



On the empirical estimator of the boundary in inverse first-exit problems

Sercan Gür¹ · Klaus Pötzelberger¹

Received: 1 July 2019 / Accepted: 18 April 2020 / Published online: 6 May 2020
© The Author(s) 2020

Abstract

First-exit problems for the Brownian motion ($W(t)$) or general diffusion processes, have important applications. Given a boundary $b(t)$, the distribution of the first-exit time τ has to be computed, in most cases numerically. In the inverse first-passage-time problems, the distribution of τ is given and the boundary b has to be found. The boundary and the density of τ satisfy a Volterra integral equation. Again numerical methods approximate the solution b for given distribution of τ . We propose and analyze estimators of b for a given sample τ_1, \dots, τ_n of first-exit times. The first estimator, the empirical estimator, is the solution of a stochastic version of the Volterra equation. We prove that it is strongly consistent and we derive an upper bound for its asymptotics convergence rate. Finally, this estimator is compared to a Bayesian estimator, which is based on an approximate likelihood function. Monte Carlo experiments suggest that the empirical estimator is simple, computationally manageable and outperforms the alternative procedure considered in this paper.

Keywords Bayes estimator · Empirical estimator · Inverse first passage times · Markov chain Monte Carlo

1 Introduction

The analysis of first-passage problems of diffusion processes and especially of the Brownian motion is a rich field in applied probability with important applications in various areas such as mathematical finance, statistics, physics, engineering or biology. For instance, in mathematical finance, a default event can be modelled by the first time a stochastic process representing firm value crosses a certain possibly time-varying barrier or a barrier option can be exercised if the underlying value process reaches a

✉ Sercan Gür
guer.sercan@gmail.com

¹ Institute of Statistics and Mathematics, WU Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria

predefined boundary. In biology, extinction of a population can be described by the event that the number of individuals first passes a threshold value (for more applications in biology, we refer the reader to Ricciardi et al. 1999). Another area of application arises in statistics, in particular in sequential analysis and change-point problems.

Let $(W(t))_{t \geq 0}$ denote a standard Brownian motion, i.e. a Gaussian process with continuous paths, $W(0) = 0$, $\mathbb{E}(W(t)) = 0$ and $\mathbb{E}(W(s)W(t)) = s \wedge t$. Let $b : [0, \infty) \rightarrow \mathbb{R}$ denote a function with $b(0) \geq 0$, the upper boundary. Define the first-exit time

$$\tau = \inf\{t > 0 \mid W(t) \geq b(t)\}. \quad (1)$$

τ is a stopping time. Denote the distribution of τ by F . Given regularity conditions, F is absolutely continuous with a density f , which is continuous and strictly positive on $(0, \infty)$.

The direct first-passage-time problem identifies F for given b . The inverse problem computes b for given F or f . The problem of finding the boundary function b is of importance in many fields, and the application areas are similar to the ones involving the direct first-passage time problems, e.g. mathematical finance, in particular credit risk modelling or in biology, in neural activity (for more details see Abundo 2015). For both problems an extensive literature exists. Since for only few boundaries the distribution F of τ can be computed in closed form, numerical procedures, approximate solutions of corresponding partial differential equations, or Monte Carlo methods are necessary. For details we refer to Durbin (1971), Lerche (1986), Salminen (1988), Novikov et al. (1999) or Pötzelberger and Wang (2001).

Inverse problems are often based on a Volterra Integral equation, the so-called master equation, for which (b, F) is a solution. Let $z \geq b(t)$. Then

$$\bar{\Phi} \left(\frac{z}{\sqrt{t}} \right) = \int_0^t \bar{\Phi} \left(\frac{z - b(u)}{\sqrt{t-u}} \right) dF(u), \quad (2)$$

with $\bar{\Phi} = 1 - \Phi$ the survival function of the standard normal distribution. $z = b(t)$ and the differentiation of (2) in $z = b(t)$ give

$$\bar{\Phi} \left(\frac{b(t)}{\sqrt{t}} \right) = \int_0^t \bar{\Phi} \left(\frac{b(t) - b(u)}{\sqrt{t-u}} \right) dF(u), \quad (3)$$

$$\phi \left(\frac{b(t)}{\sqrt{t}} \right) \frac{1}{\sqrt{t}} = \int_0^t \phi \left(\frac{b(t) - b(u)}{\sqrt{t-u}} \right) \frac{1}{\sqrt{t-u}} dF(u). \quad (4)$$

See Peskir (2002), Zucca and Sacerdote (2009) or Abundo (2015) for thorough discussion of the inverse problem.

If f is given, an approximation of $b(t)$ on $\{t_i \mid t_i = hi\}$ is the solution of the system of equations

$$\bar{\Phi} \left(\frac{b(t_i)}{\sqrt{t_i}} \right) = \sum_{j=1}^i \bar{\Phi} \left(\frac{b(t_i) - b(t_j)}{\sqrt{t_i - t_j}} \right) f(t_j) h, \quad (i = 1, \dots, n). \quad (5)$$

In this paper we analyze the statistical inverse first-passage-time problem: Given a sample τ_1, \dots, τ_n of independent first-exit times, we approximate the unknown boundary $b(t)$ by an estimator $\hat{b}_n(t)$. We propose the empirical estimator, which is the solution of (2), when F is replaced by the empirical distribution \hat{F}_n .

This paper is organized as follows. In Sect. 2 we prove that the empirical estimator is strongly consistent with rate $o((\log n + \eta \log \log n)^{1/2}n^{-1/2})$ for every $\eta > 1/2$, uniformly on $t \in [0, T]$, for all $T > 0$.

We compare the performance of the empirical estimator to an approximate conditional likelihood method, namely a Bayes estimator. The approximate conditional likelihood is the density of the first-exit time τ , when the boundary b is approximated by a piecewise linear boundary b_m , i.e. a boundary that is linear on intervals $[t_{i-1}, t_i]$, with $0 = t_0 < t_1 < \dots < t_m = T$ a partition of $[0, T]$. For b_m , the density of τ , given $W(t_1), \dots, W(t_m)$, can be computed in closed form. In Sect. 3 we compute this approximate and conditional density. Section 4 concludes with the results of Monte Carlo experiments for the empirical estimator and a Bayes estimator derived from the approximate likelihood.

2 Empirical estimator

Let τ_1, \dots, τ_n be an i.i.d. sample of first-exit times corresponding to the boundary b . Note that $\tau_i = \infty$ if the Brownian motion $(W(t))$ never crosses the boundary b . We denote the empirical distribution of $\tau_i, i \in \{1, \dots, n\}$ by \hat{F}_n . The empirical estimator $\tilde{b}_n(t)$ of the boundary $b(t)$ is the solution of

$$\bar{\Phi} \left(\frac{\tilde{b}_n(t)}{\sqrt{t}} \right) = \int_0^t \bar{\Phi} \left(\frac{\tilde{b}_n(t) - \tilde{b}_n(u)}{\sqrt{t-u}} \right) d\hat{F}_n(u). \tag{6}$$

The empirical estimator is consistent. Equation (6) has a solution for all t . However, it is convenient to solve a corresponding system of equations at the sample. Let $\tau_{(1)} \leq \tau_{(2)} \leq \dots \leq \tau_{(n)}$ denote the order statistics of the sample. Note that for finite order statistics $\tau_{(i)}, \tau_{(i)} < \tau_{(i+1)}$ a.s.

We define the estimator \hat{b}_n at $\tau_{(i)}$, i.e. $\hat{b}_n(\tau_{(1)}), \dots, \hat{b}_n(\tau_{(n)})$, as the solution of

$$\begin{aligned} \bar{\Phi} \left(\frac{\hat{b}_n(\tau_{(1)})}{\sqrt{\tau_{(1)}}} \right) &= \frac{1}{2n} \\ \bar{\Phi} \left(\frac{\hat{b}_n(\tau_{(k)})}{\sqrt{\tau_{(k)}}} \right) &= \sum_{i=1}^{k-1} \bar{\Phi} \left(\frac{\hat{b}_n(\tau_{(k)}) - \hat{b}_n(\tau_{(i)})}{\sqrt{\tau_{(k)} - \tau_{(i)}}} \right) \frac{1}{n}. \end{aligned}$$

In case not all τ_i 's are finite, we define $\hat{b}_n(\infty) = \infty$. For $t \notin \{\tau_1, \dots, \tau_n\}$ we interpolate $\hat{b}_n(t)$ linearly.

Theorem 1 *Let b be continuously differentiable with $b(0) > 0$. The empirical estimator is strongly consistent:*

Let for $\eta > 1/2$, $\epsilon_n = (\log n + \eta \log \log n)^{1/2} n^{-1/2}$. Then for all $0 < T$,

$$P \left(\limsup_{n \rightarrow \infty} \sup_{\theta \leq T} |\hat{b}_n(\theta) - b(\theta)| / \epsilon_n = 0 \right) = 1. \quad (7)$$

Proof The empirical estimator can be considered as a discretization scheme of the master equation with the order statistics $\tau_{(i)}$ as random knots.

Zucca and Sacerdote (2009) proved the consistency of the Euler scheme for the (deterministic) master equation. We follow their lines, with the necessary modifications indicated. Define for $T > 0$ fixed and knots $0 = t_0 < t_1 < \dots < t_n = T$ the solution of the Euler scheme (5) by $b^*(t_k)$ and the local consistency error by

$$\begin{aligned} \delta(h, t_k) &= \int_0^{t_k} \bar{\Phi} \left(\frac{b(t_k) - b(u)}{\sqrt{t_k - u}} \right) dF(u) \\ &\quad - \sum_{1 \leq j < k} \bar{\Phi} \left(\frac{b(t_k) - b(t_j)}{\sqrt{t_k - t_j}} \right) (F(t_j) - F(t_{j-1})), \end{aligned}$$

with $h = \max_{1 \leq j < k} (t_j - t_{j-1})$. The proof of Theorem 6.2. in Zucca and Sacerdote (2009) shows that there is a constant $\tilde{c} > 0$ (depending on b and T only), such that for all i ,

$$|b(t_i) - b^*(t_i)| \leq \tilde{c} \delta(h, t_i).$$

Define for $\theta > 0$,

$$Z_j(\theta) = \bar{\Phi} \left(\frac{b(\theta) - b(\tau_{(j)})}{\sqrt{\theta - \tau_{(j)}}} \right) I_{[0, \theta]}(\tau_{(j)})$$

and

$$\begin{aligned} \delta_n(\theta) &= \int_0^{t_k} \bar{\Phi} \left(\frac{b(\theta) - b(u)}{\sqrt{\theta - u}} \right) dF(u) - \sum_{\tau_{(j)} \leq \theta} \bar{\Phi} \left(\frac{b(\theta) - b(\tau_{(j)})}{\sqrt{\theta - \tau_{(j)}}} \right) \frac{1}{n} \\ &= \sum_{j \leq n} (\mathbb{E}(Z_j(\theta)) - Z_j(\theta)). \end{aligned}$$

Since $|Z_j(\theta)| \leq 1$ for all j , Hoeffding's inequality gives for all $\epsilon > 0$,

$$P(|\delta_n(\theta)| > \epsilon) \leq 2e^{-2n\epsilon^2}.$$

For $\epsilon_n = (\log n + \eta \log \log n)^{1/2} n^{-1/2}$ we get for $\eta' = 2\eta - 1 > 0$ and fixed θ ,

$$P(|\delta_n(\theta)| > \epsilon_n) \leq \frac{1}{n^2 \log n^{1+\eta'}}.$$

Then

$$P(\max_{1 \leq k \leq n} |\delta_n(\tau_{(k)})| > \epsilon_n) \leq \frac{1}{n \log n^{1+\eta'}}. \tag{8}$$

Since

$$\sum_{n=1}^{\infty} \frac{1}{n \log n^{1+\eta'}} < \infty,$$

the Theorem of Borel–Cantelli implies (7). □

Remark In case of censoring at T , the empirical estimator is still consistent, if a consistent estimator of $F(T)$ is available. If this estimator is even strongly consistent with rate ϵ'_n , then the empirical estimator is strongly consistent with rate $\epsilon'_n \vee \epsilon_n$, with ϵ_n defined in Theorem 1.

Let us briefly comment on the asymptotic distribution of the residuals $(\hat{\epsilon}_n(t))$ defined as

$$\hat{\epsilon}_n(t) = \sqrt{n}(\hat{b}_n(t) - b(t)). \tag{9}$$

Denote by $\hat{U}_n(t)$ the empirical process

$$\hat{U}_n(t) = \sqrt{n}(\hat{F}_n(t) - F(t)).$$

$(\hat{U}_n(t)) \rightarrow (U^F(t))$, with $(U^F(t))$ a Brownian bridge, a Gaussian process with continuous paths, $\mathbb{E}(U(t)) = 0$ and $\text{Cov}(U(s), U(t)) = F(s \wedge t) - F(s)F(t)$. There are processes (\hat{U}_n^*) and a Brownian bridge (U^{F*}) with the same distributions as (\hat{U}_n) and (U^F) , such that with probability 1, $\|\hat{U}_n^* - U^{F*}\| \rightarrow 0$ (see Shorack and Wellner 2009). To simplify the exposition, and since we are interested in the asymptotic distribution of the residuals only, we may assume that a.s. $\|\hat{U}_n - U^F\| \rightarrow 0$.

We have

$$\begin{aligned} \bar{\phi} \left(\frac{b(t) + n^{-1/2}\hat{\epsilon}_n(t)}{\sqrt{t}} \right) &= \int_0^t \bar{\phi} \left(\frac{b(t) - b(s) + n^{-1/2}(\hat{\epsilon}_n(t) - \hat{\epsilon}_n(u))}{\sqrt{t-u}} \right) dF(u) \\ &\quad + \frac{1}{\sqrt{n}} \int_0^t \bar{\phi} \left(\frac{b(t) - b(s) + n^{-1/2}(\hat{\epsilon}_n(t) - \hat{\epsilon}_n(u))}{\sqrt{t-u}} \right) d\hat{U}_n(u), \\ \bar{\phi} \left(\frac{b(t)}{\sqrt{t}} \right) - \phi \left(\frac{b(t)}{\sqrt{t}} \right) \frac{\hat{\epsilon}_n(t)}{\sqrt{n}\sqrt{t}} &= \int_0^t \bar{\phi} \left(\frac{b(t) - b(u)}{\sqrt{t-u}} \right) dF(u) \\ &\quad + \frac{1}{\sqrt{n}} \int_0^t \bar{\phi} \left(\frac{b(t) - b(u)}{\sqrt{t-u}} \right) d\hat{U}_n(u) \\ &\quad - \frac{1}{\sqrt{n}} \int_0^t \phi \left(\frac{b(t) - b(u)}{\sqrt{t-u}} \right) \frac{\hat{\epsilon}_n(t) - \hat{\epsilon}_n(u)}{\sqrt{t-u}} dF(u) + o \left(\frac{1}{\sqrt{n}} \right). \end{aligned}$$

Therefore

$$\phi \left(\frac{b(t)}{\sqrt{t}} \right) \frac{\hat{\epsilon}_n(t)}{\sqrt{t}} = \int_0^t \phi \left(\frac{b(t) - b(u)}{\sqrt{t-u}} \right) \frac{\hat{\epsilon}_n(t) - \hat{\epsilon}_n(u)}{\sqrt{t-u}} dF(u)$$

$$- \int_0^t \bar{\Phi} \left(\frac{b(t) - b(u)}{\sqrt{t - u}} \right) d\hat{U}_n(u) + o(1).$$

Assume that for all t , $\hat{\epsilon}_n(t)$ would converge to a limit $\epsilon(t)$. The process $(\epsilon(t))$ would solve

$$\begin{aligned} \phi \left(\frac{b(t)}{\sqrt{t}} \right) \frac{\epsilon(t)}{\sqrt{t}} &= \int_0^t \phi \left(\frac{b(t) - b(u)}{\sqrt{t - u}} \right) \frac{\epsilon(t) - \epsilon(u)}{\sqrt{t - u}} dF(u) \\ &\quad - \int_0^t \bar{\Phi} \left(\frac{b(t) - b(u)}{\sqrt{t - u}} \right) dU^F(u). \end{aligned}$$

With (4), $(\epsilon(t))$ would solve the stochastic linear Abel integral equation

$$\int_0^t \phi \left(\frac{b(t) - b(u)}{\sqrt{t - u}} \right) \frac{\epsilon(u)}{\sqrt{t - u}} dF(u) = - \int_0^t \bar{\Phi} \left(\frac{b(t) - b(u)}{\sqrt{t - u}} \right) dU^F(u). \tag{10}$$

However, there is no “classical” solution $(\epsilon(t))$ of (10). To see this, let $b(t) = b$ be constant. Recall that the density f of τ is continuous and bounded. Then Eq. (10) is

$$\int_0^t \frac{\epsilon(u) f(u)}{\sqrt{t - u}} du = - \sqrt{\frac{\pi}{2}} U^F(t). \tag{11}$$

Applying the Abel transform, (see Gorenflo and Vessella 1991), we get

$$\int_0^t \epsilon(u) f(u) du = H_t := - \frac{1}{\sqrt{2\pi}} \int_0^t \frac{U^F(u)}{\sqrt{t - u}} du. \tag{12}$$

The process (H_t) is a Gaussian process with $\mathbb{E}(H_t) = 0$ and for $s \leq t$,

$$K(s, t) = \text{Cov}(H_s, H_t) = 2 \int_0^s \left(\sqrt{\frac{s - u}{t - u}} + \sqrt{\frac{t - u}{s - u}} \right) F(u) du - g(s)g(t), \tag{13}$$

with

$$g(s) = \int_0^s \frac{1}{\sqrt{s - u}} F(u) du.$$

Remark 1. If $\int_0^t \epsilon(u) f(u) du$ would be of bounded variation with $\epsilon(t) f(t) = \frac{d}{dt} H_t$, then

$$\begin{aligned} \text{Var}(\epsilon(t) f(t)) &= \frac{\partial^2}{\partial s \partial t} K(s, t) |_{t=s} \\ &= \int_0^s \frac{1}{\sqrt{(s - u)(t - u)}} f(v) du - g'(s)g'(t) \end{aligned}$$

$$= \int_0^t \frac{1}{t-u} f(u) du - \left(\int_0^t \frac{1}{\sqrt{t-u}} f(u) du \right)^2 = \infty.$$

2. Note that Lévy’s theorem on the modulus of continuity implies that the Brownian bridge U^F has modulus of continuity $\sqrt{2h \log(1/h)}$, i.e. it holds

$$\lim_{h \rightarrow 0^+} \sup_{|s| \leq h} \frac{|U^F(t+s) - U^F(t)|}{\sqrt{2h \log(1/h)}} = 1, \text{ a.s. .}$$

It follows that the modulus of continuity of (H_t) is $ch\sqrt{\log(1/h)}$, with $c > 0$ a constant.

Proposition 1 *Let $b(t)$ be constant. The process $\left(\int_0^t \hat{\epsilon}_n(u) dF(u)\right)$ of integrated weighted residuals converges in distribution to the centered Gaussian process (H_t) defined by (12).*

3 Approximate likelihood

For piecewise linear boundaries the following conditional boundary crossing probability allows the computation of an approximate conditional likelihood function. Let b_m be continuous and linear on intervals $[t_i, t_{i+1}]$, where $0 = t_0 < t_1 < \dots < t_m = T < \infty$. Let τ^m denote the corresponding first-exit time and $W^m = (W(t_i))_{i \leq m}$ a discrete Brownian motion and $w^m = (w_1, \dots, w_m) \in \mathbb{R}^m$. Wang and Pötzelberger (1997) prove that

$$P(\tau^m > T) = \mathbb{E}(v_m(W(t_1), \dots, W(t_m), T)), \tag{14}$$

with

$$v_m(w_1, \dots, w_m, T) = \prod_{i=1}^m \left(1 - \exp \left\{ -\frac{2(b_m(t_{i-1}) - w_{i-1})(b_m(t_i) - w_i)}{\Delta t_i} \right\} \right) \times I_{\{w_i < b_m(t_i)\}}. \tag{15}$$

Let $f(t | b_m)$ denote the density of τ^m , the first-exit time for the boundary b_m and $f(t | b_m, w^m)$ is the conditional density of τ^m given $(W^m = w^m)$. Then $f(t | b_m) = \mathbb{E}(f(t | b_m, W^m))$.

Proposition 2 *Define for given t_d : $t_u = t_{d+1}$ and $\Delta = t_u - t_d$. Let*

$$\mu_t = \frac{w_d(t_u - t) + w_u(t - t_d)}{\Delta}, \quad \sigma_t^2 = \frac{(t_u - t)(t - t_d)}{\Delta}. \tag{16}$$

1. For $t_d < t < t_u$ and $w_u \geq b(t_u)$,

$$f(t | b_m, w^m) = v_{m-1}(w_1, \dots, w_d, t_d)$$

$$\times \phi \left(\frac{b_m(t) - \mu_t}{\sigma_t} \right) \frac{\Delta^{1/2}(b(t_d) - w_d)}{(t_u - t)^{1/2}(t - t_d)^{3/2}} \quad (17)$$

2. For $t_d < t < t_u$ and $w_u < b(t_u)$, define

$$\begin{aligned} \eta_t &= \mu_t + \frac{2(b(t_u) - w_u)(t - t_d)}{\Delta}, \\ \theta_t &= \mu_t + \frac{2(b(t_u) - w_u)(t - t_d)}{\Delta} + \frac{2(b(t_d) - w_d)(t_u - t)}{\Delta}. \end{aligned}$$

Then $f(t \mid b_m, w^m) = v_{m-1}(w_1, \dots, w_d, t_d) g(t \mid b_m, w^m)$, where

$$\begin{aligned} g(t \mid b_m, w^m) &= \phi \left(\frac{b_m(t) - \eta_t}{\sigma_t} \right) \\ &\times \left(\frac{\Delta^{1/2}(b(t_d) - w_d)}{2(t_u - t)^{1/2}(t - t_d)^{3/2}} + \frac{\Delta^{1/2}(b(t_u) - w_u)}{2(t_u - t)^{3/2}(t - t_d)^{1/2}} \right) \\ &+ \exp \left[\frac{2(b(t_d) - w_d)(b(t_u) - w_u)}{\Delta} \right] \phi \left(\frac{b_m(t) - \theta_t}{\sigma_t} \right) \\ &\times \left(\frac{\Delta^{1/2}(b(t_d) - w_d)}{2(t_u - t)^{1/2}(t - t_d)^{3/2}} - \frac{\Delta^{1/2}(b(t_u) - w_u)}{2(t_u - t)^{3/2}(t - t_d)^{1/2}} \right) \quad (18) \end{aligned}$$

Proof Let $\tau^m = t \in [t_d, t_u]$. Conditional on $W(t_d) = w_d$ and $W(t_u) = w_u$, $(W(s))_{t_d \leq s \leq t_u}$ is a Brownian bridge, for which the crossing probabilities are given in closed form. The conditional distribution of $W(t)$ is Gaussian with parameters (16). To compute $P(\tau^m > t \mid \tau^m \in [t_d, t_u])$, condition on $W(t) = v$ with $v < b(t)$. There is no crossing in $[t_d, t]$ and a crossing in $[t, t_u]$. Note that in case $W(t_u) \geq b(t_u)$ the latter conditional probability is 1. Taking expectation w.r.t. $W(t)$ and finally the derivative w.r.t. t gives (17) and (18). \square

Approximate likelihood inference replaces the exact likelihood function by the approximate one, i.e. the boundary b is approximated by a piecewise linear boundary b_m . Estimates for errors, especially on $|P(\tau > t) - P(\tau^m > t)|$ are derived in Pötzelberger and Wang (2001), Borovkov and Novikov (2005), Zucca and Sacerdote (2009) and Pötzelberger (2012), among others.

4 Monte Carlo experiments

Monte Carlo simulation experiments were performed to evaluate the performance of the empirical estimator for finite sample sizes. Since for the Bayes estimator no theoretical result on its properties is available, the Monte Carlo experiments can indicate whether likelihood-based methods have the potential to outperform the empirical estimator. We estimate four boundaries—a constant boundary, a linear increasing, a linear decreasing and a Daniel's boundary—for which the first-exit time distribution is known in closed form, on $[0, T]$ with $T = 1$. The fifth boundary corresponds to exponentially distributed first-exit times (see Abundo 2015).

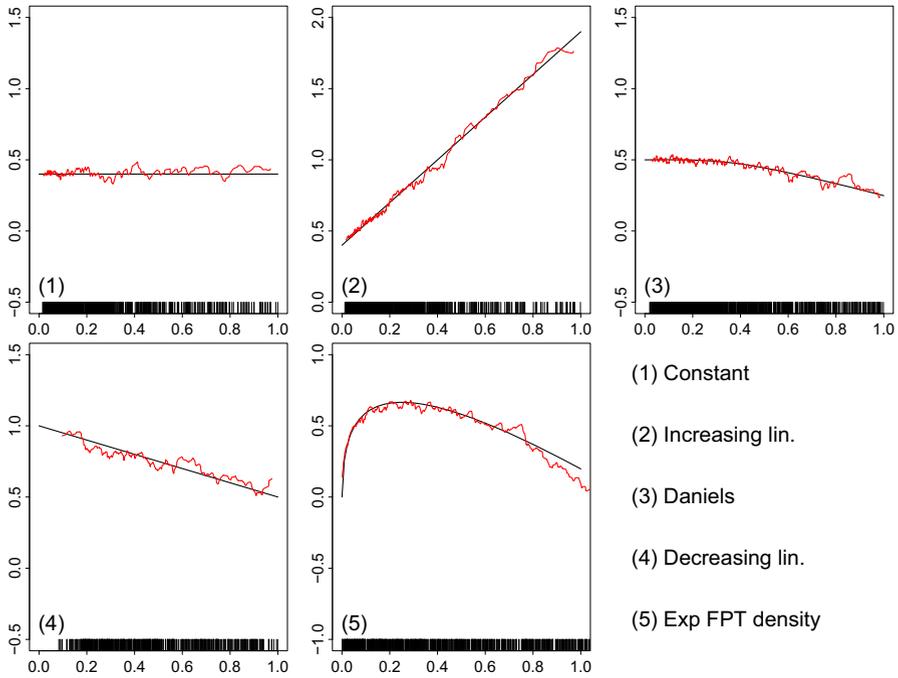


Fig. 1 Empirical estimator, $n = 10^3$. The true boundaries are shown by the black solid lines. The estimated boundaries are shown by the red solid lines (color figure online)

Table 1 Empirical estimator: MISE for $K = 100$ replications

	$n = 10^2$	$n = 10^3$	$n = 10^4$
Constant	0.0280	0.00528	0.000644
Linear (Inc.)	0.0083	0.00328	0.000534
Daniels	0.0145	0.00232	0.000393
Linear (Dec.)	0.0216	0.00388	0.000790
Exp. FPT Den.	0.0266	0.00382	0.000480

The results for the empirical estimator are given in Fig. 1. The mean-integrated-squared errors reported in Tables 1 and 2 are an estimate of $\int_0^T (\hat{b}_n(t) - b(t))^2 dt$. For the empirical estimator, we generate $K = 100$ samples of first-exit times of size n . For each sample

$$\sum_{i=1}^n (\hat{b}_n(\tau_{(i)}) - b(\tau_{(i)}))^2 (\tau_{(i)} - \tau_{(i-1)}) \tag{19}$$

(with $\tau_{(0)} = 0$) is computed. The MISE is the mean over these $K = 100$ samples.

We compare the performance of the empirical estimator to an approximate Bayes estimator: Let $t_i = Ti/m$. The boundary $b(t)$ is approximated by $b_m(t)$, which is linear on the intervals $[t_{i-1}, t_i]$. For parameter $\mathbf{b} = (b(0), b(t_1), \dots, b(t_m))$ and $\mathbf{w} = (w_1, \dots, w_m) \in \mathbb{R}^m$, the conditional approximate density of τ_i is given by (17) (if

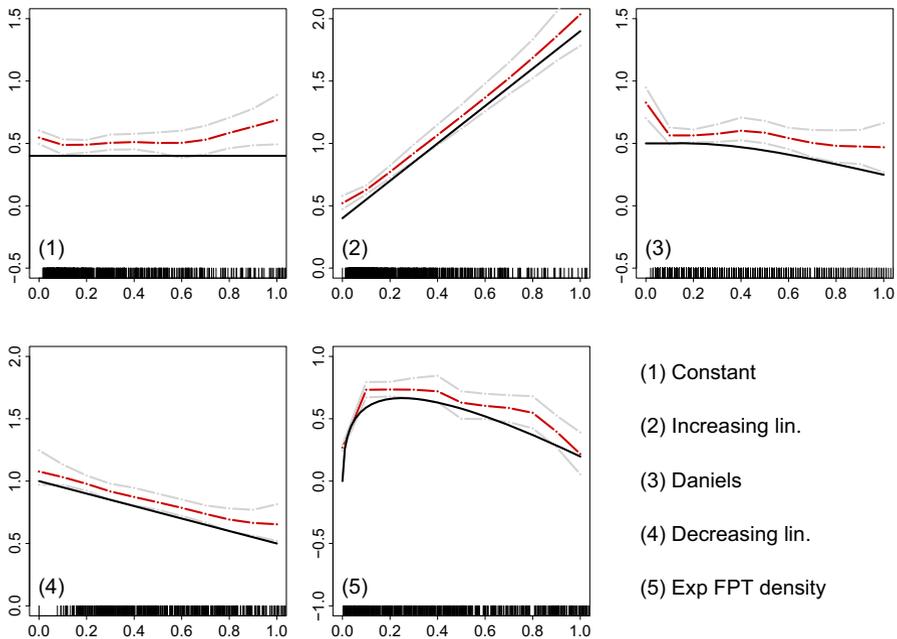


Fig. 2 Bayes estimator, $n = 10^3$. The true boundaries are shown by the black solid lines. The posterior means are shown by the red dashed lines together with 95%-credible intervals in gray (color figure online)

Table 2 Bayes—posterior mean: MISE for $K = 100$ replications

	$n = 10^2$	$n = 10^3$
Constant	0.0409	0.0174
Linear (Inc.)	0.0698	0.0148
Daniels	0.0379	0.0252
Linear (Dec.)	0.0443	0.0107
Exp. FPT Den.	0.0532	0.0262

$w_u \geq b(t_u)$), (18) (if $w_u < b(t_u)$) or (15) (if $\tau \geq T$). The product of these conditional approximate likelihoods is denoted by $L(\mathbf{b}, \mathbf{w}^1, \dots, \mathbf{w}^n, \tau_1, \dots, \tau_n)$. The problem can be formulated as a latent space model with a suitably chosen prior for the parameter \mathbf{b} . The Bayes estimator is the posterior mean of a sample of parameters \mathbf{b} generated through a Markov Chain Monte Carlo scheme:

- *Parameter \mathbf{b} .*
- *Data* A sample of i.i.d. first hitting times $\tau = (\tau_1, \dots, \tau_n)$.
- *Latent state space $\mathbf{W}^1, \dots, \mathbf{W}^n$.*
- *Prior* We assume that the slopes of \mathbf{b} follow a random walk with the double gamma shrinkage prior on the process variances, (see Bitto and Frühwirth-Schnatter 2019)

$$d_i = \frac{b_m(t_{i+1}) - b_m(t_i)}{t_{i+1} - t_i},$$

$$\begin{aligned}
 d_{t_i} &= d_{t_{i-1}} + \epsilon_{t_i}, \\
 \epsilon_{t_i} &\sim \mathcal{N}(0, \theta_j^2), \\
 \theta_j^2 &\sim \mathcal{G}(0.5, 0.5/\xi^2) \\
 \xi^2 &\sim \mathcal{G}(a^\xi, a^\xi \kappa^2).
 \end{aligned}$$

- *Computational details* The estimation is performed in JAGS (see Plummer 2015) and the results shown are based on one chain, burn-in 5000, 10,000 iterations with a thinning of 10 and hyperparameters $a^\xi = 0.1$, $\kappa^2 = 1$.

Results are given in Fig. 2 and Table 2. The MISE in Table 2 is defined analogously to the case of the empirical estimator, with (19) replaced by

$$\sum_{i=1}^n (\hat{b}_n(t_i) - b(t_i))^2 (t_i - t_{i-1}). \quad (20)$$

Remark and Conclusion The computation of the empirical estimator is straightforward and in negligible time. A tight upper bound for its asymptotic error is available. The Bayes estimator based on the approximate likelihood could incorporate prior knowledge and has its potential if the class of boundaries can be parametrized by a finite-dimensional parameter. In the nonparametric case, the numerical experiments revealed drawbacks, at least compared to the empirical estimator. The computation was costly, considering time. The study was performed on 10 nodes using a cluster of workstations. Each node on the cluster has 2 six core Intel Xeon X5670 @ 2.93 GHz processor and was used for one boundary with a given n . The execution times are around 7 min and 2 h 20 min for the Bayesian estimator (while they are only around 1 s and 9 s for the empirical estimator) with $n = 10^2$, $n = 10^3$ respectively.

As can be seen in Fig. 2, in all cases considered the Bayes estimator showed a strong positive bias. This bias should be a result of the data-augmentation procedure. Note that the discrete Brownian motion is always below the boundary up to the observed exit-time. Then, conditional on the n discrete Brownian motions, the newly sampled boundary is above all these discrete Brownian motions, which have not crossed the boundary up to t . These findings do not depend on the chosen prior. Alternative priors, such as (discrete) Ornstein–Uhlenbeck processes have been considered with qualitatively the same result.

Acknowledgements Open access funding provided by Vienna University of Economics and Business (WU).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If

material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abundo M (2015) An overview on inverse first-passage-time problems for one-dimensional diffusion processes. *Semin Interdiscip Mat* 12:1–44
- Bitto A, Frühwirth-Schnatter S (2019) Achieving shrinkage in a time-varying parameter model framework. *J Econom* 210(1):75–97. <https://doi.org/10.1016/j.jeconom.2018.11.006>
- Borovkov K, Novikov A (2005) Explicit bounds for approximation rates of boundary crossing probabilities for the Wiener process. *J Appl Probab* 42(1):82–92. <https://doi.org/10.1239/jap/1110381372>
- Durbin J (1971) Boundary-crossing probabilities for the Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov-Smirnov test. *J Appl Probab* 8(3):431–453. <https://doi.org/10.2307/3212169>
- Gorenflo R, Vessella S (1991) Abel integral equations, vol 1461. Springer, Berlin. <https://doi.org/10.1007/BFb0084665>
- Lerche HR (1986) Boundary crossing of Brownian motion, vol 40. Springer, Berlin. <https://doi.org/10.1007/978-1-4615-6569-7>
- Novikov A, Frishling V, Kordzakhia N (1999) Approximations of boundary crossing probabilities for a Brownian motion. *J Appl Probab* 36(4):1019–1030. <https://doi.org/10.1239/jap/1032374752>
- Peskir G (2002) On integral equations arising in the first-passage problem for Brownian motion. *J Integral Equ Appl* 14(4):397–423. <https://doi.org/10.1216/jiea/1181074930>
- Plummer M (2015) JAGS version 4.0. user manual asd. <https://sourceforge.net/projects/mcmc-jags/files/Manuals/4.x/>. Accessed 5 Jan 2019
- Pötzelberger K (2012) Improving the Monte Carlo estimation of boundary crossing probabilities by control variables. *Mt Carlo Methods Appl* 18:353–377. <https://doi.org/10.1515/mcma-2012-0013>
- Pötzelberger K, Wang L (2001) Boundary crossing probability for Brownian motion. *J Appl Probab* 38(1):152–164. <https://doi.org/10.1239/jap/996986650>
- Ricciardi L, Crescenzo A, Giorno V, Nobile A (1999) An outline of theoretical and algorithmic approaches to first passage time problems with applications to biological modeling. *Math Jpn* 50:247–322
- Salminen P (1988) On the first hitting time and the last exit time for a Brownian motion to/from a moving boundary. *Adv Appl Probab* 20(2):411–426. <https://doi.org/10.2307/1427397>
- Shorack GR, Wellner JA (2009) Empirical processes with applications to statistics, vol 59. SIAM, Philadelphia
- Wang L, Pötzelberger K (1997) Boundary crossing probability for Brownian motion and general boundaries. *J Appl Probab* 34(1):54–65. <https://doi.org/10.2307/3215174>
- Zucca C, Sacerdote L (2009) On the inverse first-passage-time problem for a wiener process. *Ann Appl Probab* 19(4):1319–1346. <https://doi.org/10.1214/08-AAP571>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.