LARGE DEVIATIONS FOR PROCESSES WITH INDEPENDENT INCREMENTS (U)

FLORIDA STATE UNIVERSITY
TALLAHASSEE

ISTATISTICS
LYNCH ET AL.

01 OCT 84

UNCLASSIFIED
FSU-STATISTICS-M87-ARO-19387.15-MA

F/G 12/1
NL

1/1
Large Deviations for Processes with Independent Increments

by

James Lynch¹ and Jayaram Sethuraman²

August, 1984

FSU Technical Report No. M687
USARO Technical Report No. D-71

¹Department of Statistics, The Pennsylvania State University. Research supported by the U.S. Army Research Office Grant No. DAAG29-84-K-0007.

²Department of Statistics, Florida State University. Research supported by the U.S. Army Research Office Grant No. DAAG29-82-K-0168.

Key Words: Large deviations, rate function, stationary and independent increments, Dirichlet process.

AMS 1980 Subject Classification Numbers: Primary 60F10, 60J30; Secondary 60E07
Large Deviations for Processes with Independent Increments

James Lynch and Jayaram Sethuraman

Abstract

Let $X$ be a topological space and $F$ denote the Borel $\sigma$-field in $X$. A family of probability measures $\{P_A\}$ is said to obey the large deviation principle (LDP) with rate function $I(\cdot)$ if $P_A(A)$ can be suitably approximated by $\exp\{-\lambda \inf_{x \in A} I(x)\}$ for appropriate sets $A$ in $F$. Here the LDP is studied for probability measures induced by stochastic processes with stationary and independent increments which have no Gaussian component. It is assumed that the moment generating function of the increments exists and thus the sample paths of such stochastic processes lie in the space of functions of bounded variation. The LDP for such processes is obtained under the weak*-topology. This covers a case which was ruled out in the earlier work of Varadhan (1966). As applications, the large deviation principle for the Poisson, Gamma and Dirichlet processes are obtained.
1. Introduction.

Let $X$ be a topological space and $F$ denote the Borel $\sigma$-field in $X$. Let $\{P_\lambda\}$ be a family of probability measures on $(X,F)$. The family $\{P_\lambda\}$ is said to obey the large deviation principle (LDP) (for a more precise definition see Section 2) with rate function $I(\cdot)$ if $P_\lambda(A)$ can be approximated by $\exp \{-\lambda \inf_{x \in A} I(x)\}$ for appropriate subsets $A$ in $F$.

Important examples of the LDP include the cases where $P_\lambda$ ($\lambda$ a positive integer) is either (i) the probability measure induced by the average of $\lambda$ i.i.d. random variables (see Chernoff, 1952; Bahadur and Zabell, 1979; Varadhan, 1983) or (ii) the probability measure of the empirical distribution of $\lambda$ i.i.d. random variables (Groeneboom, Oosterhoff and Ruymgaart, 1979). In an important paper, Ellis (1984) has elegantly shown how to establish the LDP when $X = \mathbb{R}^k$, solely in terms of the moment generating functions of $P_\lambda$. Further examples may be found in the recent surveys on large deviations by Azencott (1980) and by Varadhan (1983).

The establishment of the LDP has had important implications in various areas in statistics. It has been used to obtain the asymptotic efficiencies of tests and estimates (Chernoff, 1952; Bahadur, 1960a,b, 1967 and 1971) and to obtain the asymptotic behavior of functional integrals associated with solutions of stochastic integrals (Varadhan, 1966 and 1983). It appears in the evaluation of the 'free' energy in statistical mechanics (Lanford, 1973; Ruelle, 1969). It is also intimately related to certain types of laws of large numbers (Shepp, 1964; Erdös and Rényi, 1970).
In this paper we study the LDP for a stochastic process with stationary independent increments with no Gaussian component and obtain complete results. The space $X$ that is appropriate here is $BV[0,1]$, the space of functions of bounded variation and the topology is that of weak*-convergence. Varadhan (1966) studied the LDP for similar processes with possible Gaussian components but satisfying the condition $J(a)/|a| \to 0$ as $|a| \to \infty$ where $J(\cdot)$ is as defined in (3.2) and is the rate function based on the distribution of the increments. Varadhan (1966) used the space $D[0,1]$ and the topology of uniform convergence. However, the condition $J(a)/|a| \to 0$ is violated for many processes of interest including the Gamma process. We illustrate our LDP results for this process and a related process called the Dirichlet process.

The organization of this paper is as follows: Preliminary definitions and general results on the LDP, which are used in later sections, are given in Section 2. A rate function on $M[0,1]$, the space of finite measures on $[0,1]$, is defined and several theorems concerning this rate function are proved in Section 3. In Section 4, the LDP is established for stochastic processes with stationary and positive independent increments which are considered as elements of $M[0,1]$. In Section 5, the general LDP results are given for stochastic processes with stationary independent increments and no Gaussian component which are considered as elements of $BV[0,1]$, the space of functions of bounded variation. The final section, Section 6, is devoted to applications to the Poisson, Gamma and Dirichlet processes.
2. Definitions and General Results.

Let $X$ be a topological space and $F$ be the Borel σ-field in $X$. Let $\{P_\lambda\}$ be a family of probability measures on $(X,F)$. The following definitions which are slight variants of those of Varadhan (1983) allow us to state many large deviation results in concise form.

Definition 2.1. A function $I(\cdot)$ on $X$ is said to be a regular rate function if

1. $0 \leq I(x) < \infty$,
2. $I(\cdot)$ is lower semi-continuous (lsc), and
3. for each $c < \infty$, $\Gamma_c = \{x : I(x) \leq c\}$ is compact.

For any subset $A$ of $X$, define

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

Definition 2.2. The measures $\{P_\lambda\}$ satisfy the large deviation principle (LDP or LD principle) with rate function $I(\cdot)$ if

1. $I(\cdot)$ is a regular rate function,
2. for each closed set $F$,
\[ \lim_{\lambda \to \infty} \frac{1}{\lambda} \log P_\lambda(F) \leq -I(F), \]
3. for each open set $G$,
\[ \lim_{\lambda \to \infty} \frac{1}{\lambda} \log P_\lambda(G) \geq -I(G), \]

where here and throughout the remainder of this paper the limits are as $\lambda \to \infty$. 
Definition 2.3. The measures \( \{P_\lambda\} \) satisfy the weak large deviation principle (WLDP or the weak LD principle) with rate function \( I(\cdot) \) if (2.5) and (2.7) of Definition 2.3 together with (2.8) below are satisfied:

\[
\lim_{\lambda \to \infty} \frac{1}{\lambda} \log P_\lambda(K) \leq -I(K).
\]

(2.8) for each compact set \( K \),

\[
\lim_{\lambda \to \infty} \frac{1}{\lambda} \log P_\lambda(K) \leq -I(K).
\]

Definition 2.4. The measures \( \{P_\lambda\} \) are large deviation tight (LD tight) if, for each \( M > 0 \), there exists a compact set \( K_M \) such that

\[
\lim_{\lambda \to \infty} \frac{1}{\lambda} \log P_\lambda(K_M^c) \leq -M.
\]

The following lemma shows the usefulness of LD tightness.

Lemma 2.5. Let \( \{P_\lambda\} \) be LD tight and satisfy the WLDP. Then it satisfies the LDP.

Proof. Let \( C \) be closed and let \( \ell < I(C) \). Let \( M > \ell \) and choose a compact set \( K_M \) to satisfy (2.9). Then \( C \cap K_M \) is compact and \( P_\lambda(C) < P_\lambda(C \cap K_M) + P_\lambda(K_M^c) \).

Thus,

\[
\lim_{\lambda \to \infty} \frac{1}{\lambda} \log P_\lambda(C) \leq \min \{I(C \cap K_M), M\} \leq -\ell.
\]

Many interesting applications in large deviations occur when \( X \) is a Polish space, that is a separable complete metric space. Accordingly, we will assume that all spaces we consider in the rest of this paper to be Polish spaces, and the corresponding \( \sigma \)-fields to be Borel \( \sigma \)-fields.
For sequences of probability measures on a Polish space the following lemma, which will not be referred to in the remainder of the paper, shows that the LDP implies LD tightness. Consequently, the LDP is equivalent to the WLDP and LD tightness along subsequences.

**Lemma 2.6.** If \( \{ P_\lambda \} \) is a sequence of probability measures which satisfies the LDP, then \( \{ P_\lambda \} \) is LD tight.

**Proof.** Let \( \{ x_i, i=1,2,\ldots \} \) be a countable dense set in \( X \). For any \( \delta > 0 \), let \( A_i(\delta) \) be the open sphere of radius \( \delta \) around \( x_i \). Then \( \bigcup A_i(1/k) = X \) for \( k = 1,2,\ldots \). Fix \( M > 0 \) and an integer \( k \). Consider the compact set \( \Gamma_{2kM} = \{ x : I(x) < 2kM \} \). There exists a finite open covering

\[
A(k) = \bigcup_{i=1}^{I_k} A_i(1/k)
\]

of \( \Gamma_{2kM} \). Thus, from (2.6)

\[
\lim \lambda^{-1} P_\lambda (A^c(k)) \leq - I(A^c(k)) \leq - I(\Gamma_{2kM}) \leq - 2kM.
\]

Since we are considering only sequences \( \{ \lambda \} \) we can find a larger finite union

\[
B(k) = \bigcup_{i=1}^{J_k} A_i(1/k)
\]

with \( J_k \geq I_k \) such that

\[
P_\lambda (B^c(k)) \leq e^{\lambda M_k}
\]
for all \( \lambda \). The set \( K = \bigcap_{k=1}^{\infty} B(k) \), where \( B(k) \) is the closure of \( B(k) \), is totally bounded and closed, and hence is compact. Furthermore

\[
P_\lambda(K^c) \leq \sum_{k=1}^{\infty} P_\lambda(B^c(k)) \leq e^{\lambda M}/(1-e^{-\lambda M})
\]

for all \( \lambda \), where \( \lambda_0 \) is the smallest index in the sequence \( \{\lambda\} \). This completes the proof of Lemma (2.6).

Let \( \{P_\lambda^i\} \) be a family of probability measures on a Polish space \( X^i, i = 1,2 \). Let \( P_\lambda^i = P_\lambda^1 \times P_\lambda^2 \) be the product measure on the product space \( X = X^1 \times X^2 \). We will now investigate whether LD properties of marginal measures carry over to the product measures.

**Lemma 2.7.** If \( \{P_\lambda^i\} \) is LD tight for \( i = 1,2 \), then \( \{P_\lambda\} \) is LD tight.

**Proof.** Obvious.

**Lemma 2.8.** Let \( \{P_\lambda^i\} \) satisfy the WLD with rate function \( I^i(x^1_i) \), \( i = 1,2 \). Then \( \{P_\lambda\} \) satisfies the WLD with rate function \( I(x^1_i, x^2_2) = I^1(x^1_1) + I^2(x^2_2) \).

**Proof.** It is easy to check the regularity of \( I(x^1_i, x^2_2) \) from the regularity of \( I^1(x^1_i) \) and \( I^2(x^2_2) \). Let \( K \subset X \) be compact and let \( \ell \leq I(K) \). For each
(x_1, x_2) \in K, since I(\cdot) is lsc, there are open sets O^i_{x_i} in X^i
containing x_i, i = 1, 2, such that

\begin{equation}
\inf \{I(y_1, y_2) : (y_1, y_2) \in O^1_{x_1} \times O^2_{x_2} \} > \ell.
\end{equation}

Furthermore, since X^i is Polish, we can find open subsets N^i_{x_i} of
O^i_{x_i} such that x_i \in N^i_{x_i} and N^i_{x_i} \subset O^i_{x_i}.

Consider the open covering \bigcup_{i=1}^{M} N^i_{x_1} \times N^i_{x_2} of K. We
can extract a finite subcovering \bigcup_{i=1}^{M} N^i_{x_1} \times N^i_{x_2} of K.

Let K^1 and K^2 be the projections of K in X^1 and X^2. Then K^1 and
M^i_{x_1} \cap N^i_{x_2} are compact, m = 1, \ldots, M and i = 1, 2.

Furthermore, K \subset \bigcup_{i=1}^{M} M^i_{x_1} \times M^i_{x_2}.

Thus, since \( M^i_{x_1} \) is compact and \{P^i_{\lambda}\} satisfies the
WLDP,

\begin{equation}
\lim_{\lambda} \frac{1}{\lambda} \log P^i_{\lambda}(K) \leq -\min \{I^1(M^1_{x_1, m}) + I^2(M^2_{x_2, m}) \}
\end{equation}

in view of (2.11). This proves (2.8).

Let 0 be an open set in X. Fix \( \varepsilon > 0 \) and choose (x_1, x_2) so that
I(x_1, x_2) < I(0) + \varepsilon. There exist open sets O_{x_i} in X^i around x_i,
i = 1, 2 such that O_{x_1} \times O_{x_2} \subset 0. Thus

\begin{equation}
\lim_{\lambda} \frac{1}{\lambda} \log P^i_{\lambda}(0) \geq \sum_{i=1}^{2} \lim_{\lambda} \frac{1}{\lambda} \log P^i_{\lambda}(O_{x_i})
\geq -I(x_1, x_2) \geq -I(0) - \varepsilon.
\end{equation}

Since \( \varepsilon > 0 \) is arbitrary, this establishes (2.7) which completes the proof
of Lemma 2.7.

\[\square\]
The following corollary follows from Lemmas 2.5, 2.7 and 2.8.

**Corollary 2.9.** Let \( \{P^i_\lambda\} \) be LD tight and satisfy the WLDP, \( i = 1,2 \).

Then \( P_\lambda = P^1_\lambda \times P^2_\lambda \) satisfies the LDP.

Two important and immediate derivatives of the LDP are the contraction principle, which is used later in this paper, and the asymptotic expression for certain integrals. These are stated below. For proofs see Varadhan (1966, 1983).

Let \( \{P^i_\lambda\} \) satisfy the LDP with rate function \( I(x) \). Let \( h \) be a continuous map from \( X \) into another Polish space \( Y \), and let

\[ Q^i_\lambda = P^i_\lambda h^{-1}. \]

**Contraction Principle.** The measures \( \{Q^i_\lambda\} \) satisfy the LDP with rate function

\[
K(y) = \inf_{x: h(x) = y} I(x).
\]

**Asymptotic expression for certain integrals.** Let \( F \) be a bounded real valued continuous function on \( X \). Then

\[
\frac{1}{\lambda} \log \int \exp(\lambda F(x)) \, dP^i_\lambda(x) = \sup_x [F(x) - I(x)].
\]

It is interesting to note the definition of the LDP and LD tightness together with their consequences, namely (2.12) and (2.13) above, run parallel to the definition of weak convergence and tightness (see Billingsley 1968) together with their consequences, namely the continuous mapping principle and convergence of integrals of bounded continuous functions.
3. The Rate Function $I(f)$ on $M[0,1]$. 

Let $X$ be a real valued random variable and let

$$\phi(\theta) = E(e^{\theta X}) < \infty$$

for $|\theta| < n$ where $n > 0$. Let $\psi(\theta) = \log \phi(\theta)$.

Define

$$J(a) = J_X(a) = \sup \{at - \psi(t)\},$$

The function $J(a)$ is loosely called the rate function associated with $X$. More precisely, let $P_n$ be the distribution of $(X_1 + \ldots + X_n)/n$ where $X_1, X_2, \ldots$ are i.i.d. copies of $X$. The following is the oldest theorem in large deviation theory and is variously referred to as Cramér's theorem and Chernoff's theorem.

**Theorem 3.1.** (Cramér, 1938; Chernoff, 1952). The distributions $\{P_n\}$ are LD tight and satisfy the LDP with rate function $J(a)$.

The following facts concerning the function $J(a)$ are easy to obtain from its definition in (3.2):

$$0 \leq J(a) \leq \infty, J(\mu) = 0 \text{ where } E(X) = \mu \text{ and } J(a) \to \infty \text{ as } |a| \to \infty.$$  

$$J(a) = \sup \{at - \psi(t)\} \text{ if } a > \mu.$$ 

$$J(a) \text{ is convex.}$$
(3.6) \[ \lim_{a \to a^-} \frac{J(a)}{a} = C_1 \text{ and } \lim_{a \to a^-} \frac{J(a)}{a} = C_2 \text{ exist, where } 0 < C_1, C_2 < \infty. \]

(3.7) The function \( g(b) \) defined by

\[
g(b) = \begin{cases} 
\text{b} J(1/b) & \text{if } 0 < b < \infty \\
C_1 & \text{if } b = 0
\end{cases}
\]

is convex on \([0, \infty)\).

(3.8) If the support of \( X \) is \([0, \infty)\), then \( J(0) < \infty \) if and only if \( P(X=0) > 0 \).

We will now obtain an illustration of the contraction principle which will be used in Section 5 to identify the LD rates. Let \( X = X(1) - X(2) \) where \( X(1) \) and \( X(2) \) are independent non-negative random variables. Under assumption (3.1), the moment generating functions \( \phi(i)(\theta) \) of \( X(i) \) exist in a neighborhood of 0, \( i = 1, 2 \). Let \( \psi(i)(\theta) = \log \phi(i)(\theta) \) and define the rate function \( J(i)(a) \) of \( X(i) \) analogously to (3.2), \( i = 1, 2 \). From Theorem 3.1 and Corollary 2.9, the distributions of the arithmetic means of i.i.d. copies of the bivariate random variable \((X(1), X(2))\) satisfy the LD principle with rate function \( J(1)(x_1) + J(2)(x_2) \). From the contraction principle we obtain the useful result

(3.9) \[ J(a) = \inf_{b} (J(1)(a+b)+J(2)(b)) \]

Let

(3.10) \[ C(i) = \lim_{a \to a^-} \frac{J(i)(a)}{a}, \quad i = 1, 2. \]
We will now show that

(3.11) \[ C_i = C^{(i)}, \quad i = 1, 2 \]

where \( C_1, C_2 \) are as defined in terms of \( J(a) \) in (3.6). Note that \( \psi(\theta) = \psi^{(1)}(\theta) + \psi^{(2)}(-\theta) \) and that \( \psi^{(2)}(\theta) \leq 0 \) for \( \theta < 0 \) since \( X^{(2)} \) is non-negative. Thus

(3.12) \[ a\theta - \psi^{(1)}(\theta) \leq J(a) \leq J^{(1)}(a+b) + J^{(2)}(b) \]

for all \( \theta > 0 \) and all \( b \). From (3.4)

\[ J^{(1)}(a) = \sup_{\theta > 0} [a\theta - \psi^{(1)}(\theta)] \]

for large \( a \). Dividing (3.12) by \( a \) and allowing \( a \) to tend to \( \infty \), we obtain \( C_1 = C^{(1)} \). Similarly \( C_2 = C^{(2)} \).

Let \( M[0,1] \) be the space of finite measures on \([0,1], B\) where \( B \) is the usual Borel \( \sigma \)-field in \([0,1]\). For any element \( f \) in \( M[0,1] \), we define its distribution function \( f(t) \) by letting \( f(0) = 0, f(t) = f([0,t]) \), \( 0 < t \leq 1 \). We also use the same symbol \( f \) to denote both the measure \( f(A) \) and the (extended) distribution function \( f(t) \).

Let \( \alpha \) be a probability measure on \([0,1] \), that is \( \alpha \in M[0,1] \) and \( \alpha(1) = 1 \). Let \( 0 = t_0 < t_1 < \ldots < t_{k-1} = 1 \). Both the collection of points \( \{t_0, t_1, \ldots, t_k\} \) and the collection of intervals \( \{[0,t_1], (t_1,t_2], \ldots, (t_{k-1},1]\} \) will be referred to as the partition \( P \). Let \( \sigma(P) \) be the \( \sigma \)-field generated by the intervals in the partition \( P \). The partitions \( \{P\} \) from a directed set under the partial order \( P' > P \) if
\( \sigma(p') \supset \sigma(P) \). We will be taking limits of functions on \( \{P\} \) and it will always be along directed nets such that \( \sigma(P) \uparrow B \).

Let \( f \in M[0,1] \) and \( P \) be a partition. We define

\[
I_p(f) = \begin{cases} 
\tau J \left( \frac{f(t_i) - f(t_{i-1})}{\alpha(t_i) - \alpha(t_{i-1})} \right) (\alpha(t_i) - \alpha(t_{i-1})) \\
\text{if } \alpha(t_i) - \alpha(t_{i-1}) = 0 \text{ implies } i = 1, \ldots, k,
\end{cases}
\]

wherein \( J(\alpha) \) is the rate function of some non-negative random variable \( X \) satisfying (3.1) and we observe the convention \( 0 \cdot (\text{undefined}) = 0 \) and \( 0 \cdot - = 0 \). The rest of this section is devoted to obtaining many important properties of \( I_p(f) \) which are useful in obtaining the LDP results of Section 4.

Denote the restriction of the measures \( \alpha \) and \( f \) to \( \sigma(P) \) by \( \alpha_p \) and \( f_p \), respectively. We may rewrite the definition in (3.13) by

\[
I_p(f) = \begin{cases} 
\int J \left( \frac{df_p}{d\alpha} \right) \, d\alpha \text{ if } f_p \ll \alpha_p \\
\text{otherwise},
\end{cases}
\]

Let \( f = f_1 + f_2 \) be the Lebesgue decomposition of \( f \) with respect to \( \alpha \), with \( f_1 \ll \alpha \) and \( f_2 \perp \alpha \). Let \( L \subset [0,1] \) be such that \( f_2(L) = f_2([0,1]) \) and \( \alpha(L) = 0 \). Similarly define \( \alpha_1, \alpha_2 \) and \( M \) by \( \alpha = \alpha_1 + \alpha_2,\, \alpha_1 \ll f,\, \alpha_2 \perp f,\, \alpha_2(M) = \alpha_2([0,1]) \) and \( f(M) = 0 \). Let \( f_1 = \frac{d\alpha_1}{d\alpha} \) and \( \dot{f}_1 = \frac{df_1}{df} \). Then

\( f_1 = 1/\dot{f}_1 > 0 \) a.e. on \((LUM)^0 \) with respect to \( f \) and \( \alpha \).
Define

\[ I(f) = \begin{cases} \int J(f) \, d\alpha + C_1 f_2([0,1]) & \text{if supp. } f \subset \text{supp. } \alpha, \\
0 & \text{otherwise} \end{cases} \tag{3.15} \]

where supp. stands for support, \( C_1 \) depends on \( J \) and is as defined in (3.6).

The following theorem relates \( I_p(f) \) to \( I(f) \).

**Theorem 3.2.** As \( c(P) \to \mathcal{B} \),

\[ I_p(f) \to I(f). \tag{3.16} \]

**Proof.** When supp. \( f \) is not contained in supp. \( \alpha \), \( I(f) = \infty \). In this case we do not have \( f_p \ll c_P \) for some \( P \). Then \( I_p(f) = \infty \) and \( I_p(f) = \infty \)

for finer partitions \( P' \). This establishes (3.16) in this case.

From now on assume that supp. \( f \subset \text{supp. } \alpha \). It follows that \( f_p \ll c_P \)

for each \( P \) and that \( \left\{ \frac{df_p}{d\alpha}, c(P) \right\} \) is a martingale. Since \( J(\alpha) \) is convex,

\[ \left\{ J\left( \frac{df_p}{d\alpha}, c(P) \right) \right\} \]

is a sub-martingale. We also have

\[ \frac{df_p}{d\alpha} + f_1 \quad \text{and} \quad J\left( \frac{df_p}{d\alpha} \right) + J(f_1) \quad \text{a.e. } \alpha. \tag{3.17} \]

Under the condition supp. \( f \subset \text{supp. } \alpha \) it may not be true that

\( c_P \ll f_p \). We will use the notation \( \frac{d\alpha_p}{df_p} \) to denote the Radon-Nikodym derivative of \( \alpha_p \), the absolutely continuous part of \( \alpha \) with respect to \( f_p \).

Then \( \left\{ \frac{d\alpha_p}{df_p}, c(P) \right\} \) is a super-martingale under \( f \) and
(3.18) \[
\frac{d\alpha_p}{df_p} + \dot{\alpha}_1 \text{ and } g\left(\frac{d\alpha_p}{df_p}\right) < g(\dot{\alpha}_1) \text{ a.e. } f.
\]

We also have

(3.19) \[
I_p(f) = \int g\left(\frac{d\alpha_p}{df_p}\right) df + \int J\left(\frac{df_p}{d\alpha_p}\right) d\alpha
\]
\[
\frac{d\alpha_p}{df_p} < \mu^{-1}, \quad \frac{df_p}{d\alpha_p} \leq \mu
\]

and

(3.20) \[
I_p(f) \geq \int g\left(\frac{d\alpha_p}{df_p}\right) df
\]

with equality in (3.20) when \(\alpha_p \ll f_p\). Similarly, we can write

(3.21) \[
I(f) = \int g(\dot{\alpha}_1) df + \int J(\dot{\nu}_1) d\alpha
\]
\[
\dot{\alpha}_1 < \mu^{-1}, \quad \dot{\nu}_1 \leq \mu
\]

and

(3.22) \[
I(f) = \int g(\alpha_1) df + J(0) \alpha_2 ([0,1]).
\]
It is possible that $K = 0$ or $J(0) = 0$ or both and so we consider the following cases to complete the proof:

(i) $I(f) = 0$

(ii) $I(f) < 0$ and $f_2([0,1]) = 0$

(iii) $I(f) < 0$ and $\alpha_2([0,1]) = 0$

(iv) $I(f) < 0$, $f_2([0,1]) > 0$ and $\alpha_2([0,1]) > 0$.

Case (i). In this case from (3.15), $\int J(f_1) \, d\alpha = 0$ or $C_1 \cdot f_2([0,1]) = 0$ or both. When $\int J(f_1) \, d\alpha = 0$, (3.17) and Fatou’s lemma imply that $I_p(f) = 0$.

When $C_1 \cdot f_2([0,1]) = 0$, we have $\int g(\alpha_1) \, df = 0$. From (3.18), (3.20) and Fatou’s lemma, we once again obtain $I_p(f) = 0$.

Case (ii). In this case $f << \alpha$ and we can adjoin the limit $(\dot{f}_1, \hat{g})$ to the martingale $\left\{ \frac{df_p}{d\alpha_p}, \sigma(P) \right\}$. The function $J$ is convex and from (3.15), $J(\dot{f}_1)$ is $\alpha$-integrable. This implies that $J(\frac{df_p}{d\alpha_p})$ is uniformly integrable. It therefore follows that $I_p(f) = I(f)$.

Case (iii). In this case $\alpha << f$ and $\left\{ \frac{d\alpha_p}{df_p}, \sigma(P) \right\}$ is a martingale under $f$ to which can be adjoined its limit $(\dot{\alpha}_1, \hat{g})$. The function $g$ is convex and from (3.22), $g(\dot{\alpha}_1)$ is $f$-integrable. This implies that $g(\frac{df_p}{d\alpha_p})$ is uniformly integrable. Again, it follows that $I_p(f) = I(f)$.

Case (iv). In this case $J(0) < 0$ and $K < 0$, hence the functions $J$ and $g$ are bounded on $[0, \mu]$ and $[0, \mu^{-1}]$, respectively. Using the definitions (3.19) and (3.21) and the bounded convergence theorem, we have $I_p(f) = I(f)$.

Remark. In Theorem 3.1 we have actually shown that
The next two lemmas establish the fact that \( I(f) \) is a regular rate function with respect to the weak*-topology on \( M[0,1] \). A sequence \( f_n \) in \( M[0,1] \) converges in the weak*-sense to \( f \) if \( f_n(t) \to f(t) \) for each \( t \) at which \( f \) is continuous. Following tradition, we will call the weak*-topology as the weak topology in the rest of this paper.

**Lemma 3.3.** The function \( I(f) \) is lsc in the weak topology.

**Proof.** Fix \( f \in M[0,1] \). Let \( f_n \to f \) weakly. We need to show that

\[
\liminf I(f_n) \geq I(f). \tag{3.24}
\]

If the support of \( f \) is not contained in the support of \( \alpha \), then \( I(f) = -\infty \) and there exists a weak open neighborhood \( G \) of \( f \) containing only measures whose supports are not included in the support of \( \alpha \). Then \( f_n \in G \) for all large \( n \) and thus \( \lim I(f_n) = -\infty \), which establishes (3.24).

If the support of \( f \) is contained in the support of \( \alpha \), choose a partition \( P = \{0=t_0, t_1, \ldots, t_k=1\} \) consisting of continuity points of \( f \).

Then \( f_n(t_i) \to f(t_i) \) for each \( i \), and thus \( \lim I_p(f_n) = I_p(f) \). From (3.23), \( I_p(f_n) \leq I(f_n) \). Thus \( \lim I(f_n) \geq I_p(f) \). By allowing \( \sigma(P) \) to tend to \( \infty \) along such partitions and using Theorem 3.2, we obtain (3.24).

**Lemma 3.4.** Let \( c < \infty \). The set

\[
\{ f : I(f) \leq c \} \tag{3.25}
\]
is compact.

Proof. Consider the partition \( P = [0,1] \). We have

\[
J(f([0,1])) = I_p(f) \leq I(f) \leq c
\]

for \( f \in \Gamma_c \). Since \( J(a) \to \infty \) as \( a \to \infty \), we can find \( d < \) such that

\[
\Gamma_c \subseteq \Delta_d \text{ where } \Delta_d = \{f: f([0,1]) \leq d\}. \text{ The set } \Delta_d \text{ is weakly compact and from Lemma 3.3 the set } \Gamma_c \text{ is weakly closed. Hence } \Gamma_c \text{ is weakly compact.} \]

The following minimax theorem is the driving force behind the upper bound of the LD results of the next section.

**Theorem 3.5.** Let \( F \) be a weakly closed subset of \( M[0,1] \). Then

\[
\sup_P I_p(F) = I(F), \tag{3.26}
\]

where for any set \( A \)

\[
I_p(A) = \inf_{f \in A} I_p(f) \text{ and } I(A) = \inf_{f \in A} I(f). \]

**Proof.** From (3.23) we immediately have

\[
\sup_P I_p(F) \leq I(F). \tag{3.27}
\]

Suppose that (3.26) were not true; then there exists an \( n < \) such that

\[
\sup_P I_p(F) < n < I(F). \tag{3.27}
\]
Thus, for each partition \( P = \{0 = t_0 < t_1 < \ldots < t_k = 1\} \), we can find \( f_p \) in \( M[0,1] \) such that \( I_p(f_p) < n \). The support of such an \( f_p \) will be contained in the support of \( a \). Let \( \hat{f}_p \), called the \( P \)-linear form of \( f_p \) with respect to \( a \), be defined by

\[
\hat{f}_p(A) = \sum_{i=2}^{k} \frac{f_p(t_i) - f_p(t_{i-1})}{a(t_i) - a(t_{i-1})} a(\bigcap(t_{i-1}, t_i])
\]

\[
+ \frac{f_p(t_1)}{a(t_1)} a(\bigcap[0, t_1]),
\]

Then \( \hat{f}_p(t_1) = f_p(t_1) \), \( 0 \leq i \leq k \), and

\[ I_p(f_p) = I_p(\hat{f}_p) = I(\hat{f}_p). \]

Hence \( \{\hat{f}_p\} \) is a net in the set \( \Gamma_n \) which is compact from Lemma 3.3. Thus, there is a cluster point \( f_0 \) of this net and \( I(f_0) \leq n \) from the lower semicontinuity of \( I \). If we can show that \( f_0 \) is a cluster point of \( \{f_p\} \), it will follow that \( f_0 \) belongs to \( F \) since \( F \) is closed. Since \( I(f_0) \leq n \), this will lead to a contradiction of (3.27), and the conclusion (3.26) would have been established.

Let \( P' = \{0 = t_1', t_2', \ldots, t_{k'}'\} \) be a partition consisting of continuity points of \( f_0 \). Fix \( \varepsilon > 0 \), and let \( N_{p',\varepsilon} \) be a weak neighborhood of \( f_0 \) defined by

\[
N_{p',\varepsilon} = \{f : \max_{i} |f(t_i') - f_0(t_i')| < \varepsilon\}.
\]

Since \( f_0 \) is a cluster point of \( \{f_p\} \), there is a partition \( P'' > P' \) such that \( \hat{f}_p \in N_{p',\varepsilon} \) if \( P > P' \). Since \( \hat{f}_p \) and \( f_p \) agree on the partition \( P \), it
follows that \( f_p \in \mathcal{N}_{p', \varepsilon} \) for \( P \succ P' \) and that \( f_0 \) is a cluster point of \( \{f_p\} \). This completes the proof of Theorem 3.4.

Theorems 3.2, 3.5 and Lemmas 3.3, 3.4 dealt with the rate function \( I(f) \) which involved the function \( J \). It was assumed that \( J \) was the rate function of a non-negative random variable \( X \) satisfying (3.1). When these results are applied in Section 4 and 5 we will restrict \( X \) to be non-negative and infinitely divisible. For this special case the following facts are noted concerning the finiteness of \( J(0) \) and \( C_1 \). From (3.8), \( J(0) \) is finite if and only if \( P(X=0) > 0 \). Thus \( J(0) = \mu \) for the Gamma distribution and \( J(0) = \mu \) for the Poisson distribution with parameter \( \mu \). On the other hand, \( C_1 = 1 \) for the Poisson distribution and \( C_1 = 1 \) for the Gamma distribution with shape parameter 1.

The results of the rest of this paper would be strengthened if we could have proved Lemmas 3.3, 3.4 and Theorem 3.5 in the Skorohod topology wherein the distribution functions \( f \) are considered as elements of \( D[0,1] \). Unfortunately certain complications occur as indicated by the following remark.

The Skorohod topology is stronger than the weak topology. Thus the rate function \( I(f) \) is Skorohod lsc, and hence \( \Gamma_c \) is Skorohod closed. However \( \Gamma_c \) is not Skorohod compact as the following example demonstrates. Let

\[
    f_n(t) = \begin{cases} 
    t & 0 \leq t \leq \frac{1}{2} - \frac{1}{n} \\
    t + n(t - \frac{1}{2} + \frac{1}{n}) & \frac{1}{2} - \frac{1}{n} < t \leq \frac{1}{2} \\
    t + 1 & \frac{1}{2} < t \leq 1. 
    \end{cases}
\]
Let \( J(a) = a - 1 - \log a \), which is the rate function corresponding to the Gamma distribution with shape parameter 1. Let \( \alpha \) be the Lebesgue measure. Then

\[
I(f_n) = 1 - \frac{1}{n} \log (1+n)
\]

and \( f_n \in \Gamma_1 \). Note that \( f_n \to f \) in the weak topology, where

\[
f(t) = \begin{cases} 
  t & t < 1/2 \\
  t + 1 & 1/2 < t \leq 1. 
\end{cases}
\]

Since \( f_n \) is continuous and \( f \) has a jump at \( t = 1/2 \), no subsequence of \( f_n \) can converge in the Skorohod topology.
4. LD Rates for Stochastic Processes with Stationary
    and Non-negative Independent Increments.

Let \( \{X(t), 0 \leq t \leq 1\} \) be a stochastic process with stationary and
non-negative independent increments and measurable sample paths with
\( X(0) = 0 \). Since the increments are non-negative, the sample paths of
\( \{X(t), 0 \leq t \leq 1\} \) can be considered as members of \( M[0,1] \). Note that \( X(1) \) is a
non-negative infinitely divisible random variable.

We will assume that

\[
\phi(\theta) = \mathbb{E}(e^{\theta X(1)}) < \infty
\]

for some \( \theta > 0 \). Let \( \psi(\theta) = \log \phi(\theta) \) and let \( J \) be the rate function of \( X(1) \)
as defined in (3.2). Let \( \alpha \) be a probability measure on \([0,1]\). Let the
rate function \( I(\alpha) \) on \( M[0,1] \) be as defined in (3.15). For \( \lambda > 0 \), define

\[
Z_\lambda(t) = \frac{1}{\lambda} X(\alpha([0,t])) \quad 0 \leq t \leq 1.
\]

Then \( \{Z_\lambda(t), 0 \leq t \leq 1\} \) is a process with values in \( M[0,1] \). Endow
\( M[0,1] \) with the weak topology and denote the induced distribution of
\( \{Z_\lambda(t), 0 \leq t \leq 1\} \) by \( P_\lambda \). In this section we show that
\( \{P_\lambda\} \) is LD tight (Lemma 4.3) and satisfies the LDP with rate
function \( I(\alpha) \). (Theorems 4.1 and 4.2).

**Theorem 4.1.** Let \( F \) be a weakly closed subset of \( M[0,1] \). Then

\[
\lim_{\lambda \to \infty} \frac{1}{\lambda} \log P_\lambda(F) \leq -I(F).
\]

**Proof.** Let \( P = \{0 = t_0 < t_1 < \ldots < t_k = 1\} \) be a partition and let
(4.4) \[ A_i = \begin{cases} [0, t_i] & \text{if } i = 1 \\ (t_{i-1}, t_i] & \text{if } i = 2, \ldots, k. \end{cases} \]

Let

(4.5) \[ W_{\lambda,i} = Z_\lambda(A_i), \quad 1 \leq i \leq k. \]

Then \( \{W_{\lambda,i}, 1 \leq i \leq k\} \) are independent, and from Theorem 3.1 and Lemma 2.8 satisfy the LDP with rate function

(4.6) \[ \sum_i \frac{x_i}{\alpha(A_i)} \alpha(A_i). \]

Now,

\[ P_\lambda(F) = P(Z_\lambda \in F) \leq P\{I_p(Z_\lambda) \geq I_p(F)\} \]

where

\[ I_p(Z_\lambda) = \sum_i W_{\lambda,i} \cdot \frac{\alpha(A_i)}{\alpha(A_i)}, \]

Since the support of \( x(1) \) is \([0,=)\) the function \( J(x) \) is continuous in \([0,=)\) and \( J(x) \to = \) as \( x \to = \). Thus the set

\[ \{(x_1, \ldots, x_k): \sum_i \frac{x_i}{\alpha(A_i)} \alpha(A_i) \geq I_p(F)\} \]

is closed in \( \mathbb{R}^k \). Using the LDP of \( \{W_{\lambda,i}, 1 \leq i \leq k\} \) and its rate function in (4.6), we obtain

\[ \lim_{\lambda} \frac{1}{\lambda} \log P_\lambda(F) \leq -I_p(F). \]
Since \( P \) is arbitrary, we can use the minimax result in Theorem 3.5 to obtain

\[
\lim \frac{1}{\lambda} \log P_\lambda(F) \leq -I(F).
\]

**Theorem 4.2.** Let \( G \) be a weakly open subset of \( M[0,1] \). Then

\[
\lim \frac{1}{\lambda} \log P_\lambda(G) \geq -I(G).
\]

**Proof.** There is nothing to prove if \( I(G) = -\infty \). Otherwise, fix \( \varepsilon > 0 \) and choose \( f \in G \) so that \( I(f) < I(G) + \varepsilon \). There is a \( \delta > 0 \) and a partition \( P = \{0 = t_0 < t_1 < \ldots < t_K = 1\} \) consisting of continuity points of \( f \) and \( \alpha \) such that the neighborhood

\[
N_{P, \varepsilon} = \{g : \max_i |g(A_i) - f(A_i)| < \delta\}
\]

of \( f \) is contained in \( G \). Here \( A_1, \ldots, A_K \) are as defined in (4.4). Thus,

\[
P_\lambda(G) \geq P_\lambda \left( \max_i |W_{\lambda, i} - f(A_i)| < \delta \right)
\]

where \( \{W_{\lambda, i}, 1 \leq i \leq K\} \) are as defined in (4.5) and satisfy the LDP with the rate function in (4.6). Furthermore the set \( G^* = \{x_1, \ldots, x_K) : \max_i |x_i - f(A_i)| < \delta\} \) is open in \( R^K \). Thus

\[
\lim \frac{1}{\lambda} \log P_\lambda(G) \geq \inf \sum_i \lambda \left( \frac{x_i}{\alpha(A_i)} \right) \alpha(A_i)
\]

where infimum is taken over the set \( G^* \). Hence,
\[
\lim_{\lambda} \frac{1}{\lambda} \log P_{\lambda}(G) \geq -I(f) \geq -I(f) \geq -I(G) - \varepsilon.
\]

This completes the proof of Theorem 4.2. \(\Box\)

**Lemma 4.3.** The family of probability measures \(\{P_{\lambda}\}\) is LD tight.

**Proof.** This follows from Lemma 2.6. A more direct proof is as follows. The sets

\[ K_L = \{f: f([0,1]) \leq L\} \]

are compact. Let \(\theta > 0\) be such that \(\phi(\theta) < \infty\). From the Markov inequality, we have

\[ P_{\lambda}(K_L^c) \leq \exp\{-[\theta L - \psi(\theta)]\} \]

which can be made as small as we please by choosing \(L\) sufficiently large. This completes the proof. \(\Box\)
5. LD Rates for Stochastic Processes with Stationary Independent Increments with no Gaussian Component.

Let \( \{X(t), 0 \leq t \leq 1\} \) be stochastic processes with stationary independent increments and measurable sample paths with \( X(0) = 0 \). Let the infinitely divisible random variable \( X(1) \) have a finite moment generating function \( \phi(\theta) \) which is finite for \( |\theta| < \eta \) for some \( \eta > 0 \). Assume that \( X(1) \) possesses no Gaussian component.

From standard results on infinitely divisible distributions (eg. Breiman, 1968, Chapter 14) it follows that

\[
\psi(\theta) = \log \phi(\theta) = \int (e^{\theta x} - 1) dv(x)
\]

where the Levy measure \( \nu \) (possibly unbounded) satisfies \( \int |x| dv(x) < \infty \) and that the sample paths of \( \{X(t), 0 \leq t \leq 1\} \) lie in \( BV[0,1] \), the space of functions of bounded variation on \([0,1]\). Thus, we can write

\[
X(t) = X^{(1)}(t) - X^{(2)}(t)
\]

where \( X^{(1)}(t) \) and \( X^{(2)}(t) \) are two independent stochastic processes with stationary and non-negative independent increments with Levy measures for \( X^{(1)}(1) \) and \( X^{(2)}(t) \) are given by \( \nu^{(1)}(A) = \nu(A \cap [0,\infty)) \) and \( \nu^{(2)}(A) = \nu(-A \cap (-\infty,0)) \), respectively.

Let \( J \), \( J^{(1)} \) and \( J^{(2)} \) denote the rate functions associated with \( X \), \( X^{(1)} \) and \( X^{(2)} \). That is, \( J(a) = \sup \{\theta a - \psi(\theta)\} \) and \( J^{(1)}(a) = \)
sup \{\alpha - \psi^{(1)}(\theta)\} \text{ where } \psi^{(1)}(\theta) = \int (e^{\theta X} - 1) d\nu^{(1)}(x) \text{ is the cumulant generating function of } X^{(1)}, \ i = 1, 2.

Let \(\alpha\) be a probability measure on \([0,1]\). Define

(5.1) \(Z_\lambda(t) = \lambda^{-1}X(\lambda \alpha([0,t])) \text{ for } 0 \leq t \leq 1\).

Let \(Z_\lambda^{(1)}\) and \(Z_\lambda^{(2)}\) be defined in terms of \(X^{(1)}(\cdot)\) and \(X^{(2)}(\cdot)\) in a fashion similar to (5.1). Then,

(5.2) \(Z_\lambda(t) = Z_\lambda^{(1)}(t) - Z_\lambda^{(2)}(t)\).

Note that \(\{Z_\lambda(t): 0 < t < 1\}\) takes values in \(BV[0,1]\) - the space of functions of bounded variation, or equivalently, signed measures on \([0,1]\).

Let \(f \in BV[0,1]\). Let its Hahn-Jordan decomposition be given by

\[ f = h^{(1)} - h^{(2)} \]

where \(h^{(1)}, h^{(2)} \in M[0,1]\). Also suppose that

\[ f = f^{(1)} - f^{(2)} \]

where \(f^{(1)}, f^{(2)} \in M[0,1]\), and for any function \(p\) in \(BV[0,1]\) let

\[ p = p_1 + p_2 \text{ where } p_1 \ll \alpha \text{ and } p_2 \perp \alpha \text{ and let } \dot{p}_1 = \frac{dp_1}{d\alpha}. \]

It is clear that

\[ f_1 = h_1^{(1)} - h_1^{(2)} = f_1^{(1)} - f_1^{(2)}, \]

(5.3) \[ \dot{f}_1 = \dot{f}_1^{(1)} - \dot{f}_1^{(2)} \]

26
and

\[ \inf \{ f_2^{(1)}[0,1] : f = f(1) - f(2) ; f(1), f(2) \in M[0,1] \} \]

\[ = h_2^{(1)}([0,1]), \ i = 1, 2. \]

The definitions of \( h_1^{(1)} \), \( h_2^{(1)} \), \( h_2^{(2)} \) above will be used in the statement of the theorem, below, which contains the main LD result of this paper.

**Theorem 5.1.** Let \( P_\lambda \) be the probability distribution of \( \{ Z(t), 0 < t < 1 \} \). Then \( P_\lambda \) satisfies the LD principle with the rate function

\[ I(f) = \int J(f_1^0) \, d\alpha + C_1 h_2^{(1)}([0,1]) + C_2 h_2^{(2)}([0,1]) \]

where \( f_1, h_2^{(1)}, h_2^{(2)} \) are as defined before and where \( C_1 \) and \( C_2 \) are given by (3.6).

**Proof.** Let \( P^{(1)}_\lambda \) be the distribution of \( Z^{(1)}(\cdot) \) in \( M[0,1], \ i = 1, 2 \). Let \( g \) be a function from \( M[0,1] \times M[0,1] \) into \( \text{BV}[0,1] \) defined by \( g(f(1), f(2)) = f(1) - f(2) \).

Then \( g \) is a continuous function and \( P^{(1)}_\lambda = (P^{(1)}_\lambda \times P^{(2)}_\lambda) g^{-1} \).

From Theorems 4.1, 4.2 and Lemma 4.3, \( P^{(1)}_\lambda \) is LD tight and satisfies the LDP with rate function

\[ I^{(1)}(f) = \int J^{(1)}(f_1) d\alpha + C^{(1)} f_2([0,1]) \]
where \( f = f_1 + f_2 \) with \( f_1 \ll \alpha \) and \( f_2 \perp \alpha \) and \( \dot{f}_1 = \frac{df_1}{d\alpha} \) and where \( C(1) \) is given by (3.10). From Corollary 2.9 \( P_{\lambda}^{(1)} \times P_{\lambda}^{(2)} \) satisfies the LDP with rate function \( I(1)(f(1)) + I(2)(f(2)) \) for \( f(1), f(2) \in M[0,1] \). From the contraction principle, \( P_{\lambda} \) satisfies the LDP with rate function

\[
\inf \left\{ \int J(1) \left( \dot{f}_1(1) \right) d\alpha + \int J(2) \left( \dot{f}_2(2) \right) d\alpha + C(1) f_2(1)([0,1]) + C(2) f_2(2)([0,1]) \right\}
\]

\[
f(1), f(2): f = f(1) - f(2)
\]

in view of (5.3), (3.11) and (5.4).
6. Applications to the Poisson, Gamma and Dirichlet Processes.

In this section we evaluate the rate functions for three processes.

Example 1 - Poisson Processes.

Let \( \{X(t), 0 \leq t \leq 1\} \) be a Poisson process with constant intensity \( \mu \).

Define the process \( \{Z\lambda(t), 0 \leq t \leq 1\} \) as in (4.2). Then

\[ \{\lambda Z\lambda(t), 0 \leq t \leq 1\} \] is a Poisson process with intensity function

\[ \lambda \mu(0,t) \].

The distribution of \( X(1) \) is Poisson with parameter \( \mu \) and thus

\[ J(a) = a \log \frac{a}{\mu} - a + \mu \text{ and } C_1 = - \]

where \( J(a) \) and \( C_1 \) are as defined in (3.2) and (3.6). Thus, as an application of Theorems 4.1 and 4.2, \( \{Z\lambda(t), 0 \leq t \leq 1\} \) satisfies the LDP with rate function

\[
(6.1) \quad I(f) = \begin{cases} 
\int_0^1 f \int_0^1 \log \left( \frac{\lambda}{\mu} \right) d\alpha + \mu - f([0,1]) \text{ if } f << \alpha \\
= \text{ otherwise}
\end{cases}
\]

This result can also be derived from Varadhan (1966) since \( C_1 = - \).

Example 2 - Gamma Processes. Let \( \{X(t), 0 \leq t \leq 1\} \) be a Gamma process, that is a stochastic process with stationary independent increments and measurable paths with \( X(0) = 0 \) and such that \( X(1) \) has a Gamma distribution with shape parameter 1. Then

\[ J(a) = a - 1 - \log a, J(0) = - \text{ and } C_1 = 1, \]

where \( J(a) \) and \( C_1 \) are as defined in (3.2) and (3.6). Then the process

\( \{Z\lambda(t), 0 \leq t \leq 1\} \) as defined in (4.2) satisfies the LDP with
Example 3 - Dirichlet Processes. Consider the process
\[ \{W_\lambda(t), 0 \leq t \leq 1\} \text{ where } W_\lambda(t) = Z_\lambda(t)/Z_\lambda(1) \]
where \( Z_\lambda \) is as defined in Example 2. Then \( \{W_\lambda(t), 0 \leq t \leq 1\} \) is
the Dirichlet process with parameter \( \lambda \alpha(\cdot) \) as defined in Ferguson (1973).

Sethuraman and Tiwari (1982) have shown that as \( \lambda \to 0 \), \( W_\lambda \) converges
in distribution to \( W_0 \) where \( W_0 \) is the random probability measure \( \delta_Y(\cdot) \)
where \( \delta_a(\cdot) \) stands for the degenerate measure at \( a \) and \( Y \) is a random
variable with distribution \( \alpha \). However, if we let \( \lambda \to \infty \), then \( W_\lambda \)
converges to the constant \( a \) in \( M[0,1] \). The contraction principle and the
LDP for the Gamma process show that the Dirichlet process with parameter \( \lambda \alpha \)
satisfies the LDP, as \( \lambda \to \infty \), with the rate function

\[
I(f) = \begin{cases} 
K(a,f) & \text{if } f(1) = 1 \text{ and } f = a \\
\infty & \text{otherwise,}
\end{cases}
\]

where \( K(a,f) \) is the Kullback-Leibler information number between two
probability measures \( a \) and \( f \) defined by

\[
K(a,f) = - \int \log \frac{df}{da} \, da.
\]
References


Let $X$ be a topological space and $F$ denote the Borel $\sigma$-field in $X$. A family of probability measures $\{P_\lambda\}$ is said to obey the large deviation principle (LDP) with rate function $I(\cdot)$ if $P_\lambda(A)$ can be suitably approximated by $\exp\{-\lambda \inf I(x)\}$ for appropriate sets $A$ in $F$. Here the LDP is studied for probability measures induced by stochastic processes with stationary and independent increments which have no Gaussian component. It is assumed that the moment generating function of the increments exists and thus the sample paths of such stochastic processes lie in the space of functions f bounded.
variation. The LDP for such processes is obtained under the weak*-topology. This covers a case which was ruled out in the earlier work of Varadhan (1966). As applications, the large deviation principle for the Poisson, Gamma and Dirichlet processes are obtained.