

Andrzej Wilkowski

O WIEŁOŚREDNIEJ

1. W zastosowaniach praktycznych statystyki często mamy do czynienia z rozkładami wielomodalnymi, szczególnie dwumodalnymi. Zwykle ma to związek z niejednorodnością populacji. Na przykład analizowana statystycznie grupa ludzi rozpada się na kobiety i mężczyzn, uprawiane przez nich zawody, czy też rasy. Towary jednej firmy mogą być wytwarzane w różnych przedsiębiorstwach lub krajach. Pojawia się wtedy zjawisko mieszania rozkładów, rozumiane bądź jako złożenie rozkładów, bądź też jako ich wypukła kombinacja. W literaturze polskiej jest ono dobrze opisane w książce W. Ostasiewicza [14]. We wspomnianych wyżej przypadkach występuje właśnie wielomodalność. Za modalną będziemy tutaj uważali maksimum funkcji gęstości (o ile istnieje). Ogólną definicję i własności modalnej można znaleźć w książce P. Hennequina, A. Tortrata [8] i pracy A. Smoluka [16]. Niniejsza praca jest kontynuacją badań zawartych w artykule R. Antoniewicza i A. Wilkowskiego [5]. Zostaną podane wzory na momenty zmiennej o rozkładzie Webera oraz przedstawione będą rozważania dotyczące funkcji charakterystycznej. Następnie wprowadzone zostanie pojęcie k -średniej [2], omówione zostaną jej własności i związek z wielomianami ortogonalnymi.

2. Jak wiadomo, populacja jest dokładnie określona przez rozkład. Kluczowe znaczenie dla jej opisu mają zatem miejsca koncentracji masy prawdopodobieństwa-mody. Modalną nie jest łatwo jednak estymować z próbki losowej [18]. Powszechnie używanym estymatorem tejże, w rozkładach jednomodalnych, jest przeciętna. Gdy populacja jest dwumodalna, to wówczas wartość przeciętna nie wyznacza na ogół miejsca koncentracji prawdopodobieństwa. W związku z tym Antoniewicz [1] wprowadził pojęcie biprzeciętnej (przykłady liczbowe pokazują, że pozwala ono wyznaczyć miejsca koncentracji w przypadku rozkładów dwumodalnych [5]. Przypomnijmy, że jeśli X jest zmienną losową mającą dwa pierwsze momenty, to oczywiście wartość przeciętna $M(X - a)^2$ osiąga minimum dla

$a = M(X)$, natomiast minimalna wartość to wariancja: $M(X - M(X))^2 = V(X)$. Niech teraz X będzie zmienną losową mającą cztery pierwsze momenty i wariancję różną od zera. Wyrażenie $M(X^2 - pX - q)^2$ osiąga minimum, jeśli

$$P = \frac{M(X^3) - M(X^2)M(X)}{V(X)},$$

$$Q = \frac{M^2(X^2) - M(X^3)M(X)}{V(X)}.$$

Niech $P^2 + 4Q \geq 0$ oraz $R = \sqrt{P^2 + 4Q}$. Przez analogię parę liczb (M_1, M_2) , gdzie

$$M_1 = \frac{P - R}{2}, \quad M_2 = \frac{P + R}{2},$$

nazywamy **biprzeciętną**. Wyrażenie

$$V_0(X) = M((X - M_1)(X - M_2))^2$$

jest analogonem wariancji. Propozycje innych miar rozrzutu oraz przykłady liczbowe podane są w pracach [4; 5]. Na podstawie tych przykładów widać, że biprzeciętna dość dobrze przybliża mody. Ma ona jednak tę przewagę nad nimi, że łatwo ją wyznaczyć. W tym celu wystarczy obliczyć trzy pierwsze momenty a następnie skorzystać z powyższych wzorów.

Poniższy przykład jest propozycją rozkładu dwumodalnego o ciągłej funkcji gęstości, określonego na prostej, nie będącego mieszkanką.

Niech $\alpha, \beta > 0$. Rozkład o funkcji gęstości g , określonej wzorem

$$g_{\alpha, \beta}(x) = \frac{1}{z(\alpha, \beta)} e^{\alpha x^2 - \beta x^4}, \quad x \in \mathfrak{R},$$

nazywamy **rozkładem Webera** z parametrami α, β i oznaczamy jako $W(\alpha, \beta)$.

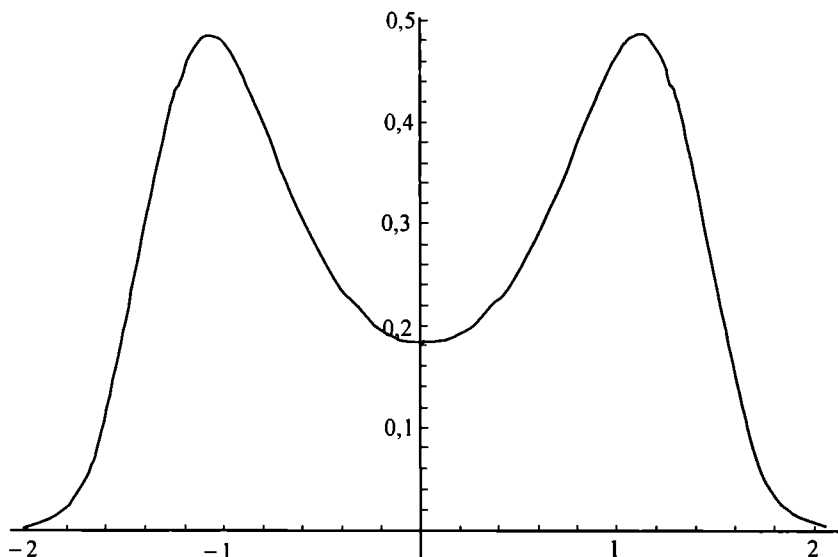
Niżej przedstawiamy funkcję gęstości rozkładu $W(2, 1)$.

Komentarza wymaga stała normalizacyjna $z(\alpha, \beta)$. Okazuje się, że tę stałą można wyrazić przez funkcje specjalne Webera [6], które są stablicowane [9]. Mianowicie:

$$z(\alpha, \beta) = \int_{\mathfrak{R}} e^{\alpha x^2 - \beta x^4} dx = \exp\left(\frac{\alpha^2}{8\beta}\right) \frac{\sqrt{\pi}}{\sqrt[4]{2\beta}} D_{-\frac{1}{2}}\left(-\frac{\alpha}{\sqrt{2\beta}}\right),$$

gdzie D jest funkcją Webera, spełniającą równanie różniczkowe:

$$\frac{d^2 D_p(x)}{dx^2} + \left(p + \frac{1}{2} - \frac{x^2}{4} \right) D_p(x) = 0, \text{ dla } p \neq 0, 1, 2, \dots$$



Rys. 1. Funkcja gęstości rozkładu $W(2, 1)$

Może być ona przedstawiona w postaci szeregu:

$$D_p(x) = \frac{2^{-\frac{p-1}{2}} e^{-\frac{x^2}{4}}}{\Gamma(-p)} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma\left(\frac{m-p}{2}\right)}{m!} (\sqrt{2}x)^m,$$

gdzie Γ jest funkcją gamma Eulera.

Na podstawie tablic lub korzystając z pakietu Mathematica 4.0, otrzymujemy:

$$z(2, 1) \approx 5,36516, \quad z(20, 10) \approx 12\,610,9, \quad z(30, 15) \approx 1,51636 \times 10^6.$$

Z definicji widać, że funkcja g ma dwie mody (maksima) w punktach

$$Mo_1 = -\sqrt{\frac{\alpha}{2\beta}}, \quad Mo_2 = \sqrt{\frac{\alpha}{2\beta}}.$$

Niech ϕ oznacza funkcję charakterystyczną rozkładu Webera. Ponieważ gęstość g rozkładu $W(\alpha, \beta)$ jest funkcją parzystą, dostajemy:

$$\phi(t) = \frac{1}{z(\alpha, \beta)} \int_{\mathfrak{R}} e^{\alpha x^2 - \beta x^4 + itx} dx = \frac{1}{z(\alpha, \beta)} \int_{\mathfrak{R}} e^{\alpha x^2 - \beta x^4} \cos tx dx, \quad t \in \mathfrak{R}.$$

Widać zatem, że gdy zróżniczkujemy tę funkcję, spełnia ona równanie różniczkowe:

$$4\beta\phi(t)^m + 2\alpha\phi(t)' - t\phi(t) = 0.$$

Równania tego nie można rozwiązać w postaci jawnej (zauważył to dr A. Misztal). Można natomiast wyznaczyć kolejne momenty zmiennej o rozkładzie Webera. Niech X będzie zmienną losową o rozkładzie Webera z parametrami α, β ($X \approx W(\alpha, \beta)$). Mamy wtedy:

$$M(X^{2n-1}) = 0,$$

$$M(X^{2n}) = \frac{1}{2z(\alpha, \beta)} \beta^{\frac{1}{4}(3-2n)} \left(\sqrt{\beta} \Gamma\left(\frac{1}{4} + \frac{n}{2}\right) + H\left(\frac{1}{4} + \frac{n}{2}, \frac{1}{2}, \frac{\alpha^2}{4\beta}\right) + \right. \\ \left. + \alpha \Gamma\left(\frac{3}{4} + \frac{n}{2}\right) H\left(\frac{3}{4} + \frac{n}{2}, \frac{3}{2}, \frac{\alpha^2}{4\beta}\right) \right),$$

gdzie $n=1, 2, \dots$, a H oznacza konfluentną funkcję hipergeometryczną [6]. Określamy ją jako sumę szeregu:

$$H(a, b, x) = \sum_{k=0}^{\infty} \frac{(a)_k x^k}{(b)_k k!}, \quad (a, b, x \in \mathfrak{R}), \quad (a)_k = (a+k-1)(a+k-2)\dots(a+2)(a+1).$$

Możemy ją także zapisać za pomocą całki:

$$H(a, b, x) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_0^1 e^{xt} t^{a-1} (1-t)^{b-a-1} dt.$$

Konfluentna funkcja hipergeometryczna jest rozwiązaniem równania różniczkowego Kummera:

$$xy'' + (b-x)y' - ay = 0$$

przy warunkach brzegowych $y(0)=1$ i $y'(0)=\frac{a}{b}$. Możemy ją traktować jako uogólnienie innych funkcji specjalnych: funkcji Webera, funkcji Bessela, wielomianów Laguerre'a i Hermite'a itd. Wartości jej są stablicowane. Obliczenia numeryczne sugerują np., że drugie momenty $M(X^2)$, w przypadku gdy $\alpha=2\beta$, zbiegają od dołu do 1, w miarę wzrostu parametru α .

3. W punkcie tym uogólnimy pojęcie biprzeciętnej zdefiniowane wcześniej. W dalszym ciągu W_k będzie oznaczało zbiór wielomianów stopnia k , z najstarszym współczynnikiem równym 1, μ niech będzie nieujemną miarą borelowską na $\mathbf{R}$.

Definicja 1. Wielomianem optymalnym $f_k \in W_k$, na przedziale $[a, b]$ (może to być przedział nieskończony), względem μ , nazywamy wielomian spełniający zależność:

$$\int_a^b f_k^2(x) \mu(dx) = \min_{w \in W_k} \int_a^b w^2(x) \mu(dx).$$

Za R. Antoniewiczem, przez analogię do biprzeciętnej, pierwiastki $M_1, M_2, \dots, M_k$ wielomianu optymalnego f_k nazywamy **k -średnią przedziału $[a, b]$ z miarą μ** . Aby wyznaczyć kolejne wielomiany optymalne $f_1, f_2, \dots, f_k$, należy obliczyć całki

$$\int_a^b x^k \mu(dx) = m_k, \quad k = 0, 1, 2, \dots$$

Całki te nazywają się momentami potęgowymi względem miary μ . Pojawiają się one przy badaniach wielomianów ortogonalnych. Przypomnijmy, że ciąg $p_1, p_2, \dots$ wielomianów jednej zmiennej o współczynnikach rzeczywistych ($\deg p_k = k$) nazywamy ciągiem wielomianów ortogonalnych, jeżeli istnieje miara borelowska μ na $[a, b]$ taka, że

$$\int_a^b p_m p_n d\mu = N_n \delta_{m,n}, \quad m, n = 0, 1, \dots, \delta_{m,n} - \text{symbol Kroneckera},$$

a wszystkie liczby N_n są dodatnie. Nie zawsze ten ciąg jest zupełny w $L^2(\mathbf{R}, \mu)$,

a także nie zawsze zbiór $A = \left\{ z \in \mathbf{C}: \sum_{n=0}^{\infty} |p_n(z)|^2 < +\infty \right\}$ jest niepusty. Ważnym

zagadnieniem związanym z ciągiem $m_1, m_2, \dots$ jest problem momentów. Mówimy, że problem momentów jest zdeterminowany, jeśli μ jest jedyną miarą reprezentującą ustalony ciąg $m_1, m_2, \dots$, natomiast w przypadku braku jedności jest on niezdedeterminowany. Zasadniczy fakt dotyczący problemu momentów mówi, że warunkiem dostatecznym niedeterminowalności problemu momentów jest, by chociaż jedno $z \in \mathbf{C}$ należało do zbioru A , warunkiem zaś koniecznym, by wszystkie $z \in \mathbf{C}$ należały do tego zbioru. Ciąg $m_1, m_2, \dots$, dla którego istnieje nieujemna miara na $\mathbf{R}$ taka, że

$$\int_{\mathbf{R}} x^k \mu(dx) = m_k, k = 0, 1, 2, \dots,$$

nazywamy ciągiem momentów Hamburgera. Wtedy wielomian ortogonalny przybiera postać

$$p_n(x) = K \begin{vmatrix} m_0 & m_1 & \dots & m_n \\ \vdots & \vdots & & \vdots \\ m_{n-1} & m_n & \dots & m_{2n-1} \\ 1 & x & \dots & x^n \end{vmatrix}, \text{ gdzie } K \neq 0.$$

Wymieńmy teraz dwie ważne własności wielomianów ortogonalnych.

Twierdzenie 1. Wielomian p_n , stopnia n , ortogonalny do wszystkich wielomianów stopnia $< n$ ma n różnych pierwiastków leżących w przedziale $[a, b]$.

Twierdzenie 2. Jedynie wielomian ortogonalny p_n , z najstarszym współczynnikiem równym 1, ma własność minimalizowania normy

$$\int_a^b p_n^2(x) \mu(dx) = \min_{w \in \mathcal{W}_n} \int_a^b w^2(x) \mu(dx).$$

Dowody tych faktów można znaleźć w książce P.-J. Laurenta [13].

Powyższe twierdzenia dowodzą, że wielośrednia (wprowadzona w definicji 1) zawsze istnieje, jest to po prostu zbiór miejsc zerowych odpowiedniego wielomianu ortogonalnego.

Przykład

Wielomiany Hermite'a, $H_n(x)$ $n \in \mathbf{N}$, są jedną z bardziej znanych rodzin wielomianów ortogonalnych. Spełniają równanie różniczkowe

$$y'' - 2xy' + 2ny = 0$$

oraz relację ortogonalności

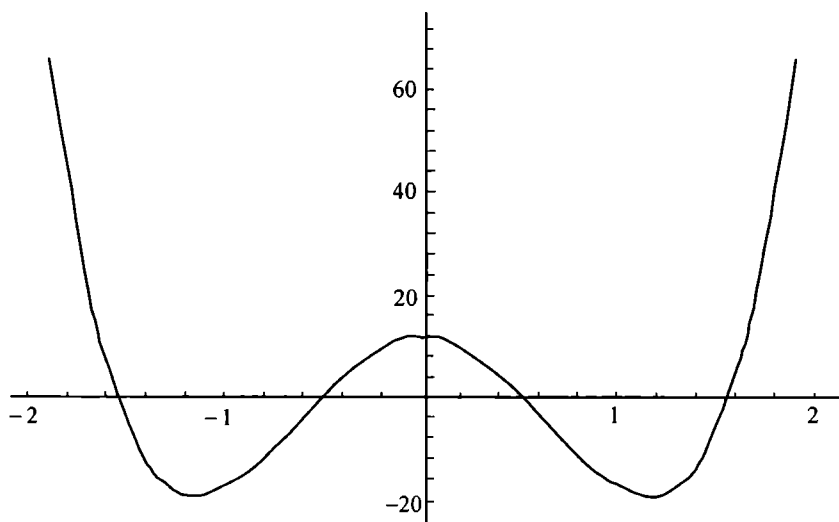
$$\int_{-\infty}^{\infty} H_m(x) H_n(x) e^{-x^2} dx = 0, m \neq n.$$

Widać zatem ich związek z rozkładem normalnym. Można zdefiniować je za pomocą wzoru

$$H_n(x) = n! \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k (2x)^{n-2k}}{k!(n-2k)!}, n \in \mathbf{N}, x \in \mathfrak{R}.$$

Aby znaleźć 4-średnią (M_1, M_2, M_3, M_4) prostej rzeczywistej z miarą $\mu(dx) = e^{-x^2} dx$, należy wyznaczyć pierwiastki wielomianu

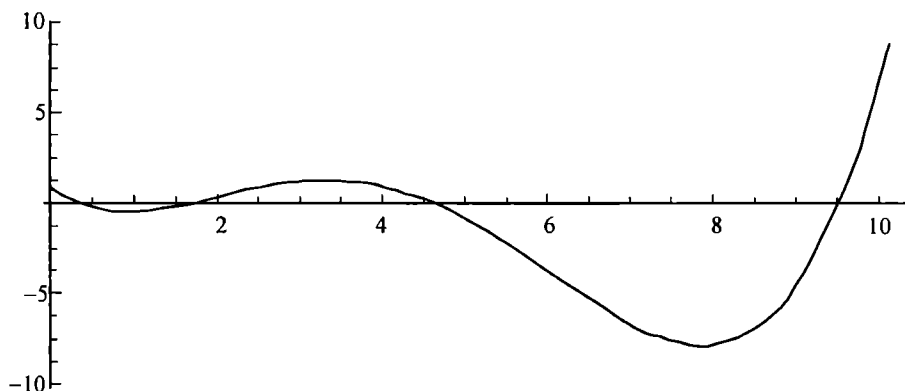
$$H_4(x) = 16x^4 - 48x^2 + 12.$$



Rys. 2. Wielomian $H_4(x)$

Otrzymujemy $(M_1, M_2, M_3, M_4) = (-1,651, -0,525, 0,525, 1,651)$. Inną powszechnie znaną rodziną wielomianów ortogonalnych są wielomiany Laguerre'a $L_n^\alpha(x)$, $n \in \mathbb{N}$, $\alpha \in \mathbb{R}$. Spełniają one równanie różniczkowe

$$xy'' + (\alpha + 1 - x)y' + ny = 0$$



Rys. 3. Wielomian Laguerre'a $L_n^\alpha(x)$

oraz relację ortogonalności

$$\int_0^{+\infty} L_m^\alpha(x) L_n^\alpha(x) x^\alpha e^{-x} dx = 0, m \neq n.$$

Wynika z niej ich związek z rozkładem gamma. Gdy $\alpha = 0$, wówczas można je zdefiniować wzorem

$$L_n(x) = \sum_{k=0}^n \frac{(-n)_k x^k}{(k!)^2}, n \in N, x \in [0, +\infty),$$

$(-n)_k = \frac{\Gamma(-n+k)}{\Gamma(-n)} = (-n+k-1)(-n+k-2)\dots(-n+2)(-n+1)$ – symbol Pochhammera.

Do znalezienia 4-średniej, dodatniej półprostej z miarą $\mu(dx) = e^{-x} dx$ trzeba obliczyć miejsca zerowe wielomianu

$$L_4(x) = \frac{1}{24}(x^4 - 16x^3 + 72x^2 - 96x + 24).$$

Otrzymujemy $(M_1, M_2, M_3, M_4) = (0,324, 1,753, 4,54, 9,402)$.

Literatura

- [1] Antoniewicz R., *O biprzeciętnej*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 857, AE, Wrocław 2000.
- [2] Antoniewicz R., *O średnich i przeciętnych*, AE, Wrocław 2005.
- [3] Antoniewicz R., Misztal A., *Biprzeciętna*, „Przegląd Statystyczny” 2001, R. XLVIII, Zeszyt 3-4.
- [4] Antoniewicz R., Misztal A., Wilkowski A., *O pewnej własności biprzeciętnej*, „Przegląd Statystyczny” 2004, R. LI, Zeszyt 4.
- [5] Antoniewicz R., Wilkowski A., *O pewnym rozkładzie dwumodalnym*, „Przegląd Statystyczny” 2004, R. LI, Zeszyt 1.
- [6] Bateman H., Erdelyi A., *Higher Transcendental Functions*, McGraw-Hill Book Company, New York 1953.
- [7] Denisov S., Simon B., *Zeros of Orthogonal Polynomials on the Real Line*, „Journal of Approximation Theory” 2003, 121.
- [8] Hennequin P., Tortrat A., *Teoria prawdopodobieństwa i niektóre jej zastosowania* (tłumaczenie z francuskiego), „Nauka”, Moskwa 1974.
- [9] Karpow K., Czystowa E., *Tablice funkcji Webera*, Centrum Obliczeniowe Akademii Nauk ZSRR, Moskwa 1968.
- [10] Kolmogoroff A., *Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung*, Math. Ann. 1931, t. 104.
- [11] Krasikov I., *Bounds for Zeros of the Laguerre Polynomials*, „Journal of Approximation Theory” 2003, 121.

-
- [12] Kung-Yu Ch., Srivastava H.M., *A Limit Relationship between Laguerre and Hermite Polynomials*, „Integral Transforms and Special Functions” 2005, vol. 16, no. 1.
 - [13] Laurent P.-J., *Aproksymacja i optymalizacja* (tłumaczenie z francuskiego), Wydawnictwo „Mir”, Moskwa 1975.
 - [14] Ostasiewicz W., *Propedeutyka probabilistyki*, AE, Wrocław 2000.
 - [15] Sikorski R., *Funkcje rzeczywiste II*, Monografie Matematyczne, tom 37, PWN, Warszawa 1959.
 - [16] Smoluk A., *Uwagi o definicji wartości modalnej*, „Przegląd Statystyczny” 1997, R. XLIV, Zeszyt 1.
 - [17] Szafraniec F.H., *Przestrzenie Hilberta z jądrem reprodukcującym*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2004.
 - [18] Wywiał J., *Estimation of Mode on the Basis of a Truncated Sample*, Acta Universitatis Lodziensis, Folia Oeconomica 152, Łódź 2000.

MULTI-AVERAGE

Summary

We present bi-average in this paper. We also give example of bi-mode distribution. We define multi-average too, and its relationship with orthogonal polynomials.

Andrzej Wilkowski – dr, adiunkt w Katedrze Matematyki i Cybernetyki Akademii Ekonomicznej we Wrocławiu.