

การวิเคราะห์ความแปรปรวน แบบทางเดียว

สถิติชั้นกลางสำหรับจิตวิทยา

สันหัตถ์ พรประเสริฐมานิต

โครงร่างการนำเสนอ

- โอกาสเกิดความผิดพลาดทั้งกลุ่ม (Familywise Error Rate)
- การปรับค่า Alpha รายคู่
- การทดสอบความแปรปรวนแบบทางเดียว
- การเปรียบเทียบรายคู่
- รูปแบบการขยายผล

โอกาสเกิดความผิดพลาดทั้งกลุ่ม

สถิติชั้นกลางสำหรับจิตวิทยา

สันหัต พรประเสริฐมานิต

แนะนำ

- ในวิชาที่ผ่านมา เราพูดถึงการเปรียบเทียบค่าเฉลี่ยระหว่าง 2 กลุ่ม
- ในบทนี้ จะแนะนำการเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มมากกว่า 2 กลุ่ม
- สถิติหลักที่ใช้ในบทนี้ คือ การทดสอบความแปรปรวน (Analysis of Variance; ANOVA)

โอกาสเกิดความผิดพลาดทั้งกลุ่ม

- บางครั้งนักวิจัยต้องการเปรียบเทียบค่าเฉลี่ยมากกว่า 2 กลุ่ม
 - เช่น การตรวจสอบประสิทธิภาพของยา โดยมีกลุ่มทดลอง 1 กลุ่ม กลุ่มควบคุม 2 กลุ่ม กลุ่มแรกทานยาหลอก (Placebo) และกลุ่มที่สองไม่ได้รับยาอะไร
 - เช่น เปรียบเทียบความรู้ทางดนตรี ระหว่าง นักร้อง นักเต้น นักดนตรี และคนธรรมดา
- การเปรียบเทียบอาจสามารถทำได้ด้วยการเปรียบเทียบรายคู่
 - 3 กลุ่ม เปรียบเทียบ 3 ครั้ง: 1-2, 1-3, 2-3
 - 4 กลุ่ม เปรียบเทียบ 6 ครั้ง: 1-2, 1-3, 1-4, 2-3, 2-4, 3-4

โอกาสเกิดความผิดพลาดทั้งกลุ่ม

- ถ้าเปรียบเทียบรายคู่ด้วย $\alpha = .05$
- หากสุ่มกลุ่มตัวอย่างจากประชากรเดียวกันออกมา 3 กลุ่ม
 - การเปรียบเทียบรายคู่ ควรจะไม่แตกต่างอย่างมีนัยสำคัญ เพราะสุ่มจากประชากรเดียวกัน ค่าเฉลี่ยของประชากรแต่ละกลุ่มต้องเท่ากัน
 - หากตั้ง $\alpha = .05$ แล้ว โอกาสที่ผลออกมาว่ามีค่าเฉลี่ยแตกต่างกัน ควรจะเท่ากับ 5%

โอกาสเกิดความผิดพลาดทั้งกลุ่ม

- แต่ทว่า

$$p_{1-2}(\text{sig}) = .05$$

$$p_{1-2}(\text{not sig}) = .95$$

$$p_{1-3}(\text{sig}) = .05$$

$$p_{1-3}(\text{not sig}) = .95$$

$$p_{2-3}(\text{sig}) = .05$$

$$p_{2-3}(\text{not sig}) = .95$$

หากการเกิดนัยสำคัญของแต่ละคู่เป็นอิสระจากกัน

$$\begin{aligned} &\text{โอกาสที่ทุกคู่ไม่แตกต่าง} \\ &\text{อย่างมีนัยสำคัญ} = \\ &(.95)^3 = .857 \end{aligned}$$

$$\begin{aligned} &\text{โอกาสที่แตกต่างอย่างมี} \\ &\text{นัยสำคัญอย่างน้อย 1 คู่} \\ &= 1 - .857 = .143 \end{aligned}$$

โอกาสที่พบความแตกต่างอย่างมีนัยสำคัญ (แตกต่างอย่างน้อย 1 คู่) = 14% > 5%

โอกาสเกิดความผิดพลาดทั้งกลุ่ม

- โอกาสที่เจอความแตกต่างอย่างน้อย 1 คู่ที่ Null Hypothesis เป็นจริง จะเรียกว่า โอกาสเจอความผิดพลาดทั้งกลุ่ม (Familywise Error Rate; α_{FW})
- หากการเกิดนัยสำคัญของแต่ละคู่เป็นอิสระจากกัน

$$\alpha_{FW} = 1 - (1 - \alpha)^C$$

- โอกาสที่แต่ละคู่เป็นอิสระจากกัน แทบเป็นไปไม่ได้
 - เช่น หากกลุ่มที่ 1 คะแนนเท่ากับกลุ่มที่ 2 การเปรียบเทียบคู่ 1-3 และคู่ 2-3 แทบจะสอดคล้องไปด้วยกัน
 - ดังนั้นสูตรข้างบนจึงไม่ถูกต้องทีเดียว

โอกาสเกิดความผิดพลาดทั้งกลุ่ม

- ประเด็นสำคัญคือ เราจะรู้ขนาดการเกิด Familywise Error Rate ได้อย่างไร และจะควบคุมอย่างไร
 - การเปลี่ยน α ในการเปรียบเทียบรายคู่ เพื่อให้ $\alpha_{FW} = .05$ (หรือค่าอื่นที่น้อยกว่า .05)
 - การวิเคราะห์การแปรปรวน (Analysis of variance)

การปรับค่า Alpha รายคู่

สถิติขั้นกลางสำหรับจิตวิทยา

สันหัด พรประเสริฐมานิต

การปรับค่า Alpha รายคู่

- วิธีการปรับค่า α ของแต่ละคู่ (Error rate per comparison; α_{PC}) เพื่อควบคุม Familywise error rate
 - การแก้ไขแบบ Sidak
 - การแก้ไขแบบ Bonferroni
 - การแก้ไขแบบ Holm

การปรับค่า Alpha รายคู่

- การแก้ไขแบบ Sidak

$$\alpha_{FW} = 1 - (1 - \alpha_{PC})^C$$



$$\alpha_{PC} = 1 - \sqrt[C]{1 - \alpha_{FW}}$$

- เช่น ในการเปรียบเทียบ 4 กลุ่ม มีการเปรียบเทียบรายคู่ 6 คู่ ดังนั้นแต่ละคู่ใช้ $\alpha_{PC} = 1 - \sqrt[C]{1 - 0.05} = .0085$
- การแก้ไขนี้ ใช้ข้อตกลงเบื้องต้นว่า การเปรียบเทียบรายคู่เป็นอิสระจากกัน
- การแก้ไขแบบ Hochberg เป็นอีกทางเลือกที่ใกล้เคียงกับของ Sidak

การปรับค่า Alpha รายคู่

- การแก้ไขแบบ Bonferroni

$$\alpha_{PC} = \frac{\alpha_{FW}}{C} < .05 \text{ (หรือค่าอื่น)}$$

- เช่น ในการเปรียบเทียบ 4 กลุ่ม มีการเปรียบเทียบรายคู่ 6 คู่ ดังนั้นแต่ละคู่ใช้
 $\alpha_{PC} = .05/6 = .0083$
- การแก้ไขนี้ การเปรียบเทียบรายคู่อาจเกี่ยวข้องกันได้

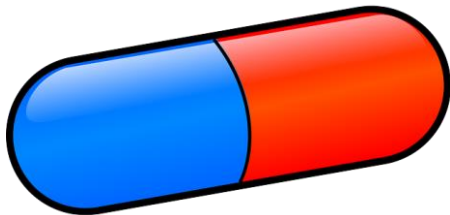
การปรับค่า Alpha รายคู่

- การแก้ไขแบบ Holm (หรือเรียกว่า Sequential Bonferroni)
 - นำ p-values ของการเปรียบเทียบทุกคู่ มาเรียงจากน้อยไปหามาก
 - นำ p-value ที่มีค่าน้อยที่สุด มาเปรียบเทียบกับ α_{FW}/C
 - หากไม่ถึงระดับนัยสำคัญ หยุด --- หากถึงระดับนัยสำคัญ เปรียบเทียบกับ p-value ที่มีค่าน้อยที่สุดลำดับที่ 2 ด้วย $\alpha_{FW}/(C - 1)$
 - หากไม่ถึงระดับนัยสำคัญ หยุด --- หากถึงระดับนัยสำคัญ เปรียบเทียบกับ p-value ที่มีค่าน้อยที่สุดลำดับที่ 3 ด้วย $\alpha_{FW}/(C - 2)$
 - ทำแบบนี้ไปเรื่อยๆ จนครบทุกคู่
- การเปรียบเทียบรายคู่อาจเกี่ยวข้อกันได้ มีกำลังทางสถิติมากกว่าวิธีของ Bonferroni

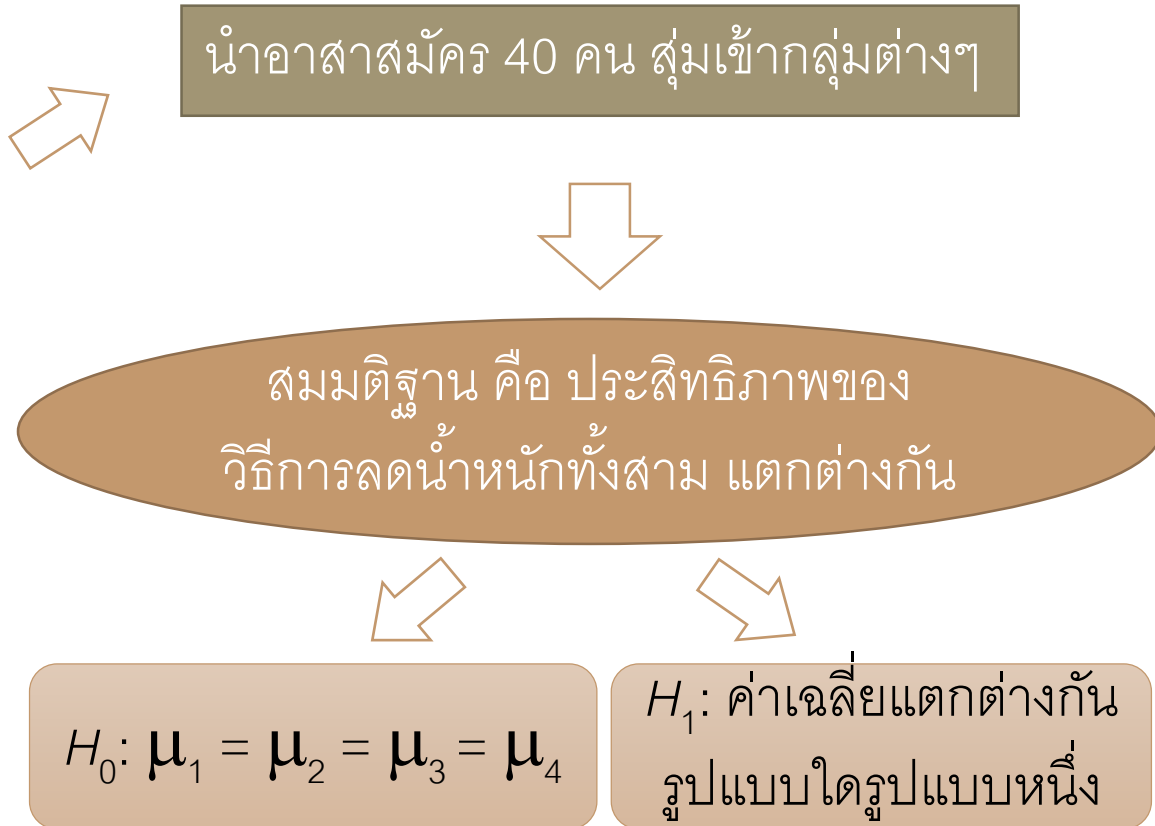
การปรับค่า Alpha รายคู่

- เช่น

เปรียบเทียบการลดน้ำหนัก
ด้วยยาลดน้ำหนักแบบใหม่,
ยาลดน้ำหนักแบบเก่า, ยาหลอก,
และไม่ทานยา กลุ่มละ 10 คน



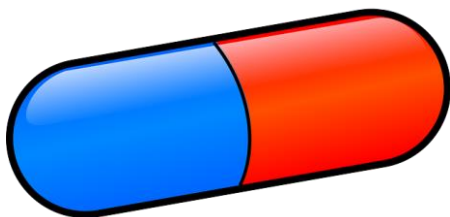
$$\alpha = .05$$



การปรับค่า Alpha รายคู่

- เช่น

เปรียบเทียบการลดน้ำหนัก
ด้วยยาลดน้ำหนักแบบใหม่,
ยาลดน้ำหนักแบบเก่า, ยาหลอก,
และไม่ทานยา กลุ่มละ 10 คน



$$\alpha = .05$$

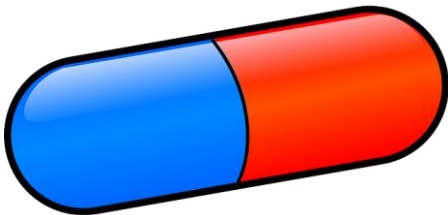
Report

weight			
group	N	Mean	Std. Deviation
New Drug	10	75.8000	10.10830
Old Drug	10	78.1000	8.14385
Placebo	10	94.4000	12.32162
No Drug	10	87.7000	12.10188
Total	40	84.0000	12.85022

การปรับค่า Alpha รายคู่

- เช่น

เปรียบเทียบการลดน้ำหนัก
ด้วยยาลดน้ำหนักแบบใหม่,
ยาลดน้ำหนักแบบเก่า, ยาหลอก,
และไม่ทานยา กลุ่มละ 10 คน



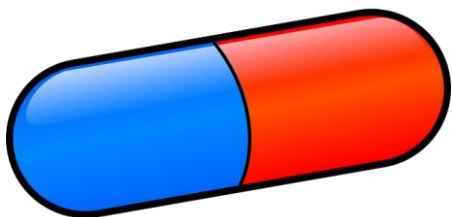
$$\alpha = .05$$

คู่	<i>t</i>	<i>df</i>	<i>p</i>
1-2	-0.56	18	.582
1-3	-3.69	18	.002
1-4	-2.39	18	.028
2-3	-3.49	18	.003
2-4	-2.08	18	.052
3-4	1.23	18	.236

การปรับค่า Alpha รายคู่

- เช่น

เปรียบเทียบการลดน้ำหนัก
ด้วยยาลดน้ำหนักแบบใหม่,
ยาลดน้ำหนักแบบเก่า, ยาหลอก,
และไม่ทานยา กลุ่มละ 10 คน



$$\alpha = .05$$

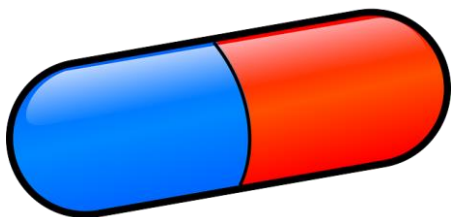
No Correction

คู่	p	α	Sig
1-2	.582	.05	NO
1-3	.002	.05	YES
1-4	.028	.05	YES
2-3	.003	.05	YES
2-4	.052	.05	NO
3-4	.236	.05	NO

การปรับค่า Alpha รายคู่

- เช่น

เปรียบเทียบการลดน้ำหนัก
ด้วยยาลดน้ำหนักแบบใหม่,
ยาลดน้ำหนักแบบเก่า, ยาหลอก,
และไม่ทานยา กลุ่มละ 10 คน



$$\alpha = .05$$

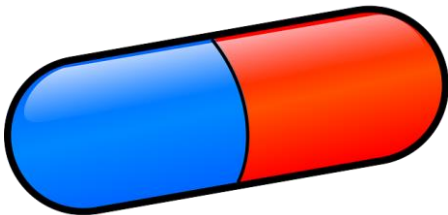
Sidak Correction

คู่	p	α	Sig
1-2	.582	.0085	NO
1-3	.002	.0085	YES
1-4	.028	.0085	NO
2-3	.003	.0085	YES
2-4	.052	.0085	NO
3-4	.236	.0085	NO

การปรับค่า Alpha รายคู่

- เช่น

เปรียบเทียบการลดน้ำหนัก
ด้วยยาลดน้ำหนักแบบใหม่,
ยาลดน้ำหนักแบบเก่า, ยาหลอก,
และไม่ทานยา กลุ่มละ 10 คน



$$\alpha = .05$$

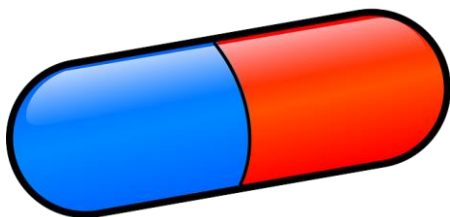
Bonferroni Correction

คู่	p	α	Sig
1-2	.582	.0083	NO
1-3	.002	.0083	YES
1-4	.028	.0083	NO
2-3	.003	.0083	YES
2-4	.052	.0083	NO
3-4	.236	.0083	NO

การปรับค่า Alpha รายคู่

- เช่น

เปรียบเทียบการลดน้ำหนัก
ด้วยยาลดน้ำหนักแบบใหม่,
ยาลดน้ำหนักแบบเก่า, ยาหลอก,
และไม่ทานยา กลุ่มละ 10 คน



$$\alpha = .05$$

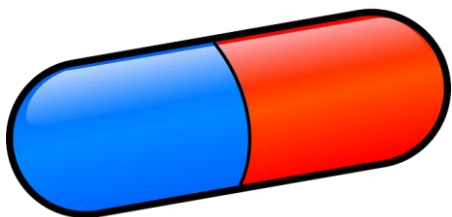
Holm Correction

คู่	p	Rank	α	Sig
1-2	.582	6		NO
1-3	.002	1	$.05/6 = .0083$	YES
1-4	.028	3	$.05/4 = .0125$	NO
2-3	.003	2	$.05/5 = .01$	YES
2-4	.052	4		NO
3-4	.236	5		NO

การปรับค่า Alpha รายคู่

- เช่น

เปรียบเทียบการลดน้ำหนัก
ด้วยยาลดน้ำหนักแบบใหม่,
ยาลดน้ำหนักแบบเก่า, ยาหลอก,
และไม่ทานยา กลุ่มละ 10 คน



$$\alpha = .05$$

สรุป

คู่	No	Sidak	Bonferroni	Holm
1-2	NO	NO	NO	NO
1-3	YES	YES	YES	YES
1-4	YES	NO	NO	NO
2-3	YES	YES	YES	YES
2-4	NO	NO	NO	NO
3-4	NO	NO	NO	NO

การทานยาลดน้ำหนักใหม่และยาลดน้ำหนักเก่า
ดีกว่ายาหลอก

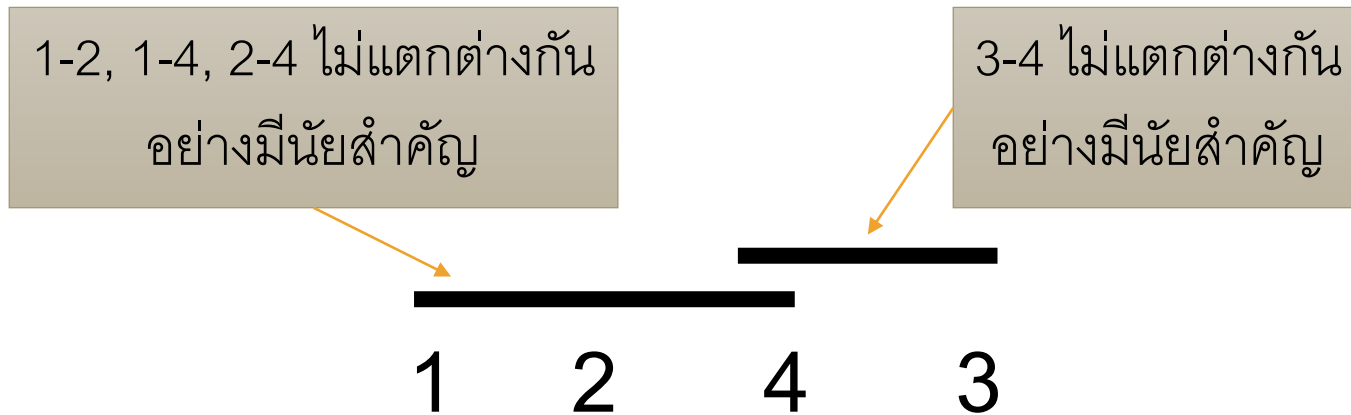
การปรับค่า Alpha รายคู่

- การเขียนรายงาน

- การวิจัยครั้งนี้ ต้องการประสิทธิภาพของยาลดน้ำหนักรใหม่ เปรียบเทียบกับกลุ่มควบคุม 3 กลุ่ม คือ ยาลดน้ำหนักรเก่า ยาหลอก และไม่ได้รับยา ทดสอบกลุ่มละ 10 คนด้วยการทดสอบ t แบบอิสระโดยปรับระดับนัยสำคัญด้วยวิธีของ Holm พบว่า การทานยาลดน้ำหนักรใหม่ ($M = 75.80$, $SD = 10.11$) และการทานยาลดน้ำหนักรเก่า ($M = 78.10$, $SD = 8.14$) ลดน้ำหนักรได้มากกว่าการทานยาหลอก ($M = 94.40$, $SD = 12.32$) อย่างมีนัยสำคัญทางสถิติ ขณะที่กลุ่มไม่ได้รับยา ($M = 87.70$, $SD = 12.10$) ไม่แตกต่างจากกลุ่มอื่นอย่างมีนัยสำคัญ

การปรับค่า Alpha รายคู่

- การเขียนสรุปการเปรียบเทียบรายคู่แบบแผนภูมิ
 - เรียงลำดับค่าเฉลี่ยของกลุ่มจากน้อยไปหามาก
 - ขีดเส้นแสดงว่า ค่าเฉลี่ยที่อยู่ในเส้นเดียวกัน ไม่แตกต่างกันอย่างมีนัยสำคัญ



เหลือ 1-3 และ 2-3 ที่ไม่ได้อยู่เส้นเดียวกัน แตกต่างกันอย่างมีนัยสำคัญ

การปรับค่า Alpha รายคู่

- วิธีการของ Sidak, Bonferroni, Holm สามารถประยุกต์ใช้ในการทำการทดสอบหลายๆ แบบพร้อมกัน เช่น เปรียบเทียบตัวแปรตามหลายตัว

บุคลิกภาพ	ชาย	หญิง	p	Rank	α	Sig
Neuroticism	45	50	.008	2	$.05 / 4 = .0125$	YES
Extroversion	55	49	.005	1	$.05 / 5 = .01$	YES
Agreeableness	49	50	.28	4		NO
Conscientiousness	48	51	.025	3	$.05 / 3 = .0167$	NO
Openness to experience	49	50	.31	5		NO

การทดสอบความแปรปรวนแบบทาง เดียว

สถิติชั้นกลางสำหรับจิตวิทยา

สันหัต พรประเสริฐมานิต

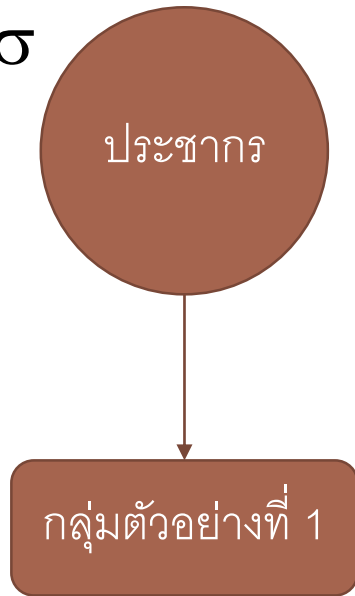
การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

- เมื่อเปรียบเทียบมากกว่า 2 กลุ่ม วิธีการอีกวิธีหนึ่งในการทดสอบความแตกต่างระหว่างกลุ่ม คือ การตรวจสอบความแปรปรวนของค่าเฉลี่ยระหว่างกลุ่ม
- กล่าวคือ หากความแปรปรวนของค่าเฉลี่ย สูงมากๆ ก็เป็นไปได้สูงที่กลุ่มทั้งหมด มาจากประชากรที่ค่าเฉลี่ยแตกต่างกัน
- วิธีการนี้เรียกว่า การวิเคราะห์ความแปรปรวน (Analysis of variance: ANOVA)

การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

- ทดสอบพฤติกรรมของค่าเฉลี่ยหากถูกสุ่มมาจากประชากรเดียวกัน

μ, σ



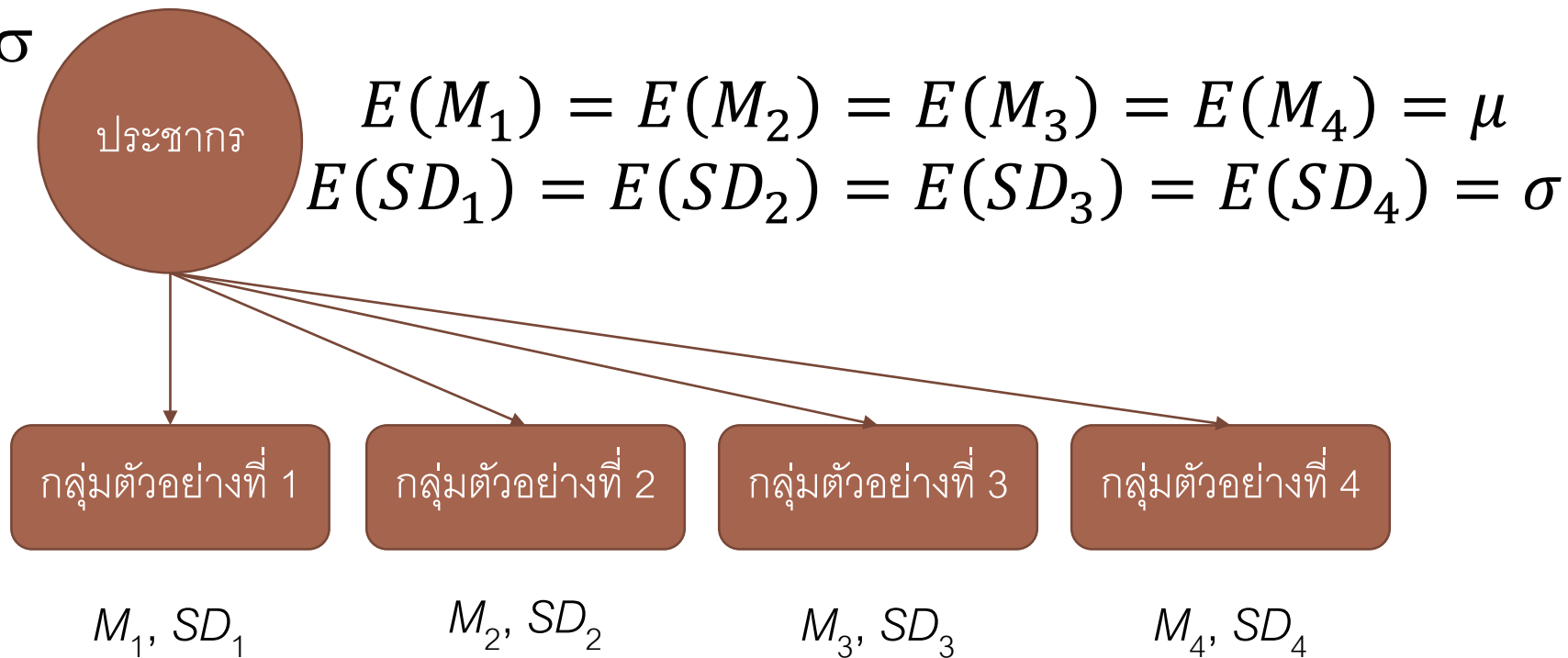
M_1, SD_1

$$E(M_1) = \mu$$
$$E(SD_1) = \sigma$$

การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

- ทดสอบพฤติกรรมของค่าเฉลี่ยหากกลุ่มมาจากประชากรเดียวกัน

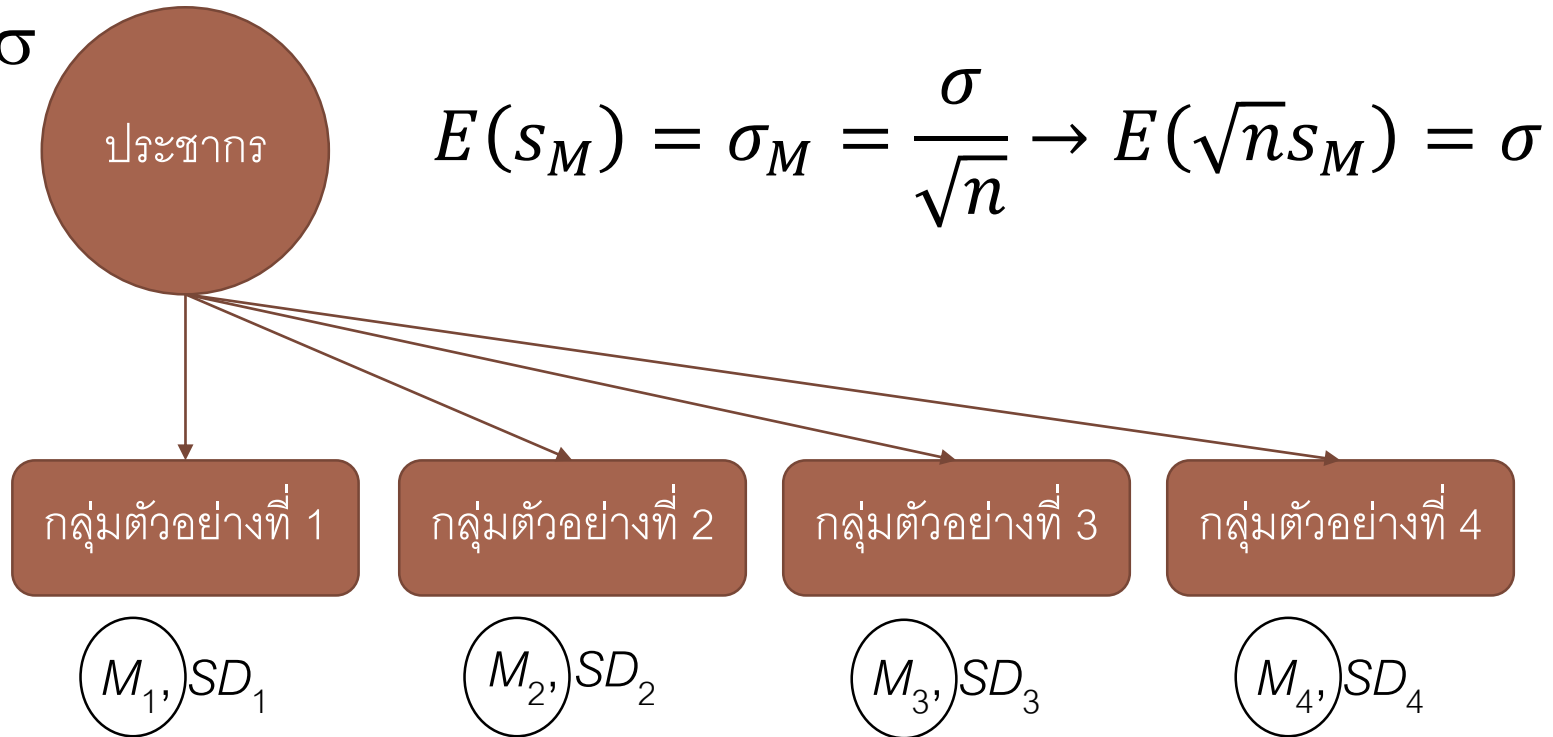
μ, σ



การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

- ทดสอบพฤติกรรมของค่าเฉลี่ยหากกลุ่มมาจากประชากรเดียวกัน

μ, σ



การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

- สังเกตว่า หากกลุ่มตัวอย่างสุ่มมาจากประชากรเดียวกันแล้ว ความแปรปรวนของประชากรสามารถทำนายด้วยสถิติสองรูปแบบด้วยกัน

$$E(SD_1^2) = E(SD_2^2) = E(SD_3^2) = E(SD_4^2) = \sigma^2$$
$$E(ns_M^2) = \sigma^2$$

- ความแปรปรวนภายในกลุ่มสามารถรวมกันได้เป็นความแปรปรวนร่วม (pooled variance)

$$SD_{\text{Pooled}} = \sqrt{\frac{\sum_{j=1}^k SS_j}{\sum_{j=1}^k df_j}}; SS_j = df_j (SD_j^2)$$

k คือจำนวนกลุ่ม

การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

- กล่าวคือ

$$\begin{aligned} E(SD_{\text{Pooled}}^2) &= \sigma^2 \\ E(ns_M^2) &= \sigma^2 \end{aligned}$$

- ดังนั้น หากกลุ่มตัวอย่างสุ่มมาจากประชากรเดียวกันจริงแล้ว

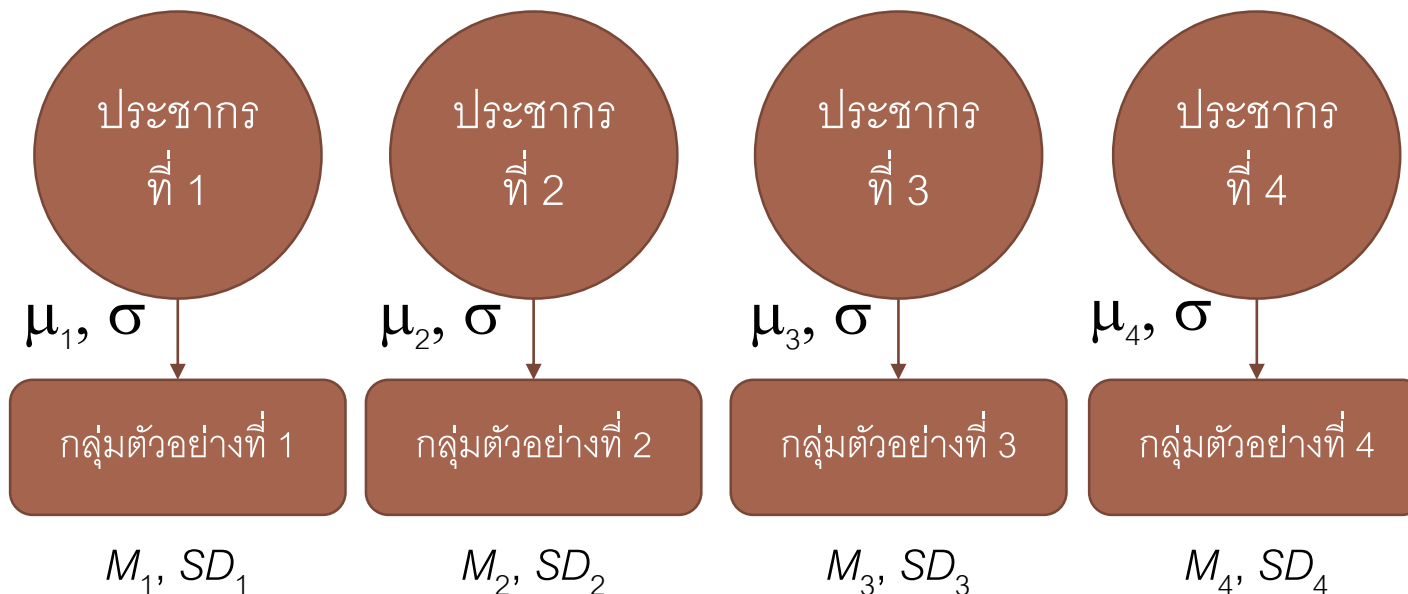
$$E\left(\frac{ns_M^2}{SD_{\text{Pooled}}^2}\right) = 1$$

- แต่ทว่า สัดส่วนดังกล่าวจากกลุ่มตัวอย่าง ไม่จำเป็นต้องเท่ากับ 1 เสมอไป อาจจะมากกว่าหรือน้อยกว่า 1 ก็ได้

การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

- หากกลุ่มตัวอย่างสุ่มมาจากประชากรคนละกลุ่ม

$$E(SD_1^2) = E(SD_2^2) = E(SD_3^2) = E(SD_4^2) = E(SD_{\text{Pooled}}^2) = \sigma^2$$
$$E(ns_M^2) = \sigma^2 + \text{ความแตกต่างระหว่าง } \mu_1, \mu_2, \mu_3, \mu_4$$



การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

- ดังนั้น หากกลุ่มตัวอย่างสุ่มมาจากประชากรที่แตกต่างกันและมีค่าเฉลี่ยแตกต่างกัน

$$E \left(\frac{ns_M^2}{SD_{\text{Pooled}}^2} \right) > 1$$

- แต่ทว่า สัดส่วนดังกล่าวจากกลุ่มตัวอย่าง อาจจะน้อยกว่า 1 ก็ได้ เนื่องจากความผิดพลาดจากการสุ่ม แต่โอกาสที่เกิดขึ้นจะยิ่งน้อย หากค่าเฉลี่ยในประชากรยังแตกต่างกัน

การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_k$$

H_1 : มีความแตกต่างระหว่างค่าเฉลี่ย

$$E \left(\frac{ns_M^2}{SD_{\text{Pooled}}^2} \right) = 1$$

$$E \left(\frac{ns_M^2}{SD_{\text{Pooled}}^2} \right) > 1$$

การกระจายของค่าเฉลี่ยระหว่างกลุ่ม

- ตัวอย่างของ H_1

$$\mu_1 = 20$$

$$\mu_2 = 20$$

$$\mu_3 = 20$$

$$\mu_4 = 25$$

$$\mu_1 = 25$$

$$\mu_2 = 20$$

$$\mu_3 = 20$$

$$\mu_4 = 25$$

$$\mu_1 = 18$$

$$\mu_2 = 19$$

$$\mu_3 = 20$$

$$\mu_4 = 91$$

$$\mu_1 = 18.1$$

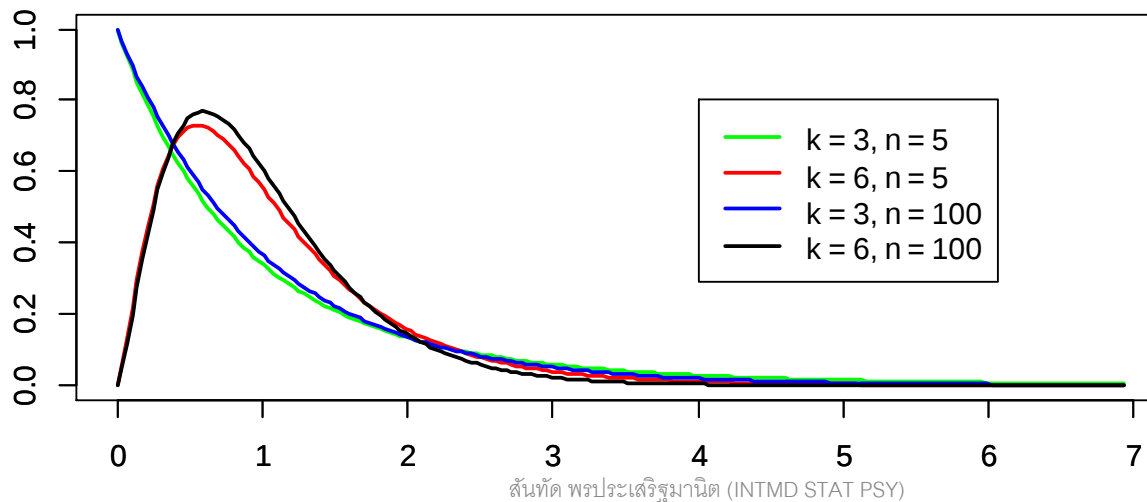
$$\mu_2 = 18.1$$

$$\mu_3 = 18.2$$

$$\mu_4 = 18.2$$

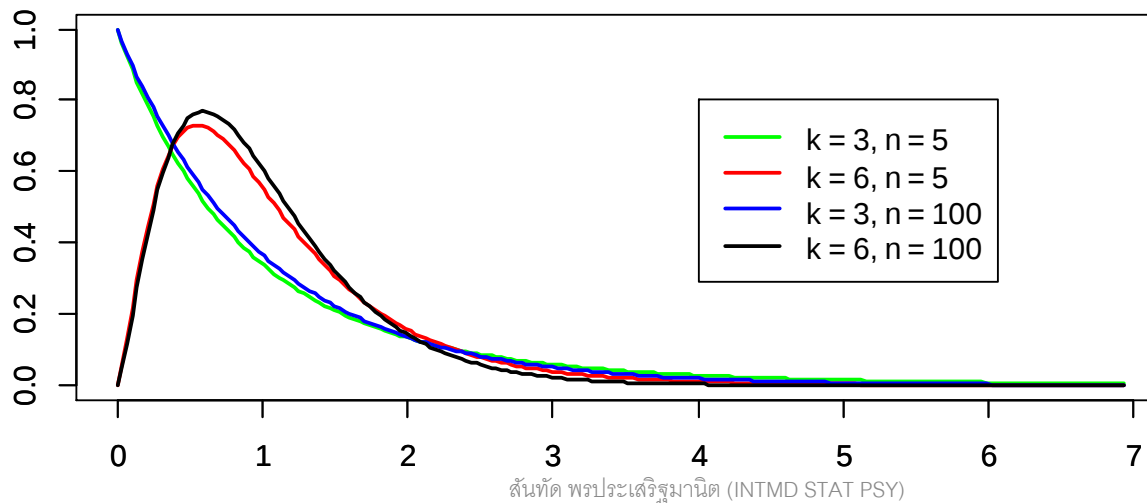
การกระจายแบบ F

- ในการทดสอบสมมติฐาน เราต้องรู้การกระจายของค่าสถิติหากสมมติฐานเป็นจริง
- นักสถิติพบว่าหาก Null Hypothesis เป็นจริง ค่าสถิติ $F = \frac{ns_M^2}{SD_{\text{Pooled}}^2}$ จะมี การกระจายแบบ F (F -distribution)



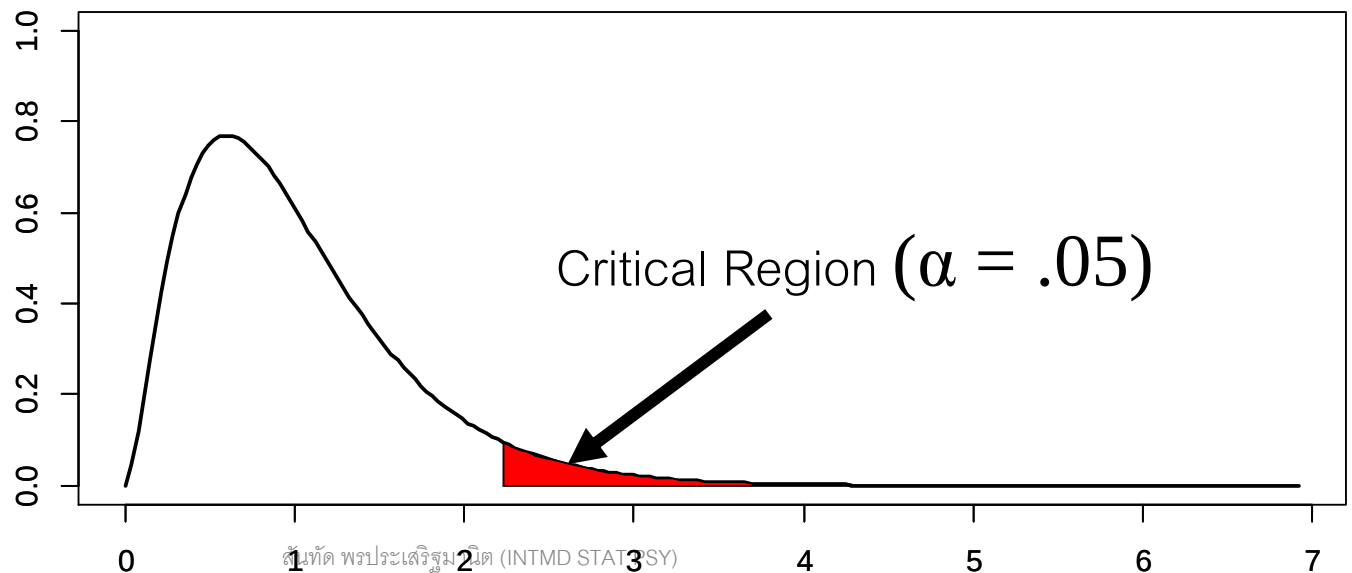
การกระจายแบบ F

- สังเกตว่าการกระจาย F จะมีรูปร่างขึ้นอยู่กับจำนวนกลุ่มที่เปรียบเทียบ (k) และจำนวนตัวอย่างในแต่ละกลุ่ม (n)
- หาก Null Hypothesis ไม่เป็นจริง ค่า F จะสูงกว่า 1 มาก



การกระจายแบบ F

- ในการทดสอบสมมติฐาน จะดูว่าค่า F มีค่าสูงกว่าที่คาดหวังเมื่อ Null Hypothesis เป็นจริงหรือไม่
- กล่าวคือ Critical Region อยู่ทางขวา
- ถ้าตกใน Critical Region $\rightarrow$ Reject H_0



การคำนวณค่า F

- ในการวิเคราะห์ความแปรปรวน จะมีรูปแบบการคำนวณค่า F ที่แสดงไว้ในโปรแกรมสถิติ

$$F = \frac{ns_M^2}{SD_{\text{Pooled}}^2}$$

โอกาสที่จะเจอค่า F ที่สูงเท่านี้
หรือมากกว่านี้ หาก Null
Hypothesis เป็นจริง

	SS	df	MS	F	p
Group				F	p
Error					
Total					

$p < \alpha \rightarrow$ Reject H_0

$p > \alpha \rightarrow$ Fail to
Reject H_0

การคำนวณค่า F

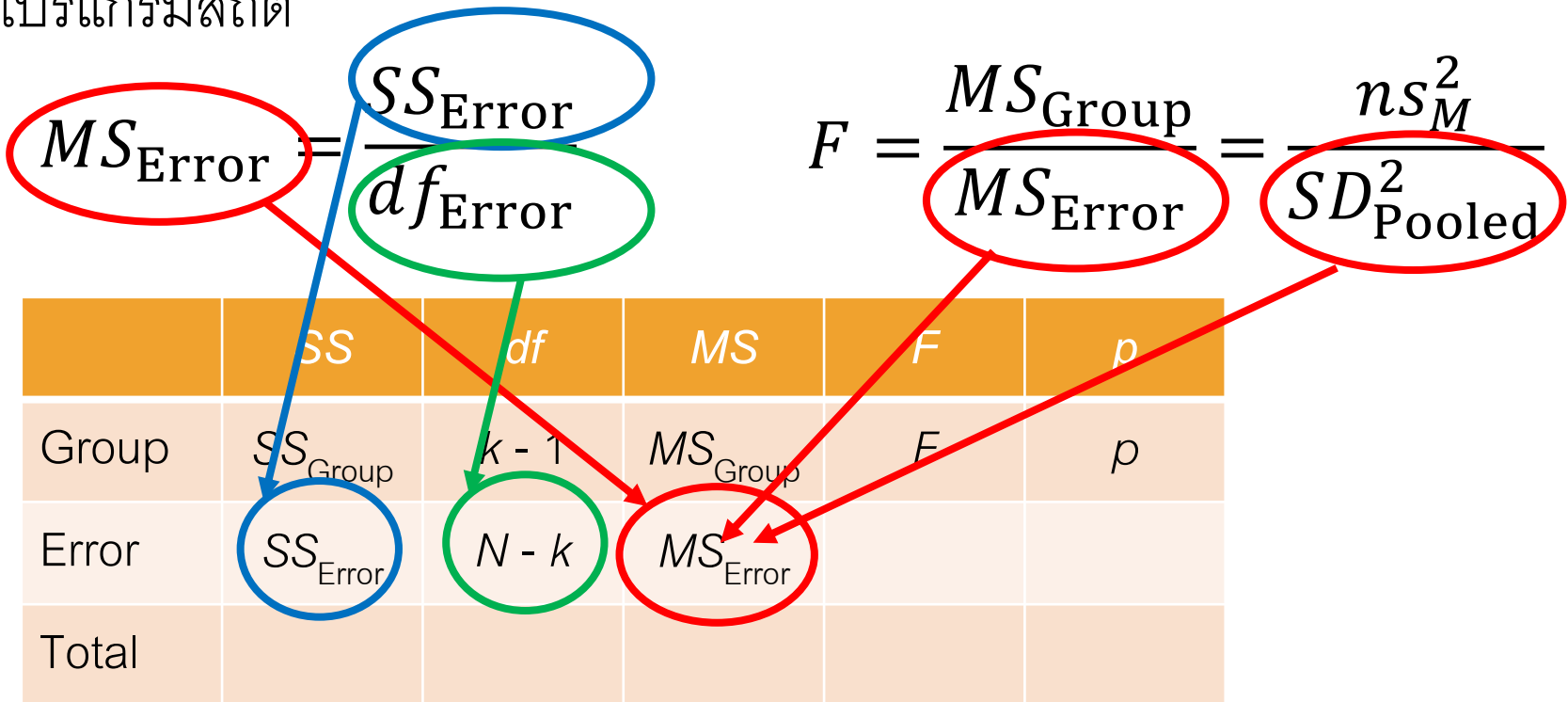
- ในการวิเคราะห์ความแปรปรวน จะมีรูปแบบการคำนวณค่า F ที่แสดงไว้ในโปรแกรมสถิติ

$$MS_{\text{Group}} = \frac{SS_{\text{Group}}}{df_{\text{Group}}}$$
$$F = \frac{MS_{\text{Group}}}{MS_{\text{Error}}} = \frac{ns_M^2}{SD_{\text{Pooled}}^2}$$

	SS	df	MS	F	p
Group	SS_{Group}	$k - 1$	MS_{Group}	F	p
Error					
Total					

การคำนวณค่า F

- ในการวิเคราะห์ความแปรปรวน จะมีรูปแบบการคำนวณค่า F ที่แสดงไว้ในโปรแกรมสถิติ



การคำนวณค่า F

- ในการวิเคราะห์ความแปรปรวน จะมีรูปแบบการคำนวณค่า F ที่แสดงไว้ในโปรแกรมสถิติ

$$SS_{\text{Total}} = SS_{\text{Group}} + SS_{\text{Error}}$$

	SS	df	MS	F	p
Group	SS_{Group}	$k - 1$	MS_{Group}	F	p
Error	SS_{Error}	$N - k$	MS_{Error}		
Total	SS_{Total}				

การคำนวณค่า F

- เช่น กลุ่มที่ 1: $X = 5, 6, 7, 8, 9 \rightarrow M_1 = 7, SD_1 = 1.58$
- กลุ่มที่ 2: $X = 4, 5, 6, 7, 8 \rightarrow M_2 = 6, SD_2 = 1.58$
- กลุ่มที่ 3: $X = 3, 4, 5, 6, 7 \rightarrow M_3 = 5, SD_3 = 1.58$

$$s_M^2 = 1$$

$$k - 1 = 3 - 1 = 2$$

$$ns_M^2 = 5$$

$$SD_{\text{Pooled}} = 1.58$$

$$SD_{\text{Pooled}}^2 = 2.5$$

	SS	df	MS	F	p
Group	10	2	5	2.00	.178
Error	30	12	2.5		
Total	40				

$$F = \frac{5}{2.5} = 2$$

$$N - k = 15 - 3 = 12$$

$$p(F_{2,12} \geq 2) = .178$$

การคำนวณค่า F

- $SS_{\text{Group}} = 10$; $df_{\text{Group}} = 2$; $MS_{\text{Group}} = 5$
- $SS_{\text{Error}} = 30$; $df_{\text{Error}} = 12$; $MS_{\text{Error}} = 2.5$
- $SS_{\text{Total}} = 40$; $F(2, 12) = 12$, $p = .178$

	SS	df	MS	F	p
Group	10	2	5	2.00	.178
Error	30	12	2.5		
Total	40				

การทดสอบสมมติฐาน

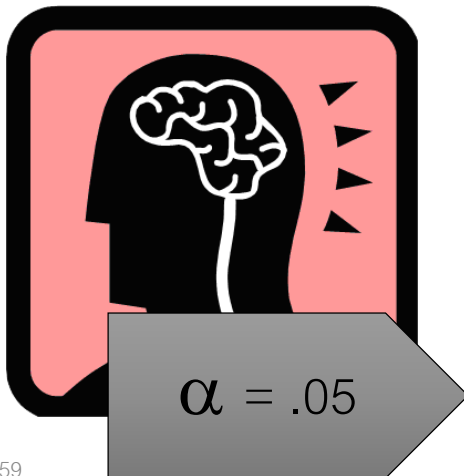
- เช่น

ทดสอบสิ่งแวดล้อมว่ามีผลต่อ
ความจำหรือไม่โดยเปิดเพลง



(1) Classic (2) Pop-dance

(3) Heavy-metal



นักวิจัยจึงนำนิสิต 15 คน
สุ่มเข้ากลุ่ม กลุ่มละ 5 คน



สมมติฐาน คือ เพลงมีอิทธิพลต่อ
ความจำ



$$H_0: \mu_1 = \mu_2 = \mu_3$$



H_1 : ค่าเฉลี่ยแตกต่างกัน
รูปแบบใดรูปแบบหนึ่ง

การทดสอบสมมติฐาน

- เช่น

ทดสอบสิ่งแวดล้อมว่ามีความจำเป็นหรือไม่โดยเปิด

(1) Classic (2) Pop-d

(3) Heavy-metal



1 = Classic

2 = Pop-dance

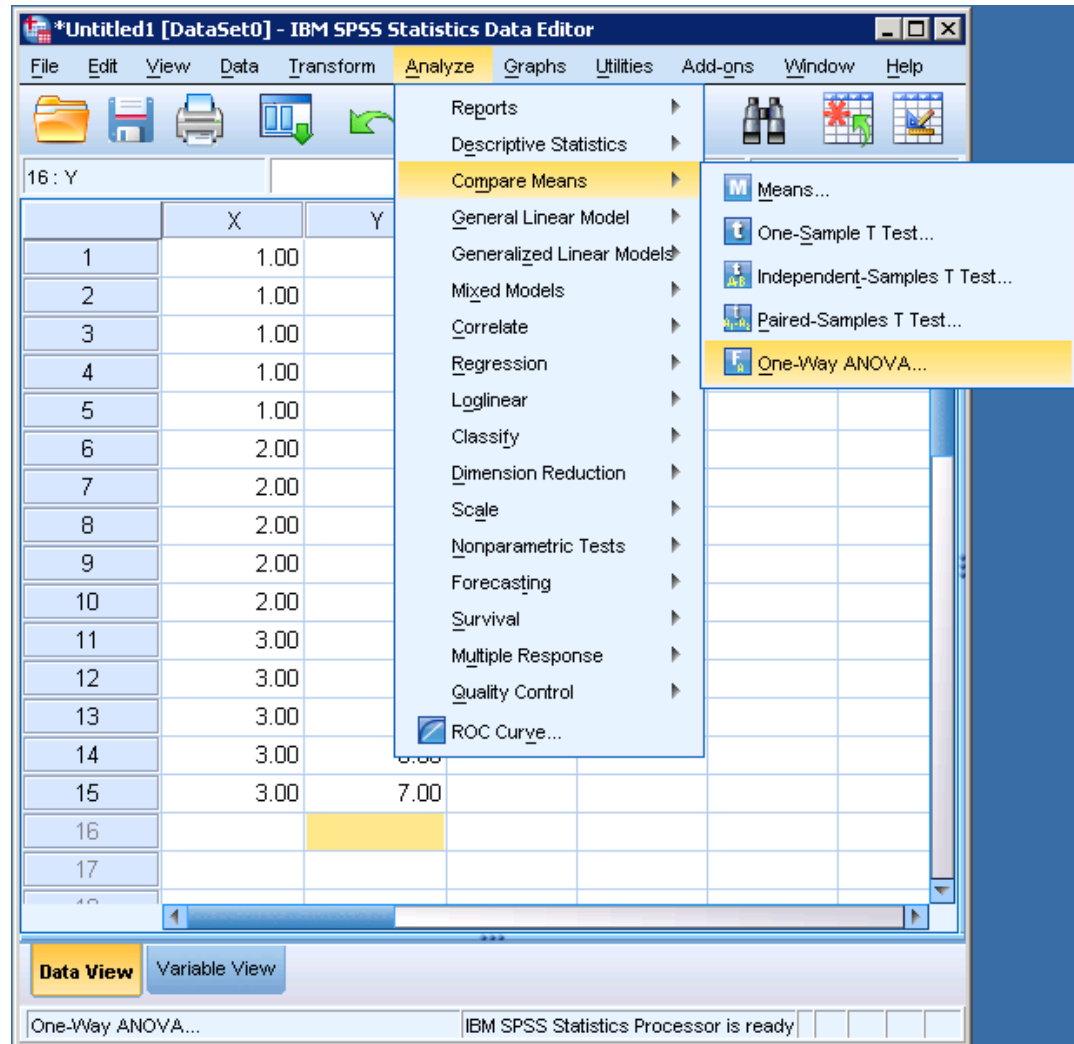
3 = Heavy-metal

	X	Y	var	var	var	var
1	1.00	5.00				
	1.00	6.00				
	1.00	7.00				
	1.00	8.00				
5	1.00	9.00				
	2.00	4.00				
	2.00	5.00				
	2.00	6.00				
	2.00	7.00				
10	2.00	8.00				
	3.00	3.00				
	3.00	4.00				
	3.00	5.00				
	3.00	6.00				
15	3.00	7.00				
16						
17						
18						

การทดสอบสมมติฐาน

- เช่น

ทดสอบสิ่งแวดล้อมว่ามีผลต่อ
ความจำหรือไม่โดยเปิดเพลง
(1) Classic (2) Pop-dance
(3) Heavy-metal



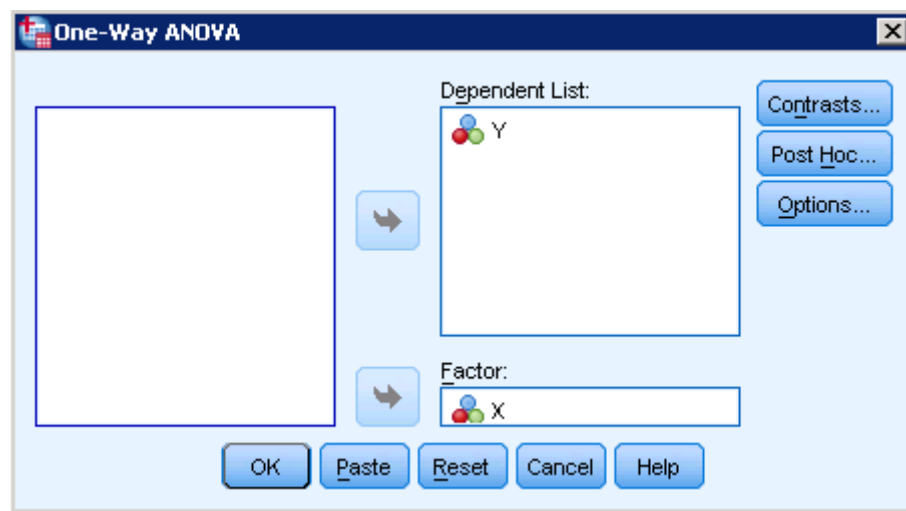
การทดสอบสมมติฐาน

- เช่น

ทดสอบสิ่งแวดล้อมว่ามีผลต่อ
ความจำหรือไม่โดยเปิดเพลง

(1) Classic (2) Pop-dance

(3) Heavy-metal



ใน Options เลือก Descriptive เพื่อหา
Descriptive Statistics

การทดสอบสมมติฐาน

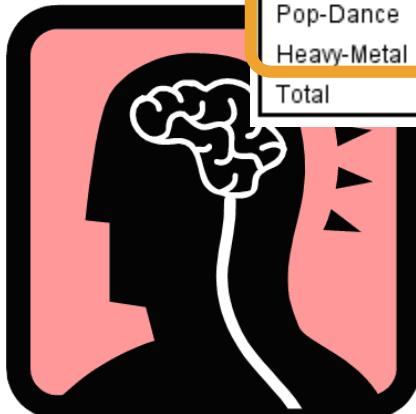
- เช่น

ทดสอบสิ่งแวดล้อมว่ามีผลต่อ
ความจำหรือไม่

Descriptive Statistic ของแต่ละกลุ่ม

(1) Classic (2)

(3) Heavy



Descriptives

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Classical	5	7.0000	1.58114	.70711	5.0368	8.9632	5.00	9.00
Pop-Dance	5	6.0000	1.58114	.70711	4.0368	7.9632	4.00	8.00
Heavy-Metal	5	5.0000	1.58114	.70711	3.0368	6.9632	3.00	7.00
Total	15	6.0000	1.69031	.43644	5.0639	6.9361	3.00	9.00

การทดสอบสมมติฐาน

- เช่น

ทดสอบสิ่งแวดล้อมว่ามีผลต่อ
ความจำหรือไม่โดยเปิดเพลง

(1) Classic (2) Pop-dance

(3) Heavy-metal



$$F(2, 12) = 2.00, p = .18$$

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	10.000	2	5.000	2.000	.178
Within Groups	30.000	12	2.500		
Total	40.000	14			

$$p = .18$$

การทดสอบสมมติฐาน

- เช่น

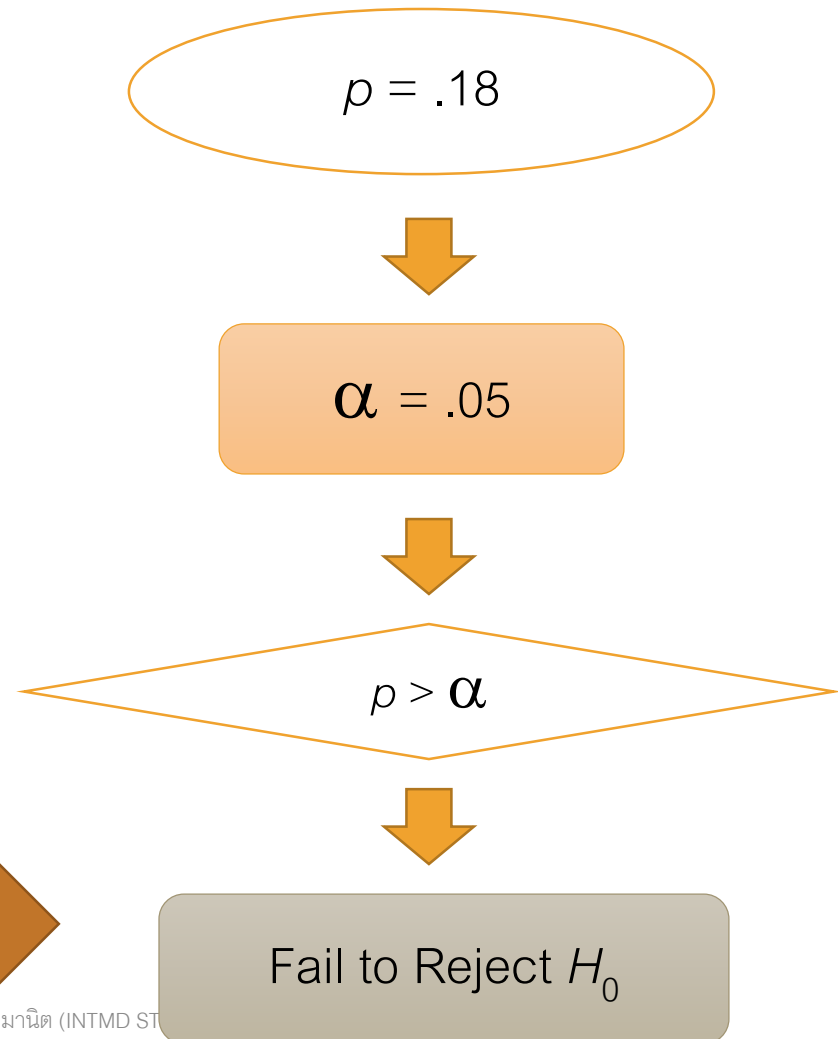
ทดสอบสิ่งแวดล้อมว่ามีผลต่อ
ความจำหรือไม่โดยเปิดเพลง

(1) Classic (2) Pop-dance

(3) Heavy-metal



ไม่สามารถสรุปได้ว่า การเปิดเพลงประเภท
ที่แตกต่างกัน มีผลต่อความจำหรือไม่



การทดสอบสมมติฐาน

- One-way ANOVA สามารถใช้ทดสอบระหว่าง 2 กลุ่มได้
- ผลที่ได้จะเหมือนกับ Independent t -test เมื่อทดสอบสองทาง
 - ค่า F จะเท่ากับ t^2
 - ค่า p ที่ได้จะเท่ากัน

การทดสอบสมมติฐาน

ความคิดเห็นต่อการมีห้องเชียร์
เปรียบเทียบกันระหว่าง 4 กลุ่ม
(ก) นักเรียนม.6 (ข) นิสิตปี 1
(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป
ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง
5 เห็นด้วยอย่างยิ่ง



เก็บข้อมูลได้จาก

- (ก) นักเรียน ม.6 จำนวน 50 คน
- (ข) นิสิตปี 1 จำนวน 200 คน
- (ค) นิสิตปี 2 จำนวน 200 คน
- (ง) ประชาชนทั่วไป จำนวน 50 คน

สมมติฐาน คือ กลุ่มทั้ง 4 กลุ่มมีความ
คิดเห็นต่อห้องเชียร์แตกต่างกัน

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

H_1 : ค่าเฉลี่ยแตกต่างกัน
รูปแบบใดรูปแบบหนึ่ง

การทดสอบสมมติฐาน

ความคิดเห็นต่อการมีห้องเชียร์
เปรียบเทียบกันระหว่าง 4 กลุ่ม

(ก) นักเรียนม.6

Descriptives

(ค) นิสิตปี 2 (ง) 1
ประเมิน 1 ไม่ให้
5 เห็นด้วย

Attitude								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
High School	50	2.58	.609	.086	2.41	2.75	1	4
Freshman	200	2.01	.340	.024	1.96	2.05	1	3
Sophomore	200	2.56	.555	.039	2.48	2.64	1	4
Adults	50	1.40	.495	.070	1.26	1.54	1	2
Total	500	2.22	.612	.027	2.17	2.28	1	4



การทดสอบสมมติฐาน

ความคิดเห็นต่อการมีห้องเชียร์
เปรียบเทียบกันระหว่าง 4 กลุ่ม
(ก) นักเรียนม.6 (ข) นิสิตปี 1
(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป
ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง
5 เห็นด้วยอย่างยิ่ง



$$F(3, 496) = 104.67, p < .001$$

ANOVA

Attitude	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	72.457	3	24.152	104.666	.000
Within Groups	114.455	496	.231		
Total	186.912	499			

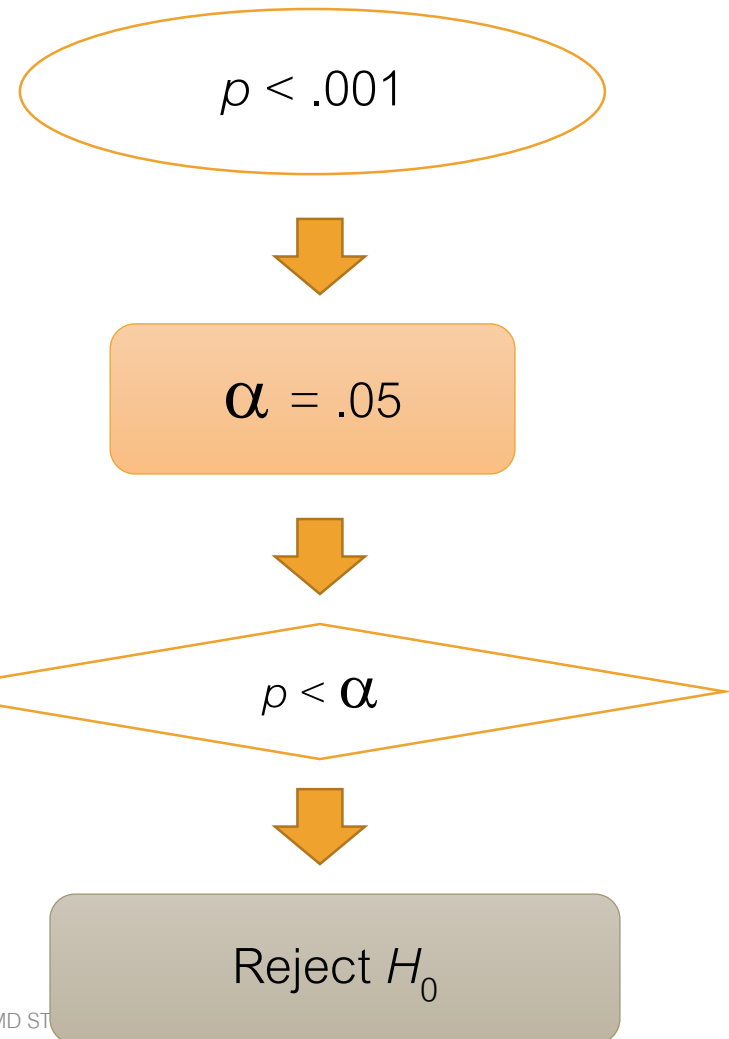
$$p < .001$$

การทดสอบสมมติฐาน

ความคิดเห็นต่อการมีห้องเชียร์
เปรียบเทียบกันระหว่าง 4 กลุ่ม
(ก) นักเรียนม.6 (ข) นิสิตปี 1
(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป
ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง
5 เห็นด้วยอย่างยิ่ง



กลุ่มทั้ง 4 กลุ่มมีความคิดเห็น
ต่อห้องเชียร์แตกต่างกัน



ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- การกระจายของประชากรภายในแต่ละกลุ่มเป็นโค้งปกติ
 - การวิเคราะห์ความแปรปรวนคงทนต่อการกระจายไม่เป็นโค้งปกติ
- ความแปรปรวนภายในแต่ละกลุ่มเท่ากัน (Homogeneity of variance)
 - กลุ่มตัวอย่างเท่ากัน ไม่ค่อยมีปัญหา
 - กลุ่มตัวอย่างไม่เท่ากัน ความแปรปรวนสูงสุดต่อต่ำสุดต่ำกว่า 10 ต่อ 1 ยังโอเค
 - สามารถใช้ Welch Test ในการทดสอบเมื่อความแปรปรวนภายในแต่ละกลุ่มไม่เท่ากัน แต่ไม่คงทนต่อการกระจายไม่เป็นโค้งปกติ
- ไม่ควรใช้ Levene test ในการทดสอบข้อตกลงเบื้องต้นก่อนการใช้สถิติ (หากไม่เข้าใจ ให้ดูวิดีโอทบทวน เรื่องการเปรียบเทียบค่าเฉลี่ยที่อิสระจากกัน)

ขนาดอิทธิพล

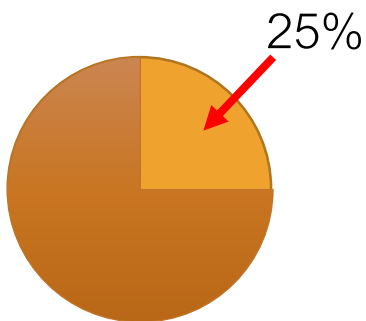
- ขนาดของความแตกต่างของคะแนนทั้งหมดอธิบายด้วยความแตกต่างระหว่างกลุ่ม
 - Eta-squared
 - Omega-squared
- ขนาดของความแตกต่างของค่าเฉลี่ยของแต่ละกลุ่มจากค่าเฉลี่ยรวมในหน่วยของส่วนเบี่ยงเบนมาตรฐาน
 - Cohen's f

ขนาดอิทธิพล

- Eta-squared (η^2)

- จาก $SS_{\text{Total}} = SS_{\text{Group}} + SS_{\text{Error}}$
- ดังนั้น $\eta^2 = SS_{\text{Group}} / SS_{\text{Total}}$
- แต่ $E(\eta^2)$ ไม่เท่ากับค่านี้ในประชากร ทำให้เป็นค่าสถิติที่ไม่ดี

$$\eta^2 = \frac{SS_{\text{Group}}}{SS_{\text{Total}}} = \frac{10}{40} = .25$$



	SS	df	MS	F	p
Group	10	2	5	2.00	.178
Error	30	12	2.5		
Total	40				

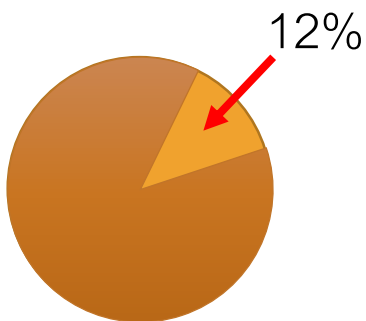
ขนาดอิทธิพล

- Omega-squared (ω^2)

- $\omega^2 = \frac{SS_{\text{Treat}} - (k-1)MS_{\text{Error}}}{SS_{\text{Total}} + MS_{\text{Error}}}$

- ω^2 ประมาณค่า η^2 ของประชากรดีกว่า η^2 ที่คำนวณจากกลุ่มตัวอย่าง

$$\omega^2 = \frac{SS_{\text{Treat}} - (k-1)MS_{\text{Error}}}{SS_{\text{Total}} + MS_{\text{Error}}} = \frac{10 - 2 \cdot 2.5}{40 + 2.5} = \frac{5}{42.5} = .12$$

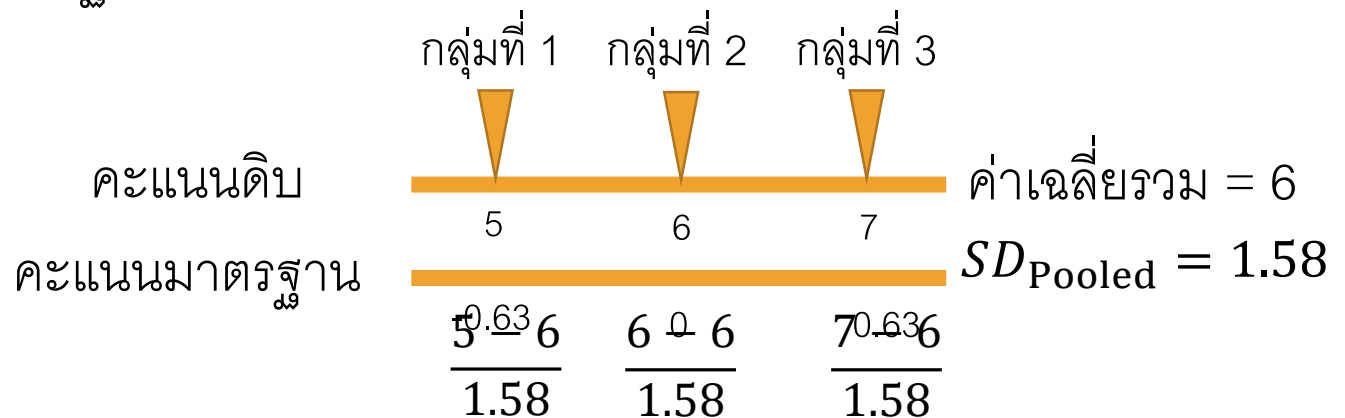


	SS	df	MS	F	p
Group	10	2	5	2.00	.178
Error	30	12	2.5		
Total	40				

ขนาดอิทธิพล

- Cohen's f

- ขนาดของความแตกต่างของค่าเฉลี่ยของแต่ละกลุ่มจากค่าเฉลี่ยรวมในหน่วยของส่วนเบี่ยงเบนมาตรฐาน



$$f = \sqrt{\frac{1}{k} \sum_{j=1}^k \left(\frac{\mu_j - \mu}{\sigma_e} \right)^2} = \sqrt{\frac{1}{3} (-0.63^2 + 0^2 + 0.63^2)} = 0.516$$

ขนาดอิทธิพล

- Cohen's f
 - ขนาดอิทธิพลน้อย:
 $f = .10$; ปานกลาง: $f = .25$; สูง: $f = .40$
 - ใช้ในการหาค่าพลังทางสถิติ

การเขียนรายงาน

- งานวิจัยนี้ต้องการทดสอบอิทธิพลของประเภทของดนตรีต่อความสามารถในการจำระยะสั้น แบ่งคนเป็นกลุ่มละ 10 คน เปรียบเทียบระหว่างเพลง Classic ($M = 7$, $SD = 1.58$) เพลง Pop-dance ($M = 6$, $SD = 1.58$) และเพลง Heavy-metal ($M = 5$, $SD = 1.58$) จากการทดสอบด้วยการวิเคราะห์ความแปรปรวนทางเดียว พบว่าความจำระยะสั้นไม่แตกต่างกันอย่างมีนัยสำคัญระหว่างประเภทของดนตรี, $F(2, 12) = 2.00$, $p = .18$, $\omega^2 = .12$

การเขียนรายงาน

- งานวิจัยนี้ เปรียบเทียบความคิดเห็นต่อห้องเชียร์ของกลุ่มคน 4 กลุ่ม คือ นักเรียนชั้นม. 6 นิสิตชั้นปีที่ 1 นิสิตชั้นปีที่ 2 และประชาชนทั่วไปที่ทำงานแล้ว จากการทดสอบด้วยการวิเคราะห์ความแปรปรวนทางเดียว กลุ่มคนทั้ง 4 กลุ่ม มีความคิดเห็นต่อห้องเชียร์แตกต่างกันอย่างมีนัยสำคัญ, $F(3, 496) = 104.67, p < .001, \omega^2 = .38$ ค่าสถิติพรรณนาแสดงไว้ในตารางที่ 1

ตารางที่ 1 สถิติพรรณนาของกลุ่มต่างๆ ในความคิดเห็นต่อห้องเชียร์

กลุ่มที่เปรียบเทียบ	<i>n</i>	<i>M</i>	<i>SD</i>
นักเรียนชั้นม. 6	50	2.58	0.61
นิสิตชั้นปีที่ 1	200	2.01	0.34
นิสิตชั้นปีที่ 2	200	2.56	0.56
ประชาชนทั่วไป	50	1.40	0.50

กำลังในการทดสอบทางสถิติ

- ใช้โปรแกรม G*POWER 3 โดย
 - เลือกกลุ่มสถิติที่ต้องการทดสอบ คือ F-tests
 - เลือกสถิติที่ใช้ทดสอบ คือ ANOVA: Fixed effects, omnibus, one-way
 - เลือกว่าจะหา กำลังในการทดสอบ คือ Post hoc: Compute achieved power – given α , sample size and effect size

กำลังในการทดสอบทางสถิติ

ขนาดอิทธิพล (หากรู้แต่ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของแต่ละกลุ่ม กด Determine ดูตัวอย่างหน้าถัดไป)

ระดับนัยสำคัญ

ขนาดกลุ่มตัวอย่าง

จำนวนกลุ่ม

Input Parameters		Output Parameters		
Test family	F tests	Statistical test	ANOVA: Fixed effects, omnibus, one-way	
Type of power analysis	Post hoc: Compute achieved power – given α , sample size, and effect size			
Determine =>	Effect size f	0.5167700	Noncentrality parameter λ	4.0057685
	α err prob	0.05	Critical F	3.8852938
	Total sample size	15	Numerator df	2
	Number of groups	3	Denominator df	12
			Power (1- β err prob)	0.3331129

ค่าพารามิเตอร์ที่เบี่ยงเบน

กำลังในการทดสอบสมมติฐานทางสถิติ

กำลังในการทดสอบทางสถิติ

ค่าเฉลี่ยของแต่ละกลุ่ม

จำนวนกลุ่ม

ส่วนเบี่ยงเบนมาตรฐานภายในแต่ละกลุ่ม หรือ SD_{Pooled}

จำนวนกลุ่มตัวอย่าง

ขนาดอิทธิพล (Cohen's f)

Group	Mean	Size
1	7	5
2	6	5
3	5	5

Equal n: 5

Total sample size: 15

Calculate Effect size f: 0.51677

Calculate and transfer to main window

Close

การกำหนดขนาดกลุ่มตัวอย่าง

- การกำหนดขนาดกลุ่มตัวอย่างใน One-way ANOVA มุ่งเน้นที่การทดสอบกำลังเท่านั้น
- การกำหนดขนาดกลุ่มตัวอย่างตามกำลังที่ต้องการ สามารถทำได้โดยใช้โปรแกรม G*POWER 3 ช่วย
 - เลือกกลุ่มสถิติที่ต้องการทดสอบ คือ F-tests
 - เลือกสถิติที่ใช้ทดสอบ คือ ANOVA: Fixed effects, omnibus, one-way
 - เลือกว่าจะหาขนาดกลุ่มตัวอย่าง คือ A Priori: Compute required sample size – given α , power and effect size

การกำหนดขนาดกลุ่มตัวอย่าง

- เช่น จงหาขนาดกลุ่มตัวอย่างที่จะทำให้กำลัง = .80 ระดับนัยสำคัญเท่ากับ .05 จากการเปรียบเทียบผลของดนตรีประเภทต่างๆ

Test family		Statistical test	
F tests		ANOVA: Fixed effects, omnibus, one-way	
Type of power analysis			
A priori: Compute required sample size - given α , power, and effect size			
Input Parameters		Output Parameters	
Determine =>	Effect size f	Noncentrality parameter λ	11.2161518
	α err prob	Critical F	3.2380961
	Power ($1 - \beta$ err prob)	Numerator df	2
	Number of groups	Denominator df	39
		Total sample size	42
		Actual power	0.8302080

ใช้กลุ่มตัวอย่างรวม 42 คน กลุ่มละ 14 คน

การเปรียบเทียบรายคู่

สถิติขั้นกลางสำหรับจิตวิทยา

สันหัด พรประเสริฐมานิต

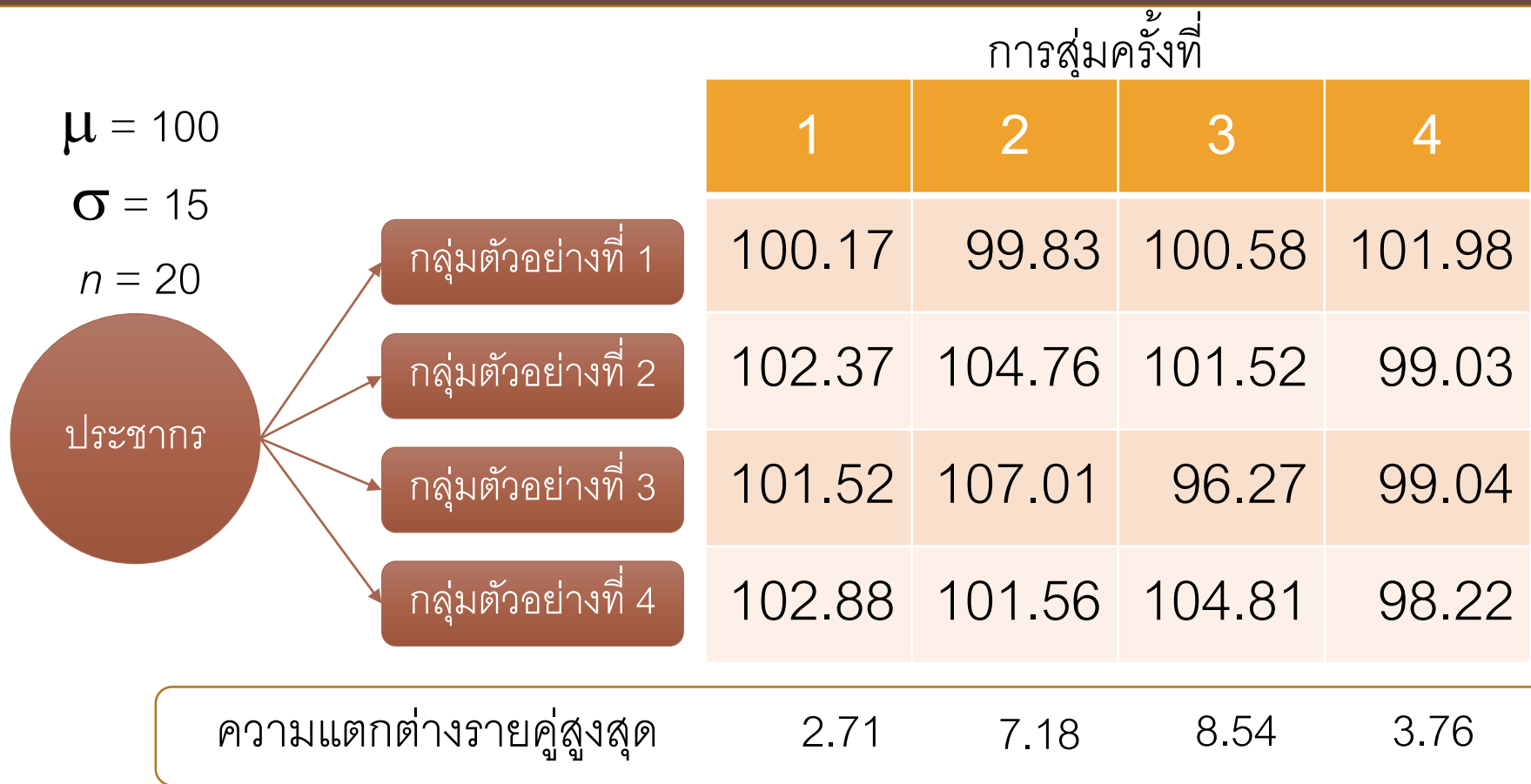
การเปรียบเทียบค่าเฉลี่ย

- One-way ANOVA เป็นการทดสอบแบบภาพรวม (Omnibus test)
- เมื่อปฏิเสธ Null Hypothesis เราจะบอกว่า Alternative Hypothesis เป็นจริง คือ มีความแตกต่างรูปแบบใดรูปแบบหนึ่ง
- ไม่ได้บอกถึงความแตกต่างแบบไหน
- วิธีการหนึ่งในการหารูปแบบความแตกต่าง คือ การเปรียบเทียบรายคู่ (Pairwise comparison) ดังที่กล่าวไปในข้างต้น

การเปรียบเทียบค่าเฉลี่ย

- วิธี Bonferroni, Sidak, หรือ Holm เป็นวิธีในการเปลี่ยนค่า α ซึ่งเป็นวิธีที่สามารถใช้กับการเปรียบเทียบรูปแบบใดก็ได้ เช่น Independent t -test กับตัวแปรตามหลายตัว
- เป็นวิธีเหล่านี้ เป็นวิธีที่ดี แต่กำลังในการทดสอบทางสถิติไม่สูงเท่ากับวิธีที่ถูกออกแบบมาสำหรับการวิเคราะห์รายคู่โดยเฉพาะ

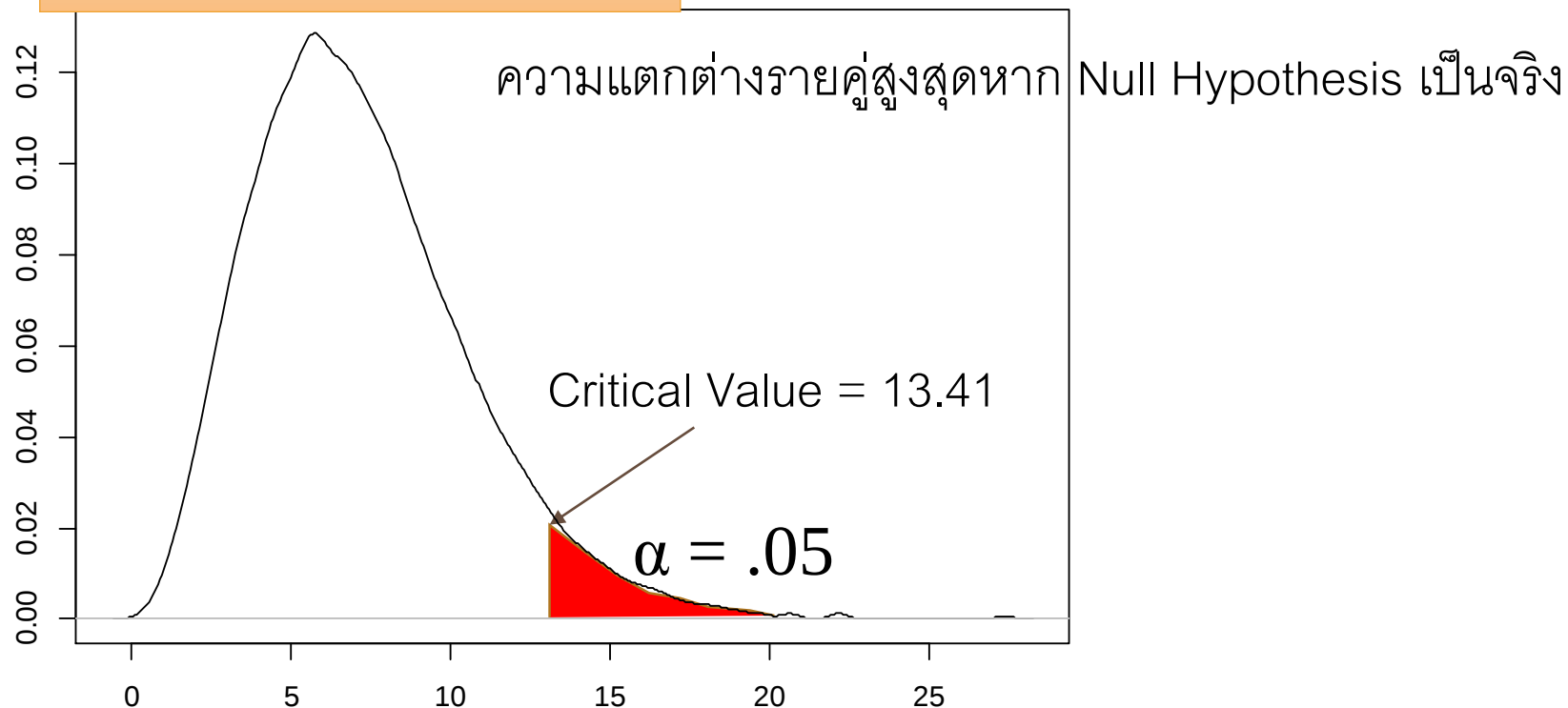
ความแตกต่างรายคู่ที่เป็นไปได้สูงสุด



นำมาสร้าง Sampling Distribution ดูว่าความแตกต่างรายคู่สูงสุดมีค่าเท่าไร ถ้า H_0 เป็นจริง

ความแตกต่างรายคู่ที่เป็นไปได้สูงสุด

Studentized Distribution



ดังนั้น หากความแตกต่างรายคู่สูงกว่า 13.41 $\rightarrow$ Reject H_0

วิธีนี้เรียกว่าวิธีการ
ของ Tukey

การเปรียบเทียบรายคู่: การทดสอบสมมติฐาน

ความคิดเห็นต่อการมีห้องเชียร์
เปรียบเทียบกันระหว่าง 4 กลุ่ม
(ก) นักเรียนม.6 (ข) นิสิตปี 1
(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป
ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง
5 เห็นด้วยอย่างยิ่ง



One-way ANOVA

Reject H_0



กลุ่มทั้ง 4 กลุ่มมีความคิดเห็น
ต่อห้องเชียร์แตกต่างกัน



ทำการเปรียบเทียบรายคู่ เพื่อค้นหา
รูปแบบของความแตกต่าง

การเปรียบเทียบรายคู่: การทดสอบสมมติฐาน

ความคิดเห็นต่อการมีห้องเชียร์

เปรียบเทียบกันระหว่าง 4 กลุ่ม

(ก) นักเรียนม.6 (ข) นิสิตปี 1

(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป

ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง

5 เห็นด้วยอย่างยิ่ง

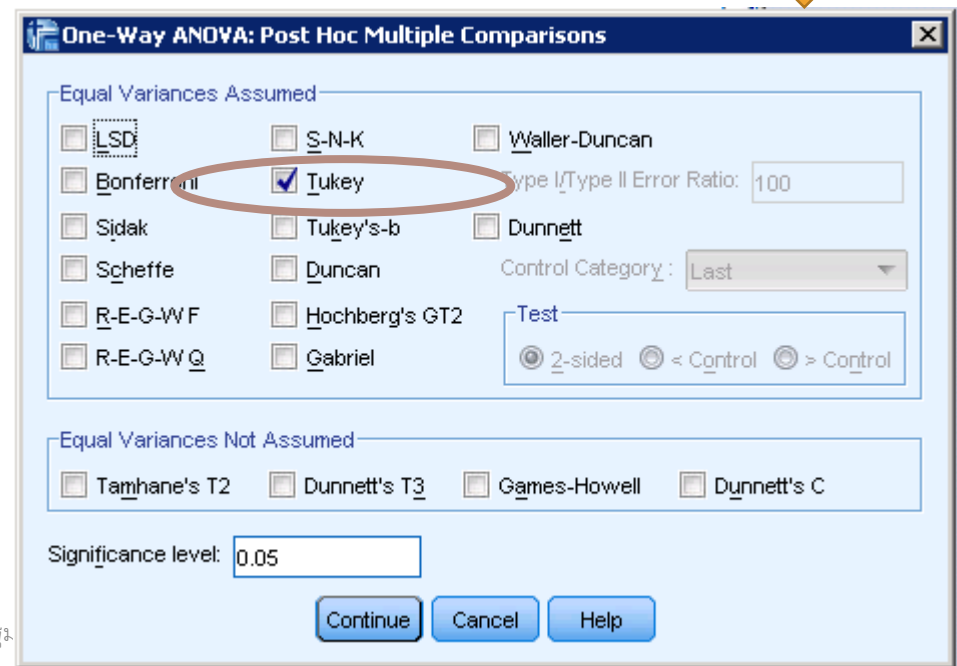
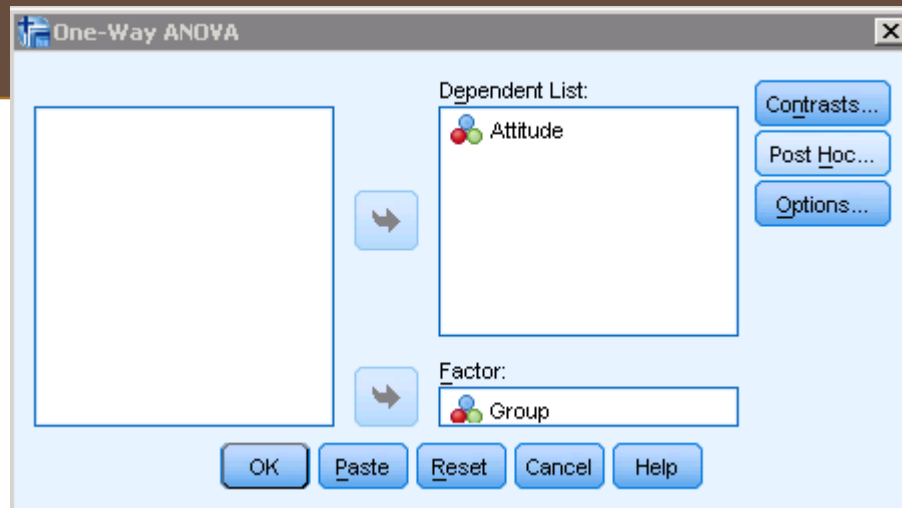
การเปรียบเทียบรายคู่ที่เป็นไปได้ทั้งหมดคือ 6 คู่



	นักเรียนม.6	นิสิตปีที่ 1	นิสิตปีที่ 2	ประชาชนทั่วไป
นักเรียนม.6				
นิสิตปีที่ 1				
นิสิตปีที่ 2				
ประชาชนทั่วไป				

การเปรียบเทียบรายคู่: การทดสอบสมมติฐาน

ความคิดเห็นต่อการมีห้องเชียร์
เปรียบเทียบกันระหว่าง 4 กลุ่ม
(ก) นักเรียนม.6 (ข) นิสิตปี 1
(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป
ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง
5 เห็นด้วยอย่างยิ่ง



การเปรียบเทียบรายคู่: การทดสอบสมมติฐาน

ความคิดเห็นต่อการมีห้องเชียร์
เปรียบเทียบกันระหว่าง 4 กลุ่ม
(ก) นักเรียนม.6 (ข) นิสิตปี 1
(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป
ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง
5 เห็นด้วยอย่างยิ่ง



Multiple Comparisons

ค่า p จาก Tukey Test

Dependent Variable: Attitude
Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
High School	Freshman	.575*	.076	.000	.38	.77
	Sophomore	.020	.076	.994	-.18	.22
	Adults	1.180*	.096	.000	.93	1.43
Freshman	Sophomore	-.555*	.048	.000	-.68	-.43
	Adults	.605*	.076	.000	.41	.80
Sophomore	Adults	1.160*	.076	.000	.96	1.36

*. The mean difference is significant at the 0.05 level.

การเปรียบเทียบรายคู่: การทดสอบสมมติฐาน

ความคิดเห็นต่อการมีห้องเชียร์
เปรียบเทียบกันระหว่าง 4 กลุ่ม
(ก) นักเรียนม.6 (ข) นิสิตปี 1
(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป
ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง
5 เห็นด้วยอย่างยิ่ง



Attitude

Tukey HSD^{a, b}

Group	N	Subset for alpha = 0.05		
		1	2	3
Adults	50	1.40		
Freshman	200		2.01	
Sophomore	200			2.56
High School	50			2.58
Sig.		1.000	1.000	.994

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 80.000.

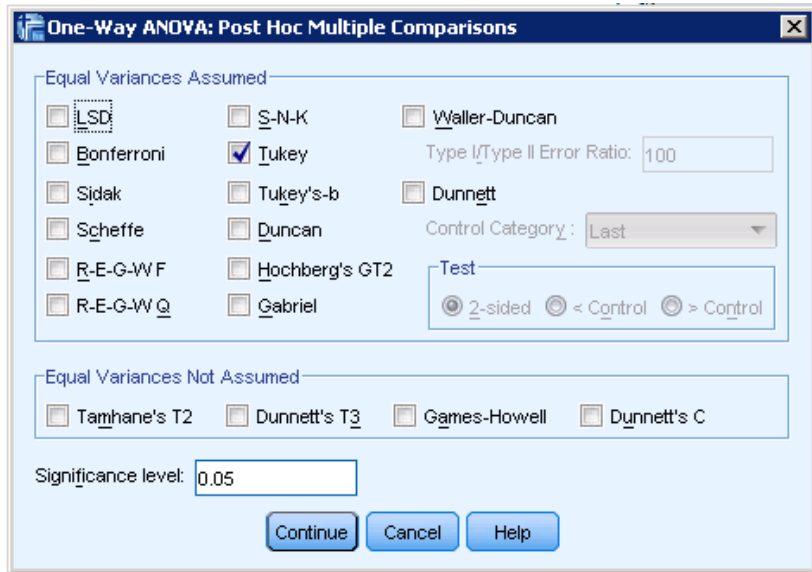
b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Sophomore vs. High School → ไม่แตกต่างกันอย่างมีนัยสำคัญ
คู่ที่เหลือ แตกต่างกันอย่างมีนัยสำคัญ

การเปรียบเทียบรายคู่: ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- เหมือนการวิเคราะห์ความแปรปรวนทุกอย่าง
- ใช้วิธีของ Games-Howell หากความแปรปรวนสูงสุดและต่ำสุด มากกว่า 10 ต่อ 1 และกลุ่มตัวอย่างไม่เท่ากัน
- แต่วิธีนี้ไม่คงทนต่อการละเมิดการกระจายภายในกลุ่มเป็นโค้งปกติ

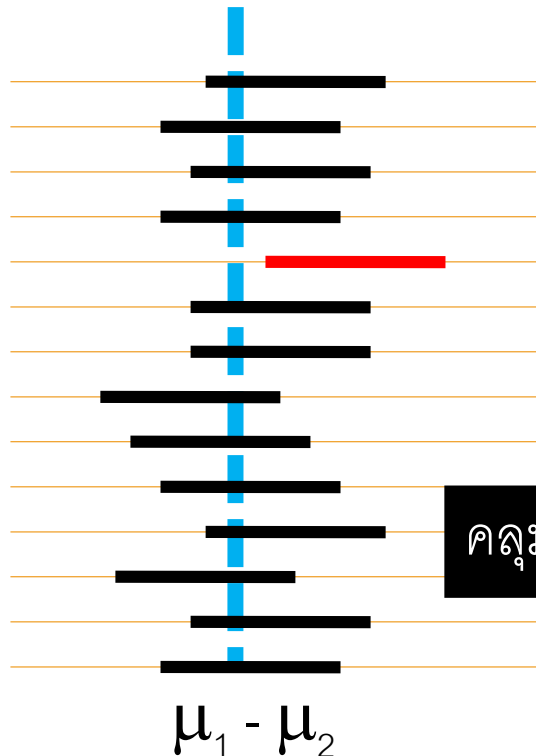
การเปรียบเทียบรายคู่: เปรียบเทียบวิธีต่างๆ



- LSD คือ การทำ Independent t -test ดังนั้นห้ามใช้
- Bonferroni, Sidak ไม่แนะนำในการเปรียบเทียบรายคู่ที่เป็นไปได้ทั้งหมด
- Tukey แนะนำมากในการเปรียบเทียบรายคู่ที่เป็นไปได้ทั้งหมด
- Games-Howell ใช้แทน Tukey เมื่อความแปรปรวนแตกต่างกันระหว่างกลุ่มสูงมาก และจำนวนกลุ่มตัวอย่างแต่ละกลุ่มไม่เท่ากัน

ช่วงเชื่อมั่นพร้อมกัน

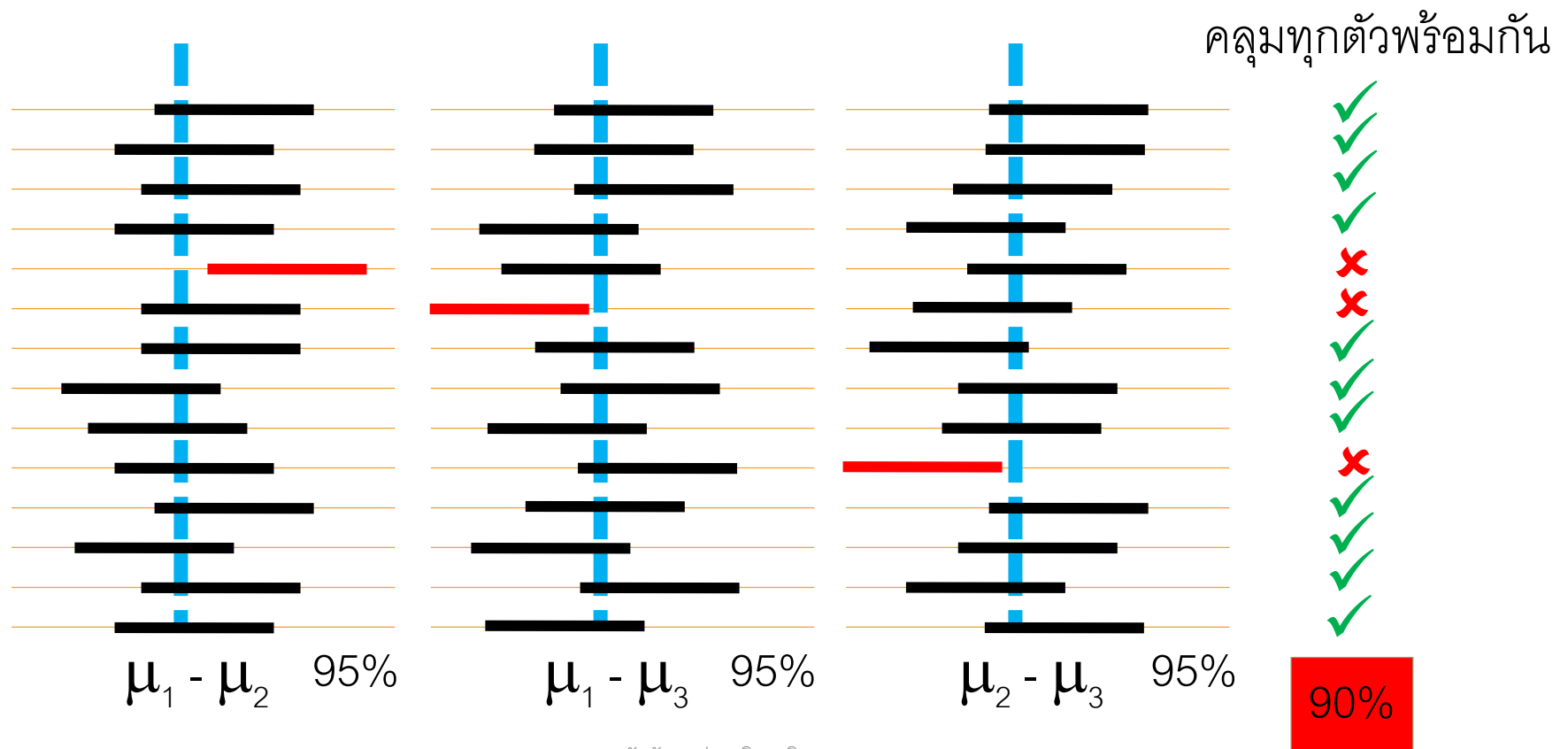
- หากสร้างช่วงเชื่อมั่นระดับ .95 โอกาสที่ช่วงเชื่อมั่นจะคลุมค่า Parameter มีทั้งหมด 95 ครั้งใน 100 ครั้ง



หากตรวจสอบช่วงเชื่อมั่นหลายช่วงพร้อมกัน
โอกาสที่จะเจอทุกช่วงที่ตรวจสอบคลุมค่า
Parameter พร้อมกัน ไม่ถึง 95%

ช่วงเชื่อมั่นพร้อมกัน

- ตรวจสอบโอกาสที่จะคลุม Parameter ทุกตัวพร้อมกัน



ช่วงเชื่อมั่นพร้อมกัน

- ปรับโอกาสให้กลุ่มทุกตัวเท่ากันให้ใกล้เคียง 95% โดยทำให้ความกว้างของช่วงแต่ละช่วงมากขึ้น
- การสร้างช่วงเชื่อมั่นแบบนี้ จะเรียกว่า ช่วงเชื่อมั่นพร้อมกัน (Simultaneous Confidence Interval)
- วิธีการปรับเหมือนการเปรียบเทียบรายคู่ เช่น วิธีของ Tukey, Sidak, Bonferroni

ช่วงเชื่อมั่นพร้อมกัน

- เช่น วิธีของ Tukey

Simultaneous Confidence Interval

Multiple Comparisons

Dependent Variable: Attitude
Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
High School	Freshman	.575*	.076	.000	.38	.77
	Sophomore	.020	.076	.994	-.18	.22
	Adults	1.180*	.096	.000	.93	1.43
Freshman	High School	-.575*	.076	.000	-.77	-.38
	Sophomore	-.555*	.048	.000	-.68	-.43
	Adults	.605*	.076	.000	.41	.80
Sophomore	High School	-.020	.076	.994	-.22	.18
	Freshman	.555*	.048	.000	.43	.68
	Adults	1.160*	.076	.000	.96	1.36
Adults	High School	-1.180*	.096	.000	-1.43	-.93
	Freshman	-.605*	.076	.000	-.80	-.41
	Sophomore	-1.160*	.076	.000	-1.36	-.96

*. The mean difference is significant at the 0.05 level.

ตรวจสอบว่าแตกต่างกัน
อย่างมีนัยสำคัญหรือไม่
โดยดูว่าช่วงเชื่อมั่นคลุม 0
หรือไม่

ใช้ SPSS คำนวณหา Simultaneous Confidence Interval ของวิธีอื่นได้

การเปรียบเทียบรายคู่: ขนาดอิทธิพล

- ในการเปรียบเทียบรายคู่นี้ สามารถ Cohen's d ได้
- หากความแปรปรวนเท่ากัน สามารถใช้ส่วนเบี่ยงเบนมาตรฐานร่วมได้

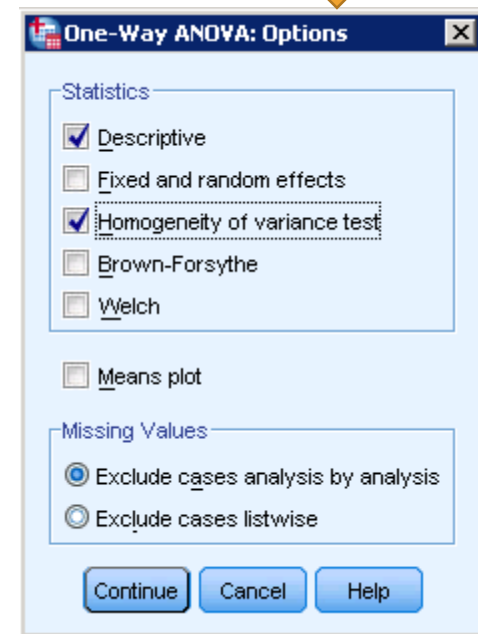
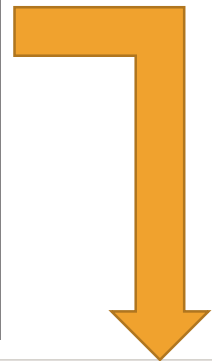
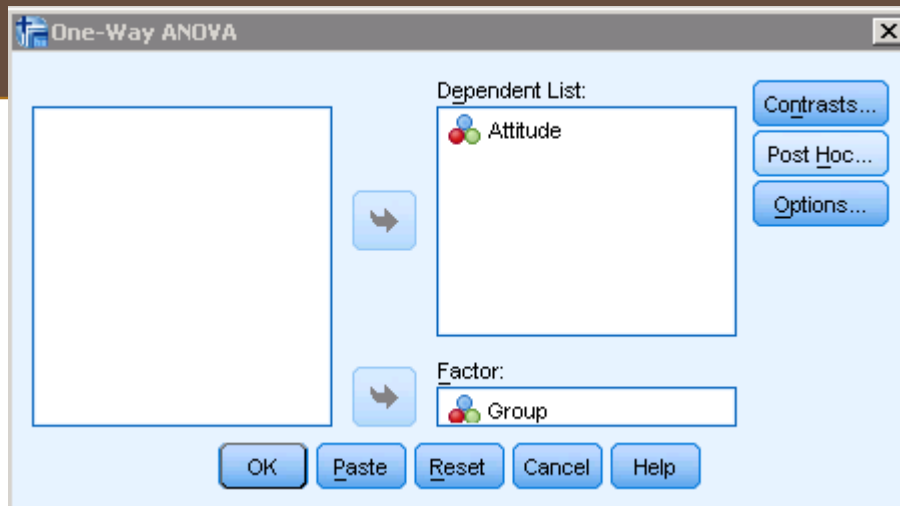
$$\delta = \frac{\mu_1 - \mu_2}{\sigma}$$

$$d = \frac{M_1 - M_2}{SD_{\text{Pooled}}} = \frac{M_1 - M_2}{\sqrt{MS_{\text{Error}}}}$$

- หากความแปรปรวนไม่เท่ากัน ใช้ความแปรปรวนของกลุ่มควบคุม หากไม่มีกลุ่มควบคุมก็อาจใช้ส่วนเบี่ยงเบนมาตรฐานร่วมได้
- ตรวจสอบว่าความแปรปรวนเท่ากันหรือไม่ ด้วย Levene Test

การเปรียบเทียบรายคู่: ขนาดอิทธิพล

ความคิดเห็นต่อการมีห้องเชียร์
เปรียบเทียบกันระหว่าง 4 กลุ่ม
(ก) นักเรียนม.6 (ข) นิสิตปี 1
(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป
ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง
5 เห็นด้วยอย่างยิ่ง



การเปรียบเทียบรายคู่: ขนาดอิทธิพล

ความคิดเห็นต่อการมีห้องเขียว
เปรียบเทียบกันระหว่าง 4 กลุ่ม
(ก) นักเรียนม.6 (ข) นิสิตปี 1
(ค) นิสิตปี 2 (ง) ประชาชนทั่วไป
ประเมิน 1 ไม่เห็นด้วยอย่างยิ่ง
5 เห็นด้วยอย่างยิ่ง



Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
106.757	3	496	.000

Levene Test ถึงระดับนัยสำคัญ ใช้
ส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุม

ไม่มีกลุ่มควบคุมชัดเจน ใช้ส่วนเบี่ยงเบนมาตรฐานร่วม

$$SD_{\text{Pooled}} = \sqrt{MS_{\text{Error}}} = \sqrt{0.231} = 0.481$$

การเปรียบเทียบรายคู่: ขนาดอิทธิพล

Descriptives

Attitude								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
High School	50	2.58	.609	.086	2.41	2.75	1	4
Freshman	200	2.01	.340	.024	1.96	2.05	1	3
Sophomore	200	2.56	.555	.039	2.48	2.64	1	4
Adults	50	1.40	.495	.070	1.26	1.54	1	2
Total	500	2.22	.612	.027	2.17	2.28	1	4

เปรียบเทียบ ม.6 และปี 1:

$$d = (2.58 - 2.01)/0.481 = 1.19$$

เปรียบเทียบ ม.6 และปี 2:

$$d = (2.58 - 2.56)/0.481 = 0.04$$

เปรียบเทียบ ม.6 และประชาชนทั่วไป:

$$d = (2.58 - 1.40)/0.481 = 2.46$$

การเปรียบเทียบรายคู่: ขนาดอิทธิพล

Descriptives

Attitude								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
High School	50	2.58	.609	.086	2.41	2.75	1	4
Freshman	200	2.01	.340	.024	1.96	2.05	1	3
Sophomore	200	2.56	.555	.039	2.48	2.64	1	4
Adults	50	1.40	.495	.070	1.26	1.54	1	2
Total	500	2.22	.612	.027	2.17	2.28	1	4

เปรียบเทียบ ปี 1 และปี 2:

$$d = (2.01 - 2.56)/0.481 = -1.14$$

เปรียบเทียบ ปี 1 และประชาชนทั่วไป:

$$d = (2.01 - 1.40)/0.481 = 1.27$$

เปรียบเทียบ ปี 2 และประชาชนทั่วไป:

$$d = (2.56 - 1.40)/0.481 = 2.41$$

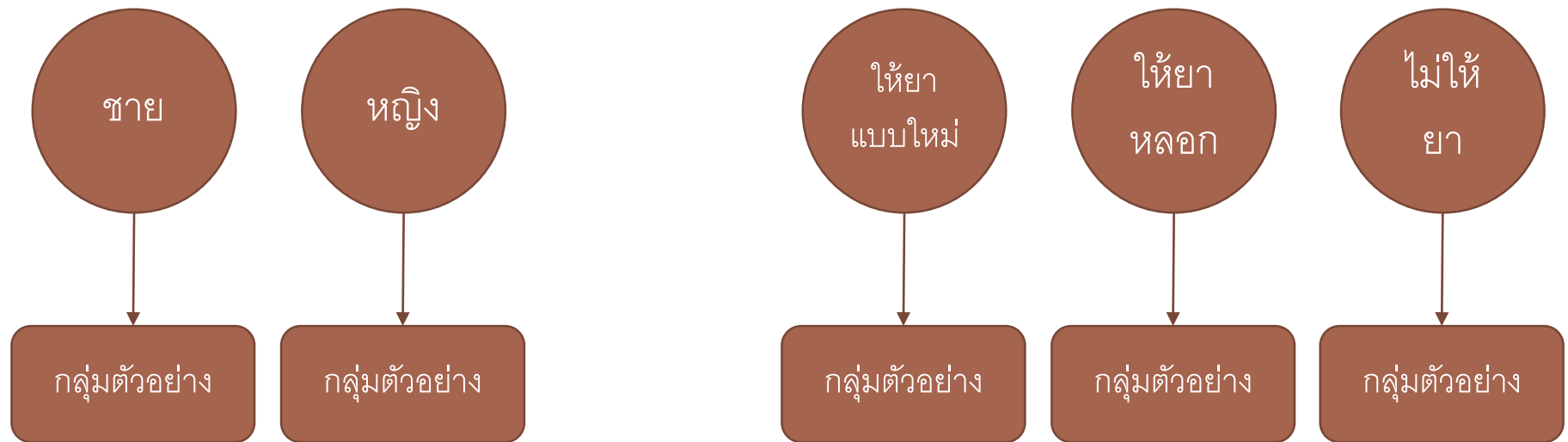
รูปแบบการขยายผล

สถิติขั้นกลางสำหรับจิตวิทยา

สันหัด พรประเสริฐมานิต

รูปแบบการขยายผล

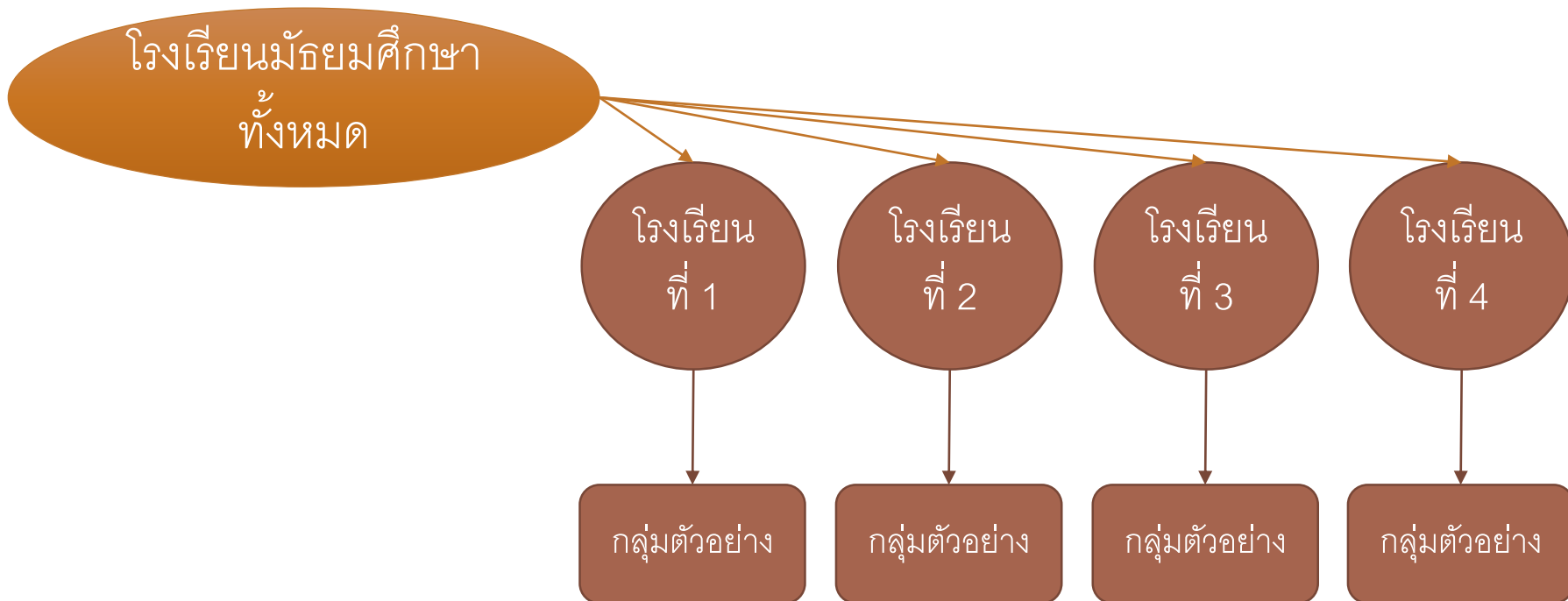
- ที่ผ่านมา กลุ่มได้ถูกกำหนดโดยผู้วิจัยไว้แล้ว ผลจากการเปรียบเทียบคือดูว่าประชากรของกลุ่มมีค่าเฉลี่ยแตกต่างกันหรือไม่



- เรียกการขยายผลแบบนี้ว่า การหาอิทธิพลของกลุ่มที่กำหนดไว้ (Fixed Effect)

รูปแบบการขยายผล

- บางครั้ง กลุ่มถูกสุ่มออกมาจากประชากรของกลุ่มอีกทีหนึ่ง



รูปแบบการขยายผล

- ในกรณีนี้ สามารถขยายผลได้ 2 รูปแบบ
 - ถ้าต้องการสรุปว่า โรงเรียนทั้ง 4 โรงเรียนนี้มีตัวแปรตามแตกต่างกัน จะเป็นการขยายผลของอิทธิพลของกลุ่มที่กำหนดไว้ (Fixed Effect)
 - ถ้าต้องการสรุปว่า โรงเรียนมัธยมศึกษาทั้งหมดมีตัวแปรตามแตกต่างกัน (คิดว่าโรงเรียนถูกสุ่มออกมาอีกทีหนึ่ง) จะเป็นการขยายผลของอิทธิพลของกลุ่มแบบสุ่ม (Random Effect)
- ทั้ง Independent t -test, One-way ANOVA, และสถิติอื่นๆ ที่สอนในวิชานี้ จะเป็น Fixed Effect
- ใช้การวิเคราะห์พหุระดับ (Multilevel Model) เพื่อทดสอบ Random Effect

คาบต่อไป

- ส่งการบ้านที่ 2 (ส่งในสัปดาห์ที่ 25-29 ม.ค. 59)
- เรียนคาบปฏิบัติการครั้งที่ 2