

สหสัมพันธ์

สถิติสำหรับจิตวิทยา 1

สันหัด พรประเสริฐมานิต

โครงร่างการนำเสนอ

- มโนทัศน์พื้นฐาน
- ค่าสหสัมพันธ์
- การทดสอบสมมติฐาน
- ข้อตกลงเบื้องต้นก่อนการใช้สถิติ
- ช่วงเชื่อมั่น
- ขนาดอิทธิพล
- กำลังและการกำหนดขนาดกลุ่มตัวอย่าง
- ความแตกต่างและความสัมพันธ์
- ปัจจัยที่เบี่ยงเบนค่าสหสัมพันธ์
- ความสัมพันธ์และการอนุมานเชิงสาเหตุ

แนะนำ

- หากตัวแปรหนึ่งเปลี่ยนแปลงแล้ว สังเกตว่าอีกตัวแปรหนึ่งเปลี่ยนแปลงด้วย จะเรียกว่าตัวแปรสองตัวสัมพันธ์กัน
 - จำนวนการสูบบุหรี่สัมพันธ์กับการลดลงของเม็ดเลือดขาว
 - ส่วนสูงสัมพันธ์กับน้ำหนัก
 - ผลการเรียนเฉลี่ยจากระดับปริญญาตรีสัมพันธ์กับผลการปฏิบัติงาน
- ในบทนี้จะกล่าวถึง วิธีการหาความสัมพันธ์ในเชิงตัวเลข และวิธีการทดสอบความสัมพันธ์ว่ามีจริงหรือไม่ในประชากร

มโนทัศน์พื้นฐาน

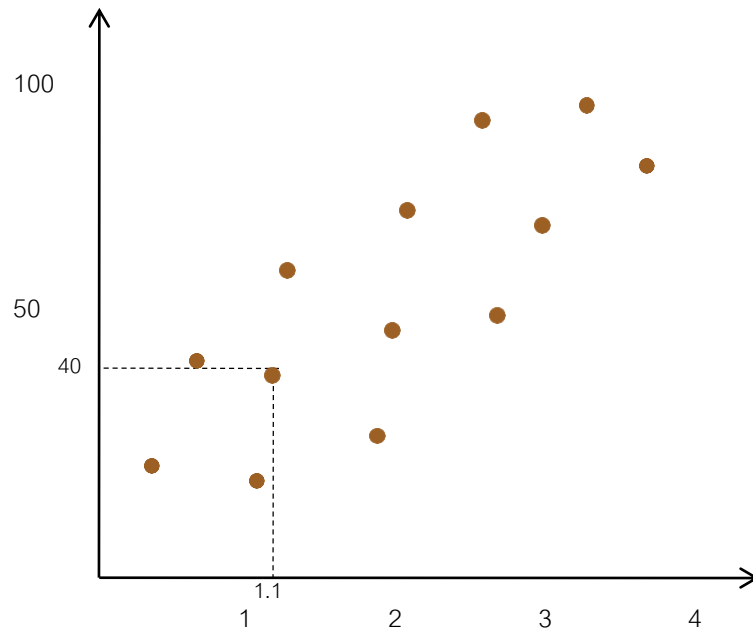
สถิติขั้นนำสำหรับจิตวิทยา

อ. สันหัต พรประเสริฐมานิต

มโนทัศน์พื้นฐาน

- แผนภาพการกระจาย (Scatterplot) เป็นการแสดงการกระจายของข้อมูลสองตัวแปรพร้อมกัน

คะแนนสอบ

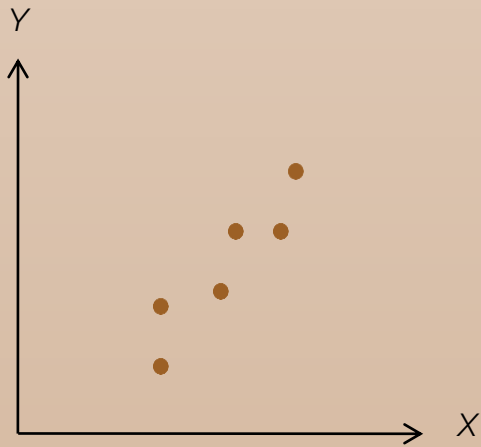


ชั่วโมงการอ่านหนังสือ (ชม.)

มโนทัศน์พื้นฐาน

- แผนภาพการกระจาย จะทำให้เห็นทิศทางของความสัมพันธ์ได้โดยง่าย

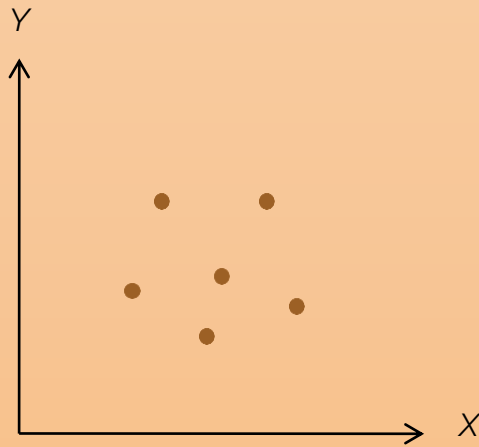
ความสัมพันธ์ทางบวก



$x \uparrow$ $y \uparrow$

$x \downarrow$ $y \downarrow$

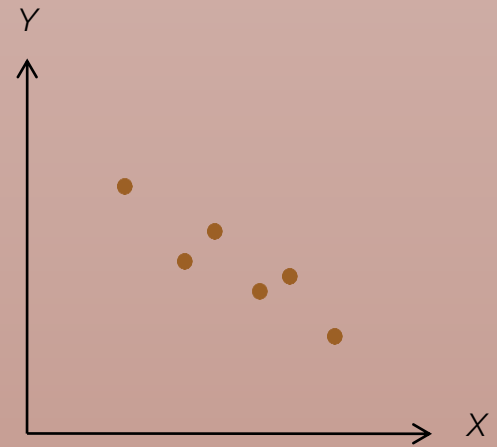
ไม่มีความสัมพันธ์



$x \uparrow$ $y \text{—}$

$x \downarrow$ $y \text{—}$

ความสัมพันธ์ทางลบ



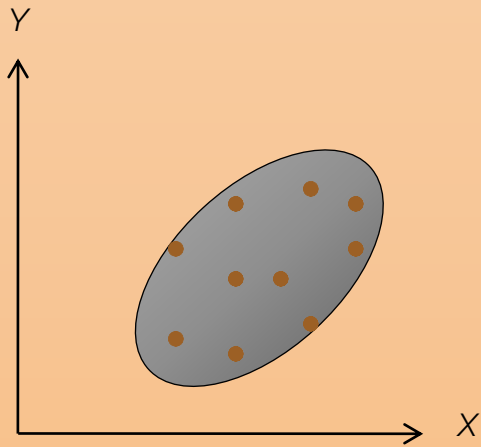
$x \uparrow$ $y \downarrow$

$x \downarrow$ $y \uparrow$

มโนทัศน์พื้นฐาน

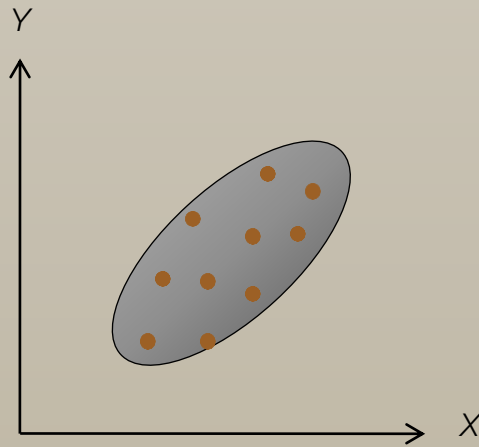
- การจับกลุ่มของจุด สามารถบอกขนาดความสัมพันธ์

ความสัมพันธ์น้อย



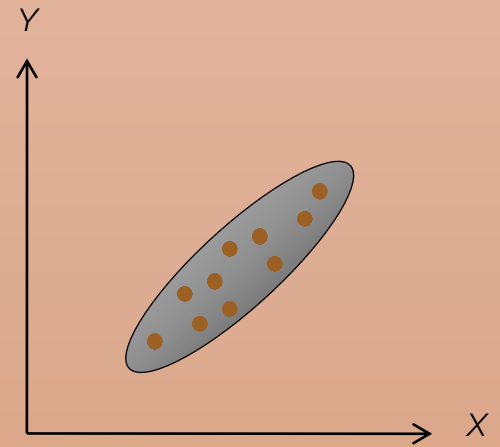
ค่าของ X ทำนาย
Y ได้ น้อย (ไม่แม่นยำ)

ความสัมพันธ์ปานกลาง



ค่าของ X ทำนาย
Y ได้ ปานกลาง

ความสัมพันธ์สูง

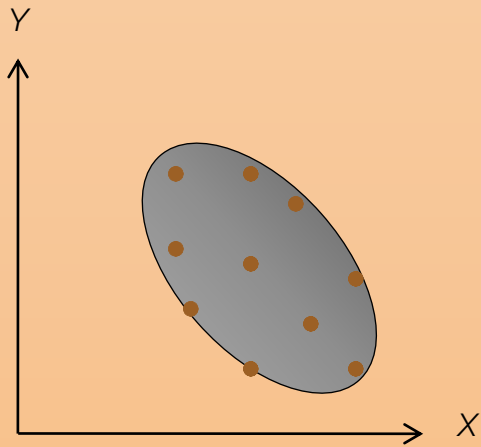


ค่าของ X ทำนาย
Y ได้ ปานสูง (แม่นยำ)

มโนทัศน์พื้นฐาน

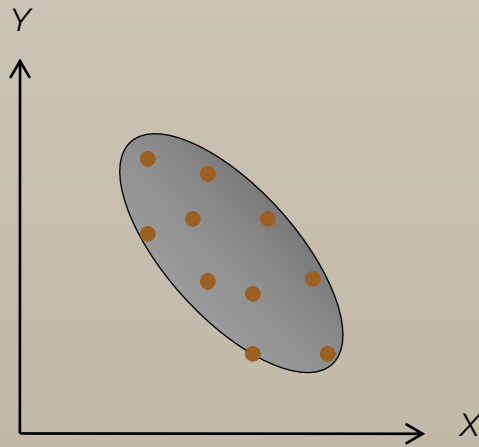
- การจับกลุ่มของจุด สามารถบอกขนาดความสัมพันธ์

ความสัมพันธ์น้อย



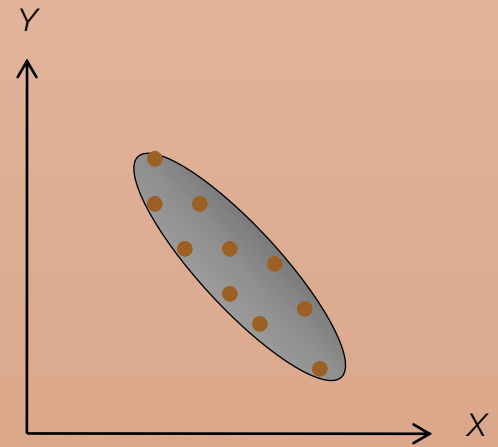
ค่าของ X ทำนาย
Y ได้น้อย (ไม่แม่นยำ)

ความสัมพันธ์ปานกลาง



ค่าของ X ทำนาย
Y ได้ปานกลาง

ความสัมพันธ์สูง

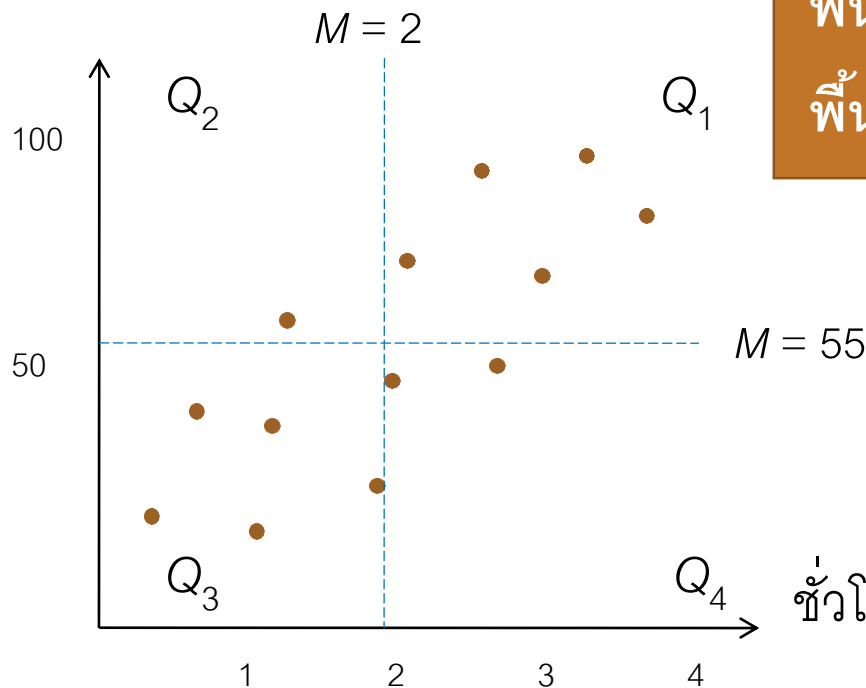


ค่าของ X ทำนาย
Y ได้ปานสูง (แม่นยำ)

มโนทัศน์พื้นฐาน

- การมองทิศทาง และขนาดของความสัมพันธ์ สามารถมองได้โดยการแบ่งแผนภาพการกระจายเป็น 4 ส่วน

คะแนนสอบ



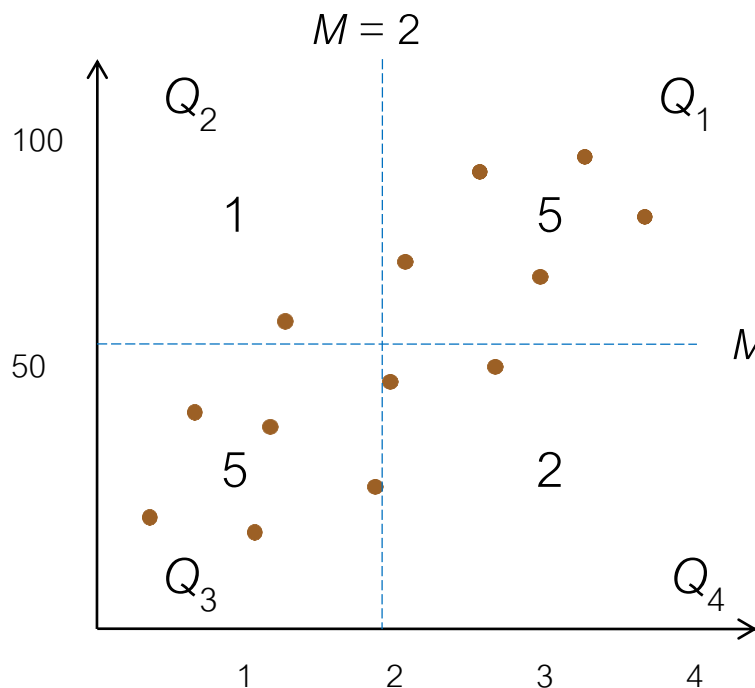
พื้นที่ 1 และ 3 → ความสัมพันธ์ทางบวก
พื้นที่ 2 และ 4 → ความสัมพันธ์ทางลบ

ชั่วโมงการอ่านหนังสือ (ชม.)

มโนทัศน์พื้นฐาน

- การมองทิศทาง และขนาดของความสัมพันธ์ สามารถมองได้โดยการแบ่งแผนภาพการกระจายเป็น 4 ส่วน

คะแนนสอบ



$(Q_1 + Q_3) = (Q_2 + Q_4) \rightarrow$ ไม่สัมพันธ์
 $(Q_1 + Q_3) > (Q_2 + Q_4) \rightarrow$ สัมพันธ์ทางบวก
 $(Q_1 + Q_3) < (Q_2 + Q_4) \rightarrow$ สัมพันธ์ทางลบ



สัมพันธ์ทางบวก

ชั่วโมงการอ่านหนังสือ (ชม.)

ค่าสหสัมพันธ์

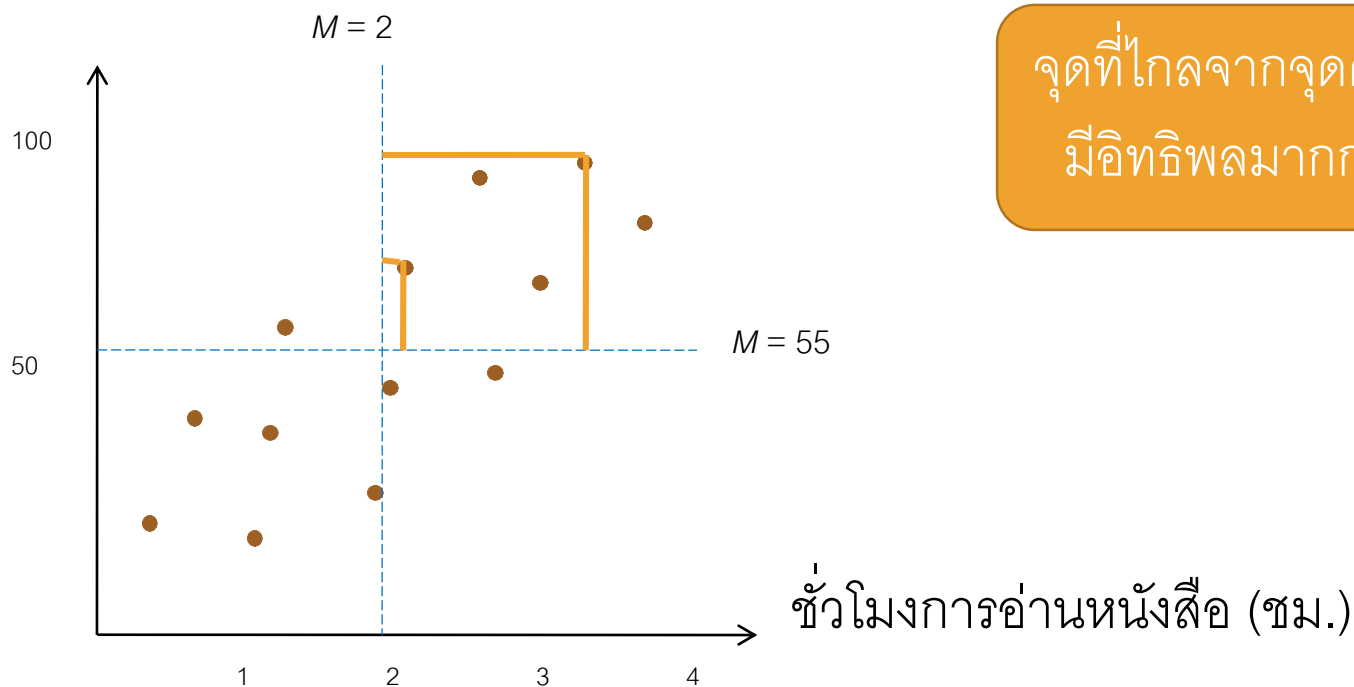
สถิติขั้นนำสำหรับจิตวิทยา

อ. สันหัต พรประเสริฐมานิต

ค่าสหสัมพันธ์

- การนับจำนวนจุด ไม่ได้คำนึงถึงว่าแต่ละจุด มีอิทธิพลต่อความสัมพันธ์ แตกต่างกัน

คะแนนสอบ



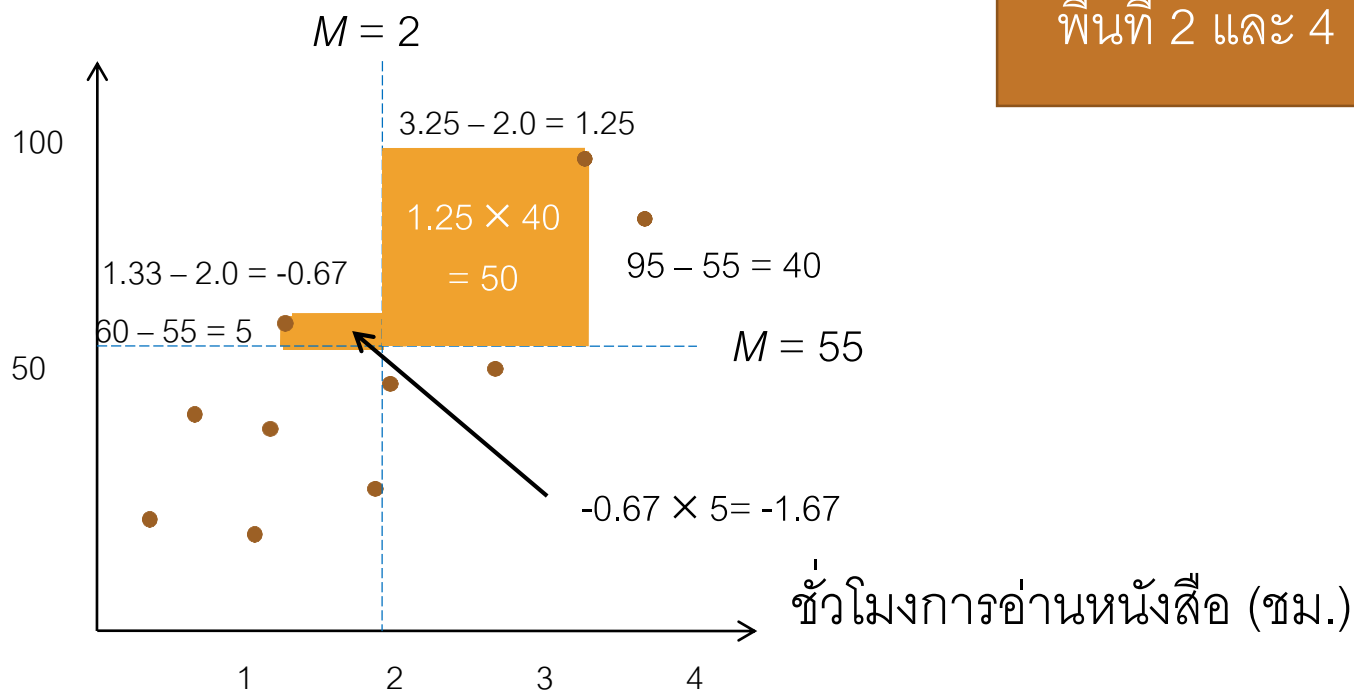
จุดที่ไกลจากจุดศูนย์กลาง จะมีอิทธิพลมากกว่าจุดที่ใกล้

ค่าสหสัมพันธ์

- วิธีการแก้ไข คือ หาค่าเบี่ยงเบนในรูปแบบพื้นที่

$$(X - M_X) \times (Y - M_Y)$$

คะแนนสอบ

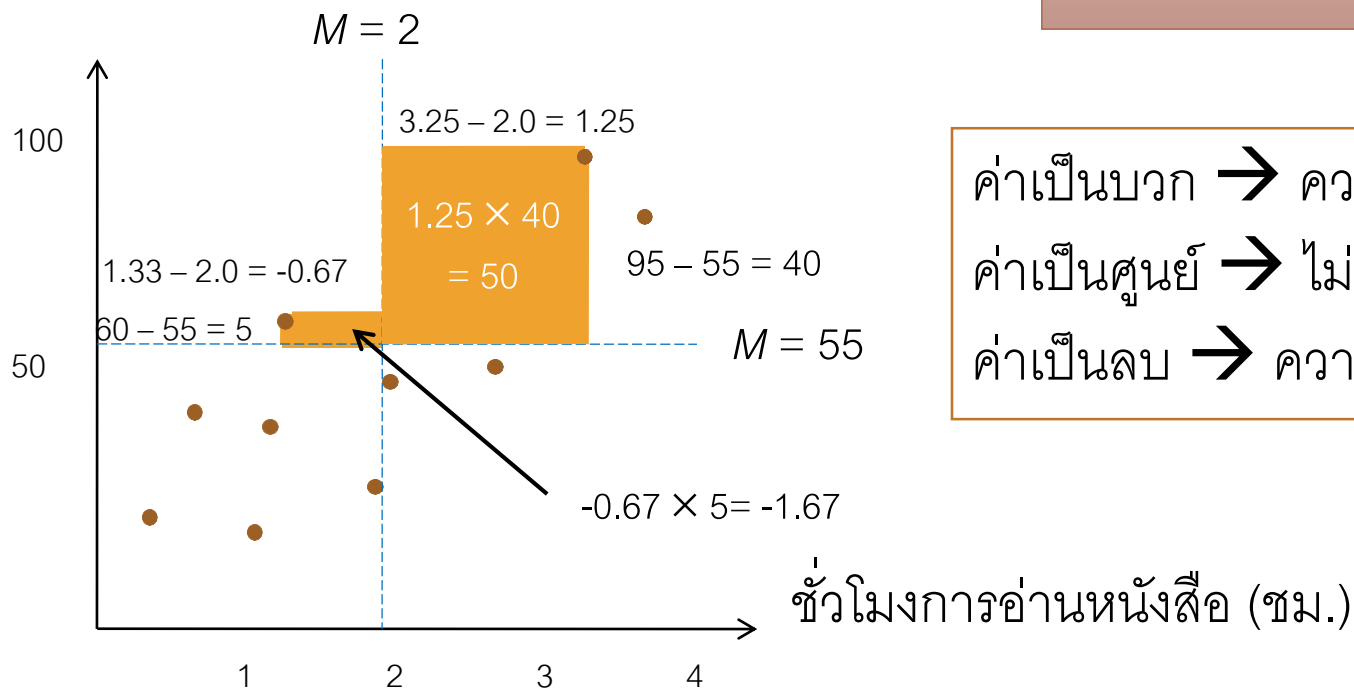


พื้นที่ 1 และ 3 ผลคูณเป็นค่าบวก
พื้นที่ 2 และ 4 ผลคูณเป็นค่าลบ

ค่าสหสัมพันธ์

- ค่าเฉลี่ยของความเบี่ยงเบนรูปพื้นที่ของทุกจุด เรียกว่า ความแปรปรวนร่วม (Covariance, s_{XY})

คะแนนสอบ



$$s_{XY} = \frac{\sum (X - M_X)(Y - M_Y)}{n - 1}$$

ค่าเป็นบวก \rightarrow ความสัมพันธ์ทางบวก
ค่าเป็นศูนย์ \rightarrow ไม่มีความสัมพันธ์
ค่าเป็นลบ \rightarrow ความสัมพันธ์ทางลบ

ค่าสหสัมพันธ์

- เนื่องจากความแปรปรวนร่วมได้รับอิทธิพลจากหน่วยการวัด ทำให้เปรียบเทียบความสัมพันธ์ต่างคู่กันไม่ได้
 - เช่น ความสัมพันธ์ระหว่างเงินเดือนและผลงาน หน่วยเป็น บาท คะแนน
 - ความสัมพันธ์ระหว่างเงินเดือนและชั่วโมงการทำงาน หน่วยเป็น บาท ชั่วโมง
- เมื่อนำความแปรปรวนร่วมมาหารด้วยส่วนเบี่ยงเบนมาตรฐานของตัวแปรทั้งสอง จะทำให้ค่าตัวเลขต่างคู่เปรียบเทียบกันได้
- เรียกค่านั้นว่า ค่าสหสัมพันธ์ (Correlation)

ค่าสหสัมพันธ์

- ค่าสหสัมพันธ์ (Correlation)

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

- ค่าที่เป็นไปได้ได้อยู่ในช่วง -1 ถึง 1
- เครื่องหมายบวกลบ บอกทิศทางความสัมพันธ์
- ค่าสัมบูรณ์ของค่าสหสัมพันธ์บอกขนาดความสัมพันธ์
 - .20 → น้อย .50 → ปานกลาง .80 → มาก
- ชื่อเต็มเรียกว่า Pearson's Product Moment Correlation

ค่าสหสัมพันธ์

- เมทริกซ์สหสัมพันธ์ (Correlation Matrix)

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

เมทริกซ์มีลักษณะสมมาตร

ความสัมพันธ์กับตนเองเท่ากับ 1

ค่าสหสัมพันธ์

- เมทริกซ์ความแปรปรวนร่วม (Covariance Matrix)

	A	B	C
A	225	75	-9
B	75	100	-9
C	-9	-9	4

เมทริกซ์มีลักษณะสมมาตร

แนวทแยงแสดงถึงความแปรปรวน หรือความแปรปรวนร่วมกับตนเอง

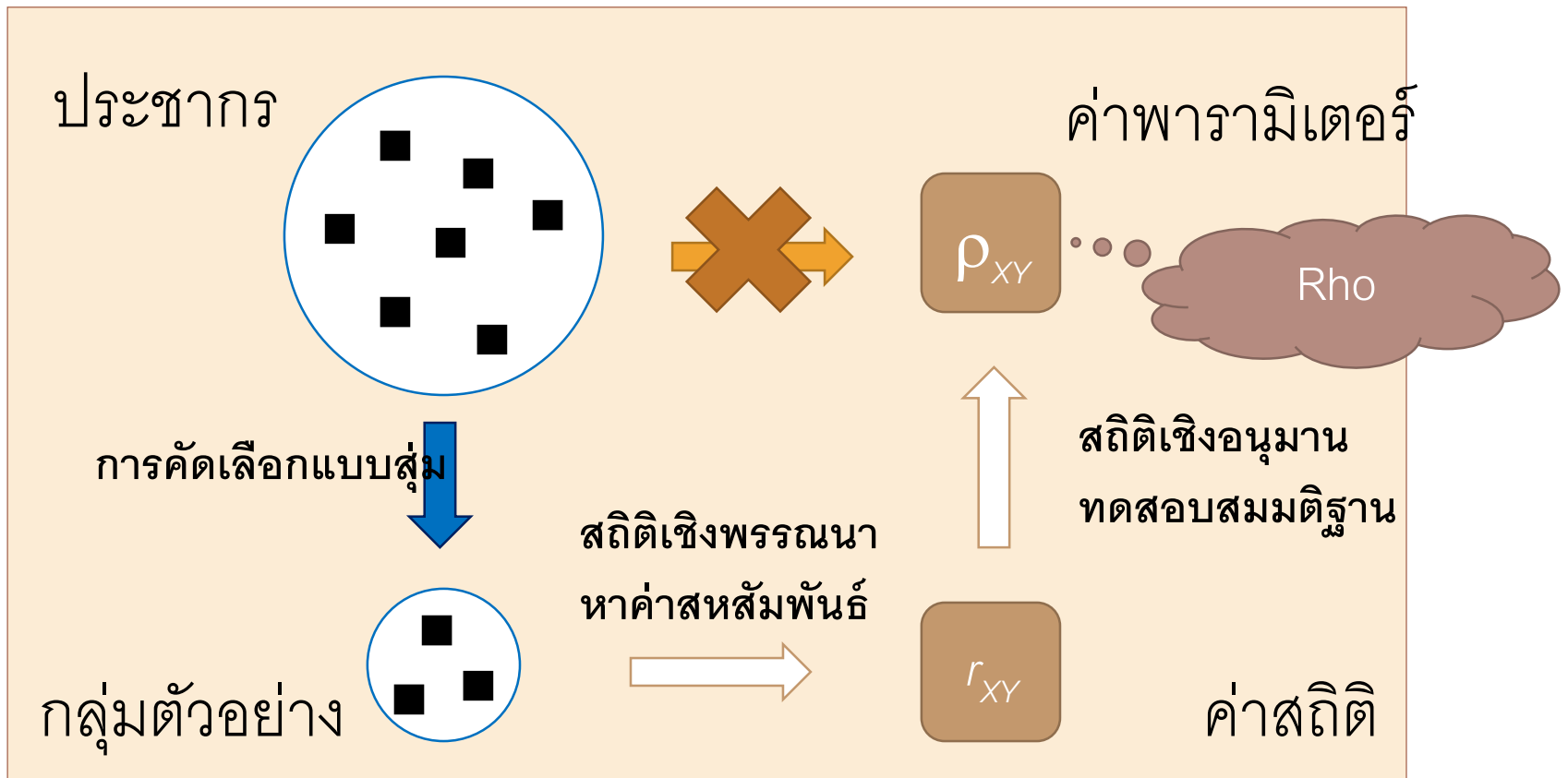
การทดสอบสมมติฐานของค่า สหสัมพันธ์

สถิติขั้นนำสำหรับจิตวิทยา

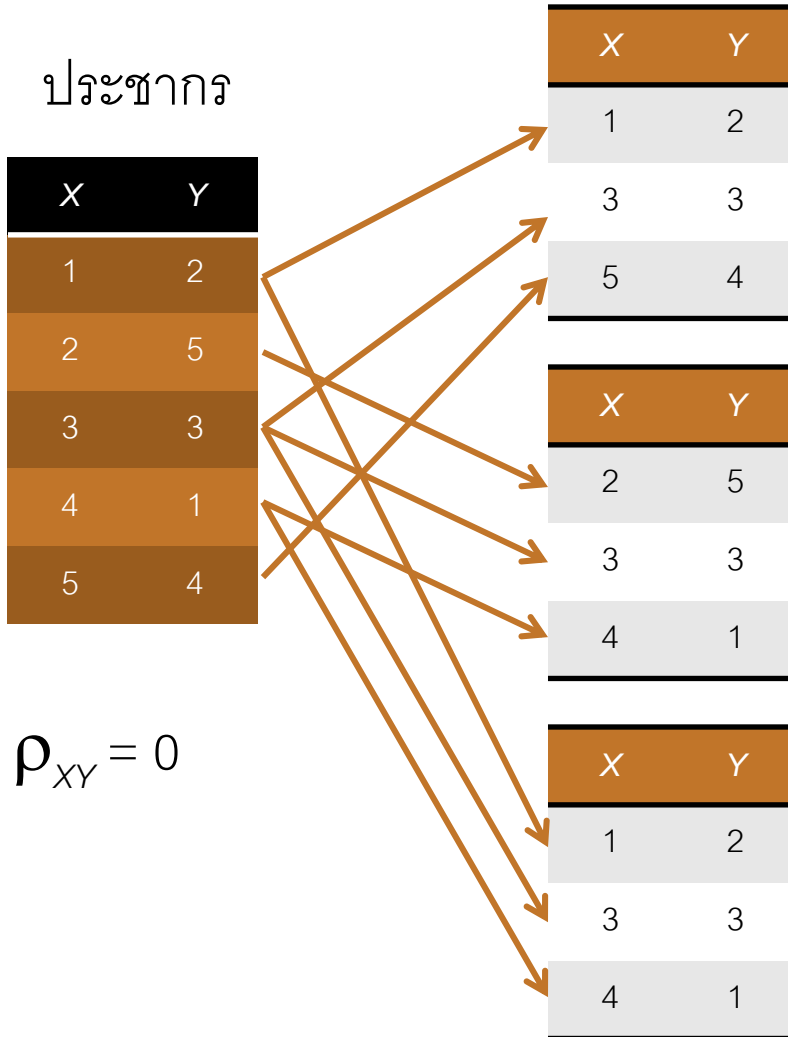
อ. สันหัต พรประเสริฐมานิต

การทดสอบสมมติฐาน

- สถิติจะมีความหมายเมื่ออ้างอิงจากกลุ่มตัวอย่างไปหาประชากร



การทดสอบสมมติฐาน



$$r_{XY} = 1$$

$$r_{XY} = -1$$

$$r_{XY} = -.33$$

ความแตกต่างที่เกิด
จากการสุ่ม
(Sampling
Error)

การทดสอบสมมติฐาน

- ความสัมพันธ์ที่ได้ในกลุ่มตัวอย่างอาจเกิดจากการสุ่ม (Sampling Error)
- เช่น ความสัมพันธ์ของตัวแปรทั้งสองมีค่าเป็นบวกในกลุ่มตัวอย่าง เพราะสุ่มเจอคนที่มีคะแนนสูงสองตัวพร้อมกัน และต่ำสองตัวพร้อมกัน
- ค่า p (p value) คือ ค่าที่บอกว่าโอกาสที่สุ่มประชากรที่ตัวแปรทั้งสองไม่สัมพันธ์กัน แล้วได้ค่าสหสัมพันธ์ขนาดนี้ มีโอกาสเท่าไร
- ถ้าน้อยกว่าระดับนัยสำคัญ \rightarrow Reject H_0
- ถ้ามากกว่าระดับนัยสำคัญ \rightarrow Fail to Reject H_0

ค่าเปลี่ยนแปลง

- การตั้งสมมติฐานทางสถิติ

สองทาง

$$H_0: \rho_{XY} = 0$$

$$H_1: \rho_{XY} \neq 0$$

ทางเดียว

$$H_0: \rho_{XY} \geq 0$$

$$H_1: \rho_{XY} < 0$$

$$H_0: \rho_{XY} \leq 0$$

$$H_1: \rho_{XY} > 0$$

การทดสอบสมมติฐาน

- เช่น

อคติต่อสาวประเภทสอง
และจำนวนครั้งในการสบตาขณะคุย



นักวิจัยจึงสุ่มนิสิตป.ตรี 10 คน ตอบแบบสอบถาม
อคติต่อสาวประเภทสอง และสังเกตจำนวนครั้ง
ในการสบตาคู่หน้าม้าที่เป็นสาวประเภทสอง

สมมติฐาน คือ อคติต่อสาวประเภท
สองสัมพันธ์กับจำนวนครั้งในการสบตา

$$H_0: \rho_{XY} = 0$$

$$H_1: \rho_{XY} \neq 0$$

การทดสอบสมมติฐาน

- เช่น

*Untitled1 [DataSet0] - SPSS Data Editor

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

17 : eyecontact Visible: 3 of 3 Variables

	id	eyecontact	att	var	var	var	var
1	1.00	3.00	2.00				
2	2.00	5.00	7.00				
3	3.00	7.00	9.00				
4	4.00	9.00	6.00				
5	5.00	0.00	3.00				
6	6.00	2.00	5.00				
7	7.00	4.00	4.00				
8	8.00	3.00	6.00				
9	9.00	8.00	4.00				
10	10.00	7.00	7.00				
11							
12							
13							
14							

Data View Variable View

SPSS Processor is ready

การทดสอบสมมติฐาน

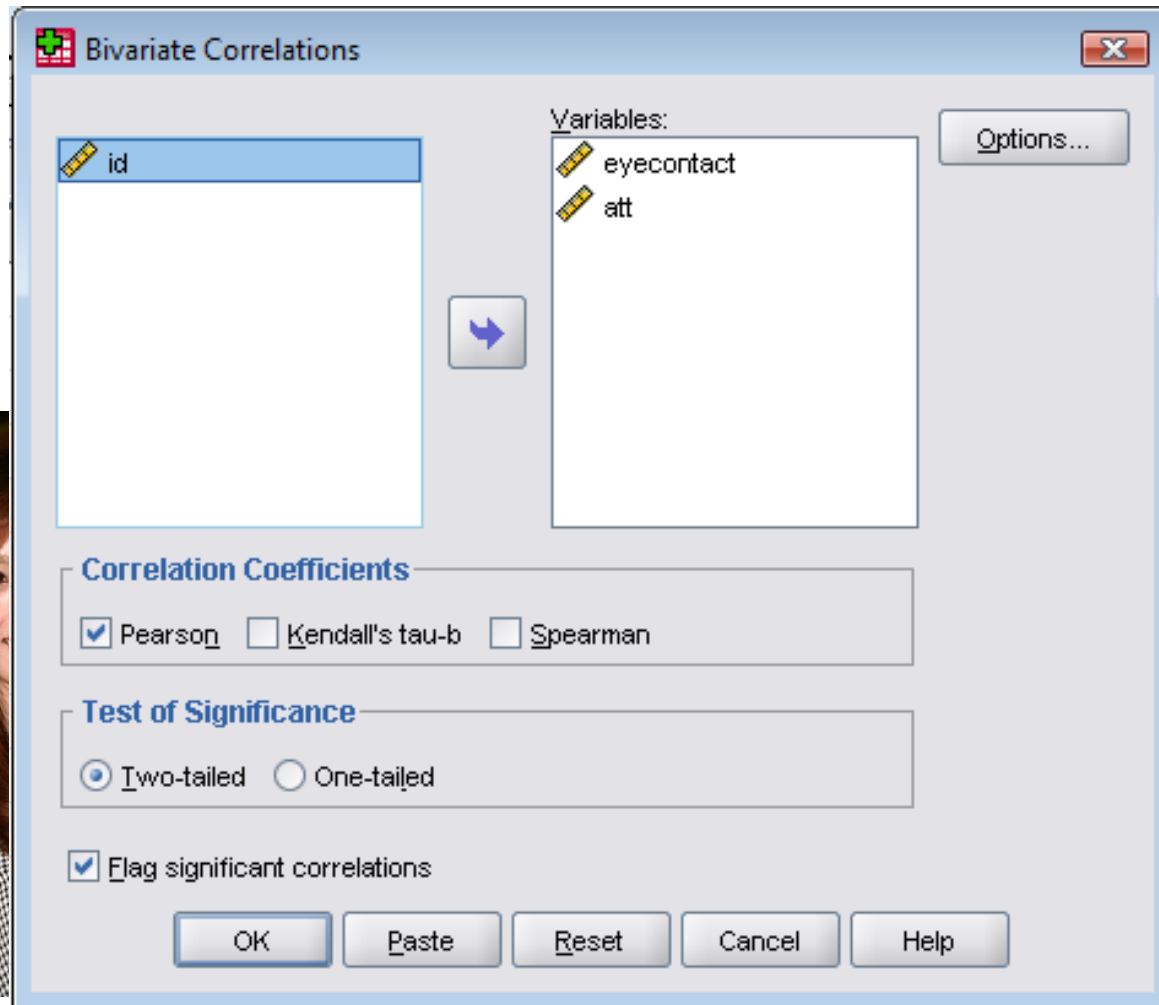
- เช่น

The screenshot shows the SPSS Data Editor interface. The main window displays a data table with 14 rows and 3 columns: 'id', 'eyecon', and an unlabeled column. The 'id' column contains values from 1.00 to 10.00. The 'eyecon' column is currently empty. The 'Analyze' menu is open, and the 'Correlate' option is selected, which has opened a submenu with three options: 'Bivariate...', 'Partial...', and 'Distances...'. The 'Bivariate...' option is highlighted. The status bar at the bottom indicates 'SPSS Processor is ready'.

	id	eyecon
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		
12		
13		
14		

การทดสอบสมมติฐาน

- เช่น



การทดสอบสมมติฐาน

- เช่น

จำนวนครั้งในการสบตา

เจตคติต่อสาวประเภทสอง

Descriptive Statistics

	Mean	Std. Deviation	N
eyecontact	4.8000	2.89828	10
att	5.3000	2.11082	10



การทดสอบสมมติฐาน

- เช่น

$r = .519, p = .124$ (สองทาง)

Correlations

		eyecontact	att
eyecontact	Pearson Correlation	1.000	.519
	Sig. (2-tailed)		.124
	N	10.000	10
att	Pearson Correlation	.519	1.000
	Sig. (2-tailed)	.124	
	N	10	10.000



เมตริกซ์สหสัมพันธ์สมมาตรกัน

สหสัมพันธ์กับตัวเองเท่ากับ 1

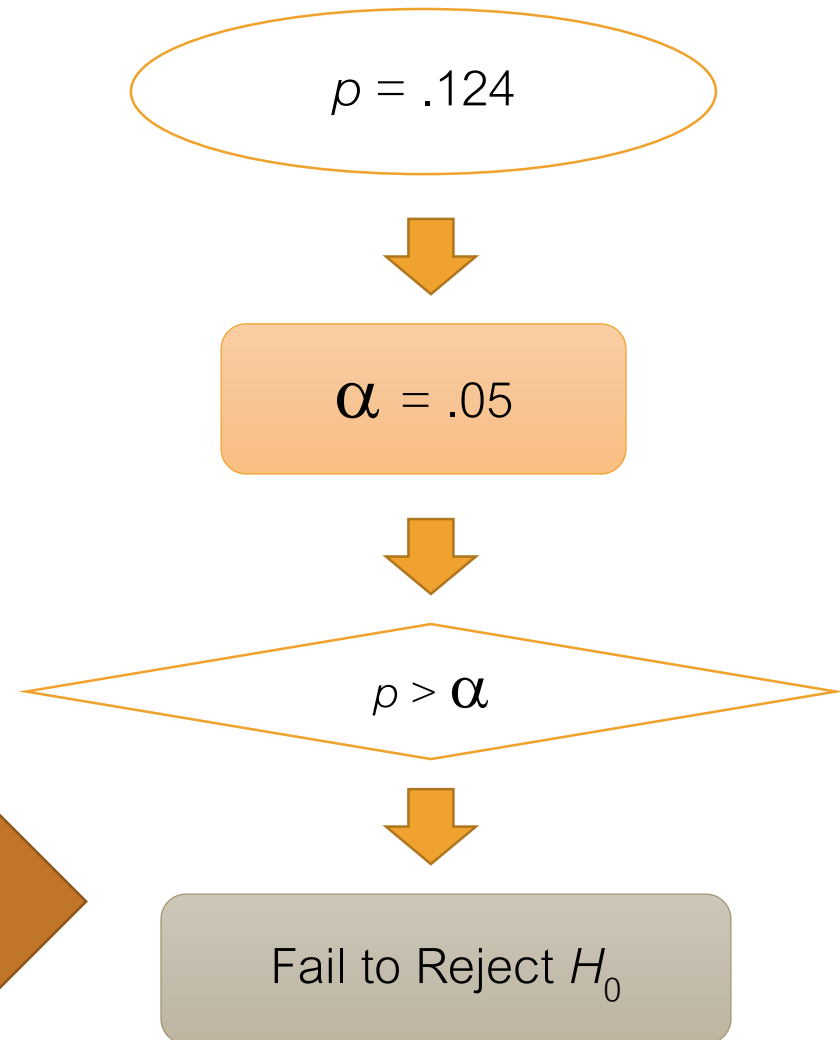
การทดสอบสมมติฐาน

- เช่น

อคติต่อสาวประเภทสอง
และจำนวนครั้งในการสบตาขณะคุย



ยังไม่สามารถสรุปได้ว่า อคติต่อสาว
ประเภทสองและจำนวนครั้งใน
การสบตาขณะคุยว่ามีความสัมพันธ์กัน



การทดสอบสมมติฐาน

- เช่น

ความขัดแย้งระหว่างบ้านและ
ที่ทำงานสัมพันธ์ทางลบ
กับผลการปฏิบัติงาน



$\alpha = .05$

นักวิจัยจึงสุ่มพนักงานบริษัทหนึ่งจำนวน 221 คน
ทำแบบสอบถามเรื่องความขัดแย้งระหว่างบ้าน
และที่ทำงาน พร้อมกับให้หัวหน้าประเมินผลงาน



สมมติฐาน คือความขัดแย้งระหว่างบ้านและ
ที่ทำงานสัมพันธ์ทางลบกับผลการปฏิบัติงาน



$H_0: \rho_{xy} \geq 0$



$H_1: \rho_{xy} < 0$

การทดสอบสมมติฐาน

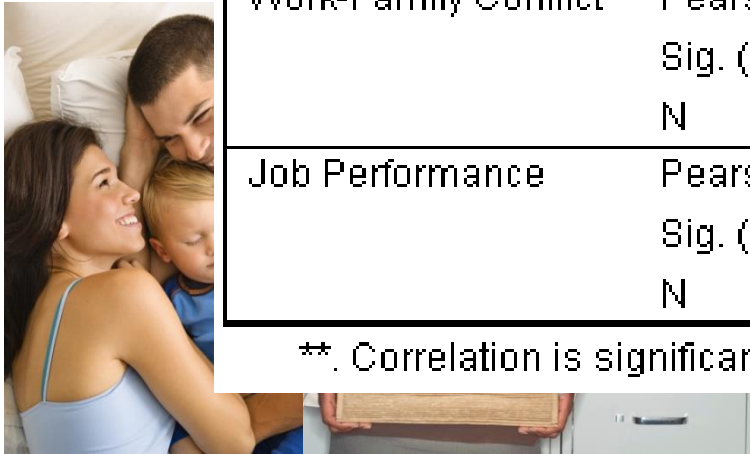
- เช่น

$r = -.214, p = .001$ (สองทาง) หรือ $p = .0005$ (ทางเดียว)

Correlations

		Work-Family Conflict	Job Performance
Work-Family Conflict	Pearson Correlation	1.000	-.214**
	Sig. (2-tailed)		.001
	N	221.000	221
Job Performance	Pearson Correlation	-.214**	1.000
	Sig. (2-tailed)	.001	
	N	221	221.000

** . Correlation is significant at the 0.01 level (2-tailed).



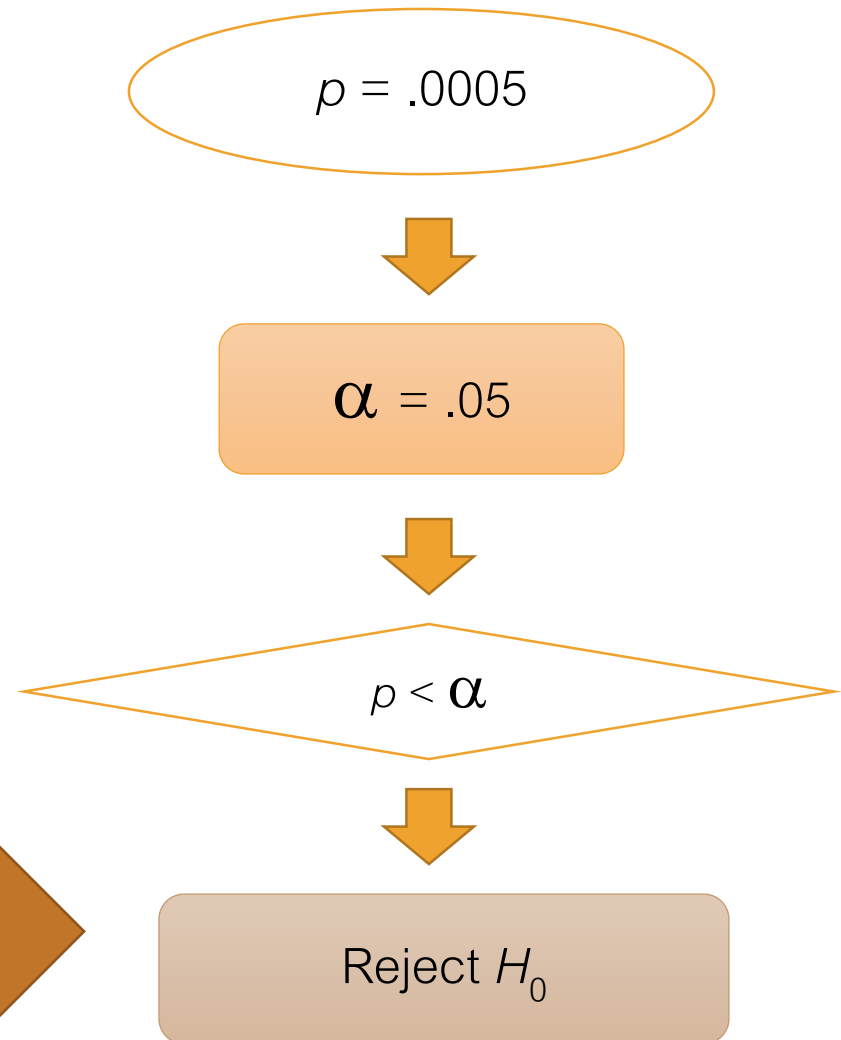
การทดสอบสมมติฐาน

- เช่น

ความขัดแย้งระหว่างบ้านและ
ที่ทำงานสัมพันธ์ทางลบ
กับผลการปฏิบัติงาน



ความขัดแย้งระหว่างบ้านและที่ทำงาน
สัมพันธ์ทางลบกับผลการปฏิบัติงานจริง



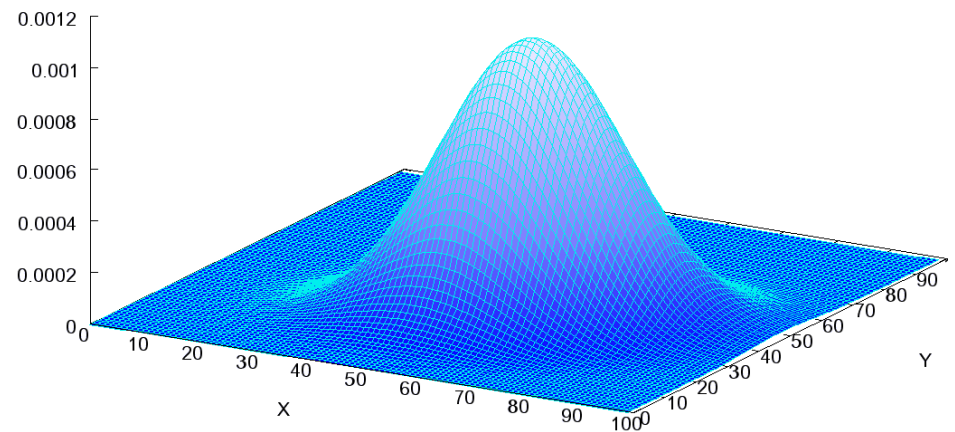
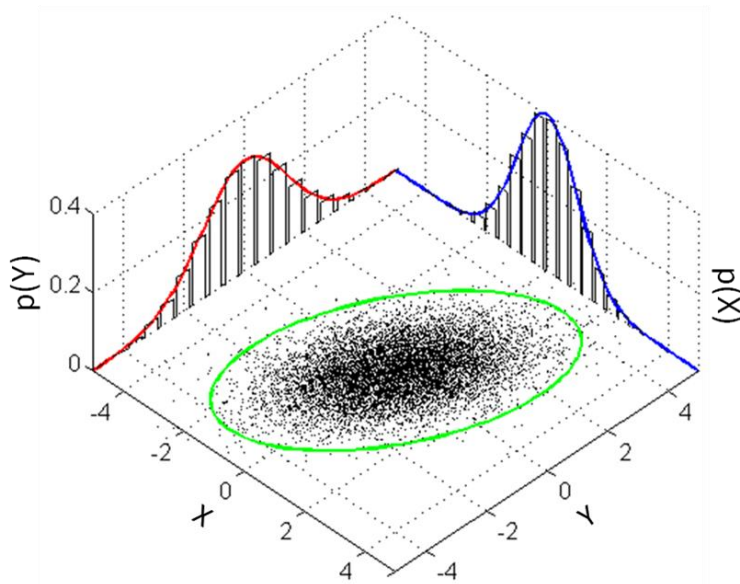
ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

สถิติขั้นนำสำหรับจิตวิทยา

อ. สันหัตถ์ พรประเสริฐมานิต

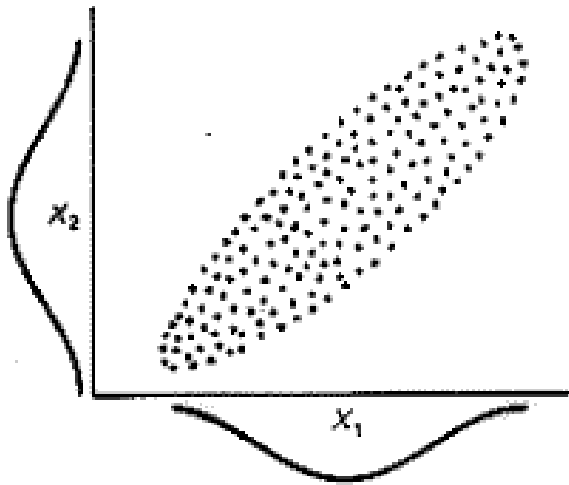
ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- การกระจายของทั้งสองตัวแปรในประชากรเป็นโค้งปกติร่วม (bivariate normality)

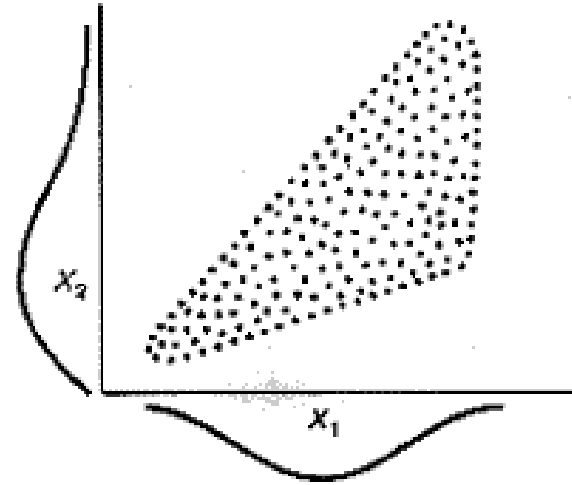


ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- อากาการเบื้องต้นที่สังเกตได้ หากข้อตกลงเบื้องต้นนี้ถูกละเมิด
 - การกระจายของแต่ละตัวแปรใดๆ (Marginal distribution) ไม่เป็นโค้งปกติ



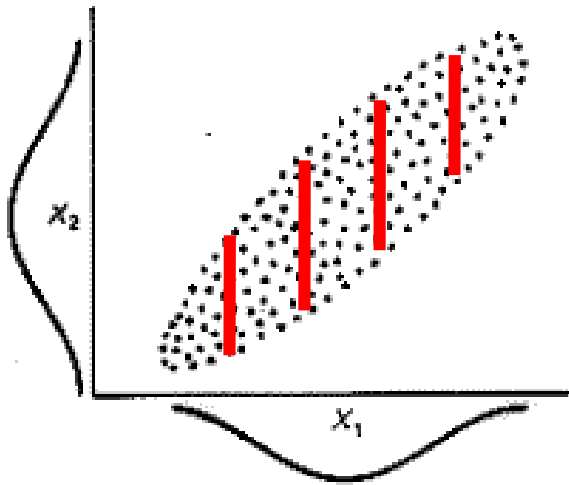
การกระจายของตัวแปร
ทั้งสองเป็นโค้งปกติ



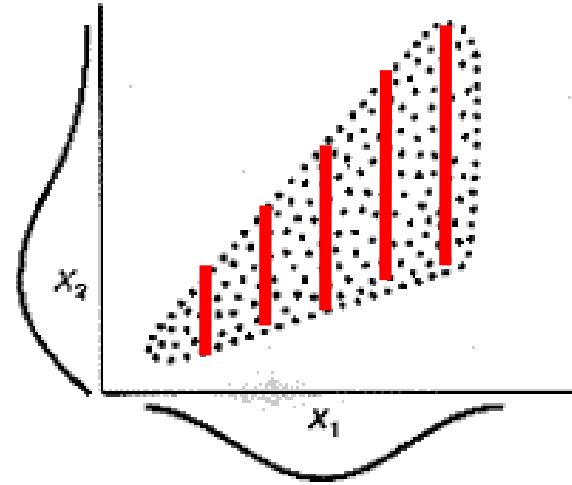
การกระจายของตัวแปร
 X_2 เบ้ขวา

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- อาการเบื้องต้นที่สังเกตได้ หากข้อตกลงเบื้องต้นนี้ถูกละเมิด
 - ความแปรปรวนของตัวแปรหนึ่ง ในทุกค่าของอีกตัวแปรหนึ่งไม่เท่ากัน เรียกว่า ความไม่เท่าเทียมกันของการกระจาย (Heteroscedasticity)



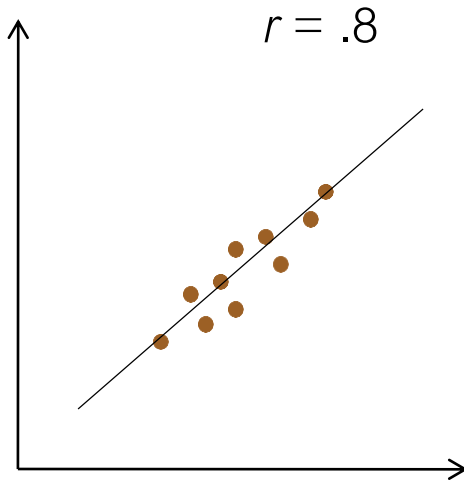
ความแปรปรวนของ X_2 ตลอด
ค่าของ X_1 เท่ากัน



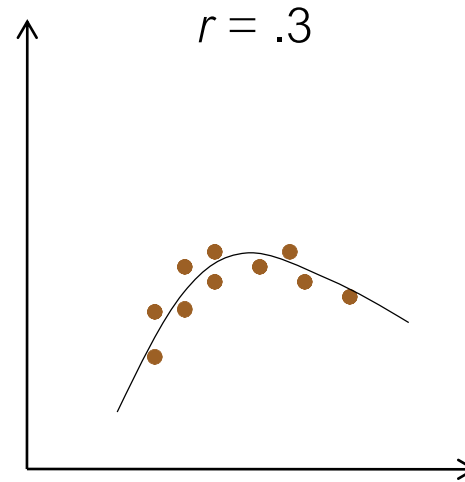
ความแปรปรวนของ X_2 เพิ่มขึ้น
หากค่าของ X_1 เพิ่มขึ้น

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- อาการเบื้องต้นที่สังเกตได้ หากข้อตกลงเบื้องต้นนี้ถูกละเมิด
 - ความสัมพันธ์ไม่ใช่เชิงเส้น (Nonlinear relationship)



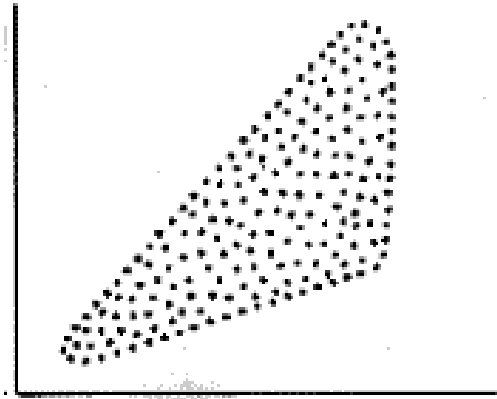
ความสัมพันธ์เชิงเส้นตรง



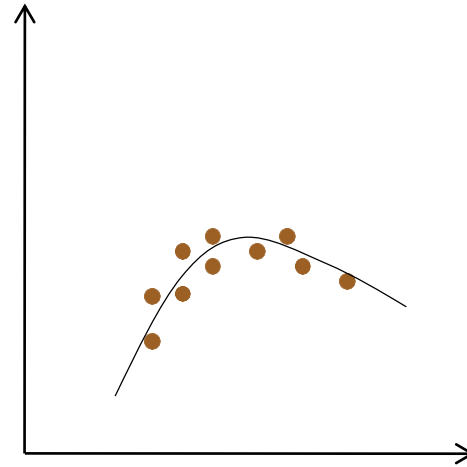
ความสัมพันธ์ที่ไม่เป็นเส้นตรง

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- ค่าสหสัมพันธ์ไม่เหมาะสมที่จะใช้ในการแสดงความสัมพันธ์หากข้อตกลงเบื้องต้นนี้ถูกละเมิด



ความสัมพันธ์สูงสำหรับค่าน้อย
แต่ความสัมพันธ์ต่ำสำหรับค่าสูง
ตัวเลขสหสัมพันธ์เพียงค่าเดียว
ไม่ได้บอกภาพของความสัมพันธ์
ตัวแปรทั้งสองได้ถูกต้อง



ความสัมพันธ์เชิงเส้นตรงไม่เหมาะสม
ในการวัดความสัมพันธ์เชิงเส้นโค้ง

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- วิธีทางแก้ไขที่ดีที่สุด คือ หาวิธีการอธิบายความสัมพันธ์ระหว่างตัวแปรให้ลึกซึ้งกว่าเดิม
 - หากเกิดความสัมพันธ์ไม่ใช่เชิงเส้น ให้หารูปแบบความสัมพันธ์ระหว่างสองตัวแปรในเชิงไม่เป็นเส้นตรง โดยใช้การวิเคราะห์ถดถอย (Regression analysis)
 - หากความแปรปรวนไม่เท่ากัน เป็นไปได้ที่ผู้วิจัยอาจลืมนำสิ่งถึงตัวแปรอื่นที่อาจมีอิทธิพลร่วม
 - เช่น ความสัมพันธ์ระหว่างความสามารถและผลการปฏิบัติงานที่มีการกระจายของผลการปฏิบัติงานในคนความสามารถมาก มากกว่าคนสามารถน้อย
 - อาจละเลยผลของแรงจูงใจ

ข้อตกลงเบื้องต้นก่อนการใช้สถิติ

- วิธีทางแก้ไขที่ดีรองลงมา คือ หาความสัมพันธ์ระหว่างตัวแปรโดยใช้การจัดอันดับ (Ranking) ซึ่งจะกล่าวถึงในภายหลัง
- วิธีทางที่แย่ที่สุด และนิยมมากที่สุด คือ ยังคงใช้ค่าสหสัมพันธ์แบบปกติต่อไป โดยไม่ตรวจสอบข้อตกลงเบื้องต้น

การเขียนรายงาน

สถิติขั้นนำสำหรับจิตวิทยา

อ. สันทัต พรประเสริฐมานิต

การเขียนรายงาน

- เช่น จากการเก็บข้อมูลจากพนักงานบริษัทหนึ่ง จำนวน 221 คน พบว่าความขัดแย้งระหว่างครอบครัวและที่ทำงานมีความสัมพันธ์ในทางลบอย่างมีนัยสำคัญทางสถิติ ในระดับน้อย ($r = -.21, p < .05$)
- จากการสุ่มนิสิตระดับปริญญาตรีจำนวน 10 คน ตอบแบบวัดเจตคติต่อสาวประเภทสอง และหลังจากนั้นให้นิสิตเหล่านั้นพูดคุยกับหน้าม้าที่เป็นสาวประเภทสอง โดยสังเกตการสบตาหน้าม้า พบว่าเจตคติต่อสาวประเภทสองและจำนวนครั้งการสบตาไม่มีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติ ($r = .52, p > .05$)

การเขียนรายงาน

- บางครั้งอาจรายงานค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และสหสัมพันธ์ของตัวแปรหลายตัวพร้อมกันในตารางเดียวกัน

	A	B	C
A			
B	.50**		
C	-.30*	-.45**	
M	100	50	5
SD	15	10	2

แนวทแยงไม่จำเป็นต้องรายงาน
เพราะเท่ากับ 1

ตัวเลขเหนือแนวทแยงก็ไม่จำเป็นต้อง
รายงาน เพราะเท่ากับตัวเลขใต้แนวทแยง

หมายเหตุ. * $p < .05$, ** $p < .01$

การหาช่วงเชื่อมั่นของค่าสหสัมพันธ์

สถิติขั้นนำสำหรับจิตวิทยา

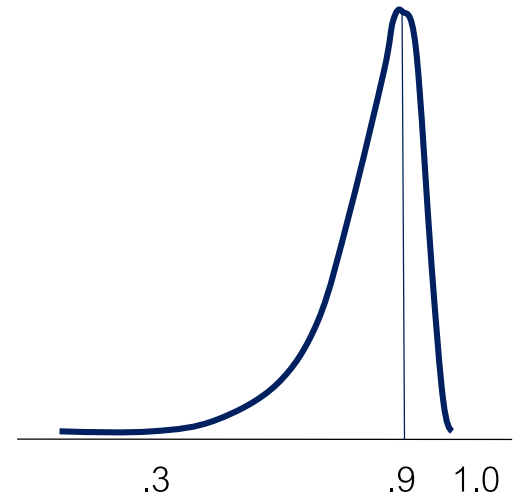
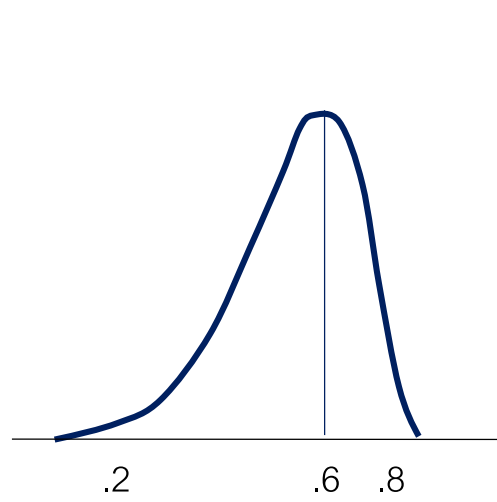
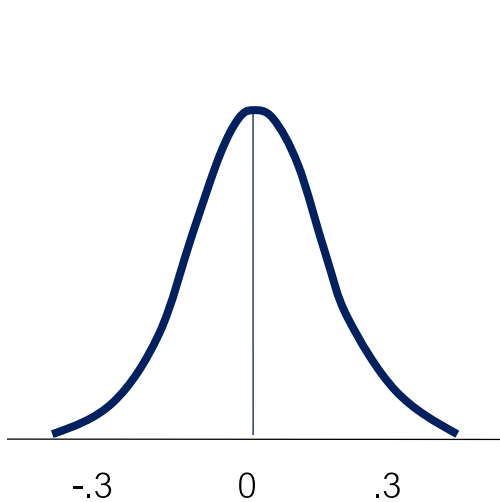
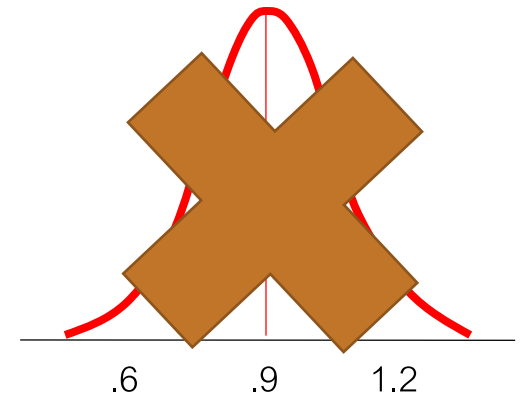
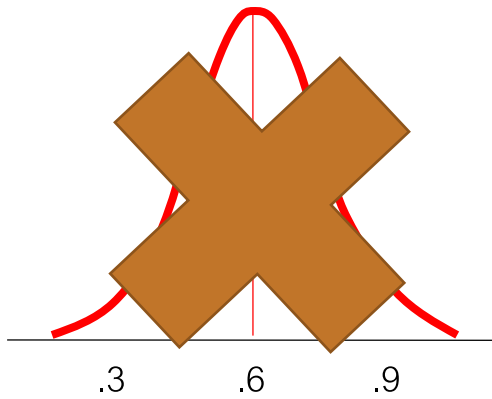
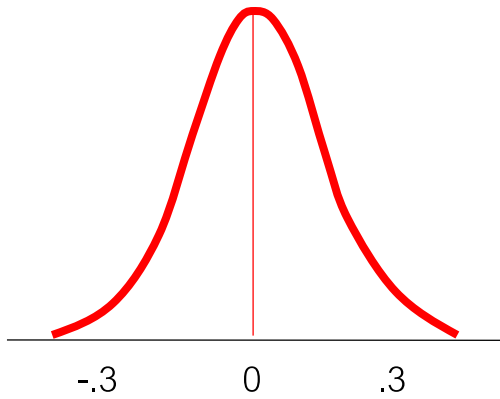
อ. สันหัต พรประเสริฐมานิต

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

- ในการหาช่วงเชื่อมั่นของค่าสหสัมพันธ์นั้น จะมีลักษณะพิเศษ
- เมื่อสุ่มกลุ่มตัวอย่างจากประชากรของตัวแปรที่สหสัมพันธ์เท่ากับ 0 จะทำให้การกระจายของการสุ่ม (Sampling Distribution) เป็นรูปสมมาตร
- หากสุ่มจากประชากรของตัวแปรที่สหสัมพันธ์ไม่เท่ากับ 0 จะทำให้การกระจายของการสุ่มไม่เป็นรูปสมมาตร

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

- สังเกต



การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

- เมื่อรูปแบบของการกระจายไม่เหมือนกันตลอดช่วงของสหสัมพันธ์แล้ว ทำให้หาช่วงเชื่อมั่นแบบปกติไม่ได้
- Fisher's z transformation เป็นการแปลงค่า r ที่ทำให้การกระจายของค่าที่แปลงแล้วมีการกระจายเป็นโค้งปกติ ตลอดช่วงของค่าทั้งหมด
- ตัวอย่าง เช่น เก็บข้อมูลจาก 28 คน พบว่ามีค่าสหสัมพันธ์ระหว่างความสุขและความภูมิใจในตนเองเท่ากับ .5

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

การหาช่วงเชื่อมั่น สามารถทำได้โดย

1. แปลงค่า r ให้เป็นค่า z'

$$z' = \tanh^{-1}(r) = 0.5 \log\left(\frac{1+r}{1-r}\right)$$

MS EXCEL
= FISHER(r)

เช่น $r = .5, n = 28$

$$z' = 0.5 \log\left(\frac{1+0.5}{1-0.5}\right) = 0.549$$

MS EXCEL
= FISHER(0.5) = 0.549

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

การหาช่วงเชื่อมั่น สามารถทำได้โดย

2. หาช่วงเชื่อมั่นของค่า z'

$$CI_{1-\alpha} \text{ for } z' = z' \pm \frac{z_{\alpha/2}}{\sqrt{n-3}}$$

เช่น $r = 0.5$, $n = 28$, $z' = 0.549$

$$CI_{.95} \text{ for } z' = 0.549 \pm \frac{1.96}{\sqrt{28-3}} = 0.157 \text{ to } 0.941$$

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

การหาช่วงเชื่อมั่น สามารถทำได้โดย

3. แปลงช่วงของค่า z' ให้เป็นช่วงของค่า r

$$r = \tanh(z') = \frac{e^{2z'} - 1}{e^{2z'} + 1} \quad \begin{array}{l} \text{MS EXCEL} \\ = \text{FISHERINV}(z') \end{array}$$

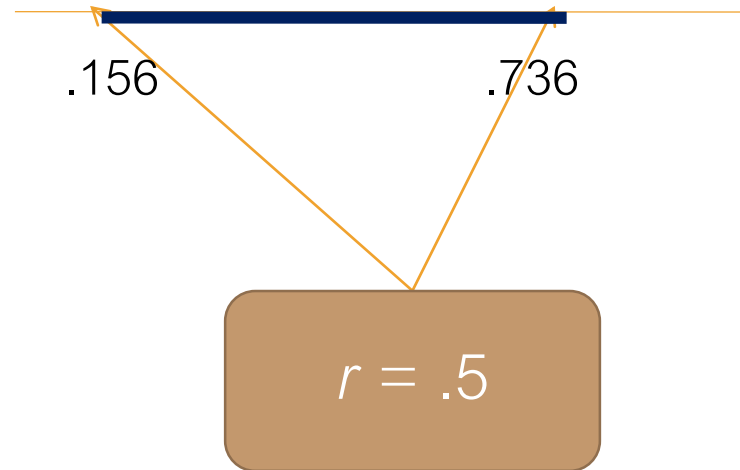
เช่น $r = 0.5$, $n = 28$, $z' = 0.549$, CI for $z' = (0.157, 0.941)$

$$r = \tanh(z') = \frac{e^{2(0.157)} - 1}{e^{2(0.157)} + 1} = 0.156 \quad \begin{array}{l} \text{MS EXCEL} \\ = \text{FISHERINV}(0.157) \end{array}$$

$$r = \tanh(z') = \frac{e^{2(0.941)} - 1}{e^{2(0.941)} + 1} = 0.736 \quad \begin{array}{l} \text{MS EXCEL} \\ = \text{FISHERINV}(0.941) \end{array}$$

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

- ช่วงเชื่อมั่น



จากข้อมูลทำให้มีความเชื่อมั่นระดับ .95 ว่า
ประชากรของค่าสหสัมพันธ์ระหว่างความสุขและ
ความภาคภูมิใจในตนเองเท่ากับ .16 ถึง .74

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

- การหาสามารถใช้โปรแกรม Excel ช่วย โดย Fisher z CI of r.xlsx
- เช่น นักจิตวิทยาอุตสาหกรรมและองค์การทดสอบความสัมพันธ์ระหว่างความพึงพอใจในงานและผลการปฏิบัติงานจากพนักงาน 250 คน ได้ค่าสหสัมพันธ์เท่ากับ .24
- จงหาช่วงเชื่อมั่นระดับ .95 ว่าช่วงค่าสหสัมพันธ์ในประชากร ของตัวแปรทั้งสองเท่ากับเท่าไร

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

- $r = .24$, $n = 250$, ช่วงเชื่อมั่นระดับ .95



จากข้อมูลทำให้มีความเชื่อมั่นระดับ .95 ว่าประชากรของค่าสหสัมพันธ์ระหว่างความพึงพอใจในงานและผลการทำงานเท่ากับ .12 ถึง .35

	A	B	C	D
1	Confidence Interval of Pearson's Correlation			
2				
3	r	0.24		ใส่ข้อมูล
4	n	250		
5	CI	0.95		
10	CI of r	0.119	0.354	แสดงผล

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

- ช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์นี้ สามารถใช้ในการทดสอบสมมติฐานว่า ค่าสหสัมพันธ์มากกว่าค่าใดค่าหนึ่งหรือไม่ (ไม่จำเป็นต้องกำหนดว่าเท่ากับ 0)
- เช่น ความสัมพันธ์ระหว่างเงินเดือนและผลการทำงานสูงกว่า 0.3 หรือไม่

สมมติฐาน คือ ความสัมพันธ์ระหว่าง
เงินเดือนและผลงานอยู่ในทางบวกระดับสูง

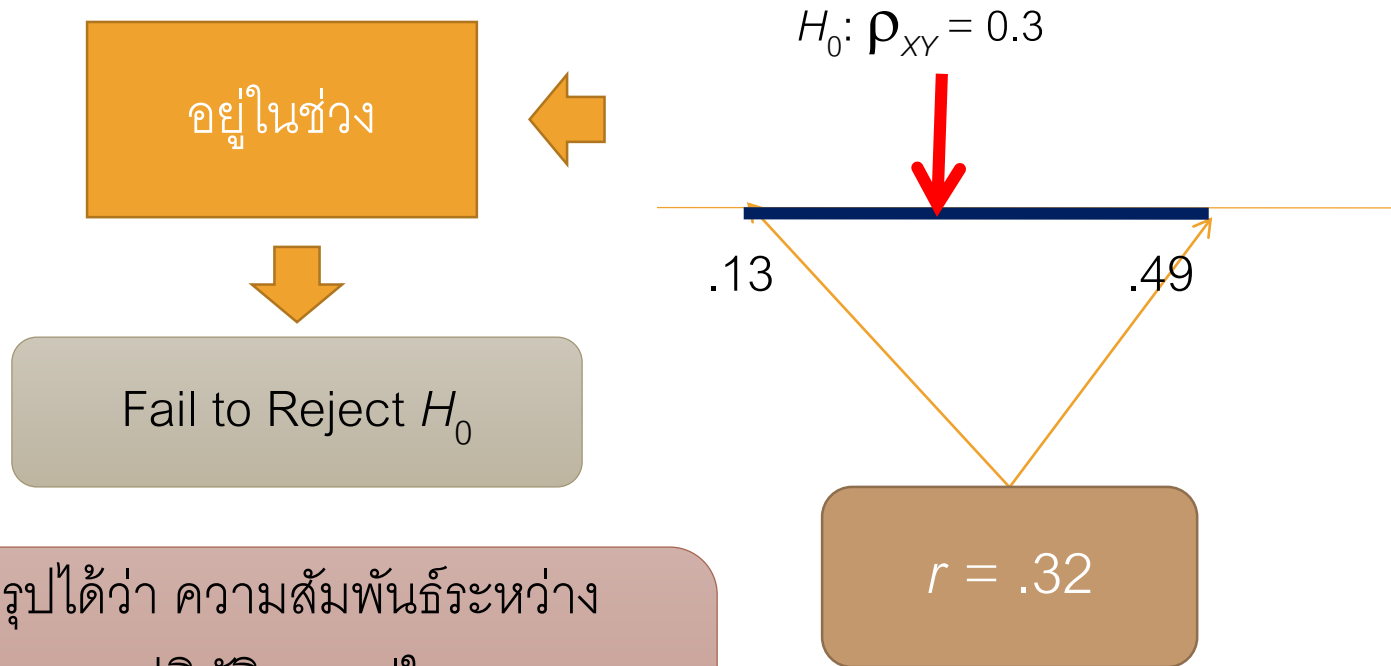
ใช้การทดสอบสองทาง

$$H_0: \rho_{XY} = 0.3$$

$$H_1: \rho_{XY} \neq 0.3$$

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

- เช่น พบ $r = .32$, $n = 100$, ช่วงเชื่อมั่นระดับ .95



ไม่สามารถสรุปได้ว่า ความสัมพันธ์ระหว่าง
เงินเดือนและผลการปฏิบัติงานอยู่ในทางบวก
ระดับสูงหรือไม่

การหาช่วงเชื่อมั่นสำหรับค่าสหสัมพันธ์

- ข้อตกลงเบื้องต้นก่อนการใช้ Fisher's z transformation
 - การกระจายเป็นแบบสุ่ม (Random sampling)
 - ค่าสหสัมพันธ์ที่ใช้ในการหาช่วงเชื่อมั่น ไม่ควรใกล้ -1 หรือ 1 มากๆ (เช่น $r = .97$)
 - ประชากรของตัวแปรทั้งสองมีการกระจายเป็นการกระจายโค้งปกติร่วม
 - การกระจายของตัวแปรแต่ละตัว (Marginal distribution) เป็นโค้งปกติ
 - ความสัมพันธ์เป็นเชิงเส้นตรง (Linear Relationship)
 - ความเท่าเทียมกันของการกระจาย (Homoscedasticity)
 - จำนวนกลุ่มตัวอย่าง มากกว่า 10

ขนาดอิทธิพล

สถิติขั้นนำสำหรับจิตวิทยา

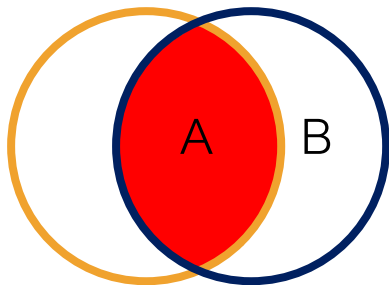
อ. สันทัด พรประเสริฐมานิต

ขนาดอิทธิพล

- ค่าสหสัมพันธ์ (Correlation) เป็นค่าที่แสดงให้เห็นถึงขนาดอิทธิพลโดยตรง
- ดังที่กล่าวไปข้างต้น
 - เครื่องหมายบวกลบ บอกทิศทางความสัมพันธ์
 - ค่าสัมบูรณ์ของค่าสหสัมพันธ์บอกขนาดความสัมพันธ์
 - .20 → น้อย .50 → ปานกลาง .80 → มาก

ขนาดอิทธิพล

- ค่าสหสัมพันธ์ (Correlation) ยังสามารถแปลความหมายได้อีกรูปแบบหนึ่ง คือ นำค่าสหสัมพันธ์มายกกำลังสอง (r^2)
- ค่านี้จะเรียกว่า สัมประสิทธิ์การทำนาย (Coefficient of Determination)
- หมายความว่า ตัวแปรหนึ่ง สามารถอธิบายความแปรปรวน (การเปลี่ยนแปลง) ของอีกตัวแปรหนึ่งได้เท่าไร



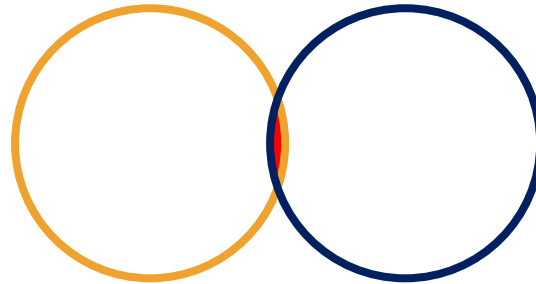
การกระจายรูปพื้นที่ของตัวแปรตาม = $A + B$

สัมประสิทธิ์การทำนาย = $A/(A + B)$

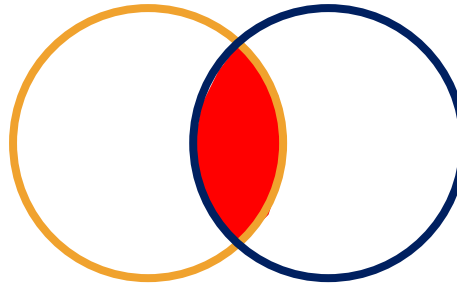
ขนาดอิทธิพล

- สัมประสิทธิ์การทำนาย (Coefficient of Determination)

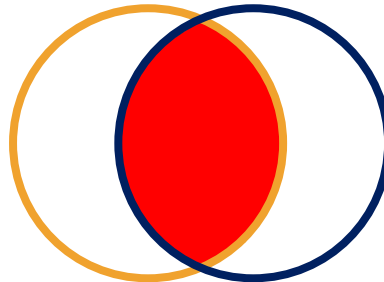
- $r = .2, r^2 = .04$



- $r = .5, r^2 = .25$



- $r = .8, r^2 = .64$



การหาค่าลึงขของค่าสหสัมพันธ์

สถิติขั้นนำสำหรับจิตวิทยา

อ. สันทัด พรประเสริฐมานิต

การหาคำสั่ง

- ใช้โปรแกรม G*POWER 3 โดย
 - เลือกกลุ่มสถิติที่ต้องการทดสอบ คือ Exact
 - เลือกสถิติที่ใช้ทดสอบ คือ Correlations: Difference from constant (One sample case)
 - เลือกว่าจะหาคำสั่งในการทดสอบ คือ Post hoc: Compute achieved power – given α , sample size and effect size

การหาคำสั่ง

- เช่น

ทดสอบทิศทาง
ค่าสหสัมพันธ์
ระดับนัยสำคัญ
ขนาดกลุ่มตัวอย่าง
ค่าสหสัมพันธ์ของ H_0

Test family: Exact
Statistical test: Correlations: Difference from constant (one sample case)
Type of power analysis: Post hoc: Compute achieved power - given α , sample size, and effect size

Input Parameters

Tail(s)	Two
Effect size r	0.285
α err prob	0.05
Total sample size	120
Population correlation ρ	0

Output Parameters

Lower critical p	-0.179343
Upper critical p	0.179343
Power ($1 - \beta$ err prob)	0.889534

กำลังในการทดสอบสมมติฐานทางสถิติ

การหาค่า

- โดยปกติ ถ้าทดสอบว่าตัวแปรทั้งสองมีความสัมพันธ์กันจริงหรือไม่ จะให้ค่า Population correlation = 0
- แต่หากทดสอบว่า ค่าสหสัมพันธ์แตกต่างกับค่าที่กำหนด (เช่น 0.2) หรือไม่ (ดังเช่นที่กล่าวในเรื่องช่วงเชื่อมั่น) ให้กำหนดค่าสหสัมพันธ์ตามที่ตั้งใน Null Hypothesis

การกำหนดขนาดกลุ่มตัวอย่าง

สถิติขั้นนำสำหรับจิตวิทยา

อ. สันหัต พรประเสริฐมานิต

การกำหนดขนาดกลุ่มตัวอย่าง

- การกำหนดขนาดกลุ่มตัวอย่าง สามารถทำได้ 2 ประเภท คือ
 - การกำหนดขนาดกลุ่มตัวอย่างจากกำลังที่ต้องการ
 - การกำหนดขนาดกลุ่มตัวอย่างจากความผิดพลาดในการประมาณค่าพารามิเตอร์ (ρ) กล่าวคือ ดูความกว้างของช่วงเชื่อมั่นของ ρ
- ในที่นี้จะเน้นเฉพาะการหาจำนวนกลุ่มตัวอย่างจากกำลังเท่านั้น

การกำหนดขนาดกลุ่มตัวอย่าง

- การกำหนดขนาดกลุ่มตัวอย่างตามกำลังที่ต้องการ สามารถทำได้โดยใช้โปรแกรม G*POWER 3 ช่วย
 - เลือกกลุ่มสถิติที่ต้องการทดสอบ คือ Exact
 - เลือกสถิติที่ใช้ทดสอบ คือ Correlations: Difference from constant (One sample case)
 - เลือกว่าจะหาลำดับในการทดสอบ คือ A Priori: Compute required sample size – given α , power and effect size

การกำหนดขนาดกลุ่มตัวอย่าง

- เช่น จงหาขนาดกลุ่มตัวอย่างที่จะทำให้กำลัง = .80 ในการทดสอบสองทาง และระดับนัยสำคัญเท่ากับ .05 เมื่อค่าสหสัมพันธ์เท่ากับ 0.285 เปรียบเทียบว่าตัวแปรทั้งสองมีความสัมพันธ์กันหรือไม่

Test family	Statistical test		
Exact	Correlations: Difference from constant (one sample case)		
Type of power analysis			
A priori: Compute required sample size - given α , power, and effect size			
Input Parameters	Output Parameters		
Determine =>	Lower critical p		
Tail(s)	Two	-0.202763	
Effect size r	0.285	Upper critical p	0.202763
α err prob	0.05	Total sample size	94
Power (1- β err prob)	0.80	Actual power	0.802759
Population correlation ρ	0		

ความแตกต่างและความสัมพันธ์

สถิติขั้นนำสำหรับจิตวิทยา

อ. สันหัต พรประเสริฐมานิต

ความแตกต่างและความสัมพันธ์

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

ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

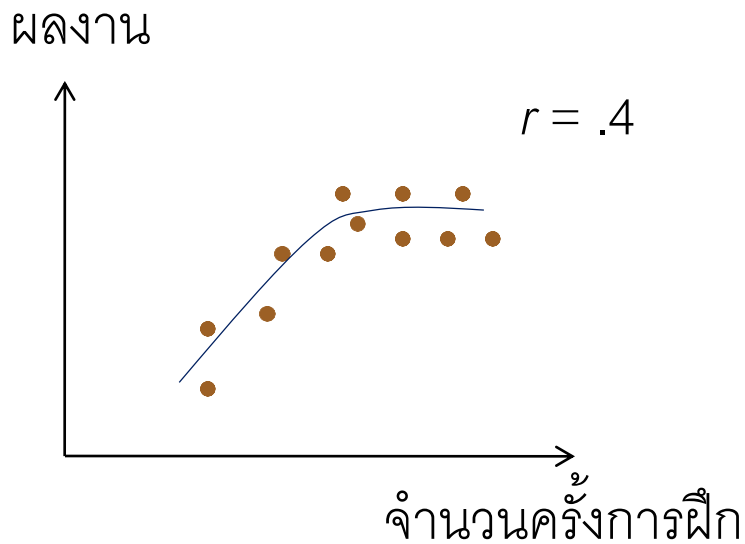
สถิติขั้นนำสำหรับจิตวิทยา

อ. สันหัตต์ พรประเสริฐมานิต

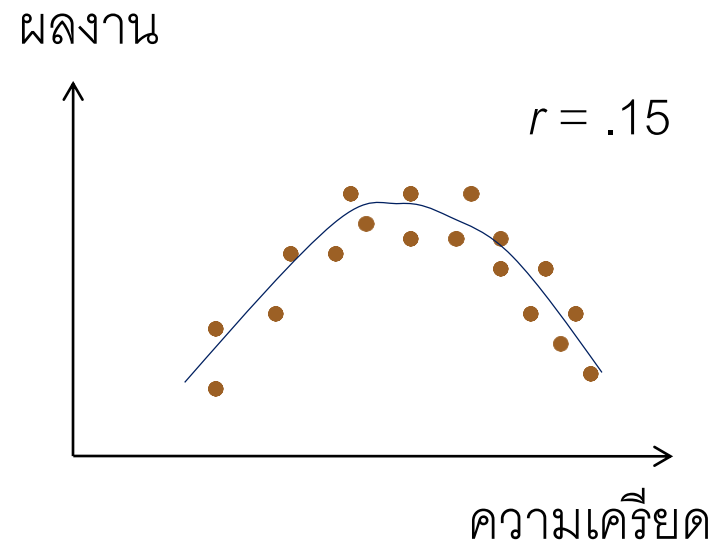
ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- ความสัมพันธ์ที่ไม่เป็นเส้นตรง

- จำนวนครั้งในการฝึก และผลงาน



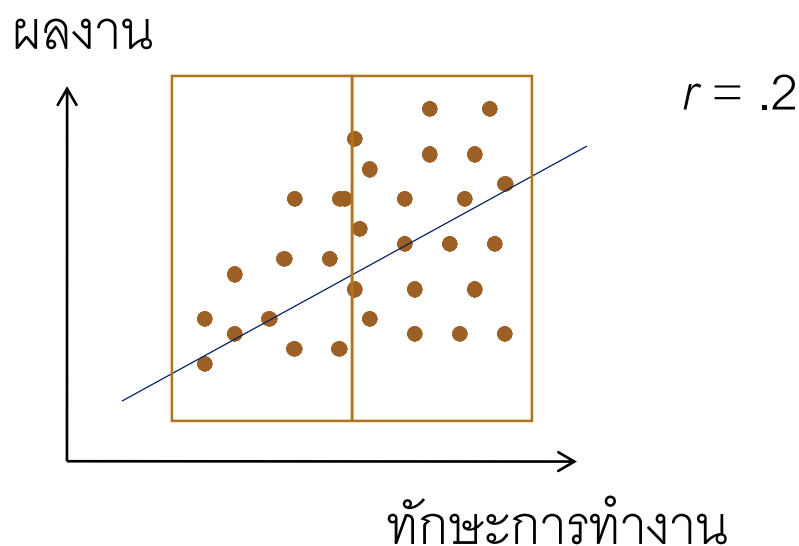
- ความเครียดและผลงาน



ความสัมพันธ์น่าจะสูงกว่านี้ เนื่องจากตัวแปรต้น สามารถทำนายค่าตัวแปรตามได้ค่อนข้างแม่นยำ (โดยไม่ได้ใช้เส้นตรง)

ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- การกระจายไม่เท่าเทียมกัน
 - ทักษะการทำงานและผลงาน



จากรูป ค่อนข้างมั่นใจว่าคนที่มีความ
ทักษะการทำงานต่ำ จะมีผลงานไม่ดี

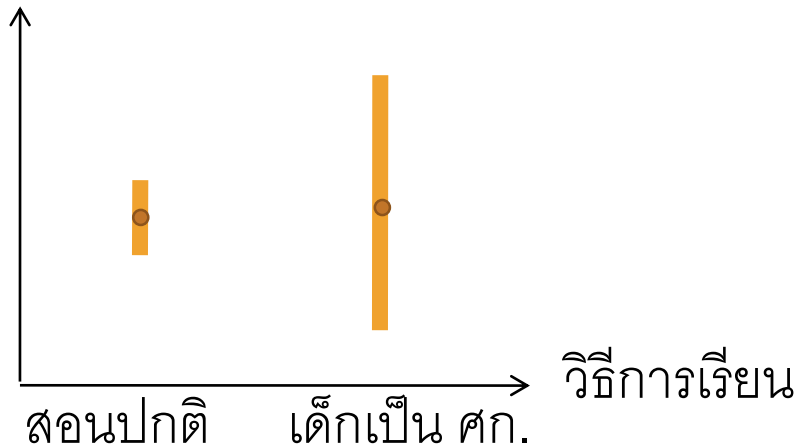
แต่คนที่มีความทักษะการทำงานสูง
อาจมีผลงานดีหรือไม่ดีก็ได้

แสดงว่าตัวแปรต้น สามารถทำนายตัวแปรตามบางช่วงได้ดี และทำนาย
ช่วงอื่นได้ไม่ค่อยดี ค่า r เพียงค่าเดียว ไม่ได้แสดงรายละเอียดตรงนี้

ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- การกระจายไม่เท่าเทียมกัน (การเปรียบเทียบกลุ่มสองกลุ่ม)
 - การสอนแบบเด็กเป็นศูนย์กลางและแบบปกติ

ผลงาน



จากรูป การสอนปกติเด็กจะมี
ผลการเรียนระดับหนึ่ง

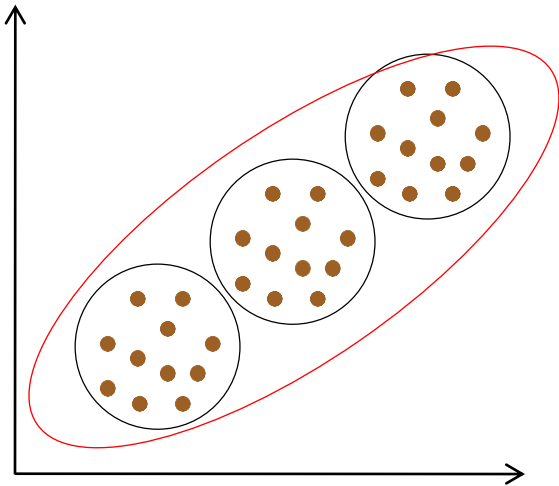
การสอนแบบเด็กเป็นศูนย์กลาง
อาจได้ผลตั้งแต่แย่งไปจนดี

แสดงว่าตัวแปรต้นบางค่าสามารถทำนายตัวแปรตามบางช่วงได้ดี และบางค่าทำนายช่วงอื่นได้ไม่ค่อยดี การเปรียบเทียบความแตกต่างอย่างเดียวกันไม่ได้แสดงรายละเอียดตรงนี้

ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- ความสัมพันธ์หลอก (Spurious Correlation)
 - จำนวนระดับเพลิง และความเสียหายของอัคคีภัย

ความเสียหาย



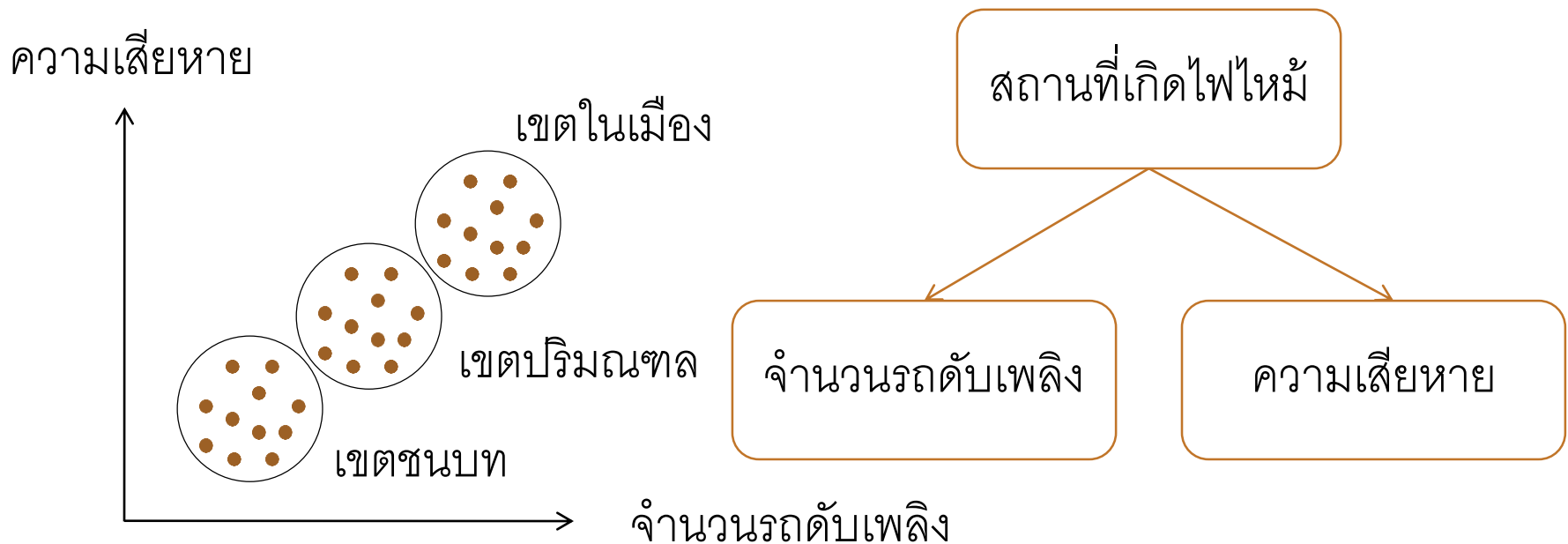
ถ้าดูในภาพรวม พบว่าความสัมพันธ์สูง

แต่เมื่อดูกลุ่มย่อยแล้ว
พบว่าตัวแปรทั้งสองไม่สัมพันธ์กัน

จำนวนระดับเพลิง

ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

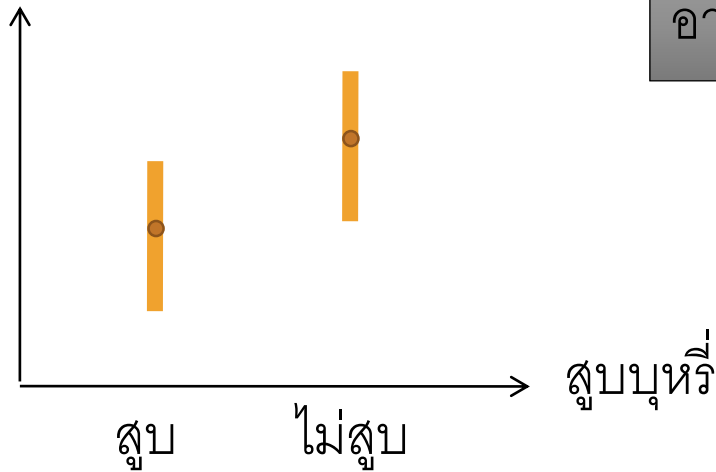
- ความสัมพันธ์หลอก (Spurious Correlation)
 - จำนวนรถดับเพลิง และความเสียหายของอัครคิภัย



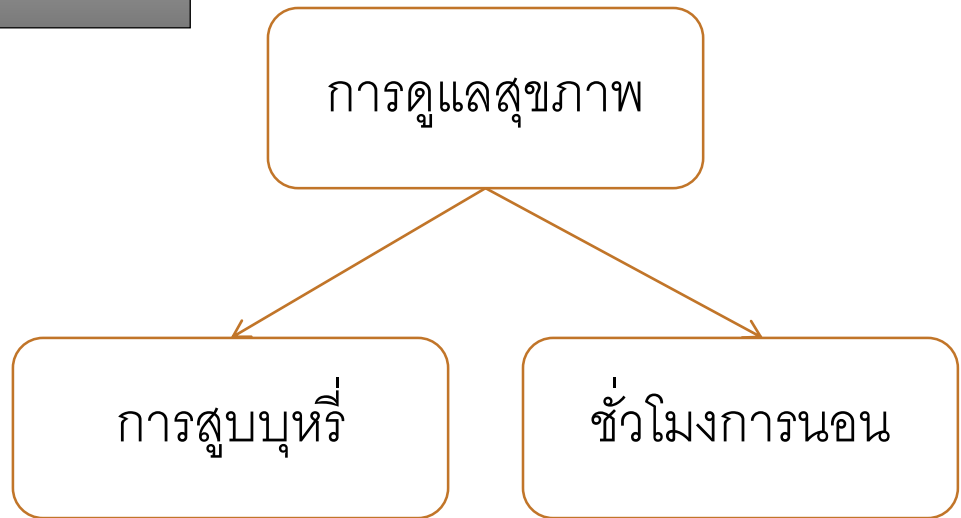
ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- ความสัมพันธ์หลอก (Spurious Correlation)
 - คนสูบบุหรี่ จะนอนน้อยกว่าคนที่ไม่สูบบุหรี่

จำนวนชั่วโมงนอน

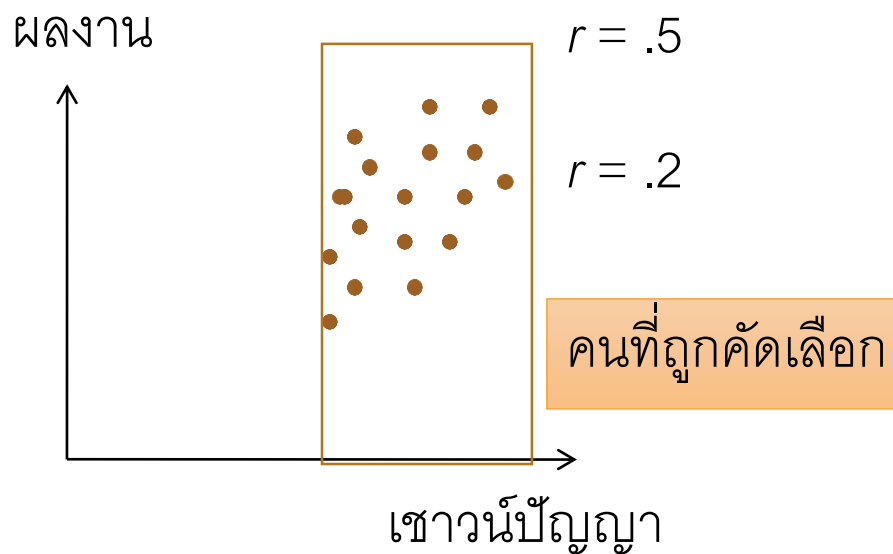


บุหรี่เป็นสิ่งที่ทำให้การนอนน้อยลงหรือไม่



ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- การจำกัดพิสัย (Range Restriction)
 - คะแนน IQ และผลการปฏิบัติงาน



จากรูป ปกติเชาวน์ปัญญาและผลงาน
มีสหสัมพันธ์ในระดับค่อนข้างสูง

เมื่อคัดเลือกเฉพาะคนที่มี
เชาวน์ปัญญาสูงเข้าทำงาน แล้ว
หาค่าสหสัมพันธ์ จะทำให้ค่า
สหสัมพันธ์ต่ำ

ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- ค่าสุดโต่ง (Outlier)
 - หาความสัมพันธ์ระหว่างรายได้และรายจ่าย

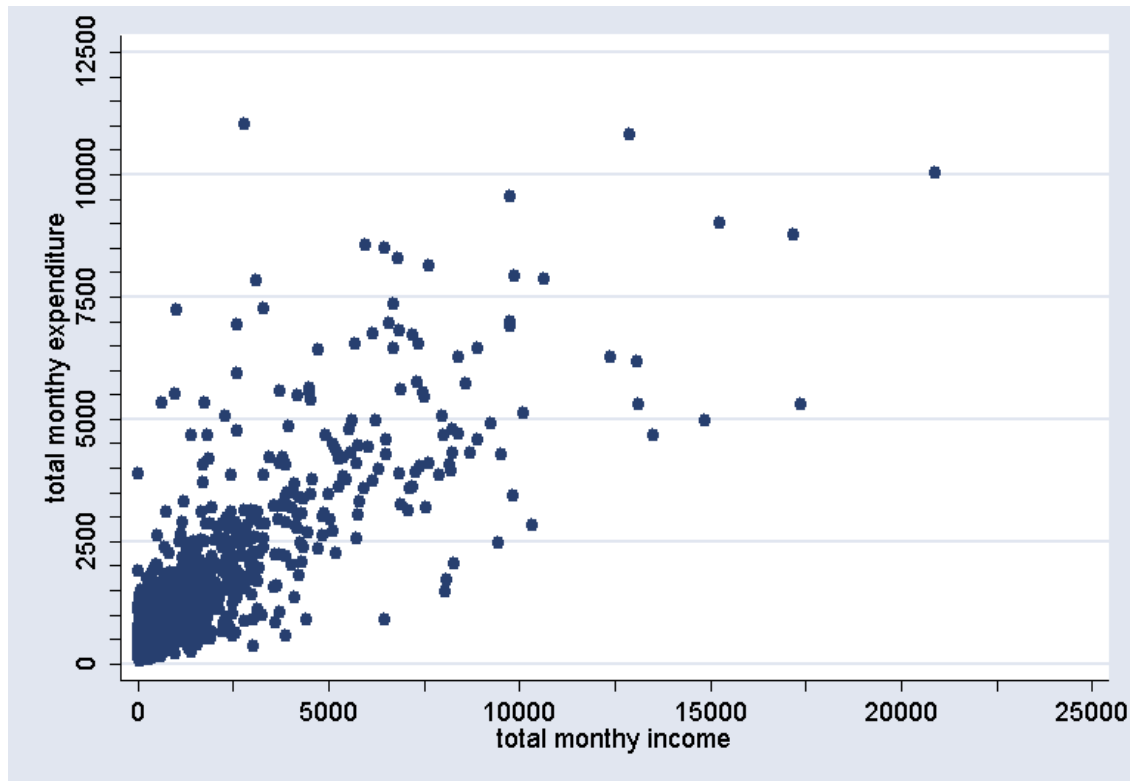


$$r = .52$$

ค่าสุดโต่ง คือ ค่าที่แตกต่างจากข้อมูลอื่น

ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- ค่าสุดโต่ง (Outlier)
 - หาความสัมพันธ์ระหว่างรายได้และรายจ่าย



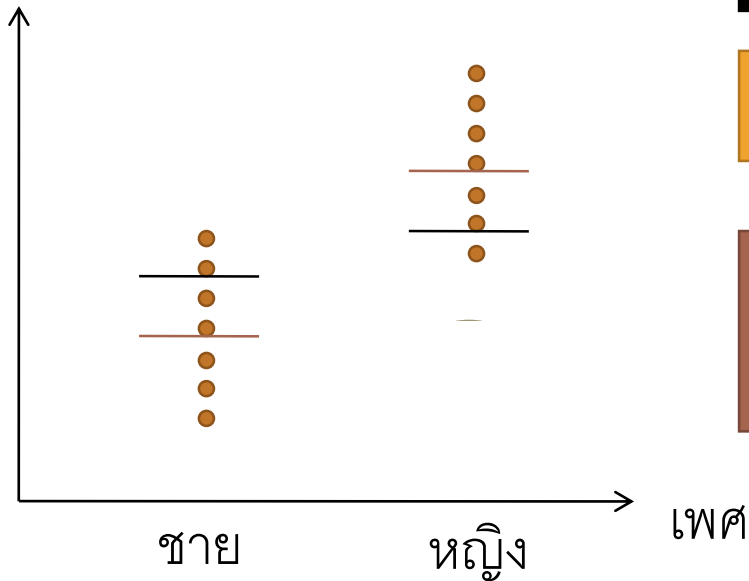
$$r = .79$$

ในกรณีนี้ ค่าสุดโต่ง
ทำให้ค่าสหสัมพันธ์
น้อยลง

ปัจจัยเบี่ยงเบนค่าสหสัมพันธ์

- ค่าสุดโต่ง (Outlier)
 - เปรียบเทียบเวลาในการอาบน้ำระหว่างเพศ

เวลาในการอาบน้ำ



ความแตกต่างดังกล่าวไม่ถึงระดับนัยสำคัญ

เนื่องจากมีค่าสุดโต่ง

หากตัดข้อมูลดังกล่าวแล้ว อาจทำให้ความแตกต่างถึงระดับนัยสำคัญ

ความสัมพันธ์และการอนุมานเชิง

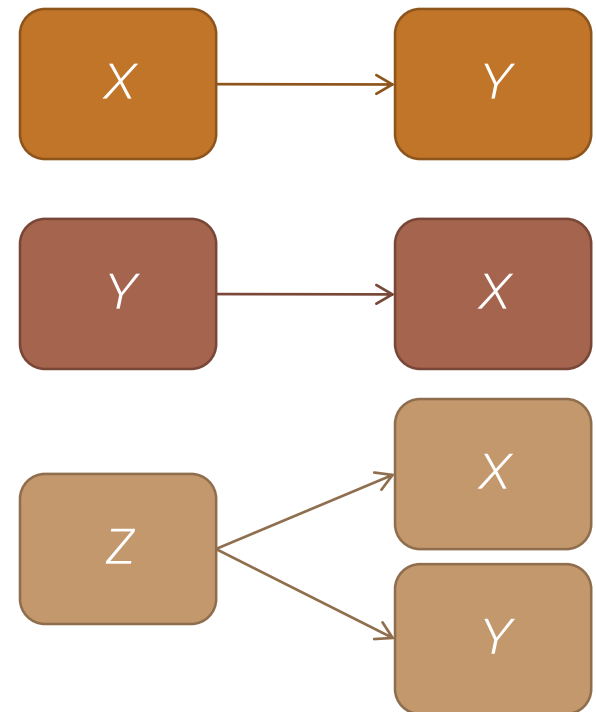
สาเหตุ

สถิติขั้นนำสำหรับจิตวิทยา

อ. สันหัตต์ พรประเสริฐสุมานิต

ความสัมพันธ์เชิงสาเหตุ

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



ความสัมพันธ์เชิงสาเหตุ

- เช่น การดูรายการโทรทัศน์ก้าวร้าวสัมพันธ์กับความก้าวร้าวในเด็ก
- สามารถอธิบายได้สามแบบ
 1. การดูรายการโทรทัศน์ก้าวร้าวทำให้เด็กเลียนแบบความก้าวร้าว
 2. เด็กที่ก้าวร้าว จะเลือกดูรายการโทรทัศน์ที่ก้าวร้าว
 3. ผู้ใหญ่ที่ปล่อยปละละเลย ทำให้เด็กแสดงความก้าวร้าว และเลือกดูรายการโทรทัศน์ที่ก้าวร้าวด้วย

ความสัมพันธ์เชิงสาเหตุ

- ด้วยเหตุนี้ การที่จะบอกว่า X เป็นสาเหตุของ Y จะต้องมีคุณสมบัติ 4 ประการด้วยกัน
 - ตัวแปร X มีความสัมพันธ์กับตัวแปร Y
 - X ต้องมาก่อน Y
 - การเกิด Y จะต้องไม่ได้รับการอธิบายด้วยตัวแปรอื่น
 - มีทฤษฎีหรือกระบวนการอธิบายว่าทำไม X ถึงก่อให้เกิด Y

ความสัมพันธ์เชิงสาเหตุ

- เช่น การอธิบายว่าการให้โบนัส จะทำให้พนักงานทำงานมากขึ้น
 - จากข้อมูล ต้องพบว่า พนักงานที่ได้รับโบนัสจะต้องทำงานมากกว่าพนักงานที่ไม่ได้รับโบนัส
 - บริษัทจะต้องจ่ายโบนัสก่อน แล้วจึงระยะเวลาก่อนวัดลักษณะการทำงาน
 - บริษัทต้องมีการควบคุมตัวแปรอื่นให้กลุ่มของพนักงานที่ได้รับ และไม่ได้โบนัสได้เท่ากัน เช่น สวัสดิการ เงินเดือนปกติ การได้รับการปฏิบัติจากเจ้านาย เป็นต้น
 - มีทฤษฎีอธิบายว่า ทำไมการให้โบนัสจึงทำให้พนักงานทำงานมากขึ้น เช่น การเสริมแรงทางบวก (Positive Reinforcement)

สหสัมพันธ์ของค่าจัดอันดับ

สถิติขั้นนำสำหรับจิตวิทยา

อ. สันหัต พรประเสริฐมานิต

สหสัมพันธ์ของค่าจัดอันดับ

ทัศนคติต่อสาวประเภทสอง

3 ???
5
7
9
0 1
2 2
4
3 ???
8
7

จำนวนครั้งที่สบตา

2
7
9
6
3
5
4
6
4
7

ความสัมพันธ์ระหว่างคะแนนดิบ

$$r = .519$$

นำตัวแปรทั้งสองไปจัดอันดับ
(Ranking)

จะใส่อันดับอะไรดี สำหรับค่าที่ซ้ำกัน

สหสัมพันธ์ของค่าจัดอันดับ

ทัศนคติต่อสาวประเภทสอง

3
5
7
9
0
2
4
3
8
7

จัดอันดับ

3
6
7
10
1
2
5
4
9
8

3.5
7.5

2. นำอันดับของ
ค่าที่เหมือนกัน
มาหาค่าเฉลี่ย

อันดับใหม่

3.5
6
7.5
10
1
2
5
3.5
9
7.5

3. แทนค่าอันดับ
ด้วยค่าเฉลี่ย

1. ค่าที่เหมือนกัน ก็จัดอันดับไป โดยเลือกตัวหนึ่งเป็นตัวน้อยกว่า อีกตัวเป็นตัวมากกว่า

สหสัมพันธ์ของค่าจัดอันดับ

อันดับของทัศนคติต่อสาวประเภทสอง

อันดับจำนวนครั้งที่สบตา

3.5

1

6

8.5

7.5

10

10

6.5

1

2

2

5

5

3.5

3.5

6.5

9

3.5

7.5

8.5

ความสัมพันธ์ระหว่างอันดับ

$$r = .486$$

เรียกว่า Spearman Rank
Correlation

เลือก **Spearman** ใน SPSS
ในหน้าต่าง Pop-up

นำอันดับมาหา Pearson Correlation

สหสัมพันธ์ของค่าจัดอันดับ

- วิธีนี้สามารถใช้แทน Pearson Correlation ได้ เมื่อข้อตกลงเบื้องต้นที่การกระจายต้องเป็นโค้งปกติร่วมถูกละเมิด
- ความสัมพันธ์รูปแบบนี้ยังสามารถใช้ได้กับข้อมูลแบบจัดอันดับ (Ordinal scale) เช่น ความสัมพันธ์ระหว่างอันดับที่ในชั้นม. 6 กับคะแนน O-NET
- สามารถทดสอบสมมติฐานได้ว่า Spearman Correlation ในประชากรเท่ากับ 0 หรือไม่ (ใช้ SPSS) แต่ผลของการทดสอบจะน่าเชื่อถือได้เมื่อกลุ่มตัวอย่างไม่น้อยกว่า 28 คน (Howell, 2007)

คาบต่อไป

- การบ้านครั้งที่ 10
- สอบครั้งที่ 2