

---

 調 査 研 究
 

---

## 安定人口増加率と姉妹数の謎

鈴木 透

## I 問題

親族数に関する形式人口学では、人口変数が親族数を決定する側面が主な関心事である。すなわち与えられた出生力・死亡力のもとで、子ども、孫、きょうだい、おじ、おば、甥・姪、いとこなどの平均数は何人か、それは本人が加齢するにつれてどう変化するか、そして人口変数が変化した場合にそれらのパターンはどのように影響されるか、といった問題に関心がもたれる<sup>1)</sup>。

しかし他方では、逆に親族数から人口動態率を決定する試みもなされている。こちらは方法論的な関心で、信頼できる年齢別データが得られない場合に、親族数データから間接的に人口動態率を推計しようというものである。

このような研究として Goldman の推計方法があり、姉妹数の比から安定人口増加率を導く方法が提示されている<sup>2)</sup>。Goldman は単性女子の古典的な安定人口モデル (Lotka モデル) に依拠し、姉数の期待値と妹数の期待値の比が人口増加率と関連していることを示した。そして前者から後者を推計するための式を導いた。

これに対して Wachter は、Goldman の推計方法の有効性をマイクロ・シミュレーションによって検証すると同時に、「姉妹数の謎 (Sisters' Riddle)」という理論的問題を提示した<sup>3)</sup>。つまり、姉が妹をひとり数えたとき、その妹は必ず姉をひとり数えるという対称性のために、姉数の期待値と妹数の期待値は一致するはずである。それなのに、なぜその比が Goldman の言うように増加率と関連するのか、という疑問である。

この問題への解答は、Wachter によって示されている。しかし Goldman や Wachter が依拠する Lotka モデルは、年齢構造のみで、既往出生児数別分布や出生順位別分布といったパリティ構造を欠いている。しかるに「姉妹数の謎」はパリティ構造と強く関連する問題であり、これを正しく理解するためには、年齢に加えてパリティ構造を持つ安定人口モデルに依拠して定式化し直した方がよい。

本稿では、まず Goldman の方法、次いで Wachter の議論を紹介する。その上で、パリティ構造を持つ安定人口モデルに依拠して姉妹数の期待値と増加率の関係を探り、「姉妹数の謎」の再考察を試みる。

---

1) 以下の拙稿およびそこで挙げた参考文献を参照。鈴木透、「直系尊属の生存確率」、『人口問題研究』、第48巻1号、1992a、pp.32-37；「親族数分布に関する解析的モデル」、『人口問題研究』、第48巻3号、1992b、pp.1-15。

2) 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.

3) Wachter, Kenneth W., "The sisters' riddle and the importance of variance when guessing demographic rates from kin counts", *Demography*, Vol.17, No.1, 1980, pp.103-114.

## II Goldman の推計方法

親族に関する人口学的モデルは、従来は出生力・死亡率によって親族数分布が決定される過程に関心があった。しかし Goldman は発想を逆転させ、親族数データから人口増加率を推計することを考えた。そこで注目したのは、任意の女子から見た姉の数と妹の数の比である<sup>4)</sup>。単純に考えて、増加中の人口では若いコーホートほど規模が大きいことから、妹の数が姉の数よりも多いことが予想される。逆に減少中の人口では、姉の数の方が多くなるだろう。

Goldman の方法は、年齢構造のみを持つ通常の安定人口モデル（女子）について定式化されている。女子の生命表人口（出生時=1）を  $l(x)$ 、 $x$  歳出生率を  $m(x)$  とする。安定人口増加率を  $r$  とし、母と本人の年齢差が  $x$  歳である確率は、 $e^{-rx}l(x)m(x)$  である。従って現在  $z$  歳の女子の妹数は、母が本人出生後  $z$  年間生存して出生を行なう確率である。再生産年齢を  $(\alpha, \beta)$  とし、妹数の期待値は次式で得られる。

$$E_z(Y) = \int_{\alpha}^{\beta} \left[ \int_0^z \frac{l(x+y)}{l(x)} m(x+y) dy \right] e^{-rx} l(x) m(x) dx$$

ここで現在  $\beta - \alpha$  歳以上の女子に限定すると、その母は全て  $\beta$  歳以上であることから、その姉妹数が確定していることになる。この場合、 $[\ ]$  内の積分範囲は  $x + y$  が  $\beta$  に達するまでになるから、整理して次式を得る。

$$E(Y) = \int_{\alpha}^{\beta} \left[ \int_x^{\beta} l(y) m(y) dy \right] e^{-rx} m(x) dx \quad (2-1)$$

一方、姉の数の期待値  $E(\Phi)$  は、本人が出生した時点で確定している。

$$E(\Phi) = \int_{\alpha}^{\beta} \left[ \int_{\alpha}^x m(y) dy \right] e^{-rx} l(x) m(x) dx$$

これは積分の順序を変えて、次のように書ける。

$$E(\Phi) = \int_{\alpha}^{\beta} \left[ \int_x^{\beta} e^{-ry} l(y) m(y) dy \right] m(x) dx \quad (2-2)$$

姉妹数の比  $Z = E(Y) / E(\Phi)$  が  $r$  と関連することは、次のようにして示すことができる。

$$Z = \frac{E(Y)}{E(\Phi)} = \frac{\int_{\alpha}^{\beta} \left[ \int_x^{\beta} l(y) m(y) dy \right] e^{-rx} m(x) dx}{\int_{\alpha}^{\beta} \left[ \int_x^{\beta} e^{-r(y-x)} l(y) m(y) dy \right] e^{-rx} m(x) dx} \quad (2-3)$$

この分母が (2-2) 式の単純な変形であることは、容易に理解できよう。比  $Z$  が 1 にならないとしたら、それは分母の  $e^{-r(y-x)}$  が 1 でないことに帰着する。この積分範囲では  $y - x$  は常に正だから、 $r$  が正であれば (2-3) 式の分母は分子よりも小さく、 $r$  が負なら分母は分子より大きくなる。つまり  $Z$  は、 $r$  が正のとき 1 より大きく、 $r$  が負なら 1 より小さいという、ちょうど NRR と同じ性質を持つことが示される。もちろん  $r=0$  のとき、分子と分母は一致して  $Z=1$  となる。

$Z$  の自然対数を取り、 $f(r) = \ln E(Y)$ 、 $g(r) = \ln E(\Phi)$  とする。 $r=0$  のとき  $f(0) = g(0)$  であることから、 $\ln Z$  の  $r=0$  でのテイラー展開は、一次導関数までしか考えないとすると、

4) 以下で姉妹数はすべて出生数を意味し、生存している姉妹の数ではない。娘数その他についても同じである。

$$\ln Z \doteq rf'(0) - rg'(0) = r \frac{\int_a^\beta \int_x^\beta (y-x) l(y)m(y)m(x) dy dx}{\int_a^\beta \int_x^\beta l(y)m(y)m(x) dy dx} \quad (2-4)$$

ここで再生産年齢での母親の死亡を無視することにし、上式から  $l(y)$  を除いて変形する。

$$\frac{\ln Z}{r} \doteq \frac{\int_a^\beta \int_x^\beta (y-x) m(y)m(x) dy dx}{\int_a^\beta \int_x^\beta m(y)m(x) dy dx} \quad (2-5)$$

右辺 =  $\kappa$  とし、積分範囲を変え絶対値表記を用いれば、次のように書ける。

$$\kappa = \frac{\int_a^\beta \int_a^\beta |y-x| m(y)m(x) dy dx}{\int_a^\beta \int_a^\beta m(y)m(x) dy dx} \quad (2-6)$$

結局  $Z$  と  $r$  の関係は、この  $\kappa$  を用いて次のように表せる。

$$Z \doteq e^{r\kappa} \quad (2-7)$$

この関係に至るまでに、テイラー展開の二次微分以降の項を無視し、また母親は再生産期間中に死亡しないなど大胆な仮定をおいているため、Goldman は実際の推計に当っては直線回帰を用いることを勧めている。

$$\hat{r} = \frac{B}{\kappa} \ln Z \quad (2-8)$$

$B$  は回帰係数で、Goldman が Coale and Trussell のモデル出生スケジュール<sup>5)</sup>から求めたところでは  $B = .980$ 、 $K$  の代わりに姉妹数の標準偏差  $\sigma$  を用いる場合は、 $B = .921$  であった。

### III Wachter の議論

Wachter はマイクロ・シミュレーションを用いて Goldman の方法を評価し、ふたつの年齢階級人口の比を使う Keyfitz の方法<sup>6)</sup>と比べると、かなり精度が劣ることを示した。しかし本稿にとって興味があるのは、むしろ Wachter が提示した理論的問題の方である。

姉妹数の比と安定人口増加率との間に一定の関係があるという Goldman の論理には説得力があり、(2-3) 式を見ると確かに  $Z$  は NRR と同様の性質を持つように思われる。しかし別の考え方をすれば、姉の数と妹の数が異なるのはおかしい。というのは、姉が妹をひとり数えるとき、その妹は必ず姉をひとり数えるという対称性があるため、両者は一致するはずだからである。

本人が一人娘の場合、姉の数も妹の数も 0 で、当然両者は一致している。ふたり姉妹の長女から見て姉の数は 0、次女から見て 1 だから期待値は 0.5 人だが、妹の数も長女は 1、次女は 0 で期待値をとれば 0.5 人となる。同様にして 3 人姉妹なら期待値は 1 人、4 人姉妹なら 1.5 人と、どこまで行っても姉の数と妹の数は期待値において一致している。つまり全女子について調査すれば姉の数と妹の

5) Coale, Ansley J. and T. J. Trussell, "Model fertility schedules: variations in the age structure of childbearing in human populations", *Population Index*, No.40, 1974, pp.185-258.

6) Keyfitz, Nathan, *Applied Mathematical Demography, Second Edition*, New York, Springer-Verlag, 1985, pp.92-96.

数は一致し、 $Z$ は常に1で $r$ とは無関係のはずである。では、なぜGoldmanの論理が成り立つのだろうか。これがWachterの言う「姉妹数の謎 (Sisters' Riddle)」である。

当然の理由として、Goldmanの方法にサンプリング・バイアスがあり、全女子が期待値の計算に含まれていない、ということが考えられる。しかしバイアスがどのようなもので、いかにして増加率との関係を生じているのかが説明されなければ、謎を解いたことにならない。

Wachterによると、姉妹数の期待値には2種類のバイアスが存在し、これらが完全に相殺し合わないために $Z$ は1以外の値をとる。第一のバイアスは年齢の下限で、Goldmanの方法では $\beta - \alpha$ 歳以上の女子についてだけ期待値が計算される。この場合、姉は $\beta - \alpha$ 歳以上で妹数を報告できるが、その妹は $\beta - \alpha$ 歳に達していないため妹数を報告できない、という事態が生じる。つまり報告される姉の数が実際より少なくなる。この年齢下限を、一般に $b$ で表す。

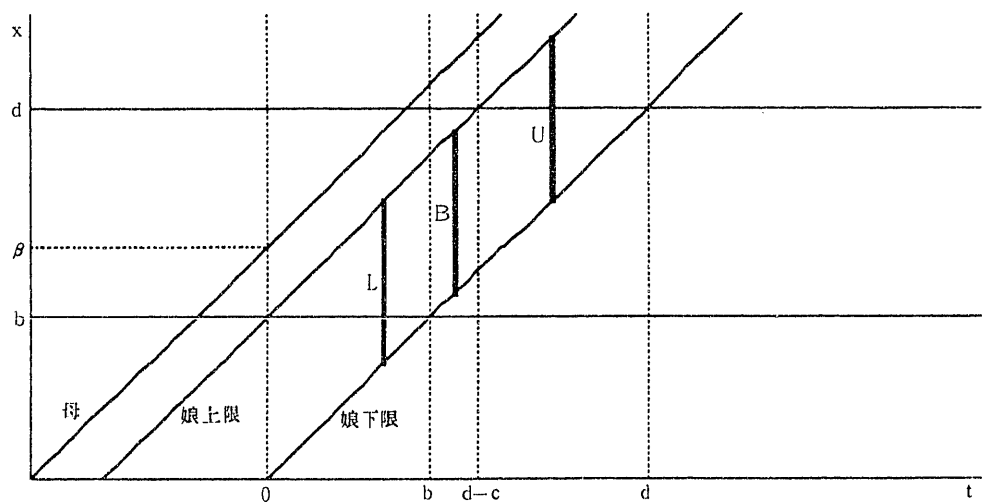
もうひとつのバイアスは、死亡に関するものである。姉の方が年長であることから、妹より先に死亡する確率が高い。この場合、生存している妹は姉の数を報告できるが、死亡している姉にとっての妹数は報告されない、という事態が生じる。つまり死亡は、年齢下限の存在とは逆に、報告される妹の数を少なくするというバイアスを生じる。Wachterは下限 $b$ に対し年齢の上限 $d$ を設定し、全女子がこの年齢で死亡するとした。

この $b$ 歳と $d$ 歳に狭まれた年齢階級が「エゴ・プール」であり、この年齢に属す女子だけが姉妹数の計算に含まれる。また、この年齢階級幅と再生産期間の長さ $c$ との間には、次の関係があると仮定する<sup>7)</sup>。

$$d - b > c \tag{3-1}$$

ひとつの母コーホートを選び、その娘の集合について考えよう。このような娘群 (Sorority) が若い時点では、図1のLのように、年少の部分が $b$ 歳未満でエゴ・プールからはずれている。これを「下またぎ (Lower Boundary)」状態と呼び、既に述べたように

図1



報告される姉の数が少なくなる。しかし再生産終了後 $b$ 年経つと、(3-1)の仮定より、図1のBのように娘群全体がエゴ・プールに含まれることになる。これを「包含 (Ballast)」状態と呼び、ここではバイアスは生じない。さらに年齢が高くなると、図1のUのように、娘群の年長の部分が死亡年齢 $d$ を越える。これが「上またぎ (Upper Boundary)」状態で、報告される妹の数が少なくなる。

これらの状態によって報告される姉妹数は異なるが、それを定式化するためにいくつか変数を準備

7) 自然な考え方では $b = c = \beta - \alpha$ だが、Wachterは $b$ と $c$ が異なり得る一般の場合について議論を展開している。

しておく必要がある。

$$p(u, v) = p(v, u) \quad \text{母 } \beta \text{ 歳時のエゴとその姉妹の年齢組合せの密度} \quad (3-2)$$

$$F(s) = \int_s^c \int_0^u p(u, v) dv du \quad \text{母 } \beta \text{ 歳時に } s \text{ 歳以上のエゴの妹数} \quad (3-3)$$

$$H(s) = \int_s^c \int_u^c p(u, v) dv du \quad \text{母 } \beta \text{ 歳時に } s \text{ 歳以上のエゴの姉数} \quad (3-4)$$

$$g = F(0) = H(0) \quad \text{娘群全体についての姉数=妹数} \quad (3-5)$$

これらを用いて、エゴが属す娘群の状態別に、報告される姉妹数は次のようになる。

	妹	姉	
下またぎ (L)	$\int_{b-c}^b F(b-t)e^{-rt} dt$	$\int_{b-c}^b H(b-t)e^{-rt} dt$	(3-6a)

包含 (B)	$\int_b^{d-c} g e^{-rt} dt$	$\int_b^{d-c} g e^{-rt} dt$	(3-6b)
--------	-----------------------------	-----------------------------	--------

上またぎ (U)	$\int_{d-c}^d [g - F(b-t)] e^{-rt} dt$	$\int_{d-c}^d [g - H(b-t)] e^{-rt} dt$	(3-6c)
----------	--	--	--------

このままでは複雑すぎるので、さらに次の変数を追加する。

$$f = \frac{1}{g} \int_0^c \int_0^u \frac{u}{c} p(u, v) dv du \quad \text{妹がいる ego の母 } \beta \text{ 歳時平均年齢 } \div c \quad (3-7)$$

$$h = \frac{1}{g} \int_0^c \int_u^c \frac{u}{c} p(u, v) dv du \quad \text{姉がいる ego の母 } \beta \text{ 歳時平均年齢 } \div c \quad (3-8)$$

$$L = g \int_{b-c}^b e^{-rt} dt \quad \text{バイアスがない場合の妹数=姉数：下またぎ} \quad (3-9)$$

$$B = g \int_b^{d-c} e^{-rt} dt \quad \text{バイアスがない場合の妹数=姉数：包含} \quad (3-10)$$

$$U = g \int_{d-c}^d e^{-rt} dt \quad \text{バイアスがない場合の妹数=姉数：上またぎ} \quad (3-11)$$

これらを用い、また  $e^{-rt}$  の部分はその区間の平均で代用すると、(3-6) 各式は次のように書き直すことができる。

	妹	姉	
下またぎ (L)	$Lf$	$Lh$	(3-12a)

包含 (B)	$B$	$B$	(3-12b)
--------	-----	-----	---------

上またぎ (U)	$U(1-f)$	$U(1-h)$	(3-12c)
----------	----------	----------	---------

ここから姉妹比  $Z$  を求めると、次のようになる。

$$Z = \frac{Lf + B + U(1-f)}{Lh + B + U(1-h)} = \frac{f}{h} + \frac{1-f}{1-h+h \frac{B+L}{B+U}} \quad (3-13)$$

このうち  $f$  と  $h$  は、娘群がどの状態にあるかに関わりなく妹（姉）がいるエゴの年齢  $\div c$  なので、 $b$  と  $d$  と関係がない。残る「包含+下またぎ」と「包含+上またぎ」の比も、実は  $b$ 、 $d$  から自由であることが示される。

$$\frac{B+L}{B+U} = \frac{\int_{b-c}^{d-c} e^{-rt} dt}{\int_b^d e^{-rt} dt} = e^{rc} \quad (3-14)$$

この式から、 $r$  が正であれば  $L > U$ 、負であれば  $L < U$  となることは明らかである。 $r=0$  であれば、もちろん  $L=U$  でなければならない。また (3-13) より、 $Z=1$  となるのは  $L=U$  のときしかないので分かるから、 $Z$  は確かに NRR と同様の性質を持つ。つまり  $r$  が 0 でなければ、下またぎ娘群の規模  $L$  と上またぎ娘群の規模  $U$  は異なり、このために  $Z$  は 1 以外の値をとるのである。

なお、(3-14) を (3-13) に戻して、Wachter による  $r$  の推計値は次のようになる。

$$\hat{r} = \frac{1}{c} \ln \left[ 1 + \frac{Z-1}{f-hZ} \right] \quad (3-15)$$

以上が Wachter の議論の格子であり、 $r$  と  $Z$  の関係が、 $L$  と  $U$  の大小関係を通じて生じていることが明快に示されている。しかしこの結論に至るまでの議論が非常に難解で、登場する変数が多い上に、中には  $f$  や  $h$  のように何を表しているのか理解し難いものも含まれている。

また、死亡年齢が  $d$  に固定されていることにも疑問を感じる。これは生命表とは何の関係もなしに設定され、娘群の全員がこの年齢で死亡するとされる。最終的には (3-14) で消去されるので、 $d$  の具体的な値は結論に影響しないものの、固定死亡年齢のような強い仮定を使わずにすればその方が好ましい。

Goldman や Wachter が、年齢構造のみを持つ単性の安定人口モデルに依拠して理論を展開しているのには、十分な理由がある。そうした Lotka モデルは、形式人口学における最も標準的でなじみ深いモデルで、これに依拠した議論は理解されやすい。しかし一方で、年齢構造しか持たない Lotka モデルは、親族の人口学的モデルに重大な制約を課している。特に「姉妹数の謎」は、姉妹数別分布や出生順位別分布のようなパリティ構造と強く関連する問題であり、この情報を欠く Lotka モデル上で議論を展開するのはかなり無理があるといえる。

一方、パリティ構造を導入した安定人口モデルに依拠すると、増加率が Lotka モデルと異なってしまい、実用上は好ましくない。そのような年齢=パリティ・モデルの増加率は、Lotka の内的自然増加率に比べてなじみが薄く、解釈が難しいため実用性に乏しいのである。

しかし理論的な関心として、年齢=パリティ・モデルに依拠して「姉妹数の謎」を考察することには、なお意義がある。そのようなアプローチは、情報不足の Lotka モデルに依拠した議論よりも明快に、より現実に近い関係を導き得るからである。次節では、この年齢=パリティ・モデルによる「姉妹数への謎」への接近を試みることにする。

#### IV 年齢=パリティ・モデルにおける姉妹数の謎

##### 1. 準備

パリティ構造を持つ安定人口モデルは別に新しいものではなく、数学的な考察も数値例の計算も既

に何度か行なわれている<sup>8)</sup>、このモデルの内的自然増加率が一般に Lotka モデルのそれと異なるのは、純再生産関数が異なるからである。Lotka モデルの  $l(x)m(x)$  に対し、年齢パリティ・モデルの純再生産関数は、次のようにパリティ別生存数と出生率の積和として定義される<sup>9)</sup>。

$$\sum_{i=1}^I l_{x,i-1} m_{x,i-1} \quad (4-1)$$

ここで  $I$  は最大パリティ、 $l_{x,i-1}$  は年齢別・パリティ別生命表人口 (基数=1)、 $m_{x,i-1}$  は年齢別・パリティ別出生率である。

任意の女子について、確定した娘数の期待値を  $N$  とする。これはつまり NRR で、年齢=パリティ・モデルの場合、次のように定義される。

$$N = \int_a^\beta \left( \sum_{i=1}^I l_{x,i-1} m_{x,i-1} \right) dx \quad (4-2)$$

娘数  $N$  と、姉妹数の期待値  $S$  の間には、次の関係があることが知られている<sup>10)</sup>。

$$S = \frac{\sigma^2}{N} + N - 1 \quad (4-3)$$

ただし  $\sigma^2$  は娘数の分散である。Lotka モデルでは情報不足のため、娘数の分散を求めることができず、上の関係を正しく表現できない。そのためポワソン分布を仮定することになり、結局娘数と姉妹数の期待値は一致することになる<sup>11)</sup>。これに対し年齢=パリティ・モデルでは、観測値にもとづく年齢別・パリティ別生命表が用意されており、それによって分散はもちろん、必要であればより高次のモメントも計算できる。このため「姉妹数の謎」に対しても、Lotka モデルに比べより現実的・合理的に接近し得る。

## 2. 特定母コーホートの娘群

まず Wachter と同様に、ひとつの母コーホートの娘群について考える。(4-2) 式は、出生順位別の娘数の合計が NRR であることを示している。従って娘群において第  $i$  子が占める比率  $O_i$  を求めるには、次のようにすればよい。

$$O_i = \frac{1}{N} \int_a^\beta l_{x,i-1} m_{x,i-1} dx \quad (4-4)$$

8) Keyfitz, Nathan, *Introduction to the Mathematics of Population*, Addison-Wesley, Reading, MA, 1968, pp.334-335; Oechsli, Frank WM., "A population model based on a life table that includes marriage and parity", *Theoretical Population Biology*, No.7, 1975, pp.229-245; Das Gupta, Prithwis, "Age-parity-nuptiality-specific stable population model that recognizes birth to single women", *Journal of American Statistical Association*, Vol.71, No.354, 1976, pp.308-314; Inaba, Hisashi, *Duration-Dependent Multistate Population Dynamics*, Working Paper Series, No.9, Institute of Population Problems, Tokyo, 1992; 鈴木 (1992b, 前掲注1) など。

9) Keyfitz (1968, 前掲注8), p.334.

10) 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; 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, p.26; Keyfitz, Nathan, *Applied Mathematical Demography, Second Edition*, Springer-Verlag, New York, 1985, p.287.

11) 鈴木 (1992b, 前掲注1, pp.2-3).

確定した姉妹数別分布を  $B_i$  とする。  $i$  人姉妹が  $B_i$  組あるということは、第1子から第  $i$  子までがそれぞれ  $B_i$  人づついるということに他ならない。従って出生順位と姉妹数の間には、次の関係が成り立つ。

$$O_i = \frac{\sum_{j=i}^I B_j}{\sum_{j=1}^I j B_j} = \frac{1}{N} \sum_{j=i}^I B_j \quad (4-5)$$

(4-4) および (4-5) 式より、姉妹数別分布を  $l_{x,i-1}$  と  $m_{x,i-1}$  で表すと次のようになる。

$$\int_a^\beta l_{x,i-1} m_{x,i-1} dx = \sum_{j=i}^I B_j$$

ここで  $i$  を重さとする加重和を考える。

$$\begin{aligned} \int_a^\beta \left( \sum_{i=1}^I i l_{x,i-1} m_{x,i-1} \right) dx &= \sum_{i=1}^I i \sum_{j=i}^I B_j \\ &= \sum_{i=1}^I \frac{i(i+1)}{2} B_i \\ &= \frac{1}{2} (\sigma^2 + N^2 + N) \end{aligned}$$

両辺を  $2/N$  倍して (4-3) 式を適用すれば、 $S$  を  $l_{x,i-1}$  と  $m_{x,i-1}$  で表すことができる。

$$S = 2 \left[ \int_a^\beta \frac{1}{N} \left( \sum_{i=1}^I i l_{x,i-1} m_{x,i-1} \right) dx - 1 \right] \quad (4-6)$$

これで年齢＝パリティ・モデルにおける姉妹数の期待値が得られた。一方、姉の数の期待値は (4-4) 式の出生順位別分布から直接求めることができる。なぜならある女子が  $i$  番目の娘であることは、すなわち姉が  $i-1$  人いることを意味するからである。従ってこの場合の姉数の期待値  $\Phi_1$  は、次のようにして得られる。

$$\begin{aligned} \Phi_1 &= \sum_{i=1}^I (i-1) O_i \\ &= \int_a^\beta \frac{1}{N} \left( \sum_{i=1}^I i l_{x,i-1} m_{x,i-1} \right) dx - 1 \end{aligned} \quad (4-7)$$

(4-6) と (4-7) 式を比較すると、姉の数が姉妹数のちょうど半分になることがわかる。つまり特定母コーホートの娘群についてみた場合、姉の数と妹の数は等しい。従ってこの場合には、Goldman の  $Z$  は常に1で  $r$  とは何の関係もない。これが「姉妹数の謎」の一方の面である。

### 3. 特定娘コーホート

次に母コーホートを特定するのではなく、ひとつの娘コーホートを直接とり出して姉妹数の期待値がどうなるか考えることにする。この娘コーホートが  $\beta - \alpha$  歳に達していれば姉妹数は確定しており、その期待値は  $S$  である。というのは、確定した娘数の期待値  $N$  と同じく、 $S$  は任意の女子についての確定した姉妹数の期待値であり、姉妹数が確定している女子を選ぶ限りその選び方には関係がない。実際、安定人口の仮定から、どの母親も平均  $N$  と分散  $\sigma^2$  を持つ同じ出生スケジュールに従っており、このため任意の女子の確定姉妹数は  $S$  で一定となるわけである。

娘コーホートの出生時の規模を  $K$  とする。安定人口の仮定より、そのとき  $x$  歳だった母コーホートの出生時規模は  $Ke^{-rx}$  である。従って当該娘コーホートのうち、この母コーホートから生れた第  $i$  子の数は  $Ke^{-rx} l_{x,i-1} m_{x,i-1}$  となり、 $x$  について積分すれば第  $i$  子全体の人数が得られる。



$$\int_a^\beta K e^{-rx} l_{x,i-1} m_{x,i-1} dx$$

これを  $i$  について合計した結果は、当然娘コーホート全体の規模  $K$  である。従って出生順位  $i$  の分布を得るには、上式を  $K$  で割ればよい。

$$O'_i = \int_a^\beta e^{-rx} l_{x,i-1} m_{x,i-1} dx \quad (4-8)$$

前述の娘群の場合と同様、特定娘コーホートの姉の数は、これから直接求めることができる。

$$\begin{aligned} \Phi_2 &= \sum_{i=1}^I (i-1) O'_i \\ &= \int_a^\beta e^{-rx} \left( \sum_{i=1}^I i l_{x,i-1} m_{x,i-1} \right) dx - 1 \end{aligned} \quad (4-9)$$

$\Phi_2$  は  $\Phi_1$  と異なり、 $S/2$  に一致するとは限らない。つまり姉数と妹数の関係は、増加率  $r$  に依存する。これが「姉妹数の謎」のもう一方の面である。

死亡が出生順位と独立であるとすれば、(4-8) および (4-9) 式は、 $\beta - \alpha$  歳以上で姉妹数が確定している全コーホートについて適用できよう。従って姉妹比  $Z$  は確かに  $r$  と連動することが、固定死亡年齢のような強い仮定を用いずに示されたことになる。残された課題は、 $Z$  と  $r$  の関係が Goldman が示したとおりのものであることを証明することである。

#### 4. 姉妹比と増加率の関係

Lotka モデルに依拠した場合、Goldman が (2-3) 式で証明したように  $r > 0$  であれば  $Z > 1$ 、つまり妹数の期待値の方が姉より多く、 $r < 0$  なら  $Z < 1$ 、つまり姉の方が多いうことがいえる。ところで増加率が正のとき姉の方が少ないということは、本稿の表記法では  $\Phi_2$  が  $S/2$  より小さいと主張することに等しい。逆に減少中の人口で姉の方が多いうことは、 $\Phi_2$  が  $S/2$  より大きいことを意味する。ところが (4-6) および (4-7) 式より、 $\Phi_1$  は常に  $S/2$  に等しいことが示されるから、姉妹数の大小関係は  $\Phi_2$  と  $\Phi_1$  の大小関係に置換えることができる。すなわち証明すべき関係は、次のとおりである。

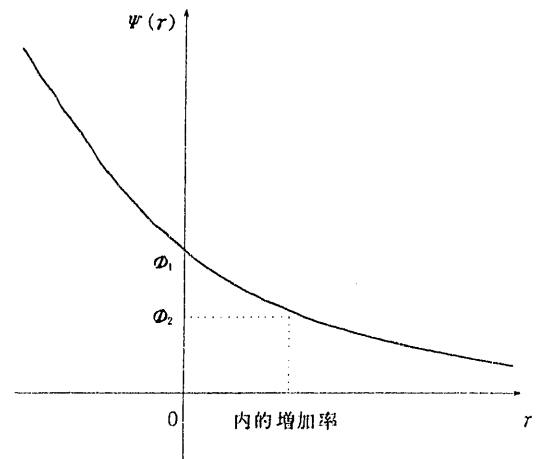
$$\begin{aligned} r > 0 \quad \text{のとき} \quad \Phi_2 &< \Phi_1 \\ r = 0 \quad \text{のとき} \quad \Phi_2 &= \Phi_1 \\ r < 0 \quad \text{のとき} \quad \Phi_2 &> \Phi_1 \end{aligned}$$

これを証明するためには、次の関数を考えるのが便利である。内的増加率としての  $r$  は、この分母の特性方程式  $= 1$  とおいたときの実根として単一の値に定まるが、ここでは一時的に  $r$  をランダム変数とみなし、自由な値をとり得るものとする。

$$\psi(r) = \frac{\sum_{i=1}^I i \left( \int_a^\beta e^{-rx} l_{x,i-1} m_{x,i-1} dx \right)}{\sum_{i=1}^I \left( \int_a^\beta e^{-rx} l_{x,i-1} m_{x,i-1} dx \right)} - 1 \quad (4-10)$$

$r=0$ であれば分母は $N$ となり、(4-7)式より $\Psi(r)$ は $\Phi_1$ に一致する。また $r$ が内的増加率であれば分母は1であり、(4-9)式より $\Psi(r)$ は $\Phi_2$ に一致する。もし図2のように $\Psi(r)$ が $r$ に関し単調減少関数であれば、予想された関係が成り立つことが理解できよう。従って、予想された関係を証明するためには、 $\Psi(r)$ の一次導関数が負であることを示せばよい。

図2



この証明には、出生に関する2種類の指標をあらかじめ用意しておくことと便利である。ひとつは出生順位別出生数の相対頻度である。

$$b_i = \int_a^\beta e^{-rx} l_{x, i-1} m_{x, i-1} dx$$

この式は(4-8)式の $O_i$ と全く同じ形だが、ここでは $r$ が内的自然増加率のとき $b_i$ は $O_i$ に一致すると考えることにする。つまり $b_i$ を $i$ について合計しても、一般には1にならない。従って $b_i$ は分布比率とは言えないが、何らかの相対頻度を表すことには変わりはない。

もうひとつは出生順位別にみた母親の出生年齢に関するものである。

$$\kappa_i = \int_a^\beta x e^{-rx} l_{x, i-1} m_{x, i-1} dx$$

この $\kappa_i$ を $b_i$ で割れば、出生順位別の平均出生年齢になる。 $b_i$ と $\kappa_i$ を用いて、 $\Psi(r)$ の導関数は次のように表せる。

$$\frac{d\Psi(r)}{dr} = \frac{\left(\sum_{i=1}^I i b_i\right) \left(\sum_{i=1}^I \kappa_i\right) - \left(\sum_{i=1}^I b_i\right) \left(\sum_{i=1}^I i \kappa_i\right)}{\left(\sum_{i=1}^I b_i\right)^2}$$

分母は問題なく正だから、この導関数の正負を知るには分子について調べればよい。これを直接展開してみることにする。

$$\begin{aligned}
 & \begin{bmatrix} b_1\kappa_1 + b_1\kappa_2 + b_1\kappa_3 + \dots + b_1\kappa_I \\ + 2b_2\kappa_1 + 2b_2\kappa_2 + 2b_2\kappa_3 + \dots + 2b_2\kappa_I \\ + 3b_3\kappa_1 + 3b_3\kappa_2 + 3b_3\kappa_3 + \dots + 3b_3\kappa_I \\ \vdots \\ + Ib_I\kappa_1 + Ib_I\kappa_2 + Ib_I\kappa_3 + \dots + Ib_I\kappa_I \end{bmatrix} - \begin{bmatrix} b_1\kappa_1 + 2b_1\kappa_2 + 3b_1\kappa_3 + \dots + Ib_1\kappa_I \\ + b_2\kappa_1 + 2b_2\kappa_2 + 3b_2\kappa_3 + \dots + Ib_2\kappa_I \\ + b_3\kappa_1 + 2b_3\kappa_2 + 3b_3\kappa_3 + \dots + Ib_3\kappa_I \\ \vdots \\ + b_I\kappa_1 + 2b_I\kappa_2 + 3b_I\kappa_3 + \dots + Ib_I\kappa_I \end{bmatrix} \\
 = & \begin{matrix} 0 - & b_1\kappa_2 - & 2 & b_1\kappa_3 - & 3 & b_1\kappa_4 - \dots - & (I-1) & b_1\kappa_I \\ + & b_2\kappa_1 + & 0 - & b_2\kappa_3 - & 2 & b_2\kappa_4 - \dots - & (I-2) & b_2\kappa_I \\ + & 2 & b_3\kappa_1 + & b_3\kappa_2 + & 0 - & b_3\kappa_4 - \dots - & (I-3) & b_3\kappa_I \\ + & 3 & b_I\kappa_1 + & 2 & b_I\kappa_2 + & b_I\kappa_3 + & 0 - \dots - & (I-4) & b_I\kappa_I \\ + & 4 & b_5\kappa_1 + & 3 & b_5\kappa_2 + & 2 & b_5\kappa_3 + & b_5\kappa_4 + \dots - & (I-5) & b_5\kappa_I \\ & & & & & & & & & \vdots \\ + & (I-1)b_I\kappa_1 + & (I-2)b_I\kappa_2 + & (I-3)b_I\kappa_3 + & (I-4)b_I\kappa_4 + \dots + & & & & & 0 \end{matrix}
 \end{aligned}$$

$$\begin{aligned}
&= (b_2\kappa_1 - b_1\kappa_2) + 2(b_3\kappa_1 - b_1\kappa_3) + 3(b_4\kappa_1 - b_1\kappa_4) + \cdots + (I-1)(b_I\kappa_1 - b_1\kappa_I) \\
&\quad + (b_3\kappa_1 - b_1\kappa_3) + 2(b_4\kappa_1 - b_1\kappa_4) + \cdots + (I-2)(b_I\kappa_1 - b_1\kappa_I) \\
&\quad \quad + (b_4\kappa_1 - b_1\kappa_4) + \cdots + (I-3)(b_I\kappa_1 - b_1\kappa_I) \\
&\quad \quad \quad \vdots \\
&\quad \quad \quad \quad + (b_I\kappa_1 - b_1\kappa_I) \\
&= \sum_{i=2}^I \sum_{j=1}^{i-1} (i-j) (b_i\kappa_j - b_j\kappa_i) \\
&= \sum_{i=2}^I \sum_{j=1}^{i-1} (i-j) b_i b_j \left( \frac{\kappa_j}{b_j} - \frac{\kappa_i}{b_i} \right)
\end{aligned}$$

この式で  $j$  は  $i-1$  までの加算だから  $i-j$  は常に正で、 $b_i b_j$  も問題なく正である。 $\kappa_i/b_i$  は第  $i$  子の平均出生年齢だから、導関数が負であるためには、「出生時の母親の平均年齢は、早い出生順位の方が若い」という関係が成り立てばよい。

この関係は自然なもので、実際、第 2 子の平均出生年齢の方が第 1 子より若くなるためには、出生年齢分布が第 1 子と第 2 子で大幅に異なっていなければならない。まして第 3 子や第 4 子の平均出生年齢が第 1 子より若いということは、まずありそうにない。つまり、仮に全てではなくても、ほとんどの出生順位組合せについて、早い順位の方が平均出生年齢は若いと考えてさしつかえない。このために上式の最後のカッコ内はほとんどが負となり、従って  $\Psi(r)$  の導関数もほぼ間違いなく負であるといえる。

以上が増加率と姉妹比について予想された関係の形式的な証明であるが、この関係が成り立つ理由はより直感的に理解することができる。つまり増加率が正の場合、静止人口の場合に比べて若い母親が多くなり、このため第 1 子、第 2 子といった早い出生順位の娘の比率が高くなる。これは妹数の期待値が大きくなることを意味する。逆に負の増加率は第 4 子、第 5 子のような遅い出生順位の娘が多くなるので、姉妹数の期待値が大きくなる。つまり人口増加率は、再生産年齢女子の年齢分布を通じて姉妹比に影響を及ぼしているのである。

## V 結語

本稿では、パリティ構造を加えた安定人口モデルにもとづき、姉妹数の期待値の比と増加率がどのように関連しているかを示した。先行研究が依拠した Lotka モデルと異なり、年齢＝パリティ・モデルは必要な情報を含んでいるため、さほど強い仮定を用いずにすんだ。すなわち、Goldman のように再生産年齢の母娘の死亡を無視したり、Wachter のように固定死亡年齢を設定するようなことはしていない。

結局、姉妹比が増加率と連動するのは、特定娘コーホートについて母コーホートの規模が  $r$  によって変化しているからである。そのため特定娘コーホートの出生順位分布は、(4-8) 式のように  $r$  の影響を受け、これが増加率と姉妹比の関係を生み出している。

具体的には、正の増加率は若い母親へのバイアスを通じて早い出生順位の娘の比率を高める効果を持ち、逆に負の増加率は遅い出生順位の娘を増やす。このため増加中の人口では妹の数、減少中の人口では姉の数の方が多くなる。

しかしこの推論は、「年長の母親は若い母親に比べ、出生順位の遅い娘を多く生む」という仮定に依存している。実際、関数  $\Psi(r)$  を用いた証明では、出生順位が遅いほど平均出生年齢が高いことが、予想された関係が成り立つ条件であることを示した。

従って出生行動が極めて特異でこの条件が満たされず、増加率と姉妹比の関係が成り立たない状況を、少なくとも想像してみることはできる。単純な例として、最大出生数が2人で、第1子と第2子の出生年齢分布が非常に異なっている場合を考えよう。たとえば、若いうちに第1子を生むことができた母親は、すぐに第2子も生むものとする。しかしこの初回のチャンスを逃した母親は、次のチャンスまで長期間待たねばならず、しかもその場合は1子しか生めないとしよう。このとき、第1子出生年齢は双峰型の分布になり、平均は第2子出生年齢の平均より高くなり得る。従って年長の母親ほど第1子より第2子を生むとはいえず、増加率と姉妹比の関係は予想とは異なるものとなる。

もちろん現実には、このような出生スケジュールはありそうにないので、増加中の人口で姉の期待値の方が大きいというようなことは考えなくてよい。しかし年齢=パリティ・モデルでは、出生順位別の出生年齢分布が観測値から与えられ、ある程度自由に動き得るという点は重要である。つまり出生順位別の構造を持たないLotkaモデルと異なり、年齢=パリティ・モデルでは、姉妹比 $Z$ の同じ値に対応する出生スケジュールはいくらでも存在し得る。このためGoldmanの(2-8)式やWachterの(3-15)式のような単純な形で $r$ の推計値を与えることは、年齢=パリティ・モデルの場合は不可能だろう。

いずれにせよ年齢=パリティ・モデルの $r$ はLotkaの内的自然増加率と一致しないので、仮に推計できたとしてもあまり意味がない。そもそも本稿の関心は、間接推計の実用的な方法論ではなく、増加率と姉妹数の関係がいかにして生じるかを理論的に考察することにある。

従来の親族数頻度に関する人口学的モデルは、ほとんどがLotkaの安定人口の範囲にとどまっていた。しかしLotkaモデルは情報不足のため、特に傍系親族を扱う場合に限界が多いことは事実である。本稿は、安定人口モデルにパリティを導入することによって、こうした限界のひとつを破る試みだった。今後は、パリティだけでなく結婚行動を含むモデルや両性モデル等を応用して、親族の人口学理論をいかに拡張し整備していくかが課題となる。

[付記] 本稿に示されたアイデアをまとめるに当たって、本研究所の稲葉寿主任研究官から貴重な示唆をいただいた。また、日本人口学会大会で報告した際には、多くの方から有益な助言を受けた。記して感謝したい。

## The Sisters' Riddle in Age-Parity-Structured Stable Population

Toru SUZUKI

N. Goldman (1978) developed a method to estimate growth rates from sibling data. She showed that the expected number of younger sister exceeds that of older when a population is growing, and the relation reverses in a decreasing population. A formula was derived to estimate  $r$  from ratio of the number of younger to older sisters.

K. Wachter (1980) evaluated Goldman's method and examined a fundamental question. When ego counts an older sister, that sister always counts ego as a younger sister. Because of this symmetry, the expected number of older sisters must be equal to younger sisters. Why the ratio of younger to older can fluctuate

in accordance with  $r$  instead of being unity? This is the “Sisters’ Riddle” presented by Wachter.

Both Goldman and Wachter based on Lotka’s age-structured stable population. However, the lack of parity structure in this model caused certain difficulties, and they had to set strong assumptions such as ignorance of death.

This article is an attempt to examine how the ratio of sisters is related with growth rate in age-parity-structured population model. Use of parity specific rates allows the expected number of sisters  $S$  which is free from Poisson distribution.

$$S = 2 \left[ \int_a^\beta \frac{1}{N} \left( \sum_{i=1}^I i l_{x,i-1} m_{x,i-1} \right) dx - 1 \right] \quad (1)$$

In above,  $N$  is the Net Reproduction Rate,  $I$  is the last parity,  $l_{x,i-1}$  is parity specific survivorship function, and  $m_{x,i-1}$  is parity specific fertility rate.

The age-parity model also enables us to obtain birth order distribution. If we consider all daughters of one mother cohort, the proportion of  $i$ -th daughter in this group is given as follows.

$$O_i = \frac{1}{N} \int_a^\beta l_{x,i-1} m_{x,i-1} dx \quad (2)$$

It is easy to calculate the expected number of older sisters from this distribution. The result is exactly a half of  $S$ , which means the number of older sisters is always same to younger sisters. This is one side of the Sisters’ Riddle.

The situation changes when we directly inspect one daughter cohort. Birth order distribution in a cohort apparently depends on the growth rate.

$$O'_i = \int_a^\beta e^{-rx} l_{x,i-1} m_{x,i-1} dx \quad (3)$$

In this case, ratio of older and younger sisters changes with  $r$ . This is the other side of the Sisters’ Riddle.

Deviation of  $O'_i$  from  $O_i$  results from growth in size of mother cohorts. Positive  $r$  causes a bias to young mothers. If younger mothers tend to bear babies of earlier birth order, which is a natural assumption, the number of younger sisters exceeds older sisters in a growing population.

Thus the relationship between ratio of sisters and growth rate depends on the fertility schedules by birth order. Though the condition which affirms the anticipated relationship is natural, it is not logically true as is in Lotka’s model.