Lab 4 - Mediation

Structural Equation Modeling ED 216F - Instructor: Karen Nylund-Gibson

Adam Garber

April 23, 2020

Contents

---

1 Lab preparation

1.1 Creating a version-controlled R-Project with Github

Download repository here: https://github.com/garberadamc/SEM-Lab4

On the Github repository webpage:

a. `fork` your own `branch` of the lab repository
b. copy the repository web URL address from the `clone or download` menu

Within R-Studio:

c. click "NEW PROJECT" (upper right corner of window)
d. choose option `Version Control`
e. choose option `Git`
f. paste the repository web URL path copied from the `clone or download` menu on Github page
g. choose location of the R-Project (too many nested folders will result in filepath error)

1.2 Load packages

```
library(datapasta)
library(mediation)
library(tidyverse)
library(MplusAutomation)
library(rhdf5)
library(here)
library(kableExtra)
library(gtsummary)
```

---

2 Lab outline

1. Estimate a mediation model using the {`mediation`} package
2. Estimate the same model using the Structural Equation Modeling (SEM) framework with {`MplusAutomation`}
3. For the second empirical example, estimate parallel models using the `mediation` and `SEM` methods

---

2.1 A quick detour - Equivalent models

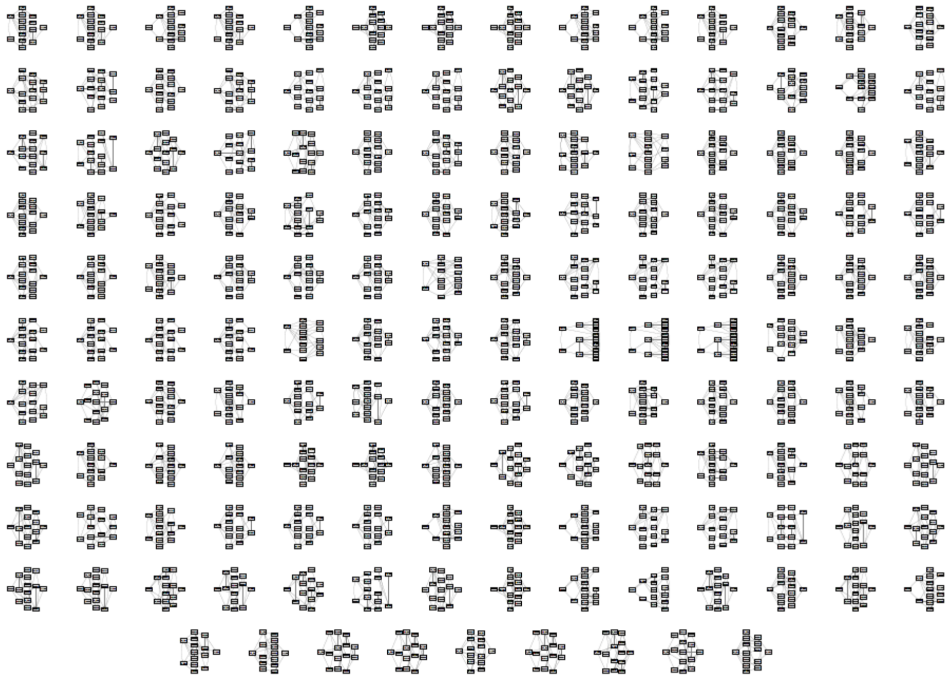


Figure. Picture adapted from SEM slides by Sacha Epskamp http://sachaepskamp.com/files/SEM22019/SEM2_2019_Week2_slides.pdf

---

2.2 Have you ever seen the perfect table and want to adapt it for your own research purposes?

Use {`datapasta`} by copying tables and pasting them automatically as `tribbles` or `dfs`

1. copy a table or data matrix
2. run the fuction `tribble_paste()` or `df_paste()`

```
tribble_paste()


pasta_table %>%
  kable(booktabs = T, linesep = "") %>%
  kable_styling(latex_options = c("striped"),
                full_width = F,
                position = "left")
```

2.3 The empirical examples of mediation used in this exercise are from the following article

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis.

https://cran.r-project.org/web/packages/mediation/vignettes/mediation.pdf

---

2.4 Data source for example 1

Brader T, Valentino NA, Suhat E (2008). **What Triggers Public Opposition to Immigration? Anxiety, Group Cues, and Immigration.** American Journal of Political Science, 52(4), 959–978.

https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-5907.2008.00353.x

To see metadata run - `?framing`

Read in the `framing` dataset

```r
set.seed(4212020)

data("framing", package = "mediation")

framing <- droplevels(framing) %>%  # drop factor levels with frequency zero
  mutate(emo = emo - 2)
```

Take a look at variables used in the mediation model

Name	Labels
emo	Measure of subjects' negative feeling during the experiment (1-10). 1 indicates the most negative feeling.
treat	Framing codition interaction term. News story with conditions tone (Negative/Positive) and ethnic identity
cong_mesg	Whether subjects requested sending an anti-immigration message to Congress on their behalf.
age	Age of subject (18-85)
educ	Education (1-4)
gender	Gender (Male/Female)
income	Subjects' income, measured as a 19-point scale.

Look at descriptives table for the framing dataset using {`gtsummary`}

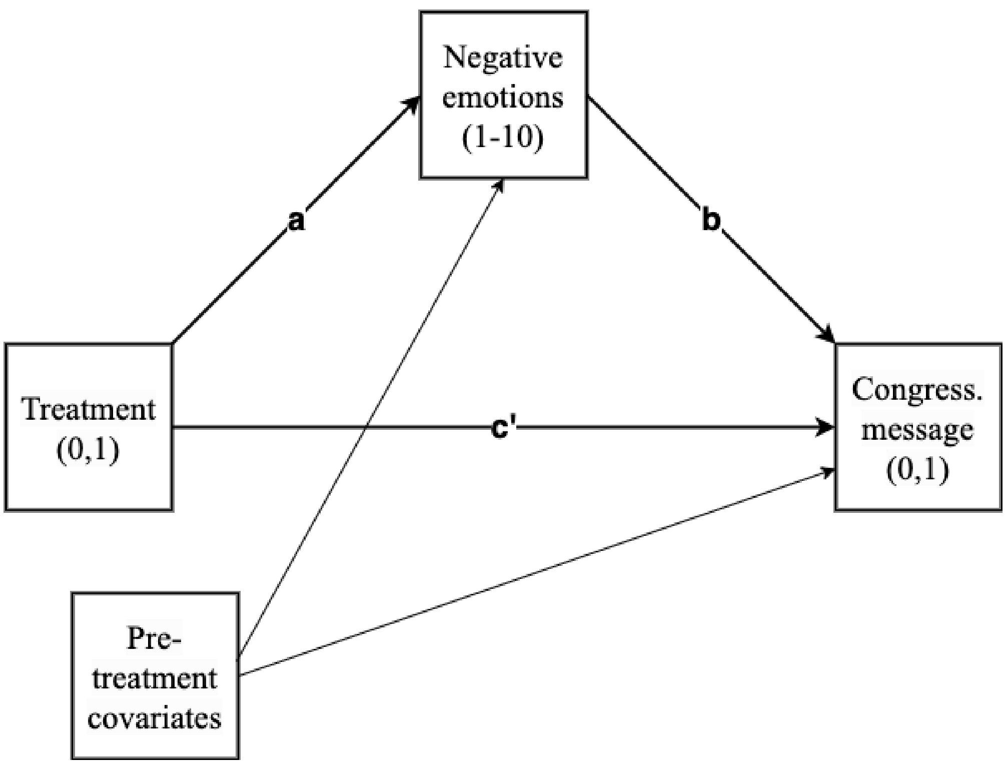
```r
table_data <- framing %>%
  select(emo, treat, cong_mesg, age, educ, gender, income)

table1 <- tbl_summary(table_data,
                      statistic = list(all_continuous() ~ "{mean} ({sd})"),
                      missing = "no") %>%
  bold_labels()

table1
```

Characteristic	**N = 265**[1]
emo	4.97 (2.77)
treat	68 (26%)
cong_mesg	88 (33%)
age	48 (16)
educ	
less than high school	20 (7.5%)
high school	92 (35%)
some college	70 (26%)
bachelor's degree or higher	83 (31%)
gender	
male	126 (48%)
female	139 (52%)
income	11 (4)

[1]Statistics presented: mean (SD); n (%)



2.5 Estimate a mediation model in R using {`mediation`}

step 1: fit a linear model of the mediator (`emo`) regressed on treatment (`treat`) and pre-treatment covariates

```r
med_fit <- lm(emo ~ treat + age + educ + gender + income,
              data = framing)
```

---

step 2: fit a general linear model (`glm`) with the binary outcome variable `cong_mesg` regressed on treatment (`treat`), mediator, and pre-treatment covariates

```r
out_fit <- glm(cong_mesg ~ emo + treat + age + educ + gender + income,
               data = framing,
               family = binomial("probit"))
```

---

step 3: estimate the mediation effects with bias corrected bootstrapped confidence intervals

```r
med_out <- mediate(med_fit, out_fit, treat = "treat", mediator = "emo",
                   boot = TRUE, boot.ci.type ="bca", sims = 100)

summary(med_out)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the BCa Method
##
##                          Estimate 95% CI Lower 95% CI Upper p-value
## ACME (control)             0.0824       0.0246         0.13  <2e-16 ***
## ACME (treated)             0.0835       0.0239         0.14  <2e-16 ***
## ADE (control)              0.0113      -0.0921         0.12    0.70
## ADE (treated)              0.0124      -0.1051         0.13    0.70
## Total Effect               0.0948      -0.0205         0.25    0.16
## Prop. Mediated (control)   0.8693     419.1265       666.19    0.16
## Prop. Mediated (treated)   0.8808     369.2224       586.64    0.16
## ACME (average)             0.0829       0.0224         0.13  <2e-16 ***
## ADE (average)              0.0118      -0.0991         0.12    0.70
## Prop. Mediated (average)   0.8751     394.1745       626.42    0.16
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 265
##
##
## Simulations: 100
```

2.6 Run mediation model 1 using the Structural Equation Modeling framework with {`MplusAutomation`}

```r
m1_mediate  <- mplusObject(
  TITLE = "m1 mediate Lab4",
  VARIABLE =
   "usevar =
    cong_mesg emo treat age
    educ gender income;

    categorical = cong_mesg; ! outcome is binary",

  ANALYSIS =  "bootstrap = 500; ! set number of bootstrap samples (500 for example purposes)" ,

  MODEL =
   "emo on treat age educ gender income;            ! mediator linear regression
    cong_mesg on emo treat age educ gender income;  ! outcome GLM regression

    Model indirect:
    cong_mesg ind treat;" ,

  OUTPUT =
    "sampstat standardized cinterval (bcbootstrap); ! bias-corrected bootstrap",

  PLOT = "type=plot2;",

  usevariables = colnames(framing),
  rdata = framing)

m1_mediate_fit <- mplusModeler(m1_mediate,
                     dataout=here("mplus_files", "Lab4.dat"),
                    modelout=here("mplus_files", "m1_mediate_Lab4.inp"),
                    check=TRUE, run = TRUE, hashfilename = FALSE)
```

2.7 Data source for example 2

Vinokur AD, Price RH, Schul Y (1995). **Impact of the JOBS Intervention on Unemployed Workers Varying in Risk for Depression.** American Journal of Community Psychology, 23(1), 39–74.

https://link.springer.com/content/pdf/10.1007/BF02506922.pdf

To see metadata run - `?jobs`

Note: For this example we will ignore the issue of non-compliance addressed in Tingley et al. (2014) as this causal inference topic is beyond the scope of this course.

Read in the data from the job search intervention study (`jobs`)

```r
data("jobs", package = "mediation")
```

Take a look at variables used in the mediation model

Name	Label
depress2 (Y)	Measure of depressive symptoms post-treatment.
treat (X)	Indicator variable for whether participant was randomly selected for the JOBS II training program. 1 = as
job_dich (Z)	The job_seek measure recoded into two categories of high and low. 1 = high job search self-efficacy.
sex	Indicator variable for sex. 1 = female
age	Age in years.
marital	Factor with five categories for marital status.
nonwhite	Indicator variable for race. 1 = nonwhite.
educ	Factor with five categories for educational attainment.
income	Factor with five categories for level of income.

Look at descriptives of the framing dataset using {`gtsummary`}

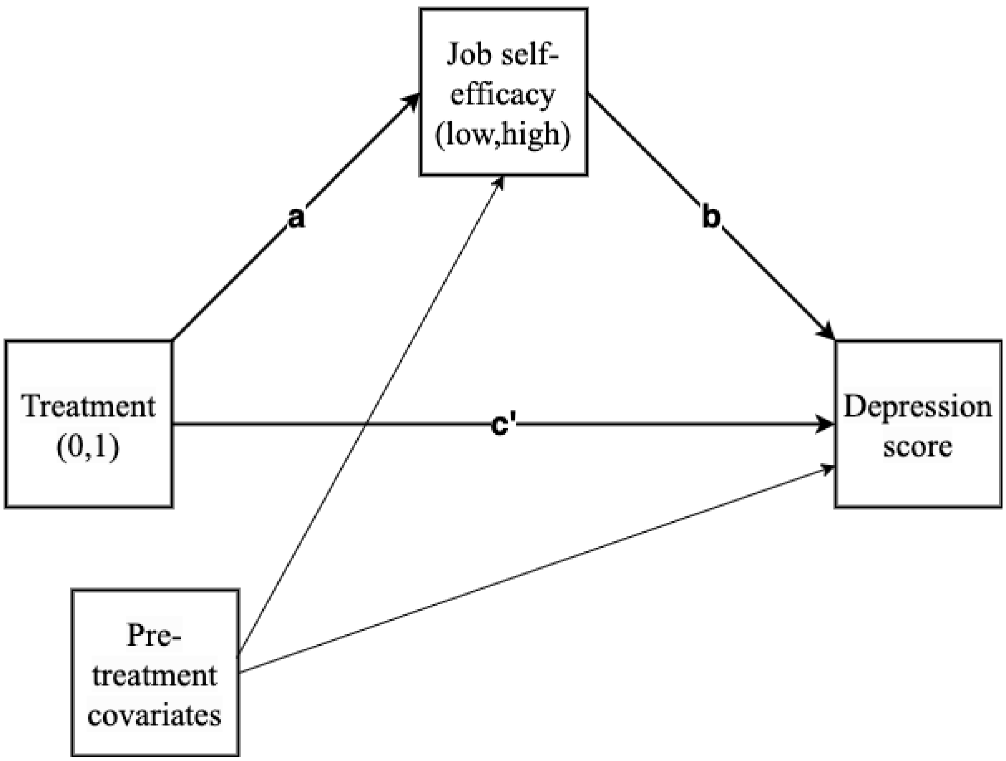
```
jobs_desc <- jobs %>%
  select(depress2, job_dich, treat, sex, age, marital, nonwhite, educ, income)

tablej <- tbl_summary(jobs_desc,
                      statistic = list(all_continuous() ~ "{mean} ({sd})"),
                      missing = "no" ) %>%
  bold_labels()

tablej
```

Characteristic	**N = 899**[1]
depress2	1.74 (0.65)
job_dich	555 (62%)
treat	600 (67%)
sex	482 (54%)
age	38 (10)
marital	
nevmarr	279 (31%)
married	408 (45%)
separtd	30 (3.3%)
divrcd	163 (18%)
widowed	19 (2.1%)
nonwhite	
white0	747 (83%)
non.white1	152 (17%)
educ	
lt-hs	50 (5.6%)
highsc	272 (30%)
somcol	319 (35%)
bach	146 (16%)
gradwk	112 (12%)
income	
lt15k	164 (18%)
15t24k	206 (23%)
25t39k	218 (24%)
40t49k	110 (12%)
50k+	201 (22%)

[1]Statistics presented: mean (SD); n (%)



---

step 1: fit a binomial logist model using `glm` with the binary mediator (`job_dich`) regressed on treatment (`treat`) and pre-treatment covariates

```
jmed_fit <- glm(job_dich ~ treat + sex + age + marital +
                   nonwhite + educ + income,
                 data = jobs, family = binomial)
```

---

step 2: fit a linear model with depression score (`depress2`) regressed on treatment, mediator, and pre-treatment covariates

```
jout_fit <- lm(depress2 ~ job_dich + treat +
                 sex + age + marital + nonwhite + educ + income,
               data = jobs)
```

---

step 3: Estimate the mediation effects with bias corrected bootstrapped confidence intervals.

```r
jmed_out <- mediate(jmed_fit, jout_fit, treat = "treat", mediator = "job_dich",
                    boot = TRUE, boot.ci.type ="bca", sims = 100)

summary(jmed_out)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the BCa Method
##
##                  Estimate 95% CI Lower 95% CI Upper p-value
## ACME              -0.0237      -0.0479         0.00    0.02 *
## ADE               -0.0306      -0.1047         0.04    0.56
## Total Effect      -0.0543      -0.1373         0.02    0.20
## Prop. Mediated     0.4359       0.2505        44.45    0.22
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 899
##
##
## Simulations: 100
```

2.8 Run mediation model 2 as a SEM model with {MplusAutomation}

```r
m2_jmediate  <- mplusObject(

  TITLE = "m2 jobs mediate Lab4",

  VARIABLE =
   "usevar = treat sex
    age marital nonwhite
    educ income depress2 job_dich;

    categorical = job_dich; ! moderator is binary",

  ANALYSIS =
   "bootstrap = 500; ! set number of bootstrap draws (500 for example purposes)" ,

  MODEL =
   "job_dich on treat sex age marital nonwhite educ income;

    depress2 on job_dich treat sex age marital nonwhite educ income;

    Model indirect:
    depress2 ind treat;" ,

  OUTPUT =
```

```
    "sampstat standardized cinterval (bcbootstrap); ! bias-corrected bootstrap",

  PLOT = "type=plot2;",

  usevariables = colnames(jobs),
  rdata = jobs)

m2_jmediate_fit <- mplusModeler(m2_jmediate,
                     dataout=here("mplus_files", "Lab4_jobs.dat"),
                    modelout=here("mplus_files", "m2_jmediate_Lab4.inp"),
                    check=TRUE, run = TRUE, hashfilename = FALSE)
```

2.9 Run model 3 including the mediator*treatement interaction (potential outcomes framework)

```
m3_jmed  <- mplusObject(

  TITLE = "m3 MX jobs mediate Lab4",

  VARIABLE =
   "usevar =
    treat sex age marital nonwhite
    educ income depress2 job_dich mx; ",

  DEFINE = "mx = job_dich*treat;",

  ANALYSIS =  "bootstrap = 500; ",

  MODEL =
   "job_dich on treat sex age marital nonwhite educ income;
    depress2 on job_dich treat mx sex age marital nonwhite educ income;

    Model indirect:
    depress2 MOD job_dich mx treat; ",

  OUTPUT =
    "sampstat cinterval(bootstrap); ",

  usevariables = colnames(jobs),
  rdata = jobs)

m3_jmed_fit <- mplusModeler(m3_jmed,
                     dataout=here("mplus_files", "Lab4_jobs.dat"),
                    modelout=here("mplus_files", "m3_jmediate_Lab4.inp"),
                    check=TRUE, run = TRUE, hashfilename = FALSE)
```

---

3 References

Brader T, Valentino NA, Suhat E (2008). What Triggers Public Opposition to Immigration? Anxiety, Group Cues, and Immigration. American Journal of Political Science, 52(4), 959–978.

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. Structural equation modeling: a multidisciplinary journal, 25(4), 621-638.

Ingels, S. J., Pratt, D. J., Herget, D. R., Burns, L. J., Dever, J. A., Ottem, R., … & Leinwand, S. (2011). High School Longitudinal Study of 2009 (HSLS: 09): Base-Year Data File Documentation. NCES 2011-328. National Center for Education Statistics.

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis.

Vinokur AD, Price RH, Schul Y (1995). Impact of the JOBS Intervention on Unemployed Workers Varying in Risk for Depression. American Journal of Community Psychology, 23(1), 39–74.

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

UC SANTA BARBARA