The Annals of Probability
1987, Vol. 15, No. 2, 610–627

LARGE DEVIATIONS FOR PROCESSES WITH INDEPENDENT INCREMENTS

BY JAMES LYNCH[1] AND JAYARAM SETHURAMAN[2]

University of South Carolina and Florida State University

Let $\mathscr{X}$ be a topological space and $\mathscr{F}$ denote the Borel σ-field in $\mathscr{X}$. 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 $\exp\{-\lambda \inf_{x \in A} I(x)\}$ for appropriate sets A in $\mathscr{F}$. 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 $\mathscr{X}$ be a topological space and $\mathscr{F}$ denote the Borel σ-field in $\mathscr{X}$. Let $\{P_\lambda\}$ be a family of probability measures on $(\mathscr{X}, \mathscr{F})$. The family $\{P_\lambda\}$ is said to obey the large deviation principle (LDP) (for a more precise definition see Section 2, Definition 2.2) 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 $\mathscr{F}$.

Important examples of the LDP include the cases where P_λ (λ a positive integer) is either (i) the probability measure induced by the average of λ i.i.d. random variables [see Cramér (1937), Chernoff (1952), Bahadur and Zabell (1979) and Varadhan (1984)] or (ii) the probability measure of the empirical distribution of λ i.i.d. random variables [Groeneboom, Oosterhoff and Ruymgaart (1979) and Bahadur and Zabell (1979)]. These authors exploited the i.i.d. property and used the techniques of moment generating functions, conjugate distributions and subadditivity to obtain these LDP results. Ellis (1984) has elegantly shown how to establish the LDP when $\mathscr{X} = R^k$, solely in terms of the moment generating functions of P_λ. Further examples may be found in the recent surveys on large deviations by Azencott (1980) and Varadhan (1984).

The establishment of the LDP has had important implications in various areas of statistics. It has been used to obtain the asymptotic efficiencies of tests and estimates [Chernoff (1952) and Bahadur (1960a, b), (1967) and (1971)] and to obtain the asymptotic behavior of functional integrals associated with solutions of stochastic integrals [Varadhan (1966) and (1984)]. It appears in the evaluation of the "free" energy in statistical mechanics [Lanford (1973) and Ruelle (1969)].

Received February 1985; revised June 1986.
[1]Research supported by U.S. Army Research Office Grant No. DAAG29-84-K-0007.
[2]Research supported by U.S. Army Research Office Grant No. DAAG29-82-K-0168.
AMS 1980 *subject classifications.* Primary 60F10, 60J30; secondary 60E07.
Key words and phrases. Large deviations, rate function, stationary and independent increments, Dirichlet process.

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

It is also intimately related to certain types of laws of large numbers [Shepp (1964) and Erdös and Rényi (1970)].

Let $X(0) = 0$ and let $\{X(t),\ t \geq 0\}$ be a stochastic process with stationary independent increments. Let P_λ be the probability measure of $\{Z_\lambda(t),\ 0 \leq t \leq 1\}$ where $Z_\lambda(t) = (1/\lambda)X(\lambda t)$. When $\{X(t),\ t \geq 0\}$ is a Brownian motion, we may consider P_λ as a probability measure on $C[0,1]$ endowed with the uniform topology. Ventsel (1976) has given LDP results for the above P_λ (and other diffusion processes obtained as solutions to stochastic differential equations) in the technical sense of Definition 2.2.

Consider once again the general case of a process $\{X(t),\ t \geq 0\}$ with stationary independent increments. Suppose that $\phi(\theta) = E(e^{\theta X(1)}) < \infty$ for θ in a neighborhood of 0. Let the rate function of $X(1)$, $J(a)$, be defined by $J(a) = \sup_\theta[a\theta - \log \phi(\theta)]$, and let

$$\frac{J(a)}{|a|} \to \infty \quad \text{as } |a| \to \infty. \tag{1.1}$$

Condition (1.1) is a growth condition on the rate function of $X(1)$ and is satisfied in many situations like the Gaussian process and the Poisson process. When this growth condition (1.1) holds, one can use the results of Varadhan (1966) and easily obtain LDP results for P_λ in the technical sense of Definition 2.2, by viewing P_λ as a probability measure on $D[0,1]$ endowed with the Skorohod topology. What happens if $X(1)$ has a moment generating function and the growth condition (1.1) is violated? The process $\{X(t),\ t \geq 0\}$ cannot contain a Gaussian component. Let us therefore consider a stochastic process $\{X(t),\ t \geq 0\}$ with stationary independent increments and with no Gaussian component.

Let $X(0) = 0$ and let $X(1)$ have a finite moment generating function. Let P_λ be the probability measure of $\{Z_\lambda(t),\ 0 \leq t \leq 1\}$, where $Z_\lambda(t) = (1/\lambda)X(\lambda t)$ or, more generally, $Z_\lambda(t) = (1/\lambda)X(\lambda\alpha[0, t])$ where α is a probability measure on $[0,1]$, and can be considered as a time deformation. We may view P_λ as a probability measure on $BV[0,1]$, the space of functions of bounded variation on $[0,1]$ endowed with the weak*-topology. In this paper we establish LDP results for this P_λ on $BV[0,1]$ in the technical sense of Definition 2.2. These LDP results are illustrated with applications to the Gamma process and 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, which are required in the proofs of the main results of this paper found in Sections 4 and 5. 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 $BV[0,1]$. Finally, Section 6 is devoted to applications to the Poisson, Gamma and Dirichlet processes.

2. Definitions and general results. Let $\mathscr{X}$ be a topological space and $\mathscr{F}$ be the Borel σ-field in $\mathscr{X}$. Let $\{P_\lambda\}$ be a family of probability measures on

$(\mathscr{X}, \mathscr{F})$. The following definitions which are slight variants of those of Varadhan (1984) allow us to state many large deviation results in concise form.

DEFINITION 2.1. A function $I(\cdot)$ on $\mathscr{X}$ is said to be a *regular rate function* if

$$0 \le I(x) \le \infty, \tag{2.1}$$

$$I(\cdot) \text{ is lower semicontinuous (lsc)} \tag{2.2}$$

and

$$\text{for each } c < \infty,\ \Gamma_c = \{x \colon I(x) \le c\} \text{ is compact.} \tag{2.3}$$

For any subset A of $\mathscr{X}$, define

$$I(A) = \inf_{x \in A} I(x). \tag{2.4}$$

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

$$I(\cdot) \text{ is a regular rate function,} \tag{2.5}$$

$$\text{for each closed set } F,\ \limsup \frac{1}{\lambda} \log P_\lambda(F) \le -I(F) \tag{2.6}$$

and

$$\text{for each open set } G,\ \liminf \frac{1}{\lambda} \log P_\lambda(G) \ge -I(G), \tag{2.7}$$

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 2.2 together with (2.8) below are satisfied:

$$\text{for each compact set } K,\ \limsup \frac{1}{\lambda} \log P_\lambda(K) \le -I(K). \tag{2.8}$$

DEFINITION 2.4. The measures $\{P_\lambda\}$ are *large deviation tight* (LD tight) if, for each $M < \infty$, there exists a compact set K_M such that

$$\limsup \frac{1}{\lambda} \log P_\lambda(K_M^c) \le -M. \tag{2.9}$$

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 $l < I(C)$. Let $M > l$ 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,

$$\limsup \frac{1}{\lambda} \log P_\lambda(C) \leq -\min\{I(C \cap K_M), M\} \leq -l. \qquad \square$$

Many interesting applications in large deviations occur when $\mathscr{X}$ 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 σ-fields to be Borel σ-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 $\{x_i,\ i = 1, 2, \dots\}$ be a countable dense set in $\mathscr{X}$. For any $\delta > 0$, let $A_i(\delta)$ be the open ball of radius δ around x_i. Then $\bigcup_i A_i(1/k) = \mathscr{X}$ for $k = 1, 2, \dots$. Fix $M > 0$ and an integer k. Consider the compact set $\Gamma_{2kM} = \{x\colon I(x) \leq 2kM\}$. There exists a finite open covering

$$A(k) = \bigcup_{i=1}^{I_k} A_i\left(\frac{1}{k}\right)$$

of Γ_{2kM}. Thus, from (2.6)

$$\limsup \lambda^{-1} P_\lambda(A^c(k)) \leq -I(A^c(k)) \leq -I(\Gamma_{2kM}^c) \leq -2kM.$$

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

$$B(k) = \bigcup_{i=1}^{J_k} A_i\left(\frac{1}{k}\right),$$

with $J_k \geq I_k$ such that

$$P_\lambda(B^c(k)) \leq e^{-\lambda M k},$$

for all λ. The set $K = \bigcap_{k=1}^{\infty} \overline{B(k)}$, where $\overline{B(k)}$ is the closure of $B(k)$, is totally bounded and closed, and hence is compact. Furthermore,

$$P_\lambda(K^c) \leq \sum_{k=1}^{\infty} P_\lambda(B^c(k)) \leq \frac{e^{-\lambda M}}{(1 - e^{-M})}, \tag{2.10}$$

for all $\lambda \geq 1$. This completes the proof of Lemma (2.6). $\square$

Let $\{P_\lambda^i\}$ be a family of probability measures on a Polish space $\mathscr{X}^i$, $i = 1, 2$. Let $P_\lambda = P_\lambda^1 \times P_\lambda^2$ be the product measure on the product space $\mathscr{X} = \mathscr{X}^1 \times \mathscr{X}^2$. 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 WLDP with rate function $I^i(x_i)$, $i = 1,2$. Then $\{P_\lambda\}$ satisfies the WLDP with rate function $I(x_1, x_2) = I^1(x_1) + I^2(x_2)$.*

PROOF. It is easy to check the regularity of $I(x_1, x_2)$ from the regularity of $I^1(x_1)$ and $I^2(x_2)$. Let $K \subset \mathscr{X}$ be compact and let $l < I(K)$. For each $(x_1, x_2) \in K$, since $I(\cdot)$ is lsc, there are open sets $O^i x_i$ in $\mathscr{X}^i$ containing x_i, $i = 1,2$, such that

$$\inf\{I(y_1, y_2): (y_1, y_2) \in O_{x_1}^1 \times O_{x_2}^2\} > l. \tag{2.11}$$

Furthermore, since $\mathscr{X}^i$ is Polish, we can find open subsets $N_{x_i}^i$ of $O_{x_i}^i$ such that $x_i \in N_{x_i}^i$ and $\overline{N}_{x_i}^i \subset O_{x_i}^i$. Consider the open covering $\bigcup_{(x_1, x_2) \in K} N_{x_1}^1 \times N_{x_2}^2$ of K. We can extract a finite subcovering $\bigcup_{m=1}^M N_{x_{1,m}}^1 \times N_{x_{2,m}}^2$ of K. Let K^1 and K^2 be the projections of K in $\mathscr{X}^1$ and $\mathscr{X}^2$. Then K^i and $M_{x_{i,m}}^i = \overline{N}_{x_{i,m}}^i \cap K^i$ are compact, $m = 1, \dots, M$, and $i = 1,2$. Furthermore, $K \subset \bigcup_{m=1}^M M_{x_{1,m}}^1 \times M_{x_{2,m}}^2$.

Thus, since $M_{x_{i,m}}^i$ is compact and $\{P_\lambda^i\}$ satisfies the WLDP,

$$\begin{aligned} \limsup \frac{1}{\lambda} \log P_\lambda(K) &\le -\min_m \left(I^1\left(M_{x_{1,m}}^1\right) + I^2\left(M_{x_{2,m}}^2\right)\right) \\ &\le -l, \end{aligned}$$

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

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

$$\begin{aligned} \liminf \frac{1}{\lambda} \log P_\lambda(O) &\ge \sum_{i=1}^{2} \liminf \frac{1}{\lambda} \log P_\lambda^i(O_{x_i}) \\ &\ge -I(x_1, x_2) \ge -I(O) - \varepsilon. \end{aligned}$$

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

The following corollary follows from Lemmas 2.5, 2.7 and 2.8.

COROLLARY 2.9. *Let $\{P_\lambda^i\}$ be LD tight and satisfy the WLDP, $i = 1,2$. Then $P_\lambda = P_\lambda^1 \times P_\lambda^2$ 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, 1984).

Let $\{P_\lambda\}$ satisfy the LDP with rate function $I(x)$. Let h be a continuous map from $\mathscr{X}$ into another topological space $\mathscr{Y}$, and let $Q_\lambda = P_\lambda h^{-1}$.

CONTRACTION PRINCIPLE. The measures $\{Q_\lambda\}$ satisfy the LDP with rate function

$$K(y) = \inf_{x:\, h(x) = y} I(x). \tag{2.12}$$

ASYMPTOTIC EXPRESSION FOR CERTAIN INTEGRALS. Let F be a bounded real valued continuous function on $\mathscr{X}$. Then

$$(2.13)\qquad \frac{1}{\lambda}\log\int \exp(\lambda F(x))\,dP_\lambda(x) \to \sup_x\,[F(x) - I(x)].$$

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 $M[0,1]$. We begin with some well known facts about LDP of sums of i.i.d. random variables and their rate functions. Let X be a real valued random variable and let

$$(3.1)\qquad \phi(\theta) = E(e^{\theta X}) < \infty,$$

for all $|\theta| < \eta$ where $\eta > 0$. Let $\psi(\theta) = \log\phi(\theta)$. Let $X_1, X_2, \dots$ be i.i.d. copies of X and let P_n be the distribution of $(X_1 + \cdots + X_n)/n$. 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 (1937) and Chernoff (1952)]. *The distributions $\{P_n\}$ are LD tight and satisfy the LDP with rate function $J(a)$ given by*

$$(3.2)\qquad J(a) = J_X(a) = \sup_t\,[at - \psi(t)].$$

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

$$(3.3)\qquad 0 \le J(a) \le \infty, \qquad J(\mu) = 0,$$
$$\text{where } E(X) = \mu, \text{ and } J(a) \to \infty \text{ as } |a| \to \infty;$$

$$(3.4)\qquad J(a) = \sup_{t\ge 0}\,[at - \psi(t)], \quad \text{if } a > \mu;$$

(3.5) $J(a)$ is convex;

$$(3.6)\qquad \lim_{a\to\infty}\frac{J(a)}{a} = C_1 \quad\text{and}\quad \lim_{a\to-\infty}\frac{J(a)}{|a|} = C_2 \text{ exist,}$$
$$\text{where } 0 < C_1, C_2 \le \infty;$$

the function $g(b)$ defined by

$$(3.7)\qquad g(b) = \begin{cases} bJ(1/b), & \text{if } 0 < b < \infty,\\ C_1, & \text{if } b = 0,\end{cases}$$
$$\text{is convex on } [0, -\infty);$$

(3.8) if X is a nonnegative random variable, 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 nonnegative 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

$$J(a) = \inf_b \left(J^{(1)}(a + b) + J^{(2)}(b) \right). \tag{3.9}$$

Let

$$C^{(i)} = \lim_{a \to \infty} \frac{J^{(i)}(a)}{a}, \qquad i = 1, 2. \tag{3.10}$$

We will now show that

$$C_i = C^{(i)}, \qquad i = 1, 2, \tag{3.11}$$

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 $\psi^{(2)}(\theta) \le 0$ for $\theta < 0$ since $X^{(2)}$ is nonnegative. Thus

$$a\theta - \psi^{(1)}(\theta) \le J(a) \le J^{(1)}(a + b) + J^{(2)}(b), \tag{3.12}$$

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

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

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

The main results of this paper which are contained in Sections 4 and 5 obtain large deviation rates for certain measures P_λ on BV[0, 1] derived from a stochastic process $\{X(t),\ t \ge 0\}$ and a time deformation α. From the Hahn–Jordan decomposition, one may consider BV[0, 1] as $M[0,1] - M[0,1]$ where $M[0,1]$ is the space of finite (nonnegative) measures on $([0,1], \mathscr{B})$, and $\mathscr{B}$ is the usual Borel σ-field in $[0,1]$. The large deviation rates for P_λ can be characterized by certain rate functions on BV[0, 1] and, more particularly, by rate functions on $M[0,1]$. These rate functions depend only on J, the rate function of the real valued random variable $X(1)$ and the time deformation measure α. In this section we define a rate function $I(f)$ depending on such a J and α and study its properties. Theorems 3.2 and 3.5 are the driving force behind the results of Sections 4 and 5. Theorem 3.2 shows that the rate function $I(f)$ can be approximated through another rate function $I_{\mathscr{P}}(f)$ defined on partitions $\mathscr{P}$ of $[0,1]$. Theorems 3.5 establishes a minimax theorem for $I_{\mathscr{P}}(f)$ over f and $\mathscr{P}$.

For any element f in $M[0,1]$, we define its distribution function $f(t)$ by letting $f(0) = 0$, $f(t) = f([0,t])$, $0 < t \le 1$. We also use the same symbol f to denote both the measure $f(A)$ and the (extended) distribution function $f(t)$. Let $J(a)$ be the rate function of a nonnegative random variable X satisfying (3.1).

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

Let $f \in M[0,1]$ and $\mathscr{P}$ be a partition. We define

$$(3.13) \qquad I_{\mathscr{P}}(f) = \begin{cases} \displaystyle\sum_i J\left(\frac{f(t_i) - f(t_{i-1})}{\alpha(t_i) - \alpha(t_{i-1})}\right)(\alpha(t_i) - \alpha(t_{i-1})), \\ \quad \text{if } \alpha(t_i) - \alpha(t_{i-1}) = 0 \text{ implies} \\ \quad f(t_i) - f(t_{i-1}) = 0,\ i = 1, \dots, k, \\ \infty, \text{ otherwise,} \end{cases}$$

where we observe the convention $0 \cdot (\text{undefined}) = 0$ and $0 \cdot \infty = 0$.

Denote the restriction of the measures α and f to $\sigma(\mathscr{P})$ by $\alpha_{\mathscr{P}}$ and $f_{\mathscr{P}}$, respectively. We may rewrite the definition in (3.13) by

$$(3.14) \qquad I_{\mathscr{P}}(f) = \begin{cases} \displaystyle\int J\left(\frac{df_{\mathscr{P}}}{d\alpha_{\mathscr{P}}}\right) d\alpha, & \text{if } f_{\mathscr{P}} \ll \alpha_{\mathscr{P}}, \\ \infty, & \text{otherwise.} \end{cases}$$

Let $f = f_1 + f_2$ be the Lebesgue decomposition of f with respect to α, 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 α_1, α_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 $\dot{f}_1 = df_1/d\alpha$ and $\dot{\alpha}_1 = d\alpha_1/df$. Then $\dot{f}_1 = 1/\dot{\alpha}_1 > 0$ a.e. on $(L \cup M)^c$ with respect to f and α.

Define

$$(3.15) \qquad I(f) = \begin{cases} \displaystyle\int J(\dot{f}_1)\, d\alpha + C_1 f_2([0,1]), & \text{if supp. } f \subset \text{supp. } \alpha, \\ \infty, & \text{otherwise,} \end{cases}$$

where supp. stands for support and C_1, which is defined in (3.6), depends on J. The following theorem relates $I_{\mathscr{P}}(f)$ to $I(f)$.

THEOREM 3.2. *As* $\sigma(\mathscr{P}) \to \mathscr{B}$,

$$(3.16) \qquad I_{\mathscr{P}}(f) \to I(f).$$

PROOF. When supp. f is not contained in supp. α, $I(f) = \infty$. In this case we do not have $f_{\mathscr{P}} \ll \alpha_{\mathscr{P}}$ for some $\mathscr{P}$. Then $I_{\mathscr{P}}(f) = \infty$ and $I_{\mathscr{P}'}(f) = \infty$ for finer partitions $\mathscr{P}'$. This establishes (3.16) in this case.

From now on assume that supp. $f \subset$ supp. α. It follows that $f_{\mathscr{P}} \ll \alpha_{\mathscr{P}}$ for each $\mathscr{P}$ and that $\{df_{\mathscr{P}}/d\alpha_{\mathscr{P}}, \sigma(\mathscr{P})\}$ is a martingale. Since $J(\alpha)$ is convex,

$\{J(df_{\mathscr{P}}/d\alpha_{\mathscr{P}}), \sigma(\mathscr{P})\}$ is a sub-martingale. We also have

$$\frac{df_{\mathscr{P}}}{d\alpha_{\mathscr{P}}} \to \dot{f}_1 \quad \text{and} \quad J\left(\frac{df_{\mathscr{P}}}{d\alpha_{\mathscr{P}}}\right) \to J(\dot{f}_1) \quad \text{a.e. } \alpha. \tag{3.17}$$

Under the condition supp. $f \subset$ supp. α it may not be true that $\alpha_{\mathscr{P}} \ll f_{\mathscr{P}}$. We will use the notation $d\alpha_{\mathscr{P}}/df_{\mathscr{P}}$ to denote the Radon–Nikodym derivative of $\alpha^*_{\mathscr{P}}$, the absolutely continuous part of $\alpha_{\mathscr{P}}$, with respect to $f_{\mathscr{P}}$. Then $\{d\alpha_{\mathscr{P}}/df_{\mathscr{P}}, \sigma(\mathscr{P})\}$ is a super-martingale under f.

Recall the function g defined in (3.7) based on the function J. The function g, is continuous when $g(0) < \infty$ and continuous in the extended sense when $g(0) = \infty$.

Thus,

$$\frac{d\alpha_{\mathscr{P}}}{df_{\mathscr{P}}} \to \dot{\alpha}_1 \quad \text{and} \quad g\left(\frac{d\alpha_{\mathscr{P}}}{df_{\mathscr{P}}}\right) \to g(\dot{\alpha}_1) \quad \text{a.e. } f. \tag{3.18}$$

Let $E(X) = \mu$. Re-examining (3.13) and (3.14) we get the alternate expression

$$I_{\mathscr{P}}(f) = \int_{d\alpha_{\mathscr{P}}/df_{\mathscr{P}} < \mu^{-1}} g\left(\frac{d\alpha_{\mathscr{P}}}{df_{\mathscr{P}}}\right) df + \int_{df_{\mathscr{P}}/d\alpha_{\mathscr{P}} \le \mu} J\left(\frac{df_{\mathscr{P}}}{d\alpha_{\mathscr{P}}}\right) d\alpha \tag{3.19}$$

and

$$I_{\mathscr{P}}(f) \ge \int g\left(\frac{d\alpha_{\mathscr{P}}}{df_{\mathscr{P}}}\right) df, \tag{3.20}$$

with equality in (3.20) when $\alpha_{\mathscr{P}} \ll f_{\mathscr{P}}$. Similarly, recalling (3.15), we can write the alternative forms

$$I(f) = \int_{\dot{\alpha}_1 < \mu^{-1}} g(\dot{\alpha}_1)\, df + \int_{\dot{f}_1 \le \mu} J(\dot{f}_1)\, d\alpha \tag{3.21}$$

and

$$I(f) = \int g(\dot{\alpha}_1)\, df + J(0)\alpha_2([0,1]). \tag{3.22}$$

It is possible that $C_1 = \infty$ or $J(0) = \infty$ or both and so we consider the following cases to complete the proof:

(i) $I(f) = \infty$,

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

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

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

CASE (i). In this case from (3.15), $\int J(\dot{f}_1)\, d\alpha = \infty$ or $C_1 \cdot f_2([0,1]) = \infty$ or both. When $\int J(\dot{f}_1)\, d\alpha = \infty$, (3.17) and Fatou's lemma imply that $I_{\mathscr{P}}(f) \to \infty$. When $C_1 \cdot f_2([0,1]) = \infty$, we have $\int g(\dot{\alpha}_1)\, df = \infty$. From (3.18), (3.20) and Fatou's lemma, we once again obtain $I_{\mathscr{P}}(f) \to \infty$.

CASE (ii). In this case $f \ll \alpha$ and we can adjoin the limit $(\dot{f}_1, \mathscr{B})$ to the martingale $\{df_{\mathscr{P}}/d\alpha_{\mathscr{P}}, \sigma(\mathscr{P})\}$. The function J is convex and from (3.15), $J(\dot{f}_1)$ is α-integrable. This implies that $\{J(df_{\mathscr{P}}/d\alpha_{\mathscr{P}})\}$ is uniformly integrable. It therefore follows that $I_{\mathscr{P}}(f) \to I(f)$.

CASE (iii). In this case $\alpha \ll f$ and $\{d\alpha_{\mathscr{P}}/df_{\mathscr{P}}, \sigma(\mathscr{P})\}$ is a martingale under f to which can be adjoined its limit $\{\dot{\alpha}_1, \mathscr{B}\}$. The function g is convex and from (3.22), $g(\dot{\alpha}_1)$ is f-integrable. This implies that $\{g(df_{\mathscr{P}}/d\alpha_{\mathscr{P}})\}$ is uniformly integrable. Again, it follows that $I_{\mathscr{P}}(f) \to I(f)$.

CASE (iv). In this case $J(0) < \infty$ and $C_1 < \infty$, 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 $I_{\mathscr{P}}(f) \to I(f)$. □

REMARK. In Theorem 3.2 we have actually shown that

$$\sup_{\mathscr{P}} I_{\mathscr{P}}(f) = I(f). \tag{3.23}$$

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 refer to 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 α, 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 α. 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 α, choose a partition $\mathscr{P} = \{0 = t_0, t_1, \dots, t_k = 1\}$ consisting of continuity points of f. Then $f_n(t_i) \to f(t_i)$ for each i, and thus $\lim I_{\mathscr{P}}(f_n) = I_{\mathscr{P}}(f)$. From (3.23), $I_{\mathscr{P}}(f_n) \leq I(f_n)$. Thus $\liminf I(f_n) \geq I_{\mathscr{P}}(f)$. By allowing $\sigma(\mathscr{P})$ to tend to $\mathscr{B}$ along such partitions and using Theorem 3.2, we obtain (3.24). □

LEMMA 3.4. *Let $c < \infty$. The set*

$$\Gamma_c = \{f \colon I(f) \leq c\} \tag{3.25}$$

is compact.

PROOF. Consider the partition $\mathscr{P} = \{0,1\}$. We have

$$J(f([0,1])) = I_{\mathscr{P}}(f) \leq I(f) \leq c,$$

for $f \in \Gamma_c$. Since $J(a) \to \infty$ as $a \to \infty$, we can find $d < \infty$ such that $\Gamma_c \subset \Delta_d$ where $\Delta_d = \{f\colon f([0,1]) \le d\}$. The set Δ_d is weakly compact and from Lemma 3.3 the set Γ_c is weakly closed. Hence Γ_c is weakly compact. □

The following is a minimax theorem for $I_{\mathscr{P}}(f)$.

THEOREM 3.5. *Let F be a weakly closed subset of $M[0,1]$. Then*

$$\sup_{\mathscr{P}} I_{\mathscr{P}}(F) = I(F), \tag{3.26}$$

where for any set A

$$I_{\mathscr{P}}(A) = \inf_{f \in A} I_{\mathscr{P}}(f) \quad \textit{and} \quad I(A) = \inf_{f \in A} I(f).$$

PROOF. From (3.23) we immediately have

$$\sup_{\mathscr{P}} I_{\mathscr{P}}(F) \le I(F).$$

Suppose that (3.26) were not true. Then there exists an $\eta < \infty$ such that

$$\sup_{\mathscr{P}} I_{\mathscr{P}}(F) < \eta < I(F). \tag{3.27}$$

Thus, for each partition $\mathscr{P} = \{0 = t_0 < t_1 < \cdots < t_k = 1\}$, we can find $f_{\mathscr{P}}$ in $\mathscr{F}$ such that $I_{\mathscr{P}}(f_{\mathscr{P}}) < \eta$. The support of such an $f_{\mathscr{P}}$ will be contained in the support of α. Let $\hat{f}_{\mathscr{P}}$, called the $\mathscr{P}$-linear form of $f_{\mathscr{P}}$ with respect to α, be defined by

$$\hat{f}_{\mathscr{P}}(A) = \sum_{i=2}^{k} \frac{f_{\mathscr{P}}(t_i) - f_{\mathscr{P}}(t_{i-1})}{\alpha(t_i) - \alpha(t_{i-1})} \alpha(A \cap (t_{i-1}, t_i]) + \frac{f_{\mathscr{P}}(t_1)}{\alpha(t_1)} \alpha(A \cap [0, t_1]),$$

Then $\hat{f}_{\mathscr{P}}(t_i) = f_{\mathscr{P}}(t_i)$, $0 \le i \le k$, and

$$I_{\mathscr{P}}(f_{\mathscr{P}}) = I_{\mathscr{P}}(\hat{f}_{\mathscr{P}}) = I(\hat{f}_{\mathscr{P}}).$$

Hence $\{\hat{f}_{\mathscr{P}}\}$ is a net in the set Γ_η which is compact from Lemma 3.4. Thus, there is a cluster point f_0 of this net and $I(f_0) \le \eta$ from the lower semicontinuity of I. If we can show that f_0 is a cluster point of $\{f_{\mathscr{P}}\}$, it will follow that f_0 belongs to F since F is closed. Since $I(f_0) \le \eta$, this will lead to a contradiction of (3.27), and the conclusion (3.26) would have been established.

Let $\mathscr{P}' = \{0 = t_1', t_2', \dots, t_l'\}$ be a partition consisting of continuity points of f_0. Fix $\varepsilon > 0$, and let $N_{\mathscr{P}', \varepsilon}$ be a weak neighborhood of f_0 defined by

$$N_{\mathscr{P}', \varepsilon} = \left\{f\colon \max_i |f(t_i') - f_0(t_i')| < \varepsilon\right\}.$$

Let $\mathscr{P}'' > \mathscr{P}'$. Since f_0 is a cluster point of $\{\hat{f}_{\mathscr{P}}\}$, there is a partition $\mathscr{P} > \mathscr{P}''$ such that $\hat{f}_{\mathscr{P}} \in N_{\mathscr{P}', \varepsilon}$. Since $\hat{f}_{\mathscr{P}}$ and $f_{\mathscr{P}}$ agree on the partition $\mathscr{P}$, it follows that $f_{\mathscr{P}} \in N_{\mathscr{P}', \varepsilon}$ and that f_0 is a cluster point of $\{f_{\mathscr{P}}\}$. This completes the proof of Theorem 3.4. □

Theorems 3.2 and 3.5 and Lemmas 3.3 and 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 nonnegative random variable X satisfying (3.1). When these results are applied in the Sections 4 and 5 we will restrict X to be nonnegative 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 $J(0) = \infty$ for the Gamma distribution and $J(0) = \mu$ for the Poisson distribution with parameter μ. On the other hand, $C_1 = \infty$ 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 Lemma 3.4 and Theorem 3.5 in the Skorohod topology wherein the distribution functions f are considered as elements of $\mathscr{D}[0,1]$. 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 Γ_c is Skorohod closed. However, Γ_c is not Skorohod compact as the following example demonstrates. Let

$$f_n(t) = \begin{cases} t, & 0 \le t \le \dfrac{1}{2} - \dfrac{1}{n}, \\ t + n\left(t - \dfrac{1}{2} + \dfrac{1}{n}\right), & \dfrac{1}{2} - \dfrac{1}{n} < t \le \dfrac{1}{2}, \\ t + 1, & \dfrac{1}{2} < t \le 1. \end{cases}$$

Let $J(a) = a - 1 - \log a$, which is the rate function corresponding to the Gamma distribution with shape parameter 1. Let α 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

$$f(t) = \begin{cases} t, & t < \frac{1}{2}, \\ t + 1, & \frac{1}{2} < t \le 1. \end{cases}$$

Since f_n is continuous and f has a jump at $t = \frac{1}{2}$, no subsequence of f_n can converge in the Skorohod topology. Thus Γ_1 is not Skorohod compact.

4. LD rates for stochastic processes with stationary and nonnegative independent increments. Let $\{X(t),\ 0 \le t \le 1\}$ be a stochastic process with stationary and nonnegative independent increments and measurable sample paths with $X(0) = 0$. Since the increments are nonnegative, the sample paths of $\{X(t),\ 0 \le t \le 1\}$ can be considered as members of $M[0,1]$. Note that $X(1)$ is a nonnegative infinitely divisible random variable.

We will assume that

$$\phi(\theta) = E(e^{\theta X(1)}) < \infty, \tag{4.1}$$

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 the rate function $I(f)$ on $M[0,1]$ be as defined in (3.15). For $\lambda > 0$, define

$$Z_\lambda(t) = \frac{1}{\lambda} X(\lambda\alpha([0,t])), \qquad 0 \le t \le 1, \tag{4.2}$$

where α is a probability measure on $[0,1]$, which may be considered as a time deformation. Then $\{Z_\lambda(t), 0 \le t \le 1\}$ is a process with values in $M[0,1]$. Endow $M[0,1]$ with the weak topology and denote the induced distribution of $\{Z_\lambda(t), 0 \le t \le 1\}$ by P_λ. In this section we show that $\{P_\lambda\}$ is LD tight (Lemma 4.3) and satisfies the LDP with rate function $I(f)$ (Theorems 4.1 and 4.2).

THEOREM 4.1. *Let F be a weakly closed subset of $M[0,1]$. Then*

$$\limsup \frac{1}{\lambda} \log P_\lambda(F) \le -I(F). \tag{4.3}$$

PROOF. Let $\mathscr{P} = \{0 = t_0 < t_1 < \cdots < t_k = 1\}$ be a partition and let

$$A_i = \begin{cases} [0, t_1], & \text{if } i = 1, \\ (t_{i-1}, t_i], & \text{if } i = 2, \dots, k. \end{cases} \tag{4.4}$$

Let

$$W_{\lambda, i} = Z_\lambda(A_i), \qquad 1 \le i \le k. \tag{4.5}$$

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

$$\sum_i J\left(\frac{x_i}{\alpha(A_i)}\right)\alpha(A_i). \tag{4.6}$$

Now,

$$P_\lambda(F) = P(Z_\lambda \in F) \le P\{I_{\mathscr{P}}(Z_\lambda) \ge I_{\mathscr{P}}(F)\},$$

where

$$I_{\mathscr{P}}(Z_\lambda) = \sum_i J\left(\frac{W_{\lambda,i}}{\alpha(A_i)}\right)\alpha(A_i).$$

Since the convex hull of the support of $X(1)$ is $[0,\infty)$ the function $J(x)$ is continuous in $[0,\infty)$ and $J(x) \to \infty$ as $x \to \infty$. Thus the set

$$\left\{(x_1, \dots, x_k): \sum_i J\left(\frac{x_i}{\alpha(A_i)}\right)\alpha(A_i) \ge I_{\mathscr{P}}(F)\right\}$$

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

$$\limsup \frac{1}{\lambda} \log P_\lambda(F) \le -I_{\mathscr{P}}(F).$$

Since $\mathscr{P}$ is arbitrary, we can use the minimax result in Theorem 3.5 to obtain

$$\limsup \frac{1}{\lambda} P_\lambda(F) \leq -I(F). \qquad \square$$

THEOREM 4.2. *Let G be a weakly open subset of $M[0,1]$. Then*

$$\liminf \frac{1}{\lambda} \log P_\lambda(G) \geq -I(G). \tag{4.7}$$

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 $\mathscr{P} = \{0 = t_0 < t_1 < \cdots < t_k = 1\}$ consisting of continuity points of f and α such that the neighborhood

$$N_{\mathscr{P}, \varepsilon} = \Big\{g\colon \max_i |g(A_i) - f(A_i)| < \delta\Big\},$$

of f is contained in G. Here $A_1, \dots, A_k$ are as defined in (4.4). Thus,

$$P_\lambda(G) \geq P_\lambda\Big\{\max_i |W_{\lambda, i} - f(A_i)| < \delta\Big\},$$

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, $G^* = \{(x_1, \dots, x_k)\colon \max_i |x_i - f(A_i)| < \delta\}$ is open in $\mathscr{R}^k$. Thus

$$\liminf \frac{1}{\lambda} \log P_\lambda(G) \geq -\inf \sum_i J\left(\frac{x_i}{\alpha(A_i)}\right)\alpha(A_i),$$

where infimum is taken over the set G^*. Hence,

$$\liminf \frac{1}{\lambda} \log P_\lambda(G) \geq -I_{\mathscr{P}}(f) \geq -I(f) \geq -I(G) - \varepsilon.$$

This completes the proof of Theorem 4.2. $\square$

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

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

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. $\square$

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 [e.g., Breiman (1968), Chapter 14] it follows that

$$\psi(\theta) = \log \phi(\theta) = \int (e^{\theta x} - 1)\, d\nu(x),$$

where the Lévy measure ν (possibly unbounded) satisfies $\int |x|\, d\nu(x) < \infty$ and that the sample paths of $\{X(t), 0 \le t \le 1\}$ 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 nonnegative independent increments. The Lévy measurements 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, $J(a) = \sup_\theta\{\theta a - \psi(\theta)\}$ and $J^{(i)}(a) = \sup_\theta\{\theta a - \psi^{(i)}(\theta)\}$ where $\psi^{(i)}(\theta) = \int (e^{\theta x} - 1)\, d\nu^{(i)}(x)$ is the cumulant generating function of $X^{(i)}$, $i = 1,2$.

Let α be a probability measure on [0,1]. For $\lambda > 0$, define

$$Z_\lambda(t) = \lambda^{-1} X(\lambda \alpha([0,t])), \quad \text{for } 0 \le t \le 1, \tag{5.1}$$

where α is a probability measure on [0,1]. This α may be considered as a time deformation. 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). We now have

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

Note that $\{Z_\lambda(t)\colon 0 \le t \le 1\}$ 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 $h^{(1)}, h^{(2)} \in M[0,1]$. Consider an arbitrary decomposition of f given by

$$f = f^{(1)} - f^{(2)},$$

where $f^{(1)}, f^{(2)} \in M[0,1]$. For any function p in BV[0,1], let $p = p_1 + p_2$ where $p_1 \ll \alpha$ and $p_2 \perp \alpha$ and let $\dot{p}_1 = dp_1/d\alpha$. It is clear that

$$\begin{aligned} f_1 &= h_1^{(1)} - h_1^{(2)} = f_1^{(1)} - f_1^{(2)}, \\ \dot{f}_1 &= \dot{f}_1^{(1)} - \dot{f}_1^{(2)} \end{aligned} \tag{5.3}$$

and

$$\begin{aligned} &\inf\{ f_2^{(i)}[0,1]\colon f = f^{(1)} - f^{(2)};\ f^{(1)}, f^{(2)} \in M[0,1]\} \\ &\quad = h_2^{(i)}([0,1]), \qquad i = 1,2. \end{aligned} \tag{5.4}$$

The definitions of $\dot{f}_1$, $h_2^{(1)}$, $h_2^{(2)}$ 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_λ be the probability distribution of $\{Z_\lambda(t),\ 0 \le t \le 1\}$. Then P_λ satisfies the LD principle with the rate function*

$$I(f) = \int J(\dot{f}_1)\, d\alpha + C_1 h_2^{(1)}([0,1]) + C_2 h_2^{(2)}([0,1]), \tag{5.5}$$

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

PROOF. Let $P_\lambda^{(i)}$ be the distribution of $Z_\lambda^{(i)}(\cdot)$ in $M[0,1]$, $i = 1,2$. Let g be a function from $M[0,1] \times M[0,1]$ into $\mathrm{BV}[0,1]$ defined by $g(f^{(1)}, f^{(2)}) = f^{(1)} - f^{(2)}$. Then g is a continuous function and $P_\lambda = (P_\lambda^{(1)} \times P_\lambda^{(2)})g^{-1}$. From Theorems 4.1, 4.2 and Lemma 4.3, $P_\lambda^{(i)}$ is LD tight and satisfies the LDP with rate function

$$I^{(i)}(f) = \int J^{(i)}(\dot{f}_1)\, d\alpha + C^{(i)} f_2([0,1]),$$

where $f = f_1 + f_2$ with $f_1 \ll \alpha$ and $f_2 \perp \alpha$ and $\dot{f}_1 = df_1/d\alpha$ and where $C^{(i)}$ 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_λ satisfies the LDP with rate function

$$\begin{aligned}
\inf_{f^{(1)}, f^{(2)}:\ f = f^{(1)} - f^{(2)}} \Big\{ \int J^{(1)}(\dot{f}_1^{(1)})\, d\alpha &+ \int J^{(2)}(\dot{f}_1^{(2)})\, d\alpha \\
&+ C^{(1)} f_2^{(1)}([0,1]) + C^{(2)} f_2^{(2)}([0,1]) \Big\} \\
= \int J(\dot{f}_1)\, d\alpha &+ C_1 h_2^{(1)}([0,1]) + C_2 h_2^{(2)}([0,1]),
\end{aligned}$$

in view of (5.3), (3.9) 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 $\{X(t),\ 0 \le t \le 1\}$ be a Poisson process with constant intensity μ. Define the process $\{Z_\lambda(t),\ 0 \le t \le 1\}$ as in (4.2). Then $\{\lambda Z_\lambda(t),\ 0 \le t \le 1\}$ is a Poisson process with intensity function $\lambda\mu\alpha([0,t])$. The distribution of $X(1)$ is Poisson with parameter μ and thus

$$J(a) = a \log\frac{a}{\mu} - a + \mu \quad \text{and} \quad C_1 = \infty,$$

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, $\{Z_\lambda(t),\ 0 \le t \le 1\}$ satisfies the LDP with rate function

$$I(f) = \begin{cases} \int \dot{f}_1 \log\left(\dfrac{\dot{f}_1}{\mu}\right) d\alpha + \mu - f([0,1]), & \text{if } f \ll \alpha, \\ \infty, & \text{otherwise.} \end{cases} \tag{6.1}$$

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

Example 2—Gamma processes. Let $\{X(t),\ 0 \le t \le 1\}$ 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

$$J(a) = a - 1 - \log a, \qquad J(0) = \infty \quad \text{and} \quad C_1 = 1,$$

where $J(a)$ and C_1 are as defined in (3.2) and (3.6). Thus the process $\{Z_\lambda(t),\ 0 \le t \le 1\}$ as defined in (4.2) satisfies the LDP with

$$I(f) = \begin{cases} f([0,1]) - 1 - \int \log \dot{f}_1 \, d\alpha, & \text{if } f_1 \equiv \alpha, \\ \infty, & \text{otherwise.} \end{cases}$$

Example 3—Dirichlet processes. Consider the process $\{W_\lambda(t),\ 0 \le t \le 1\}$ where $W_\lambda(t) = Z_\lambda(t)/Z_\lambda(1)$ where Z_λ is as defined in Example 2. Then $\{W_\lambda(t),\ 0 \le t \le 1\}$ 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_λ 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 α. However, if we let $\lambda \to \infty$, then W_λ converges to the constant α 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

$$I(f) = \begin{cases} K(\alpha, f), & \text{if } f(1) = 1 \text{ and } f_1 \equiv \alpha, \\ \infty, & \text{otherwise,} \end{cases}$$

where $K(\alpha, f)$ is the Kullback–Leibler information number between two probability measures α and f defined by

$$K(\alpha, f) = -\int \log \frac{df}{d\alpha}\, d\alpha.$$

REFERENCES

Azencott, R. (1980). Grandes deviations et applications. *École d'Été de Probabilités de Saint-Flour VIII, 1978. Lecture Notes in Math.* **774** 1–176. Springer, Berlin

Bahadur, R. R. (1960a). Stochastic comparison of tests. *Ann. Math. Statist.* **31** 276–295.

Bahadur, R. R. (1960b). Asymptotic efficiency of tests and estimates. *Sankhyā* **22** 229–252.

Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38** 303–324.

Bahadur, R. R. (1971). *Some Limit Theorems in Statistics*. SIAM, Philadelphia.

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

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

BREIMAN, L. (1968). *Probability.* Addison-Wesley, Reading, Mass.

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.

CRAMÉR, H. (1937). Sur un nouveau theoreme limite de la theorie des probabilites. *Colloquium on Theory of Probability.* Hermann, Paris.

ELLIS, R. S. (1984). Large deviations for a general class of random vectors. *Ann. Probab.* **12** 1–12.

ERDÖS, P. and RÉNYI, A. (1970). On a new law of large numbers. *J. Analyse Math.* **23** 103–111.

FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.

GROENEBOOM, P., OOSTERHOFF, J. and RUYMGAART, F. H. (1979). Large deviation theorems for empirical probability measures. *Ann. Probab.* **7** 553–586.

LANFORD, O. E. (1973). Entropy and equilibrium states in statistical mechanics. In *Statistical Mechanics and Mathematical Problems. Battelle Seattle 1971 Recontres* (A. Lenard, ed.). *Lecture Notes in Phys.* **20** 1–113. Springer, Berlin.

RUELLE, D. (1969). *Statistical Mechanics. Rigorous Results.* Benjamin, New York.

SETHURAMAN, J. and TIWARI, R. C. (1982). Convergence of Dirichlet measures and the interpretation of their parameter. In *Statistical Decision Theory and Related Topics III* (S. S. Gupta and J. O. Berger, eds.) **2** 305–315. Academic, New York.

SHEPP, L. A. (1964). A limit theorem concerning moving averages. *Ann. Math. Statist.* **35** 424–428.

VARADHAN, S. R. S. (1966). Asymptotic probabilities and differential equations. *Comm. Pure Appl. Math.* **19** 261–286.

VARADHAN, S. R. S. (1984). *Large Deviations and Applications.* SIAM, Philadelphia.

VENTSEL, A. D. (1976). Rough limit theorems on large deviations for Markov processes, I, II. *Theory Probab. Appl.* **21** 227–242, 499–512.

DEPARTMENT OF STATISTICS
UNIVERSITY OF SOUTH CAROLINA
COLUMBIA, SOUTH CAROLINA 29208

DEPARTMENT OF STATISTICS
FLORIDA STATE UNIVERSITY
TALLAHASSEE, FLORIDA 32306