"Lynch & Sethuraman (circa 1989) Rough Draft of "Functional Erdös-Renyi LLN." (unpublished)

1. Introduction Let $X_1, X_2, \ldots$ be i.i.d. random variables and let $S_0 = 0, S_n = X_1 + \cdots + X_n, n \geq 1$. The step function and broken line versions of the sample average process may be defined as follows:

$$\{U_n(t) = \frac{S_{[nt]}}{n}, t \geq 0\}$$

and

$$\{V_n(t) = \frac{S_{[nt]}}{n} + \frac{(nt - [nt])X_{[nt]+1}}{n}, t \geq 0\}.$$

These sample average processes can be considered as processes in $BV[0,1]$, the space of functions of bounded variation in $[0,1]$.

Let $\phi(\theta) = E(exp(\theta X_1)$ be finite for θ in a neighborhood of 0 and let $\psi(\theta) = \log\phi(\theta)$. Let $J(a) = \sup[\theta a - \psi(\theta)]$ be the large deviation rate function of X_1. Let $C_1 = \lim_{a\to\infty} \frac{J(a)}{a}$ and $C_2 = \lim_{a\to-\infty} \frac{J(a)}{|a|}$. One can view $\{U_n(t)\}$ and $\{V_n(t)\}$ as processes in $BV[0,1]$.

The purpose of this paper is to prove that the large deviation principle (LDP) holds for both these sample average processes with the same large deviation rate (LDR) function $I(f) = \int J(\dot{f})dt + C_1 f_{1s}[0,1] + C_2 f_{2s}[0,1]$ where $\dot{f} = \frac{df_a}{dt}$, f_a is the absolutely continuous part of f, $f = f_1 - f_2$ is the Hahn decomposition of f, and f_{1s} and f_{2s} are the singular parts of f_1 and f_2, respectively. The appropriate topology in $BV[0,1]$ that is necessary for the LDP depends on whether C_1 and C_2 are finite or infinite.

These results are used to obtain the functional forms of the Erdõs- Rènyi and Shepp laws for the sample average processes. More specifically, Let

$$\Delta_{m,n,a}(s) = \frac{S_{m+[s\log n/a]} - S_m}{[\log n/a]}$$

and

$$\frac{\Delta'_{m,n,a}(s) = S_{m+[s\log n/a]} + (s\log n/a - [s\log n/a])X_{m+[s\log n/a+1]} - S_m}{[\log n/a]}$$

for $0 \leq s \leq 1$. Let

$$\Gamma_a = \{f : I(f) \leq a\}.$$

We prove that with probability 1 the set of cluster points of $\Delta_{m,n,a}(\cdot)$ is Γ_a, and for each $\epsilon > 0$

$$\{\Delta_{m,n,a}(\cdot), m \leq n\} \subset (\Gamma_a)_\epsilon \text{ eventually}$$

and

$$(\Gamma_a) \subset (\Delta_{m,n,a}(\cdot), m \leq n\})_\epsilon.$$

These results also hold for $\Delta'_{m,n,a}(\cdot)$.

Sethu,
my grant #'s for this are

ARO DAAG29-~~[illegible]~~ 84-K-0007
ARO DAAL03-86-K-0010
ARO DAAL03-87-K-0101

2. The Large Deviation Principle - Some General Results Let $\mathcal{X}$ be a topological space and $\mathcal{F}$ be the Borel σ - field in $\mathcal{X}$. Let $\{P_n\}$ be a family of probability measures on $(\mathcal{X}, \mathcal{F})$. The following definitions which are slight variants of Varadhan (1984) and which may be found in Lynch and Sethuraman (1987) allow us to state many large deviation results in a concise form.

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

(2.1) $$0 \leq I(x) \geq \infty,$$

(2.2) $$I(\cdot) \text{ is lower semi}-\text{continuous (lsc) and}$$

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

For any subset $A \subset \mathcal{X}$, define

(2/4) $$I(A) = \inf_{x \epsilon A} I(x).$$

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

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

(2.6) $$\text{for each closed set } F, \limsup(1/n) \log P_n(F) \leq -I(F), \text{ and}$$

(2.7) $$\text{for each open set } G, \liminf(1/n) P_n(G) \geq -I(G).$$

Definition 2.3 The measures $\{P_n\}$ satisfy the *weak large deviation principle (WLDP or WLD principle)* with rate function $I(\cdot)$ if (2.5) and (2.7) of Definition 2.2 together with (2.8) below are satisfied:

(2.8) $$\text{for each compact set } K, \limsup(1/n) \log P_n(K) \leq -I(K).$$

Definition 2.4 The measures $\{P_n\}$ are *large deviation tight (LD tight)* if, for each $M > \infty$, ther exists a compact set K_M such that

(2.9) $$\limsup(1/n) \log P_n(K_M^c) \leq -M.$$

The following lemma found in Lynch and Sethuraman (1987) demonstrates the usefulness of LD tightnesss.

Lemma 2.5. Let $\{P_n\}$ be LD tight and satisfy the WLDP. Then it satisfies the LDP.

A useful consequence of the LDP is the contraction principle stated below.

The contraction principle Let $\{P_n\}$ satisfy the LDP with rate function $I(\cdot)$. Let h_n be a continuous map from $\mathcal{X}$ into a topological space $\mathcal{Y}$ and let $Q_n = P_n h_n^{-1}$. If $h_n \to h$ uniformly on compact subsets of $\mathcal{X}$, then the measures Q_n satisfy the LDP with rate function $K(y) = \inf_{x:h(x)=y} I(x)$.

The LDP, along with the above results, has the flavor of weak convergence of probability measures (Theorems 2.1 and 5.5 of Billingsley(1968)). The following lemma is the analogue of the converse part of Prohorov's theorem (Billingsley (1968) Theorem 6.2) and with Lemma 2.5 shows that for Polish spaces the LDP is equivalent to the WLdo and LD tightness. The proof is similar to Billingsley's proof of Prohorov's theorem and can be found in Lynch and Sethuraman (1987).

Lemma 2.6 If $\{P_n\}$ is a sequence of probability measures which satisfies the LDP, then $\{P_n\}$ is LD tight.

The following results (see Lynch and Sethuraman (1987)) show how the LD properties for marginal measures carry over to product spaces. These results are used in Section XX. Let $\{P_n^i\}$ be a sequence of probability measures on a topological space $\mathcal{X}^{(i)}$, $i = 1.2$. Let $Pn = Pn^1 \times P_n^2$ be the product measure on $\mathcal{X} = \mathcal{X}^1 \times \mathcal{X}^2$.

Lemma 2.8 If $\{P_n^i\}$ is LD tight for $i = 1, 2$, then $\{P_n\}$ is LD tight.

Lemma 2.9 Let $\{P_n^i\}$ satisfy the WLDP with rate function $I^i(x^i), i = 1, 2$. Then $\{P_n\}$ satisfies the WLDP with rate function $I(x_1, x_2) = i^1(x_1) + I^2(x_2)$.

Corollary 2.10 Let $\{P_n^i\}$ be LD tight and satisfy the WLDP, $i = 1, 2$. Then P_n satisfies the LDP with rate function $I(x_1, x_2) = i^1(x_1) + I^2(x_2)$.

3. The LDP for averages of i.i.d. random variables We begin with some well known facts about the LDP of averages of i.i.d. random variables. Let X be a real valued random variable and let

$$\phi(\theta) = E(exp(\theta x)) < \infty \tag{3.1}$$

for θ in some open interval about zero. Let $\psi(\theta) = \log \phi(\theta)$. Since the mean of X exists and is finite, we will without loss of generality assume that $E(X) = 0$. Let $X_1, X_2, \ldots$ be i.i.d. copies of X and let P_n be the distribution of $\overline{X}_n = (X_1 + \cdots + X_n)/n$. The following is the oldest theorem in large deviation theory and is variously referred to as Cramer's Theorem and Chernoff's Theorem.

Theorem 3.1 (Cramer (1937), Chernoff (1952)). The distribution $\{P_n\}$ satisfy the LDP with rate function $J(a)$ given by

$$J(a) = \sup\{\theta a - \psi(\theta)\} \tag{3.2}$$

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

$$0 \le J(a) \le \infty, J(0) = 0 \text{ (since } E(X) = 0\text{), and } J(a) \to \infty \text{ as } |a| \to \infty, \tag{3.3}$$

$$J(a) = \sup_{\theta \ge 0}[\theta a - \psi(\theta)], \text{ for } a > 0 \tag{3.4}$$

$$J(a) \text{ is convex.} \tag{3.5}$$

$$J(a)/a \text{ is nondecreasing.} \tag{3.6}$$

$$\lim_{a \to \infty} J(a)/a = C_+ \text{ and } \lim_{a \to -\infty} J(a)/|a| = C_- \tag{3.7}$$
exists, where $0 < C_+, C_- \le \infty$ (assume $X_1 \not\equiv 0$).

$$J(a) \text{ is continuous on } \Gamma = \overline{\{a : J(a) < \infty\}}. \tag{3.8}$$

The endpoints of Γ are $-A =$ ess inf X and $B =$ ess sup X and if $A < \infty (B < \infty)$ then $J(-A) = -\log P(X = -A)(J(B) = -\log P(X = B))$ and J is right (left) continuous at $-A(B)$.

Since $X_1 = X_1^+ - X_1^-$ the LDP for X_1 can also be obtained from the LDP for the bivariate (X_1^+, X_1^-) and the contradiction principle. This leads to the LD decon in 1 din. In one dimension?

In addition to Theorem 3.1, we will also need to establish the LDP for the positive and negative parts of $\overline{X}_n$.

Let $X^+ = \max(0, X)$ and $X^- = \max(0, -X)$. Let $\phi(\theta_1, \theta_2) = E(e^{\theta_1 X^+ + \theta_2 X^-})$ and $\psi(\theta_1, \theta_2) = \log \phi(\theta_1, \theta_2)$. Let Q_n be distribution of $M_n, N_n)$ where $M_n = (X_1^+ + \ldots + X_n^+)/n$ and $N_n = (X_1^- + \ldots + X_n^-)/n$.

We now establish the multivariate version of Theorem 3.1.

Let $\mathbf{Y} = (Y_1, \ldots, Y_k)$ and let $\phi(\theta) = E(\exp \sum_1^k \theta_i Y_i) < \infty$ for θ in some open set about 0. Let $\psi(\theta) = \log \phi(\theta)$ and let

$$L(\mathbf{a}) = \sup_\theta \{\sum_1^k \theta_i a_i - \psi(\theta)\}. \tag{3.9}$$

Let $\mathbf{Y}_1, \mathbf{Y}_2, \ldots$ be iid copies of $\mathbf{Y}$ and let Q_n denote the probability measure induced on R^k by $(\mathbf{Y}_1 + \cdots + \mathbf{Y}_n)/n$.

Theorem 3.2. The measures Q_n satisfy the LDP with rate function L.

Proof. By the results in Sections 1, 2, and 3 of Bahadur and Zabell (1979) (in particular, see Theorems 2.1, 2.2, 2.3, 3.1, and 3.2, Lemma 2.5 and (1.6), (1.7), (1.12), (3.3) and (3.4), there) it follows that $\{Q_n\}$

satisfies the weak LD principle (in fact5, $\lim \frac{1}{n} \log Q_n(G) = -\mathbf{L}(G) = -\inf_G L(a)$ if G is open and convex). Since by Theorem 3.1, the marginals satisfy the LDP, it follows from Lemma 2.9 and 2.6 that $\{Q_n\}$ is LD tight. The theorem follows ~~thus and~~ from Lemma 2.5 since $\{Q_n\}$ satisfies the weak LDP and is LD tight.

Let $X^+ = \max(0, X)$ and $X^- = \max(0, -X)$. Let $\phi(\theta_1, \theta_2) = E(\exp \theta_1 X^+ + \theta_2 X^-)$, $\psi(\theta_1, \theta_2) = \log \phi(\theta_1, \theta_2)$.

Let Q_n be the distribution of (M_n, N_n) where $M_n = (X_1^+ + \dots X_n^+)/n$ and $N_n = (X_1^- + \dots, X_n^-)/n$. Then, by Theorem 3.2, $\{Q_n\}$ satisfies the LDP. In addition to this we will need to relate the LD rate of $\{Q_n\}$ to J given in (3.2). This is given next.

Theorem 3.3 The distributions $\{Q_n\}$ satisfy the LDP with rate function $L(a_1, a_2)$ given by

$$L(a_1, a_2) = \sup_{\theta_1, \theta_2} \{\theta_1 a_1 + \theta_2 a_2 - \psi(\theta_1, \theta_2)\}. \tag{3.9}$$

Furthermore, for $J(a)$ given by (3.2),

$$J(a) = \inf\{L(b, c) \colon \ b, c \geq 0 \text{ and } b - c = a\} \tag{3.10}$$

and there exists $b \geq 0$ and $c \geq 0$ and $b - c = a$ with

$$J(a) = L(b, c). \tag{3.11}$$

Proof We've already indicated that $\{Q_n\}$ satisfies the LDP. The rate is given by (3.9).

To see (3.10) and (3.11), without loss of generality let $a \geq 0$. Let P_n denote the distribution of (M_n, N_n). Since $O_a = \{(b, c) \colon \ b - c > a\}$ and (a, ∞) are open convex sets and since there is equality in (2.7) for such sets for both $\{P_n\}$ and $\{Q_n\}$,

$$\begin{aligned} -L(O_a) &= \lim \frac{1}{n} \log Q_n(O_a) \\ &= \lim \frac{1}{n} \log P_n((a, \infty)) \\ &= -J((a, \infty)). \end{aligned}$$

Let $a \in [0, B]$ from (3.3) and (3.5), $J(a)$ is strictly increasing on $[0, B)$. Consequently, by (3.8), $J(a) = J((a, \infty)) = L(O_a)$. Thus for $a \in [0, B)$, there exists a_n, b_n and c_n with $a_n \downarrow a, b_n - c_n = a_n, b_n$ and $c_n \geq 0$ and $L(b_n, c_n) \to J(a)$. Since $\{(b, c) \colon \ L(b, c) \leq J(a) + 1\}$ is compact there exists b and c such that $b_n \to b$ and $c_n \to c$. So $b - c = a$ and $J(a) \leq L(b, c) \leq \lim L(b_n, c_n) = J(a)$ since L is lsc.

If $a = B < \infty$, then $J(a) = \log P(x = a) = \log P(X^+ = a) = L(a, 0)$. Finally, if $B < \infty$ and $a > B$, then $J(a) = \infty = K(b, c)$ for all $b - c = a$.

The final result of this section establishes a decomposition $a = (a + c(a)) - c(a)$ of a which is referred to as the (univariate) LD decomposition of a. This decomposition is the basis for the LD decomposition in Section 5 for the sample average process.

Lemma 3.3 Let $\Gamma = \{a \colon \ J(a) < \infty\}$. There exists a Borel measurable function $c(a)$ on Γ such that $c(a)$ and $a + c(a)$ are nonnegative and satisfy

$$L(a + c(a), c(a)) = J(a). \tag{3.12}$$

Proof. Let $U_n = \{\dots < a_j^{(n)} < a_{j+1}^{(n)} < \dots\}$ $n = 1, 2, \dots$ be a sequence of partitions with $U_n \subset U_{n+1}$. Let $A_n(a) = a_j^{(n)}$ for $a \in [a_j^{(n)}, a_{j+1}^{(n)})$ and let $C_n(a) = c_j^{(n)}$ where $0 \leq b = a_j^{(n)} + c_j^{(n)}$ is a solution of (3.11) with $a = a_j^{(n)}$. Furthermore, since $U_n \subset U_{n+1}$ we can choose $c_j^{(n+1)} = c_j^{(n)}$ if $a_j^{(n+1)} = a_j^{(n)}$.

Let $c(a) = \lim C_n(a)$ and note that $\lim A_n(a) = a$. Thus $c(a)$ and $a + c(a)$ are nonnegative. Since $J(a)$ is continuous on Γ and L is lsc,

$$\begin{aligned} J(a) &= \lim_n J(A_n(a)) \\ &= \lim L(A_n(a) + C_n(a), C_n(a)) \\ &\geq L(a + c(a), c(a)) \geq J(a) \end{aligned}$$

where the inequality follows from (3.10).

Theorems 3.1 and 3.2 establish the LDP for sums of iid random variables and random vectors. In Section , the LDP is established for the sample average process. The rate is defined in terms of J, C_+ and C_- (see (3.2) and (3.7)) as follows: For f of bounded variation,

$$I(f) = \int_0^1 J(\dot{f})\partial t + C_+ F_s^+[0,1] + C_- f s^-[0,1]$$

where $\dot{f} = \frac{\partial f_a}{\partial t}, f = f_a + f_s$ is the Lebesque decomposition of f and $f_s = f_s^+ - f_s^-$ is the Hahn-Jordan decomposition of f_s, the singular part of f.

In the next section, useful properties of the point function $I(f)$ and of the set function $I(A) = \inf_{f \in A} I(f)$ are established.

4. The Rate Function I. In This section we enumerate some facts about the rate function $I(\cdot)$ defined in (4.3) below. These will be needed in susbsequent sections. To properly define $I(\cdot)$ we need the following digression.

Let $BV[0,1]$ denote the space of finite signed measures on $([0,1], B)$ where B is the usual Borel θ-field in $[0,1]$. For each $f \in BV[0,1]$ associate the function $f(t)$ of bounded variation on $[0,1]$ given by letting $f(0)=0, f(t)=f([0,t))$ for $0<t\le 1$, and $f(1)=f([0,1))$. We use the same symbol f to denote both the measure $f(A)$ and the function $f(t)$. Note that with this identification, $BV[0,1] \subset D[0,1]$, the space of functions which are left continuous on $(0,1]$ and have right hand limits. Here for convenience we have slightly altered the usual definition of $D[0,1]$. (See Billingsley (1968), Section 14.) We note, though, that all results concerning the Skorahod topology hold with this definition of $D[0,1]$. Let $C[0,1]$ denote the space of continuous functions on $[0,1]$.

For $f \in BV[0,1]$, $f = f_+ - f_-$ will denote any arbitrary decomposition of f into the difference of two finite measures while $f = h_+ - h_-$ will denote the Hahn - Jordan decomposition. Let α denote Lebesgue measure on $[0,1]$ and for $j = +$ or $-$, let $h_j = h_j^a + h_j^s$ be Lebesgue decomposition of h_j with respect to α with $h_j^a << \alpha$ and $h_j^s \perp \alpha$. Let $\dot f = \dot h_+ - \dot h_-$ where $\dot h_j$ is the Radon-Nikodym derivative of h_j^a.

Let $0 = t_0 < t_1 < \ldots < t_k = 1$. Both the collections of points $\{t_0, \ldots, t_k\}$ 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 $\theta(P)$ denote θ-field generated by P. The collection of partitions $\{P\}$ form a directed set under the partial ordering $P' > P$ if $\theta(P)' \supset \theta(P)$. Throughout we will be taking limits of functions indexed by P. These limits will always be along directed nets such that $\theta(P) \to B_0$ (set limit) with $\theta(B_0) = B$. With an abuse of notation we shall write this as $\theta(P) \to B$.

For a function f and a partition $P = \{0 = t_0, t_1, \ldots, t_k = 1\}$, let $\Delta_i t = t_i - t_{i-1}$ and $\Delta_i f = f(t_i) - f(t_{i-1})$. For $\overline{\Gamma}$ as defined in (3.8), let $BV = \{f \in BV[0,1]\colon \ \Delta_i f \in \overline{\Gamma}$ for all partitions $P\}$. For $f \in BV$, let

$$I_p(f) = \Sigma J(\Delta_i f/\Delta_i t)\Delta_i t. \tag{4.1}$$

Denote the restriction of the measure α and the signed measure f to $\theta(P)$ by α_p and f_p, respectively. We may rewrite the definition in (4.1) as

$$I_p(f) = \int J\left(\frac{df_p}{d\alpha_p}\right) d\alpha. \tag{4.2}$$

Let C_+ and C_- be as given in (3.7). For $f \in BV$, let

$$I(f) = \int J(\dot f)d\alpha + C_+ h_+^s[0,1] + C_- H_s^s[0,1]. \tag{4.3}$$

where we adapt the convention that $0 \cdot \infty = 0$. In particular, if $C_+ = \infty = C_-, I(f) = \int J(\dot f) d\alpha$ if f is absolutely continous and $= \infty$ if not.

The following lemma relates $I_p(f)$ to $I(f)$.

Lemma 4.1. $I_p(f) \to I(f)$ as $\theta(P) \to B$.

Proof. For $f = h_+ - h_-$, the Hahn-Jordan decomposition of f, let $b = h_+ + h_-$. Since $h_i << b$ and $h_+ \perp h_-$,

$$\frac{dh_i}{db} = \begin{cases} 1 & \text{a·e·b on } \{\frac{dh_i}{db} > 0\} \\ 0 & \text{a·e·b on } \{\frac{dh_i}{db} > 0\} \quad j \neq i. \end{cases} \tag{4.4}$$

Let $\alpha = \alpha^a + \alpha^s$ be the Lebesgue decomposition of α with respect to b where $\alpha^a << b$ and $\alpha^s \perp b$. For A and B defined in (3.8) let

$$g_+(d) = \begin{cases} dJ(d^{-1}) & \text{if } d > B^{-1} \\ B^{-1}J(B) & \text{if } d = B^{-1} > 0 \\ C_+ & \text{if } d = B^{-1} = 0 \end{cases}$$

and

$$g_-(\partial) = \begin{cases} -\partial J(\partial^{-1}) & \text{if } \partial < -A^{-1} \\ A^{-1}J(-A) & \text{if } \partial = -A^{-1} < 0 \\ C_- & \text{if } \partial = -A^{-1} = 0 \end{cases}$$

Then, since $J(0) = 0$ and $f_p << b_p$,

Since $\frac{\partial f}{\partial b} = \frac{\partial h_+}{\partial b} - \frac{\partial h_-}{\partial b}$, it follows similarly from (4.4) that

$$I(f) = \int_{\frac{\partial f}{\partial b}>0} g_+(\dot{\alpha})\partial b + \int_{\frac{\partial f}{\partial b}>0} g_-(-\dot{\alpha})\partial b \tag{4.6}$$

where $\dot{\alpha}$ is the Radon-Nikodyn derivative of α^a with respect to b.

The following facts are needed concerning $\{\partial h_i p/\partial b_p\}, \{\partial f_p/\partial b_p\}, \{\partial \alpha_p/\partial b_p\}$ and $\{\partial f_p/\partial \alpha_p\}$. Since $h_i << b$, $\{\frac{\partial h_i p}{\partial b_p}, \theta(P); \frac{\partial h_i}{\partial b}, B\}$ is a martingale with $\frac{\partial h_{ip}}{\partial b_p} \to \frac{\partial h_i}{\partial b}$ a.e. b. So, $\{\frac{df_p}{\partial b_p}, \theta(P); \frac{\partial f}{\partial b}, B\}$ is a martingale with

$$\frac{\partial f_p}{\partial b_p} = \frac{\partial H_{+p}}{\partial b_p} - \frac{\partial h_{-p}}{\partial b_p} \to \frac{\partial f}{\partial b} \text{ a.e.} b \tag{4.7}$$

where $\frac{\partial f}{\partial b} = \frac{\partial h_+}{\partial b} - \frac{\partial h_-}{\partial b}$. The collections $\{\frac{\partial h_{ip}}{\partial \alpha_p}, \theta(P)\}$ are also martingales with

$$\frac{\partial H_{ip}}{\partial \alpha_p} \to \dot{h}_i \text{ and } \frac{\partial b_p}{\partial \alpha_p} \to \dot{b} \text{ a.e. } \alpha. \tag{4.8}$$

So, from (4.8), $\{\frac{\partial f_p}{\partial \alpha_p}, \theta(P)\}$ is a martingale

$$\frac{\partial f_p}{\partial \alpha_p} \to \dot{f} \text{ a.e. } \alpha. \tag{4.9}$$

Finally, $\{\frac{\partial \alpha_p}{\partial b_p}, \theta(P)\}$ is a martingale with

$$\frac{\partial \alpha_p}{\partial b_p} \to \dot{\alpha} \text{ a.e. } b. \tag{4.10}$$

Now if $I(f) = \infty$, then by (4.3) either $\int J(\dot{f})\partial\alpha = \infty$ or $C_+h_+^\infty[0,1] + C_-h_-^s[0,1] = \infty$. Since $J(\cdot)$ is lsc, it follows from (4.9) that $\lim J(\frac{\partial f_p}{\partial \alpha_p}) \geq J(\dot{f})$ a.e. α. So in the former case it follows from (4.2) and Fatow's lemma that $\lim I_p(f) = \infty$.

In the latter case, by (4.6),

$$\infty = I(f) = \int_{\frac{\partial f}{\partial b}>0} g_+(\dot{\alpha})\partial b + \int_{\frac{\partial f}{\partial b}<0} g_-(-\dot{\alpha})db$$

Since g_+ is continuous (in the extended sense) on $[B^{-1}, \infty)$ and g_- on $(-\infty, -A^{-1}]$, by another application of Fatow's lemma it follows from (4.4), (4.5), (4.7) and (4.10) that $\lim I_p(f) = \infty$.

Thus to complete the proof we may assume that $I(f) < \infty$ and consider the following three cases:

(i) $C_+ = \infty = C_-, h_+^s = h_-^s = 0$

(ii) $C_+ < \infty$ and $C_- = \infty$ or $C_+ = \infty$ and $C_- < \infty$,

(iii) $C_+ < \infty$ and $C_- < \infty$.

Bold E in Script

Case (i). Since $C_+ = \infty = C_-, h_i << \alpha$ for $i = +$ and $-$. Thus we can adjoin $(\dot{f}, B)$ to the martingale $\{\frac{\partial f_p}{\partial \alpha_p}, \theta(P)\}$ to form a martingale. It follows from (3.5), (3.8) and Jensen's inequality that

$$\int J(\frac{\partial f_p}{\partial \alpha_p})\partial\alpha \leq \int J(\dot{f})\partial\alpha.$$

while from (4.9), the lower semicontinuity of J and Fatow's Lemma

$$\int J(\dot{f})\partial\alpha \leq \lim \int J(\frac{\partial f_p}{\partial \alpha_p})\partial\alpha$$

Case (ii). Without loss of generality assume that $C_+ = \infty$ and $C_- < \infty$. Let $J_+(a) = J(a)$ if $a \geq 0$ and $= 0$ if $a < 0$. Note that

$$J_+ \text{ is nondecreasing and convex on } (-\infty, B], \text{ continuous on } (-\infty, B) \text{ and left continuous at } B \text{ where } B \text{ may equal } \infty. \tag{4.11}$$

Since $C_+ = \infty, h_+ << \alpha$. So, similar to (4.5) and (4.6),

$$I_p(f) = \int J_+(\frac{\partial f_p}{\partial \alpha_p})\partial\alpha + \int -\frac{\partial f_p}{\partial b_p} g_- \left(\left(\frac{\partial f_p}{\partial b_p}\right)^{-1} \frac{\partial \alpha_p}{\partial b_p}\right)\partial b, \frac{\partial f_p}{\partial b_p} < 0 \tag{4.12}$$

and

$$I(f) = \int J_+(\dot{f})\partial\alpha + \int_{\frac{\partial f}{\partial b} < 0} g_-(\dot{\alpha})\partial b. \tag{4.13}$$

Since $h_+ \perp h_-, B \geq \frac{\partial f_p}{\partial \alpha_p} \to \dot{f}$ a.e. α and $\dot{f} = h_+$ a.e. α on $\{\dot{f} > 0\}, \dot{h}_+ \leq B$, and so, $\{\frac{\partial h_{+p}}{\partial \alpha_p}, \theta(P); \dot{h}_+, B\}$ is a martingale because $h_+ << \alpha$, by (4.11) an argument like that in Case (i) shows that

$$\begin{aligned}
\int J_+\left(\frac{\partial f_p}{\partial \alpha_p}\right) &\leq \lim \int J_+\left(\frac{\partial h_{+p}}{\partial \alpha_p}\right)\partial\alpha \\
&= \int J_+(\dot{h}_+)\partial\alpha \\
&= \int J_+(\dot{f})\partial\alpha
\end{aligned}$$

since $\frac{\partial f_p}{\partial \alpha_p} \leq \frac{\partial h_{+p}}{\partial \alpha_p}, \dot{f} = \dot{h}_+$ a.e. α on $\{\dot{f} > 0\} = \{\dot{h}_+ > 0\}$ a.e. α and $J_+(a) = 0$ for $a \leq 0$. By (4.9), (4.11) and Fatow's lemma

$$\int J_+(\dot{f})\partial\alpha \leq \lim \int J_+\left(\frac{\partial f_p}{\partial \alpha_p}\right)\partial\alpha.$$

Thus,

$$\lim \int J_+\left(\frac{\partial f_p}{\partial \alpha p}\right)\partial\alpha = \int J_+(\dot{f})\partial\alpha. \tag{4.14}$$

Since $C_- < \alpha, A = \infty$. Thus,

$$g_- \text{ is continuous on } (-\infty, 0) \text{ left continuous at } 0 \text{ with } 0 \leq g_- \leq C_- \text{ by (3.6).} \tag{4.15}$$

Thus, since $-1 \leq \frac{\partial f_p}{\partial b_p} \leq 0$ with $\frac{\partial f_p}{\partial b_p} \to \frac{\partial f}{\partial b} = -1$ a.e. b on $\{\frac{\partial f}{\partial b} < 0\}$, it follows from (4.10), (4.15) and the bounded convergence theorem that

$$\lim \int_{\frac{\partial f_p}{\partial b_p} < 0} -\frac{\partial f_p}{\partial b_p} g_- \left(\left(\frac{\partial f_p}{\partial b_p}\right)^{-1} \frac{\partial \alpha_p}{\partial b_p}\right)\partial b = \int_{\frac{\partial f}{\partial b} < 0} g_-(-\dot{\alpha})\partial b.$$

Combining (4.14) and (4.16) with the identities in (4.12) and (4.13) completes the proof for Case (ii).

Case (iii). The proof follows from (4.5) and (4.6) by using the argument in the last paragraph of Case (ii).

As an immediate corollary we have

Corollary 4.2. $I(f) = \sup_p I_p(f)$.

Another consequence of Lemma 4.2 is

Lemma 4.3. Let $\{f_n\}_{n=0}^{\infty} \subset BV$. If $f_n \to f_0$ at continuity points of f_0 and $f_n(1) \to f_0(1)$, then $\lim I(f_n) \geq I(f)$.

Proof. Let $P = \{0 = t_0, t_1, \ldots, t_k = 1\}$ be a partition such that $t_1, \ldots t_k$ are continuity points of f_0. Since $f_n(0) = 0$ for $n = 0, 1, 2, \ldots$, it follows that $\Delta_i f_n \to \Delta_i f_0$ for $i = 1, 2, \ldots, k$. Thus, since J is lsc., it follows from Lemma 4.1 and (4.1) that

$$\begin{aligned} \lim I(f_n) &\geq \lim I_p(f_n) \\ &\geq I_p(f_0) \to I(f_0) \end{aligned}$$

as $\theta(P) \to B$ through partitions P with $t_1, \ldots, t_{k-1}$ continuity points of f_0.

The next series of results involve topological considerations. In addition to the appearance or nonappearance of h_j^s in the rate function I when C_j is finite or infinite, the finiteness of C_j will also play a role in these considerations.

Let $M[0,1] \subset BV[0,1]$ denote the space of finite measure on $[0,1]$. Let $\mu_0, \mu_1, \ldots$ be a sequence of measures in $M[0,1]$. We say that $\underline{\mu}_n$ coverges to μ_0 in the weak topology if

$$\int f \partial\mu_n \to \int f \partial\mu_0 \text{ for all } f \in C[0,1].$$

We note that this is equivalent to $\mu_n(+) \to \mu_0(+)$ at continuity points of μ_0 and is better known as convergence in "distribution" amongst statisticians and probabilists.

Let $\Omega = \{\lambda : \lambda \text{ is continuous and strictly increasing on } [0,1] \text{ with } \lambda(0) = 0 \text{ and } \lambda(1) = 1\}$. We say that μ_n converges to μ_0 in the Skorohod topology if there is a sequence $\lambda_1, \lambda_2, \ldots$ in Ω for which

$$\lambda_n(t) \to t \text{ uniformily}$$

and

$$\mu_n(\lambda_n(t)) \to m_0(t) \text{ uniformily.}$$

It is well known that $M[0,1]$ with either topology is separable and metrizable and that the metric can be chosen so that $M[0,1]$ is complete. (See Billinsley (1968), Section 14 for the case of the Skorohod topology.) Let d_w and d_s denote, respectively, these two metrics and note that d_s is a stronger metric than d_w.

Using d_w and d_s we define a metric d on $BV[0,1]$ as follows: For $j = +$ or $-$, let

$$d_j = \begin{cases} d_w & \text{if } C_j < \infty \\ d_s & \text{if } C_j = \infty \end{cases}$$

For $f, g \in BV$, let

$$d(f,g) = \inf\{\max_{j=+,-} d_j(f_j, g_j) : f = f_+ - f_- \text{ and } g = g_+ - g_-\}$$

We mpte tjat d is a metric and that BV with this metric is a Polish space since $M[0,1]$ with the metric d_j is a Polish space.

Since $f_n \to f_0$ in the metric d implies that $f_n \to f_0$ at continuity points of f_0, we have an immediate consequence of Lemma 4.3. that

Lemma 4.4. $I(\cdot)$ is d-lsc.

Let $\Gamma_c = \{f : I(f) \leq c\}$.

Lemma 4.5. Γ_c is d compact.

Proof. We only consider the case when $C_+ = \infty$ and $C_- < \infty$. The proofs of the other cases will follow from the technique of proof for this case.

Let $f \in \Gamma_c$. Then, since $h_+ \perp h_-$ and j is convex with $J(0) = 0$,

$$\begin{aligned} c \geq I(f) &= \int J(\dot{h}_+)d\alpha + \int J(\dot{h}_-)d\alpha + C_- h_-^s[0,1] \\ &\geq \int J(\dot{h}_+)d\alpha + J(-h_-^a[0,1]) + C_- h_-^s[0,1]. \end{aligned} \tag{4.18}$$

-9-

Let $M = M[0,1] \bigcap BV$. Let

$$\Delta_1 = \{\mu \in M : \mu << \alpha \text{ and} \int J(\dot{\mu}) \leq c\},$$
$$\Delta_2 = \{\mu \in M : J(-\mu[0,1]) \leq c\} \text{ and}$$
$$\Delta_3 = \{\mu \in M : C_- \mu[0,1] \leq c\}$$

Clearly Δ_3 is weak compact since $C_- \in (0,\infty)$ while Δ_2 is weak compact since $\Gamma_c^J = \{a : J(a) \leq c\}$ is compact. If we show that Δ_1 is Skorohov compact, then, since Δ_2 and Δ_3 are weak compact, it will follow from (4.18) that any sequence in Γ_C has a d-convergent subsequence. Since Γ_C is d-cloased by Lemma 4.3, the limit of the subsequence is in Γ_C. This suffices to show that Γ_C is d-compact.

Let $\mu \in \Delta_1$. For $d > 0$ and $A \in B$, let $A_d = \{t : \dot{\mu}(t) \leq d\}$ and $B_d = A - A_d$. Then, sionce $J(a)/a$ is nondecreasing in $a \leq 0$,

$$\begin{aligned}\mu(A) &= \int_{A_d} \dot{\mu} d\alpha + \int_{B_d} \dot{\mu} d\alpha \\ &\leq d\alpha(A) + \frac{d}{J(d)} \int J(\dot{\mu}) d\alpha \\ &\leq d\alpha(A) + \frac{dc}{J(d)}.\end{aligned}$$

Since $d/J(d) \to 0$ as $d \to \infty$, it follows from (4.19) that Δ_1 is uniformily absolutely continuous and hence equicontinuous. Since $\mu(0) = 0$ for every $\mu \in \Delta_1$, it follows from the ~~Argela0Asols~~ Theorem that Δ_1 is compact in the uniform topology and hence in the Skorohov topology. Arzela-Ascoli

The following is a minimax theorem for $I_p(f)$.

Thereom 4.6. If F is d-closer, then

$$\sup_p I_p(F) = I(F) \tag{4.20}$$

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 Corollary 4.2 we immediately have

$$\sup_p I_p(F) \leq I(F).$$

Suppose (4.20) were not true. Then there exists and $N < \infty$ such that

$$\sup_p I_p(F) < N < I(F). \tag{4.21}$$

Thus, for each partition $P = \{0 = t_0, t_1, \ldots, t_k = 1\}$, we can find $f_p \in F$ such that $I_p(f_p) < N$. For a finite measure μ, the $\hat{\mu}$, called the $\underline{P}$-linear form of μ, be given by

$$\hat{\mu}(A) = \sum_1^k \frac{\Delta_i \mu}{\Delta_i t} \alpha(A \bigcap (t_{i-1}, t_1)).$$

decomposition

Note that $\hat{\mu}_p(t)$ agrees with $\mu(t)$ for $t \in P$. For $f_p = h_{p+} - h_{p-}$, the Hahn-Jordan ~~decomposiont~~ of f_p, let

$$\hat{f}_p = \hat{h}_{p+} - \hat{h}_{p-}$$

where $\hat{h}_{pj}$ is the P-linear form of h_{pj}.

Since

$$N > I_p(f_p) = I_p(\hat{f}_p) = I(\hat{f}_p),$$

$\{\hat{f}_p\}$ is a set in the d-compact set Γ_n, it will follow that there is a cluster point f_0 of $\{\hat{f}_p\}$ in Γ_n. If we can show that f_0 is a cluster point of $\{f_p\}$; then it will follow that $f_0 \in F$ since F is d-closed. Since $I(f_0 \leq N$, this will contradict (4.21) and will establish (4.20).

As in the proof of Lemma 4.5, we shall only consider the case when $C_+ = \infty$ and $C_- < \infty$. The arguments given for this case can be used to prove the other cases.

From the proof of Lemma 4.5, there a decomposition of f_0, $f_0 = f_{0+} - f_{0-}$, with $f_{0+} << \alpha$ such that f_{0+} is a c=luster point of $\{\hat{h}_{p+}\}$ in the sup norm topology and f_0 is d_w-cluster point of $\{\hat{h}_{p-}\}$ when these two sets are viewed as sets in M.

Fix $\epsilon > 0$. Since f_{0+} is continuous, there is a partition $P' = \{0 = t'_0, t'_1, \ldots, t'_i = 1\}$ with $t'_1, \ldots, t'_{i-1}$ continuity points of f_{0-} such that $\max_i\{f_{0+}(t_i) - f_{0+}(t_{i-1}) < \epsilon/2$. Note that for $h \in M$,

$$\sup_t |h(t) - f_{0+}(t)| \leq max_i|f_{0+}(t_i) - f_{0+}(t_{i-1})| + \max_i |h(t_i) - f_{0+}9t_i)|$$

$$\leq \frac{\epsilon}{2} + \max_i |h(t_i) - f_{0+}(t_i)|.$$

It is easy from the above to see from this that the sets

$$N_{p:\epsilon} = \{f : \max_{i,j} |h_j(t_i) - f_{0j}(t_i)| < \epsilon\}$$

from a basis at f_0 for the d-open neighborhoods. of f_0 for any decomposition $f_0 = f_{0+} - f_{0-}$. Thus it suffices to show that for each $P'' > P'$, there is a $P > P''$ with $f_p \in N_{p',\epsilon}$.

Since f_0 is a cluster point of $\{\hat{f}_p\}$, ther is a $P > P''$ such that $\hat{f}_p \in N_{P',\epsilon}$. Since $\hat{h}_{pj}(t) = h_{pj}(t)$ for $t \in P$, it follows from the definition of $N_{P',\epsilon}$ that $f_p \in N_{P',\epsilon}$.

5. The Large Deviation Decomposition. Let $f \in BV$ and $P = \{0 = t_0, t_1, \dots, t_k = 1\}$ be a partition. For a decomposition $f = f_+ - f_-$ of f, let

$$(5.1) \qquad \begin{aligned} K_p(f_+, f_-) &= \sum L\Big(\frac{\Delta_i f_+}{\Delta_i t}, \frac{\Delta_i f_-}{\Delta_i t}\Big)\Delta_i t \\ &= \int L\Big(\frac{\partial f_{+p}}{\partial \alpha p}, \frac{\partial f_{-p}}{\partial \alpha p}\Big)\partial \alpha, \end{aligned}$$

where $L(\cdot,\cdot)$ is given in (3.9). Let

$$K(f_+, f_-) = \sup_p{}' K_p(f_+, f_-)$$

where the supremum is over all partitions P where $t_1, \dots, t_{k-1}$ are continuity points of both f_+ and f_-. Let

$$(5.3) \qquad K(f) - \inf\{K(f_+, f_-):\ f = f_+ - f_-\}.$$

In this section we establish some important relationships between the rate function I defined in (4.3) and $K_p(\cdot), K(\cdot,\cdot)$ and $K(\cdot)$ defined above. In particular it is shown that

$$(5.4) \qquad I(f) = K(f) = K(l_+, l_-)$$

where $f = l_+ - l_-$ is a particular decomposition of f where in light of (5.4) it will be referred to later as the large deviation (LD) decomposition. These results will be crucial to obtaining the lower bound (2.7) in the LDP for $\overline{X}_n(\cdot)$ and $\overline{Y}_n(\cdot)$.

Note that by (3.10), Lemma 4.1 and (5.1) $I(f) \le K(f_+, f_-)$. So

$$(5.5) \qquad I(f) \le K(f).$$

Consequently, (5.4) holds whenever $I(f) = \infty$ for any decomposition $f = f_+ - f_-$.

To establish (5.4) and identify the decomposition $f = l_+ - l_-$ in general, we need the following. Let $c(\cdot)$ be the Borel measurable function given in Lemma 3.3. When $f << \alpha$ with $I(f) < \infty$, let $\dot{l}_- = c(\dot{f}) \ge 0$ and $l_+ = \dot{f} + c(\dot{f}) \ge 0$. When $f << \alpha$ with $I(f) = \infty$, let $f = l_+ - l_-$ be any decomposition of f for which $l_j << \alpha$ for $j = +$ and $-$.

Lemma 5.1. Let $f << \alpha$ and let l_+ and l_- be as defined above. Then l_+ and l_- are finite absolutely continuous measures for which $f = l_+ - l_-$.

Proof. We only need to prove that $l_+[0,1]$ and $l_-[0,1]$ are finite when $I(f) < \infty$, since the other statements follow immediately from the definitions of l_+ and l_-. It follows from Lemma 3.3 and (4.3),

$$(5.6) \qquad \int L(\dot{l}_+, \dot{l}_-)\partial\alpha = \int J(\dot{f})\partial\alpha = I(f) < \infty.$$

Since $L(\cdot,\cdot)$ is convex by Theroem 3.2, the set $C = \{(b,c):\ L(b,c) < \infty\}$ is convex. By (5.6), $\alpha(\{t:\ (\dot{l}_+(t), \dot{l}_-(t) \in C) = 1$, so Jensen's inequality is in action and it follows by (5.6) that

$$(l_+[0,1], l_-[0,1]) \in \{(b,c):\ L(b,c) \le I(f)\}.$$

which is compact since $L(\cdot,\cdot)$ is a regular rate function by Theorem 3.2. This proves the finitines if l_+ and l_-.

Now for $f \in BV$, let $f = h_+ - h_-$ denote the Hahn-Jordan decomposition of f. Let $h_j = h_j^a + h_j^s$, $h_j^a << \alpha$ and $h_j^s \perp \alpha$, be the Lebesque deomposition of h_j. For $h^a = h_+^a - h_-^a$, let $h^a = l_+^a - l_-^a$ be the decomposition of h^a given by Lemma 5.1. Let $l_+ = l_+^a + h_+^s$ and $l_- = l_-^a + h_-^s$. Then $f = l_+ - l_-$. It follows from the following lemma that $K(f_+, f_-), f = f_+ - f_-$, attains its mimimum at (l_+, l_-). It is because of this that $f = l_+ - l-$ is referred to as the large deviation decomposition of f.

Lemma 5.2. Let $f = l_+ - l_-$ be the LD decomposition of f. Then $I(f) = K(f) = K(l_+, l_-)$.

Proof. From the discussion above we know that lemma is true if $I(f) = \infty$. Thus we assume that $I(f) < \infty$ and consider first the case when $f << \alpha$.

Bold or Script

Since $l_+ << \alpha$ and $l_- << \alpha$, the collection $\{\left(\frac{\partial l_{+p}}{\partial \alpha_p}, \frac{\partial l_{-p}}{\partial \alpha_p}\right), \overset{\sigma}{\cancel{\wp}}(P); (\dot{l}_+, \dot{l}_-); B\}$ is a vector-valued martingale. Consequently, by a careful application of Jensen's inequality (see the proof of Lemma 5.1),

$$K_p(l_+, l_-) = \int L\Big(\frac{\partial l_{+p}}{\partial \alpha_p}, \frac{\partial l_{-p}}{\partial \alpha_p}\Big)\partial\alpha$$

$$\le \int L(\dot{l}_+, \dot{l}_-)\partial\alpha \tag{5.7}$$

$$= \int J(\dot{f})\partial\alpha = I(f),$$

since $L(\cdot, \cdot)$ is convex by Theorem 3.2 and where the second identity follows since $I(f) < \infty$ and $f = l_+ - l_-$ is the LD decomposition. This shows that

$$K(l_+, l_-) \le I(f).$$

Combining them with (5.3) and (5.5) proves the lemma when $f << \alpha$.

Now suppose $I(f) \ne K(f)$. Then, by (5.3) and (5.5) ther is an $\eta < \infty$ with

$$I(f) < \eta < K(f) \le K(l_+, l_-). \tag{5.8}$$

For each partition P, let $\hat{l}_{jp}$ denote the P-linear form of $l_j, j = +$ and $-$. Let $\hat{f}_p = \hat{l}_{+p} - \hat{l}_{-p}$ and note that $\hat{f}_p << \alpha$. Let $\hat{f}_p = l_{p+} - l_{p-}$ denote the LD decomposition of $\hat{f}_p$. It is easy to see that $\dot{l}_{+p}$ and $\dot{l}_{p-}$ are $\overset{\sigma}{\cancel{\wp}}(P)$ measurable. So, for $P \supset P'$

$$K_{p'}(l_{p+}, l_{p-}) \le K_p(l_{p=}, l_{p-}) \tag{5.9}$$

$$K(l_{p+}, l_{p-}) = I(\hat{f}_p) = I_p(f) \le I(f) \le \eta,$$

where the second identity follows from the first part of the proof since $\hat{f}_p << \overset{\alpha}{\cancel{\partial}}$ and the third identity and the inequality follow from (4.2) and Lemma 4.1, respectively.

Since L is convex, another (careful) application of Jensen's inequality with (5.9) gives

$$L(l_{p+}[0,1], l_{p-}[0,1] \le \int L(\dot{l}_{p+}, \dot{l}_{p-})\partial\alpha = K$$

$$= K(l_{p+}, l_{p-}) \le \eta.$$

Since Γ_η^L is compact, it follows that the net $\{(l_{p+}, l_{p-})\}$ is compact in the product weak* topology. Let (l_{0+}, l_{0-}) denote a cluster point of this net. Let $P' = \{0 = t'_0, t'_1, \ldots, t'_l\}$ denote a partition where $t'_1, \ldots, t'_l$ are continity points of both l_{0+} and l_{0-}. Note that P' are continuity points of l_{p+} and l_{p-} since these measures are absolutely continuous. Thus, since L is lsc, it follows from (5.1) and (5.9) that

$$K_{p'}(l_{0+}, l_{0-}) \le \eta,$$

and hence $K(l_{0+}, l_{0-}) \le \eta$ by (5.2). Further since $\hat{f}_p(t)$ agrees with $f(t)$ for $t \in P, f = l_{0+} - l_{0-}$. So $K(f) \le K(l_{0+}, l_{0-}) \le \eta$ which contradicts (5.8).

-13-

6. The LDP for the Sample Average Process $\overline{X}_n(\cdot)$. As in Section 3, throughout the remaining sections $X_1, X_2, \ldots$ are iid with mean $E(X_1) = 0$ and m.g.f. $\phi(\theta)$ finite in some open interval about zero. Recall that from (3.2), $J(a) = \sup\{\theta_a - \log\phi(\theta)\}$ is the large deviation rate of $\overline{X}_n$.

Let $\overline{X}_n(t) = S(nt)/n$. We view $\overline{X}_n(\cdot)$ as taking values in $BV[0,1]$ with the metric ∂ defined in (4.7). With this setting we establish the LDP for $\{\overline{X}_n(\cdot)\}$.

Theorem 6.1. Let P_n denote the probability measure induced on the Polish space BV with metric ∂. Then $\{P_n\}$ satisfies the LDP with rate function

$$I(f) = \int J(\dot{f})\partial\alpha + C_+ h_+^s[0,1] + c_- h_-^s[0,1], \tag{6.1}$$

where $\dot{f}, h_+^s, h_-^s, C_+, C_-$ are defined in Section 4.

Proof. If follows from Lemmas 4.4 and 4.5 that $I(f)$ is a regular rate function. Thus to complete the proof we only need to establish the upper and lower bounds ((2.6) and (2.7), respectively) in the LDP.

Proof of the upper bound. Let F be $\cancel{\partial}$-closed d and let $P = \{0 = t_0, t_1, \ldots, t_k = 1\}$ be a partition. Then, by the definition of $I_p(F)$,

$$P(\overline{X}_n(\cdot) \in F) \le P(I_p(\overline{X}_n(\cdot)) \ge I_p(F)). \tag{6.2}$$

Let $Z_{1n} = \overline{X}_n(t_1)$ and $Z_{in} = \overline{X}_n(t_{i-1}, t_i]$ for $i = 1, 2, \ldots, k$. It follows from the Contraction Principle that $\{Z_{1n}\}$ satisfies the LDP with rate $J(a/\Delta_i t)\Delta_i t$. Thus, since $Z_{1n}, \ldots, Z_{kn}$ are independent, it follows from Lemma 2.6 and Corollary 2.10 that, for $\mathbf{Z}_n = (Z_{1n}, \ldots, Z_{kn})$, $\{\mathbf{Z}_n\}$ satisfies the LDP with rate $I(a_1, \ldots, a_k) = \sum_1^k J(a_i/\Delta_i t)\Delta_i t$.

For $\overline{\Gamma}$ as defined in (3.8), let $\Gamma_i = \Delta_i t\overline{\Gamma}$. Then, $\Gamma_0 = \Gamma_1 \times \Gamma_2 \times \ldots \times \Gamma_k$. Endow Γ_0 with the relative topology and note that there is a topologically equivalent metric under which Γ_0 is complete. (See (3.8)). Thus, $I(a_1, \ldots, a_n)$ is continuous (in the extended sense) on Γ_0 and hence, the set, $G = \{\mathbf{a} \in \Gamma_0\colon\ I(\mathbf{a}) \ge I_p(F)\}$ is a closed subset of Γ_0. Thus, by (6.2)

$$\begin{aligned} &\overline{\lim}\frac{1}{n}\log P(\overline{X}_n(\cdot) \in F) \\ &\le \overline{\lim}\frac{1}{n}\log P(\mathbf{Z}_n \in G) \\ &\le -\inf_{\mathbf{a}\in G} I(\mathbf{a}) \le -I_p(F) \end{aligned} \tag{6.3}$$

since G is closed and $\{\mathbf{Z}_n\}$ satisfies the LDP with rate $I(\mathbf{a})$. The upper bound then follows from the minimax theorem, Theorem 4.6, by~~y~~ taking the supremum over P in (6.3).

Proof of the lower bound. As in Section 5 we will only consider the case when $C_+ = \infty$ amd $C_- < \infty$. The other cases will follow from the arguments for this case.

d Let G be $\cancel{\partial}$-open set and let $f \in G$. Let $f = l_+ - l_-$ be the LD decomposition of f. Since G is ∂-open, there is a neighborhood of f, $N_{p,\epsilon} \subset G$, given by (4.22) with $f_{0+} = l_+$ and $f_{0-} = l_-$. Let $G^* = \{(\mathbf{b}, \mathbf{c}) \in R^{2k}\colon\ \max_i |b_i - l_+(t_i)| < \epsilon$ and $\max_i |c_i - l_-(t_i)| < \epsilon$. Note that G^* is an open subset of R^{2k}.

For $\overline{X}_n(\cdot) = \overline{X}_n^+(\cdot) - \overline{X}_n(\cdot)$ the Hahn-Jordan decompostion of $\overline{X}_n(\cdot)$, let $M_{1n} = \overline{X}_n^+(t_1)$ and $N_{1n} = \overline{X}_n^-(t_1)$ and for $i = 2, \ldots, k$, let $\overline{X}_n^+\Big((t_{i-1}, t_1]\Big)$ and $N_{1n} = \overline{X}_n^-\Big((t_{i-1}, t_i]\Big)$. Let $(\mathbf{M}_n, \mathbf{N}_n) = (M_{1n}, \ldots, M_{kn}, N_{1n}, \ldots, N_{kn})$. Since $(M_{1n}, N_{1n}), \ldots, (M_{kn}, N_{kn})$ are independent, it follows from Theorem 3.2 and the contraction principle that $\{(\mathbf{M}_n, \mathbf{N}_n)\}$ satisfies the LDP with rate $\sum_1^k L(\frac{b_i}{\Delta_i t}, \frac{c_i}{\Delta_i t})\Delta_i t$. Thus

$$\begin{aligned} &\lim\frac{1}{n}\log P(\overline{X}_n(\cdot) \in G) \\ &\ge \lim\frac{1}{n}\log P(\overline{X}_n(\cdot) \in N_{p,\epsilon}) \\ &= \lim\frac{1}{n}\log P\Big((\mathbf{M}_n, \mathbf{N}_n) \in G^*\Big) \\ &\ge -\inf_{(\mathbf{b},\mathbf{c})\in G^*} \sum_1^k L(\frac{b_i}{\Delta_i t}, \frac{c_i}{\Delta_i t})\Delta_i t \\ &\ge -K_p(l_+, l_-). \end{aligned} \tag{6.4}$$

where the last inequality follows by (5.1). The lower bound follows from (5.2) and Lemma (5.2) by taking the inf in (6.4) over partitions P, with $t_1, \dots, t_{k-1}$ continuity points of both l_+ and l_-.

7. The LDP for $\overline{Y}_n(\cdot)$ when $\lim_{|a|\to\infty} J(a)/|a| = \infty$. Let $\overline{Y}_n(\cdot)$ be the sample average process given by $\overline{Y}_n(t) = P(nt)/n$. Then $\overline{Y}_n(\cdot)$ takes values in $C = C[0,1] \cap BV$. Endow C with the topology given by the metric definded in (4.17) when $C_+ = \infty = C_-$. We note that this metric is topologically equivalent to the metric (actually a norm)

(7.1)
$$\partial(f,g) = \inf\{\max_{j=+,-} \|f_i - g_i\|: \ f = f_+ - f_- \text{ and } g = g_+ - g_-\}$$

where $\|\cdot\|$ is the sup norm. (See Billingsley (1968), Section 14.) Note that with this metric, C is a Polish space.

In this section, ~~we established~~, we established the LDP for $\{\overline{Y}_n(\cdot)\}$ under the strongest that

(7.2)
$$\lim_{|a|\to\infty} J(a)/|a| = \infty.$$

under the strongest condition that ?

The proof hinges on the following lemma.

Lemma 7.1 Under 7.1

$$\lim \frac{1}{n} \log P(|X_1| \geq n\epsilon) = -\infty \text{ for } \epsilon > 0.$$

Proof. For $\theta > 0$, $\frac{1}{n}\log P(X_1 \geq n\epsilon) \leq -\frac{1}{n}[n\epsilon\theta - \log\phi(\theta)]$ by Markov's inequality. Taking the inference over $\theta \geq 0$, by (3.4) we get

$$\frac{1}{n}\log P(X_1 \geq n\epsilon) \leq -J(n\epsilon)/n \to -\infty$$

under condition (7.1). Similarly,

$$\lim \frac{1}{n}\log P(X_1 \leq -n\epsilon) = -\infty.$$

As an immediate corollary we have that

Corollary 7.2. Let P_n be the distribution of X_1/n. Then, $\{P_n\}$ satisfies the LDP with rate

$$J_0(a) = \begin{cases} \infty & \text{if } a \neq 0 \\ 0 & \text{if } a = 0. \end{cases}$$

Theorem 7.3. Let P_n be the probability measure induced on C by $\overline{Y}_n(\cdot)$. Then, under condition (7.1), $\{P_n\}$ satisfies the LDP with rate

$$I(f) = \begin{cases} \int J(\dot{f})d\alpha & \text{if } f << \alpha, \text{ and} \\ \infty & \text{if not.} \end{cases}$$

Proof. It follows from Lemmas 4.4 and 4.5 that $I(f)$ is a regular rate function. Thus to complete the proof we only need to establish the upper and lower bounds in the LDP.

Proof of the upper bound. Let F be closed and let $P = \{0 = t_0, t_1, \ldots, t_k = 1\}$ be a partition. Let $Y_{jn} = 0$ if $[nt_j]$ equals an integer and $X_{[nt_j]+1}/n$ if not. Then it follows from the Contradiction Principle, Theorem 3.1, Corollary 7.2, Lemma 2.6 and Corollary 2.10 that the vectors $\{\mathbf{Z}'_n\} = \{(Z'_{1n}, Z'_{2n}, \ldots, Z'_{2k-1,n})\} \equiv \{(\overline{Y}_n((0, \frac{[nt_1]}{n}], Y_{1n}, \overline{Y}_n((\frac{[nt_1]+1}{n}, \frac{[nt_2]}{n}], Y_{2n}, \ldots,)\}$ satisfy the LDP with rate $\sum_1^k J(a_i/\Delta_i t)\Delta_i t + \sum_1^{k-1} J_0(b_i)$ since the components of the vector $\mathbf{Z}_n$ are independent. Another application of the Contraction Principle to $\{\mathbf{Z}'\}$ shows that the vectors $\{\mathbf{Z}_n\} = \{(Z_{1n}, \ldots, Z_{kn})\}\{(\overline{Y}_n([0,t_1]), \overline{Y}_n((t_1,t_2]), \ldots, \overline{Y}_n(t_{k-1}, 1])$ satisfy the LDP with rate

$$\sum_1^k J(a_i/\Delta_i t)\Delta)it.$$

Now the rest of the proof is exactly the same as the proof of the upper bound in Theorem 6.1.

Proof of the lower bound. Let G be an open set and $f \in G$. It follows that there is a neighborhood of f, $N_{p,\epsilon} \subset G$, given by (4.22).

Let $\overline{Y}_n(\cdot) = \overline{Y}_n^+(\cdot) - \overline{Y}_n^-$ denote the Hahn-Jordan decomposition of $\overline{Y}_n(\cdot)$. Let $M'_{jn} = \overline{Y}_n^+((\frac{[nt_{i-1}]}{n}, \frac{[nt_i]+1}{n}])$, and $N'_{jn} = \overline{Y}_n^-((\frac{[nt_{i-1}}{n}, \frac{[nt_i]+1}{n}])$ and Y_{jn} be as defined in the proof of the upper bound. Then, it follows from the Contraction Principle, Theorem 3.2, Corollary 7.2, Lemma 2.6 and Corollary 2.10 that the vectors $\{(\mathbf{M}'_n, \mathbf{N}'_n, \mathbf{Y}_n)\}$ satisfy the LDP with rate $\sum_1^k L(\frac{b_i}{\Delta_i t}, \frac{c_i}{\Delta_i t}) + \sum^{k-1} J_0(a_i)$ since$(M'_{jn}, N'_{jn}) j = 1, \ldots, k$ and $Y_{jn}, j = 1, \ldots, k-1$ are independent. Another application of the Contraction Principle to the above sequence of vectors shows that the vector $\{(M_{1n}, N_{1n}, M_{2n}, N_{2n}, \ldots, M_{kn}, N_{kn})\}$ satisfies the LDP with rate

$$\sum_1^k L(\frac{b_i}{\Delta_i t}, \frac{c_i}{\Delta_i t})\Delta_i t$$

where $M_{jn} = Y_n^+((t_{j-1}, t_j])$ and $N_{jn} = Y_n^-((t_{j-1}, t_j]), j = 1, \ldots, k$. Now the rest of the proof is exactly the same as the proof of the lower bound in Theorem 6.1.

8. Functional E-R Laws. Recall the definitions of $\Delta_{m,n,a}(\cdot)$ and $\Delta'_{m,n,a}(\cdot)$ in the Introduction. In this section, we establish versions of the Shepp's law and of the Erdős and ~~Reyi's~~ Rényi's law for the functionals $\Delta_{m,n,a}(\cdot)$ and $\Delta'_{m,n,a}(\cdot)$.

Note that $\Delta_{m,n,a}(\cdot)$ takes values in $BV[0,1]$ while $\Delta'_{m,n,a}(\cdot)$ can be viewed as taking values in $BV[0,1]$ or $C[0,1]$. Throughout d will denote either the metric defined in (4.17) when the relevant space is $BV[0,1]$ or the sup norm when the relevant space is $C[0,1]$.

The key to proving these functional laws are having large deviation approximations for $\{\Delta_{m,n,a}(\cdot)\}$ and $\{\Delta'_{m,n,a}(\cdot)\}$. These are available from the results in Sections 6 and 7 since $\Delta_{m,n,a}(\cdot)$ and $\overline{X}_{[a^{-1}\log n]}(\cdot)$ or $\Delta'_{m,n,a}(\cdot)$ and $\overline{Y}_{[a^{-1}\log n]}(\cdot)$ have the same distributions. These are stated below.

Let P_n denote the probability measure induced on $BV[0,1]$ (with metric d) by $\overline{X}_n(\cdot)$. Let P'_n denote the measure induced on $BV[0,1]$ by $\overline{Y}_n(\cdot)$. Recall the definition of BV given in Section 4 and of C given in Section 7.

Theorem 8.1. The sequence $\{P_n\}$ satisfies the LDP with rate function

$$I(f) = \begin{cases} \int J(\dot f) + C_+ h^s_+[0,1] + C_- h^s_-[0,1] & \text{if } f \in BV, \text{ and} \\ \infty & \text{otherwise.} \end{cases}$$

If, in addition, (7.2) is satisfied, then $\{P'_n\}$ and $\{Q_n\}$ satisfy the LDP with rate function

$$I(f) = \begin{cases} \int J(\dot f)d\alpha & \text{if } f \in C, \text{ and } f << \alpha \\ \infty & \text{otherwise.} \end{cases}$$

Proof. This is immediate from the Contraction Principle and Theorem 6.1 and 7.3 since the identity maps from $BV \to BV[0,1]$ and from $C \to C[0,1]$ are "continuous". In the latter case ~~not~~ note that the metric given by (7.1) is stronger than the sup norm.

As before let $\Gamma_a = \{f : I(f) \le a\}$ where $I(\cdot)$ one of the rate functions defined in Theorem 8.1. For a set A, let $A_\epsilon = \{g : d(f,g) < \epsilon \text{ for some } f \in A\}$. The functional analogues of Shepp's (1964) and Erdős and ~~Réníi's~~ Rényi's (1970) Laws can be stated as follows ~~where the functional an Erdős Rényi laws given~~

Theorem 8.2. The set of cluster points of $\{\Delta_n, n, a(\cdot)\}$ is Γ_a. If, in addition, condition (7.2) is satisfied, then the set of cluster points of $\{\Delta'_{n,n,a}(\cdot)\}$ is also Γ_a.

Theorem 8.3. Let $\epsilon > 0$ with probability one,

Theorem 8.3 is the functional analogue of da Acosta and Kuelbs (1983) version of the Erdős-Renyi Law.

(i) $\{\Delta_{m,n,a}(\cdot) : m \le n\} \subset (\Gamma_a)_\epsilon$ eventually

and

(ii) $\Gamma_a \subset \{\Delta_{m,n,a}(\cdot) : m \le n\}_\epsilon$.

(iii) In particular, the set of cluster points of the triangular array $\{\Delta_{m,n,a}(\cdot) : m \le n\}$ is Γ_a.

If, in addition, (7.2) is satisfied then (i), (ii), and (iii) also hold for $\{\Delta'_{m,n,a}(\cdot) : m \le n\}$.

To prove these theorems we need the following lemma.

Lemma 8.4. (i) For each $\epsilon > 0$ ther exists a $c > a$ for which $\Gamma_c \subset (\Gamma_a)_\epsilon$ and (ii) Γ_a equals the closure of $\{f : I(f) < a\}$.

Proof. (i) Suppose not. Then for every $c > a$ there exists an $f_c \in \Gamma_c$ with $f_c \in (\Gamma_a)\epsilon$. Fix $b > a$. Since Γ_b is compact the net $\{f_c : a < c < b\}$ has a subnet $\{f_d\}$ which converges, say, to f_0, as $d \downarrow a$. By lsc of $I(\cdot)$ it follows that $I(f_0) \le \lim_{d \downarrow a} I(f_d) \le a$, and so $f_0 \in \Gamma_a$. But $f_c \notin (\Gamma_a)_\epsilon$ implies that $d(f_0, f_c) > \epsilon$ for every c which contradicts that $\{f_d\}$ converges to f_0.

(ii) Fix $f \in \Gamma_a$ with $I(f) = a$. (If no such t there is nothing to prove.) It suffices to show that $I(\alpha f) < \alpha$ for $0 < \alpha < 1$. This is immediate since I is convex with $I(0) = 0$, where 0 denotes the function which is identically zero. ||

Proof of Theorems 8.2 and 8.3. We only prove the theorems for $\{\Delta_{m,n,a}\}$ since the proof for $\{\Delta'_{m,n,a}\}$ is the same.

Fix $a \ge 0$ and $\epsilon > 0$. We first show that

$$\text{(8.1)} \qquad \text{all any } f \in \Gamma_a, d((\Delta_{n,n.a}(\cdot), f < \epsilon) \text{ infinitely often a.s..}$$

Let $0 = \{g : d(g,f) < \epsilon\}$. Note that since 0 is open, $\alpha \cdot f \in 0$ for some $\alpha \in (0,1)$. By the LDP,

$$\lim n^{-1} \log P(\overline{X}_n \in 0) \ge -I(\alpha f).$$

Since $I(\cdot)$ is convex with $I(0) = 0$ and $\alpha \in (0,1), I(\alpha f) \leq \alpha I(f) \leq \alpha a$. Thus

$$(8.2) \qquad P(\Delta_{mn,n,a}(\cdot) \in 0) \geq n^{-d(1+\sigma(1))}.$$

Since $\Sigma n^{\alpha'} = \infty$, for $\alpha' \epsilon (\alpha, 1)$ fixed it follows from Lemma (3.1) of Shepp (1964) that there exists a sequence $\{n(k)\}$ with $n(k+1) = n(k) + [a^{-1}\log n(k)]$ such that $\Sigma n(k)^{-\alpha'} = \infty$. This with (8.2) shows that $\Sigma P(\Delta_{n(k),n(k),a}(\cdot) \in 0) = \infty$. Statement (8.1) follows from this and the divergent part of the Borel-Cantelli lemma since the events $\{\Delta_{n(k),n(k),a}(\cdot) \in 0\}, k = 1, 2, \ldots$ are independent.

We now show that, with probability 1,

$$(8.3) \qquad \{\Delta_{m,n,a}(\cdot) : m \leq n\} \subset (\Gamma_a)_\epsilon \text{ eventually.}$$

For all sufficiently large integer k, let $K = e^{ak+a}$ and let $1(k) = K - 1$ if K is an integer and $= [K]$ if not. Then, since $k = [a^{-1}\log l(k)]$, for $1(k-1) < n \leq 1(k), \Delta_{m,n,a}(\cdot) = \Delta_{m,l(k),a}(\cdot)$ for $m \leq n$. Thus, to prove (8.3), it suffices to show that

$$(8.4) \qquad P(\Delta_{m,l(k),a}(\cdot) \notin (\Gamma_a)_\epsilon \text{ for some } m \leq 1(k) \text{ i.o.}) = 0.$$

By Lemma 8.4, (i) there exists $\cancel{c} > a$ such that $\Gamma_c \subset \Gamma_{a,\epsilon}$. Thus $I((\Gamma_a)^c_\epsilon) \geq c$. So by the LDP,

$$\overline{\lim} n^{-1} \log P(\overline{X}_n(\cdot) \epsilon ((\Gamma_a)^\epsilon)^c \leq -I((\Gamma_a)_\epsilon)^c$$

lower case c

since $(\Gamma_a)^c_\epsilon$ is closed with $((\Gamma_a)_\epsilon)^c \subset \Gamma^C_c$. Thus, for k' sufficiently large,

$$\sum_{k \geq k'} P(\Delta_{m,l(k),a}(\cdot) \epsilon ((\Gamma_a)_\epsilon)^c \text{ for some } m \leq 1(k))$$
$$\sum_{k \geq k'} 1(k) e^{-ck(1+0(1))} = \sum_{k \geq k'} e^a e^{(a-c)(1+0(1))k} < \infty$$

So (8.4) follows from the convergent part of Borel-Cantelli lemma. This completes the proof of (8.3).

Since Γ_a is closed, $\Gamma_{a,\epsilon} \downarrow \Gamma_a$ as $\epsilon \downarrow 0$. This with (8.1) and (8.3) completes the proof of Theorem 8.2 and (i) of Theorem 8.3.

To prove (ii) of Theorem 8.3, let $\epsilon > 0$. Then, by Lemma 8.4 (ii), there is a finite collection $\{f_1, \ldots, f_k\} subset\{f$: $I(f) < a\}$ such that $\Gamma_a \subset \cup_{i=i}^k \{f_i\}_{\epsilon/2}$.

Since $\{f_i\}_{\epsilon/2}$ is an open ~~et~~ set, by the LDP,

$$\lim n^{-1} \log P(\overline{X}_n \in \{f_i\}_{\epsilon/2}) \geq -I(\{f_i\}_{\epsilon/2}) \geq -I(f_i).$$

So, ?

Since $I(f_i) < a$ for $i = i, \ldots, k$, it follows that $\Sigma P(\Gamma_a \not\subset \{\Delta_{m,n,a} : m \leq n\}_{\epsilon/2}) < \infty$. This with the convergent part of the borel-Cantelli lemma completes the proof.

9. Some Final Comments. (i) We note that the metric d defined in (4.17) induces a finer topology on $BV[0,1]$ than the weak* topology (which is, by the way, not metrizable). Furthermore, when $C_+ = \infty = C_-$, the topology induced on $BV[0,1]$ or on C by d is stronger than the Skorohod topology on either of these spaces. (Note that in the latter case this is just the topology of uniform convergence). Thus, the large deviation results in Sections 6, 7, and 8 imply the LDP for $\{\overline{X}_n(\cdot)\}$ or $\{\overline{Y}_n(\cdot)\}$ when taking values in these topological spaces.

(ii) We rate that lower bound for open sets in the LDP established in Sections 6 and 7 actually holds for open sets in the finer topology given by the metric

$$d(f_1, f_2) = \max_{j=+,-} d_j(l_j^1, l_j^2)$$

where $f_i = l_+^i - l_-^i$ is the LD decomposition of f_i and d_j is the metric defined in Section 4.

(iii) As with ~~Strassén~~ Strassen's LIL, it is important for the E-R functional laws that Γ_a be "compact". The example at the end of Section 3 in Lynch and Sethuraman (1987) shows that if C_+ or C_- is finite Γ_a cannot be compact in the stronger topology induced by the metric d defined in (4.17) when $C_+ = \infty = C_-$. In particular, if C_+ or C_- is finite, the LDP holds for $\{\overline{Y}_n(\cdot)\}$ with the weak* topology on $BV[0,1]$ but not for the Skorhod topology.

(iv) It would appear that the work of Azencott and Ruget (1977) (see also Azencott (1980)) could be used to establish the LDP for $\{\overline{X}_n(\cdot)\}$ or $\{\overline{Y}_n(\cdot)\}$. It would seem for $\overline{Y}_n(\cdot)$ when $C_+ = \infty = C_-$ (Condition (7.2)). However, if C_+ or C_- is finite, the singular part of the rate function is crucial (especially for compactness). This part does not appear in Azencott and Ruget's rate function which would be in our case

$$I_0(f) = \begin{cases} \int J(\dot{f})d\alpha & f << \alpha \\ \infty & \text{if not.} \end{cases}$$

The same example cited in Comment (iii) gives a situation where $\Gamma_a^0 = \{f : I_0(f) \le a\}$ is not weak* closed. Hence Γ_a^0 is not weak* compact on ~~Skorhod~~ Skorohod compact.

Acknowledgements. The authors would like to thank Alex de Acosta for ~~point~~ pointing out the de Acouts-Kuebbs (1983) version of the Erdős-Renyi Law. They would also like to thank the referee of an earlier version of this paper for reminding them that $BV[0,1]$ with the weak* topology is ~~and~~ not metrizable. Finally, the authors would like to thank Steven Orey and Hans Weinberger for inviting them to spend some time at the Institute of Mathematics and Its Applications where some of this research was done.

References

Azencott, R. (1980 Grand deviations et applications. Lecture Notes in Math, 774, 1 - 176.

Azencptt, R. and Ruget, G. (1977) Mélanges d'équations différentielles et grands écarts a la loi des grands nombre. Z Wahrscheinlichkeit.verw.Cebiete 28, 1 - 54.

Bahadur R.R. and Zabell, S.L., Large deviations of the sample mean in general vector spaces (a paraître).

Billingsley, Patrick (1968). Convergence of Probability Measures. Wiley, New York.

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.

Cramer, H. (1937). Sur un nouveau theoreme limite de la theorie des probabilities. Colloquium on theorey of probability. Paris-Herman.

de Acousta, A. and Kuebbs, J. (1983). Limit Theorems for moving averages of independent random variables. Z.Wahrscheinlichkeitstheorie Verw. Gebiete. 64, 67 - 123.

Erdős, P. and Rényi, A. (1970). On a new law of large numbers. J. Analyse Math., 23, 103 - 111.

Groeneboom, P., Oosterhoff, J., and Ruymgaart, F.H. (1979). Large deviation theorems for empirical probability measures. Ann.Probab., 7, 553 - 586.

Lynch, J. and Sethuraman, J. (198~~4~~7). Large deviations for processes with independent increments. ~~FSU Tech Report Number M687~~. Ann. Probab., 15, 610 - 627.

Shepp, L.A. (1964). A limit theorem concerning moving averages. Ann. Math. Statists., 35, 424 - 428.

Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie and Verw. Gebiete, 3, 211 - 226.

Stroock, D. and Varadhan, S.R.S. (1973). Topics in Probability Theory. Courant Institute of Mathematical Sciences Lecture Notes.

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, CBMS/NSF Regional Conference in Applied Math., 46 SIAM, Philadelphia.