GC3: Intro to SEM

Jimmy Zhang

2022-10-06

```
#install.packages("lavaan")
library(lavaan)
```

```
## This is lavaan 0.6-12
## lavaan is FREE software! Please report any bugs.
```

```
# Read the dataset
df <- read.csv("data/hh_sem.csv", header = FALSE)

# Change the variables names
names(df) <- c("ID", "schoolid", "W5tot6m", "W5alc6m", "W5ecig6m", "W5cig6m",
               "W5nic6m", "CESD1", "CESD2", "CESD3", "CESD4", "W5adhdat",
               "W5adhdim", "W5imp1", "W5imp2", "W5imp3", "W5imp4", "W5imp5",
               "W5imp6", "W5imp7", "W5imp8", "W5imp9", "SUBUSE1", "SUBUSE2",
               "SUBUSE3", "SUBUSE4", "W6tot6m", "W6alc6m", "W6ecig6m" ,"W6cig6m",
               "W6nic6m", "AR1", "W5adecig2", "W5adhkah", "AR2", "AR3",
               "AR4", "W5admje")

# Replace the NA values
df[df == -999] <- NA
```

Example 1: Confirmatory Factor Analysis (CFA)

```
model_1 <- 'impulsivity =~ W5imp1 + W5imp2 + W5imp3 +
                           W5imp4 + W5imp5 + W5imp6 +
                           W5imp7 + W5imp8 + W5imp9'
```

```
model_1_fit <- cfa(model_1, data = df)
summary(model_1_fit, fit.measures = TRUE)
```

```
## lavaan 0.6-12 ended normally after 25 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        18
##
##                                                   Used       Total
##   Number of observations                          2710        3396
```

```
##
## Model Test User Model:
##
##   Test statistic                                573.088
##   Degrees of freedom                                27
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                               9086.667
##   Degrees of freedom                                36
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.940
##   Tucker-Lewis Index (TLI)                       0.920
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -24379.627
##   Loglikelihood unrestricted model (H1)     -24093.082
##
##   Akaike (AIC)                               48795.253
##   Bayesian (BIC)                             48901.538
##   Sample-size adjusted Bayesian (BIC)        48844.346
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.086
##   90 Percent confidence interval - lower         0.080
##   90 Percent confidence interval - upper         0.093
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.043
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   impulsivity =~
##     W5imp1            1.000
##     W5imp2            0.774    0.031   25.081    0.000
##     W5imp3            0.931    0.042   21.919    0.000
##     W5imp4            0.920    0.036   25.913    0.000
##     W5imp5            0.870    0.037   23.339    0.000
##     W5imp6            1.053    0.043   24.374    0.000
##     W5imp7            0.911    0.035   26.141    0.000
```

```
##     W5imp8            1.037    0.038   27.023    0.000
##     W5imp9            0.953    0.036   26.740    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .W5imp1            0.725    0.021   34.787    0.000
##    .W5imp2            0.224    0.007   32.872    0.000
##    .W5imp3            0.626    0.018   34.779    0.000
##    .W5imp4            0.256    0.008   31.931    0.000
##    .W5imp5            0.416    0.012   34.133    0.000
##    .W5imp6            0.489    0.015   33.470    0.000
##    .W5imp7            0.236    0.007   31.610    0.000
##    .W5imp8            0.232    0.008   29.962    0.000
##    .W5imp9            0.215    0.007   30.577    0.000
##     impulsivity       0.295    0.021   14.294    0.000
```

```
# Standardized version
summary(model_1_fit, standardized = TRUE)
```

```
## lavaan 0.6-12 ended normally after 25 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        18
##
##                                                   Used       Total
##   Number of observations                          2710        3396
##
## Model Test User Model:
##
##   Test statistic                               573.088
##   Degrees of freedom                               27
##   P-value (Chi-square)                          0.000
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   impulsivity =~
##     W5imp1            1.000                               0.544    0.538
##     W5imp2            0.774    0.031   25.081    0.000    0.420    0.664
##     W5imp3            0.931    0.042   21.919    0.000    0.506    0.539
##     W5imp4            0.920    0.036   25.913    0.000    0.500    0.703
##     W5imp5            0.870    0.037   23.339    0.000    0.473    0.592
##     W5imp6            1.053    0.043   24.374    0.000    0.573    0.634
##     W5imp7            0.911    0.035   26.141    0.000    0.495    0.714
##     W5imp8            1.037    0.038   27.023    0.000    0.563    0.760
##     W5imp9            0.953    0.036   26.740    0.000    0.518    0.745
##
## Variances:
```

```
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .W5imp1            0.725    0.021   34.787    0.000    0.725    0.711
##    .W5imp2            0.224    0.007   32.872    0.000    0.224    0.559
##    .W5imp3            0.626    0.018   34.779    0.000    0.626    0.710
##    .W5imp4            0.256    0.008   31.931    0.000    0.256    0.506
##    .W5imp5            0.416    0.012   34.133    0.000    0.416    0.650
##    .W5imp6            0.489    0.015   33.470    0.000    0.489    0.599
##    .W5imp7            0.236    0.007   31.610    0.000    0.236    0.490
##    .W5imp8            0.232    0.008   29.962    0.000    0.232    0.422
##    .W5imp9            0.215    0.007   30.577    0.000    0.215    0.445
##     impulsivity       0.295    0.021   14.294    0.000    1.000    1.000
```

Example 2: Full SEM structure

```
model_2 <- '
            # Measurement Model
            Impulsivity =~ W5imp1 + W5imp2 + W5imp3 +
                           W5imp4 + W5imp5 + W5imp6 +
                           W5imp7 + W5imp8 + W5imp9
            Addiction =~ AR1 + AR2 + AR3 + AR4
            Depression =~ CESD1 + CESD2 + CESD3 + CESD4
            Substance =~ SUBUSE1 + SUBUSE2 + SUBUSE3 + SUBUSE4

            # Structural Model
            Addiction ~ Impulsivity + Depression
            Depression ~ Addiction
            Substance ~ Addiction + Depression
            '
```

```
model_2_fit <- sem(model_2, data = df)
summary(model_2_fit, fit.measures = TRUE)
```

```
## lavaan 0.6-12 ended normally after 110 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        47
##
##                                                   Used       Total
##   Number of observations                          2282        3396
##
## Model Test User Model:
##
##   Test statistic                              1901.303
##   Degrees of freedom                              184
##   P-value (Chi-square)                          0.000
##
## Model Test Baseline Model:
##
##   Test statistic                             22501.830
##   Degrees of freedom                               210
```

```
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.923
##   Tucker-Lewis Index (TLI)                       0.912
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -53031.148
##   Loglikelihood unrestricted model (H1)     -52080.496
##
##   Akaike (AIC)                              106156.295
##   Bayesian (BIC)                            106425.737
##   Sample-size adjusted Bayesian (BIC)       106276.410
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.064
##   90 Percent confidence interval - lower         0.061
##   90 Percent confidence interval - upper         0.067
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.059
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Impulsivity =~
##     W5imp1            1.000
##     W5imp2            0.736    0.032   22.778    0.000
##     W5imp3            0.979    0.047   21.015    0.000
##     W5imp4            0.922    0.038   24.102    0.000
##     W5imp5            0.837    0.040   21.042    0.000
##     W5imp6            1.031    0.046   22.348    0.000
##     W5imp7            0.879    0.037   23.788    0.000
##     W5imp8            1.025    0.041   25.012    0.000
##     W5imp9            0.932    0.038   24.569    0.000
##   Addiction =~
##     AR1               1.000
##     AR2               1.443    0.027   53.327    0.000
##     AR3               1.288    0.025   52.356    0.000
##     AR4               1.202    0.024   50.456    0.000
##   Depression =~
##     CESD1             1.000
##     CESD2             1.179    0.027   43.871    0.000
##     CESD3             0.908    0.029   30.960    0.000
```

```
##     CESD4             1.011    0.029   34.941    0.000
##   Substance =~
##     SUBUSE1           1.000
##     SUBUSE2           0.581    0.029   19.996    0.000
##     SUBUSE3           0.680    0.038   18.119    0.000
##     SUBUSE4           1.477    0.068   21.618    0.000
##
## Regressions:
##                   Estimate  Std.Err  z-value  P(>|z|)
##   Addiction ~
##     Impulsivity      5.552    1.699    3.267    0.001
##     Depression     -12.169    3.795   -3.207    0.001
##   Depression ~
##     Addiction         1.923    0.431    4.462    0.000
##   Substance ~
##     Addiction         0.097    0.012    7.876    0.000
##     Depression        0.077    0.026    2.989    0.003
##
## Variances:
##                   Estimate  Std.Err  z-value  P(>|z|)
##    .W5imp1            0.723    0.023   31.907    0.000
##    .W5imp2            0.225    0.007   30.513    0.000
##    .W5imp3            0.593    0.019   31.593    0.000
##    .W5imp4            0.247    0.008   29.092    0.000
##    .W5imp5            0.432    0.014   31.580    0.000
##    .W5imp6            0.490    0.016   30.837    0.000
##    .W5imp7            0.246    0.008   29.506    0.000
##    .W5imp8            0.226    0.008   27.428    0.000
##    .W5imp9            0.218    0.008   28.343    0.000
##    .AR1               0.758    0.026   29.584    0.000
##    .AR2               0.546    0.026   21.193    0.000
##    .AR3               0.510    0.022   22.985    0.000
##    .AR4               0.574    0.022   25.545    0.000
##    .CESD1             0.175    0.007   24.204    0.000
##    .CESD2             0.108    0.007   14.597    0.000
##    .CESD3             0.396    0.013   30.710    0.000
##    .CESD4             0.340    0.012   29.212    0.000
##    .SUBUSE1           0.366    0.017   21.548    0.000
##    .SUBUSE2           0.231    0.008   27.403    0.000
##    .SUBUSE3           0.461    0.016   29.485    0.000
##    .SUBUSE4           0.785    0.037   21.325    0.000
##     Impulsivity       0.299    0.023   13.209    0.000
##    .Addiction        41.583   24.871    1.672    0.095
##    .Depression        5.339    2.297    2.325    0.020
##    .Substance         0.308    0.021   14.801    0.000
```