บทนำ

โมเดลพหุระดับ (Multilevel Modeling)

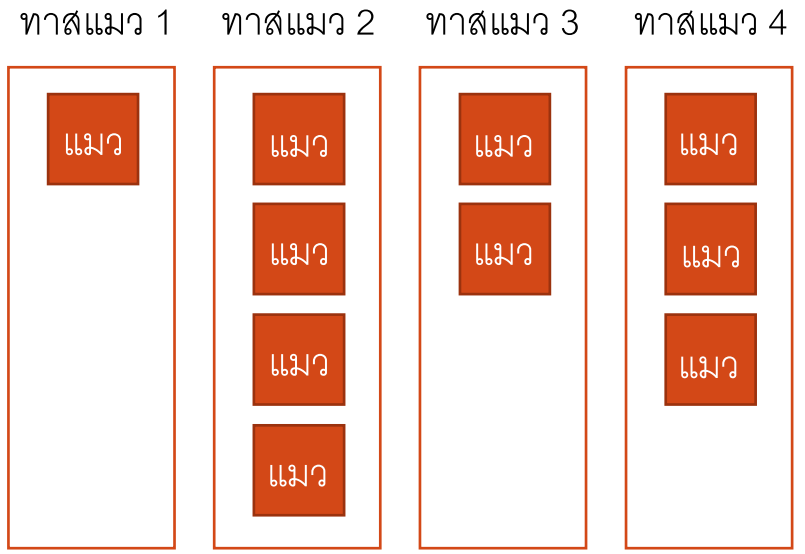
สันทัด พรประเสริฐมานิต

โครงร่างการนำเสนอ

- ระดับชั้นของข้อมูล
- รูปแบบความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ
- วิธีการวิเคราะห์ข้อมูล
- รูปแบบของตัวแปรในการวิเคราะห์พหุระดับ
- รูปแบบความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ (ต่อ)
- ความสัมพันธ์ภายในกลุ่ม
- ค่าเฉลี่ยของตัวแปรระดับกลุ่ม
- โครงสร้างของข้อมูลรูปแบบอื่นๆ

ระดับชั้นของข้อมูล

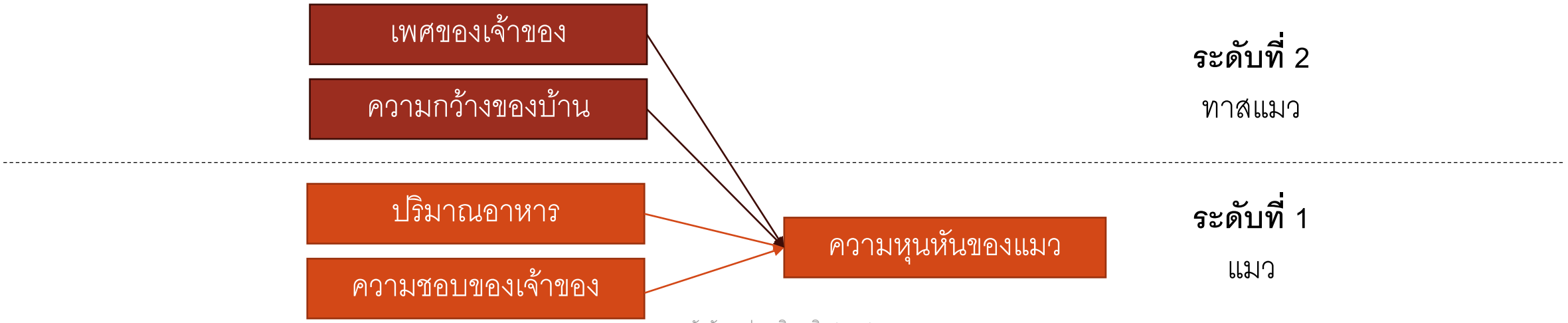
- ข้อมูลโครงสร้างซ้อนกัน (Nested Data Structure)
  - นักเรียนซ้อนในโรงเรียน
  - พนักงานซ้อนในบริษัท
  - ผู้ถูกสัมภาษณ์ซ้อนในผู้สัมภาษณ์
  - ผู้รับการบำบัดซ้อนในผู้บำบัด
  - บุคคลซ้อนในครอบครัว
  - บุคคลซ้อนในกลุ่ม
  - การวัดหลายๆ ครั้ง ซ้อนในผู้ถูกวัด



แมวซ้อนในทาสแมว

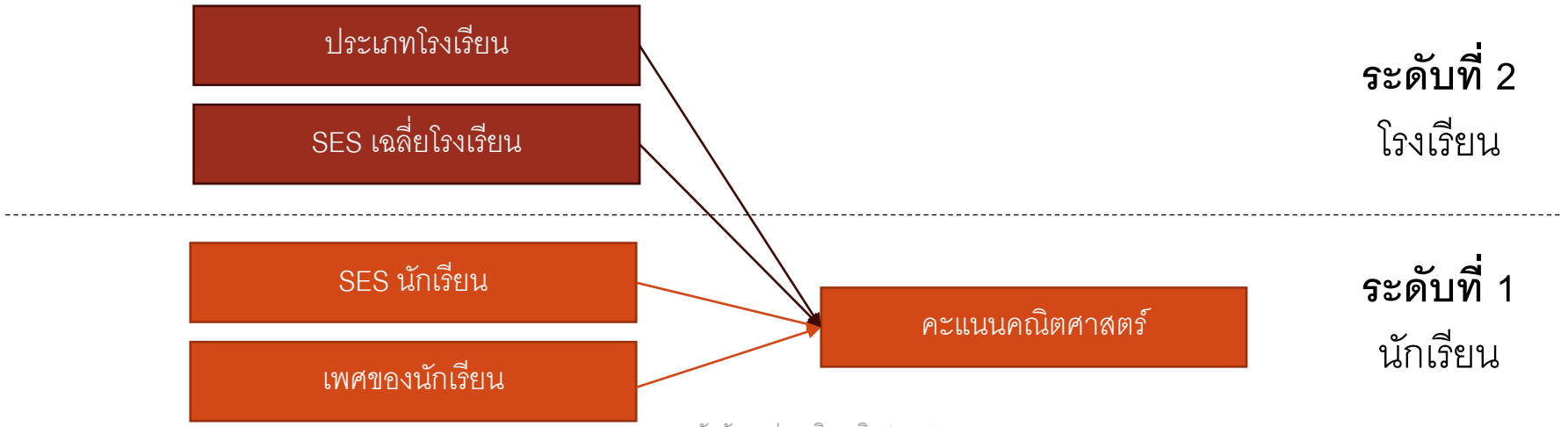
ระดับชั้นของข้อมูล

- ความหุนหันของแมว (Cat Impulsiveness) เป็นตัวแปรตาม
  - ปริมาณอาหารที่แมวกิน -> ตัวแปรต้นระดับแมว
  - ความชอบของทาสแมวต่อแมวแต่ละตัว -> ตัวแปรต้นระดับแมว
  - เพศของทาสแมว -> ตัวแปรต้นระดับทาสแมว
  - ความกว้างของบ้าน -> ตัวแปรต้นระดับทาสแมว



ระดับชั้นของข้อมูล

- คะแนนสอบวิชาคณิตศาสตร์ (Math Achievement) เป็นตัวแปรตาม
  - เพศของนักเรียน -> ตัวแปรต้นระดับนักเรียน
  - สถานะทางเศรษฐกิจของนักเรียน -> ตัวแปรต้นระดับนักเรียน
  - สถานะทางเศรษฐกิจของนักเรียนเฉลี่ยในโรงเรียน -> ตัวแปรต้นระดับโรงเรียน
  - โรงเรียนรัฐหรือเอกชน -> ตัวแปรต้นระดับโรงเรียน



ตัวแประระดับที่ 1

ตัวแประระดับที่ 2

Level 1 ID Level 2 ID

นักเรียน	โรงเรียน	คะแนนคณิตศาสตร์	เพศ	SES	$\overline{SES}$	ประเภทโรงเรียน
1	1	70	ชาย	80	85	รัฐ
2	1	80	หญิง	85	85	รัฐ
3	1	90	ชาย	90	85	รัฐ
4	2	60	ชาย	50	55	เอกชน
5	2	70	หญิง	60	55	เอกชน
6	2	80	หญิง	55	55	เอกชน
7	3	50	หญิง	30	35	เอกชน
8	3	60	ชาย	40	35	เอกชน
9	3	70	ชาย	35	35	เอกชน
10	4	40	ชาย	10	15	รัฐ
11	4	50	หญิง	15	15	รัฐ
12	4	60	หญิง	20	15	รัฐ

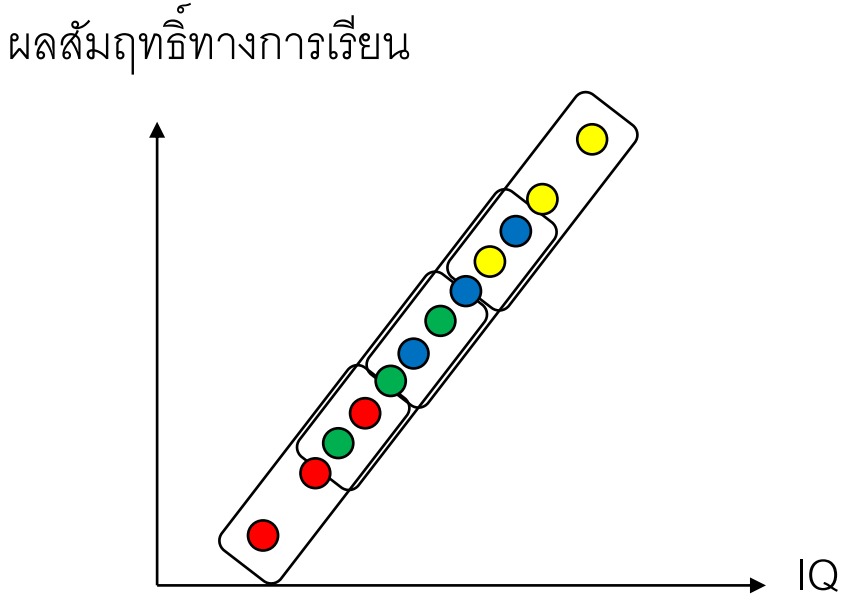
ระดับชั้นของข้อมูล

- หากข้อมูลมีสองระดับ จะเรียกชื่อระดับดังนี้

Lower Level	Upper Level
Micro Level	Macro Level
Individual Level	Group Level Cluster Level
Level 1	Level 2

รูปแบบความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ

- ความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับอาจมีค่าเท่ากัน เช่น
  - ความสัมพันธ์ระหว่าง IQ และผลสัมฤทธิ์ทางการเรียน

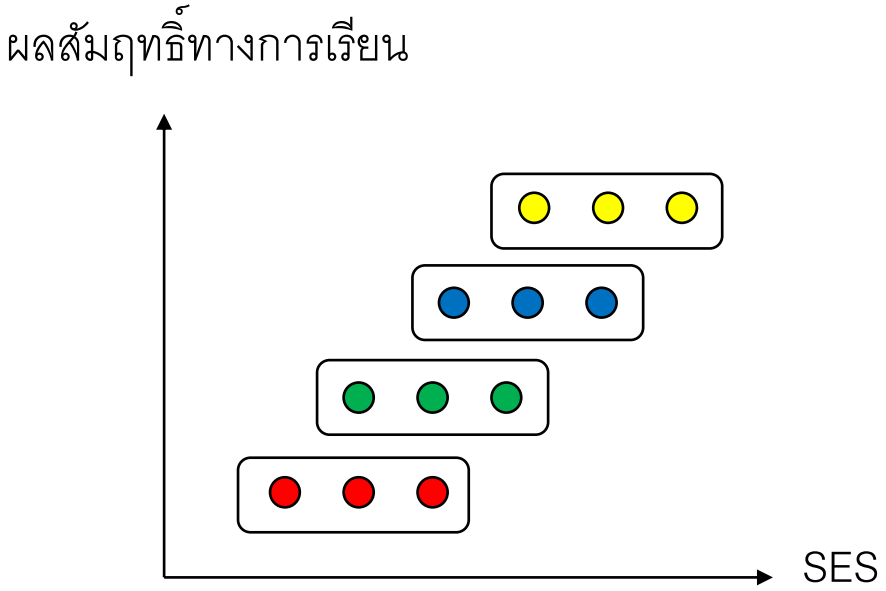


ภายในโรงเรียนเดียวกัน IQ และผลสัมฤทธิ์ทางการเรียนมีความสัมพันธ์กันทางบวก

โรงเรียนที่มีนักเรียน IQ สูง จะมีแนวโน้มที่โรงเรียนนั้นจะมีผลสัมฤทธิ์ทางการเรียนเฉลี่ยที่ดี

รูปแบบความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ

- ความสัมพันธ์ระหว่างตัวแปร ในแต่ละระดับอาจไม่เหมือนกัน เช่น
  - ความสัมพันธ์ระหว่าง SES และผลสัมฤทธิ์ทางการเรียน

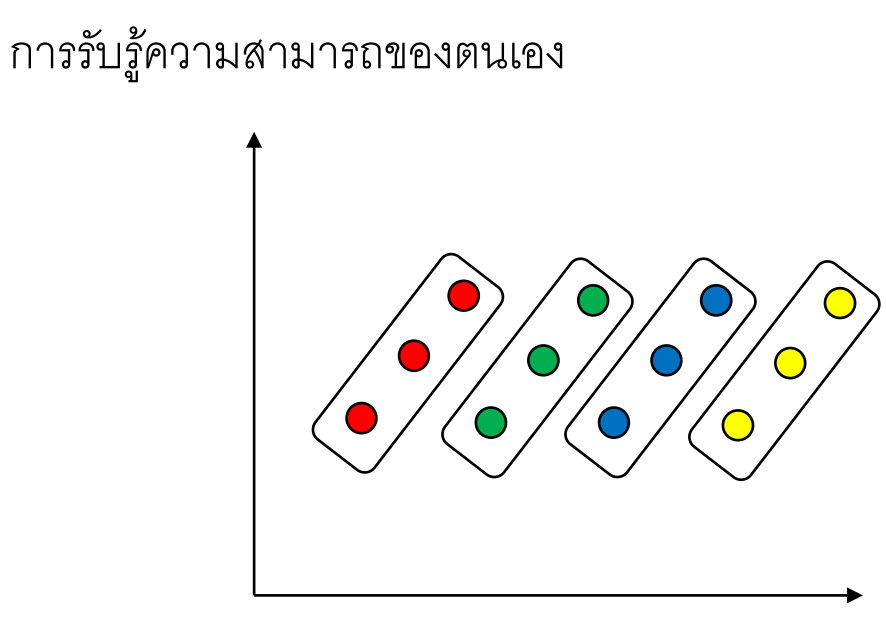


ภายในโรงเรียนเดียวกัน SES และผลสัมฤทธิ์ทางการเรียนไม่มีความสัมพันธ์กัน

โรงเรียนที่มีนักเรียน SES สูง จะมีแนวโน้มที่จะมีผลสัมฤทธิ์ทางการเรียนที่ดี

รูปแบบความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ

- ความสัมพันธ์ระหว่างตัวแปร ในแต่ละระดับอาจไม่เหมือนกัน เช่น
  - ความสัมพันธ์ระหว่างผลสัมฤทธิ์ทางการเรียนและการรับรู้ความสามารถของตนเอง

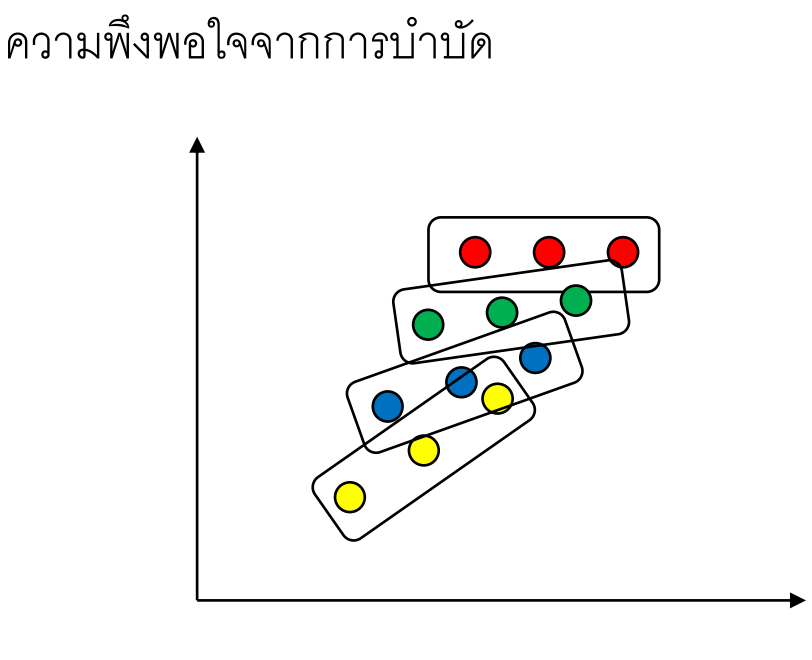


ภายในโรงเรียนเดียวกัน นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง จะมีความรับรู้ความสามารถของตนเองสูง

ผลสัมฤทธิ์ทางการเรียนระหว่างโรงเรียนไม่มีผลต่อการรับรู้ความสามารถของตนเอง

รูปแบบความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ

- ความสัมพันธ์ระหว่างตัวแปร ในแต่ละกลุ่มอาจไม่เหมือนกัน เช่น
  - ความสัมพันธ์ระหว่างความถี่เข้ารับการบำบัดและความพึงพอใจการบำบัดในการให้คำปรึกษาแบบกลุ่ม



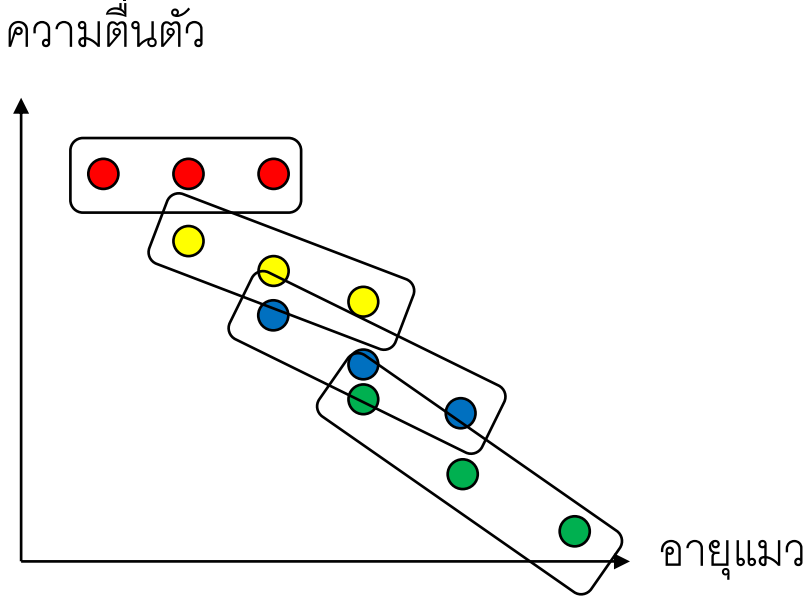
กลุ่มที่ผู้บำบัดเข้าร่วมกลุ่มมาก มีแนวโน้มพึงพอใจมากจากการบำบัด

กลุ่มที่มีการเข้าร่วมโดยเฉลี่ยสูง ผู้บำบัดที่เข้าน้อยหน่อยและมากหน่อย ไม่มีผลต่อความพึงพอใจ

กลุ่มที่มีการเข้าร่วมโดยเฉลี่ยต่ำ ผู้บำบัดที่เข้าบ่อยหน่อยจะพึงพอใจมากกว่าผู้บำบัดที่เข้าน้อย

รูปแบบความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ

- ความสัมพันธ์ระหว่างตัวแปร ในแต่ละกลุ่มอาจไม่เหมือนกัน เช่น
  - ความสัมพันธ์ระหว่างอายุของแมวและความตื่นตัวของแมว



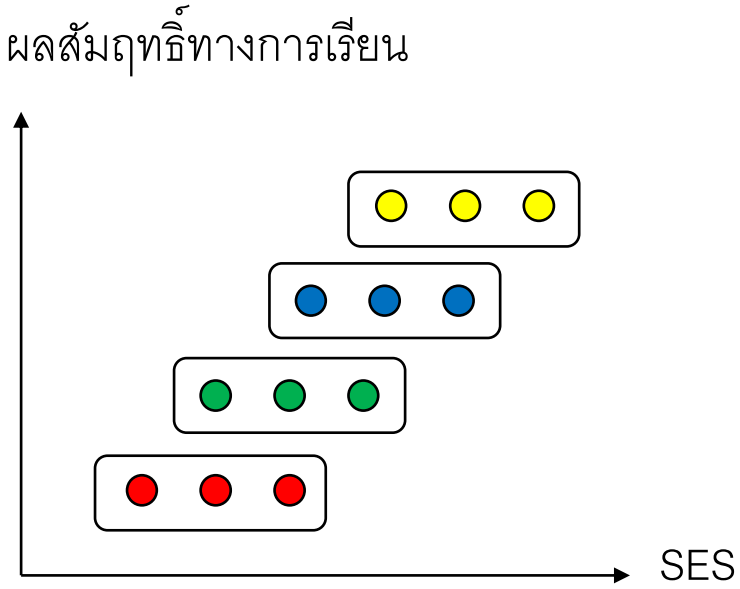
บ้านที่แมวอายุเฉลี่ยมาก จะมีความตื่นตัวน้อยกว่าบ้านที่แมวอายุเฉลี่ยน้อย

ภายในบ้านเดียวกัน ถ้าเป็นบ้านที่อายุเฉลี่ยน้อย ความตื่นตัวของแมวเด็กและแมวโตกว่าไม่ต่างกัน

ภายในบ้านเดียวกัน ถ้าเป็นบ้านที่อายุเฉลี่ยมาก แมวโตกว่าจะตื่นตัวน้อยกว่าแมวเด็กกว่ามาก

รูปแบบความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ

- ความผิดพลาดจากบริบท (Ecological Fallacy) เป็นการนำความสัมพันธ์ระดับที่สูงกว่าไปใช้ในระดับที่ต่ำกว่า

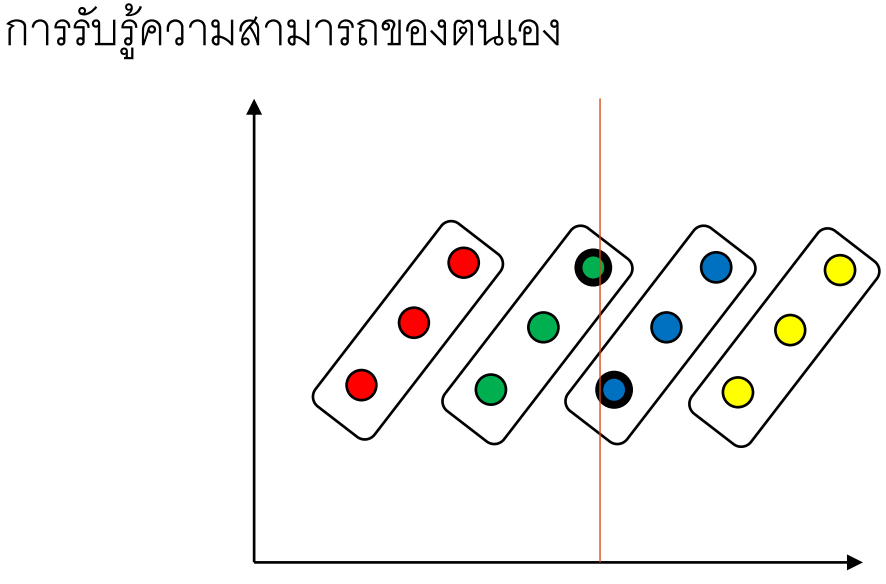


เช่น การตีความหมายว่า SES ของนักเรียน มีผลต่อผลสัมฤทธิ์ทางการเรียน

ทั้งที่แท้จริงแล้ว เกิดจาก SES ของโรงเรียน ซึ่งอาจมีผลต่อสิ่งอำนวยความสะดวกของโรงเรียน แล้วเอื้อให้นักเรียนในโรงเรียนนั้นมีผลสัมฤทธิ์ทางการเรียนที่ดี

รูปแบบความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ

- ทฤษฎีกบในหนองน้ำ (Frog Pond Theory) เป็นทฤษฎีที่บอกว่าตำแหน่งของบุคคลในกลุ่ม ที่มีผลต่อตัวแปรตาม ไม่ใช่ค่าของตัวแปรต้นโดยตรง



สังเกตที่จุดบนเส้นสีเหลือง ผลสัมฤทธิ์ทางการเรียนเท่ากัน แต่ในโรงเรียนหนึ่ง จะรับรู้ความสามารถตนเองสูง แต่อีกโรงเรียนหนึ่ง จะรับรู้ความสามารถตนเองต่ำ สิ่งที่ส่งผลจริงๆ คือเทียบกับคนในโรงเรียนแล้ว ผลสัมฤทธิ์การเรียนของตนเองอยู่ในระดับใด

วิธีการวิเคราะห์ข้อมูล

- สมมติว่าท่านต้องการวิเคราะห์อิทธิพลของสถานะทางเศรษฐกิจ (Socioeconomic Status; SES) ต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ (Math Achievement) โดยเก็บข้อมูลนักเรียน 1000 คน ใน 100 โรงเรียน
- ข้อมูลนี้เป็นข้อมูลพหุระดับ เพราะนักเรียนซ้อนในโรงเรียน
- ถ้าท่านยังรู้จักแค่การวิเคราะห์ถดถอย (Multiple Regression) ท่านอาจวิเคราะห์ได้ 2 แนวทาง
  - เอาคะแนน SES และผลสัมฤทธิ์จากนักเรียน 1000 คนมาวิเคราะห์ถดถอยตรงๆ เลย
  - หาค่าเฉลี่ย SES และ ผลสัมฤทธิ์ภายในแต่ละโรงเรียน ได้คะแนน SES และผลสัมฤทธิ์การเรียนคณิตศาสตร์ระดับโรงเรียน 100 แห่ง แล้วมาวิเคราะห์ถดถอย

วิธีการวิเคราะห์ข้อมูล

- วิธีแรก เรียกว่า การวิเคราะห์แบบแตกแยก (Disaggregation)
- วิธีที่สอง เรียกว่า การวิเคราะห์แบบรวบยอด (Aggregation)
- ทั้งสองวิธี ไม่ใช่วิธีการที่ดี

วิธีการวิเคราะห์ข้อมูล

- การวิเคราะห์แบบแตกแยก (Disaggregation) ที่ไม่ดีเพราะละเมิดข้อตกลงเบื้องต้นที่บอกว่า คะแนนทุกคะแนนจะต้องเป็นอิสระจากกัน (Independence of Observations)
- เนื่องจากว่า ภายในแต่ละกลุ่ม คะแนนจะมีความคล้ายคลึงกัน เช่น ในโรงเรียนเดียวกัน หากคะแนนผลการเรียนของนักเรียนคนนึงสูง คะแนนของนักเรียนอีกคนในโรงเรียนเดียวกันจะสูงเช่นกัน
- ข้อมูลการสุ่มสองขั้นนี้ (สุ่มโรงเรียนแล้วสุ่มนักเรียนอีกที) จะทำให้จำนวนข้อมูล (Information) น้อยกว่าการสุ่มนักเรียนอย่างง่าย 1,000 คนตรงๆ เนื่องจากนักเรียนในโรงเรียนเดียวกัน จะให้ข้อมูลที่ซ้ำซ้อนกัน
  - ยกตัวอย่างในการประมาณค่าเฉลี่ยจากข้อมูลโครงสร้างซ้อนกัน (Nested Data Structure)

ข้อมูลประชากรแบ่งเป็น 11 กลุ่ม กลุ่มละ 5 คน

0	1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10	11
2	3	4	5	6	7	8	9	10	11	12
3	4	5	6	7	8	9	10	11	12	13
4	5	6	7	8	9	10	11	12	13	14

$\mu = 7$

ประมาณค่าเฉลี่ยประชากร จาก 4 คน ด้วยวิธีสุ่มเป็น 2 วิธี

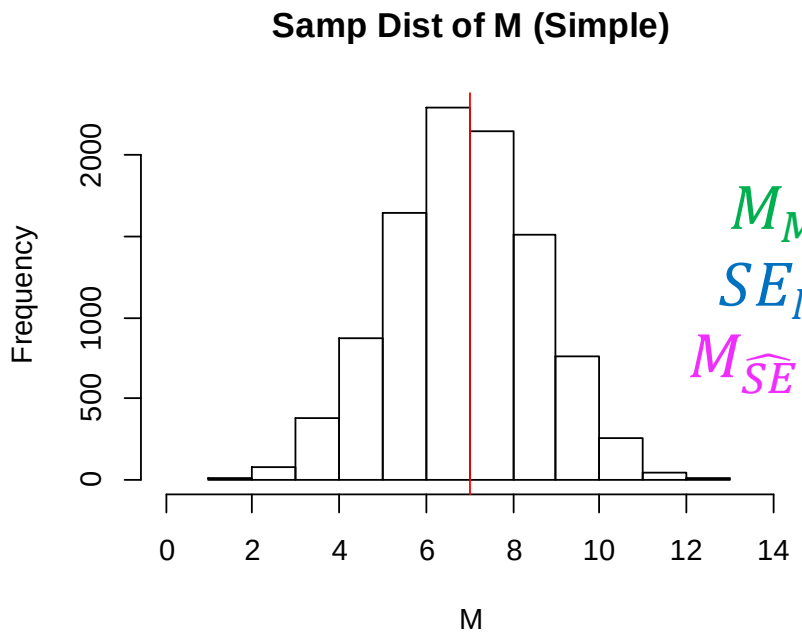
1. สุ่มอย่างง่าย (Simple Random Sampling) 4 คนโดยไม่สนใจกลุ่ม
2. สุ่มแบบลำดับขั้น (Multistage Sampling) โดยสุ่มมา 2 กลุ่มก่อน แล้วสุ่มคนด้านในกลุ่มละ 2 คน

ขั้นตอนการเปรียบเทียบจากการสุ่มทั้ง 2 วิธี

1. ค่าเฉลี่ยจากการสุ่มรูปแบบใด แม่นยำในการทำนาย μ มากกว่ากัน วิธีดังกล่าวต้องทำให้
   1) นำค่าเฉลี่ยจากการสุ่มหลายๆ ครั้ง มาเฉลี่ยอีกครั้ง ต้องตรงกับ μ
   2) นำค่าเฉลี่ยจากการสุ่มหลายๆ ครั้ง มาหาส่วนเบี่ยงเบนมาตรฐาน ต้องมีค่าต่ำ ซึ่งส่วนเบี่ยงเบนมาตรฐานนี้เรียกว่า ความผิดพลาดมาตรฐาน (Standard Error; *SE*)
2. จากแต่ละกลุ่มตัวอย่าง สามารถประมาณค่าความผิดพลาดมาตรฐานได้ด้วย $\widehat{SE}(M) = SD/\sqrt{N}$ ดูว่าการสุ่มแบบใดที่ค่า $\widehat{SE}$ ทำนาย *SE* ที่แท้จริงได้ดีที่สุด

N = จำนวนกลุ่มตัวอย่าง

ดู lecture1.R



$M_M = 7.01$

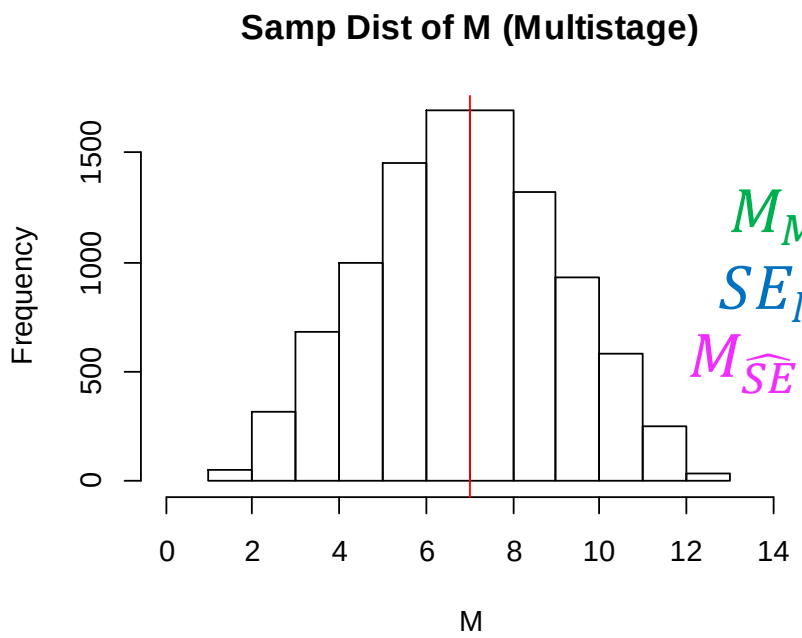
$SE_M = 1.67$

$M_{\widehat{SE}(M)} = 1.66$

ค่าเฉลี่ยของค่าเฉลี่ยจากกลุ่มตัวอย่างทั้งหมด เป็นตัวบอกว่าเราทำนายได้แม่นยำหรือไม่

ค่าส่วนเบี่ยงเบนมาตรฐานของค่าเฉลี่ย เป็นตัวบอกว่าการทำนายค่าเฉลี่ยของเราคาดเคลื่อนแค่ไหน

ค่า *SE* ที่คำนวณมาเป็นการประมาณค่าความคาดเคลื่อนของสถิติ (ค่าเฉลี่ย) ของแต่ละกลุ่มตัวอย่าง โดยค่าเฉลี่ยของ SE จากกลุ่มตัวอย่างทั้งหมด ควรใกล้เคียงกับค่าส่วนเบี่ยงเบนมาตรฐานที่เราคำนวณได้ข้างต้น



$M_M = 7.02$

$SE_M = 2.19$

$M_{\widehat{SE}(M)} = 1.40$

อคติ (Bias) ของ *M* = $M_M - \mu$

อคติของ $\widehat{SE}(M) = M_{\widehat{SE}(M)} - SE_M$

อคติสัมพัทธ์ (Relative Bias) ของ *M* = $\frac{M_M - \mu}{\mu}$

อคติสัมพัทธ์ของ $\widehat{SE}(M) = \frac{M_{\widehat{SE}(M)} - SE_M}{SE_M}$

อคติ	Simple	Multistage
M	0.01 (0.1%)	-0.016 (-1%)
SE	0.02 (0.3%)	-0.791 (-36%)

วิธีการวิเคราะห์ข้อมูล

- SE_M ของการสุ่มตัวอย่างแบบง่าย น้อยกว่า SE_M ของการสุ่มแบบลำดับขั้น
- ค่า $\widehat{SE}(M) = SD/\sqrt{N}$ ไม่เหมาะสมกับการสุ่มตัวอย่างแบบลำดับขั้น หากคำนวณค่านี้ในการสุ่มตัวอย่างแบบลำดับขั้นแล้ว ค่าที่ประมาณการได้จะต่ำกว่าความเป็นจริง
- เมื่อ *SE* ที่คำนวณได้น้อยกว่าความเป็นจริง ค่า z ที่ได้จากการทำ One-sample *z* test ก็จะสูงกว่าความจริง ทำให้ค่า *z* เกินค่าวิกฤต (Critical Value) บ่อยกว่าความจริง
- หาก H_0 เป็นจริง จะทำให้ผล sig บ่อยกว่าค่า alpha ที่ตั้งไว้ หรือเป็นปรากฏการณ์ที่เรียกว่าการเฟ้อของความผิดพลาดแบบที่ 1 (Inflated Type I Error)

วิธีการวิเคราะห์ข้อมูล

- ในการวิเคราะห์ถดถอย จะซับซ้อนกว่าการทำนายค่าเฉลี่ย เนื่องจากเป็นการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร 2 ตัว
- Moerbeek (2004) พบปรากฏการณ์คล้ายกันในการวิเคราะห์ถดถอย กล่าวคือ การวิเคราะห์ถดถอยแบบแตกแยก (Disaggregated Regression) ก็ทำให้ *SE* ต่ำกว่าที่ควรเป็น และการทดสอบทางสถิติเกิด Type I error สูงกว่าค่า Alpha ที่กำหนดไว้
- อย่างไรก็ตาม หากกลุ่มอธิบายความแปรปรวนของของทั้งสองตัวแปรน้อย (η^2 ต่ำมากหรือ ICC ต่ำมาก ซึ่งจะกล่าวถึงภายหลัง) อคติที่เกิดขึ้นจะไม่สูงมาก

วิธีการวิเคราะห์ข้อมูล

- การวิเคราะห์แบบรวบยอด (Aggregation) ที่ไม่ดีเพราะ
  - ไม่ได้สะท้อนคะแนนความสัมพันธ์ของตัวแปรในระดับนักเรียน แต่สะท้อนความสัมพันธ์ของคะแนนในระดับโรงเรียน
  - หากต้องการสะท้อนความสัมพันธ์ระดับโรงเรียนจริง การวิเคราะห์นี้ไม่ได้คำนึงถึงความเที่ยงของคะแนน (Reliability of group means)
  - กล่าวคือ คะแนนค่าเฉลี่ยกลุ่มที่แท้จริง เหมือนตัวแปรแฝงที่เราไม่ทราบค่า แต่เราใช้ค่าเฉลี่ยจากกลุ่มตัวอย่างที่เราเก็บได้ เป็นตัวทำนายค่าเฉลี่ยกลุ่มที่แท้จริง การวิเคราะห์แบบรวบยอดไม่ได้คำนึงถึงความคาดเคลื่อนตรงจุดนี้

วิธีการวิเคราะห์ข้อมูล

- ลองนึกภาพว่าต้องการหาความสัมพันธ์ระหว่างความพึงพอใจของนักเรียนต่อครูประจำชั้น ต่อความพึงพอใจในงานของครูประจำชั้น
  - ข้อมูลเป็นแบบนักเรียนซ้อนในครูประจำชั้น
  - ความพึงพอใจในงานของครูประจำชั้น เป็นข้อมูลระดับครูประจำชั้น
  - ความพึงพอใจของนักเรียนต่อครู จะเก็บข้อมูลจากนักเรียนในห้องเรียน แล้วนำมาเฉลี่ยภายในห้องเรียน ซึ่งเป็นตัวแปรระดับนักเรียนมาแปลงให้เป็นตัวแปรระดับครู
  - ถ้าห้องมี 30 คน การสุ่มเก็บข้อมูลจากนักเรียน 25 คน ย่อมมีความเที่ยงสูงกว่า (ความคาดเคลื่อนน้อยกว่ากว่า) การเก็บข้อมูลจากนักเรียน 5 คน
- ตัวอย่างต่อไปนี้ จะแสดงให้เห็น ความคาดเคลื่อนของค่าเฉลี่ยภายในกลุ่ม

ข้อมูลประชากรแบ่งเป็น 11 กลุ่ม กลุ่มละ 5 คน

0	1	2	3	4	5	6	7	8	9	10
3	4	5	6	7	8	9	10	11	12	13
4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
5	6	7	8	9	10	11	12	13	14	15
8	9	10	11	12	13	14	15	16	17	18

$\mu = 9$

แบ่งเป็น 2 สถานการณ์ เพื่อทำนาย μ คือ

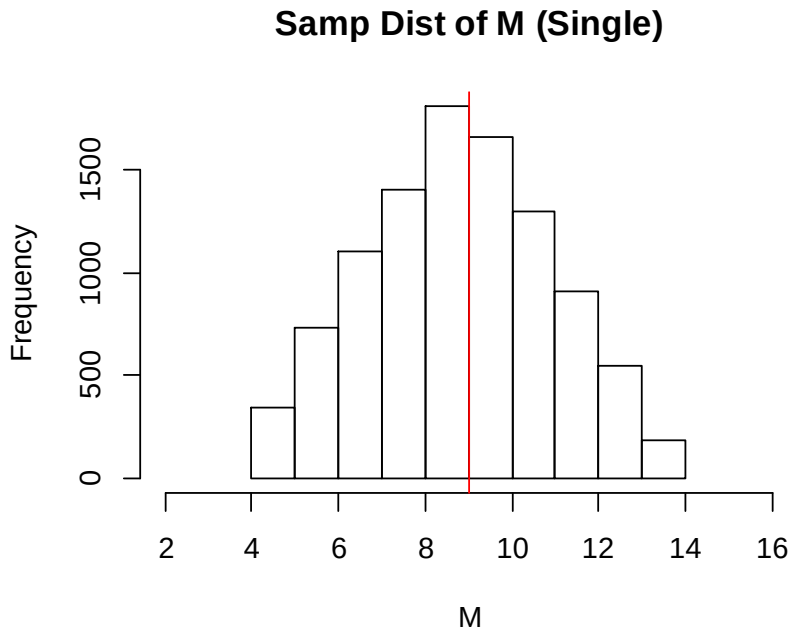
1. ใช้ตัวแปรระดับที่ 2 โดยตรง : สุ่มสีเหลือง 2 ตัวจาก 11 ตัวออกมาเฉลี่ย เพื่อทำนาย μ
2. ใช้ตัวแปรระดับที่ 1 มาแปลงเป็นตัวแปรระดับที่ 2 : สุ่ม 2 กลุ่ม ออกมา แล้วสุ่ม 2 คนมาจากแต่ละกลุ่ม หาค่าเฉลี่ยภายในแต่ละกลุ่ม แล้วน่าค่าเฉลี่ยของทั้งสามกลุ่มมาเฉลี่ยอีกครั้งเพื่อทำนาย μ

คำถามที่ใช้ในการเปรียบเทียบ

1. ค่าเฉลี่ยจากการสุ่มแบบที่สอง แม่นยำน้อยกว่าแบบแรกหรือไม่ กล่าวคือส่วนเบี่ยงเบนมาตรฐานของค่าเฉลี่ย (หรือความผิดพลาดมาตรฐาน) จากการสุ่มที่ได้จากแบบที่สองมากกว่าแบบแรกหรือไม่
2. ให้ $\widehat{SE}(M) = SD/\sqrt{k}$ แล้วค่า $\widehat{SE}$ ที่ได้จากการประมาณค่าเฉลี่ยแต่ละวิธี สามารถทำนาย SE ที่แท้จริงได้หรือไม่

k = จำนวนกลุ่ม

ดู lecture1.R



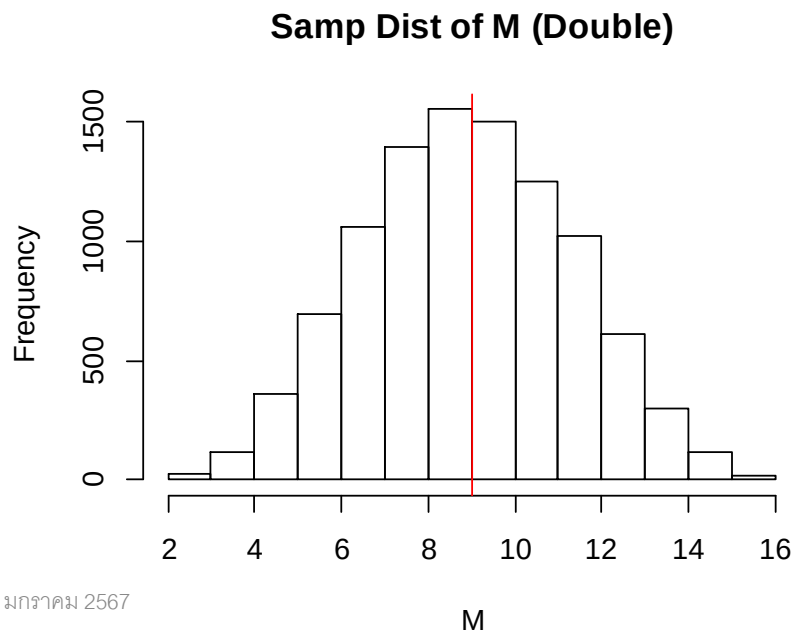
$$M_M = 9.02$$
$$SE_M = 2.12$$
$$M_{\widehat{SE}(M)} = 2.00$$

อคติ (Bias) ของ M = $M_M - \mu$

อคติของ $\widehat{SE}(M) = M_{\widehat{SE}(M)} - SE_M$

อคติสัมพัทธ์ (Relative Bias) ของ M = $\frac{M_M - \mu}{\mu}$

อคติสัมพัทธ์ของ $\widehat{SE}(M) = \frac{M_{\widehat{SE}(M)} - SE_M}{SE_M}$



$$M_M = 9.02$$
$$SE_M = 2.38$$
$$M_{\widehat{SE}(M)} = 2.13$$

อคติ	สุ่มชั้นเดียว	สุ่มสองชั้น
M	0.017 (0.1%)	0.021 (0.2%)
SE	-0.12 (-5.7%)	-0.25 (-10.4%)

วิธีการวิเคราะห์ข้อมูล

- ถ้าค่าเฉลี่ยกลุ่มไม่สามารถหาได้โดยตรง แต่ต้องประมาณการจากสมาชิกในกลุ่ม(สุ่ม 2 ชั้น) *SE* จะสูงขึ้นมากกว่าเดิม เพราะมีความแปรปรวนจากการสุ่มเพิ่มขึ้นอีกขั้น
- ค่า $\widehat{SE}(M) = SD/\sqrt{k}$ ไม่เหมาะสมกับการสุ่มตัวอย่างแบบลำดับขั้น หากคำนวณค่านี้ในการสุ่มตัวอย่างแบบลำดับขั้นแล้ว ค่าจากสูตรนี้จะต่ำกว่าความเป็นจริง
- เมื่อ *SE* ที่คำนวณได้น้อยกว่าความเป็นจริง โอกาสการเกิด Type I Error ก็จะสูงขึ้นกว่าที่ควรเป็น

วิธีการวิเคราะห์ข้อมูล

- หากความเที่ยงของคะแนนเฉลี่ยกลุ่มสูงอยู่แล้ว (กล่าวถึงในภายหลัง) อคติที่จะเกิดขึ้นจากการวิเคราะห์ถดถอยรวบยอด (Aggregated Regression) จะน้อย ซึ่งเกิดเมื่อ
  - ความแปรปรวนภายในกลุ่มต่ำ เทียบกับความแปรปรวนระหว่างกลุ่ม (η^2 หรือ ICC สูงมาก)
  - ใช้ข้อมูลนักเรียนทุกคนหรือเกือบทุกคนภายในกลุ่ม หรือใช้สมาชิกกลุ่มจำนวนสูงมาก
- อย่างไรก็ตาม การวิเคราะห์ถดถอยแบบแตกแยกและรวบยอด มีปัญหาที่ไม่สามารถเห็นความสัมพันธ์ของตัวแปรแต่ละระดับได้ ซึ่งความสัมพันธ์ของแต่ละระดับอาจไม่เท่ากัน ดังที่กล่าวไปข้างต้น

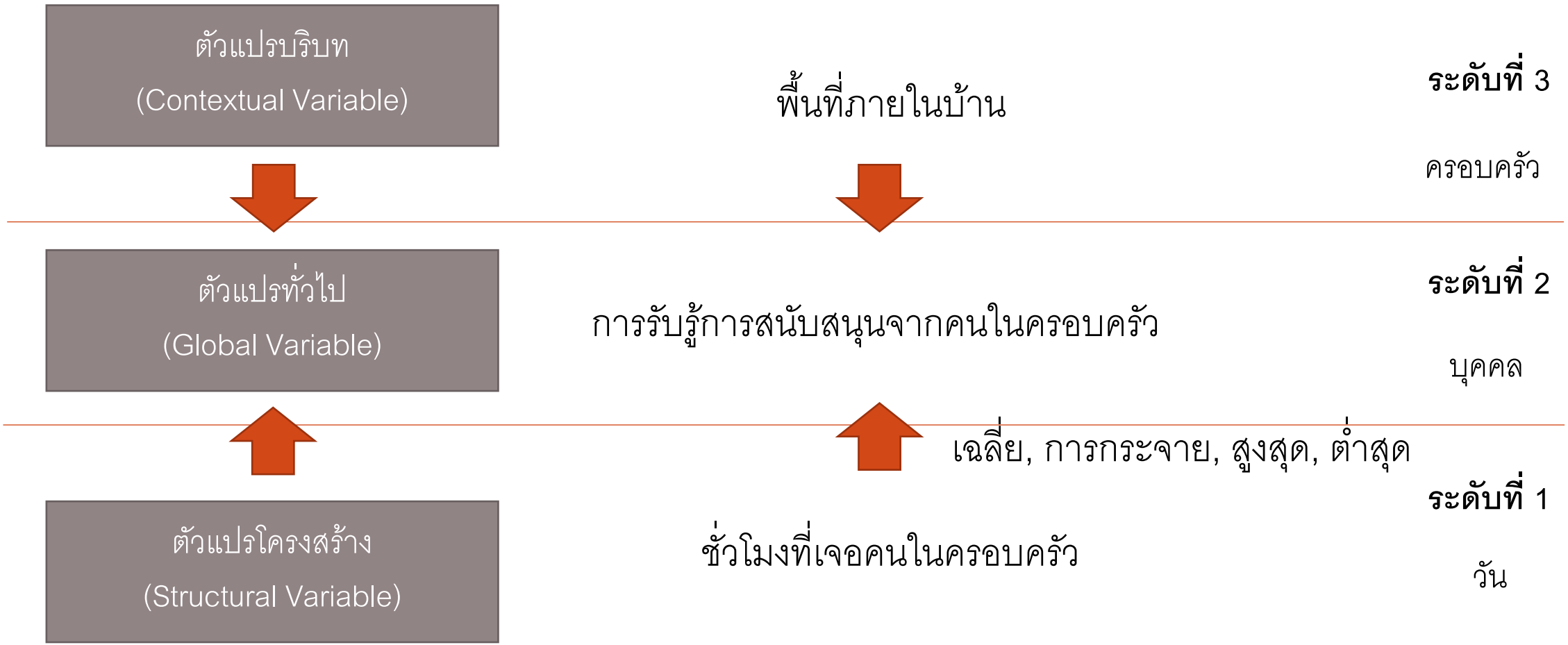
วิธีการวิเคราะห์ข้อมูล

- แน่นอนว่าการวิเคราะห์ถดถอย (หรือการวิเคราะห์ความแปรปรวน) แบบแตกแยกและแบบรวบยอด ไม่เหมาะสม ซึ่งวิธีการวิเคราะห์ที่เหมาะสม มีดังต่อไปนี้
  - การปรับความแปรปรวน (Variance adjustment) ในการประมาณค่าเฉลี่ย
  - การวิเคราะห์ความแปรปรวนแบบสุ่ม (Random-effect ANOVA)
  - การใช้น้ำหนักการสุ่ม (Sampling weight)
  - โมเดลเชิงเส้นแบบลำดับขั้น (Hierarchical Linear Modeling; HLM)
  - โมเดลสมการเชิงโครงสร้างแบบพหุระดับ (Multilevel Structural Equation Modeling; MSEM)

รูปแบบของตัวแปรในการวิเคราะห์พหุระดับ

- ตัวแปรทั่วไป (Global variable) คือตัวแปรที่เกิดการวัดในระดับที่ตัวแปรนั้นเกิดขึ้นจริงๆ เช่น IQ เกิดขึ้นในตัวแปรระดับนักเรียน
- ตัวแปรที่สร้างขึ้น (Structural variable) คือตัวแปรที่สร้างขึ้นมาจากระดับที่ต่ำกว่า เช่น ค่าเฉลี่ยของ IQ ในห้องเรียน, ส่วนเบี่ยงเบนมาตรฐานของ SES ในห้องเรียน
- ตัวแปรบริบท (Contextual variable) คือตัวแปรในระดับที่สูงกว่า แต่ถูกทำให้แตกแยกมาใช้ในระดับที่ต่ำกว่า แต่บางครั้งอาจหมายความว่าตัวแปรในระดับที่สูงกว่ามีผลต่อตัวแปรในระดับที่ต่ำกว่า เช่น ความสะอาดของที่อยู่แมว มีผลต่อความสุขของแมวแต่ละตัว

รูปแบบของตัวแปรในการวิเคราะห์พหุระดับ



รูปแบบของตัวแปรในการวิเคราะห์พหุระดับ

- ตัวแปรที่สร้างขึ้น (Structural variable) สามารถแบ่งได้ออกเป็น 2 ประเภท ตามธรรมชาติของการรวมตัวกัน
  - ตัวแปรก่อสร้างขึ้น (Formative) หรือตัวแปรจากการรวบรวม (Compilation) คือ ธรรมชาติของตัวแปรอยู่ชั้นล่าง และค่าสถิติของตัวแปรชั้นล่าง ได้สร้างขึ้นมาเป็นตัวแปรใหม่ในชั้นบน เช่น
    - สัดส่วนของเพศภายในชั้นเรียน
    - ค่าเฉลี่ยของคะแนนเชาวน์ปัญญาของนักเรียนภายในชั้นเรียน
    - ค่าสูงสุด ของอารมณ์โกรธ ตลอด 7 วันภายในบุคคล

รูปแบบของตัวแปรในการวิเคราะห์พหุระดับ

- ตัวแปรที่สร้างขึ้น (Structural variable) สามารถแบ่งได้ออกเป็น 2 ประเภท ตามธรรมชาติของการรวมตัวกัน
  - ตัวแปรสะท้อน (Reflective) หรือตัวแปรจากการประกอบ (Composition) คือ ธรรมชาติของตัวแปรอยู่ชั้นบน และคะแนนชั้นล่างที่เกิดขึ้น เป็นการวัดตัวแปรที่อยู่ชั้นบน เช่น
    - ผลการปฏิบัติงานของหัวหน้า สะท้อนจากการประเมินของผู้ใต้บังคับบัญชา
    - ความสวยของบุคคล สะท้อนจากการประเมินของผู้ประเมิน

รูปแบบของตัวแปรในการวิเคราะห์พหุระดับ

- ตัวแปรที่สร้างขึ้น (Structural variable) สามารถแบ่งได้ออกเป็น 2 ประเภท ตามธรรมชาติของการรวมตัวกัน
  - ตัวแปรจากการรวบรวม (Composition) สามารถแบ่งออกเป็น 2 ประเภทย่อย คือ
    - ตัวแปรจากการรวบรวมบริสุทธิ์ (Pure Composition) ความแตกต่างในการประเมินระดับล่าง เป็นความแตกต่างที่เกิดจากการสุ่มเท่านั้น
    - ตัวแปรจากการรวบรวมคลุมเครือ (Fuzzy Composition) ความแตกต่างในการประเมินระดับล่าง เป็นความแตกต่างจากความแตกต่างรายบุคคลรวมกับความแตกต่างที่เกิดจากการสุ่ม
  - เช่น การประเมินความสวยของเป้าหมายโดยผู้ประเมิน ซึ่งเป็นตัวแปรจากการรวบรวม ถ้าผู้ประเมินทุกคนไม่มีอคติอะไรเลย ไม่มีสเปครายบุคคลเลย ก็จะเป็นแบบบริสุทธิ์ แต่ถ้ามีอคติรายบุคคลบ้าง มีสเปครายบุคคลบ้าง ก็จะเป็นแบบคลุมเครือ

รูปแบบของตัวแปรในการวิเคราะห์พหุระดับ

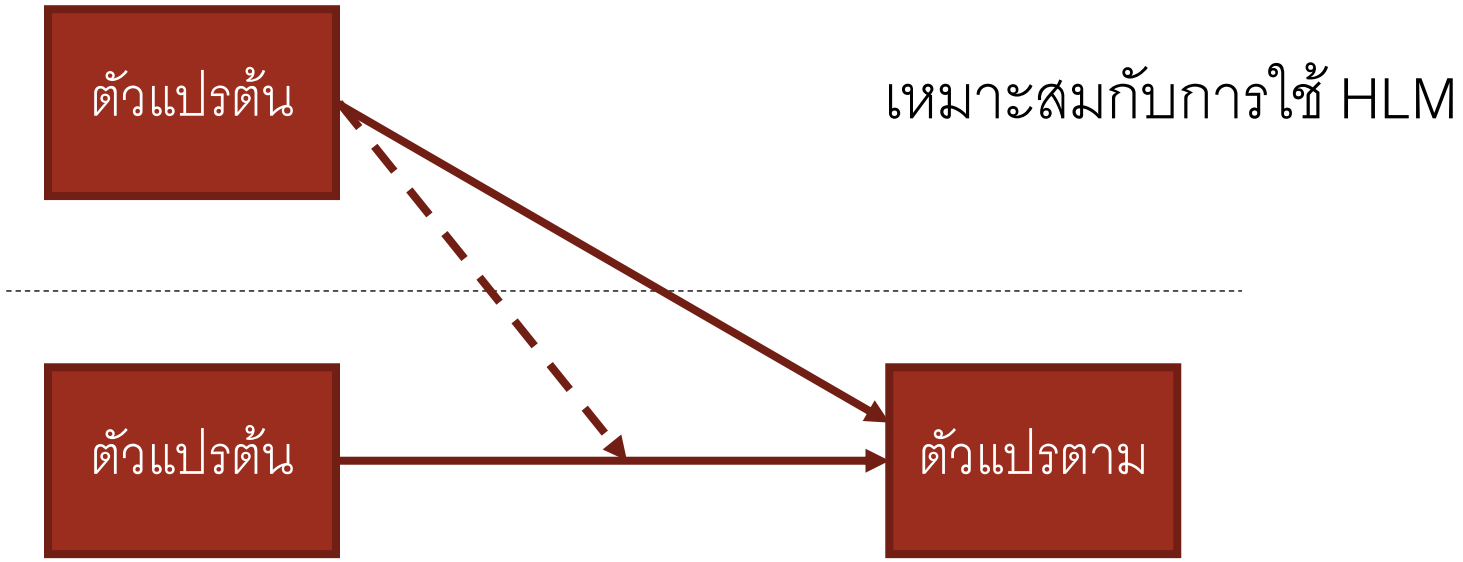
- ตัวแปรที่สร้างขึ้น (Structural variable) ทั้งหมด จะถือว่าตัวแปรทั้งสองระดับ เป็นตัวแปรคนละตัวแปรกัน มีคุณสมบัติตามทฤษฎีที่แตกต่างกัน
  - เช่น เพศ และสัดส่วนเพศ สัดส่วนเพศแสดงถึงความหลากหลายทางเพศภายในกลุ่ม ซึ่งไม่เท่ากับเพศโดยตรง
  - ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ของชั่วโมงการทำงาน ก็สะท้อนถึงวัฒนธรรม และการกระจายงานของพนักงานในแผนก ซึ่งการตีความต่างจากชั่วโมงการทำงานของแต่ละบุคคล

รูปแบบของตัวแปรในการวิเคราะห์พหุระดับ

- ข้อแตกต่างระหว่างตัวแปรจากการประกอบและตัวแปรจากการรวบรวม คือ
  - ตัวแปรจากการประกอบ จะนำข้อมูลจากสมาชิกกลุ่มที่มีจำนวนจำกัด (Finite Population) มาหาค่าสถิติ
  - ตัวแปรจากการรวบรวม จะอนุมานว่าข้อมูลจากสมาชิกกลุ่ม มีไม่จำกัด (Infinite Population) สิ่งที่เราเก็บข้อมูลมา คือการสุ่มสมาชิกกลุ่มมาเป็นตัวแทนบ่งชี้ตัวแปรระดับบน
- ตัวแปรจากการรวบรวม เหมือนค่าจริง (True Score) ในโมเดลการวัด (Measurement Model) ที่มีค่าความผิดพลาด (Error) มาเกี่ยวข้องทั้งแบบสุ่ม (Random) และแบบเป็นระบบ (Systematic)

รูปแบบความสัมพันธ์ (ต่อ)

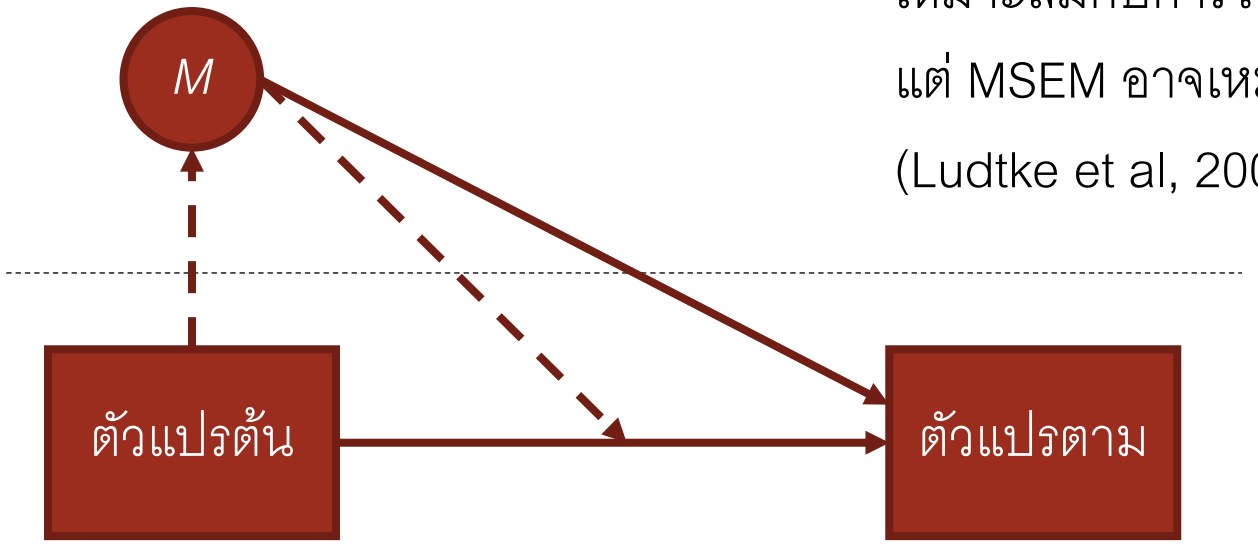
- ที่ผ่านมาได้กล่าวถึง ความสัมพันธ์ระหว่างตัวแปรระหว่างระดับ โดยตัวแปรตามเป็นตัวแปรที่อยู่ในระดับล่างสุด



เหมาะสมกับการใช้ HLM

รูปแบบความสัมพันธ์ (ต่อ)

- หากนำตัวแปรโครงสร้างมาเกี่ยวข้องแล้ว ความสัมพันธ์อาจมีได้อีกหลายๆ รูปแบบ เช่น



เหมาะสมกับการใช้ HLM
แต่ MSEM อาจเหมาะสมกว่า
(Ludtke et al, 2008)

เช่น ความสัมพันธ์ระหว่างสถานะทางเศรษฐกิจและสังคมต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนในโรงเรียน

รูปแบบความสัมพันธ์ (ต่อ)

- หากนำตัวแปรโครงสร้างมาเกี่ยวข้องแล้ว ความสัมพันธ์อาจมีได้อีกหลายๆ รูปแบบ เช่น

ปกติใช้ Aggregated Regression

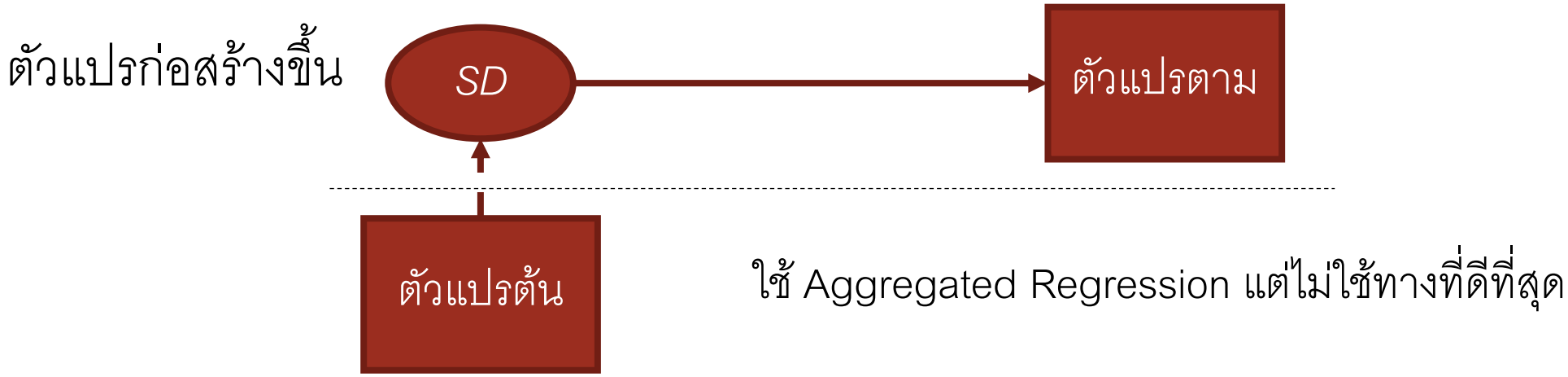
แต่ MSEM เหมาะสมกว่า (Ludtke et al, 2008)

เช่น ภาวะผู้นำของหัวหน้าที่ถูกประเมินจากลูกน้อง ส่งผลต่อผลงานของทีม

สัดส่วนของเพศหญิงภายในกลุ่ม ส่งผลต่อความสวยงามของกระทงที่ทำขึ้น

รูปแบบความสัมพันธ์ (ต่อ)

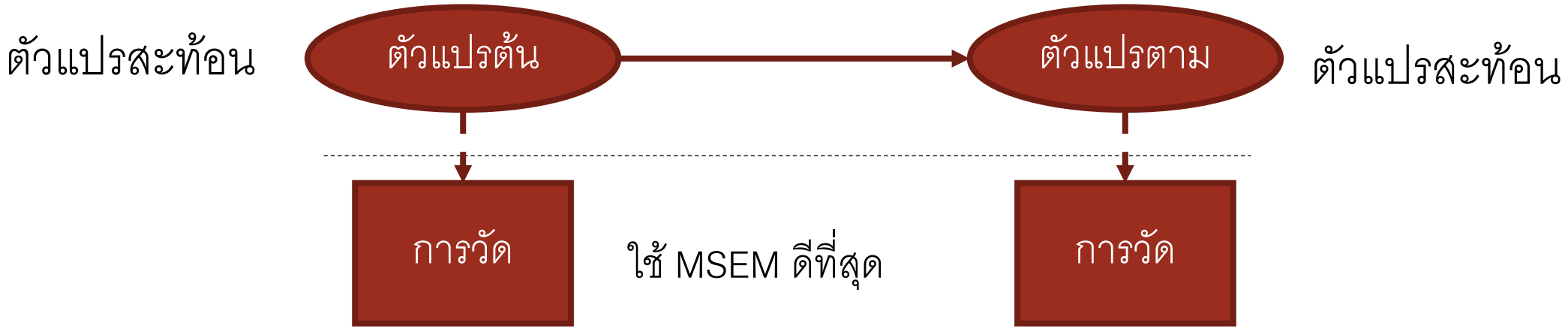
- หากนำตัวแปรโครงสร้างมาเกี่ยวข้องแล้ว ความสัมพันธ์อาจมีได้อีกหลายๆ รูปแบบ เช่น



เช่น ความแตกต่างระหว่างช่วงเวลาการนอนในแต่ละวัน ส่งผลต่อสุขภาพ
การแกว่งตัวของราคาหุ้นของแต่ละวัน ส่งผลต่อการรับรู้ความน่าเชื่อถือของหุ้น
ความแตกต่างด้านฐานะในชุมชน ส่งผลต่ออัตราอาชญากรรมภายในชุมชน

รูปแบบความสัมพันธ์ (ต่อ)

- หากนำตัวแปรโครงสร้างมาเกี่ยวข้องแล้ว ความสัมพันธ์อาจมีได้อีกหลายๆ รูปแบบ เช่น



เช่น ความสวยส่งผลต่อความน่าเชื่อถือในการพูด ตัวแปรทั้งสองประเมินจากผู้ประเมิน

ความสัมพันธ์ภายในกลุ่ม

- ความสัมพันธ์ภายในกลุ่ม (Intraclass Correlation; ICC หรือ ρ) เป็นความคล้ายคลึงกันของสมาชิกภายในกลุ่มในตัวแปรใดตัวแปรหนึ่ง
- ในการวิเคราะห์ความแปรปรวน (ANOVA) หลักการนี้จะคล้ายกับสัดส่วนความแปรปรวนระหว่างกลุ่ม (η^2 หรือ ω^2)
- ความแปรปรวนของตัวแปรใดตัวแปรหนึ่ง จะแบ่งออกเป็น 2 ส่วน คือ
  - ความแปรปรวนระหว่างกลุ่ม (Between-group variance; τ)
  - ความแปรปรวนภายในกลุ่ม (Within-group variance; σ^2)

ความสัมพันธ์ภายในกลุ่ม

- ICC จะมีค่าดังนี้

$$\rho = \frac{\tau}{\tau + \sigma^2} = \frac{\text{ความแปรปรวนระหว่างกลุ่ม}}{\text{ความแปรปรวนของตัวแปรทั้งหมด}}$$

- อย่างไรก็ตาม หากเอาค่าเฉลี่ยของแต่ละกลุ่มจากข้อมูลที่เก็บมา มาหาความแปรปรวนโดยตรง ค่าที่ได้นี้จะไม่เท่ากับ τ เนื่องจากค่าเฉลี่ยของแต่ละกลุ่มนั้น ได้รับอิทธิพลจากความผิดพลาดในการสุ่ม ในการคำนวณค่าเฉลี่ยของแต่ละกลุ่ม
- ในการคำนวณ ให้คำนวณผ่านโปรแกรมวิเคราะห์โมเดลพหุระดับโดยตรง แล้วเอาค่า τ และ σ^2 มาใส่ในสูตรจะง่ายกว่า

ความสัมพันธ์ภายในกลุ่ม

- ICC ตามทฤษฎีอาจหมายความว่า ถ้าสุ่มคนสองคนจากกลุ่มเดียวกัน คะแนนสองคนนั้นจะมีความสัมพันธ์เท่ากับค่าของ ICC
- ให้

$$Y_{ij} = \mu + U_{.j} + R_{ij}$$

- Y_{ij} คือ คะแนนของคนที่ *i* ในกลุ่มที่ *j*
- μ คือ ค่าเฉลี่ยรวม (Grand Mean) ของตัวแปรจากทุกคนในทุกกลุ่ม
- $U_{.j}$ คือ ค่าความเบี่ยงเบนจากคะแนนเฉลี่ยกลุ่มที่ *j* จากค่าเฉลี่ยรวม
  - ค่าเฉลี่ยของกลุ่ม คือ $\mu + U_{.j}$
- R_{ij} คือ ค่าความเบี่ยงเบนจากคะแนนคนที่ *i* จากค่าเฉลี่ยของกลุ่มที่ *j*

ความสัมพันธ์ภายในกลุ่ม

- หากเป็นสมการระดับประชากรแล้ว

$$\begin{aligned}\mathrm{Var}(Y_{ij}) &= \mathrm{Var}(\mu + U_{.j} + R_{ij}) \\ &= \mathrm{Var}(U_{.j} + R_{ij}) \\ &= \mathrm{Var}(U_{.j}) + \mathrm{Var}(R_{ij}) \\ &= \tau + \sigma^2\end{aligned}$$

μ เป็นค่าคงที่

$U_{.j}$ และ R_{ij} ไม่สัมพันธ์กันตามทฤษฎี

- ตรงกับความหมาย ICC ตามนิยาม คือ

$$\mathrm{ICC} = \frac{\mathrm{Var}(U_{.j})}{\mathrm{Var}(Y_{ij})} = \frac{\tau}{\tau + \sigma^2}$$

ความสัมพันธ์ภายในกลุ่ม

- ดังนั้น

$$\begin{aligned}\mathrm{Cov}\left(Y_{ij}, Y_{i'j}\right) &= \mathrm{Cov}\left(\mu + U_{.j} + R_{ij}, \mu + U_{.j} + R_{i'j}\right) \\ &= \mathrm{Cov}\left(U_{.j} + R_{ij}, U_{.j} + R_{i'j}\right) \\ &= \mathrm{Cov}\left(U_{.j}, U_{.j}\right) = \tau\end{aligned}$$

μ เป็นค่าคงที่

R_{ij} และ $R_{i'j}$ ไม่สัมพันธ์กันตามทฤษฎี และไม่สัมพันธ์กับ $U_{.j}$

$$\mathrm{Cor}\left(Y_{ij}, Y_{i'j}\right) = \frac{\mathrm{Cov}\left(Y_{ij}, Y_{i'j}\right)}{\sqrt{\mathrm{Var}\left(Y_{ij}\right) \cdot \mathrm{Var}\left(Y_{i'j}\right)}}$$

$$= \frac{\tau}{\sqrt{(\tau + \sigma^2) \cdot (\tau + \sigma^2)}} = \frac{\tau}{\tau + \sigma^2} = \mathrm{ICC}$$

ความสัมพันธ์ภายในกลุ่ม

- Hedges & Hedberg (2007) ได้รายงานว่า ในบริบทนักเรียนซ้อนในชั้นเรียน ตัวแปรทางจิตวิทยา จะมี ICC ประมาณ .05 และตัวแปรผลสัมฤทธิ์ทางการเรียน มี ICC ประมาณ .25
- หาก ICC เท่ากับหรือน้อยกว่า .05 แล้ว การวิเคราะห์แบบแตกแยกจะสร้างอคติไม่มาก พออนุโลมใช้การวิเคราะห์แบบแตกแยกได้ (Moerbeek, 2004)
- การวิเคราะห์รวบยอด สามารถอนุโลมใช้ได้ในกรณีที่
  - ตัวแปรจากการรวบรวม (Formative) ใช้ได้เมื่อ (ก) ข้อมูลภายในกลุ่มถูกเก็บข้อมูลมาทั้งหมดหรือเกือบทั้งหมด หรือ (ข) ICC สูงมาก (เช่น > .9) จนความผิดพลาดจากการสุ่ม แทบไม่มีผล
  - ตัวแปรจากการประกอบ (Reflective) ใช้ได้เมื่อ ICC สูงมาก (เช่น .9) โดยตัวแปรจากการประกอบ จะไม่สนใจจำนวนกลุ่มตัวอย่างภายในกลุ่ม แม้เก็บทุกคนก็ตาม เพราะข้อมูลจากแต่ละคน ถือเป็นตัวบ่งชี้ (Indicators) หนึ่งของภาวะสันนิษฐาน ที่ตัวบ่งชี้มีจำนวนไม่จำกัด (Infinite)

ค่าเฉลี่ยของตัวแปรระดับกลุ่ม

- ให้ Y_{ij} เป็นตัวแปรจากการรวบรวม (Composition) สามารถแปลความหมายเหมือนโมเดลการวัด (Measurement Model) ได้ดังนี้

$$O = T + E$$

$$Y_{ij} = \mu + U_{.j} + R_{ij}$$

- Y_{ij} คือ คะแนนที่สังเกตได้ (Observed Score) ของคนที่ *i* เพื่อวัดคะแนนตัวแปรจากการรวบรวมของกลุ่มที่ *j*
- $\mu + U_{.j}$ คือ ค่าคะแนนที่แท้จริง (True Score) ของตัวแปรจากการรวบรวมของกลุ่มที่ *j*
- R_{ij} คือ ค่าความเบี่ยงเบนจากคะแนนแต่ละคน ซึ่งผสมระหว่างคะแนนจำเพาะ (Systematic Error) ของคนที่ *i* และคะแนนที่เกิดจากการสุ่ม (Random Error) ของคนที่ *i*

ค่าเฉลี่ยของตัวแปรระดับกลุ่ม

- ดังนั้น ความเที่ยงของการใช้คะแนนสังเกตได้จากคนเดียว ในการประเมินตัวแปรจากการรวบรวม จึงเท่ากับ ICC

$$\rho = \frac{\tau}{\tau + \sigma^2} = \frac{Var(U_{.j})}{Var(Y_{ij})}$$

- บางครั้งในบริบทนี้ จะเรียก ICC ว่า ICC1 (Bliese, 2000)
- แต่ในการประมาณค่า เราจะได้คะแนนสังเกตได้จากหลายๆ คนมาหาค่าเฉลี่ย เพื่อประมาณตัวแปรจากการรวบรวม ความเที่ยงของค่าเฉลี่ยนี้จะเรียกว่า ความเที่ยงของค่าเฉลี่ยกลุ่ม หรือบางครั้งจะเรียกว่า ICC2 (Bliese, 2000)

- สมมติว่าได้รับการประเมินจาก n คน ความเที่ยงของค่าเฉลี่ยจะคำนวณได้ดังนี้

$$\bar{Y}_{.j} = \frac{\sum_{i=1}^{n} Y_{ij}}{n} = \frac{\sum_{i=1}^{n}(\mu + U_{.j} + R_{ij})}{n}$$

$$Var(\bar{Y}_{.j}) = Var\left(\frac{\sum_{i=1}^{n}(\mu + U_{.j} + R_{ij})}{n}\right)$$

$$Var(\bar{Y}_{.j}) = \frac{1}{n^2} Var\left(nU_{.j} + \sum_{i=1}^{n} R_{ij}\right)$$

$$Var(\bar{Y}_{.j}) = \frac{1}{n^2}\left(n^2 Var(U_{.j}) + nVar(R_{ij})\right)$$

$$Var(\bar{Y}_{.j}) = Var(U_{.j}) + \frac{Var(R_{ij})}{n}$$

$$Var(\bar{Y}_{.j}) = \tau + \frac{\sigma^2}{n}$$

μ เป็นค่าคงที่

R_{ij} และ $R_{i'j}$ ไม่สัมพันธ์กันตามทฤษฎี และไม่สัมพันธ์กับ $U_{.j}$

ค่าเฉลี่ยของตัวแปรระดับกลุ่ม

- สูตรของ ICC2 จะเป็นดังนี้

$$\text{ICC2} = \frac{Var(U_{.j})}{Var(\bar{Y}_{.j})} = \frac{\tau}{\tau + \frac{\sigma^2}{n}}$$

- ยิ่งจำนวนผู้ประเมินมาก ความเที่ยงของค่าเฉลี่ยก็จะยิ่งมากขึ้น
- การคำนวณ ICC1 และ ICC2 สามารถคำนวณได้ผ่าน ANOVA โดยสูตรสามารถดูได้จาก Bliese (2000)
- ICC2 อาจเลือกใช้จำนวนสมาชิกกลุ่มเฉลี่ยในการคำนวณ เพื่อหาค่าในภาพรวม

ค่าเฉลี่ยของตัวแปรระดับกลุ่ม

- ไม่ควรใช้ค่า ICC1 และ ICC2 แปลความหมายความเที่ยงของค่าเฉลี่ยตัวแปรระดับกลุ่มของตัวแปรจากการประกอบ (Compilation) เพราะความหมายของตัวแปรคือการประกอบกันของสมาชิกในกลุ่มที่มีจำกัด
- หากสุ่มสมาชิกทุกคนในกลุ่มมาสร้างตัวแปรจากการประกอบ คะแนนที่ได้ย่อมไม่มีความคาดเคลื่อนจากการสุ่ม อาจเปรียบเปรยว่าความเที่ยง 100%
- แต่หากไม่ได้สุ่มสมาชิกทุกคนในกลุ่ม คะแนนที่ได้ย่อมมีความคาดเคลื่อนจากการสุ่ม เช่น ในการประมาณค่าเฉลี่ย

$$SE(\bar{Y}) = \frac{SD_Y}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

ในที่นี้ N คือจำนวนคนในกลุ่มทั้งหมด และ n คือจำนวนสมาชิกที่สุ่มมา

Finite Population Correction Factor

โครงสร้างของข้อมูลรูปแบบอื่นๆ

- โมเดลสามระดับ (Three-level Model)
  - การวัด ซ้อนในนักเรียน ซ้อนในโรงเรียน
  - ครั้งที่ให้การปรึกษา ซ้อนในผู้รับคำปรึกษา ซ้อนในผู้ให้คำปรึกษา
  - พนักงานซ้อนในแผนกซ้อนในองค์กร

โครงสร้างของข้อมูลรูปแบบอื่นๆ

- โมเดลแบ่งแยกแบบซ้อนกัน (Cross-classified Model)
  - ราคาสินค้า ได้รับอิทธิพลจาก (ก) ประเภทสินค้า และ (ข) ร้านค้าปลีกที่ขายสินค้า
  - ผลสัมฤทธิ์ทางการเรียนของนักเรียน ได้รับอิทธิพลจาก (ก) โรงเรียน และ (ข) หมู่บ้าน
  - การประเมินความร้ายแรงของสถานการณ์ ได้รับอิทธิพลจาก (ก) ผู้ประเมิน และ (ข) สถานการณ์

โครงสร้างของข้อมูลรูปแบบอื่นๆ

- โมเดลเชิงซ้อนบางส่วน (Partially Nested Model)
  - การเปรียบเทียบระหว่างคนที่ได้รับการบำบัดแบบกลุ่ม กับและคนที่ไม่ได้รับการบำบัด
  - คนที่ได้รับการบำบัดแบบกลุ่ม จะมีโครงสร้างข้อมูลเป็นบุคคลซ้อนในกลุ่ม
  - คนที่ไม่ได้รับการบำบัด จะไม่มีกลุ่มตามธรรมชาติ
  - ดู Sterba, Preacher, Forehand, Hardcastle, Cole, & Compas (2014), Sterba (2017)

กิจกรรม

- แบ่งกลุ่ม คิดว่าตัวแปรต้นใด ในระดับต่างๆ มีผลต่อตัวแปรตามที่กำหนด ในโครงสร้างข้อมูลดังต่อไปนี้
  - ผลการปฏิบัติงานระดับพนักงาน (พนักงานซ้อนในองค์การ)
  - ผลการปฏิบัติงานระดับทีม (พนักงานซ้อนในทีม)
  - จำนวนการ Comments ใน Posts (Posts ซ้อนใน Pages)
  - BMI ของบุคคล (บุคคลซ้อนในครอบครัว)
  - ปริมาณการใช้ไฟฟ้าในบ้าน (บ้านซ้อนในจังหวัด)
  - Price to Earning Ratio ของบริษัทในตลาดหลักทรัพย์ (พนักงานซ้อนในองค์การ)
  - เงินที่ใช้ในการซื้อหวยของบุคคล (บุคคลซ้อนในหมู่บ้าน)
  - การเลี่ยงภาษีของบุคคล (บุคคลซ้อนในครอบครัว)

กิจกรรม

- แบ่งกลุ่ม คิดว่าตัวแปรต้นใด ในระดับต่างๆ มีผลต่อตัวแปรตามที่กำหนด ในโครงสร้างข้อมูลดังต่อไปนี้
  - การได้รับวัคซีนของเด็ก (เด็กซ้อนในครอบครัว)
  - ความเร็วในการตอบสนองต่อสิ่งเร้าที่ขึ้นบนจอ (สิ่งเร้าซ้อนในบุคคล)
  - ความถูกต้องในการทำ digit span (ข้อคำถามในบุคคล)
  - การถูก bully (เด็กซ้อนในชั้นเรียน)
  - ความรุนแรงของโทษอาชญากรรม (คดีซ้อนในหัวหน้าผู้พิพากษา)
  - คะแนนการสัมภาษณ์แบบไม่มีโครงสร้าง (ผู้ถูกสัมภาษณ์ซ้อนในผู้สัมภาษณ์)
  - ความรู้คำศัพท์ในแต่ละช่วงเวลา (การวัดซ้อนในเด็ก)

คาบต่อไป

- ให้ทบทวนการวิเคราะห์ถดถอยหรือไม่
- ให้ทบทวนการใช้ R หรือไม่