

## Author's Accepted Manuscript

Large deviations for the empirical mean of associated random variables

Carla Henriques, Paulo Eduardo Oliveira

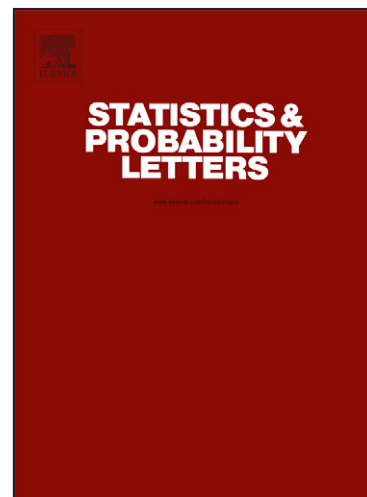
PII: S0167-7152(07)00300-8  
DOI: doi:10.1016/j.spl.2007.09.020  
Reference: STAPRO 4752

To appear in: *Statistics & Probability Letters*

Received date: 9 January 2007  
Accepted date: 7 September 2007

Cite this article as: Carla Henriques and Paulo Eduardo Oliveira, Large deviations for the empirical mean of associated random variables, *Statistics & Probability Letters* (2007), doi:10.1016/j.spl.2007.09.020

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



[www.elsevier.com/locate/stapro](http://www.elsevier.com/locate/stapro)

# Large deviations for the empirical mean of associated random variables

Carla Henriques\*<sup>†</sup>

Dep. Matemática, Escola Sup. Tecnologia de Viseu  
Campus Politécnico de Repeses, 3504-510 Viseu, Portugal

Paulo Eduardo Oliveira<sup>‡</sup>

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,  $\{\bar{X}_n, n \geq 1\}$ , satisfies the large deviation principle.

*Keywords:* association, large deviations, stationarity.

*AMS Classification:* 60E15, 60F10.

## 1 Introduction, assumptions and auxiliary results

Consider a sequence  $\{X_n, n \geq 1\}$  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  $\{X_n, n \geq 1\}$  is said to be associated if for any  $m \in \mathbb{N}$  and any two real-valued coordinatewise nondecreasing

---

\*Corresponding author. Present address: Carla Henriques, Dep. Matemática, Escola Sup. Tecnologia de Viseu, Campus Politécnico de Repeses, 3504-510 Viseu, Portugal. Email: carlahenriq@mat.estv.ipv.pt

<sup>†</sup>Author supported by Centro de Matemática da Universidade de Coimbra and Fundo Social Europeu (FSE) — PRODEP III (Action 5.3)

<sup>‡</sup>Author supported by Centro de Matemática da Universidade de Coimbra, Fundação para a Ciência e Tecnologia (FCT) and POCTI.

functions  $f$  and  $g$  it holds

$$\text{Cov}\left(f(X_1, \dots, X_m), g(X_1, \dots, X_m)\right) \geq 0,$$

whenever this covariance exists. This definition was introduced in statistics by Esary, Proschan and Walkup (1967) who were motivated from applications in reliability theory. Independently, this notion appeared also in statistical mechanics, referred as the KFG inequalities after the work of Fortuin, Kasteleyn and Ginibre (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.

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  $\text{Cov}(X_1, X_n)$  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, and 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  $\phi$ -mixing and  $\alpha$ -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.

- (A1) The sequence  $\{X_n, n \geq 1\}$  is associated and strictly stationary.
- (A2) The variables of the sequence  $\{X_n, n \geq 1\}$  are uniformly bounded, that is,  $|X_n| \leq M, n \geq 1$ .
- (A3) For each  $n \geq 1$ ,  $\bar{X}_n$  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 (A3). Concerning this last assumption, we remark that when we consider a sequence of independent and identically distributed random variables, say  $Y_1, Y_2, \dots$ , having a common bounded density with compact support, and take  $X_n = c(Y_n + \dots + Y_{n+m})$ , for some fixed  $m \in \mathbb{N}$  and  $c \in \mathbb{R}$ , then these variables verify (A3). 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

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

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

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

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  $\mathbb{E}(f(X_1)^2) < +\infty$  and  $\mathbb{E}(g(X_2)^2) < +\infty$ , then*

$$\text{Cov}(f(X_1), g(X_2)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f'(x_1)g'(x_2)H_{1,2}(x_1, x_2) dx_1 dx_2,$$

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

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  $x_1, x_2 \in \mathbb{R}$ ,*

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

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  $n, m \geq 1$ , define  $\bar{X}_n^m = \frac{1}{n-m} \sum_{i=m+1}^n X_i$ . For the sake of simplicity, we write  $\bar{X}_n$  instead of  $\bar{X}_n^0$ . The following two results are key tools to prove the LDP for the empirical mean  $\bar{X}_n$ . 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 (A1), (A3) and (H) are satisfied. Then, for each function  $g : \mathbb{R} \rightarrow \mathbb{R}$  concave, continuous and bounded above, the following limit exists*

$$\Lambda_g = \lim_{n \rightarrow +\infty} \frac{1}{n} \log \mathbb{E} \left( e^{ng(\bar{X}_n)} \right).$$

*Proof:* Let  $g : \mathbb{R} \rightarrow \mathbb{R}$  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  $x, y \in [-M, M]$ ,  $|g(x) - g(y)| \leq L|x - y|$ . Without loss of generality, we assume that  $-\infty < -B \leq g(x) \leq 0$ , for all  $x \in [-M, M]$ .

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

$$h(n+m) \leq 2LLM - \log \mathbb{E} \left( e^{ng(\bar{X}_n)} e^{mg(\bar{X}_{n+m+l}^{n+l})} \right). \quad (1)$$

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

$$\begin{aligned} & \left| \text{Cov} \left( e^{ng(\bar{X}_n)}, e^{mg(\bar{X}_{n+m+l}^{n+l})} \right) \right| \\ &= \left| \int_{[-M, M]^2} f'_n(x) f'_m(y) \text{Cov} \left( I_{(-\infty, x]}(\bar{X}_n), I_{(-\infty, y]}(\bar{X}_{n+m+l}^{n+l}) \right) dx dy \right| \\ &\leq nmL^2 \text{Cov} \left( \bar{X}_n, \bar{X}_{n+m+l}^{n+l} \right), \end{aligned}$$

remembering that all the covariances above are non-negative by association. Using the stationarity assumption (A1), we get

$$\left| \text{Cov} \left( e^{ng(\bar{X}_n)}, e^{mg(\bar{X}_{n+m+l}^{n+l})} \right) \right| \leq L^2 n \sum_{i=l+2}^{\infty} \text{Cov}(X_1, X_i) = L^2(n+m)u(l),$$

and then

$$\frac{\mathbb{E}\left(e^{ng(\bar{X}_n)}e^{mg(\bar{X}_{n+m+l}^{n+l})}\right)}{\mathbb{E}\left(e^{ng(\bar{X}_n)}\right)\mathbb{E}\left(e^{mg(\bar{X}_{n+m+l}^{n+l})}\right)} \geq 1 - L^2(n+m)u(l)e^{(n+m)B},$$

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

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

$$h(n+m) \leq 2LLM + h(n) + h(m) - \log(\Theta(l, n+m) \vee 0). \quad (2)$$

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

$$\lim_{n \rightarrow +\infty} nu \left( \frac{n}{\log^{1+\kappa} n} \right) e^{cn} = 0. \quad (3)$$

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

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

**Lemma 2.2** *Suppose that (A1), (A3), (A4) and (H) are satisfied. If  $x_1, x_2 \in \mathbb{R}$  are such that, for each  $\delta > 0$ ,*

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

then

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

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

$$\mathbb{P}\left(\bar{X}_n \in S_{x_1}^{\delta/2}\right)\mathbb{P}\left(\bar{X}_n \in S_{x_2}^{\delta/2}\right) \geq \exp(-nc_1). \quad (4)$$

By (A4), we may apply Lemma 1.2 to obtain

$$\begin{aligned} & \left| \mathbb{P}\left(\bar{X}_n \in S_{x_1}^{\delta/2}, \bar{X}_{2n+l}^{n+l} \in S_{x_2}^{\delta/2}\right) - \mathbb{P}\left(\bar{X}_n \in S_{x_1}^{\delta/2}\right)\mathbb{P}\left(\bar{X}_{2n+l}^{n+l} \in S_{x_2}^{\delta/2}\right) \right| \\ & \leq 4B_n \text{Cov}^{1/3}\left(\bar{X}_n, \bar{X}_{2n+l}^{n+l}\right), \end{aligned}$$

where  $B_n = 2 \max(2/\pi^2, 45a_1B_1^n)$ . Therefore, by the stationarity assumption (A1),

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

which yields

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

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  $l = \frac{\delta n}{2M}$ ,

$$\mathbb{P}\left(\bar{X}_{2n} \in S_{\frac{x_1+x_2}{2}}^{\delta}\right) \geq \mathbb{P}\left(\bar{X}_n \in S_{x_1}^{\delta/2}, \bar{X}_{2n+l}^{n+l} \in S_{x_2}^{\delta/2}\right),$$

so that, from (5),

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

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

We may now formulate the main result of this section.

**Theorem 2.3** *Under (A1), (A2), (A3) and (H), the sequence  $\{\bar{X}_n, n \geq 1\}$  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

$$\Lambda(t) = \lim_{n \rightarrow +\infty} \frac{1}{n} \log \mathbb{E} \left( e^{n t \bar{X}_n} \right).$$

*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  $\{\bar{X}_n, n \geq 1\}$  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  $I(x_2) < +\infty$ ,

$$\begin{aligned}
 -I\left(\frac{x_1 + x_2}{2}\right) &\geq \inf_{\delta > 0} \left\{ \liminf_{n \rightarrow +\infty} \frac{1}{2n} \log \left( \frac{\mathbb{P}\left(\bar{X}_{2n} \in S_{\frac{x_1+x_2}{2}}^\delta\right)}{\mathbb{P}\left(\bar{X}_n \in S_{x_1}^{\delta/2}\right) \mathbb{P}\left(\bar{X}_n \in S_{x_2}^{\delta/2}\right)} \right) \right\} \\
 &\quad + \inf_{\delta > 0} \left\{ \liminf_{n \rightarrow +\infty} \frac{1}{2n} \log \left( \mathbb{P}\left(\bar{X}_n \in S_{x_1}^{\delta/2}\right) \right) \right\} \\
 &\quad \quad + \inf_{\delta > 0} \left\{ \liminf_{n \rightarrow +\infty} \frac{1}{2n} \log \left( \mathbb{P}\left(\bar{X}_n \in S_{x_2}^{\delta/2}\right) \right) \right\} \\
 &\geq -\frac{1}{2}I(x_1) - \frac{1}{2}I(x_2).
 \end{aligned}$$

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



## References

- [1] Bryc, W. (1992), On large deviations for uniformly strong mixing sequences, *Stochastic Processes Appl.* 41, 191–202.
- [2] Bryc, W. and Dembo, A. (1996), Large deviations and strong mixing, *Ann. Inst. Henri Poincaré* 32, 549–569.
- [3] Dembo, A. and Zeitouni, O. (1998), *Large Deviations Techniques and Applications*, Springer-Verlag, New York.
- [4] Dewan, I. and Prakasa Rao, B.L.S. (2001), Associated sequences and related inference problems, in: D.N. Shanbhag and C.R. Rao, eds. *Stochastic Processes: Theory and Methods*, Handbook of Statistics 19, North-Holland, Amsterdam, 693–731.
- [5] Esary, J.D., Proschan, F. and Walkup, D.W. (1967), Association of random variables, with applications, *Ann. Math. Statist.* 38, 1466–1474.
- [6] Fortuin, C. M., P. W. Kasteleyn and J. Ginibre (1971), Correlation inequalities on some partially ordered sets, *Communications in Mathematical Physics* 22, 89–103.
- [7] Newman, C.M. (1980), Normal fluctuations and the FKG inequalities, *Comm. Math Phys.* 74, 119–128.
- [8] Newman, C.M. (1984), Asymptotic independence and limit theorems for positively and negatively dependent random variables, in: Y.L. Tong, ed. *Inequalities in Statistics and Probability*, IMS Lecture Notes - Monograph Series, Vol. 5, 127–140.
- [9] Nummelin, E. (1990), Large deviations for functionals of stationary processes, *Probab. Th. Rel. Fields* 86, 387–401.
- [10] Roussas, G.G. (1995), Asymptotic normality of a smooth estimate of a random field distribution function under association, *Statist. Probab. Letters* 24, 77–90.
- [11] Roussas, G.G. (1999), Positive and negative dependence with some statistical applications, in: S. Ghosh, ed. *Asymptotics, Nonparametrics and Time Series*, Statist. Textbooks Monogr. 158, Dekker, New York, 757–788.
- [12] 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.