Other Models Analyzed by R Package "Mediation"

July 6, 2017 H. Seltman

I. Moderated mediation

a. Definition: the magnitude of the Average Causal Mediated Effect depends on (i.e., varies with) a pre-treatment covariate

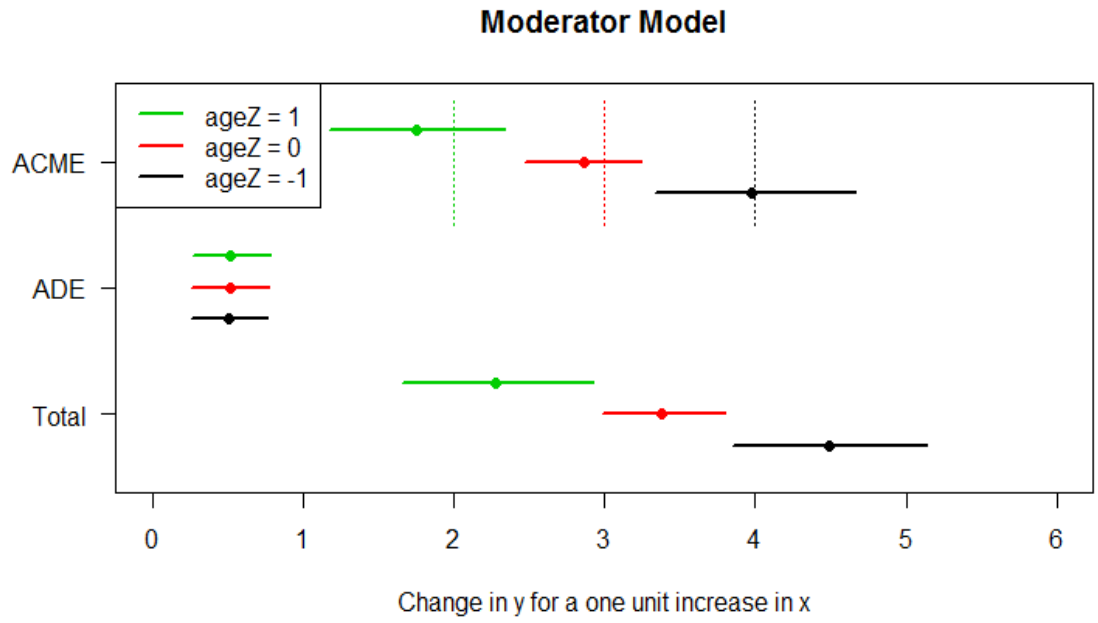
b. First approach: Include `moderator*treatment` in both the mediator and the outcome models. Use the `covariates=list()` parameter of the mediate() function of the "mediation" package to specify at what level of the moderator we want to evaluate the ACME.

c. Second approach: Directly test the difference in ACME at two levels of a covariate using the `test.modmed()` function.

d. Example from simModerate.R

```
# E.g., simulate data where older people have less effect
# of x on m
N = 250
ageZ = rnorm(N)
bxm = 1.5
bxagem = -0.5
x = rnorm(N)
m = rnorm(N, 5 + bxm*x + bxagem*x*ageZ, 1.5)
bmy = 2
y = rnorm(N, 20 + 0.5*x + bmy*m, 1.5)
dtfMod = data.frame(ageZ, x, m, y)

# Analyze at three different ages
library(mediation)
mModMod = lm(m ~ x*ageZ, dtfMod)
mModOut = lm(y ~ x + m*ageZ, dtfMod)
mModL = mediate(mModMod, mModOut, treat="x", mediator="m",
                covariates = list(ageZ=-1))
print(summary(mModL))

Causal Mediation Analysis
Quasi-Bayesian Confidence Intervals
(Inference Conditional on the Covariate Values Specified in
`covariates')
               Estimate 95% CI Lower 95% CI Upper p-value
ACME              3.975        3.339         4.66  <2e-16 ***
ADE               0.510        0.265         0.76  <2e-16 ***
Total Effect      4.484        3.854         5.14  <2e-16 ***
Prop. Mediated    0.886        0.826         0.94  <2e-16 ***

mModM = mediate(mModMod, mModOut, treat="x", mediator="m",
                covariates=list(ageZ=0))
print(summary(mModM))

               Estimate 95% CI Lower 95% CI Upper p-value
ACME              2.861        2.476         3.25  <2e-16 ***
ADE               0.515        0.266         0.77  <2e-16 ***
Total Effect      3.376        2.984         3.80  <2e-16 ***
Prop. Mediated    0.848        0.779         0.92  <2e-16 ***
```

```
mModH = mediate(mModMod, mModOut, treat="x", mediator="m",
                covariates = list(ageZ=+1))

print(summary(mModH))

               Estimate 95% CI Lower 95% CI Upper p-value
ACME              1.758        1.177         2.34  <2e-16 ***
ADE               0.518        0.271         0.79  <2e-16 ***
Total Effect      2.276        1.658         2.93  <2e-16 ***
Prop. Mediated    0.774        0.646         0.88  <2e-16 ***

# Combined plot
plot(NA, xlab="Change in y for a one unit increase in x",
     ylab="", xlim=c(0, 6), ylim=c(0.5,3.5), axes=FALSE,
     main="Moderator Model")
box(); axis(1)
axis(2, at=3:1, labels=c("ACME","ADE","Total"), las=1)
results = list(mModL, mModM, mModH)
ageZVals = sapply(results, function(x) x$covariates$ageZ)
nAgeZVals = length(ageZVals)
for (i in 1:nAgeZVals) {
  result = results[[i]]
  lines(result$d0.ci, c(2.5, 2.5)+0.25*i, col=i, lwd=2)
  points(result$d0, 2.5+0.25*i, col=i, pch=16)
  lines(result$z0.ci, c(1.5, 1.5)+0.25*i, col=i, lwd=2)
  points(result$z0, 1.5+0.25*i, col=i, pch=16)
  lines(result$tau.ci, c(0.5, 0.5)+0.25*i, col=i, lwd=2)
  points(result$tau.coef, 0.5+0.25*i, col=i, pch=16)
  lines(rep((bxm+ageZVals[i]*bxagem)*bmy, 2), c(2.5,3.5),
        col=i, lty=3)
}
legend("topleft", paste("ageZ =", rev(ageZVals)),
       col=rev(1:nAgeZVals), lwd=2)
```



```
# Alternate analysis: gives p-value and CI for moderation of
# mediation at two chosen levels of the moderator
mMod = mediate(mModMod, mModOut, treat="x", mediator="m")
aa = test.modmed(mMod, covariates.1=list(ageZ=-1),
                 covariates.2=list(ageZ=+1))
print(aa)

      Test of ACME(covariates.1) - ACME(covariates.2) = 0
data:  estimates from mMod
ACME(covariates.1) - ACME(covariates.2) = 2.2258,
  p-value < 2.2e-16
alternative hypothesis: true ACME(covariates.1) -
                        ACME(covariates.2) is not equal to 0
95 percent confidence interval: 1.320358 3.133411

      Test of ADE(covariates.1) - ADE(covariates.2) = 0
data:  estimates from mMod
ADE(covariates.1) - ADE(covariates.2) = -0.008458,
  p-value = 0.938
alternative hypothesis: true ADE(covariates.1) –
                        ADE(covariates.2) is not equal to 0
95 percent confidence interval: -0.3588943  0.3503244
```

II. Hierarchical mediation

a. Two levels: level 1 (e.g., student) is nested in level 2(e.g., classroom)

b. X is assigned at level 2 (classroom)

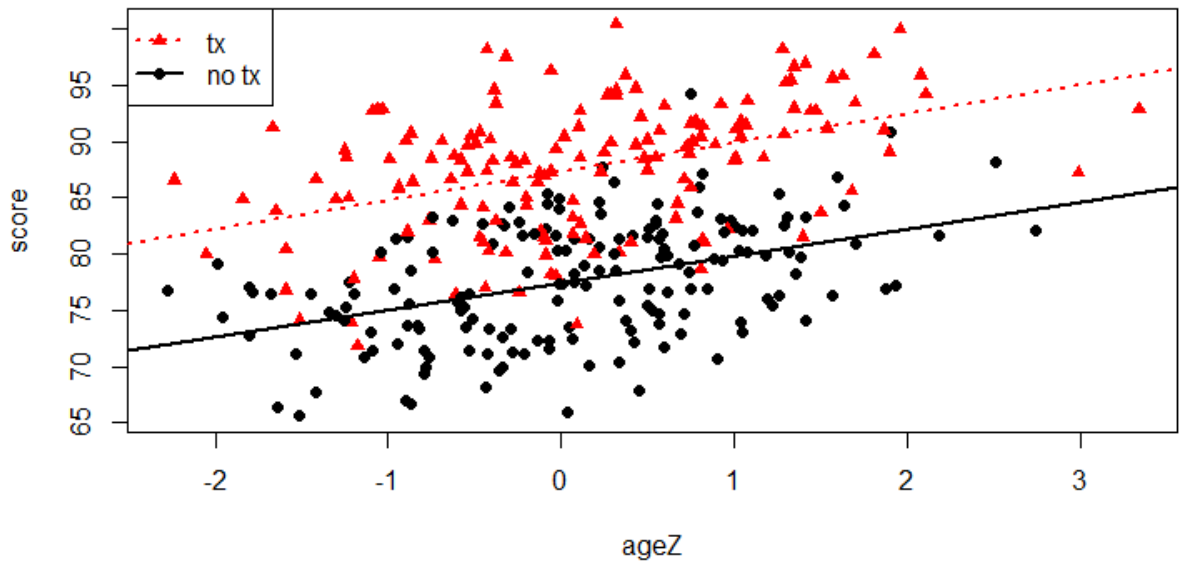
c. Y is measured at level 1 (student)

d. M is measured at level 1 or level 2

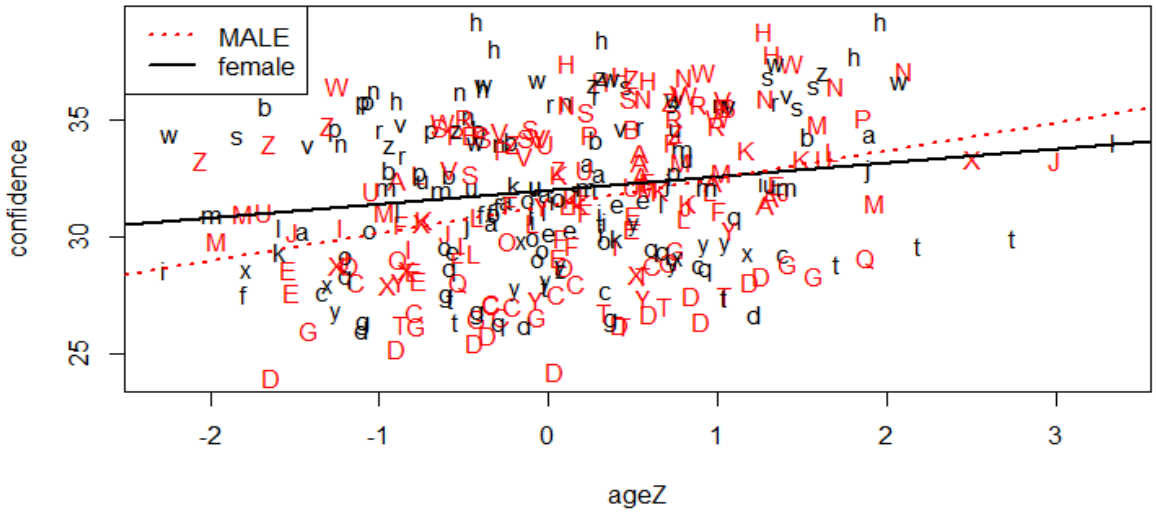
e. Example 1 from simHierarchical.R

```
# Simulate data with mediator measured at individual level (level 1)
M = 26 # classrooms
n = sample(10:15, M, replace=TRUE) # students per classroom
N = sum(n)
bTxM = 4
bMY = 2

class = rep(LETTERS[1:M], n)
ageZ = round(rnorm(N), digits=2)
male = rbinom(N, 1, 0.5)
teacherExperZ = rep(round(rnorm(M), digits=1), n)
tx = rep(sample(rep(0:1, c(M/2, M/2))), n)
RIM = rep(rnorm(M, 0, 1.5), n)
# mediator:
confidence = round(rnorm(N, 30 + bTxM*tx + 0.8*ageZ +
                           1.0*teacherExperZ + RIM, 1),
                   digits=2)
RIY = rep(rnorm(M, 0, 1), n)
score = round(rnorm(N, 20 + bMY*confidence + 1*tx - 3*male +
                       1*ageZ + RIY,
                    sd=1.3),
              digits=2)
dtf = data.frame(class, ageZ, male, teacherExperZ, tx, confidence,
                 score)
```

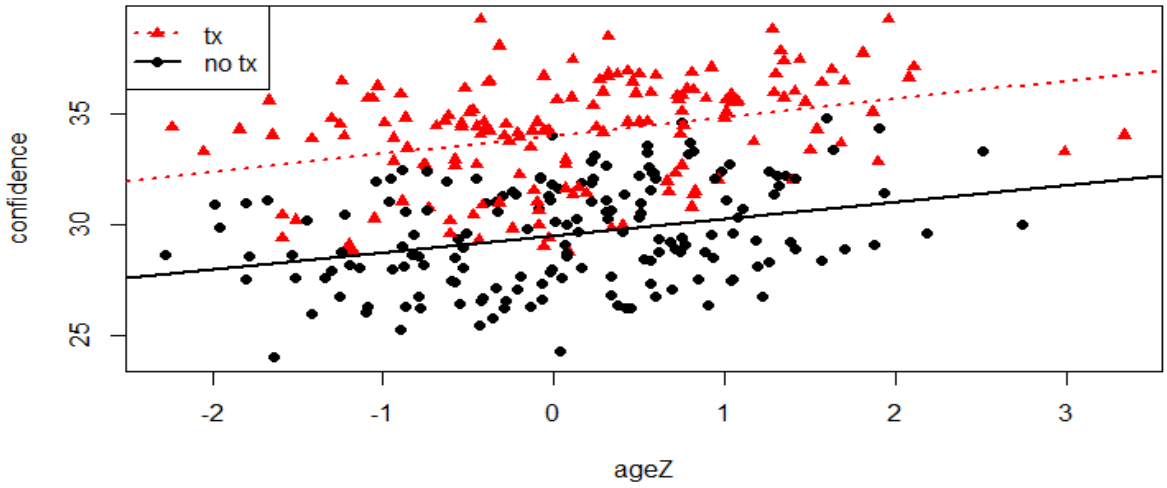
```
head(dtf)
  class  ageZ male teacherExperZ tx confidence score
1     A -0.24    0          -0.1  0      31.53 82.77
2     A  1.91    0          -0.1  0      34.41 90.88
3     A  0.56    1          -0.1  0      32.65 84.49
4     A  0.31    0          -0.1  0      32.67 86.41
5     A  1.29    1          -0.1  0      31.41 82.50
6     A  0.55    1          -0.1  0      33.26 82.96
> tail(dtf)
    class  ageZ male teacherExperZ tx confidence score
322     Z -0.20    1           0.7  1      34.00 84.98
323     Z -0.94    0           0.7  1      33.92 85.95
324     Z  0.07    1           0.7  1      32.94 81.73
325     Z  0.32    0           0.7  1      36.87 94.08
326     Z  0.76    1           0.7  1      34.50 89.96
327     Z -2.05    1           0.7  1      33.30 79.91

### EDA ###
# score on age and treatment
plot(score ~ ageZ, pch=16+(tx==1), col=1+tx, data=dtf)
abline(lm(score~ageZ, data=dtf[tx==0,]), lwd=2)
abline(lm(score~ageZ, data=dtf[tx==1,]), col=2, lwd=2, lty=3)
legend("topleft", c("tx", "no tx"), lwd=2, col=2:1, pch=17:16,
       lty=c(3,1))
```



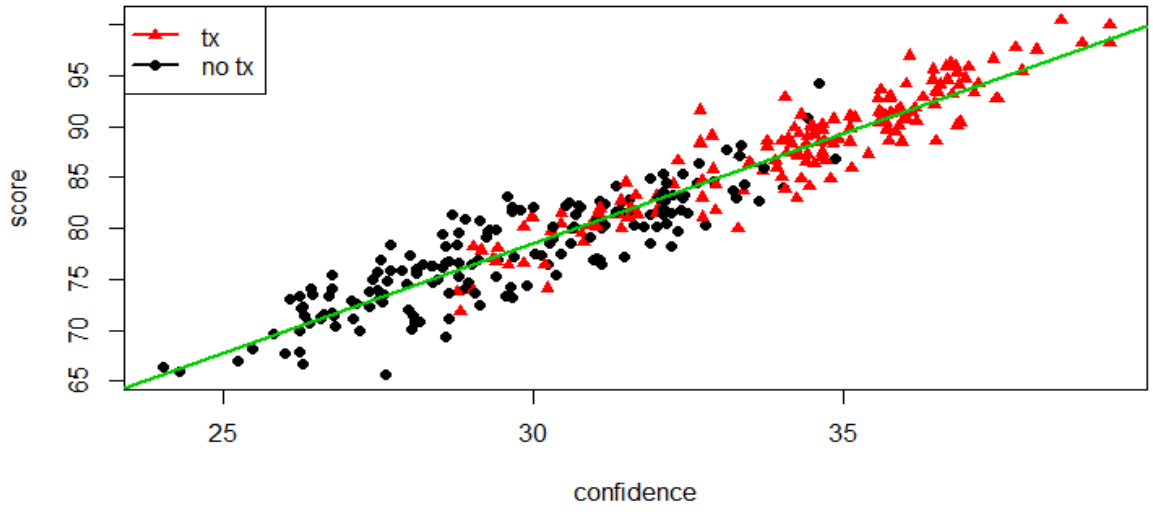
```
# mediator (confidence) on age and gender
plot(confidence ~ ageZ, col=1+male,
     pch=ifelse(male==1, LETTERS[as.numeric(class)],
                         letters[as.numeric(class)]),
     data=dtf)
abline(lm(confidence~ageZ, data=dtf[male==0,]), lwd=2)
abline(lm(confidence~ageZ, data=dtf[male==1,]), col=2, lwd=2, lty=3)
coef(summary(lm(confidence ~ ageZ*male, data=dtf))) # p.IA=0.10
legend("topleft", c("MALE", "female"), lwd=2, col=2:1, lty=c(3, 1))
```




```
# mediator (confidence) on age and treatment
plot(confidence ~ ageZ, pch=16+(tx==1), col=1+tx, data=dtf)
abline(lm(confidence~ageZ, data=dtf[tx==0,]), lwd=2)
abline(lm(confidence~ageZ, data=dtf[tx==1,]), col=2, lwd=2, lty=3)
legend("topleft", c("tx", "no tx"), lwd=2, col=2:1, pch=17:16,
       lty=c(3,1))
```

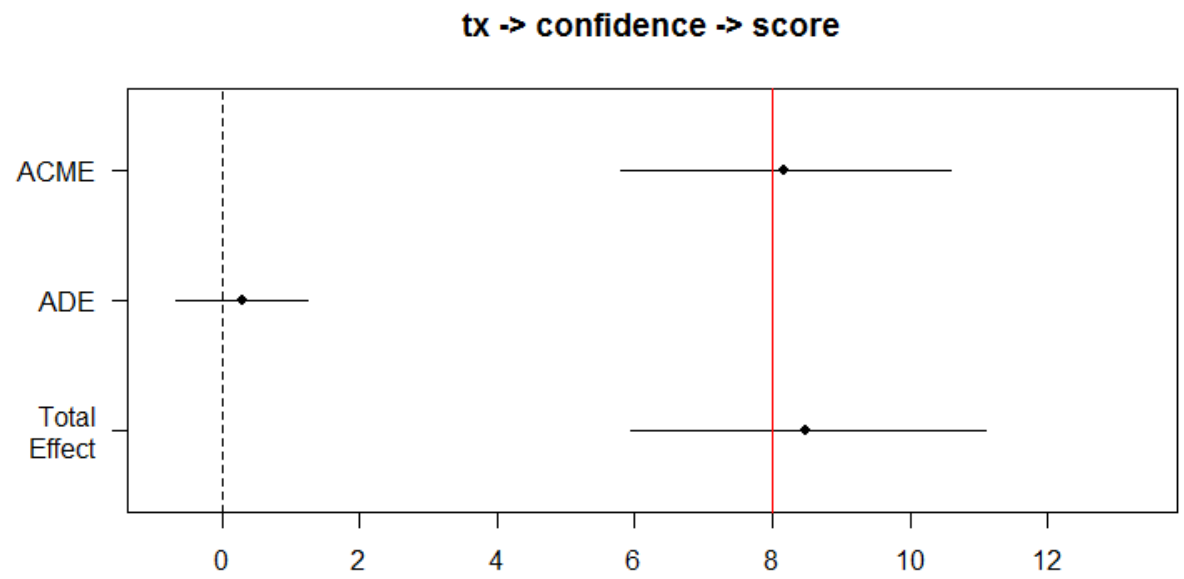


```
# score on mediator with treatment
plot(score ~ confidence, pch=16+(tx==1), col=1+tx, data=dtf)
abline(lm(score~confidence, data=dtf), lwd=2, col=3)
legend("topleft", c("tx", "no tx"), lwd=2, col=2:1, pch=17:16)
```

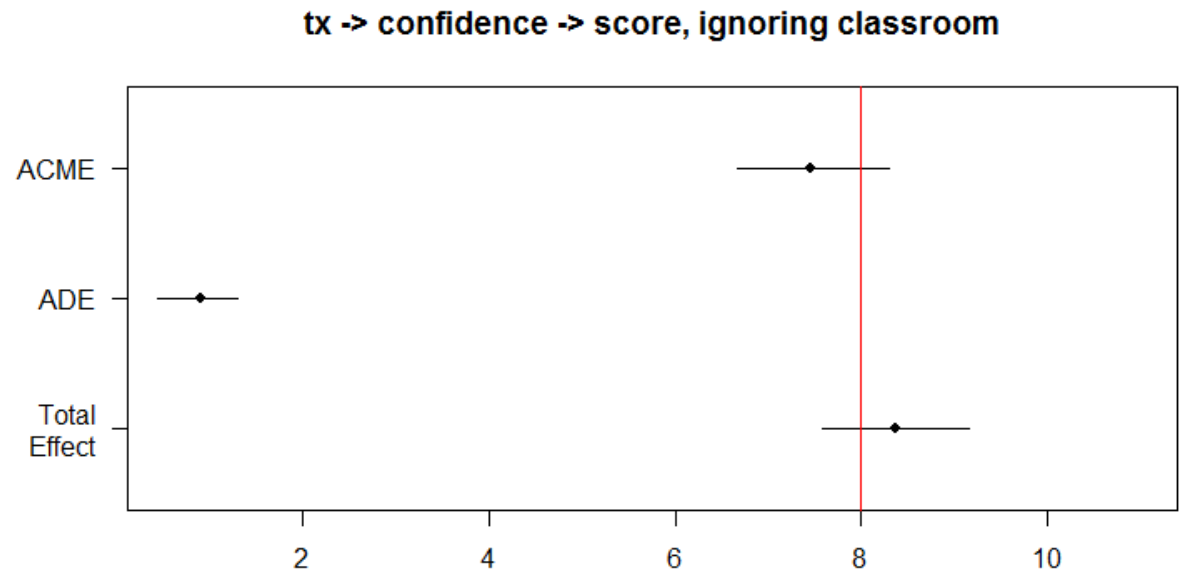



```
# Modeling
library("mediation") # for mediate()
library("lme4") # for lmer() (mixed models)

med.fit = lmer(confidence ~ tx + male + ageZ + teacherExperZ +
                 (1|class), data=dtf)
summary(med.fit)

Linear mixed model fit by REML ['lmerMod']
Formula: confidence ~ tx + male + ageZ + teacherExperZ + (1 | class)
   Data: dtf
REML criterion at convergence: 1068.8

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.78867 -0.64108 -0.04233  0.66539  2.84152

Random effects:
 Groups   Name        Variance Std.Dev.
 class    (Intercept) 1.539    1.240
 Residual             1.065    1.032
Number of obs: 341, groups:  class, 26

Fixed effects:
             Estimate Std. Error t value
(Intercept)   30.79761    0.35886   85.82
tx             3.51405    0.53054    6.62
male          -0.07478    0.11678   -0.64
ageZ           0.86705    0.06067   14.29
teacherExperZ  1.34514    0.32780    4.10
```

```
y.fit = lmer(score ~ tx + confidence + male + (1|class), data=dtf)
summary(y.fit)
Random effects:
 Groups   Name        Variance Std.Dev.
 class    (Intercept) 1.373    1.172
 Residual             1.758    1.326
Number of obs: 341, groups:  class, 26

Fixed effects:
            Estimate Std. Error t value
(Intercept)  9.39312    1.66868    5.63
tx           0.31834    0.50357    0.63
confidence   2.33703    0.05283   44.23
male        -3.15254    0.14987  -21.03

med.out = mediate(med.fit, y.fit, treat="tx", mediator="confidence")
print(summary(med.out))
Output Based on Overall Averages Across Groups
              Estimate 95% CI Lower 95% CI Upper p-value
ACME             8.177        5.790        10.59  <2e-16 ***
ADE              0.312       -0.666         1.26    0.51
Total Effect     8.489        5.944        11.10  <2e-16 ***
Prop. Mediated   0.962        0.857         1.09  <2e-16 ***

Sample Size Used: 341            Simulations: 1000

plot(med.out, main="tx -> confidence -> score")
abline(v=bTxM*bMY, col=2)
```



```
# Wrong analysis ignoring classroom
med.fit2 = lm(confidence ~ tx + male + ageZ + teacherExpZ, data=dtf)
y.fit2 = lm(score ~ tx + confidence + male, data=dtf)
med.out2 = mediate(med.fit2, y.fit2, treat="tx",
                   mediator="confidence")
plot(med.out2, main="tx -> confidence -> score, ignoring classroom")
abline(v=bTxM*bMY, col=2)
```

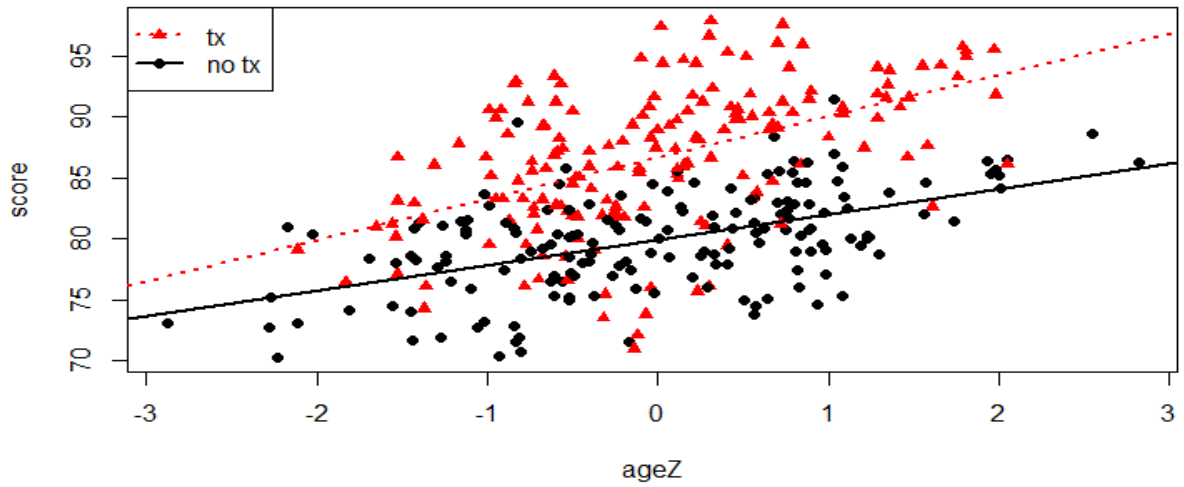


f. Example 2 from simHierarchical.R

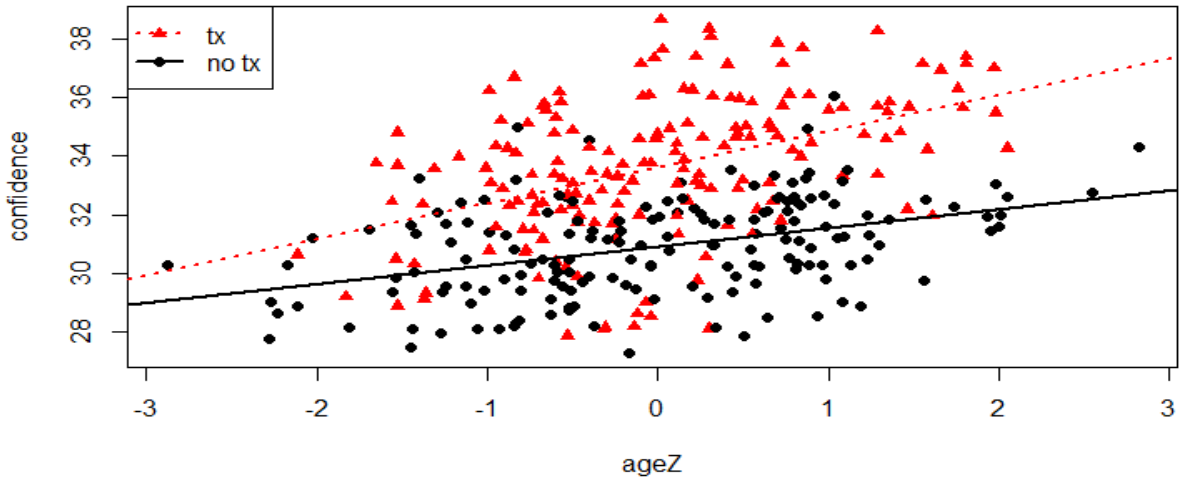
```
# Simulate data with moderator measured at level 2 (class level)
M = 46
n = sample(10:15, M, replace=TRUE)
N = sum(n)
bTxM = 4
bMY = 2

#class = rep(LETTERS[1:M], n) # Fails -- cannot be a factor!!
class = rep(1:M, n)
ageZ = round(rnorm(N), digits=2)
male = rbinom(N, 1, 0.5)
teacherExperZM = round(rnorm(M), digits=1)
teacherExperZ = rep(teacherExperZM, n)
txM = sample(rep(0:1, c(M/2, M/2)))
tx = rep(txM, n)
RIM = rnorm(M, 0, 1.5)
# mediator:
confidence = rep(round(rnorm(M, 30 + bTxM*txM +
                               1.0*teacherExperZM + RIM, 1),
                       digits=2),
                 times=n)
RIY = rep(rnorm(M, 0, 1), n)
```

```
score = round(rnorm(N, 20 + bMY*confidence + 1*tx - 3*male +
                       1*ageZ + RIY,
                    sd=1.3),
              digits=2)
dtf = data.frame(class, ageZ, male, teacherExperZ, tx, confidence,
                 score)
> head(dtfH2)
  class  ageZ male teacherExperZ tx confidence score
1     1 -1.02    0           0.1  0      33.43 86.25
2     1 -0.95    0           0.1  0      33.43 83.46
3     1  0.79    0           0.1  0      33.43 87.91
4     1  0.15    0           0.1  0      33.43 84.96
5     1 -1.43    1           0.1  0      33.43 82.20
6     1  0.46    0           0.1  0      33.43 88.66
> tail(dtfH2)
    class  ageZ male teacherExperZ tx confidence score
577    46 -1.34    0           0.9  0      28.85 75.72
578    46  2.05    1           0.9  0      28.85 74.99
579    46  1.11    1           0.9  0      28.85 75.19
580    46 -1.23    1           0.9  0      28.85 72.43
581    46  0.26    0           0.9  0      28.85 78.45
582    46 -0.21    0           0.9  0      28.85 76.44

dtfC = dtf[!duplicated(dtf$class), c("class","teacherExperZ","tx",
                                     "confidence")]
head(dtfC)
   class teacherExperZ tx confidence
1      1           0.1  0      33.43
12     2          -0.5  0      32.48
24     3           0.6  0      28.40
39     4           0.0  1      33.98
52     5          -0.3  0      30.85
66     6          -0.1  0      30.25

# EDA
```

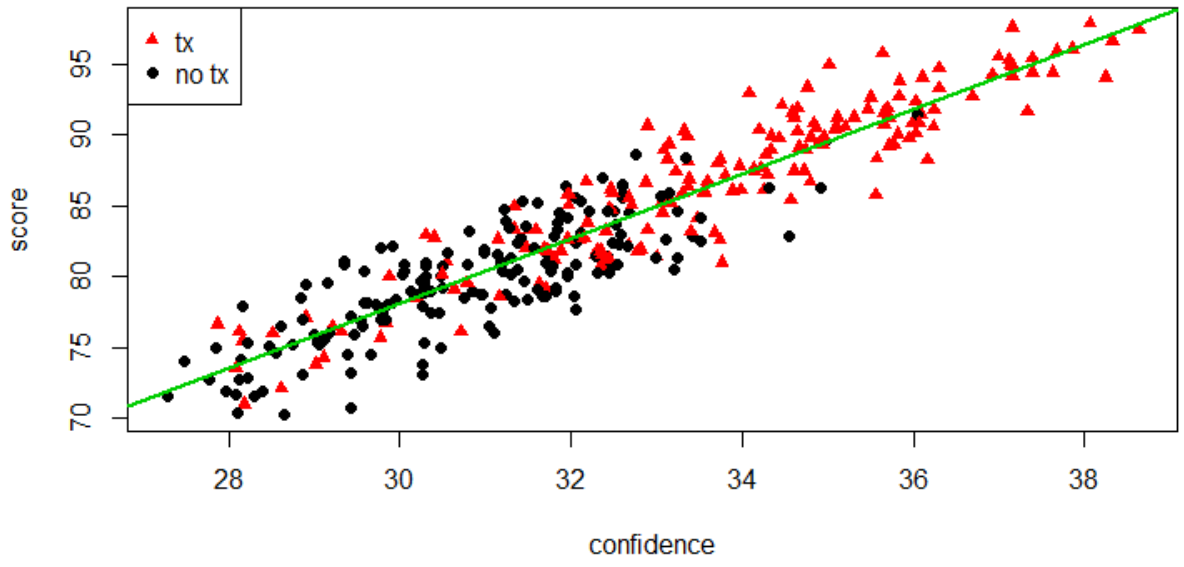
```
# score on age and treatment
plot(score ~ ageZ, pch=16+tx, col=1+tx, data=dtf)
abline(lm(score~ageZ, data=dtf[tx==0,]), lwd=2)
abline(lm(score~ageZ, data=dtf[tx==1,]), col=2, pch=17, lwd=2,
       lty=3)
legend("topleft", c("tx", "no tx"), lwd=2, col=2:1, pch=17:16,
       lty=c(3,1))
```




```
# mediator (confidence) on age and treatment
plot(confidence ~ ageZ, pch=16+tx, col=1+tx, data=dtf)
abline(lm(confidence~ageZ, data=dtf[tx==0,]), lwd=2)
abline(lm(confidence~ageZ, data=dtf[tx==1,]), col=2, pch=17, lwd=2,
       lty=3)
legend("topleft", c("tx", "no tx"), lwd=2, col=2:1, pch=17:16,
       lty=c(3,1))
```

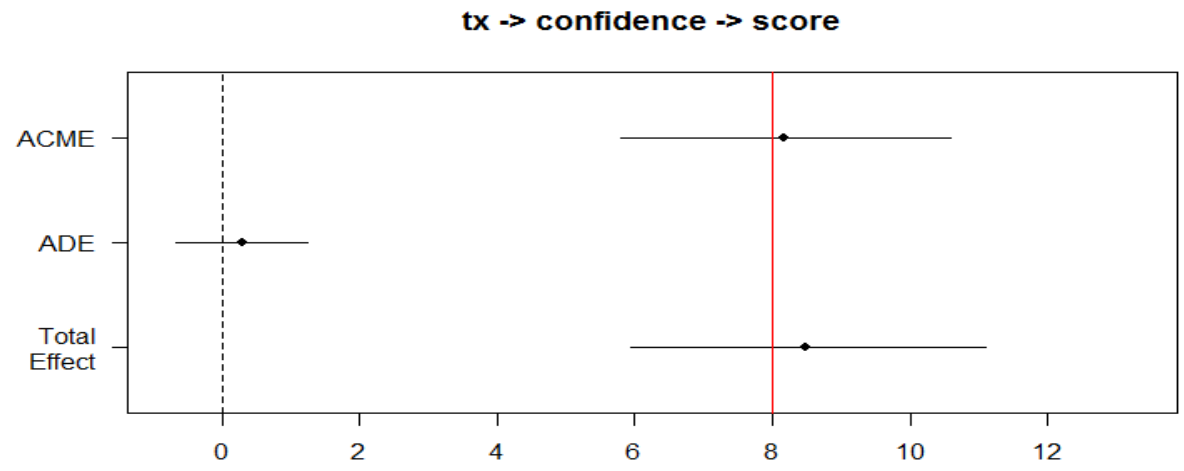


```
# score on mediator with treatment
plot(score ~ confidence, pch=16+tx, col=1+tx, data=dtf)
abline(lm(score~confidence, data=dtf), lwd=2, col=3)
legend("topleft", c("tx", "no tx"), col=2:1, pch=17:16)
```




```
# Modeling
library("mediation") # for mediate
library("lme4") # for lmer (mixed models)

# Mediator model is non-hierarchical with one line per
# group (using !duplicated())
med.fit = lm(confidence ~ tx + teacherExperZ, data=dtfC)
summary(med.fit)
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    30.9903     0.5433  57.044  < 2e-16 ***
tx              3.1014     0.8156   3.803 0.000917 ***
teacherExperZ   1.1125     0.5038   2.208 0.037462 *

Residual standard error: 1.958 on 43 degrees of freedom
Multiple R-squared:  0.4014,  Adjusted R-squared:  0.3494
F-statistic: 7.713 on 2 and 43 DF,  p-value: 0.002733
```

```
y.fit = lmer(score ~ tx + confidence + male + (1|class), data=dtf)
summary(y.fit)
Random effects:
 Groups   Name        Variance Std.Dev.
 class    (Intercept) 1.373    1.172
 Residual             1.758    1.326
Number of obs: 582, groups:  class, 46

Fixed effects:
            Estimate Std. Error t value
(Intercept)  9.39312    1.66868    5.63
tx           0.31834    0.50357    0.63
confidence   2.33703    0.05283   44.23
male        -3.15254    0.14987  -21.03

med.out = mediate(med.fit, y.fit, treat="tx", mediator="confidence")
print(summary(med.out))
               Estimate 95% CI Lower 95% CI Upper p-value
ACME              8.177        5.790        10.59  <2e-16 ***
ADE               0.312       -0.666         1.26    0.51
Total Effect      8.489        5.944        11.10  <2e-16 ***
Prop. Mediated    0.962        0.857         1.09  <2e-16 ***

plot(med.out, main="tx -> confidence -> score")
abline(v=bTxM*bMY, col=2)
```



```
# Wrong analysis ignoring classroom
med.fit2 = lm(confidence ~ tx + teacherExperZ, data=dtf)
y.fit2 = lm(score ~ tx + confidence + male, data=dtf)
med.out2 = mediate(med.fit2, y.fit2, treat="tx",
                   mediator="confidence")
plot(med.out2, main="tx -> confidence -> score, ignoring classroom")
abline(v=bTxM*bMY, col=2)
```

g. Summary

i. Moderated mediation occurs when the size of the mediated effect depend on the value of some other variable (neither treatment nor moderator)

ii. Hierarchical mediation occurs when unmeasured group level (level 2, e.g., classroom) variables are common across each set of level 1 (e.g., students) units. This affects all students in a given class similarly, and thus induces correlation that invalidates the usual calculation of standard errors (and therefore, CIs and p-values). Typically, treatments are applied at level 2 and observation of outcome are at level 1. The mediator may be measured at either level.