The Annals of Probability
1973, Vol. 1, No. 6, 910-925

ANOTHER NOTE ON THE BOREL-CANTELLI LEMMA AND THE STRONG LAW, WITH THE POISSON APPROXIMATION AS A BY-PRODUCT[1]

BY DAVID FREEDMAN

University of California, Berkeley

Here is another way to prove Lévy's conditional form of the Borel-Cantelli lemmas, and his strong law. Consider a sequence of dependent variables, each bounded between 0 and 1. Then the sum S of the variables tends to be close to the sum T of the conditional expectations. Indeed, the chance that S is above one level and T is below another is exponentially small. So is the chance that S is below one level and T is above another.

The inequalities also show that for a sequence of dependent events, such that each has uniformly small conditional probability given the past, and the sum of the conditional probabilities is nearly constant at a, the number of events which occur is ncarly Poisson with parameter a.

1. Introduction. There is a theorem in [2] which includes both the Borel-Cantelli Lemma and the Strong Law. I recently came across two inequalities which do a similar job, and I would like to present them here. Let $(\Omega, \mathscr{F}, P)$ be a probability triple, and let $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots$ be increasing sub σ-fields of $\mathscr{F}$. Let $X_1, X_2, \cdots$ be random variables, such that

$$0 \leqq X_n \leqq 1 \tag{1}$$

$$X_n \quad \text{is} \quad \mathscr{F}_n\text{-measurable.} \tag{2}$$

Let

$$M_n = E(X_n \mid \mathscr{F}_{n-1}) , \tag{3}$$

so M_n is an $\mathscr{F}_{n-1}$-measurable function, with $0 \leqq M_n \leqq 1$. Say τ is a *stopping time* iff τ is a function on Ω taking the values $0, 1, 2, \cdots, \infty$ with $\{\tau = n\} \in \mathscr{F}_n$ for all $n = 0, 1, 2, \cdots$. This allows $P\{\tau = \infty\} > 0$. Usually, one starts with X_n satisfying (1), and defines $\mathscr{F}_n$ as the σ-field generated by $X_1, \cdots, X_n$. Allowing more general $\mathscr{F}_n$ is sometimes helpful.

The main result of this note is a pair of inequalities embodying the fact that $\sum X_i$ is around the same size as $\sum M_i$; large deviations have exponentially small probability. More exactly,

---

Received August 20, 1972; revised April 5, 1973.

[1] Research sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, USAF, under Grant AFOSR-71-2100A. The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation hereon.

AMS 1970 *subject classifications.* 60F05, 60F10, 60F15, 60G40, 60G45.

Key words and phrases. Borel-Cantelli lemmas, strong law, Poisson approximation.

(4) THEOREM. *Suppose* $X_1, X_2, \cdots$ *satisfy* (1) *and* (2). *Define* M_n *by* (3). *Let* τ *be a stopping time.*

(a) *If* $0 \leqq a \leqq b$, *then*

$$P\{\textstyle\sum_1^\tau X_n \leqq a \text{ and } \sum_1^\tau M_n \geqq b\} \leqq \left(\frac{b}{a}\right)^a e^{a-b},$$

where the bound reduces to e^{-b} *for* $a = 0$.

(b) *If* $0 \leqq b \leqq a$, *then*

$$P\{\textstyle\sum_1^\tau X_n \geqq a \text{ and } \sum_1^\tau M_n \leqq b\} \leqq \left(\frac{b}{a}\right)^a e^{a-b},$$

where the bound reduces to 1 *for* $a = b = 0$.

NOTE. Let $c = \max\{a, b\}$. By (9) below,

$$\left(\frac{b}{a}\right)^a e^{a-b} \leqq \exp\left[-\frac{(a-b)^2}{2c}\right],$$

where $\exp x = e^x$.

I will prove (4) in Section 2, using a standard martingale argument: this general approach to proving inequalities was suggested by Theorem 2.12.1 of Dubins and Savage (1965).

I came across (4) while trying to understand this elementary fact. If A_1, $A_2, \cdots$ are independent events, and $P(\bigcup A_n)$ is small, then $b = \sum P(A_n)$ is small. To get this from (4a), make X_n the indicator of A_n, and put $a = 0$: so $\{\sum X_n \leqq a\} = \Omega \backslash \bigcup A_n$, and $P(\Omega \backslash \bigcup A_n) \leqq e^{-b}$. That is,

$$P(\bigcup A_n) \geqq 1 - e^{-b},$$

as is easy to prove directly. Of course, (4) also gives more general inequalities of the same type. Suppose $A_1, A_2, \cdots$ are independent events, and a is a nonnegative integer. If $0 \leqq a \leqq b \leqq \sum P(A_n)$, then (4a) shows:

$$P\{\text{no more than } a \text{ of } A_1, A_2, \cdots \text{ occur}\} \leqq \left(\frac{b}{a}\right)^a e^{a-b}.$$

If $a \geqq b \geqq \sum P(A_i)$, then (4b) shows:

$$P\{\text{at least } a \text{ of } A_1, A_2, \cdots \text{ occur}\} \leqq \left(\frac{b}{a}\right)^a e^{a-b}.$$

Here is a numerical illustration: For a sequence of independent events, such that the sum of their probabilities is $b = 10{,}000$ or more, the chance that $a = 9{,}000$ or fewer occur is at most $e^{-50} \approx 2 \times 10^{-22}$. For a sequence of independent events, such that the sum of their probabilities is $b = 10{,}000$ or less, the chance that $a = 11{,}000$ or more occur is at most $e^{-45} \approx 3 \times 10^{-20}$.

There are classical methods for computing (say) the probability of at least a occurrences exactly, in terms of $P(A_{i_1} \cap A_{i_2} \cap \cdots)$. (For instance, see Feller (1968) Chapter IV or Fréchet (1940); (1943).) There is also a very interesting inequality due to Hoeffding (1956): if $A_1, \cdots, A_n$ are independent, and n as well as $b = P(A_1) + \cdots + P(A_n)$ are fixed, then $P(\sum_1^n 1_{A_i} \leqq a)$ for $a < b$ as well as $P(\sum_1^n 1_{A_i} \geqq a)$ for $a > b$ are maximized when the $P(A_i)$ are all equal. If the X_i's are independent, then (4) is equivalent to well-known inequalities of Bernstein and Kolmogorov. Hoeffding (1963) has further inequalities and a review of the literature. The present inequalities seem to be new in two respects: there are no independence conditions on the X_n; the number of summands is variable, and does not appear in the bound. I hope to discuss a similar inequality for conditional variances in another paper.

Inequality (4) can be used to prove Lévy's conditional form of the Borel–Cantelli lemmas as follows. Let $b \to \infty$ in (4a) to see

$$P\{\textstyle\sum_1^\infty X_n \leqq a \text{ and } \sum_1^\infty M_n = \infty\} = 0\,.$$

Let $a \to \infty$ to see

$$(5) \qquad \textstyle\sum_1^\infty X_n = \infty \qquad \text{a.e. on } \{\sum_1^\infty M_n = \infty\}\,.$$

Similarly, (4b) shows

$$(6) \qquad \textstyle\sum_1^\infty X_n < \infty \qquad \text{a.e. on } \{\sum_1^\infty M_n < \infty\}\,.$$

A harder argument based on (4) proves a variant of Lévy's strong law:

$$\frac{X_1 + \cdots + X_n}{M_1 + \cdots + M_n} \to 1 \qquad \text{a.e. on } \{\textstyle\sum_1^\infty M_n = \infty\}\,;$$

this is sharper than (5). These results will all be proved for unbounded variables subject to a growth condition, in Section 4. There are examples in Section 5 to show that the growth conditions are sharp.

Section 3 concerns the Poisson approximation. For a sequence of small independent sets, the number which occur is approximately Poisson, as is well known. This is still true for dependent events. Suppose $X_1, X_2, \cdots$ are 0–1 variables. Define M_n by (3), so M_n is the conditional probability that X_n is 1, given the past. The process which equals X_1 at time M_1, equals $X_1 + X_2$ at time $M_1 + M_2$, equals $X_1 + \cdots + X_n$ at time $M_1 + \cdots + M_n$, and so on, is approximately Poisson, provided M_n is uniformly small and $\sum M_n$ is large. I hope to explore this elsewhere. Section 3 reports certain inequalities which are by-products of the argument for (4). Sections 4 and 5 do not depend on Section 3.

2. The basic inequalities. Here are some estimates.

(7) LEMMA. (a) *Let* $0 < a \leqq b$. *The function* $\exp[ha - (1 - e^{-h})b]$ *of* $h \geqq 0$ *has a minimum of* $(b/a)^a \exp(a - b)$ *at* $h = \log(b/a)$. *If* $a = 0$, *the* inf *is* $\exp(-b)$, *at* $h = \infty$.

(b) *Let* $0 < b \leqq a$. *The function* $\exp[-ha + (e^h - 1)b]$ *of* $h \geqq 0$ *has a minimum of* $(b/a)^a \exp(a - b)$ *at* $h = \log(a/b)$. *If* $0 = b < a$, *the* inf *is* 0, *at* $h = \infty$. *If* $0 = b = a$, *the function is always* 1.

PROOF. Calculus. □

(8) LEMMA. *Let* $0 < \varepsilon < 1$. *Let*

$$f(\varepsilon) = \left(\frac{1}{1-\varepsilon}\right)^{1-\varepsilon} e^{-\varepsilon} \quad \textit{and} \quad g(\varepsilon) = (1-\varepsilon)e^{\varepsilon}.$$

Then

$$f(\varepsilon) < \exp\left(-\frac{\varepsilon^2}{2}\right) < 1 \quad \textit{and} \quad g(\varepsilon) < \exp\left(-\frac{\varepsilon^2}{2}\right) < 1.$$

PROOF. Take logs, and use Taylor's expansion. For example

$$\log f(\varepsilon) = -(1-\varepsilon)\log(1-\varepsilon) - \varepsilon$$
$$= \varepsilon + \frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3} + \cdots - \varepsilon^2 - \frac{\varepsilon^3}{2} - \cdots - \varepsilon < -\frac{\varepsilon^2}{2}. \qquad \square$$

(9) COROLLARY. *Suppose a and b are nonnegative. Let* $c = \max\{a, b\}$. *Then*

$$(b/a)^a e^{a-b} \leqq \exp\left[-\frac{(a-b)^2}{2c}\right].$$

PROOF. Suppose $a \neq b$ and both are positive: the other cases are easy.

The case $0 < a < b$. Let $\varepsilon = (b-a)/b$. Then

$$\left(\frac{b}{a}\right)^a e^{a-b} = \left[\left(\frac{1}{1-\varepsilon}\right)^{1-\varepsilon} e^{-\varepsilon}\right]^b$$
$$< \exp\left(-\frac{b\varepsilon^2}{2}\right) \qquad \text{by (8)}$$
$$= \exp\left[-\frac{(b-a)^2}{2b}\right].$$

The case $0 < b < a$. Let $\varepsilon = (a-b)/a$. Then

$$\left(\frac{b}{a}\right)^a e^{a-b} = [(1-\varepsilon)e^{\varepsilon}]^a$$
$$< \exp\left[-\frac{a\varepsilon^2}{2}\right] \qquad \text{by (8)}$$
$$= \exp\left[-\frac{(a-b)^2}{2a}\right]. \qquad \square$$

The basic probability estimate is

(10) LEMMA. *Suppose X is a random variable on* $(\Omega, \mathscr{F}, P)$, *with* $0 \leqq X \leqq 1$. *Let* Σ *be a sub-σ-field of* $\mathscr{F}$, *and let* $M = E\{X \mid \Sigma\}$. *Let h be real. Then*

$$E\{\exp(hX) \mid \Sigma\} \leqq \exp[M(e^h - 1)].$$

PROOF. Let $f(x) = \exp(hx)$. Let l be the linear function which agrees with f at 0 and 1:

$$l(x) = 1 + x(e^h - 1).$$

Then $f(X) \leqq l(X)$, so

$$E\{f(X) \mid \Sigma\} \leqq E\{l(X) \mid \Sigma\}$$
$$= l(E\{X \mid \Sigma\}) = l(M) .$$

But $l(M) \leqq \exp[M(e^h - 1)]$. □

For real h, define

(11) $$R_h(m, x) = \exp[hx - (e^h - 1)m] .$$

If $X_1, X_2, \cdots$ satisfy (1)–(2), and $M_1, M_2, \cdots$ are defined by (3), let

(12) $$S_n = X_1 + \cdots + X_n \quad \text{and} \quad T_n = M_1 + \cdots + M_n ,$$

so $S_0 = T_0 = 0$.

(13) PROPOSITION. *R_h is excessive. More precisely, define R_h by* (11). *Then $R_h \geqq 0$. Suppose $X_1, X_2, \cdots$ satisfy* (1)–(2). *Define $\{M_n\}$ by* (3), *and $\{S_n\}$ and $\{T_n\}$ by* (12). *Then the process*

$$R_h(T_n, S_n): \quad n = 0, 1, 2, \cdots$$

is an expectation-decreasing martingale relative to the σ-fields $\mathscr{F}_n$: $n = 0, 1, 2, \cdots$.

PROOF. Use (10). □

It is convenient to extend exp to $\pm\infty$, as follows: $\exp(\infty) = \infty$ and $\exp(-\infty) = 0$. Then exp is continuous on $[-\infty, \infty]$. This makes $R_h(m, x)$ defined and continuous for $0 \leqq m \leqq \infty$ and $0 \leqq x \leqq \infty$, except at $m = x = \infty$. It allows for infinite values of the variables τ, S_τ and T_τ.

(14) COROLLARY. *Suppose $X_1, X_2, \cdots$ satisfy* (1)–(2), *and τ is a stopping time. Define $\{M_n\}$ by* (3), *and $\{S_n\}$ and $\{T_n\}$ by* (12), *and R_h by* (11). *Let*

$$G = \{T_\tau < \infty \ \text{ or } \ S_\tau < \infty\} .$$

Then

$$\int_G R_h(T_\tau, S_\tau)\, dP \leqq 1 .$$

PROOF. From (13),

$$\int R_h(T_{\tau\wedge n}, S_{\tau\wedge n})\, dP \leqq 1 ,$$

where $\tau \wedge n = \min(\tau, n)$. But $R_h(T_{\tau\wedge n}, S_{\tau\wedge n}) \to R_h(T_\tau, S_\tau)$ on G as $n \to \infty$. Use Fatou's lemma. □

This is the main technical inequality in the paper.

THE PROOF OF (4a). Introduce a utility function $u(m, x)$ which is 1 when both $m \geqq b$ and $x \leqq a$, and is 0 otherwise. This defines $u(m, x)$ for $0 \leqq m \leqq \infty$ and $0 \leqq x \leqq \infty$. Remember (11) and (12). Keep

$$G = \{T_\tau < \infty \ \text{ or } \ S_\tau < \infty\} .$$

For $h \geqq 0$, and $0 \leqq m, x \leqq \infty$ except $m = x = \infty$, let

$$Q_h(m, x) = \exp[ha - (1 - e^{-h})b] R_{-h}(m, x)$$
$$= \exp[-h(x - a) + (1 - e^{-h})(m - b)] .$$

Now

$$\begin{aligned} P\{S_\tau \leqq a \ \text{ and } \ T_\tau \geqq b\} &= \textstyle\int u(T_\tau, S_\tau)\, dP \\ &= \textstyle\int_G u(T_\tau, S_\tau)\, dP \\ &\leqq \textstyle\int_G Q_h(T_\tau, S_\tau)\, dP \\ &\leqq Q_h(0, 0): \end{aligned}$$

the first equality is easy; the second holds because $u(T_\tau, S_\tau) = 0$ off G; the third holds because $u(m, x) \leqq Q_h(m, x)$ provided m or x is finite; the last inequality follows from (14). To complete the proof, use (7a) to choose the h minimizing $Q_h(0, 0)$. □

THE PROOF OF (4b). Introduce a utility function $u(m, x)$, which is 1 when both $m \leqq b$ and $x \geqq a$, and is 0 otherwise. This defines $u(m, x)$ for $0 \leqq m \leqq \infty$ and $0 \leqq x \leqq \infty$. Remember (11) and (12). Keep

$$G = \{T_\tau < \infty \ \text{ and } \ S_\tau < \infty\}.$$

For $h \geqq 0$, and $0 \leqq m, x \leqq \infty$ except $m = x = \infty$, let

$$\begin{aligned} Q_h(m, x) &= \exp\,[-ha + (e^h - 1)b]R_h(m, x) \\ &= \exp\,[h(x - a) + (e^h - 1)(b - m)]. \end{aligned}$$

As before

$$\begin{aligned} P\{S_\tau \geqq a \ \text{ and } \ T_\tau \leqq b\} &= \textstyle\int u(T_\tau, S_\tau)\, dP \\ &= \textstyle\int_G u(T_\tau, S_\tau)\, dP \\ &\leqq \textstyle\int_G Q_h(T_\tau, S_\tau)\, dP \\ &\leqq Q_h(0, 0). \end{aligned}$$

Use (7b) to minimize $Q_h(0, 0)$. □

If $0 \leqq b \leqq a$, the bound for the probability in (4a) is 1. For instance, let $X_i = M_i = b/n$, and $\tau = n$. Similarly, the bound in (4b) is 1 for $0 \leqq a \leqq b$. In the range $0 \leqq a \leqq b$, the bound in (4a) is not sharp, but it is of the right order of magnitude for fixed a, as $b \to \infty$. To see this, make $M_i = b/n$ and $\tau = n$. Let the X_i be independent $0 - 1$ variables, taking the value 1 with probability b/n. So $S = X_1 + \cdots + X_n$ is essentially Poisson with parameter b, and

$$P(S \leqq a) \doteq P(S = a) = e^{-b}\,\frac{b^a}{a!} \doteq (2\pi a)^{-\frac{1}{2}} \left(\frac{b}{a}\right)^a e^{a-b}.$$

Similarly for (4b).

The game implicit in (4) is as follows. From (m, x) you can move to $(m + M, x + X)$, where X is any random variable with $0 \leqq X \leqq 1$, and $M = E(X)$. In the variant for (4a), you win \$ 1 if your m-coordinate reaches b before your x-coordinate exceeds a. In the variant for (4b), you win \$ 1 if your x-coordinate reaches a before your m-coordinate exceeds b. The Poisson strategy of using 0–1 variables with constant small expectation is sensible but not optimal. In (4a), for instance, you should change strategies if you ever reach a position on the 45 degree line through $x = a$ and $m = b$. From such a position, you have a sure win by using constant X. This strategy is probably optimal for integer a. For other a, let a^* be the largest integer less than a. When you reach a^*, you should probably switch to gambles which are $a - a^*$ with probability ε and 0 with probability $1 - \varepsilon$. The payoff from this strategy does not seem to have a simple explicit formula. Similar remarks apply to (4b).

For use later, it is helpful to extend (4) as follows.

(15) COROLLARY. *Suppose* $K > 0$. *Suppose* $0 \leqq X_n \leqq K$, *and suppose* (2). *Define* M_n *by* (3). *Let* τ *be a stopping time.*

(a) *If* $0 \leqq a \leqq b$, *then*

$$P\{\textstyle\sum_1^\tau X_n \leqq a \text{ and } \sum_1^\tau M_n \geqq b\} \leqq [(b/a)^a e^{a-b}]^{1/K} .$$

(b) *If* $0 \leqq b \leqq a$, *then*

$$P\{\textstyle\sum_1^\tau X_n \geqq a \text{ and } \sum_1^\tau M_n \leqq b\} \leqq [(b/a)^a e^{a-b}]^{1/K} .$$

PROOF. Use (4) on the variables X_n/K, with a/K for a and b/K for b. □

(16) COROLLARY. *Suppose* $K > 0$. *Suppose* $X_n \geqq 0$ *satisfies* (2). *Define* M_n *by* (3). *Let* τ *be a stopping time. Suppose that for almost all* ω,

$$X_n(\omega) \leqq K \qquad \text{for all} \quad n \leqq \tau(\omega) .$$

(a) *If* $0 \leqq a \leqq b$, *then*

$$P\{\textstyle\sum_1^\tau X_n \leqq a \text{ and } \sum_1^\tau M_n \geqq b\} \leqq [(b/a)^a e^{a-b}]^{1/K} .$$

(b) *If* $0 \leqq b \leqq a$, *then*

$$P\{\textstyle\sum_1^\tau X_n \geqq a \text{ and } \sum_1^\tau M_n \leqq b\} \leqq [(b/a)^a e^{a-b}]^{1/K} .$$

PROOF. Let $X_n^* = X_n$ for $n \leqq \tau$, and $X_n^* = 0$ for $n > \tau$. Let $M_n^* = E(X_n^* \mid \mathscr{F}_{n-1})$. Check that $M_n^* = M_n$ for $n \leqq \tau$ and $M_n^* = 0$ for $n > \tau$. Use (15) on the starred process. □

3. The Poisson approximation. Suppose for now that X_n takes only two values, 0 and 1. Suppose (2)–(3). If the M_n are uniformly small, and $\sum M_n$ is nearly constant at a, then $\sum X_n$ is nearly Poisson with parameter a. The main results (17) and (28) of this section are upper and lower bounds on the generating function $\exp(h \sum X_n)$, which imply this Poisson approximation result.

(17) PROPOSITION. *Suppose* $X_1, X_2, \cdots$ *satisfy* (1)–(2). *Define* M_n *by* (3). *Let* τ *be a stopping time, and suppose* $\sum_1^\tau M_n \leqq b$ *almost everywhere, where* b *is a non-negative real number. For* $h \geqq 0$,

$$E[\exp h(\textstyle\sum_1^\tau X_n)] \leqq \exp[b(e^h - 1)] ,$$

the generating function for a Poisson variable with parameter b. *This bound is sharp.*

PROOF. Remember (11) and (12). Let $h \geqq 0$. If $T_\tau \leqq b$ almost surely, then (14) shows

$$1 \geqq E\{\exp(hS_\tau - (e^h - 1)T_\tau)\} \geqq E\{\exp(hS_\tau)\} \exp[-b(e^h - 1)] .$$

To see the bound is sharp, fix a large positive integer N; make $M_n = b/N$ and let the X_n be independent 0–1 variables, taking the value 1 with chance b/N. So $S_\tau = X_1 + \cdots + X_N$ is essentially Poisson with parameter b. □.

The argument for (17) can also be used to prove

(18) PROPOSITION. *Suppose $X_1, X_2, \cdots$ satisfy (1)–(2). Define M_n by (3), and S_n and T_n by (12). Let τ be a stopping time.*

(a) *If $T_\tau \geqq b$ almost surely, then*

$$E[\exp(-hS_\tau)] \leqq \exp[-b(1 - e^{-h})] \qquad \text{for } h \geqq 0 .$$

(b) *If $S_\tau \leqq a$ almost surely, then*

$$E[\exp(\lambda T_\tau)] \leqq (1 - \lambda)^{-a} \qquad \text{for } 0 \leqq \lambda < 1 .$$

(c) *If $S_\tau \geqq a$ almost surely, then*

$$E[\exp(-\lambda T_\tau)] \leqq (1 + \lambda)^{-a} \qquad \text{for } \lambda \geqq 0 .$$

These bounds are sharp.

To get (b) from (14), make the change of variables $e^h = 1/(1 - \lambda)$. For (c), put $e^h = 1 + \lambda$. To interpret these inequalities, remember that $(1 + \lambda)^{-a}$ is the Laplace transform of the sum of a independent exponential variables for $\lambda > -1$. So (b) and (c) relate T_τ to the time it takes for a Poisson process to reach the level a, at least for positive integer a.

There is a lower bound (28) corresponding to (17). Here are the preliminaries. Let

(19) $$\phi(x) = \frac{1}{x}\log(1 + x) \qquad \text{for } x > 0 .$$

Clearly,

(20) $$\lim_{x \to 0} \phi(x) = 1$$

and

(21) $\phi(x)$ decreases as x increases.

By (21),

(22) If $0 \leqq x \leqq \alpha$, then $(1 + x) \geqq \exp[\phi(\alpha)x]$.

(23) LEMMA. *Fix ε with $0 \leqq \varepsilon < 1$. Let $\theta = \phi(\varepsilon(e - 1))$, where ϕ was defined in (19). Let $0 \leqq h \leqq 1$. Let X be a 0–1 valued random variable on $(\Omega, \mathscr{F}, P)$. Let Σ be a sub-σ-field of $\mathscr{F}$, and let $M = P\{X = 1 \mid \Sigma\}$. Suppose $M \leqq \varepsilon$ almost surely. Then*

$$E\{\exp(hX) \mid \Sigma\} \geqq \exp[\theta M(e^h - 1)] .$$

PROOF. Check that

$$E\{\exp(hX) \mid \Sigma\} = 1 + M(e^h - 1) .$$

Then use (22). □

The regularity condition for (28) is

(24) Fix a positive $\varepsilon < 1$. Suppose X_n is 0–1 and $\mathscr{F}_n$-measurable, and $M_n = P\{X_n = 1 \mid \mathscr{F}_{n-1}\} \leqq \varepsilon$ for all n.

(25) DEFINITION. Let $\theta = \phi(\varepsilon(e-1))$, where ϕ was defined in (19). For $0 \leqq h \leqq 1$, let

$$R_{h,\varepsilon}(m, x) = \exp[hx - \theta(e^h - 1)m].$$

(26) PROPOSITION. *$R_{h,\varepsilon}$ is defective. More precisely, let $0 < \varepsilon < 1$ and $0 \leqq h \leqq 1$. Define $R_{h,\varepsilon}$ by (25). Then $R_{h,\varepsilon} \geqq 0$. Suppose $X_1, X_2, \cdots, M_1, M_2, \cdots$ satisfy (24). Define $\{S_n\}$ and $\{T_n\}$ by (12). Then the process*

$$R_{h,\varepsilon}(T_n, S_n): \quad n = 0, 1, 2, \cdots$$

is an expectation-increasing martingale relative to the σ-fields $\mathscr{F}_0, \mathscr{F}_1, \cdots$.

PROOF. *Use* (23). □

(27) COROLLARY. *Let $0 < \varepsilon < 1$ and $0 \leqq h \leqq 1$. Define $R_{h,\varepsilon}$ by (25). Suppose $X_1, X_2, \cdots$ and $M_1, M_2, \cdots$ satisfy (24). Define $\{S_n\}$ and $\{T_n\}$ by (12). Let σ be a stopping time. Then*

$$E\{R_{h,\varepsilon}(T_{\sigma\wedge n}, S_{\sigma\wedge n})\} \geqq 1.$$

(28) PROPOSITION. *Let $\varepsilon > 0$. Suppose X_n and M_n satisfy (24), and τ is a stopping time. Suppose $\sum_1^\tau M_i \geqq b$ almost surely. Define θ as in (25). Then*

$$E[\exp h(\textstyle\sum_1^\tau X_n)] \geqq \exp[\theta b(e^h - 1)] \qquad \textit{for } 0 \leqq h \leqq 1.$$

PROOF. Let σ be the least n if any with $T_n \geqq b$, and $\sigma = \infty$ if none. So

(29) $$T_\sigma \geqq b \quad \text{almost surely}$$

(30) $$T_\sigma \leqq b + 1 \quad \text{almost surely}.$$

Of course, $\sigma \leqq \tau$: so $S_\sigma \leqq S_\tau$, and it is only necessary to prove the proposition with σ in place of τ. Write $U_n = R_{h,\varepsilon}(T_{\sigma\wedge n}, S_{\sigma\wedge n})$. So $U_n^2 \leqq \exp(2hS_\sigma)$. Then (17) and (30) show that $E(U_n^2)$ is uniformly bounded. This makes U_n uniformly integrable. But $U_n \to R_{h,\varepsilon}(T_\sigma, S_\sigma)$ as $n \to \infty$, whether σ is finite or infinite: use (30) to ensure $T_\sigma < \infty$. And $E(U_n) \geqq 1$ by (27). So

$$E\{R_{h,\varepsilon}(T_\sigma, S_\sigma)\} \geqq 1.$$

Using (29),

$$R_{h,\varepsilon}(T_\sigma, S_\sigma) \leqq \exp(hS_\sigma) \cdot \exp(-\theta b(e^h - 1)). \qquad □$$

At first sight, (17) and (28) look very special. However, as I hope to show elsewhere, it is possible to prove Dvoretzky's general central limit theorem for dependent summands (unpublished) by the same technique.

The lower bounds corresponding to (18) can be obtained in a similar way. Remember

$$\theta = [\varepsilon(e-1)]^{-1}\log[1 + \varepsilon(e-1)],$$

so $\theta < 1$ and $\theta \to 1$ as $\varepsilon \to 0$. Let

$$\theta' = -[\varepsilon(1 - e^{-1})]^{-1}\log[1 + \varepsilon(1 - e^{-1})],$$

so $\theta' > 1$ and $\theta' \to 1$ as $\varepsilon \to 0$. Now $\exp[-hx + \theta'(1 - e^{-h})m]$ is defective for $0 \leqq h \leqq 1$, in the sense of (26).

(31) PROPOSITION. *Suppose* $X_1, X_2, \cdots$ *satisfy* (24). *Define* $\{S_n\}$ *and* $\{T_n\}$ *by* (12). *Let* τ *be a stopping time.*

(a) *If* $T_\tau \leqq b$ *almost surely, then*

$$E\{\exp(-hS_\tau)\} \geqq \exp[-\theta' b(1 - e^{-h})] \qquad \textit{for } 0 \leqq h \leqq 1 .$$

(b) *If* $S_\tau \geqq a$ *almost surely, then*

$$E\{\exp(\lambda T_\tau)\} \geqq \left(\frac{\theta'}{\theta' - \lambda}\right)^a \qquad \textit{for } 0 \leqq \lambda \leqq \theta'(1 - e^{-1}) .$$

(c) *If* $S_\tau \leqq a$ *almost surely, then*

$$E\{\exp(-\lambda T_\tau)\} \geqq \left(\frac{\theta'}{\theta' + \lambda}\right)^a \qquad \textit{for } 0 \leqq \lambda \leqq \theta'(e - 1) .$$

In particular: if $X_1, X_2, \cdots$ satisfy (14), the amount of conditional expectation used by the sum $X_1 + X_2 + \cdots$ in getting to a is nearly the sum of a independent exponential variables.

4. Some almost-sure results.

(32) PROPOSITION. *Suppose the* X_n *are nonnegative, satisfy* (2), *and the* M_n *are defined by* (3). *Then*

$$\textstyle\sum_1^\infty X_n < \infty \qquad \textit{almost surely on} \quad \{\sum_1^\infty M_n < \infty\} .$$

PROOF. Let

$$X_n^* = X_n \quad \text{when} \quad 0 \leqq X_n < 1$$
$$= 0 \quad \text{when} \quad X_n \geqq 1 .$$

Let $M_n^* = E\{X_n^* \mid \mathscr{F}_{n-1}\}$. So $M_n^* \leqq M_n$, and (6) on the starred process shows:

(33) $\sum X_n^* < \infty$ almost surely on $\{\sum M_n < \infty\}$.

By Chebychev's inequality, $P\{X_n \geqq 1 \mid \mathscr{F}_{n-1}\} \leqq M_n$. Use (6) on the indicator functions of $\{X_n \geqq 1\}$:

(34) $X_n \neq X_n^*$ only finitely often, almost surely on $\{\sum M_n < \infty\}$.

Combine (33)-(34). □

The next main result is (39), which extends (5) to certain unbounded variables. Here are the preliminaries.

(35) LEMMA. *Let* ϕ *be a non-decreasing function on* $(0, \infty)$, *with* $\phi(0) > 0$. *Let* x_n *be nonnegative real numbers, with* $\sum x_n = \infty$. *Let* $s_n = x_1 + \cdots + x_n$, *so* $s_0 = 0$.

(a) *If* $\int^\infty 1/\phi(t)\,dt = \infty$, *then* $\sum x_{n+1}/\phi(s_n) = \infty$.
(b) *If* $\int^\infty 1/\phi(t)\,dt < \infty$, *then* $\sum x_n/\phi(s_n) < \infty$.

PROOF. Let $f(u) = x_n$ for $n - 1 \leqq u < n$, and let $F(t) = \int_0^t f(u)\,du$. Then $s_{n-1} \leqq F(t) \leqq s_n$ for $n - 1 \leqq t \leqq n$, so

$$\int_1^{n+1} f(u)/\phi[F(u)]\,du \leqq \sum_1^n x_{i+1}/\phi(s_i)$$
$$\int_0^n f(u)/\phi[F(u)]\,du \geqq \sum_1^n x_i/\phi(s_i)\,.$$

But

$$\int_a^b f(u)/\phi[F(u)]\,du = \int_a^b 1/\phi(F)\,dF = \int_{F(a)}^{F(b)} 1/\phi(t)\,dt\,. \qquad \square$$

NOTE. In claim (a), if $\phi(t)/t$ is unbounded at infinity, and x_n increases rapidly, x_{n+1} cannot be replaced by x_n: for $\sum x_n/\phi(x_n)$ can be made to converge, and then $\sum x_n/\phi(s_n)$ will converge *a fortiori*.

In claim (b), if x_n increases rapidly, it cannot be replaced by x_{n+1}.

The next result is probably known.

(36) COROLLARY. *Suppose the x_n are nonnegative real numbers, with $\sum x_n = \infty$. Let $s_n = x_1 + \cdots + x_n$. Then $\sum x_n/s_n = \infty$.*

PROOF. Use (35a) with $\phi(t) = t$, to see $\sum x_{n+1}/s_n = \infty$. If $x_{n+1} = O(s_n)$, the ratio test shows $\sum x_{n+1}/s_{n+1} = \infty$. If $x_{n+1} \neq O(s_n)$, a direct argument shows $\sum x_{n+1}/s_{n+1} = \infty$. $\square$

The result (5) does not apply to all unbounded variables; a growth condition is needed. To state the condition,

(37) let τ_t be the sup of n with $M_1 + \cdots + M_n \leqq t$, and let $L(t) = \sup_\omega \sup_{n \leqq \tau_t(\omega)} X_n(\omega)$.

So $L(t)$ is not random, $0 \leqq L(t) \leqq \infty$, and $L(t)$ is non-decreasing with t.

(38) FACT. Suppose $X_n \geqq 0$ satisfies (2). Define M_n by (3) and L by (37). Then

$$X_n \leqq L(M_1 + \cdots + M_n)\,.$$

PROOF. Let $t = M_1(\omega) + \cdots + M_n(\omega)$. Then $\tau_t(\omega) \geqq n$, so $X_n(\omega) \leqq L(t)$. $\square$

The next result generalizes (5) to variables satisfying a growth condition.

(39) PROPOSITION. *Suppose $X_n \geqq 0$ satisfies* (2). *Define M_n by* (3) *and L by* (37). *Suppose $L(t) = O(t)$ as $t \to \infty$. Then*

$$\sum_n X_n = \infty \qquad \textit{almost surely on} \quad \{\sum_n M_n = \infty\}\,.$$

PROOF. As in (32), let

$$X_n^* = X_n \qquad \text{for } 0 \leqq X_n < 1$$
$$= 0 \qquad \text{for } X_n \geqq 1\,.$$

Let $M_n^* = E\{X_n^* \mid \mathscr{F}_{n-1}\}$. Now

$$\{\sum M_n = \infty\} = \{\sum M_n^* = \infty\} \cup \{\sum M_n^* < \infty \text{ and } \sum M_n = \infty\}\,.$$

By (5) on the starred process, $\sum X_n^* = \infty$ almost surely on $\{\sum M_n^* = \infty\}$. But $X_n \geqq X_n^*$. Next, confine ω to the set $A = \{\sum M_n^* < \infty$ and $\sum M_n = \infty\}$. Let

$$p_n = P\{X_n \geqq 1 \mid \mathscr{F}_{n-1}\}\,.$$

Remember $T_n = M_1 + \cdots + M_n$ from (12). Then

$$\begin{aligned} M_n &= M_n^* + E\{X_n - X_n^* \mid \mathscr{F}_{n-1}\} \\ &\leqq M_n^* + p_n L(T_n), \end{aligned}$$

because $0 \leqq X_n - X_n^* \leqq X_n \leqq L(T_n)$ by (37). So

$$p_n \geqq \frac{T_n}{L(T_n)} \cdot \frac{M_n}{T_n} - \frac{M_n^*}{L(T_n)}.$$

But $L(t) = O(t)$ by assumption, so $T_n/L(T_n)$ is bounded below as n increases to infinity. And (36) shows $\sum M_n/T_n = \infty$, so $\sum [T_n/L(T_n)] \cdot [M_n/T_n] = \infty$. Finally, $\sum M_n^* < \infty$; so $\sum M_n^*/L(T_n) < \infty$. The upshot is, $\sum p_n = \infty$ on A. Use (5) on the indicator functions of the sets $\{X_n \geqq 1\}$, to see that $X_n \geqq 1$ infinitely often, almost surely on A. □

The condition that $L(t) = O(t)$ can not be weakened, as shown in (48) below.

The next result is a variant of Lévy's strong law for dependent variables subject to a growth condition.

(40) THEOREM. *Suppose the X_n are nonnegative random variables satisfying* (2). *Define the M_n by* (3), *and L by* (37). *Suppose $L(t) = o(t/\log\log t)$ as $t \to \infty$. Then*

$$\frac{X_1 + \cdots + X_n}{M_1 + \cdots + M_n} \to 1 \qquad \textit{almost surely on} \quad \{\textstyle\sum_1^\infty M_n = \infty\}.$$

PROOF. Remember from (12) that $S_n = X_1 + \cdots + X_n$ and $T_n = M_1 + \cdots + M_n$. Let

(41) $$G = \{\textstyle\sum_1^\infty M_n = \infty\}.$$

Fix $r > 1$. Let $\tau(k) = \tau_{r,k}$ be the sup of n with $T_n \leqq r^k$. So $\tau(k)$ is a stopping time, because T_n is $\mathscr{F}_{n-1}$-measurable. Clearly, $\tau(k)$ is finite on G. Check that

(42) $$\tau(k) + 1 \leqq n \leqq \tau(k+1) \qquad \text{iff} \quad r^k < T_n \leqq r^{k+1}.$$

Let

$$\begin{aligned} A_k &= \{G \text{ and } S_n < T_n/r^2 \text{ for some } n \text{ with } \tau(k) + 1 \leqq n \leqq \tau(k+1)\} \\ B_k &= \{G \text{ and } S_n > r^2 T_n \text{ for some } n \text{ with } \tau(k) + 1 \leqq n \leqq \tau(k+1)\}. \end{aligned}$$

It is possible that $\tau(k+1) = \tau(k) < \tau(k) + 1$.

I claim

(43) $$\{G \text{ and } \liminf_{n\to\infty} S_n/T_n < 1/r^2\} \subset \limsup_{k\to\infty} A_k$$

(44) $$\{G \text{ and } \limsup_{n\to\infty} S_n/T_n > r^2\} \subset \limsup_{k\to\infty} B_k.$$

For instance, fix an ω in the set on the left side of (43). Then $T_n \to \infty$, and $S_n/T_n < 1/r^2$ for infinitely many n. So there are infinitely many pairs n and k with $r^k < T_n \leqq r^{k+1}$ and $S_n/T_n < 1/r^2$; use (42) to get ω into the right side.

Next, I will argue that

(45) $$\textstyle\sum_k P(A_k) < \infty$$

and

(46) $$\sum_k P(B_k) < \infty .$$

Given (45) and (46), the usual Borel–Cantelli lemma shows that for each r,

$$P\{G \text{ and } \liminf S_n/T_n < 1/r^2\} = 0$$
$$P\{G \text{ and } \limsup S_n/T_n > r^2\} = 0 .$$

Let r decrease to 1 through a sequence, to see

$$P\{G \text{ and } \liminf S_n/T_n < 1\} = 0$$
$$P\{G \text{ and } \limsup S_n/T_n > 1\} = 0 ;$$

that is,

$$P\{G \text{ and } \lim S_n/T_n \neq 1\} = 0 ,$$

proving (40) from (45)–(46).

Of (45)–(46), the first one is a bit more delicate. To prove it, fix k; then (15) can be used to estimate $P(A_k)$, as follows. Let $X_n^* = X_n$ for $n \leqq \tau(k+1)$ and $X_n^* = 0$ for $n > \tau(k+1)$. So $0 \leqq X_n^* \leqq L(r^{k+1})$, and X_n^* is $\mathscr{F}_n$-measurable. Let $M_n^* = E\{X_n^* \mid \mathscr{F}_{n-1}\}$. As before,

$$\begin{aligned} M_n^* &= M_n && \text{for } n \leqq \tau(k+1) \\ &= 0 && \text{for } n > \tau(k+1) . \end{aligned}$$

Let τ^* be the least n if any with $M_1^* + \cdots + M_n^* > r^k$, and $\tau^* = \infty$ if none. Let

$$A_k^* = \{\textstyle\sum_1^{\tau^*} X_n^* \leqq r^{k-1} \text{ and } \sum_1^{\tau^*} M_n^* \geqq r^k\} .$$

I claim

(47) $$A_k \subset A_k^* .$$

Indeed, fix an ω in A_k; find $n = n(\omega)$ with $\tau(k) + 1 \leqq n \leqq \tau(k+1)$ and $S_n < T_n/r^2$. Then $X_i = X_i^*$ and $M_i = M_i^*$ for $i \leqq \tau(k+1)$, so $\tau^* = \tau(k) + 1$ and

$$\textstyle\sum_1^{\tau^*} X_i^* = S_{\tau(k)+1} \leqq S_n < T_n/r^2 \leqq T_{\tau(k)+1}/r^2 \leqq r^{k+1}/r^2 \leqq r^{k-1} .$$

Similarly, $\sum_1^{\tau^*} M_i^* = T_{\tau(k)+1} \geqq r^k$, proving (47). So $P(A_k) \leqq P(A_k^*)$, which can be estimated from (9) and (15a) on the starred process with $K = L(r^{k+1})$ and $\tau = \tau^*$ and $a = r^{k-1}$ and $b = r^k$. This comes out as follows:

$$P(A_k^*) \leqq \exp(-\alpha_k) , \qquad \text{where}$$
$$\alpha_k = \frac{(r^k - r^{k-1})^2}{2r^k L(r^{k+1})}$$
$$= \frac{(r-1)^2}{2r^3} \cdot \frac{r^{k+1}}{L(r^{k+1}) \log\log r^{k+1}} \cdot \log\log r^{k+1} .$$

In α_k, the first factor is constant; the second goes to infinity, because $L(t) = o(t/\log\log)$; the third is around $\log k$. Consequently, $P(A_k)$ is of the order $1/k^{\theta(k)}$, where $\theta(k) \to \infty$. This proves (45).

The argument for (46) is similar. First, $B_k \subset B_k^*$, where

$$B_k^* = \{S_{\tau(k+1)} \geqq r^{k+2} \text{ and } T_{\tau(k+1)} \leqq r^{k+1}\}.$$

Second, $P(B_k^*)$ can be estimated from (9) and (16b). This comes out as follows:

$$P(B_k) \leqq \exp(-\beta_k), \qquad \text{where}$$

$$\beta_k = \frac{(r^{k+2} - r^{k+1})^2}{2r^{k+2}L(r^{k+1})}$$

$$= \frac{(r-1)^2}{2r} \cdot \frac{r^{k+1}}{L(r^{k+1}) \log\log r^{k+1}} \cdot \log\log r^{k+1}$$

$$\sim 1/k^{\theta(k)},$$

where $\theta(k) \to \infty$. □

5. Examples. The first result shows that the condition $L(t) = O(t)$ in (39) is sharp.

(48) EXAMPLE. Let $\phi(t)$ be a non-decreasing function of t, such that $\limsup_{t\to\infty} \phi(t)/t = \infty$. Then there is a sequence of positive numbers $M_1, M_2, \cdots$ and a sequence of independent random variables $X_1, X_2, \cdots$ such that:

$$\sum M_n = \infty$$

$$E(X_n) = M_n$$

$$0 < X_n \leqq \phi(M_n), \qquad \text{so} \quad X_n \leqq \phi(M_1 + \cdots + M_n)$$

$$\sum X_n < \infty \quad \text{a.e.}$$

PROOF. Choose $M_n > 0$ so large that $\sum M_n = \infty$ and $M_n/\phi(M_n) < 1$ and $\sum M_n/\phi(M_n) < \infty$. Let X_n take only the values 0 and $\phi(M_n)$, with

$$P\{X_n = \phi(M_n)\} = M_n/\phi(M_n).$$

The usual Borel–Cantelli lemma shows

$$P\{X_n \neq 0 \text{ i.o.}\} = 0. \qquad □$$

Definition (37) could be normalized a little differently, by letting τ_t be the inf of n with $M_1 + \cdots + M_n > t$. This normalization leads to a growth condition of the more restrictive form

(49) $$X_{n+1} \leqq L(M_1 + \cdots + M_n),$$

and with this form of the growth condition, (39) changes a bit, as follows.

(50) PROPOSITION. *Suppose the X_n are nonnegative and satisfy* (2). *Define M_n by* (3). *Suppose*

$$X_{n+1} \leqq \phi(M_1 + \cdots + M_n),$$

where ϕ is a non-decreasing function on $[0, \infty)$ with $\phi(0)$ positive.

(a) *If $\int^\infty 1/\phi(t)\,dt = \infty$, then*

$$\sum X_n = \infty \qquad \text{a.e. } \textit{on} \quad \{\sum M_n = \infty\}.$$

(b) *Suppose* $\int^{\infty} 1/\phi(t)\,dt < \infty$. *Let* $M_1, M_2, \cdots$ *be any sequence of positive real numbers, with*

$$\sup_n M_{n+1}/M_n < \infty \qquad \text{and} \qquad \sum_n M_n = \infty\ .$$

There is a sequence of independent random variables $X_1, X_2, \cdots$ *with respective expectations* $M_1, M_2, \cdots$ *such that*

$$0 \leqq X_{n+1} \leqq \phi(M_1 + \cdots + M_n) \qquad \text{and} \qquad \sum X_n < \infty \quad \text{a.e.}$$

PROOF. You can argue the first assertion like (39), getting

$$p_{n+1} \geqq \frac{M_{n+1}}{\phi(M_1 + \cdots + M_n)} - \frac{M^*_{n+1}}{\phi(M_1 + \cdots + M_n)}\ .$$

The sum over n of the first term on the right diverges, by (35 a).

For the second assertion, let X_{n+1} be 0 or $\phi(M_1 + \cdots + M_n)$, with

$$P\{X_{n+1} = \phi(M_1 + \cdots + M_n)\} = \frac{M_{n+1}}{\phi(M_1 + \cdots + M_n)}\ .$$

Then $\sum_1^\infty M_{n+1}/\phi(M_1 + \cdots + M_n) < \infty$ by (35 b), so the ordinary Borel–Cantelli lemma shows

$$P\{X_n \neq 0 \ \text{ i.o.}\} = 0\ . \qquad \square$$

Finally, I would like to show that the growth condition in (40) is sharp.

(51) EXAMPLE. Let $\varepsilon > 0$. There is a positive integer $N = N(\varepsilon)$ and a sequence $X_1, X_2, \cdots$ of independent nonnegative random variables, each having mean 1, such that

$$X_n \leqq \varepsilon n/\log\log n \qquad \text{for } n > N$$

and

$$P\{(X_1 + \cdots + X_n)/n \to 1\} = 0\ .$$

PROOF. Choose N so large that $p_n = (\log\log n)/(\varepsilon n)$ satisfies

$$p_n > 0 \qquad \text{and} \qquad 1 - p_n > e^{-2p_n} \qquad \text{for } n > N\ .$$

Let $X_1, \cdots, X_N$ be 0 or 2 with chance $\frac{1}{2}$ each. For $n > N$, let X_n be 0 or $\varepsilon n/\log\log n$, with $P(X_n = 0) = 1 - p_n$.

Choose $r > 1$ but so close that $\log r < \frac{1}{4}\varepsilon$. Abbreviate $S_n = X_1 + \cdots + X_n$.

Keep k so large that $r^k > N$ and $\log\log r^{k+1} < 2\log k$. Let A_k be the event that $X_n = 0$ for all n with $r^k < n \leqq r^{k+1}$. So

$$\begin{aligned}
P(A_k) &= \Pi\{(1 - p_n) : n = r^k + 1, \cdots, r^{k+1}\} \\
&\geqq \exp\left(-2 \sum \{p_n : n = r^k + 1, \cdots, r^{k+1}\}\right) \\
&\geqq \exp\left(-2\varepsilon^{-1} \log\log r^{k+1} \cdot \sum \{n^{-1} : n = r^k + 1, \cdots, r^{k+1}\}\right) \\
&\geqq \exp\left[-4\varepsilon^{-1} \log k \cdot \log (r^{k+1}/r^k)\right] \\
&\geqq \exp\left(-4\varepsilon^{-1} \log r \cdot \log k\right)\ .
\end{aligned}$$

The upshot is $\sum P(A_k) = \infty$, so $P(\limsup A_k) = 1$. On A_k,

$$S_{r^{k+1}}/r^{k+1} = (1/r) S_{r^k}/r^k\ . \qquad \square$$

NOTE. The argument for (40) shows that $\limsup S_n/n$ and $\liminf S_n/n$ are $O(\varepsilon^{\frac{1}{2}})$ away from 1, almost surely on $\{\sum_i M_i = \infty\}$.

Incidentally, you can also prove

(52) PROPOSITION. *Suppose the X_n are nonnegative random variables satisfying (2). Define the M_n by (3), and L by (37). Suppose the growth condition $L(b) = o(b)$. Then $(X_1 + \cdots + X_n)/(M_1 + \cdots + M_n)$ converges to 1 in probability, given $\{\sum M_i = \infty\}$.*

If $\varepsilon > 0$, there is a sequence of independent nonnegative random variables $X_1, X_2, \cdots$ each having mean 1, such that

$$X_n \leqq \varepsilon n \qquad \text{for } n > 1/\varepsilon$$

$$(X_1 + \cdots + X_n)/n \quad \text{does not converge to} \quad 1 \quad \text{in probability.}$$

REFERENCES

[1] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493-507.

[2] DUBINS, L. E. and FREEDMAN, D. A. (1965). A sharper form of the Borel-Cantelli lemmas and the strong law. *Ann. Math. Statist.* **36** 800-807.

[3] DUBINS, L. E. and SAVAGE, L. J. (1965). *How to Gamble if you Must.* McGraw-Hill, New York.

[4] FELLER, W. (1968). *An Introduction to Probability Theory and its Applications.* **1**, 3rd. ed. Wiley, New York.

[5] FRÉCHET, M. (1940); (1943). Les probabilités associées à un système d'évenements compatibles et dépendants. *Actualités scientifiques et industrielles.* Nos. 859 and 942, Paris.

[6] HOEFFDING, W. (1956). On the distribution of the number of successes in independent trials. *Ann. Math. Statist.* **27** 713-721.

[7] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13-30.

[8] LÉVY, P. (1965). Théorie de l'addition des variables aléatoires. Gauthiers-Villers, Paris.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720