Inequalities in Statistics and Probability
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 78–83

SOME SHARP MARTINGALE INEQUALITIES RELATED TO DOOB'S INEQUALITY

BY David C. Cox
Battelle Memorial Institute

Let $p > 1$. The best constant $C = C_{n,p}$ in the inequality $E(\max_{1 \le i \le n} |Y_i|)^p \le C\, E|Y_n|^p$, where $Y_1, \ldots, Y_n$ is a martingale, is determined. For each n and p, the method allows one to construct a martingale attaining equality. As $n \to \infty$, $K_p n^{2/3}(q^p - C_{n,p}) \to 1$, where K_p is a known constant. As an application, the classical inequality of Doob is sharp. It is shown that equality cannot be attained by a non-zero martingale.

1. Introduction. Let $Y_1, Y_2, \ldots$ be a martingale with difference sequence $X_1 = Y_1$, $X_i = Y_i - Y_{i-1}$, $i = 2, 3, \ldots$ Thus, $E(X_i|X_1, \ldots, X_{i-1}) = 0$, $i = 2, 3, \ldots$. Let $p > 1$ and define $q = p/(p-1)$. The principal purpose of this paper is to determine the best constant $C = C_{n,p}$ in the inequality

$$E(\max_{1 \le i \le n} |Y_i|)^p \le CE|Y_n|^p. \tag{1.1}$$

Although $C_{n,p}$ is found in implicit form, it can be easily approximated. For each n and p, the method allows one to construct a martingale attaining equality in (1.1), with $C = C_{n,p}$. Once the distribution of Y_1 is fixed, such a martingale is uniquely determined.

Furthermore, as $n \to \infty$, $C_{n,p} \to q^p$ at a rate proportional to $n^{-2/3}$. Specifically, $K_p n^{2/3}$ $(q^p - C_{n,p}) \to 1$, where K_p is a known constant. As an application, this provides a new proof that Doob's inequality (1953, p. 317)

$$E(\sup_{i \ge 1} |Y_i|)^p \le q^p \sup_{i \ge 1} E|Y_i|^p \tag{1.2}$$

is sharp. An example to that effect was given previously by Dubins and Gilat (1978). Inequality (1.2) is strengthened to

$$E(\sup_{i \ge 1} |Y_i|)^p \le q^p \sup_{i \ge 1} E|Y_i|^p - q\, E|Y_1|^p. \tag{1.3}$$

It follows from (1.3) that equality cannot be attained in (1.2) by a non-zero martingale. The sharpness of Doob's inequality for $p = 1$ (1953, p. 317)

$$E(\sup_{i \ge 1} |Y_i|) \le [e/(e-1)](1 + E(\sup_{i \ge 1} |Y_i| \log^+ \sup_{i \ge 1} |Y_i|)],$$

is still an open question.

The method of this paper is based on results from the theory of moments (Kemperman (1968)), together with induction and the device of conditioning. Where applicable, it always leads to a sharp inequality and provides an example of a martingale attaining equality or nearly so. In principle, the method can be applied to many other martingale inequalities. For example, the author used it (Cox (1982)) to find the best constant in Burkholder's weak-L^1 inequality (Burkholder (1979)) for the martingale square function. The method does have the drawback of computational complexity, which sometimes makes it difficult or impossible to push the calculations through.

Section 2 contains statements of the results, together with comments and some proofs. In section 3, some needed analytic lemmas are established. Section 4 contains the main proofs, and an example for the case $p = 2$, $n = 3$ of (1.1).

AMS 1980 subject classifications. Primary 60E15, 60G42; secondary, 42B25.
Key words and phrases: Martingales, sharp inequalities, theory of moments, Doob's inequality.

2. Results. Throughout the paper, p is fixed. Dependence on p will often be suppressed in the notation. Let s, t be real numbers with $|t| \leqslant |s|$. For $0 < A \leqslant 1$ and $n = 1, 2, 3, \ldots$, define

$$\phi_n(s, t, A) = \inf E[|t + \Sigma_{i=1}^n X_i|^p - A \max(|s|^p, |t + X_1|^p, \ldots, |t + \Sigma_{i=1}^n X_i|^p)],$$

where the infimum is taken over all martingale difference sequences $X_1, \ldots, X_n$ with $EX_1 = 0$. The idea here is that $C_{n,p}^{-1}$ can be defined as the largest A for which $\phi_{n-1}(t, t, A) \geqslant 0$ for all t. Thus, to determine the best constant in (1.1), only the case $s = t$ of ϕ_n is needed. However, the inductive step (2.2) below requires knowing the value of ϕ_n for $|s| > |t|$ also. Note that $\phi_n(s, t, A) = |t|^p \phi_n(s/t, 1, A)$, for $t \neq 0$. This reduction does not, however, simplify the calculations made in the paper.

One has

$$\phi_1(s, t, A) = \inf\{E[|t + X|^p - A \max(|s|^p, |t + X|^p)]: EX = 0\}, \tag{2.1}$$

and by induction, conditioning on $X_1 = X$,

$$\phi_{n+1}(s, t, A) = \inf\{E\phi_n(\max(|s|, |t + X|), t + X, A): EX = 0\}, \tag{2.2}$$

for $n = 1, 2, \ldots$. Both (2.1) and (2.2) involve evaluating $\inf Ef(X)$ over all random variables X with $EX = 0$, where f is a given function. This is a standard problem of the theory of moments (Kemperman (1978), Cox (1982)) and can be solved graphically as shown in the proof of Theorem 1 given in Section 4.

THEOREM 1. *For $n = 1, 2, \ldots$, there exists $A_n \in (q^{-p}, 1]$, together with a function $g_n: (0, A_n] \to [0, 1)$ for which*

$$\begin{aligned} \phi_n(s, t, A) &= |t|^p - A|s|^p \quad \text{if } |t| \leqslant g_n(A)|s|, \\ &= p\, g_n(A)^{p-1}|t|\,|s|^{p-1} - [A + (p-1)g_n(A)^p]|s|^p \quad \text{if } g_n(A)|s| \leqslant |t| \leqslant |s|, \end{aligned}$$

for $0 \leqslant A \leqslant A_n$, while $\phi_n(s, t, A) = -\infty$ if $A > A_n$.

The constant A_n and the function g_n are defined inductively as follows. Let

$$\phi(y,x) = \{x/((p-1)[(py^{p-1} - (p-1)y^p - x)^{1-q} - 1])\}^{1/p} \tag{2.3}$$

for $0 \leqslant x, y \leqslant 1$ and $0 \leqslant \gamma(y,x) \equiv py^{p-1} - (p-1)y^p - x < 1$. Define $g_0(x) \equiv 1$, $g_{n+1}(x) = \phi(g_n(x), x)$, $n = 0, 1, \ldots$. Then, for $n = 1, 2, \ldots$, there is a unique $q^{-p} < A_n \leqslant 1$ with $g_n(A_n) = 0$; the domain of g_n is precisely $(0, A_n]$. One has $A_n > A_{n+1} > q^{-p}$ and $\lim_{n\to\infty} A_n = q^{-p}$. More precisely,

THEOREM 2. $\lim_{n\to\infty} n^{2/3}(A_n - q^{-p}) = (2\pi^2 q^{1-3p})^{1/3}$.

For $0 < x \leq q^p$, the sequence $\{g_n(x)\}$ is strictly decreasing with limit $g(x) = y$, the larger of the two roots of the equation

$$\phi(y, x) \equiv (p-1)(y^p - y^{p-1}) + x = 0.$$

In particular, $g(q^{-p}) = q^{-1}$.

COROLLARY 1. *Let $n \geqslant 2$ and $0 < A \leqslant A_{n-1}$. Suppose that $Y_1, \ldots, Y_n$ is a martingale. The following inequality is sharp*

$$E|Y_n|^p \geqslant AE(\max_{1\leqslant i\leqslant n}|Y_i|)^p + \gamma(g_{n-1}(A), A)E|Y_1|^p.$$

Proof: Let $X_i = Y_{i+1} - Y_i$, $i = 1, \ldots, n-1$. Then, conditional on $Y_1 = t$, $X_1, \ldots X_{n-1}$ is a martingale difference sequence with $EX_1 = 0$. Now apply Theorem 1 with $s = t$, and then integrate with respect to the distribution of Y_1. □

COROLLARY 2. *The best constant $C = C_{n,p}$ in (1.1) is $C_{n,p} = A_n^{-1}$.*

Proof. One has $\gamma(g_{n-1}(A), A) \geqslant 0$ for $0 < A \leqslant A_n$ with equality iff $g_n(A) = 0$, i.e., iff $A = A_n$. Now apply Corollary 1. □

The proof of Theorem 1, presented in Section 4, shows how a martingale attaining equality in (1.1), with $C = A_n^{-1}$, may be constructed. Moreover, once the distribution of Y_1 is fixed, such a martingale is uniquely determined. An example is given after the proof of Theorem 1.

From Theorem 2, the asymptotic behavior of $C_{n,p}$ can be characterized.

COROLLARY 3. $\lim_{n\to\infty} n^{2/3}(q^p - C_{n,p}) = (2\pi^2 q^{3p+1})^{1/3}$.

Letting $n\to\infty$ in Corollary 1, one obtains

COROLLARY 4. *Let $Y_1, Y_2, \ldots$ be a martingale, and $0 < A \leqslant q^{-p}$. The following inequality is sharp*

$$\sup_{i\geqslant 1} E|Y_i|^p \geqslant A\, E(\sup_{i\geqslant 1}|Y_i|)^p + g(A)^{p-1}E|Y_1|^p.$$

In particular, letting $A = q^{-p}$, one obtains (1.3).

Proof. Just note that $\gamma(g(A), A) = g(A)^{p-1}$, see Theorem 2. □

COROLLARY 5. *Doob's inequality (1.2) is sharp. However, equality cannot be attained by a non-zero martingale.*

Proof. Sharpness follows from $C_{n,p} \to q^p$. Equality in (1.2) forces $Y_1 \equiv 0$, from (1.3). Applying the same argument to the martingale $Y_2, Y_3, \ldots$, one finds $Y_2 = 0$, etc. □

3. Analytic Preliminaries. The object of this section is to establish some needed results concerning the functions g_n.

LEMMA 1. *The function ϕ, defined by (2.3), has the following properties.*

(3.1) $$\phi(y, x) \leqslant y \text{ with equality iff } \theta(y, x) = 0$$

(3.2) $$\delta\phi/\delta y > 0, \text{ for } 0 < x, y < 1, 0 < \gamma(y, x) < 1.$$

(3.3) $$\delta\phi/\delta x < 0, \text{ for all } y, \text{ if } q^{-p} < x < 1.$$

Proof. First consider (3.1). One has $\phi(y, x) \leqslant y$ iff

(3.4) $$(p-1)y^p[\gamma(y, x)^{1-q} - 1] - x \geqslant 0.$$

The derivative of the LHS of (3.4) with respect to x is given by $y^p\gamma(y, x)^{-q} - 1$. Since $\gamma(y, x) + \theta(y, x) = y^{p-1}$, it follows that the minimum value of the LHS of (3.4) is 0, taken when $\theta(y, x) = 0$. Since $\theta(y, x) > 0$ for all $y > 0$ when $x > q^{-p}$, one has $\phi(y, x) < y$ for all y in this case. Next, a straightforward calculation gives

(3.5) $$\delta\phi/\delta y = [(q-1)x y^{p-2}(1-y)] / [\phi^{p-1}\gamma^q(\gamma^{1-q} - 1)^2]$$

where $\phi = \phi(y, x)$, $\gamma = \gamma(y, x)$, which establishes (3.2). Finally,

(3.6) $$\delta\phi/\delta x = [\gamma - \gamma^q - (q-1)x] / [p(p-1)\phi^{p-1}\gamma^q(\gamma^{1-q} - 1)^2]$$

The numerator in (3.6) is $(1 - q)\theta(\gamma^{q-1}, x)$, which yields (3.3). □

LEMMA 2. *There is a unique $q^{-p} < A_n \leqslant 1$ with $g_n(A_n) = 0$, $n = 1, 2, \ldots$; the domain of g_n is $(0, A_n]$. One has $\theta(g_n(x), x) > 0$ for $0 < x \leqslant A_n$. For $0 < x \leqslant q^{-p}$, $g_n(x) \downarrow g(x)$ as $n\to\infty$.*

Proof. First consider $0 < x \leqslant q^{-p}$. I claim that $1 \geqslant g_n(x) > g(x)$ for all $n = 0, 1, 2, \ldots$. Since this is trivial for $n = 0$, assume that it holds for some $n \geqslant 0$. Then,

$$1 > \gamma(g_n(x), x) > \gamma(g(x), x) = g(x)^{p-1} > 0,$$

so that $g_{n+1}(x)$ is defined. Next,

$$(p-1)g(x)^p[\gamma(g_n(x),x)^{1-q}-1] < (p-1)g(x)^p[\gamma(g(x),x)^{1-q}-1] = x.$$

It follows that $g_{n+1}(x) > g(x)$. From Lemma 1, $g_{n+1}(x) < g_n(x)$, so that $g_n(x)\downarrow g(x)$ as $n\to\infty$, since $y = \lim_{n\to\infty} g_n(x)$ must satisfy $\phi(y,x) = y$. Suppose next that, for some $n \geqslant 1$, it has been established that the domain of g_n is $(0, A_n]$ with $g_n(A_n) = 0$, where $q^{-p} < A_n \leqslant 1$. I claim that $g_n'(x) < 0$ for $q^{-p} < x < A_n$. This is clear from (3.3) for $n = 1$, since $g'_1(x) = \delta\phi/\delta x$. Since $g'_{j+1}(x) = \delta\phi/\delta y\, g_j'(x) + \delta\phi/\delta x$, for $j = 1, 2, \ldots$, the claim follows by induction from (3.2) and (3.3). By the same argument, $g_{n+1}(x)$ is strictly decreasing on its domain, for $x > q^{-p}$. Since $g_{n+1}(x) < g_n(x)$ where both are defined, and $g_{n+1}(q^{-p}) > g(q^{-p}) > 0$, the existence and uniqueness of $A_{n+1} < A_n$ follow. Finally, $\theta(g_n(x), x) > 0$ for $0 < x \leqslant A_n$ follows from Lemma 1. □

4. Main Proofs.

Proof of Theorem 1. The properties of g_n and A_n a relevant to this proof have been established in Section 3. If one defines $\phi_0(s, t, A) = |t|^p - A|s|^p$, then the theorem holds for $n = 0$. Moreover, see (2.1) and (2.2), the inductive relation between ϕ_n and ϕ_{n+1} remains valid for $n = 0$. Assume by induction, therefore, that the theorem is true for some $n \geqslant 0$. Let $0 < A \leqslant A_n$ (where $A_0 = 1$), and, without loss of generality, $t \geqslant 0$. From (2.2) one finds

$$\phi_{n+1}(s, t, A) = \inf\{Eh(X): EX = t\},$$

where $h(x)$ is given by

$$\begin{array}{ll} |x|^p - A|s|^p & \text{if } |x| \leqslant g_n(A)|s| \\ p\, g_n(A)^{p-1}|s|^{p-1}|x| - [A + (p-1)g_n(A)^p]\,|s|^p & \text{if } g_n(A)|s| \leqslant |x| \leqslant |s| \\ \gamma(g_n(A), A)|x|^p & \text{if } |x| > |s|. \end{array}$$

It is well-known (Kemperman (1968), Cox (1982)) that the required infimum is given by the height, at location $x = t$, of the lower boundary of the convex hull of the graph of h. For $A_{n+1} < A \leqslant A_n$, $\gamma(g_n(A), A) < 0$ so the infimum is $-\infty$. Now suppose $0 < A \leqslant A_{n+1}$. Clearly, $h'(x)$ is continuous at $x = \pm\, g_n(A)|s|$ so that $h(x)$ is convex for $|x| < |s|$, and also for $|x| > |s|$. Moreover, $h'_+(|s|) = \gamma(g_n(A), A)\, p\, |s|^{p-1} < p\, g_n(A)^{p-1}\, |s|^{p-1} = h'_-(|s|)$, since $\theta(g_n(A), A) > 0$. It follows that the convex hull of the graph of $h(x)$ for $x \geqslant 0$ is formed by drawing a common tangent from the part for $0 \leqslant x \leqslant g_n(A)|s|$ to the part for $x > |s|$. The tangent to $y = |x|^p - A|s|^p$ at $x = x_0 > 0$ has equation

$$y = x_0^p(1-p) - A|s|^p + p\, x_0^{p-1}x. \tag{4.1}$$

The slope of $h(x)$ for $x > |s|$ is $p\, \gamma(g_n(A), A)x^{p-1}$, which coincides with the slope of (4.1) iff $x_0 = \gamma(g_n(A), A)^{q-1}x$. It follows that the required common tangent has a point of tangency at $x_0 = \phi(g_n(A), A)|s| = g_{n+1}(A)|s|$ with the graph of $h(x)$ for $0 \leqslant x \leqslant g_n(A)|s|$. Using (4.1) one immediately obtains the required formula for $\phi_{n+1}(s, t, A)$. This completes the inductive step and proves Theorem 1. □

Remark 1. It is clear from the above proof that $\phi_{n+1}(s, t, A) = \inf\{Eh(X): EX = t\}$ is attained by a unique random variable X, for each s and t. Specifically, $X \equiv t$ if $|t| \leqslant g_{n+1}(A)|s|$, while X takes the two values $g_{n+1}(A)|s|$ sgn t, $\gamma(g_n(A), A)^{1-q}\, g_{n+1}(A)|s|$ sgn t, if $g_{n+1}(A)|s| \leqslant |t| \leqslant |s|$. By working backwards, then, the unique martingale attaining the value $\phi_n(s, t, A)$ can always be constructed, see Example 1 below. Further, once the distribution of Y_1 is fixed, a unique martingale attaining equality in (1.1) with $C = C_{n,p}$ is determined.

Example 1. Let $p = 2$, so that $\phi(y, x) = [x(x + (1-y)^2)^{-1} - x)]^{1/2}$. A calculation shows that $A_3 = 16/25$. Hence, if X_1, X_2, X_3 is a martingale difference sequence, the following inequality is sharp.

(4.2) $$E\max[X_1^2, (X_1 + X_2)^2, (X_1 + X_2 + X_3)^2] \leqslant (25/16)\, E(X_1 + X_2 + X_3)^2.$$

The following martingale difference sequence attains equality. Let

$X_1 \equiv 1, P[X_2 = 1] = 3/8, P[X_2 = -3/5] = 5/8.$ Then

$P[X_3 = 4/3 | X_2 = 1] = 3/8, P[X_3 = -4/5 | X_2 = 1] = 5/8,$

$P[X_3 = 0 | X_2 = -3/5] = 1.$ □

Note that equality can be attained in (4.2) with an *arbitrary* distribution for X_1. Namely, multiply the difference sequence given above by any variable X independent of (X_1, X_2, X_3). However, once the distribution of X_1 is fixed, a unique martingale attaining equality in (4.2) is defined.

Proof of Theorem 2. From results of Section 2, it is clear that $A_n \to q^{-p}$. After all, $\lim_{n\to\infty} A_n \geqslant q^{-p}$ exists. Moreover, $\lim_{n\to\infty} A_n > q^{-p}$ is impossible because the equation $\phi(y,x) = y$ has no solution if $x > q^{-p}$

It follows from (3.5) that $\delta\phi/\delta x$ is continuous at the point (q^{-1}, q^{-p}), where it takes the value 1. Let $0 < \epsilon < 1/4$ be otherwise arbitrary and choose $\delta > 0$ such that $|y - q^{-1}| < \delta$, $|x - q^{-p}| < \delta \Rightarrow |\delta\phi/\delta y - 1| < \epsilon$. Choose n_0 so that $n \geqslant n_0 \Rightarrow A_n - q^{-p} < \delta$. Then, for $n \geqslant n_0, j = 0, \ldots, n$, one has $g_j(A_n) \leqslant g_j(q^{-p})$. Also, $g_j(A_n) \geqslant g_j(A_{n_0})$, for $j = 0, \ldots, n_0$. Since $g_j(q^{-p}) \downarrow q^{-1}$ as $j \to \infty$, the above two inequalities taken together imply that there exists n_1, independent of n, such that $|g_j(A_n) - q^{-1}| < \delta$, $j = 0, \ldots, n$, with the possible exception of n_1 values of j, i.e., all but finitely many members of the sequence $g_j(A_n)$, $j = 0, \ldots, n$, lie within δ of q^{-1} independently of n. Now fix $n \geqslant n_0$ and let $y_j = g_j(A_n)$. Thus,

$$(y_j - y_{j-1})/(\phi(y_{j-1}, A_n) - y_{j-1}) = 1, j = 1, \ldots n.$$

Now

$$\textstyle\int_{y_{j-1}}^{y_j} dy/(\phi(y, A_n) - y) = 1 + \rho_j, j = 1, \ldots n{-}1.$$

(where $j = n$ is excluded since the corresponding integral is not finite). One has $|\rho_j| \leqslant 1/2\, M_j(1 - M_j)^{-2}$, provided $M_j = \sup\{|\delta\phi/\delta y - 1| : y_j \leqslant y \leqslant y_{j-1}\} < 1$. Hence, all but n_1 of the $|\rho_j|$ are smaller than ϵ. Since ϵ is arbitrary it follows that

$$\lim_{n\to\infty} 1/n \textstyle\int_{g_{n-1}(A_n)}^{1} dy/(y - \phi(y, A_n)) = 1$$

As $n \to \infty$, $g_{n-1}(A_n) \to U$, where $U < q^{-1}$ is the solution of the equation $\gamma(U, q^{-p}) = 0$. Therefore,

(4.3) $$\lim_{n\to\infty} 1/n \textstyle\int_U^1 dy/(y - \phi(y, A_n)) = 1,$$

Next, examine the asymptotic behavior of the integral $I(x) = \int_U^1 dy/(y - \phi(y, x))$, as $x \downarrow q^{-p}$. Clearly, $I(x) \to \infty$ as $x \downarrow q^{-p}$. It is well-known that its asymptotic behavior is determined by the behavior of $y - \phi(y, x)$ near its minimum (as a function of y). For x close to q^{-p}, this minimum is attained at a value of y close to q^{-1}. Recalling the definition of $\theta = \theta(y, x)$, one has, for (y, x) close to (q^{-1}, q^{-p}),

$$\begin{aligned}(y - \phi(y, x))^{-1} &= 2\, q^{2-2p}(q-1)^{-1}\, \theta^{-2} + o(\theta^{-2}) \\ &= 8/(q p^3[(y - q^{-1})^2 + 2(x - q^{-p})(p-1)^{-2} q^{p-3}]^2)\end{aligned}$$

on expanding θ in a Taylor series about (q^{-1}, q^{-p}). It follows that

(4.4) $$\lim_{x \downarrow q^{-p}} (2\pi^2)^{-1/2} (q^{3p-1}(x - q^{-p})^3)^{1/2} I(x) = 1.$$

The conclusion of Theorem 2 follows from (4.3) and (4.4). □

REFERENCES

BURKHOLDER, D. L. (1979). A sharp inequality for martingale transforms. *Ann. Probability* 7 858–863.

COX, D. C. (1982). The best constant in Burkholder's weak-L^1 inequality for the martingale square function. *Proc. Amer. Math. Soc.* 85 427–433.

DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

DUBINS, L. E., and GILAT, D. (1978). On the distribution of maxima of martingales. *Proc. amer. Math. Soc.* 68 337–338.

KEMPERMAN, J. H. B. (1968). The general moment problem, a geometric approach. *Ann. Math. Statist.* 39 93–122.