

Variability in centred house-of-cards mutation models

SMAIL MAHDI†

Department of Computer Science, Mathematics and Physics, University of the West Indies, Cave Hill Campus, Barbados, West Indies

AND

SABIN LESSARD

Department of Mathematics and Statistics, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec, H3C 3J7, Canada

[Received 17 November 1998 and in revised form 15 November 1999]

Convergence of variability in phenotypic models with balance between selection and mutation is analysed. The mutation assumed occurs with weak probability and brings down the evolutionary process built up by selection around the mean in the population. Gaussian approximations are used.

Keywords: house-of-cards mutation; convergence of variability; Gaussian approximation.

1. Introduction

Mutation-selection balance for quantitative characters has received a lot of attention since it can explain the maintenance of variability in natural populations. In the classical model, the ‘random-walk’ mutation model, mutation is assumed to cause a random additive change (increase or decrease) on the character of the mutant individual. Another kind of model has been proposed, the ‘house-of-cards’ mutation model, in which mutation brings down the structure built up by evolution so that the character of the mutant individual takes a random value completely independent of the previous value.

Exact dynamical results have been obtained for haploid (asexual) or phenotypic models (see, e.g., Slatkin, 1970; Roughgarden, 1972; Karlin, 1979, for Gaussian models with random-walk mutation, Eshel, 1971; Karlin, 1988, for non-Gaussian models with random-walk mutation, Kingman, 1978, for non-Gaussian models with house-of-cards mutation, and Bürger, 1988, for non-Gaussian models with either type of mutation). Polygenic models with additive effects of genes at several diploid loci subject to random-walk mutation have also been studied. Although Gaussian distributions are not generally preserved under genetic mechanisms present in polygenic models, as segregation and recombination, not to mention partial mutation, Gaussian approximations have been used (see, e.g., Kimura, 1965; Bulmer, 1972; Lande, 1976, 1977, 1980; Fleming, 1979; Nagylaki, 1984). A house-of-cards approximation for a ‘random-walk’ mutation model

†Email: smahdi@uwichill.edu.bb

has even been proposed (Turelli, 1984) to model rare mutant alleles with relatively large effect; see, for instance, the discussion in Barton (1990). Approximations based on the recurrence equations for the moments of the allelic effects have also been considered (Turelli & Barton, 1990). Note that, assuming a multinormal distribution for the allelic effects at every stage of the life cycle, convergence of the mean and covariance matrix has been proved (Lessard & Mahdi, 1995; Mahdi & Lessard, 1996).

In this paper, we analyse a phenotypic model which combines random-walk mutation and house-of-cards mutation, called the centred house-of-cards mutation model. The mutation rate will be assumed small although it may be much larger than the rate that is usually used in genetic models which turns around 10^{-4} per individual per generation. When a mutation occurs, the character takes as its value the mean of the character in the population plus or minus a random effect which is completely independent of the previous value. This makes sense if, for example, the character is given by the mean in the population plus or minus an individual effect, as may occur for behavioural traits. Our mutation model can also be seen as a regulation process in a population in which a small fraction obeys a group law based exclusively on the mean. As in the standard house-of-cards mutation model, Gaussian distributions are not preserved from one generation to the next and we have to turn to approximations. Three approximations methods are proposed and compared. Convergence results are proved and numerical iterations are performed to study the rate of convergence. Furthermore, simulation studies are performed to compare the limiting variances obtained with the proposed approximation methods.

2. Centred house-of-cards mutation model

Suppose that, at generation t of an effectively infinite population, the distribution of a quantitative character is given by a Gaussian probability function $p(x, t)$ with mean $m(t)$ and variance $\sigma^2(t)$, that is,

$$p(x, t) = \frac{1}{\sigma(t)\sqrt{2\pi}} \exp\left[-\frac{(x - m(t))^2}{2\sigma^2(t)}\right]. \quad (1)$$

Let the fitness value of an individual whose character is x be given by a Gaussian function evaluated at x , namely,

$$\omega(x) = \exp\left[-\frac{(x - \theta)^2}{2\gamma^2}\right] \quad (2)$$

where θ and γ^2 represent respectively the centre and the amplitude of selection. Characters are favourably selected with a strength proportional to their distance to θ . This strength is much higher in the vicinity of θ especially when γ^2 is small.

Then, after selection, the probability density function of the character is

$$p_s(x, t) = \frac{\omega(x)p(x, t)}{\int_{-\infty}^{\infty} \omega(\xi)p(\xi, t) d\xi}, \quad (3)$$

whose mean and variance will be denoted by $m_s(t)$ and $\sigma_s^2(t)$, respectively. Actually, we have a Gaussian probability function

$$p_s(x, t) = \frac{1}{\sigma_s(t)\sqrt{2\pi}} \exp\left[-\frac{(x - m_s(t))^2}{2\sigma_s^2(t)}\right], \quad (4)$$

where

$$m_s(t) = \frac{\theta\sigma^2(t) + \gamma^2 m(t)}{\gamma^2 + \sigma^2(t)}, \quad (5)$$

and

$$\sigma_s^2(t) = \frac{\gamma^2\sigma^2(t)}{\gamma^2 + \sigma^2(t)}. \quad (6)$$

It is worth noting from equations (5) and (6) that the mean and the variance of the character are not significantly modified by selection when γ is large, say $\gamma > 5$. This case is referred to as *weak selection*. But, when γ is small, the selection changes these two parameters significantly. This case is referred to as *strong selection*.

Now assume that a fraction μ of the population mutates (or is subject to some regulation) so that the character of the mutants follows a Gaussian distribution centred at $m_s(t)$, described by the probability density function

$$g(x, m_s(t)) = \frac{1}{\delta\sqrt{2\pi}} \exp\left[-\frac{(x - m_s(t))^2}{2\delta^2}\right]. \quad (7)$$

If the individuals produce exact copies of themselves and the generations are discrete and separated, then the probability density function of the character at the beginning of the next generation will be

$$p(x, t + 1) = (1 - \mu)p_s(x, t) + \mu g(x, m_s(t)), \quad (8)$$

whose mean $m(t + 1)$ satisfies

$$|m(t + 1) - \theta| = |m_s(t) - \theta| = \frac{\gamma^2}{\gamma^2 + \sigma^2(t)} |m(t) - \theta| < |m(t) - \theta| \quad (9)$$

since $\sigma^2(t) > 0$. This is a mixture of Gaussian density functions, which is not generally a Gaussian density function. This makes an exact analysis of the recurrence system (8) difficult. However, we may have recourse to approximations.

3. Approximations

A classical approximation method suggests replacing the mixture of the distributions describing the character at any generation by a Gaussian distribution having the same

mean and the same variance. In such a case, the dynamics is completely determined by the following recurrence equations for the mean and variance of the character:

$$m(t+1) = \frac{\theta\sigma^2(t) + \gamma^2 m(t)}{\gamma^2 + \sigma^2(t)}, \quad (10)$$

$$\sigma^2(t+1) = (1-\mu) \frac{\gamma^2 \sigma^2(t)}{\gamma^2 + \sigma^2(t)} + \mu\delta^2. \quad (11)$$

We propose an alternative approximation method. Since the distributions in the mixture (8) have the same mean given by (10), it is natural to take this mean as the mean of the Gaussian distribution function approximating the mixture. Without loss of generality, we may assume this mean to be 0. It remains to determine the variance.

Given a probability density function in the form

$$p(x) = (1-\mu) \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_1^2}\right) + \mu \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_2^2}\right), \quad (12)$$

where $0 \leq \mu \leq 1$, we want to approximate this function by a Gaussian function:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (13)$$

where the value of σ is to be chosen such that some distance between f and p is minimum. We consider three possible metrics:

$$D_1(f, p) = \left\{ \int_{-\infty}^{\infty} (f(x) - p(x))^2 dx \right\}^{1/2}, \quad (14)$$

$$D_2(f, p) = \left\{ \int_{-\infty}^{\infty} [f(x) - p(x)]^2 f(x) dx \right\}^{1/2}, \quad (15)$$

$$D_3(f, p) = \int_{-\infty}^{\infty} \ln\left(\frac{p(x)}{f(x)}\right) f(x) dx. \quad (16)$$

The metrics D_1 and D_2 are L^2 distances, the first one with respect to the Lebesgue measure dx , and the second one with respect to the measure $f(x) dx$, while the metric D_3 corresponds to the Kullback's cross-entropy functional (see, e.g., Bickel & Doksum (1977) for its use in statistics).

Approximation using metric D_1

This metric does not give an explicit value for σ . Integrating carefully, we find that

$$D_1(f, p) = \left[\frac{1}{2\sqrt{\pi}} (\Phi(\sigma) + K) \right]^{\frac{1}{2}} \quad (17)$$

where

$$\Phi(\sigma) = \frac{1}{\sigma} - 2\sqrt{2} \left(\frac{1-\mu}{\sqrt{\sigma^2 + \sigma_1^2}} + \frac{\mu}{\sqrt{\sigma^2 + \sigma_2^2}} \right), \quad (18)$$

and

$$K = \frac{(1-\mu)^2}{\sigma_1} + \frac{\mu^2}{\sigma_2} + \frac{2\sqrt{2}\mu(1-\mu)}{\sqrt{\sigma_1^2 + \sigma_2^2}} > 0. \quad (19)$$

Therefore, minimizing $D_1(f, p)$ with respect to σ is equivalent to minimizing, with respect to σ , the function $\Phi(\sigma)$. The first derivative of $\Phi(\sigma)$ with respect to σ is given by

$$\Phi'(\sigma) = \frac{1}{\sigma^2} \left(-1 + \frac{2\sqrt{2}(1-\mu)}{\left(1 + \frac{\sigma_1^2}{\sigma^2}\right)^{3/2}} + \frac{2\sqrt{2}\mu}{\left(1 + \frac{\sigma_2^2}{\sigma^2}\right)^{3/2}} \right). \quad (20)$$

For $\mu = 0$, we have $\Phi'(\sigma) = 0$ when $\sigma = \sigma_1$, while, for $\mu = 1$, we have $\Phi'(\sigma) = 0$ when $\sigma = \sigma_2$. But in general, there is no explicit solution to $\Phi'(\sigma) = 0$. We have to consider other metrics. Nevertheless, there exists a unique solution which can be determined numerically. To prove this, let us consider the real positive value function

$$\zeta(\sigma) = \frac{1-\mu}{\left(1 + \frac{\sigma_1^2}{\sigma^2}\right)^{3/2}} + \frac{\mu}{\left(1 + \frac{\sigma_2^2}{\sigma^2}\right)^{3/2}} - \frac{1}{2\sqrt{2}}. \quad (21)$$

When $\sigma \rightarrow 0$, we have $\zeta(\sigma) \rightarrow -\frac{1}{2\sqrt{2}} < 0$ and, when $\sigma \rightarrow \infty$, we have $\zeta(\sigma) \rightarrow 1 - \frac{1}{2\sqrt{2}} > 0$. Therefore, the equation $\zeta(\sigma) = 0$ admits at least one positive solution. Furthermore, the derivative

$$\zeta'(\sigma) = \frac{3}{\sigma^3} \left(\frac{(1-\mu)\sigma_1^2}{\left(1 + \frac{\sigma_1^2}{\sigma^2}\right)^{5/2}} + \frac{\mu\sigma_2^2}{\left(1 + \frac{\sigma_2^2}{\sigma^2}\right)^{5/2}} \right) > 0, \quad (22)$$

for $\sigma > 0$. This shows that $\zeta(\sigma)$ is an increasing function and, therefore, the equation $\zeta(\sigma) = 0$ admits a unique positive solution. Consequently, equation (20) admits a unique positive solution as well.

Approximation using metric D_2

This metric leads to an approximate value for σ given by

$$\hat{\sigma} = \frac{\sigma_1\sigma_2}{\mu\sigma_1 + (1-\mu)\sigma_2}. \quad (23)$$

As a matter of fact, using the first-order approximation

$$E(\omega(X)) \simeq \omega(E(X)), \quad (24)$$

where

$$\omega(X) = (f(X) - p(X))^{1/2}, \quad (25)$$

the procedure of minimizing $D_2(f, p)$ with respect to σ is equivalent to minimizing, with respect to σ , the function

$$\Psi(\sigma) = \omega(0) = \left(\frac{1}{\sigma\sqrt{2\pi}} - \frac{\mu}{\sigma_2\sqrt{2\pi}} - \frac{(1-\mu)}{\sigma_1\sqrt{2\pi}} \right)^2. \quad (26)$$

Taking the first derivative with respect to σ , we get

$$\Psi'(\sigma) = -\frac{1}{\pi\sigma^2} \left(\frac{1}{\sigma} - \frac{\mu}{\sigma_2} - \frac{(1-\mu)}{\sigma_1} \right), \quad (27)$$

whose root, denoted by $\hat{\sigma}$, is the harmonic mean of σ_1 and σ_2 with respective weights $1 - \mu$ and μ . In particular, $\hat{\sigma} = \sigma_1$ for $\mu = 0$ and $\hat{\sigma} = \sigma_2$ for $\mu = 1$. Moreover, we have $\Psi'(\sigma) > 0$ for $\sigma > \hat{\sigma}$ and $\Psi'(\sigma) < 0$ for $\sigma < \hat{\sigma}$. We conclude that $\Psi(\sigma)$ is minimum at $\sigma = \hat{\sigma}$.

Therefore, the recurrence equations for the changes in the mean and standard deviation of the character from generation t to generation $t + 1$ in model (8) with the approximation using metric D_2 are

$$m(t+1) = \frac{\theta\sigma^2(t) + \gamma^2 m(t)}{\gamma^2 + \sigma^2(t)}, \quad (28)$$

$$\sigma(t+1) = \frac{\delta\gamma\sigma(t)}{\mu\gamma\sigma(t) + (1-\mu)\delta(\gamma^2 + \sigma^2(t))^{1/2}}. \quad (29)$$

Approximation using metric D_3

This metric gives the value of σ obtained by the classical approximation. To check that, we evaluate the Kullback's cross-entropy functional, which gives

$$D_3(f, p) = \ln\left(\frac{\mu(1-\mu)\sigma^2}{\sigma_1\sigma_2}\right) + \left[\frac{1}{\sigma^2} - \frac{1}{2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right) \right] (\mu\sigma_2^2 + (1-\mu)\sigma_1^2), \quad (30)$$

and we differentiate with respect to σ^2 to get

$$\frac{dD_3(f, p)}{d\sigma^2} = \frac{1}{\sigma^2} \left[1 - \frac{1}{\sigma^2} (\mu\sigma_2^2 + (1-\mu)\sigma_1^2) \right]. \quad (31)$$

This derivative admits a unique positive root

$$\hat{\sigma}^2 = (1-\mu)\sigma_1^2 + \mu\sigma_2^2, \quad (32)$$

which corresponds to a minimum point of $D_3(f, p)$ since the derivative is positive for $\sigma^2 > \hat{\sigma}^2$ and negative for $\sigma^2 < \hat{\sigma}^2$.

This approximation leads to the recurrence equations (10) and (11) for the mean and variance of the character.

4. Convergence

We are now ready to state and prove the main result of this section.

RESULT 1 The mean and variance of the character converge under the classical approximation (obtained by using metric D_3) leading to the recurrence equations (10) and (11), and under the approximation using metric D_2 leading to the recurrence equations (28) and (29).

Proof. With the classical approximation, the recurrence equation for the variance is given by the transformation

$$H(\sigma^2) = (1 - \mu) \frac{\gamma^2 \sigma^2}{\gamma^2 + \sigma^2} + \mu \delta^2. \quad (33)$$

This transformation admits a unique positive fixed point,

$$\hat{\sigma}^2 = \frac{\mu(\delta^2 - \gamma^2) + \sqrt{\mu^2(\delta^2 - \gamma^2)^2 + 4\delta^2\gamma^2\mu}}{2}. \quad (34)$$

Moreover, the transformation $H(\sigma^2)$ is increasing and concave for $\sigma^2 > 0$ since then

$$H'(\sigma^2) = (1 - \mu) \frac{\gamma^4}{(\sigma^2 + \gamma^2)^2} > 0 \quad (35)$$

and

$$H''(\sigma^2) = -2 \frac{(\sigma^2 + \gamma^2)(1 - \mu)\gamma^4}{(\sigma^2 + \gamma^2)^4} < 0. \quad (36)$$

Therefore the sequence $\sigma^2(t)$ converges to $\hat{\sigma}^2$ as t tends to ∞ from any starting variance $\sigma^2(0) > 0$. This is to say, that $\hat{\sigma}^2$ is globally stable.

The recurrence equation (10) for the mean $m(t)$ possesses a unique fixed point $\hat{m} = \theta$, and this fixed point is also globally stable. To prove this, let us rewrite equation (10) in the form

$$m(t+1) - m(t) = \frac{\sigma^2(t)}{\sigma^2(t) + \gamma^2} (\theta - m(t)). \quad (37)$$

If $\theta - m(t) \leq 0$, then $m(t+1) \leq m(t)$ and

$$m(t+1) = \frac{\gamma^2}{\gamma^2 + \sigma^2(t)} m(t) + \frac{\theta \sigma^2(t)}{\gamma^2 + \sigma^2(t)} \geq \frac{\gamma^2 \theta}{\gamma^2 + \sigma^2(t)} + \frac{\theta \sigma^2(t)}{\gamma^2 + \sigma^2(t)} = \theta. \quad (38)$$

Thus the sequence $m(t)$ converges since it is monotone decreasing and bounded below by θ . Similarly, if $\theta - m(t) \geq 0$, then $m(t+1) \geq m(t)$ and $m(t+1) \leq \theta$, from which the sequence $m(t)$ converges since it is monotone increasing and bounded above by θ . Then the sequence $m(t)$ converges. Finally, since the limit of any convergent sequence $m(t)$ must

be a fixed point, we conclude that $m(t)$ converges to the unique fixed point θ whatever the initial value is.

In the case of the approximation based on metric D_2 , it remains to prove convergence of $\sigma(t)$ since the recurrence equation (10) for $m(t)$ is the same.

The dynamics of $\sigma(t)$ is described by the transformation

$$G(\sigma) = \frac{\gamma\delta\sigma}{\mu\gamma\sigma + (1-\mu)\delta(\gamma^2 + \sigma^2)^{1/2}}. \quad (39)$$

This transformation is strictly increasing and concave for $\sigma > 0$ since

$$G'(\sigma) = \frac{\gamma(1-\mu)\delta^2(\gamma^2 + \sigma^2)^{1/2} \left(1 - \frac{\sigma^2}{\sigma^2 + \gamma^2}\right)}{[\mu\gamma\sigma + (1-\mu)\delta(\gamma^2 + \sigma^2)^{1/2}]^2} > 0 \quad (40)$$

and

$$\frac{d(\ln G'(\sigma))}{d\sigma} = \frac{G''(\sigma)}{G'(\sigma)} = - \left[\frac{\sigma}{\gamma^2 + \sigma^2} + 2 \frac{\mu\gamma + (1-\mu)\delta(\gamma^2 + \sigma^2)\sigma}{\mu\gamma\sigma + (1-\mu)\delta(\gamma^2 + \sigma^2)^{1/2}} \right] < 0. \quad (41)$$

Since $G'(\sigma) > 0$ and $G''(\sigma) < 0$ for $\sigma > 0$, with $G'(0) = 1/(1-\mu) > 1$ and $\lim_{\sigma \rightarrow \infty} G'(\sigma) = 0$, there exists a unique fixed point $\hat{\sigma} > 0$ which is globally stable from any starting $\sigma > 0$. \square

5. Rate of convergence

We present below a study on the rate of convergence towards equilibrium for the variance and the mean in both models (10) and (11), and (28) and (29).

Variance in model (10) and (11)

The sequence of the variances satisfies the inequality

$$|\sigma^2(t+1) - \hat{\sigma}^2| \leq \frac{(1-\mu)\gamma^4}{\gamma^2(\gamma^2 + \hat{\sigma}^2)} |\sigma^2(t) - \hat{\sigma}^2|. \quad (42)$$

Thus, we conclude that the sequence $\sigma^2(t)$ converges to $\hat{\sigma}^2$ geometrically fast with factor $\rho = \frac{(1-\mu)\gamma^4}{\gamma^2(\gamma^2 + \hat{\sigma}^2)} < 1$, since $\hat{\sigma}^2 > 0$.

Mean in model (10) and (11)

The sequence $m(t)$ satisfies the equation

$$m(t+1) - \hat{m} = \frac{\gamma^2}{(\gamma^2 + \sigma^2(t))} (m(t) - \hat{m}) + \frac{\gamma^2(\theta - \hat{m})}{(\gamma^2 + \sigma^2(t))(\gamma^2 + \hat{\sigma}^2)} (\sigma^2(t) - \hat{\sigma}^2). \quad (43)$$

Let $\eta(t) = m(t) - \hat{m}$, $\epsilon(t) = \sigma^2(t) - \hat{\sigma}^2$, $a(t) = \frac{\gamma^2}{(\gamma^2 + \sigma^2(t))}$ and $b(t) = \frac{\gamma^2(\theta - \hat{m})}{(\gamma^2 + \sigma^2(t))(\gamma^2 + \hat{\sigma}^2)}$. Then, equation (43) takes the form

$$\eta(t + 1) = a(t)\eta(t) + b(t)\epsilon(t). \tag{44}$$

The sequence $\sigma^2(t)$, being convergent to $\hat{\sigma}^2 > 0$, implies that $\epsilon(t)$ converges to 0. Furthermore, the sequence $b(t)$ is bounded in absolute value by $\frac{|\theta - \hat{m}|}{(\gamma^2 + \hat{\sigma}^2)}$. Therefore, for any given $\epsilon > 0$, we can find t_1 such that

$$|b(t)\epsilon(t)| < \epsilon \quad \text{for all } t > t_1.$$

Moreover, the sequence $a(t)$ converges to a limit $\hat{a} = \frac{\gamma^2}{\gamma^2 + \hat{\sigma}^2}$ satisfying $0 < \hat{a} < 1$. Thus, for any given r satisfying $0 < \hat{a} < r < 1$, we can find t_2 such that

$$|a(t)| < r < 1 \quad \text{for all } t > t_2.$$

Consequently, for all $t > \max(t_1, t_2)$, the sequence $\eta(t)$ satisfies the inequality

$$|\eta(t + 1)| \leq r|\eta(t)| + \epsilon. \tag{45}$$

We conclude that, in the vicinity of the fixed point \hat{m} , the rate of convergence of the sequence $m(t)$ is approximately geometric with a factor of the order of $\frac{\gamma^2}{(\gamma^2 + \hat{\sigma}^2)}$.

Variance in model (28) and (29)

The equilibrium expression in this case is derived as follows. The equilibrium standard deviation corresponds to the positive fixed point of the transformation G . The equation $G(\hat{\sigma}) = \frac{\gamma\delta\hat{\sigma}}{\mu\gamma\hat{\sigma} + (1 - \mu)\delta(\gamma^2 + \hat{\sigma}^2)^{1/2}} = \hat{\sigma}$ possesses the obvious solution $\hat{\sigma}_1 = 0$ and the solutions

$$\hat{\sigma}_2 = \frac{-\delta\gamma^2\mu - \sqrt{\delta^2\gamma^2(\mu - 1)^2\mu(d^2(2 - \mu) + \gamma^2\mu)}}{\delta^2(1 - \mu)^2 - \gamma^2\mu^2} \tag{46}$$

and

$$\hat{\sigma}_3 = \frac{-\delta\gamma^2\mu + \sqrt{\delta^2\gamma^2(\mu - 1)^2\mu(d^2(2 - \mu) + \gamma^2\mu)}}{\delta^2(1 - \mu)^2 - \gamma^2\mu^2} \tag{47}$$

obtained by solving the equation

$$(1 - \mu)^2\delta^2(\delta^2 + \hat{\sigma}^2) = (\gamma\delta - \mu\gamma\hat{\sigma})^2 \tag{48}$$

with the condition

$$\gamma\delta - \mu\gamma\hat{\sigma} \geq 0. \quad (49)$$

However, $\hat{\sigma} = \hat{\sigma}_3$ is the only admissible solution when $\delta^2(1 - \mu)^2 - \gamma^2\mu^2 \neq 0$. To prove this, let us rewrite equation (47) in the form

$$\hat{\sigma}_3 = \frac{-\delta\gamma^2\mu + \sqrt{(\delta\gamma^2\mu)^2 + Q}}{\delta^2(1 - \mu)^2 - \gamma^2\mu^2}, \quad (50)$$

where

$$Q = \delta^2\gamma^2(2 - \mu)\mu[\delta^2(1 - \mu)^2 - \gamma^2\mu^2]. \quad (51)$$

If $\delta^2(1 - \mu)^2 - \gamma^2\mu^2 > 0$, then $\hat{\sigma}_3$ is positive because the numerator and the denominator are both positive, while it is clear that $\hat{\sigma}_2$ cannot be positive. On the other hand, if $\delta^2(1 - \mu)^2 - \gamma^2\mu^2 < 0$, then both the numerator and the denominator in equations (47) and (50) are negative. Consequently, $\hat{\sigma}_2$ and $\hat{\sigma}_3$ are both positive. However, $\hat{\sigma}_3$ is the only admissible one with the constraint (49).

In the case of $\delta^2(1 - \mu)^2 - \gamma^2\mu^2 = 0$, that is, $\mu = \frac{\delta}{\delta + \gamma}$, the equation $G(\sigma) = \sigma$ becomes

$$-\sigma + \frac{\sigma(\delta + \gamma)}{\sigma + \sqrt{\gamma^2 + \sigma^2}} = 0, \quad (52)$$

and admits the unique positive solution $\hat{\sigma} = \frac{\delta(\delta + 2\gamma)}{2(\delta + \gamma)}$.

The analytical study of the mode of convergence of the variance sequence in this case is not obvious. Therefore, we focus our attention only on a numerical study of the behaviour of the iterates of $\sigma(t)$ in the vicinity of the equilibrium $\hat{\sigma}$. To this end, we compute the derivative of the transformation G at $\hat{\sigma}$ which gives

$$\begin{aligned} G'(\hat{\sigma}) = & (\delta\gamma^3(1 - \mu)) \left(\sqrt{\left[\gamma^2 + \frac{(-\delta\gamma^2\mu) + \sqrt{[\delta^2\gamma^2(-1 + \mu)^2\mu(2\delta^2 - \delta^2\mu + \gamma^2\mu)]}}{(\delta^2 - 2\delta^2\mu + \delta^2\mu^2 - \gamma^2\mu^2)^2} \right]^2} \right. \\ & \times \left\{ \frac{\gamma\mu(\delta\gamma^2\mu - \sqrt{[\delta^2\gamma^2(-1 + \mu)^2\mu(2\delta^2 - \delta^2\mu + \gamma^2\mu)])}}{(\delta^2 - 2\delta^2\mu + \delta^2\mu^2 - \gamma^2\mu^2)} \right. \\ & - \delta \sqrt{\left[\gamma^2 + \frac{(-\delta\gamma^2\mu) + \sqrt{[\delta^2\gamma^2(-1 + \mu)^2\mu(2\delta^2 - \delta^2\mu + \gamma^2\mu)]}}{(\delta^2 - 2\delta^2\mu + \delta^2\mu^2 - \gamma^2\mu^2)^2} \right]^2} \\ & \left. \left. + \delta\mu \sqrt{\left[\gamma^2 + \frac{(-\delta\gamma^2\mu) + \sqrt{[\delta^2\gamma^2(-1 + \mu)^2\mu(2\delta^2 - \delta^2\mu + \gamma^2\mu)]}}{(\delta^2 - 2\delta^2\mu + \delta^2\mu^2 - \gamma^2\mu^2)^2} \right]^2} \right\}^{-1} \right). \quad (53) \end{aligned}$$

By substituting different values of μ , γ and δ into equation (53), as illustrated in Appendix A., it appears that the rate of convergence near the equilibrium increases when μ increases. In particular, notice that this rate rapidly increases when the variance of the mutants is increased.

6. Approximation of the true limit variance

The approximation of the true limit variance in model (8) is done by simulation. Without loss of generality, we choose for selection a Gaussian probability density function $\omega(x)$ with mean θ and standard deviation σ . The first step in the simulation consists of subdividing the support of $\omega(x)$ into S classes with the same width around the mean. We denote by C_h , for $h = 1, \dots, S$, the class h with relative frequency $\omega(h)$. The main procedure is outlined in the following steps for each fixed μ , δ and γ .

- Step 1. Generate a N -dimensional random vector X from a normal distribution with mean $m(0)$ and standard deviation $\sigma(0)$.
- Step 2. Evaluate the variance of X .
- Step 3. For each $X(i)$, $i = 1, \dots, N$, find the selective class C_h , for $h = 1, \dots, S$, which contains $X(i)$. Then generate M random values from a uniform distribution on $[0,1]$. Let r_i denote the number of uniform values generated which are less than or equal to $\omega(h)$. Store r_i values $X(i)$ in a database Y . Repeat the process for all $X(i)$, $i = 1, \dots, N$.
- Step 4. Sample randomly $\lceil N(1 - \mu) \rceil$ observations from the database Y . ($\lceil \xi \rceil$ stands for the largest integer not exceeding the real ξ).
- Step 5. Sample randomly $N - \lceil N(1 - \mu) \rceil$ observations from a normal population with the mean of Y and variance δ^2 . Merge the two samples to create a new N -dimensional vector X' .
- Step 6. Substitute X' to X and repeat steps 1–5 until $|\text{var}(X) - \text{var}(X')|$ is less than a fixed tolerance level ϵ or a fixed large number of iterations.

7. Numerical results and discussion

We have conducted a simulation study to compare the variance limits, obtained with metrics D_1 , D_2 and D_3 , to the approximate true limit variance ν . To this end, we have considered the values $\mu = 10^{-2}$ to 10^{-10} and four classes for γ and δ , that is, $0 < \delta \leq 5, 0 < \gamma \leq 5$; $0 < \delta \leq 5, 5 < \gamma \leq 10$; $5 < \delta \leq 10, 0 < \gamma \leq 5$ and $5 < \delta \leq 10, 5 < \gamma \leq 10$. For each fixed value μ and each selected class, we have randomly generated 1000 pairs (δ, γ) . Then, from each random pair (δ, γ) and random initial variance $\sigma^2(0)$, we have computed the variance limits VD_1 , VD_2 and VD_3 using metrics D_1 , D_2 and D_3 , respectively. VD_1 , VD_2 and VD_3 are respectively obtained by repeating the minimization process of formula (17), with respect to σ , and by iterating the dynamical systems (29) and (11) until equilibrium. Next, we have computed, in each of the four classes, the average of the 1000 limit variances obtained. Similarly, we have computed the approximate true limit variance ν of the original model (8) according to the procedure outlined in section 6. The asymptotic variances are found to be independent of the initial

variances. In the case $\mu \leq 10^{-7}$, more than 10^6 iterations are needed to evaluate the limit ν . The computed asymptotic variances with metrics D_2 and D_3 agree with equilibria given in formulas (47) and (34), respectively. Note that all variance limits decline uniformly when μ decreases and become very small when both γ and δ are small (see, for instance, Tables 6 and 7 in Appendix A:). Furthermore, for each fixed μ and fixed class for γ and δ , the asymptotic variance VD_2 is always smaller than VD_1 and VD_3 which are closer to ν . Consequently, metrics D_1 and D_3 are the most suitable. However, VD_1 is generally closer to ν when μ is small while VD_3 is closer to ν when μ is large. This is illustrated in Tables 5–7 in Appendix A: . Furthermore, we have conducted another simulation to compare the rate of convergence of the mean and variance in models (10) and (11), and (28) and (29) . The study showed that equilibrium in models (10) and (11) is always reached earlier than equilibrium in models (28) and (29). Moreover, the convergence of the mean is only completed after the stability of the variance process. In Appendix B:, figures illustrating the speed of convergence of the variance sequences in both of the models (11) and (29) are presented. The vertical axis is log scaled because the curves are indistinguishable using a linear scale. The curves describing the dynamic of the variance with metric D_1 are also given. The graphs show that the transformation of the variance based on metric D_1 reaches its fixed point rapidly. It is worth nothing that some variance limits are reached after more than 10^5 iterations. Table 5 in Appendix A: gives the limiting values corresponding to Figs 1–4.

In real world situations, the measurements are often not normally distributed. In such cases, to use the approximations presented in this paper, we recommend the measurements be transformed to normality using, for example, the Box & Cox (1964) transformation. This useful transformation has recently been applied successfully by Fatti *et al.* (1998) for updating in reference centile charts.

Finally, we may say, from what we have observed, that the Gaussian approximation based on metric D_1 better targets the variance of the original mixture model when we use genetic mutation rates while the classical approximation gives better results for phenotypic mutation rates. The approximation based on metric D_2 underestimates the variability in the original model.

Acknowledgments

Thanks are due to the referees for useful comments on a first draft of this paper. This research was partially supported by UWI, FCAR and NSERC grants. The numerical studies have been mainly carried out with Mathematica and Gauss (Unix Version 3.2.42).

REFERENCES

- BARTON, N. H. 1990 Pleiotropic models of quantitative variation. *Genetics* **124**, 773–782.
- BICKEL, P. J. & DOKSUM, K. A. 1977 *Mathematical Statistics: Basic Ideas and Selected Topics*. Oakland, CA: Holden-Day Inc.
- BOX, G. E. P. & COX, D. R. 1964 An analysis of transformations (with discussion). *J. R. Statist. Soc. B* **26**, 211–252.
- BULMER, M. G. 1972 The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genet. Res.* **19**, 17–25.

- BÜGER, R. 1988 Mutation-selection balance and continuum-of-alleles models. *Math. Biosci.* **91**, 67–83.
- ESHEL, I. 1971 On evolution in a population with an infinite number of types. *Theor. Pop. Biol.* **2**, 209–236.
- FATTI, L. P., SENAOANA, E. M., & THOMPSON, M. L. 1998 Bayesian updating in reference centile charts. *J. R. Statist. Soc. A* **161**, 103–115.
- FLEMING, W. H. 1979 Equilibrium distributions of continuous polygenic traits. *SIAM J. Appl. Math.* **36**, 148–168.
- KARLIN, S. 1979 Models of multifactorial inheritance: I, multivariate formulations and basic convergence results. *Theor. Pop. Biol.* **15**, 308–355.
- KARLIN, S. 1988 Non-Gaussian phenotypic models of quantitative traits. *Proceedings of the Second International Conference on Quantitative Genetics, Raleigh, NC*, (B. S. Weir, E. J. Eisen, M. M. Goodman & G. Namkoong eds). Sinauer, pp. 123–144.
- KIMURA, M. 1965 A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci. USA* **54**, 731–736.
- KINGMAN, J. F. C. 1978 A simple model for the balance between selection and mutation. *J. Appl. Prob.* **15**, 1–12.
- LANDE, R. 1976 The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* **26**, 221–235.
- LANDE, R. 1977 The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genetics* **86**, 485–498.
- LANDE, R. 1980 The genetic covariance between traits maintained by linked pleiotropic mutations. *Genetics* **91**, 203–215.
- LESSARD, S. & MAHDI, S. 1995 Convergence de la variabilité dans les modèles polygéniques gaussiens. *Genetics Selection Evolution* **27**, 395–421.
- MAHDI, S. & LESSARD, S. 1996 Convergence of covariance structures in additive Gaussian polygenic models. *Biometrics* **52**, 833–845.
- NAGYLAKI, T. 1984 Selection on a quantitative character. *Human Population Genetics: The Pittsburg Symposium*. (A. Chakravarti ed). New York: Van Nostrand Reinhold, pp. 216–237.
- ROUGHGARDEN, J. 1972 Evolution of niche width. *Am. Nat.* **106**, 683–718.
- SLATKIN, M. 1970 Selection and polygenic characters. *Proc. Natl. Acad. Sci. USA* **66**, 87–93.
- TURELLI, M. 1984 Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theor. Pop. Biol.* **25**, 138–193.
- TURELLI, M. & BARTON, N. H. 1990 Dynamics of polygenic characters under selection. *Theor. Pop. Biol.* **38**, 1–57.

Appendix A.: Tables

TABLE 1
 Values of $\hat{\sigma}$ and $G'(\hat{\sigma})$ for various mutation rates μ in a case of strong selection ($\gamma = 1$) and mutants with small variance ($\delta = 1$)

$\delta = 1.0$ and $\gamma = 1.0$						
μ	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-10}
$\hat{\sigma}$	0.0140	0.0045	0.0014	0.00044	0.00014	0.000014
$G'(\hat{\sigma})$	0.9997	0.9998	0.9999	0.9999	0.9999	0.9999

TABLE 2
Values of $\hat{\sigma}$ and $G'(\hat{\sigma})$ for various mutation rates μ in a case of weak selection ($\gamma = 100$) and mutants with small variance ($\delta = 1$)

$\delta = 1.0$ and $\gamma = 100$						
μ	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-10}
$\hat{\sigma}$	0.7318	0.3582	0.1317	0.043	0.014	0.0014
$G'(\hat{\sigma})$	0.9997	0.9998	0.9999	0.9999	0.9999	0.9999

TABLE 3
Values of $\hat{\sigma}$ and $G'(\hat{\sigma})$ for various mutation rates μ in a case of strong selection ($\gamma = 1$) and mutants with large variance ($\delta = 100$)

$\delta = 100$ and $\gamma = 1.0$						
μ	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-10}
$\hat{\sigma}$	0.0142	0.0045	0.0014	0.00044	0.00014	0.00014
$G'(\hat{\sigma})$	0.0099	0.0099	0.0099	0.01	0.01	0.01

TABLE 4
Values of $\hat{\sigma}$ and $G'(\hat{\sigma})$ for various mutation rates μ in a case of weak selection ($\gamma = 100$) and mutants with large variance ($\delta = 100$)

$\delta = 100$ and $\gamma = 100$						
μ	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-10}
$\hat{\sigma}$	1.4040	0.4462	0.1413	0.044	0.014	0.0014
$G'(\hat{\sigma})$	0.0099	0.0099	0.0099	0.01	0.01	0.01

TABLE 5
Variance limits in the case $\mu = 10^{-7}$, $\gamma = 1, 10$ and $\delta = 1, 10$. v represents the approximate true limit variance

Variance limits	VD_1	VD_2	VD_3	v
$\delta = 1$ and $\gamma = 1$	0.00215	0.000002	0.00031	0.0011
$\delta = 1$ and $\gamma = 10$	0.058	0.000019	0.00315	0.035
$\delta = 10$ and $\gamma = 1$	0.010	0.000002	0.00316	0.022
$\delta = 10$ and $\gamma = 10$	0.215	0.00009	0.0316	0.13

TABLE 6
Simulated average limit variances for $0 < \delta \leq 10$ and $0 < \gamma \leq 10$ in the case $\mu = 10^{-4}$. v represents the approximate true limit variance

Variance limits	VD_1	VD_2	VD_3	v
$\delta \leq 5$ and $\gamma \leq 5$	0.0899	0.00126	0.062	0.056
$\delta \leq 5$ and $5 < \gamma \leq 10$	0.38	0.01010	0.188	0.276
$5 < \delta \leq 10$ and $\gamma \leq 5$	0.2042	0.0013	0.1864	0.221
$5 < \delta \leq 10$ and $5 < \gamma \leq 10$	0.7947	0.012	0.56	0.611

TABLE 7
 Simulated average limit variances for $0 < \delta \leq 10$ and $0 < \gamma \leq 10$ in the case $\mu = 10^{-2}$.
 v represents the approximate true limit variance

Variance limits	VD_1	VD_2	VD_3	v
$\delta \leq 5$ and $\gamma \leq 5$	0.1175	0.09615	0.6194	0.467
$\delta \leq 5$ and $5 < \gamma \leq 10$	0.62085	0.6116	1.67726	1.5742
$5 < \delta \leq 10$ and $\gamma \leq 5$	0.4140	0.16452	2.176	1.0071
$5 < \delta \leq 10$ and $5 < \gamma \leq 10$	0.84	0.55	5.709	4.4100

Appendix B.: Graphics

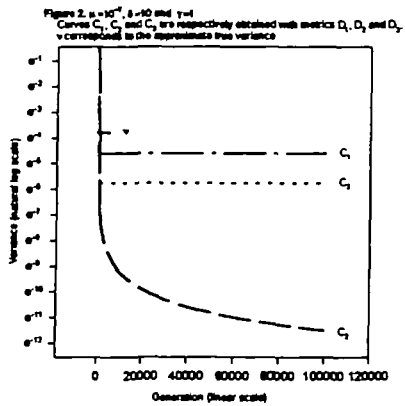
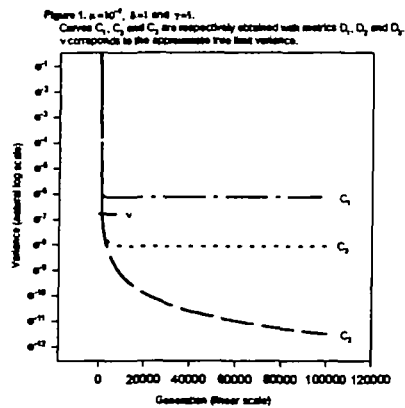


Figure 3. $\mu = 10^{-6}$, $\delta = 1$ and $\gamma = 10$
 Curves C_1 , C_2 and C_3 are respectively obtained with metrics D_1 , D_2 and D_3
 v corresponds to the approximate true fixed variance

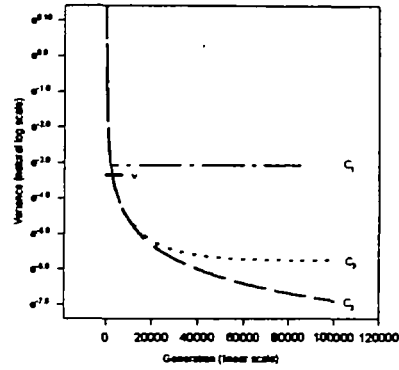


Figure 4. $\mu = 10^{-7}$, $\delta = 10$ and $\gamma = 10$
 Curves C_1 , C_2 and C_3 are respectively obtained with metrics D_1 , D_2 and D_3
 v corresponds to the approximate true fixed variance

