Hitting probabilities and large deviations
Collamore, Jeffrey F.

Published in:
Annals of Probability

Publication date:
1996

Document Version
Early version, also known as pre-print

Citation for published version (APA):
HITTING PROBABILITIES AND LARGE DEVIATIONS

BY JEFFREY F. COLLAMORE

University of Wisconsin–Madison

Let \( Y_n \) be a sequence of random variables in \( \mathbb{R}^d \) and let \( A \subset \mathbb{R}^d \) be a sequence of random variables in \( \mathbb{R}^d \). Then \( P(Y_n \in A \text{ for some } n) \) is the hitting probability of the set \( A \) by the sequence \( \{Y_n\} \). We consider the asymptotic behavior, as \( m \to \infty \), of \( P(Y_n \in mA, \text{ some } n) = P(\text{hitting } mA) \) whenever (1) the probability law of \( Y_n/n \) satisfies the large deviation principle and (2) the central tendency of \( Y_n/n \) is directed away from the given set \( A \). For a particular function \( I \), we show \( P(Y_n \in mA, \text{ some } n) \sim \exp(-mI(A)) \).

1. Introduction. Let \( Y_1, Y_2, \ldots \) be a sequence of random variables on a probability space \( (\Omega, \mathcal{F}, P) \), and let \( A \subset \Omega \). Then \( P(Y_n \in A \text{ for some } n) \) is the hitting probability of the set \( A \) by the sequence \( \{Y_n\} \).

Historically, the asymptotic properties of a sequence of hitting probabilities

\[ P\{Y_n \in A, \text{ some } n\}, \; P\{Y_n \in 2A, \text{ some } n\}, \; P\{Y_n \in 3A, \text{ some } n\}, \ldots \]

were first studied by Lundberg (1909). In an effort to analyze the capital fluctuations incurred by an insurance company, he considered the following problem. Suppose that we let \( Y_t \) represent the capital available to an insurance company at time \( t \). Then \( \{Y_t\}_{t \geq 0} \) can be modelled as the sum of two processes: a compound Poisson process, resulting from claims being paid by the insurance company which are i.i.d. and arrive at a Poisson rate, and a deterministic linear drift, resulting from a constant inflow of premiums from the policy holders. The second is assumed to grow faster than the first, implying \( E(Y_t/t) = \text{const.} > 0 \). In short, \( \{Y_t\}_{t \geq 0} \subset \mathbb{R} \) is a Lévy process having positive drift. Then \( P\{Y_t < -m, \text{ some } t \geq 0\} \) is less than 1 for any positive \( m \), and in the ruin problem one studies how \( P\{Y_t < -m, \text{ some } t \geq 0\} \) varies with \( m \). It is known that asymptotically

\[ P\{Y_t < -m, \text{ some } t \geq 0\} \sim Ce^{-Rm} \quad \text{as } m \to \infty \]

for certain constants \( C \) and \( R \). See Cramér (1954).

Recently, there has been an interest in refining the above estimate. See, for example, Lalley (1984), Martin-Löf (1986) and Nyrhinen (1994). However, the multidimensional problem of first passage of general processes into general regions in the plane—a problem of theoretical and of applied interest, for example, in risk theory, sequential analysis and queueing theory—has not been considered.
Here, we address this multidimensional problem for discrete time. Letting $Y_1, Y_2, \ldots$ be a general sequence of random variables for which the probability law of $Y_n/n$ satisfies the large deviation principle and letting $A \subset \mathbb{R}^d$ be a general set which avoids the central tendency of $\{Y_n\}$, we consider the asymptotic behavior of

$$\frac{1}{m} \log P\{Y_n \in mA, \text{ some } n \geq N\}$$

as $m \to \infty$. Under some natural conditions on the set $A$ and the sequence $\{Y_n\}$, we show that there exists a rate function $I$ such that the equation

(1.1) \quad \lim_{m \to \infty} \frac{1}{m} \log P\{Y_n \in mA, n \geq N\} = -\inf_{x \in A} I(x)

holds for suitably large $N$, provided that the set $A$ is convex and open, and under slightly stronger conditions if $A$ is a general open set. Thus—subject to appropriate restrictions on the set $A$—we show that (1.1) is valid for any sequence $Y_1, Y_2, \ldots \subset \mathbb{R}^d$ which is either the sums of an i.i.d. sequence of random variables or the additive functionals of a Markov chain or a general sequence satisfying the conditions of the Gärtner–Ellis theorem. [As discussed below, $N$ can be taken to be 1 in (1.1) under an additional regularity condition on $\{Y_n\}$ satisfied, e.g., in the i.i.d. or Markov cases.]

A related work is Dembo, Karlin and Zeitouni (1994). Letting $\{Y_t\}_{t \geq 0}$ be a multidimensional Lévy process and defining

$$\bar{T}(mA) = \inf\{t : \exists s \in [0, t) \ni Y_t - Y_s \in mA\},$$

they have derived, among other things, a law of large numbers for $\{\bar{T}(mA)\}_{m \in \mathbb{Z}_+}$. Their results are not strictly comparable to ours. In the analog of their problem in our setting we would have $Y_n = X_1 + \ldots + X_n$, where $\{X_i\}_{i \in \mathbb{Z}_+}$ is an i.i.d. sequence of random variables. Then, for example,

$$P\{\bar{T}(mA) < \infty\} \geq P\{X_i \in mA \text{ for some } i \in \mathbb{Z}_+\},$$

implying $P\{\bar{T}(mA) < \infty\} = 1$ for all fixed $m$, whenever the support of $X_i$ is $\mathbb{R}^d$. In contrast, we consider $P\{T(mA) < \infty\}$, where $T(mA) = \inf\{n : Y_n \in mA\}$, and show that this probability exhibits exponential decay in $m$, and that such decay occurs quite generally whenever the sequence $Y_1, Y_2, \ldots$ satisfies the large deviation principle. [Some of their results have also been derived in the context of finite state Markov chains. See Dembo and Karlin (1991) and Karlin and Dembo (1992).]

In our problem it is possible to obtain, in addition, a qualitative description for the manner in which the process $\{Y_n\}$ first enters the set $mA$. Namely, we can find conditional laws of large numbers describing the place and time
of first passage, in the event that \( \{Y_n\} \) enters the set \( mA \) in finite time. A discussion of these conditional laws is given at the end.

2. Statement of results. Let \( Y_1, Y_2, \ldots \) be a sequence of random variables taking values in \( \mathbb{R}^d \). For each \( n \), let \( \mu_n \) denote the probability law of \( Y_n/n \).

Our objective is to study the asymptotic behavior, as \( m \to \infty \), of

\[
T_N(mA) = \inf \{ n \geq N: Y_n \in mA \}
\]

for open sets \( A \), and particularly of \( \mathbb{P}\{T_N(mA) < \infty\} \), that is, \( \mathbb{P}\{ \text{hitting } mA \text{ after time } N\} \).

We will characterize this behavior whenever \( \{\mu_n\} \) satisfies the large deviation principle (as defined in Dembo and Zeitouni (1993)) with a convex, lower semicontinuous rate function \( I \) and some regularity conditions hold for the set \( A \) and the sequence \( \{Y_n\} \). To describe these conditions let

\[
\Lambda(\alpha) = \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E} \exp(\alpha, Y_n), \quad \forall \alpha \in \mathbb{R}^d,
\]

\[
\Lambda_N(\alpha) = \sup_{n \geq N} \frac{1}{n} \log \mathbb{E} \exp(\alpha, Y_n), \quad \forall \alpha \in \mathbb{R}^d,
\]

\[
\mathcal{L}_a f = \{x: f(x) \leq a\} \quad \text{for any } f: \mathbb{R}^d \to \mathbb{R},
\]

\[
f^*(x) = \sup_{a \in \mathbb{R}^d} \{\langle \alpha, x \rangle - f(\alpha)\} \quad \forall x \in \mathbb{R}^d \quad \text{(the convex conjugate of } f),
\]

\[
\text{cone } S = \{\lambda x: \lambda > 0, \ x \in S\} \quad \text{for each set } S \subset \mathbb{R}^d,
\]

\[
\text{ri } S = \text{the interior of } S \text{ relative to the subspace generated by } S,
\]

\[
\mathcal{D} = \text{cone(dom } I),
\]

and, if the level sets of \( I \) are compact,

\[
\mathcal{E}_\delta = \left\{ x: \inf_{y \in \mathcal{L} I} \|x - y\| < \delta \right\}.
\]

Let \( \mathcal{E}_\delta \) be the slightly larger set \( \{x: \|x - y\| < (1 + \|y\|)\delta, \ \text{some } y \in \mathcal{L} I\} \) if the level sets of \( I \) are not compact. It is assumed \( \mathcal{L} I = \{x: I(x) = 0\} \neq \emptyset \).

Hypotheses:

(H1) \( \Lambda(\alpha) < \infty \) for all \( \alpha \in \mathcal{L} I^* \).

(H2) For some \( \delta > 0 \), \( A \cap \text{cone } \mathcal{E}_\delta = \emptyset \).
Hypotheses (H1) and (H2) are natural and widely satisfied whenever \( \{\mu_n\} \) satisfies the large deviation principle and the central tendency of \( Y_n/n \) is directed away from the set \( A \). Consider, for example, the case when \( \{X_i\} \) is an i.i.d. sequence of random variables and \( Y_n = X_1 + \cdots + X_n \). The fundamental large deviations result in this setting is Cramér’s theorem, which states that if \( \Lambda \) is finite in a neighborhood of the origin, then \( \{\mu_n\} \) satisfies the large deviation principle with rate function \( I = \Lambda^* \). Then \( I^* = \Lambda^{**} \). Since \( \Lambda \) is convex and lower semicontinuous, it follows that \( I^* = \Lambda \). Hence \( \Lambda(\alpha) = I^*(\alpha) \) for all \( \alpha \) and therefore (H1) always holds. Also, in this situation \( \mathcal{L}_0I = \{x: I(x) = 0\} = \{E X_1\} \).

Thus (H2) holds if \( E X_1 \neq 0 \) and the set \( A \) stays away from a thin cone containing the mean ray \( \{\lambda(E X_1): \lambda > 0\} \). That is, if for some \( \delta > 0 \),

\[
A \cap \{\lambda x: \lambda > 0, \ x \in B_{E X_1, \delta}\} = \emptyset,
\]

where \( B_{E X_1, \delta} \) is a \( \delta \)-ball about \( E X_1 \). Similarly, if \( Y_1, Y_2, \ldots \) is a Markov-additive chain as in Ney and Nummelin (1987a, b) or a general sequence satisfying the conditions of the Gartner–Ellis theorem, then (H1) always holds, for the same reasons as for i.i.d. sums, whereas (H2) holds under a condition very similar to (2.1).

**Theorem 2.1.** Suppose \( A \) is a convex open set intersecting \( \text{ri} \mathcal{D} \) and (H1) and (H2) are satisfied. Then for sufficiently large \( N \),

\[
\lim_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(mA) < \infty\} = -\inf_{x \in A} \bar{I}(x),
\]

where \( \bar{I}(x) = \sup_{\alpha \in \mathcal{L}_0I^*} \langle \alpha, x \rangle \) is the support function of \( \mathcal{L}_0I^* \).

**Remark.** By Rockafellar [(1970), Theorem 13.5], \( \bar{I} = \text{cl} J \), where \( J(x) = \inf_{\tau > 0} \tau^{-1} I(\tau x) \) is the positively homogeneous function generated by \( I \). Note that the domain of \( J \) is \( \mathcal{D} = \text{cone}(\text{dom} I) \). Hence, from the equation \( \bar{I} = \text{cl} J \) we infer that the relative interior of the domain of \( \bar{I} \) is \( \text{ri} \mathcal{D} \).

On the other hand, if \( A \subset \mathbb{R}^d \) is a general open set, then (2.2) need no longer govern the asymptotic behavior of \( \mathbb{P}\{T_N(mA) < \infty\} \). Examples exist of sequences \( Y_1, Y_2, \ldots \) satisfying the Gärtnner–Ellis theorem, where the hypotheses of Theorem 2.1 hold, yet the conclusion of that theorem fails for certain sets \( A \). In these situations the strongest upper bound which can be established is

\[
\lim_{N \to \infty} \left\{ \limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(mA) < \infty\} \right\} = -\inf_{x \in A} \bar{I}(x).
\]

In particular, we do not obtain (2.2) for any fixed value of \( N \). See Collamore (1996).
Thus, if \( A \) is not convex we need a further hypothesis; namely, strengthen (H1) as follows:

\[
\text{(H1')} \quad \Lambda_N(\alpha) < \infty \text{ for all } \alpha \in \mathcal{A}_0 I^* \text{ and } N \geq \text{ some } N_0.
\]

Now (H1') always holds whenever one of the following statements hold:

1. \( Y_n \) is the \( n \)th sum of an i.i.d. sequence of random variables satisfying the condition given in Cramer’s theorem or the weaker condition given in Ney and Robinson (1995). Then \( n^{-1} \log \mathbf{E} \exp(\alpha, Y_n) \) does not depend on \( n \); consequently (H1') holds with \( N_0 = 1 \).
2. \( Y_1, Y_2, \ldots \) is a Markov-additive process satisfying the uniform recurrence condition (6.2) in Ney and Nummelin (1987a). Then (H1') holds with \( N_0 = 1 \).
3. \( Y_1, Y_2, \ldots \) is a general sequence satisfying the conditions of the Gärtner–Ellis theorem and (i) \( \Lambda(\alpha) \) is finite in a neighborhood of each \( \alpha \in \mathcal{A}_0 \Lambda \); (ii) the level sets of \( \Lambda \) are compact. (However, it can be shown that the latter condition is unnecessary when \( \text{cl } A \subset \text{ int } \mathcal{O} \).)

**Theorem 2.2.** Suppose \( A \subset \mathbb{R}^d \) and (H1') and (H2) are satisfied. Then for any \( N \geq N_0 \),

\[
\liminf_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{ T_N(mA) < \infty \} \geq \inf_{x \in \text{int } A} \bar{I}(x)
\]

and

\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{ T_N(mA) < \infty \} \leq \inf_{x \in \text{cl } A} \bar{I}(x),
\]

where \( \bar{I}(x) = \sup_{\alpha \in \mathcal{A}_0 I^*} \langle \alpha, x \rangle \) is the support function of \( \mathcal{A}_0 I^* \).

**3. Some preliminary results from convex analysis.**

**Notation.** For any \( \mathcal{S} \subset \mathbb{R}^d \), \( x \in \mathbb{R}^d \) and \( f: \mathbb{R}^d \to \mathbb{R} \), let

\[
\mathcal{N}_{\mathcal{S}}(x) = \text{normal cone to } \mathcal{S} \text{ at } x,
\]

\(\text{rc } \mathcal{S} = \text{recession cone of } \mathcal{S},\)

\(\text{rec } f = \text{recession function of } f,\)

\(\text{rc } f = \text{recession cone of } f\)

\(\partial f(x) = \text{set of subgradients of } f \text{ at } x.\)

[For definitions, see Rockafellar (1970).] Let \( \bar{I}(x) = \sup_{\alpha \in \mathcal{A}_0 I^*} \langle \alpha, x \rangle \) be the support function of \( \mathcal{A}_0 I^* \).

**Lemma 3.1.** Under (H2), the function \( \bar{I} \) is strictly positive on \( \mathcal{S} = (\text{cone } \mathcal{S})^\circ - \{0\}.\)
**Proof.** Note that \( I = \text{cl} \, J \), where \( J(x) = \inf_{\tau \geq 0} \tau^{-1} I(\tau x) \) is the positively homogeneous function generated by \( I \) \cite[Theorem 13.5]{Rockafellar1970}. Hence \( \overline{I}(x) = \min \{ \text{rec} \, I(x), J(x) \} \) \cite[Theorem 9.7]{Rockafellar1970}. That is, \( \overline{I}(x) > 0 \) if and only if \( \text{rec} \, I(x) > 0 \) and \( J(x) > 0 \).

Let \( x \in \mathcal{D} \). Then we assert \( \text{rec} \, I(x) > 0 \) and \( J(x) > 0 \). To show \( \text{rec} \, I(x) > 0 \), observe that

\[
\text{rc} \, A_0 I \subset \text{rc} \, (\text{cl cone} \, A_0 I) \subset \text{cl cone} \, A_0 I \subset (\text{cone} \, \mathcal{D}_0 \cup \{0\}) = \mathcal{D}^\circ.
\]

Since \( \text{rc} \, I = \text{rc} \, A_0 I \) \cite[Theorem 8.7]{Rockafellar1970}, this implies \( \text{rc} \, I \subset \mathcal{D}^\circ \). From the definition of \( \text{rc} \, I \) it follows that \( \{ x : \text{rc} \, I \leq 0 \} \subset \mathcal{D}^\circ \). Then \( x \in \mathcal{D} \Rightarrow \text{rc} \, I(x) > 0 \).

Next, to show \( J(x) > 0 \), observe \( x \notin \text{cone} \, A_0 I \Rightarrow \tau^{-1} I(\tau x) > 0 \) for any given \( \tau > 0 \). Moreover \( \tau^{-1} I(\tau x) \nrightarrow 0 \) as \( \tau \to 0 \), because \( H2 \Rightarrow I(0) \neq 0 \); \( \tau^{-1} I(\tau x) \nrightarrow 0 \) as \( \tau \to \infty \), because either \( \tau^{-1} I(\tau x) \equiv \infty \), \( \forall \tau \), or \( I(\tau_0 x) < \infty \) for some \( \tau_0 \) and then \( \lim_{\tau \to \infty} \tau^{-1} I(\tau x) = \lim_{\tau \to \infty} \tau^{-1} \{ I((\tau + \tau_0) x) - I(\tau_0 x) \} = \text{rec} \, I(x) > 0 \) \cite[Theorem 8.5]{Rockafellar1970}. \( \square \)

**Lemma 3.2.** Assume \( A \subset \mathbb{R}^d \) is a convex open set intersecting \( \text{ri} \, \mathcal{D} \) and \( H2 \) is satisfied. Let \( a = \inf_{x \in A} \overline{I}(x) \). Then there exists \( \alpha_0 \in A_0 I^* \) such that the hyperplane

\[
\mathcal{H} = \{ x : \langle \alpha_0, x \rangle = a \}
\]

separates \( A \) and \( A_0 \overline{I} \), with \( A \subset \{ x : \langle \alpha_0, x \rangle > a \} \) and \( A_0 \overline{I} \subset \{ x : \langle \alpha_0, x \rangle \leq a \} \).

**Proof.** Since \( A \) and \( A_0 \overline{I} \) are convex sets with no common points in their relative interiors, there exists a separating hyperplane

\[
\mathcal{H} = \{ x : \langle \alpha, x \rangle = t \}, \quad \text{where} \, \alpha \in \mathbb{R}^d - \{0\} \text{ and } t \geq 0.
\]

If \( t > 0 \), then because \( A_0 \overline{I} \) is a convex set containing \( \{0\} \), we must have \( A \subset \{ x : \langle \alpha, x \rangle > t \} \) and \( A_0 \overline{I} \subset \{ x : \langle \alpha, x \rangle \leq t \} \). These inclusions also hold when \( t = 0 \), possibly after replacing \( \alpha \) with \( -\alpha \).

To complete the proof, we now identify the quantities \( \alpha \) and \( t \). Here we argue indirectly, first establishing the existence of a minimizing element \( x_0 \in \text{cl} \, A \) such that

\[
\overline{I}(x_0) = \inf_{x \in A} \overline{I}(x) = a.
\]

Then \( x_0 \in A_0 \overline{I} \) and \( x_0 \in \text{cl} \, A \Rightarrow x_0 \in \mathcal{H} \). Next we show that for some \( \alpha_0 \in A_0 I^* \), which is just a constant multiple of \( \alpha \), we have

\[
a = \overline{I}(x_0) = \sup_{\tilde{\alpha} \in A_0 \overline{I}^*} \langle \tilde{\alpha}, x_0 \rangle = \langle \alpha_0, x_0 \rangle.
\]

Because \( \alpha = \lambda \alpha_0 \), some \( \lambda > 0 \), it follows that

\[
\mathcal{H} = \{ x : \langle \alpha, x \rangle = t \} = \left\{ x : \langle \alpha_0, x \rangle = \frac{t}{\lambda} \right\},
\]

and since \( x_0 \in \mathcal{H} \), it follows from (3.3) that, in fact, \( t \lambda^{-1} = a \).
Step 1. There exists \( x_0 \in \text{cl} A \cap \mathcal{A} \). Suppose, to the contrary, that \( a = \inf_{x \in A} I(x) \) is not attained at any point of \( \text{cl} A \). Then there exists an unbounded sequence \( \{x_i\}_{i \in \mathbb{Z}_+} \subset A - \{0\} \) such that \( \lim_{i \to \infty} I(x_i) = a \). Now \( \{x_i\} \subset A - \{0\} \Rightarrow \{x_i\} \subset (\text{cone } \epsilon_d^d)^c - \{0\} \), by (H2). Hence \( \{x_i/\|x_i\|\} \subset (\text{cone } \epsilon_d^d)^c \cap S^{d-1} \). Since \( (\text{cone } \epsilon_d^d)^c \cap S^{d-1} \) is compact, it follows that a subsequence \( \{x_{i_j}/\|x_{i_j}\|\}_{j \in \mathbb{Z}_+} \) converges to some element \( z \in (\text{cone } \epsilon_d^d)^c \cap S^{d-1} \). Moreover, from the lower semicontinuity of \( \hat{I}, \hat{I}(z) \leq \liminf_{j \to \infty} \hat{I}(x_{i_j}/\|x_{i_j}\|) = \liminf_{j \to \infty} \|x_{i_j}\|^{-1} \hat{I}(x_{i_j}) = 0 \). Hence \( \hat{I}(z) = 0 \) for some \( z \in (\text{cone } \epsilon_d^d)^c - \{0\} \), a contradiction to Lemma 3.1.

Step 2. \( \partial \hat{I}(x_0) \neq \emptyset \). It suffices to show \( \hat{I}(x_0; y) \geq 0 \) for some \( y \in \text{ri } \mathcal{D} - x_0 \) [Rockafellar (1970), Theorem 23.3]. This can be verified when \( x_0 \in \text{ri } \mathcal{D} \) by setting \( y = x_0 \), and when \( x_0 \in \text{cl } \mathcal{D} - \text{ri } \mathcal{D} \) by setting \( y = z - x_0 \), where \( z \in A \cap \text{ri } \mathcal{D} \).

Step 3. In (3.2), \( \alpha = \lambda \alpha_0 \), where \( \lambda > 0 \) and \( \alpha_0 \in \mathcal{A}^{\ast} \) satisfies \( I(x_0) = \langle \alpha_0, x_0 \rangle \). Note \( x_0 \in \text{cl } A \cap \mathcal{A} \Rightarrow x_0 \in \mathcal{H} \). Hence \( \langle \alpha, x - x_0 \rangle \leq 0, \forall x \in \mathcal{A} \). That is, \( \alpha \in \mathcal{A}^{\ast} \). Since \( \partial \hat{I}(x_0) \neq \emptyset \) (Step 2), it follows that \( \alpha \in \text{cl } K \), where \( K \) is the cone generated by \( \partial \hat{I}(x_0) \) [Rockafellar (1970), Theorem 23.7].

Now
\[
\text{cl } K = K \cup \text{rc } \partial \hat{I}(x_0) = K \cup \mathcal{A}^{\ast}(x_0)
\]
[Rockafellar (1970), Theorem 9.6 and the proof of Theorem 25.6]. Hence \( \alpha \in \text{cl } K \) \( \Rightarrow \) either \( \alpha \in K \) or \( \alpha \in \mathcal{A}^{\ast}(x_0) \).

However, \( \alpha \notin \mathcal{A}^{\ast}(x_0) \). Otherwise we would have \( \langle \alpha, x - x_0 \rangle \leq 0, \forall x \in \text{dom } \hat{I} \), implying \( \langle \alpha, x \rangle \leq 0, \forall x \in \text{dom } \hat{I} \), because \( \text{dom } \hat{I} \) is a cone. Since \( A \cap \{x: \langle \alpha, x \rangle \geq t \} \), where \( t \geq 0 \), it would follow that \( A \cap \text{ri(dom } \hat{I}) = \emptyset \). Then \( \text{ri(dom } \hat{I}) = \text{ri } \mathcal{D} \Rightarrow A \cap \text{ri } \mathcal{D} = \emptyset \), contrary to hypothesis.

We conclude \( \alpha \notin K \). Therefore, \( \alpha = \lambda \alpha_0 \) for some \( \lambda > 0 \) and some \( \alpha_0 \in \partial \hat{I}(x_0) \). Then, using the definition of \( \hat{I} \), we have \( \alpha_0 \in \partial \delta(x_0|\mathcal{A}^{\ast}) \), where \( \delta(\cdot|\mathcal{A}^{\ast}) \) is the indicator function on \( \mathcal{A}^{\ast} \). By Rockafellar ([1970], Theorem 23.5), it follows that \( \langle \alpha, x_0 \rangle - \delta(\alpha|\mathcal{A}^{\ast}) \) achieves its maximum in \( \tilde{\alpha} \) when \( \tilde{\alpha} = \alpha_0 \) and therefore \( \hat{I}(x_0) = \sup_{\tilde{\alpha} \in \mathcal{A}^{\ast}} \langle \tilde{\alpha}, x_0 \rangle = \langle \alpha_0, x_0 \rangle \). \( \square \)

4. Proofs of Theorems 2.1 and 2.2. For \( \alpha \in \mathbb{R}^d \) and \( a \in \mathbb{R} \), let
\[
\mathcal{H}(\alpha, a) = \text{the open half-space } \{x: \langle \alpha, x \rangle > a \}.
\]

The following lemma describes the first passage times of a sequence \( Y_1, Y_2, \ldots \) into half-spaces \( \mathcal{H}(\alpha, a) \), where \( \alpha \in \mathcal{A}^{\ast} \).

**Lemma 4.1.** Let \( \alpha \in \mathcal{A}^{\ast} \) and suppose \( \Lambda_N(\alpha) < \infty \). Then
\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(m\mathcal{H}(\alpha, a)) \leq km\} \leq -a
\]
holds for any integer \( k \).
Proof. Let \( r < 1 \), and define a sequence of measures \( \{ \tilde{\mu}_n \} \) by

\[
\tilde{\mu}_n(F) = \int_F \exp(n \langle ra, z \rangle) \, d\mu_n(z)
\]

for all measurable sets \( F \).

From this definition it follows that

\[
P\{ Y_n \in m\mathcal{H}(\alpha, a) \}
= \int_{(m/n)\mathcal{H}(\alpha, a)} \exp(-n \langle ra, z \rangle) \, d\tilde{\mu}_n(z)
= \exp(-m(ra)) \int_{(m/n)\mathcal{H}(\alpha, a)} \exp(mr(a - \langle \alpha, (n/m)z \rangle)) \, d\tilde{\mu}_n(z).
\]

Since \( \mathcal{H}(\alpha, a) = \{ x: \langle \alpha, x \rangle > a \} \), the integrand in this last expression is never greater than 1. As a consequence, we obtain

\[
P\{ Y_n \in m\mathcal{H}(\alpha, a) \} \leq \exp(-m(ra))\tilde{\mu}_n(\mathbb{R}^d).
\]

Then

\[
P\{ T_N(m\mathcal{H}(\alpha, a)) \leq km \} \leq \exp(-m(ra))\{ \tilde{\mu}_N(\mathbb{R}^d) + \cdots + \tilde{\mu}_km(\mathbb{R}^d) \}
\]

and, therefore,

\[
\limsup_{m \to \infty} \frac{1}{m} \log P\{ T_N(m\mathcal{H}(\alpha, a)) \leq km \}
\leq -ra + \limsup_{m \to \infty} \max_{N \leq n \leq km} \left\{ \frac{1}{m} \log \tilde{\mu}_n(\mathbb{R}^d) \right\}.
\]

(4.2)

To deal with the last term on the right, apply Theorem 2.2 of Varadhan (1984) \[\text{(This holds even if the level sets of} \ I \ \text{are not compact. The fact that the given function is not bounded is handled by a simple truncation, as in Theorem 4.3.1 of Dembo and Zeitouni (1993). Since} \ \Lambda_N(\alpha) < \infty, \ \text{Varadhan's theorem yields}\]

\[
\limsup_{n \to \infty} \frac{1}{n} \log \tilde{\mu}_n(\mathbb{R}^d) \equiv \limsup_{n \to \infty} \frac{1}{n} \log \int_{\mathbb{R}^d} \exp(n \langle ra, z \rangle) \, d\mu_n(z)
\leq \sup_{z \in \mathbb{R}^d} \{ \langle ra, z \rangle - I(z) \} = I^*(ra) \leq 0,
\]

since \( 0 \in \mathcal{I}_0^* I^* \) and \( \alpha \in \mathcal{I}_0^* I^* \Rightarrow ra \in \mathcal{I}_0 I^* \), and it follows that

\[
\limsup_{m \to \infty} \max_{\sqrt{m} \leq n \leq km} \left\{ \frac{1}{m} \log \tilde{\mu}_n(\mathbb{R}^d) \right\}
\leq \limsup_{m \to \infty} \max_{\sqrt{m} \leq n \leq km} \left\{ \frac{k}{n} \log \tilde{\mu}_n(\mathbb{R}^d) \right\} \lor 0 = 0.
\]

(4.3)

It remains to prove a similar inequality for

\[
\limsup_{m \to \infty} \max_{N \leq n \leq \sqrt{m}} \left\{ \frac{1}{m} \log \tilde{\mu}_n(\mathbb{R}^d) \right\}.
\]
To this end, observe that for \( n \geq N \),
\[
\frac{1}{n} \log \mu_n(\mathbb{R}^d) \leq \Lambda_N(r\alpha) \leq r\Lambda_N(\alpha) < \infty.
\]

Hence
\[
(4.4) \quad \limsup_{m \to \infty} \max_{N \leq n \leq \sqrt{m}} \left\{ \frac{1}{m} \log \mu_n(\mathbb{R}^d) \right\} \leq \limsup_{m \to \infty} \frac{1}{\sqrt{m}} \left\{ r\Lambda_N(\alpha) \vee 0 \right\} = 0.
\]

Upon substituting (4.3) and (4.4) into (4.2) we obtain
\[
\limsup_{m \to \infty} \frac{1}{m} \log P\{T_N(m\mathcal{M}(\alpha, a)) \leq km\} \leq -ra
\]
and the lemma follows by letting \( r < 1 \).

**Proof of Theorem 2.1.** *Upper bound.* Let \( a = \inf_{x \in A} \tilde{I}(x) \). By Lemma 3.2, there exists \( \alpha_0 \in \mathcal{A}_0 I^* \) such that
\[
(4.5) \quad A \subset \mathcal{M}(\alpha_0, a) = \{x: (\alpha_0, x) > a\}.
\]

Because \( \alpha_0 \in \mathcal{A}_0 I^* \), hypothesis (H1) implies that \( \Lambda(\alpha_0) < \infty \). That is, \( \Lambda_N(\alpha_0) < \infty \) for suitably large \( N \). Consequently, the hypotheses of Lemma 4.1 are satisfied for some constant \( N \). Hence, by Lemma 4.1 and (4.5) we obtain
\[
(4.6) \quad \limsup_{m \to \infty} \frac{1}{m} \log P\{T_N(mA) \leq km\} \leq -a
\]
for any integer \( k \).

To complete the proof of the upper bound, we need to show that for some \( k \),
\[
(4.7) \quad \limsup_{m \to \infty} \frac{1}{m} \log P\{km < T_N(mA) < \infty\} \leq -a.
\]

Here we observe that by (H2), \( (m/n)A \subset (m/n)(\text{cone } \mathcal{E}_\delta)^C = (\text{cone } \mathcal{E}_\delta)^C \). Hence
\[
\begin{align*}
P\{km < T_N(mA) < \infty\} &\leq \sum_{n > km} P\{Y_n \in mA\} \\
&\leq \sum_{n > km} P\left\{ \frac{Y_n}{n} \in (\text{cone } \mathcal{E}_\delta)^C \right\}.
\end{align*}
\]

By the large deviation principle, the central tendency of \( Y_n/n \) is near the set \( \mathcal{A}_0 I \), and since \( \mathcal{A}_0 I \) and \( (\text{cone } \mathcal{E}_\delta)^C \) are disjoint, the central tendency is therefore directed away from the set \( (\text{cone } \mathcal{E}_\delta)^C \). For this reason, the terms in the above sum are becoming increasingly small. More precisely, by the large deviation upper bound,
\[
\limsup_{n \to \infty} \frac{1}{n} \log P\left\{ \frac{Y_n}{n} \in (\text{cone } \mathcal{E}_\delta)^C \right\} \leq - \inf_{x \in (\text{cone } \mathcal{E}_\delta)^C} \tilde{I}(x).
\]

Also, \( \inf_{x \in (\text{cone } \mathcal{E}_\delta)^C} \tilde{I}(x) > 0 \). [Since \( \tilde{I}(0) \neq 0 \), by (H2), we would otherwise have \( J(x) = \inf_{\tau > 0} \tau^{-1} \tilde{I}(\tau x) \) equal to zero at some point of \( (\text{cone } \mathcal{E}_\delta)^C \cap S^{d-1} \), and]
since $I = \text{cl} J$, this would contradict Lemma 3.1.] Therefore, for some integer $k$,
\[
\limsup_{n \to \infty} \frac{1}{n} \log P \left\{ \frac{Y_n}{n} \in (\text{cone } \ell_\delta)^c \right\} \leq -\frac{2a}{k}.
\]
That is, for sufficiently large $n_0$, $P \left\{ \frac{Y_n}{n} \in (\text{cone } \ell_\delta)^c \right\} \leq e^{-na/k}$, $\forall n \geq n_0$. Substituting this last inequality into (4.8) when $m \geq n_0/k$, we obtain
\[
P \{ km < T_N(mA) < \infty \} \leq \sum_{n > km} e^{-na/k} \leq \text{const.} \cdot e^{-ma},
\]
establishing (4.7) and hence the required upper bound.

**Lower bound.** Let $\tau > 0$ and $x \in A$, and for each integer $m$, let $n(m) = \lceil m/\tau \rceil$. Since $A$ is an open set, there exists an $\varepsilon > 0$ such that $B_{x, 2\varepsilon} \subset A$, and for sufficiently large $m$,
\[
P \{ T_N(mA) < \infty \} \geq P \{ Y_{n(m)} \in mA \}
\geq P \left\{ \frac{Y_{n(m)}}{n(m)} \in \frac{m}{n(m)} B_{x, 2\varepsilon} \right\} \geq P \left\{ \frac{Y_{n(m)}}{n(m)} \in \tau B_{x, \varepsilon} \right\}.
\]
Hence, by the large deviation lower bound
\[
\liminf_{m \to \infty} \frac{1}{m} \log P \{ T_N(mA) < \infty \} \geq \lim_{n \to \infty} \left( \frac{n(m)}{m} \right) \liminf_{n \to \infty} \frac{1}{n} \log P \left\{ \frac{Y_n}{n} \in \tau B_{x, \varepsilon} \right\}
\geq -\tau^{-1} \inf_{z \in \tau B_{x, \varepsilon}} I(z) \geq -\tau^{-1} I(\tau x).
\]
Taking the supremum over all $\tau > 0$ and $x \in A \cap \partial \Omega$, we obtain
\[
\liminf_{m \to \infty} \frac{1}{m} \log P \{ T_N(mA) < \infty \} \geq -\inf_{\tau > 0} \left\{ \inf_{x \in A \cap \partial \Omega} \tau^{-1} I(\tau x) \right\}
= -\inf_{x \in A \cap \partial \Omega} \hat{I}(x) [\text{Rockafellar (1970), Theorem 13.5}]
\]
for any integer $N$.

Now suppose $\hat{x} \in A \cap \partial \Omega$, where $\partial \Omega = \text{cl} \Omega - \text{ri} \Omega$ is the relative boundary of the domain of $\hat{I}$. On the one hand, since $\hat{x}$ is an element of the open set $A$, $B_{\hat{x}, \varepsilon} \subset A$ for some $\varepsilon > 0$. On the other hand, since $\hat{x} \in \partial \Omega$, the set $B_{\hat{x}, \varepsilon}$ intersects $\text{ri} \Omega$. That is, $B_{\hat{x}, \varepsilon} \cap \text{ri} \Omega \neq \emptyset$. Let $\gamma$ be a line segment which joins $\hat{x}$ to some $y \in B_{\hat{x}, \varepsilon} \cap \text{ri} \Omega$. Then $\gamma - \hat{x} \subset B_{\hat{x}, \varepsilon} \cap \text{ri} \Omega$ [Rockafellar (1970), Theorem 6.1], implying $\gamma - \hat{x} \subset A \cap \text{ri} \Omega$. Consequently, $\inf_{x \in A \cap \text{ri} \Omega} \hat{I}(x) \leq \inf_{x \in \gamma - \hat{x}} \hat{I}(x) \leq \hat{I}(\hat{x})$, where the last step follows from the convexity of $\hat{I}$. As this inequality holds for each $\hat{x} \in A \cap \partial \Omega$, we conclude $\inf_{x \in A \cap \text{ri} \Omega} \hat{I}(x) = \inf_{x \in A} \hat{I}(x)$, and this, together with (4.9), establishes the lower bound and hence the theorem. $\square$
Proof of Theorem 2.2. The proofs of the lower bound and (4.7) of the upper bound of Theorem 2.1 apply equally well here. Thus we have only to establish a corresponding result to (4.6), namely, for all \( N \geq N_0 \),

\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(mA) \leq km\} \leq - \inf_{x \in \text{cl} A} \bar{I}(x)
\]

holds for any integer \( k \).

Case 1. \( A \) is bounded. Let \( a < \inf_{x \in \text{cl} A} \bar{I}(x) \). Then for \( x \in \text{cl} A \) we have \( \sup_{\alpha \in J^*} \langle \alpha, x \rangle \equiv \bar{I}(x) > a \). For each such \( x \), we can therefore choose \( \alpha_x \in J^* \) such that \( \langle \alpha_x, x \rangle > a \), giving

\[
x \in \mathcal{H}(\alpha_x, a) = \{ z : \langle \alpha_x, z \rangle > a \}.
\]

Note \( x \in \mathcal{H}(\alpha_x, a) \Rightarrow \{ \mathcal{H}(\alpha_x, a) \}_{x \in \text{cl} A} \) is an open cover for \( \text{cl} A \), and since \( \text{cl} A \) is compact, there exists a finite subcover: \( \mathcal{H}(\alpha_{x_1}, a), \ldots, \mathcal{H}(\alpha_{x_l}, a) \). Then

\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(mA) \leq km\}
\]

(4.11)

\[
\leq \max_{1 \leq i \leq l} \left\{ \limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(m\mathcal{H}(\alpha_{x_i}, a)) \leq km\} \right\}.
\]

Consider

\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(m\mathcal{H}(\alpha_{x_i}, a)) \leq km\}
\]

for any given \( i \). Since \( \alpha_{x_i} \in J^* \), hypothesis (H1') implies that \( \Lambda_N(\alpha_{x_i}) < \infty \) for all \( N \geq N_0 \). Hence, from Lemma 4.1 we obtain

\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(m\mathcal{H}(\alpha_{x_i}, a)) \leq km\} \leq -a \quad \text{for all } N \geq N_0.
\]

(4.12)

Substituting (4.12) into (4.11) yields

\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(mA) \leq km\} \leq -a \quad \text{for all } N \geq N_0.
\]

Then (4.10) is established by letting \( a = \inf_{x \in \text{cl} A} \bar{I}(x) \).

Case 2. \( A \) is unbounded. Let \( K < \infty \) and \( a = \inf_{x \in \text{cl} A} \bar{I}(x) \wedge K \). We need to show that independent of \( K \),

\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(mA) \leq km\} \leq -a \quad \text{for all } N \geq N_0.
\]

(4.13)

If \( R > 0 \), then the event \( \{T_N(mA) \leq km\} \) is the union of the events \( \{T_N(m(B_{0,R} \cap A)) \leq km\} \) and \( \{T_N(m(B_{0,R}^c \cap A)) \leq km\} \). Hence,

\[
\limsup_{m \to \infty} m^{-1} \log \mathbb{P}\{T_N(mA) \leq km\}
\]

is the larger of

\[
\limsup_{m \to \infty} \frac{1}{m} \log \mathbb{P}\{T_N(m(B_{0,R} \cap A)) \leq km\}
\]
and
\[ \limsup_{m \to \infty} \frac{1}{m} \log P \{ T_N(m(B_{0,R}^c \cap A)) \leq km \}. \]

Since \( B_{0,R} \cap A \) is bounded,
\[ \limsup_{m \to \infty} \frac{1}{m} \log P \{ T_N(m(B_{0,R} \cap A)) \leq km \} \leq -\inf_{x \in \partial(B_{0,R} \cap A)} \tilde{I}(x) \leq -a \quad \text{for all } N \geq N_0 \]
for any positive value of \( R \). Thus, to establish (4.13) it suffices to show that, for some \( R > 0 \),
\[ \limsup_{m \to \infty} \frac{1}{m} \log P \{ T_N(m(B_{0,R}^c \cap A)) \leq km \} \leq -a \quad \text{for all } N \geq N_0. \] (4.14)

We now demonstrate that (4.14) holds for sufficiently large \( R \). First, observe that the set \( L^c \) is compact. [Since \( L^c \) is positively homogeneous, it would otherwise approach zero at some point of \( \partial \tilde{I} \cap (\text{cone } \mathcal{L})^c \). Using the lower semicontinuity of \( L^c \), this would give a contradiction to Lemma 3.1.] Thus, there exists a large ball \( B_{0,R} \) which contains \( L^c \). Then
\[ B_{0,R}^c \subset (\mathcal{L})^c \cup \text{cone } \mathcal{L}, \]
implying
\[ B_{0,R}^c \cap (\text{cone } \mathcal{L})^c \subset (\mathcal{L})^c. \]
Next, observe that from the very definitions of \( \mathcal{H}(\cdot, a) \) and \( \tilde{I} \) we have
\[ (\mathcal{L})^c = \bigcup_{a \in \mathcal{L}^c} \mathcal{H}(a, a). \]
From the last two equations it follows that
\[ B_{0,R}^c \cap (\text{cone } \mathcal{L})^c \subset \bigcup_{a \in \mathcal{L}^c} \mathcal{H}(a, a). \]
Therefore \( \{ \mathcal{H}(a, a) \}_{a \in \mathcal{L}^c} \) is an open cover for \( B_{0,R}^c \cap (\text{cone } \mathcal{L})^c \), hence also an open cover for the smaller set \( \partial B_{0,R}^c \cap (\text{cone } \mathcal{L})^c \), where \( \partial B_{0,R}^c \) is the boundary of \( B_{0,R}^c \). As \( \partial B_{0,R}^c \cap (\text{cone } \mathcal{L})^c \) is compact, there exists a finite subcollection \( \{ \mathcal{H}(\cdot, a_1), \ldots, \mathcal{H}(\cdot, a_l) \} \) which covers \( \partial B_{0,R}^c \cap (\text{cone } \mathcal{L})^c \), and from the definition of \( \mathcal{H}(\cdot, a) \) it is evident that this collection \( \{ \mathcal{H}(\cdot, a_1), \ldots, \mathcal{H}(\cdot, a_l) \} \) covers the larger set \( B_{0,R}^c \cap (\text{cone } \mathcal{L})^c \). Therefore, in view of (H2) we obtain
\[ \bigcup_{i=1}^l \mathcal{H}(a_i, a) \supset B_{0,R}^c \cap (\text{cone } \mathcal{L})^c \supset B_{0,R}^c \cap A. \]
It follows that
\[
\limsup_{m \to \infty} \frac{1}{m} \log P\{T_N(m(B_0^c \cap A)) \leq km\}
\leq \max_{1 \leq i \leq l} \left( \limsup_{m \to \infty} \frac{1}{m} \log P\{T_N(mH(a_i, a)) \leq km\} \right).
\]

(4.15)

To determine the quantity on the right-hand side of (4.15), apply Lemma 4.1. Since \(a_i \in \mathcal{H}^*\), from (H1') we have \(\Lambda_N(a_i) < \infty\) for all \(N \geq N_0\). Hence, by Lemma 4.1,
\[
\limsup_{m \to \infty} \frac{1}{m} \log P\{T_N(mH(a_i, a)) \leq km\} \leq -a \quad \text{for all } N \geq N_0.
\]
(4.16)

Finally, substituting (4.16) into (4.15) yields (4.14), as required. \(\square\)

5. Some further results. Under additional conditions [sufficient conditions are (i) \(A \subset \text{int} \mathcal{D}\) is open, (ii) \(I = \Lambda^*\), where \(\Lambda\) is differentiable on \(\mathcal{H}^*\), and (iii) \(I\) attains its infimum over \(\text{cl } A\) at a unique point \(x_0\)], then
\[
P\left( \left\| \frac{Y_{T_N(mA)}}{m} - x_0 \right\| > \varepsilon \left| T_N(mA) < \infty \right| \right) \leq e^{-mK(\varepsilon)} \quad \text{for large } m
\]
(5.1)

and
\[
P\left( \left| \frac{T_N(mA)}{m} - \rho \right| > \varepsilon \left| T_N(mA) < \infty \right| \right) \leq e^{-mL(\varepsilon)} \quad \text{for large } m,
\]
(5.2)

where \(\rho\) is a positive constant. Roughly speaking, (5.1) asserts that the place of first passage of \(\{Y_n\}_{n \geq N}\) into \(mA\) is near a point \(mx_0\), while (5.2) asserts that the time of first passage is near \(m\rho\). In particular, the time of first passage grows roughly linearly in \(m\). For details, see Collamore (1996).

Acknowledgments. The author is grateful to Professor P. Ney for his invaluable assistance, to Professor S. Robinson for helpful discussions on convex analysis and to the referee for a careful reading of the original manuscript.

REFERENCES


