A LARGE DEVIATION INEQUALITY FOR CONTINUOUS-TIME MARTINGALES WITH APPLICATIONS (U) MARYLAND UNIV COLLEGE PARK DEPT OF MATHEMATICS E V SLUD 18 JAN 87

UNCLASSIFIED ND87-20-EVS N00014-86-K-0007 F/G 12/1 NL
This paper first presents a discrete-time martingale exponential bound due to W. Steiger (1969) and further developed by D. Freedman (1975), and then extends it straightforwardly to a large class of continuous-time (local) martingales. The resulting inequality yields many known estimates, and some new ones, on the growth and fluctuations of processes which can be expressed as stochastic integrals.
A LARGE DEVIATION INEQUALITY FOR CONTINUOUS-TIME
MARTINGALES, WITH APPLICATIONS

Eric V. Slud

University of Maryland
Department of Mathematics
College Park, Maryland
20742

TR87-20
MD87-20-EVS

January 10, 1987
A LARGE DEVIATION INEQUALITY FOR CONTINUOUS-TIME MARTINGALES, WITH APPLICATIONS

Eric V. Slud, Mathematics Department
University of Maryland, College Park, MD 20742

Abstract. This paper first presents a discrete-time martingale exponential bound due to W. Steiger (1969) and further developed by D. Freedman (1975), and then extends it straightforwardly to a large class of continuous-time (local) martingales. The resulting inequality yields many known estimates, and some new ones, on the growth and fluctuations of processes which can be expressed as stochastic integrals.

AMS 1980 subject classifications: Primary 60G44, Secondary 60E15.

Key words and phrases: exponential bounds, Doob-Meyer compensator, calculable variance-process, maximal inequalities, stochastic integral representation.

Research supported by the Office of Naval Research under contract N00014-86-K-0007.

1. INTRODUCTION

Some of the most important and useful tools of martingale theory are the inequalities bounding tail-probabilities for the supremum of a (sub-)martingale \( X(\cdot) \) over a time-interval \([0,T]\) in terms of expectations involving \( X(\cdot) \) and related stochastic processes evaluated at the single time-endpoint \( T \). The most famous inequality of this type is Doob's (1953, Theorem 3.4); another, less less well-known but more general, is due to Lenglart (1977); see Burkholder (1973) for other "distribution function inequalities" of this type. The submartingale
maximal inequality of Birnbaum and Marshall (1961, Theorem 5.1) is a closely related result, which Slud (1986) has recently generalized and shown to follow from Lenglart's result. The present paper develops an exponential-bound result for continuous-time martingales, which in many examples is much more informative than the previously mentioned inequalities but has the disadvantage that it applies only to martingales, not submartingales.

2. EXPONENTIAL BOUNDS FOR MARTINGALES

The present Section develops an exponential inequality generalizing to continuous-time martingales Kolmogorov's famous upper exponential bound [Loeve, 1955, pp. 254-5] for sums of uniformly bounded independent summands. The inequality is due [in the discrete-time martingale setting] to Steiger (1969) and was re-proved and exploited by Freedman (1975, pp. 102-4). The following restatement of Freedman's version is given here without proof.

**Proposition.** Let \( M(\cdot) \) be a \( \{F_t\} \)-adapted martingale on the probability space \((\Omega,F,P)\) with parameter-set \([0,\infty)\), and assume that \( M(\cdot) \) is a.s. in \( D[0,t] \) as a random function for each \( t < \infty \). Also let \( \{t_i: 0 \leq i \leq L\} \) be a nondecreasing sequence of stopping times with \( t_0 = 0 \), such that for each \( i=1,2,\ldots,L \) and a finite constant \( K \), \( |M(t_i) - M(t_{i-1})| \leq K \) a.s. Then for all \( \alpha \) and \( \beta > 0 \),

\[
P\left( M(t_i) \geq \alpha \text{ and } \sum_{j=1}^{i} E\left\{ \frac{M^2(t_j) - M^2(t_{j-1})}{F_{t_{j-1}}} \right\} \leq \beta \text{ for some } i=1,\ldots,L \right) \leq \frac{\beta e^{\alpha/K}}{(K\alpha + \beta)^2} \leq \exp\left[-\frac{\alpha^2}{2(K\alpha + \beta)}\right].
\]
In stating the foregoing Proposition, we used the idea of conditioning on past history $F_\tau$ up to a stopping time $\tau$. The definition is

$$F_\tau = \sigma(A : \text{for } t \in [0,\infty), A \cap [\tau \leq t] \in F_t).$$

See Liptser and Shiryaev (1977, vol. 1, pp. 25-29) for further background concerning $\sigma$-fields $F_\tau$. The importance for us of partitioning the interval $[0,T)$ by means of increasing stopping-time sequences $\{t_i\}$ is that the uniform bounds on $|M(t_i) - M(t_{i-1})|$ are not very restrictive when the times $t_i$ are allowed to be random.

We next restrict the continuous-time martingales $M(\cdot)$ under consideration to have calculable variance-processes (cf. Brown 1978; Helland 1982) in the following strong sense:

we assume for any nested increasing sequence of partitions $Q(k) = \{t_{jk}\}_j$ of $[0,\infty)$ consisting of a.s. nondecreasing sequences of stopping times $t_{jk}$ satisfying ($t_{0k} = 0$ and)

(i) $E M_{jk}^2 < \infty$ for each $j$ and $k$,

(ii) $\max\{j : t_{jk} \leq t\} < \infty$ a.s. for each $k$ and each $t < \infty$,

and

(iii) $\text{mesh}(Q(k)) = \max_j (t_{j+1,k} - t_{jk}) \xrightarrow{P} 0$ as $k \to \infty$,

that for each $t$ the $L^1$-limit

$$\mathcal{V}(t) = \lim_{k \to \infty} \mathcal{V}_k(t) = \lim_{k \to \infty} \sum_{j:t_{jk} \leq t} E\left(\frac{(M(t_{j+1,k}) - M(t_{jk}))^2}{F_{t_{jk}}}ight)$$

exists. When we discuss local martingales $M(\cdot)$, we implicitly restrict attention to $\mathcal{D}[0,\infty)$ processes for which some increasing
sequences \( \{ \tau_n \} \) of stopping-times yield martingales \( M(\cdot, \tau_n) \) with calculable variance-processes.

The special class of martingales which have calculable variance-processes according to the foregoing definition is well known (Brown 1978; Jacod 1979; Slud 1987) to include all continuous-path martingales and martingales whose squares are "quasi-left-continuous" (i.e., have a.s. continuous Doob-Meyer compensators); all martingales adapted to a \( \sigma \)-field family \( F_t \equiv F_0 \vee \sigma( N(s) : 0 < s \leq t ) \) where \( N(\cdot) \) is a simple multivariate counting-process; and all (finite sums of) stochastic integrals of predictable processes of the preceding types. Therefore, although not all square-integrable martingales have calculable variance-processes (see the discussion of Helland 1982), the class of processes which do seems to be quite large enough for most applications.

An important feature of the inequality (2.1) is that the upper bound does not depend on \( L \). Therefore a limit can be taken over a sequence of partitions \( Q(k) \) satisfying (i)-(iii) as above. The result obtained in this way seems to be one of the first in which the concept of calculable variance plays a crucial role.

**Theorem 2.1.** Let \( M(\cdot) \) be a \( \{ F_t \} \)-adapted locally square-integrable martingale in \( D[0,T) \) with calculable variance-process, and let \( \tau \) be a stopping-time \( < T \) a.s. Then for the calculable variance-process \( V(\cdot) \) of \( M(\cdot) \) and any \( \alpha, \beta > 0 \),

\[
P\left( M(t) \geq \alpha \quad \text{and} \quad V(t) \leq \beta \quad \text{for some} \quad t \in [0,\tau] \right) \leq e^{-\frac{1}{2} \alpha^2 / (K\alpha + \beta)} \quad (2.2)
\]

where \( K \equiv \text{ess. sup. sup} \{ |\Delta M(t)| : 0 \leq t \leq \tau \} \).
Proof. For arbitrarily small \( \delta > 0 \), define the increasing sequence \( \{ \sigma_n^\delta \} \) of \( \{ F_t \} \) stopping-times by \( \sigma_0^\delta \equiv 0 \) and

\[
\sigma_n \equiv \sigma_n^\delta \equiv \tau \wedge \min \{ t > \sigma_{n-1}^\delta : |M(t) - M(\sigma_{n-1}^\delta)| \geq \delta \}.
\]

Right-continuity of \( M(\cdot) \) implies that such a sequence exists and increases a.s. to \( \tau \), and that

\[
\sup \{ |M(t) - M(\sigma_{n-1}^\delta)| : \sigma_{n-1}^\delta \leq t < \sigma_n^\delta \} \leq \delta \quad \text{a.s.} \quad (2.3)
\]

Now let \( Q(k) = \{ t_{jk} \} \) be nested increasing random partitions of \([0,T)\) by stopping times, such that

for each \( k \geq 1 \), \( \{ \sigma_n^{k-1} \} \subset Q(k) \) and (i)-(iii) hold \( (2.4) \)

Then by construction, for every \( k \geq 1 \),

\[
\max_j |M(t_{j+1,k}) - M(t_{jk})| \leq K + k^{-1} \quad \text{a.s.}
\]

Impose for this paragraph the auxiliary assumption that \( M(\cdot) \) itself is a square-integrable martingale. It is not hard to show from the calculable-variance property of \( M(\cdot) \) that

\[
\sup_j |\sum \left\{ E \left( M^2(s,t_{j+1,k}) - M^2(s,t_{jk}) | F_{t_{jk}} \right) \right\} - V(s)| \xrightarrow{P} 0 \quad (2.5)
\]

[To see this, observe first that the processes \( V_m(\cdot) - V_k(\cdot) \) for \( m \geq k \) are each \( \{ F_t \} \) martingales for the time-index \( s \) in \( Q(k) = \{ t_{ik} \} U \{ t \} \), so that by Doob's inequality, for each \( t \) and each \( c > 0 \)

\[
P \left( \max_i |V_m(t_{ik}) - V_k(t_{ik})| \geq c \right) \leq c^{-1} E |V_m(t) - V_k(t)|
\]
which converges to 0 as \( k, m \) go to \( \infty \), by the calculable-variance assumption. Since it is easy to check from (2.3) and (2.4) that 

\[ \sup \{ V_m(t) - V_m(t_{ik}); \ t_{ik} \leq t < t_{i+1,k}, \ i \geq 0 \} \leq k^{-2}, \]

which converges in probability to 0 as \( k \to \infty \), the assertion (2.5) follows from the a.s. monotonicity of \( V(\cdot) \), the property (iii), and the fact that a.s. \( V_k(s) \geq V_k(t_{jk}) \) whenever \( s \geq t_{jk} \).

The Proposition of this Section applied to the martingale \( M(t_{rn}) \) with calculable variance-process yields for each fixed integer \( n \) and each constant \( \gamma > 0 \),

\[
P\left\{ \text{for some } i \geq 1, \ M(t_{ik}^n) \geq \alpha - \gamma \text{ and} \right. \\
\left. \sum_{j=1}^{i} E\left\{ [M(t_{j+1,k}) - M(t_{jk})]^2 \right| F_{t_{jk}} \right\} \leq \beta + \gamma \right\} \\
\leq \exp\left\{ -\frac{1}{2} (\alpha - \gamma)^2 / [(K + \gamma)(\alpha - \gamma) + \beta + \gamma] \right\} 
\]

But (2.5) together with right-continuity of \( M(\cdot) \) implies for each \( n \),

\[
P\left\{ \text{for some } t \text{ with } V(t^n) \leq \beta \right\} \setminus \left\{ \text{for some } i \right. \\
M(t_{ik}^n) \geq \alpha - \gamma \text{ and } \sum_{j=0}^{i} E\left\{ M^2(t_{j+1,k}) - M^2(t_{jk}) \right| F_{t_{jk}} \right\} \leq \beta + \gamma \right\} \\
\rightarrow P 0 \text{ as } k \to \infty. 
\]

Combining (2.6) and (2.7) and letting \( k \to \infty \) gives

\[
P\left\{ \text{for some } t \text{ with } V(t^n) \leq \beta \right\} \\
\leq \exp\left\{ -\frac{1}{2} (\alpha - \gamma)^2 / [(K + \gamma)(\alpha - \gamma) + \beta + \gamma] \right\}.
\]
Finally, let $n \to \infty$ and $\gamma \to 0$ to complete the proof of (2.1). □

As a first illustration of Theorem 2.1, consider the case of Wiener process $M(\cdot) = W(\cdot)$ on $[0,T]$, where $T < \infty$. The variance process $V(\cdot)$ for $W(\cdot)$, or equivalently the compensator for $W^2(\cdot)$, is simply $V(t) = t$; and of course, continuity of $W(\cdot)$ implies that the number $K$ in Theorem 2.1 is 0. Therefore Theorem 2.1 says in this context that

$$
P \left( \sup_{0 \leq t \leq T} W(t) \geq \alpha \right) \leq e^{-\frac{1}{2} \alpha^2 / T} \quad (2.8)
$$

Of course, more exact information exists about the probability distribution of $\sup_{t \in [0,T]} W(t)$ [see Feller, 1971, vol. 2, pp. 340-1, or Karlin and Taylor, 1975, pp. 345-7, where it is shown that the left-hand side of (2.8) is exactly equal to $2 \cdot \{1 - \Phi(\alpha / T^{1/4})\}$], but the Theorem gets the correct order of magnitude for the logarithm of the tail-probability for large $\alpha$.

Thus Theorem 2.1 can be thought of as a generalization of the known distributional bound (2.8) for the supremum of a Wiener process, controlling the supremum of a general local martingale in terms of the intrinsic time-scale given by its variance-process. Let us consider as a second application of the Theorem the case $M(\cdot) = N(\cdot) - A(\cdot)$, where $N$ is a simple counting-process on $[0,\infty)$, and $A(\cdot)$ is its compensator with respect to a $\sigma$-field family $F_t \equiv F_0 \vee \sigma(N(u): u \leq t)$. Then the variance-process $V(\cdot)$ for $N(\cdot)$ [the compensator for $M^2(\cdot)$] is described by Liptser and Shiryaev (1977, vol. 2, Theorem 18.2) and has the property $V(\cdot) \leq A(\cdot)$ a.s., with a.s. equality in case all the
conditional distributions of times $T_{n+1} - T_n$ between successive jumps given $F_n$ are nonatomic a.s. Now the quantity $K$ in the Theorem is 1. Take $\alpha = c\beta$, and apply Theorem 2.1 to conclude

$$\Pr\left\{ |N(t) - A(t)| \geq c\beta \text{ for some } t \text{ with } V(t) \leq \beta \right\} \leq 2e^{-\frac{1}{2} c\beta/(c+1)}.$$

3. APPLICATIONS TO SOLUTIONS OF STOCHASTIC EQUATIONS

We apply Theorem 2.1 next in a statistical setting: consider a finite population of $n$ individuals, each member of which comes equipped with a latent random survival-time $X_i$ and with a left-continuous $\{0, 1\}$-valued process $r_i(\cdot)$ on $[0, \infty)$ which indicates at time $t$ whether the death of individual $i$ at time $t$ would be observable. Let $N_i(t) = I[X_i \leq t]r_i(t)$ indicate whether the death of $i$ is observed by time $t$; define $F_t = \sigma(\{N_i(s), r_i(s) : 0 \leq s \leq t, i=1, \ldots, n\})$, and assume that

for each $i$, $N_i(t) - \int_0^t r_i(s) \, dH(s)$ is a $\{F_t\}$-martingale \quad (3.1)

where $H(\cdot)$ is some nonrandom continuous nondecreasing function on $[0, \infty)$, not depending on $i$, such that $H(0) = 0$ and $H(\infty) = \infty$. The statistical purpose of observing $N_i$ and $r_i$ is to estimate the distribution function $F(\cdot) \equiv 1 - \exp\{-H(\cdot)\}$ uniquely associated with the cumulative hazard function $H(\cdot)$, that is, to produce a $\{F_t\}$-adapted functional $\hat{F}(t)$ of $\{N_i(\cdot), r_i(\cdot)\}_i$ which is close to $F(t)$ (uniformly
for all t, if possible) when n is large and no other assumption than (3.1) is made. It is well known that the product-limit or Kaplan-Meier estimator \( \hat{F} \), which can be defined through the stochastic equation (Gill 1983)

\[
Z_n(t) = \frac{\hat{F}(t)-F(t)}{1-F(t)} = \int_0^t (1-Z_n(u^-)) \frac{dN(u^-)-r(u)H(u)}{r(u)}
\]

(3.2)

where

\[
N(t) = \sum_{i=1}^n N_i(t) \quad \text{and} \quad r_n(t) = \sum_{i=1}^n r_i(t)
\]

has excellent properties in this regard. Note that while the unique locally bounded solution \( Z_n(\cdot) \) of (3.2) does depend on \( F \), it is easy to check that

\[
1 - \hat{F}(t) = \frac{1 - \Delta N(s)}{r(s)} \quad \text{for all} \quad s \leq t, \Delta N(s) > 0
\]

does not. We do not motivate the estimator \( \hat{F} \) here apart from the remark that it coincides with the usual empirical distribution function in the special case when \( r_i(t) = 1_{[X_i \geq t]} \) for all \( i \) (i.e., in the case where all the \( X_i \) can be observed). Our purpose is to show that exponential (in n) bounds on tail-probabilities for \( \sup|Z_n(t)| : 0 \leq t \leq T \) can be simply derived via Theorem 2.1.

It is easy to check that the martingales (3.1) and therefore also the martingales (3.2) have calculable variance-processes. By standard theorems on stochastic integrals, on the event \( [r_n(s) > 0 \text{ for all } s \leq t] \)

\[
\langle Z_n \rangle(t) = \int_0^t \left[1-Z_n(u^-)\right]^2 r_n(u)^{-1} H(u) \quad .
\]

Noting that \( H(u) \leq C \) implies \( F(u) \leq 1-e^{-C} \) and \( |1-Z_n(u^-)| \leq e^C \), we
find for the stopping-time $\sigma_n = \sup\{ t : H(t) < C \} \text{ and } \gamma(t) \geq n/C \}$
defined in terms of an arbitrary but fixed constant $C > 0$, that a.s.

$$\langle Z_n \rangle (\sigma_n) \leq C^{-2} e^{-2C} n^{-1} \beta_n \equiv \beta \quad \text{and} \quad \sup_{0 \leq t \leq \sigma_n} |\Delta Z_n(t)| \leq C e^{-C n^{-1}} \xi_n \equiv K_n \equiv K.$$  

By Theorem 2.1,

$$P\left( \sup_{0 \leq t \leq \sigma_n} |Z_n(t)| \geq x \right) \leq 2 \exp \left\{ -\frac{1}{2} D n x^2 / (1 + x) \right\} \quad (3.3)$$

where $D \equiv C^{-2} e^{-2C}$. In the special case where $r_i(\cdot) \equiv 1$ for all $i$, the result (3.3) gives an upper bound related to bounds of Hoeffding (1963); in case $r_i(t) \equiv I[\min(X_i, Y_i) \geq t]$, where the random variables $\{Y_i\}$ form an i.i.d. sequence independent of the i.i.d. sequence $\{X_i\}$, the result (3.3) yields exponential bounds derived by Foldes and Rejtő (1981) and by Csorgő and Horváth (1983). See Slud (1987) for further discussion of the bearing of (3.3) on Estimation in Survival Analysis, as well as of applications of the Theorem in bounding probabilities of large deviations for compound-renewal processes.

Another arena of possible application of Theorem 2.1 is the study of hitting- and occupation-times for the solutions of stochastic differential equations. Such applications apparently depend heavily on detailed estimates for solutions of associated parabolic partial differential equations. For illustration, we sketch here an application to large-deviation estimates for hitting times of d-vector Wiener process $W(\cdot)$. Let $D$ denote a closed domain in $\mathbb{R}^d \times [0, \infty)$, containing $(0,0)$ and with a smooth boundary. Define for each
For any piecewise-smooth function \( f(x,t) \), Ito's Lemma says for
\[
M(t) = f(W(t),t) - f(0,0) - \int_0^t \{ \frac{\partial f}{\partial t}(W(s),s) + \frac{1}{2} \Delta f(W(s),s) \} \, ds \tag{3.4}
\]
that \( M(\cdot \wedge r) \) is a martingale with respect to the \( \sigma \)-field family \( F_t^W \) generated by \((W(s): 0 \leq s \leq t)\), where \( \Delta \) denotes the Laplacian on \( \mathbb{R}^d \). Moreover, the variance-process \( \langle M \rangle \) is calculable since \( f(W(\cdot),\cdot) \) is continuous, and (if we use \( \nabla \) to denote gradient in \( x \)-variables)
\[
\langle M \rangle(t) = \int_0^t \| \nabla f(W(s),s) \|^2 \, ds
\]
The particular choice of function \( f(x,t) \equiv \mathbb{E}^{(x,t)} \{ \tau_{(x,t)} \wedge (T-t) \} \) for \( 0 \leq t \leq T \), where \( T > 0 \) is fixed, is readily seen to solve the Partial Differential Equation
\[
\frac{\partial f}{\partial t} + \frac{1}{2} \Delta f = -1 \quad \text{on } D
\]
\[
f = 0 \quad \text{on } \partial D \cup \{ (x,T) : x \in \mathbb{R}^d \}
\]
Here we have adopted the standard notation \( \mathbb{E}^{(x,t)} \) to indicate that expectations are taken conditionally given \( W(t)=x \). Now, if we let
\[
L(t) \equiv \int_0^t I \{ (W(s),s) \in D \} \, ds
\]
denote the total occupation-time for \( D \) by \( W(\cdot) \) up to \( t \), then Theorem 2.1 applied to the martingale \( M(\cdot) \) up to stopping-time \( \tau(0,0) \) says
\[ P\{ f(W(t),t) - f(0,0) + L(t) \geq \alpha \text{ for some } t \leq r^T \text{ satisfying } \]
\[ \int_0^t \| \nabla f(W(s),s) \|^2 ds \leq \beta \} \leq e^{-\frac{1}{2} \alpha^2 / \beta} \tag{3.5} \]

Effective application of (3.5) would require a good bound on \( \| \nabla f \| \). For instance, if we could show directly or via a comparison method that \( \| \nabla f(x,t) \|^2 \leq C \) for all \((x,t) \in D\) and for all \( T \), then (3.5) implies that

\[ P(0,0)\{ L(\tau) \geq \alpha + E(0,0)\tau \text{ and } \tau^T \leq \beta/C \} \leq e^{-\frac{1}{2} \alpha^2 / \beta} \]

By letting \( T \) increase to \( \infty \), we would then conclude that

\[ P(0,0)\{ L(\tau) \geq \alpha + E(0,0)\tau \text{ and } \tau \leq \beta/C \} \leq e^{-\frac{1}{2} \alpha^2 / \beta} . \]

REFERENCES.


END
3-87
DTIC