การวิเคราะห์ข้อมูลคู่ (Dyadic Data Analysis)

โมเดลพหุระดับ (Multilevel Modeling)

สันทัด พรประเสริฐมานิต

โครงร่างการนำเสนอ

- การแบ่งแยกได้ของข้อมูลคู่
- รูปแบบการจัดวางข้อมูล
- ความสัมพันธ์ภายในกลุ่ม
- โมเดลความยึดโยงระหว่างผู้แสดงและคู่

บทนำ

- การวิเคราะห์คู่ เป็นข้อมูลเชิงซ้อนรูปแบบหนึ่ง ที่สมาชิกซ้อนอยู่ในคู่
- แต่อย่างไรก็ตาม ข้อมูลคู่บางประเภท แม้สามารถวิเคราะห์ด้วย MLM ได้ แต่การวิเคราะห์รูปแบบอื่นเช่น การวิเคราะห์สมการเชิงโครงสร้าง (Structural Equation Modeling: SEM) จะง่ายกว่า
- นอกจากนี้ การวิเคราะห์ข้อมูลคู่ จะมีปัญหาเมื่อความสัมพันธ์ภายในกลุ่ม (Intraclass Correlation; ICC) ของข้อมูลคู่ไปในทางลบ

การแบ่งแยกได้ (Distinguishability)

- การแบ่งแยกได้ของคู่ หมายถึง สมาชิกของคู่สามารถแบ่งแยกได้ด้วยตัวแปรใดตัวแปรหนึ่งหรือไม่
  - คู่ที่แบ่งแยกได้ (Distinguishable Dyads) เช่น สามี-ภรรยา, พี่-น้อง, พนักงาน-ลูกค้า, หัวหน้า-ลูกน้อง
  - คู่ที่แบ่งแยกไม่ได้ (Indistinguishable Dyads) เช่น คู่แฝด, เพื่อนร่วมห้อง, เพื่อนร่วมงาน
- สถิติที่เหมาะสมต่อการวิเคราะห์ต่อสถานการณ์ทั้งสองแตกต่างกัน
  - MLM เหมาะสมกับคู่ที่แบ่งแยกไม่ได้
  - คู่ที่แบ่งแยกได้ แม้จะวิเคราะห์ด้วย MLM ได้ แต่วิเคราะห์ด้วย SEM จะง่ายกว่ามาก

การแบ่งแยกได้ (Distinguishability)

- ตัวแปรที่ใช้แบ่งแยกคู่ ต้องเป็นตัวแปรที่ “ทุกคู่” ต้องมีหน่วยหนึ่งให้ค่าหนึ่ง และอีกคนให้อีกค่าหนึ่งได้
  - เช่น A ประเมินบุคลิกภาพตนเอง และ A ประเมินบุคลิกภาพลูกน้องตนเองที่ชื่อ B และ B ก็ประเมินบุคลิกภาพของตนเองและหัวหน้าที่ชื่อ A ด้วย แบบนี้จะมีบทบาทชัดเจน ว่าเป็น “หัวหน้า” และ “ลูกน้อง”
  - แต่หาก A และ B เป็นเพื่อนสนิทกัน ไม่สามารถให้บทบาทกับใครคนใดคนหนึ่งได้ จะเป็นคู่ที่แบ่งแยกไม่ได้ แนวทางวิเคราะห์ต่างกับกรณีบน

การแบ่งแยกได้ (Distinguishability)

- คู่ที่แบ่งแยกได้ สามารถแบ่งแยกได้ทั้งทางทฤษฎี (Theoretical) ตามตัวแปรที่กล่าวไปข้างต้น แต่อย่างไรก็ตาม การแบ่งแยกดังกล่าวอาจไม่สลักสำคัญ ตามข้อมูลจริง ซึ่งสามารถทดสอบเชิงประจักษ์ (Empirical) ได้
- กล่าวคือ หากทดสอบแล้วว่า ตัวแปรที่นำมาแบ่งแยกคู่ที่แบ่งแยกได้นั้น ไม่มีผลทำให้บทบาททั้งสองบทบาทแตกต่างกัน ให้นำตัวแปรนั้นออก แล้วถือว่าเป็นคู่ที่แบ่งแยกไม่ได้
- เราจะกล่าวถึงการทดสอบเชิงประจักษ์ในการทดสอบการแบ่งแยกได้ (Test of Distinguishability) ในภายหลัง

ประเภทตัวแปร

- การวิเคราะห์คู่ จะเป็นข้อมูลเชิงซ้อน ที่สมาชิกซ้อนอยู่ในคู่
- ดังนั้น ตัวแปรสามารถแบ่งเป็น 3 ประเภท
  - ตัวแปรระหว่างคู่ (Between-Dyads Variables) คือ ตัวแปรระดับที่ 2 เช่น แต่งงานมาแล้วกี่ปี ในสามี-ภรรยา, อายุของฝาแฝด
  - ตัวแปรภายในคู่ (Within-Dyads Variables) คือ ตัวแปรระดับที่ 1 ซึ่ง ICC = 0 ซึ่งก็คือตัวแปรที่ใช้แบ่งสมาชิก ในคู่ที่แบ่งแยกได้ เช่น พ่อ-ลูก, สามี-ภรรยา
  - ตัวแปรผสม (Mixed Variables) คือ ตัวแปรระดับที่ 1 ที่ ICC ไม่เท่ากับ 0 เช่น ความพึงพอใจในชีวิตคู่ของสมาชิก (สามี-ภรรยา), บุคลิกภาพของฝาแฝดแต่ละคน

รูปแบบการจัดวางข้อมูล

- แบบยาว (Long) หรือแบบแต่ละบุคคล (Individual)
- แบบกว้าง (Wide) หรือแบบคู่ (Dyads) ซึ่งใช้ได้เฉพาะคู่ที่แบ่งแยกได้เท่านั้น
- แบบผสมคู่ (Pairwise)

แบบยาว

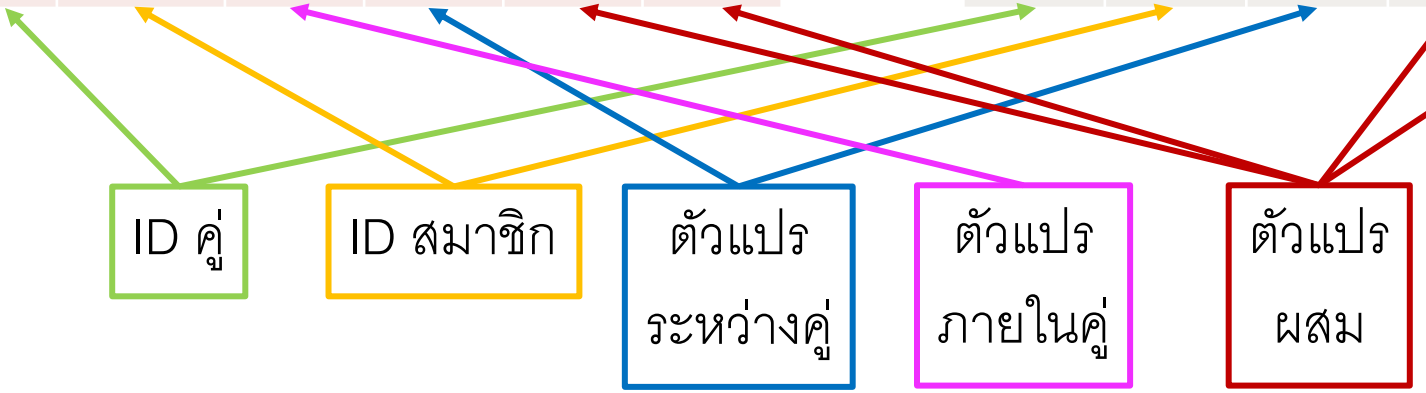
แถวแต่ละแถว คือ สมาชิกแต่ละคน

คู่ที่แบ่งแยกได้ (สามี-ภรรยา)

คู่	สมาชิก	หน้าที่	ปีแต่ง	พอใจ	รัก
1	1	สามี	2554	8	7
1	2	ภรรยา	2554	9	7
2	1	สามี	2558	6	5
2	2	ภรรยา	2558	7	7
3	1	สามี	2545	5	6
3	2	ภรรยา	2545	4	3

คู่ที่แบ่งแยกไม่ได้ (คู่แฝด)

คู่	สมาชิก	อายุ	เปิดตัว	ขี้กังวล
1	1	19	45	55
1	2	19	42	64
2	1	16	56	46
2	2	16	58	54
3	1	22	46	46
3	2	22	40	37



แบบกว้าง

แถวแต่ละแถว คือ คู่แต่ละคู่

คู่	สมาชิก	หน้าที่	ปีแต่ง	พอใจ	รัก
1	1	สามี	2554	8	7
1	2	ภรรยา	2554	9	7
2	1	สามี	2558	6	5
2	2	ภรรยา	2558	7	7
3	1	สามี	2545	5	6
3	2	ภรรยา	2545	4	3

ตัวแปรระหว่างคู่ จะมีคอลัมน์เดียว

คู่	ปีแต่ง	พอใจ สามี	พอใจ ภรรยา	รัก สามี	รัก ภรรยา
1	2554	8	9	7	7
2	2558	6	7	5	7
3	2545	5	4	6	3

ตัวแปรผสม จะแบ่งเป็นสองคอลัมน์ตามบทบาท

สามารถเปลี่ยนรูปแบบข้อมูลจากกว้างไปยาว หรือยาวไปกว้าง ด้วยคำสั่ง `reshape` ใน R

แบบผสมคู่

แถวแต่ละแถว คือ สมาชิกแต่ละคน

คอลัมน์ที่เพิ่มขึ้นมา คือ ค่าของตัวแปรผสม ของสมาชิกอีกคน

คู่ที่แบ่งแยกได้ (สามี-ภรรยา)

คู่	สมาชิก	หน้าที่ตนเอง	ปีแต่ง	พอใจตนเอง	รักตนเอง	พอใจคู่	รักคู่
1	1	สามี	2554	8	7	9	7
1	2	ภรรยา	2554	9	7	8	7
2	1	สามี	2558	6	5	7	7
2	2	ภรรยา	2558	7	7	6	5
3	1	สามี	2545	5	6	4	3
3	2	ภรรยา	2545	4	3	5	6

แบบผสมคู่

แถวแต่ละแถว คือ สมาชิกแต่ละคน

คอลัมน์ที่เพิ่มขึ้นมา คือ ค่าของตัวแปรผสม ของสมาชิกอีกคน

คู่ที่แบ่งแยกไม่ได้ (คู่แฝด)

คู่	สมาชิก	อายุ	เปิดตัวตนเอง	ขี้กังวลตนเอง	เปิดตัวคู่	ขี้กังวลคู่
1	1	19	45	55	42	64
1	2	19	42	64	45	55
2	1	16	56	46	58	54
2	2	16	58	54	56	46
3	1	22	46	46	40	37
3	2	22	40	37	46	46

ความสัมพันธ์ภายในกลุ่ม

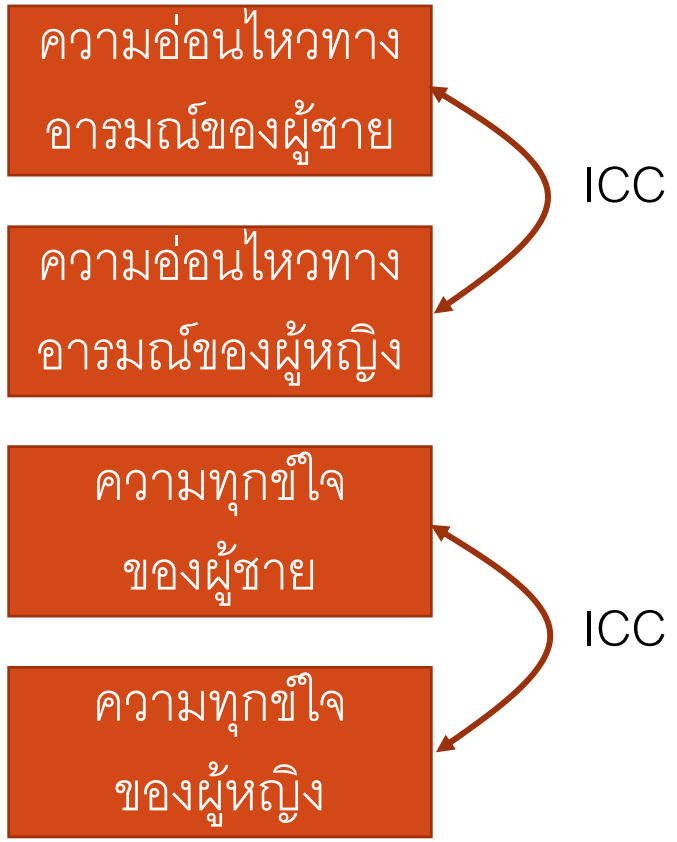
- ความสัมพันธ์ภายในกลุ่ม (Intraclass Correlation) ใน MLM จะหมายถึงสัดส่วนความแปรปรวนของคะแนนในระดับที่ 2 ต่อความแปรปรวนของตัวแปรทั้งหมด ซึ่งมีค่าตั้งแต่ 0 ถึง 1
- จากในบทที่ 1 เคยพิสูจน์แล้วว่า แท้จริงแล้ว ICC ก็คือค่าสหสัมพันธ์ระหว่างคะแนนของสมาชิกในภายในกลุ่มหรือ $Cor\left(Y_{ij}, Y_{i'j}\right)$ ซึ่งบริบทของคู่ ก็คือ สหสัมพันธ์ระหว่างสมาชิกในคู่เดียวกันนั่นเอง
- แม้ตามนิยาม ICC ในบริบท MLM จะมีค่าเป็นบวกเท่านั้น แต่หากพิจารณาว่าเป็นค่าสหสัมพันธ์ระหว่างคะแนนสมาชิก ICC ก็สามารถมีค่าเป็นลบได้

ความสัมพันธ์ภายในกลุ่ม

- ค่าทางลบ หมายถึง ถ้าค่าของคนหนึ่งสูง ค่าอีกคนหนึ่งจะต่ำ ค่าอีกคนหนึ่งต่ำ ค่าของอีกคนสูง เช่น การทำงานบ้านของเพื่อนร่วมห้องเช่า หากคนหนึ่งทำงานบ้านเยอะ อีกคนมีแนวโน้มจะทำงานบ้านน้อย
- ต่างกับ ค่าทางบวก ที่ค่าของคนหนึ่งสูง ค่าอีกคนหนึ่งจะสูง ค่าอีกคนหนึ่งต่ำ ค่าของอีกคนต่ำ เช่น สามีมีความคิดเห็นทางการเมืองในรูปแบบใด ภรรยาจะมีความคิดเห็นทางการเมืองใกล้เคียงกัน
- วิธีการที่เหมาะสมในการหา ICC จะแตกต่างกันตามประเภทของคู่

ความสัมพันธ์ภายในกลุ่ม

เก็บข้อมูลคู่ชีวิต 250 คู่ โดยวัดอิทธิพลต่อความอ่อนไหวทางอารมณ์ (Neuroticism) ของแต่ละคนที่มีต่อความทุกข์ใจ (Distress) ของแต่ละคน



แบ่งบทบาทได้ชัดเจน จึงเป็น**คู่ที่แบ่งแยกได้**

ความสัมพันธ์ภายในกลุ่ม

ID ของแต่ละคู่

```
> dat1 <-read.table("lecture12ex1.csv", sep=",", header=TRUE)
> head(dat1)
  coupleid menneuro mendistress womenneuro womendistress lengthrela
1        1 57.96475    46.85858   46.86245       41.20560   2.756603
2        2 59.67379    51.74511   45.94048       59.98045   4.868493
3        3 55.40466    41.23618   35.45253       44.24412   2.329553
4        4 62.87687    48.03789   51.41035       48.63686   3.428704
5        5 37.22799    48.83937   64.84604       43.73324   2.295287
6        6 53.68804    48.67451   39.46492       54.74921   3.663715
> psych::describe(dat1)
             vars   n   mean    sd median trimmed   mad   min    max  range  skew kurtosis   se
coupleid        1 250 125.50 72.31 125.50  125.50 92.66  1.00 250.00 249.00  0.00    -1.21 4.57
menneuro        2 250  50.83 10.27  51.75   51.16  9.88 17.31  76.30  58.99 -0.35     0.02 0.65
mendistress     3 250  49.78  9.23  50.39   50.07  9.40 25.78  71.54  45.76 -0.26    -0.30 0.58
womenneuro      4 250  50.00 10.24  50.29   49.94 10.76 20.67  77.09  56.42 -0.02    -0.30 0.65
womendistress   5 250  50.40  8.76  51.13   50.41  8.85 27.55  73.31  45.76 -0.03    -0.23 0.55
lengthrela      6 250   2.59  1.67   2.30    2.37  1.25  0.10  10.58  10.48  1.54     3.33 0.11
```

ข้อมูลเป็นแบบกว้าง

ตัวแปรผสมมี ความอ่อนไหวทางอารมณ์ของชายและหญิง และความทุกข์ใจของชายและหญิง

ตัวแปรระหว่างคู่มี อายุของความสัมพันธ์

ความสัมพันธ์ภายในกลุ่ม

- คู่ที่แบ่งแยกได้ หาได้ด้วยการหาสหสัมพันธ์ระหว่างตัวแปรเดียวกัน ของทั้งสองคน

```
> with(dat1, cor(menneuro, womenneuro))
[1] -0.06424998
> with(dat1, cor(mendistress, womendistress))
[1] 0.7704816
```

- ความอ่อนไหวทางอารมณ์มีค่า ICC = -.06
- ความทุกข์ใจมีค่า ICC = .77

ความสัมพันธ์ภายในกลุ่ม

เก็บข้อมูลเพื่อนคู่หู 500 คู่ โดยวัดบุคลิกภาพของคู่หูทั้งสองคน มีความอ่อนไหวทางอารมณ์ (Neuroticism) และการเปิดตัว (Extroversion)

คนที่ 1 และ 2 สามารถสลับกันได้ จึงเป็น**คู่ที่แบ่งแยกไม่ได้**

ความสัมพันธ์ภายในกลุ่ม

ID ของแต่ละคู่ (ระดับที่ 2)

ID ของแต่ละคน (ระดับที่ 1)

```
> dat2 <-read.table("lecture12ex2.csv", sep=",", header=TRUE)
> head(dat2)
  coupleid subjectid neuro extro
1        1         1    49    40
2        1         2    62    58
3        2         3    37    54
4        2         4    53    45
5        3         5    47    56
6        3         6    56    62
> psych::describe(dat2)
           vars    n   mean     sd median trimmed    mad min  max range skew kurtosis   se
coupleid      1 1000 250.50 144.41  250.5  250.50 185.32   1  500   499 0.00    -1.20 4.57
subjectid     2 1000 500.50 288.82  500.5  500.50 370.65   1 1000   999 0.00    -1.20 9.13
neuro         3 1000  49.61  10.02   49.0   49.43  10.38  16   80    64 0.15    -0.15 0.32
extro         4 1000  49.58   9.68   49.0   49.48  10.38  20   75    55 0.03    -0.09 0.31
```

ข้อมูลเป็นแบบยาว เหมือนข้อมูล MLM ปกติ

ตัวแปรผสมมี ความอ่อนไหวทางอารมณ์ และการเปิดตัว

- คู่ที่แบ่งแยกไม่ได้ แม้จะสามารถหา ICC ได้จากการวิเคราะห์พหุระดับ แต่จะมีปัญหากรณีที่ ICC มีค่าลบ

```
> library(lme4)
> out2a <- lmer(extro ~ 1 + (1|coupleid), data=dat2, REML=
> summary(out2a)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: extro ~ 1 + (1 | coupleid)
   Data: dat2

     AIC      BIC   logLik deviance df.resid
  7325.6   7340.3  -3659.8   7319.6      997

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.1928 -0.5824 -0.0212  0.5686  2.6089

Random effects:
 Groups   Name        Variance Std.Dev.
 coupleid (Intercept) 30.70    5.541
 Residual             62.86    7.929
Number of obs: 1000, groups:  coupleid, 500

Fixed effects:
            Estimate Std. Error t value
(Intercept)  49.5760     0.3525   140.6
```

$$ICC = \frac{30.70}{30.70 + 62.86} = .33$$

```
> out2b <- lmer(neuro ~ 1 + (1|coupleid), data=dat2, REML=FALSE)
boundary (singular) fit: see ?isSingular
> summary(out2b)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: neuro ~ 1 + (1 | coupleid)
   Data: dat2

     AIC      BIC   logLik deviance df.resid
  7452.8   7467.5  -3723.4   7446.8      997

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.3547 -0.6598 -0.0609  0.6378  3.0333

Random effects:
 Groups   Name        Variance Std.Dev.
 coupleid (Intercept)   0.0     0.00
 Residual             100.4    10.02
Number of obs: 1000, groups:  coupleid, 500

Fixed effects:
            Estimate Std. Error t value
(Intercept)  49.6100     0.3168   156.6
optimizer (nloptwrap) convergence code: 0 (OK)
boundary (singular) fit: see ?isSingular
```

$$ICC = \frac{0}{0 + 100.4} = 0$$

จะเห็นว่าเกิดปัญหาในการประมาณค่าจากโมเดลนี้ เนื่องจากข้อจำกัดที่ความแปรปรวนต้องมากกว่า 0

ความสัมพันธ์ภายในกลุ่ม

- มีโมเดลพหุระดับอีกรูปแบบหนึ่ง คือ แทนที่จะสร้างโมเดลสองระดับ ที่สร้างค่าคงเหลือระดับคู่ (ระดับที่ 2, u_{0j}) โมเดลนี้จะวิเคราะห์ถดถอยปกติ โดยยอมให้คะแนนของสมาชิกในแต่ละตัว มีค่าคงเหลือสัมพันธ์กัน (Error Correlation) ซึ่งคล้ายกับกรณีความสัมพันธ์ในตัว
- เมื่อคู่ที่เป็นข้อมูลเชิงซ้อน ถูกจัดการด้วยค่าคงเหลือที่สัมพันธ์กัน จะสามารถจัดการกับข้อมูลที่ ICC ทางลบได้ เพราะค่าสหสัมพันธ์มีค่าตั้งแต่ -1 ถึง 1
- สามารถใช้คำสั่ง `gls` ใน package `nlme` ซึ่งมีลักษณะคล้ายกับคำสั่ง `lm` สำหรับการวิเคราะห์ถดถอยปกติ แต่อนุญาตให้ความแปรปรวนของค่าคงเหลือแตกต่างกัน และอนุญาตให้ค่าคงเหลือสัมพันธ์กัน

gls จะคล้ายกับคำสั่ง lme แต่ไม่มีให้ใส่ random

```
> library(nlme)
> out2c <- gls(extro ~ 1, data = dat2,
+              correlation = corCompSymm(form = ~1|coupleid))
> summary(out2c)
Generalized least squares fit by REML
  Model: extro ~ 1
  Data: dat2
       AIC      BIC    logLik
  7325.848 7340.569 -3659.924

Correlation Structure: Compound symmetry
 Formula: ~1 | coupleid
 Parameter estimate(s):
      Rho
0.3290423

Coefficients:
             Value Std.Error  t-value p-value
(Intercept) 49.576 0.3528763 140.4912       0

Standardized residuals:
        Min          Q1         Med          Q3         Max
-3.05552852 -0.67937367 -0.05950718  0.66367038  2.62658092

Residual standard error: 9.679504
Degrees of freedom: 1000 total; 999 residual
```

บอกว่า ค่าคงเหลือมีโครงสร้างแบบ Compound Symmetry ที่ค่าคงเหลือมีความแปรปรวนเท่ากัน และความแปรปรวนร่วมมีค่าเท่ากัน ซึ่งยอมให้คนที่อยู่ในคู่เดียวกัน มีค่าคงเหลือสัมพันธ์กัน

ค่าสหสัมพันธ์ของค่าคงเหลือ คือ ค่า ICC (ตรงกับการวิเคราะห์ MLM ปกติ)

ค่าเฉลี่ยของการเปิดตัวทุกคนในทุกคู่เท่ากับ 49.58 คะแนน

ส่วนเบี่ยงเบนมาตรฐานของค่าคงเหลือ

```
> out2d <- gls(neuro ~ 1, data = dat2,
+              correlation = corCompSymm(form = ~1|coupleid))
> summary(out2d)
Generalized least squares fit by REML
  Model: neuro ~ 1
  Data: dat2
       AIC      BIC    logLik
  7332.788 7347.508 -3663.394

Correlation Structure: Compound symmetry
 Formula: ~1 | coupleid
 Parameter estimate(s):
      Rho
-0.462961
```

ค่าสหสัมพันธ์ของค่าคงเหลือ คือ ค่า ICC เท่ากับ -.46

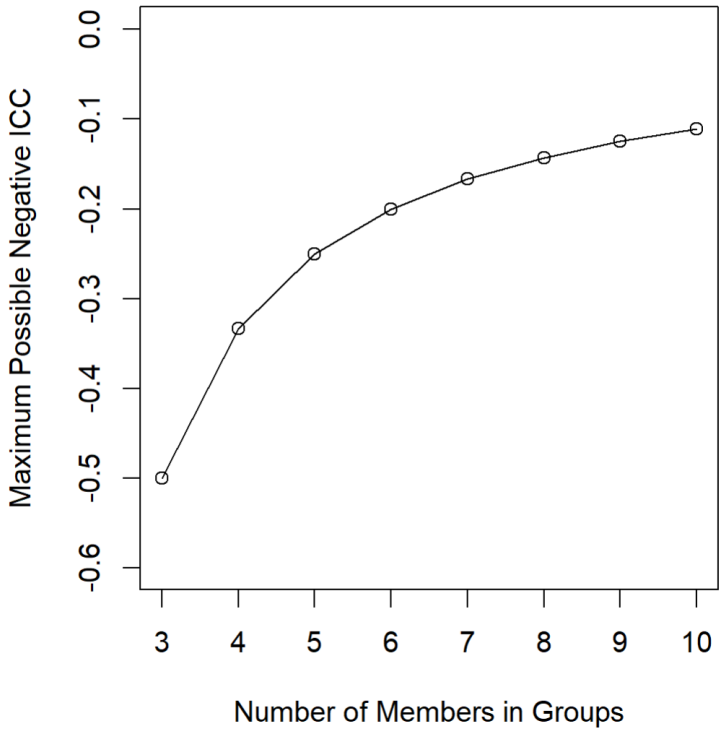
```
Coefficients:
            Value Std.Error  t-value p-value
(Intercept) 49.61 0.2322362 213.6187       0

Standardized residuals:
        Min          Q1         Med          Q3         Max
-3.35383324 -0.65959053 -0.06086993  0.63763744  3.03251985

Residual standard error: 10.02137
Degrees of freedom: 1000 total; 999 residual
```

ค่าเฉลี่ยของความอ่อนไหวทางอารมณ์ทุกคนในทุกคู่เท่ากับ 49.61 คะแนน

สังเกตว่า หากวิเคราะห์ด้วย MLM ปกติ จะประมาณค่าความแปรปรวนระดับคู่ไม่ได้ ซึ่งเป็นสาเหตุจาก ICC ทางลบ ที่ค่าความอ่อนไหวทางอารมณ์คนหนึ่งสูง อีกคนจะมีค่าต่ำ ราวกับว่าคู่หูต้องมีคนหนึ่งอารมณ์อ่อนไหว และอีกคนอารมณ์มั่นคง

ความสัมพันธ์ภายในกลุ่ม

- ดังนั้น หากมีข้อมูลแบบคู่ จึงควรใช้คำสั่ง `gls` แทนที่จะใช้ MLM ปกติ
- เราจำเป็นต้องกังวลกับค่า ICC ทางลบ ในกรณีที่มีสมาชิกกลุ่มมากกว่า 2 คนหรือไม่
  - ยิ่งสมาชิกกลุ่มยิ่งเพิ่ม ICC ทางลบจะยิ่งเกิดยากมากขึ้น
  - ลองนึกภาพว่า ถ้ามี 4 คนในกลุ่ม การจับคู่ 6 คู่ ต้องมีความสัมพันธ์ทางลบทั้งหมด ซึ่งเป็นไปได้ยากมาก หากความสัมพันธ์ทางลบอยู่ในระดับสูง



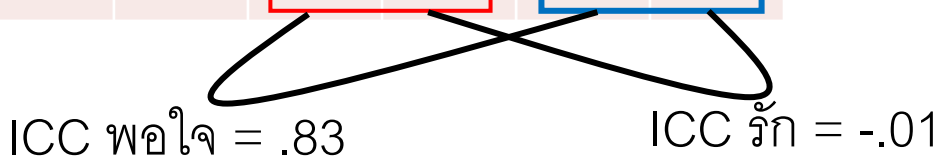
ความสัมพันธ์ภายในกลุ่ม

- ยิ่งสมาชิกกลุ่มยิ่งเพิ่ม โอกาสเกิด ICC ทางลบยิ่งยาก และค่าอาจมีค่าน้อย จนมองว่าเป็น 0 ก็ไม่ได้ทำให้การวิเคราะห์ผิดไปมากนัก
- นอกจากนี้ `gls` ยังไม่สามารถจัดการความชันสุ่มได้ ซึ่งกรณีข้อมูลแบบคู่ยิ่งเหมาะสมกับ `gls` เพราะจะไม่สามารถมีความชันสุ่มได้อยู่แล้ว
  - จำนวนอิทธิพลสุ่มที่เป็นไปได้ คือ จำนวนสมาชิกกลุ่ม – 1 ซึ่งถูกนำไปใช้ในการประมาณค่าจุดตัดสุ่ม (หรือความสัมพันธ์ระหว่างค่าคงเหลือ)
- ดังนั้น หากเป็นข้อมูลที่สมาชิกมากกว่าคู่ เช่น คู่สาม (Triad) เราอาจทดสอบด้วย `gls` ก่อนว่ามี ICC ทางลบหรือไม่ ถ้าไม่ก็ใช้ MLM แบบปกติ

ความสัมพันธ์ภายในกลุ่ม

- อีกวิธีการหนึ่ง ที่สามารถหา ICC ของคู่ที่แบ่งแยกไม่ได้ คือ การสร้างข้อมูลผสมคู่ (Pairwise) แล้วหาความสัมพันธ์จากตัวแปรเดียวกัน

คู่	สมาชิก	หน้าที่ตนเอง	ปีแต่ง	พอใจตนเอง	รักตนเอง	พอใจคู่	รักคู่
1	1	สามี	2554	8	7	9	7
1	2	ภรรยา	2554	9	7	8	7
2	1	สามี	2558	6	5	7	7
2	2	ภรรยา	2558	7	7	6	5
3	1	สามี	2545	5	6	4	3
3	2	ภรรยา	2545	4	3	5	6



```
> dat2a <- dat2[order(dat2$coupleid, dat2$subjectid),]
> head(dat2a)
  coupleid subjectid neuro extro
1        1         1    49    40
2        1         2    62    58
3        2         3    37    54
4        2         4    53    45
5        3         5    47    56
6        3         6    56    62
```

เรียงข้อมูลโดยให้เรียง ID คู่ก่อนจากน้อยไปมาก แล้วเรียง ID สมาชิกจากน้อยไปมาก

```
> dat2b <- dat2[order(dat2$coupleid, -dat2$subjectid),]
> head(dat2b)
  coupleid subjectid neuro extro
2        1         2    62    58
1        1         1    49    40
4        2         4    53    45
3        2         3    37    54
6        3         6    56    62
5        3         5    47    56
```

เรียงข้อมูลโดยให้เรียง ID คู่ก่อนจากน้อยไปมาก แล้วเรียง ID สมาชิกจากมากไปน้อย

เอาสองคอลัมน์นี้ไปแปะข้อมูลด้านบน

```
> dat2a$extropartner <- dat2b$extro
> dat2a$neuropartner <- dat2b$neuro
> head(dat2a)
  coupleid subjectid neuro extro extropartner neuropartner
1        1         1    49    40           58           62
2        1         2    62    58           40           49
3        2         3    37    54           45           53
4        2         4    53    45           54           37
5        3         5    47    56           62           56
6        3         6    56    62           56           47
```

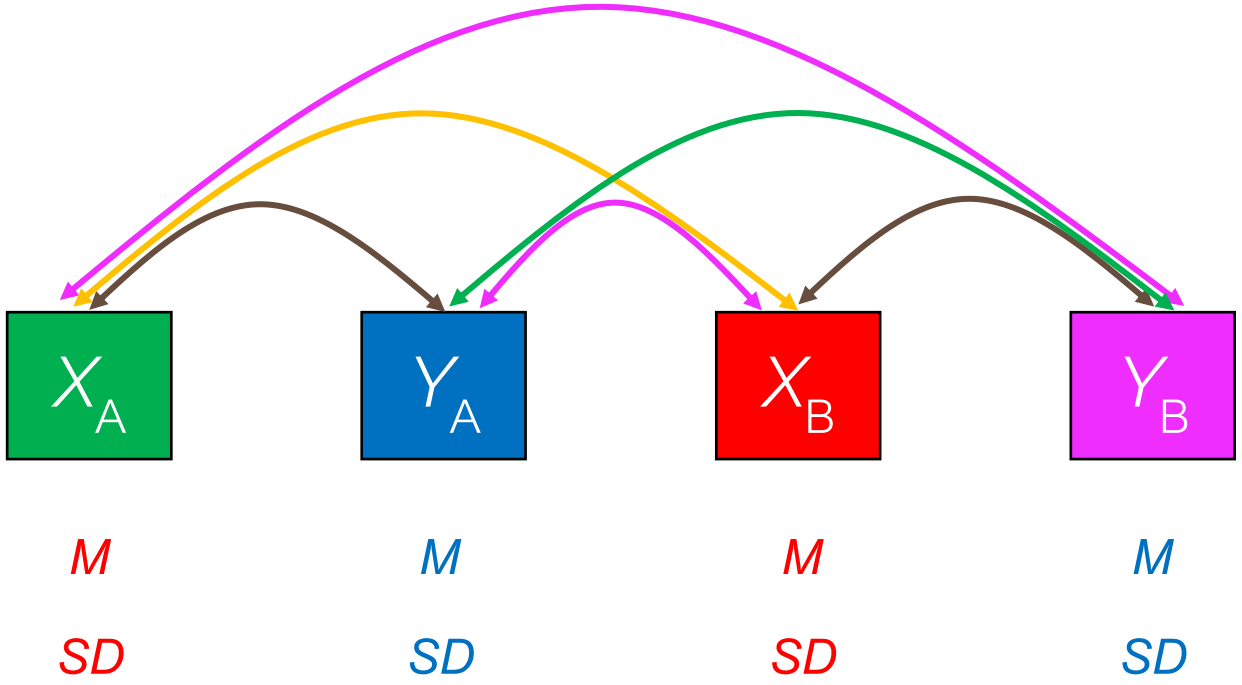
ได้ข้อมูลแบบผสมคู่

```
> with(dat2a, cor(extro, extropartner))
[1] 0.328148
> with(dat2a, cor(neuro, neuropartner))
[1] -0.463747
```

หา ICC ซึ่งได้ค่าตรงกับผลที่คำนวณจากคำสั่ง `gls`

การแบ่งแยกได้ (Distinguishability)

- ให้ A และ B เป็นค่าของตัวแปรที่กำหนดคู่ที่แบ่งแยกได้ และให้ X และ Y เป็นตัวแปรที่วัดจากสมาชิกทั้งสอง <u>คู่ที่แบ่งแยกได้ต้องมีข้อใดข้อหนึ่งด้านล่างเป็นเท็จ</u>
  - $E(X_A) = E(X_B)$ และ $E(Y_A) = E(Y_B)$
  - $Var(X_A) = Var(X_B)$ และ $Var(Y_A) = Var(Y_B)$
  - $Cor(X_A, Y_A) = Cor(X_B, Y_B)$
  - $Cor(X_A, Y_B) = Cor(X_B, Y_A)$
- หากทุกข้อด้านบนเป็นจริง แสดงว่าตัวแปรที่กำหนดคู่ที่แบ่งแยกได้ไม่มีผลในการแบ่งแยก ทำให้คู่เป็นแบบแบ่งแยกไม่ได้
- ตัวอย่างการทดสอบ เช่น แฝดพี่ กับแฝดน้อง สามารถใช้เป็นตัวแปรแบ่งแยกได้หรือไม่?



$Cor(X_A, X_B)$ และ $Cor(Y_A, Y_B)$ จะเท่ากับ ICC เมื่อเป็นคู่ที่แบ่งแยกได้เท่านั้น หากเป็นคู่ที่แบ่งแยกไม่ได้ ต้องจัดเรียงข้อมูลเป็นแบบผสมคู่ (Pairwise) ก่อน

การแบ่งแยกได้ (Distinguishability)

- จากข้อมูลข้างต้น สามารถทดสอบความเท่ากันของค่าพารามิเตอร์ด้วย `lavaan`
- การใส่ข้อจำกัดให้ค่าพารามิเตอร์มีค่าเท่ากัน สามารถทำได้โดยใส่ label เดียวกัน ในพารามิเตอร์เหล่านั้น เช่น

```
mendistress ~ m1*1
womendistress ~ m1*1
```

- ในที่นี้ คือให้ค่าเฉลี่ยของ `mendistress` และ `womendistress` มีค่าเท่ากัน โดยใส่ label เดียวกันว่า `m1`

```
> library(lavaan)
> dat1 <-read.table("lecture12ex1.csv", sep=",", header=TRUE)
> scriptparent <- '
+ # Mean
+ mendistress ~ 1
+ womendistress ~ 1
+ menneuro ~ 1
+ womenneuro ~ 1
+
+ # Variance
+ mendistress ~~ mendistress
+ womendistress ~~ womendistress
+ menneuro ~~ menneuro
+ womenneuro ~~ womenneuro
+
+ # Covariances
+ mendistress ~~ womendistress
+ menneuro ~~ womenneuro
+ mendistress ~~ menneuro
+ womendistress ~~ womenneuro
+ mendistress ~~ womenneuro
+ womendistress ~~ menneuro
+ '
```

```
> scriptnested <- '
+ # Mean
+ mendistress ~ m1*1
+ womendistress ~ m1*1
+ menneuro ~ m2*1
+ womenneuro ~ m2*1
+
+ # Variance
+ mendistress ~~ v1*mendistress
+ womendistress ~~ v1*womendistress
+ menneuro ~~ v2*menneuro
+ womenneuro ~~ v2*womenneuro
+
+ # Covariances
+ mendistress ~~ dis*womendistress
+ menneuro ~~ neuro*womenneuro
+ mendistress ~~ same*menneuro
+ womendistress ~~ same*womenneuro
+ mendistress ~~ diff*womenneuro
+ womendistress ~~ diff*menneuro
+ '
```

การแบ่งแยกได้ (Distinguishability)

- ประมาณค่าโมเดลทั้งสอง แล้วเปรียบเทียบด้วย Likelihood Ratio Test

```
> outparent <- sem(scriptparent, data=dat1)
> outnested <- sem(scriptnested, data=dat1)
> anova(outnested, outparent)
Chi-Squared Difference Test

          Df    AIC    BIC  Chisq Chisq diff Df diff Pr(>Chisq)
outparent  0 6941.8 6991.1  0.000
outnested  6 7023.4 7051.5 93.601     93.601       6  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

ทั้งสองโมเดลแตกต่างกันอย่างมีนัยสำคัญ แสดงว่าตัวแปร ผู้ชาย-ผู้หญิง สามารถแบ่งแยกบทบาทของสมาชิกแต่ละคนภายในคู่ได้

```
> summary(outparent, standardize=TRUE)
lavaan 0.6-9 ended normally after 94 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                        14

  Number of observations                           250

Model Test User Model:

  Test statistic                                 0.000
  Degrees of freedom                                 0

Parameter Estimates:

  Standard errors                             Standard
  Information                                 Expected
  Information saturated (h1) model          Structured

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  mendistress ~~
    womendistress    62.025    6.427    9.650    0.000   62.025    0.770
  menneuro ~~
    womenneuro       -6.731    6.639   -1.014    0.311   -6.731   -0.064
  mendistress ~~
    menneuro         39.720    6.478    6.132    0.000   39.720    0.421
  womendistress ~~
    womenneuro       41.093    6.219    6.608    0.000   41.093    0.460
  mendistress ~~
    womenneuro       53.721    6.856    7.836    0.000   53.721    0.571
  womendistress ~~
    menneuro         13.432    5.729    2.345    0.019   13.432    0.150

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    mendistress      49.779    0.583   85.448    0.000   49.779    5.404
    womendistress    50.400    0.553   91.181    0.000   50.400    5.767
    menneuro         50.830    0.648   78.413    0.000   50.830    4.959
    womenneuro       50.001    0.646   77.349    0.000   50.001    4.892

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    mendistress      84.844    7.589   11.180    0.000   84.844    1.000
    womendistress    76.381    6.832   11.180    0.000   76.381    1.000
    menneuro        105.053    9.396   11.180    0.000  105.053    1.000
    womenneuro      104.469    9.344   11.180    0.000  104.469    1.000
```

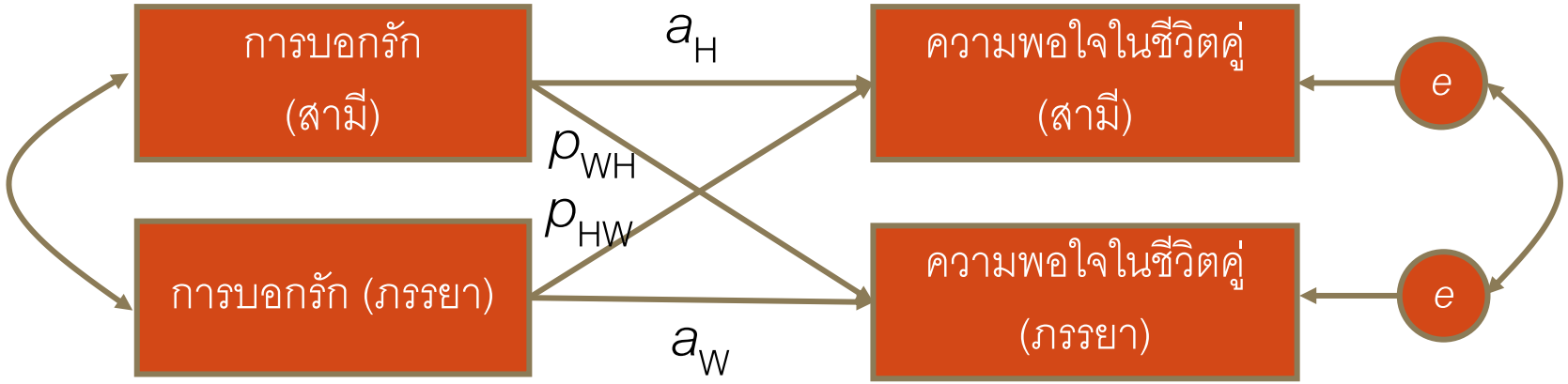
สหสัมพันธ์

ค่าเฉลี่ย

ความแปรปรวน

โมเดลความยึดโยงระหว่างผู้แสดงและคู่

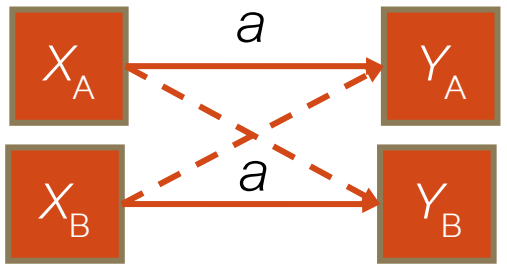
- โมเดลความยึดโยงระหว่างผู้แสดงและคู่ (Actor-Partner Interdependence Model; APIM) เป็นโมเดลที่แสดงอิทธิพลของตัวแปรที่มีซึ่งกันและกัน
- ยกตัวอย่างคู่ที่แบ่งแยกได้ ในคู่สามี-ภรรยา โดยหาอิทธิพลจากการบอกรัก ที่มีต่อความพึงพอใจในชีวิตคู่



โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- อิทธิพลจะแบ่งออกเป็น 2 ประเภท คือ
  - อิทธิพลของผู้แสดง (Actor Effect) อิทธิพลของตัวแปร ที่เกิดขึ้นภายในบุคคล ควบคุมอิทธิพลของคู่
    - a_H คือ อิทธิพลของการบอกรักของสามี ที่มีต่อความพึงพอใจในชีวิตคู่ของสามีเอง ควบคุมอิทธิพลการบอกรักของภรรยาให้คงที่
    - a_W คือ อิทธิพลของการบอกรักของภรรยา ที่มีต่อความพึงพอใจในชีวิตคู่ของภรรยาเอง ควบคุมอิทธิพลการบอกรักของสามีให้คงที่
  - อิทธิพลของคู่ (Partner Effect)
    - p_{HW} คือ อิทธิพลของการบอกรักของภรรยา ที่มีต่อความพึงพอใจในชีวิตคู่ของสามี ควบคุมอิทธิพลการบอกรักของสามีให้คงที่
    - p_{WH} คือ อิทธิพลของการบอกรักของสามี ที่มีต่อความพึงพอใจในชีวิตคู่ของภรรยาเอง ควบคุมอิทธิพลการบอกรักของภรรยาให้คงที่

ผู้แสดงเป็นหลัก (Actor-oriented)

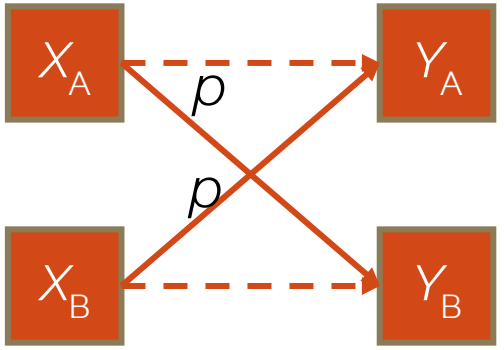


$a \neq 0$

$p = 0$

เช่น การประเมินผลได้เสียจากการมีความสัมพันธ์ต่อไปต่อเจตนาการคงความสัมพันธ์ (Social Exchange Model)

คู่เป็นหลัก (Partner-oriented)

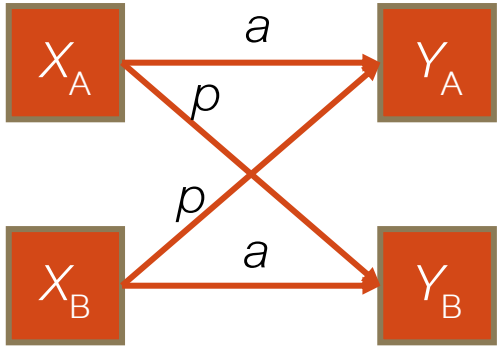


$a = 0$

$p \neq 0$

เช่น ความน่าดึงดูด ต่อความพึงพอใจในความสัมพันธ์

คนทั้งสองเป็นหลัก (Couple-oriented)



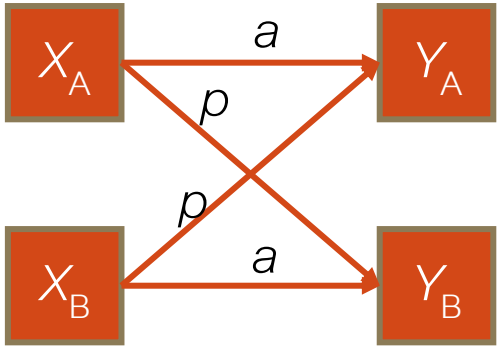
$a = p$

$a \neq 0$

$p \neq 0$

เกิดขึ้นเมื่อสนใจผลของทั้งตนเองและคู่ของตน เช่น ผลงานการเล่นเทนนิสคู่ต่อความพึงพอใจจากการเล่น

เปรียบเทียบทางสังคม (Social Comparison)



$a + p = 0$

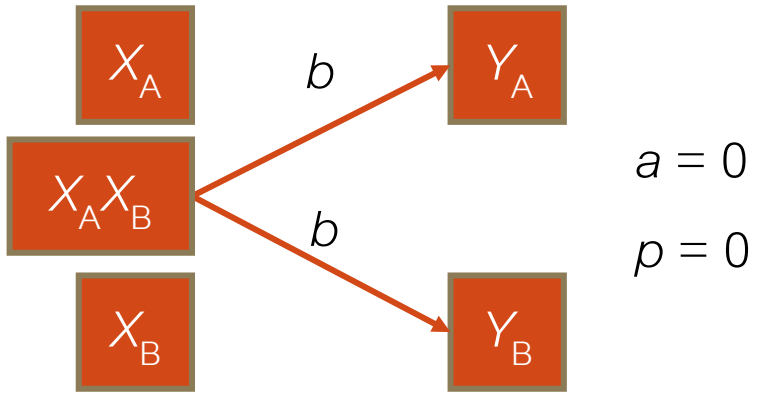
$a \neq 0$

$p \neq 0$

เกิดจากการเปรียบเทียบตนเองกับคู่ เช่น ผลงานของคู่แข่งกีฬา ที่มีผลต่อความพึงพอใจในผลการแข่ง

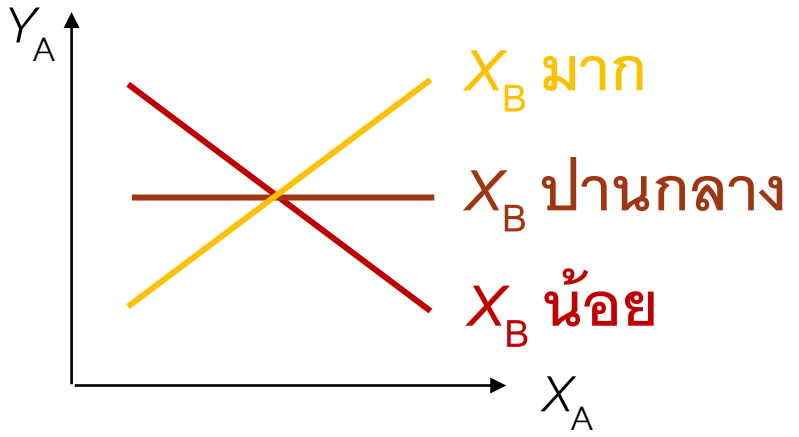
โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- ยิ่งอิทธิพลจากคู่ (p) ยิ่งสูง ยิ่งแสดงว่าผู้แสดงและคู่ยึดโยงกันมาก
- ตัวอย่างที่ผ่านมา ตัวแปรต้นจากทั้งสองคน (X_A, X_B) มีผลเชิงเส้นต่อตัวแปรตามของทั้งสองคน (Y_A, Y_B) แต่ตัวแปรต้นอาจมีปฏิสัมพันธ์กันได้ เช่น เจตคติทางการเมือง ที่มีผลต่อความพึงพอใจในความสัมพันธ์ ถ้าเจตคติทั้งสองคนเหมือนกัน จะมีผลต่อความพึงพอใจ



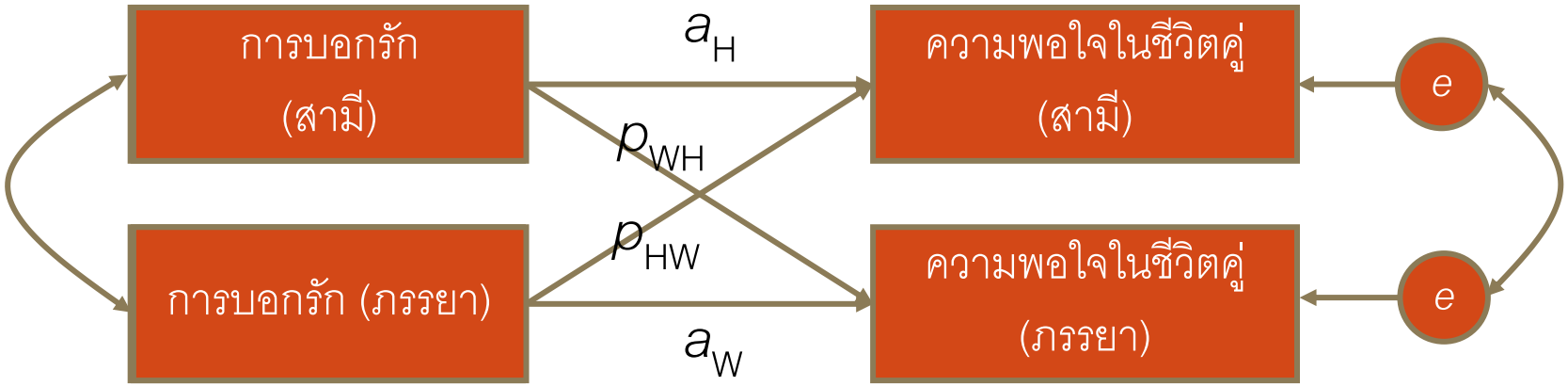
$a = 0$

$p = 0$



โมเดลความยึดโยงระหว่างผู้แสดงและคู่

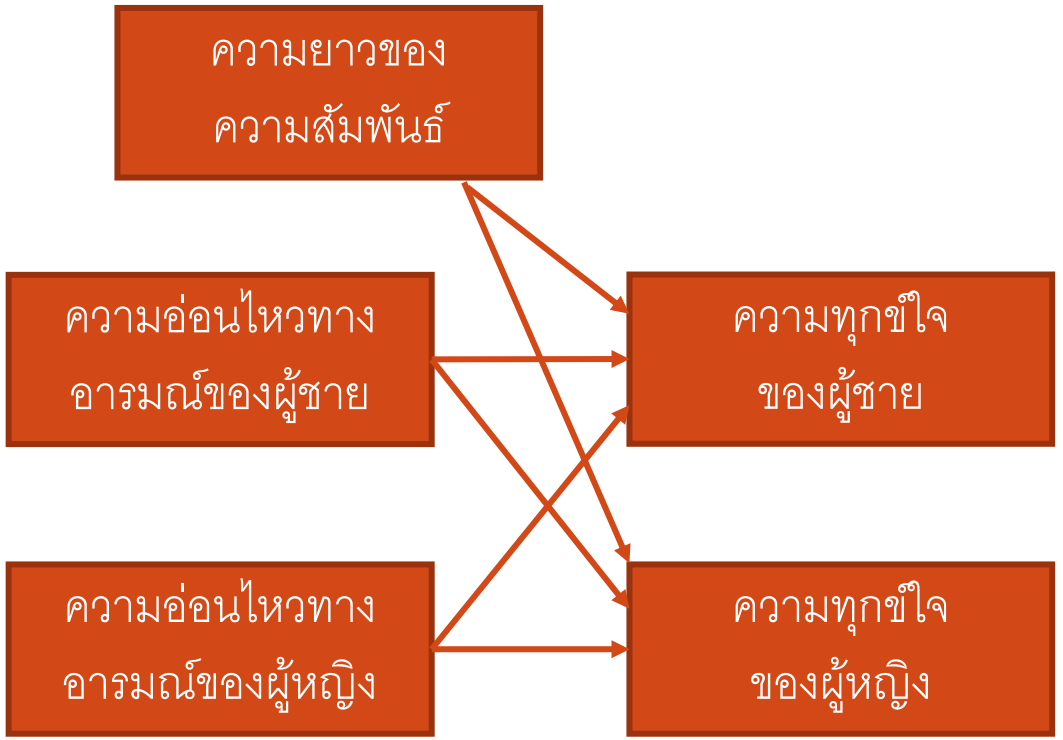
- สำหรับคู่ที่แบ่งแยกได้ แม้จะวิเคราะห์ด้วย MLM ได้ วิธีที่ง่ายที่สุดคือใช้ SEM
- วิธีการวิเคราะห์ สามารถทำได้ตามรูปความสัมพันธ์นี้ได้เลย



- สามารถใส่ตัวแปรผสมเพิ่มในโมเดลได้ (เพิ่มสี่เหลี่ยมของแต่ละบทบาท) หรือใส่ตัวแปรคู่ในโมเดลได้ (ซึ่งจะเป็นสี่เหลี่ยมกล่องเดียว)

โมเดลความยึดโยงระหว่างผู้แสดงและคู่

เก็บข้อมูลคู่ชีวิต 250 คู่
โดยวัดอิทธิพลต่อความอ่อนไหวทางอารมณ์ (Neuroticism) ของแต่ละคนที่มีต่อความทุกข์ใจ (Distress) ของแต่ละคน



```
> library(lavaan)
> script1 <- '
+ mendistress ~ menneuro + womenneuro + lengthrela
+ womendistress ~ menneuro + womenneuro + lengthrela
+ mendistress ~~ womendistress
+ '
> out1 <- sem(script1, data=dat1)
> summary(out1, standardize=TRUE, rsquare=TRUE)
```

ใส่สมการทำนายตัวแปรตามแต่ละตัว

ให้ค่าคงเหลือของตัวแปรตามทั้งสองสัมพันธ์กัน

สัมประสิทธิ์ถดถอย

p value

```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  mendistress ~
    menneuro          0.357    0.035   10.196    0.000    0.357    0.398
    womenneuro        0.451    0.036   12.456    0.000    0.451    0.500
    lengthrela       -1.850    0.224   -8.260    0.000   -1.850   -0.336
  womendistress ~
    menneuro          0.085    0.042    2.014    0.044    0.085    0.099
    womenneuro        0.291    0.043    6.697    0.000    0.291    0.341
    lengthrela       -2.300    0.269   -8.544    0.000   -2.300   -0.440

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
 .mendistress ~~
   .womendistress    23.757    2.791    8.512    0.000   23.757    0.639

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .mendistress      30.949    2.768   11.180    0.000   30.949    0.365
   .womendistress    44.694    3.998   11.180    0.000   44.694    0.585

R-Square:
                   Estimate
    mendistress       0.635
    womendistress     0.415
```

สัมประสิทธิ์ถดถอยมาตรฐาน (β)

ค่าสหสัมพันธ์ของความผิดพลาดในการทำนาย

ความแปรปรวนร่วม

p value

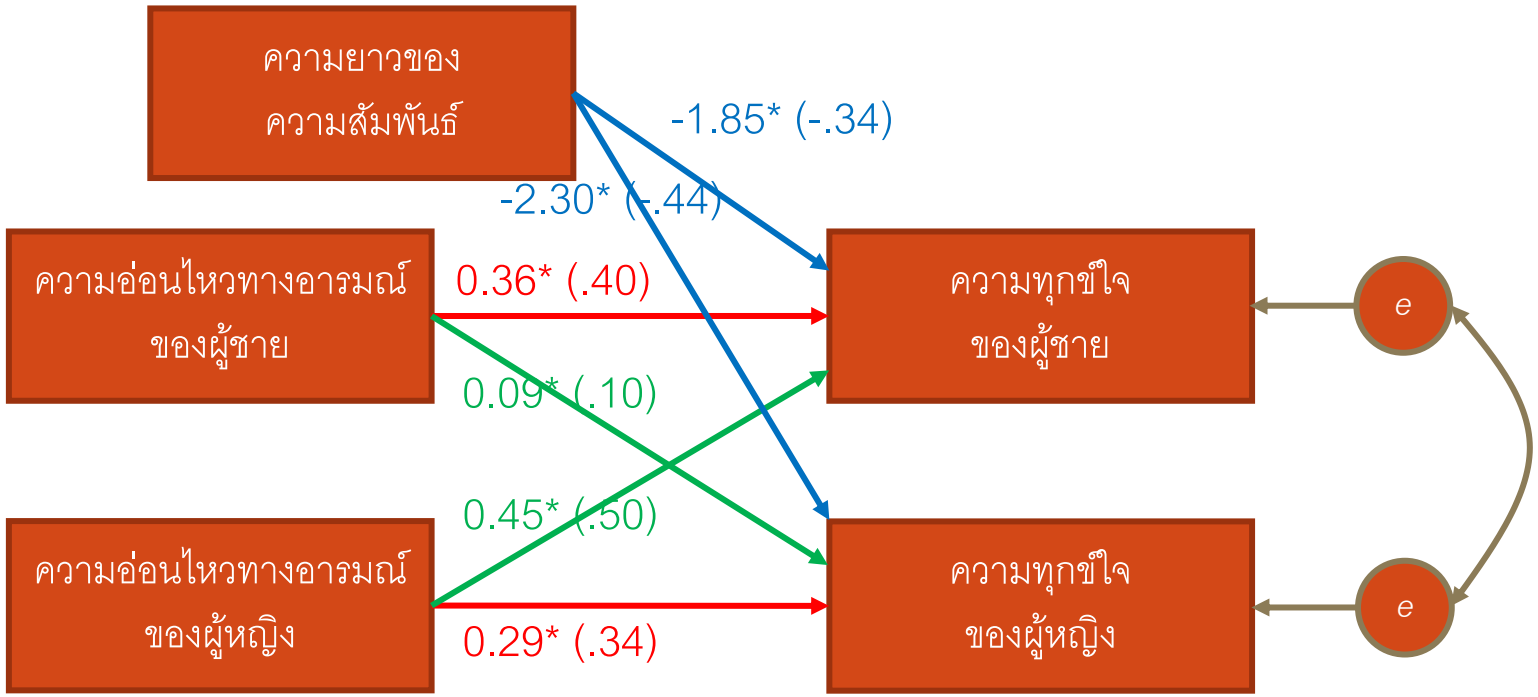
สัมประสิทธิ์การทำนายตัวแปรตามแต่ละตัว

```
Regressions:
                   Estimate  Std.all
  mendistress ~
    menneuro          0.357    0.398
    womenneuro        0.451    0.500
    lengthrela       -1.850   -0.336
  womendistress ~
    menneuro          0.085    0.099
    womenneuro        0.291    0.341
    lengthrela       -2.300   -0.440

Covariances:
                   Estimate  Std.all
 .mendistress ~~
   .womendistress    23.757    0.639
```

ทั้งผู้ชายและผู้หญิง จะมีอิทธิพลของผู้แสดง (Actor Effect) ระดับที่ใกล้เคียงกัน

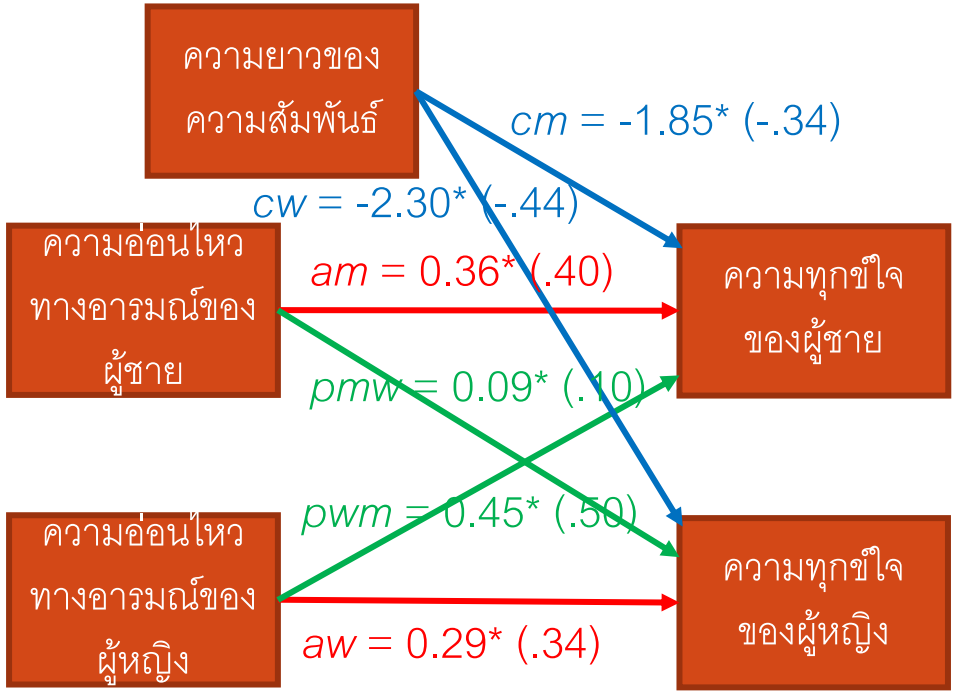
อิทธิพลคู่ (Partner Effect) จากผู้หญิงไปหาผู้ชาย มีค่าสูง เมื่อเทียบจากผู้ชายไปหาผู้หญิง

ยิ่งความยาวของความสัมพันธ์สูง ความทุกข์ใจของทั้งชายและหญิงจะน้อยลง โดยเหมือนมีผลกับความทุกข์ใจของผู้หญิงมากกว่า



โมเดลความยืดโยงระหว่างผู้แสดงและคู่

- เพื่อความสะดวกในการเปรียบเทียบสัมประสิทธิ์ถดถอย `lavaan` สามารถให้ label กับพารามิเตอร์แต่ละตัว หลังจากนั้นนำค่ามาบวกลบกันได้

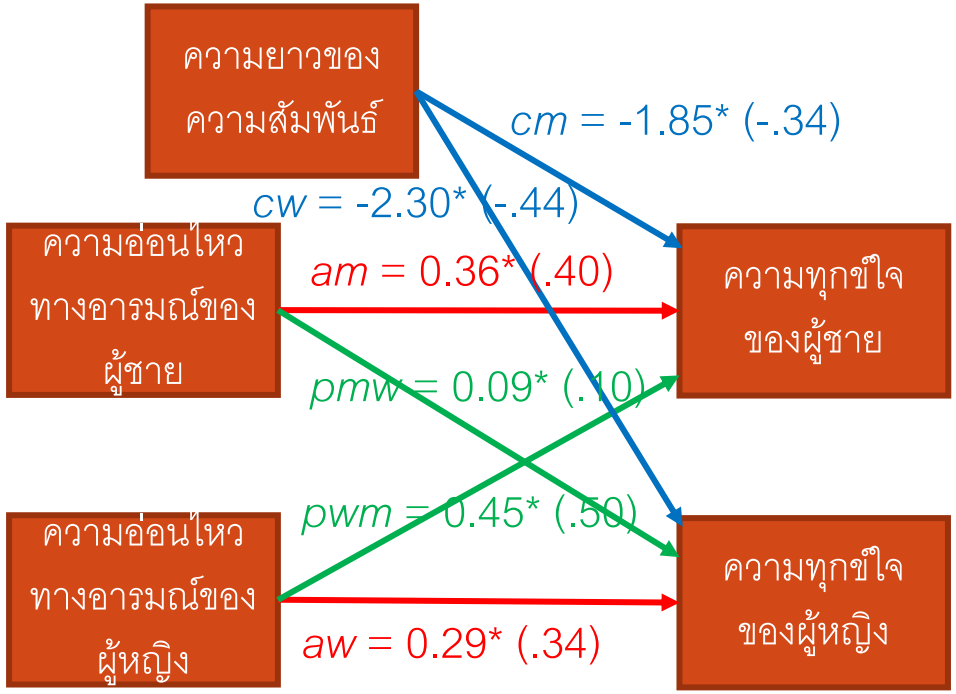


การเปรียบเทียบที่ทำได้

- เปรียบเทียบว่าในการทำนายผู้ชาย ผลของผู้แสดงหรือคู่มากกว่ากัน : *am* – *pwm*
- เปรียบเทียบว่าในการทำนายผู้หญิง ผลของผู้แสดงหรือคู่มากกว่ากัน : *aw* – *pmw*
- เปรียบเทียบผลของผู้แสดงเทียบผู้ชายและหญิง : *am* – *aw*
- เปรียบเทียบผลของคู่เทียบในการทำนายผู้ชายและหญิง : *pmw* – *pwm*

โมเดลความยืดโยงระหว่างผู้แสดงและคู่

- เพื่อความสะดวกในการเปรียบเทียบสัมประสิทธิ์ถดถอย `lavaan` สามารถให้ label กับพารามิเตอร์แต่ละตัว หลังจากนั้นนำค่ามาบวกลบกันได้

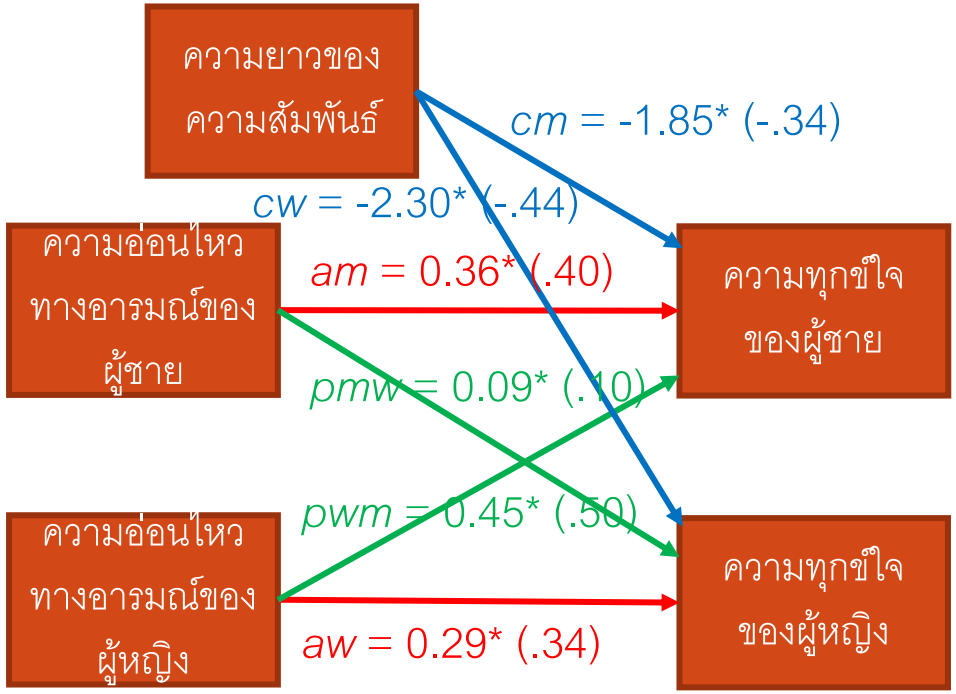


การเปรียบเทียบที่ทำได้

- เปรียบเทียบว่าความอ่อนไหวทางอารมณ์ของผู้ชายทำนายความทุกข์ใจในชายหรือหญิงได้ดีกว่ากัน : *am* - *pmw*
- เปรียบเทียบว่าความอ่อนไหวทางอารมณ์ของผู้หญิงทำนายความทุกข์ใจในชายหรือหญิงได้ดีกว่ากัน : *pwm* - *aw*
- เปรียบเทียบว่าความยาวของความสัมพันธ์ทำนายความทุกข์ใจในชายหรือหญิงได้ดีกว่ากัน : *cm* - *cw*

โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- เพื่อความสะดวกในการเปรียบเทียบสัมประสิทธิ์ถดถอย `lavaan` สามารถให้ label กับพารามิเตอร์แต่ละตัว หลังจากนั้นนำค่ามาบวกลบกันได้



ค่าพารามิเตอร์เพิ่มเติมที่ทำได้

- หาค่าอิทธิพลผู้แสดงเฉลี่ย : (*am* + *aw*) / 2
- หาค่าอิทธิพลคู่เฉลี่ย : (*pmw* + *pwm*) / 2
- เปรียบเทียบอิทธิพลผู้แสดงและอิทธิพลคู่เฉลี่ยว่าอิทธิพลใดมีค่ามากกว่ากัน :

  ((*am* + *aw*) / 2) – ((*pmw* + *pwm*) / 2)

```
> script2 <- '
+ mendistress ~ am*menneuro + pmw*womenneuro + cm*lengthrela
+ womendistress ~ pwm*menneuro + aw*womenneuro + cw*lengthrela
+ mendistress ~~ womendistress
+
+ # Define New Parameters
+ apmen := am - pmw                          # Compare actor and partner on mendistress
+ apwomen := aw - pwm                        # Compare actor and partner on womendistress
+ diffa := am - aw                           # Compare actor effects on each role
+ diffp := pmw - pwm                         #  cts on each role
+ diffmenneuro := am - pwm                   #  on menneuro on men and women distress
+ diffwomenneuro := pmw - aw                 #  on womenneuro on men and women distress
+ difflengthrela := cm - cw                  #  on lengthrela on men and women distress
+ aveactor := (am + aw) / 2                  #  s
+ avepartner := (pmw + pwm) / 2              # Average partner effects
+ diffactorpartner := aveactor - avepartner # Compare actor and partner effects
+ '
> out2 <- sem(script2, data=dat1)
> summary(out2)

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  mendistress ~
    menneuro  (am)    0.357    0.035   10.196    0.000
    womenner (pmw)    0.451    0.036   12.456    0.000
    lengthrl  (cm)   -1.850    0.224   -8.260    0.000
  womendistress ~
    menneuro (pwm)    0.085    0.042    2.014    0.044
    womenner  (aw)    0.291    0.043    6.697    0.000
    lengthrl  (cw)   -2.300    0.269   -8.544    0.000
```

ใส่ label ลงพารามิเตอร์แต่ละตัว

สร้างพารามิเตอร์ใหม่จากค่าพารามิเตอร์เดิม

มี label หลังค่าพารามิเตอร์

Defined Parameters:

	Estimate	Std.Err	z-value	P(>|z|)
apmen	-0.093	0.047	-1.972	0.049
apwomen	0.206	0.057	3.629	0.000
diffa	0.066	0.054	1.231	0.218
diffp	0.366	0.053	6.848	0.000
diffmenneuro	0.273	0.033	8.156	0.000
diffwomenneuro	0.159	0.034	4.624	0.000
difflengthrela	0.450	0.214	2.106	0.035
aveactor	0.324	0.029	11.208	0.000
avepartner	0.268	0.029	9.304	0.000
diffactorprtnr	0.057	0.023	2.506	0.012

apmen := *am* – *aw* อิทธิพลจากความอ่อนไหวทางอารมณ์ของตัวผู้ชายเองมีอิทธิพลต่อความทุกข์ใจน้อยกว่าความอ่อนไหวทางอารมณ์จากผู้หญิงอย่างมีนัยสำคัญ (Partner Effect > Actor Effect in Males)

apwomen := *aw* – *pwm* อิทธิพลจากความอ่อนไหวทางอารมณ์ของตัวผู้หญิงเองมีอิทธิพลต่อความทุกข์ใจมากกว่าความอ่อนไหวทางอารมณ์จากผู้ชายอย่างมีนัยสำคัญ (Actor Effect > Partner Effect in Females)

Defined Parameters:

	Estimate	Std.Err	z-value	P(>|z|)
apmen	-0.093	0.047	-1.972	0.049
apwomen	0.206	0.057	3.629	0.000
diffa	0.066	0.054	1.231	0.218
diffp	0.366	0.053	6.848	0.000
diffmenneuro	0.273	0.033	8.156	0.000
diffwomenneuro	0.159	0.034	4.624	0.000
difflengthrela	0.450	0.214	2.106	0.035
aveactor	0.324	0.029	11.208	0.000
avepartner	0.268	0.029	9.304	0.000
diffactorprtnr	0.057	0.023	2.506	0.012

diffa := *am* - *aw* อิทธิพลผู้แสดงของผู้ชายและผู้หญิงไม่แตกต่างกันอย่างมีนัยสำคัญ

diffp := *pmw* - *pwm* อิทธิพลคู่จากผู้หญิงไปหาผู้ชายมากกว่าอิทธิพลคู่จากผู้ชายไปหาผู้หญิงอย่างมีนัยสำคัญ

diffmenneuro := *am* - *pwm* อิทธิพลความอ่อนไหวทางอารมณ์ของผู้ชายมีผลต่อความทุกข์ใจของผู้ชายมากกว่าความทุกข์ใจของผู้หญิง

diffwomenneuro := *pmw* – *aw* อิทธิพลความอ่อนไหวทางอารมณ์ของผู้หญิงมีผลต่อความทุกข์ใจของผู้ชายมากกว่าความทุกข์ใจของผู้หญิง

Defined Parameters:

	Estimate	Std.Err	z-value	P(>|z|)
apmen	-0.093	0.047	-1.972	0.049
apwomen	0.206	0.057	3.629	0.000
diffa	0.066	0.054	1.231	0.218
diffp	0.366	0.053	6.848	0.000
diffmenneuro	0.273	0.033	8.156	0.000
diffwomenneuro	0.159	0.034	4.624	0.000
difflengthrela	0.450	0.214	2.106	0.035
aveactor	0.324	0.029	11.208	0.000
avepartner	0.268	0.029	9.304	0.000
diffactorprtnr	0.057	0.023	2.506	0.012

difflengthrela := *cm* - *cw* อิทธิพลของความยาวของความสัมพันธ์มีผลทางลบในความทุกข์ใจของผู้หญิงมากกว่าความทุกข์ใจของผู้ชายอย่างมีนัยสำคัญ

aveactor := (*am* + *aw*) / 2 อิทธิพลผู้แสดงเฉลี่ยเท่ากับ 0.324 ซึ่งต่างจากศูนย์อย่างมีนัยสำคัญ

avepartner := (*pmw* + *pwm*) / 2 อิทธิพลคู่เฉลี่ยเท่ากับ 0.268 ซึ่งต่างจากศูนย์อย่างมีนัยสำคัญ

diffactorpartner := *aveactor* – *avepartner* อิทธิพลผู้แสดงเฉลี่ยมีค่ามากกว่าอิทธิพลคู่เฉลี่ยอย่างมีนัยสำคัญ

โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- สำหรับคู่ที่แบ่งแยกไม่ได้ สามารถหา APIM ได้ด้วย MLM แต่ต้องใช้ข้อมูลผสมคู่ (Pairwise) เพื่อให้ได้คะแนนของคู่มาทำนายตัวแปรที่ต้องการ

		Y	X_A		X_P
คู่	สมาชิก	ความเชื่อใจตนเอง	ความขี้โมโหตนเอง	ความเชื่อใจคู่	ความขี้โมโหคู่
1	1	45	55	42	64
1	2	42	64	45	55
2	1	56	46	58	54
2	2	58	54	56	46
3	1	46	46	40	37
3	2	40	37	46	46

เช่น ทำนายความรู้สึกเชื่อใจต่อคู่แฝดของตน (Y)
ความรู้สึกนี้สามารถทำนายได้จาก

- ความขี้โมโหของตนเอง (Actor Effect: X_A)
- ความขี้โมโหของคู่แฝดของตน (Partner Effect : X_P)

โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- ให้ Y_{ij} เป็นตัวแปรตามของสมาชิกที่ *i* ในคู่ที่ *j*, ให้ X_{ij} เป็นค่าตัวแปรอิสระของสมาชิกคนที่ *i* ในคู่ที่ *j*, และ $X_{i'j}$ เป็นค่าตัวแปรอิสระของสมาชิกอีกคนหนึ่งในคู่ที่ *j* (คู่หนึ่งมีสมาชิก 2 คน) โมเดล MLM จะเป็นดังนี้

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \beta_{2j}X_{i'j} + e_{ij}$$ ระดับที่ 1

$$\beta_{0j} = \gamma_{00} + u_{0j}$$ ระดับที่ 2

$\beta_{1j} = \gamma_{10}$ อิทธิพลของผู้แสดง (Actor Effect)

$\beta_{2j} = \gamma_{20}$ อิทธิพลของคู่ (Partner Effect)

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{20}X_{i'j} + u_{0j} + e_{ij}$$ สมการรวม

โมเดลความยืดโยงระหว่างผู้แสดงและคู่

- ในสมการนี้จัดการความสัมพันธ์ภายในคู่ด้วยค่าคงเหลือระดับคู่ u_{0j} แต่จากที่กล่าวไปข้างค้น ค่าคงเหลือระดับคู่อาจจัดการด้วยค่าสหสัมพันธ์ระหว่างค่าคงเหลือภายในคู่แทน ซึ่งสามารถจัดการ ICC ทางลบได้
- โมเดล APIM ที่ใส่ค่าตัวแปรต้นของตนเองและคู่ของตนเอง แท้จริงแล้วเป็นเพียงการเปลี่ยนรูปการประมาณค่าพารามิเตอร์ (Reparameterization) ของโมเดลที่ใส่ตัวแปรระดับที่ 1 (ทั้งแบบย้ายและไม่ย้ายศูนย์กลางที่ค่าเฉลี่ยกลุ่ม) และค่าเฉลี่ยกลุ่ม

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}\bar{X}_{.j} + u_{0j} + e_{ij}$$

ไม่ย้ายศูนย์กลางไป
ที่ค่าเฉลี่ยกลุ่ม

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}\left(\frac{X_{ij} + X_{i'j}}{2}\right) + u_{0j} + e_{ij}$$

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \frac{\gamma_{01}}{2}X_{ij} + X_{i'j} + u_{0j} + e_{ij}$$

$$Y_{ij} = \gamma_{00} + \left(\gamma_{10} + \frac{\gamma_{01}}{2}\right)X_{ij} + \frac{\gamma_{01}}{2}X_{i'j} + u_{0j} + e_{ij}$$

$$Y_{ij} = \gamma_{00} + \gamma_{10}(X_{ij} - \bar{X}_{.j}) + \gamma_{01}\bar{X}_{.j} + u_{0j} + e_{ij}$$

ย้ายศูนย์กลางไป
ที่ค่าเฉลี่ยกลุ่ม

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + (\gamma_{01} - \gamma_{10})\bar{X}_{.j} + u_{0j} + e_{ij}$$

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + (\gamma_{01} - \gamma_{10})\left(\frac{X_{ij} + X_{i'j}}{2}\right) + u_{0j} + e_{ij}$$

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \left(\frac{\gamma_{01} - \gamma_{10}}{2}\right)X_{ij} + \left(\frac{\gamma_{01} - \gamma_{10}}{2}\right)X_{i'j} + u_{0j} + e_{ij}$$

$$Y_{ij} = \gamma_{00} + \left(\frac{\gamma_{10} + \gamma_{01}}{2}\right)X_{ij} + \left(\frac{\gamma_{01} - \gamma_{10}}{2}\right)X_{i'j} + u_{0j} + e_{ij}$$

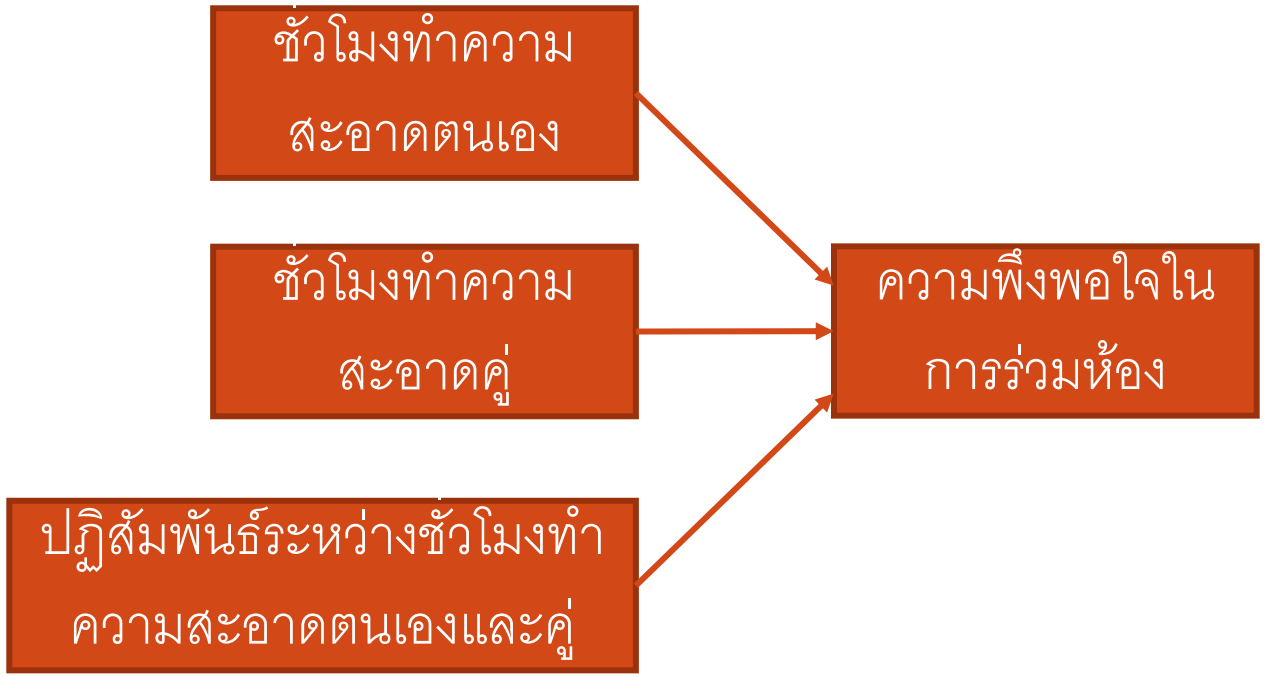
โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- อิทธิพลผู้แสดงและอิทธิพลคู่ จะไม่สามารถมีความชันสุ่มได้ เนื่องจากมีสมาชิกภายในกลุ่มแค่ 2 คน ทำให้ประมาณค่าอิทธิพลสุ่ม (Random Effect) ได้เพียงเดียว คือ จุดตัดสุ่ม (Random Intercept) ซึ่งก็คือความสัมพันธ์ของค่าคงเหลือของสมาชิกในคู่เดียวกัน
- อย่างไรก็ตาม โมเดลยังมีปฏิสัมพันธ์ระหว่างชั้นได้อยู่ (แม้ไม่มีความชันสุ่ม) เช่น ความยาวของความสัมพันธ์ (W_{1j}) ไปกำกับอิทธิพลผู้แสดงและอิทธิพลคู่

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_{1j} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_{1j}$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21}W_{1j}$$

โมเดลความยึดโยงระหว่างผู้แสดงและคู่

เก็บข้อมูลเพื่อนร่วมห้อง 500 คู่ โดยวัดชั่วโมงการทำความสะอาดห้องต่อความพึงพอใจในการใช้ชีวิตร่วมห้องของแต่ละคน



โมเดลความยืดโยงระหว่างผู้แสดงและคู่

ID คู่ ID สมาชิก

```
> dat3 <- read.table("lecture12ex3.csv", sep=",", header=TRUE)
> psych::describe(dat3)
          vars    n   mean     sd median trimmed    mad min    max range  skew kurtosis   se
coupleid     1 1000 250.50 144.41  250.5  250.50 185.32   1  500.0 499.0  0.00    -1.20 4.57
subjectid    2 1000 500.50 288.82  500.5  500.50 370.65   1 1000.0 999.0  0.00    -1.20 9.13
house        3 1000   3.26   1.00    3.2    3.24   1.04   0    6.3   6.3  0.15    -0.14 0.03
sat          4 1000  50.02  12.70   50.0   50.06  13.34   6   89.0  83.0 -0.05    -0.14 0.40
```

ชั่วโมงการทำงานบ้านต่อเดือน ความพึงพอใจในการใช้ชีวิตร่วมห้อง

```
> dat3a <- dat3[order(dat3$coupleid, dat3$subjectid),]
> dat3b <- dat3[order(dat3$coupleid, -dat3$subjectid),]
> dat3a$housepartner <- dat3b$house
```

สร้างข้อมูลผสมคู่ ที่ตัวแปรใหม่ชื่อ `housepartner` คือชั่วโมงการทำงานบ้านของคู่

ยังไม่มีปฏิสัมพันธ์

```
> out3a <- gls(sat ~ 1 + house + housepartner,
+              data = dat3a, correlation = corCompSymm(form = ~1|coupleid))
> summary(out3a)
Generalized least squares fit by REML
  Model: sat ~ 1 + house + housepartner
  Data: dat3a
       AIC      BIC    logLik
  7339.369 7363.893 -3664.685

Correlation Structure: Compound symmetry
 Formula: ~1 | coupleid
 Parameter estimate(s):
      Rho
0.3421043

Coefficients:
                Value Std.Error   t-value p-value
(Intercept)  34.12828 2.1089358 16.182701       0
house        -2.20285 0.3532515 -6.235925       0
housepartner  7.08331 0.3532515 20.051761       0
```

ICC

ความพึงพอใจของการใช้ชีวิตร่วมห้อง ถ้าตนเองและคู่ไม่ทำงานบ้านเลยเท่ากับ 34.13 คะแนน

หากตนเองทำงานบ้านขึ้น 1 ชั่วโมงต่อเดือน ควบคุมการทำงานบ้านของคู่ให้คงที่แล้ว ความพึงพอใจในการใช้ชีวิตร่วมห้องจะลดลง 2.20 คะแนน

หากคู่ทำงานบ้านขึ้น 1 ชั่วโมงต่อเดือน ควบคุมการทำงานบ้านของตนเองให้คงที่แล้ว ความพึงพอใจในการใช้ชีวิตร่วมห้องจะเพิ่มขึ้น 7.08 คะแนน

สามารถเปรียบเทียบอิทธิพลผู้แสดงและอิทธิพลคู่ ว่าค่าใดมีค่ามากกว่ากัน ด้วยการตั้งสัมประสิทธิ์ความแตกต่าง (Contrasts) ได้ คล้ายกับการตรวจสอบปฏิสัมพันธ์

กล่าวคือ เราต้องการทดสอบสมมติฐานหลักว่า $H_0: \gamma_{10} - \gamma_{20} = 0$

```
Coefficients:
               Value Std.Error   t-value p-value
(Intercept)  40.04536 2.0854973  19.20183       0
house        -6.02874 0.3500972 -17.22018       0
housepartner  9.22705 0.3500972  26.35569       0
```

0
1
-1

ชุดของตัวเลขนี้จะเรียกว่า สัมประสิทธิ์ความแตกต่าง (Contrast Coefficients) สามารถใช้ multcomp package ในการทดสอบสมมติฐานรูปแบบนี้ได้

สร้างสัมประสิทธิ์ความแตกต่างเท่ากับ $[0 \quad 1 \quad -1]$

```
> library(multcomp)
> diffhouse <- matrix(c(0, 1, -1), nrow=1, ncol=3)
> ct <- glht(out3a, linfct = diffhouse)
> summary(ct)

         Simultaneous Tests for General Linear Hypotheses

Fit: gls(model = sat ~ 1 + house + housepartner, data = dat3a, correlation = corCompSymm
(form = ~1 |
    coupleid))

Linear Hypotheses:
       Estimate Std. Error z value Pr(>|z|)
1 == 0   -9.286      0.303  -30.65   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Adjusted p values reported -- single-step method)
```

ใช้คำสั่ง `glht` โดยใส่ผลการวิเคราะห์ และสัมประสิทธิ์ความแตกต่าง

ในการทำนายชั่วโมงการทำงานบ้าน ต่อความพึงพอใจในการอยู่ร่วมห้อง อิทธิพลผู้แสดงน้อยกว่าอิทธิพลคู่อย่างมีนัยสำคัญ

มีปฏิสัมพันธ์

```
> out3b <- gls(sat ~ 1 + house*housepartner,
+              data = dat3a, correlation = corCompSymm(form = ~1|coupleid))
> summary(out3b)
Generalized least squares fit by REML
  Model: sat ~ 1 + house * housepartner
  Data: dat3a
       AIC      BIC    logLik
  7331.238 7360.661 -3659.619

Correlation Structure: Compound symmetry
 Formula: ~1 | coupleid
 Parameter estimate(s):
      Rho
0.3337279

Coefficients:
                     Value Std.Error   t-value p-value
(Intercept)       45.53177  4.091386 11.128691  0.0000
house             -5.71393  1.138436 -5.019105  0.0000
housepartner       3.57223  1.138436  3.137843  0.0018
house:housepartner 1.11993  0.345484  3.241623  0.0012
```

ปฏิสัมพันธ์ถึงระดับนัยสำคัญ

$$Y_{ij} = 45.53 - 5.71X_{ij} + 3.57X_{i'j} + 1.12X_{ij}X_{i'j} + u_{0j} + e_{ij}$$

```
> mean(dat3$house)
[1] 3.2566
> sd(dat3$house)
[1] 0.9981647
```

ชั่วโมงทำงานบ้านที่ใช้ตรวจสอบปฏิสัมพันธ์
คือ 2.25, 3.25, และ 4.25 ชั่วโมง

`interactions` package ไม่รองรับการวิเคราะห์จากฟังก์ชั่น `gls`

ใช้คำสั่ง `quick_sim_slopes` ที่ผมเขียนขึ้นมา ในการหาจุดตัดย่อย และความชันย่อย

```
> source("quicksimslopeslme4.R")
> housevalue <- c(2.25, 3.25, 4.25)
> quick_sim_slopes(out3b, pred="house", modx="housepartner", modx.values=housevalue)
$`Conditional Intercept`
  modx.values      est       se        z p
1        2.25 53.56930 1.796044 29.82628 0
2        3.25 57.14153 1.196587 47.75378 0
3        4.25 60.71376 1.493298 40.65749 0

$`Conditional Slope`
  modx.values       est       se         z        p
1        2.25 -3.1940930 0.4651709 -6.866494 0.000000
2        3.25 -2.0741653 0.3527814 -5.879464 0.000000
3        4.25 -0.9542376 0.5208103 -1.832217 0.066919
```

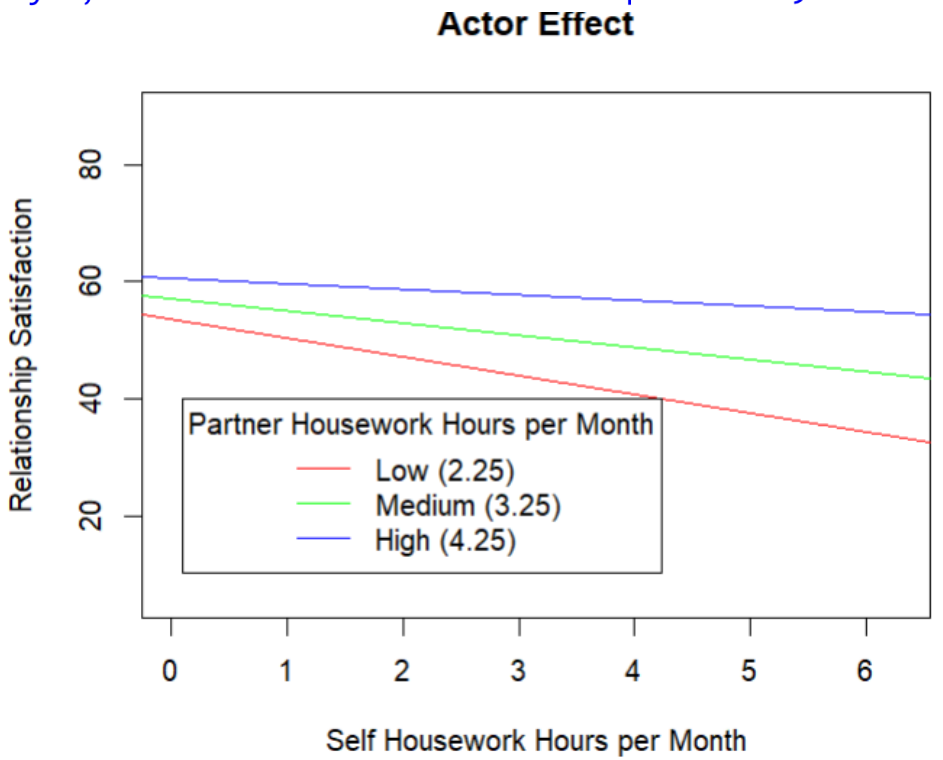
หากคู่ทำงานบ้านน้อย ยิ่งตนเองทำงานบ้านมากขึ้น ยิ่งทำให้ความพึงพอใจในความสัมพันธ์น้อยลง

หากคู่ทำงานบ้านปานกลาง ยิ่งตนเองทำงานบ้านมากขึ้น ยิ่งทำให้ความพึงพอใจในความสัมพันธ์น้อยลง

หากคู่ทำงานบ้านมาก ชั่วโมงการทำงานบ้านของตนเองไม่มีผลต่อความพึงพอใจในความสัมพันธ์อย่างมีนัยสำคัญ

```
> cond_house <- quick_sim_slopes(out3b, pred="house", modx="housepartner", modx.values=housevalue)
> cond_house_int <- cond_house[[1]][,"est"]
> cond_house_slope <- cond_house[[2]][,"est"]
> plot(dat3a$house, dat3a$sat, type="n", xlab="Self Housework Hours per Month",
+      ylab="Relationship Satisfaction", main="Actor Effect")
> abline(a = cond_house_int[1], b = cond_house_slope[1], col="red")
> abline(a = cond_house_int[2], b = cond_house_slope[2], col="green")
> abline(a = cond_house_int[3], b = cond_house_slope[3], col="blue")
> legend(0.1, 40, c("Low (2.25)", "Medium (3.25)", "High (4.25)"),
+        col=c("red", "green", "blue"), lty=1, title="Partner Housework Hours per Month")
```

ยิ่งคู่ทำงานบ้านมากขึ้น อิทธิพลของการทำงานบ้านของตนเองต่อความพึงพอใจในความสัมพันธ์จะเป็นทางลบในระดับที่น้อยลง



```
> quick_sim_slopes(out3b, pred="housepartner", modx="house", modx.values=housevalue)
$`Conditional Intercept`
  modx.values      est       se        z p
1        2.25 32.67542 1.796044 18.19300 0
2        3.25 26.96149 1.196587 22.53200 0
3        4.25 21.24756 1.493298 14.22861 0

$`Conditional Slope`
  modx.values      est       se        z p
1        2.25 6.092072 0.4651709 13.09642 0
2        3.25 7.211999 0.3527814 20.44325 0
3        4.25 8.331927 0.5208103 15.99801 0
```

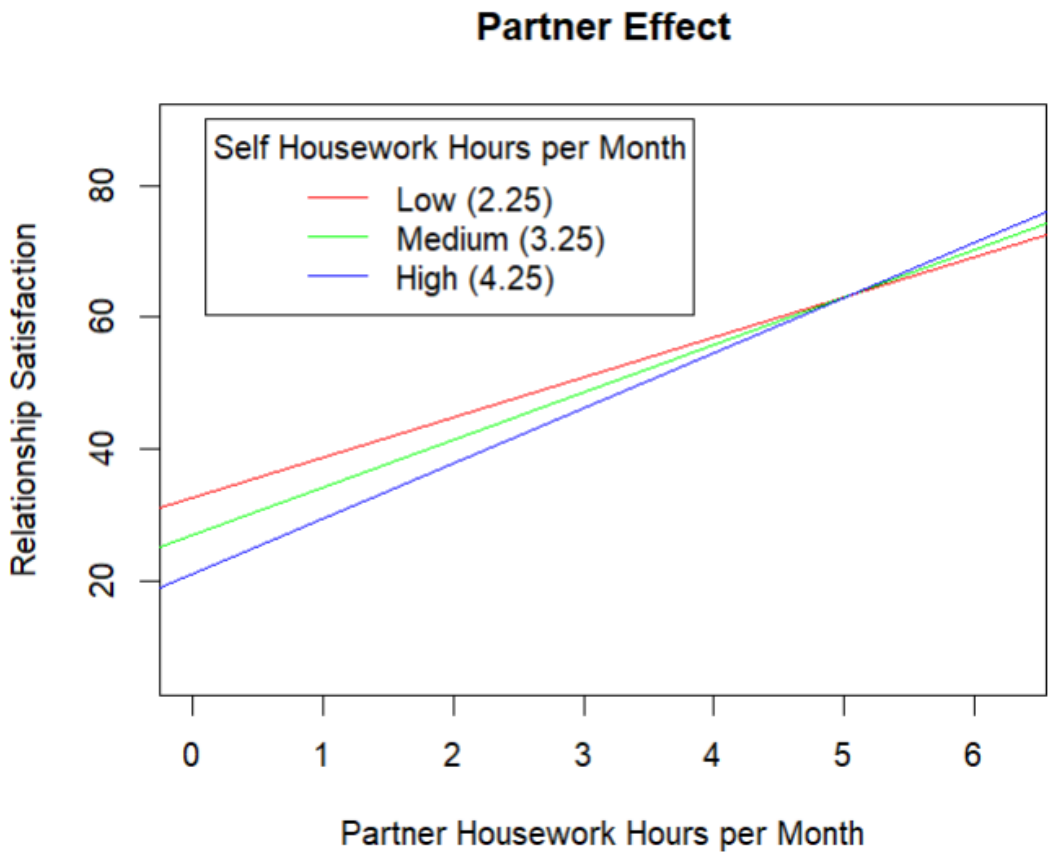
หากตนเองทำงานบ้านน้อย ยิ่งคู่ทำงานบ้านมากขึ้น ยิ่งทำให้ความพึงพอใจในความสัมพันธ์มากขึ้น

หากตนเองทำงานบ้านปานกลาง ยิ่งคู่ทำงานบ้านมากขึ้น ยิ่งทำให้ความพึงพอใจในความสัมพันธ์มากยิ่งขึ้น

หากตนเองทำงานบ้านมาก ยิ่งคู่ทำงานบ้านมากขึ้น ยิ่งทำให้ความพึงพอใจในความสัมพันธ์มากยิ่งขึ้น อิทธิพลผู้แสดงจะสูงที่สุด

```
> cond_housep <- quick_sim_slopes(out3b, pred="housepartner", modx="house", modx.values=housevalue)
> cond_housep_int <- cond_housep[[1]][,"est"]
> cond_housep_slope <- cond_housep[[2]][,"est"]
> plot(dat3a$house, dat3a$sat, type="n", xlab="Partner Housework Hours per Month",
+      ylab="Relationship Satisfaction", main="Partner Effect")
> abline(a = cond_housep_int[1], b = cond_housep_slope[1], col="red")
> abline(a = cond_housep_int[2], b = cond_housep_slope[2], col="green")
> abline(a = cond_housep_int[3], b = cond_housep_slope[3], col="blue")
> legend(0.1, 90, c("Low (2.25)", "Medium (3.25)", "High (4.25)"),
+        col=c("red", "green", "blue"), lty=1, title="Self Housework Hours per Month")
```

ยิ่งตนเองทำงานบ้านมากขึ้น อิทธิพลของการทำงานบ้านของคู่ต่อความพึงพอใจในความสัมพันธ์จะเป็นในทางบวกในระดับที่มากขึ้น



โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- หลีกเลี่ยงการย้ายศูนย์กลางไปที่ค่าเฉลี่ยกลุ่ม (Group-mean centering) เนื่องจากเป็นโมเดลเดียวกับ APIM และ APIM สามารถแปลความหมายได้ดีกว่า เลยใช้ APIM เป็นหลัก
- ตัวแปรผสม (Mixed variables) ซึ่งเป็นตัวแปรระดับสมาชิก จะต้องใส่ทั้งค่าของตนเอง และค่าของคู่เสมอ ไม่ว่าจะเป็นตัวแปรอิสระหรือตัวแปรควบคุม เพื่อหลีกเลี่ยงความผิดพลาดในการระบุโมเดล (Model misspecification)
  - ไม่ต่างกับการใส่ค่าเฉลี่ยกลุ่มเข้าไปเสริมตัวแปรระดับที่ 1 ที่ต้องใส่เสมอใน MLM

โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- ลองพิจารณาว่า ถ้าไม่ได้เก็บข้อมูลทั้ง 2 คนในแต่ละคู่ เก็บได้เพียงคนเดียว แล้วประมาณค่าอิทธิพลผู้แสดง (Actor Effect) จะเป็นอย่างไร

ใช้ข้อมูลคนเดียวในแต่ละคู่

```
> dat3c <- dat3[!duplicated(dat3$coupleid),]
> summary(lm(sat ~ house, data=dat3c))

Call:
lm(formula = sat ~ house, data = dat3c)

Residuals:
    Min      1Q  Median      3Q     Max
-29.690  -8.277  -0.051   7.874  34.159

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   64.306      1.768   36.38  < 2e-16 ***
house         -4.302      0.524   -8.21 1.91e-15 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.68 on 498 degrees of freedom
Multiple R-squared:  0.1192,    Adjusted R-squared:  0.1174
F-statistic:  67.4 on 1 and 498 DF,  p-value: 1.912e-15
```

อิทธิพลผู้แสดง = -4.30 ซึ่งเทียบกับ APIM แล้ว อิทธิพลผู้แสดง = -2.20

อิทธิพลตัวหลังเป็นอิทธิพลเมื่อควบคุมการทำงานบ้านของคู่แล้ว ซึ่งทำให้เห็นภาพทั้งหมดชัดกว่า

โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- สมมติว่าเก็บข้อมูลได้ทั้ง 2 คนในแต่ละคู่ แต่ประมาณค่าอิทธิพลผู้แสดง (Actor Effect) เพียงอย่างเดียว ไม่ใส่อิทธิพลคู่ จะเป็นอย่างไร

```
> out3bias <- gls(sat ~ 1 + house,
+                 data = dat3a, correlation = corCompSymm(form = ~1|coupleid))
> summary(out3bias)
Generalized least squares fit by REML
  Model: sat ~ 1 + house
  Data: dat3a
       AIC      BIC    logLik
  7651.409 7671.032 -3821.704

Correlation Structure: Compound symmetry
 Formula: ~1 | coupleid
 Parameter estimate(s):
      Rho
0.5441211

Coefficients:
               Value Std.Error   t-value p-value
(Intercept) 74.74862 1.0658673  70.12939       0
house       -7.59277 0.2930558 -25.90896       0

 Correlation:
      (Intr)
house -0.895
```

เมื่อเปลี่ยนไปใช้คำสั่ง `gls` ผลของผู้แสดงยิ่งรุนแรงมากกว่าการใช้ `lm` เสียอีก โมเดลนี้จะเป็นโมเดลที่เหมาว่าอิทธิพลคู่ (Partner Effect) = 0 ซึ่งไม่เป็นความจริง

โมเดลความยึดโยงระหว่างผู้แสดงและคู่

- วิธีการหาสัมประสิทธิ์การทำนาย (R^2) จากฟังก์ชั่น `gls` จะคล้ายกับการหาค่าดังกล่าวในการวิเคราะห์ถดถอยหรือฟังก์ชั่น `lm` ปกติ
- กล่าวคือ สามารถคำนวณได้จากนำค่าสหสัมพันธ์ของค่าที่ทำนายได้ $(\hat{Y})$ และค่าที่แท้จริงของตัวแปรตาม (Y) มายกกำลังสองดังนี้

```
> cor(predict(out3b), dat3$sat)^2
[1] 0.4194868
```

โมเดลที่มีค่าปฏิสัมพันธ์ระหว่างการทำงานบ้านของตนเองและคู่ มีค่าสัมประสิทธิ์การทำนายเท่ากับ 42%

โมเดลความยืดโยงระหว่างผู้แสดงและคู่

- เมื่อเปรียบเทียบกับการ R^2 จากฟังก์ชั่น `lme` ด้วย `r2mlm` package แล้ว จะพบว่าได้ค่าเท่ากัน

```
> out3c <- lme(sat ~ 1 + house*housepartner, random = ~1 |coupleid, data = dat3a)
> library(r2mlm)
> r2mlm(out3c)
$Decompositions
                    total
fixed           0.4185914
slope variation 0.0000000
mean variation  0.1940320
sigma2          0.3873767

$R2s
       total
f   0.4185914
v   0.0000000
m   0.1940320
fv  0.4185914
fvm 0.6126233
```

$$R_t^{2(f)} = 42\%$$

ตัวแปรอิสระทั้งหมด สามารถทำนายความแปรปรวนของตัวแปรตามได้ 42%

อย่างไรก็ตาม `r2mlm` ไม่สามารถใช้ได้ทุกสถานการณ์ เพราะ `lme` จะไม่เหมาะสมเมื่อ ICC มีค่าลบ