『学習院大学　経済論集』第37巻　第3・4号（2001年1月）

ジニ係数の変化を判断する尺度について

竹内　俊子　　新居　玄武

1　はじめに

近年になって，所得分配の不平等化傾向に関する議論が再燃している。平等とは何かという形而上学的な議論は抜きにして，不平等を，一様分配あるいは所得分布の一様分布からの乖離ととらえれば，不平等度を何らかの形で数量化し示す指標はいくつも考えられる。数学的には関数空間における距離を定義することになる。しかし，直感的にわかりやすいこと，いろいろなよい性質を持つことにより，実際にはジニ係数がほとんどの実証研究で使われている。そこで本稿では，不平等度の尺度としてジニ係数を用いることを前提として議論をすすめていきたい。

そもそも2時点間の所得分布が変化したことをどう判断し，どのように結論付けたらよいだろうか。従来，実証分析で用いられるジニ係数は記述統計量としてとらえられていたため，その数値の大小関係をもって所得分配の平等化，不平等化を判断し結論付けていた。ただし，サンプル数，サンプリングバイアス，サンプル誤差を慮って，かなり大きな変化があったときにのみ結論を出していた傾向がある。その際でも，変化の大きさによって，「やや」あるいは「若干」といったレトリックでその信頼性を伝えていたと考えられる。

ジニ係数を，ある確率分布に従う確率変数の実現値から計算された統計量と考えれば，2時点間の比較は，2つの確率変数（必ずしも独立ではない）の大小比較となり，確率的な表現，たとえば$P(X_t < X_{t+1}) = p$という表現しかありえないであろう。また単一時点のジニ係数を見るにしても，その統計学としての精度が問題となろう。

推定されたジニ係数の信頼性あるいは精度に関する研究はほとんど見当たらない。この方面の最近の分析としては小林[7]によるものがある。彼はブートストラップ法を用いることによって，ジニ係数の精度は，標本数が7000と9000のとき，その標準偏差は0.003から0.005程度と推定し，ジニ係数の値は小数点以下2桁までが信頼できるとしている。

また過去の実証分析例を見ると，日本の場合，ジニ係数は用いられるデータによって異なり，0.2から0.4の間に位置している。そして，1つの分析例をとったとき，その時系列の変化はあまり大きくない。すなわち，所得分配構造の変化についての結論を下すには，より慎重になるべきことが予想される。

本稿では，所得分布にパレート分布または対数正規分布を仮定したとき，前述の$P(X_t < X_{t+1}) = p$に準ずる尺度を，尤度の考え方をもとに「変化の信頼係数」として導入する。この新しい尺度によって，所得分配構造の変化を判断することができる。

2 定義

ここでは，後で用いるジニ係数に関する各種の定義について述べる。

所得分布における所得格差の大きさを表わすには，ランダムに取り出した2つの標本の差がどの位であるかを見るのが直感的にわかりやすい。そこでジニ平均差を定義する。

定義1 ジニ平均差（Gini's mean difference） ジニ平均差 $\delta(\mathrm{X})$ は，同一分布に従う2つの独立な確率変数X_1とX_2の差の絶対値の期待値$E[|X_1-X_2|]$で定義される。

$$\delta(X)=E[|X_1-X_2|] \tag{1}$$

n個の確率変数$X_1, X_2, \ldots, X_n$がi.i.d.ならば，統計量

$$d(X)=\binom{n}{2}^{-1}\sum_{i<j}^{n}\sum^{n}|X_i-X_j| \tag{2}$$

は$\delta(X)$の不偏推定量である。

ジニ平均差は，Xの絶対的な大きさに依存するので，格差を示す指標としてジニ係数を定義する。

定義2 ジニ係数（Gini index） ジニ係数は，定義1のジニ平均差を用いて次のように定義される。

$$\gamma(X)=\frac{\frac{1}{2}\delta(X)}{E[X]} \tag{3}$$

3 パレート分布におけるジニ係数の変化の信頼係数

所得分布がパレート（pareto）分布に従うことについては，Champernown[2]，Hayakawa[4]等でふれられている。本節では，所得分布をパレート分布であると仮定した場合におけるジニ係数の信頼性について考察する。

3.1 パレート分布からのジニ係数の導出

パレート分布にはいくつかの類型（Johnson, Kotz and Balakrishnan[5]）があるが，ここでは，最も基本的なP(I)(a,k)について考えることにする。すなわち，分布関数は，

$$F(x)=1-(\frac{k}{x})^a \quad k>0, a>0, x\geq k \tag{4}$$

密度関数は，

$$f(x)=ak^a x^{-(a+1)} \quad k>0, a>0, x\geq k \tag{5}$$

で与えられる。

その期待値は，

$$E[X]=\frac{ak}{a-1} \qquad a>1 \tag{6}$$

となる。

次に，ジニ係数を求めよう。

ジニ平均差は，定義1より，

$$\begin{aligned}\delta(X) &= E[|X_1 - X_2|] \\ &= \frac{2ak}{(a-1)(2a-1)}\end{aligned} \tag{7}$$

したがって，ジニ係数は，

$$\begin{aligned}\gamma(X) &= \frac{\frac{1}{2}\delta(X)}{E[X]} \\ &= \frac{1}{2a-1}\end{aligned} \tag{8}$$

となる。これがパレート分布P(I)(a,k)のパラメータaとジニ係数を関係づける基本式である。位置パラメータkは捨象される。

ジニ係数の最尤推定量を求めるために，まず，パレート分布のパラメータの最尤推定量を求めよう。

パラメータkが既知の場合と未知の場合では，どちらも同じように考えることができるので，ここでは，パラメータkが既知の場合について述べることにする。

kが既知の場合のaの最尤推定量は，

$$\hat{a} = n[\sum_{j=1}^{n} \log(\frac{x_j}{k})]^{-1}$$

で与えられる。

$\hat{a}$の分布について考えてみよう。

定理1　$\frac{2na}{\hat{a}}$は，自由度$2n$のχ^2分布に従う。（証明は付録を参照。）

したがって，$\frac{2na}{\hat{a}}$は自由度$2n$のχ^2分布に従う。これより，$\hat{a}$の期待値は，

$$\begin{aligned}E[\hat{a}] &= \int_0^\infty \frac{2na}{z} f_{\chi^2_{2n}(z)} dz \\ &= \frac{na}{n-1} \qquad n > 1\end{aligned} \tag{9}$$

となる。

また，分散は，

$$\begin{aligned}E[\hat{a}^2] &= \int_0^\infty (\frac{2na}{z})^2 f_{\chi^2_{2n}(z)} dz \\ &= \frac{n^2a^2}{(n-1)(n-2)}\end{aligned} \tag{10}$$

より，

$$\begin{aligned}V[\hat{a}] &= E[\hat{a}^2] - E[\hat{a}]^2 \\ &= \frac{n^2a^2}{(n-1)^2(n-2)} \qquad n > 2\end{aligned} \tag{11}$$

となる。

次に，ジニ係数の最尤推定量を求めよう。

補題1　βをθの関数($\beta = f(\theta)$)とする。$\hat{\theta}$をθの最尤推定量とすると，関数βが単調関数な

らば，$\hat{\beta} = f(\hat{\theta})$も最尤推定量である。(証明は付録を参照。)

$\hat{a}$をaの最尤推定量とすると，ジニ係数Gはパラメータaに関して単調関数なので，この補題1により，次の定理が得られる。

定理 2　パレート分布のジニ係数の最尤推定量は，

$$\hat{G} = \frac{1}{2\hat{a}-1} \tag{12}$$

である。

系　ジニ係数の最尤推定量$\hat{G}$の密度関数$f(\hat{G})$は次式で与えられる。 (証明は付録を参照。)

$$f(\hat{G}) = \frac{(2na)^n}{\Gamma(n)}\left(\frac{\hat{G}}{1+\hat{G}}\right)^{n-1}\frac{1}{(1+\hat{G})^2}\exp\left(-2na\frac{\hat{G}}{1+\hat{G}}\right) \qquad 0 < \hat{G} < 1 \tag{13}$$

ジニ係数の最尤推定量$\hat{G}$はパレート分布のパラメータaの最尤推定量$\hat{a}$について非線型であり，また不偏性は一般に保証されない。

定理2の系より，$\hat{G}$の期待値を級数展開の形で求めることはできるが，実用的ではない。ただし，$\hat{a}$はaの十分統計量であり，$\hat{G}$は$\hat{a}$の関数であるので，$\hat{G}$もGの十分統計量である。

次に，ジニ係数の分散を求めよう。

一般に，確率変数Xの関数$g(X)$の分散は，

$$var\{g(X)\} \doteqdot \left(\frac{dg}{dX}\right)^2_\theta varX \tag{14}$$

と一次近似できる。

この結果を用いると，ジニ係数(G)の分散は次の通りである。

$$var(\hat{G}) \doteqdot \frac{4n^2a^2}{(n-1)^2(n-2)(2a-1)^4} \tag{15}$$

a ジニ係数		n 100	1000	10000
1.75	標準偏差	0.057	0.018	0.006
0.40	変動係数	0.143	0.044	0.014
1.93	標準偏差	0.048	0.015	0.005
0.35	変動係数	0.138	0.043	0.013
2.17	標準偏差	0.040	0.012	0.004
0.30	変動係数	0.133	0.041	0.013
2.50	標準偏差	0.032	0.010	0.003
0.25	変動係数	0.128	0.040	0.013
3.00	標準偏差	0.024	0.008	0.002
0.20	変動係数	0.122	0.038	0.012
3.83	標準偏差	0.018	0.005	0.002
0.15	変動係数	0.117	0.036	0.012

表 1　パレート分布のジニ係数の標準偏差と変動係数

表1より，近似値ではあるが，標本数nが小さいほど，また，ジニ係数が大きくなるほど，あるいはパラメータaが小さくなるほど標準偏差が大きくなっていることがわかる。また，変動係数を調べてみると，ジニ係数が大きくなるほどその値は大きくなっていく。このことは，ジニ係数が大きい領域にある場合，その変化の判断についてはより慎重にならなければいけないことを示している。

3.2 パレート分布における変化の信頼係数

いま，パラメータの異なる2つの分布が得られた標本を考えよう。真のパラメータの大小関係がはっきり分かっているとき，標本から推定されたパラメータの推定値もまたその大小関係を保つことが期待される。しかし，その関係が保たれるかどうかは，2つのパラメータの開き（パラメータの差，あるいは比）に依存するだろう。その関係を調べよう。

パラメータの真の値が$a_1 > a_2$という関係であることがわかっているとき，最尤推定値$\hat{a}_1$と$\hat{a}_2$が得られ，それらが$\hat{a}_1 > \hat{a}_2$という関係になる確率を求めよう。

$\frac{2n_1a_1}{\hat{a}_1}$と$\frac{2n_2a_2}{\hat{a}_2}$はそれぞれ独立に自由度$2n_1, 2n_2$の$\chi^2$分布に従うことから，

$$
\begin{aligned}
f(\hat{a}_1|a_1) &= \frac{1}{2^{n_1}\Gamma(n_1)}\left(\frac{2n_1a_1}{\hat{a}_1}\right)^{n_1-1}\exp\left\{-\frac{1}{2}\left(\frac{2n_1a_1}{\hat{a}_1}\right)\right\}\frac{2n_1a_1}{\hat{a}_1^2} \\
&= c\left(\frac{1}{\hat{a}_1}\right)^{n_1+1}\exp\left\{-\frac{1}{2}\left(\frac{2n_1a_1}{\hat{a}_1}\right)\right\} \\
&= c\left(\frac{w_1}{2n_1a_1}\right)^{n_1+1}\exp\left\{-\frac{1}{2}w_1\right\} \\
&\qquad \text{ここで}, w_1 = \frac{2n_1a_1}{\hat{a}_1} \\
&\propto w_1^{n_1+1}\exp\left\{-\frac{1}{2}w_1\right\}
\end{aligned}
$$

となり，これは自由度$2(n_1+2)$のχ^2分布に従う。

同様に，$w_2 = \frac{2n_2a_2}{\hat{a}_2}$とおくと，$f(\hat{a}_2|a_2)$は自由度$2(n_2+2)$の$\chi^2$分布に従うことがわかる。

$\hat{a}_1 > \hat{a}_2$は，$\frac{w_1}{w_2} < \frac{a_1}{a_2}\frac{n_1}{n_2}$と同値であるので，不等式の左辺$(\frac{w_1}{w_2})$の分子・分母をそれぞれの自由度で割ると，

$$
\frac{w_1/2(n_1+2)}{w_2/2(n_2+2)} < \frac{a_1}{a_2}\frac{n_1(n_2+2)}{n_2(n_1+2)}
$$

は自由度（$2(n_1+2), 2(n_2+2)$）のF分布に従う。

したがって，

$$
\begin{aligned}
P(\hat{a}_1 > \hat{a}_2|a_1,a_2) &= \int_0^{\frac{a_1}{a_2}\frac{n_1(n_2+2)}{n_2(n_1+2)}} \frac{\{2(n_1+2)\}^{(n_1+2)}\{2(n_2+2)\}^{(n_2+2)}}{B(n_1+2,n_2+2)} \\
&\qquad \times\frac{x^{n_1+1}}{\{2(n_2+2)+2(n_1+2)x\}^{n_1+n_2+4}}dx
\end{aligned} \tag{16}
$$

となる。ここで，$B(u,v) = \int_0^1 x^{u-1}(1-x)^{v-1}dx$はベータ関数である。

以下の表は，式(16)より，2つの推定されたジニ係数$(\hat{G} = \frac{1}{2\hat{a}-1})$の大小関係が，真のジニ係数の大小関係を保つ確率$P(\hat{G}_1 < \hat{G}_2 | G_1, G_2)$を示す。また，標本数の差を見るため，標本数が10000の場合（表2）と100の場合（表3）の2通り行なった。

たとえば，表2において，0.25の行と0.001の列の交叉しているところの値0.5893は$G_1 = 0.250$と

$G_2=0.251$のときの$P(\hat{G}_1<\hat{G}_2|G_1, G_2)$ の値が約59%であることを示している。すなわち，表側はG_1の値を，表頭はG_2の増分（G_2-G_1）を表している。

表2，表3を見てわかるとおり，ジニ係数の差が大きくなるほど，真の値の関係と推定値の関係が高い確率で保たれていることがわかる。また，ジニ係数が大きくなるほど，差が等しくても真の値の関係と推定値の関係が正しく示される確率は小さくなっている。標本数については，たとえば，ジニ係数の差が0.010の場合，真の値の関係と推定値の関係が正しく示される確率はn_1, n_2=10000では90%以上であるが，n_1, n_2=100では55%以上でしかない。

ジニ	増分										
係数	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009	0.010
0.20	0.5000	0.6155	0.7210	0.8096	0.8780	0.9268	0.9589	0.9784	0.9894	0.9951	0.9979
0.21	0.5000	0.6093	0.7100	0.7962	0.8645	0.9150	0.9497	0.9720	0.9853	0.9927	0.9966
0.22	0.5000	0.6036	0.6999	0.7835	0.8515	0.9032	0.9401	0.9649	0.9805	0.9897	0.9949
0.23	0.5000	0.5984	0.6905	0.7716	0.8388	0.8914	0.9302	0.9572	0.9750	0.9861	0.9926
0.24	0.5000	0.5937	0.6819	0.7604	0.8267	0.8797	0.9200	0.9490	0.9688	0.9818	0.9898
0.25	0.5000	0.5893	0.6738	0.7499	0.8150	0.8682	0.9096	0.9404	0.9622	0.9769	0.9864
0.26	0.5000	0.5853	0.6663	0.7399	0.8039	0.8570	0.8993	0.9315	0.9550	0.9715	0.9826
0.27	0.5000	0.5815	0.6593	0.7306	0.7932	0.8461	0.8890	0.9224	0.9475	0.9656	0.9782
0.28	0.5000	0.5780	0.6528	0.7218	0.7830	0.8355	0.8787	0.9132	0.9396	0.9592	0.9733
0.29	0.5000	0.5748	0.6467	0.7135	0.7733	0.8252	0.8686	0.9039	0.9315	0.9525	0.9679
0.30	0.5000	0.5718	0.6410	0.7056	0.7641	0.8153	0.8588	0.8946	0.9232	0.9455	0.9622
0.31	0.5000	0.5690	0.6357	0.6982	0.7552	0.8057	0.8491	0.8854	0.9149	0.9382	0.9561
0.32	0.5000	0.5664	0.6306	0.6912	0.7468	0.7965	0.8396	0.8762	0.9064	0.9307	0.9497
0.33	0.5000	0.5639	0.6259	0.6846	0.7388	0.7876	0.8304	0.8672	0.8979	0.9230	0.9431
0.34	0.5000	0.5616	0.6215	0.6784	0.7312	0.7791	0.8215	0.8583	0.8895	0.9153	0.9363
0.35	0.5000	0.5594	0.6172	0.6724	0.7239	0.7709	0.8128	0.8496	0.8811	0.9075	0.9293
0.36	0.5000	0.5573	0.6133	0.6668	0.7169	0.7630	0.8044	0.8411	0.8728	0.8997	0.9221
0.37	0.5000	0.5554	0.6095	0.6614	0.7103	0.7554	0.7963	0.8327	0.8645	0.8919	0.9149
0.38	0.5000	0.5535	0.6060	0.6564	0.7040	0.7481	0.7885	0.8246	0.8565	0.8841	0.9076
0.39	0.5000	0.5518	0.6026	0.6515	0.6979	0.7412	0.7809	0.8167	0.8485	0.8764	0.9003
0.40	0.5000	-	-	-	-	-	-	-	-	-	-

表2　$P\ (\hat{G}_1<\hat{G}_2\ /\ G_1,G_2)$（パレート分布，$n_1,n_2$=10000の場合）

ジニ	増分										
係数	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009	0.010
0.20	0.5000	0.5118	0.5236	0.5352	0.5468	0.5582	0.5696	0.5808	0.5919	0.6028	0.6136
0.21	0.5000	0.5112	0.5223	0.5333	0.5442	0.5550	0.5658	0.5764	0.5869	0.5973	0.6076
0.22	0.5000	0.5106	0.5211	0.5315	0.5419	0.5522	0.5623	0.5724	0.5824	0.5923	0.6021
0.23	0.5000	0.5100	0.5200	0.5299	0.5398	0.5495	0.5592	0.5688	0.5783	0.5878	0.5971
0.24	0.5000	0.5095	0.5190	0.5284	0.5378	0.5471	0.5563	0.5655	0.5746	0.5836	0.5925
0.25	0.5000	0.5091	0.5181	0.5271	0.5360	0.5449	0.5537	0.5624	0.5711	0.5797	0.5882
0.26	0.5000	0.5087	0.5173	0.5259	0.5344	0.5428	0.5513	0.5596	0.5679	0.5761	0.5843
0.27	0.5000	0.5083	0.5165	0.5247	0.5329	0.5410	0.5490	0.5570	0.5649	0.5728	0.5806
0.28	0.5000	0.5079	0.5158	0.5237	0.5315	0.5392	0.5469	0.5546	0.5622	0.5697	0.5772
0.29	0.5000	0.5076	0.5151	0.5227	0.5301	0.5376	0.5450	0.5523	0.5596	0.5669	0.5741
0.30	0.5000	0.5073	0.5145	0.5217	0.5289	0.5361	0.5432	0.5502	0.5572	0.5642	0.5711
0.31	0.5000	0.5070	0.5140	0.5209	0.5278	0.5346	0.5415	0.5483	0.5550	0.5617	0.5684
0.32	0.5000	0.5067	0.5134	0.5201	0.5267	0.5333	0.5399	0.5464	0.5529	0.5594	0.5658
0.33	0.5000	0.5065	0.5129	0.5193	0.5257	0.5321	0.5384	0.5447	0.5510	0.5572	0.5634
0.34	0.5000	0.5062	0.5124	0.5186	0.5248	0.5309	0.5370	0.5431	0.5491	0.5551	0.5611
0.35	0.5000	0.5060	0.5120	0.5180	0.5239	0.5298	0.5357	0.5416	0.5474	0.5532	0.5589
0.36	0.5000	0.5058	0.5116	0.5173	0.5231	0.5288	0.5345	0.5401	0.5458	0.5514	0.5569
0.37	0.5000	0.5056	0.5112	0.5167	0.5223	0.5278	0.5333	0.5388	0.5442	0.5496	0.5550
0.38	0.5000	0.5054	0.5108	0.5162	0.5215	0.5269	0.5322	0.5375	0.5428	0.5480	0.5532
0.39	0.5000	0.5052	0.5105	0.5157	0.5208	0.5260	0.5312	0.5363	0.5414	0.5464	0.5515
0.40	0.5000	-	-	-	-	-	-	-	-	-	-

表3　$P\ (\hat{G}_1<\hat{G}_2\ /\ G_1,G_2)$（パレート分布，$n_1,n_2$=100の場合）

一方，現実の状況を考えると，真のパラメータはわからず，得られるのは標本から推定されたパラメータの推定値のみである。そして，標本から推定されたパラメータの推定値の大小関係が真のパラメータの大小関係にも保たれていることが期待される。前と同様に，その関係が保たれるかどうかは，推定された2つのパラメータの開き（パラメータの差，あるいは比）に依存するだろう。したがって，標本から推定されたパラメータの推定値がどのくらいの精度をもっているかは重要なことである。そこで，その関係を調べてみよう。

ところで，$P(\hat{G}_1<\hat{G}_2|G_1,G_2)$ を別の角度から考えてみれば，$\hat{G}_1,\hat{G}_2$が得られたとき，$G_1<G_2$となっている可能性の大きさと解釈することもできる。そこでまず，$\hat{G}_1,\hat{G}_2$に対応する$\hat{a}_1,\hat{a}_2$に戻って考える。標本から2つのパレート分布のパラメータの最尤推定値$\hat{a}_1$と$\hat{a}_2$が推定され，それらが$\hat{a}_1>\hat{a}_2$という関係であるとき，真の値も$a_1>a_2$という関係である可能性の大きさを求めよう。この考え方は，データが得られたときのパラメータに関する一種の尤度に該当する。しかも，パラメータがある領域におちる尤度である。

$\frac{2n_1a_1}{\hat{a}_1}$ と $\frac{2n_2a_2}{\hat{a}_2}$ はそれぞれ独立に自由度$2n_1,2n_2$のχ^2分布に従うことから，

$$
\begin{aligned}
g(a_1|\hat{a}_1) &= \frac{1}{2^{n_1}\Gamma(n_1)}(\frac{2n_1a_1}{\hat{a}_1})^{n_1-1}\exp\{-\frac{1}{2}(\frac{2n_1a_1}{\hat{a}_1})\}\frac{2n_1}{\hat{a}_1} \\
&= ca_1^{n_1-1}\exp\{-\frac{1}{2}(\frac{2n_1a_1}{\hat{a}_1})\} \\
&= c(\frac{\hat{a}_1}{2n_1}w_1)^{n_1-1}\exp\{-\frac{1}{2}w_1\} \\
&\qquad \text{ここで}, w_1=\frac{2n_1a_1}{\hat{a}_1} \\
&\propto w_1^{n_1-1}\exp\{-\frac{1}{2}w_1\}
\end{aligned}
$$

となり，これは自由度$2n_1$のχ^2分布に従う。

同様に，$w_2=\frac{2n_2a_2}{\hat{a}_2}$ とおくと，$g(a_2|\hat{a}_2)$ は自由度$2n_2$のχ^2分布に従うことがわかる。

$a_1>a_2$は，$\frac{w_1}{w_2}>\frac{\hat{a}_2}{\hat{a}_1}\frac{n_1}{n_2}$ と同値であるので，不等式の左辺$(\frac{w_1}{w_2})$の分子・分母をそれぞれの自由度で割ると，

$$\frac{w_1/2n_1}{w_2/2n_2}>\frac{\hat{a}_2}{\hat{a}_1}$$

は自由度（$2n_1,2n_2$）のF分布に従う。

したがって，$a_1>a_2$である可能性の大きさ（相対尤度）を$R(a_1>a_2|\hat{a}_1>\hat{a}_2)$ とすれば，

$$R(a_1>a_2|\hat{a}_1>\hat{a}_2)=1-\int_0^{\frac{\hat{a}_2}{\hat{a}_1}}\frac{(2n_1)^{n_1}(2n_2)^{n_2}}{B(n_1,n_2)}\frac{x^{n_1-1}}{(2n_2+2n_1x)^{n_1+n_2}}dx \tag{17}$$

となる。ここで，$B(u,v)=\int_0^1 x^{u-1}(1-x)^{v-1}dx$はベータ関数である。

パラメータaとジニ係数Gの変換式[1]により，$R(a_1>a_2|\hat{a}_1>\hat{a}_2)$を$R(G_1<G_2|\hat{G}_1<\hat{G}_2)$と表現し，

$$R(G_1<G_2|\hat{G}_1<\hat{G}_2)=1-\int_0^{\frac{1+\frac{1}{\hat{G}_2}}{1+\frac{1}{\hat{G}_1}}}\frac{(2n_1)^{n_1}(2n_2)^{n_2}}{B(n_1,n_2)}\frac{x^{n_1-1}}{(2n_2+2n_1x)^{n_1+n_2}}dx \tag{18}$$

となる。

これを確率と考えることはもはやできないが，最大値が1となる一種の尤度でもあるこの値を「変化の信頼係数」と呼ぶことにする。この変化の信頼係数についてはベイズ的な解釈も可能で

[1] $\hat{G}=\frac{1}{2\hat{a}-1}$

あるがここでは触れない。

この変化の信頼係数は，表2とは自由度で若干異なるが，標本数が多い場合はほとんど差がないため，読み替えることが可能である。こうして，再び，表2を見なおしてみる。表の0.25の行と0.001の列の交叉しているところの値0.5893は$\hat{G}_1$=0.250と$\hat{G}_2$=0.251のときの変化の信頼係数$R(G_1<G_2,\ \hat{G}_1<\hat{G}_2)$が約0.59であることを示している。すなわち，表側は$\hat{G}_1$の値を，表頭は$\hat{G}_2$の増分$(\hat{G}_2-\hat{G}_1)$を表していることになる。表3も同様に読み替えることができる。

この表を用いて不平等度に変化があったと結論づけるためには，この変化の信頼係数が少なくても0.90以上あることが望ましいだろう。そこで，推定されたジニ係数の値をもとにして，変化の信頼係数が0.90を越えるためにはジニ係数の差がどの位必要であるかを調べればよい。すなわち，所得分布がパレート分布に従うことが適合度検定によって判定されたら，この表を用いることによって，標本から推定されたパラメータの推定値がどのくらいの精度をもって真のパラメータを推定しているかを調べることができるのである。

4 対数正規分布におけるジニ係数の変化の信頼係数

3節では，ジニ係数の背後にある分布としてパレート分布を仮定した。ここでは，対数正規分布（lognormal）分布を仮定する。対数正規分布は，Gibrat [3] が1931年にこれで所得分布を近似できると主張したことに始まり，日本では，高橋他 [8]，[9] 等で分析されている。そこで，本節では，所得分布が対数正規分布であると仮定した場合について3節と同様に検討する。

4.1 対数正規分布からのジニ係数の導出

$Z=log(X-\theta)$が正規分布するようなθが存在するならば，Xの分布は対数正規分布である。

密度関数は,

$$f(x)=\frac{1}{(x-\theta)\sqrt{2\pi}\sigma}\exp[-\frac{1}{2}\{\frac{\log(x-\theta)-\mu}{\sigma}\}^2] \qquad x>\theta \tag{19}$$

で与えられる。パラメータθの値の変化は分布の位置にのみ影響を与え，分散や形には影響を与えない。そこで，ここでは，$\theta=0$（既知）としてtwo-parameter 対数正規分布（パラメータ：μ，σ），

$$U=\frac{\log X-\mu}{\sigma} \tag{20}$$

について考えることにする。

期待値は,

$$\begin{aligned} E[X] &= E[\exp(\mu+U\sigma)] \\ &= \exp(\mu+\frac{1}{2}\sigma^2) \end{aligned} \tag{21}$$

となる。

ジニ係数を求める準備として以下の定義，および定理を示す。

対数正規分布の分布関数を$\Lambda(\mu,\ \sigma^2)$と表す。すなわち，Xを対数正規分布$\Lambda(\mu,\ \sigma^2)$に従う確率変数とすると，$Z=\log X$は正規分布$N(\mu,\ \sigma^2)$に従っている。

分布関数$F(x)$をもつ非負の確率変数Xのk次のモーメント分布関数を$F_k(x)$を次のように定義する。

定義3　モーメント分布関数

$$F_k(x) = \frac{1}{E[X^k]} \int_0^x v^k dF(v) \tag{22}$$

モーメント分布関数 $F_k(x)$ は，確率変数 X のとる値の範囲が x で切断されているときの k 次のモーメントの，X の k 次のモーメントに対する比率である。

この定義を用いて，対数正規分布の1次のモーメント分布関数を次のように表すことにする。

$$\Lambda_1(x|\mu, \sigma^2) = \frac{1}{E[X]} \int_0^x v d\Lambda(v|\mu, \sigma^2) \tag{23}$$

定理3　対数正規分布の1次のモーメント分布関数 $\Lambda_1(x|\mu, \sigma^2)$ は対数正規分布 $\Lambda(x|\mu+\sigma^2, \sigma^2)$ である。（証明は付録を参照。）

定理4　X が Λ 変数に従い，a,b が定数のとき，

$$\int_0^\infty \Lambda(ax^b|\mu_1, \sigma_1^2) d\Lambda(x|\mu_2, \sigma_2^2) = \Lambda(a|\mu_1 - b\mu_2, \sigma_1^2 + b^2\sigma_2^2)$$

が成り立つ。

ジニ平均差は，

$$\begin{aligned} \delta(X) &= E[|X_1 - X_2|] \\ &= 2E[X]\{2\Phi(\frac{\sigma}{\sqrt{2}}) - 1\} \\ &\text{ここで, } \Phi(v) = \int_{-\infty}^v \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2) dx \end{aligned} \tag{24}$$

となる。

したがって，ジニ係数は，

$$\begin{aligned} \gamma(X) &= \frac{\frac{1}{2}\gamma(X)}{E[X]} \\ &= 2\Phi(\frac{\sigma}{\sqrt{2}}) - 1 \end{aligned} \tag{25}$$

となる。これが対数正規分布のパラメータ σ とジニ係数を関係づける基本式である。パラメータ μ は捨象される。

ジニ係数の最尤推定量を求めるために，まず，対数正規分布のパラメータの最尤推定量を求めることから始めよう。

θ が既知の場合のパラメータ μ と σ の最尤推定について考える。

$$Z_i = \log(X_i - \theta)$$

と変換すれば，正規分布の最尤推定量に帰着される。したがって，μ と σ の最尤推定量は，

$$\hat{\sigma} = s = [\frac{1}{n} \sum_{j=1}^n (z_j - \bar{z})^2]^{\frac{1}{2}} \tag{26}$$

となる。

$\hat{\sigma}$ の期待値と分散は，

$$E[\hat{\sigma}] = \sqrt{\frac{n-1}{n}}\sigma \tag{27}$$

$$V[\hat{\sigma}] = \frac{\mu_4 - (\sigma^2)^2}{4n\sigma^2} \tag{28}$$

ここで，$\mu_r = E[(z - E[z])^r]$

となる。

$\hat{\sigma}$ を σ の最尤推定量とする。ジニ係数Gはパラメータ σ に関して単調関数なので，補題1により，次の定理が得られる。

定理5 対数正規分布のジニ係数の最尤推定量は，

$$\hat{G} = 2\Phi(\frac{\hat{\sigma}}{\sqrt{2}}) - 1 \tag{29}$$

である。ただし $\hat{\sigma} = [\frac{1}{n}\sum_{j=1}^{n}(z_j - \bar{z})^2]^{\frac{1}{2}}$ である。

次に，ジニ係数の分散について考える。

確率変数の関数の分散についての近似式(14)を用いれば，ジニ係数 $\hat{G}$ の分散は，

$$\begin{aligned} var(\hat{G}) &\doteqdot \frac{(\mu_4 - (\sigma^2)^2)\phi^2(\frac{\sigma}{\sqrt{2}})}{8n\sigma^2} \\ &= \frac{\sigma^2}{4n}\phi^2(\frac{\sigma}{\sqrt{2}}) \end{aligned} \tag{30}$$

となる。ここで，$\mu_4 = E[(z - E[z])^4] = 3\sigma^4$，$\phi(x) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)$ である。

σ / ジニ係数		n		
		100	1000	10000
0.27	標準偏差	0.010	0.003	0.001
0.15	変動係数	0.051	0.016	0.005
0.36	標準偏差	0.014	0.004	0.001
0.20	変動係数	0.054	0.017	0.005
0.45	標準偏差	0.018	0.006	0.002
0.25	変動係数	0.056	0.018	0.006
0.54	標準偏差	0.022	0.007	0.002
0.30	変動係数	0.059	0.019	0.006
0.64	標準偏差	0.026	0.008	0.003
0.35	変動係数	0.062	0.020	0.006
0.74	標準偏差	0.031	0.010	0.003
0.40	変動係数	0.065	0.021	0.006

表4　対数正規分布のジニ係数の標準偏差と変動係数

表4を見ると，表1と同様に近似値ではあるが，標準偏差，変動係数ともにパレート分布の結果と同じ傾向が見られる。ただし，パレート分布の場合と比較すると，その値は小さくなっている。

4.2　対数正規分布における変化の信頼係数

ここでは，所得分布を対数正規分布$\Lambda(0, \sigma^2)$に仮定した場合におけるパラメータσの感度について考えよう。

σの最尤推定量は，

$$\hat{\sigma} = [\frac{1}{n}\sum_{j=1}^{n}(z_j - \bar{z})^2]^{\frac{1}{2}}$$
$$z_j = \log x_j$$

であり，$\frac{n\hat{\sigma}^2}{\sigma^2}$は自由度$n-1$の$\chi^2$分布に従うことがわかっている。

いま，パラメータの異なる2つの分布から得られた標本を考えよう。真のパラメータの大小関係がはっきり分かっているとき，標本から推定されたパラメータの推定値もまたその大小関係を保つことが期待される。しかし，その関係が保たれるかどうかは2つのパラメータの開き（パラメータの差，あるいは比）に依存するだろう。その関係を調べよう。

パラメータの真の値が$\sigma_1^2 < \sigma_2^2$という関係であることがわかっているとき，最尤推定値$\hat{\sigma}_1^2$と$\hat{\sigma}_2^2$が得られ，それらが$\hat{\sigma}_1^2 < \hat{\sigma}_2^2$という関係になる確率を求めよう。

$\frac{n_1\hat{\sigma}_1^2}{\sigma_1^2}$と$\frac{n_2\hat{\sigma}_2^2}{\sigma_2^2}$はそれぞれ独立に自由度$n_1-1, n_2-1$の$\chi^2$分布に従うことから，

$$\begin{aligned}
f(\hat{\sigma}_1^2|\sigma_1^2) &= \frac{1}{2^{\frac{n_1-1}{2}}\Gamma(\frac{n_1-1}{2})}(\frac{n_1\hat{\sigma}_1^2}{\sigma_1^2})^{\frac{n_1-1}{2}-1}\exp\{-\frac{1}{2}(\frac{n_1\hat{\sigma}_1^2}{\sigma_1^2})\}\frac{n_1}{\sigma_1^2} \\
&= c(\hat{\sigma}_1^2)^{\frac{n_1-1}{2}-1}\exp\{-\frac{1}{2}(\frac{n_1\hat{\sigma}_1^2}{\sigma_1^2})\} \\
&= c(\frac{\sigma_1^2}{n_1}w_1)^{\frac{n_1-1}{2}-1}\exp\{-\frac{1}{2}w_1\} \\
&\qquad \text{ここで}, w_1 = \frac{n_1\hat{\sigma}_1^2}{\sigma_1^2} \\
&\propto w_1^{\frac{n_1-1}{2}-1}\exp\{-\frac{1}{2}w_1\}
\end{aligned}$$

となり，これは自由度n_1-1のχ^2分布に従う。

同様に，$w_2 = \frac{n_2\hat{\sigma}_2^2}{\sigma_2^2}$とおくと，$f(\hat{\sigma}_2^2|\sigma_2^2)$は自由度$n_2-1$の$\chi^2$分布に従うことがわかる。

$\hat{\sigma}_1^2 < \hat{\sigma}_2^2$は，$\frac{w_1}{w_2} < \frac{\sigma_2^2}{\sigma_1^2}\frac{n_1}{n_2}$と同値であるので，不等式の左辺$(\frac{w_1}{w_2})$の分子・分母をそれぞれの自由度で割ると，

$$\frac{w_1/n_1-1}{w_2/n_2-1} < \frac{\sigma_2^2}{\sigma_1^2}\frac{n_1(n_2-1)}{n_2(n_1-1)}$$

は自由度(n_1-1, n_2-1)のF分布に従う。

したがって，

$$\begin{aligned}
P(\hat{\sigma}_1^2 < \hat{\sigma}_2^2|\sigma_1^2, \sigma_2^2) &= \int_0^{\frac{\sigma_2^2}{\sigma_1^2}\frac{n_1(n_2-1)}{n_2(n_1-1)}} \frac{(n_1-1)^{\frac{n_1-1}{2}}(n_2-1)^{\frac{n_2-1}{2}}}{B(\frac{n_1-1}{2}, \frac{n_2-1}{2})} \\
&\qquad \times \frac{x^{\frac{n_1-1}{2}-1}}{\{(n_2-1)+(n_1-1)x\}^{\frac{n_1+n_2-2}{2}}}dx
\end{aligned} \tag{31}$$

となる。ここで，$B(u,v) = \int_0^1 x^{u-1}(1-x)^{v-1}dx$はベータ関数である。

以下の表は，式(31)より，2つの推定されたジニ係数$(\hat{G} = 2\Phi(\frac{\hat{\sigma}}{\sqrt{2}})-1, \Phi(v) = \int_{-\infty}^{v}\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)dx)$の大小関係が，真のジニ係数の大小関係を保つ確率$P(\hat{G}_1 < \hat{G}_2|G_1, G_2)$を示す。また，標本数の差を見るため，標本数が10000の場合（表5）と100の場合（表6）の2通り行なった。

たとえば，表5において，0.25の行と0.001の列の交叉しているところの値0.6602は$G_1=0.250$と$G_2=0.251$のときの$P(\hat{G}_1 < \hat{G}_2|G_1, G_2)$の値が約66%であることを示している。

ジニ	増分										
係数	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009	0.010
0.20	0.5000	0.6949	0.8454	0.9360	0.9785	0.9942	0.9987	0.9998	1.0000	1.0000	1.0000
0.21	0.5000	0.6866	0.8340	0.9268	0.9734	0.9920	0.9980	0.9996	0.9999	1.0000	1.0000
0.22	0.5000	0.6791	0.8235	0.9178	0.9678	0.9895	0.9971	0.9994	0.9999	1.0000	1.0000
0.23	0.5000	0.6722	0.8135	0.9089	0.9620	0.9866	0.9960	0.9990	0.9998	1.0000	1.0000
0.24	0.5000	0.6660	0.8041	0.9001	0.9560	0.9833	0.9946	0.9985	0.9996	0.9999	1.0000
0.25	0.5000	0.6602	0.7952	0.8915	0.9498	0.9798	0.9930	0.9979	0.9994	0.9999	1.0000
0.26	0.5000	0.6547	0.7867	0.8831	0.9435	0.9760	0.9911	0.9971	0.9992	0.9998	1.0000
0.27	0.5000	0.6497	0.7788	0.8750	0.9371	0.9720	0.9890	0.9962	0.9988	0.9997	0.9999
0.28	0.5000	0.6451	0.7713	0.8671	0.9307	0.9678	0.9866	0.9951	0.9984	0.9995	0.9999
0.29	0.5000	0.6408	0.7643	0.8596	0.9244	0.9634	0.9841	0.9938	0.9978	0.9993	0.9998
0.30	0.5000	0.6367	0.7576	0.8522	0.9182	0.9589	0.9814	0.9924	0.9972	0.9991	0.9997
0.31	0.5000	0.6330	0.7513	0.8453	0.9120	0.9544	0.9785	0.9908	0.9964	0.9987	0.9996
0.32	0.5000	0.6295	0.7453	0.8385	0.9059	0.9498	0.9755	0.9891	0.9956	0.9984	0.9995
0.33	0.5000	0.6262	0.7397	0.8320	0.8999	0.9452	0.9724	0.9872	0.9946	0.9979	0.9993
0.34	0.5000	0.6229	0.7343	0.8258	0.8942	0.9405	0.9692	0.9853	0.9935	0.9974	0.9990
0.35	0.5000	0.6201	0.7293	0.8199	0.8885	0.9359	0.9659	0.9832	0.9924	0.9968	0.9988
0.36	0.5000	0.6173	0.7246	0.8143	0.8830	0.9314	0.9626	0.9810	0.9911	0.9961	0.9984
0.37	0.5000	0.6147	0.7200	0.8088	0.8776	0.9268	0.9592	0.9788	0.9897	0.9954	0.9981
0.38	0.5000	0.6124	0.7159	0.8036	0.8725	0.9224	0.9558	0.9765	0.9883	0.9946	0.9977
0.39	0.5000	0.6101	0.7117	0.7986	0.8675	0.9180	0.9524	0.9741	0.9868	0.9937	0.9972
0.40	0.5000	-	-	-	-	-	-	-	-	-	-

表5　P $(\hat{G}_1<\hat{G}_2 \mid G_1,G_2)$（対数正規分布，$n_1,n_2$=10000の場合）

ジニ	増分										
係数	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009	0.010
0.20	0.5000	0.5202	0.5402	0.5600	0.5796	0.5989	0.6179	0.6365	0.6547	0.6725	0.6898
0.21	0.5000	0.5192	0.5384	0.5573	0.5761	0.5945	0.6127	0.6306	0.6481	0.6653	0.6820
0.22	0.5000	0.5184	0.5367	0.5549	0.5728	0.5906	0.6081	0.6253	0.6421	0.6587	0.6748
0.23	0.5000	0.5177	0.5352	0.5527	0.5699	0.5870	0.6038	0.6204	0.6366	0.6526	0.6683
0.24	0.5000	0.5170	0.5339	0.5506	0.5672	0.5837	0.5999	0.6159	0.6316	0.6471	0.6622
0.25	0.5000	0.5163	0.5326	0.5488	0.5648	0.5806	0.5963	0.6117	0.6269	0.6419	0.6566
0.26	0.5000	0.5158	0.5314	0.5470	0.5625	0.5778	0.5929	0.6079	0.6226	0.6371	0.6514
0.27	0.5000	0.5152	0.5304	0.5454	0.5604	0.5752	0.5899	0.6044	0.6186	0.6327	0.6466
0.28	0.5000	0.5147	0.5294	0.5440	0.5584	0.5728	0.5870	0.6011	0.6149	0.6286	0.6421
0.29	0.5000	0.5143	0.5285	0.5426	0.5567	0.5706	0.5844	0.5980	0.6115	0.6248	0.6380
0.30	0.5000	0.5138	0.5276	0.5413	0.5550	0.5685	0.5819	0.5952	0.6083	0.6213	0.6340
0.31	0.5000	0.5134	0.5268	0.5402	0.5534	0.5666	0.5796	0.5925	0.6053	0.6179	0.6304
0.32	0.5000	0.5131	0.5261	0.5391	0.5520	0.5648	0.5775	0.5901	0.6025	0.6148	0.6270
0.33	0.5000	0.5127	0.5254	0.5380	0.5506	0.5631	0.5755	0.5877	0.5999	0.6119	0.6238
0.34	0.5000	0.5124	0.5248	0.5371	0.5493	0.5615	0.5736	0.5856	0.5974	0.6092	0.6208
0.35	0.5000	0.5121	0.5242	0.5362	0.5481	0.5600	0.5718	0.5835	0.5951	0.6066	0.6180
0.36	0.5000	0.5118	0.5236	0.5353	0.5470	0.5586	0.5702	0.5816	0.5930	0.6042	0.6154
0.37	0.5000	0.5115	0.5231	0.5345	0.5460	0.5573	0.5686	0.5798	0.5909	0.6019	0.6128
0.38	0.5000	0.5113	0.5226	0.5338	0.5450	0.5561	0.5671	0.5781	0.5890	0.5998	0.6105
0.39	0.5000	0.5111	0.5221	0.5331	0.5440	0.5549	0.5658	0.5765	0.5872	0.5978	0.6083
0.40	0.5000	-	-	-	-	-	-	-	-	-	-

表6　P $(\hat{G}_1<\hat{G}_2 \mid G_1,G_2)$（対数正規分布，$n_1,n_2$=100の場合）

表5と表6を見てわかるとおり，所得分布にパレート分布を仮定した場合と同様に，ジニ係数の差が大きくなるほど，真の値の関係と推定値の関係が高い確率で保たれていることがわかる。また，ジニ係数が大きくなるほど，差が等しくても真の値の関係と推定値の関係が正しく示される確率は小さくなっている。標本数については，たとえば，ジニ係数の差が0.010の場合，真の値の関係と推定値の関係が正しく示される確率はn_1, n_2=10000ではかなり高く，99%以上であるが，n_1, n_2=100では60%以上でしかない。

対数正規分布の場合もパレート分布と同様，変化の信頼係数を定義することができる。

$$R(G_1 < G_2 | \hat{G}_1 < \hat{G}_2) = 1 - \int_0^{\{\frac{\Phi^{-1}(\frac{\hat{G}_1+1}{2})}{\Phi^{-1}(\frac{\hat{G}_2+1}{2})}\}^2 \cdot \frac{n_1(n_2-1)}{n_2(n_1-1)}} \frac{(n_1-1)^{\frac{n_1-1}{2}}(n_2-1)^{\frac{n_2-1}{2}}}{B(\frac{n_1-1}{2}, \frac{n_2-1}{2})} \times \frac{x^{\frac{n_1-1}{2}-1}}{\{(n_2-1)+(n_1-1)x\}^{\frac{n_1+n_2-2}{2}}} dx \tag{32}$$

ここで，$B(u,v) = \int_0^1 x^{u-1}(1-x)^{v-1}dx$はベータ関数，$\Phi(v) = \int_{-\infty}^{v} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)dx$ である。

パレート分布の場合と同様，この変化の信頼係数は，表5とは若干異なるが，標本数が多い場合はほとんど差がないため，読み替えることが可能である。そこで，表5を見なおしてみる。表の0.25の行と0.001の列の交叉しているところの値0.6602は$\hat{G}_1$=0.250と$\hat{G}_2$=0.251のときの変化の信頼係数$R(G_1 < G_2, \hat{G}_1 < \hat{G}_2)$ が約0.66であることを示している。すなわち，この場合は表側は$\hat{G}_1$の値を，表頭は$\hat{G}_2$の増分（$\hat{G}_2 - \hat{G}_1$）を表していることになる。表6も同様に読み替えることができる。そして，パレート分布の場合と手続きによって変化の判断をすることができる。

5 結論

2時点間の所得分布に変化があると判断するためには，何らかの指標を求め比較する必要がある。しかし，そこでは指標に対する信頼性が問題となろう。従来はジニ係数を記述統計量としてとらえていたため，その数値の単なる変化をもって所得分配の平等化，不平等化を判断していたが，そこには若干無理があったと思われる。

ジニ係数を，確率変数の実現値から計算された統計量と考えれば，2時点間の比較は，2つの確率変数の大小比較となり，確率的な表現をすることになる。ただし，背後にある所得分布をパラメトリックな分布とすれば，所得分配の変化はパラメータの変化であり，現実問題としては，2時点のデータが与えられたとき，パラメータの変化の信頼性を何らかの指標で表すことになる。

本稿では，尤度の考え方をもとに所得分布の変化の信頼性を示す「変化の信頼係数」を導入し，対象となる所得分布をパレート分布，および対数正規分布に限って，数表を作成した。この手続きは同様に他の所得分布についても行なうことができる。「変化の信頼係数」は新しい概念であるが，これにより所得分配構造の変化を判断する1つの基準が与えられたと考える。

付録

定理1　$\frac{2na}{\hat{a}}$ は，自由度$2n$のχ^2分布に従う。

（証明）

$$\begin{aligned} \hat{a} &= n[\sum_{j=1}^{n} \log(\frac{X_j}{k})]^{-1} \\ &= 2an[\sum_{j=1}^{n} 2a\log(\frac{X_j}{k})]^{-1} \end{aligned}$$

ここで，$2a\log(\frac{X}{k})$ の分布を考える。

$$
\begin{aligned}
F(z) &= P(Z \leq z) \\
&= P(2a \log \frac{X}{k} \leq z) \\
&= P(X \leq k \exp(\frac{z}{2a})) \\
&= \int_{k}^{k\exp(\frac{z}{2a})} ak^a x^{-(a+1)} dx \\
&= 1 - \exp(-\frac{z}{2}) \sim \chi_2^2
\end{aligned}
$$

したがって，$2a \log(\frac{X}{k}) \sim \chi_2^2$ と各X_jは独立であることから，

$$
\hat{a} = 2an\frac{1}{Z} \qquad where \quad Z \sim \chi_{2n}^2
$$

したがって，$\frac{2na}{\hat{a}}$ は，自由度$2n$のχ^2分布に従う。

（証明終）

補題1　βをθの関数（$\beta = f(\theta)$）とする。$\hat{\theta}$をθの最尤推定量とすると，関数βが単調関数ならば，$\hat{\beta} = f(\hat{\theta})$ も最尤推定量である。

（証明）

$\ell(x \mid \beta)$ と $L(x \mid \theta)$ をβとθのそれぞれの尤度関数とする。

$$
\begin{aligned}
\ell(x|\beta) &= L(x|\theta) \\
&= L(x|f^{-1}(\beta)) \quad where \quad \beta = f(\theta)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \ell}{\partial \beta} &= \frac{\partial}{\partial \beta} L(x \mid f^{-1}(\beta)) \\
&= \frac{\partial}{\partial \theta} L(x \mid \theta) \times \frac{\partial}{\partial \beta} f^{-1}(\beta) \\
&= 0 \qquad (\frac{\partial L}{\partial \theta} = 0 \text{ より})
\end{aligned}
$$

f^{-1}の存在のためには単調性が必要である。したがって，最尤推定量を代入した関数が単調ならばその関数も最尤推定量である。

（証明終）

系　ジニ係数の最尤推定量の密度関数$f(\hat{G})$ は次式で与えられる。

$$
f(\hat{G}) = \frac{z^{n-1}e^{-\frac{z}{2}}}{2^n\Gamma(n)} \cdot \frac{4na}{(1+\hat{G})^2} \frac{(2na)^n}{\Gamma(n)} (\frac{\hat{G}}{1+\hat{G}})^{n-1} \frac{1}{(1+\hat{G})^2} \exp(-2na\frac{\hat{G}}{1+\hat{G}}) \qquad 0 < \hat{G} < 1
$$

（証明）

$\frac{2na}{\hat{a}} \sim \chi_{2n}^2$ と定理2の $\hat{G} = \frac{1}{2\hat{a}-1}$ より，$z = 2na\frac{2\hat{G}}{1+\hat{G}}$ とおくと，ジニ係数の最尤推定量$\hat{G}$の密度関数は，

$$
f(\hat{G}) = \frac{(2na)^n}{\Gamma(n)} (\frac{\hat{G}}{1+\hat{G}})^{n-1} \frac{1}{(1+\hat{G})^2} \exp(-2na\frac{\hat{G}}{1+\hat{G}})
$$

となる。

（証明終）

定理3 対数正規分布の1次のモーメント分布関数$\Lambda_1(x\mid\mu,\sigma^2)$は対数正規分布$\Lambda(x\mid\mu+\sigma^2,\sigma^2)$である。

（証明）

$$
\begin{aligned}
\Lambda_1(x|\mu,\sigma^2) &= \frac{1}{E[X]}\int_0^x v d\Lambda(x|\mu,\sigma^2) \\
&= \exp(-\mu-\frac{1}{2}\sigma^2)\int_0^x \exp(\log v)\frac{1}{v\sigma\sqrt{2\pi}}\exp\{-\frac{1}{2\sigma^2}(\log v-\mu)^2\}dv \\
&= \int_0^x \frac{1}{v\sigma\sqrt{2\pi}}\exp\{-\frac{1}{2\sigma^2}(\log v-\mu-\sigma^2)^2\}dv \\
&= \Lambda(x|\mu+\sigma^2,\sigma^2)
\end{aligned}
$$

（証明終）

参考文献

[1] Aitchison,J. and Brown,J.A.C. （1957), *The Lognormal Distribution,* Cambridge Univ. Press

[2] Champernown,D.G. (1952), "The graduation of income distribution", *Econometrica,* 20, 591-615

[3] Gibrat,R. （1931), *Les inégalités économiques,* Paris:Sirey

[4] Miyoji Hayakawa （1951), "The Application of Pareto's Law of Income to Japanese Data", *Econometrica,* 19, 2, 174-183

[5] Johnson,N.L., Kotz,S. and Balakrishnan,N. （1994）, *Continuous Univariate Distribution,* vol.1, second edition, John Wiley & Sons, Inc.

[6] Pareto,V. (1896-1897), *Cours d'économie politique,* 2 vols., Lausanne: Rouge

[7] 小林正人（1998),「所得不平等の計測と精度評価」,『所得再分配の評価手法に関する研究』, 平成9年度厚生科学研究費補助金・厚生行政科学研究事業

[8] 高橋長太郎,藤野正三郎,伊大知良太郎,宮川公男,江見康一（1958),「戦後所得分布の変遷」,『経済研究』, 9, 1, 36-51

[9] 高橋長太郎,伊大知良太郎（1954),「日本の所得分布」,『経済研究』, 5, 2, 120-135