

# Položková analýza

PSYb2590: Základy psychometriky (Seminář 4)

15./22. 4. 2024

# Položková analýza (PA)

## Metoda empirického hodnocení a výběru položek.

- *Jak moc platí pro každou položku, že je dobrým indikátorem měřeného konstruktů? Jak dobře položka „drží“ s ostatními položkami?*
- *Máme položky měřící různé úrovně konstruktů?*
- **Kriteriální/externí**
  - Predikuje položka sama o sobě nějaké kritérium?
  - PA za účelem zvýšení (kriteriální) validity.
- **Content sampling/interní**
  - Souvisí položka s ostatními položkami?
  - PA za účelem zvýšení reliability (vnitřní konzistence).
- V CTT je PA vždy vztažena k fungování položky v daném vzorku - deskriptivní statistiky (bez statistické inference).

# Položková analýza 1: deskriptivy

- Deskriptivy položek
  - Průměr, SD, rozložení/zešikmení.
  - Frekvence distraktorů/možností.
  - Srovnání rozdělení skupin.
- Deskriptivy celkového skóre
  - Průměr, SD, rozložení/zešikmení, normalita.
  - Efekt stropu/podlahy.

# Položková analýza 2: popularita

- Jak je položka obtížná/snadná?
- Obtížnost/popularita položky.
- Průměr „správných“ odpovědí – u binárních položek tedy % osob, které odpověděly „správně“. Např.:
  - středně obtížná položka:  $p = 0,5$
  - celkem snadná položka:  $p = 0,7$
  - obtížná položka:  $p = 0,25$
  - extrémně obtížná položka:  $p = 0,05$

# Položková analýza 2: popularita

- Jak je položka obtížná/snadná?
- Některé položky jsou skórovány více body.
  - Pokud se liší maxima, nelze přímo srovnávat průměry.
  - Lze ale převést na popularitu v rozmezí 0–1:

$$p = \frac{M - \min}{\max - \min}$$

(M = průměrné skóre, min/max = minimální/maximální možné skóre)

- Jak interpretovat popularitu u dotazníků (bez správné odpovědi)?
  - Např. Likertova škála nebo ano/ne?

# Položková analýza 3: diskriminace

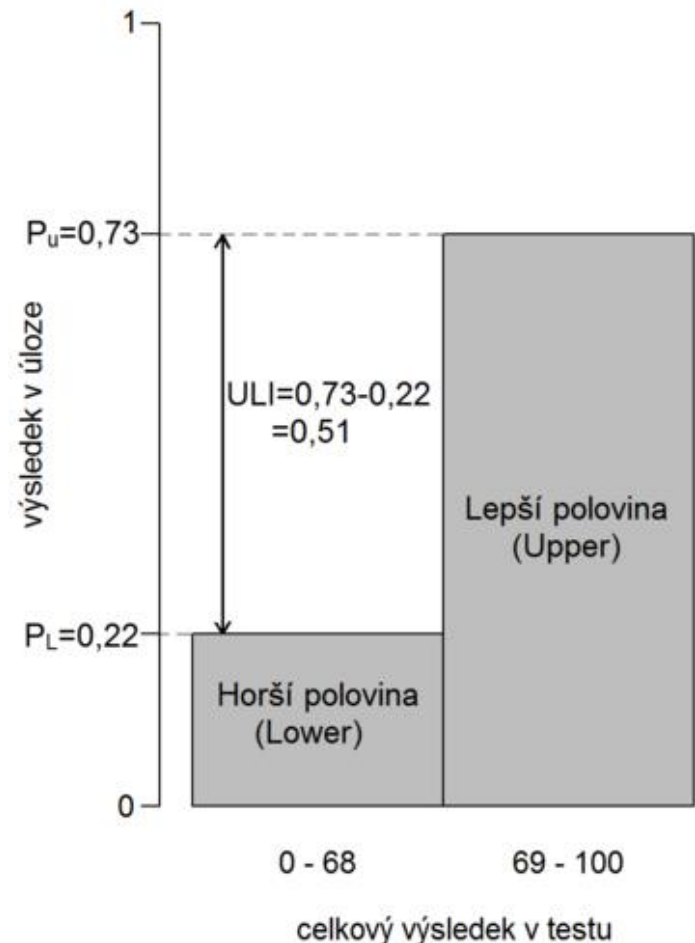
- Jak těsně položka souvisí s celkovým skórem?
- Celkové skóre je funkcí všech položek – jde tedy (i) o to, jak položka souvisí s ostatními položkami „dohromady“ či „průměrně vzato“ (řečeno hodně zjednodušeně)
- Respondenti s vyšším skóre by měli odpovídat spíše správněji i na každou z položek.

# Položková analýza 3: diskriminace

- Jak těsně položka souvisí s celkovým skórem?

- ULI: upper-lower index

- Někdy též „citlivost“ položky.
- $ULI = M_{upper} - M_{lower}$  - tedy rozdíl popularity lépe a hůře úspěšných respondentů.
- rozdělení: podle mediánu, horní a spodní tercil...
- Výsledkem je hodnota 0–1 udávající diskriminační účinnost položky.



# Položková analýza 3: diskriminace

- Korelace položky s testem (s celkovým skóre).
  - Item-total correlation.
  - Ale! - položky s vysokým rozptylem ovlivňují více celkový skór, a proto s ním i těsněji korelují
  - Korelace bude také vždy nadhodnocena kvůli tomu, že skóre položky je obsaženo v celkovém skóre
- **Korigovaná korelace položky s testem**
  - Corrected item-total correlation.
    - Někdy též „item-rest correlation“.
  - Korelace položky s celkovým skóre bez této položky, tedy se součtem všech ostatních položek.
- Existují i jiné typy korekce (viz `psych` package v R).



# Korelace?

- Co všechno ovlivňuje korelaci?
  - Rozptyl, úroveň měření, počet kategorií
  - Strop/podlaha
- U dichotomických položek se někdy uvádí bodově biseriální korelace (*point-biserial correlation*) – je ale ekvivalentní Pearsonově korelaci
- Kategorické položky – korelaci nahradí rozdíl průměrů celkového skóru (T)
  - Rozdíl průměrů lidí s odpovědí s diagnostickou odpovědí a ostatních
  - U multiple-choice – analýza distraktorů

# Položková analýza 3: diskriminace

- Reliability if item deleted
  - Jak se změní vnitřní konzistence, když položku vyřadíme?
  - Závislé na délce škály, používat jen jako doplňkový index!
- Cronbach alpha if item deleted / omega if item deleted atd.
  - Cronbachova alfa po vyřazení položky.
- Pozor! Vyšší ukazatel → horší položka!

# Položková analýza 3: **diskriminace**

- Explorační faktorová analýza (EFA)
  - Faktorové náboje (loadings) položek v EFA.
  - Jde o odhad standardizovaného regresního koeficientu („korelace“) položky s pravým skóre.
- Analýza hlavních komponent (PCA)
  - Postup blízký faktorové analýze – položková analýza je jedno z mála smysluplných použití.
- PCA s extrakcí jediného faktoru lze použít i pro empirickou identifikaci reverzně kódovaných položek.

# Položková analýza 4: Doporučení

## **Obecná pravidla neexistují!**

...Vždy záleží na účelu metody, měřeném atributu, dimenzionalitě a dalších okolnostech.

# Položková analýza 4: Doporučení

- Rozptyl/směrodatná odchylka:
  - Čím větší, tím lepší.
- Korigovaná korelace/faktorové náboje:
  - Alespoň  $>0,3$  (10 % rozptylu).
- Popularita:
  - Efekt podlahy/stropu snižuje rozptyl.
  - V osobnostních testech tedy zpravidla  $0,2 < p < 0,8$ .
  - Ale občas chceme i položky extrémně snadné/těžké.
    - Např. u inteligenčních testů položky bývají rovnoměrně rozmístěny na škále latentního rysu od nejsnazších po nejtěžší.
- Je vhodné podívat se i přímo na korelační matici.

# Položková analýza 4: Doporučení

Typický postup:

- Zkontroluji celkovou reliabilitu škály, zda mi stačí.
- Pak srovnávám položky „vzájemně“.
  - Pokud všechny položky mají  $r^* > 0,7$ , tak  $r^* = 0,5$  bude málo.
  - Pokud mají všechny položky  $r^* < 0,5$ , tak mi asi bude stačit i  $r^* = 0,3$ .
- **V žádném případě nevyřazuji položky jen kvůli číslům, dbám i na jejich obsahovou validitu!**
- „Klinická zkušenost“ 😊

# Položková analýza 4: Doporučení

Účelem PA může být:

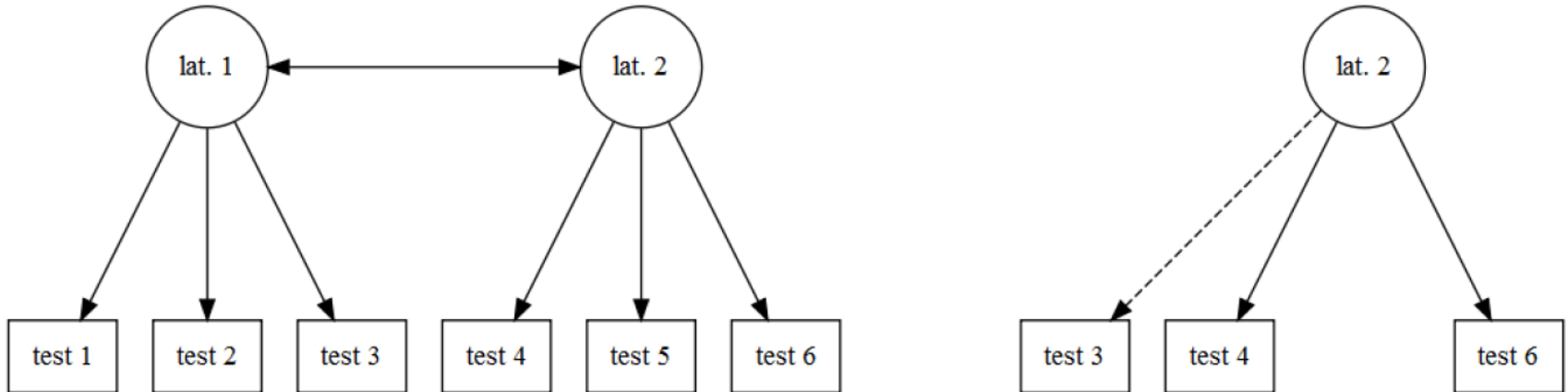
- Vývoj nové metody.
- Ověření, že zavedená metoda funguje dobře.
  - Že se „nic nepokazilo“ např. v mém vlastním výzkumu, BP a podobně.
- Ve druhém případě je velmi špatný nápad vyřazovat položky bez jednoznačného důvodu!
  - Do složení testu sahám jen tehdy, když nemám na výběr.

# Software

- Excel
- SPSS
- R (zejm. psych package)
- **JASP**
- JAMOVI
- ShinyItemAnalysis
  - <https://shiny.cs.cas.cz/ShinyItemAnalysis/>



# LOVE!



- Specifikační chyba, L.O.V.E., OVE („/left/ out variable error“).
- Nezahrnutí „všech možných“ indikátorů mění význam měřeného konstruktů.
  - A nebo náhodného vzorku z obsahového univerza.
- Výběr se nemusí týkat jen položek, ale i celých subtestů.
  - (CHC teorie, inteligenční testy).
- Každá část univerza může mít různě těsný vztah s „průměrem“ univerza.
- Operacionalismus, CTT: Indexy v PA dávají smysl jen v dané populaci a v kontextu dalších konkrétních položek!