Przegląd Statystyczny T. I Nr 1
Warszawa 1938

Jan Wiśniewski
Warszawa
1904-1939

Uwagi o definicji przeciętnej

Referat wygłoszony na zebraniu naukowym sekcji statystyki matematycznej Polskiego Towarzystwa Statystycznego dn. 13 grudnia 1937.

Zadaniem niniejszej pracy nie jest wyczerpujące przedstawienie zagadnienia definicji przeciętnej[1]), ani też danie takiej definicji, która by zaspokoiła wszelkie wymagania. Autor ogranicza się do wskazania niektórych trudności, związanych z definiowaniem przeciętnych, oraz konsekwencji pewnych szerzej znanych definicji.

Jako przeciętne będziemy pojmowali w niniejszej pracy klasę miar statystycznych (cech zbiorczych w rozumieniu Orzęckiego), zawierającą przynajmniej jeden element, a mianowicie średnią arytmetyczną. Bliższe sprecyzowanie własności elementów tej klasy stanowi właśnie definicję przeciętnej (uogólnionej). Forma definicji może się zmieniać w zależności od tego, czy mamy do czynienia z przeciętną skończonej liczby zmiennych, czy też z przeciętną rozkładu ciągłego. Ponieważ traktujemy tutaj przeciętne jako miary czysto opisowe, nie sprawia różnicy, czy chodzi o rozkład teoretyczny czy też zaobserwowany. Stale też będziemy używali wyrażenia „wielkość zaobserwowana" lub po prostu „obserwacja" dla oznaczenia jednej z wielkości, których przeciętnej szukamy.

W podręcznikach statystyki znajdujemy werbalne definicje przeciętnej. Yule[2]) pisze: „Miary, które się odnoszą do... po-

---

[1]) Bogatą literaturę przedmiotu podaje E. L. Dodd, Colorado College Publication, General Series Nr 208, Colorado Springs, Sept. 1936, str. 89 sq.

[2]) „Wstęp do teorii statystyki", tłum. polskie, str. 127.

łożenia, t.z... do wartości zmiennej, w pobliżu której spostrzeżenia sięześrodkowują, znane są ogólnie pod nazwą przeciętnych". Jakkolwiek definicja taka może być bardzo nawet intuicyjna, to jednak obecnie zajmiemy się definicjami o charakterze formalnym.

Najogólniejszą zapewne definicję przytacza E. L. Dodd[1]): funkcję $f(x_1, x_2, \ldots, x_n)$ uważa się za przeciętną wielkości $x_1, x_2, \ldots, x_n$, jeżeli choć jedna wartość $f(c, c, \ldots c)$ jest równa c.

Edward V. Huntington[2]) ustala pięć postulatów, które powinna spełniać przeciętna; ogranicza się on do rozpatrywania przeciętnych liczb dodatnich.

$$\left.\begin{array}{ll}
\text{I.} & f(x_1, x_2, \ldots, x_i, x_j, \ldots, x_n) = \\
& \qquad f(x_2, x_1, \ldots, x_j, x_i, \ldots, x_n). \\
\text{II.} & f(x_1, x_2, x_3, \ldots x_n) = f(m, m, x_3, \ldots x_n) \\
& \qquad \text{gdzie } m = f(x_1, x_2). \\
\text{III.} & f(kx_1, kx_2, \ldots, kx_n) = \\
& \qquad kf(x_1, x_2, \ldots, x_n),\ k > 0. \\
\text{IV.} & f(a, a, \ldots, a) = a. \\
\text{V.} & f(x_1, x_2, \ldots, x_n) > 0.
\end{array}\right\} \quad (1)$$

System Huntingtona jest zasadniczo stosowalny do rozkładów arytmetycznych (nieciągłych), lecz niektóre postulaty, np. III i V, można — z drobnymi modyfikacjami — dostosować i do rozkładów ciągłych.

G. Darmois[3]) pisze: „On peut définir une moyenne relative à la fonction $f(x)$ par la formule:

$$f(\xi) = \frac{\Sigma p_i f(x_i)}{\Sigma p_i} \qquad (2)$$

gdzie p_i oznacza liczebność zaobserwowanej wartości x_i, zaś ξ przeciętną oznaczoną przy pomocy funkcji f, którą nazwiemy funkcją przekształcającą. Wprowadzając symbol $A(z)$ na oznaczenie średniej arytmetycznej z-ów, możnaby też napisać (2) w ten sposób:

[1]) Cowles Commission for Research in Economics, Report of Third Annual Research Conference on Economics and Statistics, June 28 to July 23, 1937, Colorado Springs; str. 15.

[2]) Transactions of the American Mathematical Society, 1927, pp. 1—22. Profesor A. Łomnicki był łaskaw zwrócić moją uwagę na tę pracę.

[3]) „Statistique mathématique", Paris, 1928, p. 31.

$$f(\xi) = A\,[f(x)] \tag{2'}$$

Oczywiście można zbudować wyrażenie analogiczne do (2), zamiast sum zawierające całki. Wyrażenie to będzie nam służyło do definiowania przeciętnych rozkładów ciągłych.

$$f(\xi) = \int_{-\infty}^{\infty} f(x)\,y(x)\,dx \tag{2''}$$

$y(x)$ oznacza tu częstość x. $f(x)$ musi być zarówno dla rozkładów ciągłych jak i nieciągłych funkcją odwracalną, aby można było jednoznacznie określić ξ.

Obecnie rozpatrzymy stosunek niektórych przeciętnych nieklasycznych do przytoczonych definicji. Przeciętne klasyczne, a mianowicie arytmetyczna, geometryczna, harmoniczna i kwadratowa podpadają pod wszystkie definicje, o których była mowa. Ponieważ definicja przytoczona przez Dodda jest identyczna z IV postulatem Huntingtona, ograniczymy się do rozpatrywania definicji Huntingtona i Darmois. Na ogół łatwiej jest okazać, że pewna przeciętna jest zgodna lub niezgodna z definicją Huntingtona niż z definicją Darmois, w stosunku zaś do tej ostatniej łatwiej jest wykazać zgodność niż niezgodność. W pierwszym bowiem wypadku wystarczy znaleźć pewną określoną funkcję, w drugim trzeba dowieść, że nie istnieje żadna funkcja, spełniająca przepisane warunki. Z tego powodu będziemy się ograniczali do stwierdzenia, że nie skonstatowano, aby ta i ta funkcja była zgodna z definicją Darmois.

M e d i a n a. Huntington wykazuje, że mediana jest niezgodna z post. II. Rzeczywiście, niech mediana M będzie środkowym wyrazem ciągu liczb uporządkowanych według wielkości, gdy ich liczba n jest nieparzysta, zaś średnią arytmetyczną dwóch środkowych wyrazów, gdy n jest parzyste. $M(2, 4, 7) = 4$, lecz $M(2, 4) = 3$ zaś $M(3, 3, 7) = 3$. Z drugiej jednak strony mediana podpada pod definicję Darmois. Weźmy najpierw pod uwagę rozkłady nieciągłe. Jako funkcję przekształcającą zastosujemy tu funkcję $N(x)$, która dla danego x_i równa się numerowi porządkowemu x_i, gdy zaobserwowane wartości x są uporządkowane wg wielkości. Jeśli x_k nie jest zaobserwowaną wartością x lecz leży między dwiema takimi wartościami x_r i x_s, wówczas

$$N(x_k) = N(x_r) + \frac{x_k - x_r}{x_s - x_r} \tag{3}$$

Średnia arytmetyczna wartości $N(x)$ równa się, oczywiście, $\frac{1}{2}(n+1)$ zaś odpowiadająca jej wartość x jest medianą. W przypadku rozkładów ciągłych jako funkcję przekształcającą weźmy

$$Y(x_i) = \int_{-\infty}^{x_i} y(x)\,dx \tag{4}$$

Średnia arytmetyczna tej funkcji równa się $\frac{1}{2}$. Jak widać, mediana daje przykład funkcji zgodnej z definicją Darmois lecz niezgodnej z definicją Huntingtona.

Na odwrót, największa lub najmniejsza wartość w skończonym zbiorze liczb odpowiada definicji przeciętnej Huntingtona, nie stwierdziliśmy natomiast, aby odpowiadała ona definicji Darmois [1].

Średnia arytmetyczna ważona. Oznaczamy ją

$$W(x_1, x_2, \ldots, x_n) = \frac{\Sigma w_i x_i}{\Sigma w_i} \tag{5}$$

Wielkości w są funkcjami x-ów o tyle, że każdemu x_i przyporządkowana jest pewna wartość w_i, jednakże z równości $x_i = x_j$ nie wynika samo przez się $w_i = w_j$. Dlatego było by może lepiej powiedzieć, że zarówno x_i jak i w_i są jednoznacznymi funkcjami wskaźnika i; $i = 1, 2, 3, \ldots n$. Widzimy od razu, że średnia arytmetyczna ważona, zdefiniowana jak powyżej, zgadza się z definicją Huntingtona, aby zaś uzgodnić ją z definicją Darmois, trzeba uzupełnić definicję średniej ważonej jak następuje: waga średniej arytmetycznej ważonej równa się średniej arytmetycznej wag. Niech teraz funkcja przekształcająca będzie

$$f(x_i) = w_i x_i \tag{6}$$

Wówczas

$$A[f(x)] = \frac{1}{n} \Sigma w_i x_i \tag{7}$$

Z drugiej strony, niech W będzie średnią arytmetyczną ważoną zdefiniowaną przez (5). Waga, którą nadamy W, jest

[1] P. dyskusja po referacie.

$$w_W = \frac{1}{n} \Sigma w_i \tag{8}$$

a więc

$$f(W) = A[f(x)] \tag{9}$$

co chcieliśmy okazać.

Sytuacja zmienia się, gdy wagi stają się określonymi funkcjami x. Jeśli, w szczególności, przyjmiemy $w(x) = x$, otrzymamy t. zw. średnią k o n t r a - h a r m o n i c z n ą [1]).

$$Ch = \frac{\Sigma x_i^2}{\Sigma x_i} \tag{10}$$

Przeciętna ta nie jest zgodna z postulatem II Huntingtona. Istotnie, $Ch(1,2,3) = \frac{14}{6}$. Lecz $Ch(1,2) = \frac{5}{3}$; $Ch\left(\frac{5}{3}, \frac{5}{3}, 3\right) = \frac{131}{57} \neq \frac{14}{6}$. Ponieważ wagi są tu określonymi funkcjami obserwacyj, byłoby nielogiczne stosować tu „uzupełnienie" definicji średniej ważonej, a więc przekształcenie (6) nie może być zastosowane. Nie stwierdzamy, aby przeciętna Ch była zgodna z definicją Darmois.

Weźmy teraz pod uwagę t. zw. ś r e d n i ą a g r e g a t y w n ą

$$Ag\left(\frac{x_1}{y_1}, \frac{x_2}{y^2}, \ldots, \frac{x_n}{y_n}\right) = \frac{\Sigma x_i}{\Sigma y_i} \tag{11}$$

Przeciętna ta ma zastosowanie, gdy liczniki i mianowniki wartości zaobserwowanych są dane przez warunki zadania, oznaczając np. zaludnienie i powierzchnię poszczególnych powiatów albo też ceny różnych artykułów w dwóch porównywanych okresach czasu. Można też powiedzieć, że współczynnik regresji częściowej jest średnią agregatywną współczynników regresji obliczonych dla różnych wartości zmiennej eliminowanej, pod warunkiem, że regresje względem tej zmiennej są liniowe [2]). Wydaje się na pierwszy rzut oka rzeczą paradoksalną, że średnia agregatywna zdefiniowana przez (11) nie jest zgodna z postulatem II Huntingtona, natomiast staje się zgodna, gdy definicję jej przedstawimy jak następuje:

---

[1]) P. H. L. Rietz „H a n d b o o k of mathematical statistics", Boston 1924, str. 5.

[2]) Por. też: K. I w a s z k i e w i c z ó w n a „Uogólnienie metody korelacji cząstkowej etc." Kwartalnik Statystyczny, zesz. 3, 1932.

$$A\,g'\left(\frac{x_1}{y_1}, \frac{x_2}{y_2}, \ldots, \frac{x_n}{y_n}\right) = \frac{\frac{1}{n}\Sigma x_i}{\frac{1}{n}\Sigma y_i}. \tag{12}$$

Nie stwierdzamy, aby przeciętna ta odpowiadała definicji Darmois.

Weźmy np. jako funkcję przekształcającą

$$f\left(\frac{x_i}{y_i}\right) = x_i \tag{13}$$

Wówczas, rzeczywiście,

$$f(A\,g') = A\left[f\left(\frac{x}{y}\right)\right] \tag{14}$$

lecz także

$$f\left(\frac{\frac{1}{n}\Sigma x_i}{Z}\right) = A\left[f\left(\frac{x}{y}\right)\right] \tag{15}$$

dla dowolnego Z, np. $Z = \Sigma x^2$.

Jeden szczególny typ przeciętnej warto jeszcze może wziąć pod rozwagę. W zagadnieniach z zakresu korelacji częściowej współczynnik korelacji między x i y dla ustalonej wartości $z = z_i$ jest

$$r_i = \frac{p_i}{\sigma(x)_i \cdot \sigma(y)_i} \tag{16}$$

gdzie symbole po prawej stronie są użyte w znaczeniu Yule'owskim. Jeśli regresja x względem z i y względem z jest liniowa, współczynnik korelacji częściowej jest

$$r_{xy.z} = \frac{\Sigma p_i}{\sqrt{\Sigma \sigma^2(x)_i \,.\, \Sigma \sigma^2(y)_i}} \tag{17}$$

Zachodzi pytanie, czy $r_{xy.z}$ jest przeciętną poszczególnych wartości r_i. Używając terminu „przeciętna" ogólnikowo [1]), można by odpowiedzieć na to pytanie twierdząco, jednakże widzimy, że np. postulat IV Huntingtona nie jest spełniony. Będzie on spełniony, jeśli pojmiemy $r_{xy.z}$ jako przeciętną nie wartości r_i lecz współczynników regresji $b(x.y)_i$ i $b(y.x)_i$, gdzie

---

[1]) Por. J. Wiśniewski, Biometrika, vol. XXVII (1935), str. 357.

$$b(x\,.\,y)_i = \frac{p_i}{\sigma^2(x)_i} \qquad b(y\,.\,x)_i = \frac{p_i}{\sigma^2(y)_i} \tag{18}$$

Lecz teraz $r_{xy.z}$ staje się niezgodne z postulatem I, gdyż nie mamy prawa przemieniać współczynników $b(x\,.\,y)$ ze współczynnikami $b(y\,.\,x)$. Funkcja $r_{xy.z}$ nie okazuje się też zgodna z definicją Darmois.

W a r t o ś ć m o d a l n a. Zwykła jej definicja nie zawsze pozwala na ustalenie jej wartości liczbowej, wtedy np. gdy wszystkie liczebności są pomiędzy sobą równe. Jeśli sztucznie rozciągniemy definicję na takie przypadki, wartość modalna okaże się niezgodna z postulatem II. Powiedzmy np., że wartość modalna równa się średniej arytmetycznej takich wartości x, których liczebności (lub częstości) są między sobą równe i większe od liczebności (częstości) wszystkich innych wartości zaobserwowanych. Wówczas *Mod* (2, 4, 6) = 4, lecz *Mod* (2, 4) = 3, zaś *Mod* (3, 3, 6) = 3. Nie stwierdziliśmy również zgodności wartości modalnej z definicją Darmois.

Okażemy teraz, że j e s t zawsze m o ż l i w e skonstruować taką funkcję przekształcającą, ż e b y otrzymana przy jej pomocy p r z e c i ę t n a (zgodna z definicją Darmois) r ó ż n i ł a s i ę d o w o l n i e m a ł o o d n a j w i ę k s z e j lub najmniejszej w a r t o ś c i zaobserwowanej. Różnica ta jednakże zawsze pozostaje różna od zera. Dowód ograniczony jest do wypadku, gdy wszystkie obserwacje są dodatnie. Weźmy pod uwagę przeciętne typu $\sqrt[k]{M_k}$, gdzie M_k oznacza k-ty moment dokoła zera. Wiadomo, że $\sqrt[k]{M_k}$ jest funkcją rosnącą indeksu k. Do każdej dodatniej liczby D, byleby mniejszej od największej wartości zaobserwowanej [1]), można zawsze dobrać taką wartość $k = k_1$, że $\sqrt[k_1]{M_{k_1}} > D$. Istotnie, musi istnieć przynajmniej jedna wartość x, np. x_j, większa od D i posiadająca częstość y_j (dla rozkładów ciągłych: przedział x_j, x_m o częstości y_j). Oczywiście $M_k > y_j\, x_j^k$. Łatwo teraz znaleźć takie k_1, aby $y_j\, x_j = D^{k_1}$ oraz

$$\sqrt[k_1]{M_{k_1}} > x_j\, \sqrt[k_1]{y_j} > D,$$

[1]) Jeżeli wartość taka istnieje. Jeśli ona nie istnieje, rozróżnimy dwa wypadki: 1) istnieje liczba G_1, najmniejsza ze wszystkich liczb klasy G, od których wszystkie wartości zaobserwowane są mniejsze. Wtedy zmodyfikujemy swe twierdzenie w ten sposób, że położymy $D < G_1$. 2) Klasa liczb G jest pusta. Wtedy powiemy wprost, że D może być dowolnie wielkie.

co było do udowodnienia. Dla uczynienia $\sqrt[k]{M_k}$ dowolnie bliskim najmniejszej wartości obserwowanej, wystarczy skorzystać z ujemnych wartości k.

Na zakończenie udowodnimy jeszcze, że dla dowolnego rozkładu ciągłego jest możliwe skonstruować taką funkcję przekształcającą, aby przeciętna zgodna z definicją Darmois była zawsze równa z góry oznaczonemu percentylowi. Niech $y(x)$ oznacza częstość x, a

$$Y(x) = \int_{x_a}^{x} y(t)\,dt \tag{19}$$

Funkcja przekształcająca niech będzie następującej postaci:

$$f(x) = (m+1)\,Y^m(x) \tag{20}$$

Przeciętna określona przy jej pomocy będzie to taka wartość $x = x_p$, dla której

$$(m+1)\,Y^m(x_p) = \int_{x_a}^{x_b} (m+1)\,Y^m(x)\,y(x)\,dx = \int_0^1 (m+1)\,Y^m\,dY = 1 \tag{21}$$

gdzie x_a i x_b oznaczają krańce obszaru zmienności x. Okażemy, że x_p może przyjmować wszelkie wartości wnętrza przedziału x_a, x_b z wyjątkiem wartości x_p, dla której zachodzi $Y_p = Y(x_p) = \frac{1}{e}$. Istotnie

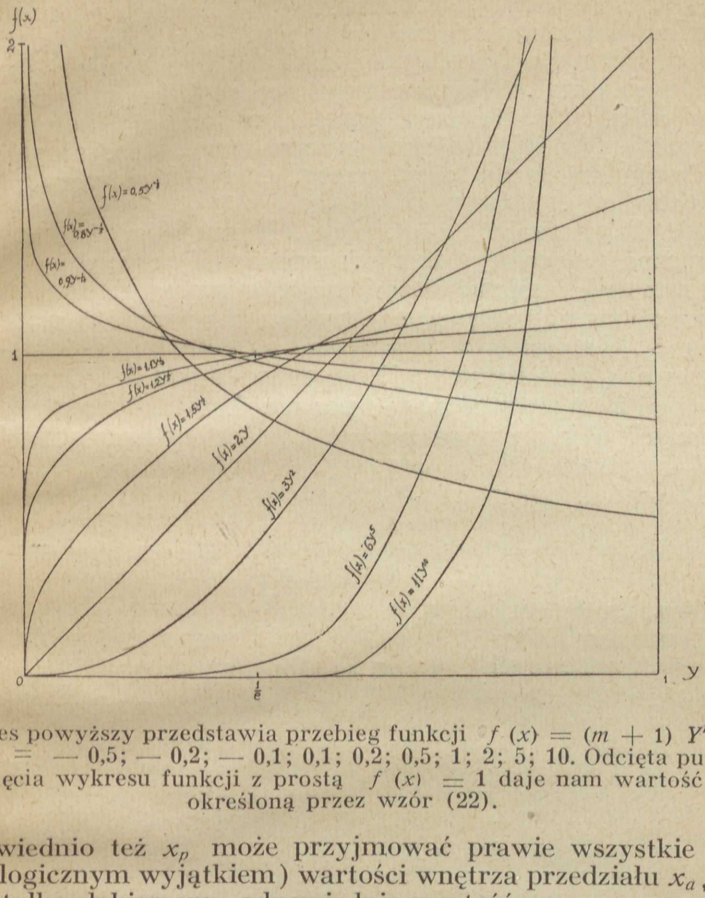
$$Y_p = \sqrt[m]{\frac{1}{m+1}} \tag{22}$$

jest ciągłą i rosnącą funkcją zmiennej m dla $-1 < m < 0$; $0 < m < \infty$; dla końców tych przedziałów zachodzi

$$\lim_{m \to -1} \sqrt[m]{\frac{1}{m+1}} = 0; \quad \lim_{m \to 0} \sqrt[m]{\frac{1}{m+1}} = \frac{1}{e}; \quad \lim_{m \to \infty} \sqrt[m]{\frac{1}{m+1}} = 1 \tag{23}$$

Widzimy więc, że Y_p może przyjmować wszelkie wartości wnętrza przedziału 0,1 z wyjątkiem ewentualnie wartości $\frac{1}{e}$ [1];

[1] Powstałą w ten sposób lukę można wypełnić przyjmując za funkcję przekształcającą $f(x) = (m+1)\,[1 - Y(x)]^m$



Wykres powyższy przedstawia przebieg funkcji $f(x) = (m+1)\, Y^m (x)$ dla $m = -0{,}5;\ -0{,}2;\ -0{,}1;\ 0{,}1;\ 0{,}2;\ 0{,}5;\ 1;\ 2;\ 5;\ 10$. Odcięta punktu przecięcia wykresu funkcji z prostą $f(x) = 1$ daje nam wartość Y_p określoną przez wzór (22).

odpowiednio też x_p może przyjmować prawie wszystkie (bo z analogicznym wyjątkiem) wartości wnętrza przedziału x_a, x_b, o ile tylko dobierzemy odpowiednią wartość na m.

Staraliśmy się przedstawić, jakie trudności nasuwa skonstruowanie jakiejś definicji przeciętnej „zdatnej do użytku“, tj. takiej, która by obejmowała to co potrzeba, a nie obejmowała zbyt wiele. Dotychczas bodajże definicji takiej nie udało się nikomu skonstruować; kto wie, czy to jest w ogóle możliwe.

Dyskusja nad referatem J. Wiśniewskiego

Dr. Michał Kerner.

1) Wyrażenie

$$-\sqrt{\frac{\Sigma x_i^2}{n}}$$

nie odpowiada definicji przeciętnej Huntingtona, gdyż przeczy nie tylko piątemu, ale także czwartemu postulatowi.

2) Definicja Darmois wymaga podania à priori funkcji $f(x)$ niezależnej od konkretnego szeregu. Dlatego zastosowanie tej definicji do mediany (przez wprowadzenie „numeru x") wykracza poza jej właściwe przeznaczenie. W szczególności def. Darmois pozwala łączyć szeregi, co się nie stosuje do przykładu mediany.

3) Jeżeli funkcja $f(x)$ w def. Darmois jest istotnie monotoniczna (czego wymaga jednoznaczność przeciętnej) i ciągła (warunek istnienia przeciętnej), wówczas łatwo dowieść, że przeciętna leży między najmniejszą i największą wartością zmiennej. Wynika stąd, że przykład na def. Huntingtona (największa z wartości zmiennej) nie podpada pod def. Darmois.

4) Rzuca się w oczy, że 4 postulat Huntingtona dla dwóch zmiennych daje się wyprowadzić z pozostałych. Istotnie:

L e m m a t. Dla każdego $m > 0$ istnieje x, że $f(x, x) = m$.

D o w ó d. Niech $z > 0$. Wówczas $f(z, z) = n > 0$. Z postulatu $f\left(z\frac{m}{n}, z\frac{m}{n}\right) = n\frac{m}{n} = m$. A więc $x = z\frac{m}{n}$. (3)

T w i e r d z e n i e. $f(m, m) = m$.

D o w ó d. Niech $m = f(x, x)$ (lemmat). Z postulatu 2 dla dwóch zmiennych $f(x, x) = f(m, m)$, gdzie $m = f(x, x)$. Stąd $m = f(m, m)$.

Jan Wiśniewski

Remarks on the definition of an average.

Read before the Mathematical Statistics Section of the Polish Statistical Society on December 13, 1937.

(Summary)

The purpose of this paper is to show the consistency or inconsistency of some types of averages, viz. of the median, of the mode, of the weighted arithmetic average, of the contraharmonic average, of the aggregatic average with two established definitions of average — that of Huntington and that of Darmois. In addition, two propositions are demonstrated: lst, that for a positive variate it is always possible to construct an average consistent with Darmois' definition which average would be as near as we please to the upper or lower bound of the variate; 2nd, that for any distribution it is always possible to construct an average consistent with Darmois' definition that would be equal to any percentile given in advance.