The Annals of Probability
1991, Vol. 19, No. 4, 1732–1736

A NOTE ON THE UPPER BOUND FOR I.I.D. LARGE DEVIATIONS[1]

BY I. H. DINWOODIE

Northwestern University

Let $\bar{X}_n$ denote the mean of an i.i.d. sequence of random vectors $X_1, X_2, X_3, \dots$ taking values in $\mathbf{R}^d$. If λ denotes the convex conjugate of the logarithm of the moment generating function for X_1, then

$$\limsup \frac{1}{n} \log P(\bar{X}_n \in C) \leq -\inf\{\lambda(v): v \in C\}$$

when $C \subset \mathbf{R}^d$ is closed and the moment generating function for X_1 is finite in a neighborhood of the origin. An example is given in which this upper bound fails for a certain closed set in $\mathbf{R}^3$ and the moment generating function for X_1 is not finite in a neighborhood of the origin. An example is also given in which this upper bound is valid for all closed sets but the moment generating function for X_1 is not finite in a neighborhood of the origin.

1. Introduction. Let $(\Omega, \mathscr{A}, P_\mu)$ be a probability triple on which is defined a sequence $X_1, X_2, X_3, \dots$ of i.i.d. random vectors in $E = \mathbf{R}^d$ with common law μ. Define $\lambda_\mu: E \to [0, \infty]$ by

$$\lambda_\mu(v_0) = \sup_{\xi \in \mathbf{R}^d} \left\{ \langle v_0, \xi \rangle - \log \int_E e^{\xi(v)} \mu(dv) \right\},$$

which is the convex conjugate of the logarithm of the moment generating function. Define $\Lambda_\mu(A) = \inf_{v \in A} \lambda_\mu(v)$ when $A \subset E$ and let $\bar{X}_n = n^{-1}(X_1 + \cdots + X_n)$. It is known that

$$\liminf \frac{1}{n} \log P_\mu(\bar{X}_n \in U) \geq -\Lambda_\mu(U)$$

when $U \subset E$ is an open set. The lower bound in this generality is due to Bahadur and Zabell (1979) who also proved that if $C \subset E$ is compact or closed and convex, then

$$\limsup \frac{1}{n} \log P_\mu(\bar{X}_n \in C) \leq -\Lambda_\mu(C). \tag{1.1}$$

It is a standard result in large deviations that if, for some $t > 0$,

$$\int e^{t\|v\|} \mu(dv) < \infty, \tag{1.2}$$

Received March 1990; revised August 1990.
[1]Supported by an NSF Graduate Fellowship.
AMS 1980 *subject classifications*. Primary 60F10.
Key words and phrases. Large deviations, random vectors.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability*. www.jstor.org

then the upper bound (1.1) holds for all closed sets $C \subset E$. A proof is given in Proposition 6.3 of Azencott (1980). The purpose here is to present an example for which (1.1) fails for some closed set C when (1.2) is not satisfied and also to present an example where (1.1) holds for all closed sets C when (1.2) is not satisfied.

It is known that the upper bound (1.1) does not always hold for open sets, even with condition (1.2). Our counterexample to (1.1) for a closed set C in $\mathbf{R}^3$ is based on Slaby's (1988) counterexample to (1.1) for an open set U in $\mathbf{R}^2$.

2. Examples. Example 2.1 will use the following lemma.

LEMMA 2.1. *Suppose that $Y_1, Y_2, \dots$ is an i.i.d. sequence of random vectors in $\mathbf{R}^d$ with law μ defined on the probability triple $(\Omega_0, \mathscr{B}_0, P_\mu)$. If $A = \bigcup_{i \ge 1} C_i \subset \mathbf{R}^d$ is a countable union of closed sets such that $C_i \subset C_{i+1}$ and if*

$$\limsup \frac{1}{n} \log P_\mu(\overline{Y}_n \in A) > -\Lambda_\mu(A),$$

then there exist an i.i.d. sequence $X_1, X_2, \dots$ of random vectors in $\mathbf{R}^{d+1}$ with law π defined on a probability triple $(\Omega, \mathscr{B}, P_\pi)$ and a closed set $C \subset \mathbf{R}^{d+1}$ such that

$$\limsup \frac{1}{n} \log P_\pi(\overline{X}_n \in C) > -\Lambda_\pi(C).$$

Furthermore, $E_\pi \|X_1\|^m < \infty$ if $E_\mu \|Y_1\|^m < \infty$, for $m \ge 1$.

REMARK. If

$$\liminf \frac{1}{n} \log P_\mu(\overline{Y}_n \in A) > -\Lambda_\mu(A),$$

then C can be chosen so that

$$\liminf \frac{1}{n} \log P_\pi(\overline{X}_n \in C) > -\Lambda_\pi(C).$$

PROOF OF LEMMA 2.1. Let $p_k = c \exp(-|k|^{1/2})$, for $k = \pm 1, \pm 2, \dots$, where $c^{-1} = \sum_{k=-\infty}^{\infty} \exp(-|k|^{1/2})$. Let $Z_1, Z_2, \dots$ be an i.i.d. sequence of real-valued random variables on a probability triple $(\Omega_1, \mathscr{B}_1, P_\nu)$ such that $P_\nu(Z_1 = k) = p_k$. Let $(\Omega, \mathscr{B}, P_\pi) = (\Omega_0 \times \Omega_1, \mathscr{B}_0 \times \mathscr{B}_1, P_\mu \times P_\nu)$ and let X_i: $\Omega_0 \times \Omega_1 \to \mathbf{R}^{d+1}$ be defined by $X_i = (Y_i, Z_i)$. The law π on $\mathbf{R}^{d+1}$ for X_1 can then be written $\pi = \sum_{k=-\infty}^{\infty} p_k \mu \circ f_k^{-1}$, where f_k: $\mathbf{R}^d \to \mathbf{R}^{d+1}$ is given by $f_k(x) = (x, k)$.

Let $A_k^\varepsilon = A \times \{x \in \mathbf{R}: |x - k| \le \varepsilon\}$ for each $k = \pm 1, \pm 2, \dots,$ where $0 < \varepsilon < \frac{1}{2}$. Now, using Chernoff's theorem,

$$\begin{aligned}\limsup \frac{1}{n} \log P_\pi(\overline{X}_n \in A_k^\varepsilon) &= \limsup \frac{1}{n} \log\left[P_\mu(\overline{Y}_n \in A) P_\nu(\overline{Z}_n \in [k - \varepsilon, k + \varepsilon])\right] \\ &= \limsup \frac{1}{n} \log P_\mu(\overline{Y}_n \in A) \\ &> -\Lambda_\mu(A) + \delta,\end{aligned}$$

for some $\delta > 0$ suitably small. Now we will show that, for every $k \ge 1$,

$$\Lambda_\mu(A) = \Lambda_\pi(A_k^\varepsilon), \tag{2.1}$$

from which it will follow that

$$\limsup \frac{1}{n} \log P_\pi(\overline{X}_n \in A_k^\varepsilon) > -\Lambda_\pi(A_k^\varepsilon) + \delta.$$

To see (2.1), let $(x_0, r_0) \in A_k^\varepsilon$. Then

$$\begin{aligned}\lambda_\pi(x_0, r_0) &= \sup_{(\xi, t) \in \mathbf{R}^d \times \mathbf{R}} \left\{\langle (x_0, r_0), (\xi, t)\rangle - \log \int_{\mathbf{R}^{d+1}} \exp\langle v, (\xi, t)\rangle \pi(dv)\right\} \\ &= \sup_{(\xi, t)} \left\{\langle (x_0, r_0), (\xi, t)\rangle - \log \sum_{k=-\infty}^{\infty} e^{kt}\left[\int_{\mathbf{R}^d} e^{\xi(x)} \mu(dx)\right] p_k\right\}.\end{aligned}$$

Now we can assume that $t = 0$, since otherwise the moment generating function is seen to be infinite; so

$$\lambda_\pi(x_0, r_0) = \sup_{\xi \in \mathbf{R}^d} \left\{\langle x_0, \xi\rangle - \log \int_{\mathbf{R}^d} e^{\xi(x)} \mu(dx)\right\} = \lambda_\mu(x_0).$$

Thus $\Lambda_\pi(A_k^\varepsilon) = \Lambda_\mu(A)$ for any k, which proves (2.1).

Hence, for $k = 1$, there exists n_1 such that

$$\frac{1}{n_1} \log P_\pi(\overline{X}_{n_1} \in A_1^\varepsilon) > -\Lambda_\pi(A_1^\varepsilon) + \delta.$$

We continue and construct a sequence $n_1 < n_2 < \cdots$ such that

$$\frac{1}{n_k} \log P_\pi(\overline{X}_{n_k} \in A_k^\varepsilon) > -\Lambda_\pi(A_k^\varepsilon) + \delta.$$

Let $C_{i,k} = C_i \times \{x \in \mathbf{R}: |x - k| \le \varepsilon\}$. Then $C_{i,k} \subset C_{i+1,k}$ and $A_k^\varepsilon = \bigcup_{i \ge 1} C_{i,k}$. Hence, for each $k \ge 1$, there exists some integer i_k such that

$$\frac{1}{n_k} \log P_\pi(\overline{X}_{n_k} \in C_{i_k,k}) > -\Lambda_\pi(A_k^\varepsilon) + \delta.$$

Let $C = \bigcup_{k \geq 1} C_{i_k, k}$. C is closed, and it is clear by construction that

$$\begin{aligned}\limsup_n \frac{1}{n} \log P_\pi(\overline{X}_n \in C) &\geq \limsup_k \frac{1}{n_k} \log P_\pi(\overline{X}_{n_k} \in C_{i_k, k}) \\ &\geq -\liminf_k \Lambda_\pi(A_k^\varepsilon) + \delta \\ &\geq -\Lambda_\pi\Big(\bigcup_{k \geq 1} A_k^\varepsilon\Big) + \delta \\ &\geq -\Lambda_\pi(C) + \delta.\end{aligned}$$

Since $\sum_{k=-\infty}^{\infty} |k|^m p_k < \infty$, for every $m \geq 0$, it follows that $E_\pi \|X_1\|^m < \infty$ if $E_\mu \|Y_1\|^m < \infty$. This proves the lemma. □

EXAMPLE 2.1. There exist a closed set C in $\mathbf{R}^3$ and a probability π on $\mathbf{R}^3$ such that an i.i.d. sequence $X_1, X_2, \ldots$ in $\mathbf{R}^3$ with law π satisfies $E_\pi \|X_1\|^m < \infty$, for all $m \geq 1$, and

$$\limsup \frac{1}{n} \log P_\pi(\overline{X}_n \in C) > -\Lambda_\pi(C).$$

PROOF. Define μ on $\mathbf{R}^2$ as in Slaby (1988). Let $f(x) = x^2 \log\log(1/x)$ and let A_t be the set $\{(x, y): 0 < x \leq t,\ 0 \leq y \leq f(x)\}$. Lemmas 2.1, 2.2 and 2.3 of Slaby show that if $Y_1, Y_2, \ldots$ is an i.i.d. sequence of $\mathbf{R}^2$-valued random vectors with common law μ and if $t > 0$ is small, then

$$\limsup \frac{1}{n} \log P_\mu(\overline{Y}_n \in A_t) > -\Lambda_\mu(A_t).$$

It is clear that $A_t = \bigcup_{i \geq 1} C_i$, where $C_i = \{(x, y): t/i \leq x \leq t,\ 0 \leq y \leq f(x)\}$. Now we can apply Lemma 2.1. Since μ is a measure of compact support, it follows that $E_\pi \|X_1\|^m < \infty$, for every $m \geq 1$. □

If $E = \mathbf{R}^1$, it is easily proved using the techniques of Bahadur and Zabell (1979) that the upper bound holds for all closed sets with no conditions on the law μ. Variations of the following example show that the upper bound can hold for all closed sets in $\mathbf{R}^d$ for any $d \geq 2$ without (1.2). The example is of interest because although the moment generating function does not exist in a neighborhood of the origin, λ_μ is nevertheless not identically zero and so the upper bound (1.1) is nontrivial.

EXAMPLE 2.2. Let the law μ on $\mathbf{R}^2$ governing the i.i.d. sequence $X_1, X_2, \ldots$ be the product of the normal distribution and the Cauchy distribution and let $C \subset \mathbf{R}^2$ be a closed set. Then $\lambda_\mu(x, y) = \frac{1}{2}x^2$, and thus the set $\{v: \lambda_\mu(v) \geq \Lambda_\mu(C)\}$ is the union of two closed half spaces H_1 and H_2 such that

$$\Lambda_\mu(H_i) = \Lambda_\mu(C).$$

Now from Lemma 2.6 of Bahadur and Zabell (1979),

$$\limsup \frac{1}{n} \log P_\mu(\overline{X}_n \in C) \le \max_{1 \le i \le 2} \limsup \frac{1}{n} \log P_\mu(\overline{X}_n \in H_i)$$
$$\le -\min_i \Lambda_\mu(H_i) = -\Lambda_\mu(C).$$

This proves that the upper bound holds for all closed sets, although (1.2) does not hold.

The preceding example is essentially one-dimensional since the set $\{v: \lambda_\mu(v) \ge \Lambda_\mu(C)\}$ can be written as the union of two closed half-spaces. We will give another example without this property in which (1.1) holds for all closed sets although (1.2) is not satisfied. Let P_1 be the normal distribution on $\mathbf{R}^1$ and let P_2 be the probability on $\mathbf{R}^1$ with density

$$f(y) = \begin{cases} y^{-2}, & y > 1, \\ 0, & y \le 1. \end{cases}$$

The law μ on $\mathbf{R}^2$ will be the product of P_1 and P_2. If $C \subset \mathbf{R}^2$ is closed, then one can show that for each $c < \Lambda_\mu(C)$, the set $\{v: \lambda_\mu(v) \ge \Lambda_\mu(C)\}$ is contained in a finite union of closed sets $\bigcup_{i=1}^k C_i$ such that $\Lambda_\mu(C_i) \ge c$ and such that each C_i is either compact or convex. This can be used to show (1.1) for the set C with Lemmas 2.5 and 2.6 of Bahadur and Zabell (1979).

An open problem is to construct a counterexample to (1.1) in $\mathbf{R}^2$. This will likely be more difficult than Example 2.1 since we avoided asymptotic calculations by using those of Slaby (1988) and stacking his example in $\mathbf{R}^3$.

Acknowledgment. I would like to thank Professor S. L. Zabell for having suggested this problem.

REFERENCES

AZENCOTT, R. (1980). Grandes déviations et applications. In *Ecole d'Eté de Probabilités de Saint-Flour VIII—1978. Lecture Notes in Math.* **774** 1–176. Springer, New York.

BAHADUR, R. R. and ZABELL, S. L. (1979). Large deviations of the sample mean in general vector spaces. *Ann. Probab.* **7** 587–621.

SLABY, M. (1988). On the upper bound for large deviations of sums of i.i.d. random vectors. *Ann. Probab.* **16** 978–990.

DEPARTMENT OF MATHEMATICS
LUNT HALL
NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS 60208