

ミトコンドリア・イブの年代推定に関わる分枝過程

金沢大学大学院教育学研究科数学教育専攻
石川県立高浜高等学校 教諭 大西 誠

1 研究の目的

分子生物学の近年の進歩は目覚ましく、遺伝現象の仕組みが解明されつつある。その一方で、対象の本質的側面を数式によって表現し解析するという数理モデルによる研究が、生命現象を解明する上で着実に重要性を増してきた。本研究では、ミトコンドリア・イブの年代を推定する分枝過程について、数理モデルの推定値とシミュレーションの結果との比較を通して、一生命現象のモデル化について考察することである。

2 ミトコンドリア・イブとは

2.1 アフリカ単一発生説

約 20 万年前にアフリカから移動した人々が全世界に広まったとする説。

2.2 多地域進化説

少なくとも 100 万年以上前にアフリカから移動した人々が各地に居住し、それぞれの場所で連続して進化したという説。

2.3 ミトコンドリア・イブ

細胞内のミトコンドリアにある DNA は母親からだけ遺伝する。母方の系統に焦点を当ててさかのぼっていけば、あるミトコンドリア DNA 配列を持つ女性に収束できると考える。Wilson と Cann は、現代人の祖先は約 20 万年前にアフリカで生存していた女性であると結論づけた。

3 アフリカ単一発生説への批判

1. 氷河期などの環境変動を考慮していない。
2. 年代推定をする際に用いる分子時計の信頼度に問題がある。

相違率は、一定か？

4 数理モデルによる年代推定

1995 年に Neil O'Connell はミトコンドリア・イブの年代を推定する数理モデルを発表した。O'Connell の数理モデルはマルコフ分枝過程を用いている。年代推定のみを行い、出生地については考えていない。

4.1 O'Connell の数理モデル

$Z_K(K)$: 現代女性人口。 σ^2 : 出生数の分散。

T_a : イブの死亡時から現在までの年数。

λ : 女性の平均寿命。 α : 成長パラメーター。

\bar{d}_n : 現代人の DNA の平均相違度。

δ : DNA の相違率。

$\hat{\alpha}$ と \hat{T}_a についての連立方程式

$$\begin{cases} Z_K(K) = \frac{\sigma^2 \hat{T}_a}{\lambda \hat{\alpha}} (e^{\hat{\alpha}} - 1), \\ \hat{T}_a = \frac{\bar{d}_n}{\delta \gamma(\hat{\alpha})} \end{cases}$$

を解けばよい。

$$\gamma(\alpha) = 1 - 2\alpha^{-1} \int_0^1 \frac{u}{(1-u)^3(u+\kappa)} [1 - u^2 + 2u \log u] du.$$

$$\kappa = \frac{e^{-\alpha}}{1 - e^{-\alpha}}.$$

4.2 O'Connell の推定値

現代女性人口: 10 億人。出生数の分散: 0.2。

女性の平均寿命: 25 年。DNA の平均相違度: 0.028。

$\delta = 1.8 \times 10^{-8}$ のとき, $\hat{T}_a = 1706103$, $\hat{\alpha} = 11.33$,

$\delta = 7.0 \times 10^{-8}$ のとき, $\hat{T}_a = 427945$, $\hat{\alpha} = 15.31$,

$\delta = 14.0 \times 10^{-8}$ のとき, $\hat{T}_a = 213282$, $\hat{\alpha} = 16.06$ 。

5 シミュレーション

O'Connell はシミュレーションを行っていないため、シミュレーションを行うことは意味のあることである。

5.1 仮定

1. 出産可能な女性だけを考える。
2. イブに 2 人の娘がいたとして、人口が 10 億人を超えるまでに何世代かかるかを調べる (1 世代 25 年)。
3. 各女性は、0 人から 3 人の娘をランダムに生み、その確率は p_0, p_1, p_2, p_3 。
4. 2002 年 出生動向基本調査を参考にし、 $p_0 = 0.0785$, $p_1 = 0.7275$, $p_2 = 0.172$, $p_3 = 0.021$ を基本。
5. 女性人口 10 億人としてシミュレーションしたいが …… 女性人口 1 千万人のとき, $\hat{T}_a = 443099$, $\hat{\alpha} = 10.27$ 。

5.2 O'Connell のモデル

途中で絶滅せずに 1 千万人を越えたのは 20000 回中 214 回で、結果は次の通りである。

	最大値	最小値	平均	標準偏差
世代数	28734	15903	20047	2151
年数	718350	397575	501175	***

5.3 成長パラメーターの揺らぎ

5.3.1 独立

氷河期などの環境変動を考慮していない。

⇒ 成長パラメーター α の揺らぎを考える。

成長パラメーター α が 1 世代ごとに独立にランダムに揺らぐ。

$\alpha = 10.27$ で固定 ⇒ $5.27 \leq \alpha \leq 15.27$ 。

結果的に、 p_0, p_1, p_2, p_3 が揺らぐ。

5.3.2 従属

成長パラメーター α が世代に従属にランダムに揺らぐ。

γ を $\{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$ に値をとる確率変数とする。

推移確率行列は、

$$\begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

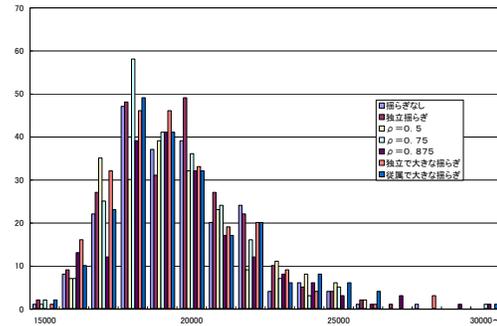
10 万回の推移で定常状態になるとし、このとき世代間の相関係数 $\rho = 0.5$ 。

$\alpha = 10.27 + \gamma$ を成長パラメーターとする。

5.3.3 結果

途中で絶滅せずに 1 千万人を越えたときの世代数についての結果は次の通りである。

成長パラメーター	揺らぎなし	独立	$\rho = 0.5$	$\rho = 0.75$	$\rho = 0.875$
20000 回中	214	237	203	225	189
最大値	28734	27317	26458	31112	31828
最小値	15903	15859	15610	15900	16261
平均	20047	20083	20065	19865	20164
標準偏差	2151	2152	2283	2089	2478



5.4 相違率の揺らぎ

分子時計 (相違率) が一定とするのは無理がある。

⇒ 相違率の揺らぎを考える。

5.4.1 結果

途中で絶滅せずに 1 千万人を越えたときの世代数についての結果は次の通りである。

相違率 ($\times 10^8$)	2.0	4.5	独立揺らぎ	従属揺らぎ	7.0
20000 回中	52	124	127	159	214
最大値	94918	45999	47183	46228	28734
最小値	57552	25056	25671	25384	15903
平均	71737	31073	31131	31349	20047
標準偏差	9522	3554	3579	4078	2151

6 結論

- O'Connell の数理モデルは、アフリカ単一発生説と多地域進化説のどちらか一方を特に支持するものではない。
- O'Connell の数理モデルの推定値はシミュレーションの結果にもほぼ一致している。
- 環境変動を考慮しても、シミュレーションの結果は平均において変化はあまり認められない。
- 相違率の揺らぎは、イブの年代の標準偏差を大きくし、分布をより平らにする。

7 課題

- 数学、生物、地理、歴史、地学、情報などの教科・科目に関わる総合的な学習の時間の教材を作成する。
- シミュレーションの結果の統計的な検討を詳しく行う。
- 成長パラメーターや置換率の揺らぎについて、数理的に考察することができるか検討する。
- O'Connell の推定値の区間推定のための極限定理を求める。