

การเปรียบเทียบค่าเฉลี่ยระหว่าง กลุ่มตัวอย่าง 2 กลุ่ม

สถิติสำหรับจิตวิทยา 1

สันทัด พรประเสริฐมานิต

โครงร่างการนำเสนอ

- รูปแบบการเปรียบเทียบค่าเฉลี่ย
 - การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน
 - การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน
- การทดสอบสมมติฐาน
- ข้อตกลงเบื้องต้นก่อนการใช้สถิติ
- ช่วงเชื่อมั่น
- ขนาดอิทธิพล
- กำลังทดสอบทางสถิติและการคำนวณขนาดกลุ่มตัวอย่าง

แนะนำ

- ในครั้งที่ผ่านๆ มา ได้กล่าวถึงการทดสอบสมมติฐานโดยนำค่าเฉลี่ยของกลุ่มเดียว ไปเปรียบเทียบกับเกณฑ์ที่ได้ตั้งเอาไว้
- ในครั้งนี้ จะเปรียบเทียบค่าเฉลี่ยระหว่างสองกลุ่มด้วยกัน เช่น
 - เพศใดจะมีความอ่อนไหวทางอารมณ์มากกว่ากัน ซึ่งเป็นการเปรียบเทียบค่าเฉลี่ยของเพศหญิง และค่าเฉลี่ยของเพศชาย
 - กลุ่มที่ได้รับการฝึกอบรม กับกลุ่มควบคุม (ทิ้งเวลาไว้เฉยๆ) กลุ่มใดจะมีทักษะในการทำงานมากกว่ากัน
 - ก่อนลดน้ำหนัก และหลังลดน้ำหนัก จะมีค่าเฉลี่ยแตกต่างกันหรือไม่

รูปแบบการเปรียบเทียบค่าเฉลี่ย

- การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน เป็นการเปรียบเทียบค่าเฉลี่ยของสองกลุ่ม โดยที่คะแนนของแต่ละกลุ่มจะมีการจับคู่กัน
 - เช่น เปรียบเทียบคะแนนก่อนหลัง คะแนนก่อนหลังของแต่ละคนจะจับคู่กัน
- การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน เป็นการเปรียบเทียบค่าเฉลี่ยของสองกลุ่ม แต่คะแนนภายในกลุ่มของแต่ละกลุ่มจะไม่จับคู่กัน
 - เช่น เปรียบเทียบความพึงพอใจในชีวิต จากกลุ่มตัวอย่างของสองหมู่บ้าน
 - เราไม่สามารถนำชาวบ้านคนที่ 1 ของหมู่บ้าน A ไปจับคู่กับชาวบ้านคนใด ในหมู่บ้าน B ได้เลย

รูปแบบการเปรียบเทียบค่าเฉลี่ย

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

รูปแบบการเปรียบเทียบค่าเฉลี่ย

- กลุ่มที่เป็นอิสระจากกัน เช่น กลุ่มที่ผ่านการฝึกอบรม และกลุ่มควบคุม กลุ่มใดมีทักษะในการทำงานมากกว่ากัน
- คนที่อยู่ในกลุ่มที่อบรม และคนที่อยู่ในกลุ่มควบคุม จะไม่เกี่ยวข้องกัน

กลุ่มอบรม

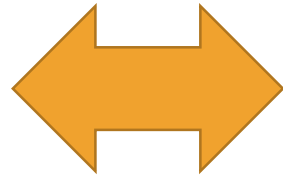
a	d
b	e
c	f

M_1

กลุ่มควบคุม

g	j
h	k
i	l

M_2



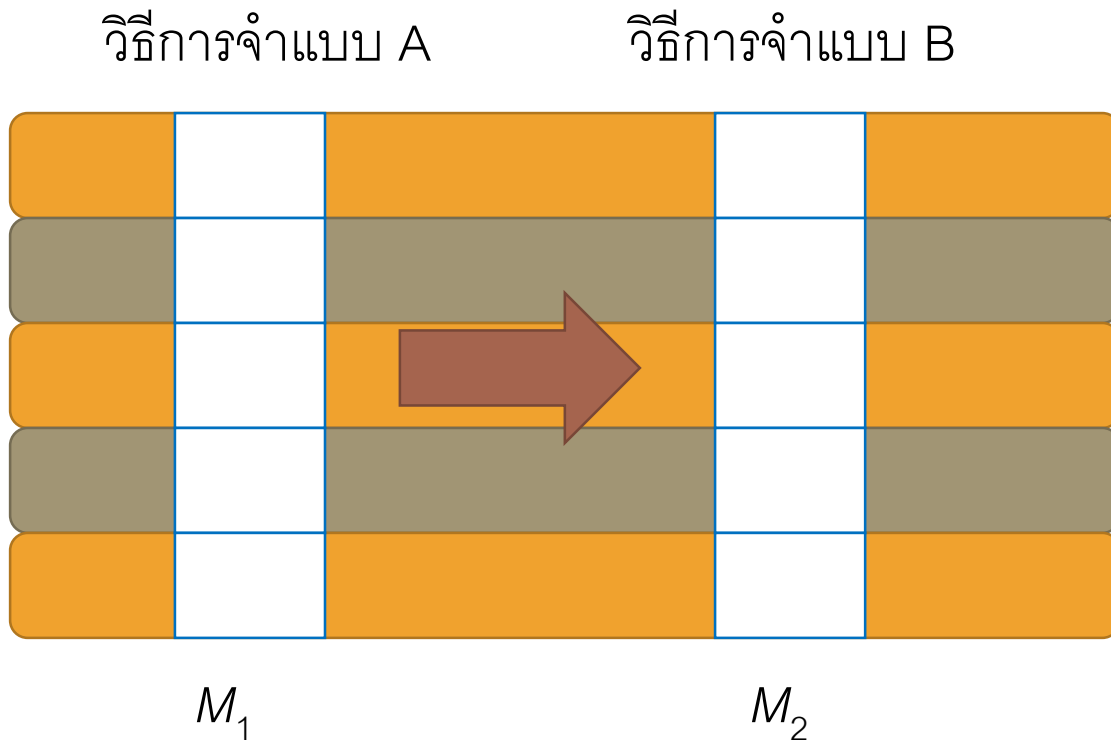
ตัวแปรที่ใช้เปรียบเทียบ
คือ ทักษะการทำงาน

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- ค่าเฉลี่ยแบบเกี่ยวข้อกัน มักจะเจออยู่ใน 4 รูปแบบนี้ คือ
 - การวัดซ้ำจากคน หรือหน่วยที่ต้องการศึกษาเดียวกัน (Repeated Measure)
 - การสุ่มเข้าการทดลองเป็นแบบการจับคู่ (Matching) หรือการสุ่มจากบล็อก (Randomized Block Design)
 - การเปรียบเทียบระหว่างฝาแฝด (Twin Studies)
 - การจับคู่ตามธรรมชาติ เช่น คู่แข่งทางการค้า สามีภรรยา ญาติ

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

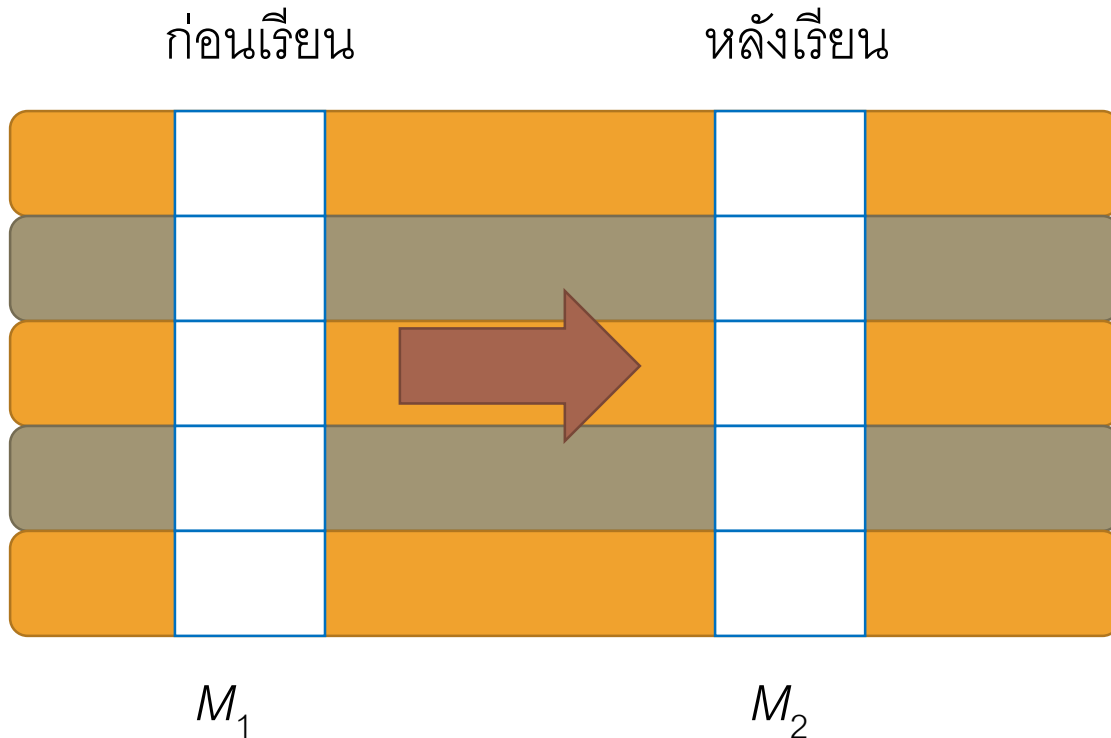
- การวัดซ้ำจากคน หรือหน่วยที่ต้องการศึกษาเดียวกัน (Repeated Measure)
เช่น การทดลอง



ตัวแปรที่ใช้เปรียบเทียบ
คือ จำนวนคำที่จำได้

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- การวัดซ้ำจากคน หรือหน่วยที่ต้องการศึกษาเดียวกัน (Repeated Measure)
เช่น การวัดก่อนหลัง



ตัวแปรที่ใช้เปรียบเทียบ
คือ คะแนนสอบ

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- การสุ่มเข้าการทดลองเป็นแบบการจับคู่ (Matching) หรือการสุ่มจากบล็อก (Randomized Block Design)
- เช่น ทดลองวิธีการสอนคอมพิวเตอร์ 2 แบบ โดยควบคุมตัวแปรความรู้ภาษาอังกฤษก่อนเข้าเรียน

อันดับแรก นำตัวแปรควบคุมมาเรียงก่อน (ในที่นี้คือภาษาอังกฤษ)

70	69	65	63	62	62	60	58	54	52
----	----	----	----	----	----	----	----	----	----

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- การสุ่มเข้าการทดลองเป็นแบบการจับคู่ (Matching) หรือการสุ่มจากบล็อก (Randomized Block Design)
- เช่น ทดลองวิธีการสอนคอมพิวเตอร์ 2 แบบ โดยควบคุมตัวแปรความรู้ภาษาอังกฤษก่อนเข้าเรียน

อันดับที่สอง จับคู่กลุ่มตัวอย่างทั้งหมด

70	69	65	63	62	62	60	58	54	52
1	2	1	2	1	2	1	2	1	2
A		B		C		D		E	

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- การสุ่มเข้าการทดลองเป็นแบบการจับคู่ (Matching) หรือการสุ่มจากบล็อก (Randomized Block Design)
- เช่น ทดลองวิธีการสอนคอมพิวเตอร์ 2 แบบ โดยควบคุมตัวแปรความรู้ภาษาอังกฤษก่อนเข้าเรียน

อันดับที่สาม นำนักเรียนภายในแต่ละคู่ สุ่มเข้าเงื่อนไขต่างๆ

70	69
1	2

A

วิธีการสอน
แบบที่ 1

วิธีการสอน
แบบที่ 2

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- การสุ่มเข้าการทดลองเป็นแบบการจับคู่ (Matching) หรือการสุ่มจากบล็อก (Randomized Block Design)
- เช่น ทดลองวิธีการสอนคอมพิวเตอร์ 2 แบบ โดยควบคุมตัวแปรความรู้ภาษาอังกฤษก่อนเข้าเรียน

อันดับที่สาม นำนักเรียนภายในแต่ละคู่ สุ่มเข้าเงื่อนไขต่างๆ

65	63
1	2

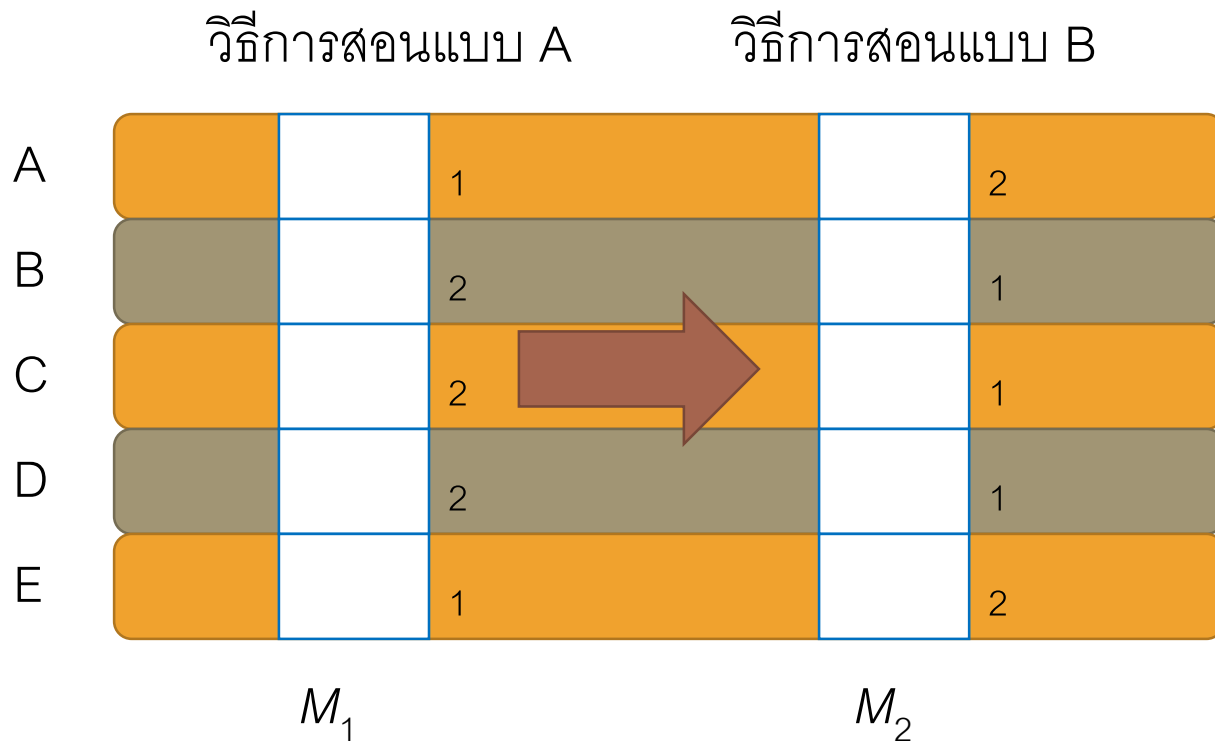
B

วิธีการสอน
แบบที่ 1

วิธีการสอน
แบบที่ 2

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

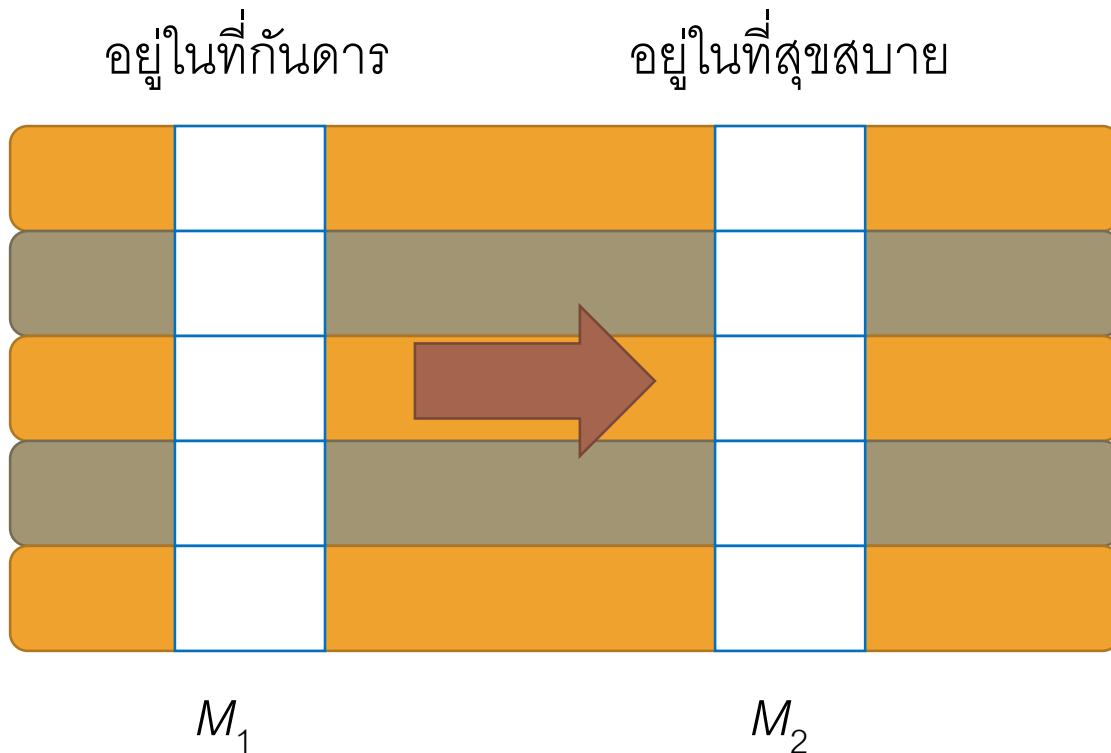
- การสุ่มเข้าการทดลองเป็นแบบการจับคู่ (Matching) หรือการสุ่มจากบล็อก (Randomized Block Design)



ตัวแปรที่ใช้เปรียบเทียบ
คือ คะแนนสอบคอมฯ

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

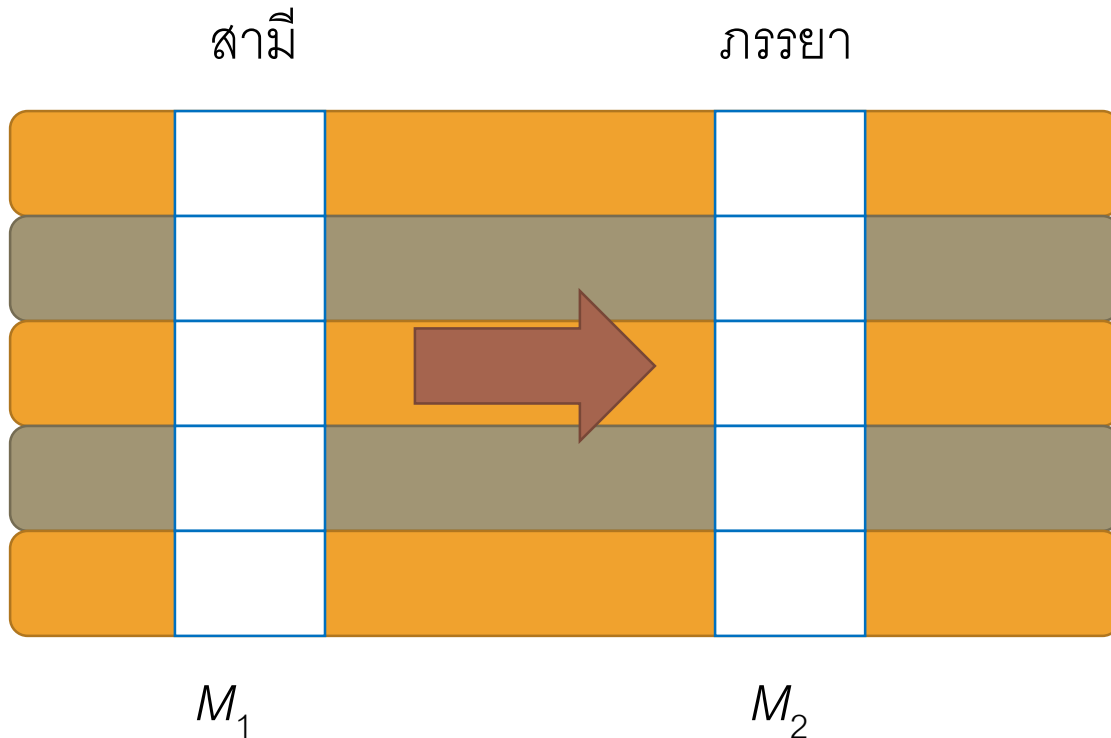
- การเปรียบเทียบระหว่างฝาแฝด (Twin Studies)



ตัวแปรที่ใช้เปรียบเทียบ
คือ IQ

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- การจับคู่ตามธรรมชาติ เช่น คู่แข่งทางการค้า สามีภรรยา ญาติ



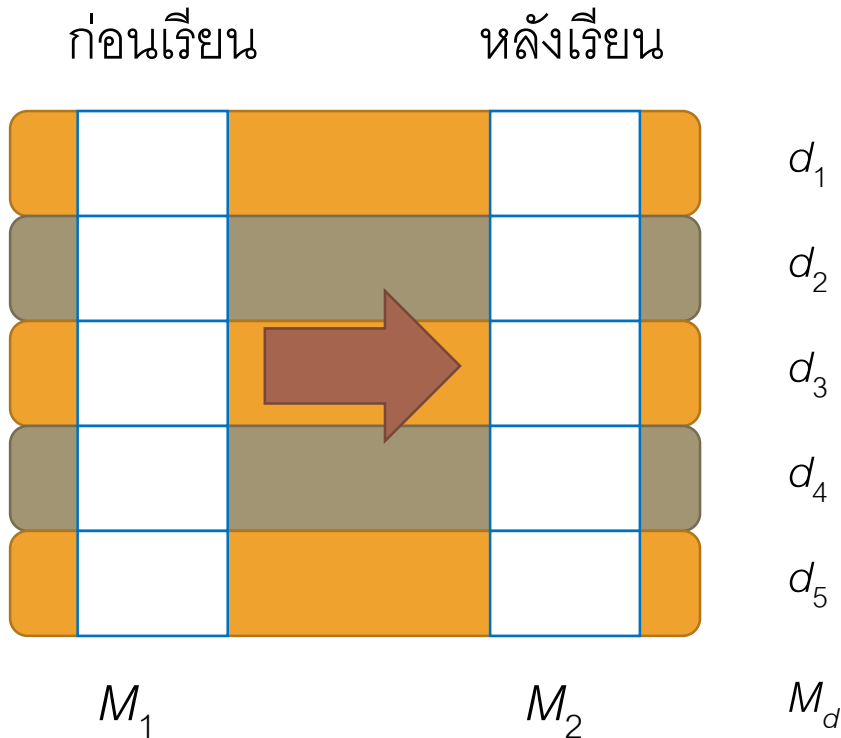
ตัวแปรที่ใช้เปรียบเทียบ
คือ ความรู้สึกรัก

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- ค่าเฉลี่ยที่เกี่ยวข้อกันอาจจะมีรูปแบบอื่นอีก สังเกตง่ายๆ โดยดูว่าข้อมูลมีลักษณะจับคู่ระหว่างกลุ่ม
- จำนวนข้อมูลในแต่ละกลุ่มต้องเท่ากัน

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- ในการทดสอบ จะมีการแปลงข้อมูลก่อนใช้ในการทดสอบ โดยหาค่าเปลี่ยนแปลง (Difference score) ของคะแนนแต่ละคู่ เช่น

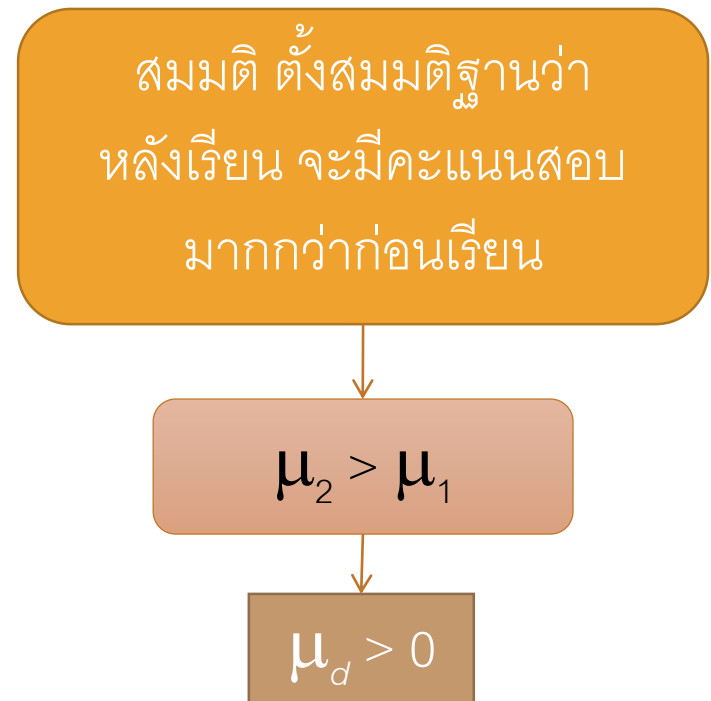
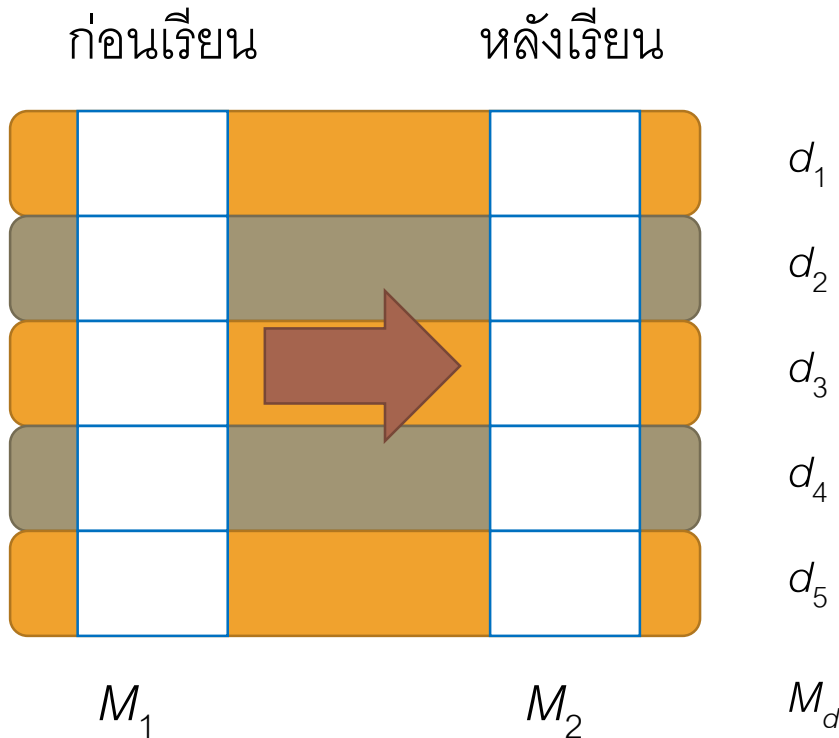


ค่าเปลี่ยนแปลง (d) = หลัง - ก่อน
(จะใช้เป็นก่อนลบหลังก็ได้)

$$M_d = M_2 - M_1$$

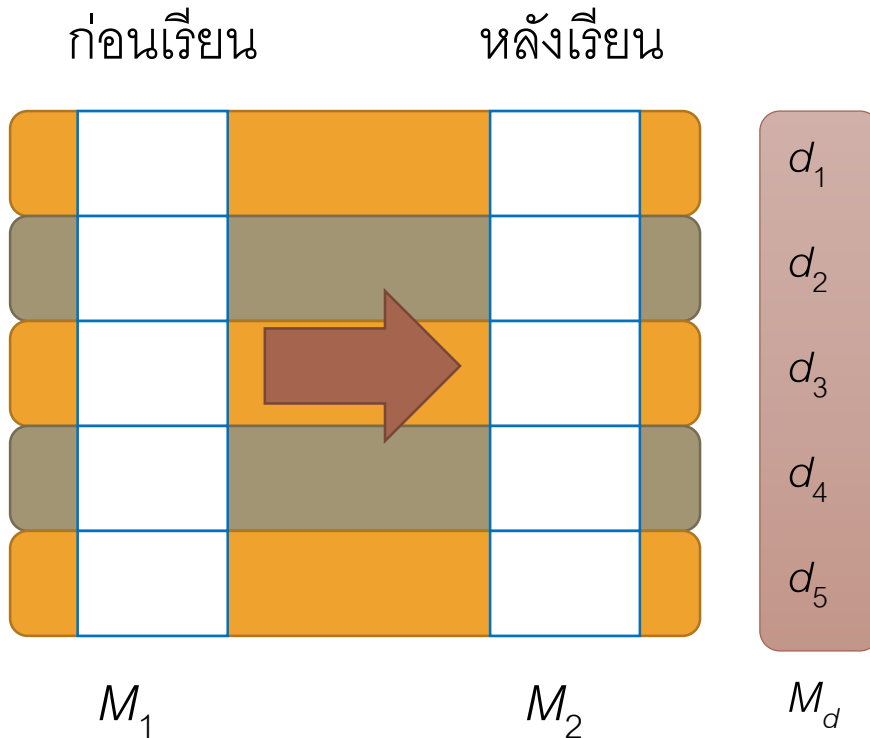
การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- ในการทดสอบ จะมีการแปลงข้อมูลก่อนใช้ในการทดสอบ โดยหาค่าเปลี่ยนแปลง (Difference score) ของคะแนนแต่ละคู่ เช่น



การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- ในการทดสอบ จะมีการแปลงข้อมูลก่อนใช้ในการทดสอบ โดยหาค่าเปลี่ยนแปลง (Difference score) ของคะแนนแต่ละคู่ เช่น



ดังนั้น ถ้าเรานำค่าเปลี่ยนแปลง
มาทดสอบด้วย
One-sample t test
(โดยตั้งสมมติว่าง ว่า $\mu_d = 0$)
ก็น่าจะตอบคำถามวิจัยได้

การเปรียบเทียบค่าเฉลี่ยแบบเกี่ยวข้อกัน

- การตั้งสมมติฐานทางสถิติ

สองทาง

$$H_0: \mu_1 = \mu_2; \mu_d = 0$$

$$H_1: \mu_1 \neq \mu_2; \mu_d \neq 0$$

ทางเดียว

$$H_0: \mu_1 \leq \mu_2; \mu_d \geq 0$$

$$H_1: \mu_1 > \mu_2; \mu_d < 0$$

$$H_0: \mu_1 \geq \mu_2; \mu_d \leq 0$$

$$H_1: \mu_1 < \mu_2; \mu_d > 0$$

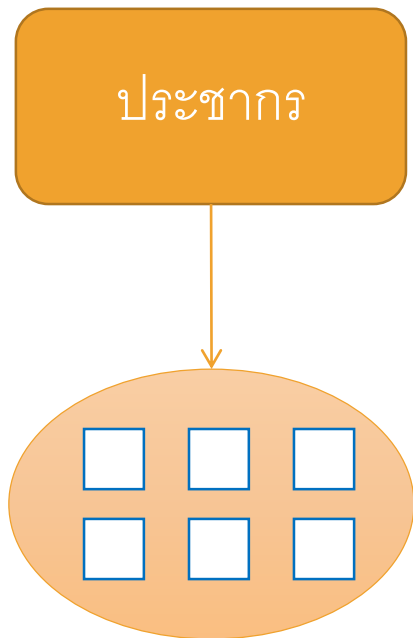
การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

- ค่าเฉลี่ยที่เป็นอิสระจากกัน มักจะเจออยู่ใน 2 รูปแบบ คือ
 - การสุ่มมาจากคนละประชากร
 - การแบ่งกลุ่มโดยผู้วิจัย

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

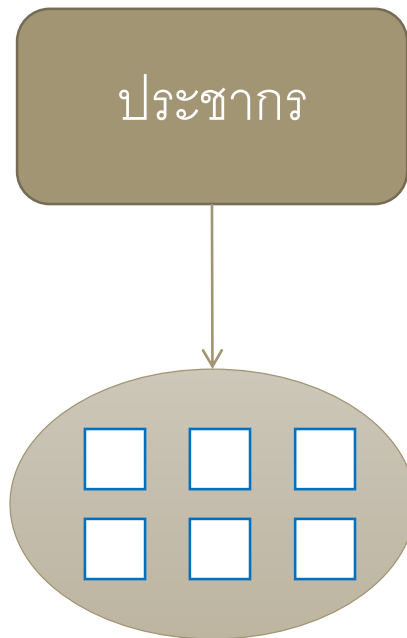
- การสุ่มมาจากคนละประชากร

คนที่มีถิ่นกำเนิดในเมือง



M_1

คนที่มีถิ่นกำเนิดในชนบท

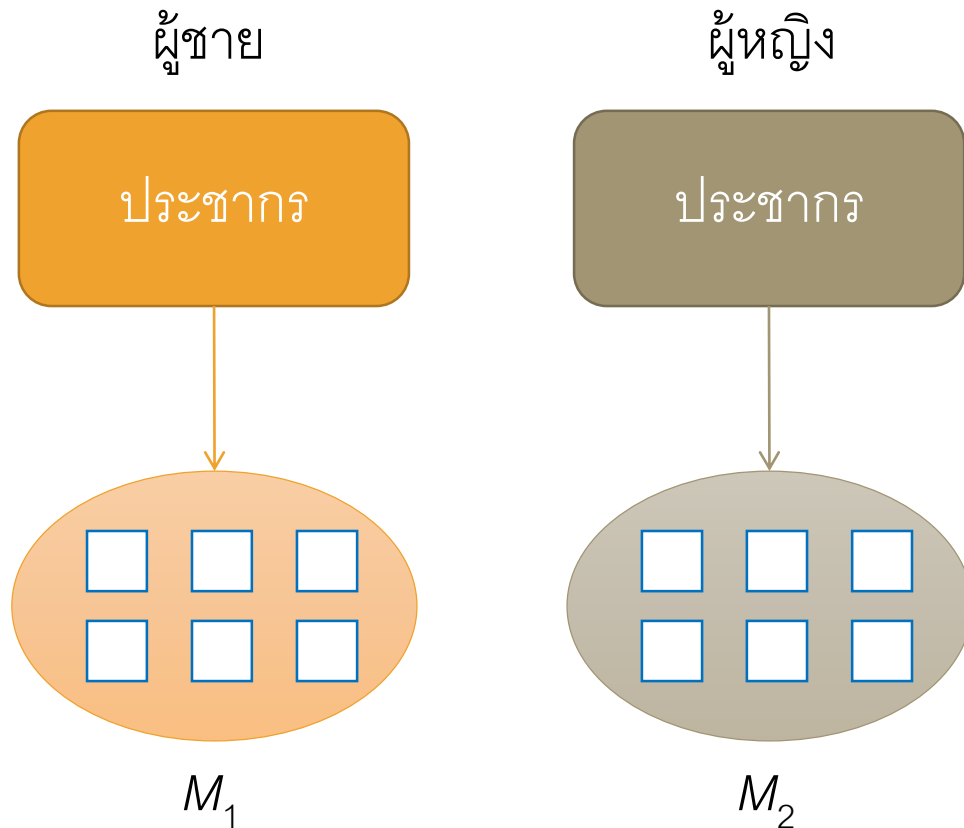


M_2

ตัวแปรที่ใช้เปรียบเทียบ
คือ ความอดอัดเมื่ออยู่ในที่แคบ

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

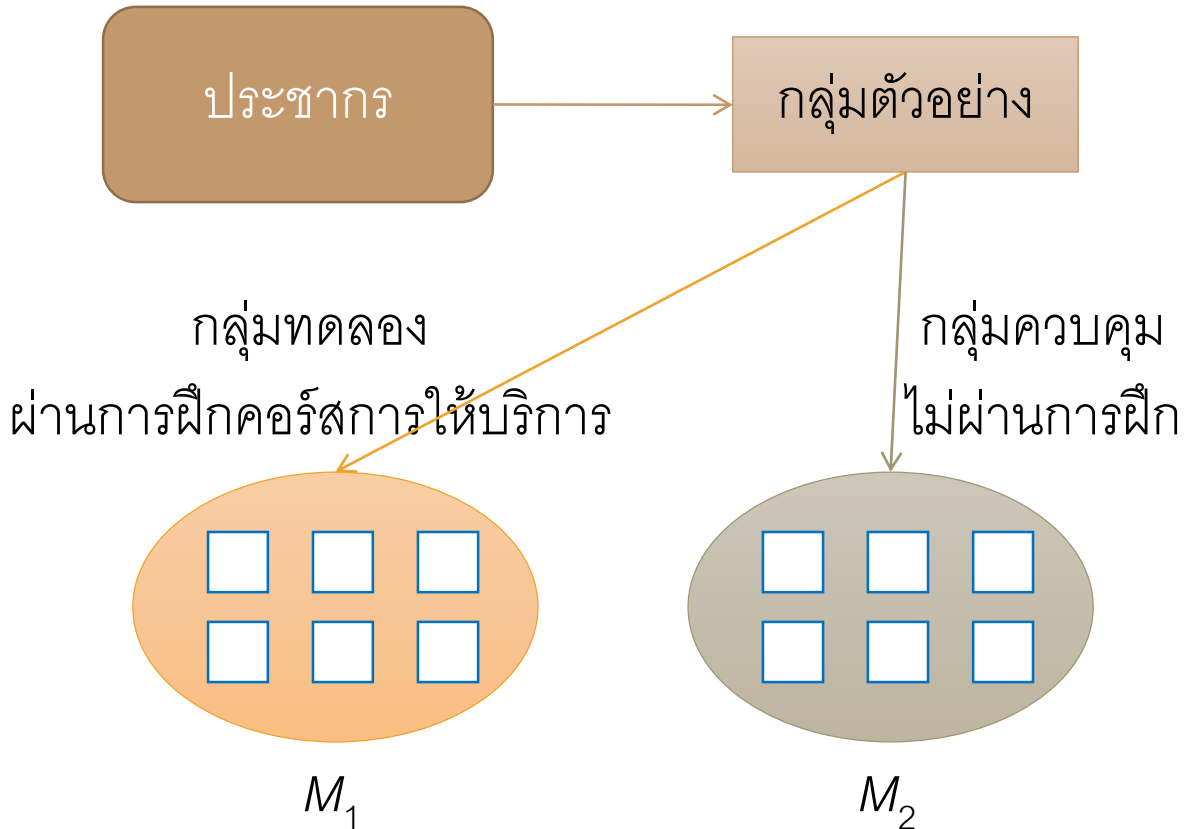
- การสุ่มมาจากคนละประชากร



ตัวแปรที่ใช้เปรียบเทียบ คือ
การให้ความสำคัญกับ
ความหล่อ/สวยของคุณครอง

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

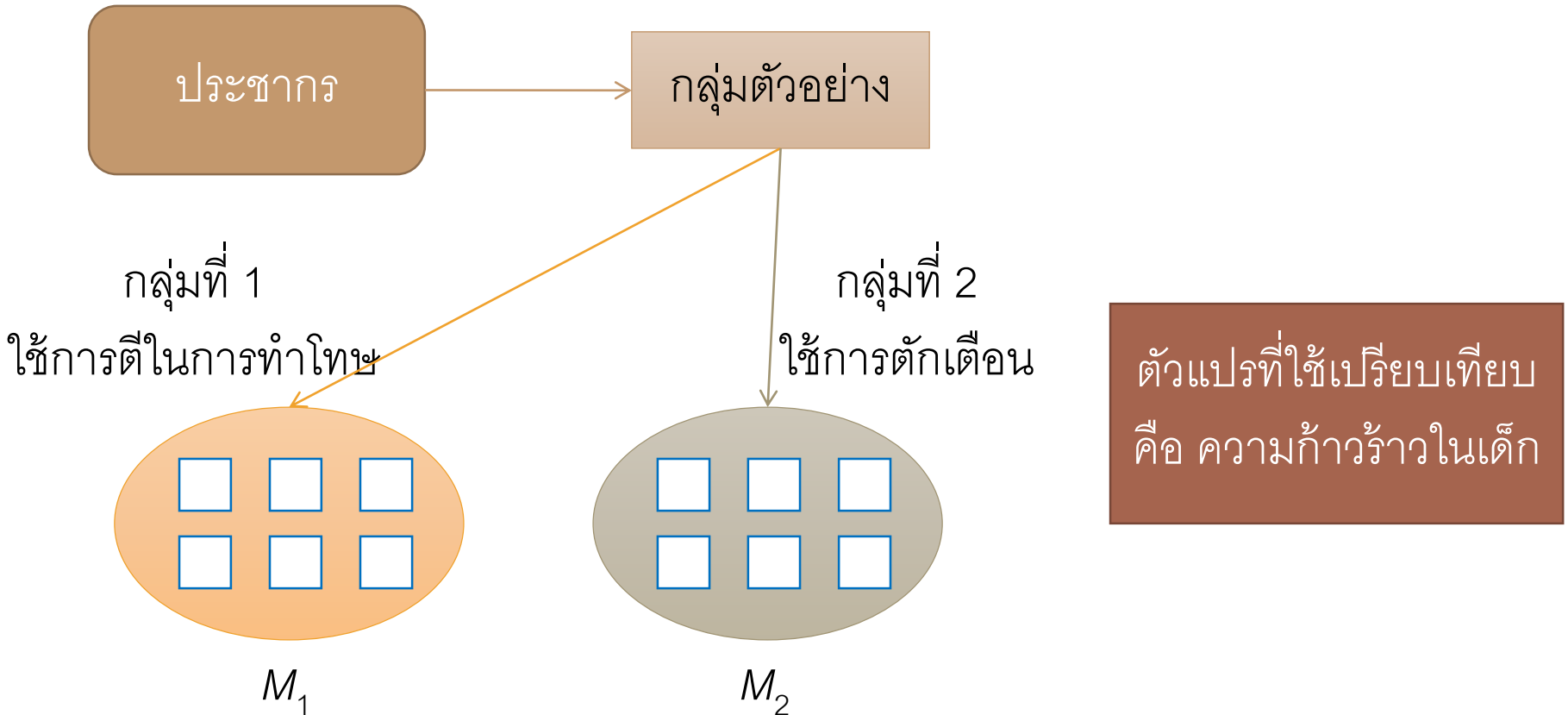
- การแบ่งกลุ่มโดยผู้วิจัย



ตัวแปรที่ใช้เปรียบเทียบ
คือ ผลการปฏิบัติงาน

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

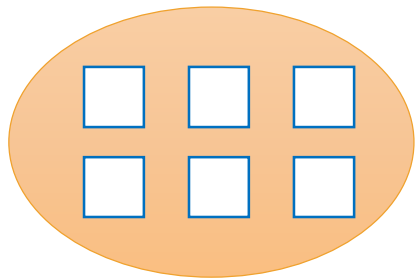
- การแบ่งกลุ่มโดยผู้วิจัย



การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

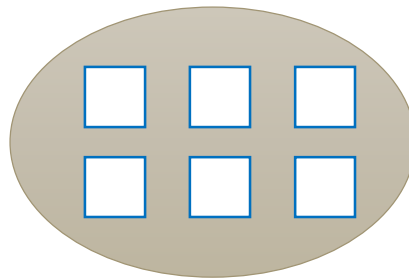
- การเปรียบเทียบค่าเฉลี่ยแบบอิสระ จะไม่สามารถสร้างค่าเปลี่ยนแปลงของคะแนนแต่ละคู่ได้ จะหาความแตกต่างของค่าเฉลี่ยโดยรวมได้โดยตรง

ผ่านการฝึก



M_1

ไม่ได้ผ่านการฝึก



M_2

$$M_d = M_2 - M_1$$

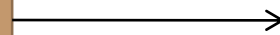
การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

- ค่าเฉลี่ยที่เป็นอิสระจากกันอาจจะมีรูปแบบอื่นอีก สังเกตง่ายๆ โดยดูว่าข้อมูลไม่สามารถจับคู่ได้ หรือไม่มีเหตุผลใดในการจับคู่ระหว่างกลุ่ม
- จำนวนข้อมูลในแต่ละกลุ่มไม่จำเป็นต้องเท่ากัน

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

- การเปรียบเทียบค่าเฉลี่ยแบบอิสระ จะไม่สามารถสร้างค่าเปลี่ยนแปลงของคะแนนแต่ละคู่ได้ จะหาความแตกต่างของค่าเฉลี่ยโดยรวมได้อย่างเดียว
- ด้วยเหตุนี้ การทดสอบจึงไม่สามารถทำการทดสอบค่า t ในกลุ่มตัวอย่างเดียว (One-sample t test) ได้
- จะใช้การประมาณค่าแบบใหม่ แล้วเรียกว่าการทดสอบค่า t ระหว่างสองกลุ่มตัวอย่างที่เป็นอิสระจากกัน (Independent t -test)
- ความแตกต่างของค่าเฉลี่ยจากสองกลุ่มตัวอย่าง จะนำไปทดสอบว่า มีความแตกต่างจริงในประชากรหรือไม่

$$M_d = M_2 - M_1$$



$$\mu_d = \mu_2 - \mu_1$$

การเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกัน

- การตั้งสมมติฐานทางสถิติ

สองทาง

$$H_0: \mu_1 = \mu_2; \mu_d = 0$$

$$H_1: \mu_1 \neq \mu_2; \mu_d \neq 0$$

ทางเดียว

$$H_0: \mu_1 \leq \mu_2; \mu_d \geq 0$$

$$H_1: \mu_1 > \mu_2; \mu_d < 0$$

$$H_0: \mu_1 \geq \mu_2; \mu_d \leq 0$$

$$H_1: \mu_1 < \mu_2; \mu_d > 0$$

ในบทนี้จะกล่าวถึงวิธีเปรียบเทียบค่าเฉลี่ยแบบอิสระจากกันเท่านั้น

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- สมมติแบ่งคนเป็นสองกลุ่ม กลุ่มที่ผ่านและไม่ผ่านการฝึกทำสมาธิ
- ถ้าการทำสมาธิไม่มีผลต่อผลการเรียน ก็ไม่ได้หมายความว่าทั้งสองกลุ่มจะไม่แตกต่างกัน
- ความแตกต่างอาจจะเกิดขึ้นจากการสุ่ม (Sampling Error)
 - เช่น บังเอิญสุ่มคนที่มีผลการเรียนดีเข้าในกลุ่มฝึก และคนที่ผลการเรียนไม่ดีเข้ากลุ่มไม่ได้ผ่านการฝึก

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- ค่า p (p value) คือ ค่าที่บอกว่าโอกาสที่สุ่มประชากรที่ไม่แตกต่างกัน แล้วได้ความแตกต่างขนาดนี้ มีโอกาสเท่าไร
- ถ้าน้อยกว่าระดับนัยสำคัญ จะปฏิเสธสมมติฐานว่าง (Reject H_0)
- ถ้ามากกว่าระดับนัยสำคัญ จะไม่สามารถปฏิเสธสมมติฐานว่าง (Fail to Reject H_0)

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น

นิสิตจิตวิทยาจะมีอารมณ์มั่นคง
แตกต่างจากนิสิตคณะอื่น



อารมณ์มั่นคง วัดโดยแบบสอบถาม
คะแนนไล่ตั้งแต่ 0-100 คะแนน



นักวิจัยจึงสุ่มนิสิตจิตวิทยา 10 คน
และนิสิตคณะอื่น 20 คน



สมมติฐาน คือ นิสิตจิตวิทยามีอารมณ์
มั่นคงแตกต่างจากนิสิตคณะอื่น



$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

●

นิสิต

อาจารย์

คณะ

*Untitled1 [DataSet0] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

30 : stability 80 Visible: 3 of 3 Variables

	id	faculty	stability	var	var
1	1.00	2.00	80.00		
2	2.00	1.00	70.00		
3	3.00	2.00	65.00		
4	4.00	2.00	88.00		
5	5.00	2.00	76.00		
6	6.00	1.00	54.00		
7	7.00	2.00	78.00		
8	8.00	2.00	67.00		
9	9.00	1.00	75.00		
10	10.00	1.00	89.00		
11	11.00	2.00	54.00		
12	12.00	2.00	65.00		
13	13.00	2.00	87.00		
14	14.00	2.00	89.00		

2 = คณะอื่นๆ

1 = คณะจิตวิทยา

Data View Variable View

SPSS Processor is ready

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

•
นิสิ
อาร
คณ

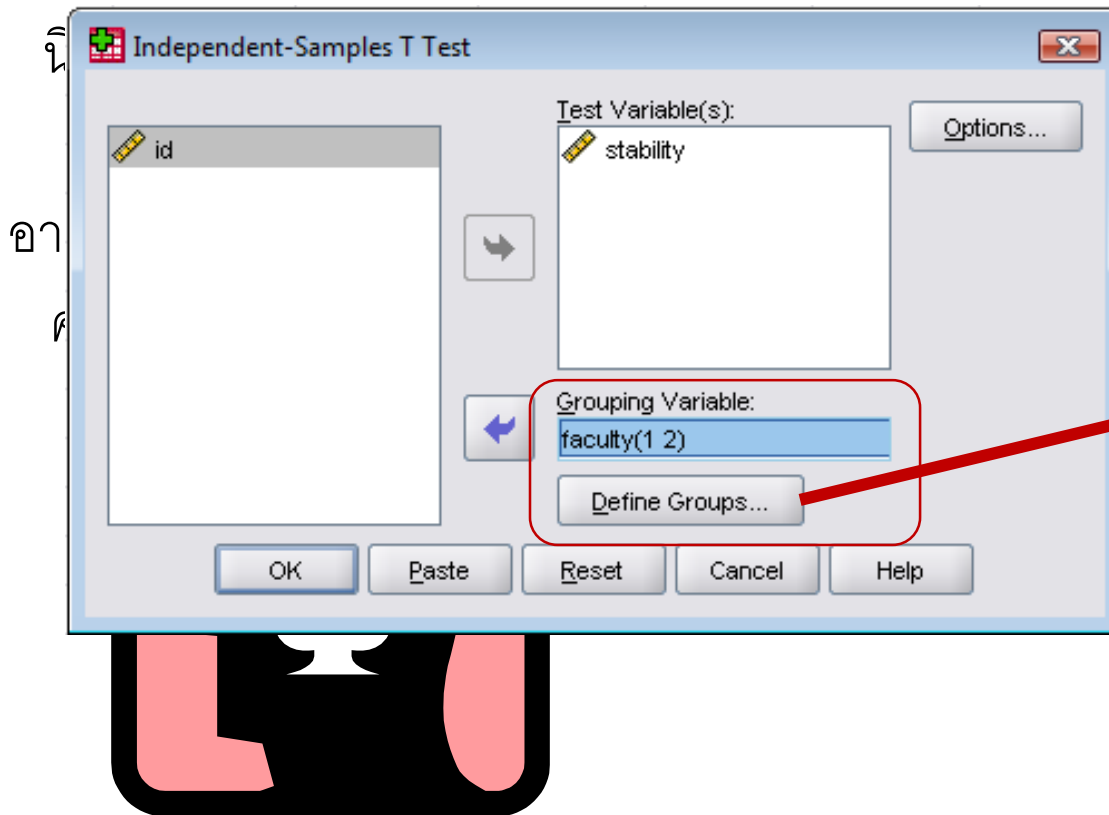
The screenshot shows the SPSS Data Editor interface. The 'Analyze' menu is open, and the 'Independent-Samples T Test...' option is highlighted. The data table shows a variable 'id' with values from 1.00 to 14.00. The status bar at the bottom indicates 'SPSS Processor is ready'.

	id	facul
1	1.00	
2	2.00	
3	3.00	
4	4.00	
5	5.00	
6	6.00	
7	7.00	
8	8.00	
9	9.00	
10	10.00	
11	11.00	
12	12.00	
13	13.00	
14	14.00	

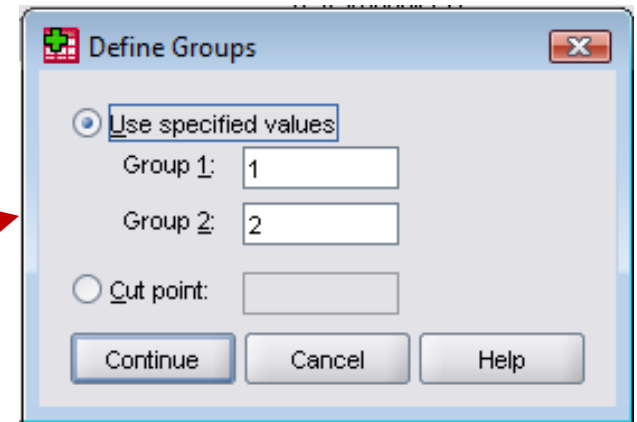
SPSS Processor is ready

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น



ใส่กลุ่มที่ต้องการเปรียบเทียบ



การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

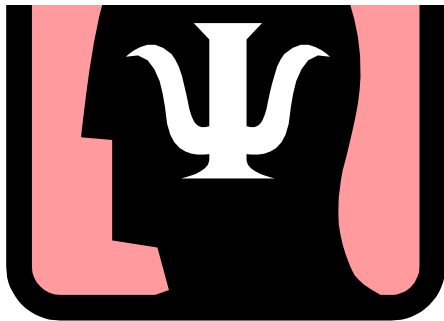
- เช่น

คณะจิตวิทยา

คณะอื่นๆ

Group Statistics

faculty	N	Mean	Std. Deviation	Std. Error Mean
Emotional Stability Psychology	10	74.0000	14.46067	4.57287
Others	20	72.4000	13.19649	2.95083



การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

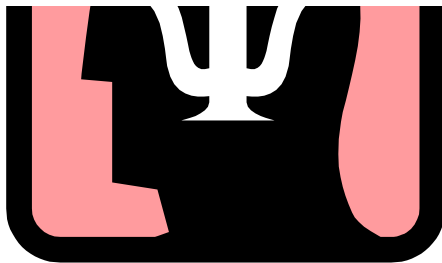
- เช่น

$$p(2\text{-tailed}) = .76$$

$$t(28) = 0.30, p > .05$$

Independent Samples Test

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Emotional Stability	Equal variances assumed	.303	28	.764	1.60000	5.27332	-9.20190	12.40190
	Equal variances not assumed							



กล่าวถึงในภายหลัง

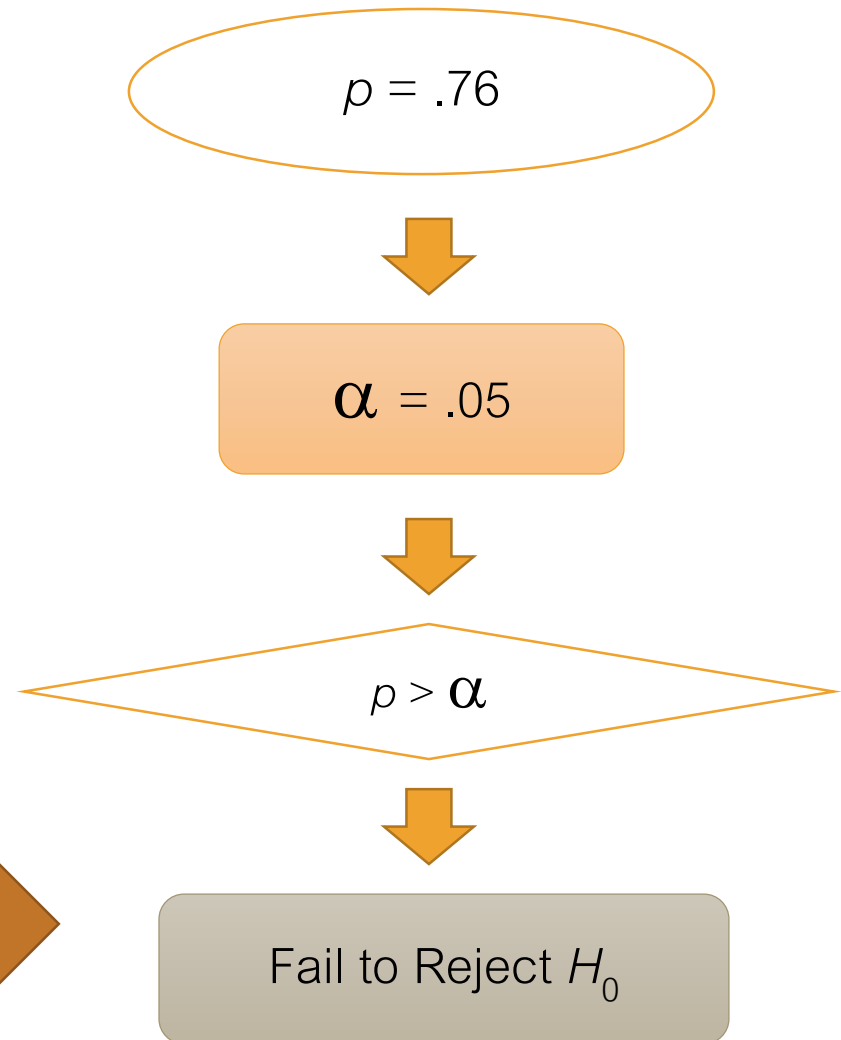
การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น

นิสิตจิตวิทยาจะมีอารมณ์มั่นคง
แตกต่างจากนิสิตคณะอื่น
อารมณ์มั่นคง วัดโดยแบบสอบถาม
คะแนนไล่ตั้งแต่ 0-100 คะแนน



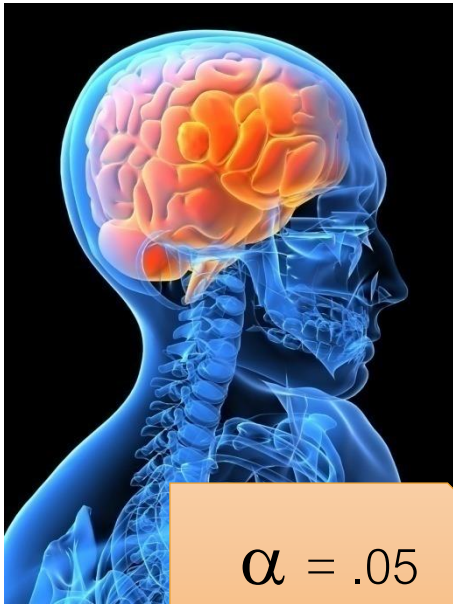
ไม่สามารถสรุปได้ว่า นิสิตคณะจิตวิทยาจะมีอารมณ์มั่นคงแตกต่างจากนิสิตคณะอื่นๆ



การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น

เด็กที่ดื่มชูปโก้ จะมีความฉลาด
สูงกว่าเด็กที่ไม่ได้ดื่ม



นักวิจัยจึงสุ่มเด็กอายุ 12 ขวบ จำนวน 20 คน
แล้วแบ่งเป็นกลุ่มที่ดื่ม และไม่ดื่มอย่างละ 10 คน



สมมติฐาน คือ เด็กดื่มชูปโก้ (μ_1)
ฉลาดกว่าเด็กที่ไม่ได้ดื่ม (μ_2)



$$H_0: \mu_1 - \mu_2 \leq 0$$



$$H_1: \mu_1 - \mu_2 > 0$$

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น

Group Statistics

Chick en ...	N	Mean	Std. Deviation	Std. Error Mean
Daily	10	101.7000	10.36072	3.27635
None	10	100.8000	11.77379	3.72320

ดื่มนมทุกวัน

ไม่ได้ดื่มนม



การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น

$$p(2\text{-tailed}) = .86$$

$$p(1\text{-tailed}) = .43$$

$$t(18) = 0.18, p > .05$$

Independent Samples Test

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
IQ	Equal variances assumed	181	18	.858	.90000	4.95950	-9.51953	11.31953
	Equal variances not assumed							

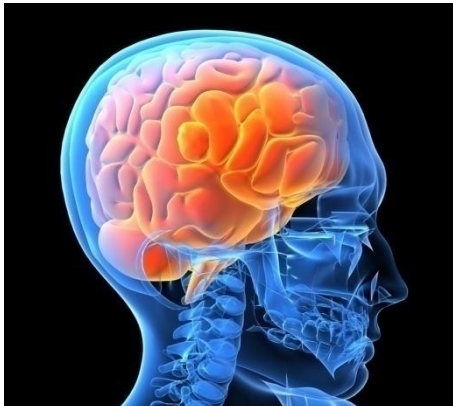


กล่าวถึงในภายหลัง

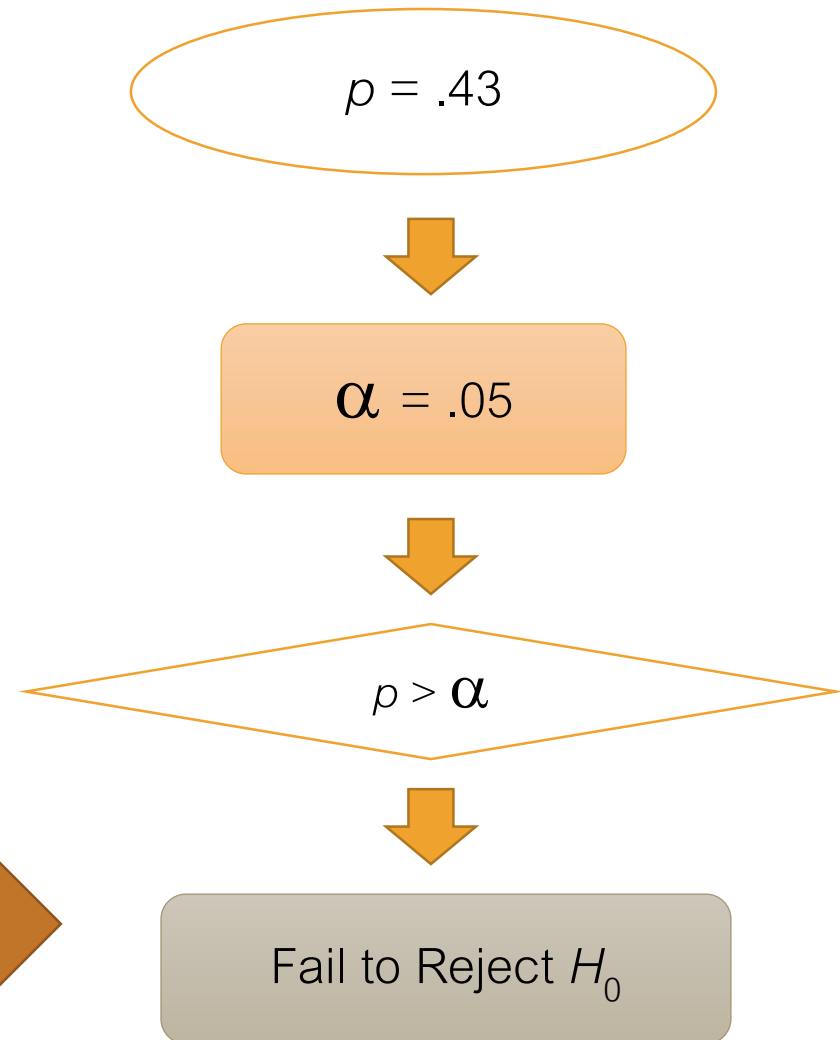
การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น

เด็กที่ดื่มชูปไก่ จะมีความฉลาด
สูงกว่าเด็กที่ไม่ได้ดื่ม



ไม่สามารถสรุปได้ว่า เด็กที่ดื่มชูปไก่จะมี
ความฉลาดสูงกว่าเด็กที่ไม่ได้ดื่ม



ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- การกระจายของคะแนนในประชากรทั้งสองกลุ่มเป็นโค้งปกติ
 - การละเมิดไม่ค่อยมีผลกระทบ
- ในประชากร กลุ่มทั้งสองกลุ่มมีความแปรปรวนเท่ากัน (Homogeneity of Variance)
 - หากจำนวนกลุ่มตัวอย่างเท่ากันหรือใกล้เคียงกัน t-test แบบปกติจะไม่ได้รับผลกระทบจากความแปรปรวนที่แตกต่างกัน
 - ในกรณีกลุ่มตัวอย่างแต่ละกลุ่มแตกต่างกันมาก หากอัตราส่วนของความแปรปรวนระหว่างสองกลุ่มมากกว่า 10 ต่อ 1 ค่า p ที่ได้จะผิดไปจากความเป็นจริง

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- แก้ไขโดยใช้ Welch Test (บรรทัดที่ 2: Equal Variance Not Assumed)
- แต่ทว่า Welch Test มีข้อตกลงเบื้องต้นว่าการกระจายของประชากรต้องเป็นโค้งปกติ หากไม่เป็นจะส่งผลกระทบต่ออย่างร้ายแรง

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
IQ	Equal variances assumed	.265	.613	.181	18	.858	.90000	4.95950	-9.51953	11.31953
	Equal variances not assumed			.181	17.714	.858	.90000	4.95950	-9.53162	11.33162

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- วิธีการทดสอบว่ากลุ่มทั้งสองกลุ่มมีความแปรปรวนเท่ากัน

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

ทดสอบด้วย

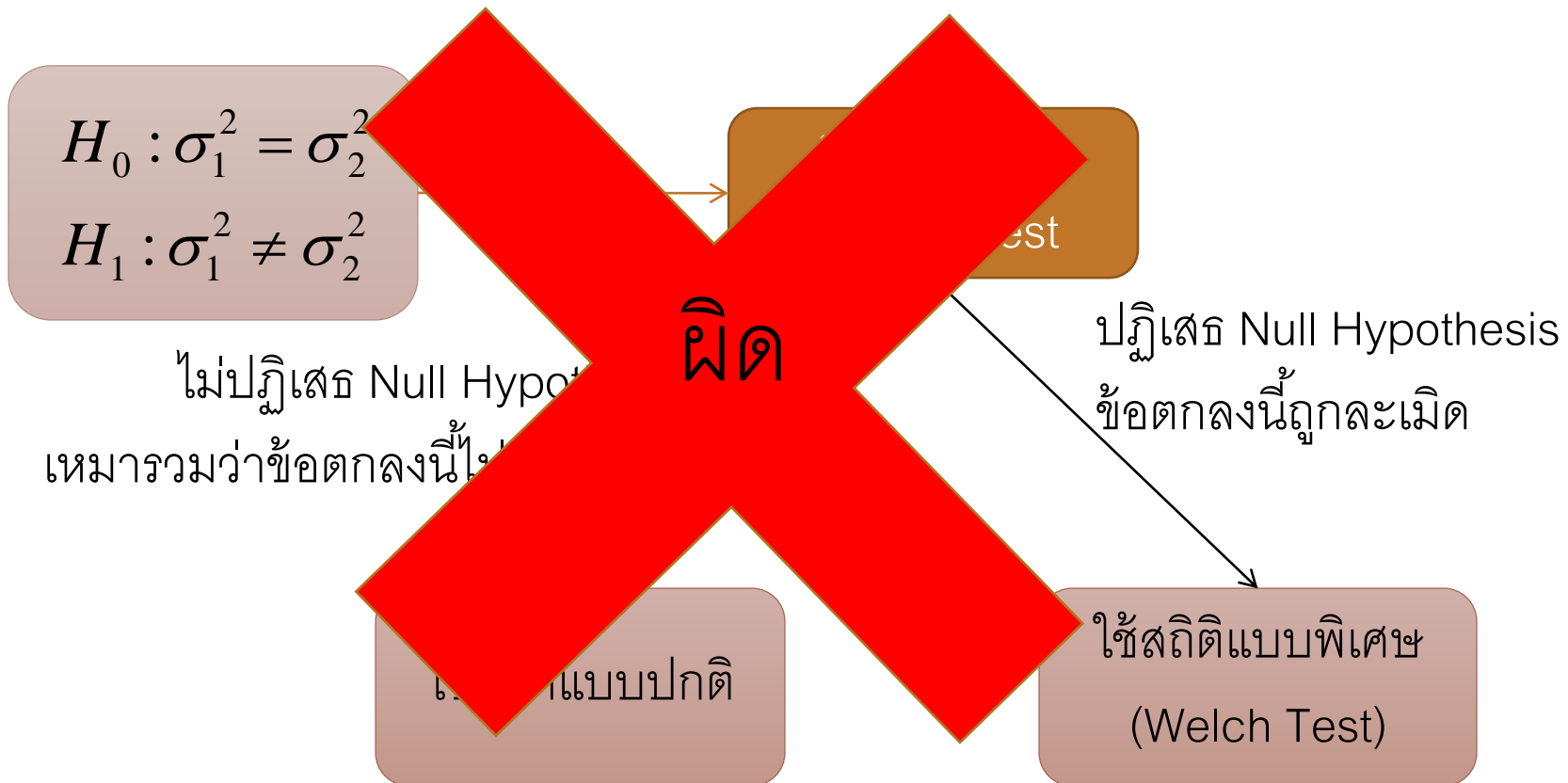
Levene Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
IQ	Equal variances assumed	.265	.613	.181	18	.858	.90000	4.95950	-9.51953	11.31953
	Equal variances not assumed			.181	17.714	.858	.90000	4.95950	-9.53162	11.33162

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- วิธีการทดสอบว่ากลุ่มทั้งสองกลุ่มมีความแปรปรวนเท่ากัน



ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- กำลังในการทดสอบ Levene Test ขึ้นอยู่กับ
 - ความแตกต่างของความแปรปรวนระหว่างกลุ่ม (ขนาดอิทธิพล)
 - ขนาดกลุ่มตัวอย่าง
- บางครั้ง กลุ่มตัวอย่างน้อย Levene Test ไม่สามารถ Reject H_0 ได้ แม้ขนาดอิทธิพลจะสูงมาก
 - ทำให้ใช้แบบปกติทั้งที่ควรใช้ Welch Test
- บางครั้ง กลุ่มตัวอย่างสูง Levene Test ทำให้ Reject H_0 ตลอดเวลา แม้ขนาดอิทธิพลจะต่ำมาก
 - ทำให้ใช้ Welch Test ตลอดเวลา บางครั้งไม่ดีเพราะการกระจายประชากรไม่เป็นโค้งปกติ

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น

การทดสอบประสิทธิภาพของ
การสอนแบบเด็กเป็นศูนย์กลาง



$\alpha = .05$



นักวิจัยจึงเลือกนักเรียน 2 ห้อง ห้องละ 30 คน
ห้องแรกสอนแบบปกติ
ห้องที่สองสอนแบบเด็กเป็นศูนย์กลาง



สมมติฐาน คือ ห้องที่สองมีคะแนน
สอบดีกว่าห้องที่หนึ่ง



$$H_0: \mu_2 - \mu_1 \leq 0$$



$$H_1: \mu_2 - \mu_1 > 0$$

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น

สอนแบบปกติ

สอนแบบนักเรียนเป็นศูนย์กลาง

Group Statistics

	Method	N	Mean	Std. Deviation	Std. Error Mean
Score	Lecture	30	6.4000	1.10172	.20115
	Child-centered	30	6.5333	1.77596	.32424



กลุ่มตัวอย่างเท่ากัน ใช้ *t*-test แบบปกติ

ตรวจสอบขนาดความแปรปรวนที่แตกต่างกัน

$$(1.776)^2 / (1.102)^2 = 2.6 \text{ (ใช้ } t\text{-test แบบปกติ)}$$

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

$p(2\text{-tailed}) = .72$

• เช่น

$p(1\text{-tailed}) = .36$

$t(58) = -0.35, p > .05$ (1-tailed)

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Score	Equal variances assumed	10.440	.002	-.349	58	.728	-.13333	.38157	-.89712	.63046
	Equal variances not assumed			-.349	48.441	.728	-.13333	.38157	-.90035	.63368



Levene Test: $p < .05$

ความแปรปรวนระหว่างกลุ่มแตกต่างกัน

ความแตกต่างยังไม่มากจนต้องใช้ Welch Test

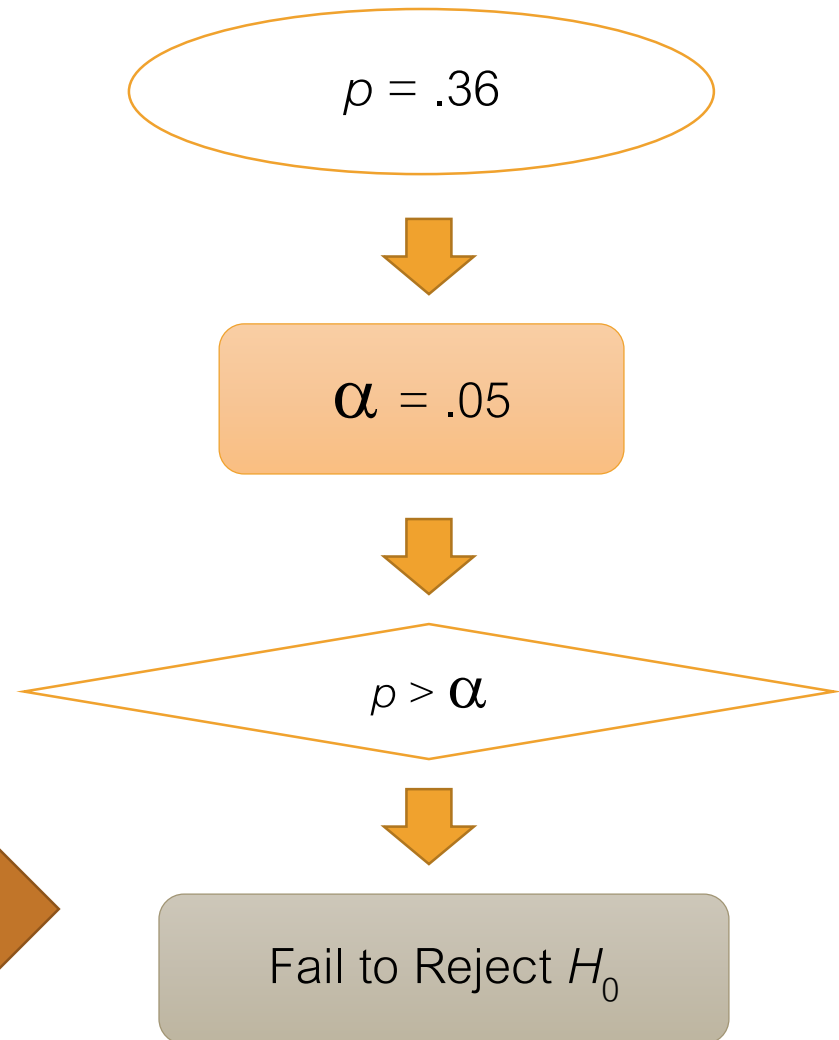
การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น

การทดสอบประสิทธิภาพของ
การสอนแบบเด็กเป็นศูนย์กลาง



ไม่สามารถสรุปได้ว่า การสอนแบบ
เด็กเป็นศูนย์กลางมีประสิทธิภาพ
สูงกว่าการสอนปกติ



การเขียนรายงาน

- เช่น งานวิจัยนี้ต้องการทดสอบประสิทธิภาพของการสอนแบบนักเรียนเป็นศูนย์กลาง การทดสอบนี้จึงใช้วิธีการสอนแบบปกติ และแบบนักเรียนเป็นศูนย์กลางกับแต่ละห้องเรียน ซึ่งมีนักเรียนห้องละ 30 คน จากการศึกษาด้วยการทดสอบ t แบบอิสระจากกัน (Independent t -test) พบว่าห้องเรียนที่ใช้การสอนแบบปกติ ($M = 6.40, SD = 1.10$) ไม่แตกต่างจากห้องเรียนใช้การสอนแบบนักเรียนเป็นศูนย์กลาง ($M = 6.53, SD = 1.76$) อย่างมีนัยสำคัญ ($t(58) = -0.35, p > .05$ [ทดสอบแบบทางเดียว])

การหาช่วงเชื่อมั่นของความแตกต่างระหว่างค่าเฉลี่ย

- ผู้วิจัยอาจต้องการทราบว่าการเปลี่ยนแปลงที่เกิดขึ้นในประชากรมีค่าเท่าไร
 - เช่น กลุ่มเด็กที่ดื่มนม จะมีความสูงมากกว่ากลุ่มเด็กที่ไม่ดื่มนมเท่าไร ในประชากร
- คำถามนี้จะสามารถใช้ช่วงเชื่อมั่นสำหรับค่าเปลี่ยนแปลงระหว่างกลุ่มสองกลุ่มในการตอบคำถามได้

การหาช่วงเชื่อมั่นของความแตกต่างระหว่างค่าเฉลี่ย

- เช่น ผู้วิจัยอาจต้องการทราบว่าวิธีการบำบัดโดยใช้หมากฝรั่งและการหักดิบจะทำให้อาการสูบบุหรี่ลดลงอย่างไร
- เปรียบเทียบโดยใช้มาตรวัดอาการติดยาสูบ (Nicotine Dependence Scale) คะแนนตั้งแต่ 20-80 คะแนน
- ผู้วิจัยเก็บข้อมูลจากกลุ่มละ 24 คน และหาช่วงเชื่อมั่นของความแตกต่างของผลการบำบัดจากทั้งสองวิธี

การหาช่วงเชื่อมั่นของความแตกต่างระหว่างค่าเฉลี่ย

เปรียบเทียบการลดบุหรี่
ด้วยหมากฝรั่งและการหักดิบ
จากกลุ่มละ 24 คน

ไม่ต้องกังวลเรื่องความแปรปรวนที่ต่างกันระหว่างกลุ่ม
เพราะจำนวนกลุ่มตัวอย่างเท่ากัน

$$M_1 - M_2 = 48.63 - 50.46 = -1.83$$



การลดด้วยหมากฝรั่ง

การลดด้วยการหักดิบ

group statistics

	Treatment Condition	N	Mean	Std. Deviation	Std. Error Mean
Post-treatment Nicotine Dependence	Gum	24	48.6250	10.90597	2.22617
	Cold Turkey	24	50.4583	7.54107	1.53931

การหาช่วงเชื่อมั่นของความแตกต่างระหว่างค่าเฉลี่ย

เปรียบเทียบการลดบุหรื
ด้วยหมากฝรั่งและการหักดิบ
จากกลุ่มละ 24 คน



Independent Samples Test

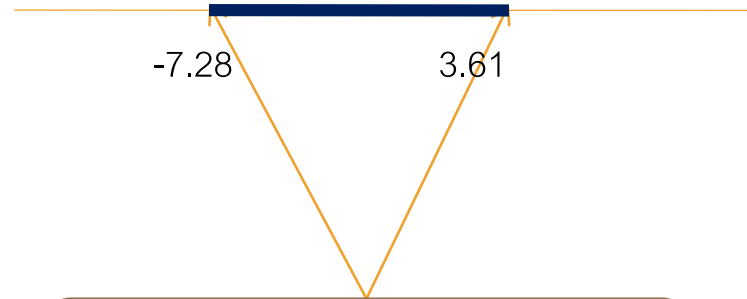
$$M_1 - M_2 = -1.83$$

$$CI_{.95} \text{ ของ } \mu_1 - \mu_2 = (-7.28, 3.61)$$

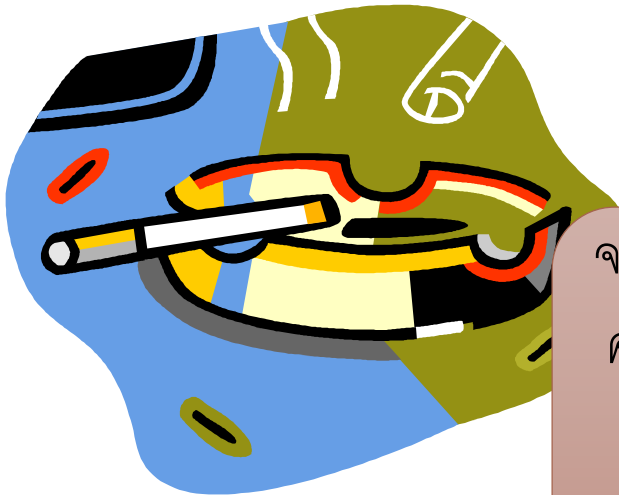
Levene's Test for Equality of Variances		t-test for Equality of Means						
F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
4.666	.036	-.677	46	.502	-1.83333	2.70654	-7.28131	3.61464
		-.677	40.901	.502	-1.83333	2.70654	-7.29969	3.63303

การหาช่วงเชื่อมั่นของความแตกต่างระหว่างค่าเฉลี่ย

เปรียบเทียบการลดบุหรี่
ด้วยหมากฝรั่งและการหักดิบ
จากกลุ่มละ 24 คน



$$M_1 - M_2 = -1.83$$



จากข้อมูลทำให้มีความเชื่อมั่นระดับ .95 ว่าประชากรของความแตกต่างระหว่างคะแนนอาการติดยาหรือ อยู่ในช่วง -7.28 (หักดิบได้ผลมากกว่าประมาณ 7 แต้ม) ถึง 3.61 (หมากฝรั่งได้ผลมากกว่าประมาณ 4 แต้ม)

การหาขนาดอิทธิพล

- การหาขนาดอิทธิพล ใช้หลักการเดียวกับการหาอิทธิพลในการทดสอบ z คือ

$$\delta = \frac{\mu_1 - \mu_2}{\sigma}$$

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

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

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

การหาขนาดอิทธิพล

- เช่น วัดขนาดอิทธิพลของการสอนแบบนักเรียนเป็นศูนย์กลาง

สอนแบบปกติ
(ให้เป็นกลุ่มควบคุม)

สอนแบบนักเรียน
เป็นศูนย์กลาง

Group Statistics

	Method	N	Mean	Std. Deviation	Std. Error Mean
Score	Lecture	30	6.4000	1.10172	.20115
	Child-centered	30	6.5333	1.77596	.32424



การหาขนาดอิทธิพล

- เช่น วัดขนาดอิทธิพลของการสอนแบบนักเรียนเป็นศูนย์กลาง

$$SS_1 = 1.10 \times 1.10 \times (30 - 1) = 35.20$$

$$SS_2 = 1.78 \times 1.78 \times (30 - 1) = 91.47$$

ics

	Method	N	Mean	Std. Deviation	Std. Error Mean
Score	Lecture	30	6.4000	1.10172	.20115
	Child-centered	30	6.5333	1.77596	.32424



$$SD_{\text{Pooled}} = \sqrt{\frac{35.20 + 91.47}{29 + 29}} = 1.48$$

$$d = \frac{6.53 - 6.40}{1.48} = 0.088$$

การทดสอบสมมติฐานของความแตกต่างระหว่างค่าเฉลี่ย

- วิธีคำนวณอีกทางหนึ่ง

$$SD_{\text{Pooled}} = \frac{SE_{M_1 - M_2}}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{0.382}{\sqrt{\frac{1}{30} + \frac{1}{30}}} = 1.48$$

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means							
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
								Lower	Upper	
Score	Equal variances assumed	10.440	.002	-.349	58	.728	-.13333	.38157	-.89712	.63046
	Equal variances not assumed			-.349	48.441	.728	-.13333	.38157	-.90035	.63368



การหาขนาดอิทธิพล

- การหาส่วนเบี่ยงเบนมาตรฐานรวมนั้น ขึ้นกับหลักการว่า ถ้าทั้งสองกลุ่มสุ่มมาจากการกระจายที่มีส่วนเบี่ยงเบนมาตรฐานเดียวกัน
- ด้วยเหตุนี้จึงใช้ข้อมูลส่วนเบี่ยงเบนมาตรฐานของทั้งสองกลุ่มมารวมกัน
- แต่หากส่วนเบี่ยงเบนมาตรฐานของทั้งสองกลุ่มในประชากรไม่เท่ากัน (Heterogeneity of Variance) แล้ว ส่วนเบี่ยงเบนมาตรฐานของทั้งสองกลุ่มไม่ควรมาใช้รวมกัน

Levene Test ได้ $p < .05$

ใช้ *SD* ของกลุ่มควบคุม

บางกรณี หากไม่มีกลุ่มควบคุมก็อาจใช้ส่วนเบี่ยงเบนมาตรฐานร่วมได้

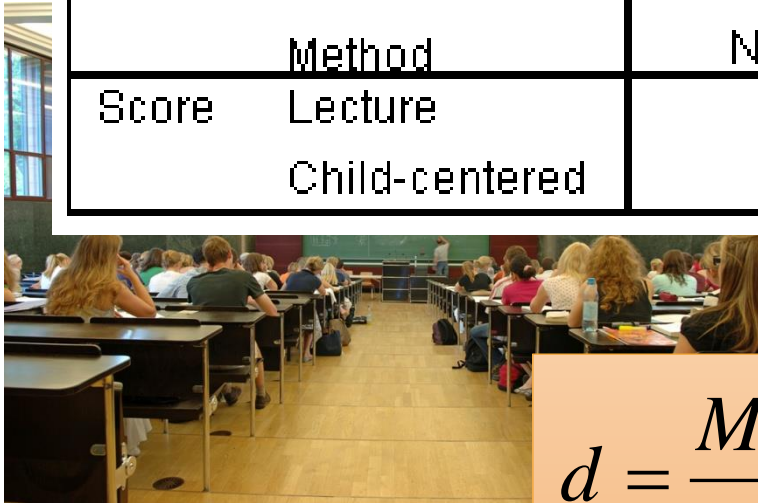
การหาขนาดอิทธิพล

- เช่น วัดขนาดอิทธิพลของการสอนแบบนักเรียนเป็นศูนย์กลาง

Levene's Test: $p = .002 \rightarrow$ ใช้ SD กลุ่มควบคุม

Group Statistics

	Method	N	Mean	Std. Deviation	Std. Error Mean
Score	Lecture	30	6.4000	1.10172	.20115
	Child-centered	30	6.5333	1.77596	.32424



$$d = \frac{M_2 - M_1}{SD_1} = \frac{6.53 - 6.40}{1.10} = 0.12$$

การหาขนาดอิทธิพล

- ช่วงเชื่อมั่นสำหรับขนาดอิทธิพล สามารถคำนวณได้จาก MBESS Package in R

การหาคำสั่ง

- ใช้โปรแกรม G*POWER 3 โดย
 - เลือกกลุ่มสถิติที่ต้องการทดสอบ คือ t-test
 - เลือกสถิติที่ใช้ทดสอบ คือ Means: Difference between two independent means (two groups)
 - เลือกว่าจะหาคำสั่งในการทดสอบ คือ Post hoc: Compute achieved power – given α , sample size and effect size

การหาคำสั่ง

- เช่น

ทดสอบกี่ทาง

ขนาดอิทธิพล

ระดับนัยสำคัญ

ขนาดกลุ่มตัวอย่าง

Test family: t tests

Statistical test: Means: Difference between two independent means (two groups)

Type of power analysis: Post hoc: Compute achieved power - given α , sample size, and effect size

Input Parameters

Determine =>	Tail(s)	Two
	Effect size d	0.04
	α err prob	0.05
	Sample size group 1	58
	Sample size group 2	62

Output Parameters

Noncentrality parameter δ	0.218967
Critical t	1.980272
Df	118
Power ($1 - \beta$ err prob)	0.055421

กำลังในการทดสอบสมมติฐานทางสถิติ

การหาค่าเฉลี่ย

- ขนาดอิทธิพลสามารถใช้ได้ทั้งสองแบบ
- สามารถหาค่า Determine แล้วใส่ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของทั้งสองกลุ่มเพื่อขนาดอิทธิพลจากส่วนเบี่ยงเบนมาตรฐานรวมได้

การกำหนดขนาดกลุ่มตัวอย่าง

- การกำหนดขนาดกลุ่มตัวอย่าง สามารถทำได้ 2 ประเภท คือ
 - การกำหนดขนาดกลุ่มตัวอย่างจากกำลังที่ต้องการ
 - การกำหนดขนาดกลุ่มตัวอย่างจากความผิดพลาดในการประมาณค่าพารามิเตอร์ (μ หรือ δ)
- ในที่นี้จะเน้นเฉพาะการหาจำนวนกลุ่มตัวอย่างจากกำลังเท่านั้น

การกำหนดขนาดกลุ่มตัวอย่าง

- การกำหนดขนาดกลุ่มตัวอย่างตามกำลังที่ต้องการ สามารถทำได้โดยใช้โปรแกรม G*POWER 3 ช่วย
 - เลือกกลุ่มสถิติที่ต้องการทดสอบ คือ t-test
 - เลือกสถิติที่ใช้ทดสอบ คือ Means: Difference between two independent means (two groups)
 - เลือกว่าจะหาลำดับในการทดสอบ คือ A Priori: Compute required sample size – given α , power and effect size

การกำหนดขนาดกลุ่มตัวอย่าง

- เช่น จงหาขนาดกลุ่มตัวอย่างที่จะทำให้กำลัง = .80 ในการทดสอบสองทาง และระดับนัยสำคัญเท่ากับ .05 เมื่อขนาดอิทธิพลเท่ากับ 0.04

Test family		Statistical test		
t tests		Means: Difference between two independent means (two groups)		
Type of power analysis				
A priori: Compute required sample size - given α , power, and effect size				
Input Parameters		Output Parameters		
Determine =>	Tail(s)	Two	Noncentrality parameter δ	2.801857
	Effect size d	0.04	Critical t	1.960085
	α err prob	0.05	Df	19624
	Power (1- β err prob)	0.80	Sample size group 1	9813
	Allocation ratio N2/N1	1	Sample size group 2	9813
			Total sample size	19626
		Actual power	0.800039	

การกำหนดขนาดกลุ่มตัวอย่าง

- ในโปรแกรม หากกำหนด Allocation Ratio = 1 หมายความว่า ให้จำนวนกลุ่มตัวอย่างของทั้งสองกลุ่มเท่ากัน
- เราสามารถกำหนดให้ Allocation Ratio ไม่เท่ากับ 1 ได้
 - เช่น Allocation Ratio = 2 หมายความว่า กลุ่มหนึ่งมีจำนวนกลุ่มตัวอย่างเป็นสองเท่าของอีกกลุ่มหนึ่ง