

Seminář z psychometrie 18./20. 3. 2024: Explorační faktorová analýza

1. Ve studijních materiálech je dataset s názvem *holzinger.csv*. Otevřete si ho v programu JASP.

Jedná se o dataset od Holzinger a Swineforda (1939), který se skládá z výsledků testů mentálních schopností sedmých a osmých tříd dětí ze dvou různých škol (Pasteur a Grant-White). V původním datasetu jsou výsledky pro 26 testů; vy máte k dispozici pouze 9. Jedná se o klasický dataset pro výukové účely. Kdo na tomto datasetu nedělal faktorku, jako by nebyl (to, že u některých položek není dohledatelné, jak vlastně úkol vypadal, nikoho netrápí).

2. Prozkoumejte dataset a proměnné v něm:

- id – identifikátor
- sex – gender
- ageyr – věk, roky
- agemo – věk, měsíce
- school – škola (Pasteur nebo Grant-White)
- grade – třída
- x1 – vizuální percepce
- x2 – kostky
- x3 – Lozenges („pastilky“) – prostorový test
- x4 – porozumění textu
- x5 – doplňování vět
- x6 – význam slov
- x7 – rychlostní sčítání
- x8 – rychlostní teček
- x9 – rychlostní rozlišování přímých a zakřivených písmen

3. Jaké jsou deskriptivy jednotlivých testů? Vizualizujte si jejich rozložení pomocí histogramů.
4. Podívejte se na korelační (resp. variačně-kovarianční) matici položek (barevně stupňovanou). Co z ní můžete vyčíst?
5. Podívejte se současně na znění položek. Mají korelující položky spolu něco společného?
6. Udělejte scree plot a Hornovu paralelní analýzu. Poté odhadněte počet faktorů podle Kaiserova pravidla. Kolik faktorů vám tyto metody naznačují? Jaká jsou slabá místa těchto metod pro odhad počtu faktorů?
7. Na základě všech dosavadních informací, k jakému modelu (tj. s kolika faktory) se přikláníte? Klidně vymyslete více variant, každou si zkuste podepřít argumenty.
8. Vyberte si jedno z řešení (tj. jeden z modelů) z předchozího úkolu a rozhodněte se pro šikmou nebo ortogonální rotaci podle toho, co vám přijde teoreticky vhodnější.
9. Realizujte EFA se zvolenou rotací. Vyberte si takovou metodu odhadu, která se lépe hodí pro Vaše data a teoretické předpoklady.

Šikmé (oblique) rotace je např. promax nebo oblimin.

Ortogonální rotace je např. varimax.

10. Zkontrolujte, jestli řešení neobsahuje Heywoodův případ. Pokud ano, zvažte, proč tomu tak je a zkuste jiný model.

Heywoodův případ nastává v případě, že je nějaký rozptylový parametr odhadnutý záporně. Stává se to hlavně v případech, kdy je model příliš složitý (máte moc faktorů), nebo když v modelu „zlobí“ nějaká položka.

11. Kolik % rozptylu vysvětluje první faktor? Kolik druhý? Kolik % rozptylu vysvětlují všechny faktory dohromady?

12. Spočítejte komunality pro položky **x2** a **x6**.

Komunalita je podíl pozorovaného rozptylu, který je způsobem obecnými (společnými) faktory.

$$\text{Komunalita} = \frac{\text{společný rozptyl}}{\text{pozorovaný rozptyl}} = 1 - \frac{\text{unikátní rozptyl}}{\text{pozorovaný rozptyl}}$$

13. Prozkoumejte faktorové náboje položek u jednotlivých faktorů. Co z nich můžete vyčíst?

14. Prozkoumejte korelace mezi faktory. Co z nich můžete vyčíst? Jak souvisí hodnota korelací s metodou rotace, kterou jste si zvolili?

Bonusy:

1. Podívejte se na ukazatele shody modelu s daty (RMSEA, SRMR, TLI, CFI). Co vám naznačují? Více o ukazatelích třeba [tady](#).
2. Nyní proveďte EFA pro model s jiným počtem faktorů. Využijte stejnou rotaci a stejný estimátor jako u předchozího modelu.
3. Porovnejte modely. K jakému řešení byste se přiklonili?