Large deviations for L-statistics

Hélène Boistard *

Summary: The purpose of this paper is to establish a functional Large Deviations Principle (LDP) for L-statistics under some new tail conditions. The method is based on Sanov's theorem and on basic tools of large deviations theory. Our study includes a full treatment of the case of the uniform law and an example in which the rate function can be calculated very precisely. We extend our result by a LDP for normalized L-statistics. The case of the exponential distribution, which is not in the scope of the previous conditions, is completely treated with another method. We provide a functional LDP obtained via Gärtner-Ellis Theorem.

1 Introduction

In this paper, we will consider L-statistics. That means that we will study the following random variable

$$A_n = \sum_{i=1}^{n} a_{n,i} X_{(i)}. \tag{1.1}$$

All along the article, $(X_i)_{i=1}^n$ is an i.i.d. sample with distribution function F, $(X_{(i)})_{i=1}^n$ is the associated order statistics, and $(a_{n,i})_{i=1}^n$ are some coefficients. It is often assumed that these coefficients are closely related to some given function $a : [0,1] \to \mathbb{R}^k$ in the following way:

$$a_{n,i} = \frac{1}{n} a\left(\frac{i}{n}\right).$$

Some examples of L-statistics include the α-trimmed mean:

$$\frac{1}{n-2[\alpha n]} \sum_{i=[\alpha n]+1}^{n-[\alpha n]} X_{(i)}, \tag{1.2}$$

or Gini's mean difference

$$\frac{1}{C_n^2} \sum_{i<j} |X_i - X_j| = \frac{1}{C_n^2} \sum_{i=1}^{n} (-n+2i-1) X_{(i)}, \tag{1.3}$$

which estimates the dispersion parameter $E\left(|X_1 - X_2|\right)$ (see Example 5.3 in Stigler [17]).

Many asymptotic results have been obtained for L-statistics. The results in the literature apply to more general L-statistics, namely:

$$A_n = \sum_{i=1}^{n} a_{n,i} b(X_{(i)}),$$

*The author has been partially supported by the Spanish Ministerio de Ciencia y Tecnología and FEDER, grant BFM2005-04430-C02-01 and 02 and by the Consejería de Educación y Cultura de la Junta de Castilla y León, grant PAPIJCL VA102/06.
A.M.S. 2000 subject classification: Primary: 60F10, secondary: 62G20
Key words and phrases: Large Deviations, L-statistics

where b is some fixed function. In general, the authors formulate conditions either on the scores $a_{n,i}$ or on the function b. We focus here on the case of b being the identity function.

A strong law of large numbers is obtained in Wellner [21], [22] and in van Zwet [19] using the Glivenko-Cantelli theorem. In Stigler [17], a central limit theorem is obtained via Hájek projections. Another way to obtain a CLT is proposed in Helmers [10], with Berry-Esseen-type bounds. The tool used there is an approximation by U-statistics. This is done also in Vandemaele [20]. A very complete version of the CLT with necessary and sufficient conditions is proved in Mason and Shorack [12], via empirical processes theory. For weaker conditions on the function b, a CLT and a LIL theorem can be found in Li et al. [11].

We refer to Shorack and Wellner [16] for an exposition of the strong LLN, LIL and CLT in a unified way. For a very clear proof of the CLT, we refer to van der Vaart [18]. Two approaches are treated: the method of Stigler [17] and the Δ-method, using the theory of empirical processes.

For LDP-type results, we cite three relevant references: Groeneboom, Oosterhoof and Ruymgaart [8] (Section 6), Groeneboom [7] (Section 1.6) and Groeneboom and Shorack [9] (Section 3). These articles give results for L-statistics written for some function $a \in L_1(0,1)$ as:

$$A_n = \sum_{i=1}^{n} X_{(i)} \int_{(i-1)/n}^{i/n} a(t)dt = \int_0^1 a(t) F_n^{-1}(t)dt, \tag{1.4}$$

where F_n^{-1} is the empirical quantile function defined as: $F_n^{-1} : t \mapsto X_{(i)}$ for $t \in (\frac{i-1}{n}, \frac{i}{n}]$. There, A_n is seen as a functional of the empirical distribution function F_n. Hence, a natural method is to use the LDP for the empirical measure given by Sanov's theorem and the contraction principle. However, Sanov's theorem cannot be used directly. The topology on the space of measures has to be strengthened into the τ-topology. Although the weak topology is generated by the continuous bounded functions, the τ-topology is generated by the measurable bounded functions (see for instance Dembo and Zeitouni [5], p263). Some hypothesis on the weight function a and the tails of the underlying distribution are introduced. In the first two references, a is asked to have bounded support. A more general result is available in Groeneboom and Shorack [9], Corollary 3. We reproduce it in Theorem 1.1 below. Before stating that result, we recall the definition for the Kullback information of some distribution function G with respect to F: it is given by

$$K(G,F) = \begin{cases} \int_{\mathbb{R}} \log \frac{dG}{dF} dG & \text{if } G \ll F \\ +\infty & \text{else.} \end{cases}$$

We define the rate function

$$I_0(C) = \inf_{G^{-1} \text{ quantile function: } \int aG^{-1} = C} K(G,F).$$

Theorem 1.1 *(Groeneboom and Shorack 1981) The weight function a is supposed to be an L_1 function satisfying for each c:*

$$\int_{1/2}^1 \left| a(t) F^{-1}\left(1 - e^{-\frac{c}{1-t}}\right)\right| dt < \infty, \textit{ and } \int_0^{1/2} \left| a(t) F^{-1}\left(1 - e^{-\frac{c}{t}}\right)\right| dt < \infty, \tag{1.5}$$

$$a \geq 0 \textit{ on an interval } (\gamma,\delta) \subset (0,1) \textit{ and } \int_\gamma^\delta a(t)dt > 0. \tag{1.6}$$

Then A_n (defined in (1.4)) satisfies for all $r \in \mathbb{R}$:

$$\lim_{n\to\infty} \frac{1}{n} \log P(A_n \leq r) = -\inf\{I_0(C) : C \leq r\}.$$

We can observe that this is not a full LDP, since the rate function is only obtained for sets which are half-lines (remark that the lower half-lines can be treated using the function $-a$). Nevertheless, under further conditions, the full LDP can be deduced. We give here some clues to derive this LDP. This proof follows the same principles as the proof of Cramér's theorem (e.g., Teorema 2.2.3 en Dembo y Zeitouni [5]). The lower bound does not need further hypothesis. Our method to prove the upper bound does require an additional condition, which is the following:

$$I_0 \text{ is decreasing on the interval } \left(-\infty, \int aF^{-1}\right) \text{ and increasing on the interval } \left(\int aF^{-1}, +\infty\right). \tag{1.7}$$

Theorem 1.2 *Let us suppose that the hypothesis of Theorem 1.1 are satisfied. Then,*
(i) for every open set $O \subset \mathbb{R}$*,*

$$\liminf \frac{1}{n} \log P(A_n \in O) \geq -\inf\{I_0(C) : C \in O\}.$$

(ii) If moreover, the monotony condition (1.7) is satisfied, then for every closed subset $U \subset \mathbb{R}$*,*

$$\limsup \frac{1}{n} \log P(A_n \in U) \leq -\inf\{I_0(C) : C \in U\}.$$

Proof: (i) Let O be some open subset. We prove that for all $x \in O$,

$$\liminf \frac{1}{n} \log P(A_n \in O) \geq -I_0(x). \tag{1.8}$$

Let $x \in O$. We assume that $x \leq \int aF^{-1}$. A similar proof can be performed when $x \geq \int aF^{-1}$. If $I_0(x) = +\infty$, then (1.8) is obvious. Hence, let us suppose that $I_0(x) < \infty$. Let $[x - \epsilon, x + \epsilon]$ be a closed neighbourhood of x included in O. By Theorem 1.1, for n large enough, $P(A_n \leq x + \epsilon) \geq P(A_n \leq x) \neq 0$.

$$\begin{aligned}\frac{1}{n} \log P(A_n \in O) &\geq \frac{1}{n} \log P\left(A_n \in [x - \epsilon, x + \epsilon]\right) \\ &\geq \frac{1}{n} \log\left(P(A_n \leq x + \epsilon) - P(A_n \leq x - \epsilon)\right) \\ &= \frac{1}{n} \log P(A_n \leq x + \epsilon) + \frac{1}{n} \log\left(1 - \frac{P(A_n \leq x - \epsilon)}{P(A_n \leq x + \epsilon)}\right)\end{aligned}$$

By Theorem 1.1,

$$\frac{1}{n} \log P(A_n \leq x + \epsilon) \to -\inf\{I_0(C) : C \leq x + \epsilon\} \geq -I_0(x).$$

Similarly, denoting $I_0(V) = \inf\{I_0(x) : x \in V\}$ for all subset V:

$$\begin{aligned}1 - \frac{P(A_n \leq x - \epsilon)}{P(A_n \leq x + \epsilon)} &= 1 - \frac{e^{n \frac{1}{n} \log P(A_n \leq x - \epsilon)}}{e^{n \frac{1}{n} \log P(A_n \leq x + \epsilon)}} \\ &= 1 - \exp\left(-nI_0\left((-\infty, x - \epsilon]\right)\left(1 - \frac{I_0\left((-\infty, x + \epsilon]\right)}{I_0\left((-\infty, x - \epsilon]\right)} + o(1)\right)\right) \\ &\to 1, \text{ when } n \to \infty.\end{aligned}$$

Therefore,

$$\liminf \frac{1}{n} \log P(A_n \in O) \geq -I_0(x).$$

(ii) Let $U \neq \emptyset$ be some closed subset and (x_-, x_+) the biggest open interval included in U^c and containing $\int aF^{-1}$. As U is closed and non-empty, either x_- or x_+ is in U and U is included in $(-\infty, x_-] \cup [x_+, \infty)$. For all $\epsilon > 0$, for n large enough,

$$\begin{aligned}P(A_n \in U) &\leq P(A_n \leq x_-) + P(A_n \geq x_+) \\ &\leq 2e^{-nI_0(U) + n\epsilon}\end{aligned} \tag{1.9}$$

Indeed, by Theorem 1.1, for n large enough, using Condition (1.7):

$$\begin{aligned}\frac{1}{n} \log P(A_n \leq x_-) &\leq -\inf\{I_0(C) : C \leq x_-\} + \epsilon \\ &= -I_0(x_-) + \epsilon \text{ by (1.7)}\end{aligned}$$

The same happens with x_+, which leads to:

$$\begin{aligned}P(A_n \leq x_-) &\leq e^{-n(I_0(x_-) - \epsilon)} \text{ and} \\ P(A_n \geq x_+) &\leq e^{-n(I_0(x_+) - \epsilon)}\end{aligned}$$

But when x_- is finite, it is an element of U and: $I_0(x_-) \geq I_0(U)$. The same occurs with x_+. When they are not finite, they do not appear in the bound (1.9). This ends up with the proof of (1.9). We deduce that: for all $\epsilon > 0$,

$$\frac{1}{n}\log P(A_n \in U) \leq \frac{1}{n}\log(2) - I_0(U) + \epsilon.$$

This implies:

$$\begin{aligned}\limsup \frac{1}{n}\log P(A_n \in U) &\leq -I_0(U) + \epsilon \\ &\leq -I_0(U), \text{ when } \epsilon \text{ tends to } 0.\end{aligned}$$

□

In this paper, we present an analogous result for L-statistics which can be written as in (1.4), under another set of conditions for the function a and the tails of the underlying distribution. The first step of our method is the obtention of a LDP result for the empirical measure in the space of probability measures with finite second moment. That space can be equipped with the L_2-Wasserstein distance. In fact, we formulate the LDP for the empirical quantile function F_n^{-1} seen as an element of $L_2(0,1)$. The conditions on F are formulated in (i) or (ii) in Theorem 2.1 below. For $a \in L_2(0,1)$, the L-statistic (1.4) is a continuous functional of F_n^{-1} for that topology. Then, a simple application of the contraction principle allows to derive a LDP result for L-statistics as in (1.4).

The main apportation of this method is that it allows a completely functional treatment. On the other hand, the underlying distribution is asked to have lighter tails.

We also relax the condition on a in the case of the exponential distribution. Indeed, Theorem 1.1 does not allow to treat the case of fonctions a which do not tend to 0 at 1 (see the remarks at the beginnig of Section 3 for more details). In Theorem 3.3 below, we obtain a functional LDP which allows to treat L-statistics for underlying exponential distribution, for the class of continuous functions.

This paper is organized as follows. Section 2 is devoted to our LDP result for L-statistics under tail conditions on the underlying distribution and on the function a. It contains some examples and an extension to the problem of large deviations for the normalized empirical quantile function, with an application to normalized L-statistics. Section 3 is dedicated to the case of the exponential distribution. Further, to be self-contained, we write an appendix where we recall useful facts on large deviations.

We introduce now some definitions and notations which will be used in the rest of the paper.

We will call $\mathcal{P}(\mathbb{R})$ the set of all probability measures on $\mathbb{R}$ equipped with the topology of convergence in distribution. $\mathcal{M}(\mathbb{R})$ will denote the set of all quantile functions of probability measures on $\mathbb{R}$. It is equipped with the topology induced by convergence in distribution. Hence, there is a topological isomorphism between $\mathcal{P}(\mathbb{R})$ and $\mathcal{M}(\mathbb{R})$.

$\mathcal{P}_2(\mathbb{R}) \subset \mathcal{P}(\mathbb{R})$ will denote the space of probability measures on $\mathbb{R}$ with a finite second moment. It is equipped with the L_2-Wasserstein distance. For P, $Q \in \mathcal{P}_2(\mathbb{R})$, this distance is defined as

$$\mathcal{W}(P,Q) = \inf\left\{\left(E(X-Y)^2\right)^{1/2}, \mathcal{L}(X) = P, \mathcal{L}(Y) = Q\right\},$$

where $\mathcal{L}(X)$ denotes the distribution of X. For a distribution function G, G^{-1} will always denote the corresponding quantile function. It is defined as the generalized inverse of G as follows:

$$G^{-1}(t) = \inf\{x : G(x) \geq t\}, t \in (0,1).$$

It is a left-continuous increasing function with range equal to the support of G.

A useful property is the expression of $\mathcal{W}(P,Q)$ in terms of the quantile functions G^{-1} and H^{-1} of P and Q:

$$\mathcal{W}(P,Q) = \left(\int (G^{-1} - H^{-1})^2\right)^{1/2}. \tag{1.10}$$

We refer to del Barrio et al. [4], Section 3.3 and the references therein for more details on the Wasserstein distance.

Naturally, $\mathcal{M}_2(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$ is defined as the set of quantile functions of probability measures on $\mathbb{R}$ with a finite second moment. $\mathcal{M}_2(\mathbb{R}) \subset L_2(0,1)$ and can be equipped with the topology inherited from the Hilbert space $L_2(0,1)$. With the help of (1.10), we see that there is a topological isomorphism between $\mathcal{P}_2(\mathbb{R})$ and $\mathcal{M}_2(\mathbb{R})$.

2 LDP for L-statistics under tail condition

In this section, we state a functional LDP for the empirical quantile function under strong and relaxed tail conditions. L-statistics can be obtained via a continuous transformation of the empirical quantile function. So, we obtain a LDP for L-statistics. The main results are presented in Subsection 2.1. Some examples of L-statistics are treated in Subsection 2.2. As a corollary of the LDP for the empirical quantile function, we obtain a LDP for a normalized empirical quantile function in Subsection 2.3. This can be applied to some normalized L-statistics. The technical proofs are postponed to Subsection 2.4.

2.1 Functional LDP for the empirical quantile function in $L_2(0,1)$

Our method to obtain a functional LDP for the quantile function is based on Sanov's theorem for the empirical measure (Theorem 6.2.10 in Dembo and Zeitouni [5]). The idea is to reinforce the topology of $\mathcal{P}(\mathbb{R})$ without losing the LDP. An appropriate topology is the one induced by the Wasserstein distance on the subspace $\mathcal{P}_2(\mathbb{R})$.

The strong tail condition we will require is the following: there exists $\varphi : \mathbb{R} \to \mathbb{R}$ with $\varphi(x) \to +\infty$ as $|x| \to \infty$, and $t > 0$ such that

$$E\left(e^{tX_1^2\varphi(X_1)}\right) < +\infty. \tag{2.1}$$

This condition is trivially satisfied for distributions with a bounded support, so that a truncation argument allows us to derive also a LDP under a relaxed tail condition (Condition (ii) in Theorem 2.1).

We now state the main theorem.

Theorem 2.1 *We assume one of the following conditions:*
(i) (2.1) is satisfied.
(ii) The r.v. X_1^2 has a Laplace transform defined on $\mathbb{R}$.
Then the empirical quantile function F_n^{-1} satisfies a LDP in $\mathcal{M}_2(\mathbb{R})$ with a good rate function

$$\begin{aligned} I_1 : \mathcal{M}_2(\mathbb{R}) &\to \mathbb{R} \\ G^{-1} &\to I_1(G^{-1}) = K(G,F) && \textit{under Condition (i),} \\ I_2 : \mathcal{M}_2(\mathbb{R}) &\to \mathbb{R} \\ G^{-1} &\to I_2(G^{-1}) = \sup_{\delta>0} \liminf_{T\to\infty} \inf_{\|H^{-1}-G^{-1}\|_2<\delta} K(H,F^T) && \textit{under Condition (ii),} \end{aligned}$$

where F^T is the distribution function of the truncated r. v.'s

$$X_i^T = -T1_{X_i<-T} + X_i 1_{|X_i|\le T} + T1_{X_i>T}.$$

The proof of this theorem can be found in Subsection 2.4.

Remark 2.2 Under Condition 2.1, the restriction to $\mathcal{M}_2(\mathbb{R})$ is not restrictive at all. Indeed, Condition (2.1) implies that if a probability measure has finite Kullback information with respect to F, then it has a finite second moment. This claim is true even for a weaker hypothesis than (2.1): suppose that there exists $t > 0$ such that

$$E\left(e^{tX_1^2}\right) < +\infty. \tag{2.2}$$

Let G be such that $K(G,F) < \infty$, then $G^{-1} \in \mathcal{M}_2(\mathbb{R})$. Indeed: recall the following duality inequality

$$ab \le a\log a + e^b \text{ for } a, b > 0.$$

Apply this to the likelihood $a(x) = \frac{dG}{dF}(x)$ and $b(x) = tx^2$ with t such that $E\left(e^{tX_1^2}\right) < \infty$. By an integration with respect to dF, it follows that $\int x^2 dG(x) < \infty$.

The last theorem allows to obtain by contraction a LDP for L-statistics with coefficients of type $a_{n,i} = a(i/n)/n$ (see Corollary 2.4 below). Assume that the support of F is included in $\mathbb{R}_+$. We now state a functional LDP for the following random measure on $[0,1]$:

$$\nu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\frac{i}{n}} X_{(i)}. \tag{2.3}$$

Here, $\delta_{i/n}$ denotes the Dirac measure at $\frac{i}{n}$. The measure ν_n gives weight $\frac{1}{n}X_{(i)}$ to the point $\frac{i}{n}$. It is seen as an element of $\mathcal{P}_+([0,1])$ which is the set of all positive measures on $[0,1]$ with the weak topology. As in Araujo and Giné [1], we define the Lipschitz Bounded metric that metrizes this topology:

$$d_{LB}(\mu,\nu) = \sup_{f\in\mathcal{F}_{LB}} \left| \int_{[0,1]} f d\nu - \int_{[0,1]} f d\mu \right|,$$

where $\mathcal{F}_{LB}$ is the class of Lipschitz continuous functions $f : [0,1] \to \mathbb{R}$, with Lipschitz constant at most 1 and uniform bound 1.

The following analogue of Theorem 2.1 holds for the random measure ν_n.

Theorem 2.3 *We assume that the support of F is included in $\mathbb{R}_+$ and that one of the following conditions holds:*
(i) (2.1) is satisfied.
(ii) The r.v. X_1^2 has a Laplace transform defined on $\mathbb{R}$.
Then the random measure ν_n satisfies a LDP on $\mathcal{P}_+([0,1])$ with good rate function

$$\begin{aligned} \tilde{I}_1 : \mathcal{P}_+(\mathbb{R}) &\to \mathbb{R} \\ P &\to \tilde{I}_1(P) = \begin{cases} I_1(G^{-1}) & \text{when } P \ll \lambda \text{ and } G^{-1} \text{ is a quantile function s.t. } \frac{dP}{d\lambda} = G^{-1}, \\ +\infty & \text{else,} \end{cases} \quad \text{under Condition (i),} \\ \tilde{I}_2 : \mathcal{P}_+(\mathbb{R}) &\to \mathbb{R} \\ P &\to \tilde{I}_2(P) = \begin{cases} I_2(G^{-1}) & \text{when } P \ll \lambda \text{ and } G^{-1} \text{ is a quantile function such that } \frac{dP}{d\lambda} = G^{-1}, \\ +\infty & \text{else,} \end{cases} \quad \text{under Condition (ii).} \end{aligned}$$

The proof is postponed to Subsection 2.4.

As a corollary of Theorems 2.1 and 2.3, we now state a LDP for L-statistics under tail conditions.

Corollary 2.4 *Let a be some function on $(0,1)$.*
(i) Under Condition (2.1), for $a \in L_2(0,1)$ (resp. for a continuous on $[0,1]$), the L-statistic $\sum_{i=1}^n X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} a(t)dt$ (resp. the L-statistic $\frac{1}{n}\sum_{i=1}^n a\left(\frac{i}{n}\right) X_{(i)}$) satisfies a LDP on $\mathbb{R}$ with good rate function $I_0(C)$.
(ii) If the r.v. X_1^2 has a Laplace transform defined on $\mathbb{R}$, for $a \in L_2(0,1)$ (resp. for a continuous on $[0,1]$), then the L-statistic $\sum_{i=1}^n X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} a(t)dt$ (resp. the L-statistic $\frac{1}{n}\sum_{i=1}^n a\left(\frac{i}{n}\right) X_{(i)}$) satisfies a LDP on $\mathbb{R}$ with good rate function

$$I_3(C) = \inf_{G^{-1} \text{ quantile function} : \int aG^{-1}=C} I_2(G^{-1}). \tag{2.4}$$

Proof: It is a direct application of the contraction principle. Let us first suppose that $a \in L_2(0,1)$. The map

$$\begin{aligned} \mathcal{M}_2(\mathbb{R}) \subset L_2(0,1) &\to \mathbb{R} \\ G^{-1} &\mapsto \int aG^{-1} \\ F_n^{-1} = \sum_{i=1}^n 1_{(\frac{i-1}{n},\frac{i}{n}]} X_{(i)} &\mapsto \sum_{i=1}^n X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} a(t)dt. \end{aligned} \tag{2.5}$$

is continuous.

Let us now suppose that a is continuous on $[0,1]$. The map

$$\begin{aligned} \mathcal{P}_+(\mathbb{R}) &\to \mathbb{R} \\ P &\mapsto \int adP \\ \nu_n = \frac{1}{n}\sum_{i=1}^n \delta_{i/n} X_{(i)} &\mapsto \frac{1}{n}\sum_{i=1}^n a(\frac{i}{n}) X_{(i)} \end{aligned}$$

is continuous.

□

2.2 Examples

We first recall the examples presented in the introduction and show how to deal with them.

Example 2.5 *The α-trimmed mean.*

Let $\alpha < 1/2$. Consider the following function a defined on $[0,1]$:

$$a(t) = \begin{cases} \frac{1}{1-2\alpha} & \text{for } t \in [\alpha, 1-\alpha] \\ 0 & \text{else.} \end{cases}$$

Let us denote by S_n^α the α-trimmed mean defined by (1.2) and by A_n^α the L-statistic $\int_0^1 a(t) F_n^{-1}(t)dt$. A straightforward calculus shows that

$$S_n^\alpha - A_n^\alpha = \begin{cases} \frac{1}{n-2\alpha n} X_{(\alpha n)} & \text{when } \alpha n \text{ is an integer,} \\ 0 & \text{else.} \end{cases}$$

Therefore, it is easy to show that S_n^α and A_n^α are exponentially equivalent under suitable conditions. The point is that the ratio by n of the order statistic near the αth quantile is exponentially equivalent to 0. For instance, it is trivial when the support of F is bounded. Indeed, suppose that it is included in $[-M, M]$. For any fixed δ,

$$\begin{aligned} P(|S_n^\alpha - A_n^\alpha| \geq \delta) &\leq P\left(\left|X_{([\alpha n])}\right| \geq \delta\left(n - 2[\alpha n]\right)\right) \\ &\leq P\left(M \geq \delta(n - 2[\alpha n])\right) \\ &= 0 \end{aligned}$$

for n large enough. This proves the exponential equivalence between S_n^α and A_n^α. Therefore, the α-trimmed mean satisfies the same LDP as A_n^α. The rate function can be calculated with Corollary 2.4.

Example 2.6 *Gini's difference mean.*

Notice that this statistic, given in (1.3), can be written as

$$\frac{4n}{n-1}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{i}{n} - \frac{1}{2} - \frac{1}{2n}\right) X_{(i)} = \frac{4}{n}\sum_{i=1}^{n} a\left(\frac{i}{n}\right) X_{(i)} + R_n,$$

where $a(t) = t - \frac{1}{2}$ and R_n satisfies, under (2.1),

$$\frac{1}{n}\log P\left(|R_n| > \delta\right) \to -\infty, n \to \infty, \forall \delta > 0.$$

Hence, it is equivalent with a L-statistic in the scope of Corollary 2.4.

Example 2.7 *Centered score function a and uniform distribution.*

This is a class of examples of L-statistics for which the rate function can be expressed as the result of an optimization problem. In a particular case, this optimization problem can be solved and the rate function can be calculated with numerical tools. Suppose that F is the uniform law on $[0,1]$. Let $a : [0,1] \to \mathbb{R}^k$ be a square integrable function such that $E[a(X_1)] = 0$. Define

$$A = (A_1, \ldots, A_k) : [0,1] \to \mathbb{R}^k \tag{2.6}$$

$$t \mapsto \int_t^1 a(s)ds. \tag{2.7}$$

By Corollary 2.4, the L-statistic $\int_0^1 a(t) F_n^{-1}(t)dt$ satisfies a LDP with good rate function expressed in terms of Kullback information. However, this expression is not explicit. The following theorem presents another formulation for the rate function expressed as the result of a more classical optimization problem. In some cases the optimization problem can be solved by numerical computation, which makes it possible to know the rate function (see Example 2.9).

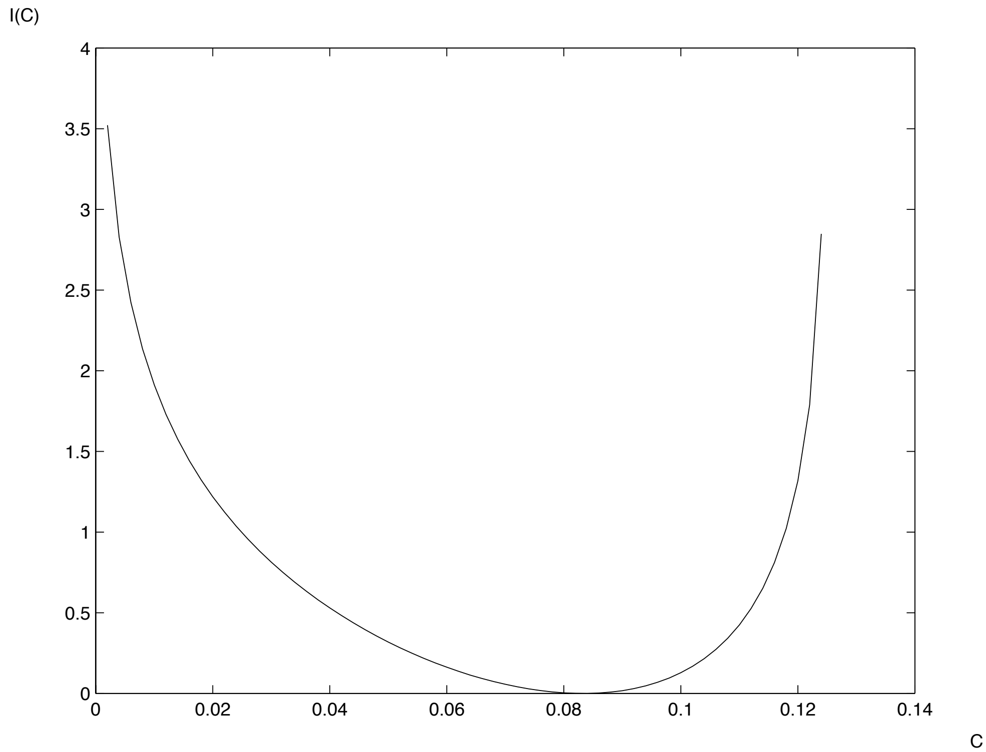

Figure 2.1 Rate function for $a(t) = t - \frac{1}{2}$

Theorem 2.8 *Suppose that F is the uniform distribution. The L-statistics $\int_0^1 a(t)F_n^{-1}(t)dt$ satisfy a LDP in $\mathbb{R}^k$, with good rate function*

$$I(C) = 1 + \sup_{\lambda \in \mathbb{R}, \tilde{\lambda} \in \mathbb{R}^{k+1}} \left\{ \lambda + \langle \tilde{\lambda}, C \rangle + \int_0^1 \log\left(-\lambda - \langle \tilde{\lambda}, A(s) \rangle \right) ds \right\}. \tag{2.8}$$

The theorem is proved in Subsection 2.4. The following example is a particular case of Example 2.7, in which the rate function is obtained by a numerical calculus.

Example 2.9 $a(t) = t - \frac{1}{2}$.

Some considerations on (2.8) lead to $I(C) = +\infty$ for $C \notin (0, \frac{1}{8})$. For $C \in (0, \frac{1}{8})$, in this particularly simple case, it is possible to calculate the quantity

$$Int(\lambda) = \int_0^1 1 + \log(\langle -\lambda, \bar{A}(s) \rangle) ds$$

in terms of elementary functions:

$$Int(\lambda) = \begin{cases} -1 + \log(-\lambda_1) + 2\sqrt{\frac{-8\lambda_1 - \lambda_2}{\lambda_2}} \arctan\left(\frac{\sqrt{\lambda_2}}{\sqrt{-8\lambda_1 - \lambda_2}} \right) & \text{for } \lambda_2 > 0, \lambda_1 < -\frac{\lambda_2}{8}, \\ -1 + \log(-\lambda_1) + \sqrt{\frac{8\lambda_1 + \lambda_2}{\lambda_2}} \left[2\log\left(1 + \sqrt{\frac{8\lambda_1 + \lambda_2}{\lambda_2}} \right) - \log\left(\frac{8\lambda_1}{\lambda_2} \right) \right] & \text{for } \lambda_2 < 0, \lambda_1 < 0, \\ +\infty & \text{else.} \end{cases}$$

In Figure 2.2, the graph of $I(C)$ has been obtained by numerical maximization with AMPL. We can check that the minimum of the rate function is attained at $C = \frac{1}{12} = \int a(t)F^{-1}(t)dt$.

2.3 Towards a LDP for normalized L-statistics

In this section, we derive a LDP for the standardized empirical quantile function. The aim is to treat some normalized L-statistics, under Condition (2.1). An example is D'Agostino's goodness of fit statistic (see D'Agostino [3]), defined as

$$D = \frac{\sum_{i=1}^{n}(i-(n+1)2^{-1})X_{(i)}}{n^2 S_n}. \tag{2.9}$$

Theorem 2.10 *Suppose that Condition (2.1) is fulfilled and that the underlying law of the sample has no atoms. Then the* normalized empirical quantile function

$$F_n^{-1,N} = \sum_{i=1}^{n} 1_{\left(\frac{i-1}{n}\,\frac{i}{n}\right]} \frac{X_{(i)}-\bar{X}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}}$$

satisfies a LDP in $\mathcal{M}_2(\mathbb{R})$ with good rate function

$$I_5(G^{-1}) = \begin{cases} \inf_{\mu\in\mathbb{R},\sigma>0} K\left(G\left(\frac{\cdot-\mu}{\sigma}\right),F\right) & \textit{when } G \textit{ has mean } 0 \textit{ and standard deviation } 1, \\ +\infty & \textit{else.} \end{cases}$$

Proof: The proof follows the arguments for Theorem 2.1; the only additional work is to take into account the normalization procedure, as follows:

$$\begin{aligned} No : \mathcal{M}_2(\mathbb{R}) &\to L_2(0,1) \\ G^{-1} &\mapsto \frac{G^{-1}-\int G^{-1}}{\left(\int\left(G^{-1}-\int G^{-1}\right)^2\right)^{\frac{1}{2}}}, \end{aligned} \tag{2.10}$$

which maps F_n^{-1} to $F_n^{-1,N}$.

This map (2.10) is continuous for the Wasserstein distance at every quantile function G^{-1} corresponding to a non-zero variance r.v.. Indeed, by the Cauchy-Schwarz inequality, the maps $G^{-1}\mapsto\int G^{-1}$ and $G^{-1}\mapsto\int\left(G^{-1}\right)^2$ are continuous in $L_2(0,1)$, so the map (2.10) is continuous as a composition of continuous maps, at quantile functions with non-zero variance.

We have supposed that the X_i have a continuous distribution. This allows to prove the continuity of the normalization map on a sufficiently large subset of quantile functions. Indeed, the continuity of F implies that any random variable which is absolutely continuous with respect to X_i has also a continuous distribution. In particular, the variance is positive. Hence, the normalization map No is continuous at any G^{-1} such that $I_1(G^{-1})<\infty$. That permits applying the contraction principle (Theorem 4.1.2, followed by Remark (c) p127 in Dembo and Zeitouni [5]), to obtain a LDP for $No(F_n^{-1})$. The good rate function is

$$I_5(G^{-1}) = \inf\{I_1(H^{-1}) : No(H^{-1}) = G^{-1}\} = \inf_{\mu\in\mathbb{R},\sigma>0} K\left(G\left(\frac{\cdot-\mu}{\sigma}\right),F\right).$$

□

Corollary 2.11 *Suppose that Condition (2.1) is satisfied and let a be some function in $L_2(0,1)$. Then the* normalized L-statistics

$$A_n^N := \sum_{i=1}^{n}\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} a(t)dt\right)\frac{X_{(i)}-\bar{X}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}} \tag{2.11}$$

satisfy a LDP with good rate function

$$I(C) = \inf_{\{G^{-1}\in\mathcal{M}_2(\mathbb{R}):\int aG^{-1}=C\}}\ \inf_{\mu\in\mathbb{R},\sigma>0} K\left(G\left(\frac{\cdot-\mu}{\sigma}\right),F\right). \tag{2.12}$$

Example 2.12 *D'Agostino's test statistic.*

Let a be the function defined on $[0,1]$ by $a(t) = t - \frac{1}{2}$. The coefficients

$$\int_{\frac{i-1}{n}}^{\frac{i}{n}} a(t)dt = \frac{i-(n+1)2^{-1}}{n^2}$$

are exactly the same as in D'Agostino's test statistics given in (2.9). Moreover, a is centered, therefore the normalized L-statistics $\int_0^1 a(t)F_n^{-1,N}(t)dt$ is equal to expression (2.9). As a consequence, Corollary (2.11) can be applied.

2.4 Proofs

Proof of Theorem 2.1. We first prove Theorem 2.1 under (i). The result can be reformulated as a LDP for the empirical measure $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ in $\mathcal{P}_2(\mathbb{R})$. Indeed, the operation which maps a measure in $\mathcal{P}_2(\mathbb{R})$ to its quantile function in $\mathcal{M}_2(\mathbb{R})$ is a topological isomorphism. Hence, we prove that μ_n satisfies a LDP in $\mathcal{P}_2(\mathbb{R})$, with good rate function

$$P \mapsto J_1(P) = K(G,F),$$

where G denotes the distribution function of $P \in \mathcal{P}_2(\mathbb{R})$.

We introduce the subsets of $\mathcal{M}_2(\mathbb{R})$:

$$K_M = \left\{ P \in \mathcal{P}_2(\mathbb{R}) : \int x^2\varphi(x)dP \le M \right\}, \text{ for } M > 0,$$

where φ is given in Condition (2.1). The LDP is based on the following facts.

$$K_M \text{ is a compact subset of } \mathcal{P}_2(\mathbb{R}). \tag{2.13}$$

$$(\mu_n) \text{ is exponentially tight in } \mathcal{P}_2(\mathbb{R}). \tag{2.14}$$

(2.13) can be proved as follows: let (Z_n) be a sequence of random variables with law $P_{Z_n} \in K_M$. Convergence in Wasserstein distance means convergence in distribution and convergence of the second moment (see del Barrio et al. [4], Proposition 3.1, for a summary of some properties of the Wasserstein distance). The hypothesis that $P_{Z_n} \in K_M$ ensures that (Z_n) is tight and that (Z_n^2) is uniformly integrable, hence we can find a random variable Z and extract a subsequence (Z_{n_k}) such that $Z_{n_k} \xrightarrow{d} Z$ and $E(Z_{n_k}^2) \to E(Z^2)$.

The tightness (2.14) of (μ_n) is proved as follows. For $t > 0$ such as in (2.1),

$$\begin{aligned} P(\mu_n \notin K_M) &= P\left(\frac{1}{n}\sum_{i=1}^n Z_i^2\varphi(Z_i) > M\right) \\ &= P\left(\sum_{i=1}^n tZ_i^2\varphi(Z_i) > tnM\right) \\ &\le e^{-ntM}\left(E\left(e^{tZ_1^2\varphi(Z_1)}\right)\right)^n, \end{aligned}$$

by Markov's exponential inequality. Therefore,

$$\frac{1}{n}\log P(\mu_n \notin K_M) \le -tM + \log E\left(e^{tZ_1^2\varphi(Z_1)}\right)$$

tends to $-\infty$ as $M \to +\infty$, which proves that (μ_n) is exponentially tight.

Now, we identify the rate function thanks to Sanov's theorem. The injection

$$i : (\mathcal{P}_2(\mathbb{R}),\ \text{Wasserstein distance}) \mapsto (\mathcal{P}(\mathbb{R}),\ \text{weak convergence})$$

is continuous because the weak topology is weaker than that given by the Wasserstein distance. Suppose that a subsequence (μ_{n_k}) of (μ_n) satisfies a LDP in $\mathcal{P}_2(\mathbb{R})$ with good rate function J. J_1 is the rate function given by

Sanov's theorem for the empirical measure. We now prove that necessarily, $J = J_1$: by the contraction principle, $i(\mu_{n_k})$ satisfies a LDP in $\mathcal{P}(\mathbb{R})$ with good rate function

$$J'(P) = \inf\{J(Q) : i(Q) = P\} = \begin{cases} J(P) \text{ if } P \in \mathcal{P}_2(\mathbb{R}) \\ +\infty \text{ else.} \end{cases}$$

But $i(\mu_{n_k}) = \mu_{n_k}$is already known to satisfy a LDP in $\mathcal{P}(\mathbb{R})$, with good rate function J_1, by Sanov's theorem. Therefore, for $P \in \mathcal{P}_2(\mathbb{R})$, $J(P) = J_1(P)$.

We can now conclude the existence of a LDP for (μ_n) in $\mathcal{P}_2(\mathbb{R})$. Let S be a measurable set in $\mathcal{M}_2(\mathbb{R})$, we want to prove that:

$$-\inf_{P\in \mathring{S}} J_1(P) \le \liminf \frac{1}{n} \log P(\mu_n \in S) \le \limsup \frac{1}{n} \log P(\mu_n \in S) \le -\inf_{P\in \bar{S}} J_1(P).$$

We shall prove here only the lower bound, since the argument for the upper bound is similar. Suppose μ_{n_k} is such that

$$\lim_{k\to\infty} \frac{1}{n_k} \log P(\mu_{n_k} \in S) = \liminf \frac{1}{n} \log P(\mu_n \in S).$$

By Lemma 4.1.23 in Dembo and Zeitouni [5] and the fact that (μ_n) is exponentially tight, we can extract a subsequence $(\mu_{n_{k_m}})_{m\in\mathbb{N}}$, that satisfies a LDP in $\mathcal{P}_2(\mathbb{R})$, with good rate function J_1. Hence in particular, the following inequality is satisfied:

$$-\inf_{P\in \mathring{S}} J_1(P) \le \liminf \frac{1}{n_k} \log P(\mu_{n_k} \in S) = \liminf \frac{1}{n} \log P(\mu_n \in S).$$

This proves the lower bound of the LDP for μ_n in $\mathcal{P}_2(\mathbb{R})$.

Under (ii), a troncation argument is involved. Let us define the *truncated empirical quantile function* as

$$F_n^{-1,T} = \sum_{i=1}^n 1_{(\frac{i-1}{n},\frac{i}{n}]} X_{(i)}^T,$$

where $\left(X_{(i)}^T\right)_{i=1}^n$ are the order statistics associated to the truncated i.i.d. variables $\left(X_i^T\right)_{i=1}^n$. By the part (i) of the theorem, $F_n^{-1,T}$ satisfies a LDP with good rate function $J_2(G^{-1}) = K(G, F^T)$ since the support of $F_n^{-1,T}$ is bounded. We now prove that it is an exponentially good approximation of F_n^{-1}. That makes is possible to apply Theorem 4.2.16 p131 of Dembo and Zeitouni [5] on exponentially good approximations and conclude. Hence, we want to prove that: $\forall \epsilon > 0$

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(\|F_n^{-1} - F_n^{-1,T}\|_2 \ge \epsilon\right) \to -\infty, T \to +\infty. \tag{2.15}$$

Notice that

$$\begin{aligned} F_n^{-1} - F_n^{-1,T} &= \sum_{i=1}^n 1_{(\frac{i-1}{n},\frac{i}{n}]} \left(X_{(i)} - X_{(i)}^T\right) \\ &= \sum_{i=1}^n 1_{(\frac{i-1}{n},\frac{i}{n}]} \left[(X_i - T)\, 1_{X_i > T} + (X_i + T)\, 1_{X_i < -T}\right] \end{aligned}$$

The square of the L_2-norm of this variable is

$$\frac{1}{n} \sum_{i=1}^n \left[(X_i - T)^2\, 1_{X_i > T} + (X_i + T)^2\, 1_{X_i < -T}\right],$$

which is bounded by

$$\frac{2}{n} \sum_{i=1}^n X_i^2 1_{X_i^2 > T^2}.$$

So, we have the following inequalities for all positive t, by Markov's exponential inequality:

$$\begin{aligned}\frac{1}{n}\log P\left(\|F_n^{-1}-F_n^{-1,T}\|_2\geq\epsilon\right) &\leq \frac{1}{n}\log P\left(\frac{2}{n}\sum_{i=1}^n X_i^2 1_{X_i^2>T^2}>\epsilon^2\right)\\ &\leq -\frac{\epsilon^2 t}{2}+\log E\left(e^{tX_i^2 1_{X_i^2>T^2}}\right)\end{aligned}$$

But $E\left(e^{tX_i^2 1_{X_i^2>T^2}}\right)=E\left(e^{tX_i^2}1_{X_i^2>T^2}\right)+P\left(X_i^2\leq T^2\right)\to 1, T\to\infty$ by Lebesgue's dominated convergence theorem.

So for all positive t,

$$\limsup_{T\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log P\left(\|F_n^{-1}-F_n^{-1,T}\|_2\geq\epsilon\right)\leq-\frac{\epsilon^2 t}{2}.$$

Hence (2.15) is satisfied, which proves that the exponential approximation of F_n^{-1} by $F_n^{-1,T}$ holds.

□

Proof of Theorem 2.3. To begin with, we suppose that (i) is satisfied. We first introduce the auxiliary measure λ_n, which is the measure on $[0,1]$ having density (with respect to the Lebesgue measure λ on $[0,1]$):

$$\frac{d\lambda_n}{d\lambda}=F_n^{-1}=\sum_{i=1}^n X_{(i)}1_{(\frac{i-1}{n},\frac{i}{n}]}. \tag{2.16}$$

Let $\mathcal{M}_{2,+}(\mathbb{R})$ denote the subset of $\mathcal{M}_2(\mathbb{R})$ formed by the positive quantile functions. By Theorem 2.1 and by the continuity of the application

$$\begin{aligned}\mathcal{M}_{2,+}(\mathbb{R}) &\to \mathcal{P}_+([0,1])\\ G^{-1} &\mapsto P \text{ with } P\ll\lambda,\ \frac{dP}{d\lambda}=G^{-1}.\end{aligned}$$

the measure λ_n satisfies a LDP on $\mathcal{P}_+(0,1)$ with good rate function $\tilde{I}_1$. The LDP for ν_n can be deduced from the fact that ν_n and λ_n are exponentially equivalent. This holds under weaker hypothesis, which we state in the following lemma.

Lemma 2.13 *Suppose that F has (non necessarily bounded) support in $\mathbb{R}_+$ and that its Laplace transform is defined for some $t>0$. Then the measures ν_n and λ_n, defined in (2.3) and (2.16) respectively, are exponentially equivalent.*

Proof: We will use the Lipschitz bounded metric. Let δ be some positive number. Let a be some continuous function on $[0,1]$ with uniform bound and Lipschitz constant bounded by 1.

$$\begin{aligned}|\nu_n(a)-\lambda_n(a)| &= \left|\frac{1}{n}\sum_{i=1}^n X_{(i)}a\left(\frac{i}{n}\right)-\sum_{i=1}^n X_{(i)}\int_{\frac{i-1}{n}}^{\frac{i}{n}}a(t)dt\right|\\ &\leq \sum_{i=1}^n X_{(i)}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\left|a\left(\frac{i}{n}\right)-a(t)\right|dt\\ &\leq \frac{1}{n^2}\sum_{i=1}^n X_i, \end{aligned} \tag{2.17}$$

where the inequality (2.17) uses the Lipschitz condition on a. Hence,

$$\begin{aligned}\frac{1}{n}\log P(d_{LB}(\nu_n,\lambda_n)>\delta) &\leq \frac{1}{n}\log P\left(\frac{1}{n^2}\sum_{i=1}^n X_{(i)}>\delta\right)\\ &\leq \frac{1}{n}\log\left(e^{-n^2t\delta}\left(e^{tX_i}\right)^n\right) \qquad (2.18)\\ &= -nt\delta+\log\psi(t)\to-\infty \text{ as } n\to\infty,\end{aligned}$$

where (2.18) holds for any $t > 0$ such that $\psi(t) = E(e^{tX_i}) < \infty$, by Markov's exponential inequality. □

Under Condition (ii), a truncation argument as in Theorem 2.1 yields the conclusion. The point is now that the truncated measure

$$\nu_n^T = \frac{1}{n}\sum_{i=1}^{n} \delta_{\frac{i}{n}} X_{(i)}^T$$

is an exponentially good approximation of ν_n for the topology of the bounded Lipschitz metric. Let a be some continuous function on $[0,1]$ with uniform bound and Lipschitz constant bounded by 1.

$$\begin{aligned} \left|\nu_n(a) - \nu_n^T(a)\right| &= \left|\frac{1}{n}\sum_{i=1}^{n} a(\frac{i}{n})(X_{(i)}^T - X_{(i)})\right| \\ &= \frac{1}{n}\sum_{i=1}^{n}(X_i - T)1_{X_i>T} + (-T - X_i)1_{X_i<-T} \\ &\leq \frac{2}{n}\sum_{i=1}^{n}|X_i|1_{|X_i|>T} \end{aligned}$$

So it is sufficient to prove that for all $\epsilon > 0$,

$$\lim_{T\to\infty}\lim_{n\to\infty}\frac{1}{n}\log P(\frac{1}{n}\sum_{i=1}^{n}|X_i|1_{|X_i|>T} > \epsilon) = -\infty.$$

But by Markov's exponential inequality and the independence of the X_i, for all positive t,

$$\begin{aligned} \frac{1}{n}\log P(\frac{1}{n}\sum_{i=1}^{n}|X_i|1_{|X_i|>T} > \epsilon) &= \frac{1}{n}\log P(e^{t\sum_{i=1}^{n}|X_i|1_{|X_i|>T}} > e^{nt\epsilon}) \\ &\leq -t\epsilon + \log\left(E\left(e^{t|X_i|1_{|X_i|>T}}\right)\right). \end{aligned}$$

We conclude as in Theorem 2.1.

□

Proof of Theorem 2.8. The theorem is proved using some convex analysis tools. Namely, we transform the rate function I_0 by some duality arguments due to Borwein and Lewis [2].

Let G be a distribution function such that $G \ll F$. This implies that G and G^{-1} are derivable almost everywhere. Let U be a random variable with uniform law on $[0,1]$. Recall that $G^{-1}(U)$ has distribution function G. Hence,

$$\begin{aligned} K(G,F) &= \int_0^1 \log\frac{G'(G^{-1}(t))}{F'(G^{-1}(t))}\mathbf{1}_{G'(G^{-1}(t))\neq 0}dt \\ &= \int_0^1 \left(-\log(G^{-1})'(t)\mathbf{1}_{G'(G^{-1}(t))\neq 0}\right)dt. \end{aligned} \tag{2.19}$$

Here, we have used that $(G^{-1})'(t)$ is defined as soon as $G'(G^{-1}(t)) \neq 0$ and that $F'(x) = 1$. At points t such that $G'(G^{-1}(t)) \neq 0$, $(G^{-1})'(t)$ is the derivative of G^{-1} in the usual meaning. Such points t are of Lebesgue measure 1 in $(0,1)$. So, we have obtained:

$$K(G,F) = -E\log\left(G^{-1}\right)'(U),$$

with the convention $\log u = -\infty, u \leq 0$. Therefore $K(G,F)$ can be expressed as the following functional of $x = G^{-1}$:

$$K(x) = \begin{cases} -\int_0^1 \log x'(t)dt & \text{for } x : (0,1] \to [0,1] \text{ strictly increasing and derivable a.e.} \\ +\infty & \text{else,} \end{cases}$$

So the problem to be solved to compute the rate function I is: minimize $K(x)$ under the k-dimensional constraint $\int_0^1 a(t)x(t)dt = C$. Remark that since a is centered, for a given x and any constant $c \geq 0$ such that $x(1) + c \leq 1$, we have $K(x) = K(x + c)$, so that $\int_0^1 a(t)x(t)dt = \int_0^1 a(t)(x(t) + c)dt$. Moreover, if x is such that $x(1) > 1$, then $K(x) = +\infty$. Hence, we can add the constraint $x(1) = 1$. Now let $y = x'$, then $x(t) = 1 - \int_t^1 y(s)ds$. Using the fact that a is centered, an integration by parts and a Fubini argument, the constraint may be rewritten as:

$$\int_0^1 a(t)x(t)dt = \int_0^1 A(s)y(s)ds,$$

so the new problem $(\tilde{P})$ is to minimize

$$\tilde{K}(y) = -\int_0^1 \log y(t)dt$$

under the constraint $\int_0^1 A(s)y(s)ds = C$, with $0 \leq 1 - x(0) = \int_0^1 y(t)dt \leq 1$. The inequality $0 \leq \int_0^1 y(t)dt \leq 1$ leads to the introduction of one more constraint, and to a new series of problems
(P_α): minimize $\tilde{K}(y)$ under

$$\int_0^1 A(s)y(s)ds = C,$$

$$\int_0^1 y(t)dt = \alpha,$$

for $0 \leq \alpha \leq 1$. The value of the infimum for problem $(\tilde{P})$, denoted by $Val(\tilde{P})$, is the infimum of the values of the infimum for (P_α), for α varying in $[0, 1]$. Denote, for each α, the value of the infimum in problem (P_α) by $Val(P_\alpha)$. The solution can be found by duality arguments such as in Borwein and Lewis [2]. First, we prove that for each α, the dual problem (P^*_α), with supremum value denoted by $Val(P^*_\alpha)$, is

$$\sup_{\lambda \in \mathbf{R}^{k+1}} \langle \lambda, \bar{C}_\alpha \rangle + \int_0^1 1 + \log(\langle -\lambda, \bar{A}(s) \rangle)ds,$$

where $\bar{A}$ and $\bar{C}_\alpha$ are defined in the following way: for $A = (A_1, \ldots, A_k)$, $\bar{A} = (1, A_1, \ldots A_k) : [0, 1] \to \mathbb{R}^{k+1}$. For $C = (C_1, \ldots, C_k) \in \mathbb{R}^k$ and $\alpha \in [0, 1]$, define $\bar{C}_\alpha = (\alpha, C_1, \ldots, C_k)$.

This can be proved as follows. The problem (P_α) is to minimize

$$\tilde{K}(y) = -\int_0^1 \log y(t)dt \text{ under}$$

$$\int_0^1 \bar{A}(s)y(s)ds = \bar{C}_\alpha.$$

We check the hypothesis of Theorem 3.4 of Borwein and Lewis [2]:

$$\Phi : u \mapsto \begin{cases} -\log u & , u > 0 \\ +\infty, & u \leq 0 \end{cases}$$

satisfies $p = \lim_{u \to -\infty} \frac{\Phi(u)}{u} = -\infty$, $q = \lim_{u \to +\infty} \frac{\Phi(u)}{u} = 0$, and its dual function is given by $\Phi^* : s \mapsto -1 - \log(-s)$. From this follows the formulation of the dual problem for given α.

Now, we prove that $Val(P_\alpha) = Val(P^*_\alpha)$. Φ is not affine and there exists $\lambda \in \mathbb{R}^{k+1}$ such that $\langle \lambda, \bar{A}(s) \rangle \in (p, q)$ $\forall s \in [0, 1]$: just take $\lambda = (-1, 0, \ldots, 0)$. Hence the *Dual Constraint Qualification* is satisfied. The *Primal Constraint Qualification* is supposed to be satisfied, i.e. we suppose that there exists $\hat{y} \in L_1([0, 1])$ such that $\hat{y}(s) \in \mathbf{R}^*_+$ a.s. and $\hat{y}$ satisfies $\int_0^1 \bar{A}(s)\hat{y}(s)ds = \bar{C}_\alpha$. When α does not satisfy this hypothesis, $Val(P_\alpha) = +\infty$ so the problem does not have to be solved for this value of α. The conclusion of the theorem of Borwein and Lewis [2] is that $Val(P_\alpha) = Val(P^*_\alpha)$.

Recall that $Val(\tilde{P}) = \inf_{\alpha\in[0,1]} Val(P_\alpha)$. Now, we prove that $\inf_{\alpha\in[0,1]} Val(P_\alpha) = \inf_{\alpha\in[0,1]} Val(P^*_\alpha) = Val P^*_1$. We use a minimax theorem for convex functions (Fan, 1953, exposed in Roberts and Varberg [13] p138). The application of this theorem gives:

$$\begin{aligned} Val(\tilde{P}) &= \inf_{\alpha\in[0,1]} \sup_{\lambda\in\mathbf{R}^{k+1}} \left\{ \langle\lambda, \bar{C}_\alpha\rangle + \int_0^1 1 + \log(-\langle\lambda, \bar{A}(s)\rangle)ds \right\} \\ &= \sup_{\lambda\in\mathbf{R}^{k+1}} \inf_{\alpha\in[0,1]} \left\{ \langle\lambda, \bar{C}_\alpha\rangle + \int_0^1 1 + \log(-\langle\lambda, \bar{A}(s)\rangle)ds \right\}. \end{aligned}$$

A discussion of the sign of the first coordinate λ_1 of λ concludes the proof, as follows.

If $\lambda_1 > 0$, $\inf_{\alpha\in[0,1]} \left\{ \langle\lambda, \bar{C}_\alpha\rangle + \int_0^1 1 + \log(-\langle\lambda, \bar{A}(s)\rangle)ds \right\} = -\infty$ because A is continuous and takes value 0 at 0, so

$$-\langle\lambda, \bar{A}(s)\rangle = -\lambda_1 - \langle\lambda_{2,k+1}, A(s)\rangle \leq 0$$

in a neighborhood of 0 and the log is not defined. Here we have used the notation $\lambda_{2,k+1} = (\lambda_2, \ldots, \lambda_{k+1})$.

If $\lambda_1 = 0$, the function to minimize in α does not depend on α and is

$$\langle\lambda_{2,k+1}, C\rangle + 1 + \int_0^1 \log\left(-\langle\lambda_{2,k+1}, A(s)\rangle\right) ds.$$

If $\lambda_1 < 0$,

$$\inf_{\alpha\in[0,1]} \left\{ \langle\lambda, \bar{C}_\alpha\rangle + \int_0^1 1 + \log(-\langle\lambda, \bar{A}(s)\rangle)ds \right\} = \lambda_1 + \langle\lambda_{2,k+1}, C\rangle + 1 + \int_0^1 \log(-\lambda_1 - \langle\lambda_{2,k+1}, A(s)\rangle)ds.$$

But for any continuous function f on $\mathbb{R}^{k+1}$,

$$\sup_{\lambda\in\mathbf{R}^{k+1},\lambda_1<0} f(\lambda) \geq \sup_{\lambda\in\mathbf{R}^{k+1},\lambda_1=0} f(\lambda).$$

Moreover, when $\lambda_1 > 0$,

$$\int_0^1 \log(-\lambda_1 - \langle\lambda_{2,k+1}, A(s)\rangle)ds = -\infty,$$

as we have already seen. Hence:

$$\begin{aligned} Val(\tilde{P}) &= \sup_{\lambda_1\in\mathbf{R},\lambda_{2,k+1}\in\mathbf{R}^k} \left\{ \lambda_1 + \langle\lambda_{2,k+1}, C\rangle + 1 + \int_0^1 \log(-\lambda_1 - \langle\lambda_{2,k+1}, A(s)\rangle)ds \right\} \\ &= \sup_{\lambda\in\mathbf{R}^{k+1}} \left\{ \langle\lambda, \bar{C}_1\rangle + 1 + \int_0^1 \log(-\langle\lambda, \bar{A}(s)\rangle)ds \right\}. \end{aligned}$$

□

3 LDP for L-statistics with exponential underlying law

We assume here that F is the exponential distribution with parameter 1. The tails of this distribution are quite heavy, so that neither the tail condition (2.1) nore the hypothesis of existence of the Laplace transform of X_1^2 at some point are satisfied. So the method employed in Section 2 does not provide any LDP result for the empirical quantile function. Let us now have a look at the contracted LDP obtained in Groeneboom and Shorack [9]. The result is part of a LDP for L-statistics $\int a(t)F_n^{-1}(t)dt$ and one key condition is that

$$\int_{1/2}^1 \left|a(t)F^{-1}\left(1 - e^{-\frac{c}{1-t}}\right)\right| dt < \infty,$$

for some c. For positive a, using that $F^{-1}(t) = -\log(1-t)$, the condition can be written as:

$$\int_{1/2}^1 c\frac{a(t)}{1-t}dt < \infty. \tag{3.1}$$

Hence, we see that the function a has to tend to 0 quite fastly near 1. The result we present in Theorem 3.3 leads to a LDP without asking for this condition. It gives a functional LDP for the measure ν_n defined in (2.3), which is at an upper level than the result of Groeneboom and Shorack. By the contraction principle, it covers all statistics $\frac{1}{n}\sum_{i=1}^n a(i/n)X_{(i)}$ when a is continuous on $(0,1)$. In Remark 3.4 below, we discuss the relationship between these two results.

Let us recall some topological results. We deal with the measure

$$\nu_n = \frac{1}{n}\sum_{i=1}^n X_{(i)}\delta_{\frac{i}{n}},$$

which lies in the set $\mathcal{P}_+([0,1])$ of all positive bounded measures on $[0,1]$. As in Theorem 2.3, $\mathcal{P}_+([0,1])$ is endowed with the weak topology. It is a closed subset of $\mathcal{P}([0,1])$ which denotes the set of all finite regular $\mathbb{R}$-valued measures on $[0,1]$. This is the dual space of $\mathcal{C}([0,1])$ for the uniform topology. We employ the abstract Gärtner-Ellis Theorem. The tool we use is a duality argument. For a in $C([0,1])$, let

$$\Lambda(a) = \begin{cases} -\int_0^1 \log\left[1 - \frac{\int_{1-t}^1 a(u)du}{t}\right] dt & \text{whenever the integral is defined} \\ +\infty & \text{else} \end{cases} \tag{3.2}$$

Denote by

$$\Lambda^*(\mu) = \sup_{a\in C([0,1])}\left[\int_0^1 a(t)d\mu(t) - \Lambda(a)\right] \tag{3.3}$$

the dual function of Λ, which may take infinite values.

The following theorem gives an explicit expression for Λ^*.

Theorem 3.1 *Let $\mu \in \mathcal{P}([0,1])$. Suppose that μ admits the decomposition $\mu = l\lambda + \mu(\{1\})\delta_1$, where $l(u) = \int_0^u dm(s)$ and the Lebesgue decomposition of m is $m = \alpha\lambda + \chi$. Moreover, suppose that the singular measure $-td\chi(1-t) + \mu(\{1\})\delta_0$ is positive. Then*

$$\Lambda^*(\mu) = \int_0^1 (t\alpha(1-t) - \log\alpha(1-t))\, dt - \int_0^1 td\chi(1-t) + \mu(\{1\}).$$

Else, $\Lambda^*(\mu) = +\infty$.

In order to understand Λ^*, let us consider a simple case: $\mu = l\lambda$, when the density l is derivable, with derivated α. Hence:

$$\mu([0,t]) = \int_0^t l(u)du,$$
$$l(u) = \int_0^u dm(s) = \int_0^u \alpha(s)ds.$$

Since all the singular measures of the decomposition are equal to zero, the rate function, in this case, has the expression:

$$\Lambda^*(\mu) = \int_0^1 (t\alpha(1-t) - \log\alpha(1-t))\, dt.$$

Proof: Remark that Λ can be decomposed in the following way: $\Lambda = \Gamma\circ T$, where

$$\begin{aligned} T : \mathcal{C}([0,1]) &\to \mathcal{C}([0,1]) \\ a &\mapsto Ta : t \mapsto \frac{1}{t}\int_{1-t}^1 a(u)du, \\ \Gamma : \mathcal{C}([0,1]) &\to \mathbb{R} \\ b &\mapsto \begin{cases} -\int_0^1 \log(1-b(t))\, dt & \text{if the integral is defined} \\ +\infty & \text{else.} \end{cases} \end{aligned}$$

Here, $Ta(0)$ is defined by continuity as $a(1)$. T is a linear, continuous function on $\mathcal{C}([0,1])$ for the uniform topology and hence has a closed graph. Γ is a proper convex function on $\mathcal{C}([0,1])$. Therefore, Theorem 19 of Rockafellar [15], can be applied. Condition (a) in Rockafellar [15] is satisfied: $\exists a \in \text{dom}\, T$ such that Γ is bounded above on a neighborhood of Ta: just take $a = 0$. The conclusion is

$$\Lambda^*(\mu) = \min\{\Gamma^*(\nu) : \nu \in \text{dom}\, T^*, T^*\nu = \mu\}$$

for $\mu \in M([0,1])$, where $\Gamma^* : M([0,1]) \to \mathbb{R}$ is the dual function of Γ, and $T^* : M([0,1]) \to M([0,1])$ is the dual function of T.

T^* can be calculated via an application of Fubini's theorem. It is defined by: $\langle Ta, \nu\rangle = \langle a, T^*\nu\rangle$ for $\nu \in M([0,1])$ and $a \in \mathcal{C}([0,1])$. We have

$$\begin{aligned}\int_0^1 \frac{\int_{1-t}^1 a(u)du}{t} d\nu(t) &= \int_{(0,1]} \frac{\int_{[1-t,1)} a(u)du}{t} d\nu(t) + a(1)\nu(\{0\}) \\ &= \int_{[0,1)} \left(\int_{[1-u,1]} \frac{1}{t} d\nu(t) \right) a(u)du + a(1)\nu(\{0\}),\end{aligned}$$

by Fubini's theorem. Hence $T^*\nu$ is the measure with Lebesgue decomposition

$$T^*\nu = \int_{1-\cdot}^1 \frac{1}{t} d\nu(t)\lambda + \nu(\{0\})\delta_1.$$

Now for $\mu = l\lambda + \mu(\{1\})\delta_1$, let us seek for ν such that $T^*\nu = \mu$. By equality between the two singular parts:

$$\nu(\{0\}) = \mu(\{1\}).$$

The densities with respect to λ of the absolutely continuous parts are equal λ-a.s., so for λ-a.e. u:

$$\int_{1-u}^1 \frac{1}{t} d\nu(t) = l(u). \tag{3.4}$$

This implies that l has bounded variation and can be written

$$l(u) = \int_0^u dm(s) \text{ with } m = \alpha\lambda + \chi.$$

So $l(u) = \int_{1-u}^1 \alpha(1-t)dt - \int_{1-u}^1 d\chi(1-t)$. By equation (3.4), ν satisfies for λ-a.e. $t \neq 0$:

$$\frac{1}{t} d\nu(t) = \alpha(1-t)dt - d\chi(1-t).$$

Hence $d\nu(t) = t\alpha(1-t)dt - td\chi(1-t) + \mu(\{1\})d\delta_0(t)$.

The last step is the calculus of $\Gamma^*(\nu)$. An application of the results of Theorem 5 of Rockafellar [14] yields the following auxilliary lemma:

Lemma 3.2 *For $b \in \mathcal{C}([0,1])$, let*

$$\Gamma(b) = \begin{cases} -\int_0^1 \log(1-b(t))dt & \textit{if the integral is defined} \\ +\infty & \textit{else}. \end{cases}$$

Let ν be a signed measure on $[0,1]$ and $\nu = \beta\lambda + \psi$ its Lebesgue decomposition. Then

$$\Gamma^*(\nu) := \sup_{b \in \mathcal{C}([0,1])} \left\{ \int_0^1 b(t)d\nu(t) - \Gamma(b) \right\}$$

is given by the expression:

$$\Gamma^*(\nu) = \begin{cases} \int_0^1 (\beta(t) - 1 - \log\beta(t))\, dt + \psi([0,1]) & \textit{if } \psi \textit{ is nonnegative,} \\ +\infty & \textit{else,} \end{cases}$$

with the convention $-\log t = +\infty$ if $t \leq 0$.

So in our case,

$$\Lambda^*(\mu) = \Gamma^*(\nu) = \int_0^1 \left(t\alpha(1-t) - 1 - \log\left(t\alpha(1-t)\right)\right) dt - \int_0^1 t d\chi(1-t) + \mu(\{1\}),$$

which gives the expression announced in Theorem 3.1. □

We now state the main result of this section.

Theorem 3.3 *ν_n satisfies a LDP with good rate function Λ^*.*

Before the proof, let us compare the LDP for L-statistics induced by Theorem 1.1 and Theorem 3.3.

Remark 3.4 *Relationship between Theorem 1.1 and by Theorem 3.3.* With both results, we can obtain LDP-type results for L-statistics $\int_0^1 a(t)F_n^{-1}(t)dt$ for suitable a. Indeed, although we have formulated Theorem 3.3 for ν_n, the exponential equivalence given in Lemma 2.13 allows to state the LDP for F_n^{-1}. Let a be some function satisfying both the decay condition of Theorem 1.1 and the continuity condition of Theorem 3.3. In order to compare the two results, let us consider the following. Let $r \in \mathbb{R}$. By Theorem 1.1, we have

$$\lim \frac{1}{n} \log P\left(\int_0^1 a(t)F_n^{-1}(t)dt \le r\right) = -\inf\left\{K(G,F) : \int aG^{-1} \le r\right\}.$$

Similarly, Theorem 3.3 implies that:

$$-\inf\left\{\Lambda^*(\mu) : \int a d\mu < r\right\} \le \liminf \frac{1}{n} \log P\left(\int_0^1 a(t)F_n^{-1}(t)dt \le r\right)$$
$$\le \limsup \frac{1}{n} \log P\left(\int_0^1 a(t)F_n^{-1}(t)dt \le r\right) \le -\inf\left\{\Lambda^*(\mu) : \int a d\mu \le r\right\}.$$

We can observe that Theorem 1.1 is slightly more precise since it gives a limit instead of the limsup and liminf. But we shall see that this contracted LDP theorem is less informative on the underlying LDP on ν_n. Let us make the link between the rate functions. Let us suppose that G^{-1} is such that $\int aG^{-1} \le r$ and $K(G,F) < \infty$. We can define a measure μ by $d\mu = G^{-1}d\lambda$. Then, obviously, $\int a d\mu \le r$. Moreover, $\Lambda^*(\mu) = K(G,F)$. Indeed, we can compute $K(G,F)$ as in (2.19):

$$\begin{aligned} K(G,F) &= \int_0^1 \log \frac{G'(G^{-1}(t))}{F'(G^{-1}(t))} \mathbf{1}_{G'(G^{-1}(t))\neq 0} dt \\ &= \int_0^1 \left(-\log(G^{-1})'(t)\mathbf{1}_{G'(G^{-1}(t))\neq 0} + G^{-1}(t)\right) dt. \end{aligned} \quad (3.5)$$

Here, we have used $F'(x) = e^{-x}$. Let us denote, as in Theorem 3.1, $l = G^{-1}$. Let m be such that $l(u) = \int_0^u dm(s)$ with $m = \alpha\lambda + \chi$ the Lebesgue decomposition of m. Again, at points t such that $G'(G^{-1}(t)) \neq 0$, $(G^{-1})'(t)$ is the derivative of G^{-1} in the usual meaning and is equal to $\alpha(t)$. Such points t are of Lebesgue measure 1 in $(0,1)$. (3.5) becomes:

$$\begin{aligned} K(G,F) &= \int_0^1 \left(-\log\alpha(t) + l(t)\right) dt \\ &= -\int_0^1 \log\alpha(t)dt + \int_0^1 \int_0^t dm(s)dt \\ &= -\int_0^1 \log\alpha(t)dt + \int_0^1 \int_s^1 dt dm(s) \\ &= -\int_0^1 \log\alpha(t)dt + \int_0^1 (1-s)\alpha(s)ds + \int_0^1 (1-s)d\chi(s) \\ &= -\int_0^1 \log\alpha(1-t)dt + \int_0^1 s\alpha(1-s)ds + \int_0^1 s d\chi(1-s). \\ &= \Lambda^*(\mu). \end{aligned} \quad (3.6)$$

This shows how the contraction step $\mu \mapsto \int a d\mu$ for functions a decaying at 1 causes a loss of information on the underlying LDP for the measure ν_n. The measures μ which have to be considered are only the particular measures $d\mu = G^{-1}d\lambda$ for some quantile function G^{-1}. The decay of a at 1 makes the possible weights of measures μ at 1 disappear from the rate function.

Proof: The proof follows the same ideas as in Gamboa et al. [6]. We will use an analogue of the techniques developed in Lemmas 7 and 8 therein to prove the lower bound. The abstract Gärtner-Ellis theorem (Theorem 4.5.3 of Dembo and Zeitouni [5]) provides the upper bound for compact sets. Exponential tension is obtained via Cramer's LDP for sums of i.i.d. real-valued random variables, which gives the upper bound for closed sets. Next, the lower bound is derived from Baldi's theorem (Theorem 4.5.20 of Dembo and Zeitouni [5]) thanks to a study of the exposed points.

Let us first check the hypothesis of the abstract Gärtner-Ellis theorem. For any function a in $C([0,1])$, define

$$\Lambda_n(a) = \log E(\exp[\nu_n(a)]). \tag{3.7}$$

We have to study the possible limit of $\frac{1}{n}\Lambda_n(na)$, which we will call $\overline{\Lambda}(a)$. The calculations are possible thanks to a suitable representation of the uniform order statistics with normalized sums of i.i.d. exponential random variables. This is a very particular and interesting case. Unfortunately, this method seems difficult to generalize.

The possible limits for $\frac{1}{n}\Lambda_n(a)$ are studied in the following lemma:

Lemma 3.5 *(i) Suppose that for all t in $(0,1]$, $\int_{1-t}^1 a(u)du < t$ and that $a(1) < 1$. Then the limit of $\frac{1}{n}\Lambda_n(a)$ is finite and coincides with $\Lambda(a)$.*
(ii) Suppose there exists t in $(0,1]$ such that $\int_{1-t}^1 a(u)du > t$ or $a(1) > 1$. Then the limit of $\frac{1}{n}\Lambda_n(a)$ is infinite and coincides with $\Lambda(a)$.

Proof: We make use of a representation of the quantiles of the uniform distribution, as follows: let $\xi_1, \dots, \xi_{n+1}$ be an i.i.d. sample of exponential law with parameter 1. Denote by $U_{(i)}$ the i-th uniform order statistics from a sample of size n. The following equality holds in distribution:

$$\left(U_{(i)}\right)_{i=1}^n =^d \left(\frac{\xi_1 + \dots + \xi_i}{\xi_1 + \dots + \xi_{n+1}}\right)_{i=1}^n.$$

Let F be the distribution function of the exponential law with parameter 1, namely $F^{-1}(t) = -\log(1-t)$ for t in $[0,1[$. Then $\left(X_{(i)}\right)_{i=1}^n = \left(F^{-1}(U_{(i)})\right)_{i=1}^n$ has the distribution of the order statistics derived from an exponential sample with parameter 1, so that we have the following distributional equality: jointly for $i = 1, \dots, n$

$$X_{(i)} = -\log\left(1 - \frac{\xi_1 + \dots + \xi_i}{\xi_1 + \dots + \xi_{n+1}}\right) = -\log\frac{\xi_{i+1} + \dots + \xi_{n+1}}{\xi_1 + \dots + \xi_{n+1}}.$$

Therefore,

$$\begin{aligned} e^{\Lambda_n(na)} &= E\left(e^{\sum_{i=1}^n a\left(\frac{i}{n}\right)X_{(i)}}\right) \\ &= \int_{\mathbb{R}_+^{n+1}} \left(\frac{x_2 + \dots + x_{n+1}}{x_1 + \dots + x_{n+1}}\right)^{-a\left(\frac{1}{n}\right)} \dots \left(\frac{x_{n+1}}{x_1 + \dots + x_{n+1}}\right)^{-a\left(\frac{n}{n}\right)} e^{-x_1 - \dots - x_{n+1}} dx_1 \dots dx_{n+1} \end{aligned}$$

Let us make the triangular change of variables:
$u_1 = x_{n+1}$
$u_2 = x_{n+1} + x_n$
$\dots$
$u_k = x_{n+1} + \dots + x_{n+2-k}$
$\dots$
$u_{n+1} = x_{n+1} + \dots + x_1$.
To simplify the notations, define

$$\kappa_k = -a\left(\frac{n-k+1}{n}\right) \tag{3.8}$$

and $T_k = \{(u_k, \dots, u_{n+1}) \in \mathbb{R}^{n-k} : 0 < u_k < \cdots < u_{n+1}\}$, for $1 \le k \le n$. So

$$e^{\Lambda_n(na)} = \int_{T_1} \left(\frac{u_1}{u_{n+1}}\right)^{\kappa_1} \cdots \left(\frac{u_n}{u_{n+1}}\right)^{\kappa_n} e^{-u_{n+1}} du_1 \dots du_{n+1}$$

Let $1 \le k \le n$ such that for every $1 \le j \le k$,

$$\kappa_1 + \cdots + \kappa_j + j - 1 > -1. \tag{3.9}$$

Then by induction,

$$e^{\Lambda_n(na)} = \frac{1}{\kappa_1 + 1} \cdots \frac{1}{\kappa_1 + \cdots + \kappa_k + k} \int_{T_{k+1}} u_{k+1}^{\kappa_1 + \cdots + \kappa_k + k} u_{k+1}^{\kappa_{k+1}} \dots u_n^{\kappa_n} u_{n+1}^{-\kappa_1 - \cdots - \kappa_n} e^{-u_{n+1}} du_{k+1} \dots du_{n+1}.$$

Therefore if (3.9) holds for $k = n$, the induction yields

$$\begin{aligned} e^{\Lambda_n(na)} &= \frac{1}{\kappa_1 + 1} \cdots \frac{1}{\kappa_1 + \cdots + \kappa_n + n} \int_{\mathbb{R}_+} u_{n+1}^{\kappa_1 + \cdots + \kappa_n + n} u_{n+1}^{-\kappa_1 - \cdots - \kappa_n} e^{-u_{n+1}} du_{n+1} \\ &= \Gamma(n+1) \prod_{j=1}^{n} \frac{1}{\kappa_1 + \cdots + \kappa_j + j} \\ &= \prod_{j=1}^{n} \frac{j}{\kappa_1 + \cdots + \kappa_j + j}. \end{aligned}$$

Now returning to expression (3.8):

$$\frac{1}{n} \Lambda_n(na) = -\frac{1}{n} \sum_{j=1}^{n} \log\left(1 - \frac{1}{j} \sum_{l=1}^{j} a\left(\frac{n-l+1}{n}\right)\right).$$

Else now suppose that for some $k \le n$, $\kappa_1 + \cdots + \kappa_k + k - 1 \le -1$. Then the k-th integral

$$\int_{o < u_k < u_{k+1}} u_k^{\kappa_1 + \cdots + \kappa_k + k - 1} du_k$$

is infinite and in that case, $e^{\Lambda_n(na)} = +\infty$. Now let us relate the satisfaction of (3.9) for $k = n$ for large n, with the following two possibilities.

In the situation described in (i): for large n, (3.9) is satisfied for $k = n$, and the limit of $\frac{1}{n}\Lambda_n(na)$ is finite and coincides with $\Lambda(a)$.

In the situation described in (ii): for large n there exists $1 \le k \le n$ such that (3.9) is not satisfied and the limit of $\frac{1}{n}\Lambda_n(na)$ is infinite and coincides with $\Lambda(a)$.

The last possible situation is: $a(1) \le 1$, and for all t in $[0, 1]$, $\int_{1-t}^{1} a(u)du \le t$, and moreover the equality holds for at least one t or $a(1) = 1$. In that case we do not know the limit but it does not matter.

To prove this, we need two technical lemmas:

Lemma 3.6 *Let b be a continuous function on $[0, 1]$ and b_{jn} be some coefficients such that*

$$\lim_{n\to\infty} \max_{j \le n} \left| b_{jn} - b\left(\frac{j}{n}\right) \right| = 0. \tag{3.10}$$

a) Suppose that $1 - b(t) > 0$ for all t. Then for large n, $b_{jn} < 1, 1 \le j \le n$ and

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=1}^{n} \log(1 - b_{jn}) = \int_0^1 \log(1 - b(t))dt$$

b) Suppose that there is some y such that: $1 - b(t) < 0$. Then for large n, $b_{jn} > 1$ for some $j \le n$.

Lemma 3.7 *If $a : [0,1] \to \mathbb{R}$ is a continuous function and b is the continuous function given by*

$$b(t) = \begin{cases} \frac{1}{t}\int_0^t a(1-u)du & \text{for } 0 < t \le 1 \\ a(1) & \text{for } t = 0, \end{cases}$$

then the coefficients $b_{jn} = \frac{1}{j}\sum_{l=1}^{j} a(\frac{n-l+1}{n})$ satisfy (3.10) of Lemma 3.6.

Now observe that condition (3.9) is just

$$\frac{1}{j}\sum_{l=1}^{j} a\left(\frac{n-l+1}{n}\right) < 1, \forall 1 \le j \le k$$

so that combining the two lemmas we treat the situations (i) and (ii).

□

Upper bound. We get the upper bound with $\overline{\Lambda}^*$ as rate function, using Theorem 4.5.3 b) of Dembo and Zeitouni [5] and the exponential tightness of (ν_n) which is proved as follows: for a a continuous function on $[0,1]$ with supremum norm bounded by 1,

$$\left|\int_0^1 a(t)d\nu_n(t)\right| = \left|\frac{1}{n}\sum_{i=1}^{n} X_{(i)} a\left(\frac{i}{n}\right)\right| \le \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Denote by $\psi^*(t) = t - 1 - \log t$ the Cramer transform of the exponential law with parameter 1. Hence for any positive α

$$\limsup \frac{1}{n}\log P\left(\left\{\sup_{a\in C([0,1]),\|a\|_\infty=1}\left|\int_0^1 a(t)d\nu_n(t)\right| > \alpha\right\}\right) \le \limsup \frac{1}{n}\log P\left(\left\{\frac{1}{n}\sum_{i=1}^{n} X_i > \alpha\right\}\right)$$
$$\le -\psi^*(\alpha)$$

Therefore the limit is $-\infty$ when $\alpha \to \infty$ and this proves the exponential tension of ν_n.

Lower bound. In order to use the same ideas as in Gamboa et al. [6], we find exposed points of Λ^* and prove that they are dense in $\mathcal{M}([0,1])$. This is done in Lemmas 3.8 and 3.9 below.

Lemma 3.8 *Let $a \in \mathcal{C}([0,1])$ be a function satisfying that for every $t > 0$, $\int_{1-t}^1 a(u)du < t$ and set*

$$\alpha(t) = \frac{1}{1 - t - \int_t^1 a(u)du}.$$

Then the measure $\mu \ll \lambda$ defined by its density $s \mapsto \int_0^s \alpha(t)dt$ is an exposed point of Λ^* with exposing hyperplane a.

Proof: Let $\mu' \neq \mu$ be a measure such that $\Lambda^*(\mu') < +\infty$. By Theorem 3.1, $\mu' = l\lambda + \mu'(\{1\})\delta_1$ and $l(s) = \int_0^s \beta(t)dt + d\chi(t)$. Moreover, the measure $-td\chi(1-t) + \mu'(\{1\})\delta_0$ is nonnegative. We have to prove the inequality:

$$\Lambda^*(\mu) - \Lambda^*(\mu') < \langle a, \mu - \mu'\rangle.$$

Let γ be the strictly convex function defined for $x > 0$ by

$$\gamma(x) = x - 1 - \log x.$$

Because of the strict convexity, for $x > 0$, $y > 0$ such that $x \neq y$,

$$\gamma(x) - \gamma(y) < (x-y)\gamma'(x).$$

Use this to bound

$$\Lambda^*(\mu) - \Lambda^*(\mu') = \int_0^1 \left(\gamma\left(t\alpha(1-t)\right) - \gamma\left(t\beta(1-t)\right)\right)dt + \int_0^1 td\chi(1-t) - \mu'\left(\{1\}\right)$$
$$< \int_0^1 \left(\left(\alpha(1-t) - \beta(1-t)\right)\int_{1-t}^1 a(u)du\right)dt + \int_0^1 td\chi(1-t) - \mu'\left(\{1\}\right).$$

An integration by parts leads to:

$$\int_0^1 \left((\alpha(1-t) - \beta(1-t)) \int_{1-t}^1 a(u)du \right) dt = \int_0^1 \left(\int_0^t (\alpha(u) - \beta(u))\, du \right) a(t)dt$$
$$= \langle a, \mu \rangle - \int_0^1 \left(\int_0^t \beta(u)du \right) a(t)dt \tag{3.11}$$

By the hypothesis on a: the function defined by $t \mapsto \frac{\int_{1-t}^1 a(u)du}{t}$ on $(0,1]$ and continuously extended by $a(1)$ at $t=0$ is always less than or equal to 1 on $[0,1]$. So by nonnegativity of the measure $-td\chi(1-t) + \mu(\{1\})\delta_0$,

$$-\int_0^1 \left(\int_{1-t}^1 a(u)du \right) d\chi(1-t) + \mu'(\{1\})a(1) \leq -\int_0^1 td\chi(1-t) + \mu'(\{1\}). \tag{3.12}$$

Combining (3.11) and (3.12), another integration by parts leads to:

$$\Lambda^*(\mu) - \Lambda^*(\mu') < \langle a, \mu \rangle - \int_0^1 \left(\int_0^t \beta(u)du \right) a(t)dt - \int_0^1 \left(\int_0^t d\chi(u) \right) a(t)dt - \mu'\left(\{1\}\right) a(1)$$
$$= \langle a, \mu \rangle - \langle a, \mu' \rangle.$$

□

The proof of the density of the exposed point concludes the demonstration as in Gamboa et al. [6].

Lemma 3.9 *Let μ be in $\mathcal{M}([0,1])$ such that $\Lambda^*(\mu) < +\infty$. Then there exists a sequence of measures μ_n which are exposed points for Λ^*, such that μ_n converges to μ in $\mathcal{M}([0,1])$ and $\lim_{n\to+\infty} \Lambda^*(\mu_n) = \Lambda^*(\mu)$.*

Proof: This proof is very similar to the proof of Lemma 8 in Gamboa et al. [6]. The following property of γ will be very useful:

$$\gamma(\tau + \tau') \leq \gamma(\tau) + \tau', \text{for } \tau > 0, \tau' \geq 0. \tag{3.13}$$

To begin with, we prove an additivity property of the set of exposed points owing to another parametrization than in Lemma 3.8. The application

$$\mathcal{A} \to \mathcal{C}([0,1])$$
$$a \mapsto c := 1 - Ta : t \mapsto 1 - \frac{1}{t}\int_{1-t}^1 a(u)du$$

has image $\mathcal{C} = \{c \in \mathcal{C}([0,1]) : c > 0 \text{ and derivable on } (0,1]\}$. The exposed point corresponding to $c \in \mathcal{C}$ is μ with density (w.r.t. λ) $s \mapsto \int_{1-s}^1 \frac{1}{tc(t)}dt$ and exposing hyperplane: $T^{-1}(1-c)$. It is straightforward to prove that if μ_i, parametrized by $c_i \in \mathcal{C}$, $i = 1, 2$ are two exposed points, then $\mu_1 + \mu_2$ is also an exposed point, parametrized by $\frac{c_1 c_2}{c_1 + c_2}$. Indeed, $\mu_1 + \mu_2$ has density w.r.t. λ:

$$s \mapsto \int_{1-s}^1 \left(\frac{1}{tc_1(t)} + \frac{1}{tc_2(t)} \right) dt = \int_{1-s}^1 \frac{1}{t\frac{c_1(t)c_2(t)}{c_1(t)+c_2(t)}} dt.$$

It is easy to see that the function $c : t \mapsto \frac{c_1(t)c_2(t)}{c_1(t)+c_2(t)}$ is also in $\mathcal{C}$, so it parametrizes an exposed point.

We now use this additivity property of the exposed points to prove their density in the set of measures μ with $\Lambda^*(\mu) < \infty$.

Step 1. We find a sequence of functions $f_M(u) = \int_{1-u}^1 \frac{1}{tc_M(t)}dt, M \in \mathbb{N}$ with $c_M \in \mathcal{C}$, such that $f_M \lambda \to \delta_1$ in $\mathcal{M}([0,1])$ and hence: the measure with density $l_M(u) = l(u) + \mu(\{1\})f_M(u)$ is an exposed point for Λ^* by the additivity property and converges to μ. Moreover, this sequence satisfies: $\Lambda^*(l_M\lambda) \to \Lambda^*(\mu)$.

A construction of (f_M) can be as follows: $c_M(u) = \frac{1}{uf'_M(1-u)}$ has to be > 0, continuous on $[0,1]$ and derivable on $(0,1]$, so f_M will have to be at least twice derivable on $[0,1)$, increasing and can be chosen in such a way that $f'_M(1-u) \sim \frac{1}{u}$ when $u \to 0$. We construct a $\mathcal{C}^2([0,1))$ strictly increasing function f_M with the pattern:

$$\begin{aligned}
\text{on } [0, 1 - \frac{2}{M^2} - \frac{1}{M}] &: f_M \leq \frac{1}{M}, \\
\text{on } [1 - \frac{2}{M^2} - \frac{1}{M}, 1 - \frac{1}{M^2} - \frac{1}{M}] &: \frac{1}{M} \leq f_M \leq M-1, \\
\text{on } [1 - \frac{1}{M^2} - \frac{1}{M}, 1 - \frac{1}{M^2}] &: M - 1 \leq f_M \leq M, \\
\text{on } [1 - \frac{1}{M^2}, 1) &: f_M(x) = M\frac{\log(1-x)}{\log(M^2)}.
\end{aligned}$$

f_M approximates δ_1, because the principal part is on the third interval. Hence: $l_M\lambda = l\lambda + \mu(\{1\})f_M\lambda \to \mu = l\lambda + \mu(\{1\})\delta_1$. Let us now prove that $\Lambda^*(l_M\lambda) \to \Lambda^*(\mu)$. Because Λ^* is lower semicontinuous, $\liminf \Lambda^*(l_M\lambda) \geq \Lambda^*(\mu)$. The other inequality is derived as follows:

$$\begin{aligned}
\Lambda^*(l_M\lambda) &= \Lambda^*\left((l + \mu(\{1\})f_M)\lambda\right) \\
&= \int_0^1 \gamma(t\alpha(1-t) + t\alpha_M(1-t))dt,
\end{aligned}$$

where $\alpha(1-t) = \frac{1}{tc(t)}$ and $\alpha_M(t) = \mu(\{1\})f'_M(t)$. By inequality (3.13), since for all t, $t\alpha(1-t) > 0$:

$$\Lambda^*(l_M\lambda) \leq \int_0^1 \left(\gamma(t\alpha(1-t)) + t\alpha_M(1-t)\right)dt.$$

But, by integration by parts:

$$\begin{aligned}
\int_0^1 t\alpha_M(1-t)dt &= \mu(\{1\})\int_0^1 tf'_M(1-t)dt \\
&= \mu(\{1\})\int_0^1 f_M(1-t)dt - \mu(\{1\})f_M(0)
\end{aligned}$$

because by choice of f_M in this construction, $\lim_{t\to 0} tf_M(1-t) = 0$. Notice that $0 \leq f_M(0) \leq \frac{1}{M}$, so $\lim_{M\to+\infty} \mu(\{1\})f_M(0) = 0$. Now as $f_M\lambda$ approximates the measure δ_1, the last expression tends to $\mu(\{1\})$. Therefore, the desired inequality is proved:

$$\limsup \Lambda^*(l_M\lambda) \leq \int_0^1 \gamma(t\alpha(1-t))dt + \mu(\{1\}) = \Lambda^*(\mu)$$

Step 2. Let $\mu = l\lambda$ be such that $\Lambda^*(\mu) < +\infty$, with $l(u) = \int_0^u dm(s)$, $m = \alpha\lambda + \chi$ and suppose that the measure with density $s \mapsto \int_0^s \alpha(t)dt$ is an exposed point of Λ^* as in Lemma 3.8.

There exists a sequence $(c_M) \subset \mathcal{C}$ such that $c_M > 0$ and $\frac{1}{c_M}\lambda$ tends to the positive measure $-td\chi(1-t)$. Define $\alpha_M(1-t) = \frac{1}{tc_M(t)}$ and μ_M the measure with density l_M with respect to λ, where $l_M : s \mapsto \int_0^s (\alpha(t) + \alpha_M(t))\,dt$. Then μ_M tends to μ: for $b \in \mathcal{C}([0,1])$, we only need to check that

$$\int_0^1 \left(\int_0^t \alpha_M(s)ds\right) b(t)dt \to \int_0^1 \int_0^t d\chi(s)b(t)dt.$$

But

$$\begin{aligned}
\int_0^1 \left(\int_0^t \alpha_M(s)ds\right) b(t)dt &= \int_0^1 \left(\int_{1-t}^1 \frac{1}{sc_M(s)}ds\right) b(t)dt \\
&= \int_0^1 \frac{\int_{1-s}^1 b(t)dt}{s} \frac{1}{c_M(s)}ds \qquad (3.14)
\end{aligned}$$

by Fubini's theorem. This operation is possible because $\|c_M\|_\infty > 0$ and $\int_0^1 \frac{\int_{1-s}^1 |b(t)|dt}{s} \frac{1}{|c_M(s)|} ds \leq \frac{\|b\|_\infty}{\|c_M\|_\infty}$. But $s \mapsto \frac{\int_{1-s}^1 b(t)dt}{s}$ is a continuous function on $[0,1]$, and $\frac{1}{c_M(s)} ds$ is a measure that tends to $-sd\chi(1-s)$. Hence: the limit of the right-hand side of (3.14) is:

$$\int_0^1 \frac{\int_{1-s}^1 b(t)dt}{s} \left(-sd\chi(1-s)\right) = -\int_0^1 \left(\int_{1-s}^1 b(t)dt\right) d\chi(1-s).$$

Note that as the measure $-sd\chi(1-s)$ does not give mass to $\{0\}$, and moreover $\int_{1-s}^1 b(t)dt = 0$ when $s = 0$, the simplification by s is allowed. Lastly, a change of variables $1-s = u$ and the use of Fubini's theorem lead to:

$$\begin{aligned} -\int_0^1 \left(\int_{1-s}^1 b(t)dt\right) d\chi(1-s) &= \int_0^1 \left(\int_u^1 b(t)dt\right) d\chi(u) \\ &= \int_0^1 \left(\int_0^t d\chi(u)\right) b(t)dt. \end{aligned}$$

We can now prove that $\Lambda^*(\mu_M) \to \Lambda^*(\mu)$: as Λ^* is lower semicontinuous, $\liminf \Lambda^*(\mu_M) \geq \Lambda^*(\mu)$. For the converse, use inequality (3.13):

$$\begin{aligned} \Lambda^*(\mu_M) &= \int_0^1 \gamma\left(t\left(\alpha(1-t) + \alpha_M(1-t)\right)\right) dt \\ &\leq \int_0^1 \gamma(t\alpha(1-t))dt + \int_0^1 t\alpha_M(1-t)dt \text{ because } t\alpha(1-t) > 0. \\ \limsup \Lambda^*(\mu_M) &\leq \int_0^1 \gamma(t\alpha(1-t)) + \limsup \int_0^1 t\alpha_M(1-t)dt \\ &\leq \int_0^1 \gamma(t\alpha(1-t)) + \limsup \int_0^1 \frac{1}{c_M(t)} dt \\ &= \int_0^1 \gamma(t\alpha(1-t)) - \int_0^1 td\chi(1-t) \\ &= \Lambda^*(\mu). \end{aligned}$$

Step 3. Let $\mu = l\lambda$ be such that $\Lambda^*(\mu) < +\infty$, with $l(u) = \int_0^u dm(s)$, and $m = \alpha\lambda$. Remark that $\Lambda^*(\mu) < +\infty$ implies that $s \mapsto \frac{1}{c(s)} := s\alpha(1-s)$ is ≥ 0 λ-a.s. and is in $L^1([0,1])$. Suppose moreover that there exists $\epsilon > 0$ such that:

$$\frac{1}{c} \geq \epsilon \ \lambda\text{-a.s.} \tag{3.15}$$

Then let $(c_M) \subset \mathcal{C}$ be such that $\forall s, \frac{1}{c_M(s)} > \frac{\epsilon}{2}$ and $\frac{1}{c_M} \to \frac{1}{c}$ in $L^1([0,1])$. Call $\alpha_M(s) = \frac{1}{(1-s)c_M(1-s)}$. Then the measure μ_M with density $l_M : s \mapsto \int_0^s \alpha_M(t)dt$ converges to μ and $\Lambda^*(\mu_M) \to \Lambda^*(\mu)$.

Indeed, for $b \in \mathcal{C}([0,1])$, we have $\int_0^1 \|b\|_\infty s\alpha_M(1-s)ds < +\infty$ so by Fubini's theorem:

$$\begin{aligned}
\int_0^1 b(t)d\mu_M(t) &= \int_0^1 b(t)\int_0^t \alpha_M(s)dsdt \\
&= \int_0^1 \left(\int_{1-s}^1 b(t)dt\right)\alpha_M(1-s)ds \\
&= \int_0^1 \frac{\int_{1-s}^1 b(t)dt}{s} s\alpha_M(1-s)ds \\
&= \int_0^1 \frac{\int_{1-s}^1 b(t)dt}{s}\frac{1}{c_M(s)}ds \\
&\to \int_0^1 \frac{\int_{1-s}^1 b(t)dt}{s}\frac{1}{c(s)}ds \\
&= \int_0^1 \left(\int_{1-s}^1 b(t)dt\right)\alpha(1-s)ds \\
&= \int_0^1 b(t)d\mu(t).
\end{aligned}$$

By lower semicontinuity of Λ^*, $\liminf \Lambda^*(\mu_M) \geq \Lambda^*(\mu)$. For the converse inequality: use that the strict convexity of γ implies:

$$\begin{aligned}
|\Lambda^*(\mu_M) - \Lambda^*(\mu)| &= \left|\int_0^1 \gamma(t\alpha_M(1-t) - t\alpha(1-t))dt\right| \\
&\leq \left|\int_0^1 |t(\alpha_M(1-t) - \alpha(1-t))| \left|1 - \frac{1}{t\alpha_M(1-t)}\right| dt\right| \\
&\leq \left(1 + \frac{2}{\epsilon}\right)\left\|\frac{1}{c_M} - \frac{1}{c}\right\|_{L^1([0,1])} \\
&\to 0.
\end{aligned}$$

Step 4. Suppose that μ is as in Step 3 but that (3.15) is not assumed any more. Define α_ϵ such that: $t\alpha_\epsilon(1-t) = t\alpha(1-t)1_{t\alpha(1-t)>\epsilon} + \epsilon 1_{t\alpha(1-t)\leq\epsilon}$ and μ_ϵ with density $u \mapsto \int_0^u \alpha_\epsilon(s)ds$. As $|t\alpha_\epsilon(1-t) - t\alpha(1-t)| \leq \epsilon$ for all t, by computations already made in step 3, $\mu_\epsilon \to \mu$. Now prove that $\Lambda^*(\mu_\epsilon) \to \Lambda^*(\mu)$:

$$\int_0^1 (\gamma(t\alpha_\epsilon(1-t)) - \gamma(t\alpha(1-t)))\,dt \leq \int_0^1 (t\alpha_\epsilon(1-t) - t\alpha(1-t))\left(1 - \frac{1}{t\alpha_\epsilon 1 - t)}\right)dt.$$

But the absolute value of this last quantity can be bounded by

$$\int_0^1 \epsilon 1_{t\alpha(1-t)\leq\epsilon}\left(1 + \frac{1}{\epsilon}\right) = (\epsilon + 1)\lambda\left(\{t : t\alpha(1-t) \leq \epsilon\}\right).$$

And the measurable function $t \mapsto t\alpha(1-t)$ is > 0 λ-almost surely, so $\lambda\left(\{t : t\alpha(1-t) \leq \epsilon\}\right) \to 0$, as $\epsilon \to 0$. So:

$$\limsup \int_0^1 \gamma(t\alpha_\epsilon(1-t))dt \leq \int_0^1 \gamma(t\alpha(1-t))dt,$$

which proves the required inequality. To conclude this step, approximate μ_ϵ by μ_M as in Step 3.

Step 5. Suppose now that μ is any measure such that $\Lambda^*(\mu) < +\infty$. Combine Steps 1, 2 and 4 and use the inequality

$$\gamma(\tau_1 + \tau_2 + \tau_3) \leq \gamma(\tau_1) + \tau_2 + \tau_3 \text{ for } \tau_1 > 0, \tau_2, \tau_3 \geq 0.$$

□

□

4 Appendix

Here we recall some basic facts of large deviations theory. For further results, we refer to Dembo and Zeitouni [5].

Définition 4.1 *Let $\mathcal{X}$ be a Hausdorff space with Borel σ-algebra $\mathcal{B}(\mathcal{X})$. $I : X \to \mathbb{R}_+ \cup \{+\infty\}$ is a lower semi-continuous function. We say that a sequence (R_n) of probability measures on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ satisfies a Large Deviations Principle (LDP) with rate function I if for any $A \in \mathcal{B}(\mathcal{X})$,*

$$-\inf_{x \in \mathring{A}} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \log R_n(A) \leq \limsup_{n\to\infty} \frac{1}{n} \log R_n(A) \leq -\inf_{x \in clo(A)} I(x).$$

Définition 4.2 *The rate function I is good if for all α, the level set $\{x : I(x) \leq \alpha\}$ is a compact set.*

Proposition 4.3 ***(Contraction principle)***
Let $\mathcal{X}$ and $\mathcal{Y}$ be two Hausdorff spaces, and $f : \mathcal{X} \to \mathcal{Y}$ be a continuous function. Suppose that (R_n) satisfies a LDP on $\mathcal{X}$ with good rate function I. Then the sequence of probability measures $\left(R_n \circ f^{-1}\right)$ satisfies a LDP on $\mathcal{Y}$ with good rate function I' defined for $y \in \mathcal{Y}$ by:

$$I'(y) = \inf\{I(x) : x \in \mathcal{X}, f(x) = y\}.$$

Proposition 4.4 ***(Exponential equivalence)***
Assume that $\mathcal{X}$ is a metric space, with distance denoted by d. Let ζ_n and ξ_n be two $\mathcal{X}$-valued r.v.'s. They are called exponentially equivalent if for all $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(d\left(\zeta_n, \xi_n\right) > \delta\right) = -\infty.$$

In that case, if (ζ_n) satisfies a LDP with good rate function, then the same LDP holds for (ξ_n).

Acknowledgements. The author is very grateful to Eustasio del Barrio for pointing the problem of Large Deviations for L-statistics. The central idea of using the topology of Wasserstein distance is also due to him. The author also wishes to thank Fabrice Gamboa for helpful discussions and for many of the ideas present in this paper.

References

[1] A. Araujo and E. Giné. *The central limit theorem for real and Banach valued random variables*. John Wiley & Sons, New York-Chichester-Brisbane, 1980. Wiley Series in Probability and Mathematical Statistics.

[2] J. M. Borwein and A. S. Lewis. Partially-finite programming in L_1 and the existence of maximum entropy estimates. *SIAM J. Optim.*, 3(2):248–267, 1993.

[3] R. B. D'Agostino. An omnibus test of normality for moderate and large size samples. *Biometrika*, 58:341–348, 1971.

[4] E. del Barrio, J. A. Cuesta-Albertos, and C. Matrán. Contributions of empirical and quantile processes to the asymptotic theory of goodness-of-fit tests. *Test*, 9(1):1–96, 2000. With discussion.

[5] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.

[6] F. Gamboa, A. Rouault, and M. Zani. A functional large deviations principle for quadratic forms of Gaussian stationary processes. *Statist. Probab. Lett.*, 43(3):299–308, 1999.

[7] P. Groeneboom. *Large deviations and asymptotic efficiencies*, volume 118 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam, 1980.

[8] P. Groeneboom, J. Oosterhoff, and F. H. Ruymgaart. Large deviation theorems for empirical probability measures. *Ann. Probab.*, 7(4):553–586, 1979.

[9] P. Groeneboom and G. R. Shorack. Large deviations of goodness of fit statistics and linear combinations of order statistics. *Ann. Probab.*, 9(6):971–987, 1981.

[10] R. Helmers. A Berry-Esseen theorem for linear combinations of order statistics. *Ann. Probab.*, 9(2):342–347, 1981.

[11] D. Li, M. B. Rao, and R. J. Tomkins. The law of the iterated logarithm and central limit theorem for L-statistics. *J. Multivariate Anal.*, 78(2):191–217, 2001.

[12] D. M. Mason and G. R. Shorack. Necessary and sufficient conditions for asymptotic normality of L-statistics. *Ann. Probab.*, 20(4):1779–1804, 1992.

[13] A. W. Roberts and D. E. Varberg. *Convex functions*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1973. Pure and Applied Mathematics, Vol. 57.

[14] R. T. Rockafellar. Integrals which are convex functionals. II. *Pacific J. Math.*, 39:439–469, 1971.

[15] R. T. Rockafellar. *Conjugate duality and optimization*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974. Lectures given at the Johns Hopkins University, Baltimore, Md., June, 1973, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 16.

[16] G. R. Shorack and J. A. Wellner. *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[17] S. M. Stigler. Linear functions of order statistics with smooth weight functions. *Ann. Statist.*, 2:676–693, 1974.

[18] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.

[19] W. R. van Zwet. A strong law for linear functions of order statistics. *Ann. Probab.*, 8(5):986–990, 1980.

[20] M. Vandemaele and N. Veraverbeke. Cramér type large deviations for linear combinations of order statistics. *Ann. Probab.*, 10(2):423–434, 1982.

[21] J. A. Wellner. A Glivenko-Cantelli theorem and strong laws of large numbers for functions of order statistics. *Ann. Statist.*, 5(3):473–480, 1977.

[22] J. A. Wellner. Correction to: "A Glivenko-Cantelli theorem and strong laws of large numbers for functions of order statistics" (Ann. Statist. **5** (1977), no. 3, 473–480). *Ann. Statist.*, 6(6):1394, 1978.

This work was performed while the author was a PhD student at the Departamento EIO, Universidad de Valladolid (Spain) and at the Institut de Mathématiques, Université Paul Sabatier, Toulouse (France).

Present address:
Hélène Boistard
Laboratoire MODAL'X
Bât. G, Bureau E08
Université Paris 10
200 avenue de la République
92001 Nanterre cedex, France
helene@boistard.fr