ISSN 2186-5647

－日本大学生産工学部第50回学術講演会講演概要（2017-12-2）－

2-8

比率の平均について

日大生産工(非常勤)　○篠原　正明

情報システム研究所　　篠原　健

1. はじめに

普通の一般的な平均（あるいは、個々のデータ値に加算が適している場合）では「算術平均」、指数、増加率など（個々のデータ値に乗算が適している場合）では「幾何平均」、速度では「調和平均」を採用すべし、等などと言われているが、その採用基準は定かでない。本論文では、2 つの物理量（例えば、距離と時間）の比率（速度＝距離／時間）に関する平均の計算（表現）式について考察する。

2．扱う問題

速度＝距離／時間の例について考える。n 区間からなる全行程において、距離 d_i〔m〕の区間 i を時間 t_i〔s〕で走行する際の全行程の距離 D〔m〕、総時間 T〔s〕は (1)、(2) となる。

$D=\Sigma d_i$ (1),　　　$T=\Sigma t_i$ (2)

区間 i の速度を v_i〔m/s〕、全行程の速度を V〔m/s〕とするならば、v_i, V は (3)、(4) となる。

$v_i = d_i/t_i$ (3),　　　$V=D/T$ (4)

本論文で扱う問題は、全区間の速度 V が区間 i の速度 $v_i (i=1,..,n)$によりどのような関数形の平均で表現できるか？である。この全区間の速度Ｖを、全区間の「平均」の速度と称する。この「平均」は、単なる算術平均 mean ではなく、より広い意味の代表値 average を 意味する。

3．物理量の比率とその平均

3.1　一般論

区間ｉの分子に相当する距離を d_i〔m〕、分母に相当する時間を t_i〔s〕とする。v_i〔m/s〕= d_i〔m〕/ t_i〔s〕(5)が区間ｉの比率データである。ここで、〔m〕,〔s〕は各物理量の単位、〔m/s〕はその比率としての単位である。又、速度＝距離／時間をイメージしたために、区間ｉと称したが、より一般的には標本ｉと考える。又、以下に一般論を展開するに当たり、分子を「距離」、分母を「時間」とする。

3.2　比率の具体例

表.1：物理量、特性量の比率の例

	比率物理量〔単位〕	分子物理量〔単位〕	分母物理量〔単位〕
*0	速度〔m/s〕	距離〔m〕	時間〔s〕
*1	処理速度〔ＭＩＰＳ〕	処理量〔命令数〕	時間〔s〕
*2	濃度〔%〕	食塩の重さ〔g〕	食塩水の重さ〔g〕
*3	人口密度〔人/m²〕	人口〔人〕	面積〔m²〕
*4	所要時間〔s/m〕	時間〔s〕	距離〔m〕
*5	平均点〔点/人〕	総点数〔点〕	人数〔人〕
*6	ドル→円為替レート	ドル額面〔ドル〕	円額面〔円〕
*7	処理時間比(j/k)	計算機ｊ上の処理時間[s]	計算機k 上の処理時間[s]

*1：単位時間当たりの物理量を○○速度と呼ぶ習慣あり。

*2：食塩を溶媒、食塩水を溶液と読めば、重量濃度となる。単位を百分率%としたが、少数表記で無単位となる。

*3：人口は人数であり、その単位は無単位である。分母物理量を距離、体積とすれば、１次元、３次元密度となる。

*4：単位距離の走行所要時間。*0 の速度を v_i, *4 の所要時間を u_iとすれば、$v_i=d_i/t_i$, $u_i= t_i/d_i$,で、$v_i\,u_i$＝１あるいは v_i＝$1/u_i$となる。

*5：(算術)平均点も総得点／人数と２つの特性量の比率として表現できる(4 章「シンプソンの逆理」参照)。

*6：為替交換上等価なドル額面Ｘ、円額面Ｙとすれば、Ｘ〔ドル〕＝Ｙ〔円〕（6）が成立する。なお、表.1 の記号に従えば、s＝ドル、m＝円であり、本例では両者共に名目単位である。この時、１〔ドル〕＝Z〔円〕＝Y/X〔円〕が成立し、Ｚをドルの対円為替レート、あるいはドル→円為替レートという(定義しよう)。具体例としては、１ドル＝120 円の時は、１〔ドル〕＝120〔円〕で、ドルの対円（ドル→円）為替レートは 120 となる。

*7：プログラム処理時間比率で言えば、分子がプログラムｉの計算機ｊ上での処理時間、分母がプログラムｉの計算機ｋ上での処理時間で、両者ともにプログラムｉという同じ処理内容に対して要した異なる計算機j,k 上での処理時間である。

On the average of ratios

Masaaki SHINOHARA and Ken SHINOHARA

3.3　物理量の比率の平均の表現式

標本(区間)数=n,　分子 d_i[m], 分母 t_i〔s〕,比率 $v_i(=d_i/t_i)$〔m/s〕(6)が与えられた時、総分子量D〔m〕＝Σd_i〔m〕(7)、総分母量T[s]＝Σt_i〔s〕(8)となり、速度V=D/T(9)が求めるべき「平均」である。この「平均」速度V〔m/s〕を個々の区間毎の平均 v_i[m/s]=d_i/t_i[m/s],i=1,..,n (10)で表現し、一般論を展開する。以降、[]内の単位は省略する。

加重平均の重みを $w_i(d)$, $w_i(t)$とする。

$w_i(d)=d_i/D$　(11),　　$w_i(t)=t_i/T$　(12)

$w_i(d)$は全距離Dに占める区間iの割合、$w_i(t)$は全時間Tに占める区間iの割合である。すると、比率の総合平均Vは(13)ならびに(14)で表現できる。

$V=D/T=\Sigma d_i/T=\Sigma(d_i/t_i)t_i/T=\Sigma w_i(t)(d_i/t_i)=\Sigma w_i(t)v_i$　(13)

$V^{-1}=T/D=\Sigma t_i/D=\Sigma(t_i/d_i)d_i/D=\Sigma w_i(d)(t_i/d_i)=\Sigma w_i(d)v_i^{-1}$　(14)

すなわち、比率の平均Vは、(13)に従えば分母物理量の区間iの重み $w_i(t)$付きの区間比率量 v_iの加重算術平均、(14)に従えば分子物理量の区間 i の重み $w_i(d)$付きの区間比率量 v_iの加重調和平均で表現できる。

3.4　考察

【1】d_i＝一定とすれば、$w_i(d)=1/n$ となり、(14)は調和平均となる。一定距離走行(例えば往復経路)の場合は、速度平均が調和平均で与えられるという既存の有名な結果と一致する。t_i＝一定とすれば、$w_i(t)=1/n$ となり、(13)は算術平均となり、一定時間走行では速度平均は算術平均で与えられる。

【2】濃度計算、密度計算においても、分子が一定ならば調和平均、分母が一定ならば算術平均となる。さらなる表現式(発展形)については、付録3を参照。

【3】(13),(14)より、$(\Sigma w_i(t)v_i)(\Sigma w_i(d)v_i^{-1})=1$ (15)

このように区間速度 v_iとその逆数 v_i^{-1}に関しては、加重算術平均と加重調和平均に関係した密接な関係がある。

4．シンプソンの逆理

4.1　単純な逆理の例

　男子学生bで数理sの人数、平均点をN(b,s)＝a, A(b,s)＝α (16)、女子学生gで数理sの人数、平均点をN(g,s)＝b, A(g,s)＝β (17)、男子学生bでマネジmの人数、平均点をN(b,m)＝c,　A(b,m)＝γ (18)、女子学生gでマネジmの人数、平均点をN(g,m)＝d,　A(g,m)＝δ (19)、とする。(i,j)層別の総得点は、S(b,s)=aα, S(g,s)＝bβ, S(b,m)＝cγ, S(g,m)=dδ (20)、と計算できる。

この時、もし「A(b,s)=α＞A(b,m)=γ (21), A(g,s)=β＞A(g,m)=δ (22)」ならば、「A(s)＝S(s)/N(s)＝(aα+bβ)/(a+b)＞A(m)＝S(m)/N(m)＝(cγ+dδ)/(c+d) (23)」が予想されるが、これが成立しないケースをシンプソンの逆理(の1例)と称する。学生を2つの添字i=b,sならびにj=s,mにより層別化したわけであるが、これに限定されない。

付録1の図A1-1にエクセル表計算シートを示すが、本例はa=8, b=2, c=1, d=9, α=20, β=90, γ=10, δ=80の例で、「α=20>10=γ, β=90>80=δ」が成立するが、「A(s)＝(aα+bβ)/(a+b)＝34<73=(cγ+dδ)/(c+d)=A(m)」となり、(23)相当の不等式が成立しない。付図A1－2にδ=36.6667でA(s)=A(m)となる例を示す。

4.2　物理量の比率への一般化

標本数N、総得点Sとした時、(算術)平均Aは, A=S/N (24)

すなわち、普通に「平均」という言葉を使用する時、それは、総得点Sと標本数Nの比率である。ここで、分子の総得点Sは総延距離,,等と問題対応に物理量として意味を持つ。さらに、分母のNも標本数に限定されない。すなわち、シンプソンの逆理は「物理量の比率」へと一般化できる。

4.3　速度版、比率データ版のシンプソン逆理

兎のA君は1日目は時速50㎞、2日目は時速100㎞で、亀のB君は1日目は時速40㎞、2日目は時速80㎞で、車を運転した。個々の運転時間、運転距離は不明である。但し、総運転距離は同じ。さて、A君とB君のどちらが速く、先に目的地に着くか?の問題を考える。………両日共に兎のA君の時速が亀のB君の時速を上回ってるので、兎のA君が亀のB君よりも先に目的地に着くと思うかもしれないが、必ずしも正しくない。B君は2日目の時速80㎞を活用すれば、A君よりも先に目的地に到着できる。例えば極端な例だが、A君は1日目は時速50㎞で8時間(400㎞走行)、2日目は時速100㎞で1時間(100㎞走行)、B君は1日目は時速40㎞で30分(20㎞走行)、2日目は時速80㎞で6時間(480㎞走行)、A君、B君走行距離は共に400㎞で同じだが、所要時間ではA君は9時間、B君は6時間半で、B君が先に目的地に到達してる。これは速度版シンプソン逆理の単純な例であり、さらなる条件が付与されても、同様な逆理例は容易に発生する。従って、算術平均値は総得点Sの標本数Nに対する比率である点を再考し、比率データを扱う際にはその分母、分子が何であるかに注意する必要があることを、シンプソン逆理は示唆する。これは、IRRのような比率かNPVのような総量で考えるかの経済性評価の議論とも共通する点がある。

５．逆比整合性と幾何平均

３章で議論したように、重み $w_i(t)$ あるいは $w_i(d)$ が既知の場合には、物理量、特性量の比率の平均として、加重算術平均あるいは加重調和平均の採用が適切である。しかし、重み情報が活用できない場合には、個別の平均 v_i の幾何平均を採用する方法が、ある種の整合性を持つ点を以下に主張する。

5.1　一般論

v_i は既知だが、d_i, $t_i(i=1,..,n)$ が入手不可の場合を考えよう。$v_i(i=1,..,n)$ の幾何平均 $G(v)$ は(25)で与えられる($v=\{v_i\}$)。

$G(v)=(\Pi v_i)^{1/n}$　　(25)

v_i を速度とすれば、その逆数 u_i は単位距離に対する所要時間を表す。　すなわち $u_i\cdot v_i=1$　(26)、これを速度と時間の逆比性と言う。所要時間 $u_i(i=1,..,n)$ の幾何平均 $G(u)$ は(27)で与えられる($u=\{u_i\}$)。　　$G(u)=(\Pi u_i)^{1/n}$　　(27)

幾何平均では、常に、$G(v)\cdot G(u)=1$　(28)が成立するが、3.3 節の適切な加重平均以外の他の平均操作では必ずしも「平均の逆比整合性：平均の逆数＝逆数の平均」は成立しない。すなわち、個別速度 v_i の幾何平均 $G(v)$ に関しては、その逆数 v_i^{-1} の幾何平均 $G(\{v_i^{-1}\})$ との間に逆比整合性 $G(v)\cdot G(\{v_i^{-1}\})=1$　(29)が成立する。

5.2　為替交換

表１の(*6)ドル→円交換において、X〔ドル〕$=\{x_i\}$〔ドル〕、Y〔円〕$=\{y_i\}$〔円〕$(i=1,..,n)$ が個別の交換額データである。区間 i(あるいは標本 i)でのドル→円為替レートは $r_i=y_i/x_i$〔円／ドル〕(30)となる。ここで、日毎の交換額$\{x_i\}$, $\{y_i\}$が既知ならば、3.3 節の加重平均式(13),(14)を使えば、正確な総合交換率が求まる。しかし、外国為替においては、個別の交換額データは入手不可だが、日々の為替レートのデータは入手可能と言う状況は十分考えられる。以下に、$r_i(i=1,..,n)$ のみが既知の場合を想定する。このような場合、日毎の為替レート$\{r_i\}$の幾何平均を注目する期間の代表値として採用する（例えば[2])。幾何平均は、正確な重み情報を用いた加重平均(3.3 節)以外では、逆比整合性を満足する唯一の平均値指標でもある。この場合の逆比整合性は、為替交換手数料を無視すれば、ドルから円へ、又その逆方向、円からドルへ、と交換した時に損得無しで同額が保持できることを意味する。この性質は、AHP の一対比較において、$a_{ij}\times a_{ji}=1$ とも対応しており、為替交換での逆比整合性と呼ぶ。

幾何平均の逆比整合性：

$(\Pi r_i)(\Pi 1/r_i)=1$ あるいは　$(\Pi(y_i/x_i))^{1/n}\cdot(\Pi(x_i/y_i))^{1/n}=1$ (31)

例えば算術平均では、$(\Sigma r_i/n)(\Sigma(1/r_i)/n)=1$ は一般には等号は不成立(不等号≧は成立)である。、

5.3　ベンチマーク性能評価

計算機 j 上でテストプログラム i を走行した時の処理時間データを $T(i,j)$ とする。プログラム $i(i=1,..,n)$ に関して何らかの平均操作 f をとった指標 $T(\cdot,j)=f(T(1,j),T(2,j),\ldots,T(n,j))$ (32)、あるいは、適当な計算機 k 上の処理時間データで割った正規化データ $T(i,j)/T(i,k)$ に対して平均操作 f をとった指標

$F(\cdot,j,k)=f(T(1,j)/T(1,k), T(2,j)/T(2,k)\ldots, T(n,j)/T(n,k))$　(33)

を考える。文献〔1〕では、この平均操作 f として幾何平均を採用することを主張している。この理由として、幾何平均のみが乗法可分離性(multiplicative property)と呼ばれる以下の Property3 を満足する点を説明している([1]の p.221 より抜粋)。

〔Property3〕

$f(a_1b_1,\ldots,a_nb_n)=f(a_1,\ldots,a_n)\cdot f(b_1,\ldots,b_n)$　　(34)

(34)は、(35)、(36)とも表現できる。

$f(a_1/c_1,\ldots,a_n/c_n)=f(a_1,\ldots,a_n)\cdot f(1/c_1,\ldots,1/c_n)$　　(35)

$f(a_1/c_1,\ldots,a_n/c_n)\cdot f(c_1/a_1,\ldots,c_n/a_n)=1$　　(36)

以上より、平均操作 f として幾何平均を採用すると、(34)～(36)の乗法可分離性が(31)の逆比整合性に一致する。

5.4　為替交換とベンチマーク性能評価の同型性

通貨 j で為替交換 i での額面を $T(i,j)$、通貨 k で為替交換 i での額面を $T(i,k)$ とすれば、$T(i,j)/T(i,k)$ は為替交換 i での通貨 k に対する通貨 j (j/k)の交換比率を表す。………このように対比すると、為替交換 i での通貨 j がベンチマーク性能評価での計算機 j、通貨 j での額面 $T(i,j)$ が計算機 j 上でのプログラム i の計算時間に対応する。すなわち、為替交換とベンチマーク性能評価の両問題の間に同型性(isomorphism)が存在する。このような観点からも、両問題において比率の平均として、適当な条件下で幾何平均の適用が推奨される。

5.5　３通貨間、３計算機間の整合性

AHP での３項目(j,k,l)間の整合性「$a_{jk}\times a_{kl}\times a_{lj}=1$」あるいは「$a_{jk}\times a_{kl}=a_{jl}$」に対応して、３通貨、例えば、円、ドル、ユーロ(j,k,l)、間の整合性は、ベンチマーク性能評価の言葉で言えば、$F(\cdot,j,k)\times F(\cdot,k,l)\times F(\cdot,l,j)=1$(37)となるが、この場合も平均操作 f として幾何平均を採用すると、３通貨間あるいは３計算機間の整合性が成立する。さらに、n 通貨間、n 計算機間の整合性へと発展する([3]参照)。

付録１　男女別・学科別平均点のエクセル集計表の例

人数N:Number　b:boy g:girl s:suuri数理 m:manageマネジ						
N(b,s)=a	8	N(g,s)=b	2	N(s)=a+b	10	
N(b,m)=c	1	N(g,m)=d	9	N(m)=c+d	10	独立・所与
N(b)=a+c	9	N(g)=b+d	11	N=a+b+c+d	20	
平均A:Average						
A(b,s)=α	20	A(g,s)=β	90	A(s)=S(s)/N(s)	34	
A(b,m)=γ	10	A(g,m)=δ	80	A(m)=S(m)/N(m)	73	独立・所与
A(b)=S(b)/N(b)	18.88889	A(g)=S(g)/N(g)	81.81818	A=S/N	53.5	
総得点S:Sum						
S(b,s)=aα	160	S(g,s)=bβ	180	S(s)=aα+bβ	340	
S(b,m)=cγ	10	S(g,m)=dδ	720	S(m)=cγ+dδ	730	
S(b)=aα+cγ	170	S(g)=bβ+dδ	900	S=aα+bβ+cγ+dδ	1070	

図 A1-1　逆理の例

人数N:Number　b:boy g:girl s:suuri数理 m:manageマネジ						
N(b,s)=a	8	N(g,s)=b	2	N(s)=a+b	10	
N(b,m)=c	1	N(g,m)=d	9	N(m)=c+d	10	独立・所与
N(b)=a+c	9	N(g)=b+d	11	N=a+b+c+d	20	
平均A:Average						
A(b,s)=α	20	A(g,s)=β	90	A(s)=S(s)/N(s)	34	
A(b,m)=γ	10	A(g,m)=δ	36.6666	A(m)=S(m)/N(m)	34	独立・所与
A(b)=S(b)/N(b)	18.88889	A(g)=S(g)/N(g)	46.36358	A=S/N	34	
総得点S:Sum						
S(b,s)=aα	160	S(g,s)=bβ	180	S(s)=aα+bβ	340	
S(b,m)=cγ	10	S(g,m)=dδ	329.9994	S(m)=cγ+dδ	340	
S(b)=aα+cγ	170	S(g)=bβ+dδ	509.9994	S=aα+bβ+cγ+dδ	680	

図 A1-2　逆理と順理の境界例

付録２　演算に注目した平均化の概念

平均化という演算・操作(operation)を、データ $a_i(i=1,..,n)$の対(a_i,a_j)に対する２項演算という観点から、見直し、整理しよう。２項演算⊕が定義される代数系 $A(\oplus)$において n 個のデータ$\{a_1,\ldots,a_n\}$について方程式(A2･1)を満足する x を、演算⊕の下での平均と定義する。

$a_1\oplus a_2\oplus\ldots\oplus a_n=x\oplus x\oplus\ldots\oplus x$　　　(A2･1)

ここで、(A2･1)の左辺では、結合律... $(a_1\oplus a_2)\oplus a_3=a_1\oplus(a_2\oplus a_3)=a_1\oplus a_2\oplus a_3$　(A2･2)が成立することが望まれる(不成立だと、(A2･1)の数式表現ができないので)。

(ⅰ)　⊕=+(通常の加算)とすれば、$x=\Sigma a_i/n$,と算術平均を得る(結合率は成立)。

(ⅱ)　⊕=x(乗算)とすれば、$x=(\Pi a_i)^{1/n}$と幾何平均を得る(結合率成立)。

(ⅲ)　$a_1\oplus a_2=a_1^p+a_2^p$ (A2･3)では、結合率(A2･2)は不成立。$p=1, p\to 0, p\to+\infty, p\to-\infty$以外では結合率が成立せず、２項演算を用いた方程式による平均定義では、調和平均さらには一般化平均は導出できない(n>3)。

(ⅳ)　$a_1\oplus a_2=a_1+a_2-a_1\cdot a_2$　(A2･4)では結合率成立し、n=2では２つの平均値を持つ。

$x=1\pm(1+a_1\cdot a_2-a_1-a_2)^{1/2}$　(A2･5)

提案する２項演算子⊕を用いた定義法は、平均化関数$f(a_1,a_2,\ldots,a_n)$の方程式：　$f(a_1,a_2,\ldots,a_n)=f(x,x,..,x)$　　(A2･6)を用いた定義法「Chisini 平均」に含まれるが、２項演算に注目した点が新しく、結合率の成立が要求され、算術平均、幾何平均は定義できるが、調和平均や一般化平均は必ずしも本アプローチでは定義できないなど、適用範囲が限定されることがわかる。

２種の演算子(⊕,⊗)を用いた(A2･7)で定義する平均概念を次に提案する。

$a_1\oplus a_2\oplus\ldots\oplus a_n=x\otimes x\otimes\ldots\otimes x$　　　(A2･7)

(A2･1)の左辺での演算子⊕、右辺での演算子⊗が異なる。

⊕を＋、⊗を×とすれば、$x=(\Sigma a_i)^{1/n}$　　　(A2･8)

⊕を×、⊗を＋とすれば、$x=\Pi a_i/n$　　　(A2･9)

と新しい平均値概念を得る。

付録３　比率の平均の表現式(発展形)

(ⅰ)　(13)において本来は$T=\Sigma t_i$　(2)であるが、この定義式を使用しなくても、例えば$T=S(t_1,t_2,..,t_n)$でも、$w_i(t)=t_i/T$と定義すれば加重算術平均式(13)は成立する。同様に、(14)において D の定義式(3)を使用しなくても、$w_i(d=d_i/D$と定義すれば加重調和平均式(14)は成立する。

(ⅱ)　$D=\Sigma d_i$　(1)　の替わりに$D=\Pi d_i$　(A3.1)とすると(他は不変)、$V=D/T=\Pi d_i/T=\Pi w_i(t)v_i$　(A3.2)、$V^{-1}=T/D=\Sigma t_i/D=\Sigma(t_i/d_i)d_i/D=\Sigma w_i(d)v_i^{-1}$　(A3.3)を得る。又、$T=\Sigma t_i$　(2)　の替わりに$T=\Pi t_i$　(A3.4)とすると(他は不変)、$V=D/T=\Sigma w_i(t)v_i$　(A3.5)、$V^{-1}=T/D=\Pi w_i(d)v_i^{-1}$　(A3.6)を得る。(A3.2) と(A3.6) は、加重算術、調和、幾何のいずれにも該当しない。

(ⅲ)　「$D=\Pi d_i$, $T=\Pi t_i$　(A3.7) (他は不変)」とすれば、$V=D/T=\Pi v_i$　(A3.8)と単純化される。

参考文献

[1] PHILIP J. FLEMING and JOHN J. WALLACE：
How not to lie with statistics: the correct way to summarize benchmark results, Communications of the ACM , 29(3):218-221 ·March 1986.

[2]伊藤雄一郎、稲場広記、尾崎直子、関根敏隆(調査統計局)　：実質実効為替レートについて、日銀レビュー2011-J-1 (2011.2)

[3]　篠原正明、篠原健　新しい整合性概念「m点整合性」の提案　日本大学生産工学部学術講演会　2-70 (2012)