

Dr Katarzyna Czainska
Wydział Ekonomii i Zarządzania
Poznańska Wyższa Szkoła Biznesu

WYKORZYSTANIE WSPÓLCZYNNIKA FLEISSA DO WERYFIKACJI ZGODNOŚCI WYNIKÓW OTRZYMYWANYCH W METODZIE DELFIKIEJ PRZY PROGNOZOWANIU ZJAWISK ZARZĄDCZYCH I GOSPODARCZYCH

Streszczenie

W artykule zaprezentowano przykłady zastosowania współczynnika zgodności Fleissa do weryfikacji wyników otrzymywanych przy wykorzystaniu metody delfickiej. Przedstawiono przyczyny wzrostu zainteresowania wspomnianą metodą ekspercką, szczególnie w kontekście projektów naukowo – badawczych typu foresight. Podkreślono zasadność stosowania ww. współczynnika w procesie diagnozowania i prognozowania gospodarczego, strategicznego i zarządczego. Ponadto, podano przykłady konkretnych układów tablic do badań eksperckich. Kluczowe słowa: foresight, metoda delficka (ekspercka), współczynnik Fleissa

1. Wprowadzenie

W greckiej miejscowości Delfy leżącej około 180 km od Aten znajdują się ruiny świątyni Apollina. W tym miejscu w czasach starożytnych mieściła się jedna z najważniejszych wyroczni, do której przychodzili z prośbą o poradę ówczesni władcy i członkowie elity międzynarodowej. Kapłanki, zwane pytiami, kierowały pytania do Apolla, który zgodnie z ówczesnymi wierzeniami, pokazywał im przyszłość w formie wizji. Wizje te nie były konkretnymi rozwiązaniami problemów. Pytie przekazywały informacje kapłanom, a ci z kolei zainteresowanym osobom, pozostawiając im możliwość interpretacji usłyszanego proroctwa. Przytoczona historia stała się podstawą nadania jednej z metod badawczych nazwy „metody delfickiej”, gdyż jej istotą jest korzystanie z opinii ekspertów w celu opracowania prognoz, np. gospodarczych, czy strategicznych.

Współcześnie w Polsce, a konkretnie od 2003 roku nastąpił ponowny wzrost zainteresowania tą metodą, głównie za sprawą funduszy Unii Europejskiej i promowanym w nich typie badań *foresight*. Foresight to forma profesjonalnego przygotowania wizji przyszłości w określonym sektorze życia, np. ekonomii, nauce, technice, itp. Opracowywana

prognoza jest zawsze średnio- lub długoterminowa. Metoda delficka jest natomiast jedną z podstawowych metod badawczych zalecanych przy realizacji tego typu projektów¹. Idea foresightu został rozpropagowana za sprawą dwóch inicjatyw:

1. Narodowego Programu Foresight „Polska 2020”² – uruchomionego w grudniu 2006 r. przez Ministra Nauki i Szkolnictwa Wyższego, obejmującego trzy pola badawcze: Zrównoważony Rozwój Polski, Technologie Informacyjne i Telekomunikacyjne oraz Bezpieczeństwo;
2. Sektorowego Programu Operacyjnego „Wzrost Konkurencyjności Przedsiębiorstw” na lata 2004 – 2006³.

Na latach 2007 – 2013 przewidziano kontynuację badań foresight i w ramach programu Innowacyjna Gospodarka⁴ przewidziano kwotę 15 milionów euro na ich finansowanie. Według danych Ośrodka Przetwarzania Informacji⁵ na dzień 3 sierpnia 2012 r., w ramach programu złożono 67 wniosków, z których 22 zatwierdzono do realizacji. Wartość dofinansowania wynikająca z podpisanych umów na realizację wybranych projektów wyniosła 45 869 078,86 zł. Fundusze otrzymały podmioty z następujących województw: wielkopolskiego (2 wnioski, 9,1% środków), pomorskiego (1 wniosek – 4,55% środków), podlaskiego (1 wniosek – 4,55% środków), mazowieckiego (1 wniosek – 4,55% środków), dolnośląskiego (2 wnioski, 9,1% środków), śląskiego (2 wnioski, 9,1% środków) i małopolskiego (2 wnioski, 9,1% środków). Pozostałe 11 wniosków (50% środków) dotyczyły projektów ogólnokrajowych. Przykłady wybranych projektów zaprezentowano w tabeli nr 1.

Tabela 1 Przykładowe projekty foresight realizowane w ramach programu Innowacyjna Gospodarka na lata 2007 - 2013

Tytuł projektu	Podmiot realizujący
Foresight regionalny dla szkół wyższych Warszawy i Mazowsza „Akademickie Mazowsze 2030”	Politechnika Warszawska
Foresight wiodących technologii kształtowania własności materiałów inżynierskich i biomedycznych FORSURF	Politechnika Śląska
Żywność i żywienie XXI wieku – wizja rozwoju polskiego sektora spożywczego	Instytut Sadownictwa i Kwiaciarstwa im. Szczepana Pieniążka
Foresight technologiczny „NT FOR PODLASIE 2010”. Regionalna	Politechnika Białostocka

¹ / zob. K. Borodako, *Foresight w zarządzaniu strategicznym*, Wydawnictwo C.H. Beck, Warszawa 2009

² / <http://foresight.polska2020.pl/cms/> - [odczytano dnia 14.08.2012]

³ / w ramach programu prowadzono również badania foresight, np. Polskie Biuro ds. Przestrzeni Kosmicznej realizowało projekt celowy Foresightu branżowego: "Ocena perspektyw i korzyści z wykorzystania technik satelitarnych i rozwoju technologii kosmicznych w Polsce"; zob. <http://www.kosmos.gov.pl/index.php?link=94> - [odczytano dnia 14.08.2012]

⁴ / Program: Program Innowacyjna Gospodarka; Priorytet: 1 Badania i rozwój nowoczesnych technologii; Działanie: 1.1 Wsparcie badań naukowych dla budowy gospodarki opartej na wiedzy; Poddziałanie: 1.1.1 Projekty badawcze z wykorzystaniem metody foresight

⁵ / <http://www.opi.org.pl/pl/stan-wdrazaniapoig/stan-wdrazania-111/> - [odczytano dnia 14.08.2012]

strategia rozwoju nanotechnologii	
POMORZE 2030 – scenariusze rozwoju i kluczowe technologie	Instytut Badań nad Gospodarką Rynkową
Zeroemisyjna gospodarka energią w warunkach zrównoważonego rozwoju Polski do 2050 roku	Główny Instytut Górnictwa
Zaawansowane technologie przemysłowe i ekologiczne dla zrównoważonego rozwoju kraju	Instytut Technologii Eksploatacji – Państwowy Instytut Badawczy w Radomiu
Foresight w zakresie priorytetowych i innowacyjnych technologii w zakresie zagospodarowywania odpadów pochodzących z górnictwa węgla kamiennego	Instytut Mechanizacji Budownictwa i Górnictwa Skalnego
Perspektywa technologiczna Kraków Małopolska 2020	Krakowski Park Technologiczny
Foresight „Sieci gospodarcze Wielkopolski” – scenariusze transformacji wiedzy wspierające innowacyjną gospodarkę	Politechnika Poznańska
Identyfikacja potencjału i zasobów Dolnego Śląska w obszarze nauka i technologie na rzecz poprawy jakości życia (Quality of Life) oraz wytyczenie przyszłych kierunków rozwoju. Badania metodami foresight	Uniwersytet Ekonomiczny we Wrocławiu
Analiza i prognozowanie ścieżek rozwoju interdyscyplinarnych nauk o poznaniu metodami foresight	Uniwersytet Jagielloński
Strategia rozwoju energetyki na Dolnym Śląsku metodami foresightowymi	Politechnika Wrocławska
Foresight dla energetyki termojądrowej	Instytut Fizyki Plazmy i Laserowej Mikrosyntezy im. S. Kaliskiego

Źródło: *Projekty badawcze z wykorzystaniem metody foresight; wybrane projekty realizowane w ramach Poddziałania 1.1.1 Programu Innowacyjna Gospodarka*, Ośrodek Przetwarzania Informacji, Warszawa 2010 oraz *Lista projektów typu foresight realizowanych w Polsce wraz z listą omawianych technologii*, http://fortech2030.pl/images/stories/downloads/pdf/3_zaacznik_lista_projektw_typu_foresight_realizowanych_w_polsce.pdf

Wraz z aktywizacją prac badawczych powrócił problem skuteczności i zasadności stosowania metody delfickiej, jej wad i zalet. Oprócz spraw organizacyjnych⁶ podkreślano, że bardzo często opinie ekspertów były zbyt rozbieżne i nie uzyskiwano jednoznacznych rozstrzygnięć. W praktyce gospodarczej, gdzie ww. metoda stosowana jest przy tworzeniu strategii, spotkać się można z błędem metodycznym polegającym na kontynuowaniu procesu zbierania opinii specjalistów, mimo dużej rozbieżności zdań w pierwszej fazie badania. Zamiast wstrzymania prac i dokonania weryfikacji pytań ankietowych lub składu grupy eksperckiej, dokonuje się obliczeń statystycznych mających skutkować podjęciem konkretnych ustaleń. Znaczną odmiennosc opinii zdiagnozować można bez obliczeń, jednak przeoczeniu mogą ulec odmiennosci umiarkowane, ale istotne z merytorycznego punktu widzenia. Dlatego też w niniejszym artykule zaproponowano zastosowanie współczynnika zgodności Fleissa do weryfikacji wyników na wstępnym etapie procesu analizy odpowiedzi

⁶ / zob. Szewczyk R., [i in.], *Foresight priorytetowych, innowacyjnych technologii na rzecz automatyki, robotyki i techniki pomiarowej. Badanie metodą Delficką.*, Tom 2; Monografie, Studia, Rozprawy - Przemysłowy Instytut Automatyki i Pomiarów; Warszawa 2010, str. 13

eksperckich. Wskazano również zalecane modyfikacje formularzy stosowanych w metodzie delfickiej, służące obiektywizacji i poprawności prac naukowo – badawczych.

Podkreśla się, że wspomniany współczynnik nie może służyć do interpretacji wyników badania eksperckiego. Poza tym, stopień zgodności opinii nie przesądza jednoznacznie o poprawności przygotowania ankiety, czy doboru respondentów. Sygnalizuje on natomiast określony fakt, który powinien autorom badania zasugerować konieczność sprawdzenia poszczególnych elementów i narzędzi projektu badawczego.

2. Metoda delficka

G. Gierszewska i M. Romanowska⁷ wskazały, że metoda delficka została opracowana i zastosowana po raz pierwszy w 1963 roku przez naukowców z RAND Corporation. Istotą badania jest kilkakrotnie powtarzany proces zbierania opinii w formie ankiet od ekspertów z danych dziedzin. Eksperti nie mogą konsultować ze sobą stanowiska, nie powinni mieć nawet kontaktu ze sobą na żadnym z etapów badania. W zależności od konstrukcji kwestionariusza ankietowego, specjaliści muszą oszacować np. prawdopodobieństwo wystąpienia konkretnego zjawiska, bądź przedział czasowy, w którym dane zjawisko będzie miało miejsce i/lub jego konsekwencje. Odpowiedzi zebrane w pierwszej turze analizowane są łącznie. Następnie każda z osób otrzymuje wyniki analizy zbiorczej i proszona jest o ponowne wyrażenie opinii. Proces weryfikacji odpowiedzi prowadzony jest aż do uzyskania pełnej zgodności wszystkich ekspertów.

E. Knosala⁸ podkreśla również, że dla skuteczności badania prowadzonego metodą delficką, zaleca się by pytania ankietowe były zamknięte, a odpowiedzi na nie formułowano w kategoriach liczbowych.

Ponadto, R. Kozielski⁹ i A. Olesiuk¹⁰ podkreślają heurystyczny aspekt tej metody. Zdaniem autorów eksperci tworząc lub przewidyując wizje przyszłości muszą przede wszystkim myśleć twórczo. Praca indywidualna i kolejne etapy analizy odpowiedzi sprzyjają natomiast powstawaniu interesujących pomysłów.

Dodatkowo J. Kuciński¹¹ zauważył, że skoro foresight ma charakter zbiorowej refleksji nad przyszłością to nosi znamiona badania ogólnie społecznego. W tym kontekście

⁷ / G. Gierszewska, M. Romanowska, *Analiza strategiczna przedsiębiorstwa*, Państwowe Wydawnictwo Ekonomiczne, Warszawa 1995, str. 41 - 42

⁸ / E. Knosala, *Zarys teorii decyzji w nauce administracji*, Wolters Kluwer Polska, Warszawa 2011, str. 119

⁹ / R. Kozielski, *Biznes nowych możliwości*, Wolters Kluwer Polska, Warszawa 2012, str. 59

¹⁰ / A. Olesiuk, *Problemy organizacji: materiały do studiowania*, Key Text Wydawnictwo, 2007, str. 41

¹¹ / J. Kuciński, *Podręcznik metodyki foresight dla ekspertów projektu Foresight regionalny dla szkół wyższych Warszawy i Mazowsza „Akademickie Mazowsze 2030”*, <http://akademickiemazowsze2030.pl/Data/File/28.pdf>

zastosowanie metody delfickiej, w której istotną rolę odgrywają opinie ekspertów oraz interesariuszy z różnych środowisk, jest jak najbardziej wymagane i uzasadnione.

Wyniki otrzymywane w procesie zbierania opinii eksperckich na każdym etapie badania są analizowane zbiorczo przy użyciu metod i technik statystycznych, z których najczęściej stosowanymi są: średnia arytmetyczna, mediana, a także kwartyle¹². Powszechność i zasadność stosowania wskazanych współczynników wynika z praktyki analizy danych uzyskiwanych w badaniach ankietowych (przy użyciu kwestionariusza) oraz wywiadu, gdy respondent ocenia prawdopodobieństwo wystąpienia danego zjawiska lub zjawisk¹³. Zgodnie z informacjami publikowanymi przez autorów, metodę delficką wraz z analizą statystyczną wykorzystano we wszystkich projektach wykazanych w tabeli 1.

Tabela 2 Przykładowe opracowania dotyczące zastosowania metody delfickiej i analizy statystycznej w projektach foresight

Tytuł projektu	Źródło szczegółowego opisu metodyki badań
Foresight regionalny dla szkół wyższych Warszawy i Mazowsza „Akademickie Mazowsze 2030”	J. Kuciński, <i>Podręcznik metodyki foresight dla ekspertów projektu Foresight regionalny dla szkół wyższych Warszawy i Mazowsza „Akademickie Mazowsze 2030”</i> , http://akademickiemazowsze2030.pl/Data/File/28.pdf
Foresight technologiczny „NT FOR PODLASIE 2010”. Regionalna strategia rozwoju nanotechnologii	<i>Metodologia i procedury badawcze w projekcie foresight technologiczny „NT for Podlaskie 2020” – regionalna strategia rozwoju nanotechnologii</i> , red. J. Kazarko, J. Ejdyś http://ntfp2020.pb.edu.pl/pliki/Metodologia_i_procedury_badawcze_w_projekcie_Foresight_tehnologiczny_NT_FOR_Podlaskie_2020.pdf
Zeroemisyjna gospodarka energią w warunkach zrównoważonego rozwoju Polski do 2050 roku	<i>Scenariusze rozwoju zeroemisyjnej gospodarki energią w Polsce w perspektywie 2050 roku</i> , red. I. Pyka, K. Czaplicka – Kolarz http://www.foresightenergetyczny.pl/cscore/files/ZII.pdf
Perspektywa technologiczna Kraków Małopolska 2020	<i>Perspektywa technologiczna Kraków - Małopolska 2020. Wyzwania rozwojowe</i> , red. E. Bendyk, S. Kisieliński http://www.sse.krakow.pl/files/common/a-help-desk/Perspektywa%20Technologiczna%20Krakow%20-%20Małopolska%202020.pdf
Foresight „Sieci gospodarcze Wielkopolski” – scenariusze transformacji wiedzy wspierające innowacyjną gospodarkę	<i>Raport z badań metodą Delhi do projektu „Foresight – sieci gospodarcze Wielkopolski – scenariusze transformacji wiedzy wspierające innowacyjną gospodarkę</i> , red. M. Wyrwicka, Wydawnictwo Politechniki Poznańskiej, Poznań 2011
Identyfikacja potencjału i zasobów Dolnego Śląska w obszarze nauka i technologie na rzecz poprawy jakości życia (Quality of Life)	<i>Identyfikacja potencjału i zasobów Dolnego Śląska w obszarze nauka i technologie na rzecz poprawy jakości życia (Quality of life) oraz wytyczenie przyszłych kierunków rozwoju. Badania metodami foresight</i> , red. J. Hanuz, E. Cibis, T. Miśkiewicz, P. Ziółkowski

¹² / szczegółową metodykę zastosowania analizy statystycznej w badaniach foresight przedstawiono m.in. w opracowaniu *Raport z badań metodą Delhi do projektu „Foresight – sieci gospodarcze Wielkopolski – scenariusze transformacji wiedzy wspierające innowacyjną gospodarkę*, red. M. Wyrwicka, Wydawnictwo Politechniki Poznańskiej, Poznań 2011

¹³ / zob. W. Starzyńska, *Statystyka praktyczna*, Wydawnictwo Naukowe PWN, Warszawa 2006; A. D. Aczel, *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa 2006; czy M. Sobczyk, *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa 2004

oraz wytyczenie przyszłych kierunków rozwoju. Badania metodami foresight

http://www.gol.ue.wroc.pl/p/ /65/identyfikacja_potencjal...jerzy_hanuz_a_m.pd

Źródło: opracowanie własne

Wielokrotne powtarzanie badania ankietowego przy użyciu tego samego kwestionariusza ma oczywiście na celu uzgodnienie rozwiązania końcowego. Należy zdawać sobie jednak sprawę z tego, że bez poprawnej weryfikacji wyników pierwszego etapu, czyli obiektywnego stwierdzenia stopnia zgodności, dalsze działania mogą mieć znamiona przekonywania ekspertów do odpowiedzi uznanych za najczęściej występujące. W końcu na skutek konformizmu, zmęczenia lub wygody respondenci zaznaczać będą te same odpowiedzi. Takie postępowanie nie można jednakże nazwać naukowym.

3. Współczynnik Fleissa

W 1971 r. Joseph L. Fleiss, profesor biostatystyki Uniwersytetu Columbia w Stanach Zjednoczonych, opublikował w *Psychological Bulletin* artykuł pt. *Measuring Nominal Scale Agreement Among Many Raters*¹⁴, w którym zaprezentował współczynnik zgodności ocen eksperckich. Koncepcja Fleissa była rozwinięciem współczynnika Cohena¹⁵, którym mierzyć można zgodność ocen wystawionych przez dwóch oceniających¹⁶. Zauważając ograniczenia modelu, Fleiss opracował współczynnik, który znajduje zastosowanie w sytuacji, gdy oceniających jest nawet więcej niż dwóch, ponadto badanie dotyczyć może kilku zmiennych.

Zakłada się, że N to łączna liczba obiektów podlegających ocenie, n – liczba oceniających, czyli liczba wykonanych pomiarów, k – ilość kategorii, do których można zaklasyfikować obiekt; ponadto, i , gdzie $i = 1, \dots, N$, reprezentuje obiekt, a j , gdzie $j = 1, \dots, k$, reprezentuje daną kategorię. Wskaźnik n_{ij} określa liczbę oceniających, którzy wystawili ocenę i -temu obiektowi w j -tej kategorii.

Dla wyjaśnienia autor zaprezentował przykład wyników badania, w którym 6 psychiatrów ($n = 6$), diagnozowało występowanie 5 rodzajów chorób ($k = 5$) wśród 30 pacjentów ($N = 30$). W tabeli 3 wyszczególniono wyniki.

¹⁴ / J. L. Fleiss, *Measuring Nominal Scale Agreement Among Many Raters*, *Psychological Bulletin*, 1971, Vol. 76, No.5

¹⁵ / zob. Cohen J. A., *A coefficient of agreement for nominal scale*, *Educational and Psychological Measurement*, 1960, 20; Cohen J. A., *Weight kappa: Nominal scale agreement with provision for scaled disagreement or partial credit*, *Psychological Bulletin*, 1968, 70

¹⁶ / opis innych współczynników służących do badania zgodności zamieściła m.in. J. Jarosz – Nowak w artykule *Modele oceny stopnia zgody pomiędzy dwoma ekspertami z wykorzystaniem współczynników kappa*, *Matematyka Stosowana*, 8/2007

Tabela 3 Przykładowe wyniki diagnozy oraz wartości wskaźników p_j i P_i

Podmiot (pacjent)	Kategoria (choroba)					P_i = 16,667
	Depresja ($j = 1$)	Problemy osobowościowe ($j = 2$)	Schizofrenia ($j = 3$)	Neuroza ($j = 4$)	Inne ($j = 5$)	
1				6		1,000
2		3			3	0,400
3		1	4		1	0,400
4					6	1,000
5		3		3		0,400
6	2		4			0,467
7			4		2	0,467
8	2		3	1		0,267
9	2			4		0,467
10					6	1,000
11	1			5		0,667
12	1	1		4		0,400
13		3	3			0,400
14	1			5		0,667
15		2		3	1	0,267
16			5		1	0,667
17	3			1	2	0,267
18	5	1				0,667
19		2		4		0,467
20	1		2		3	0,267
21					6	1,000
22		1		5		0,667
23		2		1	3	0,267
24	2			4		0,467
25	1			4	1	0,400
26		5		1		0,667
27	4				2	0,467
28		2		4		0,467
29	1		5			0,667
30					6	1,000
Łącznie = 180	26	26	30	55	43	
p_j	0.144	0.144	0.167	0.306	0.239	

Źródło: J. L. Fleiss, *Measuring Nominal Scale Agreement Among Many Raters*, Psychological Bulletin, 1971, Vol. 76, No.5

Obliczenia rozpocząć należy od wyliczenia wskaźnika p_j i P_i . Wychodząc z podstawowego wzoru:

$$[1] \quad p_j = \frac{1}{Nn} \sum_{i=1}^N n_{ij}$$

Wskaźnik p_j dla danej kolumny danych wylicza się zatem następująco:

$$p_1 = \frac{26}{180} = 0,144$$

Wskaźnik P_i dla danego wiersza wylicza się zgodnie ze wzorem:

$$[2] \quad P_i = \frac{1}{n(n-1)} \sum_{j=1}^k [n_{ij}(n_{ij}-1)]$$

Zatem przykładowo dla drugiego wiersza należy przeprowadzić następujące obliczenia:

$$P_2 = \frac{1}{6(6-1)} (0^2 + 3^2 + 0^2 + 0^2 + 3^2 - 6) = 0,400$$

Następnie wylicza się średnią wartość współczynnika P_i :

$$\bar{P} = 16,667 : 30 = 0,556$$

W celu określenia współczynnika zgodności należy przeprowadzić jeszcze kolejne obliczenia:

$$[3] \quad \bar{P}_e = \sum_{j=1}^k p_j^2$$

Dla prezentowanego przykładu będzie to:

$$\bar{P}_e = 0,144^2 + 0,144^2 + 0,167^2 + 0,306^2 + 0,239^2 = 0,220$$

Wstawiając otrzymane dane do wzoru:

$$[4] \quad \kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

otrzymuje się wartość współczynnika Fleissa, który dla danych prezentowanych w tabeli 3 wynosi 0,431.

$$\kappa = \frac{0,556 - 0,22}{1 - 0,22} = 0,431$$

W celu ułatwienia interpretacji wyliczanych wartości autor koncepcji zaproponował następujące zakresy¹⁷:

¹⁷ / J. Jarosz – Nowak, *Modele ...*, *op.cit.*

- $< 0,40$ słaba zgodność
- $0,40 - 0,74$ umiarkowana lub dobra
- $0,75 - 1,00$ perfekcyjna

J. R. Landis, G. G. Koch zasugerowali bardziej szczegółową klasyfikację przedziałów, a mianowicie¹⁸:

- $< 0,00$ brak zgodności
- $0,01 - 0,20$ bardzo słaba zgodność
- $0,21 - 0,40$ zauważalna zgodność
- $0,41 - 0,60$ średnia zgodność
- $0,61 - 0,80$ duża zgodność
- $0,81 - 1,00$ prawie idealna zgodność

Wybór skali interpretacji pozostawia się autorom badania w zależności od złożoności badania, istotności wnioskowania, czy zastosowania wyników.

4. Przykłady zastosowań

Metoda delficka, z uwagi na swój uniwersalny charakter, może być stosowana zarówno w makroekonomicznych badaniach prognostycznych, projektach typu foresight, jaki również w zarządzaniu strategicznym. Natomiast współczynnik zgodności Fleissa znaleźć może zastosowanie w procesie weryfikacji wyników ww. badań, a ponadto w wielu analizach prowadzonych nawet na poziomie zarządzania operacyjnego. Dotychczas omawianym współczynnikiem posługiwano się w badaniach medycznych, psychologicznych, a nawet w sądownictwie przy ocenie zgodności opinii sędziów.

Istotnym kryterium wykorzystania omawianego współczynnika jest takie przygotowanie badania, aby: (1) brały w nim udział osoby oceniające, (2) sformułowane były konkretne kategorie oceny, (3) wyszczególnione zostały podmioty lub przedmioty oceny. Zauważyć można zatem szerokie spektrum zastosowań. W dalszej części artykułu wskazano przykłady użycia współczynnika Fleissa na poziomie mikro i makro analiz.

Przykład pierwszy zaprezentowano posługując się analizą fikcyjnych wyników badania ankietowego zawierającego autentyczne pytania opracowane przez Przemysłowy Instytut Automatyki i Pomiarów (PIAP), Politechnikę Warszawską oraz Instytut Badań Systemowych Polskiej Akademii Nauk (IBS PAN). Założono, że w badaniu uczestniczyło 15 ekspertów ($n = 15$), określano

¹⁸ / J. R. Landis, G. G. Koch, *The measurement of observer agreement for categorical data*, Biometrics. Vol. 33, 1977

pięciostopniową skalę wpływu i znaczenia ($k = 5$) wybranych czterech czynników ($N = 4$). W tabeli 4 wyszczególniono rezultaty.

Tabela 4 Przykład nr 1

Pytanie		Skala wpływu / znaczenia					P_i = 1,314
		Bardzo wysoki	Wysoki	Średni	Niski	Bardzo niski	
		$j = 1$	$j = 2$	$j = 3$	$J = 4$	$j = 5$	
Proszę określić wpływ grupy technologii na zrównoważony rozwój ekonomiczno – społeczny Polski		1	4		10		0,486
Proszę określić znaczenie grupy technologii za pomocą podanych cech / wskaźników	Skala przewidywanych korzyści z zastosowania	3	3	5	1	3	0,181
	Priorytetowość działania na rzecz grupy technologii	1	8			6	0,410
	Zapotrzebowanie	2	2	2	2	7	0,238
Łącznie = 60		7	17	7	13	16	
p_j		0,117	0,283	0,117	0,217	0,267	

Źródło: opracowanie własne z wykorzystaniem wybranych pytań ankietowych opublikowanych w raporcie *Foresight ARP. Foresight priorytetowych, innowacyjnych technologii na rzecz automatyki, robotyki i techniki pomiarowej. Badanie metodą delficką*, Przemysłowy Instytut Automatyki i Pomiarów (PIAP), Politechnika Warszawska, Instytut Badań Systemowych Polskiej Akademii Nauk (IBS PAN), str. 29

Dla wartości wskazanych w tabeli 4 współczynnik Fleissa κ wynosi 0,133, czyli zgodnie z zakresami interpretacji jest to bardzo słaba zgodność odpowiedzi. Zamiast dalszej interpretacji należałoby zweryfikować poprawność odpowiedzi, czy innych czynników.

Przykład nr 2 dotyczy wyznaczania kierunków zmian podstawowych wskaźników makroekonomicznych. W tym przypadku przyjęto, że w badaniu uczestniczyło 100 ekspertów ($n = 100$), wyznaczono trzy kierunki zmiany ($k = 3$) w stosunku do czterech czynników ($N = 4$). W tabeli 5 wyszczególniono fikcyjne dane.

Tabela 5 Przykład nr 2

Czynnik	Kierunek zmiany			$P_i = 2,848$
	Wzrost ($j = 1$)	Stagnacja ($j = 2$)	Spadek ($j = 3$)	
1	40	40	20	0,354
2			100	1,000
3			100	1,000
4		50	50	0,495
Łącznie = 400	40	90	270	
p_j	0,100	0,225	0,675	

Źródło: opracowanie własne

Dla wartości wskazanych w powyższej tabeli współczynnik Fleissa κ wynosi 0,405 czyli eksperci wykazują zauważalną, ale nadal słabą zgodność odpowiedzi.

W przykładzie nr 3 zaprezentowano wykorzystanie analizy zgodności na poziomie zarządzania operacyjnego do oceny kandydatów do pracy. W zadaniu założono, że 20 ekspertów ($n = 20$), wskazywało przydatność 30 kandydatów ($N = 30$) do pracy w jednym z pięciu stanowisk pracy ($k = 5$).

Tabela 6 Przykład nr 3

Numer kandydata	Stanowisko pracy					$P_i = 23,30$
	Sprzedawca ($j = 1$)	Magazynier ($j = 2$)	Kierownik sklepu ($j = 3$)	Kasjer ($j = 4$)	Przedstawiciel handlowy ($j = 5$)	
1		13		6	1	0,489
2	20					1,000
3	20					1,000
4		20				1,000
5		20				1,000
6			18	2		0,811
7		18		2		0,811
8	7			13		0,521
9		4	16			0,663
10		6	14			0,558
11			20			1,000
12		1			19	0,900
13	2			18		0,811
14		1	19			0,900
15	1	1	1	3	14	0,495
16		2	2	5	11	0,353
17	18		1		1	0,805
18		17	1	1	1	0,716
19		2		18		0,811
20			20			1,000
21		20				1,000
22				20		1,000
23		1	19			0,900
24			1	17	2	0,721
25		20				1,000
26					20	1,000
27			20			1,000
28	1	1	4	4	10	0,300
29		12		3	5	0,416
30	11	2	2	3	2	0,321
Łącznie = 600	80	161	158	115	86	
p_j	0,1333	0,2683	0,2633	0,1917	0,1433	

Źródło: opracowanie własne

W powyższym przykładzie współczynnik κ wynosi 0,715, czyli eksperci wykazali dużą zgodność, tym samym badanie można uznać za prawidłowe, a analizie poddać wyłącznie pozycje przy których $P_i < 1$.

5. Zakończenie

Zaprezentowane przykłady stanowią wyjaśnienie wykorzystania współczynnika Fleissa. Jednocześnie celem autorki było zainspirowanie naukowców do poszukiwania innych zastosowań ww. współczynnika, z całą świadomością wad i zalet, z którymi jest on związany. Do pozytywnych cech wykorzystania powyższego sposobu obliczeń zaliczyć należy:

- możliwość szybkiej i kwantyfikowalnej weryfikacji zgodności odpowiedzi eksperckich, a przez to uniknięcia błędów interpretacyjnych w dalszych etapach badania w metodzie delfickiej;
- możliwość zastosowania w przypadku analiz wielokryterialnych, w których udział bierze kilkusobowa grupa ekspertów,
- łatwość zastosowania i przeprowadzenia obliczeń, np. z wykorzystaniem odpowiednio przygotowanego arkusza kalkulacyjnego takiego jak Microsoft Excel (Microsoft przygotował nawet oprogramowanie o nazwie AgreeStat, które współpracuje z Excelem; oczywiście funkcję wyliczania współczynnika Fleissa mają także programy specjalistyczne stosowane do obliczeń statystycznych, np. program STATISTICA).

Głównymi mankamentami i ograniczeniami zastosowania współczynnika Fleissa są następujące właściwości:

- możliwość wskazania przez ekspertów tylko jednego kryterium (j) dla danego podmiotu/przedmiotu (i),
- brak możliwości jakiegokolwiek stopniowości, czy wskazania prawdopodobieństwa wystąpienia danego kryterium (j).

Podstawowy problem polega zatem na tym, że omawiany współczynnik nie może być zastosowany, w przypadku gdy specjaliści nie tylko diagnozują wystąpienie danego zjawiska, ale określają również prawdopodobieństwo jego wystąpienia lub jakąkolwiek miarę szacują jego wartość.

Literatura:

1. Aczel A. D., *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa 2006
2. Borodako K., *Foresight w zarządzaniu strategicznym*, Wydawnictwo C.H. Beck, Warszawa 2009
3. Cohen J. A., *A coefficient of agreement for nominal scale*, Educational and Psychological Measurement, 1960, 20
4. Cohen J. A., *Weight kappa: Nominal scale agreement with provision for scaled disagreement or partial credit*, Psychological Bulletin, 1968, 70
5. Fleiss J. L., *Measuring Nominal Scale Agreement Among Many Raters*, Psychological Bulletin, 1971, Vol. 76, No.5

6. *Foresight ARP. Foresight priorytetowych, innowacyjnych technologii na rzecz automatyki, robotyki i techniki pomiarowej. Badanie metodą delficką*, Przemysłowy Instytut Automatyki i Pomiarów (PIAP), Politechnika Warszawska, Instytut Badań Systemowych Polskiej Akademii Nauk (IBS PAN)
7. Gierszewska G., Romanowska M., *Analiza strategiczna przedsiębiorstwa*, Państwowe Wydawnictwo Ekonomiczne, Warszawa 1995
8. *Identyfikacja potencjału i zasobów Dolnego Śląska w obszarze nauka i technologie na rzecz poprawy jakości życia (Quality of life) oraz wytyczenie przyszłych kierunków rozwoju. Badania metodami foresight*, red. J. Hanuz, E. Cibis, T. Miśkiewicz, P. Ziółkowski,
http://www.qol.ue.wroc.pl/p/ /65/identyfikacja_potencjal_u...jerzy_hanuza_m.pdf
9. Jarosz – Nowak J., *Modele oceny stopnia zgody pomiędzy dwoma ekspertami z wykorzystaniem współczynników kappa*, *Matematyka stosowana* 8, 2007,
http://www.matstos.pjwstk.edu.pl/no8/no8_jarosz-nowak.pdf
10. Knosala E., *Zarys teorii decyzji w nauce administracji*, Wolters Kluwer Polska, Warszawa 2011
11. Kozielski R., *Biznes nowych możliwości*, Wolters Kluwer Polska, Warszawa 2012
12. Kuciński J., *Podręcznik metodyki foresight dla ekspertów projektu Foresight regionalny dla szkół wyższych Warszawy i Mazowsza „Akademickie Mazowsze 2030”*,
<http://akademickiemazowsze2030.pl/Data/File/28.pdf>
13. Landis J. R., Koch G. G., *The measurement of observer agreement for categorical data*, *Biometrics*. Vol. 33, 1977
14. *Lista projektów typu foresight realizowanych w Polsce wraz z listą omawianych technologii*,
http://fortech2030.pl/images/stories/downloads/pdf/3_zaaecznik_lista_projektu_typ_u_foresight_realizowanych_w_polsce.pdf
15. *Metodologia i procedury badawcze w projekcie foresight technologiczny „NT for Podlaskie 2020” – regionalna strategia rozwoju nanotechnologii*, red. J. Kazarko, J. Ejdyś
http://ntfp2020.pb.edu.pl/pliki/Metodologia_i_procedury_badawcze_w_projekcie_Foresight_tehnologiczny_NT_FOR_Podlaskie_2020.pdf
16. Olesiuk A., *Problemy organizacji: materiały do studiowania*, Key Text Wydawnictwo, 2007
17. *Perspektywa technologiczna Kraków - Małopolska 2020. Wyzwania rozwojowe*, red. E. Bendyk, S. Kisieliński <http://www.sse.krakow.pl/files/common/a-help-desk/Perspektywa%20Technologiczna%20Krakow%20-%20Malopolska%202020.pdf>
18. *Projekty badawcze z wykorzystaniem metody foresight; wybrane projekty realizowane w ramach Poddziałania 1.1.1 Programu Innowacyjna Gospodarka*, Ośrodek Przetwarzania Informacji, Warszawa 2010
19. *Raport z badań metodą Delhi do projektu „Foresight – sieci gospodarcze Wielkopolski – scenariusze transformacji wiedzy wspierające innowacyjną gospodarkę*, red. M. Wyrwicka, Wydawnictwo Politechniki Poznańskiej, Poznań 2011
20. *Scenariusze rozwoju zeroemisyjnej gospodarki energią w Polsce w perspektywie 2050 roku*, red. I. Pyka, K. Czaplicka – Kolarz
<http://www.foresightenergetyczny.pl/cscore/files/ZII.pdf>
21. Sobczyk M., *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa 2004
22. Starzyńska W., *Statystyka praktyczna*, Wydawnictwo Naukowe PWN, Warszawa 2006

23. Szewczyk R., [i in.], *Foresight priorytetowych, innowacyjnych technologii na rzecz automatyki, robotyki i techniki pomiarowej. Badanie metodą Delficką., Tom 2; Monografie, Studia, Rozprawy - Przemysłowy Instytut Automatyki i Pomiarów; Warszawa 2010*
24. <http://foresight.polska2020.pl/cms/>
25. <http://www.opi.org.pl/pl/stan-wdrazaniapoig/stan-wdrazania-111/>
26. <http://www.kosmos.gov.pl/index.php?link=94>
27. <http://www.akademickiemazowsze2030.pl>
28. <http://www.forsurf.pl>
29. <http://zywnoscizywienie.spoleczna.pl/>
30. <http://ntfp2020.pb.edu.pl>
31. <http://pomorze2030.pl>
32. <http://www.foresightenergetyczny.pl>
33. <http://www.foresight.itee.radom.pl>
34. <http://foresight-ogwk.pl>
35. <http://www.fsgw.put.poznan.pl>
36. <http://www.qol.ue.wroc.pl>

FLEISS KAPPA AS A TOOL OF RESEARCH RESULTS VERIFICATION OBTAINED IN DELPHI METHODS OF MANAGEMENT AND ECONOMIC PHENOMENONS PROGNOSIS

Summary

In the paper author presented examples of the application of Fleiss kappa (a coefficient of compliance) for verification of results obtained using the Delphi method. The growth of interest of mentioned method is explained, particularly in the context of scientific and research foresight type of projects. The relevance of the use of Fleiss kappa in the process of diagnosis and forecasting economic, strategic and management measure is highlighted. Moreover, examples of specific tables for experts' research are presented. Key words: foresight, the Delphi method, Fleiss kappa