1998年度日本オペレーションズ・リサーチ学会
春季研究発表会

各種の分布のパラメータとパレート曲線

―― 高額所得納税額を例として ――

01600120 東京理科大学 牧野都治 MAKINO Toji

1．まえがき

貯蓄保有額などには対数正規分布がよくあてはまるが，高額所得納税額(1000万円以上)には，対数正規分布よりもパレート分布の方がよいといわれている．しかし，1000万円以上という条件つき対数正規分布ならばどうかを調べてみたところ，多くの税務署のデータによく適合することがわかった．それでは，このような分布のパレート曲線はどんな性質を持っているか．とくに高額所得納税額の格差の分析を念頭におきながら調べて，その結果を［3］で報告した．さらに，たとえば1000万円以上，3000万円未満というような条件をつけて各税務署，各年度でのパレート曲線の性質を調べてみると，それらがあまり大きくは異ならないことがわかった．そこで，このことを考慮し，とくにパラメータとの関連に重点をおいて，格差の分析を行ってみた．

2．パレート曲線とカイ離点

非負の値をとる確率変数をT，その分布の密度関数をf(t)とする．具体的には高額所得納税額をTとして考え，これを連続的に扱うことにするが，そのパレート曲線y＝g(x)を次のように書くものとする。いま，値（納税金額）の大きい方から累積することにして，横軸（x軸）に累積人数率，縦軸（y軸）に累積金額率をとって，データから計算した点を結ぶ．従って，厳密には折れ線になるはずであるが，実際には曲線とみなしてもよいものが得られる。これをパレート曲線（またはパレート図）とよぶ。このような曲線を囲っている正方形の２つの対角線のうち，原点を通る方を均等線というが，ここでは単に対角線といえば，x＋y＝1 の方をさすものとする．また，パレート曲線上の点で，均等線からのカイ離が最大の点 $P(x_0, y_0)$ のことをカイ離点といい，$\gamma=y_0-x_0$をカイ離係数とよぶ．パレート曲線と均等線との間の弓形の部分を不平等度を表わす弓形というが，その面積を２倍するとジニ係数になる．ジニ係数は格差の尺度としてよく用いられているが，カイ離点は，より多くの情報をふくんでいて，使いようによってはたいへん役立つものであると考える．

3．両端を削除した各種の分布のパレート曲線

各種の分布とその条件つき分布のパレート曲線について考察する．

パレート曲線は，Tの分布のもつ尺度パラメータの値には依存しないが，条件つき分布についてはどうか．これを，指数分布，パレート分布，対数正規分布の場合について調べてみる．これらの分布を扱う理由の１つはカイ離点の位置の違いに注目したいからである．すなわち，指数分布，パレート分布，対数正規分布のカイ離点はそれぞれ，対角線の上方，下方，線上に位置する代表的な分布だからである．

3.1 指数分布の場合

確率変数Tが指数分布に従うとき，密度関数は次のように書ける．

$$f(t)=\lambda e^{-\lambda t} \quad (t \geqq 0)$$

この分布のパレート曲線はλには無関係でただ１つ定まる．ここで，条件つきの確率変数 $(T \mid \alpha \leqq T < \beta)$ の分布のパレート曲線y＝g(x)を求めてみると，媒介変数tを用いて次のように表わされる．

$$x=\frac{e^{-\lambda t}-e^{-\lambda\beta}}{e^{-\lambda\alpha}-e^{-\lambda\beta}}, \quad y=\frac{\left(t+\frac{1}{\lambda}\right)e^{-\lambda t}-\left(\beta+\frac{1}{\lambda}\right)e^{-\lambda\beta}}{\left(\alpha+\frac{1}{\lambda}\right)e^{-\lambda\alpha}-\left(\beta+\frac{1}{\lambda}\right)e^{-\lambda\beta}}$$

また，カイ離点の座標 (x_0, y_0) は上の (x, y) の式のtに $E(T \mid \alpha \leqq T < \beta)$ をいれた値で，いまの場合

$$E(T \mid \alpha \leqq T < \beta)=\frac{\left(\alpha+\frac{1}{\lambda}\right)e^{-\lambda\alpha}-\left(\beta+\frac{1}{\lambda}\right)e^{-\lambda\beta}}{e^{-\lambda\alpha}-e^{-\lambda\beta}}$$

である．つぎに，上部（右裾）の制限βをとり外して（T｜T$\geqq\alpha$）の分布を考えてみると，これはちょうどT＋αの分布になる．このようにαだけシフトした分布のパレート曲線はふくらみの小さいものになり，Tの分布のパレート曲線上の任意の点P_1（x，y_1）と，それに対応する点P_2（x，y_2）に関して，

$$\frac{y_2-x}{y_1-x}=\frac{\nu}{\nu+\alpha}$$

という関係がでてくる．ただし，ν＝E（T）である．ここでカイ離点に関して注意したいことがある．それは，Tの分布のカイ離点はP（e^{-1},2e^{-1}）なので，対角線の上方にあるが，（T｜T$\geqq\alpha$）の分布のカイ離点は，$\lambda\cdot\alpha$の値によってさまざまであるということである．すなわちこの場合のカイ離点は，対角線の上方に限らず，線上，下方のいずれにも位置し，λはもはや尺度パラメータでない．

3.2 パレート分布の場合

確率変数Tが，密度関数

$$f(t)=a\,t_0^{a}\,t^{-(a+1)} \qquad (t\geqq t_0)$$

をもつパレート分布にしたがうとき，t_0は尺度パラメータなので，パレート曲線はt_0に依存しない．また条件つき確率変数（T｜$\alpha\leqq$T＜β）の分布のパレート曲線もt_0には無関係で，次のようになる．

$$x=\frac{\beta^{-a}-t^{-a}}{\beta^{-a}-\alpha^{-a}}\quad,\qquad y=\frac{\beta^{-a+1}-t^{-a+1}}{\beta^{-a+1}-\alpha^{-a+1}}\ .$$

とくに，カイ離点P(x_0，y_0）に関しては，

① $\alpha=1$，$\beta=3$，$a=1.5$ のとき $x_0=0.393$，$y_0=0.527$ で $x_0+y_0=0.920$ ．

② $\alpha=1$，$\beta=3$，$a=2$ のとき $x_0=0.375$，$y_0=0.500$ で $x_0+y_0=0.825$ ．

（注．上の例では1000万円を単位として，1000万円以上3000万円未満の条件つき分布を考え，
$\alpha=1$，$\beta=3$ とした．以下の例でも同様の扱いをする．）

ここで上部の制限βを外した分布のパレート曲線を考えると，もとの（条件なしの）分布のそれと一致する．

3.3 対数正規分布の場合

確率変数Tが，密度関数

$$f(t)=\frac{1}{\sqrt{2\pi}\,\sigma\,t}\cdot\exp\left[-\frac{(\log t-\mu)^2}{2\sigma^2}\right] \qquad (t>0)$$

をもつ対数正規分布にしたがうとき，μは尺度パラメータなので，Tの分布のパレート曲線はμに無関係である．一方，条件つき対数正規分布においてはパレート曲線は次のようになり，μに依存している．したがってこの分布においては，μはもはや尺度パラメータではない．

$$x=\frac{\Phi\left(\frac{\log\beta-\mu}{\sigma}\right)-\Phi\left(\frac{\log t-\mu}{\sigma}\right)}{\Phi\left(\frac{\log\beta-\mu}{\sigma}\right)-\Phi\left(\frac{\log\alpha-\mu}{\sigma}\right)}\ ,\quad y=\frac{\Phi\left(\frac{\log\beta-\mu}{\sigma}-\sigma\right)-\Phi\left(\frac{\log t-\mu}{\sigma}-\sigma\right)}{\Phi\left(\frac{\log\beta-\mu}{\sigma}-\sigma\right)-\Phi\left(\frac{\log\alpha-\mu}{\sigma}-\sigma\right)}\ .$$

ただしΦ（x）は標準正規分布の分布関数である．

またとくに $\alpha=1$，$\beta=3$ で

① $\sigma=1.0$，$\mu=0.0$ のとき，$x_0=0.340$，$y_0=0.512$ で $x_0+y_0=0.852$ ．

② $\sigma=0.8$，$\mu=0.0$ のとき，$x_0=0.421$，$y_0=0.552$ で $x_0+y_0=0.973$ ．

この例からもわかるように，条件つき対数正規分布のパレート曲線は，対角線に関して対称でない．

参考文献

[1]国土開発出版社；長者番付，平成2年版〜平成9年版
[2]牧野都治；高額所得納税額に基づく格差の分析，オペレーションズ・リサーチ，1995年8月号
[3] ―― ；条件つき対数正規分布とそのパレート図，日本統計学会1997年研究発表会予稿集