

การทดสอบสมมติฐานสำหรับ ค่าเฉลี่ยจากกลุ่มตัวอย่างเดียว

สถิติสำหรับจิตวิทยา 1

สันทัด พรประเสริฐมานิต

โครงร่างการนำเสนอ

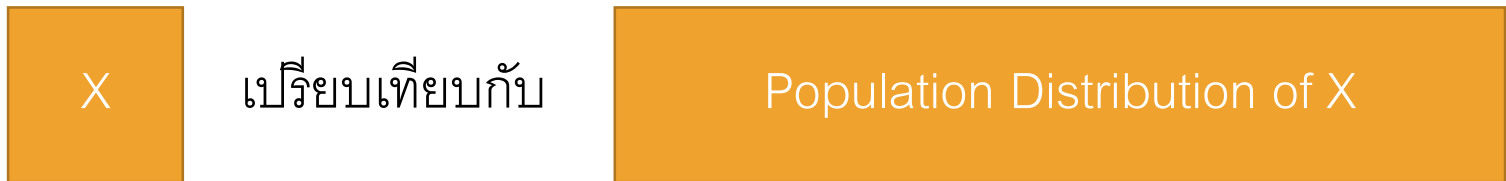
- การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง (Sampling Distribution of Means)
- การทดสอบสมมติฐานสำหรับค่าเฉลี่ย
- การประมาณค่า
- ข้อตกลงเบื้องต้นก่อนการใช้สถิติ
- การกระจายของสถิติอื่นของกลุ่มตัวอย่าง
- สรุป

แนะนำ

- ในครั้งที่ผ่านมา ได้กล่าวถึงการทดสอบสมมติฐาน ในกรณีที่เรเก็บข้อมูลจากคนเดียว
 - เช่น คอรัสฝึกอบรมทำให้มี IQ มากขึ้นหรือไม่ เราได้เก็บข้อมูลจากคนเดียว
 - สมมติได้คะแนน 130
- การใช้ข้อมูลจากหลายคน ทำให้เกิดความแม่นยำมากกว่าการใช้ข้อมูลจากคนเดียว
 - เช่น แทนที่จะเก็บข้อมูลคอรัสฝึกอบรมจากคนเดียว เราก็เก็บข้อมูลจากหลายคนพร้อมกัน แล้วดูว่าค่าเฉลี่ยของ IQ ที่ได้เป็นอย่างไร
 - สมมติ รวม 30 คน ได้คะแนนเฉลี่ยเท่ากับ 130

แนะนำ

- ค่าคะแนนของคนเดียว (X) เปรียบเทียบกับการกระจายประชากร



- ค่าเฉลี่ยจากหลายคน เปรียบเทียบการกระจายของค่าเฉลี่ย



การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

- สมมติว่าประชากรมีทั้งหมด N คน ถ้าสุ่มกลุ่มตัวอย่างออกมา ทีละ n คน จะได้ค่าเฉลี่ยอย่างไรบ้าง
- การสุ่มออกมาแต่ละครั้ง ไม่จำเป็นจะต้องมีค่าเฉลี่ยเท่ากันระหว่างกลุ่มตัวอย่าง และไม่จำเป็นจะต้องมีค่าเฉลี่ยเท่ากับค่าเฉลี่ยของประชากรเสมอไป

หมายเหตุ ในที่นี้ n แทนจำนวนคนในกลุ่มตัวอย่าง
และ N แทนจำนวนคนในประชากร

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง



$\mu = 165 \text{ cm}$
 $\sigma = 10 \text{ cm}$
 $N = 67,000,000$

$n = 100$

กลุ่มตัวอย่างที่ 1

กลุ่มตัวอย่างที่ 2

กลุ่มตัวอย่างที่ 3

...

$M = 162 \text{ cm}$

$M = 170 \text{ cm}$

$M = 165 \text{ cm}$

ความคาดเคลื่อนจากการสุ่มกลุ่มตัวอย่าง (Sampling Error)

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

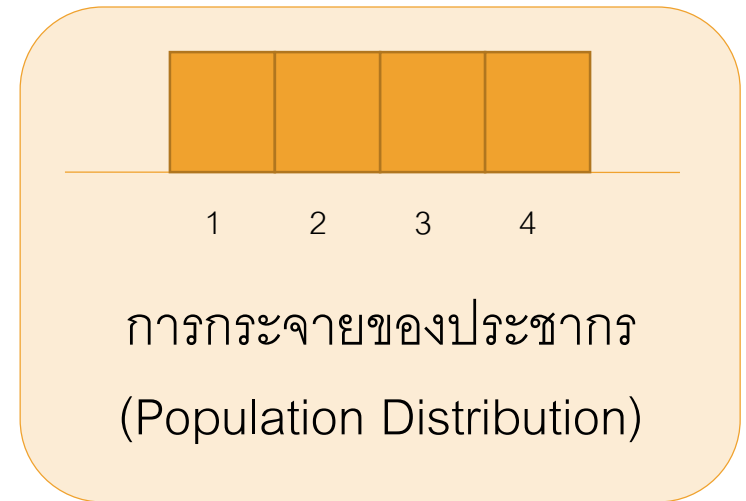
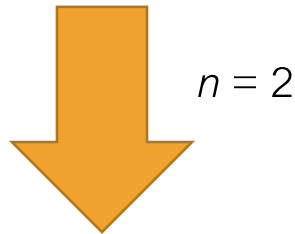
- สังเกตได้ว่า แม้ว่ากลุ่มตัวอย่างจะสุ่มมาจากประชากรเดียวกัน ก็ไม่ได้หมายความว่า ต้องได้ค่าสถิติ (Statistic) เหมือนกัน
- ความคลาดเคลื่อนจากการสุ่มกลุ่มตัวอย่าง (Sampling Error) คือ การเปลี่ยนแปลงของสถิติ (ในที่นี้คือค่าเฉลี่ย) ระหว่างกลุ่มตัวอย่างที่เกิดจากการสุ่ม
- เนื่องจากสถิติเปลี่ยนแปลงไปตามกลุ่มตัวอย่าง ถ้าสามารถสุ่มกลุ่มตัวอย่างแบบเป็นไปได้อย่างทั้งหมด (All possible samples) แล้วนำค่าสถิติของกลุ่มตัวอย่างเหล่านั้นมาสร้างเป็นกราฟการกระจาย จะสามารถนำไปใช้ในการทดสอบสมมติฐานได้

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

- เช่น ประชากร คือ เลข 1, 2, 3, 4

$$\mu = 2.5; \sigma = 1.118; N = 4$$

สุ่มกลุ่มตัวอย่างแบบทดแทน
(Sampling with replacement)



X	M
1, 1	1
1, 2	1.5
1, 3	2
1, 4	2.5

X	M
2, 1	1.5
2, 2	2
2, 3	2.5
2, 4	3

X	M
3, 1	2
3, 2	2.5
3, 3	3
3, 4	3.5

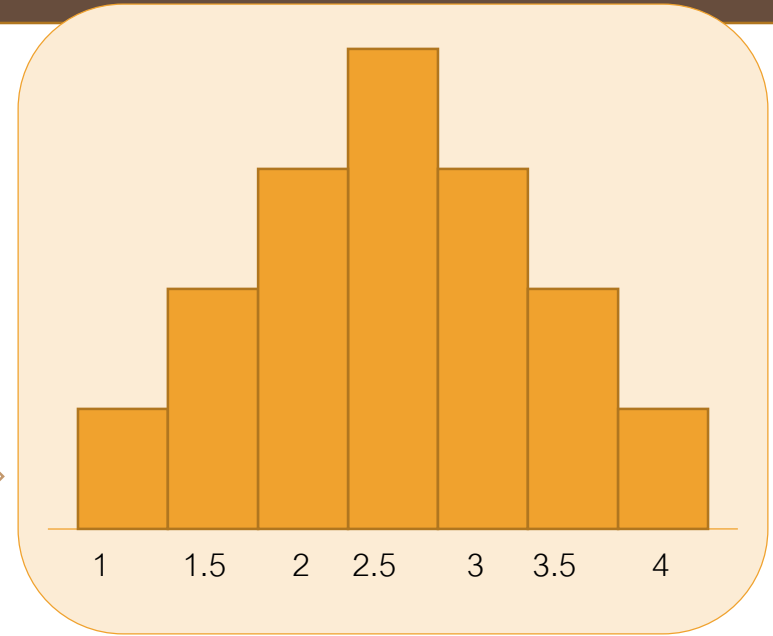
X	M
4, 1	2.5
4, 2	3
4, 3	3.5
4, 4	4

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

- เช่น ประชากร คือ เลข 1, 2, 3, 4

$$\mu = 2.5; \sigma = 1.118; N = 4$$

การเอาค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นไปได้
(Sample Mean) ทั้งหมดมาสร้างกราฟ



X	M
1, 1	1
1, 2	1.5
1, 3	2
1, 4	2.5

X	M
2, 1	1.5
2, 2	2
2, 3	2.5
2, 4	3

X	M
3, 1	2
3, 2	2.5
3, 3	3
3, 4	3.5

X	M
4, 1	2.5
4, 2	3
4, 3	3.5
4, 4	4

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

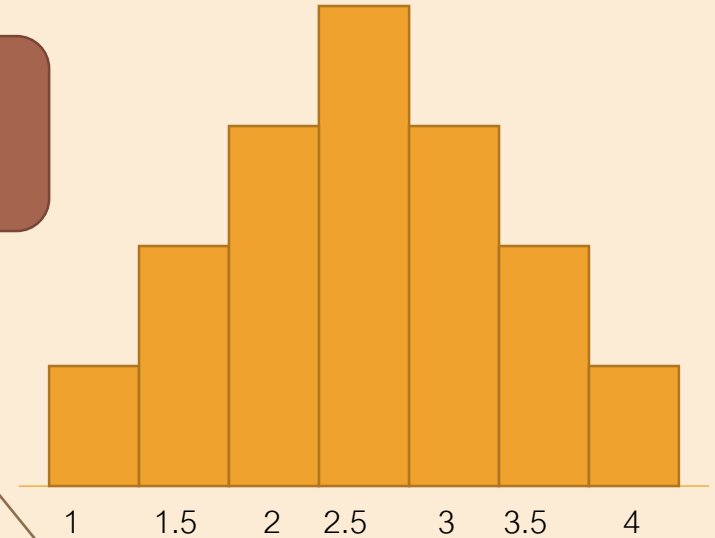
เป็นไปได้ 16
รูปแบบ

$n = 2$

ส่วนเบี่ยงเบนมาตรฐานของค่าเฉลี่ย
ของทั้ง 16 กลุ่มตัวอย่างที่ได้

ค่าเฉลี่ยของค่าเฉลี่ยของทั้ง
16 กลุ่มตัวอย่างที่ได้

$$\mu = 2.5; \sigma = 1.118$$



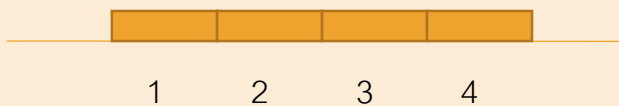
การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง
(Sampling Distribution of Means)

$$\mu_M = 2.5; \sigma_M = 0.791$$

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

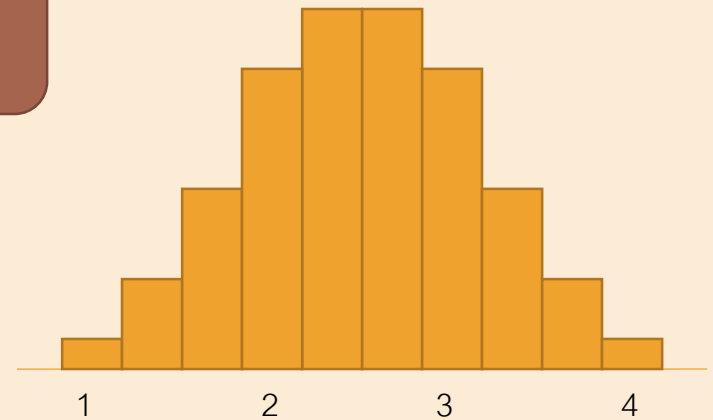
เป็นไปได้ 64
รูปแบบ

$n = 3$



การกระจายของประชากร
(Population Distribution)

$$\mu = 2.5; \sigma = 1.118$$



การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง
(Sampling Distribution of Means)

$$\mu_M = 2.5; \sigma_M = 0.645$$

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

- การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง (Sampling Distribution of Means) คือ ลักษณะการกระจายของค่าเฉลี่ยเมื่อสุ่มกลุ่มตัวอย่างขนาด n ที่เป็นไปได้ทั้งหมดออกมา

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

- ลักษณะของการกระจายค่าเฉลี่ยของกลุ่มตัวอย่างจะมี 3 ลักษณะ คือ
 1. ค่าเฉลี่ยของค่าเฉลี่ยจากกลุ่มตัวอย่างทั้งหมด (Mean of Sample Means; μ_M) จะมีค่าเท่ากับค่าเฉลี่ยของประชากร (μ)

$$\mu_M = \mu$$

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

- ลักษณะของการกระจายค่าเฉลี่ยของกลุ่มตัวอย่างจะมี 3 ลักษณะ คือ
 1. ค่าเฉลี่ยของค่าเฉลี่ยของกลุ่มตัวอย่างจะเท่ากับค่าเฉลี่ยของประชากร
 2. ความแปรปรวนของค่าเฉลี่ยจากกลุ่มตัวอย่างทั้งหมด (Variance of Sample Means; σ_M^2) จะมีค่าเท่ากับความแปรปรวนของประชากร (σ^2) หารด้วยจำนวนกลุ่มตัวอย่าง

$$\sigma_M^2 = \sigma^2 / n$$

หรือ เปลี่ยนสูตรเป็นส่วนเบี่ยงเบนมาตรฐานของค่าเฉลี่ยจากกลุ่มตัวอย่าง ซึ่งจะเรียกว่า ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย (Standard Error of Mean)

$$\sigma_M = \sigma / \sqrt{n}$$

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

- ลักษณะของการกระจายค่าเฉลี่ยของกลุ่มตัวอย่างจะมี 3 ลักษณะ คือ
 3. การกระจายของค่าเฉลี่ย จะเข้าใกล้โค้งปกติ โดยทั่วไปจะสามารถเหมารวมว่าเป็นโค้งปกติ ถ้าเข้าเงื่อนไขอันใดอันหนึ่งใน 2 เงื่อนไข
 - จำนวนกลุ่มตัวอย่างมากกว่า 30
 - การกระจายของประชากรเป็นโค้งปกติอยู่แล้ว
- ลักษณะทั้งสามนี้ จะเป็นส่วนหนึ่งของทฤษฎีทางสถิติที่เรียกว่า **ทฤษฎี**
แนวโน้มเข้าสู่ศูนย์กลาง (Central Limit Theorem)

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

- จาก Central Limit Theorem ทำให้สามารถหาโอกาสในการสุ่มเจอค่าเฉลี่ยช่วงต่างๆ ได้
- เช่น ถ้าสุ่มประชากรไทยจำนวน 25 คน ที่มีค่าเฉลี่ยของความสูง 165 cm และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 cm มีการกระจายเป็นโค้งปกติ โอกาสที่จะสุ่มเจอคนที่มีความสูงต่ำกว่า 160 cm มีเท่าไร

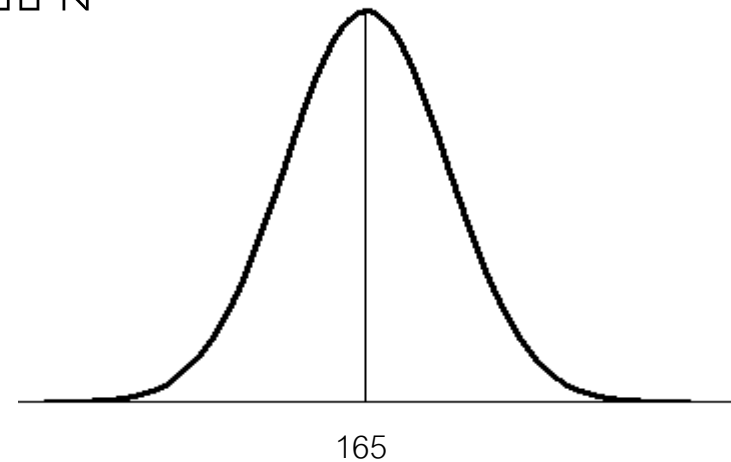
การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

- อันดับแรก ให้วาดภาพการกระจาย เมื่อสุ่มกลุ่มตัวอย่างคนไทย จำนวน 25 คน หรือสร้างการกระจายค่าเฉลี่ยของกลุ่มตัวอย่างนั่นเอง
 - ค่าเฉลี่ยของการกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

$$\mu_M = \mu = 165$$

- ส่วนเบี่ยงเบนมาตรฐาน

$$\sigma_M = \sigma / \sqrt{n} = 10 / \sqrt{25} = 2$$



- การกระจายประชากรไทยเป็นโค้งปกติ ดังนั้นการกระจายค่าเฉลี่ยก็เป็นโค้งปกติเช่นกัน

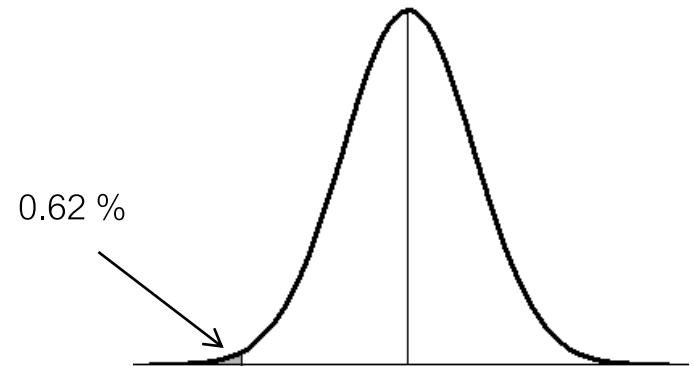
การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

2. ต่อไป หาค่ามาตรฐานของการสุ่มได้ค่าเฉลี่ย 160 cm

$$z = (160-165)/2 = -2.5$$

3. หาโอกาสเจอ z น้อยกว่า -2.5 เท่ากับ 0.00621 (0.62 %)

สรุป โอกาสในการเจอกลุ่มตัวอย่างจาก
คนไทย 25 คนที่มีค่าเฉลี่ยต่ำกว่า 160
cm มี 0.62 %



การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

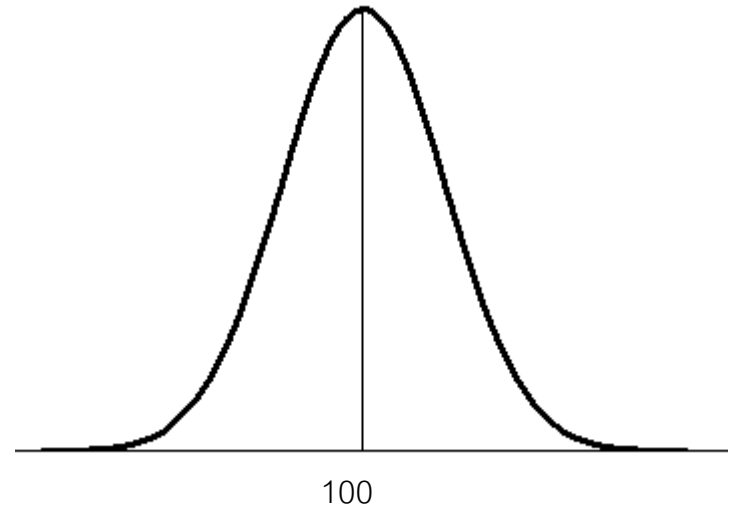
- ตัวอย่างที่ 2 ถ้าห้องเรียนหนึ่ง มีจำนวน 36 คนมี IQ เฉลี่ย 105 เมื่อเทียบกับประชากรปกติแล้ว ($\mu = 100$; $\sigma = 15$) โอกาสที่สุ่มเจอ IQ ห่างจากค่าเฉลี่ยของประชากรปกติมากขนาดนี้มีเท่าไร (ห่างในที่นี้ หมายถึง ห่างทั้งด้านบวกและลบ)

การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

1. สร้างการกระจายค่าเฉลี่ย IQ ของกลุ่มตัวอย่างที่สุ่มจากประชากรปกติ จะเป็นการกระจายโค้งปกติ ที่มี

$$\mu_M = \mu = 100$$

$$\sigma_M = \sigma / \sqrt{n} = 15 / \sqrt{36} = 2.5$$



การกระจายค่าเฉลี่ยของกลุ่มตัวอย่าง

2. หาค่ามาตรฐานของค่าเฉลี่ยจากห้องเรียนนี้เปรียบเทียบกับประชากรปกติเท่ากับ

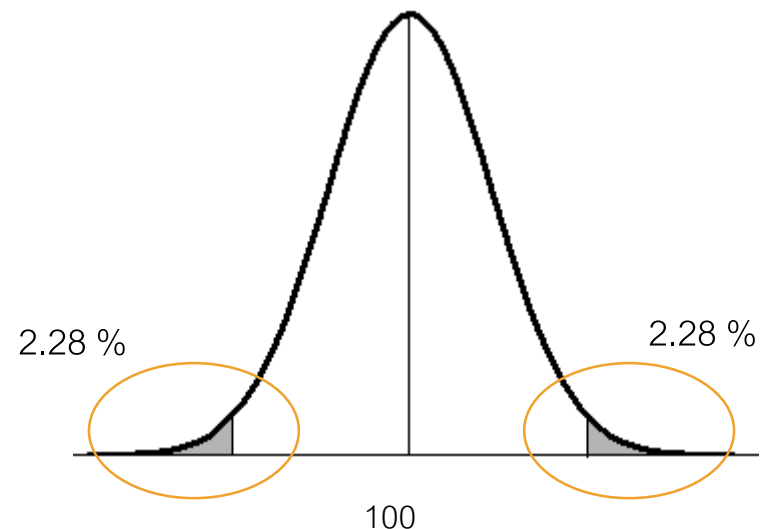
$$z = (105-100)/2.5 = 2$$

3. หาโอกาสในการเจอคนที่มีค่า z มากกว่า 2 รวมกับคนที่มีค่า z น้อยกว่า 2 (ดูทั้งสองข้าง)

○ ได้โอกาสข้างเดียว = .0228 (2.28 %)

○ รวม 2 ข้าง = .0456 (4.56 %)

โอกาสที่เจอห้องเรียนขนาด 36 คนที่มี IQ เฉลี่ยห่างจากค่าเฉลี่ยของประชากรมากกว่า 105 เท่ากับ 4.56 %

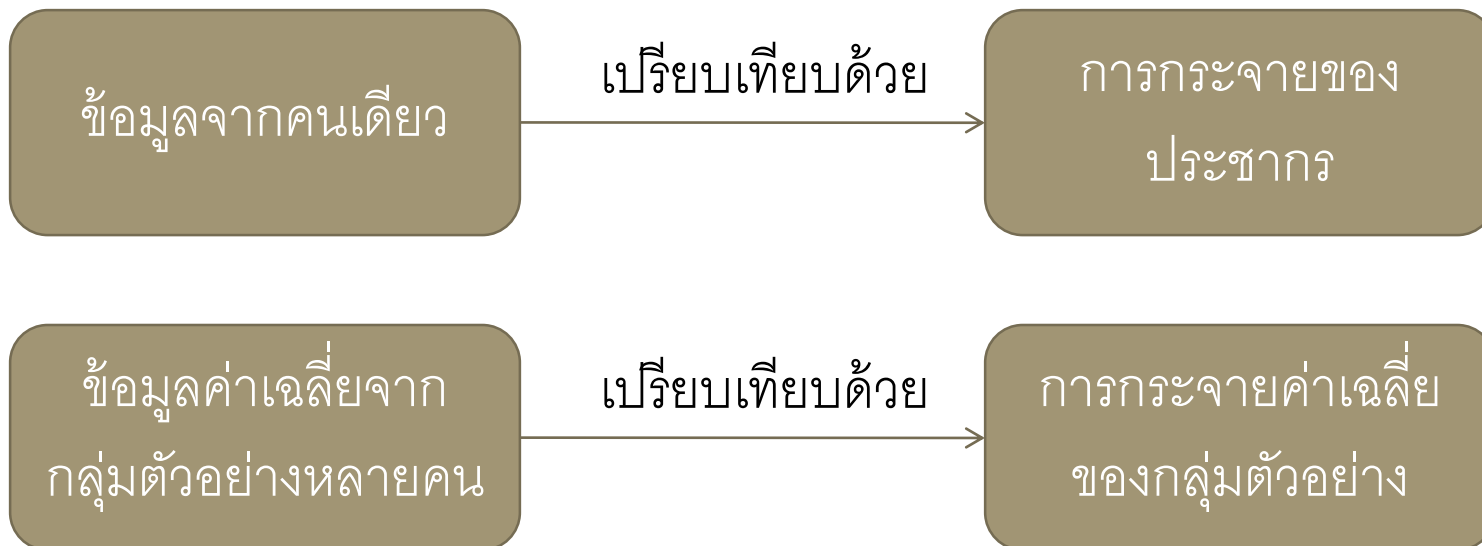


การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- กระบวนการทดสอบสมมติฐานแบบใช้ค่า p แบ่งได้เป็น 5 ขั้นตอน
 1. เขียนสมมติฐานภายในใจ (สมมติฐานวิจัย) ใหม่ ให้อยู่ในรูปของสมมติฐานทางสถิติ
 2. กำหนดลักษณะของการกระจายที่ใช้ในการเปรียบเทียบ
 3. หาโอกาสที่จะเกิดเหตุการณ์ที่ท่านสนใจ
 4. ตัดสินว่าโอกาสที่เกิดขึ้นนั้น น้อยเกินไป (จนไม่น่าจะเป็นไปได้) หรือยังเป็นไปได้
 5. สรุปว่าจะสนับสนุนสมมติฐานภายในใจ (สมมติฐานวิจัย) หรือไม่

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- สิ่งที่จะแตกต่างระหว่างการทดสอบสมมติฐานสำหรับข้อมูลจากคนเดียว และ ข้อมูลจากค่าเฉลี่ยของหลายคนรวมกันคือ
- ขั้นตอนที่ 2 การกำหนดลักษณะการกระจายของการเปรียบเทียบ



การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- ขั้นที่หนึ่ง กำหนดสมมติฐานภายในใจ (ทดสอบทางเดียว)

- เช่น คนหนึ่งสูง 180 cm

คน 4 คนชาติเดียวกันค่าเฉลี่ย
180 cm

คนนี้น่าจะเป็นคนไทย

4 คนนี้น่าจะเป็นคนไทย

ประชากรไทย: $\mu = 165$ cm, $\sigma = 10$ cm

$$H_0: \mu \leq 165$$

$$H_1: \mu > 165$$

$$H_0: \mu \leq 165$$

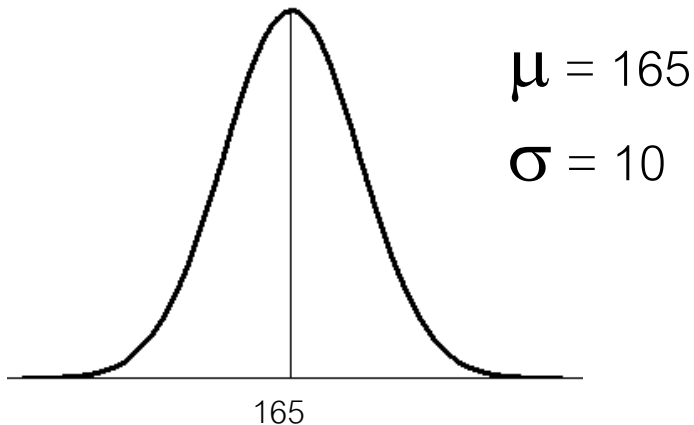
$$H_1: \mu > 165$$

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

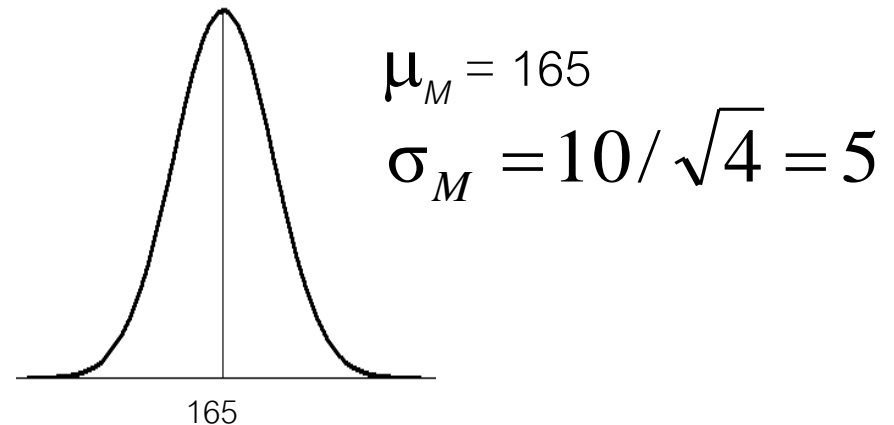
- ขั้นที่สอง ลักษณะการกระจายที่ใช้ในการเปรียบเทียบ

ประชากรไทย: $\mu = 165$ cm, $\sigma = 10$ cm

การกระจายประชากรไทย

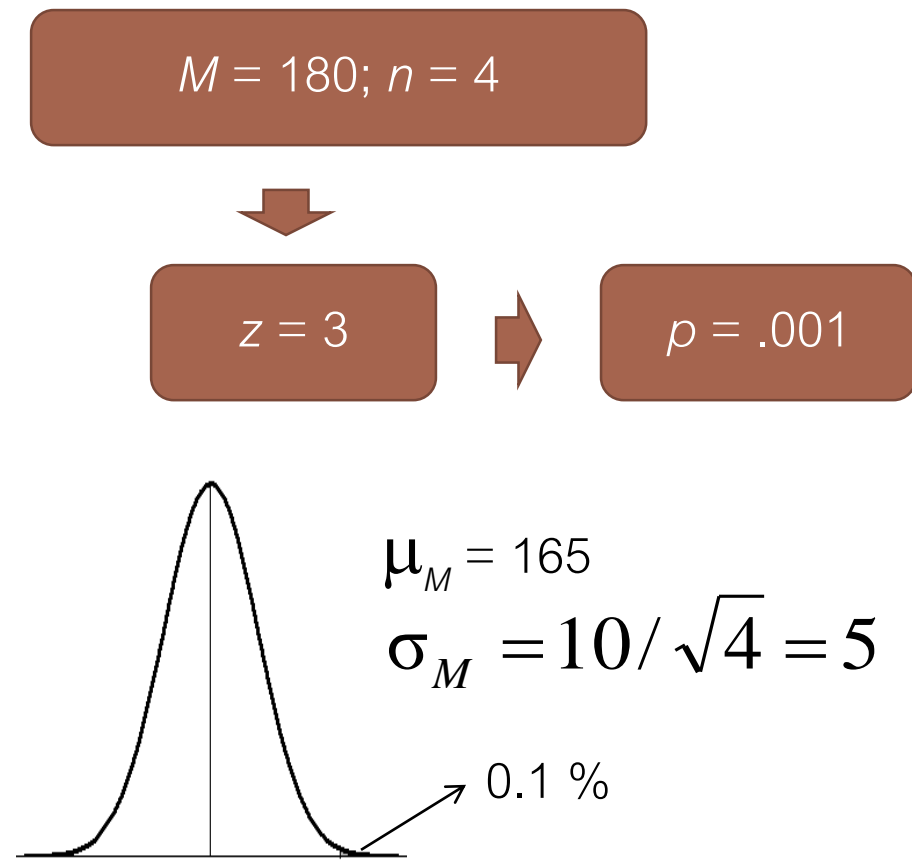
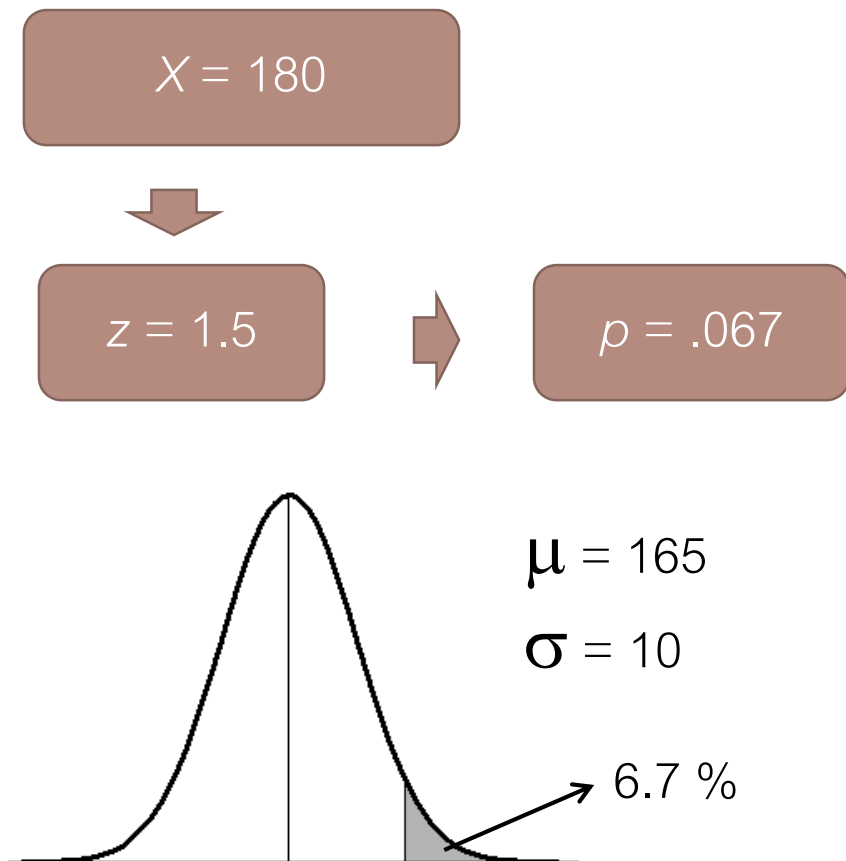


การกระจายของค่าเฉลี่ยจากการสุ่มกลุ่มตัวอย่างขนาด 4 คนจากประชากรไทย



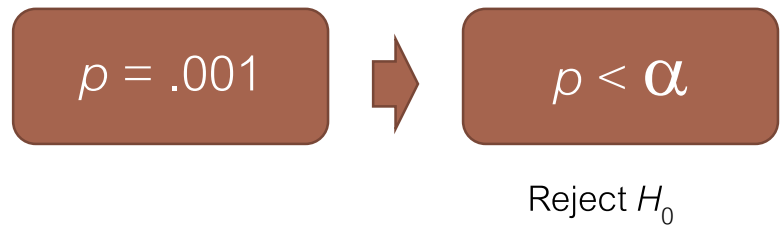
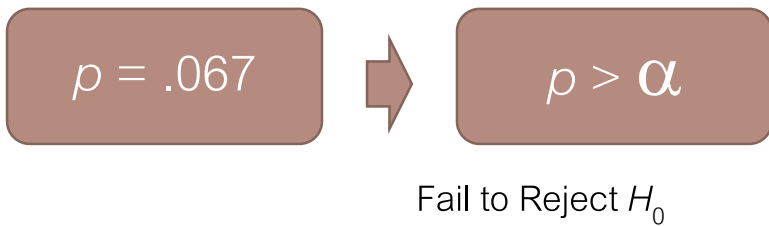
การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- ขั้นที่สาม หาโอกาสเกิดเหตุการณ์ที่สนใจ

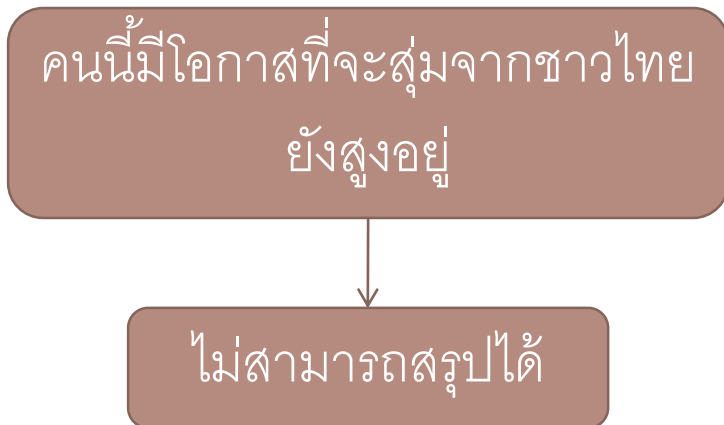


การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- **ขั้นที่สี่** ดูว่าโอกาสที่เกิดขึ้นเกินกว่าระดับนัยสำคัญหรือไม่ ($\alpha = .05$)



- **ขั้นที่ห้า** สรุปว่าสนับสนุนสมมติฐานวิจัยหรือไม่



การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- โอกาสที่จะเจอคนไทยคนเดียว ความสูง 180 cm ยังพอมี
- แต่โอกาสที่จะสุ่มคนไทย 4 คน แล้วได้ค่าเฉลี่ยความสูง 180 cm มีน้อยมากๆ

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- กระบวนการทดสอบสมมติฐานแบบใช้ Critical Region แบ่งได้เป็น 5 ขั้นตอน
 1. เขียนสมมติฐานภายในใจ (สมมติฐานวิจัย) ใหม่ ให้อยู่ในรูปของสมมติฐานทางสถิติ
 2. กำหนดลักษณะของการกระจายที่ใช้ในการเปรียบเทียบ
 3. หา Critical Region จากระดับนัยสำคัญที่ได้ตั้งเอาไว้
 4. ตัดสินว่าเหตุการณ์อยู่ใน Critical Region หรือไม่
 5. สรุปว่าจะสนับสนุนสมมติฐานภายในใจ (สมมติฐานวิจัย) หรือไม่

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- ขั้นที่หนึ่ง กำหนดสมมติฐานภายในใจ (ทดสอบสองทาง)

- เช่น คนผ่านการฝึกอบรม IQ ได้ IQ 105 เต็ม

คนนี้น่าจะมี IQ ต่างจากปกติ

- เช่น คนผ่านการฝึกอบรม IQ 25 ได้ IQ เฉลี่ย 105 เต็ม

25 คนนี้น่าจะมี IQ ต่างจากปกติ

ประชากรปกติ : $\mu = 100, \sigma = 15$

$$H_0: \mu = 100$$

$$H_1: \mu \neq 100$$

$$H_0: \mu = 100$$

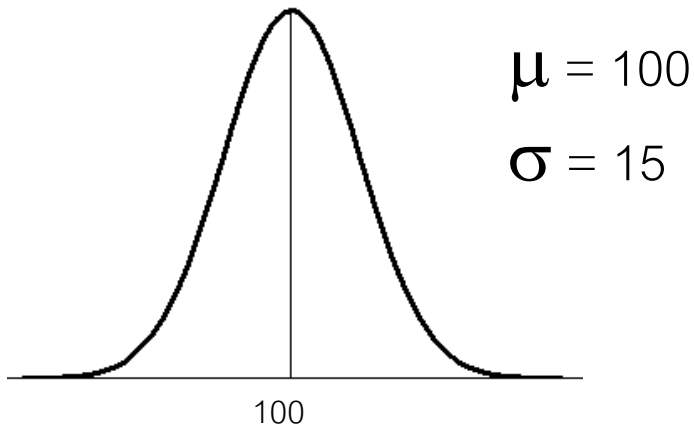
$$H_1: \mu \neq 100$$

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

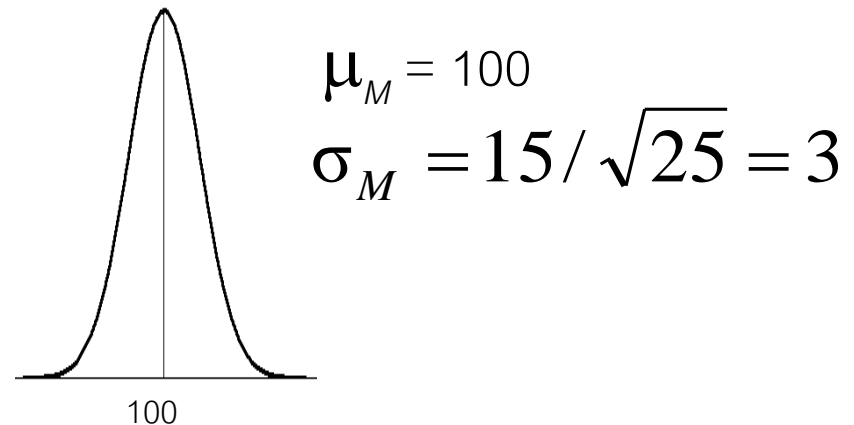
- ขั้นที่สอง ลักษณะการกระจายที่ใช้ในการเปรียบเทียบ

ประชากรปกติ : $\mu = 100$, $\sigma = 15$

การกระจายประชากรปกติ



การกระจายของค่าเฉลี่ยจากการสุ่มกลุ่มตัวอย่างขนาด 25 คนจากประชากรปกติ



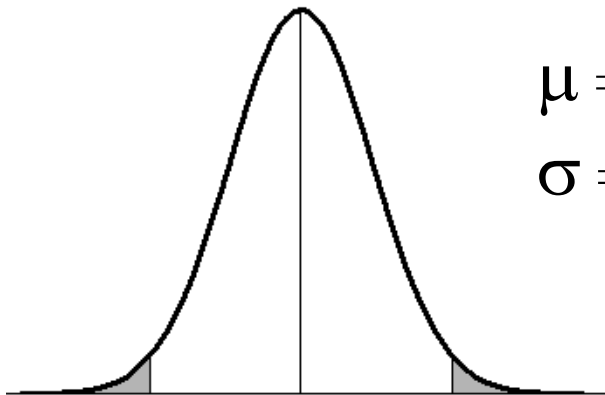
การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- ขั้นที่สาม หาค่าบริเวณวิกฤตจากระดับนัยสำคัญที่ตั้งไว้ ($\alpha = .05$)

$$\alpha = .05$$



$$\text{ค่าวิกฤต} = -1.96 \text{ และ } 1.96$$



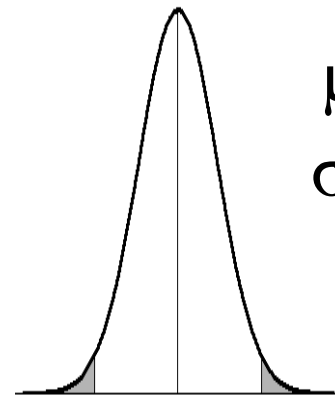
$$\mu = 100$$

$$\sigma = 15$$

$$\alpha = .05$$



$$\text{ค่าวิกฤต} = -1.96 \text{ และ } 1.96$$

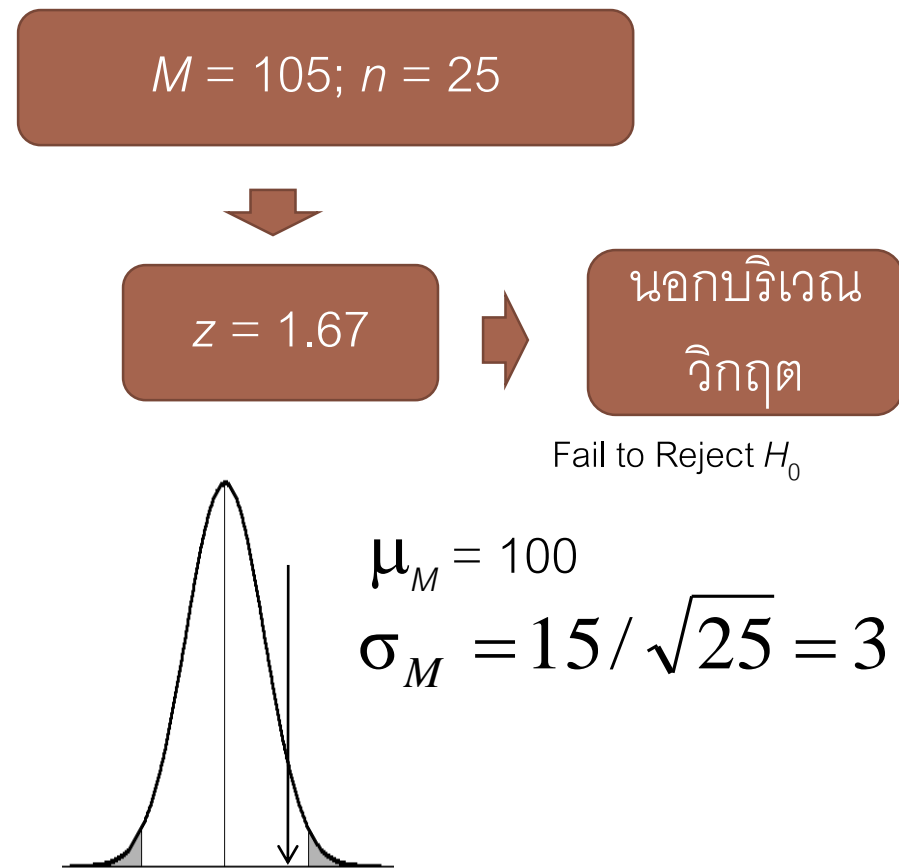
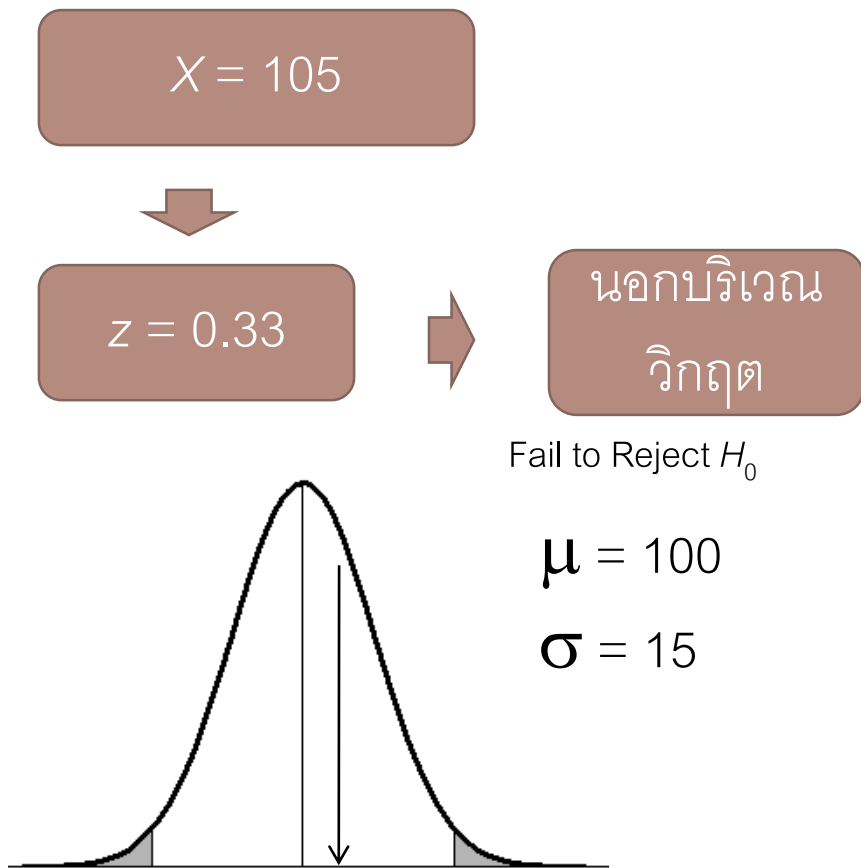


$$\mu_M = 100$$

$$\sigma_M = 15 / \sqrt{25} = 3$$

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- ขั้นที่สี่ ตัดสินว่าค่าอยู่ในบริเวณวิกฤตหรือไม่



การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- ขั้นที่ห้า สรุปว่าสนับสนุนสมมติฐานวิจัยหรือไม่



การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- อีกหนึ่งตัวอย่าง

คนที่เรียนจิตวิทยาจะดูน่ารักขึ้น
นำนิสิต 120 คนมาวัดหลังจาก
เรียน 1 ปี ได้ค่าเฉลี่ย 54 แต้ม



สมมติฐาน คือ นิสิตทั้งหมดมี
ความน่ารักมากกว่าคนปกติ

สมมติว่า ประชากรปกติมีความน่ารัก

$$\mu = 50, \sigma = 10$$

คำถาม: โอกาสที่จะสุ่มกลุ่มตัวอย่าง 120 คน
เจอค่าเฉลี่ยสูงขนาดนี้มีเท่าไร

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

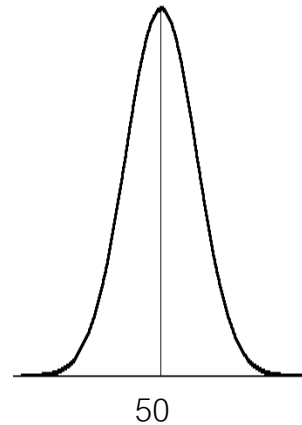
- อธิบายรายละเอียดเพิ่มเติม
คนที่เรียนจิตวิทยาจะดูน่ารักขึ้น
นำนิสิต 120 คนมาวัดหลังจาก
เรียน 1 ปี ได้ค่าเฉลี่ย 54 แต้ม



วาดภาพการสุ่มคนปกติทีละ 120 คน
แล้วดูว่าค่าเฉลี่ยมีการกระจายอย่างไร



การกระจายของค่าเฉลี่ยความน่ารัก



$$\mu_M = 50$$

$$\sigma_M = 10 / \sqrt{120} = 0.913$$

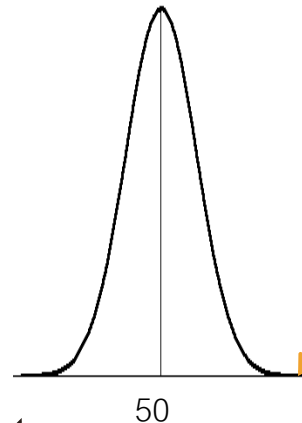
การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- อธิบายรายละเอียดเพิ่มเติม
คนที่เรียนจิตวิทยาจะดูน่ารักขึ้น
นำนิสิต 120 คนมาวัดหลังจาก
เรียน 1 ปี ได้ค่าเฉลี่ย 54 แต้ม

ถ้าสุ่มจากประชากรปกติ โอกาส
ที่จะสุ่มคน 120 คน แล้วเจอ
ค่าเฉลี่ยความน่ารักเท่ากับ 54
แต้มขึ้นไป เท่ากับ 0.0006 %

หาโอกาสในการเจอค่าเฉลี่ยสูง
กว่า 54 แต้มจากประชากรปกติ

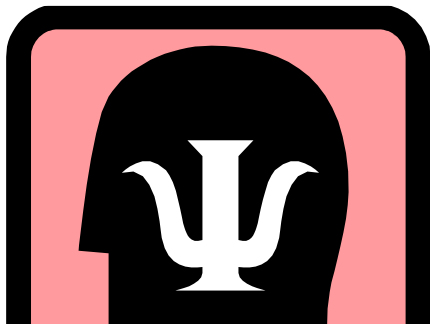
$$z = \frac{54 - 50}{0.913} = 4.382$$



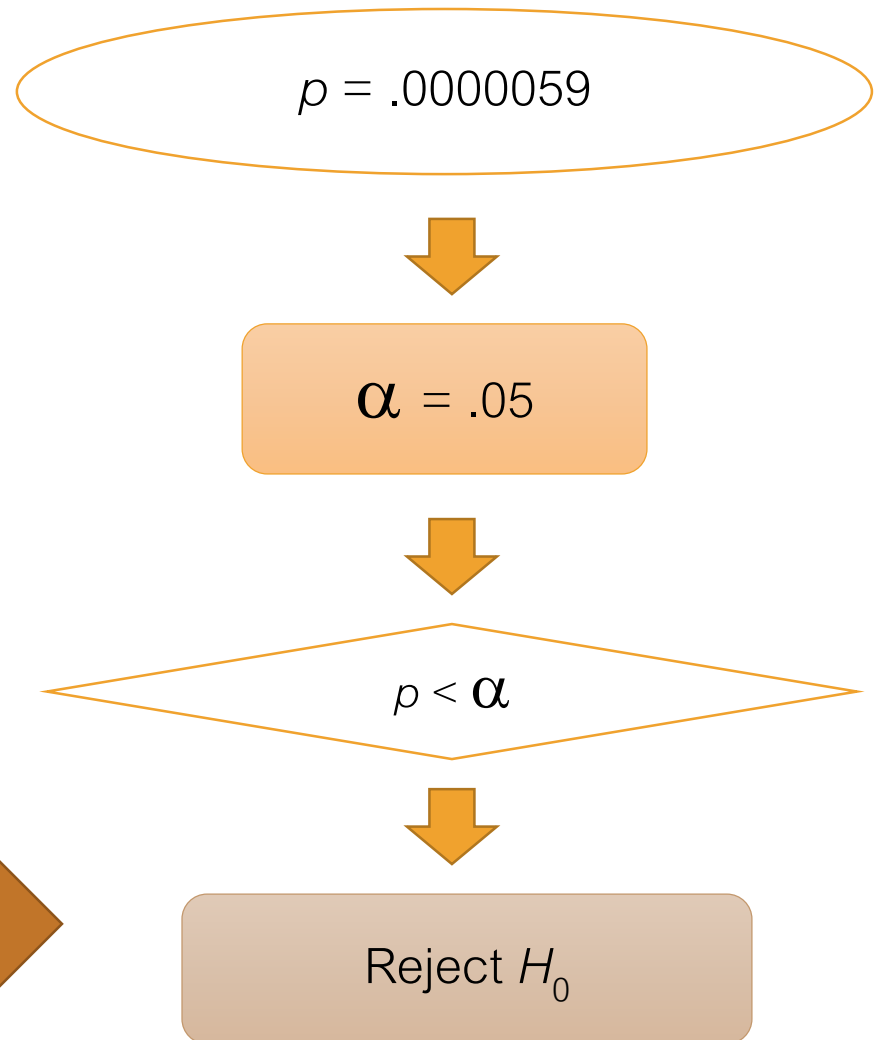
$p = .0000059$
หรือ 0.0006 %

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- อธิบายรายละเอียดเพิ่มเติม
คนที่เรียนจิตวิทยาจะดูน่ารักขึ้น
นำนิสิต 120 คนมาวัดหลังจาก
เรียน 1 ปี ได้ค่าเฉลี่ย 54 แต้ม



คนที่เรียนจิตวิทยาไป 1 ปีจะน่ารัก
กว่าคนทั่วไปจริง



การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- อธิบายรายละเอียดเพิ่มเติม

คนที่เรียนจิตวิทยาจะดูน่ารักขึ้น
นำนิสิต 120 คนมาวัดหลังจาก
เรียน 1 ปี ได้ค่าเฉลี่ย 54 แต้ม



คนที่เรียนจิตวิทยาอาจจะหน้าตาดีกว่าคนทั่วไปตั้งแต่แรกก็ได้
คนที่เรียนจิตวิทยาอาจจะอยู่ไกลผู้คน เลยไม่เจอมลภาวะ
คนอาจเหมารวมว่า คณะจิตวิทยาหน้าตาดี (ทั้งที่จริงๆ ก็พอๆ กับคณะอื่น)

คำถาม

สามารถสรุปได้หรือไม่ว่า
การเรียนจิตวิทยาทำให้คนน่ารักขึ้น

ไม่ได้

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

- กระบวนการทดสอบสมมติฐานสำหรับค่าเฉลี่ยที่ผ่านมาทั้งหมด จะเรียกอีกชื่อหนึ่งว่า การทดสอบ z (z test)

การทดสอบสมมติฐานสำหรับค่าเฉลี่ย

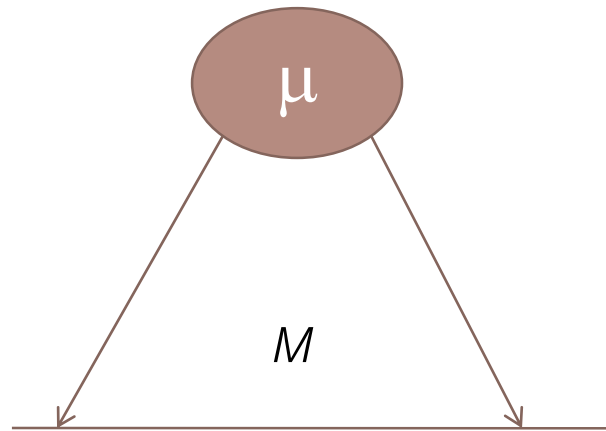
- การเขียนรายงาน

- สมมติฐานงานวิจัย คือ นิสิตจิตวิทยาที่เรียนจิตวิทยาไปแล้วอย่างน้อย 1 ปีจะมีความสุขมากขึ้น จากการทดสอบค่า z (One sample z -test) พบว่า นิสิตที่ผ่านการเรียนจิตวิทยา 1 ปีจำนวน 120 คน ($M = 54$) มีค่าเฉลี่ยของความสุขแตกต่างจากประชากรไทยทั่วไป ($\mu = 50, \sigma = 10$) อย่างมีนัยสำคัญทางสถิติ ($z = 4.38, p < .001$ [ทางเดียว])

การประมาณค่า

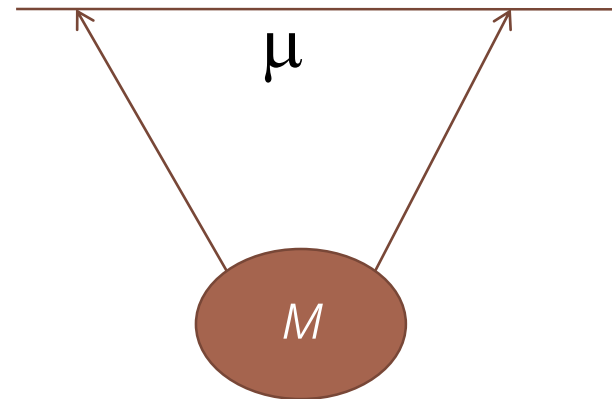
- การประมาณค่า เป็นการนำค่าสถิติไปทำนายค่าพารามิเตอร์
- เหมือนเป็นส่วนกลับ ของการสร้างการกระจายของกลุ่มตัวอย่าง

การกระจายของกลุ่มตัวอย่าง



สุ่มแล้วจะได้ค่าสถิติอะไรบ้าง

การประมาณค่า



Parameter อะไรที่น่าจะสุ่มได้สถิติตัวนี้

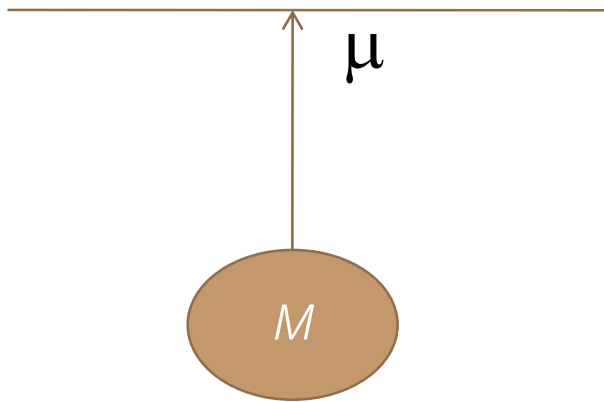
การประมาณค่า

- การประมาณค่า สามารถทำได้ 2 รูปแบบด้วยกัน
 - การประมาณค่าแบบจุด (Point Estimation) เป็นการประมาณจากสถิติ ไปหา Parameter ค่าเดียว
 - การประมาณค่าแบบช่วง (Interval Estimation) เป็นการประมาณจากสถิติ ไปหา ช่วงของ Parameter ที่น่าจะสุ่มสถิติตัวนี้ออกมาได้

การประมาณค่า

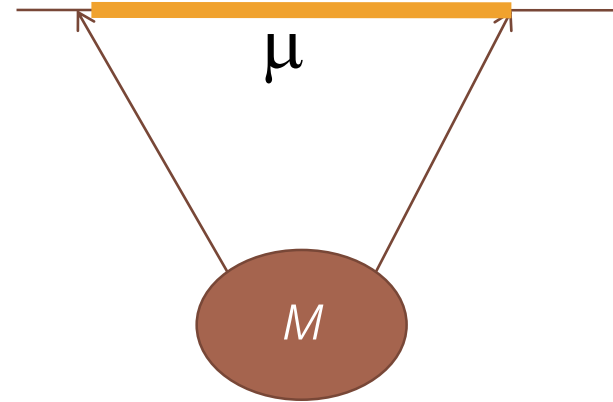
- การประมาณค่า สามารถทำได้ 2 รูปแบบด้วยกัน

การประมาณค่าแบบจุด



Parameter อะไรที่มีโอกาสสุ่ม
ได้สถิติตัวนี้มากที่สุด

การประมาณค่าแบบช่วง



ช่วง Parameter อะไร
ที่น่าจะสุ่มได้สถิติตัวนี้สูง

การประมาณค่าแบบจุด

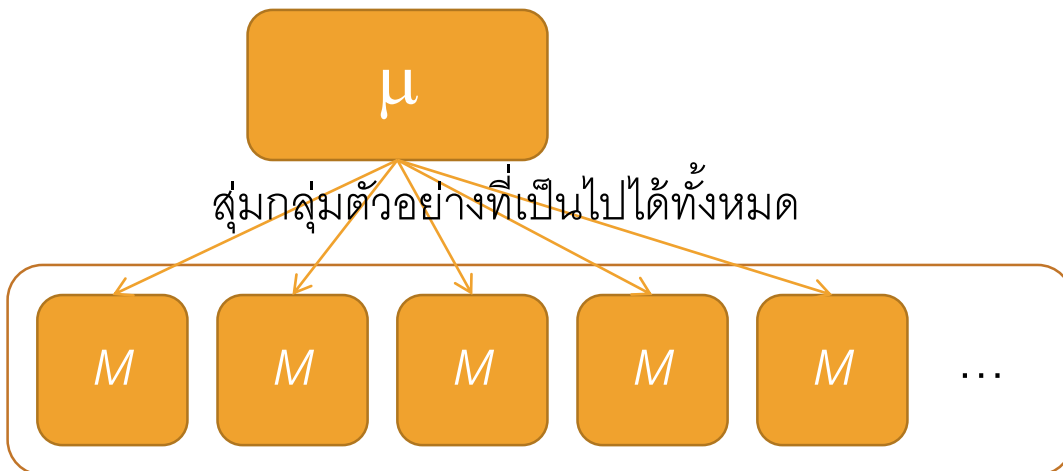
- โดยปกติแล้ว Descriptive Statistics ที่ใช้ในการอธิบายกลุ่มตัวอย่าง จะใช้ในการทำนายค่า Parameter เช่นกัน



- รู้ได้อย่างไร ว่าสถิติตัวนั้นเป็นตัวที่ทำนายค่า Parameter ได้ดี

การประมาณค่าแบบจุด

- ข้อที่ 1 สถิติตัวนั้นจะต้องทำนายค่าพารามิเตอร์ได้แม่นยำ (Precision)
- อย่างไรก็ตาม ค่าสถิติไม่สามารถทำนายค่าพารามิเตอร์ได้ตรงอยู่แล้ว เพราะเกิดความผิดพลาดจากการสุ่ม (Sampling Error)
- แต่ในระยะยาว สถิติตัวนั้น ต้องทำนายค่าพารามิเตอร์ได้ตรง



ค่าเฉลี่ยของสถิติจากกลุ่มตัวอย่าง
ที่เป็นไปได้ทั้งหมด จะต้องเท่ากับ
ค่าพารามิเตอร์

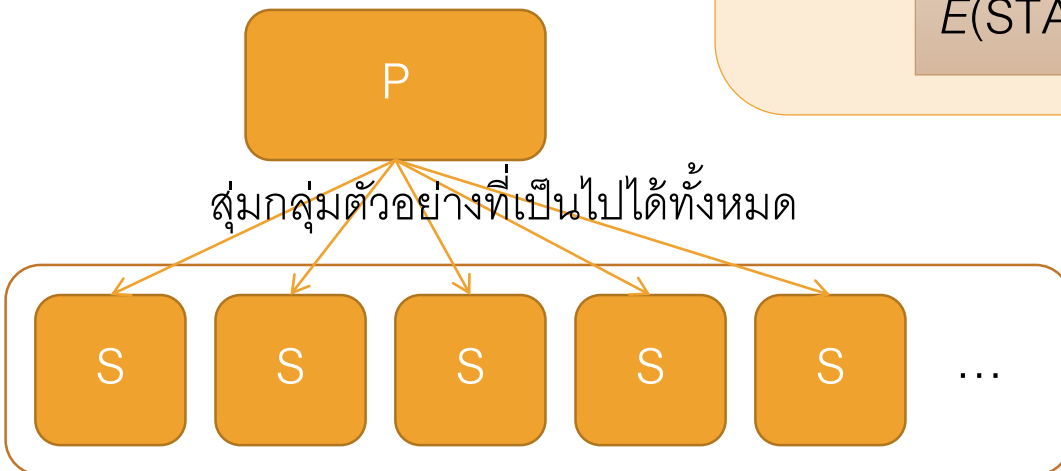
การประมาณค่าแบบจุด

- ข้อที่ 1 สถิติตัวนั้นจะต้องทำนายค่าพารามิเตอร์ได้แม่นยำ (Precision)
- เขียนสัญลักษณ์ได้ว่า

$$\mu_{\text{STATISTICS}} = \text{PARAMETER}$$

หรือ

$$E(\text{STATISTICS}) = \text{PARAMETER}$$



ค่าเฉลี่ยของสถิติจากกลุ่มตัวอย่าง
ที่เป็นไปได้ทั้งหมด จะต้องเท่ากับ
ค่าพารามิเตอร์

การประมาณค่าแบบจุด

- ข้อที่ 1 สถิติตัวนั้นจะต้องทำนายค่าพารามิเตอร์ได้แม่นยำ (Precision)
- เขียนสัญลักษณ์ได้ว่า

$$\mu_{\text{STATISTICS}} = \text{PARAMETER}$$

หรือ

$$E(\text{STATISTICS}) = \text{PARAMETER}$$

- ค่าคาดหวังในระยะยาว (Expected Value) ของสถิติตัวนั้น จะต้องเท่ากับค่า Parameter (ค่าคาดหวัง คือ ค่าเฉลี่ยของสถิติในระยะยาว)

การประมาณค่าแบบจุด

- ยกตัวอย่าง
 - การทำนายค่าความแปรปรวนของประชากร

$SS/n-1$ เป็นตัวทำนาย
ความแปรปรวนของ
ประชากรได้ดีกว่า SS/n

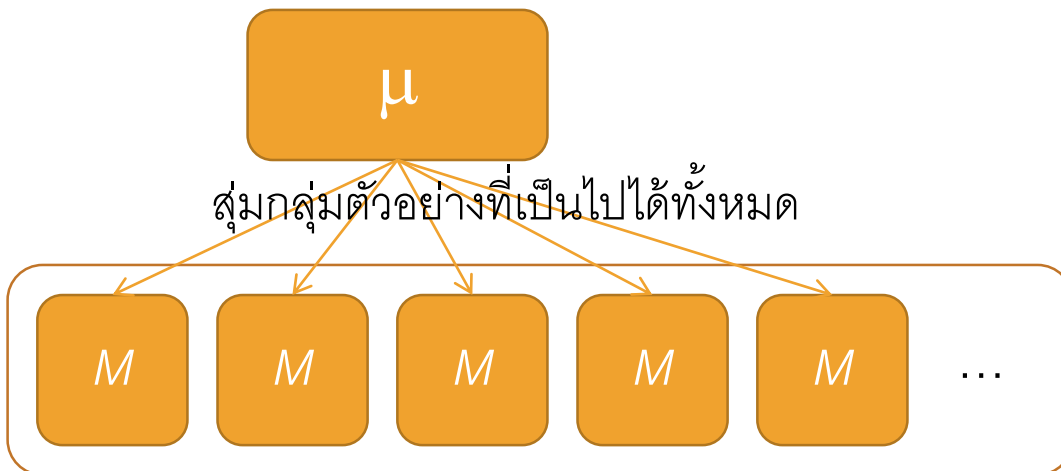
$$E\left(\frac{\sum (X - M)^2}{n}\right) \neq \frac{\sum (X - \mu)^2}{N}$$

$$E\left(\frac{\sum (X - M)^2}{n-1}\right) = \frac{\sum (X - \mu)^2}{N}$$

เหตุผล
ติดตามในบทที่ 6

การประมาณค่าแบบจุด

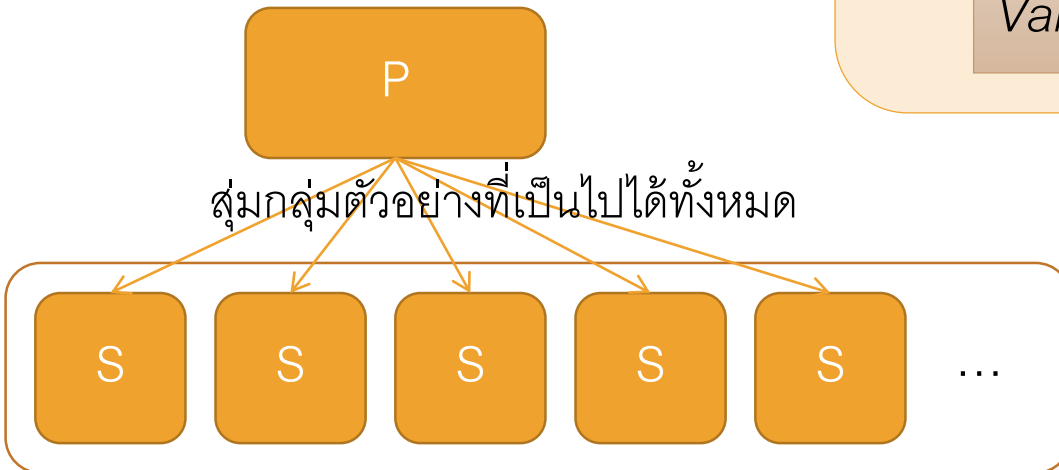
- ข้อที่ 2 สถิติตัวนั้นจะต้องมีการแกว่งไปมาต่ำ (Efficiency)
- ถึงแม้ว่าค่าคาดหวังของสถิติตัวนั้นเท่ากับค่า Parameter
- แต่ค่าสถิติจากกลุ่มตัวอย่างแต่ละกลุ่ม ไม่ควรเบี่ยงเบนไปจากค่า Parameter มากนัก ยิ่งการกระจายต่ำ ยิ่งดี



การกระจายของค่าเฉลี่ยระหว่าง
กลุ่มตัวอย่าง ควรมีค่าต่ำ

การประมาณค่าแบบจุด

- ข้อที่ 2 สถิติตัวนั้นจะต้องมีการแกว่งไปมาต่ำ (Efficiency)
- เขียนสัญลักษณ์ได้ว่า



การกระจายของค่าเฉลี่ยระหว่าง
กลุ่มตัวอย่าง ควรจะมีค่าต่ำ

การประมาณค่าแบบจุด

- ข้อที่ 2 สถิติตัวนั้นจะต้องมีการแกว่งไปมาต่ำ (Efficiency)
- เขียนสัญลักษณ์ได้ว่า

$\sigma_{\text{STATISTICS}}$ มีค่าต่ำ

หรือ

$\text{Var}(\text{STATISTICS})$ มีค่าต่ำ

- Standard Error ของสถิติตัวนั้น ควรต่ำ เมื่อเปรียบเทียบกับค่าสถิติตัวอื่น

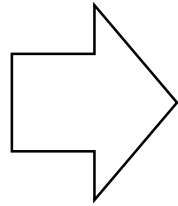
การประมาณค่าแบบจุด

- ยกตัวอย่าง
 - การทำนายค่าเฉลี่ยของประชากร

$$E(\text{Mode}) = \mu$$

$$E(\text{Mdn}) = \mu$$

$$E(M) = \mu$$



$$\text{Var}(M) < \text{Var}(\text{Mdn}) < \text{Var}(\text{Mode})$$

เหตุผล

M ใช้ค่าจากกลุ่มตัวอย่างทุกตัว

M เป็นตัวทำนายค่าเฉลี่ยของประชากร
ได้ดีกว่า Mdn และ Mode

การประมาณค่าแบบช่วง

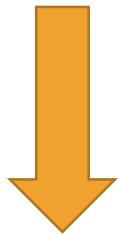
- ตัวประมาณค่าแบบจุดที่ดี จึงต้องทำนายได้ตรง แล้วแกว่งไปแกว่งมาน้อย
- แต่ทว่า การประมาณค่าแบบจุด ส่วนใหญ่ได้แค่ใกล้เคียงค่า Parameter โอกาสที่จะตรงพอดีมีน้อยมาก
- จึงมีแนวคิดว่าจะมีการประมาณค่าเป็นช่วง ว่าช่วงที่ Parameter อยู่ น่าจะเป็นช่วงไหน จึงเป็นการประมาณค่าแบบช่วง

การประมาณค่าแบบช่วง

- คำถาม Parameter ตัวไหน ที่น่าจะสุ่มได้ค่าสถิติเช่นนี้
- สิ่งแรกที่ต้องคำนึงถึง คือ ค่า Parameter จะมีการสุ่มกลุ่มตัวอย่างออกมาได้ค่าใดบ้าง
- สมมติว่า

$$\mu = 140 \text{ cm}; \sigma = 10 \text{ cm}$$

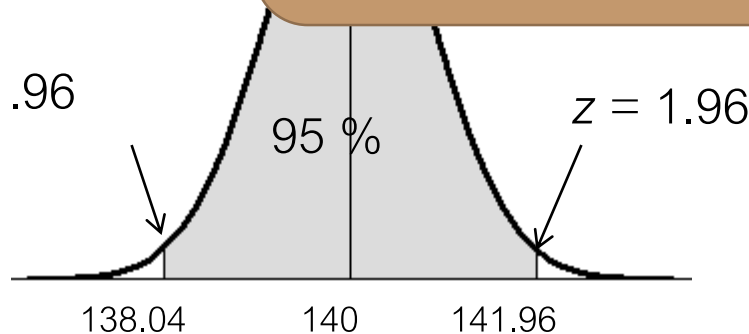
สุ่ม $n = 100$



$$\begin{aligned} \mu_M &= 140 \\ \sigma_M &= 1 \end{aligned}$$



$z = -1.96$



ถ้าสุ่มคนจากประชากรที่สูง 140 cm
จะมี 95 % ที่ได้กลุ่มตัวอย่างสูงเฉลี่ย
138.04 ถึง 141.96

การประมาณค่าแบบช่วง

- คำถาม Parameter ตัวไหน ที่น่าจะสุ่มได้ค่าสถิติเช่นนี้
- ค่า Parameter อะไรบ้าง ที่จะสุ่มได้กลุ่มตัวอย่าง 100 คนมีค่าเฉลี่ยเท่ากับ 160 cm (สมมติส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 cm)

■ $\mu = 158$ สุ่ม 95 % อยู่ในช่วง (156.04, 159.96)

■ $\mu = 159$ สุ่ม 95 % อยู่ในช่วง (157.04, 160.96)

■ $\mu = 160$ สุ่ม 95 % อยู่ในช่วง (158.04, 161.96)

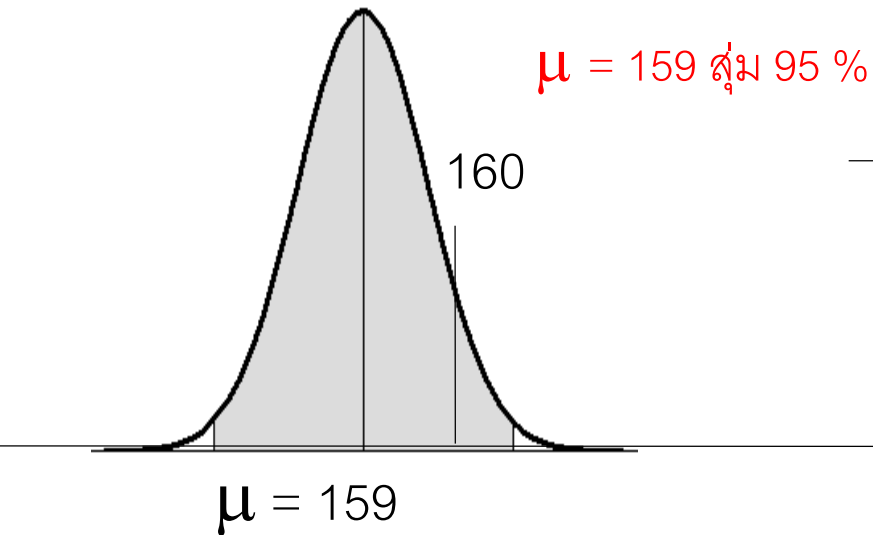
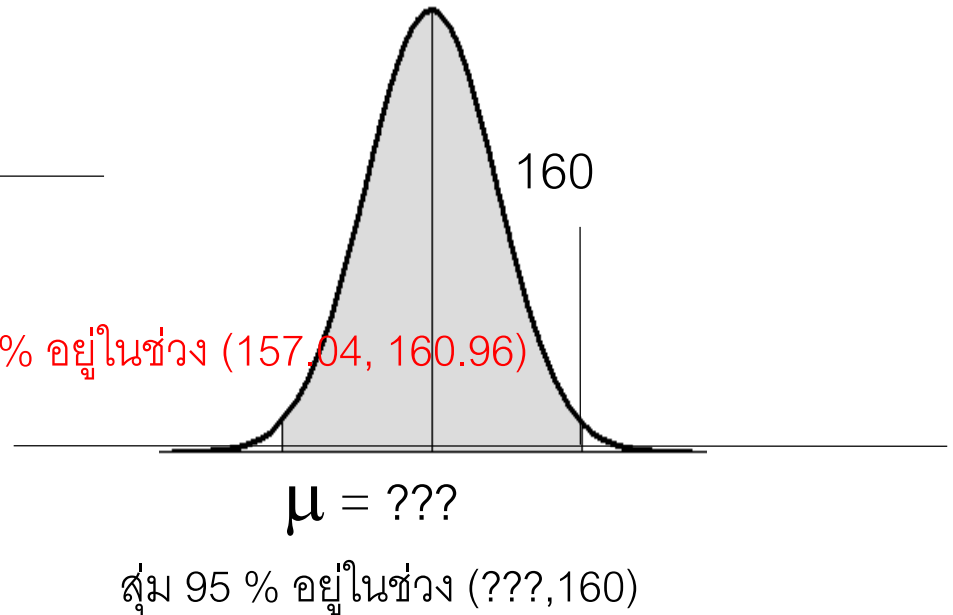
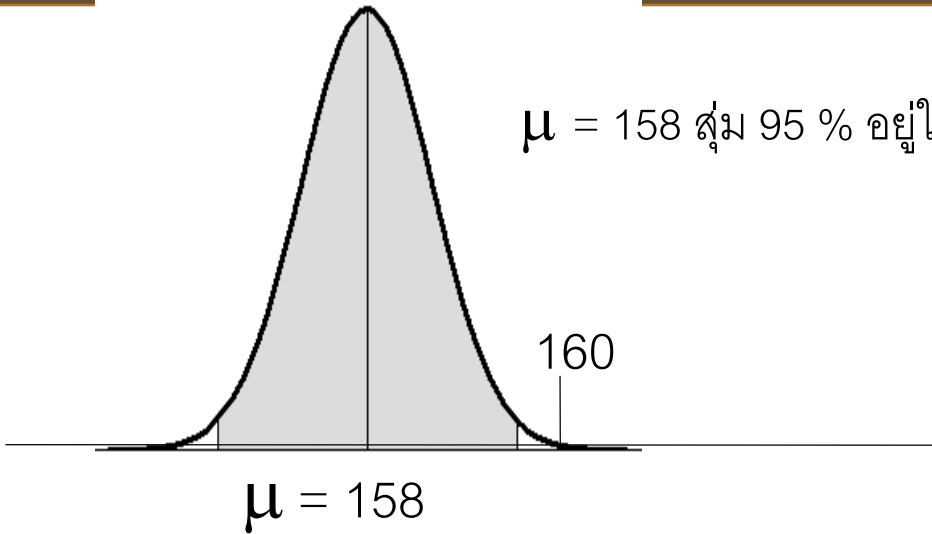
■ $\mu = 161$ สุ่ม 95 % อยู่ในช่วง (159.04, 162.96)

■ $\mu = 162$ สุ่ม 95 % อยู่ในช่วง (160.04, 163.96)

พารามิเตอร์เหล่านี้
จะสุ่มกลุ่มตัวอย่าง
ได้ค่าเฉลี่ย 160 cm

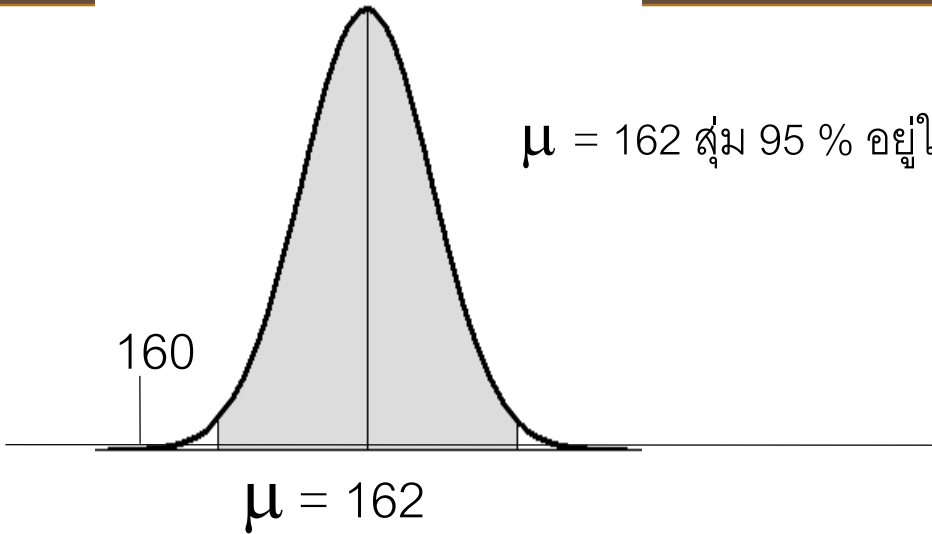
ช่วง 158.XX ถึง 161.XX

การประมาณค่าแบบช่วง

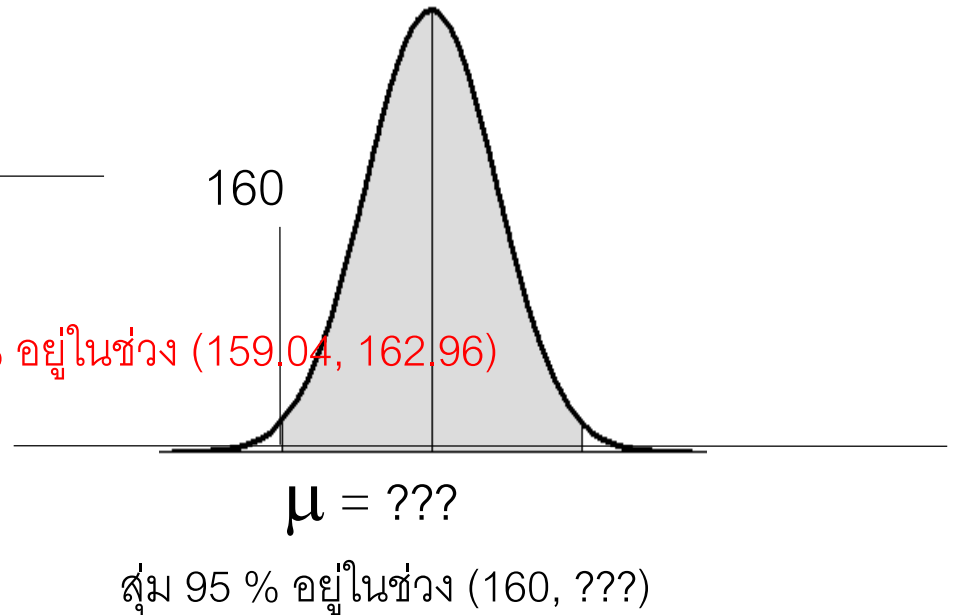
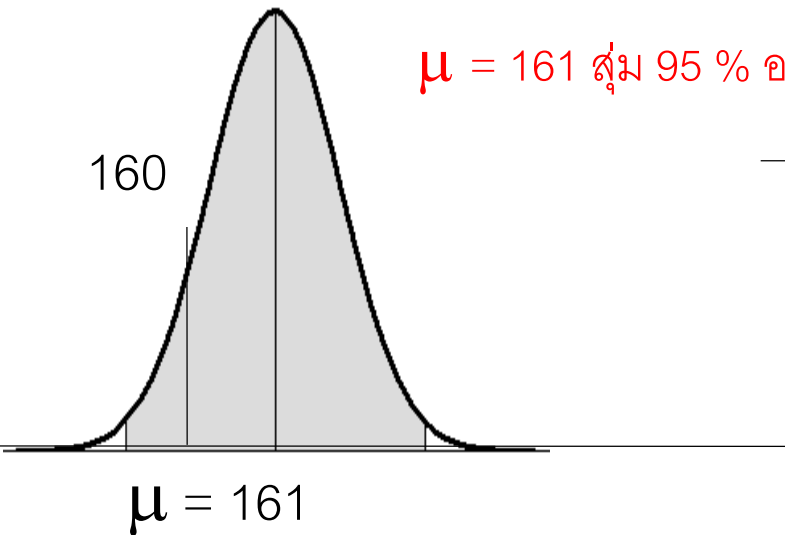


น่าจะอยู่ในช่วง 158-159

การประมาณค่าแบบช่วง



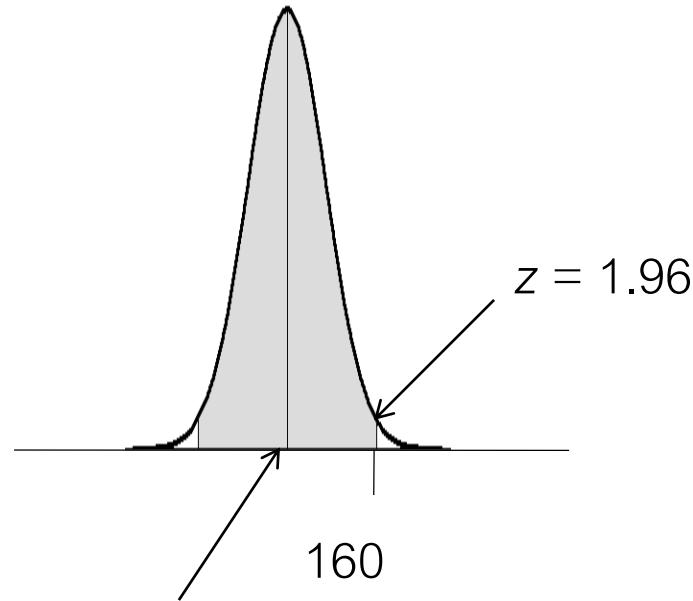
$\mu = 161$ สุ่ม 95 % อยู่ในช่วง (159.04, 162.96)



น่าจะอยู่ในช่วง 161-162

การประมาณค่าแบบช่วง

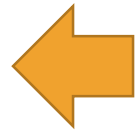
- ดูจากภาพ : $M = 160$; $N = 100$



ถ้าขยับ μ ไป
ทางซ้าย
มากกว่านี้ จะไม่
คลุม 160 แล้ว



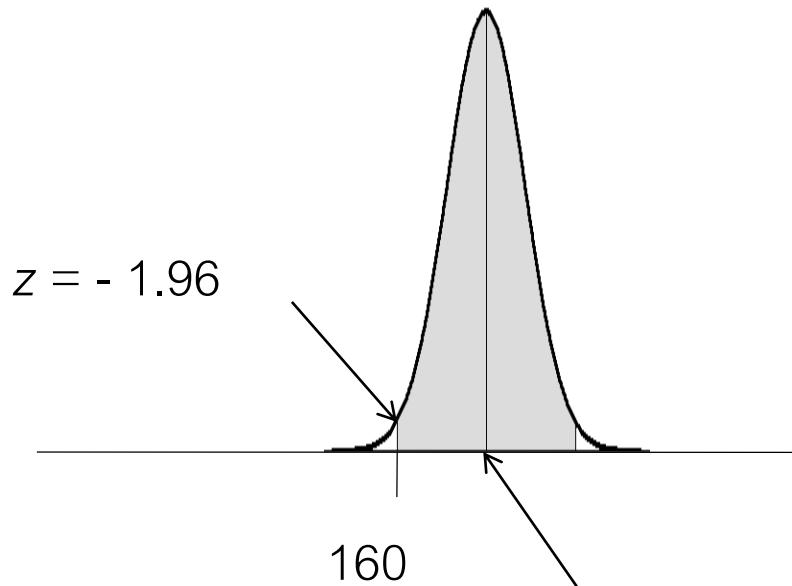
$$\mu = ???$$



$$1.96 = \frac{160 - \mu}{1}; \mu = 158.04$$

การประมาณค่าแบบช่วง

- ดูจากภาพ : $M = 160$; $N = 100$



$$-1.96 = \frac{160 - \mu}{1}; \mu = 161.96$$

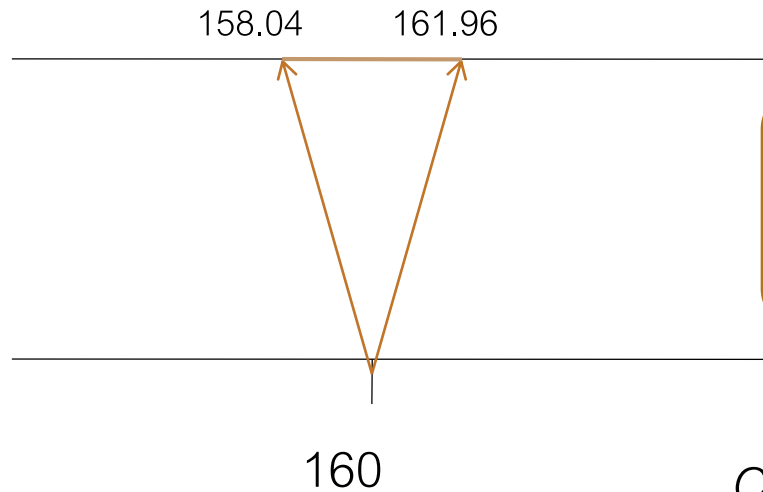
$\mu = ???$

ถ้าขยับ μ ไป
ทางขวา
มากกว่านี้ จะไม่
คลุม 160 แล้ว

การประมาณค่าแบบช่วง

- ดูจากภาพ

ช่วงนี้เรียกว่า ช่วงเชื่อมั่น
(Confidence Interval)



ช่วงเชื่อมั่น คือ ช่วงที่น่าจะ
สุ่มได้ค่าสถิติเป้าหมาย

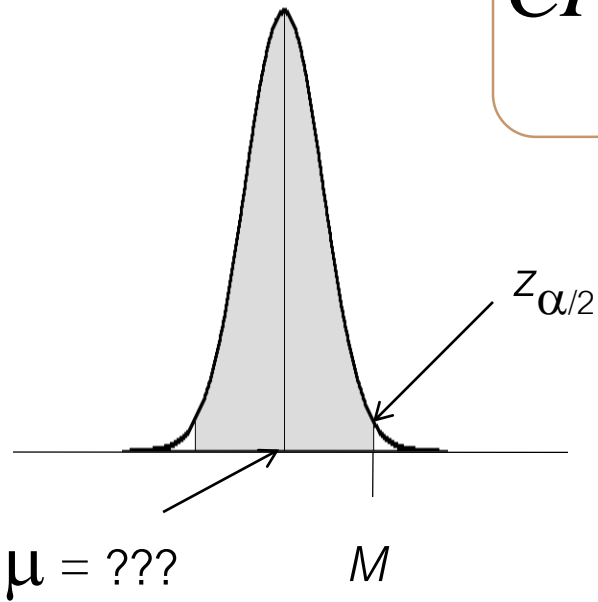
$$CI_{.95} = (158.04, 161.96)$$

ในที่นี้ จะเรียกว่า ช่วงเชื่อมั่นระดับ .95 เพราะสร้างจากการ
กระจายค่าเฉลี่ยของกลุ่มตัวอย่าง โดยตั้งช่วงกลาง 95 %

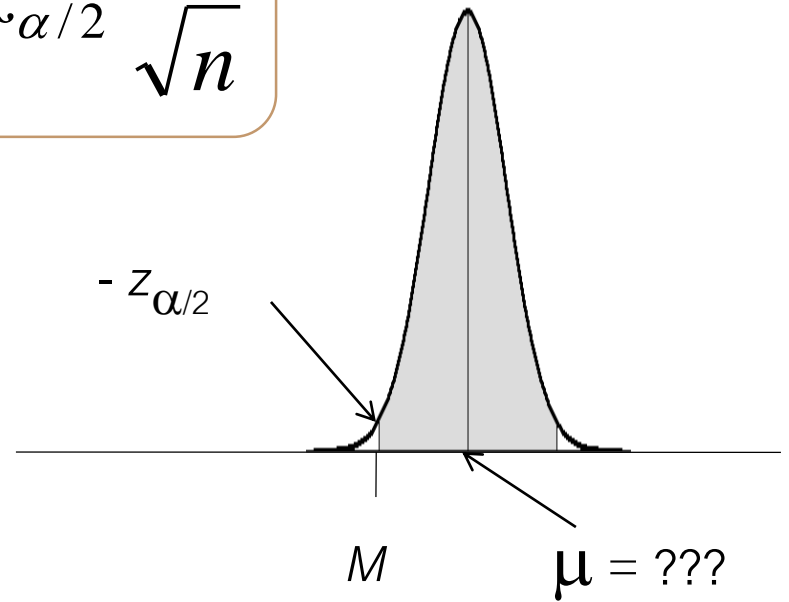
การประมาณค่าแบบช่วง

- สูตร

$$CI_{1-\alpha} = M \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$



$$z_{\alpha/2} = \frac{M - \mu}{\frac{\sigma}{\sqrt{n}}}; \mu = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} + M$$



$$-z_{\alpha/2} = \frac{M - \mu}{\frac{\sigma}{\sqrt{n}}}; \mu = -z_{\alpha/2} \frac{\sigma}{\sqrt{n}} + M$$

การประมาณค่าแบบช่วง

- ตัวอย่าง ต้องการทราบว่าประชากรไทยมีน้ำหนักโดยเฉลี่ยเท่าไร หากสุ่มกลุ่มตัวอย่างจากประชากรจำนวน 1,000 คน ได้ค่าเฉลี่ย 60 กิโลกรัม (ส่วนเบี่ยงเบนมาตรฐานของประชากรเท่ากับ 12 กิโลกรัม)
- ต้องการทราบช่วงเชื่อมั่นระดับ .95

การประมาณค่าแบบช่วง

- ประชากรไทย

กลุ่มตัวอย่าง 1,000 คน
ค่าเฉลี่ยน้ำหนัก 60 กิโลกรัม



ประชากรมีส่วนเบี่ยงเบนมาตรฐาน
เท่ากับ 12 กิโลกรัม

หาช่วงเชื่อมั่นระดับ .95

ค่า z คือ 1.96

การประมาณค่าแบบช่วง

- ประชากรไทย

กลุ่มตัวอย่าง 1,000 คน
ค่าเฉลี่ยน้ำหนัก 60 กิโลกรัม



ประชากรมีส่วนเบี่ยงเบนมาตรฐาน
เท่ากับ 12 กิโลกรัม



$$CI_{1-\alpha} = M \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 60 \pm 1.96(12 / \sqrt{1000})$$



$$CI_{.95} = 60 \pm 0.74$$

จากข้อมูลทำให้มีความเชื่อมั่นระดับ .95
ว่าประชากรจะมีค่าเฉลี่ยอยู่ในช่วง 59.26
ถึง 60.74 กิโลกรัม

การประมาณค่าแบบช่วง

- บอกได้หรือไม่ว่ามีความเป็นไปได้ 95 % ที่ประชากรอยู่ในช่วงนี้

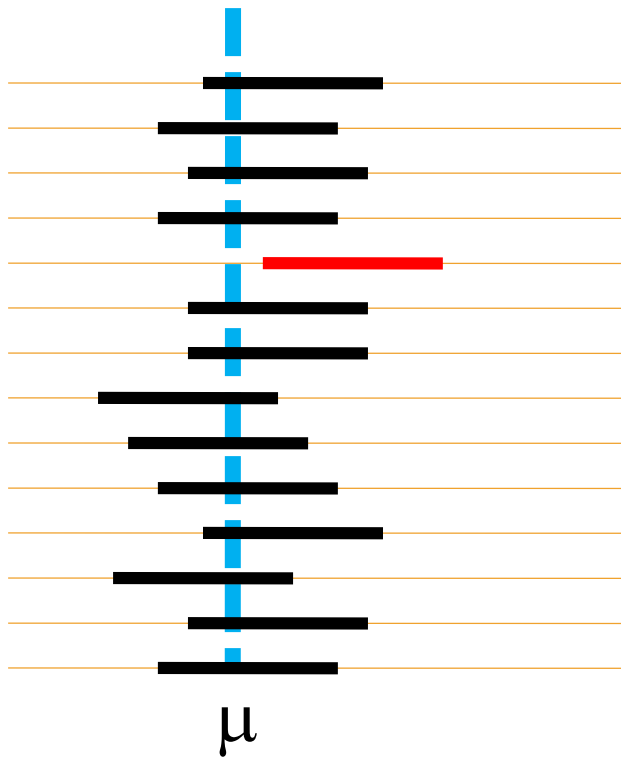
บอกไม่ได้



ค่าเฉลี่ยของประชากรมีเพียงแค่ 2 ประเภท คือ อยู่ในช่วงนี้ (100 %) หรือไม่อยู่ในช่วงนี้ (0 %) เท่านั้น

การประมาณค่าแบบช่วง

- แล้วช่วงเชื่อมั่นระดับ .95 คืออะไร
- สมมติว่าสุ่มกลุ่มตัวอย่างขนาด 1,000 คน มา 100 ครั้ง



สร้างช่วงเชื่อมั่นทั้งหมด 100 ครั้ง
จากค่าเฉลี่ยของกลุ่มตัวอย่าง 100 กลุ่ม
พบว่ามึค่าเฉลี่ยของประชากรอยู่ในช่วง
ทั้งหมด 95 ช่วง จาก 100 ช่วง

การประมาณค่าแบบช่วง

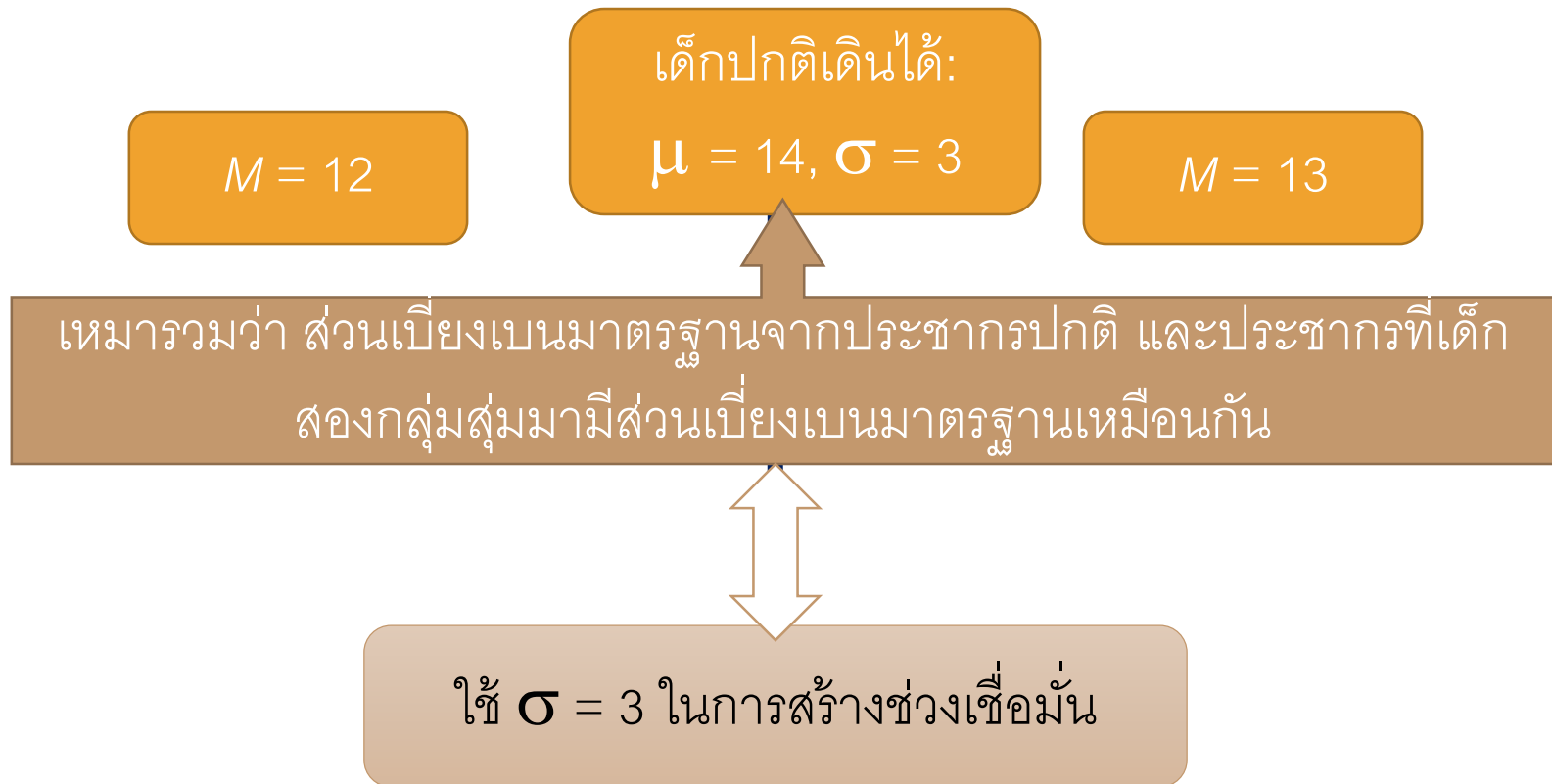
- จะสังเกตว่า การประมาณค่าแบบช่วงโดยใช้หลักการของการทดสอบซี จะต้องทราบส่วนเบี่ยงเบนมาตรฐานของประชากร (σ) เพื่อหาความผิดพลาดมาตรฐานของค่าเฉลี่ย (σ_M)
- ส่วนใหญ่เราจะไม่ทราบส่วนเบี่ยงเบนมาตรฐานของประชากร
- บทที่ 6 จะกล่าวถึงวิธีการประมาณค่าเฉลี่ย โดยไม่ทราบส่วนเบี่ยงเบนมาตรฐานของประชากร แต่รู้ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

การประมาณค่าแบบช่วง

- การประมาณค่าแบบช่วงสามารถใช้ในการทดสอบสมมติฐานได้ (ในที่นี้จะดูได้เฉพาะการทดสอบสมมติฐานแบบสองทาง)
- ถ้าช่วงเชื่อมั่น .95 คลุม Null Hypothesis หมายความว่า มีโอกาสที่ Null Hypothesis เป็นจริง ทำให้ไม่สามารถปฏิเสธ Null Hypothesis (Fail to Reject H_0) ได้ โดยตั้งระดับนัยสำคัญเท่ากับ .05
- แต่ถ้าช่วงเชื่อมั่น .95 ไม่คลุม Null Hypothesis หมายความว่า มีโอกาสน้อยที่ Null Hypothesis เป็นจริง ทำให้ปฏิเสธ Null Hypothesis (Reject H_0) ได้ โดยตั้งระดับนัยสำคัญเท่ากับ .05

การประมาณค่าแบบช่วง

- เช่น ทดสอบเด็กสองกลุ่มกลุ่มละ 16 คน มาจากสถานเลี้ยงเด็ก A และ B แบบ A เด็กเดินได้ตอนอายุ 12 เดือน แบบ B เด็กเดินได้ตอนอายุ 13 เดือน



การประมาณค่าแบบช่วง

- เช่น ทดสอบเด็กสองกลุ่มกลุ่มละ 16 คน มาจากสถานเลี้ยงเด็ก A และ B แบบ A เด็กเดินได้ตอนอายุ 12 เดือน แบบ B เด็กเดินได้ตอนอายุ 13 เดือน

เด็กปกติเดินได้:

$$\mu = 14, \sigma = 3$$

$$M = 12$$

$$M = 13$$

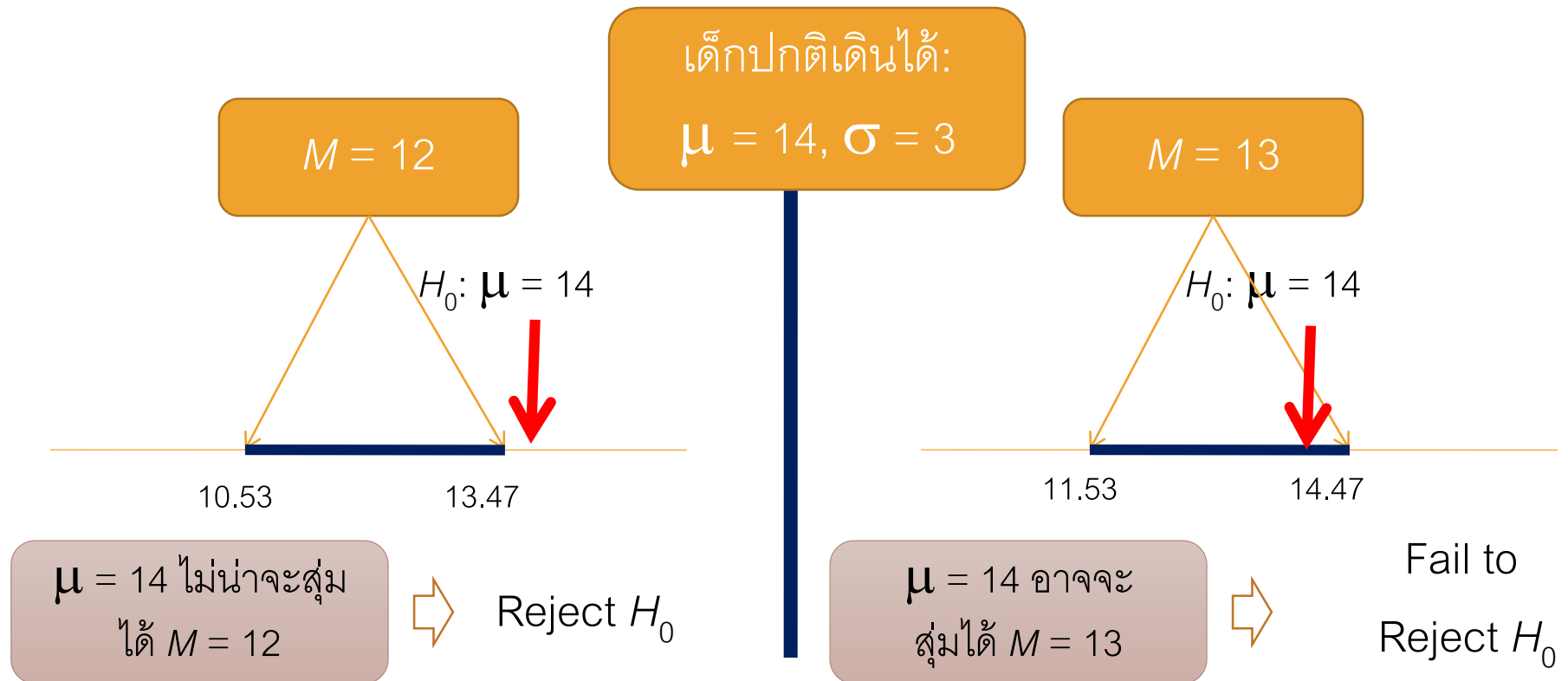
$$CI_{1-\alpha} = M \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 12 \pm 1.96(3/\sqrt{16}) \quad CI_{1-\alpha} = M \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 13 \pm 1.96(3/\sqrt{16})$$

$$CI_{.95} = (10.53, 13.47)$$

$$CI_{.95} = (11.53, 14.47)$$

การประมาณค่าแบบช่วง

- เช่น ทดสอบเด็กสองกลุ่มกลุ่มละ 16 คน มาจากสถานเลี้ยงเด็ก A และ B แบบ A เด็กเดินได้ตอนอายุ 12 เดือน แบบ B เด็กเดินได้ตอนอายุ 13 เดือน



การประมาณค่าแบบช่วง

- เช่น ทดสอบเด็กสองกลุ่มกลุ่มละ 16 คน มาจากสถานเลี้ยงเด็ก A และ B แบบ A เด็กเดินได้ตอนอายุ 12 เดือน แบบ B เด็กเดินได้ตอนอายุ 13 เดือน

เด็กที่อยู่ในสถานเลี้ยงเด็ก
A เดินได้เร็วกว่าเด็กปกติ

$\mu = 14$ ไม่น่าจะสும்
ได้ $M = 12$



Reject H_0

เด็กที่อยู่ในสถานเลี้ยงเด็ก B
ไม่สามารถสรุปได้ว่าเดินได้เร็ว
กว่าเด็กปกติหรือไม่

$\mu = 14$ อาจจะ
สும்ได้ $M = 13$



Fail to
Reject H_0

การประมาณค่าแบบช่วง

- ตัวอย่าง โรงงานแห่งหนึ่งโดยปกติผลิตฝากระป๋อง ขนาด 6.500 cm และมีส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.01 cm แต่วันหนึ่งคนคุมได้สุ่มกระป๋องเพื่อมาตรวจคุณภาพจำนวน 20 กระป๋อง พบว่ากระป๋องมีขนาดเฉลี่ย 6.497 cm คนคุมสายการผลิตคิดว่าเครื่องอาจจะผิดพลาด โดยตั้งระดับนัยสำคัญเท่ากับ .01
- จงใช้วิธีการประมาณค่าแบบช่วงในการทดสอบสมมติฐาน
- หมายเหตุ $\alpha = .01$ (สองทาง) จะได้จุดวิกฤตเท่ากับ -2.58 และ 2.58
- $\sqrt{20} = 4.5$

การประมาณค่าแบบช่วง

- ตัวอย่าง

กระป๋องที่สุ่มออกมา 20 กระป๋อง
ค่าเฉลี่ยเท่ากับ 6.497 cm



สมมติว่า ส่วนเบี่ยงเบน
มาตรฐานไม่เปลี่ยนแปลง



สมมติฐาน คือ เครื่องจักรใน
การผลิตกระป๋องผิดปกติ



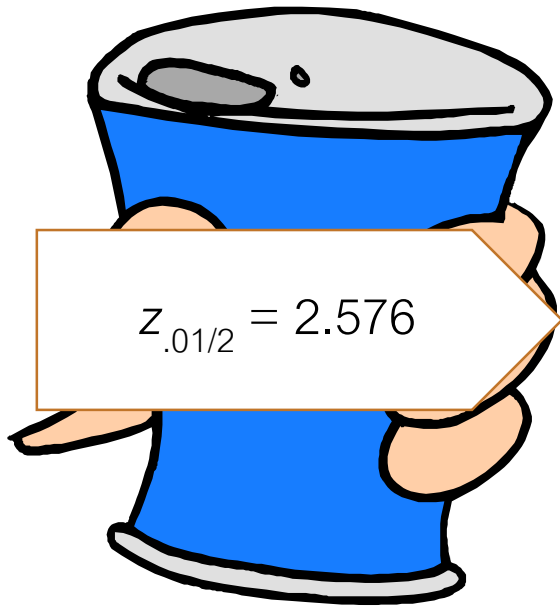
เครื่องจักรผลิตกระป๋องปกติ
 $\mu = 6.5, \sigma = 0.01$

คำถาม: กระป๋องทั้ง 20 กระป๋องนี้ น่าจะสุ่ม
ออกมาจากเครื่องจักรผลิตกระป๋องขนาดค่าเฉลี่ยเท่าไร

การประมาณค่าแบบช่วง

- ตัวอย่าง

กระป๋องที่สุ่มออกมา 20 กระป๋อง
ค่าเฉลี่ยเท่ากับ 6.497 cm



$M = 6.497$

หา $CI_{.99}$

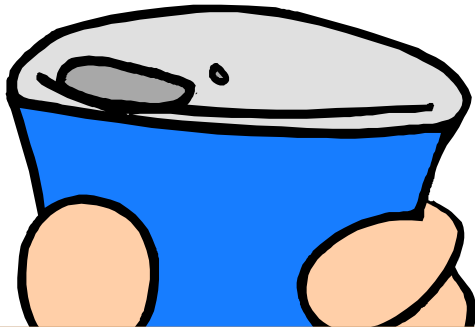
$$CI_{1-\alpha} = M \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 6.497 \pm 2.576(0.01 / \sqrt{20})$$

$CI_{.99} = (6.491, 6.502)$

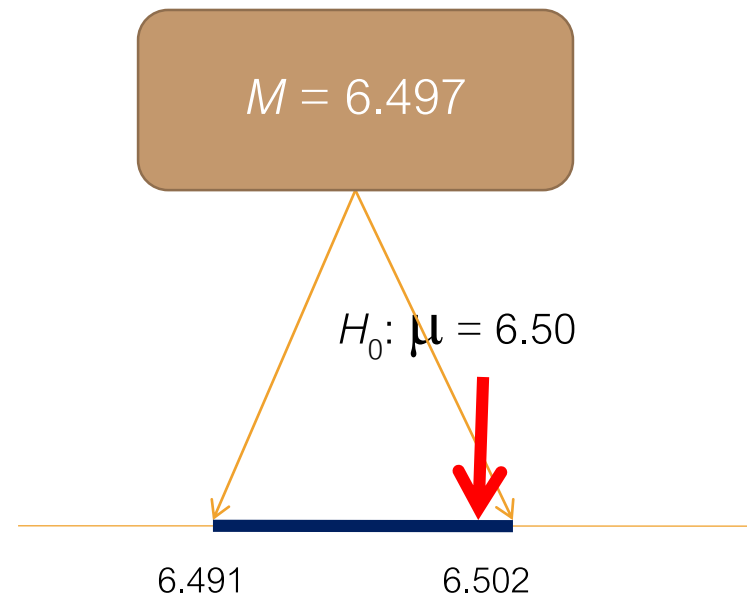
การประมาณค่าแบบช่วง

- ตัวอย่าง

กระป๋องที่สุ่มออกมา 20 กระป๋อง
ค่าเฉลี่ยเท่ากับ 6.497 cm



สมมติฐานว่างอาจจะเป็นจริงก็ได้
กล่าวคือ กระบวนการผลิตกระป๋อง
อาจจะปกติหรือไม่ก็ได้



H_0 อยู่ใน CI

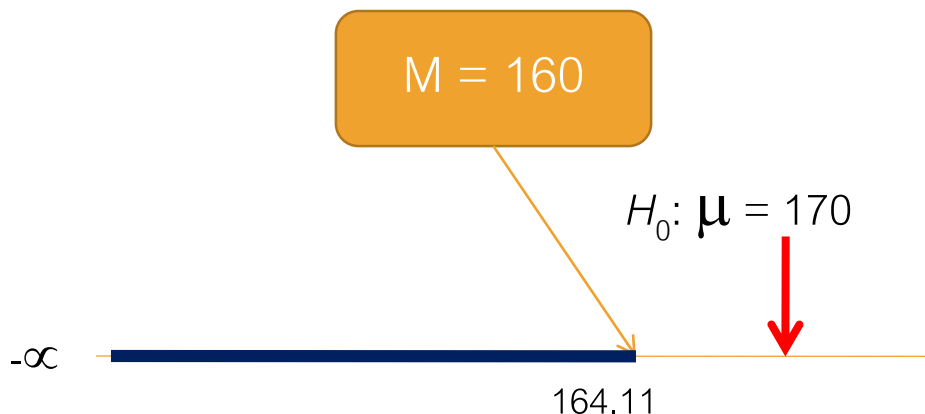
Fail to Reject H_0

การประมาณค่าแบบช่วง

- ช่วงเชื่อมั่นทำให้สามารถทดสอบ Null Hypothesis หลายสมมติฐานพร้อมกันได้
 - เช่น คน 4 คน เพศเดียวกัน ชาติเดียวกัน สูงเฉลี่ย 160 cm ได้ช่วงเชื่อมั่น .95 เท่ากับ 155.4 ถึง 164.9
 - สมมติ ผู้ชายไทยสูงเฉลี่ย 170 cm → 4 คนนี้ไม่น่าจะเป็น **นอกช่วง**
 - ผู้หญิงไทยสูงเฉลี่ย 157 cm → 4 คนนี้อาจจะเป็น **ในช่วง**
 - ผู้ชายยุโรปสูงเฉลี่ย 180 cm → 4 คนนี้ไม่น่าจะเป็น **นอกช่วง**
 - ผู้หญิงญี่ปุ่นสูงเฉลี่ย 156 cm → 4 คนนี้อาจจะเป็น **ในช่วง**

การประมาณค่าแบบช่วง

- ความจริงแล้ว การทดสอบสมมติฐานแบบทิศทางเดียว ก็สามารถใช้ช่วงเชื่อมั่นในการทดสอบได้
- จะต้องใช้วิธีการสร้างช่วงเชื่อมั่นอีกรูปแบบหนึ่ง
- แต่ช่วงเชื่อมั่นที่ได้จากวิธีนี้ ไม่ค่อยเหมือนกับความเป็นจริงมากนัก
- และวิธีนี้ไม่ค่อยได้รับความนิยม จึงไม่ได้แสดงในที่นี้



ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- ข้อตกลงเบื้องต้นก่อนการใช้สถิติ (Statistical Assumption) คือ ลักษณะของการออกแบบการทดลอง หรือลักษณะของข้อมูลที่เราควรจะเป็นก่อนการใช้สถิติ เพื่อให้ค่าสถิติที่ออกมานั้นถูกต้อง
- หากละเมิดข้อตกลงเบื้องต้นก่อนการใช้สถิติ อาจส่งผลกระทบมากหรือน้อย ขึ้นอยู่กับว่า เป็นข้อตกลงเบื้องต้นอะไร

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ z-test หรือช่วงเชื่อมั่นที่สร้างจาก z-test

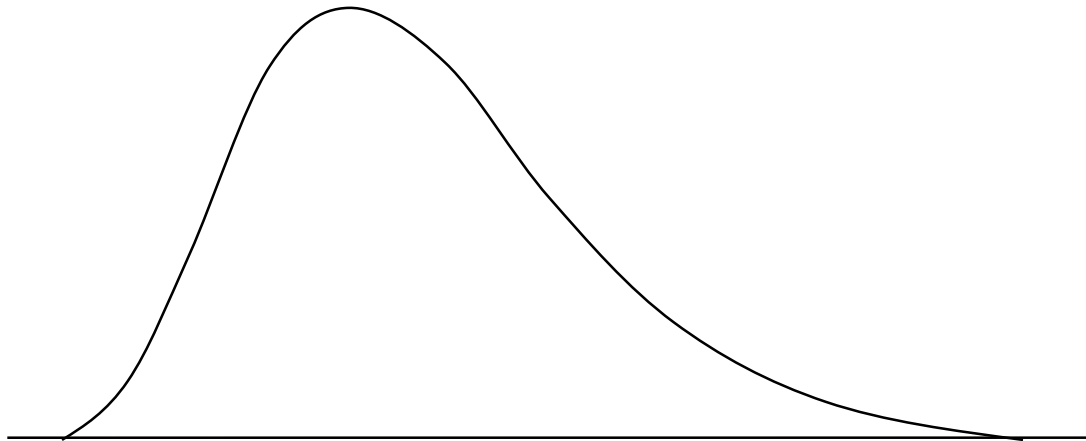
- การกระจายค่าเฉลี่ยของกลุ่มตัวอย่างจะต้องเป็นโค้งปกติ
 - ได้จากการกระจายของประชากรเป็นโค้งปกติ
 - หรือ จำนวนกลุ่มตัวอย่างที่สุ่มมากกว่า 30
- ต้องรู้ค่าส่วนเบี่ยงเบนมาตรฐาน (หรือความแปรปรวน) ของประชากร

การกระจายของกลุ่มตัวอย่าง

- อาจจะมีการกระจายของค่าสถิติรูปแบบอื่นได้ เช่น มัธยฐาน ความแปรปรวน
- โดยรวมแล้ว การสุ่มกลุ่มตัวอย่างแบบเป็นไปได้อย่างทั้งหมด มาหาค่าสถิติ แล้วดูว่าค่าสถิติเหล่านั้นมีการกระจายอย่างไร จะเรียกว่า การกระจายของกลุ่มตัวอย่าง (Sampling Distribution)
- การกระจายของค่าสถิติ ไม่จำเป็นต้องเป็นโค้งปกติเสมอไป อาจมีการกระจายรูปแบบอื่นได้ (ซึ่งจะเรียนในบทต่อไป)

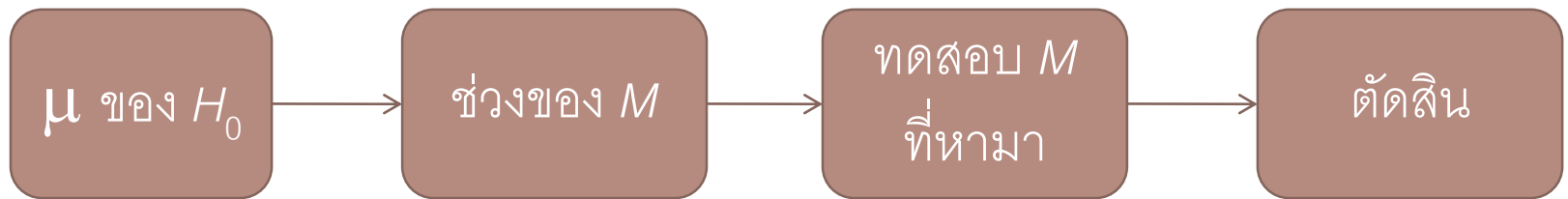
การกระจายของกลุ่มตัวอย่าง

- เช่น การกระจายของความแปรปรวนจากกลุ่มตัวอย่างขนาด 11 คน
 - ลักษณะจะเป็นการกระจายเบ้ขวา
 - การกระจายนี้ จะอยู่ในกลุ่มการกระจายที่เรียกว่า การกระจายแบบไคสแควร์ (Chi-square Distribution)

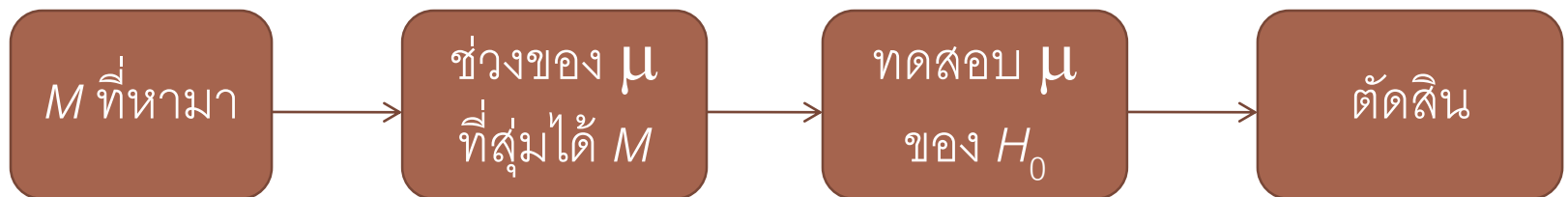


สรุปวิธีการพื้นฐานในสถิติอ้างอิง

- การทดสอบสมมติฐาน



- การประมาณค่าพารามิเตอร์



ทั้งสองแนวคิด เป็นส่วนกลับซึ่งกันและกัน

สรุปวิธีการพื้นฐานในสถิติอ้างอิง

การทดสอบสมมติฐาน

- ตั้งสมมติฐานวิจัย
- หา Null Hypothesis
- สร้างการกระจายค่าเฉลี่ยจาก Null Hypothesis
- เก็บข้อมูล
- ดูว่าค่าสถิติที่เก็บข้อมูลมา จะมีโอกาสเกิดขึ้นเมื่อ Null Hypothesis เป็นจริงกี่ %
- น้อยกว่าเกณฑ์ที่เรียกว่าน้อยหรือไม่
- ตัดสินว่าสนับสนุนสมมติฐานวิจัยหรือไม่

การประมาณค่าพารามิเตอร์

- ตั้งสมมติฐานวิจัย
- เก็บข้อมูล
- ดูว่าค่าสถิติที่เก็บข้อมูลมา น่าจะมาจากค่า Parameter อะไร โดยตั้งขนาดความเชื่อมั่น
- นำสมมติฐานวิจัย สร้าง Null Hypothesis
- ดูว่า Null Hypothesis อยู่ในช่วงเชื่อมั่นหรือไม่
- ตัดสินว่าสนับสนุนสมมติฐานวิจัยหรือไม่

คาบต่อไป

- ส่งการบ้านที่ 4
- เรียนคาบปฏิบัติการครั้งที่ 2
- ปรับปรุง Course Syllabus