

CR_Portfolio_6

Seunghun Lee

2020-11-22 16:09:29

Contents

Research Question	1
Hypotheses	1
Critical Test Statistics	2
Sample test statistic results	2
Confidence Interval	3
Conclusion	4
Bootstrapping	4

Research Question

- A researcher was interested in investigating whether mother's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education or 4 - higher education) affected final grade on Portuguese class (G3) through the indirect effect of mother's education on weekly study time. The researcher tested the hypothesis using a random sample of 649 students (see student-por.csv).

```
# https://archive.ics.uci.edu/ml/datasets/Student+Performance
perf <- read.csv("./data/student-por.csv", sep = ';') %>%
  dplyr::select(Medu, studytime, G3)
str(perf)
```

```
## 'data.frame': 649 obs. of 3 variables:
## $ Medu : int 4 1 1 4 3 4 2 4 3 3 ...
## $ studytime: int 2 2 2 3 2 2 2 2 2 2 ...
## $ G3 : int 11 11 12 14 13 13 13 13 17 13 ...
```

Hypotheses

- H_0 : Mother's education affects final grade on Portuguese class through the indirect effect of mother's education on weekly study time.
- H_1 : Mother's education does not affect final grade on Portuguese class through the indirect effect of mother's education on weekly study time.

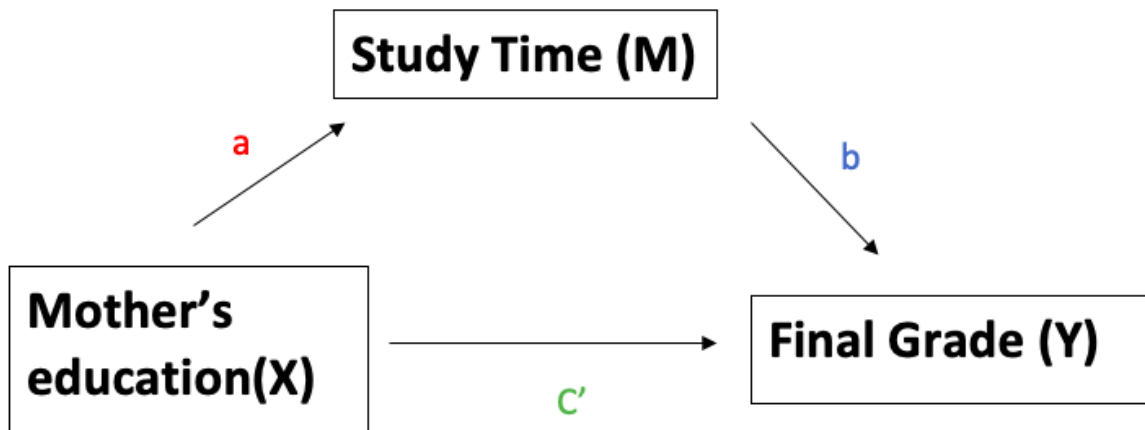


Figure 1: mediation

Critical Test Statistics

- $\alpha = 0.05$, two-tailed, $df = 649 - 2 - 1 = 646$

Sample test statistic results

- a coefficient (capturing the “effect” of the exogenous variable, Medu, on the mediator, studytime):

```

mod1 <- lm(studytime ~ Medu, data = perf)
summary(mod1)

##
## Call:
## lm(formula = studytime ~ Medu, data = perf)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.03601 -0.82324  0.03491  0.17676  2.17676
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.75231    0.07891  22.206  <2e-16 ***
## Medu         0.07092    0.02861   2.479   0.0134 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8262 on 647 degrees of freedom
## Multiple R-squared:  0.00941,    Adjusted R-squared:  0.007879
## F-statistic: 6.146 on 1 and 647 DF,  p-value: 0.01342

```

- $a = 0.071$, $S_a = 0.029$
- b coefficient from the slope for the mediator, `studytime`, in the model predicting the distal outcome, `G3`, using both the mediator and the exogenous variable, `Medu`:

```
mod2 <- lm(G3 ~ Medu + studytime, data = perf)
summary(mod2)
```

```
##
## Call:
## lm(formula = G3 ~ Medu + studytime, data = perf)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.8897  -1.6483   0.0819   2.0008   6.9723
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   8.6260      0.3871  22.282 < 2e-16 ***
## Medu          0.6207      0.1062   5.843 8.13e-09 ***
## studytime     0.8905      0.1453   6.129 1.54e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.053 on 646 degrees of freedom
## Multiple R-squared:  0.1095, Adjusted R-squared:  0.1067
## F-statistic: 39.7 on 2 and 646 DF, p-value: < 2.2e-16
```

- $b = 0.891$, $S_b = 0.145$

```
a <- unname(mod1$coefficients[2])
b <- unname(mod2$coefficients[3])
print(a*b)
```

```
## [1] 0.06315743
```

- The indirect effect estimate, $ab = 0.063$

Confidence Interval

- Then, I used Meeker & Escobar's (1994) distribution of products procedure to calculate the 95% confidence interval estimate of the indirect effect using R's `RMediation` package (Tofighi & MacKinnon, 2011) and the `medci` function as follows:

```
medci(mu.x = mod1$coefficients[2],
      mu.y = mod2$coefficients[3],
      se.x = coefficients(summary(mod1))[2,2],
      se.y = coefficients(summary(mod2))[3,2],
      type = 'dop'
)
```

```
## $'97.5% CI'
## [1] 0.01263747 0.12177032
##
## $Estimate
##      Medu
## 0.06315743
##
## $SE
##      Medu
## 0.02779311
```

Conclusion

- Reject H_0 and infer that mother's education significantly affects final grade on Portuguese class through the indirect effect of mother's education on weekly study time [$ab = 0.063$, distribution of products 95% CI = (0.013, 0.122), $p < 0.05$].
- In addition to the indirect effect that was detected, there is also a significant direct effect of mother's education (Medu) on the outcome (G3) that captures the mother's education's effect on final grade over and above its indirect effect through weekly study time [$C' = 0.621$, $t(646) = 5.843$, $p < 0.05$].

Bootstrapping

```
mediation(x = perf$Medu,
          mediator = perf$studytime,
          dv = perf$G3,
          bootstrap = TRUE)
```

```
## [1] "Bootstrap resampling has begun. This process may take a considerable amount of time if the number of
```

```
##
##      Estimate CI.Lower_Percentile
## Indirect.Effect      0.0631574307      1.461964e-02
## Indirect.Effect.Partially.Standardized 0.0195494123      4.608304e-03
## Index.of.Mediation    0.0221798248      5.175220e-03
## R2_4.5                0.0106081466      2.268614e-03
## R2_4.6                0.0005171401      2.908631e-05
## R2_4.7                0.0047245253      2.762089e-04
## Ratio.of.Indirect.to.Total.Effect      0.0923579217      2.253448e-02
## Ratio.of.Indirect.to.Direct.Effect      0.1017558836      2.305399e-02
## Success.of.Surrogate.Endpoint      9.6417408970      5.090043e+00
## Residual.Based_Gamma    0.0099471838      6.417600e-04
## Residual.Based.Standardized_gamma      0.0048170057     -1.393594e-02
## SOS                   0.1839380572      4.539826e-02
##
##      CI.Upper_Percentile  CI.Lower_BCa
## Indirect.Effect      0.119617463  1.718494e-02
## Indirect.Effect.Partially.Standardized 0.036637610  5.237099e-03
## Index.of.Mediation    0.041791417  5.998211e-03
## R2_4.5                0.021950351  2.693168e-03
## R2_4.6                0.001817683  3.768395e-05
## R2_4.7                0.014549449  3.498547e-04
```

```

## Ratio.of.Indirect.to.Total.Effect          0.181785188  2.621935e-02
## Ratio.of.Indirect.to.Direct.Effect          0.222172938  2.692614e-02
## Success.of.Surrogate.Endpoint              35.235927362  5.144261e+00
## Residual.Based_Gamma                      0.019934194  2.501036e-03
## Residual.Based.Standardized_gamma          0.015615637 -1.124973e-02
## SOS                                         0.346675236  5.064799e-02
##
## CI.Upper_BCa
## Indirect.Effect                           0.122638309
## Indirect.Effect.Partially.Standardized     0.037532649
## Index.of.Mediation                        0.042706637
## R2_4.5                                    0.023097551
## R2_4.6                                    0.001886525
## R2_4.7                                    0.015366970
## Ratio.of.Indirect.to.Total.Effect          0.187408946
## Ratio.of.Indirect.to.Direct.Effect          0.230631443
## Success.of.Surrogate.Endpoint              36.570040304
## Residual.Based_Gamma                      0.023662597
## Residual.Based.Standardized_gamma          0.017177696
## SOS                                         0.352943444

```

- With results leading to the same inference statistically (statistically significant indirect effect) with the same point estimate ($ab = 0.0632$) and different 95% confidence interval values (0.013, 0.118) for regular bootstrapped estimate and (0.015, 0.121) for the accelerated, bias-corrected bootstrapped estimate.