



Masarykova univerzita  
Fakulta sociálních studií



# PSY252 Statistická analýza dat II

Seminář 7 - Reliabilita



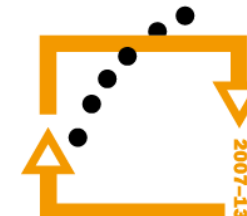
evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,  
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání  
pro konkurenceschopnost



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

# Reliabilita obecně

- Míra zatížení měření nesystematickou chybou
- Obecně  $r$ , které vyjadřuje, jak velká část rozptylu měřené proměnné je rozptyl pravých skóre (true scores).

- Jednotlivá měření

$$X_i = \tau_i + e_i$$

- Variabilita jednotlivých měření

$$\sigma^2_x = \sigma^2_\tau + \sigma^2_e$$

- Reliabilita

$$r_{xx} = \sigma^2_\tau / \sigma^2_x = \sigma^2_\tau / (\sigma^2_\tau + \sigma^2_e)$$

# Obecná užití reliability

## Stanovení intervalu spolehlivosti pro individuální skór

- Náhodná standardní chyba měření  $s_e = s_x \sqrt{1 - r_{xx}}$
- Normální rozložení náhodných chyb  $\gg$  95% interval spolehlivosti pro pravý skór je  $X_i \pm 1,96s_e$  \*

## Úvaha nad vlivem náhodné chyby měření na výsledky analýz

- Korekce korelace na nereliabilitu měření *correction for attenuation*
- $r'_{xy} = r_{xy} / \sqrt{(r_{xx}r_{yy})}$

\* Přesněji se interval staví nikoli kolem pozorovaného skóru  $X_i$ , ale kolem odhadu pravého skóru daného jedince  $T_i = r_{xx}X_i + (1 - r_{xx})M_X$

# Reliabilita specificky

## Reliabilita jako stabilita v čase

- test-restestová reliabilita - korelace opakovaných měření

## Reliabilita paralelních (ekvivalentních, alternativních) forem

## Reliabilita jako vnitřní konzistence

- Split-half
- Cronbachova alfa a Kuder-Richardsonovy koeficienty

## Reliabilita jako shoda posuzovatelů

- rodina koeficientů kappa - Cohenova kappa

# Reliabilita jako vnitřní konzistence

S náhodnou chybou měření bojujeme opakováním měření

- **opakovanou administrací testu -  $r_{tt}$**

*Nelze-li opakovat - test rozpůlíme a považujeme to za opakování :: split-half*

- **korelováním polovin testu -  $r$**  (pomocí Spearman-Brownova vzorce přepočítáme korelaci na reliabilitu)

*Půlit lze mnoha způsoby a lze i čtvrtit ... až ho rozdělíme na jednotlivé položky.*

- **korelováním položek testu - konzistence -  $\alpha$**

Uvažujeme-li test=položka, administrace sady položek je opakovaným administrováním testu

- Administrace v jednom čase > řada náhodných vlivů zůstává konstantní > primárním uvažovaným zdrojem chyby jsou **specifika/nedokonalosti jednotlivých položek** ( $\approx$ unicity).
- Do jaké míry měří testy-položky totéž?

# Cronbachova alfa

## Interpretace

- Alfa je odhadem reliability (jedním z mnoha, zatíženým nejméně výběrovou chybou)
- Podíl rozptylu hypotetické true-scorové škály (složené ze všech myslitelných položek), který by naše škála vysvětlila.
- Korelace škály se všemi možnými stejně dlouhými škálami měřícími týž konstrukt
- Průměrná korelace všech myslitelných půlení testu
- Alfa je projevem unidimenzionality (nikoli důkazem)

Obvykle požadovaná hodnota pro výzkum  $>0,7$ , pro diagnostiku  $>0,9$  (záleží na tom, jak velkou šíři intervalu spolehlivosti můžeme tolerovat).

## Cronbachova alfa je primárně funkcí

- průměrné korelace mezi položkami (čím vyšší, tím vyšší alfa)
- počtu položek (čím vyšší, tím vyšší alfa)

# Cronbachova alfa v SPSS

- **Analyze > Scale > Reliability analysis**
- Vybrat položky, z nichž se má škála skládat.
- Zvolit model „Alpha“
- Ve „Statistics“ zvolit následující volby
  - **Descriptives for:** Item, Scale, Scale if item deleted
  - **Inter-Item:** Correlations
  - **Summaries:** Correlations
- **Syntax**

## RELIABILITY

```
/VARIABLES= seznam položek  
/SCALE('POJMENOVANI SKALY') ALL  
/MODEL=ALPHA  
/STATISTICS=DESCRIPTIVE SCALE CORR  
/SUMMARY=TOTAL CORR.
```

# Tun(z)ing škál

Zjistíme-li, že některá z položek existující škály nepřispívá k vnitřní konzistenci...

- zkontrolujeme, zda nejsou někde chyby
- můžeme to pouze konstatovat (nemáme-li dostatečně reprezentativní vzorek)
- můžeme položku odstranit a analýzu reliability provést znovu
- je-li takových položek více --> FA
- průběžně hypotézujeme o změnách obsahové validity, můžeme-li tak i korelujeme
- **!Příliš mnoho tunění vede k tuzingu: vysoce konzistentní škála z tak mála položek, že nemá žádnou validitu (psychometrický paradox)**



# Reportování reliability

V běžných empirických studiích

- Alfa se udává v sekci *Metody* jako součást informace o metodě měření.

Self-Control. At age 11, self-control was measured using parent ratings of 13 items from a Czech translation of the Carey Temperament Scales (Carey & McDevitt, 1995). Items rated on a six-point scale ranging from almost always (1) to almost never (6) are listed in Appendix A. Internal consistency of the scale was acceptable (Cronbach alpha = .79 ).

# Zadání seminární práce

Proveďte analýzu reliability zvolené škály.

Zamyslete se nad funkčností jednotlivých položek.

Lze z položek, které máte k dispozici, zkonstruovat lepší škálu?

# Děkujeme za pozornost



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ