

集団遺伝学の基礎

大槻 久

(2024年9月5日版)

目次

0 基本事項	2
0.1 生物学の基本用語	2
0.2 よく用いられる数学	2
1 対立遺伝子の進化ダイナミクス	5
1.1 一倍体の進化ダイナミクス	5
1.1.1 自然選択の効果	5
1.1.2 突然変異の効果	8
1.1.3 自然選択と突然変異の効果	9
1.2 二倍体の進化ダイナミクス	12
1.2.1 Hardy-Weinberg 平衡	12
1.2.2 自然選択の効果	13
1.2.3 自然選択と突然変異の効果	16
1.3 二座位一倍体モデル	17
1.3.1 連鎖不平衡	18
1.3.2 自然選択とヒッチハイク	20
1.3.3 相加性とエピスタシス	23
2 離散表現型の進化モデル	24
2.1 表現型 gambit	24
2.2 レプリケーター方程式	25
2.3 具体例	28
2.3.1 タカ・ハトゲーム	28
2.3.2 一般の2戦略ゲーム	30
2.3.3 非線形公共財ゲーム	31
2.4 突然変異がある場合	33
2.4.1 Replicator-mutator 方程式	33
2.4.2 エラーカタストロフ理論	34
3 量的遺伝学入門	36
3.1 QTL と量的遺伝学	36
3.2 相加遺伝値	37
3.3 顕性偏差	39
3.4 表現型分散の分解	42
3.5 表現型の類似度	43
3.5.1 中間親への回帰と遺伝率	43
3.5.2 きょうだい間の相関と双生児法	46
3.6 選択差と選択応答	50
3.7 多座位の場合	52
4 連続形質の進化モデル	55
4.1 Price 方程式	55
4.2 Price 方程式の分解	59
4.3 分布の発展方程式	62
4.4 正規分布を用いた理論	63
4.5 頻度依存選択の扱い	65

0 基本事項

0.1 生物学の基本用語

- 遺伝子 (gene) とは、生物学的機能を果たす遺伝情報の単位で、その実態は ATCG の塩基配列、出力はタンパク質である。
- 座位 (locus) とは、ある遺伝子が位置する染色体上の位置である。
- 対立遺伝子もしくはアレル (allele) とは、同じ座位に存在し得る複数の遺伝子のことである。例えば ABO 式血液型で、遺伝子 A と B と O は対立遺伝子である。
- ゲノム (genome) とは、染色体のセットの最小単位であり、同時に遺伝情報の総体の最小単位を指す。ヒトの場合 23 本の染色体 (およびそこに存在する遺伝情報) を指す。
- 一倍体もしくは半数体 (haploid) とは、ゲノム 1 セットからなる生命体のことである。
- 二倍体 (diploid) とは、ゲノム 2 セットからなる生命体のことである。ヒトは二倍体である。
- 減数分裂 (meiosis) とは、染色体数が半分になるような細胞分裂のことである。ヒトでは卵や精子を作る時に行われる。
- 配偶子 (gamete) とは生殖細胞であって、接合して次世代を作るもののことを指す。ヒトの配偶子は卵や精子である。これらは一倍体である。
- 接合子もしくは接合体 (zygote) とは、配偶子が融合してできた細胞である。ヒトの接合体は狭義には受精卵であり、これは二倍体である。ヒト個体も広義には接合体である。
- 相同遺伝子 (homologous gene) とは、相同な染色体に存在する遺伝子のことである。例えば ABO 式血液型で、AB 型の人は相同遺伝子 A と B を持つ。
- 遺伝子型 (genotype) とは、個体が持つ遺伝子の組み合わせのことを言う。前述の血液型の例で、AO の所有者と OA の所有者は、ともに同じ AO の遺伝子型を持つと言う。特に遺伝子頻度 (gene frequency) と遺伝子型頻度 (genotype frequency) の違いに注意。
- ホモ接合体 (homozygote) とは、二倍体生物において相同遺伝子が等しいもののことを指す。前述の血液型の例では、AA や BB の個体がそうである。
- ヘテロ接合体 (heterozygote) とは、二倍体生物において相同遺伝子が異なるもののことを指す。前述の血液型の例では、AB や AO の個体がそうである。
- 形質 (trait) とは生物が示す生物学的特徴のことを言う。
- 表現型 (phenotype) とは、生物に実際に現れる形質を指す。前述の血液型の例では、AA 個体ならびに AO 個体の表現型はともに A 型である。
- ヘテロ接合体 XY の表現型が、ホモ接合体 XX のそれと同じ場合、遺伝子 X を顕性 (dominant) (優性)、Y を潜性 (recessive) と言う。
- ハプロタイプ (haplotype) とは遺伝子の並びにおいて、片親から受け継いだ並びのことを言う。
- 組み換え (recombination) もしくは遺伝的組み換え (genetic -) とは、遺伝子の並びがその途中で組み変わることを指す。主に相同染色体の間で行われる。
- 個体の進化 (evolution) とは、集団内の遺伝子頻度が変化することを指す。
- 個体の適応度 (fitness) とは、その個体が一生のうちに残す子の数のうち、次世代の繁殖プールに参加できる個体の数を指す。
- 自然選択もしくは自然淘汰 (natural selection) とは、適応度の差に起因して起きる進化のことを指す。

0.2 よく用いられる数学

- x の関数 $f(x)$ の x_0 の周りの Taylor 展開 (Taylor expansion) とは、

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2!}f''(x_0)(x - x_0)^2 + \frac{1}{3!}f'''(x_0)(x - x_0)^3 + \dots \quad (0.1)$$

のことで、 $|x - x_0|$ が微小量の時に用いられる。特に $x_0 = 0$ の時は

$$f(x) = f(0) + f'(0)x + \frac{1}{2!}f''(0)x^2 + \frac{1}{3!}f'''(0)x^3 + \dots \quad (0.2)$$

の形をしており (Maclaurin 展開 (マクローリン展開; Maclaurin expansion) と呼ばれる)、 $|x|$ が微小量の時に用いられる。

これを用いることで、 $|x|$ が小さい時

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots \quad (0.3)$$

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots \quad (0.4)$$

$$\frac{1}{1+bx} = 1 - bx + b^2x^2 - b^3x^3 + \dots \quad (0.5)$$

$$\frac{1+ax}{1+bx} = 1 + (a-b)x + (-ab+b^2)x^2 + (ab^2-b^3)x^3 + \dots \quad (0.6)$$

などが導かれる。応用として、例えば $|x|, |y|, |z|, |w|$ が同程度に微小量なら

$$\frac{1+x+y}{1+z+w} \approx 1 + (x+y-z-w) \quad (0.7)$$

(\approx は近似の意味) なども分かる。なぜなら、各変数の「小ささ」の部分抜き出して、 $x = \delta x', y = \delta y', z = \delta z', w = \delta w'$ ($|\delta|$ は微小量、 x', y, z', w' はそうではない) とすると、

$$\frac{1+x+y}{1+z+w} = \frac{1+(x'+y')\delta}{1+(z'+w')\delta} \quad (0.8)$$

であり、この δ に対し公式 (0.6) を用いれば良いからである。

- n 個の量 X_1, \dots, X_n の平均 (average, mean) とは

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (0.9)$$

のことであり、分散 (variance) とは

$$\text{var}[X] = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (= \overline{(X - \bar{X})^2}) \quad (0.10)$$

のことである。分散は 2 乗平均と平均を用いて

$$\text{var}[X] = \left(\frac{1}{n} \sum_{i=1}^n X_i^2 \right) - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2 \quad (= \overline{X^2} - \bar{X}^2) \quad (0.11)$$

とも書ける。

- n 個の量の組 $(X_1, Y_1), \dots, (X_n, Y_n)$ の共分散 (covariance) とは

$$\text{cov}[X, Y] = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \quad (= \overline{(X - \bar{X})(Y - \bar{Y})}) \quad (0.12)$$

のことである。共分散は積の平均と平均の積を用いて

$$\text{cov}[X, Y] = \left(\frac{1}{n} \sum_{i=1}^n X_i Y_i \right) - \left(\frac{1}{n} \sum_{i=1}^n X_i \right) \left(\frac{1}{n} \sum_{i=1}^n Y_i \right) \quad (= \overline{XY} - \bar{X}\bar{Y}) \quad (0.13)$$

とも書ける。 X と X 自身の共分散は、 X の分散

$$\text{cov}[X, X] = \text{var}[X] \quad (0.14)$$

に等しい。

- a, b, c, d を定数とする時、以下の公式が成り立つ。

$$\overline{aX + b} = a\bar{X} + b \quad (0.15)$$

$$\text{var}[aX + b] = a^2 \text{var}[X] \quad (0.16)$$

$$\text{cov}[aX + b, cY + d] = ac \text{cov}[X, Y] \quad (0.17)$$

$$\overline{X + Y} = \bar{X} + \bar{Y} \quad (0.18)$$

$$\text{var}[X + Y] = \text{var}[X] + 2\text{cov}[X, Y] + \text{var}[Y] \quad (0.19)$$

$$\text{cov}[X + Y, Z + W] = \text{cov}[X, Z] + \text{cov}[X, W] + \text{cov}[Y, Z] + \text{cov}[Y, W] \quad (0.20)$$

- Y の X に対する**回帰係数 (regression coefficient)** とは

$$b_{YX} = \frac{\text{cov}[X, Y]}{\text{var}[X]} \quad (0.21)$$

で定義され、その時の**回帰直線 (regression line)** は

$$y - \bar{Y} = b_{YX}(x - \bar{X}) \quad (0.22)$$

で表される。

- X と Y の**相関係数 (correlation coefficient)** とは

$$\rho_{XY} = \frac{\text{cov}[X, Y]}{\sqrt{\text{var}[X]}\sqrt{\text{var}[Y]}} \quad (0.23)$$

で定義される。

- 平均 μ 、分散 σ^2 の**正規分布 (normal distribution)** の**確率密度関数 (probability density function)** は

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right] \quad (0.24)$$

で与えられる。

- 二つの**確率変数 (stochastic variable)** X, Y が独立であるとは、

$$\text{Prob}(X \leq x \text{ and } Y \leq y) = \text{Prob}(X \leq x) \cdot \text{Prob}(Y \leq y) \quad (0.25)$$

が全ての x, y に対して成り立つことを言う。

- 二つの確率変数 X, Y が独立ならば、期待値 (記号 \mathbb{E} で表す) に関して

$$\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y] \quad (0.26)$$

が成り立つ。さらに任意の関数 f, g に対して

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)] \cdot \mathbb{E}[g(Y)] \quad (0.27)$$

も成り立つ。

1 対立遺伝子の進化ダイナミクス

(本章の概要と目的) 本章では対立遺伝子が複数(主に二つ)ある場合に、それらの遺伝子の頻度が時間とともにどのように変わるかを見ていく。本章で扱うトピックは Fisher、Wright、Haldane らを中心に 1920-30 年代頃に形作られ、集団遺伝学の最も基礎的で古典的な部分であると言って良い。

時間を t とすると、離散時間 $t = 0, 1, 2, \dots$ の場合には遺伝子頻度の変化は漸化式の形で記述される。かたや連続時間 $t \geq 0$ の場合には、遺伝子頻度の変化は微分方程式で記述される。本章では漸化式を正面から扱うことはあまりせず、漸化式の近似として得られた微分方程式を主に解析していくことにする。

最初に最も単純な一倍体生物を考え、次により複雑な二倍体生物の理論を展開する。対立遺伝子が A_1, A_2 の二種類の場合、一倍体では個体は二種類 (A_1 個体と A_2 個体) しかいないが、二倍体では個体の種類は三種類 (A_1A_1 個体、 A_1A_2 個体、 A_2A_2 個体) になるため理論はより複雑になる。古典的な集団遺伝学がいつも二倍体を考えるのは、ヒトが二倍体であるからである。その後で二座位の問題、すなわち別の遺伝子座に対立遺伝子 B_1, B_2 が存在する場合の理論を展開する。可能な組み合わせはさらに多くなり、理論はより複雑になっていく。

本章の目的はまず数理モデルに親しむことにある。基本的な集団遺伝学の考え方を身につけ、その現象を数式で表せるようになることが第一目標である。また、自然選択、突然変異、遺伝的組換えといった現象が、遺伝子の頻度にどのような影響を与えるかをひと通り学ぶ。mutation-selection balance、Hardy-Weinberg 平衡、連鎖不平衡などの基礎概念も学ぶ。

1.1 一倍体の進化ダイナミクス

1.1.1 自然選択の効果

一倍体、離散世代モデルを考え、対立遺伝子 A_1 および A_2 の遺伝子頻度をそれぞれ $p, q (= 1 - p)$ とする。 p のダイナミクスについて考えよう。

世代 t における p の値を $p(t)$ と書く。以降、 $p(t)$ と $p(t+1)$ の関係を考えていることが明白である場合、それらを p, p' と書くことがある。

一倍体の場合は A_1, A_2 が遺伝子型そのものである。遺伝子型 A_1, A_2 の個体の適応度をそれぞれ w_1, w_2 と書く。この時、次世代における遺伝子 A_1 の頻度は

$$p' = \frac{w_1 p}{w_1 p + w_2 q} \quad (1.1)$$

と書ける。

解説 1.1 頻度とは「全個体数に占める、注目している性質を持つ個体数の割合」だから、分からなくなったら世代 t の総個体数を N とおいてしまって構わない(この N は最後には計算から消える)。この時、

世代 t の個体数		→	世代 $t+1$ の個体数
遺伝子型 A_1	Np	→	$Np \times w_1$
遺伝子型 A_2	Nq	→	$Nq \times w_2$

だから、次世代の A_1 の遺伝子頻度は

$$p' = \frac{(\text{世代 } t+1 \text{ の } A_1 \text{ 個体数})}{(\text{世代 } t+1 \text{ の } A_1 \text{ 個体数}) + (\text{世代 } t+1 \text{ の } A_2 \text{ 個体数})} = \frac{Npw_1}{Npw_1 + Nqw_2}$$

であり、 N を消すと式 (1.1) を得る。

式 (1.1) は適応度 w_1, w_2 を一斉に定数倍しても不変である。別の言い方をすれば、式 (1.1) は適応度の比 w_1/w_2 のみに依存している。これは

$$p' = \frac{\left(\frac{w_1}{w_2}\right) p}{\left(\frac{w_1}{w_2}\right) p + 1 \cdot q} \quad (1.2)$$

と変形できることから分かる。したがって適応度そのものではなく、適応度の相対値を用いても結果は変わらない。このような相対値を**相対適応度 (relative fitness)**と呼ぶ。集団遺伝学ではしばしば w_2 を基準に取り、相対適応度を $w_1 = 1 + s, w_2 = 1$ とすることがある。この s を遺伝子型 A_1 の**選択有利度 (selective advantage)**と呼ぶ。式 (1.1) にこれら相対適応度を代入して

$$p' = \frac{(1+s)p}{(1+s)p+q} = \frac{(1+s)p}{1+sp} \quad (1.3)$$

を得る。

例 1.2 $s = 0$ ならば式 (1.3) は

$$p' = p \quad (1.4)$$

となりこれは、対立遺伝子 A_1 と A_2 が中立である時に、遺伝子頻度が世代に渡って変化しないことを表す。これはある意味当然の帰結であるが、上のモデルでは暗に集団の大きさが無限であることを仮定しているため、確率的ゆらぎが完全に無視でき、次世代の遺伝子頻度が決定論的に決まるからである。これに対し**中立進化 (neutral evolution)**の理論では有限集団を仮定するためこの事実は成り立たず、次世代の遺伝子頻度 p' は一意に決まるのではなく、ある確率分布に従う。

式 (1.3) は陽に解くことのできる数少ない差分方程式である。両辺の逆数を取って

$$\frac{1}{p(t+1)} = \left(\frac{1}{1+s}\right) \frac{1}{p(t)} + \frac{s}{1+s} \quad (1.5)$$

であるので、

$$\frac{1}{p(t+1)} - 1 = \left(\frac{1}{1+s}\right) \left(\frac{1}{p(t)} - 1\right) \quad (1.6)$$

とでき、したがって数列 $\{\frac{1}{p(t)} - 1\}_{t=0,1,\dots}$ は等比数列なので

$$\frac{1}{p(t)} - 1 = \left(\frac{1}{1+s}\right)^t \left(\frac{1}{p(0)} - 1\right) \quad (1.7)$$

である。整理して

$$p(t) = \left[\left(\frac{1}{1+s}\right)^t \left(\frac{1}{p(0)} - 1\right) + 1 \right]^{-1} \quad (1.8)$$

を得る。

式 (1.3) において $p(t)$ は、 $s > 0$ の時 $p(\infty) = 1$ に収束する。反対に $s < 0$ の時 $p(\infty) = 0$ に収束する。これらは期待通りの結果と言える。

例 1.3 数値例を考えよう。遺伝子 A_1 の初期頻度を 10% とし、選択有利度を 1% とする。すなわち $p(0) = 0.1, s = 0.01$ である。この時

$$\begin{aligned} p(10) &= 0.109318 & p(100) &= 0.231086 & p(200) &= 0.448395 \\ p(300) &= 0.687375 & p(400) &= 0.856056 & p(500) &= 0.941472 \end{aligned} \quad (1.9)$$

となる。

解説 1.4 式 (1.3) を s の絶対値が小さい時に微分方程式で近似することを考えてみよう。まず差分の形

$$p' - p = \frac{(1+s)p}{1+sp} - p = \frac{sp(1-p)}{1+sp} \quad (1.10)$$

にする。 $|s| \ll 1$ の時、式 (1.10) を s に関して Taylor 展開し、 s の 1 次の項までを考える。式 (0.4) より

$$\frac{1}{1+sp} = 1 - sp + (sp)^2 - \dots \quad (1.11)$$

だから

$$\begin{aligned} p(t+1) - p(t) &= sp(1-p) \{1 - sp + (sp)^2 - \dots\} \\ &= sp(1-p) - s^2p^2(1-p) + s^3p^3(1-p) - \dots \\ &\approx sp(1-p) \end{aligned} \quad (1.12)$$

を得る。ただし、上の最後の式変形で、 s^2 や s^3 などのつく項は全て無視した。なぜなら $|s|$ が小さい時には $|s^2|$ や $|s^3|$ などは s に比べて十分に無視できるほど小さいからである。

ここで、式 (1.12) の左辺を

$$\frac{p(t + \Delta t) - p(t)}{\Delta t} \quad (\text{ただし } \Delta t = 1) \quad (1.13)$$

とみなすと、これは微分

$$\lim_{\Delta t \rightarrow 0} \frac{p(t + \Delta t) - p(t)}{\Delta t} = \frac{dp(t)}{dt} \quad (1.14)$$

の近似であると考えられるから、式 (1.3) は最終的に微分方程式

$$\frac{dp(t)}{dt} = sp(1 - p) \quad (1.15)$$

で近似される。微分方程式 (1.15) においてはもはや t は整数とは限らず、実数の値を取り得ることに注意せよ。この微分方程式は変数分離法により陽に解け、解は

$$p(t) = \left[e^{-st} \left(\frac{1}{p(0)} - 1 \right) + 1 \right]^{-1} \quad (1.16)$$

となる。この式に $p(0) = 0.1, s = 0.01$ を代入して計算すると

$$\begin{aligned} p(10) &= 0.109367 & p(100) &= 0.231969 & p(200) &= 0.450853 \\ p(300) &= 0.690568 & p(400) &= 0.858486 & p(500) &= 0.942826 \end{aligned} \quad (1.17)$$

を得る。これは式 (1.9) の良い近似値を与えている。

解説 1.5 式 (1.15) をより一般の適応度関数に関して一般化してみることを考えよう。適応度を $(w_1, w_2) = (1 + s_1, 1 + s_2)$ とする。この時、先程と同じ考え方により差分方程式

$$p' = \frac{(1 + s_1)p}{(1 + s_1)p + (1 + s_2)q} = \frac{(1 + s_1)p}{1 + s_1p + s_2q} \quad (1.18)$$

が得られる。ここで $|s_1|, |s_2|$ は十分小さいとして右辺を Taylor 展開する。式 (0.4) より

$$\frac{1}{1 + s_1p + s_2q} = 1 - (s_1p + s_2q) + (s_1p + s_2q)^2 - \dots \quad (1.19)$$

だから

$$\begin{aligned} p' &= p(1 + s_1) \{ 1 - (s_1p + s_2q) + (s_1p + s_2q)^2 - \dots \} \\ &= p \{ 1 - (s_1p + s_2q) + (s_1^2p^2 + 2s_1s_2pq + s_2^2q^2) - \dots \\ &\quad + s_1 - (s_1^2p + s_1s_2q) + (s_1^3p^2 + 2s_1^2s_2pq + s_1s_2^2q^2) - \dots \} \\ &\approx p \{ 1 + s_1 - (s_1p + s_2q) \} \end{aligned} \quad (1.20)$$

を得る。最後の式変形で s_1^2 や s_1s_2 や s_2^2 など、 s が 2 回以上かけ算として現れる項は全て無視した。上式の右辺から左辺に p の項を移項して微分方程式近似をすることで

$$\frac{dp}{dt} = p \{ s_1 - (s_1p + s_2q) \} \quad (1.21)$$

もしくは中括弧の中をさらに計算して $s_1 - (s_1p + s_2q) = (1 - p)s_1 - qs_2 = (1 - p)(s_1 - s_2)$ だから

$$\frac{dp}{dt} = p(1 - p)(s_1 - s_2) \quad (1.22)$$

を得る。式 (1.22) に $(s_1, s_2) = (s, 0)$ を代入すれば、式 (1.15) が得られる。

式 (1.21) の中括弧の中の中 $\bar{s} \equiv s_1p + s_2q$ は集団平均の「 s の値」を表しており、 s_1 が集団平均の「 s の値」より大きい時に、遺伝子 A_1 の頻度が増加することを表している。同じことだが、式 (1.22) は $s_1 > s_2$ である時に遺伝子 A_1 の頻度が増加することを表している。これらの形は後ほど学ぶレプリケーター方程式 (2.6) にも現れる。

1.1.2 突然変異の効果

自然選択がなく、突然変異がある状況を考える。

遺伝子 A_1 を持つ個体は複製時に確率 u で A_2 の子を生じるとする。反対に、 A_2 を持つ個体は複製時に確率 v で A_1 の子を生じるとする。以下、 $0 < u, v < 1$ を仮定する。互いに適応度に差はなく、したがって相対適応度は $w_1 = w_2 = 1$ とする。この時、子の数は以下のように計算できる。

世代 t の個体数		→	世代 $t+1$ の個体数		子の遺伝子型
遺伝子型 A_1	Np	→	$Np \times 1$	$\left\{ \begin{array}{l} \xrightarrow{\text{確率 } u} \\ \xrightarrow{\text{確率 } 1-u} \end{array} \right.$	A_2
					A_1
遺伝子型 A_2	Nq	→	$Nq \times 1$	$\left\{ \begin{array}{l} \xrightarrow{\text{確率 } v} \\ \xrightarrow{\text{確率 } 1-v} \end{array} \right.$	A_1
					A_2

上の図から p のダイナミクスは

$$p' = \frac{p(1-u) + qv}{p+q} = (1-u)p + v(1-p) \quad (1.23)$$

と書ける。

簡単な計算から p は初期値にかかわらず

$$p^* = \frac{v}{u+v} \quad (1.24)$$

に収束することが分かる。 p^* は

$$p^* = \frac{(\text{A}_1 \text{ を生じる突然変異率})}{(\text{全ての突然変異率の和})} \quad (1.25)$$

という形をしており、平衡頻度は突然変異率のバランスで決まることが分かる。

解説 1.6 式 (1.23) を

$$p' - \frac{v}{u+v} = (1-u-v) \left(p - \frac{v}{u+v} \right) \quad (1.26)$$

と変形すると、数列 $\{p(t) - \frac{v}{u+v}\}_{t=0,1,\dots}$ は等比数列であることが分かるので、

$$p(t) - \frac{v}{u+v} = (1-u-v)^t \left(p(0) - \frac{v}{u+v} \right) \quad (1.27)$$

である。 $-1 < 1-u-v < 1$ であるので、 $t \rightarrow \infty$ の極限を取ることで、

$$p(\infty) = \frac{v}{u+v} \quad (= p^*) \quad (1.28)$$

が分かる。

解説 1.7 式 (1.23) の平衡点の位置 p^* を求めるだけならば、式 (1.23) の p, p' を p^* で置き換えた

$$p^* = (1-u)p^* + v(1-p^*) \quad (1.29)$$

を解けばよく、式 (1.24) を得る。

次に平衡点 p^* の局所安定性解析を行う。式 (1.23) を

$$p' = F(p) \quad (1.30)$$

と見る。ただし $F(p) = (1-u)p + v(1-p)$ である。 p^* が局所安定であるための十分条件は

$$-1 < F'(p^*) < 1 \quad (1.31)$$

であることが知られている（離散力学系における局所安定性解析の理論）。今の場合 $F'(p^*) = 1-u-v$ であり、これは $-1 < F'(p^*) < 1$ を満たすから、 p^* は局所安定である。

今の場合、 $F(p)$ は p の線型関数であるからさらに強いことが言え、 p^* は局所安定だけでなく大域安定である。つまり、どんな初期値から始めても p の値は p^* に収束する。

解説 1.8 先程と同じように式 (1.23) は微分方程式で近似できる。実際、

$$p' - p = -up + v(1 - p) \quad (1.32)$$

であるから、これを先程と同じように微分方程式で近似して

$$\frac{dp}{dt} = -up + v(1 - p) \quad (1.33)$$

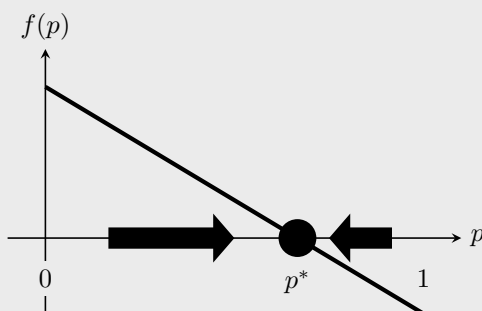
である。この微分方程式を $\frac{dp}{dt} = f(p)$ と見る。平衡点 p^* の位置は $f(p^*) = 0$ を解けばよく、 $p^* = v/(u+v)$ である。平衡点が局所安定であるための十分条件は

$$f'(p^*) < 0 \quad (1.34)$$

であることが知られている（力学系における局所安定性解析の理論）。今の場合

$$f'(p^*) = -(u+v) < 0 \quad (1.35)$$

だから p^* は局所安定である。さらに、 $f(p)$ は p の線型関数だから、 p^* は大域安定であることも言える。



1.1.3 自然選択と突然変異の効果

前述の自然選択と突然変異の効果を同時に考える。基本的な仮定は上と同じであるが、（慣習による理由から）先程と違い適応度を $w_1 = 1, w_2 = 1 - s$ とおくことにする。

解説 1.9 $(w_1, w_2) = (1 + s, 1)$ とおく場合と、 $(w_1, w_2) = (1, 1 - s)$ とおく場合とでは、 $|s|$ の値が小さい時は本質的には変わらない。

これがなぜかを理解するために、 $(w_1, w_2) = (1 + s, 1)$ の場合を考える。これらが相対適応度だったことを思い出すと、

$$w_1 = 1 + s, \quad w_2 = \frac{1}{1 + s} \quad (1.36)$$

としても同値である。しかしながら $|s|$ が小さい時は式 (0.4) で与えられる Taylor 展開

$$\frac{1}{1 + s} = 1 - s + s^2 - s^3 + \dots \quad (1.37)$$

が成り立つので、 s の一次の項まで考えれば結局は $(w_1, w_2) = (1, 1 - s)$ を考えているのと同じことになるからである。したがって、自然選択の効果を表す式は、 s の一次の項までを考える限りは、今までと同じになると予測できる。

子の数は以下のように計算できる。

世代 t の個体数		→	世代 $t+1$ の個体数		子の遺伝子型
遺伝子型 A_1	Np	→	$Np \times 1$	$\left\{ \begin{array}{l} \xrightarrow{\text{確率 } u} \\ \xrightarrow{\text{確率 } 1-u} \end{array} \right.$	遺伝子 A_2
					遺伝子 A_1
遺伝子型 A_2	Nq	→	$Nq \times (1-s)$	$\left\{ \begin{array}{l} \xrightarrow{\text{確率 } v} \\ \xrightarrow{\text{確率 } 1-v} \end{array} \right.$	遺伝子 A_1
					遺伝子 A_2

したがって上の図から p のダイナミクスは

$$p' = \frac{p(1-u) + q(1-s)v}{p + q(1-s)} = \frac{p(1-u) + (1-p)(1-s)v}{1-s(1-p)} \quad (1.38)$$

と書ける。

解説 1.10 差分方程式 (1.38) を陽に解くこともできるが煩雑であるから、微分方程式による近似を用いる。 $|s|$ と u と v は 1 より十分小さいと仮定し、式 (1.38) の右辺を s と u と v について Taylor 展開し、 s と u と v に関する 1 次の項までを残す。具体的には、式 (0.3) より

$$\frac{1}{1-s(1-p)} = 1 + s(1-p) + \underbrace{s^2}_{\approx 0}(1-p) + \dots \approx 1 + s(1-p) \quad (1.39)$$

であるから式 (1.38) は近似的に

$$\begin{aligned} p' &= \frac{p(1-u) + (1-p)(1-s)v}{1-s(1-p)} \approx \{p - up + (1-p)(v - vs)\} \{1 + s(1-p)\} \\ &= p - up + (1-p)(v - \underbrace{vs}_{\approx 0}) + sp(1-p) - \underbrace{us}_{\approx 0}p(1-p) + (1-p)^2(\underbrace{vs}_{\approx 0} - \underbrace{vs^2}_{\approx 0}) \\ &\approx p + sp(1-p) - up + v(1-p) \end{aligned} \quad (1.40)$$

とできることが分かった。さらに、微分方程式にすると

$$\frac{dp}{dt} = \underbrace{sp(1-p)}_{\text{自然選択}} - \underbrace{up}_{\text{変異 } A_1 \rightarrow A_2} + \underbrace{v(1-p)}_{\text{変異 } A_2 \rightarrow A_1} \quad (1.41)$$

となる。式 (1.41) の右辺は、自然選択のみを考えた式 (1.15) の右辺と突然変異のみを考えた式 (1.33) の右辺の、単なる足し算になっている。

解説 1.11 微分方程式 (1.41) は $|s|$ や u や v が十分小さいという仮定の下に導かれたのだから、例えば $(s, u, v) = (0.1, 0.05, 0.05)$ というパラメータを代入することはあっても、その 100 倍大きい $(s, u, v) = (10, 5, 5)$ などというパラメータを代入することは望ましくない (そもそも u, v は確率であったのだから、1 を超えることはナンセンスである)。

一方で、微分方程式の定性的な振る舞いのみに興味があり、時間がどれだけ進んだかにはあまり興味がない場合がある。例えば時間 t に対して 100 倍遅く進む時計 t' で微分方程式を考えることにしよう。すると $t' = t/100$ であり、この時計で見た時間微分は、chain-rule から

$$\frac{d}{dt'} = \frac{dt}{dt'} \frac{d}{dt} = 100 \frac{d}{dt} \quad (1.42)$$

となり、元の時間 t による微分の 100 倍の大きさになる。したがって、時間 t' で見た微分方程式 (1.41) は

$$\frac{dp}{dt'} = (100s)p(1-p) - (100u)p + (100v)(1-p) \quad (1.43)$$

であり、係数 (s, u, v) は $(100s, 100u, 100v)$ に変化する。この $(100s, 100u, 100v)$ に $(10, 5, 5)$ という値を入れることは何ら差し支えないので、以下ではこの事実を踏まえた上で、断りなしにこのような値を代入することがある。

解説 1.12 差分方程式と微分方程式では、一般に微分方程式のほうが解析が簡単である。なぜなら差分方程式では $p(t)$ と $p(t+1)$ の値は一般に離れているので、 $p(t+1)$ の値そのものを知る必要があるが、微分方程式では $p(t)$ と $p(t+\Delta t)$ の値は近いので、

$p(t)$ に比べて $p(t + \Delta t)$ の値が「増えたか減ったか」を知れば、あとは解の連続性により、方程式の（特に十分時間が経った後の）振る舞いを定性的には十分に理解できるからである。

例えば差分方程式では 1 次元でカオスが生じるが、微分方程式でカオスが生じるためには（自励系に限れば）少なくとも 3 次元が必要である。

以降、式 (1.41) の微分方程式モデルを考える。右辺を $f(p)$ とおく。

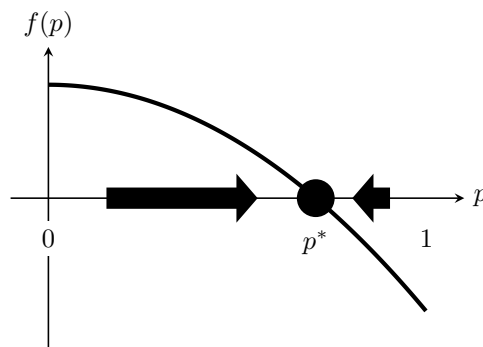
はじめに $u, v > 0$ の時を考えると、 $0 \leq p \leq 1$ での $f(p)$ の符号の変化は

p	0	...	p^*	...	1
$f(p)$	+	+	0	-	-

となる。平衡点 p^* の位置は

$$p^* = \frac{(s - u - v) + \sqrt{(s - u - v)^2 + 4sv}}{2s} \quad (1.44)$$

であり、増減表からこれは大域安定平衡点である（下図）。



次に、遺伝子 A_1 が遺伝子 A_2 に比べて機能的に優れ ($s > 0$)、しかしながら遺伝子 A_1 は突然変異により機能が損なわれ ($u > 0$)、反対に遺伝子 A_2 が機能を獲得するような突然変異は無視できる ($v = 0$) 場合を考える。突然変異は一般に機能を壊すほうに向かうので、これらの仮定は生物学的に妥当かつ重要である。この場合、式 (1.41) は

$$\frac{dp}{dt} = sp(1 - p) - up = p\{s(1 - p) - u\} \quad (1.45)$$

と書ける。再び右辺を $f(p)$ とおくと、 $u < s$ の時 $f(p)$ の符号は

p	0	...	p^*	...	1
$f(p)$	0	+	0	-	-

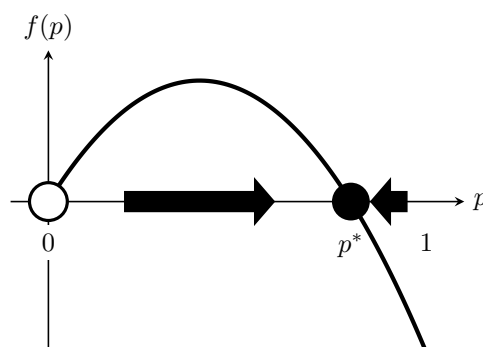
のように変化する。ここで

$$p^* = \frac{s - u}{s} \quad (1.46)$$

つまり、遺伝子 A_2 の頻度 q の言葉に直すと

$$q^* = \frac{u}{s} \quad (1.47)$$

であり、これは大域安定平衡点である（下図）。



遺伝子 A_2 は自然選択により排除されるものの、突然変異により共有されることで、頻度 q^* で集団中に保たれることになる。このようにして生まれる均衡を **mutation-selection balance** と呼ぶ。

対立遺伝子が複数ある時に、集団平均の適応度 \bar{w} が、集団で可能な適応度の最大値 w_{\max} よりもどのくらい低下しているかを考えることができる。以下で定義される低下の相対量

$$L = \frac{w_{\max} - \bar{w}}{w_{\max}} \quad (1.48)$$

を**遺伝的荷重 (genetic load)** と呼ぶ。

上の $w_1 = 1, w_2 = 1 - s$ ($s > 0$) なる例では、荷重を引き起こしているのは遺伝子 A_1 から A_2 が生じる有害突然変異 (deleterious mutation) である。集団で可能な適応度の最大値は $w_{\max} = w_1 = 1$ である。mutation-selection balance では遺伝子 A_2 を持つ個体が割合 q^* だけ存在し、その適応度は s だけ低下しているから、集団平均適応度は $\bar{w} = 1 - sq^*$ である。したがって遺伝的荷重は

$$L = \frac{1 - (1 - sq^*)}{1} = sq^* = u \quad (1.49)$$

となり、 s の値に依存しない。この事実を **Haldane-Muller の原理 (ホールデン・マラーの原理; Haldane-Muller's principle)** と呼ぶ。

1.2 二倍体の進化ダイナミクス

次に二倍体、離散世代モデルを考え、2 対立遺伝子 A_1 と A_2 を考える。それぞれの遺伝子頻度を p, q ($= 1 - p$) とする。

遺伝子型は A_1A_1, A_1A_2, A_2A_2 の 3 種類である。それぞれの遺伝子型頻度を $P_{11}, 2P_{12}, P_{22}$ とする (ヘテロ接合体の遺伝子型頻度に 2 がついていることに注意! これは、以降の計算を簡単にするためである)。

時刻 t における集団の状態は $(P_{11}(t), 2P_{12}(t), P_{22}(t))$ の三つ組で記述できる。遺伝子頻度は

$$\begin{aligned} p &= P_{11} + P_{12} \\ q &= P_{12} + P_{22} \end{aligned} \quad (1.50)$$

と計算できる。

解説 1.13 逆に、遺伝子頻度 (p, q) が分かっても、遺伝子型頻度 $(P_{11}, 2P_{12}, P_{22})$ は復元できないことに注意する。なぜなら前者は自由度が 1 (2 変数あるが、 $p + q = 1$ なる制約がある) であるのに対し、後者は自由度が 2 (3 変数あるが、 $P_{11} + 2P_{12} + P_{22} = 1$) であるからである。

それぞれの遺伝子型を持つ個体の適応度を w_{11}, w_{12}, w_{22} とおく。次のような生活環を考える。

- (1) 各個体は適応度に応じた数の一倍体の配偶子を生産し、自らは死ぬ。作られた配偶子は一つの配偶子プールに集められる。
 - (i) 配偶子は同形配偶子 (isogamete)、つまり卵や精子といった区別のない単一種類の配偶子とする。
 - (ii) ヘテロ接合体 A_1A_2 が作る配偶子は、その半分が A_1 、残り半分が A_2 とする (Mendel の分離の法則)。
- (2) 配偶子プールの中で任意交配 (random mating) が行われる。すなわち配偶子がランダムに出会い二倍体の接合子を形成する。このようにして作られた二倍体個体が次世代を構成する。

1.2.1 Hardy-Weinberg 平衡

適応度に差がない $w_{11} = w_{12} = w_{22} = 1$ の場合を考えていこう。

世代 t の個体数		→	世代 $t + 1$ に残す配偶子数		配偶子
遺伝子型 A_1A_1	NP_{11}	→	$NP_{11} \times 1$	$\xrightarrow{\text{割合 } 1}$	A_1
遺伝子型 A_1A_2	$N(2P_{12})$	→	$N(2P_{12}) \times 1$	$\left\{ \begin{array}{l} \xrightarrow{\text{割合 } 1/2} \\ \xrightarrow{\text{割合 } 1/2} \end{array} \right.$	A_1 A_2
遺伝子型 A_2A_2	NP_{22}	→	$NP_{22} \times 1$	$\xrightarrow{\text{割合 } 1}$	A_2

次世代の遺伝子頻度は、配偶子プールの遺伝子頻度と同じである。これらは上の表から以下のように計算できる。

$$\begin{aligned} p' &= \frac{P_{11} + P_{12}}{P_{11} + 2P_{12} + P_{22}} \\ q' &= \frac{P_{12} + P_{22}}{P_{11} + 2P_{12} + P_{22}} \end{aligned} \quad (1.51)$$

式 (1.51) に式 (1.50) を代入して

$$\begin{aligned} p' &= \frac{p}{p+q} = p \\ q' &= \frac{q}{p+q} = q \end{aligned} \quad (1.52)$$

を得る。つまり次世代の遺伝子頻度は現世代の遺伝子頻度に等しい。これは中立であることを考えると当然の結果である。

次世代の遺伝子型頻度は、上記配偶子プール内の任意交配の結果として以下の式で表される。

$$\begin{aligned} P'_{11} &= p'^2 \\ 2P'_{12} &= 2p'q' \\ P'_{22} &= q'^2 \end{aligned} \quad (1.53)$$

式 (1.53) は

- ホモ接合体の遺伝子型頻度は、対応する遺伝子の頻度の 2 乗に等しい。
- ヘテロ接合体の遺伝子型頻度は、対応する二つの遺伝子の頻度の積の 2 倍に等しい。

ことを表している。このような状態を **Hardy-Weinberg 平衡** (ハーディ-ワインベルグ平衡; **Hardy-Weinberg equilibrium**) (略して HW 平衡) と呼ぶ。式 (1.53) は、任意交配があると一世代後には HW 平衡が実現することを表している。

問題 1.14 n 対立遺伝子 A_1, \dots, A_n の時にも同様の結果、つまり一世代の任意交配によって HW 平衡が成り立つこと、を示せ。

1.2.2 自然選択の効果

次に適応度に差がある場合を考える。 $w_{11} = 1 + s, w_{22} = 1$ とおく。 $s (> 0)$ は遺伝子 A_1 のホモ接合体の、遺伝子 A_2 のホモ接合体に対する選択有利度を表す。

ヘテロ接合体 A_1A_2 の適応度を $w_{12} = 1 + hs$ とおく。以下の場合には特に大切である。

- (1) $h = 1/2$ の時：これは w_{12} が w_{11} と w_{22} のちょうど中間の値を取る時であり、あたかも遺伝子 A_1 一つあたり有利度 $s/2$ が個体にもたらされ、 A_1 を二つ持つとその効果が足し算 $(s/2) + (s/2) = s$ でもたらされる状況に対応する。これを遺伝子の効果が**相加的 (additive)** な場合と呼ぶ。
- (2) $h = 1$ の時： $w_{11} = w_{12} > w_{22}$ が成り立つ。したがって A_1 は顕性、 A_2 は潜性遺伝子である。
- (3) $h = 0$ の時： $w_{11} > w_{12} = w_{22}$ が成り立つ。したがって先程と逆で、 A_2 は顕性、 A_1 は潜性遺伝子である。
- (4) $h > 1$ の時：遺伝子 A_1 は、顕性であるとき以上の効果をヘテロ接合体 A_1A_2 にもたらし適応度を高めるので、この現象を**超顕性 (overdominance)** という。 w_{12} は w_{11}, w_{12}, w_{22} の中で最大値を取る。

(5) $h < 0$ の時：遺伝子 A_2 は、潜性であるとき以上の効果をヘテロ接合体 A_1A_2 にもたらし適応度を下げるので、この現象を負の超顕性 (negative overdominance) もしくは underdominance という。 w_{12} は w_{11}, w_{12}, w_{22} の中で最小値を取る。

次世代の遺伝子頻度、遺伝子型頻度を考える。

世代 t の個体数		→	世代 $t+1$ に残す配偶子数		配偶子
遺伝子型 A_1A_1	NP_{11}	→	$NP_{11} \times (1+s)$	$\xrightarrow{\text{割合 } 1}$	A_1
遺伝子型 A_1A_2	$N(2P_{12})$	→	$N(2P_{12}) \times (1+hs)$	$\left\{ \begin{array}{l} \xrightarrow{\text{割合 } 1/2} \\ \xrightarrow{\text{割合 } 1/2} \end{array} \right.$	A_1 A_2
遺伝子型 A_2A_2	NP_{22}	→	$NP_{22} \times 1$	$\xrightarrow{\text{割合 } 1}$	A_2

先程と同様に、次世代の遺伝子頻度は

$$p' = \frac{P_{11}(1+s) + P_{12}(1+hs)}{P_{11}(1+s) + 2P_{12}(1+hs) + P_{22}} \quad (1.54)$$

$$q' = \frac{P_{12}(1+hs) + P_{22}}{P_{11}(1+s) + 2P_{12}(1+hs) + P_{22}}$$

であり、遺伝子型頻度は HW 平衡 (1.53) が実現する。

したがって、世代 $t \geq 1$ では HW 平衡を仮定できるので、 $t \geq 1$ において

$$p' = \frac{p^2(1+s) + pq(1+hs)}{p^2(1+s) + 2pq(1+hs) + q^2} \quad (1.55)$$

である。

解説 1.15 世代 $t \geq 1$ では HW 平衡が仮定できるので、 $p(1)$ が分かると $p(2)$ が分かり、 $p(2)$ が分かると $p(3)$ が分かる、... という構造になっている。

一方で、世代 $t=0$ では HW 平衡が成り立っていないので、 $p(0)$ が分かっても $p(1)$ を知ることはできない。 $p(1)$ を知るためには $p(0)$ の代わりに $(P_{11}(0), 2P_{12}(0), P_{22}(0))$ を知らなければならない。ただし、 $|s|$ が小さければ $p(0)$ と $p(1)$ の差は小さいと予測できるので、この初期の効果はその後起きる遺伝子頻度の変化に比べて無視でき、したがって式 (1.55) が進化ダイナミクスであると最初から仮定して差し支えない。

式 (1.55) を微分方程式で近似するために、 s に関して Taylor 展開し s の 1 次の項までを考えることで、

$$p' - p = sp(1-p)\{p + h(1-2p)\} \quad (1.56)$$

を得るので、微分方程式近似は

$$\frac{dp}{dt} = sp(1-p)\{p + h(1-2p)\} \quad (1.57)$$

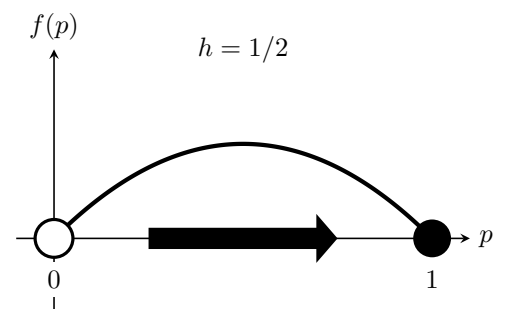
となる (以下、式 (1.57) の右辺を $f(p)$ とおく)。

(1) $h = 1/2$ の時：

式 (1.57) は

$$\frac{dp}{dt} = \frac{s}{2}p(1-p) \quad (1.58)$$

となる。したがって進化ダイナミクスは右図のようになる。



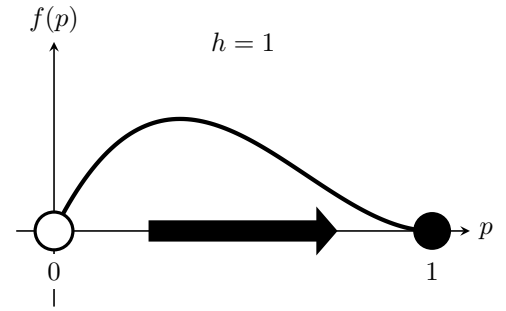
(2) $h = 1$ の時:

式 (1.57) は

$$\frac{dp}{dt} = sp(1-p)^2 \quad (1.59)$$

となる。したがって進化ダイナミクスは右図のようになる。

$p = 0$ の近辺での遺伝子 A_1 の広がる速度は、 $p = 1$ の近辺に比べて速い。これは A_1 が顕性であるために、遺伝子 A_2 で占められた集団に侵入した直後に、ヘテロ接合体 A_1A_2 が有利度 s を得られるからである。



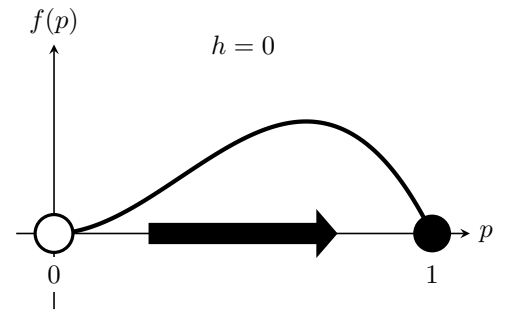
(3) $h = 0$ の時:

式 (1.57) は

$$\frac{dp}{dt} = sp^2(1-p) \quad (1.60)$$

となる。したがって進化ダイナミクスは右図のようになる。

$p = 0$ の近辺での遺伝子 A_1 の広がる速度は、 $p = 1$ の近辺に比べて遅い。これは A_1 が潜性であるために、遺伝子 A_2 で占められた集団に侵入した直後は、ヘテロ接合体 A_1A_2 が特に選択上の有利さを享受できないからである。



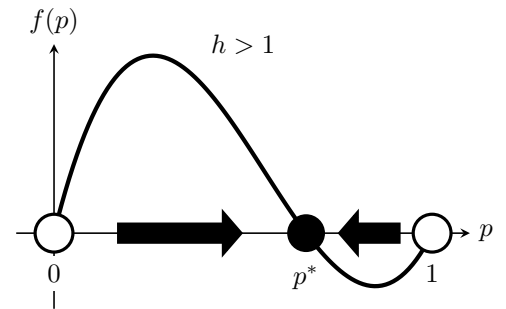
(4) $h > 1$ の時:

式 (1.57) は安定内部平衡点を

$$p^* = \frac{h}{2h-1} \quad (1.61)$$

に持つ。進化ダイナミクスは右図のようになる。

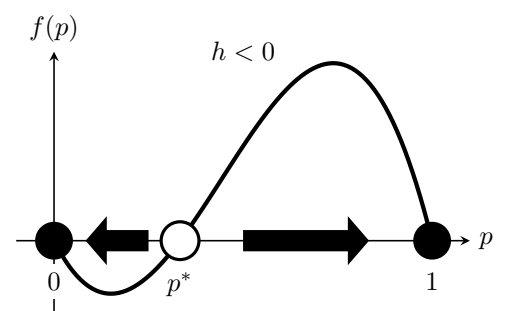
超顕性の状況では、ヘテロ接合体の適応度が一番高いので、遺伝子 A_1 と A_2 の両方が集団に保たれる**多型 (polymorphism)** が実現される。



(5) $h < 0$ の時:

式 (1.57) は不安定内部平衡点を持ち、その位置は式 (1.61) で与えられる。進化ダイナミクスは右図のようになり、**双安定 (bistability)** を示す。

負の超顕性の状況では、ヘテロ接合体の適応度が一番低いので、例えば遺伝子 A_2 で占められた集団に遺伝子 A_1 が少数侵入してもそれは必ずヘテロ接合体 A_1A_2 を作ることとなり、自然選択により遺伝子 A_1 は排除される。逆もまた同じである。



解説 1.16 式 (1.57) をより一般の適応度 $(w_{11}, w_{12}, w_{22}) = (1 + s_{11}, 1 + s_{12}, 1 + s_{22})$ に関して一般化することを考える。ただし、 $|s_{11}|, |s_{12}|, |s_{22}|$ は十分小さいものとする。式 (1.55) と同じように考えれば、差分方程式は

$$\begin{aligned} p' &= \frac{p^2(1+s_{11}) + pq(1+s_{12})}{p^2(1+s_{11}) + 2pq(1+s_{12}) + q^2(1+s_{22})} \\ &= p \cdot \frac{1 + s_{11}p + s_{12}q}{1 + s_{11}p^2 + s_{12} \cdot 2pq + s_{22}q^2} \\ &\approx p \cdot (1 + s_{11}p + s_{12}q - s_{11}p^2 - s_{12} \cdot 2pq - s_{22}q^2) \\ &= p + p(1-p)\{(s_{11} - s_{12})p + (s_{12} - s_{22})(1-p)\} \end{aligned} \quad (1.62)$$

となるから、微分方程式近似は

$$\frac{dp}{dt} = p(1-p)\{(s_{11} - s_{12})p + (s_{12} - s_{22})(1-p)\} \quad (1.63)$$

となる。式 (1.63) に $(s_{11}, s_{12}, s_{22}) = (s, hs, 0)$ を代入すれば式 (1.57) が得られる。

1.2.3 自然選択と突然変異の効果

一倍体の時と同じように、配偶子形成の際に $A_1 \rightarrow A_2$ なる突然変異が確率 u で起き、 $A_2 \rightarrow A_1$ なる突然変異が確率 v で起きるモデルを考える。適応度は（慣習の理由により）先程に代えて $w_{11} = 1, w_{12} = 1 - hs, w_{22} = 1 - s$ の場合を考える。

解説 1.17 先程までのモデルの h を h' と書くことにする。先程までのモデルでは適応度は $w_{11} = 1 + s, w_{12} = 1 + h's, w_{22} = 1$ と仮定していた。これらは相対適応度であるから、

$$w_{11} = 1, \quad w_{12} = \frac{1 + h's}{1 + s}, \quad w_{22} = \frac{1}{1 + s} \quad (1.64)$$

と書いても同じことである。ここで $|s|$ が小さいことを利用して s に関する Taylor 展開を行うと

$$\begin{aligned} \frac{1 + h's}{1 + s} &= 1 - (1 - h')s + \dots \\ \frac{1}{1 + s} &= 1 - s + \dots \end{aligned} \quad (1.65)$$

である。したがって $|s|$ が小さい時には先程のモデルは近似的に

$$w_{11} = 1, \quad w_{12} = 1 - (1 - h')s, \quad w_{22} = 1 - s \quad (1.66)$$

と書けるので、これから考えるモデルの h とは $h = 1 - h'$ の関係がある。特に $h = 1/2$ のとき効果は相加的、 $h = 0$ のとき A_1 は顕性、 $h = 1$ のとき A_1 は潜性である。

ここまで述べてきた説明から、遺伝子 A_1 の遺伝子頻度 p の進化ダイナミクスは

$$\frac{dp}{dt} = \underbrace{sp(1-p)\{p + (1-h)(1-2p)\}}_{\text{自然選択}} \underbrace{-up}_{\text{変異 } A_1 \rightarrow A_2} + \underbrace{v(1-p)}_{\text{変異 } A_2 \rightarrow A_1} \quad (1.67)$$

と書ける。

以下特に $s > 0, v = 0, 0 \leq h \leq 1$ の場合を考える。進化ダイナミクスは

$$\frac{dp}{dt} = sp(1-p)\{p + (1-h)(1-2p)\} - up \quad (1.68)$$

である。

突然変異がない場合は $0 \leq h \leq 1$ であれば $p = 1$ は安定平衡点であった。突然変異がある場合はこの平衡点は不安定化し、 u が小さい時は代わりに

$$p^* \simeq \begin{cases} 1 - \frac{u}{sh} & (\sqrt{\frac{u}{s}} \ll h \leq 1) \\ 1 - \sqrt{\frac{u}{s}} & (h = 0) \end{cases} \quad (1.69)$$

なる平衡点が安定になる。

解説 1.18 式 (1.67) の右辺を $f(p)$ とおき、方程式 $f(p) = 0$ を解くことを考える。 $h = 1/2$ の時、この方程式は

$$p = 0, \quad 1 - \frac{2u}{s} \quad (1.70)$$

を解に持つが、 $u = 0$ を代入して $p = 1$ となるのは後者なので、 $p^* = 1$ が変化した解は後者であり、特に

$$p^* = 1 - \frac{u}{hs} \quad (1.71)$$

と書ける。

$0 < h < 1/2$ もしくは $1/2 < h \leq 1$ の時は、 $f(p) = 0$ は 3 次方程式であり、 $u = 0$ を代入して $p = 1$ となる解は直接計算から

$$p^* = 1 - \frac{h}{2(1-2h)} \left\{ \sqrt{1 + \frac{4(1-2h)}{sh^2}u} - 1 \right\} \quad (1.72)$$

いま $\frac{u}{sh^2} \ll 1$ (つまり $\sqrt{\frac{u}{s}} \ll h$) を仮定すれば、 $|x|$ が小さい時に成り立つ公式 $\sqrt{1+x} \simeq 1 + \frac{x}{2}$ を用いて、

$$p^* \simeq 1 - \frac{h}{2(1-2h)} \left\{ 1 + \frac{2(1-2h)}{sh^2}u - 1 \right\} = 1 - \frac{u}{hs} \quad (1.73)$$

を得る。最後に、 $h = 0$ の時は方程式 $f(p) = 0$ の 3 解は陽に

$$p = 0, \quad 1 \pm \sqrt{\frac{u}{s}} \quad (1.74)$$

であり、このうち

$$p^* = 1 - \sqrt{\frac{u}{s}} \quad (1.75)$$

が安定平衡解である。

式 (1.69) を遺伝子 A_2 の言葉に直すと、

$$q^* \simeq \begin{cases} \frac{u}{sh} & (\sqrt{\frac{u}{s}} \ll h \leq 1) \\ \sqrt{\frac{u}{s}} & (h = 0) \end{cases} \quad (1.76)$$

となる。したがって特に、遺伝子 A_2 が顕性である ($h = 1$) 時は集団中に $q^* = u/s$ の頻度で、潜性である ($h = 0$) 時は集団中に $q^* = \sqrt{u/s}$ の頻度で保たれ、後者の頻度のほうが高い。これは、 A_2 が潜性である場合にはヘテロ接合体 A_1A_2 に存在する遺伝子 A_2 の効果は隠蔽され、集団から除去されにくいからである。

この mutation-selection balance における遺伝的荷重を考える。HW 平衡 $(P_{11}, 2P_{12}, P_{22}) = ((1-q^*)^2, 2(1-q^*)q^*, q^{*2})$ において、平均適応度の低下度は

$$w_{\max} - \bar{w} = hs \cdot 2(1-q^*)q^* + s \cdot q^{*2} \quad (1.77)$$

である。 $\sqrt{u/s} \ll h \leq 1$ のとき、式 (1.77) において q^{*2} は q^* より十分小さいのでこれを無視すると、

$$L = \frac{w_{\max} - \bar{w}}{w_{\max}} \simeq hs \cdot 2q^* = hs \cdot 2 \frac{u}{sh} = 2u \quad (1.78)$$

となる。一方で $h = 0$ の時、式 (1.77) は単に sq^{*2} なので、遺伝的荷重は

$$L = \frac{w_{\max} - \bar{w}}{w_{\max}} \simeq s \cdot q^{*2} = s \cdot \frac{u}{s} = u \quad (1.79)$$

となる。式 (1.78) や (1.79) は再び選択有利度 s に依存せず、したがって二倍体の場合にも Haldane-Muller の原理が成り立つ。

解説 1.19 式 (1.78) と式 (1.79) は一見 h に関して不連続に変化しているように見えるが、 h が 0 に十分小さく、もはや $\sqrt{\frac{u}{s}} \ll h$ を満たさなくなる範囲においては、遺伝的荷重は h が上から 0 に近づくにつれて $2u$ から u に連続的に変化する。

1.3 二座位一倍体モデル

一倍体生物の無限集団を考え、二座位 A, B を考える。それぞれの座位には二対立遺伝子 A_1, A_2 、および B_1, B_2 があるとする。

この一倍体生物は有性生殖を以下のように行う。

- (1) 各個体は適応度に応じた数の、自らと同じ遺伝的組成を持つ一倍体の配偶子を生産し、自らは死ぬ。作られた配偶子は一つの配偶子プールに集められる。
 - 配偶子は同形配偶子 (isogamete)、つまり卵や精子といった区別のない単一種類の配偶子とする。
- (2) 配偶子は配偶子プールの中で任意交配 (random mating) が行い、二倍体の接合子が作られる。この二倍体の接合子はただちに減数分裂を行い一倍体個体を生成する。この際、ある確率で座位 A と B の間で組み換えが行われる。こうやって作られた一倍体個体が次世代を構成する。

一倍体個体の遺伝子型は $A_1B_1, A_1B_2, A_2B_1, A_2B_2$ の 4 つが考えられる。これらの遺伝子型頻度をそれぞれ $p_{1|1}, p_{1|2}, p_{2|1}, p_{2|2}$ とおく (以下、異なる座位は記号 | で区切ることに約束する)。各遺伝子の頻度は

$$\begin{aligned}
 \text{遺伝子 } A_1 \text{ の頻度: } p_{1|\bullet} &= p_{1|1} + p_{1|2} \\
 \text{遺伝子 } A_2 \text{ の頻度: } p_{2|\bullet} &= p_{2|1} + p_{2|2} \\
 \text{遺伝子 } B_1 \text{ の頻度: } p_{\bullet|1} &= p_{1|1} + p_{2|1} \\
 \text{遺伝子 } B_2 \text{ の頻度: } p_{\bullet|2} &= p_{1|2} + p_{2|2}
 \end{aligned} \tag{1.80}$$

と書ける。

配偶子 A_xB_y と配偶子 A_zB_w が融合してできた二倍体 D を考える。D は減数分裂の際に、一倍体 A_xB_y と A_zB_w を作るのみならず、一定の確率で組み換えを起こし、一倍体 A_xB_w および A_zB_y も作る。組み換えの確率を r とおくと、二倍体 D が作る一倍体個体は

一倍体個体の遺伝子型	頻度
A_xB_y	$\frac{1}{2}(1-r)$
A_zB_w	$\frac{1}{2}(1-r)$
A_xB_w	$\frac{1}{2}r$
A_zB_y	$\frac{1}{2}r$

とまとめられる。

組み換え率 r は、遺伝子座 A, B が同一染色体上で隣接している場合が最小であると考えられる。反対に遺伝子座 A, B が異なる染色体上にある時は、その値は $r = 1/2$ (最大値) である。

個体 $A_1B_1, A_1B_2, A_2B_1, A_2B_2$ の適応度をそれぞれ $w_{1|1}, w_{1|2}, w_{2|1}, w_{2|2}$ とおこう。

1.3.1 連鎖不平衡

まず $w_{1|1} = w_{1|2} = w_{2|1} = w_{2|2} = 1$ である中立な場合を考える。

任意交配下で、次世代の個体の遺伝子型頻度 $p'_{1|1}, p'_{1|2}, p'_{2|1}, p'_{2|2}$ がどうなるかを考えよう。

解説 1.20 $i = 1, 2, j = 1, 2$ とし、個体 A_xB_y と A_zB_w の交配に起因して、次世代に個体 A_iB_j がどれだけの数で生産されるかを考える。

上の表から、第一に、 $(x, y) = (i, j)$ のとき、確率 $(1/2)(1-r)$ で個体 A_iB_j が作られる。 $(x, y) = (i, j)$ である確率は $p_{i|j}$ である。

第二に、 $(z, w) = (i, j)$ のとき、確率 $(1/2)(1-r)$ で個体 A_iB_j が作られる。 $(x, y) = (i, j)$ である確率は $p_{i|j}$ である。

第三に、 $(x, w) = (i, j)$ のとき、確率 $(1/2)r$ で個体 A_iB_j が作られる。任意交配の仮定より、 $x = i$ である確率は $p_{i|\bullet}$ であり、 $w = j$ である確率は $p_{\bullet|j}$ である。

第四に、 $(z, y) = (i, j)$ のとき、確率 $(1/2)r$ で個体 A_iB_j が作られる。任意交配の仮定より、 $y = j$ である確率は $p_{\bullet|j}$ であり、 $z = i$ である確率は $p_{i|\bullet}$ である。

これらをまとめて、 $i = 1, 2, j = 1, 2$ に対し、

$$\begin{aligned} p'_{i|j} &= \frac{1}{2}(1-r) \cdot p_{i|j} + \frac{1}{2}(1-r) \cdot p_{i|j} + \frac{1}{2}r \cdot p_{i|\bullet}p_{\bullet|j} + \frac{1}{2}r \cdot p_{i|\bullet}p_{\bullet|j} \\ &= (1-r)p_{i|j} + rp_{i|\bullet}p_{\bullet|j} \end{aligned} \quad (1.81)$$

を得る。

上の計算から、具体的に書くと

$$\begin{aligned} p'_{1|1} &= (1-r)p_{1|1} + rp_{1|\bullet}p_{\bullet|1} \\ p'_{1|2} &= (1-r)p_{1|2} + rp_{1|\bullet}p_{\bullet|2} \\ p'_{2|1} &= (1-r)p_{2|1} + rp_{2|\bullet}p_{\bullet|1} \\ p'_{2|2} &= (1-r)p_{2|2} + rp_{2|\bullet}p_{\bullet|2} \end{aligned} \quad (1.82)$$

を得る。

ここで

$$D \equiv p_{1|1}p_{2|2} - p_{1|2}p_{2|1} \quad (1.83)$$

なる量を定義する。この量を連鎖不平衡 (linkage disequilibrium; LD) と呼ぶ。

解説 1.21 連鎖不平衡は二座位の独立性からずれを測る量である。もし二座位が独立ならば、配偶子の遺伝子型頻度は、各座位の遺伝子頻度の積として

$$\begin{aligned} p_{1|1} &= p_{1|\bullet}p_{\bullet|1} \\ p_{1|2} &= p_{1|\bullet}p_{\bullet|2} \\ p_{2|1} &= p_{2|\bullet}p_{\bullet|1} \\ p_{2|2} &= p_{2|\bullet}p_{\bullet|2} \end{aligned} \quad (1.84)$$

と書けているはずなので、この時の連鎖不平衡の値は

$$D = (p_{1|\bullet}p_{\bullet|1}) \cdot (p_{2|\bullet}p_{\bullet|2}) - (p_{1|\bullet}p_{\bullet|2}) \cdot (p_{2|\bullet}p_{\bullet|1}) = 0 \quad (1.85)$$

となる。この状態を連鎖平衡 (linkage equilibrium) と呼ぶ。

$D > 0$ ならば独立時に期待される時よりも遺伝子型 A_1B_1 および A_2B_2 の頻度が大きい状態を、逆に $D < 0$ ならば独立時に期待される時よりも遺伝子型 A_1B_2 および A_2B_1 の頻度が大きい状態を表している。

D をどちら向きに定義するか、すなわち式 (1.83) に代えて $p_{1|2}p_{2|1} - p_{1|1}p_{2|2}$ を連鎖不平衡の定義としたいかどうかは、純粋に座位 A と B の対立遺伝子のラベル付けの問題である。すなわちどちらの向きを採用したいかは定義する側の都合による。

一般に、異なる座位の遺伝子の組み合わせが、独立性で期待される確率からずれている現象を連鎖 (linkage) という。

このとき、直接計算により

$$\begin{cases} p_{1|\bullet}p_{\bullet|1} = p_{1|1} - D \\ p_{1|\bullet}p_{\bullet|2} = p_{1|2} + D \\ p_{2|\bullet}p_{\bullet|1} = p_{2|1} + D \\ p_{2|\bullet}p_{\bullet|2} = p_{2|2} - D \end{cases} \iff \begin{cases} p_{1|1} = p_{1|\bullet}p_{\bullet|1} + D \\ p_{1|2} = p_{1|\bullet}p_{\bullet|2} - D \\ p_{2|1} = p_{2|\bullet}p_{\bullet|1} - D \\ p_{2|2} = p_{2|\bullet}p_{\bullet|2} + D \end{cases} \quad (1.86)$$

が分かるので、式 (1.86) を式 (1.82) に代入して、

$$\begin{aligned} p'_{1|1} &= p_{1|1} - rD \\ p'_{1|2} &= p_{1|2} + rD \\ p'_{2|1} &= p_{2|1} + rD \\ p'_{2|2} &= p_{2|2} - rD \end{aligned} \quad (1.87)$$

が分かる。特に、次世代の連鎖不平衡 $D' = p'_{1|1}p'_{2|2} - p'_{1|2}p'_{2|1}$ は式 (1.87) より、

$$D' = (1 - r)D \quad (1.88)$$

と計算される。したがって $r > 0$ の時、選択がないと、連鎖不平衡は毎世代 $(1 - r)$ 倍に縮小し、やがて 0 に収束する。

解説 1.22 例えば式 (1.86) の第一式は直接計算により以下のように示される。

$$p_{1|\bullet}p_{\bullet|1} + D = (p_{1|1} + p_{1|2})(p_{1|1} + p_{2|1}) + (p_{1|1}p_{2|2} - p_{1|2}p_{2|1}) = p_{1|1} \underbrace{(p_{1|1} + p_{1|2} + p_{2|1} + p_{2|2})}_{=1} + \underbrace{(p_{1|2}p_{2|1} - p_{1|2}p_{2|1})}_{=0} = p_{1|1} \quad (1.89)$$

他の三式も同様に示せる。式 (1.86) は、系の状態を $(p_{1|1}, p_{1|2}, p_{2|1}, p_{2|2})$ のうちの三変数で記述する代わりに、例えば三変数 $(p_{1|\bullet}, p_{\bullet|1}, D)$ (つまり、「遺伝子 A_1 の頻度」「遺伝子 B_1 の頻度」「連鎖不平衡」) で記述することもできることを意味している。

式 (1.87) から式 (1.88) は、

$$\begin{aligned} D' &= (p_{1|1} - rD)(p_{2|2} - rD) - (p_{1|2} + rD)(p_{2|1} + rD) \\ &= \underbrace{p_{1|1}p_{2|2} - p_{1|2}p_{2|1}}_{=D} - rD \underbrace{(p_{1|1} + p_{1|2} + p_{2|1} + p_{2|2})}_{=1} \end{aligned} \quad (1.90)$$

より従う。

1.3.2 自然選択とヒッチハイク

次に一般の適応度 $(w_{1|1}, w_{1|2}, w_{2|1}, w_{2|2})$ に対し、各遺伝子型頻度 $p_{i|j}$ の時間変化を考えよう。

適応度の差によって、配偶子プールにおける各配偶子の頻度は

配偶子の遺伝子型	配偶子プールにおける頻度
A_1B_1	$w_{1 1}p_{1 1}/\bar{w} (= g_{1 1})$
A_1B_2	$w_{1 2}p_{1 2}/\bar{w} (= g_{1 2})$
A_2B_1	$w_{2 1}p_{2 1}/\bar{w} (= g_{2 1})$
A_2B_2	$w_{2 2}p_{2 2}/\bar{w} (= g_{2 2})$

とまとめられる。ただしここで \bar{w} は平均適応度

$$\bar{w} = w_{1|1}p_{1|1} + w_{1|2}p_{1|2} + w_{2|1}p_{2|1} + w_{2|2}p_{2|2} \quad (1.91)$$

である。また、各周辺頻度を

$$\begin{aligned} g_{1|\bullet} &= g_{1|1} + g_{1|2} \\ g_{2|\bullet} &= g_{2|1} + g_{2|2} \\ g_{\bullet|1} &= g_{1|1} + g_{2|1} \\ g_{\bullet|2} &= g_{1|2} + g_{2|2} \end{aligned} \quad (1.92)$$

のように定義しておく。

この時、次世代の遺伝子頻度の計算は、**解説 1.20** で p_{ij} を全て g_{ij} で置き換えたものになる。したがって式 (1.82) に代えて

$$\begin{aligned} p'_{1|1} &= (1-r)g_{1|1} + rg_{1|\bullet}g_{\bullet|1} \\ p'_{1|2} &= (1-r)g_{1|2} + rg_{1|\bullet}g_{\bullet|2} \\ p'_{2|1} &= (1-r)g_{2|1} + rg_{2|\bullet}g_{\bullet|1} \\ p'_{2|2} &= (1-r)g_{2|2} + rg_{2|\bullet}g_{\bullet|2} \end{aligned} \quad (1.93)$$

を得る。

より見やすい式を得るために、 $w_{i|j} = 1 + s_{i|j}$ とおき、各 $|s_{i|j}|$ は十分に小さいと仮定しよう。また r も同程度に十分小さいと仮定しよう。するといくらかの計算の後に微分方程式近似

$$\begin{aligned} \frac{dp_{1|1}}{dt} &= p_{1|1}(s_{1|1} - \bar{s}) - rD \\ \frac{dp_{1|2}}{dt} &= p_{1|2}(s_{1|2} - \bar{s}) + rD \\ \frac{dp_{2|1}}{dt} &= p_{2|1}(s_{2|1} - \bar{s}) + rD \\ \frac{dp_{2|2}}{dt} &= p_{2|2}(s_{2|2} - \bar{s}) - rD \end{aligned} \quad (1.94)$$

を得る。ただし $\bar{s} \equiv s_{1|1}p_{1|1} + s_{1|2}p_{1|2} + s_{2|1}p_{2|1} + s_{2|2}p_{2|2}$ であり、集団の平均適応度 \bar{w} とは $\bar{w} = 1 + \bar{s}$ の関係にある。

解説 1.23 式 (1.94) の第一式を導出しよう。まず s に関わる項に関する Taylor 展開より

$$g_{i|j} = \frac{(1 + s_{i|j})p_{i|j}}{1 + \bar{s}} \approx p_{i|j}(1 + s_{i|j} - \bar{s}) \quad (1.95)$$

である。これを用いると

$$(1-r)g_{1|1} \approx p_{1|1}(1 + s_{1|1} - \bar{s} - r) = p_{1|1} + p_{1|1}(s_{1|1} - \bar{s}) - rp_{1|1} \quad (1.96)$$

である。また、 $rg_{1|\bullet}g_{\bullet|1}$ の項は、 r が既に微小量であることを考えれば、

$$\begin{aligned} rg_{1|\bullet}g_{\bullet|1} &= r \{p_{1|1}(1 + s_{1|1} - \bar{s}) + p_{1|2}(1 + s_{1|2} - \bar{s})\} \{p_{1|1}(1 + s_{1|1} - \bar{s}) + p_{2|1}(1 + s_{2|1} - \bar{s})\} \\ &\approx r(p_{1|1} + p_{1|2})(p_{1|1} + p_{2|1}) = rp_{1|\bullet}p_{\bullet|1} \end{aligned} \quad (1.97)$$

である。したがって

$$p'_{1|1} \approx p_{1|1} + p_{1|1}(s_{1|1} - \bar{s}) - \underbrace{r(p_{1|1} + p_{1|\bullet}p_{\bullet|1})}_{=D} \quad (1.98)$$

であり、これを微分方程式近似すれば求める式が得られる。他の三式も同様である。

例として、座位 A のみが適応度を完全に決定している状況、 $(s_{1|1}, s_{1|2}, s_{2|1}, s_{2|2}) = (s, s, 0, 0)$ ($s > 0$) を考えよう。

式 (1.94) にこの仮定を代入して

$$\begin{aligned} \frac{dp_{1|1}}{dt} &= p_{1|1}(s - \bar{s}) - rD \\ \frac{dp_{1|2}}{dt} &= p_{1|2}(s - \bar{s}) + rD \\ \frac{dp_{2|1}}{dt} &= p_{2|1}(0 - \bar{s}) + rD \\ \frac{dp_{2|2}}{dt} &= p_{2|2}(0 - \bar{s}) - rD \\ \bar{s} &= s(p_{1|1} + p_{1|2}) \end{aligned} \quad (1.99)$$

を得る。

より直観的に分かりやすくするため、これを三変数 $(p_{1|\bullet}, p_{\bullet|1}, D)$ に関するより微分方程式に変換する。式を見やすくするため、

以下では $(p_{1|\bullet}, p_{\bullet|1}, D) = (x, y, D)$ と書こう。互いに移り合うための公式は

$$\begin{cases} x = p_{1|1} + p_{1|2} \\ y = p_{1|1} + p_{2|1} \\ D = p_{1|1}p_{2|2} - p_{1|2}p_{2|1} \end{cases} \iff \begin{cases} p_{1|1} = xy + D \\ p_{1|2} = x(1-y) - D \\ p_{2|1} = (1-x)y - D \\ p_{2|2} = (1-x)(1-y) + D \end{cases} \quad (1.100)$$

であり、特に時間微分は

$$\begin{aligned} \frac{dx}{dt} &= \frac{dp_{1|1}}{dt} + \frac{dp_{1|2}}{dt} \\ \frac{dy}{dt} &= \frac{dp_{1|1}}{dt} + \frac{dp_{2|1}}{dt} \\ \frac{dD}{dt} &= \frac{dp_{1|1}}{dt}p_{2|2} + p_{1|1}\frac{dp_{2|2}}{dt} - \frac{dp_{1|2}}{dt}p_{2|1} + p_{1|2}\frac{dp_{2|1}}{dt} \end{aligned} \quad (1.101)$$

で得られる。いくつかの計算の後、

$$\frac{dx}{dt} = sx(1-x) \quad (1.102a)$$

$$\frac{dy}{dt} = sD \quad (1.102b)$$

$$\frac{dD}{dt} = -rs(1-2x)D \quad (1.102c)$$

を得る（計算は複雑なので、数式計算ソフトウェアを用いるとよい）。

式 (1.102a) は式 (1.15) と全く同じ形をしており、遺伝子 A_1 が頻度を増していくことを述べている。これは仮定から期待できる当然の帰結である。

一方で対立遺伝子 B_1, B_2 は中立であったことを考えると、式 (1.102b) の右辺が 0 でないことはある種の驚きである。例えば初期状態で $D(0) > 0$ なる正の連鎖不平衡が存在したとする。すると式 (1.102c) より、 D の符号は常に正であることが分かる（最終的には $D(\infty) = 0$ となる）。したがって遺伝子 B_1 には常に正の選択が働くことになる。

一見この現象は奇妙にも思えるが、 $D > 0$ であるということは A_1 と B_1 が正に連鎖しており、 B_1 はランダムで期待されるよりも A_1 との結びつきが強いので、 A_1 が有利であることに「ヒッチハイク」して、 B_1 自身には有利さがないにもかかわらず、 A_1 との連鎖の恩恵を受けて正の選択を受けていることを意味する。このような現象を**遺伝的ヒッチハイク (genetic hitchhiking)** と呼ぶ。

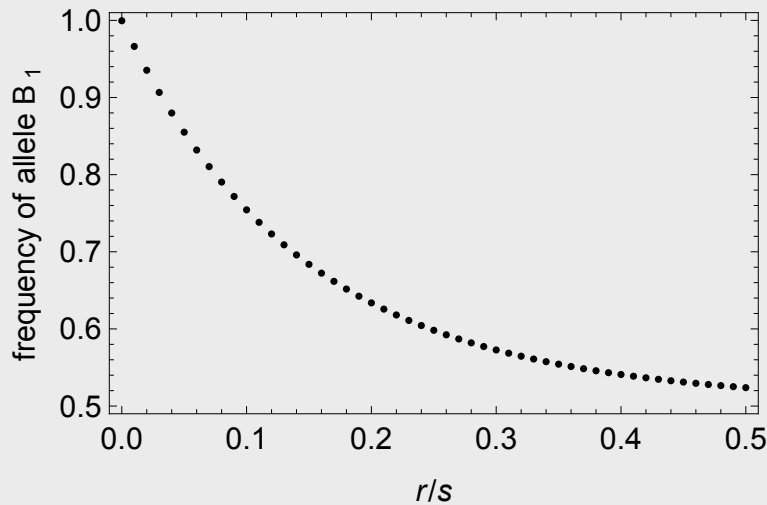


図 1.1: 初期値 $(p_{1|1}, p_{1|2}, p_{2|1}, p_{2|2}) = (1, 0, 499, 500)/1000$ から始めて $p_{1|1} = 999/1000$ に至る時刻 t_{end} まで式 (1.102) を計算した時の、時刻 t_{end} での座位 B における対立遺伝子 B_1 の頻度 $y (= p_{\bullet|1})$ を縦軸にプロットした。横軸は組み換え率の相対的大きさである r/s の値。遺伝的ヒッチハイクによって $y(t_{\text{end}}) > 1/2$ となっていることが分かる。

図 1.1 の横軸は、座位 A との相対的な組み換え率を表すので、これは座位 A, B 間のゲノム上の距離のように思えば良い。すると座位 A で自然選択が起きると座位 A における多様性が失われると、座位 A の近隣の座位においても遺伝的ヒッチハイクによって多様性が失われてしまう（注：つまりこの例では、対立遺伝子 B₁ の頻度が増し、B₂ が集団から失われる）。このように自然選択が働いた座位の近隣でも多様性が失われる現象を**選択的一掃 (selective sweep)** と呼ぶ。

1.3.3 相加性とエピスタシス

次に、

- 座位 A では、遺伝子 A₁ が遺伝子 A₂ に対して $s_A (> 0)$ の有利度を持つ
- 座位 B では、遺伝子 B₁ が遺伝子 B₂ に対して $s_B (> 0)$ の有利度を持つ
- これらの効果は相加的である

ような状況を考えよう。これらは $(s_{1|1}, s_{1|2}, s_{2|1}, s_{2|2}) = (s_A + s_B, s_A, s_B, 0)$ とまとめられる。

式 (1.94) に代入して整理すると、最終的に

$$\frac{dx}{dt} = s_A x(1-x) + s_B D \quad (1.103)$$

$$\frac{dy}{dt} = s_A D + s_B y(1-y) \quad (1.104)$$

$$\frac{dD}{dt} = \{-r + s_A(1-2x) + s_B(1-2y)\} D \quad (1.105)$$

を得る。これは唯一の安定平衡点 $(x, y, D) = (1, 1, 0)$ を持つ。これは期待される当然の結果である。

問題 1.24 上を示せ（数式処理ソフトウェアを使うと良い）。

次に、A₁B₁ に非相加的な効果 s_{AB} を導入し、 $(s_{1|1}, s_{1|2}, s_{2|1}, s_{2|2}) = (s_A + s_B + s_{AB}, s_A, s_B, 0)$ の場合を考えよう。効果 s_{AB} は対立遺伝子 A₁ と B₁ の組み合わせで初めて生じる効果であり、統計学で言うところの交互作用に対応する効果である。このような異なる座位間の相乗効果をもたらす非相加的效果を**エピスタシス (epistasis)** と呼ぶ。このモデルの場合、進化ダイナミクスは

$$\frac{dx}{dt} = s_A x(1-x) + s_B D + s_{AB}(1-x)(xy + D) \quad (1.106)$$

$$\frac{dy}{dt} = s_A D + s_B y(1-y) + s_{AB}(1-y)(xy + D) \quad (1.107)$$

$$\frac{dD}{dt} = \{-r + s_A(1-2x) + s_B(1-2y)\} D + s_{AB}\{(1-x)(1-y) - D\}(xy + D) \quad (1.108)$$

となる。

エピスタシスの存在する簡単な例として、 $(s_{1|1}, s_{1|2}, s_{2|1}, s_{2|2}) = (0, s, s, 0)$ ($s > 0$) の場合を考える。言い換えれば、A₁ と B₁ のうちどちらか片方だけを持っている時にのみ適応度が s 増加する場合である。

これは例えば以下のような状況で起こりうる。A と B はともに体サイズを決める座位で、A₁ や B₁ は、A₂ や B₂ に比べて体サイズを大きくする遺伝子であるとする。環境には適切な体サイズが存在し、遺伝子型 A₂B₂ では小さすぎ、遺伝子型 A₁B₁ では大きすぎ、遺伝子型 A₁B₂ もしくは遺伝子型 A₂B₁ が最適な体サイズをもたらすとする。

この時、安定な平衡点は $(x, y, D) = (1, 0, 0)$ と $(x, y, D) = (0, 1, 0)$ のみであることが分かり、中間に生じる平衡点

$$(x, y, D) = \left(\frac{1}{2}, \frac{1}{2}, \frac{r - \sqrt{r^2 + s^2}}{4s} \right) \quad (1.109)$$

は不安定である。したがって、両遺伝子座で多型が維持されることはなく、A₁ と B₂ に固定されるか、A₂ と B₁ に固定されるかのどちらかになることが分かる。最終状態としてどちらが実現するかは初期頻度に依存する。

問題 1.25 上を示せ（数式処理ソフトウェアを使うと良い）。

2 離散表現型の進化モデル

(本章の概要と目的) 本章ではレプリケーター方程式を中心に、表現型が離散的に与えられる場合についてその基本的な進化ダイナミクスの記述方法を学ぶ。本章の内容は集団遺伝学で扱われることは少なく、主に進化生物学で用いられる。レプリケーター方程式は1970年代に導入された比較的新しい道具であり、進化生物学のみならず理論経済学においても幅広く用いられている。

特に本章では連続時間レプリケーター方程式の数理を学ぶ。 n 個の異なる戦略が存在する時、レプリケーター方程式は $x_1 + \dots + x_n = 1$ を満たす可能な頻度 (x_1, \dots, x_n) の集合上で定義される微分方程式となり、様々な良い数学的性質を持つ。この性質を利用して、特に一般の対称2戦略ゲームの完全分類を行い、どのような進化の最終状態が可能であるかを理解する。また、応用例としてエラーカタストロフ理論に触れる。

前章と対照的なのは、本章では表現型を実現する遺伝学的詳細には一切踏み込まず、むしろその議論を省略して最も単純化された遺伝様式を仮定し、その代わりにより自由な論理展開を目指すところにある。この表現型 *gambit* と呼ばれる方法論は、進化生物学で広く用いられ、より直観的な現象の理解を可能にした。

本章の目的はレプリケーター方程式を用いた自由なモデリングの方法論を体得し、その構成方法を理解し、単純な例を自分自身の力で解析できるようになることである。そのためにはレプリケーター方程式の基本的な数学的性質の理解も必要となる。

2.1 表現型 *gambit*

例えば行動生態学においてある脊椎動物の行動を分析することを考える。その行動を引き起こす遺伝子が一遺伝子座によって決定されていることはごく稀で、複数の遺伝子座の遺伝子が統合的に働いてその行動が実現されていると考えるのが自然であろう。また多くの脊椎動物は二倍体で有性生殖を行う。したがって注目する行動を、その遺伝システムに忠実に、集団遺伝学的に分析するには、

- 二倍体
- 有性生殖
- 複数座位

のモデルを考えなければならない。しかしながら前章で見てきたように、このようなモデルは数学的に複雑である。

そこで、実際の遺伝システムには目をつぶって、なるべく単純なモデルを用いてこのような行動を分析するという代替案が考えられる。具体的には

- 一倍体
- 無性生殖
- 単一座位

のモデルが最も簡単である。そしてしばしばこのようなモデルは、上に述べたような詳細モデルの良い近似になっていることが多い。

生物の表現型を分析する場合には、現実を単純化したモデルを用いることで、集団遺伝学的詳細を犠牲にする代わりにより単純明快な結論を引き出そうとすることが多い。このような方法論を**表現型 *gambit* (ひょうげんがたギャンビット) (phenotypic gambit)** という (Grafen 1984)。

表現型 *gambit* は例えば**進化的に安定な戦略 (Evolutionarily stable strategy; ESS)** の理論において進化生態学分野で70年代から広く用いられ成功を収めた。表現型 *gambit* で得られた結論は、実はしばしば集団遺伝学の詳細モデルで得られる結論と同一になり、そのような場合は、表現型 *gambit* を行ったモデルのほうが好まれる傾向がある。しかしより複雑な状況下では、表現型 *gambit* が誤りを導くこともあるので、その場合は集団遺伝学モデルに立ち返って分析をしなければならない。

解説 2.1 *gambit* はチェス用語で、駒を犠牲にする代わりに有利な盤面を得ようとする戦略のことを指す。表現型 *gambit* とは、集団遺伝学詳細を犠牲にする代わりに単純明快な結論を得ようとすることにかけての用語である。

例 2.2 タカ・ハトゲーム (hawk-dove game) では戦略は「タカ」と「ハト」の二種類であるが、古典的な枠組みでは

- 「タカ」遺伝子と「ハト」遺伝子の両方を持つヘテロ接合体
- 「タカ」親と「ハト」親が子を作ったら何が生まれるか
- 「タカ」「ハト」的行動を生み出す複数の遺伝子座

などは一切考えない。つまりこれは表現型 gambit の典型例である。

例 2.3 局所配偶者競争 (local mate competition; LMC) の理論で、局所的に競争する母親の数が N である場合に、Hamilton は ESS 性比

$$x^* = \frac{N-1}{2N} \quad (2.1)$$

を導いたが、なぜか彼のコンピュータシミュレーションの結果とは合わなかった (Hamilton 1967)。しかし後日、より詳細な集団遺伝学的分析が行われ、この公式は二倍体生物に関しては正しいが、半倍数性生物に対しては誤りであり、半倍数性下での正しい公式は

$$x^* = \frac{(N-1)(2N-1)}{N(4N-1)} \quad (2.2)$$

であることが分かった (Hamilton 1979; Taylor & Bulmer 1980)。

2.2 レプリケーター方程式

以下では特に断りのない限り、一倍体、無性生殖、単一座位のモデルを考える。各個体は s_1 から s_n のうちどれか一つの形質を持つとし、それらは座位 A における対立遺伝子 A_1, \dots, A_n が決定しているとする。この状況において形質はゲーム理論との対比でしばしば**戦略 (strategy)** と呼ばれる。

戦略 s_i を持つ個体の数を N_i 、その集団中における頻度を x_i で表すことにする。各 N_i は十分に大きく、したがってその有限性は数学的に無視できるとする。

x_i が従う微分方程式を導くために以下のような生活環を考える。

- (1) 戦略 s_i を持つ個体は、出生率 b_i でクローンの子を生む。
 - 繁殖率 b_i は時刻 t に依存して変化してよく、また戦略頻度 x_1, \dots, x_n に依存して変化してもよく、添字 i に依存してもよい。
- (2) 戦略 s_i を持つ個体は、死亡率 d_i で死亡する。
 - 死亡率 d_i は、時刻 t に依存して変化してよく、また戦略頻度 x_1, \dots, x_n に依存して変化してもよく、添字 i に依存してもよい。

出生率と死亡率の差

$$f_i \equiv b_i - d_i \quad (2.3)$$

は、戦略 s_i の個体の**増殖率 (growth rate)** と呼ばれる (個体群の増殖率を指す場合は**マルサス係数 (Malthusian coefficient)**とも呼ばれる)。

以上の仮定から次の式を得る。

$$\dot{N}_i = (b_i - d_i)N_i = f_i N_i \quad (2.4)$$

ただしここで $\dot{\cdot}$ は時間微分を表す (以下、この記法を頻繁に用いる)。戦略 s_i の頻度 x_i はその定義から

$$x_i = \frac{N_i}{N} \quad (2.5)$$

である。ただし $N = N_1 + \dots + N_n$ である。式 (2.4) と式 (2.5) を用いて、

$$\dot{x}_i = x_i(f_i - \bar{f}) \quad (2.6a)$$

が得られる。ここで \bar{f} は平均増殖率

$$\bar{f} = f_1x_1 + \cdots + f_nx_n = \sum_{j=1}^n f_jx_j \quad (2.6b)$$

である。式 (2.6) をレプリケーター方程式 (replicator equation) と呼ぶ (Taylor & Jonker, 1978)。式 (1.21) や式 (1.94) ($r = 0$ とおく) などとの類似に注目せよ。

解説 2.4 レプリケーター方程式 (2.6) を導こう。

まず、式 (2.4) を導く。出生率が b_i であるとは、微小時間 Δt 内に注目しているある個体が子を産む確率が $b_i\Delta t$ であることを意味する。いま、個体数は N_i なので、微小時間 Δt に生まれる子の総数は $N_ib_i\Delta t$ となる。同様に、死亡率が d_i であるとは、微小時間 Δt 内に注目しているある個体が子を産む確率が $d_i\Delta t$ であることを意味する。いま、個体数は N_i なので、微小時間 Δt に死亡する個体の総数は $N_id_i\Delta t$ となる。この微小時間内に、注目しているある個体が子を産むイベントと死亡するイベントが両方を経験する確率は $b_i\Delta t \times d_i\Delta t$ であるが、これは $(\Delta t)^2$ を含む二次の微小量であるので、無視しても構わない。

以上のことから、時刻 t から Δt だけ経過した後の N_i の値は

$$N_i(t + \Delta t) = N_i(t) + N_ib_i\Delta t - N_id_i\Delta t \quad (2.7)$$

と計算でき、整理することで

$$\frac{N_i(t + \Delta t) - N_i(t)}{\Delta t} = (b_i - d_i)N_i = f_iN_i \quad (2.8)$$

を得る。両辺の $\Delta t \rightarrow 0$ なる極限を取ると、左辺は $N_i(t)$ の t に関する微分

$$\dot{N}_i(t) = \lim_{\Delta t \rightarrow 0} \frac{N_i(t + \Delta t) - N_i(t)}{\Delta t} \quad (2.9)$$

に等しいから、式 (2.4) が得られる。

次に x_i の時間微分を求める。式 (2.5) に商の微分公式を用いることで

$$\dot{x}_i = \frac{\dot{N}_iN - N_i\dot{N}}{N^2} = \frac{\dot{N}_i}{N} - \underbrace{\frac{N_i}{N}}_{x_i} \frac{\dot{N}}{N} = \frac{\dot{N}_i}{N} - x_i \sum_{j=1}^n \frac{\dot{N}_j}{N} \quad (2.10)$$

である。ここに式 (2.4) を代入して

$$\dot{x}_i = \frac{f_iN_i}{N} - x_i \sum_{j=1}^n \frac{f_jN_j}{N} = f_ix_i - x_i \underbrace{\sum_{j=1}^n f_jx_j}_{=\bar{f}} = x_i(f_i - \bar{f}) \quad (2.11)$$

を得る。

解説 2.5 進化では戦略の頻度のみが重要だが、総個体数 N はどうなっているだろうか。上の議論で既に

$$\dot{N} = \bar{f}N \quad (2.12)$$

を示した。もし \bar{f} が正の時間が長く続けば、 N は限りなく増殖するし、逆に \bar{f} が負である時間が長く続けば、 N は 0 に近づいてしまう。前者は生物学的に妥当ではなく、後者も、絶滅した集団の中の戦略の頻度を分析することはナンセンスに思われる。

これを解決する一つの「言い訳」は、出生率や死亡率が時間とともにうまく変化し、集団サイズ N は一定になっているという考え方である。例えば各 b_i の値が大きいので、集団では頻繁に出生が行われているとしよう。集団の平均出生率は \bar{b} である。この時、各戦略 s_i にはこれを打ち消すだけの新たな死亡率が生じていると仮定する。出生が多いということはそれだけ環境が混んでいるということだから、これはある程度妥当な仮定である。また、その効果はどの戦略も等しく受けることとする。これも一つの集団における環境は皆が共有するものだから妥当な仮定である。結果として各戦略の死亡率は d_i から $d'_i = d_i + \bar{f}$ へと変更を受けるとしよう。すると変更後の集団の新たな平均増殖率 \bar{f}' は

$$\bar{f}' = \bar{b} - \bar{d}' = \bar{b} - \sum_{j=1}^n (d_j + \bar{f})x_j = \bar{b} - (\bar{d} + \bar{f}) = 0 \quad (2.13)$$

となり、集団サイズ N は一定であることになる。この時、新しい増殖率 f'_i を用いたレプリケーター方程式は

$$\dot{x}_i = x_i \left(\underbrace{f'_i}_{=b_i - (d_i + \bar{f}) = f_i - \bar{f}} - \underbrace{\bar{f}'}_{=0} \right) = x_i (f_i - \bar{f}) \quad (2.14)$$

となり、変更前のレプリケーター方程式と変わらない。逆に、死亡率 d_i が大きくそのままだと集団が絶滅に向かう時は、出生の側にフィードバックがあると考えれば以上の議論は同様に成り立つ。

まとめると、集団サイズ N の増殖がどのようにになっているかを考える時に、各戦略に共通の死亡率もしくは出生率を新たに導入しても、元のレプリケーター方程式は引き続き成り立つ。言い換えれば、レプリケーター方程式は戦略に固有の出生率や死亡率の差には敏感であるが、そうでない要因には無関係であるため、戦略の頻度変化を考える問題と集団全体の増殖率を考える問題は、前者は進化学の問題として、後者は生態学の問題として、それぞれ独立に考えても差し支えない。

n 戦略レプリケーター方程式は n 次元ユークリッド空間の $(n-1)$ 次元閉部分集合

$$S_n = \{ \mathbf{x} = (x_1, \dots, x_n) \mid x_i \geq 0, x_1 + \dots + x_n = 1 \} \quad (2.15)$$

上で定義される。 S_n を $(n-1)$ 次元単体 (simplex) と呼ぶ。

解説 2.6 イメージしやすいように図形で言えば、 S_2 は長さ 1 の線分、 S_3 は正三角形の周及び内部、 S_4 は正四面体の周及び内部、などとなる。

定理 2.1 n 次元レプリケーター方程式 (2.6) は以下の性質を持つ。

- (1) S_n は不変 (invariant) である。すなわち、初期ベクトル $\mathbf{x}(0)$ が S_n 上であれば、全ての t (正でも負でも良い) に対して $\mathbf{x}(t)$ も S_n 上にある。
- (2) S_n の部分単体も不変である。つまりある添字の集合 $I \subset \{1, \dots, n\}$ に対し、 S_n のうち、添字が I に含まれる成分が 0 であるものを S_n^I と書くことにすると、 S_n^I も不変である。
 - 例えば $n=4$ で $I = \{2, 3\}$ ならば、

$$S_n^I = \{ \mathbf{x} = (x_1, 0, 0, x_4) \mid x_1, x_4 \geq 0, x_1 + x_4 = 1 \} \quad (2.16)$$

が不変である、ということである。

- (3) n 次元単体の内部

$$\text{int}S_n = \{ \mathbf{x} = (x_1, \dots, x_n) \mid \underline{x_i > 0}, x_1 + \dots + x_n = 1 \} \quad (2.17)$$

も不変である。

- (4) f_i の代わりに、 f_i に正の線形変換を施した

$$g_i = cf_i + c' \quad (c > 0) \quad (2.18)$$

を用いても、解の軌道は変わらない。

解説 2.7 定理 2.1 を証明する。ただしレプリケーターの方程式の解の存在と一意性 (同じ初期条件を満たす異なる二つの解は存在しないこと) は所与のものとしておく。

- (1) はレプリケーター方程式 (2.6) を $i = 1, \dots, n$ に渡って足し合わせて

$$\dot{x}_1 + \dots + \dot{x}_n = \underbrace{\sum_{i=1}^n x_i f_i}_{=\bar{f}} - \left(\sum_{i=1}^n x_i \right) \bar{f} \quad (2.19)$$

より、 $y = x_1 + \dots + x_n$ とおけば、

$$\dot{y} = (1 - y)\bar{f} \quad (2.20)$$

であるが、 $y(t) = 1$ (定数) は、 $y(0) = 1$ を満たすこの微分方程式の解なので、解の一意性より常に $y(t) = 1$ が分かる。

次に $x_i(t) \geq 0$ を示せばよいが、もし $x_i(t_1) < 0$ なる時刻 t_1 があったとすると、連続性より $x_i(t_2) = 0$ となる時刻 t_2 もあったことになる。しかし $x_i(t) = 0$ (定数) は式 (2.6a) の解であるので、 $x_i(t_2) = 0$ なる初期値から始めた異なる二つの解が存在したことになる、これは解の一意性に反する。よって、 $x_i(t) \geq 0$ である。

(2) は、上の議論と同様に、 I に含まれる添字 i に対して、 $x_i(t) = 0$ は、 $x_i(0) = 0$ を満たすレプリケーター方程式 (2.6) の解であるから、解の一意性より従う。

(3) は、(1) より S_n が不変で、(2) よりどんな添字集合 I に対しても部分単体 S_n^I が不変なのだから、 S_n から部分単体を除いて残った部分である $\text{int}S_n$ も不変である。

(4) は式 (2.18) を式 (2.6) に代入して

$$\dot{x}_i = x_i(g_i - \bar{g}) \quad (2.21)$$

だが、

$$\begin{aligned} g_i &= cf_i + c' \\ \bar{g} &= \sum_{j=1}^n g_j x_j = \sum_{j=1}^n (cf_j + c') x_j = c \left(\sum_{j=1}^n f_j x_j \right) + c' = c\bar{f} + c' \end{aligned} \quad (2.22)$$

より

$$\dot{x}_i = x_i \{ (cf_i + c') - (c\bar{f} + c') \} = cx_i(f_i - \bar{f}) \quad (2.23)$$

であり、解説 1.11 と同じ論法を用いて時刻を元より c 倍速く測る時間単位系 $t' = ct$ を用いれば

$$\frac{dx_i}{dt'} = x_i(f_i - \bar{f}) \quad (2.24)$$

と変形できるから、結論が従う。

2.3 具体例

しばらくレプリケーター方程式で戦略が2つの場合 ($n = 2$) について調べていく。

$n = 2$ の時の場合は $x_1 + x_2 = 1$ であるので、レプリケーター方程式 (2.6) の自由度は $2 - 1 = 1$ であり、したがって、 $(x_1, x_2) = (x, 1 - x)$ として、 x の方程式を求めれば良く、それは

$$\dot{x} = x(1 - x)(f_1 - f_2) \quad (2.25)$$

となる。

解説 2.8 平均適応度は

$$\bar{f} = f_1 x + f_2(1 - x) \quad (2.26)$$

と計算できるので、特に

$$f_1 - \bar{f} = f_1 - \{f_1 x + f_2(1 - x)\} = (1 - x)(f_1 - f_2) \quad (2.27)$$

である。したがって式 (2.6) に代入して式 (2.25) を得る。

2.3.1 タカ・ハトゲーム

以下の利得行列で定まるゲームを考える。

自分\相手	タカ	ハト
タカ	$\frac{v-c}{2}$	v
ハト	0	$\frac{v}{2}$

二個体が価値 $v (> 0)$ の資源を巡って競争をしている。「タカ」戦略は闘争を好み、「ハト」戦略は平和を好む。利得は以下の仮定を満たす。

- (1) タカ同士が出会うと闘争を行い、勝者は v を得て、敗者は戦いのコスト $c (> 0)$ を負う。勝つ確率は $1/2$ なので、両者の期待利得は $(v - c)/2$ である。
- (2) タカとハトが出会うとタカは資源 v を独占し、ハトは資源を譲るため利得は 0 である。
- (3) ハト同士が出会うと両者は平和的に資源を分割し、両者の利得は $v/2$ である。

このゲームを**タカ・ハトゲーム (hawk-dove game)** と呼ぶ。タカ・ハトゲームは動物の儀礼的な闘争を説明するために Maynard-Smith & Price (1973) によって導入された。

集団の大きさを無限とし、集団中からランダムに選ばれた相手と対戦し利得 u を得るとする。この時、タカ、ハトそれぞれの平均利得は

$$\begin{aligned} u_1 &= \frac{v-c}{2}x_1 + vx_2 \\ u_2 &= 0 \cdot x_1 + \frac{v}{2}x_2 \end{aligned} \tag{2.28}$$

と計算できる (以下、タカの添字を 1、ハトの添字を 2 とする)。得られた利得 u は線形変換を

$$f = cu + c' \quad (c > 0) \tag{2.29}$$

で増殖率 f に換算されるとする。

定理 2.1 の (4) より、レプリケーター方程式の増殖率に式 (2.29) の f を用いても式 (2.28) の u を用いても結果は変わらないので、以降始めから「増殖率=利得」であるとして差し支えない。したがって

$$\begin{aligned} f_1 &= \frac{v-c}{2}x_1 + vx_2 \\ f_2 &= 0 \cdot x_1 + \frac{v}{2}x_2 \end{aligned} \tag{2.30}$$

である。これを式 (2.25) に代入して、 $(x_1, x_2) = (x, 1-x)$ を使って整理して

$$\dot{x} = x(1-x) \cdot \frac{1}{2}(v-cx) \tag{2.31}$$

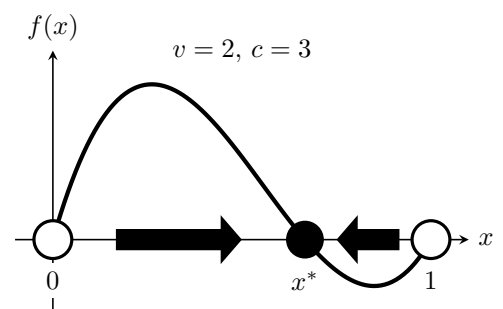
を得る。ただしここで x はタカ戦略の頻度である。

問題 2.9 式 (2.31) を導出せよ。

以降 $c > v$ (闘争で負けるコストは、闘争で勝って得る資源量より大きい) と仮定する。このとき方程式 (2.31) の右辺を $f(x)$ とおくと、 $f(x)$ の符号は

x	0	...	v/c	...	1
$f(x)$	0	+	0	-	0

のように変化する。したがって、 $x^* = v/c$ が大域的に安定な平衡点であり、タカとハトの多型が実現する。



問題 2.10 $c < v$ の時の式 (2.31) の振る舞いを調べよ。

2.3.2 一般の 2 戦略ゲーム

$n = 2$ 戦略 s_1, s_2 からなる一般の 2 人ゲームを考える。自身を行、相手プレイヤーを列に書いたこのゲームの利得行列は

$$\begin{matrix} & s_1 & s_2 \\ s_1 & \begin{pmatrix} a & b \end{pmatrix} \\ s_2 & \begin{pmatrix} c & d \end{pmatrix} \end{matrix} \quad (2.32)$$

と書ける。ここでこの行列の各成分は自身の利得を表す。

先程同じようにランダム相互作用を仮定すると、各戦略の平均利得は

$$\begin{aligned} f_1 &= ax_1 + bx_2 \\ f_2 &= cx_1 + dx_2 \end{aligned} \quad (2.33)$$

と書けるから、式 (2.25) に代入して、 $(x_1, x_2) = (x, 1-x)$ を使って整理して、レプリケーター方程式

$$\begin{aligned} \dot{x} &= x(1-x)(\alpha x + \beta) \\ \begin{pmatrix} \alpha \equiv a - b - c + d \\ \beta \equiv b - d \end{pmatrix} \end{aligned} \quad (2.34)$$

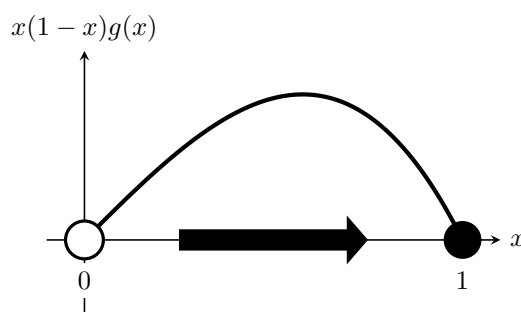
を得る。

問題 2.11 式 (2.34) を導け。

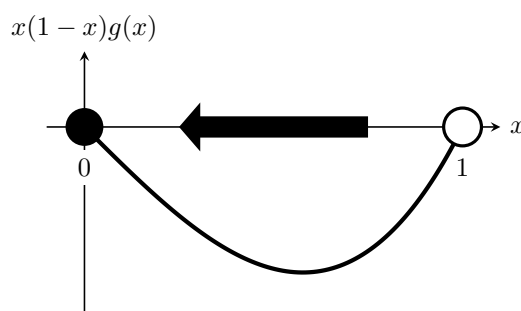
式 (2.34) で $g(x) = \alpha x + \beta$ とおく。 $g(x)$ は線型関数なので、 $g(0) = b - d$ 、 $g(1) = a - c$ の符号で、進化ダイナミクスの性質は以下の表のように完全分類できる。

$g(0) \setminus g(1)$	正 ($a > c$)	0 ($a = c$)	負 ($a < c$)
正 ($b > d$)	(i)	(i)	(iv)
0 ($b = d$)	(i)	(iii)	(ii)
負 ($b < d$)	(v)	(ii)	(ii)

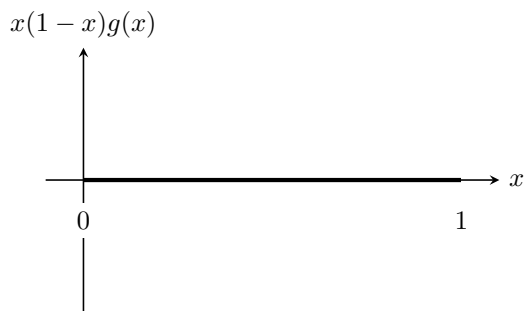
- (i) $0 < x < 1$ で常に $g(x) > 0$ であり戦略 s_1 が常に有利である。 x は時間とともに単調に増加し、 $x(\infty) = 1$ である。進化ダイナミクスは右図のようになる。



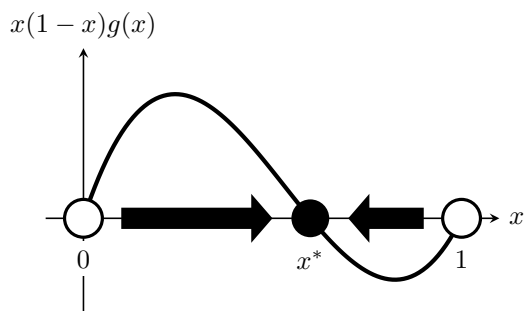
- (ii) $0 < x < 1$ で常に $g(x) < 0$ であり戦略 s_1 が常に不利である。 x は時間とともに単調に減少し、 $x(\infty) = 0$ である。進化ダイナミクスは右図のようになる。



- (iii) $0 < x < 1$ で常に $g(x) = 0$ であり戦略 s_1 と s_2 は常に中立である。
 x は初期値から動かず $x(t) = x(0)$ である。



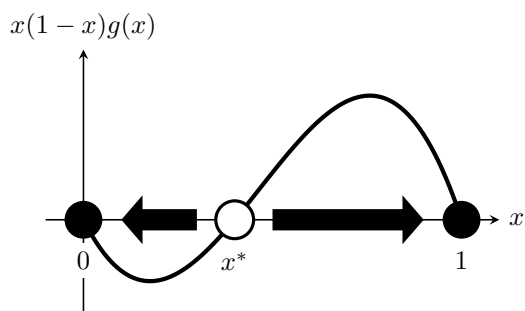
- (iv) 内部に $g(x^*) = 0$ なる安定平衡点 $x^* = -(\alpha/\beta)$ が存在し、任意の初期値 $0 < x(0) < 1$ から始めた時、 x はこの安定平衡点に収束する ($x(\infty) = x^*$)。進化ダイナミクスは右図のようになる。



- (v) 内部に $g(x^*) = 0$ なる不安定平衡点 $x^* = -(\alpha/\beta)$ が存在し、以下のような双安定性 (bistability) を示す。

- 初期値が $x(0) < x^*$ であれば、 x は時間とともに単調に減少し、 $x(\infty) = 0$ である。
- 初期値が $x(0) > x^*$ であれば、 x は時間とともに単調に増加し、 $x(\infty) = 1$ である。

進化ダイナミクスは右図のようになる。



このように自然選択の方向は、集団の現在の構成によって変化する。このような自然選択を**頻度依存選択 (frequency-dependent selection)** と呼ぶ。

上の (v) のような場合は、ある戦略が集団で頻度を増せば増すほど、その戦略がより**有利**になる。このような頻度依存選択を特に**正の頻度依存選択 (positive frequency-dependent selection)** と呼ぶ。

反対に、上の (iv) のような場合は、ある戦略が集団で頻度を増せば増すほど、その戦略がより**不利**になる。このような頻度依存選択を特に**負の頻度依存選択 (negative frequency-dependent selection)** と呼ぶ。

2.3.3 非線形公共財ゲーム

2人ゲームでない場合として、 $N (> 2)$ 人で行われる非線形公共財ゲームを考える。戦略は協力 (s_1) と非協力 (s_2) とする。

まず次の「全員参加型」公共財ゲームを考えよう。

- 集団から N 人のグループがランダムに形成される。協力者はコスト $c (> 0)$ を払って協力するが非協力者は何も支払わない。 N 人全員が協力者であったグループは、協力によって各自が利益 $b (> c)$ を得るものとする。それ以外のグループでは利益は生じない。

このゲームにおいて、各戦略の利得は

$$\begin{aligned} f_1 &= -c + x^{N-1}b \\ f_2 &= 0 \end{aligned} \tag{2.35}$$

と計算できる。なぜなら協力者のみが、グループの残り $(N-1)$ 人が全員協力者であるときに利益 b をあげるチャンスがあり、その確率は x^{N-1} であるからである。この利得をレプリケーター方程式 (2.25) に代入して、

$$\dot{x} = x(1-x)(-c + x^{N-1}b) \tag{2.36}$$

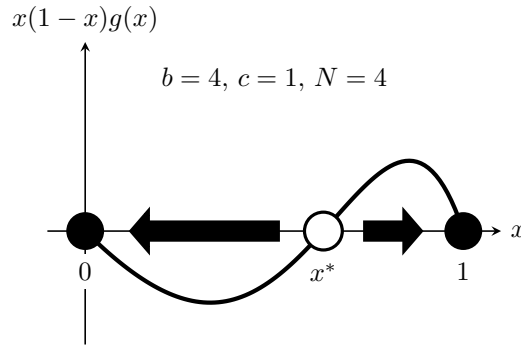
を得る。 $g(x) = -c + x^{N-1}b$ とおくと、 $g(x)$ の符号は以下のように変化する。

x	0	...	$\sqrt[N-1]{\frac{c}{b}}$...	1
$g(x)$	-	-	0	-	0

したがって内部平衡点

$$x^* = \sqrt[N-1]{\frac{c}{b}} \quad (2.37)$$

は不安定で、この点を境として進化ダイナミクスは双安定性を見せる。



次に「最低一人参加型」公共財ゲームを考えよう。

- 集団から N 人のグループがランダムに形成される。協力者はコスト $c(> 0)$ を払って協力するが非協力者は何も支払わない。一人でも協力者がいたグループは、協力によって各自が利益 $b(> c)$ を得るものとする。全員が非協力者のグループでは利益は生じない。

このゲームにおいて、各戦略の利得は

$$\begin{aligned} f_1 &= -c + b \\ f_2 &= b\{1 - (1-x)^{N-1}\} \end{aligned} \quad (2.38)$$

と計算できる。なぜなら非協力者が利益 b を得るには、グループの残り $(N-1)$ 人の中に少なくとも一人協力者が含まれれば良い。その確率は、1 から、全員が非協力者である確率 $(1-x)^{N-1}$ を引いて求められる。この利得をレプリケーター方程式 (2.25) に代入して、

$$\dot{x} = x(1-x)\{-c + (1-x)^{N-1}b\} \quad (2.39)$$

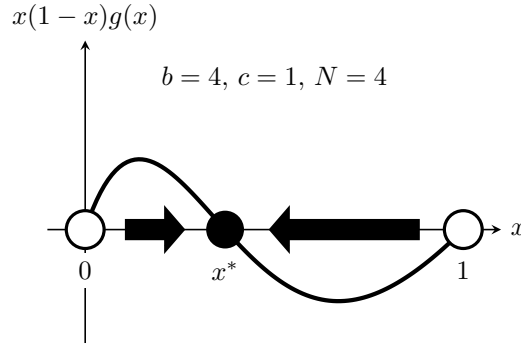
を得る。 $g(x) = -c + (1-x)^{N-1}b$ とおくと、 $g(x)$ の符号は以下のように変化する。

x	0	...	$1 - \sqrt[N-1]{\frac{c}{b}}$...	1
$g(x)$	+	+	0	-	-

したがって内部平衡点

$$x^* = 1 - \sqrt[N-1]{\frac{c}{b}} \quad (2.40)$$

は安定で、任意の初期値 $0 < x(0) < 1$ から始めて、集団は $x = x^*$ の多型状態に収束する。



2.4 突然変異がある場合

2.4.1 Replicator-mutator 方程式

2.2 章の仮定に代えて、突然変異があるために個体は必ずしもクローナルな子孫を生み出さないとする。特に戦略 s_i の親が戦略 s_j の子を産む確率を $q_{i \rightarrow j}$ とおく。確率なので $\sum_{j=1}^n q_{i \rightarrow j} = 1$ である。

以下のような生活環を考える。

- (1) 戦略 s_i を持つ個体は、出生率 b_i で子を産み、確率 $q_{i \rightarrow j}$ でその子は戦略 j を取る。
 - 繁殖率 b_i は時刻 t に依存して変化してよく、また戦略頻度 x_1, \dots, x_n に依存して変化してもよく、添字 i に依存してもよい。
- (2) 戦略 s_i を持つ個体は、死亡率 d_i で死亡する。
 - 死亡率 d_i は、時刻 t に依存して変化してよく、また戦略頻度 x_1, \dots, x_n に依存して変化してもよく、添字 i に依存してもよい。

以上の仮定より、戦略 s_i を持つ個体数 N_i に対して以下の微分方程式が成り立つ。

$$\dot{N}_i = \sum_{j=1}^n q_{j \rightarrow i} b_j N_j - d_i N_i \quad (2.41)$$

問題 2.12 式 (2.41) を導け。

これを頻度 $x_i = N_i/N$ の微分方程式に書き換えて

$$\dot{x}_i = \left(\sum_{j=1}^n q_{j \rightarrow i} b_j x_j - \bar{b} x_i \right) - (d_i - \bar{d}) x_i \quad (2.42)$$

を得る。

解説 2.13 式 (2.42) を導こう。式 (2.10) に式 (2.41) を代入して

$$\begin{aligned} \dot{x}_i &= \frac{\sum_{j=1}^n q_{j \rightarrow i} b_j N_j - d_i N_i}{N} - x_i \frac{\sum_{k=1}^n \sum_{\ell=1}^n q_{\ell \rightarrow k} b_{\ell} N_{\ell} - d_k N_k}{N} \\ &= \sum_{j=1}^n q_{j \rightarrow i} b_j x_j - x_i \sum_{k=1}^n \sum_{\ell=1}^n q_{\ell \rightarrow k} b_{\ell} x_{\ell} - \left(d_i x_i - x_i \underbrace{\sum_{k=1}^n d_k x_k}_{\bar{d}} \right) \\ &= \sum_{j=1}^n q_{j \rightarrow i} b_j x_j - x_i \underbrace{\sum_{\ell=1}^n \left(\underbrace{\sum_{k=1}^n q_{\ell \rightarrow k}}_{=1} \right) b_{\ell} x_{\ell}}_{=\bar{b}} - (d_i - \bar{d}) x_i \end{aligned} \quad (2.43)$$

から結論を得る。

特に死亡率 d_i が戦略によらず一定値 $d_i \equiv d$ であるときには、 $d_i - \bar{d} = 0$ となるから、式 (2.42) は

$$\dot{x}_i = \sum_{j=1}^n q_{j \rightarrow i} b_j x_j - \bar{b} x_i \quad (2.44)$$

となる。式 (2.42) や式 (2.44) を replicator-mutator 方程式と呼ぶ。

解説 2.14 Page & Nowak (2002) は式 (2.44) の b を f で置き換えたものを replicator-mutator と呼んでいるが、突然変異は出生時に起こり死亡時には起こらないため、 q は b にはかかっても、 d にはかからない。したがって式 (2.42) や式 (2.44) が適切である。

2.4.2 エラーカストロフ理論

Replicator-mutator 方程式の応用として、配列の進化を考えよう。機能的な DNA や RNA 配列は効率的な自己複製を可能とする一方で、そうでない junk 配列は自己複製の効率が悪かったり、正確に自己複製をする能力を持たないかもしれない。以下のような仮定をおく。

- 塩基配列の空間を考え、ある自己複製子 s_1 を考えよう。自己複製子の複製率を $r > 1$ とする。自己複製子は長さ L となる配列で、自己複製時に各塩基あたり q の確率で突然変異を起こし別の塩基に置き換わってしまうとする。配列長を変化させる突然変異は以降は考えない。以下で L は十分大きく、反対に q は十分小さく、積 Lq は 1 と比べてあまり大きくも小さくもない有限値であるとする。
- 自己複製子以外の長さ L の配列を全て junk な配列 s_2 とみなす。junk 配列の複製率はどの配列も 1 とする。junk 配列は複製時に突然変異を起こし別の配列を生み出すかもしれないが、その配列もまた junk 配列であるとする。とても微小な確率で junk 配列が突然変異で自己複製子 s_1 の配列を取り戻すことも理論的にはあり得るが、 L は十分に大きいのでこの確率は無視できることにする。
- 自己複製子 s_1 も junk 配列 s_2 も、等しい速さで分解されるとする。

このようなプロセスの下で集団には様々な配列が分布することになる。このような配列の集まりを種になぞらえて擬種 (quasispecies) (もしくは準種) と呼ぶ。以上の仮定の下、自己複製子および junk 配列の頻度 $(x_1, x_2) = (x, 1 - x)$ の進化ダイナミクスを考えよう。

突然変異確率は以下で与えられる。

$$\begin{pmatrix} q_{1 \rightarrow 1} & q_{1 \rightarrow 2} \\ q_{2 \rightarrow 1} & q_{2 \rightarrow 2} \end{pmatrix} = \begin{pmatrix} 1 - Q & Q \\ 0 & 1 \end{pmatrix} \quad (2.45)$$

ただし、 $1 - Q = (1 - q)^L$ は複製時にどの塩基も突然変異を受けない確率である。

仮定 $(b_1, b_2) = (r, 1)$ を replicator-mutator 方程式 (2.44) に代入して

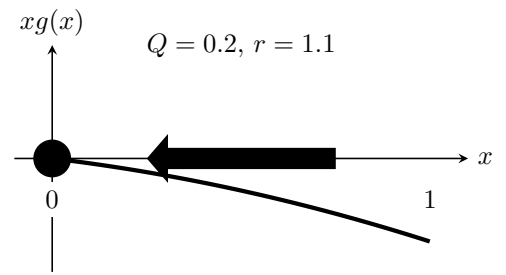
$$\dot{x} = (1 - Q)rx - \underbrace{\{rx + 1 \cdot (1 - x)\}}_{=\bar{b}} x = x\{(1 - Q)r - 1 - (r - 1)x\} \quad (2.46)$$

を得る。

$g(x) = (1 - Q)r - 1 - (r - 1)x$ とおくと、 $g(x)$ は減少する線型関数で、 $g(1) = -Qr < 0$ であるから、 $g(0) = (1 - Q)r - 1$ の符号で以下のように場合分けできる。

(1) $g(0) \leq 0$ の時：

$0 < x < 1$ で常に $g(x) < 0$ であり、進化ダイナミクスは右図のようになる。すなわち自己複製子は消滅し、junk 配列だけになる。

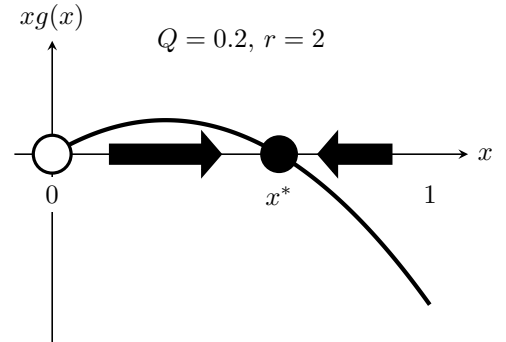


(2) $g(0) > 0$ の時:

内部に $g(x^*) = 0$ を満たす安定平衡点

$$x^* = \frac{(1-Q)r-1}{r-1} \quad (2.47)$$

が存在し、自己複製子が多くある初期値 $x(0) \approx 1$ から始めた場合、自己複製子は頻度 x^* で集団に保たれる。



以上より、自己複製子が集団に保たれる条件は $g(0) > 0$ すなわち

$$(1-Q)r > 1 \quad (2.48)$$

である。ここで近似公式 $1-Q = (1-q)^L \approx \exp(-Lq)$ を用い、また自己複製子 s_1 の複製効率 r はたとえば自然対数 $e = 2.718 \dots$ 程度であるとすると、式 (2.48) は

$$\exp(-Lq + 1) > 1 \quad (2.49)$$

と整理され、これは

$$q < \frac{1}{L} \quad (2.50)$$

と書ける。不等式 (2.50) の右辺は、ある配列長 L を持つ複製子にとって、自己と同じ複製子を集団に保持することができるような突然変異率の上限を表す。

不等式 (2.50) が破れることによって自己複製子が集団から消失する現象をエラーカタストロフ (error catastrophe) と呼ぶ (Eigen 1971)。ウイルスの塩基あたりの突然変異率がヒトのそれと比べて大きく、進化したウイルスが毎年我々を苦しめるのは、このエラーカタストロフ理論によりその理由の一部が説明される。なぜならウイルスの DNA/RNA 全長はヒトのそれよりも圧倒的に短く、したがってエラーカタストロフ理論によると、ウイルスのほうがヒトよりも圧倒的に高い突然変異率を持つことができると予測できるからである。

3 量的遺伝学入門

(本章の概要と目的) 本章では古典的量的遺伝学の概要を学ぶ。基礎的理論の体系は Fisher の 1930 年の著書に遡る。本章では二倍体有性生殖生物に焦点をおく。量的遺伝学では一般に多くの遺伝子座が個体の表現型に関わっていると考え、遺伝によって親と子が似る程度や、自然選択によって集団の表現型平均が変化する程度について議論する。ここで重要なのは親は自らの表現型を子に渡すことはできず、遺伝子しか渡せないという点である。二倍体生物では父と母から受け継いだ遺伝子の組み合わせによって表現型が決定するので、親は結果として表現型の一部しか子に渡すことができない。

この議論を重回帰、分散分析の枠組みを用いて統計的に展開したのが Fisher の功績である。特に、どのような親を掛け合わせるとどのような子が生まれるかを知ることは農学、特に育種学において重要であるため、その理論体系は古くから発展した。これらの体系をまとめた良書として Falconer(1960) がある。

本章で使われる数理は主に統計学であり、特に平均、分散、共分散の計算が頻出する。最も重要な概念として表現型の分解があり、表現型は遺伝成分と環境成分に分かれること、そしてその遺伝成分はさらに相加的な部分と非相加的な部分に分解されることを学ぶ。

本章の目的の一つは表現型を統計的に分解する考え方と、その具体的計算法を学ぶことである。もう一つの目的は、その中でも相加的な部分が本質的な役割を果たし、特に子への遺伝や自然選択に対する集団の応答においては、遺伝率と呼ばれる概念が中心的役割を果たすことを理解することである。

3.1 QTL と量的遺伝学

身長のような量的な表現型を考えよう。これは個体の持つ遺伝子型とどのような関係があるのだろうか。

身長のような表現型は

- 複数の遺伝子座における遺伝子
- 個体の置かれた環境

などが総合的に合わさった結果として決定していると考えられる。以下、このような量的形質 (quantitative trait) についてその統計学を展開していこう。

個体の量的な表現型の値、つまり**表現型値 (phenotypic value)** を P とする。最も単純なモデルは、 P が遺伝子の効果と環境の効果により決定していると考えられるモデルである。そこで

$$P = G + E \quad (3.1)$$

のように和の形で決定していると考えられる。 G は個体の遺伝子型から決定される表現型の部分で、**遺伝子型値 (genotypic value)** と呼ぶ。それに対し E は個体の置かれている (いた) 環境から決定される表現型の部分で、**環境偏差 (environmental deviation)** と呼ぶ。

解説 3.1 記号 $indv$ で個体のラベルを、記号 $geno$ で遺伝子型のラベルを表すことにすると、式 (3.1) は

$$P_{indv} = G_{geno} + E_{indv} \quad (3.2)$$

と言っている。つまり言葉に直せば、「 $indv$ さんの表現型値は、($indv$ さんの持つ) 遺伝子型 $geno$ から決まる遺伝子型値と、 $indv$ さんの置かれた環境から決まる環境偏差で決まる」ということが言いたい。

以下で環境偏差の集団平均は 0 であると仮定する。つまり式で書けば

$$\bar{E} = 0 \quad (3.3)$$

である (以降バーをつけたら、特別の断りがない限り、それは集団平均を表すこととする)。また遺伝子型値と環境偏差は無相関であるとする。共分散の言葉で書けば

$$\text{cov}[G, E] = 0 \quad (3.4)$$

である。

解説 3.2 実際には式 (3.1) や式 (3.4) が満たされている保証はない。例えば遺伝子型と環境の相乗効果、つまり統計学で言うところの両者の交互作用、が存在すれば式 (3.1) のように P は G と E の単純な足し算では書けない。また、特定の遺伝子を持っていると特定の環境に晒されやすい場合は、式 (3.4) の共分散は 0 ではない。しかしながら以降ではそのようなことはない仮定する。

問題 3.3 特定の遺伝子を持つと特定の環境に晒されやすいような場合とは、どのような例があるか。

遺伝子型値 G に貢献しているのは、実際には複数の遺伝子座である。量的形質に影響を与えるこれらの遺伝子座を **QTL (Quantitative Trait Loci; 量的形質遺伝子座)** と呼ぶ。以下、QTL が合計 L 座位あるとし、それらを $\ell = 1, \dots, L$ と名付ける。

3.2 相加遺伝値

本章の以下では二倍体で任意交配をしている無限集団を考える。しばらくは $L = 1$ (単一座位) の場合を考えよう。

個体の遺伝子型値 G はその個体の持つ遺伝子からどのように説明できるのだろうか？注目している座位には対立遺伝子 A_1, A_2 があると、それぞれの遺伝子頻度を p, q ($p + q = 1$) とおく。

個体が持つ遺伝子を表す以下のような指示変数 X を導入する。

- 注目する遺伝子が A_1 ならば、 $X = 1$ の値を取る。
- 注目する遺伝子が A_2 ならば、 $X = 0$ の値を取る。

各個体は注目する座位に二つの遺伝子を持つので、それらの指示変数を X_1, X_2 とおく。

このとき G を X_1 と X_2 で説明することを試みる。一番簡単なモデルは、 G を X_1 と X_2 の足し算で

$$G = \alpha X_1 + \alpha X_2 + \beta + \epsilon \quad (3.5)$$

と説明するモデルである (相加的モデル)。ここで α, β は何らかの定数である。 ϵ は相加的モデルからのずれを表し (注: 環境偏差の値ではないので混乱しないように)、遺伝子型に依存してよいが、 $\bar{\epsilon} = 0$ とする。 X_1 と X_2 の係数 α が共通しているのは、表現型への遺伝子の効果はその遺伝子をどちらの親から受け継いだかに依らず等しいとする仮定に基づく。式 (3.5) の集団平均 $\bar{G} = \alpha \bar{X}_1 + \alpha \bar{X}_2 + \beta$ を両辺から差し引けばこれは

$$G - \bar{G} = \alpha(X_1 - \bar{X}_1) + \alpha(X_2 - \bar{X}_2) + \epsilon \quad (3.6)$$

と書ける (特に $\bar{X} = \bar{X}_1 = \bar{X}_2 = p$)。 $X_{1+2} = X_1 + X_2$ と書けば、相加的モデルはさらにシンプルに

$$G - \bar{G} = \alpha(X_{1+2} - \bar{X}_{1+2}) + \epsilon \quad (3.7)$$

と書ける (特に $\bar{X}_{1+2} = 2p$)。

解説 3.4 先程のラベル (記号 $indv$ で個体のラベルを、記号 $geno$ で遺伝子型のラベル) で言えば、式 (3.7) は

$$G_{geno} - \bar{G} = \alpha(X_{geno,1+2} - \bar{X}_{1+2}) + \epsilon_{geno} \quad (3.8)$$

と言っている。 ϵ の添字は $geno$ であって $indv$ でない。つまり言葉に直せば、 ϵ_{geno} はある遺伝子型 $geno$ を持つ各個体に共通の値である。

式 (3.7) の誤差 ϵ を考え、その二乗平均 J を最小にするような α を求めることにしよう (最小二乗法)。形式的に書くと

$$J = \sum_{geno} P_{geno} \epsilon_{geno}^2 \rightarrow \min \quad (3.9)$$

である。ここで、 P_{geno} は遺伝子型 $geno$ を持っている個体の頻度である。

解説 3.5 式 (3.9) の J を最小化する α を求めよう。

- 遺伝子型 A_1A_1 の遺伝子型値を G_{11}
- 遺伝子型 A_1A_2 の遺伝子型値を G_{12}
- 遺伝子型 A_2A_2 の遺伝子型値を G_{22}

とおく。HW 平衡を仮定すると、式 (3.9) に現れる各量は

遺伝子型	頻度	$G - \bar{G}$	$X_{1+2} - \bar{X}_{1+2}$
A_1A_1	p^2	$G_{11} - \bar{G}$	$2 - 2p$
A_1A_2	$2pq$	$G_{12} - \bar{G}$	$1 - 2p$
A_2A_2	q^2	$G_{22} - \bar{G}$	$0 - 2p$

と計算できる。ただし

$$\bar{G} = p^2G_{11} + 2pqG_{12} + q^2G_{22} \quad (3.10)$$

である。この時、式 (3.9) は

$$J = p^2\{G_{11} - \bar{G} - \alpha(2 - 2p)\}^2 + 2pq\{G_{12} - \bar{G} - \alpha(1 - 2p)\}^2 + q^2\{G_{22} - \bar{G} - \alpha(0 - 2p)\}^2 \quad (3.11)$$

となり、これを最小にする α は

$$\alpha = p(G_{11} - G_{12}) + q(G_{12} - G_{22}) \quad (3.12)$$

と求まる。この時の J の最小値は

$$J_{\min} = p^2q^2(G_{11} - 2G_{12} + G_{22})^2 \quad (3.13)$$

である。

問題 3.6 式 (3.12) および式 (3.13) を示せ。

上の計算より、相加的モデル (3.7) の下では、個体の遺伝子型値 G は、遺伝子型値の集団平均 \bar{G} に比べて

- 遺伝子 A_1 を一つ持つ毎に $\alpha(1 - \bar{X}) = \alpha(1 - p) = q\alpha$ ($\equiv \alpha_1$) だけ増加する
- 遺伝子 A_2 を一つ持つ毎に $\alpha(0 - \bar{X}) = \alpha(-p) = -p\alpha$ ($\equiv \alpha_2$) だけ増加する

と言うことが妥当である (式 (3.6) を参照のこと)。これら α_1, α_2 をそれぞれ遺伝子 A_1, A_2 の平均効果 (average effect) と呼ぶ。また $\alpha = \alpha_1 - \alpha_2$ を遺伝子置換の平均効果 (average effect of the gene substitution) と呼ぶ。したがって、

- 遺伝子型 A_1A_1 に値 $A = 2\alpha_1 = 2q\alpha$
- 遺伝子型 A_1A_2 に値 $A = \alpha_1 + \alpha_2 = (q - p)\alpha$
- 遺伝子型 A_2A_2 に値 $A = 2\alpha_2 = -2p\alpha$

をそれぞれ割り振ると、これらが相加的モデル (3.5) の下で、遺伝子型値の集団平均 \bar{G} からのずれをもっともうまく言い表していることになる。この値 A のことを各遺伝子型の相加遺伝値 (additive genetic value) もしくは育種価 (breeding value) と呼ぶ。

定義より $A = \alpha(X_{1+2} - \bar{X}_{1+2})$ なので、両辺の集団平均を取ることで $\bar{A} = 0$ が分かる。

例 3.7 数値例として、遺伝子型値が $(G_{11}, G_{12}, G_{22}) = (36, 11, 11)$ で与えられ、遺伝子 A_1, A_2 の頻度が $(p, q) = (0.6, 0.4)$ で与えられる場合を考えてみよう。

遺伝子型値の集団平均は式 (3.10) より

$$\bar{G} = (0.6)^2 \cdot 36 + 2(0.6)(0.4) \cdot 11 + (0.4)^2 \cdot 11 = 20 \quad (3.14)$$

である。また遺伝子置換の平均効果 α は式 (3.12) より

$$\alpha = (0.6)(36 - 11) + (0.4)(11 - 11) = 15 \quad (3.15)$$

であり、遺伝子 A_1, A_2 の平均効果は

$$\alpha_1 = 0.4 \cdot 15 = 6, \quad \alpha_2 = -0.6 \cdot 15 = -9 \quad (3.16)$$

である。したがって遺伝子型 A_1A_1, A_1A_2, A_2A_2 の遺伝子型値はそれぞれ、

$$A_{11} = 2\alpha_1 = 12, \quad A_{12} = \alpha_1 + \alpha_2 = -3, \quad A_{22} = 2\alpha_2 = -18 \quad (3.17)$$

となる。この時、遺伝子型値の集団平均は

$$\bar{A} = (0.6)^2 \cdot 12 + 2(0.6)(0.4) \cdot (-3) + (0.4)^2 \cdot (-18) = 0 \quad (3.18)$$

となり、確かに $\bar{A} = 0$ を満たしている。

解説 3.8 定義から明らかのように、相加遺伝値は遺伝子頻度 p, q に依存しており、したがってどのような集団で考えるかに依存して変わる値である。

例えば集団が対立遺伝子 A_1 で占められている場合、 $q = 0$ であるから、 $\alpha_1 = 0$ であり、したがって遺伝子型 A_1A_1 の相加遺伝値は $A = 0$ である。相加遺伝値は集団平均からのずれを表す指標であると考えれば、これは納得がいく。なぜならこの場合、集団の「平均的な」個体は他ならぬ A_1A_1 であるからである。

反対に集団が対立遺伝子 A_2 で占められている場合、 $q = 1$ であるから、 $\alpha_1 = \alpha = G_{12} - G_{22}$ であり、したがって遺伝子型 A_1A_1 の相加遺伝値は $A = 2(G_{12} - G_{22})$ である。対立遺伝子 A_2 で占められている集団においてある一つの A_2 が突然 A_1 に置換されれば、それを持つ個体の遺伝子型は A_2A_2 から A_1A_2 に変わるはずだから、この置換がもたらす遺伝子型値への効果は $G_{12} - G_{22}$ であり、したがって、遺伝子型 A_1A_1 の相加遺伝値はこの 2 倍と計算される。

この段階では相加遺伝値の定義は天下りでありがたみを全く感じないが、以下で議論を進めるとこれが重要な概念であることが徐々に分かる。

もし遺伝子型値が相加的に決まっており、ヘテロ接合体の遺伝子型値が二つのホモ接合体の遺伝子型値の平均値として

$$G_{12} = \frac{G_{11} + G_{22}}{2} \quad (3.19)$$

と決まっているならば、式 (3.13) の J_{\min} はちょうど 0 になる。すなわち相加的モデル (3.7) の誤差項は常に $\epsilon \equiv 0$ である。言い換えれば

$$G - \bar{G} = A \quad (3.20)$$

である。

3.3 顕性偏差

引き続き $L = 1$ の場合を考える。

もし遺伝子型値が式 (3.19) を満たさず、したがって相加的に決まっていなければ、式 (3.13) で $J_{\min} > 0$ である。この非相加性を説明するため、以下の非相加的モデルを考える。

$$G - \bar{G} = \underbrace{\alpha(X_1 - \bar{X}_1) + \alpha(X_2 - \bar{X}_2)}_{=\alpha(X_{1+2} - \bar{X}_{1+2})} + \delta(X_1 - \bar{X}_1)(X_2 - \bar{X}_2) + \epsilon \quad (3.21)$$

ここで項 $(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ は回帰分析で言うところの $X_1 - \bar{X}_1$ と $X_2 - \bar{X}_2$ の交互作用である。また、 ϵ は (遺伝子型に依存してよい) 新しい誤差である。この新しい誤差に関して再び最小二乗法 (3.9) を適用しよう。

解説 3.9 非相加的モデル (3.21) で最小二乗法で推定される α の値と、相加的モデル (3.7) で最小二乗法で推定される α の値は異なるかもしれないと心配になるが、投入される新たな変数 $(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ が変数 $(X_{1+2} - \bar{X}_{1+2})$ と直交 (orthogonal) していれば推定値は変わらないことが重回帰分析の一般論から分かっている。

ここで変数 Y と Z が直交しているとは、 Y と Z の相関係数が 0 であることをいう。これは Y と Z の共分散が 0 であると言っても同じである。またこれは、

$$\overline{Y \cdot Z} = \bar{Y} \cdot \bar{Z} \quad (3.22)$$

(バーは集団平均を表す) ということにも等しい。

今は任意交配をしている無限集団を考えているから、HW 平衡のもとで X_1 と X_2 は確率変数として見た時独立であり、特にデータとして $X_1 - \bar{X}_1$ と $X_2 - \bar{X}_2$ は直交している (注: 無限集団の平均値の話をしているので、大数の法則より、平均値と期待値は特に問題なく同一視できる)。したがって例えば $X_1 - \bar{X}_1$ と $(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ が直交していることは

$$\overline{(X_1 - \bar{X}_1) \cdot (X_1 - \bar{X}_1)(X_2 - \bar{X}_2)} = \overline{(X_1 - \bar{X}_1)^2} \cdot \underbrace{\overline{(X_2 - \bar{X}_2)}}_{=\bar{X}_2 - \bar{X}_2 = 0} = 0 \quad (3.23)$$

から分かる (最初の等号で X_1 と X_2 の独立性を使った)。同じように $X_2 - \bar{X}_2$ と $(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ も直交している。したがって $X_{1+2} - \bar{X}_{1+2} = (X_1 - \bar{X}_1) + (X_2 - \bar{X}_2)$ も $(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ と直交している。

非相加的モデル (3.21) に対し最小二乗法を適用しよう。

解説 3.10 α は先程と同じだから δ を求めれば良い。

遺伝子型	頻度	$G - \bar{G}$	$X_{1+2} - \bar{X}_{1+2}$	$(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$
A_1A_1	p^2	$\bar{P}_{11} - \bar{P}$	$2 - 2p$	$(1 - p)^2$
A_1A_2	$2pq$	$\bar{P}_{12} - \bar{P}$	$1 - p$	$(1 - p)(0 - p)$
A_2A_2	q^2	$\bar{P}_{22} - \bar{P}$	$0 - p$	$(0 - p)^2$

であり、平均二乗誤差は

$$J = p^2 \{G_{11} - \bar{G} - \alpha(2 - 2p) - \delta(1 - p)^2\}^2 + 2pq \{G_{12} - \bar{G} - \alpha(1 - 2p) - \delta(1 - p)(0 - p)\}^2 + q^2 \{G_{22} - \bar{G} - \alpha(0 - 2p) - \delta(0 - p)^2\}^2 \quad (3.24)$$

となり、これを最小にする α は引き続き式 (3.12) で与えられ、 δ は

$$\delta = G_{11} - 2G_{12} + G_{22} \quad (3.25)$$

である。この α, δ のもとで、平均二乗誤差 (3.24) の最小値は

$$J_{\min} = 0 \quad (3.26)$$

である。

そこでこの δ を用いて、各遺伝子型に非相加的モデル (3.21) の $\delta(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ の値、すなわち

- 遺伝子型 A_1A_1 に値 $D = \delta(1 - p)^2 = q^2\delta$
- 遺伝子型 A_1A_2 に値 $D = \delta(1 - p)(0 - p) = -pq\delta$
- 遺伝子型 A_2A_2 に値 $D = \delta(0 - p)^2 = p^2\delta$

を割り振ると、 $J_{\min} = 0$ であることから非相加的モデル (3.21) の誤差項は常に $\epsilon \equiv 0$ であり、したがって $G - \bar{G}$ は相加遺伝値

A とこの D によって

$$G - \bar{G} = A + D \quad (3.27)$$

と完全に説明できることになる。この D を **顕性偏差 (dominance deviation)** と呼ぶ。

定義より $D = \delta(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ だが、両辺の集団平均を取れば、右辺は $X_1 - \bar{X}_1$ と $X_2 - \bar{X}_2$ の直交性より 0 なので、 $\bar{D} = 0$ である。

解説 3.11 もし遺伝子型値が相加的で条件 (3.19) を満たすならば、式 (3.25) より $\delta = 0$ であって、したがってすべての遺伝子型に対して顕性偏差は $D \equiv 0$ である。よって式 (3.20) が再現でき、先程の議論と一貫性がある。

例 3.12 例 3.7 の数値例を引き続き解析しよう。遺伝子型値が $(G_{11}, G_{12}, G_{22}) = (36, 11, 11)$ で与えられ、遺伝子 A_1, A_2 の頻度が $(p, q) = (0.6, 0.4)$ で与えられる場合である。

δ の値は式 (3.25) より

$$\delta = 36 - 2 \cdot 11 + 11 = 25 \quad (3.28)$$

である。したがって遺伝子型 A_1A_1, A_1A_2, A_2A_2 の顕性偏差はそれぞれ、

$$D_{11} = (0.4)^2 \cdot 25 = 4, \quad D_{12} = -(0.6)(0.4) \cdot 25 = -6, \quad D_{22} = (0.6)^2 \cdot 25 = 9 \quad (3.29)$$

となる。

この時、顕性偏差の集団平均は

$$\bar{D} = (0.6)^2 \cdot 4 + 2(0.6)(0.4) \cdot (-6) + (0.4)^2 \cdot 9 = 0 \quad (3.30)$$

となり、確かに $\bar{D} = 0$ を満たしている。

したがって例 3.7 の結果と合わせて、以下の結論を得る。

遺伝子型	頻度	$G - \bar{G}$	=	A	+	D
A_1A_1	0.36	$36 - 20$	=	12	+	4
A_1A_2	0.48	$11 - 20$	=	(-3)	+	(-6)
A_2A_2	0.16	$11 - 20$	=	(-18)	+	9
		$\bar{G} = 20,$	$\alpha = 15,$	$\delta = 25$		

解説 3.13 ここまでをまとめる。常に $G - \bar{G} = A + D$ が成り立っている。

遺伝子型	頻度	$G - \bar{G}$	=	A	+	D
A_1A_1	p^2	$G_{11} - \bar{G}$	=	$2q\alpha$	+	$q^2\delta$
A_1A_2	$2pq$	$G_{12} - \bar{G}$	=	$(q - p)\alpha$	+	$-pq\delta$
A_2A_2	q^2	$G_{22} - \bar{G}$	=	$-2p\alpha$	+	$p^2\delta$
		$\bar{G} = p^2G_{11} + 2pqG_{12} + q^2G_{22}$				
		$\alpha = p(G_{11} - G_{12}) + q(G_{12} - G_{22})$				
		$\delta = G_{11} - 2G_{12} + G_{22}$				

この分解の直観的な説明は以下の通りである。相加遺伝値 A は各親の表現型のうち「子に伝えることができる部分」と考えて良い。それに対して顕性偏差 D は遺伝子の組み合わせによって決まる部分であり、各親の表現型のうち、単一の親の力だけでは「子に伝えることができない部分」と考えて良い。

3.4 表現型分散の分解

以下、任意の変数 Y について、その分散 $\text{var}[Y]$ をしばしば V_Y などと略記する。

式 (3.1) の両辺の分散を考えると $V_P = V_G + V_E + 2\text{cov}[G, E]$ だが、仮定 (3.4) より $\text{cov}[G, E] = 0$ なので、

$$V_P = V_G + V_E \quad (3.31)$$

を得る。

次に式 (3.27) の両辺の分散を考えると $V_G = V_A + V_D + 2\text{cov}[A, D]$ だが、 A と D の定義を思い出すとこれらは直交していたので $\text{cov}[A, D] = 0$ である。したがって、

$$V_G = V_A + V_D \quad (3.32)$$

を得る。したがって式 (3.31) と併せて、分散の分解

$$V_P = \underbrace{V_A + V_D}_{V_G} + V_E \quad (3.33)$$

を得た。

解説 3.14 定義より V_A および V_D は X の言葉で以下のように書ける。

$$\begin{aligned} V_A &= \text{var}[\alpha(X_1 - \bar{X}) + \alpha(X_2 - \bar{X})] \\ &= \alpha^2 \text{var}[X_1 - \bar{X}] + 2\alpha^2 \underbrace{\text{cov}[X_1 - \bar{X}, X_2 - \bar{X}]}_{=0} + \alpha^2 \text{var}[X_2 - \bar{X}] = 2\alpha^2 \text{var}[X - \bar{X}] = 2\alpha^2 \text{var}[X] = 2\alpha^2 pq \end{aligned} \quad (3.34a)$$

$$V_D = \text{var}[\delta(X_1 - \bar{X})(X_2 - \bar{X})] = \delta^2 \overline{(X_1 - \bar{X})^2(X_2 - \bar{X})^2} = \delta^2 \overline{(X_1 - \bar{X})^2} \cdot \overline{(X_2 - \bar{X})^2} = \delta^2 \text{var}[X]^2 = \delta^2 p^2 q^2 \quad (3.34b)$$

ここで $\text{var}[X] = \overline{(X - \bar{X})^2} = pq$ であることを用いた。

問題 3.15 $\text{var}[X] = \overline{(X - \bar{X})^2} = pq$ であることを示せ。

解説 3.16 分散の分解 (3.32) が本当に正しいかを、**解説 3.13** の表を元に具体的に計算してみよう。まず

$$V_G = p^2(G_{11} - \bar{G})^2 + 2pq(G_{12} - \bar{G})^2 + q^2(G_{22} - \bar{G})^2 \quad (3.35)$$

であるが、表にある分解を使って

$$\begin{aligned} V_G &= p^2(A_{11} + D_{11})^2 + 2pq(A_{12} + D_{12})^2 + q^2(A_{22} + D_{22})^2 \\ &= \underbrace{(p^2 A_{11}^2 + 2pq A_{12}^2 + q^2 A_{22}^2)}_{=V_A} + \underbrace{(p^2 D_{11}^2 + 2pq D_{12}^2 + q^2 D_{22}^2)}_{=V_D} + 2(p^2 A_{11} D_{11} + 2pq A_{12} D_{12} + q^2 A_{22} D_{22}) \end{aligned} \quad (3.36)$$

である。ここで $\bar{A} = \bar{D} = 0$ であることを使った。式 (3.36) の最後のカッコの中は直接計算すれば

$$p^2 A_{11} D_{11} + 2pq A_{12} D_{12} + q^2 A_{22} D_{22} = p^2 \cdot 2q\alpha \cdot q^2\delta + 2pq \cdot (q-p)\alpha \cdot (-pq\delta) + q^2 \cdot (-2p\alpha) \cdot p^2\delta = 0 \quad (3.37)$$

であるので、分散の分解 (3.32) が成り立っていることが確かめられた。

以下の用語を定義する。

- V_P を表現型分散 (phenotypic variance) と呼ぶ。
- V_G を遺伝分散 (genetic variance) と呼ぶ。
- V_E を環境分散 (environmental variance) と呼ぶ。
- V_A を相加遺伝分散 (additive genetic variance) と呼ぶ。
- V_D を顕性分散 (dominance variance) と呼ぶ。

このとき、表現型分散に占める遺伝分散の比率、

$$H^2 \equiv \frac{V_G}{V_P} \quad (3.38)$$

を**広義の遺伝率 (broad-sense heritability)**と呼ぶ。

また、表現型分散に占める相加遺伝分散の比率、

$$h^2 \equiv \frac{V_A}{V_P} \quad (3.39)$$

を**狭義の遺伝率 (narrow-sense heritability)**と呼ぶ。

解説 3.17 本章で延々と計算をしてきた目的は、この狭義の遺伝率 h^2 が量的遺伝学を支配していることを以下で示すためにある。

例 3.18 **例 3.12**の数値例を引き続き解析しよう。遺伝子型値が $(G_{11}, G_{12}, G_{22}) = (36, 11, 11)$ で与えられ、遺伝子 A_1, A_2 の頻度が $(p, q) = (0.6, 0.4)$ で与えられる場合である。

新たな仮定を加える。各個体の環境偏差 E は平均 0、分散 12^2 のある分布から取り出された値であるとする。したがって環境分散は $V_E = 12^2 = 144$ である。

相加遺伝分散 V_A および顕性分散 V_D は、式 (3.34) よりそれぞれ

$$V_A = 2 \cdot 15^2 \cdot (0.6)(0.4) = 108, \quad \delta = 25^2 \cdot (0.6)^2(0.4)^2 = 36 \quad (3.40)$$

と求められる。したがって、遺伝分散は $108 + 36 = 144$ であると期待されるが、これは式 (3.35) による直接計算

$$V_G = (0.6)^2 \cdot (36 - 20)^2 + 2(0.6)(0.4) \cdot (11 - 20)^2 + (0.4)^2 \cdot (11 - 20)^2 = 144 \quad (3.41)$$

によっても確かめられる。したがって、以下の表現型分散の分解を得る。

$$\underbrace{288}_{V_P} = \underbrace{144}_{V_G} + \underbrace{144}_{V_E} = (\underbrace{108}_{V_A} + \underbrace{36}_{V_D}) + \underbrace{144}_{V_E} \quad (3.42)$$

したがって、広義の遺伝率は

$$H^2 = \frac{144}{288} = \frac{1}{2} \quad (3.43)$$

と計算でき、狭義の遺伝率は

$$h^2 = \frac{108}{288} = \frac{3}{8} \quad (3.44)$$

と計算できる。

3.5 表現型の類似度

3.5.1 中間親への回帰と遺伝率

父親の持つ遺伝子の指示変数を (X_{p1}, X_{p2}) 、母親の持つ遺伝子の指示変数を (X_{m1}, X_{m2}) とおく (p は “paternal”、m は “maternal” の略)。また、この両者の子の遺伝子の指示変数は (X_{o1}, X_{o2}) とおく。(o は “offspring” の略)。一般性を失わずに

$$\begin{aligned} X_{o1} &= X_{p1} \\ X_{o2} &= X_{m1} \end{aligned} \quad (3.45)$$

と仮定してよい。

このとき、父の表現型 P_p 、母の表現型 P_m 、子の表現型 P_o はどのような関係にあるだろうか。単一の親子の組に対してではなく、親子関係一般に関して統計学を展開しよう。

子の表現型 P_o の、一方の親 (たとえばここでは父親) の表現型 P_p への回帰係数を求める。すなわち

$$b_{op} = \frac{\text{cov}[P_p, P_o]}{\text{var}[P_p]} \quad (3.46)$$

を求める。計算の結果、

$$b_{op} = \frac{1}{2} \frac{V_A}{V_P} \quad \left(= \frac{h^2}{2} \right) \quad (3.47)$$

を得る。母親の表現型 P_m への回帰も同様に

$$b_{om} = \frac{1}{2} \frac{V_A}{V_P} \quad \left(= \frac{h^2}{2} \right) \quad (3.48)$$

である。すなわち、横軸に片親の表現型、縦軸に子の表現型をプロットすると、回帰直線の傾きは $h^2/2$ である。ここで h^2 は狭義の遺伝率である。

次に仮想的な個体である**中間親 (midparent)** を考える。この仮想的な個体の表現型は、父親と母親の表現型の平均値

$$P_{mid} \equiv \frac{P_p + P_m}{2} \quad (3.49)$$

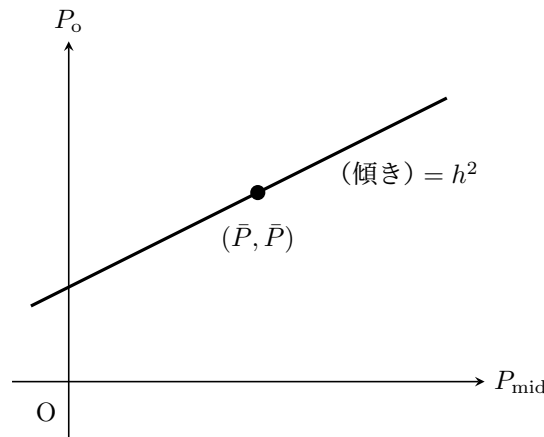
であると仮定する。すると、子の表現型 P_o の中間親の表現型 P_{mid} への回帰係数は

$$b_{o,mid} = \frac{V_A}{V_P} \quad (= h^2) \quad (3.50)$$

となり、これは狭義の遺伝率 h^2 に等しい。すなわち、横軸に中間親の表現型 P_{mid} を、縦軸に子の表現型 P_o をプロットすると、回帰直線の式は

$$P_o - \bar{P} = h^2(P_{mid} - \bar{P}) \quad (3.51)$$

でありその傾きは h^2 である。 h^2 は一般に 1 より小さいので、言い方を変えれば「中間親から期待されるほどの表現型は子の表現型には現れない」ことが分かった。その「縮小係数」こそが狭義の遺伝率 h^2 でなのである (!)



解説 3.19 式 (3.47) を導出する。

表現型は式 (3.1) および式 (3.27) より

$$\begin{aligned} P_p &= G_p + E_p = \bar{G} + A_p + D_p + E_p \\ P_f &= G_f + E_f = \bar{G} + A_f + D_f + E_f \\ P_o &= G_o + E_o = \bar{G} + A_o + D_o + E_o \end{aligned} \quad (3.52)$$

もしくは、さらに詳しく式 (3.21) より

$$\begin{aligned} P_p &= \bar{G} + \alpha(X_{p1} - \bar{X}) + \alpha(X_{p2} - \bar{X}) + \delta(X_{p1} - \bar{X})(X_{p2} - \bar{X}) + E_p \\ P_f &= \bar{G} + \alpha(X_{f1} - \bar{X}) + \alpha(X_{f2} - \bar{X}) + \delta(X_{f1} - \bar{X})(X_{f2} - \bar{X}) + E_f \\ P_o &= \bar{G} + \alpha(X_{o1} - \bar{X}) + \alpha(X_{o2} - \bar{X}) + \delta(X_{o1} - \bar{X})(X_{o2} - \bar{X}) + E_o \end{aligned} \quad (3.53)$$

と書ける。

まず式 (3.46) に現れる分散 $\text{var}[P_p]$ は単に V_P と書くのであった。

次に、共分散 $\text{cov}[P_p, P_o]$ は

$$\begin{aligned} \text{cov}[P_p, P_o] &= \text{cov}[G_p + E_p, G_o + E_o] \\ &= \text{cov}[G_p, G_o] + \text{cov}[G_p, E_o] + \text{cov}[E_p, G_o] + \text{cov}[E_p, E_o] \end{aligned} \quad (3.54)$$

である。ここで異なる個体間の遺伝子型値と環境偏差の共分散ならびに、環境偏差同士の共分散は0であると仮定する。すると

$$\begin{aligned}\text{cov}[P_p, P_o] &= \text{cov}[G_p, G_o] \\ &= \text{cov}[A_p + D_p, A_o + D_o] \\ &= \text{cov}[A_p, A_o] + \text{cov}[A_p, D_o] + \text{cov}[D_p, A_o] + \text{cov}[D_p, D_o]\end{aligned}\quad (3.55)$$

となる。以下、式 (3.55) の各項を詳細に調べよう。

式 (3.55) の第一項は

$$\begin{aligned}\text{cov}[A_p, A_o] &= \text{cov}[\alpha(X_{p1} - \bar{X}) + \alpha(X_{p2} - \bar{X}), \alpha(X_{o1} - \bar{X}) + \alpha(X_{o2} - \bar{X})] \\ &= \sum_{i,j=1}^2 \alpha^2 \text{cov}[X_{pi} - \bar{X}, X_{oj} - \bar{X}] = \sum_{i,j=1}^2 \alpha^2 \overline{(X_{pi} - \bar{X})(X_{oj} - \bar{X})}\end{aligned}\quad (3.56)$$

と計算できる。もし添字 pi と添字 oj に対応する遺伝子が同一祖先由来（これを**同祖的 (identical by descent; IBD)**である、という）であれば、上式で

$$\alpha^2 \overline{(X_{pi} - \bar{X})(X_{oj} - \bar{X})} = \alpha^2 \overline{(X_{pi} - \bar{X})^2} = \alpha^2 \text{var}[X] = \frac{V_A}{2}\quad (3.57)$$

である。ここで最後の等式は式 (3.34a) を用いた。かたや、もし添字 pi と添字 oj に対応する遺伝子が同祖的でなければ、

$$\overline{(X_{pi} - \bar{X})(X_{oj} - \bar{X})} = \underbrace{\overline{(X_{pi} - \bar{X})}}_{=0} \cdot \underbrace{\overline{(X_{oj} - \bar{X})}}_{=0} = 0\quad (3.58)$$

である。今の場合、親子関係 (3.45) より、添字 $p1$ と添字 $o1$ に対応する遺伝子のみ同祖的であるから、

$$\text{cov}[A_p, A_o] = \frac{V_A}{2}\quad (3.59)$$

を得る。

次に式 (3.55) の第二項は

$$\begin{aligned}\text{cov}[A_p, D_o] &= \text{cov}[\alpha(X_{p1} - \bar{X}) + \alpha(X_{p2} - \bar{X}), \delta(X_{o1} - \bar{X})(X_{o2} - \bar{X})] \\ &= \sum_{i=1}^2 \alpha \delta \text{cov}[X_{pi} - \bar{X}, (X_{o1} - \bar{X})(X_{o2} - \bar{X})] = \sum_{i=1}^2 \alpha \delta \overline{(X_{pi} - \bar{X})(X_{o1} - \bar{X})(X_{o2} - \bar{X})}\end{aligned}\quad (3.60)$$

と計算できる。次の二通りの場合を考える。

- (1) 添字 pi に対応する遺伝子と、添字 $o1$ もしくは $o2$ に対応する遺伝子いずれかが同祖的である場合。例えば、添字 pi と添字 $o1$ に対応する遺伝子が同祖的であるとしてよい。すると

$$\overline{(X_{pi} - \bar{X})(X_{o1} - \bar{X})(X_{o2} - \bar{X})} = \overline{(X_{pi} - \bar{X})^2(X_{o2} - \bar{X})} = \overline{(X_{pi} - \bar{X})^2} \cdot \underbrace{\overline{(X_{o2} - \bar{X})}}_{=0} = 0\quad (3.61)$$

- (2) 添字 pi に対応する遺伝子と、添字 $o1$ もしくは $o2$ に対応する遺伝子の間に同祖性がない場合。

$$\overline{(X_{pi} - \bar{X})(X_{o1} - \bar{X})(X_{o2} - \bar{X})} = \underbrace{\overline{(X_{pi} - \bar{X})}}_{=0} \cdot \underbrace{\overline{(X_{o1} - \bar{X})}}_{=0} \cdot \underbrace{\overline{(X_{o2} - \bar{X})}}_{=0} = 0\quad (3.62)$$

となる。

したがっていずれにせよ $\text{cov}[A_p, D_o] = 0$ である。同様の理由で $\text{cov}[D_p, A_o] = 0$ である。

最後に式 (3.55) の第四項を考える。

$$\begin{aligned}\text{cov}[D_p, D_o] &= \text{cov}[\delta(X_{p1} - \bar{X})(X_{p2} - \bar{X}), \delta(X_{o1} - \bar{X})(X_{o2} - \bar{X})] \\ &= \delta^2 \overline{(X_{p1} - \bar{X})(X_{p2} - \bar{X})(X_{o1} - \bar{X})(X_{o2} - \bar{X})}\end{aligned}\quad (3.63)$$

と計算できる。次の三通りの場合を考える。

- (1) 添字 p1, p2 に対応する遺伝子二つと、添字 o1, o2 遺伝子二つの間いずれも同祖性があるとき（つまり「p1 = o1 かつ p2 = o2」、もしくは「p1 = o2 かつ p2 = o1」のとき）。

$$\begin{aligned} \overline{\delta^2(X_{p1} - \bar{X})(X_{p2} - \bar{X})(X_{o1} - \bar{X})(X_{o2} - \bar{X})} &= \overline{\delta^2(X_{p1} - \bar{X})^2(X_{p2} - \bar{X})^2} \\ &= \overline{\delta^2(X_{p1} - \bar{X})^2} \cdot \overline{(X_{p2} - \bar{X})^2} = \delta^2 \text{var}[X]^2 = V_D \end{aligned} \quad (3.64)$$

である。ここで最後の等式は式 (3.34b) を用いた。

- (2) 添字 p1, p2 に対応する遺伝子二つと、添字 o1, o2 遺伝子二つの中の片方にのみ同祖性があるとき（つまり「p1 = o1 かつ p2 ≠ o2」、もしくは「p1 ≠ o2 かつ p2 = o1」のとき）。前者が成り立っているとして一般性を失わず、

$$\begin{aligned} \overline{(X_{p1} - \bar{X})(X_{p2} - \bar{X})(X_{o1} - \bar{X})(X_{o2} - \bar{X})} &= \overline{(X_{p1} - \bar{X})^2(X_{p2} - \bar{X})(X_{o2} - \bar{X})} \\ &= \overline{(X_{p1} - \bar{X})^2} \cdot \underbrace{\overline{(X_{p2} - \bar{X})}}_{=0} \cdot \underbrace{\overline{(X_{o2} - \bar{X})}}_{=0} = 0 \end{aligned} \quad (3.65)$$

である。

- (3) 添字 p1, p2 に対応する遺伝子二つと、添字 o1, o2 遺伝子二つの間のどれにも同祖性がないとき。

$$\overline{(X_{p1} - \bar{X})(X_{p2} - \bar{X})(X_{o1} - \bar{X})(X_{o2} - \bar{X})} = \underbrace{\overline{(X_{p1} - \bar{X})}}_{=0} \cdot \underbrace{\overline{(X_{p2} - \bar{X})}}_{=0} \cdot \underbrace{\overline{(X_{o1} - \bar{X})}}_{=0} \cdot \underbrace{\overline{(X_{o2} - \bar{X})}}_{=0} = 0 \quad (3.66)$$

である。

今の場合、親子関係 (3.45) より上の (2) の場合に該当するから

$$\text{cov}[D_p, D_o] = 0 \quad (3.67)$$

を得る。

以上より、式 (3.46) に現れる共分散 $\text{cov}[P_p, P_o]$ は

$$\text{cov}[P_p, P_o] = \frac{V_A}{2} + 0 + 0 + 0 = \frac{V_A}{2} \quad (3.68)$$

であるので、式 (3.46) に代入して結論を得る。

解説 3.20 次に式 (3.50) を導出する。定義より

$$b_{o,\text{mid}} = \frac{\text{cov}[P_{\text{mid}}, P_o]}{\text{var}[P_{\text{mid}}]} \quad (3.69)$$

である。任意交配の仮定より、父親と母親の表現型に相関はないので、式 (3.69) の分散は

$$\text{var}[P_{\text{mid}}] = \text{var}\left[\frac{P_p + P_m}{2}\right] = \frac{1}{4} \underbrace{\text{var}[P_p]}_{=V_P} + \frac{1}{2} \underbrace{\text{cov}[P_p, P_m]}_{=0} + \frac{1}{4} \underbrace{\text{var}[P_m]}_{=V_P} = \frac{1}{2} V_P \quad (3.70)$$

となる。一方で式 (3.69) の共分散は式 (3.47) および式 (3.48) を用いて

$$\text{cov}[P_{\text{mid}}, P_o] = \text{cov}\left[\frac{P_p + P_m}{2}, P_o\right] = \frac{1}{2} \underbrace{\text{cov}[P_p, P_o]}_{=V_A/2} + \frac{1}{2} \underbrace{\text{cov}[P_m, P_o]}_{=V_A/2} = \frac{1}{2} V_A \quad (3.71)$$

となり、これらを式 (3.69) に戻して結果を得る。

3.5.2 きょうだい間の相関と双生児法

次にきょうだい関係を持つ二個体 u, v 間の類似度について調べる。きょうだいには

- 片親のみ共通する半きょうだい (half sib)
- 両親が共通する全きょうだい (full sib)

が存在し、双生児は全きょうだいであるが、さらに

- 二卵性双生児 (DZ; dizygotic twin)
- 一卵性双生児 (MZ; monozygotic twin)

に分けられる。

きょうだい u, v 間の表現型の共分散 $\text{cov}[P_u, P_v]$ は以下のように計算できる。

$$\text{半きょうだい: } \text{cov}^{(\text{HS})}[P_u, P_v] = \frac{1}{4}V_A + V_{EC}^{(\text{HS})} \quad (3.72a)$$

$$\text{全きょうだい: } \text{cov}^{(\text{FS})}[P_u, P_v] = \frac{1}{2}V_A + \frac{1}{4}V_D + V_{EC}^{(\text{FS})} \quad (3.72b)$$

$$\text{二卵性双生児: } \text{cov}^{(\text{DZ})}[P_u, P_v] = \frac{1}{2}V_A + \frac{1}{4}V_D + V_{EC}^{(\text{DZ})} \quad (3.72c)$$

$$\text{一卵性双生児: } \text{cov}^{(\text{MZ})}[P_u, P_v] = V_A + V_D + V_{EC}^{(\text{MZ})} \quad (3.72d)$$

ここで

- $V_{EC}^{(\text{HS})}$ は、環境偏差のうち半きょうだいが共有する部分の分散
- $V_{EC}^{(\text{FS})}$ は、環境偏差のうち全きょうだいが共有する部分の分散
- $V_{EC}^{(\text{DZ})}$ は、環境偏差のうち二卵性双生児が共有する部分の分散
- $V_{EC}^{(\text{MZ})}$ は、環境偏差のうち一卵性双生児が共有する部分の分散

をそれぞれ表す。

一般に、一卵性双生児が環境（特に親の子育て）から受ける影響と、二卵性双生児が環境（特に親の子育て）から受ける影響の間に差はないと考えられるから、 $V_{EC}^{(\text{DZ})} = V_{EC}^{(\text{MZ})}$ であろう。また、双子ではない全きょうだいは双子のきょうだいよりも環境の共通部分が少ないから $V_{EC}^{(\text{FS})} < V_{EC}^{(\text{DZ})}$ であろう。そして、半きょうだいは全きょうだいに比べて、別の家庭環境で育つなど環境の共通部分が少ないから $V_{EC}^{(\text{HS})} < V_{EC}^{(\text{FS})}$ と考えられる。これらを合わせて

$$V_{EC}^{(\text{HS})} < V_{EC}^{(\text{FS})} < V_{EC}^{(\text{DZ})} = V_{EC}^{(\text{MZ})} \quad (3.73)$$

と考えて良い。

狭義の遺伝率 h^2 (式 (3.39) 参照) を推定する一つの方法は、式において、二卵性双生児の共分散 (3.72c) と一卵性双生児の共分散 (3.72d) を比較することである。式 (3.73) のように $V_{EC}^{(\text{DZ})} = V_{EC}^{(\text{MZ})}$ が仮定できるから、

$$\text{cov}^{(\text{MZ})} - \text{cov}^{(\text{DZ})} = \frac{1}{2}V_A + \frac{3}{4}V_D \quad (3.74)$$

である。もし顕性偏差が無視でき $V_D = 0$ が仮定できるならば、

$$\frac{2(\text{cov}^{(\text{MZ})} - \text{cov}^{(\text{DZ})})}{V_P} = \frac{V_A}{V_P} = h^2 \quad (3.75)$$

となるので、式 (3.75) の左辺の値を狭義の遺伝率 h^2 の推定値とできる。このような方法を**双生児法 (twin method)**と呼ぶ。しかし実際には顕性偏差が無視できないこともあるため、その場合には式 (3.75) の左辺の値は狭義の遺伝率を過大に推定することになる。

問題 3.21 式 (3.72a) の半きょうだいの表現型分散 $\text{cov}^{(\text{HS})}$ を用いて狭義の遺伝率 h^2 を推定しようとする場合、どのようにすれば良いか。

解説 3.22 式 (3.72) を導出する。全ては導出せず式 (3.72b) の導出を例にとる。基本的には**解説 3.19**と同じである。

きょうだい二個体 u, v の表現型を

$$\begin{aligned} P_u &= G_u + E_u = \underbrace{\bar{G} + A_u + D_u}_{=G_u} + \underbrace{EC + EW_u}_{=E_u} \\ P_v &= G_v + E_v = \underbrace{\bar{G} + A_v + D_v}_{=G_v} + \underbrace{EC + EW_v}_{=E_v} \end{aligned} \quad (3.76)$$

と書く。ただし EC は二個体に共通の環境（例えば家庭環境）、 EW は二個体に共通でない環境とする（ EC や EW はかけ算を表しているのではなくそれぞれ記号である、ここで C は common の、 W は within の、それぞれ頭文字である）。

表現型共分散 $\text{cov}[P_u, P_v]$ は

$$\begin{aligned} \text{cov}[P_u, P_v] &= \text{cov}[G_u + E_u, G_v + E_v] \\ &= \text{cov}[G_u, G_v] + \text{cov}[G_u, E_v] + \text{cov}[E_u, G_v] + \text{cov}[E_u, E_v] \end{aligned} \quad (3.77)$$

と計算でき、 $\text{cov}[G_u, E_v]$ および $\text{cov}[E_u, G_v]$ は遺伝子型値と環境偏差の共分散であるから、それらはともに 0 であると仮定する。一方で $\text{cov}[E_u, E_v]$ はさらに

$$\begin{aligned} \text{cov}[E_u, E_v] &= \text{cov}[EC + EW_u, EC + EW_v] \\ &= \underbrace{\text{cov}[EC, EC]}_{=V_{EC}} + \text{cov}[EC, EW_v] + \text{cov}[EW_u, EC] + \text{cov}[EW_u, EW_v] \end{aligned} \quad (3.78)$$

と計算でき、後半の三項は 0 であると仮定する。以上より $\text{cov}[P_u, P_v] = \text{cov}[G_u, G_v] + V_{EC}$ まで分かった。

残りの $\text{cov}[G_u, G_v]$ の計算は、**解説 3.19** にある $\text{cov}[G_p, G_o]$ の計算の方法と同じである。

(i) $\text{cov}[G_u, G_v]$ に相加遺伝分散 V_A 成分がどれだけ含まれるかは、式 (3.57) および式 (3.58) の議論を繰り返せば良い。すなわち個体 u と v が同祖的な遺伝子を持つごとに、分散 $V_A/2$ が生じる。

今、 u と v は全きょうだいであるので、 u, v の持つ父方由来の遺伝子が同祖的になる確率は、父の持つ二つの相同遺伝子のうち同じものを両者が受け継ぐ確率であり、それは $1/2$ である。母方遺伝子についても同様である。したがって生じる相加遺伝分散成分の総量は

$$\frac{1}{2} \cdot \frac{V_A}{2} + \frac{1}{2} \cdot \frac{V_A}{2} = \frac{1}{2} V_A \quad (3.79)$$

となる。

(ii) $\text{cov}[G_u, G_v]$ に顕性分散 V_D 成分がどれだけ含まれるかは、式 (3.64)、式 (3.65) および式 (3.66) の議論を繰り返せば良い。すなわち個体 u と v のそれぞれが持つ相同遺伝子二つがともにきょうだい間で同祖的であれば、分散 V_D が生じる。

今、 u と v は全きょうだいであるので、 u, v の持つ父方由来の遺伝子が同祖的であり、かつ母方由来の遺伝子も同祖的である確率は、 $(1/2) \cdot (1/2) = 1/4$ である。したがって生じる顕性分散成分の総量は

$$\frac{1}{4} V_D \quad (3.80)$$

となる。

以上の議論より、遺伝子型値の共分散は

$$\text{cov}[G_u, G_v] = \frac{1}{2} V_A + \frac{1}{4} V_D \quad (3.81)$$

であり、上記環境偏差の共分散と合わせて式 (3.72b) を得る。

解説 3.23 きょうだい間の遺伝共分散を $\text{cov}[G_u, G_v]$ を直接計算することもできる（が、たいへん煩雑である）。

たとえば全きょうだいを考えよう。そのためには以下の mating table を用いる。

父 \ 母	$A_1A_1 (p^2)$	$A_1A_2 (2pq)$	$A_2A_2 (q^2)$
$A_1A_1 (p^2)$	子 $A_1A_1 : 1$ 子 $A_1A_2 : 0$ 子 $A_2A_2 : 0$	子 $A_1A_1 : 1/2$ 子 $A_1A_2 : 1/2$ 子 $A_2A_2 : 0$	子 $A_1A_1 : 0$ 子 $A_1A_2 : 1$ 子 $A_2A_2 : 0$
$A_1A_2 (2pq)$	子 $A_1A_1 : 1/2$ 子 $A_1A_2 : 1/2$ 子 $A_2A_2 : 0$	子 $A_1A_1 : 1/4$ 子 $A_1A_2 : 1/2$ 子 $A_2A_2 : 1/4$	子 $A_1A_1 : 0$ 子 $A_1A_2 : 1/2$ 子 $A_2A_2 : 1/2$
$A_2A_2 (q^2)$	子 $A_1A_1 : 0$ 子 $A_1A_2 : 1$ 子 $A_2A_2 : 0$	子 $A_1A_1 : 0$ 子 $A_1A_2 : 1/2$ 子 $A_2A_2 : 1/2$	子 $A_1A_1 : 0$ 子 $A_1A_2 : 0$ 子 $A_2A_2 : 1$

きょうだいともに遺伝子型が A_1A_1 であり、したがって二人共遺伝子型値 G_{11} を持つ確率は、上の mating table より

$$\text{Prob}[(G_{11}, G_{11})] = \underbrace{p^4 \cdot 1^2 + 4p^2q^2 \cdot \left(\frac{1}{4}\right)^2 + q^4 \cdot 0^2}_{=\text{対角線}} + 2 \underbrace{\left\{ 2p^3q \cdot \left(\frac{1}{2}\right)^2 + p^2q^2 \cdot 0^2 + 2pq^3 \cdot 0^2 \right\}}_{=\text{非対角線}} \quad (3.82)$$

である。また、きょうだいの遺伝子型の組み合わせが A_1A_1 と A_1A_2 であり、したがって二人の遺伝子型値の組み合わせが G_{11} と G_{12} である確率は

$$\text{Prob}[(G_{11}, G_{12})] = 2 \left[\underbrace{p^4 \cdot (1 \cdot 0) + 4p^2q^2 \cdot \left(\frac{1}{4} \cdot \frac{1}{2}\right) + q^4 \cdot (0 \cdot 0)}_{=\text{対角線}} + 2 \underbrace{\left\{ 2p^3q \cdot \left(\frac{1}{2} \cdot \frac{1}{2}\right)^2 + p^2q^2 \cdot (0 \cdot 1) + 2pq^3 \cdot \left(0 \cdot \frac{1}{2}\right) \right\}}_{=\text{非対角線}} \right] \quad (3.83)$$

である。ここで先頭の 2 はきょうだいの並べ替えに対応する。

このように計算すると、最終的に

$$\begin{aligned} \text{Prob}[(G_{11}, G_{11})] &= p^4 + p^3q + \frac{1}{4}p^2q^2 \\ \text{Prob}[(G_{11}, G_{12})] &= + 2p^3q + p^2q^2 \\ \text{Prob}[(G_{11}, G_{22})] &= + \frac{1}{2}p^2q^2 \\ \text{Prob}[(G_{12}, G_{12})] &= + p^3q + 3p^2q^2 + pq^3 \\ \text{Prob}[(G_{12}, G_{22})] &= + p^2q^2 + 2pq^3 \\ \text{Prob}[(G_{22}, G_{22})] &= + \frac{1}{4}p^2q^2 + pq^3 + q^4 \end{aligned} \quad (3.84)$$

を得る。これらの確率を用いて $\text{cov}[G_u, G_v]$ を定義から直接計算をすると

$$\begin{aligned} \text{cov}[G_u, G_v] &= \frac{1}{2} \cdot 2 \underbrace{\{p(G_{11} - G_{12}) + q(G_{12} - G_{22})\}^2 pq}_{=\alpha} + \frac{1}{4} \underbrace{(G_{11} - 2G_{12} + G_{22})^2 p^2 q^2}_{=\delta} \\ &= \frac{1}{2} \cdot \underbrace{2\alpha^2 pq}_{=V_A \text{ (式 (3.34a) を参照)}} + \frac{1}{4} \cdot \underbrace{\delta^2 p^2 q^2}_{=V_D \text{ (式 (3.34b) を参照)}} \end{aligned} \quad (3.85)$$

となるので、式 (3.81) が確かめられる。

例 3.24 例 3.18 の数値例を引き続き解析しよう。遺伝子型値が $(G_{11}, G_{12}, G_{22}) = (36, 11, 11)$ で与えられ、遺伝子 A_1, A_2 の頻度が $(p, q) = (0.6, 0.4)$ で与えられる場合である。また、 $V_A = 108, V_D = 36, V_E = 144, V_P = 288$ であった。

例えば、全きょうだいである二個体 u, v の表現型 P_u, P_v の相関係数を求めてみよう（相関係数の定義は式 (0.23) を参照）。

例えば式 (3.72b) において、共通環境に由来する分散 $V_{EC}^{(FS)}$ は全環境分散 V_E の半分であると仮定しよう。したがって $V_{EC}^{(FS)} = V_E/2 = 72$ となる。

この時、式 (3.72b) より P_u と P_v の共分散は

$$\text{cov}^{(FS)} [P_u, P_v] = \frac{1}{2} \cdot 108 + \frac{1}{4} \cdot 36 + 72 = 135 \quad (3.86)$$

となる。したがって $V_{P_u} = V_{P_v} = V_P$ であることに注意すると、求める相関係数は

$$\rho^{(FS)} = \frac{135}{\sqrt{288}\sqrt{288}} = \frac{15}{32} \quad (3.87)$$

となる。

3.6 選択差と選択応答

今まで議論してきた集団に、表現型 P に応じて自然選択（もしくは人為選択）がかかる状況を考えよう。

具体的に、個体の適応度 W はその表現型 P から

$$W = 1 + s(P - \bar{P}) \quad (3.88)$$

で決定すると仮定する。以下で $|s|$ は十分小さいとする。パラメータ s は、表現型が 1 増加した時の適応度の増分であり、**選択勾配 (selection gradient)** と呼ばれる。

もし $s > 0$ ならば、表現型が集団平均 \bar{P} より大きい個体が有利である。反対に $s < 0$ ならば、表現型が集団平均 \bar{P} より小さい個体が有利である。

ここで次の量を考える。

- 親世代から適応度に応じた重みでランダムにサンプルした個体は、繁殖を行った平均的な個体を表す。この「平均繁殖個体」の表現型を \bar{P}_{parent} で表す。
- 親世代の任意交配の後の、子世代の表現型の平均値を \bar{P}' で表す。

この時、差分 $\Delta S \equiv \bar{P}_{\text{parent}} - \bar{P}$ は、「平均繁殖個体」と「親世代の平均値」の差であり、表現型にかけようとしている選択の大きさを表す。この量を**選択差 (selection differential)**（もしくは**選抜差**）と呼ぶ。

一方で、差分 $\Delta P \equiv \bar{P}' - \bar{P}$ は実際に表現型にかかった選択の大きさを表す。この量を**選択応答 (response to selection)**と呼ぶ。

選択差と選択応答はどのような関係にあるだろうか？

解説 3.25 選択差と選択応答を計算する。

まず選択差を計算する。定義より

$$\bar{P}_{\text{parent}} = \frac{\overline{WP}}{\bar{W}} \quad (3.89)$$

である。式 (3.89) の分母は、式 (3.88) の両辺の集団平均をとることで $\bar{W} = 1$ となる。一方、式 (3.89) の分子は

$$\overline{WP} = \overline{\{1 + s(P - \bar{P})\}P} = \bar{P} + s(\overline{P^2} - \bar{P}^2) = \bar{P} + sV_P \quad (3.90)$$

である。したがって選択差は

$$\Delta S = \bar{P}_{\text{parent}} - \bar{P} = (\bar{P} + sV_P) - \bar{P} = sV_P \quad (3.91)$$

となる。

次に選択応答を計算する。現世代および次世代の対立遺伝子 A_1, A_2 の遺伝子頻度をそれぞれ p, q ($p + q = 1$) および p', q' ($p' + q' = 1$) とすると、HW 平衡の下で

$$\begin{aligned}\bar{P} &= p^2\bar{P}_{11} + 2pq\bar{P}_{12} + q^2\bar{P}_{22} \\ \bar{P}' &= p'^2\bar{P}_{11} + 2p'q'\bar{P}_{12} + q'^2\bar{P}_{22}\end{aligned}\quad (3.92)$$

である。ここで \bar{P}_{ij} は遺伝子型 A_iA_j を持つ個体だけ集団内で集め、その表現型の平均を取った値である。しかしながらそれらの個体 indv は $P_{\text{indv}} = G_{ij} + E_{\text{indv}}$ を満たすから (つまり、 G_{ij} は indv に依らない)、両辺の平均を取って

$$\bar{P}_{ij} = G_{ij} + \bar{E}_{ij}\quad (3.93)$$

である。ここでここで \bar{E}_{ij} は遺伝子型 A_iA_j を持つ個体だけ集団内で集め、その環境偏差の平均を取った値であり、 \bar{E} の時と同じように 0 であると仮定する。したがって

$$\bar{P}_{ij} = G_{ij}\quad (3.94)$$

であり、式 (3.92) は

$$\begin{aligned}\bar{P} &= p^2G_{11} + 2pqG_{12} + q^2G_{22} \\ \bar{P}' &= p'^2G_{11} + 2p'q'G_{12} + q'^2G_{22}\end{aligned}\quad (3.95)$$

と書ける。

いま、 p' と p の差が小さく、 $p' = p + \epsilon$ と書けているとする ($|\epsilon|$ は微小量)。この時、 ϵ^2 以上の微小量を全て無視することで、

$$\begin{aligned}p'^2 &= (p + \epsilon)^2 \approx p^2 + 2p\epsilon \\ 2p'q' &= 2(p + \epsilon)(q - \epsilon) \approx 2pq + 2(q - p)\epsilon \\ q'^2 &= (q - \epsilon)^2 \approx q^2 - 2q\epsilon\end{aligned}\quad (3.96)$$

と書けるので、これらを式 (3.95) に代入することで、選択応答は

$$\begin{aligned}\Delta P &= (p'^2 - p^2)G_{11} + (2p'q' - 2pq)G_{12} + (q'^2 - q^2)G_{22} \\ &\approx \{2pG_{11} + 2(q - p)G_{12} - 2qG_{22}\}\epsilon = 2 \underbrace{\{p(G_{11} - G_{12}) + q(G_{12} - G_{22})\}}_{=\alpha \text{ (式 (3.12) を参照)}}\epsilon = 2\alpha\epsilon\end{aligned}\quad (3.97)$$

となる。ここで α は遺伝子置換の平均効果である。

では ϵ はどのような値であろうか。遺伝子型 A_iA_j の平均適応度 w_{ij} は式 (3.88) を各個体 indv に適用した式 $W_{\text{indv}} = 1 + s(P_{\text{indv}} - \bar{P})$ の、遺伝子型 A_iA_j を持つ個体に渡る平均をとって、

$$w_{ij} = 1 + s(G_{ij} - \bar{G})\quad (3.98)$$

であるから (注：再び先ほどの議論、つまり $\bar{P}_{ij} = G_{ij}$ を使った)、式 (1.62) が使えて

$$p' = p + spq \underbrace{\{(G_{11} - G_{12})p + (G_{12} - G_{22})q\}}_{=\alpha}\quad (3.99)$$

である (注：式 (1.62) において $s_{ij} = s(G_{ij} - \bar{G})$ とすれば良い)。

よって、式 (3.97) に $\epsilon = spq\alpha$ を代入すればよいので、選択応答は

$$\Delta P = 2\alpha \cdot spq\alpha = s \cdot \underbrace{2\alpha^2 pq}_{=V_A \text{ (式 (3.34a) を参照)}} = sV_A\quad (3.100)$$

となる。

計算の結果を以下にまとめる。

$$\text{選択差： } \Delta S = sV_P \quad (3.101)$$

$$\text{選択応答： } \Delta P = sV_A \quad (3.102)$$

したがって、選択差と選択応答の関係は、狭義の遺伝率 $h^2 = V_A/V_P$ を用いて

$$\Delta P = h^2 \Delta S \quad (3.103)$$

と表すことができる。式 (3.103) を **育種家方程式 (breeder's equation)** と呼ぶ。

解説 3.26 二倍体にかかる自然選択に興味がある場合、選択応答を表す式 (3.102) と、式 (3.103) の育種家方程式が最も基本的かつ重要な式となる。遺伝子型値 G をわざわざ相加遺伝値 A と顕性偏差 D に分解したのは、これらの式を導出するためと言っても過言ではない。

解説 3.27 育種家方程式 (3.103) で選択応答 ΔP が選択差 ΔS の h^2 倍になる理由は、3.5.1 節の式 (3.50) で示したように、子の表現型の中間親の表現型に対する回帰係数が h^2 であるからにほかならない。つまり直観的な言葉で言えば「中間親は子に対し、自分の表現型のうち割合 h^2 の部分しか伝えられないから」である。

3.7 多座位の場合

QTL が $L > 1$ 座位ある場合は、 $L = 1$ の時の議論を次のように拡張する。以降、座位間での連鎖平衡を仮定する。座位 ℓ に関する量には上添字 $[\ell]$ をつけることと約束する。

分かりやすいように、 $L = 2$ の時を説明する。非相加的モデル (3.21) を 2 座位に拡張して

$$\begin{aligned} G - \bar{G} &= \alpha^{[1]}(X_{1+2}^{[1]} - \overline{X_{1+2}^{[1]}}) + \delta^{[1]}(X_1^{[1]} - \overline{X_1^{[1]}})(X_2^{[1]} - \overline{X_2^{[1]}}) \\ &+ \alpha^{[2]}(X_{1+2}^{[2]} - \overline{X_{1+2}^{[2]}}) + \delta^{[2]}(X_1^{[2]} - \overline{X_1^{[2]}})(X_2^{[2]} - \overline{X_2^{[2]}}) + \epsilon \end{aligned} \quad (3.104)$$

とすれば、この誤差 ϵ に関する誤差の二乗平均 J (式 (3.9)) を最小にする $\alpha^{[1]}, \alpha^{[2]}, \delta^{[1]}, \delta^{[2]}$ が求まる。これらを用いて 各座位における相加遺伝値 $A^{[\ell]}$ および顕性偏差 $D^{[\ell]}$ を

$$\begin{aligned} A^{[\ell]} &= \alpha^{[\ell]}(X_{1+2}^{[\ell]} - \overline{X_{1+2}^{[\ell]}}) \\ D^{[\ell]} &= \delta^{[\ell]}(X_1^{[\ell]} - \overline{X_1^{[\ell]}})(X_2^{[\ell]} - \overline{X_2^{[\ell]}}) \end{aligned} \quad (3.105)$$

で定め、また (合計の)相加遺伝値 A および顕性偏差 D を

$$\begin{aligned} A &= \sum_{\ell=1}^L A^{[\ell]} \\ D &= \sum_{\ell=1}^L D^{[\ell]} \end{aligned} \quad (3.106)$$

で定めることと約束する。すなわち式 (3.104) は

$$G - \bar{G} = A^{[1]} + D^{[1]} + A^{[2]} + D^{[2]} + \epsilon = A + D + \epsilon \quad (3.107)$$

と書けたことになる。

しかしながら、座位間にエピスタシスがある（つまり、表現型に対する各座位の効果が相加的でない）場合、誤差の二乗平均の最小値 J_{\min} は 0 にならない。したがって、モデル (3.104) に投入した変数の座位間での交互作用も投入する必要があり、式の

上ではそれらは積の形で表される。数式で書けばフルモデルは

$$\begin{aligned}
G - \bar{G} = & \alpha^{[1]}(X_{1+2}^{[1]} - \bar{X}_{1+2}^{[1]}) + \delta^{[1]}(X_1^{[1]} - \bar{X}_1^{[1]})(X_2^{[1]} - \bar{X}_2^{[1]}) \\
& + \alpha^{[2]}(X_{1+2}^{[2]} - \bar{X}_{1+2}^{[2]}) + \delta^{[2]}(X_1^{[2]} - \bar{X}_1^{[2]})(X_2^{[2]} - \bar{X}_2^{[2]}) \\
& + \iota_{AA} \left\{ (X_{1+2}^{[1]} - \bar{X}_{1+2}^{[1]}) \times (X_{1+2}^{[2]} - \bar{X}_{1+2}^{[2]}) \right\} \\
& + \iota_{AD} \left\{ (X_{1+2}^{[1]} - \bar{X}_{1+2}^{[1]}) \times (X_1^{[2]} - \bar{X}_1^{[2]})(X_2^{[2]} - \bar{X}_2^{[2]}) \right\} \\
& + \iota_{DA} \left\{ (X_1^{[1]} - \bar{X}_1^{[1]})(X_2^{[1]} - \bar{X}_2^{[1]}) \times (X_{1+2}^{[2]} - \bar{X}_{1+2}^{[2]}) \right\} \\
& + \iota_{DD} \left\{ (X_1^{[1]} - \bar{X}_1^{[1]})(X_2^{[1]} - \bar{X}_2^{[1]}) \times (X_1^{[2]} - \bar{X}_1^{[2]})(X_2^{[2]} - \bar{X}_2^{[2]}) \right\} \\
& + \epsilon
\end{aligned} \tag{3.108}$$

となる (ι はギリシャ文字で「イオタ」と読む)。このモデルの誤差の最小二乗値は $J_{\min} = 0$ となる。そこでここで求めた推定値 $\iota_{AA}, \iota_{AD}, \iota_{DA}, \iota_{DD}$ を用いて

- ι_{AA} を伴う項を $AA^{[12]}$ とおく (AA は一つの記号とみなす)。これは座位 1 の相加遺伝値と座位 2 の相加遺伝値の交互作用である。
- ι_{AD} を伴う項を $AD^{[12]}$ とおく (AD は一つの記号とみなす)。これは座位 1 の相加遺伝値と座位 2 の顕性偏差の交互作用である。
- ι_{DA} を伴う項を $DA^{[12]}$ とおく (DA は一つの記号とみなす)。これは座位 1 の顕性偏差と座位 2 の相加遺伝値の交互作用である。
- ι_{DD} を伴う項を $DD^{[12]}$ とおく (DD は一つの記号とみなす)。これは座位 1 の顕性偏差と座位 2 の顕性偏差の交互作用である。

とおくと、式 (3.108) は最終的に

$$\begin{aligned}
G - \bar{G} = & \underbrace{A^{[1]} + D^{[1]}}_{\text{座位 1 の主効果}} + \underbrace{A^{[2]} + D^{[2]}}_{\text{座位 2 の主効果}} + \underbrace{AA^{[12]} + AD^{[12]} + DA^{[12]} + DD^{[12]}}_{\text{座位 1 と 2 の交互作用}} \\
= & \underbrace{A^{[1]} + A^{[2]}}_{=A \text{ (相加遺伝値)}} + \underbrace{D^{[1]} + D^{[2]}}_{=D \text{ (顕性偏差)}} + \underbrace{AA^{[12]} + AD^{[12]} + DA^{[12]} + DD^{[12]}}_{=I \text{ (相互作用偏差)}}
\end{aligned} \tag{3.109}$$

と書ける。式 (3.109) にあるように、座位間のエピスタシスから定まる、相加遺伝値 A および顕性偏差 D に含まれない項の和 I を、**相互作用偏差 (interaction deviation)** もしくは**エピスタシス偏差 (epistatic deviation)** と呼ぶ。

解説 3.28 式 (3.109) の右辺には合計 8 項が含まれているが、これは必然である。なぜならば式 (3.109) で左辺の \bar{G} を右辺に移項すると、この式は遺伝子型値 G を $8 + 1 = 9$ 項で説明しようとすることになるが、座位 1 にも 2 にも異なる遺伝子型は 3 通りあるので、座位 1, 2 を合わせた遺伝子型は合計 $3^2 = 9$ 通りあるから、ちょうど項の数が釣り合うからである。

座位が 3 つある場合、同様の議論によって $G - \bar{G}$ を説明する項は全部で $3^3 - 1 = 26$ 項あると考えられるが、実際並べてみると

$$\begin{aligned}
G - \bar{G} = & A^{[1]} + D^{[1]} + A^{[2]} + D^{[2]} + A^{[3]} + D^{[3]} \\
& + AA^{[12]} + AD^{[12]} + DA^{[12]} + DD^{[12]} \\
& + AA^{[13]} + AD^{[13]} + DA^{[13]} + DD^{[13]} \\
& + AA^{[23]} + AD^{[23]} + DA^{[23]} + DD^{[23]} \\
& + AAA^{[123]} + AAD^{[123]} + ADA^{[123]} + ADD^{[123]} \\
& + DAA^{[123]} + DAD^{[123]} + DDA^{[123]} + DDD^{[123]}
\end{aligned} \tag{3.110}$$

であり、右辺には確かに 26 項が出現する。このうち 3 項の和 $A = A^{[1]} + A^{[2]} + A^{[3]}$ が相加遺伝値、3 項の和 $D = D^{[1]} + D^{[2]} + D^{[3]}$ が顕性偏差、そして残り 20 項全ての和が相互作用偏差 I である。

よって表現型値 P は最終的に以下のような分解を受ける。

$$P = \underbrace{\bar{G} + A + D + I}_{=G} + E \tag{3.111}$$

この時、表現型分散の分解は式 (3.33) に代えて

$$V_P = \underbrace{V_A + V_D + V_I}_{V_G} + V_E \quad (3.112)$$

となることが知られている。 V_I を**相互作用分散 (interaction variance)** もしくは**エピスタシス分散 (epistatic variation)** と呼ぶ。相互作用偏差、相互作用分散の存在下でも、前節までの結果は基本的には変わらない。

4 連続形質の進化モデル

(本章の概要と目的) 本章は連続値で表される形質にはたらく自然選択がどのように記述されるかを学ぶ。2章で用いた表現型 *gambit* を適用し、基本的には一倍体無性生殖の生物の無限集団に焦点を当て議論を進める。基本的になるのは1970年に発表されたPrice方程式であり、これを元にして様々な現代流の理論(例: adaptive dynamics 理論)も発展した。本章で扱う内容も集団遺伝学で扱われることは少なく、どちらかという進化生物学の内容である。

Priceの慧眼は、連続形質の進化の方向を適応度と形質値の共分散として捉えた点にある。現代ではこれは当然の原理として理解されているが、共分散を用いた定式化はその後の血縁淘汰理論の発展にも大きく貢献した。

2章と異なり、表現型は離散的ではなく連続的に分布しているため、集団状態の完全な記述には表現型の頻度分布 *p* の発展方程式が必要になるが、これは特殊な場合を除いて計算することが難しい。そのため、表現型平均や表現型分散などに限ってそれらの時間発展を記述する方法論が発展した。

本書ではPrice方程式の様々な形に対して、その構成法と含意を理解し、単純な例に関して進化の方向を自らの力で導出できるようにすることを目標とする。

4.1 Price 方程式

連続値をとる形質の進化を一般的に分析しよう。以下の仮定を置く。

- (1) 集団の各個体はある属性 *z* を備えているとする(例: 身長)。
- (2) 各個体には適応度 *w* が定まっており、適応度の数だけ次世代に子を残すとする。
- (3) 生まれた子は親の属性 *z* から Δz だけ変化した属性 $z + \Delta z$ を持つとする。

解説 4.1 Δz は各親のラベル *i* に依存して良い。すなわち、親の属性が z_i であり、その子の属性が $z_i + (\Delta z)_i$ であることを言っている。

このとき、現世代と次世代の間で、属性の集団平均 \bar{z} はどのように変化するであろうか? 以下の記号を準備する。

- \bar{z} で、注目している属性の現世代における集団平均を表す。
- \bar{z}' で、注目している属性の次世代における集団平均を表す。
- $\Delta \bar{z}$ で、その差 $\bar{z}' - \bar{z}$ を表す。

現世代の個体のラベルを *i* とすると、

$$\bar{z}' = \frac{\sum_i w_i \{z_i + (\Delta z)_i\}}{\sum_i w_i} = \frac{\overline{wz} + \overline{w\Delta z}}{\bar{w}} \quad (4.1)$$

であるから、

$$\Delta \bar{z} = \bar{z}' - \bar{z} = \frac{\overline{wz} + \overline{w\Delta z}}{\bar{w}} - \bar{z} = \frac{\overbrace{\overline{wz} - \bar{w} \cdot \bar{z}}^{(\text{分子}) = \text{cov}[w, z]}}{\bar{w}} + \frac{\overline{w\Delta z}}{\bar{w}} \quad (4.2)$$

より、整理して

$$\bar{w} \Delta \bar{z} = \text{cov}[w, z] + \overline{w\Delta z} \quad (4.3)$$

を得る。式(4.3)を**Price方程式 (Price equation)**と呼ぶ(Price, 1970)。常に $\Delta z = 0$ である時には、Price方程式はより単純に

$$\bar{w} \Delta \bar{z} = \text{cov}[w, z] \quad (4.4)$$

となる。

例 4.2 適応度が属性 *z* に線形に依存し、例えば

$$w = 1 + s(z - \bar{z}) \quad (4.5)$$

と書ける場合を考えよう。ここで *s* は選択勾配である。

Price方程式(4.4)において

$$\bar{w} = \overline{1 + s(z - \bar{z})} = 1 + s(\bar{z} - \bar{z}) = 1 \quad (4.6)$$

である。また、共分散は

$$\text{cov}[w, z] = \text{cov}[1 + s(z - \bar{z}), z] = \underbrace{\text{cov}[1, z]}_{=0} + s(\underbrace{\text{cov}[z, z]}_{=\text{var}[z]} - \underbrace{\text{cov}[\bar{z}, z]}_{=0}) = s \cdot \text{var}[z] \quad (4.7)$$

であるので、これらを Price 方程式 (4.4) に代入して

$$\Delta \bar{z} = s \cdot \text{var}[z] \quad (4.8)$$

を得る。

式 (4.8) は、量的遺伝学における式 (3.101)、(3.102)、(3.103) において $V_P = \text{var}[z]$ 、 $h^2 = 1$ とおいた特別な場合に対応している。

例 4.3 属性 z として、適応度 w 自体をとってみよう。すなわち $z = w$ とする。この時 Price 方程式 (4.4) は

$$\bar{w} \Delta \bar{w} = \text{cov}[w, w] = \text{var}[w] \quad (\geq 0) \quad (4.9)$$

となる。これを **Fisher の自然選択の基本定理 (Fisher's fundamental theorem of natural selection)** と呼ぶ。基本定理 (4.9) より集団平均適応度 \bar{w} は自然選択によって常に増加（厳密には、非減少）することが分かる。

しかしながら、基本定理 (4.9) を導出するにあたっては、子も親と同じ適応度を受け継ぐという強い仮定がなされていることに注意しよう。自然選択によって集団の構成が変わり、集団の構成の変化によって各個体の適応度が変わってしまうようなゲーム理論的な状況ではこの仮定は当てはまらないので、そのような場合では集団平均適応度 \bar{w} は単調に増加するとは限らない。

解説 4.4 本章では表現型 *gambit* を用い一倍体無性生殖生物を考えているので基本定理は式 (4.9) の形をしているが、もし表現型 *gambit* を放棄し、二倍体有性生殖生物を考えた場合には、前章の知見（量的遺伝学）を用いると、 $\text{var}[w]$ は「適応度の相加遺伝分散」に置き換えられる。

例 4.5 以下では、 i を個体の index とし、 j を戦略の index とする。

各個体は s_1 から s_n のうちのどれか一つの戦略を持つとし、戦略 s_j を持つ個体の適応度を w_j とおこう。この状況において個体の属性 $z^{(j)}$ を次のように定める。

$$z_i^{(j)} = \begin{cases} 1 & \text{個体 } i \text{ が戦略 } s_j \text{ を持つ場合} \\ 0 & \text{個体 } i \text{ が戦略 } s_j \text{ を持たない場合} \end{cases} \quad (4.10)$$

属性 $z^{(j)}$ に関する Price 方程式 (4.4) は

$$\bar{w} \Delta \bar{z}^{(j)} = \text{cov}[w, z^{(j)}] \quad (4.11)$$

となる。ところで $\bar{z}^{(j)}$ の意味を考えると、属性 $z^{(j)}$ の集団平均とは、戦略 s_j を持つ個体の集団中の頻度 x_j にほかならないから、

$$\Delta \bar{z}^{(j)} = \Delta x_j = x'_j - x_j \quad (4.12)$$

である。一方で

$$\text{cov}[w, z^{(j)}] = \overline{wz^{(j)}} - \bar{w} \cdot \bar{z}^{(j)} \quad (4.13)$$

である。式 (4.13) 内の項 $\overline{wz^{(j)}}$ について考えると $wz^{(j)}$ の値は、戦略 s_j を持つ個体（頻度 x_j ）の場合は $wz^{(j)} = w_j \cdot 1 = w_j$ であり、戦略 s_j を持たない個体の場合は $wz^{(j)} = 0$ であるので、

$$\overline{wz^{(j)}} = w_j x_j \quad (4.14)$$

と計算できる。これを式 (4.13) に代入して

$$\text{cov}[w, z^{(j)}] = w_j x_j - \bar{w} x_j = x_j (w_j - \bar{w}) \quad (4.15)$$

である。式 (4.12)、(4.15) を元の Price 方程式 (4.11) に代入すると、

$$\bar{w}(x'_j - x_j) = x_j(w_j - \bar{w}) \iff x'_j = x_j \cdot \frac{w_j}{\bar{w}} \quad (4.16)$$

を得る。式 (4.16) は離散時間版のレプリケーター方程式である (式 (2.6) の連続時間版のレプリケーター方程式と比較せよ)。

解説 4.6 離散時間版レプリケーター方程式 (4.16) から連続時間版レプリケーター方程式 (2.6) を導出するには次のようにすれば良い。

離散世代で 1 世代進んだことを、連続時間で Δt 時間進んだことに対応させる。また離散世代モデルで適応度が w_j であることを、連続時間モデルでは出生率が w_j であることに対応させる。さらに、離散世代モデルで次の世代には親が必ず死んでいることを、連続時間モデルでは死亡率が $1 \cdot \Delta t = \Delta t$ であることに対応させる。

以上より、式 (4.16) において

$$\begin{aligned} x_j &\longrightarrow x_j(t) \\ x'_j &\longrightarrow x_j(t + \Delta t) \\ w_j &\longrightarrow 1 + \underbrace{(w_j - 1) \Delta t}_{=f_j} \\ \bar{w} &\longrightarrow 1 + \underbrace{(\bar{w} - 1) \Delta t}_{=\bar{f}} \end{aligned} \quad (4.17)$$

なる置換を行えば良い。置換後の式は

$$\begin{aligned} (1 + \bar{f} \Delta t) \{x_j(t + \Delta t) - x_j(t)\} &= x_j(t)(f_j - \bar{f}) \Delta t \\ \iff (1 + \bar{f} \Delta t) \frac{x_j(t + \Delta t) - x_j(t)}{\Delta t} &= x_j(t)(f_j - \bar{f}) \end{aligned} \quad (4.18)$$

なので、式 (4.18) において $\Delta t \rightarrow 0$ の極限を取ること、連続時間版レプリケーター方程式 (2.6) が得られる。

例 4.7 利他行動の進化を議論しよう。個体の属性 $z (\geq 0)$ とは「利他度」を表し、利他度が z の個体は

- 利他度に比例したコスト cz を支払って、
- 自分のパートナーに、利他度に比例した利益 bz を与える。

と仮定する。パートナーの「利他度」を z_p と書き、パートナーの自分に対する行動も同様に定められているとする。

このとき、注目する個体の適応度 w は

$$w = 1 - cz + bz_p \quad (4.19)$$

と書ける (b や c は 1 に比べて十分小さく、したがっていつも $w > 0$ が満たされていると仮定する)。式 (4.19) を Price 方程式 (4.4) に代入して

$$\bar{w} \Delta \bar{z} = \text{cov}[1 - cz + bz_p, z] = -c \text{var}[z] + b \text{cov}[z, z_p] \quad (4.20)$$

を得る。したがって集団平均利他度が増加する条件は、 $\Delta \bar{z} > 0$ 、すなわち

$$-c + b \cdot \frac{\text{cov}[z, z_p]}{\text{var}[z]} > 0 \quad (4.21)$$

である。ここで共分散と分散の比 $\text{cov}[z, z_p] / \text{var}[z]$ は、パートナーの利他度 z_p の、注目する個体の利他度 z への回帰係数であることに注意せよ。すなわち比 $\text{cov}[z, z_p] / \text{var}[z]$ は個体とそのパートナーの利他度の類似度を表しており、この類似度が高いほど利他行動の進化はしやすくなる。

式 (4.21) は、回帰係数を血縁度だと思えば Hamilton 則に似ている (が、これは Hamilton 則とは呼ばれない)。

例 4.8 例 4.7 を元に、引き続き利他行動の進化を議論しよう。今度は個体の属性 g は、利他的個体かそうでないかを表す 2 値変数とし、

- 利他的個体ならば $g = 1$
- 非利他的個体ならば $g = 0$

を表すことにする。これは利他遺伝子 A_1 と非利他遺伝子 A_2 が存在し、 g が個体の遺伝子型値に対応している状況を表していることにほかならない。パートナーの遺伝子型値を g_p と書く。

このとき、注目する個体の適応度 w は

$$w = 1 - cg + bg_p \quad (4.22)$$

と書けると仮定する (b や c は 1 に比べて十分小さく、したがっていつも $w > 0$ が満たされていると仮定する)。式 (4.22) を Price 方程式 (4.4) に代入して

$$\bar{w}\Delta\bar{g} = \text{cov}[1 - cg + bg_p, g] = -c \text{var}[g] + b \text{cov}[g_p, g] \quad (4.23)$$

を得る。したがって利他遺伝子の頻度が増加する条件は、 $\Delta\bar{g} > 0$ 、すなわち

$$-c + b \cdot \frac{\text{cov}[g, g_p]}{\text{var}[g]} > 0 \quad (4.24)$$

である。

式 (4.24) に現れた回帰係数

$$r = \frac{\text{cov}[g, g_p]}{\text{var}[g]} \quad (4.25)$$

は、個体とパートナーの遺伝的類似度を表し、回帰係数で表されていることからこれを回帰血縁度 (regression relatedness) と呼ぶ。回帰血縁度 (4.25) を用いれば、式 (4.24) は簡単に

$$-c + br > 0 \quad (4.26)$$

と書かれる。式 (4.26) を Hamilton 則 (Hamilton's rule) と呼ぶ。

例 4.9 Price 方程式 (4.4) は属性 z の集団平均の変化しか予測できないと思われがちだが、例えば属性 z の集団内分散の変化も予測できる。そのためには z の分散 $\text{var}[z]$ が

$$\text{var}[z] = \overline{z^2} - \bar{z}^2 \quad (4.27)$$

と書けていたことを思い出そう。したがって分散の変化は

$$\Delta\text{var}[z] = \text{var}[z'] - \text{var}[z] = (\overline{z'^2} - \bar{z}'^2) - (\overline{z^2} - \bar{z}^2) = \left\{ \overline{z^2} + \Delta\overline{z^2} - (\bar{z} + \Delta\bar{z})^2 \right\} - (\overline{z^2} - \bar{z}^2) = \Delta\overline{z^2} - 2\bar{z}\Delta\bar{z} - (\Delta\bar{z})^2 \quad (4.28)$$

である。ここで

$$q = z^2 \quad (4.29)$$

も個体の持つ属性であるから、 q に Price 方程式 (4.4) を用いて

$$\bar{w}\Delta\bar{q} = \text{cov}[w, q] \quad (4.30)$$

であり、式 (4.29) の q の定義を上式に代入して

$$\bar{w}\Delta\overline{z^2} = \text{cov}[w, z^2] \quad (4.31)$$

を得る。したがって式 (4.28) に、通常の Price 方程式 (4.4) と z^2 に関する Price 方程式 (4.31) を代入して

$$\Delta\text{var}[z] = \frac{\text{cov}[w, z^2]}{\bar{w}} - 2\bar{z}\frac{\text{cov}[w, z]}{\bar{w}} - \left(\frac{\text{cov}[w, z]}{\bar{w}}\right)^2 \quad (4.32)$$

を得る。もしくは、両辺に $\bar{z}^2(\text{cov}[w, 1]/\bar{w})$ (これは $= 0$ である) を加えて整理することにより、より見やすい式

$$\Delta\text{var}[z] = \frac{\text{cov}[w, (z - \bar{z})^2]}{\bar{w}} - \left(\frac{\text{cov}[w, z]}{\bar{w}}\right)^2 \quad (4.33)$$

に変形もできる。

4.2 Price 方程式の分解

Δz 付きの Price 方程式 (4.3) の応用例として、**構造のある集団 (structured population)** を考えよう。

集団はラベル $i = 1, \dots, n_i$ で添字付けられたいくつかの**分集団 (subdivided population)**に分かれており、各分集団 i の中の個体はラベル $j = 1, \dots, n_j$ で添字付けられているとする (n_j は i に依らず一定とする)。すなわちラベル (i, j) で、分集団 i 中の個体 j を表すものとする。

以下の仮定を置く。

- (1) 個体 (i, j) はある属性 z_{ij} を備えているとする。
- (2) 個体 (i, j) には適応度 w_{ij} が定まっており、適応度の数だけ次世代に子を残すとする。
- (3) 生まれた子は親の属性 z_{ij} をそのまま引き継ぐものとする。

以下で様々な平均を表すために記号 avg を以下のように定義する。すなわち (i, j) -添字を持つ任意の記号 x_{ij} に対して、

$$\begin{aligned} \text{avg}_{ij} [x_{ij}] &= \frac{1}{n_i n_j} \sum_i \sum_j x_{ij} && (\bar{x} \text{ とも書く}) \\ \text{avg}_j [x_{ij}] &= \frac{1}{n_j} \sum_j x_{ij} && (x_i \text{ とも書く}) \\ \text{avg}_i [x_i] &= \frac{1}{n_i} \sum_i x_i && (\text{定義より} = \bar{x} \text{ である}) \end{aligned} \quad (4.34)$$

と定める。共分散も同様に、 (i, j) -添字を持つ任意の記号 x_{ij}, y_{ij} に対して、

$$\begin{aligned} \text{cov}_{ij} [x_{ij}, y_{ij}] &= \frac{1}{n_i n_j} \sum_i \sum_j (x_{ij} - \bar{x})(y_{ij} - \bar{y}) \\ \text{cov}_j [x_{ij}, y_{ij}] &= \frac{1}{n_j} \sum_j (x_{ij} - x_i)(y_{ij} - y_i) \\ \text{cov}_i [x_i, y_i] &= \frac{1}{n_i} \sum_i (x_i - \bar{x})(y_i - \bar{y}) \end{aligned} \quad (4.35)$$

と定める。

まず、集団構造を意識しない Price 方程式を考える。モデルの仮定 (3) より $(\Delta z)_{i,j} = 0$ だから、これは Price 方程式 (4.4) を用いて

$$\bar{w} \Delta \bar{z} = \text{cov}_{ij} [w_{ij}, z_{ij}] \quad (4.36)$$

である。

次に集団構造を意識した Price 方程式を考える。分集団 i の個体は子も分集団 i の中に産むと仮定する。次世代において分集団 i には様々な子が存在するが、分集団 i のことは分かるが、分集団 i 内の各個体については分からない観察者を想定してみよう。具体的に、この観察者は

- 次世代の分集団 i のサイズ
- 次世代の分集団 i 内の属性 z の平均値

しか分からないとしよう。

まず次世代の分集団 i のサイズは、個体 (i, j) が w_{ij} の子を産むことを考えると、 $\sum_j w_{ij}$ である。もともとのサイズは n_j であったから、分集団 i は

$$\frac{\sum_j w_{ij}}{n_j} = w_i \quad (4.37)$$

倍になったことになる。これを、分集団を一つの**超個体 (superorganism)** と見た時に、「超個体 i が『子』を w_i 個体産んだ」もしくは「超個体 i の『適応度』が w_i であった」とみなすこととする。

次に、個体 (i, j) が属性 z_{ij} を持つ子を w_{ij} 個体産んだのだから、次世代の分集団 i 内の属性 z の平均値は

$$z'_i = \frac{\sum_j w_{ij} z_{ij}}{\sum_j w_{ij}} \quad (4.38)$$

である。現代の分集団 i 内の属性 z の平均値は z_i であったから、観察者から見れば「超個体は『属性』が z_i だったが、超個体の『子』は『属性』が z'_i である」もしくは「超個体の『属性』が『子』に受け継がれる際、『属性』が $\Delta z_i = z'_i - z_i$ だけ変化した」とみなすことができる。

したがって、分集団 i を超個体とみなす観察者の視点では、Price 方程式 (4.3) を用いることができ

$$\bar{w} \Delta \bar{z} = \text{cov}_i [w_i, z_i] + \text{avg}_i [w_i \Delta z_i] \quad (4.39)$$

が成り立つ。

一方で、先ほどとは反対に分集団 i 内の個体のみを知る別の観測者の視点に立てば、分集団 i 内では Price 方程式 (4.4) が成り立ち、それは

$$w_i \Delta z_i = \text{cov}_j [w_{ij}, z_{ij}] \quad (4.40)$$

である。

つまり Price 方程式 (4.39) は分集団間 (between) にはたらく自然選択を、一方で Price 方程式 (4.40) は分集団内 (within) にはたらく自然選択を表すと考えることができ、式 (4.40) を (4.39) の右辺第二項に代入することで、Price 方程式の分解

$$\bar{w} \Delta \bar{z} = \underbrace{\text{cov}_i [w_i, z_i]}_{\text{between}} + \underbrace{\text{avg}_i [\text{cov}_j [w_{ij}, z_{ij}]]}_{\text{within}} \quad (4.41)$$

が得られた (Price, 1972)。

集団構造がさらに入れ子状になっている場合 (例：分集団 i の中に小分集団 j がありその中に個体 k がいる場合) なども添え字 (i, j, k) の各レベルの Price 方程式を組み合わせれば、Price 方程式の分解を得ることができる。

解説 4.10 Price 方程式の分解 (4.41) は特別な観察者を考えなくとも、式 (4.35) を機械的に分解することで得られる。その導出を下に示す。

$$\begin{aligned} & \text{cov}_{ij} [x_{ij}, y_{ij}] \\ &= \frac{1}{n_i n_j} \sum_i \sum_j (x_{ij} - \underbrace{x_i + x_i - \bar{x}}_{=0}) (y_{ij} - \underbrace{y_i + y_i - \bar{y}}_{=0}) \\ &= \frac{1}{n_i n_j} \sum_i \sum_j \{ (x_{ij} - x_i)(y_{ij} - y_i) + (x_{ij} - x_i)(y_i - \bar{y}) + (x_i - \bar{x})(y_{ij} - y_i) + (x_i - \bar{x})(y_i - \bar{y}) \} \\ &= \frac{1}{n_i} \sum_i \text{cov}_j [x_{ij}, y_{ij}] + \frac{1}{n_i} \sum_i (y_i - \bar{y}) \underbrace{\frac{1}{n_j} \sum_j (x_{ij} - x_i)}_{=0} + \frac{1}{n_i} \sum_i (x_i - \bar{x}) \underbrace{\frac{1}{n_j} \sum_j (y_{ij} - y_i)}_{=0} + \frac{1}{n_i} \sum_i (x_i - \bar{x})(y_i - \bar{y}) \\ &= \text{avg}_i [\text{cov}_j [x_{ij}, y_{ij}]] + \text{cov}_i [x_i, y_i] \end{aligned} \quad (4.42)$$

例 4.11 Price 方程式の分解 (4.41) の例として、分集団内の利他行動を考える。各分集団には $n_j = n$ 個体があり、利他遺伝子を持っている ($g = 1$) か持っていない ($g = 0$) かのどちらかであるとする。各個体の適応度は

$$w_{ij} = 1 - cg_{ij} + bg_i \quad (4.43)$$

で与えられるとする (g_i は分集団 i 内の利他的個体の頻度である)。すなわち、利他行動は行為者にコスト c がかかり、その代わり自分を含めた分集団内の全員にそれぞれ利益 b/n を生む。利他行動の定義から $c > (b/n)$ を、また協力の利益とコストの関係について $b > c$ を、以下仮定する。また b, c は十分小さく、したがって適応度 w は常に正であると仮定する。

この適応度を Price 方程式の分解 (4.41) に代入する。within 項の中身について、

$$\text{cov}_j [w_{ij}, g_{ij}] = \text{cov}_j [1 - cg_{ij} + bg_i, g_{ij}] = \underbrace{\text{cov}_j [1, g_{ij}]}_{=0} - c \underbrace{\text{cov}_j [g_{ij}, g_{ij}]}_{=\text{var}_j [g_{ij}]} + b \underbrace{\text{cov}_j [g_i, g_{ij}]}_{=0} \quad (4.44)$$

である。また式 (4.43) の分集団 i 内の平均 avg_i を取って、

$$w_i = 1 - cg_i + bg_i = 1 + (b - c)g_i \quad (4.45)$$

であるから、Price 方程式の分解 (4.41) の between 項について

$$\text{cov}_i [w_i, g_i] = \text{cov}_i [1 + (b - c)g_i, g_i] = \underbrace{\text{cov}_i [1, g_i]}_{=0} + (b - c) \underbrace{\text{cov}_i [g_i, g_i]}_{=\text{var}_i [g_i]} \quad (4.46)$$

である。以上より、Price 方程式の分解 (4.41) は

$$\bar{w}\Delta\bar{g} = \underbrace{(b - c)\text{var}_i [g_i]}_{\text{between}} + \underbrace{(-c)\text{avg}_i [\text{var}_j [g_{ij}]]}_{\text{within}} \quad (4.47)$$

となる。

式 (4.47) の between 項は一般に正であり、その中の分散 $\text{var}_i [g_i]$ は分集団間でどれだけ利他的個体の頻度にばらつきがあるか (分集団間分散) を表す。一方で within 項は一般に負であり、その中の分散 (の平均) $\text{avg}_i [\text{var}_j [g_{ij}]]$ は分集団内でどれだけ利他的個体の頻度にばらつきがあるか (分集団内分散) を表す。

つまり分集団間で利他個体の頻度にばらつきがあるほど利他行動は進化しやすく、分集団内で利他個体と非利他個体の共存があればあるほど利他行動は進化しにくい。これは直観的にも理解できる。

例 4.12 さらに式 (4.47) を理解するために具体的な数値例を考える。

分集団は $i = 1, 2$ の二つしかないとし、各分集団のサイズを 10 とする。集団内には利他的個体が 10 個体、非利他的個体が 10 個体の、合計 20 個体がいるとする。

(1) 第一の例では、各分集団の組成は以下の表で与えられているとする。

	利他的個体の数	非利他的個体の数	総個体数
分集団 1	10	0	10
分集団 2	0	10	10
集団全体	10	10	20

この時

$$g_{1,1} = \dots = g_{1,10} = 1, \quad g_{2,1} = \dots = g_{2,10} = 0 \quad (4.48)$$

としてよく、分集団内平均は

$$g_1 = 1, \quad g_2 = 0 \quad (4.49)$$

であり、分集団間分散は

$$\text{var}_i [g_i] = \text{avg}_i [g_i^2] - (\text{avg}_i [g_i])^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4} \quad (4.50)$$

である。また各分集団内分散は

$$\begin{aligned} \text{var}_j [g_{1j}] &= \text{avg}_j [g_{1j}^2] - (\text{avg}_j [g_{1j}])^2 = 1 - 1^2 = 0 \\ \text{var}_j [g_{2j}] &= \text{avg}_j [g_{2j}^2] - (\text{avg}_j [g_{2j}])^2 = 0 - 0^2 = 0 \end{aligned} \quad (4.51)$$

より分集団内分散の平均は

$$\text{avg}_i [\text{var}_j [g_{ij}]] = 0 \quad (4.52)$$

である。以上を式 (4.47) に代入して

$$\bar{w}\Delta\bar{g} = \underbrace{\frac{b-c}{4}}_{\text{between}} + \underbrace{0}_{\text{within}} \quad (4.53)$$

を得る。この例では利他的個体が分集団 1 だけに集中して存在しており、このような場合利他行動は進化する。

(2) 第二の例では、各分集団の組成は以下の表で与えられているとする。

	利他的個体の数	非利他的個体の数	総個体数
分集団 1	5	5	10
分集団 2	5	5	10
集団全体	10	10	20

この時

$$\begin{aligned} g_{1,1} = \dots = g_{1,5} = 1, \quad g_{1,6} = \dots = g_{1,10} = 0 \\ g_{2,1} = \dots = g_{2,5} = 1, \quad g_{2,6} = \dots = g_{2,10} = 0 \end{aligned} \quad (4.54)$$

としてよく、分集団内平均は

$$g_1 = \frac{1}{2}, \quad g_2 = \frac{1}{2} \quad (4.55)$$

であり、分集団間分散は

$$\text{var}_i [g_i] = \text{avg}_i [g_i^2] - (\text{avg}_i [g_i])^2 = \frac{1}{4} - \left(\frac{1}{2}\right)^2 = 0 \quad (4.56)$$

である。また各分集団内分散は

$$\begin{aligned} \text{var}_j [g_{1j}] &= \text{avg}_j [g_{1j}^2] - (\text{avg}_j [g_{1j}])^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4} \\ \text{var}_j [g_{2j}] &= \text{avg}_j [g_{2j}^2] - (\text{avg}_j [g_{2j}])^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4} \end{aligned} \quad (4.57)$$

より分集団内分散の平均は

$$\text{avg}_i [\text{var}_j [g_{ij}]] = \frac{1}{4} \quad (4.58)$$

である。以上を式 (4.47) に代入して

$$\bar{w}\Delta\bar{g} = \underbrace{0}_{\text{between}} + \underbrace{\left(-\frac{c}{4}\right)}_{\text{within}} \quad (4.59)$$

を得る。この例では利他的個体は両方の分集団にまんべんなく散らばっているため、利他行動は進化できない。

問題 4.13 例 4.12 に倣って、以下の状況を解析せよ。

	利他的個体の数	非利他的個体の数	総個体数
分集団 1	8	2	10
分集団 2	2	8	10
集団全体	10	10	20

4.3 分布の発展方程式

ここまでは Price 方程式 (4.3) を用いて主に属性の平均値 \bar{z} の変化を調べてきたが、平均ではなく頻度分布自体はどのように変化しているだろうか？

現世代の属性 z の頻度分布を $p(z)$ とおき、属性 z を持つ個体の適応度を $w(z)$ とおく。子は親の属性 z をそのまま受け継ぐと仮定した時、次世代の z の分布関数 $p_{\text{next}}(z)$ はどのような形だろうか？（注：次世代の分布を表すのに記号 $p'(z)$ を使いたところだが、この記号は微分と混同しやすいので、あえて分布だけは記号 $p_{\text{next}}(z)$ を用いる）。

$p(z), p_{\text{next}}(z)$ は頻度分布なので、全域での積分は

$$\int_{-\infty}^{\infty} p(z)dz = 1, \quad \int_{-\infty}^{\infty} p_{\text{next}}(z)dz = 1 \quad (\text{規格化条件}) \quad (4.60)$$

を満たしていなければならない。 z を持つ個体数は次世代では $w(z)$ 倍されるので、 $p_{\text{next}}(z)$ はある定数 C を用いて

$$p_{\text{next}}(z) = Cw(z)p(z) \quad (4.61)$$

と書かれる。ここで $p_{\text{next}}(z)$ は規格化条件 (4.60) を満たさねばならないので

$$\int_{-\infty}^{\infty} p_{\text{next}}(z)dz = C \int_{-\infty}^{\infty} w(z)p(z)dz = 1 \quad (4.62)$$

であることから

$$C = \frac{1}{\int_{-\infty}^{\infty} w(z)p(z)dz} \quad (4.63)$$

が分かる。ここで C の分母

$$\bar{w} = \int_{-\infty}^{\infty} w(z)p(z)dz \quad (4.64)$$

は現世代の集団平均適応度である。したがって式 (4.63) を式 (4.61) に戻して

$$p_{\text{next}}(z) = \frac{w(z)p(z)}{\int_{-\infty}^{\infty} w(z)p(z)dz} = \frac{w(z)p(z)}{\bar{w}} \quad (4.65)$$

である。

現世代の分布の平均と分散を \bar{z}, σ^2 、次世代の分布の平均と分散を $\bar{z}', (\sigma^2)'$ とそれぞれおくと、それらは

$$\begin{aligned} \bar{z} &= \int_{-\infty}^{\infty} zp(z)dz & \sigma^2 &= \int_{-\infty}^{\infty} (z - \bar{z})^2 p(z)dz \\ \bar{z}' &= \int_{-\infty}^{\infty} zp_{\text{next}}(z)dz & (\sigma^2)' &= \int_{-\infty}^{\infty} (z - \bar{z}')^2 p_{\text{next}}(z)dz \end{aligned} \quad (4.66)$$

と書けるが、適応度関数 $w(z)$ の具体的な形が分からないと、これら平均と分散の時間変化はこれ以上は記述できない。

4.4 正規分布を用いた理論

現世代の分布に $p(z)$ に正規分布を、適応度関数 $w(z)$ に正規分布型の関数を仮定した場合の理論を展開しよう (Lande 1981)。

以下を仮定する。

- (1) 現世代の分布 $p(z)$ は平均 \bar{z} 、分散 σ^2 の正規分布で与えられる。すなわち

$$p(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(z - \bar{z})^2}{2\sigma^2} \right] \quad (4.67)$$

とする。

- (2) 適応度関数 $w(z)$ を、平均 θ 、分散 ω^2 の正規分布型の関数 (Gauss 関数)

$$w(z) = \exp \left[-\frac{(z - \theta)^2}{2\omega^2} \right] \quad (4.68)$$

で与える。この適応度関数 $w(z)$ は生物学の言葉で言えば、 $z = \theta$ が最適な形質であり、適応度のピークの幅が ω 程度であることを表している。

解説 4.14 式 (4.67) において厳密に正規分布を仮定した一方で、式 (4.68) ではそうしなかった理由は、分布 $p(z)$ だけが規格化条件 (4.60) を満たさなければならないからである。一方で適応度関数は定数倍しても進化の結果にはなんら影響を及ぼさない（これは式 (4.65) の形からも分かる）、式 (4.68) の先頭には式 (4.67) にあるような定数が付かない。もちろん式 (4.68) を厳密な正規分布の式にははいけないという理由もなく、その時は

$$w(z) = \frac{1}{\sqrt{2\pi\omega^2}} \exp\left[-\frac{(z-\theta)^2}{2\omega^2}\right] \quad (4.69)$$

とすれば良いだけである。

ここで二つの Gauss 関数の積は再び Gauss 関数になるという数学的事実を使う。したがって式 (4.65) の右辺の分子 $w(z)p(z)$ は Gauss 関数である。また $p_{\text{next}}(z)$ はすでに式 (4.65) のように規格化されているから、再び厳密に正規分布である。その平均と分散は以下のように与えられることが数学的事実として知られている。

$$\bar{z}' = \frac{\omega^2\bar{z} + \sigma^2\theta}{\sigma^2 + \omega^2} \quad (4.70a)$$

$$(\sigma^2)' = \frac{\sigma^2\omega^2}{\sigma^2 + \omega^2} \quad (4.70b)$$

式 (4.70) は平均と分散の組に関する更新則 $(\bar{z}_t, \sigma_t^2) \rightarrow (\bar{z}_{t+1}, \sigma_{t+1}^2)$ を与えるから、初期値 (\bar{z}_0, σ_0^2) を与えれば解くことができる。

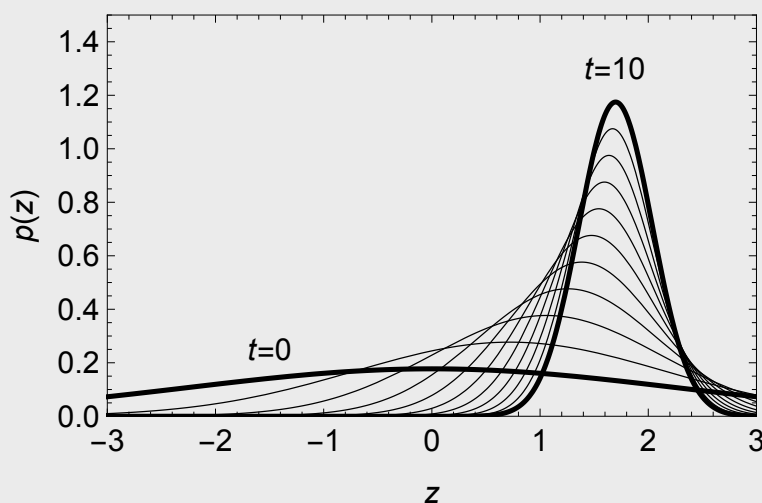


図 4.1: 漸化式 (4.70) の解を分布の形で描いた。初期条件は $(\bar{z}_0, \sigma_0^2) = (0, 1.5)$ 、適応度関数のパラメータは $(\theta, \omega^2) = (2, 2)$ を用いた。時刻 $t = 0$ から $t = 10$ までのプロットで、世代が進むにつれ、 $\theta = 2$ に鋭いピークを持つ分布に近づいていることが分かる。

解説 4.15 漸化式 (4.70) を陽に解いてみよう。まず (4.70b) の両辺の逆数を取ることで

$$\frac{1}{\sigma_{t+1}^2} = \frac{1}{\sigma_t^2} + \frac{1}{\omega^2} \quad (4.71)$$

であるので、数列 $\{1/\sigma_t^2\}_{t=0, \dots}$ は初項 $1/\sigma_0^2$ 、公差 $1/\omega^2$ の等差数列であり、

$$\frac{1}{\sigma_t^2} = \frac{1}{\sigma_0^2} + \frac{t}{\omega^2} \quad (4.72)$$

である。これを用いて

$$\sigma_t^2 = \frac{\sigma_0^2\omega^2}{\sigma_0^2 t + \omega^2} \quad (4.73)$$

が分かる。したがって分散 σ_t^2 は時間とともに単調減少する。

次に式 (4.70a) の両辺から θ を引いて整理することで

$$\bar{z}_{t+1} - \theta = \frac{\omega^2}{\sigma_t^2 + \omega^2} (\bar{z}_t - \theta) \quad (4.74)$$

が分かる。よって式 (4.74) を繰り返し用いて

$$\bar{z}_t - \theta = \left(\frac{\omega^2}{\sigma_{t-1}^2 + \omega^2} \cdot \frac{\omega^2}{\sigma_{t-2}^2 + \omega^2} \cdots \frac{\omega^2}{\sigma_0^2 + \omega^2} \right) (\bar{z}_0 - \theta) \quad (4.75)$$

である。ここで式 (4.73) から直接計算で

$$\frac{\omega^2}{\sigma_t^2 + \omega^2} = \frac{\sigma_0^2 t + \omega^2}{\sigma_0^2 (t+1) + \omega^2} \quad (4.76)$$

が分かるので式 (4.76) を式 (4.75) に用いると

$$\bar{z}_t - \theta = \left(\frac{\sigma_0^2 (t-1) + \omega^2}{\sigma_0^2 t + \omega^2} \cdot \frac{\sigma_0^2 (t-2) + \omega^2}{\sigma_0^2 (t-1) + \omega^2} \cdots \frac{\omega^2}{\sigma_0^2 + \omega^2} \right) (\bar{z}_0 - \theta) = \frac{\omega^2}{\sigma_0^2 t + \omega^2} (\bar{z}_0 - \theta) \quad (4.77)$$

を得る。これを整理して

$$\bar{z}_t = \frac{\sigma_0^2 \theta t + \omega^2 \bar{z}_0}{\sigma_0^2 t + \omega^2} \quad (4.78)$$

を得る。

特に解 (4.78) および (4.73) で $t \rightarrow \infty$ として、

$$\bar{z}_\infty = \theta \quad \text{および} \quad \sigma_\infty^2 = 0 \quad (4.79)$$

が分かる。すなわち分布 $\{p_t(z)\}_{t=0, \dots}$ は時間とともに $z = \theta$ に尖ったピークをもつ関数に近づく（極限はデルタ関数）。適応度関数 $w(z)$ が $z = \theta$ で最大値を取ることを考えると、これは当然の結果と言える。

解説 4.16 本章で仮定していた表現型 *gambit* を一旦放棄して、前章で得た量的遺伝学の知見と本節の結果を組み合わせることを考えてみよう。具体的には二倍体有性生殖生物を考え、属性 z とは表現型値であると考え。そして表現型値の分解 (3.111) を考える。

式 (4.70a) を書き換えると

$$\bar{z}' - \bar{z} = \frac{\sigma^2}{\sigma^2 + \omega^2} (\theta - \bar{z}) \quad (4.80)$$

を得るが、右辺は単に選択差を表しているにすぎない。育種家方程式 (3.103) より、選択応答はこれに狭義の遺伝率 h^2 をかけたものにほかならないから、二倍体有性生殖生物を考えた結果、式 (4.80) は次のような修正を受ける。

$$\bar{z}' - \bar{z} = h^2 \frac{\sigma^2}{\sigma^2 + \omega^2} (\theta - \bar{z}) \quad (4.81)$$

ここで σ^2 は表現型分散であり、したがって積 $h^2 \sigma^2$ は相加遺伝分散である。

4.5 頻度依存選択の扱い

前二節では適応度関数は $w(z)$ の形で与えられ、これは時間とともに不変だった。しかし生物学のモデルではしばしば適応度関数は「集団の状態」に依存して変わることがある。つまり $w(z)$ 自体が $p(z)$ の形に依存している場合がある。

例 4.17 例えば z が動物の体サイズを表すとしよう。他者との力比べで繁殖成功の結果が決まるとすると、「 $z = 1$ だから弱い、 $z = 2$ だから中程度、 $z = 3$ だから強い」のように絶対的な体サイズが適応度を決定するのではなく、「集団の平均体サイズ \bar{z} に比べてどれだけ z が大きいか」のような相対的な体サイズがものを言うかもしれない。このような状況では、適応度関数 w は z のみならず \bar{z} の関数にもなりうる。

そこで本節では属性 z を持つ個体の適応度 w が

- 個体の属性 z
- 属性の集団平均 \bar{z}

の二つに依存する二変数関数 $w(z, \bar{z})$ として書ける状況を解析しよう。

4.4 節では分布 $p(z)$ と適応度関数 $w(z)$ を正規分布 (型の関数) としておくことで解析的な扱いが可能になった。本節ではそうはせず、代わりに分布 $p(z)$ に特殊な仮定を置くことで、進化ダイナミクスを得ることを目指す。

式 (4.73) で見たように、自然選択がかかると有利な個体のみが生き残る影響で分布 $p(z)$ の分散が時間とともに小さくなることがある。例えば、式 (4.79) では分散は時間 $t \rightarrow \infty$ で 0 になった。しかしながら実際の生物集団では様々な理由により分散は 0 にはならない。例えば親から子に形質が伝わる際に突然変異が起きると、集団には新たな分散が生み出されることになる。

そこで以降では、次を仮定する。

- 分布 $p(z)$ は常にその平均値 \bar{z} に鋭いピークを持つ単峰 (unimodal) 型をしている。
- 分布 $p(z)$ の分散 σ^2 は小さな値だが 0 ではなく、モデルに含まれない効果などによって常に一定値に保たれている。

この仮定の下で、平均の時間変化 $\Delta \bar{z} = \bar{z}' - \bar{z}$ を求めてみよう。

解説 4.18 個体の属性 z は \bar{z} の近辺に集中しているはずであるから、分布 p に属するほとんどの個体に対して絶対値 $|z - \bar{z}|$ の値は小さい。そこで適応度関数 $w(z, \bar{z})$ を変数 z に関して \bar{z} の周りで Taylor 展開すると

$$w(z, \bar{z}) = w(z, \bar{z})|_{z=\bar{z}} + (z - \bar{z}) \left. \frac{\partial w(z, \bar{z})}{\partial z} \right|_{z=\bar{z}} + \frac{1}{2} (z - \bar{z})^2 \left. \frac{\partial^2 w(z, \bar{z})}{\partial z^2} \right|_{z=\bar{z}} + \dots \quad (4.82)$$

である。ここで式 (4.82) の右辺第一項は

$$w(z, \bar{z})|_{z=\bar{z}} = w(\bar{z}, \bar{z}) \approx \bar{w} \quad (4.83)$$

である。なぜなら分布 p は $z = \bar{z}$ の近辺に集中しているからである。Price 方程式 (4.4) に式 (4.82) および (4.83) を適用して

$$\begin{aligned} \bar{w} \Delta \bar{z} &= \text{cov}[w, z] \\ &= \overline{\{w(z, \bar{z}) - \bar{w}\}(z - \bar{z})} \\ &= \int_{-\infty}^{\infty} \left\{ (z - \bar{z}) \left. \frac{\partial w(z, \bar{z})}{\partial z} \right|_{z=\bar{z}} + \frac{1}{2} (z - \bar{z})^2 \left. \frac{\partial^2 w(z, \bar{z})}{\partial z^2} \right|_{z=\bar{z}} + \dots \right\} (z - \bar{z}) p(z) dz \\ &= \left(\int_{-\infty}^{\infty} (z - \bar{z})^2 p(z) dz \right) \left. \frac{\partial w(z, \bar{z})}{\partial z} \right|_{z=\bar{z}} + \left(\int_{-\infty}^{\infty} \frac{1}{2} (z - \bar{z})^3 p(z) dz \right) \left. \frac{\partial^2 w(z, \bar{z})}{\partial z^2} \right|_{z=\bar{z}} + \dots \end{aligned} \quad (4.84)$$

となる。ここで式 (4.84) において、分散の定義より

$$\int_{-\infty}^{\infty} (z - \bar{z})^2 p(z) dz = \sigma^2 \quad (4.85)$$

であり、また分布 p が \bar{z} の近辺に集中しているという仮定から、それより高次のモーメントは σ^2 の大きさに比べて無視できるとして良い。つまり

$$\int_{-\infty}^{\infty} (z - \bar{z})^n p(z) dz \approx 0 \quad (n \geq 3) \quad (4.86)$$

である。式 (4.85) および (4.86) を式 (4.84) に代入して、近似

$$\bar{w} \Delta \bar{z} = \sigma^2 \left. \frac{\partial w(z, \bar{z})}{\partial z} \right|_{z=\bar{z}} \quad (4.87)$$

を得る。もしくは

$$\left. \frac{\partial \log w(z, \bar{z})}{\partial z} \right|_{z=\bar{z}} = \frac{1}{w(\bar{z}, \bar{z})} \left. \frac{\partial w(z, \bar{z})}{\partial z} \right|_{z=\bar{z}} = \frac{1}{\bar{w}} \left. \frac{\partial w(z, \bar{z})}{\partial z} \right|_{z=\bar{z}} \quad (4.88)$$

を式 (4.87) に用いて整理すると

$$\Delta \bar{z} = \sigma^2 \left. \frac{\partial \log w(z, \bar{z})}{\partial z} \right|_{z=\bar{z}} \quad (4.89)$$

を得る (Iwasa & Pomiankowski, 1991)。

解説 4.19 解説 4.16 の時と同じように、表現型 gambit を放棄し二倍体有性生殖生物を考えたら式 (4.89) はどのように変更を受けるだろうか。式 (4.89) の右辺は選択差を表しているので、それに狭義の遺伝率 h^2 を乗じたものが選択応答となる。仮に h^2 を時間に依らない定数だと考えれば、式 (4.89) は

$$\Delta \bar{z} = h^2 \sigma^2 \left. \frac{\partial \log w(z, \bar{z})}{\partial z} \right|_{z=\bar{z}} \quad (4.90)$$

と修正を受ける。ここで σ^2 は表現型分散であり、したがって $h^2 \sigma^2$ は相加遺伝分散である。

例 4.20 体サイズの例に戻って、体サイズ $z (> 0)$ を持つ個体の適応度が

$$w(z, \bar{z}) = \underbrace{\exp[(z/\bar{z})^n]}_{\text{他者との競争に勝つ利益}} \times \underbrace{\exp[-z]}_{\text{体を大きくするコスト}} \quad (4.91)$$

と書けているとしよう ($n \geq 1$)。競争の結果は体サイズの相対値 z/\bar{z} で決まり、コストは体サイズ z そのもので決まるというこれらの仮定は生物学的にある程度妥当である。式 (4.91) を式 (4.89) に代入する。 $\log w(z, \bar{z}) = (z/\bar{z})^n - z$ であるから

$$\Delta \bar{z} = \sigma^2 \left\{ n (z/\bar{z})^{n-1} \cdot (1/\bar{z}) - 1 \right\} \Big|_{z=\bar{z}} = \sigma^2 \{ (n/\bar{z}) - 1 \} \quad (4.92)$$

となる。したがって最終的に $\bar{z}^* = n$ なる体サイズが実現すると予測される。