Tandetna seria tekstów matemetycznych

Ciekawe rzeczy o średnich

19 sierpnia 2024

1 Wstęp

Ala pije po 2 piwa w soboty i niedziele a Bob — 1 piwo w patrzyste dni i pół w nieparzyste. Kto pije więcej?
Rozwiązanie. Ala pije średnio $\frac{0+0+0+0+0+2+2}{7} = \frac{4}{7}$ piwa dziennie a Bob — $\frac{1+0.5}{2} = \frac{3}{4}$. Oczywiście $\frac{3}{4} > \frac{4}{7}$ więc Bob pije więcej.
Ten przykład pokazuje do czego można użyć średnich w prawdziwym życiu. Teraz do rzeczy:

2 Średnia

To słowo bez kontekstu oznacza średnią arytmetyczną, czyli

$$\bar{a} = \frac{a_1 + a_2 + \cdots + a_n}{n}$$

albo $\frac{1}{n} \sum a_i$. Elementy ciągu a nazywamy *danymi*. Ewentualnie jeśli nie chodzi nam o ciąg tylko o całą funkcję o dziedzinie $(0, \Delta x)$:

$$\bar{f} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} f(x_i) = \frac{1}{\Delta x} \int_0^{\Delta x} f(x)dx = \frac{\Delta F}{\Delta x}$$

dla $x_i = \frac{i}{n} \cdot \Delta x$ i $F'(x) = f(x)$. Przedział $(0, \Delta x)$ można zastąpić przez (a, b) ale wtedy wzory będą brzydsze. Zwykle nazywa się to wartością średnią funkcji $f(x)$, np jeśli przyjmiemy taką zmianę oznaczeń:

$$\text{współrzędna } x \to \text{czas } t$$
$$\text{funkcja } f(x) \to \text{prędkość } v(t)$$
$$\text{funkcja } F(x) \text{ której pochodną jest } F'(x) = f(x) \to \text{droga } s(t) \text{ której pochodną jest } s'(t) = v(t)$$
$$\bar{f} = \frac{1}{\Delta x} \int_0^{\Delta x} f(x)dx = \frac{\Delta F}{\Delta x} \to v_{\text{śr}} = \frac{1}{\Delta t} \int_0^{\Delta t} v(t)dt = \frac{\Delta s}{\Delta t}$$

. . . to będziemy ją nazywać średnią wartością prędkości albo prędkością średnią. Tak samo działa to dla średniego ciśnienia ($\frac{\Delta F}{\Delta S}$), natężenia prądu ($\frac{\Delta q}{\Delta t}$), gęstości ($\frac{\Delta m}{\Delta V}$) itd.
Dalej będziemy oznaczać średnie funkcji przez $\bar{a}$, a średnie danych $a_1, a_2, \cdots, a_n$ przez $\mu(a_1, a_2, \cdots, a_n)$

2.1 Średnia ważona

Kiedy chcemy żeby ta sama dana trafiła do średniej ileś razy, tak jak trójka przy $\frac{3+3+3+3+2+2+5}{7}$ to przypisujemy jej wagę, w tym wypadku 4 bo są 4 trójki. Ogólnie wzór na średnią ważoną wygląda tak:

$$\bar{a} = \frac{\sum w_i a_i}{\sum w_i}$$

Gdzie w_i jest wagą danej a_i, która domyślnie wynosi 1.

2.2 Łączenie średnich

jeśli weźmiemy ciąg $(a_1, a_2, \cdots)$, po czym podzielimy go na n podciągów równej długości $(a_1, \cdots, a_n), (a_{n+1}, \cdots, a_{2n}), \cdots$ o średnich odpowiednio $s_1, s_2, \cdots$, to wartość średnia całego ciągu wynosi

$$\bar{a} = \frac{s_1 + s_2 + \cdots}{n}$$

Jeżeli ich długości nie będą takie same, tylko i-ty podciąg będzie miał długość ℓ_i, to

$$\bar{a} = \frac{\ell_1 s_1 + \ell_2 s_2 + \cdots}{\ell_1 + \ell_2 + \cdots}$$

Jeżeli dodatkowo dodamy wagi, a suma wag w i-tym podciąg będzie miał wynosić W_i, to

$$\bar{a} = \frac{W_1 s_1 + W_2 s_2 + \cdots}{W_1 + W_2 + \cdots}$$

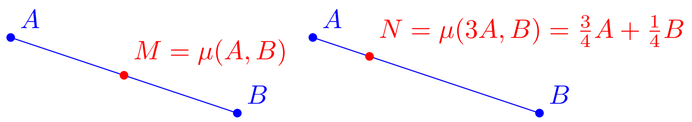
W skrócie chodzi o to, że średnia średnich jest średnią, tak samo jak suma sum jest sumą. W ten sposób można definiować średnie dla rzeczy których nie da się zsumować.

2.3 Uśrednianie punktów

Średnia punktów A i B to środek odcinka AB. Za to ich średnia ważona o wagach w_A i w_B to taki punkt M na tym odcinku że sotsunek jego odległości od każdego z końców jest taki sam jak stosunek wag, tj.

$$\frac{MB}{MA} = \frac{w_A}{w_B}$$

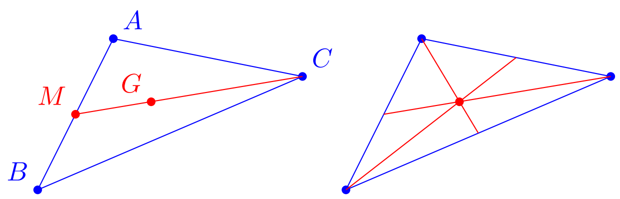
Obrazek:



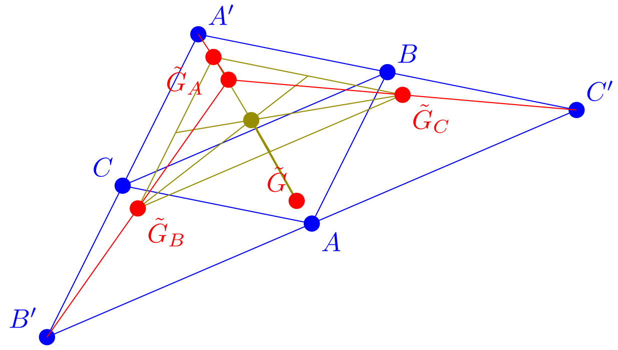
Jeśli stosunek $\frac{w_A}{w_B}$ jest wymierny, to można równie dobrze korzystać ze wzorów na nie-ważoną średnią i łączenie średnich, np N można policzyć jako $\mu(A, M)$. Oczywiście, to samo wyszłoby gdybyśmy uśrednili współrzędne punktów (tzn. współrzędna x punktu m to śrenia współrzędnych x punktów A i B), o ile mamy jakiś układ współrzędnych.
Czasami średnią 2 punktów nazywa się (uwaga długie słowo) interpolacją liniową (LERP-em od *linear interpolation*); wtedy zamiast wag podaje się ile % ich sumy zajmuje 1 z nich.
Za to średnią 3 punktów liczymy tak:
Twierdzenie 1. Odcinki łączące środki boków trójkąta z wierzchołkami naprzeciwko nich przecinają się w 1 punkcie, który:
1) istnieje,
2) dzieli każdy z nich w stosunku 1:2,
3) jest średnią wierzchołków trójkąta (barycentrum).
Dowód. Oznaczmy wierzchołki A, B, i C w dowolnej kolejności. Śrenia wierzchołków A i B to środek odcnika AB (dalej M), a średnia wszystkich wierzchołków (G) to średnia M z wagą 2 i C z wagą 1, tj.

$$G = \frac{2M + C}{3}$$

A więc leży na odcinku MC, 2 razy bliżej do M niż do C. Ponieważ wierzchołki oznaczyliśmy dowolnie a punkt G jest tylko 1, to odcinki z MC i spółka przecinają się w 1 punkcie $\square$.



Podobnie to liczymy dla n punktów; tutaj ciekawym faktem jest że ich średnia zawsze będzie we wnętrzu ich otoczki wypukłej, tzn figury o najmniejszym obwodzie która je zawiera. Dowód jest łatwy przez indukcję. Poza tym:
Twierdzenie 1 ciąg dalszy. Punkt G:
4) jest średnią wszystkich jego punktów (tzn z wewnętrzem).
Dowód.
Niech trójkąt ABC przystaje do trójkątów $A'CB$, $B'AC$ i $C'BA$. Oznaczmy średnią jego punktów przez $\tilde{G}$, a średnie punktów jego kopii — $\tilde{G}_A$, $\tilde{G}_B$ i $\tilde{G}_C$. ABC jest podobny w skali 2 do $A'B'C'$ a więc średnia jego punktów musi leżeć 2 razy dalej od A' niż $\tilde{G}_A$, 2 razy dalej od B' niż $\tilde{G}_B$ i 2 razy dalej od C' niż $\tilde{G}_C$, czyli na przecięciu prostych $A'\tilde{G}_A$, $B'\tilde{G}_B$ i $C'\tilde{G}_C$. Jednak musi też leżeć na odcinku łączącym barycentrum trójkąta $\tilde{G}_A\tilde{G}_B\tilde{G}_C$ i punkt $\tilde{G}$ — a to jest niemożliwe, a dlaczego proszę samemu się domyśleć :)



d-wymiarowe uogólnienie trójkąta (jako otoczki wypukłej $d + 1 = 3$ niewspółliniowych punktów) nazywamy d-sympleksem. Ogólniejsza wersja twierdzenia 1:
Twierdzenie 1.1. Średnia wierzchołków (przy okazji wszystkich punktów) d-sympleksu leży na przecięciu odcinków łączących jego wierzchołki z barycentrami ścian naprzeciwko, i dzieli te odcinki w stosunku $1 : d$.
Dowód bo jest taki sam jak dla trójkątów.

3 f-średnia

Czasami trzeba uśrednić nie zmienną a coś co od niej zależy, np często używa się wartości napięcia (U) dla którego moc ($P = U^2/R$) będzie miała swoją średnią wartość (akurat tu nazywa się to wartością skuteczną napięcia). W tym wypadku uśredniamy (na przedziale długości T) kwadrat wartości funkcji $u(t)$, a potem bierzemy odpowiednią wartość u, tj

$$U_s = \sqrt{\overline{u^2(t)}} = \sqrt{\frac{1}{T}\int_0^T (u(t))^2 dt}$$

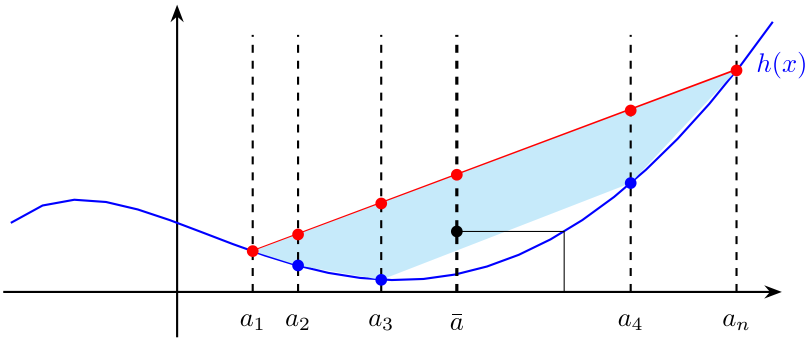
i ogólnie oprócz kwadratowania można to robić z każdą inną funkcją (o ile jest odwracalna):

$$\mu_f(a) = f^{-1}\left(\overline{f(a)}\right)$$

3.1 Nierówności między f-średnimi

Na potrzeby tego pdf-u $\mu_f > \mu_g$ będzie oznaczać $\mu_f(a_1, a_2, \cdots, a_n) > \mu_g(a_1, a_2, \cdots, a_n)$ dla dowolnej ilości dowolnych danych.
Twierdzenie 2. $\mu_f > \mu_g$ **jeśli f jest funkcją rosnącą a $g(f^{-1}(x))$ — wypukłą**
Dowód. Niech:
1) danymi będą $a_1, \cdots, a_n$,
2) $\alpha_i = f^{-1}(a_i)$,
3) $g(f^{-1}(x)) = h(x)$ (oczywiście jest funkcją wypukłą).

Wtedy $f(\mu_f(\alpha_1, \cdots, \alpha_n)) = \mu(a_1, \cdots, a_n)$, a $f(\mu_g(\alpha_1, \cdots, \alpha_n)) = \mu_h(a_1, \cdots, a_n)$. Rysujemy:
1) wykres funkcji h,
2) n prostych, gdzie i-ta z nich ma równanie $x = a_i$,
3) punkty (w sensie kropki) $A_1, \cdots, A_n$ w przecięciach prostych z wykresem,
4) punkty $B_1, \cdots, B_n$ w przecięciach prostych z odcinkiem A_1A_n,
5) prostą o równaniu $x = \mu(a_1, \cdots, a_n)$ przecinającą odcinek w punkcie M.



Punkt M jest śrenią punktów $B_1, \cdots, B_n$, bo leży na odcinku B_1B_n i jego współrzędna x się zgadza (jest tylko 1 taki punkt). Liczba $\mu_h(a_1, \cdots, a_n)$ to taka x że $h(x)$ jest współrzędną y średniej punktów $A_1, \cdots, A_n$. Ta średnia leży pod punktem M (bo ma tą samą współrzędną x i leży w otoczce wypukłej punktów $A_1, \cdots, A_n$), a więc prosta pozioma przechodząca przez nią przecina wykres funkcji h gdzieś po prawo, tj. dla większej wartości x niż współrzędna x punktu M. A ponieważ ta pierwsza wynosi $f(\mu_g(\alpha_1, \cdots, \alpha_n))$ a druga — $f(\mu_f(\alpha_1, \cdots, \alpha_n))$ (a f jest rosnąca), to istotnie $\mu_f(\alpha_1, \cdots, \alpha_n)) > \mu_g(\alpha_1, \cdots, \alpha_n)$ □
To samo zadziała jeśli f będzie malejąca a h — wklęsła; za to oba warunki na raz można zapisać $f' \cdot (g \circ f^{-1})' > 0$. Jeśli f i g są ciągłe (zazwyczaj są) i odwracalne (muszą takie być żeby dało się robić f-średnią) to muszą być albo rosnące albo malejące w całej dziedzinie, a warunek $f' \cdot (g \circ f^{-1})' > 0$ robi się (tak to się ładnie nazywa) konieczny i wystarczający.
Teraz kilka uszczególnień, na ich potrzeby będziemy używali takiego symbolu:
• $\mu_a := \mu_{x^a}$ dla $a \neq 0$,
• $\mu_0 := \lim_{a \to 0} \mu_a = \mu_{\ln(x)} = \sqrt[n]{a_1 a_2 \cdots a_n}$ (średnia geometryczna),
• $\mu_{+\infty} := \lim_{a \to +\infty} \mu_a = \max a_1, a_2, \cdots, a_n$ (maksimum),
• $\mu_{-\infty} := \lim_{a \to -\infty} \mu_a = \min a_1, a_2, \cdots, a_n$ (minimum).
Twierdzenie 3. $\mu_a > \mu_b$ **iff** $a > b$
Dowód dla $b \neq 1$**:**
oczywiście funkcja x^a **jest rosnąca a funkcja** x^{a-b} **jest wklęsła iff** $a > b$ □
Dowód dla $b = 1$**:**
funkcja e^{ax} **jest wklęsła iff** $a > 0$ □
Dowód dla $b = +\infty$**:**
niech ciąg $a_1, a_2, \cdots, a_n$ **będzie nierosnący. Wtedy** $\frac{1}{\sqrt{n}} \sqrt[a]{a_1, a_2 \cdots, a_n} < \frac{1}{\sqrt[a]{n}} \sqrt[a]{a_1, a_1, \cdots, a_1} = a_1$ **□.**
Dowód dla $b = -\infty$**:**
niech ciąg $a_1, a_2, \cdots, a_n$ **będzie niemalejący. Wtedy** $\frac{1}{\sqrt{n}} \sqrt[a]{a_1, a_2 \cdots, a_n} > \frac{1}{\sqrt[a]{n}} \sqrt[a]{a_1, a_1, \cdots, a_1} = a_1$ **□.**
Jeśli $a_1 = a_2 = \cdots = a_n$ to zamiast znaku $<$ mamy znak $=$ (uprzednio zakładaliśmy że dane nie są wszystkie takie same).
W szczególności:

$$\mu_{x^2} > \mu_x > \mu_{\ln(x)} > \mu \frac{1}{x}$$

innymi słowy

$$K > A > G > H$$