2007 年 9 月 4 日

遺伝の比率の数列について

新潟工科大学 情報電子工学科 竹野茂治

1 はじめに

ある本を読んでいたら、性染色体によって引き起されるある遺伝的な性質があり、その説明があった上で「よって女性にはほとんど起こらない」と書いてあった。

それが何となく気になったのであるが、この比率は、世代毎の数列として考えることができるので、実際に計算してみる。

2 問題

その本によると、問題の遺伝的な性質は性染色体 X の中にある因子によって引き起され、男性の性染色体 XY の X にその因子がある場合は、その男性は必ずその性質を持つが、女性の性染色体 XX の一方の X にその因子があっただけではその性質は出ず、両方の X にその因子があるときだけその性質が出る、ということであった。

その性質を持つ人の割合自体は、全体に比べてかなり小さいのであるが、上のような状況により、女性にその性質が出る割合は特に小さい、のだということである。

今、その因子を持つ X を $\bar{X}$ と書くこととし、その因子を持たない X は、そのまま X と書くことにすると、男性、女性の性染色体には、それぞれ

- 男性: XY, $\bar{X}Y$
- 女性: XX, $X\bar{X}$, $\bar{X}\bar{X}$

のように 2 種類、3 種類の状態があることになるが、その性質が発現するのは $\bar{X}Y$ と $\bar{X}\bar{X}$ だけ、ということである。

この因子は、親から子へは次のように遺伝する。例えば、父親が $\bar{X}_1Y_1$, 母親が $X_2\bar{X}_2$ (区別がつきやすいように、ここでは添え字 1,2 をつけておくことにする) の場合、その子には父親の一つの染色体と母親の一つの染色体が受け継がれるので、

- $\bar{X}_1 X_2$: 性質を持たない女性
- $\bar{X}_1 \bar{X}_2$ 性質を持つ女性
- $Y_1 X_2$: 性質を持たない男性
- $Y_1 \bar{X}_2$: 性質を持つ男性

の 4 通りが起こりうる。これらはランダムに、それぞれ 1/4 の確率で起こると考えられる。

3 各状態の比率の継承

次に、各状態の分布 (比率) が、どのように次の世代に継承されるか、ということを考えてみる。

簡単のため、男女は同数であるとし、各夫婦は $X, \bar{X}$ の状態にかかわらず同程度の子を作ると考えることにする。

男性の $\bar{X}Y$ の割合を A とし、よって XY の割合を $1 - A$ とする。同様に、女性の $\bar{X}\bar{X}$ の割合を B, $\bar{X}X$ の割合を C とし、よって、XX の割合を $1 - B - C$ とする。

このとき、次の子の世代のそれぞれの割合を求めることにする。子の世代の割合を、ここでは A', B', C' のように書くことにする。

1. $\bar{X}Y$ と $\bar{X}\bar{X}$ の子の場合

   まず、このような親の組み合わせの割合は、男女それぞれの比率が A,B なので、すべての夫婦に対して AB の割合で起こることに注意する。

   そして、この場合は、子には

   $$\bar{X}\bar{X}, \quad \bar{X}\bar{X}, \quad Y\bar{X} \quad ,Y\bar{X}$$

   の 4 通り、つまり実質 2 通りが現れる。つまり、子の世代での男性のうち、$Y\bar{X}$ の割合 A' には、この親の組合せの割合である AB が、女子のうち $\bar{X}\bar{X}$ の割合 B' にも AB が含まれることになる。

   なお、$Y\bar{X}$ になるか $\bar{X}\bar{X}$ になるかは 1/2 であるから、AB ではなく $AB/2$ ではないかと思うかもしれないが、A' は、「男性のうちの」$Y\bar{X}$ の割合であるから、1/2 をつける必要はない。

2. $\bar{X}Y$ と $\bar{X}X$ の子の場合 (夫婦の割合は AC)

   この場合は、子供は

$$\bar{X}\bar{X}, \quad \bar{X}X, \quad Y\bar{X}, \quad YX$$

   の 4 種類が同程度に起こりうる。この夫婦の割合は AC なので、B' に $AC/2$, C' に $AC/2$, A' に $AC/2$, $1-A'$ に $AC/2$ が含まれることになる。

3. $\bar{X}Y$ と XX の子の場合 (夫婦の割合は $A(1-B-C)$)

   この場合は、

$$\bar{X}X, \quad \bar{X}X, \quad YX, \quad YX$$

   の実質 2 種類なので、C' と $1-A'$ に $A(1-B-C)$ が含まれる。

4. XY と $\bar{X}\bar{X}$ の子の場合 (夫婦の割合は $(1-A)B$)

   この場合は、

$$X\bar{X}, \quad X\bar{X}, \quad Y\bar{X}, \quad Y\bar{X}$$

   の実質 2 種類なので、C' と A' に $(1-A)B$ が含まれる。

5. XY と $\bar{X}X$ の子の場合 (夫婦の割合は $(1-A)C$)

   この場合は、

$$X\bar{X}, \quad XX, \quad Y\bar{X}, \quad YX$$

   の 4 種類なので、C', $1-B'-C'$, A', $1-A'$ に $(1-A)C/2$ が含まれる。

6. XY と XX の子の場合 (夫婦の割合は $(1-A)(1-B-C)$)

   この場合は、

$$XX, \quad XX, \quad YX, \quad YX$$

   の 2 種類なので、$1-B'-C'$, $1-A'$ に $(1-A)(1-B-C)$ が含まれる。

結局以上により、

$$
\begin{aligned}
A' &= AB + \frac{1}{2}AC + (1-A)B + \frac{1}{2}(1-A)C, \\
1 - A' &= \frac{1}{2}AC + A(1-B-C) + \frac{1}{2}(1-A)C + (1-A)(1-B-C) \\
B' &= AB + \frac{1}{2}AC, \\
C' &= \frac{1}{2}AC + A(1-B-C) + (1-A)B + \frac{1}{2}(1-A)C \\
1 - B' - C' &= \frac{1}{2}(1-A)C + (1-A)(1-B-C)
\end{aligned}
$$

が得られることになる。それぞれ展開すると、

$$
\begin{aligned}
A' &= B + \frac{1}{2}C && (1) \\
1 - A' &= 1 - B - \frac{1}{2}C && (2) \\
B' &= AB + \frac{1}{2}AC && (3) \\
C' &= A + B + \frac{1}{2}C - 2AB - AC && (4) \\
1 - B' - C' &= 1 - A - B - \frac{1}{2}C + AB + \frac{1}{2}AC && (5)
\end{aligned}
$$

となる。(1), (2) の右辺の和、(3), (4), (5) の右辺の和がそれぞれ 1 となることは容易に確認できるだろう。

この (1)-(5) を用いれば、親から次の世代への割合の継承が計算できることになる。

例えば、$A = 0.4$, $B = 0.6$, $C = 0.2$ の場合、

$$
\begin{aligned}
A' &= 0.6 + 0.1 = 0.7, \\
B' &= 0.24 + 0.04 = 0.28, \\
C' &= 0.4 + 0.6 + 0.1 - 0.48 - 0.08 = 0.54
\end{aligned}
$$

のようになる。同様に、孫の世代まで計算すれば、表 1 のようになる。

	男性		女性		
	$A\ (Y\bar{X})$	$1-A\ (YX)$	$B\ (\bar{X}\bar{X})$	$C\ (\bar{X}X)$	$1-B-C\ (XX)$
親	0.4	0.6	0.6	0.2	0.2
子	0.7	0.3	0.28	0.54	0.18
孫	0.55	0.45	0.385	0.48	0.135

表 1: $A=0.4$, $B=0.6$, $C=0.2$ の場合

4 漸化式

今度は、(1)-(5) を数列の漸化式とみて、その数列の一般の式を求めることを考えてみることにする。つまり、n 世代目の割合を、A_n, $1-A_n$, B_n, C_n, $1-B_n-C_n$ としてその漸化式を立てると、(1)-(5) により、次が成り立つ。

$$\begin{cases} A_{n+1} &= B_n+\frac{1}{2}C_n, \\ B_{n+1} &= A_nB_n+\frac{1}{2}A_nC_n, \\ C_{n+1} &= A_n+B_n+\frac{1}{2}C_n-2A_nB_n-A_nC_n \end{cases} \tag{6}$$

もし、この右辺が A_n, B_n, C_n の一次式であれば、$\boldsymbol{X}_n={}^t(A_n,B_n,C_n)$ とすることで、係数行列 P によって

$$\boldsymbol{X}_{n+1}=P\boldsymbol{X}_n$$

と書けるので、

$$\boldsymbol{X}_n=P^n\boldsymbol{X}_0$$

となり、線形代数の知識を用いれば、この行列 P の固有値を求めることで P^n の成分を n の式で表すことができ、それにより $\boldsymbol{X}_n$ を、そして A_n, B_n, C_n を n の式で表すことが可能となる。

しかし、(6) の右辺は A_n, B_n, C_n の 2 次式で非線形なので、一般にはこのようなことは行えず、n の式で表すことは難しい。ただし、この漸化式 (6) の場合は、これが特殊な形をしているので、そこに着目してそれを求めることができる。

(6) の右辺には、B_n, C_n が、いずれも

$$B_n + \frac{1}{2}C_n$$

の形で入っているので、

$$D_n = B_n + \frac{1}{2}C_n \tag{7}$$

とすると、(6) は、

$$\left\{\begin{array}{rcl} A_{n+1} & = & D_n, \\ B_{n+1} & = & A_n D_n, \\ C_{n+1} & = & A_n + D_n - 2A_n D_n \end{array}\right. \tag{8}$$

と書け、これにより、

$$\begin{array}{rcl} D_{n+1} & = & B_{n+1} + \frac{1}{2}C_{n+1} = A_n D_n + \frac{1}{2}(A_n + D_n - 2A_n D_n) \\ & = & \frac{1}{2}A_n + \frac{1}{2}D_n \end{array}$$

となり、よって、A_n と D_n について

$$\left\{\begin{array}{rcl} A_{n+1} & = & D_n, \\ D_{n+1} & = & \frac{1}{2}A_n + \frac{1}{2}D_n \end{array}\right. \tag{9}$$

という漸化式を導くことができる。

この (9) の右辺は、A_n, D_n の 1 次式なので、前の方針に従って、

$$\left[\begin{array}{c} A_{n+1} \\ D_{n+1} \end{array}\right] = \left[\begin{array}{cc} 0 & 1 \\ 1/2 & 1/2 \end{array}\right] \left[\begin{array}{c} A_n \\ D_n \end{array}\right] \tag{10}$$

から A_n, D_n を求めることができる。ただ、ここでは 2 本の方程式なので、行列を用いずに、より素朴な方法 (特性方程式法) により A_n, D_n を求めてみる。

(9) の 1 本目から $D_n = A_{n+1}$ であり、よって $D_{n+1} = A_{n+2}$ であるから、これらを (9) の 2 本目に代入すれば

$$A_{n+2} = \frac{1}{2}A_{n+1} + \frac{1}{2}A_n \tag{11}$$

という A_n に対する 3 項漸化式が得られる。

なお、この式 (11) は、A_{n+2} が、その前の 2 つの項 A_{n+1}, A_n の平均であることも意味しているので、これにより、他の数列を使わなくても A_n を単独で容易に順次計算できることになる。

一般に、3 項漸化式

$$a_{n+2} + pa_{n+1} + qa_n = 0 \tag{12}$$

に対して、2 次方程式

$$t^2 + pt + q = 0 \tag{13}$$

を (12) の **特性方程式** という。この特性方程式の解が $t = \alpha, \beta$ であるとき、3 項漸化式 (12) は、

$$\begin{aligned} a_{n+2} - \alpha a_{n+1} &= \beta(a_{n+1} - \alpha a_n), \\ a_{n+2} - \beta a_{n+1} &= \alpha(a_{n+1} - \beta a_n) \end{aligned}$$

の形に変形できることが、2 次方程式の解と係数の関係によって容易にわかる。

この 3 項漸化式 (11) の場合は、特性方程式は

$$t^2 = \frac{1}{2}t + \frac{1}{2}$$

であるが、これを解くと

$$2t^2 - t - 1 = 0$$

より、$t = 1, -1/2$ と求まる。(11) は、

$$\begin{aligned} A_{n+2} - A_{n+1} &= -\frac{1}{2}(A_{n+1} - A_n), \\ A_{n+2} + \frac{1}{2}A_{n+1} &= A_{n+1} + \frac{1}{2}A_n \end{aligned}$$

の 2 通りの形に変形できることになる。これにより、$A_{n+1} - A_n$ と $A_{n+1} + A_n/2$ は、それぞれ公比が $(-1/2), 1$ である等比数列であることになり、よって、

$$A_{n+1} - A_n = \left(-\frac{1}{2}\right)^n (A_1 - A_0), \tag{14}$$

$$A_{n+1} + \frac{1}{2}A_n = A_1 + \frac{1}{2}A_0 \tag{15}$$

となるので、(15) から (14) を引いて 2/3 倍すれば、

$$A_n = \frac{2}{3}\left\{1 - \left(-\frac{1}{2}\right)^n\right\} A_1 + \frac{2}{3}\left\{\frac{1}{2} + \left(-\frac{1}{2}\right)^n\right\} A_0$$

となり、$A_1 = D_0$ より、結局

$$\begin{aligned} A_n &= \frac{2}{3}\left\{1 - \left(-\frac{1}{2}\right)^n\right\} D_0 + \frac{2}{3}\left\{\frac{1}{2} + \left(-\frac{1}{2}\right)^n\right\} A_0 \\ &= \frac{A_0 + 2D_0}{3} + \frac{2A_0 - 2D_0}{3}\left(-\frac{1}{2}\right)^n \end{aligned} \tag{16}$$

が得られる。

D_n は、$D_n = A_{n+1}$ であるから、(16) より、

$$\begin{aligned} D_n &= \frac{1}{3}\left\{2 + \left(-\frac{1}{2}\right)^n\right\} D_0 + \frac{1}{3}\left\{1 - \left(-\frac{1}{2}\right)^n\right\} A_0 \\ &= \frac{A_0 + 2D_0}{3} + \frac{D_0 - A_0}{3}\left(-\frac{1}{2}\right)^n \end{aligned} \tag{17}$$

となる。B_n, C_n は、(8), (7) により、

$$B_n = A_{n-1}D_{n-1}, \quad C_n = 2D_n - 2B_n \tag{18}$$

より求めることができる (B_n, C_n の一般形は複雑なので省略)。

5 漸近的な様子

さて、n が大きくなると、$(-1/2)^n$ は 0 に収束していくから、(16), (17) より、$n \to \infty$ のときに

$$A_n \to \alpha, \quad D_n \to \alpha \quad \left(\alpha = \frac{A_0 + 2D_0}{3}\right)$$

と収束することがわかる。B_n, C_n は (18) により、

$$B_n \to \alpha^2, \quad C_n \to 2\alpha - 2\alpha^2$$

のようになる。これらの極限を $A_\infty = \alpha$ のように書くことにすると、結局以下のようになる。

$$\begin{cases} A_\infty = \alpha, \\ 1 - A_\infty = 1 - \alpha, \\ B_\infty = \alpha^2, \\ C_\infty = 2\alpha(1-\alpha), \\ 1 - B_\infty - C_\infty = (1-\alpha)^2 \end{cases} \tag{19}$$

$(-1/2)^n$ は、かなり速く 0 に近づくので、数世代でほぼ (19) の値になる。α は、

$$\alpha = \frac{A_0 + 2D_0}{3} = \frac{A_0 + 2B_0 + C_0}{3} \tag{20}$$

であり、つまりこの最初の世代の比率で与えられる α の値によって、安定的な比率 (19) が決定することになる。

なお、$0 \leq B_0 \leq 1$, $0 \leq C_0 \leq 1$, $0 \leq B_0 + C_0 \leq 1$ なので、$0 \leq 2B_0 + C_0 \leq 2$ (2 になるのは $B_0 = 1$, $C_0 = 0$ のとき) であり、$0 \leq A_0 \leq 1$ であるから α は 0 から 1 までの値を取り得る。

安定的な比率の式 (19) を見ると、α は $0 \leq \alpha \leq 1$ であるから、

$$A_\infty \geq B_\infty, \quad 1 - A_\infty \geq 1 - B_\infty - C_\infty$$

は言える。つまり、男性のその性質を持つ割合 A_∞ は、女性のその性質を持つ割合 B_∞ よりは確かに多くなる。しかし、その他の大小関係は、α の値によって色々な上下は

ありえて、グラフを書いてみればわかるが、それらは $\alpha = 1/3, 1/2, 2/3$ でそれぞれ大小が変化する。例えば $0 < \alpha < 1/3$ であれば、

$$B_\infty < A_\infty < C_\infty < 1 - B_\infty - C_\infty < 1 - A_\infty$$

のようになる。

元々の話の場合は、α が 0 に近い場合であるが、この場合は B_∞ は 2 次なのでかなり小さくなる。例えば

$$A_\infty = \alpha = 0.02 = 2\%$$

であったとすると、

$$B_\infty = \alpha^2 = 0.0004 = 0.04\%$$

であるから、確かに女性のその性質を持つ割合は、男性に比べてはるかに小さくなる。ただし、この場合は

$$C_\infty = 2\alpha(1-\alpha) = 0.04 \times 0.98 = 0.0392 = 3.92\%$$

となるので、女性のうち潜在的にその因子を持つ割合 C_∞ は、A_∞ の倍位いることになる。

6 最後に

高校生の頃、生物の授業で優性遺伝と劣性遺伝というものを習ったが、その割合の推移の話を聞いたときに、なんとなく納得しづらかったような記憶がある。

上のような数列での解析は、高校生にも可能であるだろうから、今にして思えば、そういう考え方で納得する方法もあったのかもしれない。