

# Faktorová analýza testu S. Harterové

## Cíl

Zabýváme se využitím metod explorativní faktorové analýzy při statistickém zpracování výsledků testu S. Harterové, v němž třídní učitelé hodnotí děti v patnácti položkách na čtyřbodové stupnici. Datový soubor byl získán v rámci grantového projektu GAČR 406/02/D143 a obsahoval údaje o šedesáti dětech.

## Úvod

Institut výzkumu dětí, mládeže a rodiny při FSS MU se mimo jiné zabývá problematikou rozumově nadaných dětí tj. těch, jejichž IQ je aspoň 130. Výzkumný tým vedený Mgr. Šárkou Portešovou, Ph.D shromáždil v rámci grantu GAČR údaje o několika desítkách nadaných dětí vybrané věkové skupiny a také kontrolní skupiny dětí, které nesplňují kritéria pro zařazení mezi rozumově nadané děti. V průběhu podrobného klinického vyšetření dítěte byly získány údaje týkající se např. intelektových schopností dětí, jejich sociálních, emocionálních a osobnostních charakteristik (sebepojetí, sebehodnocení a vědomí vlastní účinnosti) a dále bylo zaznamenáno hodnocení dětí učiteli a matkami.

V tomto příspěvku se zaměříme na jednu oblast tohoto rozsáhlého a v ČR ojedinělého výzkumu, a to na oblast hodnocení dětí jejich třídními učiteli. Pro tyto účely byl použit v USA velmi rozšířený test S. Harterové nazvaný Teacher's Rating Scale of Child's Actual Behavior (Harter, 1985). V 15 položkách hodnotil učitel každé dítě na škále od 1 do 4 bodů. Metodami faktorové analýzy se prokázalo, že v pozadí těchto 15 položek stojí 5 faktorů, které odpovídají školním a pohybovým dovednostem dítěte, jeho sociální akceptaci, fyzickému vzhledu a chování.

## Popis datového souboru

Datový soubor obsahuje údaje o 60 dětech. Průměrný věk dětí je 8 let a 9 měsíců, směrodatná odchylka je 9 měsíců, minimum je 7 let a 1 měsíc, maximum 10 let. V souboru je 28 chlapců, 32 dívek.

Sledované proměnné jsou:

- U1 ... schopnost zvládnout školní práci
- U2 ... schopnost snadno se skamarádit
- U3 ... schopnost dobrého zvládnání sportů
- U4 ... pěkný vzhled
- U5 ... schopnost dobře se chovat
- U6 ... kvalita paměti
- U7 ... množství přátel
- U8 ... schopnost hrát dobře sportovní hry
- U9 ... hezký fyzický vzhled
- U10 ... vhodnost chování
- U11 ... schopnost odpovídat na otázky ve škole
- U12 ... oblíbenost mezi vrstevníky
- U13 ... šikovnost při hraní nových her
- U14 ... nepěkné vizuální působení
- U15 ... potíže kvůli provozování nevhodných činností

## Výsledky faktorové analýzy a jejich psychologická interpretace

Při výpočtech byly použity statistické programové systémy SPSS a STATISTICA. Prvotní informace o korelační struktuře datového souboru získáme z realizace výběrové korelační matice uvažovaných patnácti proměnných.

Prom.	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	U11	U12	U13	U14	U15
U1	1,00	0,13	0,15	0,48	0,47	0,77	0,08	0,08	0,39	0,43	0,59	0,20	0,25	0,37	0,30
U2	0,13	1,00	0,31	0,18	-0,03	0,32	0,67	0,21	0,04	0,02	0,29	0,31	0,34	0,14	0,06
U3	0,15	0,31	1,00	0,15	-0,00	0,12	0,44	0,85	0,15	0,09	0,05	0,36	0,49	-0,11	0,09
U4	0,48	0,18	0,15	1,00	0,49	0,47	0,24	0,10	0,81	0,47	0,39	0,38	0,22	0,55	0,38
U5	0,47	-0,03	-0,00	0,49	1,00	0,26	0,18	-0,02	0,49	0,88	0,33	0,47	0,06	0,28	0,78
U6	0,77	0,32	0,12	0,47	0,26	1,00	0,08	0,13	0,33	0,23	0,66	0,16	0,31	0,43	0,15
U7	0,08	0,67	0,44	0,24	0,18	0,08	1,00	0,35	0,14	0,22	0,20	0,64	0,35	-0,01	0,18
U8	0,08	0,21	0,85	0,10	-0,02	0,13	0,35	1,00	0,09	0,01	0,01	0,29	0,43	-0,23	-0,04
U9	0,39	0,04	0,15	0,81	0,49	0,33	0,14	0,09	1,00	0,48	0,29	0,36	0,20	0,42	0,47
U10	0,43	0,02	0,09	0,47	0,88	0,23	0,22	0,01	0,48	1,00	0,33	0,53	0,13	0,22	0,78
U11	0,59	0,29	0,05	0,39	0,33	0,66	0,20	0,01	0,29	0,33	1,00	0,22	0,32	0,27	0,25
U12	0,20	0,31	0,36	0,38	0,47	0,16	0,64	0,29	0,36	0,53	0,22	1,00	0,31	0,18	0,46
U13	0,25	0,34	0,49	0,22	0,06	0,31	0,35	0,43	0,20	0,13	0,32	0,31	1,00	0,12	0,15
U14	0,37	0,14	-0,11	0,55	0,28	0,43	-0,01	-0,23	0,42	0,22	0,27	0,18	0,12	1,00	0,31
U15	0,30	0,06	0,09	0,38	0,78	0,15	0,18	-0,04	0,47	0,78	0,25	0,46	0,15	0,31	1,00

Vidíme, že korelace kolísají od -0,23 (korelace mezi U8 a U14) až po 0,88 (korelace mezi U5 a U10). Vysoké vzájemné korelace mají tyto skupiny proměnných: U1, U6, U11; U2, U7; U3, U8; U4, U9, U14; U5, U10, U15. Proměnná U12 sice vysoce koreluje s U7 a U10, avšak korelace mezi U7 a U10 je velmi slabá. Proměnná U13 má nízké korelace se všemi ostatními proměnnými. Zdá se tedy možné, že v pozadí těchto 15 proměnných stojí 5 společných

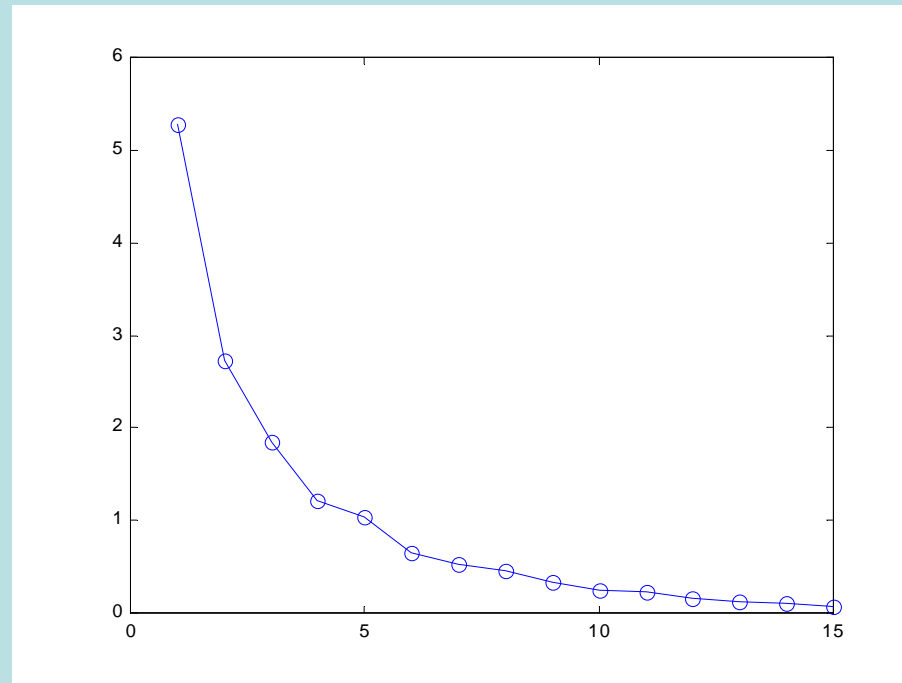
Pro posouzení, zda má vůbec smysl provádět faktorovou analýzu, použijeme Kaiserovu – Meierovu – Olkinovu statistiku. Její hodnota je 0,73, tedy provedení faktorové analýzy se jeví jako středně užitečné.

Testová statistika Bartlettova testu sféricity nabývá hodnoty 588,274, počet stupňů volnosti je 105, odpovídající p-hodnota je velmi blízká 0, tedy hypotézu, že realizace výběrové korelační matice 15 uvažovaných proměnných je jednotková matice, zamítáme na asymptotické hladině významnosti 0,05.

Podívejme se nejprve na vlastní čísla realizované výběrové korelační matice **R** a na procento vysvětleného rozptylu:

vl. číslo	5,273	2,719	1,845	1,218	1,046	0,646	0,531	0,449	0,339	0,242	0,227	0,154	0,124	0,109	0,078
% celk. roz.	35,15	18,13	12,30	8,12	6,98	4,31	3,54	2,99	2,26	1,61	1,51	1,03	0,83	0,73	0,52
kumul. %	35,15	53,28	65,58	73,70	80,68	84,98	88,52	91,51	93,77	95,39	96,90	97,93	98,75	99,48	100,00

Pro větší názornost ještě zakreslíme sutinový graf:



Počet  $m$  společných faktorů stanovíme pomocí Kaiserova kritéria : za  $m$  zvolíme počet těch vlastních čísel realizace matice  $\mathbf{R}$ , která jsou větší než 1, v našem případě 5.

Pro extrakci faktorů zvolíme metodu hlavních faktorů. Vlastní čísla výběrové redukované korelační matice  $\widehat{\Lambda}\widehat{\Lambda}'$  a procento vysvětlené variability máme uvedeny v následující tabulce:

vl. číslo	5,028	2,499	1,609	1,004	0,801
% celk. roz.	33,52	16,66	10,73	6,69	5,34
kumul. %	33,52	50,18	60,91	67,60	72,94

První společný faktor tedy vysvětluje 33,52% variability obsažené v patnácti sledovaných proměnných  $U_1, \dots, U_{15}$ , druhý 16,66% atd. Celkové procento vysvětlené variability je 72,94%. Rotaci faktorů provedeme metodou varimax a opět uvedeme vlastní čísla rotované výběrové redukované korelační matice a procento vysvětlené variability:

vl. číslo	2,861	2,303	2,099	1,868	1,810
% celk. roz.	19,08	15,35	13,99	12,45	12,07
kumul. %	19,08	34,43	48,42	60,88	72,94

Odhady komunalit: .

proměnné	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	U11	U12	U13	U14	U15
poč. odhad	0,722	0,636	0,807	0,777	0,841	0,771	0,737	0,795	0,711	0,825	0,523	0,629	0,376	0,522	0,724
po extr. faktorů	0,710	0,600	0,836	0,930	0,892	0,928	0,934	0,911	0,726	0,898	0,539	0,587	0,354	0,423	0,672

Např. odhad komunality proměnné U1 je 0,71, což lze interpretovat tak, že 71 % variability proměnné U1 lze vysvětlit působením pěti společných faktorů. U jednotlivých proměnných procento variability vysvětlené působením společných faktorů značně kolísá – od 35,4 % u proměnné U13 až po 93,4 % u proměnné U7.



## Odhad matice rotovaných faktorových zátěží:

Proměnná	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5
U1	0,276	0,104	0,758	-0,027	0,218
U2	-0,085	0,132	0,239	0,719	0,039
U3	0,041	0,881	0,027	0,237	0,031
U4	0,267	0,102	0,274	0,125	0,871
U5	0,896	-0,033	0,208	-0,007	0,211
U6	0,021	0,092	0,927	0,088	0,230
U7	0,187	0,263	-0,028	0,910	0,040
U8	-0,016	0,947	0,012	0,110	-0,032
U9	0,354	0,129	0,152	0,009	0,749
U10	0,913	0,031	0,178	0,051	0,171
U11	0,183	0,005	0,669	0,197	0,139
U12	0,496	0,260	0,023	0,484	0,194
U13	0,047	0,427	0,271	0,294	0,099
U14	0,124	-0,200	0,328	0,067	0,505
U15	0,781	-0,006	0,094	0,081	0,217
Výkl.roz	2,861	2,303	2,099	1,868	1,810
Prp.celk	0,191	0,153	0,140	0,125	0,121

(Červeně jsou zvýrazněny zátěže větší než 0,4.)

Vidíme, že první faktor má vysoké korelace s proměnnými U5, U10, U15 a značně nižší s proměnnou U12. Druhý faktor vysoce koreluje s proměnnými U3 a U8, slaběji s U13. U třetího faktoru pozorujeme vysoké korelace s U1, U6 a U11. Čtvrtý faktor má vysoké korelace s U2 a U7, slabší s U12 a pátý faktor vysoce koreluje s U4 a U9, slaběji s U14. Po rotaci připadá na první faktor 2,861 celkového rozptylu (ten je roven 15, tj. počtu uvažovaných proměnných), což je 19,1 %. Na druhý faktor připadá 2,303 celkového rozptylu, tj. 15,3 % atd.

## **Interpretace faktorů:**

1. faktor ... oblast chování
2. faktor ... pohybové dovednosti
3. faktor ... školní dovednosti
4. faktor ... sociální akceptace
5. faktor ... fyzický vzhled

Poněkud problematická se jeví proměnná U12, která vyjadřuje oblíbenost dítěte v kolektivu posouzenou učitelem. Tato proměnná by podle psychologických kritérií měla spíše korelovat s faktorem sociální akceptace než s faktorem vztahujícím se k oblasti chování, jak se to projevilo v našem konkrétním datovém souboru. Zřejmě by otázka týkající se položky U12 měla být formulována jinak.

Kvalitu získaného faktorového modelu posoudíme též pomocí odhadnuté korelační matice a reziduální korelační matice.

Odhadnutá korelační matice:

Prom.	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	U11	U12	U13	U14	U15
U1	0,71	0,16	0,12	0,48	0,45	0,77	0,04	0,09	0,39	0,43	0,58	0,21	0,28	0,37	0,33
U2	0,16	0,60	0,29	0,18	-0,03	0,30	0,67	0,21	0,06	0,01	0,29	0,35	0,33	0,11	0,02
U3	0,12	0,29	0,84	0,16	0,02	0,13	0,46	0,86	0,16	0,09	0,08	0,37	0,46	-0,13	0,05
U4	0,48	0,18	0,16	0,93	0,48	0,48	0,22	0,08	0,80	0,45	0,38	0,40	0,25	0,55	0,43
U5	0,45	-0,03	0,02	0,48	0,89	0,26	0,16	-0,05	0,50	0,89	0,33	0,48	0,10	0,29	0,77
U6	0,77	0,30	0,13	0,48	0,26	0,93	0,09	0,10	0,33	0,23	0,67	0,14	0,34	0,41	0,16
U7	0,04	0,67	0,46	0,22	0,16	0,09	0,93	0,35	0,13	0,23	0,20	0,61	0,39	0,04	0,22
U8	0,09	0,21	0,86	0,08	-0,05	0,10	0,35	0,91	0,10	0,02	0,03	0,29	0,44	-0,20	-0,02
U9	0,39	0,06	0,16	0,80	0,50	0,33	0,13	0,10	0,73	0,48	0,27	0,36	0,19	0,45	0,45
U10	0,43	0,01	0,09	0,45	0,89	0,23	0,23	0,02	0,48	0,90	0,32	0,52	0,14	0,26	0,77
U11	0,58	0,29	0,08	0,38	0,33	0,67	0,20	0,03	0,27	0,32	0,54	0,23	0,26	0,33	0,25
U12	0,21	0,35	0,37	0,40	0,48	0,14	0,61	0,29	0,36	0,52	0,23	0,59	0,30	0,15	0,47
U13	0,28	0,33	0,46	0,25	0,10	0,34	0,39	0,44	0,19	0,14	0,26	0,30	0,35	0,08	0,11
U14	0,37	0,11	-0,13	0,55	0,29	0,41	0,04	-0,20	0,45	0,26	0,33	0,15	0,08	0,42	0,24
U15	0,33	0,02	0,05	0,43	0,77	0,16	0,22	-0,02	0,45	0,77	0,25	0,47	0,11	0,24	0,67

(Diagonální prvky této matice jsou komunity příslušných proměnných.)

## Odhadnutá reziduální korelační matice:

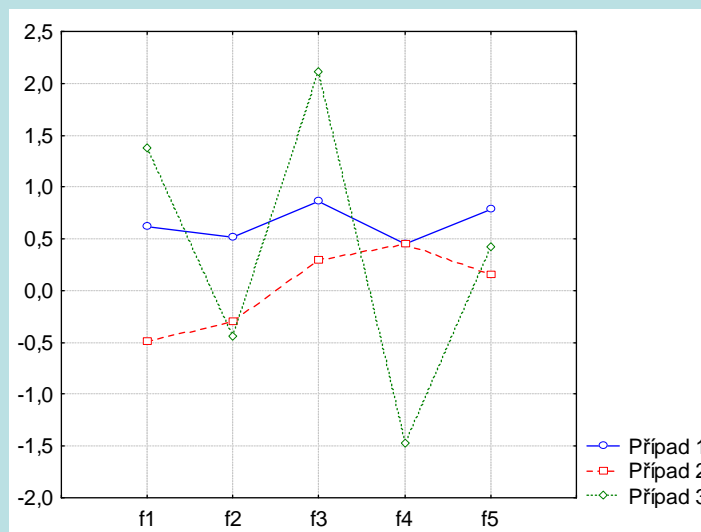
Proměnná	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	U11	U12	U13	U14	U15
U1		-0,04	0,02	0,00	0,02	0,01	0,04	-0,01	0,00	0,00	0,01	-0,01	-0,03	0,00	-0,03
U2	-0,04		0,02	-0,00	-0,00	0,02	0,01	-0,00	-0,02	0,00	-0,00	-0,04	0,01	0,03	0,04
U3	0,02	0,02		-0,01	-0,02	-0,02	-0,02	-0,01	-0,00	-0,00	-0,03	-0,01	0,03	0,02	0,03
U4	0,00	-0,00	-0,01		0,02	-0,01	0,02	0,02	0,01	0,02	0,02	-0,02	-0,03	0,00	-0,05
U5	0,02	-0,00	-0,02	0,02		0,00	0,02	0,03	-0,01	-0,01	-0,00	-0,01	-0,04	-0,01	0,01
U6	0,01	0,02	-0,02	-0,01	0,00		-0,01	0,03	-0,00	0,00	-0,01	0,01	-0,03	0,02	-0,01
U7	0,04	0,01	-0,02	0,02	0,02	-0,01		0,00	0,01	-0,01	-0,00	0,03	-0,03	-0,05	-0,05
U8	-0,01	-0,00	-0,01	0,02	0,03	0,03	0,00		-0,01	-0,01	-0,01	0,01	-0,01	-0,03	-0,02
U9	0,00	-0,02	-0,00	0,01	-0,01	-0,00	0,01	-0,01		-0,01	0,01	-0,00	0,01	-0,02	0,01
U10	0,00	0,00	-0,00	0,02	-0,01	0,00	-0,01	-0,01	-0,01		0,01	0,01	-0,00	-0,03	0,01
U11	0,01	-0,00	-0,03	0,02	-0,00	-0,01	-0,00	-0,01	0,01	0,01		-0,01	0,06	-0,05	-0,01
U12	-0,01	-0,04	-0,01	-0,02	-0,01	0,01	0,03	0,01	-0,00	0,01	-0,01		0,01	0,03	-0,01
U13	-0,03	0,01	0,03	-0,03	-0,04	-0,03	-0,03	-0,01	0,01	-0,00	0,06	0,01		0,04	0,05
U14	0,00	0,03	0,02	0,00	-0,01	0,02	-0,05	-0,03	-0,02	-0,03	-0,05	0,03	0,04		0,07
U15	-0,03	0,04	0,03	-0,05	0,01	-0,01	-0,05	-0,02	0,01	0,01	-0,01	-0,01	0,05	0,07	

Rezidua jsou vcelku malá, kolísají od  $-0,05$  po  $0,07$ . Pouze dvě z nich (tj. 1 %) jsou v absolutní hodnotě větší než  $0,05$ .

Pro několik prvních dětí, které jsou zařazené v datovém souboru, uvedeme ještě odhad faktorového skóre získaný regresní metodou:

	f1	f2	f3	f4	f5
objekt 1	0,6171	0,5131	0,8618	0,4504	0,7838
objekt 2	-0,4883	-0,2984	0,2990	0,4514	0,1564
objekt 3	1,3841	-0,4361	2,1124	-1,4735	0,4277
...	...	...	...	...	...

Spojnicový graf pro první tři děti:



Vidíme, že první dítě má vcelku vyrovnaná skóre u všech pěti faktorů. Druhé dítě se vyznačuje nízkým skóre v oblasti chování, nejvyššího skóre dosahuje v sociální akceptaci. Třetí dítě má nízké skóre u faktoru 4, který reprezentuje sociální akceptaci dítěte jeho vrstevníky a rovněž u faktoru 2 – pohybové dovednosti. Naopak dosahuje vysokého skóre u faktoru 3 – školní dovednosti.

### **Zhodnocení nalezených společných faktorů:**

Metoda hlavních faktorů našla pět společných faktorů stojících v pozadí patnácti sledovaných proměnných. Faktory vztahující se k pohybovým dovednostem, školním dovednostem a fyzickému vzhledu jsou velmi dobře určeny lineárními kombinacemi vždy tří skupin proměnných, zatímco faktory vztahující se k oblasti chování a sociální akceptace nejsou tak jednoznačně určeny kvůli proměnné U12, která vyjadřuje oblíbenost dítěte v kolektivu posouzenou učitelem. Jako vhodné řešení tohoto problému se jeví jiná formulace této položky.

## Závěr

Posoudíme kvalitu faktorového modelu z několika různých hledisek:

- KMO statistika = 0,733 (provedení faktorové analýzy se jeví jako středně užitečné).
- Testová statistika Bartlettova testu sféricity = 588,274, počet stupňů volnosti = 105,  $p$ -hodnota velmi blízká 0 (korelační matice sledovaných 15 proměnných je s rizikem omylu nejvýše 5 % různá od jednotkové matice).
- Odhad modelu založený na metodě hlavních faktorů vysvětluje 72,94 % variability obsažené v datovém souboru.
- Komunality (tj. podíly variability jednotlivých proměnných vysvětlené působením pěti společných faktorů) kolísají od 35,4 % u proměnné U13 až po 93,4 % u proměnné U7.
- Reziduální korelační matice má velmi malé prvky, pouze dva z nich jsou v absolutní hodnotě větší než 0,05.

## Literatura

- [1] J. Anděl. *Matematická statistika*. SNTL/ALFA Praha, 1978.
- [2] S. Harter. *Self-Perception Profile for Children*. Denver. University of Denver, 1985.
- [3] P. Hebák, J. Hustopecký. *Vícerozměrné statistické metody s aplikacemi*. SNTL/Alfa, Praha 1987.
- [4] R. A. Johnson, D. W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall International, Inc. 1992.
- [5] R.P. McDonald. *Faktorová analýza a příbuzné metody v psychologii*. Academia Praha, 1991.
- [6] SPSS/PC+ Statistics™ 4.0. SPSS Inc. 1990.
- [7] STATISTICA for Windows. StatSoft, Inc. 2000.
- [8] K. Überla. *Faktorová analýza*. ALFA Bratislava, 1974.