PSY.308d.DA5

**Prompt**

Parents with anxiety disorders tend to have children with anxiety disorders as well (see everything ever written by Bogels, Borelli, Wood, or Rapee). This has led some to hypothesize that anxiety is genetically transmitted. However, you think there might be something else going on. Parental overcontrol, the excess regulation of a child’s emotion, cognition, and behavior is strongly related to child anxiety. You wonder if anxious parents are more overcontrolling, leading to child anxiety. You decided to conduct a study to figure this out.

*Variables:* Parent\_Anx: 1-20 (higher scores indicating higher anxiety symptoms) Child\_Anx: 1-20 (higher scores indicating higher anxiety symptoms) Parent\_OC: 1-20 (higher scores indicating greater use of overcontrol)

**Research Question:** Does parental overcontrol mediate the relationship between parent anxiety and child anxiety?

**(1)** *General assignment:* Conduct the appropriate analysis using bootstrapping techniques. Report the proper assumptions and statistics in the results section. In the discussion section, provide a summary of what you found, discuss the implications, and give at least one limitation and future research direction. Don’t forget to include a table for descriptives, correlations, and regression models.

**(2)** *Conceptual component:* Conduct the same analysis, but use the Sobel test instead of bootstrapping to test the indirect effect. *Report the statistics in your results section.* Provide a summary of what of the findings were for this analysis and compare it to your bootstrapped analysis findings. If the two analyses found different outcomes, determine which results are more appropriate and justify your decision (Hint: What is a known issue with the Sobel test?).

#Prep

library(pacman)  
p\_load(psych, jmv, medmod, lavaan, multilevel)

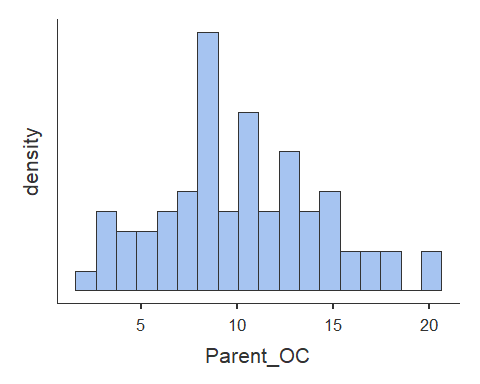
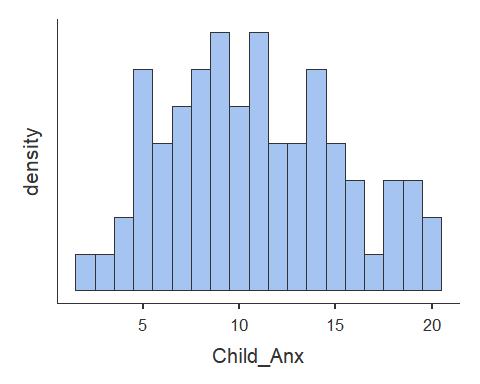
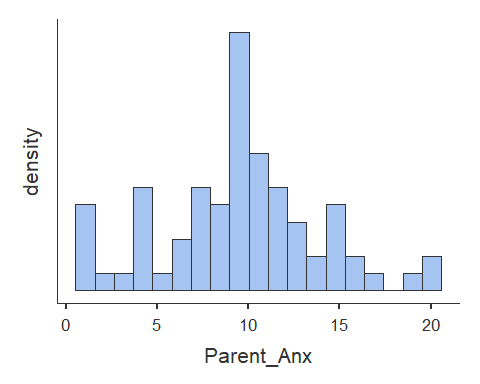
#Read in your data

datboot <- read.csv("https://www.dropbox.com/s/dsplwa2mppfhmu7/Psy.308d.DA5.csv?dl=1")

**Descriptives**

desc <- descriptives(datboot,   
 vars = c('Parent\_Anx', 'Child\_Anx', 'Parent\_OC'),   
 hist = TRUE,   
 sd = TRUE,   
 min = TRUE,   
 max = TRUE,   
 skew = TRUE,   
 kurt = TRUE)  
desc

##   
## DESCRIPTIVES  
##   
## Descriptives   
## ---------------------------------------------------------------   
## Parent\_Anx Child\_Anx Parent\_OC   
## ---------------------------------------------------------------   
## N 74 74 74   
## Missing 0 0 0   
## Mean 9.54 10.8 10.3   
## Median 10.0 10.5 10.5   
## Standard deviation 4.53 4.49 4.31   
## Minimum 1 2 2   
## Maximum 20 20 20   
## Skewness 0.0261 0.254 0.130   
## Std. error skewness 0.279 0.279 0.279   
## Kurtosis -0.195 -0.721 -0.542   
## Std. error kurtosis 0.552 0.552 0.552   
## ---------------------------------------------------------------



cat("\n")

corr <- corrMatrix(datboot,   
 vars = c('Parent\_Anx', 'Child\_Anx', 'Parent\_OC'),   
 flag = TRUE)  
corr

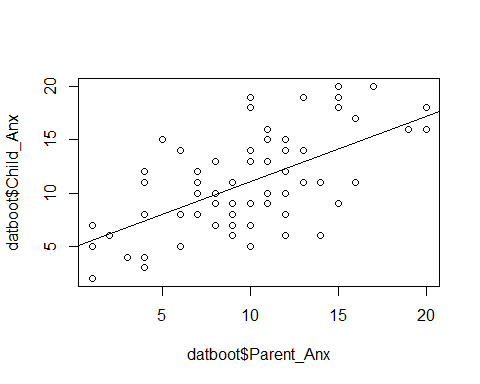
##   
## CORRELATION MATRIX  
##   
## Correlation Matrix   
## ---------------------------------------------------------------------   
## Parent\_Anx Child\_Anx Parent\_OC   
## ---------------------------------------------------------------------   
## Parent\_Anx Pearson's r — 0.618 0.602   
## p-value — < .001 < .001   
##   
## Child\_Anx Pearson's r — 0.675   
## p-value — < .001   
##   
## Parent\_OC Pearson's r —   
## p-value —   
## ---------------------------------------------------------------------   
## Note. \* p < .05, \*\* p < .01, \*\*\* p < .001

Assumptions:

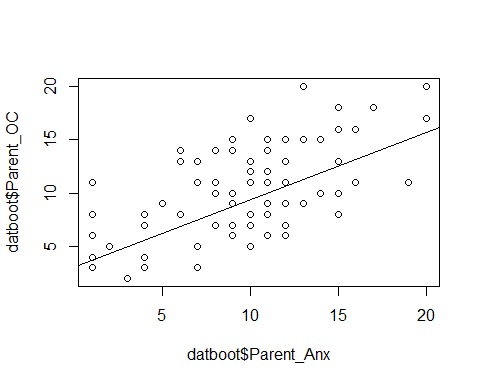
1. Missing Data - **NONE**
2. Univariate a. Normality **PASSED**, b. Linearity and c. Outliers
3. Multivariate a. Normality **PASSED** and b.Outliers **REMOVED**
4. Heteroscedsticity **PASSED**
5. Multi-collinearity **EXPECTED**

**2b. Univariate Linearity**

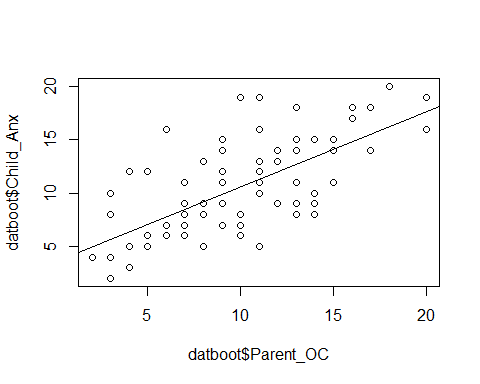
# Scatterplots [Assumption 2b]  
  
# Y ~ X [path c]  
plot(datboot$Parent\_Anx, datboot$Child\_Anx, abline(lm(datboot$Child\_Anx ~ datboot$Parent\_Anx)))



# M ~ X [path a]   
plot(datboot$Parent\_Anx, datboot$Parent\_OC, abline(lm(datboot$Parent\_Anx ~ datboot$Parent\_OC)))



# Y ~ M [path b]   
plot(datboot$Parent\_OC, datboot$Child\_Anx, abline(lm(datboot$Child\_Anx ~ datboot$Parent\_OC)))



**2c. Univariate Outliers**

#Identify outliers  
#scale() converts to z scores - "3" refers to standard deviations  
datboot[abs(scale(datboot$Parent\_Anx)) > 3, ]

## [1] Participant Parent\_Anx Parent\_OC Child\_Anx   
## <0 rows> (or 0-length row.names)

datboot[abs(scale(datboot$Child\_Anx)) > 3, ]

## [1] Participant Parent\_Anx Parent\_OC Child\_Anx   
## <0 rows> (or 0-length row.names)

datboot[abs(scale(datboot$Parent\_OC)) > 3, ]

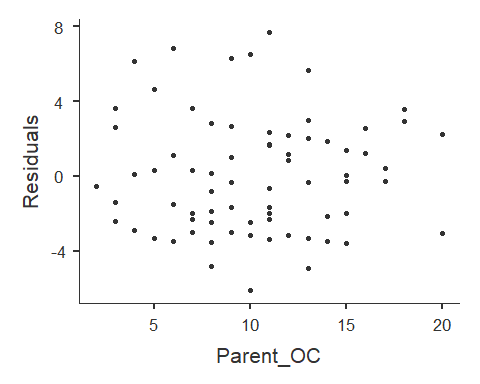
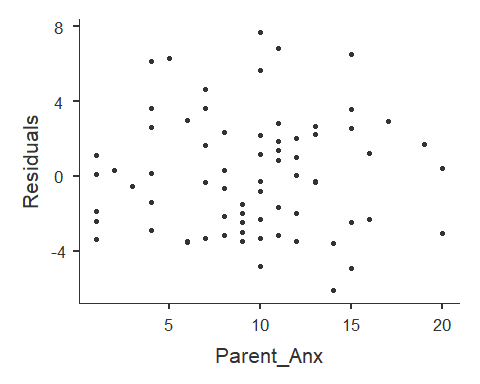
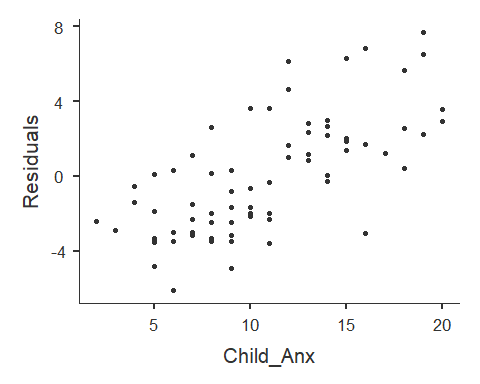
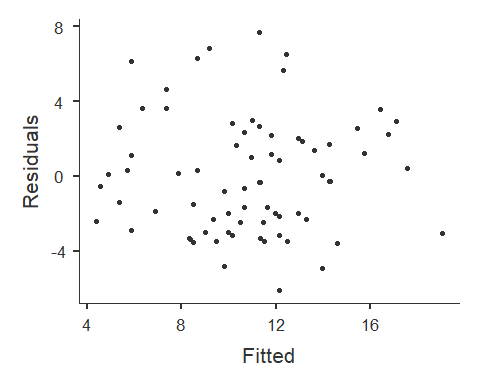
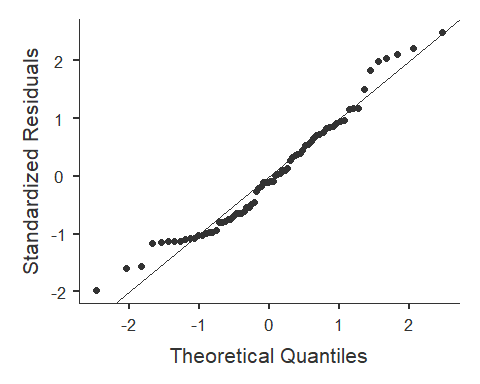
## [1] Participant Parent\_Anx Parent\_OC Child\_Anx   
## <0 rows> (or 0-length row.names)

#There are a total of 0 independent observations that contain outliers

**3a. Multivariate Normality**

#look at residuals and the Q-Q plot for independent variables (should be no relationship)  
#Observe Leverage (Mahalanobis' Distance) + Discrepancy (= Influence; Cook's Distance)  
  
# Y ~ X + M  
model.multi\_norm <- linReg(data = datboot,   
 dep = 'Child\_Anx',   
 covs = c('Parent\_Anx', 'Parent\_OC'),  
 blocks = list(c('Parent\_Anx', 'Parent\_OC')),   
 modelTest = TRUE,   
 r2Adj = TRUE,   
 stdEst = TRUE,   
 ciStdEst = TRUE,  
 qqPlot = TRUE, ##QQ plot  
 resPlots = TRUE) ##residuals plot   
  
model.multi\_norm

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## --------------------------------------------------------------------------   
## Model R R² Adjusted R² F df1 df2 p   
## --------------------------------------------------------------------------   
## 1 0.725 0.526 0.512 39.4 2 71 < .001   
## --------------------------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## ------------------------------------------------------------------------------------------   
## Predictor Estimate SE t p Stand. Estimate Lower Upper   
## ------------------------------------------------------------------------------------------   
## Intercept 2.575 0.998 2.58 0.012   
## Parent\_Anx 0.328 0.101 3.24 0.002 0.331 0.127 0.535   
## Parent\_OC 0.496 0.107 4.65 < .001 0.476 0.272 0.680   
## ------------------------------------------------------------------------------------------   
##   
##   
## ASSUMPTION CHECKS



#Alternate not using jvm library  
#model <- lm(model, data)  
#plot(model)

**3b. Multivariate Outliers**

#Check and remove multivariate outliers based on Cook's distance (CD)  
#for Mahalanobis' Distance (leverage only), see /Regression/Regression\_Diagnostics.Rmd for how-to  
#CD = Influence = Leverage + Discrepancy (Discrepancy = how much an observation deviates from the overall pattern of the model)  
  
#create model  
model.cook <- lm(datboot$Child\_Anx ~ datboot$Parent\_Anx + datboot$Parent\_OC)  
model.cook

##   
## Call:  
## lm(formula = datboot$Child\_Anx ~ datboot$Parent\_Anx + datboot$Parent\_OC)  
##   
## Coefficients:  
## (Intercept) datboot$Parent\_Anx datboot$Parent\_OC   
## 2.5754 0.3284 0.4962

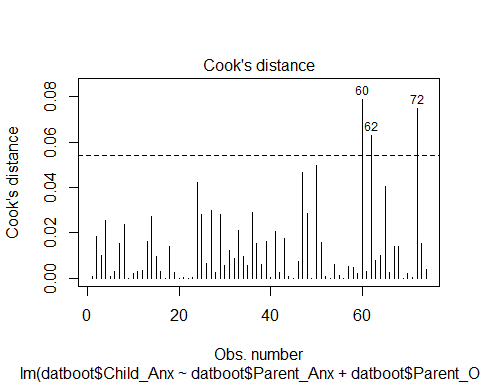
summary(model.cook)

##   
## Call:  
## lm(formula = datboot$Child\_Anx ~ datboot$Parent\_Anx + datboot$Parent\_OC)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.134 -2.451 -0.302 2.148 7.683   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.5754 0.9978 2.581 0.01192 \*   
## datboot$Parent\_Anx 0.3284 0.1015 3.236 0.00184 \*\*   
## datboot$Parent\_OC 0.4962 0.1067 4.650 1.5e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.137 on 71 degrees of freedom  
## Multiple R-squared: 0.5258, Adjusted R-squared: 0.5124   
## F-statistic: 39.36 on 2 and 71 DF, p-value: 3.145e-12

#find cook's distance for that model  
datboot$cook <- cooks.distance(model.cook)  
  
#create the cutoff [> 4/N]  
cook.cutoff <- 4/nrow(datboot)   
print(paste("Our Cook cutoff =", round(cook.cutoff, 3), "- anything above this value will be removed"))

## [1] "Our Cook cutoff = 0.054 - anything above this value will be removed"

#plot it out  
plot(model.cook, which = 4, cook.levels = cook.cutoff)  
  
#Add a cutoff line  
abline(h = cook.cutoff, lty = 2)



#Show and remove all outliers above your cutoff line  
datboot.no.outliers <- datboot[!(datboot$cook) > cook.cutoff, ]  
  
print(paste("There were", nrow(datboot), "observations before removing multivariate outliers"))

## [1] "There were 74 observations before removing multivariate outliers"

print(paste("We removed outlier observation #:", datboot[(datboot$cook) > cook.cutoff, 1], "with a Cook's distance =", round(datboot[(datboot$cook) > cook.cutoff, 5], 3)))

## [1] "We removed outlier observation #: 60 with a Cook's distance = 0.079"  
## [2] "We removed outlier observation #: 62 with a Cook's distance = 0.063"  
## [3] "We removed outlier observation #: 72 with a Cook's distance = 0.075"

print(paste("We now have", nrow(datboot.no.outliers), "total observations saved in the new dataset after removing", nrow(datboot[(datboot$cook) > cook.cutoff, ]), "outliers"))

## [1] "We now have 71 total observations saved in the new dataset after removing 3 outliers"

#N is now 39 after removing 1 multivariate outlier observation(s)  
 #was 40 after removing 0 univariate outlier observation(s)  
 #was 40 after removing 0 observation(s) with missing parameters  
 #was 40 originally (total 1 observation(s) removed from orginal dataset - 3%)

**4. Heteroscedasticity**

p\_load(car)  
#Breusch-Pagan test   
#H0 = no change in variance across residuals.  
model.breusch\_pagan <- lm(datboot.no.outliers$Child\_Anx ~ datboot.no.outliers$Parent\_Anx + datboot.no.outliers$Parent\_OC)  
ncvTest(model.breusch\_pagan)

## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 0.0003392379, Df = 1, p = 0.98531

#not significant = homoscedastic  
#If violated use Box-cox transformation [boxcox(model)] in library MASS

**5. Multi-collinearity**

#Multicollinearity is expected in mediation  
  
model.multicoll <- linReg(data = datboot.no.outliers,   
 dep = 'Child\_Anx',   
 cov = c('Parent\_Anx', 'Parent\_OC'),  
 blocks = list(c('Parent\_Anx', 'Parent\_OC')),   
 modelTest = TRUE,  
 r2Adj = TRUE,  
 stdEst = TRUE,  
 ciStdEst = TRUE,   
 collin = TRUE) #this line does the thing  
model.multicoll

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## --------------------------------------------------------------------------   
## Model R R² Adjusted R² F df1 df2 p   
## --------------------------------------------------------------------------   
## 1 0.771 0.594 0.582 49.7 2 68 < .001   
## --------------------------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## --------------------------------------------------------------------------------------------   
## Predictor Estimate SE t p Stand. Estimate Lower Upper   
## --------------------------------------------------------------------------------------------   
## Intercept 1.899 0.9375 2.03 0.047   
## Parent\_Anx 0.263 0.0954 2.76 0.007 0.269 0.0747 0.464   
## Parent\_OC 0.594 0.1004 5.92 < .001 0.576 0.3820 0.771   
## --------------------------------------------------------------------------------------------   
##   
##   
## ASSUMPTION CHECKS  
##   
## Collinearity Statistics   
## -----------------------------------   
## VIF Tolerance   
## -----------------------------------   
## Parent\_Anx 1.59 0.629   
## Parent\_OC 1.59 0.629   
## -----------------------------------

#Tolerance for all variables indicates low/no multicollinearity

*Mediation* 1. Regression a. Y ~ X b. M ~ X c. Y ~ X + M

1. No bootstrapping
2. Model
3. Sobel Test
4. Bootstrapping

**1. Regression** a. Y ~ X

model1 <- linReg(data = datboot.no.outliers,   
 dep = 'Child\_Anx',   
 covs = 'Parent\_Anx',  
 blocks = list(c('Parent\_Anx')),   
 modelTest = TRUE,   
 stdEst = TRUE,  
 ci = TRUE,  
 ciWidth = 95)  
model1

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.620 0.385 43.2 1 69 < .001   
## -----------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## -------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## -------------------------------------------------------------------------------------------   
## Intercept 4.826 0.9730 2.885 6.767 4.96 < .001   
## Parent\_Anx 0.607 0.0924 0.423 0.791 6.57 < .001 0.620   
## -------------------------------------------------------------------------------------------

1. M ~ X

model2 <- linReg(data = datboot.no.outliers,   
 dep = 'Parent\_OC',  
 covs = 'Child\_Anx',  
 blocks = list(c('Child\_Anx')),   
 modelTest = TRUE,   
 stdEst = TRUE,  
 ci = TRUE,  
 ciWidth = 95)  
model2

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.740 0.548 83.7 1 69 < .001   
## -----------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## ------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## ------------------------------------------------------------------------------------------   
## Intercept 2.814 0.9018 1.015 4.613 3.12 0.003   
## Child\_Anx 0.719 0.0785 0.562 0.875 9.15 < .001 0.740   
## ------------------------------------------------------------------------------------------

1. Y ~ X + M

model3 <- linReg(data = datboot.no.outliers,   
 dep = 'Child\_Anx',   
 covs = c('Parent\_Anx', 'Parent\_OC'),  
 blocks = list(c('Parent\_Anx', 'Parent\_OC')),   
 modelTest = TRUE,   
 stdEst = TRUE,  
 ci = TRUE,   
 ciWidth = 95)  
model3

##   
## LINEAR REGRESSION  
##   
## Model Fit Measures   
## -----------------------------------------------------------   
## Model R R² F df1 df2 p   
## -----------------------------------------------------------   
## 1 0.771 0.594 49.7 2 68 < .001   
## -----------------------------------------------------------   
##   
##   
## MODEL SPECIFIC RESULTS  
##   
## MODEL 1  
##   
## Model Coefficients   
## --------------------------------------------------------------------------------------------   
## Predictor Estimate SE Lower Upper t p Stand. Estimate   
## --------------------------------------------------------------------------------------------   
## Intercept 1.899 0.9375 0.0287 3.770 2.03 0.047   
## Parent\_Anx 0.263 0.0954 0.0731 0.454 2.76 0.007 0.269   
## Parent\_OC 0.594 0.1004 0.3936 0.794 5.92 < .001 0.576   
## --------------------------------------------------------------------------------------------

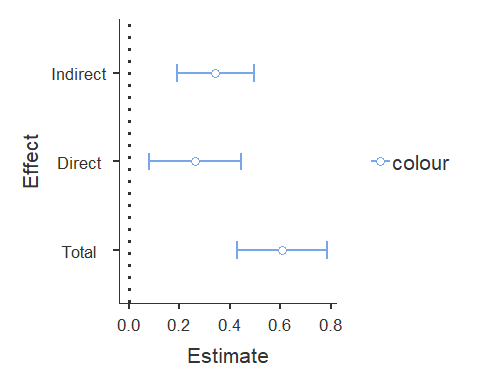
**2. No bootstrapping**

1. Model b. Sobel Test

# (a \* b) = (c - c') = Indirect Effect [i.e., amount of mediation]  
# Z = Sobel test  
# a = Path Estimate from X to M  
# b = Path Estimate from M to Y  
# c = Total Estimate (Direct Estimate + Indirect Estimate)  
# c' = Direct Estimate  
  
med <- medmod::med(datboot.no.outliers,   
 dep = 'Child\_Anx',   
 pred = 'Parent\_Anx',   
 med = 'Parent\_OC',   
 pm = TRUE,   
 ci = TRUE,  
 paths = TRUE,  
 label = TRUE,  
 estPlot = TRUE)  
med

##   
## MEDIATION  
##   
## Mediation Estimates   
## ---------------------------------------------------------------------------------------------------   
## Effect Label Estimate SE Lower Upper Z p % Mediation   
## ---------------------------------------------------------------------------------------------------   
## Indirect a × b 0.344 0.0778 0.1912 0.496 4.42 < .001 56.6   
## Direct c 0.263 0.0933 0.0804 0.446 2.82 0.005 43.4   
## Total c + a × b 0.607 0.0911 0.4285 0.786 6.66 < .001 100.0   
## ---------------------------------------------------------------------------------------------------   
##   
##   
## Path Estimates   
## ----------------------------------------------------------------------------------------------------   
## Label Estimate SE Lower Upper Z p   
## ----------------------------------------------------------------------------------------------------   
## Parent\_Anx <U+2192> Parent\_OC a 0.579 0.0894 0.4035 0.754 6.47 < .001   
## Parent\_OC <U+2192> Child\_Anx b 0.594 0.0983 0.4013 0.787 6.04 < .001   
## Parent\_Anx <U+2192> Child\_Anx c 0.263 0.0933 0.0804 0.446 2.82 0.005   
## ----------------------------------------------------------------------------------------------------

## Scale for 'colour' is already present. Adding another scale for  
## 'colour', which will replace the existing scale.



# Look at this in comparison to the indirect mediation estimate.  
# sobel(X, M, Y)  
cat("\n")

sobel(datboot.no.outliers$Parent\_Anx, datboot.no.outliers$Parent\_OC, datboot.no.outliers$Child\_Anx)

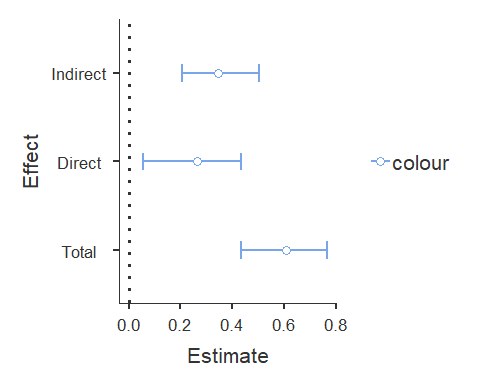
## $`Mod1: Y~X`  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 4.825577 0.97296433 4.959665 4.863136e-06  
## pred 0.607077 0.09240246 6.569922 7.931925e-09  
##   
## $`Mod2: Y~X+M`  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 1.8994837 0.93752054 2.026072 4.668296e-02  
## pred 0.2633798 0.09537577 2.761496 7.391486e-03  
## med 0.5939162 0.10040604 5.915144 1.195656e-07  
##   
## $`Mod3: M~X`  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 4.9267776 0.95486459 5.159661 2.264567e-06  
## pred 0.5786964 0.09068353 6.381494 1.723080e-08  
##   
## $Indirect.Effect  
## [1] 0.3436972  
##   
## $SE  
## [1] 0.07922674  
##   
## $z.value  
## [1] 4.338146  
##   
## $N  
## [1] 71

**3. Bootstrapping**

boot <- medmod::med(datboot.no.outliers,   
 dep = 'Child\_Anx',   
 pred = 'Parent\_Anx',   
 med = 'Parent\_OC',   
 estMethod = 'bootstrap',   
 bootstrap = 1000,   
 pm = TRUE,  
 ci = TRUE,  
 paths = TRUE,  
 label = TRUE,  
 estPlot = TRUE)  
  
boot

##   
## MEDIATION  
##   
## Mediation Estimates   
## ---------------------------------------------------------------------------------------------------   
## Effect Label Estimate SE Lower Upper Z p % Mediation   
## ---------------------------------------------------------------------------------------------------   
## Indirect a × b 0.344 0.0785 0.2045 0.502 4.38 < .001 56.6   
## Direct c 0.263 0.0937 0.0574 0.433 2.81 0.005 43.4   
## Total c + a × b 0.607 0.0827 0.4323 0.766 7.34 < .001 100.0   
## ---------------------------------------------------------------------------------------------------   
##   
##   
## Path Estimates   
## ----------------------------------------------------------------------------------------------------   
## Label Estimate SE Lower Upper Z p   
## ----------------------------------------------------------------------------------------------------   
## Parent\_Anx <U+2192> Parent\_OC a 0.579 0.0874 0.3993 0.748 6.62 < .001   
## Parent\_OC <U+2192> Child\_Anx b 0.594 0.0915 0.4186 0.780 6.49 < .001   
## Parent\_Anx <U+2192> Child\_Anx c 0.263 0.0937 0.0574 0.433 2.81 0.005   
## ----------------------------------------------------------------------------------------------------

## Scale for 'colour' is already present. Adding another scale for  
## 'colour', which will replace the existing scale.



# Check the z-values, SE, and p-values for pathways to see differences (X -> Y for this one).