

所得分布とジニ係数に関する一考察*

河野敬雄

(受付 2005年9月27日)

要 旨

国民所得分布の理論近似として、従来から高額所得者の範囲ではパレート分布が、その他の部分では対数正規分布がよい近似を与えているといわれている。本来、理論分布は一定の条件を満たす範囲で所得分布全体を近似できるものでなくてはならない。本稿ではそのような分布として逆数ガンマ分布を提案する。また、非負値で正の平均が存在するすべての分布に対して統一的にローレンツ曲線の定義を与える。平均が存在しない分布（たとえば、ある範囲のパレート分布）に対してはジニ係数が求められない。この困難を回避するためにジニ係数の修正を提案する。

キーワード：ローレンツ曲線，ジニ係数，所得分布の近似，自己分解可能分布，逆数ガンマ分布

§1. 問題設定

所得分布が如何なる理論分布で近似できるかについては古来いろいろ議論されている。依田（1997，152頁）によると現実の所得分布を近似する分布として、パレート，ジブラ（対数正規），ガンマ，対数ロジスティック，一般ガンマ，ベータ，シン・マダラの各種分布が利用されている，とのことである。この本でもそうであるがこれらの分布のうち，高額所得者の範囲ではパレート分布が，それ以外の範囲では対数正規分布が現実の分布をよく近似しているということが多くの論者に支持されているようである。ここで，パレート分布とは密度関数 $f_p(x)$ がパラメータ $\alpha > 0$ ， $x_0 > 0$ に対して

$$f_p(x; \alpha) = c_\alpha x^{-\alpha-1}, \quad x_0 \leq x < +\infty, \quad c_\alpha = \alpha x_0^\alpha$$

で表される分布であり，対数正規分布とは密度関数 $f_\ell(x)$ がパラメータ (μ, σ^2) に対して

$$f_\ell(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma x}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right), \quad 0 < x < +\infty$$

で表される分布である^{注1)}。

* 本稿の内容は第39回数理社会学会（2005年3月4日新潟国際情報大学）における萌芽セッション発表「所得分布は逆数ガンマ分布に従うこと，並びにジニ係数の修正について」に基づいて大幅に修正，加筆したものである。

注1) 対数正規分布の密度関数については，著者によって，若干の混乱がみられる。たとえば，依田（[11]，157頁）に書いてある密度関数には $1/x$ のファクターが抜けている。ミスプリントであると考えるのが常識的であるが，Gumbel（翻訳18頁）には「 $1/x$ のない密度関数を考えている著者もあるが，積分すると1より大きくなるから（当然である）分布とはならない，この形式は以

ここで、対数正規分布について分布の裾 (x が大きい範囲) のオーダーを調べると、関数の形から容易に、

$$\forall k > 0, \lim_{x \rightarrow +\infty} x^k f_l(x) = 0$$

となって、パレート分布の裾のオーダーとは決定的に異なることがわかる。

一方、パレート分布は高額所得者の範囲しか近似し得ないことは自明である。つまり、現実の所得分布全体をパレート分布ないし対数正規分布ひとつで近似することは原理的にできないのである。

§2. 所得の理論分布に要請されること

現実の所得分布をよりよく近似する理論分布は如何なる分布であるべきかについて、モデルからパレート分布、対数正規分布を導出することが種々試みられている (例えば、青木 1979, 121–130頁参照)。しかし、上述したようにパレート分布も対数正規分布も単独では所得分布全体を近似することは出来ないことから、モデルそのものの説得力は必ずしも十分ではないように思われる。従って、本稿においては、高額所得者の分布はパレート分布と同じ減少のオーダーであること (弱パレート性)、低額所得者の分布においては対数正規分布に近い (数学的には表現しない) 分布のなかで、かつ理論的に意味のある極限分布の中で所得分布を近似できそうな分布を探すことにする。以上をまとめて、所得の理論分布の密度関数 $f(x)$ に次のことを要請する。

所得の理論分布に対する要請

- (i) ひとつの初等関数で表されること。
- (ii) (弱パレート性) 高額所得者の範囲での分布のオーダーがべき法則に従うこと (パレート分布と同じオーダー) (Mandelbrot, B. (1960) にある“The weak Pareto law”を要請することと同値)。
- (iii) 自己分解可能分布 (= L 分布) であること。

(iii) については数学 (確率論) の一部専門家には周知の概念であるが、一般には (特に応用上) なじみ深い分布とは言い難いので若干の解説が必要である。無限分解可能分布、自己

下では用いない], という趣旨のことがわざわざ書かれてあるところを見るとミスプリントとは必ずしも断定できないのかもしれない。正規分布の密度関数の x を単に $\log x$ に置き換えればよい、と誤解している著者もあるということらしい。分布関数の定義からすぐ導けることではあるが、基本的には合成関数の微分の公式で $\log x$ を x で微分することによって $1/x$ が出てくる。ついでにいうと、標準偏差 σ^* が根号記号の中に含まれているがこれもミスプリントで $\sqrt{2\pi}\sigma^*$ が正しい。

分解可能分布, 安定分布に関する詳しい定理については佐藤 (1990) および Sato (1999) を参照されたい。なお, (iii) をパラメータ $1 < \alpha < 2$ の安定分布 (Mandelbrot はパレート・レヴィ分布と命名している) である, としたのが Mandelbrot, B. (1960) である。従って, 彼の論文と対比しながら解説すると理解しやすいであろう。

安定分布はよく知られているように自己分解可能分布であり (佐藤1990 111頁, 命題 4.3.9), かつ(ii)の弱パレート原理を満たしている (Mandelbrot)。しかし, 残念ながら指数 $1/2$ ($\alpha = 1/2$) の片側安定分布, コーシー分布 ($\alpha = 1$), 正規分布 ($\alpha = 2$) 以外には密度関数の具体的表現は知られていない^{注2)}。もちろん, 安定分布の特性関数 (分布のフーリエ変換) はよく知られているから, いわゆる反転公式によって理論的には特性関数から密度関数を求めることは出来る。しかし, 理論関数を数値計算で各点毎に求めなければならない, というのは実用上いかにも不都合であると思われる^{注3)}。Mandelbrot の提案する安定分布によって近

注2) 初等関数では表されないと思われるが証明は知られていないようである。Mandelbrot 1960, 87 頁, 11-12行目にも指摘がある。

注3) 青木 (1979, 68頁) では特性関数から反転公式を用いて密度関数を求めるよう要請している。もっとも, 同書のグラフはどうか Mandelbrot の論文の Figure 1 の3つのグラフをひとつの座標に写しただけのように見える。彼自身が実際に反転公式を用いて数値計算して描いたグラフではないであろう。ついでに指摘しておく, 彼は安定分布の特性関数のパラメータの範囲を $0 < \alpha < 2$ としているが, 正しくは $0 < \alpha < 1$ または $1 < \alpha < 2$ である。 $(\alpha = 1$ の場合は別の表現をとる) Mandelbrot は平均の存在を仮定しているから $1 < \alpha < 2$ としてある。さらにいうと, 実は Mandelbrot の原論文に (従って, 当然青木のそれにも) 書いてある安定分布の特性関数の表記にはミスプリントがあると思われる。彼は特性関数を

$$\varphi(\zeta) = \exp[-iM\zeta - |\zeta|^\alpha (u^*)^\alpha (1 - \frac{i\beta\zeta}{|\zeta|} \tan \frac{\alpha\pi}{2}) \cos \frac{\alpha\pi}{2}]$$

としてあるが, 正しくは, $\cos \alpha\pi/2$ のところを, $|\cos \alpha\pi/2|$ とすべきであろう。もちろん, 数学的には前の符号を+にしても同じことであるが, $0 < \alpha < 1$ の場合との整合性を考えると \cos の項に絶対値をつける方がベターであろう。なお, 多くのテキストでは \cos は u^* の項に繰り込んであって表面上現れない形にしてある。青木 (1979, 68頁) に参照せよとの指示のある Gnedenko-Kolmogorov ([9], ch. 7) の本では表現上は \cos の項はないが, 計算途中には現れている。彼等の場合, $0 < \alpha < 1$ と $1 < \alpha < 2$ の場合に分けて計算してあり (収束条件が異なるために分けて計算せざるを得ない), 符号をよく調べてみると, 上記のように訂正する必要があることがわかる。なお, 何故ミスプリントに気がついたか, というと, 実数値確率変数 X (の分布) の特性関数とは, 任意の実数 ζ に対して, $\varphi_X(\zeta) = E[\exp(i\zeta X)]$ で定義される複素数値連続関数で, 定義式から直ちに $|\varphi_X(\zeta)| \leq \varphi_X(0) = 1$ となるはずであるが, Mandelbrot の表記では $1 < \alpha < 2$ に対して, $\cos \alpha\pi/2$ は負の値をとり, $u^* \geq 0$ を仮定しているから, 特性関数の絶対値が 1 より大きくなる可能性に気がついたからである。原則として, 数学公式ないし定理を利用しないし引用する場合は, 人間は間違いを犯すものであるという自覚の下に, 自分で正しいことを確認した上で利用しないし引用するのが数学者の慣習である。もちろん, 深刻な間違いの場合はそれを指摘すること自体がオリジナルな論文になり得る。例えば, L 分布の単峰性は ([9]) のロシア語のオリジナル版で「証明」されていたが, その証明は途中で利用した命題に反例が挙げられたことによって, あやまりであることが証明され, 「定理」は予想にすぎなくなった。その後25年間の紆余曲折があった末に最終的に正しい証明を与えたのは山里真氏 (1978) である (詳しくは山里 (1978) を参照されたい)。そのため, 大学院生のゼミのテキストとしては完全無欠な教科書よりはアイデアは豊富であるが, 少々ミスプリントや証明にアナがある方が教育上好ましい \wedge

似することのもうひとつの欠点は、平均が有限に存在することを仮定するとパラメータの範囲が $1 < \alpha \leq 2$ に限定されることであり、その場合、分布の台（密度関数が正となる範囲）が実数全体となることである。負の無限大の方向には確かに任意のべきより早いオーダーでゼロに収束することが知られている（佐藤 1990, 133頁）が、分布のモードを現実の所得分布のモードに一致させた場合に、負の半直線上の確率が無視できるほど小さいのか、という点について Mandelbrot (1960) では何も検討されていないようである。

正規分布や対数正規分布が自己分解可能分布であることはよく知られているが、依田 (1997, 152頁) に挙げられている前述のほとんどの分布は自己分解可能分布である。しかし、高額所得者あたりをパレート分布で、それ以外を対数正規分布で近似して途中でそれらをつなげても無限分解可能であるとは限らない（多分証明できない）。ひとつの初等関数で表すことも出来ないであろう（このような命題を「証明」することは難しい）。

所得分布が自己分解可能分布であることを要請する理由は、それが、小さな独立確率変数を多数足し合わせて得られる極限分布として得られる無限分解可能分布に含まれ、さらにある種の自己相似的性質を満たす分布であるからである。この事実は、個人のミクロな経済活動の結果として全体の所得分布が得られるはずであるという方法的個人主義を前提にした場合都合がよい。個人のミクロな経済活動の数学モデルから全体の所得分布を数学的に導く、という研究はいろいろなされているが完全な成功を見ているとは言い難い。本稿においても、上記の条件を満たす分布をモデルから特定する、という考察には立ち入らない。しかしながら逆に考えれば、現実の所得分布が自己分解可能分布で近似できるという事実は、独立した個人のミクロな経済活動の集積としてマクロな経済活動が説明できる可能性がある、ということを示唆しているとも考えられるであろう。

ここで、上記の考察で述べた分布の分類について簡単に解説しておく。詳細は佐藤 (1990), Sato (1999) その他の確率論の教科書を参照されたい。

確率分布を特徴付ける方法は種々あるが、もっとも一般的方法は特性関数（分布のフーリエ変換）を用いる方法である。しかし、特性関数による表現は理論的簡潔さがある反面、実態がわかりにくい。本稿では確率変数を用いた表現によって定義を与える。同値な表現は一方を定義に採用すれば他方は定理となる。1次元分布のみを考えるから、本稿で言う確率変数とはすべて実数値をとる確率変数である。よく知られているように実数値確率変数 X の分布関数 F_X とは、任意の実数 x に対して X が x 以下の値を取る確率 $F_X(x) = P(X \leq x)$ のことである。

（院生が自力で間違いを発見、訂正して、厳密な証明を与える、という研究者としての基礎的訓練になるから）、と考えている指導教授は多いのである。

定義 1. 任意の実数 x に対して、 $F_X(x) = F_Y(x)$ が成り立つとき、 $X \stackrel{L}{=} Y$ と記し、確率変数 X と Y は同じ分布をもつ（同分布である）という。

定義 2. 確率変数 X の分布が無限分解可能であるとは、任意の自然数 n に対して、独立な確率変数列 $\{X_{n,j}; j=1, 2, \dots, n\}$ が存在して、

1. $X \stackrel{L}{=} \sum_{j=1}^n X_{n,j}$,
2. $\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \sup_{j \leq n} P(|X_{n,j}| > \varepsilon) = 0$

が成り立つことである。通常の設定では分布の convolution（独立確率変数の和の分布）の概念を用いて次のように定義する。つまり、同値な定義として

定義 2*. 確率分布 μ が無限分解可能分布であるとは、任意の自然数 n に対して確率分布 μ_n が存在して、 $\mu = \mu_n^{*n}$ (n 個の分布の convolution) が成り立つことである。

定義 3. (増田 定義2.2.) (佐藤 1990, 補題4.3.3 では特性関数の形で表現してある)。確率変数 X の分布が自己分解可能であるとは、任意の定数 $0 < c < 1$ に対して X とは独立な確率変数 Y_c が存在して

$$X \stackrel{L}{=} cX + Y_c$$

が成り立つことである。なお、 Y_c の分布は結果として無限分解可能である（佐藤 1990, 114 頁）

定義 4. (佐藤 1990, 94頁. 命題4.1.15) 確率変数 X の分布が安定分布であるとは、 X と同じ分布を持つ独立なふたつの確率変数を X_1, X_2 とするとき、任意の $a_1 > 0, a_2 > 0$ に対して、 $b > 0$ と実数 c が存在して

$$a_1 X_1 + a_2 X_2 \stackrel{L}{=} bX + c$$

が成り立つことである。

ここで、単位分布でない安定分布に対しては

$$b = (a_1^\alpha + a_2^\alpha)^{1/\alpha}$$

と表される。ここで、 α は $0 < \alpha \leq 2$ に限定され、安定分布の指数と呼ばれている。ただし、安定分布は指数だけでは決まらない。 $\alpha = 2$ の場合は正規分布であり、平均が有限に存在するのは $1 < \alpha \leq 2$ の場合に限る。分散および任意の高次モーメントが存在するのは $\alpha = 2$ つまり正規分布の場合に限る。

よく知られているように、以上の分布に関しては次のような包含関係がある。

定理 1. 正規分布 \Rightarrow 安定分布 \Rightarrow 自己分解可能分布 \Rightarrow 無限分解可能分布。

この定理によって、Mandelbrot の主張する安定分布が適当でないと考えた場合、次に検討すべき分布の型は自己分解可能分布である、ということが自然に導かれる。

なお、以上の分布族はすべて何らかの意味で独立な確率変数族の極限分布として得られる。正規分布は有限の分散を持つ分布の極限として得られるから、応用上極めて重要ではあるが、(それらの定理は中心極限定理と総称される) 極限分布が必ず正規分布になるわけではない。安定分布は分散有限を仮定しない独立同分布確率変数列の部分和から極限分布として得られ、自己分解可能分布 (L 分布) は同分布を仮定しない独立確率変数列の部分和の極限分布として得られる (佐藤 1990, 113頁注意4.3.13)。このあたりのことは現在詳細に知られているが、条件および結果の表現は極めてデリケートであるから利用する場合には細心の注意が必要である。

§3. 逆数ガンマ分布

対数正規分布はもともになる分布が正規分布であること、正規分布は極限分布として広範囲の分野で利用されていることから、正の値を取る確率変数の分布に対してはその対数をとったものが正規分布に従うに違いない、と考えるのは極めて自然な発想であろう。和を積に、算術平均を幾何平均にそれぞれ置き換えて極限分布を考えると中心極限定理が成り立つ状況ではつねに対数正規分布が現れるのは当然である。しかしながら、必ずしも対数正規分布がよい近似を与えない、ということは応用の分野では古くから経験的に知られているようである。代替りの候補は一般逆正規分布 (generalized inverse Gaussian distribution, GIG) である。その密度関数は

$$f(x) = cx^{\alpha-1} \exp(-\beta/x - \gamma x), \quad 0 < x < +\infty, \quad \beta, \gamma \geq 0, \quad \beta + \gamma > 0.$$

で与えられる自己分解可能分布である (増田 2002, 177頁)。ただし、 α は関数 f が正の半直線上で可積分となる範囲に定め、 $c > 0$ は密度関数となる条件、つまり積分が 1 となるように定められる。一般には、第 3 種変形ベッセル関数という特殊関数を用いて表される定数である。特に、 $\alpha = -1/2$ の場合に inverse Gauss 分布といい、さらに、 $\gamma = 0$ の場合は one-sided $1/2$ stable 分布と呼ばれている。

GIG 分布をはじめとする自己分解可能分布は最近流行の数理ファイナンスの分野で注目を集めているようである。詳しくは増田 (2002) を参照されたい。

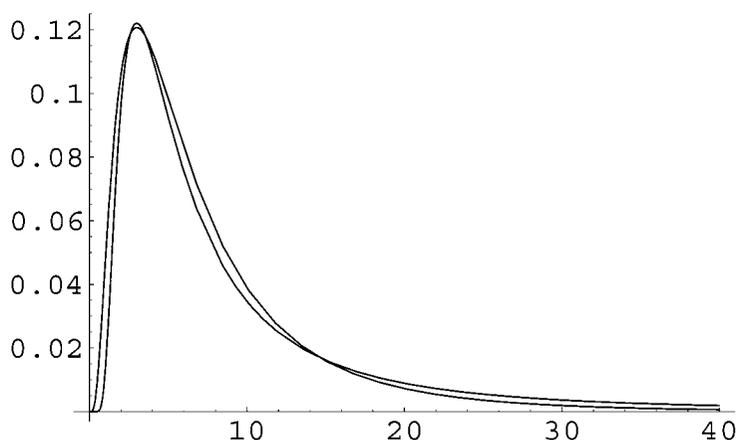
この分布を眺めてみると、 $\gamma > 0$ の場合は確かに分布の裾が任意のべき乗関数 x^{-k} , $k > 0$ より早くゼロに収束して対数正規分布に近い。しかし、 $\gamma = 0$ とすると、分布であるためには α

< 0 の条件が付くが正しく条件 (ii) (弱パレート性) を満たす^{注4)}。この特別な場合は、逆数ガンマ分布と呼ばれている。確率変数 X の分布がガンマ分布に従うとき、確率変数 $Y = 1/X$ の分布を逆数ガンマ分布という。すなわち、 X, Y の分布の密度関数を $f_{\Gamma}, f_{1/\Gamma}$ とすると、

$$f_{\Gamma}(x; \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad 0 < x < +\infty, \quad \alpha, \beta > 0,$$

$$f_{1/\Gamma}(x; a, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\beta/x}, \quad 0 < x < +\infty, \quad \alpha, \beta > 0.$$

で表される分布のことである。



対数正規分布の密度関数と逆数ガンマ分布の密度関数のグラフ

上図のグラフは対数正規分布の密度関数 $f_{\ell}(x; m, 0.8)$ と逆数ガンマ分布の密度関数 $f_{1/\Gamma}(x; 3/2, 7.5)$ のグラフを重ねた図である。ただし、モードを一致させるために $m = \log(7.5 \times e^{0.64}/2.5)$ としてある。裾のオーダーの違いをみれば、どちらのグラフであるかを特定することは容易である。(モードを一致させてもなお互いにひとつのパラメータの自由性が残るから、この図をさらに改良することは可能である。)

重ね合わせたグラフを眺める限りでは、所得分布の近似としてどちらの分布を用いてもよいと思われまいだろうか。しかし、理論的には、分布の裾(高額所得者の分布)に関して決定的な違いがあることは前述した。さらにジニ係数(または今回提案する修正ジニ係数)についても対数正規分布と逆数 Γ 分布とは決定的な違いがある。すなわち、対数正規分布のジニ係数と修正ジニ係数は一致し、 μ には無関係で σ にのみ依存する。 σ はもとの正規分布で考えると社会のある種の分散を表していると思われるから、 σ が大きくなるとジニ係数すな

注4) 増田(2002)には、このことは明示的に指摘してある。175頁。

わち不平等指数が大きくなるのは自然に思われたかもしれない。しかしながら、社会の不平等性増加の責任が構成員全員に等しくあるのであろうか。もし逆数 Γ 分布で全体の国民所得が近似できるとした場合、この分布のジニ係数（修正ジニ係数でも同様）は高額所得者の分布に関係している（パレート分布と同じ）パラメータ α のみに依存して決まるのである（次節参照）。社会の不平等性増加の責任はもっぱら高額所得者の増加に原因がある、という事実の方が庶民感覚としてはより納得がゆくであろう。つまり、対数正規分布で国民所得全体の分布を近似するのは理論的にも「まずい」のである。

§4. ジニ係数およびその修正

所得の不平等性をどのように指標化するべきかについて、経済学の分野ではいろいろな基準を設けて議論されている（例えば高山（1974）を参照されたい）が、本稿では純粋に確率論的に考察する。所得の不平等さを所得分布の平均からの散らばり具合、と理解する場合、よく知られている指標は分散 $E[(X - \mu)^2] = \sigma^2$ ないしその平方根をとった標準偏差である。貨幣単位の変更に対して不変である、つまり、 X の指標と任意の定数 $a > 0$ に対して aX の指標が一致するべきである、ということのを要請するならば、平均との比 σ/μ （変動係数という）を用いればよい。あるいは、平均偏差 $E[|X - \mu|]$ と平均との比、相対平均偏差を考えてもよい。しかし、いずれの場合もこれらの量に理論的上限が存在しない。その点、ジニ係数は都合がよい。取りあえず、定義を与える。ジニ係数は、平均が有限に存在し、非負の半直線上に台を持つすべての確率分布に対して定義される量であるが、そのような分布を持つ確率変数を用いて表現する方が解析には都合がよい。なお、平均の存在を仮定しない一般の1次元分布に対して散らばり具合を表す dispersion という量が知られているが残念ながら上記の意味の不変性を満たさない。

定義 5. 非負の値を取り、正の平均 $E[X] = \mu$ が有限に存在する確率変数 X のジニ係数 G_X を次の式で定義する。

$$G_X = \frac{E[|X_1 - X_2|]}{2\mu}.$$

ここで、 X_1, X_2 は X と同じ分布を持つ独立な確率変数である。定義式から容易に

$$0 \leq G_X = \frac{E[|X_1 - X_2|]}{2\mu} \leq \frac{E[|X_1|] + E[|X_2|]}{2\mu} = \frac{2E[X]}{2\mu} = 1$$

を得る。ただし、後述するように第2項の不等式において、等号となることはないから、ジニ係数の範囲は $0 \leq G_X < 1$ であり、 X_1 と X_2 が独立であることを考慮すると、 $G_X = 0$ となる

のは X が単位分布（完全平等）を持つ場合に限る^{注5)}。

n 人の個人所得のデータ, $\{x_1, x_2, \dots, x_n\}$ が与えられた場合は, 確率変数 X の分布として $P(X=x_k) = 1/n$ という離散分布を想定すればよいから, この場合のジニ係数はよく知られているように

$$\mu = E[X] = \frac{\sum_{j=1}^n x_j}{n}, G_X = \frac{E[|X_1 - X_2|]}{2\mu} = \frac{1}{2\mu n^2} \sum_{j,k=1}^n |x_j - x_k|$$

となる。

ただし, データがランダムサンプルの場合, それらを独立同分布をもつ確率変数列 (i.i.d.) の実現値とみなす, という数理統計学の立場に立った場合,

$$E\left[\sum_{j,k=1}^n |x_j - x_k|\right] = 2n(n-1)\mu G_X$$

だから, 平均 μ を既知のパラメータとした場合

$$\hat{G}_X = \frac{1}{2\mu n(n-1)} \sum_{j,k=1}^n |x_j - x_k|$$

が G_X の不偏推定量である^{注6)}。

このように一般的に表現して理論分布に適用すると直ちに問題があることがわかる。すなわち, 平均が発散する分布に対しては定義できないのである。もちろん, 現実のデータでは平均所得は必ず有限であるからこのような心配をする必要はないが, 現実の所得分布を要請(ii)の弱パレート性を満たす理論分布で近似して, パラメータが $1 \geq \alpha > 0$ の範囲にある場合, ジニ係数を理論的に算出できない, という不都合が生じる。

そこで, 所得分布の理論分布が有限な平均を持たない場合にもジニ係数に相当する量を定

注5) 高山(1974)45頁には, 完全独占, つまり一人の人間が全所得を独占した場合にジニ係数が1である, と解説してあるが数学的には正しくない。この場合, 正確なジニ係数は $1 - 1/n$ である。ここで n は構成員の人数。この場合, $n \rightarrow +\infty$ としても確率分布としては収束しない。もし, n を無限大にすれば1に近づくからよい, というのであれば連続分布でもパレート分布においてパラメータを1に近づけるとジニ係数はいくらでも1に近づけることが出来る。つまり, ジニ係数を不平等の指標とする限り, 一人の人間が富を独占しなくても極端な不平等は起こり得る。この事実は理論と直感との齟齬というよりむしろ, 直感的には捕らえにくい極端な不平等さがあり得るという意味でむしろ重要ではなからうか。

なお, たとえば A. Sen (1976) でも $1/n$ を無視ないし省略してある議論もあるが, その場合は命題の記述にかならず, 「大きな n に対して,」 というような条件が付してあり, 決して論理的飛躍はしていない。

青木(1979)95頁にも「完全不平等の場合, (ジニ係数は)1となる」とある。「完全不平等」なる言葉の定義がないので意味不明であるが, 本稿の定義に従う限り, ジニ係数が1となる非負値平均有限な確率分布は存在しない。

注6) Kendall-Stuart (1963) p. 46 には n^2 で割る場合と, $n(n-1)$ で割る場合が理由もなく併記しており, n が大きければたいした違いはないと書いてあるが, 理論的には大きな違いである。

義するために次のようなジニ係数の修正を提案する。

定義 6. 正の値をとる確率変数 X に対して $\mu_m \equiv E[1/X] < +\infty$ を仮定する（現実の所得分布を議論する場合，この仮定は不自然ではないであろう）。この時，確率変数 $1/X$ の分布に対するジニ係数 $G_{1/X}$ を X の分布の修正ジニ係数（modified Gini coefficient）と呼び， mG_X で表す。つまり，

$$mG_X \equiv G_{1/X} = \frac{1}{2\mu_m} E\left[\left|\frac{1}{X_1} - \frac{1}{X_2}\right|\right] = \frac{1}{2\mu_m} E\left[\frac{|X_1 - X_2|}{X_1 X_2}\right].$$

ここで， X_1, X_2 は X と同じ分布を持つ独立な確率変数である。定義から直ちにわかるように，ジニ係数，修正ジニ係数ともに貨幣単位の変更，つまり定数 $a > 0$ に対して， $X = aY$ なる関係にある場合には Y のジニ係数，修正ジニ係数はそれぞれ X のジニ係数，修正ジニ係数と一致する。 X が連続分布をもつ場合，その密度関数を f_X とすると，

$$\mu_m = \int_0^{+\infty} \frac{f_X(x)}{x} dx, \quad mG_X = \frac{1}{2\mu_m} \int_0^{+\infty} \int_0^{+\infty} \left|\frac{1}{x} - \frac{1}{y}\right| f_X(x) f_X(y) dx dy$$

である。なお， n 人の個人所得のデータ， $\{x_1, x_2, \dots, x_n\}$ が与えられた場合の修正ジニ係数は

$$\mu_m = \sum_{j=1}^n \frac{1}{x_j}, \quad mG_X = \frac{1}{2\mu_m} \sum_{j,k=1}^n \left| \frac{1}{x_j} - \frac{1}{x_k} \right|$$

で求められる。

なお，田口（1979）では単調増加正値連続関数 $\zeta(x)$ を導入して確率変数 X を $\zeta(X)$ と変換して一般化する方法が研究されている。

ジニ係数および修正ジニ係数の計算例：

(1) パレート分布（パラメータ $\alpha > 0$ ）：

$$G_p(\alpha) = \frac{1}{2\alpha - 1}, \quad (\alpha > 1). \quad mG_p = G_p(\alpha + 1) = \frac{1}{2\alpha + 1}, \quad (\alpha > 0).$$

(2) 対数正規分布（パラメータ μ, σ^2 ）：ジニ係数 = 修正ジニ係数である。

$$G_\ell(\mu, \sigma^2) = mG_\ell(\mu, \sigma^2) = 2N\left(\frac{\sigma}{\sqrt{2}}\right) - 1, \quad N(x) \text{ は標準正規分布の分布関数：}$$

(3) 定義から明らかに，

ガンマ分布のジニ係数 $G_\Gamma(\alpha, \beta) =$ 逆数ガンマ分布の修正ジニ係数 $mG_{1/\Gamma}(\alpha, \beta) (\alpha > 0, \beta > 0)$,

ガンマ分布の修正ジニ係数 $mG_\Gamma(\alpha, \beta) =$ 逆数ガンマ分布のジニ係数 $G_{1/\Gamma}(\alpha, \beta) (\alpha > 1,$

$\beta > 0$) であって、

$$G_{\Gamma}(\alpha, \beta) = mG_{1/\Gamma}(\alpha, \beta) = \frac{\Gamma(2\alpha+1)}{2^{2\alpha}\Gamma^2(\alpha+1)}, \quad (\alpha > 0)$$

である (付録参照)。

ガンマ関数は連続で $\Gamma(1) = 1$ だから

$$\lim_{\alpha \downarrow 0} G_{\Gamma}(\alpha, \beta) = 1$$

に注意しておく。つまり、完全独占でなくてもジニ係数を 1 に近づけることはできるのである。

逆数ガンマ分布のジニ係数は $\alpha > 1$ の場合しか定義できず、

$$G_{1/\Gamma}(\alpha, \beta) = mG_{\Gamma}(\alpha, \beta) = G_{\Gamma}(\alpha-1, \beta), \quad (\alpha > 1)$$

である。

(1) の計算は初等的である。(2) 対数正規分布 $\Lambda(\mu, \sigma^2)$ のジニ係数はよく知られていて、結果をみると分かるように μ に関係していない。確率変数 X の分布が対数正規分布 $\Lambda(\mu, \sigma^2)$ に従うと仮定すると定義から $1/X$ の分布は対数正規分布 $\Lambda(-\mu, \sigma^2)$ に従う。よって、修正ジニ係数はジニ係数に一致する。所得分布が逆数ガンマ分布をしていると仮定した場合、ジニ係数にしろ、修正ジニ係数にしろパラメータ α にのみ関係している。つまり、不平等増加の原因は高額所得者がますます増えるからである。対数正規分布で全体を近似した場合のジニ係数より庶民感覚にはよりフィットした結果である。なお、(2)、(3) の導出は付録を参照されたい。

コメント：ヒストグラムを逆数ガンマ分布の密度関数で近似する場合、高額所得者の分布のオーダー (パレート分布のパラメータ) とモードを一致させると、パラメータの α と β が決まってしまう。(モードの x 座標は $\beta/(\alpha+1)$) しかし、このときに密度関数がよい近似を与えない場合はもうひとつのパラメータを導入して、密度関数として

$$f(x) = \frac{\gamma\beta^{\alpha/\gamma}}{\Gamma(\alpha/\gamma)} x^{-\alpha-1} e^{-\beta/x^\gamma}, \quad \alpha, \beta, \gamma, x > 0$$

を考えればよいであろう。つまり、 X^γ の分布が逆数ガンマ分布に従う、と考えることと同値である。この場合、密度関数のモードは $x = (\beta\gamma/(\alpha+1))^{1/\gamma}$ であるから、高額所得者の分布のオーダー (α) とモードまで一致させてもなおひとつの自由度が残るから、分布形をさらに調節することが出来る。この関数に対して我々の要請 (i)、(ii) を満たすことは明らかであるが、(iii) については $\gamma \leq 1$ の場合は自己分解可能であることが知られているとのことである (佐藤健一氏からの私信)。

§5. ローレンツ曲線とジニ係数の関係について

通常、ジニ係数はローレンツ曲線と直線 $x=y$ で囲まれる面積の 2 倍で定義する場合が多い。では、ローレンツ曲線はどう定義するか。解説書はもちろん、学術論文を参照しても相当厳密に考察してある文献においてすら、離散分布の場合と正の連続分布を持つ場合の定義が別々に与えられている。本来ローレンツ曲線の定義は、非負値、平均有限の確率変数（確率分布）に対して統一的に与えられるべきものである（従って、もちろん特異分布に対しても定義できる）。以下でこのことを試みる。ただし、厳密な数学的証明にはいわゆる $\varepsilon-\delta$ 論法が必要であるので一部分の証明を省略した。

ローレンツ曲線の一般的定義

非負値確率変数を X とし、 $0 < \mu = E[X] < +\infty$ を仮定する。さらに N を X とは独立な標準正規分布を持つ確率変数とする。 $\varepsilon > 0$ に対して $X_\varepsilon = X + \varepsilon N$ とおく。 $E[X_\varepsilon] = E[X] = \mu$ であることに注意する。よく知られているように^{注7)} このとき、 X の分布の如何に関わらず、 X_ε は至るところ正で、無限回微分可能な密度関数 f_ε を持つ。すなわち、 $-\infty < t < +\infty$ に対して

$$\begin{aligned} F_\varepsilon(t) &= P(X + \varepsilon N \leq t) \\ &= \int_{-\infty}^t \frac{1}{\sqrt{2\pi\varepsilon}} \exp\left(-\frac{u^2}{2\varepsilon^2}\right) F_X(t-u) du \end{aligned} \quad (1)$$

$$= \int_0^{+\infty} \left(\int_{-\infty}^{t-v} \frac{1}{\sqrt{2\pi\varepsilon}} \exp\left(-\frac{u^2}{2\varepsilon^2}\right) du \right) dF_X(v). \quad (2)$$

ただし、(2)式の積分は一般にはルベーグ・スティルチェス積分といわれるものである。離散分布の場合は積分の代わりに Σ で表せばよい。また、連続な密度関数を持つ連続分布の場合は通常の広義リーマン積分である。また、(2)式から X_ε の分布は連続分布であって、その密度関数は

$$f_\varepsilon(t) = \int_0^{+\infty} \frac{1}{\sqrt{2\pi\varepsilon}} \exp\left(-\frac{(t-v)^2}{2\varepsilon^2}\right) dF_X(v)$$

と与えられ、この表現から至るところ正の値を取る無限回微分可能な関数であることがわかる。

一方、よく知られているように分布関数 F_X は右連続、左極限をもつ単調非増加な関数であ

注7) 偏微分方程式論ではよく知られている。たとえば、溝畑茂 (1965) 28頁。彼はこのような操作（作用素）を Friedrichs に従って mollifier（軟化子）と呼んでいる。

るから、 $\lim_{\varepsilon \downarrow 0} F_X(t - \varepsilon) = F_X(t - 0)$ と書くと、(1)式から

$$\lim_{\varepsilon \downarrow 0} F_\varepsilon(t) = \frac{F_X(t) + F_X(t - 0)}{2}$$

であることが容易にわかる。従って、 F_X の連続点では F_ε は各点収束する。

さて、 $-\infty < t < +\infty$ をパラメータとして次のような曲線 $(x_\varepsilon(t), y_\varepsilon(t))$ が定義できる。

$$x_\varepsilon(t) = F_\varepsilon(t), \quad y_\varepsilon(t) = \frac{1}{\mu} \int_{-\infty}^t u f_\varepsilon(u) du, \quad -\infty < t < +\infty.$$

任意の $-\infty < t < +\infty$ に対して、 $f_\varepsilon(t) > 0$ だから $F_\varepsilon(t)$ は真に単調増加関数である。従って、 y_ε は x_ε の関数であり、かつその導関数と2階の導関数は合成関数の微分法によって、

$$\frac{dy_\varepsilon}{dx_\varepsilon} = \frac{dy_\varepsilon}{dt} \frac{dt}{dx_\varepsilon} = \frac{t}{\mu}, \quad \frac{d^2 y_\varepsilon}{dx_\varepsilon^2} = \frac{1}{\mu f_\varepsilon(t)} > 0$$

を得る。さらに、

$$\lim_{t \rightarrow -\infty} x_\varepsilon(t) = 0, \quad \lim_{t \rightarrow +\infty} x_\varepsilon(t) = 1, \quad \lim_{t \rightarrow -\infty} y_\varepsilon(t) = 0, \quad \lim_{t \rightarrow +\infty} y_\varepsilon(t) = 1$$

を得るから、 (x, y) 座標上では、関数 $y = y_\varepsilon(x)$ のグラフは $0 \leq x \leq 1$ で下に凸な連続関数であり、 $-\infty < t < 0$ で減少、 $0 < t < +\infty$ で増加関数である。こうして得られる曲線を X_ε に対応するローレンツ曲線と呼ぶことにする。この定義は多くの文献で連続分布に対するローレンツ曲線の定義と一致している（非負値確率変数で正の半直線上至るところ正の密度関数をもつということを仮定すれば、上記の考察で $-\infty < t < 0$ の部分を考える必要はない）。

定義7. 上記の X_ε に対応するローレンツ曲線において、 $\varepsilon \downarrow 0$ として得られる極限の曲線を X （の分布）に対応するローレンツ曲線と呼ぶ。

もちろん、この定義が well defined であるためには、 $\varepsilon \downarrow 0$ としたときに X_ε に対応するローレンツ曲線がひとつの曲線に収束することを証明する必要があるが、いわゆる ε - δ 論法を駆使する必要があるので省略する。離散分布の場合に、通常の場合と一致することのみを例示しておく。

$0 = t_0 < t_1 < \dots < t_n$ に対して、 $P(X = t_k) = p_k > 0, k = 1, 2, \dots, n.$

$$\mu = \sum_{j=1}^n t_j p_j, \quad x_k = \sum_{j=1}^k p_j, \quad y_k = \frac{1}{\mu} \sum_{j=1}^k t_j p_j,$$

とおく。点、 $(0, 0), (x_1, y_1), \dots, (x_n, y_n) = (1, 1)$ を線分で結んで得られる折れ線が離散分布の場合のローレンツ曲線の通常の場合の定義である。

この場合、分布関数は

$$F_X(t) = \begin{cases} 0 & ; 0 \leq t < t_1, \\ \sum_{j=1}^{k-1} p_j & ; t_{k-1} \leq t < t_k, \quad k = 2, \dots, n \\ 1 & ; t_n \leq t < +\infty \end{cases}$$

である。\$F_X(t_{k-1}) < x < F_X(t_k)\$ とする。\$F_\varepsilon(t)\$ は真に単調増加関数であるから、\$x = F_\varepsilon(t)\$ を満たす \$t = t_\varepsilon(x)\$ が一意に定まる。このとき、部分積分を用いて、

$$y_\varepsilon(x) = \frac{1}{\mu} \int_{-\infty}^{t_\varepsilon(x)} u f_\varepsilon(u) du = \frac{t_\varepsilon(x) F_\varepsilon(t_\varepsilon(x))}{\mu} - \frac{1}{\mu} \int_{-\infty}^{t_\varepsilon} F_\varepsilon(u) du$$

と表せる。

$$\lim_{\varepsilon \downarrow 0} t_\varepsilon(x) = t_k \tag{3}$$

を考慮すると (\$x \neq F_X(t_j)\$ の点のみで考察していることに注意されたい) (3) より

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} y_\varepsilon(x) &= \frac{t_k x}{\mu} - \frac{1}{\mu} \int_0^{t_k} F_X(u) du = \frac{t_k x}{\mu} - \frac{1}{\mu} \sum_{j=2}^k \int_{t_{j-1}}^{t_j} F_X(u) du \\ &= \frac{t_k x}{\mu} - \frac{1}{\mu} \sum_{j=2}^k \sum_{i=1}^{j-1} p_i (t_j - t_{j-1}) = \frac{t_k x}{\mu} - \frac{1}{\mu} \sum_{i=1}^{k-1} \sum_{j=i+1}^k p_i (t_j - t_{j-1}) \\ &= \frac{t_k x}{\mu} - \frac{1}{\mu} \sum_{i=1}^{k-1} p_i (t_k - t_i) = \frac{t_k}{\mu} (x - x_{k-1}) + y_{k-1} \end{aligned}$$

となって、点 \$(x_{k-1}, y_{k-1})\$ と \$(x_k, y_k)\$ を結ぶ直線の方程式が得られるから従来の定義と一致することがわかる。

次に、ローレンツ曲線と直線 \$x = y\$ で囲まれる面積の 2 倍がジニ係数と一致することを証明しよう注8)。そのためには、\$X_\varepsilon\$ に関するローレンツ曲線と \$x\$ 軸、直線 \$x = 1\$ で囲まれる面積を \$S_\varepsilon\$ とおく。ただし、\$y_\varepsilon < 0\$ の部分の面積は負とする。密度関数 \$f_\varepsilon\$ は十分滑らかで、\$t \to \pm\infty\$ のときに十分早くゼロに収束するから通常の微積分の公式が安心して使えることに注意する。

$$\begin{aligned} S_\varepsilon &= \int_0^1 y_\varepsilon(x) dx \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^t u f_\varepsilon(u) du \right) f_\varepsilon(t) dt / \mu \\ &= \left(\int_{-\infty}^t u f_\varepsilon(u) du F_\varepsilon(t) \right) \Big|_{t=-\infty}^{+\infty} - \int_{-\infty}^{+\infty} t f_\varepsilon(t) F_\varepsilon(t) dt / \mu \\ &= 1 - \int_{-\infty}^{+\infty} t f_\varepsilon(t) F_\varepsilon(t) dt / \mu \end{aligned} \tag{4}$$

注 8) Kendall-Stuart (1963) pp. 46-49 に同様のことが書いてある。ただし、彼の場合は連続分布の場合のみを扱っている。しかも密度関数に何らの仮定がないので、通常の微積分の公式が使えるかどうか若干不安である。なお、彼はローレンツ曲線のことを the curve of concentration と呼んでいる。

従って,

$$\int_{-\infty}^{+\infty} t f_{\varepsilon}(t) F_{\varepsilon}(t) dt = \mu(1 - S_{\varepsilon}) \quad (5)$$

一方, 独立で X_{ε} と同分布を持つ確率変数を $X_{\varepsilon,1}, X_{\varepsilon,2}$ とすると,

$$\begin{aligned} E[|X_{\varepsilon,1} - X_{\varepsilon,2}|] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |t - s| f_{\varepsilon}(t) f_{\varepsilon}(s) dt ds \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^s (s - t) f_{\varepsilon}(t) dt + \int_s^{+\infty} (t - s) f_{\varepsilon}(t) dt \right) f_{\varepsilon}(s) ds \\ &= \int_{-\infty}^{+\infty} s F_{\varepsilon}(s) f_{\varepsilon}(s) ds - \int_{-\infty}^{+\infty} \left(\int_{-\infty}^s t f_{\varepsilon}(t) dt \right) f_{\varepsilon}(s) ds \\ &\quad + \int_{-\infty}^{+\infty} \left(\mu - \int_{-\infty}^s t f_{\varepsilon}(t) dt - s + s F_{\varepsilon}(s) \right) f_{\varepsilon}(s) ds \\ &= 2 \left(\int_{-\infty}^{+\infty} s F_{\varepsilon}(s) f_{\varepsilon}(s) ds - \int_{-\infty}^{+\infty} \left(\int_{-\infty}^s t f_{\varepsilon}(t) dt \right) f_{\varepsilon}(s) ds \right) \\ &= 2 \int_{-\infty}^{+\infty} s F_{\varepsilon}(s) f_{\varepsilon}(s) ds - 2\mu S_{\varepsilon} \quad ((4) \text{より}), \\ &= 2\mu(1 - S_{\varepsilon}) - 2\mu S_{\varepsilon} \quad ((5) \text{より}), \\ &= 2\mu(1 - 2S_{\varepsilon}) \end{aligned}$$

故に

$$G_{X_{\varepsilon}} = \frac{E[|X_{\varepsilon,1} - X_{\varepsilon,2}|]}{2\mu} = 1 - 2S_{\varepsilon}$$

を得る。ここで, $\varepsilon \downarrow 0$ とすると, $X_{\varepsilon,i}$ は X_i にいわゆる L^1 収束するから $G_{X_{\varepsilon}}$ は G_X に収束する。ローレンツ曲線が収束してその面積も収束したローレンツ曲線の面積に収束する, ということは厳密な証明を要するが省略する (曲線が同様収束することを証明することは結構大変である)。ローレンツ曲線と直線 $x = y$ で囲まれる面積は $1/2$ より真に小さいからジニ係数が 1 となることは理論的にあり得ない。

付録 ガンマ分布, 逆数ガンマ分布および対数正規分布のジニ係数と修正ジニ係数の計算。
まず, 平均 $\mu_{\Gamma}(\alpha, \beta)$, $\mu_{1/\Gamma}(\alpha, \beta)$ はよく知られた Γ 関数の計算であるから, 容易に

$$\mu_{\Gamma}(\alpha, \beta) = \frac{\alpha}{\beta}, \quad (\alpha > 0), \quad \mu_{1/\Gamma}(\alpha, \beta) = \frac{\beta}{\alpha - 1}, \quad (\alpha > 1)$$

を得る。次に, ガンマ分布を持つ独立な確率変数を X_{Γ}, Y_{Γ} として, ジニ係数の定義通りに計算する。

$$\begin{aligned}
 E[|X_\Gamma - Y_\Gamma|] &= \int_0^{+\infty} \int_0^{+\infty} |x-y| f_\Gamma(x; \alpha, \beta) f_\Gamma(y; \alpha, \beta) dx dy \\
 &= \frac{\beta^{2\alpha}}{\Gamma^2(\alpha)} \int_0^{+\infty} \int_0^{+\infty} |x-y| x^{\alpha-1} y^{\alpha-1} e^{-\beta(x+y)} dx dy \\
 &= \frac{1}{\beta \Gamma^2(\alpha)} \int_0^{+\infty} \int_0^{+\infty} |s-t| s^{\alpha-1} t^{\alpha-1} e^{-(s+t)} ds dt \quad (s = u^2, s = v^2) \\
 &= \frac{4}{\beta \Gamma^2(\alpha)} \int_0^{+\infty} \int_0^{+\infty} |u^2 - v^2| u^{2\alpha-1} v^{2\alpha-1} e^{-(u^2+v^2)} dudv \\
 &\quad (u = r \cos \theta, v = r \sin \theta, dudv = r dr d\theta) \\
 &= \frac{4}{\beta \Gamma^2(\alpha)} \int_{0 \leq r, 0 \leq \theta \leq \pi/2} r^{4\alpha+1} e^{-r^2} |\cos^2 \theta - \sin^2 \theta| \cos^{2\alpha-1} \theta \sin^{2\alpha-1} \theta dr d\theta \\
 &= \frac{2^{3-2\alpha}}{\beta \Gamma^2(\alpha)} \int_0^{+\infty} r^{4\alpha+1} e^{-r^2} dr \int_0^{\pi/2} |\cos 2\theta| \sin^{2\alpha-1} 2\theta d\theta \quad (r^2 = t, 2\theta = \phi) \\
 &= \frac{2^{1-2\alpha} \Gamma(2\alpha+1)}{\beta \Gamma^2(\alpha)} \left(\int_0^{\pi/2} \cos \phi \sin^{2\alpha-1} \phi d\phi - \int_{\pi/2}^{\pi} \cos \phi \sin^{2\alpha-1} \phi d\phi \right) \\
 &= \frac{2^{2-2\alpha} \Gamma(2\alpha+1)}{\beta \Gamma^2(\alpha)} \int_0^1 t^{2\alpha-1} dt \\
 &= \frac{2^{1-2\alpha} \Gamma(2\alpha+1)}{\alpha \beta \Gamma^2(\alpha)}.
 \end{aligned}$$

この式が $2\mu_\Gamma(\alpha, \beta)G_\Gamma(\alpha, \beta)$ に等しいわけであるから、よく知られた公式 $\alpha\Gamma(\alpha) = \Gamma(\alpha+1)$ を用いて、

$$G_\Gamma(\alpha, \beta) = \frac{\Gamma(2\alpha+1)}{2^{2\alpha} \Gamma^2(\alpha+1)}$$

を得る。

$G_{1/\Gamma}(\alpha, \beta)$ の計算も同様であるが、関係式

$$f_\Gamma(x; \alpha, \beta) / x = \frac{\beta \Gamma(\alpha-1)}{\Gamma(\alpha)} f_\Gamma(x; \alpha-1, \beta), \quad (\alpha > 1)$$

に注意すると容易に $G_{1/\Gamma}(\alpha, \beta) = G_\Gamma(\alpha-1, \beta)$, $(\alpha > 1)$ を得る。

(2) の対数正規分布のジニ係数に関してはほとんどの文献が Aitchison-Brown (1957) p. 13 Theorem 2.7 を根拠にしていると思われる。しかし、この定理の証明をさかのぼってゆくと、キーになる Theorem 2.2 はさらに Cramér (1946) p. 190 参照となっている。内容的には確率論の基本的なことであるが定理の形に表現していないようなので、簡潔に説明しておく。

有限次元確率分布は有名な Bochner (ボホナー) の定理 (たとえば、伊藤 1953, 96頁定理 8.1) によって、特性関数と一対一に対応している。これまた有名な Kac (カツツ) の定理 (たとえば、伊藤 1953, 64頁定理13.7) によって、有限個の確率変数が独立であるための必

要十分条件は、有限次元特性関数が 1 次元特性関数の積で表されることである。正規分布 $N(\mu, \sigma^2)$ の特性関数が $\exp(i\mu\zeta - \sigma^2\zeta^2/2)$ であることを知っている、二つの独立な確率変数の分布がそれぞれ、正規分布 $N(\mu_1, \sigma_1^2)$ と $N(\mu_2, \sigma_2^2)$ である場合、その和の分布が正規分布 $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ であることは Bochner の定理と Kac の定理から暗算で証明できる（直接密度関数の計算を実行すると結構大変である）。さて、正值確率変数 X の対数をとった確率変数の分布が正規分布 $N(\mu, \sigma^2)$ であるときの X の分布を対数正規分布というのであった。この分布を記号 $\Lambda(\mu, \sigma^2)$ と記し、分布関数を $\Lambda(x; \mu, \sigma^2)$ と書く（Aitchison, J. and Brown, J. A. C. (1957) と同じ記号である）。 X, Y をそれぞれ対数正規分布 $\Lambda(\mu_1, \sigma_1^2), \Lambda(\mu_2, \sigma_2^2)$ に従う独立な確率変数とする。定義と正規分布の性質から $Z = X/Y^b$ は対数正規分布 $\Lambda(\mu_1 - b\mu_2, \sigma_1^2 + b^2\sigma_2^2)$ に従う。従って、 $P(Z \leq a) = \Lambda(a; \mu_1 - b\mu_2, \sigma_1^2 + b^2\sigma_2^2)$ であるが、具体的に積分で表すと、

$$P(Z \leq a) = P(X \leq aY^b) = \int_0^{+\infty} \Lambda(ax^b; \mu_1, \sigma_1^2) d\Lambda(x; \mu_2, \sigma_2^2)$$

である。従って、Aitchison, J. and Brown, J. A. C. (1957) p. 11 の Corollary 2.2.b

$$\Lambda(a; \mu_1 - b\mu_2, \sigma_1^2 + b^2\sigma_2^2) = \int_0^{+\infty} \Lambda(ax^b; \mu_1, \sigma_1^2) d\Lambda(x; \mu_2, \sigma_2^2) \quad (6)$$

が得られる。なお、この等式を密度関数を用いて直接計算すると相当大変である。もちろん、大学初年級の微分積分であるから時間をかければ誰にでも出来る。

次の等式は対数正規分布特有の性質である。（Aitchison, J. and Brown, J. A. C. (1957) p. 12 Theorem 2.6 の特別な場合）対数正規分布の平均を $m = \exp(\mu + \sigma^2/2)$ とおいて、

$$\Lambda_1(x; \mu, \sigma^2) = \frac{1}{m} \int_0^x y d\Lambda(y; \mu, \sigma^2) = \int_0^x \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\mu - \frac{\sigma^2}{2} - \frac{(\log y - \mu)^2}{2\sigma^2}\right) dy$$

とおく。

$$u + \frac{\sigma^2}{2} + \frac{(\log y - \mu)^2}{2\sigma^2} = \frac{(\log y - \mu - \sigma^2)^2}{2\sigma^2} + 2\sigma^2 \log y$$

に注意すると、

$$\begin{aligned} \Lambda_1(x; \mu, \sigma^2) &= \int_0^x \frac{1}{\sqrt{2\pi}\sigma y} \exp\left(-\frac{(\log y - \mu - \sigma^2)^2}{2\sigma^2}\right) dy \\ &= \Lambda(x; \mu + \sigma^2, \sigma^2) \end{aligned} \quad (7)$$

が得られる。つぎに、対数正規分布のローロンツ曲線と x 軸、直線 $x = 1$ で囲まれる面積を S とおくと、

$$S = \int_0^1 y dx, \quad dx = \Lambda(t; \mu, \sigma^2), \quad y = \int_0^t s d\Lambda(s; \mu, \sigma^2) / m$$

だから,

$$\begin{aligned}
 S &= \int_0^{+\infty} \left(\int_0^t s d\Lambda(s; \mu, \sigma^2) \right) d\Lambda(t; \mu, \sigma^2) / m \\
 &= \int_0^{+\infty} \Lambda_1(t; \mu, \sigma^2) d\Lambda(t; \mu, \sigma^2) \\
 &= \int_0^{+\infty} \Lambda(t; \mu + \sigma^2, \sigma^2) d\Lambda(t; \mu, \sigma^2) \quad ((7)より,) \\
 &= \Lambda(1, \sigma^2, 2\sigma^2) \\
 &\quad (\text{公式(6)において } a=1, b=1, \mu_1 = \mu + \sigma^2, \mu_2 = \mu, \sigma_1^2 = \sigma_2^2 = \sigma^2 \text{とした}) \\
 &= \int_0^1 \frac{1}{2\sqrt{\pi}\sigma x} \exp\left(-\frac{(\log x - \sigma^2)^2}{4\sigma^2}\right) dx \\
 &\quad \left(\frac{\log x - \sigma^2}{\sqrt{2}\sigma} = y \text{とおくと, } \frac{dx}{\sqrt{2}\sigma x} = dy \text{ だから}\right) \\
 &= \int_{-\infty}^{-\sigma/\sqrt{2}} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = 1 - \int_{-\infty}^{\sigma/\sqrt{2}} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\
 &= 1 - N(\sigma/\sqrt{2})
 \end{aligned}$$

従って, $G_\ell(\mu, \sigma^2) = 1 - 2S = 2N(\sigma/\sqrt{2}) - 1$ が得られる。

謝辞：無限分解可能分布，L分布，安定分布およびL分布の単峰性証明の経緯等に関して，佐藤健一氏より種々ご教示頂いた。記して感謝の意を表したい。もちろん，本稿の内容に関する責任は全面的に著者にある。

参 考 文 献

- [1] Aitchison, J. and Brown, J. A. C. (1954) "On criteria for description of income distribution" *Metroeconomica* vol. VI, 88-107.
- [2] Aitchison, J. and Brown, J. A. C. (1957) *Lognormal distribution*, Cambridge at the University Press.
- [3] 青木昌彦 (1979) 『分配理論』第二版経済学全集13筑摩書房.
- [4] Champnowne, D. G. (1953) "A model of income distribution" *The Economic Journal*, vol. LXIII, 318-351.
- [5] Champnowne, D. G. (1973) *The Distribution of Income Between Persons*. University Press.
- [6] Cramér, H. (1946) *Mathematical methods of statistics* Princeton Univ. Press.
- [7] Gastwirth, J. L. (1972) "The estimation of the Lorenz curve and Gini index" *The Review of Economics and Statistics*. 306-316.
- [8] Gumbel, E. J. (1958) *Statistics of extremes*, Columbia Univ. Press. 『極値統計学』河田龍夫他監訳 (1963) 広川書店。
- [9] Gnedenko, B. V., Kolmogorov, A. N. (1954) *Limit distributions for sums of independent random variables*, (translated by K. L. Chung). Addison-Wesley.
- [10] 浜田宏 (1999) 「所得分布の生成モデルー反復投資ゲームによる定式化」『理論と方法』vol. 14, no. 2, 3-17.

- [11] 依田高典 (1997) 『不確実性と意思決定の経済学』日本評論社.
- [12] 伊藤清 (1953) 『確率論』現代数学14 岩波書店。
- [13] Kakwani, N. C. and Podder, N. (1973) “On the estimation of Lorenz curves from grouped observations” *International Economic Review*, vol. 14, no. 2, 278–292.
- [14] Kalecki, M. (1945) “On the Gibrat distribution” *Econometrica* vol. 13, no. 2, 161–170.
- [15] Kendall, M. G. and Stuart, A. (1963) *The advanced theory of statistics. vol. 1* second ed. Charles Griffin.
- [16] Kimura, K. (1994) “A micro-macro linkage in the measurement of inequality: another look at the Gini coefficient. *Quality & Quantity* 28. 83-97.
- [17] 河野 敬雄 (2005) 「所得分布は逆数ガンマ分布に従うこと、並びにジニ係数の修正について」第39回数理社会学会2005年3月4日新潟国際情報大学ポスターセッション。
- [18] 倉林義正 (1973) 「わが国における所得分布と富の階層別分布」『週刊東洋経済臨時増刊』69–76.
- [19] Mandelbrot, B. (1960) “The Pareto-Lévy law and the distribution of income” *International Economic Review* vol. 1, No. 2, 79-106.
- [20] 増田弘毅 (2002) 「GIG 分布とGH 分布に関する解析」『統計数理』第50 巻第2 号165-199.
- [21] 宮沢健一 (1976) 『国民所得理論三訂版』第二版経済学全集11 筑摩書房.
- [22] 溝畑茂 (1965) 『偏微分方程式論』現代数学9, 岩波書店.
- [23] Morgan, J. (1962) “The anatomy of income distribution” *The Review of Economics and Statistics*, vol. 44, 270–282.
- [24] Sahota, G. S. (1978) “Theories of personal income distribution: Survey, *The Journal of Economic Literature*, vol. XVI, no. 1. 1–55.
- [25] 佐藤健一 (1990) 『加法過程』紀伊国屋数学叢書33。
- [26] Sato, K. (1999) *Lévy processes and infinitely divisible distribution*, Cambridge University Press, Cambridge.
- [27] Sen, A. (1976) “Poverty: An ordinal approach to measurement” *Econometrica*, Vol. 44, No. 2, 219–231.
- [28] Sen, A. (1997) *On economic inequality*, Oxford University Press. 『不平等の経済学』(2000) 鈴木・須賀訳, 東洋経済新報社
- [29] 田口時夫 (1979) 「Amoroso 型所得分布の解析——一般化ガンマ分布とローレンツ曲線の一般化——」*統計数理*26 巻 1, 2 号81–106.
- [30] 高山憲之 (1974) 「所得不平等の尺度：再検討」*国民経済*No. 131, 41-69.
- [31] Takayama, N. (1979) “Poverty, income inequality, and their measures: Professor Sen’s axiomatic approach reconsidered” *Econometrica*, Vol. 47, No. 3, 747–759.
- [32] Theil, H. (1967) *Economics and information theory*. *Studies in Mathematical and Managerial Economics*, vol. 7. North-Holland Pub.co.
- [33] 綿貫伸一郎 (1977) 「所得不平等の測定に関するノート——アトキンソンの命題をめぐる——」*季刊理論経済学* Vol. XXVIII, No. 2, 155–159.
- [34] ——— (1978) 「ローレンツ曲線とジニ係数」『経済論叢』第121 巻第45号京都大学経済学会 36–60.
- [35] 山本啓三・宮島佐介 (2001) 「日本の高額所得納税額のベキ乗則について」『理論と方法』vol. 16, no. 1, 133–138.
- [36] Yamazato, M. (1978) Unimodality of infinitely divisible distribution functions of class L, *Ann. Probab.* 6, 523–531.