Mediation Analysis for Binary Outcome in R

Sample Data

```
# set the size of the sample
n=2000
# set seed
set.seed(123)

# simulate x (normal distribution; mean = 1, SD=1)
X <- rnorm(n,1,1)

# simulate M, linear relationship with X
M<-0.3+0.8*X+rnorm(n)

# simulate Y
Y_temp<-0.5+0.66*X+0.5*M
p <- 1/(1 + exp(-Y_temp))
Y<-rbinom(n, size = 1, prob = p)

# combine into a dataframe and print out the first 6 rows
df <- data.frame(X=X, M=M, Y=Y)
head(df)
```

```
##           X         M Y
## 1 0.4395244 0.1400158 1
## 2 0.7698225 1.1527959 1
## 3 2.5587083 1.8053775 1
## 4 1.0705084 2.3756344 1
## 5 1.1292877 1.3775661 1
## 6 2.7150650 1.8567837 1
```

a and b paths

```
# a path
a_path<-lm(M~X, data=df)
summary(a_path)
```

```
##
## Call:
## lm(formula = M ~ X, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1161 -0.6343 -0.0159  0.6644  3.4210
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.29854    0.03161   9.444   <2e-16 ***
## X            0.78719    0.02203  35.740   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9853 on 1998 degrees of freedom
## Multiple R-squared:   0.39,  Adjusted R-squared:  0.3897
## F-statistic:  1277 on 1 and 1998 DF,  p-value: < 2.2e-16
```

```
# b path
b_path <- glm(Y ~X+M, data=df,family = "binomial")
summary(b_path)
```

```
##
## Call:
## glm(formula = Y ~ X + M, family = "binomial", data = df)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.7961   0.2016   0.4002   0.6242   1.8259
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.49941    0.07539   6.625 3.48e-11 ***
## X            0.73248    0.08577   8.540  < 2e-16 ***
## M            0.57406    0.06737   8.521  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1944.9  on 1999  degrees of freedom
## Residual deviance: 1565.3  on 1997  degrees of freedom
## AIC: 1571.3
##
## Number of Fisher Scoring iterations: 5
```

Apply the function (Wrong way)

```
library(mediation)
```

```
## Warning: package 'mediation' was built under R version 4.1.3
```

```
## Loading required package: MASS
```

```
## Loading required package: Matrix
```

```
## Loading required package: mvtnorm
```

```
## Loading required package: sandwich
```

```
## Warning: package 'sandwich' was built under R version 4.1.2
```

```
## mediation: Causal Mediation Analysis
## Version: 4.5.0
```

```
Mediation_results <- mediate(a_path, b_path, sims=1000, boot=TRUE,treat="X", mediator="M")
```

```
## Running nonparametric bootstrap
```

```
summary(Mediation_results)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##                          Estimate 95% CI Lower 95% CI Upper p-value
## ACME (control)             0.0896       0.0721         0.11  <2e-16 ***
## ACME (treated)             0.0637       0.0468         0.08  <2e-16 ***
## ADE (control)              0.1376       0.1043         0.17  <2e-16 ***
## ADE (treated)              0.1117       0.0827         0.14  <2e-16 ***
## Total Effect               0.2014       0.1725         0.23  <2e-16 ***
## Prop. Mediated (control)   0.4451       0.3514         0.54  <2e-16 ***
## Prop. Mediated (treated)   0.3165       0.2289         0.41  <2e-16 ***
## ACME (average)             0.0767       0.0596         0.09  <2e-16 ***
## ADE (average)              0.1247       0.0935         0.16  <2e-16 ***
## Prop. Mediated (average)   0.3808       0.2901         0.48  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 2000
##
##
## Simulations: 1000
```

Apply the function (Correct way)

```
mean(X)
```

```
## [1] 1.029297
```

```
sd(X)
```

```
## [1] 1.000561
```

```
# define a control condition value
X_control=mean(X)-sd(X)

# define a treatment condition value
X_treatment=mean(X)+sd(X)
```

```
library(mediation)
Mediation_results <- mediate(a_path, b_path, sims=1000, boot=TRUE, treat="X",
                             mediator="M",control.value=X_control,treat.value = X_treatment)
```

```
## Running nonparametric bootstrap
```

```
summary(Mediation_results)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##                          Estimate 95% CI Lower 95% CI Upper p-value
## ACME (control)             0.1586       0.1289         0.19  <2e-16 ***
## ACME (treated)             0.0621       0.0426         0.09  <2e-16 ***
## ADE (control)              0.2242       0.1816         0.27  <2e-16 ***
## ADE (treated)              0.1278       0.0946         0.17  <2e-16 ***
## Total Effect               0.2864       0.2569         0.33  <2e-16 ***
## Prop. Mediated (control)   0.5539       0.4566         0.65  <2e-16 ***
## Prop. Mediated (treated)   0.2170       0.1410         0.32  <2e-16 ***
## ACME (average)             0.1104       0.0869         0.14  <2e-16 ***
## ADE (average)              0.1760       0.1400         0.22  <2e-16 ***
## Prop. Mediated (average)   0.3854       0.3015         0.49  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 2000
##
##
## Simulations: 1000
```

Double Check

```
  # a path
a_path <- lm(M ~ X, data = df)
a_0<-a_path$coefficients[1]
a_1<-a_path$coefficients[2]

  # b path
b_path <- glm(Y ~X+M, data=df,family = "binomial")
b_0<-b_path$coefficients[1]
b_1<-b_path$coefficients[3]
c_1_apostrophe<-b_path$coefficients[2]

# calculate estimated mediator for control and treatment conditions
M_est_control=a_0+a_1*X_control
M_est_treatment=a_0+a_1*X_treatment

# Indirect Effect - control
IE_control=1/(1+exp(-(b_0+b_1*M_est_treatment+c_1_apostrophe*X_control)))-
  1/(1+exp(-(b_0+b_1*M_est_control+c_1_apostrophe*X_control)))
print(IE_control)
```

```
## (Intercept)
##   0.1640327
```

```
# Indirect Effect - treatment
IE_treatment=1/(1+exp(-(b_0+b_1*M_est_treatment+c_1_apostrophe*X_treatment)))-
  1/(1+exp(-(b_0+b_1*M_est_control+c_1_apostrophe*X_treatment)))
print(IE_treatment)
```

```
## (Intercept)
##  0.05826663
```