



## Mirosław Wójciak

Uniwersytet Ekonomiczny w Katowicach  
Wydział Zarządzania  
Katedra Ekonometrii  
mwojciak@ue.katowice.pl

# METODY OCENY ZGODNOŚCI OPINII EKSPERTÓW NA POTRZEBY BADANIA FORESIGHT

**Streszczenie:** Nadrzędnym celem foresightu jest konstrukcja scenariuszy rozwoju sytuacji w stosunkowo dalekiej perspektywie (zwykle 20-25 lat). W celu ich opracowania foresight korzysta z wielu różnych metod, wśród których dużym uznaniem cieszą się działania oparte na pozyskiwaniu wiedzy eksperckiej, głównie panele eksperckie, burze mózgów oraz metoda Delphi. W celu sformułowania jednoznacznych sądów konieczne jest uzyskanie zgodności odpowiedzi ekspertów. W przypadku skal słabych stosuje się m.in. współczynnik: dyspersji, wariacji, rangowej korelacji parami czy konkordancji.

W badaniu przedstawiono stosowane miary zgodności ekspertów wraz z ich własnościami. Umożliwi to wyznaczenie granicznych wartości miar zgodności ekspertów, których osiągnięcie pozwoli uznać, że eksperci są zgodni.

**Słowa kluczowe:** foresight, zgodność opinii ekspertów, prognozowanie heurystyczne.

## Wprowadzenie

Nadrzędnym celem foresightu jest konstrukcja scenariuszy rozwoju sytuacji w stosunkowo dalekiej perspektywie oraz w każdym przypadku, gdy nie jest możliwa ekstrapolacja posiadanej wiedzy. W celu ich opracowania foresight korzysta z wielu różnych metod, które są ciągle modyfikowane. Dużym uznaniem cieszą się działania oparte na pozyskiwaniu wiedzy eksperckiej, głównie panele eksperckie, burze mózgów oraz metoda Delphi. Dlatego też podstawowymi narzędziami są tu metody bazujące na wiedzy ekspertów merytorycznych, których kompetencje powinny być wysoko ocenione. W celu sformułowania jednoznacznych sądów konieczne jest uzyskanie zgodności odpowiedzi ekspertów.

## 1. Problematyka badań typu foresight

Najczęściej cytowaną definicją foresightu jest definicja B. Martina, który określa go jako „proces zaangażowany w systematyczne próby spojrzenia na długoterminową przyszłość nauki, technologii gospodarki oraz społeczeństwa, mający na celu identyfikację obszarów badań strategicznych oraz powstających technologii genetycznych, które mają potencjał przyniesienia najwyższych korzyści gospodarczych i społecznych” [Martin, 1995]. Cele foresightu są następujące:

- przemyślenie przyszłości,
- przeprowadzenie na jej temat specjalistycznej publicznej debaty,
- podjęcie w obecnym czasie działań na rzecz odpowiedniego ukształtowania przyszłości.

Pierwsze badania foresightowe przeprowadziła Japonia na początku lat 70. [zob. Kuwahara, 2002], ale prawdziwą popularność projekty te zyskały na całym świecie w latach 90. ubiegłego wieku. Obecnie w literaturze światowej rozróżnia się dwa główne rodzaje foresightów: regionalny (*Regional Foresight*) i technologiczny (*Technology Foresight*) oraz ich kombinacje [Czaplicka-Kolarz (red.), 2007, tom 1, s. 13]. W foresighcie regionalnym identyfikuje się kluczowe kierunki rozwoju dla regionu, uzyskuje się w tym obszarze konsensus społeczny oraz tworzy sieć współpracy między organizacjami i jednostkami uczestniczącymi w pracach nad rozwojem. Natomiast foresight technologiczny ma na celu identyfikację ważnych, wschodzących technologii, które są istotne dla poprawy warunków ekonomicznych oraz jakości życia społeczeństwa [Halal, 1997]. Badanie foresight jest zadaniem dla specjalistów. Stanowi ono dogodną okazję do przedstawienia swego stanowiska przez szeroki krąg osób w społeczeństwie, dla którego wyniki takiej analizy nie powinny być obojętne.

W celu kreowania przyszłości foresight korzysta z wielu różnych metod, wśród których jako główne narzędzie pozyskiwania wiedzy można wyróżnić metodę delficką. Polega ona na przeprowadzeniu kilkakrotnego ankietowania wybranej grupy anonimowych ekspertów, którzy nie mogą się ze sobą w tej sprawie komunikować i naradzać. Aby wyniki badania można uznać za wiarygodne, ankietowana grupa powinna być liczna, reprezentatywna, a eksperci powinni posiadać dużą wiedzę merytoryczną i doświadczenie w tematyce będącej przedmiotem badania.

Zebranie dużej grupy ekspertów wystarczająco kompetentnych w danej dziedzinie bywa problematyczne. W realizowanych w Polsce foresightach zdarza się, że w niektórych panelach tematycznych uczestniczy tylko kilku ekspertów, np. gdy przedmiotem badania są nowoczesne, zaawansowane technologie.

## 2. Rekrutacja ekspertów

Wykorzystanie wiedzy ekspertów przy budowie prognoz czy też scenariuszy rozwoju niesie ze sobą niebezpieczeństwo uzyskania błędnych wyników, gdyż prognozy będą tak samo trafne, jak eksperci biorący udział w badaniu. Nie zawsze osoby uważające siebie za ekspertów są nimi w rzeczywistości, a także osoby mające bogate doświadczenie w danej dziedzinie nie będą dobrymi prognostami przyszłości, np. mogą być zbyt optymistyczne w swych poglądach. Dlatego odpowiedni dobór ekspertów jest kluczowym problemem badań typu foresight, czy też badań Delphi.

W projektach typu foresight stosuje się kilka sposobów zapraszania ekspertów do udziału w panelach badawczych. W przeciwieństwie do prób sondażowych, stosuje się dobór celowy, dzięki któremu uzyskujemy osoby najbardziej kompetentne z punktu widzenia badania. Najczęściej stosuje się metodę rekomendacji (nominacji czy wspólnominacji), kiedy to osoby powszechnie uważane za autorytet lub eksperci już uczestniczący w pracach w projekcie wskazują specjalistów w danej dziedzinie. W celu wyboru kompetentnych ekspertów w metodzie rekomendacji można kierować się np. analizą publikacji w danej dziedzinie. Jednak stosując analizę bibliografii, można pominąć osoby mające doświadczenie i kompetencje eksperta, którzy nie publikują swoich prac, np. przedstawiciele biznesu. Niewątpliwą zaletą tej metody jest gwarancja wyboru osób kompetentnych w danej dziedzinie i uniknięcie osób przypadkowych. Jednak ten typ rekrutacji ekspertów posiada także wady. Wysokiej klasy specjaliści mający uznanie w środowisku, tzw. top-eksperci, mogą być zbyt optymistyczni w ocenie szans rozwoju własnej dziedziny, nie doceniając jej barier. W związku z tym powinno się dążyć do zróżnicowania poziomu wiedzy eksperckiej uczestników badania. Chcąc zapewnić zróżnicowany poziom wiedzy eksperckiej uczestników, należy wziąć pod uwagę ich reprezentatywność, czemu sprzyja otwarta rekrutacja ekspertów. Drugą wadą omawianej metody jest dobór współpracowników przez już wybrane osoby, tzw. wspólnominacja. Eksperci najczęściej wskazują osoby, z którymi im się dobrze pracuje, i unikają osób o odmiennych poglądach. W efekcie eksperci mają ograniczone spojrzenie na dane zagadnienie, co wprawdzie ułatwia wypracowanie konsensusu, ale jednocześnie uniemożliwia pojawianie się nowych koncepcji i alternatywnych pomysłów [Gajda, 2011, s. 32].

Innym stosowanym sposobem zapraszania ekspertów do prac w projekcie jest rekrutacja otwarta, kiedy każda osoba może się zgłosić do realizujących projekt i po spełnieniu pewnych kryteriów zostać ekspertem panelu badawczego. Metoda ta wprawdzie zmniejsza możliwość pojawienia się w panelu przedstawi-

cieli reprezentujących tylko jeden pogląd na dane zagadnienie, jednakże niesie ze sobą ryzyko pojawienia się wśród ekspertów osób o niewystarczającej wiedzy i doświadczeniu. Można oczywiście zastosować także system mieszany, czyli zarówno metodę rekomendacji, jak i rekrutacji otwartej [Kowalewska i Głuszyński (red.), 2009, s. 22].

W realizowanych projektach korzystających z metodyki foresight panele składają się głównie z ekspertów wywodzących się ze środowiska naukowego [Kowalewska i Głuszyński (red.), 2009, s. 22; Gajda, 2011, s. 32]. Przedstawiciele biznesu znacznie rzadziej deklarują chęć uczestniczenia w badaniach. Spowodowane jest to m.in. długim horyzontem badań nieprzynoszących szybkich i wyraźnych rezultatów, których oczekują w codziennej praktyce. Dodatkowo eksperci ze świata biznesu często deklarują brak czasu i w niektórych przypadkach oczekują gratyfikacji finansowej.

Dodatkowym czynnikiem decydującym o wyborze metody doboru ekspertów jest ich dostępność. W takich przypadkach konieczne jest zastosowanie metody nominacji i współnominacji ekspertów [Gajda, 2011, s. 32].

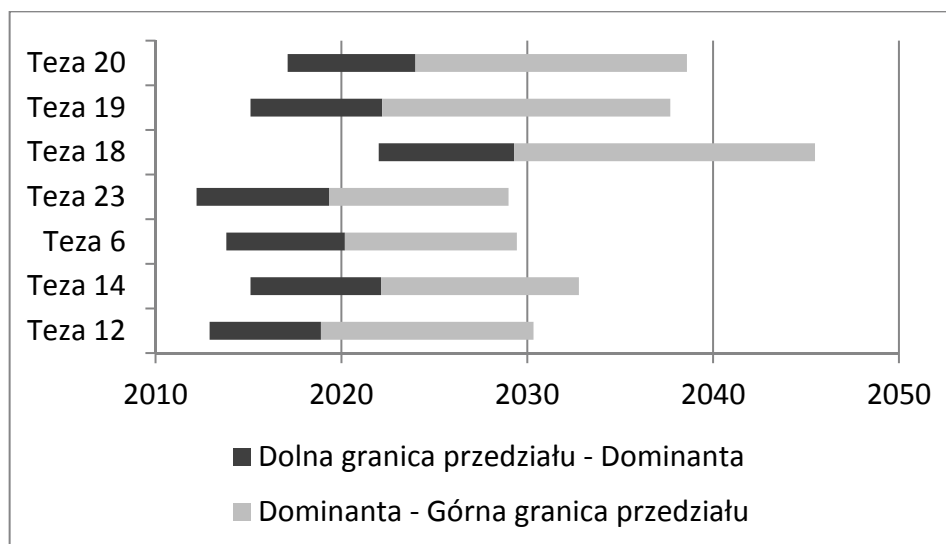
### 3. Mierniki oceny zgodności opinii ekspertów

Metody heurystyczne opierają się na założeniu, że trafność sądów grupowych jest wyższa niż indywidualnych ekspertów. Informacje dotyczące przedmiotu badań otrzymane od grupy są co najmniej tak wartościowe, jak od każdego z uczestników grupy z osobna. Szeroka i wyjątkowa wiedza eksperta w określonej dziedzinie może rekompensować niewiedzę i ignorancję innych ekspertów, mających doświadczenie w innych dziedzinach. Jeśli uczestnicy badania zostaną rozważnie dobrani, to będą wymieniać swoją wyjątkową wiedzę i zrównoważą poglądy – ma to szczególnie znaczenie w metodzie delfickiej, w której dąży się do uzyskania zgodności ekspertów. Zgodność ekspertów może być także wymagana w celu:

1. Oceny wpływu czynników/zdarzeń na rozwój danego zjawiska.
2. Oceny barier/stymulatorów rozwoju danego zjawiska.
3. Budowy precyzyjniejszych prognoz.

Ocena wpływu zdarzeń na rozwój analizowanego zjawiska, czy też ocena barier i stymulatorów może pomóc w budowie scenariuszy jego rozwoju. Dotyczy to zarówno rozwoju nowoczesnych technologii, jak i ich otoczenia. W tym przypadku zgodność ekspertów jest wymagana w celu określenia „czynników kluczowych” i ich wpływu na rozwój analizowanego zjawiska. Pozwoli to nie tylko ograniczyć liczbę scenariuszy rozwoju, ale także precyzyjniej je opisać [Gajda, 2011, s. 35].

Na podstawie skonstruowanych scenariuszy budowane są prognozy rozwoju nowych technologii lub prognozy czasu realizacji tez. W przypadku, gdy eksperci nie wykażą się dostateczną zgodnością swoich opinii, prognozy te będą obarczone wysokim stopniem niepewności, a rozpiętość zbudowanych przedziałów prognoz będzie zbyt szeroka [por. Poradowska i Wójciak, 2009; Poradowska i in., 2011].



**Rys. 1.** Przykładowe przedziały prognoz dla czasu realizacji wybranych tez

Źródło: Poradowska i Wójciak [ 2011].

Pozyskane od ekspertów dane poddaje się analizie statystycznej, która obejmuje wyznaczenie miar położenia oraz ocenę zgodności opinii. Wśród miar położenia najczęściej stosuje się miary pozycyjne, tj. medianę i dominantę. Zgodność ekspertów można scharakteryzować za pomocą miar zmienności [Cieślak (red.), 1997, s. 180]. Wybór miernika oceny zgodności opinii ekspertów dokonuje się w zależności od skali, na której są mierzone odpowiedzi ekspertów.

1. Jeśli mamy do czynienia ze skalami mocnymi, tj. skalą przedziałową lub ilorazową, to do oceny zgodności stosuje się rozstęp międzykwartyłowy:

$$\Delta = Q_3 - Q_1, \quad (1)$$

$\Delta \leq \Delta^*$  to eksperci są zgodni, czyli rozstęp międzykwartyłowy nie przekroczy jakiegś z góry ustalonej liczby (wartości progowej) określonej przez organizatorów badania. Im niższa wartość miary 1, tym większa zgodność ekspertów.

2. Jeśli odpowiedzi ekspertów są mierzone na skali nominalnej, ich zgodność można sprawdzić współczynnikiem dyspersji względnej klasyfikacji:

$$h = \frac{k}{k-1} \left( 1 - \sum_j f_j^2 \right), \quad h_r \in [0,1], \quad (2)$$

gdzie:

$k$  – liczba wariantów dla  $r$ -tego pytania,

$f_j$  – częstość występowania  $j$ -tego wariantu wśród odpowiedzi.

Miara 2 jest unormowana w przedziale  $[0; 1]$  i im wartość miary  $h$  bliższa zero, tym większa zgodność ekspertów.

3. Przeciętny stopień zgodności poglądów wszystkich ekspertów przy stosowaniu skali porządkowej można zbadać współczynnikiem konkordancji Kendalla i Smitha:

$$W = \frac{12S}{n^2(k^3 - k)}, \quad (3)$$

gdzie:

$n$  – jest liczbą ekspertów,  $k$  – liczbą wariantów, a  $S$  wyraża się wzorem:

$$S = \sum_{j=1}^k \left( \sum_{i=1}^n x_{ij} - \bar{x} \right)^2, \quad (4)$$

gdzie:

$$\bar{x} = \frac{1}{k} \sum_{i=1}^n \sum_{j=1}^k x_{ij}.$$

Współczynnik konkordancji przybiera wartości z przedziału  $(0,1]$ ; duża wartość świadczy o zgodności ekspertów. Do oceny istotności współczynnika  $W$  może posłużyć statystyka  $\chi^2$  o  $k-1$  stopniach swobody

$$\chi^2 = \frac{12S}{nk(k+1)}, \quad (5)$$

dla której weryfikowana jest hipoteza zerowa, że eksperci nie są zgodni wobec hipotezy alternatywnej, która stanowi, iż eksperci są zgodni. W przypadku braku podstaw do odrzucenia hipotezy zerowej można wnioskować, że odpowiedzi ekspertów są przypadkowe, a więc nie są zgodne. Współczynnik konkordancji służy do pomiaru zgodności ekspertów dla zmiennych mierzonych na skali po-

rządkowej, gdzie eksperci przyporządkowują rangi poszczególnym czynnikom. Taka sytuacja ma miejsce podczas oceny wpływu czynników/zdarzeń na rozwój danego zjawiska, czy też oceny barier lub stymulatorów.

Jeśli eksperci są wystarczająco zgodni w swoich opiniach, to na podstawie ich odpowiedzi można skonstruować prognozy wraz z miernikami ich dopuszczalności. Jeśli jednak nie uzyskano wystarczająco zgodnych opinii, należy przygotować następną ankietę i wraz z wynikami poprzedniej tury rozesłać do ekspertów [Dittmann, 2003, s. 168].

Oczywiście nie zawsze wymagana jest zgodność ekspertów. W przypadku burzy mózgów wymagane jest wręcz twórcze myślenie. Pomysł, który na początku wydawał się absurdalny, po dopracowaniu i późniejszej ocenie może się okazać bardzo korzystny.

#### 4. Analiza własności miar oceny zgodności opinii ekspertów

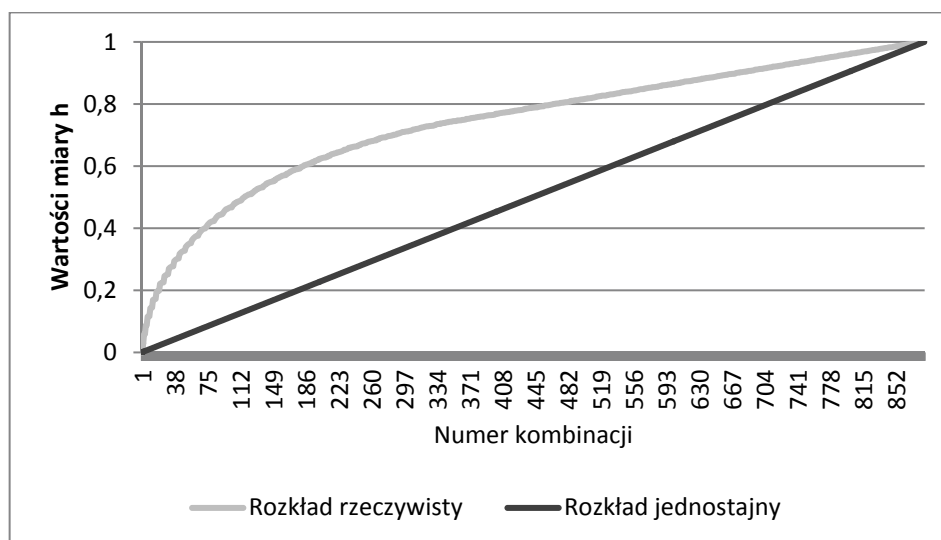
Ponieważ w badaniach opartych na sądach ekspertów w głównej mierze korzysta się ze skal nominalnych lub porządkowych, analizę własności miar oceny zgodności opinii ekspertów przeprowadzono w stosunku do współczynnika dyspersji (miara 2) oraz współczynnika konkordancji (miara 3). Współczynnik dyspersji jest bardzo wrażliwy na zmiany w rozkładzie odpowiedzi [por. z tabelą 1].

**Tabela 1.** Wartości współczynnika dyspersji (miara h) a rozkład odpowiedzi ekspertów

Kategoria	P1	P2	P3
1	0	0	0
2	0	0	0
3	0	0	0
4	0	1	2
5	10	9	8
h	0,000	0,225	0,400

W tabeli 1 przedstawiono rozkład wyników 10 ekspertów dla pojedynczego pytania z pięcioma kategoriami odpowiedzi. W przypadku pierwszym (P1) wszyscy eksperci zgodnie wskazali piątą kategorię, a więc wartość miary h wyniosła zero. W drugim przypadku jeden z ekspertów wskazał kategorię czwartą (pozostali piątą), a wartość miary wzrosła do 0,225, a więc ponad 1/5 jej rozpiętości. Jeśli dwóch ekspertów wskazało kategorię czwartą, a pozostali piątą, to wartość miary h wyniosła 0,4. Na podstawie przedstawionych wyników można

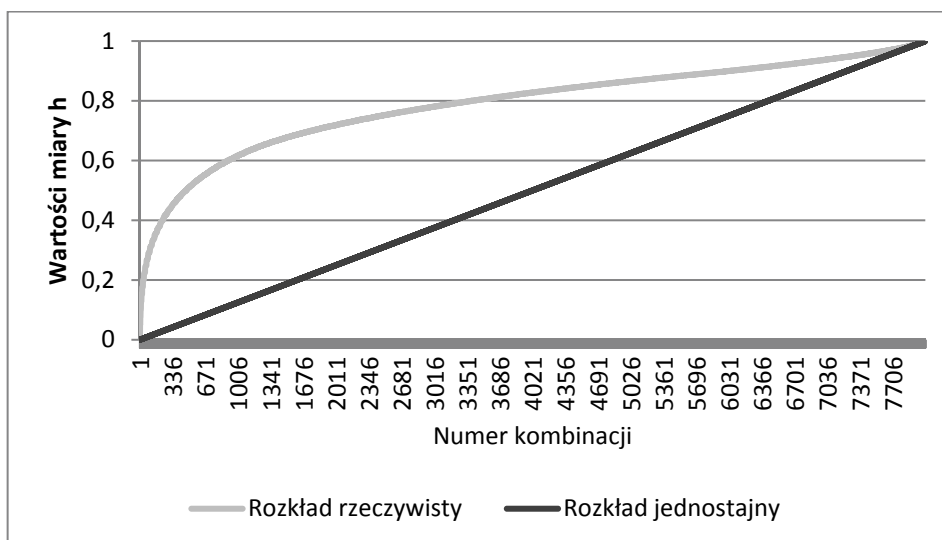
wysnuć wniosek, iż miara  $h$  jest bardzo wrażliwa nawet na niewielkie różnice w odpowiedziach ekspertów. W związku z tym sprawdzono rozkład wartości tej miary dla trzech, czterech i pięciu kategorii odpowiedzi. W celu uniknięcia wpływu liczby ekspertów na rozkład wyników posłużono się odsetkiem ekspertów, którzy wskazali daną kategorię odpowiedzi<sup>1</sup>. Rozpatrzono wszystkie możliwe kombinacje układu wyników odpowiedzi ekspertów z dokładnością do jednego procenta. W przypadku pytań z trzema możliwymi kategoriami odpowiedzi możliwych kombinacji jest 884, dla czterech kategorii – 8037, a dla pięciu – 46 262. Wyniki przedstawiono na rys. 2, 3 i 4.



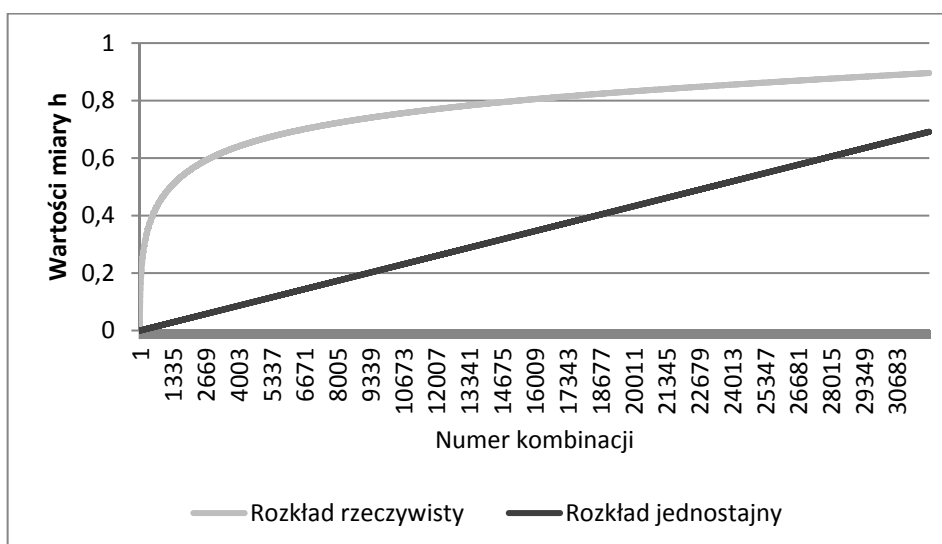
**Rys. 2.** Rozkład wyników współczynnika dyspersji dla trzech kategorii odpowiedzi – skala nominalna

<sup>1</sup> Dla odpowiedzi z tabeli 1 dla piątej kategorii odsetek ten wyniósłby dla P1 – 1; P2 – 0,9 i 0,8 dla P3.





**Rys. 3.** Rozkład wyników współczynnika dyspersji dla czterech kategorii odpowiedzi – skala nominalna



**Rys. 4.** Rozkład wyników współczynnika dyspersji dla pięciu kategorii odpowiedzi – skala nominalna

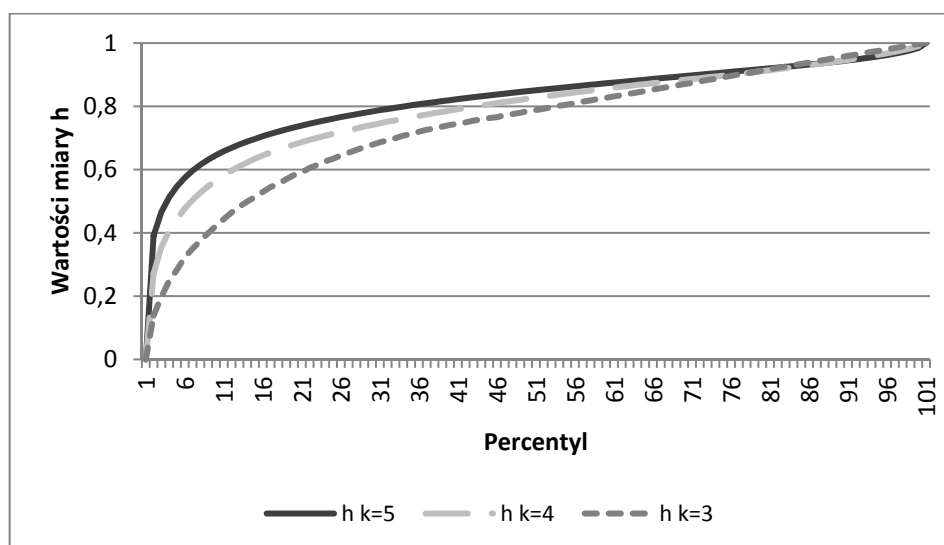
Na rys. 2-4 można zauważyć, że empiryczny rozkład odpowiedzi znacznie odchyła się od rozkładu jednostajnego. Oznacza to, że niewielkie różnice w odpowiedziach ekspertów powodują szybki wzrost wartości miary  $h$ , który następ-

nie charakteryzuje się zmniejszającą dynamiką. Można zauważyć, że początkowa dynamika wzrostu wartości współczynnika dyspersji rośnie wraz ze wzrostem liczby kategorii odpowiedzi. Stąd pytanie, przy jakiej wartości miary  $h$  można uznać wyniki ekspertów za zgodne? Twierdzenie, iż wartości bliskie zeru świadczą o zgodności opinii ekspertów jest niewystarczające, bo czy wartość miary  $h = 0,3$  można uznać za zgodną opinię?

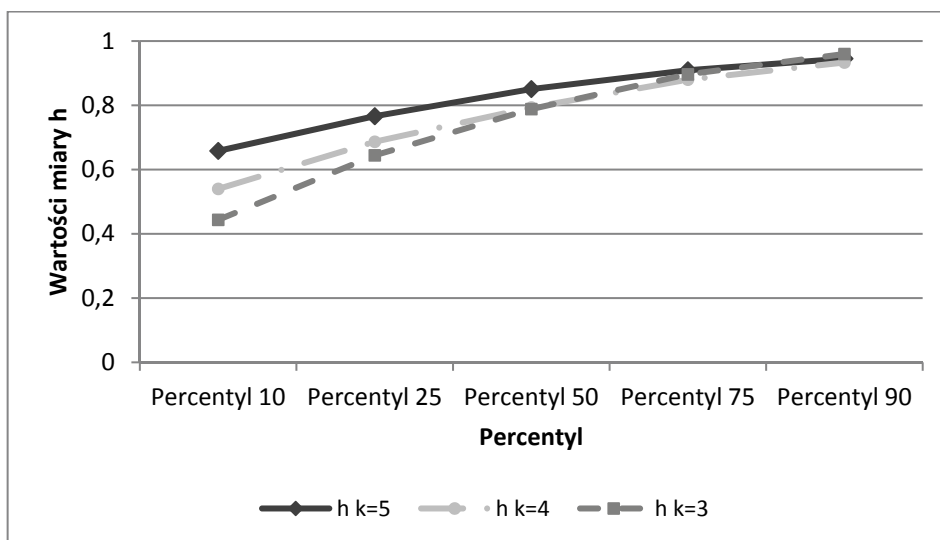
W związku z tym dylematem sprawdzono wartości progowe współczynnika dyspersji, obliczając percentyle wyników [por. tabelę 2 oraz rys. 5 i 6].

**Tabela 2.** Percentyle rozkładu wyników współczynnika dyspersji – skala nominalna

	$h k = 5$	$h k = 4$	$h k = 3$
Percentyl 10	0,658	0,540	0,444
Percentyl 25	0,766	0,687	0,644
Percentyl 50	0,851	0,793	0,788
Percentyl 75	0,909	0,880	0,896
Percentyl 90	0,946	0,933	0,960



**Rys. 5.** Percentyle rozkładu wyników współczynnika dyspersji – skala nominalna



**Rys. 6.** Główne percentyle rozkładu wyników współczynnika dyspersji – skala nominalna

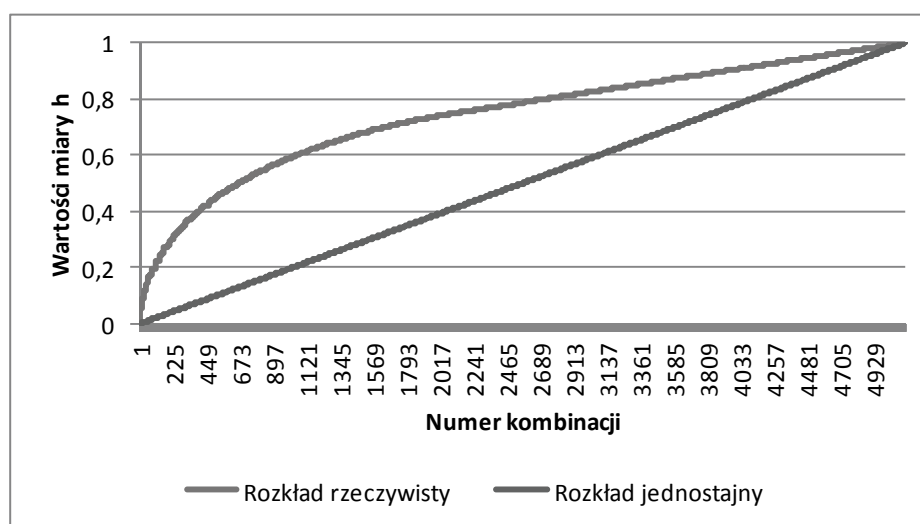
Na potrzeby badań przyjęto, że wartość miary poniżej wartości progowej na poziomie percentyla 10 będzie oznaczała, iż eksperci są zgodni (w wersji bardziej rygorystycznej). Oznacza to, że empiryczny układ odpowiedzi ekspertów jest poniżej 10% najbardziej zgodnych odpowiedzi spośród wszystkich możliwych kombinacji wyników<sup>2</sup>. W przypadku trzech kategorii percentyl 10 wyniósł 0,444, ale w przypadku pięciu kategorii już 0,658. Stąd wniosek, że w przypadku trzech kategorii odpowiedzi sądy ekspertów można uznać za zgodne, jeśli wartość współczynnika dyspersji będzie równa lub niższa od 0,444; w przypadku czterech kategorii – 0,540, a dysponując pięcioma kategoriami odpowiedzi – 0,658. Oczywiście można przyjąć mniej rygorystyczną wersję wartości progowej na poziomie percentyla 25. Ogólnie można sklasyfikować zgodność ekspertów według schematu:

- wartości miary niższe niż percentyl 10 – wysoka zgodność ekspertów,
- wartości miary mieszczące się w przedziale ⟨percentyl 10; percentyl 25⟩ – umiarkowana zgodność ekspertów,
- wartości miary mieszczące się w przedziale ⟨percentyl 25; percentyl 50⟩ – niska zgodność ekspertów,
- wartości miary większe niż percentyl 50 – brak zgodności ekspertów.

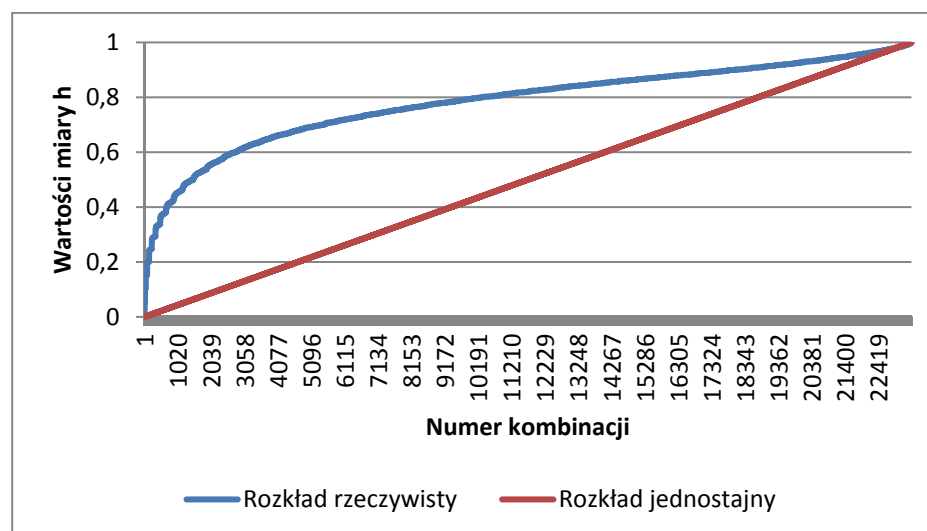
Powyższa analiza dotyczy odpowiedzi ekspertów mierzonych na skali nominalnej, a więc poszczególnych kategorii nie można uporządkować. W przy-

<sup>2</sup> W przypadku, gdyby wyniki miały rozkład jednostajny, odpowiadałoby to wartości miary  $h = 0,1$ .

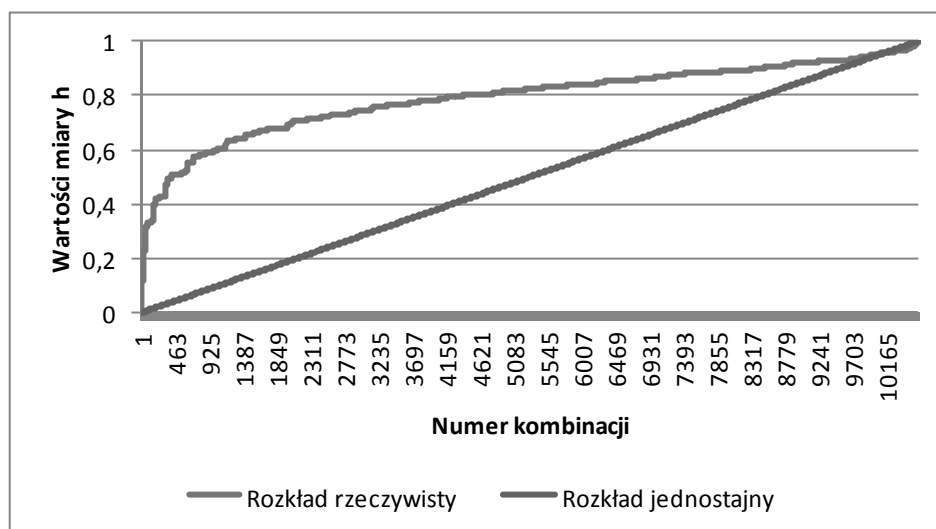
padku, gdy odpowiedzi są mierzone na skali porządkowej, np. skali Likerta, wartości progowe ulegną zmianie ze względu na zwiększenie liczby możliwych kombinacji wyników. W związku z tym analizę przeprowadzono, rozróżniając kolejność kategorii, a wyniki przedstawiono na rys. 7, 8 i 9. Na rys. 10 i 11 oraz w tabeli 3 przedstawiono obliczone percentyle – wartości progowe.



Rys. 7. Rozkład wyników współczynnika dyspersji dla trzech kategorii odpowiedzi – skala porządkowa



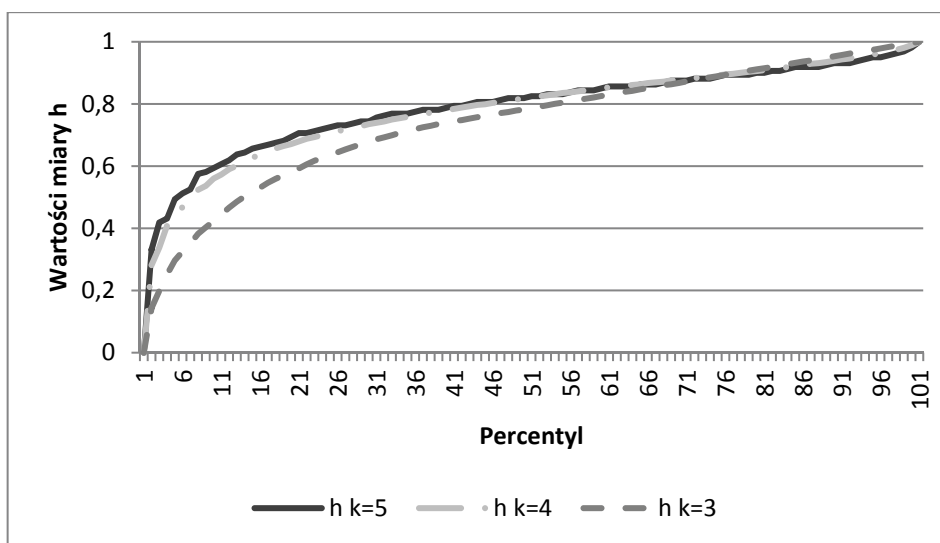
Rys. 8. Rozkład wyników współczynnika dyspersji dla czterech kategorii odpowiedzi – skala porządkowa



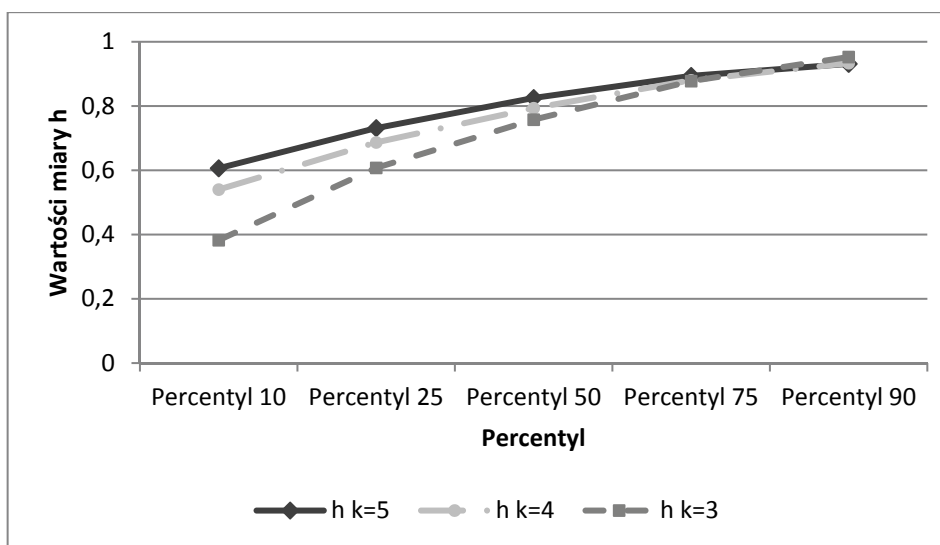
**Rys. 9.** Rozkład wyników współczynnika dyspersji dla pięciu kategorii odpowiedzi – skala porządkowa

**Tabela 3.** Percentyle rozkładu wyników współczynnika dyspersji – skala porządkowa

	h k = 5	h k = 4	h k = 3
Percentyl 10	0,606	0,573	0,445
Percentyl 25	0,731	0,713	0,644
Percentyl 50	0,825	0,820	0,787
Percentyl 75	0,894	0,895	0,894
Percentyl 90	0,931	0,943	0,957



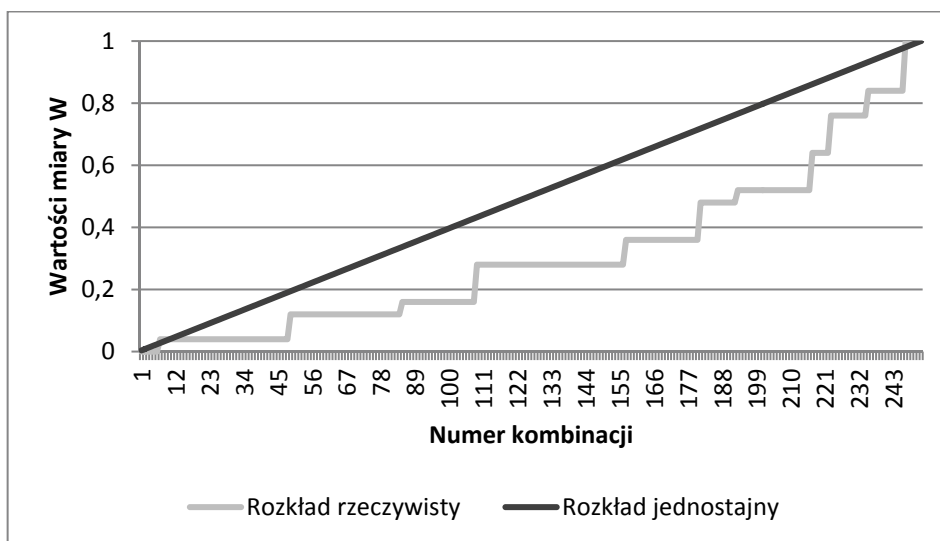
Rys. 10. Percentyle rozkładu wyników współczynnika dyspersji – skala porządkowa



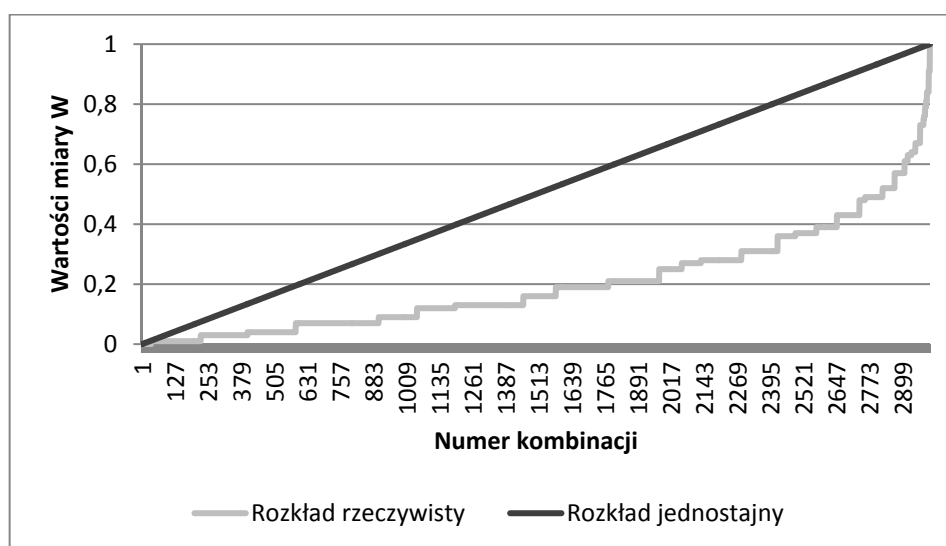
Rys. 11. Główne percentyle rozkładu wyników współczynnika dyspersji – skala porządkowa

Na rys. 7-9, podobnie jak w przypadku skali nominalnej, można zauważyć, że empiryczny rozkład odpowiedzi znacznie odchyła się od rozkładu jednostajnego. W przypadku trzech kategorii odpowiedzi wartości percentyli dla skali porządkowej są zbliżone do wartości percentyli dla skali nominalnej. Dla czterech kategorii odpowiednie percentyle są wyższe dla skali porządkowej, np. percentyl 10 dla skali porządkowej wyniósł 0,573, a dla skali nominalnej – 0,540 [por. tabele 2 i 3]. W przypadku pięciu kategorii odpowiedzi odpowiednie percentyle są niższe niż dla skali nominalnej (percentyl 10 wyniósł odpowiednio 0,606 i 0,658).

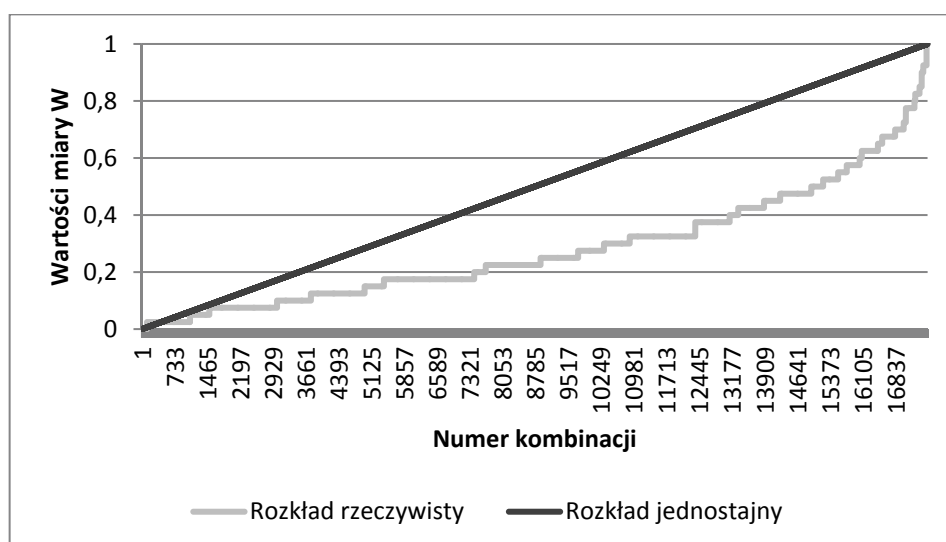
Podobną analizę przeprowadzono dla współczynnika konkordancji (3). Wyniki przedstawiono na rys. 12-14.



**Rys. 12.** Rozkład wyników współczynnika konkordancji dla trzech kategorii odpowiedzi i pięciu ekspertów



**Rys. 13.** Rozkład wyników współczynnika konkordancji dla trzech kategorii odpowiedzi i dziesięciu ekspertów

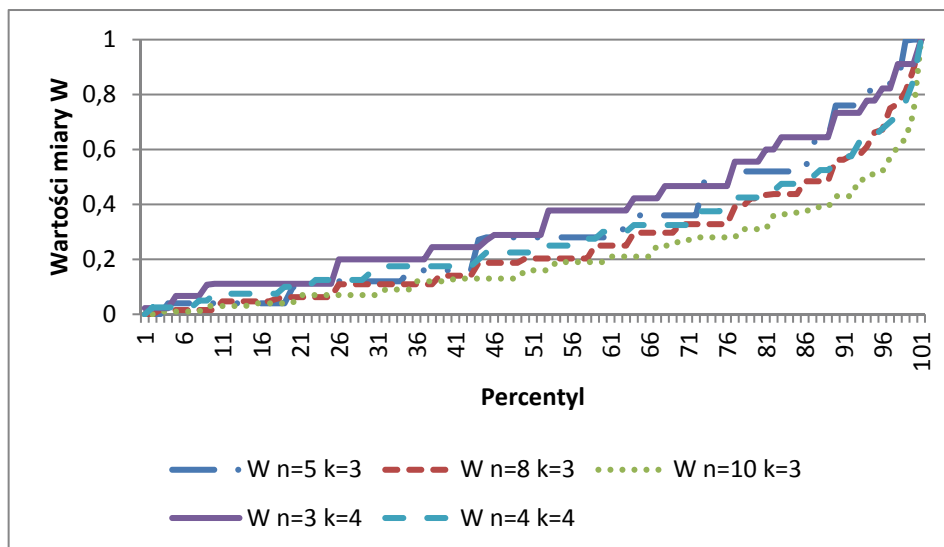


**Rys. 14.** Rozkład wyników współczynnika konkordancji dla czterech kategorii odpowiedzi i czterech ekspertów

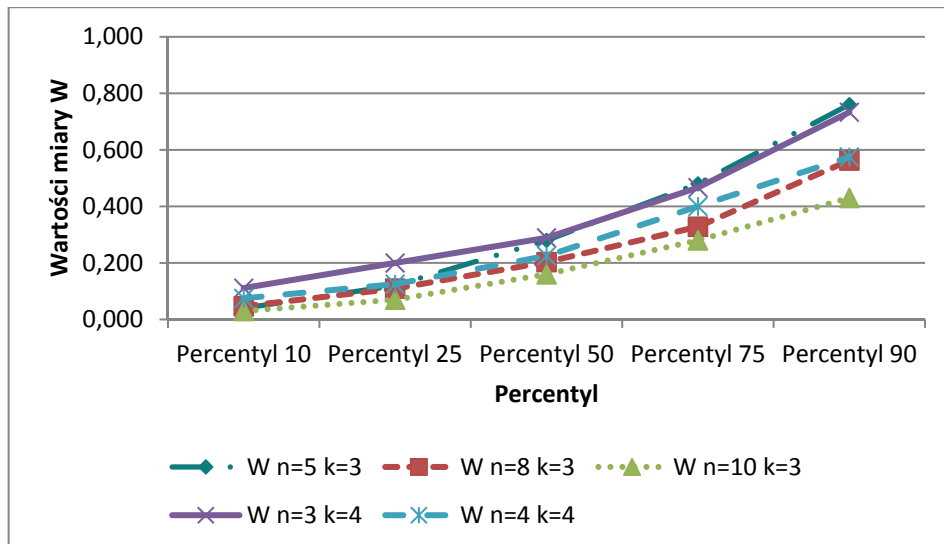
Podobnie jak dla współczynnika dyspersji otrzymane wyniki nie mają rozkładu jednostajnego. W pobliżu zera (brak zgodności ekspertów) wartości miary W rosną wolniej niż w pobliżu jedności. Oznacza to, iż podobnie jak dla współ-



czynnika dyspersji większość możliwych wyników wskazuje raczej na brak zgodności ekspertów niż na ich zgodność. Na rys. 15 i 16 oraz w tabeli 4 przedstawiono obliczone percentyle, które mogą stanowić wartości progowe przy ocenie zgodności ekspertów.



Rys. 15. Percentyle rozkładu wyników współczynnika konkordancji



Rys. 16. Główne percentyle rozkładu wyników współczynnika konkordancji

**Tabela 4.** Percentyle rozkładu wyników współczynnika konkordancji

	$W_{n=5 k=3}$	$W_{n=8 k=3}$	$W_{n=10 k=3}$	$W_{n=3 k=4}$	$W_{n=4 k=4}$
Percentyl 10	0,040	0,047	0,030	0,111	0,075
Percentyl 25	0,120	0,109	0,070	0,200	0,125
Percentyl 50	0,280	0,203	0,160	0,289	0,225
Percentyl 75	0,480	0,328	0,280	0,467	0,400
Percentyl 90	0,760	0,563	0,430	0,733	0,575

Wartość miary powyżej wartości progowej na poziomie percentyla 90 będzie oznaczała, iż eksperci są zgodni. Oznacza to, że empiryczny układ odpowiedzi ekspertów jest powyżej 10% najbardziej zgodnych odpowiedzi spośród wszystkich możliwych kombinacji wyników. W przypadku trzech kategorii percentyl 90 dla pięciu ekspertów wyniósł 0,760, a dla dziesięciu 0,43. Oznacza to, że wraz ze wzrostem liczby ekspertów maleje wartość progowa. Ogólnie można sklasyfikować zgodność ekspertów według schematu:

- wartości miary wyższe niż percentyl 90 – wysoka zgodność ekspertów,
- wartości miary mieszczące się w przedziale ⟨percentyl 75; percentyl 90⟩ – umiarkowana zgodność ekspertów,
- wartości miary mieszczące się w przedziale ⟨percentyl 50; percentyl 75⟩ – niska zgodność ekspertów,
- wartości miary niższe niż percentyl 50 – brak zgodności ekspertów.

## Podsumowanie

Przeprowadzona analiza wskazała, że możliwe wartości miar dyspersji i konkordancji odbiegają od rozkładu jednostajnego, co stwarza problemy interpretacyjne. Trudno w tej sytuacji określić, czy dana wartość miary, np.  $h = 0,2$ , wskazuje na zgodność ekspertów, czy też nie. W związku z tym należy wyznaczyć wartości progowe tych miar za pomocą statystyk pozycyjnych. Wartości te są zależne zarówno od liczby kategorii odpowiedzi i skali pomiaru zmiennej (współczynnik dyspersji), jak też od liczby kategorii i liczby ekspertów (współczynnik konkordancji). Można zauważyć, że wzrost liczby klas odpowiedzi zwiększa asymetrię możliwych wyników. Analogiczna sytuacja występuje w przypadku współczynnika konkordancji – wraz ze wzrostem liczby ekspertów rośnie asymetria wyników. Dalsze prace nad miarami zgodności ekspertów powinny obejmować próbę modyfikacji omawianych miar w celu uzyskania bardziej jednostajnego rozkładu wyników. Umożliwiłoby to pominięcie wyznaczania wartości progowych przed każdym badaniem. W przypadku większej liczby klas odpowiedzi

czy ekspertów uczestniczących w badaniu nie jest możliwe wyznaczenie wszystkich możliwych wyników. W tym przypadku pomocne byłyby testy permutacyjne, które umożliwiłyby ocenę istotności współczynnika dyspersji i konkordancji.

## Literatura

- Cieślak M. (red.) (1997), *Prognozowanie gospodarcze – metody i zastosowania*, Wydawnictwo Naukowe PWN, Warszawa.
- Czaplicka-Kolarz K. (red.) (2007), *Scenariusze rozwoju technologicznego kompleksu paliwowo-energetycznego dla zapewnienia bezpieczeństwa energetycznego kraju – część 1. Studium gospodarki paliwami i energią dla celów opracowania foresightu energetycznego dla Polski na lata 2005-2030*, Główny Instytut Górnictwa, Katowice.
- Czaplicka-Kolarz K. (red.) (2007), *Scenariusze rozwoju technologicznego kompleksu paliwowo-energetycznego dla zapewnienia bezpieczeństwa energetycznego kraju – część 2. Scenariusze opracowane na podstawie foresightu energetycznego dla Polski na lata 2005-2030*, Główny Instytut Górnictwa, Katowice.
- Dittmann P. (2003), *Prognozowanie w przedsiębiorstwie – metody i ich zastosowanie*, Oficyna Ekonomiczna, Kraków.
- Gajda A. (2011), *Doświadczenia i metody pozyskania danych eksperckich na potrzeby badań z wykorzystaniem metody foresight* [w:] P. Dittmann, A. Szpulak (red.), *Prognozowanie w zarządzaniu firmą*, Prace Naukowe UE we Wrocławiu nr 185, Wrocław.
- Halal W.E., Kull M.D., Leffmann A. (1997), *Emerging Technologies. What's Ahead for 2001-2030*, „The Futurist”, Nov.-Dec.
- Kowalewska A., Głuszyński J. (red.) (2009), *Zastosowanie metody Delphi w Narodowym Programie Foresight Polska 2020 – Główne wyniki, doświadczenia i wnioski*, Pentor Research International, Warszawa.
- Kuwahara T. (2001), *Technology Foresight in Japan – The Potential and Implications of DELPHI Approach*. NISTEP Study Material 77, [www.nistep.go.jp](http://www.nistep.go.jp).
- Martin B.R. (1995), *Foresight in Science and Technology*, „Technology Analysis & Strategic Management”, No. 7(2).
- Poradowska K., Wójciak M. (2009), *Uogólniony rozkład trójkątny w analizie wyników badania foresight* [w:] *Dynamiczne modele ekonometryczne*, Acta Universitatis Nicolai Copernici, *Ekonomia* XXXIX, zeszyt specjalny, Wydawnictwo UMK, Toruń.
- Poradowska K., Wójciak M. (2011), *Stymulatory i bariery rozwoju gospodarki zeroemisyjnej w opinii społeczeństwa* [w:] *Modelowanie i prognozowanie gospodarki narodowej*, Prace i Materiały Wydziału Zarządzania Uniwersytetu Gdańskiego, zeszyt 4/8, Wydział Zarządzania Uniwersytetu Gdańskiego, Fundacja Rozwoju Uniwersytetu Gdańskiego, Sopot.
- Poradowska K., Szkutnik T., Wójciak M. (2011), *Scenariusze rozwoju wybranych technologii oszczędności energii w życiu codziennym* [w:] P. Dittmann, A. Szpulak (red.), *Prognozowanie w zarządzaniu firmą*, Prace Naukowe UE we Wrocławiu nr 185, Wrocław.

---

**METHODS USED WHILE ASSESSING COHERENCE  
OF EXPERTS' OPINIONS IN FORESIGHT**

**Summary:** Foresight primary objective is to construct scenarios of situation development in a relatively remote perspective (usually 20-25 years). In order to elaborate such scenarios, foresight applies numerous methods including highly appreciated activities that involve acquisition of expertise. The activities in question predominantly refer to expert panels, brainstorming and Delphi method. To formulate unambiguous judgements, it is necessary to obtain some conformity of experts' responses. In case of variables that are interval or ratio scale measured, location or dispersion measures are calculated. In case of nominal and ordinal scales, the following coefficients are, inter alia, applied: dispersion, variance, pairwise rank correlation or concordance.

In the research, tested measures of coherence of experts' opinions are presented along with their properties. This should allow for setting border values of the measures in question, which in turn should lead to recognition of experts' coherence. In the second part of the research some suggestions of the way the applied measures should be modified are presented.

**Keywords:** foresight, coherence of experts' opinions, heuristic forecasting.