

Confirmatory Factor Analysis

Statsomat.com

28 December 2020

Warning: The automatic computation and interpretation delivered by the Statsomat should not completely replace the classical, made by humans graphical exploratory data analysis and statistical analysis. There may be data cases for which the Statsomat does not deliver the most optimal solution or output interpretation.

Basic Information

Automatic statistics for the file:

File
case13.csv

Your selection for the encoding: Auto

Your selection for the decimal character: Auto

Observations (rows with at least one non-missing value): 2571

Variables (columns with at least one non-missing value): 7

Variables considered continuous: Numerical variables considered binary or ordinal: 7

Numerical variables considered binary or ordinal
q01
q02
q03
q04
q05
q06
q07

Character variables considered binary: 0

Character variables considered nominal and transformed to binary: 0

Model Syntax

The following table describes the applied model equations in lavaan model syntax, either as entered by you in the text area (denoted by User=1) or established internally (User=0). The last column numbers the free parameters which are estimated.

Left hand side	Operator	Right hand side	User	Free parameter
f	=~	q03	1	0
f	=~	q04	1	1
f	=~	q05	1	2
q03	~~	q03	0	3
q04	~~	q04	0	4
q05	~~	q05	0	5
f	~~	f	0	6
q03	~1		0	7
q04	~1		0	8
q05	~1		0	9
f	~1		0	0

Assumptions

Open issue

Model Settings

Outputs

Model Fit Summary

lavaan 0.6-7 ended normally after 23 iterations

Estimator	ML
Optimization method	NLMINB
Number of free parameters	9
Number of observations	2571
Number of missing patterns	1

Model Test User Model:

Test statistic	0.000
Degrees of freedom	0

Model Test Baseline Model:

Test statistic	942.750
Degrees of freedom	3
P-value	0.000

User Model versus Baseline Model:

Comparative Fit Index (CFI)	1.000
Tucker-Lewis Index (TLI)	1.000

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-10429.128
Loglikelihood unrestricted model (H1)	-10429.128
Akaike (AIC)	20876.256
Bayesian (BIC)	20928.924
Sample-size adjusted Bayesian (BIC)	20900.329

Root Mean Square Error of Approximation:

RMSEA	0.000
90 Percent confidence interval - lower	0.000
90 Percent confidence interval - upper	0.000
P-value RMSEA <= 0.05	NA

Standardized Root Mean Square Residual:

SRMR	0.000
------	-------

Parameter Estimates:

Standard errors	Standard
Information	Observed
Observed information based on	Hessian

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)
f =~				
q03	1.000			
q04	-1.139	0.073	-15.652	0.000
q05	-0.945	0.056	-16.840	0.000

Intercepts:

	Estimate	Std.Err	z-value	P(> z)
.q03	2.585	0.021	121.968	0.000
.q04	2.786	0.019	148.960	0.000
.q05	2.722	0.019	143.114	0.000

f 0.000

Variances:

	Estimate	Std.Err	z-value	P(> z)
.q03	0.815	0.031	26.484	0.000
.q04	0.458	0.030	15.359	0.000
.q05	0.626	0.025	24.599	0.000
f	0.340	0.031	11.034	0.000

Completely Standardized Parameter Estimates

Latent Variables:

	est.std	Std.Err	z-value	P(> z)	ci.lower	ci.upper
f =~						
q03	0.543	0.022	25.120	0.000	0.500	0.585
q04	-0.701	0.024	-29.710	0.000	-0.747	-0.655
q05	-0.572	0.022	-26.105	0.000	-0.615	-0.529

Intercepts:

	est.std	Std.Err	z-value	P(> z)	ci.lower	ci.upper
.q03	0.000	0.020	0.000	1.000	-0.039	0.039
.q04	0.000	0.020	0.000	1.000	-0.039	0.039
.q05	-0.000	0.020	-0.000	1.000	-0.039	0.039
f	0.000				0.000	0.000

Variances:

	est.std	Std.Err	z-value	P(> z)	ci.lower	ci.upper
.q03	0.705	0.023	30.066	0.000	0.659	0.751
.q04	0.509	0.033	15.384	0.000	0.444	0.574
.q05	0.673	0.025	26.888	0.000	0.624	0.722
f	1.000				1.000	1.000

Communality

Table 4: Communality

Variable	Communality
q03	0.29
q04	0.49
q05	0.33

Factor Discriminant Validity

Error in discriminantVal(fit, cutoff = 0.85, merge = FALSE, level = 0.95, : The model has only one e

Error in eval(expr, envir, enclos): object 'discr' not found

Error in lapply(X = X, FUN = FUN, ...): object 'discr' not found

Error in knitr::kable(discr, caption = "Factor Discriminant Validity Test at Cutoff 0.85", : object

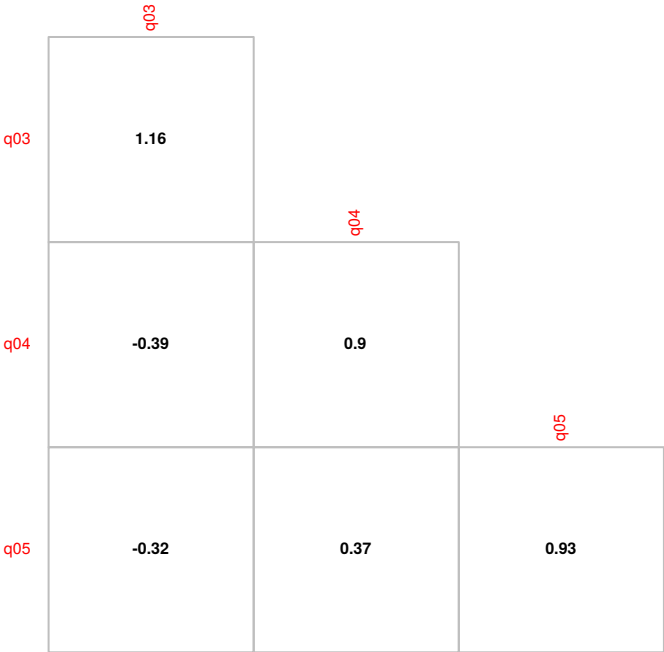
Factor Reliability

Error in `rownames<-`(`*tmp*`, value = c("Omega (Bentler)", "Omega (McDonald)", : attempt to set 'r

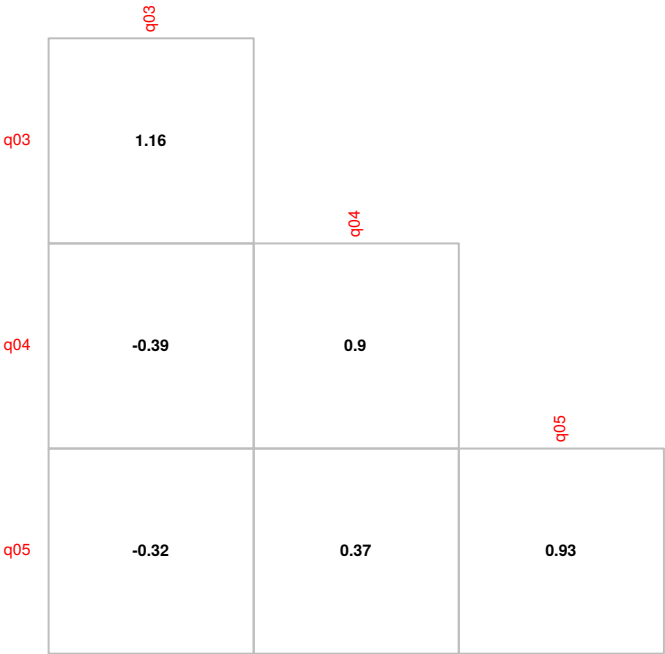
Table 5: Factor Reliability

	x
omega2	0.17
omega3	0.17
avevar	0.36

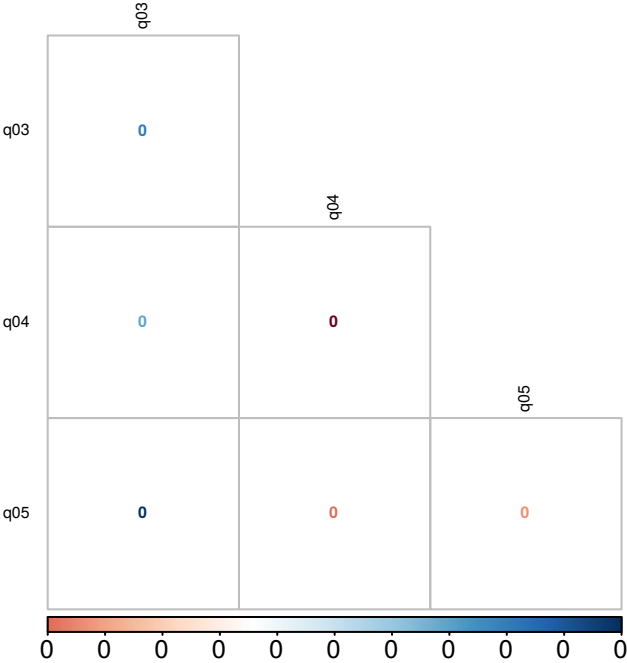
Observed Covariance Matrix



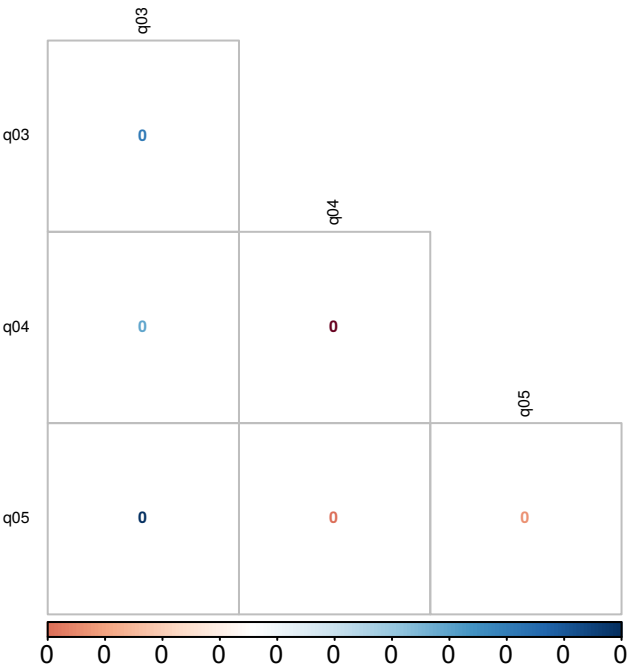
Model-Implied Covariance Matrix



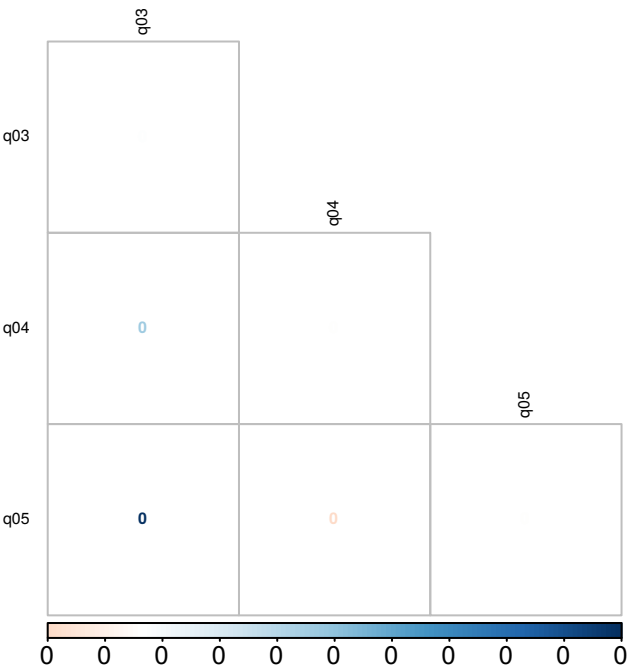
Residual Covariance Matrix



Standardized Residual Matrix



Residual Correlation Matrix



Modification Indices

Interpretation

Goodness of Fit Indices

We consider some of the model fit indices from the Model Fit Summary section to check the goodness-of-fit of the model. To decide for an acceptable or non-acceptable model, we apply thresholds considered in the References: [Brown], [Kline].

Model Test User Model

The degrees of freedom are calculated as the number of known parameters minus the number of free parameters: $9 - 9 = 0$. The 0 degrees of freedom are indicating a just-identified or saturated model.

The test statistic with the value 0 is called the Chi-square model fit index and represents the difference between summaries of the model-implied covariance matrix and the observed covariance matrix which is hypothesized and desirable to be zero. In general, if the p-value is larger than 0.01 then the test is not statistically significant at 1 % error, the hypothesis cannot be rejected, which would be in favour of the model.

In our case, the p-value is NA

```
Error in if (round(fit@fit@test$standard$pvalue, 3) > cutchi2(nrow(df_work))) {: missing value where
```

Model Test Baseline Model

The test statistic with the value 942.75 represents the difference between summaries of the baseline model (an alternative model-implied covariance matrix having zero covariances, i.e. a worst fitting model assuming independent variables) and the observed covariance matrix. The p-value of the test of a zero difference is < 0.001 suggesting that the baseline model does not fit good to the data. This result is used indirectly in the construction of other model fit indices.

Root Mean Square Error of Approximation:

The Root Mean Square Error of Approximation (RMSEA) is a fit index based on the chi-square test statistic, which corrects for parsimony, i.e. overly complex models are penalized. RMSEA can be greater or equal than zero, with values close to zero suggesting an acceptable model fit.

In our case, the RMSEA is 0. The upper bound of the 90% confidence interval of the RMSEA is 0 and smaller than the threshold value 0.05, suggesting an excellent model fit.

Standardized Root Mean Square Residual:

The Standardized Root Mean Square Residual (SRMR) is a fit index derived from the residual correlation matrix with a range between zero and one with values close to zero suggesting an acceptable model fit.

In our case, the SRMR is 0 which is smaller than the threshold value 0.06 suggesting an acceptable model fit.

User Model versus Baseline Model

The Comparative Fit Index (CFI), evaluates the fit of the the model in relation to the worst-fitting baseline model described above. It ranges between zero and one, with values close to one suggesting good models (in the sense of departure from the baseline model).

In our case, the CFI is 1 which is greater or equal than the threshold value 0.95, suggesting a good model fit.

Similarly to the CFI, the Tucker-Lewis Index (TLI) evaluates the fit of the model in relation to the worst-fitting baseline model described above. Moreover, overly complex models are penalized. Values can range outside zero and one but the index is interpreted similarly to the CFI.

In our case, the TLI is 1 which is greater or equal than the threshold value 0.95, suggesting a good model fit.

Summary of the Goodness of Fit Indices

```
Error in eval(expr, envir, enclos): object 'chi2' not found
```

Residuals

We analyze the residual matrices from the Outputs chapter. The residual covariance matrix represents the difference between the observed covariance matrix and the fitted model-implied covariance matrix. Large absolute values indicate local areas of misfit. However, the residuals are affected by the raw metric and are difficult to interpret more precisely.

A better interpretation allows the standardized residual matrix (residuals divided by their estimated asymptotic standard error) and the residual correlation matrix.

There are no variable pairs with standardized residuals which are larger or equal than the considered threshold 2.58 [brown] or correlation residuals which are larger or equal than the considered threshold 0.1 [kline]. Therefore, no relationships among the variables are substantially underestimated by the model. There are no variable pairs with standardized residuals which are smaller or equal than the considered threshold -2.58 [brown] or correlation residuals which are smaller or equal than the considered threshold -0.1 [kline]. Therefore, no relationships among the variables are substantially overestimated by the model.

Modification Indices

In the interpretation of the modification indices we rely mostly on [brown] and [mi]. We cite from [brown]: “The modification index reflects an approximation of how much the overall model χ^2 will decrease if the fixed or constrained parameter is freely estimated.” In other words, if adding a line with a high modification index to the model, i.e. if adding a parameter, the overall goodness-of-fit may be improved. Nevertheless, this should be done only under certain conditions, described in the sequel.

We consider only modification indices greater or equal than 3.84 (which are statistically significant at 5% type I error). Next, we search only for modification indices which achieve a power of minimum 75% in detecting a (relevant) misspecification of at least 0.1 for (residual) correlations, respectively 0.4 for factor loadings. These are characterized in the decision column by the label “epc:m”.

We remark that these conditions are not fulfilled for modification indices with respect to (residual) correlations. Therefore, there exist no significant and relevant modification indices with respect to (residual) correlations.

We remark that these conditions are not fulfilled for modification indices with respect to factor loadings. Therefore, there exist no significant and relevant modification indices with respect to factor loadings.

NULL

Parameter Estimates

Factor Loadings

We remark that the completely standardized factor loadings (section “Completely Standardized Parameter Estimates”) are all statistically significant at 5% type I error. Moreover, in absolute value they are all greater than 0.4. This cutoff-value is considered in some CFA research areas a magnitude that is substantively meaningful [brown]. Please consider also cutoff-values from your particular research area when interpreting the factor loadings. We summarize the interpretation of the completely standardized factor loadings in the next table:

Table 6: Check Completely Standardized Factor Loadings

Latent Variable	Observed Variable	Loading ¹	P-Value	Significant? ²	Relevance ³	Sign ⁴	Check
f	q03	0.54	<0.001	Yes	*	—	Ok
f	q04	-0.70	<0.001	Yes	**	—	Ok
f	q05	-0.57	<0.001	Yes	*	—	Ok

¹ The completely standardized factor loading can be interpreted as the correlation with the factor.

² 5% type I error.

³ Stars correspond to factor loadings cutoff-values: 0.4, 0.6, 0.8.

⁴ No (correct) information available. We assume the signs of the factor loadings correspond to your expectation.

⁵ _____

⁶ _____

⁷ _____

Next, we inspect the unstandardized factor loadings (section “Model Fit Summary”). We remark that the unstandardized factor loadings are all statistically significant at 5% type I error. Therefore, the significance test results for standardized and unstandardized factor loadings coincide (for non-marker variables). We summarize the interpretation of the unstandardized factor loadings in the next table(s):

Table 7: Check Unstandardized Factor Loadings

Latent Variable	Observed Variable	Factor Loading	P-Value	Significant? ¹	Direction
f	q03	1.00			Ok
f	q04	-1.14	<0.001	Yes	Ok
f	q05	-0.95	<0.001	Yes	Ok

¹ 5% type I error.

Table 8: Interpretation of Unstandardized Factor Loadings

Interpretation of Unstandardized Factor Loadings
q03 is marker variable for f
A 1-unit increase in f leads to a -1.14 -unit decrease in the q04
A 1-unit increase in f leads to a -0.95 -unit decrease in the q05

Error in eval(expr, envir, enclos): object 'discr' not found

Factor Discriminant Validity

As noted by [Brown], “the interpretability of the size and statistical significance of factor intercorrelations depends on the specific research context.”

Error in eval(expr, envir, enclos): object 'discr' not found

Error Variances

Table 9: Completely Standardized Error Variances and Communality

Observed Variable	Error Variance ¹	Communality ²³	P-Value	Significant Error Variance? ⁴
q03	0.71	0.29	<0.001	Yes
q04	0.51	0.49	<0.001	Yes
q05	0.67	0.33	<0.001	Yes

¹ Can be interpreted as proportion of unexplained variance by the latent factor(s) (%).

² Corresponds to the squared factor loading.

³ Can be interpreted as proportion of explained variance by the latent factor(s) (%).

⁴ 5% type I error. Typically significant since a large portion of variance is not explained by the latent variable.

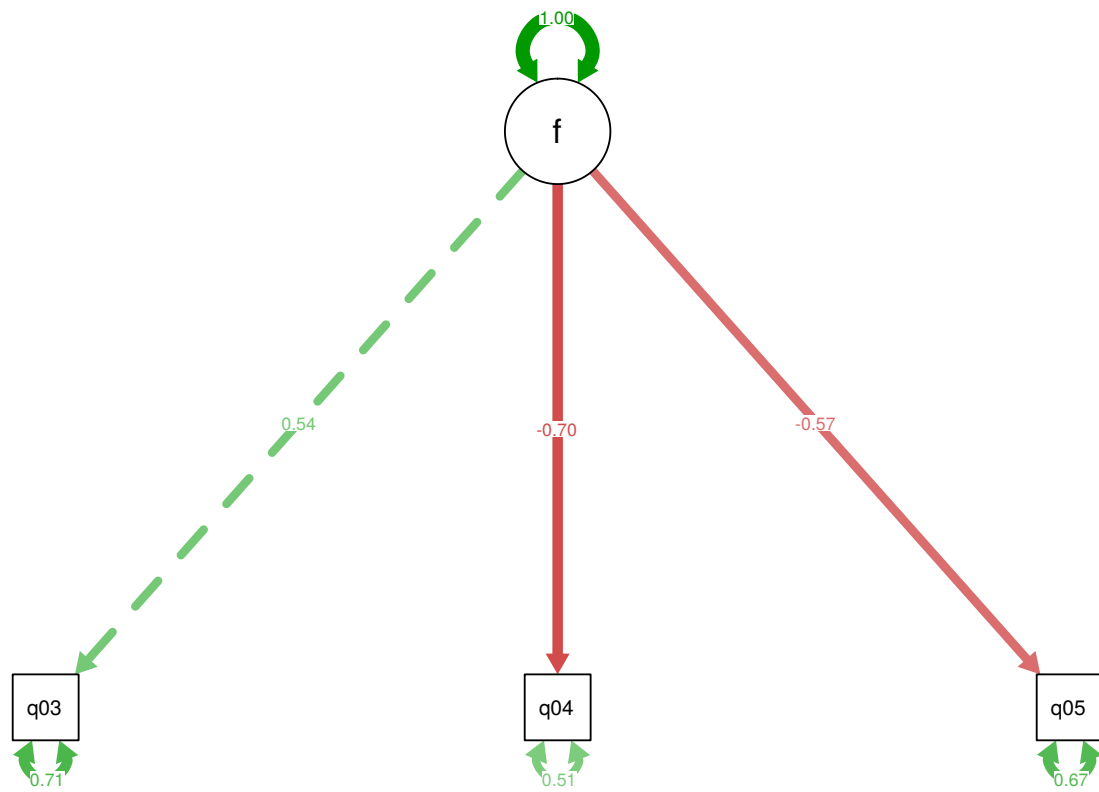
Intercepts

In case of missing values and estimation via FIML, a meanstructure i.e. the intercepts of the observed variables are added to the model. The means of the latent factors are fixed to zero. Therefore, the estimated intercepts within the section “Model Fit Summary” are just the means of the observed variables.

Factor Reliability

We provide the omega measures of factor reliability given by Bentler (Bentler, 1972, 2009) and McDonald (McDonald, 1999) and the average variance extracted (AVE). The interpretability of the reliability measures depend on the specific research context. Nevertheless, omega values below 0.6 or AVE values below 0.5 should be regarded with criticism.

Path Diagram . . . with completely standardized values



References