How to estimate a cumulative process's rate-function

Ken Duffy[(1)] and Anthony P. Metcalfe[(2)]

(1) Hamilton Institute, National University of Ireland, Maynooth, Ireland

(2) Boole Centre for Research in Informatics, University College Cork, Co. Cork, Ireland

22^{nd} June, 2004

Abstract

Consider two sequences of bounded random variables, a value and a timing process, that satisfy the large deviation principle (LDP) with rate-function $J(\cdot\,,\cdot)$ and whose cumulative process satisfies the LDP with rate-function $I(\cdot)$. Under mixing conditions, an LDP for estimates of I constructed by transforming an estimate of J is proved. For the case of cumulative renewal processes it is demonstrated that this approach is favorable to a more direct method as it ensures the laws of the estimates converge weakly to a Dirac measure at I.

1 Introduction, setup and estimation schemes.

Let $\{(X_n, \tau_n), n \in \mathbb{N}\}$ be a stationary process of (not necessarily independent) bounded random random variables taking values in $[a, b] \times [\alpha, \beta] \subset \mathbb{R} \times (0, \infty)$. Defining $S_n := X_1 + \cdots + X_n$ and $T_n := \tau_1 + \cdots + \tau_n$, assume $\{(S_n, T_n)/n, n \in \mathbb{N}\}$ satisfies the Large Deviation Principle (LDP) on the scale $1/n$ with rate-function J that is the Legendre-Fenchel transform of its scaled Cumulant Generating Function (sCGF)

$$J(x_1, x_2) = \sup_{(\theta_1,\theta_2)\in\mathbb{R}^2} (\theta_1 x_1 + \theta_2 x_2 - M(\theta_1, \theta_2)), \qquad (1)$$

where

$$M(\theta_1, \theta_2) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[\exp(\theta_1 S_n + \theta_2 T_n)].$$

[1]Estimating Large Deviations, Cumulative Process, Renewal Process

[2]American Mathematical Society 2000 subject classifications: Primary 60F10; Secondary 60K05

Treating τ_i as the inter-arrival time between volumes X_{i-1} and X_i, the (pure, zero-delayed) cumulative process $\{A_t, t \in \mathbb{R}^+\}$ is defined by

$$A_t := \sum_{i=1}^{N(t)} X_i,$$

where $N(t) := \sup\{n : T_n \leq t\}$ is the counting process associated with $\{T_n, n \in \mathbb{N}\}$ and the empty sum with $N(t) = 0$ is defined to have value zero. Cumulative processes arise naturally in many applications, including queueing theory and risk theory. Assume $\{A_t/t, t \in \mathbb{R}^+\}$ satisfies the LDP on the scale $1/t$ with rate-function I that is the Legendre-Fenchel transform of its sCGF

$$I(x) = \sup_{\theta \in \mathbb{R}} (\theta x - \lambda(\theta)), \text{ where } \lambda(\theta) = \lim_{t \to \infty} \frac{1}{t} \log \mathbb{E}[\exp(\theta A_t)]. \tag{2}$$

If we are given an observation $(X_1, \tau_1), (X_2, \tau_2) \ldots$, but the statistics of the process $\{(X_n, \tau_n), n \in \mathbb{N}\}$ are unknown, how would we estimate the rate-function I? The purpose of this work is to consider the large deviations of estimating I by transforming an estimate of the sCGF M. Although a more direct approach based on estimating λ and taking its Legendre-Fenchel transform is possible, this scheme performs better in certain, typical circumstances. For instance in the case when $\{A_t, t \in \mathbb{R}^+\}$ is a cumulative renewal process, that is $\{(X_n, \tau_n), n \in \mathbb{N}\}$ forms an i.i.d. sequence. This is demonstrated in section 3.

The transformation is based on the following observation. For a broad class of processes it is known that J and I from equations (1) and (2) are related by

$$I(x) = \inf_{y>0} yJ\left(\frac{x}{y}, \frac{1}{y}\right) \tag{3}$$

(see [19, 26, 11]). Having estimated J, our estimate of I is constructed by the transforming the estimates through equation (3).

In order to discuss the LDP for estimates of I and J we must select topological spaces for I, J and their estimates. The spaces of lower semi-continuous (lsc) convex functions from $[a/\beta, b/\alpha]$ and $[a, b] \times [\alpha, \beta]$, respectively, taking values in $\mathbb{R} \cup \{+\infty\}$ and denoted $\text{Conv}[a/\beta, b/\alpha]$, $\text{Conv}([a, b] \times [\alpha, \beta])$, are natural. However, the usual topologies, such as uniform convergence on bounded subsets, are not suitable. For example we want a topology in which estimates such as $I_n = n|x|$ converge to I which is 0 at $x = 0$ and $+\infty$ otherwise.

The topology we employ is the Attouch-Wets topology [1, 2], a metrizable topology which identifies the convergence of functions with the convergence of their epigraphs. For an extended real-valued function $f : Z \mapsto \bar{\mathbb{R}}$ on a metric space (Z, d), the epigraph of f is defined to be $\text{epi } f := \{(z, \alpha) \in Z \times \mathbb{R} : \alpha \geq f(z)\}$. Equipping $Z \times \mathbb{R}$ with the box metric D, f is lsc if and only if $\text{epi } f$ is a closed subset of $Z \times \mathbb{R}$ (for example, see Theorem 1.6 of [25]). Therefore a topology for lsc functions is inherited from a topology on the closed subsets of the metric space.

Let $\mathrm{CL}(Z)$ denote the collection of nonempty closed subsets of Z. A sequence $\{A_n, n \in \mathbb{N}\} \subset \mathrm{CL}(Z)$ converges in the Attouch-Wets topology to $A \in \mathrm{CL}(Z)$, denoted $A = \tau_{\mathrm{AW}d} - \lim A_n$, if given any bounded set $B \subset Z$ and any $\epsilon > 0$, there exists N_ϵ such that

$$\sup_{z \in B} |d(z, A_n) - d(z, A)| < \epsilon \text{ for all } n > N_\epsilon.$$

Letting $\mathrm{Conv}(Z)$ denote the set of convex lsc functions from Z to $\mathbb{R} \cup \{+\infty\}$, a sequence $\{f_n, n \in \mathbb{N}\} \subset \mathrm{Conv}(Z)$ converges to $f \in \mathrm{Conv}(Z)$ if and only if $\{\mathrm{epi}\, f_n, n \in \mathbb{N}\}$ converges to $\mathrm{epi}\, f$ in the Attouch-Wets topology on $\mathrm{CL}(Z \times \mathbb{R})$, i.e., $f = \tau_{\mathrm{AW}D} - \lim f_n$ if and only if $\mathrm{epi}\, f = \tau_{\mathrm{AW}D} - \lim \mathrm{epi}\, f_n$.

Whenever Z is a subset of $\mathbb{R}^n$ (for any finite n) equipped with the Euclidean distance, d, Theorem 3.3.3 of Beer [3] can be used to show that the topologies $\tau_{\mathrm{AW}D}$ and $\tau_{\mathrm{AW}d}$ on $\mathrm{CL}(Z \times \mathbb{R})$ coincide. Thus either can be used to form the Attouch-Wets topology on $\mathrm{Conv}(Z)$, and we denote this topology simply by τ_{AW} from now on.

A necessary and sufficient condition for checking convergence in $\mathrm{CL}(Z)$ is the following, which we shall use in our proofs:

Theorem 1 (Theorem 3.1.7 of Beer [3]) *Let (Z, d) be a metric space, and let $A, A_1, A_2, \ldots$ be a sequence in $\mathrm{CL}(Z)$. Then $A = \tau_{\mathrm{AW}d} - \lim A_n$ if and only if for each bounded subset B of Z we have both $\lim_{n\to\infty} e_d(A \cap B, A_n) = 0$ and $\lim_{n\to\infty} e_d(A_n \cap B, A) = 0$, where for any two nonempty subsets E and F, $e_d(E, F)$, the excess of E over F is given by $e_d(E, F) = \sup\{d(e, F) : e \in E\}$.*

A good reference for the Attouch-Wets topology is Beer [3]. Its suitability for our purposes will be demonstrated by the continuity of the transformation in equation (3).

The sCGF estimation scheme we adopt for M was proposed by A. Dembo in a private communication to Duffield et al. [8]. The scheme is this: select a block-length $B \in \mathbb{N}$ sufficiently large that you believe the blocked sequence $\{\vec{Y}_n, n \in \mathbb{N}\}$, where $\vec{Y}_n := (X_{(n-1)B+1}, \tau_{(n-1)B+1}) + \cdots + (X_{nB}, \tau_{nB})$, can be treated as approximately i.i.d; then use the estimator:

$$M_n(\vec{\theta}) = \frac{1}{B} \log \frac{1}{n} \sum_{i=1}^{n} \exp(\langle \vec{\theta}, \vec{Y}_i \rangle). \tag{4}$$

We propose taking the Legendre-Fenchel transform of M_n to form an estimate J_n of J,

$$J_n(\vec{x}) = \sup_{\vec{\theta}} (\langle \vec{\theta}, \vec{x} \rangle - M_n(\vec{\theta})), \tag{5}$$

and then transform J_n by equation (3) to form an estimate I_n of I,

$$I_n(x) = \inf_{y>0} y J_n\left(\frac{x}{y}, \frac{1}{y}\right). \tag{6}$$

In section 2, Theorem 7, the LDP for $\{I_n, n \in \mathbb{N}\}$ in $\mathrm{Conv}[a/\beta, b/\alpha]$ is established via the contraction principle. In section 3 it is explained why this scheme can be preferable to the more direct approach of estimating I through the Legendre-Fenchel transform of estimates of λ. In particular, using this scheme if $\{A_t, t \in \mathbb{R}^+\}$ is a cumulative renewal process, we prove the laws of $\{I_n, n \in \mathbb{N}\}$ converge weakly to a Dirac measure at a I, which is not typically the case using the direct method. In section 4 we present a guide to related work and an indication of the practical significance of the proposed estimation scheme.

2 Main results.

Let $\mathcal{M}_1([Ba, Bb]\times[B\alpha, B\beta])$ denote the space of probability measures on $[Ba, Bb]\times[B\alpha, B\beta]$ equipped with the τ topology. In [4], Bryc and Dembo introduce a mixing condition, condition (S), for stationary processes. For example it is satisfied by hypercontractive Markov chains. If $\{(X_n, \tau_n), n \in \mathbb{N}\}$ satisfies (S) then by inclusion of σ-algebras the blocked process $\{\vec{Y}_n, n \in \mathbb{N}\}$ also does. Under condition (S), Theorem 1 of [4] proves that the empirical laws

$$L_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{\vec{Y}_i}$$

satisfy the LDP in $\mathcal{M}_1([Ba, Bb]\times[B\alpha, B\beta])$ with good, convex rate-function H. The rate-function H can be calculated by through the following representation. For $\nu \in \mathcal{M}_1([Ba, Bb]\times[B\alpha, B\beta])$,

$$H(\nu) = \sup_{f\in M([Ba,Bb]\times[B\alpha,B\beta],\mathbb{R})} \left\{\int f d\nu - \Lambda(f)\right\},$$

where $M([Ba, Bb]\times[B\alpha, B\beta], \mathbb{R})$ is the set of bounded, measurable functions from $[Ba, Bb]\times[B\alpha, B\beta]$ to $\mathbb{R}$ and

$$\Lambda(f) = \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\left[\exp\left(\sum_{i=1}^{n} f(\vec{Y}_i)\right)\right].$$

For each $\nu \in \mathcal{M}_1([Ba, Bb]\times[B\alpha, B\beta])$ define M_ν by

$$M_\nu(\vec{\theta}) = \frac{1}{B}\log\int_{\vec{x}\in[Ba,Bb]\times[B\alpha,B\beta]} e^{\langle\vec{\theta},\vec{x}\rangle}\nu(d\vec{x}),$$

and J_ν to be the Legendre-Fenchel transform of M_ν. Note that when $\nu = L_n$, the empirical law, $J_{L_n} = J_n$ defined in equation (5). For a one dimensional process, assuming condition (S) holds, Duffy and Metcalfe [10] prove that rate-function estimates satisfy the LDP. The arguments in [10] generalize verbatim to processes taking values in $\mathbb{R}^d$, for any finite d, giving us the following theorem.

Theorem 2 (Thm. 1, [10]) *If* $\{(X_n,\tau_n), n \in \mathbb{N}\}$ *satisfies condition (S) of Bryc and Dembo [4], then the estimates* $\{J_n, n \in \mathbb{N}\}$ *satisfy the LDP in* $\mathrm{Conv}([a,b]\times[\alpha,\beta])$ *with good rate-function*

$$K(\phi) = \begin{cases} H(\nu) & \text{if } \phi = J_\nu, \text{ where } \nu \in \mathcal{M}_1([Ba,Bb]\times[B\alpha,B\beta]), \\ +\infty & \text{otherwise.} \end{cases}$$

The arguments presented in [10] used to prove Theorem 2 rely on the assumption that the random variables are bounded. Based on evidence from numerical experiments we conjecture that Theorem 2 also holds for (an appropriate class of) processes with unbounded random variables, but proving this is a mathematically challenging problem. As there are is no constraint on the value of the bound, for practical applications we believe the boundedness assumption posses no serious difficulty.

In light of Theorem 2, in order to prove the LDP for $\{I_n, n \in \mathbb{N}\}$, where I_n is defined in equation (6), it suffices to prove that the transformation (3) from $\mathrm{Conv}([a,b]\times[\alpha,\beta]) \mapsto \mathrm{Conv}[a/\beta, b/\alpha]$ is continuous and invoke the contraction principle (Theorem 4.2.1 of [6]). To prove continuity, first we prove a Proposition and Corollary that provide a checkable condition for the continuity of a function $f : \mathrm{Conv}(Z_1) \mapsto \mathrm{Conv}(Z_2)$, where (Z_1, d_1) and (Z_2, d_2) are metric spaces and both $\mathrm{Conv}(Z_1)$ and $\mathrm{Conv}(Z_2)$ are equipped with the Attouch-Wets topology. The Proposition, which may be of independent interest, establishes conditions on a function $g : Z_2 \mapsto Z_1$ that ensure its pullback g^*, where $g^*(B_1) := \{x \in Z_2 : g(x) \in B_1\}$, is continuous in the Attouch-Wets topology.

Proposition 3 *Let* (Z_1, d_1) *and* (Z_2, d_2) *be metric spaces. Let* $g : Z_2 \mapsto Z_1$ *be a bijective, continuous function that maps bounded sets to bounded sets and whose inverse* g^{-1} *is uniformly continuous on bounded sets. Then* $g^* : \mathrm{CL}(Z_1) \mapsto \mathrm{CL}(Z_2)$ *is continuous.*

Proof: Let $A, A_1, A_2, \dots$ denote a convergent sequence in $\mathrm{CL}(Z_1)$ with $A = \tau_{\mathrm{AW}_{d_1}} - \lim A_n$. Fix $\gamma > 0$, and let B_2 be a bounded subset of Z_2. As g maps bounded subsets to bounded subsets, $g(B_2) := \{g(x) : x \in B_2\} =: B_1 \subset Z_1$ is bounded. Then as $A = \tau_{\mathrm{AW}_{d_1}} - \lim A_n$, there exists an N such that $n > N$ implies that $e_{d_1}(B_1 \cap A_n, A) < \gamma$, so that $d_1(x, A) < \gamma$ for all $x \in B_1 \cap A_n$. Fix $n > N$ and $y \in B_2 \cap g^*(A_n)$. Then $y \in g^*(B_1) \cap g^*(A_n)$, and so $d_1(g(y), A) < \gamma$. Thus, since g is surjective, there must exist $y' \in g^*(A)$ with $d_1(g(y), g(y')) < \gamma$.

Fix $\epsilon > 0$. Since $B_2 \cap g^*(A_n)$ is bounded and g^{-1} is uniformly continuous on bounded subsets we can choose γ sufficiently small that $d_2(y, y') < \epsilon$. Thus, since $y' \in g^*(A)$, we have $d_2(y, g^*(A)) < \epsilon$. This is true for all $y \in B_2 \cap g^*(A_n)$ with $n > N$, and so $e_{d_2}(B_2 \cap g^*(A_n), g^*(A)) < \epsilon$. Thus $\lim_{n\to\infty} e_{d_2}(B_2 \cap g^*(A_n), g^*(A)) = 0$. Similarly $\lim_{n\to\infty} e_{d_2}(B_2 \cap g^*(A), g^*(A_n)) = 0$, and so $g^*(A) = \tau_{\mathrm{AW}_{d_2}} - \lim g^*(A_n)$.

■

In particular, the following Corollary states that if we wish to check the continuity of $f : \mathrm{Conv}(Z_1) \mapsto \mathrm{Conv}(Z_2)$ and $\mathrm{epi}\, f(\psi) = g^*(\mathrm{epi}\, \psi)$, then it suffices to check the conditions on g in Proposition 3.

Corollary 4 *Let* (Z_1, d_1) *and* (Z_2, d_2) *be metric spaces and let* $f : \text{Conv}(Z_1) \mapsto \text{Conv}(Z_2)$ *be a function. Suppose that* $\text{epi } f(\psi) = g^*(\text{epi } \psi)$ *for all* $\psi \in \text{Conv}(Z_1)$*, where* $g : Z_2 \times \mathbb{R} \mapsto Z_1 \times \mathbb{R}$ *is a function that satisfies the conditions of Lemma 3. Then* f *is continuous.*

We prove that the transformation (3) is continuous by considering it as the composition of two maps. The first takes $J(x, y)$ to $yJ(x/y, 1/y)$ and is proved via Corollary 4; the second takes $yJ(x/y, 1/y)$ to $\inf_y yJ(x/y, 1/y)$ and is proved by direct arguments.

Lemma 5 *For any* $\phi \in \text{Conv}([a, b] \times [\alpha, \beta])$*, define its transform* $f_1(\phi)$ *by*

$$(f_1(\phi))(x, y) := y\phi\left(\frac{x}{y}, \frac{1}{y}\right), \text{ for all } (x, y) \in [a/\beta, b/\alpha] \times [1/\beta, 1/\alpha].$$

Then $f_1(\phi) \in \text{Conv}([a/\beta, b/\alpha] \times [1/\beta, 1/\alpha])$ *and* f_1 *is continuous.*

Proof: Let $\phi \in \text{Conv}([a, b] \times [\alpha, \beta])$ and fix $p \in [0, 1]$ and $(x_1, y_1), (x_2, y_2) \in [a, b] \times [\alpha, \beta]$. Consider $(f_1(\phi))(px_1 + (1 - p)x_2, py_1 + (1 - p)y_2)$. Noting that

$$\left(\frac{px_1 + (1-p)x_2}{py_1 + (1-p)y_2}, \frac{1}{py_1 + (1-p)y_2}\right) = q\left(\frac{x_1}{y_1}, \frac{1}{y_1}\right) + (1-q)\left(\frac{x_2}{y_2}, \frac{1}{y_2}\right),$$

and $q = py_1/(py_1 + (1 - p)y_2) \in [0, 1]$ as $y_1, y_2 > 0$, the convexity of $f_1(\phi)$ follows from the convexity of ϕ. To show that $f_1(\phi)$ is lsc note that $\text{epi } f_1(\phi) = g^*(\text{epi } \phi)$, where $g : [a/\beta, b/\alpha] \times [1/\beta, 1/\alpha] \times \mathbb{R} \mapsto [a, b] \times [\alpha, \beta] \times \mathbb{R}$ is given by $g(x, y, z) = (x/y, 1/y, z/y)$. Since $\text{epi } \phi$ is closed and g is continuous, $\text{epi } f_1(\phi)$ is closed. Thus $f_1(\phi)$ is lsc and $f_1(\phi) \in \text{Conv}([a/\beta, b/\alpha] \times [1/\beta, 1/\alpha])$.

The function g is bijective with inverse $g^{-1} : [a, b] \times [\alpha, \beta] \times \mathbb{R} \mapsto [a/\beta, b/\alpha] \times [1/\beta, 1/\alpha] \times \mathbb{R}$ given by $g^{-1}(x, y, z) = (x/y, 1/y, z/y)$. Both g and g^{-1} are uniformly continuous on bounded sets. Therefore g satisfies the conditions of Proposition 3 and thus Corollary 4 implies that f_1 is continuous.

■

Lemma 6 *For any function* $\phi \in \text{Conv}([a/\beta, b/\alpha] \times [1/\beta, 1/\alpha])$*, define its transform* $f_2(\phi)$ *by*

$$(f_2(\phi))(x) := \inf_{y \in [1/\beta, 1/\alpha]} \phi(x, y), \text{ for all } x.$$

Then $f_2(\phi) \in \text{Conv}([a/\beta, b/\alpha])$ *and* f_2 *is continuous.*

Proof: As $\phi \in \text{Conv}([a/\beta, b/\alpha] \times [1/\beta, 1/\alpha])$, it's global infimum is attained which ensures the conditions of Corollary 3.32 of Rockafellar and Wets [25] hold and $f_2(\phi) \in \text{Conv}([a/\beta, b/\alpha])$.

As ϕ is convex and lsc, it is finite on a convex closed subset of $[a/\beta, b/\alpha] \times [1/\beta, 1/\alpha]$ and is continuous on the interior of this set. Therefore for any fixed x, $\inf_{y \in [1/\beta, 1/\alpha]} \phi(x, y)$ is attained at some point in $[1/\beta, 1/\alpha]$. Thus $z \geq \inf_{y \in [1/\beta, 1/\alpha]} \phi(x, y)$ if and only if there exists a $y \in [1/\beta, 1/\alpha]$ with

$z \geq \phi(x,y)$, i.e., $(x,z) \in \text{epi } f_2(\phi)$ if and only if there exists a $y \in [1/\beta, 1/\alpha]$ with $(x,y,z) \in \text{epi } \phi$.

Let $\phi, \phi_1, \phi_2, \dots$ be a sequence in $\text{Conv}([a/\beta, b/\alpha] \times [1/\beta, 1/\alpha])$. Let $B \subset [a/\beta, b/\alpha] \times \mathbb{R}$ be the bounded set $B = [a/\beta, b/\alpha] \times [-c, c]$. Fix n and $(x,z) \in B \cap \text{epi } f_2(\phi_n)$. Then there must exist $y \in [1/\beta, 1/\alpha]$ with $(x,y,z) \in B' \cap \text{epi } \phi_n$, where $B' = [a/\beta, b/\alpha] \times [1/\beta, 1/\alpha] \times [-c, c]$. Thus, since $(x', z') \in \text{epi } f_2(\phi)$ if and only if there exists y' with $(x', y', z') \in \text{epi } \phi$, and since $d((x,z),(x',z')) \leq d((x,y,z),(x',y',z'))$, we have

$$\inf_{(x',z') \in \text{epi } f_2(\phi)} d((x,z),(x',z')) \leq \inf_{(x',y',z') \in \text{epi } \phi} d((x,y,z),(x',y',z')),$$

that is

$$d((x,z), \text{epi } f_2(\phi)) \leq d((x,y,z), \text{epi } \phi).$$

This is true for any $(x,z) \in B \cap \text{epi } f_2(\phi)$, and so

$$e_d(B \cap \text{epi } f_2(\phi_n), \text{epi } f_2(\phi)) \leq e_d(B' \cap \text{epi } \phi_n, \text{epi } \phi) \ \ \forall\, n.$$

Similarly $e_d(B \cap \text{epi } f_2(\phi), \text{epi } f_2(\phi_n)) \leq e_d(B' \cap \text{epi } \phi, \text{epi } \phi_n)$ for all n. Thus $\phi = \tau_{\text{AW}} - \lim \phi_n$ implies that $f_2(\phi) = \tau_{\text{AW}} - \lim f_2(\phi_n)$, and so f_2 is continuous.

■

The transformation in equation (3) is the composition $f = f_2 \circ f_1$, and so Lemmas 5 and 6 demonstrate its continuity. As we defined the estimates I_n of I in equation (5) by $I_n = f(J_n)$ and Theorem 2 proves $\{J_n, n \in \mathbb{N}\}$ satisfies an LDP in $\text{Conv}([a,b] \times [\alpha, \beta])$, the following Theorem follows from an application of the contraction principle.

Theorem 7 *The sequence $\{I_n, n \in \mathbb{N}\}$ satisfies an LDP in $\text{Conv}([a/\beta, b/\alpha])$ with good rate function*

$$K'(\phi) = \inf \left\{ K(\psi) : \psi \in \text{Conv}([a,b] \times [\alpha,\beta]) \text{ and } \inf_{y>0} y\psi\left(\frac{x}{y}, \frac{1}{y}\right) = \phi(x) \text{ for all } x \right\},$$

where K is defined in the statement of Theorem 2.

3 The case of cumulative renewal processes.

In Duffy and Metcalfe [10] it is shown that whenever the sequence $\{\vec{Y}_n, n \in \mathbb{N}\}$ is i.i.d., the laws of the sequence of empirical estimates $\{J_n, n \in \mathbb{N}\}$ converge weakly to the Dirac measure δ_J. An analogous result holds for the sequence $\{I_n, n \in \mathbb{N}\}$ of estimates of I.

Theorem 8 *Assume there exists a unique $\mu \in \mathcal{M}_1([Ba, Bb] \times [B\alpha, B\beta])$ with $H(\mu) = 0$, then the sequence of laws of $\{I_n, n \in \mathbb{N}\}$ converge weakly to the Dirac measure $\delta_{f(J_\mu)}$.*

Proof: Clearly $K'(f(J_\mu)) = 0$. A simple adaption of Corollary 27.2.2 of [24] ensures that $K'(\psi) > 0$ for any $\psi \neq f(J_\mu)$. The result now follows from Theorems 2.1 and 2.2 of Lewis et al. [20].

■

Thus for a cumulative renewal process, the blocked sequence $\{\vec{Y}_n, n \in \mathbb{N}\}$ is i.i.d. when $B = 1$ and by Sanov's Theorem the conditions of Theorem 8 are satisfied. In this case H is the relative entropy $H(\nu) = H(\nu|\mu)$, where μ is the common distribution of $\{(X_n, \tau_n), n \in \mathbb{N}\}$. Thus Theorem 8 implies that the laws of the sequence of estimates $\{I_n, n \in \mathbb{N}\}$ converge weakly to the Dirac measure $\delta_{f(J_\mu)} = \delta_{f(J)} = \delta_I$, which is clearly desirable.

The direct adaption of the estimation scheme from [8] to estimate I would be to select a time-scale $T \in \mathbb{R}^+$ such that you believe $\{A_{Tn} - A_{T(n-1)}, n \in \mathbb{N}\}$ forms an i.i.d. sequence and estimate λ in equation (2) by

$$\lambda_t(\theta) = \frac{1}{T} \log \frac{1}{[t/T]} \sum_{i=1}^{[t/T]} \exp(\theta(A_{Ti} - A_{T(i-1)})). \tag{7}$$

and take its Legendre-Fenchel transform. However, if $\{(X_n, \tau_n), n \in \mathbb{N}\}$ is i.i.d., $\{N_{Tn} - N_{T(n-1)}, n \in \mathbb{N}\}$ is typically not i.i.d. for any T, leading to bias in the estimate of λ. See [8, 15] for a discussion on bias in sCGF estimation.

In the simplest scenario possible, we illustrate the advantage of using the apparently more complex approach proposed in this paper. If $X_i = 1$ with probability 1, then the relationship in equation (3) reduces to $I(x) = xJ(1, 1/x)$, relating the large deviations of $\{T_n/n, n \in \mathbb{N}\}$ and $\{N(t)/t, t \in \mathbb{R}^+\}$; a relationship known to hold for a broad class of processes (see [16, 26, 23, 9, 11]). Setting $\{\tau_n, n \in \mathbb{N}\}$ to be i.i.d. taking the values $1, 2$ each with probability $1/2$, one thousand (X_n, τ_n) pairs were generated and both schemes were implemented and run on the same data-set. The results shown in Figure 1 are typical for a broad class of cumulative renewal processes. The real I was determined explicitly and is plotted for reference. Although the results of just a single data-set are shown, they're representative. The scheme described in this paper nearly perfectly overlays I. The more direct scheme is shown for three block-sizes $5, 10, 20$. When the block-size is 5 or 10, underestimation occurs. When the block-size is 20, the estimate does not span the whole x range of I as the full range of possible $A_{Tn} - A_{T(n-1)}$ blocks has not been observed.

4 A rough guide to related work.

The estimation scheme in equation (4) for a one-dimensional version of M was proposed by A. Dembo and used by Duffield et al. [8] for a problem in Asynchronous Transfer Mode (ATM) networking. When combined with theorems of Glynn and Whitt [17], it provided an online measurement-based mechanism for estimating the tail of queue-length distributions. For the success of this approach see, for example, Crosby et al. [5] and Lewis et al. [21].

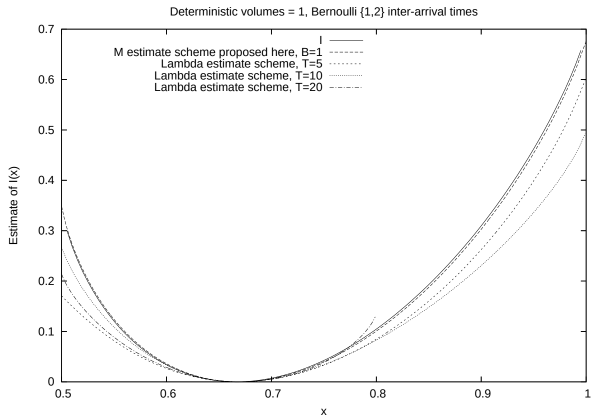

Figure 1: Estimating I via the scheme advocated in this paper with a block-size of 1 and via the more direct scheme for a range of block-sizes. The processes are: $X_n = 1$ for all n and $\{\tau_n, n \in \mathbb{N}\}$ is Bernoulli $\{1, 2\}$.

In ATM networking, discrete time models are appropriate. This makes the need for rate-function estimators of continuous time processes unnecessary. In IP networks, which are more prevalent, it is more appropriate to model data-traffic by a cumulative process where X_n is a packet-size and τ_n is an inter packet arrival time. Thus to adapt for IP networking the successful Connection Admission Control framework based on rate-function estimation, we propose the estimator described in this paper.

For fixed θ, the estimator in equation (4) is known to be biased for correlated data, with the bias tending to zero as block-size B increases. In [15] Ganesh analyses this bias for auto-regressive and finite-state Markovian sources. He shows that it is of similar form for both classes and demonstrates how to compensate for the bias.

If extra structure of the input process is known a priori, more accurate estimates can be made for $M(\theta)$ at fixed θ. If the source is known to be Markovian, for example, see Paschalidis and Vassilaras [22], Eichelsbacher and Ganesh [12], and references therein.

In other analysis of the estimator (4) the existence of B such that $\{Y_n, n \in \mathbb{N}\}$ is i.i.d. is usually assumed. See Györfi et al. [18] for distribution-free confidence intervals for measurement of $\lambda(\theta)$ for fixed θ. For a related question, in the Bayesian context, see Ganesh et al. [13], and Ganesh and O'Connell [14] and references therein. For a large deviations analysis of a connection admission control algorithm based on estimating sCGFs see Duffield [7].

Acknowledgments

Work supported by Science Foundation Ireland grants IN3/03/I346 and 00/PI.1/C067.

References

[1] H. Attouch and R.J.-B. Wets, *A convergence theory for saddle functions*, Trans. Amer. Math. Soc. **280** (1983), no. 1, 1–41.

[2] ———, *Isometries for the Legendre-Fenchel transform*, Trans. Amer. Math. Soc. **296** (1986), no. 1, 33–60.

[3] G. Beer, *Topologies on closed and closed convex sets*, Kluwer Academic, 1993.

[4] W. Bryc and A. Dembo, *Large deviations and strong mixing*, Ann. Inst. H. Poincaré Probab. Statist. **32** (1996), no. 4, 549–569.

[5] S. Crosby, I. Leslie, B. McGurk, J.T. Lewis, R. Russell, and F. Toomey, *Statistical properties of a near-optimal measurement-based CAC algorithm*, Proceedings of IEEE ATM '97 (Lisbon, Portugal), 1997.

[6] A. Dembo and O. Zeitouni, *Large deviation techniques and applications*, Springer, 1998.

[7] N. G. Duffield, *A large deviation analysis of errors in measurement based admission control to buffered and bufferless resources*, Queueing Systems Theory Appl. **34** (2000), no. 1-4, 131–168.

[8] N. G. Duffield, J.T. Lewis, N. O'Connell, R. Russell, and F. Toomey, *Entropy of ATM traffic streams: a tool for estimating QoS parameters*, IEEE Journal of Selected Areas in Communications (special issue on Advances in the Fundamentals of Networking) **13** (1995), 981–990.

[9] N. G. Duffield and W. Whitt, *Large deviations of inversely related processes with non-linear scalings*, Annals of Applied Probability **8** (1998), 995–1026.

[10] K. Duffy and A.P. Metcalfe, *The large deviations of estimating rate-functions*, J. Appl. Probab. **42** (2005), no. 1, 267–274.

[11] K. Duffy and M. Rodgers-Lee, *Some useful functions for functional large deviations*, Stoch. Stoch. Rep. **76** (2004), no. 3, 267–279.

[12] P. Eichelsbacher and A. Ganesh, *Bayesian inference for Markov chains*, J. Appl. Probab. **39** (2002), no. 1, 91–99.

[13] A. Ganesh, P. Green, N. O'Connell, and S. Pitts, *Bayesian network management*, Queueing Systems Theory Appl. **28** (1998), no. 1-3, 267–282.

[14] A. Ganesh and N. O'Connell, *An inverse of Sanov's theorem*, Statist. Probab. Lett. **42** (1999), no. 2, 201–206.

[15] A.J. Ganesh, *Bias correction in effective bandwidth estimation*, Performance Evaluation **27** (1996), no. 8, 319–330.

[16] P. Glynn and W. Whitt, *Large deviations behaviour of counting processes and their inverses*, Queueing Systems **17** (1994), 107–128.

[17] ______, *Logarithmic asymptotics for steady-state tail probabilities in a single-server queue*, J. Appl. Probab. **31A** (1994), 413–430.

[18] L. Györfi, A. Rácz, K. Duffy, J.T. Lewis, and F. Toomey, *Distribution-free confidence intervals for measurement of effective bandwidths*, J. Appl. Probab. **37** (2000), 1–12.

[19] T. Kuczek and K.N. Crank, *A large-deviation result for regenerative processes*, J. Theoret. Probab. **4** (1991), no. 3, 551–561.

[20] J.T. Lewis, C.-E. Pfister, and W.G. Sullivan, *Entropy, concentration of probability and conditional limit theorems*, Markov Process. Related Fields **1** (1995), 319–386.

[21] J.T. Lewis, R. Russell, F. Toomey, B. McGurk, S. Crosby, and I. Leslie, *Practical connection admission control for ATM networks based on on-line measurements*, Computer Communications **21** (1998), no. 17, 1585–1596.

[22] I.Ch. Paschalidis and S. Vassilaras, *On the estimation of buffer overflow probabilities from measurements*, IEEE Trans. Inform. Theory **47** (2001), no. 1, 178–191.

[23] A. Puhalskii and W. Whitt, *Functional large deviation principles for first-passage-time processes*, Ann. Appl. Probab. **7** (1997), no. 2, 362–381.

[24] R.T. Rockafellar, *Convex analysis*, Princeton University Press, 1970.

[25] R.T. Rockafellar and R.J.-B. Wets, *Variational analysis*, Springer, 1998.

[26] R. Russell, *The large deviations of random time changes*, Ph.D. thesis, Trinity College Dublin, 1997.