The Maximum Maximum of a Martingale

David G. Hobson

Abstract

Let $(M_t)_{0 \leq t \leq 1}$ be any martingale with initial law $M_0 \sim \mu_0$ and terminal law $M_1 \sim \mu_1$ and let $S \equiv \sup_{0 \leq t \leq 1} M_t$. Then there is an upper bound, with respect to stochastic ordering of probability measures, on the law of S.

An explicit description of the upper bound is given, along with a martingale whose maximum attains the upper bound.

21 May 1997

1 Introduction

Let μ_0 and μ_1 be probability measures on $\mathbb{R}$ with associated distribution functions $F_i[x] \equiv \mu_i((-\infty, x])$. Now let $\mathcal{M} \equiv \mathcal{M}(\mu_0, \mu_1)$ be the space of all martingales $(M_t)_{0\leq t\leq 1}$ with initial law μ_0 and terminal law μ_1. For such a martingale $M \in \mathcal{M}$ let $S \equiv \sup_{0\leq t\leq 1} M_t$ and denote the law of S by ν. In this short article we are interested in the set $\mathcal{P} \equiv \mathcal{P}(\mu_0, \mu_1) \equiv \{\nu; M \in \mathcal{M}\}$ of possible laws ν, and in particular we find a least upper bound for $\mathcal{P}$. The fact that M is a martingale imposes quite restrictive conditions on ν.

Clearly $\mathcal{M}$ is empty unless the random variables corresponding to the laws μ_i have the same finite mean, and henceforth we will assume without loss of generality that this mean is zero. Moreover a simple application of Jensen's inequality shows that a further necessary condition for the space to be non-empty is that

$$\int_x^\infty (y-x)\mu_0(dy) \leq \int_x^\infty (y-x)\mu_1(dy) \qquad \forall x. \tag{1}$$

These conditions are also sufficient, see for example Strassen [16, Theorem 2] or Meyer [10, Chapter XI].

The question described in the opening paragraph is a special case of a problem first considered in Blackwell and Dubins [4] and Dubins and Gilat [7]. There the authors derive conditions on the possible laws $\tilde{\nu}$ of the supremum $\tilde{S}$ of a martingale $(\tilde{M}_t)_{0\leq t\leq 1}$ whose terminal distribution $\tilde{\mu}_1$ is given, but whose initial law $\tilde{\mu}_0$ is *not* specified. Let $\preceq$ denote stochastic ordering on probability measures, (so that $\rho \preceq \pi$ if and only if $F_\rho[x] \geq F_\pi[x] \quad \forall x$, with the obvious notational convention) and let ρ^* denote the Hardy transform of a probability measure ρ. Then it follows from [4] and [7] that

$$\tilde{\mu}_1 \preceq \tilde{\nu} \preceq \tilde{\mu}_1^*. \tag{2}$$

Indeed, Kertz and Rösler [9] have shown that the converse to (2) also holds: for any probability measure ρ satisfying $\tilde{\mu}_1 \preceq \rho \preceq \tilde{\mu}_1^*$, there is a martingale with terminal distribution $\tilde{\mu}_1$ whose maximum has law ρ. If, moreover, ρ is concentrated on $[0, \infty)$ then the martingale $\tilde{M}$ can be taken to have initial law consisting of the unit mass at 0. See also Rogers [14] for a proof of these results based on excursion theory, and Vallois [17] for a discussion of the case where $\tilde{M}$ is a continuous martingale. Thus if $\mu_0 \equiv \delta_0$ (the unit mass at 0) then our problem is solved and

$$\mathcal{P}(\delta_0, \mu_1) \equiv \{\nu : \delta_0 \vee \mu_1 \preceq \nu \preceq \mu_1^*\}.$$

Otherwise, as Kertz and Rösler [9, Remark 3.3] observe,

$$\mathcal{P}(\mu_0, \mu_1) \subseteq \{\nu : \mu_0 \vee \mu_1 \preceq \nu \preceq \mu_1^*\}. \tag{3}$$

In a sense Kertz and Rösler [9, Theorem 3.4] answer our question of interest also. They describe necessary and sufficient conditions for a candidate probability measure ν to be a member of

$\mathcal{P}(\mu_0, \mu_1)$. These conditions involve displaying a pair of bivariate densities with marginals (μ_0, μ_1) and (μ_0, ν) and may be thought of as a restatement of the problem. In contrast the solution presented here is both explicit and constructive.

The main results of this article are that the set $\mathcal{P}(\mu_0, \mu_1)$ is bounded above by a probability measure $\mu^*_{0,1}$ (in the sense that if $\nu \in \mathcal{P}$ then $\nu \preceq \mu^*_{0,1}$), and that this upper bound is attained. Moreover we provide an explicit construction of this upper bound: we do so now for the nice special case where μ_1 has a continuous distribution. For $i = 0, 1$ define $\eta_i = \int_x^\infty (y - x)\mu_i(dy)$ and let $a(z)$ be the solution with $a(z) \geq z$ to the equation

$$\eta_0(a(z)) = \eta_1(z) + (a(z) - z)\eta_1'(z).$$

Pictorially $a(z)$ is the x-co-ordinate of the point where the tangent to η_1 at z intersects the graph of η_0. See Figure 1. Then $\mu^*_{0,1}$ is defined by $\mu^*_{0,1}((-\infty, x]) = \mu_1((-\infty, a^{-1}(x)])$. In Section 2 we prove this result and extend to arbitrary measures μ_1.

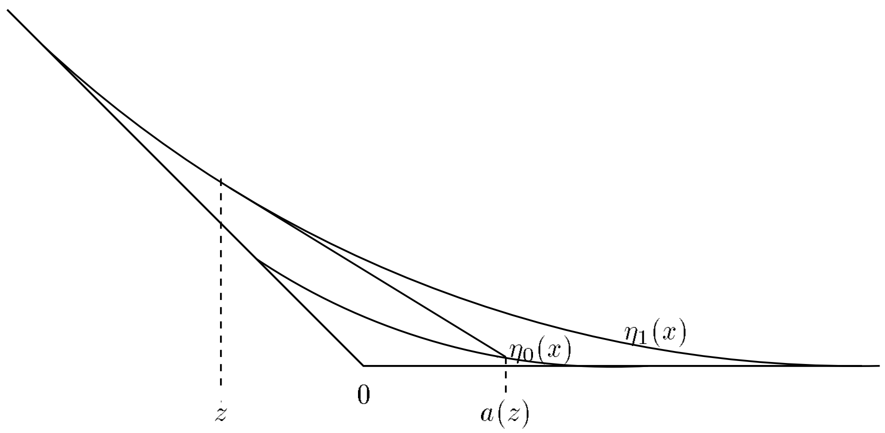


Figure 1: The functions $\eta_i(x)$ and $a(z)$.

Clearly one non-constructive definition of $\mu^*_{0,1}$ is via its distribution function $F^*_{0,1}$:

$$F^*_{0,1}[x] \equiv \inf_{\nu \in \mathcal{P}} F_\nu[x]. \tag{4}$$

There seems no reason *a priori* why the measure corresponding to (4) should be an element of $\mathcal{P}(\mu_0, \mu_1)$. Indeed if the greatest lower bound $\mu_*^{0,1}$ is defined via its distribution function $F_*^{0,1}$:

$$F_*^{0,1}[x] \equiv \sup_{\nu \in \mathcal{P}} F_\nu[x]; \tag{5}$$

then $\mu_*^{0,1}$ need not be an element of $\mathcal{P}$; in particular it is not in general true that $\mu_0 \vee \mu_1 \in \mathcal{P}$. See Section 3.4 for a simple example showing non-attainment of the lower bound.

The Skorokhod embedding theorem concerns the embedding of a given law in Brownian motion by construction of a suitably minimal stopping time. (Skorokhod embeddings for

Brownian motion and other processes remains an active area of research; see the recent paper by Bertoin and Le Jan [3] for a new class of suitable stopping rules.) We show that one martingale whose maximum attains the upper bound is a (time-change of) Brownian motion, and this explains why the prescient reader will recognise in the arguments expounded below elements of the Chacon and Walsh [5] and Azéma and Yor [2] proofs of the Skorokhod theorem (and also the Rogers [12] excursion theoretic version of the Azéma-Yor argument). An incidental remark in Section 3.3 indicates how these alternative derivations of the Skorokhod theorem are in fact closely related.

Finally some brief words on a motivation for studying this problem. Let M_t be the price process of a financial asset, and suppose that interest rates are zero. Then standard arguments from the theory of complete markets show that when pricing contingent claims or derivative securities it is natural to treat M as if it were a martingale. The simplest and most liquidly traded contingent claims are European call options which, at maturity T, have payoff $(M_T - k)^+$. Suppose now that instead of attempting to model M_t and thence to predict the prices of call options, we assume that the prices of calls are fairly determined by the market. From knowledge of call prices for all strikes k it is possible to infer the law (at least under the measure used for pricing derivatives) of M_T. Bounds on the prices of 'exotic' derivatives can be obtained by maximising the expected payoff of the exotic option over the space of martingales with the given (or rather the inferred) terminal distribution. These bounds depend on the market prices of call options, but they do not rely on any modelling assumptions which attempt to describe the underlying price process.

As an example, the lookback option is a security which at maturity T has value S_T, the maximum price attained by the asset over the interval $[0, T]$. Given the set of prices of call options with maturity T we can deduce the (implied) law of M_T under the pricing measure. Since M_0 is fixed and M_t is a martingale under the pricing measure, the problem of characterising the possible prices of a lookback security is solved once the possible laws of the maximum S_T have been determined. This is the problem under consideration in Blackwell and Dubins [4], and more generally here. See Hobson [8] for a more detailed analysis of the lookback option, the derivation of non-parametric bounds on the lookback price and the description of an associated hedging strategy.

Acknowledgement It is a pleasure to thank David Marles, Uwe Rösler and an anonymous referee for helpful and insightful comments on a previous version of this paper.

2 Main results

In this section we derive conditions relating the distribution of the maximum ν to the initial and terminal distributions μ_0, μ_1. First we recall some simple bounds which do not depend on the initial law μ_0.

Clearly $\mathbb{P}[M_1 > x] \leq \mathbb{P}[S > x]$ so that $\mu_1 \preceq \nu$. Define the non-decreasing barycentre function b_1 by

$$b_1(x) = \frac{\mathbb{E}[M_1; M_1 \geq x]}{\mathbb{P}[M_1 \geq x]}$$

for all x such that $\mathbb{P}[M_1 \geq x] > 0$, and $b_1(x) = x$ otherwise. By Doob's submartingale inequality

$$c\mathbb{P}[S \geq c] \leq \mathbb{E}[M_1; S \geq c].$$

Fixing c, then at least in the case where μ_1 has no atoms, there is some d with $\mathbb{P}[S \geq c] = \mathbb{P}[M_1 \geq d]$. Moreover it is trivial that $\mathbb{E}[M_1 - d; A] \leq \mathbb{E}[M_1 - d; M_1 \geq d]$ for all sets A, so that $\mathbb{E}[M_1; S \geq c] \leq \mathbb{E}[M_1; M_1 \geq d] = b_1(d)\mathbb{P}[M_1 \geq d]$. Thus $c \leq b_1(d)$ and it follows that $\nu \preceq \mu_1^*$, where μ_1^*, the Hardy transform of μ_1, has associated distribution function F_1^* given by $F_1^*[x] = F_1[b_1^{-1}(x)]$. The proof in the case where μ_1 contains atoms requires only minor modifications.

The above paragraph, which follows Blackwell and Dubins [4] closely, contains a proof of (2), and provides many of the essential arguments we will use in Proposition 2.1 to find an upper bound for $\mathcal{P}$. Our purpose is to consider the effect of fixing the initial law of the martingale.

For $i = 0, 1$ define the functions $\eta_i(x)$ by

$$\eta_i(x) = \int_x^\infty (1 - F_i[y])dy = \int_x^\infty (y - x)\mu_i(dy) = \mathbb{E}[(M_i - x)^+].$$

The functions η_i are positive, decreasing and convex with $\eta_i(x) \geq -x$. Recall from the introduction that a necessary and sufficient condition for $\mathcal{M}(\mu_0, \mu_1)$ to be non-empty is that $\eta_0(x) \leq \eta_1(x)$ for all x. Henceforth we assume that this condition is satisfied. Furthermore, if $\eta_1(x) = \eta_0(x)$ then

$$\begin{aligned}
\mathbb{E}((M_1 - x)^+; M_0 \leq x) &= \mathbb{E}((M_1 - x)^+) - \mathbb{E}((M_1 - x)^+; M_0 > x) \\
&\leq \eta_1(x) - \mathbb{E}((M_0 - x)^+; M_0 > x) \\
&= 0,
\end{aligned}$$

so that $\mathbb{P}(M_1 > x, M_0 \leq x) = 0$. Similarly $\mathbb{P}(M_1 < x, M_0 \geq x) = 0$ so that if μ_0 associates any mass with a point $x \in \{z : \eta_1(z) = \eta_0(z)\}$ then we must have that the martingale M is constant on the set $M_0 = x$, and μ_1 must include a corresponding atom. By considering such atoms separately, and by dividing the set $I = \{x : \eta_1(x) > \eta_0(x)\}$ into its constituent intervals we can reduce to the case where I takes the form of a single interval $I \equiv (i^-, i^+) \subseteq (-\infty, \infty)$. and assume further that μ_0 places no mass at the endpoints of I.

We now construct functions a, α and β which will play a crucial role in subsequent analysis. As motivation, suppose temporarily that μ_1 has no atoms. Note that the derivative of η_1 is given by $\eta_1'(x) = -\mathbb{P}[M_1 > x] \equiv F_1[x] - 1$. For $z \le i^-$ and $z \ge i^+$ let $a(z) = z$ and otherwise define $a(z)$ to be the unique solution with $a(z) > z$ to the equation

$$\eta_0(a(z)) = \eta_1(z) + (a(z) - z)\eta_1'(z). \tag{6}$$

$a(z)$ is the x-coordinate of the point where the tangent to η_1 at z, taken in the direction of increasing z, intersects with the function $\eta_0(z)$. Recall Figure 1. The function a is non-decreasing, and on I it satisfies

$$a(dz)(\eta_0'(a(z)) - \eta_1'(z)) = \mu_1(dz)(a(z) - z).$$

Define $F^*_{0,1}$ to be the distribution function given by

$$F^*_{0,1}[x] = F_1[a^{-1}(x)].$$

(The composition $F_1 \circ a^{-1}$ is well defined since μ_1 assigns no mass to intervals where a is constant.) Let $\mu^*_{0,1}$ be the associated probability measure; $\mu^*_{0,1}$ will be the upper bound on $\mathcal{P}(\mu_0, \mu_1)$.

For the general case where μ_1 has atoms we let the above argument guide our intuition. For $u \in (0,1)$ define $\beta(u) = \inf\{x : F_1[x] \ge u\}$. The parameter u will play the role of defining the slope of the relevant tangent. Define h_u via

$$h_u(z) = \eta_0(z) + z(1-u) - \eta_1(\beta(u)) - \beta(u)(1-u).$$

Then h_u is a convex function; if $h_u(x) < 0$ then h_u has a unique root in (x, ∞). On $\eta_1(\beta(u)) = \eta_0(\beta(u))$ set $\alpha(u) = \beta(u)$, and on $\eta_1(\beta(u)) > \eta_0(\beta(u))$ let $\alpha(u)$ be the root in $(\beta(u), \infty)$ of h_u; then the function α satisfies

$$\eta_0(\alpha(u)) = \eta_1(\beta(u)) - (\alpha(u) - \beta(u))(1-u). \tag{7}$$

Informally, $\alpha(u)$ is the x-coordinate of the point where the tangent to η_1 with gradient $-(1-u)$ intersects the graph of η_0. If F_1 has an atom of mass v at i^+ then for all $u > 1 - v$ we have $\alpha(u) = \beta(u)$; meanwhile $\alpha(u) > \beta(u)$ on $(0, 1-v)$. The function α is continuous and has (left)-derivative

$$\frac{d\alpha}{du} = \frac{\alpha(u) - \beta(u)}{\eta_0'(\alpha(u)) + 1 - u}, \tag{8}$$

where η_0' is again a left-derivative. By the convexity of η_0, and the definition of α as the x-coordinate of the point where a line with slope $-(1-u)$ intersects η_0, it is clear that for $u \in (0,1)$ the denominator must be positive. Finally define the measure $\mu^*_{0,1}$ via its distribution function

$$F^*_{0,1}[x] \equiv \inf\{u : \alpha(u) > x\}. \tag{9}$$

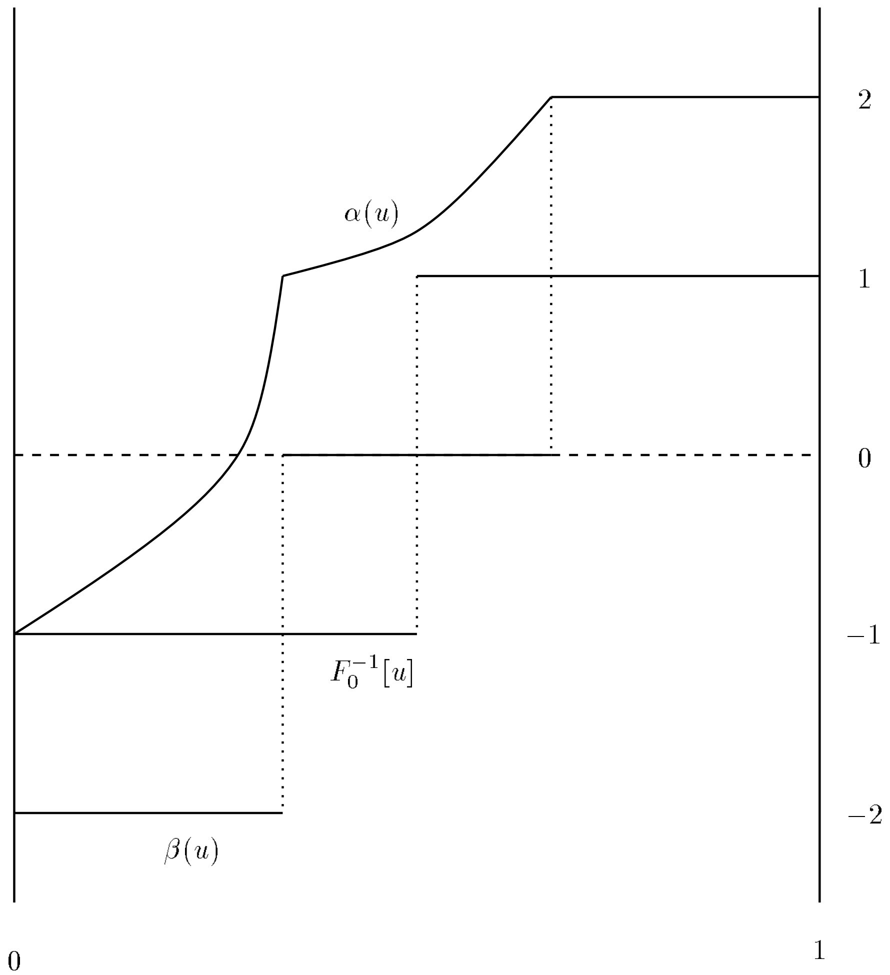
Where defined we have that $\alpha(u) = a(\beta(u))$, and the two definitions of $\mu^*_{0,1}$ agree.

Example 2.1 Examples always help to make things clearer...
Let μ_0, μ_1 be the uniform measures on $\{-1, 1\}$, $\{-2, 0, 2\}$ respectively. In particular μ_1 is discrete so that a is not well defined, and this example illustrates the general method. Then $\eta_0(x) = \max\{-x, 0, (1-x)/2\}$ and $\eta_1(x) = \max\{-x, 0, 2(1-x)/3, (2-x)/3\}$. Further

$$\alpha(u) = 2I_{(u>2/3)} + 2/(3(1-u))I_{(1/3<u\leq 2/3)} + (4u-1)/(1-2u)I_{(u\leq 1/3)}.$$

See Figure 2 for a pictorial representation of the functions α, β and F_0^{-1}. Note that α is continuous and non-decreasing, and that for $u \in (0,1)$, $\alpha(u) \geq \beta(u+) \vee F_0^{-1}[u+]$.

Figure 2: α, β and F_0^{-1} for the distributions in Example 2.1.



Proposition 2.1 *Let M be a martingale with the desired initial and terminal distributions. Denote by ν the law of the maximum process S. Then*

$$\nu \preceq \mu^*_{0,1} \qquad \forall \nu \in \mathcal{P}(\mu_0, \mu_1),$$

so that $\mu^*_{0,1}$ is an upper bound for $\mathcal{P}(\mu_0, \mu_1)$.

Proof

We prove that $F_\nu^{-1}[u] \leq \alpha(u)$ for all $u \in (0,1)$, where F_ν is the distribution function associated with the law ν. Since F_ν^{-1} and α are increasing functions it is sufficient to prove that if $\mathbb{P}(S > c) = 1 - F_\nu[c] = 1 - u$, then $c \leq \alpha(u)$.

Fix c, then by Doob's submartingale inequality

$$c\mathbb{P}[S > c, M_0 \leq c] \leq \mathbb{E}[M_1; S > c, M_0 \leq c].$$

Similarly, by the martingale property,

$$\mathbb{E}[M_0; M_0 > c] = \mathbb{E}[M_1; S > c, M_0 > c].$$

Adding these two expressions yields after some elementary manipulations

$$c\mathbb{P}[S > c] + \eta_0(c) \leq \mathbb{E}[M_1; S > c].$$

Now suppose $\mathbb{P}[S > c] = 1 - u$, then, since for any set A it is true that $\mathbb{E}[Y; A] \leq \mathbb{E}[Y; Y \geq 0]$,

$$\begin{aligned} \mathbb{E}[M_1; S > c] &= \mathbb{E}[M_1 - \beta(u); S > c] + (1-u)\beta(u) \\ &\leq \mathbb{E}[M_1 - \beta(u); M_1 \geq \beta(u)] + (1-u)\beta(u). \end{aligned}$$

Using (7) we can summarise these inequalities:

$$\begin{aligned} c(1-u) + \eta_0(c) &\leq \eta_1(\beta(u)) + (1-u)\beta(u) \\ &= \eta_0(\alpha(u)) + \alpha(u)(1-u). \end{aligned}$$

It is easy to see by the convexity of η_0 that $x(1-u) + \eta_0(x)$ is increasing on $(F_0^{-1}[u+], \infty)$. Since $\alpha(u) \geq F_0^{-1}[u+]$ it follows that $c \leq \alpha(u)$.

$\square$

Inspection of the proof of Proposition 2.1 reveals that the upper bound is attained if both the martingale is continuous, (so that there is equality in Doob's inequality), and the sets $(S > c)$ can be identified with sets of the form $(M_1 > d)$. Guided by these observations our goal now is to construct a martingale M with initial law μ_0 and terminal law μ_1, whose maximum has law $\mu^*_{0,1}$. We ensure continuity of the martingale M by basing the construction on a stopping time for a Brownian motion. Moreover the stopping rule is a function of the current maximum of the Brownian motion and its current value.

Rösler [15] provides an alternative construction of a martingale whose maximum attains the upper bound. This construction is in the spirit of arguments given by Blackwell and Dubins [4] and Kertz and Rösler [9]. In some respects the Rösler construction is simpler than the methods presented below; the advantage of the methods we use is that they provide a solution of the Skorokhod problem.

Suppose that an initial point x_0 is chosen according to the distribution μ_0. For motivation consider first the case where μ_1 has a continuous distribution function and the function a and its inverse are well defined. Let B be a Brownian motion started at x_0 and define

$$\begin{aligned} S_t^B &= \sup_{0 \le u \le t} B_u \\ \tau &= \inf\{u : B_u \le a^{-1}(S_u^B)\} \end{aligned}$$

Note that τ is almost surely finite for if $y > x_0$ and H_y denotes the first hitting time of the Brownian motion B at level y then $\tau \le \inf\{t > H_y : B_t \le a^{-1}(y)\}$.

We show that, when averaged over the law of the starting point, B_τ has the law μ_1. Then also

$$\mathbb{P}[S_\tau^B > x] = \mathbb{P}[a(B_\tau) > x],$$

and the distribution function of S^B is given by

$$F_1[a^{-1}(x)] \equiv F_{0,1}^*[x]. \tag{10}$$

It will follow (in Corollary 2.1 below) that $\mu_{0,1}^*$ defined via (9) *is* an element of $\mathcal{P}(\mu_0, \mu_1)$. Our goal now is to prove the above claim that B_τ has law μ_1, in the setting of a general probability measure μ_1.

Proposition 2.2 *For a Brownian motion B with initial law μ_0 define $S_t^B = \sup_{0 \le u \le t} B_u$ and $\tau = \inf\{u : F_1[B_u] \le F_{0,1}^*[S_u^B]\}$. Then $B_\tau \sim \mu_1$ and $S_\tau^B \sim \mu_{0,1}^*$.*

Proof

This follows directly from Chacon and Walsh [5] although here we provide a direct proof, similar in spirit to Azéma and Yor [2]. For a connection between these two approaches, see Section 3.3.

Suppose that μ_1 has an atom of size v at i^+, and thence that $\mu_{0,1}^*$ has an atom of at least this size there also.

With α, β and τ all as above define the random variable Z via $Z = F_{0,1}^*(S_\tau^B)$. Then $F_1(B_\tau) \ge Z \ge F_1(B_\tau -)$ so that $\beta(Z) \equiv B_\tau$ and, since α is continuous, $\alpha(Z) \equiv S_\tau^B$. We find the law of Z: it is sufficient to show that Z has the uniform distribution on $[0, 1]$, or more particularly, and to allow for atoms in $\mu_{0,1}^*$, it is sufficient to show that $\mathbb{P}(Z \le u) = u$, for $0 < u < 1 - v$. Then

$$\mathbb{P}(B_\tau \le x) = \mathbb{P}(\beta(Z) \le x) = \mathbb{P}(Z \le F_1[x]) = F_1[x],$$

and similarly $\mathbb{P}(S_\tau^B \le x) = F_{0,1}^*[x]$.

Return to the consideration of the law of Z: for a test function Φ (continuous and compactly supported) set $g = \Phi \circ \alpha^{-1}$ and define $G(x) = \int_0^x g(u)du$. Note that $G(\alpha(x)) = \int_0^x \Phi(u)d\alpha(u)$. Itô calculus shows that $N_t \equiv G(S_t^B) - (S_t^B - B_t)g(S_t^B)$ is a continuous local martingale, moreover the stopped martingale N^τ is bounded (see [2] for details). Therefore one has

$$
\begin{aligned}
0 = \mathbb{E}[N_\tau - N_0] &= \mathbb{E}[G(S_\tau^B) - (S_\tau^B - B_\tau)g(S_\tau^B) - G(S_0^B)] \\
&= \mathbb{E}[G(\alpha(Z)) - G(B_0)] - \mathbb{E}[(\alpha(Z) - \beta(Z))g(\alpha(Z))]. \qquad (11)
\end{aligned}
$$

Let π denote the law of Z, and let $\overline{F}_\pi[x] \equiv 1 - F_\pi[x] = \pi((x, \infty))$. Note that π has support contained in the interval $[0, 1]$. Then

$$\mathbb{E}[G(\alpha(Z))] = \int_{\mathbb{R}} \pi(dz) \int_0^z \Phi(u)d\alpha(u) = \int_0^1 \overline{F}_\pi[u]\Phi(u)d\alpha(u).$$

For the second term a simple transformation of variable yields

$$\mathbb{E}[G(B_0)] = \int_0^1 \overline{F}_0[\alpha(u)]\Phi(u)d\alpha(u)$$

and (11) becomes

$$\int_0^1 (\overline{F}_\pi[u] - \overline{F}_0[\alpha(u)])\Phi(u)d\alpha(u) = \int_0^1 (\alpha(u) - \beta(u))\Phi(u)\pi(du).$$

Since Φ is arbitrary, π must satisfy the identity

$$(\alpha(u) - \beta(u))\pi(du) = (\overline{F}_\pi[u] - \overline{F}_0[\alpha(u)])d\alpha(u). \qquad (12)$$

Substituting from (8) gives that at least for $x < 1 - v$, (whence $\alpha(u) > \beta(u)$ for all $u \in (0, x)$)

$$F_\pi[x] - x = \int_0^x \left(\frac{F_0[\alpha(u)] - F_\pi[u]}{F_0[\alpha(u)] - u} - 1 \right) du = - \int_0^x \frac{F_\pi[u] - u}{F_0[\alpha(u)] - u} du.$$

and it follows that $F_\pi[x] = x$ for $x < 1 - v$. □

Corollary 2.1 *The measure $\mu_{0,1}^*$ is an element of $\mathcal{P}(\mu_0, \mu_1)$.*

Proof

It suffices to show that $M_t \equiv B(\tau \wedge (t/(1-t))$ is a true martingale and not just a local martingale, or equivalently that $(B_{t\wedge\tau})_{t\geq 0}$ is uniformly integrable. This follows by a straightforward extension to Lemma 2.3 in Rogers [14] and an appeal to Theorem 1 in Azéma, Gundy and Yor [1]. □

3 Remarks

3.1 The case $\mu_0 = \delta_0$ and reduction to previous results.

If $\mu_0 \equiv \delta_0$ then $\eta_0(x) = x^- \equiv (-x) \vee 0$. Suppose that μ_1 has a continuous distribution function then since $a(x) \geq 0$ the formula (6) becomes

$$a(x) = \frac{\eta_1(x) - x\eta_1'(x)}{|\eta_1'(x)|}.$$

The function $a(x)$ is easily shown to equal the barycentre function $b_1(x)$ and, for general μ_1 also, the results of the previous section become the twin statements that μ_1^* is both an element of, and an upper bound for, $\mathcal{P}(\delta_0, \mu_1)$.

3.2 Excursion-theoretic arguments

The defining equation for π given in (12) can be derived using excursion arguments. Readers who would like an introduction to excursion theory are referred to Rogers [13]. For simplicity consider the case where μ_1 has no atoms and as before let B be a Brownian motion with maximum process S^B such that the initial point B_0 is chosen according to the law μ_0. Define $\tau = \inf\{u : a(B_u) \leq S_u^B\}$.

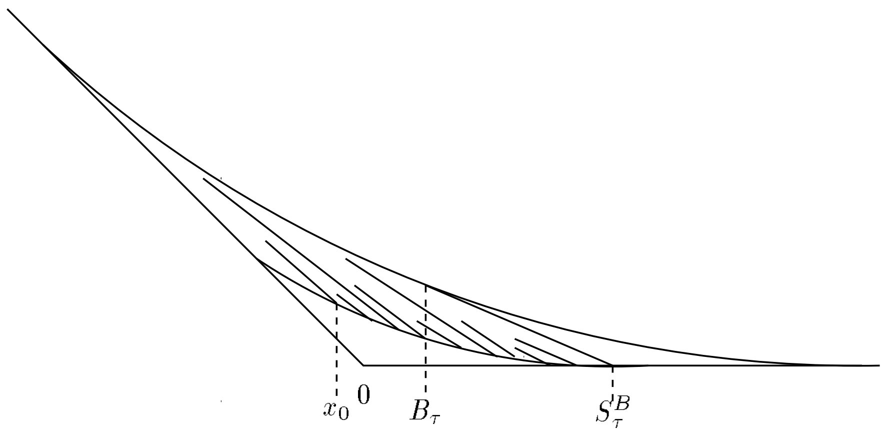


Figure 3: Some of the excursions below $S_t^B = s$ plotted along the tangents to η_1 at $a^{-1}(s)$

Consider splitting the Brownian path into excursions below its maximum. Imagine plotting the excursions from the maximum $S_t^B = s$, in xy space, in such a way that the x-component is always B_t, and the y-component is chosen to keep the two-dimensional process on the tangent to η_1 which joins $(s, \eta_0(s))$ with $(a^{-1}(s), \eta_1(a^{-1}(s)))$. As the maximum s increases, so the line along which excursions are plotted changes. See Figure 3. τ is then the first time that one of these excursions first meets the curve $(x, \eta_1(x))$.

By construction $\mathbb{P}[B_\tau \le y] = \mathbb{P}[B_0 \le a(y)] - \mathbb{P}[B_0 \le a(y), S^B_\tau \ge a(y)]$. Now for the event $(B_\tau \in dy)$ to occur it must be true that both $(B_0 \le a(y))$ and $(S^B_\tau \ge a(y))$, and then, before the maximum rises to $a(y+dy)$, there must be an excursion down from the maximum of relative depth at least $a(y) - y$. Using Lévy's identity in law of the pairs $(S^B, S^B - B)$ and $(L^B, |B|)$, and the fact that the local time rate of excursions of height in modulus at least x is x^{-1}, it follows that

$$\mathbb{P}[B_\tau \in dy] = \mathbb{P}[B_0 \le a(y), S^B_\tau \ge a(y)] \frac{a(dy)}{a(y) - y}.$$

Define $\psi(y) = \mathbb{P}[B_\tau \le y] - F_1[y]$. We wish to show that $\psi(y) \equiv 0$. Now

$$\begin{aligned} \psi(y+dy) - \psi(y) &= \mathbb{P}[B_\tau \in dy] - \mu_1(dy) \\ &= (F_0[a(y)] - F_1[y] - \psi(y)) \frac{a(dy)}{a(y) - y} - \mu_1(dy) \\ &= -\frac{\psi(y)}{F_0[a(y)] - F_1[y]} \mu_1(dy), \end{aligned}$$

where this last line follows from the identity

$$\frac{a(dy)}{a(y) - y} = \frac{\mu_1(dy)}{F_0[a(y)] - F_1[y]}.$$

It must follow that $\psi(y) \equiv 0$.

3.3 The Skorokhod Embedding Theorem

The proof of the Skorokhod embedding theorem given in Chacon and Walsh [5] can be expressed pictorially in a similar manner to Sections 2 and 3.2. One version of their algorithm, (see in particular Dubins [6]), involves a sequence of the following steps: firstly choose a value x and draw the tangent to η_1 at x; and secondly run a Brownian motion until it leaves some interval defined via this tangent (and the history of the construction to date). Our construction is a special case in which the values x chosen at each step are as small as possible. By the remarks in Section 3.1, when $\mu_0 \equiv \delta_0$ the function $a(x)$ is equivalent to the barycentre function $b_1(x)$ and the argument of Section 2 reduces to the Azéma-Yor proof. Thus the Azéma and Yor [2] and Rogers [14] proofs of the Skorokhod embedding theorem are seen to be special cases of the proof due to Chacon and Walsh [5].

3.4 The minimum maximum of a martingale

A *maximal* (respectively *minimal*) element of a set of measures $\mathcal{S}$ is a measure $\rho \in \overline{\mathcal{S}}$ for which there does *not* exist $\nu \in \mathcal{S}$ with $\nu \succ \rho$ (respectively $\nu \prec \rho$). The results of Section 2 show that

$\mu^*_{0,1}$ is the *unique* maximal element of $\mathcal{P}(\mu_0, \mu_1)$. However it is not in general true that $\mathcal{P}$ has a unique minimal element.

It is clear that if M is any martingale with the desired initial and terminal laws, and if $\tilde{M}$ is the martingale consisting of a single jump such that $\tilde{M}_t \equiv M_0$ for $0 \leq t < 1$ and $\tilde{M}_1 \equiv M_1$, then the law of the maximum of M stochastically dominates that of $\tilde{M}$. Thus the study of minimal elements reduces to a study of discrete parameter martingales at the timepoints 0 and 1.

Suppose the probability measures μ_0 and μ_1 are the Uniform measures on $\{-1, 1\}$ and $\{-2, 0, 2\}$ respectively. The joint laws of all (single jump) martingales with these initial and terminal distributions are parameterised by θ, $(0 < \theta < 1/12)$, in the following table:

		M_0^θ	
		-1	+1
M_1^θ	-2	(1/4) + θ	(1/12) - θ
	0	(1/4) - 2θ	(1/12) + 2θ
	2	θ	(1/3) - θ

Table 1: The joint law of M^θ

If S^θ denotes the supremum of the martingale M^θ parameterised by θ then $\mathbb{P}[S^\theta > x] = \mathbb{P}[M_0^\theta \vee M_1^\theta > x]$. In particular $\mathbb{P}[S^\theta \geq 0] = (3/4) - \theta$ and $\mathbb{P}[S^\theta > 0] = (1/2) + \theta$. It is impossible to minimise both these expressions simultaneously. For this simple example there is a non-degenerate family of minimal elements of $\mathcal{P}(\mu_0, \mu_1)$ and the greatest lower bound of $\mathcal{P}(\mu_0, \mu_1)$ is not attained. Further $\mu_0 \vee \mu_1$ is not an element of $\mathcal{P}(\mu_0, \mu_1)$.

3.5 The minimum maximum of a continuous martingale

If the martingales M are further constrained to be continuous, then as Perkins [11] has shown, if the initial law is trivial, then there is a unique minimal element to the set of possible laws of the maximum.

Specifically, let μ_0 have zero mean and let $\mathcal{M}_C \equiv \mathcal{M}_C(\delta_0, \mu_1)$ be the space of all continuous martingales $(M_t)_{0 \leq t \leq 1}$ which are null at 0, and have terminal law μ_1. Let $\mathcal{P}_C \equiv \mathcal{P}_C(\delta_0, \mu_1)$ be the set of laws of the associated maxima. Then $\mathcal{P}_C$ has a unique minimal element, which can be represented as a Skorokhod embedding of a Brownian motion. See [11] for further details of this construction.

References

[1] Azéma, J., Gundy, R.F. and Yor, M; Sur l'integrabilité uniforme des

martingales continues, *Séminaire de Probabilités*, **XIV**, 53–61, 1980.

[2] Azéma, J and Yor, M; Une solution simple au problème de Skorokhod, *Séminaire de Probabilités*, **XIII**, 90–115, 1979.

[3] Bertoin, J. and Le Jan, Y.; Representation of measures by balayage from a regular recurrent point, *Annals of Probability*, **20**, 538–548, 1992.

[4] Blackwell, D. and Dubins, L.D.; A Converse to the Dominated Convergence Theorem, *Illinois Journal of Mathematics*, **7**, 508–514, 1963.

[5] Chacon, R. and Walsh, J.B.; One-dimensional potential embedding, *Séminaire de Probabilités*, **X**, 19–23, 1976.

[6] Dubins, L.D.; On a theorem of Skorokhod, *Annals of Mathematical Statistics*, **39**, 2094–2097, 1968.

[7] Dubins, L.D. and Gilat, D.; On the distribution of maxima of martingales, *Proceedings of the American Mathematical Society*, **68**, 337–338, 1978.

[8] Hobson, D.G.; Robust hedging of the lookback option, Submitted to *Finance and Stochastics*, 1996.

[9] Kertz, R.P. and Rösler, U.; Martingales with given maxima and terminal distributions, *Israel J. Math.*, **69**, 173–192, 1990.

[10] Meyer, P.A.; *Probability and Potentials*, Blaidsell, Waltham, Mass., 1966.

[11] Perkins, E.; The Cereteli-Davis solution to the H^1-embedding problem and an optimal embedding in Brownian motion, *Seminar on Stochastic Processes, 1985*, Birkhauser, Boston, 173–223, 1986.

[12] Rogers, L.C.G.; Williams' characterisation of the Brownian excursion law: proof and applications, *Séminaire de Probabilités*, **XV**, 227–250, 1981.

[13] Rogers, L.C.G.; A guided tour through excursions, *Bull. London Math. Soc.*, **21**, 305–341, 1989.

[14] Rogers, L.C.G.; The joint law of the maximum and the terminal value of a martingale, *Prob. Th. Rel. Fields*, **95**, 451–466, 1993.

[15] Rösler, U.; Personal communication.

[16] Strassen, V.; The existence of probability measures with given marginals, *Ann. Math. Statist.*, **36**, 423–439, 1965.

[17] Vallois, P.; Sur la loi du maximum et du temps local d'une martingale continue uniformément intégrable, *Proc. London Math. Soc.*, **3**, 399–427, 1994.