

การทดสอบค่าเฉลี่ยที่สัมพันธ์กัน

สถิติสำหรับจิตวิทยา 1

สันทัด พรประเสริฐมานิต

โครงร่างการนำเสนอ

- ค่าเปลี่ยนแปลง
- การทดสอบสมมติฐาน
- ช่วงเชื่อมั่น
- ขนาดอิทธิพล
- กำลังและการกำหนดขนาดกลุ่มตัวอย่าง
- การเปรียบเทียบการทดสอบที่ทั้งสองแบบ

แนะนำ

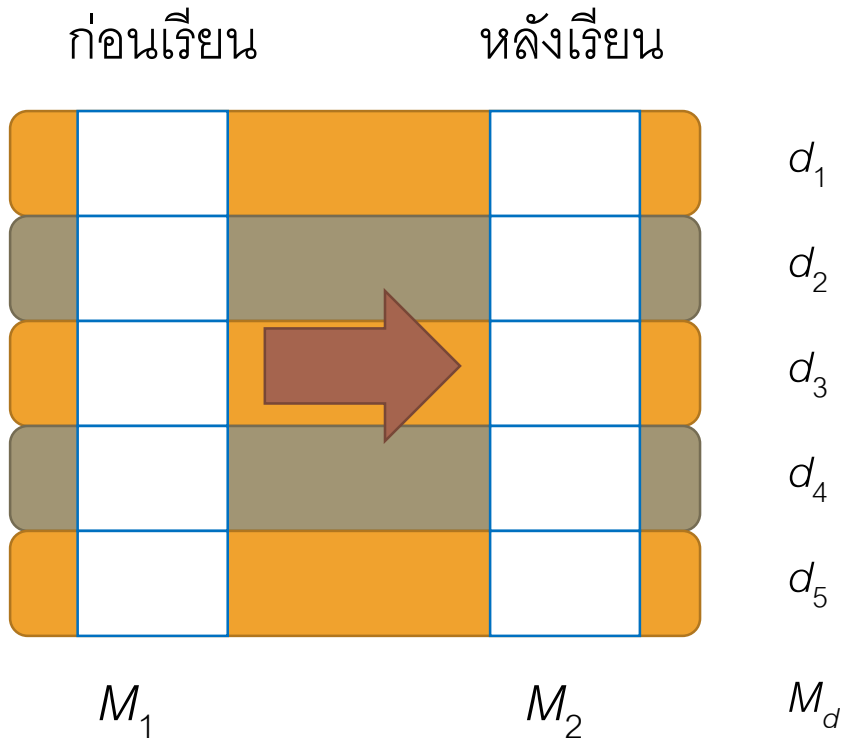
- เราได้กล่าวถึงการเปรียบเทียบค่าเฉลี่ยจากกลุ่มที่เป็นอิสระจากกันโดยใช้
Independent *t*-test
- ในครั้งนี้ จะเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มที่เกี่ยวข้องกันโดยใช้
Dependent *t*-test

แนะนำ

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

ค่าเปลี่ยนแปลง

- ในการทดสอบ จะมีการแปลงข้อมูลก่อนใช้ในการทดสอบ โดยหาค่าเปลี่ยนแปลง (Difference score) ของคะแนนแต่ละคู่ เช่น

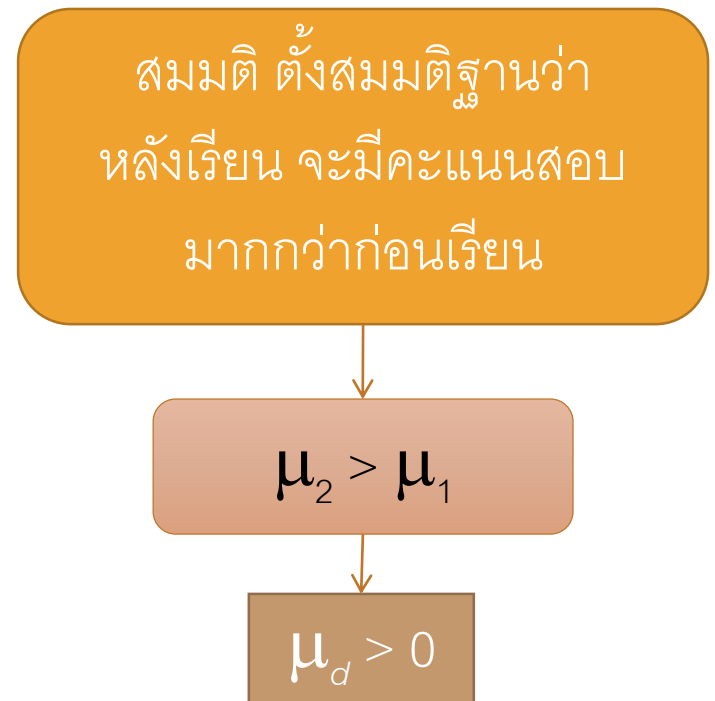
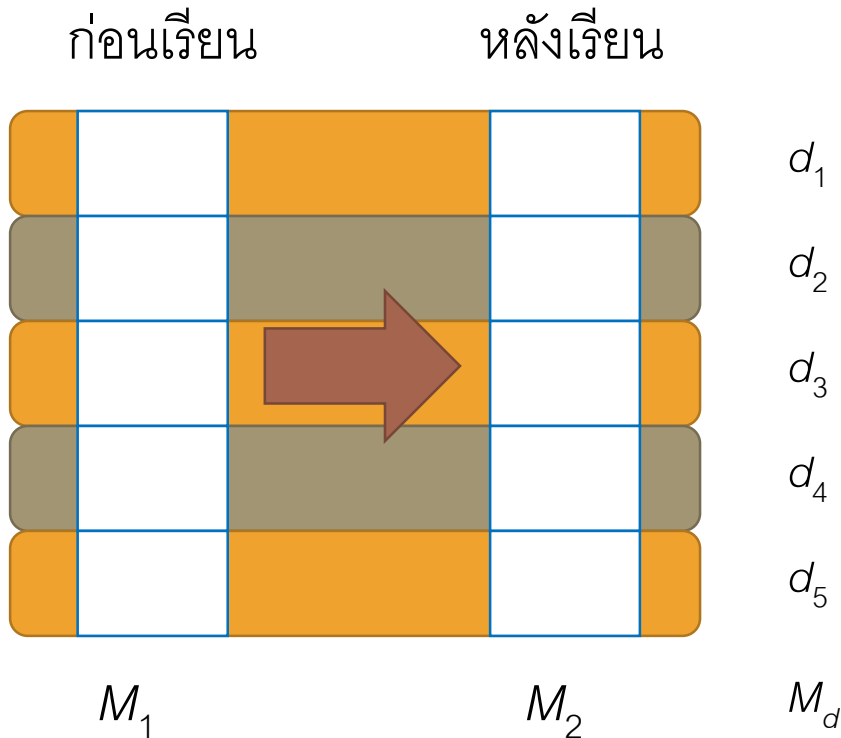


ค่าเปลี่ยนแปลง (d) = หลัง - ก่อน
(จะใช้เป็นก่อนลบหลังก็ได้)

$$M_d = M_2 - M_1$$

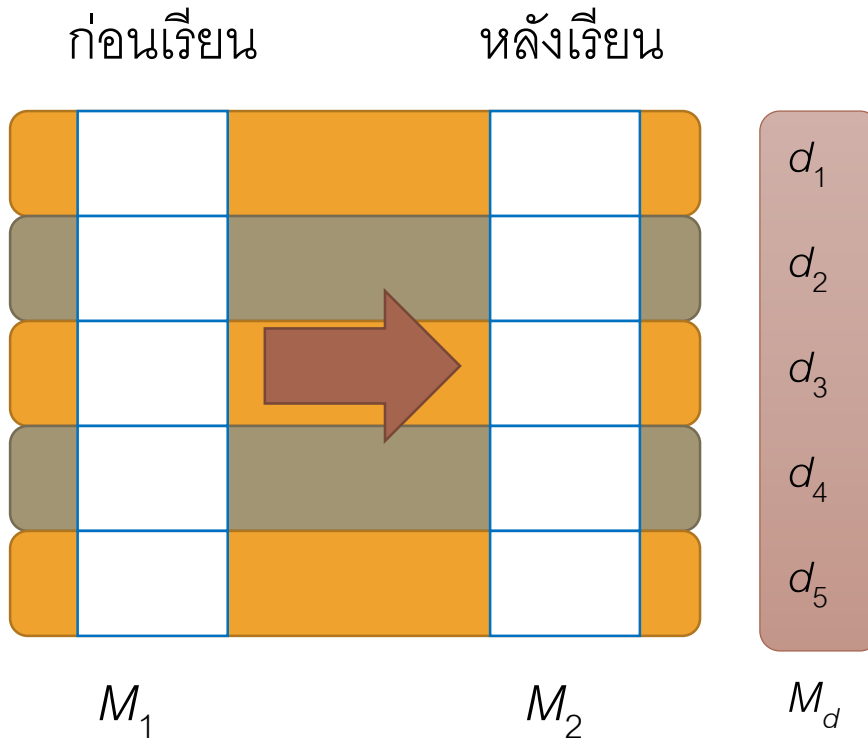
ค่าเปลี่ยนแปลง

- ในการทดสอบ จะมีการแปลงข้อมูลก่อนใช้ในการทดสอบ โดยหาค่าเปลี่ยนแปลง (Difference score) ของคะแนนแต่ละคู่ เช่น



ค่าเปลี่ยนแปลง

- ในการทดสอบ จะมีการแปลงข้อมูลก่อนใช้ในการทดสอบ โดยหาค่าเปลี่ยนแปลง (Difference score) ของคะแนนแต่ละคู่ เช่น



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

ค่าเปลี่ยนแปลง

- การตั้งสมมติฐานทางสถิติ

สองทาง

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

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

ทางเดียว

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

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

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

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

การทดสอบสมมติฐาน

- ตัวอย่าง ในการวัดน้ำหนัก ก่อนและหลังโปรแกรมลดน้ำหนัก
- ถ้าโปรแกรมการลดน้ำหนักไม่มีผล ไม่ได้หมายความว่าทุกคนไม่เปลี่ยนแปลงเลย
- บางคนน้ำหนักเพิ่ม บางคนน้ำหนักลด แต่โดยเฉลี่ยแล้วไม่มีผล
- ดังนั้น ตอนสุ่มกลุ่มตัวอย่าง อาจได้คนที่น้ำหนักเพิ่มเยอะ ทำให้ค่าเปลี่ยนแปลงเป็นบวก
- หรืออาจได้คนที่น้ำหนักลดเยอะ ทำให้ได้ค่าเปลี่ยนแปลงเป็นลบ

การทดสอบสมมติฐาน

- ดังนั้น ความแตกต่างที่ได้แบบนี้ เรียกว่า ความแตกต่างที่เกิดจากการสุ่ม (Sampling Error)
- ประชากรไม่ได้มีความเปลี่ยนแปลง แต่บังเอิญสุ่มได้คนสูงขึ้นบ้าง หรือ ต่ำลงบ้าง
- ค่า p (p value) คือ ค่าที่บอกว่าโอกาสที่สุ่มประชากรที่ไม่แตกต่างกัน แล้ว ได้ความแตกต่างขนาดนี้ มีโอกาสเท่าไร
- ถ้าน้อยกว่าระดับนัยสำคัญ จะปฏิเสธ Null Hypothesis (Reject H_0)
- ถ้ามากกว่าระดับนัยสำคัญ จะไม่สามารถปฏิเสธ Null Hypothesis (Fail to Reject H_0)

การทดสอบสมมติฐาน

- สถิติที่ใช้ทดสอบ คือ การทดสอบที่แบบเกี่ยวข้อกัน (Dependent t -test)
- ซึ่งแท้จริงแล้ว ก็คือการทดสอบที่สำหรับกลุ่มตัวอย่างเดียว (One sample t -test) เพื่อทดสอบว่าค่าเฉลี่ยของค่าเปลี่ยนแปลงแตกต่างจาก 0 ในประชากรหรือไม่

การทดสอบสมมติฐาน

- เช่น

นั้สิตจุฬายา ๒
จะมีความสวยเพิ่มขึ้นจากปี 1
การวัดความสวยไล่ตั้งแต่ 0-10



$\alpha = .05$

นักวิจัยจึงเก็บข้อมูลจากนั้สิตจุฬายา 20 คน
ตอนปี 1 รอดจนกระทั่งปี 2 แล้ววัดอีกครั้งหนึ่ง

สมมติฐาน คือ นั้สิตจุฬายา ๒
มีความสวยแตกต่างจากปี 1

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

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

การทดสอบสมมติฐาน

*Untitled1 [DataSet0] - SPSS Data Editor

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

18 : year2 5 Visible: 3 of 3 Variables

	id	year1	year2	var	var	var	var	var	var
1	1.00	4.00	5.00						
2	2.00	5.00	4.00						
3	3.00	6.00	8.00						
4	4.00	4.00	7.00						
5	5.00	3.00	3.00						
6	6.00	5.00	3.00						
7	7.00	6.00	7.00						
8	8.00	8.00	10.00						
9	9.00	8.00	8.00						
10	10.00	4.00	4.00						
11	11.00	5.00	6.00						
12	12.00	7.00	6.00						
13	13.00	5.00	3.00						
14	14.00	4.00	8.00						
15	15.00	5.00	5.00						

Data View Variable View

SPSS Processor is ready

การทดสอบสมมติฐาน

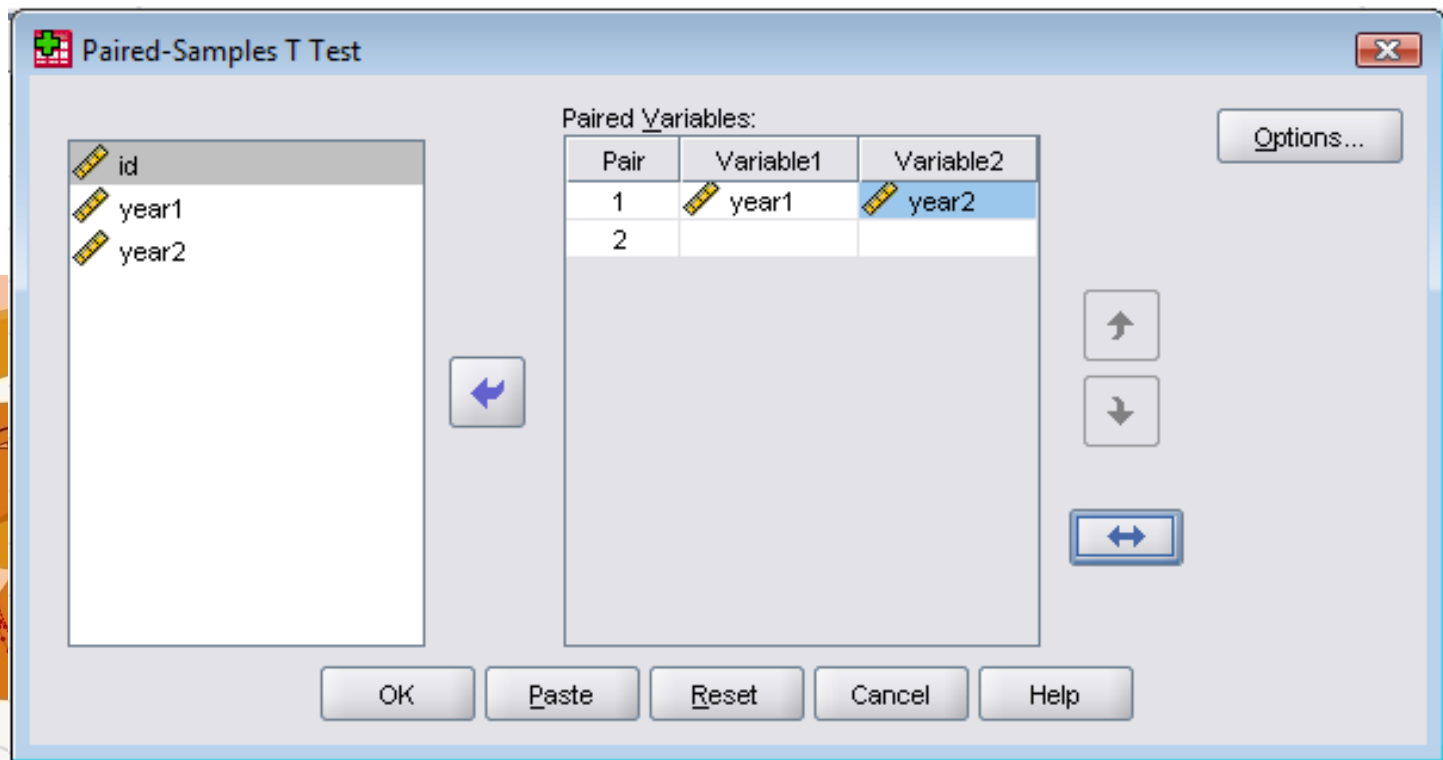
The screenshot shows the SPSS Data Editor interface. The 'Analyze' menu is open, and the 'Compare Means' option is selected, which has opened a sub-menu. In this sub-menu, 'Paired-Samples T Test...' is highlighted. The background shows a data table with columns 'id' and 'year'.

	id	year
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

การทดสอบสมมติฐาน

- เช่น



การทดสอบสมมติฐาน

- เช่น

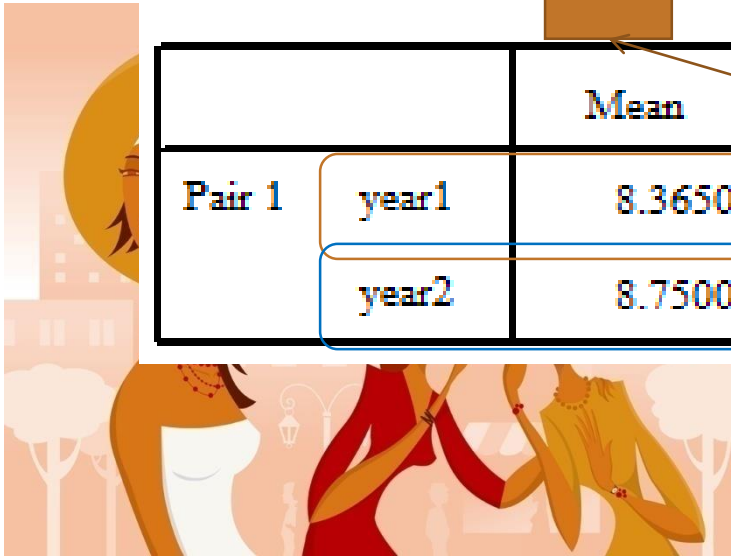
ค่าสถิติพรรณนาของนิสิตจุฬาฯ 20 คน

ปี 1

Paired Samples Statistics

ปี 2

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	year1	8.3650	20	.91610	.20485
	year2	8.7500	20	.80033	.17896



การทดสอบสมมติฐาน

- เช่น

หาโอกาสที่ความแตกต่างเกิดจากการสุ่ม

ค่าเฉลี่ยของความแตกต่าง (M_d)

$$t(19) = -4.50, p < .001$$

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 year1 - year2	-0.38500	.38289	.08562	-.56420	-.20580	-4.497	19	.000

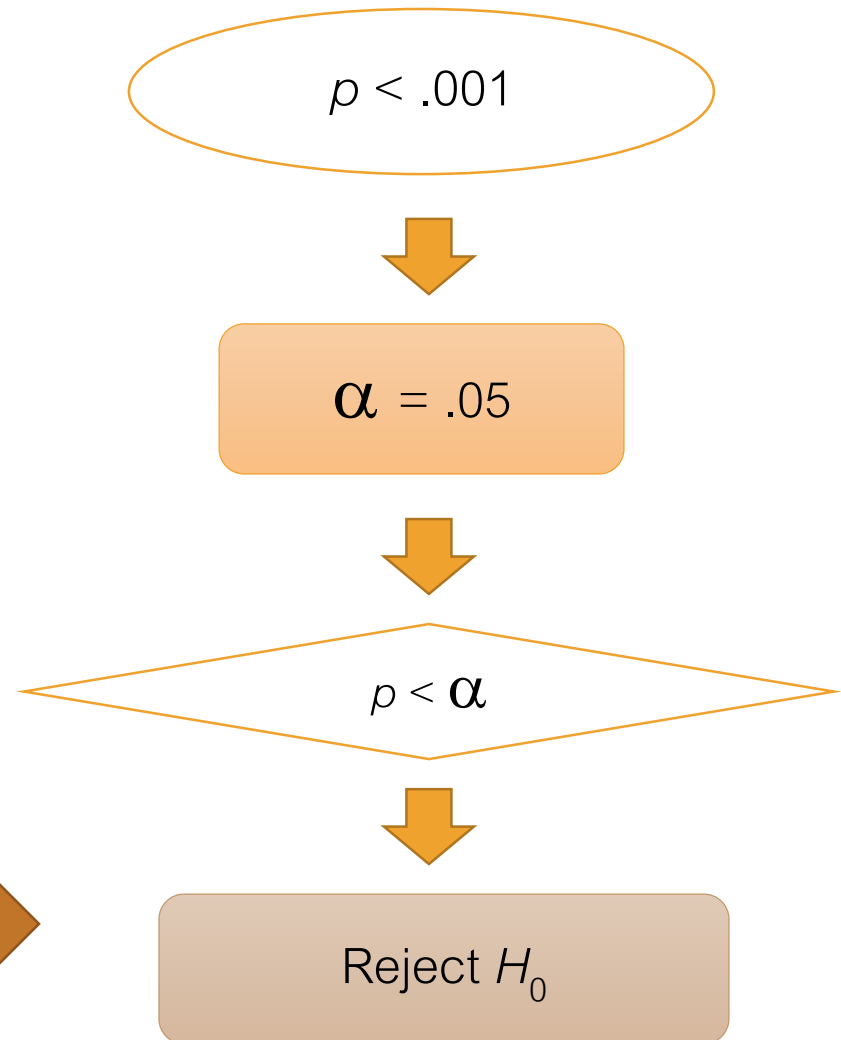
การทดสอบสมมติฐาน

- เช่น

นิติตจุฟ้าฯ เข้าปี 2
จะมีความสุขเพิ่มขึ้นจากปี 1
การวัดความสุขไล่ตั้งแต่ 0-10



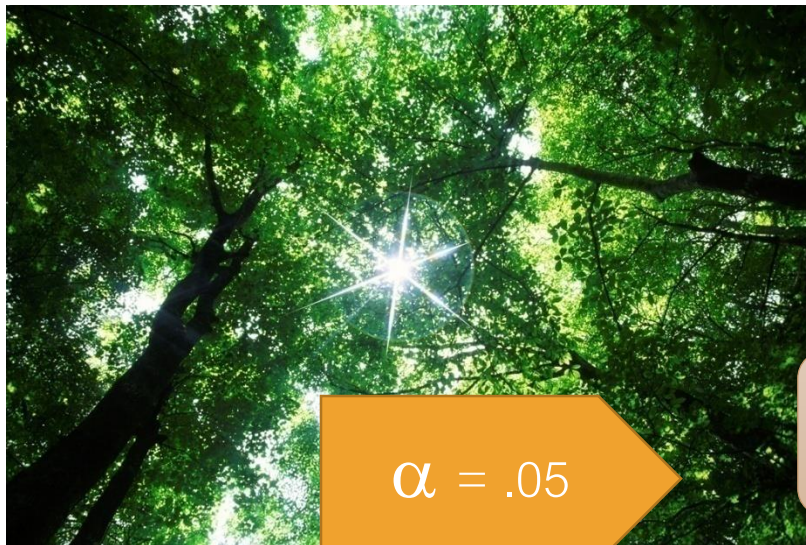
นิติตจุฟ้าฯ พอขึ้นปี 2 จะมีความสุขมากขึ้นกว่าตอนปี 1



การทดสอบสมมติฐาน

- เช่น

ฝาแฝดที่ได้รับการเลี้ยงดู
ในสภาพแวดล้อมต่างกัน (ธรรมชาติ,
มลพิษ) จะมี EQ ต่างกัน



นักวิจัยจึงเก็บข้อมูลจากฝาแฝดที่ถูกแยกเลี้ยงดู
10 คู่ แล้ววัด EQ

สมมติฐาน คือ ฝาแฝดที่เลี้ยงดู
ในธรรมชาติจะมี EQ สูงกว่า
เลี้ยงดูในที่ที่มีมลพิษ

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

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

การทดสอบสมมติฐาน

- เช่น

ค่าสถิติพรรณนาของฝาแฝดทั้ง 10 คู่

เสียงดูในสถานที่ที่มีมลพิษ

เสียงดูในสถานที่ธรรมชาติ

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	pollute	55.4400	10	8.35320	2.64151
	nature	55.6700	10	6.76883	2.14049



การทดสอบสมมติฐาน

- เช่น

หาโอกาสที่ความแตกต่างเกิดจากการสุ่ม

ค่าเฉลี่ยของความแตกต่าง (M_d)

$p = .788$ (two-tailed)

$p = .399$ (one-tailed)

$t(9) = -0.28, p > .05$ (one-tailed)



Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	pollute - nature	-0.23000	2.62978	.83161	-2.11123	1.65123	-0.277	9	.788



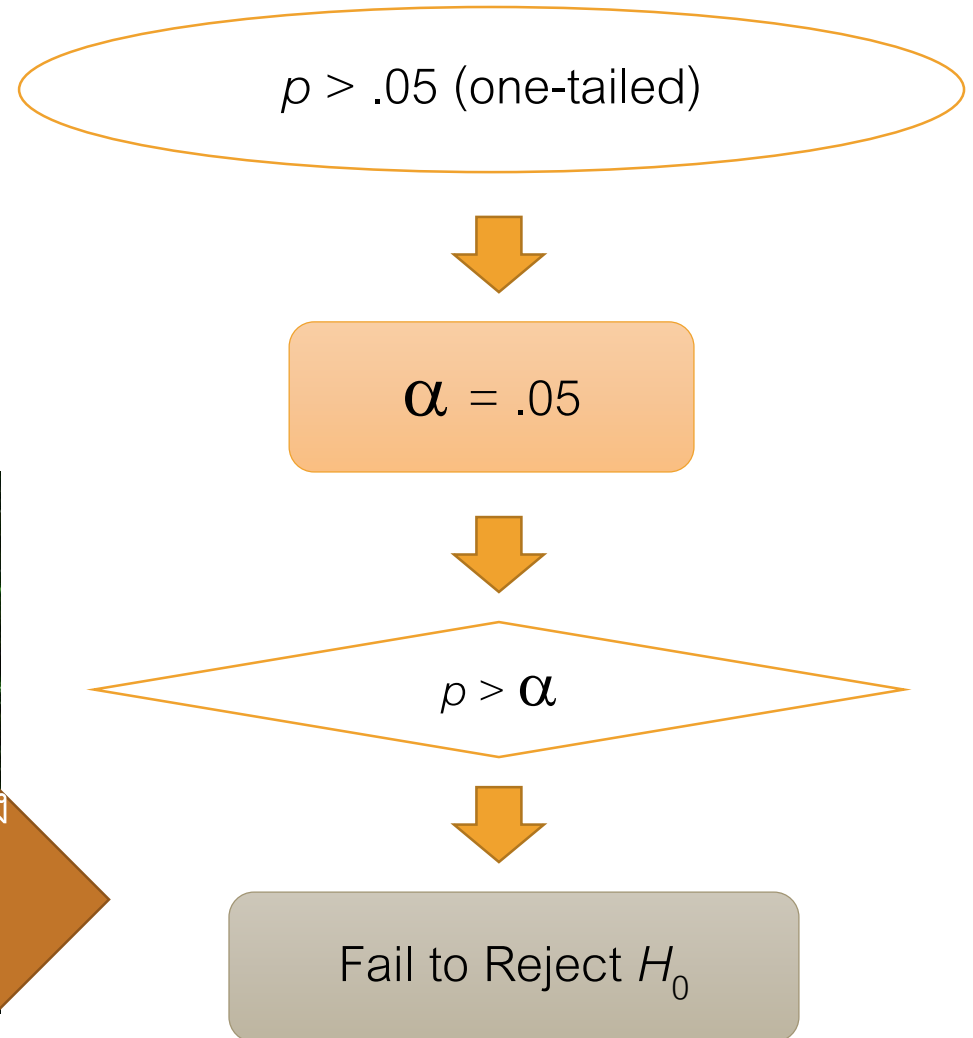
การทดสอบสมมติฐาน

- เช่น

ฝาแฝดที่ได้รับการเลี้ยงดู
ในสภาพแวดล้อมต่างกัน (ธรรมชาติ,
มลพิษ) จะมี EQ ต่างกัน



ไม่สามารถสรุปได้ว่า ฝาแฝดคนที่เลี้ยงดู
ตามธรรมชาติ จะมี EQ สูงกว่าคนที่
เลี้ยงดูในสถานที่ที่มีมลพิษ



ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- ประชากรของค่าเปลี่ยนแปลงมีการกระจายเป็นโค้งปกติ
 - สถิติมีความคงทนต่อการละเมิดข้อนี้
- ในกรณีการออกแบบการทดลองแบบจับคู่ (Matching Design) การคัดเข้ากลุ่มทั้งสองจะต้องเป็นแบบสุ่ม (Random Assignment)

การหาช่วงเชื่อมั่นสำหรับค่าเปลี่ยนแปลง

- ผู้วิจัยอาจต้องการทราบว่าการเปลี่ยนแปลงที่เกิดขึ้นในประชากรมีค่าเท่าไร
 - เช่น คอรัลการลดน้ำหนักนี้ ทำให้ประชากรกลุ่มเป้าหมายมีน้ำหนักลดลงกี่กิโลกรัม
- คำถามนี้จะสามารถใช้ช่วงเชื่อมั่นสำหรับค่าเปลี่ยนแปลงระหว่างคะแนนสองชุด (เช่น ก่อนและหลัง) ในการตอบคำถามได้

การหาช่วงเชื่อมั่นสำหรับค่าเปลี่ยนแปลง

- ช่วงเชื่อมั่นสำหรับค่าเปลี่ยนแปลง สามารถใช้ในการทดสอบสมมติฐานได้
- หากช่วงเชื่อมั่นครอบคลุมค่า 0 แสดงว่าคะแนนสองชุดไม่แตกต่างกันอย่างมีนัยสำคัญ (Fail to Reject H_0)
- แต่เมื่อช่วงเชื่อมั่นไม่ครอบคลุมค่า 0 แสดงว่าคะแนนสองชุดแตกต่างกันอย่างมีนัยสำคัญ (Reject H_0)
- ผลการวิเคราะห์ข้อมูลจากการทดสอบค่าที่แบบไม่อิสระ (Dependent t -test) โดย SPSS มีข้อมูลช่วงเชื่อมั่นด้วย

การหาช่วงเชื่อมั่นสำหรับค่าเปลี่ยนแปลง

- เช่น นักจิตวิทยาตรวจสอบการนอน 4 ชั่วโมงแรก
 - ร้อยละของการนอนแบบคลื่นสมองช้า (Slow Wave Sleep; SWS)
 - ร้อยละของการนอนแบบตาเคลื่อนไหว (Rapid Eye Movement Sleep; REM)
- นักจิตวิทยาต้องการทราบว่าในช่วงดังกล่าว ร้อยละของการนอนทั้งสองแบบแตกต่างกันมากน้อยเพียงใด ในประชากร

การหาช่วงเชื่อมั่นสำหรับค่าเปลี่ยนแปลง

นักจิตวิทยาเก็บข้อมูลจาก 58 คน
วัดคลื่นสมองในการนอน 4 ชั่วโมง

แรกและตรวจสอบว่ามี SWS
และ REM อย่างไร

$$M_{\text{SWS}} - M_{\text{REM}} = 24.04 - 6.56 = 17.48$$

ร้อยละการนอนแบบ SWS

ร้อยละการนอนแบบ REM

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Slow-Wave Sleep Percentage During Sleep	24.0442	58	12.40575	1.62895
	Rapid Eye Movement Percentage During Sleep	6.5622	58	4.86524	.63884



การหาช่วงเชื่อมั่นสำหรับค่าเปลี่ยนแปลง

นักจิตวิทยาเก็บข้อมูลจาก 58 คน
วัดคลื่นสมองในการนอน 4 ชั่วโมง
แรกและตรวจสอบว่ามี SWS
และ REM อย่างไร



$$M_{\text{SWS}} - M_{\text{REM}} = 17.48$$

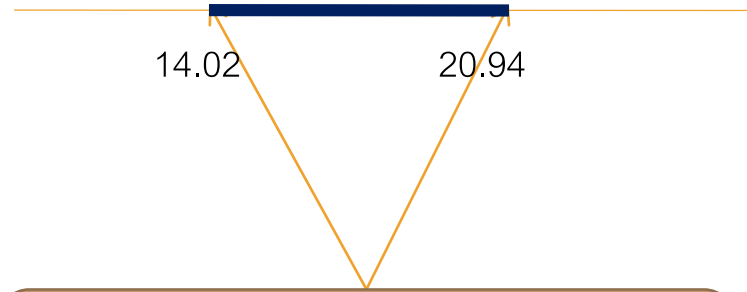
$$CI_{.95} \text{ ของ } \mu_{\text{SWS}} - \mu_{\text{REM}} \\ = (14.02, 20.94)$$

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Slow-Wave Sleep Percentage During Sleep - Rapid Eye Movement Percentage During Sleep	17.48198	13.14999	1.72668	14.02437	20.93960	10.125	57	.000

การหาช่วงเชื่อมั่นสำหรับค่าเปลี่ยนแปลง

นักจิตวิทยาเก็บข้อมูลจาก 58 คน
วัดคลื่นสมองในการนอน 4 ชั่วโมง
แรกและตรวจสอบว่ามี SWS
และ REM อย่างไร



จากข้อมูลทำให้มีความเชื่อมั่นระดับ .95 ว่าประชากร
ของความแตกต่างระหว่างร้อยละการเกิด SWS และ
REM ในช่วงการนอน 4 ชั่วโมงแรก อยู่ในช่วง 14.02 %
ถึง 20.94 %

การหาขนาดอิทธิพล

- การหาขนาดอิทธิพล ใช้หลักการเดียวกับการหาอิทธิพลในการทดสอบ z คือ

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

- ค่าเฉลี่ยจะแทนด้วยค่าจากกลุ่มตัวอย่างจากคะแนนทั้งสองชุด

การหาขนาดอิทธิพล

- อย่างไรก็ตาม ส่วนเบี่ยงเบนมาตรฐานจากประชากร จะมีแนวคิดอยู่สามรูปแบบด้วยกัน
 1. ใช้ส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุม หรือกลุ่มที่ไม่ได้รับการแทรกแซง (SD_C)
 2. ใช้ส่วนเบี่ยงเบนมาตรฐานรวมของคะแนนทั้งสองชุด (SD_{Pooled})
 3. ใช้ส่วนเบี่ยงเบนมาตรฐานของค่าเปลี่ยนแปลง (SD_D)

การหาขนาดอิทธิพล (SD กลุ่มควบคุม)

- เช่น ความแตกต่างของความสวยของปีหนึ่งและปีสอง
- คะแนนของปี 1 เหมาะสมที่จะใช้เป็นกลุ่มควบคุม

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	year1	8.3650	20	.91610	.20485
	year2	8.7500	20	.80033	.17896

$$SD_1 = 0.92$$

$$d = \frac{\mu_1 - \mu_2}{\sigma} = \frac{8.75 - 8.36}{0.92} = 0.42$$

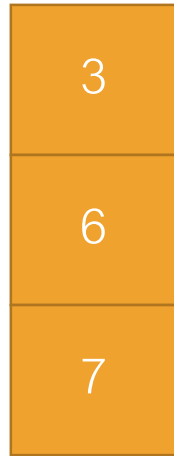
การหาขนาดอิทธิพล (SD ร่วม)

- ส่วนเบี่ยงเบนมาตรฐานรวม (Pooled SD) คือ นำส่วนเบี่ยงเบนมาตรฐานของทั้งสองกลุ่มมาถ่วงเฉลี่ย



$$SD_1 = 1.53$$

$$df_1 = 2$$



$$SD_2 = 2.08$$

$$df_2 = 2$$

$$SD_{\text{Pooled}} = \sqrt{\frac{SS_1 + SS_2}{df_1 + df_2}}$$

$$SS_1 = SD_1^2 \times df_1 = (1.53)^2 \times 2 = 4.67$$

$$SS_2 = SD_2^2 \times df_2 = (2.08)^2 \times 2 = 8.67$$

การหาขนาดอิทธิพล (SD ร่วม)

- ส่วนเบี่ยงเบนมาตรฐานรวม (Pooled SD) คือ นำส่วนเบี่ยงเบนมาตรฐานของทั้งสองกลุ่มมาถ่วงเฉลี่ย



$$SD_1 = 1.53$$

$$df_1 = 2$$



$$SD_2 = 2.08$$

$$df_2 = 2$$

$$SD_{\text{Pooled}} = \sqrt{\frac{SS_1 + SS_2}{df_1 + df_2}}$$

$$SD_{\text{Pooled}} = \sqrt{\frac{4.67 + 8.67}{2 + 2}} = 1.83$$

การหาขนาดอิทธิพล (SD ร่วม)

- เช่น ความแตกต่างของความสวยของปีหนึ่งและปีสอง

$$SS_1 = 0.92 \times 0.92 \times 19 = 15.95$$

$$SS_2 = 0.80 \times 0.80 \times 19 = 12.17$$

Two-Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	year1	8.3650	20	.91610	.20485
	year2	8.7500	20	.80033	.17896

$$SD_{\text{Pooled}} = \sqrt{\frac{15.95 + 12.17}{19 + 19}} = 0.86$$

$$d = \frac{8.75 - 8.36}{0.86} = 0.45$$

การหาขนาดอิทธิพล (SD ของค่าเปลี่ยนแปลง)

- เช่น ความแตกต่างของความสวยของปีหนึ่งและปีสอง

ส่วนเบี่ยงเบนมาตรฐาน
ของความแตกต่าง (SD_d)

$$d = \frac{8.75 - 8.36}{0.38} = 1.02$$

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 year1 - year2	-.38500	.38289	.08562	-.56420	-.20580	-4.497	19	.000

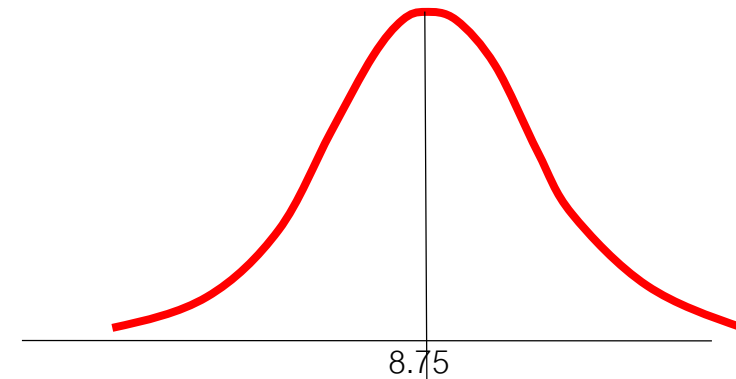
การหาขนาดอิทธิพล

- จะเห็นว่า การใช้ส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุม (SD_C) หรือส่วนเบี่ยงเบนมาตรฐานรวม (SD_{Pooled}) จะมีขนาดอิทธิพลใกล้เคียงกัน
- สองตัวนี้จะบอกว่า ขนาดความแตกต่างนั้นเป็นกี่เท่าของการกระจายของคะแนนปกติ
- เช่น การลดน้ำหนัก ทำให้ลดน้ำหนักได้เป็นกี่เท่าของส่วนเบี่ยงเบนมาตรฐานของน้ำหนักเดิม
- ขนาดอิทธิพลนี้ เหมือนกับขนาดอิทธิพลในการเปรียบเทียบค่าเฉลี่ยที่เป็นอิสระจากกัน

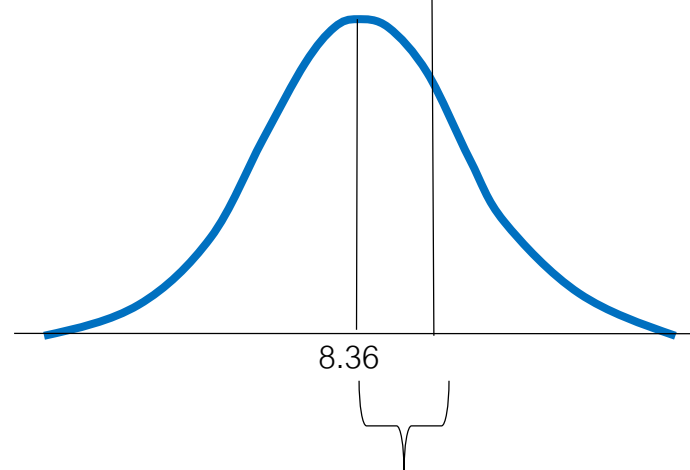
การหาขนาดอิทธิพล

ขนาดอิทธิพลแบบ
ที่ 1 และ 2

ปีสอง



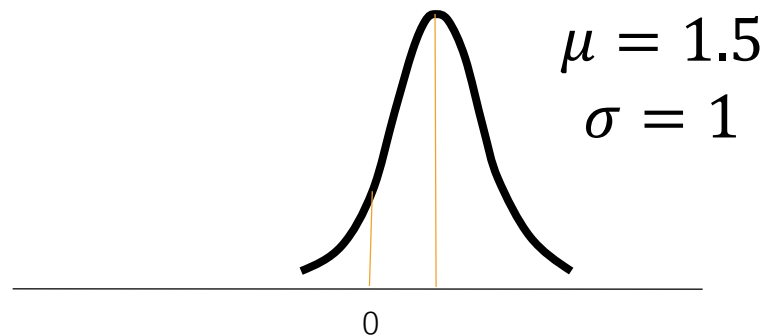
ปีหนึ่ง



ความแตกต่างเทียบเป็น 0.42 เท่าของ SD กลุ่มควบคุม

การหาขนาดอิทธิพล

- แต่ การใช้ส่วนเบี่ยงเบนมาตรฐานจากการเปลี่ยนแปลงนั้น จะบอกว่า ขนาดความแตกต่างนั้น เป็นกี่เท่าของการกระจายคะแนนเปลี่ยนแปลง (Difference Score)
- จริงๆ แล้ว ก็คือบอกว่า 0 คิดเป็นค่า z score เท่ากับเท่าไร ในการกระจายของค่าเปลี่ยนแปลง



$$z_d = \frac{0 - 1.5}{1} = -1.5$$

d โดยใช้ SD ของค่าเปลี่ยนแปลง = 1.5

การหาขนาดอิทธิพล

- ปกติ จะใช้ส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุม หรือส่วนเบี่ยงเบนมาตรฐานร่วม (แบบที่ 1 หรือ 2)
- ระดับของขนาดอิทธิพลเท่ากับ 0.2, 0.5, และ 0.8 เทียบเท่ากับขนาดอิทธิพลน้อย, ปานกลาง, มาก สามารถนำไปใช้ได้
- แต่ทว่า การแปลความหมายโดยใช้ส่วนเบี่ยงเบนมาตรฐานของค่าเปลี่ยนแปลงนั้น ทำค่อนข้างยาก
- ส่วนใหญ่จะใช้ในการหาลำดับในการทดสอบทางสถิติเท่านั้น

การเขียนรายงาน

- เช่น งานวิจัยนี้ต้องการทดสอบว่า มลภาวะส่งผลให้ความฉลาดทางอารมณ์ (Emotional Intelligence; EQ) เปลี่ยนแปลงหรือไม่ จากการทดสอบจากฝาแฝดจำนวน 10 คู่ ที่อยู่แยกเลี้ยงดูแตกต่างกัน พบว่าฝาแฝดคนที่เลี้ยงดูในสิ่งแวดล้อมสะอาดและมีมลภาวะมีคะแนนความฉลาดทางอารมณ์เท่ากับ 55.67 และ 55.64 ตามลำดับ ($SD = 6.77$ และ 8.35) จากการทดสอบ t แบบเกี่ยวข้อกัน (Dependent t -test) พบว่าค่าเฉลี่ยไม่แตกต่างกันอย่างมีนัยสำคัญ ($t(9) = -0.28, p > .05, d^a = 0.004$)

^a การคำนวณขนาดอิทธิพลใช้วิธีการหารด้วย SD ของกลุ่มสิ่งแวดล้อมสะอาด

การหาคำสั่งของความเปลี่ยนแปลง

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

การหาค่ากำลัง

- เช่น ทดสอบกี่ทาง ขนาดอิทธิพลแบบที่สาม ระดับนัยสำคัญ ขนาดกลุ่มตัวอย่าง

Test family: t tests

Statistical test: Means: Difference between two dependent means (matched pairs)

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

Input Parameters

Tail(s)	Two
Effect size dz	0.49
α err prob	0.05
Total sample size	120

Output Parameters

Noncentrality parameter δ	5.367681
Critical t	1.980100
Df	119
Power ($1 - \beta$ err prob)	0.999616

ค่าพารามิเตอร์ที่เบี่ยงเบน

กำลังในการทดสอบสมมติฐานทางสถิติ

การหาค่า d_z

- สังเกตว่าขนาดอิทธิพลที่ใช้ คือ ขนาดอิทธิพลแบบที่สาม ที่ใช้ส่วนเบี่ยงเบนมาตรฐานของการเปลี่ยนแปลง (SD_d)
- ในกรณีที่ไม่สามารถคำนวณได้ ให้กด Determine และใส่ M และ SD ของคะแนนทั้งสองชุด รวมถึงค่าสหสัมพันธ์ของคะแนนทั้งสองชุด

The screenshot shows a dialog box with two main sections: "From differences" and "From group parameters". The "From group parameters" section is selected with a radio button. It contains the following input fields and values:

Parameter	Value
Mean of difference	0
SD of difference	1
Mean group 1	0
Mean group 2	0.5
SD group 1	1
SD group 2	1
Correlation between groups	0.48

At the bottom, there is a "Calculate" button, a text field for "Effect size dz" containing the value 0.4902903, a "Calculate and transfer to main window" button, and a "Close" button.

การหาค่าล้ง

- สังเกตว่า สิ่งที่มีผลกระทบต่อขนาดอิทธิพลมี 3 ปัจจัย
 - ความแตกต่างระหว่างค่าเฉลี่ย
 - ส่วนเบี่ยงเบนมาตรฐานของแต่ละกลุ่ม
 - ความสัมพันธ์ของคะแนนทั้งสอง

ลองเปลี่ยนแปลงค่าสหสัมพันธ์ ขณะที่ $\mu_1 = 0, \mu_2 = 0.5, \sigma_1 = 1, \sigma_2 = 1$

ρ	d	ρ	d	ρ	d
.8	.79	.2	.40	-.4	.30
.6	.56	0	.35	-.6	.28
.4	.46	-.2	.32	-.8	.26

การหากำลัง

- ดังนั้น ความสัมพันธ์ระหว่างตัวแปรทั้งสองยิ่งสูง ขนาดอิทธิพลยิ่งสูง ส่งผลให้กำลังในการทดสอบทางสถิติสูง
- ในการออกแบบการทดลอง การทดสอบค่าเฉลี่ยแบบเกี่ยวข้งกัน มักมีกำลังในการทดสอบทางสถิติมากกว่าการทดสอบค่าเฉลี่ยที่ไม่เกี่ยวข้ง
 - โดยเฉพาะอย่างยิ่ง ความสัมพันธ์สูง
 - รายละเอียดกล่าวในภายหลัง

การกำหนดขนาดกลุ่มตัวอย่าง

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

การกำหนดขนาดกลุ่มตัวอย่าง

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

การกำหนดขนาดกลุ่มตัวอย่าง

- เช่น จงหาขนาดกลุ่มตัวอย่างที่จะทำให้กำลัง = .80 ในการทดสอบสองทาง และระดับนัยสำคัญเท่ากับ .05 เมื่อขนาดอิทธิพลแบบที่สามเท่ากับ 0.49

The screenshot shows a statistical software interface for power analysis. It is divided into several sections:

- Test family:** t tests
- Statistical test:** Means: Difference between two dependent means (matched pairs)
- Type of power analysis:** A priori: Compute required sample size - given α , power, and effect size
- Input Parameters:**
 - Tail(s): Two
 - Effect size dz: 0.49
 - α err prob: 0.05
 - Power (1- β err prob): 0.8
- Output Parameters:**
 - Noncentrality parameter δ : 2.898879
 - Critical t: 2.032245
 - Df: 34
 - Total sample size: 35** (circled in orange)
 - Actual power: 0.804049

A "Determine =>" button is located in the Input Parameters section.

การกำหนดขนาดกลุ่มตัวอย่าง

- อย่างไรก็ตาม เกณฑ์ขนาดอิทธิพล 0.2, 0.5, และ 0.8 ในขนาดอิทธิพลที่คำนวณจากส่วนเบี่ยงเบนมาตรฐานของความเปลี่ยนแปลง อาจไม่ตรงกับที่คำนวณจากคะแนนดิบ
- แนะนำให้กด Determine และใส่ค่า M , SD ของทั้งสองกลุ่ม (อาจใส่ค่าให้ได้เทียบเท่ากับขนาดอิทธิพลแบบที่หนึ่ง) และค่าสหสัมพันธ์ลงไป

The screenshot shows a dialog box with two main sections: "from differences" and "from group parameters". The "from group parameters" section is selected. It contains input fields for Mean group 1 (5), Mean group 2 (0), SD group 1 (10), SD group 2 (10), and Correlation between groups (0). Below these fields are buttons for "Calculate", "Calculate and transfer to main window", and "Close". The "Effect size dz" is displayed as 0.3535534.

Parameter	Value
Mean of difference	0
SD of difference	1
Mean group 1	5
Mean group 2	0
SD group 1	10
SD group 2	10
Correlation between groups	0
Effect size dz	0.3535534

การกำหนดขนาดคกลุ่มตัวอย่าง

- หากไม่ทราบค่าสหสัมพันธ์ ให้ใส่ 0 ลงไป
เปรียบเทียบได้ว่าตัวแปรทั้งสองชุดไม่สัมพันธ์กัน
 - ซึ่งเป็นเกณฑ์ที่เผื่อจำนวนกลุ่มตัวอย่างมากกว่าปกติ
 - ในการทดสอบส่วนใหญ่ คะแนนทั้งสองจะ
ไม่มีความสัมพันธ์เป็นลบ

from differences

Mean of difference 0

SD of difference 1

from group parameters

Mean group 1 5

Mean group 2 0

SD group 1 10

SD group 2 10

Correlation between groups 0

Calculate Effect size dz 0.353534

Calculate and transfer to main window

Close

การเปรียบเทียบการทดสอบที่ทั้งสองแบบ

- การเก็บข้อมูลแบบวัดซ้ำมีข้อดี คือ
 - ใช้จำนวนกลุ่มตัวอย่างน้อยกว่า
 - กำลังในการทดสอบสูงกว่า
- การเก็บข้อมูลแบบวัดซ้ำมีข้อเสีย คือ
 - ใช้ระยะเวลานาน
 - กลุ่มตัวอย่างอาจถอนตัวจากงานวิจัย ทำให้ข้อมูลสูญหาย (Attrition)
 - ผลจากการจำ ความเหนื่อยล้า (Carry over effect)

การเปรียบเทียบการทดสอบที่ทั้งสองแบบ

- ทำไมกำลังของ Dependent t -test สูงกว่า
 - เปรียบเทียบค่าความผิดพลาดมาตรฐาน (Standard Error)

สูตรทั่วไปของค่า t

$$t = \frac{\text{Mean Diff}}{\text{Standard Error}}$$

Independent t -test

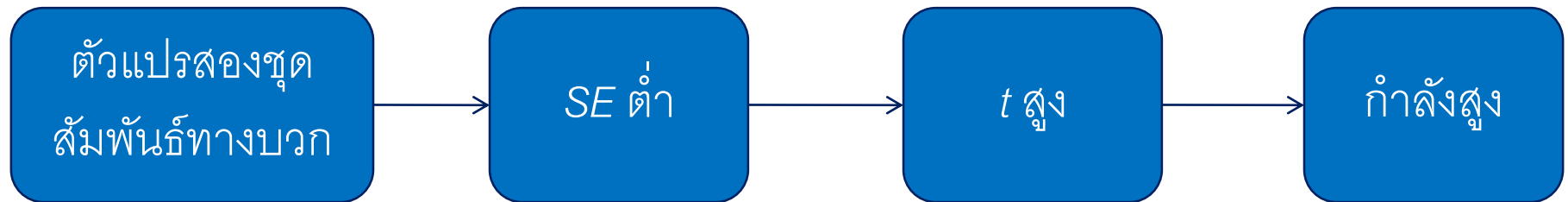
$$SE = \frac{1}{n} \sqrt{s_1^2 + s_2^2}$$

Dependent t -test

$$SE = \frac{1}{n} \sqrt{s_1^2 + s_2^2 - 2s_1s_2r_{12}}$$

การเปรียบเทียบการทดสอบที่ทั้งสองแบบ

- ทำไมกำลังของ Dependent t -test สูงกว่า
 - เปรียบเทียบค่าความผิดพลาดมาตรฐาน (Standard Error)



Dependent t -test

$$SE = \frac{1}{n} \sqrt{s_1^2 + s_2^2 - 2s_1s_2r_{12}}$$

การเปรียบเทียบการทดสอบที่ทั้งสองแบบ

- ผลจากการจำ ความเหนื่อยล้า (Carry over effect)
 - เช่น ผู้ร่วมการทดลองตั้งใจทำงานที่หนึ่ง แต่ไม่ตั้งใจทำงานที่สอง
 - ผู้ร่วมการทดลองสามารถจำคำที่ใช้จากงานที่หนึ่ง จึงทำให้คะแนนของงานที่สองสูงกว่า
- วิธีแก้ คือ การสลับลำดับ (Counter Balancing)
 - เช่น บางคนให้ทำงานที่หนึ่งก่อน และบางคนให้ทำงานที่สองก่อน