

Cheat Sheet: CFA

confirmatory factor analysis

Measurement and Evaluation of HCC Systems

Scenario

Use CFA if you measured one or more subjective traits with multiple items. For this cheat sheet we assume three factors (Fa, Fb, Fc) each measured with 5 items (a1, a2, ..., c4, c5) that are measured on 5- or 7-point scales.

Power analysis

- There is no power analysis for CFA, but a general rule of thumb is to have at least 5 participants per item in the model. So for 15 items, one would need an N of at least 75.

Running the test

- Set up the model:

```
model <- '
  Fa =~ a1+a2+a3+a4+a5
  Fb =~ b1+b2+b3+b4+b5
  Fc =~ c1+c2+c3+c4+c5'
```

- Run the model with categorical indicators (`ordered`) and Unit Variance Identification

(`std.lv=T`):

```
fit <- cfa(model, data=data, ordered=names(data), std.lv=T)
```

- Get the model summary, including the communalities (`rsquare=T`) and the alternative fit measures (`fit.measures=T`):

```
summary(fit, rsquare=T, fit.measures=T)
```

Improving the model

- One by one, remove items with low communality (r -square < 0.40). If you have many items, you can use r -square < 0.50 . If you have only few, you can use r -square < 0.30 . Re-run the model every time you remove an item.
- Get the modification indices, keep the ones that are significant, and sort them by value:

```
mods <- modindices(fit)
mods <- mods[mods$mi > 3.84 & !is.na(mods$mi),]
mods[order(mods$mi),]
```

- One by one, remove items with high modification indices ($mi > 8$). If you have only few items left, you can use $mi > 20$. Re-run the model and the modification indices every time you remove an item.
- Go back and forth between checking communalities and modification indices until every items meets the criteria. Try to keep at least 3 items for each factor (2 is the bare minimum).

Inspecting the factor and model fit

- To test the factor fit, calculate the Average Variance Extracted (AVE) for each factor. The AVE of a factor is the average r-squared of the items of that factor.
 - o You have convergent validity if $AVE > 0.50$.
 - o You have discriminant validity if the \sqrt{AVE} is higher than the largest correlation of that factor with any other factor.
- To test model fit, look at the “robust” version of the following indicators:
 - o For perfect model fit, the chi-square test of model fit should be non-significant ($p > .05$). This rarely happens.
 - o An alternative metric is to use: $\chi^2 / df < 3$ (good fit) or < 2 (great fit).
 - o CFI should be > 0.96 and TLI should be > 0.95 .
 - o RMSEA should be < 0.05 , and the upper bound of its confidence interval should not exceed 0.10. The test of $RMSEA \leq 0.05$ should not be significant.

(optional) Calculating Cronbach’s alpha

- CFA is better than Cronbach’s alpha, but sometimes we compute it anyway to show that it is > 0.70 (acceptable fit), > 0.80 (good fit), or > 0.90 (great fit). If you do Chronbach’s alpha after you have done a CFA, you only have to report the `raw_alpha`; you do not have to inspect the “Reliability if an item is dropped”.
- Conduct Cronbach’s alpha for each factor on the items that remain. E.g. if we removed item a2 from factor Fa:


```
alpha(data[,c("a1", "a3", "a4", "a5")])
```

Reporting

- “We conducted a CFA and examined the validity and reliability scores of the constructs measured in our study. Upon inspection of the CFA model, we removed item [item] (communality: 0.xxx) and item [item] (high cross-loadings with several other factors). The remaining items share at least xx% of their variance with their designated construct. The final model has a good model fit: $\chi^2(xxx) = xxx.xxx$, $p = .xxxx$; $RMSEA = 0.xxx$, 90% CI: [0.xxx, 0.xxx]; $CFI = 0.xxx$; $TLI = 0.xxx$ ”

- “To ensure the convergent validity of constructs, we examined the average variance extracted (AVE) of each construct. The AVEs were all higher than the recommended value of 0.50, indicating adequate convergent validity. To ensure discriminant validity, we ascertained that the square root of the AVE for each construct was higher than the correlations of the construct with other constructs.”
- Report the loadings as follows (in this example a2 and b3 are removed):

Construct	Item	Loading
[Fa] Alpha: 0.xx AVE: 0.xxx	[a1]	[estimate]
	[a2]	
	[a3]	[estimate]
	[a4]	[estimate]
	[a5]	[estimate]
[Fb] Alpha: 0.xx AVE: 0.xxx	[b1]	[estimate]
	[b2]	[estimate]
	[b3]	
	[b4]	[estimate]
	[b5]	[estimate]
[Fc] Alpha: 0.xx AVE: 0.xxx	[c1]	[estimate]
	[c2]	[estimate]
	[c3]	[estimate]
	[c4]	[estimate]
	[c5]	[estimate]

- Also report the correlations between factors. The diagonal of this table has $\sqrt{\text{AVE}}$:

	Alpha	AVE	[Fa]	[Fb]	[Fc]
[Fa]	0.xx	0.xxx	.xxx	.xxx	.xxx
[Fb]	0.xx	0.xxx	.xxx	.xxx	.xxx
[Fc]	0.xx	0.xxx	.xxx	.xxx	.xxx