Large Deviations on One Dimensional Hypergroups

Abstract

Let S_n be a random walk on a Sturm-Liouville hypergroup $(\mathbb{R}_+, *)$, i.e. a Markov chain on $\mathbb{R}_+$ with stationary transition kernel $P(x; A) = \delta_x * \mu(A)$ where $\mu \in M^1(K)$. The principle of large deviations is shown for the distributions of S_n/n under the assumption that μ is compactly supported.

1 Introduction

The interest in random walks on one dimensional hypergroups originates in Kingman's [7] work on rotation invariant random vectors on $\mathbb{R}^n$. More generally this concept can be used to study isotropic random walks on algebraic structures such as Riemannian symmetric spaces of noncompact type and rank one (see [1]). For random walks on one dimensional hypergroups several probabilistic limit theorems have been proven including laws of large numbers and central limit theorems (see [1, 11, 12]).

We shall study here the large deviation principle associated with the law of large numbers for these random walks that is the large deviation principle for the distributions of S_n/n. This leads to an analogue of Cramér's theorem concerning large deviations for sums of independent identical distributed random variables (as given for instance in Section 3 of [9]).

For the convenience of the reader we start by recalling some basic facts concerning random walks and moment functions on hypergroups and the abstract large deviation principle. In section 3 the large deviation upper bound is proven under a suitable condition on the law μ. The idea of the proof of the large deviation lower bound is to replace the exponential function by the multiplicative functions of the hypergroup. We shall see in section 4 that this is indeed possible whenever the law μ has bounded support.

Finally let us remark that corresponding results are also available for random walks on polynomial hypergroups on $\mathbb{N}_0$ (see [3]). Furthermore large deviation principles of three levels for random walks on the dual Jacobi and on the dual disk hypergroup can be found in [6].

2 Random walks on one dimensional hypergroups

2.1 A class of one dimensional hypergroups

Let $\mathbb{R}_+ = [0, \infty[$ be a *hypergroup*; this means that there exists an associative *convolution* $(x, y) \mapsto \delta_x * \delta_y \in M^1(\mathbb{R}_+)$ satisfying certain conditions such that $M(\mathbb{R}_+)$ is a Banach algebra with this convolution. For details and an exhaustive list of examples we refer the interested reader to [1].
Now consider functions $A : \mathbb{R}_+ \to \mathbb{R}$ with the properties

(A_0) $A \in C(\mathbb{R}_+),\ A(x) > 0$ for $x > 0$, and A restricted to $]0, \infty[$ is continuously differentiable.

(A_1) One of the following two conditions is valid.

(A_{1a}) $A(0) = 0$ and for all x in a neighborhood of 0

$$\frac{A'(x)}{A(x)} = \frac{\alpha_0}{x} + \alpha_1(x)$$

where $\alpha_0 > 0$ and $\alpha_1 \in C^\infty(\mathbb{R})$ satisfies $\alpha_1(-x) = -\alpha_1(x)$ for $x \in \mathbb{R}$.

(A_{1b}) $A(0) > 0$ and $A \in C^1(\mathbb{R}_+)$.

(A_2) There exists a function $\beta \in C^1(\mathbb{R}_+)$ with $\beta(0) \geqslant 0$ and $\frac{A'(x)}{A(x)} - \beta(x) \geqslant 0$. Furthermore the functions $\frac{A'(x)}{A(x)} - \beta(x)$ and $q := \frac{1}{2}\beta' - \frac{1}{4}\beta\frac{A'(x)}{A(x)}\beta$ are decreasing on $]0, \infty[$.

A hypergroup $(\mathbb{R}_+, *)$ is called *Sturm-Liouville hypergroup (associated with A)*, if there is a function A on $\mathbb{R}_+$ satisfying (A_0) such that for every real valued function f on $\mathbb{R}_+$, which is the restriction of an even C^∞-function on $\mathbb{R}$, the function $u_f : \mathbb{R}_+^2 \to \mathbb{R}$ defined by $u_f(x, y) := \int f(z)\delta_x * \delta_y(z)$ belongs to $C(\mathbb{R}_+^2)$ and is a solution of the partial differential equation

$$u_{xx} + \frac{A'(x)}{A(x)}u_x = u_{yy} + \frac{A'(y)}{A(y)}u_y$$

and

$$u_y(x, 0) = 0 \text{ for } x \in]0, \infty[.$$

Theorem 3.11 in [11] shows that to every function A with the properties (A_0)–(A_2) there exists a Sturm-Liouville hypergroup associated with A. Note that all known examples of hypergroups on $\mathbb{R}_+$ fall within this framework.
In the sequel we assume without mention that every hypergroup $(\mathbb{R}_+, *)$ is a Sturm-Liouville hypergroup associated with a function A with the properties (A_0)–(A_2).

A function $\varphi : \mathbb{R}_+ \to \mathbb{C}$ is called *multiplicative* if $\int \varphi(z)d(\delta_x * \delta_y)(z) = \varphi(x)\varphi(y)$ for all $x, y \in \mathbb{R}_+$. The multiplicative functions of Sturm-Liouville hypergroups are the solutions φ_λ of the Sturm-Liouville differential equation

$$\varphi_\lambda'' + \frac{A'(x)}{A(x)}\varphi_\lambda' + (\lambda + \rho)\varphi_\lambda = 0 \tag{2.1}$$

$$\varphi_\lambda(0) = 1 \text{ and } \varphi_\lambda'(0) = 0 \quad (\lambda \in \mathbb{C}) \tag{2.2}$$

where

$$\rho := \lim_{x\to\infty} \frac{A'(x)}{2A(x)}. \tag{2.3}$$

The existence of ρ is proven in [11].
Example:
Let $\alpha > -1/2$ and $A_\alpha(x) = x^{2\alpha+1}$. The Sturm-Liouville hypergroup associated with $A_\alpha(x)$ is called *Bessel hypergroup to the parameter α*. Its multiplicative functions are given by $\varphi_\lambda(x) = \Lambda_\alpha(\lambda x)$ where Λ_α is the Bessel function to the parameter α normed by $\Lambda_\alpha(0) = 1$.

2.2 Random walks

Let $(K, *)$ be a hypergroup and let $\mathcal{B}(K)$ be its Borel σ-algebra. Any Markov chain S_n with state space K and stationary transition kernel is called *random walk with law μ* if $S_0 = 0$ and its transition kernel is homogeneous with respect to the convolution of the hypergroup in the following sense:

$$P(x; A) := \delta_x * \mu(A) \qquad (x \in K; A \in \mathcal{B}(K)) \tag{2.4}$$

with a probability measure μ on K.
For general properties of such random walks we refer to [1]. It is immediate from the definition that the distribution of the variables S_n is given by the n-fold convolution product $\mu^{(n)}$.

2.3 The modified moments

For $n \geqslant 0$, $\lambda \in \mathbb{C}$ and $x \in \mathbb{R}_+$ define functions $\phi_{n,\lambda}(x)$ and $m_n(x)$ by

$$\phi_{n,\lambda}(x) := \left(\frac{\partial}{\partial s}\right)^n \varphi_{(\lambda+s)}(x)|_{s=0} \qquad \text{and} \qquad m_n(x) = \phi_{n,i\rho}(x). \tag{2.5}$$

It can be shown that $m_n(x) \geqslant 0$ for all $n \geqslant 1$. For $\mu \in M^1(\mathbb{R}_+)$ the *modified expectation* is defined as

$$E_*(\mu) := \int_0^\infty m_1(x)d\mu(x). \tag{2.6}$$

The modified expectation plays an important role in the laws of large numbers for random walks on $\mathbb{R}_+$ (see [1, 11]). It will also occur in our large deviation result below.

2.4 The abstract large deviation principle

Consider a sequence F_n of probability measures on a polish space E converging weakly to a degenerate distribution at some point $x_0 \in E$ (in our main result (Theorem 4.3) F_n will be the distribution of S_n/n and E the interval $[0, \infty[$). The abstract definition of the large deviation principle is (see [2, 4] or [9]):

Definition 1. Let $\{F_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures on a polish space E and $\{a_n\}_{n\in\mathbb{N}}$ a divergent sequence of positive numbers. We say that $\{F_n\}$ satisfies the *large deviation principle* with constants $\{a_n\}_{n\in\mathbb{N}}$ and rate function $I : E \to [\,0,\,\infty\,]$, if the following conditions hold:

(i) I is lower semicontinuous and has compact level sets, i.e. for each $m \geqslant 0$ $\{x \mid I(x) \leqslant m\}$ is compact.

(ii) For each closed subset A of E

$$\limsup_{n\to\infty} \frac{1}{a_n} \log F_n(A) \leqslant -\inf_{x\in A} I(x).$$

This is called the large deviation upper bound.

(iii) For each open subset G of E

$$\liminf_{n\to\infty} \frac{1}{a_n} \log F_n(G) \geqslant -\inf_{x\in G} I(x).$$

This is called the large deviation lower bound.

Here and in the following log denotes the natural logarithm.

To prove these large deviation bounds we apply the following theorem due to Ellis (Theorem II.6.1 in [4]).

Theorem 1 (Ellis). *Let W_n be an arbitrary sequence of random variables with values in $\mathbb{R}$ and $\{a_n\}_{n\in\mathbb{N}}$ a divergent sequence of positive numbers. Define for $t \in \mathbb{R}$*

$$c_n(t) := \frac{1}{a_n} \log E(\exp(tW_n)),$$

where E denotes the usual expectation of a random variable.
Assume that

(a) Each $c_n(t)$ is finite for every $t \in \mathbb{R}$.

(b) $c(t) := \lim_{n\to\infty} c_n(t)$ exists and is finite for every $t \in \mathbb{R}$.

Let F_n be the distribution of W_n/a_n. Then the following conclusions hold.

(i) The function $I(x) = \sup_{t\in\mathbb{R}}\{tx - c(t)\}$ is convex, lower semicontinuous, and nonnegative. $I(x)$ has compact level sets.

(ii) *The large deviation upper bound is valid with constants $\{a_n\}_{n\in\mathbb{N}}$ and rate function $I(x)$.*

(iii) *Assume in addition that $c(t)$ is continuously differentiable for all $t \in \mathbb{R}$. Then the large deviation lower bound is valid with constants $\{a_n\}_{n\in\mathbb{N}}$ and rate function $I(x)$.*

3 The large deviation upper bound

Let $(\mathbb{R}_+, *)$ be a Sturm-Liouville hypergroup as above. For $\mu \in M^1(\mathbb{R}_+)$ its *(usual) moment generating function* $f_\mu(t)$ is defined by

$$f_\mu(t) = \int_0^\infty e^{tx}\, d\mu(x). \tag{3.1}$$

Theorem 2. *Let $(\mathbb{R}_+, *)$ be as above and let $\mu \in M^1(\mathbb{R}_+)$ be a probability measure with $f_\mu(t) := \int_0^\infty e^{tx} d\mu(x) < \infty$ for $t > 0$.*
Denote the random walk with law μ by S_n and let $c_n(t) = \frac{1}{n}\log E(\exp(tS_n))$. Then

(i) *$c_n(t)$ exists for all $t \in \mathbb{R}$ and is finite. Furthermore $c(t) := \lim_{n\to\infty} c_n(t)$ exists for all $t \in \mathbb{R}$ and is finite.*

(ii) *The large deviation upper bound is valid for the distributions of S_n/n, i.e. for every closed set $A \subseteq \mathbb{R}$*

$$\limsup_{n\to\infty} \frac{1}{n} \log P(S_n/n \in A) \leqslant -\inf_{x\in A} I(x)$$

where

$$I(x) \begin{cases} +\infty & x < 0 \\ \sup_{t\in\mathbb{R}}\{tx - c(t)\} & x \geqslant 0 \end{cases}$$

Proof. We have for $m, n \in \mathbb{N}$

$$\int_0^\infty e^{tx}\, d\mu^{(n+m)}(x) = \int_0^\infty \int_0^\infty \exp(t(x*y))\, d\mu^{(m)}(x) d\mu^{(n)}(y), \tag{3.2}$$

where $\exp(t(x*y)) := \int e^{tz}\, d(\delta_x * \delta_y)(z)$.
It is shown in [11] Proposition 3.9 that for all $x, y \in \mathbb{R}_+$

$$\mathrm{Tr}\, \delta_x * \delta_y \subseteq [|x-y|, x+y] \tag{3.3}$$

and consequently

$$\exp(t(x*y)) \leqslant \exp(t(x+y)) \qquad t \geqslant 0, \tag{3.4}$$

$$\exp(t(x*y)) \geqslant \exp(t(x+y)) \qquad t \leqslant 0. \tag{3.5}$$