

# Day 1: Test for Distinguishability

Back to schedule

---

(Material in this handout is not in Kenny, Kashy, and Cook, 2006)

Advantages of Treating Dyad Members as Indistinguishable:

- Simpler model with fewer parameters
- More power in tests of actor and partner effects

Disadvantages of Treating Dyad Members as Indistinguishable:

- If distinguishability makes a difference, then the model is wrong.
- Sometimes the focus is on distinguishing variable and it is lost.
- Some editors or reviewer will not allow you to do it.

Two Runs:

- Distinguishable (either interaction or two-intercept, results are the same)
  - Different Actor and Partner Effects
  - Main Effect of Distinguishing Factor
  - Heterogeneity of Variance: Add a `weights = option`
- Indistinguishable (4 fewer parameters)
  - Same Actor and Partner Effects
  - No Main Effect of Distinguishing Factor
  - No `weights = option`

Run using `method = "ML"`, not `method = "REML"` (or blank because REML is the default).

Note the number of parameters:

- There should be 4 more than for the distinguishable run.

Note the  $-2\text{LogLikelihood}$  (deviance) for each model. Subtract the deviances and number of parameters to get a  $\chi^2$  with 4 df. Or simply have R do it for you with `anova()`.

If  $\chi^2$  is not significant, then the data are consistent with the null hypothesis that the dyad members are indistinguishable. If however,  $\chi^2$  is significant, then the data are inconsistent with the null hypothesis that the dyad members are indistinguishable (i.e., dyad members are distinguishable in some way).

Read in the individual data (or a pairwise dataset)

```
library(tidyr)
library(dplyr)
library(nlme)

acitelli_ind <- read.csv(file.choose(), header=TRUE)
```

Convert individual data to pairwise.

```
tempA <- acitelli_ind %>%
  mutate(genderE = gender, partnum = 1) %>%
  mutate(gender = ifelse(gender == 1, "A", "P")) %>%
  gather(variable, value, self_pos:genderE) %>%
  unite(var_gender, variable, gender) %>%
  spread(var_gender, value)
```

```

tempB <- acitelli_ind %>%
  mutate(genderE = gender, partnum = 2) %>%
  mutate(gender = ifelse(gender == 1, "P", "A")) %>%
  gather(variable, value, self_pos:genderE)%>%
  unite(var_gender, variable, gender) %>%
  spread(var_gender, value)

acitelli_pair <- bind_rows(tempA, tempB) %>%
  arrange(cuplid) %>%
  mutate(gender_A = ifelse(genderE_A == 1, "hus", "wife"), gender_A = as.factor(gender_A))

rm(tempA, tempB)

```

## Indistinguishable dyads.

Make sure to use method = "ML".

```

apim_in <- gls(satisfaction_A ~ other_pos_A + other_pos_P,
  data = acitelli_pair,
  method = "ML",
  correlation = corCompSymm(form=~1|cuplid),
  na.action = na.omit)

summary(apim_in)

```

```

## Generalized least squares fit by maximum likelihood
## Model: satisfaction_A ~ other_pos_A + other_pos_P
## Data: acitelli_pair
##      AIC      BIC    logLik
## 292.8841 311.3359 -141.442
##
## Correlation Structure: Compound symmetry
## Formula: ~1 | cuplid
## Parameter estimate(s):
##      Rho
## 0.4666759
##
## Coefficients:
##              Value Std.Error  t-value p-value
## (Intercept) 0.6697541 0.3218874  2.080709  0.0383
## other_pos_A 0.4004231 0.0472926  8.466931  0.0000
## other_pos_P 0.2879705 0.0472926  6.089124  0.0000
##
## Correlation:
##      (Intr) oth__A
## other_pos_A -0.792
## other_pos_P -0.792  0.264
##
## Standardized residuals:
##      Min      Q1      Med      Q3      Max
## -4.8047105 -0.4682679  0.1310436  0.6452902  1.7824857
##

```

```
## Residual standard error: 0.4149128
## Degrees of freedom: 296 total; 293 residual
```

## Distinguishable Dyads

Make sure to use method = "ML".

```
apim_di <- gls(satisfaction_A ~ gender_A + other_pos_A:gender_A + other_pos_P:gender_A - 1,
  data = acitelli_pair,
  method = "ML",
  correlation = corCompSymm(form=~1|cuplid),
  weights = varIdent(form=~1|genderE_A),
  na.action = na.omit)

summary(apim_di)
```

```
## Generalized least squares fit by maximum likelihood
## Model: satisfaction_A ~ gender_A + other_pos_A:gender_A + other_pos_P:gender_A - 1
## Data: acitelli_pair
## AIC BIC logLik
## 293.607 326.8203 -137.8035
##
## Correlation Structure: Compound symmetry
## Formula: ~1 | cuplid
## Parameter estimate(s):
## Rho
## 0.4751094
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | genderE_A
## Parameter estimates:
## 1 -1
## 1.000000 1.203895
##
## Coefficients:
## Value Std.Error t-value p-value
## gender_Ahus 0.6904589 0.3429034 2.013567 0.0450
## gender_Awife 0.6112485 0.4128195 1.480667 0.1398
## gender_Ahus:other_pos_A 0.4243866 0.0677157 6.267184 0.0000
## gender_Awife:other_pos_A 0.3777000 0.0739222 5.109428 0.0000
## gender_Ahus:other_pos_P 0.2616498 0.0614025 4.261222 0.0000
## gender_Awife:other_pos_P 0.3214782 0.0815225 3.943427 0.0001
##
## Correlation:
## gndr_Ah gndr_Aw gndr_Ah:__A gndr_Aw:__A
## gender_Awife 0.475
## gender_Ahus:other_pos_A -0.667 -0.317
## gender_Awife:other_pos_A -0.267 -0.562 -0.111
## gender_Ahus:other_pos_P -0.562 -0.267 -0.234 0.475
## gender_Awife:other_pos_P -0.317 -0.667 0.475 -0.234
## gndr_Ah:__P
## gender_Awife
## gender_Ahus:other_pos_A
```

```
## gender_Awife:other_pos_A
## gender_Ahus:other_pos_P
## gender_Awife:other_pos_P -0.111
##
## Standardized residuals:
##      Min      Q1      Med      Q3      Max
## -4.7303571 -0.4647126  0.1311508  0.6387054  1.9631065
##
## Residual standard error: 0.3744831
## Degrees of freedom: 296 total; 290 residual
```

The following function call conducts the  $\chi^2$  test for distinguishability.

```
anova(apim_in, apim_di)
```

```
##      Model df      AIC      BIC    logLik    Test  L.Ratio p-value
## apim_in    1   5 292.8841 311.3359 -141.4420
## apim_di    2   9 293.6070 326.8203 -137.8035 1 vs 2  7.277053   0.122
```

But, you can always hand calculate it. For simplicity, in these calculations we use the customary  $-2 \times \log \text{Lik}$ , or the “deviance”.

$$\chi^2(\text{parameters}_{dist} - \text{parameters}_{indist}) = \text{deviance}_{indist} - \text{deviance}_{dist}$$

$$\chi^2(9 - 5) = 282.884 - 275.607 = 7.277, p = .122$$

The null hypothesis is that the dyads are indistinguishable. We cannot reject the null hypothesis, so we conclude that there is no empirical evidence that dyad members should be differentiated by their gender.

In general, any two multilevel models can be compared with `anova()` if one model is a nested version of the other. Model A is said to be a nested version of Model B, if Model A is a simpler version of Model B, i.e., one with fewer parameters. One computes estimates each model (using ML if the models have different fixed effects or REML if the two models have the same fixed effects) and subtracts the deviance of Model B from Model A. Under the null hypothesis that the simplifications of Model A are true, that difference is distributed as chi square where the degrees of freedom is the number of extra parameters that Model B has.