

Inferenční statistika a modelování vazeb latentních proměnných ve strukturálních modelech.

Lubor Homolka

28. května 2014

Obsah semináře

1. Inferenční statistika
2. Analýza χ^2 (s aplikací v R)
3. Strukturální modelování:
 - Postup
 - Ukázka v článku

Inferenční statistika

- Proces zevšeobecňování. Z výběrové statistiky \bar{x} na populační charakteristiku μ .
- Klíčovým bodem je p-value.

$$P(H_0 = \text{PRAVDA} | \text{DATA})$$

vs.

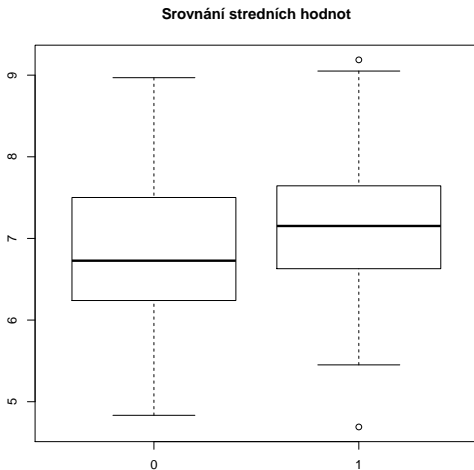
$$P(\text{DATA} | H_0 = \text{PRAVDA})$$

p-value: Jak nepravděpodobná máme data, pokud ve skutečnosti platí H_0 ?

Je průměrná výše u žen vyšší než u mužů?

```
muzi <- rnorm(50, mean=6.8, sd=1)
```

```
zeny <- rnorm(50, mean=7, sd=1)
```



$H_0 : \mu_m = \mu_{\bar{z}}$ oproti alternativní:

$H_0 : \mu_m \neq \mu_{\bar{z}}$

```
> t.test(hodnoty~pohlavi, data=data.all)
```

Welch Two Sample t-test

```
data: hodnoty by pohlavi
```

```
t = -1.7036, df = 97.951, p-value = 0.09162
```

```
alternative hypothesis: true difference in  
means is not equal to 0
```

```
95 percent confidence interval:
```

```
-0.67544851  0.05143907
```

```
sample estimates:
```

```
mean in group 0 mean in group 1
```

```
6.834404
```

```
7.146408
```

Nezamítáme H_0 . Nemáme důkaz pro tvrzení, že ženy mají *jinou* průměrnou hodnotu než muži.

$H_0 : \mu_m = \mu_{\bar{z}}$ oproti alternativní:

$H_0 : \mu_m < \mu_{\bar{z}}$

```
> t.test(hodnoty~pohlavi, data=data.all, alternative="l")
```

Welch Two Sample t-test

```
data: hodnoty by pohlavi
```

```
t = -1.7036, df = 97.951, p-value = 0.04581
```

```
alternative hypothesis: true difference in means  
is less than 0
```

```
95 percent confidence interval:  
-Inf -0.007884767
```

```
sample estimates:
```

```
mean in group 0 mean in group 1  
6.834404        7.146408
```

Na hladině významnosti 5% zamítáme H_0 ve prospěch alternativní hypotézy. Máme v rukou důkaz pro tvrzení, že ženy mají vyšší průměrnou hodnotu než muži.

$H_0 : \mu_m = \mu_{\bar{z}}$ oproti alternativní:

$H_0 : \mu_m > \mu_{\bar{z}}$

```
> t.test(hodnoty~pohlavi, data=data.all, alternative="g")
```

Welch Two Sample t-test

```
data: hodnoty by pohlavi
```

```
t = -1.7036, df = 97.951, p-value = 0.9542
```

```
alternative hypothesis: true difference in means  
is greater than 0
```

```
95 percent confidence interval:
```

```
-0.6161247      Inf
```

```
sample estimates:
```

```
mean in group 0 mean in group 1
```

```
6.834404      7.146408
```

Na hladině významnosti 5% zamítáme H_0 ve prospěch alternativní hypotézy. Máme v rukou důkaz pro tvrzení, že ženy mají vyšší průměrnou hodnotu než muži.

Proporční test

Testuje hodnotu nebo shodu podílů π . Nulová hypotéza tvrdí, že neexistuje rozdíl $H_0 : \pi_1 = \pi_2$

Zeptali jsme se 100 mužů a 100 žen na preferenci daného výrobku.

- Možné odpovědi: líbí/nelíbí
- 75% Čechů se líbí, Slovákům pouze 70 %
- Je možné tvrdit, že existují různé preference?

```
> prop.test(x=c(75, 70), n=c(100, 100))
```

Jak by se situace změnila, kdyby náš vzorek byl 10× větší? (a proporce se zachovaly)

2-sample test for equality of proportions with continuity correction

```
data:  c(75, 70) out of c(100, 100)
X-squared = 0.4013, df = 1, p-value = 0.5264
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.08357103  0.18357103
sample estimates:
prop 1 prop 2
 0.75   0.70
```

```
> prop.test(x=c(750, 700), n=c(1000, 1000))
```

2-sample test for equality of proportions
with continuity correction

```
data:  c(750, 700) out of c(1000, 1000)
X-squared = 6.0213, df = 1, p-value = 0.01413
alternative hypothesis: two.sided
95 percent confidence interval:
 0.00992341 0.09007659
sample estimates:
prop 1 prop 2
 0.75   0.70
```

Analýza kontingenční tabulky

Zjišťujeme, zda struktura kontingenční tabulky získané z našich dat odpovídá té, kterou bychom očekávali, kdyby neexistoval vztah mezi proměnnými.

$$CH = \sum \sum \frac{(n_{ij} - n_{ij}^*)^2}{n_{ij}^*}$$

- V případě splnění podmínek má testové kritérium CH rozdělení χ^2
- Porušení podmínek (min. počet pozorování v buňce 5) neumožňuje srovnání CH s χ^2 .
- V případě porušení podmínek buď `fisher.test`, nebo parametr `simulate.p.value=TRUE` ve funkci `chisq.test`.

Kontingenční tabulka - příklad

Je prokazatelné, že se BSC firmy více hlásí k CSR?

```
> data.chi <- as.data.frame(matrix(c(30, 90, 60, 40),  
byrow=TRUE, ncol=2))
```

```
> data.chi
```

	CSR-ne	CSR-ano
BSC	30	90
non-BSC	60	40

```
> chisq.test(data.chi)
```

Pearson's Chi-squared test with Yates'
continuity correction

```
data: data.chi
```

```
X-squared = 26.2121, df = 1, p-value = 3.059e-07
```

modelování vazeb latentních proměnných ve strukturálních modelech

Šest fází SEM analýzy

1. Definice konstruktů.
2. Návrh studie → empirické výsledky.
3. Vytvoření *measurement* modelu.
4. Určení validity *measurement* modelu.
5. Specifikace *structural* modelu.
6. Určení validity *structural* modelu.
7. Prezentace výsledků.

Fáze I + Pilotní studie

- latent variable, construct - nepřímo měřitelná proměnná,
- measured variable, manifest nebo indicator - změřená hodnota, například otázkou na Likrtově škále,
- kauzalita, korelace a kovariance - kauzalita je přísný koncept vymezující podmíněné chování. Korelace je mírnější forma vztahu. Korelační koeficient je standardizovaná kovariance.

Na základě teorie vymezíme konstrukt (Definice loajality) a vazby mezi konstrukty (orientace a typ).

The items for e-service quality were adopted from the original E-S-QUAL... (p.776)

Šest fází SEM analýzy

1. Definice konstruktů.
2. Návrh studie → empirické výsledky.
3. Vytvoření *measurement* modelu.
4. Určení validity *measurement* modelu.
5. Specifikace *structural* modelu.
6. Určení validity *structural* modelu.
7. Prezentace výsledků.

Fáze II - Návrh studie

- Odlehlé hodnoty, MCAR \sim MAR a řešení (*pairwise* \sim *listwise deletion*), *imputation* metody,
- volba mezi korelací a kovariancí,
- metoda odhadu (OLS \sim LM),
- velikost vzorku.

min	Konstrukty	Indikátory	Komunalita
100	5	3+	0.6
150	7	3+	0.5
300	7	3+	0.45
500	7+	3+	<0.5

- *After refusing some incomplete or invalid questionnaires, 428 valid... p.774*
- *To congregate external validity conditions... p.774*
- *The PLS procedure is able to model latent constructs under conditions of non-normality and small-to-medium samplesize p.779*

Šest fází SEM analýzy

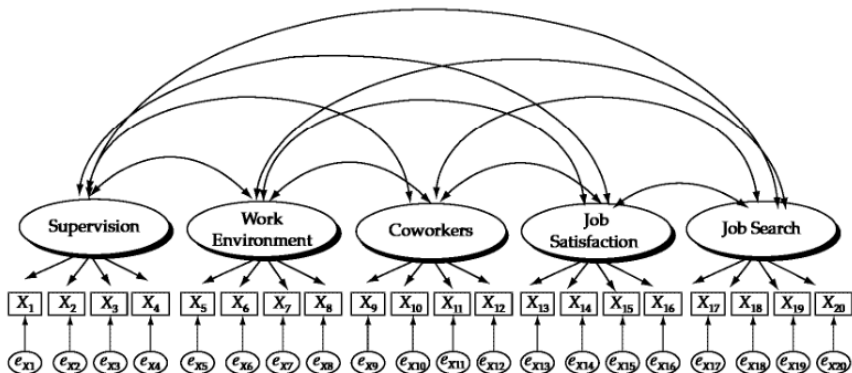
1. Definice konstruktů.
2. Návrh studie → empirické výsledky.
3. Vytvoření *measurement* modelu.
4. Určení validity *measurement* modelu.
5. Specifikace *structural* modelu.
6. Určení validity *structural* modelu.
7. Prezentace výsledků.

Fáze III: Measurement model

- Vytvoření grafického schématu modelu.
- SPSS umožňuje grafický zápis, R vyžaduje zápis textový:

```
konstrukt1 =~ konstrukt2 # kauzální vazba  
konstrukt1 ~~ konstrukt2 # korelační vazba
```

- Existuje několik druhů zápisů tzv. path diagramu.



Šest fází SEM analýzy

1. Definice konstruktů.
2. Návrh studie → empirické výsledky.
3. Vytvoření *measurement* modelu.
4. Určení *validity measurement* modelu.
5. Specifikace *structural* modelu.
6. Určení *validity structural* modelu.
7. Prezentace výsledků.

Fáze IV: Validita Measurement modelu - 1. část

1. Nutné ověřit validitu a reliabilitu konstruktů.
2. Úprava počtu indikátorů.
 - **validita** - míra správnosti otázek ve vztahu k předmětu.
Rozlišujeme několik pohledů,
 - **reliabilita** - interní konzistence mezi indikátory daného konstruktu.

Tato fáze se provádí metodou **Exploratory factor analysis**.
Závisí pouze na datech, nikoli na testované teorii!

Příklad - předměty

```
data.in <- read.csv("data.csv")
```

id	BIO	GEO	CHEM	ALG	CALC	STAT
1	1	1	1	1	1	1
2	4	4	3	4	4	4
3	2	1	3	4	1	1
4	2	3	2	4	4	3
5	3	1	2	2	3	4
6	1	1	1	4	4	4

```
cor(data.in)
```

	BIO	GEO	CHEM	ALG	CALC	STAT
BIO	1.00	0.68	0.75	0.12	0.21	0.20
GEO	0.68	1.00	0.68	0.14	0.20	0.23
CHEM	0.75	0.68	1.00	0.08	0.14	0.17
ALG	0.12	0.14	0.08	1.00	0.77	0.41
CALC	0.21	0.20	0.14	0.77	1.00	0.51
STAT	0.20	0.23	0.17	0.41	0.51	1.00

Ukázka faktorové analýzy

Je možné nalézt \sim tolik faktorů, kolik je proměnných.

	F1	F2		F1	F2
BIO	0.86	0.03	ot1-spok	0.86	0.03
GEO	0.78	0.04	ot2-spok	0.78	0.04
CHEM	0.88	-0.05	ot3-spok	0.88	-0.05
ALG	-0.04	0.80	ot4-spok	-0.05	0.88
CALC	0.01	0.97	ot1-loa	-0.04	0.80
STAT	0.13	0.49	ot2-loa	0.01	0.97
			ot3-loa	0.13	0.49
			ot4-loa	0.2	0.25

Ukázka faktorové analýzy - kolik faktorů?

```
> factanal(data.in,factors=3)
```

Loadings:

	Factor1	Factor2	Factor3
BIO	0.850	0.153	
GEO	0.775	0.151	
CHEM	0.868		
ALG		0.696	0.714
CALC		0.891	0.208
STAT	0.157	0.545	

	Factor1	Factor2	Factor3
SS loadings	2.112	1.624	0.557
Proportion Var	0.352	0.271	0.093
Cumulative Var	0.352	0.623	0.715

```
> factanal(data.in,factors=2)
```

Loadings:

	Factor1	Factor2
BIO	0.855	0.133
GEO	0.779	0.135
CHEM	0.865	
ALG		0.791
CALC		0.971
STAT	0.170	0.506

	Factor1	Factor2
SS loadings	2.124	1.863
Proportion Var	0.354	0.311
Cumulative Var	0.354	0.665

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 2.94 on 4 degrees of freedom.
The p-value is 0.568

4. Results

4.1. Scales to assess e-quality and e-recovery

A preliminary survey of scale dimensionality was undertaken by exploratory factor analyses with varimax rotation (Hair, Anderson, Tatham, & Black, 1998) using the Kaiser criteria of eigenvalues greater than 1.

The *first* of those exploratory factor analyses involved the items for e-quality (as listed above). The Kaiser–Meyer–Olkin (KMO) measure was 0.935. Bartlett’s sphericity test was 5,125.3 ($df = 171$) with a significance of 0.000. Three dimensions, which accounted for 64.11% of the variability of the sample, were identified.

The *first factor* (labelled ‘efficiency’) gathered seven of the eight efficiency items noted above. Only item EFF5 (‘the website loads its pages quickly’) migrated from this factor to the second factor.

The *second factor* was labelled ‘system availability’. In all, seven items loaded on this ‘new’ factor of system availability: (i) the original four items of ‘system availability’ noted above; (ii) the first two items of ‘fulfilment’; and (iii) the item EFF5 (previously noted).

The *third factor* neatly included all three items of ‘privacy’ noted above.

As a result of this analysis, two of the original ‘fulfilment’ items were included in ‘system availability’, and the other two were discarded because they loaded equally on

Validita konstruktů

Míra správnosti: *Indikátory* → *latent construct*

- **Convergent validity:** Indikátory sdílejí velkou část rozptylu s daným faktorem. Testujeme pomocí: výše Loadings, Average Variance Extracted (AVE) a míry reliability (Cronbach α).
 - ▶ Reliabilita - jak je odolná otázka proti nesystematickým vlivům.
Chápou všichni stejně: *výhodný úvěr*?
- **Discriminant validity:** Míra rozdílnosti konstruktů v modelu. Dokážeme odlišit odpovědi spokojenosti od loajality? Testujeme v rámci CFA. Testujeme, zda korelace mezi konstrukty=1. Pokud ano, nedokážeme rozlišit.
- **Nomological, Face validity:** Prokazuje chování modelu předpokládané výstupy (z hypotéz)?

Table 2. Reliability analysis of adapted E-S-Qual and E-RecS-Qual scales.

Factor	Items	Cronbach's alpha	Range for Cronbach's alpha removing one item	Range for correlations of the items and the sum of the subscale
Efficiency	EFF1, EFF2, EFF3, EFF4, EFF6, EFF7, EFF8	0.906	0.884–0.904	0.624–0.801
System availability	SYA1, SYA2, SYA3, SYA4, FUL1, FUL2, EFF5	0.887	0.857–0.884	0.574–0.794
Privacy	PRI1, PRI2, PRI3	0.890	0.819–0.871	0.752–0.813
Responsiveness	RES1, RES2, RES3, COM1	0.835	0.747–0.835	0.573–0.763
Contact	CON1, CON2, CON3	0.771	0.647–0.741	0.559–0.643

4.2.4. *Evaluation of measurement model*

Evaluation of the measurement model involved assessment of: (i) the validity of individual items; (ii) the reliability of subscales (internal consistency); (iii) convergent validity; and (iv) discriminant validity of the constructs.

The *validity of individual items* within the constructs was confirmed by load values greater than 0.707 (Carmines & Zeller, 1979). Table 4 shows that only item 'EFF 5' did not satisfy this criterion, but because it was so close to the threshold it was decided that it could be retained, in accordance with the relaxed criterion suggested by Barclay, Higgins, and Thompson (1995). The robustness of these loads was analysed by a 'bootstrapping' process; that is, taking 1000 subsamples of a hundred cases of each. All the *p*-values were found to be greater than the 1.96 threshold.

The *internal consistency* (reliability) of the subscales had been previously demonstrated (see Tables 2 and 3). In addition, assessment of the reliability of the reflective constructs was checked by composite reliability. As shown in Table 6, all values exceeded the recommended value of 0.8 (Nunnally & Bernstein, 1994).

With regard to *convergent validity*, the average variance extracted (AVE) for all scales was greater than Fornell and Larcker (1981) recommended value of 0.5 (see Table 6).

To assess *discriminant validity*, the criterion of Fornell and Larcker (1981) was adopted – that the square root of the AVE should be greater than the correlations presented by each construct with other constructs. Table 5 shows that this was so.

Fáze 4: Určení validity measurement modelu 2. část

Srovnání očekávané a skutečné kovarianční matice. → testy založené na χ^2 Několik typů testů:

- **Absolute Fit Indices** – pouze jeden model, Goodness-of-fit Index (GFI), RMSEA - směrodatná odchylka rezidua,
- **Incremental Fit Indices** – srovnání výchozí modelu s upraveným modelem, Normed-fit Index (NFI), Tucker-Lewis index (TLI), Comparative Fit Index (CFI).
- **Parsimony Fit Indices** – srovnání více modelů vzhledem k možné komplexitě, Adjusted GOF (AGFI)

In the first of these analyses (for the amended E-S-QUAL), the comparative fit index (CFI) was 0.933 and the root mean-square error of approximation (RMSEA) was 0.065. The Bentler–Bonett non-normed fit index (NNFI) was 0.921. The Satorra–Bentler scaled chi-square was 323.63 on 116 degrees of freedom, and its probability value for the chi-square statistic was 0.000. These results indicate that global fit was acceptable (Byrne, 1994; Hu & Bentler, 1999). The loads were all high (at a significance level of 0.05). The amended model was therefore shown to be an acceptable fit for the data.

Fáze 4: Určení validity measurement modelu 2. část

$$R = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0.3 & 0.3 \\ 0.3 & 1 & 0.3 \\ 0.3 & 0.3 & 1 \end{pmatrix}$$

Je inter-korelace dostatečná?

- $r_{ij} > 0.3$
- Anti-Image correlation matrix $r_{(i,j)k} > 0.7$
- Bartlettův test sféricity matice
- Measure of sampling adequacy (MSA, KMO): Když je $MSA = 1$ tak je možné vysvětlit proměnnou ostatními proměnnými. $(MSA) > 0.6$

Strana 777:

above). The Kaiser–Meyer–Olkin (KMO) measure was 0.935. Bartlett’s sphericity test was 5,125.3 (df = 171) with a significance of 0.000. Three dimensions, which accounted

Strana 782:

Table 5. Discriminant validity.

	Efficiency	System availability	Privacy	Responsi- veness	Contact	Perceived value	Loyalty
Efficiency	<i>0.82024</i>						
System availability	0.685571	<i>0.77666</i>					
Privacy	0.483266	0.532644	<i>0.90559</i>				
Responsiveness	0.323561	0.282561	0.221888	<i>0.82041</i>			
Contact	0.223523	0.162847	0.212933	0.542431	<i>0.82216</i>		
Perceived Value	0.642428	0.614633	0.561749	0.374284	0.294245	<i>0.81747</i>	
Loyalty	0.604303	0.548655	0.497310	0.443064	0.276656	0.710634	<i>0.84343</i>

Note: Correlations between latent variables under the main diagonal. On the diagonal the square roots of AVE in italics.

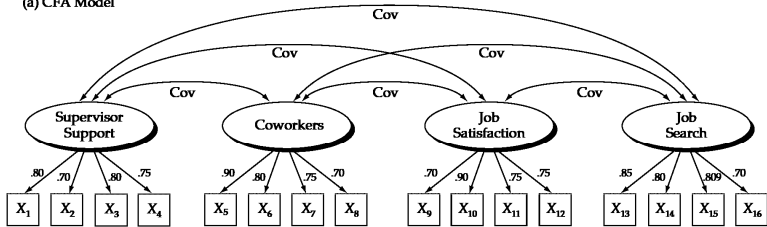
Šest fází SEM analýzy

1. Definice konstruktů.
2. Návrh studie → empirické výsledky.
3. Vytvoření *measurement* modelu.
4. Určení validity *measurement* modelu.
5. Specifikace *structural* modelu.
6. Určení validity *structural* modelu.
7. Prezentace výsledků.

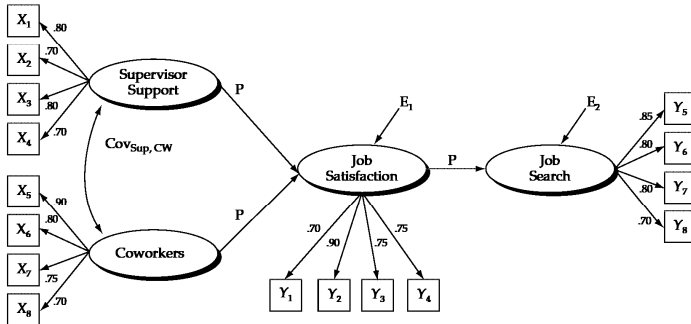
Fáze V: Specifikace *structural* modelu.

- Vytvoření modelu z validních dat na základě našich hypotéz.
- Fixed parameters - nejsou odhadnuty SEM.
- CFA a její speciální případ SEM.

(a) CFA Model



(b) Structural Model



Šest fází SEM analýzy

1. Definice konstruktů.
2. Návrh studie → empirické výsledky.
3. Vytvoření *measurement* modelu.
4. Určení validity *measurement* modelu.
5. Specifikace *structural* modelu.
6. Určení validity *structural* modelu.
7. Prezentace výsledků.

- Využijeme všechny metody z druhého kroku fáze IV
 - **Absolute Fit Indices** – pouze jeden model, Goodness-of-fit Index (GFI), RMSEA - směrodatná odchylka rezidua,
 - **Incremental Fit Indices** – srovnání výchozí modelu s upraveným modelem, Normed-fit Index (NFI), Tucker-Lewis index (TLI), Comparative Fit Index (CFI).
 - **Parsimony Fit Indices** – srovnání více modelů vzhledem k možné komplexitě, Adjusted GOF (AGFI)
- Obvykle stanovíme několik modelů (model-competing strategy) a hledáme například ten nejjednodušší, který je akceptovatelný vzhledem k *face* validitě.
- Inferenční statistika - hypotézy o významnosti vazeb.

Strana 780-781

4.2.5. *Evaluation of structural model*

The goodness-of-fit index proposed by Tenenhaus, Amato, and Esposito (2004), which takes into account both the explained variances for the latent dependent variables and

their commonalities, was 0.3235 (see Table 6). The model explained more than half of the variance of each independent variable.

All the paths, with the exception of that between 'contact' and 'loyalty' (*H6*), were significant ($p < 0.01$) (see Table 7). In other words, five of the six hypotheses were confirmed. The lack of evidence for the direct influence of 'contact' on loyalty (*H6*) in the present study is in accordance with Akinci et al. (2010), who suggested that online service customers might be reluctant to experience direct interpersonal interaction, even when a problem occurs.

Strana 782

Hypothesis	Path coefficient (<i>t</i> -statistic) (*)	Hypothesis testing
<i>H1</i> : Efficiency → Perceived value	0.357304 (3.346563) (*)	Accepted
<i>H2</i> : System availability → Perceived value	0.226775 (2.192262) (*)	Accepted
<i>H3</i> : Privacy → Perceived value	0.268287 (3.026427) (*)	Accepted
<i>H4</i> : Perceived Value → Loyalty	0.636828 (9.646206) (*)	Accepted
<i>H5</i> : Responsiveness → Loyalty	0.221439 (2.251563) (*)	Accepted
<i>H6</i> : Contact → Loyalty	−0.030843 (0.310610)	Refused

Šest fází SEM analýzy

1. Definice konstruktů.
2. Návrh studie → empirické výsledky.
3. Vytvoření *measurement* modelu.
4. Určení validity *measurement* modelu.
5. Specifikace *structural* modelu.
6. Určení validity *structural* modelu.
7. **Prezentace výsledků.**

Prezentace výsledků

- Přesná definice konstruktů, hypotéz a metod sběru dat. → návaznost na dřívější výzkum.
- Klíčová fáze je stanovení validity *measurement* modelu.
 - 1 reliabilita indikátorů a validita konstruktů,
 - 2 GOF testy.
- Testy strukturálního modelu.
- Grafická nebo tabulková interpretace strukturálního modelu.
- Velmi často je dáván k dispozici dotazník → REsearch.

Prezentace je k dispozici:



<https://github.com/luboRprojects>