

Fluctuation of Population Size and Effective Size of Population

Masaru Iizuka

*Division of Mathematics, Kyushu Dental College
2-6-1 Manazuru, Kokurakita-ku, Kitakyushu 803-8580, Japan*

集団の確率変動と有効個体数

飯塚 勝

九州歯科大学 数学教室

Abstract

The effective size of population has played an important role in population genetics. We consider a Wright-Fisher model whose population size is a simple Markov chain. For this model, we define inbreeding effective size and variance effective size. These effective sizes are turned out to be the same for this model. Effects of fluctuation of population size on the effective size are investigated. The effective size is not the same as the harmonic mean of population size unless fluctuation of population size is uncorrelated. The effective size is larger than the harmonic mean when the fluctuation of population size is positively autocorrelated and smaller than the harmonic mean when the fluctuation is negatively autocorrelated.

1. Effective size of population

In population genetics theory, a traditional formulation of simple stochastic haploid model is the following Wright-Fisher model. Consider a population of N haploid individuals of which i are of type A_1 and $N - i$ are of type A_2 . The population reproduces itself in discrete generations. The whole of the next generation is formed by N independent repetitions of the sampling with replacement. The probability that the next generation will contain j members of type A_1 and $N - j$ of type A_2 is

$$p_{ij} = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}. \quad (1)$$

Let $Z(t)$ be the number of type A_1 in generation t . The process $\{Z(t)\}_{t=0,1,2,\dots}$ is a discrete time Markov chain on $\{0, 1, \dots, N - 1, N\}$ with $p_{ij} = P(Z(t+1) = j | Z(t) = i)$. Let $x(t) = \frac{Z(t)}{N}$ be the gene frequency of type A_1 in generation t in the population. The process $\{x(t)\}_{t=0,1,2,\dots}$ is a discrete time Markov chain on $\{0, \frac{1}{N}, \dots, \frac{N-1}{N}, 1\}$ with the transition matrix $T = (p_{ij})_{i,j=0,1,2,\dots,N}$. This Markov chain is referred to as a Wright-Fisher model which has the following properties (Ethier and Kurtz (1986) and Ewens (1979)). The maximum eigenvalue of the transition matrix T is 1 and the maximum non-unit eigenvalue is $\lambda_2 = 1 - \frac{1}{N}$. The

probability that randomly sampled two genes from the population have the same parent is $\pi_2 = \frac{1}{N}$. The quantity π_2 is referred to as the inbreeding coefficient. The conditional variance of $x(t+1)$ conditional on $x(t) = x$ is $Var[x(t+1)|x(t) = x] = E[\{x(t+1) - x(t)\}^2|x(t) = x] = \frac{x(1-x)}{N}$. The population size N can be expressed as

$$N = \frac{1}{1 - \lambda_2}, \quad (2)$$

$$N = \frac{1}{\pi_2}, \quad (3)$$

and

$$N = \frac{x(1-x)}{Var[x(t+1)|x(t) = x]}. \quad (4)$$

For stochastic models that are more complicated than the Wright-Fisher model, the effective size of population plays an important role for the population size N (Crow (1954), Kimura and Crow (1963) and Wright (1938)). There are several ways to define the effective size of population. The eigenvalue effective size $Ne^{(e)}$ is defined by

$$Ne^{(e)} = \frac{1}{1 - \lambda_2}, \quad (5)$$

where λ_2 is the maximum non-unit eigenvalue of the transition matrix of the stochastic model. The inbreeding effective size $Ne^{(i)}$ is defined by

$$Ne^{(i)} = \frac{1}{\pi_2}, \quad (6)$$

where π_2 is the inbreeding coefficient of the stochastic model. The variance effective size $Ne^{(v)}$ is defined by

$$Ne^{(v)} = \frac{x(1-x)}{Var[x(t+1)|x(t) = x]}, \quad (7)$$

where $x(t)$ is the gene frequency of type A_1 in generation t of the stochastic model. For concrete examples of these effective sizes, see Crow and Kimura (1970), Ewens (1979) and Nagylaki (1992).

2. Effective size of fluctuating population

There is a lot of ecological data to the effect that the numbers of individuals in natural populations fluctuate considerably in each epoch and from generation to generation (see Andrewartha and Birch (1954), Elton and Nicholson (1942) and Odum (1959)). The variations in population size are influenced by such factors as climate, the abundance of available resources, fluctuation in prey-predator balance, competition with other species using the same habitat (Nicholson (1957)).

One of recent interests in theoretical population genetics is to find a mechanism for evaluating overdispersed molecular evolution (Gillespie (1991, 1993, 1994a, 1994b), Iwasa (1993), Ohta and Kimura (1971), Ohta and Tachida (1990), Tachida (1991) and Takahata (1987)). Overdispersed molecular evolution is a phenomena that the ratio of the variance

to the mean number of substitution of mutants among species is larger than one. This ratio is referred to as the dispersion index. Gillespie (1989), using data of 20 protein loci from three species of mammals, obtained 6.95 as an estimate of the dispersion index. Ohta (1995) analyzed 49 mammalian protein data and obtained an estimate of 5.6. Note that the dispersion index is equal to one for the neutral model since the substitution process is a Poisson process (Kimura (1983)). One of the candidates of the mechanism for evaluating the dispersion index being much larger than one is nearly neutral mutation model with fluctuating population size (Araki and Tachida (1997)).

Fluctuation of population size is not independent from generation to generation in general as in the case of stochastic selection (Guess and Gillespie (1977), Iizuka (1987), Iizuka and Matsuda (1982), Seno and Shiga (1984) and Takahata, Ishii and Matsuda (1975)). An appropriate concept for autocorrelated stochastic processes is mixing processes (Billingsley (1968)). In this paper, we will consider a simple case of mixing processes, that is, two-valued Markov chain as a model of autocorrelated fluctuation of population size. Our interest is how the effective size of population depends on the degree of autocorrelation of the fluctuation.

3. Model

Let $N(t) = N(t, \omega_1)$ be the population size in generation t . In this paper, we assume that $\{N(t, \omega_1)\}_{t=0, \pm 1, \pm 2, \dots}$ is a stationary Markov chain on $\{N_1, N_2\}$ such that

$$P_{\omega_1}(N(t+1, \omega_1) \neq N(t, \omega_1) | N(t, \omega_1) = N_j) = q_j, \quad (8)$$

$$P_{\omega_1}(N(t+1, \omega_1) = N(t, \omega_1) | N(t, \omega_1) = N_j) = 1 - q_j, \quad (9)$$

and

$$P_{\omega_1}(N(0, \omega_1) = N_j) = p_j^{(st)}, \quad (10)$$

where $1 < N_1 < N_2 < \infty$, $0 \leq q_j \leq 1$, $q_1 + q_2 > 0$ and

$$(p_1^{(st)}, p_2^{(st)}) = \left(\frac{q_2}{q_1 + q_2}, \frac{q_1}{q_1 + q_2} \right) \quad (11)$$

is the stationary distribution of the Markov chain. The parameter q_j is the probability of changing size for N_j ($j = 1, 2$).

Note that

$$\tau_j = \sum_{k=1}^{\infty} k(1 - q_j)^{k-1} q_j = \frac{1}{q_j} \quad (12)$$

is the mean persistence time for population size N_j and

$$\tau = p_1^{(st)} \tau_1 + p_2^{(st)} \tau_2 = \frac{q_1^2 + q_2^2}{q_1 q_2 (q_1 + q_2)} \quad (13)$$

is the mean persistence time of the Markov chain. The mean of $N(t, \omega_1)$ is

$$E_{\omega_1}[N(t, \omega_1)] = p_1^{(st)} N_1 + p_2^{(st)} N_2 = \frac{q_2 N_1 + q_1 N_2}{q_1 + q_2}. \quad (14)$$

The covariance of $N(t, \omega_1)$ and $N(t+k, \omega_1)$ is

$$\text{Cov}_{\omega_1}[N(t, \omega_1), N(t+k, \omega_1)] = \frac{q_1 q_2 (1 - q_1 - q_2)^k}{(q_1 + q_2)^2} (N_2 - N_1)^2. \quad (15)$$

Eq.(15) means that constant population size, periodic change per generation, positively autocorrelated fluctuation, uncorrelated fluctuation, and negatively autocorrelated fluctuation correspond to $q_1 q_2 = 0$, $q_1 = q_2 = 1$, $0 < q_1 + q_2 < 1$ with $q_1 q_2 \neq 0$, $q_1 + q_2 = 1$ with $q_1 q_2 \neq 0$, and $1 < q_1 + q_2 < 2$, respectively. The case of q_2 is much smaller than q_1 and N_1 is much smaller than N_2 corresponds to a model of population bottleneck since the mean persistence time τ_1 for N_1 is much shorter than τ_2 for N_2 . The mean of $\frac{1}{N(t, \omega_1)}$ is

$$E_{\omega_1}\left[\frac{1}{N(t, \omega_1)}\right] = \frac{1}{N_H}, \quad (16)$$

where

$$N_H = \left\{ \frac{p_1^{(st)}}{N_1} + \frac{p_2^{(st)}}{N_2} \right\}^{-1} = (q_1 + q_2) \left(\frac{q_2}{N_1} + \frac{q_1}{N_2} \right)^{-1} \quad (17)$$

is the harmonic mean of $N(t, \omega_1)$.

For fixed ω_1 , we consider a haploid population with type A_1 and A_2 . The population size in generation t is $N(t, \omega_1)$. The number of type A_1 in generation t which is denoted by $Z(t, \omega_1, \omega_2)$ is a discrete time Markov process with

$$\begin{aligned} P_{\omega_2}(Z(t+1, \omega_1, \omega_2) = j | Z(t, \omega_1, \omega_2) = i) \\ = \binom{N(t+1, \omega_1)}{j} \left(\frac{i}{N(t, \omega_1)} \right)^j \left(1 - \frac{i}{N(t, \omega_1)} \right)^{N(t+1, \omega_1) - j}, \end{aligned} \quad (18)$$

$0 \leq i \leq N(t, \omega_1)$, $0 \leq j \leq N(t+1, \omega_1)$. Let $x(t) = x(t, \omega_1, \omega_2) = \frac{Z(t, \omega_1, \omega_2)}{N(t, \omega_1)}$ be the gene frequency of type A_1 in generation t . The process $\{x(t, \omega_1, \omega_2)\}_{t=0,1,2,\dots}$ is a Wright-Fisher model with variable population size $\{N(t, \omega_1)\}$ for fixed ω_1 . Incorporating stochastic effects by fluctuating population size, this model is referred to as a Wright-Fisher model with fluctuating population size. Note that $x(t, \omega_1, \omega_2)$ is $\sigma(x(t-1, \omega_1, \omega_2), N(t-1, \omega_1), N(t, \omega_1))$ measurable and $N(t, \omega_1)$ is $\sigma(N(s, \omega_1), s \leq t-1)$ measurable where $\sigma(x(t-1, \omega_1, \omega_2), N(t-1, \omega_1), N(t, \omega_1))$ is a σ -field generated by $x(t-1, \omega_1, \omega_2), N(t-1, \omega_1), N(t, \omega_1)$ and $\sigma(N(s, \omega_1), s \leq t-1)$ is a σ -field generated by $N(s, \omega_1), s \leq t-1$.

Because of the autocorrelation of fluctuation of population size, we must extend the definition of effective size, which will be done for inbreeding effective size and variance effective size in the following. Let $\pi_2(t)$ be the probability that randomly sampled two genes from the population in generation t have the same ancestral gene. We have

$$1 - \pi_2(t) = E_{\omega_1} \left[\prod_{k=0}^{t-1} \left(1 - \frac{1}{N(k, \omega_1)} \right) \right] \{1 - \pi_2(0)\}. \quad (19)$$

Since

$$1 - \pi_2(t) = \left(1 - \frac{1}{N} \right)^t \{1 - \pi_2(0)\} \quad (20)$$

in the case of constant population size ($N(t, \omega_1) = N$), the inbreeding effective size $Ne^{(i)} = Ne^{(i)}(q_1, q_2)$ can be defined by

$$1 - \frac{1}{Ne^{(i)}} = \lim_{t \rightarrow \infty} \left\{ E_{\omega_1} \left[\prod_{k=0}^{t-1} \left(1 - \frac{1}{N(k, \omega_1)} \right) \right] \right\}^{\frac{1}{t}}. \quad (21)$$

Let $V_t(x)$ be the conditional variance of $x(t, \omega_1, \omega_2)$ conditional on $x(0, \omega_1, \omega_2) = x$, that is,

$$\begin{aligned} V_t(x) &= \text{Var}_{(\omega_1, \omega_2)} [x(t, \omega_1, \omega_2) | x(0, \omega_1, \omega_2) = x] \\ &= E_{(\omega_1, \omega_2)} [\{x(t, \omega_1, \omega_2) - x(0, \omega_1, \omega_2)\}^2 | x(0, \omega_1, \omega_2) = x]. \end{aligned} \quad (22)$$

Since

$$V_t(x) = \left\{ 1 - \left(1 - \frac{1}{N} \right)^t \right\} x(1-x) \quad (23)$$

in the case of constant population size ($N(t, \omega_1) = N$), the variance effective size $Ne^{(v)} = Ne^{(v)}(q_1, q_2)$ can be defined by

$$1 - \frac{1}{Ne^{(v)}} = \lim_{t \rightarrow \infty} \left\{ 1 - \frac{V_t(x)}{x(1-x)} \right\}^{\frac{1}{t}}. \quad (24)$$

We can show that the inbreeding effective size is the same as the variance effective size for this model.

Lemma 1 For this model,

$$Ne^{(i)} = Ne^{(v)}. \quad (25)$$

Proof. It is enough to show that

$$1 - \frac{1}{Ne^{(v)}} = \lim_{t \rightarrow \infty} \left\{ E_{\omega_1} \left[\prod_{k=0}^{t-1} \left(1 - \frac{1}{N(k, \omega_1)} \right) \right] \right\}^{\frac{1}{t}}. \quad (26)$$

Taking a conditional expectation, we have

$$\begin{aligned} &E_{(\omega_1, \omega_2)} [x(t, \omega_1, \omega_2) \{1 - x(t, \omega_1, \omega_2)\} | x(0, \omega_1, \omega_2) = x] \\ &= E_{(\omega_1, \omega_2)} [E_{\omega_2} [x(t, \omega_1, \omega_2) \{1 - x(t, \omega_1, \omega_2)\} | N(t-1, \omega_1)] | x(0, \omega_1, \omega_2) = x] \\ &= E_{(\omega_1, \omega_2)} \left[\left(1 - \frac{1}{N(t-1, \omega_1)} \right) x(t-1, \omega_1, \omega_2) \{1 - x(t-1, \omega_1, \omega_2)\} | x(0) = x \right]. \end{aligned} \quad (27)$$

Iterating this operation, we have

$$E_{(\omega_1, \omega_2)} [x(t, \omega_1, \omega_2) \{1 - x(t, \omega_1, \omega_2)\} | x(0, \omega_1, \omega_2) = x] = E_{\omega_1} \left[\prod_{k=0}^{t-1} \left(1 - \frac{1}{N(k, \omega_1)} \right) \right] x(1-x). \quad (28)$$

Since $\{x(t, \omega_1, \omega_2)\}$ is a martingale ($E_{(\omega_1, \omega_2)} [x(t, \omega_1, \omega_2) | x(0, \omega_1, \omega_2) = x] = x$),

$$\begin{aligned} &E_{(\omega_1, \omega_2)} [\{x(t, \omega_1, \omega_2) - x(0, \omega_1, \omega_2)\}^2 | x(0, \omega_1, \omega_2) = x] \\ &= -E_{(\omega_1, \omega_2)} [x(t, \omega_1, \omega_2) \{1 - x(t, \omega_1, \omega_2)\} | x(0, \omega_1, \omega_2) = x] + x(1-x) \\ &= \left\{ 1 - E_{\omega_1} \left[\prod_{k=0}^{t-1} \left(1 - \frac{1}{N(k, \omega_1)} \right) \right] \right\} x(1-x). \end{aligned} \quad (29)$$

We have the conclusion by the definition of the variance effective size (Eq.(24)). \square

We will use a notation $Ne = Ne(q_1, q_2)$ for $Ne^{(i)}$ and $Ne^{(v)}$. Our interests are as follows.

Is the effective size Ne equal to the harmonic mean N_H ?

How does the effective size $Ne(q_1, q_2)$ depend on q_1 and q_2 ?

We will consider these problems in the next section. For models with fluctuating population, see Chia and Pollak (1974), Donnelly (1986), Heyde and Seneta (1975), Karlin (1968), Klebaner (1988) and Seneta (1974).

4. Results

First, we will present a concrete expression for the effective size Ne . For this end, we prepare two lemmas.

Lemma 2 For $i = 1, 2$, the conditional expectation

$$B_i(t) = E_{\omega_1} \left[\prod_{k=0}^{t-1} \left(1 - \frac{1}{N(k, \omega_1)} \right) \middle| N(0, \omega_1) = N_i \right] \quad (30)$$

satisfies

$$\begin{aligned} B_i(t+2) - \left\{ \left(1 - \frac{1}{N_1} \right) (1 - q_1) + \left(1 - \frac{1}{N_2} \right) (1 - q_2) \right\} B_i(t+1) \\ + \left(1 - \frac{1}{N_1} \right) \left(1 - \frac{1}{N_2} \right) (1 - q_1 - q_2) B_i(t) = 0. \end{aligned} \quad (31)$$

Proof. For $i, j = 1, 2$ ($i \neq j$), we have

$$\begin{aligned} B_i(t+1) &= E_{\omega_1} \left[\prod_{k=0}^t \left(1 - \frac{1}{N(k, \omega_1)} \right) \middle| N(1, \omega_1) = N_i \right] P(N(1, \omega_1) = N_i | N(0, \omega_1) = N_i) \\ &+ E_{\omega_1} \left[\prod_{k=0}^t \left(1 - \frac{1}{N(k, \omega_1)} \right) \middle| N(1, \omega_1) = N_j \right] P(N(1, \omega_1) = N_j | N(0, \omega_1) = N_i) \\ &= \left(1 - \frac{1}{N_i} \right) \{ (1 - q_i) B_i(t) + q_i B_j(t) \}, \end{aligned} \quad (32)$$

which implies the conclusion. \square

Let α_+ and α_- ($\alpha_+ \geq \alpha_-$) be solutions to

$$\begin{aligned} f(\alpha) &= \alpha^2 - \left\{ \left(1 - \frac{1}{N_1} \right) (1 - q_1) + \left(1 - \frac{1}{N_2} \right) (1 - q_2) \right\} \alpha \\ &+ \left(1 - \frac{1}{N_1} \right) \left(1 - \frac{1}{N_2} \right) (1 - q_1 - q_2) = 0. \end{aligned} \quad (33)$$

Note that α_+ and α_- are real and $\alpha_+ = \alpha_-$ if and only if $1 - \frac{1}{N_1} = \left(1 - \frac{1}{N_2} \right) (1 - q_2)$.

Lemma 3 For $q_1 = 0$ and $1 - \frac{1}{N_1} = (1 - \frac{1}{N_2})(1 - q_2)$,

$$E_{\omega_1}[\prod_{k=0}^{t-1} (1 - \frac{1}{N(k, \omega_1)})] = (1 - \frac{1}{N_1})^t, \quad (34)$$

and

$$E_{\omega_1}[\prod_{k=0}^{t-1} (1 - \frac{1}{N(k, \omega_1)})] = \frac{c_1 q_2 + c_2 q_1}{q_1 + q_2} \alpha_+^{t-1} + \frac{d_1 q_2 + d_2 q_1}{q_1 + q_2} \alpha_-^{t-1} \quad (35)$$

otherwise. Here

$$c_i = \frac{1 - \frac{1}{N_i}}{\alpha_+ - \alpha_-} \{ (1 - \frac{1}{N_i})(1 - q_i) + (1 - \frac{1}{N_j})q_i - \alpha_- \}, \quad (36)$$

and

$$d_i = -\frac{1 - \frac{1}{N_i}}{\alpha_+ - \alpha_-} \{ (1 - \frac{1}{N_i})(1 - q_i) + (1 - \frac{1}{N_j})q_i - \alpha_+ \}, \quad (37)$$

$i, j = 1, 2$ and $i \neq j$.

Proof. Since

$$E_{\omega_1}[\prod_{k=0}^{t-1} (1 - \frac{1}{N(k, \omega_1)})] = B_1(t)p_1^{(st)} + B_2(t)p_2^{(st)}, \quad (38)$$

we have the conclusion by solving the recurrence equation for $B_1(t)$ and $B_2(t)$. \square

The following theorem presents a concrete expression for the effective size.

Theorem 1 For $q_1 q_2 \neq 0$,

$$Ne = \frac{1}{1 - \alpha_+}. \quad (39)$$

For $q_j = 0$,

$$Ne = N_H = N_j. \quad (40)$$

Proof. Note that for real numbers a, b, A and B with $A > |B|$, $a > 0$

$$\lim_{t \rightarrow \infty} (aA^t + bB^t)^{\frac{1}{t}} = A, \quad (41)$$

and for for real numbers c, d and C with $C > 0$, $c > |d|$,

$$\lim_{t \rightarrow \infty} \{cC^t + d(-C)^t\}^{\frac{1}{t}} = C, \quad (42)$$

We have the conclusion by Lemma 3. \square

Next, we consider the relation between the effective size Ne and the harmonic mean N_H . From the sign of $f(1 - \frac{1}{N_1})$, $f(1 - \frac{1}{N_2})$ and $f(1 - \frac{1}{N_H})$, we can obtain the following result (Note that $f(1 - \frac{1}{N_e}) = 0$).

Theorem 2 For positively autocorrelated fluctuation ($0 < q_1 + q_2 < 1$ and $q_1 q_2 \neq 0$),

$$N_1 < N_H < Ne < N_2. \quad (43)$$

For uncorrelated fluctuation ($q_1 + q_2 = 1$ and $q_1 q_2 \neq 0$),

$$N_1 < N_H = Ne < N_2. \quad (44)$$

For negatively autocorrelated fluctuation ($q_1 + q_2 > 1$),

$$N_1 < Ne < N_H < N_2. \quad (45)$$

By this result, the effective size is equal to the harmonic mean if and only if the fluctuation of population size is uncorrelated.

We consider the dependence of q_1 and q_2 on the effective size Ne . Differentiating $\alpha_+ = \alpha_+(q_1, q_2)$ by q_1 and q_2 , we can obtain the following result.

Theorem 3 For fixed q_2 ($q_2 \neq 0$), $Ne = Ne(q_1, q_2)$ is an increasing function of q_1 . For fixed q_1 ($q_1 \neq 0$), $Ne = Ne(q_1, q_2)$ is a decreasing function of q_2 .

In the next theorem, we consider the case where q_2 is proportional to q_1 . We can obtain the following result in the same way as Theorem 3.

Theorem 4 For fixed c ($c > 0$), we set $q_1 = q$ and $q_2 = cq$. For $0 < q < \min\{1, \frac{1}{c}\}$, $Ne = Ne(q)$ is a decreasing function of q for fixed c .

J. H. Gillespie performed computer simulations for the case of $c = 1$ where mutation and selection are incorporated. He found that average heterozygosity (a measure of genetic diversity) is an increasing function of the mean persistence time τ . Since the values of parameters in his computer simulations are restricted, he is interested in whether this is a general phenomena or not (Gillespie:personal communication). Theorem 4 is consistent with this phenomena, since $\tau = \frac{1}{q}$ when $c = 1$ and average heterozygosity is an increasing function of Ne . This means that his finding by computer simulations seems to be a general phenomena. Indeed, this is a motivation of the present paper.

Theorem 4 implies that the weaker the autocorrelation of fluctuation of population size is, the smaller the effective size is. An explanation of this result is as follows. When the autocorrelation is weak, it is difficult to predict what will happen to changes in population size in the next several generation. It may be deleterious to the population. On the other hand, when the effective size is small, the population has little genetic variation which may cause the extinction of the population by a sudden environmental change that has a harmful effect to the majority (wild type) of individuals in the population.

5. Asymptotic relations

The sizes of population N_1 and N_2 may be very large in natural populations. Further, the autocorrelation of $\{N(t, \omega_1)\}$ may be very strong. In such cases, we can obtain asymptotic behavior of the effective size. For this end, we parameterize N_1 , N_2 , Ne , N_H , q_1 and q_2 by ε such as N_1^ε , N_2^ε , Ne^ε , N_H^ε , q_1^ε and q_2^ε ($N_1^\varepsilon < N_2^\varepsilon$).

In the following theorem, we consider the case where N_1^ε and N_2^ε are very large and the ratio of N_2^ε to N_1^ε is finite.

Theorem 5 Assume that $\lim_{\varepsilon \rightarrow 0} N_1^\varepsilon = \infty$ with $\lim_{\varepsilon \rightarrow 0} \frac{N_1^\varepsilon}{N_2^\varepsilon} > 0$, then

$$\lim_{\varepsilon \rightarrow 0} \frac{Ne^\varepsilon}{N_2^\varepsilon} = 1, \quad (46)$$

if $\lim_{\varepsilon \rightarrow 0} N_1^\varepsilon(q_1^\varepsilon + q_2^\varepsilon) = 0$ and

$$\lim_{\varepsilon \rightarrow 0} \frac{Ne^\varepsilon}{N_H^\varepsilon} = 1, \quad (47)$$

if $\lim_{\varepsilon \rightarrow 0} N_1^\varepsilon(q_1^\varepsilon + q_2^\varepsilon) = \infty$.

The following result is more general than Theorem 5, since it is not necessarily assumed that N_1^ε and N_2^ε are very large.

Theorem 6 Assume that $\lim_{\varepsilon \rightarrow 0} N_1^\varepsilon(q_1^\varepsilon + q_2^\varepsilon) = 0$ and $0 < \lim_{\varepsilon \rightarrow 0} \frac{N_1^\varepsilon}{N_2^\varepsilon} < 1$, then

$$\lim_{\varepsilon \rightarrow 0} \frac{Ne^\varepsilon}{N_2^\varepsilon} = 1. \quad (48)$$

Assume that $\lim_{\varepsilon \rightarrow 0} N_2^\varepsilon(q_1^\varepsilon + q_2^\varepsilon) = 0$ and $\lim_{\varepsilon \rightarrow 0} \frac{N_1^\varepsilon}{N_2^\varepsilon} = 0$, then

$$\lim_{\varepsilon \rightarrow 0} \frac{Ne^\varepsilon}{N_2^\varepsilon} = 1. \quad (49)$$

Assume that $\lim_{\varepsilon \rightarrow 0} N_1^\varepsilon(q_1^\varepsilon + q_2^\varepsilon) = 0$ and $0 < \lim_{\varepsilon \rightarrow 0} N_2^\varepsilon(q_1^\varepsilon + q_2^\varepsilon) \leq \infty$, then

$$\lim_{\varepsilon \rightarrow 0} \frac{Ne^\varepsilon}{N_2^\varepsilon} = \frac{1}{1 + \lim_{\varepsilon \rightarrow 0} N_2^\varepsilon q_2^\varepsilon}. \quad (50)$$

Assume that $\lim_{\varepsilon \rightarrow 0} N_1^\varepsilon(q_1^\varepsilon + q_2^\varepsilon) = \infty$, then

$$\lim_{\varepsilon \rightarrow 0} \frac{Ne^\varepsilon}{N_H^\varepsilon} = 1. \quad (51)$$

The autocorrelation of fluctuation of population size can be classified as follows. The case of $\lim_{\varepsilon \rightarrow 0} N_2^\varepsilon(q_1^\varepsilon + q_2^\varepsilon) = 0$ is referred to as strong autocorrelation. The case of $\lim_{\varepsilon \rightarrow 0} N_1^\varepsilon(q_1^\varepsilon + q_2^\varepsilon) = \infty$

is referred to as weak autocorrelation. The other cases are classified as moderate autocorrelation (See Gillespie (1991) for such a classification in stochastic selection models). Theorem 5 and Theorem 6 imply that the effective size is very close to the harmonic mean when the fluctuation has weak autocorrelation. On the other hand, it is very close to the larger population size when the fluctuation has strong autocorrelation.

The author is grateful to J. H. Gillespie for suggesting the problem of this paper. This research was partially supported by a grant-in-aid from the Ministry of Education, Science and Culture of Japan.

References

- Andrewartha, H.G. and L.C. Birch (1954) *The Distribution and Abundance of Animals*. University of Chicago Press, Chicago.
- Araki, H. and H. Tachida (1997) Bottleneck effect on evolutionary rate in nearly neutral mutation model. *Genetics* **147**, 907-914.
- Billingsley, P. (1968) *Convergence of Probability Measures*. Wiley, New York.
- Chia, A.B. and E. Pollak (1974) The inbreeding effective number and the effective number of alleles in a population that varies in size. *Theoretical Population Biology* **6**, 149-172.
- Crow, J.F. (1954) Breeding structure of population. II. Effective population number, p.543-556. in O. Kempthorne [ed.], *Statistics and Mathematics in Biology*. Hafner, New York.
- Crow, J.F. and M. Kimura (1970) *An Introduction to Population Genetics Theory*. Harper & Row, New York.
- Donnelly, P. (1986) A genealogical approach to variable-population-size models in population genetics. *J. Appl. Prob.* **23**, 283-296.
- Elton, C. and M. Nicholson (1942) The ten-year cycle in numbers of the lynx in Canada. *J. Anim. Ecol.* **11**, 215-244.
- Ethier, S.N. and T.G. Kurtz (1986) *Markov Processes: Characterization and Convergence*. Wiley, New York.
- Ewens, W. (1979) *Mathematical Population Genetics*. Springer-Verlag, New York.
- Gillespie, J.H. (1989) Lineage effects and the index of dispersion of molecular evolution. *Mol. Biol. Evol.* **6**, 636-647.
- Gillespie, J.H. (1991) *The Causes of Molecular Evolution*. Oxford University Press, New York.
- Gillespie, J.H. (1993) Substitution processes in molecular evolution. I. Uniform and clustered substitutions in a haploid model. *Genetics* **134**, 971-981.

- Gillespie, J.H. (1994a) Substitution processes in molecular evolution. II. Exchangeable models from population genetics. *Evolution* **48**, 1101-1113.
- Gillespie, J.H. (1994b) Substitution processes in molecular evolution. III. Deleterious alleles. *Genetics* **138**, 943-952.
- Guess, H.A. and J.H. Gillespie (1977) Diffusion approximations to linear stochastic difference equations with stationary coefficients. *J. Appl. Prob.* **14**, 58-74.
- Heyde, C.C. and E. Seneta (1975) The genetic balance between random sampling and random population size. *J. Math. Biol.* **1**, 317-320.
- Iizuka, M. (1987) Weak convergence of a sequence of stochastic difference equations to a stochastic ordinary differential equation. *J. Math. Biol.* **25**, 643-652.
- Iizuka, M. and H. Matsuda (1982) Weak convergence of discrete time non-Markovian processes related to selection models in population genetics. *J. Math. Biol.* **15**, 107-127.
- Iwasa, Y. (1993) Overdispersed molecular evolution in constant environments. *J. Theor. Biol.* **164**, 373-393.
- Karlin, S. (1968) Rates of approach to homozygosity for finite stochastic models with variable population size. *The American Naturalist* **102**, 443-455.
- Kimura, M. (1983) *The Neutral Allele Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
- Kimura, M. and J.F. Crow (1963) The measurement of effective population number. *Evolution* **17**, 278-288.
- Klebaner, P.C. (1988) Conditions for fixation of an allele in the density-dependent Wright-Fisher model. *J. Appl. Prob.* **25**, 247-256.
- Nagylaki, T. (1992) *Introduction to Theoretical Population Genetics*. Springer-Verlag, New York.
- Nicholson, A.J. (1957) The self adjustment of populations to change. *Cold Spring Harbor Symp. Quant. Biol.* **22**, 153-173.
- Odum, E.P. (1959) *Fundamentals of Ecology*. Saunders, Philadelphia.
- Ohta, T. (1995) Synonymous and nonsynonymous substitutions in mammalian genes and the nearly neutral theory. *J. Mol. Evol.* **40**, 56-63.
- Ohta, T. and M. Kimura (1971) On the constancy of the evolutionary rate of cistrons. *J. Mol. Evol.* **1**, 18-25.
- Ohta, T. and H. Tachida (1990) Theoretical study of near neutrality. I. Heterozygosity and rate of mutant substitution. *Genetics* **126**, 219-229.
- Seneta, E. (1974) A note on the balance between random sampling and population size. *Genetics* **77**, 607-610.

- Seno, S. and T. Shiga (1984) Diffusion models of temporally varying selection in population genetics. *Adv. Appl. Prob.* **16**, 260-280.
- Tachida, H. (1991) A study on a nearly neutral mutation model in finite populations. *Genetics* **128**, 183-192.
- Takahata, N. (1987) On the overdispersed molecular clock. *Genetics* **116**, 169-179.
- Takahata, N., K. Ishii and H. Matsuda (1975) Effect of temporal fluctuation of selection coefficient on gene frequency in a population. *Proc. Natl. Acad. Sci. USA* **72**, 4541-4545.
- Wright, S. (1938) Size of population and breeding structure in relation to evolution. *Science* **87**, 430-431.