Wstęp do numeru specjalnego *Delty*

Dane to jeden z surowców, na których oparta jest współczesna cywilizacja. Jak niemal każdy surowiec, wymagają pewnego przetworzenia, zanim staną się w pełni użyteczne. Intelektualną rafinerią, którą wypracowała ludzkość, by radzić sobie z bogatymi złożami wydobywanej z wielką determinacją informacji, jest szeroko rozumiana statystyka. Stanowi ona również wspólny mianownik artykułów zamieszczonych w tym wydaniu *Delty*. Mamy nadzieję, że każdy Czytelnik (nie tylko *Statystyczny*) znajdzie w tym zeszycie coś odpowiadającego jego naukowym zainteresowaniom. Życzymy pouczającej lektury!

Powstanie niniejszego numeru jest skorelowane z odbywającą się w dniach 3–7 lipca 2023 roku w Warszawie konferencją *The European Meeting of Statisticians*, organizowaną przez Polskie Towarzystwo Matematyczne, Politechnikę Warszawską i Uniwersytet Warszawski. Uczestnicy konferencji otrzymają angielską wersję wydania, którą można również odnaleźć (i polecać wśród swoich niepolskojęzycznych znajomych) na stronie `deltami.edu.pl`.

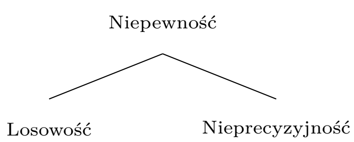
Redakcja

O statystyce, braku precyzji i serach

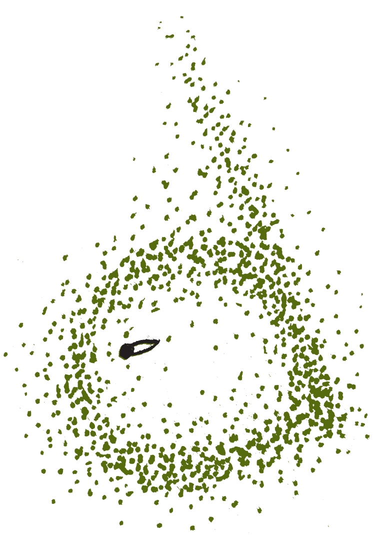
*Przemysław GRZEGORZEWSKI**

* Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Statystyka może być postrzegana jako sztuka podejmowania decyzji w warunkach niepewności. Dostarcza narzędzi do opisu, wyjaśniania i prognozowania zjawisk i procesów zachodzących w realnym świecie oraz weryfikowania hipotez ich dotyczących. Przez długi czas niepewność była utożsamiana z losowością, a w konsekwencji rachunek prawdopodobieństwa był uznawany za jedyną solidnie ugruntowaną teorię pozwalającą opisać niepewność. W ciągu ostatnich pięćdziesięciu lat pojawiły się jednak pewne podejścia rozszerzające klasyczną teorię prawdopodobieństwa lub do niej ortogonalne. Ich wspólną cechą jest próba złagodzenia założeń standardowych metod, aby mogły radzić sobie z innymi rodzajami niepewności, takimi jak nieprecyzyjność.



Należy pamiętać, że sama nieprecyzyjność nie jest (*nomen omen*) precyzyjnie zdefiniowana. Dość często wyniki eksperymentów są nieprecyzyjne z powodu niedokładności aparatury lub błędów popełnionych przez osoby dokonujące pomiary. Czasami pożądany pomiar jest tak trudny do przeprowadzenia, że jego wynik, z reguły, powinien być traktowany jako bardzo niepewny. Może również zdarzyć się, że dokładna wartość zmiennej jest celowo ukrywana ze względu na wrażliwy charakter gromadzonych danych. We wszystkich tych sytuacjach obserwacje często są rejestrowane jako zbiory (np. przedziały) zawierające ich dokładne wartości, nie dające jednak pełnej informacji o należącej do nich punktowej obserwacji x. Każdy taki zbiór reprezentuje pewien **epistemiczny** stan obserwacji. Istnieją jednak sytuacje, gdy dane eksperymentalne są nieprecyzyjne w samej swej istocie. Typowym przykładem są ludzkie odczucia, które nie mają obiektywnej wartości (jak smak czy nastrój). Z tego typu sytuacją mamy również do czynienia, gdy dane dotyczą zjawiska ze swej natury zmiennego, z rozmytymi lub zmieniającymi się granicami, elastycznymi przedziałami czasowymi lub skalami ocen itp. Każda taka obserwacja, nawet jeśli nie jest precyzyjna, odpowiada pewnemu istniejącemu stanowi rzeczywistości, a zatem ma charakter nieprecyzyjności **ontycznej**.



Wygodną metodę matematycznego reprezentowania nieprecyzyjności zaproponował Lotfi A. Zadeh (1921–2017), który wprowadził **teorię zbiorów rozmytych** stanowiącą rozszerzenie klasycznej teorii zbiorów. Zadeh, jeden z najwybitniejszych myślicieli współczesności, zdał sobie sprawę, że chociaż jesteśmy przyzwyczajeni do dzielenia wszystkiego na „tak” i „nie” lub na czarne i białe, to nasz świat rysuje się w odcieniach szarości. Jego słynna teza, że *wszystko jest kwestią stopnia,* stała się główną myślą stojącą za logiką rozmytą i jej imponującymi zastosowaniami. Warto w tym miejscu wyjaśnić, że logika rozmyta w rzeczywistości nie jest „rozmytą” czy „niedokładną” logiką, ale logiką, która opisuje i ujarzmia nieprecyzyjność.

L.A. Zadeh, *Fuzzy sets*, Information and Control 8 (1965), 338–353.

Podstawowym pojęciem wprowadzonym przez Zadeha jest **zbiór rozmyty**. Niech $\mathbb{U}$ będzie ustaloną przestrzenią rozważań. **Zbiór rozmyty** A w $\mathbb{U}$ jest utożsamiany z odwzorowaniem $A : \mathbb{U} \to [0, 1]$, nazywanym **funkcją przynależności**, przypisującym każdemu obiektowi $x \in \mathbb{U}$ liczbę rzeczywistą $A(x)$ z przedziału $[0, 1]$, wskazującą stopień przynależności x do A. Zatem zbiór rozmyty A może być postrzegany jako (standardowy) podzbiór $\mathbb{U} \times [0, 1]$, tj.:

$$A = \{(x, A(x)) : x \in \mathbb{U}, A(x) \in [0, 1]\}.$$

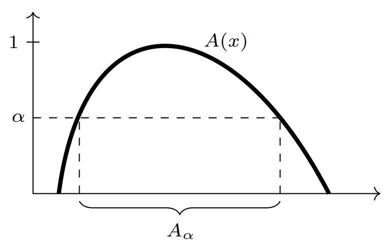
Zauważmy, że w teorii mnogości funkcje $f\colon X \to Y$ są zazwyczaj definiowane jako zbiory; dokładniej, podzbiory $X \times Y$ takie, że dla każdego $x \in X$ istnieje dokładnie jedno $y \in Y$ spełniające $(x, y) \in f$. W tym sensie $f(x)$ to jedynie konwencja notacyjna oznaczająca jedyne y, dla którego $(x, y) \in f$.

Interpretacja funkcji przynależności jest naturalna: jeśli $A(x) = 1$, to jesteśmy pewni, że element x należy do A, podczas gdy $A(x) = 0$ oznacza, że x nie należy do A. We wszystkich innych przypadkach, tj. gdy $A(x) \in (0, 1)$, mamy częściową przynależność do A. Oznacza to, że jeśli $A(x)$ jest bliskie 1, to stopień przynależności x do A jest wysoki, a jeśli $A(x)$ jest bliskie 0, to stopień przynależności x do A jest niski. Jeśli $A(x) \in \{0, 1\}$ dla wszystkich $x \in \mathbb{U}$, to A jest zbiorem w klasycznym rozumieniu (i odwrotnie, każdy „zwykły" zbiór może być traktowany jako zbiór rozmyty, którego funkcją przynależności jest funkcja charakterystyczna).

Innym ważnym pojęciem związanym ze zbiorami rozmytymi jest tzw. **α-cięcie**. Dla każdego $\alpha \in [0, 1]$ α-cięcie zbioru rozmytego A, oznaczane jako A_α, jest zdefiniowane jako

$$A_\alpha = \begin{cases} \{x \in \mathbb{U} : A(x) \geqslant \alpha\} & \text{dla } \alpha \in (0, 1], \\ \mathrm{cl}\{x \in \mathbb{U} : A(x) > 0\} & \text{dla } \alpha = 0, \end{cases}$$

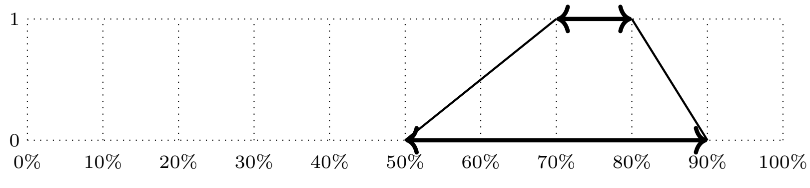
gdzie cl oznacza operację domknięcia (od tej pory zakładamy, że taka operacja jest zdefiniowana na $\mathbb{U}$). Innymi słowy, α-cięcie to „zwykły" podzbiór tych elementów $\mathbb{U}$, których stopień przynależności do A jest nie mniejszy niż α. Można pokazać, że każdy zbiór rozmyty jest jednoznacznie wyznaczony przez rodzinę wszystkich swoich α-cięć $\{A_\alpha\}_{\alpha \in [0,1]}$. Dwa α-cięcia zasługują na szczególną uwagę: A_1 nazywane **jądrem**, które zawiera wszystkie wartości w pełni zgodne z opisywanym przez A pojęciem, oraz A_0 nazywane **nośnikiem**, zawierające elementy zgodne z tym pojęciem w jakimś stopniu.



Rys. 1. Wykres przykładowej funkcji przynależności $A(x)$ liczby rozmytej A

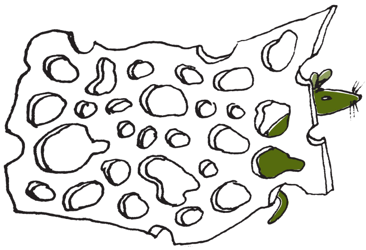
Ważną podrodziną zbiorów rozmytych są liczby rozmyte. Mówimy, że A jest **liczbą rozmytą**, jeśli $A : \mathbb{R} \to [0, 1]$ i wszystkie jego α-cięcia są niepustymi przedziałami domkniętymi. Przykład liczby rozmytej jest przedstawiony na rysunku 1.

Przykład. Ser Gamonedo to rodzaj sera pleśniowego produkowanego w Asturii (północna Hiszpania). Jest on wędzony, a następnie pozostawiany do dojrzewania w jaskiniach lub innych suchych miejscach. W trosce o zapewnienie odpowiedniej jakości produktu eksperci (degustatorzy) oceniają różne cechy sera: wizualne (kształt, skórka, wygląd), teksturę (twardość i kruchość), zapachowo-smakowe (intensywność i jakość), a także ogólne o nim wrażenie. Ostatnio zapostulowano, aby do opisu swych subiektywnych ocen degustatorzy posłużyli się trapezoidalnymi liczbami rozmytymi. Ten właśnie typ liczb rozmytych jest często stosowany ze względu na łatwość interpretacji oraz przetwarzania. Ocena sera dokonywana jest w skali od 0% (najniższa jakość) do 100% (najwyższa jakość). Na poziomie 0 (por. rys. 2) degustator zaznacza zakres wartości, który jest w jakimś stopniu zgodny z jego opinią (tzn. jakość rozważanej cechy z pewnością nie wychodzi poza ten zakres). Natomiast na poziomie 1 ekspert zaznacza zakres wartości, który w pełni odzwierciedla jego opinię odnośnie jakości. Przykład z rysunku 2 ilustruje sytuację, w której tester uważa, że dany ser spełnia wymagania jakościowe pod względem badanej cechy w 70–80%. Jednocześnie jest on przekonany, że wymagania jakościowe są spełnione w stopniu nie niższym niż 50%, ale też nie wyższym niż 90%.

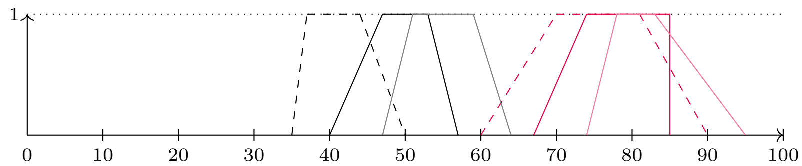


Rys. 2. Przykład eksperckiej opinii reprezentowanej przez trapezoidalny zbiór rozmyty

Rozwiązanie zadania M 1750.
Rozważmy dowolne rozmieszczenie liczb. Zauważmy, że liczby od 1 do 2022 reprezentują co najwyżej 2022 wiersze i 2022 kolumny. Istnieją zatem wiersz i kolumna zawierające wyłącznie liczby większe od 2022. Iloczyn dowolnych dwóch z nich jest równy co najmniej $2023 \cdot 2024$, a to więcej niż dowolna liczba wpisana w tablicę. Oznacza to, że nie istnieje rozmieszczenie liczb opisane w zadaniu.



Zakresy wartości oceny na pozostałych poziomach (poza 0 i 1) są następnie interpolowane liniowo, tak aby otrzymać trapezoidaną liczbę rozmytą (por. rys. 2). Przykład ocen wyrażonych przez sześciu degustatorów, które zostały zamodelowane przy użyciu trapezoidalnych liczb rozmytych, przedstawiono na rysunku 3.



Rys. 3. Próbka sześciu eksperckich opinii reprezentowanych przez trapezoidalne zbiory rozmyte

Bardziej formalnie, mówimy, że A jest **trapezoidalną** liczbą rozmytą, jeśli jej funkcja przynależności jest dana wzorem

$$A(x) = \begin{cases} \frac{x-a_1}{a_2-a_1} & \text{gdy } a_1 \leqslant x < a_2, \\ 1 & \text{gdy } a_2 \leqslant x \leqslant a_3, \\ \frac{a_4-x}{a_4-a_3} & \text{gdy } a_3 < x \leqslant a_4, \\ 0 & \text{w przeciwnym przypadku,} \end{cases} \tag{1}$$

gdzie $a_1, a_2, a_3, a_4 \in \mathbb{R}$ spełniają $a_1 \leqslant a_2 \leqslant a_3 \leqslant a_4$. Ponieważ trapezoidalna liczba rozmyta (1) jest charakteryzowana przez cztery liczby rzeczywiste, często oznacza się ją przez $A = (a_1, a_2, a_3, a_4)_T$.

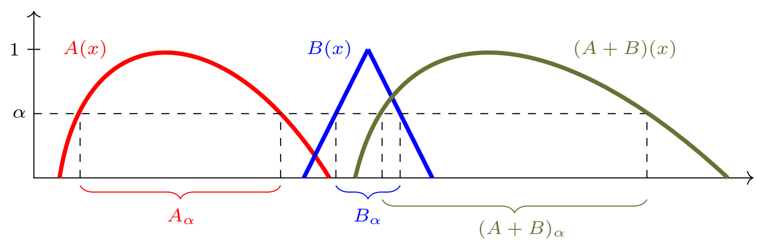
Zdefiniujemy teraz podstawowe operacje arytmetyczne na liczbach rozmytych. Chociaż można je wprowadzić, operując bezpośrednio na funkcjach przynależności, łatwiejsza wydaje się definicja wykorzystująca α-cięcia. Suma dwóch liczb rozmytych A i B jest określona przez tzw. *sumę Minkowskiego* α-cięć składników (patrz rys. 4), tj. dla każdego $\alpha \in [0, 1]$

$$(A + B)_\alpha = \big[\inf A_\alpha + \inf B_\alpha, \sup A_\alpha + \sup B_\alpha\big], \tag{2}$$

gdzie $\inf I$ oraz $\sup I$ oznaczają, odpowiednio, lewy i prawy koniec przedziału I.

Warto podkreślić, że suma trapezoidalnych liczb rozmytych jest trapezoidalną liczbą rozmytą, tzn. jeśli $A = (a_1, a_2, a_3, a_4)_T$ i $B = (b_1, b_2, b_3, b_4)_T$, to

$$A + B = \\ = (a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4)_T.$$



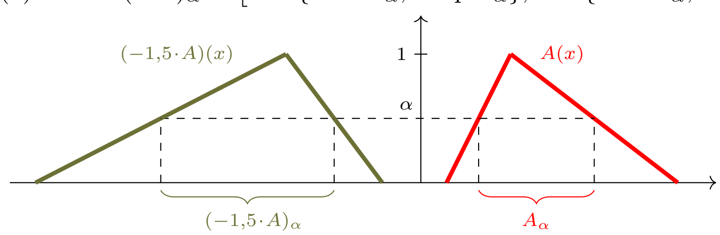
Rys. 4. Suma liczb rozmytych A i B

Podobnie, mnożenie liczby rozmytej A przez skalarną wartość $\theta \in \mathbb{R}$ jest określone poprzez tzw. iloczyn Minkowskiego α-cięć (patrz rys. 5), tzn. dla każdego $\alpha \in [0, 1]$

$$(\theta \cdot A)_\alpha = \big[\min\{\theta \inf A_\alpha, \theta \sup A_\alpha\}, \max\{\theta \inf A_\alpha, \theta \sup A_\alpha\}\big]. \tag{3}$$

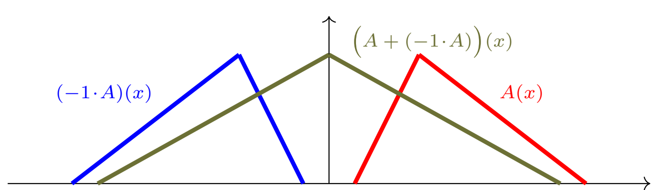
Pomnożenie trapezoidalnej liczby rozmytej $A = (a_1, a_2, a_3, a_4)_T$ przez skalar θ daje trapezoidalną liczbę rozmytą:

$$\theta \cdot A = \begin{cases} (\theta a_1, \theta a_2, \theta a_3, \theta a_4)_T & \text{gdy } \theta \geqslant 0, \\ (\theta a_4, \theta a_3, \theta a_2, \theta a_1)_T & \text{gdy } \theta < 0. \end{cases}$$



Rys. 5. Mnożenie liczby rozmytej A przez skalar

Niestety, w ogólności $A + (-1 \cdot A) \neq \mathbb{1}_0$ (patrz rys. 6). W konsekwencji, dla różnicy opartej na operacji Minkowskiego nie zawsze zachodzi $(A + (-1 \cdot B)) + B = A$.



Rys. 6. Ilustracja problemów z odejmowaniem liczb rozmytych

Aby sprostać niektórym trudnościom związanym z brakiem „porządnej" operacji odejmowania liczb rozmytych, zwłaszcza w kontekście konstruowania narzędzi do wnioskowania statystycznego, często rozważa się *odległość* między liczbami rozmytymi. Tak jak poprzednio, wygodnie będzie do jej opisu wykorzystać α-cięcia. Odległość między dwoma odcinkami $[a,b]$, $[c,d]$ możemy określić jako odległość między punktami (a,b), (c,d) na płaszczyźnie (czyli $\sqrt{(a-c)^2+(b-d)^2}$. Następnie odległość między liczbami rozmytymi A, B zdefiniujemy jako średnią kwadratową z odległości między A_α i B_α po wszystkich α, co prowadzi do następującego wzoru:

$$D(A,B) = \sqrt{\int_0^1 \big[(\inf A_\alpha - \inf B_\alpha)^2 + (\sup A_\alpha - \sup B_\alpha)^2\big]\, d\alpha}. \tag{4}$$

Odnotujmy, że rzeczywiście jest to odległość, czyli *metryka* (patrz np. artykuł Jarosława Górnickiego z Δ^5_{21}). Oczywista jest własność $D(A,B) \geqslant 0$, jak i to, że $D(A,B) = 0$ tylko w przypadku $A = B$. Nieco trudniejsza w uzasadnieniu jest nierówność trójkąta $D(A,B) + D(B,C) \geqslant D(A,C)$, którą pozostawiamy jako ćwiczenie.

Załóżmy teraz, że mamy do czynienia z niezależnymi, rozmytymi obserwacjami dwóch populacji; n obserwacji z populacji pierwszej, $\boldsymbol{x} = (x_1,\ldots,x_n)$, oraz m obserwacji z drugiej, $\boldsymbol{y} = (y_1,\ldots,y_m)$ (wszystkie x_i oraz y_i są liczbami rozmytymi). Chcemy przekonać się, czy występuje znacząca różnica między tymi populacjami. W tym celu mierzymy odległość (4) między średnimi arytmetycznymi wyznaczonymi dla obu próbek. Zauważmy, że faktycznie potrafimy już obliczać średnią arytmetyczną liczb rozmytych (która sama jest liczbą rozmytą), gdyż mamy narzędzia do ich dodawania i mnożenia przez liczby rzeczywiste. Pozostaje pytanie, czy konkretna odległość między średnimi, np. 3,14, jest duża, czy mała? W tym miejscu na scenę wkracza statystyka matematyczna!

W żargonie statystyków naszym celem jest weryfikacja hipotezy zerowej H_0, mówiącej, że obie próbki pochodzą tak naprawdę z tej samej populacji, przeciw hipotezie alternatywnej, że populacje się różnią. Jeśli hipoteza zerowa jest prawdziwa, oczekujemy, że średnie z obu próbek nie będą się zbytnio różnić. Natomiast znacząca różnica między średnimi wskazuje na to, że badane próbki pochodzą z różnych populacji.

Aby zdecydować, czy mierzona odległość jest wystarczająco duża, by uznać ją za znaczącą, statystycy często odwołują się do pojęcia *p-wartości.* W naszym przypadku będzie to prawdopodobieństwo (przy założeniu, że hipoteza zerowa jest prawdziwa) otrzymania co najmniej tak dużej odległości między średnimi, jak ta zaobserwowana. Zgodnie z intuicją, jeśli to prawdopodobieństwo jest niskie, mamy dobry powód, by odrzucić hipotezę zerową. Problem polega na tym, że w naszym przypadku nie możemy dokładnie wyznaczyć tego prawdopodobieństwa, ponieważ hipoteza zerowa orzeka jedynie, że dwie populacje można traktować jak jedną, ale nie podaje nam konkretnego opisu probabilistycznego tej populacji. Musimy zatem odwołać się do innego, sprytnego pomysłu.

Niech $\boldsymbol{v}$ będzie *złączeniem* dwóch próbek, tj. $v_i = x_i$ dla $1 \leqslant i \leqslant n$ oraz $v_i = y_{i-n}$ dla $n+1 \leqslant i \leqslant N$, gdzie $N = n+m$. Niech teraz $\boldsymbol{v}^*$ oznacza permutację początkowego zbioru danych $\boldsymbol{v}$. Z pierwszych n elementów ciągu $\boldsymbol{v}^*$ tworzymy próbkę $\boldsymbol{x}^*$, natomiast z pozostałych m elementów, próbkę $\boldsymbol{y}^*$. Innymi słowy, dokonujemy losowego podziału elementów $\boldsymbol{v}$ na dwie próbki o rozmiarach, odpowiednio, n i m. Zauważmy, że jeśli hipoteza H_0 jest spełniona, tj. obie próbki pochodzą z tego samego rozkładu, to taka operacja nie zmieni losowości stojącej za obserwowanymi próbkami (z punktu widzenia losowości dowolny taki podział jest w pewnym sensie równoprawny). W konsekwencji możemy *oszacować* prawdziwą p-wartość jako frakcję wszystkich możliwych permutacji $\boldsymbol{v}^*$, które prowadzą do większej różnicy między średnimi $\boldsymbol{x}^*$ i $\boldsymbol{y}^*$ niż ta zaobserwowana.

Rozwiązanie zadania M 1752.
Zauważmy, że dla dowolnych $M, a, b \in [0,1]$ takich, że $M \geqslant a, b$, zachodzi nierówność:

$$(M-a)(M-b)(1-bM) \geqslant 0,$$

co po przekształceniach daje

$$(1-bM+b^2)(1-aM+M^2) \geqslant 1-ab+b^2.$$

Aby udowodnić nierówność z zadania dla $n=2$, wystarczy w powyższej nierówności wziąć $a = b = x_1$ oraz $M = x_2$. Załóżmy, że teza zadania zachodzi dla pewnego n; wywnioskujemy stąd jej słuszność dla $n+1$. Bez straty ogólności załóżmy, że $x_{n+1} = \max\{x_1,\ldots,x_{n+1}\}$. Używając powyższej nierówności dla $b = x_n$, $M = x_{n+1}$ oraz $a = x_1$, otrzymujemy:

$$(1-x_nx_{n+1}+x_n^2)(1-x_{n+1}x_1+x_{n+1}^2) \geqslant \geqslant (1-x_nx_1+x_n^2).$$

Dlatego

$$\prod_{\text{cykl}}^{n+1}(1-x_ix_{i+1}+x_i^2) \geqslant \geqslant \prod_{\text{cykl}}^{n}(1-x_ix_{i+1}+x_i^2) \geqslant 1,$$

gdzie przez „$\prod_{\text{cykl}}$" rozumiemy „zacyklony" iloczyn. Powołanie się na zasadę indukcji kończy dowód.

Równanie (5) można traktować jako definicję *prawdziwej* p-wartości, pod warunkiem zdarzenia, że nasze próbki sumują się (jako zbiory swoich wyrazów) do $\boldsymbol{v}$. Takie podejście jest dość standardowe przy projektowaniu tzw. *testów nieparametrycznych*, np. *testu serii* opisanego szczegółowo w artykule o tej samej nazwie w Δ^9_{17}.

Poniższa tabela przedstawia opinie dwóch ekspertów dotyczące ogólnego wrażenia na temat łącznej liczby 78 próbek sera Gamonedo, por. Ramos-Guajardo i in. (2019). Każdy ekspert oceniał inne próbki.

Ekspert 1	Ekspert 2
$(65, 75, 85, 85)$	$(50, 50, 63, 75)$
$(35, 37, 44, 50)$	$(39, 47, 52, 60)$
$(66, 70, 75, 80)$	$(60, 70, 85, 90)$
$(70, 74, 80, 84)$	$(50, 56, 64, 74)$
$(65, 70, 75, 80)$	$(39, 45, 53, 57)$
$(45, 50, 57, 65)$	$(55, 60, 70, 76)$
$(60, 66, 70, 75)$	$(50, 50, 57, 67)$
$(65, 65, 70, 76)$	$(65, 67, 80, 87)$
$(60, 65, 75, 80)$	$(50, 50, 65, 75)$
$(55, 60, 66, 70)$	$(50, 55, 64, 70)$
$(60, 65, 70, 74)$	$(39, 46, 53, 56)$
$(30, 46, 44, 54)$	$(19, 29, 41, 50)$
$(60, 65, 75, 75)$	$(40, 47, 52, 56)$
$(70, 75, 85, 85)$	$(54, 55, 65, 76)$
$(44, 45, 50, 56)$	$(59, 65, 75, 85)$
$(51, 56, 64, 70)$	$(50, 52, 57, 60)$
$(40, 46, 54, 60)$	$(60, 60, 70, 80)$
$(55, 60, 65, 70)$	$(50, 54, 61, 67)$
$(80, 85, 90, 94)$	$(40, 46, 50, 50)$
$(80, 84, 90, 90)$	$(44, 50, 56, 66)$
$(65, 70, 76, 80)$	$(60, 64, 75, 85)$
$(75, 80, 86, 90)$	$(54, 56, 64, 75)$
$(65, 70, 73, 80)$	$(50, 50, 60, 66)$
$(70, 80, 84, 84)$	$(44, 46, 55, 57)$
$(55, 64, 70, 70)$	$(59, 63, 74, 80)$
$(64, 73, 80, 84)$	$(49, 50, 54, 58)$
$(50, 56, 64, 70)$	$(55, 60, 70, 75)$
$(55, 55, 60, 70)$	$(44, 47, 53, 60)$
$(60, 70, 75, 80)$	$(19, 20, 30, 41)$
$(64, 71, 80, 80)$	$(40, 44, 50, 60)$
$(50, 50, 55, 65)$	$(50, 50, 59, 66)$
$(50, 54, 60, 65)$	$(50, 53, 60, 66)$
$(65, 75, 80, 86)$	$(50, 52, 58, 61)$
$(50, 55, 60, 66)$	$(60, 65, 72, 80)$
$(40, 44, 50, 50)$	$(50, 50, 55, 60)$
$(70, 76, 85, 85)$	$(30, 34, 43, 47)$
$(44, 50, 53, 60)$	$(19, 25, 36, 46)$
$(34, 40, 46, 46)$	$(53, 63, 74, 80)$
$(40, 45, 51, 60)$	
$(84, 90, 95, 95)$	

Formalnie można wyrazić tę ideę wzorem:

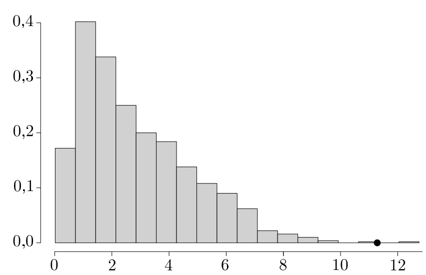
$$\text{p-wartość} \simeq \frac{1}{N!} \sum_{\boldsymbol{v}^*} \mathbb{1}(T(\boldsymbol{v}^*) \geqslant t_0), \tag{5}$$

w którym suma obejmuje wszystkie permutacje $\boldsymbol{v}^*$ zbioru $\boldsymbol{v}$, $T(\boldsymbol{v}^*)$ oznacza odległość między średnimi z próbek $\boldsymbol{x}^*$ i $\boldsymbol{y}^*$, zaś t_0 jest obserwowaną odległością między średnimi z próbek $\boldsymbol{x}$ i $\boldsymbol{y}$. Wartość wyrażenia $\mathbb{1}(\texttt{warunek})$ jest równa 1, jeśli dany warunek jest spełniony, a 0 w przeciwnym przypadku.

Wzór (5) może być dalej uproszczony dzięki wykorzystaniu faktu, że permutacje można podzielić na grupy rozmiaru $n!m!$, w których permutacje dają te same średnie z próbek $\boldsymbol{x}^*$ i $\boldsymbol{y}^*$ (są to permutacje, z których każdą można uzyskać z innej poprzez permutowanie pierwszych n i ostatnich m obserwacji). Jednak nawet po dokonaniu takiego uproszczenia liczba składników rozważanej sumy, wynosząca wówczas $\binom{N}{n}$, rośnie wykładniczo wraz z N (utrzymując stosunek n/N na ustalonym poziomie). Dlatego, zamiast uwzględniać wszystkie możliwe permutacje, ograniczamy się do średniej obliczonej na podstawie (możliwie wielu) losowo wybranych permutacji. Konkretniej, niech $\boldsymbol{v}_1^*, \boldsymbol{v}_2^*, \ldots, \boldsymbol{v}_K^*$ będą pewnymi losowymi permutacjami $\boldsymbol{v}$ (gdzie K jest odpowiednio duże, np. większe niż 1000). Wtedy przybliżenie p-wartości naszego testu jest określone przez:

$$\text{p-wartość} \simeq \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}(T(\boldsymbol{v}_k^*) \geqslant t_0). \tag{6}$$

Przykład. Wykorzystamy teraz dane z pracy Ramos-Guajardo i in. (2019), aby porównać ogólne wrażenia dwóch degustatorów na temat serów Gamonedo. Trapezoidalne liczby rozmyte odpowiadające 40 opiniom wydanym przez pierwszego eksperta oraz 38 opiniom drugiego są zebrane w tabeli na marginesie. Liczby w nawiasach odpowiadają notacji używanej do opisu trapezoidalnych liczb rozmytych, np. $x_1 = (65, 75, 85, 85)_T$, $y_1 = (50, 50, 63, 75)_T$ itd. Chcemy sprawdzić, czy istnieje zgodność między tymi dwoma ekspertami co do ogólnego wrażenia, jakie robią oceniane sery. Aby osiągnąć ten cel, weryfikujemy hipotezę zerową H_0 głoszącą, że nie ma istotnej różnicy między opiniami ekspertów, przeciwko temu, że ich opinie na temat jakości sera różnią się. Na podstawie danych z tabeli można wyznaczyć średnie wartości $\overline{x} = (57{,}65;\ 63{,}20;\ 69{,}18;\ 73{,}48)_T$ i $\overline{y} = (47{,}34;\ 51{,}21;\ 59{,}87;\ 66{,}84)_T$. Podstawiając te wyniki do (4), otrzymujemy wartość statystyki testowej $t_0 = D(\overline{x}, \overline{y}) = 11{,}26$. Następnie, po złączeniu próbek i wygenerowaniu $K = 1000$ losowych permutacji, stosując (6), otrzymujemy przybliżenie p-wartości równe 0,002. Jej interpretacja jest pokazana na rysunku 7, przedstawiającym histogram wszystkich uzyskanych symulacyjnie różnic $D(\overline{x^*}, \overline{y^*})$. Czarny punkt oznacza obserwowaną wartość t_0 statystyki testowej. Słabo widoczny szary obszar na prawo od tego punktu odpowiada prawdopodobieństwu uzyskania odległości między średnimi $\boldsymbol{x}^*$ a $\boldsymbol{y}^*$ nie mniejszej niż t_0. Dlatego też możemy bez większego wahania odrzucić hipotezę zerową i wnioskować, że nie ma ogólnej zgody między opiniami rozważanych ekspertów na temat sera Gamonedo.



Rys. 7

Przedstawiony permutacyjny test zgodności dwóch próbek zawierających nieprecyzyjne informacje jest jednym z wielu przykładów, jak modelowanie rozmyte może być połączone z klasycznym wnioskowaniem statystycznym. Pomimo początkowego sceptycyzmu wobec prób połączenia obu teorii z czasem stało się jasne, że zarówno statystyka, jak i teoria zbiorów rozmytych nie powinny być uważane za konkurencyjne i mogą skutecznie się uzupełniać. Co więcej, rozszerzenie statystyki o analizę zbiorów rozmytych nie tylko pomaga rozwiązać pewne problemy, ale jednocześnie stawia nowe pytania badawcze. W szczególności, rozróżnienie między tzw. zbiorami ontologicznymi i epistemicznymi prowadzi do różnych definicji nawet tak podstawowych pojęć jak wariancja, a co za tym idzie – różnych narzędzi wnioskowania.