

Course: Multivariate statistics (AUT23)

Chapter 8: Mediation

8.14 Time to practice on your own

8.15.1 Exercise 1: Mediation analysis (using lavaan)

In this exercise, we will use the data “protest.sav” (Hayes, 2022) which can be downloaded here under “data files and code”. Especially, we will focus on the following variables:

- Protest (independent variable): A lawyer protests against gender discrimination (experimental group, dichotomous 0 = no and 1 = yes)
- Respappr (mediator): Perceived adequacy of response (scale 1-7)
- Like (dependent variable): assessment of the lawyer (scale 1-7)

Hypothesis: If the lawyer protests, her reaction will be perceived as more appropriate, and therefore the lawyer will be evaluated more favorably.

Start by drawing the regression equations.

- Solution: equation

Now, we want to calculate the mediation model and to answer the following questions:

- Is the a-path significant? If so, how much variance does it explain?
- Are the b-path and the c'-path significant? If so, how much variance do they explain together?
- Is there a total effect of X on Y?
- Is there a significant mediation effect? If so, is there partial or total mediation?

Start by loading and selecting the data:

- Show the code

Now, define the model in lavaan and give the results for the c path:

- Show the code
- Solution: Interpretation

Now get the model for the c' path:

- Show the code
- Solution: Interpretation

The output is a little better understandable if you take the indirect and the overall effect on Y with the names of the paths:

- Show the code

Give the full interpretation of the model:

- Solution: Interpretation

8.14.2 Exercise 2: Moderated mediation analysis (using lavaan)

In this exercise, we will use the data “protest.sav” (Hayes, 2022) which can be downloaded [here](#) under “data files and code”. Especially, we will focus on the following variables:

- Protest (independent variable): A lawyer protests against gender discrimination (experimental group, dichotomous 0 = no and 1 = yes)
- Respappr (mediator): Perceived adequacy of response (scale 1-7)
- Like (dependent variable): assessment of the lawyer (scale 1-7)
- Sexism (moderator): perception of sexism as a ubiquitous problem in society (scale 1-7)

We want to test the assumption that if the lawyer protests, her response will be judged more appropriate by women who perceive sexism as a problem (moderator: dichotomous variable “sexism”), and therefore the lawyer will be judged better.

We want to test the following hypothesis: If the lawyer protests against gender discrimination, her response is perceived as more appropriate and therefore the lawyer is judged better.

Start by drawing the regression equations.

- Solution: equation

Now, we want to calculate the mediation model and to answer to following questions:

- Is the a-path moderated?
- How much mediator variance does regression explain for the overall a-path and how much of that is explained by moderation?

Illustrate the moderation of the a-path.

- What do the results mean in terms of content?
- Are the b-path and the c’-path significant?
- If so, how much variance do they explain together?

- Show the code

What do the results mean in terms of content?

- Solution: Interpretation

Finally, we want to know whether there is a significant moderated mediation effect? If so, how can this be described and interpreted in terms of content?

- Solution: Interpretation

Chapter 8: Mediation (answers)

8.14 Time to practice on your own

Chapter 8: Mediation

8.14 Time to practice on your own

8.15.1 Exercise 1: Mediation analysis (using lavaan)

In this exercise, we will use the data “protest.sav” (Hayes, 2022) which can be downloaded here under “data files and code”. Especially, we will focus on the following variables:

- Protest (independent variable): A lawyer protests against gender discrimination (experimental group, dichotomous 0 = no and 1 = yes)
- Respappr (mediator): Perceived adequacy of response (scale 1-7)
- Like (dependent variable): assessment of the lawyer (scale 1-7)

Hypothesis: If the lawyer protests, her reaction will be perceived as more appropriate, and therefore the lawyer will be evaluated more favorably.

Start by drawing the regression equations.

➤ Solution: equation

The regression equations go as:

$$Y_i = \beta_0 + \beta_1 * Protest_i + \beta_2 * Respappr_i + \epsilon_i = c' + b$$

$$M_i = \beta_0 + \beta_1 * Protest_i + \epsilon_i = a$$

$$Y_i = \beta_0 + \beta_1 * Protest_i + \epsilon_i = c = c' + ab$$

Now, we want to calculate the mediation model and to answer the following questions:

- Is the a-path significant? If so, how much variance does it explain?
- Are the b-path and the c'-path significant? If so, how much variance do they explain together?
- Is there a total effect of X on Y?
- Is there a significant mediation effect? If so, is there partial or total mediation?

Start by loading and selecting the data:

➤ Show the code

```
# load the data
```

```
library(foreign)
```

```
db <- read.spss(file=paste0(getwd(),
```

```
  "/data/protest.sav"),
```

```
  use.value.labels = F,
```

```

        to.data.frame = T)

# get the data
sel <- db |>
  dplyr::select(protest, respappr, liking) |>
  stats::na.omit()

```

Now, define the model in lavaan and give the results for the c path:

➤ Show the code

```

# define the model
modell.c = "liking ~ protest"

# get the complete output
fit.c = lavaan::sem(modell.c, data=sel)

lavaan::parameterestimates(fit.c, standardized=T)[1:8]

##   lhs op  rhs  est  se   z pvalue ci.lower
## 1 liking ~ protest 0.479 0.193 2.478 0.013 0.100
## 2 liking ~~ liking 1.044 0.130 8.031 0.000 0.789
## 3 protest ~~ protest 0.217 0.000 NA NA 0.217

lavaan::inspect(fit.c,"r2")

## liking
## 0.045

lavaan::fitMeasures(fit.c)[c("chisq","df","aic","cfi","rmsea")]

##  chisq  df  aic  cfi  rmsea
## 0.0000 0.0000 375.5983 1.0000 0.0000

```

➤ Solution: Interpretation

The C path is significant and explains 4.5% of the variance of liking. Linear regression makes sense for this data.

Now get the model for the c' path:

➤ Show the code

```

# define the model
modell.mediation = "

```

```

## Direct effect
liking ~ protest

## mediation path
respappr ~ protest
liking ~ respappr
"

# get the complete output
fit.med = lavaan::sem(modell.mediation, data=sel)
lavaan::parameterestimates(fit.med, standardized=T)[1:8]

##   lhs op   rhs  est  se   z pvalue ci.lower
## 1 liking ~ protest -0.101 0.198 -0.508 0.611 -0.489
## 2 respappr ~ protest 1.440 0.220 6.544 0.000 1.008
## 3 liking ~ respappr 0.402 0.069 5.857 0.000 0.268
## 4 liking ~~ liking 0.824 0.103 8.031 0.000 0.623
## 5 respappr ~~ respappr 1.354 0.169 8.031 0.000 1.024
## 6 protest ~~ protest 0.217 0.000 NA NA 0.217

lavaan::inspect(fit.med,"r2")

## liking respappr
## 0.246 0.249

lavaan::fitMeasures(fit.med)[c("chisq","df","aic","cfi","rmsea")]

## chisq  df  aic  cfi  rmsea
## 0.000 0.000 756.355 1.000 0.000

```

➤ Solution: Interpretation

The c' path is not significant, but both a and b are significant and the variance explained by liking is significantly higher (24.6%) than in the last model.

The output is a little better understandable if you take the indirect and the overall effect on Y with the names of the paths:

➤ Show the code

```
# define the model
```

```

modell.complete = "
## direct effect
liking ~ c*protest
## mediation path
respappr ~ a*protest
liking ~ b*respappr
## indirect effect (a*b)
ab := a*b
## total effect (c+a*b)
total := c+a*b
"

# get the complete output
fit.complete = lavaan::sem(modell.complete, data=sel)
lavaan::parameterestimates(fit.complete, standardized=T)[1:8]

##   lhs op   rhs label  est  se   z  pvalue
## 1 liking ~ protest  c -0.101 0.198 -0.508 0.611
## 2 respappr ~ protest  a  1.440 0.220  6.544 0.000
## 3 liking ~ respappr  b  0.402 0.069  5.857 0.000
## 4 liking ~~ liking    0.824 0.103  8.031 0.000
## 5 respappr ~~ respappr  1.354 0.169  8.031 0.000
## 6 protest ~~ protest   0.217 0.000   NA   NA
## 7  ab :=  a*b  ab  0.579 0.133  4.364 0.000
## 8  total := c+a*b total 0.479 0.193  2.478 0.013

lavaan::inspect(fit.complete,"r2")

##  liking respappr
##  0.246  0.249

lavaan::fitMeasures(fit.complete)[c("chisq","df","aic","cfi","rmsea")]

##  chisq  df  aic  cfi  rmsea
##  0.000  0.000 756.355  1.000  0.000

```

Give the full interpretation of the model:

➤ Solution: Interpretation

- The indirect effect (ab) of X via the mediator on Y is significant and positively directed.
- If the lawyer protests, her reaction will be perceived as more appropriate than if she does not protest. An appropriate perception of the reaction then leads to a better assessment of the lawyer.
- The direct effect (c), however, is not significant. Since the indirect effect is significant at the same time, there is complete mediation.
- The connection between the lawyer's protesting and the lawyer's evaluation is fully mediated by the lawyer's reaction being perceived as more appropriate when she protests.

8.14.2 Exercise 2: Moderated mediation analysis (using lavaan)

In this exercise, we will use the data "protest.sav" (Hayes, 2022) which can be downloaded [here](#) under "data files and code". Especially, we will focus on the following variables:

- Protest (independent variable): A lawyer protests against gender discrimination (experimental group, dichotomous 0 = no and 1 = yes)
- Respappr (mediator): Perceived adequacy of response (scale 1-7)
- Like (dependent variable): assessment of the lawyer (scale 1-7)
- Sexism (moderator): perception of sexism as a ubiquitous problem in society (scale 1-7)

We want to test the assumption that if the lawyer protests, her response will be judged more appropriate by women who perceive sexism as a problem (moderator: dichotomous variable "sexism"), and therefore the lawyer will be judged better.

We want to test the following hypothesis: If the lawyer protests against gender discrimination, her response is perceived as more appropriate and therefore the lawyer is judged better.

Start by drawing the regression equations.

➤ Solution: equation

The regression equations go as:

$$Y_i = \beta_0 + \beta_1 * Protest_i + \beta_2 * Respappr_i + \epsilon_i = c' + b$$
$$M_i = \beta_0 + \beta_1 * Protest_i + \beta_2 * Sexism_i + \beta_3 * (Protest_i * Sexism_i) + \epsilon_i = a$$
$$Y_i = \beta_0 + \beta_1 * Protest_i + \epsilon_i = c = c' + ab$$

Now, we want to calculate the mediation model and to answer to following questions:

- Is the a-path moderated?

- How much mediator variance does regression explain for the overall a-path and how much of that is explained by moderation?

Illustrate the moderation of the a-path.

- What do the results mean in terms of content?
- Are the b-path and the c'-path significant?
- If so, how much variance do they explain together?

➤ Show the code

```
# load the data
```

```
library(foreign)
```

```
db <- read.spss(file=paste0(getwd(),
                             "/data/protest.sav"),
                use.value.labels = F,
                to.data.frame = T)
```

```
# get the data
```

```
sel <- db |>
```

```
  dplyr::select(protest, respappr, liking, sexism) |>
```

```
  stats::na.omit()
```

```
# construct the interaction variable
```

```
sel$protest.sexism = sel$protest*sel$sexism
```

```
# define the model
```

```
modell.mod = "
```

```
  liking ~ c*protest ## direct effect
```

```
  respappr ~ a*protest + sexism + protest.sexism ## moderation/mediation paths
```

```
  liking ~ b*respappr
```

```
  protest ~~ sexism ## covariances
```

```
  protest ~~ protest.sexism
```

```
  sexism ~~ protest.sexism
```

```
  respappr ~1 ## Intercepts
```

```
  liking ~1
```



```

ab := a*b ## indirect effect
total := c+a*b ## total effect
"

# get the complete output
fit.mod = lavaan::sem(modell.mod, data=sel)
lavaan::parameterestimates(fit.mod, standardized=T)[1:8]

##      lhs op      rhs label  est  se   z pvalue
## 1   liking ~    protest  c -0.101 0.198 -0.508 0.611
## 2   respappr ~  protest  a -2.687 1.429 -1.880 0.060
## 3   respappr ~  sexism   -0.529 0.232 -2.278 0.023
## 4   respappr ~ protest.sexism  0.810 0.278 2.919 0.004
## 5   liking ~    respappr  b 0.402 0.069 5.857 0.000
## 6   protest ~~  sexism    0.015 0.032 0.456 0.648
## 7   protest ~~  protest.sexism  1.114 0.141 7.886 0.000
## 8   sexism ~~  protest.sexism  0.501 0.176 2.846 0.004
## 9   respappr ~1          6.567 1.191 5.516 0.000
## 10  liking ~1          3.747 0.302 12.400 0.000
## 11  liking ~~    liking    0.824 0.103 8.031 0.000
## 12  respappr ~~  respappr   1.269 0.158 8.031 0.000
## 13  protest ~~  protest    0.217 0.027 8.031 0.000
## 14  sexism ~~    sexism    0.610 0.076 8.031 0.000
## 15  protest.sexism ~~ protest.sexism  6.151 0.766 8.031 0.000
## 16  protest ~1          0.682 0.041 16.640 0.000
## 17  sexism ~1          5.117 0.069 74.441 0.000
## 18  protest.sexism ~1    3.505 0.218 16.053 0.000
## 19   ab :=      a*b  ab -1.081 0.604 -1.790 0.073
## 20   total :=    c+a*b total -1.182 0.664 -1.781 0.075

lavaan::inspect(fit.mod,"r2")

##  liking respappr

```

```
## 0.246 0.296
```

```
lavaan::fitMeasures(fit.mod)[c("chisq","df","aic","cfi","rmsea")]
```

```
## chisq df aic cfi rmsea
```

```
## 6.5581814 2.0000000 1345.5708600 0.9921174 0.1329187
```

What do the results mean in terms of content?

➤ Solution: Interpretation

The overall model for the mediator significantly explains 29.6% of the mediator's variance.

There is a significant conditional effect of the independent variable on the mediator when the moderator has a value of 0 (= with a moderate level of sexism, the lawyer's protest leads to her reaction being perceived as more appropriate).

The a-path from the independent variable to the mediator is significantly moderated by the sexism attitude. This moderation alone explains 4.7% of the variance of the mediator.

The extent to which protesting influences the perceived appropriateness of the reaction thus varies depending on the subjects' sexism attitude.

However, the strong correlations between the independent variables are a problem for the model! The standardized coefficients are outside the natural limits of -1 to +1 and RMSEA has risen above 0.1. The model represents the data poorly. Moderation is better studied independently of mediation.

Finally, we want to know whether there is a significant moderated mediation effect? If so, how can this be described and interpreted in terms of content?

➤ Solution: Interpretation

Interaction is significant: The moderator has an influence on the a-path.

Path ab is not significant: There is not an indirect effect for all values of the moderator.

RMSEA is too high and there are beta values above 1.0: The data are not suitable for this evaluation and the estimators cannot be fully trusted.