

# Mediation Lab - Regression approach for causal mediation

12/07/2021

```
#####  
## Statistical Methods for Causal Inference ##  
## R Application ##  
## Lung cancer example ##  
#####  
  
## load in data  
data_sens <- read.csv("sim_data_lung.csv")  
  
#####  
### QUESTION 1C 2 AND 4#####  
#####  
  
#install.packages("CMAverse")  
library(CMAverse)  
  
## run mediation analysis with interaction using cmest()  
mediation.int.m1 <- cmest(data = data_sens, model = "rb", casecontrol = TRUE, yrare = TRUE,  
  outcome = "case",  
  exposure = "snp", mediator = "smoking", EMint = TRUE,  
  prec = c("sex", "colgrad", "age"),  
  mreg = list("logistic"), yreg = "logistic",  
  a = 1, astar = 0, mval = list(1),  
  estimation = "paramfunc", inference = "delta", full=FALSE)  
  
mediation.int.m0 <- cmest(data = data_sens, model = "rb", casecontrol = TRUE, yrare = TRUE,  
  outcome = "case",  
  exposure = "snp", mediator = "smoking", EMint = TRUE,  
  prec = c("sex", "colgrad", "age"),  
  mreg = list("logistic"), yreg = "logistic",  
  a = 1, astar = 0, mval = list(0),  
  estimation = "paramfunc", inference = "delta", full=FALSE)  
  
summary(mediation.int.m1)  
  
## Causal Mediation Analysis  
##  
## # Outcome regression:  
##  
## Call:  
## glm(formula = case ~ snp + smoking + snp * smoking, family = binomial(),  
## data = getCall(x$reg.output$yreg)$data, weights = getCall(x$reg.output$yreg)$weights)  
##  
## Deviance Residuals:
```

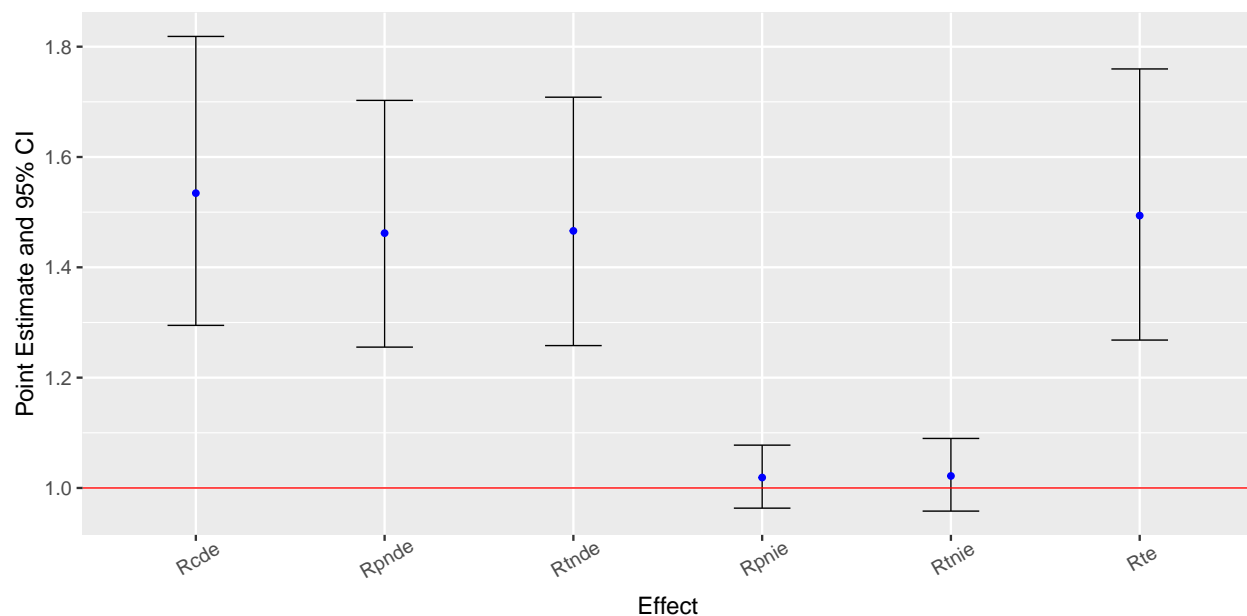
```

##      Min      1Q   Median      3Q      Max
## -1.5152 -1.3239  0.8738   0.8738   1.6529
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.07158    0.11586  -9.249  <2e-16 ***
## snp          0.07014    0.15581   0.450  0.6526
## smoking      1.40955    0.13378  10.536  <2e-16 ***
## snp:smoking  0.35804    0.17829   2.008  0.0446 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 4547.5  on 3299  degrees of freedom
## Residual deviance: 4138.8  on 3296  degrees of freedom
## AIC: 4146.8
##
## Number of Fisher Scoring iterations: 4
##
##
## # Mediator regressions:
##
## Call:
## glm(formula = smoking ~ snp, family = binomial(), data = getCall(x$reg.output$mreg[[1L]])$data,
##      weights = getCall(x$reg.output$mreg[[1L]])$weights)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.325  -1.295   1.037   1.065   1.065
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.27128    0.07769   3.492  0.00048 ***
## snp          0.06893    0.10500   0.656  0.51152
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 2044.0  on 1499  degrees of freedom
## Residual deviance: 2043.6  on 1498  degrees of freedom
## AIC: 2047.6
##
## Number of Fisher Scoring iterations: 4
##
##
## # Effect decomposition on the risk ratio scale for a case control study via the regression-based app
##
## Closed-form parameter function estimation with
## delta method standard errors, confidence intervals and p-values
##
##      Estimate Std.error 95% CIL 95% CIU    P.val
## Rcde    1.53446   0.13300 1.29473   1.819 7.80e-07 ***

```

```
## Rpnde 1.46196 0.11365 1.25535 1.703 1.03e-06 ***
## Rtnde 1.46608 0.11443 1.25811 1.708 9.50e-07 ***
## Rpnie 1.01890 0.02913 0.96338 1.078 0.512
## Rtnie 1.02177 0.03358 0.95802 1.090 0.512
## Rte 1.49379 0.12483 1.26811 1.760 1.57e-06 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Rcde: controlled direct effect risk ratio; Rpnde: pure natural direct effect risk ratio; Rtnde: tot
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $yval
## [1] "1"
##
## $mval
## $mval[[1]]
## [1] 1
```

```
ggcmest(mediation.int.m1) +
  ggplot2::theme(axis.text.x = ggplot2::element_text(angle = 30, vjust = 0.8))
```



```
summary(mediation.int.m0)
```

```
## Causal Mediation Analysis
##
## # Outcome regression:
##
## Call:
## glm(formula = case ~ snp + smoking + snp * smoking, family = binomial(),
```

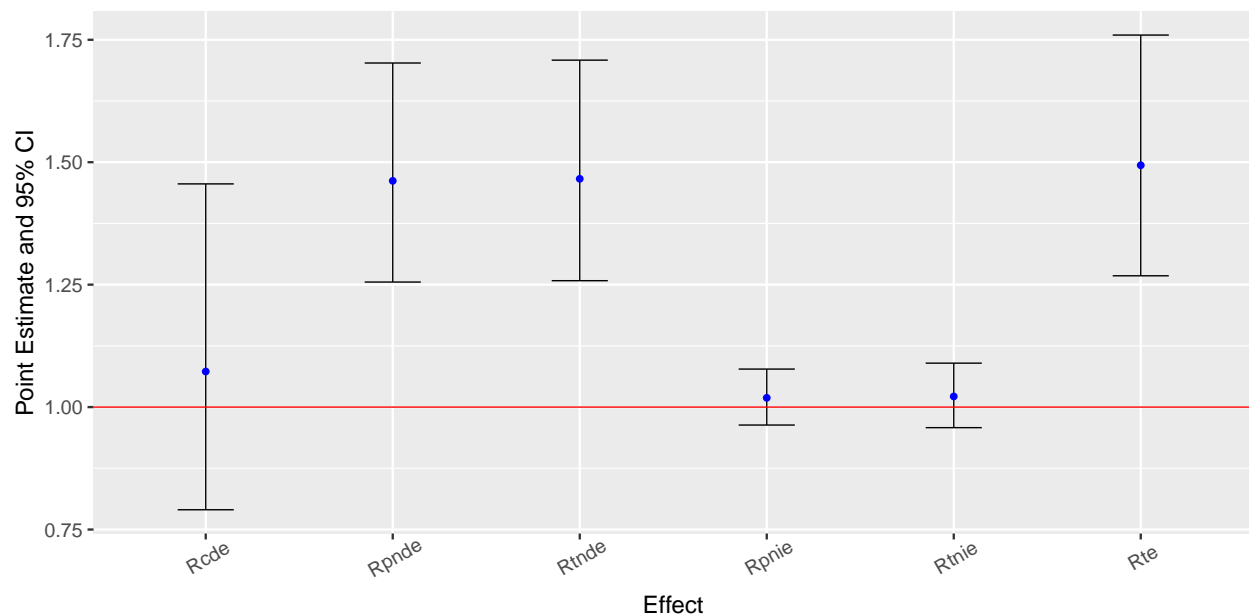
```

##      data = getCall(x$reg.output$yreg)$data, weights = getCall(x$reg.output$yreg)$weights)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5152  -1.3239   0.8738   0.8738   1.6529
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.07158    0.11586  -9.249  <2e-16 ***
## snp          0.07014    0.15581   0.450  0.6526
## smoking      1.40955    0.13378  10.536  <2e-16 ***
## snp:smoking  0.35804    0.17829   2.008  0.0446 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 4547.5  on 3299  degrees of freedom
## Residual deviance: 4138.8  on 3296  degrees of freedom
## AIC: 4146.8
##
## Number of Fisher Scoring iterations: 4
##
##
## # Mediator regressions:
##
## Call:
## glm(formula = smoking ~ snp, family = binomial(), data = getCall(x$reg.output$mreg[[1L]])$data,
##      weights = getCall(x$reg.output$mreg[[1L]])$weights)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.325  -1.295   1.037   1.065   1.065
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.27128    0.07769   3.492  0.00048 ***
## snp          0.06893    0.10500   0.656  0.51152
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 2044.0  on 1499  degrees of freedom
## Residual deviance: 2043.6  on 1498  degrees of freedom
## AIC: 2047.6
##
## Number of Fisher Scoring iterations: 4
##
##
## # Effect decomposition on the risk ratio scale for a case control study via the regression-based app
##
## Closed-form parameter function estimation with
## delta method standard errors, confidence intervals and p-values

```

```
##
##      Estimate Std.error 95% CIL 95% CIU    P.val
## Rcde    1.07265   0.16713 0.79038   1.456    0.653
## Rpnede   1.46196   0.11365 1.25535   1.703 1.03e-06 ***
## Rtnede   1.46608   0.11443 1.25811   1.708 9.50e-07 ***
## Rpnide   1.01890   0.02913 0.96338   1.078    0.512
## Rtnide   1.02177   0.03358 0.95802   1.090    0.512
## Rte     1.49379   0.12483 1.26811   1.760 1.57e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Rcde: controlled direct effect risk ratio; Rpnede: pure natural direct effect risk ratio; Rtnede: tot
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $yval
## [1] "1"
##
## $mval
## $mval[[1]]
## [1] 0
```

```
ggcmest(mediation.int.m0) +
  ggplot2::theme(axis.text.x = ggplot2::element_text(angle = 30, vjust = 0.8))
```



```
## CDE(1) = 1.53
## CDE(0) = 1.07
## NIE = 1.02
## NDE = 1.45
```