Answers to Exercises Basic CFA

```r
library("lavaan")
```

Additional Exercise 3

```r
data(HolzingerSwineford1939)
HS.mod.1 <- '
  IQ =~ x1 + x2 + x3 + x4 + x5 + x6
'
HS.fit.1 <- cfa(HS.mod.1, data = HolzingerSwineford1939, estimator = "MLR")
summary(HS.fit.1, standardized = TRUE, fit.measures = TRUE)
```

```
## lavaan 0.6-5 ended normally after 34 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         12
##
##   Number of observations                           301
##
## Model Test User Model:
##                                               Standard      Robust
##   Test Statistic                               103.230     100.487
##   Degrees of freedom                                 9           9
##   P-value (Chi-square)                           0.000       0.000
##   Scaling correction factor                                  1.027
##     for the Yuan-Bentler correction (Mplus variant)
##
## Model Test Baseline Model:
##
##   Test statistic                               668.643     605.920
##   Degrees of freedom                                15          15
##   P-value                                        0.000       0.000
##   Scaling correction factor                                  1.104
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.856       0.845
##   Tucker-Lewis Index (TLI)                       0.760       0.742
##
##   Robust Comparative Fit Index (CFI)                         0.856
##   Robust Tucker-Lewis Index (TLI)                            0.760
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -2559.686   -2559.686
```

```
##   Scaling correction factor                                    1.117
##       for the MLR correction
##   Loglikelihood unrestricted model (H1)      -2508.071   -2508.071
##   Scaling correction factor                                    1.079
##       for the MLR correction
##
##   Akaike (AIC)                                5143.372    5143.372
##   Bayesian (BIC)                              5187.857    5187.857
##   Sample-size adjusted Bayesian (BIC)         5149.800    5149.800
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.187       0.184
##   90 Percent confidence interval - lower         0.155       0.153
##   90 Percent confidence interval - upper         0.220       0.217
##   P-value RMSEA <= 0.05                          0.000       0.000
##
##   Robust RMSEA                                              0.186
##   90 Percent confidence interval - lower                    0.154
##   90 Percent confidence interval - upper                    0.220
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.114       0.114
##
## Parameter Estimates:
##
##   Information                                 Observed
##   Observed information based on                Hessian
##   Standard errors                    Robust.huber.white
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   IQ =~
##     x1                1.000                               0.488    0.419
##     x2                0.511    0.152    3.354    0.001    0.250    0.212
##     x3                0.468    0.128    3.657    0.000    0.229    0.203
##     x4                2.028    0.322    6.303    0.000    0.990    0.852
##     x5                2.234    0.374    5.974    0.000    1.091    0.847
##     x6                1.882    0.295    6.375    0.000    0.919    0.840
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .x1                1.120    0.109   10.321    0.000    1.120    0.824
##    .x2                1.319    0.128   10.272    0.000    1.319    0.955
##    .x3                1.223    0.078   15.715    0.000    1.223    0.959
##    .x4                0.370    0.050    7.341    0.000    0.370    0.274
##    .x5                0.470    0.058    8.132    0.000    0.470    0.283
##    .x6                0.351    0.046    7.635    0.000    0.351    0.294
##     IQ                0.238    0.077    3.113    0.002    1.000    1.000
```

a) The model does not fit very well. Also, the standardized loadings of X1, X2 and X3 are much smaller than those of X4, X5 and X6, so do not seem well explained by the model.

b) The robust ML χ^2 value is lower than the standard ML χ^2, but the difference is small. The standard and robust CFI are identical, standard and robust RMSEA differ only slightly. The SRMR is the same between standard and robut ML, as it is not based on the χ^2 value, but based on the residual correlations. These are identical between standard and robust ML estimation.

c) We inspect residuals and modification indices:

```
residuals(HS.fit.1, type = "cor")
```

```
## $type
## [1] "cor.bollen"
##
## $cov
##        x1     x2     x3     x4     x5     x6
## x1  0.000
## x2  0.208  0.000
## x3  0.356  0.297  0.000
## x4  0.016 -0.028 -0.014  0.000
## x5 -0.061 -0.040 -0.094  0.012  0.000
## x6  0.005  0.014  0.027 -0.012  0.008  0.000
```

```
modificationIndices(HS.fit.1, sort. = TRUE)
```

```
##    lhs op rhs     mi    epc sepc.lv sepc.all sepc.nox
## 15  x1 ~~  x3 49.835  0.484   0.484    0.414    0.414
## 19  x2 ~~  x3 29.298  0.399   0.399    0.314    0.314
## 14  x1 ~~  x2 17.170  0.295   0.295    0.243    0.243
## 24  x3 ~~  x5 14.774 -0.206  -0.206   -0.271   -0.271
## 17  x1 ~~  x5  7.829 -0.149  -0.149   -0.205   -0.205
## 26  x4 ~~  x5  7.807  0.260   0.260    0.623    0.623
## 27  x4 ~~  x6  7.107 -0.207  -0.207   -0.573   -0.573
## 28  x5 ~~  x6  3.327  0.154   0.154    0.380    0.380
## 21  x2 ~~  x5  2.742 -0.092  -0.092   -0.117   -0.117
## 20  x2 ~~  x4  1.399 -0.059  -0.059   -0.085   -0.085
## 25  x3 ~~  x6  1.179  0.050   0.050    0.076    0.076
## 16  x1 ~~  x4  0.540  0.035   0.035    0.055    0.055
## 23  x3 ~~  x4  0.342 -0.028  -0.028   -0.042   -0.042
## 22  x2 ~~  x6  0.308  0.026   0.026    0.039    0.039
## 18  x1 ~~  x6  0.042  0.009   0.009    0.015    0.015
```

Residuals among X1, X2 and X3 are largest. Highest modification indices are for correlations between X1, X2 and X3. This matches what we already expected based on the standardized loadings: perhaps the model can maybe be improved by adding a separate factor for X1, X2 and X3:

```
HS.mod.2 <- '
  IQ1 =~ x1 + x2 + x3
  IQ2 =~ x4 + x5 + x6
'
HS.fit.2 <- cfa(HS.mod.2, data = HolzingerSwineford1939, estimator = "MLR")
summary(HS.fit.2, standardized = TRUE, fit.measures = TRUE)
```

```
## lavaan 0.6-5 ended normally after 28 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of free parameters                         13
##
##   Number of observations                           301
```

```
##
## Model Test User Model:
##                                               Standard      Robust
##   Test Statistic                                24.361      24.373
##   Degrees of freedom                                 8           8
##   P-value (Chi-square)                           0.002       0.002
##   Scaling correction factor                                  1.000
##     for the Yuan-Bentler correction (Mplus variant)
##
## Model Test Baseline Model:
##
##   Test statistic                               668.643     605.920
##   Degrees of freedom                                15          15
##   P-value                                        0.000       0.000
##   Scaling correction factor                                  1.104
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.975       0.972
##   Tucker-Lewis Index (TLI)                       0.953       0.948
##
##   Robust Comparative Fit Index (CFI)                         0.975
##   Robust Tucker-Lewis Index (TLI)                            0.953
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -2520.252   -2520.252
##   Scaling correction factor                                  1.127
##       for the MLR correction
##   Loglikelihood unrestricted model (H1)      -2508.071   -2508.071
##   Scaling correction factor                                  1.079
##       for the MLR correction
##
##   Akaike (AIC)                                5066.503    5066.503
##   Bayesian (BIC)                              5114.696    5114.696
##   Sample-size adjusted Bayesian (BIC)         5073.467    5073.467
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.082       0.082
##   90 Percent confidence interval - lower         0.046       0.046
##   90 Percent confidence interval - upper         0.121       0.121
##   P-value RMSEA <= 0.05                          0.067       0.067
##
##   Robust RMSEA                                               0.082
##   90 Percent confidence interval - lower                     0.046
##   90 Percent confidence interval - upper                     0.121
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.047       0.047
##
## Parameter Estimates:
##
```

```
##   Information                                        Observed
##   Observed information based on                       Hessian
##   Standard errors                           Robust.huber.white
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   IQ1 =~
##     x1                1.000                               0.907    0.778
##     x2                0.559    0.163    3.436    0.001    0.507    0.431
##     x3                0.708    0.162    4.369    0.000    0.642    0.568
##   IQ2 =~
##     x4                1.000                               0.991    0.852
##     x5                1.111    0.066   16.910    0.000    1.101    0.854
##     x6                0.925    0.062   14.966    0.000    0.917    0.838
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   IQ1 ~~
##     IQ2               0.414    0.106    3.889    0.000    0.461    0.461
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .x1                0.536    0.194    2.766    0.006    0.536    0.395
##    .x2                1.125    0.120    9.401    0.000    1.125    0.814
##    .x3                0.863    0.110    7.832    0.000    0.863    0.677
##    .x4                0.369    0.051    7.311    0.000    0.369    0.274
##    .x5                0.449    0.057    7.830    0.000    0.449    0.270
##    .x6                0.356    0.047    7.639    0.000    0.356    0.298
##     IQ1               0.822    0.215    3.831    0.000    1.000    1.000
##     IQ2               0.981    0.122    8.053    0.000    1.000    1.000
```

d) CFI and SRMR indicate good model fit. CFI is $> .95$, SRMR is $< .08$. RMSEA is too high according to the Hu and Bentler criteria ($> .08$). However, the p-value for the test of the RMSEA being $< .05$ is $> .05$ (indicating that the hypothesis of close fit is not rejected), and the confidence interval for the RMSEA includes .05, which seems acceptable.

Looking at the estimated parameters, the standardized factor loadings of X1, X2 and X3 have substantially increased. Looking at model fit, that definitely improved. Chi-square value is much smaller (although still significant, but that is to be expected with $N = 300$).