UDC 519.21

R. W. Coad (Great Britain)

DIFFUSION APPROXIMATION OF THE WRIGHT – FISHER MODEL OF POPULATION GENETICS: SINGLE-LOCUS TWO ALLELE

ДИФУЗІЙНА АПРОКСИМАЦІЯ МОДЕЛІ РАЙТА – ФІШЕРА ПОПУЛЯЦІЙНОЇ ГЕНЕТИКИ: ОДНОЛОКУСНИЙ ДВОАЛЕЛЬНИЙ ВИПАДОК

We investigate an autoregressive diffusion approximation method applied to the Wright – Fisher model in population genetics by considering a Markov chain with Bernoulli distributed independent variables. The use of an autoregressive diffusion method and an averaged allelic frequency process lead to an Ornstein – Uhlenbeck diffusion process with discrete time. The normalized averaged frequency process possesses independent allele frequency indicators with constant conditional variance at equilibrium. In a monoecious diploid population size N with r generations we consider the time to equilibrium of averaged allele frequency in a single-locus two allele pure sampling model.

Досліджується авторегресивна дифузійна апроксимація моделі Райта – Фішера популяційної генетики. Для цього використовується марковський ланцюг з незалежними змінними, підпорядкованими розподілу Бернуллі. Методи авторегресивної дифузії та усереднення алельних частот дозволяють звести проблему, що вивчається, до дифузійного процесу Орнстейна – Уленбека з дискретим часом. Нормований процес усереднених частот має в стані рівноваги незалежні індикатори алельних частот з постійною умовною дисперсією. Встановлюється час, потрібний однодомній деплоїдній популяції розміром N, що складається з r поколінь, для того щоб досягти стану рівноваги усереднених алельних частот в однолокусній двоалельній моделі.

1. Background. Diffusion approximations have been, and still are, a popular method for obtaining values of fixation probabilities, mean absorption or fixation times as well as other values of interest for the study of evolutionary dynamics of genes in finite populations (see [1 – 5]).

The representation of random genetic drift by differential equation was first applied by R. A. Fisher (see [6]), in his implicit version of the model, who noted that the equation describing the diffusion of heat through a solid bar applies to random genetic drift. This pattern of change in allele frequency is a good approximation to that expected theoretically for an ideal population and iterations of the discrete Wright – Fisher model using a Markov chain give the expected outcomes of a pure drift process [7 – 10].

Kimura [11] used the recurrence relation $p' = p + (2U - 1)\sqrt{3pq/2N}$, where U is a random variable uniformly distributed between 0 and 1, to demonstrate the diffusion approximation. Each generation U is reselected and a sample realization of the allelic frequency of the next generation is obtained. The variance of the uniform random variable in the recurrence relation is found to be $pq/2N$, the same as that for a binomial sampling distribution. Kimura concluded that even though the distribution of change in allele frequency is uniform, the process reproduces the desired results for fewer iterations. The best understood case is that of the neutral genes first argued by Kimura, those genes which have no advantageous or deleterious effect in the population, an argument which has had an crucial impact on evolutionary theory [12, 13]. We however consider only the existence of the non-neutral cases. Burger and Ewens [14] demonstrated that for advantageous alleles, for a large population N, the diffusion approximation is more accurate than for deleterious alleles. Other relevant works on diffusion approximations include Ethier and Nagylaki [15] and Shiga [16]. It is useful to refer also to the diffusion approximations of non-Markovian models of Watterson (1962) and Norman (1975).

The mathematical theory describing gene frequency under selection in finite populations is much less developed than in the neutral case, but it has been shown that for advantageous (as for neutral) genes the well-known diffusion approximation for the fixation probability and the fixation absorption time are quite accurate (see [17 – 20]). In this paper we use an autoregressive diffusion approximation method and an averaged allelic frequency process in discrete time.

2. Introduction. The Wright – Fisher model of population genetics in a monoecious diploid population can be formalized, after mutation, as a stochastic system with states denoted by

$$\eta_r^{(N)} := \frac{1}{N}\sum_{i=1}^{N}\xi_i(r), \quad r \geq 0, \tag{1}$$

where $\eta_r^{(N)}$ is the allelic frequency for one fixed allele after r generations with subpopulation size N and $\xi_i(r)$ is the indicator of the fixed allele for the i-th individual after r generations. On a single-locus two allele pure sampling model, the corresponding Markov chain fixing the frequency of the first allele can be defined as (1) with Bernoulli distributed independent random variables. These random variables take two values such that

$$\Pr\{\xi_i(r+1)=1 \mid \eta_r^{(N)}=p\} = 1 - \Pr\{\xi_i(r+1)=0 \mid \eta_r^{(N)}=p\} = C(p), \tag{2}$$

where the regression function of the Markov chain in the model is

$$C(p) = E[\eta_{r+1}^{(N)} \mid \eta_r^{(N)}=p]. \tag{3}$$

The Wright – Fisher model is determined by the following form of the regression function

$$C(p) = p(w_1 p + q)/w(p),$$

and

$$w(p) = w_1 p^2 + 2pq + w_2 q^2, \quad q := 1-p, \tag{4}$$

are the genotype frequencies with a selective influence in random mating. In this paper, we investigate the diffusion approximation in an attractive case (see [21 – 23]):

$$0 < w_k < 1, \quad k = 1, 2,$$

where w_k are the viability coefficients of $A_k A_k$ allele individuals and the regression function (4) can be transformed into the following expression

$$C(p) = p + C_0(p)/w(p), \tag{5}$$

where

$$C_0(p) = \nu p(1-p)(\rho - p) \tag{6}$$

such that $\nu_k := 1 - w_k$ and $\nu = \nu_1 + \nu_2$ are the selection coefficients*. The cubic parabola (6) has three real roots $p_0 = 0, p_1 = 1$ and, at the equilibrium point, $\rho = \nu_2/\nu, 0 < \rho < 1$, such that

$$C_0'(0) > 0, \quad C_0'(1) > 0 \quad \text{and} \quad C_0'(\rho) < 0.$$

Moreover, by using the definitions of the regression function (4) to (6) we obtain the following important linkage representation

$$C(p) - \rho = (p - \rho)[1 - \nu(p)], \tag{7}$$

where

$$\nu(p) := \nu p(1-p)/w(p) \tag{8}$$

* Selection coefficient values less than zero i.e $\nu_1, \nu_2 < 0$ describe advantageous alleles and values greater than zero describe deleterious alleles in the population. The neutral case $\nu_1, \nu_2 = 0$ is excluded here.

is the recombination function such that

$$0 < b := C'(\rho) = 1 - \nu(p) = 1 - \nu_0 < 1, \tag{9}$$

where $\nu(p) = \Delta/(1-\Delta) =: \nu_0,\ 0 < \nu_0 < 1$ and b, the drift coefficient, is a fixed value depending on the viability coefficients.

From equation (5), the genotype frequency $w(p)$ has the following representation

$$w(p) = 1 - \nu_1 p^2 - \nu_2 q^2, \tag{10}$$

which leads us to the inequality

$$0 \le \nu(p) \le 1.$$

Furthermore, by considering the continuity property of $w(p)$, and for small enough $\delta > 0$, there exist the stronger inequalities

$$0 < \nu(p) - c \le \nu(p) \le \nu(p) + c < 1,$$
$$c := \max_{|p-\rho|<\delta} |\nu(p) - \nu(\rho)|; \tag{11}$$

where $\nu(p) + c = b$. Finally, by considering the recurrence relation

$$\rho_{r+1} = C(\rho_r), \quad r \ge 0, \tag{12}$$

for $\rho_0 \in [0, 1]$, also fixed, then the unique equilibrium point of the recurrence relation (12) is $\rho : \rho = C(\rho)$.

3. Theorem 1. *The notation and definition for the normalized process of allele frequency is introduced*

$$\zeta_r^{(N)} := \sqrt{N}\left(\eta_r^{(N)} - \rho\right) \quad for \quad r \ge r_0. \tag{13}$$

The following convergence of the normalized process takes place

$$\zeta_{r+r_0}^{(N)} \Rightarrow \zeta_r \quad for \quad r \ge 0,\ N \to \infty, \tag{14}$$

where r_0 is sufficiently large and the limiting process ζ_r, $r \ge 0$, is determined by the following autoregressive diffusion approximation

$$\zeta_{r+1} = b\zeta_r + \sigma\omega_{r+1}, \quad r \ge 0, \tag{15}$$

such that ω_r, $r \ge 1$ is a sequence of independent and identically distributed random variables and $\sigma^2 = B(\rho)$. Furthermore the initial condition for ζ_0 is defined as follows

$$\zeta_{r_0}^{(N)} \Rightarrow \zeta_0, \quad N \to \infty.$$

***Remark* 1.** The autoregressive diffusion approximation (13) may be extended on the process

$$\zeta_{r+R}^{(N)} = \sqrt{N}\left(\eta_{r+R}^{(N)} - \rho\right), \quad R \ge 0,$$

for $r \to \infty$ such that

$$E_\pi \eta_r^{(N)} = \rho + \frac{c}{\sqrt{N}}.$$

Proof of Theorem* 1.** ***3.1. Conditional variance. Notations and definitions are introduced for the conditional variance between the Markov chain $\eta_r^{(N)}$ and the equilibrium point ρ using the function $C(p)$ with variance denoted as $\psi_r^{(N)}$ and variance of the frequency indicator in the i-th individual in the r-th generation, $\xi_i(r)$, where

$$\psi_r^{(N)} := \left[\eta_r^{(N)} - \rho\right]^2, \quad r \geq 0; \tag{16}$$

and

$$B(p) := C(p)(1 - C(p)) \leq 1/4. \tag{17}$$

Then we obtain the following definition for the conditional variance using (16),

$$E\left[\psi_{r+1}^{(N)} \mid \eta_r^{(N)}\right] = E\left[\eta_{r+1}^{(N)}\right]^2 - 2\rho C\left(\eta_r^{(N)}\right) + \rho^2 \tag{18}$$

such that

$$E\left[\eta_{r+1}^{(N)}\right]^2 = E\left[\xi_i^2(r+1) \mid \eta_r^{(N)}\right]/N + \left(E\left[\xi_i(r+1) \mid \eta_r^{(N)}\right]\right)^2 N(N-1)/N^2 =$$

$$= C\left(\eta_r^{(N)}\right)/N + \left(1 - \frac{1}{N}\right)C^2\left(\eta_r^{(N)}\right) = B\left(\eta_r^{(N)}\right)/N + C^2\left(\eta_r^{(N)}\right),$$

by (17), and since $E\left[\xi_i(r+1) \mid \eta_r^{(N)}\right] := C\left(\eta_r^{(N)}\right)$. Relation (18) can now be represented as

$$E\left[\psi_{r+1}^{(N)} \mid \eta_r^{(N)}\right] = B\left(\eta_r^{(N)}\right)/N + C^2\left(\eta_r^{(N)}\right) - 2\rho C\left(\eta_r^{(N)}\right) + \rho^2 =$$

$$= \left[C\left(\eta_r^{(N)}\right) - \rho\right]^2 + B\left(\eta_r^{(N)}\right)/N. \tag{19}$$

Moreover, by considering equations (7) and (19) together, for an attractive case we obtain the following definitions for the conditional variance:

$$E\left[\psi_{r+1}^{(N)} \mid \eta_r^{(N)}\right] = \psi_r^{(N)}\left(1 - \nu\left(\eta_r^{(N)}\right)\right)^2 + B\left(\eta_r^{(N)}\right)/N \tag{20}$$

or alternatively

$$E\left[\psi_{r+1}^{(N)} \mid \eta_r^{(N)}\right] = \psi_r^{(N)} - \left[\phi\left(\eta_r^{(N)}\right) - B\left(\eta_r^{(N)}\right)/N\right], \tag{21}$$

where

$$\phi(p) = (p - \rho)\nu(p)(2 - \nu(p)) \quad \forall\, 0 \leq p \leq 1. \tag{22}$$

3.2. Estimation of variance convergence. 3.2.1. **Lemma 1.**

$$\Pr\left\{\left|\eta_r^{(N)} - \rho\right| > \delta\right\} \leq 1/9N\delta^2, \quad N, r \to \infty,$$

for a sufficiently large subpopulation N and for a sufficient number of generations r.

***Proof of Lemma* 1.** There exists sufficiently large r and N such that the following inequality is valid:

$$\left|\rho_r^{(N)} - \rho\right| \leq \frac{\delta}{2},$$

Hence we can use the following inequality for N and r such that

$$\Pr\left\{\left|\eta_r^{(N)} - \rho\right| > \delta\right\} \leq \Pr\left\{\left|\eta_r^{(N)} - \rho\right| > \delta/2\right\}$$

and by Chebyshev's inequality we obtain the following:

$$\Pr\left\{\left|\eta_r^{(N)} - \rho_r^{(N)}\right| > 3\delta/2\right\} \leq E_\pi B\left(\eta_{r-1}^{(N)}\right)/9N\delta^2 \leq 1/9N\delta^2. \tag{23}$$

Proposition. *The conditional variance*

$$E\left[\psi_{r+1}^{(N)} \mid \eta_r^{(N)}\right] = \psi_r^{(N)} - \left[\phi\left(\eta_r^{(N)}\right) - B\left(\eta_r^{(N)}\right)/N\right], \quad r \geq r_0,$$

and for large enough N is a supermartingale.

Proof. By considering the property of a supermartingale (see [8, 24] such that

$$E\left[\psi_{r+1}^{(N)} \mid \eta_r^{(N)}\right] = \psi_r^{(N)} + E\left[\eta_{r+1}^{(N)}\right]^2 \le \psi_r^{(N)}$$

we require that

$$\phi\left(\eta_r^{(N)}\right) \ge B\left(\eta_r^{(N)}\right)/N$$

and by using the representations for $B\left(\eta_r^{(N)}\right)/N$ from (17), and $\phi\left(\eta_r^{(N)}\right)$ from (22), we have

$$\psi_r^{(N)} \nu\left(\eta_r^{(N)}\right)\left(2-\nu\left(\eta_r^{(N)}\right)\right) \ge C\left(\eta_r^{(N)}\right)\left(1-C\left(\eta_r^{(N)}\right)\right)/N$$

and it follows that

$$\lim_{N, r\to\infty} \psi_r^{(N)} \ge (1-\Delta)^2 / \nu N(2-3\Delta), \quad r \ge r_0,$$

and sufficiently large N. By further considering that $\psi_r^{(N)} \le 1$ and $\Delta \le 1/2$ for N, $r \ge 1$ then the following inequality

$$\nu \ge 1/2N, \quad N \ge 1,$$

holds. Hence the conditional mean square value is a supermartingale.

3.3. **Lemma 2.** *For an attractive equilibrium point* ρ, $0 < \rho < 1$, *the expected variance* $\Delta_r^{(N)} := E_\pi\left[\eta_r^{(N)} - \rho\right]^2 = E_\pi \psi_r^{(N)}$ *satisfies the following inequality:*

$$\Delta_{r+1}^{(N)} \le \tilde{B}/N + b^2 \Delta_r^{(N)}, \quad r \ge 0, \tag{24}$$

where $0 < b < 1$ *depends on selection coefficients and* $\tilde{B} = B_1 + B_2$.

Proof. By considering the variance convergence with stationary distribution of frequency indicators in a subpopulation N, we have

$$E_\pi \psi_r^{(N)} = E_\pi\left[\eta_r^{(N)} - E\left[\eta_r^{(N)} \mid \eta_{r-1}^{(N)}\right]\right]^2 + E_\pi\left[E\left[\eta_r^{(N)} \mid \eta_{r-1}^{(N)}\right] - \rho\right]^2 \tag{25}$$

and by considering the definition of the regression function $C(p)$ in (3), we obtain the following indicators for averaged frequency

$$I_r^{(N)} := E_\pi\left[\eta_r^{(N)} - E\left[\eta_r^{(N)} \mid \eta_{r-1}^{(N)}\right]\right]^2 = E_\pi\left[\eta_r^{(N)} - C\left(\eta_{r-1}^{(N)}\right)\right]^2 \tag{26}$$

and by definition of ρ, we have

$$II_r^{(N)} := E_\pi\left[E\left[\eta_r^{(N)} \mid \eta_{r-1}^{(N)}\right] - \rho\right]^2 = E_\pi\left[C\left(\eta_{r-1}^{(N)}\right) - \rho\right]^2. \tag{27}$$

The variance (or mean square distance between the averaged frequency and frequency equilibrium in a subpopulation N) $\Delta_r^{(N)}$ can be represented as follows, using (26) and (27),

$$\Delta_{r+1}^{(N)} = E_\pi\left[\eta_r^{(N)} - C\left(\eta_{r-1}^{(N)}\right)\right]^2 + E_\pi\left[C\left(\eta_{r-1}^{(N)}\right) - \rho\right]^2 \tag{28}$$

and by definition of the variance, in (17), we obtain for the first term in (28), an upper bound,

$$I_r^{(N)} := E_\pi B\left(\eta_{r-1}^{(N)}\right)/N \le 1/4N\delta^2 \tag{29}$$

from (26). Now we consider the second term in (28) by using the representation of the regression function from (7). Hence

$$\Pi_r^{(N)} := E\left[\eta_{r-1}^{(N)} - \rho\right]^2 \left(1 - \nu\left(\eta_{r-1}^{(N)}\right)\right)^2,$$

where

$$\dot{\Pi}_r^{(N)} := I\left(\left|\eta_r^{(N)} - \rho\right| \le \delta\right)$$

and

$$\ddot{\Pi}_r^{(N)} := I\left(\left|\eta_r^{(N)} - \rho\right| > \delta\right)$$

are the internal and external averaged frequency indicators for

$$E\left[1 - \nu\left(\eta_r^{(N)}\right)\right]^2 := E\left(1 - \nu\left(\eta_r^{(N)}\right)\right)^2 \left[\dot{\Pi}_r^{(N)} + \ddot{\Pi}_r^{(N)}\right]. \tag{30}$$

By using the inequality from (9) we can estimate the first term in (30) to be

$$E\left(1 - \nu\left(\eta_r^{(N)}\right)\right)^2 I\left(\left|\eta_r^{(N)} - \rho\right| \le \delta\right) \le b^2 < 1 \tag{31}$$

and for the second term we get

$$E\left(1 - \nu\left(\eta_r^{(N)}\right)\right)^2 I\left(\left|\eta_r^{(N)} - \rho\right| > \delta\right) \le \Pr\left\{\left|\eta_r^{(N)} - \rho\right| > \delta\right\}.$$

Hence, by recalling inequality (23) for the second term $\ddot{\Pi}_r^{(N)}$, in (30) we get

$$E_\pi\left[1 - \nu\left(\eta_r^{(N)}\right)\right]^2 I\left(\left|\eta_r^{(N)} - \rho\right| > \delta\right) \le 1/9N\delta^2. \tag{32}$$

By then uniting the inequalities (31) and (32) we obtain,

$$\Pi_r^{(N)} \le 1/9N\delta^2 + b^2 E_\pi \psi_r^{(N)} \tag{33}$$

and by substituting (29) and (33) into (28) we have

$$\Delta_{r+1}^{(N)} \le 1/4N\delta^2 + 1/9N\delta^2 + b^2 \Delta_r^{(N)}, \tag{34}$$

where $\Delta_r^{(N)} := E_\pi \psi_r^{(N)}$, or equivalently,

$$\Delta_{r+1}^{(N)} \le \tilde{B}/N + b^2 \Delta_r^{(N)}, \quad \tilde{B} = 1/4\delta^2 + 1/9\delta^2. \tag{35}$$

By induction from (35), we obtain the following condition

$$\Delta_{r+1}^{(N)} \le \frac{\tilde{B}}{(1-b)N} + \frac{b^{2r}}{\left(1 - b^{2r}\right)} \Delta_0^{(N)},$$

where $\Delta_0^{(N)} := E_\pi \psi_0^{(N)*}$. Note that in the case $E_\pi \eta_0^{(N)} \ne \rho$ we use the fact that $b \le$ ≤ 1, hence $b^{2r} \to 0$ as $r \to \infty$ such that the variance convergence

$$\lim_{N \to \infty} E_\pi \psi_r^{(N)} = 0 \quad \forall \;\; r, N \to \infty,$$

follows from Lemma 2.

3.4. **Lemma 3.** *The limit process ζ_t, $t \ge 0$, is a diffusion process of Ornstein – Uhlenbeck type with discrete time and has the following representation*

$$W_t = \sum_{r=1}^{t} \omega_r, \quad t \ge 1, \tag{36}$$

where ω_r are independent normally distributed random variables with variance

---

* In the trivial case $E_\pi \eta_0^{(N)} = \rho$, the following initial condition is obtained $\Delta_0^{(N)} := \rho(1-\rho)/N =$ $= \Delta / N\nu$ and since $0 < \Delta < 1/2$, we can say that $\Delta_0^{(N)} < 1/2N\nu$.

$$E\omega_r^2 = \sigma^2 = B(\rho).$$

Proof of Lemma **3.** By the central limit theorem for normalized sums of frequency indicators in a subpopulation N after r generations the following weak convergence takes place:

$$\omega_r^{(N)} := \zeta_{r+1}^{(N)} - \zeta_r^{(N)} \Rightarrow \omega_r, \quad 1 < r < t \quad \text{as} \quad N \to \infty.$$

We may now calculate the standard deviation and variance of $\omega_r^{(N)}$ as follows

$$E_\pi \omega_r^{(N)} = E_\pi\left[\zeta_{r+1}^{(N)} - b\zeta_r^{(N)}\right] = E_\pi\left[E\left[\zeta_{r+1}^{(N)} \mid \zeta_r^{(N)}\right] - b\zeta_r^{(N)}\right]. \tag{37}$$

By expectation (37) we obtain

$$E_\pi \omega_r^{(N)} = E_\pi\left[\zeta_r^{(N)}\left(\nu(\rho) - \nu\left(\eta_r^{(N)}\right)\right)\right] \to 0, \quad N \to \infty, \tag{38}$$

by the continuity property and hence

$$E_\pi \omega_r = 0, \quad r \geq 1.$$

By using equation (38) the expected variance of $\omega_r^{(N)}$ may be calculated

$$E_\pi\left(\omega_r^{(N)}\right)^2 = E_\pi\left[\zeta_{r+1}^{(N)} - b\zeta_r^{(N)}\right]^2 \tag{39}$$

and by again considering (37) then this equals

$$E_\pi\left[\zeta_{r+1}^{(N)} - E\left[\zeta_{r+1}^{(N)} \mid \zeta_r^{(N)}\right] + \zeta_r^{(N)}\left(\nu(\rho) - \nu\left(\eta_r^{(N)}\right)\right)\right]^2 =$$

$$= E_\pi\left[\zeta_{r+1}^{(N)} - E\left[\zeta_{r+1}^{(N)} \mid \zeta_r^{(N)}\right]\right]^2 + E_\pi\left[\zeta_r^{(N)}\left(\nu(\rho) - \nu\left(\eta_r^{(N)}\right)\right)\right]^2.$$

Finally, by recalling the definition of conditional variance we may transform (39) into the following form

$$E_\pi\left(\omega_r^{(N)}\right)^2 = B\left(\eta_r^{(N)}\right) + E_\pi\left[\zeta_r^{(N)}\left(\nu(\rho) - \nu\left(\eta_r^{(N)}\right)\right)\right]^2$$

such that the second term on the RHS tends to zero by the continuity property. Therefore we can conclude that

$$\lim_{N \to \infty} E_\pi\left(\omega_r^{(N)}\right)^2 = B(\rho) = \sigma^2 \tag{40}$$

and that

$$E_\pi \omega_r^2 = \Delta / \nu.$$

By limit (40), the martingale limit has variance equal to the sum of individual variances $\omega_r^2, r \geq 1$. Hence, we may conclude, that the random variables ω_r, are independent (see [7, 8, 24]). This completes the proof of Theorem 1.

4. Process of averaged frequency.

4.1. **Theorem 2.** *The regression function $C(p)$ has the property of a covergent averaged frequency such that*

$$\sum_{r=0}^{Nt} \eta_r^{(N)} / N - \rho t \Rightarrow 0, \quad N \to \infty, \tag{41}$$

takes place by probability.

Martingale representation of the process of averaged frequency. By considering the sum and square characteristics of the martingale-differences we may demonstrate the process in which the averaged allele frequencies in the subpopulation size N converge to an equilibrium.

The convergent averaged frequency is derived as follows:

$$\mu_t^{(N)} := \sum_{r=0}^{Nt-1}\left[\eta_{r+1}^{(N)} - E\left[\eta_{r+1}^{(N)} \mid \eta_r^{(N)}\right]\right]/N = \sum_{r=0}^{Nt-1}\left[\eta_{r+1}^{(N)} - C\left(\eta_r^{(N)}\right)\right]/N \quad (42)$$

by definition of the martingale-differences (see [8, 25]). The square characteristic of the martingale in (42) is as follows:

$$\langle\mu^{(N)}\rangle_t = \sum_{r=0}^{Nt-1} E\left[\left[\eta_{r+1}^{(N)} - C\left(\eta_r^{(N)}\right)\right]^2 \mid \eta_r^{(N)}\right]/N^2 = \sum_{r=0}^{Nt-1} B\left(\eta_r^{(N)}\right)/N^2. \quad (43)$$

There then exists the following asymptotic estimation from the continuity property such that

$$\langle\mu^{(N)}\rangle_t \le B(\rho)t/N = \Delta t/\nu N \to 0, \quad N\to\infty,$$

hence by the limit theorem for martingales [24] we obtain

$$\mu_t^{(N)} \Rightarrow 0 \quad \text{as} \quad N\to\infty \quad (44)$$

with probability one. The martingale in (42) can now be transformed in the following way:

$$\mu_t^{(N)} = \sum_{r=0}^{Nt-1}\eta_{r+1}^{(N)}/N - \sum_{r=0}^{Nt-1} C\left(\eta_r^{(N)}\right)/N =$$

$$= \sum_{r=1}^{Nt-1}\eta_r^{(N)}/N - \rho t - \sum_{r=0}^{Nt-1}\left[C\left(\eta_r^{(N)}\right) - \rho\right]/N. \quad (45)$$

Consider now (7) and inequality (9) by substituting we have the following representation

$$C\left(\eta_r^{(N)}\right) - \rho = b\left[\eta_r^{(N)} - \rho\right] + \left(v(\rho) - v\left(\eta_r^{(N)}\right)\right)\left[\eta_r^{(N)} - \rho\right]. \quad (46)$$

By substituting (46) into (45) we obtain the following form

$$\mu_t^{(N)} = (1-b)\left(\sum_{r=1}^{Nt-1}\eta_r^{(N)}/N - \rho t\right) -$$

$$-\frac{1}{N}\sum_{r=1}^{Nt-1}\left(v(\rho) - v\left(\eta_r^{(N)}\right)\right)\left[\eta_r^{(N)} - \rho\right] - \frac{b}{N}\left(\eta_{Nt}^{(N)} - \eta_0^{(N)}\right).$$

***Proof of Theorem* 2.** By the continuity property of the regression function $C(p)$ and (9) we obtain the following representation for the process of averaged frequency $\mu_t^{(N)}$,

$$\mu_t^{(N)} = (1-b)\left(\sum_{r=1}^{Nt-1}\eta_r^{(N)}/N - \rho t\right) + \frac{b}{N}\left(\eta_0^{(N)} - \eta_{Nt}^{(N)}\right) \quad (47)$$

hence

$$\mu_t^{(N)} = v_0\sum_{r=1}^{Nt-1}\eta_r^{(N)}/N - \rho t \Rightarrow 0, \quad N\to\infty,$$

by probability.

By recalling the normalized process $\zeta_r^{(N)}$ for the process of averaged frequency such that

$$\mu_t^{(N)} := \sum_{r=0}^{t-1}\left[\zeta_{r+1}^{(N)} - E\left[\zeta_{r+1}^{(N)} \mid \eta_r^{(N)}\right]\right], \tag{48}$$

where

$$E\left[\zeta_{r+1}^{(N)} \mid \eta_r^{(N)}\right] = \sqrt{N}\left(b\left[E\left[\eta_{r+1}^{(N)} \mid \eta_r^{(N)}\right] - \rho\right]\right) = \sqrt{N}\left[C\left(\eta_r^{(N)}\right) - \rho\right]. \tag{49}$$

Then by (46) and then substituting into (49) we have the following representation

$$E\left[\zeta_{r+1}^{(N)} \mid \eta_r^{(N)}\right] = \sqrt{N}\left(b\left[\eta_r^{(N)} - \rho\right] + \left(\nu(\rho) - \nu\left(\eta_r^{(N)}\right)\right)\left[\eta_r^{(N)} - \rho\right]\right). \tag{50}$$

Then by substituting (50) into equation (48) we obtain

$$\mu_t^{(N)} = \sum_{r=0}^{t-1}\left[\zeta_{r+1}^{(N)} - \sqrt{N}\left(b\left[\eta_r^{(N)} - \rho\right]\right)\right] + \alpha_t^{(N)}, \tag{51}$$

where

$$\alpha_t^{(N)} = \sum_{r=0}^{t-1}\zeta_r^{(N)}\left(\nu(\rho) - \nu\left(\eta_r^{(N)}\right)\right),$$

and we have the following

$$\mu_t^{(N)} = \sum_{r=0}^{t-1}\left[\zeta_{r+1}^{(N)} - b\zeta_r^{(N)}\right] + \alpha_t^{(N)} \tag{52}$$

and by the continuity property $\alpha_t^{(N)} \to 0, N, r \to \infty$, so (52) reduces to

$$\mu_t^{(N)} = \sum_{r=0}^{t-1}\left[\zeta_{r+1}^{(N)} - b\zeta_r^{(N)}\right].$$

The square characteristic of martingale (52) according to (17) and (49) has the following representation

$$\langle \mu^{(N)} \rangle_t = \sum_{r=0}^{t-1} B\left(\eta_r^{(N)}\right)$$

and by the continuity property we have

$$\langle \mu^{(N)} \rangle_t \overset{N\to\infty}{\Rightarrow} tB(\rho) = \Delta t / \nu.$$

Then by the limit theorem for martingales in series scheme [23] we conclude that the following limit by convergence takes place

$$W_r^{(N)} := \sum_{r=0}^{t-1}\left[\zeta_{r+1}^{(N)} - b\zeta_r^{(N)}\right] \Rightarrow W_t, \quad N \to \infty,$$

where W_t is a limiting process of normalized averaged frequency with independent increments. This completes the proof of Theorem 2.

***Remark* 2.** The diffusion approximation in discrete form considered here may be investigated with multidimensional stochastic Markov models with persistent regression.

Acknowledgements. To Professor V. S. Korolyuk for his consideration and criticism.

1. *Feller W.* Diffusion processes in genetics // Princeton Univ., Second Berkeley Symposium, 1968.
2. *Kimura M.* Diffusion models in population genetics. – London: Methuen, 1964.

3. *Kimura M., Ohta T.* Theoretical aspects of population genetics. – Princeton, NJ, PUP, 1971.
4. *Crow J.F., Kimura M.* An introduction to population genetics theory. – New York: Harper Row, 1970.
5. *Wright S.* Evolution and the genetics of populations. Vol. 2: The theory of gene frequencies. – Chicago: CUP, 1969.
6. *Fisher R.A.* On the dominance ratio // Proc. Roy. Soc. Edinburgh. – 1922. – **52**. – P. 321 – 431.
7. *Doob J.L.* Stochastic processes. – New York: Wiley, 1953.
8. *Grimmett G.R., Stirzaker D.R.* Probability and random processes. – Oxford: Oxford Sci. Publ., 1995.
9. *Kimura M.* Solution of a process of random genetic drift with a continuous model // Proc. Nat. Acad. Sci. USA. – 1955. – **41**. – P. 144 – 150.
10. *Kimura M.* Population genetics and molecular evolution // Johns Hopkins Med. J. – 1976. – **138**. – P. 253 – 261.
11. *Kimura M.* Average time until fixation of a mutant allele in a finite population under continued mutation pressure: studies by analytical, numerical and pseudosampling methods // Proc. Nat. Acad. Sci. USA. – 1980. – **77**. – P. 522 – 526.
12. *Kimura M.* The neutral theory of molecular evolution. – CUP, 1983.
13. *Ewens W. J.* Mathematical population genetics. – Berlin etc.: Springer, 1979.
14. *Burger R.,Ewens W. J.* Fixation probabilities of additive alleles in diploid populations // J. Math. Biol. – 1995. – **55**. – P. 557 – 575.
15. *Ethier S.N., Nagylaki T.* Diffusion approximations of Markov chains with two time-scales and applications to population genetics // Adv. Appl. Probab. – 1980. – **12**. – P. 14 – 49.
16. *Shiga T.* Diffusion processes in population genetics // J. Math. Kyoto Univ. – 1981. – **21**. – P. 133 – 151.
17. *Ewens W.J.* Numerical results and diffusion approximations in a genetic process // Biometrika. – 1963. – **50**. – P. 241 – 249.
18. *Ewens W.J.* The pseudo-transient distribution and its uses in genetics // J. Appl. Probab. – 1964. – **1**. – P. 141 – 156.
19. *Ewens W.J.* Correcting diffusion approximations in finite genetic models // Technical report 4. – Stanford Univ., 1964.
20. *Kimura M., Ohta T.* The average number of generations until fixation of a mutant gene in a finite population // Genetics. – 1069. – **61**. – P. 763 – 771.
21. *Ethier S.N., Kurtz T.G.* Markov processes: characterization and convergence. – New York: John Wiley and Sons, 1986.
22. *Korolyuk V.S., Coad R.W.* Equilibrium points with persistent regression: Wright – Fisher model of population genetics on a diallelic locus. – Kyiv: Inst. Math. Nat. Acad. Sci. Ukraine, 1997.
23. *Korolyuk V.S., Korolyuk D.* Diffusion approximation of stochastic models with persistent regression // Ukr. Math. J. – 1995.– **48**, № 7. – P. 928 – 930.
24. *Hall P., Heyde C.C.* Martingale limit theory and its applications // Ann. Probab. – 1980. – 308 p.
25. *Fukushima M., Stroock D.* Reversibility of solutions to martingale problems. Probability, statistical mechanics and number theory // Adv. Math. Suppl. Stud. – 1986. – **9**. – P. 107 – 123.

Received 05.03.99