

Taxa de mutação e seleção dos mais aptos em um modelo diferencial de evolução

Raul A. Assis¹,

Departamento de Matemática, UNEMAT, 78555-000, Sinop/MT.

Wilson C. Ferreira Jr.²,

DMA, IMECC, UNICAMP, 13.083-970, Campinas/SP.

Resumo. Neste artigo apresentamos uma análise da relação entre velocidade de mutação e a efetiva seleção dos indivíduos mais aptos em uma população através de um modelo diferencial do processo de evolução de uma população em um espaço de aspecto abstrato. Resultados indicam se a taxa de mutação na escala de tempo da reprodução for suficiente para produzir uma grande diferença no valor de adaptação do indivíduos a seleção natural não será eficiente na seleção dos indivíduos mais aptos.

Palavras-chave: *Evolução; Modelo ; Equações Diferenciais; Taxa de Mutação.*

1. Introdução

Dadas as escalas de tempo em que ocorrem processos evolutivos, a observação direta de certos fenômenos se torna inviável. Dessa forma, modelos matemáticos se apresentam como alternativas para a análise de processos evolutivos em diversas escalas de tempo. Diversos tipos de modelos foram desenvolvidos para a análise de fenômenos evolutivos, alguns baseados em sequências que simulam o código genético (Kauffman e Levin, 1987; Taylor e Higgs, 2000), outros buscam descrever a mudança na frequência de genótipos uma população

¹raul@unemat-net.br

²wilson@ime.unicamp.br

(Kimura, 1983) e outros ainda utilizam conceitos de teoria dos jogos para descrever processos evolutivos (Maynard-Smith, 1982). Cada uma dessas diferentes abordagens possui suas vantagens e qual é a mais eficiente depende do contexto de aplicação do modelo.

Assis e Ferreira Jr. (2005) propuseram um modelo para descrever processos evolutivos em um espaço de aspecto abstrato. A natureza desse espaço de aspecto também é dependente do contexto em que se aplique o modelo, mas pode ser entendido como um espaço que representa alguma característica (ou aspecto) do indivíduos, podendo ser uma medida física ou biológica (altura, peso, curvatura do bico, etc.) ou simplesmente uma “posição” abstrata dos indivíduos com relação a alguma característica (dessa forma pode-se até mesmo incluir características subjetivas através de uma quantificação *fuzzy*).

Nosso interesse no presente artigo é analisar como a “taxa de mutação” nesse espaço de aspecto abstrato influencia na seleção dos indivíduos da população. Resultados clássicos obtidos por J. S. Haldane e R. A. Fisher (Gillespie, 1998; Hartl e Clark, 2007; Fisher, 1930) relacionam a frequência de uma mutação deletéria com a taxa de mutação e os valores de adaptação relativos do genótipo selvagem (*wild type*) e da mutação e indicam que no caso de valores de mutação “moderados” a população chega a um equilíbrio nas frequências dos genótipos. Em geral, esse tipo de análise do equilíbrio entre mutação e seleção supões que a frequência do genótipo deletério permanece pequena, desprezando fatores como geração de novas mutações deletérias e mutação de volta ao genótipo selvagem.

De fato, esses fatores podem ser desprezados no caso de uma taxa de mutação pequena e de uma desvantagem seletiva apreciável, entretanto, quando a taxa de mutação for muito alta, ou a desvantagem (efeito deletério da mutação) for muito pequena, os resultados não se aplicam. Dessa forma, utilizaremos um modelo de espaço de aspecto para analisar essas relações, inclusive quando a taxa de mutação for elevada.

2. O modelo

A seguir descrevemos as hipóteses que descrevem o modelo:

1. A população de distribui em um espaço de aspecto Ω .
2. Existe uma função $f : \Omega \rightarrow [0, 1]$ que descreve o *valor de adaptação* do

indivíduo, relacionado com sua taxa de reprodução.

3. A meio em que se encontra a população possui recursos limitados e todos indivíduos competem pelos recursos.
4. Ao se reproduzir, indivíduos produzem descendentes iguais ou semelhantes às suas características descritas no espaço de aspecto.

As hipóteses acima são uma breve descrição dos mecanismos de seleção natural propostos por Darwin, no século XIX. Dessa forma, nosso modelo estará trabalhando com as hipóteses descritas acima, apesar de que as mesmas podem ser modificadas e/ou relaxadas (por exemplo, são conhecidos exemplos em que a hipótese (4) acima não é válida).

De forma a tornar a análise mais simples adotaremos o espaço de aspecto como sendo $\Omega = [-L, L] \subset \mathbb{R}$. Assim, uma dinâmica que simula as hipóteses acima é:

$$\frac{\partial u}{\partial t} = rD \frac{\partial^2 [f(x)u]}{\partial x^2} + ru \left(f(x) - \int_{\Omega} u(x)dx/K \right) \quad (2.1)$$

onde r é a taxa máxima de reprodução, D o coeficiente mutação e K a capacidade máxima de suporte do meio. A equação (2.1) foi obtida inicialmente através de um processo de discretização do espaço de aspecto e do tempo, sendo descritas em detalhes em Assis e Ferreira Jr. (2005). A seguir apresentamos uma rápida interpretação do modelo.

Separando os termos do lado direito da equação (2.1) temos:

$$\frac{\partial u}{\partial t} = rD \underbrace{\frac{\partial^2 [f(x)u]}{\partial x^2}}_I + ru \underbrace{f(x)}_{II} - ru \underbrace{\int_{\Omega} u(x)dx/K}_{III}$$

O termo I representa a geração de indivíduos com características semelhantes a partir da densidade de indivíduos $u(x, t)$. O termo II representa a reprodução dos indivíduos de acordo com a densidade $u(x, t)$, seus valores de adaptação $f(x)$ e a taxa máxima de reprodução r . O termo III representa uma mortalidade do tipo Verhulst, só que, neste caso o número total de indivíduos $N(t)$ é dado pela integral $\int_{\Omega} u(x, t)dx$. São considerados dados do problema a distribuição inicial dos indivíduos $u_0(x)$, os parâmetros r, D, K e a função de adaptação $f(x)$.

Utilizando uma adimensionalização na forma $t^* = rt$, $x^* = x/\sqrt{D}$ e $u^* = u\sqrt{D}/K$ a equação (2.1) fica:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 [f(x)u]}{\partial x^2} + u \left(f(x) - \int_{\Omega} u(x)dx \right) \quad (2.2)$$

onde abandonamos os asteriscos nas variáveis por simplicidade de notação e $\bar{\Omega} = [-L/\sqrt{D}, L/\sqrt{D}]$. De agora em diante também escreveremos apenas Ω para $\bar{\Omega}$.

3 Descrição do caso a ser analisado

O modelo descrito pela equação (2.2) é bastante geral e pode descrever diversos casos relativos à evolução de uma população em um espaço de aspecto. Nosso interesse é discutir uma situação específica em que existe uma posição no espaço de aspecto que possui um valor de máximo de adaptação e analisar a influência do coeficiente de mutação D no estabelecimento dos mais aptos.

Para simular esse tipo de cenário, utilizaremos uma função $f : \Omega \rightarrow [0, 1]$ com as seguintes características:

1. $x = 0$ é ponto de máximo global de f .
2. f é par (simétrica em torno do ponto de maior adaptação)
3. $f(x) \approx 0$ se $x > \delta$.

com as características acima, f descreve uma situação onde o valor de adaptação é máximo em $x = 0$ e decai, sendo próximo de zero quando $x \rightarrow \delta$. Isso inclui uma importante escala no modelo, pois estamos informando que uma modificação de δ no espaço de aspecto resulta em uma mudança significativa no valor de adaptação do indivíduo.

Queremos também que as condições de contorno não influenciem no processo de evolução, de modo que, idealmente, $L = \infty$.

4 Análise do modelo

4.1 Dinâmica para $N(t)$

A equação (2.2) é não linear, o que dificulta sua análise. Entretanto, simulações numéricas e a coerência com o fenômeno biológico indicam que a solução $u(x, t)$ deve tender à uma solução estacionária $u^*(x)$, o que também nos leva a um ponto estacionário N^* para $N(t) = \int_{\Omega} u(x, t) dx$.

Integrando ambos lados da equação (2.2) em relação a x :

$$\frac{\partial [\int_{\Omega} u dx]}{\partial t} = \int_{\Omega} \frac{[\partial^2 f(x) u]}{\partial x^2} dx + \int_{\Omega} u f(x) dx - \int_{\Omega} u dx \int_{\Omega} u(x) dx$$

como $\int_{\Omega} u dx = N(t)$, temos:

$$\frac{dN}{dt} = \int_{\Omega} \frac{\partial^2 [f(x)u]}{\partial x^2} dx + \int_{\Omega} u f(x) dx - N^2$$

mas

$$\int_{\Omega} \frac{\partial^2 [f(x)u]}{\partial x^2} dx = \lim_{b \rightarrow \infty} \frac{\partial [uf(x)]}{\partial x} \Big|_{-b}^b$$

simulações numéricas e o formato esperado para $u(x)$ indicam que $\frac{\partial [u(x)f(x)]}{\partial x} \rightarrow 0$ se $|x| \rightarrow \infty$. Assim, temos a seguinte dinâmica para $N(t)$:

$$\frac{dN}{dt} = \int_{\Omega} u(x, t) f(x) dx - N^2 \quad (4.3)$$

utilizando a hipótese de que $u(x, t) \rightarrow u^*(x)$ temos que N^* deve satisfazer:

$$N^* = \sqrt{\int_{\Omega} u^*(x) f(x) dx} \quad (4.4)$$

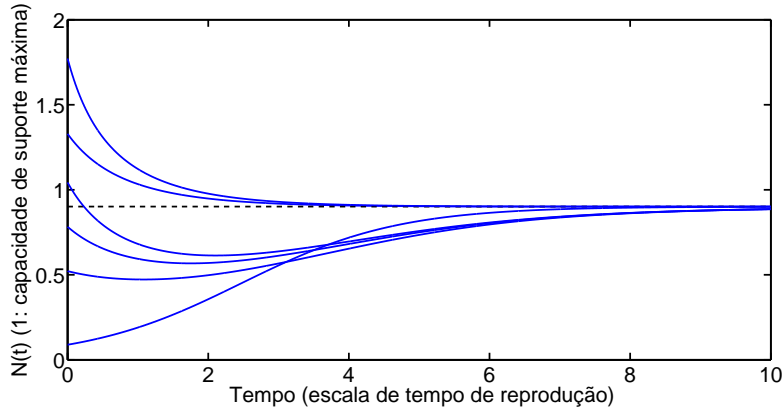


Figura 1: Simulações numéricas do modelo. Gráfico de $N(t)$ para diferentes valores de $N(0) = \int_{\Omega} u_0(x) dx$. Observe como todas tendem ao valor estacionário $N^* = \sqrt{\int_{\Omega} u^*(x) f(x) dx}$.

Na figura 1 apresentamos trajetórias de $N(t)$ calculadas através da aproximação numérica da solução $u(x, t)$ de (2.2) com diferentes valores de $N(0) = \int_{\Omega} u_0(x) dx$, a figura ilustra como as trajetórias tendem ao valor estacionário $N^* = \sqrt{\int_{\Omega} u^*(x) f(x) dx}$

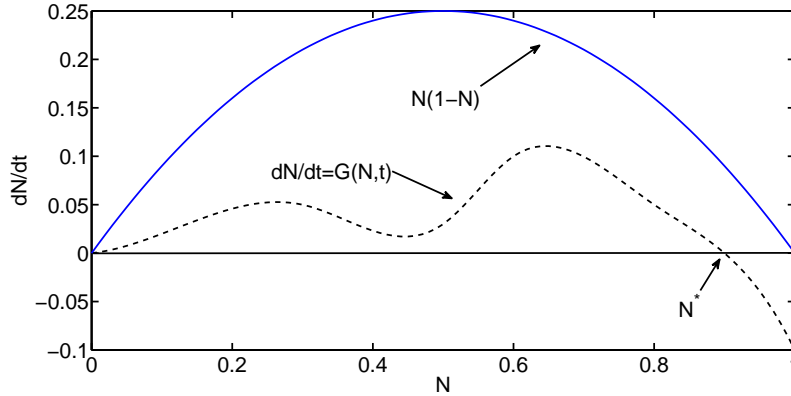


Figura 2: Esboço de $G(N, t) = dN/dt$ em função de N . Na verdade $G(N, t)$ depende de $u(x, t)$, mas será positiva quando $N \approx 0$ e negativa se $N > 1$, gerando um ponto estacionário N^* entre 0 e 1. Esse ponto é estacionário porque $u(x, t)$ tende à uma distribuição estacionária $u^*(x)$.

A garantia da existência do ponto estacionário pode ser deduzida observando-se que, como $f(x) \leq 1$ temos que $\int_{\Omega} u(x, t) f(x) dx \leq N(t)$, logo se $N(t) > 1$, $\frac{dN}{dt} \leq N - N^2 < 0$. Além disso, se $N \approx 0$, temos que:

$$\frac{dN}{dt} = \int_{\Omega} u(x, t) f(x) dx - \int_{\Omega} u(x, t) N(t) dx = \int_{\Omega} u(x, t) (f(x) - N(t)) dx$$

mas $N(t)$ não depende de x , logo, se $N(t) \approx 0$, o termo $f(x) - N(t)$ será positivo na maior parte do domínio, desde que $N(t)$ seja pequeno o bastante, dessa forma a integral $\int_{\Omega} u(x, t) (f(x) - N(t)) dx$ será positiva e dN/dt também. Levando essas informações em conta, podemos esboçar um gráfico de dN/dt em função de N , que apresentamos na figura 2.

4.2 Distribuição estacionária

Simulações numéricas e a natureza biológica particular da formulação do modelo sugerem que a solução $u(x, t)$ deve tender a uma distribuição estacionária $u^*(x)$. Dessa forma, levando-se em conta as equações (2.2) – equação de dinâmica $u(x, t)$ – e (4.4) – equação para N^* – temos que u^* deve satisfazer:

$$\frac{d^2 [f(x)u]}{dx^2} + u(f(x) - N^*) = 0 \quad (4.5)$$

que é uma equação linear em u .

A análise do modelo pode tomar diversas direções, uma vez que o mesmo representa um processo generalizado de evolução de uma população em um espaço de aspecto. Neste trabalho em particular, estamos abordando a questão da relação entre taxa de mutação (no modelo representada pelo coeficiente de mutação D) e seleção dos indivíduos mais aptos (neste caso representados pelos indivíduos com característica $x = 0$, que é o ponto de máximo de f). Dessa forma, é de nosso interesse determinar se a distribuição estacionária terá o ponto de máximo em $x = 0$, coincidindo com o ponto de máximo de f ou se tomará outra forma, dependendo do coeficiente de mutação.

Uma vez que estamos interessados no comportamento assintótico da solução estacionária $u^*(x)$ com $x \rightarrow 0$, utilizaremos a expansão em séries de u^* e f em torno do ponto $x = 0$ para obter uma aproximação da solução. Por hipótese, f é par, de forma que todos os termos de potência ímpar são nulos, além disso, u^* também é par, pois a equação (4.5) é simétrica em torno da origem (fazendo $y = -x$, e utilizando o fato de f ser par obtemos exatamente a mesma equação para $u(y)$). Dessa forma, a expansão para f e u^* fica:

$$f(x) = \sum_{n=0}^{\infty} f_{2n} \frac{x^{2n}}{(2n)!}$$

$$u^*(x) = \sum_{n=0}^{\infty} a_{2n} x^{2n}$$

onde $f_{2n} = f^{(2n)}(0)$. Utilizando a equação (4.5), obtemos as relações de recursão:

$$a_0 f_0 + 2a_2 f_0 + a_0 f_2 - a_0 N^* = 0 \quad (4.6)$$

$$a_2 f_0 + 12a_4 f_0 + (a_0 f_2)/2 + 6a_2 f_2 + (a_0 f_4)/2 - a_2 N^* = 0 \quad (4.7)$$

Como estamos interessados no comportamento assintótico de u^* quando $x \rightarrow 0$, utilizaremos apenas a relação (4.6), obtendo.

$$a_2 = -a_0 \frac{(f_0 + f_2 - N^*)}{2f_0} \quad (4.8)$$

como $2a_2 = (d[u^*(x)]/dx^2)|_{x=0}$, a solução terá um ponto de máximo em $x = 0$ se $a_2 < 0$, ou seja:

$$f_0 + f_2 > N^* \quad (4.9)$$

Para entender o significado da inequação (4.9) devemos observar a escala δ de variação de f . O significado biológico de δ é a distância no espaço de

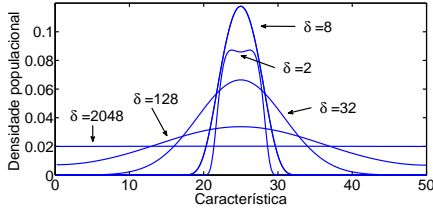


Figura 3: Gráficos de $u^*(x)$ para diferentes valores de δ . A população tende a se distribuir uniformemente se f tende a uma constante ($\delta \rightarrow \infty$).

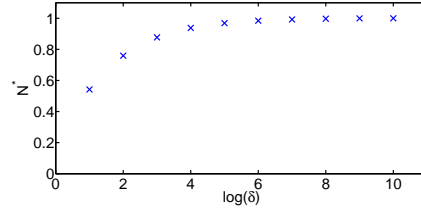


Figura 4: N^* em função de δ . Se $f(0) = 1$, $N^* \rightarrow 1$. Conforme $\delta \rightarrow \infty$, N^* tende a $f(0)$. Neste caso $f(x) = e^{-x^2/\delta^2}$.

aspecto que a população deve percorrer para que ocorra uma grande variação no valor de adaptação ($f(0)$ ponto de máximo, $f(\pm\delta) \approx 0$). Tanto f_2 como N^* são funções de δ . f_2 descreve como o valor de adaptação decai em torno do ponto de máximo em função de x . Para notar que N^* também depende de δ , basta tomar o caso extremo em que $\delta \rightarrow \infty$, neste caso f tende a um valor constante, e a distribuição da população tende a uma distribuição uniforme com $N(t) \rightarrow f(0)$, como na figura 4. Na figura 3 apresentamos alguns gráficos que representam qualitativamente essas relações.

O que pode ser observado através da figura 3 é que, conforme a escala δ tende a infinito, a seleção é muito fraca e a população tende a se distribuir uniformemente pelo espaço de aspecto, com N^* atingindo o valor máximo de f . Portanto, a inequação (4.9) representa uma condição sobre a escala δ de mudança no valor de adaptação e a estabelecimento dos mais aptos.

Para ilustrar a validade do critério, nas figuras 5 e 6 apresentamos duas aproximações numéricas da solução estacionária $u^*(x)$ e da aproximação $\bar{u}(x) = a_0 + a_2x^2$. Observamos como a mudança de concavidade coincide com a aparição de dois picos na solução estacionária.

A região de validade da inequação $f_0 + f_2 > N^*$ depende de como f e N variam em função de δ , na seção a seguir, analisamos mais a fundo o caso particular em que $f(x) = f_0 e^{-(x/\delta)^2}$.

4.3 Um caso particular

Nesta seção, analisaremos o caso em que $f(x) = f_0 e^{-(\epsilon x)^2}$ onde $\epsilon = 1/\delta$. Neste caso, temos que $f_2 = f''(0) = -2f_0\epsilon^2$, dessa forma, para melhor

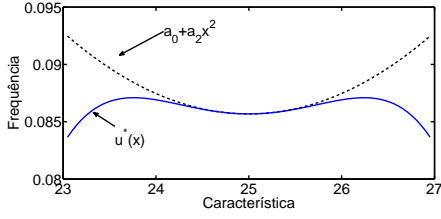


Figura 5: Se $a_2 > 0$, $u^*(x)$ tem um ponto de mínimo em $x = 0$ e os mais aptos não aparecem em maior frequência na população. Aproximação numérica para $u^*(x)$.

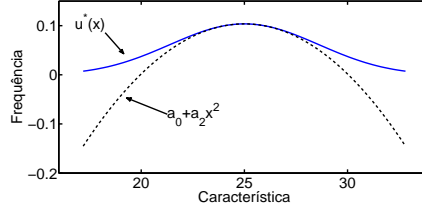


Figura 6: Se $a_2 < 0$, $u^*(x)$ tem um ponto de máximo em $x = 0$ e os mais aptos são os mais frequentes na população. O ponto de máximo de $u^*(x)$ e $f(x)$ coincidem.

interpretar a inequação (4.9) necessitamos de uma aproximação para N^* em função de ϵ . Nesta seção trabalharemos com aproximações de N^* e $u^*(x)$ mas, para não sobrecarregar a notação utilizaremos as próprias variáveis para denotar suas aproximações.

Para obter tal aproximação utilizaremos a suposição de que a solução estacionária pode ser aproximada por $u(x) = u_0 e^{-(\theta x)^2}$, de forma que as variáveis N^* , θ e ϵ estão relacionadas. Para obter uma aproximação para essas relações, utilizamos a equação para a solução estacionária:

$$\frac{d^2 [f(x)u(x)]}{dx^2} + u(x)(f(x) - N^*) = 0$$

como $N^* = \int_{\Omega} u(x) dx = u_0 \sqrt{\pi}/\theta$, temos:

$$\left(u_0 f_0 / \epsilon x^2 \epsilon^2 - \sqrt{\pi} u_0 / \theta \right) / e^{x^2 \theta^2} + f_0 u_0 e^{-x^2 (\epsilon^2 + \theta^2)} (-2\epsilon^2 - 2\theta^2 + (-2x\epsilon^2 - 2x\theta^2)^2) = 0 \quad (4.10)$$

desenvolvendo a equação acima em séries de potências e agrupando os termos de ordem um e dois obtemos duas equações:

$$f_0 u_0 - 2f_0 u_0 \epsilon^2 - \sqrt{\pi} u_0^2 / \theta - 2f_0 u_0 \theta^2 = 0 \quad (4.11)$$

$$f_0 u_0 (6\epsilon^4 - \epsilon^2 - \theta^2 + 12\epsilon^2 \theta^2 + 6\theta^4) + \sqrt{\pi} u_0^2 \theta = 0 \quad (4.12)$$

da equação (4.11) obtemos u_0 em função de ϵ e θ :

$$u_0 = \frac{f_0 \theta (1 - 2\epsilon^2 - 2\theta^2)}{\sqrt{\pi}} \quad (4.13)$$

utilizando a equação de ordem 2 (4.12) obtemos a relação de θ e ϵ :

$$\theta = \frac{\sqrt{-5\epsilon^2 + \sqrt{4\epsilon^2 + \epsilon^4}}}{2} \quad (4.14)$$

Finalmente, da relação $N^* = \sqrt{\pi}u_0/\theta$, obtemos uma aproximação para N^* :

$$N^* \cong \frac{f_0 \left(2 + \epsilon^2 - \sqrt{\epsilon^2(4 + \epsilon^2)} \right)}{2} \quad (4.15)$$

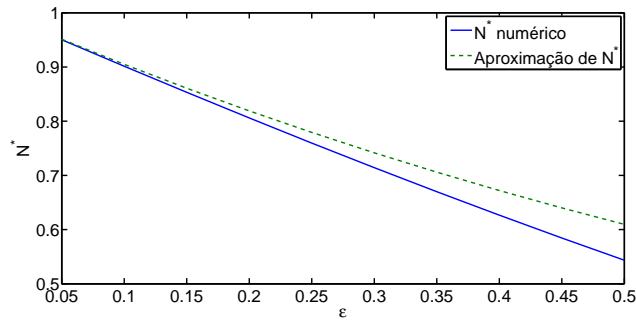


Figura 7: Comparação entre a aproximação dada por $N^* \approx f_0 \left(2 + \epsilon^2 - \sqrt{\epsilon^2(4 + \epsilon^2)} \right) / 2$ e valores numéricos de N^* . Conforme $\epsilon \rightarrow 0$ o erro tende a zero.

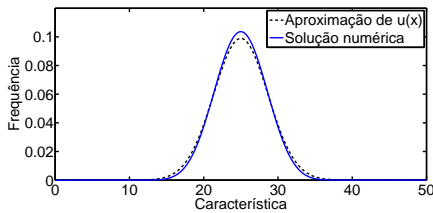


Figura 8: Comparação entre aproximação numérica e aproximação na forma $u(x) = u_0 e^{-(\epsilon x)^2}$. $\epsilon = 0,1$.

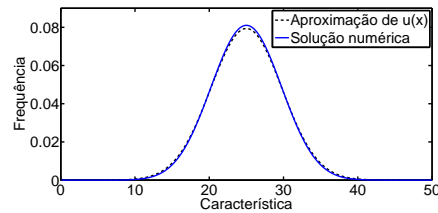


Figura 9: Comparação entre aproximação numérica e aproximação na forma $u(x) = u_0 e^{-(\epsilon x)^2}$. $\epsilon = 0,05$.

Na figura 7 apresentamos uma comparação entre os valores de N^* obtidos através de simulação numérica do modelo e aqueles obtidos através da aproximação fornecida pela equação (4.15). Nas figuras 8 e 9 apresentamos uma

comparação entre a solução obtida através da aproximação $u(x) = u_0 e^{-(\epsilon x)^2}$ e a solução numérica.

Voltando à inequação que fornece a condição para que $u^*(x)$ tenha um ponto de máximo em $x = 0$, temos que se $f_0 - 2f_0\epsilon^2 > N^*(\epsilon)$ então ocorre a seleção dos mais aptos. Desenvolvendo N^* em torno de $\epsilon = 0$, temos

$$N^* \cong f_0 (1 - \epsilon + \epsilon^2/2 + O(\epsilon^3)) \quad (4.16)$$

de onde obtemos a condição $\epsilon < 2/5$ para que ocorra seleção dos mais aptos. Voltando ao parâmetro adimensional δ , temos a condição $\delta > 5/2$, onde δ é a escala da variação da função de adaptação f em termos da variável adimensional. Se voltarmos à variável original temos que $\delta = \Delta/\sqrt{D}$, onde Δ é a escala de mudança de valor de adaptação de f e \sqrt{D} é a distância de modificação das características dos indivíduos na escala de tempo da reprodução. Assim, a condição para seleção dos mais aptos fica:

$$\Delta > 5\sqrt{D}/2 \quad (4.17)$$

que relaciona o coeficiente de mutação com a escala de mudança no valor de adaptação. Um coeficiente de mutação muito elevado impede que o processo de seleção natural selecione os indivíduos mais aptos.

5. Conclusões

O modelo de espaço de aspecto do tipo descrito pela equação (2.1) é uma forma distinta de abordagem da dinâmica de processos evolutivos e ainda deve ser explorado com mais profundidade de forma que os paralelos entre os resultados obtidos com o mesmo e os modelos clássicos sejam esclarecidos. Neste estudo em particular, o modelo parece confirmar o equilíbrio entre mutação e seleção previsto pelos modelos de frequências de genótipos quando a taxa de mutação não é excessivamente alta, apontando para uma falha na seleção dos mais aptos quando a taxa de mutação é muito elevada.

A relação em (4.17) foi obtida para um caso particular da função de adaptação e através do uso de aproximações. É esperado que, com modificações da relação de f com δ (isto é, a ordem com que f decai em função de δ) ou mudanças nos métodos de aproximação, existam modificações no formato específico da relação entre as escalas Δ e \sqrt{D} . Reservamos um estudo mais detalhado dessas relações para trabalhos futuros mas acreditamos que o mesmo

resultado geral deve permanecer, isto é, se a taxa de mutação for muito alta em comparação com a mudança nos valores de adaptação, não há seleção dos mais aptos.

Referências

- Assis, R. A. e Ferreira Jr., W. C. (2005). Um modelo diferencial de macroevolução: primeiros resultados. *Biomatemática*, 15:1–12.
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Dover Publications, New York.
- Gillespie, J. H. (1998). *Population Genetics: A Concise Guide*. John Hopkins University Press, London.
- Hartl, D. L. e Clark, A. G. (2007). *Principles of Population Genetics*. Sinauer Associates Publishers Inc., Sunderland.
- Kauffman, S. e Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology*, 128:11–45.
- Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge.
- Maynard-Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- Taylor, C. e Higgs, P. (2000). A population genetics model for multiple quantitative traits exhibiting pleiotropy and epistasis. *Journal of Theoretical Biology*, 224:241–247.