

MODEL POMIAROWY SATYSFAKcji I LOJALNOŚCI

Adam Sagan

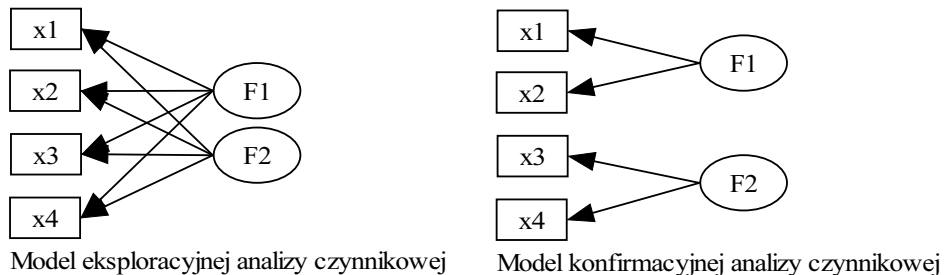
Akademia Ekonomiczna w Krakowie, Katedra Analizy Rynku i Badań Marketingowych

Wstęp

Zaletą stosowania *konfirmacyjnej analizy czynnikowej (CFA)* w porównaniu z omówioną poprzednio eksploracyjną *analizą czynnikową (EFA)* jest możliwość sprawdzenia dopasowania hipotetycznego modelu czynnikowego do macierzy kowariancji zmiennych obserwowalnych i estymacji parametrów modelu czynnikowego. Proces dopasowania modelu do danych empirycznych można ująć w następujący sposób:

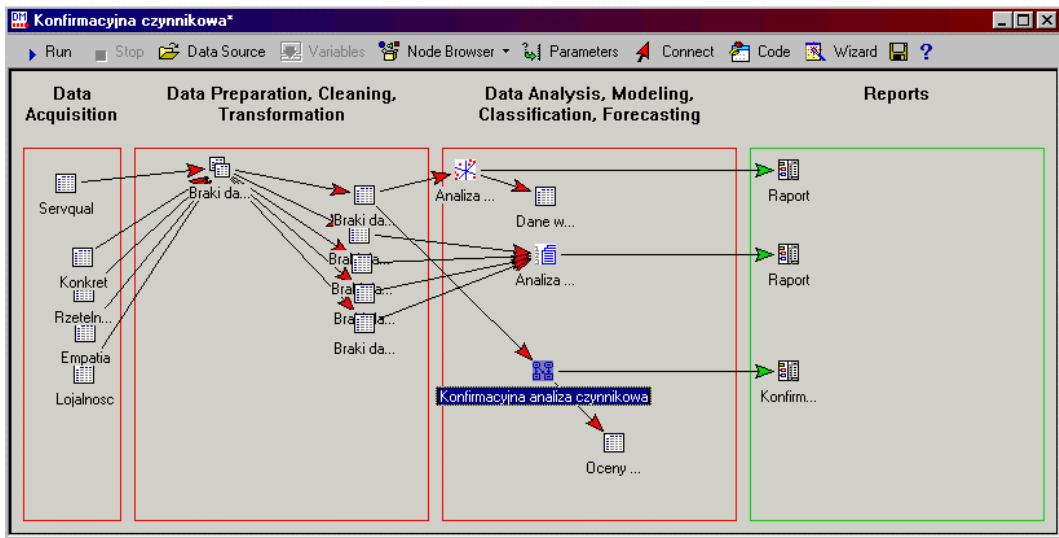
$$\text{Dane} = \text{Model} + \text{Błąd}$$

CFA umożliwia również porównywanie różnych konkurencyjnych modeli między sobą i obliczenie różnych wskaźników dopasowania do danych. Najważniejsza różnica między EFA i CFA sprowadza się do tego, że w modelu EFA wszystkie czynniki główne są skorelowane z wszystkimi zmiennymi obserwowalnymi, a w CFA model teoretyczny decyduje, które wybrane zmienne obserwowalne korelują z określonymi czynnikami.



Rys. 1. Eksploracyjna i konfirmacyjna analiza czynnikowa

Należy jednak silnie podkreślić, że konfirmacyjna analiza czynnikowa powinna być oparta na pewnej spójnej teorii lub co najmniej wcześniejsze wnioski z przeprowadzonej tradycyjnej, eksploracyjnej analizy czynnikowej. W programie *STATISTICA* konfirmacyjna analiza czynnikowa znajduje się w module *Modelowanie równań strukturalnych (SEPATH)*.



W wyniku przeprowadzonej eksploracyjnej analizy czynnikowej uzupełnionej o ocenę rzetelności skal wyodrębnione zostały podstawowe wymiary satysfakcji i lojalności konsumenta. W celu potwierdzenia uzyskanych rezultatów i oceny stopnia rzetelności pomiaru mierzonych konstruktów została zastosowana confirmacyjna analiza czynnikowa. Proces analizy składa się z następujących etapów: a/ specyfikacja modelu pomiarowego, b/ ocena identyfikacji modelu, c/ wybór metody estymacji modelu, d/ ocena wskaźników dopasowania modelu, e/ ocena parametrów modelu.

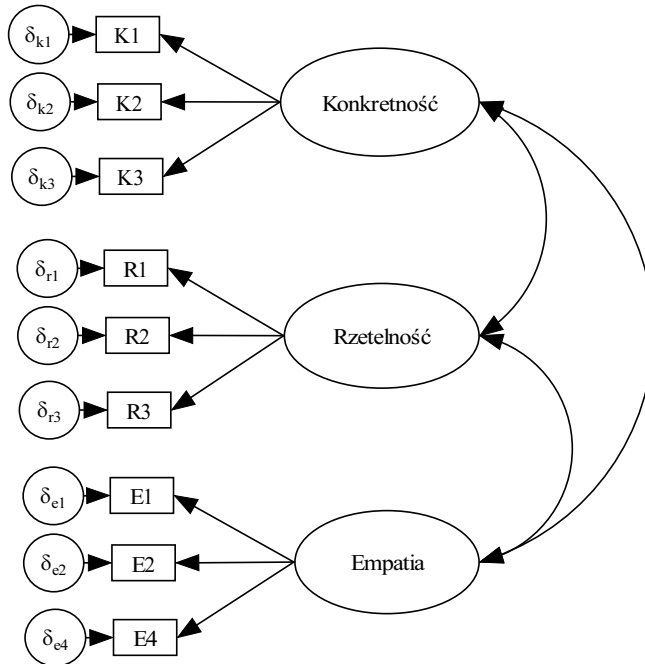
Specyfikacja modelu pomiarowego

Konfirmacyjna analiza czynnikowa ma swoje zastosowanie w ocenie rzetelności skal pomiarowych. W naszym przykładzie oceny rzetelności skali satysfakcji z produktu i lojalności należy określić relacje między pozycjami skali a czynnikami ukrytymi oraz korelacje między czynnikami. Graficzny obraz poszczególnych czynników jest przedstawiony na rys. 2 (poniżej). Zazwyczaj zmienne obserwowalne (pozycje skali) są reprezentowane za pomocą prostokątów, a zmienne ukryte (wymiary) za pomocą owali. Relacje regresyjne (przyczynowe) przedstawia się za pomocą strzałek jednokierunkowych, a zależności korelacyjne lub kowariancyjne – strzałek dwukierunkowych.

Identyfikacja modelu

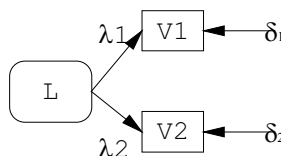
Jednym z podstawowych problemów jest kwestia identyfikacji modelu pomiarowego. Pierwszym zagadnieniem jest określenie jednostek miar dla zmiennych obserwowalnych i ukrytych. Jeżeli dane wejściowe stanowią macierz wariancji-kowariancji, wówczas interpretacji parametrów dokonuje się w jednostkach miar zmiennych wejściowych (są to

zazwyczaj punkty na skali Likerta). Jeżeli natomiast danymi wejściowymi jest macierz korelacji Pearsona, to interpretacja parametrów jest dokonywana na podstawie odchyłeń standardowych. Ważnym zagadnieniem jest ustalenie jednostki miary dla zmiennych ukrytych. Jednostki miary takich zmiennych mają charakter umowy. Wykorzystuje się w tym celu dwa podejścia: w pierwszym ustala się ładunek czynnikowy pierwszej zmiennej obserwowalnej jako 1 (jednostki miary zmiennej ukrytej są jednostkami tej zmiennej), w drugim podejściu ustala się wariancję zmiennej ukrytej na poziomie 1. Zabieg ten w różnych programach modelowania strukturalnego jest najczęściej dokonywany automatycznie. W programie *STATISTICA* stosowana jest druga strategia identyfikacji zmiennych ukrytych (wariancja zmiennych ukrytych wynosi 1).



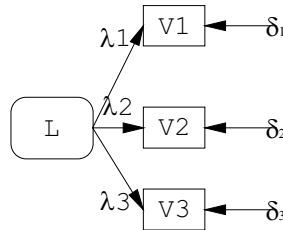
Rys. 2. Model pomiarowy satysfakcji z produktu

Drugie zagadnienie związane z identyfikacją modelu dotyczy warunków koniecznych dla rozwiązania matematycznego modelu konfirmacyjnej analizy czynnikowej. Można w tym względzie mówić o tzw. niedoidentyfikacji, identyfikacji i nadidentyfikacji modelu pomiarowego [Bollen 1989].



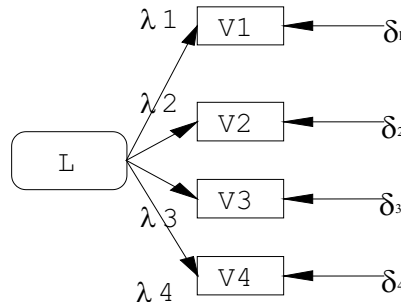
Rys. 3. Niedoidentyfikacja modelu pomiarowego

Niedoidentyfikacja modelu występuje wtedy, gdy liczba parametrów do oszacowania jest większa niż wiadomych danych wejściowych. W sytuacji z rys. 3 liczba wolnych parametrów do oszacowania wynosi 4 (2 ładunki czynnikowe λ oraz 2 wariancje błędu δ), natomiast wejściowa macierz danych obejmuje 2 wariancje $V1$ i $V2$ oraz 1 kowariancję $V1V2$.



Rys. 4. Identyfikacja modelu pomiarowego

Identyfikacja modelu występuje, jeżeli liczba parametrów jest równa liczbie danych wejściowych. W przypadku przedstawionym na rys. 4 liczba niewiadomych parametrów wynosi 6, a wejściowa macierz kowariancji obejmuje 3 wariancje ($V1$, $V2$, $V3$) oraz 3 kowariancje między zmiennymi ($V1V2$, $V1V3$, $V2V3$).



Rys. 5. Nadidentyfikacja modelu pomiarowego

Nadidentyfikacja modelu oznacza sytuację, gdy liczba parametrów jest mniejsza od liczby danych wejściowych. Model przedstawiony na rys. 5. zawiera 8 parametrów modelu do oszacowania, macierz wariancji-kowariancji zawiera 4 wariancje oraz 6 kowariancji między zmiennymi. Ogólny warunek konieczny do identyfikacji modelu jest określony wzorem:

$$t \leq \frac{p(p+1)}{2},$$

gdzie: t – liczba wolnych parametrów modelu,

p – rząd macierzy kowariancji.



Na podstawie powyższych założeń często mówi się o tzw. regułach identyfikacji modelu. Pierwsza z nich, zwana „regułą trzech mierników”, mówi, że model jest identyfikowalny, jeżeli każda zmienna ukryta jest związana z co najmniej trzema zmiennymi obserwowalnymi, a zmienne ukryte są ortogonalne (niezależne). Druga reguła „dwóch mierników” głosi, że model jest identyfikowalny, jeżeli każda zmienna ukryta jest związana z co najmniej dwiema zmiennymi obserwowalnymi i jest jednocześnie skorelowana z co najmniej jedną inną zmienną ukrytą. W przypadku modelu pomiarowego satysfakcji i lojalności formalne warunki identyfikacji są spełnione.

Estymacja modelu

Celem procedury estymacyjnej jest uzyskanie, na podstawie wejściowej macierzy kowariancji, szacunków parametrów, które w taki sposób odtwarzają macierz kowariancji zakładaną przez model (Σ), że wartości te są możliwie najbliższe wartościom macierzy wariancji-kowariancji zmiennych obserwowalnych (S). Proces estymacji wykorzystuje odpowiednią funkcję rozbieżności (lub funkcję dopasowania) w celu minimalizacji różnicy między Σ a S . W programie SEPATH wykorzystywane są następujące metody estymacji parametrów modelu:

- ◆ **metoda największej wiarygodności (NW)** wykorzystuje estymację największej wiarygodności Wisharta w przypadku analizy korelacji lub kowariancji, a estymację normalną największej wiarygodności w przypadku analizy momentów. Jest to najczęściej spotykana metoda estymacji parametrów modelu. Wymaga ona jednak spełnienia założenia wielowymiarowej normalności rozkładu,
- ◆ **metoda uogólniona najmniejszych kwadratów (UNK)** - program wykonuje estymację metodą najmniejszych kwadratów; polega ona na minimalizacji ważonych kwadratów odchyłeń śladu empirycznej i teoretycznej macierzy kowariancji. Do jej zalet należy ważenie elementów macierzy kowariancji wg ich wariancji i kowariancji z innymi elementami macierzy i niezależność od skali pomiaru zmiennych obserwowalnych. Wadą jest wymóg dużych prób przekraczających 200 obserwacji,
- ◆ **metoda UNK-NW** - jest to opcja domyślna. SEPATH wykonuje 5 iteracji przy użyciu UNK, a następnie przechodzi do estymacji metodą NW,
- ◆ **metoda zwykłych najmniejszych kwadratów (ZNK)** - program wykonuje estymację metodą ZNK,
- ◆ **metoda asymptotycznie nieparametrycznej macierzy Grama** - program wykonuje estymację asymptotycznie nieparametryczną. Polega ona na minimalizacji ważonej sumy kwadratów odchyłeń empirycznej i teoretycznej macierzy kowariancji. Jej zaletami są minimalne założenia dotyczące rozkładu (w odróżnieniu od metod NW i UNK nie wymaga ona spełnienia założeń o wielowymiarowej normalności rozkładu). Wadą tej metody są wymagania bardzo dużych prób (ponad 2000 obserwacji) i występujące problemy obliczeniowe przy dużych macierzach korelacji.



Ocena wskaźników dopasowania modelu do danych

Dopasowanie modelu do danych w modelowaniu strukturalnym jest rozumiane bardzo specyficznie. Przyjmuje się bowiem, że w zasadzie każdy model jest jedynie uproszczonym obrazem rzeczywistości i zamiast mówić o *goodness of fit* tych modeli należy raczej posługiwać się terminem *badness of fit*. Inaczej mówiąc, każdy model źle pasuje do danych i należy zaakceptować taki model, który pasuje najmniej źle! Z powodu częstej nadidentyfikacji modelu (wówczas liczba stopni swobody jest większa od zera) należy również podkreślić, że dobre dopasowanie modelu świadczy jedynie o tym, że jest to jeden z reprezentantów modeli w miarę dobrze dopasowanych do danych (istnieje bowiem wiele alternatywnych modeli strukturalnych także dobrze dopasowanych). Należy więc pamiętać, że istnienie dobrego dopasowania nie oznacza jeszcze prawdziwości danego modelu, pozwala jedynie na nieodrzućenie go w oparciu o dane empiryczne. Do tych danych pasować może równie dobrze wiele konkurujących ze sobą modeli. Miary dopasowania mają właściwie jedynie moc odrzucania modeli teoretycznych i falsyfikowania istniejącej u ich podstaw teorii. Stąd na podstawie przeglądu wyników analiz z wykorzystaniem modelowania strukturalnego wskaźniki dopasowania mogą być podobne dla wielu konkurencyjnych modeli i dostarczają jedynie subiektywnej weryfikacji prawdziwości modelu. Wskaźniki dopasowania pozwalają jedynie na falsyfikację modelu i nie dają żadnych podstaw do przyjęcia danego modelu jako "prawdziwego".

Najczęściej spotykanym wskaźnikiem dopasowania modelu do danych jest wartość statystyki χ^2 . Test χ^2 dotyczy hipotezy zerowej, głoszącej że reszty standaryzowane macierzy empirycznej i teoretycznej (odtworzonej przez model) wynoszą 0, tzn. że ograniczenia nałożone przez badacza wynikające z przyjętego modelu teoretycznego są trafne. Stosowanie tego testu jest właściwe, jeżeli zmienne mają wielowymiarowy rozkład normalny oraz próba jest liczna. Jednakże przy bardzo dużych próbach ufność, że macierze są równe rośnie, ale istotność owych różnic może być niewielka. Z drugiej strony przy bardzo małych próbach nie można poprawnie zdiagnozować nawet dużych różnic między macierzami. Wrażliwość testu χ^2 na wielkość próby spowodowała, że w literaturze spotyka się liczne alternatywne testy tzw. dobroci dopasowania modelu do danych.

Program SEPATH wyposażony jest w szereg takich testów. Można je podzielić na trzy podstawowe grupy, tj.: 1/ absolutne wskaźniki dopasowania, 2/ porównawcze wskaźniki dopasowania oraz 3/ oszczędne wskaźniki dopasowania. Pierwsze z nich odpowiadają na pytania, jak poprawnie model pomiarowy odtwarza obserwowalną macierz kowariancji S. Drugie określają stopień, w jakim dany model jest lepszy niż tzw. model zerowy (lub podstawowy), w którym zakładamy brak zmiennych ukrytych i występowanie relacji jedynie między zmiennymi obserwowalnymi. Trzeci typ wskaźników dopasowania ukazuje stopnie swobody wymagane dla osiągnięcia odpowiedniego dopasowania modelu do danych. W tab.1. zaprezentowane są najczęściej wykorzystywane wskaźniki dopasowania modelu występujące również w programie *STATISTICA*.



Tabela 1. Wskaźniki dopasowania modelu konfirmacyjnej analizy czynnikowej

Kryterium dobroci dopasowania	Dobre dopasowanie	Intepretacja	Typ dopasowania
Chi-Kwadrat	$p > 0.05$	Istotne wartości χ^2 wskazują, że macierze S i Σ różnią się.	Absolutne
$\chi^2 / \text{s.s.}$)	Wartość < 5		Absolutne
GFI (<i>Goodness of Fit</i>)	> 0.95	GFI mierzy wielkość wariancji-kowariancji macierzy S, która jest przewidywana przez odtworzoną macierz Σ .	Absolutne
Indeks Gamma populacji	> 0.95	Rozwinięcie wskaźnika GFI.	Absolutne
AGFI	> 0.95	Wartości GFI dostosowane do liczby stopni swobody (df) modelu względem liczby zmiennych. Wartości > 0.90 wskazują na dobre dopasowanie.	Absolutne
Skorygowany indeks Gamma populacji	> 0.95	Rozwinięcie wskaźnika AGFI.	Absolutne
RMSEA Steigera-Linda	< 0.05 dobre dopasowanie, < 0.01 znakomite dopasowanie	Wartości mniejsze niż 0.05 wskazują na dobre dopasowanie.	Absolutne
Kryterium informacyjne Akaike	bliskie 0	Porównuje wartości alternatywnych modeli.	Oszczędne
Kryterium bayesowskie Schwarz (S _k)		Porównanie kilku modeli zagnieżdżonych (nested). Wybór tego modelu, którego S _k jest najmniejsze.	Przyrostowe
Indeks oceny krzyżowej Browne'a-Cudecka (C _k)		Porównanie kilku modeli zagnieżdżonych (nested). Wybór tego modelu, którego C _k jest najmniejsze.	Przyrostowe
Unormowany indeks Bentlera-Bonetta	1	Indeks mierzy względny spadek funkcji rozbieżności spowodowany przejściem od modelu zerowego do modelu bardziej złożonego. Nie kompensuje oszczędności modelu.	Przyrostowe
Nieunormowany indeks Bentlera-Bonetta	1	Indeks mierzy względny spadek funkcji rozbieżności spowodowany przejściem od modelu zerowego do modelu bardziej złożonego. Kompensuje oszczędność modelu.	Przyrostowe

Wyniki konfirmacyjnej analizy czynnikowej wymiarów satysfakcji zaprezentowanych na rys. 2. są przedstawione poniżej. Zastosowano w niej metodę łączoną uogólnioną najmniejszych kwadratów i największej wiarygodności.



Wyniki modelowania równań strukturalnych: statistica

Metoda estymacji:	GLS -> ML	Statystyka chi-kwadrat:	53.5226
Funkcja rozbieżności:	0.251	Liczba stopni swobody:	24
Maks. cosinus resztowy:	3.96E-005	Poziom p chi-kwadrat:	0.000493
Maks.abs. gradient:	2.87E-005	Steiger-Lind RMSEA	
Kryterium ICSF:	3.18E-008	--->Estymator punktowy:	0.0731
Kryterium ICS:	1.04E-005	-->Dolna granica 90:	0.0452
Warunki brzegowe:	0	-->Górna granica 90:	0.101
		RMS stand. reszty:	0.0565

Podstawowe Więcej Założenia Reszty

Podsumowanie modelu

Podsumowanie: Podstawowe statystyki

Podsum.

Anuluj

Opcje

Poziom p dla podświetlania: 0.05

Wartość statystyki χ^2 jest istotna na poziomie $p=0.0005$, stąd należy odrzucić hipotezę zerową o równości empirycznej i odtworzonej przez model macierzy kowariancji (reszty standaryzowane nie wynoszą zero, stąd model nie jest dobrze dopasowany do danych). Z powodów wymienionych poprzednio, nie należy poprzestawać na ocenie dopasowania modelu na podstawie jedynie wartości χ^2 . Dodatkowe absolutne wskaźniki dopasowania takie jak $\chi^2/s.s$ ($=53.52/24$) oraz wartość pierwiastka średniokwadratowego błędu aproksymacji (RMSEA = 0.07), wskazują na akceptowalny (w miarę) poziom dopasowania modelu. Spośród dodatkowych wskaźników dopasowania również GFI oraz AGFI wskazują na raczej dobre dopasowanie.

Skoroszyt1* - Wskaźniki dla pojednych prób (statistica)

	Wskaźniki dla pojednych prób
	Wartość
GFI Joreskoga	0.949
AGFI Joreskoga	0.905
Kryt. informacyjne Akaike'a	0.448
Kryt. bayesowskie Schwarza	0.780
In. atestacji krzyż. Browne'a-Cudecka	0.458
Model niezal. chi-kwadrat	643.229
Model niezal. ss	36.000
Unorm. wskaźnik Bentlera-Bonetta	0.917
Nieunorm. wskaźnik Bentlera-Bonetta	0.927
Porówn. wskaźnik dop. Bentlera	0.951
Oszczędny wskaźnik Jamesa-Mulaika-Bretta	0.611
Ro Bollena	0.875
Delta Bollena	0.952

Wskaźniki dla pojednych prób (statistica)



Ocena parametrów modelu

Ocena parametrów modelu pomiarowego pozwala na oszacowanie prawdopodobieństwa, z jakim dany parametr w populacji ma wartość zero. Istotność parametru jest mierzona na podstawie ilorazu krytycznego wartości parametru podzielonego przez błąd standardowy (jest to wartość statystyki T). Ocena ta dotyczy zarówno ładunków czynnikowych, jak i wariacji błędów pomiaru danego konstrukt.

Wyniki analizy zaprezentowane w tabeli wskazują, że parametry modelu są istotne na poziomie $p=0.05$. Istotna jest również kowariancja między wymiarami „empatia” i „konkretność”, która wynosi 0.75. Należy zwrócić uwagę, że jest to model zbudowany na podstawie macierzy kowariancji, stąd ładunki czynnikowe odzwierciedlają współczynniki regresji między zmiennymi obserwowalnymi a czynnikami (współczynniki te mogą być większe od 0), a wariacje każdej zmiennej ukrytej wynoszą 1.

	Oceny modelu (statistica)			
	Ocena parametru	Błąd standard.	Statystyka T	Poziom prawdop.
(Konkret)-1->[K1]	1.245	0.108	11.536	0.000
(Konkret)-2->[K2]	1.020	0.099	10.333	0.000
(Konkret)-3->[K3]	1.487	0.114	13.052	0.000
(Rzetel)-4->[R1]	0.628	0.073	8.560	0.000
(Rzetel)-5->[R2]	0.748	0.075	9.965	0.000
(Rzetel)-6->[R3]	0.444	0.070	6.369	0.000
(Empatia)-7->[E1]	1.209	0.130	9.281	0.000
(Empatia)-8->[E2]	1.029	0.085	12.071	0.000
(Empatia)-9->[E4]	1.056	0.088	11.960	0.000
(DELTA1)->[K1]				
(DELTA2)->[K2]				
(DELTA3)->[K3]				
(DELTA4)->[R1]				
(DELTA5)->[R2]				
(DELTA6)->[R3]				
(DELTA7)->[E1]				
(DELTA8)->[E2]				
(DELTA9)->[E4]				
(DELTA1)-10-(DELTA1)	1.263	0.167	7.561	0.000
(DELTA2)-11-(DELTA2)	1.206	0.143	8.434	0.000
(DELTA3)-12-(DELTA3)	1.093	0.186	5.861	0.000
(DELTA4)-13-(DELTA4)	0.450	0.075	6.020	0.000
(DELTA5)-14-(DELTA5)	0.218	0.088	2.463	0.014
(DELTA6)-15-(DELTA6)	0.670	0.072	9.292	0.000
(DELTA7)-16-(DELTA7)	2.234	0.254	8.785	0.000
(DELTA8)-17-(DELTA8)	0.692	0.105	6.599	0.000
(DELTA9)-18-(DELTA9)	0.757	0.112	6.729	0.000
(Rzetel)-19-(Konkret)	-0.100	0.085	-1.171	0.241
(Empatia)-20-(Konkret)	0.752	0.051	14.875	0.000
(Empatia)-21-(Rzetel)	0.086	0.087	0.995	0.320

Oceny parametrów modelu pozwalają również na oszacowanie rzetelności pomiaru zmiennych ukrytych. Analiza rzetelności jest dokonywana w obrębie tzw. klasycznej teorii testu. Według tej teorii wartość obserwowana empirycznie (x) jest funkcją wartości prawdziwej pomiaru (T) oraz błędu pomiaru (e) [Peter 1979]:

$$x = T + e$$



Wartość prawdziwa pomiaru wynika z modelu confirmacyjnej analizy czynnikowej i dana jest wzorem:

$$T = \Lambda_x \xi + s ,$$

gdzie s oznacza wariancję specyficzną dla danej zmiennej x .

Możemy wyrazić więc, że:

$$x = \Lambda_x \xi + s + e$$

Wariancja błędu pomiaru (e) oraz wariancja specyficzna (s) jest nazywana wariancją swoistą δ . Wówczas otrzymujemy równanie confirmacyjnej analizy czynnikowej:

$$x = \Lambda_x \xi + \delta$$

Wyrażając to w kategoriach wariancji można napisać, że

$$\text{var}(x_1) = \lambda^2_1 \Phi + \text{var}(\delta_1) ,$$

gdzie: λ – ładunek czynnikowy,

Φ – wariancja zmiennej ukrytej,

$\text{var}(\delta)$ – wariancja swoista danej zmiennej.

Rzetelność narzędzia pomiaru jest określana jako wariancja wyniku prawdziwego i wariancji specyficznnej dla danej zmiennej w stosunku do całkowitej wariancji zmiennej (odpowiada na pytanie na ile wynik prawdziwy „tłumaczy” całkowitą wariancję danej zmiennej). W praktyce bardzo trudno jest wyodrębnić wariancję specyficzną dla danej zmiennej, dlatego „konserwatywna” (rozumiana jako dolna granica) wartość współczynnika rzetelności jest dana wzorem:

$$\rho = \frac{\sum (\lambda^2_x \Phi)}{\sum \text{var}(x)}$$

Zmienna	Ładunek czynnikowy	Wariancja zmiennej	Współczynnik rzetelności ρ
K1	1.245	2.813	0.57
K2	1.020	2.236	0.57
K3	1.487	3.304	0.57
R1	0.628	0.844	0.46
R2	0.748	0.776	0.46
R3	0.444	0.867	0.46
E1	1.209	3.695	0.50
E2	1.029	1.750	0.50
E4	1.056	1.873	0.50



Na podstawie danych z powyższej tabeli ładunków czynnikowych i wariancji swoistej (DELTA) można obliczyć współczynniki rzetelności pomiaru poszczególnych zmiennych ukrytych dotyczących satysfakcji z produktu.

Z tabeli wynika, że zmienna ukryta „konkretność” jest mierzona skalą o najwyższej rzetelności, natomiast skala pomiaru „rzetelność personelu” jest skalą o najniższej rzetelności pomiaru.

Konfirmacyjna analiza czynnikowa pozwala na pełną diagnozę modelu pomiarowego danych zmiennych ukrytych. Umożliwia sprawdzenie trafności teoretycznego modelu wysuniętego przez badacza oraz rzetelności wskaźników użytych do pomiaru wyodrębnionych konstruktów.

Literatura

1. Bollen, K., *Structural Equations with Latent Variables*, New York, Wiley 1989.
2. Peter, J.P., *Reliability: A Review of Psychometric Basics and Recent Marketing Practices*, *Journal of Marketing Research*, 1979/Feb.