Predikce profilu řidičských chyb v provozu na základě inventáře stylu jízdy: kanonická korelační analýza[1]

Úvod a cíl

V dopravněpsychologickém výzkumu se často hodí mít rychlé vodítko: dotazník máme k dispozici hned, ale chyby v reálném provozu vidíme až časem. Napadlo nás proto zkusit jednoduchou věc – jestli se z krátkého inventáře stylu jízdy dá aspoň orientačně poznat, jaký typ chyb se u řidiče v provozu objevuje častěji. V našem souboru máme dvě celé sady proměnných: deset položek inventáře (p1–p10) a pět typů chyb z delšího monitoringu. Místo desítek jednotlivých korelací, ve kterých se člověk snadno ztratí, používáme kanonickou korelační analýzu (CCA). Ta z obou sad poskládá souhrnná skóre tak, aby si co nejvíc odpovídala, a vytáhne z dat několik čitelných vzorců, které má smysl věcně interpretovat (Sherry & Henson, 2005).

Data

Analyzovali jsme data 600 řidičů (N = 600). Chybovost v provozu je popsána pěti proměnnými: riskantní předjíždění, jízda na červenou, překročení povolené rychlosti, dezorientace a nedostatečná plynulost jízdy. Styl jízdy měří deset položek inventáře (p1–p10) na škále 0–10. V datech nebyly chybějící hodnoty. Před CCA jsme proměnné standardizovali (z-skóry).

Metoda

CCA vytváří lineární kombinace proměnných z obou sad (kanonické variáty) a postupně hledá několik nezávislých párů variát, které maximalizují jejich vzájemnou korelaci. Prakticky to znamená, že z položek inventáře vznikne jedno souhrnné skóre, z chyb druhé – a metoda najde takové složení, aby spolu souvisela co nejvíc. Počet funkcí, které dává smysl interpretovat, určujeme postupným testováním přes Wilksovu lambdu. Ta shrnuje, kolik společné vazby mezi sadami ještě zůstává nevysvětleno. Čím je λ menší, tím silnější je celkový vztah. Postupné testování pak ukáže, jestli další funkce přidává smysluplnou informaci, nebo už jen drobné dobarvení (Pennsylvania State University, n.d.).

Pro věcnou interpretaci se opíráme hlavně o strukturální koeficienty, tedy korelace původních proměnných s kanonickými variáty. Ty se čtou podobně jako běžné korelace a často bývají stabilnější než samotné váhy ve vážených součtech, zvlášť když se proměnné v jedné sadě překrývají (Sherry & Henson, 2005). Mahalanobisova vzdálenost nenaznačila extrémní odlehlé hodnoty a multikolinearita byla nízká (VIF inventář 1.014–1.119; chyby 1.027–1.150).

[1] *Data a další informace o této zprávě jsou dostupné na adrese https://dostal.vyzkum-psychologie.cz/stat4?i=631*

Prakticky tím říkáme, že v datech nebyl žádný řidič, který by byl jako celek tak výjimečný, aby výsledky nepřiměřeně táhl jedním směrem, a že proměnné v každé sadě nejsou mezi sebou natolik překryté, aby se kvůli tomu výsledné kombinace stávaly nestabilní nebo špatně interpretovatelné.

Výsledky CCA

Tabulka 1 shrnuje dvě kanonické funkce, které vyšly statisticky významné. První funkce má kanonickou korelaci r = 0.509 (sdílená variance $r^2 = 0.259$) a druhá r = 0.388 ($r^2 = 0.150$) – obě vyšly významné při postupném testu ($p < .001$). Pro praktičtější představu uvádíme i redundanci (kolik rozptylu jedné sady se přes funkci přenese do druhé) – u první funkce je redundance pro chyby 0.077, u druhé 0.035.

Redundance se dá číst hodně prakticky: říká, kolik reálné informace se přes danou funkci přenese z jedné sady do druhé. Když je redundance pro chyby 0.077, znamená to zhruba 7.7 % rozptylu v pěti chybách, který souvisí se souhrnným skórem z inventáře v první funkci (u druhé funkce je to asi 3.5 %). Není to přesná předpověď jednotlivce, spíš orientační síla vztahu.

Tabulka 1: Souhrn kanonických funkcí (CCA)

Funkce	R	r^2	Wilksova λ (od funkce)	χ^2	df	P	Var. extr. X	Redund. X	Var. extr. Y	Redund. Y
1	0.509	0.259	0.619	283.145	50	< .001	0.132	0.034	0.296	0.077
2	0.388	0.150	0.836	105.836	36	< .001	0.147	0.022	0.231	0.035

Poznámka: r = kanonická korelace; r^2 = sdílená variance. Wilksova λ je uvedena pro postupné testování od dané funkce. Var. extr. = průměrná vysvětlená variance v sadě (průměr z r^2 strukturálních koeficientů). Redund. = Var. extr. × r^2 (kolik rozptylu jedné sady se přes funkci přenese do druhé).

Tabulky 2a a 2b ukazují strukturální koeficienty pro obě funkce. Zvýrazňujeme hodnoty s $|r| \geq 0.400$, aby bylo na první pohled vidět, které proměnné danou funkci nejvíc táhnou.

Tabulka 2a: Strukturální koeficienty pro chyby (Y)

Proměnná	F1	F2
Riskantní předjíždění	**0.784**	-0.055
Jízda na červenou	**0.540**	0.361
Překročení povolené rychlosti	**0.746**	0.044
Dezorientace	-0.138	**0.851**
Nedostatečná plynulost jízdy	-0.012	**0.544**

Poznámka: Uvedeny jsou strukturální koeficienty (korelace proměnných s kanonickými variáty). Tučně jsou zvýrazněny hodnoty s $|r| \geq 0.400$.

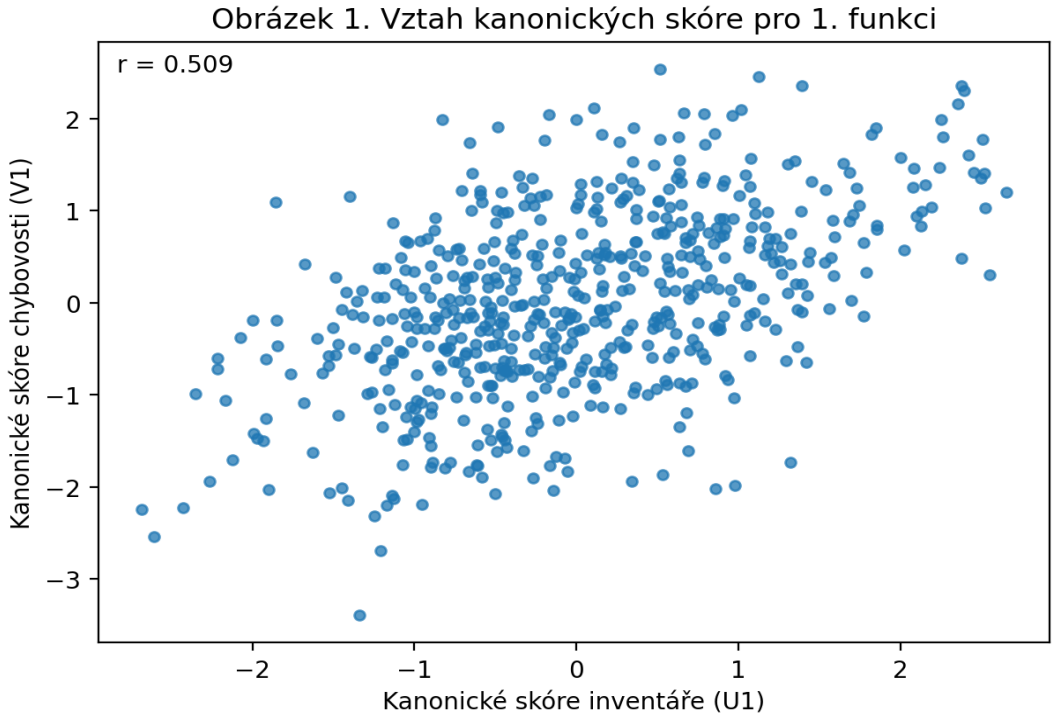
Tabulka 2b: Strukturální koeficienty pro inventář (X)

Proměnná	**F1**	**F2**
p1	**0.570**	0.073
p2	-0.056	**0.458**
p3	-0.048	-0.067
p4	**0.852**	0.074
p5	0.019	0.235
p6	-0.278	**0.732**
p7	**0.411**	0.332
p8	0.116	-0.107
p9	0.000	**0.424**
p10	0.042	**0.592**

Poznámka: Uvedeny jsou strukturální koeficienty (korelace proměnných s kanonickými variáty). Tučně jsou zvýrazněny hodnoty s $|r| \geq 0.400$.

Postupné testování přes Wilksovu lambdu ukázalo, že vztahy mezi sadami nejsou náhodné, a vymezilo dvě interpretovatelné funkce. První funkce spojuje část položek inventáře (zejména p4, p1 a p7) s rizikovějšími chybami (předjíždění, jízda na červenou, překročení rychlosti). Druhá funkce se váže spíše k dezorientaci a plynulosti a v inventáři ji nejlépe vystihují p6, p10 a p2. Nejsilněji se k první funkci váže položka p4.

Obrázek 1: Vztah mezi prvním kanonickým variátem inventáře (U1) a prvním kanonickým variátem chybovosti (V1)



Poznámka: Každý bod představuje jednoho řidiče; osy jsou kanonické variáty (standardizované souhrnné skóre).

Psychometrická kontrola inventáře (IRT)

CCA ukázala, že z odpovědí v inventáři se dá poskládat profil, který souvisí s chybovostí v provozu. Abychom ale nezůstali jen u samotné existence vztahu, doplnili jsme krátkou psychometrickou kontrolu pomocí teorie odpovědi na položku (IRT). Ta odpovídá na praktickou otázku, jak pevný je inventář jako měřicí nástroj, tzn. které položky skutečně rozlišují mezi řidiči a ve které části škály inventář měří nejpřesněji. Pro ordinální položky s více kategoriemi je vhodný model odstupňovaných odpovědí (graded response model; Samejima, 1969). A protože v CCA vyšla jako klíčová zejména položka p4, ukazujeme na ní i křivky odpovědi, aby bylo zřejmé, co přesně znamená, že položka rozlišuje (Embretson & Reise, 2000; Samejima, 1969).

Protože původní škála 0–10 měla v horních hodnotách u některých položek velmi málo odpovědí, sloučili jsme odpovědi do čtyř kategorií (1 = 0–2, 2 = 3–4, 3 = 5–6, 4 = 7–10). Tím jsme zajistili, že každá kategorie má dostatek pozorování (minimum = 24), což je běžný postup, když jsou některé kategorie nepoužívané (Linacre, 2002).

Odhadli jsme jednorozměrný model odstupňovaných odpovědí (Samejima, 1969). Výsledky interpretujeme opatrně, jelikož inventář je zjevně vícerozměrný (Cronbachovo α na původních položkách = 0.268; po sloučení kategorií = 0.254). Takže jednorozměrný IRT model zde spíše shrnuje obecný sklon volit vyšší odpovědi než jednu jasně ohraničenou vlastnost.

Celková přesnost inventáře vychází spíše střední až nižší: marginální reliabilita odvozená z testové informace je 0.443 (Nicewander, 2018). Obrázek 2 naznačuje, že inventář měří nejpřesněji ve střední části škály (I(θ) zhruba kolem 0.796) a směrem k extrémům přesnost klesá.

Stručně řečeno – inventář měří nejlépe ve střední části škály a položky se liší v tom, jak dobře rozlišují respondenty. To pomáhá chápat, proč některé položky v CCA nesou většinu predikční informace.

Tabulka 3: Parametry položek v modelu odstupňovaných odpovědí (IRT)

Položka	Diskriminace a	Prahy b1	Prahy b2	Prahy b3	Informace I(θ=0)
p1	0.370	-8.354	-1.574	4.497	0.039
p2	0.702	-4.581	-0.895	2.891	0.134
p3	0.203	-14.169	-2.717	8.117	0.012
p4	0.451	-7.210	-1.152	3.791	0.058
p5	0.209	-12.907	-2.536	8.232	0.013
p6	0.567	-5.301	-1.024	3.243	0.090
p7	0.730	-4.333	-0.804	2.570	0.148

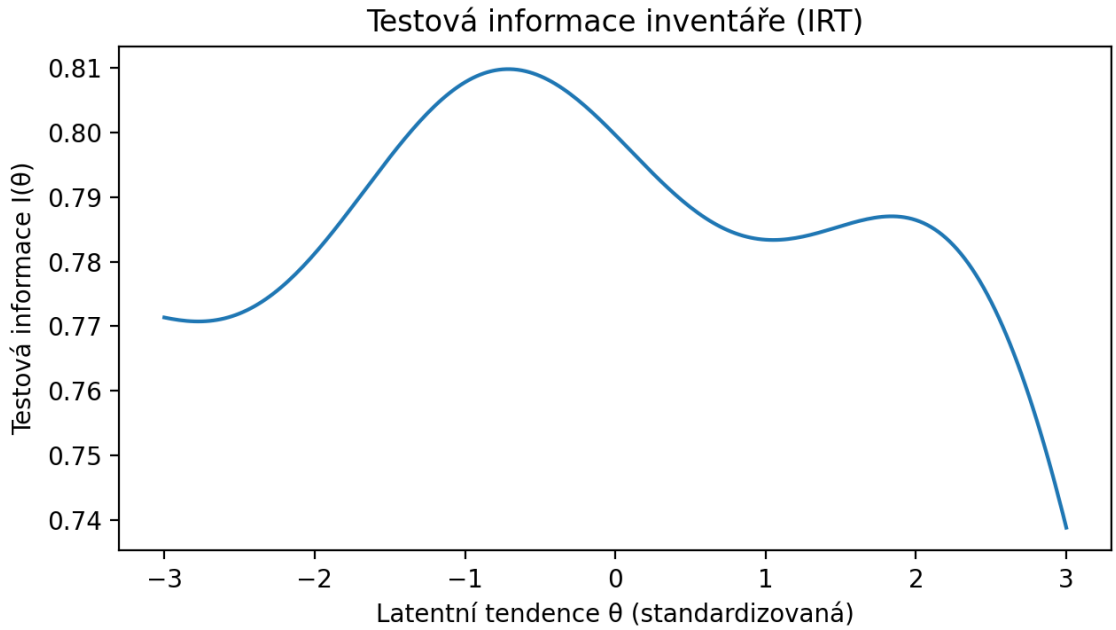
p8	0.200	-15.009	-2.527	8.282	0.011
p9	0.850	-3.814	-0.699	2.246	0.199
p10	0.581	-5.426	-0.808	3.066	0.096

Poznámka: Odhad je proveden na sloučených kategoriích (1 = 0–2, 2 = 3–4, 3 = 5–6, 4 = 7–10). Diskriminace *a* vyjadřuje, jak ostře položka rozlišuje; prahy b1–b3 říkají, kde na škále θ roste šance volit vyšší kategorie. I(θ=0) ukazuje příspěvek položky kolem průměrné úrovně; reliabilitu v bodě θ lze aproximovat jako I(θ)/(I(θ)+1).

Nejpřímější ukazatel síly položky je diskriminace *a*. Čím je vyšší, tím citlivěji položka reaguje na rozdíly mezi lidmi. V našem inventáři vycházejí jako nejvíce rozlišující p9, p7 a p2, zatímco p3 a p5 jsou slabší. Zároveň p4, která v CCA nejvíc nese 1. funkci, není psychometricky mezi nejsilnějšími – může být věcně důležitá pro vztah k chybám, aniž by sama o sobě byla nejpřesnější měřicí položkou.

V Tabulce 3 jsou ještě IRT prahy (b1–b3). Dá se na ně dívat jako na tři mezníky na skryté škále θ – ukazují, při jaké úrovni této tendence se člověk začne častěji přesouvat z nižších odpovědí do vyšších kategorií. Čím jsou prahy posunuté víc doprava, tím náročnější je položka na vysoké odpovědi (vyšší hodnoty volí až lidé s vyšší θ). Proto prahy nečteme jako body na škále 0–10, ale hlavně relativně, tedy pro srovnání položek mezi sebou.

Obrázek 2: Testová informace inventáře napříč latentní tendencí θ

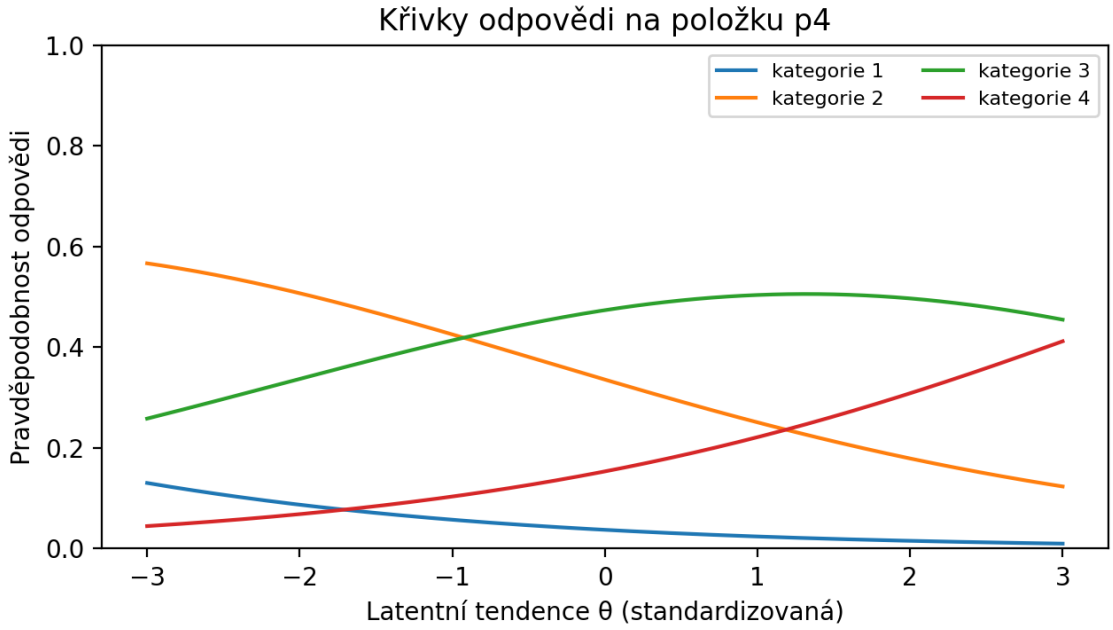


Poznámka: Testová informace I(θ) vyjadřuje přesnost měření; vyšší křivka znamená menší chybu odhadu.

U p4 ukazujeme i křivky odpovědi (Obrázek 3): střední kategorie jsou málo pravděpodobné a odpovědi se soustředí spíše do krajů škály. Prakticky to naznačuje, že by inventáři mohlo prospět

zjednodušení odpovědní škály – méně kategorií může paradoxně znamenat čitelnější a přesnější měření.

Obrázek 3: Křivky odpovědi na položku pro p4, která v CCA nejvíce přispívá k 1. kanonické funkci (profil riskantních chyb)



Poznámka: Křivky ukazují pravděpodobnosti čtyř odpovědních kategorií v závislosti na latentní tendenci θ.

Diskuze a závěr

Našli jsme dva relativně srozumitelné vzorce mezi inventářem a chybami v provozu. První vzorec spojuje zejména p4, p1 a p7 s rizikovými chybami (předjíždění, jízda na červenou, překročení rychlosti). Druhý vzorec se pojí spíše s dezorientací a ztrátou plynulosti a v inventáři je zachycen hlavně p6, p10 a p2. V tomto smyslu má inventář potenciál naznačit, jaký typ chybovosti může být u řidiče pravděpodobnější – zatím ale mluvíme o vztahu v jednom souboru, ne o ověřené predikci na nových datech.

IRT doplňuje, že inventář jako jednorozměrná škála vychází spíše slabě (nízké α i marginální reliabilita), což podporuje dojem, že měří více dimenzí. Pro praktické použití by proto dávalo smysl inventář rozdělit na dílčí škály (např. faktorovou analýzou) a pak ověřit měření i predikční výkon na nezávislém vzorku.

Položka p4 je v CCA klíčová pro 1. funkci, ale v IRT vychází její diskriminační schopnost spíše slabší. Ukazuje to, že položka může být důležitá pro vztah inventáře k profilu chybovosti, aniž by sama o sobě byla ideální měřítko. Pokud bychom inventář chtěli zpřesnit jako nástroj, měli bychom zvážit úpravu odpovědních kategorií nebo přeformulování p4 tak, aby lépe zachycovala střed škály.

Zdroje

Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. Lawrence Erlbaum Associates.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. Journal of Applied Measurement, 3(1), 85–106.

Nicewander, W. A. (2018). Conditional reliability coefficients for test scores. Psychological Methods, 23(2), 351–362. https://doi.org/10.1037/met0000132

Pennsylvania State University. (n.d.). Lesson 13: Canonical correlation analysis. STAT 505: Applied multivariate statistical analysis. https://online.stat.psu.edu/stat505/lesson/13

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph Supplement, 34(4, Pt. 2).

Sherry, A., & Henson, R. K. (2005). Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. Journal of Personality Assessment, 84(1), 37–48. https://doi.org/10.1207/s15327752jpa8401_09