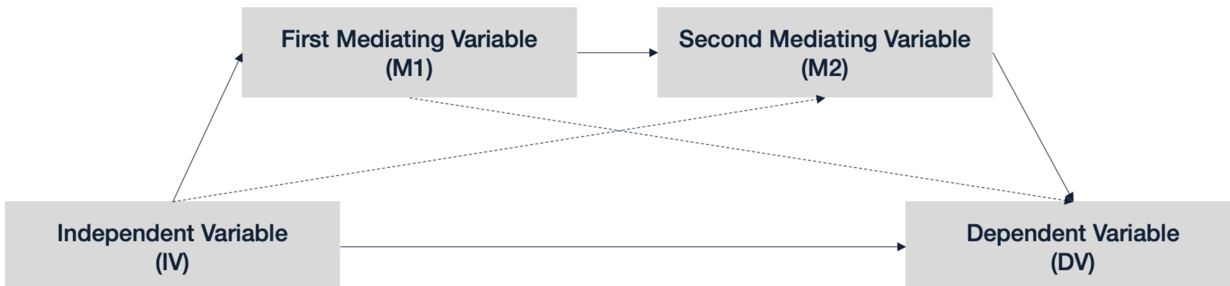
Serial Mediation in R

Amanda Danielle Perez

Serial Mediation



In our serial mediation. There are an independent variable and a dependent variable, and two mediators. The two mediators do an increasingly better job at predicting the dependent variable. The schematic above depicts how it works. In everyday research, one should deduct this serial mediation model from theory or previous findings.

Loading Libraries

We begin by loading the *lavaan* and *lavaanPlot* packages.

```
library(lavaan)
library(lavaanPlot)
library(tidyverse)
```

Loading Simulated data

IV1 is food: *how much foot the cat intakes a day*

M1 is weight: *how much the cat weighs*

M2 is lazy: *how lazy the cat is*

DV is naps: *how much the cat sleeps*

CV1 is cattitude: *how sassy the cat is*

CV2 is temperature: *how hot it is*

use read.csv() to load in the data

```
cat_nap <- read.csv("cat_naps.csv")
```

Is there a relationship?

In general, you want to first establish that there is a significant direct effect of your IV on your DV. The hope is that when you run a mediation, adding in the mediator will partial out the significance of that direct effect. Meaning, the *why* cat naps increase is explained through food increasing weight, which increased laziness, which then in turn increases cat naps.

```
  lm(dv ~ iv, cat_nap) %>%
  summary()
```

```
##
## Call:
## lm(formula = dv ~ iv, data = cat_nap)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.4672  -4.3277  -0.0632   4.3764  20.7661
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 46.45486    1.04258  44.558  < 2e-16 ***
## iv           0.07146    0.02055   3.477  0.00053 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.479 on 998 degrees of freedom
## Multiple R-squared:  0.01197,    Adjusted R-squared:  0.01098
## F-statistic: 12.09 on 1 and 998 DF,  p-value: 0.0005298
```

We do have a significant IV DV relationsip. Now, we will see if this effect becomes null when taking into account the mediators.

Configuring the SEM for the serial mediation analysis

Now, we need to communicate our model configuration to Lavaan. Therefore, we create a new object, "model" in which we save the configuration.

```
model <-
  '
  #Regressions
  m1 ~ a*iv
  m2 ~ b*m1 + iv
  dv ~ c*m2 + m1 + d*iv + cv1 + cv2

  #Defined Parameters:
  ie := a*b*c
  de := d
'
```

We first communicate the individual regressions.

Then, we define parameters within the SEM model using the `:=` operator.

We first define the **indirect effect** (ie) by multiplying the three coefficients a, b, and c. Those are the three coefficients that connect the IV to the DV via M1 and M2.

Then we extract the **direct effect**, which is simply the one coefficient d.

Using the following `sem()` command, we run the SEM. We then run the familiar "summary()" on the resulting object, which yields the below output.

```
fit <- sem(model, data = cat_nap, se = "bootstrap", bootstrap = 1000)
summary(fit, standardized = TRUE)
```

```
## lavaan 0.6-7 ended normally after 22 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         11
##
##   Number of observations                          1000
##
## Model Test User Model:
##
##   Test statistic                                 1.838
##   Degrees of freedom                                 4
##   P-value (Chi-square)                           0.766
##
## Parameter Estimates:
##
##   Standard errors                            Bootstrap
##   Number of requested bootstrap draws             1000
##   Number of successful bootstrap draws            1000
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   m1 ~
##     iv         (a)    0.433    0.019   23.373    0.000    0.433    0.592
##   m2 ~
##     m1         (b)    0.363    0.032   11.412    0.000    0.363    0.396
##     iv                0.025    0.024    1.070    0.285    0.025    0.038
##   dv ~
##     m2         (c)    0.391    0.032   12.378    0.000    0.391    0.402
##     m1                0.003    0.033    0.086    0.932    0.003    0.003
##     iv         (d)   -0.001    0.023   -0.038    0.970   -0.001   -0.001
##     cv1               0.016    0.019    0.821    0.412    0.016    0.024
##     cv2               0.009    0.019    0.467    0.641    0.009    0.013
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .m1               34.513    1.512   22.825    0.000   34.513    0.649
##    .m2               36.801    1.541   23.886    0.000   36.801    0.824
##    .dv               35.483    1.564   22.686    0.000   35.483    0.837
##
## Defined Parameters:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##     ie                0.062    0.008    8.028    0.000    0.062    0.094
##     de               -0.001    0.023   -0.038    0.970   -0.001   -0.001
```

Finally, we use *lavaanPlot()* to visualize the serial mediation.

```
lavaanPlot(model = fit, edge_options = list(color = "grey"), coefs = TRUE, sig = .05)
```

As in the model, you can see individual regressions presented in a pretty familiar way. The named/saved coefficients (a-d) are highlighted in brackets. At the very end, we see our defined parameters: We have our indirect effect and the direct effect. The convenient thing is that lavaan also tests the significance of our parameters.

The indirect effect (ie) is significant (b=.062, p<.001), and the direct effect (de) is insignificant (b= -.001, p>.05) - implying that the effect is fully mediated. We already knew this because I simulated it to be in this way.

Reporting the serial mediation as a SEM

Food intake had a significant positive effect on cat naps (b=.07, t(998) = 3.48, p<.001). As theorized, this effect was serially mediated by weight and laziness. The indirect pathway of the effect of food intake on cat naps via weight and laziness was significant (b[indirect]= .062, p<.001). This pathways fully accounted for the overall impact of food intake on cat naps with the direct effect becoming insignificant (bdirect= -.001, p>.05).

Sources

Hayes, A. F. (2013). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. New York: The Guilford Press.

Michalak, N. (2016, July 29). Reproducing Hayes' PROCESS models' results in R. Retrieved from https://nickmichalak.blogspot.com/2016/07/reproducing-hayess-process-models.html

Rosseel, Y. (2017, February 24). Package 'lavaan'. Retrieved from https://cran.r-project.org/web/packages/lavaan/lavaan.pdf

Sales, A. C. (2017). Review: Mediation package in R. Journal of Educational and Behavioral Statistics, 42, 1, 69-84.

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

MacKinnon, D. P. (2008). Introduction to statistical mediation analysis. New York, NY: Lawrence Erlbaum Associates.

Revelle, W. (2017) How to use the psych package for mediation/moderation/regression analysis.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). Journal of statistical software, 48(2), 1-36.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. Social and Personality Psychology Compass, 5(6), 359-371.

http://www.stat.cmu.edu/~hseltman/PIER/SemAndLavaan.pdf

https://cran.r-project.org/web/packages/lavaanPlot/vignettes/Intro_to_lavaanPlot.html

https://towardsdatascience.com/doing-and-reporting-a-serial-mediation-model-with-two-mediators-in-r-with-lavaan-739ed3155814

https://nmmichalak.github.io/nicholas_michalak/blog_entries/2018/nrg01/nrg01.html

https://ademos.people.uic.edu/Chapter15.html