สมมติว่าระยะเวลา(นาทึ) ที่ลูกค้าของธนาคารสาขาหนึ่งใช้ในการติดต่อกับพนักงานของธนาคารมีการแจกแจงแบบปกติ ที่มีค่าเบี่ยงเบนมาตรฐานเท่ากับ 3 นาที ถ้าสุ่มเลือกลูกค้าของธนาคารสาขานั้นจำนวน 16 คน พบว่า เวลาเฉลี่ยที่ลูกค้าใช้ในการติดต่อกับพนักงานธนาคารเท่ากับ 10 นาที ให้คำนวณช่วงความเชื่อมั่นที่ระดับ 96 %ของเวลาเฉลี่ยที่ลูกค้าให้ติดต่อกับพนักงานธนาคาร

$$\mu = \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$



ตัวอย่าง

กรณี 2
ไม่ทราบความแปรปรวนหรือส่วน
เบี่ยงเบนมาตรฐานของประชากร
 $n \geq 30$

- สุ่มตัวอย่างบุหรี่ยี่ห้อหนึ่งจำนวน 36 ราย พบว่ามีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของจำนวนนิโคตินที่มีอยู่ในบุหรี่ยี่ห้อเป็น 22 และ 4 mg. ตามลำดับ จงหาขอบเขตที่เชื่อมั่นได้ 90 % ของปริมาณนิโคตินที่แท้จริงที่มีอยู่ในบุหรี่ยี่ห้อนี้

$$\mu = \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$



ตัวอย่าง

กรณี 3
ไม่ทราบความแปรปรวนหรือส่วน
เบี่ยงเบนมาตรฐานของประชากร
 $n < 30$

- สุ่มตัวอย่างคะแนนสอบวิชาคณิตศาสตร์ของนักศึกษาจำนวน 25 คน ได้ค่าเฉลี่ยคะแนน 75 คะแนน และส่วนเบี่ยงเบนมาตรฐาน 10 คะแนน จงหาช่วงความเชื่อมั่นที่ 98 % ของค่าเฉลี่ยคะแนนสอบวิชาคณิตศาสตร์นี้

$$\mu = \bar{x} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

