Ann. Inst. Statist. Math.
Vol. 55, No. 3, 487–497 (2003)

CHARACTERIZATION OF THE LEAST CONCAVE MAJORANT OF BROWNIAN MOTION, CONDITIONAL ON A VERTEX POINT, WITH APPLICATION TO CONSTRUCTION

CHRIS CAROLAN[1] AND RICHARD DYKSTRA[2]

[1] *Department of Mathematics, East Carolina University, 129 Austin, Greenville, NC 27858, U.S.A.*
[2] *Department of Statistics and Actuarial Science, University of Iowa, 241 Schaeffer, Iowa City, IA 52242, U.S.A.*

(Received August 15, 2001; revised May 1, 2002)

Abstract. The characterization of the least concave majorant of Brownian motion by Pitman (1983, *Seminar on Stochastic Processes, 1982* (eds. E. Cinlar, K. L. Chung and R. K. Getoor), 219–228, Birkhäuser, Boston) is tweaked, conditional on a vertex point. The joint distribution of this vertex point is derived and is shown to be generated with extreme ease. A procedure is then outlined by which one can construct the least concave majorant of a standard Brownian motion path over any finite, closed subinterval of $(0, \infty)$. This construction is exact in distribution. One can also construct a linearly interpolated version of the Brownian motion path (i.e. we construct the Brownian motion path over a grid of points and linearly interpolate) corresponding to this least concave majorant over the same finite interval. A discussion of how to translate the aforementioned construction to the least concave majorant of a Brownian bridge is also presented.

Key words and phrases: Brownian motion, least concave majorant.

1. Introduction

The least concave majorant of Brownian motion captures the limiting behavior of estimators in a few order restricted problems when the parameters fall on a boundary (equality of some or all the parameters). These boundaries are typically the null hypotheses in statistical tests involving order restrictions of the parameters. As such, the least concave majorant of Brownian motion captures the limiting null behavior of these order restricted test statistics. The least concave majorant of Brownian motion is fairly complex as a process and great strides have been made in understanding its behavior. Most notable are the papers by Groeneboom (1983) and Pitman (1983) which characterize this process. Groeneboom (1983) and Carolan and Dykstra (2001) also discuss the marginal behavior at a fixed point of this process. These results notwithstanding, any particular functional of the least concave majorant of Brownian motion may simply be too complex to convey its distribution in a closed form. Statisticians may then be forced to simulate these functionals in an attempt to table their (estimated) distributions.

We discuss later in the paper how the least concave majorant of Brownian motion maps to the least concave majorant of a Brownian bridge. Functionals of the least concave majorant of a Brownian bridge are typically what statisticians are concerned about. The steps by which the least concave majorant of a Brownian bridge is presently constructed (approximately) typically consist of constructing Brownian bridge values over a finite grid, linearly interpolating, and then determining the least concave majorant of this interpolated path. The "appropriateness" of this type of construction procedure depends on the spacing of the grid points, the number of grid points, and the functional desired. Clearly, the more grid points, the better. As an example, suppose the functional is the area under the least concave majorant of a Brownian bridge. The procedure would most likely do well in capturing the behavior of this area for a large grid of equally spaced times, but it is clearly biased downward. Adding more points to the grid can only serve to raise (or maintain) the least concave majorant and thus perhaps increase the area. Another example of a functional is the longest horizontal run over which the least concave majorant of a Brownian bridge is linear. The common procedure outlined above should not perform well for this particular functional as any additional grid points have the chance to "break" a linear segment into two linear segments (or even consolidate linear segments).

This paper addresses the construction (exactly) of the least concave majorant of Brownian motion over any closed subinterval of $(0,\infty)$. This in turn can be used to construct the least concave majorant of a Brownian bridge over any closed subinterval of $(0,1)$. If this procedure were used for the purposes of simulating the area under the least concave majorant of a Brownian bridge, one would still be forced to approximate the least concave majorant near time zero and time one, but a bound on how far our estimated area is off could be obtained and one could choose the interval sufficiently large as to guarantee our simulated area is correct to any epsilon amount. If this procedure were used for the purposes of simulating the longest horizontal run over which the least concave majorant of a Brownian bridge is linear, the distribution could be simulated exactly, choosing the interval sufficiently large so that the lengths of the unconstructed regions are shorter than the longest horizontal run in the constructed region.

Our procedure for the construction of the least concave majorant of a Brownian bridge over any closed subinterval of $(0,1)$ will generally result in better estimates of functionals of the least concave majorant of a Brownian bridge and hence is quite desirable. Two specific examples in the literature where this procedure could have been utilized are in the papers by Woodroofe and Sun (1999) and Wu *et al.* (2001). In these papers, the authors obtain a penalized, order restricted test statistic whose null limiting distribution is $\int [K'_{0,c}(t)]^2 dt$, where $K'_{0,c}$ is the left-hand slope corresponding to $K_{0,c}$ which is the least concave majorant of "a Brownian bridge process minus the indicator function $cI_{(0,1)}$". Thus, Woodroofe and Sun lower a Brownian bridge by $c > 0$ everywhere except the endpoints and then take the least concave majorant in order to obtain $K_{0,c}$. $K_{0,c}$ can equivalently obtained by taking the least concave majorant of the Brownian bridge, lowering the function by c everywhere except the endpoints, and taking the least concave majorant again. Our construction procedure outlined in this paper can be used to construct $K_{0,c}$ over the entire interval $[0,1]$ simply by constructing the least concave majorant of the Brownian bridge over a sufficiently large interval so that our constructed process is below c at our endpoints of construction and neither of our endpoints is a maximum.

In Section 2, we discuss Pitman's (1983) characterization of the least concave majorant of Brownian motion. It is this characterization on which we base our construction procedure. The following three sections discuss how one can construct the least concave majorant of a Brownian motion over any closed, subinterval of $(0,\infty)$. In Section 6, we discuss how one can (approximately) construct the associated Brownian motion path over the same interval of construction as the least concave majorant. We conclude with Section 7 which discusses how to construct the least concave majorant of a Brownian bridge over any closed, subinterval of $(0,1)$ and how one can (approximately) construct the associated Brownian bridge path over the same interval of construction as the least concave majorant.

2. Pitman's characterization of $\{K(t) : t \geq 0\}$

Define the process $\{K(t) : t \geq 0\}$ to be the least concave majorant of a standard Brownian motion process $\{W(t) : t \geq 0\}$. Pitman (1983) notes early in his paper that the process $\{K(t) : t \geq 0\}$ will almost surely form a piecewise linear curve. Pitman denotes the random set of all *vertex times* by $V \subset (0,\infty)$, a vertex time being the location where a change in slope occurs in the process $\{K(t) : t \geq 0\}$. Groeneboom (1983) showed with probability one, for $0 < s < t < \infty$, the set of vertex times V has a finite number of points in (s,t) and a countably infinite number of points in each of the intervals $(0,s)$ and (t,∞). Pitman states in order to isolate a point in V, we could fix $b > 0$ and consider the unique line with slope b that is tangent to $\{K(t) : t \geq 0\}$. Pitman defines τ_b to be the last time this line touches $\{K(t) : t \geq 0\}$. Given τ_b is a point in V, Pitman proceeds to index the points of V relative to τ_b, defining V_0 to be τ_b, V_n to be the n-th point in V after V_0, and V_{-n} to be the n-th point in V before V_0. Thus, $V = \{V_i : i \in \boldsymbol{Z}\}$ almost surely.

For $i \in \boldsymbol{Z} - \{0\}$, Pitman defines (we alter indices/notation slightly)

$$T_i = \begin{cases} V_i - V_{i-1} & i \geq 1 \\ V_{i+1} - V_i & i \leq -1 \end{cases}$$

and

$$\alpha_i = \begin{cases} [K(V_i) - K(V_{i-1})]/T_i & i \geq 1 \\ [K(V_{i+1}) - K(V_i)]/T_i & i \leq -1 \end{cases}.$$

Thus, α_i is the slope of the $|i|$-th linear segment of $\{K(t) : t \geq 0\}$ after V_0 (or before V_0 if i is negative) with horizontal run T_i, see Fig. 1. Pitman notes that the entire process $\{K(t) : t \geq 0\}$ is characterized by the set of random variables $\{(T_i, \alpha_i) : i \in \boldsymbol{Z} - \{0\}\}$. Pitman gives the joint distribution of these random pairs (his Theorem 1.1) as the following:

- α_1 is uniformly distributed on $(0,b)$ and conditional on $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, α_{n+1} is uniformly distributed on $(0, \alpha_n)$.
- α_{-1} has density bx^{-2} on the interval (b,∞) and conditional on $\{\alpha_{-1}, \alpha_{-2}, \ldots, \alpha_{-n}\}$, $\alpha_{-(n+1)}$ has density $\alpha_{-n}x^{-2}$ on the interval (α_{-n}, ∞).
- The sequences $(\alpha_n : n \geq 1)$ and $(\alpha_n : n \leq -1)$ are independent.
- Conditional on $\{\alpha_i : i \in \boldsymbol{Z} - \{0\}\}$, $\{\alpha_i^2 T_i : i \in \boldsymbol{Z} - \{0\}\}$ is a set of independent, chi-square random variables each with one degree of freedom.

The preceding characterization is very attractive. However, this characterization is unconditional of $(V_0, K(V_0))$. Thus, in order to directly use it to construct the process $\{K(t) : t \geq 0\}$, one would need to obtain the entire sequence (T_i, α_i) for $i = -1, -2, -3, \ldots$ and construct the vertex points based on the knowledge that $K(0) = 0$, noting that

$$(V_0, K(V_0)) = \left(\sum_{i=-1}^{-\infty} T_i, \sum_{i=-1}^{-\infty} \alpha_i T_i \right).$$

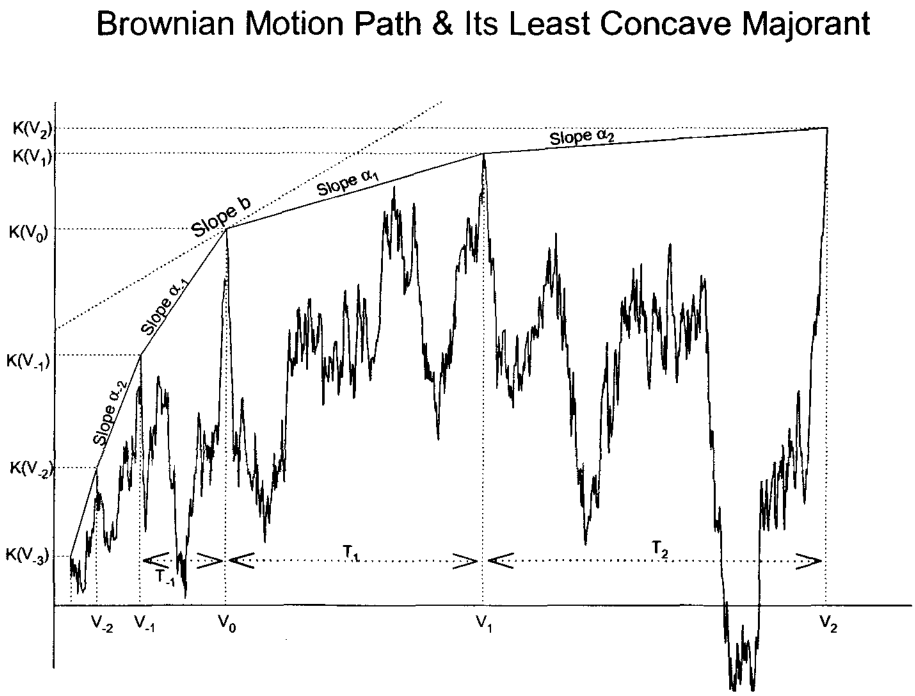

Fig. 1. Graph of a Brownian motion realization along with its least concave majorant. The graph was simulated using the procedures outlined in this paper. The slopes and vertex points are labelled.

The following 3 sections tweak Pitman's characterization of $\{K(t) : t \geq 0\}$, in order to allow one to construct $\{K(t) : t \geq 0\}$ over any closed subinterval of $(0, \infty)$.

3. Choosing a starting point

Pitman indexes the vertex points of $\{K(t) : t \geq 0\}$ by fixing $b > 0$ and considering the unique line with slope b that is tangent to $\{K(t) : t \geq 0\}$. Pitman defines τ_b to be the last time this line touches $\{K(t) : t \geq 0\}$. The point $(\tau_b, K(\tau_b))$ is a vertex point of $\{K(t) : t \geq 0\}$ such that the slopes of all linear segments of $\{K(t) : t \geq 0\}$ before τ_b are at least b and after τ_b are strictly smaller than b. We will use the point $(\tau_b, K(\tau_b))$ as a starting point for our construction of $\{K(t) : t \geq 0\}$.

Recall $\{K(t) : t \geq 0\}$ is defined to be the least concave majorant of a standard Brownian motion process $\{W(t) : t \geq 0\}$. The vertex times of $\{K(t) : t \geq 0\}$ are the locations where these two processes touch. Since the least concave majorant of the sum of a line and an arbitrary function is the sum of the line and the least concave majorant of the arbitrary function, then $\{K(t) - bt : t \geq 0\}$ is the least concave majorant of the process $\{W(t) - bt : t \geq 0\}$. The processes $\{K(t) : t \geq 0\}$ and $\{K(t) - bt : t \geq 0\}$ have the same vertex times and corresponding left-hand slopes of the processes differ by b. Thus, if we wish to determine the location τ_b where the left-hand slopes of the process $\{K(t) : t \geq 0\}$ change from being at least b to being strictly less than b, we can determine the location where the left-hand slopes of the process $\{K(t) - bt : t \geq 0\}$

change from being at least 0 to being strictly less than 0. This change in left-hand slopes from nonnegative to negative clearly occurs where $\{K(t) - bt : t \geq 0\}$ achieves its last global maximum, which is also the location where $\{W(t) - bt : t \geq 0\}$ achieves its last global maximum. Hence,

$$\tau_b = \sup\left\{x : W(x) - bx = \sup_{u\geq 0}\{W(u) - bu\}\right\}.$$

The associated value of $K(\tau_b)$ is given by

$$K(\tau_b) = \sup_{u\geq 0}\{W(u) - bu\} + b\tau_b.$$

We fix $b = 1$ and give the joint density of $(\tau_1, K(\tau_1))$ in the following theorem.

THEOREM 3.1. *The joint density of* $(\tau_1, K(\tau_1))$ *is given by*

$$g_{\tau_1,K(\tau_1)}(t, y) = \frac{2}{t^{3/2}}(y - t)\phi\left(\frac{y}{\sqrt{t}}\right) \qquad 0 < t \leq y < \infty$$

where ϕ *is the standard normal density function. The proof can be found in the Appendix.*

The marginal density of the vertex time τ_1 is given in Carolan and Dykstra (2001), where τ_1 is denoted by X_1 in that paper. Amazingly, observations from the bivariate density given in Theorem 3.1 can be generated with great ease, as described in the following lemma.

LEMMA 3.1. *If* U *is a standard uniform random variable and independently,* T *is a chi-square random variable with* 3 *degrees of freedom, then*

$$(\tau_1, K(\tau_1)) \stackrel{d}{=} (T(1 - \sqrt{U})^2, T(1 - \sqrt{U})).$$

The proof can be found in the Appendix.

Of course, one may not wish to fix b at 1. Since $\{K(t) : t \geq 0\} \stackrel{d}{=} \{bK(\frac{t}{b^2}) : t \geq 0\}$, see Carolan and Dykstra (2001) for details on proving this rescaling property of the least concave majorant of Brownian motion as well as for details on proving the time reversed property and Doob's transformation discussed later in this paper, then

$$(\tau_b, K(\tau_b)) \stackrel{d}{=} \left(\frac{1}{b^2}\tau_1, \frac{1}{b}K(\tau_1)\right) \stackrel{d}{=} \left(\frac{1}{b^2}T(1 - \sqrt{U})^2, \frac{1}{b}T(1 - \sqrt{U})\right).$$

where U is a standard uniform random variable and independently, T is a chi-square random variable with 3 degrees of freedom. Thus, we can easily simulate $(\tau_b, K(\tau_b))$ for any arbitrary b. Just as Pitman does in his characterization, we will now define V_0 as τ_b and continue our construction conditional on knowing $(V_0, K(V_0))$.

4. Constructing $\{K(t) : t \geq 0\}$ beyond V_0

Pitman utilized the path decomposition given in Williams (1974) in the derivation of his characterization of $\{K(t) : t \geq 0\}$. According to Williams' path decomposition,

$$\{W(t) : 0 \leq t \leq V_0\} \quad \text{and} \quad \{W(V_0 + s) - W(V_0) - bs : s \geq 0\} \tag{4.1}$$

are independent. Pitman was able to show from the process on the right of (4.1) that

- the slope α_1 corresponding to the process $\{K(t) : t \geq 0\}$ is uniformly distributed on $(0, b)$ and conditional on the slopes $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, the slope α_{n+1} is uniformly distributed on $(0, \alpha_n)$.
- conditional on the slopes $\{\alpha_1, \alpha_2, \ldots\}$, the horizontal runs between vertex points $T_i = V_i - V_{i-1}$ are independent with $\alpha_i^2 T_i$ having a chi-square distribution with 1 degree of freedom.

Notice that the point $(V_0, K(V_0))$ is the last point of the process on the left of (4.1). Thus, the process on the right is independent of our starting vertex point $(V_0, K(V_0))$. Hence, the information necessary to conditionally construct the vertex points to the right of our starting vertex point $(V_0, K(V_0))$, is outlined in the preceding paragraph.

In practice, the next n vertex points to the right of V_0 can be obtained very quickly. We recommend the following steps:

1. Create a vector of n independent uniform $(0, 1)$ realizations. Form the partial products of this vector followed by multiplying the vector by b; the result is the vector of slopes $(\alpha_1, \alpha_2, \ldots, \alpha_n)$.
2. Create a vector of n independent chi-square with one degree of freedom realizations. Take this vector of chi-square realizations and divide it (coordinate-wise) by the vector $(\alpha_1^2, \alpha_2^2, \ldots, \alpha_n^2)$; the result is the $(T_1, T_2, \ldots, T_n)$ vector.
3. Form the partial sums of the $(T_1, T_2, \ldots, T_n)$ vector and add V_0; the result is the vector $(V_1, V_2, \ldots, V_n)$.
4. Form the partial sums of the product (coordinate-wise) of the vector $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and the vector $(T_1, T_2, \ldots, T_n)$, followed by adding $K(V_0)$; the result is the vector $(K(V_1), K(V_2), \ldots, K(V_n))$.
5. In practice, if n is not chosen large enough initially, we may need to continue to add $R \geq 0$ (random) coordinates to the end of the same uniform and chi-square vectors in steps one and two until we have constructed enough of the path for some criteria to be satisfied. For example, if we desire to construct $\{K(t) : t \geq 0\}$ over the interval $[c_1, c_2]$, then we choose R large enough so that (or until) $V_{n+R} \geq c_2$.

Remark 4.1. Construction before V_0, presented in the next section, can be avoided if b is initially chosen large enough so that it is highly LIKELY that V_0 falls to the left of c_1, the left endpoint of the desired interval of construction $[c_1, c_2]$. We recommend choosing $b^2 = 20/c_1$.

5. Constructing $\{K(t) : t \geq 0\}$ before V_0

Given the independence of the two processes in the path decomposition of Williams, it follows that given the slope b and the point $(V_0, K(V_0))$, the behavior of $\{K(t) : t \geq 0\}$ before V_0 can be obtained independently of the behavior after V_0. The key to constructing $\{K(t) : t \geq 0\}$ before V_0 is to translate $\{K(t) : t \geq 0\}$ to its time-reversed process, which itself is distributed as the least concave majorant of a Brownian motion. It can be demonstrated that

$$\{K(t) : t \geq 0\} \stackrel{d}{=} \left\{tK\left(\frac{1}{t}\right) : t \geq 0\right\}.$$

Thus, we can construct $\{K(t) : t \geq 0\}$ before V_0 by constructing its time-reversed version $\{K^*(t) : t \geq 0\}$ defined by $\{tK(\frac{1}{t}) : t \geq 0\}$ beyond some point and then mapping back. This is accomplished in the following manner.

1. The translated time-reversed problem has initial vertex point $(V_0^*, K^*(V_0^*)) = (\frac{1}{V_0}, \frac{K(V_0)}{V_0})$ with initial slope $b^* = K(V_0) - bV_0$.
2. Get the vectors $(V_1^*, V_2^*, \ldots, V_m^*)$ and $(K^*(V_1^*), K^*(V_2^*), \ldots, K^*(V_m^*))$ for the time-reversed process $\{K^*(t) : t \geq 0\}$ in the same manner as in the previous section.
3. We obtain the vector $(V_{-m}, V_{-(m-1)}, \ldots, V_{-1})$ by taking the (coordinate-wise) reciprocal of the $(V_1^*, V_2^*, \ldots, V_m^*)$ vector and reversing the order of the vector.
4. We obtain the $(K(V_{-m}), K(V_{-(m-1)}), \ldots, K(V_{-1}))$ vector by dividing (coordinate-wise) the vector $(K^*(V_1^*), K^*(V_2^*), \ldots, K^*(V_m^*))$ by the vector $(V_1^*, V_2^*, \ldots, V_m^*)$ and reversing the order.
5. In practice, if m is not chosen large enough initially, we may need to continue to add $R^* \geq 0$ (random) coordinates to the same m-length vectors of uniforms and chi-squares formed in step two until we have constructed enough of the path for some criteria to be satisfied. For example, if we desired to construct $\{K(t) : t \geq 0\}$ over the interval $[c_1, c_2]$, then we choose R^* large enough so that (or until) $V_{m+R^*}^* \geq 1/c_1$.

Remark 5.1. Interestingly, given b and conditional on $(V_0, K(V_0))$, α_{-1} is uniformly distributed over $(b, \frac{K(V_0)}{V_0})$ and conditional on the vertex points

$$(V_0, K(V_0)), (V_{-1}, K(V_{-1})), \ldots, (V_{-n}, K(V_{-n})),$$

$\alpha_{-(n+1)}$ is uniformly distributed over $(\alpha_{-n}, \frac{K(V_{-n})}{V_{-n}})$. This falls out from the transformation procedure.

6. Constructing (approximately) the corresponding $\{W(t) : t \geq 0\}$

Groeneboom (1983) states that the behavior of the process $\{K(t) - W(t) : t \geq 0\}$ between successive zeros (or vertex locations of K) V_i and V_{i+1}, $i \in \boldsymbol{Z}$, is as independent, linearly time-transformed, rescaled Brownian excursions. Specifically,

$$\begin{aligned}&\{K(t) - W(t) : V_i \leq t \leq V_{i+1}, i \in \boldsymbol{Z}\}\\ &\stackrel{d}{=} \left\{\sqrt{V_{i+1} - V_i}\, W_0^{*(i)}\left(\frac{t - V_i}{V_{i+1} - V_i}\right) : V_i \leq t \leq V_{i+1}, i \in \boldsymbol{Z}\right\}\end{aligned}$$

where $\{W_0^{*(i)}(t) : 0 \leq t \leq 1\}$, $i \in \boldsymbol{Z}$, are independent standard Brownian excursion processes.

Itô and McKean (1974) note that a Brownian excursion process, denoted by $\{W_0^*(t) : 0 \leq t \leq 1\}$, is a nonhomogeneous Markov process over the unit interval with $W_0^*(0) = W_0^*(1) = 0$, marginal densities for $0 < t < 1$ given by

$$f_{W_0^*(t)}(y) = \frac{2y^2}{\sqrt{2\pi t^3(1-t)^3}} \exp\left[-\frac{y^2}{2t(1-t)}\right] \qquad y > 0$$

and transition densities for $0 < s < t < 1$ given by $(x > 0)$

$$f_{W_0^*(t)|W_0^*(s)}(y \mid x) = \frac{1}{\sqrt{t-s}}\left[\phi\left(\frac{y-x}{\sqrt{t-s}}\right) - \phi\left(\frac{y+x}{\sqrt{t-s}}\right)\right]\left(\frac{1-s}{1-t}\right)^{3/2}\frac{y}{x}\frac{\exp\left[-\dfrac{y^2}{2(1-t)}\right]}{\exp\left[-\dfrac{x^2}{2(1-s)}\right]} \qquad y > 0$$

where ϕ is the standard normal density. Note that, for $t \in [0,1]$, $W_0^*(t) \stackrel{d}{=} \sqrt{t(1-t)\chi_3^2}$ where χ_3^2 is a chi-square distribution with three degrees of freedom. Also, the transition density for $0 < s < t < 1$ can be simplified to $(x > 0)$

$$f_{W_0^*(t)|W_0^*(s)}(y \mid x) = \frac{\sqrt{c}}{u}y\{\phi(\sqrt{c}(y-u)) - \phi(\sqrt{c}(y+u))\} \qquad y > 0$$

where $c = \frac{(1-s)}{(1-t)(t-s)}$ and $u = \frac{(1-t)x}{(1-s)}$. The corresponding transition distribution function is given by

$$F_{W_0^*(t)|W_0^*(s)}(y \mid x) = P(W_0^*(t) \le y \mid W_0^*(s) = x) = \begin{cases} 1 - \{\overline{\Phi}(\sqrt{c}(y-u)) + \overline{\Phi}(\sqrt{c}(y+u))\} \\ -\dfrac{1}{u\sqrt{c}}\{\phi(\sqrt{c}(y-u)) - \phi(\sqrt{c}(y+u))\} \end{cases} \qquad y \ge 0.$$

where $\overline{\Phi}$ is the standard normal survival function.

Thus, for t in between constructed vertex times V_i and V_{i+1},

$$\begin{aligned} W(t) &= K(t) - \sqrt{V_{i+1} - V_i}\,W_0^{*(i)}\left(\frac{t - V_i}{V_{i+1} - V_i}\right) \\ &= \frac{[t - V_i]K(V_{i+1}) + [V_{i+1} - t]K(V_i)}{V_{i+1} - V_i} - \sqrt{V_{i+1} - V_i}\,W_0^{*(i)}\left(\frac{t - V_i}{V_{i+1} - V_i}\right). \end{aligned}$$

So, to approximate the corresponding $\{W(t) : t \ge 0\}$ over the same interval we have constructed $\{K(t) : t \ge 0\}$, we have the options of constructing $\{W(t) : t \ge 0\}$ over a predetermined, equally spaced grid or a grid that accounts for the differing distances between neighboring vertex locations. One would typically construct the corresponding $\{W(t) : t \ge 0\}$ over the chosen grid and linearly interpolate.

7. Constructing $\{K_0(t) : 0 \le t \le 1\}$ and $\{W_0(t) : 0 \le t \le 1\}$

The procedure to construct $\{K(t) : t \ge 0\}$ can also be used to construct the least concave majorant of a Brownian bridge over any closed subinterval of $(0,1)$. We define $\{K_0(t) : 0 \le t \le 1\}$ to be the least concave majorant of a standard Brownian bridge process $\{W_0(t) : 0 \le t \le 1\}$. By Doob's transformation, we can define $\{W_0(t) : 0 \le t \le 1\}$ to be $\{(1-t)W(\frac{t}{1-t}) : 0 \le t \le 1\}$. It thus follows that $\{K_0(t) : 0 \le t \le 1\}$ is given by $\{(1-t)K(\frac{t}{1-t}) : 0 \le t \le 1\}$. Thus, the set of constructed vertex points $\{(V_i, K(V_i)) : i = -m, \ldots, n\}$ of the process $\{K(t) : t \ge 0\}$ maps to the set of vertex points $\{(\frac{V_i}{1+V_i}, \frac{1}{1+V_i}K(V_i)) : i = -m, \ldots, n\}$ of the process $\{K_0(t) : 0 \le t \le 1\}$. Thus, if one wishes to construct $\{K_0(t) : 0 \le t \le 1\}$ over $[c_3, c_4]$, we choose n so that $V_n \ge c_4/(1 - c_4)$ and m such that $V_m^* \ge (1 - c_3)/c_3$.

In order to construct $\{W_0(t) : 0 \le t \le 1\}$ associated with $\{K_0(t) : 0 \le t \le 1\}$, one could map the gridded $\{W(t) : t \ge 0\}$ to $\{W_0(t) : 0 \le t \le 1\}$ in the same fashion as the vertex points of $\{K(t) : t \ge 0\}$ were mapped to $\{K_0(t) : 0 \le t \le 1\}$ in the above paragraph. Another option is to construct $\{W_0(t) : 0 \le t \le 1\}$ after the transformation. From the transformation procedure, it follows that the behavior of the process $\{K_0(t) - W_0(t) : t \ge 0\}$ between successive zeros (or vertex locations of K_0) $\frac{V_i}{1+V_i}$ and $\frac{V_{i+1}}{1+V_{i+1}}$, $i \in \boldsymbol{Z}$, is also as independent, linearly time-transformed, rescaled Brownian excursions. Specifically,

$$\left\{K_0(t) - W_0(t) : \frac{V_i}{1+V_i} \le t \le \frac{V_{i+1}}{1+V_{i+1}}, i \in \boldsymbol{Z}\right\}$$
$$\stackrel{d}{=} \left\{\sqrt{\frac{V_{i+1}}{1+V_{i+1}} - \frac{V_i}{1+V_i}}\, W_0^{*(i)}\left(\frac{t - \dfrac{V_i}{1+V_i}}{\dfrac{V_{i+1}}{1+V_{i+1}} - \dfrac{V_i}{1+V_i}}\right) : \right.$$
$$\left. \frac{V_i}{1+V_i} \le t \le \frac{V_{i+1}}{1+V_{i+1}}, i \in \boldsymbol{Z}\right\}.$$

So, to approximate the corresponding $\{W_0(t) : 0 \le t \le 1\}$ over the same interval we have constructed $\{K_0(t) : 0 \le t \le 1\}$, we have the options of constructing the process $\{W_0(t) : 0 \le t \le 1\}$ over a predetermined, equally spaced grid or a grid that accounts for the differing distances between neighboring vertex locations. Again, one would typically construct the corresponding $\{W_0(t) : 0 \le t \le 1\}$ over the chosen grid and linearly interpolate.

Appendix

PROOF OF THEOREM 3.1. We seek to find the joint distribution of $(\tau_1, K(\tau_1))$ where

$$\tau_1 = \sup\left\{x : W(x) - x = \sup_{u \ge 0}\{W(t) - u\}\right\} \quad \text{and} \quad K(\tau_1) = \sup_{u \ge 0}\{W(u) - u\} + \tau_1.$$

We begin by defining $M = \sup_{u \ge 0}\{W(u) - u\}$ and look to derive the joint density of (τ_1, M) using linear boundary crossing probabilities of Brownian motion and the fact that Brownian motion is a Gaussian process with the Markov property. Define $M_t = \sup_{0 \le u \le t}\{W(u) - u\}$. The conditional survival function of $M_t \mid W(t)$ is given by

$$\begin{aligned}
&\overline{H}_{M_t|W(t)}(m_t \mid w)\\
&\quad = P(M_t > m_t \mid W(t) = w)\\
&\quad = P(W(x) \text{ crosses } m_t + x \text{ over } (0, t) \mid W(t) = w)\\
&\quad = \exp\left[-\frac{2}{t}(m_t)^+(m_t + t - w)^+\right]
\end{aligned}$$

with conditional density of $M_t \mid W(t)$ given by

$$\begin{aligned} &h_{M_t|W(t)}(m_t \mid w) \\ &= \frac{2}{t}(2m_t + t - w) \exp\left\{-\frac{2}{t} m_t(m_t + t - w)\right\} \qquad m_t > \max(0, w - t). \end{aligned}$$

The density of $W(t)$ is that of a normal distribution with mean 0 and variance t. The joint survival function of (τ_1, M) is given by

$$\begin{aligned} \overline{F}_{\tau_1,M}(t,m) &= P(\tau_1 \geq t, M \geq m) \\ &= P\left(\sup_{u\geq t}\{W(u)-u\} \geq \sup_{0\leq u\leq t}\{W(u)-u\}, M \geq m\right) \\ &= \int_{-\infty}^{\infty}\int_{(w-t)^+}^{\infty} P\left(\sup_{u\geq t}\{W(u)-u\} \geq m_t, M \geq m \mid M_t = m_t, W(t) = w\right) \\ &\qquad \cdot h_{M_t|W(t)}(m_t \mid w) f_{W(t)}(w) dm_t dw. \end{aligned}$$

The key to solving this double integral is to break the double integral into subregions. These subregions are chosen based upon their effect on the support of the conditional density of $M_t \mid W(t)$ and the simplification of the conditional probability statement. We break the double integral into the following 4 regions

$$\int_{-\infty}^{t}\int_{0}^{m}, \quad \int_{t}^{t+m}\int_{w-t}^{m}, \quad \int_{-\infty}^{t+m}\int_{m}^{\infty}, \quad \text{and} \quad \int_{t+m}^{\infty}\int_{w-t}^{\infty}.$$

Over the first 2 regions, the conditional probability statement reduces to

$$\begin{aligned} &P\left(\sup_{u\geq t}\{W(u)-u\} \geq m_t, M \geq m \mid M_t = m_t, W(t) = w\right) \\ &\quad = P\left(\sup_{u\geq t}\{W(u)-u\} \geq m \mid M_t = m_t, W(t) = w\right) \\ &\quad = \exp\{-2(m+t-w)\} \end{aligned}$$

and over the second 2 regions, the conditional probability statement reduces to

$$\begin{aligned} &P\left(\sup_{u\geq t}\{W(u)-u\} \geq m_t, M \geq m \mid M_t = m_t, W(t) = w\right) \\ &\quad = P\left(\sup_{u\geq t}\{W(u)-u\} \geq m_t \mid M_t = m_t, W(t) = w\right) \\ &\quad = \exp\{-2(m_t+t-w)\}. \end{aligned}$$

We do not provide the tedious integration needed to obtain the joint survival function of (τ_1, M) which is given by

$$\begin{aligned} \overline{F}_{\tau_1,M}(t,m) &= P(\tau_1 \geq t, M \geq m) \\ &= \exp\{-2m\}\overline{\Phi}\left(\frac{t-m}{\sqrt{t}}\right) - 2\sqrt{t}\phi\left(\frac{t+m}{\sqrt{t}}\right) \\ &\quad + [1 + 2(t+m)]\overline{\Phi}\left(\frac{t+m}{\sqrt{t}}\right) \end{aligned}$$

where ϕ is the standard normal density and $\overline{\Phi}$ is the standard normal survival function. Differentiating, we obtain the joint density of (τ_1, M) given as the following

$$\begin{aligned} g_{\tau_1,M}(t,m) &= \frac{1}{2t^{3/2}}\left[\frac{1}{t}(t+m)^2-1\right]\exp\{-2m\}\phi\left(\frac{t-m}{\sqrt{t}}\right) \\ &- \frac{1}{2t^{3/2}}\left[\frac{1}{t}(t-m)^2-1\right]\phi\left(\frac{t+m}{\sqrt{t}}\right). \end{aligned}$$

A transformation is necessary to obtain the joint density of $(\tau_1, K(\tau_1)) \equiv (\tau_1, M+\tau_1)$. The Jacobian for this transformation is 1. Much simplification results and the joint density of $(\tau_1, K(\tau_1))$ is given by

$$g_{\tau_1,K(\tau_1)}(t,y) = \frac{2}{t^{3/2}}(y-t)\phi\left(\frac{y}{\sqrt{t}}\right) \qquad 0 < t \le y < \infty$$

where ϕ is the standard normal density function. □

Proof of Lemma 3.1. We have that the joint density of $(\tau_1, K(\tau_1))$ is given by

$$g_{\tau_1,K(\tau_1)}(t,y) = \frac{2}{t^{3/2}}(y-t)\phi\left(\frac{y}{\sqrt{t}}\right) \qquad 0 < t \le y < \infty$$

where ϕ is the standard normal density function. Consider the transformation where $(R,S) = (\frac{K(\tau_1)}{\sqrt{\tau_1}}, \sqrt{\tau_1})$ or $(\tau_1, K(\tau_1)) = (S^2, RS)$. The Jacobian of this transformation is $2s^2$. Thus, the joint density of (R,S) is

$$g_{R,S}(r,s) = 4(r-s)\phi(r) \qquad 0 < s \le r < \infty.$$

We can determine that $R \stackrel{d}{=} \sqrt{T}$ where T is a chi-square random variable with 3 degrees of freedom. Conditional on R, we see that S has a linear density over $[0,R]$ passing through the points $(0, \frac{2}{R})$ and $(R, 0)$, and so R is simply a scale parameter. Thus, $S \stackrel{d}{=} R(1-\sqrt{U})$. Therefore, $(R,S) \stackrel{d}{=} (\sqrt{T}, \sqrt{T}(1-\sqrt{U}))$. Finally, we have,

$$(\tau_1, K(\tau_1)) = (S^2, RS) \stackrel{d}{=} (T(1-\sqrt{U})^2, T(1-\sqrt{U})). \qquad \square$$

References

Carolan, C. and Dykstra, R. (2001). Marginal distributions of the least concave majorant of Brownian motion, *Ann. Statist.*, **29**, 1732–1750.

Groeneboom, P. (1983). The concave majorant of Brownian motion, *Ann. Probab.*, **11**, 1016–1027.

Itô, K. and McKean, H. P., Jr. (1974). *Diffusion Processes and Their Sample Paths*, 2nd ed., Springer, Berlin.

Pitman, J. W. (1983). Remarks on the convex minorant of Brownian motion, *Seminar on Stochastic Processes, 1982* (eds. E. Cinlar, K. L. Chung and R. K. Getoor), 219–228, Birkhäuser, Boston.

Williams, D. (1974). Path decomposition and continuity of local time for one-dimensional diffusions, I, *Proc. London Math. Soc. (3)*, **28**, 738–768.

Woodroofe, M. and Sun, J. (1999). Testing uniformity versus a monotone density, *Ann. Statist.*, **27**, 338–360.

Wu, W. B., Woodroofe, M. and Mentz, G. (2001). Isotonic regression: Another look at the changepoint problem, *Biometrika*, **88**, 793–804.