

ジニ係数とカイ離係数

01600120 牧野都治 MAKINO Toji

1. カイ離係数

集団内のデータのもつ格差を表わす尺度として、よく用いられるものに、ジニ係数があるが、われわれは、類似の尺度として、カイ離係数の使用をすすめたい。それは次のような意味をもつ。

いま T を、正の値をとる連続的確率変数として、その密度関数を $f(t)$ 、期待値を ν とする。このとき、 T の分布のパレート曲線 $y = g(x)$ は、 t を媒介変数として、次式で表わすことができる。

$$x = \int_t^{\infty} f(t) dt, \quad y = \int_t^{\infty} t f(t) dt / \nu.$$

この曲線 $y = g(x)$ 上にあつて、均等線とよばれる直線 $y = x$ とのカイ離が最大になる点、つまりふくらみ最大の点 $P(x_0, y_0)$ のことをカイ離点とよび、 $d = y_0 - x_0$ をカイ離係数という。

表1は千葉県柏税務署管内での高額所得納税(1000万円以上)についてのデータから計算した例である。この表には納税金額を T (万円)として、

$\{T \geq 1000\}$ 、 $\{T \geq 2000\}$ 、 $\{T \geq 3000\}$ の人数およびその分布のカイ離係数、ならびに
 $\{5000 > T \geq 1000\}$ 、 $\{3000 > T \geq 1000\}$ の人数およびその分布のカイ離係数

が示されている。

表1. 柏税務署における納税者数とカイ離係数

年次	項目	1000万円以上	2000万円以上	3000万円以上	5000万円未満	3000万円未満
平成 3 年	人数	902	356	194	821	708
	カイ離係数	0.308	0.271	0.254	0.186	0.169
平成 4 年	人数	638	210	93	605	545
	カイ離係数	0.267	0.242	0.233	0.183	0.134
平成 12 年	人数	417	106	41	404	376
	カイ離係数	0.228	0.220	0.213	0.169	0.132

バブル崩壊前(平成3年分)と崩壊後(平成4年分)の1000万円以上について比べてみると、崩壊後は記載者数が激減し、カイ離係数も小さくなっている。しかし、平成12年分は、それよりもさらに落ち込んでいる。経済活性化の道遠しとみるべきかもしれない。

2. ジニ係数とカイ離係数

次の説明図は、平成14年8月10日の読売新聞に掲載された記事である。

説明図. 読売新聞に載った解説「ジニ係数」

日本の世帯の所得格差が拡大している。総務省によると、所得分布のばらつきを表す「ジニ係数」という指標は、1999年に0.273となり、年々上昇する傾向が出ている。ジニ係数は、所得格差が小さいほど0に近づき、格差が大きいほど1に近づく。ただ、所得格差は年齢が上がるにつれて拡大する性質がある。ジニ係数の上昇は、日本社会の不平等の度合いが増していると言うより、「日本社会の高齢化が進んでいる影響」(総務省統計局)と見られる。

しかし、1999年にいきなり高齢化社会を迎えたわけではないので、この説明図にある、統計局の説明というのは納得できない。ただ、それはそれとして、ジニ係数が0.273になったということに注目してみたい。ジニ係数 M_c はパレート図での、不平等度を表わす弓形の面積 S の2倍の値になるので、 $M_c = 0.273$ は $S = 0.1365$ ということと同じである。次に、この値はカイ離係数でいうと、どの程度になるかを調べてみよう。いま、所得納税金額の分布 T が次のような密度関数 $f(t)$ をもつパレート分布に従うとしてみる。

$$f(t) = a t^{-(a+1)} \quad (t \geq 1).$$

このとき、 T の分布のパレート図での弓形の面積は

$$S = 1 / \{2(2a - 1)\}$$

となり、パレート曲線のカイ離点の座標は

$$x_0 = (1 - 1/a)^a, \quad y_0 = (1 - 1/a)^{a-1}$$

となるので、それを用いてカイ離係数 $d = y_0 - x_0$ を算出することができる。

いまの場合、 $S = 0.1365$ なので、 $a = 2.3315$ となり

$$d = 0.2068$$

になる。この値は、感覚的には表1の平成12年(1000万円以上)での格差よりも、いくぶん低めとみることができよう。

(注. 上では T がパレート分布にしたがうとして d の値を求めたが、そうでなく、例えば対数正規分布にしたがうとしても、 d は似たような値になると思われる。)

なお、平成3年分～5年分の国税務署での高額所得納税金額のデータについて、ジニ係数とカイ離係数とを比べてみると表2のようになる。ただし表2のカイ離係数の推定値 δ は、上で述べたように、 T がパレート分布にしたがうとして計算して求めた値である。

表2. ジニ係数とカイ離係数の比較

係数		平成3年分	平成4年分	平成5年分
ジニ係数		0.417	0.356	0.391
カイ離	推定値 δ	0.316	0.268	0.296
係数	真の値 d	0.308	0.267	0.290

参考文献

- [1] 日本工業標準調査会「品質管理用語JISZ8101(パレート図)」, 日本規格協会(1999)
- [2] ——— 「オペレーションズ・リサーチ用語JISZ8121(ABC分析)」, 日本規格協会(1967)
- [3] ——— 「生産管理用語JISZ8141(ABC分析)」, 日本規格協会(2001)
- [4] 田口時夫「日本における資本の集中形態と構造分析の方法としてのローレンツ・カーブについて」, 統計数理研究所イ報12巻(1964)
- [5] ——— 「パレート分布とパレート曲線の分析」, 統計数理研究所イ報12巻(1964)
- [6] ——— 「集中多様体と集中解析のシステム(1) —ジニの統計方法論の幾何学的展開—」, 統計数理36巻1号(1988)
- [7] 豊田敬 「所得分布の不平等度 — 不平等度の比較と尺度 — 」, 国民経済134(1975)
- [8] 牧野都治「分布の特性の表現に対するパレート図の利用について」, 統計数理研究所イ報, 26巻(1979)
- [9] ——— 「ABC分析とその区分線の設定」, 統計数理研究所イ報, 30巻(1983)
- [10] ——— 「格差・パレート図・ABC分析」, 日本評論社(1984)
- [11] ——— 「パレート図とABC分析」, オペレーションズ・リサーチ(1987年6月号)
- [12] 日本オペレーションズ・リサーチ学会編「OR事典2000」, (2000)