ZESZYTY NAUKOWE UNIWERSYTETU SZCZECIŃSKIEGO

NR 643 STUDIA INFORMATICA NR 27 2011

ANDRZEJ PIEGAT
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

FUNKCJA $R_1(X)$ REPREZENTATYWNOŚCI HIPOTEZ I JEJ ZASTOSOWANIE DO AGREGACJI OCEN EKSPERCKICH

Wprowadzenie

W systemach technicznych do pomiarów różnych wielkości jednowymiarowych, takich jak napięcie, natężenie prądu, szybkość samochodu *etc*., stosowane są techniczne urządzenia pomiarowe, takie jak mierniki elektryczne, szybkościomierze, wagi, ciśnieniomierze itd. Pomiar wielkości występujących w systemach ekonomicznych, takich jak wartość firmy, posiadłości, budynku, atrakcyjności oferty złożonej w przetargu itd., jest zadaniem o wiele trudniejszym. Są to bowiem wielkości nie jedno-, lecz wielowymiarowe. Przykładowo, wartość firmy zależy od jej zysków w ostatnim okresie, od wartości terenów, na których firma leży, od jakości i liczności kadry inżynierskiej, zarządzającej i robotniczej, od stanu zadłużenia, od liczności, wieku i nowoczesności maszyn i urządzeń należących do firmy, od aktualnego portfela zamówień produkcyjnych, od atrakcyjności branży, w której firma działa, i innych atrybutów. Jednak nie ma urządzeń technicznych, które byłyby w stanie pomierzyć tak złożone wielkości wielowymiarowe. Mogą tego dokonać jedynie eksperci. Eksperci jednak, podobnie jak i mierniki techniczne, także mniej lub bardziej różnią się w swych „pomiarach", czyli ocenach. Stąd istnieje konieczność angażowania do oceny zmiennych wielu ekspertów, a następnie dokonywania agregacji (podsumowywania) ich ocen po to, aby uzyskać jedną łączną ocenę wynikową, na której podstawie można by podjąć rozsądna decyzję ekonomiczną.

Agregacja ocen eksperckich jest zadaniem bardzo trudnym. Świadczy o tym fakt, że badania w tym zakresie prowadzone są od kilkudziesięciu co najmniej lat, opracowano dużo metod agregacji, a badania nie ustają, co świadczy o ciągłym naukowym niezadowoleniu badaczy. Obecnie trudno jest powiedzieć, która z istniejących metod agregacji jest obiektywnie najlepsza. Przyczyną tego stanu rzeczy jest brak jednolitego stanowiska świata naukowego w kwestii modelowania niepewności (oceny eksperckie są niepewne). Do modelowania niepewności, z którą w życiu codziennym i w działalności firm mamy nieustannie do czynienia, stosuje się różne podejścia. Najbardziej znane jest podejście probabilistyczne[1], następnie oparte na teorii wiary i niewykluczalności (*belief and plausibility*) Dempstera-Shafera[2], posybilistyczne – oparte na pojęciu możliwości i konieczności (*possibility and necessity*) Duboisa i Prade[3] – i inne. Metoda przedstawiona w niniejszym artykule oparta jest na podejściu probabilistycznym, które jest chyba najbardziej rozbudowaną i przebadaną teorią naukową w zakresie niepewności zdarzeń. Również w ramach teorii prawdopodobieństwa opracowano znaczną liczbę metod agregacji ocen eksperckich. Przegląd tych metod został przedstawiony w książce A. O'Hagana i C. Bucka[4]. Autorzy tej pracy, renomowani specjaliści z zakresu teorii prawdopodobieństwa, przeanalizowali w ramach dużego grantu rządowego wszystkie najbardziej znane i stosowane metody probabilistyczne, takie jak metody Bayesowskie, liniowa, ważona agregacja ocen (*opinion pooling*), agregacja behawioralna, agregacja logarytmiczna, metoda Cooka (*seeding variables*) i inne. Jedną z ważnych miar rzeczywistej wartości metod agregacji jest częstość ich praktycznego stosowania. Badania wykazały,

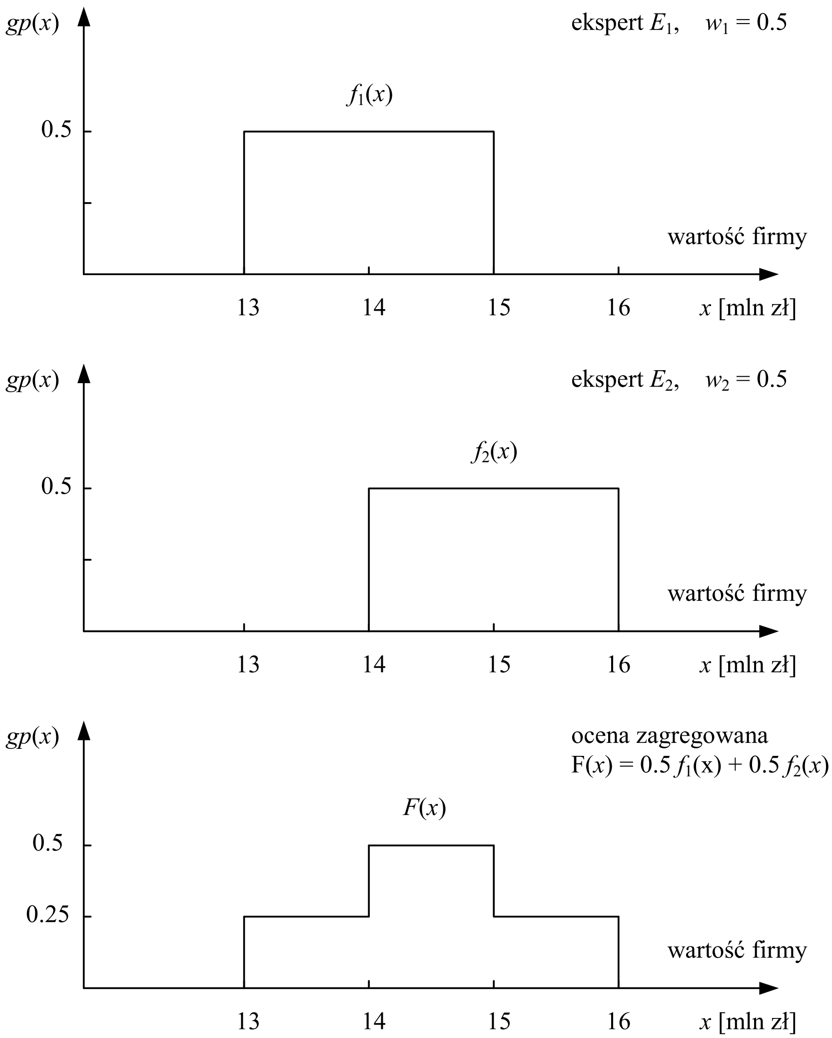
[1] R. Clemen, R. Winkler, *Combining probability distributions from experts in risk analysis*, "Risk Analysis", vol. 19, no. 2, s. 187–203; A. O'Hagan, C.A. Buck i in., *Uncertain judgements-eliciting experts' probabilities*, John Wiley & Sons LTD, Chichester, England 2006.

[2] G. Shafer, *A mathematical theory of evidence*, Princeton, NJ, Princeton University Press 1976.

[3] S. Destercke, D. Dubois, E. Chojnacki, *Possibilistic information fusion using maximal coherent subsets*, "IEEE Transactions on Fuzzy Systems" 2009, vol. 17, no. 1, s. 79–92; D. Dubois, H. Prade, R. Yager, *Merging fuzzy information*, w: *Fuzzy sets in approximate reasoning and information systems*, Kluwer, Boston, MA 1999, s. 335–401; S. Sandri, D. Dubois, H. Kalfsbeek, *Elicitation assesment and pooling of expert judgments using possibility theory*, "IEEE Transactions on Fuzzy Systems" 1995, vol. 3, no. 3, s. 313–335.

[4] A. O'Hagan, C.A. Buck i in., *Uncertain judgements-eliciting experts' probabilities*...

że metodą najczęściej stosowaną i zweryfikowaną w praktyce jest metoda liniowej agregacji ważonej (*linear opinion pooling*)[5].



Rys. 1. Przykład użycia najczęściej w praktyce stosowanej metody agregacji liniowej do zagregowania dwóch spójnych opinii eksperckich, $gp(x)$ – gęstość prawdopodobieństwa ocenianej zmiennej X

Źródło: opracowanie własne.

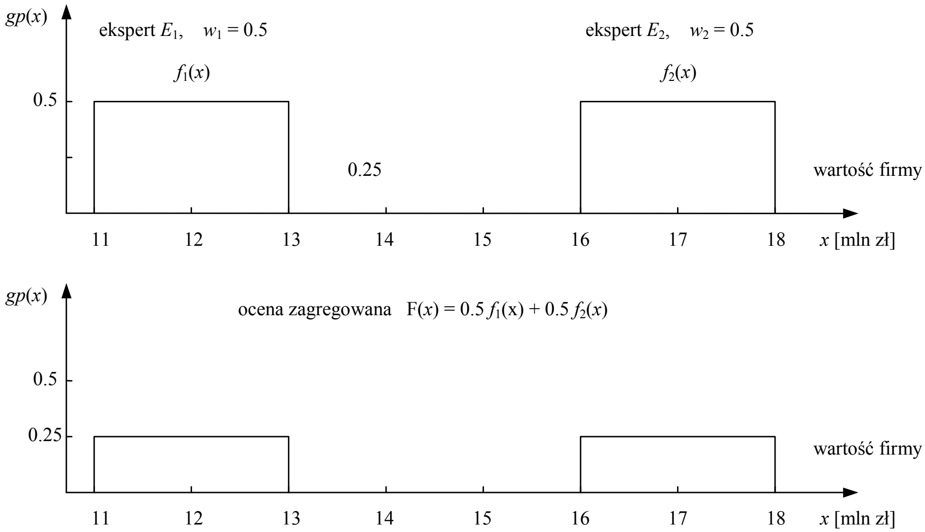
[5] Tamże.

Załóżmy, że danych jest n eksperckich ocen $f_i(x)$ wartości x niepewnej zmiennej X (na przykład wartości firmy), gdzie $f_i(x)$ mają formę rozkładów gęstości prawdopodobieństwa. Wówczas zagregowaną ocenę łączną $F(x)$ określa się według wzoru (1):

$$F(x) = \sum_{i=1}^{n} w_i f_i(x) \tag{1}$$

gdzie w_i oznacza współczynnik zaufania (współczynnik jakości) eksperta.

Zwykle zakłada się, że $\sum_{i=1}^{n} w_i = 1$. Na rysunku 1 przedstawiono przykład zastosowania tej metody do agregacji dwóch różnych ocen wartości firmy, uzyskanych od dwóch ekspertów E_1 i E_2. Oceny te mają formę równomiernych rozkładów gęstości prawdopodobieństwa. Eksperci nieco różnią się w swych opiniach, jednak istnieje między nimi pewien zakres konsensusu (zgodności). Jak pokazano to na rysunku 1, metoda agregacji liniowej daje bardzo wiarygodny i przekonujący wynik, stąd jej popularność. Wiarygodność ta jest jednak wysoka tylko wówczas, gdy eksperckie opinie $f_i(x)$ są spójne. Sytuacja zmienia się natomiast, gdy mamy do czynienia z ocenami niespójnymi (rysunek 2).



Rys. 2. Przykład agregacji dwóch niespójnych eksperckich ocen wartości firmy najbardziej popularną metodą liniowej agregacji ważonej

Źródło: opracowanie własne.

Uzyskany wynik agregacji przedstawiony na rysunku 2 trudno zaakceptować. Sugeruje on, że wartość firmy możliwa jest tylko w zakresach $x \in [11, 13]$ oraz $x \in [16, 18]$. Wartość firmy leżąca w zakresie pośrednim $x \in [13, 16]$ jest natomiast nieprawdopodobna. Jest to oczywiście niezgodne ze zdrowym rozsądkiem. Przykład ten pokazuje, że metoda agregacji liniowej nie obejmuje swym zakresem pewnych ważnych przypadków występujących w praktyce. Z tego względu nie jest ona metodą generalną i istnieje potrzeba opracowania metody bardziej ogólnej, która by umożliwiła agregację zarówno spójnych, jak i niespójnych opinii eksperckich. Przykład takiej metody opublikowano w artykule D. Duboisa, H. Prade'a i R. Yagera[6]. Jest to jednak metoda oparta na posybilistycznej teorii niepewności. W niniejszym artykule przedstawiona natomiast zostanie teoria oparta na rachunku prawdopodobieństwa, będącego najstarszą i najbardziej ugruntowaną teorią badania niepewności.

Oceny eksperckie $f_i(x)$ mogą w ogólnym przypadku być ocenami punktowymi lub mieć formę rozkładów gęstości prawdopodobieństwa. W rozdziale 2 przedstawiona zostanie ze względu na dopuszczalny wolumen artykułu tylko wersja metody realizująca agregację ocen punktowych, które są zawsze ocenami niespójnymi, poza przypadkiem idealnego pokrywania się tych ocen. Inne wersje proponowanej metody zostaną przedstawione w kolejnych publikacjach autora.

1. Agregacja punktowych ocen eksperckich metodą optymalnej reprezentacji

Ze względu na konieczne ograniczenie objętości artykułu metoda zostanie opisana skrótowo. Składa się ona z czterech kroków A, B, C, D.

A. Określenie oddzielnych rozkładów $S_{1Ei}(x)$ bezwzględnych błędów reprezentacji x względem poszczególnych eksperckich ocen x^*_{Ei}.

B. Zsumowanie wszystkich oddzielnych rozkładów $S_{1Ei}(x)$ błędów bezwzględnych w jeden łączny, ważony rozkład $S_1(x)$ według wzoru (2).

$$S_1(x) = \sum_{i=1}^{n} S_{1Ei}(x) \tag{2}$$

[6] D. Dubois, H. Prade, R. Yager, *Merging fuzzy information...*, s. 335–401.

C. Określenie rozkładu $r_1(x)$ dokładności reprezentowania ocen eksperckich x_i^* przez wszystkie możliwe hipotezy ($X = x$) dotyczące ocenianej zmiennej – wzór (3).

$$r_1(x) = 1 - \frac{S_1(x)}{S_{1\max}} \tag{3}$$

gdzie $S_{1\max}$ oznacza maksymalny sumaryczny błąd reprezentacji występujący w zbiorze możliwych hipotez $x_{min} \leq x \leq x_{max}$.

D. Określenie rozkładu gęstości prawdopodobieństwa $gpr_1(x)$ zagregowanej oceny według wzoru (4).

$$gpr_1(x) = \frac{r_1(x)}{\int\limits_{x\min}^{x\max} r_1(x)\,dx} \tag{4}$$

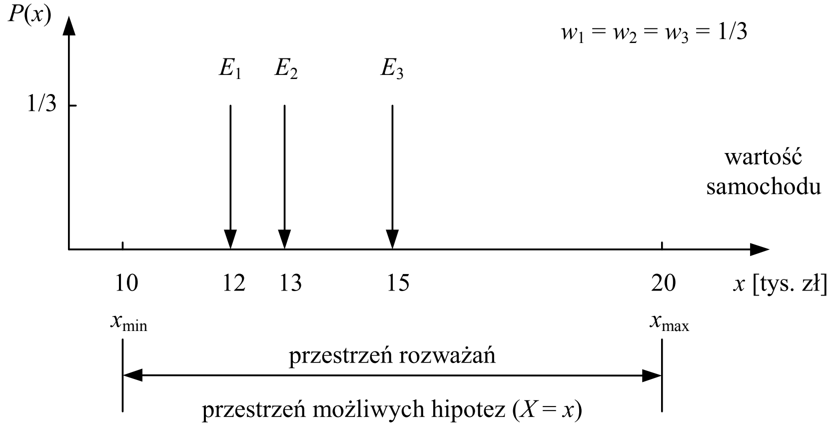
W następnym punkcie zostanie przedstawiony przykład praktycznego użycia metody.

2. Przykład zastosowania metody optymalnej reprezentacji do agregacji punktowych eksperckich ocen wartości samochodu

Pewna osoba zamierza sprzedać swój używany samochód. Z analizy ogłoszeń wynika, że samochody tej marki i wieku osiągają na rynku cenę w zakresie od 10 do 20 tys. zł ($10 \leq x \leq 20$) zależnie od swego przebiegu, stanu technicznego, stopnia zużycia, wyposażenia, koloru *etc*. Aby dokładniej poznać aktualną wartość swego samochodu, jego właściciel zlecił ocenę tej wartości trzem ekspertom. Eksperci – po dokładnym zapoznaniu się z samochodem i na podstawie swej głębokiej znajomości rynku samochodowego – podali punktowe, według nich najbardziej prawdopodobne, wartości samochodu:

E_1: $x_{E1}^* = 12$ tys. zł, E_2: $x_{E2}^* = 13$ tys. zł, E_3: $x_{E3}^* = 15$ tys. zł.

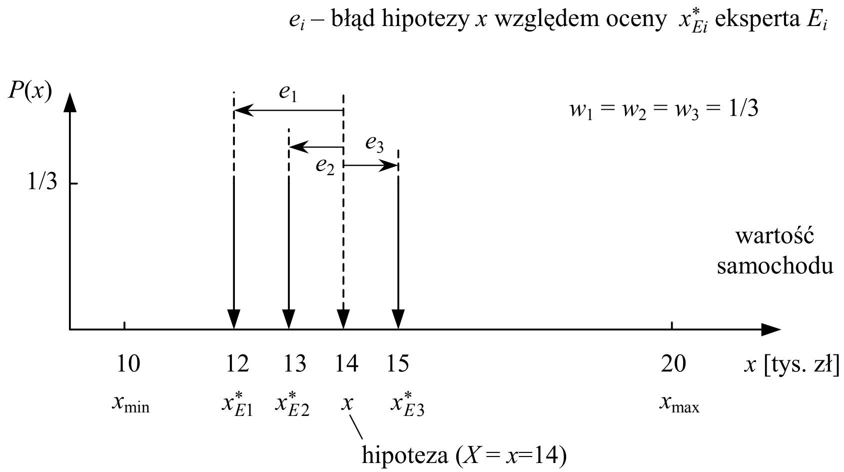
Oceny te przedstawiono graficznie na rysunku 3.



Rys. 3. Wizualizacja punktowych ocen wartości samochodu dostarczona przez trzech ekspertów o jednakowym współczynniku zaufania (jakości) $w_1 = w_2 = w_3 = 1/3$, P oznacza prawdopodobieństwo oceny, $10 \leq x \leq 20$ oznacza zakres możliwych wartości samochodu (przestrzeń rozważań, przestrzeń hipotez)

Źródło: opracowanie własne.

Oceny eksperckie oraz przestrzeń rozważań ceny jest to jedyna wiedza, jaką dysponujemy. Należy też zauważyć, że punktowe oceny ekspertów są przykładem całkowitej niespójności wiedzy. Oceny te nie mają żadnego wspólnego zakresu konsensusu, czyli zgodności. Z tego względu metoda agregacji liniowej nie może być tu zastosowana. W trakcie wyjaśniania metody stosowane będzie pojęcie możliwych hipotez ($X = x$), gdzie X oznacza nazwę zmiennej ocenianej, a x oznacza liczbową wartość samochodu w tys. zł. Pojęcie wartość samochodu może być też interpretowane jako cena, za jaką właścicielowi faktycznie uda się sprzedać samochód. Zastanówmy się teraz, jaka wartość x samochodu najlepiej reprezentuje wszystkie trzy oceny eksperckie. Przyjmijmy na przykład hipotezę, że wartość samochodu x = 14 tys. zł. Zauważmy, że hipoteza ta nie jest zgodna z żadną z opinii eksperckich (rysunek 4).



Rys. 4. Ilustracja pojęcia błędu hipotezy ($X = x$) względem opinii x^*_{Ei} pojedynczych ekspertów

Źródło: opracowanie własne.

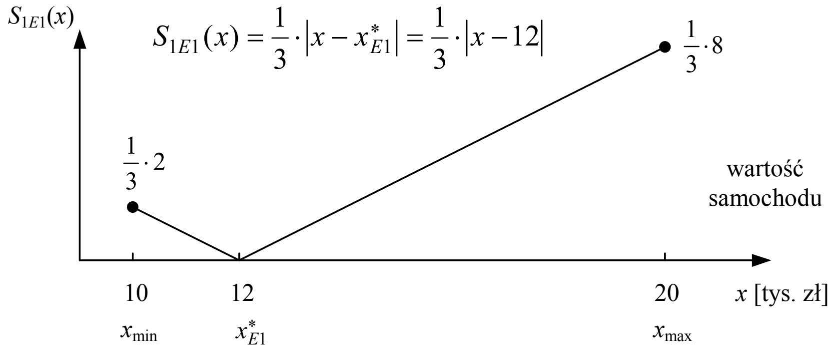
Hipoteza ($X = x = 14$) ma błąd (niezgodność) z każdą z opinii eksperckich. Zastosowane zostanie tutaj pojęcie błąd w sensie błędu bezwzględnego. Można oczywiście użyć innych matematycznych form błędu. Poniżej podane są wartości błędów hipotezy ($X = x = 14$) względem poszczególnych ocen eksperckich oraz ich ważona suma $S_1(x)$.

$$|e_1(x)| = 2, \qquad |e_2(x)| = 1, \qquad |e_3(x)| = 1, \qquad x = 14$$

$$S_1(x) = \sum_{i=1}^{3} w_i |e_i(x)| = \frac{1}{3} \cdot 2 + \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 1 = \frac{4}{3}.$$

Ważona suma błędów bezwzględnych $S_1(x)$ jest dobrą miarą niezgodności hipotezy z posiadanymi ocenami eksperckimi. Zastanówmy się teraz, czy w ogóle istnieje jakakolwiek hipotetyczna wartość x samochodu, która byłaby całkowicie zgodna ze wszystkimi ocenami jednocześnie. Jeden rzut oka na rysunek 4 uzmysławia, że hipoteza taka nie istnieje. W związku z tym jedynym rozsądnym postępowaniem może być znalezienie takiej wartości x samochodu, która najlepiej będzie reprezentować wszystkie oceny eksperckie. W tym celu należy określić rozkłady niezgodności $S_1(x)$ dla wszystkich możliwych hipotez wartości

x samochodu w ramach możliwego zakresu wartości $x_{min} \le x \le x_{max}$. Na rysunku 5 przedstawiono rozkład niezgodności $S_{1E1}(x)$ hipotez x względem jednej tylko oceny eksperckiej x^*_{E1} = 12 tys. zł.



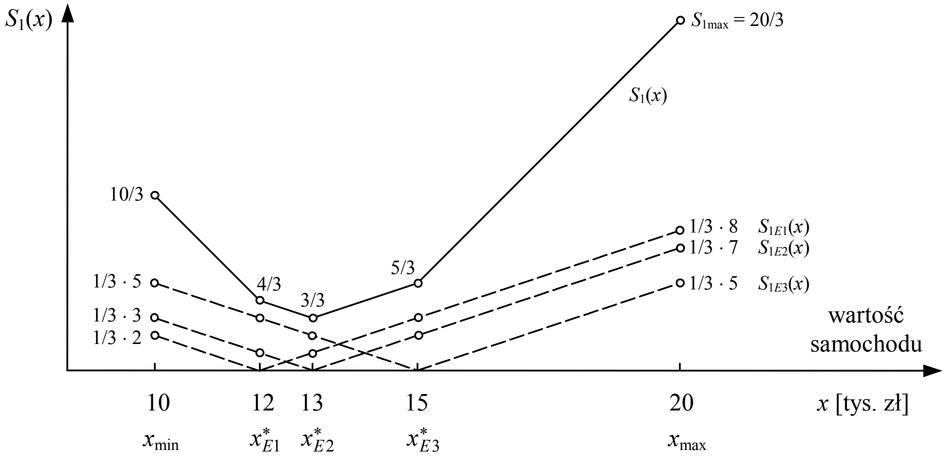
Rys. 5. Rozkład niezgodności $S_1(x)$ poszczególnych możliwych hipotez ($X = x$) wartości samochodu z oceną x^* = 12 tys. zł pierwszego eksperta E_1

Źródło: opracowanie własne.

Jak łatwo zauważyć na rysunku 5, w przypadku pojedynczej tylko oceny eksperckiej istnieje jedna jedyna hipoteza (x = 12), która jest całkowicie zgodna z opinią eksperta E_1. Pozostałe hipotezy są mniej lub bardziej niezgodne. Najwyższą niezgodność wykazuje hipoteza x = 20, dla której wartość kryterium niezgodności $S_{1E1}(x)$ wynosi 8/3. Na rysunku 6 przedstawiono trzy rozkłady $S_{1Ei}(x)$ ważonej niezgodności możliwych hipotez x z poszczególnymi ocenami eksperckimi oraz sumaryczny rozkład niezgodności $S_1(x)$.

Jak widać na rysunku 6, żadna możliwa hipotetyczna wartość x samochodu nie ma idealnej zgodności z ocenami wszystkich ekspertów jednocześnie. Najlepsza pod tym względem jest hipoteza x = 13 tys. zł, która ma najmniejszą niezgodność $S_1(x)$ = 3/3 = 1.

$$S_1(x) = \sum_{i=1}^{3} w_i S_{1Ei}(x) = \frac{1}{3} \cdot |x - 12| + \frac{1}{3} \cdot |x - 13| + \frac{1}{3} \cdot |x - 15|.$$



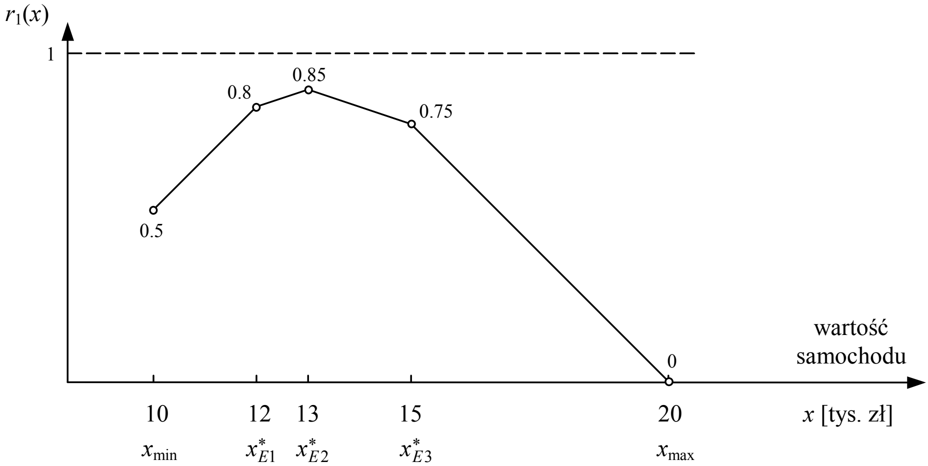
Rys. 6. Rozkłady niezgodności $S_{1E1}(x)$ możliwych hipotez ($X = x$) ceny samochodu z ocenami poszczególnych ekspertów x^*_{Ei} oraz łączny rozkład $S_1(x)$ bezwzględnej niezgodności hipotez z ocenami ekspertów E_i

Źródło: opracowanie własne.

W kolejnym kroku metody określony zostanie rozkład $r_1(x)$ zgodności możliwych hipotez ($X = x$) z ocenami wszystkich ekspertów, według wzoru (5).

$$r_1(x) = 1 - \frac{S_1(x)}{S_{1max}} \tag{5}$$

W rozpatrywanym przykładzie $S_{1max} = 20/3$ (rysunek 6). Na rysunku 7 przedstawiono rozkład $r_1(x)$ zgodności hipotez ($X = x$) z ocenami x^*_{E1} ekspertów.



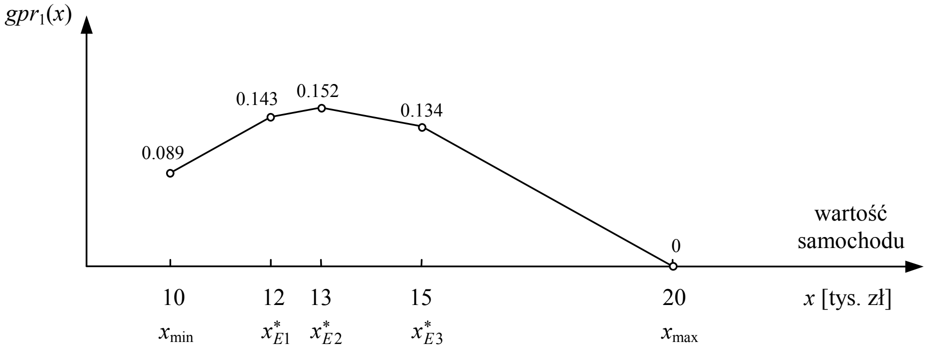
Rys. 7. Funkcja dokładności reprezentowania ocen eksperckich przez poszczególne możliwe hipotezy ($X = x$), $x_{min} \le x \le x_{max}$ lub inaczej – funkcja zgodności hipotez z ocenami eksperckimi

Źródło: opracowanie własne.

Funkcja $r_1(x)$ dokładności reprezentacyjnej przedstawiona na rysunku 7 informuje nas, jak dokładnie poszczególne hipotezy ($X = x$) wartości samochodu reprezentują trzy eksperckie oceny dotyczące tej wartości. Żadna z możliwych wartości x nie reprezentuje tych ocen w stopniu idealnym $r_1(x) = 1$, gdyż jest to niemożliwe. Jednak hipoteza $X = x = 13$ tys. zł osiąga najwyższą dokładność reprezentacyjną, gdyż jest najbardziej zgodna ($r_1(x) = 0{,}85$) ze wszystkimi trzema ocenami ekspertów. Wartość $x = 13$ tys. zł jest więc najlepszym reprezentantem posiadanej wiedzy o wartości sprzedawanego samochodu. Wartość ta może być wykorzystana na przykład do określenia wywoławczej ceny w ofercie sprzedaży samochodu. Na podstawie rozkładu dokładności reprezentacyjnej $r_1(x)$ można uzyskać rozkład gęstości prawdopodobieństwa $gpr_1(x)$ dokładności hipotezy ($X = x$) wartości samochodu, stosując wzór (6).

$$gpr_1(x) = \frac{r_1(x)}{\int_{x\min}^{x\max} r_1(x)\, dx} \tag{6}$$

Całka występująca w mianowniku wzoru (6) określa powierzchnię A rozkładu dokładności reprezentacji. W rozpatrywanym przypadku powierzchnia rozkładu z rysunku 7 wynosi A = 5,6. Na podstawie wzoru (6) uzyskuje się rozkład przedstawiony na rysunku 8.



Rys. 8. Rozkład gęstości prawdopodobieństwa wartości x samochodu wynikający z posiadanej wiedzy (trzy oceny eksperckie dotyczące tej wartości)

Źródło: opracowanie własne.

Rozkład z rysunku 8 informuje nas o szansach uzyskania poszczególnych cen za samochód. Najbardziej prawdopodobna jest cena w wysokości około 13 tys. zł. Bardziej prawdopodobne jest uzyskanie ceny w zakresie od 10 do 15 tys. zł niż w zakresie powyżej 15 tys. zł *etc*.

Wnioski

Najczęściej stosowane probabilistyczne metody agregacji ocen eksperckich, na które istnieje duże zapotrzebowanie w zakresie nauk ekonomicznych i zarządzania, dają wiarygodne wyniki tylko w przypadku spójnych, niesprzecznych ocen eksperckich. Nie potrafią one jednak wiarygodnie agregować ocen niespójnych, sprzecznych. W artykule przedstawiono metodę opartą na funkcji dokładności reprezentowania ocen eksperckich przez możliwe hipotezy ($X = x$) dotyczące ocenianej zmiennej. Metoda ta może być stosowana zarówno w przypadku spójnych, jak i niespójnych ocen eksperckich. Dostarcza ona wiarygodne wyniki, zgodne z ludzką intuicją i zdrowym rozsądkiem, i jest dość łatwa pod względem obliczeniowym. W niniejszej publikacji, ze względu na jej ograniczenie objętościowe, przedstawiono jedynie wersję metody opartą na kryterium sumarycznego błędu bezwzględnego $S_1(x)$ dla przypadku punktowych ocen eksperckich, które są najostrzejszym przypadkiem niespójności (oceny punktowe są zwykle niespójne). Metoda oparta na błędzie bezwzględnym ma dużą odporność na opinie eksperckie silnie różniące się od większości opinii eksperckich (tak zwanych outliersów). Cechy takiej nie mają natomiast metody agregacji, oparte na błędzie kwadratowym, w których wpływ opinii nietypowych jest niezwykle silny ze względu na stosowanie operacji podnoszenia błędu do kwadratu. Metoda oparta na błędzie bezwzględnym ma charakter medianowy, to znaczy maksimum funkcji zgodności $r_1(x)$ występuje zawsze albo dla opinii eksperta środkowego (w przypadku nieparzystej liczby ekspertów), albo też leży między opiniami dwóch ekspertów środkowych (w przypadku parzystej liczby ekspertów). W następnych publikacjach autora przedstawione zostaną wersje metody reprezentacyjnej dla niepunktowych ocen eksperckich w formie rozkładów gęstości prawdopodobieństwa) oraz dla innych kryteriów niezgodności.

Literatura

Clemen R., Winkler R., *Combining probability distributions from experts in risk analysis*, “Risk Analysis”, vol. 19, no. 2.

Destercke S., Dubois D., Chojnacki E., *Possibilistic information fusion using maximal coherent subsets*, IEEE Transactions on Fuzzy Systems 2009, vol. 17, no. 1.

Dubois D., Prade H., Yager R., *Merging fuzzy information*, w: *Fuzzy sets in approximate reasoning and information systems*, Kluwer, Boston, MA 1999.

O’Hagan A., Buck C.A. i in., *Uncertain judgements-eliciting experts’ probabilities*, John Wiley&Sons, LTD, Chichester 2006.

Sandri S., Dubois D., Kalfsbeek H., *Elicitation assesment and pooling of expert judgments using possibility theory*, “IEEE Transactions on Fuzzy Systems” 1995, vol. 3, no. 3.

Shafer G., *A mathematical theory of evidence*, Princeton University Press, Princeton, NJ 1976.

REPRAESENTATIONSFUNKTION $R_1(X)$ VON HYPHOTHESEN UND IHRE ANWENDUNG ZUR AGGREGATION VON EXPERTENSCHAETZUNGEN

Zusammenfassung

Integrierung numerischer Daten mit Expertenschätzungen ermöglicht bessere Lösung ökonomischer (und nicht nur) Problemen. Deswegen ist die Entwicklung der Aggregationsmethoden von Expertenschätzungen sehr wichtig. Es gibt bereits eine Reihe der Aggregationsmethoden die im Rahmen unterschiedlicher Theorien der Unsicherheit erarbeitet wurden. In dem Artikel ist eine probabilistische Methode dargestellt. Die meisten Methoden aggregieren gut kohärente Expertenschätzungen. Die im Artikel vorgeschlagene Methode dient zur Aggregation aller Schätzungstypen: sowohl kohärenter wie auch akohärenter, widersprüchlichen Schatzungen. Die Methode basiert auf dem Konzept der Repräsentationsgenauigkeit von Expertenschätzungen durch alle möglichen Hypothesen ($X = x$) bettreffens der geschätzten Variable X. Nach dem Wissen des Authors ist die Methode neu.

Übersetzt von Andrzej Piegat