ELSEVIER

Large deviations for the empirical mean of associated random variables

Carla Henriques[a],[*],[1], Paulo Eduardo Oliveira[b],[2]

[a]*Dep. Matemática, Escola Sup. Tecnologia de Viseu, Campus Politécnico de Repeses, 3504-510 Viseu, Portugal*
[b]*Dep. Matemática, Univ. Coimbra, Apartado 3008, 3001-454 Coimbra, Portugal*

Abstract

We find conditions under which the sequence of empirical means of associated random variables, $\{\overline{X}_n, n \geqslant 1\}$, satisfies the large deviation principle.
© 2007 Elsevier B.V. All rights reserved.

MSC: 60E15; 60F10

Keywords: Association; Large deviations; Stationarity

1. Introduction, assumptions and auxiliary results

Consider a sequence $\{X_n, n \geqslant 1\}$ of associated and strictly stationary real valued random variables. In this paper, we prove a large deviation principle (LDP) for the sequence of partial means.

We recall that a sequence of random variables $\{X_n, n \geqslant 1\}$ is said to be associated if for any $m \in \mathbb{N}$ and any two real-valued coordinatewise nondecreasing functions f and g it holds

$$\text{Cov}(f(X_1, \ldots, X_m), g(X_1, \ldots, X_m)) \geqslant 0,$$

whenever this covariance exists. This definition was introduced in statistics by Esary et al. (1967) who were motivated from applications in reliability theory. Independently, this notion appeared also in statistical mechanics, referred as the FKG inequalities after the work of Fortuin et al. (1971). There has been an intense research on the subject matter of association. For a brief review of the relevant literature on this subject, we suggest Roussas (1999) and Dewan and Prakasa Rao (2001), which include, besides numerous references, several applications of this concept.

[*]Corresponding author. Present address: Carla Henriques, Dep. Matemática, Escola Sup. Tecnologia de Viseu, Campus Politécnico de Repeses, 3504-510 Viseu, Portugal.

E-mail address: carlahenriq@mat.estv.ipv.pt (C. Henriques).

As mentioned before, we will find conditions under which the large deviation principle (LDP) holds for the sequence of empirical means of associated random variables. When the random variables are independent and identically distributed (i.i.d.), this is a classical result known in the literature of large deviations as the Cramér's theorem. As is well known, this result has been extended in several directions, one of which is the case of non-independence of the underlying random variables. The literature on this subject is now extensive. For an account of the relevant results on LDPs, see, for example, Dembo and Zeitouni (1998) and the references therein. We will prove a LDP assuming a hyper-geometric decrease rate on the covariances $\mathrm{Cov}(X_1, X_n)$ of associated variables. This is in accordance with analogous results obtained under mixing assumptions, where deviations from independence were conveniently controlled (see, for example, Nummelin, 1990; Bryc, 1992; Bryc and Dembo, 1996). In fact, as is well known, the covariance structure of a collection of associated random variables highly determines its approximate independence (cf. Newman, 1984). So, the referred condition on the decrease rate of the covariances is, for association, the counterpart of the hyper-geometric mixing rates assumed to establish the LDP under ϕ-mixing and α-mixing in Bryc (1992) and Bryc and Dembo (1996) (see Theorem 1 of Bryc, 1992, and Proposition 2 of Bryc and Dembo, 1996).

We now present the assumptions to be considered in this paper.

(A.1) The sequence $\{X_n, n \geqslant 1\}$ is associated and strictly stationary.

(A.2) The variables of the sequence $\{X_n, n \geqslant 1\}$ are uniformly bounded, that is, $|X_n| \leqslant M, n \geqslant 1$.

(A.3) For each $n \geqslant 1$, $\overline{X}_n$ has density function bounded by $a_1(B_1)^n$, for some $a_1 > 0$ and some $B_1 > 1$.

These assumptions are usual in associated literature, except for (A.3). Concerning this last assumption, we remark that when we consider a sequence of independent and identically distributed random variables, say $Y_1, Y_2, \ldots$, having a common bounded density with compact support, and take $X_n = c(Y_n + \cdots + Y_{n+m})$, for some fixed $m \in \mathbb{N}$ and $c \in \mathbb{R}$, then these variables verify (A.3). Notice further that this procedure is a commonly used method to generate associated sequences.

In addition, we will consider an hyper-geometric decay rate of the covariances. We state this condition in terms of the sequence

$$u(n) = \sum_{j=n+1}^{\infty} \mathrm{Cov}(X_1, X_j).$$

(H) $u(n) \leqslant a_0 \exp(-n \log^{1+a} n)$, for some $a_0 > 0$ and $a > 0$.

We notice that (H) holds true if $\mathrm{Cov}(X_1, X_n) = a_1 \exp(-n \log^{1+a} n)$, for some $a_1 > 0$ and $a > 0$. In fact, if we take $v_K(x) = \int_x^{+\infty} K(t) \, \mathrm{d}t$, with $K(t) = \exp(-t \log^{1+a} t)$, from the l'Hospital rule it easily follows that $v_K(t)/K(t) \to 0$, as $t \longrightarrow +\infty$. Now just recall that $u(n) \leqslant a_1 v_K(n)$.

The following two lemmas will be needed for the proof of the main result. First we state a result contained in Newman (1980), which generalizes the classical Hoeffding identity (see relation (2.2) in Newman, 1980).

Lemma 1.1. *Let f and g be two absolutely continuous functions. If X_1 and X_2 are random variables such that $E(f(X_1)^2) < +\infty$ and $E(g(X_2)^2) < +\infty$, then*

$$\mathrm{Cov}(f(X_1), g(X_2)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f'(x_1)g'(x_2)H_{1,2}(x_1, x_2) \, \mathrm{d}x_1 \, \mathrm{d}x_2,$$

where $H_{1,2}(x_1, x_2) = \mathrm{Cov}(I_{(x_1, +\infty)}(X_1), I_{(x_2, +\infty)}(X_2))$.

The next result follows from relation (21) in Newman (1980) and Corollary to Theorem 1 in Sadikova (1996) (see Lemma 2.6 in Roussas, 1995, for details).

Lemma 1.2 (*Roussas, 1995*). *Let X_1 and X_2 be two associated random variables having density functions bounded by B_0. Then, for all $x_1, x_2 \in \mathbb{R}$,*

$$\mathrm{Cov}(I_{(-\infty, x_1]}(X_1), I_{(-\infty, x_2]}(X_2)) \leqslant B_1 \mathrm{Cov}^{1/3}(X_1, X_2),$$

where $B_1 = 2 \max(2/\pi^2, 45B_0)$.

2. Large deviation principle

For the proof of the LDP, we follow the methodology of proof of Theorem 6.4.4 of Dembo and Zeitouni (1998), which deals with the large deviation principle of the empirical mean, under a certain mixing assumption.

For $n, m \geqslant 1$, define $\overline{X}_n^m = (1/(n-m))\sum_{i=m+1}^{n} X_i$. For the sake of simplicity, we write $\overline{X}_n$ instead of $\overline{X}_n^0$. The following two results are key tools to prove the LDP for the empirical mean $\overline{X}_n$. These results are the analogues for our framework of Lemmas 6.4.6 and 6.4.7 of Dembo and Zeitouni (1998). The proofs follow the same sort of arguments, so we will concentrate only on the technical details that are specific to associated variables.

Lemma 2.1. *Suppose that* (A.1), (A.3) *and* (H) *are satisfied. Then, for each function* $g : \mathbb{R} \longrightarrow \mathbb{R}$ *concave, continuous and bounded above, the following limit exists*

$$\Lambda_g = \lim_{n \to +\infty} \frac{1}{n} \log E(\mathrm{e}^{n\,g(\overline{X}_n)}).$$

Proof. Let $g : \mathbb{R} \longrightarrow \mathbb{R}$ be concave, continuous and bounded above. Being concave and continuous, g is also Lipschitz continuous on $[-M, M]$, that is, there exists $L > 0$ such that, for all $x, y \in [-M, M]$, $|g(x) - g(y)| \leqslant L|x - y|$. Without loss of generality, we assume that $-\infty < -B \leqslant g(x) \leqslant 0$, for all $x \in [-M, M]$.

As in the proof of Lemma 6.4.6 of Dembo and Zeitouni (1998), define $h(n) = -\log E(\mathrm{e}^{ng(\overline{X}_n)})$ and obtain

$$h(n + m) \leqslant 2lLM - \log E(\mathrm{e}^{ng(\overline{X}_n)}\mathrm{e}^{mg(\overline{X}_{n+m+l}^{n+l})}). \tag{1}$$

For each $n \in \mathbb{N}$, define $f_n(x) = \mathrm{e}^{ng(x)}$, $x \in [-M, M]$, which are Lipschitz continuous and almost everywhere differentiable with $|f_n'(x)| \leqslant nL$. Applying Lemma 1.1, we get

$$|\mathrm{Cov}(\mathrm{e}^{ng(\overline{X}_n)}, \mathrm{e}^{mg(\overline{X}_{n+m+l}^{n+l})})|$$
$$= \left| \int_{[-M,M]^2} f_n'(x)f_m'(y)\mathrm{Cov}\left(I_{(-\infty,x]}(\overline{X}_n), I_{(-\infty,y]}(\overline{X}_{n+m+l}^{n+l})\right) \mathrm{d}x\,\mathrm{d}y \right|$$
$$\leqslant nmL^2\,\mathrm{Cov}(\overline{X}_n, \overline{X}_{n+m+l}^{n+l}),$$

remembering that all the covariances above are non-negative by association. Using the stationarity assumption (A.1), we get

$$|\mathrm{Cov}(\mathrm{e}^{ng(\overline{X}_n)}, \mathrm{e}^{mg(\overline{X}_{n+m+l}^{n+l})})| \leqslant L^2 n \sum_{i=l+2}^{\infty} \mathrm{Cov}(X_1, X_i) = L^2(n+m)u(l),$$

and then

$$\frac{E(\mathrm{e}^{ng(\overline{X}_n)}\mathrm{e}^{mg(\overline{X}_{n+m+l}^{n+l})})}{E(\mathrm{e}^{ng(\overline{X}_n)})E(\mathrm{e}^{mg(\overline{X}_{n+m+l}^{n+l})})} \geqslant 1 - L^2(n+m)u(l)\mathrm{e}^{(n+m)B},$$

as $g(x) \geqslant -B$, for all $x \in [-M, M]$.

Now, define $\Theta(l, n) = 1 - L^2 nu(l)\mathrm{e}^{nB}$, $l, n \in \mathbb{N}$. As in the proof of Lemma 6.4.6 of Dembo and Zeitouni (1998), by means of the preceding inequality, (1) yields

$$h(n + m) \leqslant 2lLM + h(n) + h(m) - \log(\Theta(l, n + m) \vee 0). \tag{2}$$

By (H), we have, for each $\kappa < a$ and for each $c \in \mathbb{R}$,

$$\lim_{n \to +\infty} n\,u\left(\frac{n}{\log^{1+\kappa} n}\right)\mathrm{e}^{cn} = 0. \tag{3}$$

Let $0 < \delta < a$, where a is given in (H). From (3) it is obvious that $\Theta(n/\log^{1+\delta}n, n) = 1 - L^2 nu(n/\log^{1+\delta}n)\mathrm{e}^{nB} \longrightarrow 1$, as $n \to +\infty$. So, we may take $l = [(n+m)/\log^{1+\delta}(n+m)]$ as in the proof of Lemma 6.4.6 of Dembo and Zeitouni (1998) and proceed in the same way to conclude the proof. $\square$

In what follows we will use the notation $S_x^\delta =]x - \delta, x + \delta[$.

Lemma 2.2. *Suppose that* (A.1), (A.3), (A.4) *and* (H) *are satisfied. If* $x_1, x_2 \in \mathbb{R}$ *are such that, for each* $\delta > 0$,

$$\liminf_{n \to +\infty} \frac{1}{n} \log P\left(\overline{X}_n \in S_{x_i}^\delta\right) > -\infty, \quad i = 1, 2,$$

then

$$\inf_{\delta > 0} \liminf_{n \to +\infty} \frac{1}{n} \log \frac{P(\overline{X}_{2n} \in S_{(x_1+x_2)/2}^\delta)}{P(\overline{X}_n \in S_{x_1}^{\delta/2})P(\overline{X}_n \in S_{x_2}^{\delta/2})} \geq 0.$$

Proof. Fix $\delta > 0$. From the assumptions of the lemma, there exists $c_1 > 0$ such that, for every sufficiently l arge n,

$$P(\overline{X}_n \in S_{x_1}^{\delta/2})P(\overline{X}_n \in S_{x_2}^{\delta/2}) \geq \exp(-nc_1). \tag{4}$$

By (A.4), we may apply Lemma 1.2 to obtain

$$|P(\overline{X}_n \in S_{x_1}^{\delta/2}, \overline{X}_{2n+l}^{n+l} \in S_{x_2}^{\delta/2}) - P(\overline{X}_n \in S_{x_1}^{\delta/2})P(\overline{X}_{2n+l}^{n+l} \in S_{x_2}^{\delta/2})|$$
$$\leq 4B_n \mathrm{Cov}^{1/3}(\overline{X}_n, \overline{X}_{2n+l}^{n+l}),$$

where $B_n = 2\max(2/\pi^2, 45a_1 B_1^n)$. Therefore, by the stationarity assumption (A.1),

$$|P(\overline{X}_n \in S_{x_1}^{\delta/2}, \overline{X}_{2n+l}^{n+l} \in S_{x_2}^{\delta/2}) - P(\overline{X}_n \in S_{x_1}^{\delta/2})P(\overline{X}_n \in S_{x_2}^{\delta/2})|$$
$$\leq 4B_n \left(\frac{1}{n^2} n \sum_{i=l+1}^{\infty} \mathrm{Cov}(X_1, X_i)\right)^{1/3} = 4B_n \left(\frac{u(l)}{n}\right)^{1/3},$$

which yields

$$\frac{P(\overline{X}_n \in S_{x_1}^{\delta/2}, \overline{X}_{2n+l}^{n+l} \in S_{x_2}^{\delta/2})}{P(\overline{X}_n \in S_{x_1}^{\delta/2})P(\overline{X}_n \in S_{x_2}^{\delta/2})} \geq 1 - 4B_n \left(\frac{u(l)}{n}\right)^{1/3} \exp(c_1 n), \tag{5}$$

for each $l \in \mathbb{N}$ and n large enough, taking into account (4).

Following the arguments used in the proof of Lemma 6.4.7 in Dembo and Zeitouni (1998), we get, for $l = \delta n/(2M)$,

$$P(\overline{X}_{2n} \in S_{(x_1+x_2)/2}^\delta) \geq P(\overline{X}_n \in S_{x_1}^{\delta/2}, \overline{X}_{2n+l}^{n+l} \in S_{x_2}^{\delta/2}),$$

so that, from (5),

$$\liminf_{n \to +\infty} \frac{1}{n} \log \frac{P\left(\overline{X}_{2n} \in S_{(x_1+x_2)/2}^\delta\right)}{P(\overline{X}_n \in S_{x_1}^{\delta/2})P(\overline{X}_n \in S_{x_2}^{\delta/2})}$$
$$\geq \liminf_{n \to +\infty} \frac{1}{n} \log \left(\left\{1 - 4B_n \left(\frac{u(\delta n/(2M))}{n}\right)^{1/3} \exp(c_1 n)\right\} \vee 0\right).$$

By (3), which is valid under (H), the right-hand side above is equal to zero, from which the desired result follows. $\square$

We may now formulate the main result of this section.

Theorem 2.3. *Under* (A.1), (A.2), (A.3) *and* (H), *the sequence* $\{\overline{X}_n, n \geq 1\}$ *satisfies the large deviation principle with rate function given by*

$$\Lambda^*(x) = \sup_{t \in \mathbb{R}} \{tx - \Lambda(t)\}, \quad x \in \mathbb{R},$$

which is the Fenchel-Legendre transform of

$$\Lambda(t) = \lim_{n \to +\infty} \frac{1}{n} \log E(e^{nt\overline{X}_n}).$$

Proof. The proof goes along the same lines as that of Theorem 6.4.4 of Dembo and Zeitouni (1998), using our Lemma 2.1, together with Lemma 4.4.8 and Theorem 4.4.10 of Dembo and Zeitouni (1998), to ensure that $\{\overline{X}_n, n \geqslant 1\}$ satisfies the large deviation principle with good rate function I. The convexity of I follows in the same way as in Dembo and Zeitouni (1998), applying our Lemma 2.2 to obtain, for $x_1, x_2 \in \mathbb{R}$ such that $I(x_1) < +\infty$ and $I(x_2) < +\infty$,

$$\begin{aligned}
-I\left(\frac{x_1 + x_2}{2}\right) \geqslant &\inf_{\delta > 0} \left\{ \liminf_{n \to +\infty} \frac{1}{2n} \log\left(\frac{P(\overline{X}_{2n} \in S^{\delta}_{(x_1+x_2)/2})}{P(\overline{X}_n \in S^{\delta/2}_{x_1}) P(\overline{X}_n \in S^{\delta/2}_{x_2})} \right) \right\} \\
&+ \inf_{\delta > 0} \left\{ \liminf_{n \to +\infty} \frac{1}{2n} \log\left(P\left(\overline{X}_n \in S^{\delta/2}_{x_1}\right) \right) \right\} \\
&+ \inf_{\delta > 0} \left\{ \liminf_{n \to +\infty} \frac{1}{2n} \log(P(\overline{X}_n \in S^{\delta/2}_{x_2})) \right\} \\
\geqslant &-\frac{1}{2} I(x_1) - \frac{1}{2} I(x_2).
\end{aligned}$$

The rest of the proof proceeds exactly as in Theorem 6.4.4 of Dembo and Zeitouni (1998). □

References

Bryc, W., 1992. On large deviations for uniformly strong mixing sequences. Stoch. Processes Appl. 41, 191–202.

Bryc, W., Dembo, A., 1996. Large deviations and strong mixing. Ann. Inst. Henri Poincaré 32, 549–569.

Dembo, A., Zeitouni, O., 1998. Large Deviations Techniques and Applications, Springer-Verlag, New York.

Dewan, I., Prakasa Rao, B.L.S., 2001. Associated sequences and related inference problems, In: Shanbhag, D.N., Rao, C.R. (Eds.), Stochastic Processes: Theory and Methods. Handbook of Statistics 19, North-Holland, Amsterdam, pp. 693–731.

Esary, J.D., Proschan, F., Walkup, D.W., 1967. Association of random variables with applications. Ann. Math. Statist. 38, 1466–1474.

Fortuin, C.M., Kasteleyn, P.W., Ginibre, J., 1971. Correlation inequalities on some partially ordered sets. Commun. Math. Phys. 22, 89–103.

Newman, C.M., 1980. Normal fluctuations and the FKG inequalities. Commun. Math. Phys. 74, 119–128.

Newman, C.M., 1984. Asymptotic independence and limit theorems for positively and negatively dependent random variables. In: Tong, Y.L. (Ed.), Inequalities in Statistics and Probability. IMS Lecture Notes—Monograph Series, Vol. 5, pp. 127–140.

Nummelin, E., 1990. Large deviations for funccionals of stationary processes. Probab. Theory. Rel. Fields 86, 387–401.

Roussas, G.G., 1995. Asymptotic normality of a smooth estimate of a random field distribution function under association. Statist. Probab. Lett. 24, 77–90.

Roussas, G.G., 1999. Positive and negative dependence with some statistical applications. In: Ghosh, S. (Ed.), Asymptotics, Nonparametrics and Time Series, Statist. Textbooks Monogr. 158, Dekker, New York, pp. 757–788.

Sadikova, S.M., 1996. Two-dimensional analogies of an inequality of Esseen with applications to the central limit theorem. Theory Probab. Appl. 11, 325–335.