Z. Wahrscheinlichkeitstheorie verw. Gebiete
36, 67–73 (1976)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Generalized Kolmogorov Inequalities for Martingales

D. Gilat* ** and W. D. Sudderth***

Department of Theoretical Statistics, University of Minnesota,
Minneapolis, Minnesota 55455, USA

The classical Čebyšev inequality leads to an inequality for martingales which is often called the Kolmogorov inequality. It is shown here that many generalized Čebyšev inequalities for random variables lead in a similar way to martingale inequalities, and that the corresponding martingale inequality is sharp when the Čebyšev inequality is.

1. The Main Result

Let R be the set of real numbers and let $\mathscr{B}$ be the collection of Borel subsets of R. As is customary, set $R^\infty = R \times R \times \cdots$ and let $\mathscr{B}^\infty$ be the product σ-field $\mathscr{B} \times \mathscr{B} \ldots$. Let $X_1, X_2, \ldots$ be the coordinate process on R^∞. It is convenient here to regard a martingale as being a probability measure P on $\mathscr{B}^\infty$ under which the process $\{X_n\}$ is a martingale in the usual sense. There is no loss of generality since every martingale on an abstract probability space has a distribution for which $\{X_n\}$ is a martingale.

Next let Φ be a set of Borel functions from R to R and suppose every member of Φ is either convex or concave. Let r be a mapping from Φ to R. Associate with Φ and r the class $C = C(\Phi, r)$ of all probability measures p on $\mathscr{B}$ such that, for every $\varphi \in \Phi$, φ is integrable with respect to p and $\int \varphi \, dp \leqq r(\varphi)$. Finally, let $M = M(C)$ be the collection of all probability measures P on $\mathscr{B}^\infty$ such that $\{X_n\}$ is a martingale under P and, such that, for every $n \geqq 1$, the distribution of X_n under P is in C.

Theorem 1. *If $B \in \mathscr{B}$, then*

$$\sup_{P \in M} P\{X_n \in B \text{ for some } n\} = \sup_{P \in M} P\{X_1 \in B\} = \sup_{p \in C} p(B). \tag{1}$$

---

* On leave from Tel-Aviv University. Presently at the University of California, Berkeley.

** To whom offprint requests should be sent

*** Research supported by National Science Foundation Grant MPS75-06173

Proof. The second equality in (1) is trivial and it is obvious that the right-hand-side of the first equality is no larger than the left. It remains to prove the reverse inequality.

For $x=(x_1, x_2, \ldots)\in R^\infty$, let

$$t(x)=\text{least } n \text{ (if any) such that } x_n\in B,$$
$$\quad =\infty \text{ if there is no such } n.$$

Then, for every P,

$$\begin{aligned} P\{X_n\in B \text{ for some } n\} &= P\{t<\infty\} \\ &= \lim_{n\to\infty} P\{t\leq n\} \\ &= \lim_{n\to\infty} P\{X_{t\wedge n}\in B\}. \end{aligned}$$

The proof will be finished once it is shown that, for $P\in M$, the distribution of $X_{t\wedge n}$ under P is in C and this latter fact is a consequence of the lemma below.

A *stopping variable* is a mapping s from R^∞ to the positive integers such that, for every n, the event $\{s\leq n\}$ is measurable with respect to $\{X_1, \ldots, X_n\}$.

Lemma 1. *If $P\in M$ and s is a bounded stopping variable, then the distribution of X_s under P is in C.*

Proof of Lemma 1. Let $\varphi\in\Phi$. Suppose first that φ is concave. Then $\{\varphi(X_n)\}$ is a supermartingale under P [3, V.T. 6, p. 79] and $\int\varphi(X_s)\,dP\leq\int\varphi(X_1)\,dP\leq r(\varphi)$, where the first inequality is by the optional sampling theorem [3, V.T 9, p. 80] and the second is by definition of M. Suppose next that φ is convex. Then $\{\varphi(X_n)\}$ is a submartingale under P [3, V.T 6, p. 79] and, if m is a positive integer such that $s\leq m$, then again by the optional sampling theorem and the definition of M, $\int\varphi(X_s)\,dP\leq\int\varphi(X_m)\,dP\leq r(\varphi)$. This completes the proof of the lemma and of Theorem 1.

Theorem 1 can be viewed as another reflection of the idea of *bold play* for gambling problems used in [1]. If a gambler, who seeks to attain a fortune in B, is allowed to select a martingale in the class M, then, according to Theorem 1, he can come as near as is possible to reaching his goal at the first stage of play.

It is easy to check that the supremum in (1) is equal to

$$\sup_{P\in M} P\{\text{for some } n \text{ and all } k\geq n,\ X_k\in B\}$$

and also equal to

$$\sup_{P\in M} P\{X_n\in B \text{ for infinitely many } n\}.$$

Roughly, the reason is that, after reaching B, the process can be stopped and thus remain in B from then on. This suggests the following generalization of Theorem 1.

Theorem 2. *If ψ is a nonnegative, Borel function on R, then*

$$\begin{aligned} \sup_{P\in M}\int\{\limsup_n \psi(X_n)\}\,dP &= \sup_{P\in M}\int\{\liminf_n \psi(X_n)\}\,dP \\ &= \sup_{P\in M}\int\psi(X_1)\,dP \qquad (2) \\ &= \sup_{p\in C}\int\psi\,dp. \end{aligned}$$

Since Theorem 2 is not used in the examples of the next section, its proof is omitted. A proof can be based on the preceeding ideas together with Fatou's lemma and [5, Theorem 1]. Like Theorem 1, Theorem 2 has an interpretation for gambling problems. In fact, if ψ is bounded and is regarded as being a utility function and if $\{X_n\}$ is viewed as the sequence of successive fortunes of a gambler, then $\int \{\limsup_n \psi(X_n)\}\, dP$ is the utility associated with the strategy P by Dubins and Savage [6, Theorem 3.2].

Remarks. 1. The results and their proofs remain valid if in the first paragraph of this section the real line R is replaced by any Borel subset S of R, $\mathscr{B}$ is taken as the collection of Borel subsets of S and $X_1, X_2, \ldots$ is the coordinate process on the product space S^∞. The processes considered are then those whose state-space is S. In particular when S is taken to be a compact interval, then, as suggested by the referee, Theorem 1 can be used to obtain sharp inequalities for uniformly bounded martingales.

2. If every function in Φ is convex (concave) increasing, then Theorems 1 and 2 remain true if, in the definition of M, "martingale" is replaced by "sub-(super)-martingale." The proofs are easily modified.

2. Some Applications

In this section Theorem 1 is applied to some known Čebyšev-type inequalities for real-valued random variables and the corresponding inequalities for martingales are obtained.

Examples 1 and 1′ present sharp upper bounds on the probability that an L_1-bounded martingale ever leaves an open interval. Examples 2 and 2′ provide analogous results for L_2-bounded martingales. The classical L_1 and L_2 forms of Kolmogorov's inequality are special cases of Examples 1′ and 2′ respectively. In addition, these Examples also lead to one-sided versions of the Kolmogorov inequality, some of which seem to be new. Finally, Example 3 considers the class of all martingales $\{X_n\}$ such that for each n, the Laplace transform of x_n is majorized by the Laplace transform of the standard normal distribution. Such martingales may conveniently be called *subnormal.* Example 3 provides an upper bound on the probability that a subnormal martingale will ever exit from an open interval. A symmetric as well as a one-sided version of the inequality is also derived for this case.

Example 1. Fix a positive number c and consider the class C of all probability measures p on $(R, \mathscr{B})$ with mean 0 and absolute first moment not exceeding c. Let b and a be positive numbers. Then according to an inequality due to Glasser [2, p. 481],

$$\sup_{p\in C} p\{(-\infty, -b] \cup [a, +\infty)\} = \min\left\{1, \frac{c}{2}\left(\frac{1}{b}+\frac{1}{a}\right)\right\}. \tag{3}$$

To adapt (3) to an application of Theorem 1, take $\Phi=\{\varphi_-, \varphi_+, \varphi_1\}$ where $\varphi_+(x)=x=-\varphi_-(x)$ and $\varphi_1(x)=|x|$; set $r(\varphi_-)=0=r(\varphi_+)$ and $r(\varphi_1)=c$, and let $B=(-\infty, -b]\cup[a, +\infty)$. Then apply Theorem 1 to (3) to obtain

$$\sup_{P\in M} P\{X_n\leqq -b \text{ or } X_n\geqq a \text{ for some } n\}=\min\left\{1, \frac{c}{2}\left(\frac{1}{b}+\frac{1}{a}\right)\right\}, \tag{4}$$

where $M=M(C)$ is the collection of all probability measures on $(R^\infty, \mathscr{B}^\infty)$ under which the coordinate process $\{X_n\}$ is a martingale with mean 0 and L_1-norm bounded by c.

Thus if $\{X_n, n\geqq 1\}$ is any L_1-bounded martingale with mean zero, then

$$\text{Prob}\{X_n\leqq -b \text{ or } X_n\geqq a \text{ for some } n\}\leqq\frac{1}{2}\left(\frac{1}{b}+\frac{1}{a}\right)\sup E|X_n|. \tag{5}$$

Letting b tend to $+\infty$ in (5) yields a one-sided version

$$\text{Prob}\{\sup X_n\geqq a\}\leqq\frac{1}{2a}\sup E|X_n|, \tag{6}$$

which holds for martingales with mean zero.

Example 1'. This example differs from the preceeding one only in that the assumption of mean zero made there is omitted. Formally, let $c>0$ and take C to be the class of all p with absolute first moment no larger than c. Let b and a be positive numbers. The role played by (3) in Example 1 is here played by

$$\sup_{p\in C} p\{(-\infty, -b]\cup[a, +\infty)\}=\min\left\{1, \frac{c}{m}\right\} \tag{7}$$

where $m=\min\{a, b\}$. Since the indicator function of $\{x: x\leqq -b \text{ or } x\geqq a\}$ is dominated by $m^{-1}|x|$, it follows that the right-hand-side of (7) majorizes the left. In fact, the supremum in (7) is attained by a p with support $\{-b, 0\}$ if $b=m$ and $\{0, a\}$ if $a=m$. (We were led to (7) by the general result of Karlin and Studden [2, Theorem 2.1, p. 472].) Apply Theorem 1 with $\Phi=\{\varphi\}$ where $\varphi(x)=|x|$ and with $r(\varphi)=c$ to get

$$\text{Prob}\{X_n\leqq -b \text{ or } X_n\geqq a \text{ for some } n\}\leqq\frac{1}{m}\sup E|X_n|. \tag{8}$$

Set $a=b$ in (8) to obtain the classical Kolmogorov inequality. Let b approach $+\infty$ in (8) to obtain the well-known inequality

$$\text{Prob}\{\sup X_n\geqq a\}\leqq\frac{1}{a}\sup E|X_n|. \tag{9}$$

In contrast to (5) and (6), which require the martingale $\{X_n\}$ to have mean zero, (8) and (9) hold for all L_1-bounded martingales.

Example 2. Replace the condition on the first absolute moment of p in Example 1 by a similar condition on the second moment. That is, consider the class C of all p with mean 0 and variance at most c. Again, let b and a be fixed positive numbers and set $m=\min\{b, a\}$. By Selberg's inequality [2, p. 475],

$$\sup_{p\in C} p\{(-\infty, -b]\cup[a, +\infty)\}=1 \text{ if } ab\leqq c$$

$$=\frac{(b-a)^2+4c}{(b+a)^2} \qquad ab-m^2\leqq 2c<2ab \tag{10}$$

$$=\frac{c}{m^2+c} \qquad 2c<ab-m^2.$$

Abbreviate the right-hand-side of (10) by $U(a,b,c)$. Then Theorem 1 with the obvious Φ, r and B, yields

$$\sup_{P\in M} P\{X_n\leqq -b \text{ or } X_n\geqq a \text{ for some } n\}=U(a,b,c), \tag{11}$$

where $M=M(C)$. Thus, for any L_2-bounded mean zero martingale,

$$\text{Prob}\,\{X_n\leqq -b \text{ or } X_n\geqq a \text{ for some } n\}\leqq U(a,b,\sup EX_n^2). \tag{12}$$

Bert Fristedt helped us to discover (12).

As $b\to+\infty$, U reduces to $U(a,+\infty,c)=U(a,c)=\frac{c}{a^2+c}$. Hence, for martingales $\{X_n\}$ with mean zero,

$$\text{Prob}\,\{\sup X_n\geqq a\}\leqq\frac{\sup EX_n^2}{a^2+\sup EX_n^2}. \tag{13}$$

Example 2'. Here, as in Example 1', the mean zero assumption is dropped to obtain sharp bounds for all L_2-bounded martingales. The notation here is the same as in Example 1' except that C is the set of p with second moment no larger than c and $\Phi=\{\varphi\}$ where $\varphi(x)=x^2$. The Čebyšev result corresponding to (7) is

$$\sup_{p\in C} p\{(-\infty, -b]\cup[a, +\infty)\}=\min\left\{1, \frac{c}{m^2}\right\}. \tag{14}$$

The proof of (14) is similar to that of (7). By Theorem 1,

$$\text{Prob}\,\{X_n\leqq -b \text{ or } X_n\geqq a \text{ for some } n\}\leqq\frac{1}{m^2}\sup EX_n^2 \tag{15}$$

for every L_2-bounded martingale $\{X_n\}$. Two well-known results can be obtained from (15) by taking $a=b$ for one and letting b tend to $+\infty$ for the other.

Example 3. Suppose p is a probability measure on $(R,\mathscr{B})$ whose Laplace transform $\lambda_p(c)=\int e^{cx}dp(x)$ is finite for all real c. Bernstein's inequality

$$p\{(-\infty, -b]\cup[a, +\infty)\}\leqq\inf_{c>0} e^{-cb}\lambda_p(-c)+\inf_{c>0} e^{-ca}\lambda_p(c) \tag{16}$$

is not hard to verify, or else its proof can be found, for example in [4, p. 86].

Consider now the class $C=C_\lambda$ of all p for which λ_p is majorized by some fixed function λ. Then when λ_p is replaced by λ on the right-hand-side of (16), an upper bound for the entire class C_λ is obtained. Since for each c, the function $\varphi_c\colon x\to e^{cx}$ is convex, Theorem 1, with $\Phi=\{\varphi_c\colon c\neq 0\}$ and $r\colon \varphi_c\to\lambda(c)$, applies to obtain the corresponding martingale inequality for the class $M_\lambda = M(C_\lambda)$. An interesting upper bound (17) is thus obtained from (16) by taking $\lambda(c) = e^{c^2/2} =$ the Laplace transform of the standard normal distribution.

$$\sup_{P \in M_\lambda} P\{X_n \leqq -b \text{ or } X_n \geqq a \text{ for some } n\} \leqq e^{-\frac{b^2}{2}} + e^{-\frac{a^2}{2}}. \tag{17}$$

The symmetric form of (17) as well as its one-sided version can be obtained as in the previous examples.

3. A Gambling Problem

As already hinted earlier, we were motivated to conjecture the various bounds of Section 2 by the idea of *bold-play* for subfair gambling problems.

To establish, for example, the bound in (13), one might proceed as follows. Consider a gambler with positive fortune x and total variance-allowance of $c > 0$, who wishes to attain the fortune 0 by means of a sequence of fair bets whose total variance does not exceed c. How should he play so as to maximize his chance of attaining the goal and what is then that chance as a function of x and c? As soon as one conjectures that it is optimal for the gambler to use up his entire variance on a single fair bet, one is led to believe, from the one-sided Čebyšev inequality ((10) with $b = 0$ and $a \to \infty$), that $Q(x, c) = c(x^2 + c)^{-1}$ is the maximal probability of attaining the goal. One then may try to verify this conjecture, using Theorem 2.12.1 of [1], by proving that the function

$$Q(x,c) = \begin{cases} \dfrac{c}{x^2 + c}, & x > 0 \\ 1, & x \leqq 0 \end{cases}$$

is *excessive* at every (x, c), for every gamble with mean x and variance at most c; that is, by proving

$$EQ(x + Z, c - EZ^2) \leqq Q(x, c) \tag{18}$$

for all x, all $c > 0$ and all random variables Z, with $EZ = 0$ and $EZ^2 \leqq c$.

Initially this approach was attempted, but we were unable to find a direct proof of (18). On the other hand, since Example 2 shows that Q is indeed the optimal utility (the "U of the house" in the terminology of [1]) for this gambling problem, it follows from Theorem 2.14.1 of [1] (which is a general version of the so called "optimality principle" of dynamic programming), that Q is in fact excessive and thus (18) holds. This idea can be used to generate some Čebyšev-type inequalities from the Kolmogorov-type examples of Section 2.

Example 4. Specialize (18) to $x = 1$, $c = 2$ and $Z \geqq -1$ with $EZ = 0$ and $EZ^2 = 1$, to obtain

$$E \frac{1}{1 + (1 + Z)^2} \leqq \frac{2}{3}. \tag{19}$$

Set $1+Z=X$, then X is an arbitrary non-negative random variable with mean 1 and variance 1. Thus any such random variable X satisfies

$$E\frac{1}{1+X^2} \leqq \frac{2}{3}. \tag{20}$$

We do not know how to prove (20) directly. Trying to maximize the left hand side of (20) over two-point distributions for X, to which the problem can be reduced by some general principles, leads to a polynomial of degree 6 which we were unable to reduce. Alternatively, fitting a parabola on top of the graph of $0 \leqq x \to (1+x^2)^{-1}$, seems likewise unmanageable. Looking at some computer data for the problem, Burt Fristedt has correctly guessed that the bound 2/3 in (20) is attained by the two-point distribution

$$P\left[X=\frac{3-\sqrt{5}}{2}\right]=\frac{1}{2}\left(1+\frac{1}{\sqrt{5}}\right), \quad P\left[X=\frac{3+\sqrt{5}}{2}\right]=\frac{1}{2}\left(1-\frac{1}{\sqrt{5}}\right).$$

It is perhaps curious to notice that the distances of the values of this X from its mean, are the Fibbonaci numbers $(\sqrt{5}\pm 1)/2$.

References

1. Dubins, L.E., Savage, L.J.: How to gamble If You Must. New York: Dover 1965
2. Karlin, S.J., Studden, W.J.: Čebyšeff systems: with applications in analysis and statistics. New York: Interscience 1966
3. Meyer, P.A.: Probability and potentials. Waltham, Mass.: Blaisdell 1966
4. Renyi, A.: Foundations of probability. San Francisco: Holden-Day 1970
5. Sudderth, W.D.: A "Fatou equation" for randomly stopped variables. Ann. Math. Statist. **42**, 2143–2146 (1971)
6. Sudderth, W.: On measurable gambling problems. Ann. Math. Statist. **42**, 260–269 (1971)

Received December 30, 1975