ISRAEL JOURNAL OF MATHEMATICS, Vol. 63, No. 3, 1988

ON THE RATIO OF THE EXPECTED MAXIMUM OF A MARTINGALE AND THE L_p-NORM OF ITS LAST TERM

BY
DAVID GILAT
School of Mathematical Sciences, Tel Aviv University, Ramat Aviv 69978, Israel

ABSTRACT

For each $p > 1$, the supremum, S, of the absolute value of a martingale terminating at a random variable X in L_p, satisfies

(1) $$ES \leqq (\Gamma(q+1))^{1/q} \| X \|_p \qquad (q = p(p-1)^{-1}).$$

The maximum, M, of a mean-zero martingale which starts at zero and terminates at X, satisfies

(2) $$EM \leqq \sigma_q \| X \|_p,$$

where σ_q is the unique solution of the equation $t = \| Z - t \|_q$ for an exponentially distributed random variable Z with mean 1. σ_p has other characterizations and satisfies $\lim_{p\downarrow 1} q^{-1}\sigma_q = c$ with c determined by $ce^{c+1} = 1$. Equalities in (1) and (2) are attainable by appropriate martingales which can be realized as stopped segments of Brownian motion. A presumably new property of the exponential distribution is obtained en route to inequality (2).

Introduction

Let S be the supremum of a nonnegative submartingale (or — which in view of Gilat [6] is equivalent — of the absolute value of a martingale) whose last term is the random variable X. According to Theorem 3.4 in Doob [3] (see also Hardy's inequality on p. 240 of Hardy, Littlewood and Polya [9]), for $p > 1$, the L_p-norm of S is no larger than $q = p(p-1)^{-1}$ times the L_p-norm of X. In fact, as pointed out by Dubins and Gilat [4] as well as by Hardy, Littlewood and Polya on p. 240 of [9], q is the (unattainable, except for the identically zero-martingale) least upper bound of the ratio $\| S \|_p / \| X \|_p$. For $p = 1$,

Received December 27, 1987

however, $\| S \|_1 = ES$ may well be infinite, unless $E|X|\log^+|X| < \infty$. It is therefore of interest to investigate how fast can the ratio $R(p) = ES/\| X \|_p$ grow when p approaches 1 from above. Of course, since $ES \leqq \| S \|_p$, q is always an upper bound on $R(p)$, but, as it turns out, q can be further reduced to yield

$$ES \leqq (q!)^{1/q} \| X \|_p, \tag{1}$$

where $q!$ stands for $\Gamma(q+1) = \int_0^\infty x^q e^{-x} dx$. Since $q^{-1}(q!)^{1/q}$ decreases to e^{-1} as $q \to \infty$ (i.e. $1 < p \to 1$), the obvious bound q is thus asymptotically improved by a factor of e^{-1}.

In Section 2 inequality (1) is proved, conditions for equality are investigated and extremal distributions for X and S, for which equality in (1) is attained, are obtained.

An analogous sharp inequality is obtained for the supremum, M, of a mean-zero martingale which starts at zero (so that $M \geqq 0$) and terminates at X. In this case the least upper bound of the ratio $r(p) = EM/\| X \|_p$ is a somewhat more complicated function of p than the one for the ratio $R(p)$. To introduce it, let Z be an exponentially distributed random variable with $EZ = 1$ and, for $q > 1$, set

$$\sigma_q = \inf_{-\infty < t < \infty} \| Z - t \|_q,$$

then

$$EM \leqq \sigma_q \| X \|_p \qquad (p > 1, q = p(p-1)^{-1}). \tag{2}$$

There seems to be no explicit expression for σ_q as a function of q. Inequality (2) and its sharpness are established in Section 4. Investigation of conditions for equality in (2) led to the discovery of a curious, presumably hitherto unnoticed, property of the exponential distribution, which yielded three additional formulations of σ_q. These results, as well as the asymptotic behaviour of σ_q, are presented in Section 3 as preliminaries for the discussion of inequality (2).

The proofs of both inequalities, (1) and (2), are based on a result of Blackwell and Dubins [2] according to which the Hardy and Littlewood [8] maximal function, corresponding to (the distribution of) the last term of a martingale, stochastically dominates the maximum of every martingale terminating with that distribution. It is an easy corollary that the Blackwell–Dubins result remains valid if in its statement, *martingale* is replaced by *submartingale*. The attainability of equality in both (1) and (2) then follows from Dubins and Gilat

[4] according to which, for every distribution with finite mean, there exists a martingale terminating with that distribution, whose maximum is distributed like the Hardy and Littlewood maximal function of its terminal distribution. These ideas are reviewed in Section 1. The rest of the proof is easy sailing through Hölder's inequality with a careful analysis of conditions for equality to derive the extremal distributions. The final Section 5 is devoted to indicating how equality in (1) and (2) can be attained by a segment of Brownian motion determined by an appropriate stopping time.

Recently, inequality (1) was independently obtained by Jacka [10] using an entirely different method of proof. For the interesting special case $p = q = 2$, both inequalities (1) and (2) were recently obtained by Dubins and Schwarz [5]. This article was motivated by the desire to extend the results of [5].

1. Review of the main tools of proof

In this section, the Hardy and Littlewood maximal function and its role in martingale theory are briefly reviewed. For further details the reader is referred to Gilat [7].

Given a random variable X with a finite mean, let $f = f_X$ be the (essentially unique) nonincreasing function on (0, 1), whose distribution (with respect to Lebesgue measure) is the same as the distribution of X. The function $F = F_X$ defined by

$$F(x) = \frac{1}{x} \int_0^x f(t)dt \qquad (0 < x \leqq 1)$$

is then the *Hardy and Littlewood* (HL for short) *maximal function* associated with (the distribution of) X.

THEOREM 1.1 (Blackwell and Dubins [2]). *The Lebesgue-distribution of F_X stochastically dominates the supremum of every martingale whose last term is distributed like X.*

COROLLARY 1.1. *Theorem* 1.1 *remains valid when, in its statement,* martingale *is replaced by* submartingale.

PROOF. If $\{X_t\}$ is a submartingale adapted to the filtration $\{\mathcal{F}_t\}$ and X is an integrable random variable such that $E(X \mid \mathcal{F}_t) \geqq X_t$ for all t, then the martingale $\{E(X \mid \mathcal{F}_t)\}$ majorizes the given submartingale $\{X_t\}$. Consequently, by Theorem 1.1, $\sup_t E(X \mid \mathcal{F}_t)$, and *a fortiori* $\sup_t X_t$, is stochastically dominated by the HL maximal function of $E(X \mid \mathcal{F}_\infty)$, where $\mathcal{F}_\infty$

is, of course, the smallest sigma algebra containing $\bigcup \mathscr{F}_t$. Applying now Theorem 1.1 to the martingale-pair $(E(X \mid \mathscr{F}_\infty), X)$ establishes the corollary. □

THEOREM 1.2 (Dubins and Gilat [4]). *For every distribution with finite mean, there exists a martingale, whose last term has the given distribution and whose supremum is distributed like the HL maximal function of that distribution.*

Furthermore, if the initially given distribution is supported in $(0, \infty)$, *then the martingale, whose existence is claimed, can be made nonnegative.*

2. Proof of inequality (1), conditions for equality and extremal distributions

Let S and X be as in the opening sentence of the introduction. Here and throughout the rest of the paper, for $a > 0$, $a!$ is used to abbreviate $\Gamma(a+1) = \int_0^\infty x^a e^{-x} dx$; and for $p > 1$, $q = p(p-1)^{-1}$.

THEOREM 2.1.

(i) *For every* $p > 1$, $ES \leqq (q!)^{1/q} \| X \|_p$.

(ii) *For each* $p > 1$, *there exists a nonnegative (proper) martingale for which equality holds in* (i).

(iii) *In the extremal case of equality in* (i), *the distributions of X and S are uniquely determined (up to a common scale factor) as follows: X must have the Weibull distribution with shape parameter* $p-1$, *i.e. for some* $\lambda > 0$,

$$P(X \geqq y) = \exp[-(\lambda y)^{p-1}], \qquad y \geqq 0.$$

To describe the corresponding extremal distribution for S, let Z be an exponential random variable with mean 1, *then S must be distributed like*

$$\lambda^{-1} \int_0^\infty (t+Z)^{q-1} e^{-t} dt.$$

NOTE. For $p = 1 + n^{-1}$ (n being a positive integer), the last integral becomes a polynomial in Z, of degree n, with easily identifiable coefficients. In the special case, $p = q = 2$, treated in [5], (i) and (ii) yield $\sqrt{2}$ as the attainable least upper bound of the ratio $ES / \| X \|_2$ whereas (iii) says that equality can be attained only if X is exponentially distributed and S is distributed like $X + EX$.

PROOF. Given a nonnegative submartingale with last term X and supremum S, let $f = f_X$ and $F = F_X$ be the functions introduced in Section 1. By Theorem 1.1,

$$(2.1)\qquad ES \leqq \int_0^1 F(x)dx = \int_0^1 \left(\frac{1}{x}\int_0^x f(t)dt\right)dx = \int_0^1 f(x)\ln\frac{1}{x}\,dx,$$

where the last equality is obtained by changing order of integration ($f \geqq 0$ because $X \geqq 0$). Assuming $X \in L_p$ (of the probability space on which it is defined), hence $f \in L_p(0, 1)$, and applying Hölder's inequality, one obtains

$$(2.2)\qquad \int_0^1 f(x)\ln\frac{1}{x}\,dx \leqq \left\| \ln\frac{1}{x} \right\|_q \| f \|_p = (q!)^{1/q} \| f \|_p = (q!)^{1/q} \| X \|_p.$$

Combining (2.2) and (2.1) establishes statement (i) of Theorem 2.1.

REMARK. Hopefully no confusion arises from using $\| \cdot \|_p$ to denote both the L_p-norm of the Lebesgue unit interval and the L_p-norm of the probability space on which X happens to be defined. This practice will be used in the sequel without further comment.

To obtain equality in (i), equalities must be attained in both (2.1) and (2.2). As for the Hölder-inequality in (2.2), recall that equality holds if f^p is proportional to $(\ln(1/x))^q$, i.e.

$$f(x) = \left(\frac{1}{\lambda}\ln\frac{1}{x}\right)^{q-1} \qquad \text{for some } \lambda > 0.$$

Since the Lebesgue-distribution of $\ln(1/x)$ is exponential, the stated Weibull distribution is obtained for f. Now use Theorem 1.2 to construct the nonnegative martingale corresponding to this Weibull distribution to also obtain equality in (2.1). According to Theorem 1.2, the S of this martingale is distributed like the HL maximal function of the said Weibull distribution. An elementary change of variable of integration then yields the representation stated in (iii) for the extremal distribution of S. □

3. A digression about the exponential distribution

Suppose Y is a random variable with finite second moment and let $L(t) = \| Y - t \|_2$, $-\infty < t < \infty$. It is then both well known and elementary that the function L has a unique minimum at $t = m \equiv$ mean of Y and $\min L = L(m) = \sigma \equiv$ standard deviation of Y. If in addition $\sigma = m$, as is for example the case when Y is exponentially distributed, then the function L has the nice property

$$(3.1)\qquad \text{minimum of } L = \text{minimizer of } L,$$

where *minimizer of* L is the point at which L attains its minimum. Geometrically, (3.1) means that the lowest point on the graph of L lies on the diagonal $D = \{(z, y) : y = x\}$ of the plane. It turns out that for the exponential distribution the validity of (3.1) is not restricted to the second moment. Formally, let Y be a random variable with finite qth moment ($q \geqq 1$) and set $L_q(t) = \| Y - t \|_q$, $-\infty < t < \infty$. It is easy to verify that L_q is convex, has no intervals of constancy and $L_q(t)$ tends to $+\infty$ as $t \to \pm\infty$. Consequently L_q has a unique minimum. Denote by m_q the value of t at which L_q attains its minimum and let $\sigma_q = L_q(m_q) = \min L_q$.

PROPOSITION 3.1. *It Y is either uniformly distributed on the two-point set $\{0, 2m\}$ with $m > 0$, or exponentially distributed, then*

$(*)$ $\sigma_q = m_q$ *for all* $q \geqq 1$; *i.e.* (3.1), *with* $L = L_q$, *holds for all* $q \geqq 1$.

REMARKS. (1) I do not know whether this property of the exponential distribution has ever been noticed before.

(2) The family of functions $\{L_q : q \geqq 1\}$ is increasing in q, therefore their minima σ_q, are increasing. Consequently, if $(*)$ is valid for Y, then m_q is increasing. If $(*)$ holds and m_q is bounded, then one can prove that m_q is necessarily constant, say m, in which case Y has the symmetric distribution on the two-point set $\{0, 2m\}$. This partial converse to Proposition 3.1 was obtained jointly with Jon Aaronson.

(3) On the other hand, I do not know to what extent does property $(*)$ characterize the exponential distribution when m_q is unbounded. This problem will be studied elsewhere.

PROOF OF PROPOSITION 3.1. Property $(*)$ is obviously invariant under (positive) scale change. Consequently it suffices to prove $(*)$ for the symmetric distribution on $\{0, 2\}$ and for the exponential distribution with mean 1. In the first case, a trivial calculation shows that $m_q = \sigma_q = 1$ for all $q \geqq 1$. For the second case, let Z be an exponentially distributed random variable with $EZ = 1$ and set $\varphi_q(t) = (L_q(t))^q$, then

$$\begin{aligned}\varphi_q(t) &= E\,|Z - t|^q \\ &= \int_t^\infty (x - t)^q e^{-x} dx + \int_0^t (t - x)^q e^{-x} dx \\ &= e^{-t}\left(q! + \int_0^t x^q e^x dx\right).\end{aligned}$$

Differentiation with respect to t yields

$$\varphi_q'(t) = t^q - \varphi_q(t).$$

Since φ_q is smooth (in addition to being convex, having no intervals of constancy and satisfying $\varphi_q(\pm\infty) = +\infty$), its unique minimum, hence also that of L_q, is attained at the unique solution of the equation $\varphi_q'(t) = 0$. By the foregoing this is equivalent to $(\varphi_q(t))^{1/q} = L_q(t) = t$. Thus m_q is the unique solution of the equation $L_q(t) = t$ and consequently $\sigma_q \equiv \min L_q = L_q(m_q) = m_q$. □

The next result shows that for an exponential distribution, m_q is symptotically a linear function of q, and identifies its asymptotic slope.

PROPOSITION 3.2. *For the exponential distribution with mean* 1, $\lim_{q\to\infty}(m_q/q) = c$, *where c is determined by the equation* $\ln c + c + 1 = 0$. *Numerically*, $c \doteq 0.27846 < e^{-1} \doteq 0.36788 = \lim_{q\to\infty} q^{-1}(q!)^{1/q}$.

PROOF. Recall from the proof of the previous proposition that m_q is defined as the unique t which satisfies the equation

$$t^q = E\,|Z - t\,|^q = e^{-t}\left(q! + \int_0^t z^q e^x dx\right)$$

or, equivalently (multiplying by e^t and integrating by parts),

$$t^q e^t = q! + t^q e^t - q\int_0^t x^{q-1}e^x dx.$$

Thus m_q can be characterized by the equation

$$q! = q\int_0^{m_q} x^{q-1}e^x dx = \int_0^{m_q} e^x d(x^q), \tag{3.2}$$

or, by putting $x = m_q y$ in the last integral,

$$q! = (m_q)^q \int_0^1 e^{m_q y} d(y^q).$$

Anticipating the conclusion of the proposition, set $m_q = qc(q)$ in the last equation and take the qth root on both of its sides to obtain

$$\begin{aligned} \frac{1}{q}(q!)^{1/q} &= c(q)\left[\int_0^1 (e^{c(q)y})^q d(y^q)\right]^{1/q} \\ &= c(q)\,\| e^{c(q)y} \|_{q;d(y^q)}, \end{aligned} \tag{3.3}$$

where $\|\cdot\|_{q;d(y^q)}$ is the L_q-norm of the measure $d(y^q)$ on the unit interval. Now, let q tend to ∞ on both sides of (3.3): the left hand side converge (downwards) to e^{-1} while the right hand side is clearly asymptotic to $c(q)\max_{0\leq y\leq 1} e^{c(q)y} = c(q)e^{c(q)}$. Consequently, $c(q)e^{c(q)}$ decreases to e^{-1} as $q \to \infty$. Since $x \to xe^x$ is monotone increasing, it follows that $c(q)$ is decreasing, and its limit c satisfies the equation $ce^c = e^{-1}$ which is equivalent to $\ln c + c + 1 = 0$. □

REMARK. If Y is exponential with $EY = \theta$, then clearly $m_q(Y) = \theta m_q(Z)$.

4. Proof of inequality (2), conditions for equality and extremal distributions

In this section, I return to the main theme of the article. Let M be the supremum of a mean-zero martingale starting at zero (so that $M \geqq 0$) and terminating at the random variable X.

THEOREM 4.1. *For every $p > 1$, $EM \leqq \sigma_q \| X \|_p$ $(q = p(p-1)^{-1})$ where, Z being an exponential random variable with $EZ = 1$, σ_q is given by the following four equivalent characterizations:*

(i) $\sigma_q = \min_{-\infty<t<\infty} \| Z - t \|_q$.

(ii) *σ_q is the unique t which minimizes the function $t \to \| Z - t \|_q$.*

(iii) *σ_q is the unique solution of the equation $\| Z - t \|_q = t$.*

(iv) *σ_q is determined by the equations $(q-1)! = \int_0^{\sigma_q} x^{q-1}e^x dx$.*

Moreover, for each $p > 1$, there is a mean-zero martingale for which equality is attained.

The extremal distributions for X and M, attaining equality, are uniquely determined (up to scale): X must be distributed like a positive multiple of $g_{\sigma_q}(Z)$, where Z is again an exponential random variable with mean 1 and the functions g_t are given by: $g_t(x) = (x-t)^{q-1}$ for $x \geqq t$ and $-(t-x)^{q-1}$ for $x < t$. Furthermore, σ_q is the unique t for which $g_t(Z)$ has mean-zero. When equality prevails, M must be distributed as the HL maximal function of $g_{\sigma_q}(Z)$.

PROOF. Given a martingale starting at zero and terminating at X, let $f = f_X$ and $F = F_X$ be, once again, the functions introduced in Section 1. Then $F(1) = \int_0^1 f(x)dx = EX = 0$, hence $F \geqq 0$. Initially proceed as in the proof of Theorem 2.1: First, apply Theorem 1.1 to obtain

$$EM \leqq \int_0^1 F(x)dx = \int_0^1 \left(\frac{1}{x}\int_0^x f(u)du\right)dx = \int_0^1 f(x)\ln\frac{1}{x}dx$$

(4.1)

$$= \int_0^1 \left(\ln\frac{1}{x} - t\right) f(x)dx \qquad \text{for all real } t,$$

where the last equality is a consequence of $\int_0^1 f(x)dx = 0$. Next, assuming $X \in L_p$, hence $f \in L_p(0, 1)$, and applying Hölder's inequality, one obtains

$$\int_0^1 F(x)dx = \int_0^1 \left(\ln\frac{1}{x} - t\right) f(x)dx \leqq \left\| \ln\frac{1}{x} - t \right\|_q \| f \|_p. \tag{4.2}$$

Since (4.2) is valid for all t and in view of (4.1), one obtains

$$EM \leqq \inf_{-\infty < t < \infty} \left\| \ln\frac{1}{x} - t \right\|_q \| f \|_p,$$

which is clearly an equivalent form of the desired inequality. That (i), (ii) and (iii) define the same constant is a consequence of Proposition 3.1. That this constant is also determined by (iv) follows from formula (3.2) in the proof of Proposition 3.2.

To obtain the extremal distribution for X or, equivalently, for f, conditions for equality in (4.2) have to be investigated. Since the functions whose product is integrated in the middle term of (4.2) are not of constant sign, to obtain equality in (4.2), the following two conditions are necessary and sufficient

(a) $|f|^p$ is proportional to $|\ln(1/x) - t|^q$, or, equivalently $|f|$ is proportional to $|\ln(1/x) - t|^{q-1}$.

(b) The integrand, $f(x)(\ln(1/x) - t)$, has to be nonnegative, or, equivalently, since both factors are decreasing, f and $(\ln(1/x) - t)$ have to change sign (from $+$ to $-$) at the same point x, namely at $x = e^{-t}$.

Consequently (a) and (b) are simultaneously satisfied iff f is a positive multiple of the function f_t given by:

$$f_t(x) = \begin{cases} \left(\ln\frac{1}{x} - t\right)^{q-1}, & 0 < x \leqq e^{-t}, \\ -\left(t - \ln\frac{1}{x}\right)^{q-1}, & e^{-t} < x \leqq 1. \end{cases} \tag{4.3}$$

It remains to find a value of t for which $\int_0^1 f_t(x)dx = 0$, where f_t is defined by (4.3). This task is equivalent to finding t such that

$$\int_0^{e^{-t}} \left(\ln\frac{1}{x} - t\right)^{q-1} dx = \int_{e^{-t}}^{1} \left(t - \ln\frac{1}{x}\right)^{q-1} dx. \tag{4.4}$$

Making the obvious change of variable ($x = e^{-y}$) in (4.4), an elementary calculation leads to

$$(4.5) \qquad (q-1)! = \int_0^t x^{q-1} e^x dx.$$

The step leading from (4.4) to (4.5) is of course reversible, hence both determine the same value of t. However, by formula (3.2), the t determined by (4.5) is precisely $m_q = \sigma_q$.

Now, if one constructs the Dubins–Gilat martingale (of Theorem 1.2) corresponding to f_{σ_q} (which clearly has the same distribution as $g_{\sigma_q}(Z)$), one gets a mean-zero martingale whose maximum, M, is the HL maximal function of f_{σ_q}, whose last term is f_{σ_q} and for which $EM = \sigma_q \parallel \text{last term} \parallel_p$. □

REMARK. It is curious to note that $\sigma_2 = 1$, thus the ratio between the expected maximum of a mean-zero martingale and the standard deviation of its last term cannot exceed one.

5. Optimal embedding in Brownian motion

In this final section, we indicate how equalities in (1) and (2) can also be attained by segments of Brownian motion determined by appropriate stopping times $\tau = \tau_p$.

Once the extremal distribution of X is determined, embedding it in Brownian motions by the Aźema and Yor [1] stopping method (for the case of a mean-zero martingale, i.e. inequality (2)), or by its modification for the absolute value of Brownian motion due to Don van der Vecht [12] (for the case of a nonnegative submartingale, i.e. inequality (1)), yields the desired segments of Brownian motion.

Acknowledgements

I wish to thank the Wiskunde en Informatica Institute of Vrije Universiteit, Amsterdam and the Vakgroep Stochastiek of the Delft University of Technology for their warm hospitality and support while parts of this research were conducted. Special thanks are due to Mike Keane of Delft for helping me obtain the asymptotic behaviour of σ_q in (2). Finally, I wish to thank Lester Dubins for calling my attention to Jacka's work after I had informed him of (1).

REFERENCES

1. J. Ażema and M. Yor, (a) *Une solution simple au problème de Skorokhod*, (b) *Le problème de Skorokhod: complements a l'exposé precedent*, Sem. Prob. **XIII**, Lecture Notes in Math. **721**, Springer-Verlag, Berlin, 1978.

2. D. Blackwell and L. E. Dubins, *A converse to the dominated convergence theorem*, Illinois J. Math. **7** (1963), 508–514.

3. J. L. Doob, *Stochastic Processes*, Wiley, New York, 1953.

4. L. E. Dubins and D. Gilat, *On the distribution of maxima of martingales*, Proc. Am. Math. Soc. **68** (1978), 337–338.

5. L. E. Dubins and G. Schwarz, *A sharp inequality for semi-martingales and stopping times*, unpublished manuscript (1987).

6. D. Gilat, *Every nonnegative submartingale is the absolute value of a martingale*, Ann. Prob. **5** (1977), 475–481.

7. D. Gilat, *The best bound in the* $L \log L$ *inequality of Hardy & Littlewood and its martingale counterpart*, Proc. Am. Math. Soc. **97** (1986), 429–436.

8. G. H. Hardy and J. E. Littlewood, *A maximal theorem with function theoretic applications*, Acta Math. **54** (1930), 81–116.

9. G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, 2nd ed., Cambridge University Press, 1952.

10. S. D. Jacka, *Optimal stopping and best constants for Doob-like inequalities*, Handwritten draft, 1987.

11. I. Meilijson and A. Nadas, *Convex majorization with an application to the length of critical paths*, J. Appl. Prob. **16** (1979), 671–677.

12. D. P. van der Vecht, *Inequalities for stopped Brownian motion*, CWI tract No. 21, Amsterdam, 1986.