

〔論文〕

参入下限値を単位としたベキ乗則生成モデル

鈴木 武

本稿では、あるシステムがどのような仕組みでベキ乗則を生成するかについて、優先的選択を仮定して考察している。これについては Simon (1955) 以来多くの研究がなされてきたが、ほとんどの場合、要素数 1 で新たなグループとしてシステムに参入することが前提とされてきた。その結果、グループの要素数が大きい場合でのみベキ乗則が成立するという結論を導いている。

本稿ではシステムへの参入下限値を設定し、その大きさを単位としてグループ要素数の確率分布を計算した。その結果、すべての範囲にわたってベキ乗則が成立することを述べている。また、従来の研究では初期条件の確率を表す式と、その後の確率の変化を表現する式に整合性がなかったが、その点も整合的になるように考察している。

【1】モデルの前提である優先的選択を述べ、各グループの時間経過による確率の変化を表す式を記述している。【2】モデルの解を段階を追って記述している。そのうち(a) t 時点における大きさ x の期待度数を求めている。(b)期待度数の関係式で用いられる比例定数を計算している。(c)システムに参入するグループ数を考察し、そこから t 時点における大きさ x の期待相対度数である確率の関係式を導いている。(d)定常状態における確率の関係式を求めている。(e)初期条件の確率を計算するさいに必要な定数を考察している。(f)確率分布を計算している。その結果、ユール分布を導いている。(g)参入下限値を単位にした確率分布を求めている。その結果、一般化ユール分布を導いている。さらにその極限としてパレート分布を求めている。

【3】先行研究の問題点を記述している。そのうち(a)Simon 論文と、(b)Kullmann & Kertesz 論文を取り上げている。(c)本稿の特徴を述べている。【4】優先的選択の代わりにランダム・モデルを仮定し、同様の手順で計算している。(a)モデルの解を計算

している。その結果、幾何分布を導いている。(b)参入下限値を単位にした確率分布を求めている。その結果、指数分布を導いている。【5】優先的選択とランダム・モデルをミックスした混合モデルを考察している。(a)モデルの解を計算している。その結果、一般的なパラメータをもつユール分布を導いている。(b)その分布の両極端なケースを計算し、それぞれ通常のユール分布と幾何分布を得ている。(c)参入下限値を単位にした確率分布を求めている。その極限として、混合分布を導いている。さらに両極端なケースとして、それぞれパレート分布と指数分布を求めている。

【1】モデルの前提

あるシステムがどのような仕組みでベキ乗則を生成するのであろうか。本稿ではそれを説明するモデルを考察する。モデルを構成するために、以下の前提をおく。

システムはいくつかのグループから構成されており、各グループの大きさは要素数で示される。グループの大きさ x は m 以上の自然数であり、 $x \geq m \geq 1$ である。初期状態で m 以上の大きさをもつグループが a 個存在し、大きさの合計を c とする。初期にシステムの中にグループが存在しないことも含め、 $a \geq 0$ とする。

毎期 1 要素ずつシステムに参入する。そのさい、既存のグループのどれかに所属するように参入する確率を $1 - a$ とする。したがって、既存のグループではなく新たに構成されるグループに参入する確率は a になる。ただし、新たなグループは m 以上の要素がなければならぬので、大きさが m になるまでシステムには参入しないで待機しているものとする。 $m = 1$ の場合には、新たなグループとして、すぐにシステムに参入する。システム

外で待機している新たなグループは1つとする。

日本の都市システムという事例を考えてみよう。現在、「市」というグループが数百存在している。市の大きさは人口という要素で表現される。市を構成するために必要最低限の人口が m に相当する。 m は数万人と考えられるので、かなり大きな数である。

都市への人口移動、あるいは、都市で生まれる人口が、都市システムへの参入になる。ただし、本モデルでは都市から出ていく人口や、都市における死亡者については考慮していない。每期1人が都市システムに参入することになるが、参入者数が多い場合には、各期の間隔を短くとればよいであろう。

既存の都市ではなく新たな都市を形成しそこに参入する場合には、人口が m 人に達するまで都市以外のところに待機しているものとしている。待機人口が m 人に達したとき都市システムに参入してくる。

特定のグループがどのような大きさをとるか、その確率分布を考えよう。 t 時点において、第 i グループが大きさ x をとる確率を $P_i(x, t)$ であらわす。 t 時点におけるグループの大きさは m から $t+c$ までの範囲内にある。したがって

$$\sum_{x=m}^{t+c} P_i(x, t) = 1$$

である。

ここで2つの仮定を設けよう。それは Simon (1955) が初めて提示した仮定である。

【仮定1】システムに参入する要素が既存のどのグループに所属するかは、各グループの大きさに誘引されるものとする。

【仮定2】システムに参入する要素が既存グループではなく、新たなグループに所属する確率は a である。ただし、 $0 \leq a \leq 1$ 。

仮定1は「優先的選択」(preferential attachment) と呼ばれているので、このモデルを「優先的選択モデル」と呼ぼう。

第 i グループにおける時間経過による確率の変化をみよう。 $(t-1)$ 時点から t 時点にかけて大き

さ x がとる確率の変化は、

$x > m$ の場合、仮定1から

$$\begin{aligned} P_i(x, t) - P_i(x, t-1) \\ = K(t-1)\{(x-1)P_i(x-1, t-1) \\ - xP_i(x, t-1)\} \end{aligned} \quad (1)$$

である。ただし、 $K(t-1)$ は時間に依存する比例定数とする。

$x = m$ の場合

$$\begin{aligned} P_i(m, t) - P_i(m, t-1) \\ = d_i(t) - K(t-1) m P_i(m, t-1) \end{aligned} \quad (2)$$

ここで $d_i(t)$ は第 i グループが t 時点でシステムに参入する確率である。

【2】モデルの解

時間が経過し t となったとき、システム全体として大きさ x に関する確率分布がどのようになるか、(1)式、(2)式をもとに算出しよう。

(a) 期待度数

t 時点において大きさ x を固定し、全グループの確率合計を求めると、 t 時点における大きさ x の期待度数 $f(x, t)$ になる。

t 時点におけるグループ数を $g(t)$ としよう。初期状態においてグループは a 個存在しているので、すべてのグループの大きさが1であるならば、グループ数は $t+a$ になる。したがって、グループ数は $a \leq g(t) \leq t+a$ の範囲にある。ただし、 $a=0$ の場合には $1 \leq g(t) \leq t$ の範囲とする。この小さな差はモデルの本質ではないので、ここでは $a+1 \leq g(t) \leq t+a$ としよう。それゆえ、 t 時点における大きさ x の期待度は

$$f(x, t) = \sum_{i=1}^{g(t)} P_i(x, t) \quad (3)$$

と表現される。また、期待度数を合計すると

$$\begin{aligned} \sum_{x=m}^{t+c} f(x, t) &= \sum_{i=1}^{g(t)} \sum_{x=m}^{t+c} P_i(x, t) \\ &= g(t) \end{aligned}$$

であり、グループ数に一致する。

(1)式を i で合計しよう。($t-1$)時点において t 時点のものは実現していない。したがって、その確率は 0 であることに注意すると

$$\sum_{i=1}^{g(t)} P_i(x, t-1) = \sum_{i=1}^{g(t-1)} P_i(x, t-1)$$

である。

(1)式を i で合計すると

$$\begin{aligned} \sum_{i=1}^{g(t)} P_i(x, t) - \sum_{i=1}^{g(t-1)} P_i(x, t-1) \\ = K(t-1) \sum_{i=1}^{g(t-1)} \{ (x-1) P_i(x-1, t-1) \\ - x P_i(x, t-1) \} \end{aligned}$$

よって

$$\begin{aligned} f(x, t) - f(x, t-1) \\ = K(t-1) \{ (x-1) f(x-1, t-1) \\ - x f(x, t-1) \} \end{aligned} \quad (4)$$

同様に(2)式についても

$$\begin{aligned} f(m, t) - f(m, t-1) \\ = \sum_{i=1}^{g(t)} d_i(t) - K(t-1) m f(m, t-1) \end{aligned} \quad (5)$$

である。

ここで $\sum_{i=1}^{g(t)} d_i(t)$ について考えてみよう。これ

は t 時点においてシステムに新たなグループが参入する確率を意味する。

新たなグループがシステムに参入するためには、要素数が m 個に達していなければならない。 m 個に達していない場合はシステム外で待機していると想定している。待機しているケースとしては、要素数が 0 個、1 個、 \dots 、 $(m-1)$ 個の m 通りの場合がある。

要素が新たなグループに所属する事象は每期ランダムに起こり、 m 個になるとクリアされ 0 個に戻るの、どのケースも同じ条件にある。したがって、各ケースの生じる確率は等しいと考えてよい。 t 時点において要素が新たなグループに所属する確率は a であり、 $(m-1)$ 個の要素がシステム外で待機している場合のみ、新たなグループが

システムに参入してくる。したがって、その確率は $\frac{a}{m}$ となる。それゆえ(5)式は

$$\begin{aligned} f(m, t) - f(m, t-1) \\ = \frac{a}{m} - K(t-1) m f(m, t-1) \end{aligned} \quad (6)$$

と書ける。

(b) 比例定数の計算

比例定数 $K(t-1)$ を求めよう。($t-1$)時点における全グループの要素数合計は $(t-1+c)$ である。ただし、新たなグループに所属するためシステムへの参入を待機している要素があるかもしれない。その数もここでは含めている。というのは、後述する定常状態における確率を考慮するさい、都合がよいからである。

グループの大きさは m から $(t-1+c)$ までの可能性がある。該当する大きさが無い場合には、その確率を 0 とする。

$K(t-1) x f(x, t-1)$ は第 t 要素がシステムに参入するとき、すでに x 回生起しているグループのどれかに属する確率である。すべての x についてその和をとると、第 t 要素が既存のどれかのグループに属する確率 $(1-a)$ になるので

$$\sum_{x=m}^{t-1+c} K(t-1) x f(x, t-1) = 1-a \quad (7)$$

また、すべての大きさを合計すると要素数全体の値になるから

$$\sum_{x=m}^{t-1+c} x f(x, t-1) = t-1+c \quad (8)$$

したがって、(7)式、(8)式から

$$K(t-1) = \frac{1-a}{t-1+c} \quad (9)$$

である。

(c) グループ数についての仮定

相対度数を求めるさい度数の合計が必要である。本モデルではグループ数と度数の合計が一致するので、グループ数の具体的な表現を考察する必要がある。そこで、グループ数は参入する要素数に比例するという仮定を設けよう。すなわち、 t 時点におけるグループ数 $g(t)$ は初期状態における a

個のグループ数と、要素数 t に比例する部分からなると仮定する。比例定数を k とすると

$$g(t) = kt + a, \quad (0 < k \leq 1) \quad (10)$$

と表される。

t 時点における大きさ x の期待相対度数、すなわち確率は

$$\begin{aligned} P(x, t) &= \frac{f(x, t)}{g(t)} \\ &= \frac{f(x, t)}{kt + a} \end{aligned} \quad (11)$$

と表現される。

もし要素数としてシステム外で待機している要素を含めないとしよう。 t 時点における待機要素数を $w(t)$ とすると、グループ数の期待値は(10)式ではなく

$$g(t) = k\{t - w(t)\} + a$$

となり、大きさ x の確率も

$$P(x, t) = \frac{f(x, t)}{k\{t - w(t)\} + a}$$

となる。

ここで t として、定常状態における大きさ x の確率 $P(x)$ を求めよう。 $w(t)$ は 0 から $(m-1)$ までたえず変化するので $P(x)$ の値も揺らぐことになる。とくに m の値が大きい場合には、 $m-1$ の状態から 0 の状態に変化するとき、 $P(x)$ の値のブレも大きい。それゆえ、システム外で待機している要素も含めて数えることにする。

(11)式を用いると、(4)式は

$$\begin{aligned} &\left(t + \frac{a}{k}\right)P(x, t) - \left(t - 1 + \frac{a}{k}\right)P(x, t-1) \\ &= (1-a) \cdot \frac{t-1+\frac{a}{k}}{t-1+c} \\ &\quad \cdot \{(x-1)P(x-1, t-1) - xP(x, t-1)\} \end{aligned} \quad (12)$$

(6)式は

$$\begin{aligned} &\left(t + \frac{a}{k}\right)P(m, t) - \left(t - 1 + \frac{a}{k}\right)P(m, t-1) \\ &= \frac{a}{km} - (1-a) \cdot \frac{t-1+\frac{a}{k}}{t-1+c} \end{aligned}$$

$$\cdot mP(m, t-1) \quad (13)$$

となる。

(d) 定常状態における確率

t として、定常状態になると仮定して、大きさ x の確率を求めよう。その準備として(12)式、(13)式に出てくる

$$\frac{t-1+\frac{a}{k}}{t-1+c}$$

の値について考えてみよう。

もし $a = c = 0$ ならば、この値は 1 である。また、グループ数が要素数に比例するという仮定に従うとすれば $a = kc$ であり、この場合も値は 1 になる。

グループ数と要素数は比例関係に近いと考えられるが、それから多少はずれたとしても、 t が大きな値をとる場合には分母・分子がほぼ等しい値になる。したがって、この値は 1 とみなしてよい。

以上のことを考慮すると、 $x > m$ の場合における(12)式は

$$\begin{aligned} &\left(t + \frac{a}{k}\right)P(x) - \left(t - 1 + \frac{a}{k}\right)P(x) \\ &= (1-a)\{(x-1)P(x-1) - xP(x)\} \end{aligned} \quad (14)$$

となる。したがって

$$\begin{aligned} &\{1 + (1-a)x\}P(x) \\ &= (1-a)(x-1)P(x-1) \end{aligned}$$

変形して

$$\begin{aligned} P(x) &= \frac{(1-a)(x-1)}{1 + (1-a)x}P(x-1) \\ &= \frac{x-1}{x + \frac{1}{1-a}}P(x-1) \end{aligned}$$

となる。ここで

$$b = \frac{1}{1-a} \quad (15)$$

とすると

$$P(x) = \frac{x-1}{x+b}P(x-1) \quad (16)$$

と表現される。 b は後述するように分布を特徴づけるパラメータとなる。

t のとき, $x=m$ の場合における(13)式は

$$\begin{aligned} & \left(t + \frac{a}{k}\right)P(m) - \left(t-1 + \frac{a}{k}\right)P(m) \\ &= \frac{a}{km} - (1-a)mP(m) \end{aligned} \quad (17)$$

したがって

$$\{1 + (1-a)m\}P(m) = \frac{a}{km}$$

変形して

$$P(m) = \frac{\frac{a}{km}}{1 + (1-a)m} \quad (18)$$

となる。

(e) 定数の計算

ここで, 定数 $\frac{a}{km}$ の値について考えてみよう。

(10)式から

$$\frac{a}{km} = \frac{t}{g(t)-a} \cdot \frac{a}{m}$$

である。

t 時点までに t 個の要素がシステムに参入している。そのうち, 新たなグループに所属する要素の期待割合は a である。グループを形成するためには m 個の要素が必要なので t 時点におけるグループ数の期待値は,

$$\left(\frac{at}{m}\right) + (\text{初期に存在している } a \text{ 個})$$

である。したがって, $g(t) - a = \frac{at}{m}$ となるので,

$\frac{a}{km}$ の期待値は

$$\frac{a}{km} = \frac{t}{\frac{at}{m}} \cdot \frac{a}{m} = 1 \quad (19)$$

それゆえ, (18)式における $P(m)$ の期待値は

$$\begin{aligned} P(m) &= \frac{1}{1 + (1-a)m} \\ &= \frac{b}{m+b} \end{aligned} \quad (20)$$

となる。

(f) 確率分布の計算

(16)式, (20)式から

$$\begin{aligned} P(x) &= \frac{x-1}{x+b} P(x-1) \\ &= \frac{x-1}{x+b} \cdot \frac{x-2}{x+b-1} \cdots \frac{m}{m+b+1} P(m) \\ &= \frac{\Gamma(x)}{\Gamma(m)} \cdot \frac{\Gamma(m+b+1)}{\Gamma(x+b+1)} \cdot \frac{b}{m+b} \\ &= \frac{\Gamma(x)\Gamma(b+1)}{\Gamma(x+b+1)} \cdot \frac{\Gamma(m+b+1)}{\Gamma(m)\Gamma(b+1)} \cdot \frac{b}{m+b} \\ &= \frac{\Gamma(x)\Gamma(b+1)}{\Gamma(x+b+1)} \cdot \frac{\Gamma(m+b)}{\Gamma(m)\Gamma(b)} \\ &= \frac{B(x, b+1)}{B(m, b)} \end{aligned}$$

したがって

$$P(x) = \frac{B(x, b+1)}{B(m, b)} \quad (21)$$

これは定義域 $x = m, m+1, m+2, \dots$ で, パラメータ b のユール分布である。

(g) 参入下限値を単位にした確率分布の表現

ここまでは x のとりうる単位を 1 としてきたが, 参入下限値 m が 1 に比べてかなり大きいときには m を単位にした方が現実的である。

m を単位にした確率変数を Y とし

$$y = \frac{x}{m}$$

とする。

y のとりうる区間幅は

$$h = \frac{1}{m}$$

であり, 定義域は $y = 1, 1+h, 1+2h, \dots$ である。

そのとき確率変数 Y の分布は

$$P(y) = \frac{B\left(\frac{y}{h}, b+1\right)}{B\left(\frac{1}{h}, b\right)} \quad (22)$$

となる。

これは「一般化ユール分布」と呼ばれるものである。 m が大きくなるほど h は 0 に漸近する。したがってパレート分布

$$f(y) = b y^{-b-1}, \quad (y > 1) \quad (23)$$

に近似する^(注1)。

【3】先行研究の問題点

ベキ乗則現象をユール分布に帰着して説明したものとしては、Simon (1955) をもって嚆矢とする。したがって、本稿では主として Simon 論文を参考にして議論を組み立てた。Simon の考え方を踏まえた論文は数多く発表されているが、そのひとつとして、本稿では Kullmann & Kertesz (2001) を取り上げている。とくにこの論文が代表的というわけではないが、一例として参考にした。

(a) Simon 論文

Simon 論文と本稿との主な相違を記述するが、そのさい本稿の記号に翻訳して述べる。

本稿では各グループの確率の変化を説明する(1)式と(2)式から始めている。しかし、Simon 論文ではいきなり仮定 1 から(4)式を導いており、各グループについての記述である(1)式と(2)式を省いている。その理由として、Simon は(1)式から(4)式は導けるが、その逆は必ずしも成りたないことを言い、それゆえ、仮定 1 を少し緩やかな条件と解釈し、モデルを組み立てている。Simon の議論では、 $f(x, t)$ は t 時点における実現値を意味している。また $f(x, t+1)$ は t 時点からみた $t+1$ 時点での度数の予測値であり、期待値でそれを表している。したがって Simon は $E\{f(x, t+1)\}$ という記号を用いている。

本稿では各グループの確率を説明する $P_i(x, t)$ は実現値ではなく、期待値として考えている。したがって各グループの確率を合計した $f(x, t)$ も期待値である。各グループごとに実現値として(1)式、(2)式を想定することはなかり厳しい仮定になるであろう。しかし、グループの確率を合計した(4)式と(6)式であれば大数法則が作用するので、より現実的である。その点で、Simon 論文の方が本稿より緩い仮定で成りたっている。

Simon 論文で不備な点は、いきなり度数を表現する(4)式から始めているので、本稿【2】(c)で記述したグループ数についての考察が欠ける結果になっていることである。それゆえ、初期条件を与える(6)式は $m=1$ のケースについて導かれているが、(4)式を解いたときの初期条件である $P(1)$ が具体的な式として与えられていない。もっとも、

(4)式を解いて $\sum_{x=1}^{\infty} P(x) = 1$ という条件を付加す

れば、初期条件としての $P(1)$ は定まる。そのときの問題点は、(6)式における $P(1)$ と整合的であるかということである。整合的にするためには、本稿【2】(e)で記述している定数の計算を考慮しなければならない。Simon 論文ではこの点が欠落している。

(4)式と(6)式を用いて定常状態における $P(x)$ を求めるさいに、何らか追加的な仮定が必要である。本稿では(10)式がそれにあたる。Simon 論文ではすべての x, t について

$$\frac{f(x, t+1)}{f(x, t)} = \frac{t+1}{t} \quad (24)$$

を仮定している。

本稿の記述からは $a=0$ のケースで(11)式を用い

$$\frac{f(x, t+1)}{f(x, t)} = \frac{k(t+1)P(x, t+1)}{ktP(x, t)}$$

となる。さらに定常状態を想定し、

$$P(x, t) = P(x, t+1)$$

を用いて(24)式を導くことができる。このような手順を踏まないと(24)式の仮定の意味は分かりにくい。

Simon 論文では(24)式を仮定することにより

$$\frac{f(x, t+1)}{f(x, t)} = \frac{f(x-1, t+1)}{f(x-1, t)} = \frac{t+1}{t}$$

であるから

$$\frac{f(x, t)}{f(x-1, t)} = \frac{f(x, t+1)}{f(x-1, t+1)} = k(x)$$

を導き、 t によらない比を求めている。 t によらない比であるので、定常状態においては期待相対度数、すなわち確率の比と書き直してもよい。

$$\frac{P(x)}{P(x-1)} = k(x)$$

と表される。これが(16)式に対応する。

(b) Kullmann & Kertesz 論文

Kullmann & Kertesz 論文では、各グループの確率の変化を説明する(1)式、(2)式から始めている。したがって本稿と同じように $P_i(x, t)$ は実現値ではなく期待値である。

ただし、その先の展開が本稿とは異なり、度数 $f(x, t)$ を求めず、確率 $P(x, t)$ を直接求めている。すなわち、 $a = c = 0$ のケースで

$$P(x, t) = \frac{1}{t} \sum_{i=1}^t P_i(x, t) \quad (25)$$

としている。これに相当する本稿の記述では、(11)式を变形した

$$P(x, t) = \frac{1}{g(t)} \sum_{i=1}^{g(t)} P_i(x, t) \quad (26)$$

である。

(25)式と(26)式の相違は、各グループを表す i が 1 から t までをとるのか、あるいは 1 から $g(t)$ までかの点にある。すべての要素が異なるグループを形成するならば前者になるが、それ以外ではグループ数は t 個より少ない。その場合に(25)式のように t で割ると、確率 $P(x, t)$ は小さめに計算されてしまう。この点が問題である。

Kullmann & Kertesz 論文において

$$f(x, t) = \sum_{i=1}^t P_i(x, t) \quad (27)$$

とおき、新たなグループを構成していない i の確率を 0 とすれば、(3)式と同等のものが得られる。したがって、(1)式と(2)式から、(4)式と(6)式を導くことができる。

(4)式から(12)式に相当するものを導くとき、 $f(x, t)$ を $g(t)$ で割るのではなく、 t で割ることになる。ただし、各項に $f(x, t)$ あるいは $f(x, t-1)$ があるので、結果は(12)式と同じものが得られる。すなわち

$$\begin{aligned} tP(x, t) - (t-1)P(x, t-1) \\ = (1-a)\{(x-1)P(x-1, t-1) \\ - xP(x, t-1)\} \end{aligned} \quad (28)$$

である。

問題は(13)式に相当する式である。 $m=1$ の場合、(6)式の右辺第 1 項は a である。そこに注意して、(13)式に相当する式を書いてみると

$$\begin{aligned} tP(1, t) - (t-1)P(1, t-1) \\ = a - (1-a)P(1, t-1) \end{aligned} \quad (29)$$

となる。ここで定常状態における式を求めると

$$P(1) = \frac{a}{2-a}$$

になる。

これは(28)式を $\sum_{x=1}^{\infty} P(x) = 1$ という制約条件で解いたときの初期値である

$$P(1) = \frac{1}{2-a}$$

とは整合的でない。

(c) 本稿の特徴

Simon (1955) を踏まえた従来の論文にはない本稿の特徴は、初期条件の式を明確にしたことに加え、 x のとりうる下限 m を明示し、 m を単位にして分布 $P(y)$ を求めた点にある。Simon 論文では、ベキ乗則が成りたつのは x が大きな値をとるようになる場合である。すなわち、(21)式において $m=1$ の場合

$$\begin{aligned} P(x) &= \frac{B(x, b+1)}{B(1, b)} \\ &= \frac{\Gamma(b+1)}{B(1, b)} \cdot \frac{\Gamma(x)}{\Gamma(x+b+1)} \\ &= \frac{\Gamma(b+1)}{B(1, b)} \cdot \frac{1}{(x+b)(x+b-1)\cdots x} \end{aligned}$$

$x \rightarrow \infty$ とすると

$$\sim \frac{\Gamma(b+1)}{B(1, b)} \cdot x^{-(b+1)}$$

である。

本稿では大きな m の値を想定しているので、 $h = \frac{1}{m} \rightarrow 0$ の場合、(22)式が(23)式に漸近する。したがって $y > 1$ のすべての範囲にわたってベキ乗則が成立する。

【4】ランダム・モデル

仮定 1 の代わりに次の仮定をおく。仮定 2 はそのままである。

【仮定 1R】システムに参入する要素が既存のどのグループに所属するかは、各グループすべて同じ確率である。

(a) モデルの解

前述した優先的選択モデルと同様の手順を踏んで、モデルの解を求めてみよう。

まず、各グループの確率の変化を記述する式で、

(1)式に対応する $x > m$ の場合は

$$\begin{aligned} P_i(x, t) - P_i(x, t-1) \\ = K(t-1)r\{P_i(x-1, t-1) \\ - P_i(x, t-1)\} \end{aligned} \quad (30)$$

(2)式に対応する $x = m$ の場合は

$$\begin{aligned} P_i(m, t) - P_i(m, t-1) \\ = d_i(t) - K(t-1)rP_i(m, t-1) \end{aligned} \quad (31)$$

である。仮定 1R から、 r はグループの大きさに関わらず一定の値とする。

期待度数を表す(4)式に対応する式は

$$\begin{aligned} f(x, t) - f(x, t-1) \\ = K(t-1)r\{f(x-1, t-1) \\ - f(x, t-1)\} \end{aligned} \quad (32)$$

(6)式に対応するのは

$$\begin{aligned} f(m, t) - f(m, t-1) \\ = \frac{a}{m} - K(t-1)rf(m, t-1) \end{aligned} \quad (33)$$

である。

比例定数 $K(t-1)$ を求める計算において、

(7)式に対応するのは

$$\sum_{x=m}^{t-1+c} K(t-1)rf(x, t-1) = 1-a \quad (34)$$

また、期待度数を合計するとグループ数になるから、(8)式に対応するのは

$$\begin{aligned} \sum_{x=m}^{t-1+c} f(x, t-1) = g(t-1) \\ = k(t-1) + a \end{aligned} \quad (35)$$

したがって、(34)式、(35)式から

$$K(t-1)r = \frac{1-a}{k(t-1)+a} \quad (36)$$

となる。

大きさに関する確率の関係式を記述しよう。

まず $x > m$ の場合、(11)式と(36)式を(32)式に代入すると

$$\begin{aligned} \left(t + \frac{a}{k}\right)P(x, t) - \left(t-1 + \frac{a}{k}\right)P(x, t-1) \\ = \frac{1-a}{k}\{P(x-1, t-1) \\ - P(x, t-1)\} \end{aligned} \quad (37)$$

になる。したがって定常状態では

$$P(x) = \frac{1-a}{k}\{P(x-1) - P(x)\}$$

よって

$$P(x) = \frac{1}{1 + \frac{k}{1-a}}P(x-1) \quad (38)$$

である。これを解くと

$$P(x) = \left(\frac{1}{1 + \frac{k}{1-a}}\right)^{x-m}P(m)$$

になる。

ここで初期条件 $P(m)$ を求めよう。

$$\begin{aligned} \sum_{x=m}^{\infty} P(x) = P(m) \sum_{x=m}^{\infty} \left(\frac{1}{1 + \frac{k}{1-a}}\right)^{x-m} \\ = P(m) \cdot \frac{1 + \frac{k}{1-a}}{\frac{k}{1-a}} \end{aligned}$$

したがって、 $\sum_{x=m}^{\infty} P(x) = 1$ から

$$P(m) = \frac{\frac{k}{1-a}}{1 + \frac{k}{1-a}} \quad (39)$$

でなければならない。

初期条件 $P(m)$ の確率を決める関係式を(33)式から直接求めよう。

$$\begin{aligned} \left(t + \frac{a}{k}\right)P(m, t) - \left(t-1 + \frac{a}{k}\right)P(m, t-1) \\ = \frac{a}{km} - \frac{1-a}{k}P(m, t-1) \end{aligned} \quad (40)$$

定常状態では

$$P(m) = \frac{a}{km} - \frac{1-a}{k}P(m)$$

ここで(19)式から $\frac{a}{km} = 1$ とすると

$$P(m) = \frac{\frac{k}{1-a}}{1 + \frac{k}{1-a}}$$

したがって(39)式で求めた $P(m)$ と整合的である。それゆえ

$$P(x) = \frac{\frac{k}{1-a}}{1 + \frac{k}{1-a}} \cdot \left(\frac{1}{1 + \frac{k}{1-a}} \right)^{x-m} \quad (41)$$

が得られる。これは幾何分布である。

(b) 参入下限値を単位にした確率分布の表現
参入下限値 m を単位とした分布を考えよう。

$$p = \frac{\frac{k}{1-a}}{1 + \frac{k}{1-a}}, \quad q = \frac{1}{1 + \frac{k}{1-a}}, \quad p + q = 1$$

とすると

$$P(x) = p q^{x-m} \quad (42)$$

と書ける。ここで m を単位とする確率変数 Y を

$$y = \frac{x-m}{m}, \quad h = \frac{1}{m}$$

とする。確率密度関数は

$$P(y) = p q^{\frac{y}{h}} \quad (43)$$

と書ける。 $x = m, m+1, m+2, \dots$ の値をとるから $y = 0, h, 2h, \dots$ の値である。

【定理】(43)式において、区間幅 h を限りなく 0 に近づけたとき $\frac{P(y)}{h}$ は指数分布

$$f(y) = g e^{-gy}, \quad (y > 0) \quad (44)$$

に収束する。ただし、 $g = \frac{p}{qh}$ である。

(証明) $p = \frac{gh}{1+gh}, q = \frac{1}{1+gh}$ であるから

$$\begin{aligned} \frac{P(y)}{h} &= \frac{gh}{1+gh} \left(\frac{1}{1+gh} \right)^{\frac{y}{h}} \frac{1}{h} \\ &= \frac{g}{1+gh} (1+gh)^{\frac{1}{h}(-gy)} \end{aligned}$$

ここで $h \rightarrow \infty$ とすると

$$f(y) = \lim_{h \rightarrow \infty} \frac{P(y)}{h} = g e^{-gy}$$

(証明終わり)

定理より、ランダム・モデルにおいて参入下限値 m を単位とした分布は、指数分布に近似される。

次の混合モデルを議論する準備として、 g を用いて分布を表現しておこう。

$$\begin{aligned} g &= \frac{p}{qh} \\ &= \frac{k}{1-a} \cdot \frac{1}{h} \end{aligned}$$

したがって

$$\frac{k}{1-a} = gh = \frac{g}{m} \quad (45)$$

である。(41)式は次のように表現される。

$$P(x) = \frac{\frac{g}{m}}{1 + \frac{g}{m}} \cdot \left(\frac{1}{1 + \frac{g}{m}} \right)^{x-m} \quad (46)$$

【5】混合モデル

仮定 1 の代わりに次の仮定をおく。仮定 2 はそのままである。

【仮定 1 M】システムに参入する要素が既存のどのグループに所属するかは、各グループの大きさに誘引される部分と、すべてのグループに共通な誘因部分とからなる。

混合モデルは優先的選択モデルとランダム・モデルをミックスしたものである。

(a) モデルの解

優先的選択モデルと同様の手順を踏んで、モデルの解を求めよう。 $x > m$ の場合、(1)式に対応するのは

$$\begin{aligned} P_i(x, t) - P_i(x, t-1) &= K(t-1) \{ (x-1)P_i(x-1, t-1) \\ &\quad - xP_i(x, t-1) + sP_i(x-1, t-1) \\ &\quad - sP_i(x, t-1) \} \quad (47) \end{aligned}$$

$x = m$ の場合、(2)式に対応するのは

$$\begin{aligned} P_i(m, t) - P_i(m, t-1) &= d_i(t) - K(t-1) \{ mP_i(m, t-1) \\ &\quad + sP_i(m, t-1) \} \quad (48) \end{aligned}$$

である。仮定 1 M から、 s はグループの大きさに関わらず一定の値とする。

(4)式に対応する期待度数の式は(47)式から

$$\begin{aligned} f(x, t) - f(x, t-1) &= K(t-1) \{ (x-1+s)f(x-1, t-1) \\ &\quad - (x+s)f(x, t-1) \} \end{aligned} \quad (49)$$

(6)式に対応するのは

$$\begin{aligned} f(m, t) - f(m, t-1) &= \frac{a}{m} - K(t-1)(m+s)f(m, t-1) \end{aligned} \quad (50)$$

である。

比例定数 $K(t-1)$ を求めよう。(7)式に対応するのは

$$\begin{aligned} \sum_{x=m}^{t-1+c} K(t-1) \cdot (x+s)f(x, t-1) &= 1-a \end{aligned} \quad (51)$$

要素数全体の値から

$$\sum_{x=m}^{t-1+c} xf(x, t-1) = t-1+c \quad (52)$$

また、期待度数の合計から

$$\sum_{x=m}^{t-1+c} f(x, t-1) = k(t-1)+a \quad (53)$$

したがって、(51)式に(52)式、(53)式を代入して

$$K(t-1)[(t-1+c)+s\{k(t-1)+a\}] = 1-a$$

変形して

$$K(t-1) = \frac{1-a}{t-1+c+s\{k(t-1)+a\}} \quad (54)$$

となる。

大きさに関する確率の関係式を記述しよう。

$x > m$ の場合、(49)式に(11)式と(54)式を代入すると

$$\begin{aligned} \left(t + \frac{a}{k}\right)P(x, t) - \left(t-1 + \frac{a}{k}\right)P(x, t-1) &= \frac{(1-a)(t-1 + \frac{a}{k})}{(t-1+c)+s\{k(t-1)+a\}} \\ &\quad \cdot [(x-1+s)P(x-1, t-1) - (x+s)P(x, t-1)] \end{aligned} \quad (55)$$

である。定常状態では

$$\left(t + \frac{a}{k}\right)P(x) - \left(t-1 + \frac{a}{k}\right)P(x)$$

$$\begin{aligned} &= \frac{(1-a)(t-1 + \frac{a}{k})}{(t-1+c)+s\{k(t-1)+a\}} \\ &\quad \cdot [(x-1+s)P(x-1) - (x+s)P(x)] \end{aligned}$$

【2】(d)の議論を踏まえて

$$P(x) = \frac{x-1+s}{x+s + \frac{1+sk}{1-a}} P(x-1)$$

となる。

(15)式および(45)式から

$$\begin{aligned} \frac{1+sk}{1-a} &= \frac{1}{1-a} + \frac{sk}{1-a} \\ &= b + \frac{sg}{m} \end{aligned} \quad (56)$$

したがって

$$P(x) = \frac{x-1+s}{x+s + b + \frac{sg}{m}} P(x-1) \quad (57)$$

と表現される。これを解くと

$$\begin{aligned} P(x) &= \frac{x-1+s}{x+s + b + \frac{sg}{m}} \cdot \frac{x-2+s}{x-1+s + b + \frac{sg}{m}} \\ &\quad \cdots \frac{m+s}{m+1+s + b + \frac{sg}{m}} \cdot P(m) \\ &= \frac{\Gamma(x+s)}{\Gamma(m+s)} \cdot \frac{\Gamma(m+s + b + \frac{sg}{m} + 1)}{\Gamma(x+s + b + \frac{sg}{m} + 1)} \\ &\quad \cdot P(m) \end{aligned}$$

ここで $\sum_{x=m}^{\infty} P(x)$ を計算するために、つぎの準備が必要である。

$$\begin{aligned} \Delta \left(\frac{\Gamma(x+s)}{\Gamma(x+s + b + \frac{sg}{m})} \right) &= - \left(b + \frac{sg}{m} \right) \cdot \frac{\Gamma(x+s)}{\Gamma(x+s + b + \frac{sg}{m} + 1)} \end{aligned}$$

以下、上式を途中の変形で用いる。

$$\begin{aligned} \sum_{x=m}^{\infty} P(x) &= P(m) \cdot \frac{\Gamma(m+s + b + \frac{sg}{m} + 1)}{\Gamma(m+s)} \end{aligned}$$

$$\begin{aligned}
 & \cdot \sum_{x=m}^{\infty} \frac{\Gamma(x+s)}{\Gamma(x+s+\mathbf{b}+\frac{sg}{m}+1)} \\
 = & -P(m) \cdot \frac{\Gamma(m+s+\mathbf{b}+\frac{sg}{m}+1)}{\Gamma(m+s)} \\
 & \cdot \frac{1}{\mathbf{b}+\frac{sg}{m}} \cdot \sum_{x=m}^{\infty} \Delta \left(\frac{\Gamma(x+s)}{\Gamma(x+s+\mathbf{b}+\frac{sg}{m})} \right) \\
 = & P(m) \cdot \frac{\Gamma(m+s+\mathbf{b}+\frac{sg}{m}+1)}{\Gamma(m+s)} \\
 & \cdot \frac{1}{\mathbf{b}+\frac{sg}{m}} \cdot \frac{\Gamma(m+s)}{\Gamma(m+s+\mathbf{b}+\frac{sg}{m})} \\
 = & P(m) \cdot \frac{m+s+\mathbf{b}+\frac{sg}{m}}{\mathbf{b}+\frac{sg}{m}}
 \end{aligned}$$

したがって

$$P(m) = \frac{\mathbf{b} + \frac{sg}{m}}{m+s+\mathbf{b}+\frac{sg}{m}} \tag{58}$$

でなければならない。

初期条件の確率 $P(m)$ を (50) 式から直接求めよう。

$$\begin{aligned}
 & \left(t + \frac{a}{k} \right) P(m, t) - \left(t - 1 + \frac{a}{k} \right) P(m, t-1) \\
 = & \frac{a}{km} - \frac{(1-a)(m+s)(t-1+\frac{a}{k})}{(t-1+c)+s\{k(t-1)+a\}} \\
 & \cdot P(m, t-1) \tag{59}
 \end{aligned}$$

定常状態では【2】(d)の議論を踏まえて

$$P(m) = \frac{a}{km} - \frac{(1-a)(m+s)}{1+sk} P(m)$$

ここで(19)式から $\frac{a}{km} = 1$ とすると

$$\begin{aligned}
 P(m) &= \frac{\frac{1+sk}{1-a}}{m+s+\frac{1+sk}{1-a}} \\
 &= \frac{\mathbf{b} + \frac{sg}{m}}{m+s+\mathbf{b}+\frac{sg}{m}}
 \end{aligned}$$

となり, (58)式に一致するので整合的である。

(57)式を解いて, 初期条件の(58)式を代入すれば

$$P(x) = \frac{B(x+s, \mathbf{b} + \frac{sg}{m} + 1)}{B(m+s, \mathbf{b} + \frac{sg}{m})} \tag{60}$$

が得られる。

(b) 両極端のケース

(60)式において $s=0$ とおけば

$$P(x) = \frac{B(x, \mathbf{b} + 1)}{B(m, \mathbf{b})}$$

となり, 優先的選択モデルの解である(21)式に一致する。

s の場合には, (46)式に一致することを述べよう。

$$\begin{aligned}
 P(x) &= \frac{B(x+s, \mathbf{b} + \frac{sg}{m} + 1)}{B(m+s, \mathbf{b} + \frac{sg}{m})} \\
 &= \frac{\Gamma(x+s)\Gamma(\mathbf{b} + \frac{sg}{m} + 1)}{\Gamma(x+s+\mathbf{b} + \frac{sg}{m} + 1)} \\
 & \cdot \frac{\Gamma(m+s+\mathbf{b} + \frac{sg}{m})}{\Gamma(m+s)\Gamma(\mathbf{b} + \frac{sg}{m})} \\
 &= \frac{\Gamma(x+s)}{\Gamma(m+s)} \cdot \frac{\Gamma(m+s+\mathbf{b} + \frac{sg}{m} + 1)}{\Gamma(x+s+\mathbf{b} + \frac{sg}{m} + 1)} \\
 & \cdot \frac{\mathbf{b} + \frac{sg}{m}}{m+s+\mathbf{b} + \frac{sg}{m}} \\
 &= \frac{(x-1+s)(x-2+s)}{(x+s+\mathbf{b} + \frac{sg}{m})(x-1+s+\mathbf{b} + \frac{sg}{m})} \\
 & \cdot \dots (m+s) \\
 & \cdot \dots (m+s+\mathbf{b} + \frac{sg}{m} + 1) \\
 & \cdot \frac{\mathbf{b} + \frac{sg}{m}}{m+s+\mathbf{b} + \frac{sg}{m}} \\
 &= \frac{(1+\frac{x-1}{s})(1+\frac{x-2}{s})}{(1+\frac{g}{m}+\frac{x+b}{s})(1+\frac{g}{m}+\frac{x+b-1}{s})} \\
 & \cdot \dots (1+\frac{m}{s}) \\
 & \cdot \dots (1+\frac{g}{m}+\frac{m+b+1}{s})
 \end{aligned}$$

$$\frac{\frac{g}{m} + \frac{b}{s}}{1 + \frac{g}{m} + \frac{m+b}{s}} \cdot \frac{\dots \left(\frac{1}{h} + s + b + \frac{sg}{m} - 1\right)}{\dots \left(\frac{y}{h} + s + b + \frac{sg}{m}\right)}$$

ここで s とする。

$$P(x) = \left(\frac{1}{1 + \frac{g}{m}}\right)^{x-m} \cdot \frac{\frac{g}{m}}{1 + \frac{g}{m}}$$

となり, ランダム・モデルの解である(46)式に一致する。

$$\begin{aligned} & \cdot \left(b + \frac{sg}{m}\right) \\ & = \frac{1}{h} \cdot \left(\frac{1}{h}\right)^{b+\frac{sg}{m}} \cdot \frac{\{1+hs\}\{1+h(s+1)\}}{\{y+hs\}\{y+h(s+1)\}} \\ & \quad \cdot \frac{\dots \{1+h(s+b+\frac{sg}{m}-1)\}}{\dots \{y+h(s+b+\frac{sg}{m})\}} \end{aligned}$$

(c) 参入下限値を単位にした確率分布の表現

参入下限値 m を単位にする確率変数 Y の分布を求めよう。(60)式において

$$y = \frac{x}{m}, \quad h = \frac{1}{m}$$

とおく。 $x = m, m+1, m+2, \dots$ であるから $y = 1, 1+h, 1+2h, \dots$ の値をとる。

(60)式は

$$P(y) = \frac{B\left(\frac{y}{h} + s, b + \frac{sg}{m} + 1\right)}{B\left(\frac{1}{h} + s, b + \frac{sg}{m}\right)} \quad (61)$$

と変形される。

(61)式で $h = 0$ としたときの確率密度関数

$$f(y) = \lim_{h \rightarrow 0} \frac{P(y)}{h}$$

を求めよう。

$$\begin{aligned} \frac{P(y)}{h} &= \frac{1}{h} \cdot \frac{B\left(\frac{y}{h} + s, b + \frac{sg}{m} + 1\right)}{B\left(\frac{1}{h} + s, b + \frac{sg}{m}\right)} \\ &= \frac{1}{h} \cdot \frac{\Gamma\left(\frac{y}{h} + s\right)\Gamma\left(b + \frac{sg}{m} + 1\right)}{\Gamma\left(\frac{y}{h} + s + b + \frac{sg}{m} + 1\right)} \\ & \quad \cdot \frac{\Gamma\left(\frac{1}{h} + s + b + \frac{sg}{m}\right)}{\Gamma\left(\frac{1}{h} + s\right)\Gamma\left(b + \frac{sg}{m}\right)} \\ &= \frac{1}{h} \cdot \frac{\Gamma\left(\frac{y}{h} + s\right)}{\Gamma\left(\frac{y}{h} + s + b + \frac{sg}{m} + 1\right)} \\ & \quad \cdot \frac{\Gamma\left(\frac{1}{h} + s + b + \frac{sg}{m}\right)}{\Gamma\left(\frac{1}{h} + s\right)} \cdot \left(b + \frac{sg}{m}\right) \\ &= \frac{1}{h} \cdot \frac{\left(\frac{1}{h} + s\right)\left(\frac{1}{h} + s + 1\right)}{\left(\frac{y}{h} + s\right)\left(\frac{y}{h} + s + 1\right)} \end{aligned}$$

$$\cdot \left(b + \frac{sg}{m}\right)$$

ここで $hs = \frac{s}{m} = l$ とおく。

$$\begin{aligned} \frac{P(y)}{h} &= \frac{\{1+l\}\{1+l+h\}}{\{y+l\}\{y+l+h\}} \\ & \quad \cdot \frac{\dots \{1+l+h(b+lg-1)\}}{\dots \{y+l+h(b+lg)\}} \cdot (b+lg) \end{aligned}$$

$h = 0, s = l$ で $l = hs$ が一定の値に収束するならば, Y の分布は

$$f(y) = (b+lg)(1+l)^{b+lg}(y+l)^{-(b+lg+1)}, \quad (y > 1) \quad (62)$$

となる。

$l = 0$ ならば

$$f(y) = b y^{-(b+1)}, \quad (y > 1)$$

である。 Y の分布はパレート分布になり, (23)式に一致する。

$l = 0$ ならば, $h = 0$ であるから

$$\begin{aligned} \frac{P(y)}{h} &= \left\{ l \left(g + \frac{b}{l} \right) \right\} \left\{ l \left(1 + \frac{1}{l} \right) \right\}^{b+lg} \\ & \quad \cdot \left\{ l \left(1 + \frac{y}{l} \right) \right\}^{-(b+lg+1)} \\ &= \left(g + \frac{b}{l} \right) \left(1 + \frac{1}{l} \right)^{b+lg} \left(1 + \frac{y}{l} \right)^{-(b+lg+1)} \\ &= \left(g + \frac{b}{l} \right) \left(1 + \frac{1}{l} \right)^{l \cdot \left(g + \frac{b}{l} \right)} \\ & \quad \cdot \left(1 + \frac{y}{l} \right)^{l \cdot y \cdot \left(-\left(g + \frac{b}{l} \right) \right)} \end{aligned}$$

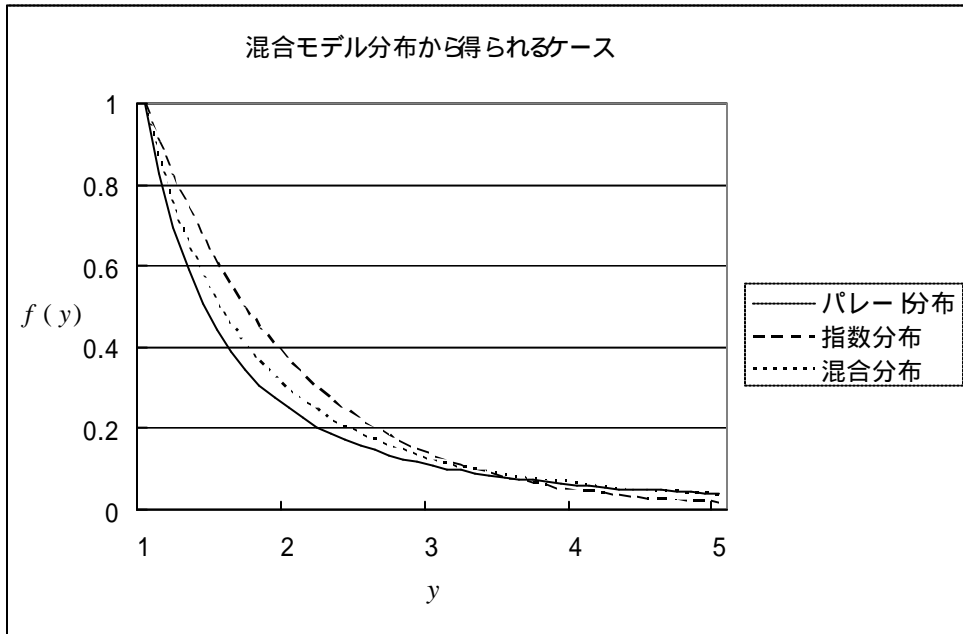
ここで l とすると

$$\begin{aligned}
 f(x) &= \lim_{l \rightarrow \infty} \frac{P(y)}{h} \\
 &= g e^g e^{-g y} \\
 &= g e^{-g(y-1)}, \quad (y > 1)
 \end{aligned}$$

である。Y の分布は指数分布になり、(44)式と一致する。ただし、(44)式は $y > 0$ の範囲で表現さ

れていることに注意する必要がある。

(62)式において $b = g = 1$ として、 $l = 0$ (パレート分布)、 $l = 1$ (指数分布)、 $l = 1$ (混合分布の一例)のケースをグラフに描いた。



(63)式において $b = g = 1$ で、
 $l = 0$ パレート分布
 $l = 1$ 指数分布
 $l = 1$ 混合分布

【注】

関係』経営志林, 第44巻第 1号, p.1-15

(1) 鈴木(2007)を参照

【参考文献】

- L. Kullmann and J. Kertesz (2001) " Preferential growth : Exact solution of the time-dependent distributions ", Physical Review E, Vol.63, 051112 p1-7
- H.A. Simon (1955) " On a Class of Skew Distribution Functions ", Biometrika, Vol.42, No.3/4, p.425-440
- 鈴木武(2007) 『パレート分布とユール分布との対応