KONFIRMATORNA FAKTORSKA ANALIZA KORIŠTENJEM MICROSOFT EXCEL-A

CONFIRMATIVE FACTOR ANALYSIS BY USING MICROSOFT EXCEL

Nezir Huseinspahić*
Džemal Kulašin*

SAŽETAK

U svrhu što potpunijeg sagledavanja korisnosti konfirmatorne faktorske analize, u ovom radu poslužili smo se alatom Microsoft Excel. Odabran je primjer faktora koji se odnose na empatiju korisnika telekomunikacijskih usluga označenog sa F4 određen pitanjima iz anketnog upitnika da osoblje BH Telekoma: poklanja individualnu pažnju korisnicima usluga (P14), osoblje udovoljavaju ličnim potrebama i zahtijevima korisnika (P15), osoblje ima strpljenja za sva pitanja korisnika (P16) i osoblje BH Telekoma osjeća najbolje korisničke interese (P17).

Ključne riječi: konfirmatorna faktorska analiza, AMOS, Miscosoft Excel

SUMMARY

For a completenes view of the utility of confirmatory factor analysis, we have used the Microsoft Excel tool in this paper. An example of the factors pertaining to the empathy of the users of telecommunications services marked with F4 identified in the questionnaire that staff of BH Telekom: gives individual attention to service users (P14), staffing to meet personal needs and user requirements (P15), staff has patience all customer questions (P16) and staff of BH Telekom feel the best user interests (P17).

Keywords: confirmatory factor analysis, AMOS, Microsoft Excel

UVOD

Konfirmatorna faktorska analiza služi za provjeru unaprijed formulisanog modela, hipoteze ili teorije. Na osnovu izvora varijanse i kovarijanse ostvarujemo razumijevanje kako naša struktura (model, hipoteza, teorija) djeluje u primjeni, tj. kako se ponaša pri empirijskoj provjeri. Osnovna je prednost CFA nad klasičnim eksploratornim faktorskim analizama je u tome što vrši provjeravanje postavljenog teorijskog modela.[1] Drugim riječima, prvo se utvrđuju pojedinačni odnosi unutar modela, obično postupcima strukturalnog modeliranja linearnim jednačinama. Zatim se procjenjuju mjere odstupanja teorijskog modela od empirijski dobivenih podataka. Ukoliko mjere ukazuju kako model značajno ne odgovara podacima, model je potrebno odbaciti ili pronaći podobniji. Iako CFA nudi mnogo prednosti, ona nije bez ograničenja. Osnovno je ograničenje CFA se ogleda u činjenici kako više različitih teoretskih modela može odgovarati istim podacima.

KONFIRMATORNA FAKTORSKA ANALIZA

Pretpostavimo da se varijansno - kovarijansna matrica S nalazi u polju ćelija B2:E5, kao na Slici 1.

1 Miles, J.N., «Confirmatory factor analysis using Microsoft Excel», Behavior Research Methods, 2005, 37 (4), 672.-676.

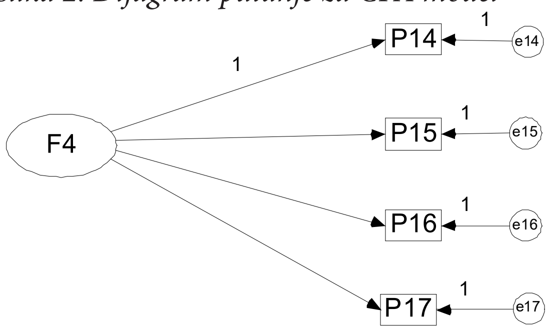
* - Fakultet za menadžment i poslovnu ekonomiju Univerziteta u Travniku

Slika 1. Sadržaj S matrice za pitanjima P14, P15, P16 i P17

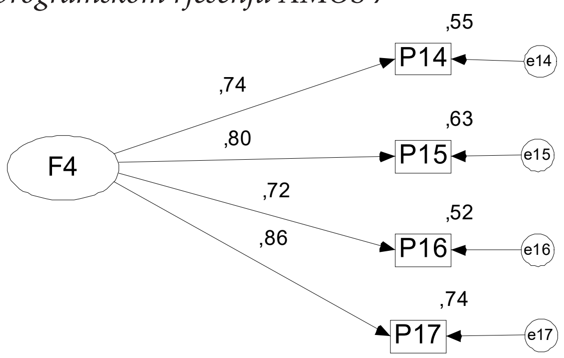
	A	B	C	D	E
1		S			
2		0,7423	0,4598	0,4808	0,5450
3		0,4598	0,8563	0,5417	0,6340
4		0,4808	0,5417	1,0290	0,6205
5		0,5450	0,6340	0,6205	0,9832
6					

Na slici 2. je specificiran model u obliku dijagrama putanje za pitanja P14, P15, P16 i P17. Podaci odgovaraju odgovorima 224 korisnika usluga fiksne telekomunikacione mreže BH Telecoma d.d.

Slika 2. Dijagram putanje za CFA model



Slika 3. Izračunati parametri modela u programskom rješenju AMOS 7



U ovom modelu imamo jedan latentni faktor prvog reda F4. Želimo da pokažemo kako se dolazi do ovih rezultata. Ciljna matrica će biti izračunata korištenjem matričnog računa, a proračunska tablica Excel raspolaže ugrađenim funkcijama koje mogu da obavljaju ovakve matrične operacije. Ako je matrica faktorskih opterećenja (Λ), matrica grešaka (Θ), a matrica varijansi i kovarijansi latentnih varijabli (F), ciljna kovarijansna matrica (Σ) je data u obliku sljedeće jednačine:

$$\Sigma = \Lambda F \Lambda^T + \Theta \tag{1}$$

Kako u ovom slučaju imamo jedan faktor (F4), radi identifikacije modela varijansu faktora F4 fiksirat ćemo na vrijednost 1. Gornja formula sada se može predstaviti kao:

$$\Sigma = \Lambda \Lambda^T + \Theta \tag{2}$$

Ocjene za tražene elemente matrica lambda i theta pronalazimo tako da minimiziramo funkciju dicrepancy između matrice Σ i matrice kovarijansi na našim podacima S. Načešće korišteni metod koji se koristi za procjenu ove razlike jeste metod maksimalne vjerodostojnosti (ML) čija discrepancy funkcija odgovara jednačini:

$$D_{ML} = \ln|\Sigma| + tr\left(S\Sigma^{-1}\right) - \ln|S| - p \tag{3}$$

gdje je p broj indikator varijabli u modelu (u ovom slučaju p = 4), $\ln|\Sigma|$ odgovara prirodnom logaritmu determinante ciljne (implied) matrice Σ, $\ln|S|$ odgovara prirodnom logaritmu determinante matrice uzorka S, $tr\left(S\Sigma^{-1}\right)$ odgovara sumi dijagonalnih elemenata matrice koja se dobija množenjem matrice S i inverzne matrice od Σ.

Kada su matrice Σ i S jednake onda je i funkcija D_{ML} jednaka nuli što znači da nema razlike između ovih matrica. Postupak izračunavanja vrijednosti parammetar je sljedeći:

Prvo ćemo odrediti polje ćelija u kome će se nalaziti faktorska opterećenja, elementi Λ matrice. Inicijalne vrijednosti smo postavili sve na 0,7. Položaj matrice Λ prikazan je na slici 4.

Slika 4. Položaj Λ matrice i transponovane (Λ^T) matrice

	A	B	C	D	E	F	G
8							
9		**Lambda**					
10		0,7000		translambda		**ime: tranlam**	
11		0,7000		0,7000	0,7000	0,7000	0,7000
12		0,7000		=INDEX(lambda; 1)			
13		0,7000					

Radi lakšeg računanja svim ovim matricama možemo dodijeliti i adekvatno ime, kako se i vidi na odgovarajućim slikama. Sljedeća matrica koju treba kreirati je matrica uniknih varijansi i kovarijansi Θ. Položaj ove matrice prikazan je na slici 5.

Slika 5. Matrica uniknih varijansi i kovarijansi Θ.

	A	B	C	D	E
16					
17		Theta		**ime: theta**	
18		0,300	0,000	0,000	0,000
19		0,000	0,300	0,000	0,000
20		0,000	0,000	0,300	0,000
21		0,000	0,000	0,000	0,300

I na kraju potrebno je kreirati i matricu $\Sigma = \Lambda\Lambda^T + \Theta$, sa inicijalnim vrijednostima kao na slici 6.

Slika 6. Inicijalna ciljna (implied) kovarijansna matrica $\Sigma = \Lambda\Lambda^T + \Theta$

	A	B	C	D	E	F
25	**ime: sigma**		1	2	3	4
26		1	0,79	0,49	0,49	0,49
27		2	0,49	0,79	0,49	0,49
28		3	0,49	0,49	0,79	0,49
29		4	0,49	0,49	0,49	0,79

Formula u ćelijima:
C26=INDEX(MMULT(lambda;tranlam);$B26;C$25)+INDEX(theta;$B26;C$25)

D26 =INDEX(MMULT(lambda;tranlam);$B26;D$25)+INDEX(theta;$B26;D$25)

E26=INDEX(MMULT(lambda;tranlam);$B26;E$25)+INDEX(theta;$B26;E$25)

F26=INDEX(MMULT(lambda;tranlam);$B26;F$25)+INDEX(theta;$B26;F$25)

C27=INDEX(MMULT(lambda;tranlam);$B27;C$25)+INDEX(theta;$B27;C$25)

D27=INDEX(MMULT(lambda;tranlam);$B27;D$25)+INDEX(theta;$B27;D$25)

E27=INDEX(MMULT(lambda;tranlam);$B27;E$25)+INDEX(theta;$B27;E$25)

F27=INDEX(MMULT(lambda;tranlam);$B27;F$25)+INDEX(theta;$B27;F$25)

C28=INDEX(MMULT(lambda;tranlam);$B28;C$25)+INDEX(theta;$B28;C$25)

D28=INDEX(MMULT(lambda;tranlam);$B28;D$25)+INDEX(theta;$B28;D$25)

E28=INDEX(MMULT(lambda;tranlam);$B28;E$25)+INDEX(theta;$B28;E$25)

F28=INDEX(MMULT(lambda;tranlam);$B28;F$25)+INDEX(theta;$B28;F$25)

C29=INDEX(MMULT(lambda;tranlam);$B29;C$25)+INDEX(theta;$B29;C$25)

D29=INDEX(MMULT(lambda;tranlam);$B29;D$25)+INDEX(theta;$B29;D$25)

E29=INDEX(MMULT(lambda;tranlam);$B29;E$25)+INDEX(theta;$B29;E$25)

F29=INDEX(MMULT(lambda;tranlam);$B29;F$25)+INDEX(theta;$B29;F$25)

Funkcija MMULT je korištena da multiplicira dvije matrice, u našem primjeru, matrice lambda i transponovane matrice lambda ($\Lambda\Lambda^T$). Ćelije $B29; F$25, u nastavku formule u ćeliji F29, su korištene da referiraju na elemente matrice. Ovo su brojevi sa krajeva sigma matrice (brojevi 1, 2, 3 i 4 sa lijeve strane matrice i brojevi 1, 2, 3 i 4 iznad matrice).

Sve ove formule, koje smo koristili za izračunavanje matrice sigma (Σ), možemo zamijeniti jednom formulom polja:

INDEX(MMULT(lambda;tranlam);B26:B29;C25:F25)+INDEX(theta;B26:B29;C25:F25).

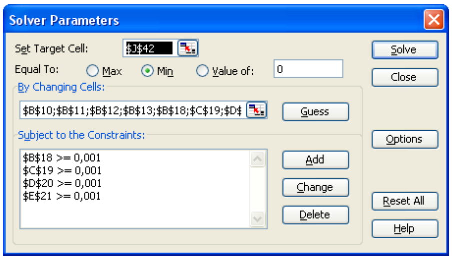
Selektovanjem polja ćelija, u koje smještamo matricu sigma, potrebno je još istovremeno pritisnuti tri tipke tastature: Ctrl+Shift+Enter. Poređenjem matrica S sa Σ posredstvom jednačine (3) izračunavamo funkciju discrepancy. Jednačina je relativno jednostavna, mada za izračunavanje sume dijagonalnih elemenata matrice $tr\left(S\Sigma^{-1}\right)$ treba referirati veći broj ćelija. Inverznu matricu, Σ^{-1}, možemo jednostavno izračunati posredstvom ugrađene funkcije MINVERSE( ). Multiplikacijom matrica S i Σ^{-1} dobijamo matricu čija suma elemenata po dijagonali odgovara izrazu: $tr\left(S\Sigma^{-1}\right)$. Formule koje odgovaraju sadržaju pojedinih ćelija za $tr\left(S\Sigma^{-1}\right)$ su:

=INDEX(MMULT(S; MINVERSE(sigma)); 1; 1)
=INDEX(MMULT(S; MINVERSE(sigma)); 2; 2)
=INDEX(MMULT(S; MINVERSE(sigma)); 3; 3)
=INDEX(MMULT(S; MINVERSE(sigma)); 4; 4)

Formula za izračunavanje funkcije discrepancy je:
= LN (MDETERM (sigma)) + SUM (trace) –LN (MDETERM (S)) - 4

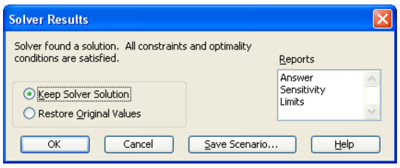
Sada ćemo koristiti Solver (Solver je add-in Excelov alat za optimizaciju i rješavanje jednačina) da pronađemo vrijednosti za faktorska opterećenja, odnosno elemente matrice Λ, i elemente matrice Θ, tako da minimiziramo vrijednost za funkciju D_{ML}. Sad ćemo odabrati jednu ćeliju u koju ćemo upisati gornju formulu za funkciju discrepancy i pokrenuti Solver, kao na slici 7.

Slika 7. Upisivanje ograničenja i referiranje na ćelije koje će se mijenjati kod traženja minimalne vrijednosti za funkciju D_{ML}



Odabrali smo za mijenjanje ćelije koje odgovaraju elementima matrica lambda i theta. Pritiskom na dugme Solve, Solver nas obavještava (Slika 8.) da je pronašao rješenje uz zadovoljenje svih ograničenja i optimalnih uslova.

Slika 8. Obavijest programa za optimizaciju rješenja



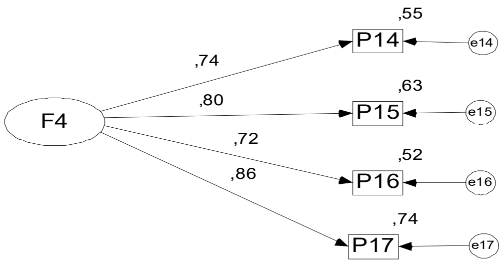
Rezultati izračuna su sljedeći:

	B	C	D	E	F	G	H	I
36	Funkcija discrepancy:		0,002541769	=LN(MDETERM(sigma))+SUM(trace)-LN(MDETERM(S))-4				
37		chi	0,57	=D36*(224-1)				
38		df	2	=4*5/2-4-4				
39		p	0,75	=CHIDIST(chisquared; 2)				

Vidimo da je D_{ML} =0,002541769, χ^2 =0,57, broj stepeni slobode, df = 2, i odgovarajuća vjerovatnoća, p = 0,75. Broj stepeni slobode df se izračunava kao razlika broja varijansi i kovarijansi (10) i broja parametara koji treba ocijeniti (8), četiri faktorska opterećenja i četiri greške. Nestandardizirane i standardizirane vrijednosti parametara za Λ, date su na Slici 9.

Slika 9. Vrijednosti faktorskih opterećenja

	B	C
9	**Lambda**	Standardizirana rješenja
10	0,6368	0,74
11	0,7361	0,80
12	0,7348	0,72
13	0,8551	0,86



Standardizirana rješenja se dobijaju tako što se elementi lambda matrice podijele sa standardnim devijacijama odgovora P14, P15, P16 i P17. Nestandardizirane vrijednosti za Θ su prikazane na slici 10.

Slika 10. Vrijednosti parametara theta matrice

	A	B	C	D	E
16					
17		Theta		**ime: theta**	
18		0,337	0,000	0,000	0,000
19		0,000	0,314	0,000	0,000
20		0,000	0,000	0,489	0,000
21		0,000	0,000	0,000	0,252

Na narednoj slici 11. prikazane su izračunate vrijednosti elemenata ciljne matrice Σ.

Slika 11. Ciljna matrica Σ

	B	C	D	E	F
25	**sigma**	1	2	3	4
26	1	0,7423	0,4688	0,4679	0,5445
27	2	0,4688	0,8563	0,5408	0,6294
28	3	0,4679	0,5408	1,0290	0,6283
29	4	0,5445	0,6294	0,6283	0,9832

ZAKLJUČAK

Konfirmatorna faktorska analiza u naučnim istraživanjima ima veliku ulogu, jer služi za provjeru unaprijed formulisanog modela, hipoteze ili teorije. No, kako se može zaključiti iz predstavljenog slučaja u ovom radu, jasno je da su softverska rješenja poput AMOS-a, koja se bave SEM analizom, mnogo prikladnija i jednostavnije za izračunavanja ocjena parametara od istih izračunavanja u proračunskoj tablici popularnog PC alata Microsoft Excel.

LITERATURA

[1] Miles, J.N., (2005), „Confirmatory factor analysis using Microsoft Excel“, Behavior Research Methods, 37 (4), 672.-676.

[2] Knusel, L., (1998), „On the accuracy of statistical distriutions in MS Exsel 97. Conceptational Statistic & Data Analiysis“, 26, 375.-377.