LIETUVOS MATEMATIKOS RINKINYS

Lietuvos matematikos rinkinys
ISSN: 0132-2818
ISSN: 2335-898X
arturas.stikonas@mif.vu.lt
Vilniaus Universitetas
Lituania

Weak approximations of Wright–Fisher equation

Mongirdaite, Gabriele; Mackevi#ius, Vigirdas
Weak approximations of Wright–Fisher equation
Lietuvos matematikos rinkinys, vol. 62 Ser. A, 2021
Vilniaus Universitetas, Lituania
Disponible en: https://www.redalyc.org/articulo.oa?id=692674114002
DOI: https://doi.org/10.15388/LMR.2021.25217

Articles

Weak approximations of Wright–Fisher equation

Wright–Fisher lygties silpnosios aproksimacijos

Gabriele Mongirdaite gabriele.mongirdaite@mif.vu.lt
Vilnius University, Lituania
https://orcid.org/0000-0003-3102-1587
Vigirdas Mackevičius vigirdas.mackevicius@mif.vu.lt
Vilnius University, Lituania
https://orcid.org/0000-0002-9792-735X

Lietuvos matematikos rinkinys, vol. 62
Ser. A, 2021

Vilniaus Universitetas, Lituania

Recepción: 10 Julio 2021
Publicación: 15 Diciembre 2021

DOI: https://doi.org/10.15388/LMR.2021.25217

Redalyc: https://www.redalyc.org/articulo.oa?id=692674114002

Abstract: We construct weak approximations of the Wright-Fisher model and illustrate their accuracy by simulation examples.
Keywords: Wright–Fisher model, simulation, weak approximation .
Summary: Sukonstruota silpnoji pirmos eile˙s aproksimacija stochastinei Wright–Fisher lygčiai. Pavyzdžiais iliustruojamas jos tikslumas.
Keywords: Wright–Fisher modelis, modeliavimas, silpnoji aproksimacija.

Introduction

We consider Wright–Fisher process defined by the stochastic differential equation

$$dX_t^x = \left(a - bX_t^x\right)dt + \sigma\sqrt{X_t^x\left(1 - X_t^x\right)}bB_t,\ X_0^x = x, \tag{1}$$

where B is a standard Brownian motion, $0 \,\square\, a \,\square\, b$, $\sigma > 0$, and $x \in [0, 1]$.

The Wright–Fisher model (Fisher 1930; Wright 1931) takes the values in the interval $[0, 1]$ and explicitly accounts for the effects of various evolutionary forces – random genetic drift, mutation, selection – on allele frequencies over time. This model can also accommodate the effect of demographic forces such as variation in population size through time and/or migration connecting populations [5].

In this note, we present a simple first-order weak approximation of the solution of Eq. (1) by discrete random variables that take two values at each approximation step.

Recall the definition of such an approximation. By a discretization scheme with time step $h > 0$ we mean any time-homogeneous Markov chain $\hat{X}^h = \left\{\hat{X}^h_{kh}, k = 0, 1, \dots\right\}$.

We say that a family of discretization schemes $\hat{X}^h$, $h > 0$, is a first-order weak approximation of the solution X^x of (1) in the interval $[0, T]$ if

$$\left|\mathrm{E}\, f\left(\hat{X}_T^h\right) - \mathrm{E}\, f\left(X_T^x\right)\right| \square\ Ch, \qquad h = \frac{T}{N} \square\ ho, \tag{2}$$

for a "sufficiently wide" class of functions $f : [0, 1] \to \mathrm{R}$ and some constants C and $h_0 > 0$ (depending on the function f), where N∈N. Note that because of the Markovity, the one-step approximation $\hat{X}_h^h$ completely defines (in distribution) a weak approximation $\hat{X}_{kh}^h$, $k = 0, 1, \ldots$ Thus, with some ambiguity, we also call it an approximation and denote it by $\hat{X}_h^x$, with x indicating its starting point.

In our context, we introduce the following "sufficiently wide" function class of infinitely differentiable functions with "not too fast" growing derivatives:

$$C_*^\infty[0, 1] := \left\{ f \in C^\infty[0, 1] : \limsup_{k\to\infty} \frac{1}{k!} \sup_{x\in[0,1]} \left| f^{(k)}(x) \right| < \infty \right\}.$$

We easily see that all functions from this class can be expanded by the Taylor series in the interval $[0, 1]$ around arbitrary $x_0 \in [0, 1]$ (which, in fact, converges on the whole real line R) and contain, for example, all polynomials and exponential functions.

Approximation

Let us first construct an approximation for the "stochastic" part of Wright–Fisher equation, that is, the solution S_t^x of Eq. (1) with $a = b = 0$. Similarly to [4] (see also [3]), we look for an approximation $\hat{X}_h^x$ as a two-valued discrete random variable taking values $x_{1,2} \in [0, 1]$ with probabilities $p_{1,2}$ such that

$$\mathrm{E}\left(\hat{S}_h^x - x\right) = 0, \ x \in [0, 1], \tag{3}$$

$$\mathrm{E}\left(\hat{S}_h^x - x\right)^2 = \sigma^2 x(1 - x)h + \mathrm{O}\left(h^2\right), \ x \in [0, 1], \tag{4}$$

$$|\mathrm{E}\left(\hat{S}_h^x - x\right)^3| = O\left(h^2\right), \ x \in [0, 1], \tag{5}$$

$$\mathrm{E}\left[\left(\hat{S}_h^x - x\right)^4 = O\left(h^2\right), \ x \in [0, 1]\right]. \tag{6}$$

By solving the equation system (3)–(4) with respect to x_1, x_2, p_1, p_2, we get the solution

$$x_1 = x + (1-x)\sigma^2 h - \sqrt{(x + (1-x)\sigma^2 h)(1-x)\sigma^2 h}, \ x \in [0, 1], \tag{7}$$

$$x_2 = x + (1-x)\sigma^2 h + \sqrt{(x + (1-x)\sigma^2 h)(1-x)\sigma^2 h}, \ x \in [0, 1] \tag{8}$$

with $p_{1,2} = \frac{x}{2x_{1,2}}$. It also satisfies conditions (5)–(6). However, for the values of x near 1, the values of x_2 a slightly greater than 1, which is unacceptable. We overcome this problem by using the symmetry of the solution of the stochastic part with respect to the point $\frac{1}{2}$; to be precise, $S_t^x \overset{d}{=} 1 - S_t^{1-x}$. Therefore, in the interval $[0, 1/2]$, we can use the values $x_{1,2}$ defined by (7)-(8), whereas in the interval $(1/2, 1)$, we use the values corresponding to the process $1 - \hat{S}_t^{1-x}$, that is,

$$\widehat{x_{1,2}} = \hat{x_{1,2}}(x, h) := 1 - x_{1,2}(1-x, h) = x - x\sigma^2 h \pm \sqrt{(1 - x + x\sigma^2 h)x\sigma^2 h} \tag{9}$$

with probabilities

with probabilities $\hat{p_{1,2}} = \frac{1-x}{2x_{1,2}(1-x, h)}$. Thus we obtain a correct (i.e., with values in $[0, 1]$) approximation $\hat{S}_h^x$ taking the values

$$\widetilde{x_{1,2}} := \begin{cases} x_{1,2}(x, h) \text{ With probabilities } p_{1,2} = \frac{x}{2x_{1,2}(x,h)}, \ x \in [0, 1/2], \\ 1 - x_{1,2}(1-x, h) \text{ with probabilities } p_{1,2} = \frac{1-x}{2x_{1,2}(1-x,h)}, \ x \in (1/2, 1] \end{cases}$$

Now for the initial equation (1), we obtain an approximation $\hat{X}_h^x$ by a simple "splitstep" procedure (again, see, e.g., [4] or [3]):

$$\hat{X}_h^x := \hat{S}_h^x e^{-bh} + \frac{a}{b}\left(1 - e^{-bh}\right). \tag{10}$$

Now we can state the following:

Theorem 1. *Let* $\hat{X}_t^x$ *be the discretization scheme defined by one-step approximation* (10). *Then* $\hat{X}_t^x$ *is a first-order weak approximation of equation* (1) *for functions* $f \in C_*^\infty[0, 1]$.

Backward Kolmogorov equation

The constructed approximation is in fact a so-called *potential* first-order weak approximation of Eq. (1) (for a definition, see, e.g., Alfonsi [1], Section 2.3.1). The proof that, indeed, it is a first-order weak approximation, is based on the following:

Theorem 2. *Let* $f \in C_*^\infty[0,1]$*The* $u(t,x) := \mathrm{E}\, f(X_t^x)$ *is a* C^∞ *function on* $[0,1]$x$\mathbb{R}$ *that solves the backward Kolmogorov equation*

$$\partial_t u(t,x) = Au(t,x),\ x \in [0,1],\ t \geq 0.$$

In particular,

$$\forall T > 0,\ \forall l, m \in N,\ \exists C_{l,m} : |\partial_t^l \partial_x^m u(t,x)| \leq C_{l,m}, t \in [0,T],\ x \in [0,1].$$

Such theorem is stated for $f \in C^\infty[0,1]$ in [1, Thm.6.1.12], based on the results of [2]. Our class of functions f is slightly narrower, but our proof of the theorem is significantly simpler and is based on the estimates of the moments of X_t^x, which show that they grow slower than factorials. The recurrent relations of the moments $E\left[(X_t^x)^k\right]$ show that they are infinitely differentiable with respect to t and x, which allows us infinitely differentiate the series

$$u(t,x) = Ef\left(X_t^x\right) = \sum_{k=0}^{\infty} c_k E\left[(X_t^x)^k\right]$$

termwise with respect to t and x, where $f(x) = \sum_{k=0}^{\infty} c_k x^k$ is the Taylor expansion of f.

Simulation examples

We illustrate our approximation for $f(x) = x^4$ and $f(x) = \exp\{-x\}$. Since we do not explicitly know the moments $E\exp\{-X_t^x\}$, we use the approximate equality $\exp\{-x\} \approx 1 - x + \frac{x^2}{2} - \frac{x^3}{6} + \frac{x^4}{24}$. In Figs. 1 and 2, we compare the moments $Ef(\hat{X}_t^x)$ and $Ef(x_t^x)$ as functions of t (left plots, $h = 0.001$) and as functions of discretization step h (right plots, $t = 1$). As expected, the approximations agree with exact values pretty well.

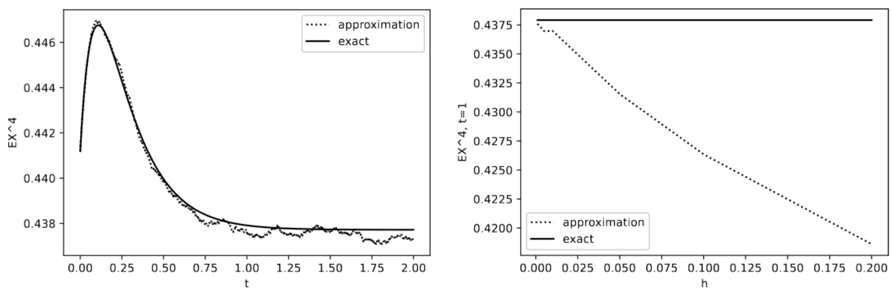

Fig. 1. Comparison of $\mathbb{E}f(\widehat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^4$: $x = 0.815$, $\sigma^2 = 0.5$, $a = 4$, $b = 5$, the number of iterations $N = 500.000$.

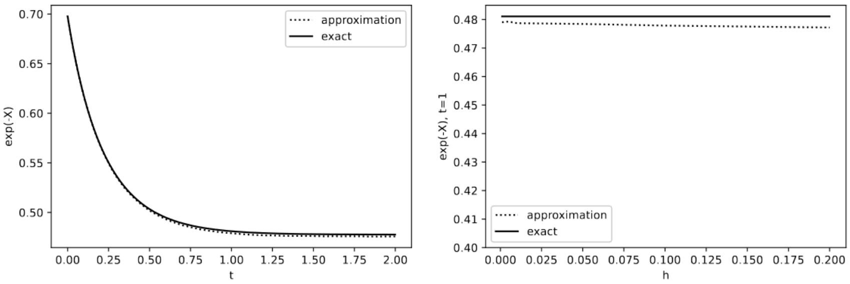

Fig. 2. Comparison of $\mathbb{E}f(\widehat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = \exp\{-x\}$: $x = 0.36$, $\sigma^2 = 0.6$, $a = 3$, $b = 4$, $N = 100.000$.

References

1. A. Alfonsi. *Affine Diffusions and Related Processes: Simulation, Theory and Applications*. Springer, 2015.
2. C. Epstein, R. Mazzeo. Wright–Fisher diffusion in one dimension. *SIAM J. Math. Anal.*, 42:568–608, 2010.
3. G. Lileika, V. Mackevičius. Weak approximation of CKLS and CEV processes by discrete random variables. *Lith. Math. J.*, 20(2):208–224, 2020.
4. V. Mackevičius. Weak approximation of CIR equation by discrete random variables. *Lith. Math. J.*, 51(3):385–401, 2011.
5. P. Tataru, M. Simonsen, T. Bataillon, A. Hobolth. Statistical inference in the Wright– Fisher model using allele frequency data. *Syst. Biol.*, 66(1):e30–e46, 2016.