

平成22年度修士論文

家系図ネットワークの構造解析

大阪府立大学大学院 工学研究科
電子・数物系専攻 数理工学分野
数理物理講座 非線形力学研究グループ

学籍番号 2090103041

堀内陽介

2011年2月

目次

第1章	序論	1
1.1	問題設定	1
第2章	先行研究	5
2.1	Derridaらのモデル	5
2.2	Plaらのモデル	7
2.3	競走馬の家系図	9
第3章	競走馬データの再解析	15
3.1	性の数比	15
3.2	親子年齢差	19
3.3	平均出産数	20
3.4	仔数分布	20
第4章	モデル	23
4.1	性数比可変モデル	23
4.2	性比-仔数分布可変モデル	24
4.3	仔数を割り振る方法	25
4.4	木の作成のアルゴリズム	26
4.5	解析結果	28
4.6	非先祖率の理論的な見積もり	29
第5章	配偶システム	35
5.1	配偶システムの紹介	35
5.2	ペア作成のアルゴリズム	36
5.3	配偶システムごとの非先祖率の比較	38

第 6 章	まとめと今後の課題	41
6.1	まとめ	41
6.2	今後の課題	42
付 録 A	Pla らによる理論	47
付 録 B	関数作成法による森	51
付 録 C	Fisher による性比の議論	53

第1章 序論

1.1 問題設定

家系図とは親子関係にある生物個体を線で結んだ図である。一口に家系図といっても、ある個体の子孫方向のみに注目した図や父子のつながりのみに注目した図など、研究の目的に応じて見方を変えられるほど家系図は多様かつ複雑に構成されている。まず、本論文で用いる用語の定義を表 1.1 に示す。本研究では家系図の中でも主個体の祖先に注目し、一本の木のネットワーク構造を解析した。

表 1.1: 本論文に登場する用語の説明

用語	説明
森	親子関係にある個体同士を線で結んだもの
主個体	森の中で着目する任意の個体
木	主個体の直接の祖先だけを取り出したもの

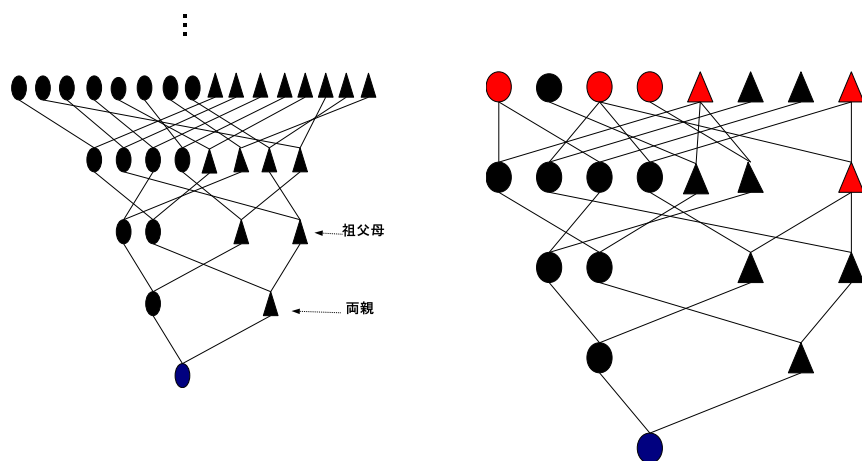


図 1.1: 主個体とその木の例。●はオス (男)、▲はメス (女) を表し、青印は主個体を表す。左: 単純な木の例。右: 複数の役割を持つ個体を含む木の例。赤色で示されたものは役割の重なる個体を表す。

有性生殖をおこなう生物における主個体の祖先数は、図 1.1 左のように親が 2、祖父母が 4...と 2 のべき乗で増加していく。つまり主個体の G 世代前の祖先数を N_G とすると、

$$N_G = 2^G, \quad (1.1)$$

となる。ヒトを例に考えると、30 世代前 (1 世代を 30 年と考えるとおよそ 900 年前) における現代の 1 個体の祖先数は $2^{30} \approx 10$ 億人存在していたことになる。もちろんこれは当時の人口の 3 億人 [1] を超えることになるので、式 (1.1) による祖先数の見積もりと実際の祖先数の増加は矛盾している。

実際の家系図では、図 1.1 右における赤印のように、ある個体が現代の複数個体の祖先となっている、といった役割の重複が存在していると考えられる。この重複した役割を持つ個体の数が世代を遡るにつれて増大していくことにより、主個体の祖先数の増加が抑制されるが、それに伴い木の構造も複雑になっていくと考えられる。

この構造に関する研究がいくつか報告されている。Derrida ら [2][3] は木の構造を特徴づける量をいくつか定義し、中立的なモデルを用いてそれらを解析した。Pla ら [4] は Derrida らのモデルを一般化したモデルを用いている。水口 [5][6]、西村 [7] や筆者 [8] が実データの家系図における木の構造を解析している。これらの中で Derrida らは、ある世代での総個体数のうち、主個体の祖先となっていない個体数の割合を非先祖率と定義した。すなわち最も若い世代を $G = 0$ と設定し、 G 世代前の非先祖率を $s(G)$ とする。 $s(G)$ は G 世代前での総個体数を $N(G)$ 、主個体の祖先となっていない個体数を $A(G)$ とすると、

$$s(G) = \frac{A(G)}{N(G)}, \quad (1.2)$$

である。式 (1.2) に従うと $s(G)$ の値が 0 であれば、 G 世代前の個体はすべて主個体の祖先である。逆に 1 に近い値となる時はほとんど全ての個体が主個体の祖先ではないことを意味している。このように非先祖率は家系図の中で、ある木がどの程度広がっているかを示しており、木の構造を特徴づける量の一つになっている。

具体的に式 (1.2) を使って図 1.2 のような木の $s(G)$ を世代ごとで計算してみると、 $s(0) = \frac{7}{8}$, $s(1) = \frac{6}{8}$, $s(2) = \frac{5}{8}$, ... と、 $s(G)$ が世代を遡るごとに変化していることがわかる。Derrida らは $s(G)$ の他にもウェイトという量も定義し、単純なモデルを用いてこれらを解析した。また Pla らは、Derrida らが用いたモデルを一般化し、非

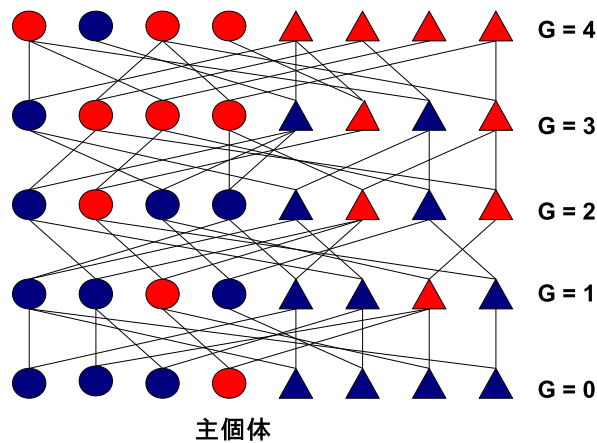


図 1.2: 森と木の例。未来方向は下向き。丸印はオス、三角印はメスを表し、親子関係にある個体同士を線で結んでいる。赤色で示されたものは主個体の木に属している個体を表す。

先祖率の解析を行った。そして西村や筆者、水口は実データとして競走馬の家系図を用いて、その非先祖率を実測したが、Derrida らの解析結果と実データを用いた解析結果は異なっていることが分かった。違いの原因は、性差や両親の組み合わせ方や設定する仔数分布の違いなど、モデルと競走馬の繁殖過程の間のさまざまな差異にあると考えられる。

本研究においても非先祖率を解析することで家系図の構造を見ていく。加えて、Derrida らのモデルに「性の数比及びオスとメスの仔数分布を変化させる」という修正を加えることで、Derrida らのモデルと競走馬の家系図をつなぐようなモデルを作ることができないかを考え、その条件を考慮したモデルを設定して森を構築し、非先祖率を解析した。

また、実在する生物集団は、1 個体のオスが 1 個体のメスとペアを作り仔を作る「一夫一妻制」や 1 個体のオスが複数個体のメスとペアを作り仔を作る「一夫多妻制」など多様なオスとメスのペア形成の形を取っている [9]。本研究ではそのペア形成のルールをまとめて「配偶システム」と名づけた。非先祖率のオスとメスの比率に対する依存性に注目し、配偶システムの違いによる森からそれぞれ非先祖率を解析、比較を行った。

本論文の構成は以下のとおりである。第 2 章で先行研究である Derrida らや Pla らのモデルによる解析結果、水口らによる競走馬の家系図の解析結果を紹介する。第

3 章で競走馬データの再解析について述べる。第 4 章では設定したモデルの紹介及びそれを用いて行った解析結果と理論的見積もりを述べる。第 5 章では配偶システムによる血縁関係の構造の違いを比較する。第 6 章では本研究のまとめと今後の課題について述べる。付録 A では Pla らの理論、付録 B では関数作成法の特徴、付録 C で Fisher 性比について紹介する。

第2章 先行研究

競走馬の家系図と Derrida らや Pla らが用いたモデルにより構成された森は、様々な点で構造が異なっている。以下の節でそれぞれを簡単に紹介し、比較する。なお、先行研究や本研究で用いられた非先祖率や世代個体数を互いに比較するために異なる記号を用いたので、それらを表 2.1 にまとめておく。

表 2.1: 各モデルやデータで使われる非先祖率と世代個体数

モデル, データ	非先祖率	世代個体数の記号
Derrida ら	$s(G)$	$N(G)$
Pla ら	$s_P(G)$	$N_P(G)$
競走馬データ	$s_T(G)$	$N_T(G)$
性比可変	$s_f(G)$	$N_f(G)$
性比-仔数分布可変	$s_h(G)$	$N_h(G)$

2.1 Derrida らのモデル

Derrida らは単純な家系図モデルを構築し、それを用いて森を作り非先祖率 $s(G)$ を解析した。まず森を構築する際に、 G 世代での総個体数 $N(G)$ を以下の式で計算した。

$$N(G) = \left(\frac{2}{m}\right)^G N(0). \quad (2.1)$$

式 (2.1) において、 m はペアの平均出産数である。式 (2.1) に従って、世代ごとに個体数を割り振り、個体集団を作る。このとき、オスとメスの数比は $1:1$ である。次に、各世代内で、オスとメスのペアをランダムに作る。孤立した個体は存在しない。 G 世代前の個体の両親を、 $G+1$ 世代前のペアからランダムに作る。このようにして作られた森は以下の 4 つの性質を有している。

1. 両親のペアはランダム

2 . 性別の数比はオス : メス = 1 : 1

3 . 世代の重複は無い

4 . 仔数は平均 m のポアソン分布に従う

例えば、 $N(0) = 10$ 、 $m = 2$ の時は図 2.1 のような森が構築される。平均 m のポアソン分布 $Q(k)$ は、

$$Q(k) = \frac{e^{-m} m^k}{k!}, \quad (2.2)$$

で表される。 $m = 2$ と $m = 2.5$ の時の $Q(k)$ を図 2.2 に示す。実例としては、1350 ~ 1986 年のイギリスにおけるヒトの夫婦が k 人の息子を持つ時の確率分布が平均 1.15 のポアソン分布になっており、図 2.3 に示されている。なお、各世代で親となるペアを作らず、父親母親をそれぞれ $G + 1$ 世代のオス・メスからランダムに選ぶという方法でも以下に述べる結果は変わらない。

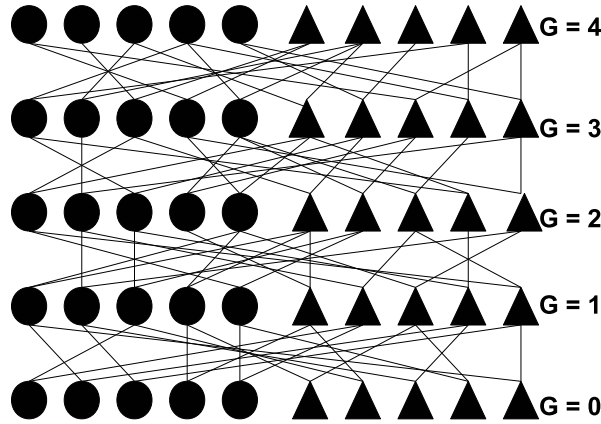


図 2.1: Derrida らのモデルによる森。未来方向は下向き。●はオス、▲はメスを表し、親子関係にある個体同士を線で結んでいる。 $N(0) = 10$, $m = 2$ 。

こうして構築された森において Derrida は、 $G = 0$ で任意に主個体を選んだ時の $s(G)$ が、

$$s(G + 1) = \exp [-m + ms(G)], \quad (2.3)$$

$$s(0) = 1 - \frac{1}{N(0)}, \quad (2.4)$$

で与えられることを理論的に示し、数値的に確かめた。式 (2.4) は $G = 0$ での非先祖率であり、主個体以外の個体は祖先でないことを意味する。式 (2.4) を式 (2.3) に

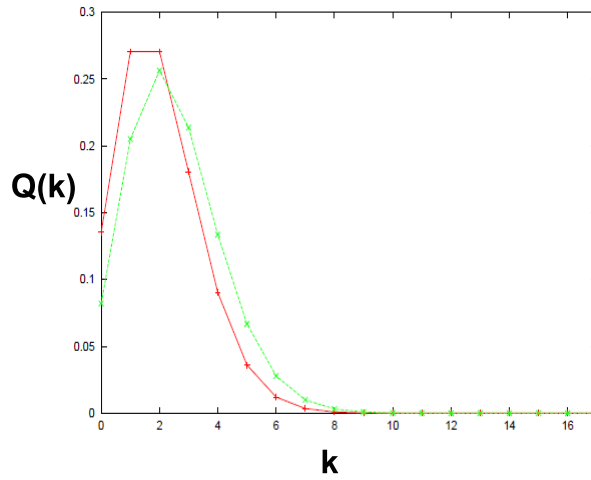


図 2.2: ポアソン分布 $Q(k)$ 。横軸は k 、縦軸は $Q(k)$ 。赤印及び赤点線は $m = 2$ 、緑印及び緑点線は $m = 2.5$ の時の $Q(k)$ を表す。

代入すると帰納的に $s(G)$ を求めることができる。例えば $m = 2.5$ の時の式 (2.3) の結果は図 2.4 の通りである。図 2.4 を見ると、 $s(G)$ は $G \rightarrow \infty$ で s^* に収束すると分かる。 $m = 2.5$ の時、 $s^* = 0.1073$ になる。この他にも Derrida らはウェイトを用いた解析や 2 本の木の関係の解析も行っているが、これらについては参考文献 [2][3] を参照されたい。

2.2 Pla らのモデル

Pla らも家系図モデルを考え、それを用いて森を構成し、非先祖率を解析した。まず森において、各世代の総個体数 $N_P(G)$ を N で固定する。次に総個体数に対するオスの数の割合を性比と呼び、それを f_P とする。世代の重複はなく、ペアの作り方は Derrida らのモデルと同じである。Pla らのモデルによる森は以下のような性質を有している。

- 1 . 両親のペアはランダム
- 2' . 性別の数比はオス : メス $= f_P : 1 - f_P$
- 3 . 世代の重複は無い
- 4' . 仔数はポアソン分布もしくは幾何分布に従う

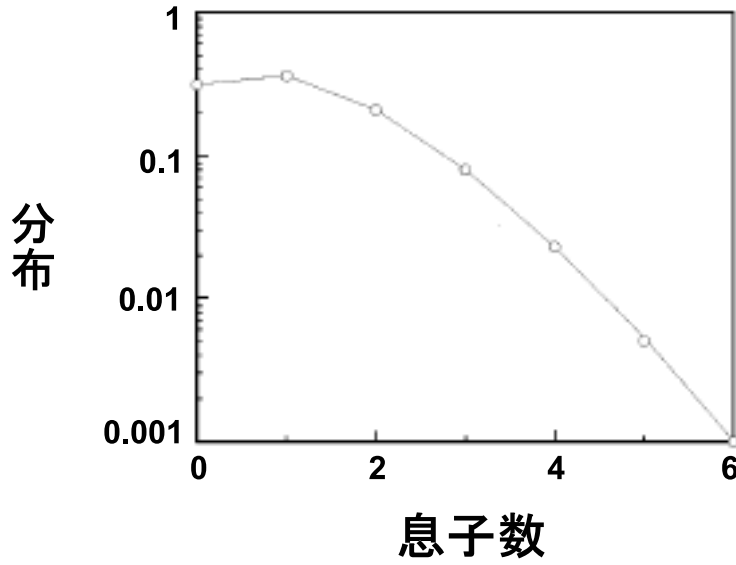


図 2.3: 1350 年～1986 年の期間におけるイギリスの夫婦が k 人の息子を持つ時の確率分布。丸印で表わしている。横軸は息子数、縦軸は分布。縦軸は $\log$ スケールである。実線は平均 1.15 のポアソン分布である。文献 [3] より転載。

具体的に Pla らのモデルによる森を構築すると、図 2.5 のようになる。また、幾何分布 $R(k)$ は、

$$R(k) = (1 - q)q^k, \quad (2.5)$$

で表わされる。仔数分布が式 (2.5) の時のペアの平均仔数を m' とすると、

$$m' = \frac{q}{1 - q}, \quad (2.6)$$

で表わされる。 $q = 0.71$ の時の $R(k)$ を図 2.6 に示す。

Pla らは Derrida 理論の $s(G)$ と対応するような母関数 $G(\lambda)$, $H(\mu)$ を定義し、解析した。この母関数の導出と解説は付録 A で詳しく述べる。 $\lambda, \mu \rightarrow \infty$ で $G(\infty) = G_0$, $H(\infty) = H_0$ とし、 G_0 と H_0 を数値的に解いた。 G_0 と H_0 はそれぞれオスとメスの非先祖率の収束値と対応している。図 2.7 は G_0 , H_0 の解析結果である。性比の値や仔数分布により G_0 , H_0 は異なる値を取る。 $f_P = 0.5$ の時は $G_0 = H_0$ となり、その値は s^* と同じ値を取る。このようにして Pla らは、前節で述べた Derrida らのモデルを性比と仔数分布について一般化し、家系図の構造は性比や仔数分布によって変化することを示した。

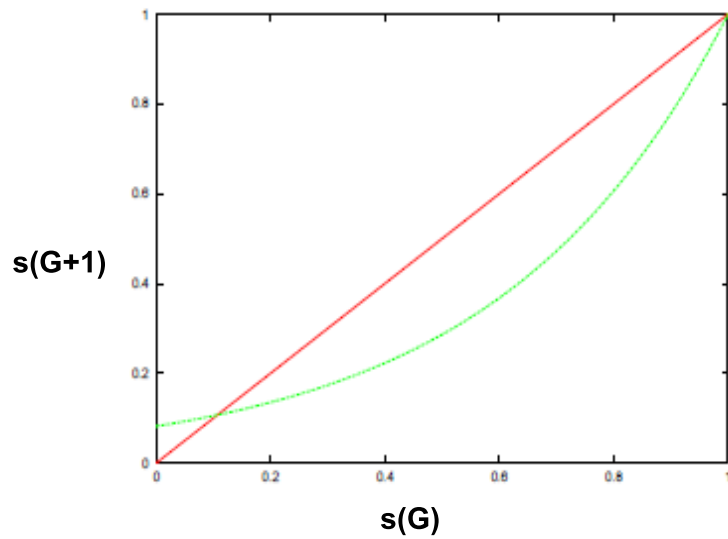


図 2.4: $s(G)$ の漸化式。 $m = 2.5$ 。横軸は $s(G)$ 、縦軸は $s(G+1)$ 。赤線は $s(G+1) = s(G)$ 、緑線は $s(G+1) = \exp[-m + ms(G)]$ を表す。この 2 関数の交点が s^* 。

2.3 競走馬の家系図

水口らは研究材料として豊富なデータ量が保存されており、長期間に渡る木を構成できることから競走馬のデータを利用した。競走馬のデータベース (Pedigree Online Thoroughbred Database) [11] 上から、各馬の名前・性別・生年・両親の名前を入手し、それらを使って家系図を作成した。また、競走馬のデータから 4349 頭をサンプリングし、競走馬の性比 p_T 、平均出産数 m_T 、オスの k 頭仔を持つ確率 $P_m(k)$ 、メスの k 頭仔を持つ確率 $P_f(k)$ 等、家系図を構築する際のパラメーターを測定した。その結果、性比は約 0.25、1 年あたりの個体増加率は約 1.02 であった。別に測定した平均出産年齢 12 年を 1 世代とすると、1 世代ごとの平均出産数 $m = 2.5$ と見積もられた。また、競走馬の繁殖過程には人の手加えられているため、競走馬特有の状況で森が構築されていると考えられる。競走馬の実データによる森は Derrida のモデルと以下の点で異なると考えられる。

- 1' . 両親のペアはランダムでない
- 2'' . 性別の数比はオス : メス = 1 : 3
- 3' . 世代の重複はある

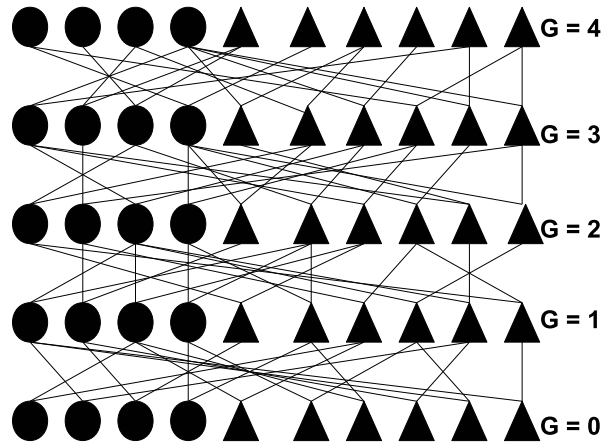


図 2.5: Pla らのモデルによりつくられた森。未来方向は下向き。●はオス、▲はメスを表し、親子関係にある個体同士を線で結んでいる。 $N_P(G) = N = 10$ で固定, $f_P = 0.4$ 。

4'' . 仔数はポアソン分布でない

一般的な競走馬の家系図の模式図を図 2.8 に示す。

水口は競走馬の実データにおける各世代の個体数 $N_T(G)$ に対して、主個体の祖先でない個体数を $A_T(G)$ とし、競走馬の実データによる非先祖率 $s_T(G)$ を

$$s_T(G) = \frac{A_T(G)}{N_T(G)}, \quad (2.7)$$

で計算した。この結果、 $s_T(G)$ は 17 世代前でおおよそ 0.9 となった。図 2.9 に $s_T(G)$ および $s(G)$ の世代依存性を表したものを示す。この値は Derrida らによる理論値 $s(G)$ のおおよそ 8.39 倍であり、Derrida らの理論と競走馬のデータは一致しているとは言えないと水口は結論付けている。

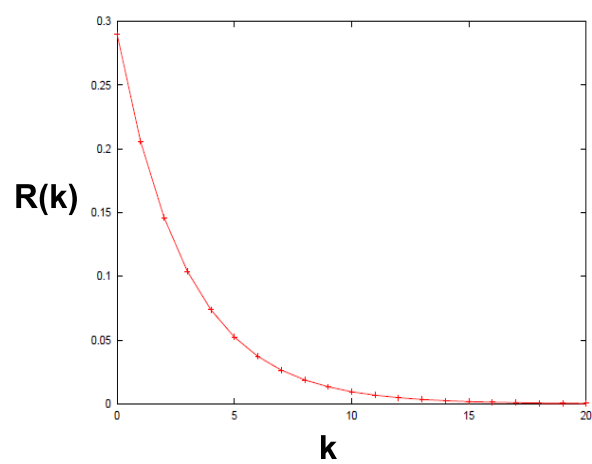


図 2.6: 幾何分布 $R(k)$ 。横軸は k 、縦軸は $R(k)$ 。赤印及び赤点線は $q = 0.71$ の時の $R(k)$ を表す。

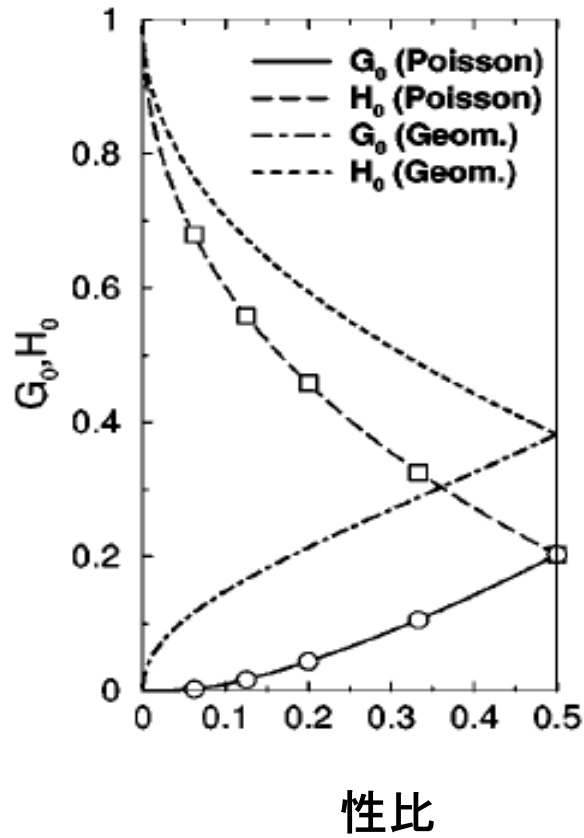


図 2.7: G_0 , H_0 の性比依存性。縦軸は G_0, H_0 、横軸は性比 f_P 。 $N_P = 20000$ 。印がついているのは $f_P = \frac{1}{2}, \frac{1}{3}, \frac{1}{5}, \frac{1}{8}, \frac{1}{16}$ の時である。実線、点線は仔数分布がポアソン分布の時の G_0, H_0 。二点鎖線、破線は幾何分布の時の G_0, H_0 。文献 [4] より転載。

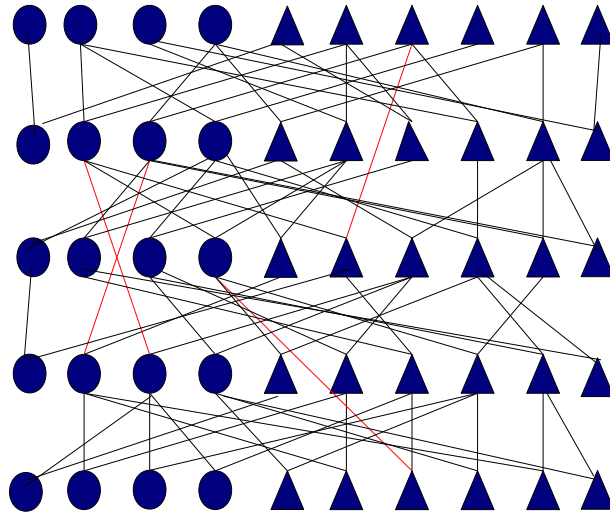


図 2.8: 競走馬の家系図の模式図。未来方向は下向き。●はオス、▲はメスを表す。親子関係にある個体同士を線で結んでおり、赤線は世代を重複している個体を結んでいる。 $N_T(G) = 10$ で固定。

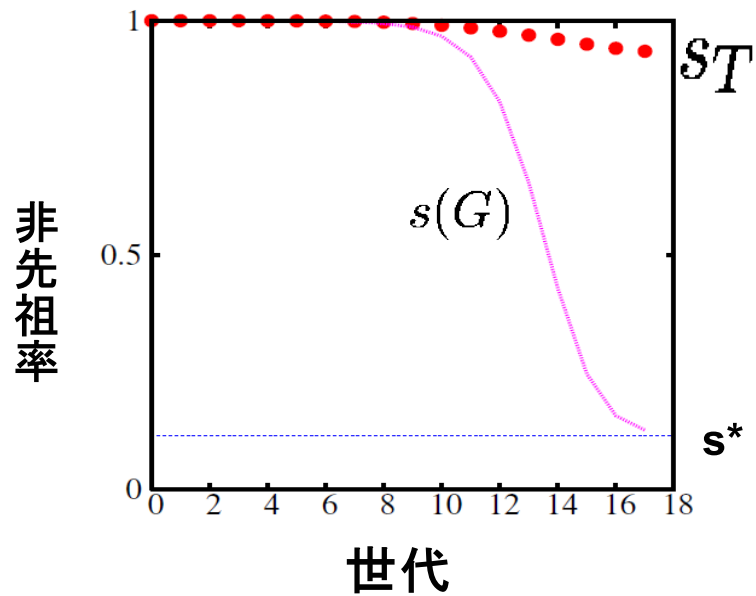


図 2.9: $s(G)$ と $s_T(G)$ の世代依存性。横軸は世代、縦軸は非先祖率。赤丸は $s_T(G)$ 、桃色点線は $s(G)$ 、水色点線は s^* を表す。

第3章 競走馬データの再解析

データベースに掲載されている競走馬データは、利用する人々のニーズを考えるとレース成績の優秀な個体が多いと推定され、水口のデータ解析は偏りがあると考えられる。しかし、先行研究ではこのことがあまり考慮されていない。例えばサンプリングのデータからオスの平均子数とメスの平均子数を計算すると、それぞれ 30.8、3.6 となったが、水口の見積もりによれば、平均出産数 $m_T = 2.5$ 、性比 $p_T = 0.25$ であり [6]、性比の偏りを考慮しても性別ごとの平均子数と m_T が一致していない。データベース上に掲載されている競走馬データに偏りがあると考えられる。本研究ではより偏りの少ない統計によるパラメーターの見積もりを考えるために、競走馬のデータの再解析を行った。

3.1 性の数比

水口が 4349 頭をサンプリングしたデータを用いた結果、 $p_T = 0.25$ と算出された [6]。この見積もりの妥当性を考える。まず、最も単純な状況として、2 世代だけで閉じている場合を考え、

「世代 $G - 1$ の全ての個体の親は全て世代 G の個体である」(仮定 A)

という仮定を設定する。世代 G での総個体数を $N(G)$ 、性の数比を p_p とすると、世代 G のオスの数、メスの数はそれぞれ、 $p_p N(G)$ 、 $(1 - p_p)N(G)$ である。また世代 $G - 1$ の性比を p_c とし、世代 G のオス 1 個体あたりの息子数を C_{ms} とすると、 $p_p N(G)$ のオスに対して $p_p N(G)C_{ms}$ の息子が存在することになる。同様に、世代 G のオス 1 個体あたりの娘数を C_{md} とすると、 $p_p N(G)$ のオスに対して $p_p N(G)C_{md}$ の娘が存在することになる。次に、世代 G のメス 1 個体あたりの息子数、娘数をそれぞれ C_{fm} 、 C_{fd} とすると、 $(1 - p_p)N(G)$ のメスに対して $(1 - p_p)N(G)C_{fd}$ の息子、 $(1 - p_p)N(G)$ の娘が存在することになる。世代 $G - 1$ の全個体の父親が全て

世代 G にいるという仮定から、

$$p_{c1} = \frac{p_p N(G) C_{ms}}{p_p N(G) C_{ms} + p_p N(G) C_{md}} = \frac{C_{ms}}{C_{ms} + C_{md}}, \quad (3.1)$$

は世代 $G - 1$ の性比 p_c を意味している。同様に世代 $G - 1$ の全個体の母親も全て世代 G の個体にいると仮定したので、

$$p_{c2} = \frac{(1 - p_p) N(G) C_{fs}}{(1 - p_p) N(G) C_{fs} + (1 - p_p) N(G) C_{fd}} = \frac{C_{fs}}{C_{fs} + C_{fd}}, \quad (3.2)$$

も p_c を示している。仮定 A が満たされており、かつ全ての個体データがあれば、

$$p_{c1} = p_{c2} = p_c, \quad (3.3)$$

が成り立つはずである。つまり、世代 $G - 1$ の性比を、「父親から見た仔の性比」と「母親から見た仔の性比」という 2 種類の方法で計算することができる。そのときに用いるのは $C_{ij}(i = m, f)(j = s, d)$ である。

次に、世代 G での性比 p_p を C_{ij} を使って計算する。仮定 A より、世代 $G - 1$ のオスは、世代 G のオスの息子であり、メスの息子でもある。このことから、

$$(\text{世代 } G - 1 \text{ のオスの数}) = p_p N(G) C_{ms} = (1 - p_p) N(G) C_{fs}, \quad (3.4)$$

が成立する。この中辺 = 右辺を解いて、

$$p_p = p_{p1} \equiv \frac{C_{fs}}{C_{ms} + C_{fs}}, \quad (3.5)$$

が成り立つ。同様に、世代 $G - 1$ のメスは、世代 G のオスの娘であり、メスの娘であることから、

$$(\text{世代 } G - 1 \text{ のメスの数}) = p_p N(G) C_{md} = (1 - p_p) N(G) C_{fd}, \quad (3.6)$$

が成立する。この中辺 = 右辺を解いて、

$$p_p = p_{p2} \equiv \frac{C_{fd}}{C_{md} + C_{fd}}, \quad (3.7)$$

が成り立つ。

p_c の場合と同様、仮定 A が満たされており、かつ全ての個体データがあれば、

$$p_p = p_{p1} = p_{p2}, \quad (3.8)$$

が成り立つ。 p_c と同様に世代 G の性比 p_p も2種類の方法で計算が可能である。もちろん世代を通じて性比が一定であれば、

$$p_p = p_c, \quad (3.9)$$

より、 C_{ij} から計算される p_{p1} 、 p_{p2} 、 p_{c1} 、 p_{c2} は全て一致するはずである。さらに、ここまでは G 、 $G - 1$ の二つの世代のみで考えたが、個体がさまざまな世代に分散していても、ある個体集団とその仔をもれなく数え上げることができれば成り立つと考えられる。性比が世代を通じて変化する場合はその平均が求まることになる。

しかし、実際は全個体のデータを得るのは困難であり、サンプリングした一部のデータから C_{ij} を推定することになるが、サンプリングしたときの偏りが問題である。まず、水口がサンプリングに用いた 4349 頭のデータを見てみる。

表 3.1: 4349 頭のデータ

	個体数	息子数 (平均息子数)	娘数 (平均娘数)	総仔数 (平均仔数)
オス	1256	14536(11.6)	24143(19.2)	38679(30.8)
メス	3093	4081(1.3)	7069(2.3)	11150(3.61)

サンプルしたデータの平均子数から、 p_{p1} 、 p_{p2} 、 p_{c1} 、 p_{c2} を見積もってみる。表 3.1 は、4349 頭の個体データから得られたオス及びメスの息子数、娘数、平均息子数、平均娘数、総仔数、平均仔数をまとめたものである。サンプリングによって得られた平均仔数測定値をそれぞれ C'_{ms} 、 C'_{md} 、 C'_{fs} 、 C'_{fd} と定義すると、

$$C'_{ms} \sim 11.6, \quad C'_{md} \sim 19.2, \quad C'_{fs} \sim 1.3, \quad C'_{fd} \sim 2.3, \quad (3.10)$$

である。これらの測定値から見積もられる仔世代の性比及び親世代の性比 p'_{c1} 、 p'_{c2} 、 p'_{p1} 、 p'_{p2} を計算してみると、

$$p'_{c1} = 0.38, \quad p'_{c2} = 0.37, \quad p'_{p1} = 0.10, \quad p'_{p2} = 0.11, \quad (3.11)$$

となる。 p' のそれぞれの値を見ると、性別による変化は小さく、親の性比と仔の性比は違いが大きい、ということが分かる。この隔たりはサンプルの偏りに起因すると思われる。データベース上の競走馬のデータがある程度有名な馬で構成されているとすれば、多くの場合「有名な馬 = 活躍した馬」と考えられるから、データベ-

ス上の競走馬は無名の馬よりも子孫を残しやすく、仔数も真の平均よりも大きい値をとっていると考えられる。

ここで、

「偏りがあるかもしれない平均仔数の測定値 C'_{ij} と真の平均仔数 C_{ij} の比は、親の性では違っているかもしれないが、仔の性では変わらない」(仮定 B)

と仮定する。つまり、

$$\frac{C'_{ms}}{C_{ms}} = \frac{C'_{md}}{C_{md}} \equiv r_m, \quad \frac{C'_{fs}}{C_{fs}} = \frac{C'_{fd}}{C_{fd}} \equiv r_f, \quad (3.12)$$

と設定する。すると、仔の性比は式 (3.1)、式 (3.2) より、

$$p_{c1} = \frac{\frac{C'_{ms}}{r_m}}{\frac{C'_{ms}}{r_m} + \frac{C'_{md}}{r_m}} = \frac{C'_{ms}}{C'_{ms} + C'_{md}}, \quad (3.13)$$

$$p_{c2} = \frac{\frac{C'_{fs}}{r_f}}{\frac{C'_{fs}}{r_f} + \frac{C'_{fd}}{r_f}} = \frac{C'_{fs}}{C'_{fs} + C'_{fd}}, \quad (3.14)$$

となり、測定値による値がそのまま算出される。一方、親の性比は式 (3.5)、式 (3.7) より、

$$p_{p1} = \frac{\frac{C'_{fs}}{r_f}}{\frac{C'_{ms}}{r_m} + \frac{C'_{fs}}{r_f}} = \frac{C'_{fs}}{C'_{ms} \frac{r_f}{r_m} + C'_{fs}}, \quad (3.15)$$

$$p_{p2} = \frac{\frac{C'_{fd}}{r_f}}{\frac{C'_{md}}{r_m} + \frac{C'_{fd}}{r_f}} = \frac{C'_{fd}}{C'_{md} \frac{r_f}{r_m} + C'_{fd}}, \quad (3.16)$$

となる。正しい性比 $p = p_{c1} = 0.38$ とすると、式 (3.15) より $\frac{r_f}{r_m} = 0.18$ となる。一方、式 (3.16) より $\frac{r_f}{r_m} = 0.20$ で、ほぼ一定の値をとっていると言える。

ここで、サンプリングで取り出した 4349 頭の仔を考える。これらの仔は、有名な個体の仔であるから、依然有名である可能性は残っているが、無名な馬が含まれていると考えられる。4349 頭の仔の個体数、息子数、平均息子数、娘数、平均娘数、

表 3.2: 4349 頭の仔のデータ

	個体数	息子数 (平均息子数)	娘数 (平均娘数)	総仔数 (平均仔数)
オス	17925	49388(2.8)	77021(4.3)	126409(7.1)
メス	30435	32178(1.3)	48744(1.6)	80922(2.7)

総仔数、平均仔数を表 3.2 にまとめた。4349 頭の仔の平均仔数の測定値をそれぞれ C''_{ms} , C''_{md} , C''_{fs} , C''_{fd} と定義すると、

$$C''_{ms} \sim 2.8, \quad C''_{md} \sim 4.3, \quad C''_{fs} \sim 1.1, \quad C''_{fd} \sim 1.6, \quad (3.17)$$

となり、4349 頭のデータに比べ、それらの仔のデータは性別間の平均仔数の差が小さくなった。これらの測定値を用いて見積もられる仔世代の性比及び親世代の性比 p''_{c1} , p''_{c2} , p''_{p1} , p''_{p2} を計算すると、

$$p''_{c1} = 0.39, \quad p''_{c2} = 0.40, \quad p''_{p1} = 0.28, \quad p''_{p2} = 0.27, \quad (3.18)$$

となる。これらの値と式 (3.11) と比べると、 p''_{p1} 、 p''_{p2} の値は大きく変化した。したがって、 p_{p1} 、 p_{p2} の測定値は信用できないと考えられる。一方で、 p''_{c1} 、 p''_{c2} の方は、式 (3.11) と比べても、値が 0.03 ほどしか変わっていないところから、 p_{p1} 、 p_{p2} に比べ p_{c1} 、 p_{c2} は信用できると考えられる。本研究では競走馬の性比 $p_T = 0.4$ 、つまりオス : メス = 2 : 3 と設定した。

3.2 親子年齢差

親子年齢差は、各個体が父親及び母親と何歳離れているかを表すものである。総個体数に対して親子年齢差の平均を取り、この値を競走馬の 1 世代と考える。水口はデータベース上のデータから作った競走馬の家系図に対して木を構成し、ある世代までの木に対して測定を行ったところ、平均親子年齢差は 12 ± 1 年となった [5]。

4349 頭の仔の各個体に対して平均親子年齢差を調べたところ、母親との平均年齢差は 11.1 年、父親との平均年齢差は 12.1 年となった。これらの値は木の測定から求めた年齢差と矛盾しておらず、競走馬の 1 世代は 11.5 年としてよいと考えられるので、本研究では競走馬の 1 世代を 11.5 年と設定した。

3.3 平均出産数

前に述べた通り 1 世代辺りの平均出産数 m は、競走馬の誕生年分布を考え、そこから 1 年での個体増加率 m_Y を計算し、ペアあたりの平均仔数 $\frac{m}{2}$ を、

$$\frac{m}{2} = m_Y^{-11.5}, \quad (3.19)$$

で求められることから見積もった。4349 頭のデータから見積もった個体増加率 $m_{Y4349} = 0.9823$ 、4349 頭の仔のデータから見積もった個体増加率 $m_{Y4349J} = 0.9663$ であったので、式 (3.19) を用いて、4349 頭のデータによる平均出産数 $m_{4349} = 2.5$ 、4349 頭の仔のデータによる平均出産数 $m_{4349J} = 3.0$ とそれぞれ見積もられた。

二つの値に違いがあったが、本研究では平均出産数 m の値は 2.5 を採用している。

3.4 仔数分布

4349 頭のサンプリングデータを用いてオス馬およびメス馬の仔数分布を見積もった。4349 頭のデータから得られた仔数をランダムにサンプリングした結果、オス及びメスの仔数分布は図 3.1 のようになった。ただし、仔の持たないオスの割合は 0.5 程あり、それらは除外している。

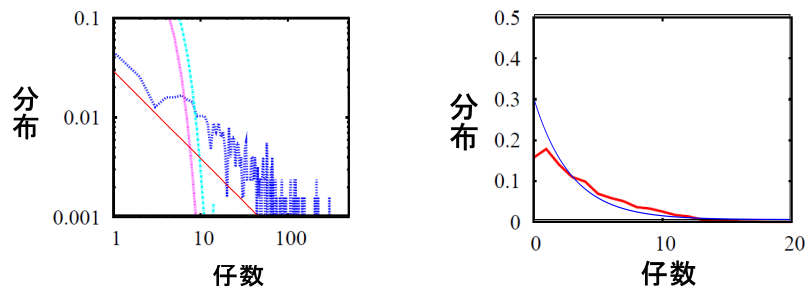


図 3.1: 左:データベースからサンプリングした 4349 頭の競走馬のうちのオスの仔数分布。青線で表わされている。横軸は仔数、縦軸は分布を表している。両軸共に log スケールである。青色がオス馬のデータをプロットしたものであり、桃色線、水色線はポアソン分布である。赤線は式 (3.20)。右:4349 頭の競走馬のうちのメスの仔数分布。赤線で表されている。横軸は仔数、縦軸は分布を表している。青線は式 (3.22)。

図 3.1 左を見ると、オスの仔数分布は揺らぎが大きい、仔を持たないオスを除くと指数 1 ~ 1.2 のべき分布に近い形をしていると見受けられる。そこでオスの仔数

分布 $P_m(k)$ を、以下の式のようにフィッティングした。

$$P_m(k) = \begin{cases} a_0 & (k = 0), \\ a_1 k^{-\gamma} & (0 < k \leq k^*), \\ 0 & (k > k^*). \end{cases} \quad (3.20)$$

式 (3.20) 内のパラメーターはそれぞれ、

$$a_0 = 0.9, \quad \gamma = 1.2, \quad a_1 = \frac{1 - a_0}{\sum_{k=1}^{k^*} k^{-\gamma}}, \quad k^* = 1000, \quad (3.21)$$

である。仔を持たないオスの割合の見積もりについては文献 [12] を参照されたい。通常のべき分布に「 $k > k^*$ では 0 の値を取る」というカットオフを加えており、式 (3.20) を「カットオフ-べき分布」と名づけた。式 (3.20) のグラフは図 3.1 左に示されている。一方でメスの仔数分布は、図 3.1 右を見ると幾何分布に近い形をしていると考えられる。メスの仔数分布 $P_f(k)$ を幾何分布、

$$P_f(k) = (1 - q)q^k, \quad (3.22)$$

とフィッティングした。パラメーター $q = 0.71$ である。式 (3.22) のグラフは図 3.1 右に示されている。

第4章 モデル

Derrida らのモデルと競走馬の家系図をつなぐモデルを作るために、Derrida らが設定したいくつかの仮定を修正した。本章では2つのモデル「性数比可変モデル」、「性比-仔数分布可変モデル」を紹介する。

4.1 性数比可変モデル

2章で述べた通り、Derrida らの単純なモデルでは競走馬のデータを説明できるとは言い難い。この原因は Derrida らがモデルを作る際に設定した仮定 1 ~ 4 のいずれかが破れていると考えられる。我々は、まず性比の効果に注目した。各世代の総個体におけるオスの数の割合を性比とし、前節で挙げた Derrida らのモデルの仮定の中で、性比を $p(0 < p < 1)$ とし、 G 世代の時の p を変化させた時の非先祖率 $s_f(G, p)$ の世代依存性を調べた [10]。

両親は1つ上の世代のオスから父親、メスから母親をランダムに選び、 $N(G)$ の計算は Derrida らのモデルと同様である。このモデルを性数比可変モデルと名づけた。性数比可変モデルが持つ性質をまとめると、

- 1 . 両親のペアはランダム
- 2''' . 性別の数比はオス:メス= $p : 1 - p$
- 3 . 世代の重複は無い
- 4 . 仔数は平均 m のポアソン分布に従う

である。このモデルの性質は、仔数分布をポアソン分布とした時の Pla らのモデルと本質的に同じである。性数比可変モデルの仔数分布に、平均 2 のポアソン分布を用いれば Pla らのモデルと対応する。

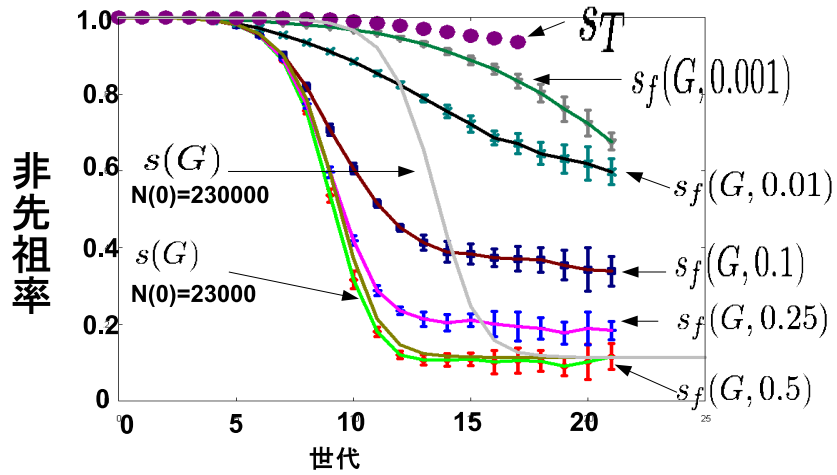


図 4.1: 性数比可変モデルによる、性比 p を変化させたときの $s_f(G, p)$ の G 依存性、競走馬の実データによる $s_T(G)$ の実測値、Derrida らのモデルによる $s(G)$ の理論値を表した図。縦軸はそれぞれ非先祖率 s , s_T , s_f 、横軸は世代を表す。丸印は s_T 。それ以外の印は $s_f(G, 0.001)$ 、 $s_f(G, 0.01)$ 、 $s_f(G, 0.1)$ 、 $s_f(G, 0.25)$ 、 $s_f(G, 0.5)$ を表す。縦線はそれぞれの値の標準偏差を表す。 $N(0) = 23000$, $N(0) = 230000$ の時の $s(G)$ も示す。

性数比可変モデルを用いて非先祖率 $s_f(G, p)$ を測定したところ、図 4.1 のようになった。 p の値が小さくなるほど世代をさかのぼるにつれて起こる $s_f(G, p)$ の値の落ち込みが小さくなり、収束する値が大きくなることが分かった。 $p = 0.4$ の時が競走馬の性比と見積もられているが、上の考察から $s_f(G, 0.4)$ は $s_f(G, 0.25)$ と $s_f(G, 0.5)$ の間の値をとると考えられる。しかし $G = 17$ で $s_T(G)$ と $s_f(G, 0.4)$ の二つの値に $0.7 \sim 0.8$ 程の差がある。これは Derrida らのモデルを性の数比について改良するだけでは競走馬のデータを説明できないことを意味する。

4.2 性比-仔数分布可変モデル

前節で述べたように、Derrida らのモデルを性の数比を可変にするという修正だけでは、競走馬のデータを説明できなかった。競走馬の家系図は、Derrida らの設定した仮定のうち、2 以外の他の 3 つの仮定も破れているからだと考えられる。そこで性の数比に加え仔数分布にも注目し、オスが k 頭の仔を持つ確率を $P_m(k)$ 、メスが k 頭の仔を持つ確率を $P_f(k)$ とし、それぞれにいろいろな仔数分布を与えて森を構築し、非先祖率 $s_h(G)$ の世代依存性を調べた。計算機上で森を構築する過程で、

Derrida らのモデルと同様に、最も若い世代を $G = 0$ とし、世代の重複はないとした。性比 p をオスが産まれる確率とし、世代を通して性比 p は固定する。

本モデルでは与えられた仔数分布を作るために森の作り方を変更した。Derrida らのように過去にさかのぼる方向で森を作る時に、両親を 1 つ上の世代からランダムに選べば、その結果各個体の仔数分布はポアソン分布になる。しかし、過去にさかのぼる方向に森を構築する方法で両親の選び方に工夫をし、任意の仔数分布を作るのは難しい。そこで本モデルでは Derrida らのモデルや性数比可変モデルとは異なり、最も昔の世代の総個体数を与え $G = 0$ に向かって未来方向に森を作る。世代ごとにオスの仔数とメスの仔数を、それぞれ $P_m(k)$ 、 $P_f(k)$ に従って割り振る。1 つ下の世代の個体数は、オスの仔数の合計とメスの仔数の合計が一致しない時は合計数が少ない性別の総仔数が採用されたとした。その時、採用されなかった性別の仔数分布は、採用した個体数を総仔数として各個体に割り振られる。両親は、1 つ上の世代のオスから一様乱数 [13] により父親、メスから母親をランダムに選んだ。本モデルは以下の特徴を有している。

1 . 両親のペアはランダム

2''' . 性別の数比はオス : メス = $p : 1 - p$

3 . 世代の重複は無い

4'' . 仔数はオスは $P_m(k)$ 、メスは $P_f(k)$ (k : 仔数) に従う

このモデルを「性比-仔数分布可変モデル」と名づけた。

4.3 仔数を割り振る方法

G 世代の各個体に与えられた仔数分布 $P_m(k)$ 、 $P_f(k)$ に従った仔数を割り振る際、2通りの方法が考えられる。第一は、オスとメスに $P_m(k)$ 、 $P_f(k)$ に従って k 頭仔を持つ個体数を決めるものである。この方法を「DCN 法」と名づけた。DCN 法に従うと、図 4.2 が示すように、 G 世代にオスが $N_m(G)$ 、メスが $N_f(G)$ 存在している時、 k 頭仔を持つ仔数はそれぞれ $N_m(G)P_m(k)$ 、 $N_f(G)P_f(k)$ である。

第二に、逆関数法 [13] により与えられた仔数分布に従う乱数を作り、その乱数を用いて各個体に仔数を割り振る方法がある。この方法を「関数作成法」と名づけた。

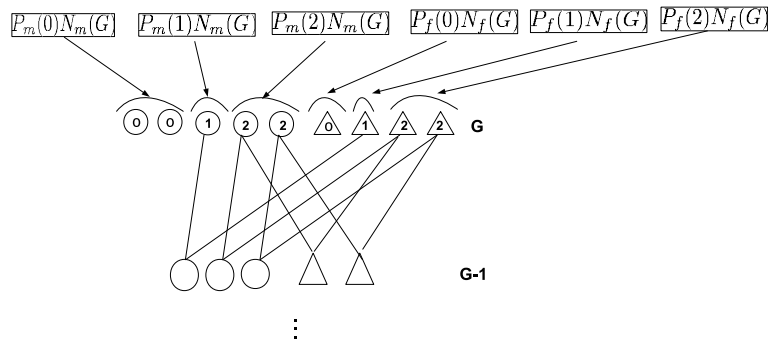


図 4.2: DCN 法による仔数の割り振り方の説明を表したもの。丸印がオス、三角印がメスを表す。枠内の数字が G 世代の各個体が持つ仔数である。

しかし、この方法では個体増加率が森ごとで大きく揺らいでしまうという欠点がある。これについては付録 B で述べる。DCN 法で仔数を割り振ると、あらかじめ世代ごとの各個体に与えたい仔数分布に従って仔数を割り振ることで、何回かの試行により作られた森ごとでの個体増加率の揺らぎを小さくすることができる。図 4.3 は DCN 法で作られた 10 個の森の世代に対する個体増加率の平均を示したものである。しかし、各個体の性別は一様乱数により割り振られており、性比が世代ごとで揺らいでいるため、世代の個体数増加率もある程度の揺らぎは存在してしまう。図 4.4 は性比を 0.5 と設定したときの 10 個の森の世代に対する性比の平均を示したものである。乱数による揺らぎが存在していることが分かる。各世代での個体数が確定してから性比に従ってオスの個体数、メスの個体数を決めると図 4.5 に示されるように性比の揺らぎがなくなる。この方法で性別を割り振り、DCN 法を用いると図 4.6 に示されるように個体増加率の揺らぎもなくなる。

仔数に関しては関数作成法よりも DCN 法の方が各世代の個体数の増減の揺らぎをより少なくできるという点から、性比-仔数分布可変モデルによる森の作成において、DCN 法を採用した。しかし性比に関しては、実生物の繁殖時の条件を考慮し一様乱数により各個体に性別を割り振った。

4.4 木の作成のアルゴリズム

木を作成する際、木に属するのべの祖先数は世代をさかのぼるにつれて 2 のべき乗で増えていくため、単純に祖先数を数え上げていくだけでは解析に時間がかかっ

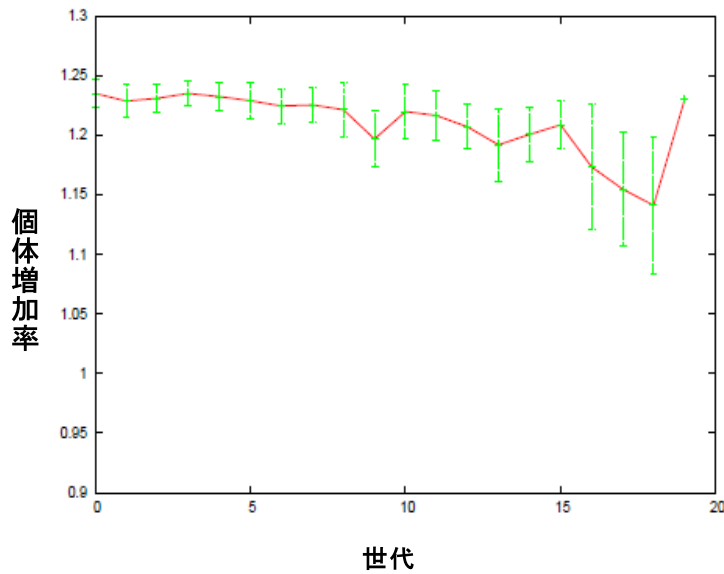


図 4.3: DCN 法で作られた森の個体増加率。横軸が世代。縦軸が個体数増加率。個体数増加率は $\frac{N(G+1)}{N(G)}$ ($G = 0, 1, \dots, 19$) で計算される。性比は 0.5, $P_m(k)$, $P_f(k)$ 共に平均 2.5 のポアソン分布である。赤線が個体増加率の平均値、緑色の縦線が標準偏差を表す。

てしまう。そこで木の作成にも工夫が必要である。

世代を遡りながら森の中の主個体の祖先を探索して木を作成する。旧来の主個体の祖先の探索法は「個体の両親をもれなく見る」というものである。図 4.7 のように主個体の祖先となる個体の両親を、役割の重複のあるなしに関わらず数え上げていく。この探索法だと、1 個体ののべの祖先数は世代をさかのぼるにつれて 2 のべき乗で増加していくので、探索回数も同じ数だけ増加していく。この方法だと探索に時間がかかってしまい、効率的とはいえない。

旧来の木の作成法に対し、新しい主個体の祖先の探索法は、図 4.8 のように「ある世代の木に属する個体の両親になっているかどうか、1 つ上の世代の個体をもれなく調べる」ものである。探している世代の個体が、1 つ下の世代の木に属している個体の親に一度でもなっていれば、次の個体を調べることにする。この方法は、木に属する個体ごとで重複度合いが分からないが、木に属する個体の探索回数は多くても世代個体数を超えないので、探索時間は旧来の方法に比べ短縮される。本研究ではこの新しい木に属する個体の探索法で木を作成した。

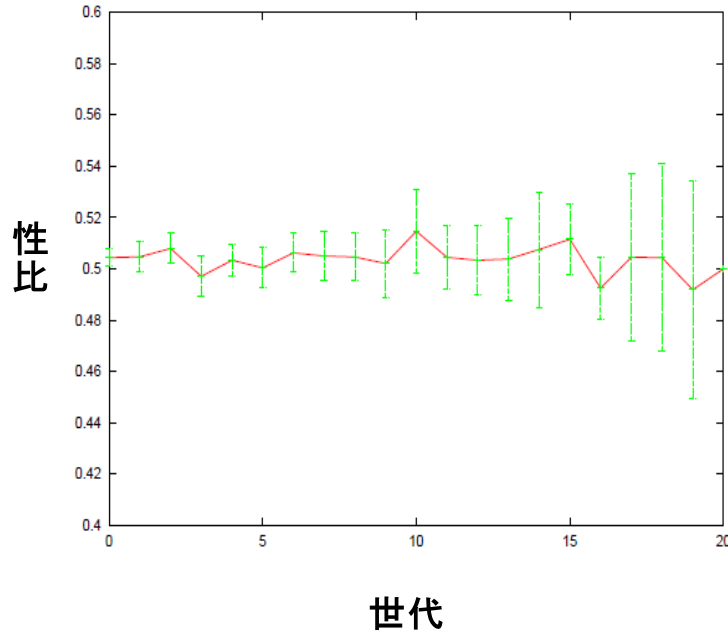


図 4.4: 世代ごとの森の性比。横軸が世代、縦軸が性比。赤線が性比の平均値、緑色の縦線が標準偏差を表す。

4.5 解析結果

競走馬の実データから $N_T(20) \sim 2600$ と見積もられたが、性比-仔数分布可変モデルにおいては、解析の便宜上 $N_h(20) = 200$ とした。性比 p は 3 章での競走馬の性比に関する見積りより $p = 0.4$ を用いた。オスの仔数分布 $P_m(k)$ 及びメスの仔数分布 $P_f(k)$ に与える分布関数には第 3 章での競走馬の個体増加率や仔数分布の見積りから式 (3.20) のカットオフ-べき分布と式 (3.22) の幾何分布、それから比較のために平均 λ のポアソン分布である、

$$P(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad (4.1)$$

を用いた。 $\lambda = 2.5$ である。 $P_m(k)$, $P_f(k)$ の組み合わせを表 4.1 に示す。各組み合わせごとに森を 10 個作り、その平均を計算した。

$s_h(G)$ の平均値 $\bar{s}_h(G)$ の世代依存性及び競走馬の実データによる $s_T(G)$ の実測値を図 4.9 に示す。図 4.9 を見ると、仔数分布によって非先祖率の振る舞いが大きく変化していることがわかる。特にカットオフ-べき分布の影響は大きく、組み合わせ 1 や組み合わせ 2 のようにオスの仔数分布にカットオフ-べき分布を導入した時は $\bar{s}_h(G)$ の値がかなり高くなる。モデルの性比と仔数分布を実データのものに合わせるこ

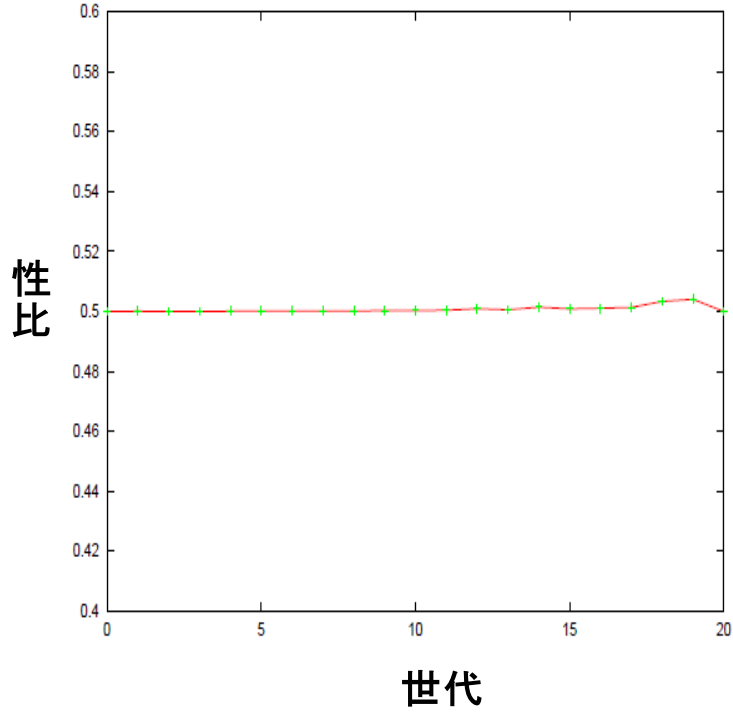


図 4.5: 性比の揺らぎを無くした森の世代ごとの性比。横軸が世代、縦軸が性比。赤線が性比の平均値、緑色の縦線が標準偏差を表す。

で、非先祖率の世代依存性が実データの振る舞いにかかなり近くなることが分かった。具体的には、組み合わせ 1 の時のオス・メスの仔数分布を採用した時の $G = 17$ の $\bar{s}_h(G)$ はおよそ 0.769 となったが、この値は $s_T(17)$ のおよそ 0.854 倍である。組み合わせ 1 の仔数分布を用いた時の性比-仔数分布可変モデルの非先祖率の振る舞いは、Derrida による見積り $s(G)$ と比較すると競走馬の実データの振る舞いにより近くなっている。

4.6 非先祖率の理論的な見積もり

Derrida らは式 (2.3) と式 (2.4) を用いて、彼らが考えたモデルでの条件下における非先祖率の世代依存性を帰納的に見積もったが、性比と仔数分布に関して一般化した家系図モデルにおいても、確率母関数を用いて非先祖率を帰納的に見積もることができる。オスの仔数分布の確率母関数を $\phi_m(x)$ とすると、

$$\phi_m(x) = \sum_{k=0}^{\infty} P_m(k) x^k, \quad (4.2)$$

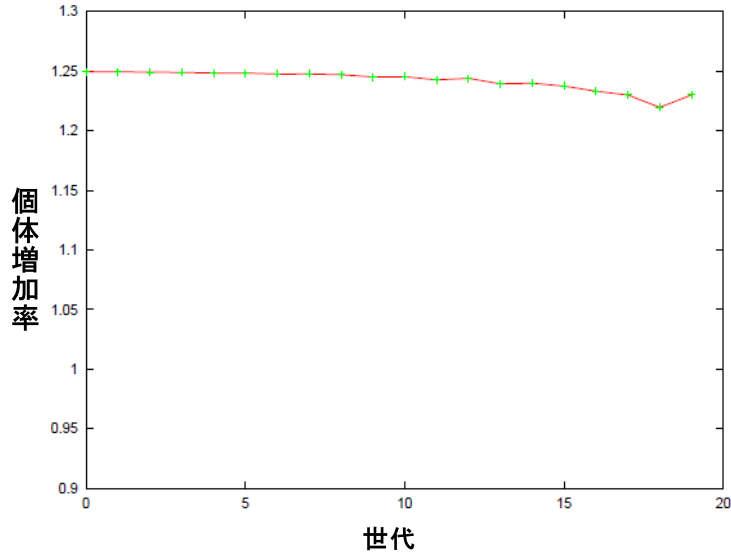


図 4.6: 性比の揺らぎを無くし、DCN 法で作られた森の個体増加率。横軸が世代。縦軸が個体数増加率。個体数増加率は $\frac{N(G+1)}{N(G)}$ ($G = 0, 1, \dots, 19$) で計算される。性比は 0.5, $P_m(k)$, $P_f(k)$ 共に平均 2.5 のポアソン分布である。赤線が個体増加率の平均値、緑色の縦線が標準偏差を表す。

であり、メスの仔数分布の確率母関数を $\phi_f(x)$ とすると、

$$\phi_f(x) = \sum_{k=0}^{\infty} P_f(k)x^k, \quad (4.3)$$

である。 G 世代の非先祖率を $s'(G)$ とし、 $s'(G+1)$ と $s'(G)$ の間に成り立つ関係を考える。 $G+1$ 世代の個体が主個体の木に属していない確率 $s'(G+1)$ は、 G 世代にいるその仔の全てがその木に属していない確率 $s'(G)$ に他ならないので、以下の漸化式が成り立つ。

$$s'(G+1) = f'\phi_m(s'(G)) + (1 - f')\phi_f(s'(G)). \quad (4.4)$$

ここで f' は性比を表わす。0 世代の総個体数を n とすれば、 $s' = 1 - \frac{1}{n}$ とおいて、式 (4.4) に代入すれば、漸化的に全世代での $s'(G)$ を求めることができる。

ところで、式 (4.4) は、式 (2.3) に対応している。Derrida らはポアソン分布の確率母関数を使って式 (2.3) を導いている。詳しい計算は参考文献 [2] を参照されたい。式 (4.4) を使った非先祖率の理論値 s' を図 4.10 に示す。競走馬の実データから見積

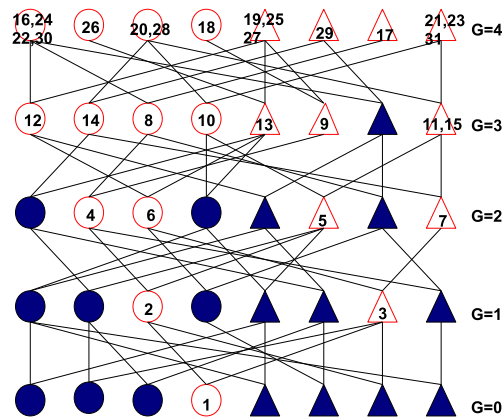


図 4.7: 旧来の木の作成法の説明を図にしたもの。丸印はオス、三角印はメスを表している。数字が振られている個体は木に属しているものであり、印内の数字は主個体から見て何番目の祖先かを表す。

表 4.1: $P_m(k)$ と $P_f(k)$ の組み合わせ

組み合わせ	$P_m(k)$	$P_f(k)$
1	カットオフ-べき分布	幾何分布
2	カットオフ-べき分布	ポアソン分布
3	ポアソン分布	幾何分布
4	ポアソン分布	ポアソン分布

もった $G = 0$ での総個体数 $N_T(0) = 230000$ を用いて、 $s'(0) = 1 - \frac{1}{230000}$ とした。

図 4.10 を見ると、理論的見積もりにおいてもカットオフ-べき分布の影響は大きく、これを用いた s' の値が大きくなり、組み合わせ 1 の時の非先祖率の収束値は実データに最も近くなっている事が分かる。

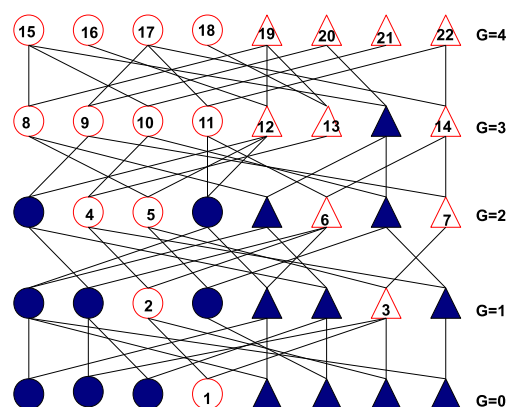


図 4.8: 新しい木の作成法の説明を図にしたもの。丸印はオス、三角印はメスを表している。数字が振られている個体は木に属しているものであり、印内の数字は主個体から見て重複を除いた何番目の祖先かを表す。

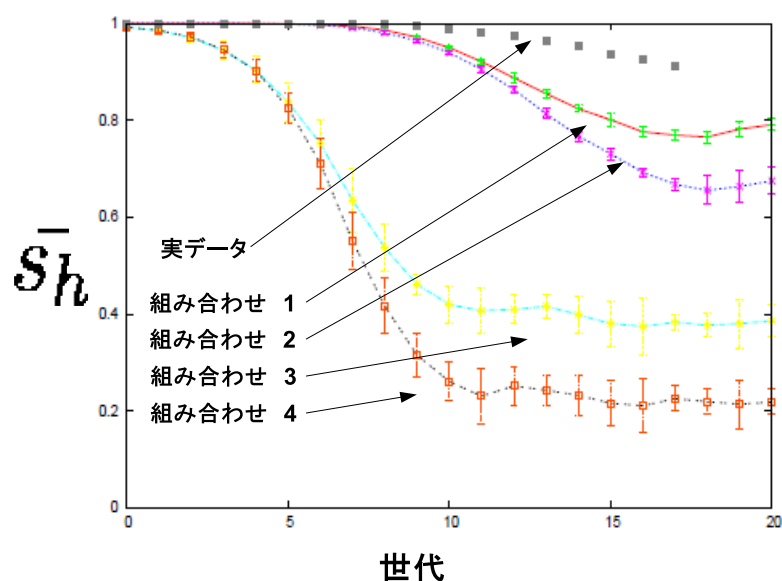


図 4.9: 性比-仔数分布可変モデルによる、オスとメスの仔数分布を変化させた時の $\bar{s}_h(G)$ の世代依存性、競走馬の実データによる $s_T(G)$ の実測値を表した図。縦軸はそれぞれの非先祖率 $\bar{s}_h, s_T(G)$ 、横軸は世代を表す。灰色印は $s_T(G)$ 。それ以外の印は表 4.1 で示した仔数分布の組み合わせの時の非先祖率の世代依存性を表している。各印の縦線は標準偏差を示している。

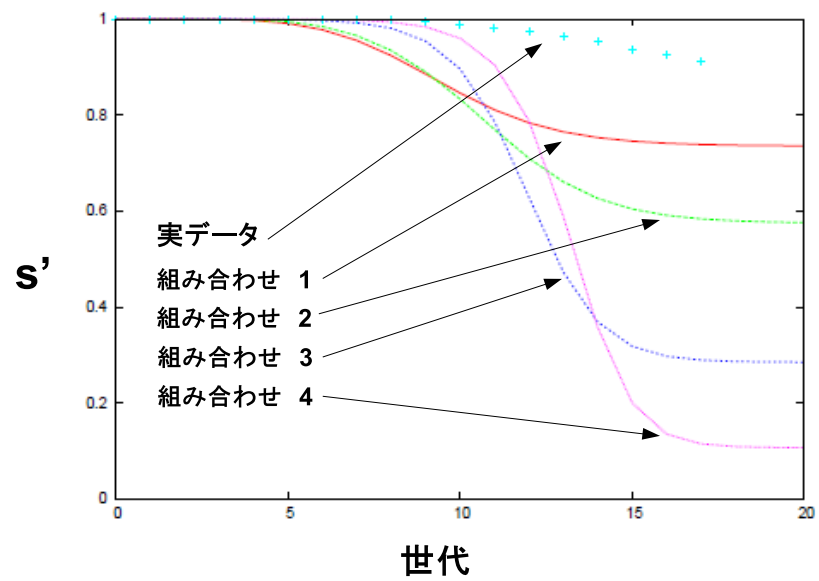


図 4.10: 式 (4.4) による非先祖率の理論的見積り。縦軸は s' 、横軸は世代を表す。水色印は $s_T(G)$ 。それ以外の実線は表 4.1 で示した仔数分布の組み合わせの時の s' の世代依存性を表している。

第5章 配偶システム

オスとメスが生涯にそれぞれ何体の異性とペアを結ぶかを「配偶システム」と呼ぶ。実在する生物は、繁殖により種族を反映させる際に様々な配偶システムを採用している。配偶システムには乱婚、一夫多妻、一妻多夫、一夫一妻などが知られている。この分類は生態学的な定義において、ペアの作り方に関するルールだけで決めたのではなく、ペア関係を持つ動物について「誰が仔の世話をするのか」といった仔の養育形態によってもわけられたものである [9]。例えばヒトの場合を考えると、日本の社会においては女性と男性のペアは、一旦ペアを形成された個体同士がその子供を育て、他の異性の個体とはペアを組めないケースがあり、その意味では一夫一妻制であると言えるが、再婚という例外もある。他の生物ではオスが複数のメスとペア形成を行う「一夫多妻制」を採用しているものもある。本研究では、配偶システムの違いをペア形成のルールのみに限定し、配偶システムによる違いから生じる非先祖率の違いを解析し、配偶システムの違いから生まれる家系図構造の違いを調べた。

5.1 配偶システムの紹介

乱婚 ある生物集団内において、オスとメスが共に複数の異性とペアを結ぶ事ができる、というペア作成のルールを持つ配偶システムである。「集団婚」とも呼ばれる。実在生物においては、ヒヒ、チンパンジーなどがこの配偶システムを採用している。両親をランダムに選んでいるという点で、前章で用いた2つのモデルはこの配偶システムを採用している。

一夫多妻 ある生物集団内においてメスは1個体のオスとしかペアを作れず、一方でオスは複数のメスとペアを結ぶことができるというペア作成のルールを持つ配偶システムである。実在生物においては、イスラム教圏内のヒト、ライオン

やアザラシなどがこの配偶システムを採用している。生態学的な観点では一夫多妻制には、資源防衛型一夫多妻、メス防衛型一夫多妻、レック型一夫多妻、スクランブル競争型一夫多妻の4種類に分けられる [9] が、本研究ではペア形成のルールのみ注目するのでこれらの小分類は考慮しない。

一妻多夫 ある生物集団内において、オスは1個体のメスとしかペアを作れず、一方でメスは複数のオスとペアを結ぶことができる、というペア作成のルールを持つ配偶システムである。一夫多妻と性別に関して対照的である。アメリカイソシギがこの配偶システムを採用していると報告されているが、一夫多妻の配偶システムをとる生物に比べ、多夫一妻の配偶システムをとる生物はまれにしか見つかっていない。本研究では、性差は性別の数比のみを考慮するので、一夫多妻と一妻多夫を同一のものとして考える。

一夫一妻 ある生物集団内で、特定のオスとメスの個体同士がペアを作り、他の異性とはペアを作らない、というルールを持つ配偶システムである。日本で社会の婚姻制度等で一般的によく知られている。「単婚」とも呼ばれ、それ以外の配偶システムをまとめて「複婚」と呼ぶこともある。DerridaらのモデルやPlaらのモデルはこの配偶システムを採用している。

5.2 ペア作成のアルゴリズム

配偶システムと性比と各性別の仔数分布は、それぞれ独立に設定することはできない。ここでは、異なる配偶システムを比較するためにメスの仔数分布を共通とした。まず、一夫多妻制の森を作る際世代 G のメスの個体全てに仔数分布を割り振り、メスの総仔数が $G - 1$ 世代の個体数になるように設定した。そしてメスは夫をオスの個体からランダムに選ぶ。オスは選ばれたメスの仔数を自分の仔数とする。複数のメスから選ばれた場合は、選ばれたメスの持つ仔数の和がそのオスの仔数となる。一夫多妻制には、ペアを組めば組むほど他の異性の個体とペアが組みやすくなるシステムや、逆にペアを組むほど他の異性の個体と組みやすくなる、いわゆるユール過程 [14] のようなシステムも考えられるが、ここではその中の一例として、上で述べたようなランダムにペアを組むという設定を採用している。一夫多妻制の森の作り方を表すと図 5.1 のようになる。

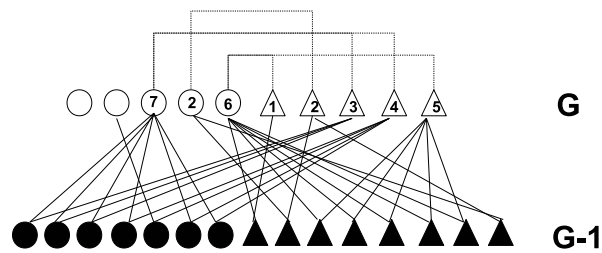


図 5.1: 一夫多妻制のペア形成による森の作り方を説明したもの。実線が親子関係を表した線であり、点線がペアを結んでいる。白抜きの印が G 世代の個体、黒塗りの個体が $G - 1$ 世代の個体であり、印の中に書かれている数字が個体の持つ仔数を表している。

こうして作られた一夫多妻制の森を元に乱婚による森を作る。一夫多妻により割り振られたオスの仔数分布は固定し、親子関係だけシャッフルすることによって乱婚制のペア形成による森を構成する。一夫多妻制による森と、乱婚制による森の構造の違いを比較するために、上述の方法で森を構成した。

一夫一妻制により作られる森は、各世代のオスとメスの総数を比較し、少ない性別がもう一方の性別からペアの相手を選ぶ。メスの個体に仔数分布を与えたが、(メスの個体数) $>$ (オスの個体数) となる場合は例外としてオスに仔数分布を与えている。一旦ペア相手として選ばれた個体は他の個体とペアを組めない。図 5.2 を例に考えると、 G 世代の個体はオスは 5 体、メスは 4 体であるから、メスに仔数分布が割り振られ、メスからペアを選ぶ。オスの中にはペアを組めない個体も存在する。ペアの相手として選ばれたオスは相手と同じ仔数を持つ。

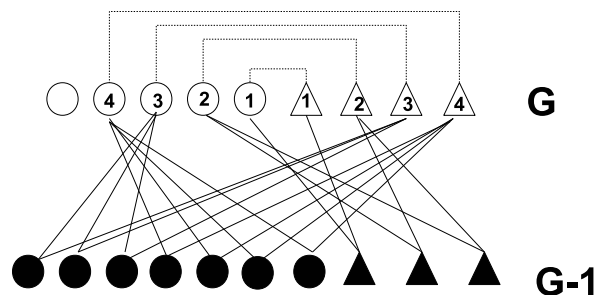


図 5.2: 一夫一妻制のペア形成による森の作り方を説明したもの。実線が親子関係を表した線であり、点線がペアを結んでいる。白抜きの印が G 世代の個体、黒塗りの個体が $G - 1$ 世代の個体であり、印の中に書かれている数字が個体の持つ仔数を表している。

5.3 配偶システムごとの非先祖率の比較

前節の方法で、乱婚制の森、一夫多妻制の森、一夫一妻制の森をそれぞれ 10 個作り、それぞれにおける木を一本ずつ作り、非先祖率の平均及び標準偏差を算出した。なお、 $G = 23$ まで森を作り、森を作る際に $N(0) = 1000$ 、性比 $f = 0.42 \sim 0.75$ の時の非先祖率を解析した。メスの各個体に与える k 頭仔を持つ確率 $P(k)$ はポアソン分布、

$$P(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad (5.1)$$

であり、パラメーター $\lambda = 2.5$ である。

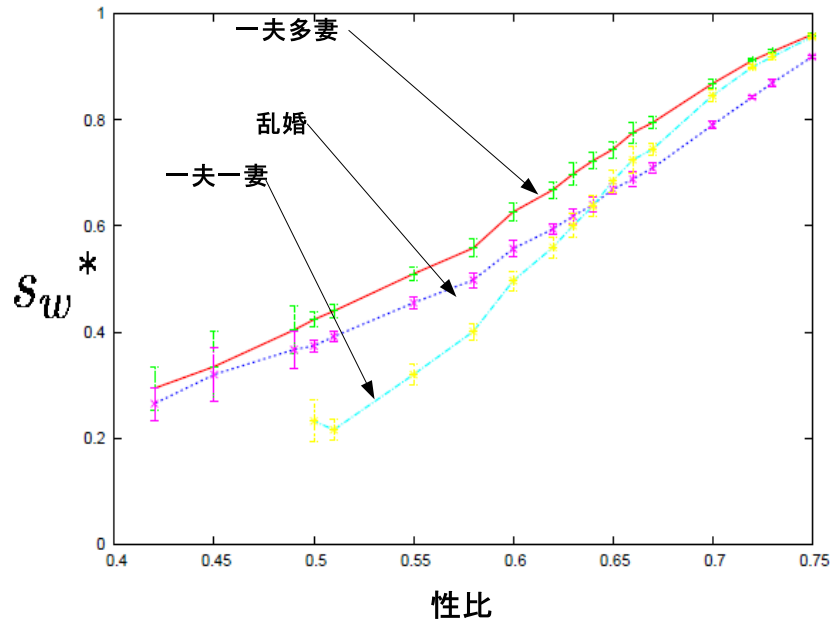


図 5.3: 0.42 ~ 0.75 の区間におけるいくつかの性比での非先祖率の収束値をプロットした図。横軸は性比で縦軸は非先祖率の収束値 s_w^* 。桃色印及び青線は乱婚制、黄色印及び水色線は一夫一妻制、緑印及び赤線は一夫多妻制の s_w^* を示し、それぞれの印についている縦線は標準偏差を表している。

図 5.3 は f ごとの非先祖率の収束値をプロットし、非先祖率の収束値 s_w^* の性比依存性を配偶システムごとで比較したものである。図 5.3 を見ると、配偶システムが変わると s_w^* も異なる値を持つことがわかる。図 5.4 は、一夫一妻制のオスとメスの仔数分布、図 5.5 は一夫多妻制及び乱婚制のオスとメスの仔数分布を示したものである。一夫多妻制により作られたオスの仔数分布を変えないように親子関係の線

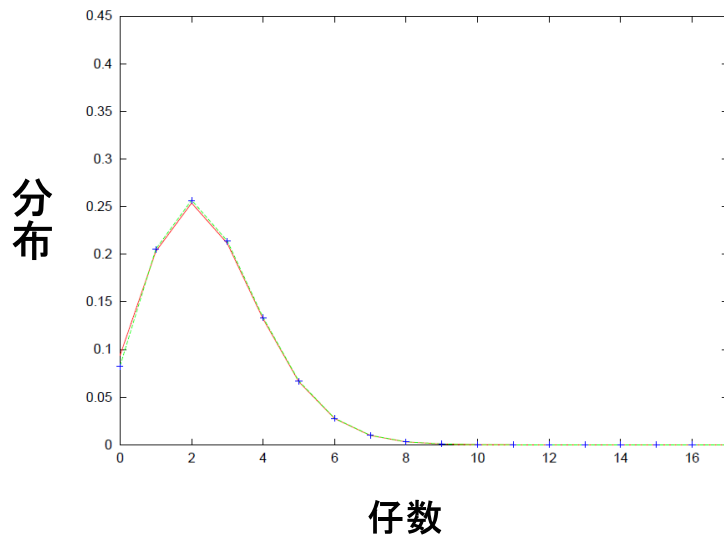


図 5.4: 性比 $f = 0.5$ の時の一夫一妻制のオスとメスの仔数分布。横軸は仔数。縦軸は分布。赤線がオスの仔数分布、緑線がメスの仔数分布、青印が平均 2.5 のポアソン分布を表す。

をシャッフルしたものが乱婚制の森なので、両配偶システムは同じ仔数分布を持っている。一夫一妻制による仔数分布と、一夫多妻制と乱婚制による仔数分布は異なる分布となっていることが図 5.4、図 5.5 を比較すると分かる。この違いが s_{ω}^* の値の差異を生んだと考えられる。一夫多妻制、乱婚制の仔数分布では、仔を持たないオスの割合が一夫一妻制のものに比べて高くなっており、そのため $f = 0.5$ 付近では一夫多妻制と乱婚制の s_{ω}^* が一夫一妻制の s_{ω}^* よりも高くなっていると考えられる。また、どの配偶システムも性比が大きくなるにつれて s_{ω}^* も大きくなっている。特に一夫一妻制ではペア作成の束縛が厳しいため、非先祖率の変化が性比の変化に敏感であると考えられる。そのため他の配偶システムと比べると s_{ω}^* の値の変化が特に大きくなっている。 $f = 0.5$ 付近では一夫一妻制の方が s_{ω}^* が最も小さいのに対し、 f が 0.64 と 0.65 の間で、一夫一妻制は乱婚制の s_{ω}^* よりも値が大きくなり、 f をさらに大きくすると一夫多妻制の s_{ω}^* に近付いていく。従って、性比によって非先祖率が最小を取る配偶システムが変わっていると解釈できる。非先祖率は森の中の一本の木の広がりを表したものである。つまり s_{ω}^* が 1 に近いほど主個体は少数の祖先の血しか引いておらず、逆に 0 に近ければ多くの祖先の血を受け継いでいると言える。図 5.3 では、 $f = 0.5$ では一夫一妻制の s_{ω}^* が最も小さく、多くの祖先の血を主個体が

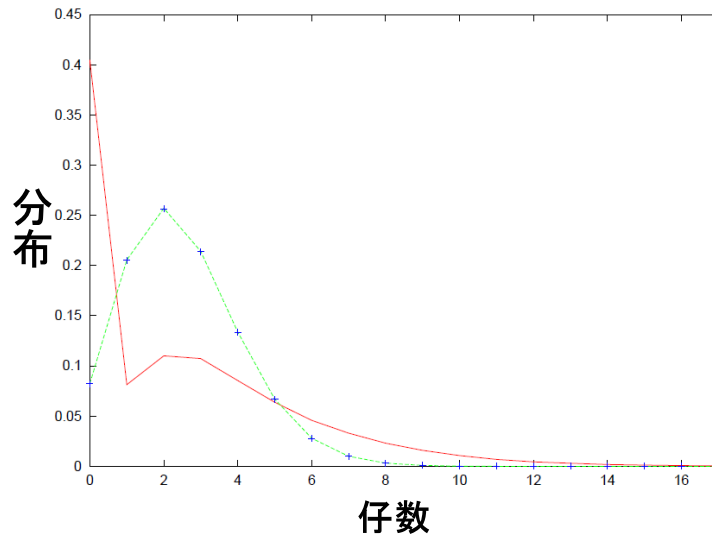


図 5.5: 性比 $f = 0.5$ の時の一夫多妻制及び乱婚制のオスとメスの仔数分布。横軸は仔数。縦軸は分布。赤線がオスの仔数分布、緑線がメスの仔数分布、青印が平均 2.5 のポアソン分布を表す。

受け継いでいると考えられる。しかし性比が大きくなると乱婚制の s_{ω}^* が最も小さくなり、乱婚制の方が血を受け継ぎやすくなっていると考えられる。まとめると、性比によって血を受け継ぎやすい配偶システムが異なっていると解釈される。

また、 $f = 0.5$ での一夫一妻制の s_{ω}^* を見てみよう。この場合は、Derrida らのモデルと同条件であるが、第 2 章で述べた Derrida らの見積もりである $s^* = 0.1073$ と異なる値をとっている。これは上で述べた一夫一妻制によるペア作成の束縛の厳しさの為に、 s_{ω}^* が本研究のシミュレーションにおける性比の揺らぎに敏感になっているため、Derrida らの見積もりからずれてしまっている。この事からも、一夫一妻制は束縛の厳しい配偶システムであることが分かる。

第6章 まとめと今後の課題

6.1 まとめ

本研究では様々な条件を考慮し、有性生殖を行う生物個体集団の家系図に対するモデルを構築して、森及び木を構成し非先祖率を解析した。各個体に与える仔数分布の関数を変化させながら非先祖率を計算したが、仔数分布の違いは非先祖率の世代依存性に大きな変化を与えることが分かった。仔数分布は家系図のネットワーク構造に大きな影響を及ぼしていると言える。また、競走馬の実データからフィッティングした性比や性別ごとの仔数分布などのパラメーターを本研究の家系図モデルに適用し、非先祖率のシミュレーション値と競走馬データによる非先祖率の実測値を比較した。Derridaらのモデルや性数比可変モデルによる非先祖率の解析と比べると、性比-仔数分布可変モデルによる非先祖率の世代依存性はかなり実データのそれに近づいており、実データの家系図構造を再現できたと考えられる。

次に家系図構造の特徴を生み出す要因としてペア形成のルールに注目し、一夫多妻制、一夫一妻制、乱婚制の三種類の配偶システムに注目して、それぞれの森から非先祖率を求め、非先祖率の世代に対する収束値 s_w^* の性比依存性を解析し、比較した。まず、配偶システムが変わると s_w^* が異なる値を持つことがわかった。一夫多妻制と乱婚制におけるオスの選び方の自由度の違いや、一夫一妻制、一夫多妻制と乱婚制の間の仔数分布に違いがあることが原因だと考えられる。一夫多妻制や乱婚制による仔数分布は、ペア相手として選ばれないオスの個体が存在するため、その結果仔数0の個体の割合が多くなり非先祖率が一夫一妻制に比べて高くなっていると考えられる。そして、どの配偶システムも性比が高くなるほど非先祖率の収束値 s_w^* が高くなることが分かった。特に一夫一妻制では、ペア作成の束縛が厳しいため、非先祖率の変化が性比の変化に敏感になっており、変化が大きい。実在する多くの生物では性比は0.5となっており、自然選択による適応の観点から性比が0.5であることが最も有利であることを、Fisher が示した [9][15]。Fisher の理論は付録

で詳述する。非先祖率が性比依存性を持っていることから、非先祖率と生物における自然選択の有利不利は何か関係があるのかもしれない。

また配偶システムごとの s_w^* を比較してみよう。乱婚制による森は、一夫多妻制による森を元に作ったので、その2つの配偶システムでは各個体の仔数分布は変わらない。一方、一夫一妻制の森における個体の仔数分布は、メスの仔数分布は共通であるが、その2種類の森のオスの個体のそれとは異なっている。それゆえに、本研究における3種類の配偶システムによる森の s_w^* は単純に比較できないが、性比の値によって配偶システムごとの s_w^* の値の大小の逆転が起こっている。非先祖率は森における一本の木の広がり具合を表したもののなので、非先祖率が低いほど個体の遺伝情報が森全体に広がりやすくなっていると言える。つまり、性比によって祖先個体の遺伝情報を現代に伝えるのに最適な配偶システムが変わると考えられる。

6.2 今後の課題

本研究では競走馬の家系図構造と、Derridaらのモデルによる家系図をつなぐようなモデルを構築することを目的の1つとしたが、第3章で紹介したような競走馬の実データから見積もられたパラメーターの設定が妥当であったかどうかは疑問である。バイアスのかかっていないデータのサンプリングを行ったり、競走馬を実際に扱う牧場などに聞き込みやフィールドワークの協力をお願いすることで、より偏りの無いデータを取ることが必要である。より正確なパラメータを用いてモデルの再構築を行い、より正確に競走馬の家系図構造に近づくことが重要だと考えられる。

また性比-仔数分布可変モデルにおいて、ある世代でのオスの持つ仔数の合計とメスの持つ仔数の合計が一致しない時は少ないほうが1つ下の世代個体数として採用されるようにしているが、それにより採用されなかった性別の仔数分布が設定時の仔数分布からずれてしまっていると考えられる。シミュレーションによる誤差をより少なくできるモデルの改良を考えることも重要であると考えられる。さらにこのモデルでは、Derridaら、Plaらのモデルと同様に世代の重複は無いという仮定のもとで森を構築しているが、競走馬の家系図には世代の重複ももちろん含まれている。世代の重複を考慮した森の構築を考えるのも重要であると考えられる。

また本研究では非先祖率の解析を通じて森の構造を見てきた。そして非先祖率の振る舞いの変化は性比や仔数分布などのパラメーターの値の変化によって引き起こ

されていることが分かった。ではその非先祖率の振る舞いの変化の要因となるこれらのパラメーターの値は、何によって決まっているのであろうか。競走馬の場合であればレースの戦績や、各個体を引き取る際の取引額の高低、生物的な特性等様々な要因が絡み合って、仔数分布や性比などのパラメーターが決まると考えられる。これらのパラメーターを決める要因をモデル設計に取り込んで森を作り、非先祖率の解析を行えるようにしたい。

さらに配偶システムの比較において、本研究では「メスがオスをランダムに選ぶ」場合の一夫多妻制だけを考えた。一夫多妻制には、5章で紹介したように様々なバリエーションが存在する。それぞれの一夫多妻制による家系図構造を解析や比較を行いたい。

非先祖率は森の中で1本の木がどれだけ広がっているかを示したものであり、主個体がどれだけ多くの祖先から血を受け継いだかの目安として考えられる量である。非先祖率をはじめとする家系図の構造を特徴づける量と生態学や遺伝学と結び付けられるようにしたい。近い将来、家系図ネットワークの構造解析が盛んになり、生態学や遺伝学が大きく発展することが望まれる。

謝辞

本研究を行うにあたり、指導教官である水口毅講師には非常に感謝しております。指導教員として私の研究の指導等に時間を割いていただき、また研究に行き詰った時に的確な助言で私の研究の方向性を示してくださいました。また本研究の共同研究者としても水口毅講師には尽力いただきました。研究に関する議論を幾度と無く交わし、私の研究内容に豊かさを与えてくださるだけでなく、研究会の発表や論文の投稿の機会等を与えていただき、私の成長の機会を与えていただきました。静岡大学工学部システム工学科所属の守田智先生にも、本研究の共同研究者として、貴重な時間を割いて私の研究発表や論文にも目を通していただき、研究に関する幾多の助言や知識を与えていただきました。厚く感謝しております。大同 明教授には、非線形力学に関する広範な知識を与えていただいただけでなく、本研究に関する有益な助言、学会発表のようなプレゼンテーションでの心構えも教えていただきました。とても感謝しております。堀田武彦准教授には、非線形力学や統計学などの、研究に関する知識をご教授頂くとともに、研究に関する理論にヒントを与えてくださりました。感謝申し上げます。福田浩昭助教には、研究に関するご指摘、助言だけでなく、研究室生活の環境面からサポートをいただきました。研究に集中できる環境を作って頂き、非常に感謝しております。そして、敬称略させていただきますが、研究室生活を公私にわたり支えてくださった研究室の学生に心から感謝いたします。学生達の勉学や研究に打ち込む姿に刺激をいただき、また公私にわたるサポートがあり、充実した研究室生活を送ることができました。誠に感謝いたします。最後になりましたが、私の学生生活に惜しみない支援をしていただきました両親に、深い感謝の念を抱いております。

付 録 A Pla らによる理論

この節では、第 2 章の 2 節で紹介された Pla らのモデルによる家系図構造の解析に用いた母関数 $G(\lambda)$, $H(\lambda)$ の導出を示していく。まず、 g 世代の総個体数を $N(g)$ 、 $N(g)$ に対するオスの総数の割合を性比とし、 $f(g)$ とおく。 $N(g) = N$, $f(g) = f$ で固定する。 g 世代前に r 回一本の木に現われる個体のオス及びメスの数をそれぞれ $m_r(g)$ 、 $f_r(g)$ とすると、次の式が成り立つ。

$$\sum_{r=0}^{\infty} m_r(g) \Delta r = fN, \quad (\text{A.1})$$

$$\sum_{r=0}^{\infty} f_r(g) \Delta r = (1 - f)N. \quad (\text{A.2})$$

ここで、 $\Delta r = 1$ である。

さらに g 世代前の主個体ののべの祖先数のうち、オス及びメスの数はどちらも 2^{g-1} であるから、以下の関係式が成り立つ。

$$\sum_{r=0}^{\infty} r m_r(g) \Delta r = \sum_{r=0}^{\infty} r f_r(g) \Delta r = 2^{g-1}. \quad (\text{A.3})$$

また、 g 世代前のオスの総個体数に対する $m_r(g)$ の確率を $M_r(g)$ 、メスの総個体数に対する $f_r(g)$ の確率を $F_r(g)$ とすると、これらは以下のように計算される。

$$M_r(g) = \frac{m_r(g)}{fN}, \quad (\text{A.4})$$

$$F_r(g) = \frac{f_r(g)}{(1 - f)N}. \quad (\text{A.5})$$

式 (A.1) 及び式 (A.2)、式 (A.3) を用いると、式 (A.4) と式 (A.5) はそれぞれ以下のように書くことができる。

$$\sum_{r=0}^{\infty} M_r(g) \Delta r = \sum_{r=0}^{\infty} F_r(g) \Delta r = 1, \quad (\text{A.6})$$

$$\sum_{r=0}^{\infty} r M_r(g) \Delta r = \frac{2^{g-1}}{fN}, \quad \sum_{r=0}^{\infty} r F_r(g) \Delta r = \frac{2^{g-1}}{(1-f)N}. \quad (\text{A.7})$$

ここで、 r 及び $F_r(g)$ 、 $M_r(g)$ を以下のようにリスケールする。

$$P_M(r, g) = \frac{2^{g-1}}{fN} M_r(g), \quad P_F(r, g) = \frac{2^{g-1}}{(1-f)N} F_r(g), \quad (\text{A.8})$$

$$\omega_M(g) = \frac{fN}{2^{g-1}}, \quad \omega_F(g) = \frac{(1-f)N}{2^{g-1}}. \quad (\text{A.9})$$

式 (A.8) と式 (A.9) を用いると、式 (A.6) は以下のように変形され、

$$\int_0^{\infty} P_M(\omega, g) d\omega_M = \int_0^{\infty} P_F(\omega, g) d\omega_F = 1, \quad (\text{A.10})$$

式 (A.7) は以下のように変形される。

$$\int_0^{\infty} \omega_M P_M(\omega_M, g) d\omega_M = \int_0^{\infty} \omega_F P_F(\omega_F, g) d\omega_F = 1. \quad (\text{A.11})$$

次に、 $g+1$ 世代のオスの個体 i の木に表れる回数 $r_{M,i}(g+1)$ は、 g 世代の自分の息子と娘の持つ登場回数の和と同じであるから、

$$r_{M,i}(g+1) = \sum_{j \text{ son of } i} r_{M,j}(g) + \sum_{j \text{ daughter of } i} r_{M,j}(g), \quad (\text{A.12})$$

となり、メスについても同様に計算できて、

$$r_{F,i}(g+1) = \sum_{j \text{ son of } i} r_{F,j}(g) + \sum_{j \text{ daughter of } i} r_{F,j}(g), \quad (\text{A.13})$$

となる。式 (A.12) を $\frac{2^{g-1}}{fN}$ で割ると、

$$\omega_{M,i}(g+1) = \frac{1}{2} \sum_{j \text{ son of } i} \omega_{M,j}(g) + \frac{f}{2(1-f)} \sum_{j \text{ daughter of } i} \omega_{M,j}(g), \quad (\text{A.14})$$

が得られ、メスの場合は式 (A.13) を $\frac{2^{g-1}}{(1-f)N}$ で割ることより、

$$\omega_{F,i}(g+1) = \frac{(1-f)}{2f} \sum_{j \text{ son of } i} \omega_{F,j}(g) + \frac{1}{2} \sum_{j \text{ daughter of } i} \omega_{M,j}(g), \quad (\text{A.15})$$

が得られる。

ここで各個体の両親は全て 1 世代昔のオスとメスからランダムに一体ずつ選ぶので、オスの個体が持つ息子数の確率分布及び娘数の確率分布は共に二項分布になる。

つまり、オスが k 頭息子を持つ確率を $p_{mm}(k)$ 、 k 頭娘を持つ確率を p_{mf} とすると、 $N(g)$ 及び性比 $f(g)$ はそれぞれ N , f で世代を通して固定されていることを考慮すると、

$$p_{mm} = {}_fN C_k \left(\frac{1}{fN} \right)^k \left(1 - \frac{1}{fN} \right)^{fN-k}, \quad (\text{A.16})$$

$$p_{mf} = {}_{(1-f)N} C_k \left(\frac{1}{fN} \right)^k \left(1 - \frac{1}{fN} \right)^{(1-f)N-k}, \quad (\text{A.17})$$

となる。同様の計算を行い、メスが k 頭息子を持つ確率 p_{fm} 、 k 頭娘を持つ確率 p_{ff} も求めることができる。

ここで、 N が非常に大きい場合、すなわち $N \rightarrow \infty$ の場合を考える。この極限において、 p_{mm} , p_{mf} , p_{fm} , p_{ff} は以下のように変形される。

$$p_{mm}(k, f) = p_{ff}(k, f) = \frac{e^{-1}}{k!}, \quad (\text{A.18})$$

$$p_{mf}(k, f) = p_{fm}(k, 1-f) = \frac{e^{-\frac{1-f}{f}}}{k!} \left(\frac{1-f}{f} \right)^k. \quad (\text{A.19})$$

$f = \frac{1}{2}$ の時は、Derrida 理論で使われている確率分布となる。

ここで、以下のような母関数を定義する。

$$G_g(\lambda) \equiv \int_0^\infty e^{-\lambda\omega_M} P_M(\omega_M, g) d\omega_M, \quad (\text{A.20})$$

$$H_g(\mu) \equiv \int_0^\infty e^{-\mu\omega_F} P_F(\omega_F, g) d\omega_F. \quad (\text{A.21})$$

この二つの関数を用いると、式 (A.14) は、

$$G_g(\lambda) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} p_{mm}(k) \left[G_g \left(\frac{\lambda}{2} \right) \right]^k p_{mf}(j) \left[H_g \left(\frac{\lambda}{2} \frac{f}{1-f} \right) \right]^j, \quad (\text{A.22})$$

と変形でき、式 (A.15) は、

$$H_g(\mu) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} p_{fm}(k) \left[G_g \left(\frac{\mu}{2} \frac{1-f}{f} \right) \right]^k p_{ff}(j) \left[H_g \left(\frac{\mu}{2} \right) \right]^j, \quad (\text{A.23})$$

と変形できる。式 (A.18) と式 (A.19) を式 (A.22) と式 (A.23) に代入、変形を行い次式を得る。

$$G_{g+1}(\lambda) = \exp \left[-\frac{1}{f} + G_g \left(\frac{\lambda}{2} \right) + \frac{1-f}{f} H_g \left(\frac{\lambda}{2} \frac{f}{1-f} \right) \right], \quad (\text{A.24})$$

$$H_{g+1}(\mu) = \exp \left[-\frac{1}{1-f} + \frac{f}{1-f} G_g \left(\frac{\mu}{2} \frac{1-f}{f} \right) + H_g \left(\frac{\mu}{2} \right) \right]. \quad (\text{A.25})$$

次に式 (A.24) と式 (A.25) を、世代 $g \rightarrow \infty$ の定常状態で考えた時、以下の式のようになる。

$$G(\lambda) = \exp \left[-\frac{1}{f} + G \left(\frac{\lambda}{2} \right) + \frac{1-f}{f} H \left(\frac{\lambda}{2} \frac{f}{1-f} \right) \right], \quad (\text{A.26})$$

$$H(\mu) = \exp \left[-\frac{1}{1-f} + \frac{f}{1-f} G \left(\frac{\mu}{2} \frac{1-f}{f} \right) + H \left(\frac{\mu}{2} \right) \right]. \quad (\text{A.27})$$

タウバー型定理を用いると、 $\lambda, \mu \rightarrow \infty$ で、 $r_M, r_F \rightarrow 0$ に対応するので、式 (A.26) と式 (A.27) において $\lambda, \mu \rightarrow \infty$ の場合を考えれば、過去の世代で主個体の木に表れない個体の割合、つまり非先祖率を考えることができる。この時、 $G_0 = G(\infty)$, $H_0 = H(\infty)$ とすると、以下の2式を得る。

$$G_0 = \exp \left(-\frac{1}{f} + G_0 + \frac{1-f}{f} H_0 \right), \quad (\text{A.28})$$

$$H_0 = \exp \left(-\frac{1}{1-f} + \frac{f}{1-f} G_0 + H_0 \right). \quad (\text{A.29})$$

この2式は数値的に解かれ、その解の挙動は図 2.7 で描かれている。

付 録 B 関数作成法による森

第 4 章の 2 節で、性比-仔数分布可変モデルによる森を作る時に、DCN 法により G 世代の個体ごとに持つ仔数を割り振り、その仔数の和で $G - 1$ 世代の個体数を決めていたが、本研究で採用された DCN 法の他に森における各個体に仔数を割り振るために関数作成法を用いる方針も考えられた。ここでは関数作成法と、その特徴を述べる。

関数作成法では、与えられた関数に従う乱数を用いられるが、この乱数は逆関数法と呼ばれる方法で一様乱数から作られる [13]。逆関数法による与えられた乱数を作成する方法を説明する。まず、 $[0, 1]$ 上の一様分布に従う一様乱数を考える。一様乱数列を $u_i (i = 0, 1, 2, \dots)$ とする。

ここで、離散確率分布関数 $f(x)$ を考える。 $f(x)$ は、

$$\sum_{x=0}^{\infty} f(x) = 1, \quad (\text{B.1})$$

を満たしている。区間 $[0, \infty)$ 内における任意の整数を $x_j, x_{j+1} (x_j < x_{j+1}) (j = 0, 1, 2, \dots)$ とする。 u_i に対して、

$$\sum_{x=0}^{x_j} f(x) \leq u_i < \sum_{x=0}^{x_{j+1}} f(x), \quad (\text{B.2})$$

を満たす x_j を v_j とすると、 v_j は関数 $f(x)$ に従う乱数列になる。この関係を図で表すと図 B.1 のようになる。これが逆関数法である。

こうして作られた与えられた仔数分布に従う乱数を用い、各個体に仔数を与えていくのが関数作成法である。関数作成法による森の一つ一つを見ると、世代に対する個体増加率等で揺らぎが大きくなってしまう。図 B.2 は関数作成法によって作られた 10 個の森の世代に対する個体増加率の平均を示したものである。本研究では個体増加率の揺らぎが少ない DCN 法を用いて森を作成した。

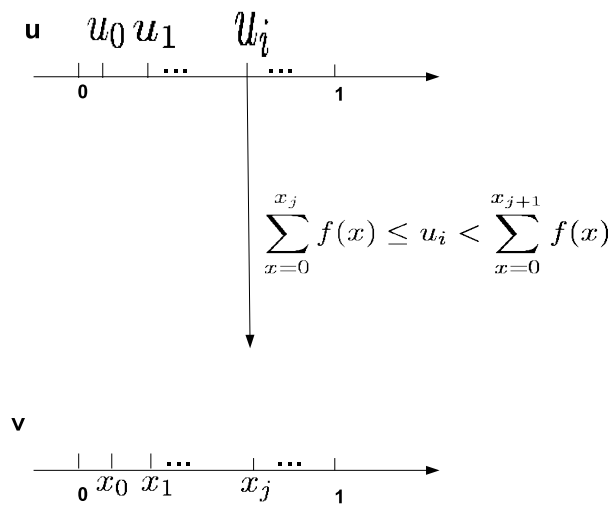


図 B.1: 逆関数法の説明。

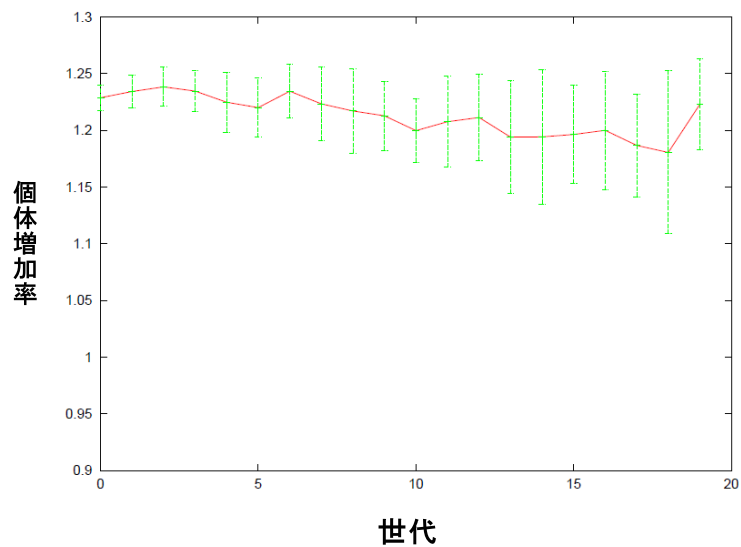


図 B.2: 乱数による仔数分布で作られた森の個体増加率。横軸が世代。縦軸が個体数増加率。個体数増加率は $\frac{N(G+1)}{N(G)}$ ($G = 0, 1, \dots, 19$) で計算される。性比は 0.5, $P_m(k)$, $P_f(k)$ 共に平均 2.5 のポアソン分布である。赤線が個体増加率の平均値、緑色の縦線が標準偏差を表す。

付 録 C Fisher による性比の議論

6 章で紹介された Fisher の性比に関する議論について述べる。多くの生物の性比が $1:1$ 、すなわち $f = 0.5$ になる理由は、1 個体当たりのオスとメスが残すことができる仔数の期待値が性比と逆の比率になるからである。ある世代のメス：オスを $1:x$ とし、総個体数を N とする。メスの総数を N_F とすると、

$$N_F = \frac{N}{1+x}, \quad (\text{C.1})$$

となり、総産仔数を C 、メスの産仔数の平均を C_F とすると、

$$C = \frac{C_F N}{1+x}, \quad (\text{C.2})$$

である。今、オスの総数を N_M とすると、

$$N_M = \frac{xN}{1+x}, \quad (\text{C.3})$$

であるから、オス 1 個体が残すことのできる仔の数の平均は、式 (C.2) と式 (C.3) から $\frac{C_F}{x}$ となる。メス 1 個体が残す仔の数は C_F なので、メスとオスが残す 1 個体当たりの仔数は、 N や C_F に関わらず $x:1$ になり、これは性比の逆になっている。つまり、メス対オスの比率に対して、メスとオスの残す仔数が逆比例する。したがって少ない性別の方が多くの仔を残すことになり、自然選択は生存または繁殖に不利な対立遺伝子は取り除くようにかかるので、自然選択は少ない方の性別を増やす方向にかかる。この結果、次第に少ない方の性別が増えていくが、これが過半数を超えると、その時点で自然選択が少数派になったもう一方の性別の数が有利になる方向にかかる。これを繰り返すことにより、オスとメスの数比が、ちょうど時計の振り子が真ん中でとまるように $1:1$ で安定するようになる。これは Fisher 性比と呼ばれている。Fisher 性比の詳しい証明は参考文献 [15] を参照されたい。

参考文献

- [1] 国連人口基金東京事務所 http://www.unfpa.or.jp/p_graph.html.
- [2] B. Derrida, S. C. Manrubia and D. H. Zanette, Phys. Rev. Lett. (1999) **82**, 1987-1990.
- [3] B. Derrida, S. C. Manrubia and D. H. Zanette, J. Theor. Biol. (2000) **203**, 303-305.
- [4] P. De L. Rios and O. Pla, Phys. Rev. E. (2000) **61**, 5620-5623.
- [5] 水口毅, 西村麻衣子, 数理解析研究所講究録 (2008) **1597**, 191-197.
- [6] 水口毅, 堀内陽介, 守田智, 数理解析研究所講究録 (2009) **1663**, 11-13.
- [7] 西村麻衣子, 大阪府立大学卒業論文 (2006).
- [8] 堀内陽介, 大阪府立大学卒業論文 (2009).
- [9] 日本生態学学会編, 生態学入門, 東京化学同人 (2004).
- [10] 堀内陽介, 水口毅, 守田智, 数理解析研究所講究録 (2010) **1704**, 61-67.
- [11] Pedigree Online Thoroughbred Database <http://www.pedigreequery.com/>.
- [12] S.Morita and T.Mizuguchi, private communication.
- [13] William H. Press, William T. Vetterling, Saul A. Teukolsky, Brian P. Flannery, *NUMERICAL RECIPES in C* [日本語版], 技術評論社 (1993).
- [14] M. E. J. NEWMAN, Contemporary Physics(2005), **46**, 323-351.
- [15] 山村則男, 動物—その適応戦略と社会 繁殖戦略の数理モデル, 東海大学出版会 (1986).