人口問題研究（J. of Population Problems）53－4（1997.12）pp.67～74

研究ノート

安定人口における姉妹数

鈴木 透

1. 問題

本稿では，単性女子の安定人口における姉妹数の期待値に関し，ふたつの問題を扱う．ひとつは，通常知られている姉妹数の期待値が，特殊な抽出方法を前提としていることを指摘する．娘数の確率分布から姉妹数の確率分布を作る場合，母親を抽出してその娘集団についての姉妹数を論じていることになる．これに対し，娘を直接抽出した場合，姉妹数の期待値がどう変わるかを示す．

もうひとつの問題は，姉妹数が確定していない若い女子について，姉妹数の実現値を推定することである．姉の数は本人出生時点で確定しているが，生まれたばかりの乳児にとって妹数はゼロである．加齢とともに妹が増え，母親が再生産年齢を超えた時点で姉妹数が確定することになる．この過程を横断的に集約し，人口全体に関する妹数の期待値を求める．

2. 抽出方法による確定姉妹数の違い

単性女子の安定人口について，再生産年齢の下限と上限を（α, β）と置く．$\beta-\alpha$ 歳以上の女子について，その母親は全て再生産を終了している．ここではそのような女子を対象に，姉妹数の確定値を考える．本稿では姉妹の出生数のみ考察の対象とし，生存しているか否かは考慮しない．

任意の女子が最終的に娘を i 人産む確率を B_i，最大娘数を I，最終的な娘数の期待値（純再生産率）を N とする．

$$N=\sum_{i=1}^{I} i\cdot B_i \qquad (2\text{-}1)$$

娘から見て，母親が i 人の娘を産んだことは，自分に i-1 人の姉妹がいることを意味する．母親ひとりに娘 N 人が対応するから，任意の娘が i-1 人の姉妹を持つ確率は $i\cdot B_i/N$ で表される．従って姉妹数の期待値 S は，娘数の分散を σ_N^2 として，

$$S = \sum_{i=1}^{I}(i-1)\frac{i \cdot B_i}{N}$$

$$= \frac{1}{N}\sum_{i=1}^{I} i^2 B_i - \frac{1}{N}\sum_{i=1}^{I} i \cdot B_i$$

$$= \frac{1}{N}(\sigma_N^2 + N^2) - 1$$

$$= \frac{\sigma_N^2}{N} + N - 1 \tag{2-2}$$

この結果はよく知られているが[1)]，再生産を終えた母親を特定した上でその娘全員についての姉妹数を考えている点に注意する必要がある．つまりある女子を抽出した場合，その姉妹全員も標本に含めなければならない．このような抽出法を，ここでは「家族抽出」と呼ぶことにする．これに対し，姉妹数が確定している女子を無作為に選び，その姉妹については関知しない抽出法を「個人抽出」と呼ぶ．

年齢と同時に既往出生娘数（パリティ）を考慮した安定人口モデルを考える．瞬間年齢xで既往出生娘数iである女子の生存率（出生時＝1）を$l_{x,i}$，その女子が新たに第$i+1$娘を産む確率を$m_{x,i}$とする．任意の女子が産む第i娘の期待値は，

$$N_i = \int_{\alpha}^{\beta} l_{x,i-1} m_{x,i-1} dx \tag{2-3}$$

これをiについて合計すると，全出生順位の娘数の期待値Nが得られる．従って，母親を特定する家族抽出において，第i娘が占める割合は，

$$O_i = \frac{N_i}{N} \tag{2-4}$$

純再生産率のキュミュラント展開[2)]より，

$$\log O_i = \log N_i - r\mu + \frac{r^2}{2!}\sigma^2 - \frac{r^3}{3!}\kappa^3 + \cdots \tag{2-5}$$

ただしrは内的自然増加率，μは全娘の出産年齢の平均，σ^2は分散，κ^3は3次モメント，logは自然対数である．次に，$\beta-\alpha$歳以上の特定のコーホートからの個人抽出にお

1）たとえば以下を参照．Goodman, Leo A., Nathan Keyfitz and Thomas W. Pullum, "Addendum to family formation and the frequency of various kinship relationships", *Theoretical Population Biology*, No. 8, 1975, p. 378; Keyfitz, Nathan, *Applied Mathematical Demography, Second Edition*, New York, Springer-Verlag, 1985, p. 287; 廣嶋清志，「戦後日本における親と子の同居率の形式人口学的分析モデル」，『人口問題研究』，第167号，1983年，26頁．

2）Keyfitz（1985，前掲注1，p. 119）．

ける出生順位別分布は[3]，

$$O_i' = \int_{\alpha}^{\beta} e^{-rx} l_{x,i-1} m_{x,i-1} dx \tag{2-6}$$

この両辺を N_i で割って対数をとると，$l_{x,i-1}m_{x,i-1}/N_i$ が第 i 娘の出生年齢の密度関数であることから，

$$\log O_i' = \log N_i - r\mu_i + \frac{r^2}{2!}\sigma_i^2 - \frac{r^3}{3!}\kappa_i^3 + \cdots \tag{2-7}$$

ただし μ_i は第 i 娘の出産年齢の平均，σ_i^2 は分散，κ_i^3 は3次モメントである．(2-5) および (2-7) より，

$$\log \frac{O_i'}{O_i} = -r(\mu_i - \mu) + \frac{r^2}{2!}(\sigma_i^2 - \sigma^2) - \frac{r^3}{3!}(\kappa_i^3 - \kappa^3) + \cdots \tag{2-8}$$

ここで $r>0$ の増加人口について考えると，長女 ($i=1$) の平均出産年齢は全体より若い ($\mu_1 < \mu$) と考えられ，これは (2-8) 式のログオッズ比を高める方向に働く．ヒトの場合2%以上の増加率が長期間続くことは稀と考えられるので，(2-8) 式の第2項以降の効果は小さい．従って増加人口の場合，個人抽出における長女の割合 O_i' は，家族抽出における O_i より大きくなることが予想される．増加率が高い場合は，次女についても同じことが言えるかも知れない．逆に三女，四女といった出生順位が高い娘の割合は，家族抽出に比べ小さくなるだろう．

このように個人抽出では長女への偏りが生じるが，それは長女が相対的に若い母コーホートから生まれており，増加人口では若い母コーホートは相対的に規模が大きいことによる．しかし母コーホートの規模に影響されるのは出生順位分布であり，出生順位 i が与えられたときの姉妹数 j は，家族抽出と個人抽出の間で等しく，従って条件付き平均 $E(j|i)$ も等しいと考えられる．この結果，個人抽出における姉妹数の期待値 S' は，家族抽出における S とは一般に異なる．

$$S = \sum_{i=1}^{I} O_i E(j|i) \neq \sum_{i=1}^{I} O_i' E(j|i) = S' \tag{2-9}$$

三女・四女といった出生順位が高い娘は，少なくとも姉妹が二人または三人以上いるため，条件付き平均 $E(j|i)$ は長女・次女に比べ大きい．つまり出生順位が遅いほど，条件付き平均は大きい．増加人口では長女をはじめ出生順位の早い娘への偏りが生じるから，$S' < S$ となるだろう．

もうひとつ興味深い点は，個人抽出で姉の数と妹の数が異なることである．家族抽出で

3) 鈴木透，「安定人口と姉妹数の謎」，『人口問題研究』，第49巻第1号，1993年，9頁．

は，姉が妹をひとり数えたとき，必ずその妹は姉をひとり数えるという対称性があるため，姉の数と妹の数は常に等しい．ところが増加人口の個人抽出では，家族抽出に比べ長女への偏りがあるため，妹の数の方が姉の数より多いことになる[4]．

ここではごく簡単な数値例を挙げておく．全女子が25歳で長女を産み，半数が30歳で次女を産み，それ以降の出生はなく，また30歳まで死亡はないものとする．この場合，長女の半数が一人娘，残る半数は妹が１人いるので$E(j|1)=0.5$，次女は全員姉が１人いるので$E(j|2)=1$である．再生産に関する指標と姉妹数は表１のようになる．内的自然増加率rは，平均と分散を係数とする二次方程式から求めた．個人抽出の出生順位分布は，いったん$O'_i=N_i exp(-r\mu_i)$を計算し，誤差$=1-(O'_1+O'_2)$を比例配分した．

この表によると，予想どおり個人抽出では長女が多く，次女が少なくなっている．このため，個人抽出における姉妹数の期待値S'は家族抽出より1.25％減少している．また，個人抽出では妹の数の方が姉の数より多くなっている．

表１　数値例

再生産の指標	家族抽出	個人抽出
出生女児数分布	出生順位分布	出生順位分布
$B_0=0$	$O_1=0.6667$	$O'_1=0.6834$
$B_1=0.5$	$O_2=0.3333$	$O'_2=0.3166$
$B_2=0.5$		
出生女児数の期待値	姉妹数の期待値	姉妹数の期待値
$N=1.5$	$S=0.6667$	$S'=0.6583$
$N_1=1.0$	姉＝0.3333	姉＝0.3166
$N_2=0.5$	妹＝0.3333	妹＝0.3417
出生年齢の平均	抽出法による違い	
$\mu=26.6667$	$(S'-S)/S=-0.0125$	
$\mu_1=25.0$		
$\mu_2=30.0$		
出生年齢の分散		
$\sigma^2=5.5556$		
$\sigma_1^2=0$		
$\sigma_2^2=0$		
内的自然増加率		
$r=0.0152$		

3.　確定前の妹数

次に，$\beta-\alpha$歳で妹数が確定していない女子の妹数の実現率について考える．図１のレキシス図で，現在y歳のコーホートについて，既に生まれた妹は(1)の領域，これから生まれる妹は(2)の領域に分布する．従って妹数の実現率を，(2)における出生数が(1)と(2)の出生数の和に占める割合で推定する．

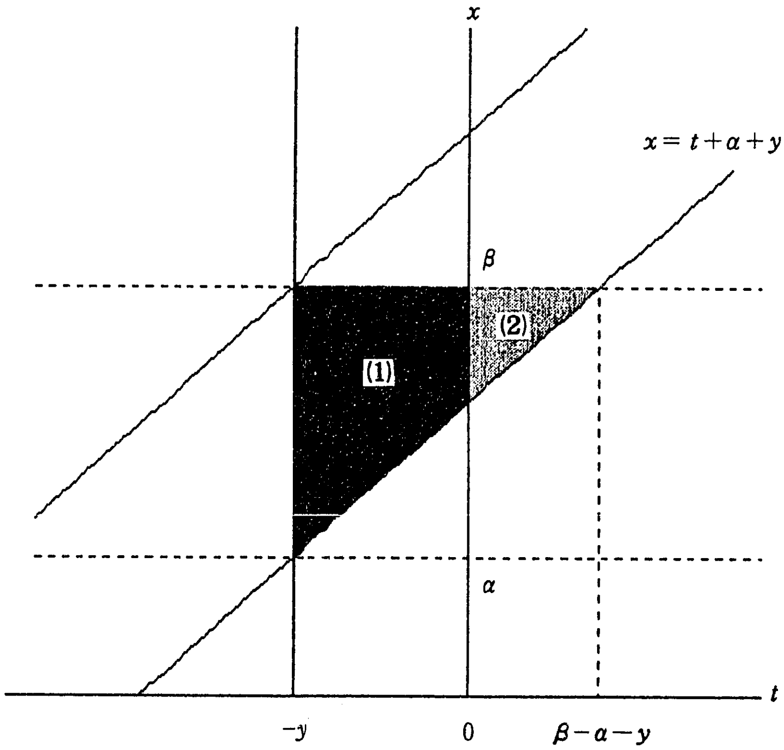
以下では繁雑な式が続くので，純再生産関数を簡単にϕ_xで表す．

$$\phi_x=\sum_{i=1}^{I} l_{x,i-1}m_{x,i-1} \tag{3-1}$$

安定人口の仮定から，t年の出生数は$B_t=B_0e^{rt}$となる．(1)における出生数は，B_0で基準化して，

4）詳しい議論は鈴木（1993，前掲注３)を参照．

図1　既に生まれた妹とこれから生まれる妹



$$\int_{-y}^{0}\int_{t+\alpha+y}^{\beta}\frac{B_t}{B_0}e^{-rx}\phi_x dxdt$$

$$=\int_{\alpha}^{\alpha+y}\int_{-y}^{x-\alpha-y} e^{rt}dt\cdot e^{-rt}\phi_x dx+\int_{\alpha+y}^{\beta}\int_{-y}^{0} e^{rt}dt\cdot e^{-rx}\phi_x dx$$

$$=\frac{1}{r}\left(e^{-r(\alpha+y)}\int_{\alpha}^{\alpha+y}\phi_x dx+\int_{\alpha+y}^{\beta}e^{-rx}\phi_x dx-e^{-ry}\right) \qquad (3\text{-}2)$$

(2)における出生数は,

$$\int_{0}^{\beta-\alpha-y}\int_{t+\alpha+y}^{\beta}\frac{B_t}{B_0}e^{-rx}\phi_x dxdt=\int_{\alpha+y}^{\beta}\int_{0}^{x-\alpha-y}e^{rt}dt\cdot e^{-rx}\phi_x dx$$

$$=\frac{1}{r}\left(e^{-r(\alpha+y)}\int_{\alpha+y}^{\beta}\phi_x dx-\int_{\alpha+y}^{\beta}e^{-rx}\phi_x dx\right) \qquad (3\text{-}3)$$

(1)と(2)の合計は,

$$\int_{-y}^{\beta-\alpha-y}\int_{t+\alpha+y}^{\beta}\frac{B_t}{B_0}e^{-rx}\phi_x dxdt=\int_{\alpha}^{\beta}\int_{-y}^{x-\alpha-y}e^{rt}dt\cdot e^{-rx}\phi_x dx$$

$$=\frac{e^{-ry}}{r}\left(Ne^{-r\alpha}-1\right) \qquad (3\text{-}4)$$

y 歳における妹数の実現率 M_y は，(3-2) を (3-4) で割った値なので，

$$M_y=\frac{e^{-r\alpha}\int_{\alpha}^{\alpha+y}\phi_x dx+e^{ry}\int_{\alpha+y}^{\beta}e^{-rx}\phi_x dx-1}{Ne^{-r\alpha}-1} \tag{3-5}$$

また，$\beta-\alpha$ 歳の中で y 歳が占める割合を c_y とする．ここでは $\beta-\alpha$ 歳までの死亡が無視でき，その年齢までの年齢構造はもっぱら人口増加率によって規定されていると仮定する．

$$c_y=\frac{e^{-ry}}{\int_0^{\beta-\alpha}e^{-rx}dx}=\frac{r\cdot e^{-ry}}{1-e^{-r(\beta-\alpha)}} \tag{3-6}$$

$\beta-\alpha$ 歳全体での妹数の実現率は，c_yMy を $(0,\ \beta-\alpha)$ に関して積分した値である．その前に，次の変数を定義しておく．

$$\Phi_z=\int_{\alpha}^{z}\phi_x dx \tag{3-7}$$

$$\Theta_z=\int_{\alpha}^{z}e^{-rx}\phi_x dx \tag{3-8}$$

部分積分法から，次の定理が成り立つ．

$$\int_{\alpha}^{\beta}z\cdot e^{-rz}\phi_z dz=[z\cdot\Theta_z]_{\alpha}^{\beta}-\int_{\alpha}^{\beta}\Theta_z dz=\beta-\int_{\alpha}^{\beta}\Theta_z dz$$

従って，安定人口の平均世代間隔を T として，

$$\int_{\alpha}^{\beta}\Theta_z dz=\beta-T \tag{3-9}$$

また，

$$\int_{\alpha}^{\beta}e^{-rz}\Phi_z dz=\left[-\frac{e^{-rz}}{r}\Phi_z\right]_{\alpha}^{\beta}-\int_{\alpha}^{\beta}\left(-\frac{e^{-rz}}{r}\right)\phi_z dz=\frac{1-Ne^{-r\beta}}{r} \tag{3-10}$$

ここで c_yM_y を Φ_z，Θ_z で表すと，

$$c_yM_y=\frac{r\cdot e^{-r(\alpha+y)}\Phi_{\alpha+y}+r(1-\Theta_{\alpha+y})-r\cdot e^{-r(\beta-\alpha)}}{(1-e^{-r(\beta-\alpha)})(Ne^{-r\alpha}-1)} \tag{3-11}$$

これを $(0,\ \beta-\alpha)$ について積分すると，(3-9) および (3-10) より，

$$\int_0^{\beta-\alpha} c_y M_y dy = \frac{(1-Ne^{-r\beta})+r\{(\beta-\alpha)-(\beta-T)\}-(1-e^{-r(\beta-\alpha)})}{(1-e^{-r(\beta-\alpha)})(Ne^{-r\alpha}-1)}$$

$$= \frac{r(T-\alpha)-e^{-r\beta}(N-e^{r\alpha})}{(e^{-r\alpha}-e^{-r\beta})(N-e^{r\alpha})}$$

ここには純再生産率Nが含まれているが，$N=e^{rT}$を用いて妹の実現率PをrとTだけの関数として表すことにする．

$$P=\int_0^{\beta-\alpha} c_y M_y dy = \frac{r(T-\alpha)-e^{-r\beta}(e^{rT}-e^{r\alpha})}{(e^{-r\alpha}-e^{-r\beta})(e^{rT}-e^{r\alpha})} \qquad (3\text{-}12)$$

図2は，$\alpha=15$，$\beta=50$，$0<r\leqq 0.02$，$25\leqq T\leqq 30$についてPの値を示したものである．増加率が大きいほど，また世代間隔が長いほど，妹の実現率は小さいことが分かる．

Pは$\beta-\alpha$歳未満の各コーホートの妹実現率の加重平均だから，視点は娘のコーホートにあり，暗黙のうちに個人抽出を前提としている．従ってPは，個人抽出における妹数の確定値にかけられるべきである．確定した姉数の期待値をS'_{old}，妹数の期待値をS'_{young}とすると，

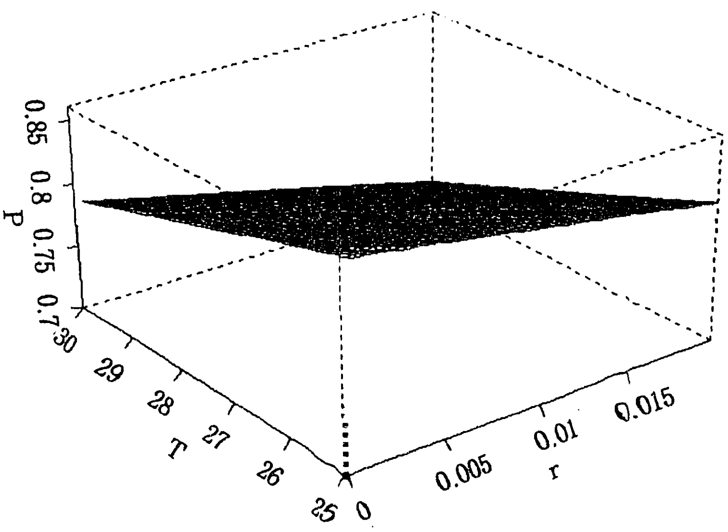
$$S'_{young}=S'-S'_{old}=S'-\sum_{i=1}^{I}(i-1)O'_i=S'-\sum_{i=1}^{I}i\cdot O'_i+1 \qquad (3\text{-}13)$$

これを用いて，$\beta-\alpha$歳未満における姉妹数の期待値は$S'_{old}+P\cdot S'_{young}$となる．

4. 結語

本稿では従来暗黙のうちに前提とされていた家族抽出の場合と比べ，個人抽出の場合に姉妹数の期待値がどう変わるかを示した．いわば分子単位で抽出する家族抽出では，都合が悪い点がいくつかある．たとえばGoldman[5]は，姉の数と妹の数の比から安定人口増加率を間接的に推定する方法を示した．これは，姉と妹の数が常に一致する家族抽出では不可能で，そ

図2　35歳未満女子の妹の実現率



5) Goldman, Noreen, "Estimating the intrinsic rate of increase of a population from the average numbers of younger and older sisters", *Demography*, Vol. 15, No. 4, 1978, pp.499-507. なお，鈴木（1993，前掲注3）も参照．

の理論的根拠は個人抽出に求めなければならない．

あるいは安定人口を離れて，たとえば出生力低下に伴いきょうだい数がどう変化するかを分析したいとする．しかし家族抽出ではきょうだい全員を含めなければならず，対象者が広い範囲のコーホートに分散してしまうことになる．この場合，比較可能なのは親コーホートについてのみで，個々のコーホートのきょうだい数を直接比較することが出来ない．

さらに本稿で示した妹の実現率も，個人抽出を前提として議論を展開している．家族抽出では問題があり，個人抽出を前提としなければならないという事態は，予想外に多いかも知れない．

本稿では単性の安定人口のみ扱い，両性への拡張は考えなかった．Pullum[6]は家族抽出を前提とする分岐過程に依拠した親族数の理論において，二項分布を用いて両性への拡張を行った．しかし個人抽出の場合，両性への拡張は簡単には行かない．たとえば娘から見たきょうだい数を求めようとしても，その母親は少なくともひとり娘を産んでいるという条件が加わるため，母親から見た息子数と娘数は二項分布しないからである．この問題については，慎重な分析が必要だろう．

6) Pullum, Thomas W., "The eventual frequencies of kin in a stable population", Demography, Vol. 19, No. 4, 1982, pp.549-565. なお，平均より高次のモメントについては，Suzuki, Toru, "A kinship model based on branching process",『人口問題研究』, 第49巻, 第2号, 17-29頁を参照.