Applied Mathematics and Computation 244 (2014) 432–446

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

More general problems on first-passage times for diffusion processes: A new version of the fptdApprox R package

P. Román-Román *, J.J. Serrano-Pérez, F. Torres-Ruiz

Departamento de Estadística e Investigación Operativa, Universidad de Granada, Spain

ARTICLE INFO

Keywords:
Diffusion processes
Conditioned and unconditioned first-passage times
Volterra integral equations
FPTL function
R

ABSTRACT

Version 1.2 of the `fptdApprox` R package implemented a general heuristic strategy for the efficient application of numerical schemes aimed at solving Volterra integral equations whose solutions are first-passage-time (f.p.t.) density functions for a diffusion process.

This paper presents a new version of this package, developed with the goal of addressing interesting f.p.t. problems that cannot be solved with the previous version. More specifically, it addresses f.p.t. problems for diffusion processes defined by means of not-closed-form time expressions (which appear in application fields such as tumor growth or the evolution of economic variables), and unconditioned f.p.t. problems for diffusion processes with non-degenerate initial distributions (of great interest in the study of population growth in Biology and Ecology).

Here, the functionality of this new version is described, with a focus on its numerical and computational aspects.

Moreover, by means of several examples, we show how the package can be used to solve problems such as those mentioned above.

1. Introduction

Stochastic processes, and diffusion processes in particular, are widely used in the modeling of phenomena arising in many disciplines. Two main aspects in the application of diffusion processes are the simulation of sample paths and the estimation from either real or simulated data. As a result, the computational treatment of these problems has undergone remarkable development.

Regarding the simulation of sample paths for stochastic process solutions to stochastic differential equations, we must name the R Sim.DiffProc package [3] and, for simulation as well as estimation, the book by Iacus [7], in which both problems are tackled by providing R-efficient codes for simulation schemes as well as for some estimation methods based on discretely sampled observations from the models.

Yet another problem of great interest is the first-passage time (f.p.t.) through time dependent boundaries, which arises in a variety of applications ranging from finance to biology, physics, or psychology (see [16,8,9], and examples cited therein). The study of f.p.t. problems has been addressed by means of analytical, numerical and simulation methods (see [10,15] for a review of these results). This paper focuses on numerical methods to approximate f.p.t. densities through the use of a R package.

* Corresponding author. Postal address: Dpto. de Estadística e I.O., Facultad de Ciencias, Avda. Fuentenueva s/n, 18071 Granada, Spain.
E-mail addresses: proman@ugr.es (P. Román-Román), jjserra@ugr.es (J.J. Serrano-Pérez), fdeasis@ugr.es (F. Torres-Ruiz).

http://dx.doi.org/10.1016/j.amc.2014.06.111
0096-3003/

Let $\{X(t);\ t_0 \leqslant t \leqslant T\}$ be a diffusion process defined on a real interval I, with infinitesimal moments $A_1(x,t)$ and $A_2(x,t)$, and let $S(t)$ be a derivable function. We assume that the extremes of the diffusion interval I are unattainable boundaries. The first-passage time (f.p.t.) of the process through boundary $S(t)$, provided $X(t_0) = x_0$ (in the following, a conditioned f.p.t.), is defined as the random variable

$$T_{S(t),x_0} = \begin{cases} \underset{t \geqslant t_0}{Inf}\{t:\ X(t) > S(t) | X(t_0) = x_0\}, & \text{if } x_0 < S(t_0) \\ \underset{t \geqslant t_0}{Inf}\{t:\ X(t) < S(t) | X(t_0) = x_0\}, & \text{if } x_0 > S(t_0). \end{cases} \tag{1}$$

This approach includes the cases of a degenerate initial distribution and of a non-degenerate initial distribution if an initial value is fixed.

The density function of $T_{S(t),x_0}$, $g(S(t),t|x_0,t_0)$, is the solution to the Volterra integral equation of the second kind

$$g(S(t),t|x_0,t_0) = \rho\left\{-2\Psi(S(t),t|x_0,t_0) + 2\int_{t_0}^{t} g(S(\tau),\tau|x_0,t_0)\Psi(S(t),t|S(\tau),\tau)d\tau\right\}, \tag{2}$$

where $\rho = \mathrm{Sgn}(S(t_0) - x_0)$,

$$\Psi(S(t),t|y,\tau) = \frac{1}{2}f(S(t),t|y,\tau)\left[S'(t) - A_1(S(t),t) + \frac{3}{4}\left.\frac{\partial A_2(x,t)}{\partial x}\right|_{x=S(t)}\right] + \frac{1}{2}A_2(S(t),t)\left.\frac{\partial f(x,t|y,\tau)}{\partial x}\right|_{x=S(t)}$$

and $f(x,t|y,s)$ is the transition probability density function of the process (see [10] for a detailed study and a review of this subject in the case of homogeneous diffusion processes and [4] for some results in the non-homogeneous case).

The `fptdApprox` R package (described in [13] and available in [12]) implements a general heuristic strategy for the efficient application of the following numerical scheme (proposed by [2] on the basis of the composite trapezoid method):

$$\begin{aligned} &g(S(t_0+h),t_0+h|x_0,t_0) = -2\rho\Psi(S(t_0+h),t_0+h|x_0,t_0), \\ &g(S(t_0+kh),t_0+kh|x_0,t_0) = \rho\left\{-2\Psi(S(t_0+kh),t_0+kh|x_0,t_0) + 2h\sum_{j=1}^{k-1} g(S(t_0+jh),t_0+jh|x_0,t_0)\right. \\ &\qquad\left.\times\Psi(S(t_0+kh),t_0+kh|S(t_0+jh),t_0+jh)\right\} \quad k = 2,3,\ldots, \end{aligned} \tag{3}$$

in order to approximate the solution of the Volterra integral equation (2).

This strategy was based on information provided by the FPTL function about the location of the first-passage-time variable, which results in (see [11,13] for more details):

- An initial time (after t_0) for the application of the numerical scheme if $g(S(t),t|x_0,t_0)$ has a negligible value between t_0 and this time.
- Potential final times for the application of the numerical scheme as an alternative stopping rule to that the cumulative integral of the approximate density is almost 1.
- Suitable integration steps in appropriate time subintervals of the interval $[t_0,T]$.

In order to approximate $g(S(t),t|x_0,t_0)$ with version 1.2 of the package, the following steps had to be carried out (see Fig. 1 and [13] for more details):

(1) Defining the diffusion process.
(2) Evaluating the FPTL function.
(3) Obtaining information from the FPTL function in order to locate the f.p.t. variable.
(4) Approximating the f.p.t. density function for suitable subintervals and integration steps.

There are, however, situations of great interest in various fields of application in which this version of the package does not allow to address the approximation of the f.p.t. density function, such as

- Diffusion processes with exogenous factors that do not have a closed-form expression:
  - Gompertz-type diffusion process in the context of tumor growth. In Albano et al. [1] a modified Gompertz diffusion process is considered in the context of tumor growth. This process includes in its drift term a time-dependent function representing the effect of a therapy. It is defined as $\{X^C(t);\ t_0 \leqslant t \leqslant T\}$, takes values on $\mathbb{R}^+$ and is characterized by infinitesimal moments

$$\begin{aligned} A_1^C(x,t) &= [\alpha - C(t)]x - \beta x\log x \\ A_2^C(x,t) &= \sigma^2 x^2, \end{aligned} \tag{4}$$