Electron. Commun. Probab. **20** (2015), no. 83, 1–8.
DOI: 10.1214/ECP.v20-4613
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

An Erdös-Rényi law for nonconventional sums

Yuri Kifer*

Abstract

We obtain the Erdös-Rényi type law of large numbers for "nonconventional" sums of the form $S_n = \sum_{m=1}^{n} F(X_m, X_{2m}, ..., X_{\ell m})$ where $X_1, X_2, ...$ is a sequence of i.i.d. random variables and F is a bounded Borel function. The proof relies on nonconventional large deviations obtained in [8].

Keywords: laws of large numbers; large deviations; nonconventional setup.
AMS MSC 2010: Primary 60F15, Secondary 60F10.
Submitted to ECP on October 7, 2015, final version accepted on November 1, 2015.

1 Introduction

Let $X_1, X_2, ...$ be a sequence of independent identically distributed (i.i.d.) random variables such that $EX_1 = 0$ and the moment generating function $\phi(t) = Ee^{tX_1}$ exists. Denote by I the Legendre transform of $\ln \phi$ and set $S_n = \sum_{m=1}^{n} X_m$ for $n \geq 1$ and $S_0 = 0$. The Erdös-Rényi law of large numbers from [4] says that with probability one

$$I(\alpha) \lim_{n\to\infty} \max_{0\leq m\leq n-[\frac{\ln n}{I(\alpha)}]} \frac{S_{m+[\frac{\ln n}{I(\alpha)}]} - S_m}{\ln n} = \alpha \tag{1.1}$$

for all $\alpha > 0$ in some neighborhood of zero.

The nonconventional limit theorems initiated in [5] and partially motivated by nonconventional ergodic theorems study asymptotic behaviors of sums of the form $S_n = \sum_{m=1}^{n} F(X_m, X_{2m}, ..., X_{\ell m})$ and more general ones where F is a Borel function. In this paper we will obtain an Erdös-Rényi law similar to (1.1) for such sums where $X_1, X_2, ...$ is again a sequence of i.i.d. random variables and F is a bounded Borel function. Observe that summands in nonconventional sums are long range dependent so this result cannot be derived directly from existing literature. On the other hand, as most proofs of the Erdös-Rényi law we will rely on large deviations which in the nonconventional setup were obtained in [8].

2 Preliminaries and main results

Let $X_1, X_2, ...$ be a sequence of i.i.d. random variables and F be a bounded Borel function on $\mathbb{R}^\ell$ such that

$$\bar{F} = EF(X_1, X_2, ..., X_\ell) = 0 \quad \text{and} \quad \sigma^2 = EF^2(X_1, X_2, ..., X_\ell) > 0. \tag{2.1}$$

*Institute of Mathematics, Hebrew University, Jerusalem, Israel. E-mail: kifer@math.huji.ac.il

The first condition in (2.1) is not a restriction since we always can consider $F - \bar{F}$ in place of F and the second condition there means that F is not a constant almost surely (a.s.) with respect to the ℓ-product measure $\mu^{(\ell)} = \mu \times \mu \times \cdots \times \mu$ on $\mathbb{R}^\ell$ where μ is the distribution of X_1. Set also $M = \|F\|_\infty$ and $M_+ = \|F_+\|_\infty$ where $F_+(x_1, ..., x_\ell) = \max(0, F(x_1, ..., x_\ell))$ and the L^∞ norm on $\mathbb{R}^\ell$ is considered with respect to the measure $\mu^{(\ell)}$. Introduce the moment generating function $\phi(t) = E \exp(tF(X_1, X_2, ..., X_\ell))$ and its Legendre transform

$$I(\alpha) = sup_t(t\alpha - \ln \phi(t)). \tag{2.2}$$

Theorem 2.1. *With I given by (2.2) the Erdös-Rényi law (1.1) holds true also for the nonconventional sums $S_n = \sum_{m=1}^{n} F(X_m, X_{2m}, ..., X_{\ell m})$ for all $\alpha \in (0, M_+)$ where we set also $S_0 = 0$.*

Our proof of Theorem 2.1 will follow the scheme of [2] but we will rely also on nonconventional large deviations results from [8]. As in some books and many papers on large deviations we did not address explicitly in [8] the crucial question when the rate function of large deviations is positive without which the large deviations principle is meaningless since it does not lead to any nontrivial estimates for the domains were the rate function is zero. We will rely on the following theorem which specifies further the results of [8] and actually provides more information than we need for the proof of Theorem 2.1.

Theorem 2.2. *The limit*

$$Q(\lambda F) = \lim_{N\to\infty} \frac{1}{N} \ln E \exp(S_N(\lambda F)) \tag{2.3}$$

exists where $Q(\lambda F)$ is a $\mathbb{C}^\infty$ function of λ with bounded derivatives and S_n, $n \geq 1$ are nonconventional sums from Theorem 2.1. The Legendre transform of Q,

$$J(u) = \sup_{\lambda}(\lambda u - Q(\lambda F)) \tag{2.4}$$

is a nonnegative, convex, lower semi-continuous function such that $J(u) = 0$ if and only if $u = 0$ and $J(u)$ is strictly increasing for $u \geq 0$ (writing for convenience $\infty > \infty$) while it is strictly decreasing for $u \leq 0$. In addition, if $M_+ = \|F_+\|_\infty > 0$ ($M_- = \|F - F_+\|_\infty > 0$) then there exists $L_+ > 0$ ($L_- > 0$) such that $J(u) < \infty$ when $u \in [0, L_+)$ ($u \in (-L_-, 0]$) and $J(u) = \infty$ when $u > L_+$ ($u < -L_-$). Furthermore, the sums S_n, $n \geq 1$ satify the large deviations principle in the form

$$\limsup_{N\to\infty} \frac{1}{N} \ln P\{\frac{1}{N} S_N \in K\} \leq -\inf_{u\in K} J(u) \tag{2.5}$$

for any closed set $K \subset \mathbb{R}$ while for any open set $U \subset \mathbb{R}$,

$$\liminf_{N\to\infty} \frac{1}{N} \ln P\{\frac{1}{N} S_N \in U\} \geq -\inf_{u\in U} J(u). \tag{2.6}$$

Remark 2.3. Theorem 2.1 shows that the Erdös-Rényi law for nonconventional sums has the same form as for sums of i.i.d. random variables having the same distribution as $F(X_1, X_2, ..., X_\ell)$. This is similar to the nonconventional strong law of large numbers proved in [6]. On the other hand, the nonconventional central limit theorem and the nonconventional large deviations estimates are somewhat different from the corresponding results for sums of i.i.d. random variables. In particular, it is shown in [7] that the nonconventional functional central limit theorem may yield in the limit a process with dependent increments while concerning large deviations it follows from [8] that the rate functions I and J above are, in general, different.

3 Proof of Theorem 2.1

Let $Y_1, Y_2, ...$ be a sequence of i.i.d. random variables which have the same distribution as $F(X_1, X_2, ..., X_\ell)$ and set $\Sigma_n = \sum_{m=1}^{n} Y_m$. We will need the classical Cramér large deviation estimates in the form (see, for instance, Section 2.2 in [3]),

$$\limsup_{N\to\infty} \frac{1}{N} \ln P\{\frac{1}{N}\Sigma_N \in K\} \leq -\inf_{u\in K} I(u) \tag{3.1}$$

for any closed set $K \subset \mathbb{R}$ while for any open set $U \subset \mathbb{R}$,

$$\liminf_{N\to\infty} \frac{1}{N} \ln P\{\frac{1}{N}\Sigma_N \in U\} \geq -\inf_{u\in U} I(u) \tag{3.2}$$

where I is given by (2.2).

It is essential to observe that $I(\alpha) > 0$ ($I(\alpha) = \infty$ is possible) unless $\alpha = 0$ which is well known and follows, in particular, from Theorem II.6.3 in [1] (which relies on general convex analysis results) but it has also a simple direct explanation in our case. Indeed, since $\ln\phi(0) = (\ln\phi(t))'_{t=0} = 0$ then $\ln\phi(t) = o(t)$ for small t. Hence, if $\alpha \neq 0$ then $t\alpha > \ln\phi(t)$ either for small positive or for small negative t, and so in view of (2.2), $I(\alpha) = 0$ only when $\alpha = 0$ and otherwise $I(\alpha)$ is positive. By (2.1) and the Jensen inequality $\ln\phi(t) \geq tEF(X_1, ..., X_\ell) = 0$, and so (see Lemma 2.2.5 in [3]),

$$I(\alpha) = sup_{t\geq 0}(t\alpha - \ln\phi(t)) \text{ if } \alpha \geq 0 \text{ and } I(\alpha) = sup_{t\leq 0}(t\alpha - \ln\phi(t)) \text{ if } \alpha \leq 0. \tag{3.3}$$

For each $\alpha > 0$ there exists a sequence $t_n \to t_0$ as $n \to \infty$ such that $I(\alpha) = \lim_{n\to\infty}(t_n\alpha - \ln\phi(t_n))$ where $t_0 > 0$ ($t_0 = \infty$ is possible) since by above $I(\alpha) > 0$. Therefore, for any $\Delta > 0$,

$$I(\alpha + \Delta) \geq \lim_{n\to\infty}(t_n(\alpha + \Delta) - \ln\phi(t_n)) = I(\alpha) + t_0\Delta$$

which means that $I(\alpha)$ is strictly increasing for $\alpha \geq 0$. Similarly, $I(\alpha)$ is strictly decreasing for $\alpha \leq 0$. Observe that, in fact, for any $\varepsilon > 0$,

$$e^{tM_+} + 1 \geq \phi(t) \geq P\{F(X_1, ..., X_\ell) \geq M_+ - \varepsilon\}e^{t(M_+ - \varepsilon)} \text{ if } t \geq 0 \text{ and}$$

$$e^{-tM_-} + 1 \geq \phi(t) \geq P\{-F(X_1, ..., X_\ell) \geq M_- - \varepsilon\}e^{-t(M_- - \varepsilon)} \text{ if } t \leq 0$$

where $M_- = \|F - F_+\|_\infty$. This together with (2.2) yields that $I(\alpha) < \infty$ if $-M_- < \alpha < M_+$ while $I(\alpha) = \infty$ if $\alpha > M_+$ or $\alpha < -M_-$. Similar arguments relying on explicit formulas from [8] yield Theorem 2.2 but for now we will take it for granted in order to prove Theorem 2.1.

Fix $\alpha \in (0, M)$ and let $b_n = [\ln n / I(\alpha)]$. Choose $\varepsilon > 0$ and define the event

$$A_n(\varepsilon) = \{\max_{0\leq m\leq n-b_n}(S_{m+b_n} - S_m) \geq (\alpha + \varepsilon)b_n\}.$$

Then

$$P(A_n(\varepsilon)) = P\{\bigcup_{0\leq m\leq n-b_n}\{S_{m+b_n} - S_m \geq (\alpha + \varepsilon)b_n\}\} \leq \sum_{m=0}^{n-b_n} P\{S_{m+b_n} - S_m \geq (\alpha + \varepsilon)b_n\}. \tag{3.4}$$

Observe that when $m > (\ell - 1)b_n$ then

$$S_{m+b_n} - S_m = \sum_{k=m+1}^{m+b_n} F(X_k, X_{2k}, ..., X_{\ell k})$$

is the sum of i.i.d. random variables having the same distribution as $F(X_1, X_2, ..., X_\ell)$. Indeed, if $\ell = 1$ this is clear and if $\ell > 1$ then the equality $ik = j\tilde{k}$ is impossible for integers $(\ell-1)b_n \le m < \tilde{k} < k \le m + b_n$ and $1 \le i < j \le \ell$ since $j/i \ge 1 + (\ell-1)^{-1}$ while $k/\tilde{k} < 1 + b_n/m \le 1 + (\ell-1)^{-1}$. Hence, we can use Cramér's upper large deviations bound (3.1) to conclude that for any $m \ge (\ell-1)b_n$,

$$P\{S_{m+b_n} - S_m \ge (\alpha+\varepsilon)b_n\} \le \exp(-b_n(I(\alpha+\varepsilon)-\delta)) \le \exp(-b_n(I(\alpha)+\delta)) \tag{3.5}$$

where $0 < \delta < \frac{1}{2}(I(\alpha+\varepsilon) - I(\alpha))$, $n \ge n(\delta)$ is large enough and we use the fact that $I(\beta)$ is strictly increasing when $\beta \ge 0$.

On the other hand, if $m \le (\ell-1)b_n$ and $\ell > 1$ then we write

$$\begin{aligned} P\{S_{m+b_n} - S_m \ge (\alpha+\varepsilon)b_n\} \le P\{S_{m+b_n} \ge \tfrac{1}{2}(\alpha+\varepsilon)b_n\} \\ +P\{-S_m \ge \tfrac{1}{2}(\alpha+\varepsilon)b_n\} \le P\{\tfrac{1}{m+b_n}S_{m+b_n} \ge \tfrac{1}{2\ell}(\alpha+\varepsilon)\} \\ +P\{-\tfrac{1}{m}S_m \ge \tfrac{1}{2m}(\alpha+\varepsilon)b_n\}. \end{aligned} \tag{3.6}$$

Applying the upper nonconventional large deviations bound (2.5) we obtain for $m \le (\ell-1)b_n$ that

$$\begin{aligned} P\{\tfrac{1}{m+b_n}S_{m+b_n} \ge \tfrac{1}{2\ell}(\alpha+\varepsilon)\} \\ \le \exp(-(m+b_n)(J(\tfrac{1}{2\ell}(\alpha+\varepsilon)) - \delta)) \le \exp(-\tfrac{1}{2}b_nJ(\tfrac{\alpha}{2\ell})) \end{aligned} \tag{3.7}$$

where $0 < \delta < J(\frac{1}{2\ell}(\alpha+\varepsilon)) - J(\frac{\alpha}{2\ell})$, $n \ge n(\delta)$ is large enough and we use that $J(\beta)$ is strictly increasing when $\beta \ge 0$.

Since $|S_m| \le mM$ a.s. then

$$P\{-\frac{1}{m}S_m \ge \frac{1}{2m}(\alpha+\varepsilon)b_n\} = 0 \text{ if } m < \frac{\alpha}{2M}b_n. \tag{3.8}$$

Now assume that $(\ell-1)b_n \ge m \ge \frac{\alpha}{2M}b_n$ and $\ell > 1$. Observe that $-S_m = \sum_{k=1}^m(-F(X_k, X_{2k}, ..., X_{\ell k}))$, and so we can consider nonconventional large deviations estimates of Theorem 2.2 for the case where F is replaced by $-F$ with a corresponding rate function $\hat{J}$ having the same properties as J. Then we obtain

$$\begin{aligned} P\{-\tfrac{1}{m}S_m \ge \tfrac{1}{2m}(\alpha+\varepsilon)b_n\} \\ \le P\{-\tfrac{1}{m}S_m \ge \tfrac{1}{2(\ell-1)}(\alpha+\varepsilon)\} \\ \le \exp(-m(\hat{J}(\tfrac{1}{2(\ell-1)}(\alpha+\varepsilon)) - \delta)) \le \exp(-b_n\tfrac{\alpha}{2M}\hat{J}(\tfrac{\alpha}{2\ell})) \end{aligned} \tag{3.9}$$

where $0 < \delta < \hat{J}(\frac{1}{2(\ell-1)}(\alpha+\varepsilon)) - \hat{J}(\frac{\alpha}{2\ell})$, $n \ge n(\delta)$ is large enough and we use that $\hat{J}(\beta)$ is strictly increasing when $\beta \ge 0$ (of course, if $\hat{J}(\frac{1}{2(\ell-1)}(\alpha+\varepsilon)) = \infty$ then any δ will do).

Set $c = c_\alpha = \frac{1}{2}\min(J(\frac{\alpha}{2\ell}), \frac{\alpha}{M}\hat{J}(\frac{\alpha}{2\ell}))$ which is a positive number. Then it follows from (3.4)–(3.9) that for δ satisfying (3.5) and for n large enough,

$$\begin{aligned} P(A_n(\varepsilon)) \le n\exp(-b_n(I(\alpha)+\delta)) + \ell b_n\exp(-cb_n) \\ \le n\exp\big(-(\tfrac{\ln n}{I(\alpha)} - 1)(I(\alpha)+\delta)\big) + \ell(\tfrac{\ln n}{I(\alpha)}+1)\exp(-c(\tfrac{\ln n}{I(\alpha)}-1)) \\ = e^{I(\alpha)+\delta}n^{-\frac{\delta}{I(\alpha)}} + \ell(\tfrac{\ln n}{I(\alpha)}+1)e^c n^{-\frac{c}{I(\alpha)}}. \end{aligned} \tag{3.10}$$

Now let $d > I(\alpha)\max(\delta^{-1}, c^{-1})$. Then

$$\sum_{n=1}^{\infty}(n^{-\frac{d\delta}{I(\alpha)}} + n^{-\frac{dc}{I(\alpha)}}\ln n) < \infty$$

which together with (3.10) and the Borel-Cantelli lemma yields that with probability one $A_{n^d}(\varepsilon)$ occurs only finitely often. Hence, setting $a_n = b_{n^d}$ we obtain

$$\limsup_{n\to\infty} \max_{0\le m\le n^d-a_n} \frac{S_{m+a_n}-S_m}{a_n} \le \alpha+\varepsilon. \tag{3.11}$$

Since for $n^d < r \le (n+1)^d$ large enough the difference $a_n - b_r$ is bounded by 1 then it follows that

$$\limsup_{r\to\infty} \max_{0\le m\le r-b_r} \frac{S_{m+b_r}-S_m}{b_r} \tag{3.12}$$
$$\le \limsup_{n\to\infty} \max_{0\le m\le (n+1)^d-a_{n+1}} \frac{S_{m+a_{n+1}}-S_m+M}{a_n} \le \alpha+\varepsilon.$$

In order to derive the lower bound choose $\varepsilon > 0$ so that $\alpha - \varepsilon > 0$ and define

$$B_n(\varepsilon) = \{\max_{0\le m\le n-b_n}(S_{m+b_n}-S_m) \le b_n(\alpha-\varepsilon)\}.$$

Let $C_m = \{S_{m+b_n} - S_m \le b_n(\alpha-\varepsilon)\}$. Then

$$P(B_n(\varepsilon)) = P(\bigcap_{0\le m\le n-b_n} C_m) \le P(\bigcap_{(1-\ell^{-1})n\le m\le n-b_n} C_m). \tag{3.13}$$

Observe that when $n-b_n \ge m, \tilde{m} \ge (1-\ell^{-1})n$ then $\frac{m}{\tilde{m}} < \frac{\ell}{\ell-1}$ if $m > \tilde{m}$, and so the equality $im = j\tilde{m}$ for integers $n \ge m > \tilde{m} \ge (1-\ell^{-1})n$ and $\ell \ge j > i \ge 1$ is impossible since then $\min \frac{j}{i} = \frac{\ell}{\ell-1}$. Hence, all $F(X_k, X_{2k}, ..., X_{\ell k})$, $(1-\ell^{-1})n \le k \le n-b_n$ are independent, and so all events C_m, $(1-\ell^{-1})n \le m \le n-b_n$ are independent. Hence by (3.13),

$$P(B_n(\varepsilon)) \le \prod_{(1-\ell^{-1})n\le m\le n-b_n} P(C_m). \tag{3.14}$$

Taking into account that $S_{m+b_n} - S_m$ is a sum of i.i.d. random variables having the same distribution as $F(X_1, X_2, ..., X_\ell)$ when $(1-\ell^{-1})n \le m \le n-b_n$ we obtain by Cramér's lower large deviations bound (3.2) that

$$P(\Omega \setminus C_m) \ge \exp(-b_n(I(\alpha-\varepsilon)+\delta)) \ge \exp(-b_n I(\alpha)(1-\delta)) \ge \exp(-(1-\delta)\ln n) \tag{3.15}$$

where we choose $\delta > 0$ so small that $(I(\alpha-\varepsilon)+\delta)/I(\alpha) < 1-\delta$ which is possible since $I(\beta)$ is strictly increasing for $\beta \ge 0$. Hence, if n is sufficiently large,

$$P(B_n(\varepsilon)) \le (1-\exp(-(1-\delta)\ln n)^{\frac{n}{2\ell}} = (1-n^{-(1-\delta)})^{\frac{n}{2\ell}} \tag{3.16}$$
$$= \left((1-n^{-(1-\delta)})^{n^{1-\delta}}\right)^{\frac{n^\delta}{2\ell}} = O(\exp(-n^{\delta/2}).$$

It follows that

$$\sum_{n=1}^{\infty} P(B_n(\varepsilon)) < \infty$$

and by the Borel-Cantelli lemma with probability one $B_n(\varepsilon)$ occurs only finitely often which implies that

$$\liminf_{n\to\infty} \max_{0\le m\le n-b_n} \frac{S_{m+b_n}-S_m}{b_n} \ge \alpha+\varepsilon. \tag{3.17}$$

Since $\varepsilon > 0$ can be chosen arbitrarily small we obtain the assertion of Theorem 2.1 from (3.12) and (3.17). □

4 Proof of Theorem 2.2

Theorem 2.2 mostly follows from the results of [8] together with Theorem II.6.3 from [1] but for reader's convenience we will give a direct argument here. First, we recall relevant notations and formulas from [8]. Let $r_1, ..., r_m \geq 2$ be all primes not exceeding ℓ. Set $A_n = \{a \leq n : a \text{ is coprime with } r_1, ..., r_m\}$ and $B_n(a) = \{b \leq n : b = ar_1^{d_1} r_2^{d_2} \cdots r_m^{d_m}$ for some nonnegative integers $d_1, ..., d_m\}$. For any function V on $\mathbb{R}^\ell$ we write

$$S_N(V) = \sum_{a \in A_N} S_{N,a}(V) \text{ where } S_{N,a}(V) = \sum_{b \in B_N(a)} V(X_b, X_{2b}, ..., X_{\ell b})$$

observing that S_n from Theorem 2.2 equals $S_n(F)$ here.

The existence of the limit (2.3) was proved in [8]. Recall, that convexity and lower semi-continuity of the Legendre transform $J(u)$ of $Q(\lambda F)$ follows from (2.3) and (2.4) automatically (see Theorem II.6.1 in [1]). Observe that by (2.1) and the Jensen inequality $Q(\lambda F) \geq 0$ and since $Q(\lambda F) \leq |\lambda| M$ then $J(u) = \infty$ when $u > M$. Note that Theorem 2.7 in [8] is formulated for continuous functions but, in fact, only boundedness of functions is used in the proof so we can apply it to our setup where $\|F\|_\infty = M < \infty$.

In order to exhibit an explicit formula for $Q(\lambda F)$ obtained in [8] introduce

$$D(\rho) = \{n = (n_1, ..., n_m) \in \mathbb{Z}^m : n_1, ..., n_m \geq 0, \text{ and } \sum_{i=1}^m n_i \ln r_i \leq \rho\}$$

and observe that $D(\ln(N/a))| = |B_N(a)| = |B_{N/a}(1)|$ where $|\Gamma|$ denotes the cardinality of a finite set Γ. Set

$$\rho_{\min}(l) = \inf\{\rho \geq 0 : |D(\rho)| = l\} \text{ and } \rho_{\max}(l) = \sup\{\rho \geq 0 : |D(\rho)| = l\}.$$

It was shown in [8] that for any $l \geq 1$,

$$\rho_{\max}(l) > \rho_{\min}(l) \geq (l^{1/m} - 1) \ln 2. \tag{4.1}$$

Set

$$Z_{n,a}(\lambda F) = E \exp S_{N,a}(\lambda F).$$

As it was explained in [8] the distribution of $S_{N,a}(\lambda F)$ depends only on $|B_n(a)|$ (in addition to λF, of course), and so $Z_{n,a}(\lambda F)$ is determined by $|B_n(a)|$. Hence, we can set $R_l(\lambda F) = Z_{n,a}(\lambda F)$ provided $|B_n(a)| = l$. Now we can write the formula for Q obtained in [8],

$$Q(\lambda F) = r \sum_{l=1}^\infty (e^{-\rho_{\min}(l)} - e^{-\rho_{\max}(l)}) \ln R_l(\lambda F) \tag{4.2}$$

where

$$r = \prod_{k=1}^m (1 - \frac{1}{r_k}) = 1 + \sum_{k=1}^m (-1)^k \sum_{i_1 < i_2 < ... < i_k \leq m} \prod_{j=1}^k \frac{1}{r_{i_j}}.$$

The series in (4.2) converges absolutely in view of (4.1) taking into account that $\ln R_l(\lambda F) \leq lM|\lambda|$. By (2.1) and the Jensen inequality we have also that $\ln R_l(\lambda F) \geq 0$.

Now observe that for any $k \geq 1$,

$$\left|\frac{d^k R_l(\lambda F)}{d\lambda^k}\right| \leq l^k M^k R_l(\lambda F), \tag{4.3}$$

and so

$$\left|\frac{d^k \ln R_l(\lambda F)}{d\lambda^k}\right| \leq C_k l^k M^k \tag{4.4}$$

for some $C_k > 0$ depending only on k. It follows that $Q(\lambda F)$ is $\mathbb{C}^\infty$ in λ and

$$|\frac{d^k Q(\lambda F)}{d\lambda^k}| \leq \hat{C}_k \tag{4.5}$$

where

$$\hat{C}_k = C_k r M^k \sum_{l=1}^{\infty} (e^{-\rho_{\min}(l)} - e^{-\rho_{\max}(l)}) l^k$$

and the latter series converges absolutely in view of (4.1). Note that existence of the first derivative of $Q(\lambda F)$ in λ already yields the large deviations bounds (2.3) and (2.4) (see Theorem II.6.1 in [1]).

Now observe that in view of (2.1),

$$\frac{d \ln R_l(\lambda F)}{d\lambda}|_{\lambda=0} = \frac{dR_l(\lambda F)}{d\lambda}|_{\lambda=0} = 0,$$

and so

$$\frac{dQ(\lambda F)}{d\lambda}|_{\lambda=0} = 0. \tag{4.6}$$

From Theorem II.6.3 in [1] it follows that (2.1), (2.3), (2.4) and (4.6) yield already that $J(u)$ attains its infimum at the unique point 0 and it is positive when $|u| > 0$. As in Section 3 the direct argument proceeds as follows. Since $Q(0) = 0$ then (4.6) implies that $Q(\lambda F) = o(\lambda)$ for small λ, and so $|\lambda u| > Q(\lambda F)$ when $|\lambda|$ is small which together with (2.4) yields the assertion above.

Taking into account that $Q(\lambda F) \geq 0$ we see that

$$J(u) = sup_{\lambda \geq 0}(\lambda u - Q(\lambda F)) \text{ if } u \geq 0 \text{ and } J(u) = sup_{\lambda \leq 0}(\lambda u - Q(\lambda F)) \text{ if } u \leq 0. \tag{4.7}$$

Similarly to Section 3 we argue that for each $u > 0$ there exists a sequence $\lambda_n \to \lambda_0$ as $n \to \infty$ such that $J(u) = \lim_{n\to\infty}(\lambda_n u - Q(\lambda F))$ where $\lambda_0 > 0$ ($t_0 = \infty$ is possible) since by above $J(u) > 0$. Therefore, for any $\Delta > 0$,

$$J(u + \Delta) \geq \lim_{n\to\infty}(\lambda_n(u + \Delta) - Q(\lambda_n F)) = J(u) + \lambda_0 \Delta$$

which means that $J(u)$ is strictly increasing for $u \geq 0$. Similarly, $J(u)$ is strictly decreasing for ≤ 0.

Observe that by Jensen's inequality $\ln R_l(\lambda F) \geq 0$ for all $l \geq 1$, and so all terms of the series in (4.2) are nonnegative. Hence, for any $\lambda > 0$ and $\varepsilon \in (0, M_+)$,

$$Q(\lambda F) \geq K \ln R_1(\lambda F) \geq L(\varepsilon)\lambda > 0$$

where $K = r(e^{-\rho_{\min}(1)} - e^{-\rho_{\max}(1)})$ and $L(\varepsilon) = KP\{F(X_1, ..., X_\ell) \geq M_+ - \varepsilon\}(M_+ - \varepsilon)$. Then, clearly, $J(u) < \infty$ for all $u \in [0, L(\varepsilon))$. If $M_+ > 0$ then by the monotonicity property of J obtained above we conclude that there exists $L_+ > 0$ such that $J(u) < \infty$ for $u \in [0, L_+)$ while $J(u) = \infty$ for $u > L_+$. Similarly, if $M_- > 0$ then there exists $L_- > 0$ such that $J(u) < \infty$ for $u \in (-L_-, 0]$ while $J(u) = \infty$ for $u < -L_-$, completing the proof of Theorem 2.2. □

References

[1] R.S. Ellis, *Entropy, large deviations and statistical mechanics*, Springer, New York (1985). MR-0793553

[2] M. Denker and Z. Kabluchko, *An Erdös-Rényi law for mixing processes*, Probab. Math. Stat. 27 (2007), 139–149. MR-2353277

[3] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd. ed., Springer, Berlin (2010). MR-2571413

[4] P. Erdös and A. Rényi, *On a new law of large numbers*, J. Anal. Math. 23(1970), 103–111. MR-0272026

[5] Yu. Kifer, *Nonconventional limit theorems*, Probab. Th. Rel. Fields, 148 (2010), 71-106. MR-2653222

[6] Yu. Kifer, *A nonconventional strong law of large numbers and fractal dimensions of some multiple recurrence sets*, Stoch. Dynam. 12 (2012), 1150023. MR-2926580

[7] Yu. Kifer and S.R.S. Varadhan, *Nonconventional limit theorems in discrete and continuous time via martingales*, Ann. Probab. 42(2014), 649–688. MR-3178470

[8] Yu.Kifer and S.R.S Varadhan, *Nonconventional large deviations theorem*, Probab. Th. Rel. Fields, 158 (2014), 197-224. MR-3152784