Statistics and Probability Letters 81 (2011) 140–145

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Test to distinguish a Brownian motion from a Brownian bridge using Polya tree process

Karthik Bharath *, Dipak K. Dey

Department of Statistics, University of Connecticut, Storrs, United States

ARTICLE INFO

Article history:
Received 14 July 2010
Received in revised form 4 October 2010
Accepted 6 October 2010
Available online 14 October 2010

Keywords:
Bayes factor
Non-subjective prior
Non-parametric Bayes
Tierney–Kadane approximation

ABSTRACT

The problem of distinguishing a Brownian bridge from a Brownian motion, both with possible drift, on the closed unit interval, is investigated via a pair of hypothesis tests. The first, tests for observations obtained at n discrete time points to be arising from a Brownian bridge with drift by embedding the Brownian bridge into a mixture of Polya trees which represents the non-parametric alternative. The second test, tests in an identical manner, for the observations to be coming from a Brownian motion with drift. The Bayes factors for the two tests are derived and then combined to obtain the Bayes factor for the test to distinguish between the two Gaussian processes. The Tierney–Kadane approximation of the Bayes factor is derived with an error approximation of order $O(n^{-4})$.

1. Introduction

The problem of modeling processes in finance, biology etc., using Brownian motion and its variants has received considerable attention, for example, see Steele (2000) and Berg (1993). For better predictive results, it is clearly important that one can fit an appropriate model from a pool of possible processes. In this paper, we specifically consider two special Gaussian processes: the Brownian motion and the Brownian bridge. For real valued i.i.d. observations $X_1, X_2, \dots, X_n$ recorded at times $t_1, t_2, \dots, t_n$ are assumed to be arising from a Brownian bridge $B(t)$ or Brownian motion $W(t)$, both with possible drift. It is further assumed that $t_1, t_2, \dots, t_n$ are deterministic. This assumption, as will be shown later, does not qualitatively affect the test statistic greatly. The Brownian bridge on the unit interval, intuitively, can be understood as a Brownian motion not only "tied down" at the origin but also at time $t = 1$. Hence, most uncertainty is present in the middle of the bridge. Suppose, we are given n observations recorded at discrete times $t_1, \dots, t_n$, $t_i \in [0, 1]$, we would be hard pressed to distinguish with certainty, between the Gaussian processes. DasGupta (1996) describes the difficulty in distinguishing the two processes by considering the L_1 mean distance between them and showing the distance on average to be quite small. Suppose $B(t)$ and $W(t)$ be a Brownian bridge and standard Brownian motion respectively. The L_1 norm between the two is given by

$$D_t = \int_0^\delta |B(t) + \mu t - (W(t) + \mu t)| \mathrm{d}t$$
$$= \int_0^\delta |B(t) - W(t)| \mathrm{d}t$$

which defines a measure on the 'difference' in their respective sample paths on the interval $[0, \delta]$. Using results from Johnson and Killeen (1983), DasGupta showed that

$$E(D_\delta) = \frac{\sqrt{\pi}}{2\sqrt{2}} \left(\frac{\delta}{2 - \delta} \right)^{3/2} \tag{1}$$

* Corresponding author.
E-mail address: karthikbharath@gmail.com (K. Bharath).

0167-7152/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2010.10.006

is small for $\delta = 0.05, 0.5, 1$. The variance of the Brownian bridge at time t is $t(1-t)$ and that of the Brownian motion is t. Hence for small t, we can intuitively understand the difficulty in distinguishing the two. The nature of the problem ensures that, approaching the problem from a model selection perspective, we no longer have one model nested within the other since both processes are considered with possible drift components. It is important to note that we assume to be observing the underlying stochastic process only at finite discrete time points. It hence becomes natural, at least from a testing perspective, to consider the problem as a multivariate testing problem.

The problem, despite being viewed as a multivariate one, is by no means routine. Since the models under the null and the alternative hypotheses are not nested, a regular likelihood ratio test for the equality of the covariance matrices is inappropriate. Also, since we are observing the process only at finite discrete time points, a finite sample test will be desirable for better accuracy, especially for predictive purposes and this necessity renders Bartlett's test for equality of covariance matrices inappropriate too. We hence adopt a Bayesian approach and derive the Bayes factor by considering a pair of hypothesis tests. We proceed in the spirit of Berger and Guglielmi (2001) in considering a null parametric model vs a non-parametric alternative. With the objective of model elaboration, for more generality, one is better off using a non-parametric alternative as opposed to a parametric model elaboration technique. We first propose a Bayes factor to distinguish between a Brownian motion with drift and a non-parametric alternative by utilizing a mixture of Polya tree process prior on the alternative. We then derive the Bayes factor to distinguish a Brownian bridge with drift and same non-parametric alternative. We then combine the two Bayes factors to obtain a Bayes factor for a test between Brownian motion and the Brownian bridge. Section 2 contains definitions and some preliminaries of Polya trees and mixed Polya trees. We then formally specify our problem and set it up in terms of two pairs of hypotheses. In Section 3 the analytical expression for the Bayes factors is derived in detail and the Tierney–Kadane approximation of the Bayes factor is obtained.

2. Polya trees and mixtures of Polya trees

2.1. Preliminaries

Polya trees form a class of distributions for a random probability measure $\mathcal{P}$ and are a special case of tail-free processes (refer to Fabius (1964) and Freedman (1963) for more details). For a rigorous treatment of Polya trees and their properties we direct the reader to Lavine (1992), Lavine (1994), Maudlin et al. (1992) and Ferguson (1974). We follow the notation used in Lavine (1992). Let $E = 0, 1$, $E^0 = \emptyset$, E^m be the m-fold product $E \times E \times \cdots \times E$ and $E^* = \bigcup_0^\infty E_m$. Let (B_0, B_1) be a partition of $\mathbb{R}$. Then let (B_{00}, B_{01}) be partition of B_0 and (B_{10}, B_{11}), a partition of B_1. Continuing in this fashion *ad infinitum*, we obtain a sequence of partitions $\prod = \{\pi_m = 0, 1, \ldots\}$ of $\mathbb{R}$ such that $\bigcup_0^\infty \pi_m$ generates the Borel sets and every $B \in \pi_{m+1}$ is obtained by splitting some $B' \in \pi_m$ into two pieces. Hence the number of partitions at level m is 2^m. Thus in general for all $\epsilon = \epsilon_1, \ldots, \epsilon_m \in E^*$, B_ϵ splits into B_{ϵ_0} and B_{ϵ_1}. Degenerate splits are permitted and hence it is possible to have a $B_{\epsilon_0} = B_\epsilon$ and $B_{\epsilon_1} = \emptyset$. A convenient way to picturise the partition is given by Muliere and Walker (1997). Imagine a particle cascading through these partitions. It starts in $\mathbb{R}$ and moves into B_{ϵ_0} with probability Y_0 or B_{ϵ_1} with probability $1 - Y_0$. In general on entering B_ϵ, the particle could either move into B_{ϵ_0} or B_{ϵ_1} with probability Y_{ϵ_0} or $1 - Y_{\epsilon_0}$. We can now formally define the Polya tree distribution.

A random probability measure $\mathcal{P}$ on $\mathbb{R}$ has Polya tree distribution with parameter $(\prod, \mathcal{A})$, denoted as $\mathcal{P} \sim \mathrm{PT}(\prod, \mathcal{A})$, if there exist nonnegative numbers $\mathcal{A} = \{\alpha_\epsilon, \epsilon \in E^*\}$ and random variables $\mathcal{Y} = \{Y_\epsilon, \epsilon \in E^*\}$, respectively, such that

- all the random variables in $\mathcal{Y}$ are independent
- for every ϵ, $Y_{\epsilon_0} \sim \mathrm{Beta}(\alpha_{\epsilon_0}, \alpha_{\epsilon_1})$
- for every $m = 1, 2, \ldots$ and every $\epsilon = \epsilon_1, \ldots, \epsilon_m \in E^*$

$$\mathcal{P}(B_{\epsilon_1,\ldots,\epsilon_m}) = \prod_{j=1,\epsilon_j=0}^{m} Y_{\epsilon_1,\ldots,\epsilon_{j-1}} \prod_{j=1,\epsilon_j=1}^{m} (1 - Y_{\epsilon_1,\ldots,\epsilon_{j-1}}).$$

The α_ϵ's determine the smoothness of the realizations of $\mathcal{P}$ and they also control how closely the distribution of $\mathcal{P}$ is concentrated around its mean (Lavine, 1992). The choice of α_ϵ is of some importance and we refer the reader to Berger and Guglielmi (2001) for further information. Dirichlet processes are special cases of Polya trees when for every $\epsilon \in E^*$, $\alpha_\epsilon = \alpha_{\epsilon_0} + \alpha_{\epsilon_1}$ as shown in Ferguson (1974). Polya trees are conjugate priors. If $\mathcal{P} | \prod, \mathcal{A} \sim \mathrm{PT}(\prod, \mathcal{A})$ and $X_1, \ldots, X_n | \mathcal{P} \sim \mathcal{P}$ i.i.d., then $\mathcal{P} | X_1, \ldots, X_n \sim \mathrm{PT}(\prod, \mathcal{A}^*)$, where $\mathcal{A}^* = \{\alpha_\epsilon^* = \alpha_\epsilon + n_\epsilon, \epsilon \in E^*\}$ and $n_\epsilon =$ number of $X_1, X_2, \ldots, X_n \in B_\epsilon$. A random probability measure is said to be mixture over θ of Polya trees, with mixing distribution π and parameters $\{\prod_\theta, \mathcal{A}_\theta\}$, $\theta \in \Theta$, if the conditional distribution of $\mathcal{P}$, given θ, is $\mathrm{PT}(\prod_\theta, \mathcal{A}_\theta)$ (Berger and Guglielmi, 2001). Our motivation for using a mixture Polya tree prior over Dirichlet process or mixture Dirichlet process priors is that the Polya tree process can give probability one to the set of continuous distributions.

2.2. Model specification

The rationale for the choice of the mixture of Polya tree prior on the alternative is multifold. As elegantly described by Berger and Guglielmi (2001), we wish to "utilize" the typically improper noninformative prior distributions for drift component. Along with the advantages offered by the mixture of Polya tree processes in modeling continuous densities, by embedding each parametric model into a mixture of Polya trees, we are able to construct a suitable probability model "around" the models of interest, i.e., the Brownian motion and the Brownian bridge, and also provide a sensitive but useful elaboration of the two models. Our proposed method leads to an elegant and computable expression for the combined Bayes factor.

We follow closely, the methodology in Berger and Guglielmi (2001, Sec. 2.2) in specifying our test(s) of hypothesis and also in notation. We embed the parametric model $f(x|\mu)$ (standard Brownian motion/Brownian bridge) into the Polya tree process. We center the Polya tree around $\mathcal{G}_\mu$ which denotes the probability measure corresponding to f. This is accomplished by choosing a mixture of Polya trees with parameters $(\prod_\mu, \mathcal{A}_\mu)$, such that

$$E(\mathcal{P}_\mu) = \mathcal{G}_\mu. \tag{2}$$

We adopt the Polya tree construction in Berger and Guglielmi (2001) and not the canonical construction with a fixed mean given in Lavine (1992). Let $W(t)$ denote a standard Brownian motion and $\mu t + W(t)$ represent the Brownian motion with drift. Let $\mu t+B(t)$ denote a Brownian bridge with drift. We are interested in the question of testing if observations $X_1, \ldots, X_n$ corresponding to time $t_1 \leq t_2 \leq, \ldots, \leq t_n$ are coming from a Brownian motion with drift or a Brownian bridge with drift. For simplicity, we consider non-random $t_i \in [0, 1]$ for all $i = 1, \ldots, n$. Let $x_1, \ldots, x_n$ represent the realization of the process at the mentioned times. Let

$$H_0^{(1)} : X(t) \equiv \mu t + B(t)$$
$$H_1^{(1)} : X(t) \equiv \mathcal{P}_\mu, \mu \in \mathbb{R}.$$

We have by virtue of the above specification, embedded the parametric Brownian bridge model within the non-parametric mixture of Polya trees by specifying the Brownian bridge as the mean of the mixture Polya process. We consider a noninformative prior $\pi(\mu)$, which is shown to be appropriately calibrated in Berger and Guglielmi (2001). Also, let,

$$H_0^{(2)} : X(t) \equiv \mu t + W(t)$$
$$H_1^{(2)} : X(t) \equiv \mathcal{P}_\mu, \mu \in \mathbb{R}.$$

We derive the Bayes factors for the pair of hypotheses, BF_1 and BF_2 respectively and then by considering the ratio $\frac{\mathrm{BF}_1}{\mathrm{BF}_2}$, we can comment on the problem of distinguishing between the two processes.

3. Bayes factor calculations

3.1. Bayes factor for $H_0^{(1)}$ vs $H_1^{(1)}$

Let $f_1(\mathbf{x}|\boldsymbol{\mu})$, where $\mathbf{x} = (x_1, \ldots, x_n)$, be the likelihood under $H_0^{(1)}$. Hence

$$f_1(\mathbf{x}|\boldsymbol{\mu}) = \frac{1}{\sqrt{2\pi}|\Sigma_1|^{1/2}} e^{\frac{-1}{2}(\mathbf{x}-\mu\mathbf{t})'\Sigma_1^{-1}(\mathbf{x}-\mu\mathbf{t})}$$

where $\mathbf{t} = (t_1, \ldots, t_n)$ and $\Sigma_1 = \big(\min(t_i, t_j) - t_i t_j\big)$. Let $\pi(\mu) = 1$, the usual noninformative prior. It is useful now to also note that $\Sigma_2 = \big(\min(t_i, t_j)\big)$ where Σ_2 is the dispersion matrix under $H_0^{(2)}$. We mention a few relations which are useful in our calculations (DasGupta, 1996).

- $\Sigma_1 = \Sigma_2 - \mathbf{tt}'$
- $\Sigma_1^{-1} = \Sigma_2^{-1} + \frac{\Sigma_2^{-1}\mathbf{tt}'\Sigma_2^{-1}}{1-\mathbf{t}\Sigma_2^{-1}\mathbf{t}}$
- $\Sigma_2^{-1}\mathbf{t} = (0, 0, \ldots, 1)'$.

Then,

$$\begin{aligned}
\frac{-1}{2}\left[(\mathbf{x}-\mu\mathbf{t})'\Sigma_1^{-1}(\mathbf{x}-\mu\mathbf{t})\right] &= \frac{-1}{2}(\mathbf{x}-\mu\mathbf{t})'\left[\Sigma_2^{-1} + \frac{\Sigma_2^{-1}\mathbf{tt}'\Sigma_2^{-1}}{1-\mathbf{t}\Sigma_2^{-1}\mathbf{t}}\right](\mathbf{x}-\mu\mathbf{t}) \\
&= \frac{-1}{2}\left[(\mathbf{x}-\mu\mathbf{t})'\Sigma_2^{-1}(\mathbf{x}-\mu\mathbf{t}) + (\mathbf{x}-\mu\mathbf{t})'\frac{\Sigma_2^{-1}\mathbf{tt}'\Sigma_2^{-1}}{1-\mathbf{t}\Sigma_2^{-1}\mathbf{t}}(\mathbf{x}-\mu\mathbf{t})\right] \\
&= \left[\frac{-\mathbf{t}'\Sigma_2^{-1}\mathbf{t}}{2}\left(\mu - \frac{\mathbf{x}'\Sigma_2^{-1}\mathbf{t}}{\mathbf{t}'\Sigma_2^{-1}\mathbf{t}}\right)^2 + \frac{(\mathbf{x}'\Sigma_2^{-1}\mathbf{t})^2}{2\mathbf{t}'\Sigma_2^{-1}\mathbf{t}}\right] - \frac{1}{2}\left[(\mathbf{x}-\mu\mathbf{t})'\frac{\Sigma_2^{-1}\mathbf{tt}'\Sigma_2^{-1}}{1-\mathbf{t}\Sigma_2^{-1}\mathbf{t}}(\mathbf{x}-\mu\mathbf{t})\right] \\
&= \frac{-t_n}{2}\left(\mu - \frac{x_n}{t_n}\right)^2 + \frac{x_n^2}{2t_n} - \frac{(x_n-\mu t_n)^2}{2(1-t_n)}.
\end{aligned}$$

Consequently, the marginal density of the sample $X_1, \ldots, X_n$ under $H_0^{(1)}$ is

$$\begin{aligned} m_0(x_1, \ldots, x_n) &= \frac{1}{\sqrt{2\pi}|\Sigma_1|^{1/2}} e^{\frac{x_n^2}{2t_n}} \int_\mu e^{\frac{-t_n}{2}\left(\mu - \frac{x_n}{t_n}\right)^2} e^{\frac{-(x_n-\mu t_n)^2}{2(1-t_n)}} \pi(\mu) \mathrm{d}\mu \\ &= \frac{1}{\sqrt{2\pi}|\Sigma_1|^{1/2}} \int_\mu e^{\frac{-t_n(x_n-\mu t_n)^2}{2(1-t_n)}} \pi(\mu) \mathrm{d}\mu \\ &= \frac{e^{\frac{x_n^2}{2t_n}}}{|\Sigma_1|^{1/2}} \sqrt{\frac{1-t_n}{t_n}}. \end{aligned}$$

Following the notation introduced in Polya tree preliminaries, from Berger and Guglielmi (2001) we obtain,

$$m_1(x_1, \ldots, x_n|\mu) = f(\mathbf{x}|\mu)\psi(\mu),$$

where

$$\psi(\mu) = \prod_{j=2}^{n} \prod_{m=1}^{m^*(x_j)} \frac{\alpha'_{\underline{\epsilon}_{m(x_j)}}(\mu)\left[\alpha_{\underline{\epsilon}_{(m-1)0}(x_j)}(\mu) + \alpha_{\underline{\epsilon}_{(m-1)1}(x_j)(\mu)}\right]}{\alpha_{\underline{\epsilon}_m(x_j)}(\mu)\left[\alpha'_{\underline{\epsilon}_{(m-1)0}(x_j)}(\mu) + \alpha_{\underline{\epsilon}'_{(m-1)1}(x_j)}(\mu)\right]} \tag{3}$$

with $\underline{\epsilon}_m(x_j)$ is the index $\epsilon_1 \cdots \epsilon_m$ such that x_j belongs to the $B_{\epsilon_1,\ldots,\epsilon_m}$, for each level m, and $\alpha'_{\underline{\epsilon}_{m(x_j)}}(\mu)$ is equal to $\alpha_{\underline{\epsilon}_{m(x_j)}}(\mu)$ plus the number of observations among $\{x_1, \ldots, x_j\}$ that belong to $B_{\epsilon_1,\ldots,\epsilon_m(x_j)}$ and for each x_j, the product in (3) is up to the smallest level $m^*(x_j)$, such that no x_i, $i < j$, belong to $B_{\underline{\epsilon}_m(x_j)}$. Now,

$$\begin{aligned} m_1(x_1, \ldots, x_n) &= \frac{1}{\sqrt{2\pi}|\Sigma_1|^{1/2}} \int_\mu e^{\frac{-1}{2}(\mathbf{x}-\mu\mathbf{t})'\Sigma_1^{-1}(\mathbf{x}-\mu\mathbf{t})} \psi(\mu)\pi(\mu)\mathrm{d}(\mu) \\ &= \frac{1}{\sqrt{2\pi}|\Sigma_1|^{1/2}} e^{\frac{x_n^2}{2t_n}} \int_\mu e^{\frac{-t_n\left(\mu-\frac{x_n}{t_n}\right)^2}{2(1-t_n)}} \psi(\mu)\mathrm{d}(\mu). \end{aligned}$$

Hence, the Bayes factor for first test is

$$\mathrm{BF}_1 = \frac{m_0(x_1, \ldots, x_n)}{m_1(x_1, \ldots, x_n)} = \frac{1}{(\Phi_n^{(1)} * \psi)\left(\frac{x_n}{t_n}\right)}, \tag{4}$$

where $\Phi_n^{(1)}$ denotes the $N\left(0, \frac{1-t_n}{t_n}\right)$ CDF and $(\Phi_n^{(1)} * \psi)$ denotes the density convolution of $\Phi_n^{(1)}$ and ψ.

3.2. Bayes factor for $H_0^{(2)}$ vs $H_1^{(2)}$

We derive BF_2 in the same manner as above. Note that

$$\begin{aligned} m_0(x_1, \ldots, x_n) &= \frac{1}{\sqrt{2\pi}|\Sigma_2|^{1/2}} e^{\frac{x_n^2}{2t_n}} \int_\mu e^{\frac{-t_n\left(\mu-\frac{x_n}{t_n}\right)^2}{2}} \pi(\mu)\mathrm{d}\mu \\ &= \frac{e^{\frac{x_n^2}{2t_n}} t_n^{-1/2}}{|\Sigma_2|^{1/2}} \end{aligned}$$

$$m_1(x_1, \ldots, x_n) = \frac{1}{\sqrt{2\pi}|\Sigma_2|^{1/2}} e^{\frac{x_n^2}{2t_n}} \int_\mu e^{\frac{-t_n\left(\mu-\frac{x_n}{t_n}\right)^2}{2}} \psi(\mu)\pi(\mu)\mathrm{d}\mu$$

which implies $\mathrm{BF}_2 = \dfrac{m_0(x_1, \ldots, x_n)}{m_1(x_1, \ldots, x_n)} = \dfrac{1}{(\Phi_n^{(2)} * \psi)\left(\frac{x_n}{t_n}\right)}$

where $\Phi_n^{(2)}$ denotes the $N\left(0, \frac{1}{t_n}\right)$ CDF and $(\Phi_n^{(2)} * \psi)$ denotes the density convolution of $\Phi_n^{(2)}$ and ψ.

3.3. Bayes factor for distinguishing $B(t)$ and $W(t)$

We can now combine BF_1 and BF_2 to obtain the required composite Bayes factor under $0 - 1$ loss for testing

$$H_0^{(1)} : X(t) \equiv \mu t + B(t)$$
$$H_0^{(2)} : X(t) \equiv \mu t + W(t).$$

It follows that the Bayes factor for $H_0^{(1)}$ vs $H_0^{(2)}$ is

$$\mathrm{BF} = \frac{\mathrm{BF}_1}{\mathrm{BF}_2} = \frac{(\Phi_n^{(2)} * \psi)\left(\frac{x_n}{t_n}\right)}{(\Phi_n^{(1)} * \psi)\left(\frac{x_n}{t_n}\right)}. \tag{5}$$

Thus, we obtain a very elegant expression for the Bayes factor of test to distinguish a Brownian motion from a Brownian bridge, both with drift on the closed unit interval. Observe that the Bayes test rejects $H_0^{(1)}$ if and only if

$$(\Phi_n^{(2)} * \psi)\left(\frac{x_n}{t_n}\right) < (\Phi_n^{(1)} * \psi)\left(\frac{x_n}{t_n}\right).$$

The test represents an exact finite sample test with some desirable properties. It is of interest to note that, owing to the structure of the covariance matrices of Brownian motion and the Brownian bridge, the Bayes factor depends *only* on x_n and t_n, the last observed value. Our assumption of deterministic times $t_i,\ i = 1, \ldots, n$, is not too stringent since we only then need to consider t_n from a random sample having an absolutely continuous distribution function on [0, 1], which becomes a part of the convolution in the numerator and denominator of the Bayes factor. As more observations are recorded, i.e. as n increases, $t_n \to 1$ and $(\Phi_n^{(1)} * \psi)\left(\frac{x_n}{t_n}\right) \to \infty$ with the inequality always satisfied and Bayes factor ruling in favor of the Brownian motion with drift. This result is intuitively appealing as the Brownian bridge in our setup is essentially a Brownian motion conditioned to assume the value zero as time $t = 1$.

3.4. Tierney–Kadane approximation

The Bayes factor BF, albeit an exact test, can be approximated with a fair degree of accuracy by using the Tierney–Kadane approximations (see Tierney and Kadane, 1986) of the integrals as posterior means in the numerator and the denominator. For instance, consider the convolution

$$(\Phi_n^{(1)} * \psi)\left(\frac{x_n}{t_n}\right) = \frac{1}{\sqrt{2\pi}}\left(\frac{t_n}{1-t_n}\right)^{1/2}\int_\mu \mathrm{e}^{\frac{-t_n\left(\mu-\frac{x_n}{t_n}\right)^2}{2(1-t_n)}}\psi(\mu)\pi(\mu)\mathrm{d}(\mu)$$
$$= \frac{\frac{1}{\sqrt{2\pi}}\left(\frac{t_n}{1-t_n}\right)^{1/2}\int_\mu \mathrm{e}^{\frac{-t_n\left(\mu-\frac{x_n}{t_n}\right)^2}{2(1-t_n)}}\psi(\mu)\pi(\mu)\mathrm{d}(\mu)}{\frac{1}{\sqrt{2\pi}}\left(\frac{t_n}{1-t_n}\right)^{1/2}\int_\mu \mathrm{e}^{\frac{-t_n\left(\mu-\frac{x_n}{t_n}\right)^2}{2(1-t_n)}}\mathrm{d}(\mu)}$$

since denominator is the $N\left(\frac{x_n}{t_n}, \frac{t_n}{1-t_n}\right)$ density and $\pi(\mu) = 1$. Let

$$L = \log\left(\frac{1}{\sqrt{2\pi}}\right) + \log\left(\frac{t_n}{1-t_n}\right)^{1/2} + \frac{-t_n\left(\mu-\frac{x_n}{t_n}\right)^2}{2(1-t_n)} \quad \text{the log-likelihood under } H_0^{(1)}$$
$$L^* = \log\left(\frac{1}{\sqrt{2\pi}}\right) + \log\left(\frac{t_n}{1-t_n}\right)^{1/2} + \frac{-t_n\left(\mu-\frac{x_n}{t_n}\right)^2}{2(1-t_n)} + \log\psi(\mu).$$

Therefore,

$$(\Phi_n^{(1)} * \psi)\left(\frac{x_n}{t_n}\right) = \frac{\int_\mu \mathrm{e}^{L^*}\mathrm{d}(\mu)}{\int_\mu \mathrm{e}^{L}}.$$

Assuming the requisite regularity conditions for the Tierney–Kadane approximations and noting our assumption of non-random $t_i,\ i = 1, \ldots, n$, we can observe that L and L^* attain maxima at $\hat{\mu} = \frac{x_n}{t_n}$. Now,

$$L''(\hat{\mu}) = \frac{-t_n}{1-t_n} \quad \text{and} \quad \sigma^2 = \frac{-1}{L''(\hat{\mu})} = \frac{1-t_n}{t_n}$$
$$L^{*''}(\hat{\mu}) = \frac{-t_n}{1-t_n} + g\left(\frac{x_n}{t_n}\right) \quad \text{and} \quad \sigma^{*2} = \frac{-1}{L^{*''}(\hat{\mu})} = \frac{(1-t_n)}{t_n\left(1+g\left(\frac{x_n}{t_n}\right)\right) - g\left(\frac{x_n}{t_n}\right)}$$

where $g(v) = \frac{\psi(v)\psi''(v)-\psi'(v)^2}{\psi'(v)^2}$. From the Tierney–Kadane approximation we obtain

$$\begin{aligned}(\Phi_n^{(1)} * \psi)\left(\frac{x_n}{t_n}\right) &= \frac{\sigma^*}{\sigma}\mathrm{e}^{(L^*(\hat{\mu})-L(\hat{\mu}))}(1+O(n^{-2}))\\ &= \left(\frac{t_n}{t_n\left(1+g\left(\frac{x_n}{t_n}\right)\right)-g\left(\frac{x_n}{t_n}\right)}\right)^{1/2}\psi\left(\frac{x_n}{t_n}\right)(1+O(n^{-2}))\end{aligned}$$

with the error in approximation being of order $O(n^{-2})$. In a similar fashion,

$$(\Phi_n^{(2)} * \psi)\left(\frac{x_n}{t_n}\right) = \left(\frac{t_n}{\left(t_n - g\left(\frac{x_n}{t_n}\right)\right)}\right)^{1/2}\psi\left(\frac{x_n}{t_n}\right)(1+O(n^{-2})).$$

Consequently,

$$\mathrm{BF} = \left(\frac{t_n\left(1+g\left(\frac{x_n}{t_n}\right)\right)-g\left(\frac{x_n}{t_n}\right)}{t_n - g(\frac{x_n}{t_n})}\right)^{1/2}(1+O(n^{-4}))$$

with an approximation error of order $O(n^{-4})$ with probability one under the true value of μ. We direct the reader to Section 5 of Berger and Guglielmi (2001) for the computation details of ψ and consequently g. By embedding each of the models under consideration into Polya trees and then taking the ratio of the resulting BF_1 and BF_2, we are able to obtain a closer approximation of the required Bayes factor BF. The approximation should not deflect attention from the fact that the Bayes factor BF is indeed an exact test for finite samples. The approximation performed in order to obtain a more "friendly" form in fact, is of order $O(n^{-4})$, which compares favorably with $O(n^{-2})$ (see Greenstreet and Connor, 1974): the order to which Bartlett's test statistic belongs in testing the equality of covariance matrices is a chi-square random variable. In general, our method does not offer an alternative to Bartlett's test in comparing covariance matrices but offers a more suitable alternative to the problem we are addressing.

4. Concluding remarks

We develop a hypothesis testing procedure using Bayes factor for distinguishing between a Brownian bridge and Brownian motion on the closed unit interval, both with drift. In order to compute the Bayes factor, we embed both hypotheses within a non-parametric class of alternatives using a mixture of Polya trees and then obtain the required Bayes factor by taking the ratio of the two Bayes factors. This approach offers certain computational as well as conceptual advantages in terms of model elaboration, complementing the use of a flat noninformative prior on the drift parameter. By directing our efforts towards deriving the Bayes factor, we exploit the rich decision–theoretic component of our approach: quantifying the utility in choosing our null hypothesis of modeling the process using a Brownian bridge with drift. Finally, the Tierney–Kadane approximation of the Bayes factor is derived, which offers a computable form of the Bayes factor with an error approximation of order $O(n^{-4})$.

Acknowledgement

We wish to thank an anonymous referee for the insightful comments, particularly with the construction of the Polya tree prior, which led to some improvements on an earlier version of the manuscript.

References

Berg, H., 1993. Random Walks in Biology. Princeton University Press.
Berger, J., Guglielmi, A., 2001. Bayesian and conditional frequentist testing of a parametric model versus nonparametric alternatives. Journal of American Statistical Association 96, 174–184.
DasGupta, A., Distinguishing a Brownian bridge from a Brownian motion with drift. Tech Report 96-8. Department of Statistics, Purdue University. 1996.
Fabius, J., 1964. Asymptotic behaviour of Bayes' estimate. Annals of Mathematical Statistics 35, 846–856.
Ferguson, T., 1974. Prior distributions of spaces of probability measures. Annals of Statistics 2, 615–629.
Freedman, D., 1963. On the asymptotic behaviour of Bayes' estimate in the discrete case. Annals of Mathematical Statistics 34, 1386–1403.
Greenstreet, R., Connor, R., 1974. Power of tests for equality of covariance matrices. Technometrics 16, 27–30.
Johnson, B., Killeen, T., 1983. An explicit formula for the CDF of the L_1 norm of the Brownian bridge. Annals of Probability 11, 807–808.
Lavine, M., 1992. Some aspects of Polya tree distributions for statistical modelling. Annals of Statistics 20, 1222–1235.
Lavine, M., 1994. More aspects of Polya tree distributions for statistical modelling. Annals of Statistics 22, 1161–1176.
Maudlin, R., Sudderath, W., Williams, S., 1992. Polya trees and random distributions. Annals of Statistics 20, 1203–1221.
Muliere, P., Walker, S., 1997. A Bayesian non-parametric approach to survival analysis using Polya trees. Scandinavian Journal of Statistics 24, 331–340.
Steele, M., 2000. Stochastic Calculus and Financial Applications. Springer.
Tierney, L., Kadane, J., 1986. Approximations for posterior moments and marginal densities. Journal of American Statistical Association 81, 82–86.