研究ノート

算術平均と幾何平均

増　田　賢　司

Ⅰ　はじめに

一様確率変数２つから作る算術平均および幾何平均の分布について，⑴同時分布，⑵差の分布，⑶相関係数とその応用を検討する．また，よく知られている性質だが，証明を見かけないので，⑷算術平均が幾何平均より大きいことの証明をメモとして与えておく．一様確率変数が３つそして４つの場合には，現在，結果は得られていない．今後の課題となる．

Ⅱ　同時分布

Ⅱ－１　一様分布を利用する方法

分布の範囲が［０，１］である互いに独立な２つの一様確率変数を X と Y とする．X の値域を x 軸にとり，Y の値域を y 軸上にとる．値域を細分化し，n^2 個の一辺の長さが $1/n$ の正方形をつくる．すなわち，

$$R_{kh}=\{(x,y)\,;k/n<x\leq(k+1)/n,h/n<y\leq(h+1)/n\},$$
$$k=0,\ldots,n-1,h=0,\ldots,n-1$$

この正方形を，算術平均と幾何平均なる領域に写像する．

$$P_{kh}=\left\{\left(\frac{x+y}{2},\sqrt{xy}\right);(x,y)\in R_{kh}\right\},\ \ k=0,\ldots,n-1,\ \ h=0,\ldots,n-1$$

$\left(\frac{x+y}{2},\sqrt{xy}\right)$ は x と y に関して連続である，したがって，P_{kh} は k と h について順番に隣り合い隙間なくつながっている．

算術平均と幾何平均を次のように表す．

$$\mathrm{A}=\frac{X+Y}{2},\ G=\sqrt{XY}$$

確率変数（X, Y）が R_{kh} の中に落ちる確率は $1/n^2$，（A, G）が P_{kh} の中に落ちる確率は $2/\mathrm{n}^2$，$k\neq h$ そして $1/n^2$，$k=h$ になる．

分割して作った各小さな正方形の中心（x_0, y_0）を代表値として選び，その各々の点に確率 $1/n^2$ を与えて，（X, Y）を近似する２次元の離散一様分布を作る．（x_0, y_0）より算術平均と幾何平均を求める．このようにして得た離散一様分布による（A, G）を近似した確率変数を（A', G'）とすると $P_{kh}=P_{hk}$ だから

$$\mathrm{P}\{(A',G')\in P_{kh}\}=\begin{cases}2/n^2,\ k\neq h\\1/n^2,\ k=h\end{cases}$$

いま，

$$\left(\frac{x_0+y_0}{2},\sqrt{x_0y_0}\right)=(a_{kh},g_{kh}),\ k>h$$

とし，$0<A'<1$，$0<G'<1$ であることを考え，x 座標と y 座標それぞれの［0，1］部分を m 等分することによって m^2 個の正方形をつくる．各正方形の中に落ちる（a_{kh}, g_{kh}），$k=0$，…，$n-1$，$h=0$，…，$n-1$ の個数を，分割された一様分布の個数，n^2 で割れば，（A', G'）の同時確率分布が得られる．適当な分割を行うことによって，（A, G）の同時確率密度関数が高い精度で得られる．既知である周辺分布を利用すれば，おおよその精度も推測できよう．

表 1．1　算術平均と幾何平均の同時分布［0，1］の一様確率変数から作成

	a-mean										
g-mean	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95	計
0.95										6	6
0.85									18	4	22
0.75								26	8		34
0.65							30	16			46
0.55						28	26				54
0.45					28	26	2				56
0.35				30	20	10					60
0.25			26	14	10	4					54
0.15		18	16	6	8	2					50
0.05	8	10									18
計	8	28	42	50	66	70	58	42	26	10	400

表 1．2　算術平均と幾何平均の同時分布［1，2］の一様確率変数から作成

	a-mean										
g-mean	1.05	1.15	1.25	1.35	1.45	1.55	1.65	1.75	1.85	1.95	計
1.95										6	6
1.85									18	4	22
1.75								30	8		38
1.65							38	12			50
1.55						44	20				64
1.45					46	30					76
1.35				40	24						64
1.25			30	14							44
1.15		18	8								26
1.05	6	4									10
計	6	22	38	54	70	74	58	42	26	10	400

Ⅱ－2　条件付き分布を使う方法（［2］を参照）

$$G(x/y)=P\left(\left\{\sqrt{XY}\leq x;\frac{X+Y}{2}=y\right\}\right)$$

算術平均の累積分布を

$$H(y)=P\left(\frac{X+Y}{2}\leq y\right)$$

その確率密度関数を

$$h(y)=\begin{cases}2y,\ 0\leq y\leq 1/2\\ -2y+1,\ 1/2<y\leq 1\end{cases}$$

で表す.

$$G(x/y)=P\left(\sqrt{y^2-(X-y)^2}\leq x\right)$$

変形して

$$=P\left(X\leq y-\sqrt{y^2-x^2}\right)$$

X ≤ y で一様に分布しているから

$$=\left(y-\sqrt{y^2-x^2}\right)/y$$

これより，確率密度関数は

$$G'(x/y)=g(x/y)=\begin{cases}\dfrac{1}{y}\dfrac{x}{\sqrt{y^2-x^2}},\quad x\neq y\\ 1,\ x=y\end{cases}$$

したがって，同時確率密度関数は

$$f(x,y)=g(x/y)\,h(y)$$

$$=\begin{cases}\dfrac{2x}{\sqrt{y^2-x^2}} & ,0\leq y\leq 1/2,\ 0\leq x<y\\ \dfrac{2x}{\sqrt{y^2-x^2}}\dfrac{1-y}{y} & ,1/2<y\leq 1,\ 0\leq x<y\end{cases}$$

Ⅲ　算術平均と幾何平均の差の分布

分布の範囲が［0，1］である互いに独立な2つの一様確率変数を X，Y とする．X と Y の算術平均を A，幾何平均を G とする．

$$\frac{1}{2}\left(\sqrt{x}-\sqrt{y}\right)^2=\frac{x+y}{2}-\sqrt{xy}$$

であるから

$$A-G=\frac{1}{2}\left(\sqrt{X}-\sqrt{Y}\right)^2$$

$\sqrt{X}$ の分布は $P\left(\left\{\sqrt{X}\leq x\right\}\right)=P\left(\left\{X\leq x^2\right\}\right)=x^2,\ 0\leq x\leq 1$ であるから確率密度関数は

$$h(x)=2x,\quad 0\leq x\leq 1$$

同様にして $-\sqrt{Y}$ の確率密度関数は

$$g(y)=-2y,\quad 0\leq y\leq 1$$

次に $Z=\sqrt{X}-\sqrt{Y}$ の確率密度関数を畳み込みの方法を使ってもとめる．
$1\ge z\ge 0$ のとき

$$f(z)=\int_z^1 h(x)\,g(x-z)\,dx=\int_z^1 4x(x-z)\,dx$$
$$=\frac{2}{3}(2-3z+z^3)$$

$-1\le z\le 0$ のとき

$$f(z)=\frac{2}{3}(2+3z-z^3)$$

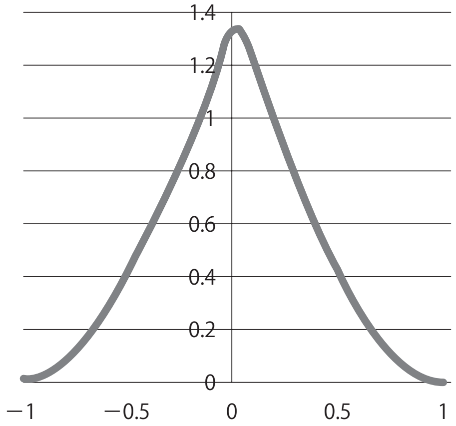

図 1　$Z=\sqrt{X}-\sqrt{Y}$ の確率密度関数

$A-G=\frac{1}{2}\left(\sqrt{X}-\sqrt{Y}\right)^2=Z^2$ の分布は

$$P\left(\left\{\frac{1}{2}Z^2\le x\right\}\right)=P\left(\left\{-\sqrt{2x}\le Z\le\sqrt{2x}\right\}\right)$$

対称性から

$$=2P\left(\left\{0\le Z\le\sqrt{2x}\right\}\right)$$
$$=2F\left(\sqrt{2x}\right)$$

したがって，確率密度関数は

$$f(x)=2F'\left(\sqrt{2x}\right)\frac{1}{\sqrt{2x}}$$
$$=\frac{4}{3}\left(2-3\sqrt{2x}+\left(\sqrt{2x}\right)^3\right)\frac{1}{\sqrt{2x}},\ 0\le x\le\frac{1}{2}$$

という形になる．（図 2 参照）

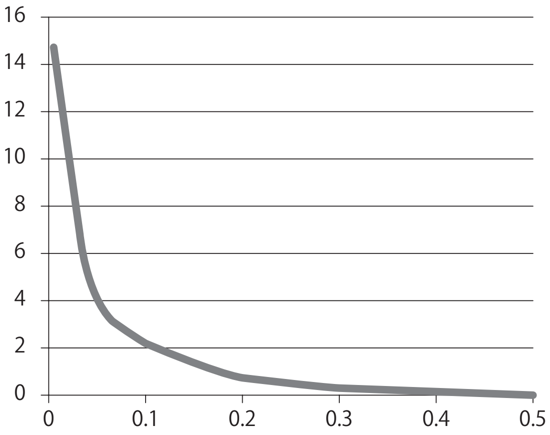

図２　幾何平均と算術平均の差の確率密度関数

Ⅳ　差の分布の拡張

⑴［０，１］の範囲の一様分布に従い，互いに独立な２つの変数 X と Y の算術平均を A，幾何平均 G とし，［０，K］の範囲の一様分布に従い，互いに独立な２つの変数の算術平均を A_K，幾何平均を G_K とすると

$$A_K - G_K = K(A - G)$$

と表現でき，確率密度関数が求まる．

⑵確率変数 X と Y が互いに独立で，［０，１］の範囲の一様分布に従い，確率変数 X_1 と Y_1 が互いに独立で，$[K-\Delta,\ K+\Delta]$ の範囲の一様分布に従うとする．確率変数

$$\frac{X_1}{K-\Delta},\ \frac{Y_1}{K-\Delta}$$

は互いに独立で，それぞれ範囲

$$\left[1, 1+\frac{2\Delta}{K-\Delta}\right]$$

で一様分布する．したがって，

$$\frac{X_1}{K-\Delta}-1,\ \frac{Y_1}{K-\Delta}-1$$

は範囲

$$\left[0, \frac{2\Delta}{K-\Delta}\right]$$

で，それぞれ一様分布に従う．簡単さのため

$$X_2 = \frac{X_1}{K-\Delta}-1,\ Y_2 = \frac{Y_1}{K-\Delta}-1$$

とおいて，確率変数 X_1 と Y_1 の算術平均と幾何平均をもとめる．

$$A_1 = \frac{X_1 + Y_1}{2} = (K - \Delta)\left(1 + \frac{X_2 + Y_2}{2}\right)$$

また，

$$G_1 = \sqrt{X_1 Y_1} = (K - \Delta)\sqrt{(1 + X_2)(1 + Y_2)}$$

テイラー展開を適用して

$$\fallingdotseq (K - \Delta)\left(1 + \frac{X_2 + Y_2 + X_2 Y_2}{2}\right)$$

上式より

$$\begin{aligned} A_1 - G_1 &\fallingdotseq (K - \Delta)\frac{X_2 Y_2}{2} \\ &= (K - \Delta)\frac{1}{2}\left(\frac{2\Delta}{K - \Delta}\right)^2 XY \\ &= \frac{2\Delta^2}{K - \Delta} XY \end{aligned}$$

ここでは確率密度関数は導出しないが，算術平均と幾何平均の差は［ 0，1 ］の範囲の値をとる一様分布を基準として，ほぼ，K に反比例することが分かる．

V　相関係数

算術平均 A と幾何平均 G の同時分布より，両確率変数の相関が高いことが推測できる．確率変数 X と Y が互いに独立で，［ 0，1 ］の範囲の一様分布に従う場合の相関係数を計算する．

$$\begin{aligned} V(A - G) &= E(A - G)^2 - \left\{E(A - G)\right\}^2 \\ &= V(A) + V(G) - 2\left\{E(AG) - E(A)E(G)\right\} \end{aligned}$$

A と G の分布は既知であり

$$E(A) = \frac{1}{2},\ V(A) = \frac{1}{24},\ E(G) = \frac{4}{9},\ V(G) = \frac{17}{324}$$

上の結果は簡単に得られる（［ 1 ］参照）．また，Ⅱで求めた $A - G$ の分布を使って

$$E(A - G) = \int_0^{1/2} \frac{4}{3}\left\{2 - 3\sqrt{2x} + \left(\sqrt{2x}\right)^3\right\}\frac{\sqrt{x}}{\sqrt{2}}dx = \frac{1}{18}$$

$$E(A - G)^2 = \int_0^{1/2} \frac{4}{3}\left\{2 - 3\sqrt{2x} + \left(\sqrt{2x}\right)^3\right\}\frac{x\sqrt{x}}{\sqrt{2}}dx = \frac{1}{120}$$

したがって，

$$V(A - G) = \frac{1}{12}\frac{17}{270}$$

A と G の相関係数 ρ_{AG} は

$$\rho_{AG} = \frac{E(AG) - E(A)E(G)}{\sqrt{V(A)}\sqrt{V(G)}} = 0.950542$$

と 1 に非常に近くなり，代用特性の可能性を探る材料になろう．

Ⅵ　A と G に関する不等式

不等式の結果はよく知られているし，証明方法の示唆も見られるが，証明自体は目にしたことがない．証明には少し工夫も必要なのでメモとして記述しておく価値はある．

$\bar{x} \geq (x_1 x_2 \ldots x_n)^{1/n}$

［証明］

正の実数　x_1，x_2 に対して

$$\frac{1}{2}(x_1 + x_2) \geq (x_1 x_2)^{\frac{1}{2}}$$

が成り立つ．

正の実数　x_1，x_2，…，x_{2^N} に対して

$$\frac{1}{2^N}(x_1 + x_2 + \cdots + x_{2^N}) \geq (x_1 x_2 \ldots x_{2^N})^{1/2^N}$$

が成り立つとする．

正の実数　x_1，x_2，…，$x_{2^{N+1}}$ に対して

$$\begin{aligned}
&\frac{1}{2}\left(\frac{1}{2^N}(x_1 + x_2 + \cdots + x_{2^N}) + \frac{1}{2^N}(x_{2^N+1} + x_{2^N+2} + \cdots + x_{2^{N+1}})\right) \\
&\geq \frac{1}{2}\left((x_1 x_2 \ldots x_{2^N})^{1/2^N} + (x_{2^N+1} x_{2^N+2} \ldots x_{2^{N+1}})^{1/2^N}\right)^{1/2} \\
&\geq (x_1 x_2 \ldots x_{2^N} x_{2^N+1} x_{2^N+2} \ldots x_{2^{N+1}})^{1/2^{N+1}}
\end{aligned}$$

したがって，不等式は 2^{N+1} 個の正数に対して成り立つことが示された．

2^N 個の正の実数 x_1，x_2，…，x_n，$\bar{x}$，$\bar{x}$，…，$\bar{x}$ にたいしても不等式は成り立つから

$$\frac{1}{2^N}\left(x_1 + x_2 + \cdots + x_n + \left(2^N - n\right)\bar{x}\right) \geq \left(x_1 x_2 \ldots x_n \bar{x}^{2^N - n}\right)^{1/2^N}$$

整理して

$$\begin{aligned}
\bar{x} &\geq \left(x_1 x_2 \ldots x_n (x_1 x_2 \ldots x_n)^{\left(2^N - n\right)/n}\right)^{1/2^N} \\
&= (x_1 x_2 \ldots x_n)^{1/n}
\end{aligned}$$

任意の個数の正の数の算術平均と幾何平均に不等式が成り立つことが証明された．

さらに，重み付平均について不等式を適用すると

n 個の正の実数 $x_1 x_1 \cdots x_1 x_2 \cdots x_2 \cdots\cdots x_k \cdots x_k$ に対して

$$\frac{n_1}{n}x_1 + \frac{n_2}{n}x_2 + \cdots + \frac{n_k}{n}x_k \geq x_1^{n_1/n} x_2^{n_2/n} \cdots x_k^{n_k/n}$$

ただし，$\sum_{i=1}^{k} n_i = n$

が得られる．

Ⅶ　おわりに

2つの一様確率変数について，算術平均と幾何平均の関係を考えた．同時分布，差の分布，相関係数等をⅡ～Ⅵで与えている．3つ以上の変数の場合，解析的に解くことができていないが，両者の差の大きさが小さくなることはシミュレーションで確認できる．そこで2つの変数の場合についての結果が，変数の数が増加していく場合の傾向をみるときの参考になろう．

Ⅷ　参考文献

［1］増田賢司（2009）："一様分布の幾何平均"，経済集志　第 78 巻 1 号.
［2］増田賢司（2011）："一様分布について　―二乗和の確率密度関数―"，経済集志　第 81 巻 2 号.