บทนำ

สถิติขั้นกลางสำหรับจิตวิทยา

สันทัด พรประเสริฐมานิต

โครงร่างการนำเสนอ

- ทบทวนสถิติขั้นนำ
- เทคนิคการสุ่มซ้ำ
- การจัดการข้อมูลสูญหาย

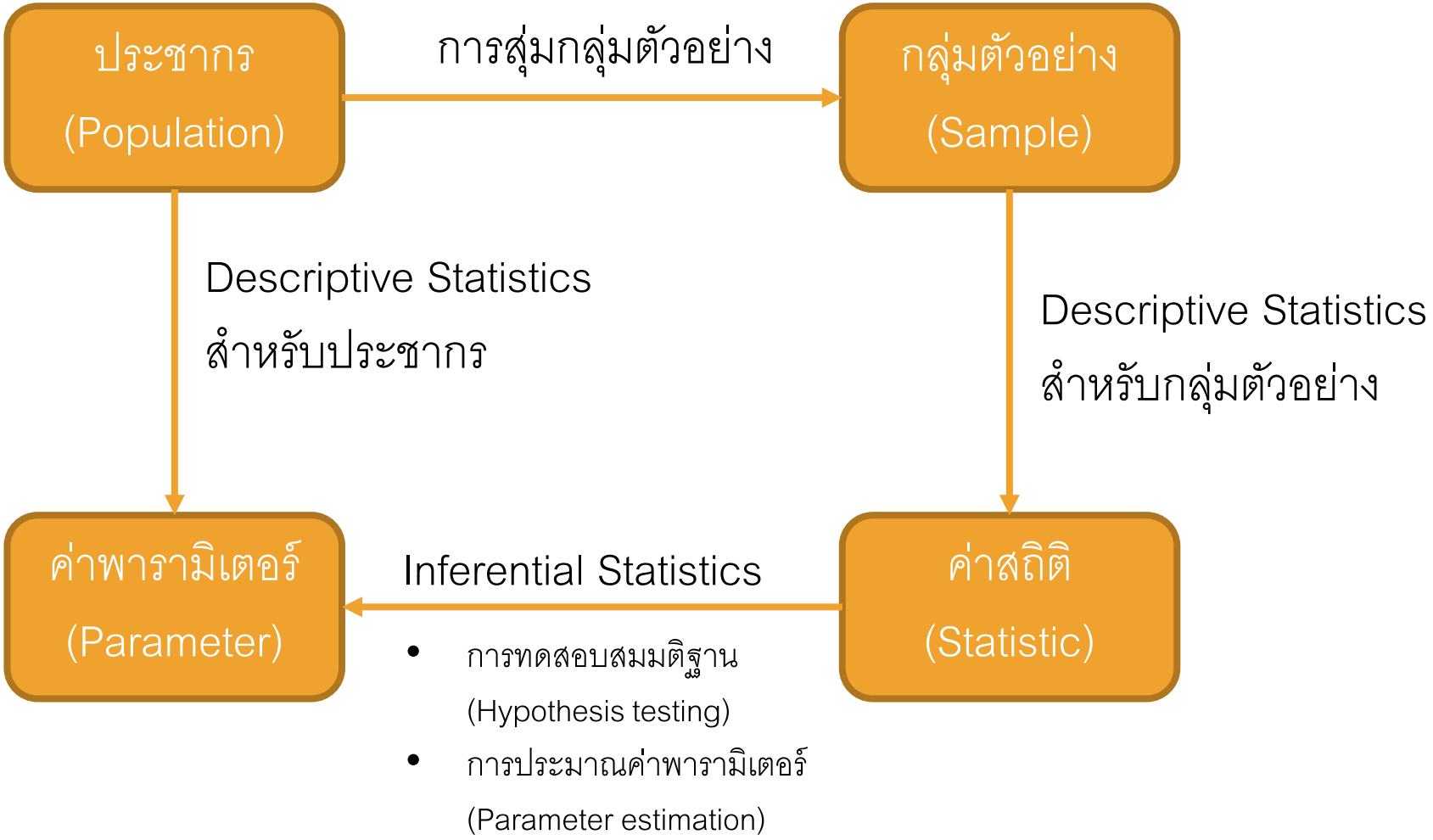
ทบทวนสถิติขั้นนำ

สถิติสำหรับจิตวิทยา 1

สันทัด พรประเสริฐมานิต

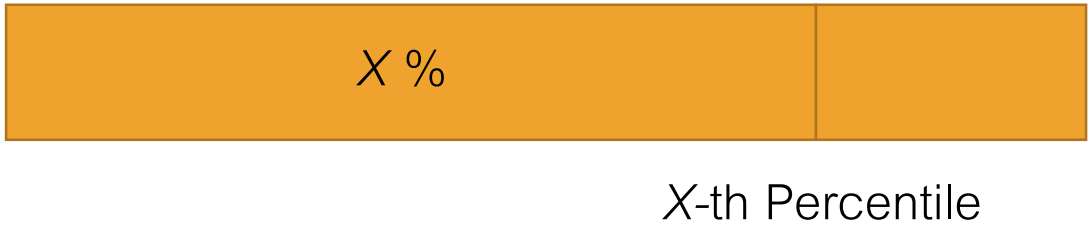
มโนทัศน์พื้นฐาน

- สถิติพรรณนา (Descriptive Statistics) เป็นสถิติที่ในการสรุปกลุ่มของข้อมูล
- สถิติอ้างอิง (Inferential Statistics) เป็นสถิติที่ใช้ในการประมาณการข้อมูลของกลุ่มที่ใหญ่กว่า จากข้อมูลจากกลุ่มที่เล็กกว่า


ประชากร
(Population)
การสุ่มกลุ่มตัวอย่าง
กลุ่มตัวอย่าง
(Sample)
Descriptive Statistics
สำหรับประชากร
Descriptive Statistics
สำหรับกลุ่มตัวอย่าง
ค่าพารามิเตอร์
(Parameter)
Inferential Statistics
ค่าสถิติ
(Statistic)
• การทดสอบสมมติฐาน
(Hypothesis testing)
• การประมาณค่าพารามิเตอร์
(Parameter estimation)

แนวโน้มสู่ศูนย์กลาง

- การวัดแนวโน้มสู่ศูนย์กลาง เป็นการหาค่ากลางของกลุ่มคะแนน ในที่นี้จะพูดถึง 3 ชนิด
  - ค่าเฉลี่ย (Mean) คือการนำคะแนนรวมมาหาด้วยจำนวนกลุ่มตัวอย่าง
  - มัธยฐาน (Median) คือ ค่าที่อยู่ตำแหน่งตรงกลางของข้อมูล
  - ฐานนิยม (Mode) คือ ค่าที่มีความถี่มากที่สุด
- ค่าเปอร์เซนต์ไทล์ (Percentile)



การกระจายของข้อมูล

- การกระจายของข้อมูล (Variability) เป็นการวัดขนาดความแตกต่างภายในกลุ่ม
  - พิสัย (Range) คือ คะแนนสูงสุดลบด้วยคะแนนต่ำสุด
  - พิสัยระหว่าง Quartile (Interquartile range) คือ Q3 – Q1
  - ความแปรปรวน (Variance)
  - ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)
- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “บทนำ”

การแปลความหมายคะแนน

- คะแนนดิบ (Raw score) คือ คะแนนที่ได้รับจากกลุ่มตัวอย่างโดยตรง
- ในบางครั้ง คะแนนดิบเพียงอย่างเดียว ไม่สามารถบอกได้ว่าคะแนนดิบนั้นสูงหรือต่ำ คะแนนเหล่านี้จะมีความหมายก็ต่อเมื่อถูกอ้างอิงกับสิ่งอื่น
- อิงเกณฑ์ (Criterion-referenced) เป็นการนำคะแนนดิบที่ได้ ไปเปรียบเทียบกับมาตรฐานจากผู้สร้างแบบทดสอบ ว่าคะแนนในแต่ละช่วง ผ่านมาตรฐานระดับใด
- อิงกลุ่ม (Norm-referenced) เป็นการนำคะแนนดิบที่ได้ ไปเปรียบเทียบกับคะแนนดิบของคนอื่นภายในกลุ่ม

ค่ามาตรฐาน

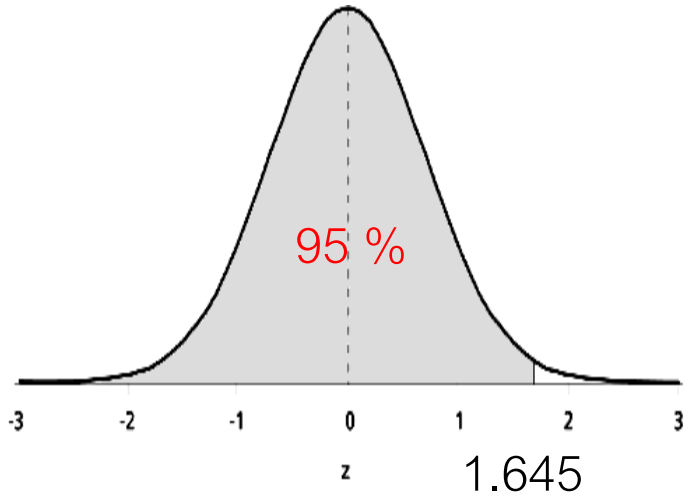
- การทำค่ามาตรฐาน (Standard score) หรือ *z*-score

$$z_i = \frac{X_i - M}{SD}$$

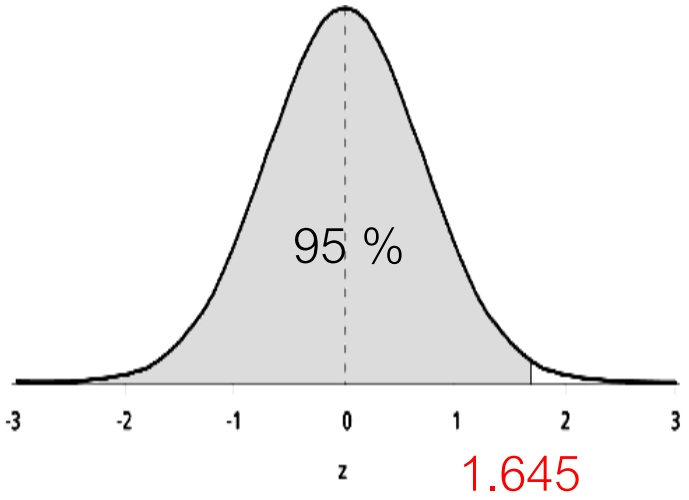
- ค่า z ก็คือการดูว่าคะแนนดังกล่าว ห่างจากค่าเฉลี่ยเป็นกี่เท่าของส่วนเบี่ยงเบนมาตรฐาน
- หากการกระจายเป็นโค้งปกติ สามารถหาเปอร์เซนต์ไทล์ของค่ามาตรฐานใต้โค้งปกติได้

ค่ามาตรฐาน

- Standard score -> Percentile
- =NORMSDIST(z)
- =NORMSDIST(1.645) = .95



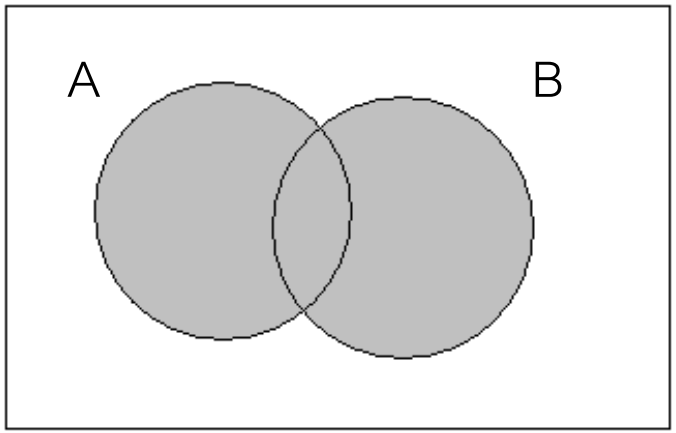
- Percentile -> Standard score
- =NORMSINV(Proportion)
- =NORMSINV(.95) = 1.645



ความน่าจะเป็น

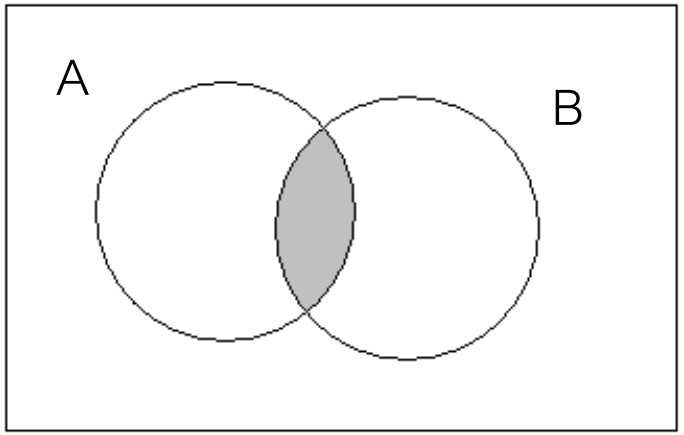
- ความน่าจะเป็น (Probability) คือ โอกาสในการเกิดเหตุการณ์ที่สนใจ

$$p(A \cup B)$$



Union

$$p(A \cap B)$$

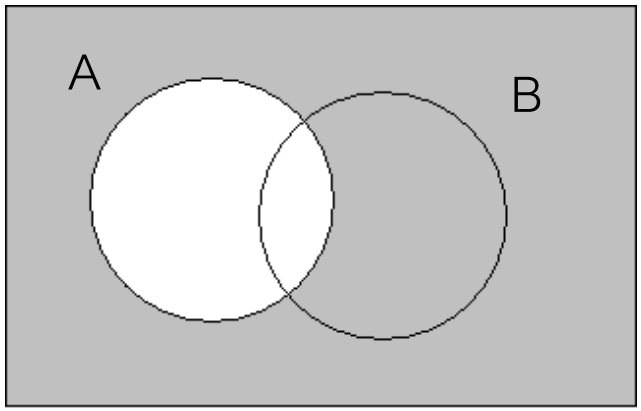


Intercept

ความน่าจะเป็น

- ความน่าจะเป็น (Probability) คือ โอกาสในการเกิดเหตุการณ์ที่สนใจ

$$p(A') = 1 - p(A)$$

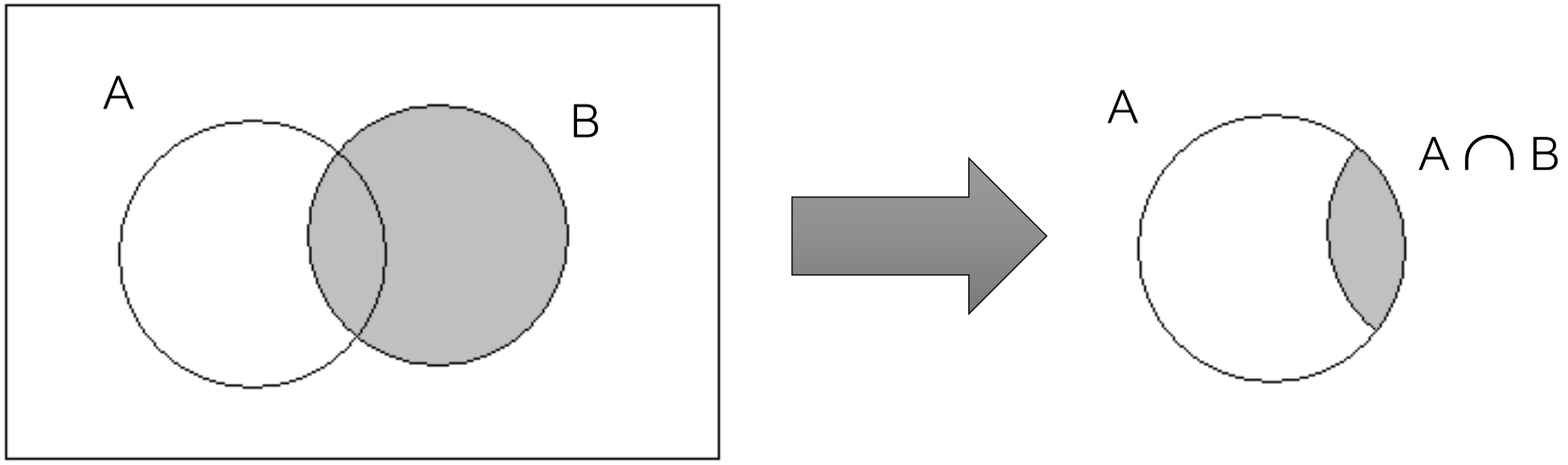


Complement

ความน่าจะเป็น

- ความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability)

$$p(B) = n(B)/n(S) \qquad p(B|A) = n(A \cap B)/n(A)$$



$$p(B|A) = \frac{n(B \cap A)/n(S)}{n(A)/n(S)} = \frac{p(B \cap A)}{p(A)}$$

ความน่าจะเป็น

- ความเป็นอิสระจากกันระหว่างตัวแปร (Statistical Independent) หมายถึง โอกาสในการเกิด A ไม่เกี่ยวข้องกับการเกิด B

$$p(A) = p(A|B) = p(A|B') \qquad p(B) = p(B|A) = p(B|A')$$

- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “แนะนำสถิติเชิงอ้างอิง”

การทดสอบสมมติฐาน

- เป็นการทดสอบว่าสมมติฐานทางสถิติเกี่ยวกับลักษณะในประชากรที่ตั้งไว้ถูกต้องหรือไม่ เช่น
  - นักเรียนคนหนึ่งเรียนไม่ค่อยดี ครูเห็นว่านักเรียนคนนี้อาจจะมีความผิดปกติด้านการเรียน (Learning Disability) จึงนำไปทดสอบแบบวัดความจำ พบว่าสามารถจำคำได้ 4 คำ
  - เด็กปกติสามารถจำได้: $\mu = 7$, $\sigma = 1.5$

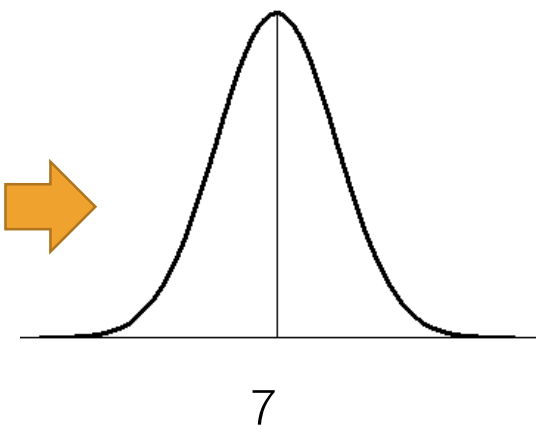
การทดสอบสมมติฐาน

- ปกติจะสร้างการกระจายของความสามารถในการจำของเด็กปกติออกมา โอกาสที่เจอเด็กที่จำได้น้อยกว่าหรือเท่ากับ 4 คำ มีเท่าไร
- ก็คือ นำ Null Hypothesis ไปสร้างการกระจายนั่นเอง โดยเลือก $\mu = 7$ ไปสร้างการกระจาย

สมมติฐานวิจัย
นร. คนนี้จำได้น้อยกว่านร. ปกติ

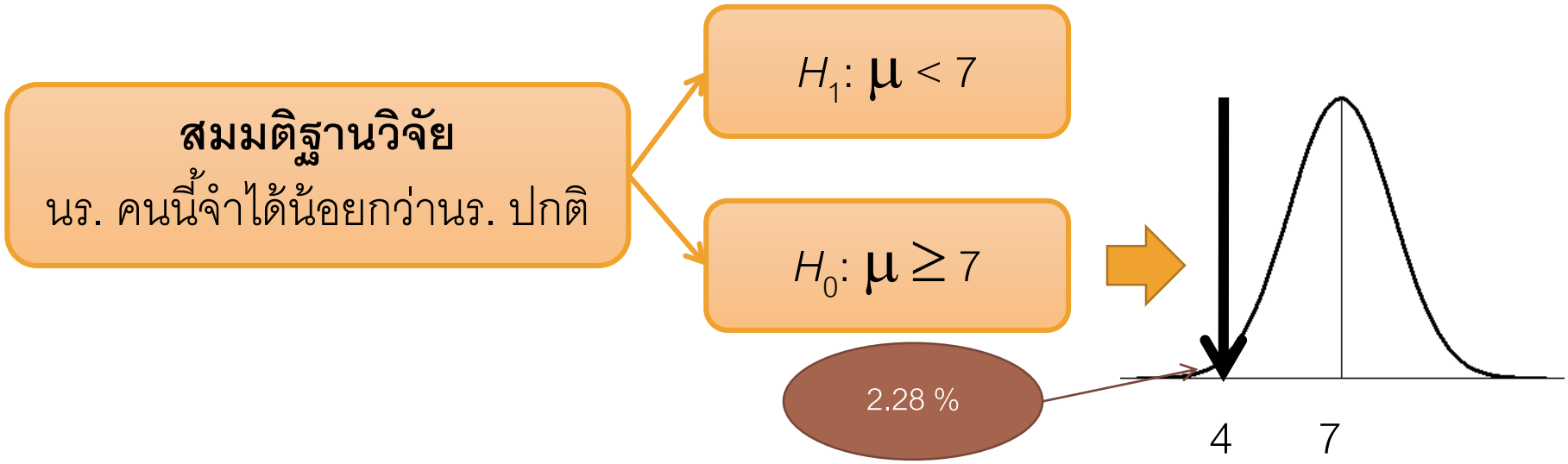
$H_1: \mu < 7$

$H_0: \mu \geq 7$



การทดสอบสมมติฐาน

- โอกาสที่จะเจอนักเรียนที่จำได้ 4 คำ หาได้โดย
  - แปลงให้เป็น *z*-score มีค่าเท่ากับ (4 – 7) / 1.5 = -2
  - % จากตังแต่ -∝ ถึง -2 เท่ากับ 2.28 %



การทดสอบสมมติฐาน

- ถ้าสมมติว่า เราตัดสินว่า ถ้าโอกาสเกิดน้อยกว่า 5 % แสดงว่าโอกาสเกิดน้อยมาก จึงน่าจะมีทักษะในการจำช้ากว่าปกติ
- 2.28 % < 5 %
- นักเรียนคนนี้น่าจะมีทักษะในการจำช้า หรือ ปฏิเสธ Null Hypothesis (Reject H_0) แล้วบอกว่า Alternative Hypothesis น่าจะเป็นจริง

จำนวนทิศทาง

- การทดสอบสมมติฐาน สามารถสร้างสมมติฐานวิจัยได้ 2 รูปแบบ
  - การทดสอบทางเดียว (One-tailed test) คือ สมมติฐานวิจัยจะกำหนดทิศทางไว้แน่นอน เช่น คนนี้มีพัฒนาการล่าช้ากว่าปกติ
  - การทดสอบสองทาง (Two-tailed test) คือ สมมติฐานวิจัยไม่ได้กำหนดทิศทางไว้แน่นอน เช่น คนนี้มีพัฒนาการที่ผิดปกติ (อาจช้าเกินไป หรือเร็วเกินไป)

จำนวนทิศทาง	Null Hypothesis	Alternative Hypothesis
ทิศทางเดียว	$H_0: \mu \geq 7$	$H_1: \mu < 7$
สองทิศทาง	$H_0: \mu = 7$	$H_1: \mu \neq 7$

ระดับนัยสำคัญทางสถิติ

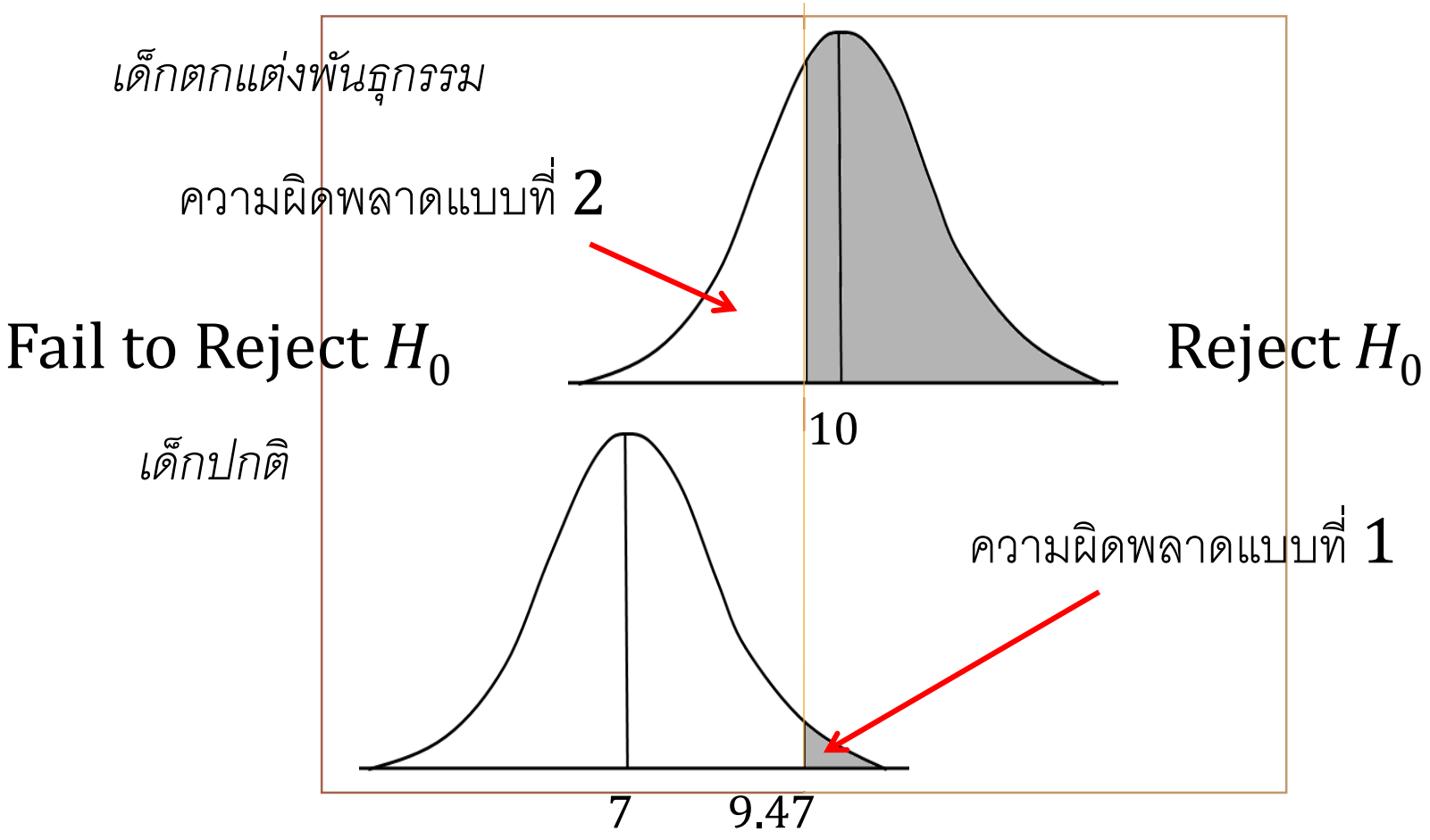
- เกณฑ์เท่าไรถึงตัดสินว่า Null Hypothesis ไม่น่าจะถูกต้อง ในตัวอย่างที่ผ่านมาได้ตั้งไว้ที่ 5 %
- เกณฑ์นี้จะเรียกว่า ระดับนัยสำคัญทางสถิติ (Significance Level) หรือ ค่าอัลฟ่า (Alpha; α)
- เราอาจเปลี่ยนระดับนัยสำคัญจาก 5 % เป็น 10%, 1% หรือระดับอื่นๆ ได้

ระดับนัยสำคัญทางสถิติ

- ความน่าจะเป็นของที่ Null hypothesis เป็นจริง (*p* value)
- หากค่านี้น้อยกว่าระดับนัยสำคัญ จะทำให้ Null Hypothesis ไม่น่าจะเป็นจริง

- จุดของคะแนนดิบหรือค่ามาตรฐานที่ตรงกับระดับนัยสำคัญ (Critical value)
- บริเวณที่โอกาสที่ Null hypothesis เป็นจริงน้อยกว่าระดับนัยสำคัญ คือ บริเวณวิกฤต (Critical region)
- คะแนนดิบตกลงบริเวณวิกฤต จะทำให้ Null Hypothesis ไม่น่าจะเป็นจริง

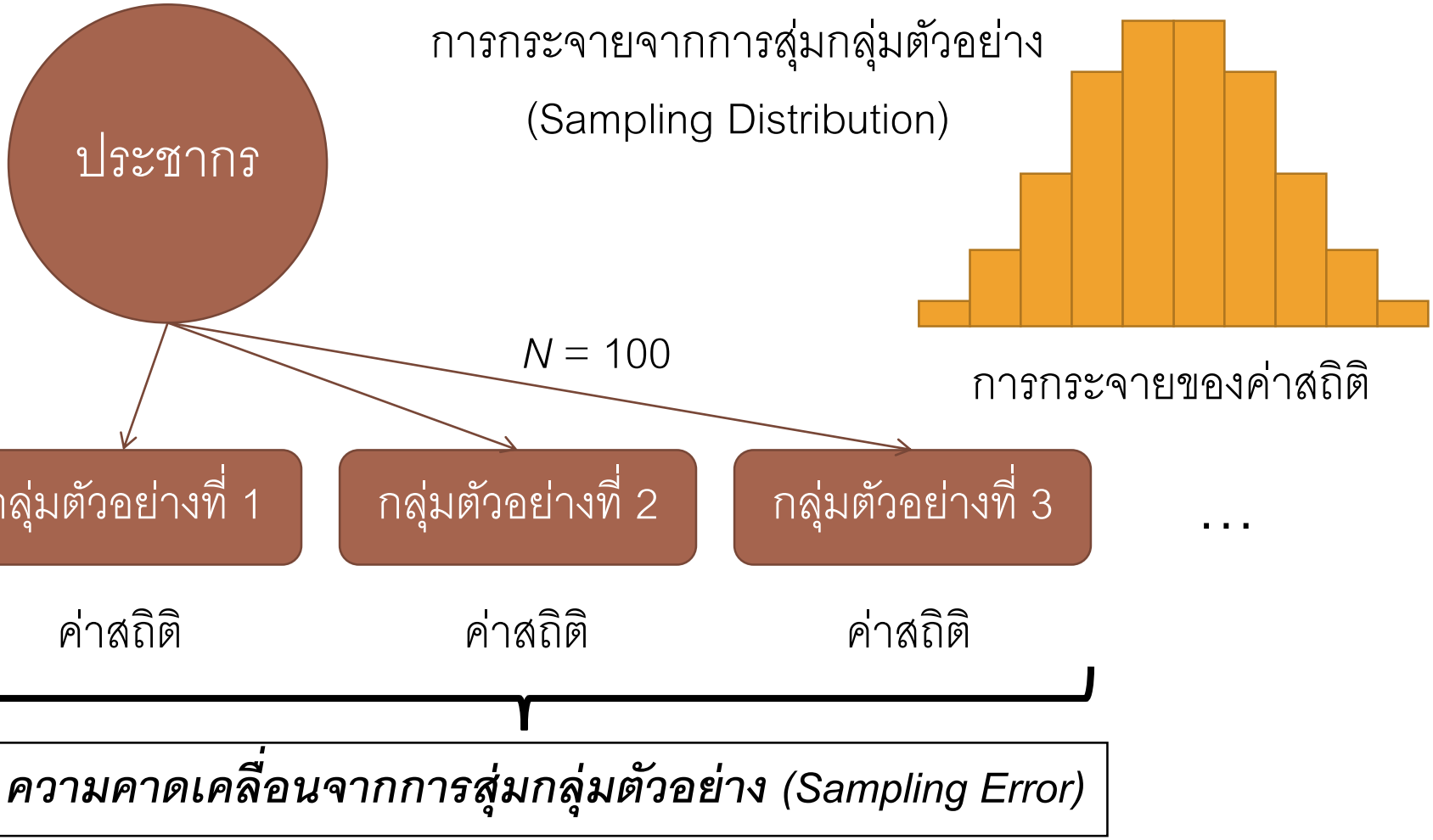
รูปแบบความผิดพลาดของกระบวนการทดสอบสมมติฐาน



รูปแบบความผิดพลาดของกระบวนการทดสอบสมมติฐาน

- หากค่าที่ได้ไม่ถึงระดับนัยสำคัญ ไม่ได้หมายความว่า Null hypothesis เป็นจริง
  - อาจเป็นเพราะไม่มีความแตกต่างจริงๆ
  - หรืออาจเกิด Type II error
- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “การทดสอบสมมติฐานเบื้องต้น”

การกระจายจากการสุ่มกลุ่มตัวอย่าง



การกระจายจากการสุ่มกลุ่มตัวอย่าง

- หากค่าสถิติเป็นค่าเฉลี่ยแล้ว จะเรียกว่าการกระจายค่าเฉลี่ยของการสุ่มกลุ่มตัวอย่าง (Sampling distribution of means) มีคุณสมบัติตามทฤษฎีแนวโน้มสู่ศูนย์กลาง (Central limit theorem) ดังนี้
  - $\mu_M = \mu$
  - $\sigma_M = \sigma^2/N$
  - การกระจายค่าเฉลี่ยจากการสุ่มกลุ่มตัวอย่าง จะเข้าใกล้โค้งปกติ เมื่อการกระจายของประชากรเป็นโค้งปกติอยู่แล้ว หรือจำนวนกลุ่มตัวอย่างสูง (เช่น สูงกว่า 30)

การทดสอบ *z*

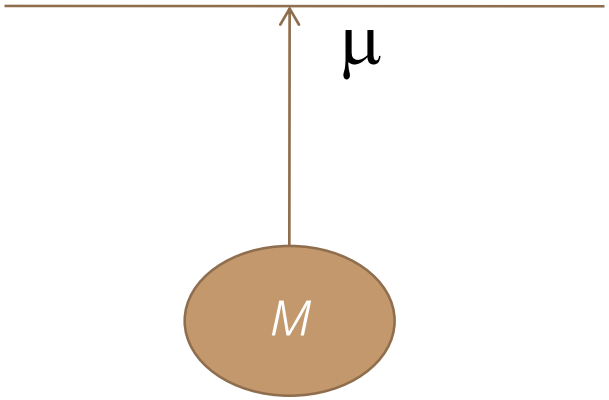
- การทดสอบ *z* (*z* test) เป็นการทดสอบว่าค่าเฉลี่ยจากกลุ่มตัวอย่างที่ได้ มาจากประชากรที่สนใจหรือไม่
- กล่าวคือ เป็นการเปรียบเทียบค่าเฉลี่ยของกลุ่มตัวอย่าง กับการกระจายค่าเฉลี่ยจากการสุ่มกลุ่มตัวอย่าง เมื่อ Null hypothesis เป็นจริง
- เขียนเป็นสูตรง่ายๆ ได้ดังนี้

$$z = \frac{M - \mu_M}{\sigma_M} = \frac{M - \mu}{\sigma / N}$$

การประมาณค่า

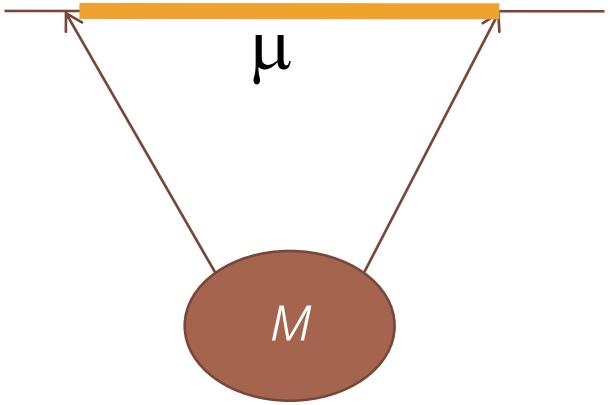
- การประมาณค่า สามารถทำได้ 2 รูปแบบด้วยกัน

การประมาณค่าแบบจุด



Parameter อะไรที่มีโอกาสสุ่มได้สถิติตัวนี้มากที่สุด

การประมาณค่าแบบช่วง



ช่วง Parameter อะไรที่น่าจะสุ่มได้สถิติตัวนี้สูง

การประมาณค่าแบบจุด

- ค่าสถิติที่ใช้ในการประมาณค่า Parameter ควรมีคุณสมบัติ 2 ประการ
  - ค่าเฉลี่ยของค่าสถิติจากการสุ่มกลุ่มตัวอย่างที่เป็นไปได้ทั้งหมด ต้องมีค่าเท่ากับค่า Parameter เรียกคุณสมบัตินี้ว่า ความแม่นยำ (Precision)

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

  - การกระจายของค่าสถิติจากการสุ่มกลุ่มตัวอย่างที่เป็นไปได้ทั้งหมด ควรมีค่าน้อยที่สุด เรียกคุณสมบัตินี้ว่า การแกว่งไปมาต่ำ (Efficiency)

$$\sigma_{STATISTICS} \text{ มีค่าต่ำ}$$

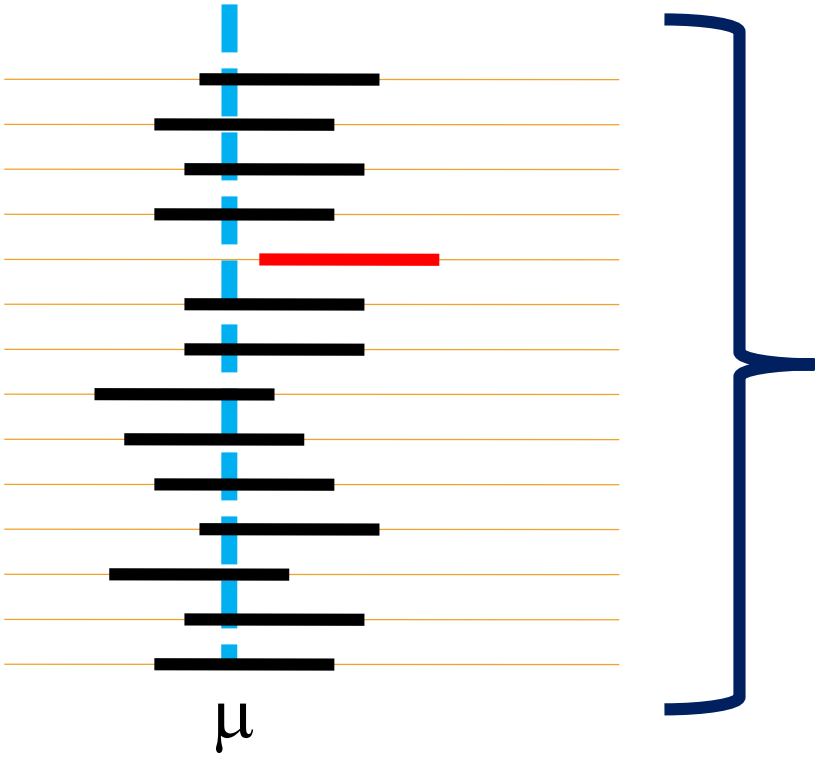
การประมาณค่าแบบช่วง

- การประมาณค่าแบบช่วง เป็นการหาว่า ช่วงใดที่น่าจะเป็นช่วง Parameter ภายใต้กลุ่มตัวอย่างนี้อยู่
- จะเรียกช่วงนั้นว่า ช่วงเชื่อมั่น (Confidence interval)
- เช่น ช่วงเชื่อมั่นของค่าเฉลี่ย

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

การประมาณค่าแบบช่วง

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



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

การประมาณค่าแบบช่วง

- บอกได้หรือไม่ว่ามีความเป็นไปได้ 95 % ที่ประชากรอยู่ในช่วงนี้

บอกไม่ได้

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

- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “การทดสอบสมมติฐานสำหรับค่าเฉลี่ยจากกลุ่มตัวอย่างเดียว”

ขนาดอิทธิพล

- “นัยสำคัญทางสถิติ (Statistical Significance)” ว่าหมายถึง Null Hypothesis มีแนวโน้มไม่เป็นจริง
  - เช่น เขาไม่ใช่คนปกติ; เขามี IQ สูงกว่าปกติ
- แต่ “นัยสำคัญทางสถิติ” ไม่ได้กล่าวถึงความรุนแรงของความแตกต่าง
  - เช่น เขาไม่ใช่คนปกติ แล้วเขาไม่ปกติขนาดไหน
  - เขามี IQ สูงกว่าปกติมากน้อยเพียงใด
- สิ่งที่จะบอก คือ “ขนาดอิทธิพล (Effect Size)”

ขนาดอิทธิพล

- ในการเปรียบเทียบค่าเฉลี่ย การวัดขนาดอิทธิพลอาจทำได้โดยการสำรวจคะแนนดิบ หรือความแตกต่างในหน่วยของส่วนเบี่ยงเบนมาตรฐาน
- ค่าหลัง จะเรียกว่า ขนาดอิทธิพลมาตรฐาน (Standardized Effect Size) หรือค่า *d* ของ Cohen (Cohen's *d*)

** ส่วนเบี่ยงเบนมาตรฐาน ไม่ใช่ Standard Error

$$\delta = \frac{\mu_1 - \mu_0}{\sigma}$$

$$d = \frac{M_1 - \mu_0}{\sigma}$$

คือ ค่าเฉลี่ยของกลุ่มที่ได้รับอิทธิพล ลบด้วยกลุ่มที่ไม่ได้รับอิทธิพล หารด้วยส่วนเบี่ยงเบนมาตรฐานของกลุ่มที่ไม่ได้รับอิทธิพล

ช่วงเชื่อมั่นของขนาดอิทธิพล

- การหาช่วงเชื่อมั่นของขนาดอิทธิพล เป็นการหาว่าค่า Parameter ของขนาดอิทธิพลอยู่ในช่วงอะไร

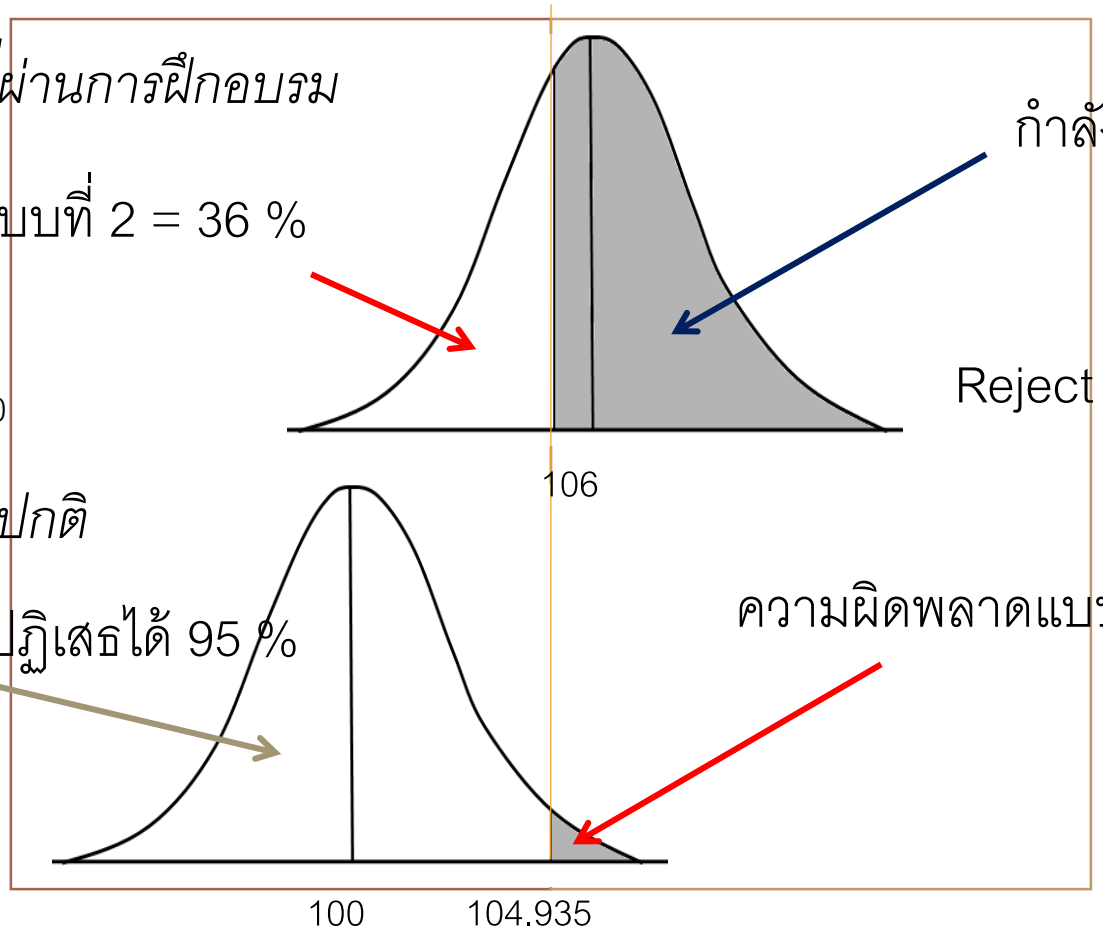
$$CI_{1-\alpha} \text{ of } \delta = d \pm \frac{z_{\alpha/2}}{\sqrt{n}}$$

กำลังทางสถิติ

- ถึงแม้ Alternative hypothesis จะเป็นจริงในประชากร การทดสอบสมมติฐานอาจไม่ถึงระดับนัยสำคัญเสมอไป
- เป็นปรากฏการณ์ที่เกิดจากความผิดพลาดจากการสุ่ม (Sampling error)
- โอกาสที่จะสุ่มมาเจอผลถึงระดับนัยสำคัญทางสถิติ จากประชากรที่ Alternative hypothesis เป็นจริง จะเรียกว่า กำลังในการทดสอบทางสถิติ (Statistical Power)
- อาจเรียกสั้นๆ ได้ว่า กำลัง (Power)

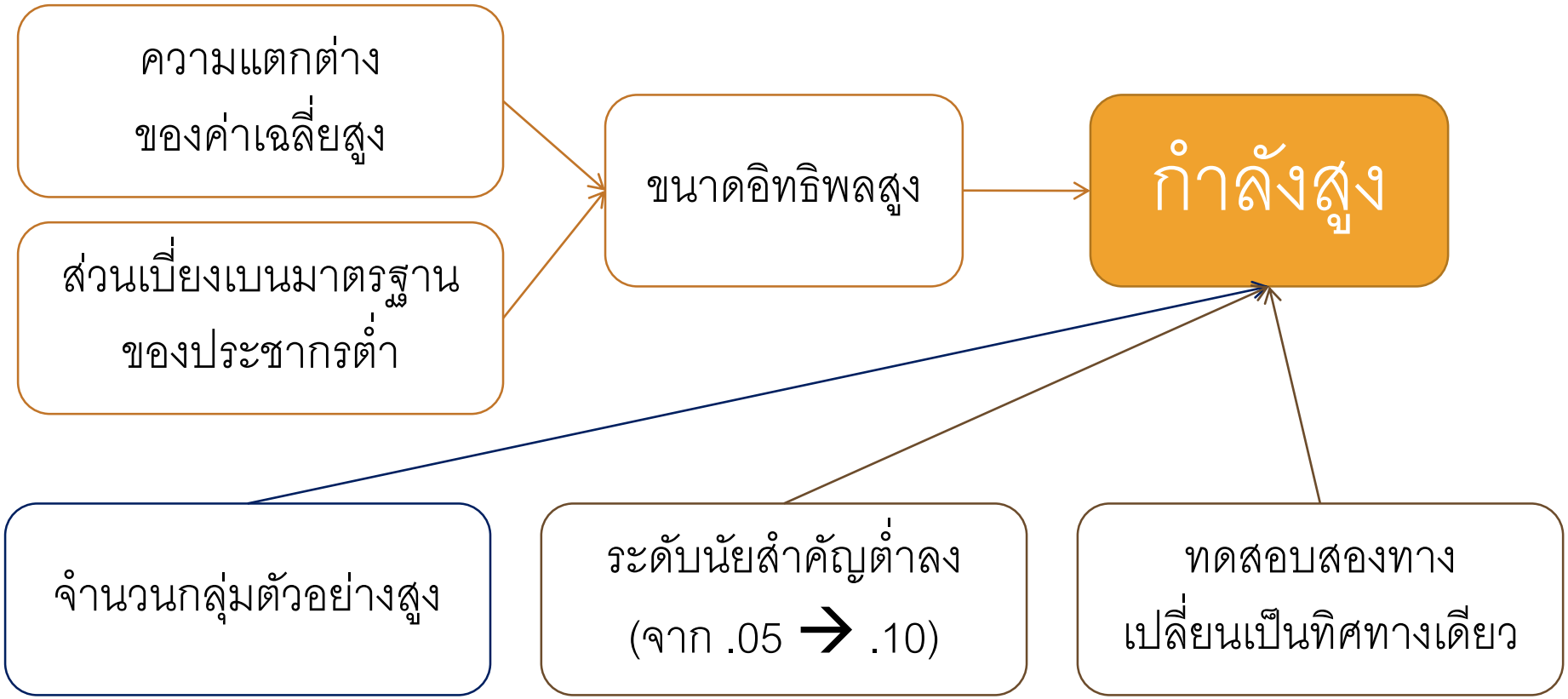
กำลังทางสถิติ

- ผลของการตัดสินใจจากการทดสอบทางสถิติ



กำลังทางสถิติ

- ปัจจัยที่มีผลต่อกำลังทางสถิติ



การกำหนดขนาดกลุ่มตัวอย่าง

- ก่อนการทดสอบทางสถิติ ควรเก็บข้อมูลให้มีขนาดมากเพียงพอที่จะทำให้โอกาสในการปฏิเสธ Null Hypothesis สูง
- กล่าวคือ มีกำลังในการปฏิเสธสมมติฐานสูง เช่น กำลังมีสูงเท่ากับ .80 หรือ .90
- จากความรู้เรื่องการหากำลัง จะทำให้หาขนาดกลุ่มตัวอย่างที่สามารถทำให้ได้กำลังขนาดที่ต้องการได้ เมื่อทราบความแตกต่างของค่าเฉลี่ย หรือขนาดอิทธิพล ว่ามีขนาดเท่าไร

การกำหนดขนาดกลุ่มตัวอย่าง

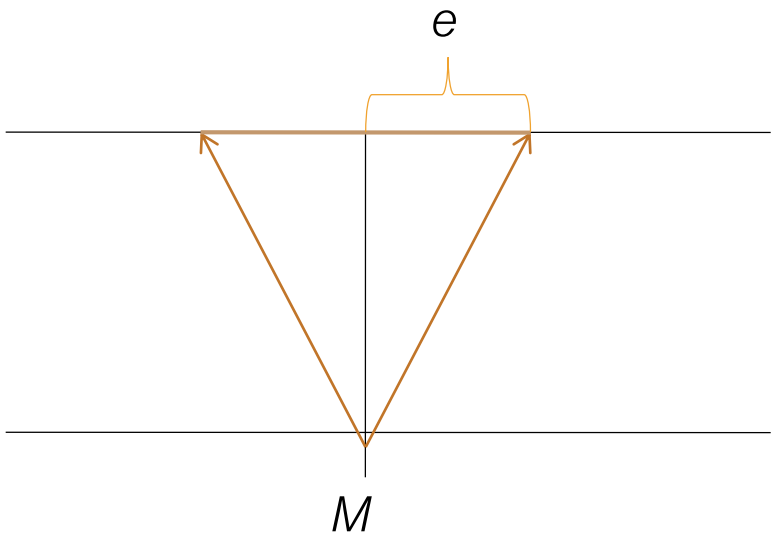
- ในการทดสอบ z สูตรลัดอย่างง่ายในการคำนวณขนาดกลุ่มตัวอย่าง เมื่อกำหนดกำลังที่ต้องการ คือ

$$n = \left(\frac{\sigma(z_{H1} - z_{H0})}{\mu_{H0} - \mu_{H1}} \right)^2$$

การกำหนดขนาดกลุ่มตัวอย่าง

- อีกแนวทางหนึ่งของการกำหนดขนาดของกลุ่มตัวอย่าง คือ การกำหนดความคาดเคลื่อน (Error; *e*)ในการประมาณค่าพารามิเตอร์

$$n = \left(\frac{z_{\alpha/2}\sigma}{e} \right)^2$$



- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “ประเด็นเพิ่มเติมสำหรับการทดสอบสมมติฐาน”

การทดสอบ *t* แบบกลุ่มตัวอย่างเดียว

- การทดสอบ t แบบกลุ่มตัวอย่างเดียว (One-sample *t*-test) ถูกสร้างขึ้นมาเพื่อนำค่าเฉลี่ยจากกลุ่มตัวอย่าง เปรียบเทียบกับค่าเฉลี่ยในประชากร
- การทดสอบนี้เหมือนกับ One-sample *z*-test เพียงแต่ว่าจะใช้ส่วนเบี่ยงเบนมาตรฐานจากกลุ่มตัวอย่าง แทนที่จะเป็นส่วนเบี่ยงเบนมาตรฐานจากประชากร
- สามารถใช้ SPSS ในการทดสอบทางสถิติและการหาช่วงเชื่อมั่น

การทดสอบ *t* แบบกลุ่มตัวอย่างเดียว

- การหาขนาดอิทธิพล ทำเหมือนกับ One-sample *z*-test

$$d = \frac{M_1 - \mu_0}{SD}$$

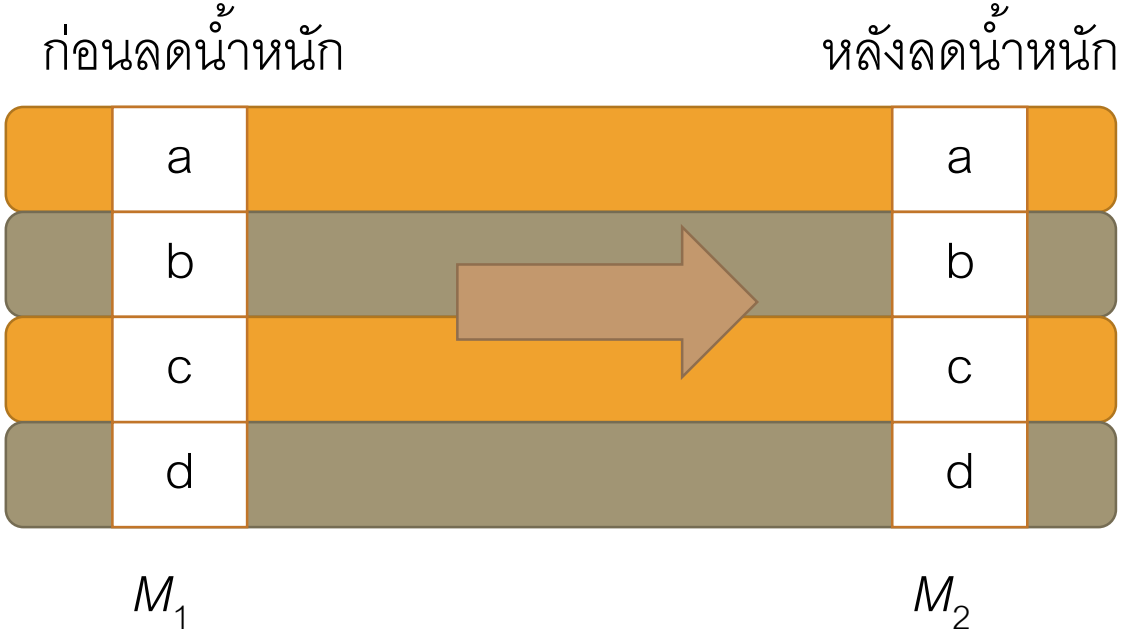
- การหากำลังและการกำหนดขนาดกลุ่มตัวอย่างจากกำลังที่กำหนด สามารถใช้โปรแกรม G*Power 3 ในการคำนวณได้
- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “การทดสอบที สำหรับกลุ่มตัวอย่างเดียว”

การเปรียบเทียบค่าเฉลี่ยระหว่าง 2 กลุ่ม

- การเปรียบเทียบค่าเฉลี่ยระหว่างสองกลุ่ม จะแบ่งการเปรียบเทียบออกเป็น 2 ประเภท คือ
  - ค่าเฉลี่ยจากกลุ่มที่เกี่ยวข้องกัน (Dependent Groups)
  - ค่าเฉลี่ยจากกลุ่มที่เป็นอิสระจากกัน (Independent Groups)
- กลุ่มที่เกี่ยวข้องกัน หมายถึง คะแนนของทุกคนในกลุ่มหนึ่ง**เกี่ยวข้อง**กับคะแนนของทุกคนในกลุ่มที่สองเป็นคู่ๆ

การเปรียบเทียบค่าเฉลี่ยระหว่าง 2 กลุ่ม

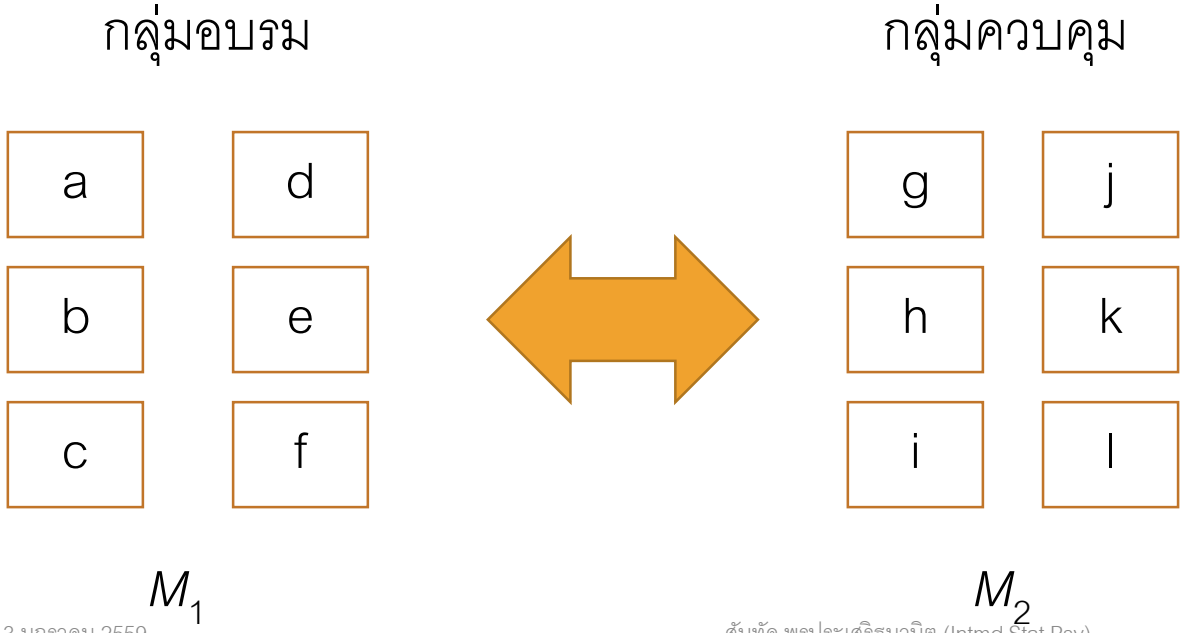
- ยกตัวอย่างเช่น การวัดก่อนลดน้ำหนักหนึ่งครั้ง แล้ววัดหลังลดอีกหนึ่งครั้ง ค่าเฉลี่ยของก่อนลด และหลังลดจะเกี่ยวข้องกัน
- เนื่องจาก ก่อนลดและหลังลด วัดจากคนเดียวกัน



ตัวแปรที่ใช้เปรียบเทียบ คือ น้ำหนัก

การเปรียบเทียบค่าเฉลี่ยระหว่าง 2 กลุ่ม

- กลุ่มที่เป็นอิสระจากกัน เช่น กลุ่มที่ผ่านการฝึกอบรม และกลุ่มควบคุม กลุ่มใดมีทักษะในการทำงานมากกว่ากัน
- คนที่อยู่ในกลุ่มที่อบรม และคนที่อยู่ในกลุ่มควบคุม จะไม่เกี่ยวข้องกัน



ตัวแปรที่ใช้เปรียบเทียบ
คือ ทักษะการทำงาน

การเปรียบเทียบค่าเฉลี่ยระหว่าง 2 กลุ่ม

- ค่าเฉลี่ยแบบเกี่ยวข้องกัน มักจะเจออยู่ใน 4 รูปแบบนี้ คือ
  - การวัดซ้ำจากคน หรือหน่วยที่ต้องการศึกษาเดียวกัน (Repeated Measure)
  - การสุ่มเข้าการทดลองเป็นแบบการจับคู่ (Matching) หรือการสุ่มจากบล็อก (Randomized Block Design)
  - การเปรียบเทียบระหว่างฝาแฝด (Twin Studies)
  - การจับคู่ตามธรรมชาติ เช่น คู่แข่งทางการค้า สามีภรรยา ญาติ
- ค่าเฉลี่ยที่เป็นอิสระจากกัน มักจะเจออยู่ใน 2 รูปแบบ คือ
  - การสุ่มมาจากคนละประชากร
  - การแบ่งกลุ่มโดยผู้วิจัย

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

- การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน จะใช้การทดสอบ t แบบกลุ่มอิสระจากกัน (Independent *t*-test)
- การทดสอบสมมติฐาน การหาช่วงเชื่อมั่น สามารถใช้ SPSS หาได้
- ข้อตกลงเบื้องต้นก่อนการใช้สถิติที่สำคัญ คือ กลุ่มทั้งสองกลุ่มมีความแปรปรวนเท่ากัน (Homogeneity of Variance)
  - หากจำนวนกลุ่มตัวอย่างเท่ากันหรือใกล้เคียงกัน t-test แบบปกติจะไม่ได้รับผลกระทบจากความแปรปรวนที่แตกต่างกัน
  - ในกรณีกลุ่มตัวอย่างแต่ละกลุ่มแตกต่างกันมาก หากอัตราส่วนของความแปรปรวนระหว่างสองกลุ่มมากกว่า 10 ต่อ 1 ค่า *p* ที่ได้จะผิดไปจากความเป็นจริง

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

- แก้ไขโดยใช้ Welch Test (บรรทัดที่ 2: Equal Variance Not Assumed)
- แต่ทว่า Welch Test มีข้อตกลงเบื้องต้นว่าการกระจายของประชากรต้องเป็นโค้งปกติ หากไม่เป็นจะส่งผลกระทบอย่างร้ายแรง
- การใช้ Levene test ในการตัดสินว่า จะใช้ *t*-test แบบปกติ หรือ Welch test เป็นวิธีการที่ไม่ถูกต้อง

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

- ขนาดอิทธิพล

$$d = \frac{M_1 - M_2}{SD_{pooled}}$$

- ส่วนเบี่ยงเบนมาตรฐานจะใช้ ส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุม (SD_C) หรือส่วนเบี่ยงเบนมาตรฐานรวม (SD_{Pooled})

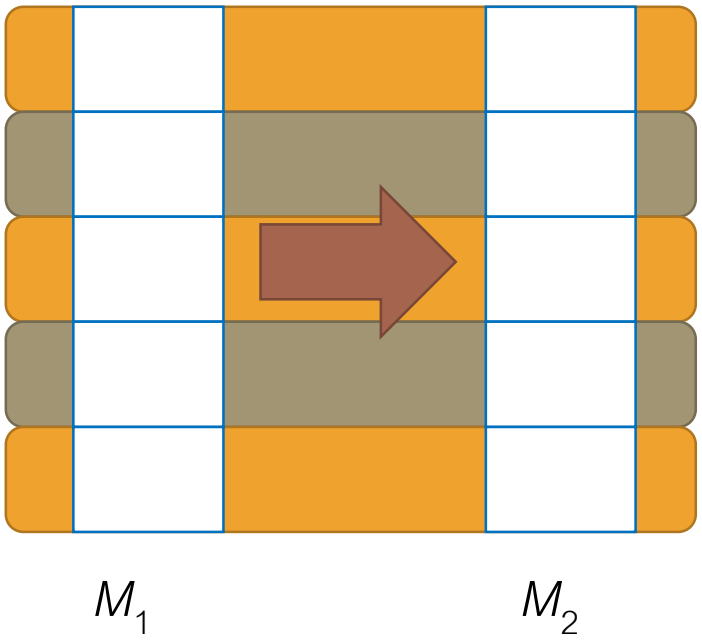
$$SD_{\text{Pooled}} = \sqrt{\frac{SS_1 + SS_2}{df_1 + df_2}}; SS_k = df_k(SD_k^2)$$

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

- การหากำลัง และการกำหนดขนาดกลุ่มตัวอย่างจากกำลัง สามารถใช้โปรแกรม G*Power 3 ในการหาได้
- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “การเปรียบเทียบค่าเฉลี่ยสำหรับกลุ่มตัวอย่าง 2 กลุ่ม”

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้องกัน

- ในการทดสอบ จะมีการแปลงข้อมูลก่อนใช้ในการทดสอบ โดยหาค่าเปลี่ยนแปลง (Difference score) ของคะแนนแต่ละคู่ เช่น



d_1

d_2

d_3

d_4

d_5

M_d

ค่าเปลี่ยนแปลง (d) = หลัง – ก่อน
(จะใช้เป็นก่อนลบหลังก็ได้)

สามารถใช้ One-sample *t*-test
ในการทดสอบว่า M_d แตกต่างจาก 0 หรือไม่

$M_d = M_2 - M_1$

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้องกัน

- การทดสอบทางสถิติแบบนี้จะเรียกว่า การทดสอบ t แบบค่าเฉลี่ยเกี่ยวข้องกัน (Dependent t-test)
- การทดสอบสมมติฐาน การหาช่วงเชื่อมั่น สามารถใช้ SPSS หาได้
- ขนาดอิทธิพล จะซับซ้อน เนื่องจากว่าส่วนเบี่ยงเบนมาตรฐานที่ใช้ในสูตรสามารถใช้ได้ 3 รูปแบบด้วยกัน
    1. ใช้ส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุม หรือกลุ่มที่ไม่ได้รับการแทรกแซง (SD_C)
    2. ใช้ส่วนเบี่ยงเบนมาตรฐานรวมของคะแนนทั้งสองชุด (SD_{Pooled})
    3. ใช้ส่วนเบี่ยงเบนมาตรฐานของค่าเปลี่ยนแปลง (SD_d)

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้องกัน

- ผลที่ได้จากการใช้ส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุม (SD_C) และส่วนเบี่ยงเบนมาตรฐานรวม (SD_{Pooled}) จะมีขนาดอิทธิพลใกล้เคียงกัน
- สองตัวนี้จะบอกว่า ขนาดความแตกต่างนั้นเป็นกี่เท่าของการกระจายของคะแนนปกติ
- เช่น การลดน้ำหนัก ทำให้ลดน้ำหนักได้เป็นกี่เท่าของส่วนเบี่ยงเบนมาตรฐานของน้ำหนักเดิม
- ขนาดอิทธิพลนี้ เหมือนกับขนาดอิทธิพลในการเปรียบเทียบค่าเฉลี่ยที่เป็นอิสระจากกัน

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้องกัน

- แต่ การใช้ส่วนเบี่ยงเบนมาตรฐานจากการเปลี่ยนแปลงนั้น จะบอกว่า ขนาดความแตกต่างนั้น เป็นกี่เท่าของการกระจายคะแนนเปลี่ยนแปลง (Difference Score)
- แต่ทว่า การแปลความหมายโดยใช้ส่วนเบี่ยงเบนมาตรฐานของค่าเปลี่ยนแปลงนั้น ทำค่อนข้างยาก
- ส่วนใหญ่จะใช้ในการหากำลังในการทดสอบทางสถิติเท่านั้น
- เช่นเดิม การหากำลัง และการกำหนดขนาดกลุ่มตัวอย่างจากกำลังสามารถใช้โปรแกรม G*Power 3 ได้

การเปรียบเทียบการทดสอบทีทั้งสองแบบ

- การเก็บข้อมูลแบบวัดซ้ำมีข้อดี คือ
  - ใช้จำนวนกลุ่มตัวอย่างน้อยกว่า
  - กำลังในการทดสอบสูงกว่า
- การเก็บข้อมูลแบบวัดซ้ำมีข้อเสีย คือ
  - ใช้ระยะเวลานาน
  - กลุ่มตัวอย่างอาจถอนตัวจากงานวิจัย ทำให้ข้อมูลสูญหาย (Attrition)
  - ผลจากการจำ ความเหนื่อยล้า (Carry over effect)
- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “การทดสอบค่าเฉลี่ยที่สัมพันธ์กัน”

สหสัมพันธ์

- ค่าสหสัมพันธ์ (Correlation) เป็นการทดสอบว่าตัวแปรระดับอันตรภาคขึ้นไป 2 ตัวแปร มีความแปรผันร่วมกันหรือไม่

$$r_{XY} = \frac{s_{XY}}{SD_X SD_Y}$$

- ค่าที่เป็นไปได้อยู่ในช่วง -1 ถึง 1
- เครื่องหมายบวกลบ บอกทิศทางความสัมพันธ์
- ค่าสัมบูรณ์ของค่าสหสัมพันธ์บอกขนาดความสัมพันธ์
  - .10 → น้อย .30 → ปานกลาง .50 → มาก
- ชื่อเต็มเรียกว่า Pearson's Product Moment Correlation

สหสัมพันธ์

- หากมีตัวแปรมากกว่า 2 ตัวแปร การรายงานอาจอยู่ในรูปเมทริกซ์สหสัมพันธ์ (Correlation Matrix)

	A	*B*	*C*
A	1	.50	-.30
B	.50	1	-.45
C	-.30	-.45	1

เมทริกซ์มีลักษณะสมมาตร

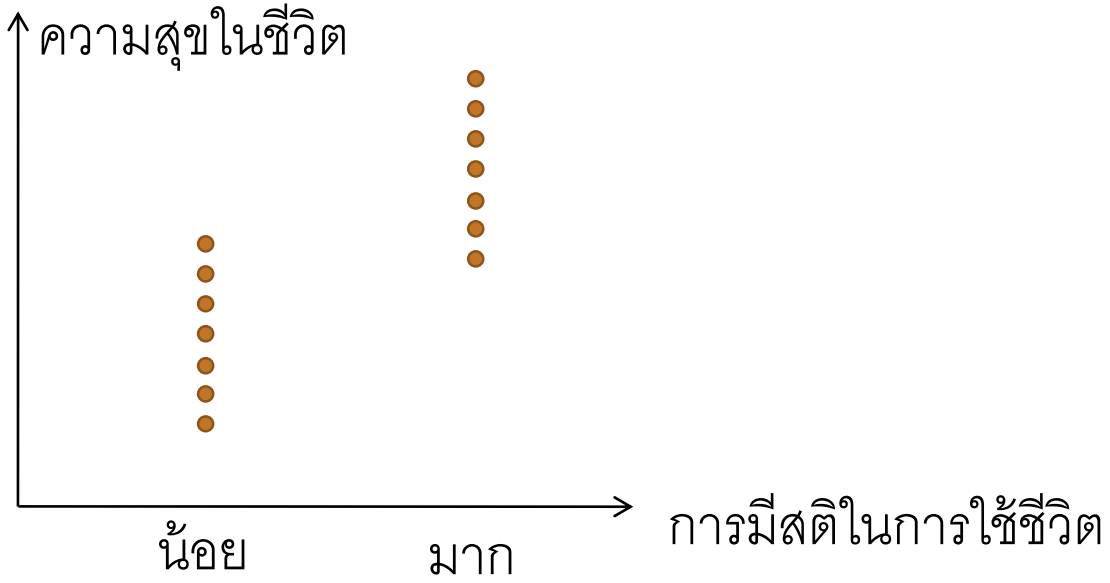
ความสัมพันธ์กับตนเองเท่ากับ 1

สหสัมพันธ์

- การทดสอบสมมติฐาน ว่าค่าสหสัมพันธ์แตกต่างจาก 0 อย่างมีนัยสำคัญหรือไม่ สามารถทดสอบด้วย SPSS
- การทดสอบสมมติฐาน ว่าค่าสหสัมพันธ์แตกต่างจากค่าอื่นที่ไม่ใช่ 0 อย่างมีนัยสำคัญหรือไม่ ต้องผ่านการหาช่วงเชื่อมั่น
- ช่วงเชื่อมั่น ไม่สามารถคำนวณผ่าน SPSS ได้ ต้องใช้วิธีการแบบ Fisher's *z* transformation
- เช่นเดิม การหากำลัง และการกำหนดขนาดกลุ่มตัวอย่างจากกำลังสามารถใช้โปรแกรม G*Power 3 ได้

ความแตกต่างและความสัมพันธ์

- ตามนิยาม ความสัมพันธ์ คือ เมื่อตัวแปรหนึ่งมีค่าเปลี่ยนแปลง อีกตัวแปรหนึ่งมีแนวโน้มจะมีค่าเปลี่ยนแปลงตาม
- เช่น การมีสติในการใช้ชีวิต (Mindfulness) มีความสัมพันธ์กับความสุขในชีวิต
- เมื่อการมีสติในการใช้ชีวิตมากขึ้น คนนั้นมีแนวโน้มจะมีความสุขในชีวิตมากขึ้น



ความแตกต่างและความสัมพันธ์

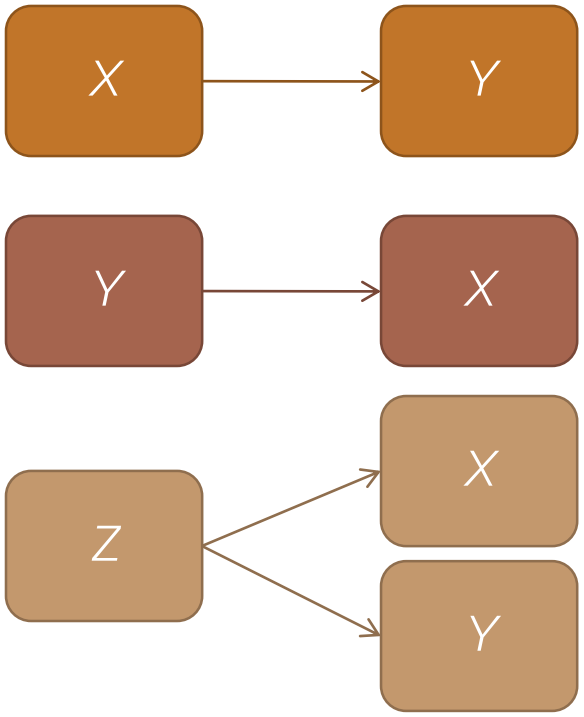
- จากนิยามนี้ ความแตกต่างก็คือความสัมพันธ์
- เช่น คนที่ประสบภัยพิบัติจะมีอาการออกจากความเป็นจริง (Derealization) มากกว่า (หรือแตกต่าง) จากคนที่ไม่ได้ประสบ
- ในที่นี้ เราอาจกล่าวได้ว่า การประสบภัยพิบัติมีความสัมพันธ์กับอาการออกจากความเป็นจริง

ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- ความสัมพันธ์หลอก (Spurious correlation)
- การจำกัดพิสัย (Range restriction)
- ค่าสุดโต่ง (Outlier)

ความสัมพันธ์เชิงสาเหตุ

- การพบความแตกต่างระหว่างกลุ่ม หรือการพบความสัมพันธ์ไม่ได้บอกว่า ตัวแปรหนึ่งเป็นสาเหตุการเกิดของอีกตัวแปรหนึ่ง
- ถ้า *X* สัมพันธ์กับ *Y* จะบอกกระบวนการเชิงสาเหตุได้สามแบบ
  - *X* เป็นสาเหตุของ *Y*
  - *Y* เป็นสาเหตุของ *X*
  - ตัวแปรอื่น เป็นสาเหตุของทั้ง *X* และ *Y*



ความสัมพันธ์เชิงสาเหตุ

- ด้วยเหตุนี้ การที่จะบอกว่า *X* เป็นสาเหตุของ *Y* จะต้องมีคุณสมบัติ 4 ประการด้วยกัน
  - ตัวแปร *X* มีความสัมพันธ์กับตัวแปร *Y*
  - *X* ต้องมาก่อน *Y*
  - การเกิด *Y* จะต้องไม่ได้รับการอธิบายด้วยตัวแปรอื่น
  - มีทฤษฎีหรือกระบวนการอธิบายว่าทำไม *X* ถึงก่อให้เกิด *Y*
- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “สหสัมพันธ์”

ค่าสถิติที่เกี่ยวข้องกับการหาสัดส่วน

- **สัดส่วน** (Proportion) **หรือ ความเสี่ยง** (Risk) คือ อัตราการเกิดเหตุการณ์ เช่น สัดส่วนของผู้ชายจากนิสิตทั้งหมด เท่ากับ .45
- **ร้อยละ** (Percentage) คือ ร้อยละของการเกิดเหตุการณ์ หรือ สัดส่วนคูณ 100 เช่น ร้อยละของแม่ที่แท้งบุตรภายใน 2 เดือนหลังจากการปฏิสนธิเท่ากับ 20%
- **อัตราส่วน** (Ratio) คือ การเปรียบเทียบปริมาณของสองสิ่ง เช่น คนทำงานไปแล้วหนึ่งปี มีคนลาออกต่อคนที่อยู่ต่อเท่ากับ 2 ต่อ 3
- **แต้มต่อ** (Odd) คือ อัตราส่วนที่เลขตัวหลัง ต่อ 1 เท่านั้น เช่น คนเป็นโรคหัวใจและออกกำลังกาย มีแต้มต่อในการรอดชีวิตในเวลา 1 ปี เท่ากับ 3 ต่อ 1

การทดสอบไคสแควร์แบบความเหมาะสมของข้อมูล

- สถิติไคสแควร์แบบความเหมาะสมของข้อมูล (Chi-square: Goodness-of-fit test) เป็นการเปรียบเทียบสัดส่วน (Proportion) ของข้อมูลกับประชากร
  - เช่น ส่วนแบ่งทางการตลาดของรถยนต์ส่วนบุคคล แตกต่างจากเมื่อปีที่ผ่านมาหรือไม่ หรือลูกเต๋าออกหน้าทุกหน้าเท่าเทียมกันหรือไม่
- การทดสอบทางสถิติ สามารถใช้ SPSS วิเคราะห์ได้โดยตรง
- การหาช่วงเชื่อมั่น สามารถทำได้ผ่านการทดสอบทวินาม (Binomial test)

การทดสอบไคสแควร์แบบความเหมาะสมของข้อมูล

- ขนาดอิทธิพล สามารถวิเคราะห์ได้หลายรูปแบบ
  - ความแตกต่างระหว่างสัดส่วน (Proportion difference) หรือความแตกต่างของความเสี่ยง (Risk difference)
  - อัตราส่วนของความเสี่ยง (Risk ratio)
  - อัตราส่วนของแต้มต่อ (Odd ratio)
  - Cohen's *w*
- การหากำลัง และการกำหนดขนาดกลุ่มตัวอย่างผ่านกำลัง สามารถใช้โปรแกรม G*Power 3 ได้

การทดสอบไคสแควร์แบบความเหมาะสมของข้อมูล

- การกำหนดขนาดกลุ่มตัวอย่าง สามารถคำนวณได้ผ่านความคาดเคลื่อนในการประมาณค่าสัดส่วนในประชากร
- สูตรของ Yamane ที่นิยมใช้กันอยู่ในงานวิจัยในไทย ก็คือการกำหนดขนาดกลุ่มตัวอย่างผ่านการประมาณค่าสัดส่วนในประชากร
- ส่วนใหญ่แล้ว นักวิจัยไม่ได้ประมาณค่าสัดส่วนในประชากร แต่ก็ใช้สูตรของ Yamane อยู่ดี

การทดสอบไคสแควร์แบบตารางไขว้

- การทดสอบไคสแควร์สามารถใช้แปลงเพื่อใช้เปรียบเทียบตัวแปรประเภทสัดส่วนระหว่างกลุ่มได้
- คล้ายกับการเปรียบเทียบค่า *t* แบบอิสระ (Independent *t*-test) แต่ตัวแปรที่เปรียบเทียบเป็นตัวแปรแบบจัดกลุ่ม (Categorical variables)
  - เช่น ผู้ชายหรือผู้หญิงชอบกินผัดไทยมากกว่ากัน
  - คนอายุต่างๆ ชอบสีต่างๆ แตกต่างกันหรือไม่
- การทดสอบสมมติฐาน สามารถใช้โปรแกรม SPSS วิเคราะห์โดยตรงได้
- การหาช่วงเชื่อมั่น จะไม่สามารถทดสอบผ่าน SPSS ได้ ต้องคำนวณมือ

การทดสอบไคสแควร์แบบตารางไขว้

- เช่นเดิม ขนาดอิทธิพล สามารถวิเคราะห์ได้ผ่าน
  - ความแตกต่างระหว่างสัดส่วน (Proportion difference) หรือความแตกต่างของความเสี่ยง (Risk difference)
  - อัตราส่วนของความเสี่ยง (Risk ratio)
  - อัตราส่วนของแต้มต่อ (Odd ratio)
  - Cohen's *w*
- การหากำลัง และการกำหนดขนาดกลุ่มตัวอย่างผ่านกำลัง สามารถใช้โปรแกรม G*Power 3 ได้

การทดสอบของ McNemar

- การทดสอบ McNemar เป็นการแปลงการทดสอบไคสแควร์เพื่อเปรียบเทียบตัวแปรแบบจัดกลุ่ม ระหว่างกลุ่มที่เกี่ยวข้องกัน คล้ายกับ Dependent *t*-test
  - เช่น คู่สามีภรรยาเลือกที่จะกินเจในช่วงเทศกาลแตกต่างกันหรือไม่
  - คนในพื้นที่เลือกนักการเมืองคนหนึ่งแตกต่างกันระหว่างช่วงก่อนการเลือกตั้ง 1 เดือนและก่อนการเลือกตั้ง 2 วัน
- การทดสอบ McNemar ใช้ได้เฉพาะการเปรียบเทียบระหว่าง 2 กลุ่ม เมื่อตัวแปรแบบจัดกลุ่มสามารถแบ่งได้แค่ 2 ประเภท
- การทดสอบสมมติฐานทางสถิติ สามารถวิเคราะห์ผ่าน SPSS ได้โดยตรง

การทดสอบของ McNemar

- การหาช่วงเชื่อมั่น ต้องใช้การคำนวณมือผ่านการทดสอบทวินาม
- การหาขนาดอิทธิพล เหมือนกับสถิติไคสแควร์ที่ผ่านมา
- การหากำลัง และการกำหนดขนาดกลุ่มตัวอย่างผ่านกำลัง สามารถใช้โปรแกรม G*Power 3 ได้
- หากเนื้อหาที่ผ่านมา คุณไม่ค่อยแม่น ให้คุณเข้าไปที่วีดีโอเรื่อง “การทดสอบไคสแควร์”

เทคนิคการสุ่มซ้ำ

สถิติสำหรับจิตวิทยา 1

สันทัด พรประเสริฐมานิต

เทคนิคการสุ่มซ้ำ

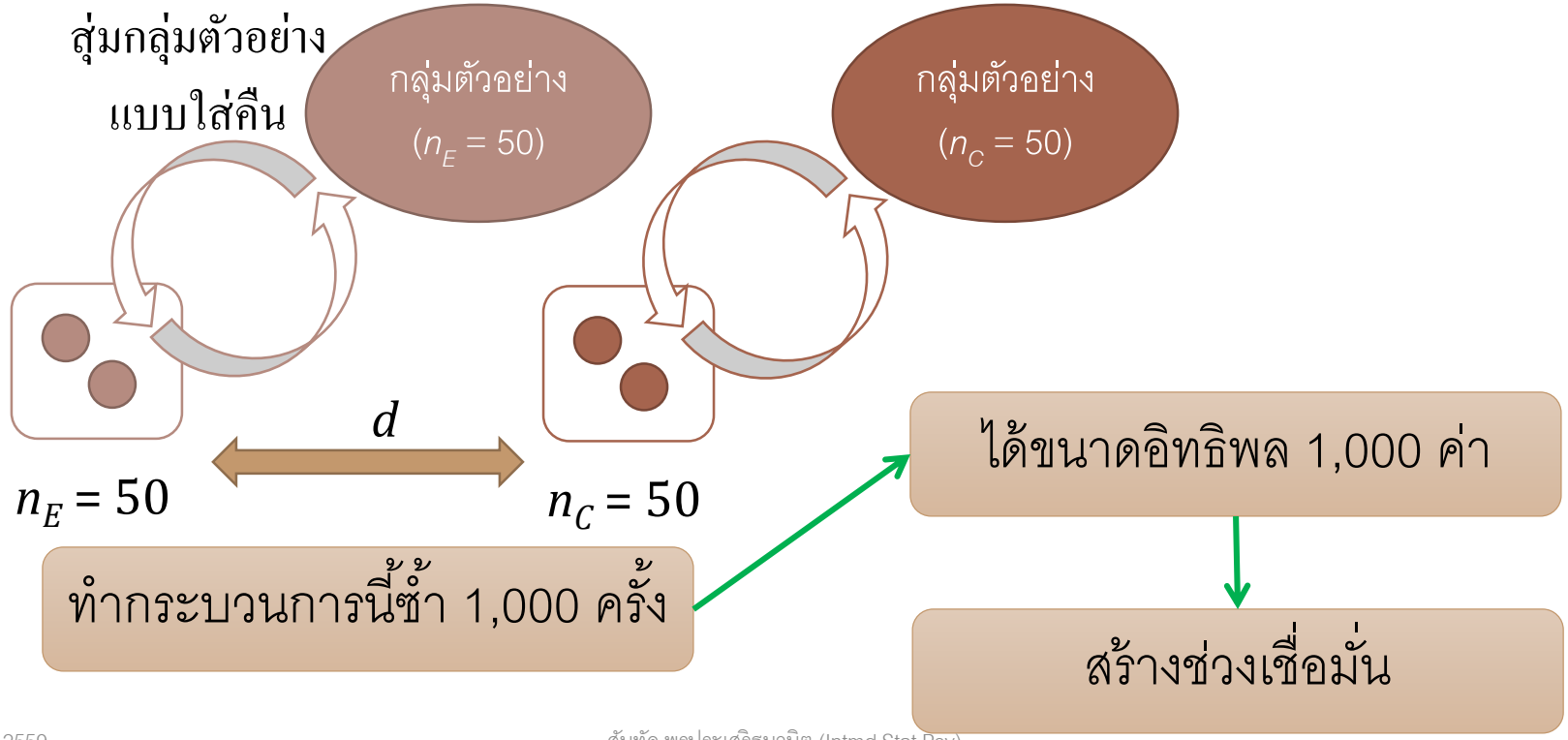
- การทดสอบสมมติฐาน หรือการสร้างช่วงเชื่อมั่นด้วยการกระจายมาตรฐาน (เช่น z, t, χ^2) ไม่ตรงกับความเป็นจริง เมื่อละเมิดข้อตกลงเบื้องต้นบางข้อ
- ค่าสถิติบางชนิด ไม่สามารถสร้างช่วงเชื่อมั่นได้ (เช่น r) เนื่องจากว่า ไม่มีการกระจายมาตรฐานที่เหมาะสม
- เทคนิคการสุ่มซ้ำ (Bootstrapping techniques) เป็นเทคนิคที่ใช้ในการประมาณค่ารูปแบบหนึ่ง โดยไม่ต้องอาศัยการกระจายมาตรฐาน

เทคนิคการสุ่มซ้ำ

- หลักการของเทคนิคการสุ่มซ้ำ คือ การนำกลุ่มตัวอย่างที่ได้นั้น เปรียบเป็นประชากรสมมติ
- หลังจากนั้น สุ่มกลุ่มตัวอย่างแบบใส่คืน (Sampling with Replacement) ตามจำนวนกลุ่มตัวอย่างที่มี
- หาค่าสถิติที่ต้องการ เช่น ค่าเฉลี่ย
- หลังจากนั้น ทวนกระบวนการที่ 2 และ 3 ใหม่ หลายๆ รอบ (เช่น 1,000 รอบขึ้นไป)
- นำค่าสถิติที่ได้ มาสร้างช่วงเชื่อมั่น เช่น หาเปอร์เซนต์ไทล์ที่ 2.5 และ 97.5 จากค่าสถิติจากกลุ่มตัวอย่างทั้งหมด เพื่อหาขอบบน ขอบล่าง

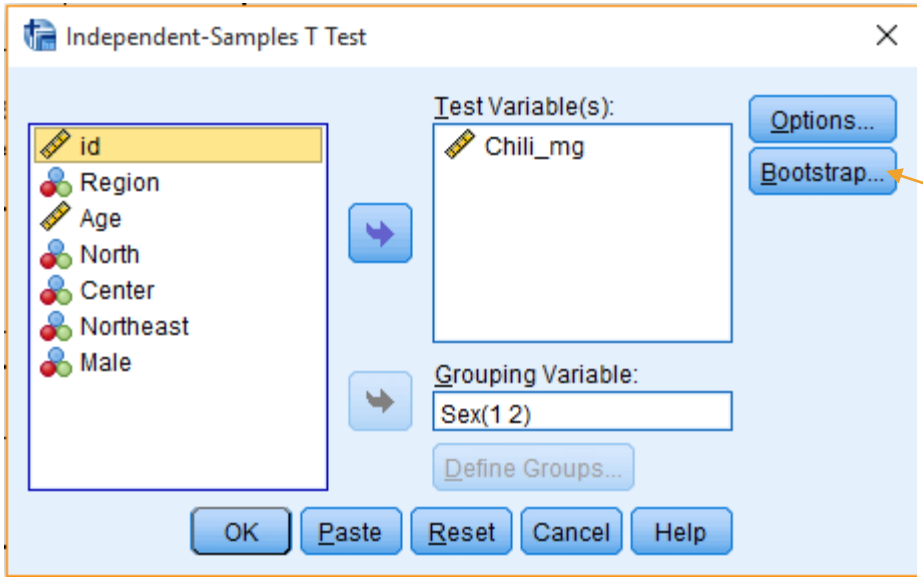
เทคนิคการสุ่มซ้ำ

- เทคนิคการสุ่มซ้ำ (Bootstrapping techniques) เช่น การหาช่วงเชื่อมั่น ของค่า *d* ใน Independent *t*-test



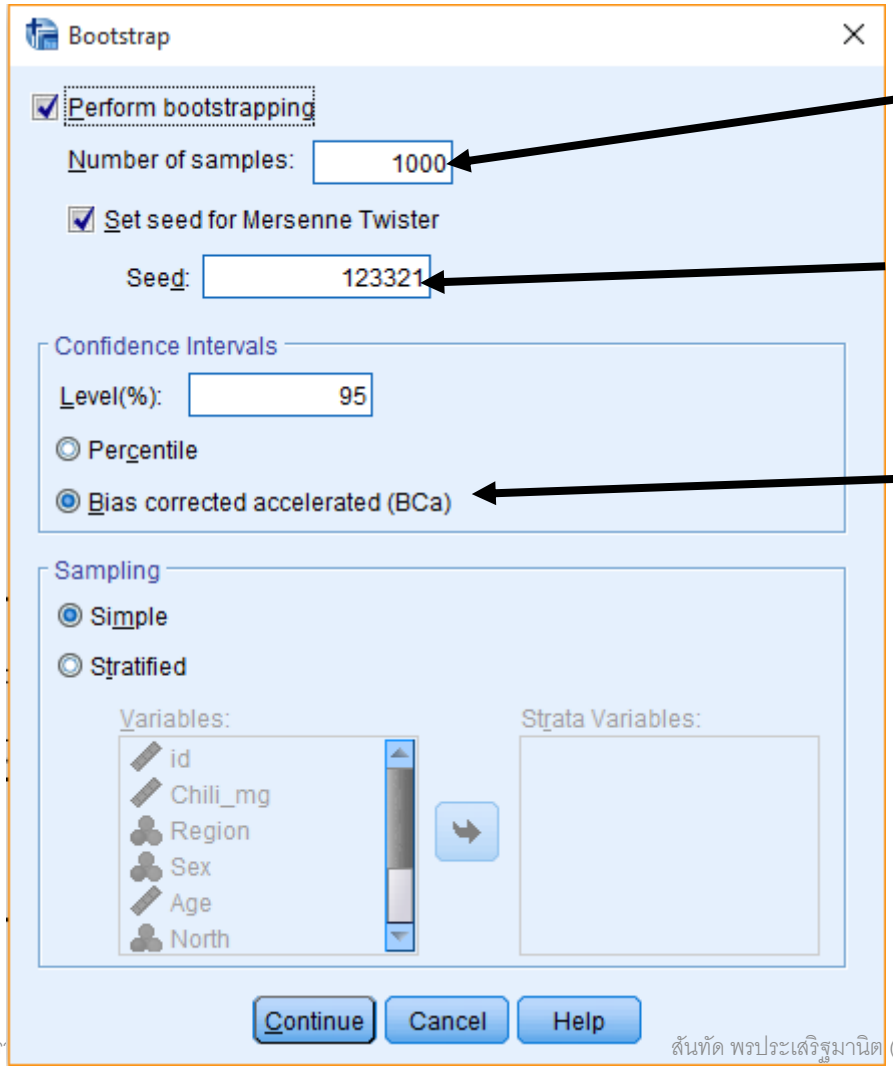
เทคนิคการสุ่มซ้ำ

- Bootstrap สามารถใช้ SPSS (v. 23) ทำได้ แต่ไม่สามารถทำได้กับค่าสถิติทุกชนิด เช่น R^2, ΔR^2
- ตัวอย่างเช่น เปรียบเทียบความแตกต่างในการใส่พริกในก๋วยเตี๋ยวระหว่างเพศ



เลือก Bootstrap

เทคนิคการสุ่มซ้ำ



จำนวนครั้งที่ใช้ในการสุ่มซ้ำ

กำหนดรหัสที่ใช้ในการสร้างตัวเลขสุ่มเสมอ เพื่อให้วิเคราะห์อีกครั้งได้ผลเหมือนเดิม

การสร้างช่วงเชื่อมั่น แนะนำให้ใช้วิธี BCa เนื่องจากช่วงเชื่อมั่นนี้จะดีกว่าเมื่อค่าสถิติทำนายค่าพารามิเตอร์ไม่แม่นยำ E(Statistics) ≠ Parameter

ช่วงเชื่อมั่นของค่าเฉลี่ย
และส่วนเบี่ยงเบนมาตรฐาน
ในแต่ละกลุ่ม

Group Statistics

	Sex		Statistic	Bootstrap[a] Bias	Bootstrap[a] Std. Error	Bootstrap[a] BCa 95% Confidence Interval Lower	Bootstrap[a] BCa 95% Confidence Interval Upper
Chili_mg	Male	N	200				
		Mean	6.9560	-.0116	.3734	6.2713	7.6789
		Std. Deviation	5.53732	-.01130	.19873	5.13136	5.87694
		Std. Error Mean	.39155				
	Female	N	200				
		Mean	7.4494	-.0128	.3608	6.7688	8.1492
		Std. Deviation	5.23180	-.02225	.24166	4.78403	5.63951
		Std. Error Mean	.36994				

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

เหมือน Independent *t*-test แบบปกติ

Independent Samples Test

		Levene's Test for Equality of Variances F	Levene's Test for Equality of Variances Sig.	t-test for Equality of Means t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper
Chili_mg	Equal variances assumed	1.826	.177	-.916	398	.360	-.49335	.53867	-1.55235	.56565
	Equal variances not assumed			-.916	396.725	.360	-.49335	.53867	-1.55236	.56566

ไม่ต้องดูผลการทดสอบทางสถิติ
เพราะช่วงเชื่อมั่น บอกผลอยู่แล้ว
ว่าแตกต่างจาก 0 อย่างมีนัยสำคัญหรือไม่

Bootstrap for Independent Samples Test

			Bootstrap[a]				
		Mean Difference	Bias	Std. Error	Sig. (2-tailed)	BCa 95% Confidence Interval	
						Lower	Upper
Chili_mg	Equal variances assumed	-.49335	.00124	.52739	.338	-1.52388	.57112
	Equal variances not assumed	-.49335	.00124	.52739	.337	-1.52388	.57112

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

ค่าเฉลี่ยที่แตกต่างกัน
ระหว่างกลุ่ม

ช่วงเชื่อมั่นด้วยวิธี BCa ซึ่งช่วงเชื่อมั่นวิธีนี้ อาจไม่สมมาตร
หากไม่คลุม 0 แสดงว่าแตกต่างจาก 0 อย่างมีนัยสำคัญ

การจัดการข้อมูลสูญหาย

สถิติสำหรับจิตวิทยา 1

สันทัด พรประเสริฐมานิต

การจัดการข้อมูลสูญหาย

- ในการเก็บข้อมูล ผู้วิจัยไม่สามารถเก็บข้อมูลทั้งหมดได้ ทำให้เป็นข้อมูลสูญหาย (Missing data)
- ข้อมูลสูญหาย สามารถเกิดได้หลายสาเหตุ
  - กลุ่มตัวอย่างไม่อยากตอบคำถามบางข้อ เช่น รายได้
  - กลุ่มตัวอย่างข้ามคำถามบางข้อโดยบังเอิญ
  - กลุ่มตัวอย่างไม่สามารถตอบคำถามบางข้อได้ เช่น เพศชายไม่สามารถตอบคำถามเกี่ยวกับการตั้งครรภ์ได้

กระบวนการเกิดข้อมูลสูญหาย

- Missing Completely at Random (MCAR) การเกิดข้อมูลสูญหาย ไม่ได้เกิดอย่างเป็นระบบ เช่น กลุ่มตัวอย่างลืมตอบคำถาม

เพศ	น้ำหนัก
ชาย	75
หญิง	45
ชาย	?
ชาย	65
หญิง	55
หญิง	50

ผู้ชายคนนี้ลืมตอบคำถามพอดี

กระบวนการเกิดข้อมูลสูญหาย

- Missing at Random (MAR) การเกิดข้อมูลสูญหาย ขึ้นอยู่กับตัวแปรหนึ่งที่ได้เก็บข้อมูลเอาไว้ หากใช้วิธีการจัดการที่ถูกต้อง จะทำให้ผลการวิเคราะห์ไม่เพี้ยนไป เช่น โอกาสในการตอบอายุ ขึ้นอยู่กับเพศ (ซึ่งได้เก็บข้อมูลเอาไว้)

เพศ	น้ำหนัก
ชาย	75
หญิง	45
ชาย	70
ชาย	65
หญิง	?
หญิง	?

ผู้หญิงมีแนวโน้มไม่ตอบน้ำหนักมากกว่า

กระบวนการเกิดข้อมูลสูญหาย

- Missing Not at Random (MNAR) การเกิดข้อมูลสูญหาย ขึ้นอยู่กับตัวแปรหนึ่งที่ไม่ได้เก็บข้อมูลเอาไว้ ทำให้ไม่สามารถจัดการได้ง่าย (ยกเว้นแต่ใช้เทคนิคบางอย่าง ซึ่งมีข้อตกลงเบื้องต้นมากมาย) เช่น โอกาสในการตอบรายได้ ขึ้นอยู่กับรายได้เอง

เพศ	น้ำหนัก
ชาย	?
หญิง	45
ชาย	70
ชาย	65
หญิง	50
หญิง	?

คนน้ำหนักเยอะ มีแนวโน้มไม่ตอบคำถามนี้

กระบวนการเกิดข้อมูลสูญหาย

- ในข้อมูลหนึ่ง ตัวแปรแต่ละตัว อาจมีกระบวนการเกิดข้อมูลสูญหายแตกต่างกัน หรือตัวแปรเดียวกัน อาจมีกระบวนการเกิดข้อมูลสูญหายหลายอย่างพร้อมกัน
- ไม่มีตัวแปรใดที่สามารถตัดสินได้ว่าเป็น MAR และ MNAR อย่างชัดเจน อาจมองได้ว่าข้อมูลสูญหายนั้น ถูกอธิบายด้วยตัวแปรอื่นที่เก็บข้อมูลมา มากน้อยเพียงใด ถ้ามากก็จะใกล้เคียง MAR ถ้าน้อยก็จะใกล้เคียง MNAR
  - เช่น ผู้หญิงที่น้ำหนักมาก มีแนวโน้มไม่ตอบน้ำหนักของตนเอง

วิธีการดั้งเดิมในการจัดการข้อมูลสูญหาย

- Listwise deletion คือ หากข้อมูลของคนใด มีข้อมูลสูญหาย จะลบข้อมูลของคนนั้นทิ้งออกไปเลย (ซึ่งเป็นวิธีการจัดการตามปกติใน SPSS หากวิเคราะห์ข้อมูลตรงๆ)

X	Y	Z
5	7	5
4	8	4
~~3~~	~~?~~	~~1~~
4	7	2
6	9	2

เช่น วิเคราะห์ถดถอย โดยใช้ข้อมูลแค่ 4 คน

ตัดข้อมูลบางส่วนทิ้ง ทั้งที่ควรจะใช้

Standard error น้อยลงกว่าความเป็นจริง

วิธีการดั้งเดิมในการจัดการข้อมูลสูญหาย

- Pairwise deletion คือ การวิเคราะห์ข้อมูลที่ใช้ข้อมูล 2 ตัวแปร จะใช้ข้อมูลคู่ที่มีทั้งหมด

X	*Y*	*Z*
5	7	5
4	8	4
3	?	1
4	7	2
6	9	2

หาความสัมพันธ์ระหว่าง

- *X* กับ *Y* ใช้ข้อมูล 4 คน
- *X* กับ *Z* ใช้ข้อมูล 5 คน
- *Y* กับ *Z* ใช้ข้อมูล 4 คน

ข้อมูลทุกคู่ไม่เท่าเทียมกัน อาจทำให้เกิดเมทริกซ์สหสัมพันธ์ที่เป็นไปไม่ได้ในทางสถิติ

วิธีการดั้งเดิมในการจัดการข้อมูลสูญหาย

- Mean imputation คือ การแทนข้อมูลสูญหายของตัวแปรนั้น ด้วยค่าเฉลี่ยของตัวแปร

X	Y	Z
5	7	5
4	8	4
3	?	1
4	7	2
6	9	2

แทนค่าด้วย (7 + 8 + 7 + 9) / 4 = 7.75

ทำให้ความแปรปรวนต่ำกว่าความเป็นจริง และความสัมพันธ์ระหว่างตัวแปรน้อยกว่าความเป็นจริง

วิธีการดั้งเดิมในการจัดการข้อมูลสูญหาย

- Regression method คือ การแทนค่าข้อมูลสูญหาย จากการทำนายผ่าน Regression ด้วยตัวแปรอื่นในข้อมูล

X	Y	Z
5	7	5
4	8	4
3	?	1
4	7	2
6	9	2

$$\hat{Y} = 5.611 + 0.583X - 0.194Z$$

$$\hat{Y} = 5.611 + 0.583(3) - 0.194(1)$$
$$= 7.17$$

ทำให้ความสัมพันธ์ระหว่างตัวแปรสูงกว่าความเป็นจริง

วิธีการดั้งเดิมในการจัดการข้อมูลสูญหาย

- Within-case mean substitution คือ การแทนค่าข้อคำถามที่สูญหาย จากคะแนนเฉลี่ยของข้อคำถามอื่นจากกลุ่มตัวอย่างเดียวกัน

X	Y_1	Y_2	Y_3	Y_4	Z
5	7	3	1	3	5
4	8	9	5	4	4
3	?	2	7	5	1
4	7	2	7	3	2
6	9	8	5	2	2

แทนค่าด้วย (2 + 7 + 5) / 3 = 4.67

ข้อตกลงเบื้องต้น คือ ค่าเฉลี่ยของแต่ละข้อคำถามเท่าเทียมกัน ซึ่งมักไม่เป็นจริง

วิธีการใหม่การจัดการข้อมูลสูญหาย

- Direct maximum likelihood หรือ Full information maximum likelihood คือ การประมาณค่าทางสถิติ ผ่านข้อมูลทั้งหมดที่มี และข้ามข้อมูลที่ไม่มี วิธีการนี้จะต้องใช้โมเดลสมการเชิงโครงสร้าง (Structural equation modeling)

X	Y	Z
5	7	5
4	8	4
3	?	1
4	7	2
6	9	2

ใช้ข้อมูลที่มีทั้งหมด ในการประมาณค่าทางสถิติ ซึ่ง 3 และ 1 ในข้อมูลของคนที่ Y สูญหายจะนำไปใช้คำนวณด้วย

วิธีการใหม่การจัดการข้อมูลสูญหาย

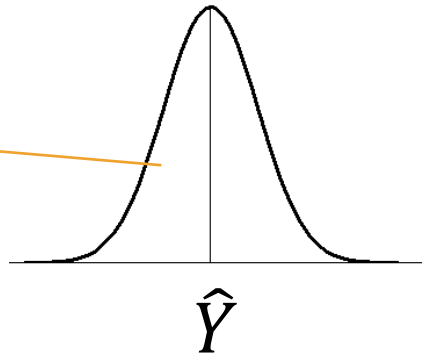
- Multiple imputation คือ การแทนค่าข้อมูลสูญหายรูปแบบหนึ่ง ซึ่งจะแทนค่าข้อมูลสูญหาย ด้วยค่าที่ทำนายได้บวกกับความผิดพลาดในการทำนาย
- เพื่อไม่ให้ผลการวิเคราะห์ได้รับอิทธิพลจากความผิดพลาดในการทำนายที่สุ่มออกมาเพียงแค่ตัวเดียว นักวิจัยจึงแทนค่าออกมาหลายครั้ง ได้ข้อมูลหลายชุด
- นำข้อมูลแต่ละชุดมาวิเคราะห์สถิติที่ต้องการ ได้ค่าสถิติที่ต้องการหลายตัว
- นำค่าสถิติที่ได้ทั้งหมด มารวมกัน และหา Standard error ที่ถูกต้อง ด้วยวิธีการของ Rubin

X	Y	Z
5	7	5
4	8	4
3	?	1
4	7	2
6	9	2

$$\hat{Y} = 5.611 + 0.583(3) - 0.194(1)$$
$$= 7.17$$

สมมติว่า $SE(\hat{Y}) = 2$

สุ่มค่าจากโค้งปกติ



X	Y	Z
5	7	5
4	8	4
3	10.02	1
4	7	2
6	9	2

X	Y	Z
5	7	5
4	8	4
3	3.48	1
4	7	2
6	9	2

X	Y	Z
5	7	5
4	8	4
3	6.15	1
4	7	2
6	9	2

X	Y	Z
5	7	5
4	8	4
3	10.02	1
4	7	2
6	9	2

$\hat{Y} = 9.8 - 0.03X - 0.53Z$

X	Y	Z
5	7	5
4	8	4
3	3.48	1
4	7	2
6	9	2

$\hat{Y} = 0.2 + 1.37X + 0.24Z$

X	Y	Z
5	7	5
4	8	4
3	6.15	1
4	7	2
6	9	2

$\hat{Y} = 4.1 + 0.8X - 0.1Z$

ข้อมูลแทนค่า	b_1	$SE(b_1)$
1	-0.028	0.631
2	1.374	0.752
3	0.801	0.425

ใช้วิธีการรวมข้อมูลของ Rubin

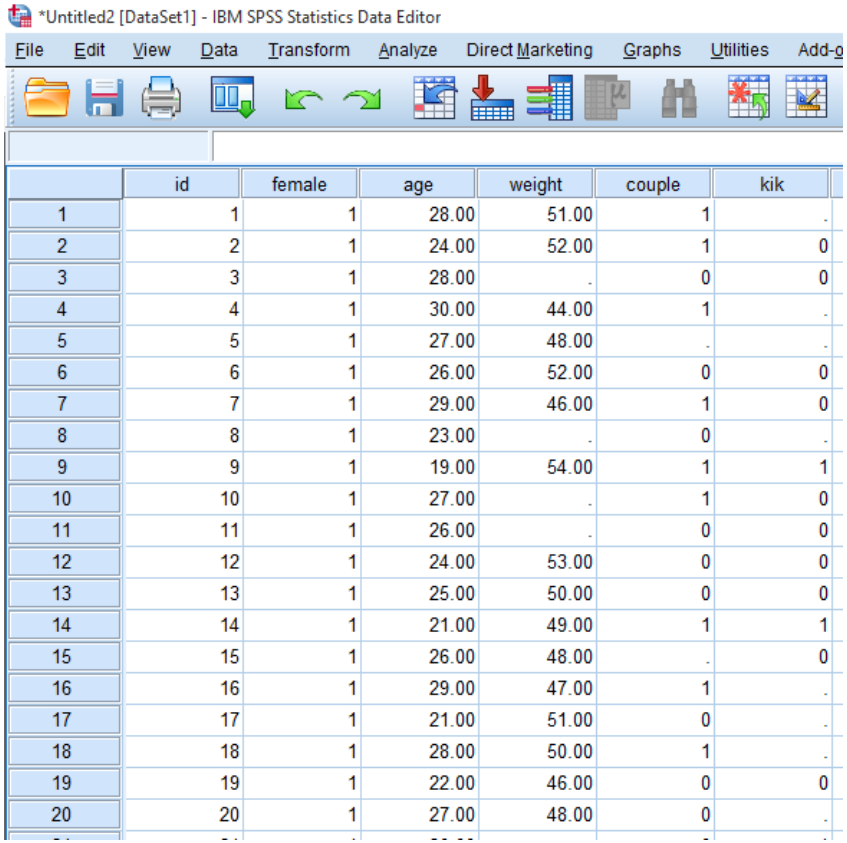
$b_1 = 0.715$

$SE(b_1) = 1.022$

วิธีการใหม่การจัดการข้อมูลสูญหาย

- ในกรณีที่ตัวแปรหนึ่ง มีข้อมูลสูญหายไม่ถึง 5% การเลือกวิธีการดั้งเดิมไม่ได้มีผลกระทบมาก แต่ถ้าใช้วิธีการใหม่ได้ ให้ใช้ดีกว่า
- แต่หากตัวแปรใด มีข้อมูลสูญหายเกิน 5% ควรใช้วิธีการใหม่
- SPSS สามารถจัดการข้อมูลสูญหายได้ด้วยวิธี Multiple imputation (MI)
- วิธี Multiple imputation ถือเป็นวิธีการเตรียมข้อมูล การนำไปวิเคราะห์จริง การแทนค่าเพียงครั้งเดียว สามารถนำไปใช้ในการวิเคราะห์หลายครั้งได้

การแทนค่าแบบพหุ



	id	female	age	weight	couple	kik
1	1	1	28.00	51.00	1	.
2	2	1	24.00	52.00	1	0
3	3	1	28.00	.	0	0
4	4	1	30.00	44.00	1	.
5	5	1	27.00	48.00	.	.
6	6	1	26.00	52.00	0	0
7	7	1	29.00	46.00	1	0
8	8	1	23.00	.	0	.
9	9	1	19.00	54.00	1	1
10	10	1	27.00	.	1	0
11	11	1	26.00	.	0	0
12	12	1	24.00	53.00	0	0
13	13	1	25.00	50.00	0	0
14	14	1	21.00	49.00	1	1
15	15	1	26.00	48.00	.	0
16	16	1	29.00	47.00	1	.
17	17	1	21.00	51.00	0	.
18	18	1	28.00	50.00	1	.
19	19	1	22.00	46.00	0	0
20	20	1	27.00	48.00	0	.

ตัวอย่างข้อมูล

Female: 1 = หญิง, 0 = ชาย

Age: อายุ

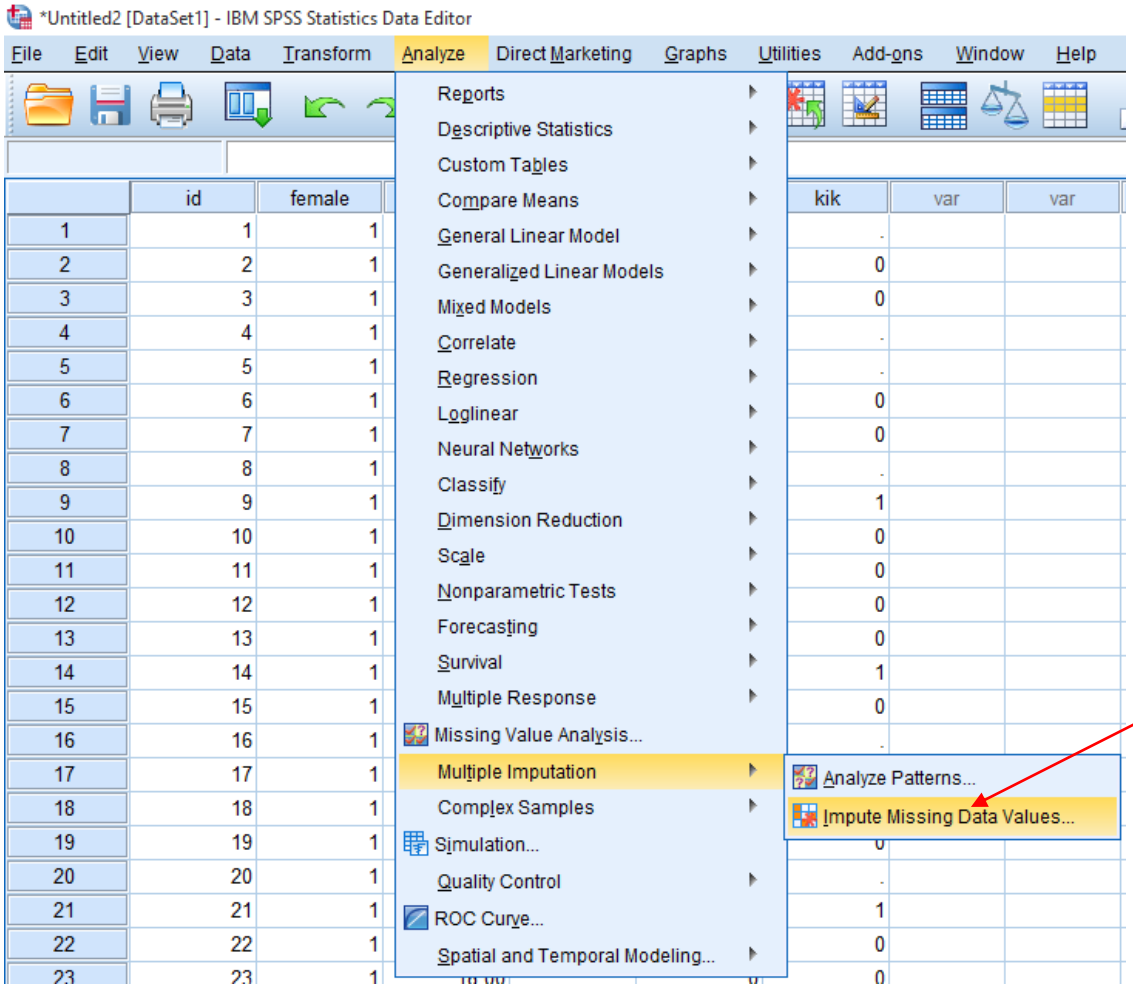
Weight: น้ำหนัก

Couple: มีแฟนอย่างเป็นทางการ

Kik: มีกิ๊ก

หมายเหตุ ตัวแปรจะต้องกำหนดระดับการวัดของแต่ละตัวแปรใน Variable View… ให้ถูกต้อง (Nominal/Ordinal/Scale)

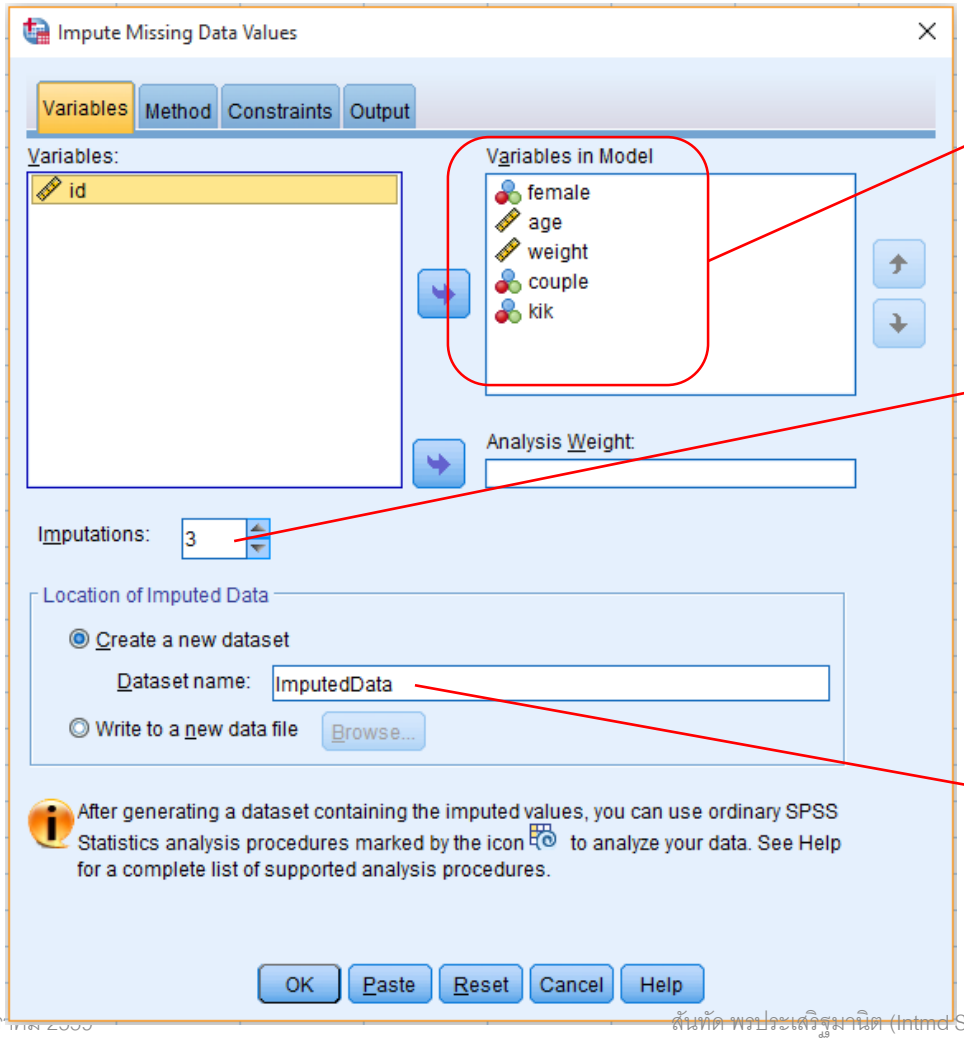
การแทนค่าแบบพหุ



เลือกเพื่อทำ

Multiple Imputation

การแทนค่าแบบพหุ

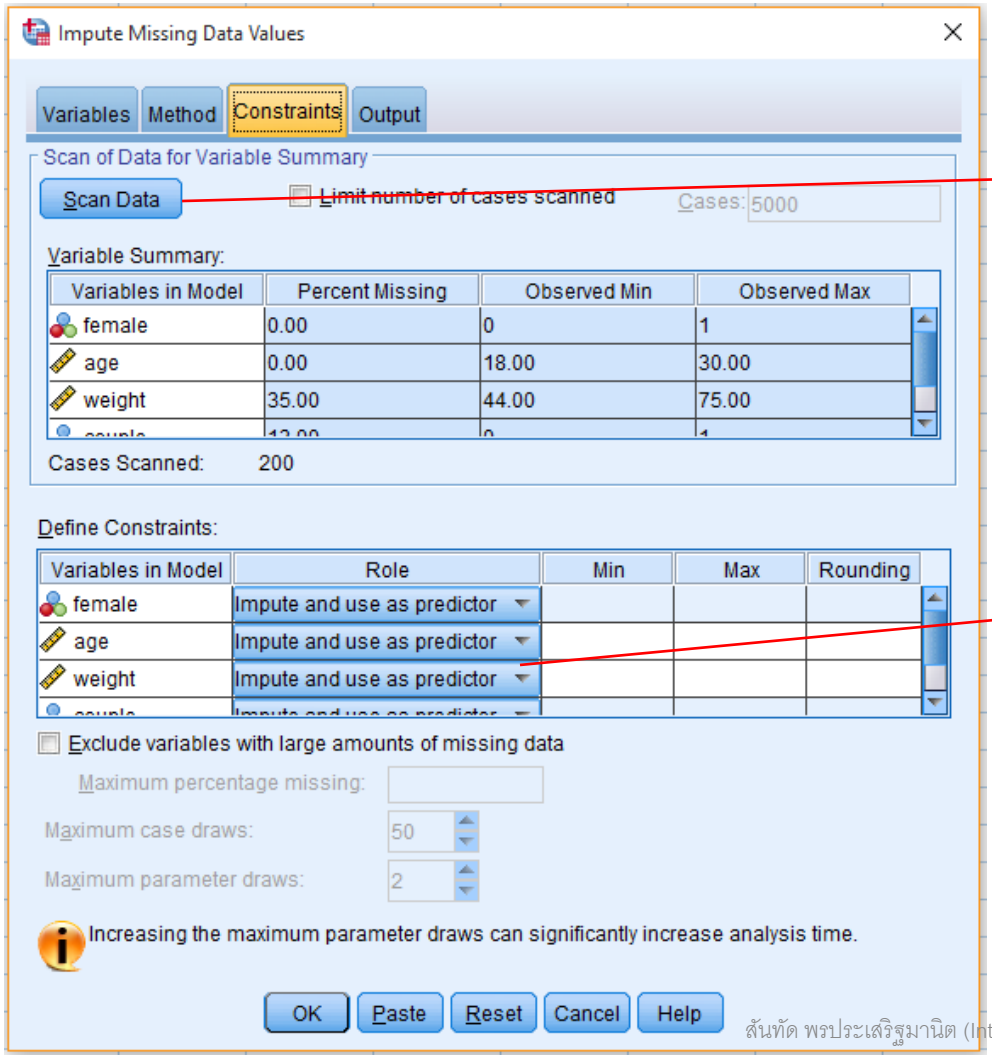


เลือกตัวแปรที่ต้องการแทนค่าสูญหายหรือตัวแปรที่ใช้ในการทำนายค่าสูญหายของตัวแปรอื่น

จำนวนครั้งในการแทนค่าสูญหาย ยิ่งเยอะยิ่งดี แต่จะเสียเวลาในการวิเคราะห์ ผมแนะนำให้ใช้ประมาณ 20 ครั้ง

ใส่ชื่อข้อมูลที่แทนค่าแล้ว

การแทนค่าแบบพหุ



สามารถกด Scan Data เพื่อตรวจสอบลักษณะของข้อมูลโดยสังเขป

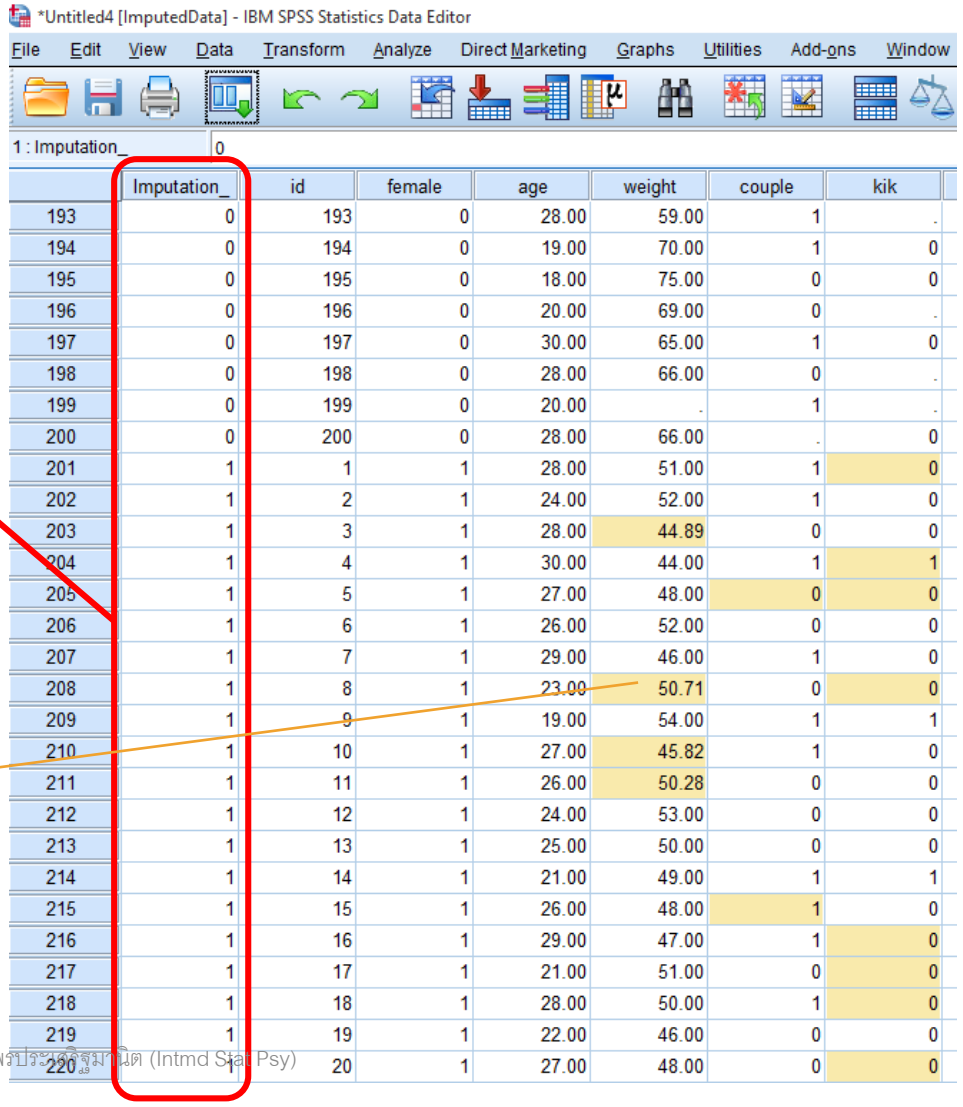
สามารถเลือกได้ว่าตัวแปรใดให้มีการแทนค่า ตัวแปรใดใช้ทำนายค่าสูญหายของตัวแปรอื่น

การแทนค่าแบบพหุ

ข้อมูลใหม่หลังจากแทนค่าแล้ว

ลำดับของครั้งที่แทนค่า
0 คือข้อมูลเดิม

สีเหลือง คือ ค่าที่โปรแกรมแทนค่าสูญหาย



	Imputation_	id	female	age	weight	couple	kik
193	0	193	0	28.00	59.00	1	.
194	0	194	0	19.00	70.00	1	0
195	0	195	0	18.00	75.00	0	0
196	0	196	0	20.00	69.00	0	.
197	0	197	0	30.00	65.00	1	0
198	0	198	0	28.00	66.00	0	.
199	0	199	0	20.00	.	1	.
200	0	200	0	28.00	66.00	.	0
201	1	1	1	28.00	51.00	1	0
202	1	2	1	24.00	52.00	1	0
203	1	3	1	28.00	44.89	0	0
204	1	4	1	30.00	44.00	1	1
205	1	5	1	27.00	48.00	0	0
206	1	6	1	26.00	52.00	0	0
207	1	7	1	29.00	46.00	1	0
208	1	8	1	23.00	50.71	0	0
209	1	9	1	19.00	54.00	1	1
210	1	10	1	27.00	45.82	1	0
211	1	11	1	26.00	50.28	0	0
212	1	12	1	24.00	53.00	0	0
213	1	13	1	25.00	50.00	0	0
214	1	14	1	21.00	49.00	1	1
215	1	15	1	26.00	48.00	1	0
216	1	16	1	29.00	47.00	1	0
217	1	17	1	21.00	51.00	0	0
218	1	18	1	28.00	50.00	1	0
219	1	19	1	22.00	46.00	0	0
220	1	20	1	27.00	48.00	0	0

การแทนค่าแบบพหุ

- จากข้อมูลนี้ ลองทำนายน้ำหนัก ด้วยเพศและอายุ โดยการทำ Multiple Regression แบบปกติ
  - Analyze → Regression → Linear
  - ใส่ Female และ Age เป็นตัวแปรอิสระ และ Weight เป็นตัวแปรตาม

การแทนค่าแบบพหุ

Model Summary

Imputation Number	Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
Original data	1	.913[a]	.834	.831	3.33576
1	1	.915[a]	.836	.835	3.32228
2	1	.911[a]	.829	.827	3.39921
3	1	.915[a]	.837	.835	3.35190

a. Predictors: (Constant), age, female

ค่า R^2 ของข้อมูลดั้งเดิม และข้อมูลที่แทนค่า

ANOVA[a]

Imputation Number	Model		Sum of Squares	df	Mean Square	F	Sig.
Original data	1	Regression	7081.360	2	3540.680	318.197	.000[b]
		Residual	1413.171	127	11.127		
		Total	8494.531	129			
1	1	Regression	11119.062	2	5559.531	503.692	.000[b]
		Residual	2174.401	197	11.038		
		Total	13293.463	199			
2	1	Regression	11047.687	2	5523.843	478.064	.000[b]
		Residual	2276.257	197	11.555		
		Total	13323.944	199			
3	1	Regression	11375.417	2	5687.709	506.240	.000[b]
		Residual	2213.335	197	11.235		
		Total	13588.752	199			

a. Dependent Variable: weight

b. Predictors: (Constant), age, female

การทดสอบ $H_0: \mathrm{P}^2 = 0$ ของข้อมูลดั้งเดิม และข้อมูลที่แทนค่า

การแทนค่าแบบพหุ

ผลการวิเคราะห์สัมประสิทธิ์ถดถอย ของข้อมูลดั้งเดิม และข้อมูลที่แทนค่า

Coefficients[a]

Imputation Number	Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Fraction Missing Info.	Relative Increase Variance	Relative Efficiency
			B	Std. Error	Beta					
Original data	1	(Constant)	67.352	1.914		35.181	.000			
		female	-14.989	.596	-.916	-25.150	.000			
		age	-.066	.075	-.032	-.884	.378			
1	1	(Constant)	68.701	1.511		45.463	.000			
		female	-14.960	.471	-.918	-31.730	.000			
		age	-.112	.060	-.054	-1.880	.062			
2	1	(Constant)	64.505	1.546		41.720	.000			
		female	-14.838	.482	-.909	-30.759	.000			
		age	.036	.061	.017	.582	.561			
3	1	(Constant)	66.314	1.525		43.495	.000			
		female	-15.096	.476	-.916	-31.735	.000			
		age	-.021	.060	-.010	-.349	.727			
Pooled	1	(Constant)	66.507	2.870		23.169	.000	.799	2.532	.790
		female	-14.965	.499		-29.971	.000	.096	.098	.969
		age	-.033	.105		-.310	.771	.759	2.036	.798

a. Dependent Variable: weight

ผลของการรวมค่าสถิติจากข้อมูลที่แทนค่าทั้งหมด

การแทนค่าแบบพหุ

- จากผลการวิเคราะห์ข้อมูล เพศมีผลต่อน้ำหนักอย่างมีนัยสำคัญ แต่อายุไม่มีผลอย่างมีนัยสำคัญ
- จะเห็นว่า การรวมผลการวิเคราะห์จะมีเฉพาะบางสถิติ เช่น สัมประสิทธิ์ถดถอย แต่สัมประสิทธิ์การทำนาย ไม่ได้ถูกรวมในที่นี้
- โปรแกรมอื่น มีความสามารถมากกว่า SPSS ในการทำ Multiple imputation เช่น Mplus, R

คาบต่อไป

- การบ้านครั้งที่ 1 (ส่งในสัปดาห์ที่ 18-22 ม.ค. 59)
- เรียนคาบปฏิบัติการครั้งที่ 1