経済主体間の相関について．（どう理解又は解釈するか？ 消費パターン？ 生産パターン，所得？）

$n=2$ の場合は Ewens SF(θ) より

$$P_2(a_1=0, a_2=2) = \frac{1}{1+\theta}$$

であり θ が大きいと2主体間の相関が小さくなる。

より一般的に $n>2$ の場合に

(n_1) $(n-n_1)$

という場合 2つの主体をサンプルした場合は

$$P_n(\text{2主体が同一タイプ} \mid n_1) = \frac{\binom{n_1}{2}+\binom{n-n_1}{2}}{\binom{n}{2}}$$

従って

$$\sum_{n_1} \frac{\binom{n_1}{2}+\binom{n-n_1}{2}}{\binom{n}{2}} \times \frac{n!}{c(n,2)\, n_1 (n-n_1)}$$

$$= \frac{\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n-1}}{1+\frac{1}{2}+\cdots+\frac{1}{n-1}} = \left(\widehat{\frac{1}{1+\theta}}\right)$$

と解釈できる．（$c(n,2) = (n-1)!\left(1+\frac{1}{2}+\cdots+\frac{1}{n-1}\right)$．）

類似の方法で $n=n_1+\cdots+n_k$, $n_i>0$, $i=1\cdots k$ の場合も考える．

青木(2003, p.152−154)の補足 ①

$E(e^{-\theta Z}) = e^{-\gamma\theta} s^{-\theta} e^{-\theta E_1(s)}$ となる訳

第152頁で 下から3行目

$\Gamma(\alpha) = \dfrac{\Gamma(1+\alpha)}{\alpha}$ に留意して

$$\left[1 - \frac{\int_s^\infty \xi^{\alpha-1} e^{-\xi} d\xi}{\Gamma(\alpha)}\right]^K = \left[1 - \frac{\alpha \int_s^\infty \xi^{\alpha-1} e^{-\xi} d\xi}{\Gamma(1+\alpha)}\right]^K \quad (*)$$

ここで

$$\frac{\alpha K \int_s^\infty \xi^{\alpha-1} e^{-\xi} d\xi}{\Gamma(1+\alpha)} \to \theta \int_s^\infty \frac{e^{-\xi}}{\xi} d\xi$$

$$= \theta E_1(s),$$

$\alpha \to 0$

$K \to \infty$

$K\alpha \to \theta$

だから

$(*) \to e^{-\theta E_1(s)}$, 全体として

$E(e^{-\theta Z_K}) \to e^{-\gamma\theta} s^{-\theta} e^{-\theta E_1(s)}$

である.

(解析的恒等式として

$$\int_0^s \frac{1-e^{-y}}{y} dy = E_1(s) + \ln s + \gamma, \quad s>0$$

を使用してもよい.)

(ii)

$$E(e^{-\alpha z}) = \exp\left[-\int_0^1 (1-e^{-\alpha x})\frac{\theta}{x}dx\right]$$

を1回 α について微分して $\alpha \to 0$ とおくと

$$E(z) = \theta$$

もう1回微分して $\alpha \to 0$ により

$$Var(z) = \frac{\theta}{2}$$

が求まる.

$$n_{(1)} \geq n_{(2)} \geq \cdots$$

$n_{(1)}$の期待値と分散

θ	$E(n_{(1)})$	$Var(n_{(1)})$	$\sigma(n_{(1)})$
0.5	0.758	~~0.04~~ 0.037	0.19
1.0	0.624	0.037	〃
2.0	0.476	0.027	0.16
5.0	0.293	0.012	0.11
10.0	0.195	0.005	0.07

Source
(Arratia et. al (2002))

$E n_{(1)}$ は $E X_{(1)}$ として (Aoki 2002, P.167) に

$$E X_{(1)} = e^{-\gamma\theta}\,\Gamma(\theta+1)\,E\left[(1+z)^{-(1+\theta)}\right]$$

として与えられている。又 z の確率密度は $p_\theta(z) = e^{-\gamma\theta}\Gamma(\theta)^{-1} z^{\theta-1}$ (Aoki 1996, P.245) $0 \le z \le 1$.

$Ez = \theta$ と Jensenの不等式より

$$E\left((1+z)^{-(1+\theta)}\right) \geq (1+Ez)^{-(1+\theta)} = (1+\theta)^{-(1+\theta)}$$

を用いて $E X_{(1)}$ の lower brend を求めうる。

又 $1 + Ez = 1 + \theta + E(z-\theta)$ として近似を改善できる (θが小さいとき)

n elementsで k-cycleをもつ

クラスター数 と θ

青木 2004 5月

置換の数 : $C(n,k)$: Stirling 数の第一種

別の定義 $x^{[n]} = \sum_{k=1}^{n} C(n,k)\, x^k$; $x^{[n]} = x(x+1)\cdots(x+n-1)$

$$C(n+1,k) = nC(n,k) + C(n,k-1) \quad (*)$$

$$C(n,n) = 1$$

$$C(n,1) = (n-1)!$$

(*)より直接的に: $C(n,2) = (n-1)!\left[1 + \frac{1}{2} + \cdots + \frac{1}{n-1}\right]$

$$= (n-1)!\left[\gamma + \ln n - \frac{1}{2n} + O\left(\frac{1}{n^2}\right)\right]$$

ここで $\gamma = \lim\left[\sum_{1}^{n} \frac{1}{j} - \ln n\right] = .5772$: オイラーの定数

より一般には
Hwangの近似式 : H.-K. Hwang (1995) "Asymptotic Expansion for the Stirling Numbers of the First Kind" J. Combinatorial Theory Series A 71 343-351,

を用いる

$$C(n,k) = \frac{(n-1)!\,(\ln(n))^{k-1}}{(k-1)!}\left\{\frac{1}{\Gamma\left(1+\frac{k-1}{\ln n}\right)} + O\left(\frac{k}{(\ln n)^2}\right)\right\}$$

ここで $\frac{k-1}{\ln n} < 1$ の場合

$$\Gamma\left(1+\frac{k-1}{\ln n}\right) = 1 + \Gamma'\left(\frac{k-1}{\ln n}\right) + \cdots$$

$$= 1 - \gamma\frac{k-1}{\ln n} + \cdots$$

$$\frac{1}{\Gamma\left(1+\frac{k-1}{\ln n}\right)} = 1 + \gamma\frac{k-1}{\ln n} + \cdots$$ となる

$k=2$ の場合には　Hwang は

$$C(n,2) = (n-1)!\,(\ln n)\left(1-\gamma\frac{1}{\ln n}\right)$$

で OK.

$c(n,k)$ を用いて

$$\Pr(K_n = k) = \frac{\theta^k}{\theta^{[n]}}\, c(n,k)$$

母関数

$$E(z^{K_n}) = \sum_{k=1}^{n} z^k P(K_n = k)$$

$$= \sum \frac{(\theta z)^k}{\theta^{[n]}}\, c(n,k)$$

$$= \frac{(\theta z)^{[n]}}{\theta^{[n]}}$$

を用いて

$$E(K_n) = \frac{d}{dz} E(z^{K_n})\Big|_{z=1}$$

$$= \sum_{j=1}^{n} \frac{\theta}{\theta + j - 1}$$

数値の計算で又

$$\mathrm{Var}(K_n) = \theta \sum_{1}^{n} \frac{j-1}{(\theta + j - 1)^2}$$

Hwang公式より

$$\frac{P(K_n=k+1)}{P(K_n=k)} = \theta\frac{c(n,k+1)}{c(n,k)} = \theta\frac{\ln n}{k}\frac{\Gamma(1+\frac{k-1}{\ln n})}{\Gamma(1+\frac{k}{\ln n})}$$

$\frac{k}{\ln n} < 1$ の場合

$$\frac{P(K_n=k+1)}{P(K_n=k)} \simeq \frac{\theta\ln n}{k}\frac{\ln n-\gamma(k-1)}{\ln n-\gamma k} \gtrless 1$$

$$\Longleftrightarrow \quad \frac{\theta\ln n}{k} \gtrless \frac{1-\gamma\frac{k}{\ln n}}{1-\gamma\frac{k-1}{\ln n}}$$

$$E(K_n) = \begin{cases} 9.8\,\theta & n=10^4 \\ 14.4\,\theta & n=10^6 \end{cases}$$

$\theta = 1$

$$E(K_n) = \begin{cases} 7.5 & n=10^3 \\ 12.1 & n=10^5 \end{cases}$$

$$Var(K_n) = \begin{cases} .75 & \theta = 0.1 \quad n=10^3 \\ 1.21 & \theta = 0.1 \quad n=10^5 \end{cases}$$

$\theta = 1$ $\quad E(K_n) = \sum_1^n \frac{1}{j} = \gamma + \ln n$

$$Var(K_n) = \sum_{j=1}^{n} \frac{j-1}{j^2} \approx \gamma + \ln n - \sum_1^n \frac{1}{j^2}$$

$$\approx \gamma + \ln n - \int_1^{n-1} \frac{dx}{x^2}$$

$$= \gamma + \ln n - 1 + \frac{1}{n-1}$$

$$= \begin{cases} 6.5 & n=10^3 \\ 11.1 & n=10^5 \end{cases}$$

MLE の θ ; $\hat{\theta}$ は

$$\sum_{j=0}^{n-1} \frac{\hat{\theta}}{\hat{\theta}+j} = K_n$$

の unique な解

$$\sqrt{\ln n}\left(\frac{\hat{\theta}_n - \theta}{\sqrt{\theta}}\right) \xrightarrow{d} N(0,1)$$

も知られている.

Herfindahl index of concentration

$x_i \geq 0$, $\sum x_i = 1$ Aoki 2002, p. 174

$Y = \sum x_i^2$ $x_i \sim \theta x^{-1}(1-x)^{\theta-1}$ (freq. spectrum)

$E\, Y = \frac{1}{1+\theta}$

$\parallel$

$\theta \int x^2 x^{-1}(1-x)^{\theta-1} dx$

$\frac{E\, j a_j}{n} = \theta \frac{(n-1)!}{(n-j)!} \frac{\Gamma(\theta+n-j)}{\Gamma(\theta+n)}$: (# of agents) share of size j clusters.

$j=1$: singleton share $= \theta \frac{\Gamma(\theta+n-1)}{\Gamma(\theta+n)} = \frac{\theta}{\theta+n-1} \leq 0.05$

$\theta \leq \frac{n-1}{19}$, $n \geq 1 + 19\theta$ (5%)

$\theta = 0.1$ $n \geq 3$

$\theta = 0.4$ $n \geq 9$

主体の相関

$$\Pi_n(\underline{a}) = \frac{n!\,\theta^{K_n}}{\theta^{[n]}} \prod \left(\frac{1}{j}\right)^{a_j} \frac{1}{a_j!} \ ; \ K_n = \sum_{1}^{n} a_j$$

より

$n=2$, $\underline{a}=(0,1)$ つまり $a_1=0$, $a_2=1$, すると

$$\Pi_2(a_1=0, a_2=1) = \frac{1}{1+\theta} \qquad \left(\frac{2!}{\theta(\theta+1)}\theta^1\left(\frac{1}{2}\right) = \frac{1}{1+\theta} \right)$$

$\theta=6$ $\quad \frac{1}{1+\theta} = 0.14$

$\theta=0.4$ $\quad \frac{1}{1+\theta} = 0.71$

$\theta=0.5$ $\quad \frac{1}{1+\theta} = .67$

2つの主体の相関?(消費パターン;所得;サイズ etc ?)

数値例

$10! = 3{,}628{,}800 \simeq 3.6\times10^6$

Stirlingの近似 $\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right)$

$3{,}598{,}695.6$

$\simeq 3.6\times10^6$

$20! \simeq 2.43\times10^{18}$

$19! \simeq \cancel{8.34\times10^{16}}$ $\longrightarrow$ $19! \simeq 1.21\times10^{17}$

$$C(n,k) \simeq \frac{(n-1)!(\ln n)^{k-1}}{(k-1)!} \frac{1}{\Gamma\left(1+\frac{k-1}{\ln n}\right)}$$

$\ln 20 = 3.0$

$C(20,1) = 19! \simeq 1.21\times10^{17}$

$$C(20,2) = \frac{19!\,(3.0)}{\Gamma(1+\frac{1}{3})} \simeq 19!\times3\left(1+\frac{\gamma}{3}\right) \simeq 4.33\times10^{17}$$

$C(20,3) \simeq 7.04\times10^{17}$

$C(20,4) \simeq 5.45\times10^{17}$

$\theta = 0.5$, $\mathbb{E}K_{20} = 3.77$

$K_n = 4$

$20 = n_1 + n_2 + n_3 + n_4$, $n_i > 1$

non-empty

\# of patterns $\binom{20-1}{3} = \frac{19!}{3!\,16!}$: # of 4 clusters

$\simeq \frac{1.21 \times 10^{17}}{6 \times 2.087 \times 10^{14}}$; $16! \simeq \sqrt{2\pi \cdot 16}\left(\frac{16}{e}\right)^{16}$

$= 2.087 \times 10^{14}$

$= \cancel{9.66 \times 10}$ 96.6

$$\frac{\cancel{\Pr(a_{10}=1, a_1=3) = \frac{20!}{\theta^{[20]}}\theta^4 \frac{1}{3!} \times \frac{1}{10}}}{\cancel{\Pr(a_{10}=1, a_8=1, a_2=2) = \frac{20!}{\theta^{[20]}}\theta^4 \frac{1}{2 \cdot 8 \cdot 10}}} = 2.[illegible]$$

@ ~~Pr(a_11 = 1, a_1 =~~

$$\frac{\Pr(a_{10}=1, a_4=2, a_1=2) = C \cdot \frac{1}{2 \cdot 4^2\, 2!\, 10}}{\Pr(a_{10}=1, a_1=3) = C \cdot \frac{1}{3!\, 10}}$$

$= 0.094$

$$\frac{\Pr(a_8=2, a_2=2) = C \cdot \frac{1}{2^2 \cdot 2!\; 8^2 \cdot 2!}}{\Pr(a_{10}=1, a_4=2, a_1=2) \qquad C \frac{1}{2 \cdot 4^2 \cdot 2!\; 10}} = 0.625$$

$\theta = 0.5$

$$P(K_{20}=2) = \frac{C(20,2)\,\theta}{\theta^{[n]}} = \frac{4.33\times 10^{17}\times 0.\cancel{25}5}{1.68\times 10^{29}} = \cancel{6.44}\,1.29\times 10^{-12}$$

$$\theta^{[n]} = \theta(\theta+1)(\theta+2)\cdots(\theta+19)\Big|_{\theta=\frac{1}{2}} = \frac{39!}{2^{20}\,2\times 19!}$$

$$\simeq \left(\frac{39}{19}\right)^{19.5}\left(\frac{39}{e}\right)^{20}$$

$$\simeq 1.68\times 10^{29}$$

$$P(K_{20}=3) = \frac{C(20,3)\,\theta^2}{\theta^{[20]}} = \frac{7.04\times 10^{17}\times 0.25}{1.68\times 10^{29}} = \cancel{1.64}\,1.04\times 10^{-12}$$

$$P(K_{20}=4) = \frac{C(20,4)\,\theta^3}{\theta^{[20]}} = \frac{5.45\times 10^{17}\times 0.125}{1.68\times 10^{29}} = 4.06\times 10^{-13}$$

$$E K_{20} = \theta\left(\frac{1}{\theta} + \frac{1}{\theta+1} + \cdots + \frac{1}{\theta+n-1}\right)_{n=20}$$

$$= 1 + \theta\left(\frac{1}{\theta+1} + \cdots + \frac{1}{\theta+n-1}\right)_{n=20}$$

$\theta = 0.5$: $E K_{20} = 1 + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{39}$

$$= 1 + \frac{1}{2} + \cdots + \frac{1}{39} - \frac{1}{2}$$

$$\simeq \gamma + \ln 40 - \frac{1}{2} = 3.77$$

$$= \cancel{3.27}$$

$n = 3$

$$P(a_1=1, a_2=1) = \frac{3!\ \theta^2}{\theta^{[3]}} \frac{1}{2 \cdot \cancel{2}}$$

$\theta = \frac{1}{2}$

$$= \frac{3\theta}{(\theta+1)(\theta+2)} \quad : \quad \frac{1{,}5}{1{,}5 \times 2{,}5} = \frac{1}{2{,}5} = \underline{0{,}4}$$

$$P(a_1 = 3) = \frac{3! \times 3 \cdot \theta^3}{\theta^{[3]}} = \frac{\cancel{2}\ \theta^2}{(\theta+1)(\theta+2)} = \underline{0{,}067}$$

$$P(a_3 = 1) = \frac{2}{(\theta+1)(\theta+2)} = \underline{0{,}53}$$

← prob. of the largest cluster

$E(K_3) = 1{,}53$

$P(\cancel{a_1 = 2},$