Large deviations of empirical processes

Miguel A. Arcones

Abstract. We give necessary and sufficient conditions for the large deviations of empirical processes and of Banach space valued random vectors. We also consider the large deviations of partial sums processes. The main tool used is an isoperimetric inequality for empirical processes due to Talagrand.

April 13, 2004

1. Introduction

We study the (LDP) large deviation principle for different types of sequences of empirical processes $\{U_n(t) : t \in T\}$, where T is an index set. General references on large deviations are Bahadur (1971), Varadhan (1984) and Deuschel and Stroock (1989). We consider stochastic processes as elements of $l_\infty(T)$, where T is an index set. $l_\infty(T)$ is the Banach space consisting of the bounded functions defined in T with the norm $\|x\|_\infty = \sup_{t\in T} |x(t)|$. We will use the following definition.

Definition 1.1. Given a sequence of stochastic processes $\{U_n(t) : t \in T\}$, a sequence of positive numbers $\{\epsilon_n\}_{n=1}^\infty$ such that $\epsilon_n \to 0$, and a function $I : l_\infty(T) \to [0, \infty]$, we say that $\{U_n(t) : t \in T\}$ satisfies the LDP with speed ϵ_n^{-1} and with good rate function I if:

(i) For each $0 \le c < \infty$, $\{z \in l_\infty(T) : I(z) \le c\}$ is a compact set of $l_\infty(T)$.

(ii) For each set $A \in l_\infty(T)$,

$$-I(A^o) \le \liminf_{n\to\infty} \epsilon_n \log(\Pr{}_*\{\{U_n(t) : t \in T\} \in A\})$$

and

$$\limsup_{n\to\infty} \epsilon_n \log(\Pr{}^*\{\{U_n(t) : t \in T\} \in A\}) \le -I(\bar{A}),$$

where for $B \subset l_\infty(T)$, $I(B) = \inf\{I(x) : x \in B\}$.

It was shown in Arcones (2002a), that this definition is equivalent to the large deviations of the finite dimension distributions plus an asymptotic equicontinuity condition. Thus, large deviations can be studied similarly to the weak convergence of empirical processes.

1991 *Mathematics Subject Classification.* Primary 62E20; Secondary 62F12.
Key words and phrases. Large deviations, empirical processes.

In Section 2, we present necessary and sufficient conditions for the large deviations of empirical processes and of sums of Banach space valued i.i.d.r.v.'s. If $\{X_j\}_{j=1}^\infty$ is a sequence of real i.i.d.r.v.'s, it follows from the results in Cramér (1937) and Chernoff (1952) that $n^{-1}\sum_{j=1}^n X_j$ satisfies the LDP with speed n if and only if for some $\lambda > 0$ $E[e^{\lambda |X_1|}] < \infty$. This is also true for r.v.'s with values in a finite dimensional vector space. Given a sequence $\{X_i\}$ of B–valued i.i.d.r.v.'s, where B is a separable Banach space, Sethuraman (1964, Theorem 7) and Donsker and Varadhan (1976, Theorem 5.3) showed that if for each $\lambda > 0$ $E[e^{\lambda |X_1|}] < \infty$, then the LDP holds for $n^{-1}\sum_{j=1}^n X_j$ with speed n and with rate function

$$I(x) = \sup\{f(x) - \log(E[e^{f(X)}]) : f \in B^*\},$$

where X is a copy of X_1 and B^* is the dual of B. We obtain that in the previous situation the LDP holds for $n^{-1}\sum_{j=1}^n X_j$ with speed n and a good rate function if and only if there exists a $\lambda > 0$ such that $E[e^{\lambda |X_1|}] < \infty$; and for each $\lambda > 0$ there exists a $\eta > 0$ such that $E[e^{\lambda W^{(\eta)}}] < \infty$, where

$$W^{(\eta)} = \sup\{|f_1(X) - f_2(X)| : f_1, f_2 \in B_1^*,\ E[|f_1(X) - f_2(X)|] \le \eta\},$$

where B_1^* is the unit ball of B^*. As a corollary, we obtain that when B is a Hilbert space $(H, < \cdot >)$, the LDP holds for $n^{-1}\sum_{j=1}^n X_j$ with speed n and a good rate function if and only if there exists a $\lambda > 0$ such that $E[\exp(\lambda |X|)] < \infty$; and for each $\lambda > 0$, there exists an integer m such that $E[\exp(\lambda |X^{(m)}|)] < \infty$, where $X^{(m)} = \sum_{k=m+1}^\infty < X, h_k > h_k$ and $\{h_k\}$ is an orthogonal basis of H. We also prove that the stochastic process $\{n^{-1}\sum_{j=1}^{[nt]} X_j : 0 \le t \le 1\}$ satisfies the LDP in $l_\infty([0,1], B)$ with speed n and a good rate function if and only if for each $\lambda > 0$ $E[e^{\lambda |X|}] < \infty$. Here, $l_\infty([0,1], B)$ denotes the Banach space consisting of the bounded functions from $[0,1]$ into B with the norm $\|x\|_{\infty,[0,1],B} = \sup_{0\le t\le 1} |x(t)|$.

We will obtain the previous results from results on empirical processes. The study of the large deviations of empirical processes started with Sethuraman (1964, 1965, 1970), where sufficient conditions were presented such that for each $\epsilon > 0$, the following limit exists,

$$\lim_{n\to\infty} n^{-1} \log(\Pr\{\sup_{f\in\mathcal{F}} n^{-1} |\sum_{j=1}^n (f(X_j) - E[f(X_j)]) \ge \epsilon\}),$$

where $\{X_j\}_{j=1}^\infty$ is sequence of i.i.d.r.v.'s with values in separable compact metric space S and $\mathcal{F}$ is a collection of functions on S

c will denote an universal constant that may vary from line to line. We will use the usual multivariate notation. For example, given $u = (u_1, \dots, u_d)' \in \mathbb{R}^d$ and $v = (v_1, \dots, v_d)' \in \mathbb{R}^d$, $u'v = \sum_{j=1}^d u_j v_j$ and $|u| = (\sum_{j=1}^n u_j^2)^{1/2}$. Given a sequence of real numbers $a = \{a_k\}$, we denote $|a|_\infty = \sup_{k\ge 1} |a_k|$ and $|a|_p = (\sum_{k=1}^\infty |a_k|^p)^{1/p}$.

2. Large deviations of empirical processes

We consider the LDP for general triangular arrays of empirical processes. Let $(\Omega_n, \mathcal{A}_n, Q_n)$ be a sequence of probability spaces. Let $(S_{n,j}, \mathcal{S}_{n,j})$ be measurable spaces for $1 \le j \le k_n$, where $\{k_n\}_{n=1}^{\infty}$ is a sequence of positive integers converging to infinity. Let $\{X_{n,j} : 1 \le j \le k_n\}$ be $S_{n,j}$–valued independent r.v.'s defined on Ω_n. To avoid measurability problems, we assume that $\Omega_n = \prod_{j=1}^{k_n} S_{n,j}$, $\mathcal{A}_n = \prod_{j=1}^{k_n} \mathcal{S}_{n,j}$ and $Q_n = \prod_{j=1}^{k_n} \mathcal{L}(X_{n,j})$. Let $f_{n,j}(\cdot, t) : S_{n,j} \to \mathbb{R}$ be a measurable function for each $1 \le j \le k_n$, each $n \ge 1$ and each $t \in T$. Let $U_n(t) := \sum_{j=1}^{k_n} f_{n,j}(X_{n,j}, t)$. To avoid measurability problems, we will assume that the probabilities of $\{U_n(t) : t \in T\}$ are determined by a countable set T_0. Alternatively, we could assume that for each $1 \le j \le k_n$, $\{f_{n,j}(x,t) : t \in T\}$ is an image admissible Suslin class of functions (see page 80 in Dudley, 1999).

First, we present a couple of lemmas that we will need later on.

Lemma 2.1. *Under the previous notation, let $\{\epsilon_n\}$ be sequence of positive numbers converging to zero. Let $0 < c_1, c_2, M_1, M_2 < \infty$. Suppose that*

$$\limsup_{n\to\infty} \epsilon_n \log \left(\Pr\{\sup_{t\in T} |\sum_{j=1}^{k_n} f_{n,j}(X_{n,j}, t)| \ge M_1\} \right) \le -c_1$$

and

$$\liminf_{n\to\infty} \min_{1\le j\le k_n} \epsilon_n \log \left(\Pr\{\sup_{t\in T} |f_{n,j}(X_{n,j}, t)| \le M_2\} \right) \ge -c_2.$$

Then,

$$\limsup_{n\to\infty} \epsilon_n \log \left(\sum_{j=1}^{k_n} \Pr\{\sup_{t\in T} |f_{n,j}(X_{n,j}, t)| \ge 2M_1 + M_2\} \right) \le -(c_1 - c_2).$$

Proof. Let $0 < c_1' < c_1$ and let $c_2' > c_2$. For n large enough,

$$\Pr\{\sup_{t\in T} |\sum_{j=1}^{k_n} f_{n,j}(X_{n,j}, t)| \ge M_1\} \le e^{-c_1' \epsilon_n^{-1}}.$$

Let $\{X'_{n,j} : 1 \le j \le k_n,\ 1 \le n\}$ be an independent copy of $\{X_{n,j} : 1 \le j \le k_n, 1 \le n\}$. Then, for n large enough,

$$\Pr\{\sup_{t\in T} |\sum_{j=1}^{k_n} (f_{n,j}(X_{n,j}, t) - f_{n,j}(X'_{n,j}, t))| \ge 2M_1\} \le 2e^{-c_1' \epsilon_n^{-1}}.$$

By the Lévy inequality (see for example Proposition 2.3 in Ledoux and Talagrand, 1991),

$$\Pr\{\max_{1\le j\le k_n} \sup_{t\in T} |f_{n,j}(X_{n,j}, t) - f_{n,j}(X'_{n,j}, t)| \ge 2M_1\} \le 4e^{-c_1' \epsilon_n^{-1}}.$$

This inequality and Lemma 2.6 in Ledoux and Talagrand (1991) imply that for n large enough,

$$\sum_{j=1}^{k_n} \Pr\{\sup_{t\in T} |f_{n,j}(X_{n,j},t) - f_{n,j}(X'_{n,j},t)| \geq 2M_1\} \leq 8e^{-c'_1\epsilon_n^{-1}}.$$

We also have that for n large enough and each $1 \leq j \leq k_n$,

$$\begin{aligned} & e^{-\epsilon_n^{-1}c'_2} \Pr\{\sup_{t\in T} |f_{n,j}(X_{n,j},t)| \geq 2M_1 + M_2\} \\ \leq\ & \Pr\{\sup_{t\in T} |f_{n,j}(X_{n,j},t)| \geq 2M_1 + M_2\} \Pr\{\sup_{t\in T} |f_{n,j}(X'_{n,j},t)| \leq M_2\} \\ \leq\ & \Pr\{\sup_{t\in T} |f_{n,j}(X_{n,j},t) - f_{n,j}(X'_{n,j},t)| \geq 2M_1\}. \end{aligned}$$

Thus, for n large enough,

$$\sum_{j=1}^{k_n} \Pr\{\sup_{t\in T} |f(X_{n,j},t)| \geq 2M_1 + M_2\} \leq 8e^{-(c'_1-c'_2)\epsilon_n^{-1}}.$$

This implies the claim. □

Lemma 2.2. *Under the notation in Lemma 2.1, let d be a pseudometric in T and let v be a function on T. Suppose that:*
(i) (T,d) is totally bounded.
(ii) For each $\tau > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \Pr\{\sup_{d(s,t)\leq\delta} |U_n(s) - U_n(t)| \geq \tau\} = 0.$$

(iii) For each $t \in T$, $U_n(t) \xrightarrow{\Pr} v(t)$.
(iv) $v(t)$ is uniformly continuous in (T,d).
Then,

$$\sup_{t\in T} |U_n(t) - v(t)| \xrightarrow{\Pr} 0.$$

Proof. We have to prove that for each $\tau > 0$,

$$\limsup_{n\to\infty} \Pr\{\sup_{t\in T} |U_n(t) - v(t)| \geq \tau\} \leq \tau.$$

Take $\delta > 0$ such that $\sup_{d(s,t)\leq\delta} |v(s) - v(t)| \leq \tau/3$ and

$$\limsup_{n\to\infty} \Pr\{\sup_{d(s,t)\leq\delta} |U_n(s) - U_n(t)| \geq \tau/3\} \leq \tau.$$

Hence,

$$\limsup_{n\to\infty} \Pr\{\sup_{d(s,t)\leq\delta} |U_n(s) - v(s) - (U_n(t) - v(t))| \geq 2\tau/3\} \leq \tau. \tag{2.1}$$

Take a function $\pi : T \to T$ with finite range such that $\sup_{t\in T} d(t,\pi(t)) \leq \delta$. By condition (iii),

$$\lim_{n\to\infty} \Pr\{\sup_{t\in T} |U_n(\pi(t)) - v(\pi(t))| \geq \tau/3\} = 0.$$

By (2.1),

$$\limsup_{n\to\infty} \Pr\{\sup_{t\in T} |U_n(t) - v(t) - (U_n(\pi(t)) - v(\pi(t)))| \ge 2\tau/3\} \le \tau.$$

The two last estimations imply the claim. □

We also will need the following exponential inequality for empirical processes

Theorem 2.3. *(Talagrand, 1996, Theorem 1.4). With the above notation, suppose that $E[f(X_i,t)] = 0$, for each $T \in T$ and each $1 \le i \le n$. Then, there exists a universal constant K such that for any $\tau > 0$,*

$$\Pr\{|Z - E[Z]| \ge \tau\} \le K \exp\left(-\frac{\tau}{KC}\log\left(1 + \frac{\tau C}{\sigma^2 + CE[Z]}\right)\right),$$

where $Z = \sup_{t\in T} |\sum_{i=1}^n f(X_i,t)|$, $C = \sup_{t\in T}\sup_{1\le i\le n} \|f(X_i,t)\|_\infty$, and $\sigma^2 = \sup_{t\in T}\sum_{i=1}^n \mathrm{Var}(f(X_i,t))$.

To get the large deviations for the considered empirical processes, we use the following theorem:

Theorem 2.4. *(Arcones, Theorem 3.1, 2002a,). Under the notation above, let $\{\epsilon_n\}$ be a sequence of positive numbers that converges to zero. Let $I : l_\infty(T) \to [0,\infty]$ and let $I_{t_1,\dots,t_m} : \mathbb{R}^m \to [0,\infty]$ be a function, where $t_1,\dots,t_m \in T$. Let d be a pseudometric in T.*

Consider the conditions:

(a.1) (T,d) is totally bounded.

(a.2) For each $t_1,\dots,t_m \in T$, $(U_n(t_1),\dots,U_n(t_m))$ satisfies the LDP with speed ϵ_n and good rate function $I_{t_1,\dots,t_m}$.

(a.3) For each $\tau > 0$,

$$\lim_{\eta\to 0}\limsup_{n\to\infty} \epsilon_n \log \Pr{}^*\{\sup_{d(s,t)\le\eta} |U_n(t) - U_n(s)| \ge \tau\} = -\infty.$$

(b.1) For each $0 \le c < \infty$, $\{z \in l_\infty(T) : I(z) \le c\}$ is a compact set in $l_\infty(T)$.

(b.2) For each $A \subset l_\infty(T)$,

$$\begin{aligned} & -I(A^o) \le \liminf_{n\to\infty} \epsilon_n \log \Pr{}_*\{U_n \in A\} \\ \le\ & \limsup_{n\to\infty} \epsilon_n \log \Pr{}^*\{U_n \in A\} \le -I(\bar{A}). \end{aligned}$$

If the set of conditions (a) is satisfied, then the set of conditions (b) holds with

$$I(z) = \sup\{I_{t_1,\dots,t_m}(z(t_1),\dots,z(t_m)) : t_1,\dots,t_m \in T, m \ge 1\}.$$

If the set of conditions (b) is satisfied, then the set of conditions (a) holds with

$$\begin{aligned} & I_{t_1,\dots,t_m}(u_1,\dots,u_m) \\ =\ & \inf\{I(z) : z \in l_\infty(T), (z(t_1),\dots,z(t_m)) = (u_1,\dots,u_m)\} \end{aligned}$$

and the pseudometric $\rho(s,t) = \sum_{k=1}^\infty k^{-2}\min(\rho_k(s,t),1)$, where $\rho_k(s,t) = \sup\{|u_2 - u_1| : I_{s,t}(u_1,u_2) \le k\}$.

To check condition (a.3) in the previous theorem, we will use the following lemma:

Lemma 2.5. *Under the previous notation, let d be a pseudometric in T such that (T,d) is totally bounded. Suppose that:*

(i)

$$\lim_{\substack{M\to\infty\\ \eta\to 0}} \limsup_{n\to\infty} \epsilon_n \log(\sum_{j=1}^{k_n} \Pr\{F_{n,j}^{(\eta)}(X_{n,j}) \ge M\}) = -\infty,$$

where $F_{n,j}^{(\eta)}(x) = \sup_{d(s,t)\le\eta} |f_{n,j}(x,s) - f_{n,j}(x,t)|$.

(ii) For each $0 < M, \lambda < \infty$,

$$\lim_{\substack{\eta\to 0\\ a\to\infty}} \limsup_{n\to\infty} \epsilon_n \log E[\exp(\epsilon_n^{-1}\lambda \sum_{j=1}^{k_n} F_{n,j}^{(\eta)}(X_{n,j}) I_{M\ge F_{n,j}^{(\eta)}(X_{n,j})\ge a\epsilon_n})] = 0.$$

(iii) For each $a, \eta_0 > 0$,

$$\lim_{\eta\to 0} \limsup_{n\to\infty} \sup_{d(s,t)\le\eta} |\sum_{j=1}^{k_n} E[(f_{n,j}(X_{n,j},s) - f_{n,j}(X_{n,j},t)) I_{F_{n,j}^{(\eta_0)}(X_{n,j})\le a\epsilon_n}]| = 0.$$

(iv) For each $a, \eta_0 > 0$,

$$E[\sup_{t\in T} |\sum_{j=1}^{k_n} (f_{n,j}(X_{n,j},t) I_{F_{n,j}^{(\eta_0)}(X_{n,j})\le a\epsilon_n} - E[f_{n,j}(X_{n,j},t) I_{F_{n,j}^{(\eta_0)}(X_{n,j})\le a\epsilon_n}])|] \to 0.$$

(v) For each $a, \eta_0 > 0$,

$$\lim_{\eta\to 0} \limsup_{n\to\infty} \sup_{d(s,t)\le\eta} \epsilon_n^{-1} \sum_{j=1}^{k_n} \mathrm{Var}((f_{n,j}(X_{n,j},s) - f_{n,j}(X_{n,j},t)) I_{F_{n,j}^{(\eta_0)}(X_{n,j})\le a\epsilon_n}) = 0.$$

Then, for each $\tau, M > 0$,

$$\lim_{\eta\to 0} \limsup_{n\to\infty} \epsilon_n \log \Pr\{\sup_{d(s,t)\le\eta} |\sum_{j=1}^{k_n} (f_{n,j}(X_{n,j},s) - f_{n,j}(X_{n,j},t))| \ge \tau\} = -\infty,$$

and

$$\lim_{\eta\to 0} \limsup_{n\to\infty} \epsilon_n \log \Pr\{\sup_{d(s,t)\le\eta} |\textstyle\sum_{j=1}^{k_n} (f_{n,j}(X_{n,j},s) - f_{n,j}(X_{n,j},t) \\ -E[(f_{n,j}(X_{n,j},s) - f_{n,j}(X_{n,j},t)) I_{F_{n,j}^{(\eta)}(X_{n,j})\le M}])| \ge \tau\} = -\infty.$$

Proof. Note that by conditions (ii) and (iii), the two limits above are equivalent. By (i) and (ii), given $c > 0$, we may take $\eta_0, \lambda, M, a > 0$ such that $\lambda\tau \ge 6c$,

$$\limsup_{n\to\infty} \epsilon_n \log(\sum_{j=1}^{k_n} \Pr\{F_{n,j}^{(\eta_0)}(X_{n,j}) \ge M\}) \le -c$$

and

$$\limsup_{n\to\infty} \epsilon_n \log E[\exp(\epsilon_n^{-1}\lambda \sum_{j=1}^{k_n} F_{n,j}^{(\eta_0)}(X_{n,j}) I_{M > F_{n,j}^{(\eta_0)}(X_{n,j}) \geq a\epsilon_n})] \leq 6^{-1}\lambda\tau.$$

Then,

$$\begin{aligned} &\epsilon_n \log(\Pr\{\sup_{d(s,t)\leq \eta_0} |\textstyle\sum_{j=1}^{k_n}(f_{n,j}(X_{n,j},s) - f_{n,j}(X_{n,j},t)) I_{F_{n,j}^{(\eta_0)}(X_{n,j})\geq M}| \geq 3^{-1}\tau\}) \\ &\leq \epsilon_n \log(\textstyle\sum_{j=1}^{k_n} \Pr\{F_{n,j}^{(\eta_0)}(X_{n,j}) \geq M\}) \leq -c. \end{aligned}$$

We also have that

$$\begin{aligned} &\Pr\{\sup_{d(s,t)\leq \eta_0} |\textstyle\sum_{j=1}^{k_n}(f_{n,j}(X_{n,j},s) - f_{n,j}(X_{n,j},t)) I_{M > F_{n,j}^{(\eta_0)}(X_{n,j}) \geq a\epsilon_n}| \geq 3^{-1}\tau\} \\ \leq\ &\Pr\{\textstyle\sum_{j=1}^{k_n} F_{n,j}^{(\eta_0)} I_{M > F_{n,j}^{(\eta_0)}(X_{n,j}) \geq a\epsilon_n} \geq 3^{-1}\tau\} \\ \leq\ &e^{-3^{-1}\epsilon_n^{-1}\lambda\tau} E[\exp(\epsilon_n^{-1}\lambda \textstyle\sum_{j=1}^{k_n} F_{n,j}^{(\eta_0)}(X_{n,j}) I_{M > F_{n,j}^{(\eta_0)}(X_{n,j}) \geq a\epsilon_n})] \leq e^{-c\epsilon_n^{-1}}. \end{aligned}$$

Hence,

$$\begin{aligned} &\limsup_{n\to\infty} \epsilon_n \log(\Pr\{\sup_{d(s,t)\leq \eta_0} |\textstyle\sum_{j=1}^{k_n}(f_{n,j}(X_{n,j},s) - f_{n,j}(X_{n,j},t)) \\ &\times I_{F_{n,j}^{(\eta_0)}(X_{n,j})\geq a\epsilon_n}| \geq (2/3)\tau\}) \leq -c. \end{aligned}$$

By previous estimations and condition (iii), it suffices to consider $\sup_{d(s,t)\leq\eta} |Z_n(s) - Z_n(t)|$, where

$$Z_n(t) = \sum_{j=1}^{k_n}(f_{n,j}(X_{n,j},t) I_{F_{n,j}^{(\eta_0)}(X_{n,j}) < a\epsilon_n} - E[f_{n,j}(X_{n,j},t) I_{F_{n,j}^{(\eta_0)}(X_{n,j}) < a\epsilon_n}]).$$

It is easy to see that by Theorem 2.3, conditions (iv) and (v) imply that

$$\begin{aligned} &\lim_{\eta\to 0} \limsup_{n\to\infty} \epsilon_n \log(\Pr\{|\sup_{d(s,t)\leq\eta} |Z_n(s) - Z_n(t)| \\ &-E[\sup_{d(s,t)\leq\eta} |Z_n(s) - Z_n(t)|]| \geq 3^{-1}\tau\}) = -\infty. \end{aligned}$$

Hence, the claim follows. □

We also will need the following lemma:

Lemma 2.6. *Let X be a r.v. with values in a measurable space $(S, \mathcal{S})$. Let $f_1, \ldots, f_m$ be measurable functions in S. Let*

$$\begin{aligned} I^{(1)}(u_1, \ldots, u_m) = \sup \quad &\left\{\textstyle\sum_{j=1}^m \lambda_j u_j - \log\left(E\left[\exp\left(\sum_{j=1}^m \lambda_j f_j(X)\right)\right]\right) : \right. \\ &\left. \lambda_1, \ldots, \lambda_m \in \mathbb{R}\right\}, \end{aligned}$$

and let

$$\begin{aligned} I^{(2)}(u_1, \ldots, u_m) = \inf \quad &\{E[h(\gamma(X))] : E[\gamma(X)] = 1 \text{ and} \\ &E[f_j(X)\gamma(X)] = u_j \text{ for each } 1 \leq j \leq m\}, \end{aligned}$$

where

$$h(x) = x\log\left(\frac{x}{e}\right) + 1, \text{ if } x \geq 0; \quad \text{and } h(x) = \infty, \quad \text{if } x < 0. \tag{2.2}$$

(i) If there exists a $\lambda > 0$ such that for each $1 \le j \le m$, $E[e^{\lambda |f_j(X)|}] < \infty$, then for each $u_1, \dots, u_m \in \mathbb{R}$,

$$I^{(1)}(u_1, \dots, u_m) \le I^{(2)}(u_1, \dots, u_m).$$

(ii) If for each $\lambda > 0$ and each $1 \le j \le m$, $E[e^{\lambda |f_j(X)|}] < \infty$, then for each $u_1, \dots, u_m \in \mathbb{R}$,

$$I^{(1)}(u_1, \dots, u_m) = I^{(2)}(u_1, \dots, u_m).$$

Proof. Fix $u_1, \dots, u_m \in \mathbb{R}$. Let $I^{(1)} = I^{(1)}(u_1, \dots, u_m)$ and let $I^{(2)} = I^{(2)}(u_1, \dots, u_m)$. Assume the hypothesis in part (i). Suppose that $E[\gamma(X)] = 1$ and $E[f_j(X)\gamma(X)] = u_j$ for each $1 \le j \le m$. Then, by the Jensen's inequality,

$$\begin{aligned}
&\textstyle\sum_{j=1}^m \lambda_j u_j - E[h(\gamma(X))] = E[\sum_{j=1}^m \lambda_j f_j(X)\gamma(X) - \gamma(X)\log(\gamma(X))] \\
= \;& E\left[\left(\textstyle\sum_{j=1}^m \lambda_j f_j(X) - \log(\gamma(X))\right)\gamma(X)\right] \\
\le \;& \log\left(E\left[\exp\left(\textstyle\sum_{j=1}^m \lambda_j f_j(X) - \log(\gamma(X))\right)\gamma(X)\right]\right) \\
= \;& \log\left(E\left[\exp\left(\textstyle\sum_{j=1}^m \lambda_j f_j(X)\right)\right]\right).
\end{aligned}$$

So, $I^{(1)} \le I^{(2)}$ and (i) follows.

Part (ii) follows from Theorem 5.2 in Donsker and Varadhan (1976). □

Next, we present necessary and sufficient conditions for the large deviations of empirical processes. The set–up for sums of i.i.d. r.v.'s is as follows. Let $(S, \mathcal{S}, \nu)$ be a probability space. Let $\Omega = S^{\mathbb{N}}$, $\mathcal{A} = \mathcal{S}^{\mathbb{N}}$, and $Q = \nu^{\mathbb{N}}$. Let X_n be the n–th projection from Ω into S. Then, $\{X_n\}_{n=1}^\infty$ is a sequence of i.i.d.r.v.'s with values in S. Let $\{f(\cdot, t) : t \in T\}$ be an image admissible Suslin class of measurable functions from S into $\mathbb{R}$. We consider the LDP for $\{n^{-1}\sum_{j=1}^n f(X_j, t) : t \in T\}$. Sethuraman (1964) got the large deviations for the empirical distribution function, that is $T = \mathbb{R}$ and $f(x,t) = I(x \le t)$. The large deviations for general empirical processes was considered by Wu (1994). He obtained necessary and sufficient conditions for a bounded set of functions (Wu, 1994, Theorem 1). But, for unbounded classes, the sufficient conditions in Theorem 4 in Wu (1994) are not necessary. Next theorem gives necessary and sufficient conditions for the large deviations of empirical processes.

Theorem 2.7. *Suppose that $\sup_{t\in T} |f(X,t)| < \infty$ a.s. Then, the following sets of conditions ((a) and (b)) are equivalent:*

(a.1) (T, d) is totally bounded, where $d(s,t) = E[|f(X,s) - f(X,t)|]$.

(a.2) There exists a $\lambda > 0$ such that

$$E[\exp(\lambda F(X))] < \infty,$$

where $F(x) = \sup_{t\in T} |f(x,t)|$.

(a.3) For each $\lambda > 0$, there exists a $\eta > 0$ such that $E[\exp(\lambda F^{(\eta)}(X))] < \infty$, where $F^{(\eta)}(x) = \sup_{d(s,t)\le\eta} |f(x,s) - f(x,t)|$.

(a.4) $\sup_{t\in T} |n^{-1}\sum_{j=1}^n (f(X_j,t) - E[f(X_j,t)])| \xrightarrow{\Pr} 0$.

(b) $\{n^{-1}\sum_{j=1}^n f(X_j,t)\,:\,t\in T\}$ *satisfies the large deviation principle in* $l_\infty(T)$ *with speed* n *and a good rate.*

Moreover, the rate function is given by

$$I(z)=\sup\{I_{t_1,\dots,t_m}(z(t_1),\dots,z(t_m)):t_1,\dots,t_m\in T, m\geq 1\},$$

where

$$\begin{aligned}(2.3)\qquad & I_{t_1,\dots,t_m}(u_1,\dots,u_m)\\ &=\sup\left\{\sum_{j=1}^m \lambda_j u_j-\log(E[\exp(\sum_{j=1}^m\lambda_j f(X,t_j))]):\lambda_1,\dots,\lambda_m\in\mathbb{R}\right\}.\end{aligned}$$

Proof. Assume the set of conditions (a). We apply Theorem 2.4. Condition (a.1) in Theorem 2.4 is obviously satisfied. Condition (a.2) in Theorem 2.4 follows from the Cramér–Chernoff theorem.

Assume (b). Since we have a good rate, for each $0<c<\infty$, there exists $M<\infty$ such that

$$\limsup_{n\to\infty} n^{-1}\log(\Pr\{\sup_{t\in T}|n^{-1}\sum_{j=1}^n f(X_j,t)|\geq M\})\leq -c.$$

By Lemma 2.1, for each $\delta>0$,

$$\limsup_{n\to\infty} n^{-1}\log(n\Pr\{n^{-1}F(X)\geq 2M+\delta\})\leq -c.$$

is easy to see that this implies that $E[e^{\lambda F(X)}]<\infty$, for each $0<\lambda<c(2M+\delta)^{-1}$. So, condition (a.2) holds. Besides the rate function for the finite dimensional distributions is given by (2.3). Hence, by Theorem 2.4, for each $k\geq 1$, $(T,\rho_k^{(1)})$ is totally bounded, where

$$(2.4)\qquad \rho_k^{(1)}(s,t)=\sup\{|u_2-u_1|:I_{s,t}^{(1)}(u_1,u_2)\leq k\},$$

and

$$\begin{aligned}& I_{s,t}^{(1)}(u,v)\\ &=\sup\left\{\lambda_1u_1+\lambda_2u_2-\log(E[\exp(\lambda_1 f(X,s)+\lambda_2 f(X,t))]):\lambda_1,\lambda_2\in\mathbb{R}\right\}.\end{aligned}$$

By Lemma 2.6, for each $s,t\in T$, and each $k>0$, $\rho_k^{(2)}(s,t)\leq\rho_k^{(1)}(s,t)$, where

$$\begin{aligned}(2.5)\qquad & \rho_k^{(2)}(s,t)\\ &=\sup\{|E[\gamma(X)(f(X,t)-f(X,s))]|:E[\gamma(X)]=1,E[h(\gamma(X))]\leq k\}.\end{aligned}$$

Hence, for each $k>0$, $(T,\rho_k^{(2)})$ is totally bounded. Given $1>\delta>0$, there exists a $b>\delta$ such that $E[F(X)I_{F(X)\geq 2^{-1}b}]<2^{-2}\delta$. Hence, for each $s,t\in T$,

$$(2.6)\qquad E[|f(X,t)-f(X,s)|I_{|f(X,t)-f(X,s)|\geq b}]\leq 2^{-1}\delta.$$

Take $k_0>h(4b\delta^{-1})$. Given $s,t,\in T$ with $\rho_{k_0}^{(2)}(s,t)<2^{-2}\delta$. We define

$$\gamma(x)=a^{-1}(f(x,t)-f(x,s))I_{b>f(x,t)-f(x,s)>0},$$

where
$$a = E[(f(X,t) - f(X,s))I_{b>f(X,t)-f(X,s)>0}].$$
If $a^{-1} \le 4\delta^{-1}$, then $E[h(\gamma(X))] \le h(\delta^{-1}4b)$. So,
$$\begin{aligned} E[(f(X,t) - f(X,s))I_{b>f(X,t)-f(X,s)>0}] \\ \le a^{-1}E[(f(X,t) - f(X,s))^2 I_{b>f(X,t)-f(X,s)>0}] \\ = E[(f(X,t) - f(X,s))\gamma(X)] \le 2^{-2}\delta. \end{aligned} \tag{2.7}$$
If $a^{-1} > 4\delta^{-1}$, then (2.7) holds obviously. Combining (2.6) and (2.7), we get that if $\rho^{(2)}_{k_0}(s,t) < 2^{-2}\delta$ then $d(s,t) \le \delta$. Therefore, (T,d) is totally bounded, that is (a.1) holds.

Since (T,d) is totally bounded, by Theorem 2.4, for each $\tau, c < \infty$, there exists $0 < \eta < \infty$ such that
$$\limsup_{n\to\infty} n^{-1} \log \Pr\{\sup_{d(s,t)\le\eta} |n^{-1}\sum_{j=1}^n f(X_j,t)| \ge \tau\}) \le -c.$$
Using a previous argument, this limit and Lemma 2.1 implies that $E[e^{\lambda F^{(\eta)}(X)}] < \infty$, for each $0 < \lambda < c2^{-1}\tau^{-1}$. So, condition (a.3) holds.

Condition (a.4) follows from Lemma 2.2. □

In condition (a.1) in the theorem above, we may use $d(s,t) = (E[|f(X,s) - f(X,t)|^p])^{1/p}$, for any $p \ge 1$, or $d(s,t) = E[|f(X,s) - f(X,t)| \wedge 1]$.

It is not sufficient to have that for some $\lambda > 0$, $E[\exp(\lambda F(X))] < \infty$ to have the large deviations for empirical processes. Let $T = \{0,1,2,\dots\}$ and let $\{\xi_n\}_{n=1}^\infty$ be a sequence of symmetric independent r.v.'s, with $\Pr\{|\xi_0| = 0\} = 1$, and for $n \ge 1$, $\Pr\{|\xi_n| \ge t\} = e^{-t}$ for each $t \ge n$ and $\Pr\{\xi_n = 0\} = 1 - e^{-n}$. Then, for each $0 < \lambda < 1$, $E[\exp(\lambda \sup_{n\ge 0} |\xi_n|)] < \infty$. However, condition (a.3) in Theorem 2.7 does not hold. We have that then $E[(\xi_n - \xi_0)^2] = 4(n_0+1)e^{-n_0}$. For each $\eta > 0$, if n_0 is an integer with $4(n_0+1)e^{-n_0} \le \eta^2$, then $\sup_{m,n,d(m,n)\le\eta} |\xi_m - \xi_n| \ge \sup_{n\ge n_0} |\xi_n|$. However, for $\lambda > 1$ and any n_0, $E[\exp(\lambda \sup_{n\ge n_0} |\xi_n|)] = \infty$.

Conditions (a.2) and (a.3) hold if for each $\lambda > 0$ $E[\exp(\lambda F(X))] < \infty$. However, however there are empirical processes for which the large deviations hold, but it is not true that for each $\lambda > 0$ $E[\exp(\lambda F(X))] < \infty$. Let $\{\xi_k\}_{k=1}^\infty$ be a sequence of symmetric i.i.d.r.v.'s with $\Pr\{|\xi_k| \ge t\} = e^{-t/a_k}$ for each $t > 0$ and each $k \ge 1$, where $\{a_k\}$ is a sequence of positive numbers such that $\sum_{k=1}^\infty a_k^2 < \infty$. Then, there exists a r.v. X and functions $f(x,k)$ such that for each $k \ge 1$ $f(X,k) = \xi_k$. Then, (a.1)–(a.4) in Theorem 2.7 hold, but for $\lambda > \sup_{k\ge1} a_k$, $E[\exp(\lambda \sup_{k\ge1} |\xi_k|)] = \infty$.

By Theorem 4.2 in Arcones (2002b), if for each $\lambda \in \mathbb{R}$ and each $t \in T$, $E[\exp(\lambda f(X,t))] < \infty$, then the rate function in the previous theorem is given by
$$\begin{aligned} I(z) = \inf\{ \quad & E[h(\gamma(X))] : E[\gamma(X)] = 1 \text{ and} \\ & z(t) = E[\gamma(X)f(X,t)] \text{ for each } t \in T\}, \end{aligned}$$

where h is as in (2.2). Observe that $h(y) = \sup_x(xy - (e^x - 1))$.

By Corollary 3.4 in Arcones (2002a), the previous theorem gives necessary and sufficient conditions for the LDP for Banach space values r.v.'s:

Corollary 2.8. *Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d.r.v.'s with values in a separable Banach space B. Then, the following conditions are equivalent:*

(a.1) There exists a $\lambda > 0$ such that

$$E[\exp(\lambda|X|)] < \infty.$$

(a.2) For each $\lambda > 0$, there exists a $\eta > 0$ such that

$$E[\exp(\lambda W^{(\eta)})] < \infty,$$

where

$$W^{(\eta)} = \sup\{|f_1(X) - f_2(X)| : f_1, f_2 \in B_1^*, E[|f_1(X) - f_2(X)|] \le \eta\}.$$

(b) $\{n^{-1}\sum_{j=1}^n X_j\}$ satisfies the LDP in B with speed n.

Proof. By Theorem 2.5, it suffices to show that (a.1) and (a.2) imply that $\{f(X) : f \in B_1^*\}$ is totally bounded in L_1. Conditions (a.1) and (a.2) imply that $E[|X|] < \infty$. Hence, given $\epsilon > 0$, there exists a r.v. $Y = \sum_{j=1}^m x_j I(E_j)$ such that $E[|X-Y|] < \epsilon$, where $x_j \in B$ and $E_1, \dots, E_m$ are disjoint Borel sets. It is easy to see that $\{f(Y) : f \in B_1^*\}$ is totally bounded in L_1. □

If B is a finite dimensional space, then (a.1) in the previous corollary implies (a.2). So, for a finite dimensional Banach space, (a.1) and (b) are equivalent.

The previous theorem relaxes the conditions in Donsker and Varadhan (1976). Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d.r.v.'s with $\Pr\{|X_j| \ge t\} = e^{-t}$, for each $t > 0$ and each $1 \le j$. Then, $f(X_j, t) = |X_j - t|, 0 \le t \le 1$, defines a r.v. with values in $C[0,1]$ with the uniform norm. It is easy to see that Corollary 2.8 applies to this example. However, it is not true that for each $\lambda > 0$, $E[\exp(\lambda \sup_{0 \le t \le 1} |X - t|)] < \infty$. So, the theorem by Donsker and Varadhan (1976) does not apply to this case. Jiang, Bhaskara Rao and Wang (1995) obtained another type of conditions for the large deviations for Banach space random variables. Their conditions are more difficult to check.

The previous corollary gives the following for Hilbert space valued r.v.'s

Corollary 2.9. *Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d.r.v.'s with values in a separable Hilbert space $(H, < \cdot >)$. Then, the following conditions are equivalent:*

(a.1) There exists a $\lambda > 0$ such that

$$E[\exp(\lambda|X|)] < \infty.$$

(a.2) For each $\lambda > 0$, there exists an integer m such that

$$E[\exp(\lambda|X^{(m)}|))] < \infty,$$

where $X^{(m)} = \sum_{k=m+1}^{\infty} < X, h_k > h_k$ and $\{h_k\}$ is an orthogonal basis of H.

(b) $\{n^{-1}\sum_{j=1}^n X_j\}$ satisfies the LDP in B with speed n.

Proof. We show that under (a.1), (a.2) above is equivalent to (a.2) in Corollary 2.8.

Suppose that (a.2) in Corollary 2.9 holds. Then, given $\lambda > 0$, there exists an integer m such that $E[e^{4\lambda|X^{(m)}|}] < \infty$. Let $Y^{(m)} = X - X^{(m)} = \sum_{k=1}^m < X, h_k > h_k$. $Y^{(m)}$ is a finite dimensional r.v. So, there exists a $\eta > 0$ such that $E[e^{2\lambda V^{(m,\eta)}}] < \infty$, where

$$V^{(m,\eta)} = \sup\{|f_1(x) - f_2(x)| : f_1, f_2 \in B_1^*, E[|f_1(X) - f_2(X)|] \leq \eta\}.$$

Since $W^\eta \leq V^{(m,\eta)} + 2|X^{(m)}|$, (a.2) in Corollary 2.8 follows.

Suppose that (a.2) in Corollary 2.8 holds. For each $\lambda > 0$, there exists a $\eta > 0$ such that $E[\exp(\lambda W^{(\eta)})] < \infty$. Take an integer m such that $E[|X^{(m)}|] \leq \eta$, Then, for any $f_1 \in B_1^*$ with $f_1(h_j) = 0$, for each $1 \leq j \leq m$, and $f_2 = 0$, we have $E[|f_1(X) - f_2(X)| \leq E[|X^{(m)}|] \leq \eta$. So, $|X^{(m)}| \geq W^{(\eta)}$ and (a.2) in Corollary 2.9 follows. $\square$

A similar result holds for r.v.'s with values in l_p, $p \geq 1$. In this case, (a.2) in Corollary 2.8 can be substituted by

(a.2)' For each $\lambda > 0$, there exists an integer m such that

$$E[\exp(\lambda|X^{(m)}|)] < \infty,$$

where $X = (Y^{(1)}, Y^{(2)}, \dots,)$ and $X^{(m)} = (0, \dots, 0, Y^{(m+1)}, Y^{(m+2)}, \dots,)$.

We must notice the conditions above are sort of compactness conditions. A set K of a separable Hilbert space is compact if and only if it is closed, bounded and

$$\lim_{m\to\infty} \sup_{x\in K} \sum_{k=m+1}^{\infty} | < x, h_k > |^2 = 0,$$

where $\{h_k\}$ is an orthogonal basis of H. For $p \geq 1$, a set K of l_p is compact if and only if it is closed, bounded and $\lim_{m\to\infty} \sup_{x\in K} \sum_{k=m+1}^{\infty} |x^{(k)}|^p = 0$, where $x = (x^{(1)}, x^{(2)}, \dots)$ (see for example page 6 in Diestel, 1984).

Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of symmetric independent r.v.'s, with $\Pr\{|\xi_n| \geq t\} = e^{-t}$ for each $t \geq n$ and $\Pr\{\xi_n = 0\} = 1 - e^{-n}$. Then, for each $p \geq 1$, $X = (\xi_1, \xi_2, \dots,)$ is a r.v. with values in l_p such that for each $0 < \lambda < 1$, $E[\exp(\lambda|X|_p)] < \infty$, where $|\cdot|_p$ is the l_p norm. However, X does not satisfy (a.2)'.

Our methods also apply to partial sums processes. First, we consider the case of a unique function.

Theorem 2.10. *Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d.r.v.'s. Then, the following conditions are equivalent:*

(a) For each $\lambda \in \mathbb{R}$, $E[\exp(\lambda X)] < \infty$.

(b) $\{n^{-1}\sum_{j=1}^{[nt]} X_j : 0 \leq t \leq 1\}$ satisfies the large deviation principle in $l_\infty([0,1])$ with speed n and a good rate.

Moreover, the rate function is given by

$$I(z) = \begin{cases} \int_0^1 \Psi(z'(t))\,dt, & \text{if } z(0)=0 \\ & \text{and } z(t) \text{ is absolutely continuous,} \\ \infty & \text{else,} \end{cases} \tag{2.8}$$

where $\Psi(x) = \sup_y (xy - \Phi(y))$ *and* $\Phi(y) = \log(E[e^{yX}])$.

Proof. Let $U_n(t) = n^{-1}\sum_{j=1}^{[nt]} X_j$. Assume condition (a). We apply Theorem 2.4 with $d(s,t) = |s-t|$. Obviously $([0,1], d)$ is totally bounded. Given $0 \le t_1 \le \cdots \le t_m \le 1$, we have that

$$\begin{aligned} &\textstyle\sum_{j=1}^m \sum_{i=1}^{[nt_j]} \lambda_j X_i = \sum_{j=1}^m \sum_{p=1}^j \sum_{i=[nt_{p-1}]+1}^{[nt_p]} \lambda_j X_i \\ &= \textstyle\sum_{p=1}^m \sum_{j=p}^m \sum_{i=[nt_{p-1}]+1}^{[nt_p]} \lambda_j X_i = \sum_{p=1}^m \sum_{i=[nt_{p-1}]+1}^{[nt_p]} \sum_{j=p}^m \lambda_j X_i, \end{aligned} \tag{2.9}$$

where $t_0 = 0$. Hence,

$$\begin{aligned} &n^{-1} \log E[\exp(\textstyle\sum_{j=1}^m \sum_{i=1}^{[nt_j]} \lambda_j X_i)] \\ = \; & n^{-1} \textstyle\sum_{p=1}^m \sum_{i=[nt_{p-1}]+1}^{[nt_p]} \log E[\exp(\sum_{j=p}^m \lambda_j X)] \\ &\to \textstyle\sum_{p=1}^m (t_p - t_{p-1}) \log E[\exp(\sum_{j=p}^m \lambda_j X)] = \int_0^1 \Phi(\sum_{j=1}^m \lambda_j I_{0\le s\le t_j})\,ds. \end{aligned} \tag{2.10}$$

This limit and Theorem II.2 in Ellis (1981) imply condition (a.2) in Theorem 2.4. To check condition (a.3) in Theorem 2.4, given $c, \tau > 0$, take $\lambda > 0$ such that $\lambda > 2^3 c\tau^{-1}$ and take an integer

$$m > \max\left(c^{-1}\log(E[\exp(\lambda|X|)]), 2^3\tau^{-1}E[|X|]\right).$$

Let $\{X_i'\}$ be a independent copy of $\{X_i\}$. Let $s_j = m^{-1}j$, for $0 \le j \le m$. Let $\pi(s) = s_j$ if $s_{j-1} \le s < s_j$ for some $j = 1, \dots, m-1$. Let $\pi(s) = s_m$ if $s_{m-1} \le s \le s_m$. By symmetrization (see Lemma 1.2.1 in Giné and Zinn, 1986) and the Lévy inequality

$$\begin{aligned} &\Pr\{\sup_{0\le s\le 1} |U_n(\pi(s)) - U_n(s)| \ge \tau\} \\ &\le m \max_{1\le j\le m} \Pr\{\sup_{s_{j-1}\le s\le s_{j-1}} |\textstyle\sum_{i=[ns_{j-1}]+1}^{[ns]} X_i| \ge n\tau\} \\ &\le 2m \max_{1\le j\le m} \Pr\{\sup_{s_{j-1}\le s\le s_{j-1}} |\textstyle\sum_{i=[ns_{j-1}]+1}^{[ns]} (X_i - X_i')| \ge 2^{-1}n\tau\} \\ &\le 4m \Pr\{|\textstyle\sum_{i=1}^{[m^{-1}n]+2} (X_i - X_i')| \ge 2^{-1}\tau n\} \\ &\le 8m \Pr\{|\textstyle\sum_{i=1}^{[m^{-1}n]+2} X_i| \ge 2^{-2}\tau n\} \\ &\le 8m e^{-\lambda 2^{-2}\tau n} (E[\exp(\lambda|X|)])^{[m^{-1}n]+2}, \end{aligned}$$

for n large enough. Observe that we may symmetrize because

$$E[n^{-1}|\sum_{i=[ns_{j-1}]+1}^{[ns]} X_i|] \le n^{-1}([m^{-1}n]+2)E[|X|] \le 2^{-3}\tau,$$

for n large enough. Hence,

$$\begin{aligned}&\limsup_{n\to\infty} n^{-1}\log(\Pr\{\sup_{0\le s\le 1}|U_n(\pi(s))-U_n(s)|\ge\tau\})\\ &\le -\lambda 2^{-2}\tau+m^{-1}\log(E[\exp(\lambda|X|)])\le -c.\end{aligned}$$

Therefore, (b) follows.

Assume (b). By Theorem 2.4, there exists a pseudometric ρ in $[0,1]$ such that $([0,1]),\rho)$ is totally bounded and for each $\tau>0$, there exists a $\delta>0$ such that

$$\limsup_{n\to\infty} n^{-1}\log\Pr\{\sup_{\substack{0\le s,t\le 1\\ \rho(s,t)\le\delta}}|U_n(s)-U_n(t)|\ge\tau\}\le -1.$$

Take $0\le s<t\le 1$ such that $\rho(s,t)\le\delta$. Then,

$$\limsup_{n\to\infty} n^{-1}\log\Pr\{|\sum_{i=[ns]+1}^{[nt]}X_i|\ge\tau n\}\le -1.$$

By Lemma 2.1, this implies that $E[e^{\lambda X}]<\infty$, for each $|\lambda|<\tau^{-1}$. So, condition (a) holds.

The results in Section 4 in Arcones (2002b) give that the rate function is the one claimed. □

Large deviations for partial sums processes in a similar set–up to the previous one have been considered by Borovkov (1967) and Mogulskii (1976).

For partial sums of empirical processes, we present the following theorem:

Theorem 2.11. *Let $\{X_j\}$ be a sequence of i.i.d.r.v.'s with values in a measurable space $(S,\mathcal{S})$. Let $\{f(x,t):t\in T\}$ be a class of measurable functions defined on S such that $\sup_{t\in T}|f(X,t)|<\infty$ a.s. Let $U_n(s,t)=n^{-1}\sum_{j=1}^{[ns]}f(X_j,t)$. Then, the following sets of conditions (a) and (b) are equivalent:*

(a.1) (T,d) is totally bounded, where $d(s,t)=E[|f(X,s)-f(X,t)|]$.

(a.2) For each $\lambda>0$,

$$E[\exp(\lambda F(X))]<\infty,$$

where $F(x)=\sup_{t\in T}|f(x,t)|$.

(a.3) $\sup_{t\in T}|n^{-1}\sum_{j=1}^{n}(f(X_j,t)-E[f(X_j,t)])|\xrightarrow{\Pr}0$.

(b) $\{U_n(s,t):t\in T\}$ satisfies the large deviation principle in $l_\infty([0,1]\times T)$ with speed n and a good rate.

Moreover, the rate function is given by

$$\begin{aligned}(2.11)\quad I(z)=\inf\{&\textstyle\int_0^1 E[h(\gamma(X,u))]\,du : E[\gamma(X,u)]=1 \text{ for each } 0\le u\le 1,\\ &\text{and } z(s,t)=\textstyle\int_0^s E[f(X,t)\gamma(X,u)]\,du \text{ for each } (s,t)\in[0,1]\times T\}.\end{aligned}$$

Proof. The proof of (a) is equivalent to (b) is similar to that of Theorem 2.7 and it is omitted. The new part is the form of the rate function. By the argument in (2.9) for $0 \le s_1 \le \cdots \le s_m \le 1$ and $t_1, \dots, t_m \in T$,

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\lambda_{j,k}\sum_{i=1}^{[ns_j]} f(X_i,t_k) = \sum_{p=1}^{m}\sum_{i=[s_{p-1}]+1}^{[ns_p]}\sum_{j=p}^{m}\sum_{k=1}^{m}\lambda_{j,k} f(X_i,t_k),$$

where $s_0 = 0$. Hence,

$$\begin{aligned}
& n^{-1}\log(E[\exp(\textstyle\sum_{j=1}^{m}\sum_{k=1}^{m}\lambda_{j,k}\sum_{i=1}^{[ns_j]} f(X_i,t_k))]) \\
= \;& n^{-1}\textstyle\sum_{p=1}^{m}\sum_{i=[s_{p-1}]+1}^{[ns_p]}\log(E[\exp(\sum_{j=p}^{m}\sum_{k=1}^{m}\lambda_{j,k}f(X_i,t_k))]) \\
\to \;& \textstyle\sum_{p=1}^{m}(s_p - s_{p-1})\log\left(E\left[\exp\left(\sum_{j=p}^{m}\sum_{k=1}^{m}\lambda_{j,k}f(X,t_k)\right)\right]\right) du \\
= \;& \textstyle\int_0^1 \log\left(E\left[\exp\left(\sum_{j=1}^{m}\sum_{k=1}^{m}\lambda_{j,k}I(0\le u\le s_j)f(X,t_k)\right)\right]\right) du.
\end{aligned}$$

Hence, by Theorem II.2 in Ellis (1981), the rate function for the LDP of

$$\{(U_n(s_1,t_1),\dots,U_n(s_1,t_m),\dots,U_n(s_m,t_1),\dots,U_n(s_m,t_m))\}$$

is given by

$$\begin{aligned}
\sup \quad & \Big\{\textstyle\sum_{j=1}^{m}\sum_{k=1}^{m}\lambda_{j,k}u_{j,k} \\
& - \textstyle\int_0^1 \log\left(E\left[\exp\left(\sum_{j=1}^{m}\sum_{k=1}^{m}\lambda_{j,k}I(0\le u\le s_j)f(X,t_k)\right)\right]\right) du : \lambda_{j,k}\in\mathbb{R}\Big\}.
\end{aligned}$$

The arguments in Lemma 2.6 give that

$$\begin{aligned}
& I_{(s_1,t_1),\dots,(s_m,t_m)}(u_{1,1},\dots,u_{m,m}) \\
:= \;& \inf\{\textstyle\sum_{p=1}^{m}(s_p-s_{p-1})E[h(\gamma_p(X))] : E[\gamma_j(X)] = 1 \text{ for each } 1\le j\le m\} \\
& \text{ and } \textstyle\sum_{p=1}^{j}(s_p-s_{p-1})E[f(X,t_k)\gamma_p(X)]\,du = u_{j,k} \text{ for each } 1\le j,k\le m\} \\
= \;& \inf\{\textstyle\int_0^1 E[h(\gamma(u,X))]\,du : E[\gamma(X,u)] = 1 \text{ for each } 0\le u\le 1, \\
& \text{ and } \textstyle\int_0^{s_j} E[f(X,t_k)\gamma(X,u)]\,du = u_{j,k} \text{ for each } 1\le j,k\le m\}.
\end{aligned}$$

It is easy to that the methods in Section 4 in Arcones (2002b), give that

$$\sup\{I_{(s_1,t_1),\dots,(s_m,t_m)}(z(s_1,t_1),\dots,z(s_m,t_m)) : s_i\in[0,1], t_i\in T, \text{ for } 1\le i\le m\},$$

is the rate in (2.11). □

For Banach space r.v.'s, we have that:

Corollary 2.12. *Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d.r.v.'s with values in a separable Banach space B. $l_\infty([0,1],B)$ is the Banach space consisting by the bounded functions from $[0,1]$ into B with the norm $\|x\|_{\infty,T,B} = \sup_{t\in T}|x(t)|$. Then, the following conditions are equivalent:*

(a) For each $\lambda > 0$ $E[\exp(\lambda|X|)] < \infty$.

(b) $\{n^{-1}\sum_{j=1}^{[nt]} X_j : 0\le t\le 1\}$ satisfies the LDP in $l_\infty([0,1],B)$ with speed n and a good rate function.

Moreover, the rate function is

$$I(z) = \inf\{\int_0^1 E[h(\gamma(X,u))]\,du : E[\gamma(X,u)] = 1 \text{ for each } 0 \le u \le 1, \text{ and } z(t) = \int_0^t E[\gamma(X,u)X]\,du \text{ for each } t \in [0,1]\}.$$

References

[1] M. A. Arcones, *The large deviation principle for stochastic processes I.* To appear in Theor. Probab. Applic. (2002a).

[2] M. A. Arcones, *The large deviation principle for stochastic processes II.* To appear in Theor. Probab. Applic. (2002b).

[3] R. R. Bahadur, *Some Limit Theorems in Statistics.* SIAM, Philadelphia, PA. 1971.

[4] R. R. Bahadur, and S. L. Zabell, *Large deviations of the sample mean in general vector spaces.* Ann. Probab. **57** (1979), 587-621.

[5] A. A. Borovkov, *Boundary–value problems for random walks and large deviations in function spaces.* Theor. Probab. Appl. **12** (1967), 575–595.

[6] H. Chernoff, *A measure of asymptotic efficiency for tests of hypothesis based on the sum of observations.* Ann. Math. Statist. **23** (1952), 493–507.

[7] H. Cramér, *Sur un nouveau théorèm limite de la théorie des probabilities.* Actualités Sci. Ind. **736** (1938), 5–23.

[8] J. D. Deuschel, and D. W. Stroock, *Large Deviations.* (1989), Academic Press, Inc., Boston, MA.

[9] J. Diestel, *Sequences and Series in Banach Spaces.* Springer–Verlag, New York. 1984.

[10] M. D. Donsker, and S. R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time III.* Comm. Pure Appl. Math. **29** (1976), 389–461.

[11] R. M. Dudley, *Uniform Central Limit Theorems.* Cambridge University Press, Cambridge. 1999.

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

[13] E. Giné, and J. Zinn, Lectures on the central limit theorem for empirical processes. Lect. Notes in Math. **1221** (1986), 50–112. Springer, New York.

[14] T. Jiang; M. Bhaskara Rao, and X. P. Wang, *Large deviations for moving average processes.* Stoch. Proc. Appl. **59** (1995), 309–320.

[15] M. Ledoux, and M. Talagrand, *Probability in Banach Spaces.* (Springer–Verlag, New York. 1991.

[16] A. A. Mogulskii, *Large deviations for trajectories of multidimensional random walks.* Theor. Probab. Appl. **21** (1976), 300–315.

[17] J. Sethuraman, *On the probability of large deviations of families of sample means.* Ann. Math. Statist. **35** (1964), 1964 1304–1316.

[18] J. Sethuraman, *On the probability of large deviations of of mean for random variables in $D[0,1]$.* Ann. Math. Statist. **36** (1965), 280–285.

[19] J. Sethuraman, *Corrections to: "On the probability of large deviations of families of sample means".* Ann. Math. Statist. **41** (1970), 1376–1380.

[20] M. Talagrand, *New concentration inequalities in product spaces.* Inv. Math. **126** (1996), 505–563.

[21] S. R. S. Varadhan, *Large Deviations and Applications*. SIAM, Philadelphia, Pennsylvania. (1984).

[22] L. Wu, *Large deviations, moderate deviations and LIL for empirical processes.* Ann. Probab. **22** (1994), 17–27.

Department of Mathematical Sciences, Binghamton University, Binghamton, NY 13902,
E-mail address: arcones@math.binghamton.edu